

# Development of scale-bridging approaches for the simulation of proton conduction and intermolecular interactions

## Dissertation

zur Erlangung des Doktorgrades der Naturwissenschaften  
(Dr. rer. nat.)

der

Naturwissenschaftlichen Fakultät II  
Chemie, Physik und Mathematik

der Martin-Luther-Universität  
Halle-Wittenberg

vorgelegt von

Herrn Christian Dreßler  
geb. am 11.10.1989 in Löbau

1. Gutachter: Prof. Dr. Daniel Sebastiani  
Martin-Luther-Universität Halle-Wittenberg
2. Gutachter: Prof. Dr. Christoph Jacob  
Technische Universität Braunschweig

Tag der Verteidigung: 11.12.2020



## Abstract

This thesis deals with the development of scale-bridging methods for the simulation of proton conduction and an efficient approach for the calculation of inter-molecular electrostatic interactions.

The starting point for the investigation of long-range proton transfer is the combined Molecular Dynamics/Lattice Monte Carlo (cMD/LMC) approach. In order to benchmark the cMD/LMC approach and to examine the effect of the rotational reorientation rate of the anions on the diffusion coefficient, *ab initio* molecular dynamics simulations of the solid acid family  $\text{CsH}_y\text{XO}_4$  ( $X = \text{S, P, Se}$ ,  $y = 1, 2$ ) were performed in this thesis. It is clearly demonstrated that only the dynamical updating of the lattice in the cMD/LMC method allows for the simulation of the high proton conductivities within the solid acid family. In addition, the cMD/LMC approach is refined in this thesis by introducing a cut-off angle for the calculation of proton transfer probabilities in order to avoid unphysical proton jumps within the oxygen grid.

By neglecting the explicit proton correlation in the cMD/LMC approach, a further drastic reduction in the dimensionality becomes possible. The resulting Molecular Dynamics/Matrix Propagation (MDM) method condenses the dynamic information on proton transport within an entire molecular dynamics simulation into an  $M \times M$  matrix where  $M$  is the number of oxygen atoms. This enormous reduction in dimensionality allows for analytical (instead of numerical) analysis of the model. A thorough mathematical discussion resulted in the verification of the uniqueness of the solutions of the MDM approach and provided the proof of its correct asymptotic behavior, i.e. all protons are equally distributed across the (chemically equivalent) oxygen atoms for large time scales. In order to highlight the increase in the length and time scales, the MDM approach is used to explicitly compute proton dynamics in  $\text{CsH}_2\text{PO}_4$  on the micrometer length scale and in a nanoporous network in agreement with experimental results.

The second part of this thesis is dedicated to the investigation of low-dimensional representations of the linear density-density response function (LDDRF). The LDDRF offers a tool to calculate the fully self-consistent density response of a molecule to arbitrary potentials (in practice: the electrostatic field arising from nearby molecules), corresponding to the exact polarization at all multipolar orders. In this thesis, a thorough derivation of the mathematical foundations of an efficient representation - the direct moment expansion - of the LDDRF is presented. A description of this representation in terms of a QR and a Cholesky decomposition allowed for its generalization to other self-adjoint, compact and positive linear operators. In a further development step, a modified representation - the reduced eigensystem representation - is derived that allows us to define a trivial criterion for the convergence of the approximation to the density response. As a consequence of its novel eigensystem-like structure, the significant reduction in dimensionality becomes apparent for the calculation of the density-density response function. In order to facilitate the application of the LDDRF in a molecular dynamics simulation, the dependence of the efficient representation on changes in the molecular geometry is investigated.

## List of peer-reviewed articles contributing to this thesis

### Efficient representations of the static linear density-density response function

1. Christian Dreßler and Daniel Sebastiani  
Reduced eigensystem representation of the linear density-density response function.  
*International Journal of Quantum Chemistry*, 120(3):e26085, **2020**.
2. Christian Dreßler, Arne Scherrer, Paul Ahlert, and Daniel Sebastiani.  
Efficient representation of the linear density-density response function.  
*Journal of Computational Chemistry*, 40(31):2712–2721, **2019**.
3. Paul Ahlert, Arne Scherrer, Christian Dreßler, and Daniel Sebastiani.  
Iterative approach for the moment representation of the density-density response function.  
*The European Physical Journal B*, 91(6):94, **2018**.
4. Arne Scherrer, Christian Dreßler, Paul Ahlert, and Daniel Sebastiani.  
Generalization of the electronic susceptibility for arbitrary molecular geometries.  
*The Journal of Chemical Physics*, 144(14):144111, **2016**.

### Multiscale approaches for long-range proton transfer

5. Christian Dreßler, Gabriel Kabbe, Martin Brehm, and Daniel Sebastiani.  
Exploring non-equilibrium molecular dynamics of mobile protons in the solid acid  $\text{CsH}_2\text{PO}_4$  at the micrometer and microsecond scale.  
*The Journal of Chemical Physics*, 152(16):164110, **2020**.
6. Christian Dreßler, Gabriel Kabbe, Martin Brehm, and Daniel Sebastiani.  
Dynamical matrix propagator scheme for large-scale proton dynamics simulations.  
*The Journal of Chemical Physics*, 152(11):114114, **2020**.
7. Maximilian Wagner, Christian Dreßler, Felix P. Lohmann-Richters, Kevin Hanus, Daniel Sebastiani, Aron Varga, and Bernd Abel.  
Mechanism of ion conductivity through polymer-stabilized  $\text{CsH}_2\text{PO}_4$  nanoparticulate layers from experiment and theory.  
*Journal of Materials Chemistry A*, 7:27367–27376, **2019**.
8. Gabriel Kabbe, Christian Dreßler, and Daniel Sebastiani.  
Proton mobility in aqueous systems: combining ab initio accuracy with millisecond timescales.  
*Physical Chemistry Chemical Physics*, 19(42):28604–28609, **2017**.
9. Gabriel Kabbe, Christian Dreßler, and Daniel Sebastiani.  
Toward realistic transfer rates within the coupled Molecular Dynamics/Lattice Monte Carlo approach.  
*The Journal of Physical Chemistry C*, 120(36):19905–19912, **2016**.

### Proton conduction in solid acids

10. Christian Dreßler and Daniel Sebastiani.  
Effect of chemical environment on proton mobility in the solid acids family

$\text{CsH}_y\text{XO}_4$  ( $x = \text{S, P, Se}$ ;  $y = 1, 2$ ) from *ab initio* molecular dynamics simulations.  
*Physical Chemistry Chemical Physics*, 22:10738–10752, **2020**.

11. Christian Dreßler, Gabriel Kabbe, and Daniel Sebastiani.  
Insight from atomistic simulations of protonation dynamics at the nanoscale.  
*Fuel Cells*, 16(6):682–694, **2016**.
12. Christian Dreßler, Gabriel Kabbe, and Daniel Sebastiani.  
Proton conductivity in hydrogen phosphate/sulfates from a coupled Molecular Dynamics/Lattice Monte Carlo (cMD/LMC) approach.  
*The Journal of Physical Chemistry C*, 120(36):19913–19922, **2016**.



## *Acknowledgements*

First of all, I would like to thank my supervisor Daniel Sebastiani for his encouragement, patience and advice. While preparing this thesis, I became a father and received every possible support from my supervisor. I also appreciate the fact that he always encouraged me to attend further mathematics lectures, which enabled me to write this thesis in its present version.

I would like to express my appreciation to all the members of the Sebastiani research group for their support over the years. I am especially grateful for the opportunity to work with Arne Scherrer and Gabriel Kabbe, who shared valuable insights into their projects with me in a very open, welcoming manner and taught me a lot during my thesis work.

Although I am very thankful that I could have been able to spend much time on my research work, life would be dull and empty without my family and friends. I am deeply indebted to my wife Annegret for not running away during the preparation of this thesis and for bringing the greatest gift, our son Theodor, into our live.

I am also very grateful for the ever-present backing of my parents and my other family members. I am aware that I have had the privilege of growing up in a stable, supportive family in a prosperous country.

I also wish to thank all the people who proofread this thesis, particularly Annika Eisenschmidt, Florian Glasneck, Pouya Partovi-Azar, Martin Brehm and last but not the least Patrick Sheehan.



# Contents

<b>Abstract</b>	<b>i</b>
<b>Acknowledgements</b>	<b>v</b>
<b>1 Preface</b>	<b>1</b>
<b>2 Theoretical Framework</b>	<b>5</b>
2.1 Molecular Dynamics	5
2.1.1 Classical vs. <i>Ab Initio</i> Molecular Dynamics	5
2.1.2 Numerical Integration of the Equations of Motion	6
2.1.3 Controlling the Temperature: Thermostats	7
2.1.4 <i>Ab Initio</i> Molecular Dynamics (AIMD)	8
2.2 Electronic Structure Methods	8
2.2.1 Hartree-Fock Approach	10
2.2.2 Density Functional Theory	11
2.2.3 Density Functional Perturbation Theory	14
2.3 The Linear Density-Density Response Function (LDDRF)	17
2.3.1 Different Representations of the LDDRF	17
2.3.2 Derivation of the LDDRF	21
2.3.3 Basic Symmetries and Analytic Properties of the LDDRF	23
2.3.4 Iterative Spectral Decomposition of the Static LDDRF	24
2.3.5 Representation of Linear Operators	24
2.4 Markov Chains	29
2.5 Dynamical Systems	32
2.6 Asymmetric Simple Exclusion Process	33
2.7 Monte Carlo Methods	33
2.8 The Combined Molecular Dynamics/Lattice Monte Carlo (cMD/LMC) Approach for Long-range Proton Transfer	35
2.9 Solid Acids	38
2.9.1 Open Questions Concerning the Conduction Mechanism in Solid Acids	42
<b>3 Summary of the Peer-reviewed Articles</b>	<b>45</b>
3.1 Multiscale Approaches for Long-range Proton Transfer	45
3.1.1 Article I: Exploring Non-equilibrium Molecular Dynamics of Mobile Protons in the Solid Acid CsH <sub>2</sub> PO <sub>4</sub> at the Micrometer and Microsecond scale	50
3.1.2 Article II: Dynamical Matrix Propagator Scheme for Large-scale Proton Dynamics Simulations	64
3.1.3 Article III: Mechanism of Ion Conductivity Through Polymer- stabilized CsH <sub>2</sub> PO <sub>4</sub> Nanoparticulate Layers from Experiment and Theory.	79

3.1.4	Article IV: Proton Mobility in Aqueous Systems: Combining <i>ab initio</i> Accuracy with Millisecond Timescales . . . . .	90
3.1.5	Article V: Toward Realistic Transfer Rates within the Coupled Molecular Dynamics/Lattice Monte Carlo Approach. . . . .	97
3.2	Efficient Representations of the Static LDDRF . . . . .	106
3.2.1	Article VI: Reduced Eigensystem Representation of the LDDRF . . . . .	109
3.2.2	Article VII: Efficient Representation of the LDDRF . . . . .	121
3.2.3	Article VIII: Iterative Approach for the Moment Representation of the LDDRF. . . . .	132
3.2.4	Article IX: Generalization of the Electronic Susceptibility for Arbitrary Molecular Geometries. . . . .	138
3.3	Proton Conduction in Solid Acids . . . . .	146
3.3.1	Article X: Effect of Anion Reorientation on Proton Mobility in the Solid Acids Family $\text{CsH}_y\text{XO}_4$ ( $X = \text{S, P, Se}$ ; $y = 1, 2$ ) from AIMD . . . . .	147
3.3.2	Article XI: Proton Conductivity in Hydrogen Phosphate/Sulfates from a Coupled Molecular Dynamics/Lattice Monte Carlo (cMD/LMC) Approach . . . . .	163
3.3.3	Review XII: Insight from Atomistic Simulations of Protonation Dynamics at the Nanoscale . . . . .	174
<b>4</b>	<b>Conclusion</b>	<b>189</b>
<b>A</b>	<b>General Linear Response Theory</b>	<b>191</b>



# List of Figures

2.1	Domain and image of the linear operator $\hat{\mathbf{T}}$ , and domain $\mathcal{P}$ and image $\mathcal{Z}$ of the restriction $\mathcal{T} := \hat{\mathbf{T}} _{\mathcal{P}}$ . Please note that $\mathcal{P}$ and $\mathcal{Z}$ have a trivial intersection ( $\mathcal{P} \cap \mathcal{Z} = \emptyset$ ).	18
2.2	Schematic illustration of the response density of the water molecule (right) due to a perturbing water molecule (left). The potential originating from the left water molecule can be expanded at the responding (right) water molecule using a few basis functions. This figure is reproduced from article [Dreßler, Sebastiani, <i>Int. J. Quantum Chem.</i> , <b>2020</b> .] in this thesis.	19
2.3	Number of proton jumps with respect to the O-O distance (histogram) occurring in a 100 ps simulation of a box containing 16 protons. Details concerning the high temperature (HTP) and low temperature phases (LTP) of $\text{CsHSO}_4$ , $\text{CsHSeO}_4$ , $\text{CsH}_2\text{PO}_4$ are given in section 2.9. Reproduced with permission from Dreßler, Sebastiani, <b>Phys. Chem. Chem. Phys.</b> , 2020, <b>22</b> , 10738-10752.	36
2.4	The combined Molecular Dynamics/ lattice Monte Carlo approach. Reproduced with permission from Dreßler et al., <i>J. Chem. Phys.</i> <b>153</b> , 114114 (2020). Copyright 2020 AIP Publishing LLC.	36
2.5	O-O distance histograms and jump rate function $\omega$ for the specific example of $\text{CsH}_2\text{PO}_4$ . The jump rate function $\omega$ (right) is obtained as point-wise quotient of the two histograms (left).	37
2.6	Flow chart diagram of the CMD/LMC approach. This figure is taken from ref. 1.	38
2.7	Eight unit cells of the crystal structure of the monoclinic phase (LTP) of $\text{CsH}_2\text{PO}_4$ are shown with different lattice orientations. Red: oxygen, green: caesium, gray: hydrogen. Reproduced with permission from Dreßler, Sebastiani, <b>Phys. Chem. Chem. Phys.</b> , 2020, <b>22</b> , 10738-10752.	40
2.8	One unit cell of the crystal structure of the cubic phase (HTP) of $\text{CsH}_2\text{PO}_4$ is shown. Partial occupancies are indicated by partial coloring of the atoms. <sup>2</sup> Red: oxygen, green: caesium, lilac: phosphorus. Reproduced with permission from Dreßler, Sebastiani, <b>Phys. Chem. Chem. Phys.</b> , 2020, <b>22</b> , 10738-10752.	41
2.9	Four unit cells of the crystal structure of the monoclinic phase (LTP) of $\text{CsHSO}_4$ are shown with different lattice orientations. red: oxygen, green: caesium, gray: hydrogen. Reproduced with permission from Dreßler, Sebastiani, <b>Phys. Chem. Chem. Phys.</b> , 2020, <b>22</b> , 10738-10752.	42

2.10	Four unit cells of the crystal structure of the tetragonal phase (HTP) of CsHSO <sub>4</sub> are shown with different lattice orientations. red: oxygen, green: caesium, gray: hydrogen. Reproduced with permission from Dreßler, Sebastiani, <b>Phys. Chem. Chem. Phys.</b> , 2020, <b>22</b> , 10738-10752. . . . .	43
3.1	Schematic illustration of the Molecular Dynamics/ Matrix propagation (MDM) approach. Reproduced with permission from Dreßler et al., <i>J. Chem. Phys.</i> <b>153</b> , 114114 (2020). Copyright 2020 AIP Publishing LLC.	47
3.2	Comparison of the AIMD, cMD/LMC and MDM approaches. . . . .	49

# List of Tables

2.1	Characteristic crystal structure parameters for the high- and low-temperature phases (HTP/LTP) of $\text{CsH}_2\text{PO}_4$ , $\text{CsHSeO}_4$ and $\text{CsHSO}_4$ <sup>3</sup> . . . . .	40
-----	---	----



# List of Abbreviations

HTP	High Temperature Phase
LTP	Low Temperature Phase
MD	Molecular Dynamics
AIMD	<i>Ab Initio</i> Molecular Dynamics
KS	Kohn-Sham
DFT	Density Functional Theory
DFPT	Density Functional Perturbation Theory
cMD/LMC	combined Molecular Dynamics/Lattice Monte Carlo approach
MDM	Molecular Dynamics/Matrix Propagation approach
LDDRF	Linear Density-Density Response Function
MC	Monte Carlo
ASEP	Asymmetric Simple Exclusion Process



# Chapter 1

## Preface

The field of (theoretical) physics aims to answer the fundamental questions of life: What is the origin of everything? What are the underlying rules of the universe? Much more applied, but still just as important is the central subject of (theoretical) chemistry: the understanding and the prediction of formation of molecules or crystals from atoms. The fight against the permanent scourges of humanity including diseases such as cancer or conditions such as hunger, the search for a sustainable supply of resources and the halting of global warming will only be successful with decisive contributions from molecular explanations and predictions of reactions or compound properties.

But in contrast with physics, the toolbox for answering the central questions in chemistry has been essentially complete since the dawn of quantum mechanics. This was expressed by Dirac back in 1929:<sup>4</sup>

The fundamental laws necessary for the mathematical treatment of a large part of physics and the whole of chemistry are thus completely known, and the difficulty lies only in the fact that application of these laws leads to equations that are too complex to be solved.

This quotation of Dirac points also to a fundamental problem which has not been resolved in the last 90 years, namely that the required computational resources for the calculation of large systems are not available. Two reasons contribute to this problem: first of all, the scaling of quantum chemistry methods is often far from linear; secondly, the dimensions of systems that can be accounted for with current *ab initio* methods (a few nanometers and less than a nanosecond) are several orders of magnitude apart from the system sizes of interest for problems in material or life sciences (e.g. millimeters and milliseconds for micro-structured materials).

Even if it were possible to systematically improve the scaling behavior of electronic structure methods towards linearity, the increase in length and time scales for the macroscopic dimensions mentioned in the last sentence would require about a  $10^{12}$ -fold increase in computational effort compared to the system sizes that are feasible nowadays. Over the past century, a computational chemist could hope that the exponential growth of computational resources would help to overcome this problem with the passage of time. However, this idea turned out to be incorrect because the development of more powerful computers is also dependent on the laws of physics and it is not possible to further minimize the dimensions of the circuits on a computer chip without enormous efforts.

Therefore, a new class of algorithms is necessary to achieve progress with time and length scales. One possible solution is provided by multiscale methods. It is immediately apparent that the entire set of atomistic information of a large system provided by quantum chemical methods is not necessary to answer the questions concerning

a single macroscopic property. This thesis deals with two general approaches for the creation of multiscale models.

The first approach identifies the characteristic length and time scales of the property of interest and develops a coarse-grained model for its calculation. The input parameters of this large-scale model are obtained from a much smaller and shorter simulation. Following this line of approach, a multiscale method for long-range proton transfer is constructed in this thesis. The second approach for the construction of multiscale methods makes use of the “divide and conquer” principle. This principle can be utilized for the reduction of the computational costs if the original algorithm possesses a poorer scaling compared to linear (which is the case for quantum chemical calculations). The efficiency gain is obtained by dividing the system into  $N$  parts and applying the original algorithm to the smaller subsystems. In a final step, the overhead costs have to be taken into account for the calculation of the property of interest from the  $N$  subsystems. In this thesis, a fragmentation approach of this type is employed for the calculation of the electrostatic interaction energy of condensed phase systems.

This cumulative thesis is composed of a number of articles,<sup>5–16</sup> which can be grouped into three parts: the investigation of efficient representations of the linear density-density response function,<sup>5–8</sup> the development of a scale bridging approach for the simulation of long-range proton transfer,<sup>9–13</sup> and the explicit simulation of proton conduction in solid acids by means of *ab initio* molecular dynamics.<sup>14–16</sup> In the following, a brief overview of these topics is presented as a condensed summary of the introductions of the constituent articles.

## Linear Density-Density Response Function

From a microscopic point of view, condensed phase systems are described by a large number of interacting particles. In principle, all properties of these large condensed matter systems are accessible from the many-particle wave function. However, the determination of the many-particle wave function (as the solution of the Schrödinger equation) is constrained to rather small systems because the scaling behavior of the quantum chemistry approaches is far from linear. For the specific case of interacting molecules, we can use a fragmentation approach in order to overcome this “curse of dimensionality”. This is already demonstrated by methods such as multipole expansions,<sup>17–23</sup> density fitting<sup>24–27</sup> or perturbation theories.<sup>28–33</sup> Two interacting molecules disturb the ground state density of each molecule. Thus, accurate electrostatic interaction energies should also take polarization effects into account. In this thesis, the static linear density-density response function (LDDRF) is used as a tool to calculate the fully self-consistent density response of a molecule to arbitrary potentials (in practice: the electrostatic field arising from nearby molecules), corresponding to the exact polarization at all multipolar orders (cf. section 3.2).<sup>34–37</sup>

The calculation of molecular density responses becomes trivial once the LDDRF is explicitly known. Due to the high dimensionality of this function, the direct calculation of the LDDRF for a large system is not possible. In order to circumvent this problem, efficient representations of the LDDRF are derived and analyzed in this thesis. It is also demonstrated how to predict these efficient representations for distorted molecular geometries to enable the principal applicability of the LDDRF scheme for the calculation of electrostatic intermolecular interaction including all orders of polarization.

Compared to the static LDDRF, which is an explicit subject of this thesis, the dynamic variant of the LDDRF is perhaps the most important response function



from the time-dependent density functional theory (TDDFT) point of view.<sup>38,39</sup> It gives the linear frequency dependent response of the density to an external scalar potential. After its explicit determination, it can then be utilized to obtain the first-order response of all properties derivable from the density with respect to any scalar field, e.g., polarizability or magnetic susceptibility.<sup>40</sup> The calculation of the poles of the dynamic LDDRF opens up a direct approach for the calculation of electronic excitations. In the theory part of this thesis, the transition from the dynamic to the static LDDRF is discussed. The decisive analytical properties for the application of the efficient transformations (derived for the static LDDRF) also hold for the dynamic LDDRF. Thus, an attempt to apply efficient representation to the dynamic LDDRF is a realistic prospect.

### A Scale-Bridging Approach for the Simulation of Long-range Proton Transfer

Proton conduction is a process of fundamental scientific interest and has been extensively studied by chemists and physicists.<sup>41–51</sup> For simulation of proton transfer, the system size is constrained to a small number of atoms, because the simulation of covalent bond breaking requires the application of *ab initio* molecular dynamics (AIMD).<sup>15,52–60</sup> While we are restricted to this expensive simulation technique, the complete transfer of a proton and the relaxation of the environment occurs typically only a few times per nanosecond within the system sizes tractable with AIMD. Following this line of approach, AIMD simulations will not lead to converged descriptors of proton conduction such as diffusion coefficients. This underlines the urgent need for scale-bridging approaches within this field.<sup>61–63</sup>

In order to speed up molecular dynamics, a multitude of methods such as coarse graining, replica exchange, and adaptive molecular dynamics schemes were developed.<sup>64–66</sup> One promising approach is the use of neural network potentials which are trained using AIMD trajectory data.<sup>67–70</sup> Once trained, fully atomistic dynamics can be generated for very long time scales. The successful application of neural networks was already demonstrated for various systems such as NaOH solutions,<sup>71</sup> *n*-alkanes,<sup>72</sup> or water on ZnO surfaces.<sup>73</sup>

Within the multiscale approaches, the combination of molecular dynamics (MD) simulations and a Markov model<sup>74–79</sup> or a Monte Carlo method<sup>80,81</sup> is widely used. A combination of an MD and a Monte Carlo (MC) algorithm was already applied to a broad range of systems/problems such as polymer materials,<sup>82–85</sup> semiconductors,<sup>86</sup> surface phenomena<sup>87–89</sup> and the folding of proteins.<sup>90–92</sup> The coupling of the two techniques can be achieved through various approaches such as the alternating application of MD and MC steps, the application of different approaches to different spatial subsystems and the algorithmic mixture of the MD and MC methods.<sup>80,81</sup>

In particular, Markov models constructed from MD simulations are well established for the description of dynamical processes of large molecular (often biomolecular) systems.<sup>74–79,93–103</sup> The applicability of Markov state models has been demonstrated for a huge variety of systems, e.g. the prediction of protein folding,<sup>104,105</sup> the binding of small molecules to proteins,<sup>106–114</sup> and the RNA folding kinetics.<sup>96,115,116</sup>

In this thesis, a multiscale Markov model for the simulation of long-range proton dynamics for extended systems and time scales is derived (section 3.1). Within this approach, the proton dynamics information of an entire molecular dynamics simulation can be condensed into a single  $M \times M$  matrix ( $M$  is the number of oxygen atoms in the simulated system). As a result of this drastic reduction in complexity,

the non-equilibrium molecular dynamics of protons in the solid acid  $\text{CsH}_2\text{PO}_4$  can be calculated on the micrometer length scale.

### **Proton Conduction in Solid Acids**

In particular, the solid acid  $\text{CsH}_2\text{PO}_4$  and in general the members of the family  $\text{XH}_y\text{ZO}_4$  (where  $X = \text{K, Rb, Cs, NH}_4$ ;  $y = 1, 2$ ;  $Z = \text{P, S, Se, As}$ ) has attracted considerable interest in recent years as alternative high temperature, water-free proton exchange fuel cell membrane materials.<sup>117-122</sup> A first fuel cell based on  $\text{CsH}_2\text{PO}_4$  is already commercially available.<sup>123</sup> In this thesis, different solid acids are investigated by means of AIMD to elucidate differences in proton conduction mechanism (section 3.3) and to generate a benchmark set for the general applicability of the scale bridging approach for long-range proton conduction.

## Chapter 2

# Theoretical Framework

### 2.1 Molecular Dynamics

Molecular Dynamics (MD) is a numerical simulation method for the solution of the classical many-body problem, i.e. the motion of several interacting bodies in a potential. The solution to this problem is obtained by numerically integrating the (Newtonian) equations of motion. Molecular dynamics is applicable to an immensely broad range of problems from crystalline and amorphous solids through to liquids and solutions. Because of its versatility, it has become indispensable in both pure and applied research.<sup>124,125</sup>

The equations of motion can only be integrated analytically for special cases. In order to tackle this problem, the time is discretized and the problem is solved successively for each time step.

#### 2.1.1 Classical vs. *Ab Initio* Molecular Dynamics

Since we wish to predict the evolution of a system, the obvious approach is to start from the equations of motion. Neglecting the quantum nature, Newton's second law describes the propagation of a particle  $I$  according to

$$m_I \ddot{\mathbf{R}}_I(t) = \mathbf{F}_I(t) = -\nabla V(\mathbf{R}_1(t), \mathbf{R}_2(t), \dots, \mathbf{R}_N(t)), \quad (2.1)$$

where  $m_I$  denotes the mass of the particle  $I$  and  $\mathbf{R}_I(t)$  its spatial coordinates at the time  $t$ .  $V$  denotes the potential energy of the system with respect to all spatial coordinates of the  $N$  particles. The forces  $\mathbf{F}_I(t)$  acting on the particles can be obtained as the gradient of the potential  $V$ .

Several different methods can be employed for the calculation of the forces  $\mathbf{F}_I(t)$ , which lead to different types of MD simulations. In classical MD, the energy of a molecular system is decomposed into different contributions and the individual contributions are parameterized with respect to the nuclear positions. Following this line of approach, the total energy of the system (and thus also the forces) is obtained as the sum of the individual contributions. The use of pre-parameterized potentials is computationally very efficient and allows for the calculation of several thousand atoms for up to hundreds of nanoseconds.

An alternative approach for the calculation of the forces acting on the atomic nuclei is the approximate solution of the Schrödinger equation. This type of method is computationally much more expensive and restricts the system size to several hundred atoms and the time scale to typically less than one nanosecond. This type of MD is referred to as *ab initio* MD (AIMD). While the dimensions of the investigated systems are much more limited in AIMD, it has powerful predictive ability because the forces are obtained from an electronic structure method. In contrast, in classical MD simulations, only previously parameterized processes contribute to the evolution

of the system. This means that the force fields used in classical MD simulations are specifically optimized for certain classes of compounds (and properties) and can not be transferred to other compound classes.

### 2.1.2 Numerical Integration of the Equations of Motion

Equation (2.1) defines a set of coupled second-order partial differential equations, which can be solved by various methods. The standard approach for the solution of such a set of coupled partial differential equations is finite-difference integration. Not all algorithms for finite-difference integration fulfill conditions such as time-reversibility and conservation of energy and momentum. An appropriate method for the numerical integration of the equations of motion is given by the Verlet algorithm:

#### The Verlet algorithm

We start with a Taylor expansion up to the second order for the position of a particle at time  $t + \Delta t$ :

$$\mathbf{r}_i(t + \Delta t) \approx \mathbf{r}_i(t) + \frac{d\mathbf{r}_i(t)}{dt} \Delta t + \frac{1}{2} \frac{d^2\mathbf{r}_i(t)}{dt^2} \Delta t^2. \quad (2.2)$$

Utilizing Newton's second law, we can replace the acceleration according to:

$$\mathbf{r}_i(t + \Delta t) \approx \mathbf{r}_i(t) + \frac{d\mathbf{r}_i(t)}{dt} \Delta t + \frac{\mathbf{F}_i(t)}{2m_i} \Delta t^2. \quad (2.3)$$

The velocity term in equation (2.3) can be eliminated by introducing a similar expression for  $\mathbf{r}_i(t - \Delta t)$ :

$$\mathbf{r}_i(t - \Delta t) \approx \mathbf{r}_i(t) - \frac{d\mathbf{r}_i}{dt} \Delta t + \frac{\mathbf{F}_i(t)}{2m_i} \Delta t^2. \quad (2.4)$$

Adding up equations (2.3) and (2.4) yields:

$$\mathbf{r}_i(t + \Delta t) \approx 2\mathbf{r}_i(t) - \mathbf{r}_i(t - \Delta t) + \frac{\mathbf{F}_i(t)}{m_i} \Delta t^2. \quad (2.5)$$

Equation (2.5) is the well-known Verlet algorithm.<sup>126</sup> Starting from an initial state, equation (2.3) can be used to generate a set of coordinates at the time  $\Delta t$ , which can then be propagated along a trajectory of arbitrary length using equation (2.5). The velocities are not explicitly calculated by the Verlet algorithm. They are accessible by post-processing of the particle positions of the trajectory according to

$$\mathbf{v}_i(t) = \frac{d\mathbf{r}_i}{dt} = \frac{\mathbf{r}_i(t + \Delta t) - \mathbf{r}_i(t - \Delta t)}{2\Delta t}. \quad (2.6)$$

#### The Velocity-Verlet algorithm

The velocity-Verlet algorithm circumvents a disadvantage of the Verlet algorithm and provides explicit formulas for both positions and velocities. We start from the time reversibility of Newton's equations of motion and perform a Taylor expansion of  $\mathbf{r}_i(t + \Delta t)$  for a time step  $-\Delta t$  as follows:

$$\mathbf{r}_i(t) \approx \mathbf{r}_i(t + \Delta t) - \frac{d\mathbf{r}_i(t + \Delta t)}{dt} \Delta t + \frac{1}{2} \frac{d^2\mathbf{r}_i(t + \Delta t)}{dt^2} \Delta t^2. \quad (2.7)$$

Inserting this expression to (2.3) yields:

$$\mathbf{v}_i(t + \Delta t) = \mathbf{v}_i(t) + \frac{\mathbf{F}_i(t) + \mathbf{F}_i(t + \Delta t)}{2m_i} \Delta t. \quad (2.8)$$

The velocity-Verlet algorithm is obtained by combining equations (2.3) and (2.8). This algorithm propagates positions and velocities simultaneously, and direct access to these quantities is useful when employing a constant temperature algorithm.

### 2.1.3 Controlling the Temperature: Thermostats

The numerical integration of the classical equations of motion using an appropriate algorithm conserves the total Hamiltonian. Employing these algorithms and starting with a number of particles  $N$  in a fixed volume  $V$ , the energy is a constant of motion. This situation corresponds to an NVE ensemble.

However, it is desirable to sample from other ensembles (constant temperature (NVT) or constant pressure (NPT) ensemble) as MD simulations usually have to be performed under experimental conditions. This requires the coupling of the simulation system to an external thermostat or barostat.

A case in point is the Nosé-Hoover thermostat. The discussion of this thermostat will follow ref. 127. The Nosé-Hoover dynamics are defined by the following set of dynamical equations:

$$\dot{q}_i = \frac{p_i}{m}, \quad (2.9)$$

$$\dot{p}_i = -\frac{\partial V(\mathbf{q})}{\partial q_i} - p_i \frac{p_\eta}{Q}, \quad (2.10)$$

$$\dot{p}_\eta = \sum_{i=1}^N \frac{p_i^2}{m_i} - NkT, \quad (2.11)$$

$$\dot{\eta} = \frac{p_\eta}{Q}. \quad (2.12)$$

Here  $p_i$  and  $q_i$  denote one-dimensional positions and momenta, respectively, and  $m_i$  denotes the mass of the  $i$ -th particle.  $V(\mathbf{q})$  is the potential acting on the particles. The basic idea of the Nosé-Hoover thermostat is to use a friction factor  $\xi := \frac{p_\eta}{Q}$  (introduced in equation (2.10)) to control the particle velocities. From equation (2.11), it is immediately apparent that the change in the friction factor  $\xi$  is coupled with the difference between the system temperature and the target temperature. This coupling is opposed to the temperature difference and counteracts excessively large deviations from the target temperature. It can be shown that equations (2.9) - (2.12) conserve the Hamiltonian in equation (2.13) for the system  $(\mathbf{q}, \mathbf{p}, p_\eta, \eta)$ . Thus, the system  $(\mathbf{q}, \mathbf{p}, p_\eta, \eta)$  evolves in a microcanonical ensemble (NVE), while the positions and momenta  $(\mathbf{q}, \mathbf{p})$  can be sampled in a canonical ensemble (NVT).

$$H(\mathbf{p}, \mathbf{q}, \eta, \mathbf{p}_\eta) = V(\mathbf{q}) + \sum_{i=1}^N \frac{p_i^2}{2m_i} + \frac{p_\eta^2}{2Q} + NkT\eta \quad (2.13)$$

Equation (2.13) supports the interpretation of  $\xi = \frac{p_\eta}{Q}$  as the scaled velocity of an additional and dimensionless degree of freedom. In this regard, we can consider the variable  $Q$  as the mass of a fictitious particle and  $p_\eta = Q\xi$  as its momentum.

The choice of the mass  $Q$  of the fictitious particle is important. If very large masses are chosen, a distribution consistent with the microcanonical ensemble may result. If very small masses are chosen, fluctuations in the momenta may be greatly inhibited.<sup>127</sup>

The presented approach only generates a canonical ensemble for ergodic systems. Ergodicity is a strong property and, particularly for small or stiff systems, the correct distributions of positions and moments are not generated with the Nosé-Hoover thermostat. A possible attempt to overcome this problem is to couple the variable  $p_\eta$  to another Nosé-Hoover-like thermostat using a new set of variables  $\tilde{\eta}$  and  $\tilde{p}_\eta$ . By connecting  $N$  Nosé-Hoover thermostats using this scheme, a Nosé-Hoover chain of length  $N$  is obtained. Even for large systems, the addition of extra thermostats is relatively inexpensive as they form a simple one dimensional chain. Only the first thermostat interacts with  $N$  particles.

#### 2.1.4 *Ab Initio* Molecular Dynamics (AIMD)

The limiting factor for the efficiency of molecular dynamics simulations is the chosen level of theory for the calculation of the forces acting on the atoms. By employing pre-parametrized potentials for the calculation of the forces, classical MD constitutes a computationally inexpensive variant. In contrast, AIMD is computationally much more demanding due to the solution of the Schrödinger equation for each time step.

This thesis will use AIMD because a more realistic description of the electronic structure is needed. In particular, the description of formation and cleavage of covalent bonds is only possible with the truly predictive power of AIMD.

We use the term AIMD to refer to Born-Oppenheimer MD. In Born-Oppenheimer MD, the static electronic structure problem is solved in each MD step for a set of fixed nuclear positions at an instant in time (cf. eq. (2.15)). The nuclei are propagated according to the classical equations of motion whereas the forces are obtained as the gradient of the potential resulting from the solution of the electronic structure problem. The electronic contribution to the total force on atom  $I$  is given according to equation (2.14).

$$M_I \ddot{\mathbf{R}}_I(t) = -\nabla_I \min_{\Psi_0} \{ \langle \Psi_0, \mathcal{H}_{el} \Psi_0 \rangle \} \quad (2.14)$$

$$\mathcal{H}_{el} \Psi_0 = E_0 \Psi_0 \quad (2.15)$$

In equation (2.15),  $\mathcal{H}_{el}$  denotes the electronic Hamiltonian,  $\Psi_0$  denotes the ground state wave function, and  $E_0$  denotes the ground state energy.

In the next section, we derive equation (2.15) from the time-independent Schrödinger equation and we discuss efficient numerical approximation schemes for the solution of this equation.

## 2.2 Electronic Structure Methods

Electronic structure methods are used in the calculation of many-body properties, in particular many-body energies, from the bare postulates of quantum physics. A many-body particle system is completely described by its many-body wave function. Neglecting relativistic effects and assuming time-independent external potentials, the many-body wave function can be obtained from the stationary many-body Schrödinger equation:

$$\mathcal{H}\Psi = E\Psi \quad (2.16)$$

Here, the time-independent Hamiltonian  $\mathcal{H}$  is defined as:

$$\mathcal{H} = \sum_{i=1}^N -\frac{1}{2}\nabla_i^2 + \sum_{I=1}^M -\frac{1}{2M_I}\nabla_I^2 + \sum_{i=1}^N \sum_{j>i}^N \frac{1}{|\vec{r}_i - \vec{r}_j|} + \sum_{A=1}^M \sum_{B>A}^M \frac{Z_A Z_B}{|\vec{R}_A - \vec{R}_B|} + \sum_{i=1}^N \sum_{A=1}^M -\frac{Z_A}{|\vec{r}_i - \vec{R}_A|} \quad (2.17)$$

where the indices  $i, j$  are used for electrons and  $A, B$  are used for atomic nuclei.  $M_I$  denotes the nuclear masses.  $\vec{R}_A$  and  $\vec{r}_i$  stand for the nuclear and electron positions, respectively, and  $Z_A$  refers to the atomic number.

Equation (2.17) is given in atomic units, i.e.  $e, \hbar, m$  and  $4\pi\epsilon_0$  are considered to be equal to one.

The Hamiltonian  $\mathcal{H}$  can be grouped into contributions from the electron and the nuclear kinetic energy operator, respectively,

$$\hat{T}_e = \sum_{i=1}^N -\frac{1}{2}\nabla_i^2, \quad \hat{T}_n = \sum_{I=1}^M -\frac{1}{2}\nabla_I^2, \quad (2.18)$$

the electron-electron repulsion and the nucleus-nucleus repulsion, respectively,

$$\hat{V}_{ee} = \sum_{i=1}^N \sum_{j>i}^N \frac{1}{|\vec{r}_i - \vec{r}_j|} \quad \hat{V}_{nn} = \sum_{A=1}^M \sum_{B>A}^M \frac{Z_A Z_B}{|\vec{R}_A - \vec{R}_B|}, \quad (2.19)$$

and the electron-nucleus interaction

$$\hat{V}_{ne} = \sum_{i=1}^N \sum_{A=1}^M -\frac{Z_A}{|\vec{r}_i - \vec{R}_A|}. \quad (2.20)$$

Additionally, another time-independent external potential  $v_{ext}$  can be included in the Hamiltonian. This case is discussed in section 2.2.3. Here, we will use a Hamiltonian of the type presented in (2.17), as we wish to solve the electronic structure problem arising from Born-Oppenheimer Molecular Dynamics given by equation (2.14).

The exact quantum mechanical treatment of a many-body system is only possible for very small system sizes. The Born-Oppenheimer approximation achieves a reduction in complexity by decomposing the full many-body wave function  $\Psi(\mathbf{r}, \mathbf{R})$  into the product of an electronic wave function  $\Psi_{el}(\mathbf{r}, \{\mathbf{R}\})$  and a nuclear wave function  $\phi_{nuc}(\mathbf{R})$

$$\Psi(\mathbf{r}, \mathbf{R}) = \Psi_{el}(\mathbf{r}, \{\mathbf{R}\})\phi_{nuc}(\mathbf{R}), \quad (2.21)$$

where  $\mathbf{r}$  denotes the degrees of freedom of the electronic subsystem and  $\mathbf{R}$  denotes the degrees of freedom of the nuclear subsystem. As a consequence of equation (2.21), electron and nuclear states can be treated separately. Thus, it is possible to calculate the electronic part of the wave function with respect to fixed nuclear positions, which is denoted by the parametric dependence of  $\Psi_{el}(\mathbf{r}, \{\mathbf{R}\})$  on the nuclear coordinates  $\{\mathbf{R}\}$ . The theoretical justification of the Born-Oppenheimer approximation is the difference of around 3 orders of magnitude between the masses of nuclei and electrons. As a result, the electrons adapt to the position of the slowly moving cores instantaneously and can be assumed to be completely relaxed at any moment of atomic movement.

Instead of calculating the full many-body wave function, one can reduce the many-body problem to a problem of electrons moving in an external field which is due to

static nuclei. The time-independent Hamiltonian assigned to this problem reads as:

$$\mathcal{H}_{el} = \sum_{i=1}^N -\frac{1}{2}\nabla_i^2 + \sum_{i=1}^N \sum_{j>i}^N \frac{1}{|\vec{r}_i - \vec{r}_j|} + \sum_{i=1}^N \sum_{A=1}^M -\frac{Z_A}{|\vec{r}_i - \vec{R}_A|} \quad (2.22)$$

with

$$\mathcal{H}_{el}\Psi_{el} = E_{el}\Psi_{el} \quad (2.23)$$

In the following, we will focus on the electronic part of the total Hamiltonian. Even for very small systems, the electron-electron term still prevents the analytical solution of the eigenvalue problem associated with the redefined operator  $\mathcal{H}_{el}$ .

Various methods have been developed for the construction of approximate solutions of equation (2.23). A short introduction to a wave-function based (Hartree-Fock) and a density based (density functional theory) approach is given in the next section.

From now on, we will use  $\mathcal{H}$ ,  $E$  and  $\Psi$  instead of  $\mathcal{H}_{el}$ ,  $E_{el}$  and  $\Psi_{el}$ .

### 2.2.1 Hartree-Fock Approach

For the derivation of the Hartree-Fock approach, we start from the construction of an antisymmetrized  $N$ -electron wave function as a Slater determinant of one-particle spin orbitals  $\chi(\mathbf{x}_i)$

$$\Psi_{SD}(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) = \frac{1}{\sqrt{N!}} \sum_{i=1}^{N!} (-1)^{p_i} \mathcal{P}_i \{ \chi_a(\mathbf{x}_1) \chi_b(\mathbf{x}_2) \dots \chi_k(\mathbf{x}_N) \}, \quad (2.24)$$

where the vector  $\mathbf{x}_i$  contains both positional coordinates  $\mathbf{r}_i$  and a spin coordinate  $\sigma$ . Decomposing the Hamiltonian  $\mathcal{H}$  into contributions from the one-particle operator  $\mathcal{O}_1$  and two-particle operator  $\mathcal{O}_2$  according to

$$\mathcal{H} = \underbrace{\mathcal{O}_1}_{\sum_{i=1}^N h_i} + \underbrace{\mathcal{O}_2}_{\sum_{i=1}^N \sum_{j>i}^N \frac{1}{r_{ij}}} \quad (2.25)$$

and

$$h_i = -\frac{1}{2}\nabla_i^2 + \sum_{A=1}^M -\frac{Z_A}{|\vec{r}_i - \vec{R}_A|}, \quad (2.26)$$

yields the following expression for the expectation value of the energy with respect to the Slater determinant  $\Psi_{SD}(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N)$ :

$$E = \langle \Psi_{SD}, \mathcal{H}_{el} \Psi_{SD} \rangle = \sum_{a=1}^N \int \chi_a^*(\mathbf{x}_1) h_1 \chi_a(\mathbf{x}_1) d\mathbf{x}_1 + \frac{1}{2} \sum_{a=1}^N \sum_{b=1}^N [\mathcal{J}_{ab} - \mathcal{K}_{ab}]. \quad (2.27)$$



In equation (2.27) the term  $\mathcal{J}_{ab}$  is referred to as the Coulomb integral, whereas the term  $\mathcal{K}_{ab}$  is referred to as the exchange integral

$$\mathcal{J}_{ab} = \int \chi_a^*(\mathbf{x}_1)\chi_b^*(\mathbf{x}_2)\frac{1}{r_{12}}\chi_a(\mathbf{x}_1)\chi_b(\mathbf{x}_2)d\mathbf{x}_1d\mathbf{x}_2, \quad (2.28)$$

$$\mathcal{K}_{ab} = \int \chi_a^*(\mathbf{x}_1)\chi_b^*(\mathbf{x}_2)\frac{1}{r_{12}}\chi_b(\mathbf{x}_1)\chi_a(\mathbf{x}_2)d\mathbf{x}_1d\mathbf{x}_2. \quad (2.29)$$

While the Coulomb term  $\mathcal{J}_{ab}$  describes the classical Coulomb repulsion between the charge densities  $|\chi_a|^2$  and  $|\chi_b|^2$ , the exchange term does not have a classical equivalent and is a purely quantum mechanical phenomenon.

The next step is to find an optimal wave function by minimizing the expectation value from equation (2.27) with respect to the spin orbitals  $\chi_a(\mathbf{x}_i)$ . To do this, we make use of the variation principle, which means that the expectation value of the Hamiltonian for any trial wave function is always equal to or greater than the ground-state energy. Following this line of approach, the “optimal” set of these spin orbitals in the Slater determinant is the one which minimizes the electronic energy. The minimization is carried out under the constraint that the spin orbitals are orthonormal, which can be achieved by means of the Lagrange multiplier method, and the result is a new eigenvalue equation for each orbital:

$$\hat{f}_i|\chi_a\rangle = \epsilon_a|\chi_a\rangle, \quad (2.30)$$

where  $\hat{f}$  is the Fock-operator:

$$\hat{f} = h_i + \underbrace{\sum_{a \neq b} \mathcal{J}_b - \sum_{a \neq b} \mathcal{K}_b}_{\nu_{HF}}. \quad (2.31)$$

Instead of determining an  $N$ -electron wave function, we now have to solve a set of  $N$  effective single electron Schrödinger equations. According to equation (2.31), it is apparent that the wave function of the  $i$ -th electron is calculated within the field generated by the remaining  $N - 1$  electrons. Thus, the essence of the Hartree-Fock approximation is to replace the complicated many-electron problem by a one-electron problem in which the electron-electron repulsion is treated in an average manner (mean field approach).<sup>128</sup> According to equation (2.31), the Fock-operator is dependent on the one-particle wave function and the one-particle wave functions are obtained from the diagonalization of the Fock-operator. This relationship turns the Hartree-Fock equations into a set of non-linear equations which have to be solved iteratively.

The Slater determinant obtained from the optimized spin orbitals yields the best single-determinant approximation for the exact ground state wave function. The difference between the exact electronic energy and the Hartree-Fock energy of a many-particle system is termed correlation error. According to the variational principle, this correlation error could be systematically reduced by adding “excited” determinants to the space of possible solutions. This type of post-Hartree-Fock approach is known as configuration interaction method. Another ansatz for the incorporation of the correlation error is density functional theory.

### 2.2.2 Density Functional Theory

Wave function based methods in particular suffer from the “curse of dimensions”. While native Hartree-Fock calculations scale with the fourth power with respect to

the number of involved particles, more sophisticated methods such as configuration interaction calculations display even worse scaling behavior.

A possible attempt to overcome this problem is Density Functional Theory (DFT). This method uses the electron density from equation (2.32) as a basic quantity which drastically reduces the number of spatial variables for an  $N$ -electron system from  $3N$  to only 3.

$$n_0(\mathbf{r}) = N \sum_i \langle \Psi_0(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N), \delta(\mathbf{r} - \mathbf{r}_i) \Psi_0(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) \rangle \quad (2.32)$$

A more comprehensive presentation of DFT can be found in in ref. 38 which we follow in this section.

### Hohenberg-Kohn Theorem

Modern DFT is based on the theorems formulated by Hohenberg and Kohn. One theorem states that the entire information of the electronic ground state is contained in its electronic density.

**First Hohenberg-Kohn Theorem:**<sup>38</sup> In a finite, interacting  $N$ -electron system with a given particle-particle interaction, there exists a one-to-one correspondence between the external potentials  $v(\mathbf{r})$  and the ground-state density  $n_0(\mathbf{r})$ . In other words, the external potential is a unique functional of the ground-state density,  $v[n_0](\mathbf{r})$ , up to an arbitrary additive constant.

The Hamiltonian  $\mathcal{H}$  for a given system is fixed up to the external potential. Using the Hohenberg-Kohn theorem, it is possible to express the full many-particle ground state as a unique functional  $E_{v_0}[n]$  of  $n_0$ . This total-energy functional  $E_{v_0}[n]$ , associated with a given external potential  $v_0(\mathbf{r})$ , can be written as:

$$E_{v_0}[n] = \langle \Psi[n], \mathcal{H} \Psi[n] \rangle \quad (2.33)$$

Here,  $n(\mathbf{r})$  is a density of an  $N$ -particle system, and  $\Psi[n]$  is the unique ground-state wave function which produces this density.

Another corollary formulated in the original work of Hohenberg and Kohn<sup>129</sup> states that the ground state energy can be obtained according to the Rayleigh-Ritz variational principle,<sup>130</sup> i.e. only the exact ground state density  $n_0$  minimizes the total energy functional  $E_{v_0}[n]$ . The latter can be written as:

$$E_{v_0}[n] \geq E_0 \quad \forall n \quad (2.34)$$

and  $E_{v_0}[n] = E_0$  if and only if  $n = n_0$ . From now on, we will omit the subscript  $v_0$  in  $E_{v_0}[n]$ .

The Hohenberg-Kohn theorem was derived for  $v$ -representable densities, that is, those that can be represented by a real physical potential.<sup>131</sup> It is in general not possible to decide for a given density whether it is  $v$ -representable or not. However, it is possible to circumvent this need by using the Levy-Lieb constrained search formalism.<sup>132,133</sup> Levy and Lieb provided a similar functional as Hohenberg and Kohn, but which requires a weaker condition for the density ( $N$ -representability) compared to being  $v$ -representable. An electron density is  $N$ -representable if it stems from an

$N$ -electron antisymmetric wave function. This ensures the general applicability of DFT.

From a theoretical point of view, the total-energy  $E[n]$  functional can be decomposed into contributions from the kinetic energy functional  $T[n]$  and the functionals of energy due to electron-electron and electron-nuclei interaction:

$$E[n] = T[n] + E_{en}[n] + E_{ee}[n]. \quad (2.35)$$

The presented concept of a total energy functional is exact. Unfortunately, the shape of this functional remains unknown.

One problem appears in terms of a density formulation of the kinetic energy: there is no adequate approximation for the kinetic energy functional of the electronic ground state density.<sup>134</sup> Kohn-Sham DFT can be utilized to overcome this problem.

### Kohn-Sham DFT

The Kohn-Sham (KS) ansatz provides a practical approach for the determination of the electronic ground state density. Kohn and Sham introduced a non-interacting single determinant reference system of KS orbitals  $\phi_m$ , which fulfills the condition that the electron density of the Kohn-Sham auxiliary system  $n_{KS}$  is equal to the ground state density  $n_0$  obtained from the fully interacting wave function:

$$n_0(\mathbf{r}) \stackrel{!}{=} n_{KS}(\mathbf{r}) = \sum_m (\phi_m^{KS}(\mathbf{r}))^* \phi_m^{KS}(\mathbf{r}). \quad (2.36)$$

By replacing the many-body system with a system of  $N$  non-interacting electrons, we can write within KS-DFT for the total-energy functional:

$$E_{KS}[\{\phi_m^{KS}\}] = \underbrace{\sum_m \langle \phi_m^{KS}, -\frac{1}{2} \nabla^2 \phi_m^{KS} \rangle}_{T_s[\{\phi_m^{KS}\}]} + E_{en}[n] + E_H[n] + E_{xc}[n], \quad (2.37)$$

where  $T_s[\{\phi_m^{KS}\}]$  is the kinetic energy functional.  $T_s[\{\phi_m^{KS}\}]$  is an explicit function of the Kohn-Sham orbitals, but an implicit functional of the density:  $T_s[n] = T_s[\{\phi_m^{KS}[n]\}]$ . Here,  $[\{\phi_m^{KS}\}]$  denotes functional dependence on the set of occupied orbitals. In equation (2.37),  $E_{en}[n]$  refers to the functional due to nucleus-electron repulsion

$$E_{en}[n] = \int \int v_{en}(\mathbf{r}) n(\mathbf{r}) d^3r. \quad (2.38)$$

The functional arising from electron-electron interaction can be split into contributions from the classical Hartree potential  $V_H$ , the energy functional  $E_H[n]$  and the exchange-correlation functional  $E_{xc}[n]$

$$E_{ee}[n] = E_H[n] + E_{xc}[n] = \frac{1}{2} \int \int \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3r' d^3r + E_{xc}[n] \quad (2.39)$$

$$= \frac{1}{2} \int \int v_H(\mathbf{r}) n(\mathbf{r}) d^3r' d^3r + E_{xc}[n]. \quad (2.40)$$

The exchange-correlation functional  $E_{xc}[n]$  contains the difference between the independent-electron system and the real system of interacting electrons. Formally, we can define:

$$E_{xc}[\{\phi_m^{KS}\}] = T[n] - T_s[\{\phi_m^{KS}\}] + E_{ee}[n] - E_H[n]. \quad (2.41)$$

An exact expression of the exchange-correlation functional would allow for the calculation of the exact ground state energies. Unfortunately, the exact shape of  $E_{xc}[n]$  is not known and considerable efforts have been made to search for efficient approximations to the exchange-correlation functional. Many approximations have been developed so far, ranging from simple local approximations based on the homogeneous electron gas to generalized gradient approximations and hybrid functionals incorporating exact Hartree-Fock exchange. These approaches can be grouped into a hierarchy that can be described as a Jacob's ladder of successive approximations climbing up to the "heaven of chemical accuracy".<sup>135</sup>

It should be noted that the variation principle with respect to ground-state energies and densities does not hold true for approximated energy functionals. Nevertheless, for practical applications the variational principle is still applied for current state-of-the-art approximated energy functionals, and the obtained electron density is considered to be an approximation for the true ground-state density.

Differentiation of equation (2.37) with respect to the  $\phi_m^{KS}$  eventually yields the KS equations:

$$\underbrace{\left(-\frac{\hbar^2}{m}\nabla^2 + v_{ne}(\mathbf{r}) + v_H(\mathbf{r}) + v_{xc}(\mathbf{r})\right)}_{\mathcal{H}_{KS}} \phi_m^{KS}(\mathbf{r}) = \epsilon_m \phi_m^{KS}(\mathbf{r}), \quad (2.42)$$

where the Kohn-Sham potential  $v_{KS}$  is defined as:

$$v_{KS}(\mathbf{r}) = v_{ne}(\mathbf{r}) + v_H(\mathbf{r}) + v_{xc}(\mathbf{r}). \quad (2.43)$$

The structure of the KS equations is similar to the Hartree-Fock equations (cf. eq. (2.30)). Like the Hartree-Fock equations, the KS equations have to be solved in a self-consistent manner, since the Kohn-Sham potential  $v_{KS}$  represents an effective potential which depends on the density itself.

### 2.2.3 Density Functional Perturbation Theory

Many physical properties of solids and molecules can be considered as the response of the system to an external perturbation. In standard perturbation theory, a small perturbation operator  $\lambda\mathcal{H}^{(1)}$  is added to the Hamiltonian  $\mathcal{H}$  of a system and its action is evaluated. Instead of the conventional Hamiltonian formulation of perturbation theory of DFT,<sup>136–138</sup> a more general variational approach is presented in this section.<sup>139</sup> The starting point is KS-DFT. The variational principle states for an unperturbed system in the electronic ground state that the ground state KS orbitals are those that minimize the KS energy. Perturbing a KS ground state system, the electronic structure adjusts in such a way that the perturbed energy is minimized again. Thus, the perturbed states are also accessible by a variational approach. This theoretical concept is referred to as density functional perturbation theory (DFPT). The discussion of the DFPT equations in this section is taken from ref. 139, which can be consulted for detailed information.

We start from the density matrix formed by non-orthogonal KS orbitals:

$$\rho(\mathbf{r}, \mathbf{r}') = \sum_{i,j} \phi_i^*(\mathbf{r}) (\mathbf{S})_{ij}^{-1} \phi_j(\mathbf{r}'), \quad (2.44)$$

where  $(\mathbf{S})^{-1}$  is the inverse of the overlap matrix  $\mathbf{S}_{ij} = \langle \phi_i, \phi_j \rangle$ . The KS density functional from equation (2.37) reads now as:

$$\varepsilon_{KS}[\rho(\mathbf{r}, \mathbf{r}')] = \frac{1}{2} \int d\mathbf{r} \int d\mathbf{r}' \delta(\mathbf{r} - \mathbf{r}') \nabla_{\mathbf{r}}^2 \rho(\mathbf{r}, \mathbf{r}') + \frac{1}{2} \int d\mathbf{r} \frac{|\rho(\mathbf{r}, \mathbf{r}')|^2}{|\mathbf{r} - \mathbf{r}'|} + \varepsilon_{xc}[\rho] + \varepsilon_{ext}[\rho], \quad (2.45)$$

where the density  $n(\mathbf{r})$  is simply:

$$n(\mathbf{r}) = \rho(\mathbf{r}, \mathbf{r}). \quad (2.46)$$

The minimum of the functional  $\varepsilon_{KS}$  is  $\varepsilon_{KS}^{(0)}[\{|\phi_i^{(0)}\rangle\}] = \varepsilon^{(0)}$ , assuming that (at the minimum) the orbitals are orthogonal ( $\langle \phi_i^{(0)}, \phi_j^{(0)} \rangle = \delta_{ij}$ ).

Starting from  $\varepsilon_{KS}^{(0)}[\{|\phi_i\rangle\}]$ , we add a perturbation functional of arbitrary form  $\varepsilon_{KS}^{pert}[\{|\phi_i\rangle\}]$  multiplied by a small perturbative parameter  $\lambda$ :

$$\varepsilon_{KS}^{tot}[\{|\phi_i\rangle\}] = \varepsilon_{KS}^{(0)}[\{|\phi_i\rangle\}] + \lambda \varepsilon_{KS}^{pert}[\{|\phi_i\rangle\}]. \quad (2.47)$$

This is meant to represent the interaction with a static, but otherwise arbitrarily complex external field. The perturbation parameter  $\lambda$  is an infinitesimally small auxiliary variable that helps separate different orders of the response with respect to the perturbation and does not occur in the final expressions.

The total functional  $\varepsilon_{KS}^{tot}[\{|\phi_i\rangle\}]$  will have a minimum which we shall expand perturbatively:

$$E_{KS}^{tot} = E^{(0)} + \lambda E^{(1)} + \lambda^2 E^{(2)} + \mathcal{O}(\lambda^3) \quad (2.48)$$

Similarly, the KS orbitals that minimize  $E_{KS}^{tot}[\{|\phi_i\rangle\}]$  can be expanded in terms of  $\lambda$

$$\phi_i = \phi_i^{(0)} + \lambda \phi_i^{(1)} + \lambda^2 \phi_i^{(2)} + \dots \quad (2.49)$$

Inserting equation (2.48) and (2.49) into (2.47) yields:

$$\varepsilon_{KS}^{tot}[\{|\phi_i\rangle\}] = \varepsilon_{KS}^{(0)}[\{\phi_i^{(0)} + \lambda \phi_i^{(1)} + \dots\}] + \lambda \varepsilon_{KS}^{pert}[\{\phi_i^{(0)} + \lambda \phi_i^{(1)} + \dots\}] \quad (2.50)$$

Following the derivations of ref. 139, we obtain an expression for the second order energy correction for the limit of orthogonal orbitals ( $\mathbf{S}_{ij} = \delta_{ij}$ ):

$$\begin{aligned} E^{(2)} &= \sum_m \sum_{m'} \left[ \langle \phi_m^{(1)}, \left( \hat{\mathcal{H}}_{KS}^{(0)} \delta_{mm'} - \langle \phi_m^{(0)}, \hat{\mathcal{H}}_{KS}^{(0)} \phi_m^{(0)} \rangle \right) \phi_{m'}^{(1)} \rangle \right] + \frac{1}{2} \int \int d^3r d^3r' K(\mathbf{r}, \mathbf{r}') n^{(1)}(\mathbf{r}) n^{(1)}(\mathbf{r}') \\ &+ \sum_m \left\langle \phi_m^{(1)}, \left( \frac{\delta E^{pert}}{\delta \langle \phi_m^{(0)} |} + \frac{\delta E^{pert}}{\delta |\phi_m^{(0)} \rangle} \right) \phi_m^{(1)} \right\rangle, \end{aligned} \quad (2.51)$$

with the Hartree-exchange correlation kernel  $K(\mathbf{r}, \mathbf{r}') = \frac{\delta(V_H(\mathbf{r}) + V_{xc}(\mathbf{r}))}{\delta n(\mathbf{r}')}$ , and the unperturbed KS Hamiltonian  $\hat{\mathcal{H}}_{KS}^{(0)}$  according to equation (2.42). In case of canonical orbitals

$$\langle \phi_{m'}^{(0)}, \hat{\mathcal{H}}_{KS}^{(0)} \phi_m^{(0)} \rangle = \varepsilon_m^{(0)} \delta_{mm'} \quad (2.52)$$

equation (2.51) reads as

$$E^{(2)} = \sum_m \langle \phi_m^{(1)}, (\hat{\mathcal{H}}_{KS}^{(0)} - \epsilon_m^{(0)}) \phi_m^{(1)} \rangle + \frac{1}{2} \int \int d^3r d^3r' K(\mathbf{r}, \mathbf{r}') n^{(1)}(\mathbf{r}) n^{(1)}(\mathbf{r}') \quad (2.53)$$

$$+ \sum_m \left\langle \phi_m^{(1)}, \left( \frac{\delta E^{pert}}{\delta \langle \phi_m^{(0)} |} + \frac{\delta E^{pert}}{\delta | \phi_m^{(0)} \rangle} \right) \phi_m^{(1)} \right\rangle. \quad (2.54)$$

In order to retain the normalization of the wave function, the perturbed  $\phi^{(1)}$  and the unperturbed  $\phi^{(0)}$  wave functions have to fulfill the following orthogonality relation:

$$\langle \phi_i^{(0)}, \phi_j^{(1)} \rangle = 0 \quad \forall i, j. \quad (2.55)$$

Minimization of the second order energy  $E^{(2)}$  with respect to this orthogonality condition using the methods of Lagrange multiplier leads to the following set of Sternheimer-like equations:<sup>139</sup>

$$- \left( \hat{\mathcal{H}}_{KS}^{(0)} - \epsilon_m^{(0)} \right) | \phi_m^{(1)} \rangle = \hat{P}_e \left( \int d^3r' K(\mathbf{r}, \mathbf{r}') n^{(1)}(\mathbf{r}') | \phi_m^{(0)} \rangle + \frac{\delta \epsilon^{pert}}{| \phi_m^{(0)} \rangle} \right), \quad (2.56)$$

where  $P_e = 1 - \sum_j | \phi_j^{(0)} \rangle \langle \phi_j^{(0)} |$ . This set of equations has to be solved iteratively because also the right-hand side depends on the  $\{ | \phi^{(1)} \rangle \}$  via the perturbation density  $n^{(1)}$ . Introducing the variations of the Hartree potential  $V_H^{(1)}$  and the exchange-correlation potential  $V_{xc}^{(1)}$  induced by the change of the density

$$\int d^3r' K(\mathbf{r}, \mathbf{r}') n^{(1)}(\mathbf{r}') = V_H^{(1)}(\mathbf{r}) + V_{xc}^{(1)}(\mathbf{r}), \quad (2.57)$$

and inserting this expression into (2.56) yields

$$- \left( \hat{\mathcal{H}}_{KS}^{(0)} - \epsilon_m^{(0)} \right) | \phi_m^{(1)} \rangle = \hat{P}_e \left( V_H^{(1)}(\mathbf{r}) | \phi_m^{(0)} \rangle + V_{xc}^{(1)}(\mathbf{r}) | \phi_m^{(0)} \rangle + \frac{\delta \epsilon^{pert}}{| \phi_m^{(0)} \rangle} \right). \quad (2.58)$$

In summary by combining equations (2.51) and (2.58), we obtain the second order corrections to the energy by applying the variation principle to the  $\{ | \phi^{(1)} \rangle \}$  under the constraints from equation (2.55). This is a special case of the more general “ $2n + 1$ ” theorem which states that using a variational approach the  $(2n)$ -th and the  $(2n + 1)$ -th order energy are accessible from the  $n$ -th order wavefunction.<sup>137</sup>

## 2.3 The Linear Density-Density Response Function (LDDRF)

From a microscopic point of view, condensed phase systems are described by a large number of interacting particles. As a result of these interactions, the electron density of the individual molecules is disturbed and differs from the ground state density of an isolated molecule. Following this line of approach, the calculation of accurate electrostatic interaction energies should take these polarization effects into account.<sup>37, 140-142</sup> The static linear density-density response function (LDDRF) from equation (2.60) provides a tool to calculate the fully self-consistent density response of a molecule to arbitrary potentials (in practice: the electrostatic field arising from nearby molecules), corresponding to exact polarization at all multipolar orders.<sup>34-37</sup>

$$\hat{\mathbf{T}} : V_{\text{pert}} \mapsto n^{\text{resp}} \quad (2.59)$$

$$\text{with } n^{\text{resp}}(\mathbf{r}) = \int \chi(\mathbf{r}, \mathbf{r}') V_{\text{pert}}(\mathbf{r}') d^3 r' \quad (2.60)$$

A diagram illustrating the principle of the distortion of the electron density of a molecule due to the perturbing potential arising from a neighboring molecule is presented in Figure 2.2. The calculation of molecular density responses becomes trivial, once the kernel  $\chi(\mathbf{r}, \mathbf{r}')$  is explicitly known. However, the explicit calculation of the six-dimensional kernel  $\chi(\mathbf{r}, \mathbf{r}')$  is not possible due to the storage limitations of real-world computing systems. A representation of  $\chi(\mathbf{r}, \mathbf{r}')$  on a grid with only 100 grid points in each dimension would already result in the necessity to store  $10^{12}$  floating point numbers. More sophisticated approaches are necessary to allow for efficient representation of  $\chi(\mathbf{r}, \mathbf{r}')$ .

The following three chapters will highlight important aspects of the LDDRF. The next section provides an overview of a recently published efficient representation of the LDDRF. In addition, a careful derivation of equation (2.60) (including a discussion of the properties of the LDDRF) is given. The articles that constitute this thesis make extensive use of the matrix representation of linear operators with respect to different (non-orthogonal) bases of the domain and the image. For this reason, the fundamentals concerning the change-of-basis and the basis representation of operators and vectors are summarized in the last chapter of this part of the introduction.

### 2.3.1 Different Representations of the LDDRF

From a mathematical point of view, equation (2.60) constitutes an operator  $\hat{\mathbf{T}}$  which is a linear, self-adjoint, positive and compact integral transform with kernel  $\chi(\mathbf{r}, \mathbf{r}')$ .<sup>1</sup> Such operators can be represented by their eigensystem (i.e. the entire set of eigenvalues  $\lambda_i$  and eigenfunctions  $\chi_i(\mathbf{r})$ ) according to

$$\hat{\mathbf{T}} = \sum_{i=1}^{\infty} |\chi_i\rangle \lambda_i \langle \chi_i|. \quad (2.61)$$

In section 2.3.4, it is shown that  $\hat{\mathbf{T}}$  is a Hilbert-Schmidt operator. For this reason, the sequence of the squared eigenvalues of this operator decays fast enough to zero ( $\lim_{i \rightarrow \infty} \lambda_i = 0$ ), ensuring that the sum of eigenvalues does not diverge

<sup>1</sup>See section 2.3.3 for the derivation of the properties of the operator  $\hat{\mathbf{T}}$ .

( $\sum_i^\infty \lambda_i)^2 < \infty$ ). Thus, the eigensystem representation of the LDDRF can be truncated to the desired numerical precision of the molecular density response. Scherrer and Sebastiani demonstrated that for the case of water, 5000 eigenfunctions are sufficient for the calculation of converged response densities. The eigenvalues and eigenfunctions can be obtained using a Lanczos algorithm (cf. section 2.3.4). In this approach, a self-consistent solution of a Sternheimer-like equation using density functional perturbation theory (cf. section 2.2.3) has to be performed for each eigenstate. For details, see references 141 and 139 and section 2.3.4.

### Efficient representations

This thesis aims to answer the question of how to derive and understand a more efficient representation of the LDDRF. The key to the dimensional reduction of the eigensystem representation emerges from the mismatch between the physically meaningful description and the formal mathematical description of the domain and image of the linear operator  $\hat{\mathbf{T}}$ . From a mathematical point of view,  $\hat{\mathbf{T}}$  is a self-adjoint operator. This means that both the domain and the image are isometrically isomorphic. This becomes immediately apparent from the eigensystem decomposition (cf. eq. (2.61)), because both the image and the domain of  $\hat{\mathbf{T}}$  are generated by the set of the eigenfunctions. Assuming the kernel  $\chi(\mathbf{r}, \mathbf{r}') \in L^2[X \times X]$  to be a square-integrable function with compact support  $X \times X \subset \mathbb{R}^3 \times \mathbb{R}^3$ , domain and image can be identified as the vector space of square-integrable functions  $D(\hat{\mathbf{T}}) := L^2[X]$  with compact support  $X \subset \mathbb{R}^3$ .

From a physical point of view, the entity of meaningful perturbing potentials generates the domain  $\mathcal{P}$  and the entity of possible response densities due to  $V_{\text{pert}} \in \mathcal{P}$  generates the image  $\mathcal{Z}$ . It is immediately apparent that perturbing potentials and molecular response densities are quite different functions. While molecular response densities are highly oscillating functions which resemble linear combinations of atomic basis functions, the perturbing potential arising from a non-overlapping charge density can be expressed efficiently by rather smooth functions, e.g. by an Taylor expansion using monomials.<sup>143</sup>

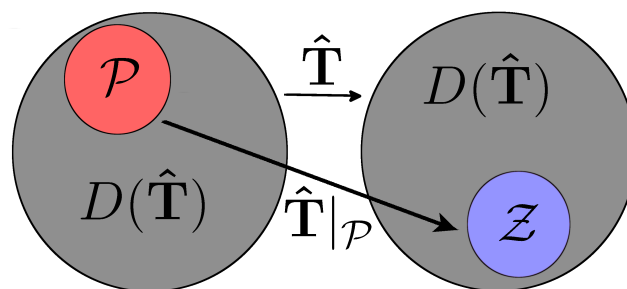


FIGURE 2.1: Domain and image of the linear operator  $\hat{\mathbf{T}}$ , and domain  $\mathcal{P}$  and image  $\mathcal{Z}$  of the restriction  $\mathcal{T} := \hat{\mathbf{T}}|_{\mathcal{P}}$ . Please note that  $\mathcal{P}$  and  $\mathcal{Z}$  have a trivial intersection ( $\mathcal{P} \cap \mathcal{Z} = \emptyset$ ).

The mismatch of the vector spaces is illustrated in Figure 2.1. In this figure, the vector spaces of physically relevant perturbing potentials  $\mathcal{P}$  and response densities  $\mathcal{Z}$  are subspaces (with trivial intersection) of the overall domain and image of  $\hat{\mathbf{T}}$ .

Two principal choices of a basis are now possible:

1. We can attempt to express vectors from the two different vector spaces  $\mathcal{P}$  and  $\mathcal{Z}$  within a single basis. This is done within the eigensystem representation given



by equation (2.61). In this case, a large number of basis functions is required to describe the perturbing potentials and the molecular response densities.

- Using information concerning the shape of the perturbing potentials and response densities, we can choose an especially adapted basis for the efficient representation of these functions. Only a small number of basis functions (compared to the first case) is necessary to describe perturbing potentials and molecular response densities alone. Instead of the calculation of an infinite-dimensional operator  $\hat{\mathbf{T}}$ , we can confine ourselves to the calculation of the restriction of the linear operator  $\hat{\mathbf{T}}$  between the two low-dimensional vector spaces  $\mathcal{P}$  and  $\mathcal{Z}$ .

The second item is the starting point for all attempts in this thesis to identify an efficient representation of the LDDRF. In this regard, we wish to formalize the idea of the second item in a more rigid mathematical framework. Perturbing potentials generated by a non-overlapping charge density can be straightforwardly expanded in a few terms within a multipole expansion (schematic illustration in Figure 2.2).<sup>143</sup> As a result, we assume that every perturbing potential can be expanded using  $N$  basis functions  $\{|P_1\rangle, |P_2\rangle, \dots, |P_N\rangle\}$  to a desired precision:

$$V_{\text{pert}}(\mathbf{r}) \approx \sum_{n=1}^N c_n P_n(\mathbf{r}). \quad (2.62)$$

This assumption will be verified for specific examples in the article [Dreßler et al., *J. Comput. Chem.*, **2019**.]<sup>6</sup> in this thesis, and special choices of  $\{|P_1\rangle, |P_2\rangle, \dots, |P_N\rangle\}$  are discussed in this reference.

In order to emphasize that the vector space of the physically meaningful perturbing potential is spanned by a small number  $N$  of basis functions, we will replace the symbol  $\mathcal{P}$  by  $\mathcal{P}[\mathbf{N}]$ .

The vector space of all possible response densities due to  $V_{\text{pert}} \in \mathcal{P}[\mathbf{N}]$  is spanned by the images  $\{|\tilde{n}_1\rangle := \hat{\mathbf{T}}|P_1\rangle, |\tilde{n}_2\rangle := \hat{\mathbf{T}}|P_2\rangle, \dots, |\tilde{n}_N\rangle := \hat{\mathbf{T}}|P_N\rangle\}$  and is at the most  $N$ -dimensional, which motivates the renaming of  $\mathcal{Z}$  as  $\mathcal{Z}[\mathbf{N}]$ .

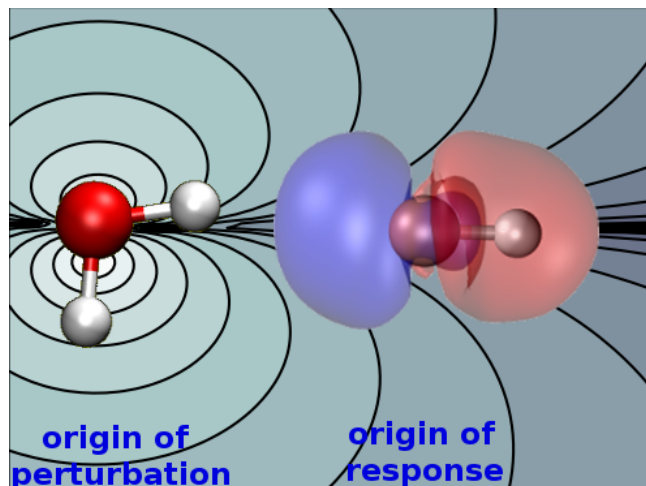


FIGURE 2.2: Schematic illustration of the response density of the water molecule (right) due to a perturbing water molecule (left). The potential originating from the left water molecule can be expanded at the responding (right) water molecule using a few basis functions. This figure is reproduced from article [Dreßler, Sebastiani, *Int. J. Quantum Chem.*, **2020**.] in this thesis.

The underlying idea of all attempts in this thesis to achieve a more efficient representation of the LDDRF is to consider the operator  $\hat{\mathbf{T}}$  not as an (infinite) endomorphism  $D(\hat{\mathbf{T}}) \rightarrow D(\hat{\mathbf{T}})$  but rather an ( $N$ -dimensional) homomorphism  $\mathcal{T} : \mathcal{P}[\mathbf{N}] \rightarrow \mathcal{Z}[\mathbf{N}]$ . In other words  $\mathcal{T} := \hat{\mathbf{T}}|_{\mathcal{P}[\mathbf{N}]}$  is the restriction of  $\hat{\mathbf{T}}$  to  $\mathcal{P}$ . Taking into account the small dimension of  $\mathcal{P}[\mathbf{N}]$  and  $\mathcal{Z}[\mathbf{N}]$  compared to the dimension of  $D(\hat{\mathbf{T}}) = L^2[X]$ , the significant reduction of the dimensionality is apparent.

Starting from this mathematical representation, the calculation of molecular response densities seems to be very simple: the density response due to perturbing potentials  $V_{\text{pert}} \in \mathcal{P}[\mathbf{N}] \Leftrightarrow V_{\text{pert}}(\mathbf{r}) = \sum_{n=1}^N c_n P_n(\mathbf{r})$  can be calculated via:

$$\hat{\mathbf{T}}|V_{\text{pert}}\rangle = \sum_{i=1}^N c_i \underbrace{\hat{\mathbf{T}}|P_i\rangle}_{:=|\tilde{n}_i\rangle} = \sum_{i=1}^N c_i |\tilde{n}_i\rangle. \quad (2.63)$$

Knowledge of the density response  $|\tilde{n}_i\rangle$  due to a specific perturbing potential  $|P_i\rangle$  is the limiting factor for the application of equation (2.63). The calculation of these density responses  $|\tilde{n}_i\rangle$  is possible via an explicit single self-consistent solution of a Sternheimer-like equation in terms of DFPT (cf. section 2.2.3).

However, it is not straightforward to use this approach for the calculation of molecular density responses. While the formal mathematical representation is correct and provides a versatile tool to understand the basic idea of the dimensionality reduction, special care has to be taken with the choice of the basis of the perturbing potentials  $\{|P_1\rangle, |P_2\rangle, \dots, |P_N\rangle\}$ .

The vector space of the physically meaningful perturbing potentials  $\mathcal{P}[\mathbf{N}]$  is a subspace of the domain  $D(\hat{\mathbf{T}})$  of  $\hat{\mathbf{T}}$  and hence a subspace of the vector space generated from the eigenfunctions  $D(\hat{\mathbf{T}}) = \text{span}(|\chi_1\rangle, |\chi_2\rangle, \dots)$ . Thus, the support  $\text{supp}(\mathcal{P}) \subset \mathbb{R}^3$  of the set of functions  $\{|P_1\rangle, |P_2\rangle, \dots, |P_N\rangle\}$  has to be a subset of the support  $X \subset \mathbb{R}^3$  of the set of eigenfunctions  $\{|\chi_1\rangle, |\chi_2\rangle, \dots\}$ .<sup>2</sup>

However, the shape of the kernel  $\chi(\mathbf{r}, \mathbf{r}')$  (and hence also the support of the eigenfunctions) is different for every molecule. In contrast with this, the set of basis functions  $\{|P_1\rangle, |P_2\rangle, \dots, |P_N\rangle\}$  can be chosen freely from suitable candidates such as Racah-normalized regular solid harmonics or other polynomials, but we want to keep this set of basis functions fixed for all molecules, i.e. the perturbing potential is (independently from the specific molecule) always described by the same set of basis functions. This means that the support  $\text{supp}(\mathcal{P})$  is fixed and constant for all molecules, while the support  $X$  of the eigenfunctions  $\{|\chi_1\rangle, |\chi_2\rangle, \dots\}$  is individual for each molecule. As a result,  $\mathcal{P}[\mathbf{N}] = \{|P_1\rangle, |P_2\rangle, \dots, |P_N\rangle\}$  can not be a subspace of  $D(\hat{\mathbf{T}}) = \text{span}(|\chi_1\rangle, |\chi_2\rangle, \dots)$ . Expanding the perturbing potentials in a different vector space compared to the eigenspace of the linear operator  $\hat{\mathbf{T}}$  will lead to altered response densities. For this reason, equation (2.63) is not suited for actual calculations of the response density.

However, the problems induced by the mismatch of the supports can be circumvented by the moment expansion of the LDDRF, which was first introduced by Scherrer and Sebastiani in 2016.<sup>144</sup> They utilized a unitary transformation of the eigenfunctions  $\{|\chi_i\rangle \mid i \in \mathbb{N}\}$  which condenses the physically relevant information into a few moment generating states  $\{|\xi_1\rangle, |\xi_2\rangle, \dots, |\xi_N\rangle\}$ . The transformed states - denoted as moment-expanded states - fulfill the following partial orthogonality condition with

<sup>2</sup>We define the support  $\text{supp}(\{|\chi_1\rangle, |\chi_2\rangle, \dots\})$  of a set of functions as the union of the supports  $\cup_{i=1} \text{supp}(\{|\chi_i\rangle\})$  of the sole functions .

respect to the basis functions of the perturbing potential  $\{|P_1\rangle, |P_2\rangle, \dots, |P_N\rangle\}$ :

$$\langle \xi_i, P_n \rangle = 0 \quad \forall i > n \quad (2.64)$$

Due to the partial orthogonality relation, only  $N$  moment-expanded states  $\{|\xi_1\rangle, |\xi_2\rangle, \dots, |\xi_N\rangle\}$  have a non-zero overlap with the  $N$ -th basis function  $|P_N\rangle$  of the perturbing potential. Thus, the full density response due to perturbing potentials  $V_{\text{pert}} \in \mathcal{P}[\mathbf{N}]$  is reduced to a sum over  $N$  terms:

$$\hat{\mathbf{T}}|V_{\text{pert}}\rangle = \sum_{i=1}^N |\xi_i\rangle \langle \xi_i, V_{\text{pert}}\rangle. \quad (2.65)$$

The moment expansion does not use a specific basis expansion of the perturbing potentials. Only the overlap of the moment-expanded states and the perturbing potential contributes to the sum in equation (2.65). The overlap integral is taken automatically at the vector space with correct support, because the moment-expanded functions are only non-zero within the support of the eigenfunctions.

### 2.3.2 Derivation of the LDDRF

The LDDRF is also known as electronic susceptibility, non-local electronic density susceptibility and the screened electronic dielectric response function. The LDDRF is a very complex and sophisticated subject and it is perhaps the most important response function from the time-dependent density functional theory (TDDFT) point of view.<sup>38,39</sup> It gives the linear frequency dependent response of the density to an external scalar potential. After its explicit determination, it can then be utilized to obtain the first-order response of all properties derivable from the density with respect to any scalar field e.g., the polarizability.<sup>40</sup> The calculation of the poles of the LDDRF enables a direct approach for the calculation of electronic excitations.

The LDDRF is also important for the adiabatic-connection fluctuation-dissipation theorem and symmetry adapted perturbation theory.<sup>145–149</sup> It is employed in G0W0,<sup>150–152</sup> fluctuation-dissipation density functional theory,<sup>153–155</sup> van-der-Waals<sup>156–159</sup> or random phase approximation<sup>160–163</sup> and beyond random phase approximation<sup>164–167</sup> calculations. In addition, polarization effects were included in QM/MM free-energy calculations.<sup>168–170</sup> In a complementary manner it is used for the qualitative prediction of properties such as aromaticity<sup>171</sup> and charge delocalization<sup>172,173</sup> within the context of conceptual chemistry.<sup>174,175</sup> The LDDRF was also used to investigate alchemical derivative evaluations.<sup>176–178</sup> In the next subsection, we will demonstrate how to derive the LDDRF from general linear response theory.

#### General linear response theory

Linear response theory is a commonly used method that is applied when considering the response to a weak perturbation, such as the density response due to the perturbing field arising from the electron density of a neighboring molecule.

The discussion of general linear response theory in this section will closely follow reference 38. We start from the expectation value of a quantum mechanical operator  $\hat{\alpha}$ :

$$\alpha_0 = \langle \Psi_0, \hat{\alpha} \Psi_0 \rangle, \quad (2.66)$$

where  $\Psi_0$  is the ground-state many-body wavefunction, which fulfills  $\hat{H}_0\Psi_0 = E_0\Psi_0$  where  $\hat{H}_0$  is the static Hamiltonian and  $E_0$  the lowest eigenvalue.

Let  $F(t)$  be an external field that couples to an observable  $\hat{\beta}$ . At time  $t_0$ , we switch on a time-dependent perturbation according to

$$\hat{H}_1(t) = F(t)\hat{\beta}, \quad t \geq t_0, \quad (2.67)$$

which leads to a perturbed system and a time-dependent expectation value of  $\hat{\alpha}$ :

$$\alpha(t) = \langle \Psi(t), \hat{\alpha}\Psi(t) \rangle, \quad t \geq t_0. \quad (2.68)$$

The response  $\alpha(t) - \alpha_0$  of  $\hat{\alpha}$  to the perturbation from equation (2.67) can be expanded in terms of powers of the field  $F(t)$ :

$$\alpha(t) - \alpha_0 = \alpha_1(t) + \alpha_2(t) + \dots, \quad (2.69)$$

where  $\alpha(t)$  is the linear response,  $\alpha_2(t)$  is the quadratic (second-order) response,  $\alpha_3(t)$  is the third-order response and so on. By introducing a linear response function  $\chi_{\alpha\beta}(t - t')$  it is demonstrated in Appendix A that the linear response  $\alpha_1(t)$  can be calculated using the integral equation

$$\alpha_1(t) = \int_{-\infty}^{\infty} dt' \chi_{\alpha\beta}(t - t') F(t'), \quad (2.70)$$

where the lower integration limit  $t_0$  is replaced by  $-\infty$  since the external field  $F(t)$  is zero for all times before  $t_0$ . By replacing the time  $t$  with the frequency  $\omega$  via a Fourier transformation, the linear-response equation in frequency space is derived as

$$\alpha_1(\omega) = \chi_{\alpha\beta}(\omega) F(\omega). \quad (2.71)$$

As a final result of Appendix A, an explicit representation of the frequency-dependent response function  $\chi_{\alpha\beta}(\omega)$  is obtained according to

$$\chi_{\alpha\beta}(\omega) = \lim_{\eta \rightarrow 0^+} \sum_{n=1}^{\infty} \left\{ \frac{\langle \Psi_0, \hat{\alpha}\Psi_n \rangle \langle \Psi_n, \hat{\beta}\Psi_0 \rangle}{\omega - \Omega_n + i\eta} - \frac{\langle \Psi_0, \hat{\beta}\Psi_n \rangle \langle \Psi_n, \hat{\alpha}\Psi_0 \rangle}{\omega - \Omega_n + i\eta} \right\}. \quad (2.72)$$

Equation (2.72) describes a representation of the linear response function  $\chi_{\alpha\beta}(\omega)$  as a sum over states. We call this expression Lehmann representation. This is an important part of response theory because it explicitly shows how a frequency-dependent perturbation couples to the excitation spectrum of a system. Please note that the response function  $\chi_{\alpha\beta}(\omega)$  depends only on properties of the system in the absence of the probe.

### Lehmann Representation of the LDDRF

The most important response function in the context of TDDFT, namely the density-density response, is obtained by setting  $\hat{\alpha}$  and  $\hat{\beta}$  to the density operator, which is defined for an  $N$ -electron system as

$$\hat{n}(\mathbf{r}) = \sum_{l=1}^N \delta(\mathbf{r} - \mathbf{r}_l). \quad (2.73)$$

In this case, the external perturbation is a scalar potential  $v_1(\mathbf{r}', t)$ , switched on at  $t_0$ , which is coupled with the density operator from equation (2.73)

$$\hat{H}_1 = \int d^3r' v_1(\mathbf{r}', t) \hat{n}(\mathbf{r}'). \quad (2.74)$$

Then, in analogy with equation (2.70), the linear response of the density is given by

$$n_1(\mathbf{r}, t) = \int_{-\infty}^{\infty} dt' \int d^3r' \chi_{nn}(\mathbf{r}, \mathbf{r}', t - t') v_1(\mathbf{r}', t'). \quad (2.75)$$

To relate equation (2.75) to the general expression (2.70), we note that the linear responses to different perturbations can simply be added independently. Taking this into account, we can relate equation (2.75) to the general expression (2.70). The perturbing Hamiltonian from equation (2.74) can be viewed as a sum of perturbations  $v_1(\mathbf{r}', t') \hat{n}(\mathbf{r}')$  that are of the form of equation (2.67). Each of these perturbations gives rise to a retarded density response at all  $\mathbf{r}$ , and all of these are then integrated over space in equation (2.75). The frequency-dependent density response can be obtained again by Fourier transformation in analogy with equation (2.71)

$$n_1(\mathbf{r}, \omega) = \int d^3r' \chi_{nn}(\mathbf{r}, \mathbf{r}', \omega) v_1(\mathbf{r}', \omega). \quad (2.76)$$

The Lehmann representation (cf. (2.72)) is given for the special case of the density-density response function by

$$\chi_{nn}(\mathbf{r}, \mathbf{r}', \omega) = \lim_{\eta \rightarrow 0^+} \sum_{n=1}^{\infty} \left\{ \frac{\langle \Psi_0, \hat{n}(\mathbf{r}) \Psi_n \rangle \langle \Psi_n, \hat{n}(\mathbf{r}') \Psi_0 \rangle}{\omega - \Omega_n + i\eta} - \frac{\langle \Psi_0, \hat{n}(\mathbf{r}') \Psi_n \rangle \langle \Psi_n, \hat{n}(\mathbf{r}) \Psi_0 \rangle}{\omega - \Omega_n + i\eta} \right\}. \quad (2.77)$$

### 2.3.3 Basic Symmetries and Analytic Properties of the LDDRF

We start this discussion of the properties of the LDDRF by considering basic symmetries. The density response  $n_1(\mathbf{r}, t)$  has to be real, which also implies that  $\chi(\mathbf{r}, t, \mathbf{r}', t')$  is a real function. As a consequence, we find from the Lehmann representation from equation (2.77):

$$\chi_{nn}(\mathbf{r}, \mathbf{r}', \omega) = \chi_{nn}^*(\mathbf{r}, \mathbf{r}', -\omega), \quad (2.78)$$

which implies that the real part of  $\chi_{nn}(\mathbf{r}, \mathbf{r}', \omega)$  is an even function of the frequency whereas the imaginary part is an odd function.

In addition, the following symmetry expression is immediately apparent from sum-over-states representation in equation (2.77)

$$\chi_{nn}(\mathbf{r}, \mathbf{r}', \omega) = \chi_{nn}^*(\mathbf{r}', \mathbf{r}, \omega). \quad (2.79)$$

In the following, we will confine ourselves to the static density response function  $\chi(\mathbf{r}, \mathbf{r}') := \chi(\mathbf{r}, \mathbf{r}', \omega = 0)$  by setting  $\omega$  to zero. In contrast with its dynamic variant, the static density response function does not possess poles. From the Lehmann representation, we can suppose  $\chi(\mathbf{r}, \mathbf{r}') \in L^2(X \times X)$  for  $X \subset \mathbb{R}^3$  to be a square-integrable function.

The physical origin for the distortion of a charge density due to a perturbing potential is the minimization of its energy. The energy difference due to polarization  $\Delta E^{pol}$  is always negative (or always positive by the opposing definition) and can be

calculated as:

$$\begin{aligned} 0 \geq \Delta E^{pol} &= \int d^3r v_{\text{pert}}(\mathbf{r}) n^{\text{resp}}(\mathbf{r}) = \int d^3r v_{\text{pert}}(\mathbf{r}) \int d^3r' \chi(\mathbf{r}, \mathbf{r}') v_{\text{pert}}(\mathbf{r}') \\ &= \langle v_{\text{pert}}, \hat{\mathbf{T}} v_{\text{pert}} \rangle \end{aligned} \quad (2.80)$$

The last equation holds for all perturbing potentials  $v_{\text{pert}} \in L^2(X)$ . As a result, the linear operator  $\hat{\mathbf{T}}$  defined in equation (2.60) is negative-definite (or positive-definite by the opposing definition).

### 2.3.4 Iterative Spectral Decomposition of the Static LDDRF

This thesis uses the static LDDRF  $\chi(\mathbf{r}, \mathbf{r}') := \chi(\mathbf{r}, \mathbf{r}', \omega = 0)$ . Its explicit non-local real-space representation has only been the subject of a small number of studies.<sup>141, 179–181</sup>

The analytical properties of the LDDRF are discussed in section 2.3.3. For  $X \subset \mathbb{R}^3$  and a square-integrable kernel  $\chi(\mathbf{r}, \mathbf{r}') \in L^2[X \times X]$ , equation (2.60) defines a linear Hilbert-Schmidt operator  $\hat{\mathbf{T}} : L^2[X] \rightarrow L^2[X]$ , which is compact.<sup>182, 183</sup>

The assumption of  $V_{\text{pert}}$  and  $n^{\text{resp}}$  as a square-integrable function is a very weak restriction and includes all physically meaningful perturbing potentials or response densities with support  $X \subset \mathbb{R}^3$ .

In section 2.3.3, we showed that the kernel is also symmetric ( $\chi(\mathbf{r}, \mathbf{r}') = \chi(\mathbf{r}', \mathbf{r}) \forall \mathbf{r}, \mathbf{r}' \in X$ ) and positive-definite ( $\int \int \chi(\mathbf{r}, \mathbf{r}') f(\mathbf{r}') f(\mathbf{r}) d^3r d^3r' \geq 0 \forall f(\mathbf{r}) \in D(\hat{\mathbf{T}})$ ), which also makes the linear operator  $\hat{\mathbf{T}}$  from equation (2.60) a self-adjoint and positive-definite one.

According to the spectral theorem for self-adjoint, positive and compact operators, the operator  $\hat{\mathbf{T}}$  can be expressed by its eigensystem given in equation (2.61). The calculation of the eigenvalues  $\lambda_i$  and the eigenfunctions  $\chi_i(\mathbf{r})$  of the linear operator  $\hat{\mathbf{T}}$  is possible via an iterative diagonalization of the static response function. Reference 141 provides a detailed description of the underlying Hermitian Lanczos algorithm for the spectral decomposition. Formally, the numerical problem to be solved is the determination of the eigenvectors for the largest eigenvalues of a Hermitian operator  $\hat{\mathbf{T}}$  with only its action on a vector available.

From a conceptual point of view, the key for the application of the Hermitian Lanczos algorithm is the possibility of calculating the explicit density response to arbitrary static perturbing potentials. This can be done in terms of DFPT.<sup>136–139</sup> In this regard, the density response can be calculated via a self-consistent solution of a Sternheimer-like equation:

$$(\hat{H}^{(0)} - \epsilon^{(0)}) \phi_i^{(1)} = -\hat{P}_e \left( V^{(1)} \phi_i^{(0)} \right), \quad (2.81)$$

where  $V^{(1)}$  is an arbitrary perturbing potential. A detailed derivation and explanation of equation (2.81) is given in section 2.2.3.

In principle, the iterative spectral decomposition can also be used to calculate the representation of the full frequency-dependent response function, which enables an efficient pathway for RPA<sup>151, 163, 184, 185</sup> and post-RPA calculations.<sup>167</sup>

### 2.3.5 Representation of Linear Operators

The efficient representations of the static LDDRF presented in this thesis can be understood in terms of their representing matrices. By careful analysis of the shape of

these matrices, the properties of the efficient representations become immediately apparent. In the next section, we will derive the framework for calculating the transition matrix elements of a linear operator with respect to non-orthogonal basis sets. First, we discuss how to calculate the coordinates of a vector according to an arbitrary basis. In this section, we will follow references 6 and 5.

### Resolution of the identity and coordinate vectors

From now on, we define  $V$  and  $W$  to be vector spaces with a scalar product  $\langle \cdot, \cdot \rangle$ . Let  $\{|u_1\rangle, |u_2\rangle, \dots\}$  be a (not necessarily orthogonal) basis of  $V$  such that every  $v \in V$  can be expressed as

$$v = a_1|u_1\rangle + a_2|u_2\rangle + \dots = \sum_{i=1} |u_i\rangle a_i, \quad (2.82)$$

with unique expansion coefficients  $\{a_1, a_2, \dots\}$ . We also refer to  $|a\rangle = (a_1, a_2, \dots)^T$  as the coordinate vector. Multiplying equation (2.82) by  $\langle u_i|$  from the left, we obtain a system of coupled linear equations

$$\langle u_i, v \rangle = a_1 \langle u_i, u_1 \rangle + a_2 \langle u_i, u_2 \rangle + \dots \quad \forall i. \quad (2.83)$$

This set of equations corresponds to the following matrix-vector multiplication:

$$|\tilde{v}\rangle = \mathbf{S}_{|\mathbf{u}\rangle} |a\rangle, \quad (2.84)$$

with

$$|\tilde{v}\rangle = \begin{pmatrix} \langle u_1, v \rangle \\ \langle u_2, v \rangle \\ \langle u_3, v \rangle \\ \dots \end{pmatrix}, \quad \mathbf{S}_{|\mathbf{u}\rangle} = \begin{pmatrix} \langle u_1, u_1 \rangle & \langle u_1, u_2 \rangle & \langle u_1, u_3 \rangle & \dots \\ \langle u_2, u_1 \rangle & \langle u_2, u_2 \rangle & \langle u_2, u_3 \rangle & \dots \\ \langle u_3, u_1 \rangle & \langle u_3, u_2 \rangle & \langle u_3, u_3 \rangle & \dots \\ \dots & \dots & \dots & \dots \end{pmatrix}, \quad |a\rangle = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ \dots \end{pmatrix}.$$

$\mathbf{S}_{|\mathbf{u}\rangle}$  is the Gramian matrix of the linear independent  $\{|u_1\rangle, |u_2\rangle, \dots\}$  and its inverse matrix  $\mathbf{S}_{|\mathbf{u}\rangle}^{-1}$  exists. Thus, the coordinate vector  $|a\rangle$  is given according to:

$$|a\rangle = \left( \mathbf{S}_{|\mathbf{u}\rangle}^{-1} \right) |\tilde{v}\rangle. \quad (2.85)$$

For a single coordinate  $a_i$ , we have:

$$a_i = \sum_{j=1} \left( \mathbf{S}_{|\mathbf{u}\rangle}^{-1} \right)_{ij} |\tilde{v}\rangle_j = \sum_{j=1} \left( \mathbf{S}_{|\mathbf{u}\rangle}^{-1} \right)_{ij} \langle u_j, \tilde{v} \rangle. \quad (2.86)$$

Combining equations (2.86) and (2.82), we can express a vector  $v \in V$  with respect to the basis  $\{|u_1\rangle, |u_2\rangle, \dots\}$  as:

$$|v\rangle = \sum_{i=1} |u_i\rangle a_i = \sum_{i=1} |u_i\rangle \sum_{j=1} \left( \mathbf{S}_{|\mathbf{u}\rangle}^{-1} \right)_{ij} \langle u_j, v \rangle = \left( \sum_{i=1} |u_i\rangle \sum_{j=1} \left( \mathbf{S}_{|\mathbf{u}\rangle}^{-1} \right)_{ij} \langle u_j| \right) |v\rangle. \quad (2.87)$$

So we can write for the resolution of the identity:

$$\hat{I}d_{|u\rangle} = \sum_{i=1} |u_i\rangle \sum_{j=1} \left( \mathbf{S}_{\{\mathbf{u}\}}^{-1} \right)_{ij} \langle u_j|, \quad (2.88)$$

and for the special case of an orthonormal basis, we have:

$$\hat{I}d_{|u\rangle} = \sum_{i=1} |u_i\rangle \langle u_i|. \quad (2.89)$$

By specifying a basis  $\{|u_1\rangle, |u_2\rangle, \dots, |u_N\rangle\}$ , a coordinate vector  $a = (a_i)_{i \leq N}$  is uniquely assigned to every vector  $v \in V$  according to equation (2.82). The isomorphism between the  $N$ -dimensional vector space  $V$  and the  $N$ -dimensional Cartesian vector space  $\mathbb{R}^N$  formed by the coordinate vectors is given by the operator  $\hat{I}d_{|u\rangle} : \mathbb{R}^N \rightarrow V$  defined in equation (2.88).

$$\hat{I}d_{|u\rangle} : \mathbb{R}^N \rightarrow V \quad (2.90)$$

$$(a_i)_{i \leq N} \mapsto v = \sum_{i=1}^N |u_i\rangle a_i \quad (2.91)$$

Our considerations are not limited to the special case of finite-dimensional vector spaces. An important special case within this thesis is the infinite-dimensional, separable Hilbert space  $L^2[X]$  formed by the square-integrable functions with support  $X \subset \mathbb{R}^3$ . By choosing an appropriate basis (e.g. expansion within a Fourier series), a separable, infinite-dimensional vector space formed by the coordinate vectors is obtained. The coefficients of the coordinate vectors are square-summable and the coordinate vector space is denoted by  $\ell^2$ . The connection of the vector space of square-integrable functions and square-summable series is based on the Riesz representation theorem, which states that all separable, infinite-dimensional Hilbert spaces are isometrically-isomorphic.

### Change-of-basis matrix

Let  $\{|u_1\rangle, |u_2\rangle, \dots\}$  and  $\{|p_1\rangle, |p_2\rangle, \dots\}$  be two orthonormal bases of the vector space  $V$ . For  $|v\rangle \in V$ , we obtain by inserting the resolution of the identity twice (cf. eq. (2.89)):

$$|v\rangle = \hat{I}d_{|p\rangle} \hat{I}d_{|u\rangle} |v\rangle = \sum_{i=1} |u_i\rangle \underbrace{\sum_{j=1} \langle u_i, p_j \rangle \langle p_j, v \rangle}_{a_i}, \quad (2.92)$$

where  $a_i$  is the  $i$ -th coefficient of the coordinate vector  $a := (a_1, a_2, \dots)^T$  of  $|v\rangle$  with respect to the basis  $\{|u_1\rangle, |u_2\rangle, \dots\}$ . We can write for  $a_i$ :

$$a_i = \sum_{j=1} \underbrace{\langle u_i, p_j \rangle}_{\left( \mathbf{W}_{\{\{\mathbf{u}\}\}}^{\{\{\mathbf{p}\}\}} \right)_{ij}} \underbrace{\langle p_j, v \rangle}_{\tilde{a}_j} \quad (2.93)$$

Equation (2.93) describes the element-wise matrix-vector multiplication of a matrix  $\mathbf{W}_{\{\{\mathbf{p}\}\}}^{\{\{\mathbf{u}\}\}}$  and a vector  $\tilde{a} = (\tilde{a}_1, \tilde{a}_2, \dots)^T$ . In this context, we identify  $\tilde{a}$  as the coordinate vector of  $|v\rangle$  with respect to the orthonormal basis  $\{|p_1\rangle, |p_2\rangle, \dots\}$ . The matrix  $\mathbf{W}_{\{\{\mathbf{p}\}\}}^{\{\{\mathbf{u}\}\}}$



describes the transformation between the two coordinate vectors. We call  $\mathbf{W}_{\{\mathbf{p}\}}^{\{\mathbf{u}\}}$  the change-of-basis matrix.

An orthogonal matrix  $\left(\mathbf{W}_{\{\mathbf{p}\}}^{\{\mathbf{u}\}}\right)$  fulfills:  $\left(\mathbf{W}_{\{\mathbf{p}\}}^{\{\mathbf{u}\}}\right)^{-1} = \left(\mathbf{W}_{\{\mathbf{p}\}}^{\{\mathbf{u}\}}\right)^T$ . The following calculation proves that  $\mathbf{W}_{\{\mathbf{p}\}}^{\{\mathbf{u}\}}$  is an orthogonal matrix, by checking element-wise that

$$\left(\mathbf{W}_{\{\mathbf{p}\}}^{\{\mathbf{u}\}}\right)^T \left(\mathbf{W}_{\{\mathbf{p}\}}^{\{\mathbf{u}\}}\right) = Id:$$

$$\sum_j \left(\mathbf{W}_{\{\mathbf{p}\}}^{\{\mathbf{u}\}}\right)_{ij} \left(\mathbf{W}_{\{\mathbf{p}\}}^{\{\mathbf{u}\}}\right)_{jl}^T = \sum_j \left(\mathbf{W}_{\{\mathbf{p}\}}^{\{\mathbf{u}\}}\right)_{ij} \left(\mathbf{W}_{\{\mathbf{p}\}}^{\{\mathbf{u}\}}\right)_{lj} = \sum_j \langle u_i, p_j \rangle \langle p_j, u_l \rangle \quad (2.94)$$

$$= \langle u_i | \underbrace{\left( \sum_j |p_j\rangle \langle p_j| \right)}_{\hat{I}_{d_{|p\rangle}}} |u_l\rangle = \langle u_i, u_l \rangle = \delta_{il} \quad (2.95)$$

Here, we used  $\langle p_j, u_l \rangle \in \mathbb{R}$  and  $\langle p_j, u_l \rangle = \langle u_l, p_j \rangle$ .

We define a linear operator  $\hat{\mathbf{U}}$  which describes the change of basis from  $\{|u_1\rangle, |u_2\rangle, \dots\rangle$  to  $\{|p_1\rangle, |p_2\rangle, \dots\rangle$ . The essence of the presented scheme for the matrix representation of a change-of-basis operator  $\hat{\mathbf{U}}$  is summarized in the following commutative diagram:

$$\begin{array}{ccc} V & \xrightarrow{\hat{\mathbf{U}}} & V \\ \hat{I}_{d_{|p\rangle}} \uparrow & & \uparrow \hat{I}_{d_{|u\rangle}} \\ K^N & \xrightarrow{\mathbf{W}_{\{\mathbf{p}\}}^{\{\mathbf{u}\}}} & K^N \end{array}$$

with  $\hat{I}_{d_{|p\rangle}} = \sum_{i=1} |p_i\rangle \langle p_i|$  and  $\hat{I}_{d_{|u\rangle}} = \sum_{i=1} |u_i\rangle \langle u_i|$ . The change-of-basis matrix  $\mathbf{W}_{\{\mathbf{p}\}}^{\{\mathbf{u}\}}$  is given by (2.96).  $K^N$  is used as a symbol for the vector spaces formed by the coordinate vectors.

Our considerations are not limited to the special case of finite-dimensional vector spaces. An important special infinite-dimensional case in this thesis is  $V \cong L^2[X]$  with  $X \subset \mathbb{R}^3$  and  $K^N \cong \ell^2$ .

The elements of the change-of-basis matrix for orthonormal bases are as follows (cf. eq. (2.93)):

$$\left(\mathbf{W}_{\{\mathbf{p}\}}^{\{\mathbf{u}\}}\right)_{ij} = \langle u_i, p_j \rangle. \quad (2.96)$$

For the general case of (not necessarily orthonormal) bases  $\{|w_1\rangle, |w_2\rangle, \dots\rangle$  and  $\{|q_1\rangle, |q_2\rangle, \dots\rangle$ , we obtain in a completely analogous manner by using the resolution of the identity defined by equation (2.88):

$$\left(\mathbf{W}_{\{\mathbf{q}\}}^{\{\mathbf{w}\}}\right)_{ij} = \sum_l \left(\mathbf{S}_{|\mathbf{w}\rangle}^{-1}\right)_{il} \langle w_l, q_j \rangle \quad (2.97)$$

where  $\mathbf{S}_{ij} = \langle w_i, w_j \rangle$ .

### Representation of linear operators

If two vector spaces  $V$  and  $W$  are finite-dimensional vector spaces and a basis is defined for each of the vector spaces, then every linear transformation from  $V$  to  $W$  can be represented by a matrix. Infinite-dimensional, compact, linear operators can also be viewed as continuous generalizations of matrices. Here, we will derive formulas for the transformation matrix elements of linear operators with respect to arbitrary bases.

Let  $\hat{\mathbf{T}} : V \rightarrow W$  be a linear map with  $\Psi \mapsto \Phi := \hat{\mathbf{T}}(\Psi)$ . First, we confine ourselves to orthonormal basis sets. Later, we will generalize our results to arbitrary bases. Let  $\{|p_1\rangle, |p_2\rangle, \dots\}$  be the orthonormal basis of the image and  $\{|\tilde{p}_1\rangle, |\tilde{p}_2\rangle, \dots\}$  the orthonormal basis of the domain. We expand the image vector  $\Phi$  and the domain vector  $\Psi$  with respect to the chosen basis by inserting the resolution of the identity from equation (2.89):

$$|\Psi\rangle = \hat{I}_{|\tilde{p}\rangle} |\Psi\rangle = \sum_i |\tilde{p}_i\rangle \langle \tilde{p}_i, \Psi\rangle \quad (2.98)$$

$$|\Phi\rangle = \hat{I}_{|p\rangle} |\Phi\rangle = \sum_i |p_i\rangle \langle p_i, \Phi\rangle = \sum_i |p_i\rangle \langle p_i, \hat{\mathbf{T}}\Psi\rangle. \quad (2.99)$$

Inserting equation (2.98) into equation (2.99), we obtain:

$$|\Phi\rangle = \sum_i |p_i\rangle \langle p_i, \hat{\mathbf{T}} \sum_j |\tilde{p}_j\rangle \langle \tilde{p}_j, \Psi\rangle \quad (2.100)$$

$$= \sum_i |p_i\rangle \sum_j \underbrace{\langle p_i, \hat{\mathbf{T}} \tilde{p}_j\rangle}_{\left(\mathbf{T}_{|\tilde{p}\rangle}^{|p\rangle}\right)_{ij}} \underbrace{\langle \tilde{p}_j, \Psi\rangle}_{\tilde{a}_j}. \quad (2.101)$$

According to equation (2.82), we can define  $a_i$  as the  $i$ -th coefficient of the coordinate vector  $a = (a_1, a_2, \dots)^T$  of  $|\Phi\rangle$  with respect to the basis  $\{|p_1\rangle, |p_2\rangle, \dots\}$ .

$$a_i = \sum_j \underbrace{\langle p_i, \hat{\mathbf{T}} \tilde{p}_j\rangle}_{\left(\mathbf{T}_{|\tilde{p}\rangle}^{|p\rangle}\right)_{ij}} \underbrace{\langle \tilde{p}_j, \Psi\rangle}_{\tilde{a}_j} \quad (2.102)$$

Equation (2.102) describes the element-wise matrix-vector multiplication of a matrix  $\mathbf{T}_{|\tilde{p}\rangle}^{|p\rangle}$  and a vector  $\tilde{a} = (\tilde{a}_1, \tilde{a}_2, \dots)$ . In this context, we identify  $\tilde{a}$  as the coordinate vector of  $|\Psi\rangle$  with respect to the basis  $\{|\tilde{p}_1\rangle, |\tilde{p}_2\rangle, \dots\}$ .  $\mathbf{T}_{|\tilde{p}\rangle}^{|p\rangle}$  corresponds to the matrix representation of the operator  $\hat{\mathbf{T}}$ . We call  $\mathbf{T}_{|\tilde{p}\rangle}^{|p\rangle}$  the transformation matrix (with respect to the basis  $\{|p_1\rangle, |p_2\rangle, \dots\}$  of the image and the basis  $\{|\tilde{p}_1\rangle, |\tilde{p}_2\rangle, \dots\}$  of the domain).

We can express the action of the linear operator  $\hat{\mathbf{T}}$  to a vector  $|\Psi\rangle$  in two equivalent ways:

$$a = \mathbf{T}_{|\tilde{p}\rangle}^{|p\rangle} \tilde{a} \quad \Leftrightarrow \quad |\Phi\rangle = \hat{\mathbf{T}}|\Psi\rangle \quad (2.103)$$

with  $|\Phi\rangle = \sum_i |p_i\rangle a_i$  and  $|\Psi\rangle = \sum_i |\tilde{p}_i\rangle \tilde{a}_i$ . We can write explicitly (cf. eq. (2.102)) for the  $ij$ -th transformation matrix element:

$$\left(\mathbf{T}_{|\tilde{p}\rangle}^{|p\rangle}\right)_{ij} = \langle p_i, \hat{\mathbf{T}} \tilde{p}_j\rangle. \quad (2.104)$$

The essence of the presented scheme for the matrix representation of linear operators is summarized in the following commutative diagram:

$$\begin{array}{ccc}
 V & \xrightarrow{\hat{\mathbf{T}}} & W \\
 \hat{I}d_{|\tilde{p}\rangle} \uparrow & & \uparrow \hat{I}d_{|p\rangle} \\
 K^M & \xrightarrow{\mathbf{T}_{|\tilde{p}\rangle}^{|\mathbf{p}\rangle}} & K^N
 \end{array}$$

with  $\hat{I}d_{|\tilde{p}\rangle} = \sum_i |\tilde{p}_i\rangle\langle\tilde{p}_i|$  and  $\hat{I}d_{|p\rangle} = \sum_i |p_i\rangle\langle p_i|$ . The transformation matrix  $\mathbf{T}_{|\tilde{p}\rangle}^{|\mathbf{p}\rangle}$  is given by (2.104).  $K^M$  and  $K^N$  are the vector spaces formed by the coordinate vectors of  $V$  and  $W$ , respectively. Our considerations are not limited to finite dimensional vector spaces. An important special infinite-dimensional case in this thesis is  $V = W = L^2[X]$  with  $X \subset \mathbb{R}^3$  and  $K^M = K^N = \ell^2$ .

The generalization of the derivation of the transformation matrix elements to arbitrary bases is straightforward. Let  $\{|q_1\rangle, |q_2\rangle, \dots\}$  be a (not necessarily orthogonal) basis of the image and  $\{|\tilde{q}_1\rangle, |\tilde{q}_2\rangle, \dots\}$  the (not necessarily orthogonal) basis of the domain.

For the general case, the resolution of the identity is given by equation (2.88). Expanding the image vector  $|\Phi\rangle$  and the domain vector  $|\Psi\rangle$  with respect to the corresponding bases yields:

$$|\Psi\rangle = \hat{I}d_{|\tilde{q}\rangle}|\Psi\rangle = \sum_i |\tilde{q}_i\rangle \sum_j (\mathbf{S}_{|\tilde{q}_i\rangle}^{-1})_{ij} \langle\tilde{q}_j, \Psi\rangle \quad (2.105)$$

$$|\Phi\rangle = \hat{I}d_{|q\rangle} \underbrace{|\Phi\rangle}_{|\hat{\mathbf{T}}\Psi\rangle} = \sum_i |q_i\rangle \sum_j (\mathbf{S}_{|q_i\rangle}^{-1})_{ij} \langle q_j, \hat{\mathbf{T}}\Psi\rangle \quad (2.106)$$

Equations (2.107) and (2.108) can be obtained in a completely analogous manner for the general case.

$$|\Phi\rangle = \hat{\mathbf{T}}\Psi = \sum_o |q_o\rangle \underbrace{\sum_n (\mathbf{S}_{|q\rangle}^{-1})_{on} \sum_m \langle q_n, \hat{\mathbf{T}}\tilde{q}_m\rangle \sum_s (\mathbf{S}_{|\tilde{q}\rangle}^{-1})_{ms} \langle\tilde{q}_s, \Psi\rangle}_{a_o} \quad (2.107)$$

$$a_o = \underbrace{\sum_n (\mathbf{S}_{|q\rangle}^{-1})_{on} \sum_m \langle q_n, \hat{\mathbf{T}}\tilde{q}_m\rangle}_{(\mathbf{T}_{|\tilde{q}\rangle}^{|\mathbf{q}\rangle})_{om}} \underbrace{\sum_s (\mathbf{S}_{|\tilde{q}\rangle}^{-1})_{ms} \langle\tilde{q}_s, \Psi\rangle}_{\tilde{a}_m}, \quad (2.108)$$

with the transformation matrix element:

$$(\mathbf{T}_{|\tilde{q}\rangle}^{|\mathbf{q}\rangle})_{om} = \sum_n (\mathbf{S}_{|q\rangle}^{-1})_{on} \langle q_n, \hat{\mathbf{T}}\tilde{q}_m\rangle. \quad (2.109)$$

## 2.4 Markov Chains

A Markov chain is a mathematical system that undergoes transitions from one state to another according to certain probabilistic rules. Markov chains are a common, and relatively simple, way to statistically model the evolution of certain systems such

as animal populations and traffic flows and they are applied in fields such as economics, game theory, queueing (communication) theory, genetics, and finance.<sup>186,187</sup> In contrast with other stochastic models Markov chains describe a sequence of possible events in which the probability of each event depends only on the previous state. In general, Markov chains can be grouped according to continuous or discretized time and a countable or continuous state space. Here, we will focus on a discrete-time Markov chain with a countable state space  $S$ . A discrete-time Markov chain is a sequence of random variables  $Y = (X_t)_{t \in \mathbb{N}}$  where  $X_t$  are elements of a countable state space  $S = \{s_1, s_2, s_3, \dots\}$ . The sequence of states satisfy the Markov property, i.e. the probability of the transition to the next state depends only on the present state and not on the previous states. This central property is also referred to using the term "memorylessness". More formally, this condition can be expressed as

$$P\{X_{n+1} = s_{j_{n+1}} | X_0 = s_{j_0}, \dots, X_n = s_{j_n}\} = P\{X_{n+1} = s_{j_{n+1}} | X_n = s_{j_n}\}. \quad (2.110)$$

where  $P\{X_{n+1} = s_{j_{n+1}} | X_0 = s_{j_0}, \dots, X_n = s_{j_n}\}$  and  $P\{X_{n+1} = s_{j_{n+1}} | X_n = s_{j_n}\}$  are conditional properties.

For a finite state space, the transition probability distribution can be represented by a matrix  $\mathbf{Q}$ , which we call the transition matrix. The elements of the quadratic transition matrix are formally given as

$$\mathbf{Q}_{ij} = P\{X_{n+1} = s_i | X_n = s_j\}. \quad (2.111)$$

From the definition of the elements of the transition matrix as probabilities given by equation (2.111), it is apparent that

1. the transition matrix is stochastic (sum of the rows of the matrix equal to one)
2. the matrix elements are greater than or equal to zero.

A transition matrix which fulfills the condition that the sum of the columns and the sum of the rows are both equal to one is called doubly stochastic.

For a time-independent transition matrix, a Markov chain  $\{X_1, X_2, X_3, \dots\}$  is obtained by successive application of the transition matrix to the state vectors

$$X_{n+1} = \mathbf{Q}X_n \quad (2.112)$$

The  $n$ -step transition probability between the  $i$ -th state at  $t = 0$  and  $j$ -th state at the  $n$ -th step is given by

$$p_{ij}^n := P(X_n = s_j | X_0 = s_i). \quad (2.113)$$

From equation (2.112), it follows that this  $n$ -step transition probability  $p_{ij}^n$  is given by the  $ij$ -th element of the  $n$ -th power of the transition matrix  $(\mathbf{Q}^n)_{ij}$ .

The connection between discrete Markov chains and the transition matrix further expand the toolbox available for the investigation of Markov models by the methods for matrix investigation and characterization provided by linear algebra. Graph theory can also be utilized for the characterization of Markov chains. In that case, the elements of the state space are considered as nodes within a directed graph. The nodes are then connected by the non-zero transition probabilities (which can be directly obtained from the elements of the transition matrix).

In the following discussion, the important properties of Markov chains are presented and the effect of these properties on the elements of the transition matrix is also highlighted.

**Reducibility/Irreducibility** A Markov chain is said to be irreducible if it is possible to get to any state from any state (possibly in more than one time step). A state  $j$  is said to be accessible from a state  $i$  if a system started in state  $i$  has a non-zero probability of transitioning to state  $j$  at some point. Formally, state  $j$  is accessible from state  $i$  if there exists an integer  $n_{ij} \geq 0$  such that

$$P\{X_{n_{ij}} = j | X_0 = i\} = p_{ij}^{n_{ij}} > 0. \quad (2.114)$$

This integer is allowed to be different for each pair of states, hence the subscripts in  $n_{ij}$ .

We derived from equation (2.112) and (2.113) that  $p_{ij}^{n_{ij}} = (\mathbf{Q}^{n_{ij}})_{ij}$ . As a result, the irreducibility of the Markov chain corresponds to the condition that every element of the  $M$ -th power of the  $M \times M$  transition matrix is greater than zero.

**Periodicity/Aperiodicity** A state  $i$  has period  $k$  if any return to state  $i$  must occur in multiples of  $k$  time steps. Formally, the period of state  $i$  is defined as

$$\text{gcd}\{n > 0 : P(X_n = i | x_0 = i) > 0\}, \quad (2.115)$$

where gcd is the greatest common divisor and it is assumed that the set is nonempty. If  $k = 1$ , then the state is said to be aperiodic. Otherwise ( $k > 1$ ), the state is said to be periodic with period  $k$ . A Markov chain is aperiodic if every state is aperiodic. An irreducible Markov chain only needs one aperiodic state to imply that all states are aperiodic. As a consequence of the latter observation, a transition matrix is aperiodic if at least one diagonal element is greater than zero.

## Convergence behavior of Markov models

In this thesis, a discrete Markov model for long-range proton transfer is constructed. In this section, we wish to examine the impact of certain properties of the transition matrix - such as being irreducible, aperiodic and doubly stochastic - on the qualitative properties of the Markov model. This section closely follows article [Dreßler et al., *J. Chem. Phys.*, **2020**, 152(11):114114.] in this thesis.

Let  $\mathbf{Q} \in \mathbb{R}^{n \times n}$  be a transition matrix of a Markov chain i.e. all entries are non-negative and the sum of each column vector of  $\mathbf{Q}$  is equal to one ( $\mathbf{Q}$  is a stochastic matrix). The orbit or trajectory of  $\mathbf{Q}$  is the set  $\{\mathbf{Q}^n x_0 \mid x_0 \in \mathbb{R}^n, n \in \mathbb{N}\}$ . Fixed points or stationary distributions of  $\mathbf{Q}$  are those vectors  $x \in \mathbb{R}^n$  with  $\mathbf{Q}x = x$ , whose orbit is the one-element set  $\{x\}$ . This is equivalent to  $\lim_{n \rightarrow \infty} \mathbf{Q}^n x = x$ . In this section, we wish to answer the following three questions:

- Under which conditions does a unique stationary distribution or a fixed point of  $\mathbf{Q}$  exist?
- Which  $x \in \mathbb{R}^n$  is the fixed point of  $\mathbf{Q}$ ?
- For which initial vectors  $x_0 \in X$  do we encounter this fixed point?

The following theorem from the literature will assist us in this regard:<sup>188</sup>

**Theorem 1** (*A part of the first ergodic theorem*)

*Every stochastic, irreducible and aperiodic matrix  $\mathbf{Q}$  has a unique fixed point or stationary distribution  $x$ . This fixed point is positive (i.e. all elements  $x_i \geq 0$ ), and the sequence of the distributions  $\mathbf{Q}^n x_0$  converges to  $x$  for every initial distribution  $x_0$ .*

It follows from this statement that a Markov matrix  $\mathbf{Q}$  always has 1 as an eigenvalue. The absolute value of all other eigenvalues are less than 1. The eigenvector  $x$  for the eigenvalue 1 is called the stationary distribution of  $\mathbf{Q}$ .

We are even able to predict the fixed point for symmetric or double-stochastic transition matrices  $\mathbf{Q}$ :<sup>188</sup>

**Theorem 2** *The uniform distribution at  $\mathbb{R}^n$ , that is, the vector  $x \in \mathbb{R}^n$  with  $x^{(0)} = x^{(1)} = \dots = x^{(n)}$ , is the stationary distribution of the stochastic transition matrix  $\mathbf{Q}$  if the sums of the columns and of the rows of  $\mathbf{Q}$  are both equal to one (i.e.  $\mathbf{Q}$  is double-stochastic).*

## 2.5 Dynamical Systems

If one is interested in the simulation of dynamic processes, starting from the formal mathematical definition of a dynamical system is a much more obvious approach compared to the concept of Markov chains. However, this contradiction can easily be resolved because a discrete Markov chain can always be considered as a dynamical system.

From a formal mathematical point of view, a discrete dynamical system is a 3-tuple  $(T, X, \Phi)$  consisting of a set  $T = \mathbb{N}_0$  or  $\mathbb{Z}$  (indexing the temporal evolution), a non-empty set  $X$  (the state space or the phase space), and an evolution function of the dynamical system  $\Phi: T \times X \rightarrow X$ , so that for all states  $x \in X$  and all times  $t, s \in T$  we have:

$$I: \Phi(0, x) = x \quad (2.116)$$

$$II: \Phi(s, \Phi(t, x)) = \Phi(s + t, x) \quad (\text{semigroup action}). \quad (2.117)$$

With respect to the time  $t \in T$ , the function  $\Phi(t, x)$  maps every point  $x \in X$  to a unique image. For fixed  $x \in X$ , the function  $\beta_x: T \rightarrow X$ ,  $t \mapsto \beta_x(t) := \Phi(t, x)$  is called the flow through  $x$ . The set  $O(x) := \{\beta_x(t) \mid t \in T\}$  is called the trajectory (or orbit or phase curve) and describes the (temporal) evolution of an initial state vector  $x$ . By specifying an elementary physical time step  $\Delta s$ , we can assign a physical time to our discrete time via the relation  $t \in T \leftrightarrow t \cdot \Delta s$ .

We wish to restrict ourselves to a specific dynamical system that fulfills the following conditions:

- The phase space  $X$  is isomorphic to a Euclidean vector space  $\mathbb{R}^M$ .
- The evolution function  $\Phi(n, x)$  of the state vector  $x \in \mathbb{R}^M$  (corresponding to  $n$  elementary time steps) can be expressed by the  $n$ -fold application of a matrix  $\mathbf{P} \in \mathbb{R}^{M \times M}$  to the state vector  $x$ :

$$\Phi(n, x) = \mathbf{P}^n x \quad (2.118)$$

with

$$\mathbf{P}^n = \mathbf{P} \times \mathbf{P} \times \dots \times \mathbf{P} \quad (n \text{ times}). \quad (2.119)$$

For this specific evolution function, the semigroup property can be expressed by matrix multiplication:

$$\Phi(s, \Phi(t, x)) = \mathbf{P}^s (\mathbf{P}^t x) = \mathbf{P}^{s+t} x = \Phi(s + t, x). \quad (2.120)$$

### Relation between Markov chains and dynamic models

By comparing equations (2.118) to (2.120) with equation (2.112), it becomes apparent that the terms discrete Markov chain and discrete dynamical system are synonyms. Starting from a state  $x$  at  $t = 0$ , an appropriately constructed evolution function  $\Phi(n, x)$  (cf. eq. (2.118)) yields the same state compared to the  $n$ -fold application of the transition matrix  $\mathbf{P}$ . Conversely, the following calculation highlights the Markov property of a dynamical system induced from the semigroup action: Let  $x_{t+\Delta t}$  be the future state of a dynamical system, let  $x_t$  be the present state of the system and let  $x_0$  be the initial state of the system (at  $t = 0$ ). According to

$$x_{t+\Delta t} = \Phi(t + \Delta t, x_0) \stackrel{\text{eq. (2.117)}}{=} \Phi(\Delta t, \Phi(t, x_0)) = \Phi(\Delta t, x_t) \quad (2.121)$$

the future state of a dynamical system is only dependent on the current state of the system.

## 2.6 Asymmetric Simple Exclusion Process

The Asymmetric Simple Exclusion Process (ASEP) is a simple model that allows for the simulation of particles moving on a grid.<sup>189</sup> The simplicity of the ASEP model has made it the default stochastic model for transport phenomena.<sup>190</sup> This method was introduced as a mathematical model in 1970 and several variants exist nowadays. The ASEP in its original version is too simple to study real world systems, but it can be regarded as an abstraction of complex processes - as demonstrated in the case of protein bio-synthesis, for example.<sup>191</sup> The ASEP method can be used to derive fundamental relations between the microscopic scale and macroscopic ensembles in these systems. In contrast with other more realistic simulation methods, the ASEP principle allows for (at least partial) analytical treatment.

All variants of the ASEP principle model transitions of particles between lattice sites. These lattice sites can be occupied by either one or zero particles. If a particle already resides at a lattice site, no further jumps to this site are possible. In a simple version of the ASEP principle, particle movement along a one-dimensional lattice was simulated in continuous time.<sup>192</sup> An “alarm clock” was used that rang after an exponentially distributed random time with mean value  $\tau = 1/(p + q)$ . The alarm initiated a particle jump with a probability of  $P_l = p/(p + q)$  to the next lattice site to the left and a particle jump with a probability of  $P_r = q/(p + q)$  to the next lattice site to the right. The jump was only possible if the corresponding adjacent lattice site was unoccupied.

In this work, a method based on the ASEP principle will be used to model the movement of protons on a (dynamically evolving) grid of oxygen atoms. A detailed description of how to derive the exact transition probabilities and the positions of the lattice sites is presented in section 2.8.

## 2.7 Monte Carlo Methods

Monte Carlo (MC) methods are employed for numerical integration, global optimization, queuing theory, structural mechanics and the solution of large systems of linear, partial differential or integral equations.<sup>124</sup> In this section, the basic idea of the MC method is introduced. In statistical physics and chemistry, MC methods are utilized to study the behavior of complex systems of thousands of atoms in space and time. In

MC methods, a sequence of points in phase space is generated from an initial geometry by adding an appropriate random perturbation to the coordinates of a randomly chosen particle (atom or molecule). The core components of an MC algorithm are a random number generator and the ability to calculate the energy of the system for a given set of coordinates.<sup>193</sup> Generated “random” numbers are typically pseudo-random because a deterministic recursion rule is used to generate a sequence of numbers starting from an initial seed. A pseudo-random number generator produces a sequence of seemingly random numbers, but the sequence is repeated after a certain period. A good pseudo-random number generator is characterized by a very long periodicity and within this periodicity the numbers do not show any systematic correlation with each other.

In the following discussion, a condensed description of the Metropolis algorithm is presented that follows the approach in reference 124. We denote the collective position vector as  $X$ . The Metropolis algorithm provides an efficient simple procedure for the sampling of the canonical ensemble. It generates a chain of molecular states  $\{X_1, X_2, X_3, \dots\}$ . From a mathematical point of view, this chain can be considered as a Markov chain since the generation of a new state  $X_{i+1}$  solely depends on the current state  $X_i$ . The Metropolis algorithm offers one possibility to calculate the transition matrix for this Markov chain and thus enables the construction of a phase space trajectory in the canonical ensemble.

We start from a chain of states  $\{X_1, X_2, X_3, \dots\}$  which is prolonged by means of the Metropolis method using the following protocol: Starting from the  $i$ -th state  $X_i$ , a trial  $\tilde{X}_{i+1}$  is generated by an appropriate random perturbation and accepted if the corresponding energy is lower. If, however, the energy of the state  $\tilde{X}_{i+1}$  is larger than the energy of the state  $X_i$  ( $E(\tilde{X}_{i+1}) > E(X_i)$ ), then the new state is accepted with a probability of  $p = \exp(-\beta\Delta E)$ , with  $\Delta E = E(\tilde{X}_{i+1}) - E(X_i) > 0$  and  $\beta = 1/(k_B T)$ . This is achieved by comparison of  $p$  with a uniformly-generated number on the interval (0,1). If  $p$  is greater than the random number, the new state  $\tilde{X}_{i+1}$  is accepted, otherwise another trial  $\tilde{X}_{i+1}$  is constructed but  $X_i$  is recounted in the chain of states. The result of this procedure is the acceptance probability at step  $i$  of:

$$p_{acc,MC} = \min[1, \exp(-\beta\Delta E)] \quad (2.122)$$

In this manner, states with lower energies are always accepted but states with higher energies have a non-zero probability of acceptance too. Consequently, the sequence tends towards regions of the configuration space with low energies, but the system can always escape to other areas of the energy surface.

In the work carried out for this thesis, a modified protocol of the Metropolis algorithm is used for the simulation of proton transfer. In the standard algorithm probabilities derived from the energy differences ( $p = \exp(-\beta\Delta E)$ ) are used, whereas the probability of proton transfer is directly accessible in the method derived in this thesis. Thus proton transfer only takes place if the uniformly-generated number on the interval (0,1) is less than the probability calculated for proton transfer between the two oxygen atoms.

In general, the Metropolis algorithm does not simulate the temporal evolution of the system. However, in the method derived in this thesis a time-dependency is introduced indirectly by relating the probabilities for proton transfer to a certain time interval. The details of this approach are given in the next section.



## 2.8 The Combined Molecular Dynamics/Lattice Monte Carlo (cMD/LMC) Approach for Long-range Proton Transfer

The simulation of proton conduction is a computationally challenging task.<sup>61-63</sup> On the one hand, we are forced to use *ab initio* MD simulations because we have to simulate the cleavage and formation of covalent bonds. On the other hand, proton transfer is a rare event at the AIMD time scale and only a few proton transfer events can be observed in an entire trajectory.

From a formal point of view, molecular kinetics derived from an MD simulation can always be described in terms of a Markov model. The evolution of a state in an MD simulation is obtained by the numerical solution of the equations of motion (cf. equation (2.1)). These equations depend only on the current state of a system (positions, velocities and accelerations at time  $t$ ) and thus fulfill the Markov property.

There are at least two crucial reasons why the construction and application of Markov models is hindered in practice:

1. An appropriate discretization of the phase space is needed.
2. The underlying transition matrix could be determined by brute force sampling, i.e. counting of interesting events from an MD trajectory. However, in order to obtain an efficiency gain for the simulation of molecular kinetics, more sophisticated methods should be employed for the calculation of the transition matrix of the Markov model.

In the following, a detailed description of the cMD/LMC method is given. We will demonstrate that the two crucial factors are implicitly fulfilled by the cMD/LMC approach. Proton transfer has been an area of interest of the group of Prof. Sebastiani for many years. A careful analysis of the AIMD trajectory of hexa(p-phosphonatophenyl)benzene (HPB) revealed an interesting Gaussian-shaped dependency between the O-O distance and the frequency of proton transfer. This functional shape of the O-O distance/proton transfer frequency relationship was confirmed for several O $\cdots$ H-O hydrogen bonded compounds (poly(vinylphosphonic acid) (PVPA), CsHSO<sub>4</sub>, CsHSeO<sub>4</sub>, CsH<sub>2</sub>PO<sub>4</sub>) and is shown in Figure 2.3. This figure presents the histogram of the proton jumps with respect to the oxygen-oxygen distance within an AIMD trajectory (normalized to 100 ps and 16 protons for reasons of comparability). In this context, we define a proton jump as a change in the next oxygen neighbor of a hydrogen atom (in two subsequent AIMD steps). All compounds share the same center ( $d_{OO} \approx 2.4$  Å) for the Gaussian-shaped functions of the proton jumps with respect to the O-O distances.

Figure 2.3 suggests that the positions of the oxygen atoms are decisive for the description of proton movement. In addition, the protons do not diffuse freely within a condensed matter system. They are always attached to an oxygen atom for the proton conducting systems within the scope of this thesis. As a result, the positions of the oxygen atoms provide an apparent discretization for the simulation of long-range proton transfer. In the Lattice Monte Carlo part of the cMD/LMC method, we reduce the system to the position of the oxygen atoms. These oxygen atoms are referred to as “sites”. A site can be occupied by one proton at the most, which corresponds to the situation of a covalently bonded proton to this oxygen atom.

In the cMD/LMC model, the movement of the protons among this oxygen lattice is governed by the oxygen-oxygen distances. In contrast with other lattice gas models, the positions of the oxygen atoms (i.e. the positions of the lattice sites) are not

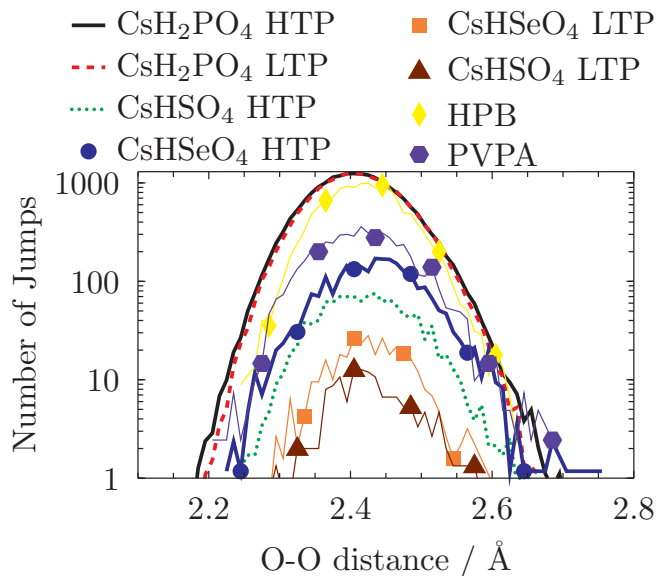


FIGURE 2.3: Number of proton jumps with respect to the O-O distance (histogram) occurring in a 100 ps simulation of a box containing 16 protons. Details concerning the high temperature (HTP) and low temperature phases (LTP) of  $\text{CsHSO}_4$ ,  $\text{CsHSeO}_4$ ,  $\text{CsH}_2\text{PO}_4$  are given in section 2.9. Reproduced with permission from Dreßler, Sebastiani, **Phys. Chem. Chem. Phys.**, 2020, **22**, 10738-10752.

fixed. Instead, the positions are updated from an AIMD trajectory. The articles that constitute this thesis clearly demonstrate that only the updating of the lattice enables the simulation of long-range proton transfer by incorporating relaxation processes of the atomistic environment (for details see section 3.1).

The actual proton dynamics are performed by the Monte Carlo part “on top” of the dynamically evolving molecular structure. In addition to the positions of the oxygen atoms and their occupations by protons, the Monte Carlo routine requires the transition probabilities for a proton between the oxygen sites. A schematic illustration of the cMD/LMC model is presented in Figure 2.4.

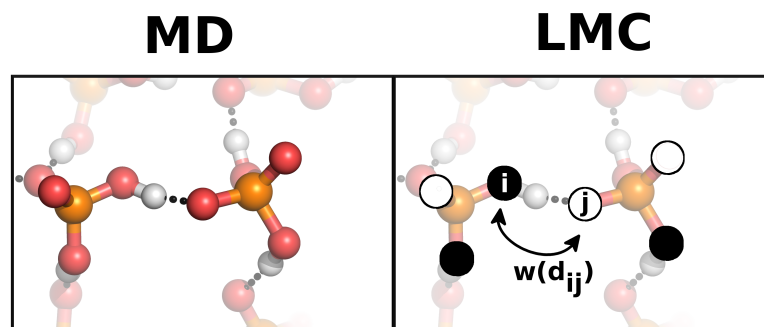
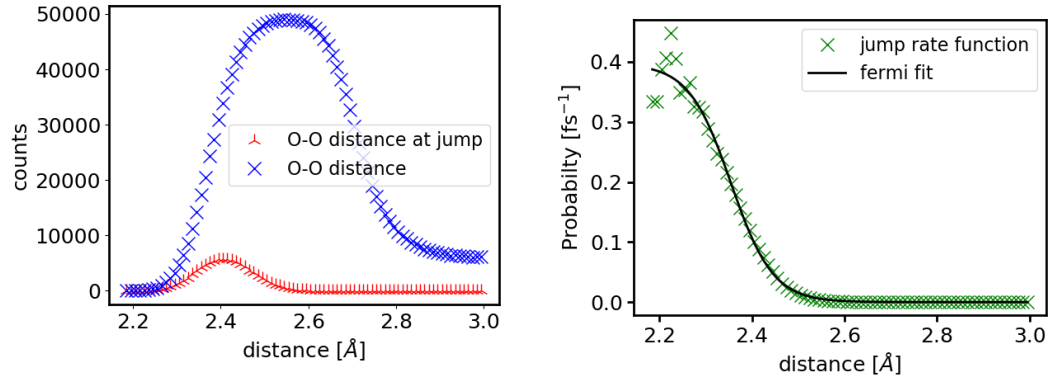


FIGURE 2.4: The combined Molecular Dynamics/ lattice Monte Carlo approach. Reproduced with permission from Dreßler et al., *J. Chem. Phys.* **153**, 114114 (2020). Copyright 2020 AIP Publishing LLC.

In order to obtain a Markovian model, the transition probabilities for a proton should only depend on geometric criterion derived from the current oxygen positions. In line with our discussion of Figure 2.3, the following protocol is used for the construction of the jump rate function  $\omega(d_{ij})$  from an AIMD trajectory: The jump rate



(A) Histograms of O-O distances and of O-O distances at a proton jump.

 (B) Jump rate function  $\omega(d)$  of  $\text{CsH}_2\text{PO}_4$ .

 FIGURE 2.5: O-O distance histograms and jump rate function  $\omega$  for the specific example of  $\text{CsH}_2\text{PO}_4$ . The jump rate function  $\omega$  (right) is obtained as point-wise quotient of the two histograms (left).

function  $\omega$  is the conditional probability of a proton jump at a given O-O distance. The value of this function for a certain oxygen-oxygen distance is calculated as the ratio of the number of proton jumps at a given oxygen-oxygen distance and the number of absolute occurrences of that oxygen-oxygen distance. Both dividend and divisor can be obtained by detecting and counting these events from the AIMD trajectory. The connection between the jump rate function  $\omega$ , the overall histogram of oxygen-oxygen distances and the histogram of the oxygen-oxygen distances *at* a proton jump is illustrated in Figure 2.5 for the special case of the compound  $\text{CsH}_2\text{PO}_4$ . The jump rate function  $\omega$  (green) is quotient of the two histograms (blue and red points). The resulting function  $\omega$  depends only on the oxygen-oxygen distances  $d_{ij}$  and is fitted to a Fermi function:

$$\omega(d_{ij}) = \frac{a}{1 + \exp\left(\frac{d_{ij}-b}{c}\right)}, \quad (2.123)$$

where  $d_{ij}$  is the distance between the oxygen atoms  $\text{O}_i$  and  $\text{O}_j$  and  $a$ ,  $b$ ,  $c$  are the fit parameters.

The existence of a jump rate function for the description of proton transfer which only depends on the current state of the system is comprehensible. However, the specific functional form of  $\omega$  obtained from the above protocol is arbitrary. It is only motivated by the conclusions from Figure 2.3 and the empirical justification from the application of the cMD/LMC scheme to a variety of compounds. Other approaches and modifications of the protocol presented above for the calculation of  $\omega$  were also investigated in the articles that constitute this thesis (for details see section 3.1).

The unit of the proton jump rate is probability per time. According to this rate function a physical time can be assigned to the LMC steps. The probability of proton transfer  $p_{ij}$  between the  $i$ -th and the  $j$ -th lattice site within the interval  $\Delta t$  is given by:

$$p_{ij} = \omega(d_{ij})\Delta t. \quad (2.124)$$

For a system of  $N$  sites with fixed positions, the transition probabilities can be stored in an  $N \times N$  jump matrix.

As the positions (and hence the distances  $d_{ij}$ ) are updated from the AIMD trajectory, the jump matrix becomes also time-dependent. The cMD/LMC approach

employs a variation of the Asymmetric Simple Exclusion Process (ASEP) principle (cf. section 2.6) for the stochastic propagation of the protons. The proton movement is modeled on a discrete lattice given by the oxygen positions and each lattice site can either be occupied or unoccupied. At each Monte Carlo step, a pair of oxygen atoms  $O_i$  and  $O_j$  is chosen randomly. If  $O_i$  is occupied by a proton, and site  $O_j$  is empty, the proton is moved to  $O_j$  with the probability  $p_{ij}$  as obtained from equation (2.124).

Simulating a system of  $N$  oxygen atoms, there are a total number of  $N(N - 1)$  connections between the oxygen atoms. Therefore,  $N(N - 1)$  Monte Carlo steps (denoted as a sweep) are associated with the time step  $\Delta t$  in (2.124). The interval  $\Delta t$  corresponds to the AIMD time between the updates of the atomic positions. Thus, the time step of the AIMD simulation is an obvious choice for  $\Delta t$  in (2.124).

A general scheme of the cMD/LMC approach, illustrating the connection between the updating of the molecular structure and the Monte Carlo part is presented in Figure 2.6.

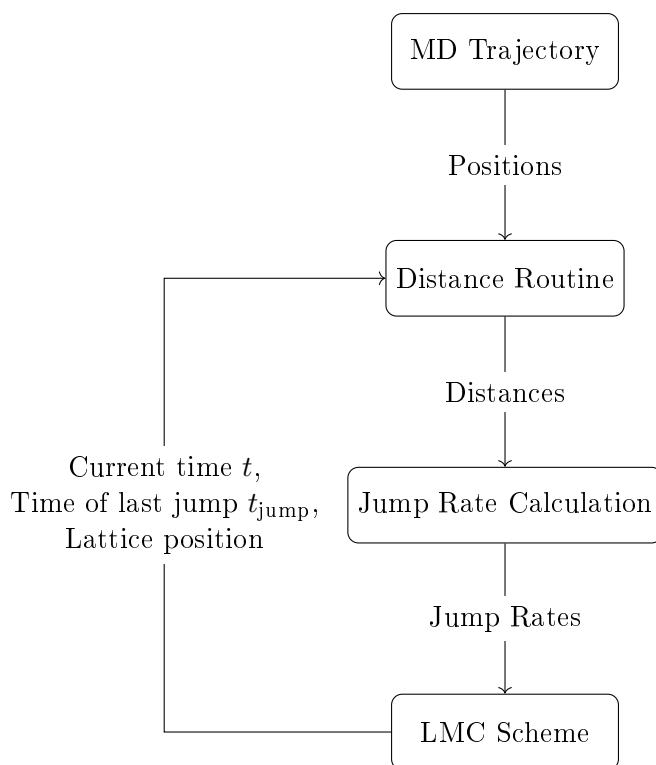


FIGURE 2.6: Flow chart diagram of the cMD/LMC approach. This figure is taken from ref. 1.

## 2.9 Solid Acids

Proton conduction is a process of fundamental scientific interest and has been extensively studied by chemists and physicists.<sup>41–51</sup> If we focus on the field of energy conversion, the core component of a hydrogen fuel cell is the proton exchange membrane. This membrane has to be chemically stable, durable and a good proton conductor.

Solid, inorganic acid compounds (or simply solid acids) of the type  $XH_yZO_4$  ( $X = K, Rb, Cs, NH_4$ ;  $y = 1, 2$ ;  $Z = P, S, Se, As$ ) often undergo a super-protonic phase transition at higher temperatures.<sup>3, 120, 194</sup> This super-protonic phase transition leads to an increase in the proton conductivity by several orders of magnitude.<sup>195–203</sup> The

first super-protonic phase transition was reported more than 30 years ago.<sup>204</sup> Since that time, solid acids have been under constant investigation, which has resulted in novel proton-conducting membrane materials for fuel cell applications. The potential use of solid acids as fuel cell membranes materials has also been demonstrated outside of the laboratory:<sup>205</sup> a first fuel cell based on  $\text{CsH}_2\text{PO}_4$  is already commercially available.<sup>123</sup> The most common fuel cell membrane material is Nafion, a perfluorinated sulfonic acid.<sup>206</sup> Water molecules are involved in the conduction mechanism, which limits the operating temperature of Nafion based fuel cells to  $100^\circ\text{C}$ .

Up to the early 2000s, there was an ongoing discussion about the existence of a super-protonic phase transition in  $\text{CsH}_2\text{PO}_4$ . The idea that water traces or thermal decomposition products are responsible for proton conduction at higher temperatures was widely discussed.<sup>207–209</sup> Several reviews and further experimental studies finally confirmed the existence of the high-temperature, super-protonic phase of  $\text{CsH}_2\text{PO}_4$ , thus avoiding the requirement for a non-genuine conduction mechanism.<sup>205,210–214</sup> This water free proton conduction in solid acids is a fundamental advantage compared to other proton conducting materials, such as Nafion.

The origin of the extraordinarily high proton conductivity in the super-protonic phases of the solid acids is a dynamically disordered hydrogen bond network.<sup>3,215,216</sup> Although these compounds are still solid crystalline materials, the reorientation dynamics of the tetrahedral anions can be characterized as “liquid”-like. The latter property is of utmost importance, as proton transfer can only take place in these crystalline materials *via* the Grotthuss mechanism, which involves proton transfer events between neighboring anions followed by structural relaxation of the molecular environment (rotation of the anions).<sup>3,215,216</sup> The super-protonic phase is always observed at elevated temperatures.<sup>3,120,194,205,217</sup> Surprisingly, this is associated with a formal increase in the symmetry of the crystal structures, e.g. transition from the monoclinic to the tetragonal or cubic crystal system. In fact, this increase in symmetry enables the dynamically disordered hydrogen bond network due to the increased disorder of the oxygen atoms indicated by only partial occupation of sites in the crystal structure. A case in point is the transition of the monoclinic phase of  $\text{CsH}_2\text{PO}_4$  to the high-temperature phase of  $\text{CsH}_2\text{PO}_4$ .<sup>2,205</sup> The orientational disorder of the phosphate groups in the super-protonic phase arises from the mismatch of the tetrahedral symmetry of the anion and the octahedral symmetry of the center of the primitive cubic cell constituted from the  $\text{Cs}^+$  ions. An alignment of the P-O bonds along the  $\langle 111 \rangle$  direction of the crystal structure would in principle retain the cubic symmetry. However, the Cs-P distance is too short for this alignment (4.30 Å).

Thus, twenty-four possible partially occupied sites of the oxygen atoms are reported in the crystal structure (cf. Figure 2.8), resulting in six different orientations of the phosphate group. At the elevated temperature of the phase transition (503 K), the different orientations are frequently inter-converted, leading to a highly dynamically disordered or “liquid”-like hydrogen bond network.<sup>205,215,216</sup>

Within the scope of this thesis, the focus is predominantly on a series of solid acid compounds: the super-protonic (cubic or tetragonal) and monoclinic phases of  $\text{CsHSO}_4$ ,  $\text{CsHSeO}_4$  and  $\text{CsH}_2\text{PO}_4$ . The temperature of the super-protonic phase transition and the space group of involved crystal structures are shown in Table 2.1. In the following, the term low-temperature phase (LTP) refers to the monoclinic phase just below the super-protonic phase transition and the term high-temperature phase (HTP) refers to the tetragonal or cubic phase just above the phase transition.

TABLE 2.1: Characteristic crystal structure parameters for the high- and low-temperature phases (HTP/LTP) of  $\text{CsH}_2\text{PO}_4$ ,  $\text{CsHSeO}_4$  and  $\text{CsHSO}_4$ <sup>3</sup>

	$\text{CsH}_2\text{PO}_4$	$\text{CsHSO}_4$	$\text{CsHSeO}_4$
HTP	$\text{Pm}\bar{3}\text{m}$ cubic	$\text{I4}_1/\text{amd}$ tetragonal	$\text{I4}_1/\text{amd}$ tetragonal
LTP	$\text{P2}_1/\text{m}$ monoclinic	$\text{P2}_1/\text{c}$ monoclinic	$\text{P2}_1/\text{c}$ monoclinic
$\lambda_c$	503 K	414 K	384 K

### Crystal structures of the HTP and LTP of $\text{CsH}_2\text{PO}_4$

The following figures and descriptions of the crystal structures are taken from the article [Dreßler, Sebastiani, *Phys. Chem. Chem. Phys.*, **2020.**] in this thesis.

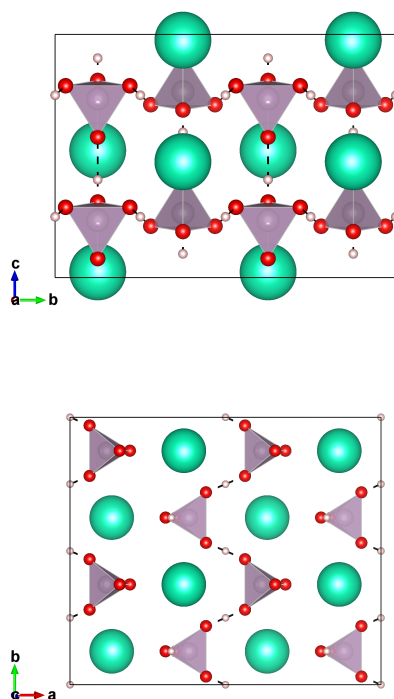


FIGURE 2.7: Eight unit cells of the crystal structure of the monoclinic phase (LTP) of  $\text{CsH}_2\text{PO}_4$  are shown with different lattice orientations. Red: oxygen, green: caesium, gray: hydrogen. Reproduced with permission from Dreßler, Sebastiani, **Phys. Chem. Chem. Phys.**, 2020, **22**, 10738-10752.

The crystal structure of the monoclinic phase (LTP) of  $\text{CsH}_2\text{PO}_4$  is shown in Figure 2.7.<sup>218</sup> The hydrogen bond network in this phase consists of two different types of hydrogen bonds.<sup>218</sup> Symmetric double-minima hydrogen bonds (2.48 Å) form one-dimensional chains of oxo-anions in the  $b$  direction. These chains are linked by asymmetric single-minimum hydrogen bonds (2.54 Å) leading to two-dimensional

layers of hydrogen bonded tetrahedra. Hydrogen atoms involved in the symmetric hydrogen bond with the double potential well undergo rapid exchange between two different oxygen sites.<sup>217,219,220</sup>

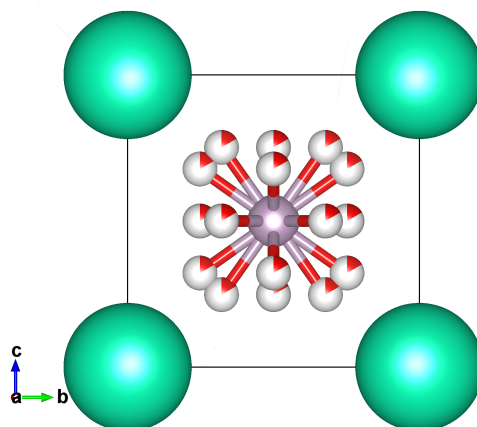


FIGURE 2.8: One unit cell of the crystal structure of the cubic phase (HTP) of  $\text{CsH}_2\text{PO}_4$  is shown. Partial occupancies are indicated by partial coloring of the atoms.<sup>2</sup> Red: oxygen, green: caesium, lilac: phosphorus. Reproduced with permission from Dreßler, Sebastiani, **Phys. Chem. Chem. Phys.**, 2020, **22**, 10738-10752.

One unit cell of the crystal structure of the cubic phase (HTP) of  $\text{CsH}_2\text{PO}_4$  is presented in Figure 2.8.<sup>2</sup> The large number of partially colored atoms (partial occupation of the crystal sites) indicates a strong disorder of the oxygen atoms. Twenty-four partially occupied oxygen sites are reported, resulting in six different orientations of the phosphate tetrahedra. The  $\text{PO}_4$  groups reside in the center of the cubic unit cell. The reorientational disorder is caused by the mismatch of the tetrahedral oxo-anion unit and the octahedral symmetry of the cubic center. An alignment of the P-O bonds with the  $\langle 111 \rangle$  body diagonal would retain the overall cubic symmetry while removing the octahedral symmetry from the center. In fact, this is not possible due to the small phosphorus-caesium distance of  $4.30 \text{ \AA}$  in this direction. Displacing the oxygen atoms from this diagonal retains the cubic symmetry and results in six different orientations of the phosphate anions depicted in Figure 2.8.<sup>205</sup> Besides a strong (single minimum) hydrogen bond of  $2.46 \text{ \AA}$ , other hydrogen bonds of  $2.76 \text{ \AA}$ ,  $2.85 \text{ \AA}$  and  $3.03 \text{ \AA}$  length are possible from the crystal structure.

### Crystal structures of the HTP and LTP of $\text{CsHSO}_4$ and $\text{CsHSeO}_4$

The low- and the high-temperature phases of  $\text{CsHSO}_4$  and  $\text{CsHSeO}_4$  are isostructural. Thus, only the crystal structure of  $\text{CsHSO}_4$  will be discussed.

Three different phases of  $\text{CsHSO}_4$  (phases I, II, and III) are reported within the temperature range between room temperature and approximately 500 K.<sup>221-223</sup> In this thesis, we are interested in the proton conduction mechanism just above and just below the super-protonic phase transition. As a result, we restrict our investigations to phase II (LTP) and phase III (HTP). The crystal structures of the monoclinic (LTP) and the tetragonal (HTP) phase of  $\text{CsHSO}_4$  are presented in Figures 2.9 and 2.10. While the monoclinic phase forms infinite chains of hydrogen-bonded sulfate tetrahedra, cyclic dimers linked by hydrogen bonds are reported for the tetragonal phase. These structural motifs are also known from other solid acids and the hydrogen bonds

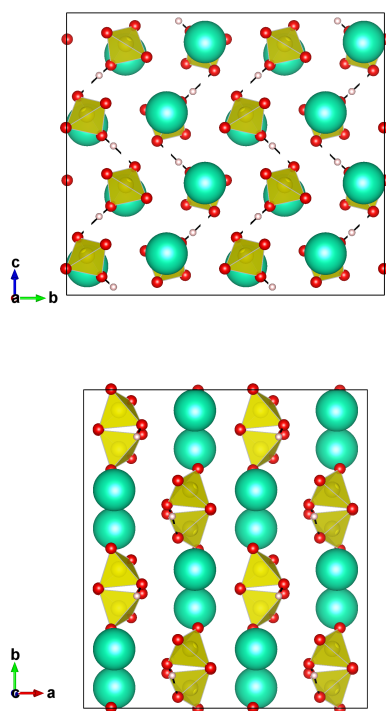


FIGURE 2.9: Four unit cells of the crystal structure of the monoclinic phase (LTP) of  $\text{CsHSO}_4$  are shown with different lattice orientations. red: oxygen, green: caesium, gray: hydrogen. Reproduced with permission from Dreßler, Sebastiani, **Phys. Chem. Chem. Phys.**, 2020, **22**, 10738-10752.

within zigzag chains tend to be stronger than hydrogen bonds formed by the cyclic dimers.<sup>217</sup> For  $\text{CsHSO}_4$ , this also holds for the monoclinic and the tetragonal phase, and hydrogen bond lengths of 2.63 Å (monoclinic phase) and 2.79 Å (tetragonal phase) can be observed from the crystal structures. Slightly different crystal structures of the tetragonal phase were proposed by Jiraak et al.,<sup>224</sup> Merinov et al.<sup>225</sup> and Belushkin et al.<sup>226</sup> These structures differ in terms of the number and the sites of the partially occupied oxygen positions. In Figure 2.10, we present the crystal structure with two possible orientations of the oxygen tetrahedron reported by Jiraak et al..

### 2.9.1 Open Questions Concerning the Conduction Mechanism in Solid Acids

Solid acids have been extensively studied from both the experimental and the theoretical perspectives. The structure and the phase transition were characterized by X-ray diffraction and calorimetry. The proton dynamics were investigated with the help of PFG-NMR techniques, neutron scattering and impedance spectroscopy.<sup>205,207,208,210–213,227–229</sup> From the theoretical point of view, static calculations of the electronic structure as well as classical and *ab initio* molecular dynamics were used to study the solid acids.<sup>16,117,118,228,230–237</sup>

There is a strong consensus in the scientific community that proton conduction in the high-temperature phases of these systems follows a general Grotthuss mechanism enabled by a dynamically disordered hydrogen bond network.<sup>3,120,194,205,217</sup> However, a large amount of (isolated) studies produced contrary findings regarding



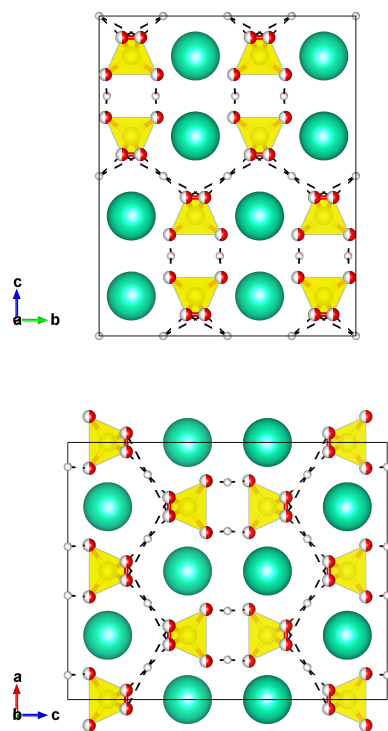


FIGURE 2.10: Four unit cells of the crystal structure of the tetragonal phase (HTP) of  $\text{CsHSO}_4$  are shown with different lattice orientations. red: oxygen, green: caesium, gray: hydrogen. Reproduced with permission from Dreßler, Sebastiani, **Phys. Chem. Chem. Phys.**, 2020, **22**, 10738-10752.

specific parts of the proton conduction mechanism in some cases: the proton conduction mechanism in these compounds can be decomposed into contributions from two underlying sub-processes: the proton transfer rate between anions and the structural reorientation rate of the molecular environment (rotation rate of the anions). For the high-temperature phase of  $\text{CsH}_2\text{PO}_4$ , it was reported that the anion reorientation rate ( $1 \cdot 10^{11} \text{ s}^{-1}$ ) compared to the proton transfer rate ( $7 \cdot 10^{11} \text{ s}^{-1}$ ) is the rate limiting step for proton conduction.<sup>117,238</sup>

Several experimental<sup>226,239-242</sup> and theoretical research groups<sup>16,230-233</sup> have reported that the situation for  $\text{CsHSO}_4$  is the opposite: the proton conductivity in the HTP of this compound is limited by the proton transfer rate compared to the frequency of the anion rotation. However, an NMR study and a more recent *ab initio* molecular dynamics (AIMD) study by Wood et al. proposed that the anion rotation rate of  $\text{CsHSO}_4$  constitutes the limiting factor for the proton conductivity in this compound.<sup>118,243</sup>

Each of the low-temperature phases has a drastically reduced proton conductivity compared to its high-temperature phase equivalent. Nevertheless, the proton conduction mechanism in the low-temperature phase of  $\text{CsH}_2\text{PO}_4$  was recently investigated by several NMR studies.<sup>2,219,220,244</sup> The reorientation rate of the anion seems to be crucial for the long-range proton transfer mechanism. The rotation rates of the  $\text{H}_2\text{PO}_4^-$  anion are not isotropic and two types of rotations were reported, differing in their correlation with long-range proton transfer. While for the faster rotational process only local fluctuations of the hydrogen atoms within a hydrogen bond are

involved, the slower rotation enables the long-range transfer of a proton.<sup>219</sup> Independently of the specific rotational process, the frequency of phosphate rotations in the LTP is reduced compared with the frequency in the disordered HTP.<sup>117,219,238</sup>

Many previous theoretical investigations focused on CsHSO<sub>4</sub>. A variety of methods ranging from static quantum chemical calculations by Te and Kanka<sup>237</sup> to classical molecular dynamics studies by Munch et al. or Haile et al. were conducted.<sup>230-232</sup> In particular, it is noteworthy that the classical MD simulations of Haile et al. were not only able to compute the structure and dynamics of CsHSO<sub>4</sub>, but that they could even predict the transition itself, without any *a priori* knowledge of the existence of the super-protonic phase.<sup>232</sup> More recent studies focus on the importance of specific effects for the simulations of the monoclinic phase of CsH<sub>2</sub>PO<sub>4</sub>, such as dispersion corrections or nuclear quantum effects.<sup>228,234</sup>

If theoretical calculations are designed to elucidate the proton conduction mechanism, the underlying method has to be capable of simulating bond breaking. Consequently AIMD has to be used for the investigation of the proton conduction mechanism by means of MD, as only this technique demonstrates a truly predictive ability. This predictive ability is achieved by an approximate solution of the Schrödinger equation for the investigated system in every time step, which makes AIMD a computationally demanding method and typically restricts the simulations to system sizes of several hundred of atoms and the timescale to less than a nanosecond. This timescale is comparable to the timescale of long-range proton transfer in many systems. Thus, long-range proton transfer is generally a rare event in AIMD simulations and the convergence of descriptors of protonic motion (such as diffusion coefficients) obtained from AIMD simulations is often questionable.<sup>61-63,245,246</sup> Also, the AIMD studies already published on CsH<sub>2</sub>PO<sub>4</sub> and CsHSO<sub>4</sub> are affected by these problems.<sup>117,118</sup> While the experimental diffusion coefficient of CsH<sub>2</sub>PO<sub>4</sub> has been reproduced within a factor of two, a ten-fold increased value of the diffusion coefficient of CsHSO<sub>4</sub> has been reported.

## Chapter 3

# Summary of the Peer-reviewed Articles

### 3.1 Multiscale Approaches for Long-range Proton Transfer

A detailed introduction in the combined Molecular Dynamics/Lattice Monte Carlo (cMD/LMC) approach is presented in section 2.8. The cMD/LMC method is a scale-bridging method for the simulation of long-range proton transfer. This approach uses the atomistic information from an AIMD simulation to model microscopic proton transfer probabilities, and the Monte Carlo part takes the dynamically evolving molecular structure into account and performs actual proton dynamics "on top" of that dynamical structure.

In 2014, Kabbe *et al.* introduced the idea of the cMD/LMC approach and demonstrated its applicability in principle for the compound hexa(p-phosphonatophenyl)benzene.<sup>247</sup> A further benchmark of the cMD/LMC method is presented in the article [Dreßler, Kabbe, Sebastiani, *J. Phys. Chem. C*, **2016**.] in this thesis. In this article, the cMD/LMC approach is applied to the solid acids CsH<sub>2</sub>PO<sub>4</sub> and CsHSO<sub>4</sub>, which are extraordinarily good proton conductors and promising candidates for fuel cell membrane materials. CsH<sub>2</sub>PO<sub>4</sub> and CsHSO<sub>4</sub> undergo a super protonic phase transition at elevated temperatures that is accompanied by significantly increased proton conductivity. The cMD/LMC approach was able to reproduce this enormous change in the proton mobility as well as the moderate increase in the conductivity due to simple variations of the temperature. Phase transitions and temperature variations are both reflected by only moderate changes in the positions of the atoms in the trajectory, but lead to a wide range of proton conductivity responses for these compounds. It is particularly noteworthy that this variety in proton conductivity responses is completely captured by the cMD/LMC approach without any reparameterization. This mainly results from the fact that the positions of the oxygen atoms (i.e. the positions of the lattice sites) are not fixed. Instead, the positions of the oxygen atoms are periodically updated from an AIMD trajectory. In the article [Dreßler, Kabbe, Sebastiani, *J. Phys. Chem. C*, **2016**.] in this thesis, it is clearly demonstrated that only the updating of the lattice enables the simulation of long-range proton transfer by incorporating relaxation processes of the atomistic environment. Several other approaches using a fixed oxygen lattice (e.g. positions from crystal structure or averaged positions from an AIMD trajectory) were tested. None of them was able to reproduce the high proton conductivity in the super-protonic phases.

Despite the reproduction of the diffusion coefficients, the mechanism of proton transfer can also be investigated using the cMD/LMC method, which is presented in the article [Dreßler, Kabbe, Sebastiani, *J. Phys. Chem. C*, **2016**.]. The investigation

of the proton conduction mechanism in this article is the reason why it is formally listed in section 3.3. Proton conduction in  $\text{CsH}_2\text{PO}_4$  and  $\text{CsHSO}_4$  follows a general Grotthuss type mechanism, which can be understood in terms of the proton transfer frequency between the anions and the reorientation rate of the  $\text{HSO}_4^-/\text{H}_2\text{PO}_4^-$  anions. While the proton transfer frequency itself is a central parameter of the cMD/LMC method (via the jump rate function), the rotation of the anions is reflected by the periodic updating of the molecular structure from the AIMD trajectory. For  $\text{CsH}_2\text{PO}_4$ , the anion rotation frequency is the limiting step for the proton conductivity and a similar limitation holds for the proton transfer frequency and  $\text{CsHSO}_4$ . This can be identified by careful analysis of the AIMD trajectories - for example, as done in article [Dreßler, Sebastiani, *Phys. Chem. Chem. Phys.*, **2020**.] in this thesis. However, the cMD/LMC approach is also able to cross-check this assumption from a different point of view. The cMD/LMC method allows for the direct simulation of proton diffusion with respect to variations of the proton jumprate. This can be used to verify the sensitivity of the proton mobility as a function of the proton transfer rate. The dependency of the proton mobility on the anion rotation rate can also be investigated within the cMD/LMC approach by means of appropriate modifications of the underlying trajectory.

The jump rate function  $\omega$  is a very important parameter in the cMD/LMC approach. The existence of such a function, which describes the probability of proton transfer and which is only dependent on the current positions of the oxygen atoms, is logical. However, the suggested functional form from section 2.8 with sole dependence on the oxygen-oxygen distance is quite arbitrary. In article [Kabbe, Dreßler, Sebastiani, *J. Phys. Chem. C*, **2016**.] an additional dependency of the jump rate function on the hydrogen bond angle is investigated. As a result, a (rather large) cut-off angle was introduced in order to avoid unphysical proton transfer events. A case in point for such unwanted events is the intramolecular proton transfer within a phosphate or sulfate anion.

The jump rate can be determined from an AIMD trajectory according to the protocol given in section 2.8. A desirable aim in order to reduce the computational effort is the application of the cMD/LMC approach without the necessity of an underlying (rather short) AIMD trajectory. In this scenario, the dynamical structure update could be carried out based on a classical MD trajectory (assuming the existence of an accurate force field parametrization for the investigated system). A solution of the remaining open question for the direct calculation of the jump rate function is presented in the article [Kabbe, Dreßler, Sebastiani, *J. Phys. Chem. C*, **2016**.]. This article demonstrates that the jump rate function for  $\text{CsH}_2\text{PO}_4$  and hexa(p-phosphonatophenyl)benzene can be derived from static calculations of the activation energies for a proton transfer between hydrogen phosphate anions.

Article [Kabbe, Dreßler, Sebastiani, *Phys. Chem. Chem. Phys.*, **2017**.] follows this line of approach and its aim is the calculation of the proton diffusion coefficient of bulk-water from a classical MD trajectory. Liquid water is a special case because changes in the solvation shell (coordination number and coordination pattern and not just the oxygen-oxygen distance) are decisive for proton transport. In water, the symmetrization of the local hydrogen bond topology between the proton-donating and proton-receiving oxygen atoms is the key step for the initialization of a proton transfer event. This symmetrization requires equalization of the local coordination numbers.<sup>248,249</sup> In the article [Kabbe, Dreßler, Sebastiani, *Phys. Chem. Chem. Phys.*, **2017**.], the different proton conduction mechanism in water is reflected by rescaling the oxygen-oxygen distances in the vicinity of an excess charge. The characteristic time scale for the relaxation of the oxygen-oxygen distances and the actual function

for the rescaling of the distances were derived from an AIMD simulation of an excess charge in water.

By coupling an MD simulation to the Lattice Monte Carlo part for the propagation of the protonation state, the scale-bridging character of the cMD/LMC simulation becomes apparent. After the calculation of the input parameters from a short AIMD trajectory, several nanoseconds of proton dynamics can be simulated within the Monte Carlo part on a single desktop PC in several hours. However, the propagation of the protonation state within a time step in the LMC part requires  $M(M - 1)$  single Monte Carlo steps for a system containing  $M$  oxygen atoms. This translates into at least linear scaling behavior with significantly reduced (compared to AIMD) but not negligible computational costs for a single time step. Questions arising from various fields of material science, such as nano- or micro-structured materials, require answers at the micrometer and millisecond time scale, which is far beyond the scope of the cMD/LMC approach.

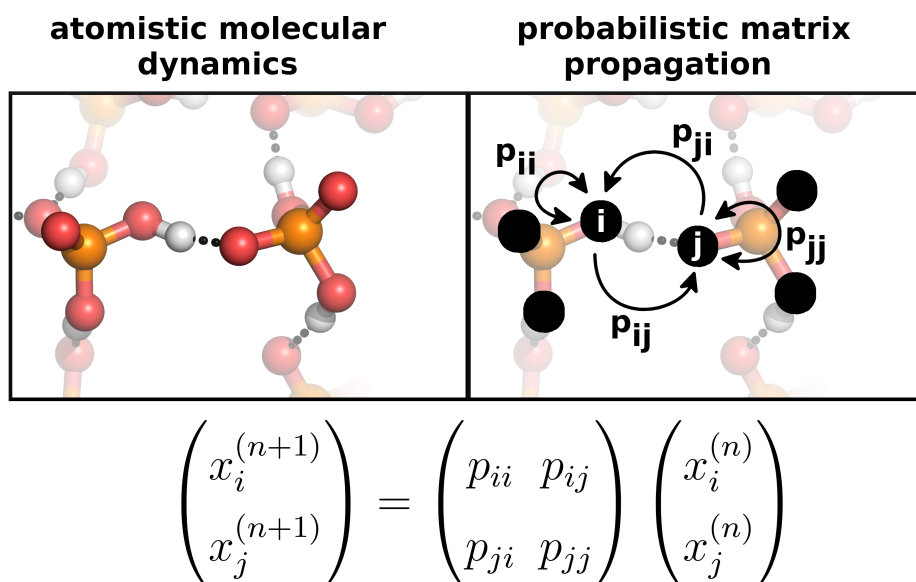


FIGURE 3.1: Schematic illustration of the Molecular Dynamics/ Matrix propagation (MDM) approach. Reproduced with permission from Dreßler et al., *J. Chem. Phys.* **153**, 114114 (2020). Copyright 2020 AIP Publishing LLC.

The cMD/LMC method is a Markovian model, because the propagation of the protonation state depends only on the current positions of the oxygen atoms and their current occupation by protons. The explicit correlation of the protons (i.e. the dependency of the jump matrix on the current proton distribution) hinders the calculation of an overall transition matrix (in terms of a Markov model) for more than one time step or even the entire trajectory.

The condensation of the information concerning proton dynamics of an entire trajectory into a single transition matrix allows for a further significant increase of the length and time scale of the evolution of the protons. The scaling behavior is still linear, but a single (transition) matrix - (state) vector multiplication propagates the system for a duration corresponding to several hundred thousand MD steps. The articles [Dreßler et al., *J. Chem. Phys.*, **2020**, 152(16):164110.] and [Dreßler et al., *J. Chem. Phys.*, **2020**, 152(11):114114.] follow this line of approach in order to further extend the length and time scale of the cMD/LMC approach. The new approach

obtained by the modification of the cMD/LMC approach is referred to as the Molecular Dynamics/Matrix propagation (MDM) method. This model condenses the dynamic information about proton transport within an entire molecular dynamics simulation into an  $M \times M$  matrix  $\mathbf{P}_{\text{MD}}$  where  $M$  is the number of oxygen atoms. The system is again reduced to the position of the  $M$  oxygen atoms and the proton distribution is described by an  $M$ -dimensional state space. In contrast to the cMD/LMC method, the state vector does not describe whether the oxygen atoms are occupied (or not) by exclusively one proton. In the MDM model, the  $i$ -th coordinate of the state vector corresponds to the fraction of a proton (from a statistical point of view) that is covalently bound to the  $i$ -th oxygen atom. A schematic illustration of the MDM model is given in Figure 3.1. The temporal evolution of the distribution of proton probabilities is achieved by means of (time-dependent) transition matrices. These transition matrices are constructed from an underlying AIMD trajectory, such that the evolution of the supramolecular structure in the molecular dynamics simulation is also reflected within the MDM model (in complete analogy to the cMD/LMC method). Thus, the relaxation of the atomistic environment induced by a proton jump is retained in the MDM model. A detailed protocol for the calculation of the elementary transition matrices (obtained from each frame of the trajectory) and their combination in order to obtain the overall transition matrix  $\mathbf{P}_{\text{MD}}$  of the entire MD simulation is given in article [Dreßler et al., *J. Chem. Phys.*, **2020**, 152(11):114114.]. Starting from common topology features of the hydrogen bond network of good proton conductors, this article derives the fundamentals of the MDM approach in terms of a dynamic model for long-range proton transfer. The MDM approach enables the condensation of the proton dynamics information of an entire molecular dynamics simulation into a single  $M \times M$  matrix ( $M$  is the number of oxygen atoms in the simulated system). This large reduction in dimensionality allows for the analytical (instead of numerical) analysis of the model using the toolbox provided by several fields such as graph theory, Markov chains and linear algebra. The article [Dreßler et al., *J. Chem. Phys.*, **2020**, 152(11):114114.] presents a thorough mathematical discussion of the MDM approach, in particular highlighting the implications of a positive, aperiodic and irreducible transition matrix. The linking of these properties of the transition matrix to common topological features of the hydrogen-bond network of good proton conductors allowed for the verification of the uniqueness of the MDM approach and facilitated the proof of its correct asymptotic behavior - i.e. for large time scales all protons are equally distributed to the (chemically equivalent) oxygen atoms. An aperiodic, positive and irreducible matrix is called ergodic. The relationship between the ergodicity of the transition matrix and ergodicity of the MD simulation is further discussed in article [Dreßler et al., *J. Chem. Phys.*, **2020**, 152(11):114114.].

While proton correlation was explicitly incorporated in the cMD/LMC method via the ASEP principle (oxygen atoms can only be occupied or unoccupied), this feature is lost in the MDM approach. In this regard, the state vector describes fractional occupations of the oxygen atoms by protons in a statistical sense.

The article [Dreßler et al., *J. Chem. Phys.*, **2020**, 152(16):164110.] incorporates the correlation of protonation dynamics with the protonation state of the neighboring proton sites into the MDM method. Although the elements of the transition matrix are modified, it is illustrated in this article that this modification conserves the Markov character of the MDM method.

On the one hand, the removal of the explicit proton correlation enables the construction of an overall transition matrix of the entire MD simulation, and hence a significant increase in the length and time scales. On the other hand, the implicit treatment of the proton correlation within the MDM approach is only approximate

and thus proton dynamics are altered compared to the cMD/LMC approach. An illustration of this interplay of accuracy and efficiency is presented in Figure 3.2.

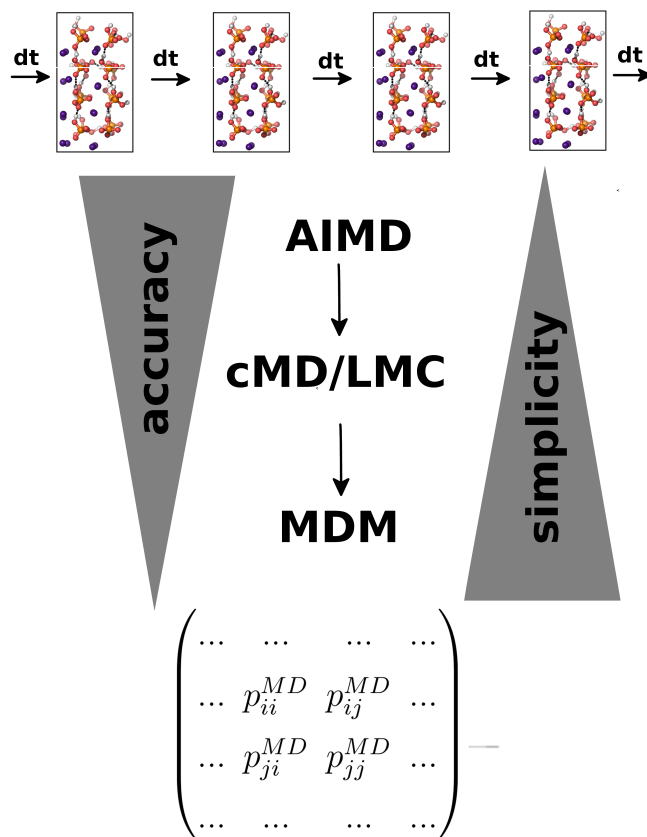


FIGURE 3.2: Comparison of the AIMD, cMD/LMC and MDM approaches.

By calculating the diffusion coefficients and other descriptors of protonic motion, article [Dreßler et al., *J. Chem. Phys.*, **2020**, 152(16):164110.] demonstrates that the relevant part of the proton dynamics is still retained within the MDM approach for at least two specific compounds.

In order to highlight the increase in the length and time scale, the MDM approach is used to explicitly compute non-equilibrium molecular dynamics of protons in the solid acid  $\text{CsH}_2\text{PO}_4$  on a micrometer length scale. It is shown that an excess proton distribution initially localized within a few nanometers diffuses through an  $8 \mu\text{m}$  sized system within 5 ms, in full agreement with the common diffusion laws.

Alongside this proof-of-principle example, the scale-bridging approach for the simulation of long-range proton transfer is used in article [Wagner et al., *J. Mater. Chem. A*, **2019**.] (still referred to as cMD/LMC there) to explain the experimentally measured conductivity trend of the solid acid electrolyte  $\text{CsH}_2\text{PO}_4$  in a composite solid acid fuel cell electrode. In order to increase the electrochemical activity of the fuel cell electrode, nanoporous  $\text{CsH}_2\text{PO}_4$  is used to maximize the surface for platination. The scale-bridging approach for the simulation of long-range proton transfer was utilized to simulate the proton conductivity of these  $\text{CsH}_2\text{PO}_4$  networks with respect to varying degrees of porosity.

### 3.1.1 Article I: Exploring Non-equilibrium Molecular Dynamics of Mobile Protons in the Solid Acid $\text{CsH}_2\text{PO}_4$ at the Micrometer and Microsecond scale

Christian Dreßler, Gabriel Kabbe, Martin Brehm, and Daniel Sebastiani.

Exploring non-equilibrium molecular dynamics of mobile protons in the solid acid  $\text{CsH}_2\text{PO}_4$  at the micrometer and microsecond scale.

*The Journal of Chemical Physics*, 152(16):164110, **2020**.

In this article, I derived the theory and performed all calculations. G. Kabbe contributed with valuable discussions and developed the preceding cMD/LMC method, which was used as the starting point for the derivation of the MDM method. I wrote the manuscript. M. Brehm proofread the manuscript and contributed with valuable discussions. D. Sebastiani supervised the project and provided me with valuable advice while I was writing the manuscript.



# Exploring non-equilibrium molecular dynamics of mobile protons in the solid acid $\text{CsH}_2\text{PO}_4$ at the micrometer and microsecond scale

Cite as: J. Chem. Phys. 152, 164110 (2020); doi: 10.1063/5.0002167

Submitted: 22 January 2020 • Accepted: 3 April 2020 •

Published Online: 28 April 2020



View Online



Export Citation



CrossMark

Christian Dreßler,  Gabriel Kabbe, Martin Brehm,  and Daniel Sebastiani<sup>a)</sup> 

## AFFILIATIONS

Institute of Chemistry, Martin Luther University Halle-Wittenberg, Von-Danckelmann-Platz 4, 06120 Halle (Saale), Germany

<sup>a)</sup>Author to whom correspondence should be addressed: [daniel.sebastiani@chemie.uni-halle.de](mailto:daniel.sebastiani@chemie.uni-halle.de)

## ABSTRACT

We explicitly compute the non-equilibrium molecular dynamics of protons in the solid acid  $\text{CsH}_2\text{PO}_4$  on the micrometer length scale via a multiscale Markov model: The molecular dynamics/matrix propagation (MDM) method. Within the MDM approach, the proton dynamics information of an entire molecular dynamics simulation can be condensed into a single  $M \times M$  matrix ( $M$  is the number of oxygen atoms in the simulated system). Due to this drastic reduction in the complexity, we demonstrate how to increase the length and time scales in order to enable the simulation of inhomogeneities of  $\text{CsH}_2\text{PO}_4$  systems at the nanometer scale. We incorporate explicit correlation of protonation dynamics with the protonation state of the neighboring proton sites and illustrate that this modification conserves the Markov character of the MDM method. We show that atomistic features such as the mean square displacement and the diffusion coefficient of the protons can be computed quantitatively from the matrix representation. Furthermore, we demonstrate the application potential of the scheme by computing the explicit dynamics of a non-equilibrium process in an  $8 \mu\text{m}$   $\text{CsH}_2\text{PO}_4$  system during 5 ms.

Published under license by AIP Publishing. <https://doi.org/10.1063/5.0002167>

## I. INTRODUCTION

Within theoretical chemistry and computational physics, there is a constant challenge to increase the system size and time scale for the calculation of the dynamics of (supra)-molecular systems.

Molecular dynamics (MD) is a very common approach for the simulation of the structure and dynamics of microscopic systems. Within this method, forces on every atom are calculated and subsequently used for the propagation of the atoms for a short time interval according to Newton's second law. By repeated application of this elementary molecular dynamics step, the length of the simulation can be systematically increased. While this concept is in principle suited for the simulation of arbitrary time scales, it suffers in practice from its enormous computational effort, because every single atom has to be considered as an elementary particle within the calculations. The bottleneck for the efficiency of the molecular dynamics simulations is the chosen level of theory for the calculation of the forces acting on the atoms. By the term *ab initio* molecular dynamics (AIMD), we refer to the calculation of the forces by the approximate solution of the Schrödinger equation (using, for example, density

functional theory), which is computationally expensive and restricts the system size to several hundreds of atoms and the time scale to typically less than 1 ns. While the dimensions of the investigated systems are limited within AIMD, it possesses a truly predictive power. This becomes important if the simulation should incorporate breaking or formation of covalent bonds (e.g., simulation of proton transfer).

Many attempts were developed, are in constant development, and will also be developed in the future in order to overcome the restrictions of molecular dynamics concerning system sizes and time scales.

One could think of, e.g., coarse graining, replica exchange, or adaptive/multiple time scale molecular dynamics schemes.<sup>1-7</sup> Another promising direction is the use of neural network potentials which are trained on AIMD trajectory data.<sup>8-11</sup> Once trained, fully atomistic dynamics can be generated for very long time scales. The successful application of neural networks was already demonstrated for various systems such as NaOH solutions,<sup>12</sup> *n*-alkanes,<sup>13</sup> or water on ZnO surfaces.<sup>14</sup> Another common ansatz is the development of explicit scale bridging approaches, which make use of the

combination of two different methods. While a molecular dynamics simulation is often used to obtain atomistic information, the efficient evolution of the system can be done via coupling to a Monte Carlo<sup>15,16</sup> or a Markov method.<sup>17–22</sup> The specific combination of a molecular dynamics simulation and a Monte Carlo algorithm was employed for the investigation of questions arising from the fields of polymer materials,<sup>23–26</sup> semiconductors,<sup>27</sup> surface phenomena,<sup>28–30</sup> and the folding of proteins.<sup>31–33</sup>

In general, a process whose evolution relies only on the current state of the system is called a Markov process. The combination of molecular dynamics simulations and a Markov model becomes immediately apparent if we recall that the calculation of new atomic positions within molecular dynamics depends only on the current positions and velocities of the atoms.<sup>34</sup> However, the construction of a Markov model involving the entirety of molecular degrees of freedom will not reduce the dimensionality of the problem. A crucial point for the setup of the Markov model is the careful discretization of the phase space, reflecting only the molecular processes of interest.<sup>17,34–36</sup>

Concerning the combination of molecular dynamics simulations and Markov models, dynamical processes of large molecular (often biomolecular) systems were addressed in the past,<sup>17–22,34–45</sup> in particular the prediction of protein folding<sup>46,47</sup> or the binding of small molecules to proteins<sup>48–56</sup> as well as the RNA folding kinetics.<sup>40,57,58</sup>

For the simulation of proton transfer, the system size is constrained to a small number of atoms because the simulation of covalent bond breaking enforces the employment of AIMD.<sup>59–69</sup> While we are restricted to this expensive simulation technique, the complete transfer of a proton and the relaxation of the environment occur typically only a few times within a nanosecond within the system sizes tractable with AIMD. Following this line, the simulations will not lead to converged descriptors of proton conduction such as diffusion coefficients, which underlines the urgent need of scale bridging approaches within this field.<sup>70–72</sup>

In a recent publication,<sup>73</sup> we derived the theoretical foundations of a Markov model—the molecular dynamics/matrix propagation (MDM) approach—for the description of long range proton transfer (cf. Figs. 1 and 2). This model condenses the dynamic information on the proton transport within an entire molecular dynamics simulation into an  $M \times M$  matrix, where  $M$  is the number of oxygen atoms. We reduce the system to the position of the  $M$  oxygen atoms, and the proton distribution is described by an  $M$ -dimensional state space. The  $i$ th coordinate of a vector from this state space corresponds to a fraction of a proton (from a statistical point of view), which is covalently bound to the  $i$ th oxygen atom. The temporal evolution of the distribution of proton probabilities is achieved by means of (time dependent) transition matrices. The construction of these matrices is the crucial element of the MDM approach and makes use of an underlying *ab initio* molecular dynamics trajectory of the proton conducting compound under investigation. While other multiscale approaches rely on *static* transition matrices, our transition matrix  $\mathbf{P}_{MD}$  within the MDM method is obtained from the elementary transition matrices  $\mathbf{P}^{(n,n+1)}$ , which are constructed from the current oxygen positions from an *ab initio* molecular dynamics trajectory at the  $n$ th time step and, thus, represent a *dynamic* quantity. Thus, the evolution of the supra-molecular structure in the molecular dynamics simulation is transferred to our MDM model

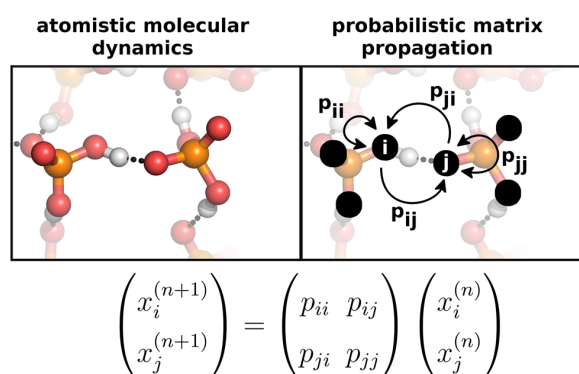


FIG. 1. The combined molecular dynamics/matrix propagation (MDM) approach. Reproduced with permission from Dreßler *et al.*, J. Chem. Phys. **153**, 114114 (2020). Copyright 2020 AIP Publishing LLC.

and enables the incorporation of the relaxation of the environment after a proton jump.

We assume that the probability of a proton transfer between two oxygen atoms is determined by the current atom positions of the system. The  $ij$ th element of the elementary transition matrix  $\mathbf{P}^{(n,n+1)}$  yields the probability of a proton transfer between the two oxygen atoms ( $O_i$  and  $O_j$ ) within the elementary time step of the MDM approach. We calculate the elements of the elementary transition matrix  $\mathbf{P}^{(n,n+1)}$  using a jump rate function  $\omega$ , which is dependent on the distance  $d_{ij}$  between the  $i$ th and the  $j$ th oxygen atoms at the  $n$ th time step. This jump rate function possesses the shape of a Fermi function and is directly accessible via post processing of the underlying *ab initio* molecular dynamics simulation. A detailed description of this function is given elsewhere.<sup>73–75</sup>

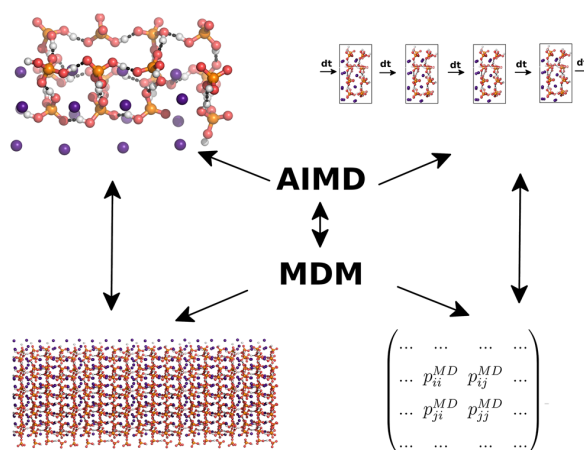


FIG. 2. Comparison of *ab initio* molecular dynamics and the combined molecular dynamics/matrix propagation (MDM) approach.

The elements of the elementary transition matrix  $\mathbf{P}^{(n,n+1)}$  are calculated according to

$$\left(\mathbf{P}^{(n,n+1)}\right)_{ij} = \omega(d_{ij}) \quad \forall i \neq j. \quad (1)$$

The diagonal elements are determined by the condition that the overall number of protons should remain constant,

$$\left(\mathbf{P}^{(n,n+1)}\right)_{jj} = 1 - \sum_{i=1, i \neq j} \left(\mathbf{P}^{(n,n+1)}\right)_{ij}. \quad (2)$$

The proton distribution vector can be propagated by successive multiplication with the elementary transition matrices  $\mathbf{P}^{(1,2)}$ ,  $\mathbf{P}^{(2,3)}$ , ...,  $\mathbf{P}^{(N-1,N)}$ . An equivalent protocol would be obtained if the elementary transition matrices are multiplied in beforehand, and subsequently, the resulting matrix  $\mathbf{P}_{\text{MD}}$  is used to propagate the proton distribution,

$$\mathbf{P}_{\text{MD}} = \mathbf{P}^{(1,N)} = \mathbf{P}^{(N-1,N)} \cdot \dots \cdot \mathbf{P}^{(2,3)} \cdot \mathbf{P}^{(1,2)}. \quad (3)$$

The transition matrix obtained from the successive multiplication of all  $N$  elementary transition matrices resulting from  $N$  frames of the molecular dynamics trajectory will be denoted as  $\mathbf{P}_{\text{MD}}$ , and its application to a proton distribution vector corresponds to a time interval of the entire length of the underlying molecular dynamics simulation. By repeated application of the transition matrix  $\mathbf{P}_{\text{MD}} \in \mathbb{R}^{M \times M}$  to an initial proton distribution  $x_0 \in \mathbb{R}^M$ , we obtain a Markov chain (also referred to as trajectory)  $\{(\mathbf{P}_{\text{MD}})^v x_0 | v \in \mathbb{N}\}$ , which describes the temporal evolution of the proton distribution. Each time step corresponds to the entire length of the underlying molecular dynamics simulation.

In addition to the enormous reduction in the dimensionality of the entire proton dynamics of the molecular dynamics trajectory to  $M \times M$  matrix elements, the MDM method benefits (in contrast to other Markov models) from the physical meaning of the  $M$ -dimensional discretization of the phase space (which translates into the statistical occupation of the  $M$  oxygen atoms by protons). Thus, we can relate the non-zero elements of the transition matrix to the topology of the hydrogen bond network, because we assume that the proton transfer between oxygen atoms does only occur for small oxygen–oxygen distances (which correspond to the formation of a hydrogen bond). On the one hand, this interplay of the topology of the hydrogen bond network and the transition matrix enables the analytical interpretation of the shape of the transition matrices so that they can also be used to investigate the qualitative picture of proton movement in these materials. On the other hand, it is possible to derive information concerning the shape of the transition matrix from the knowledge of the structure and dynamics of the hydrogen bond network. The qualitative picture of an efficient proton long range transfer is (at least partially) governed by the Grotthuss mechanism, i.e., a proton transfer between oxygen atoms is followed by structural reorientation of the atomistic neighborhood. This involves a strong as well as fluctuating hydrogen bond network. For sufficient long simulation times, we can assume that all oxygen atoms are connected via a chain of hydrogen bonds (at different times). This translates into the picture of an irreducible (as well as doubly stochastic and aperiodic) transition matrix. A detailed derivation of the general properties of the transition matrix of a good proton conductor from the topology

of the hydrogen bond network is given in Ref. 73. An irreducible as well as aperiodic transition matrix is called ergodic, and the shortest duration of the underlying molecular dynamics simulation which leads to ergodicity of the transition matrix can be used to determine the effective equilibration time scale of the systems as a whole.

For all “good” proton conducting systems, there is a unique stationary distribution of the protons to all chemically equivalent oxygen atoms: the uniform distribution. In Ref. 73, we show that the transition matrices within the MDM model, obtained from sufficient long molecular dynamics trajectories, have only one stationary distribution (or fixed point): the uniform distribution. This correct asymptotic behavior confirms the conceptual consistency of the MDM method. In this article, we put this approach to the next level. We will demonstrate that the MDM model is able to predict descriptors of long range proton transfer such as the mean square displacement as well as the diffusion coefficient for the real proton conducting materials hexakis(*p*-phosphonatophenyl)benzene (HPB) and  $\text{CsH}_2\text{PO}_4$ . Furthermore, we will make use of the low dimensionality of our approach in order to give an example for the extended system sizes and time scales which are accessible within our scale bridging method. Therefore, we will simulate the micrometer and millisecond proton dynamics of a  $\text{CsH}_2\text{PO}_4$  lattice, which will give us access to the investigation of inhomogeneities at the nanometer scale. In a first step, we explain how to obtain observables from the Markov chain constructed within our model, and we benchmark the MDM approach for ideal test systems.

## II. MEAN SQUARE DISPLACEMENT FROM THE TRANSITION MATRIX

While we reduce the system to the positions of the oxygen atoms, it is more challenging to calculate the macroscopic observables within the MDM approach compared with the analysis of standard MD simulations because we can no longer explicitly follow a particular atom.

An important descriptor for the long range proton conduction is the diffusion coefficient. Using the Einstein relation, the diffusion coefficient can be obtained as the asymptotic slope of the mean square displacement (MSD) as

$$D = \frac{1}{6} \lim_{t \rightarrow \infty} \frac{d}{dt} \text{MSD}(t). \quad (4)$$

In the following, we outline how to derive the diffusion coefficient from our MDM model. Two types of information are available to calculate the diffusion coefficient within the MDM method: the probability  $(\mathbf{P}^{(0,t)})_{ij}$  for a proton transfer from the  $i$ th to the  $j$ th oxygen atom within the time interval  $t$  and the positions of the oxygen atoms. As a first approximation, we identify the positions of the protons with the positions of the corresponding oxygen atoms because the exact proton positions cannot be retrieved from the proposed model. This introduces displacements that are too large for small time intervals of the proton evolution because the smallest movement is the entire distance between two oxygen atoms. Nevertheless, the intermediate behavior of the mean square displacement remains untouched, which should be used to determine the diffusion coefficient. For relatively long proton evolution periods, the

proton motion seems too slow because the transition matrix only describes proton transfer within the given oxygen grid and does not take into account periodic boundary conditions. In Sec. VI, we will demonstrate that it is straightforward to increase the size of the transition matrix. Thus, in order to obtain a uniform slope of the mean square displacement, we can systematically enlarge the intermediate regime.

In an explicit molecular dynamics simulation, the numerical value of the mean square displacement depends on (a) the exact initial conformation  $R(t = 0)$  and (b) the exact evolution of the trajectory, which is formally deterministic albeit in practice always chaotic/random. In such an explicit simulation, the diffusion is, therefore, always computed on the basis of an averaged mean square displacement, using many different initial configurations [which implies an averaging over both  $R(t)$  for different times  $t$  and the actual trajectory evolution].

In the context of our MDM approach, this double averaging is automatically incorporated by considering (a) a probability vector for the initial configuration, i.e., an initial state which contains probabilities for having a particular oxygen site protonated, and (b) using transition probabilities instead of the deterministic integration of the equation of motion, i.e., using the fluxes  $(\mathbf{P}^{(0,t)})_{ij}$ .

The mean square displacement at a macroscopic time  $t$  can be decomposed into contributions from every possible pair of oxygen sites  $(i, j)$ . Each of these pairs contributes the squared distance between the two oxygen atoms  $d_{ij} = \Delta R(O_i, O_j)$ , weighted with the integrated flux  $(\mathbf{P}^{(0,t)})_{ij}$  between those sites since  $t = 0$ ,

$$[\text{MSD}(t)]_{ij} = (\mathbf{P}^{(0,t)})_{ij} \cdot (d_{ij})^2. \quad (5)$$

Summing over all elementary proton fluxes (i.e., all matrix elements) and dividing by the number of particles, we obtain the overall mean square displacement  $\text{MSD}(t)$  for the interval  $t$ ,

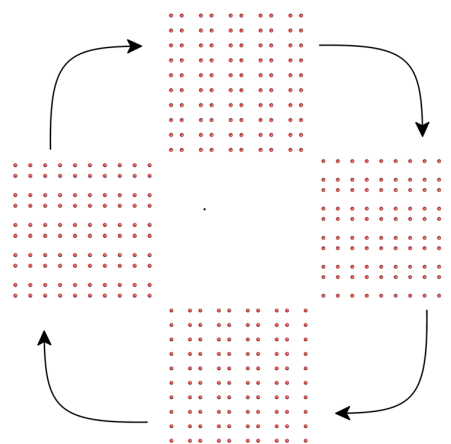
$$\text{MSD}(t) = \frac{1}{M} \sum_{i,j=1}^M [\text{MSD}(t)]_{ij}. \quad (6)$$

### III. APPLICATION OF THE MDM TRANSITION MATRIX TO MODEL SYSTEMS

#### A. Description of the test systems

Before we apply our dynamic model to calculate diffusion coefficients of real proton conducting materials, we study two well-defined model systems of proton conductors in order to better understand the properties of our approach. For each test system, fictitious trajectories were created, with typical oxygen–oxygen distances resulting in typical jump probabilities for the protons (when compared to real proton conducting materials):

- Test system 1: A single O–O dimer with a fixed distance.
- Test system 2: A dynamically evolving, two-dimensional oxygen lattice with periodic boundary conditions consisting of  $10 \times 10$  oxygen atoms. The evolution of the positions of the oxygen atoms is repeated periodically after four frames of the artificial trajectory (see Fig. 3). Each oxygen atom has four neighbors at different distances. For each time step, there is only one neighbor that is sufficiently close to



**FIG. 3.** The four periodically repeated frames of the trajectory of test system 2. In each frame, any given site has only one neighbor that is sufficiently close to allow proton transfer.

allow proton transfer. This nearest oxygen neighbor, however, changes in the next time step, imitating a fluctuating hydrogen bond network.

Since the fictitious trajectories of our test systems were not generated by *ab initio* molecular dynamics, a reference for the comparison of the results obtained by the MDM approach is required. Therefore, we used a simple Metropolis algorithm for the description of explicit stochastic jumps between the sites  $i$  and  $j$ . In this reference calculation, the sites can be occupied or unoccupied by protons. Proton jumps are only possible between pairs of an occupied and an unoccupied site, and the probability of a jump is determined by the distance between the sites. This type of calculation can be done within a recently developed program package in our group, published as the combined molecular dynamics (cMD)/lattice Monte Carlo approach (LMC).<sup>74–77</sup> We will refer to the reference calculations by the term “Metropolis reference.” Furthermore, we will denote our MDM method (as described in the Sec. I) as the “native MDM method.”

#### B. Results of the (native) MDM transition matrices

In Fig. 4, we show the mean square displacement for the oxygen–oxygen dimer with a fixed distance, calculated with the Metropolis reference and the (native) MDM approach using the transition matrices obtained by Eqs. (1) and (2). Both methods yield the same mean square displacement.

In Figs. 5 and 6, we show the mean square displacement for the fluctuating oxygen lattice calculated with the Metropolis reference and the (native) MDM approach using the transition matrices obtained by Eqs. (1) and (2). For one proton, both methods agree. If 50 protons are distributed among the oxygen atoms, the slope of the mean square displacement calculated with the Metropolis reference is significantly lower compared to the slope obtained with the MDM method. The probabilistic description of the protonation state and protonation dynamics of the MDM scheme lacks any explicit

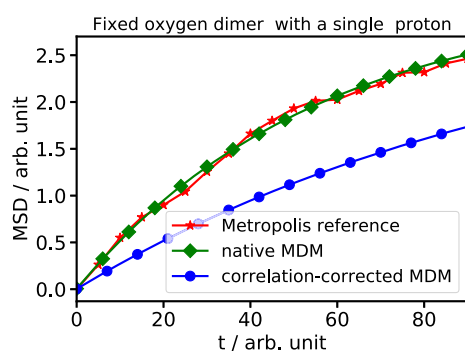


FIG. 4. Mean square displacement of a single proton for a fixed oxygen dimer.

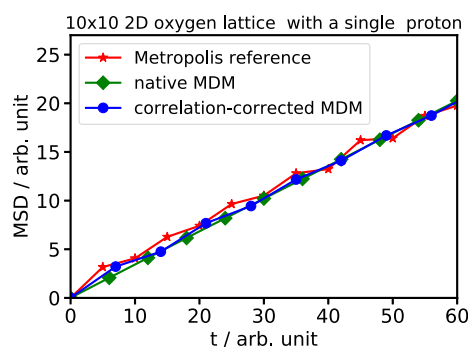


FIG. 5. Mean square displacement of a single proton for a dynamically evolving  $10 \times 10$  oxygen lattice.

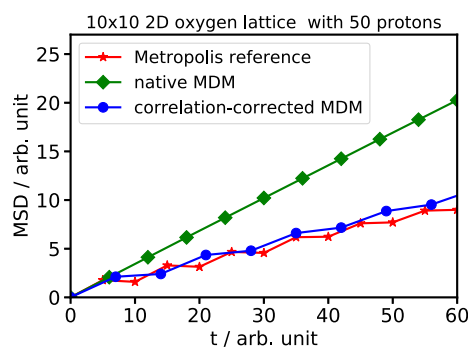


FIG. 6. Mean square displacement of 50 protons for a dynamically evolving  $10 \times 10$  oxygen lattice.

correlation between actual particles. The probability for the transfer of a proton from one oxygen atom to another is independent of the (partial) occupation number of either oxygen atom. This is different from the situation in the reference simulation. Here, any oxygen site

is either occupied or unoccupied, and a jump can only occur between a pair of an occupied and an unoccupied site.

#### IV. INCORPORATION OF PROTON CORRELATION WITHIN THE MDM APPROACH (CORRELATION-CORRECTED MDM)

In order to include proton correlation into our model, it is sufficient to modify the matrix elements of the transition matrix with the current occupation of the respective oxygen sites. In each time step, the probability  $(\mathbf{P}^{(n,n+1)})_{ij}$  of a proton transfer between the  $j$ th and  $i$ th oxygen atoms ( $i \neq j$ ) has to be weighted by the factor  $(1 - x_i^{(n)})$ ,

$$(\tilde{\mathbf{P}}^{(n,n+1)})_{ij} = (\mathbf{P}^{(n,n+1)})_{ij} \cdot (1 - x_i^{(n)}). \quad (7)$$

The diagonal elements are determined by the condition that the overall number of protons should remain constant,

$$(\tilde{\mathbf{P}}^{(n,n+1)})_{jj} = 1 - \sum_{i=1, i \neq j} (\tilde{\mathbf{P}}^{(n,n+1)})_{ij}. \quad (8)$$

#### A. Application of the correlation-corrected MDM transition matrix to model systems

In Figs. 5 and 6, we show the mean square displacement for the fluctuating oxygen lattice, calculated with the Metropolis reference and the correlation-corrected MDM approach obtained by Eqs. (7) and (8).

For a single proton, both methods still agree. If 50 protons are distributed among the oxygen atoms, only the correlation-corrected MDM approach yields a mean square displacement in agreement with the results of the Metropolis reference. The native MDM variant results in a far too diffusive behavior. This result illustrates the importance of considering explicit correlations between the charge carriers, which effectively limits their mobility when the occupation density is increased.

In Fig. 4, we show the mean square displacement for the oxygen-oxygen dimer with a fixed distance, calculated with the Metropolis reference and the correlation-corrected MDM approach obtained by Eqs. (7) and (8). The inclusion of proton correlation leads to a decrease in the slope of the mean square displacement. Here, the results for the first generation of the MDM approach were already correct, and thus, the results for the correlation-corrected MDM approach are somewhat less accurate. This is expected because the transfer of a single proton within the oxygen dimer is trivially uncorrelated. However, in the correlation-corrected MDM approach, any proton density reduces the probability of a proton jump to an oxygen atom, regardless of whether this density originated from the original proton itself. This residual error which can be characterized as a kind of “self-correlation” could in principle be corrected in order to bring also the single-proton-diffusion value to an agreement with the Metropolis reference simulation. For simulations containing essentially uncorrelated charge carriers at high occupation densities (in this example, one proton and two sites, yielding 50% density), this should actually be done. However, our aim is to address more “crowded” situations in which the protons are actually correlated so that the correlation correction is essential.



## B. Properties of the correlation-corrected MDM transition matrix

In order to incorporate the proton correlation, we slightly changed the protocol for the determination of the elements of the transition matrix [cf. Eqs. (7) and (8)]. In this section, we verify formally that these modifications do not deteriorate the qualitative properties of the MDM model such as its asymptotic behavior.

Within the correlation-corrected MDM approach, the elements of the transition matrices are scaled/modified. However, no zero entries are added or removed from the matrix. Thus, properties of the transition matrix such as being irreducible, non-negative, and aperiodic are transferred from the native to the correlation-corrected MDM model. The correlation-corrected transition matrices are also stochastic by construction [cf. Eqs. (7) and (8)]. These properties ensure the existence of a stationary distribution.

Adding the condition of being doubly stochastic to the transition matrix enforces the stationary distribution (or fixed point) to be the uniform distribution of the protons.

Within the native MDM approach, the elementary transition matrices were stochastic and symmetric (and, thus, also doubly stochastic) due to the definition of the matrix elements according to Eqs. (1) and (2). In contrast, the elementary correlation-corrected transition matrices are neither symmetric nor doubly stochastic, and thus, their successive products are also not doubly stochastic for arbitrary initial proton distributions  $x_0$ . Therefore, it is not clear if the correlation-corrected transition matrices also possess the uniform distribution as stationary distribution.

In order to answer this question, we will prove by an explicit calculation that the Markov chains generated from the elementary native transition matrices  $\mathbf{P}^{(n, n+1)}$  and the elementary correlation-corrected transition matrices  $\tilde{\mathbf{P}}^{(n, n+1)}$  are identical. To do so, we will apply the two different matrices to an arbitrary vector  $x^{(n)} \in \mathbb{R}^M$  of the proton distribution to the oxygen atoms at the  $n$ th time step,

$$x^{(n+1)} = \mathbf{P}^{(n, n+1)} x^{(n)}, \quad (9)$$

$$\tilde{x}^{(n+1)} = \tilde{\mathbf{P}}^{(n, n+1)} x^{(n)}. \quad (10)$$

By the following calculation, we will show that the difference  $\Delta x^{(n+1)} := \tilde{x}^{(n+1)} - x^{(n+1)}$  between the two resulting proton distributions is equal to zero:

$$\Delta x^{(n+1)} = \tilde{\mathbf{P}}^{(n, n+1)} x^{(n)} - \mathbf{P}^{(n, n+1)} x^{(n)} = \underbrace{\left[ \tilde{\mathbf{P}}^{(n, n+1)} - \mathbf{P}^{(n, n+1)} \right]}_{\mathbf{P}^\Delta} x^{(n)}. \quad (11)$$

The off-diagonal elements of the matrix  $\mathbf{P}^\Delta x^{(n)}$  are read as

$$\left( \mathbf{P}^\Delta \right)_{ij} = \left( \tilde{\mathbf{P}}^{(n, n+1)} \right)_{ij} - \left( \mathbf{P}^{(n, n+1)} \right)_{ij} \quad (12)$$

$$\stackrel{\text{Eq. (7)}}{=} \left[ \left( \mathbf{P}^{(n, n+1)} \right)_{ij} \cdot \left[ 1 - \left( x^{(n)} \right)_i \right] \right] - \left[ \left( \mathbf{P}^{(n, n+1)} \right)_{ij} \right] \\ = - \left( \mathbf{P}^{(n, n+1)} \right)_{ij} \left( x_i^{(n)} \right). \quad (13)$$

For the diagonal elements, we have

$$\left( \mathbf{P}^\Delta \right)_{jj} = \left( \tilde{\mathbf{P}}^{(n, n+1)} \right)_{jj} - \left( \mathbf{P}^{(n, n+1)} \right)_{jj} \quad (14)$$

$$\stackrel{\text{Eq. (2)}}{\stackrel{\text{Eq. (8)}}{=}} \left[ 1 - \sum_{\substack{i=1 \\ i \neq j}}^M \left( \mathbf{P}^{(n, n+1)} \right)_{ij} \right] \cdot \left[ 1 - \left( x^{(n)} \right)_i \right] \\ - \left[ 1 - \sum_{\substack{i=1 \\ i \neq j}}^M \left( \mathbf{P}^{(n, n+1)} \right)_{ij} \right], \quad (15)$$

$$= \sum_{\substack{i=1 \\ i \neq j}}^M \left( \mathbf{P}^{(n, n+1)} \right)_{ij} \left( x_i^{(n)} \right). \quad (16)$$

Now, Eq. (11) reads in element form as

$$\left( \Delta x^{(n+1)} \right)_i = \sum_{j=1}^M \left( \mathbf{P}^\Delta \right)_{ij} \left( x^{(n)} \right)_j = \left( \mathbf{P}^\Delta \right)_{ii} \left( x^{(n)} \right)_i + \sum_{\substack{j=1 \\ j \neq i}}^M \left( \mathbf{P}^\Delta \right)_{ij} \left( x^{(n)} \right)_j \quad (17)$$

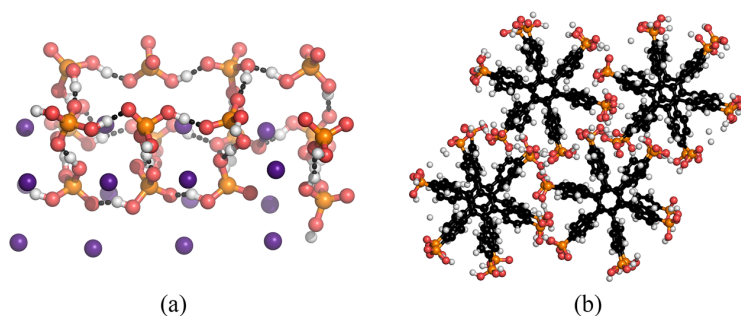
$$\stackrel{\text{Eq. (13)}}{\stackrel{\text{Eq. (16)}}{=}} \sum_{\substack{j=1 \\ j \neq i}}^M \left( \mathbf{P}^{(n, n+1)} \right)_{ji} \left( x^{(n)} \right)_j \left( x^{(n)} \right)_i \\ - \left( \mathbf{P}^{(n, n+1)} \right)_{ij} \left( x^{(n)} \right)_i \left( x^{(n)} \right)_j = 0. \quad (18)$$

In the last step, we used the symmetry of the elementary native transition matrices  $\left[ \left( \mathbf{P}^{(n, n+1)} \right)_{ji} = \left( \mathbf{P}^{(n, n+1)} \right)_{ij} \right]$ . We have shown that the proton distributions obtained by the propagation of the initial proton distribution by the elementary native transition matrices  $\mathbf{P}^{(n, n+1)}$  and the elementary correlation-corrected transition matrices  $\tilde{\mathbf{P}}^{(n, n+1)}$  are equal. It is worth noting that we obtain the same Markov chain regardless of the changes in the transition matrix introduced by the correlation correction. Thus, the stationary distribution (obtained by repeated applications of the elementary transition matrices) is the same for the native as well as the correlation-corrected MDM model: the uniform distribution.

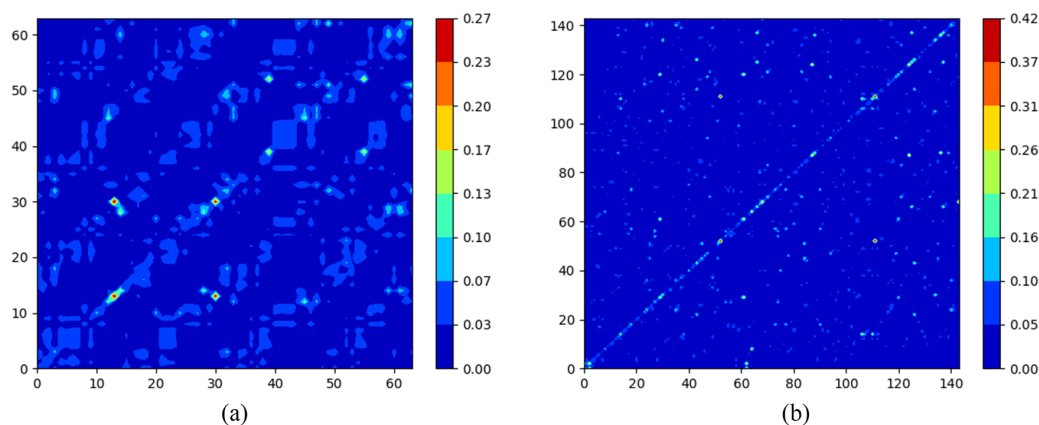
The rescaling of the off-diagonal elements of the native transition matrix [cf. Eq. (7)] does not alter the evolution of the proton distribution but represents a direct rescaling of proton fluxes between oxygen atoms, which are decisive for the calculation of the mean square displacement [cf. Eq. (5)].

## V. APPLICATION OF THE MDM METHOD TO REAL PROTON CONDUCTING COMPOUNDS

We have applied our new MDM simulation scheme to two different compounds which represent a “soft” and a “hard” proton conductor: Hexakis(*p*-phosphonatophenyl)benzene and the high-temperature cubic phase of  $\text{CsH}_2\text{PO}_4$  (an inorganic crystal). Both compounds are interesting from an application perspective, as they combine high proton conductivity with interesting structural features, and are suitable candidates for future low-humidity membrane materials for proton-exchange membrane fuel cells. Snapshots of the local molecular structure for both systems are shown in Fig. 7. The proton dynamics in both systems have been previously



**FIG. 7.** Potential membrane materials for proton-exchange membrane fuel cells. (a) CsH<sub>2</sub>PO<sub>4</sub>. (b) Hexakis(*p*-phosphonatophenyl)benzene. Figure 7(a) is reproduced with permission from Dreßler *et al.*, J. Chem. Phys. **152**, 114114 (2020). Copyright 2020 AIP Publishing LLC.

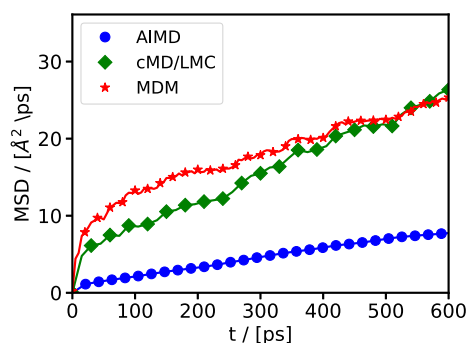


**FIG. 8.** Assembled transition matrices. (a) Visualization of the transition matrix  $\mathbf{P}_{\text{MD}}$  of CsH<sub>2</sub>PO<sub>4</sub> for an underlying 850 ps AIMD simulation. (b) Visualization of the transition matrix  $\mathbf{P}_{\text{MD}}$  of hexakis(*p*-phosphonatophenyl)benzene for an underlying 115 ps AIMD simulation. Figure 8(a) is reproduced with permission from Dreßler *et al.*, J. Chem. Phys. **152**, 114114 (2020). Copyright 2020 AIP Publishing LLC.

simulated by means of conventional *ab initio* molecular dynamics (AIMD) simulations.<sup>78,79</sup> We already combined these calculations with the explicit Monte Carlo variant of our stochastic long-range simulation method (termed cMD/LMC).<sup>74–76</sup> Thus, they represent ideal candidates for the quantitative validation and transferability demonstration of the MDM approach. Computational details of the underlying AIMD simulations are given in Appendixes A and B. Assembled transition matrices for the investigated compounds obtained by the MDM method are shown in Fig. 8.

We use the MDM and the Monte Carlo reference cMD/LMC method to simulate the proton dynamics of hexakis(*p*-phosphonatophenyl)benzene and CsH<sub>2</sub>PO<sub>4</sub>. The resulting mean square displacement functions are depicted in Figs. 9 and 10. Comparing the mean square displacement obtained from the MDM method and that from the reference Monte Carlo propagation (cMD/LMC), we observe a certain offset for CsH<sub>2</sub>PO<sub>4</sub>, which originates from the initial simulation period ( $t = 0–25$  ps). Apart from this offset, the time evolution agrees well, corresponding to a good reproduction of the diffusion constant, as illustrated in Table I. The same holds true for the organic compound (Fig. 10). In comparison to the diffusion constants obtained in the explicit AIMD simulations, the agreement is again very good, the deviations stay within a factor of two. Similarly, the final comparison to experimental

reference data confirms the good overall agreement and, thus, illustrates that all levels of coarse graining are within the acceptable range. The diffusion coefficient obtained from an AIMD simulation is highly sensitive to the length of the trajectory and the actual protocol for the determination of the diffusion coefficient (e.g., the fitting



**FIG. 9.** Comparison of the mean square displacement of CsH<sub>2</sub>PO<sub>4</sub> for AIMD, MDM, and cMD/LMC methods.

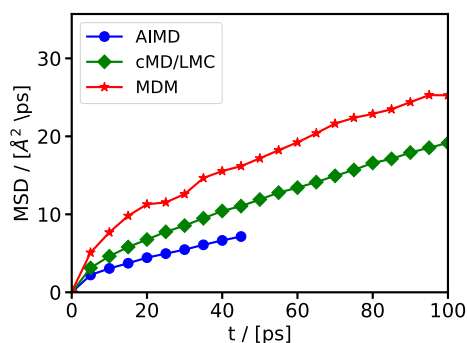


FIG. 10. Comparison of the mean square displacement of hexakis (*p*-phosphonatophenyl)benzene for AIMD, MDM, and cMD/LMC methods.

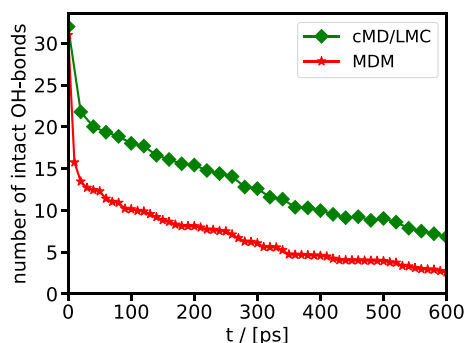


FIG. 11. Comparison of covalent O–H bond autocorrelation functions of  $\text{CsH}_2\text{PO}_4$  for the MDM and the cMD/LMC methods.

procedure to the MSD). Details for the calculation of the diffusion coefficients from the AIMD trajectory are given in Appendix B.

Apart from the diffusion coefficient, another observable of the proton dynamics can be extracted from the MDM model. The  $i$ th diagonal element of the transition matrix corresponds to the fraction of the protons which remain at the  $i$ th oxygen atom within the corresponding time interval described by the transition matrix. Hence, the trace of the transition matrix describes the overall number of protons which remain at their initially covalently bound oxygen atoms. By successive multiplication of the elementary transition matrices  $\mathbf{P}^{(1,2)}$ ,  $\mathbf{P}^{(2,3)}$ , ...,  $\mathbf{P}^{(N-1,N)}$ , we obtain transition matrices  $\mathbf{P}^{(1,N)}$  which describe the integrated proton dynamics for  $N$  elementary time steps, i.e., for total times of  $N \cdot \Delta t$ . The trace of these matrices yields the autocorrelation function of the covalent O–H bonds, which describes the fraction of O–H bonds that remain intact during a period of  $N \cdot \Delta t$ .<sup>74</sup> Thus, we have an easy access to this autocorrelation function of the covalent O–H bonds using the MDM approach. In Fig. 11, we show the time dependent trace of the transition matrix obtained by the MDM method and the corresponding autocorrelation function of the covalent O–H bonds calculated using the reference cMD/LMC approach. Both functions are similar to each other, and their long term decay matches very well.

In this section, we have calculated several quantitative descriptors of long range proton transfer (diffusion coefficients and

covalent O–H bond autocorrelation functions) for the compounds hexakis(*p*-phosphonatophenyl)benzene and  $\text{CsH}_2\text{PO}_4$ . The MDM approach yields the same results compared to the reference cMD/LMC method and AIMD simulations. While we have illustrated the correct asymptotic behavior of our model previously,<sup>73</sup> we have demonstrated here the general applicability of our model toward the prediction of proton dynamics in real proton conducting compounds. The MDM approach further condenses (compared to the cMD/LMC approach) the entire proton dynamics of the underlying molecular dynamics trajectory into an  $M \times M$  matrix. Here, we verified that despite this considerable coarse graining, the method still yields a similarly accurate description. All important features of proton transfer are integrated at high accuracy within the MDM propagation model.

## VI. INCREASING THE SYSTEM SIZE AND TIME SCALE USING THE MDM METHOD

In this section, we demonstrate how to use the remarkable reduction of the dimensionality within the MDM model to increase the system size and time scale of the simulations. The MDM approach enables the storage of the proton dynamics of the entire underlying *ab initio* molecular dynamics simulation within a single transition matrix  $\mathbf{P}_{\text{MD}}$ . The application of the transition matrix  $\mathbf{P}_{\text{MD}}$  to a proton distribution vector  $x$  will correspond to the evolution of the proton distribution for a time interval of the full length of the underlying molecular dynamics simulation. It is immediately apparent that we can simulate huge periods of proton movement by repeated application of the transition matrix  $\mathbf{P}_{\text{MD}}$ , i.e., repeated simple matrix–vector multiplications.

Due to the limitations of the system sizes to the nanometer scale within *ab initio* molecular dynamics simulations, only inhomogeneities at the subnanometer scale can be incorporated in such simulations. However, the properties of many systems interesting for material sciences or engineering are governed by inhomogeneities at the micro- or nanometer scales. Within the MDM approach, the information about the proton dynamics of the entire underlying *ab initio* molecular dynamics simulation can be stored within  $M \times M$  elements of the transition matrix  $\mathbf{P}_{\text{MD}}$  ( $M$  is the number of oxygen atoms of the system under investigation). This allows for a

TABLE I. Diffusion coefficients [ $10^{-3} \text{Å}^2/\text{ps}$ ] from AIMD, cMD/LMC, and MDM.

	$\text{CsH}_2\text{PO}_4$ 510 K	Hexakis( <i>p</i> -phosphonatophenyl) benzene 600 K
$D_k$ cMD/LMC	6.0	29.8
$D_k$ MDM	4.0	37.1
$D_k$ AIMD	2.1	18.7
Experimental	2.9–25 <sup>a</sup> 0.5–6.5 <sup>b</sup>	

<sup>a</sup>Reference 107.

<sup>b</sup>See Refs. 80–84.



drastic increase in the system size. The MDM method is capable of handling several hundred thousands of atoms instead of several hundred. Assuming the length of the entire *ab initio* molecular dynamics as an elementary time step of the MDM approach, proton transfer will only occur in a nanoscopic environment with respect to the initial proton positions, which will be reflected by a sparse transition matrix. Due to this sparsity, the storage and application of transition matrices describing several millions of atoms is now possible, which corresponds to characteristic lengths in the millimeter scale.

The remaining crucial point for the construction of transition matrices of much larger systems is the determination of the positions of the oxygen atoms. These positions serve as a basis for the construction of the hopping rate function and, thus, directly determine the values of the propagation matrices  $\mathbf{P}^{(n,n+1)}$ . The straightforward idea of using *ab initio* molecular dynamics simulations (as done here for the proof-of-principle validations) quickly reaches its limits when large-scale inhomogeneities are involved. A natural further choice is atomistic force field based simulations, which can extend the length scale by several orders of magnitude. A complementary way is the fragmentation of the entire system into (partially overlapping) domains whose atomistic dynamical structure can be determined individually. The resulting fragment trajectories can then be re-assembled for use within the MDM method, yielding a natural path toward actual microporous systems.

## VII. SIMULATION OF PROTON DYNAMICS FOR THE REAL WORLD SYSTEMS AT THE MICROMETER AND MICROSECOND SCALES

We give a proof-of-principle example for the extension of the time and size scales within the MDM method. To do so, we resort

to a bulk system of the high-temperature cubic phase of  $\text{CsH}_2\text{PO}_4$  constructed from about 500 000 atoms. This results in cell parameters of  $8 \mu\text{m} \times 1 \text{nm} \times 1 \text{nm}$ . The transition matrix was constructed using the following protocol, which is also illustrated in Fig. 12: First, we construct the transition matrix  $\mathbf{P}_{\text{MD}}$  of the small system from the originally *ab initio* molecular dynamics simulation (containing 16 ion pairs/64 oxygen atoms/system size:  $2 \times 1 \times 1 \text{nm}^3$ ). According to the periodic boundary conditions used for the *ab initio* molecular dynamics (AIMD) simulation, we can decompose the large bulk phase system into repeated periodic images of the original simulated small AIMD system. By adapting this translational symmetry, we construct the transition matrix of the larger system as a block matrix with block sizes according to the number  $M$  of oxygen atoms in the original AIMD simulation ( $M = 64$ ). The elements of the diagonal blocks describe the proton transfer within the subsystems corresponding to the size of the original systems. The elements of the off-diagonal blocks of the block matrix correspond to proton transfer between the sole subsystems. The entire protocol for the construction of the large transition matrix is also illustrated in Fig. 12.

The initial distribution was built by setting about 64 (0.02%) of the accessible oxygen sites to “occupied.” We used the resulting transition matrix to propagate a specific initial proton distribution for several microseconds. The evolution of this initial proton distribution vector is presented in Fig. 13. The proton distributions for the different evolution intervals correspond to the broadening of a Gaussian function and can be described mathematically by the diffusion distribution,

$$P(r_t, t) = (4\pi Dt)^{(-3/2)} \exp\left(-\frac{(r_t - r_0)^2}{4Dt}\right). \quad (19)$$

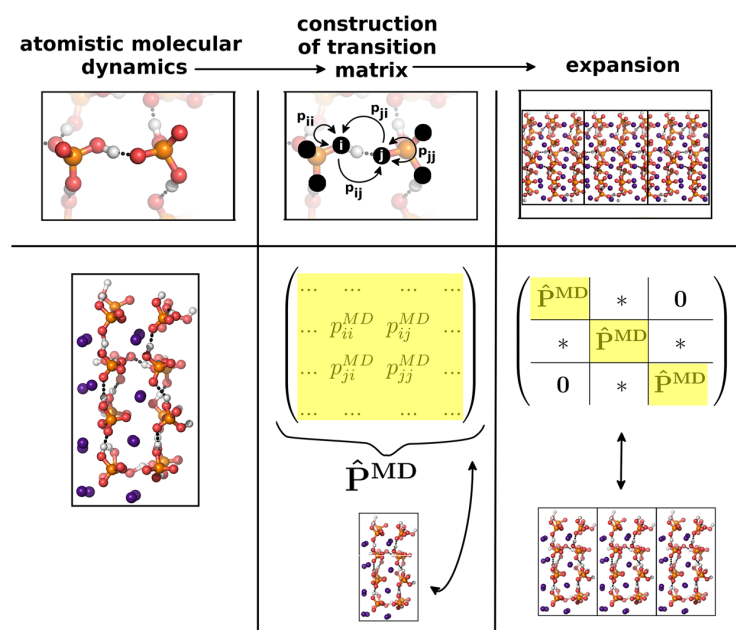
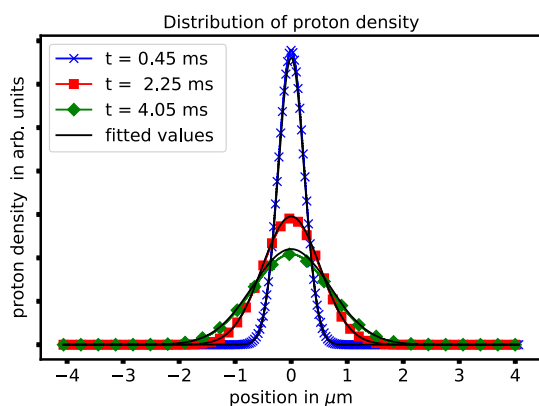


FIG. 12. Construction of huge transition matrices by assembling the propagation matrices from smaller fragments of the oxygen topology.



**FIG. 13.** Distribution of protons with respect to time and position in a  $8 \mu\text{m} \times 1 \text{ nm} \times 1 \text{ nm}$   $\text{CsH}_2\text{PO}_4$  bulk system. The resulting distributions were fitted according to the diffusion equation [Eq. (19)] in order to determine the diffusion coefficient.

Herein,  $P(r_t, t)$  denotes the probability to find a particle with the initial position  $r_0$  at position  $r_1$  at time  $t$ . Via fitting Eq. (19) to the resulting proton distributions of the MDM model, we determined the diffusion coefficient of the protons to be  $5.9 \cdot 10^{-3} \text{ \AA}^2/\text{ps}$ .

The time dependent mean square displacement  $\langle (\vec{r}_t - \vec{r}_0)^2 \rangle$  [see Eq. (4)] can be derived from the diffusion distribution [Eq. (19)] via

$$\langle (\vec{r}_t - \vec{r}_0)^2 \rangle = \int P(r_t, t) (\vec{r}_t - \vec{r}_0)^2 d\vec{r}_t. \quad (20)$$

There is an alternative expression for the determination of the diffusion coefficient, which is derived from the Kubo transform, involving the velocity–velocity autocorrelation function,

$$D = \frac{1}{3} \int_0^\infty \langle \vec{v}(\tau) \cdot \vec{v}(0) \rangle d\tau. \quad (21)$$

In principle, all discussed approaches (fitting to the diffusion equation, taking the slope of the mean square displacement as well as the integration of the velocity autocorrelation function) are equivalent methods for the calculation of the diffusion coefficients and should yield the same value.<sup>85,86</sup> However, due to the limitation to short time scales, which arise from the AIMD simulations, the diffusion coefficients obtained by the mean square displacement and velocity autocorrelation function may differ. The MDM method enables the determination of the diffusion coefficient via a third approach: the direct fit of the numerically obtained proton distribution to the analytical diffusion distribution [cf. Eq. (19)]. The resulting value can be employed in order to judge which of the diffusion coefficients obtained by the other two methods is more trustworthy. For the specific case of the very long AIMD simulation of  $\text{CsH}_2\text{PO}_4$ , the diffusion coefficients obtained from the mean square displacement of the AIMD and the MDM method (cf. Table 1) and the diffusion coefficient obtained from the fit to the diffusion equation agree quite well.

### A. Limitations of the MDM model

In its present form, the MDM model is not capable of simulating proton transport in more complex hydrogen bonded systems such as liquid water. In such systems, the proton transfer probability is not a function of a simple (local) geometric criterion alone. In the particular case of water, changes in the solvation shell (coordination number and coordination pattern) of the molecules are decisive for proton transport. In water, the symmetrization of the local hydrogen bond topology between the proton-donating and proton-receiving oxygen atoms is the key step for the initialization of a proton transfer event. This symmetrization requires equalization of the local coordination numbers.<sup>87,88</sup> In fact, the hydrogen bond length between hydronium ions and water, for example, tends to be between 2.4 Å and 2.5 Å; however, the proton generally cannot hop until the proton-receiving oxygen atom reduces its coordination number from four to three. Several studies on kinetic models for proton transport in aqueous systems have already been published, based, for example, on population correlation functions of the protonic defect.<sup>63,89,90</sup> However, these models are only suitable for rather short time scales (i.e., shorter than a typical AIMD simulation length).

In the future, we are planning to extend our model to the description of more complex hydrogen bonded systems such as liquid water. The criteria based on coordination numbers (as described above) can fortunately be incorporated into the construction of the transition matrix.

At the moment, the construction of the transition matrix is based on *ab initio* molecular dynamics simulations and, therefore, does not take into account any nuclear quantum effects. One approach to account for such effects would be to construct the transition matrix from path integral molecular dynamics (PIMD) simulations, which would, however, significantly increase the required computational time. Another alternative is to include the contribution of nuclear quantum effects via post-processing of the transition matrix. We are planning to investigate this in future studies.

### VIII. CONCLUSION AND OUTLOOK

In this work, we explicitly computed the non-equilibrium molecular dynamics of protons in the solid acid  $\text{CsH}_2\text{PO}_4$  on the micrometer length scale on a quasi-atomistic level based on a probabilistic matrix propagation approach. We have shown that an excess proton distribution which is initially localized within a few nanometers diffuses through the  $8 \mu\text{m}$  sized system within 5 ms, in full agreement with the common diffusion laws. The probabilistic propagation scheme utilizes an  $M \times M$  transition matrix whose dimension equals the number of possible proton sites in the system, which can be of the order of several millions due to the sparsity of the matrix. This allows the realistic (quasi-atomistic) treatment of nanometer-sized morphological inhomogeneities within a limited computational effort.

The use of *ab initio* molecular dynamics simulations as a basis for the probabilistic transition matrix ensures a highly accurate description of elementary chemical facets such as hydrogen bond structure and hydrogen bond network dynamics. A special methodological focus of the present work is the incorporation of explicit correlation between the instantaneous protonation states of the system during the propagation.

Our new approach allows the explicit simulation of non-equilibrium phenomena in systems of direct experimental interest, touching the length and time scales of real proton conducting devices such as advanced fuel cell proton exchange membranes.<sup>91</sup>

#### ACKNOWLEDGMENTS

This work was supported by the Deutsche Forschungsgemeinschaft (Project-ID No. 435886714).

#### APPENDIX A: COMPUTATIONAL DETAILS

We applied Born–Oppenheimer Molecular Dynamics (BOMD) using the CP2K<sup>92</sup> program package to simulate the proton transfer in the solid acid CsH<sub>2</sub>PO<sub>4</sub> and hexakis(*p*-phosphonatophenyl)-benzene (HPB). We utilized the Quickstep module<sup>93</sup> and orbital transformation<sup>94</sup> for faster convergence. The electronic structure was calculated with density functional theory utilizing the PBE<sup>95–97</sup> functional for CsH<sub>2</sub>PO<sub>4</sub> and BLYP<sup>98,99</sup> for HPB. A basis set of the type DZVP-MOLOPT-SR-GTH<sup>100</sup> and GTH pseudopotentials<sup>101,102</sup> were applied. Furthermore, we used the empirical dispersion correction (D3) from Grimme.<sup>103</sup> The temperature was set by using a Nosé–Hoover chain thermostat<sup>104–106</sup> (NVT ensemble). The time step was chosen as 0.5 fs. A system of 16 formula units (128 atoms) of the high-temperature cubic phase of CsH<sub>2</sub>PO<sub>4</sub> was investigated at 510 K for 900 ps. Concerning the HPB system, 816 atoms were calculated for 115 ps at 600 K. The dimensions of the simulation box and the starting configurations of the CsH<sub>2</sub>PO<sub>4</sub> system were obtained from crystal structure data.<sup>85</sup> For the high-temperature structure of CsH<sub>2</sub>PO<sub>4</sub>, partial occupancies of the oxygen atoms have been reported in the literature. From the set of partial occupancies, we chose one possible set of coordinates and used them as the initial structure for the geometry optimization. Afterward, we performed a 10 ps equilibration run using massive and global thermostatting. It turned out that the choice of the initial set of coordinates from the partial occupancies did not affect the resulting trajectory due to the large amount of orientational disorder of the oxygen tetrahedra.

#### APPENDIX B: CALCULATION OF DIFFUSION COEFFICIENTS FROM AIMD TRAJECTORIES

The diffusion coefficient obtained from an *ab initio* molecular dynamics simulation is highly sensitive to the length of the trajectory and the actual protocol for the determination of the diffusion coefficient (e.g., the fitting procedure to the mean square displacement). According to Eq. (4), the diffusion coefficient is obtained from the slope of a linear equation which is fitted to the linear regime of the mean square displacement. On the one hand, it is often hard to judge (for a short *ab initio* molecular dynamics simulation) if the linear part of the mean square displacement is reached. On the other hand, the statistics of particle displacements (in particular, for proton transfer as a rare event) is not sufficient for larger times of the mean square displacement.

Following this line, special care has to be taken in order to report proton diffusion coefficients. In awareness of these problems, we applied the following protocol for the determination of the diffusion coefficients for CsH<sub>2</sub>PO<sub>4</sub>. We calculated the diffusion

coefficients for several lengths of the trajectory and monitored its convergence behavior: For a 60 ps trajectory, we reported a diffusion coefficient of  $5.9 \cdot 10^{-3} \text{ \AA}^2/\text{ps}$  in a previous article.<sup>76</sup> With increasing trajectory length, we observed a decrease in the diffusion coefficient (e.g.,  $3.2 \cdot 10^{-3} \text{ \AA}^2/\text{ps}$  for 200 ps trajectory length). For trajectory lengths between 700 ps and 950 ps, the diffusion coefficients remained constant at  $2.1 \cdot 10^{-3} \text{ \AA}^2/\text{ps}$ . We used this value as the diffusion coefficient of CsH<sub>2</sub>PO<sub>4</sub> in Table I. We performed a linear fit of the mean square displacement between 25 ps and 50 ps for all trajectory lengths of more than 150 ps.

#### REFERENCES

- 1 L. Delle Site, “What is a multiscale problem in molecular dynamics?,” *Entropy* **16**, 23–40 (2014).
- 2 M. Praprotnik, L. Delle Site, and K. Kremer, “Multiscale simulation of soft matter: From scale bridging to adaptive resolution,” *Annu. Rev. Phys. Chem.* **59**, 545–571 (2008).
- 3 K. Wendler, F. Dommert, Y. Y. Zhao, R. Berger, C. Holm, and L. Delle Site, “Ionic liquids studied across different scales: A computational perspective,” *Faraday Discuss.* **154**, 111–132 (2012).
- 4 J. A. Morrone, R. Zhou, and B. J. Berne, “Molecular dynamics with multiple time scales: How to avoid pitfalls,” *J. Chem. Theory Comput.* **6**, 1798–1804 (2010).
- 5 J. A. Morrone, T. E. Markland, M. Ceriotti, and B. J. Berne, “Efficient multiple time scale molecular dynamics: Using colored noise thermostats to stabilize resonances,” *J. Chem. Phys.* **134**, 014103 (2011).
- 6 A. Perez, J. A. Morrone, E. Brini, J. L. MacCallum, and K. A. Dill, “Blind protein structure prediction using accelerated free-energy simulations,” *Sci. Adv.* **2**, e1601274 (2016).
- 7 G. Tabacchi, C. J. Mundy, J. Hutter, and M. Parrinello, “Classical polarizable force fields parametrized from *ab initio* calculations,” *J. Chem. Phys.* **117**, 1416–1433 (2002).
- 8 J. Behler, “First principles neural network potentials for reactive simulations of large molecular and condensed systems,” *Angew. Chem., Int. Ed.* **56**, 12828–12840 (2017).
- 9 J. Behler, “Constructing high-dimensional neural network potentials: A tutorial review,” *Int. J. Quantum Chem.* **115**, 1032–1050 (2015).
- 10 B. Nebgen, N. Lubbers, J. S. Smith, A. E. Sifain, A. Likhov, O. Isayev, A. E. Roitberg, K. Barros, and S. Tretiak, “Transferable dynamic molecular charge assignment using deep neural networks,” *J. Chem. Theory Comput.* **14**, 4687–4698 (2018).
- 11 K. T. Butler, D. W. Davies, H. Cartwright, O. Isayev, and A. Walsh, “Machine learning for molecular and materials science,” *Nature* **559**, 547 (2018).
- 12 M. Hellström and J. Behler, “Concentration-dependent proton transfer mechanisms in aqueous NaOH solutions: From acceptor-driven to donor-driven and back,” *J. Phys. Chem. Lett.* **7**, 3302–3306 (2016).
- 13 M. Gastegger, C. Kauffmann, J. Behler, and P. Marquetand, “Comparing the accuracy of high-dimensional neural network potentials and the systematic molecular fragmentation method: A benchmark study for all-trans alkanes,” *J. Chem. Phys.* **144**, 194110 (2016).
- 14 V. Quaranta, M. Hellström, and J. Behler, “Proton-transfer mechanisms at the water–ZnO interface: The role of presolvation,” *J. Phys. Chem. Lett.* **8**, 1476–1483 (2017).
- 15 L. J. LaBerge and J. C. Tully, “A rigorous procedure for combining molecular dynamics and Monte Carlo simulation algorithms,” *Chem. Phys.* **260**, 183–191 (2000).
- 16 E. C. Neyts and A. Bogaerts, “Combining molecular dynamics with Monte Carlo simulations: Implementations and applications,” *Theor. Chem. Acc.* **132**, 1320 (2012).

- <sup>17</sup>F. Noé, J. Chodera, G. Bowman, V. Pande, and F. Noé, *An Introduction to Markov State Models and Their Application to Long Timescale Molecular Simulation*, Advances in Experimental Medicine and Biology Vol. 797 (Springer, 2014).
- <sup>18</sup>G. R. Bowman, E. R. Bolin, K. M. Hart, B. C. Maguire, and S. Marqusee, "Discovery of multiple hidden allosteric sites by combining Markov state models and experiments," *Proc. Natl. Acad. Sci. U. S. A.* **112**, 2734–2739 (2015).
- <sup>19</sup>B. E. Husic and V. S. Pande, "Markov state models: From an art to a science," *J. Am. Chem. Soc.* **140**, 2386–2396 (2018).
- <sup>20</sup>V. S. Pande, K. Beauchamp, and G. R. Bowman, "Everything you wanted to know about Markov State Models but were afraid to ask," *Methods* **52**, 99–105 (2010).
- <sup>21</sup>C. Schütte and M. Sarich, "A critical appraisal of Markov state models," *Eur. Phys. J.: Spec. Top.* **224**, 2445–2462 (2015).
- <sup>22</sup>C. R. Schwantes, R. T. McGibbon, and V. S. Pande, "Perspective: Markov models for long-timescale biomolecular dynamics," *J. Chem. Phys.* **141**, 090901 (2014).
- <sup>23</sup>B. M. Forrest and U. W. Suter, "Hybrid Monte Carlo simulations of dense polymer systems," *J. Chem. Phys.* **101**, 2616–2629 (1994).
- <sup>24</sup>D. G. Gromov and J. J. de Pablo, "Structure of binary polymer blends: Multiple time step hybrid Monte Carlo simulations and self-consistent integral-equation theory," *J. Chem. Phys.* **103**, 8247–8256 (1995).
- <sup>25</sup>A. Irbäck, "Hybrid Monte Carlo simulation of polymer chains," *J. Chem. Phys.* **101**, 1661–1667 (1994).
- <sup>26</sup>D. W. Heermann and L. Yixue, "A global-update simulation method for polymer systems," *Macromol. Chem. Phys.* **2**, 299–308 (1993).
- <sup>27</sup>I. Martin-Bragado, R. Borges, J. P. Balbuena, M. Jaraiz, "Kinetic Monte Carlo simulation for semiconductor processing: A review," *Prog. Mater. Sci.* **92**, 1–32 (2018).
- <sup>28</sup>G. Betz and W. Husinsky, "A combined molecular dynamics and kinetic Monte Carlo calculation to study sputter erosion and beam assisted deposition," *Nucl. Instrum. Methods Phys. Res., Sect. B* **193**, 352–358 (2002).
- <sup>29</sup>A. Ghoufi and G. Maurin, "Hybrid Monte Carlo simulations combined with a phase mixture model to predict the structural transitions of a porous metal-organic framework material upon adsorption of guest molecules," *J. Phys. Chem. C* **114**, 6496–6502 (2010).
- <sup>30</sup>A. A. Knizhnik, A. A. Bagaturyants, I. V. Belov, B. V. Potapkin, and A. A. Korokin, "An integrated kinetic Monte Carlo molecular dynamics approach for film growth modeling and simulation: ZrO<sub>2</sub> deposition on Si(100) surface," *Comput. Mater. Sci.* **24**, 128–132 (2002).
- <sup>31</sup>U. H. E. Hansmann and Y. Okamoto, "New Monte Carlo algorithms for protein folding," *Curr. Opin. Struct. Biol.* **9**, 177–183 (1999).
- <sup>32</sup>E. K. Peter and J.-E. Shea, "A hybrid MD-kMC algorithm for folding proteins in explicit solvent," *Phys. Chem. Chem. Phys.* **16**, 6430–6440 (2014).
- <sup>33</sup>H. Zhang, "A new hybrid Monte Carlo algorithm for protein potential function test and structure refinement," *Proteins: Struct., Funct., Genet.* **34**, 464–471 (1999).
- <sup>34</sup>J.-H. Prinz, H. Wu, M. Sarich, B. Keller, M. Senne, M. Held, J. D. Chodera, C. Schütte, and F. Noé, "Markov models of molecular kinetics: Generation and validation," *J. Chem. Phys.* **134**, 174105 (2011).
- <sup>35</sup>F. Nüske, B. G. Keller, G. Pérez-Hernández, A. S. J. S. Mey, and F. Noé, "Variational approach to molecular kinetics," *J. Chem. Theory Comput.* **10**, 1739–1752 (2014).
- <sup>36</sup>A. M. Berezhkovskii and A. Szabo, "Committors, first-passage times, fluxes, Markov states, milestones, and all that," *J. Chem. Phys.* **150**, 054106 (2019).
- <sup>37</sup>F. Noé and E. Rosta, "Markov models of molecular kinetics," *J. Chem. Phys.* **151**, 190401 (2019).
- <sup>38</sup>M. Dibak, M. J. del Razo, D. De Sancho, C. Schütte, and F. Noé, "MSM/RD: Coupling Markov state models of molecular kinetics with reaction-diffusion simulations," *J. Chem. Phys.* **148**, 214107 (2018).
- <sup>39</sup>B. G. Keller, J.-H. Prinz, and F. Noé, "Markov models and dynamical fingerprints: Unraveling the complexity of molecular kinetics," *Chem. Phys.* **396**, 92–107 (2012), special issue: Experimental and theoretical studies of protein dynamics and function: From femtoseconds to milliseconds.
- <sup>40</sup>B. G. Keller, A. Kobitski, A. Jäschke, G. U. Nienhaus, and F. Noé, "Complex RNA folding kinetics revealed by single-molecule FRET and hidden Markov models," *J. Am. Chem. Soc.* **136**, 4534–4543 (2014).
- <sup>41</sup>F. Nüske, H. Wu, J.-H. Prinz, C. Wehmeyer, C. Clementi, and F. Noé, "Markov state models from short non-equilibrium simulations—Analysis and correction of estimation bias," *J. Chem. Phys.* **146**, 094104 (2017).
- <sup>42</sup>C. Schütte, F. Noé, J. Lu, M. Sarich, and E. Vanden-Eijnden, "Markov state models based on milestoneing," *J. Chem. Phys.* **134**, 204105 (2011).
- <sup>43</sup>B. Trendelkamp-Schroer, H. Wu, F. Paul, and F. Noé, "Estimation and uncertainty of reversible Markov models," *J. Chem. Phys.* **143**, 174101 (2015).
- <sup>44</sup>F. Vitalini, A. S. J. S. Mey, F. Noé, and B. G. Keller, "Dynamic properties of force fields," *J. Chem. Phys.* **142**, 084101 (2015).
- <sup>45</sup>H. Wu, F. Paul, C. Wehmeyer, and F. Noé, "Multiensemble Markov models of molecular thermodynamics and kinetics," *Proc. Natl. Acad. Sci. U. S. A.* **113**, E3221–E3230 (2016).
- <sup>46</sup>E. Hruska, J. R. Abella, F. Nüske, L. E. Kavrakı, and C. Clementi, "Quantitative comparison of adaptive sampling methods for protein dynamics," *J. Chem. Phys.* **149**, 244119 (2018).
- <sup>47</sup>U. Sengupta, M. Carballo-Pacheco, and B. Strodel, "Automated Markov state models for molecular dynamics simulations of aggregation and self-assembly," *J. Chem. Phys.* **150**, 115101 (2019).
- <sup>48</sup>F. Noé, C. Schütte, E. Vanden-Eijnden, L. Reich, and T. R. Weikl, "Constructing the equilibrium ensemble of folding pathways from short off-equilibrium simulations," *Proc. Natl. Acad. Sci. U. S. A.* **106**, 19011–19016 (2009).
- <sup>49</sup>T. J. Lane, G. R. Bowman, K. Beauchamp, V. A. Voelz, and V. S. Pande, "Markov state model reveals folding and functional dynamics in ultra-long MD trajectories," *J. Am. Chem. Soc.* **133**, 18413–18419 (2011).
- <sup>50</sup>G. Pérez-Hernández, F. Paul, T. Giorgino, G. De Fabritiis, and F. Noé, "Identification of slow molecular order parameters for Markov model construction," *J. Chem. Phys.* **139**, 015102 (2013).
- <sup>51</sup>N. Stanley, S. Esteban-Martín, and G. De Fabritiis, "Kinetic modulation of a disordered protein domain by phosphorylation," *Nat. Commun.* **5**, 5272 (2014).
- <sup>52</sup>I. Buch, T. Giorgino, and G. De Fabritiis, "Complete reconstruction of an enzyme-inhibitor binding process by molecular dynamics simulations," *Proc. Natl. Acad. Sci. U. S. A.* **108**, 10184–10189 (2011).
- <sup>53</sup>M. Held, P. Metzner, J.-H. Prinz, and F. Noé, "Mechanisms of protein–ligand association and its modulation by protein mutations," *Biophys. J.* **100**, 701–710 (2011).
- <sup>54</sup>D.-A. Silva, G. R. Bowman, A. Sosa-Peinado, and X. Huang, "A role for both conformational selection and induced fit in ligand binding by the LAO protein," *PLoS Comput. Biol.* **7**, 1–11 (2011).
- <sup>55</sup>G. R. Bowman and P. L. Geissler, "Equilibrium fluctuations of a single folded protein reveal a multitude of potential cryptic allosteric sites," *Proc. Natl. Acad. Sci. U. S. A.* **109**, 11681–11686 (2012).
- <sup>56</sup>N. Plattner and F. Noé, "Protein conformational plasticity and complex ligand-binding kinetics explored by atomistic simulations and Markov models," *Nat. Commun.* **6**, 7653 (2015).
- <sup>57</sup>D. Chakraborty and D. J. Wales, "Dynamics of an adenine-adenine RNA conformational switch from discrete path sampling," *J. Chem. Phys.* **150**, 125101 (2019).
- <sup>58</sup>G. Pinamonti, F. Paul, F. Noé, A. Rodriguez, and G. Bussi, "The mechanism of RNA base fraying: Molecular dynamics simulations analyzed with core-set Markov state models," *J. Chem. Phys.* **150**, 154123 (2019).
- <sup>59</sup>J. A. Morrone and M. E. Tuckerman, "Ab initio molecular dynamics study of proton mobility in liquid methanol," *J. Chem. Phys.* **117**, 4403–4413 (2002).
- <sup>60</sup>J. A. Morrone, K. E. Haslinger, and M. E. Tuckerman, "Ab initio molecular dynamics simulation of the structure and proton transport dynamics of methanol–water solutions," *J. Phys. Chem. B* **110**, 3712–3720 (2006).
- <sup>61</sup>C. Dreßler, G. Kabbe, and D. Sebastiani, "Insight from atomistic simulations of protonation dynamics at the nanoscale," *Fuel Cells* **16**, 682–694 (2016).
- <sup>62</sup>R. Vuilleumier and D. Borgis, "Proton conduction: Hopping along hydrogen bonds," *Nat. Chem.* **4**, 432 (2012).



- <sup>63</sup>R. A. Krueger, L. Vilčiauskas, J.-P. Melchior, G. Bester, and K.-D. Kreuer, "Mechanism of efficient proton conduction in diphosphoric acid elucidated via first-principles simulation and NMR," *J. Phys. Chem. B* **119**, 15866–15875 (2015).
- <sup>64</sup>T. C. Berkelbach, H.-S. Lee, and M. E. Tuckerman, "Concerted hydrogen-bond dynamics in the transport mechanism of the hydrated proton: A first-principles molecular dynamics study," *Phys. Rev. Lett.* **103**, 238302 (2009).
- <sup>65</sup>L. Vilčiauskas, M. E. Tuckerman, G. Bester, S. J. Paddison, and K.-D. Kreuer, "The mechanism of proton conduction in phosphoric acid," *Nat. Chem.* **4**, 461 (2012).
- <sup>66</sup>R. Devanathan, N. Idupulapati, M. D. Baer, C. J. Mundy, and M. Dupuis, "Ab initio molecular dynamics simulation of proton hopping in a model polymer membrane," *J. Phys. Chem. B* **117**, 16522–16529 (2013).
- <sup>67</sup>V.-A. Glezakou, M. Dupuis, and C. J. Mundy, "Acid/base equilibria in clusters and their role in proton exchange membranes: Computational insight," *Phys. Chem. Chem. Phys.* **9**, 5752–5760 (2007).
- <sup>68</sup>M. D. Baer, J. L. Fulton, M. Balasubramanian, G. K. Schenter, and C. J. Mundy, "Persistent ion pairing in aqueous hydrochloric acid," *J. Phys. Chem. B* **118**, 7211–7220 (2014).
- <sup>69</sup>M. D. Baer, D. J. Tobias, and C. J. Mundy, "Investigation of interfacial and bulk dissociation of HBr, HCl, and HNO<sub>3</sub> using density functional theory-based molecular dynamics simulations," *J. Phys. Chem. C* **118**, 29412–29420 (2014).
- <sup>70</sup>J. Schmidt, J. VandeVondele, I.-F. W. Kuo, D. Sebastiani, J. I. Siepmann, J. Hutter, and C. J. Mundy, "Isobaric–isothermal molecular dynamics simulations utilizing density functional theory: An assessment of the structure and density of water at near-ambient conditions," *J. Phys. Chem. B* **113**, 11959–11964 (2009).
- <sup>71</sup>J. C. Grossman, E. Schwegler, E. W. Draeger, F. Gygi, and G. Galli, "Towards an assessment of the accuracy of density functional theory for first principles simulations of water," *J. Chem. Phys.* **120**, 300–311 (2004).
- <sup>72</sup>E. Schwegler, J. C. Grossman, F. Gygi, and G. Galli, "Towards an assessment of the accuracy of density functional theory for first principles simulations of water. II," *J. Chem. Phys.* **121**, 5400–5409 (2004).
- <sup>73</sup>C. Drefler, G. Kabbe, M. Brehm, and D. Sebastiani, "Dynamical matrix propagator scheme for large-scale proton dynamics simulations," *J. Chem. Phys.* **152**, 114114 (2020).
- <sup>74</sup>G. Kabbe, C. Wehmeyer, and D. Sebastiani, "A coupled molecular dynamics/kinetic Monte Carlo approach for protonation dynamics in extended systems," *J. Chem. Theory Comput.* **10**, 4221 (2014).
- <sup>75</sup>G. Kabbe, C. Drefler, and D. Sebastiani, "Toward realistic transfer rates within the coupled molecular dynamics/lattice Monte Carlo approach," *J. Phys. Chem. C* **120**, 19905–19912 (2016).
- <sup>76</sup>C. Drefler, G. Kabbe, and D. Sebastiani, "Proton conductivity in hydrogen phosphate/sulfates from a coupled molecular dynamics/lattice Monte Carlo (cMD/LMC) approach," *J. Phys. Chem. C* **120**, 19913–19922 (2016).
- <sup>77</sup>G. Kabbe, C. Drefler, and D. Sebastiani, "Proton mobility in aqueous systems: Combining *ab initio* accuracy with millisecond timescales," *Phys. Chem. Chem. Phys.* **19**, 28604–28609 (2017).
- <sup>78</sup>H.-S. Lee and M. E. Tuckerman, "The structure and proton transport mechanisms in the superprotonic phase of CsH<sub>2</sub>PO<sub>4</sub>: An *ab initio* molecular dynamics study," *J. Phys. Chem. C* **112**, 9917–9930 (2008).
- <sup>79</sup>C. Wehmeyer, M. Schrader, D. Andrienko, and D. Sebastiani, "Water-free proton conduction in hexakis(*p*-phosphonatophenyl)benzene nanochannels," *J. Phys. Chem. C* **117**, 12366–12372 (2013).
- <sup>80</sup>A. I. Baranov, V. P. Khiznichenko, and L. A. Shuvalov, "High temperature phase transitions and proton conductivity in some kdp-family crystals," *Ferroelectrics* **100**, 135–141 (1989).
- <sup>81</sup>S. M. Haile, H. Liu, and R. A. Secco, "High-temperature behavior of CsH<sub>2</sub>PO<sub>4</sub> under both ambient and high pressure conditions," *Chem. Mater.* **15**, 727–736 (2003).
- <sup>82</sup>E. Ortiz, R. A. Vargas, and B.-E. Mellander, "On the high-temperature phase transitions of CsH<sub>2</sub>PO<sub>4</sub>: A polymorphic transition? A transition to a superprotonic conducting phase?" *J. Chem. Phys.* **110**, 4847–4853 (1999).
- <sup>83</sup>K. Yamada, T. Sagara, Y. Yamane, H. Ohki, and T. Okuda, "Superprotonic conductor CsH<sub>2</sub>PO<sub>4</sub> studied by <sup>1</sup>H, <sup>31</sup>P NMR and X-ray diffraction," *Solid State Ionics* **175**, 557–562 (2004), special issue: Fourteenth International Conference on Solid State Ionics.
- <sup>84</sup>J. Otomo, N. Minagawa, C.-j. Wen, K. Eguchi, and H. Takahashi, "Protonic conduction of CsH<sub>2</sub>PO<sub>4</sub> and its composite with silica in dry and humid atmospheres," *Solid State Ionics* **156**, 357–369 (2003).
- <sup>85</sup>R. C. Desai and M. Nelkin, "Atomic motions in a rigid sphere gas as a problem in neutron transport," *Nucl. Sci. Eng.* **24**, 142–152 (1966).
- <sup>86</sup>O. J. Eder, "The velocity autocorrelation function and the diffusion coefficient for a dilute hard sphere gas," *J. Chem. Phys.* **66**, 3866–3870 (1977).
- <sup>87</sup>D. Marx, A. Chandra, and M. E. Tuckerman, "Aqueous basic solutions: Hydroxide solvation, structural diffusion, and comparison to the hydrated proton," *Chem. Rev.* **110**, 2174–2216 (2010).
- <sup>88</sup>D. Marx, "Proton transfer 200 years after von Grothuss: Insights from *ab initio* simulations," *ChemPhysChem* **7**, 1848–1870 (2006).
- <sup>89</sup>M. E. Tuckerman, A. Chandra, and D. Marx, "A statistical mechanical theory of proton transport kinetics in hydrogen-bonded networks based on population correlation functions with applications to acids and bases," *J. Chem. Phys.* **133**, 124108 (2010).
- <sup>90</sup>A. Chandra, M. E. Tuckerman, and D. Marx, "Connecting solvation shell structure to proton transport kinetics in hydrogen-bonded networks via population correlation functions," *Phys. Rev. Lett.* **99**, 145901 (2007).
- <sup>91</sup>M. Wagner, C. Drefler, F. P. Lohmann-Richters, K. Hanus, D. Sebastiani, A. Varga, and B. Abel, "Mechanism of ion conductivity through polymer-stabilized CsH<sub>2</sub>PO<sub>4</sub> nanoparticulate layers from experiment and theory," *J. Mater. Chem. A* **7**, 27367–27376 (2019).
- <sup>92</sup>J. Hutter, M. Iannuzzi, F. Schiffmann, and J. VandeVondele, "cp2k: atomistic simulations of condensed matter systems," *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **4**, 15–25 (2014).
- <sup>93</sup>J. VandeVondele, M. Krack, F. Mohamed, M. Parrinello, T. Chassaing, and J. Hutter, "Quickstep: Fast and accurate density functional calculations using a mixed Gaussian and plane waves approach," *Comput. Phys. Commun.* **167**, 103–128 (2005).
- <sup>94</sup>J. VandeVondele and J. Hutter, "An efficient orbital transformation method for electronic structure calculations," *J. Chem. Phys.* **118**, 4365–4369 (2003).
- <sup>95</sup>Y. Zhang and W. Yang, "Comment on 'generalized gradient approximation made simple'," *Phys. Rev. Lett.* **80**, 890 (1998).
- <sup>96</sup>J. P. Perdew, A. Ruzsinszky, G. I. Csonka, O. A. Vydrov, G. E. Scuseria, L. A. Constantin, X. Zhou, and K. Burke, "Restoring the density-gradient expansion for exchange in solids and surfaces," *Phys. Rev. Lett.* **100**, 136406 (2008).
- <sup>97</sup>J. P. Perdew, K. Burke, and M. Ernzerhof, "Generalized gradient approximation made simple," *Phys. Rev. Lett.* **77**, 3865–3868 (1996).
- <sup>98</sup>C. Lee, W. Yang, and R. G. Parr, "Development of the Colle–Salvetti correlation-energy formula into a functional of the electron density," *Phys. Rev. B* **37**, 785–789 (1988).
- <sup>99</sup>A. D. Becke, "Density-functional exchange-energy approximation with correct asymptotic behavior," *Phys. Rev. A* **38**, 3098–3100 (1988).
- <sup>100</sup>J. VandeVondele and J. Hutter, "Gaussian basis sets for accurate calculations on molecular systems in gas and condensed phases," *J. Chem. Phys.* **127**, 114105 (2007).
- <sup>101</sup>C. Hartwigsen, S. Goedecker, and J. Hutter, "Relativistic separable dual-space Gaussian pseudopotentials from H to Rn," *Phys. Rev. B* **58**, 3641–3662 (1998).
- <sup>102</sup>M. Krack, "Pseudopotentials for H to Kr optimized for gradient-corrected exchange-correlation functionals," *Theor. Chem. Acc.* **114**, 145–152 (2005).
- <sup>103</sup>S. Grimme, J. Antony, S. Ehrlich, and H. Krieg, "A consistent and accurate *ab initio* parametrization of density functional dispersion correction (DFT-D) for the 94 elements H–Pu," *J. Chem. Phys.* **132**, 154104 (2010).
- <sup>104</sup>S. Nosé, "A unified formulation of the constant temperature molecular dynamics methods," *J. Chem. Phys.* **81**, 511–519 (1984).
- <sup>105</sup>G. J. Martyna, M. L. Klein, and M. Tuckerman, "Nosé–Hoover chains: The canonical ensemble via continuous dynamics," *J. Chem. Phys.* **97**, 2635–2643 (1992).
- <sup>106</sup>S. Nosé, "A molecular dynamics method for simulations in the canonical ensemble," *Mol. Phys.* **52**, 255–268 (1970).
- <sup>107</sup>A. Ishikawa, H. Maekawa, T. Yamamura, Y. Kawakita, K. Shibata, and M. Kawai, *Solid State Ionics* **179**, 2345–2349 (2008).

### 3.1.2 Article II: Dynamical Matrix Propagator Scheme for Large-scale Proton Dynamics Simulations

Christian Dreßler, Gabriel Kabbe, Martin Brehm, and Daniel Sebastiani.  
Dynamical matrix propagator scheme for large-scale proton dynamics simulations.  
*The Journal of Chemical Physics*, 152(11):114114, **2020**.

In this article, I derived the theory and performed all calculations. G. Kabbe contributed with valuable discussions and developed the preceding cMD/LMC method, which was used as the starting point for the derivation of the MDM method. I wrote the manuscript. M. Brehm proofread the manuscript and contributed with valuable discussions. D. Sebastiani supervised the project and provided me with valuable advice while I was writing the manuscript.

# Dynamical matrix propagator scheme for large-scale proton dynamics simulations

Cite as: J. Chem. Phys. 152, 114114 (2020); doi: 10.1063/1.5140635

Submitted: 29 November 2019 • Accepted: 10 February 2020 •

Published Online: 20 March 2020



Christian Dreßler,  Gabriel Kabbe, Martin Brehm,  and Daniel Sebastiani<sup>a)</sup> 

## AFFILIATIONS

Institute of Chemistry, Martin Luther University Halle-Wittenberg, Von-Danckelmann-Platz 4, 06120 Halle (Saale), Germany

<sup>a)</sup>Author to whom correspondence should be addressed: [daniel.sebastiani@chemie.uni-halle.de](mailto:daniel.sebastiani@chemie.uni-halle.de)

## ABSTRACT

We derive a matrix formalism for the simulation of long range proton dynamics for extended systems and timescales. On the basis of an *ab initio* molecular dynamics simulation, we construct a Markov chain, which allows us to store the entire proton dynamics in an  $M \times M$  transition matrix (where  $M$  is the number of oxygen atoms). In this article, we start from common topology features of the hydrogen bond network of good proton conductors and utilize them as constituent constraints of our dynamic model. We present a thorough mathematical derivation of our approach and verify its uniqueness and correct asymptotic behavior. We propagate the proton distribution by means of transition matrices, which contain kinetic data from both ultra-short (sub-ps) and intermediate (ps) timescales. This concept allows us to keep the most relevant features from the microscopic level while effectively reaching larger time and length scales. We demonstrate the applicability of the transition matrices for the description of proton conduction trends in proton exchange membrane materials.

Published under license by AIP Publishing. <https://doi.org/10.1063/1.5140635>

## I. INTRODUCTION

In mathematics, a dynamical system is a system in which a fixed rule or function describes the evolution of a point (state) in a geometrical space (state space). One could think of the swinging of a pendulum or the predator–prey equations for the simulation of population dynamics.<sup>1,2</sup> Typically, such a dynamical system is low-dimensional and possesses a rigorously determined behavior.

In physics or chemistry, a dynamical system is described as a “particle or ensemble of particles whose state varies over time and thus obeys differential equations involving time derivatives.”<sup>3</sup> An example is molecular dynamics (MD) simulations in which a set of Newtonian equations of motion is integrated for a multi-atomic system. Typically, these high-dimensional problems have only numerical solutions (trajectories).

Our present work is an attempt to reconcile the properties of these two extremal cases of dynamical systems. We aim at keeping as many atomistic details as possible (corresponding to the high dimensionality of the problem) while establishing a mathematical framework with a highly rigorous character.

We start from the *ab initio* molecular dynamics (AIMD) method, which can be used to simulate the dynamics of atomistic processes including covalent bond formation or breaking, such as proton diffusion via a Grotthuss mechanism. Due to the approximate solution of the Schrödinger equation in each time step, *ab initio* molecular dynamics is a computationally very demanding approach, which limits the simulations to several hundreds of atoms and picoseconds.

In order to speed up molecular dynamics, a multitude of methods such as coarse graining, replica exchange, or adaptive MD schemes were developed.<sup>4–6</sup> A promising direction is the use of neural network potentials, which are trained on AIMD trajectory data.<sup>7–10</sup> Once trained, fully atomistic dynamics can be generated for very long time scales. The successful application of neural networks was already demonstrated for various systems such as NaOH solutions,<sup>11</sup> *n*-alkanes,<sup>12</sup> or water on ZnO surfaces.<sup>13</sup>

For the speed up of MD simulations, multiscale methods also play an important role. They make use of the atomistic details from a molecular dynamics simulation and bridge the scale toward a significantly increased system size/simulation time by employing a second approach.

Within the multiscale approaches, the combination of molecular dynamics simulations and a Markov model<sup>14–19</sup> or a Monte Carlo method<sup>20,21</sup> is widely used. A combination of MD and a Monte Carlo algorithm was already applied to a broad range of systems/problems such as polymer materials,<sup>22–25</sup> semiconductors,<sup>26</sup> surface phenomena,<sup>27–29</sup> and the folding of proteins.<sup>30–32</sup> The coupling of the two techniques can be done via different approaches such as the alternating application of MD and MC steps, the application of different approaches to different spatial subsystems, or the algorithmic mixture of the MD and the MC methods.<sup>20,21</sup>

In particular, for the description of dynamical processes of large molecular (often biomolecular) systems, Markov models constructed from MD simulations are well established.<sup>14–19,33–43</sup> The applicability of Markov state models has been demonstrated for a huge variety of systems, e.g., the prediction of protein folding<sup>44,45</sup> or the binding of small molecules to proteins<sup>46–54</sup> as well as the RNA folding kinetics.<sup>36,55,56</sup>

It was shown that MD simulations can be understood, under certain conditions, as a discrete model of continuous dynamic processes, which depend only on the current degrees of freedom (positions and velocities).<sup>39</sup> The latter corresponds to the Markov property such that the combination of a Markov state model and MD simulations becomes immediately apparent. Thus, after an appropriate discretization of the relevant dynamic processes, the important part of the long term kinetics can be extracted from an MD simulation within a discrete Markov state model.<sup>14,37,39,57</sup>

In this work, we aim toward a new discrete Markov state multiscale model, which is presently focused on describing condensed-phase ion diffusion, in particular, proton transport. First, we discuss the important properties of proton conductors, which are used as initial prerequisites for our model. In a second part, we review a recently developed scale bridging approach for proton transfer from a conceptual perspective. Finally, we present a thorough mathematical derivation of our new Markov model and verify its uniqueness and correct asymptotic behavior.

### A. General properties of proton conductors

In contrast to the transport of other types of ions, the transport of hydrogen ions, i.e., individual protons, can often not be understood as the diffusion of an independent particle within a homogeneous medium. A proton possesses a high charge-to-volume ratio compared to other ions. It is, therefore, strongly polarizing and has only extraordinary short lifetimes as a free particle. Proton transport is either bound to other particles (e.g., vehicle diffusion in the  $\text{H}_3\text{O}^+$  ion) or takes place via a Grotthuss-type mechanism.<sup>58–60</sup> By the term Grotthuss-like mechanism, we refer to a two-step process: protons transfer between molecules (“proton jumps”) with subsequent relaxation of the molecular environment. By means of the Grotthuss mechanism, a very high diffusivity of the protons is achieved, since the transport of the protons is independent of the displacement of the surrounding particles. Thus, such proton conductors are of particular interest for applications in engineering or material science (e.g., for proton conducting fuel cell membranes).

In any case, the Grotthuss mechanism requires a strong (to facilitate the proton transfer) but also fluctuating (to allow

the structural relaxation of the environment) hydrogen bond network.<sup>58–62</sup>

The prerequisite of strong as well as fluctuating hydrogen bonds leads to the following qualitative topological picture of the hydrogen bond network of a good proton conductor.<sup>62</sup> We assume that only oxygen atoms occur as hydrogen bond donors and acceptors. Then, at any point in time, the oxygen atoms are involved in a (strong) hydrogen bond, but after a certain time interval (governed by the relaxation time), the hydrogen bonds break and the oxygen atoms form new hydrogen bonds with (other) neighboring oxygen atoms. For larger timescales, any pair of oxygen atoms is linked together (via a sequence of hydrogen bonds at different times).

While this general description of the evolution of the hydrogen bond networks holds for many hydrogen bonded systems, there are specific other systems for which special additional care will be necessary. Here, a case in point is liquid water, where changes in the solvation shell (coordination number and coordination pattern) of the molecules are decisive for proton transport. In water, the symmetrization of the local hydrogen bond topology between the proton-donating and proton-receiving oxygen atoms is the key step for the initialization of a proton transfer event. This symmetrization requires equalization of the local coordination numbers.<sup>63,64</sup> Kinetic models for proton transport in aqueous systems have already been introduced based, for example, on population correlation functions of the protonic defect.<sup>58,65,66</sup> In summary, the most relevant single factor for the possibility of proton transfer in water is the coordination number and not a geometric criterion.

A computational approach for the simulation of proton transfer has to be able to take bond breaking into account. Under this prerequisite (for the usage of molecular dynamics), a quantum chemical method is, in principle, necessary for the calculation of the forces acting on the nuclei (*ab initio* MD). However, in this case, the simulation time and the system size are drastically restricted and only partial aspects of the proton transport can be described. Characteristic observables that describe the proton transport as a whole, such as diffusion coefficients, can only be roughly estimated by using *ab initio* MD.<sup>67–69</sup>

### B. Review of a recently developed scale bridging approach for proton transfer

In order to compensate for the drastic restriction of system size and simulation time, an attempt was made in our group to develop a hybrid simulation method, which is based on the data from an AIMD simulation of a proton conductor. The resulting combined Molecular Dynamics/Lattice Monte Carlo (cMD/LMC) approach uses the atomistic information from the MD simulations to model the microscopic proton transfer probabilities, and the Monte Carlo part takes the dynamically evolving molecular structure into account and performs the actual proton dynamics “on top” of that dynamical structure (for a graphic scheme of the method, see Fig. 1).<sup>70–73</sup>

The cMD/LMC method requires a short AIMD simulation of the proton conductor as an initial requirement. The resulting trajectory is then reduced to the evolution of oxygen coordinates, which can be either unoccupied or occupied by protons.



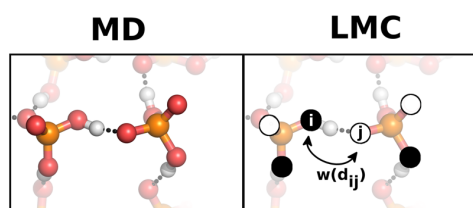


FIG. 1. The combined molecular dynamics/lattice Monte Carlo approach.

The probability for a proton transfer is described by a distance-dependent jump rate, which itself can also be determined directly from the AIMD trajectory. The probability of a proton transfer at a certain oxygen–oxygen distance and per time step of the AIMD trajectory is determined as the ratio of the number of proton jumps at a given oxygen–oxygen distance and the number of absolute occurrences of that oxygen–oxygen distance. Both the dividend and divisor can be obtained by detecting and counting these events from the AIMD trajectory. The dynamics of the proton motion is then simulated using the Monte Carlo method on the dynamically updated oxygen grid. The current oxygen grid is obtained from an MD trajectory. The simulated time scale of the proton dynamics can exceed the length of the underlying MD simulation because we can pass several times through the trajectory. The length of the Monte Carlo time step is related to the length of the MD time step. For each oxygen atom and each time step, the algorithm checks all jump possibilities to neighboring oxygen atoms in a certain radius. This procedure still scales at least linearly with the number of atoms in the system but comes along with a drastic reduced computational effort for the elementary step in the Monte Carlo part (compared to an elementary MD step). Following this protocol, the proton dynamics can be simulated for many nanoseconds on a single desktop PC (instead of several 100 ps on a large computing cluster). We already demonstrated that the cMD/LMC approach allows the description of long range proton transfer in several potential proton exchange membrane (PEM) materials for fuel cell applications [hexakis(*p*-phosphonatophenyl)benzene,  $\text{CsH}_2\text{PO}_4$ ,  $\text{CsHSO}_4$ , and  $\text{H}_2\text{O}$ ].<sup>70–73</sup>

In this work, we develop a model in which the overall dynamics of the proton motion of a molecular dynamics simulation is consolidated in a single  $M \times M$  matrix (where  $M$  is the number of oxygen atoms). As a further simplification, correlation between proton occupation numbers and proton hopping probabilities is treated implicitly. This allows the simulation of even larger time scales and a drastic reduction of storage requirements, which, in turn, enables an analytical interpretation of the high dimensional trajectory. Of course, the level of atomistic detail of the new method is not equivalent to the underlying AIMD simulation, but all important parts of the information governing the long range proton transfer are conserved in the new method.

We will first construct a heuristic model describing proton dynamics in good proton conductors. Afterward, we confirm that our model is well defined and we point out under which conditions our model predicts the long range proton dynamics accurately.

## II. DERIVATION OF A MARKOV MODEL FOR PROTON TRANSFER

### A. Definition of a dynamical system and Markov chains

Here, we review the general mathematical model for the description of a discrete dynamic process using (operator) semigroups: A discrete dynamical system is a 3-tuple  $(T, X, \Phi)$  consisting of a set  $T = \mathbb{N}_0$  or  $\mathbb{Z}$  (indexing the temporal evolution), a non-empty set  $X$  (the state space or the phase space), and an evolution function of the dynamical system  $\Phi: T \times X \rightarrow X$  so that for all states  $x \in X$  and all times  $t, s \in T$  we have

$$I: \quad \Phi(0, x) = x, \quad (1)$$

$$II: \quad \Phi(s, \Phi(t, x)) = \Phi(s + t, x) \quad (\text{semigroup action}). \quad (2)$$

With respect to the time  $t \in T$ , the function  $\Phi(t, x)$  maps every point  $x \in X$  to a unique image. For fixed  $x \in X$ , the function  $\beta_x: T \rightarrow X, t \mapsto \beta_x(t) := \Phi(t, x)$  is called flow through  $x$ . The set  $O(x) := \{\beta_x(t) | t \in T\}$  is called the trajectory (or orbit or phase curve) and describes the (temporal) evolution of an initial state vector  $x$ . By specification of an elementary physical time step  $\Delta s$ , we can assign a physical time to our discrete time via the relation  $t \in T \leftrightarrow t \cdot \Delta s$ .

We want to restrict ourselves to a specific dynamical system, which fulfills the following conditions:

- The phase space  $X$  is isomorphic to an Euclidean vector space  $\mathbb{R}^M$ .
- The evolution function  $\Phi(n, x)$  of the state vector  $x \in \mathbb{R}^M$  (corresponding to  $n$  elementary time steps) can be expressed by the  $n$ -fold application of a matrix  $\mathbf{P} \in \mathbb{R}^{M \times M}$  to the state vector  $x$ ,

$$\Phi(n, x) = \mathbf{P}^n x \quad (3)$$

with

$$\mathbf{P}^n = \mathbf{P} \times \mathbf{P} \times \dots \times \mathbf{P} \quad (n \text{ times}). \quad (4)$$

For this specific evolution function, the semigroup property can be expressed via matrix multiplication,

$$\Phi(s, \Phi(t, x)) = \mathbf{P}^s (\mathbf{P}^t x) = \mathbf{P}^{s+t} x = \Phi(s + t, x). \quad (5)$$

### B. Relation between Markov chains and dynamic models

A discrete Markov chain is a stochastic model of a discrete dynamical process. It describes a sequence of possible events that fulfill the Markov property. A process satisfies the Markov property if the probability of each event depends only on the current state of the system and is not dependent on previous events.

The Markov property corresponds to the semigroup action of the dynamical system: Let  $x_{t+\Delta t}$  be the future state of a dynamical system, let  $x_t$  be the present state of the system, and let  $x_0$  be the initial state of the system (at  $t = 0$ ). According to

$$x_{t+\Delta t} = \Phi(t + \Delta t, x_0) \stackrel{\text{Eq. (2)}}{=} \Phi(\Delta t, \Phi(t, x_0)) = \Phi(\Delta t, x_t), \quad (6)$$

the future state of a dynamical system is only dependent on the present state of the system.

## C. Discrete dynamical model for proton transfer

The choice of an appropriate discretization of the phase space is a crucial point for many mesoscale simulation methods. A proton in a hydrogen bonded system is covalently attached to an oxygen atom at an initial time  $t_0$ . At another time  $t_1 > t_0$ , the covalent bond has remained intact or the proton has changed its covalently bonded partner. The possible pathways for the evolution of the proton position are determined by the hydrogen bond network, which can be described by the positions of the oxygen atoms. Following this line, we obtain the discretization of the phase space by reducing the chemical structure of the proton conductor to the coordinates of the oxygen atoms or, more generally, to the donors and acceptors involved in the hydrogen bonds, hereinafter referred to as “sites.” Each site can be assigned a number between zero and one, which is the probability for the occupation of the site by a proton. Thus, if  $M$  is the number of sites, the model has an  $M$ -dimensional state space  $X = [0, 1]^M \subset \mathbb{R}^M$ .

The temporal evolution of the occupation of the oxygen atoms or sites by protons, starting from an initial proton distribution  $x$ , is described by an  $M \times M$  transition matrix. This approach would constitute a formal discrete dynamical model as defined by Eqs. (3)–(5).

The actual elements of the transition matrix are determined from the molecular dynamics simulation of the proton conductor by extracting the positions of the  $M$  oxygen atoms from the MD trajectory. Following the evolution of that trajectory, these positions are updated periodically, which relates the time step of the MD and the time step of our dynamical model. Thus, our oxygen topology will also *dynamically evolve* and incorporate atomistic fluctuations of the hydrogen bond network as simulated via MD.

Based on the information from the current MD step, the transition matrix  $\mathbf{P}^{(n,n+1)}$  follows as an explicit representation of the transformation

$$\hat{\mathbf{P}}^{(n,n+1)} : X \rightarrow X, \quad (7)$$

$$x \mapsto \hat{\mathbf{P}}^{(n,n+1)}(x), \quad (8)$$

which describes the evolution of the proton distribution from the  $n$ th to the  $(n+1)$ th step of the dynamical model. Here,  $\hat{\mathbf{P}}^{(n,n+1)} : X \rightarrow X$  with  $X = \mathbb{R}^M$  maps a point  $x$  of the phase space back into the phase space and can thus be understood as an  $M \times M$  matrix  $\mathbf{P}^{(n,n+1)}$ . The matrix element  $(\mathbf{P}^{(n,n+1)})_{ij}$  describes the probability of a proton transfer from the  $j$ th to the  $i$ th oxygen atom within the time interval from the  $n$ th to the  $(n+1)$ th time step. An illustration of this matrix formalism is given in Fig. 2.

Two consecutive temporal evolution processes are represented by multiplying the elementary transition matrices  $\mathbf{P}^{(n+1,n+2)}$  and  $\mathbf{P}^{(n,n+1)}$ ,

$$\mathbf{P}^{(n,n+2)} = \mathbf{P}^{(n+1,n+2)} \cdot \mathbf{P}^{(n,n+1)}. \quad (9)$$

Thus, the transition matrix  $\mathbf{P}^{(n,n+2)}$  describes the proton dynamics for the sum of the elementary time intervals. Eventually, the transition matrix corresponding to the full length of the MD simulation ( $N$  MD steps) can be obtained by successive multiplication of every transition matrix corresponding to an elementary time step,

$$\mathbf{P}_{\text{MD}} = \mathbf{P}^{(1,N)} = \mathbf{P}^{(N-1,N)} \cdot \dots \cdot \mathbf{P}^{(2,3)} \cdot \mathbf{P}^{(1,2)}. \quad (10)$$

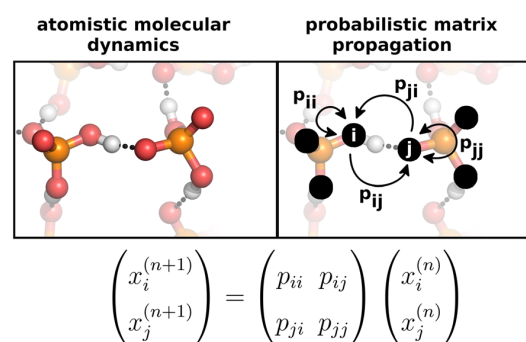


FIG. 2. The molecular dynamics/matrix propagation (MDM) approach. The transition matrix  $\mathbf{P}$ , which is obtained from the current position of the oxygen atoms, links the proton distribution  $x_n$  in the  $n$ -th step and the proton distribution  $x_{n+1}$  in the  $(n+1)$ th step.

With this procedure, it is possible to describe the entire dynamics of the proton transport of an MD simulation via a single  $M \times M$  matrix  $\mathbf{P}_{\text{MD}}$  (cf. Fig. 3). By applying this transformation matrix of the entire molecular dynamics simulation (of duration  $N \cdot \Delta t$ ) to the vector of the distribution of the protons  $x_v \in X$ , a new state vector  $x_{v+1} := \mathbf{P}_{\text{MD}}(x_v)$  is obtained, which describes the proton distribution after a time interval corresponding to the length of the entire molecular dynamics simulation. Note that the index  $n$  describes single MD time steps, while  $v$  enumerates the passes of the entire trajectory ( $N \cdot \Delta t$ ).

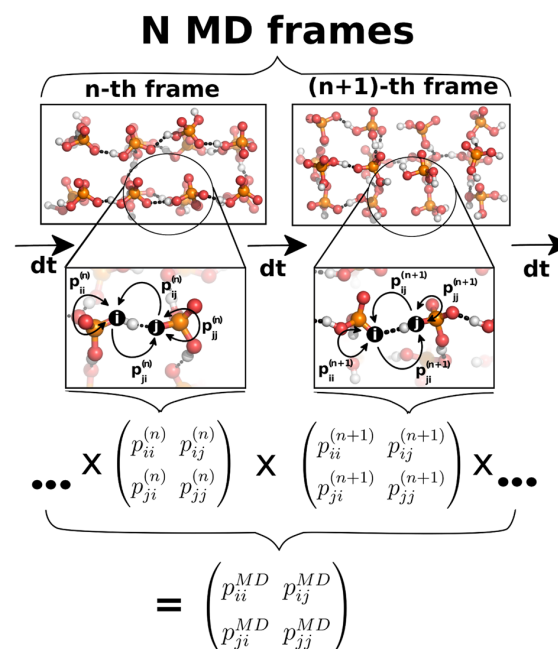
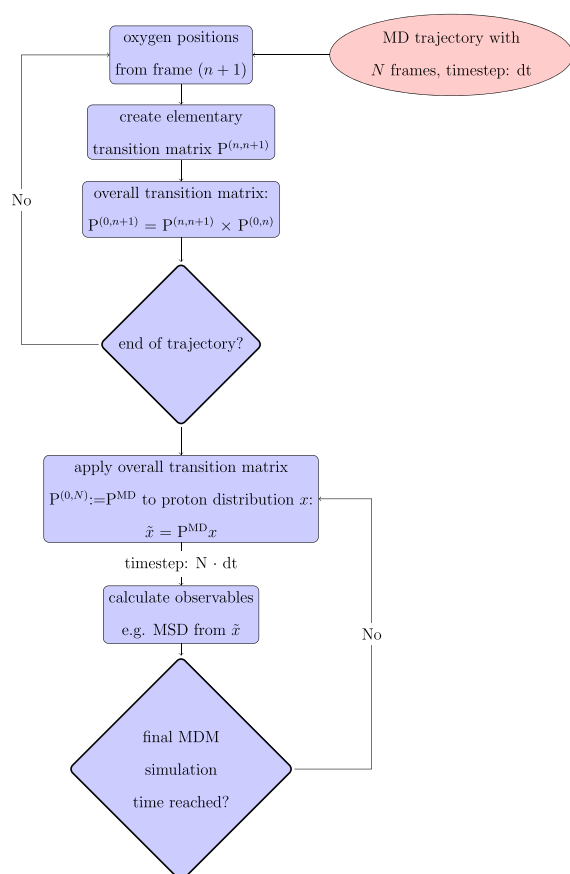


FIG. 3. Condensation of the proton dynamics of the entire MD trajectory into a single matrix.

By iterated application of the transition matrix  $\mathbf{P}_{\text{MD}}$  on  $x_0 \in X$ , we obtain a dynamical model  $\{(\mathbf{P}_{\text{MD}})^v x_0 \mid v \in \mathbb{N}\}$  according to the initial (mathematical) definition from Eqs. (3)–(5). The associated elementary time step of this dynamic model corresponds to the entire length of the MD. A scheme of the propagation of the proton distribution state via the transition matrix  $\mathbf{P}_{\text{MD}}$  is presented in Fig. 4. We refer to this hybrid simulation framework for the evolution of the proton dynamics as the molecular dynamics/matrix propagation (MDM) approach. Using a shorter associated time step of the dynamic model, e.g., the elementary time step of the MD simulation, we obtain a different model, which does not fulfill the regular evolution function from Eq. (3). In the [supplementary material](#), we discuss an adapted evolution function  $\tilde{\Phi}(t, \Delta t, x)$ , which would circumvent this problem.

Up to now, we claimed a formal mathematical model for the description of proton transfer, which is dependent on the current positions of the oxygen atoms updated from an MD trajectory. The question if such transition matrices for the description of proton transfer exist and if we can calculate them in an efficient manner is



**FIG. 4.** A scheme of the dynamic model for the propagation of the proton distribution using the length of the entire MD trajectory as elementary time step.

completely open. In Sec. II D, we present a concrete protocol for the construction of these matrices and apply this protocol to the AIMD trajectory of a typical proton conductor.

#### D. Calculation of the transition matrix—A concrete protocol

In order to calculate the elements  $(\mathbf{P}^{(n,n+1)})_{ij}$  of the elementary proton transition matrix  $\mathbf{P}^{(n,n+1)}$  at time step  $n$ , we have to specify a relation that links the positions of the  $i$ th oxygen atom  $O_i$  and  $j$ th oxygen atom  $O_j$  to the probability of a proton transfer between these atoms.

Within the framework of our cMD/LMC approach, we have recently developed a protocol for the calculation of the proton transfer probability as a function of the oxygen–oxygen distance (referred to as “jump rate function”).<sup>70–72</sup> The distance dependent jump rate function is obtained point-wise (i.e., for a list of discrete distances) from post-processing of an AIMD trajectory of the proton conductor.

The determination of this function for a certain oxygen–oxygen distance is based on the following idea: The frequency of occurrence of proton jumps (i.e., changes of the nearest oxygen neighbor of a proton) at a given oxygen–oxygen distance is divided by the overall frequency of occurrence of that oxygen–oxygen distance.

We fit the point-wise calculated jump rate function to a Fermi function according to

$$\omega(d_{ij}) = \frac{a}{1 + \exp\left(\frac{d_{ij}-b}{c}\right)}, \quad (11)$$

with  $d_{ij}$  being distance between the oxygen atoms  $O_i$  and  $O_j$ , and  $a$ ,  $b$ ,  $c$  being the fit parameters. Detailed information concerning the jump rate function is given in our previous publications.<sup>70–72</sup>

Now, the matrix element  $(\mathbf{P}^{(n,n+1)})_{ij}$  of the transition matrix  $\mathbf{P}^{(n,n+1)}$  is given by the product of the jump rate function  $\omega(d_{ij})$  at the current distance  $d_{ij}$  between the oxygen atoms  $O_i$  and  $O_j$  and the Heaviside step function  $\Theta(\alpha - \alpha_0)$  with  $\alpha$  being the X–O··O angle (with X = P in  $\text{CsH}_2\text{PO}_4$ , for example),

$$(\mathbf{P}^{(n,n+1)})_{ij} = \omega(d_{ij})\Theta(\alpha - \alpha_0), \quad \forall i \neq j, \quad (12)$$

with  $\alpha_0 = \pi/2$ . The Heaviside step function  $\Theta(\alpha - \alpha_0)$  is zero for  $\alpha < \alpha_0$  and one for  $\alpha \geq \alpha_0$ . This modification avoids nonphysical proton transfer events and suppresses any proton jumps between oxygen pairs with a geometry that does not allow for a linear hydrogen bond. In a previous paper, we have already demonstrated the necessity of this cut-off value in order to obtain realistic proton transfer rates between the oxygen atoms.<sup>71</sup>

Employing Eq. (12), we can assign a jump probability to every non-diagonal element of the transition matrix  $\mathbf{P}^{(n,n+1)}$ . The total number of protons must also be preserved after a proton transfer, so the sum of the rows of the transition matrices must be equal to one. This condition is enforced by setting the diagonal transition matrix elements according to

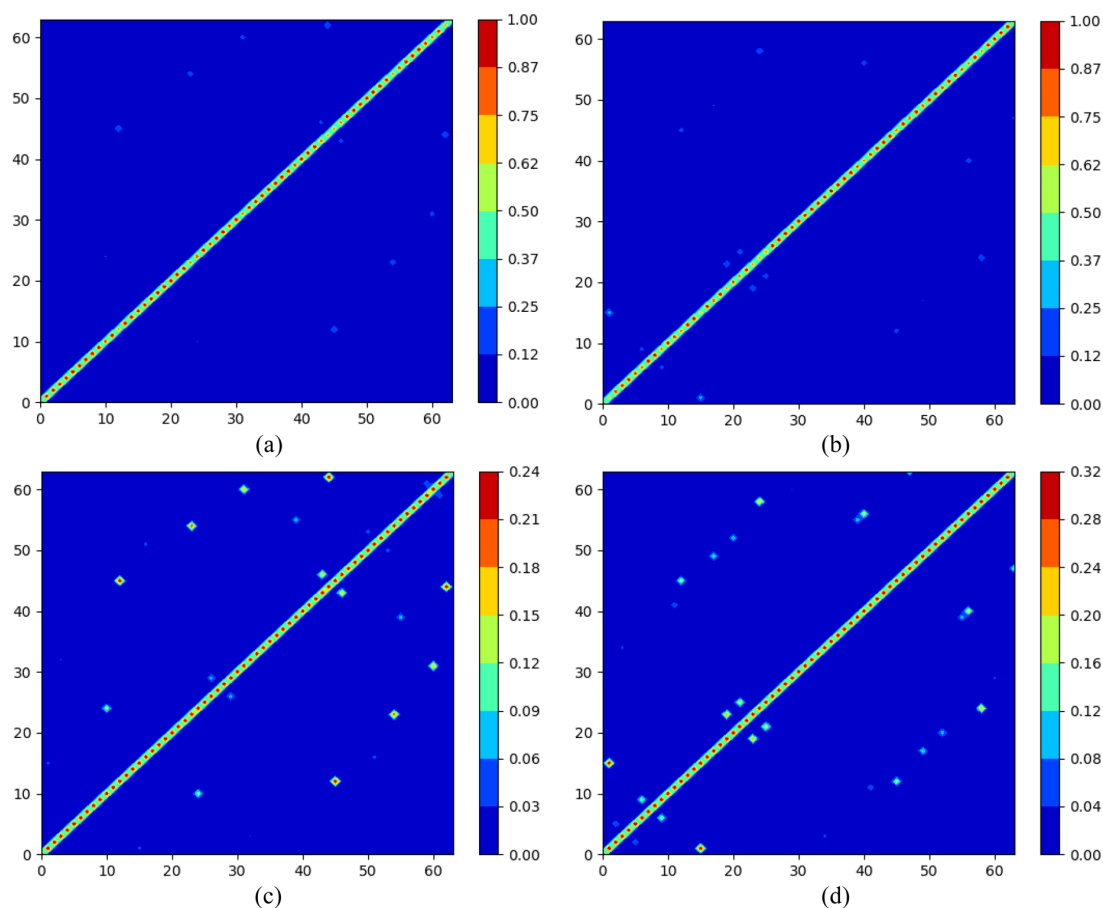
$$(\mathbf{P}^{(n,n+1)})_{jj} = 1 - \sum_{i=1, i \neq j} (\mathbf{P}^{(n,n+1)})_{ij} \quad \forall j. \quad (13)$$

### E. Transition matrices of the real proton conductor $\text{CsH}_2\text{PO}_4$

The high-temperature phase of  $\text{CsH}_2\text{PO}_4$  is an excellent proton conductor.<sup>74</sup> We have already performed an AIMD simulation with an elementary MD time step of 0.5 fs of this compound (and applied our cMD/LMC approach for the calculation of long range proton transfer).<sup>70-73</sup> Therefore, the jump rate function for  $\text{CsH}_2\text{PO}_4$  is already known (and shown in the [supplementary material](#)). Details on the preparation of the 900 ps AIMD simulation and processing of the data are given in the [Appendix](#). In the following, we restrict the processing of the AIMD trajectory to every 100th frame, i.e., the elementary time step of the MDM model is 50 fs. We calculate explicitly elementary transition matrices from the 900 ps trajectory of  $\text{CsH}_2\text{PO}_4$  via Eq. (12). Two representative elementary transition matrices for an early ( $t = 5$  ps) and a late ( $t = 450$  ps) time are shown in [Fig. 5](#).

We can report the following observations from the entirety of elementary transition matrices:

1. By construction [cf. Eq. (13)], the transition matrices for the elementary time steps are stochastic (sum of the rows of the matrix equal to one) and the matrix elements are greater than or equal to zero.
2. Due to the symmetry of the O–O distances (the distance between  $O_i$  and  $O_j$  is equal to the distance between  $O_j$  and  $O_i$ ), the transition matrices are also symmetric. Symmetric stochastic matrices are also doubly stochastic (the sum of the columns and the sum of the rows of a matrix is equal to one).
3. The elementary transition matrices are sparse. Only 1.7% of the oxygen atoms are involved in more than one hydrogen bond, i.e., in 98.3%, there is at the most one non-zero matrix element per row or column besides the diagonal element.



**FIG. 5.** Elementary transition matrices of  $\text{CsH}_2\text{PO}_4$ . (a) Elementary transition matrix obtained from the oxygen positions after 5 ps of the AIMD simulation. (b) Elementary transition matrix obtained from the oxygen positions after 450 ps of the AIMD simulation. (c) The same transition matrix as in (a), but using a smaller cutoff for the color scale, thus amplifying smaller transition probabilities. (d) The same transition matrix as in (b), but using a smaller cutoff for the color scale, thus amplifying smaller transition probabilities.

- Each of the diagonal elements of the elementary transition matrices is greater than zero.
- The elementary transition matrices are reducible. This property will be explained in the next paragraph.

Considering proton transfer in a hydrogen bonded system during short periods, the first four properties are confirmed by chemical intuition. While the fifth property is rather formal, it becomes apparent if we take another interpretation of a transition matrix of the protons into account: The term transition matrix within our MDM approach is a synonym for the Markov matrix. The repeated application of the Markov matrix to an initial vector constitutes a Markov chain. It is a known property that a Markov matrix is reducible if there is least one transition from a single state to another single state with probability equal to zero for an arbitrary length of the Markov chain.

A non-zero probability for a transition between two oxygen atoms corresponded to a small oxygen–oxygen distance [cf. Eq. (11)] and, hence, the formation of a hydrogen bond. Then, the reducibility of the transition matrix translates into the picture that not all oxygen atoms are connected via hydrogen bonds, which again fits to the chemical intuition for the short period of proton dynamics described by an elementary transition matrix. Vice versa, an irreducible transition matrix is equivalent to a non-zero probability for every transition from any state to any other state (even if in more than one step). This translates into the picture that all oxygen atoms are connected via a chain of hydrogen bonds (which may, of course, exist at different times). In Secs. III A and III B, we will demonstrate that the property of being reducible/irreducible is of great importance for the asymptotic convergence behavior of our dynamic model.

First, we want to focus on the properties of transition matrices, which corresponds to larger time intervals compared to the elementary time step of the MD.

According to Eq. (10), by successively multiplying the elementary transition matrices, transition matrices are obtained for the

description of larger time intervals. We want to refer to such a transition matrix as assembled transition matrix. We calculated assembled transition matrices from the  $\text{CsH}_2\text{PO}_4$  trajectory corresponding to intervals ranging from the initial time step of the MD (50 fs) to the entire length of the MD simulation (900 ps). These calculations confirm that properties 1 and 4 of the elementary transition matrices also hold for assembled transition matrices. Surprisingly, assembled transition matrices are no longer symmetric. The chemical interpretation of this observation is that transition processes, which occur on the timescale of the evolution of the heavy-atom conformation, are no longer reversible because of the corresponding change in lattice topology leading to a modified transition matrix [Eqs. (11) and (12)].

In Fig. 6, specific assembled transition matrices are shown for the interval of 25 ps and 850 ps (obtained by the multiplication of 500 and 17 000 elementary transition matrices). The transition matrix according to the time interval of 25 ps is less sparse compared to the elementary transition matrix, and the same holds for the comparison of the transition matrices according to the time interval of 25 ps and 900 ps. This is the expected behavior because more and more oxygen atoms are connected via hydrogen bonds (at different times) when considering increasing periods of the trajectory. The last property motivates the hypothesis that, for sufficient long intervals, the transition matrix may become irreducible. Therefore, we checked the assembled transition matrices for reducibility/irreducibility. For the compound  $\text{CsH}_2\text{PO}_4$ , we can state that transition matrices corresponding to a time interval of 60 ps or more are indeed irreducible. The details of determination of the reducibility/irreducibility are discussed in Sec. IV B.

The results of the evaluation of the AIMD transition matrices of  $\text{CsH}_2\text{PO}_4$  are consistent with the atomistic picture that we expect for a good proton conductor, and it is in accordance with the Grotthuss mechanism: The oxygen atoms act as hydrogen-bond donors and acceptors and are involved in one (strong) hydrogen bond at any time. As the evolution period of the system is increased, the topology/connectivity of the hydrogen bond network changes, i.e.,

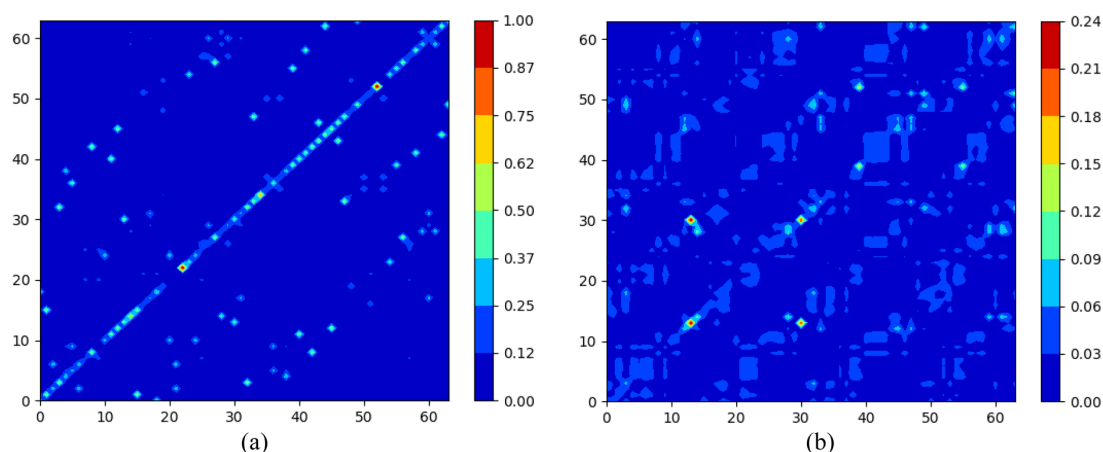


FIG. 6. Assembled transition matrices of  $\text{CsH}_2\text{PO}_4$ . (a) Assembled transition matrix for a 25 ps interval of the underlying MD trajectory. (b) Assembled transition matrix for a 850 ps interval of the underlying MD trajectory.



hydrogen bonds break and the oxygen atoms form new hydrogen bonds with (other, newly) adjacent oxygen atoms. Over a very long period of time, the hydrogen bond network evolves in such a way that all the oxygen atoms are “linked” together (via hydrogen bonds at different times).

From this extensive description of the specific proton conductor  $\text{CsH}_2\text{PO}_4$ , we want to generalize a set of properties of the transition matrices of good proton conductors:

#### F. General properties of the transition matrix

*Property 1.* The transition matrices are always stochastic matrices and have only positive entries.

The  $(ij)$ th matrix element of the transition matrix describes the probability of a proton transfer from the  $j$ th to the  $i$ th oxygen atom within the associated time interval. These probabilities are always greater than zero. Furthermore, the total number of protons must remain constant, i.e., the sum of each column has to be equal to one, so that the transformation matrix is stochastic.

*Property 2.* The transition matrices for elementary time steps are sparse and determined by the hydrogen bond network.

Proton transfer takes place only along small oxygen–oxygen distances, i.e., hydrogen bonds. Thus, the elementary transition matrices (corresponding to short periods of proton dynamics) are sparse, i.e., there are only a few non-zero entries in each row and column of the matrix. In particular, in most cases, the number of non-zero entries per column or row will be two because a hydrogen bond typically involves only two oxygen atoms (bifurcations are rare).

*Property 3.* After a sufficiently long simulation time, the hydrogen bond network links all oxygen atoms.

After a sufficiently long period of proton dynamics, the hydrogen bond network interconnects all oxygen atoms (via a series of hydrogen bonds at different times). This corresponds to the irreducibility of the transition matrix associated with sufficiently large time intervals.

If this condition is not fulfilled, the investigated compound cannot be considered as a good proton conductor in terms of the Grotthuss mechanism.

Furthermore, we demand the following:

*Property 4.* The transition matrices for elementary time steps are symmetric.

The probability of a proton transfer within an elementary time step from oxygen atom  $i$  to oxygen atom  $j$  should be the same as that of oxygen atom  $j$  to oxygen atom  $i$ .

If these four properties constitute a reasonable model of proton transfer, any compound that fulfills these properties will exhibit a high proton conductivity. For such a compound, it is immediately apparent that

*Hypothesis 5.* The protons, regardless of their initial position, should be distributed uniformly to the “chemically identical” oxygen atoms after an infinite period of evolution of the proton dynamics.

This statement is equivalent to the existence of a unique fixed point of the proton distribution (the uniform distribution) and will be formally derived in Sec. III. Chemically, it reflects the well-known

thermodynamic relationship between population distribution and chemical potential,

$$x_i = \frac{\exp(-\beta\mu_i)}{\sum_j \exp(-\beta\mu_j)}, \quad (14)$$

since, in our case, all proton sites were assumed to be of the same species, thus having equal chemical potential  $\mu_i = \mu_j \forall i, j$ .

In Sec. III, we will demonstrate that Hypothesis 5 is correct for all compounds that fulfill the first four general properties listed in this section.

### III. ANALYSIS OF THE ASYMPTOTIC BEHAVIOR OF THE DYNAMICAL MODEL

#### A. Convergence behavior and fixed points of transition matrices

Let  $\mathbf{Q} \in \mathbb{R}^{M \times M}$  be a transition matrix of a Markov Chain, i.e., all entries are non-negative and the sum of each column vector of  $\mathbf{Q}$  is equal to one ( $\mathbf{Q}$  is a stochastic matrix). The orbit or trajectory of  $\mathbf{Q}$  is the set  $\{\mathbf{Q}^v x_0 \mid x_0 \in \mathbb{R}^M, v \in \mathbb{N}\}$ . Fixed points or stationary distributions of  $\mathbf{Q}$  are those vectors  $x \in \mathbb{R}^m$  with  $\mathbf{Q}x = x$ , whose orbit is the one-element set  $\{x\}$ . This is equivalent to  $\lim_{v \rightarrow \infty} \mathbf{Q}^v x = x$ . In this section, we want to answer the following three questions:

- Under which conditions exists a unique stationary distribution or a fixed point of  $\mathbf{Q}$ ?
- Which  $x \in \mathbb{R}^M$  is the fixed point of  $\mathbf{Q}$ ?
- For which initial vectors  $x_0 \in X$  do we encounter this fixed point?

The following theorem from the literature will help us:<sup>75</sup>

**Theorem 1.** (A part of the first ergodic theorem)

Every stochastic, irreducible, and aperiodic matrix  $\mathbf{Q}$  has a unique fixed point or stationary distribution  $x$ . This fixed point is positive (i.e. all elements  $x_i \geq 0$ ), and the sequence of the distributions  $\mathbf{Q}^v x_0$  converges to  $x$  for every initial distribution  $x_0$ .

It follows from this statement that a Markov matrix  $\mathbf{Q}$  always has an eigenvalue 1. All other eigenvalues are in absolute value smaller than 1. The eigenvector  $x$  to the eigenvalue 1 is called the stationary distribution of  $\mathbf{Q}$ .

We are even able to predict the fixed point for symmetric or doubly stochastic transition matrices  $\mathbf{Q}$ .<sup>75</sup>

**Theorem 2.** The uniform distribution at  $\mathbb{R}^M$ , that is the vector  $x \in \mathbb{R}^M$  with  $x^{(0)} = x^{(1)} = \dots = x^{(M)}$ , is the stationary distribution of the stochastic transition matrix  $\mathbf{Q}$ , if the sums of the columns as well as of the rows of  $\mathbf{Q}$  are equal to one (i.e.,  $\mathbf{Q}$  is doubly stochastic).

A stochastic matrix, which is also symmetric, is always doubly stochastic.

#### B. Transfer of the results to the discrete dynamic model for proton transfer

We want to apply the theorems for a general transition matrix  $\mathbf{Q}$  to the assembled transition matrix  $\mathbf{P}_{\text{MD}}$  of the MDM approach, which describes the largest possible time interval, i.e., the assembled transition matrix that corresponds to the entire MD

length. Therefore, we have to check if the transition matrix  $\mathbf{P}_{\text{MD}}$  fulfills the prerequisites of the first theorem from Sec. III A:

First of all, the transition matrix  $\mathbf{P}_{\text{MD}}$  is stochastic and positive according to property one in Sec. II F. Second, the transition matrix  $\mathbf{P}_{\text{MD}}$  is irreducible for a sufficiently long length of the underlying MD simulation, as we discussed in detail in “property three” in Secs. II E and II F. Third, we mentioned in Sec. II E that the diagonal elements are always greater than zero. Thus, the transition matrix  $\mathbf{P}_{\text{MD}}$  is also aperiodic.

In conclusion, the transition matrix  $\mathbf{P}_{\text{MD}}$  based on sufficiently long intervals of the underlying MD simulation is stochastic, aperiodic, and irreducible, which ensures the existence of a stationary distribution, and this fixed point can be reached from every initial vector.

Unfortunately, the transition matrix  $\mathbf{P}_{\text{MD}}$  is not symmetric, and hence, it is not clear if the stationary distribution of our dynamic model is equal to the uniform distribution (second theorem of Sec. III A). However, we can show that  $\mathbf{P}_{\text{MD}}$  is doubly stochastic by recalling the construction of  $\mathbf{P}_{\text{MD}}$  by the successive multiplication of the elementary transition matrices  $\mathbf{P}^{(1,2)}, \mathbf{P}^{(2,3)}, \dots, \mathbf{P}^{(N-1,N)}$  according to Eq. (10). The elementary transition matrices are by construction stochastic and symmetric [cf. Eqs. (13) and (12)] and, thus, also doubly stochastic. The symmetry is not retained by multiplication of two symmetric matrices, but the product of two doubly stochastic matrices is again doubly stochastic. Thus,  $\mathbf{P}_{\text{MD}}$  is [as a product of doubly stochastic matrices, cf. Eq. (10)] a doubly stochastic matrix. Hence, we can also apply the second theorem from Sec. III A to the transition matrix  $\mathbf{P}_{\text{MD}}$  and the Markov chain, constructed by repeated applications of  $\mathbf{P}_{\text{MD}}$  to an initial proton distribution will always converge to the uniform distribution of the protons (to the oxygen atoms).

Following this line, the four general properties (see Sec. II F) of the transition matrix are sufficient to uniquely define the MDM framework, which possesses the correct physical asymptotic behavior for the simulation of proton transfer.

In summary, we can answer the three questions listed at the beginning of Sec. III A:

- The general properties of a transition matrix  $\mathbf{P}_{\text{MD}}$  listed in Sec. II F are sufficient to ensure the existence of unique stationary distribution of the protons.
- This stationary distribution is the uniform distribution.
- The uniform distribution of the protons is encountered for any initial proton distribution.

An aperiodic, irreducible, and double stochastic matrix (such as  $\mathbf{P}_{\text{MD}}$ ) is called ergodic. The ergodicity of the transition matrix and the ergodicity of the MD trajectory are not equivalent, but they are connected in a certain point of view. We discuss their relationship in the [supplementary material](#).

#### IV. INVESTIGATION OF THE PROTON CONDUCTION MECHANISM OF $\text{CsH}_2\text{PO}_4$ AND HPB WITH THE DYNAMIC MODEL

##### A. Insights from AIMD simulations

In addition to the extended AIMD simulations of  $\text{CsH}_2\text{PO}_4$  (128 atoms, 900 ps, 510 K), we also performed MD simulations

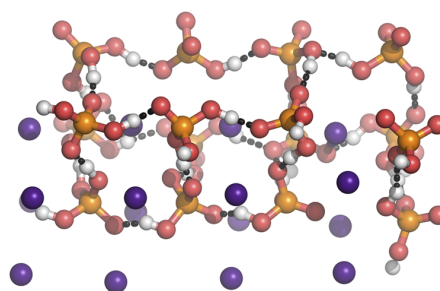


FIG. 7. Snapshot from the MD simulation of  $\text{CsH}_2\text{PO}_4$ .

of Hexa(*p*-phosphonatophenyl)benzene (HPB) (816 atoms, 115 ps at 600 K and 33 ps at 400 K). A snapshot from the simulation of  $\text{CsH}_2\text{PO}_4$  is shown in Fig. 7, and a snapshot from the simulation of Hexa(*p*-phosphonatophenyl)benzene (HPB) is shown in Fig. 8. We have previously analyzed the mechanism of proton transfer within these compounds, which turns them into ideal testcases for our dynamic model.<sup>72,76</sup>  $\text{CsH}_2\text{PO}_4$  is an inorganic salt, and HPB is a disk-shaped organic compound that possesses a supramolecular columnar structure. Despite their different chemical nature, the relevant groups for proton transfer (hydrogen phosphate anions, respectively, phosphonic acid groups) are immobile at our timescales and do not diffuse. Thus, the proton conduction mechanism is similar within both compounds and can be understood in terms of the Grotthuss mechanism.<sup>74</sup> It can be decomposed into two parts: rapid proton transfer (proton rattling) between the hydrogen phosphate anions/phosphonic groups and slow rotations of the hydrogen phosphate anions/phosphonic groups, which enables the long range proton transfer. In the [supplementary material](#), we report the jump rate functions of  $\text{CsH}_2\text{PO}_4$  and HPB, which describe the rapid proton transfer (proton rattling). The jump rate functions do not differ significantly for both compounds, and hence, the frequency of proton transfer between adjacent oxygen atoms is the same. The second part of the proton transfer mechanism (reorientation of the hydrogen phosphate anions or phosphonic groups) can be studied by the autocorrelation function of the P–O vector of the hydrogen phosphate

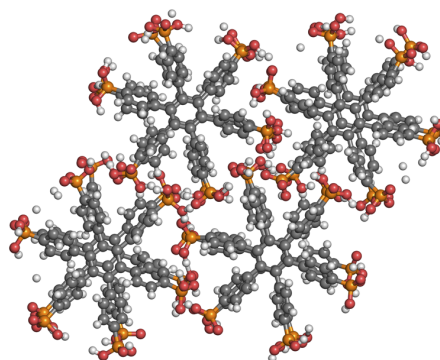


FIG. 8. Snapshot from the MD simulation of Hexa(*p*-phosphonatophenyl)benzene (HPB).

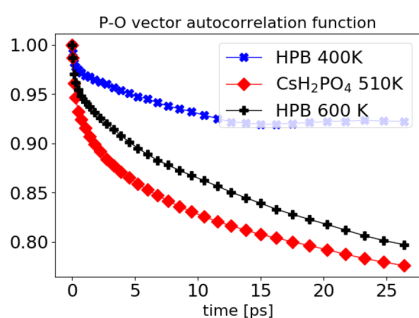


FIG. 9. Autocorrelation functions of the P–O vector.

anions or phosphonic groups. In Fig. 9, we present the autocorrelation function of these P–O vectors. For CsH<sub>2</sub>PO<sub>4</sub>, this function decays somewhat faster than for HPB at 600 K.

Nevertheless, from the analysis of the AIMD trajectory, we obtain a very similar qualitative picture for proton conduction in CsH<sub>2</sub>PO<sub>4</sub> at 510 K and HPB at 600 K. The quantitative descriptors of the contributing processes of the proton transfer slightly favor proton conduction in CsH<sub>2</sub>PO<sub>4</sub>. In order to verify our predictions, we calculated the diffusion coefficients for CsH<sub>2</sub>PO<sub>4</sub> and HPB directly from the AIMD trajectories via the slope of the mean square displacements. It turns out that both diffusion coefficients are quite large:  $D[\text{CsH}_2\text{PO}_4] = 23 \frac{\text{pm}^2}{\text{ps}}$  and  $D[\text{HPB}] = 180 \frac{\text{pm}^2}{\text{ps}}$ . The significantly higher diffusion coefficient of HPB is in contradiction to the trend expected from the slower decay of the autocorrelation function of the P–O vector of HPB compared to CsH<sub>2</sub>PO<sub>4</sub>. Thus, the simple decomposition of the proton conduction mechanism into rapid proton transfer (proton rattling) between the hydrogen phosphate anions/phosphonic groups and slow rotations of the hydrogen phosphate anions/phosphonic groups provides an useful pictorial understanding of the proton dynamics in these materials but is not able to explain differences for the diffusion coefficient. In Sec. IV B, we demonstrate that the last building block for the understanding of the proton dynamics is provided by the transition matrix obtained within the MDM model.

## B. Insights from the transition matrix

Our dynamic model can also be understood in terms of graph theory. In this case the set of oxygen atoms is interpreted as an undirected graph. Then, the (assembled) transition matrix obtained from the dynamic model corresponds to an adjacency matrix (of oxygen atoms that are connected by a possible proton transfer events). The adjacency matrix can be used to derive connected components of the graph. These connected components correspond to sets of oxygen atoms that are connected by hydrogen bonds (at different times). A criterion for a good proton conductor is that all oxygen atoms will be connected by hydrogen bonds (at different times) for a sufficiently long period of proton dynamics. The assembled transition matrix describing this period will possess only one connected component, consisting of the entire set of oxygen atoms. Thus, the following two statements concerning an adjacency matrix and an undirected graph are equivalent:

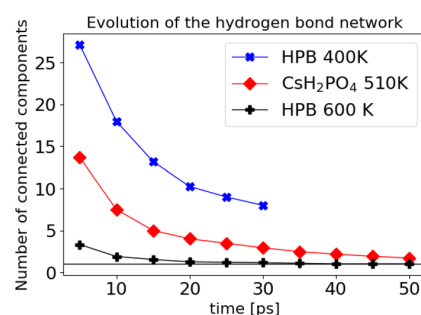


FIG. 10. Averaged number of connected components obtained from the assembled transition matrices.

- The undirected graph possesses only one connected component.
- The transition matrix is ergodic.

In Fig. 10, we show the number of connected components obtained from assembled transition matrices. The abscissa of Fig. 10 indicates the period of proton dynamics described by the assembled transition matrices. The number of connected components decays much faster for the assembled transition matrices of HPB at 600 K compared to the connected components of assembled transition matrices of CsH<sub>2</sub>PO<sub>4</sub>. After 25 ps, HPB possesses only one connected component (which corresponds to an ergodic transition matrix), whereas this is the case for CsH<sub>2</sub>PO<sub>4</sub> after 60 ps. Now, we can correlate the increased diffusion coefficient of HPB at 600 K compared to CsH<sub>2</sub>PO<sub>4</sub> with the faster evolution of the hydrogen bond network.

In addition, we also show the number of the connected components obtained from assembled transition matrices of HPB at 400 K in Fig. 10. The decay of connected components for HPB at 400 K is drastically slower compared to 600 K or the decay for CsH<sub>2</sub>PO<sub>4</sub> at 510 K. The transition matrix of HPB at 400 K does not become ergodic for the whole period of proton transfer described by the underlying AIMD simulation. Using the specific example of HPB at 400 K, we want to demonstrate that only systems with extraordinary high proton conductivities possess fast decays of the number of connected components comparable to HPB at 600 K or CsH<sub>2</sub>PO<sub>4</sub>.

Although the analysis of the connected components enabled us to explain the quantitative trend of the proton diffusion coefficient in HPB at 600 K and CsH<sub>2</sub>PO<sub>4</sub>, it does not make use of the entire information contained in the transition matrix. The connected component analysis takes only into account if a matrix element is greater than zero but does not make use of its magnitude.

Following this line, we want to extract another descriptor of the proton dynamics from the transition matrix, which takes also into account the magnitude of transferred protons.

We start from the interpretation of the  $\nu$ th power  $(\mathbf{P}_{\text{MD}})^\nu$  of the transition matrix  $\mathbf{P}_{\text{MD}}$  obtained from an AIMD trajectory with  $N$  time steps. The  $j$ -th element of the  $i$ -th column of  $(\mathbf{P}_{\text{MD}})^\nu$  gives the fraction of a proton at the  $j$ -th oxygen at the time  $\nu \cdot N \cdot \Delta t$  if a single proton was at the  $i$ -th oxygen atom at the time  $t = 0$ . Thus,  $(\mathbf{P}_{\text{MD}})^\nu$  will converge to a matrix, where every element has the same



value,

$$(\mathbf{P}_{\text{MD}})^{\nu} \xrightarrow{\nu \rightarrow \infty} \mathbf{J}; \quad \mathbf{J}_{ij} = 1/M. \quad (15)$$

We recall that this specific uniform matrix  $\mathbf{J}$  is only adequate for chemical systems in which all proton acceptor sites have identical chemical potential; otherwise,  $\mathbf{J}$  will be non-uniform.

Setting  $\|\cdot\|$  as the Frobenius norm, we obtain for a matrix  $\mathbf{A}$ ,

$$\|\mathbf{A}\| = \sqrt{\sum_{ij} (a_{ij})^2}. \quad (16)$$

Using Eq. (15), we have  $\|(\mathbf{P}_{\text{MD}})^{\nu} - \mathbf{J}\| \rightarrow 0$  for  $\nu \rightarrow \infty$ . At time  $t = 0$ , each proton is bonded to an oxygen atom, which can be expressed by  $(\mathbf{P}_{\text{MD}})^0 = \mathbf{Id}$  with  $\mathbf{Id}$  being the identity matrix. Now, we can define a descriptor  $O[\nu]$ ,

$$O[\nu] = \frac{\|(\mathbf{P}_{\text{MD}})^{\nu} - \mathbf{J}\|}{\|(\mathbf{P}_{\text{MD}})^0 - \mathbf{J}\|} = \frac{\|(\mathbf{P}_{\text{MD}})^{\nu} - \mathbf{J}\|}{\|\mathbf{Id} - \mathbf{J}\|}, \quad (17)$$

which fulfills  $O[0] = 1$  and  $\lim_{\nu \rightarrow \infty} O[\nu] = 0$ . The new descriptor  $O[\nu]$  is a measure for the relaxation time of any proton distribution to the uniform distribution. While we derived the descriptor  $O[\nu]$  via a rather formal definition, it has an important physical interpretation originated from the statistical interpretation of entropy: The application of the  $\nu$ -th power of  $\mathbf{P}_{\text{MD}}$  corresponds to mixing of “pure states” (protons remaining on a single oxygen atom). Mixing of “pure states” increases the entropy. According to the second law of thermodynamics, the entropy of a system will increase with increasing time until it has reached a maximum. This state with maximal entropy does not change any more and is the equilibrium state. The state of a maximal entropy of the proton distributions is the state of maximal mixing of the protons and corresponds to the uniform distribution of the protons among all oxygen atoms. We already demonstrated that this relaxation behavior is ensured for any initial proton distribution by the general properties of the transition matrix (cf. Sec. III B). Thus,  $O[\nu]$  describes (for any initial proton distribution) the length of the interval (relaxation time) for reaching the equilibrium regarding the proton distribution.

In Fig. 11, we show the relaxation of the proton distribution to the uniform distribution for the investigated systems. The relaxation times are at least several nanoseconds for all investigated

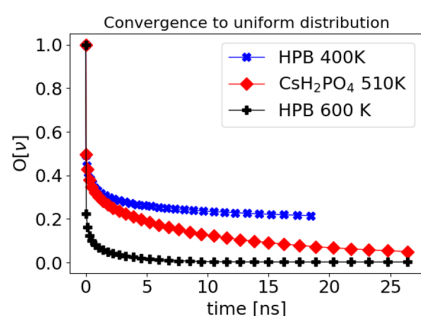


FIG. 11. Relaxation of the proton density obtained by repeated applications of the assembled transition matrix  $\mathbf{P}_{\text{MD}}$ .

compounds. It is immediately apparent that only the drastic condensation of the entire proton dynamics into an  $M \times M$  matrix  $\mathbf{P}_{\text{MD}}$  enables the simulation of such very large timescales (by simple matrix multiplications).

The decay of  $O[\nu]$  is much faster for HPB at 600 K compared to  $\text{CsH}_2\text{PO}_4$ . Thus, the relaxation of the proton distribution of HPB at 600 K is significantly faster compared to the simulation of  $\text{CsH}_2\text{PO}_4$ . We can correlate the increased diffusion coefficient of HPB at 600 K compared to  $\text{CsH}_2\text{PO}_4$  with the faster relaxation of the proton distribution, i.e., the more efficient long range proton transfer. Due to the short underlying AIMD length of the simulation of HPB at 400 K, the transition matrix is not yet ergodic and the relaxation behavior is very slow compared to  $\text{CsH}_2\text{PO}_4$  or the simulation at increased temperatures.

## V. CONCLUSION AND OUTLOOK

In this article, we have presented a new dynamic model—the molecular dynamics/matrix propagation (MDM) method—for the simulation of long range proton transfer, which condenses the entire proton dynamics into an  $M \times M$  matrix (where  $M$  is the number of the oxygen atoms of the system). We reduce the proton conducting system to the  $M$  coordinates of the oxygen atoms and describe the proton distribution by an  $M$ -dimensional state vector. This proton distribution vector is evolved by means of a transition matrix, which is constructed (with respect to different intervals of proton dynamics) from the positions of the oxygen atoms obtained from an *ab initio* MD of the proton conductor.

We calculated transition matrices for the excellent proton conductor  $\text{CsH}_2\text{PO}_4$  and verified our conceptual assumptions for the MDM model. Within a thorough mathematical discussion, we proved that our MDM model is well defined and possesses the correct asymptotic behavior for the distribution of the protons. We elucidated the relation between the ergodicity of the transition matrix and the ergodicity of the underlying MD simulation. Furthermore, we demonstrated the usage of the transition matrices in order to understand and predict quantitatively proton conduction trends in the promising proton exchange membrane materials for fuel cell reactions hexakis(*p*-phosphonatophenyl)benzene and  $\text{CsH}_2\text{PO}_4$ .

The MDM model allows the condensation of the proton dynamics of an entire MD simulation into a small matrix. Our overall attempt is to use the drastically reduced dimensionality of our approach to increase the system size and timescales of the simulation of long range proton transfer. So far, we have focused on the validation. Now, our tool can be used to compute protonation dynamics on micrometer/microsecond length- and time-scales for realistic compounds including nanoscale inhomogeneities.<sup>77</sup>

Toward a more realistic description of long range proton transfer, we are also planning to include the contribution of nuclear quantum effects via post-processing of the transition matrix in our MDM scheme.

## SUPPLEMENTARY MATERIAL

See the [supplementary material](#) for the jump rate functions of  $\text{CsH}_2\text{PO}_4$  and hexa(*p*-phosphonatophenyl) benzene, a discussion of the relationship between ergodicity of the transition matrix and the

ergodicity of the MD simulation, and an alternative scheme of the dynamic model for the propagation of the proton distribution.

## ACKNOWLEDGMENTS

This work was supported by the Deutsche Forschungsgemeinschaft (Project-ID 435886714).

## APPENDIX: COMPUTATIONAL DETAILS

We applied Born–Oppenheimer Molecular Dynamics (BOMD) using the CP2K<sup>78</sup> program package to simulate the proton transfer in a superconducting phase of the solid acid CsH<sub>2</sub>PO<sub>4</sub> and Hexa(*p*-phosphonatophenyl)benzene (HPB). We utilized the Quickstep module<sup>79</sup> and orbital transformation<sup>80</sup> for faster convergence. The electronic structure was calculated with density functional theory utilizing the PBE<sup>81–83</sup> functional for CsH<sub>2</sub>PO<sub>4</sub> and BLYP<sup>84,85</sup> for HPB. A basis set of the type DZVP-MOLOPT-SR-GTH<sup>86</sup> and GTH pseudopotentials<sup>87,88</sup> were applied. Furthermore, we used the empirical dispersion correction (D3) from Grimme.<sup>89</sup> The temperature was set by a Nosé–Hoover chain thermostat<sup>90–92</sup> (NVT ensemble). The time step was chosen as 0.5 fs. A system of 16 formula units (128 atoms) of CsH<sub>2</sub>PO<sub>4</sub> were investigated at 510 K for 900 ps. Concerning the HPB system, 816 atoms were calculated for 38 ps at 400 K and 115 ps at 600 K.

## REFERENCES

- 1 A. Katok and B. Hasselblatt, *Introduction to the Modern Theory of Dynamical Systems* (Cambridge University Press, 1997), Vol. 54.
- 2 S. H. Strogatz, *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering* (CRC Press, 2018).
- 3 See <https://www.nature.com/subjects/dynamical-systems> for a definition of a dynamical system.
- 4 L. Delle Site, “What is a multiscale problem in molecular dynamics?,” *Entropy* **16**, 23–40 (2014).
- 5 M. Praprotnik, L. Delle Site, and K. Kremer, “Multiscale simulation of soft matter: From scale bridging to adaptive resolution,” *Annu. Rev. Phys. Chem.* **59**, 545–571 (2008).
- 6 K. Wendler, F. Dommert, Y. Y. Zhao, R. Berger, C. Holm, and L. Delle Site, “Ionic liquids studied across different scales: A computational perspective,” *Faraday Discuss.* **154**, 111–132 (2012).
- 7 J. Behler, “Constructing high-dimensional neural network potentials: A tutorial review,” *Int. J. Quantum Chem.* **115**, 1032–1050 (2015).
- 8 J. Behler, “First principles neural network potentials for reactive simulations of large molecular and condensed systems,” *Angew. Chem., Int. Ed.* **56**, 12828–12840 (2017).
- 9 B. Nebgen, N. Lubbers, J. S. Smith, A. E. Sifain, A. Lokhov, O. Isayev, A. E. Roitberg, K. Barros, and S. Tretiak, “Transferable dynamic molecular charge assignment using deep neural networks,” *J. Chem. Theory Comput.* **14**, 4687–4698 (2018).
- 10 K. T. Butler, D. W. Davies, H. Cartwright, O. Isayev, and A. Walsh, “Machine learning for molecular and materials science,” *Nature* **559**, 547 (2018).
- 11 M. Hellström and J. Behler, “Concentration-dependent proton transfer mechanisms in aqueous NaOH solutions: From acceptor-driven to donor-driven and back,” *J. Phys. Chem. Lett.* **7**, 3302–3306 (2016).
- 12 M. Gastegger, C. Kauffmann, J. Behler, and P. Marquetand, “Comparing the accuracy of high-dimensional neural network potentials and the systematic molecular fragmentation method: A benchmark study for all-trans alkanes,” *J. Chem. Phys.* **144**, 194110 (2016).
- 13 V. Quaranta, M. Hellström, and J. Behler, “Proton-transfer mechanisms at the water–ZnO interface: The role of presolvation,” *J. Phys. Chem. Lett.* **8**, 1476–1483 (2017).
- 14 F. Noé, J. Chodera, G. Bowman, V. Pande, and F. Noé, *An Introduction to Markov State Models and Their Application to Long Timescale Molecular Simulation*, Advances in Experimental Medicine and Biology Vol. 797 (Springer, 2014).
- 15 G. R. Bowman, E. R. Bolin, K. M. Hart, B. C. Maguire, and S. Marqusee, “Discovery of multiple hidden allosteric sites by combining Markov state models and experiments,” *Proc. Natl. Acad. Sci. U. S. A.* **112**(9), 2734–2739 (2015).
- 16 B. E. Husic and V. S. Pande, “Markov state models: From an art to a science,” *J. Am. Chem. Soc.* **140**, 2386–2396 (2018).
- 17 V. S. Pande, K. Beauchamp, and G. R. Bowman, “Everything you wanted to know about Markov State Models but were afraid to ask,” *Methods* **52**, 99–105 (2010), Protein Folding.
- 18 C. Schütte and M. Sarich, “A critical appraisal of Markov state models,” *Eur. Phys. J. Spec. Top.* **224**, 2445–2462 (2015).
- 19 C. R. Schwantes, R. T. McGibbon, and V. S. Pande, “Perspective: Markov models for long-timescale biomolecular dynamics,” *J. Chem. Phys.* **141**, 090901 (2014).
- 20 L. J. LaBerge and J. C. Tully, “A rigorous procedure for combining molecular dynamics and Monte Carlo simulation algorithms,” *Chem. Phys.* **260**, 183–191 (2000).
- 21 E. C. Neyts and A. Bogaerts, “Combining molecular dynamics with Monte Carlo simulations: Implementations and applications,” *Theor. Chem. Acc.* **132**, 1320 (2012).
- 22 B. M. Forrest and U. W. Suter, “Hybrid Monte Carlo simulations of dense polymer systems,” *J. Chem. Phys.* **101**, 2616–2629 (1994).
- 23 D. G. Gromov and J. J. de Pablo, “Structure of binary polymer blends: Multiple time step hybrid Monte Carlo simulations and self-consistent integral-equation theory,” *J. Chem. Phys.* **103**, 8247–8256 (1995).
- 24 A. Irbäck, “Hybrid Monte Carlo simulation of polymer chains,” *J. Chem. Phys.* **101**, 1661–1667 (1994).
- 25 D. W. Heermann and L. Yixue, “A global-update simulation method for polymer systems,” *Macromol. Chem. Phys.* **2**, 299–308 (1993).
- 26 I. Martín-Bragado, R. Borges, J. P. Balbuena, and M. Jaraiz, “Kinetic Monte Carlo simulation for semiconductor processing: A review,” *Prog. Mater. Sci.* **92**, 1–32 (2018).
- 27 G. Betz and W. Husinsky, “A combined molecular dynamics and kinetic Monte Carlo calculation to study sputter erosion and beam assisted deposition,” *Nucl. Instrum. Methods Phys. Res., Sect. B* **193**, 352–358 (2002).
- 28 A. Ghoufi and G. Maurin, “Hybrid Monte Carlo simulations combined with a phase mixture model to predict the structural transitions of a porous metal-organic framework material upon adsorption of guest molecules,” *J. Phys. Chem. C* **114**, 6496–6502 (2010).
- 29 A. A. Kizhnik, A. A. Bagaturyants, I. V. Belov, B. V. Potapkin, and A. A. Korkin, “An integrated kinetic Monte Carlo molecular dynamics approach for film growth modeling and simulation: ZrO<sub>2</sub> deposition on Si(100) surface,” *Comput. Mater. Sci.* **24**, 128–132 (2002).
- 30 U. H. E. Hansmann and Y. Okamoto, “New Monte Carlo algorithms for protein folding,” *Curr. Opin. Struct. Biol.* **9**, 177–183 (1999).
- 31 E. K. Peter and J.-E. Shea, “A hybrid MD-kMC algorithm for folding proteins in explicit solvent,” *Phys. Chem. Chem. Phys.* **16**, 6430–6440 (2014).
- 32 H. Zhang, “A new hybrid Monte Carlo algorithm for protein potential function test and structure refinement,” *Proteins: Struct., Funct., Genet.* **34**, 464–471 (1999).
- 33 F. Noé and E. Rosta, “Markov models of molecular kinetics,” *J. Chem. Phys.* **151**, 190401 (2019).
- 34 M. Dibak, M. J. del Razo, D. De Sancho, C. Schütte, and F. Noé, “MSM/RD: Coupling Markov state models of molecular kinetics with reaction-diffusion simulations,” *J. Chem. Phys.* **148**, 214107 (2018).
- 35 B. G. Keller, J.-H. Prinz, and F. Noé, “Markov models and dynamical fingerprints: Unraveling the complexity of molecular kinetics,” *Chem. Phys.* **396**, 92–107 (2012), part of special issue: Experimental and theoretical studies of protein dynamics and function: From femtoseconds to milliseconds.
- 36 B. G. Keller, A. Kobitski, A. Jäschke, G. U. Nienhaus, and F. Noé, “Complex RNA folding kinetics revealed by single-molecule FRET and hidden Markov models,” *J. Am. Chem. Soc.* **136**, 4534–4543 (2014).

- <sup>37</sup>F. Nüske, B. G. Keller, G. Pérez-Hernández, A. S. J. S. Mey, and F. Noé, "Variational approach to molecular kinetics," *J. Chem. Theory Comput.* **10**, 1739–1752 (2014).
- <sup>38</sup>F. Nüske, H. Wu, J.-H. Prinz, C. Wehmeyer, C. Clementi, and F. Noé, "Markov state models from short non-equilibrium simulations—Analysis and correction of estimation bias," *J. Chem. Phys.* **146**, 094104 (2017).
- <sup>39</sup>J.-H. Prinz, H. Wu, M. Sarich, B. Keller, M. Senne, M. Held, J. D. Chodera, C. Schütte, and F. Noé, "Markov models of molecular kinetics: Generation and validation," *J. Chem. Phys.* **134**, 174105 (2011).
- <sup>40</sup>C. Schütte, F. Noé, J. Lu, M. Sarich, and E. Vanden-Eijnden, "Markov state models based on milestoning," *J. Chem. Phys.* **134**, 204105 (2011).
- <sup>41</sup>B. Trendelkamp-Schroer, H. Wu, F. Paul, and F. Noé, "Estimation and uncertainty of reversible Markov models," *J. Chem. Phys.* **143**, 174101 (2015).
- <sup>42</sup>F. Vitalini, A. S. J. S. Mey, F. Noé, and B. G. Keller, "Dynamic properties of force fields," *J. Chem. Phys.* **142**, 084101 (2015).
- <sup>43</sup>H. Wu, F. Paul, C. Wehmeyer, and F. Noé, "Multiensemble Markov models of molecular thermodynamics and kinetics," *Proc. Natl. Acad. Sci. U. S. A.* **113**(23), E3221–E3230 (2016).
- <sup>44</sup>E. Hruska, J. R. Abella, F. Nüske, L. E. Kavrakci, and C. Clementi, "Quantitative comparison of adaptive sampling methods for protein dynamics," *J. Chem. Phys.* **149**, 244119 (2018).
- <sup>45</sup>U. Sengupta, M. Carballo-Pacheco, and B. Strodel, "Automated Markov state models for molecular dynamics simulations of aggregation and self-assembly," *J. Chem. Phys.* **150**, 115101 (2019).
- <sup>46</sup>F. Noé, C. Schütte, E. Vanden-Eijnden, L. Reich, and T. R. Weikel, "Constructing the equilibrium ensemble of folding pathways from short off-equilibrium simulations," *Proc. Natl. Acad. Sci. U. S. A.* **106**, 19011–19016 (2009).
- <sup>47</sup>T. J. Lane, G. R. Bowman, K. Beauchamp, V. A. Voelz, and V. S. Pande, "Markov state model reveals folding and functional dynamics in ultra-long MD trajectories," *J. Am. Chem. Soc.* **133**, 18413–18419 (2011).
- <sup>48</sup>G. Pérez-Hernández, F. Paul, T. Giorgino, G. De Fabritiis, and F. Noé, "Identification of slow molecular order parameters for Markov model construction," *J. Chem. Phys.* **139**, 015102 (2013).
- <sup>49</sup>N. Stanley, S. Esteban-Martín, and G. De Fabritiis, "Kinetic modulation of a disordered protein domain by phosphorylation," *Nat. Commun.* **5**, 5272 (2014).
- <sup>50</sup>I. Buch, T. Giorgino, and G. De Fabritiis, "Complete reconstruction of an enzyme-inhibitor binding process by molecular dynamics simulations," *Proc. Natl. Acad. Sci. U. S. A.* **108**, 10184–10189 (2011).
- <sup>51</sup>M. Held, P. Metzner, J.-H. Prinz, and F. Noé, "Mechanisms of protein-ligand association and its modulation by protein mutations," *Biophys. J.* **100**, 701–710 (2011).
- <sup>52</sup>D.-A. Silva, G. R. Bowman, A. Sosa-Peinado, and X. Huang, "A role for both conformational selection and induced fit in ligand binding by the LAO protein," *PLoS Comput. Biol.* **7**, 1–11 (2011).
- <sup>53</sup>G. R. Bowman and P. L. Geissler, "Equilibrium fluctuations of a single folded protein reveal a multitude of potential cryptic allosteric sites," *Proc. Natl. Acad. Sci. U. S. A.* **109**, 11681–11686 (2012).
- <sup>54</sup>N. Plattner and F. Noé, "Protein conformational plasticity and complex ligand-binding kinetics explored by atomistic simulations and Markov models," *Nat. Commun.* **6**, 7653 (2015).
- <sup>55</sup>D. Chakraborty and D. J. Wales, "Dynamics of an adenine-adenine RNA conformational switch from discrete path sampling," *J. Chem. Phys.* **150**, 125101 (2019).
- <sup>56</sup>G. Pinamonti, F. Paul, F. Noé, A. Rodriguez, and G. Bussi, "The mechanism of RNA base fraying: Molecular dynamics simulations analyzed with core-set Markov state models," *J. Chem. Phys.* **150**, 154123 (2019).
- <sup>57</sup>A. M. Berezhkovskii and A. Szabo, "Committers, first-passage times, fluxes, Markov states, milestones, and all that," *J. Chem. Phys.* **150**, 054106 (2019).
- <sup>58</sup>J. A. Morrone, K. E. Haslinger, and M. E. Tuckerman, "Ab initio molecular dynamics simulation of the structure and proton transport dynamics of methanol–water solutions," *J. Phys. Chem. B* **110**, 3712–3720 (2006).
- <sup>59</sup>T. C. Berkelbach, H.-S. Lee, and M. E. Tuckerman, "Concerted hydrogen-bond dynamics in the transport mechanism of the hydrated proton: A first-principles molecular dynamics study," *Phys. Rev. Lett.* **103**, 238302 (2009).
- <sup>60</sup>L. Vilčiauskas, M. E. Tuckerman, G. Bester, S. J. Paddison, and K.-D. Kreuer, "The mechanism of proton conduction in phosphoric acid," *Nat. Chem.* **4**, 461 (2012).
- <sup>61</sup>C. Drefßler, G. Kabbe, and D. Sebastiani, "Insight from atomistic simulations of protonation dynamics at the nanoscale," *Fuel Cells* **16**, 682–694 (2016).
- <sup>62</sup>R. Vuilleumier and D. Borgis, "Proton conduction: Hopping along hydrogen bonds," *Nat. Chem.* **4**, 432 (2012).
- <sup>63</sup>D. Marx, A. Chandra, and M. E. Tuckerman, "Aqueous basic solutions: Hydroxide solvation, structural diffusion, and comparison to the hydrated proton," *Chem. Rev.* **110**, 2174–2216 (2010).
- <sup>64</sup>D. Marx, "Proton transfer 200 years after von Grothuss: Insights from ab initio simulations," *ChemPhysChem* **7**, 1848–1870 (2006).
- <sup>65</sup>M. E. Tuckerman, A. Chandra, and D. Marx, "A statistical mechanical theory of proton transport kinetics in hydrogen-bonded networks based on population correlation functions with applications to acids and bases," *J. Chem. Phys.* **133**, 124108 (2010).
- <sup>66</sup>A. Chandra, M. E. Tuckerman, and D. Marx, "Connecting solvation shell structure to proton transport kinetics in hydrogen-bonded networks via population correlation functions," *Phys. Rev. Lett.* **99**, 145901 (2007).
- <sup>67</sup>J. Schmidt, J. VandeVondele, I.-F. W. Kuo, D. Sebastiani, J. I. Siepmann, J. Hutter, and C. J. Mundy, "Isobaric-isothermal molecular dynamics simulations utilizing density functional theory: An assessment of the structure and density of water at near-ambient conditions," *J. Phys. Chem. B* **113**, 11959–11964 (2009).
- <sup>68</sup>J. C. Grossman, E. Schwegler, E. W. Draeger, F. Gygi, and G. Galli, "Towards an assessment of the accuracy of density functional theory for first principles simulations of water," *J. Chem. Phys.* **120**, 300–311 (2004).
- <sup>69</sup>E. Schwegler, J. C. Grossman, F. Gygi, and G. Galli, "Towards an assessment of the accuracy of density functional theory for first principles simulations of water. II," *J. Chem. Phys.* **121**, 5400–5409 (2004).
- <sup>70</sup>G. Kabbe, C. Wehmeyer, and D. Sebastiani, "A coupled molecular dynamics/kinetic Monte Carlo approach for protonation dynamics in extended systems," *J. Chem. Theory Comput.* **10**, 4221 (2014).
- <sup>71</sup>G. Kabbe, C. Drefßler, and D. Sebastiani, "Toward realistic transfer rates within the coupled molecular dynamics/lattice Monte Carlo approach," *J. Phys. Chem. C* **120**, 19905–19912 (2016).
- <sup>72</sup>C. Drefßler, G. Kabbe, and D. Sebastiani, "Proton conductivity in hydroxide phosphate/sulfates from a coupled molecular dynamics/lattice Monte Carlo (cMD/LMC) approach," *J. Phys. Chem. C* **120**, 19913–19922 (2016).
- <sup>73</sup>G. Kabbe, C. Drefßler, and D. Sebastiani, "Proton mobility in aqueous systems: Combining ab initio accuracy with millisecond timescales," *Phys. Chem. Chem. Phys.* **19**, 28604–28609 (2017).
- <sup>74</sup>H.-S. Lee and M. E. Tuckerman, "The structure and proton transport mechanisms in the superprotonic phase of CsH<sub>2</sub>PO<sub>4</sub>: An ab initio molecular dynamics study," *J. Phys. Chem. C* **112**, 9917–9930 (2008).
- <sup>75</sup>T. Müller-Gronbach, E. Novak, and K. Ritter, *Monte Carlo-Algorithmen* (Springer-Verlag, 2012).
- <sup>76</sup>C. Wehmeyer, M. Schrader, D. Andrienko, and D. Sebastiani, "Water-free proton conduction in hexakis(*p*-phosphonatophenyl)benzene nanochannels," *J. Phys. Chem. C* **117**, 12366–12372 (2013).
- <sup>77</sup>M. N. Garaga, V. Dracopoulos, U. Werner-Zwanziger, J. W. Zwanziger, M. Maréchal, M. Persson, L. Nordstierna, and A. Martinelli, "A long-chain protic ionic liquid inside silica nanopores: Enhanced proton mobility due to efficient self-assembly and decoupled proton transport," *Nanoscale* **10**, 12337–12348 (2018).
- <sup>78</sup>J. Hutter, M. Iannuzzi, F. Schiffmann, and J. VandeVondele, "cp2k: atomistic simulations of condensed matter systems," *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **4**, 15–25 (2014).
- <sup>79</sup>J. VandeVondele, M. Krack, F. Mohamed, M. Parrinello, T. Chassaing, and J. Hutter, "Quickstep: Fast and accurate density functional calculations using a mixed Gaussian and plane waves approach," *Comput. Phys. Commun.* **167**, 103–128 (2005).
- <sup>80</sup>J. VandeVondele and J. Hutter, "An efficient orbital transformation method for electronic structure calculations," *J. Chem. Phys.* **118**, 4365–4369 (2003).

- <sup>81</sup>Y. Zhang and W. Yang, "Comment on "Generalized gradient approximation made simple",  
*Phys. Rev. Lett.* **80**, 890 (1998).
- <sup>82</sup>J. P. Perdew, A. Ruzsinszky, G. I. Csonka, O. A. Vydrov, G. E. Scuseria, L. A. Constantin, X. Zhou, and K. Burke, "Restoring the density-gradient expansion for exchange in solids and surfaces," *Phys. Rev. Lett.* **100**, 136406 (2008).
- <sup>83</sup>J. P. Perdew, K. Burke, and M. Ernzerhof, "Generalized gradient approximation made simple," *Phys. Rev. Lett.* **77**, 3865–3868 (1996).
- <sup>84</sup>C. Lee, W. Yang, and R. G. Parr, "Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density," *Phys. Rev. B* **37**, 785–789 (1988).
- <sup>85</sup>A. D. Becke, "Density-functional exchange-energy approximation with correct asymptotic behavior," *Phys. Rev. A* **38**, 3098–3100 (1988).
- <sup>86</sup>J. VandeVondele and J. Hutter, "Gaussian basis sets for accurate calculations on molecular systems in gas and condensed phases," *J. Chem. Phys.* **127**, 114105 (2007).
- <sup>87</sup>C. Hartwigsen, S. Goedecker, and J. Hutter, "Relativistic separable dual-space Gaussian pseudopotentials from H to Rn," *Phys. Rev. B* **58**, 3641–3662 (1998).
- <sup>88</sup>M. Krack, "Pseudopotentials for H to Kr optimized for gradient-corrected exchange-correlation functionals," *Theor. Chem. Acc.* **114**, 145–152 (2005).
- <sup>89</sup>S. Grimme, J. Antony, S. Ehrlich, and H. Krieg, "A consistent and accurate *ab initio* parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu," *J. Chem. Phys.* **132**, 154104 (2010).
- <sup>90</sup>S. Nosé, "A unified formulation of the constant temperature molecular dynamics methods," *J. Chem. Phys.* **81**, 511–519 (1984).
- <sup>91</sup>G. J. Martyna, M. L. Klein, and M. Tuckerman, "Nosé-Hoover chains: The canonical ensemble via continuous dynamics," *J. Chem. Phys.* **97**, 2635–2643 (1992).
- <sup>92</sup>S. Nosé, "A molecular dynamics method for simulations in the canonical ensemble," *Mol. Phys.* **52**, 255–268 (1970).

### 3.1.3 Article III: Mechanism of Ion Conductivity Through Polymer-stabilized $\text{CsH}_2\text{PO}_4$ Nanoparticulate Layers from Experiment and Theory.

Maximilian Wagner, Christian Dreßler, Felix P. Lohmann-Richters, Kevin Hanus, Daniel Sebastiani, Aron Varga, and Bernd Abel.

Mechanism of ion conductivity through polymer-stabilized  $\text{CsH}_2\text{PO}_4$  nanoparticulate layers from experiment and theory.

*Journal of Materials Chemistry A* , 7:27367–27376, **2019**.

In this article, I derived the theory and performed all numerical calculations. D. Sebastiani supervised the project and provided me with valuable advice. M. Wagner, L. Richter, K. Hanus, A. Varga and B. Abel conducted and analyzed all experiments. I wrote the section of the manuscript concerning the theoretical calculations. The main section of the article was written by M. Wagner.

Cite this: *J. Mater. Chem. A*, 2019, 7, 27367

## Mechanism of ion conductivity through polymer-stabilized $\text{CsH}_2\text{PO}_4$ nanoparticulate layers from experiment and theory†

Maximilian Wagner,<sup>a</sup> Christian Dreßler,<sup>b</sup> Felix P. Lohmann-Richters,<sup>id</sup><sup>ac</sup> Kevin Hanus,<sup>a</sup> Daniel Sebastiani,<sup>id</sup><sup>b</sup> Aron Varga<sup>a</sup> and Bernd Abel<sup>id</sup><sup>\*a</sup>

Electrodes are currently the primary performance-limiting component in low and intermediate temperature fuel cells. A proven method for improving electrode performance in solid acid fuel cells is to create ever finer nanostructures and thus increase the electrochemically-active surface area. However, this performance enhancement is limited by issues of long-term stability, as well as increasing both the electronic and ionic conduction pathways. Here, we combine a systematic experimental study with a computational model to quantify the effect of (1) the stabilizing polymer polyvinylpyrrolidone as well as (2) the porosity and electrode layer thickness on the average ionic conductivity of the solid acid electrolyte  $\text{CsH}_2\text{PO}_4$  in a composite solid acid fuel cell electrode. With a multiscale simulation approach using a combined molecular dynamics and lattice Monte Carlo method, proton conduction through a porous electrode is simulated at mesoscopic timescales while retaining near-atomistic structured evolution. Electrochemical impedance spectroscopy is used to evaluate the porous electrodes obtained via spray drying. Both approaches reveal a similar and significant contribution of the porous electrolyte layer to the overall cell resistance. This indicates that geometrical parameters, as well as stabilizing materials may play an essential role when designing a high-performance solid acid fuel cell.

Received 24th April 2019  
Accepted 13th November 2019

DOI: 10.1039/c9ta04275j

rsc.li/materials-a

### Introduction

As environmental awareness increases, fuel cells are attracting growing interest due to their high energy conversion efficiency. Fuel cells based on solid acid proton conductors, known as solid acid fuel cells (SAFCs), are a relatively new technology that carry several fundamental advantages.<sup>1</sup> An intermediate operating temperature of 513 K, at which the electrolyte  $\text{CsH}_2\text{PO}_4$  exists in its super protonic phase, allows for fuel flexibility and excellent resistance to catalyst poisoning.<sup>2</sup> To improve the performance of such devices, it has been shown that increasing the density of catalytically-active sites by reducing the size of both the catalyst and electrolyte particles leads to a significant decrease in electrode impedance.<sup>3–6</sup> The impedance spectroscopy of symmetrical electrochemical cells in a symmetric gas configuration of humidified hydrogen atmosphere is often employed to analyze anodic processes.<sup>3,7–9</sup> However, since the

oxygen reduction reaction is the rate limiting step, using the results from such a symmetrical measurement to predict performance under fuel cell conditions should be treated carefully. To compare the key characteristics of fuel cells, the United States Department of Energy (DOE) regularly publishes guidelines for direct hydrogen fuel cells.<sup>10</sup> Platinum group metals (PGMs) are an important cost driver and the DOE goal for 2020 is a PGM utilization of 0.125 g kW<sup>-1</sup>. Since Haile *et al.* introduced SAFCs in 2001 as promising intermediate temperature models, PGM utilization in a sintered powder electrode improved from 37 g kW<sup>-1</sup> to 3.2 g kW<sup>-1</sup> in 2011 by reducing the platinum loading and improving its spatial distribution using metal-organic chemical vapor deposition (MOCVD).<sup>5,11,12</sup> Fairly recently, further improvement, to 2.5 g kW<sup>-1</sup>, was accomplished by using carbon nanotubes as the support material.<sup>13–15</sup> These advances are attributed to the increasing fraction of electrocatalytically-active Pt particles through nanostructuring, showing an impressive but still insufficient improvement to meet the 2020 DOE goal. So far, research has mainly focused on enhancing the activity and interconnectivity of the catalyst, overlooking the resistance of the electrolyte network within a porous powder electrode. Which, however, may be limiting the performance, independent of the catalyst utilization. In this combined computational and experimental study, we explore the influence of porous electrodes on cell performance. Therefore, we synthesize three-dimensional, nanostructured, porous

<sup>a</sup>Leibniz Institute for Surface Engineering, Department Functional Surfaces, Permoserstraße 15, 04318 Leipzig, Germany

<sup>b</sup>Martin-Luther-Universität Halle-Wittenberg, Chemical Department, von-Danckelmann-Platz 4, 06120 Halle, Germany

<sup>c</sup>Forschungszentrum Jülich GmbH, Institute of Energy and Climate Research, IEK-3, Electrochemical Process Engineering, 52425, Jülich, Germany. E-mail: bernd.abel@iom-leipzig.de; Fax: +49 341 235 2584; Tel: +49 341 235 2373

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c9ta04275j



electrodes with varying thicknesses using a combination of spray drying and sputtering techniques to determine the impact of a porous electrolyte structure on cell performance. The influence of the stabilizing polymer additive on the average ionic resistivity is then evaluated. AC impedance measurements enable us to observe the impact of a  $\text{CsH}_2\text{PO}_4$  network on electrode impedance. Measurements in a symmetric hydrogen atmosphere and under fuel cell conditions have been conducted and the results compared. Based on an analytical model developed through scanning microscope imaging, we adopted a combined molecular dynamics and lattice Monte Carlo approach to the specific morphological situation.

## Experimental

### Materials

The solid acid electrolyte  $\text{CsH}_2\text{PO}_4$  was synthesized as previously described.<sup>16</sup> In brief, 25 g of  $\text{H}_3\text{PO}_4$  (Roth, Stk# 6366.1, 85 wt%) was diluted with 80 ml of methanol (Alfa Aesar, Stk# E30Z829, 99.9%). A stoichiometric quantity of  $\text{Cs}_2\text{CO}_3$  (Alfa Aesar, Stk# J02Z042, 99.9%) was then dissolved in 300 ml of methanol and quickly added to the  $\text{H}_3\text{PO}_4$ -methanol mixture. Rapid precipitation occurred and the suspension was stirred for 20 minutes until the reaction was complete. The reaction product,  $\text{CsH}_2\text{PO}_4$ , was filtered through a glass frit (pore size: 0.45  $\mu\text{m}$ ) and washed three times with 100 ml of methanol to remove residual water and reactants. The  $\text{CsH}_2\text{PO}_4$  powder was then dried overnight at 353 K. The crystal structure of the product was subsequently confirmed by X-ray powder diffraction measurements.

### Cell and electrode fabrication

Solid, dense electrolyte disks (20 mm in diameter, 1 mm in thickness) were pressed from the as-synthesized  $\text{CsH}_2\text{PO}_4$  powder with an automatic, uniaxial hydraulic press (GS25440-AtlasTM Automatic 15T) at 250 MPa in a 'maintain load' setting for 5 minutes. DC magnetron sputtering (Edwards Auto 306) was used to deposit 10, 20 and 30 nm Pt films on polished electrolyte disks. To synthesize the porous nanostructured electrodes, a spray dryer (Büchi B90) was used. During the spray drying process, a previously optimized precursor solution, consisting of  $\text{CsH}_2\text{PO}_4$ , methanol (Alfa Aesar, E30Z829, 99.9%), DI-water and polyvinylpyrrolidone (PVP, Alfa Aesar, F06Z045, 1 300 000  $\text{g mol}^{-1}$ ), (ESI, S1†), was vaporized by a piezoelectric spray head. The aerosol was transported by a hot gas stream (mixture of  $\text{N}_2$  and  $\text{CO}_2$ ) towards a deposition area for electrophoretic deposition. During transport, the solvent evaporates and the solid acid electrolyte material precipitates along with the PVP, forming dry, stabilized nanoparticles. Before entering the deposition area, the gas stream is concentrated by a funnel for an increased deposition rate on the electrolyte disk. A custom-made pellet holder allows the deposition of nanoparticles on both sides of the electrolyte without modifying the other. The deposition time was varied between 5 and 120 min in order to achieve different porous layer thicknesses. DC magnetron sputtering was used to deposit a 30 nm Pt thin film

on the porous electrolyte nanostructures, creating a symmetrical electrochemical cell (Pt-thin film +  $\text{CsH}_2\text{PO}_4^{\text{porous}}|\text{CsH}_2\text{PO}_4|\text{CsH}_2\text{PO}_4^{\text{porous}}$  + Pt thin-film). The thickness  $d$  of the porous electrolyte layers was varied between 0  $\mu\text{m}$  and 18  $\mu\text{m}$  for characterization in a symmetric cell measurement using anodic conditions and between 5  $\mu\text{m}$  and 6  $\mu\text{m}$  for characterization under fuel cell conditions. In both cases, the platinum group metal (PGM) content was calculated using the nominal Pt layer thickness set during sputter deposition to be 0.064  $\text{mg cm}^{-2}$  for each electrode. To determine the layer thickness and area coverage of the porous electrolyte layer for each cell, four cross section as well as top view SEM pictures were taken from different points: three relatively close to the edge and one from the middle of the cell. The area coverage was checked by SEM top view pictures prior to the cell splitting. The standard deviation given in Fig. 6 is based on these measurements. To measure the effect of PVP on the average ionic conductivity of the spray-dried  $\text{CsH}_2\text{PO}_4$ -PVP structure, a composite pellet was fabricated in two steps. First, a 2 cm-diameter, 500  $\mu\text{m}$ -thick pure  $\text{CsH}_2\text{PO}_4$  pellet was pressed as previously described. Then, a spray-dried  $\text{CsH}_2\text{PO}_4$  + PVP powder, obtained from the deposition area, was pressed on top, forming an additional 100  $\mu\text{m}$ -thick uniform layer. Neat carbon paper (Toray Inc. TGP-H-60) was used as the electrode on both sides, creating a cell for impedance measurements (carbon paper| $\text{CsH}_2\text{PO}_4$ | $\text{CsH}_2\text{PO}_4$  + PVP|carbon paper).

### Physical and electrochemical characterization

The electrode morphology was characterized using scanning electron microscopy (Ultra 55 SEM, Carl Zeiss Ltd.). Thermogravimetric measurements (Pyris 1-TGA, PerkinElmer) were performed on mixtures of  $\text{CsH}_2\text{PO}_4$ , PVP and Pt nanoparticles (Sigma Aldrich, MKBL4444, heat rate: 10  $\text{K min}^{-1}$ ) in a nitrogen atmosphere as well as in air. An overview of samples and their composition can be found in the ESI, S2.† Electrochemical measurements were recorded with a potentiostat (Biologic VSP 300). Biologic EC-Lab software V10.33 was used for the data analysis. All measurements were carried out at the solid acid fuel cell operating temperature of 513 K. Symmetrical impedance measurements were conducted in a humidified hydrogen atmosphere (flowrate of 75 sccm, dew point of 353 K) and 10 mV perturbation amplitude over a frequency range of 1 MHz to 100 mHz. The cells (diameter of 20 mm) were placed between two carbon paper spacer rings (Toray Inc. TGP-H-60) and stainless steel gas diffusion layers (PACOPOR ST 60 AL3, PACO Paul GmbH und Co. KG). Dry argon gas (flowrate of 100 sccm) was supplied whenever the temperature was below 423 K to avoid water condensation. Prior to the measurements a 12 h stabilization period was waited after reaching operating temperature. For simplicity and ease of comparison, the electrolyte resistance, observed as a shift of the impedance arc along the real axis in a Nyquist plot, was subtracted in all impedance measurements. The width of the impedance arcs was taken as the electrode resistance. As the external resistance of 0.025  $\Omega \text{ cm}^2$  can be neglected, the electrolyte resistance is obtained from the high frequency intercept of the real axis. Fuel cell

measurements were performed following the same heat up procedure with humidified hydrogen (flowrate of 75 sccm, dew point of 353 K) at the anode and humidified oxygen (flowrate of 75 sccm, dew point of 353 K) at the cathode. The cells (diameter = 20 mm) were placed between two carbon paper discs (Toray Inc. TGP-H-60) and stainless steel gas diffusion layers (PACOPOR ST 60 AL3, PACO Paul GmbH und Co. KG).

Each cell was investigated following a repeated testing protocol of three steps. (1) a 15 s relaxation period to the open circuit voltage (2) measuring a potentiostatic electrochemical impedance spectra (PEIS) in a frequency range of 1 MHz to 100 MHz at an open circuit voltage (OCV) with a single sine perturbation amplitude of 10 mV; and (3) scanning the working electrode potential with a scan speed of 10 mV s<sup>-1</sup> from OCV to a cell potential of 0 V to record a polarization curve. To obtain iR-free polarization curves, the voltage was corrected for ohmic resistance as determined from the high frequency intercept in the impedance spectra.

### Computational modeling – the cMD/LMC approach

**Kinetic rate model for proton conduction in nano-porous systems.** Our description of proton conduction is based on atomistic structural details, covering mesoscopic length and time scales. We selected a recently developed simulation technique, a combined molecular dynamics and lattice Monte Carlo approach cMD/LMC, which is described in detail by Kabbe *et al.*<sup>17,18</sup> With this technique, the molecular dynamics trajectory provides an accurate atomistic picture and its temporal evolution, while the Monte Carlo section models the long-range motion of the acidic protons. The suitability of the approach for the modeling of proton transfer has already been demonstrated for the bulk phase of CsH<sub>2</sub>PO<sub>4</sub>.<sup>19,20</sup> Here, we use the cMD/LMC approach to calculate the average proton resistivity of a nano-porous composite CsH<sub>2</sub>PO<sub>4</sub> system and compare it to experimental data. For the validation, the model system's size has a dimension of 1 × 1 × 100 nm. To simulate the actual nanostructured network, a system size of 8 × 8 × 20 nm was used. The total system was represented by 1 × 1 × 2 nm CsH<sub>2</sub>PO<sub>4</sub> boxes, each containing 16 formula units of CsH<sub>2</sub>PO<sub>4</sub>. The boxes could be removed separately to achieve different porosities. Within each box, proton dynamics are simulated in the atomistic model based on the existing cMD/LMC scheme, while the proton flux in the total system is constructed by coupling the proton rates of the adjacent box surfaces. In the *x*- and *y*-dimensions, we used periodic boundary conditions. Summation of all proton fluxes between the boxes leads to the total resistivity of the system. The approach is first reported by Drefßler *et al.* and described in detail.<sup>20</sup> To provide an accurate atomistic structure and its temporal evolution, an *ab initio*-molecular-dynamics simulation (AIMD) is necessary for the application of the cMD/LMC model. We applied Born–Oppenheimer Molecular Dynamics (BOMD), using the CP2K<sup>21</sup> program package, to simulate the proton transfer in CsH<sub>2</sub>PO<sub>4</sub>. The electronic structure was calculated with density functional theory utilizing the PBE functional.<sup>22–24</sup> A basic set of the type DZVP-MOLOPT-SR-GTH45 and GTH pseudo potentials was

applied.<sup>25–27</sup> Furthermore, we used the empirical dispersion correction (D3) from Grimme *et al.*<sup>28</sup> The temperature was set by a Nosé–Hoover chain thermostat (NVT ensemble)<sup>29–31</sup> and a time step of 0.5 fs chosen. AIMD simulations of a 1 × 1 × 2 nm box were then executed for 450 ps.

**Details of the cMD/LMC method.** While a molecular dynamics simulation provides a representative conformational ensemble of the supramolecular structure at the accuracy level of density functional theory (DFT) under periodic boundary conditions, the Monte Carlo part models the long-range mobility of the acidic protons considering the dynamically-evolving molecular structure of the system. In the Lattice Monte Carlo (LMC) part, the system is reduced to a lattice of oxygen atoms that are treated as nodes and can hold one proton. Proton jumps between the oxygen atoms occur, according to a distance-dependent jump rate  $\omega(d_{ij})$ , as in Fig. 1. The jump rate is obtained from the CsH<sub>2</sub>PO<sub>4</sub> AIMD trajectory and is a physical rate for the frequency of a proton transfer between two oxygen atoms. Therefore, it is dependent on the O–O distance. The actual network of transition rates between all LMC lattice sites is obtained from the specific molecular geometry of the system at a given snapshot of the MD trajectory. The propagation of the LMC state is then followed by an update of the transition/jump rate matrix corresponding to the evolution of the MD trajectory. The details of this simulation technique are extensively described by Drefßler *et al.*<sup>20</sup> As a result of the application of the cMD/LMC method we obtain proton fluxes between the AIMD simulation box and its neighboring images. Using these boxes (1 × 1 × 2 nm) as the smallest building block, we can construct the nano-porous models for

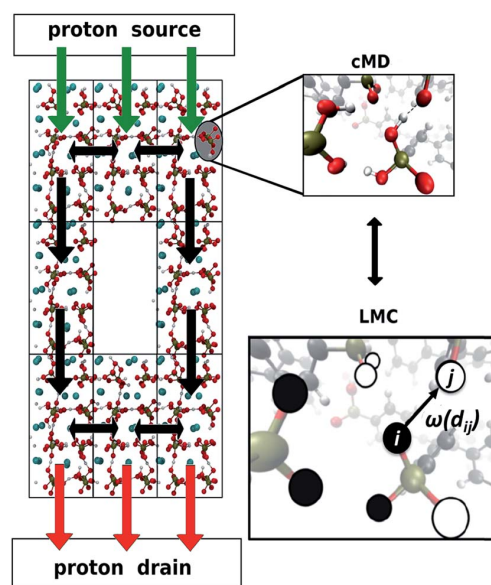


Fig. 1 Schematic of the applied cMD/LMC approach. The molecular dynamics trajectory provides an accurate atomistic structure and its temporal evolution, while the Monte Carlo part models the long-term motion of the acidic protons. The combination enables proton transport calculations for large systems.



Table 1 Parameters for calculations carried out using the cMD/LMC approach

	Excess proton fraction	System size		
		x-Dimension	y-Dimension	z-Dimension
EPF vs. current density	0.001 to 0.5	1 nm	1 nm	100 nm
Current vs. system size	0.5	1 nm	1 nm	12 nm to 2044 nm
Impact of porosity	0.5	8 nm	8 nm	20 nm
Impact of PVP	0.5	8 nm	8 nm	20 nm

the investigation of proton conduction. The proton fluxes for the entire nano-porous system are accessible by combining all elementary proton fluxes from the AIMD boxes. For a nonbiased simulation,  $\text{CsH}_2\text{PO}_4$  possess 0.5 protons per oxygen atom. To simulate the proton current during operation, we insert excess protons on one side of the system (anode) and remove them on the opposite side (cathode). The number of excess protons per oxygen atom at the proton insertion area (anode) is denoted by the excess proton fraction (EPF). The EPF can vary between 0 and 0.5, referring to a neutral simulation or maximally protonated oxygen atoms. The final proton current is controlled by the gradient of the excess protons within the system. An increased excess proton gradient leads to an increased proton current. A summary of the system sizes and excess proton fractions of the performed calculations for this work is provided in Table 1.

## Results and discussion

### Experimental characterization of porous electrodes

Representative SEM micrographs of  $\text{CsH}_2\text{PO}_4$  electrolyte discs at various stages of sample preparation are presented in Fig. 2. The surface roughness decreased significantly from before polishing (a) to afterwards (b), providing a smooth surface for Pt deposition *via* sputtering. Fig. 2(d) shows  $\text{CsH}_2\text{PO}_4$  particles deposited onto an electrolyte pellet *via* spray drying without the stabilizing polymer PVP. The particles agglomerate rapidly, even in ambient conditions. In contrast, the  $\text{CsH}_2\text{PO}_4$  structure can be effectively stabilized by adding PVP to the precursor solution, as reported previously.<sup>16</sup> Fig. 2(e and f) shows a PVP stabilized nanostructure on a dense  $\text{CsH}_2\text{PO}_4$  electrolyte pellet. In general, particle loading on the surface of the electrolyte pellet increases linearly with the duration of spray drying, with

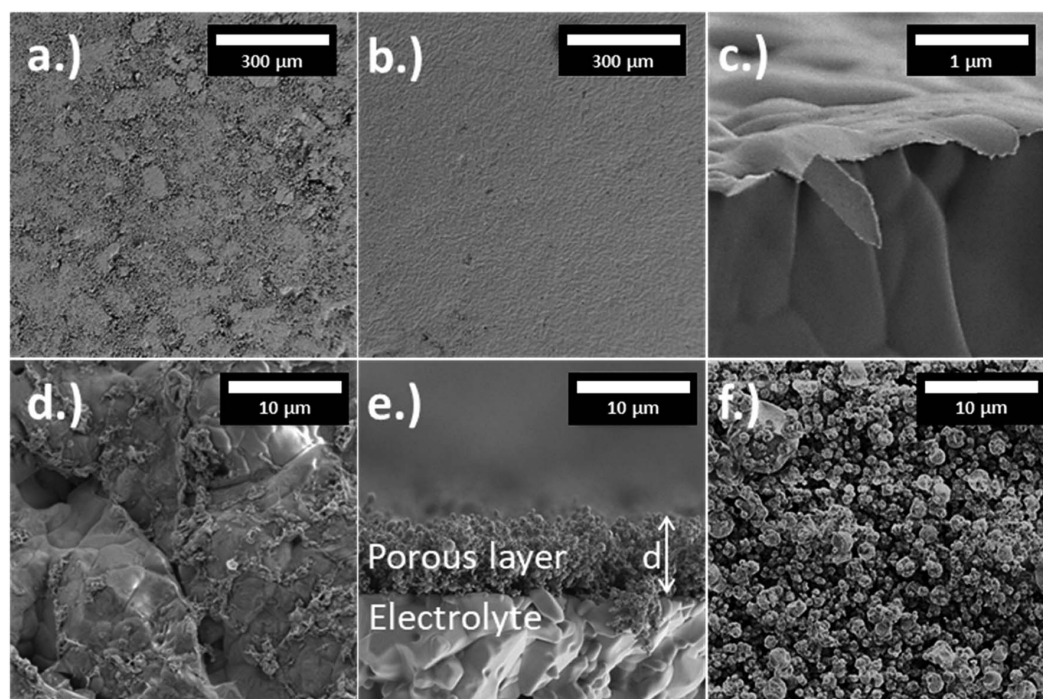


Fig. 2 Scanning electron micrographs of: (a) as pressed; and (b) a polished  $\text{CsH}_2\text{PO}_4$  electrolyte pellet; (c) a cross section of a sputtered Pt thin film; (d) agglomerated pure  $\text{CsH}_2\text{PO}_4$  nanostructure deposited *via* spray drying; (e) a cross section; and (f) a top view of a scanning electron micrograph showing a nanostructured  $\text{CsH}_2\text{PO}_4$  electrode with thickness  $d$ , stabilized by the polymer PVP and deposited on a dense  $\text{CsH}_2\text{PO}_4$  electrolyte pellet.

a slight variation across the pellet. In the initial stage, only the deposited particle density increases on the pellet surface, as depicted in Fig. 3(a). Here, we define the area coverage of the electrolyte surface as the ratio of the projected area of the deposited particles and the free surface area (between 0, for no coverage and 1 for full coverage). After a mixed stage, when both incomplete coverage and growth of the layer thickness  $d$  occurs, a thick porous layer is established, as can be seen in Fig. 3(b and c), respectively. To quantify all stages of nanoparticle deposition, the area coverage was measured with SEM and multiplied with the average thickness of deposited particle clusters, yielding the effective layer thickness. The thermal stability of PVP and mixtures of PVP with  $\text{CsH}_2\text{PO}_4$  and Pt nanoparticles was confirmed using TGA measurements (S2) for temperatures below 600 K in both air and nitrogen atmospheres.

Representative impedance spectra in the Nyquist form are shown in Fig. 4 for a symmetrical cell with an effective layer thickness of  $2.5\ \mu\text{m}$  and a Pt film thickness of 30 nm. The width of the impedance arcs is taken as the electrode resistance. The electrolyte resistance, observed as a shift of the impedance arc along the real axis, is subtracted. For ease of comparison, the results are converted to area-normalized electrode resistances (ANRs). To verify the setup used and ensure comparability to the literature, Pt thin film electrodes were fabricated and characterized as described (ESI, S3†). The electrode resistances of the thin film electrodes are in good agreement with the data published by Louie and Haile.<sup>7</sup> Symmetric measurements in

a hydrogen environment have often been employed to evaluate electrode structures, since they require less construction effort.<sup>3</sup> However, the rate-limiting step for a hydrogen-powered fuel cell is the oxygen reduction reaction at the cathode. We compare structured and thin film cells in symmetrical and fuel cell measurements to analyze whether the structuring is yielding similar activity increases for both techniques. Therefore two cell types, differing only by the electrode type, are compared. “Thin film cells” have 30 nm platinum thin film electrodes on both sides whereas “porous cells” have two structured electrodes with platinum sputtered on top. All cells have the same

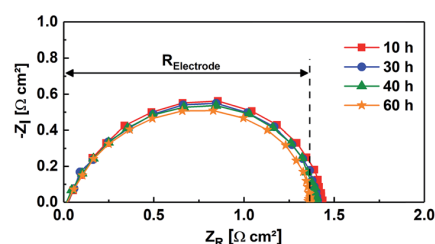


Fig. 4 Electrode impedance arcs as a function of time for a representative symmetric electrochemical cell: Pt thin-film +  $\text{CsH}_2\text{PO}_4$  |  $\text{CsH}_2\text{PO}_4$  |  $\text{CsH}_2\text{PO}_4^{\text{porous}}$  + Pt thin-film showing good stability over a 60 h time period. The Pt film thickness was 30 nm and the effective layer thickness,  $2.5\ \mu\text{m}$ . The electrolyte impedance is subtracted and the area normalized electrode resistance,  $R_{\text{Electrode}}$ , is indicated as the width of the impedance arc.

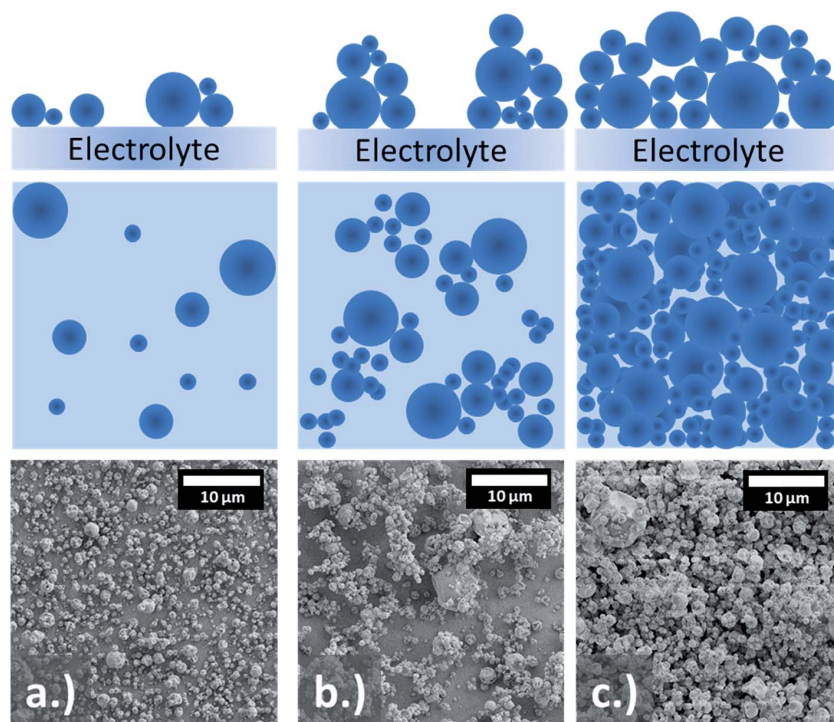


Fig. 3 Schematic drawing (side and top view) of  $\text{CsH}_2\text{PO}_4$  particles deposited onto a dense  $\text{CsH}_2\text{PO}_4$  pellet and corresponding scanning electron micrographs at (a) low; (b) intermediate; and (c) high particle loading.

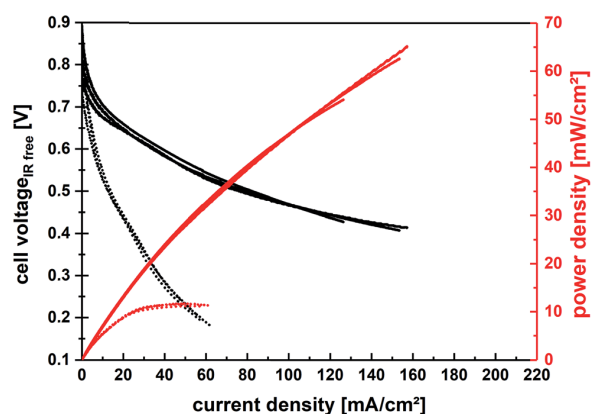


Fig. 5 IR-free polarization plots, as well as the corresponding power density, are shown for three Pt thin film cells (dotted lines) and three cells with ca. 5 μm-thick structured electrodes (solid lines). The platinum content is the same in all presented cells. The results for both electrode types are very well reproducible and a significant increase in the power density by means of electrode structuring is evident. For the porous electrodes, no peak power maximum is reached due to the high electrolyte resistance.

platinum content of 0.128 mg cm<sup>-2</sup>. Fig. 5 compares the IR-corrected polarization plots for cells with Pt thin film electrodes, as well as porous electrodes under fuel cell conditions. The cells show high reproducibility among themselves. To roughly compare the two characterization techniques, the current density at a voltage of 0.5 V is taken for the fuel cell measurements and the mass normalized activity (MNA) is calculated for the symmetric measurement, which represents only the hydrogen oxidation reaction. Both the current densities as well as the MNA depend directly on the catalyst utilization. Table 2 compares the average current density at 0.5 V obtained in fuel cell measurements and the increase of the MNA obtained from symmetrical measurements in hydrogen for both electrode types. The standard deviation of the relative increase of the MNA was calculated as described in Table 2(b). First it becomes clear that structuring an electrode considerably increases its activity. Through an area increase, the sputtered platinum is distributed over a larger area, increasing the active site density. As mentioned earlier, it is widely assumed, but not proven, that structure-induced improvements of the anode

Table 2 Comparison of MNA and current density ( $j$ ) of the thin film and porous cells with the same PGM loading at IR-free voltages of 0.5 V, rounded to two decimal places. The margin of error represents the standard deviation

	Thin film cell	Porous cell	Relative increase
$j@0.5\text{ V (mA cm}^{-2}\text{)}$	$13 \pm 0.7$	$79 \pm 2.3$	$6.1 \pm 0.5$
MNA ( $\text{S mg}_{\text{Pt}}^{-1}$ )	$3.9 \pm 0.1$	$21^a$	$5.5 \pm 0.3^b$

<sup>a</sup> For the MNA at a layer thickness of 5 μm, no standard deviation can be calculated due to a lack of data points. <sup>b</sup> The standard deviation was calculated using the relative error of the current density for the MNA of the porous cell.

performance observed under symmetrical conditions are equally viable for the cathode. Since the reaction mechanism of the oxygen reduction is not yet fully understood, this assumption is not trivial. Second, a similar percentual increase resulting from structuring can be seen for both the MNA and current densities. This implies that the electrode structure affects the anodic hydrogen oxidation and rate-limiting cathodic oxygen reduction in a similar way, confirming symmetric impedance measurements as well-suited method for analyzing improvements to the electrode due to spray drying. The ANR for cells with varying effective layer thicknesses and constant 30 nm Pt thin films, are shown in Fig. 6. The obtained effective layer thickness depends on the spraying time and the exact position of the sample within the deposition section relative to the gas inlet. Therefore, the deposited layer thickness varies from cell to cell, as well as over each sample's surface. The layer thickness of each cell shown in Fig. 6 was measured at four different points and the standard deviation was calculated as described in the experimental section. With the increasing amount of spray dried particles, the electrode impedance decreased sharply as a result of the increase in the absolute electrochemically-active surface area. However, after reaching the lowest ANR at a layer thickness of about 5 μm and complete area coverage, electrode resistance increased rapidly. One should keep in mind that this represents only the optimum for these model cells which can't compete with high platinum loading state of the art cells in terms of power density. As reported by Suryaprakash *et al.*, the platinum deposition *via* sputtering only penetrates the top 5 μm of the structure.<sup>16</sup> It seems that the porous layer significantly increases the ANR, given the fact that a maximum 18 μm-thick layer was applied. As CsH<sub>2</sub>PO<sub>4</sub> itself serves as a frequency-independent ohmic resistor within the measurement conditions employed, the contribution of the porous electrolyte layer to the ANR is surprising. A detailed equivalent circuit based description is available in the ESI S5.† In the following sections, an analytical and cMD/LCM model is introduced showing that the observed effect is mainly caused by the porosity and stabilizing agent PVP. By generating constrictions between the sprayed particles, the porous layer develops a frequency depending resistance affecting the ANR rather than the purely ohmic electrolyte resistance.

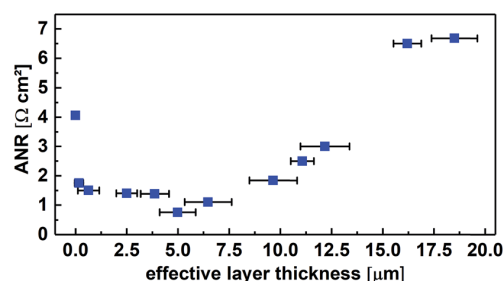


Fig. 6 Area-normalized electrode resistance (ANR) obtained *via* AC impedance spectroscopy plotted vs. electrode layer thickness  $d$ , showing an optimal layer thickness of ca. 5 μm, followed by an unexpected increase for larger thicknesses. All electrodes had a 30 nm Pt thin film catalyst layer deposited *via* DC magnetron sputtering.



### Analytical model

For increasing layer thicknesses a rise in the electrode resistance from  $0.75 \Omega \text{ cm}^2$  to  $7 \Omega \text{ cm}^2$  is observed, Fig. 6. Based on SEM images of the porous structure, an analytical model is developed, not to provide precise data but to estimate the impact of the induced porosity on the resistivity of the electrode. It will be shown, that the resistance of a pure  $\text{CsH}_2\text{PO}_4$  electrolyte network while taking into account the porosity and constriction effects fails to explain the observed resistance increase by one order of magnitude. For this model, a constant resistivity for  $\text{CsH}_2\text{PO}_4$  at 513 K of  $\rho = 45.45 \Omega \text{ cm}$  is assumed.<sup>3</sup> A  $20 \mu\text{m}$ -thick dense electrode therefore has a resistance of  $0.09 \Omega \text{ cm}^2$ . For a porous structure, the tortuosity as well as the porosity influences the resistance.

Analyzing SEM micrographs of the two highest layer thicknesses shown in Fig. 6, the porosity was estimated to be around 0.5 and the extended proton pathway due to tortuosity to be 1.4 times the layer thickness. For this rough estimate four different regions were analyzed as described in the experimental section. Based on this, a  $20 \mu\text{m}$ -thick porous structure is modeled by circular columns with a total area filling of 50%. According to eqn (1),

$$R_{\text{area}} = \frac{\rho}{N_{\text{Col}}} \frac{20 \mu\text{m}}{A_{\text{Col}}} x_t = 0.25 \Omega \text{ cm}^2 \quad (1)$$

the resulting area normalized resistance is  $0.25 \Omega \text{ cm}^2$ . With the resistivity " $\rho$ ", the area of a single column " $A_{\text{Col}}$ ", the number of columns per  $\text{cm}^2$  " $N_{\text{Col}}$ " and the estimated tortuosity " $x_t$ " (ESI, S4†).

Geometric constrictions created between two adjoining particles of different sizes increase the area-normalized resistance. Taking Fig. 7 as a representative image, the smallest constrictions have a diameter of around 300 nm. Assuming ten spherical 300 nm diameter constrictions within a  $20 \mu\text{m}$  layer as an upper limit, the layer resistance increases to  $0.56 \Omega \text{ cm}^2$ , as in eqn (2),

$$R_{\text{area}} = \frac{\rho \left( \frac{20 \mu\text{m}}{A_{\text{Col}}} x_t + \frac{l_{\text{const}}}{A_{\text{const}}} 10 \right)}{N_{\text{Col}}} = 0.56 \Omega \text{ cm}^2 \quad (2)$$

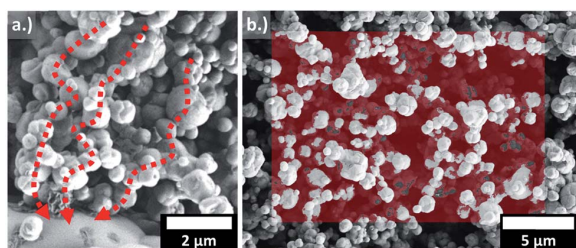


Fig. 7 Representative SEM micrographs of the generated porous  $\text{CsH}_2\text{PO}_4$  electrode network: (a) the dotted lines indicate possible pathways through the porous network. The length is roughly 1.4 times the linear distance. (b) Top view on the porous network. The space between  $\text{CsH}_2\text{PO}_4$  particles, which are in roughly the same plane, are colored in red. The area of  $\text{CsH}_2\text{PO}_4$  particles compared to the void in one plane is about 50%.

with the length of the constriction " $l_{\text{const}}$ " and the area of the constriction " $A_{\text{const}}$ ". The calculations conducted represent a worst-case estimation for a porous system without PVP, based on a SEM characterization of the structure. The extent of the measured increase in electrode impedance from  $0.75 \Omega \text{ cm}^2$  to  $7 \Omega \text{ cm}^2$ , as seen in Fig. 6, cannot be explained by mere electrode structuring and would rather be a result of the used polymer PVP.

Suryaprakash *et al.* report the essential effect of the PVP for the structure's stabilization, which forms a visible shell around the  $\text{CsH}_2\text{PO}_4$  particles during the spraying process for a very high PVP concentration of  $10 \text{ g l}^{-1}$ .<sup>16</sup> As PVP is not proton conducting, such a shell would act as a barrier obstructing proton transport. To prove the significance of PVP for the overall resistivity, the electrolyte resistances of  $500 \mu\text{m}$ -thick pure  $\text{CsH}_2\text{PO}_4$  pellets are compared to a composite cell containing a  $500 \mu\text{m}$ -thick pure  $\text{CsH}_2\text{PO}_4$  layer and a  $100 \mu\text{m}$ -thick dense layer of spray-dried  $\text{CsH}_2\text{PO}_4$ -PVP (ESI, S6†). The electrolyte resistance increases from  $2.5 \Omega \text{ cm}^2$  for the pure  $\text{CsH}_2\text{PO}_4$  pellet to about  $7.2 \Omega \text{ cm}^2$  for the composite  $\text{CsH}_2\text{PO}_4$  pellet. The resistivity of the additional  $100 \mu\text{m}$  thick PVP-containing layer is thus  $5.0 \Omega \text{ cm}$ , representing an almost 11 times increase compared to pure  $\text{CsH}_2\text{PO}_4$ . Assuming this increased resistivity for the calculation according to eqn (2), we achieve a resistance increase comparable to the observed. This shows that the ANR's increase of  $6.25 \Omega \text{ cm}^2$  in Fig. 6 between an effective layer thickness of  $5 \mu\text{m}$  and  $20 \mu\text{m}$  can be caused by the additional effect of the PVP stabilized porous electrolyte network. It becomes clear that the porous structure itself makes a significant contribution to the cell's performance, especially when stabilized with PVP. Although further investigations are necessary, it is important to take the thickness of the porous structure into consideration when engineering a high-performance cell. As spectroscopic analysis of a thin polymer film on a nano-size powder is highly challenging, a computational cMD/LMC model was used to understand the nature of the observed PVP restrictions.

### Theoretical description of a porous $\text{CsH}_2\text{PO}_4$ network using a cMD/LMC approach

Using a cMD/LMC approach, the current density was calculated as a function of first, the proton insertion rate and second, the system length. The excess proton fraction (EPF) scales with the amount of additional hydrogen atoms attached to near-surface O atoms. Higher EPF therefore represents a more active surface. For  $\text{CsH}_2\text{PO}_4$ , the possible number of protons per oxygen atom range from 0.5 to 1, resulting in a maximum EPF value of 0.5. Within this range, the current density shows a linear dependence, depicted in Fig. 8. According to Ohm's law, the current density is proportional to the inverse length of the system. The calculated data fits well ( $R = 0.9992$ ) to the expected behavior, as in Fig. 8.

Porous systems are generated randomly by removing single  $1 \times 1 \times 2 \text{ nm}$  boxes from an  $8 \times 8 \times 20 \text{ nm}$   $\text{CsH}_2\text{PO}_4$  system until a defined porosity is reached. The random removal is restricted in such a way that no percolation interruption is generated.

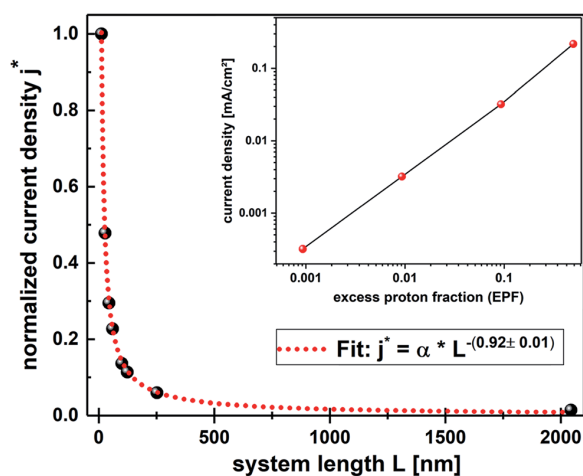


Fig. 8 Calculated current density as a function of system length. The exponent of  $-0.92$  is in good agreement with Ohm's law. Inset: computational model based on a cMD/LMC approach, showing the corresponding current density of a 100 nm-long system as a function of the excess proton fraction (EPF) inserted into the system per time.

Examples for  $\text{CsH}_2\text{PO}_4$  systems with different porosities are depicted in Fig. 9. The conductivity simulations are performed for a fixed EPF, corresponding to a constant platinum surface and voltage. Periodic boundary conditions in  $x$  and  $y$  directions were used for all calculations. For every porosity, three different calculations of randomly obtained systems were performed and the standard deviation for both the conductivity and resulting porosity calculated. In the analytical model, the resistance triples for a porosity of 0.5 without additional geometric constrictions, and an eight-fold increase is observed with these constrictions, keeping in mind that the number of geometric constrictions were set as an upper limit. The cMD/LMC model shows a relative increase in the ohmic resistance of 4-fold for a porosity of 0.5, as

in Fig. 9. The analytical model was closely aligned to the observed porous structure, whereas the cMD/LMC approach aimed at random porous structures. Geometrical constrictions are the main reason for the large error, due to the strong negative effect on the resistivity. It can be concluded that the computational model is able to sufficiently characterize a porous system, as used in this work within the given margin of error.

#### Analyzing the effect of PVP on $\text{CsH}_2\text{PO}_4$ using a cMD/LMC approach

As mentioned previously, the PVP was found to have a high negative impact on the resistivity of  $\text{CsH}_2\text{PO}_4$ . Suryaprakash *et al.* reported the tendency of PVP to form a shell around the  $\text{CsH}_2\text{PO}_4$  particle during the spray drying process.<sup>16</sup> With a concentration ratio of  $1 : 33 \text{ g l}^{-1}$  between PVP and  $\text{CsH}_2\text{PO}_4$ , it is reasonable to conclude that PVP does not form a uniform outer layer around the  $\text{CsH}_2\text{PO}_4$  particles but reduces the contact area between two adjoining  $\text{CsH}_2\text{PO}_4$  particles. The obstructing effect caused by the PVP was modelled by selecting four layers of  $\text{CsH}_2\text{PO}_4$  along the  $z$  axis in an  $8 \times 8 \times 20 \text{ nm}$  system and randomly removing boxes to simulate the area of  $\text{CsH}_2\text{PO}_4$  covered by the PVP. For every surface-coverage in Fig. 10, three different calculations of randomly obtained systems were performed and the averaged values are shown. The resulting proton currents are shown in Fig. 10, using the cMD/LMC approach. An increase of the  $\text{CsH}_2\text{PO}_4$  surface coverage of PVP leads to an exponential increase in the resistance. According to this model, the increase of the resistivity by a factor of over 10, as observed in our cells, would represent a PVP surface coverage of 94%. Following this argumentation, at the boundary between PVP and  $\text{CsH}_2\text{PO}_4$ , a chemical double layer can develop, generating a frequency-dependent resistance. Such a resistance develops an own arc in an impedance spectrogram. In our case the impedance arc of the charge transfer reaction at the platinum- $\text{CsH}_2\text{PO}_4$  interface and from the PVP induced constrictions overlap, leading to a flattened overall

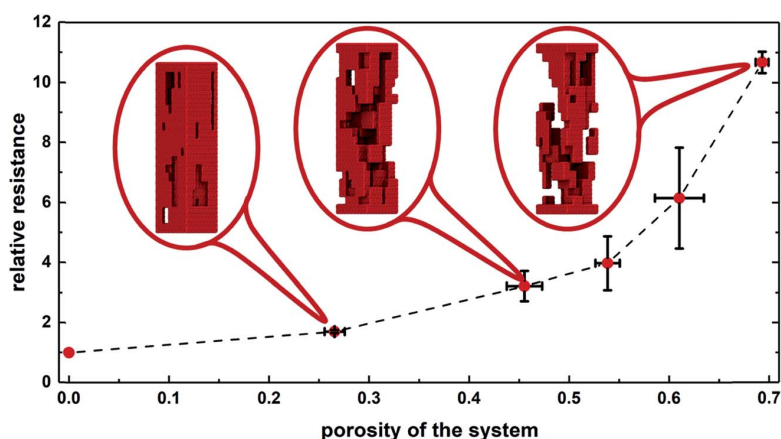


Fig. 9 Porous model cells with varying porosity are generated by randomly deleting single boxes. No dead ends were generated in this way, taking into account the periodic boundary conditions. The relative resistance, calculated by the cMD/LMC model, is plotted against the porosity with the associated margin of error shown for each system. The dashed line is a visual guideline.

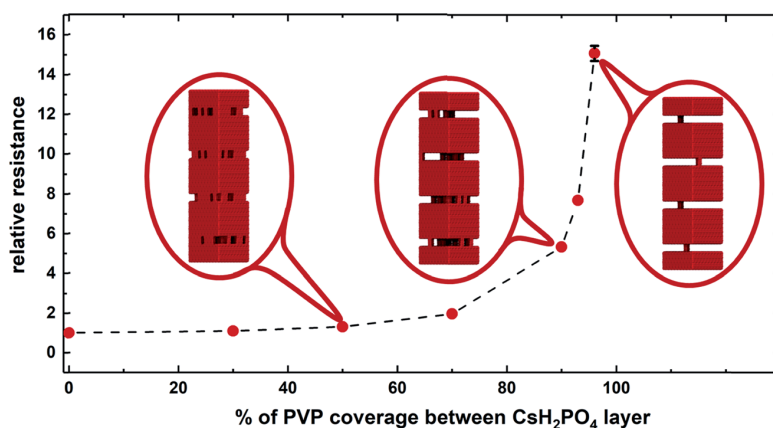


Fig. 10 To simulate the area of  $\text{CsH}_2\text{PO}_4$  covered by PVP, four layers of  $\text{CsH}_2\text{PO}_4$  boxes with equidistant spacing along the  $z$  axis were randomly removed to different degrees from an  $8 \times 8 \times 20$  nm  $\text{CsH}_2\text{PO}_4$  system. The relative resistance calculated by the cMD/LMC model is plotted against the percentage of the area covered by PVP with the associated margin of error shown for each system. The dashed line is a visual guideline and shows exponential growth for a high degree of coverage.

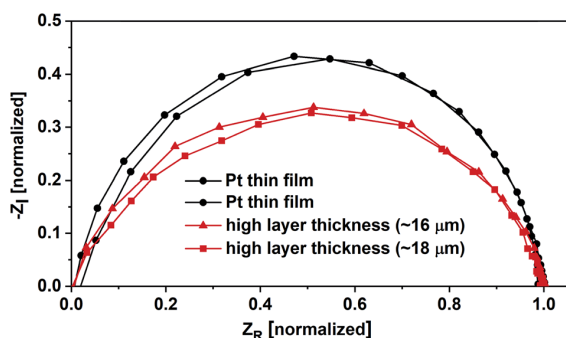


Fig. 11 Normalized impedance arcs of platinum thin film electrodes and porous electrodes are compared. While thin film electrodes generate an almost semi-circle like symmetry, the symmetry of the porous electrode arcs is significantly deformed due to PVP introduced constrictions which obstruct the proton movement.

impedance arc ( $\text{ESI S5}^\dagger$ ). This effect is shown in Fig. 11, where the impedance arc symmetry of platinum thin film electrodes and porous electrodes are compared. All arcs are normalized for better comparison. While thin film electrodes show an almost semicircle like symmetry, the arc's associated with the porous electrodes are deformed due to the additional frequency dependent resistance.

## Conclusion

Solid acid fuel cell electrodes consisting of a  $\text{CsH}_2\text{PO}_4$  nanostructured layer and a Pt thin film were fabricated *via* spray drying and DC magnetron sputtering. The porous layer thickness was systematically varied between 0 and 18  $\mu\text{m}$ , while the Pt thin film was kept constant at 30 nm. The applicability of symmetric impedance measurements in hydrogen as a method to determine the activity of an electrode under fuel cell conditions was confirmed. The electrode impedance for different

layer thicknesses was determined *via* AC impedance spectroscopy in a humidified hydrogen environment, showing a fast decrease when increasing the layer thickness to 5  $\mu\text{m}$  which is associated with an increase of the active site density. After complete area coverage and a layer thickness above 5  $\mu\text{m}$ , magnetron sputtering couldn't penetrate the whole porous structure anymore resulting in a constant active site density. The followed step increase of the measured electrode impedance for thicknesses greater than 7.5  $\mu\text{m}$  therefore results from the negative impact of the PVP containing porous  $\text{CsH}_2\text{PO}_4$  network. The lowest symmetrical electrode impedance corresponds to a layer thickness of about 5  $\mu\text{m}$ , resulting in a 5.5-fold increase in the Pt-mass normalized activity compared to thin film electrodes. The impact of a polymer-stabilized porous  $\text{CsH}_2\text{PO}_4$  network was investigated using a multiscale simulation approach (cMD/LMC) as well as an analytical model. Especially for highly porous systems, a significant increase in the resistance was shown. Even if a high porosity is favorable in terms of active site optimization, the resulting electrolyte network resistance can quickly become rate limiting.

To develop a deeper understanding of the polymer induced resistance increase, a partial proton blocking layer was implemented in the cMD/LMC approach, simulating the incomplete PVP cover around the sprayed  $\text{CsH}_2\text{PO}_4$  particles. It was shown experimentally that the stabilizing agent PVP reduces the conductivity by a factor of more than ten, corresponding to a PVP surface coverage of 94% in the cMD/LMC model. Beyond the presented data, our results may enable researchers to further improve the performance of solid acid fuel cells. Current state-of-the-art SAFC electrodes work with electrodes as thick as 100  $\mu\text{m}$ , possibly limited by the proton conduction resistance of the porous  $\text{CsH}_2\text{PO}_4$  electrode network.

## Conflicts of interest

There are no conflicts of interest to declare.

## Acknowledgements

The authors thank Dietmar Hirsch and Andrea Prager for assistance with SEM imaging, Petra Hertel for sputter deposition and Ingrid Reinhardt for TGA measurements. Financial support for this work was provided through internal funds of the Leibniz Institute for Surface Modification and by the European Social Fund (ESF) of the European Union (EU), the Free State of Saxony and the German Federal Environmental Organization (DBU) and the Volkswagen Foundation. The authors also thank the reviewers for important comments that improved the manuscript significantly. Funding from the German Science Foundation (DFG Project-ID 189853844 – TRR 1029 is also gratefully acknowledged.

## Notes and references

- D. A. Boysen, T. Uda, C. R. I. Chisholm and S. M. Haile, *Science*, 2004, **303**, 68–70.
- C. R. I. Chisholm and S. M. Haile, *Solid State Ionics*, 2000, **136–137**, 229–241.
- S. M. Haile, C. R. I. Chisholm, K. Sasaki, D. A. Boysen and T. Uda, *Faraday Discuss.*, 2007, **134**, 17–39.
- C. R. I. Chisholm, D. A. Boysen, A. B. Papandrew, S. Zecevic, S. Yal Cha, K. A. Sasaki, A. Varga, K. P. Giapis and S. M. Haile, *Electrochem. Soc. Interface*, 2009, 53–59.
- A. B. Papandrew, C. R. I. Chisholm, R. A. Elgammal, M. M. Ozer and S. K. Zecevic, *Chem. Mater.*, 2011, **23**, 1659–1667.
- A. B. Papandrew, C. R. I. Chisholm, S. K. Zecevic, G. M. Veith and T. A. Zawodzinski, *J. Electrochem. Soc.*, 2013, **160**, F175–F182.
- M. W. Louie and S. M. Haile, *Energy Environ. Sci.*, 2011, **4**, 4230.
- V. S. Thoi, R. E. Usiskin and S. M. Haile, *Chem. Sci.*, 2015, **6**, 1570–1577.
- A. Varga, N. A. Brunelli, M. W. Louie, K. P. Giapis and S. M. Haile, *J. Mater. Chem.*, 2010, **20**, 6309–6315.
- International Energy Outlook 2016*, ed. J. Conti, *et al.*, U.S. Department of Energy, 2016, [www.eia.gov/forecasts/ieo/pdf/0484\(2016\).pdf](http://www.eia.gov/forecasts/ieo/pdf/0484(2016).pdf).
- S. M. Haile, D. A. Boysen, C. R. I. Chisholm and R. B. Merle, *Nature*, 2001, **410**, 910–913.
- T. Uda and S. M. Haile, *Electrochem. Solid-State Lett.*, 2005, **8**, A245–A246.
- A. B. Papandrew, R. A. Elgammal, M. Tian, W. D. Tennyson, C. M. Rouleau, A. A. Poretzky, G. M. Veith, D. B. Geohegan and T. A. Zawodzinski, *J. Power Sources*, 2017, **337**, 145–151.
- F. P. Lohmann, P. S. C. Schulze, M. Wagner, O. Naumov, A. Lotnyk, B. Abel and A. Varga, *J. Mater. Chem. A*, 2017, **5**, 15021–15025.
- O. Naumov, S. Naumov, R. Flyunt, B. Abel and A. Varga, *Chemoschem*, 2016, **9**, 3298–3306.
- R. C. Suryaprakash, F. P. Lohmann, M. Wagner, B. Abel and A. Varga, *RSC Adv.*, 2014, **4**, 60429–60436.
- G. Kabbe, C. Wehmeyer and D. Sebastiani, *J. Chem. Theory Comput.*, 2014, **10**, 4221–4228.
- G. Kabbe, C. Drefßler and D. Sebastiani, *Phys. Chem. Chem. Phys.*, 2017, **19**, 28604–28609.
- G. Kabbe, C. Drefßler and D. Sebastiani, *J. Phys. Chem. C*, 2016, **120**, 19905–19912.
- C. Drefßler, G. Kabbe and D. Sebastiani, *J. Phys. Chem. C*, 2016, **120**, 19913–19922.
- J. Hutter, M. Iannuzzi, F. Schiffmann and J. VandeVondele, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2014, **4**, 15–25.
- Y. Zhang and W. Yang, *Phys. Rev. Lett.*, 1998, **80**, 890.
- J. P. Perdew, A. Ruzsinszky, G. I. Csonka, O. A. Vydrov, G. E. Scuseria, L. A. Constantin, X. Zhou and K. Burke, *Phys. Rev. Lett.*, 2008, **100**, 136406.
- J. P. Perdew, K. Burke and M. Ernzerhof, *Phys. Rev. Lett.*, 1996, **77**, 3865–3868.
- J. VandeVondele and J. Hutter, *J. Chem. Phys.*, 2007, **127**, 114105.
- C. Hartwigsen, S. Goedecker and J. Hutter, *Phys. Rev. B*, 1998, **58**, 3641–3662.
- M. Krack, *Theor. Chem. Acc.*, 2005, **114**, 145–152.
- S. Grimme, J. Antony, S. Ehrlich and H. Krieg, *J. Chem. Phys.*, 2010, **132**, 154104.
- S. Nose, *J. Chem. Phys.*, 1984, **81**, 511–519.
- G. J. Martyna, M. L. Klein and M. Tuckerman, *J. Chem. Phys.*, 1992, **97**, 2635–2643.
- S. Nose, *Mol. Phys.*, 1970, **52**, 255–268.

### 3.1.4 Article IV: Proton Mobility in Aqueous Systems: Combining *ab initio* Accuracy with Millisecond Timescales

Gabriel Kabbe, Christian Dreßler, and Daniel Sebastiani.

Proton mobility in aqueous systems: combining *ab initio* accuracy with millisecond timescales.

*Physical Chemistry Chemical Physics*, 19(42):28604–28609, **2017**.

For this publication, I provided *ab initio* molecular dynamics trajectories of  $\text{H}_3\text{O}^+$  in water. G. Kabbe incorporated the dielectric relaxation effects that can be observed within the first solvation shell of  $\text{H}_3\text{O}^+$  into the cMD/LMC scheme. D. Sebastiani supervised the project and gave valuable advice during the development and while G. Kabbe was writing the manuscript.





# Proton mobility in aqueous systems: combining *ab initio* accuracy with millisecond timescales†

Cite this: *Phys. Chem. Chem. Phys.*, 2017, 19, 28604

Gabriel Kabbe,  Christian Dreßler and Daniel Sebastiani \*

Received 17th August 2017,  
Accepted 10th October 2017

DOI: 10.1039/c7cp05632j

rsc.li/pccp

We present a multiscale simulation of proton transport in liquid water, combining *ab initio* molecular dynamics simulations with force-field ensemble averaging and kinetic Monte-Carlo simulations. This unique Ansatz allows for *ab initio* accuracy incorporating the femtosecond dielectric relaxation dynamics of the aqueous hydrogen bonding network, and bridges the time-scale gap towards the explicit simulation of millisecond diffusion dynamics.

## 1 Introduction

Proton transfer in water plays an important role in many chemical and biological processes.<sup>1–3</sup> Especially with regard to Proton Exchange Membrane (PEM) fuel cells, the high conductivity of water is a desirable property and the reason for the efficiency of Nafion-based membranes.<sup>4–8</sup> These membranes enclose water channels, in which proton conduction is possible. In this context, a lot of research has been conducted to further clarify the proton conduction mechanisms both experimentally and theoretically.

A model picture of the underlying mechanisms which explain the anomalously high proton conduction in water was actually first proposed by Grotthuss more than 200 years ago.<sup>9</sup> It comprises two main elements: a hopping step, which is later followed by reorientation. The hopping step can actually happen as a series of proton transfers along a chain of water molecules, or rather “water wire”,<sup>10,11</sup> which leads to the propagation of the excess charge along the chain without any diffusive motion of the molecules themselves, and a simultaneous reversal of the chain’s dipole moment. Consequently, the subsequent rotation of the water molecules is necessary in order to enable further proton transfer along the same “water wire”.

In recent years, this idealized model picture has been refined further. In particular, the water reorientation by rotation is believed to be an oversimplification as a water molecule in bulk water is hydrogen-bonded to four neighbor molecules, which hinders free rotation. Instead, the breaking of a hydrogen bond, which is necessary to reduce the hydrogen-bond coordination

number of an accepting H<sub>2</sub>O from four to three, is considered the rate-determining step.<sup>12,13</sup>

With the continuously increasing computing capabilities, *ab initio* calculations of proton transfer in water have become feasible. It is now possible to investigate proton transfer on the atomistic level.<sup>13–19</sup> However, while offering a high accuracy, these methods are still computationally expensive and give only access to timescales of tens to hundreds of picoseconds. Especially with regard to the simulation of macroscopic properties of proton exchange membranes, they often do not allow statistically relevant sampling of the excess proton movement in a reasonable amount of time. Similarly, the adequate incorporation of co-solvents, geometric confinement, effects, or the influence of functional interfaces is computationally challenging.

In this work, we present a coupled Molecular Dynamics/Lattice Monte Carlo (cMD/LMC) approach, which aims at simulating large-scale excess charge diffusion in aqueous systems. This combined method has already been applied successfully to PEM materials<sup>20,21</sup> and solid acids.<sup>22</sup> Based on structural dynamics calculated from molecular dynamics (MD), it builds a time-dependent excess charge transition matrix for the Lattice Monte Carlo (LMC) scheme. The mapping between the topology of the system and the resulting excess charge transfer rate is determined with statistical analysis from *ab initio* MD.

On the basis of a neutral classical MD trajectory of water, we show that our cMD/LMC scheme is able to determine dynamical quantities of the excess charge in best agreement with *ab initio* methods over large time scales.

## 2 Method

For the propagation of an excess proton through a liquid water system, we have developed a hybrid scheme which combines aspects from molecular dynamics (MD), and lattice/kinetic

Institute of Chemistry, Martin-Luther-University Halle-Wittenberg, von-Danckelmann-Platz 4, 06120 Halle, Germany.

E-mail: daniel.sebastiani@chemie.uni-halle.de

† Electronic supplementary information (ESI) available: Proton jump rate functions determined from AIMD, illustration of the linear distance rescaling. See DOI: 10.1039/c7cp05632j

Monte Carlo methods.<sup>23–25</sup> Protons can jump between lattice sites with transfer rates which are extracted on-the-fly from an underlying MD simulation (see also ref. 20–22). Each lattice site represents an oxygen atom of the considered system and can be in one of two states: occupied or unoccupied. The trajectory-based evolution of the pairwise oxygen distances in the MD results in a change of the proton transition rate in the LMC scheme. This scheme has been shown to yield a realistic picture of condensed-phase systems.

In order to take the characteristic hydrogen-bond network of  $\text{H}_3\text{O}^+$  in liquid water into account, we construct the LMC topology such that each oxygen site is connected to its closest three neighbors.

### 2.1 Lattice Monte Carlo scheme

The LMC system consists of a discrete range of states and is described by a master equation with time-dependent transition rates:

$$\frac{dp_i(t)}{dt} = \sum_j [p_j(t)\omega_{ji}(t) - p_i(t)\omega_{ij}(t)] \quad (1)$$

where  $p_i(t)$  is the probability to find the system in state  $i$ , and  $\omega_{ij}$  is the transition rate from state  $i$  to state  $j$ .

The scheme first translates the transfer rate  $\sum_j \omega_{ij}$  from a lattice site  $i$  into a “lifetime” of this very site. In this context, the lifetime is defined as the duration  $\tau$ , after which the excess charge at lattice site  $i$  is transferred to an adjacent site. In our specific algorithm, this lifetime is calculated by drawing a random number  $x$  in the interval  $(0, 1]$  and solving the equation

$$-\ln(x) = \int_0^\tau \sum_j \omega_{ij}(t') dt' \quad (2)$$

for  $\tau$ .

Once this specific lifetime has passed, the algorithm determines the target site for the excess charge by drawing another random number  $x$  from the interval  $\left[0, \sum_j \omega_{i_0,j}(t)\right]$  and choosing the second oxygen site with index  $j_0$  such that

$$\sum_0^{j_0-1} \omega_{i_0,j}(t) \leq x < \sum_0^{j_0} \omega_{i_0,j}(t) \quad (3)$$

For a more detailed mathematical explanation, see Prados *et al.*<sup>23</sup>

### 2.2 Determination of the excess charge diffusion

The challenge in determining the excess charge motion from an atomic trajectory is the fact that there is no “labeled” excess charge particle whose position can be tracked. Instead, similar to the Grotthuss picture, a multitude of protons are involved during a (long-range) excess charge transfer.

Within the *ab initio* molecular dynamics (AIMD) simulations, the excess charge diffusion is therefore determined by

means of a collective variable specifically designed for this task, with an effective coordinate given by:

$$\mathbf{R}^{\text{excess}} = -2 \sum_{\text{O}} \mathbf{R}_{\text{O}} + \sum_{\text{H}} \mathbf{R}_{\text{H}} \quad (4)$$

$\mathbf{R}_{\text{O}}$  and  $\mathbf{R}_{\text{H}}$  in eqn (4) stand for the positions of the oxygen atoms and hydrogen atoms, respectively.  $\mathbf{R}^{\text{excess}}$  has the property that it is invariant under translations of neutral  $\text{H}_2\text{O}$  molecules while being sensitive to the motion of single (excess) protons and  $\text{H}_3\text{O}^+$  complexes. This effective variable can easily be adjusted for more complex chemical systems (as has been demonstrated for R-PH<sub>2</sub>O<sub>3</sub>/R-PH<sub>3</sub>O<sub>3</sub><sup>26</sup>).

In contrast to MD, the cMD/LMC scheme represents the excess charge position explicitly as an overprotonated molecular state that is assigned to a water molecule at each time step. It is therefore trivial to follow the excess charge over time and determine its diffusion.

Unlike MD, the proton motion in the cMD/LMC scheme is not continuous: when a water molecule transfers its excess charge to another nearby molecule, the excess charge position changes instantaneously from the position of the donor molecule to the position of the acceptor molecule. As a consequence, the mean squared displacement function is discontinuous and changes stepwise by the full oxygen–oxygen distance.

### 2.3 Excess charge transfer rate determination

In order to obtain a realistic transfer rate matrix between individual lattice sites for the cMD/LMC scheme, the instantaneous atomic geometry around a protonated water molecule must be mapped onto the probability of an excess charge transfer for a given constellation of the  $\text{H}_3\text{O}^+$ 's surroundings. We determine this mapping by statistically analysing the excess charge transfer events within the *ab initio* MD trajectory and determining the transfer rate as a function of the oxygen–oxygen distance between  $\text{H}_3\text{O}^+$  and the surrounding  $\text{H}_2\text{O}$  molecules by means of the following algorithm for each time step  $t$ :

(1) Determine the distance distribution  $\rho^{\text{OO}}(d^{\text{OO}}, t)$  of potential  $\text{H}^+$  acceptor water molecules around the  $\text{H}_3\text{O}^+$ .

(2) Lookup the distance distribution  $\rho^{\text{jump}}(d^{\text{OO}}, t)$  of actual  $\text{H}^+$  acceptor molecules, *i.e.* consider only those waters to which a proton jump has actually occurred in the next MD step at  $t + \Delta t$ .

(3) The proton jump rate at time  $t$  is determined *via*

$$\omega^{\text{jump}}(d^{\text{OO}}, t) = \frac{1}{\Delta t} \frac{\rho^{\text{jump}}(d^{\text{OO}}, t)}{\rho^{\text{OO}}(d^{\text{OO}}, t)} \quad (5)$$

The final jump rate is obtained by averaging over all time steps

$$\omega^{\text{jump}}(d^{\text{OO}}) = \langle \omega^{\text{jump}}(d^{\text{OO}}, t) \rangle_t \quad (6)$$

We fit this jump rate to a function of the form

$$f(d) = \frac{a}{1 + \exp\left(\frac{d - d_0}{c}\right)} \quad (7)$$

The resulting fit parameters, from our *ab initio* molecular dynamics simulation are listed in the ESI.†

## 3 Results

### 3.1 Analysis of *ab initio* molecular dynamics of protonated water

**3.1.1 Trajectory setup.** For this work, *ab initio* trajectories of protonated H<sub>2</sub>O and a classical trajectory of neutral H<sub>2</sub>O were simulated. We used the software CP2K<sup>27,28</sup> for the *ab initio* simulations, and Gromacs<sup>29–31</sup> for the classical MD simulation.

While the *ab initio* trajectories are used both as a reference system and for analysis of the dynamical properties of H<sub>3</sub>O<sup>+</sup>, the classical trajectory provides the underlying topology for calculating the excess charge jump rates of the cMD/LMC scheme.

AIMD trajectories of 100 water molecules and an excess proton in a periodic box of length  $14.405 \times 14.405 \times 14.405 \text{ \AA}^3$  were used to determine the proton jump rates and analyze the radial distribution function (RDF) between hydronium and neutral water. In order to investigate the temperature dependence of the observables, three trajectories at 300 K, 350 K and 400 K were run. We used the BLYP functional<sup>32,33</sup> with a DZVP-MOLOPT-SR-GTH basis and the GTH pseudopotential.<sup>34</sup> The DFT-D3 dispersion correction was used.<sup>35</sup> The molecular dynamics simulation was run with a timestep of 0.5 fs using the Nosé–Hoover thermostat.<sup>36</sup>

At the force field level, an MD trajectory consisting of 216 water molecules in a periodic box of length  $18.626 \times 18.626 \times 18.626 \text{ \AA}^3$  was run on a simulation time scale of 10 ns at 350 K. For the classical MD, we used the TIP4P water model with the Amber force field.<sup>37</sup>

**3.1.2 *Ab initio* excess charge diffusion.** We determine the mean squared displacement (Fig. 1) from  $\mathbf{R}(t)$  (see Section 2.2). After about 5 to 10 ps, the mean squared displacement (MSD) reaches an almost linear regime. The remaining fluctuations are due to the limited statistics (only one excess proton is present in the water box). From a linear fit we get a diffusion coefficient  $D = (1.48 \pm 0.01) \text{ \AA}^2 \text{ ps}^{-1}$ . We use this value as a reference to judge the results of the cMD/LMC approach.

**3.1.3 Excess charge transfer rate.** We determine the excess charge transfer rate according to Section 2.3 from the available AIMD trajectories at 300 K, 350 K and 400 K. The resulting transfer rates as functions of the donor–acceptor distance are shown in the ESI.† We observe a noisy development of the jump rates at the

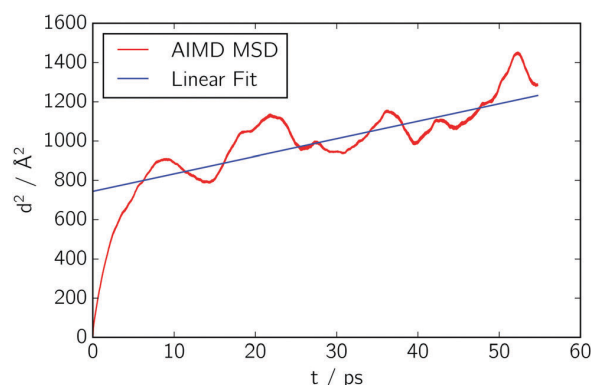


Fig. 1 Mean squared displacement of the excess charge within the *ab initio* trajectory. The slope of the linear fit is used to determine the diffusion coefficient.

borders of the distance axis. This is due to the fact that only few O–O distances in these regions are sampled within the AIMD trajectories. In the middle region, on the other hand, the better sampling leads to a more steady course. We take the varying sampling into account by weighting the jump rate results with the number of observed hydronium–H<sub>2</sub>O pairs  $N^{\text{OO}}$ . All three jump rates yield very similar results, which is in agreement with previous studies.<sup>20</sup> In the following sections, the trajectory at 400 K will be used.

### 3.2 Dielectrical relaxation following an elementary proton hopping step

The presence of an excess proton in water leads to considerable electrostatic forces, which influence the structure of water molecules in the vicinity of the H<sub>3</sub>O<sup>+</sup> ion. This effect is incorporated automatically by AIMD, which is known to yield a precise description of both structure and femto/picosecond dynamics around a solvated charge,<sup>15,16,38</sup> but is missing within the cMD/LMC scheme, as the excess charge state within the LMC is not returned as feedback to the MD simulation. As a result, the acceptor–donor distances in the LMC scheme are overestimated, which leads to a diminished proton hopping rate and therefore an underestimation of the overall excess charge diffusion.

We propose a scheme which adjusts the trajectory data to take into account this dielectric response of the water upon protonation changes. Specifically, we analyze how the presence of an excess charge in the AIMD influences the RDF of the first solvation shell. This change is subsequently used to construct a conversion function that provides a mapping between the pairwise oxygen distances of an uncharged water system and those of a protonated system within the first solvation shell.

**3.2.1 Construction of a conversion function.** The micro-solvation of an H<sub>3</sub>O<sup>+</sup> cation differs in two ways from that of a neutral H<sub>2</sub>O molecule: first, the three donor hydrogen bonds are shorter, and second, the acceptor bridge is elongated. Since our force-field MD only contains H<sub>2</sub>O molecules, we have to consider the different solvation structure within the Monte Carlo part, in order to yield more realistic excess charge transfer rates.

To address this issue, we construct a conversion function which maps the distance distribution of an uncharged water and its first solvation shell to the distance distribution of a hydronium ion and its first solvation shell. This allows us to take the structural effect of a charged particle within the LMC scheme into account by effectively rescaling the distances between an occupied oxygen site and its neighbors.

For the construction of the conversion function, we use the integrated RDF, which yields the number of particles within a radius  $r$ . It is defined as

$$n(r) = \int_0^r 4\pi x^2 \rho g(x) dx \quad (8)$$

with the particle density  $\rho$  and the radial distribution function  $g(x)$ .

We also define  $n_{\text{hydronium}}(r)$  as the integrated RDF of the desired distribution (*i.e.* the hydronium–H<sub>2</sub>O distribution), and  $n_{\text{neutral}}(r)$  as the integrated RDF of the original distribution in the neutral H<sub>2</sub>O system.

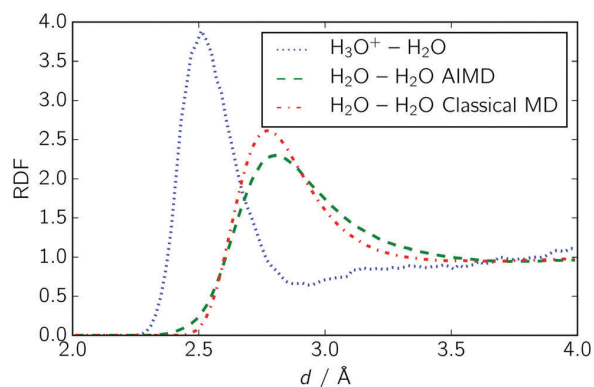


Fig. 2 Comparison of the radial distribution functions between neutral water molecules, and between a hydronium ion and water. The dotted blue line with a peak at 2.5 Å shows the radial distribution function of a hydronium ion and water. The dashed green line and the dashdotted red line show the radial distribution functions of water from our *ab initio* and classical MD simulations.

The conversion function  $c(d)$  is then defined as:

$$c(d) = n_{\text{hydronium}}^{-1}(n_{\text{neutral}}(d)) \quad (9)$$

using the inverse function of  $n_{\text{hydronium}}$ . The lower picture of Fig. 3 shows the resulting conversion.

In our implementation, we calculate the conversion function at discrete points, and use a linear interpolation scheme to provide a continuous mapping.

Note that the conversion function effectively shortens the donor hydrogen bonds by around 0.25 Å in the range between 2.5 Å and 3 Å. This agrees well with the distance of the peaks of the  $\text{H}_3\text{O}^+-\text{H}_2\text{O}$  versus the  $\text{H}_2\text{O}-\text{H}_2\text{O}$  (see Fig. 2).

### 3.2.2 Modeling relaxation time within the LMC scheme.

Upon the jump of a proton from one water to another, the hydrogen bond network responds in the way described above (Section 3.2.1). However, this response is not instantaneous, but has a characteristic relaxation time. In order to simulate this relaxation of the water molecules in the vicinity of the newly emerged  $\text{H}_3\text{O}^+$  ion, we rescale the O–O distances when an excess proton has just jumped (at time  $t_0$ ) linearly in time from the unrescaled ( $d^u$ ) to the rescaled ( $d^r$ ) distances.

$$d^{\text{OO}}(t) = d_{\text{neutral}}^{\text{OO}}(t) + \frac{t - t_0}{t_{\text{relax}}} \left[ d_{\text{hydronium}}^{\text{OO}}(t) - d_{\text{neutral}}^{\text{OO}}(t) \right] \quad (10)$$

with  $t_0 < t < t_0 + t_{\text{relax}}$ .

The choice of the dielectric relaxation time parameter  $t_{\text{relax}}$  determines the time span, in which the linear rescaling takes place. It turns out that an adequate range for  $t_{\text{relax}}$  is found between 2 and 10 ps.

### 3.3 Proton dynamics in the cMD/LMC approach

In the following, we investigate how the cMD/LMC scheme reproduces dynamical quantities such as MSD and the excess charge autocorrelation. For this, we compare the dynamical properties of the LMC scheme with the reference AIMD calculation and experimental results.

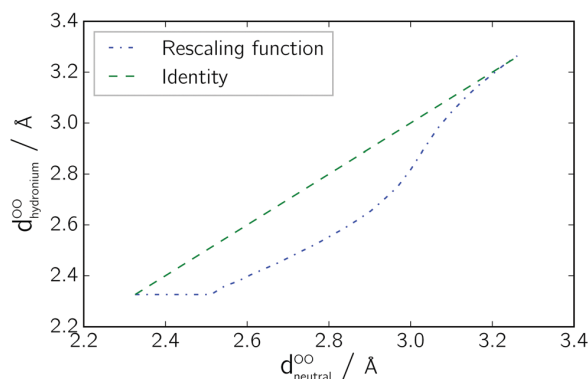
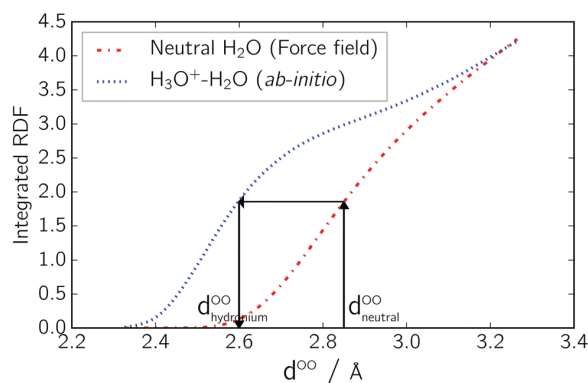


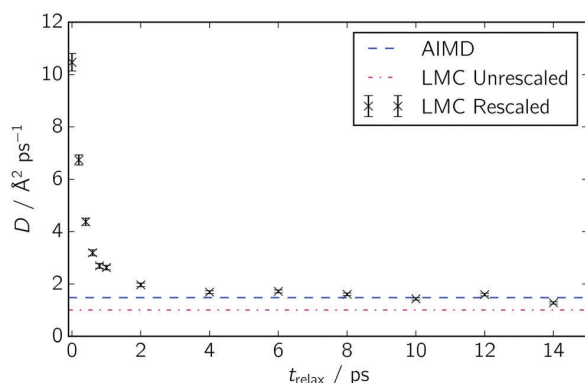
Fig. 3 Upper picture: Cumulative distribution functions of the oxygen–oxygen distribution for neutral water, and for neutral water and a hydronium ion. The arrows indicate how, given an unrescaled distance of a neutral system  $d_{\text{neutral}}^{\text{OO}}$ , one arrives at the corresponding rescaled distance  $d_{\text{hydronium}}^{\text{OO}}$ . Lower picture: Resulting conversion function, which maps a distance  $d_{\text{neutral}}^{\text{OO}}$  from the oxygen distance distribution of uncharged water molecules onto the distance  $d_{\text{hydronium}}^{\text{OO}}$  from a distance distribution between an  $\text{H}_3\text{O}^+$  and a neutral water molecule.

**3.3.1 Diffusion coefficients.** Table 1 lists the resulting diffusion coefficients of the LMC scheme with different parameters, the reference AIMD trajectory, and experimental data. *Ab initio* and experimental results differ by about a factor of two. This is a good agreement, as routinely factors between 0.1 and 10 are reported.<sup>40,41</sup>

Next, we compare the results of the cMD/LMC scheme, and *ab initio* results. Fig. 4 shows the variation of the excess charge diffusion coefficient with respect to the relaxation time parameter. We see that the diffusion coefficient is in best agreement with the reference AIMD for values  $t_{\text{relax}}$  between 4 and around 12 ps. Values of  $t_{\text{relax}}$  below 2 ps, on the other hand, lead to unrealistically high diffusion coefficients which differ by an order of magnitude. Using the oxygen distances of the neutral water system yields a lower diffusion coefficient (around  $1 \text{ \AA}^2 \text{ ps}^{-1}$ ) than our reference AIMD. The reason for this behavior can be seen from Fig. 2: the neutral oxygen–oxygen distances lie around a distance of 1.7 Å, which corresponds to a proton jump rate of less than one jump per ps. It should be noted that the lowest possible diffusion rate in our scheme occurs, if the excess charge does not

**Table 1** Comparison of the diffusion coefficients of the cMD/LMC scheme and a reference AIMD simulation

Method	$D/\text{\AA}^2 \text{ps}^{-1}$
LMC (w/rescaling, $t_{\text{relax}} = 0$ ps)	10.5(3)
LMC (w/rescaling, $t_{\text{relax}} = 4$ ps)	1.69(6)
LMC (w/rescaling, $t_{\text{relax}} = 6$ ps)	1.72(5)
LMC (w/o rescaling, ( $t_{\text{relax}} \rightarrow \infty$ ))	1.01(3)
Reference AIMD (this work)	1.48(1)
Experimental	0.865 <sup>39</sup>



**Fig. 4** Diffusion coefficients of the cMD/LMC scheme against the relaxation time parameter  $t_{\text{relax}}(x)$ . The dashdotted line shows the diffusion coefficient of the cMD/LMC scheme with unrescaled distance (representing the limiting case for  $t_{\text{relax}} \rightarrow \infty$ ). The dashed line represents the diffusion coefficient of the *ab initio* simulation at  $1.48 \text{\AA}^2 \text{ps}^{-1}$  as a reference.

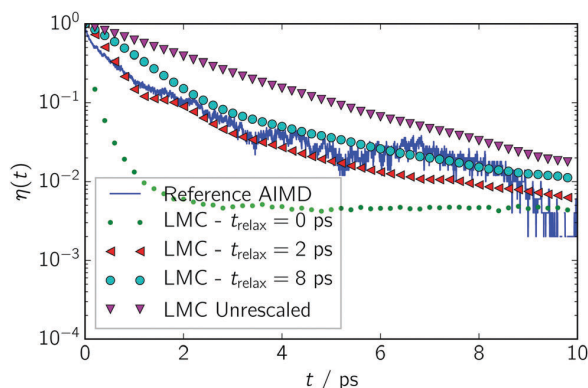
jump at all. In this case, only the oxygen diffusion contributes with  $D = 0.6 \text{\AA}^2 \text{ps}^{-1}$ .

While the concept of the relaxation time parameter  $t_{\text{relax}}$  is physically motivated, the range of 4 to 12 ps obtained by comparison of the final diffusion coefficient to the reference AIMD simulation may appear a quasi-empirical choice. However, this range can also be derived from an analytical consideration: assuming an activation energy of  $2.6 \text{ kcal mol}^{-1}$  for the cleavage and reformation of two hydrogen bonds,<sup>42</sup> and using a wavenumber of  $170 \text{ cm}^{-1}$  for the hydrogen bond stretching frequency,<sup>43</sup> application of the Arrhenius equation yields an average time of around five picoseconds before a hydrogen bond breaks. This quantitative agreement validates the choice of  $t_{\text{relax}} = 5$  ps as a universal value (for liquid water).

**3.3.2 Excess state lifetime.** A complementary parameter for the characterization of the dynamics of the excess proton is the lifetime of an (excess) protonation state of a water molecule. In order to compare our LMC scheme against our reference AIMD simulation, we first define a state function  $i^+(t)$ , which yields the oxygen index of the current  $\text{H}_3\text{O}^+$  ion at time  $t$ . We then define the excess charge autocorrelation function  $\eta$  as

$$\eta(t) = \langle \delta_{i^+(t_0), i^+(t_0+t)} \rangle_{t_0} \quad (11)$$

where  $\delta$  is the Kronecker delta. The excess charge autocorrelation function can be interpreted as the probability that an oxygen atom that was threefold protonated at time  $t_0$  is still threefold protonated at time  $t_0 + t$ .



**Fig. 5** Autocorrelation of the  $\text{H}_3\text{O}^+$  state for different values of  $t_{\text{relax}}$ . The upper triangular line shows the autocorrelation of the cMD/LMC scheme with unrescaled distances as the limiting case for  $t_{\text{relax}} \rightarrow \infty$ . The dotted line shows the autocorrelation for  $t_{\text{relax}} = 0$ . The solid line shows the autocorrelation of the *ab initio* simulation as a reference.

**Fig. 5** shows the results for the reference AIMD and LMC scheme for a range of values for  $t_{\text{relax}}$ . Qualitatively, the results are similar to the diffusion coefficients: the best agreement with the reference *ab initio* trajectory is found for values of  $t_{\text{relax}}$  between 2 and 8 ps. The LMC scheme with unrescaled distances shows the lowest decline of  $\eta$ , caused by the larger oxygen–oxygen distances and the resulting decreased excess charge transfer rates. The LMC scheme with instantaneous relaxation, on the other hand, shows the fastest decline.

## 4 Conclusion

We have investigated proton dynamics in liquid water on extended time scales using a multiscale scheme which combines *ab initio* and classical molecular dynamics with kinetic Monte-Carlo. By virtue of deriving all relevant energetic/kinetic parameters needed for the KMC part from *ab initio* calculations, we are able to reach an accuracy comparable to quantum chemical methods on mesoscopic time scales.

The diffusion constants and the excess state lifetimes obtained from our multiscale Ansatz agree quantitatively with reference AIMD simulations. Our studies show that the dielectrical relaxation of the immediate surroundings of the hydronium ion needs to be explicitly incorporated into the scheme for a realistic description of the excess proton diffusion process. The typical relaxation time in our scheme agrees well with hydrogen bond lifetimes, which we see as a confirmation of the importance of hydrogen bond cleavage as the rate determining step for excess proton transfer.

Our approach is easily applicable to more complex multi-component systems such as Nafion, water channels in proteins, and other nano structured molecular systems.

## Conflicts of interest

There are no conflicts to declare.



## References

- 1 M. Rini, B.-Z. Magnes, E. Pines and E. T. J. Nibbering, *Science*, 2003, **301**, 349–352.
- 2 O. F. Mohammed, D. Pines, J. Dreyer, E. Pines and E. T. J. Nibbering, *Science*, 2005, **310**, 83–86.
- 3 S. Meiboom, *J. Chem. Phys.*, 1961, **34**, 375–388.
- 4 S. Dokmaisrijan and E. Spohr, *J. Mol. Liq.*, 2006, **129**, 92.
- 5 M. K. Petersen and G. A. Voth, *J. Phys. Chem. B*, 2006, **110**, 18594–18600.
- 6 D. Seeliger, C. Hartnig and E. Spohr, *Electrochim. Acta*, 2005, **50**, 4234–4240.
- 7 S. Feng and G. A. Voth, *J. Phys. Chem. B*, 2011, **115**, 5903–5912.
- 8 M. Cappadonia, *Solid State Ionics*, 1995, **77**, 65–69.
- 9 C. J. T. De Grotthuss, *Mémoire sur la décomposition de l'eau et des corps qu'elle tient en dissolution à l'aide de l'électricité galvanique*, 1805.
- 10 G. Bekçioğlu, C. Allolio and D. Sebastiani, *J. Phys. Chem. B*, 2015, **119**, 4053–4060.
- 11 G. Bekçioğlu, C. Allolio, M. Ekimova, E. T. J. Nibbering and D. Sebastiani, *Phys. Chem. Chem. Phys.*, 2014, **16**, 13047–13051.
- 12 N. Agmon, *Chem. Phys. Lett.*, 1995, **244**, 456–462.
- 13 D. Marx, *ChemPhysChem*, 2006, **7**, 1848–1870.
- 14 L. Vilčiauskas, M. E. Tuckerman, J. P. Melchior, G. Bester and K.-D. Kreuer, *Solid State Ionics*, 2013, **252**, 34–39.
- 15 M. Tuckerman, K. Laasonen, M. Sprik and M. Parrinello, *J. Chem. Phys.*, 1995, **103**, 150–161.
- 16 M. Tuckerman, K. Laasonen, M. Sprik and M. Parrinello, *J. Phys. Chem.*, 1995, **99**, 5749–5752.
- 17 M. E. Tuckerman, D. Marx and M. Parrinello, *Nature*, 2002, **417**, 925–929.
- 18 D. Marx, M. E. Tuckerman, J. Hutter and M. Parrinello, *Nature*, 1999, **397**, 601–604.
- 19 P. Intharathap, A. Tongraar and K. Sagarik, *J. Comput. Chem.*, 2006, **27**, 1723–1732.
- 20 G. Kabbe, C. Wehmeyer and D. Sebastiani, *J. Chem. Theory Comput.*, 2014, **10**, 4221–4228.
- 21 G. Kabbe, C. Drefßler and D. Sebastiani, *J. Phys. Chem. C*, 2016, **120**, 19905–19912.
- 22 C. Drefßler, G. Kabbe and D. Sebastiani, *J. Phys. Chem. C*, 2016, **120**, 19913–19922.
- 23 A. Prados, J. J. Brey and B. Sánchez-Rey, *J. Stat. Phys.*, 1997, **89**, 709–734.
- 24 Y. Cao, H. Li and L. Petzold, *J. Chem. Phys.*, 2004, **121**, 4059–4067.
- 25 D. T. Gillespie, *J. Phys. Chem.*, 1977, **81**, 2340–2361.
- 26 G. A. Ludueña, T. D. Kühne and D. Sebastiani, *Chem. Mater.*, 2011, **23**, 1424–1429.
- 27 J. Hutter, M. Iannuzzi, F. Schiffmann and J. VandeVondele, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2014, **4**, 15–25.
- 28 J. VandeVondele, M. Krack, F. Mohamed, M. Parrinello, T. Chassaing and J. Hutter, *Comput. Phys. Commun.*, 2005, **167**, 103–128.
- 29 B. Hess, C. Kutzner, D. van der Spoel and E. Lindahl, *J. Chem. Theory Comput.*, 2008, **4**, 435–447.
- 30 D. van der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark and H. J. C. Berendsen, *J. Comput. Chem.*, 2005, **26**, 1701–1718.
- 31 H. J. C. Berendsen, D. van der Spoel and R. van Drunen, *Comput. Phys. Commun.*, 1995, **91**, 43–56.
- 32 A. D. Becke, *Phys. Rev. A: At., Mol., Opt. Phys.*, 1988, **38**, 3098–3100.
- 33 C. Lee, W. Yang and R. G. Parr, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1988, **37**, 785–789.
- 34 S. Goedecker, M. Teter and J. Hutter, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1996, **54**, 1703–1710.
- 35 S. Grimme, *J. Comput. Chem.*, 2006, **27**, 1787–1799.
- 36 D. J. Evans and B. L. Holian, *J. Chem. Phys.*, 1985, **83**, 4069–4074.
- 37 J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *J. Comput. Chem.*, 2004, **25**, 1157–1174.
- 38 C. Allolio and D. Sebastiani, *Phys. Chem. Chem. Phys.*, 2011, **13**, 16395–16403.
- 39 J. H. Simpson and H. Y. Carr, *Phys. Rev.*, 1958, **111**, 1201–1202.
- 40 J. C. Grossman, E. Schwegler, E. W. Draeger, F. Gygi and G. Galli, *J. Chem. Phys.*, 2004, **120**, 300–311.
- 41 E. Schwegler, J. C. Grossman, F. Gygi and G. Galli, *J. Chem. Phys.*, 2004, **121**, 5400–5409.
- 42 S. Cukierman, *Biophys. J.*, 2000, **78**, 1825–1834.
- 43 W. B. Bosma, L. E. Fried and S. Mukamel, *J. Chem. Phys.*, 1993, **98**, 4413–4421.

### 3.1.5 Article V: Toward Realistic Transfer Rates within the Coupled Molecular Dynamics/Lattice Monte Carlo Approach.

Gabriel Kabbe, Christian Dreßler, and Daniel Sebastiani.

Toward realistic transfer rates within the coupled Molecular Dynamics/Lattice Monte Carlo approach.


*The Journal of Physical Chemistry C*, 120(36):19905–19912, **2016**.

Several aspects of the coupled Molecular Dynamics / Lattice Monte Carlo scheme were refined within this publication. I provided the AIMD trajectories of  $\text{CsH}_2\text{PO}_4$  and benchmarked the influence of the additional cut-off parameter for the hydrogen bond angle on the proton conductivity. G. Kabbe derived proton transfer rates from separate quantum chemical calculations and extended the functional form of the proton jump rates used within the cMD/LMC scheme. D. Sebastiani supervised the project and gave valuable advice during the project work and while G. Kabbe was writing the manuscript.

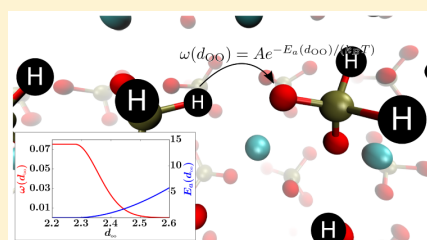
# Toward Realistic Transfer Rates within the Coupled Molecular Dynamics/Lattice Monte Carlo Approach

Gabriel Kabbe, Christian Dreßler, and Daniel Sebastiani\*

Institute of Chemistry, Martin-Luther-University Halle-Wittenberg, von-Danckelmann-Platz 4, 06120 Halle, Germany

 Supporting Information

**ABSTRACT:** We refine our recently developed coupled molecular dynamics/lattice Monte Carlo (cMD/LMC) scheme for the simulation of protonation dynamics in complex hydrogen-bonded solids in view of improving the resulting transport processes. The distance dependency of the proton jump rate between lattice sites and its dependence on additional geometric criteria (bond angles) are derived in a systematic and consistent way. The distance dependency follows an accurate potential energy surface (PES) scan from quantum chemical calculations. The novel geometric criterion takes into account that proton hopping occurs almost exclusively along linear hydrogen bonds. We illustrate the capabilities and the versatility of our scheme on the example of two chemically quite different condensed phase systems: a crystalline solid acid compound and a liquid crystal. Surprisingly, we find that our cMD/LMC scheme yields converged mobility parameters even when based on underlying *ab initio* molecular dynamics (AIMD) trajectories which themselves are not fully converged. Our method yields more accurate values for the mean square displacement, the OH bond autocorrelation function and the proton jump frequencies in agreement with both reference AIMD simulations and experimental values.



## INTRODUCTION

Atomistic simulations play an important role for the determination of dynamical properties and structural features of complex materials. The observation of mesoscale phenomena, however, is often hindered by the limited time scales accessible to these methods, which necessitates the use of sophisticated scale-bridging approaches.

The simulation of proton transport in proton exchange membrane (PEM) fuel cells is a typical example of a phenomenon that needs a high level of structural accuracy over a large time interval in order to sufficiently sample the relevant processes. High level *ab initio* methods such as density functional theory (DFT) can provide the structural accuracy needed for a quantitative description of the transport processes, but are computationally expensive. Nevertheless, proton transport is a phenomenon which has been studied under quantum chemical aspects before.<sup>1–9</sup> Lower level methods on the other hand allow the exploration of time scales several magnitudes above those of DFT, but they often deliver either unsatisfactory results or need additional fine-tuning (e.g., force field parametrization) specifically for this observable.

Our cMD/LMC approach<sup>10</sup> aligns with several methods that have combined molecular dynamics and Monte Carlo like schemes in order to extend the accessible time scales in diverse fields such as solid state physics,<sup>11–15</sup> polymer sciences<sup>16–19</sup> or biophysics.<sup>20–23</sup> The general vision is the simulation of ion transport phenomena on mesoscopic time scales while preserving the structural accuracy delivered by high level methods of theory. For this, we combine AIMD with a Lattice Monte Carlo (LMC) scheme, which is discrete both in time

and space. The AIMD delivers the structural features of the examined system, whereas the LMC algorithm is responsible for the simulation of the ion transport over significantly extended time scales compared to pure AIMD simulations. In this work, we considerably enhance the accuracy of the cMD/LMC scheme by incorporating additional parameters for the structure dependent jump rates that dynamically adapt to the constantly changing molecular structure. Our enhanced scheme is able to predict the response of ionic conductivity to phase transitions and chemical substitutions in organic and inorganic proton conducting compounds.<sup>24</sup>

## METHODS

**Summary of the cMD/LMC Method.** Our cMD/LMC method consists of an LMC-based propagation of protons on lattice sites which are extracted on-the-fly from an underlying molecular dynamics simulation.<sup>10</sup> Lattice Monte Carlo in this context refers to a random walk Monte Carlo method, which comprises a discrete lattice whose sites can be in one of two states: occupied or unoccupied. In contrast to Kinetic Monte Carlo schemes like the residence time algorithm, where the KMC time is incremented by drawing  $\Delta t$  from an exponential distribution, our scheme evolves the occupation states of the lattice in discrete time steps  $\Delta t = \text{const.}$  In our case, the lattice sites correspond to the oxygen atoms of the considered system. Their distances fluctuate during the MD run, and so does the

Received: June 9, 2016

Revised: August 19, 2016

Published: August 22, 2016



corresponding hopping rate within the LMC run. The specific details of the implementation have been reported previously.<sup>10</sup> Here, we review only specific aspects which are relevant for our improvements.

Kinetic Monte Carlo methods often employ predefined reaction rates,<sup>25–27</sup> which are suitable in solid state physics thanks to the fixed atomic structure in a solid state lattice. The higher flexibility of soft matter systems, however, demands a scheme that adapts dynamically to changes in the molecular structure over time. In general, the jump rate of a proton (or any other particle) is determined by the energy barrier, which it needs to overcome, in order to move to another site. The free energy on the other hand, depends on the locations of the remaining atoms, which “form” the potential energy surface, on which the proton moves. Thus, a realistic jump rate between two sites requires (a) sufficient supramolecular sampling to incorporate entropic contributions and (b) an adequate way to determine the energy barriers for a given molecular geometry. In our approach, the entropy contribution is taken into account implicitly by means of the atomic/molecular motion within the MD simulation.

**Determination of Accurate Proton Jump Rates.** The matrix of jump rates between the proton acceptor sites is the most critical input parameter of the cMD/LMC scheme. In our approach, this matrix is repeatedly derived from the instantaneous molecular geometry of the system.

There exist two distinct ways to determine a rate function that links the molecular geometry with the jump rate matrix: Postprocessing of the AIMD trajectory directly results in a distance dependent jump rate.<sup>10</sup> More information on this method can be found in the [Supporting Information](#). The other method, on which we will focus here, obtains the proton jump rates from quantum chemical PES scans of a proton transfer reaction.

The proton transfer from one site to another is the reaction whose activation energy we want to determine. The systems considered in this work contain hydrogen bonds of type P–O–H···O=P. Hence, we have chosen a phosphoric acid dimer to compute the PES via quantum chemical calculations.

Calculating the activation energy for a proton transfer at different oxygen distances allows us to determine the proton transfer activation energy  $E_a^{\text{PES}}(d_{\text{OO}})$  as a function of the O–O distance  $d_{\text{OO}}$ , which in turn allows us to determine a distance-dependent proton jump rate as a function of the oxygen distance and temperature via the Arrhenius law:

$$\omega^{\text{PES}}(d_{\text{OO}}) = A \exp\left(-\frac{E_a^{\text{PES}}(d_{\text{OO}})}{k_{\text{B}}T_{\text{LMC}}}\right) \quad (1)$$

It should be noted that the temperature parameter  $T_{\text{LMC}}$  in eq 1 is an additional temperature, which is formally independent of the temperature at which the MD simulation has been performed. The coexistence of these two temperatures and their impact on protonation dynamics will be discussed further down in the [Results](#).

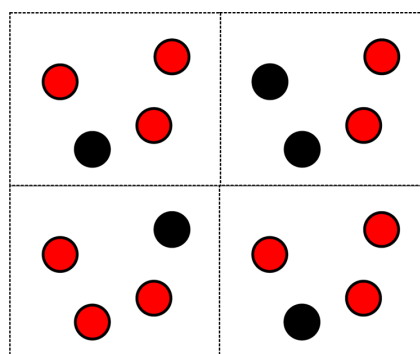
#### Simulation Time and Box Size in the LMC Scheme.

Thanks to the parametrically determined hopping probabilities in the LMC scheme, the cMD/LMC approach reaches a number of proton propagation steps which is several orders of magnitude larger than the corresponding number of MD steps. As our LMC scheme uses an underlying AIMD trajectory to generate the proton jump rates, the cMD/LMC algorithm

needs to sample the jump rates and oxygen positions multiple times from the trajectory.

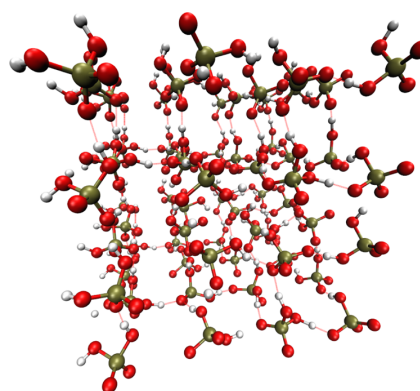
The LMC algorithm updates its jump rates during its run by loading the next frame from the trajectory. As soon as the trajectory has reached its end, it starts from the beginning again. We have deliberately chosen this continuous update scheme, as previous studies showed that it generates the most accurate proton dynamics, compared to other schemes such as random update schemes.<sup>10</sup>

The cMD/LMC algorithm allows the simulation of considerably larger boxes compared to AIMD. In our cMD/LMC scheme, we exploit this by replicating the MD box along its periodic axes. [Figure 1](#) shows a schematic drawing, where the MD box is duplicated along two axes, still allowing different protonation patterns in the four replica.

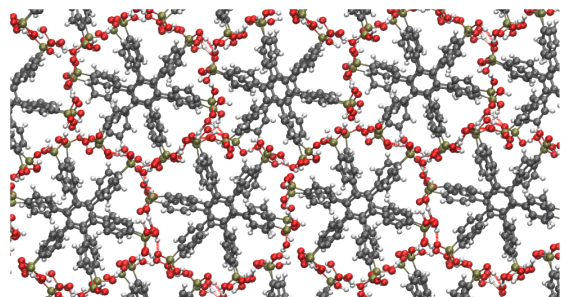


**Figure 1.** Schematic picture of the LMC box enlargement. The oxygens of the AIMD trajectory (red circles) are copied along the periodic axes, whereas the LMC protons (black circles) can move individually in the newly created large box.

**Proton Conducting Compounds as Benchmark Systems.** We have chosen two chemically very different compounds for the cMD/LMC based simulation of proton conduction: a cesium type solid acid  $\text{CsH}_2\text{PO}_4$  (see [Figure 2](#)), and a discotic liquid crystal Hexakis(*p*-phosphonatophenyl)-benzene (*p*-6 PA-HPB) (see [Figure 3](#)), both containing a phosphate/phosphonic acid group. Computational details of



**Figure 2.** Crystal structure of  $\text{CsH}_2\text{PO}_4$ , a solid acid compound, in the high temperature phase. Note the three-dimensional hydrogen bond network.



**Figure 3.** Sketch of the local structure of hexakis(*p*-phosphonatophenyl)benzene (*p*-6 PA-HPB). The disk shaped molecules form columns, which assemble in a hexagonal lattice structure. Hydrogen bonds form between the columns enabling three-dimensional proton transport.

the previously performed simulations are given in<sup>24</sup> and<sup>28</sup> The dimensions of the simulation box and the starting configurations of the systems were obtained from crystal structure data from the literature.<sup>29,30</sup>

## RESULTS

### Jump Rate Determination from Energy Surface Scans.

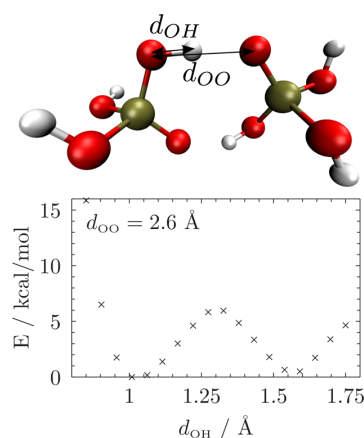
The central input for the MC part of the cMD/LMC scheme is the jump rate  $\omega(d_{OO})$ , which defines the proton hopping frequency between a pair of oxygens, as a function of their distance  $d_{OO}$ . Using the Arrhenius law (eq 1), we can write it in terms of the activation energy  $E_a(d_{OO})$  a proton needs to overcome to move between the two corresponding oxygens atoms.

For both parameters  $A$  and  $E_a$  we will show two independent ways of calculation. The activation energy  $E_a$  can be determined statistically from AIMD trajectories at different temperatures (denoted here as  $E_a^{\text{aimd}}$ ) or from PES scans (denoted as  $E_a^{\text{PES}}$ ). The Arrhenius prefactor  $A$  can be determined both from theoretical considerations (denoted as  $A^{\text{vibrational}}$ ) and from statistical analysis of the jump rates in AIMD simulations (denoted as  $A^{\text{aimd}}$ ).

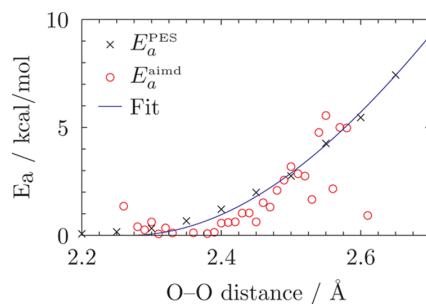
**Determination of the Activation Energy.** We determine the activation energies for a proton transfer between two phosphoric acids for different distances of the donor and acceptor oxygens in a range between 2.2 and 3.5 Å by means of constrained geometry optimizations. The potential energy surface (PES) were determined using the *ab initio* quantum chemistry program Orca<sup>31</sup> with closed shell SCF and the DEF2-TVZP basis set<sup>32,33</sup> with the BLYP functional.<sup>34,35</sup> Fixing the distance  $d_{OO}$  between donor and acceptor, a PES scan is conducted by varying the OH distance  $d_{OH}$ . Figure 4 shows the resulting potential energy surfaces (PES) for  $d_{OO} = 2.6$  Å. We obtain a very low barrier of only 6 kcal/mol, corresponding to about the energy of a hydrogen bond.

We determine the energy barrier the proton has to overcome by calculating the difference between the left PES minimum and the local maximum in the center. This procedure is repeated for values of  $d_{OO}$  in the range between 2.2 and 3.5 Å.

Figure 5 shows the resulting activation energy of a proton transfer within the phosphoric acid dimer as a function of the distance  $d_{OO}$  between donor and acceptor oxygen. The black crosses show the results from the PES scans of the phosphoric acid dimer, whereas the red circles show results from a statistical analysis of the AIMD trajectories (for a detailed



**Figure 4.** Upper picture: Phosphoric acid dimer used for the quantum chemical calculation of the activation energy at different distances between the donor and acceptor oxygens. Lower picture: Potential energy surface of a proton transfer within the phosphoric acid dimer.



**Figure 5.** Calculated activation energies for the proton transfer between phosphonate groups. Black crosses: Results from the energy surface scan of the phosphoric acid dimer (Figure 4). Red circles: Statistical determination from jump rates in an AIMD simulation.<sup>10</sup>

description see Supporting Information and ref 10). Both methods are in good agreement, which we believe is a good indication of their validity.

It turns out that the distance dependence of the activation energy can be fitted with good accuracy to a function of the form

$$E_a(d) = \begin{cases} 0 & d \leq d_0 \\ \frac{a(d - d_0)}{\sqrt{b + \frac{1}{(d - d_0)^2}}} & d > d_0 \end{cases} \quad (2)$$

Note that the function  $E_a(d)$  has a quadratic increase when  $d$  is close to  $d_0$  and becomes linear for large  $d$ .

The resulting fit parameters for  $E_a^{\text{PES}}(d_{OO})$  are listed in Table 1. The functional form of  $E_a^{\text{PES}}(d_{OO})$  is shown in Figure 5 as a solid blue line.

Additionally, we determined the fit parameters for a PES scan of a proton transfer in a sulfur dimer, and between a phosphorus monomer and a sulfur monomer. The results can be found in the Supporting Information.

Finally, we want to make some remarks about the reasoning behind this procedure. For the geometry optimizations, only the two internal coordinates  $d_{OO}$  and  $d_{OH}$  are constrained. As a

**Table 1. Functional Parameters for the Activation Energy Function  $E_a^{\text{PES}}(d_{\text{OO}})$  (See eq 2) for a Phosphoric Acid Dimer**

parameter	value	unit
$a$	$(36 \pm 5)$	kcal/mol/Å <sup>2</sup>
$b$	$(-0.7 \pm 0.5)$	Å <sup>-2</sup>
$d_0$	$(2.22 \pm 0.02)$	Å

consequence, we observe that during the proton scan, a second proton moves from one phosphoric acid to the other, conserving the charge neutrality of the two molecules. We believe that allowing this (inverse) proton transfer is the correct approach because we observe in our MD simulation that the lifetime of a  $P(\text{OH})_3^+$  group (see Supporting Information) is below 70 fs in 90% of all occurrences. Therefore, we conclude that in the vast majority of proton transfers, a protonation event will almost immediately be followed by deprotonation at the same phosphonic group.

Actually, an almost identical PES is obtained in the case of a phosphoric acid dimer with an additional excess proton, which hinders the transfer of a second proton (see Supporting Information).

**Arrhenius Prefactor.** According to eq 1, the determination of the activation energy for a proton jump is not yet sufficient to determine an absolute jump rate; in addition, the Arrhenius prefactor  $A$  needs to be determined. There are two distinct ways: one is a statistical determination from AIMD trajectories, and the other one an estimation from the vibrational frequency of the relevant proton.

The statistical approach yields a jump frequency  $\omega_T^{\text{AIMD}}$  via an explicit counting of hopping events (see Supporting Information and <sup>10</sup>). The resulting rates  $\omega_T^{\text{AIMD}}$  can be used to fit the free parameter  $A$  of  $\omega^{\text{PES}}$  by minimizing the sum of the squared deviations  $\Delta$  between  $\omega^{\text{PES}}$  and  $\omega^{\text{AIMD}}$  with respect to  $A$ :

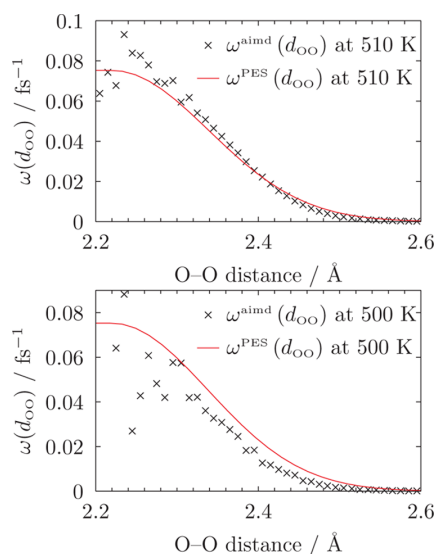
$$\Delta(A) = \sum_{i,T} (\omega_T^{\text{AIMD}}(d_i) - \omega^{\text{PES}}(A, d_i, T))^2 \quad (3)$$

We find the minimum of  $\Delta$  for  $A = 0.0753 \text{ fs}^{-1}$ . This value has been used for all of the following calculations.

The Arrhenius prefactor  $A$  in eq 1 can also be estimated in an alternative way from a simple physical consideration: the Arrhenius equation for a proton hopping event can be interpreted by an escape-attempt frequency and a success probability factor. The former corresponds to the Arrhenius prefactor, while the latter is represented by the Boltzmann factor  $\exp(-\beta E_a)$ . Within this interpretation, it is straightforward to use the vibrational frequency (of the stretch mode) of the mobile proton as the attempt frequency. Hence, we use the experimentally determined frequency of phosphoric acid (2280–2380  $\text{cm}^{-1}$ ),<sup>36</sup> which yields a value  $A^{\text{vibrational}} \approx 0.07 \text{ fs}^{-1}$ .

It is reassuring that the Arrhenius prefactor obtained via the statistical analysis of the AIMD trajectory and the Arrhenius prefactor obtained from the vibrational frequency show such good agreement.

Figure 6 shows a comparison of the jump rates  $\omega^{\text{PES}}(d_{\text{OO}})$  with the AIMD jump rates for  $\text{CsH}_2\text{PO}_4$  at 510 K and  $p$ -6 PA-HPB at 500 K. The statistical errors of  $\omega^{\text{AIMD}}$  are of the order  $\sigma = 0.1 \text{ fs}^{-1}$ , but are not shown in Figure 6 for a better overview (The jump rate comparison including error bars can be found in the Supporting Information).  $\omega^{\text{PES}}$  shows a good agreement with  $\omega^{\text{AIMD}}$  in the range from 2.3 and 2.6 Å. Most O–O distances occur in this range, and therefore the proton jump rates are more converged there. Below 2.3 Å, the sampling



**Figure 6.** Comparison of proton jump rates obtained from AIMD (crosses) and the new jump rates obtained from quantum chemical activation energy calculations. Upper picture: jump rates of  $\text{CsH}_2\text{PO}_4$  at 510 K. Lower picture: jump rates of  $p$ -6 PA-HPB at 500 K.

worsens, however. In this range,  $\omega^{\text{PES}}$  shows a considerably smoother line.

**Angular Dependence of the Jump Rate.** For a realistic description of proton transfer between hydrogen bonds within the LMC scheme, a purely distance-based description of the proton jump rate is not sufficient, as it allows proton jumps for hydrogen bonds of arbitrary geometry. We therefore improve our LMC algorithm by refining the criterion that is used to determine the probability of a proton jump for a given molecular structure. While the purely distance-dependent jump criterion in the LMC scheme already results in a reasonable description of the proton dynamics on longer time scales, it tends to overestimate jump rates between certain oxygen pairs that show little to no proton jumps at all in the AIMD. Therefore, we integrate an angle criterion that filters out fictitious OH bonds that only fulfill the distance criterion.

While the existence of an H-bond can be verified in an MD simulation by geometric examination, this is not possible within the LMC scheme because of the nonatomistic nature of the protons: they are modeled as particles “sitting” on an oxygen, but do not have a real position in space. We therefore use the angle  $\alpha = \angle \text{PO}_1\text{O}_2$  between the phosphorus atom  $P$  next to oxygen  $\text{O}_1$ , which hosts a proton, and  $\text{O}_2$ , which is an unoccupied oxygen atom, from the MD trajectory as an additional geometric parameter for the LMC.

The previous distance-dependent jump rate function is then modified with a Heaviside step function  $\Theta(\alpha - \alpha_0)$ , according to

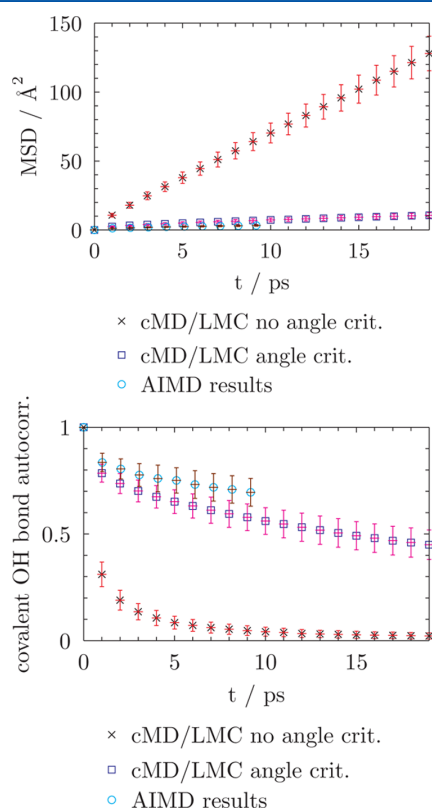
$$\omega(d, \alpha) = \omega(d)\Theta(\alpha - \alpha_0) \quad (4)$$

As angle cutoff parameter, we choose  $\alpha_0 = \pi/2$ , as an analysis of the angle dependency of the proton jump rate showed that nonzero probabilities for proton jumps are mostly found at angles  $\alpha > \pi/2$ . This modification suppresses any proton jumps between oxygen pairs with a geometry that does not allow for a linear hydrogen bond.

We benchmark the results of the newly introduced angle criterion by calculating the mean squared displacement and the covalent OH bond auto correlation function of the compounds  $\text{CsH}_2\text{PO}_4$  and *p*-6 PA-HPB from the cMD/LMC method without angle criterion and the cMD/LMC method with angle criterion, and comparing it with the AIMD results.

For the determination of the LMC diffusion coefficients, we average each run over 100 intervals with length 500000 sweeps ( $\approx 250$  ps MD time).

Figure 7 shows the resulting diffusion parameters of *p*-6 PA-HPB at 600 K. The cMD/LMC algorithm strongly over-

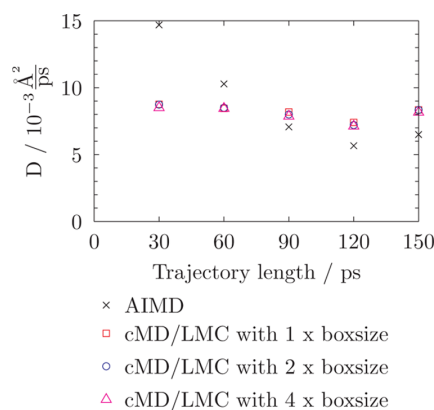


**Figure 7.** Comparison of the proton movement in *p*-6 PA-HPB at 600 K using the cMD/LMC algorithm without an additional angle criterion (crosses), with angle criterion (squares), and the AIMD (circles) as reference. Upper picture: Comparison of the mean squared displacement. Lower picture: Comparison of the covalent O–H bond auto correlation function.

estimates the MSD without the angle criterion. Enabling it reduces the MSD and the resulting slope. Similarly, the covalent OH bond auto correlation function shows a strong decline with the angle criterion turned off, whereas the LMC run with activated angle criterion produces an auto correlation in much better agreement with the AIMD results. Table 2 shows the resulting diffusion coefficients for all compounds and temperatures. All further graphical comparisons can be found in the Supporting Information.

**Convergence Behavior in Length and Time.** On the basis of a trajectory of  $\text{CsH}_2\text{PO}_4$  at 510 K, we investigate the convergence behavior of the cMD/LMC algorithm as well as the AIMD upon variation of the length of the AIMD trajectories, and compare the resulting proton diffusion coefficients of the cMD/LMC scheme with the corresponding AIMD results. Complementary to this, we have analyzed the convergence in terms of simulation box size within the cMD/LMC approach.

Figure 8 shows the results for trajectory lengths from 30 to 150 ps. Starting at an AIMD trajectory length of 30 ps, the



**Figure 8.** Comparison of the diffusion coefficients obtained from AIMD and LMC for different trajectory lengths.

diffusion coefficient of the AIMD (black crosses) is at about  $15 \times 10^{-3} \text{\AA}^2/\text{ps}$ . At this point, the AIMD trajectory has not converged yet, as its diffusion coefficient declines quickly for simulation times up to 150 ps, and approaches a value of about  $6 \text{\AA}^2/\text{ps}$ .

The cMD/LMC method, on the other hand, is much less affected by the length of the underlying AIMD trajectory, showing reasonably converged results when applied on top of an AIMD trajectory of 30 ps total duration. It should be recalled that the cMD/LMC algorithm loops continuously

**Table 2.** Diffusion Coefficients of the Proton Motion in the Two Compounds *p*-6 PA-HPB and  $\text{CsH}_2\text{PO}_4$  for Different Temperatures<sup>a</sup>

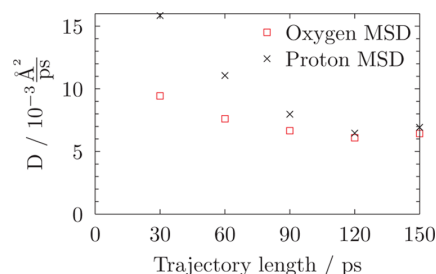
compound	cMD/LMC w/o cutoff	cMD/LMC w/cutoff	AIMD result
<i>p</i> -6 PA-HPB 400 K	$(3.0 \pm 0.2) \times 10^{-1}$	$(4 \pm 1) \times 10^{-3}$	$(1.3 \pm 0.2) \times 10^{-2}$
<i>p</i> -6 PA-HPB 500 K	$(6.1 \pm 0.3) \times 10^{-1}$	$(1.0 \pm 0.2) \times 10^{-2}$	$(1.5 \pm 0.3) \times 10^{-2}$
<i>p</i> -6 PA-HPB 600 K	$(1.05 \pm 0.05)$	$(4.6 \pm 0.4) \times 10^{-2}$	$(3.8 \pm 0.5) \times 10^{-2}$
$\text{CsH}_2\text{PO}_4$ 510 K	$(1.6 \pm 0.2)$	$(9 \pm 2) \times 10^{-3}$	$(5 \pm 3) \times 10^{-3}$
$\text{CsH}_2\text{PO}_4$ 560 K	$(1.9 \pm 0.2)$	$(1.4 \pm 0.3) \times 10^{-2}$	$(1.2 \pm 0.4) \times 10^{-2}$

<sup>a</sup>Middle column: cMD/LMC with angle cutoff. Right column: cMD/LMC without angle cutoff. All values in  $\frac{\text{\AA}^2}{\text{ps}}$ .



through copies of the underlying AIMD trajectory, yielding much larger effective time scales for the protonation dynamics.

We further analyzed the convergence of the diffusion coefficients for the oxygens versus hydrogens, respectively (Figure 9). The data shows that the oxygen diffusivity



**Figure 9.** Comparison of the diffusion coefficients of the proton and oxygen atoms obtained from AIMD.

converges considerably faster (after around 50 ps of AIMD sampling) than that of the protons (after around 100 ps). This illustrates why the cMD/LMC approach is able to yield converged proton diffusivities earlier than a corresponding explicit AIMD simulation. On a more technical level, we observe that the oxygen–oxygen radial distribution function is virtually constant during the entire trajectory (see Figure 34 in the Supporting Information), which in turn results in a practically constant proton hopping rate at the LMC level and thus a very quick convergence of the LMC proton diffusion constant.

Regarding the influence of the simulation box size, Figure 8 shows that the protonation dynamics are already uncorrelated at distances of 10–20 Å.

In summary, the cMD/LMC is able to yield stable values of the diffusion coefficient, even when the diffusive dynamics in the underlying AIMD trajectory have not fully converged yet.

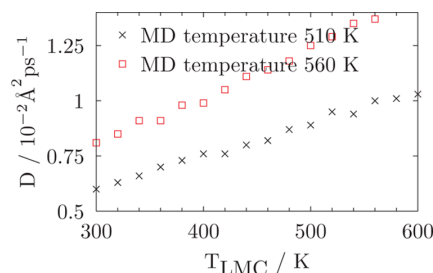
**Relevance of MD versus Lattice Monte Carlo (LMC) Temperature to Proton Conduction.** A special aspect of our cMD/LMC scheme is the coexistence of two formally independent temperatures:  $T_{MD}$  of the underlying AIMD simulation, and  $T_{LMC}$  used in the Arrhenius factor for the LMC proton jump rate. Naturally, both temperatures have their specific influence on the outcome of a cMD/LMC run.

While  $T_{MD}$  determines the temperature of the degrees of freedom associated with the heavy atom structure,  $T_{LMC}$  affects the local hopping dynamics of the protons only.

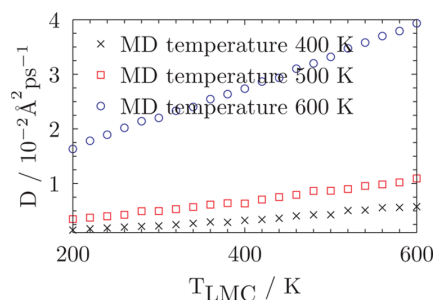
We have computed the proton diffusion coefficients for  $CsH_2PO_4$  and *p*-6 PA-HPB at different values of  $T_{MD}$  and  $T_{LMC}$ . The results are shown in Figure 10 and Figure 11. For each data point, 1000 cMD/LMC runs were conducted over a time of 250 ps and averaged.

For a fixed MD temperature  $T_{MD}$ , the diffusion coefficient shows a linear increase with  $T_{LMC}$ . In all observed cases, an increase of  $T_{MD}$  leads to a higher increase of the diffusion coefficient than an equal increase of  $T_{LMC}$ . This initially appears surprising. Actually, we observe in all cases that the diffusion coefficient is more sensitive to  $T_{MD}$  than it is to  $T_{LMC}$ . This shows that the dynamics of the heavy atom structure play a larger role for the proton transport than the local proton dynamics, at least for all systems considered here.

The simultaneous variation of both  $T_{MD}$  and  $T_{LMC}$  leads to a conventional Arrhenius-like temperature behavior of the

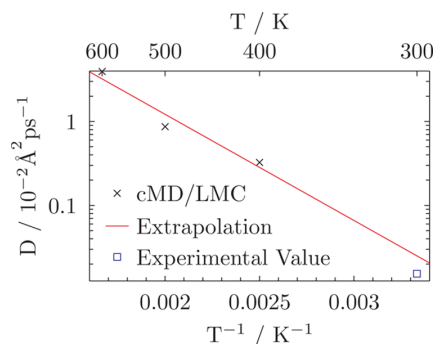


**Figure 10.** Diffusion coefficient of  $CsH_2PO_4$  in the cMD/LMC scheme for varying LMC temperatures.



**Figure 11.** Diffusion coefficient of *p*-6 PA-HPB in the cMD/LMC scheme for varying LMC temperatures.

diffusion coefficient. Experimentally, the diffusion coefficient of dehydrated *p*-6 PA-HPB is available for  $T = 300$  K.<sup>37</sup> Hence, we have extrapolated our cMD/LMC simulations to ambient temperature (see Figure 12), yielding almost quantitatively the corresponding experimental diffusion coefficient.



**Figure 12.** Comparison of the diffusion coefficient of *p*-6 PA-HPB determined via cMD/LMC simulations and extrapolated from experimental results.

## CONCLUSION

In this paper, we have presented an important step in the development of our cMD/LMC algorithm in view of its general applicability to proton conduction in complex disordered condensed phase systems.

The proton jump rate function, which is the most important aspect of the LMC part, has been improved in two aspects: We show, how to derive its distance dependency from *ab initio* calculations on the specific example of a discotic organic and a

crystalline anorganic compound containing phosphonic and phosphoric acid groups. Further, a geometric cutoff criterion for the proton jump rates in the LMC scheme has been added to account for the fact that most proton jumps occur only along linear hydrogen bonds. Benchmarks of the improved cMD/LMC scheme show good agreement with conventional AIMD simulations.

The cMD/LMC scheme has enabled us to decompose the proton diffusion rates from heavy atom motion and (more local) proton hopping phenomena, respectively. In our context, we are able to assign independent characteristic temperatures to these separate degrees of freedom, providing new insight into their competition as rate-determining processes. Our results indicate that for both considered systems, slow modes (i.e., heavy atom motion) are more relevant than the local proton mobility.

We have shown additionally that the cMD/LMC scheme is able to yield a converged diffusion coefficient even if only a short AIMD trajectory is available.

## ■ ASSOCIATED CONTENT

### ● Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: [10.1021/acs.jpcc.6b05821](https://doi.org/10.1021/acs.jpcc.6b05821).

Lifetime of a  $\text{P}(\text{OH})_3$  group, activation energy functions of a  $\text{CsH}_2\text{PO}_4$  dimer and between  $\text{CsHSO}_4$  and  $\text{CsH}_2\text{PO}_4$ , comparison of a proton transfer surface scan in  $\text{CsH}_2\text{PO}_4$  with and without excess protons, free energy calculation from  $E_a^{\text{PES}}(d_{\text{OO}})$ , angular dependence of the proton jump rate—benchmarks, determination of proton jump rates from AIMD, jump rates determined from AIMD and jump rates from the new Arrhenius method in comparison, OH vibration frequencies, oxygen and proton MSD over trajectory length, and comparison of the proton jump frequency in MD and cMD/LMC (PDF)

## ■ AUTHOR INFORMATION

### Corresponding Author

\*(D.S.) E-mail: [daniel.sebastiani@chemie.uni-halle.de](mailto:daniel.sebastiani@chemie.uni-halle.de).

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

The authors gratefully acknowledge the computing time granted by the John von Neumann Institute for Computing (NIC) and provided on the supercomputer JUROPA and JURECA at Jülich Supercomputing Centre (JSC). This work has been supported by the German Research Foundation (DFG) under Grant Se 1008/6-2

## ■ REFERENCES

- (1) Krueger, R. A.; Vilčiauskas, L.; Melchior, J.-P.; Bester, G.; Kreuer, K.-D. Mechanism of Efficient Proton Conduction in Diphosphoric Acid Elucidated via First-Principles Simulation and NMR. *J. Phys. Chem. B* **2015**, *119*, 15866–15875.
- (2) Choe, Y.-K.; Tsuchida, E.; Ikeshoji, T.; Yamakawa, S.; Hyodo, S.-A. Nature of Proton Dynamics in a Polymer Electrolyte Membrane, Nafion: a First-Principles Molecular Dynamics Study. *Phys. Chem. Chem. Phys.* **2009**, *11*, 3892–3899.
- (3) Devanathan, R.; Idupulapati, N.; Baer, M. D.; Mundy, C. J.; Dupuis, M. Ab Initio Molecular Dynamics Simulation of Proton Hopping in a Model Polymer Membrane. *J. Phys. Chem. B* **2013**, *117*, 16522–16529.
- (4) Wood, B. C.; Marzari, N. Proton Dynamics in Superprotonic  $\text{CsHSO}_4$ . *Phys. Rev. B: Condens. Matter Mater. Phys.* **2007**, *76*, 134301.
- (5) Münch, W.; Kreuer, K.-D.; Silvestri, W.; Maier, J.; Seifert, G. The Diffusion Mechanism of an Excess Proton in Imidazole Molecule Chains: First Results of an Ab Initio Molecular Dynamics Study. *Solid State Ionics* **2001**, *145*, 437–443.
- (6) Habenicht, B. F.; Paddison, S. J.; Tuckerman, M. E. Ab Initio Molecular Dynamics Simulations Investigating Proton Transfer in Perfluorosulfonic Acid Functionalized Carbon Nanotubes. *Phys. Chem. Chem. Phys.* **2010**, *12*, 8728–8732.
- (7) Habenicht, B. F.; Paddison, S. J.; Tuckerman, M. E. The Effects of the Hydrophobic Environment on Proton Mobility in Perfluorosulfonic Acid Systems: an Ab Initio Molecular Dynamics Study. *J. Mater. Chem.* **2010**, *20*, 6342.
- (8) Marx, D. Proton Transfer 200 Years After Von Grothuss: Insights from Ab Initio Simulations. *ChemPhysChem* **2006**, *7*, 1848–1870.
- (9) Vilčiauskas, L.; Tuckerman, M. E.; Bester, G.; Paddison, S. J.; Kreuer, K.-D. The Mechanism of Proton Conduction in Phosphoric Acid. *Nat. Chem.* **2012**, *4*, 461–466.
- (10) Kabbe, G.; Wehmeyer, C.; Sebastiani, D. A Coupled Molecular Dynamics/Kinetic Monte Carlo Approach for Protonation Dynamics in Extended Systems. *J. Chem. Theory Comput.* **2014**, *10*, 4221–4228.
- (11) Hiwatari, Y.; Kaneko, Y.; Mikami, T.; Ohara, K.; Asa, F. Molecular Dynamics–Monte Carlo Hybrid Simulation of Thin Film Growth and Void Formation in Electrodeposition Process. *Mol. Simul.* **2007**, *33*, 133–138.
- (12) Betz, G.; Husinsky, W. A Combined Molecular Dynamics and Kinetic Monte Carlo Calculation to Study Sputter Erosion and Beam Assisted Deposition. *Nucl. Instrum. Methods Phys. Res., Sect. B* **2002**, *193*, 352–358.
- (13) Knizhnik, A. A.; Bagaturyants, A. A.; Belov, I. V.; Potapkin, B. V.; Korokin, A. A. An Integrated Kinetic Monte Carlo Molecular Dynamics Approach for Film Growth Modeling and Simulation:  $\text{ZrO}_2$  Deposition on  $\text{Si}(100)$  Surface. *Comput. Mater. Sci.* **2002**, *24*, 128–132.
- (14) Kaneko, Y.; Mikami, T.; Hiwatari, Y.; Ohara, K. Computer Simulation of Electrodeposition: Hybrid of Molecular Dynamics and Monte Carlo. *Mol. Simul.* **2005**, *31*, 429–433.
- (15) Ghoufi, A.; Maurin, G. Hybrid Monte Carlo Simulations Combined with a Phase Mixture Model to Predict the Structural Transitions of a Porous Metal–Organic Framework Material upon Adsorption of Guest Molecules. *J. Phys. Chem. C* **2010**, *114*, 6496–6502.
- (16) Forrest, B. M.; Suter, U. W. Hybrid Monte Carlo Simulations of Dense Polymer Systems. *J. Chem. Phys.* **1994**, *101*, 2616–2629.
- (17) Gromov, D. G.; de Pablo, J. J. Structure of Binary Polymer Blends: Multiple Time Step Hybrid Monte Carlo Simulations and Self-Consistent Integral-Equation Theory. *J. Chem. Phys.* **1995**, *103*, 8247–8256.
- (18) Irbäck, A. Hybrid Monte Carlo Simulation of Polymer Chains. *J. Chem. Phys.* **1994**, *101*, 1661–1667.
- (19) Heermann, D. W.; Yixue, L. A Global-Update Simulation Method for Polymer Systems. *Makromol. Chem., Theory Simul.* **1993**, *2*, 299–308.
- (20) Hansmann, U. H. E.; Okamoto, Y. New Monte Carlo Algorithms for Protein Folding. *Curr. Opin. Struct. Biol.* **1999**, *9*, 177–183.
- (21) Zhang, H. A New Hybrid Monte Carlo Algorithm for Protein Potential Function Test and Structure Refinement. *Proteins: Struct., Funct., Genet.* **1999**, *34*, 464–471.
- (22) Peter, E. K.; Shea, J.-E. A Hybrid MD-KMC Algorithm for Folding Proteins in Explicit Solvent. *Phys. Chem. Chem. Phys.* **2014**, *16*, 6430–6440.
- (23) Yang, J.; Monine, M. I.; Faeder, J. R.; Hlavacek, W. S. Kinetic Monte Carlo Method for Rule-Based Modeling of Biochemical Networks. *Phys. Rev. E Stat Nonlin Soft Matter Phys.* **2008**, *78*, 031910.

(24) Dreßler, C.; Kabbe, G.; Sebastiani, D. Proton Conductivity in Hydrogen-Phosphate/-Sulfates from a Coupled Molecular Dynamics/Lattice Monte Carlo (cMD/LMC) Approach. *J. Phys. Chem. C* **2016**, DOI: 10.1021/acs.jpcc.6b05822.

(25) Battaile, C. C.; Srolovitz, D. J.; Butler, J. E. A Kinetic Monte Carlo Method for the Atomic-Scale Simulation of Chemical Vapor Deposition: Application to Diamond. *J. Appl. Phys.* **1997**, *82*, 6293–6300.

(26) Zhang, P.; Zheng, X.; Wu, S.; Liu, J.; He, D. Kinetic Monte Carlo Simulation of Cu Thin Film Growth. *Vacuum* **2004**, *72*, 405–410.

(27) Adams, J.; Wang, Z.; Li, Y. Modeling Cu Thin Film Growth. *Thin Solid Films* **2000**, *365*, 201–210.

(28) Wehmeyer, C.; Schrader, M.; Andrienko, D.; Sebastiani, D. Water-Free Proton Conduction in Hexakis(p-Phosphonatophenyl)-benzene Nanochannels. *J. Phys. Chem. C* **2013**, *117*, 12366–12372.

(29) Bronowska, W. Comment on “Does the Structural Superionic Phase Transition At 231 °C in CsH<sub>2</sub>PO<sub>4</sub> Really Not Exist?” [*J. Chem. Phys.* **110**, 4847 (1999)]. *J. Chem. Phys.* **2001**, *114*, 611–612.

(30) Preisinger, A.; Mereiter, K.; Bronowska, W. The Phase Transition of CsH<sub>2</sub>PO<sub>4</sub> (CDP) At 505 K. *Mater. Sci. Forum* **1994**, *166–169*, 511–516.

(31) Neese, F. The ORCA Program System. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012**, *2*, 73–78.

(32) Schäfer, A.; Horn, H.; Ahlrichs, R. Fully Optimized Contracted Gaussian Basis Sets for Atoms Li to Kr. *J. Chem. Phys.* **1992**, *97*, 2571–2577.

(33) Weigend, F.; Ahlrichs, R. Balanced Basis Sets of Split Valence, Triple Zeta Valence and Quadruple Zeta Valence Quality for H to Rn: Design and Assessment of Accuracy. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297–3305.

(34) Becke, A. D. Density-Functional Exchange-Energy Approximation with Correct Asymptotic Behavior. *Phys. Rev. A: At., Mol., Opt. Phys.* **1988**, *38*, 3098–3100.

(35) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula Into a Functional of the Electron Density. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1988**, *37*, 785–789.

(36) Chapman, A. C.; Thirlwell, L. E. Spectra of Phosphorus Compounds—I the Infra-Red Spectra of Orthophosphates. *Spectrochim. Acta* **1964**, *20*, 937–947.

(37) Jiménez-García, L.; Kaltbeitzel, A.; Pisula, W.; Gutmann, J. S.; Klapper, M.; Müllen, K. Phosphonated Hexaphenylbenzene: a Crystalline Proton Conductor. *Angew. Chem., Int. Ed.* **2009**, *48*, 9951–9953.

## 3.2 Efficient Representations of the Static LDDRF

A central aim of this thesis is the derivation and investigation of efficient representations of the linear density-density response function (LDDRF). The LDDRF is given in equation (3.2) and links arbitrary perturbing potentials  $V_{\text{pert}}$  to molecular response densities  $n^{\text{resp}}$  via a linear integral transform. We describe this transformation using the linear operator  $\hat{\mathbf{T}} : V_{\text{pert}} \mapsto n^{\text{resp}}$ .

$$\hat{\mathbf{T}} : V_{\text{pert}} \mapsto n^{\text{resp}} \quad (3.1)$$

$$\text{with } n^{\text{resp}}(\mathbf{r}) = \int \chi(\mathbf{r}, \mathbf{r}') V_{\text{pert}}(\mathbf{r}') d^3r' \quad (3.2)$$

In section 2.3 of the theory part of this thesis, a detailed derivation of equation (3.2) and a discussion of the properties of the linear operator  $\hat{\mathbf{T}}$  are presented. It was shown that the LDDRF essentially possesses an infinite-dimensional eigensystem representation according to

$$\hat{\mathbf{T}} = \sum_{i=1}^{\infty} |\chi_i\rangle \lambda_i \langle \chi_i|, \quad (3.3)$$

with eigenvalues  $\lambda_i$  for the eigenfunctions  $\chi_i(\mathbf{r})$ .

The starting point of the investigations in this thesis is a much more efficient representation of the LDDRF that was published in 2016 by Scherrer and Sebastiani.<sup>144</sup> the moment expansion.

The assumption that a significant part of the perturbing potential (caused by neighboring molecules) can be expanded at the responding molecule using only a few basis functions  $\{|P_1\rangle, |P_2\rangle, \dots, |P_N\rangle\}$  (e.g. in terms of a multipole or Taylor expansion) is a crucial prerequisite of the moment expansion. Thus, we assume for the perturbing potential *at* the responding molecule:

$$V_{\text{pert}}(\mathbf{r}) \approx \sum_{n=1}^N c_n P_n(\mathbf{r}). \quad (3.4)$$

Article [Dreßler et al., *J. Comput. Chem.*, **2019**.] in this thesis demonstrates that this assumption is a very good approximation for the specific case of water. It is demonstrated there that the perturbing potential arising from a neighboring water molecule at different chemically relevant distances can be described with high accuracy using 34 monomials or Racah-normalized regular solid harmonic functions. From here on,  $\mathcal{P}[\mathbf{N}] := \text{Span}(P_1, P_2, \dots, P_N)$  denotes the vector space spanned by  $\{|P_1\rangle, |P_2\rangle, \dots, |P_N\rangle\}$ .

Scherrer and Sebastiani converted the eigenstates  $\{\chi_i \mid i \in \mathbb{N}\}$  into a new set of states  $\{\xi_i \mid i \in \mathbb{N}\}$  by using unitary transformations. We refer to these functions  $\xi_i$  as moment-expanded states. These states fulfill a partial orthogonality relation with respect to the basis functions  $\{|P_1\rangle, |P_2\rangle, \dots\}$  of the expansion of the perturbing potential:

$$\langle \xi_i, P_n \rangle = 0 \quad \forall i > n \quad (3.5)$$

Due to the partial orthogonality relation, only  $N$  moment-expanded states  $\{|\xi_1\rangle, |\xi_2\rangle, \dots, |\xi_N\rangle\}$  have a non-zero overlap with the  $N$ -th basis function  $|P_N\rangle$  of the perturbing potential. The full density response due to perturbing potentials  $V_{\text{pert}} \in \mathcal{P}[\mathbf{N}]$  is reduced to a



sum over  $N$  terms:

$$\hat{\mathbf{T}}|V_{\text{pert}}\rangle = \sum_{i=1}^N |\xi_i\rangle \langle \xi_i, V_{\text{pert}}\rangle. \quad (3.6)$$

In article [Dreßler et al., *J. Comput. Chem.*, **2019**.], a thorough mathematical derivation of the moment expansion and a careful investigation of its properties is presented. The moment expansion of the LDDRF using  $N$  moment-expanded states is equivalent to a restriction  $\hat{\mathbf{T}}|_{\mathcal{P}[\mathbf{N}]}$  of the operator  $\hat{\mathbf{T}}$  to the vector space  $\mathcal{P}[\mathbf{N}] := \text{span}(P_1, P_2, \dots, P_N)$ . The article [Dreßler et al., *J. Comput. Chem.*, **2019**.] starts from the (infinite) matrix representation  $\mathbf{T}_{\{\mathbf{P}_k\}}^{\{\sqrt{\lambda_k}\chi_k\}}$  of the linear operator  $\hat{\mathbf{T}}$  with respect to the  $\{P_n \mid n \in \mathbb{N}\}$  as the basis of the domain and the eigenvalue-weighted eigenfunctions  $\{\sqrt{\lambda_i}\chi_i \mid i \in \mathbb{N}\}$  of  $\hat{\mathbf{T}}$  as the basis of the image. Utilizing this matrix formalism, it is proved that the moment expansion can be interpreted as a QR decomposition of the (infinite) matrix representation  $\mathbf{T}_{\{\mathbf{P}_k\}}^{\{\sqrt{\lambda_k}\chi_k\}}$  of  $\hat{\mathbf{T}}$ :

$$\mathbf{T}_{\{\mathbf{P}_k\}}^{\{\sqrt{\lambda_k}\chi_k\}} = \mathbf{Q}\mathbf{R}. \quad (3.7)$$

The orthogonal matrix  $\mathbf{Q}$  is the change-of-basis matrix that describes the unitary transformation of the eigenfunctions  $\{\chi_i \mid i \in \mathbb{N}\}$  to the moment-expanded states  $\{\xi_i \mid i \in \mathbb{N}\}$ :

$$\xi_i = \sum_{l=1}^{\infty} \mathbf{Q}_{li} \sqrt{\lambda_l} \chi_l, \quad (3.8)$$

The elements  $\mathbf{R}_{ij}$  of the matrix  $\mathbf{R}$  are obtained as the overlap integrals  $\mathbf{R}_{ij} = \langle \xi_i, P_j \rangle$  of the moment-expanded states  $\xi_i$  and the basis functions of the perturbing potential  $P_j$ . Taking into account that  $\mathbf{R}$  is an upper triangular matrix, we can directly deduce the partial orthogonality condition from equation 3.5.

The description of the moment expansion in terms of an (infinite) QR decomposition includes the initially published algorithm of the moment expansion, which is based on the calculation and transformation of thousands of eigenstates.<sup>144</sup>

In the article [Ahlert et al., *Eur. Phys. J. B*, **2018**.], another algorithm based on the initial idea of Scherrer is published that allows for the calculation of the first  $N$  moment-expanded states within only  $N$  self-consistent DFPT calculations, instead of several thousands to reach the same level of accuracy. We refer to this algorithm, which enables the calculation of the moment expansion itself with a significantly reduced computationally effort, as direct moment expansion. A more sophisticated understanding of the algorithm of the direct moment expansion is provided in the article [Dreßler et al., *J. Comput. Chem.*, **2019**.] in this thesis. It is proved here that the first  $N$  basis functions of the expansion of the perturbing potential  $\{|P_1\rangle, |P_2\rangle, \dots, |P_N\rangle\}$  and their explicitly calculated associated density responses  $\{|\tilde{n}_1\rangle := \hat{\mathbf{T}}|P_1\rangle, |\tilde{n}_2\rangle := \hat{\mathbf{T}}|P_2\rangle, \dots, |\tilde{n}_N\rangle := \hat{\mathbf{T}}|P_N\rangle\}$  are related via the overlap matrix  $\mathbf{R}$  (which was calculated in terms of the QR decomposition). It is also proved that the matrix  $\mathbf{R}$  can also be obtained by a Cholesky decomposition of a matrix  $\Theta$ :

$$\Theta = (\mathbf{R}^T \mathbf{R}), \quad (3.9)$$

where  $\Theta$  is a (small)  $N \times N$  matrix with elements  $\Theta_{ij} = \langle P_i, \tilde{n}_j \rangle$ .

Following this line of approach, we can formulate an efficient protocol for the calculation of the first  $N$  moment-expanded states  $\{|\xi_1\rangle, |\xi_2\rangle, \dots, |\xi_N\rangle\}$ :

- calculate  $\{|\hat{\mathbf{T}}P_1\rangle, |\hat{\mathbf{T}}P_2\rangle, \dots, |\hat{\mathbf{T}}P_N\rangle\}$  via  $N$  explicit self-consistent solutions of a Sternheimer-like equation using DFPT (cf. section 2.2.3).  

$$\underbrace{|\hat{\mathbf{T}}P_1\rangle}_{|\tilde{n}_1\rangle}, \underbrace{|\hat{\mathbf{T}}P_2\rangle}_{|\tilde{n}_2\rangle}, \dots, \underbrace{|\hat{\mathbf{T}}P_N\rangle}_{|\tilde{n}_N\rangle}$$
- calculate the  $N \times N$  matrix  $\Theta$  with elements  $\Theta_{ij} = \langle P_i, \tilde{n}_j \rangle$
- calculate  $\mathbf{R}$  using a Cholesky decomposition  $\Theta = \mathbf{R}^T \mathbf{R}$  (using library algorithms, e.g. lapack)
- the inverse matrix  $\mathbf{R}^{-1}$  transforms the set  $\{\tilde{n}_i \mid i \leq N\}$  into the moment-expanded states  $\{\xi_i \mid i \leq N\}$  via

$$|\xi_i\rangle = \sum_{j=1}^i |\tilde{n}_j\rangle (\mathbf{R}^{-1})_{ji}. \quad (3.10)$$

In conclusion, the numerical determination of the first  $N$  moment-expanded states requires of the order of  $N$  applications of  $\hat{\mathbf{T}}$  (and linear algebra operations with negligible effort). The matrix  $\mathbf{R}$  is an upper triangular matrix, which ensures that all equations used in the protocol for the determination of the first  $N$  moment-expanded states are exact and can be truncated for every  $N \in \mathbb{N}$  without the loss of any information. The description of the moment expansion and the direct moment expansion in terms of a QR decomposition and a Cholesky decomposition, respectively, puts these approaches on a solid mathematical foundation and allows for their generalization to other self-adjoint, compact and positive linear operators.

The derivation of another efficient representation of the LDDRF is motivated by a comparison of the domain and the image of the restriction  $\mathcal{T} := \hat{\mathbf{T}}|_{\mathcal{P}[N]}$  of the linear operator  $\hat{\mathbf{T}}$  to  $\mathcal{P}[N]$ . It is immediately apparent that perturbing potentials and molecular response densities are quite different functions. While molecular response densities are highly oscillating functions which resemble linear combinations of atomic basis functions, the perturbing potential arising from a non-overlapping charge density can be expressed efficiently by smooth functions, e.g. by a Taylor expansion using monomials. The eigensystem representation resolves this mismatch of the domain and the image of  $\mathcal{T}$  by means of a large number of basis functions (eigenfunctions) resulting in a high-dimensional matrix representation. In contrast, the dimensionality of the moment expansion in equation (3.6) is significantly reduced, but molecular response densities and perturbing potentials have to be expressed by a small number of  $N$  moment-expanded states. This contradiction is resolved by the fact that the moment-expanded states only form a basis for the response densities. According to equation (3.6), only the overlap of the non-orthogonal moment-expanded states and the perturbing potential is taken into account for the calculation of the response densities. The overlap of two functions can be non-zero even if the functions are from distinct vector spaces with trivial intersection. In the article [Drefler, Sebastiani, *Int. J. Quantum Chem.*, 2020.], it is proved that the expression of perturbing potentials within a basis of the moment-expanded states does not retain the functions. However, these “wrong” perturbing potentials expanded in moment-expanded states can still be inserted as  $|V_{\text{pert}}\rangle$  into equation (3.6) without any loss of information concerning the molecular responses densities. While the expression of the perturbing potentials in the moment-expanded states will alter the function itself, it retains the overlap with the moment-expanded states which corresponds to unchanged response densities.

Starting from this idea, it is straightforward to create a new representation with even more desirable properties compared to the moment expansion. By diagonalization of the overlap matrix of the moment-expanded states, it is possible to obtain a new representation with orthonormal eigenstates and eigenvalues that possesses the same dimensionality and contains the same information for the calculation of response densities compared to the moment expansion.

$$\hat{\mathbf{T}}|_{\mathcal{P}[\mathbf{N}]} = \sum_{n=1}^N |\tilde{\chi}_n\rangle \tilde{\lambda}_n \langle \tilde{\chi}_n|. \quad (3.11)$$

We refer to this new representation obtained from the reconstructed eigenstates and eigenvalues as the reduced eigensystem representation. The convergence behavior of this reduced eigensystem representation is now apparent from the decay of the eigenvalues because the eigenstates are orthonormal. In the article [Dreßler, Sebastiani, *Int. J. Quantum Chem.*, **2020**.], it is demonstrated that the eigenvalue spectrum decays very fast for small molecules, which in turn allows for a further reduction of the dimensionality of this representation by neglecting the eigenfunctions corresponding to small eigenvalues.

The efficient representation of the LDDRF is particularly useful for the repeated calculation of molecular density responses due to varying perturbing potentials. As a result, molecular dynamics simulation is an ideal candidate for the application of the efficient representations of the LDDRF. In molecular dynamics, the electron density of a molecule is perturbed by the potential of the charge density from neighboring molecules. Due to the fluctuation of the atomic coordinates, the perturbing potentials are constantly changing for subsequent MD steps. In the article [Scherrer et al., *J. Chem. Phys.*, **2016**.], the geometry dependence of the moment-expanded states is investigated by means of a Taylor expansion. Employing the efficient representation of the LDDRF, it is possible to calculate polarizabilities of water molecules by post processing of an MD trajectory. The Raman spectra are accessible by means of autocorrelation and Fourier transformation of these polarizabilities. These spectra were compared to Raman spectra obtained from explicit AIMD simulations for the validation of our approach.

### 3.2.1 Article VI: Reduced Eigensystem Representation of the LDDRF

Christian Dreßler and Daniel Sebastiani

Reduced eigensystem representation of the linear density-density response function. *International Journal of Quantum Chemistry*, 120(3):e26085, **2020**.

In this article, I derived and implemented the theory. I prepared the manuscript. D. Sebastiani supervised the project and provided me with valuable advice during the derivation of the equations and while I was writing the manuscript.

# Reduced eigensystem representation of the linear density-density response function

Christian Dreßler  | Daniel Sebastiani

Institute of Chemistry, Martin-Luther-University Halle-Wittenberg, Halle, Germany

## Correspondence

Daniel Sebastiani, Institute of Chemistry, Martin-Luther-University Halle-Wittenberg, Von-Danckelmann-Platz 4, 06120 Halle (Saale), Germany.  
Email: daniel.sebastiani@chemie.uni-halle.de

## Abstract

The linear density-density response function represents a formulation of the generalized density response of a molecular (or extended) system to arbitrary perturbing potentials. We have recently established an approach for reducing the dimension of the (in principle infinite) eigenspace representation (the moment expansion) and generalized it to arbitrary self-adjoint, positive-definite, and compact linear operators. Here, we present a modified representation—the reduced eigensystem representation—which allows to define a trivial criterion for the convergence of the approximation to the density response. By means of this novel eigensystem-like structure, the remarkable reduction of the dimensionality becomes apparent for the calculation of the density-density response function.

## KEYWORDS

density-density response function, density functional perturbation theory, eigensystem representation, molecular interaction

## 1 | INTRODUCTION

For the description of large systems, quantum mechanical electronic structure methods suffer from the “curse of dimensionality,” because their scaling behavior is far from linear. The decomposition of large systems into interacting smaller subsystems is an obvious attempt in order to solve this problem. Following these ideas, many embedding or subsystem density functional theory methods<sup>[1]</sup> like frozen density embedding<sup>[2–8]</sup> or embedded correlated wave functions methods<sup>[8–11]</sup> were developed in the literature.

For the specific case of interacting molecules, a fragmentation approach which considers the individual molecules as elementary subsystems is suited. Then, the intermolecular interaction energy has to be obtained via post processing of the results for the isolated molecules. In particular for the electrostatic interaction energy, methods like multipole expansions,<sup>[12–18]</sup> density fitting,<sup>[19–22]</sup> or perturbation theories<sup>[23–28]</sup> are utilized.

Accurate electrostatic interaction energies should take polarization effects into account. The linear density-density response function from Equation (2) offers a tool to calculate the fully self-consistent density response of a molecule to arbitrary potentials (in practice: the electrostatic field arising from nearby molecules), corresponding to the exact polarization at all multipolar orders.<sup>[29–32]</sup>

$$\hat{T}: V_{\text{pert}} \mapsto n^{\text{resp}} \quad (1)$$

$$\text{with } n^{\text{resp}}(\mathbf{r}) = \int \chi(\mathbf{r}, \mathbf{r}') V_{\text{pert}}(\mathbf{r}') d^3r' \quad (2)$$

A less formal access to the linear density-density response function is possible by comparison of the unperturbed and perturbed Schrödinger equation. For an isolated molecule the Schrödinger equation is given by:

-----  
This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2019 The Authors. *International Journal of Quantum Chemistry* published by Wiley Periodicals, Inc.

$$\hat{H}|\Psi\rangle = E|\Psi\rangle \quad (3)$$

$$n(\mathbf{r}) = |\Psi(\mathbf{r})|^2 \quad (4)$$

Subsequently, we have for an additional external potential:

$$(\hat{H} + \hat{V}_{\text{pert}})|\Psi^{\text{polarized}}\rangle = E^{\text{polarized}}|\Psi^{\text{polarized}}\rangle \quad (5)$$

$$n^{\text{polarized}}(\mathbf{r}) = |\Psi^{\text{polarized}}(\mathbf{r})|^2 \quad (6)$$

$$n^{\text{resp}}(\mathbf{r}) = n^{\text{polarized}}(\mathbf{r}) - n(\mathbf{r}) \quad (7)$$

On the one hand, Equation 2 is the shorthand notation for applying this protocol (Equations 3-7). On the other hand it enables a possibility to drastically speed up calculations of the response density: Once the electronic susceptibility  $\chi(\mathbf{r}, \mathbf{r}')$  is known, the computation of the density response breaks down to integration according to Equation (2).

Following this line, the density-density response function can be seen as a formalism to express the specific response to any specific perturbing potential in an abstract way. Due to its non-locality, any straightforward representation of  $\chi(\mathbf{r}, \mathbf{r}')$  on a grid would require excessive storage dimensions.

In this publication, we aim toward a novel representations (reduced eigensystem  $\tilde{\chi}_i$ ) of the regular density-density response function. Before we start to derive the reduced eigensystem representation, we have to discuss the properties of the regular eigensystem decomposition using the regular eigenvalues  $\lambda_i$  and eigenstates  $\chi_i$ :

## 2 | GENERAL PROPERTIES OF THE LINEAR DENSITY-DENSITY RESPONSE FUNCTION

For  $X \subseteq \mathbb{R}^3$  and  $\chi(\mathbf{r}, \mathbf{r}') \in L^2[X \times X]$  we have  $\hat{T}: L^2[X] \rightarrow L^2[X]$  and Equation (2) defines a (compact) Hilbert-Schmidt operator.<sup>[33]</sup> The operator  $\hat{T}$  is self-adjoint due to the symmetry of the kernel:  $\chi(\mathbf{r}, \mathbf{r}') = \chi(\mathbf{r}', \mathbf{r}) \forall \mathbf{r}, \mathbf{r}' \in X$ .<sup>[29,30]</sup>

Furthermore, the kernel  $\chi(\mathbf{r}, \mathbf{r}')$  is positive definite because the energy due to the polarization of a molecule is always positive. Hence the following condition is fulfilled:

$$\iint \chi(\mathbf{r}, \mathbf{r}') V_{\text{pert}}(\mathbf{r}') V_{\text{pert}}(\mathbf{r}) d^3r d^3r' \geq 0 \quad \forall V_{\text{pert}}(\mathbf{r}) \in L^2[X] \quad (8)$$

and the operator  $\hat{T}$  is positive-definite. According to the spectral theorem for self-adjoint and compact operators, there exists an eigensystem representation of the operator  $\hat{T}$ :

$$n^{\text{resp}}(\mathbf{r}) = \hat{T}[V_{\text{pert}}(\mathbf{r})] = \sum_{i=1}^{\infty} \chi_i(\mathbf{r}) \lambda_i \langle \chi_i, V_{\text{pert}} \rangle \quad (9)$$

with eigenvalue  $\lambda_i$  to the eigenfunction  $\chi_i(\mathbf{r})$ . The function  $\chi(\mathbf{r}, \mathbf{r}')$  can be written as follows:

$$\chi(\mathbf{r}, \mathbf{r}') = \sum_{i=1}^{\infty} \chi_i(\mathbf{r}) \lambda_i \chi_i(\mathbf{r}') \quad (10)$$

The eigensystem can be computed using an iterative Lanczos diagonalization technique for the Operator  $\hat{T}$  within density functional perturbation theory (DFPT).<sup>[32,34-39]</sup> For a detailed description, see References 29 and 30.

Since the sequence of eigenvalues converges to zero ( $\lim_{i \rightarrow \infty} \lambda_i = 0$ ), the eigensystem representation can be truncated according to the desired numerical precision at a finite index.<sup>[29-31,40-46]</sup> In particular, the sequence of eigenvalues is square summable ( $\sum_{i=1}^{\infty} (\lambda_i)^2 < \infty$ ), since the Operator  $\hat{T}$  belongs to the second shadow class as a Hilbert-Schmidt operator.<sup>[33]</sup> This ensures the fast decay of the eigenvalues. Nevertheless, thousands of scalar products with the eigenstates have to be determined in Equation (9) for each perturbing potential, which turns the eigensystem into an inefficient representation.

The spatial origin of the perturbation in Equation (9) is determined by the geometry of the specific system. But independent of the origin of the perturbation, it is often possible to expand the perturbing potential at the position of the responding molecule (as opposed to the location of the perturbing object, cf. Figure 1). In particular a multipole expansion is suited for this because it can often be truncated, meaning that only the first few terms need to be employed for a good approximation to the regular function. We can employ Racah normalized solid harmonics  $\{\tilde{R}_l^m \mid l \in \mathbb{N}, m \in \{-l, -l+1, \dots, l-1, l\}\}$ <sup>[47]</sup>

$$V_{\text{pert}}(\mathbf{r}) = \sum_{l=0}^{\infty} \sum_{m=-l}^l \langle \tilde{R}_l^m, V_{\text{pert}} \rangle \tilde{R}_l^m(\mathbf{r}) = \sum_{l=0}^{\infty} \sum_{m=-l}^l V_l^m \tilde{R}_l^m(\mathbf{r}). \quad (11)$$

for such an expansion. Note that the two indices  $l$  and  $m$  of the  $\tilde{R}_l^m$  can be converted into a single index  $n$  by the unique assignment  $n = (l^2 + l) + m$ . We define the space spanned by the first  $N$  ( $= L^2 + 2L$ ) Racah normalized solid harmonics as  $\mathcal{R}[\mathbf{N}] := \text{span}\left(\{\tilde{R}_l^m \mid l \leq L \in \mathbb{N}, m \in \{-l, \dots, l\}\}\right) \cong \mathbb{R}^N$ . In a recent article, we demonstrated that the perturbing potential  $|V_{\text{pert}}|$  (originated from a neighboring molecule) at a responding molecule can be approximated using only 35 Racah normalized solid harmonics as basis functions ( $|V_{\text{pert}}| \in \mathcal{R}[\mathbf{N}]$  with  $N = 35$ ).

With respect to this basis of the perturbing potential, Scherrer presented in 2016 an algorithm for a transformation of the eigenstates, which condenses the full information of the linear map into a few moment generating states.<sup>[47]</sup> We refer to the new set of states  $\{\xi_l^m \mid l \in \mathbb{N}, m \in \{-l, \dots, l\}\}$ , which are obtained by a transformation of the  $\{\chi_i(\mathbf{r}) \mid i \in \mathbb{N}\}$ , as moment expanded states.

Then, the full density response due to  $V_{\text{pert}} \in \mathcal{R}[\mathbf{N}]$  (with  $N = L^2 + 2L$ ) can be expressed as:

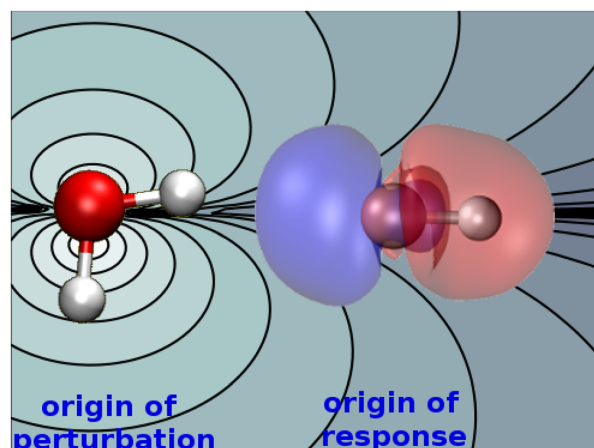
$$n^{\text{resp}}(\mathbf{r}) = \sum_{l=1}^L \sum_{m=-l}^l \xi_l^m(\mathbf{r}) \langle \xi_l^m, V_{\text{pert}} \rangle. \quad (12)$$

The moment expanded states are in general not normalized, as opposed to the regular eigenfunctions  $\chi_i(\mathbf{r})$  of  $\hat{T}$ . Hence there is no trivial quantity which characterizes the numeric convergence of this expansion. The  $\{\xi_l^m \mid l \in \mathbb{N}, m \in \{-l, \dots, l\}\}$  fulfill:

$$\langle \xi_l^m, \tilde{R}_{l'}^{m'} \rangle = 0 \quad \forall l > l' \quad (13)$$

Scherrer already demonstrated the use of the  $\{\xi_l^m \mid l \in \mathbb{N}, m \in \{-l, \dots, l\}\}$  for the efficient calculation of the response density of a water molecule disturbed by neighboring water molecules.<sup>[47]</sup> Also the dependency of the  $\{\xi_l^m \mid l \in \mathbb{N}, m \in \{-l, \dots, l\}\}$  on geometric distortions was investigated for the application of the linear density-density response function toward molecular dynamics simulations.<sup>[48]</sup> The principal applicability for the calculation of Raman spectra resulting from a post processed ab initio molecular dynamics trajectory was demonstrated.<sup>[48]</sup>

The transformation of the eigenstates into the  $\{\xi_l^m \mid l \in \mathbb{N}, m \in \{-l, \dots, l\}\}$  is algorithmically involved and takes place by iterative Givens rotations of thousands of eigenstates. The calculation of eigenstates (without transformations) requires an ab initio perturbation calculation per eigenstate.<sup>[47]</sup> In order to tackle this problem, we developed a more efficient, iterative algorithm (referred to as direct moment expansion), which needs a single DFPT calculation per moment expanded state.<sup>[49,50]</sup>



**FIGURE 1** Principal illustration of the response density of the water molecule (right) due to a perturbing water molecule (left). The potential originated from the left water molecule can be expanded at the responding (right) water molecule within a few basis functions

In this article, we present the reduced eigensystem representation as a completion of our development effort toward an efficient representation of the static linear density-density response function. The reduced eigensystem representation will allow to define a trivial criterion for the convergence of the approximation to the density response. Due to its truncated eigensystem like shape, the enormous reduction of the dimensionality will become apparent in comparison to the regular eigensystem representation.

### 3 | REPRESENTATION OF LINEAR OPERATORS WITHIN NON-ORTHOGONAL BASIS FUNCTIONS

In this section, we recall the specific notation for the expression of linear operators that we recently introduced.<sup>[50]</sup> The derivation closely follows the previous version, but differs in the final expression because the basis functions are now allowed to be non-orthogonal for the image as well as the domain of the linear operator.

The Dirac notation is a standard tool in quantum chemistry in order to express manipulations of linear operators and wave functions in an abstract formalism. In particular it allows the efficient transformations of the representation with respect to different bases. The theoretical justification of the Dirac notation is the Riesz representation theorem, which states that all separable infinite Hilbert spaces are isometric isomorph.

We defined the operator  $\hat{T}$  as linear map between two infinite separable Hilbert spaces, which justifies the use of the Dirac notation even through the physical meaning of a bracket vector differs from the conventional one (potentials and densities as opposed to wave functions). Here we will briefly review the Dirac notation, in particular highlighting deviations from standard notation for non-orthogonal basis sets.

The resolution of identity  $Id_{|p\rangle}$  can be expressed for an orthogonal basis set  $\{|p_n\rangle|n \in \mathbb{N}\}$  as:

$$Id_{|p\rangle} = \sum_n |p_n\rangle\langle p_n| \quad (14)$$

and for a non-orthogonal basis  $\{|q_n\rangle|n \in \mathbb{N}\}$  we have to write:

$$Id_{|q\rangle} = \sum_n |q_n\rangle \sum_m \left(\mathbf{S}_{|q\rangle}^{-1}\right)_{nm} \langle q_m| \quad (15)$$

with  $\mathbf{S}_{|q\rangle}$  being the overlap matrix with the elements  $(\mathbf{S}_{|q\rangle})_{nm} = \langle q_n, q_m \rangle$ . The matrix  $\mathbf{S}_q$  is known as Gramian matrix of the vectors  $\{|q_n\rangle|n \in \mathbb{N}\}$  and it is a non-singular matrix, if the  $\{|q_n\rangle|n \in \mathbb{N}\}$  are linearly independent. Thus the existence of the inverse  $\mathbf{S}_{|q\rangle}^{-1}$  is always ensured. A derivation of Equation (15) and an explanation of the treatment of non-orthogonal basis sets is given in the Supporting Information.

Within an orthogonal basis  $\{|p_n\rangle|n \in \mathbb{N}\}$ , the linear operator  $\hat{T}$  can be expressed as

$$\hat{T} = Id_{|p\rangle} \hat{T} Id_{|p\rangle} = \sum_n \sum_m |p_n\rangle\langle p_n, \hat{T} p_m \rangle \langle p_m| \quad (16)$$

with the element of the transformation matrix:

$$\left(\mathbf{T}_{\{p_k\}}^{\{p_k\}}\right)_{nm} = \langle p_n, \hat{T} p_m \rangle. \quad (17)$$

Herein,  $\left(\mathbf{T}_{\{p_k\}}^{\{p_k\}}\right)_{nm}$  is the  $(n, m)$ -th element of the transformation matrix. The elements of the transformation matrix depend on the choice of the basis of the domain as well as of the basis of the image of the linear operator  $\hat{T}$ . In order to distinguish transformation matrices expressed in different bases, we assign the super-/subscript indices to the transformation matrix  $\mathbf{T}_{\{p_k\}}^{\{p_k\}}$  to define the basis  $\{p_k\} := \{|p_k\rangle|k \in \mathbb{N}\}$  of the image/domain. The index  $k$  in  $\{p_k\}$  will only be used to refer to the entity of basis vectors  $\{|p_k\rangle|k \in \mathbb{N}\}$ , but is not for a specific indexing.

Within a non-orthogonal basis  $\{|q_n\rangle|n \in \mathbb{N}\}$  of the image as well as the domain the linear operator  $\hat{T}$  can be expressed as

$$\hat{T} = Id_{|q\rangle} \hat{T} Id_{|q\rangle} = \sum_n \sum_m \sum_o \sum_s |q_o\rangle \left(\mathbf{S}_{|q\rangle}^{-1}\right)_{on} \langle q_n, \hat{T} q_m \rangle \left(\mathbf{S}_{|q\rangle}^{-1}\right)_{ms} \langle q_s| \quad (18)$$

with the transformation matrix element:

$$\left(\mathbf{T}_{\{\mathbf{q}_k\}}^{\{\mathbf{q}_k\}}\right)_{om} = \sum_n \left(\mathbf{S}_{\{\mathbf{q}\}}^{-1}\right)_{on} \langle q_n, \hat{\mathbf{T}} q_m \rangle. \quad (19)$$

#### 4 | DERIVATION OF THE REDUCED EIGENSYSTEM REPRESENTATION

We recently generalized the moment expansion to any compact, self-adjoint, and positive-definite linear operator  $\hat{\mathbf{T}}: L^2[X] \rightarrow L^2[X]$ .<sup>[50]</sup> Assuming a finite (orthonormal) basis  $\{|P_1\rangle, |P_2\rangle, \dots, |P_N\rangle\}$  for a subspace of the domain of  $\hat{\mathbf{T}}$ , we showed recently that there is a specific basis set  $\{|\xi_1\rangle, |\xi_2\rangle, \dots, |\xi_N\rangle\}$  for the (finite) image of  $\hat{\mathbf{T}}$  which fulfills a partial orthogonality condition with respect to the domain basis:

$$\langle \xi_i, P_n \rangle = 0 \quad \forall i > n \quad (20)$$

with the interesting property that the restriction  $\mathcal{T} := \hat{\mathbf{T}}|_{\mathcal{P}[\mathbf{N}]}$  of  $\hat{\mathbf{T}}$  to the subspace  $\mathcal{P}[\mathbf{N}] := \text{span}(P_1, P_2, \dots, P_N)$  can be expressed according to

$$\mathcal{T} := \hat{\mathbf{T}}|_{\mathcal{P}[\mathbf{N}]} = \sum_{i=1}^N |\xi_i\rangle \langle \xi_i|. \quad (21)$$

Now, we use the expression for the transformation matrix element  $\left(\mathcal{T}_{\{\xi_k\}}^{\{\xi_k\}}\right)_{om}$  from Equation (19):

$$\left(\mathcal{T}_{\{\xi_k\}}^{\{\xi_k\}}\right)_{om} = \sum_{n=1}^N \left(\mathbf{S}_{\{\xi\}}^{-1}\right)_{on} \langle \xi_n, \mathcal{T} \xi_m \rangle = \sum_{n=1}^N \left(\mathbf{S}_{\{\xi\}}^{-1}\right)_{on} \left\langle \xi_n, \sum_{i=1}^N |\xi_i\rangle \langle \xi_i, \xi_m \rangle \right\rangle \quad (22)$$

$$= \sum_{n=1}^N \sum_{i=1}^N \left(\mathbf{S}_{\{\xi\}}^{-1}\right)_{on} \langle \xi_n, \xi_i \rangle \langle \xi_i, \xi_m \rangle = \sum_{n=1}^N \sum_{i=1}^N \left(\mathbf{S}_{\{\xi\}}^{-1}\right)_{on} \left(\mathbf{S}_{\{\xi\}}\right)_{ni} \left(\mathbf{S}_{\{\xi\}}\right)_{im}. \quad (23)$$

The last equation yields in matrix notation:

$$\mathcal{T}_{\{\xi_k\}}^{\{\xi_k\}} = \mathbf{S}_{\{\xi\}}^{-1} \mathbf{S}_{\{\xi\}} \mathbf{S}_{\{\xi\}} = \mathbf{S}_{\{\xi\}}. \quad (24)$$

The transformation matrix  $\mathcal{T}_{\{\xi_k\}}^{\{\xi_k\}} = \mathbf{S}_{\{\xi\}}$  is symmetric and finite, which in turn always allows its diagonalization:

$$\mathbf{U}^\dagger \mathbf{S}_{\{\xi\}} \mathbf{U} = \mathbf{D} \quad (25)$$

with an unitary matrix  $\mathbf{U}$  and a matrix  $\mathbf{D} = \text{diag}(\{\tilde{\lambda}_i | i \leq N\})$  with  $\tilde{\lambda}_i$  being the  $i$ -th eigenvalue. We apply the unitary matrix  $\mathbf{U}$  to the moment expanded states  $\{|\xi_1\rangle, |\xi_2\rangle, \dots, |\xi_N\rangle\}$ :

$$\tilde{\chi}_j(\mathbf{r}) = \sum_{i=1}^N \frac{\mathbf{U}_{ij}}{\sqrt{\tilde{\lambda}_j}} \xi_i(\mathbf{r}) \quad (26)$$

yielding a new set of functions  $\{\tilde{\chi}_1(\mathbf{r}), \tilde{\chi}_2(\mathbf{r}), \dots, \tilde{\chi}_N(\mathbf{r})\}$ . It can be shown (see Supporting Information for the full derivation) that these  $\{\tilde{\chi}_1(\mathbf{r}), \tilde{\chi}_2(\mathbf{r}), \dots, \tilde{\chi}_N(\mathbf{r})\}$  and  $\{\tilde{\lambda}_1, \tilde{\lambda}_2, \dots, \tilde{\lambda}_N\}$  are actually eigen-functions/-values of  $\mathcal{T} := \hat{\mathbf{T}}|_{\mathcal{P}[\mathbf{N}]}$  such that

$$\mathcal{T} = \sum_{i=1}^N |\tilde{\chi}_i\rangle \tilde{\lambda}_i \langle \tilde{\chi}_i|. \quad (27)$$

Thus, we denote the  $\{\tilde{\chi}_1(\mathbf{r}), \tilde{\chi}_2(\mathbf{r}), \dots, \tilde{\chi}_N(\mathbf{r})\}$  as reduced eigenstates and the  $\{\tilde{\lambda}_1, \tilde{\lambda}_2, \dots, \tilde{\lambda}_N\}$  as reduced eigenvalues of  $\hat{\mathbf{T}}$ . The reduced eigenstates are orthonormal ( $\langle \tilde{\chi}_i, \tilde{\chi}_j \rangle = \delta_{ij}$ ). The convergence behavior of the reduced eigensystem representation given by Equation (27) is controlled by the decay of the eigenvalues  $\{\tilde{\lambda}_1, \tilde{\lambda}_2, \dots, \tilde{\lambda}_N\}$ .



#### 4.1 | Chemical interpretation of the symmetrized representation

While the derivation of the symmetrized representation (Equation (27)) appears mathematically straightforward, we have so far ignored a subtle issue: Our starting point was the expansion of the perturbing potential  $|V_{\text{pert}}\rangle$  in Racah normalized harmonics, that is, within the space  $\mathcal{P}[\mathbf{N}] := \text{span}(|P_1\rangle, |P_2\rangle, \dots, |P_N\rangle)$ . In the eigensystem-like representation (Equation (27)), however, this has become an expansion within  $\mathcal{Z}[\mathbf{N}] := \text{span}(|\xi_1\rangle, |\xi_2\rangle, \dots, |\xi_N\rangle)$ , because the  $\{|\tilde{\chi}_i\rangle \mid i \in \{1, 2, \dots, N\}\}$  in Equation (27) are obtained by a unitary transformation of the moment expanded states  $\{|\xi_i\rangle \mid i \in \{1, 2, \dots, N\}\}$ , cf. Equation (26). The two spaces  $\mathcal{P}[\mathbf{N}]$  and  $\mathcal{Z}[\mathbf{N}]$  are in general not identical (despite their equivalence in rank and the partial orthogonality condition in Equation (20)).

Chemically, this corresponds to the different nature of potential functions and electronic densities: Perturbing potentials may have (remote) singularities but are spatially smooth, while densities tend to oscillate locally (think of orthogonality wiggles, for instance). Note that this corresponds to our choice for the treatment of the respective basis functions: The  $\{|P_i\rangle \mid i \in \{1, 2, \dots, N\}\}$  which span  $\mathcal{P}[\mathbf{N}]$  can be chosen freely (guided by chemical intuition), while the functions  $\{|\xi_i\rangle \mid i \in \{1, 2, \dots, N\}\}$  that span  $\mathcal{Z}[\mathbf{N}]$  are determined intrinsically by the actual result of the quantum chemical response calculations for the specific molecule under consideration. The conceptual aim of our approach is to use an optimized set  $\{|P_1\rangle, |P_2\rangle, \dots\}$  that is tuned for optimal representation of the space of physically meaningful perturbation potentials.

However, the two spaces  $\mathcal{Z}[\mathbf{N}]$  and  $\mathcal{P}[\mathbf{N}]$  being mathematically disjunct does not prevent the actual functions contained within them to have a large overlap. This is because the overlap is actually determined in the (much larger) underlying function space  $L^2[X]$  with  $X \subset \mathbb{R}^3$ . In the end, it is only this overlap which determines the actual density response to a perturbation. Actually, our previous work<sup>[47,49,50]</sup> shows that the overlap matrix  $\Theta^{\tilde{\chi}}$  with elements  $(\Theta^{\tilde{\chi}})_{ij} := \langle \tilde{\chi}_i, P_j \rangle$  is an upper triangular matrix and hence possess full rank. Therefore, the (small) part of a function from  $\mathcal{P}[\mathbf{N}]$  which exceeds the support of  $\mathcal{Z}[\mathbf{N}]$  does actually *not* contribute *any* response density (within  $\mathcal{Z}[\mathbf{N}]$ ). Hence, it is accurate to restrict the representation of the perturbation potential to the space  $\mathcal{Z}[\mathbf{N}]$ . In the Supporting Information, we demonstrate (using a detailed calculation) that the overlap integral  $\langle \xi_n, V_{\text{pert}} \rangle$  of  $|V_{\text{pert}}\rangle$  with  $|\xi_n\rangle$  is not effected by the expansion of a perturbing potential within the non-orthogonal  $\{|\xi_n\rangle \mid n \in \{1, \dots, N\}\}$ .

## 5 | DENSITY RESPONSE OF A H<sub>2</sub>O DIMER

We have computed the moment expanded states and the reduced eigenstates of a water molecule, and subsequently used those once-computed states to determine the density response of that water molecule to the perturbing potential of an adjacent second water molecule. For comparison, the exact density response was also computed for the dimer in the conventional self-consistent framework of density functional perturbation theory.

Three selected moment states and reduced eigenstates of the first water molecule are plotted in Figure 2. Neither representation bears any resemblance to the common molecular orbital shapes, and the moment expanded functions also do not appear to follow the spatial symmetry that is intuitively expected from the form of the underlying molecule. The reduced eigenstates exhibit a slightly more apparent symmetry, which however is lost for higher orders. It must be concluded that even for the comparably simple case of a water molecule, no intuitive understanding of the density-density response functions along the lines of frontier molecular orbitals is possible.

In a recent paper, we demonstrated that the density response of a molecule due to the perturbing potential of a neighboring molecule can be described within a few moment expanded states.<sup>[50]</sup> In this article, we reuse the design of the experiment in order to compare the convergence of the density response with respect to different numbers of employed moment expanded and reduced eigenstates: We choose the perturbative effect of one water molecule on an adjacent one as an elementary example. For simplicity, we use the Hartree potential using the partial charges of one of the most common force field water models (TIP3P) as perturbation. The Hartree potential of the TIP3P partial charges reads<sup>[51]</sup>:

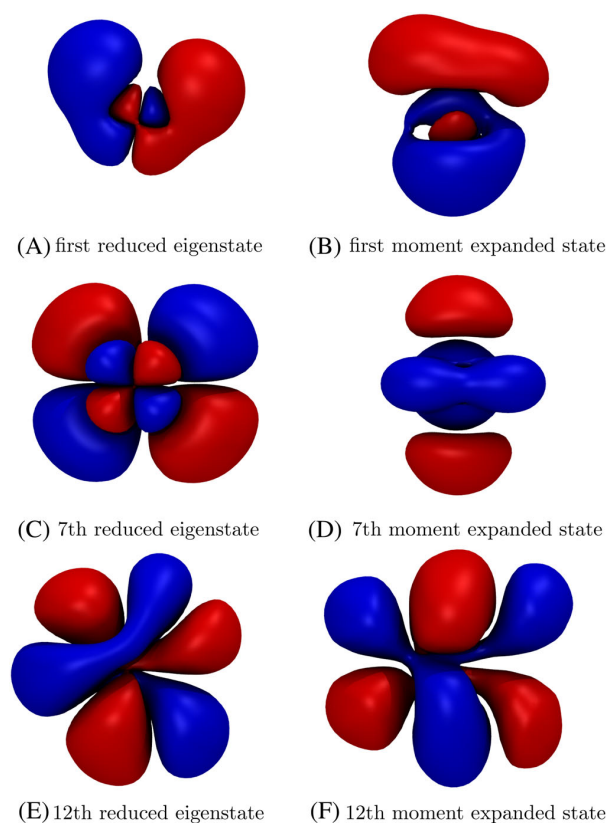
$$V_{\text{TIP3P}}^{\text{frag}}(\mathbf{r}) = \sum_{\lambda=1}^{N_n} \frac{Q_{\lambda}^{\text{partial}}}{|\mathbf{r} - \mathbf{R}_{\lambda}|} \text{erf}\left(\frac{|\mathbf{r} - \mathbf{R}_{\lambda}|}{\sqrt{2}\sigma}\right). \quad (28)$$

We created 12 water dimer configurations by placing the perturbing TIP3P water on six octahedral coordinates around the responding water molecule for two oxygen-oxygen distances of 2.5 Å and 2.75 Å. These distances were chosen because they represent typical hydrogen bonding distances, with 2.5 Å being virtually the closest possible distance that is thermodynamically relevant at common molecular dynamics simulation temperatures. At both distances, the density response was computed with our density-density response function formalism (in moment expansion and reduced eigensystem representation and using different numbers of states), as well as conventional self-consistent density functional perturbation theory calculations.

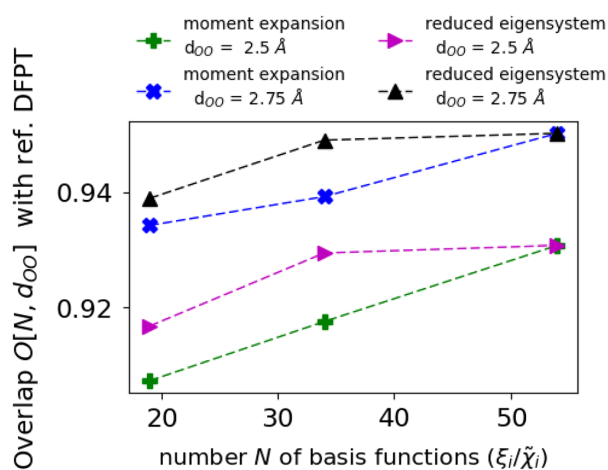
The overlap  $O[N, d_{OO}]$  of functions  $|\delta n_N\rangle$  and  $|\delta n_{\text{ref}}\rangle$  was taken according to:

$$O[N, d_{OO}] = \frac{\langle \delta n_N, \delta n_{\text{ref}} \rangle}{\langle \delta n_{\text{ref}}, \delta n_{\text{ref}} \rangle} \quad (29)$$

**FIGURE 2** Comparison of the moment expanded states and the reduced eigenstates of a water molecule



**FIGURE 3** Convergence of the density response employing moment expanded states ( $\xi_i$ ) and reduced eigenstates ( $\tilde{\chi}_i$ ) (with respect to different number  $N$  of states). The explicit density response was obtained from density functional perturbation theory (DFPT) calculations.  $d_{OO}$  denotes the oxygen-oxygen distance of the perturbing and the responding molecule

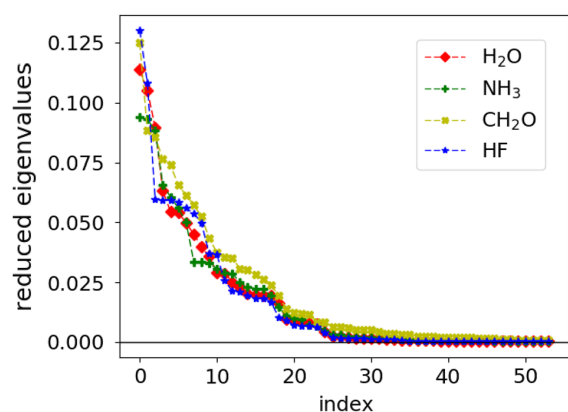


with  $|\delta n_N\rangle$  the calculated response density employing reduced eigenstates:

$$\delta n_N(\mathbf{r}) = \sum_{j=1}^N \tilde{\chi}_j(\mathbf{r}) \tilde{\lambda}_j \langle \tilde{\chi}_j, V_{\text{pert}} \rangle. \quad (30)$$

or moment expanded states:

$$\delta n_N(\mathbf{r}) = \sum_{j=1}^N \xi_j(\mathbf{r}) \langle \xi_j, V_{\text{pert}} \rangle. \quad (31)$$



**FIGURE 4** Decay of the reduced eigenvalues with respect to different molecules

The convergence of the calculation of  $\delta n_N$  with respect to the number of employed (moment expanded or reduced eigen-) states  $N$  is shown in Figure 3. We deduce two important statements from this graph:

Both the moment expansion and the reduced eigensystem representation of the density-density response function formalism have a similar convergence behavior, with the reduced eigensystem variant performing gradually better. Only 34 reduced eigenstates instead of 54 moment expanded states are necessary in order to obtain converged response densities. For the smaller distance (2.5 Å), which represents an already quite repulsive hydrogen bond configuration, the deviation from the explicit density functional perturbation theory calculation is about 7% to 8%, which improves to only 5% for the relaxed hydrogen bond conformation. Thus, both representations also yield a very accurate density response.

In order to determine how the reduced eigenvalues spectrum changes for different molecules, we calculated the reduced eigensystem representation of the linear density-density response function for H<sub>2</sub>O, NH<sub>3</sub>, CH<sub>2</sub>O, and HF. The resulting reduced eigenvalues are presented in Figure 4. We can report an exponential decay of the reduced eigenvalues for all investigated molecules, which motivates a further reduction of the number of significantly contributing states by around 50%.

## 6 | SUMMARY: ON THE DIFFERENT FLAVORS OF THE DENSITY-DENSITY RESPONSE FUNCTION

In this publication, we have introduced a new representation (the reduced eigensystem representation) of the linear density-density response function. The new representation is strongly connected to our recently developed moment expansion of the linear density-density response function.

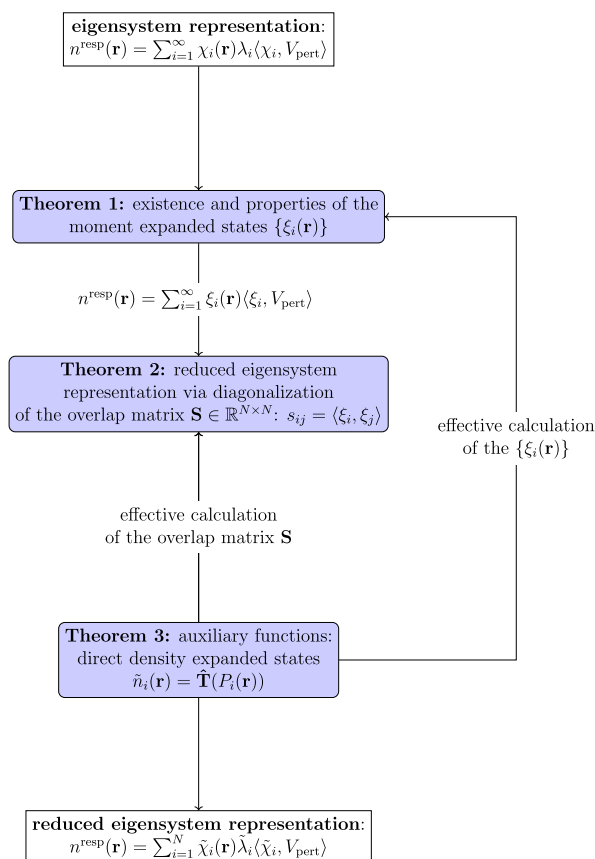
In order to facilitate a better understanding of the overall framework of this tool, we give a scheme of the underlying theory of the moment expansion and the reduced eigensystem representation in Figure 5. Starting from the regular (infinite) eigensystem representation, we derived the moment expansion which allowed a drastic reduction of the dimensionality (theorem 1 in Figure 5). By diagonalization of the overlap matrix of the moment expanded states, we developed the reduced eigensystem representation within this publication (theorem 2 in Figure 5). Both of these novel expansions originally required the calculation of several thousands of eigenstates, followed by adequate transformations. The introduction of the direct density expanded states allowed the calculation of the first  $N$  moment expanded /reduced eigenstates within  $N$  self-consistent DFPT calculations (theorem 3 in Figure 5).

The regular eigensystem representation as well as theorem 1 and 3 were recently published by our group.<sup>[29,30,47–50]</sup> In particular, we generalized these theorems to arbitrary linear, self-adjoint, positive-definite, and compact operators.<sup>[50]</sup> Within this article, we derived the reduced eigensystem representation as a completion of our development effort toward an efficient representation combined with an efficient determination of the static linear density-density response function. Due to its truncated eigensystem like shape, the enormous reduction of the dimensionality becomes obvious in comparison to the regular eigensystem representation. With this increased applicability, together with its more controlled convergence behavior, governed by the decay of the eigenvalues, we believe that the density-density response function formalism can soon be applied to “real” extended molecular systems.

## 7 | REMARKS FOR PRACTICAL APPLICATION OF THE REDUCED EIGENSYSTEM REPRESENTATION

Once the reduced eigensystem representation is calculated, the calculation of actual molecular response densities boils down to calculation of simple overlap integrals of reduced eigenfunctions and the perturbing potential. The computational costs of this procedure are typically two to three orders of magnitude smaller compared to a single self-consistent quantum chemistry calculation.

**FIGURE 5** A scheme of the structure of the underlying theory of the moment expansion  $\xi_i$  and the reduced eigensystem representation  $\tilde{\chi}_i$



The computational bottleneck for the calculation of the reduced eigensystem representation itself is the efficient calculation of the moment expanded states. In a recent article we stated an algorithm, which allows the calculation of the first  $N$  moment expanded states within  $N$  ab initio quantum chemistry calculations.<sup>[49,50]</sup> The accuracy of the reduced eigensystem representations can be adapted to the level of electronic structure theory utilized for the calculation of the moment expanded states.

A repeated application of the reduced eigensystem representation leads to the maximal efficiency gain. The core application for this approach will be in the field of molecular dynamics simulations, where the same kind of intermolecular interactions has to be computed multiple times at every single MD time step.

## 8 | CONCLUSION

In a previous article,<sup>[47,50]</sup> we demonstrated that the perturbing potential originated from neighboring molecules can be expressed within a few basis functions  $\{|P_1\rangle, |P_2\rangle, \dots, |P_N\rangle\}$  (choosing for example a polynomial expression of the potential) to a very high precision. In this article, we have started from an efficient representation of the electronic susceptibility (the moment expansion) for the calculation of molecular density responses due to perturbing potentials spanned by the  $\{|P_1\rangle, |P_2\rangle, \dots, |P_N\rangle\}$ . Here, we presented a modified representation—the reduced eigensystem representation—which allows the calculation of molecular response densities due to arbitrary perturbing potentials at the same very low storage and computational requirements compared to the moment expansion. The reduced eigensystem representation includes now a trivial criterion for the convergence of the approximation of the density response and by the eigensystem like structure of our new representation, the remarkable reduction of the dimensionality becomes immediately apparent for the calculation of the density-density response function. We demonstrated the exponential decay of the reduced eigenvalues for several molecules and we have illustrated our approach for the calculation of the response density on a simple water dimer, where 34 reduced eigenfunctions are sufficient to represent the density response due to the intermolecular interaction within an accuracy of a few percent. Employing the moment expanded states, we would need 54 moment expanded states in order to obtain the same accuracy.

## ORCID

Christian Dreßler  <https://orcid.org/0000-0002-0145-1263>

## REFERENCES

- [1] A. Goez, J. Neugebauer, Embedding Methods in Quantum Chemistry. in *Frontiers of Quantum Chemistry* (Eds: M. J. Wójcik, H. Nakatsuji, B. Kirtman, Y. Ozaki), Springer Singapore, Singapore **2018**, p. 139.
- [2] C. R. Jacob, J. Neugebauer, L. Visscher, *J. Comput. Chem.* **2008**, *29*, 1011.
- [3] T. A. Wesolowski, A. Warshel, *J. Phys. Chem.* **1993**, *97*, 8050.
- [4] D. G. Artiukhin, C. R. Jacob, J. Neugebauer, *J. Chem. Phys.* **2015**, *142*, 234101.
- [5] T. A. Wesolowski, *Computational Chemistry: Reviews of Current Trends*, World Scientific, Singapore **2006**, p. 1.
- [6] C. R. Jacob, J. Neugebauer, *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2014**, *4*, 325.
- [7] P. Cortona, *Phys. Rev. B* **1991**, *44*, 8454.
- [8] F. R. Manby, M. Stella, J. D. Goodpaster, T. F. Miller, *J. Chem. Theor. Comput.* **2012**, *8*, 2564.
- [9] F. Libisch, C. Huang, E. A. Carter, *Acc. Chem. Res.* **2014**, *47*, 2768.
- [10] X. Zhang, E. A. Carter, *J. Chem. Theor. Comput.* **2019**, *15*, 949.
- [11] S. Sharifzadeh, P. Huang, E. A. Carter, *Chem. Phys. Lett.* **2009**, *470*, 347.
- [12] A. D. Buckingham, *Advances in Chemical Physics*, John Wiley and Sons, Ltd, Hoboken, New Jersey **2007**, p. 107.
- [13] A. Stone, *Chem. Phys. Lett.* **1981**, *83*, 233.
- [14] J.-P. Piquemal, N. Gresh, C. Giessner-Prettre, *J. Phys. Chem. A* **2003**, *107*, 10353.
- [15] G. Naray-Szabo, G. G. Ferenczy, *Chem. Rev.* **1995**, *95*, 829.
- [16] J. G. Ángyán, C. Chipot, F. Dehez, C. Hättig, G. Jansen, C. Millot, *J. Comput. Chem.* **2003**, *24*, 997.
- [17] D. Elking, T. Darden, R. J. Woods, *J. Comput. Chem.* **2007**, *28*, 1261.
- [18] R. J. Wheatley, *Mol. Phys.* **1993**, *79*, 597.
- [19] G. G. Hall, C. M. Smith, *Int. J. Quantum. Chem.* **1984**, *25*, 881.
- [20] D. M. Elking, G. A. Cisneros, J.-P. Piquemal, T. A. Darden, L. G. Pedersen, *J. Chem. Theor. Comput.* **2010**, *6*, 190.
- [21] K. Eichkorn, O. Treutler, H. Öhm, M. Häser, R. Ahlrichs, *Chem. Phys. Lett.*, **1995**, *1995*(242), 652.
- [22] G. A. Cisneros, J.-P. Piquemal, T. A. Darden, *J. Chem. Phys.* **2005**, *123*, 044109.
- [23] B. Jeziorski, R. Moszynski, K. Szalewicz, *Chem. Rev.* **1887–1930**, *1994*, 94.
- [24] A. Heßelmann, *J. Chem. Theor. Comput.* **1943–1959**, *2018*, 14.
- [25] D. M. Benoit, D. Sebastiani, M. Parrinello, *Phys. Rev. Lett.* **2001**, *87*, 226401.
- [26] A. Heßelmann, G. Jansen, *Chem. Phys. Lett.* **2003**, *367*, 778.
- [27] A. J. Misquitta, B. Jeziorski, K. Szalewicz, *Phys. Rev. Lett.* **2003**, *91*, 033201.
- [28] A. Hesselmann, *J. Phys. Chem. A* **2011**, *115*, 11321.
- [29] A. C. Ihrig, A. Scherrer, D. Sebastiani, *J. Chem. Phys.* **2013**, *139*, 094102.
- [30] A. Scherrer, V. Verschinin, D. Sebastiani, *J. Chem. Theor. Comput.* **2012**, *8*, 106.
- [31] H. F. Wilson, F. M. C. Gygi, G. Galli, *Phys. Rev. B* **2008**, *78*, 113303.
- [32] X. Gonze, D. C. Allan, M. P. Teter, *Phys. Rev. Lett.* **1992**, *68*, 3603.
- [33] see standard textbooks for functional analysis, e.g. Dirk Werner, *Funktionalanalysis*.
- [34] X. Gonze, J.-P. Vigneron, *Phys. Rev. B* **1989**, *39*, 13120.
- [35] P. Giannozzi, S. de Gironcoli, P. Pavone, S. Baroni, *Phys. Rev. B* **1991**, *43*, 7231.
- [36] X. Gonze, *Phys. Rev. A* **1995**, *52*, 1086.
- [37] X. Gonze, *Phys. Rev. A* **1995**, *52*, 1096.
- [38] S. Baroni, S. de Gironcoli, A. del Corso, P. Giannozzi, *Rev. Mod. Phys.* **2001**, *73*, 515.
- [39] A. Putrino, D. Sebastiani, M. Parrinello, *J. Chem. Phys.* **2000**, *113*, 7102.
- [40] S. Hamel, A. J. Williamson, H. F. Wilson, F. Gygi, G. Galli, E. Ratner, D. Wack, *Appl. Phys. Lett.* **2008**, *92*, 043115.
- [41] D. Lu, F. Gygi, G. Galli, *Phys. Rev. Lett.* **2008**, *100*, 147601.
- [42] H. F. Wilson, D. Lu, F. M. C. Gygi, G. Galli, *Phys. Rev. B* **2009**, *79*, 245106.
- [43] D. Lu, H.-V. Nguyen, G. Galli, *J. Chem. Phys.* **2010**, *133*, 154110.
- [44] T. A. Pham, T. Li, S. Shankar, F. Gygi, G. Galli, *Appl. Phys. Lett.* **2010**, *96*, 062902.
- [45] D. Rocca, D. Lu, G. Galli, *J. Chem. Phys.* **2010**, *133*, 164109.
- [46] T. A. Pham, T. Li, S. Shankar, F. Gygi, G. Galli, *Phys. Rev. B* **2011**, *84*, 045308.
- [47] A. Scherrer, D. Sebastiani, *J. Comput. Chem.* **2016**, *37*, 665.
- [48] A. Scherrer, C. Dreßler, P. Ahlert, D. Sebastiani, *J. Chem. Phys.* **2016**, *144*, 144111.
- [49] P. Ahlert, A. Scherrer, C. Dreßler, D. Sebastiani, *Eur. Phys. J. B* **2018**, *91*, 94.
- [50] C. Dreßler, A. Scherrer, P. Ahlert, D. Sebastiani, *J. Comput. Chem.* **2019**, *40*, 2712.
- [51] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, *J. Chem. Phys.* **1983**, *79*, 926.
- [52] CPMD 3.15.3, <http://www.cpmid.org/>, Copyright IBM Corp 1990–2008, Copyright MPI für Festkörperforschung Stuttgart 1997–2001 (accessed: May, 2016).
- [53] W. Kohn, L. J. Sham, *Phys. Rev.* **1965**, *140*, A1133.
- [54] P. Hohenberg, W. Kohn, *Phys. Rev.* **1964**, *136*, B864.
- [55] N. Troullier, J. L. Martins, *Phys. Rev. B* **1991**, *43*, 1993.
- [56] A. D. Becke, *Phys. Rev. A* **1988**, *38*, 3098.

[57] C. Lee, W. Yang, R. G. Parr, *Phys. Rev. B* **1988**, 37, 785.

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Dreßler C, Sebastiani D. Reduced eigensystem representation of the linear density-density response function. *Int J Quantum Chem.* 2019;e26085. <https://doi.org/10.1002/qua.26085>

## APPENDIX

### COMPUTATIONAL DETAILS


The moment expansion has been implemented in our development version of the CPMD<sup>[52]</sup> electronic structure package. The transformation of the moment expanded states into the reduced eigenstates was done using own python scripts. The calculations have been performed using Density Functional Perturbation Theory<sup>[37-39,53,54]</sup> with Troullier-Martins<sup>[55]</sup> pseudo potentials in the Becke,<sup>[56]</sup> and Lee-Yang-Parr<sup>[57]</sup> approximation for the exchange correlation kernel. The perturbing potential was expressed by a polynomial expansion. The basis was formed by the set of polynomials  $x^l y^m z^n$  with  $l, m, n \in \mathbb{N}$ , which fulfill the condition  $l + m + n \leq i_{\max}$ . The largest basis set consisted out of 54 polynomials ( $i_{\max} = 5$ ). The calculation of the reference density response with CPMD was done for a relaxed water geometry. We used a plane wave basis for all involved quantities, be it the electronic density, the perturbing potentials or the moment expanded states. The real space representation naturally gives rise to a regular grid. For our chosen parameters, the grid increment is about 0.08 Angstroms, whereas the spread of the Gaussians in Equation (28) is chosen as  $\sigma = 0.5a_0$ .

**3.2.2 Article VII: Efficient Representation of the LDDRF**

Christian Dreßler, Arne Scherrer, Paul Ahlert, and Daniel Sebastiani.  
Efficient representation of the linear density-density response function.  
*Journal of Computational Chemistry*, 40(31):2712–2721, **2019**.

In this article, I generalized the concept of the moment expansion to arbitrary compact, positive and self-adjoint linear operators. By means of a thorough mathematical investigation, I proved that the moment expansion can be understood in terms of a QR decomposition of a matrix. I derived that a computationally cheaper algorithm for the calculation of the moment-expanded representation of the linear density-density response function can be explained within a Cholesky decomposition of a small matrix. I prepared the manuscript. D. Sebastiani supervised the project and provided me with valuable advice during the derivation of the equations and while I was writing the manuscript. Arne Scherrer contributed the initial derivation of the (Direct-) moment expansion and Paul Ahlert helped to prepare the manuscript.

# Efficient Representation of the Linear Density-Density Response Function

Christian Dreßler , Arne Scherrer, Paul Ahlert, and Daniel Sebastiani\*

We present a thorough derivation of the mathematical foundations of the representation of the molecular linear electronic density-density response function in terms of a computationally highly efficient moment expansion. Our new representation avoids the necessities of computing and storing numerous eigenfunctions of the response kernel by means of a considerable dimensionality reduction about from  $10^3$  to  $10^1$ . As the scheme is applicable to any compact, self-adjoint, and positive definite linear operator, we present a general formulation,

which can be transferred to other applications with little effort. We also present an explicit application, which illustrates the actual procedure for applying the moment expansion of the linear density-density response function to a water molecule that is subject to a varying external perturbation potential. © 2019 The Authors. *Journal of Computational Chemistry* published by Wiley Periodicals, Inc.

DOI: 10.1002/jcc.26046

## Introduction

Almost any situation in physics and chemistry can be described by a state function and its temporal evolution. The latter is generally governed by a differential equation, for example, the Newton equation of motion, the time-dependent Schrödinger or Dirac equation, the Navier–Stokes equation, the Maxwell equations, and others. In many of these situations, a viable strategy for simplifying the numerical solution of the underlying differential equations is the separation of both equation and solution into two parts of “large” and “small” amplitudes, respectively, which are then technically easier to deal with. The formal framework of this idea is commonly called perturbation theory, and the formal solution to this type of problem is often called Green’s function.

The standard perturbation theory ansatz is based on the assumption that there is one given “large” part and one given “small” contribution of the differential operator, so that two new specific differential equations result from this large/small separation. However, there are situations in which a whole variety of “small” contributions exists (but only one “large” differential equation). Examples for these situations from molecular quantum chemistry are the calculation of vibrational modes, where there is one perturbation Hamiltonian for the displacement of each atom in each Cartesian direction or molecular dynamics simulations within embedding schemes, that is, where a given molecule is subject to varying chemical environments. In such situations, the same perturbation theory equation has to be solved repeatedly for numerous different perturbations.

In the specific case of quantum chemistry, the formal solution to this problem is called the density-density response function, which means a linear operator that acts on the perturbation and yields the electronic density response to that perturbation. This function is very general, as it yields the response density for any perturbation potential, but in turn, it is not readily available explicitly in numerical form. Nevertheless,

there is a path for the explicit calculation of this complex quantity by means of a particular mathematical representation (using the eigenfunctions and eigenvalues of the density-density-response function seen as a matrix operator).<sup>[1,2]</sup> The straightforward calculation has been shown to be accurate yet computationally cumbersome. Here, we provide a mathematical study of a recently developed, considerably more efficient variant for the computation and the storage of this quantity (called the moment representation of this response function<sup>[3,4]</sup>).

Compact linear operators often appear as integral transforms and can be viewed as continuous generalizations of matrices, where the corresponding integral kernel fulfills some weak conditions, that is, the kernel must not be singular and must decay to zero fast enough at infinity. The compactness of these operators ensures that the matrix elements are well defined.

For  $X \subseteq \mathbb{R}^3$  and  $\chi(\mathbf{r}, \mathbf{r}') \in L^2(X \times X)$  eq. (1) defines a linear Hilbert-Schmidt operator  $\hat{\mathbf{T}}: L^2(X) \rightarrow L^2(X)$ , which is in particular compact.<sup>[5]</sup>

$$\hat{\mathbf{T}}: f(\mathbf{r}) \mapsto \int \chi(\mathbf{r}, \mathbf{r}') f(\mathbf{r}') d^3 r' \quad (1)$$

If the kernel is furthermore symmetric ( $\chi(\mathbf{r}, \mathbf{r}') = \chi(\mathbf{r}', \mathbf{r}) \forall \mathbf{r}, \mathbf{r}' \in X$ ) and positive

( $\int \int \chi(\mathbf{r}, \mathbf{r}') f(\mathbf{r}) f(\mathbf{r}') d^3 r d^3 r' \geq 0 \forall f(\mathbf{r}) \in L^2(X)$ ) the operator is self-adjoint and positive.

The properties of this class of operators are well-understood from a mathematical point of view. In physics or quantum chemistry, these integral transforms are often derived from the

[a] C. Dreßler, A. Scherrer, P. Ahlert, D. Sebastiani  
Institute of Chemistry, Martin-Luther-University Halle-Wittenberg, 06120,  
Halle (Saale), Germany  
E-mail: daniel.sebastiani@chemie.uni-halle.de

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2019 The Authors. *Journal of Computational Chemistry* published by Wiley Periodicals, Inc.



inverse Hamiltonian  $\hat{H}^{-1}$  and have many applications, for example, the Lippmann-Schwinger equation for the calculation of scattering in quantum mechanics. The Lippmann-Schwinger equation relates the scattered wave function with the interaction that produces the scattering (the scattering potential). Another example is the frequency dependent density-density response function, which includes an externally defined energy shift of the molecular orbital spectrum (originating from a photon of given frequency), which in turn leads to a breakdown of the linear response regime when the photon energy matches an electronic excitation energy. While this latter situation is within our view, we are still working on how to adapt the present representation of the linear density-density response function to this energy-dependent version.

In this article, the operator defined by the static linear density-density response function  $\hat{T}: V_{\text{pert}}(\mathbf{r}) \mapsto n^{\text{resp}}(\mathbf{r})$  will be considered as an example for the illustration of our theorems for linear, self-adjoint and positive definite operators.

For any given specific perturbation potential  $V_{\text{pert}}^{xy}(\mathbf{r})$  the corresponding electronic response density  $n_{xy}^{\text{resp}}(\mathbf{r})$  can be calculated via standard quantum chemical approaches, for example, by means of a perturbation-theory based calculation.<sup>[6–12]</sup> The challenge, however, is to develop a formalism for any (arbitrary) perturbation potential to a level where no explicit quantum chemical calculation is needed any more once the perturbation potential is specified. The core application for this approach will be molecular dynamics simulations, where the potential experienced by a given molecule varies slightly from any molecular dynamics step to the next.

Once the kernel  $\chi(\mathbf{r}, \mathbf{r}')$  is calculated, the calculation of molecular response density breaks down to integration according to eq. (1). The computational costs of this procedure are negligible compared to a single quantum chemistry calculation.

Unfortunately, there is no closed expression for the calculation of the kernel  $\chi(\mathbf{r}, \mathbf{r}')$  and hence for the calculation of density responses. The (molecular density-density) response function is a six dimensional nonlocal kernel  $\chi(\mathbf{r}, \mathbf{r}')$  and its storage on a grid would also require excessive storage dimensions, which decrease the efficiency of any real application.

A possible attempt to overcome these problems is to apply the spectral theorem of compact and self-adjoint operators. Thus, we can express the action of the operator  $\hat{T}$  to a vector  $f(\mathbf{r})$  as:

$$\hat{T}(f(\mathbf{r})) = \sum_{i=1}^{\infty} \chi_i(\mathbf{r}) \lambda_i \langle \chi_i, f \rangle \quad (2)$$

using eigenvalues  $\lambda_i$  and eigenfunctions  $\chi_i(\mathbf{r})$ , which can be stored without difficulties.<sup>[1,2,13,14]</sup> The eigenfunctions and eigenvalues satisfy  $\hat{T}\chi_i(\mathbf{r}) = \lambda_i \chi_i(\mathbf{r})$ .  $\langle \chi_i, f \rangle$  denotes, the inner product of functions  $\chi_i(\mathbf{r})$  and  $f(\mathbf{r})$  in  $L^2(X)$  and is defined as  $\langle \chi_i, f \rangle = \int_X \overline{\chi_i(\mathbf{r})} f(\mathbf{r}) d^3r$ .

According to the spectral theorem, the sequence of the eigenvalues converges to zero ( $\lim_{i \rightarrow \infty} \lambda_i = 0$ ), which allows

approximation of eq. (2) via a finite sum, that is, the truncation of the eigensystem representation at a finite index to the desired numerical precision. In particular, the sequence of eigenvalues is square summable, as the operator  $\hat{T}$  belongs to the second shadow class as a Hilbert-Schmidt operator.<sup>[15]</sup> Thus the eigenvalues decay fast enough to satisfy the condition  $\sum_{i=1}^{\infty} (\lambda_i)^2 < \infty$ .

For our specific example of  $\hat{T}: V_{\text{pert}}(\mathbf{r}) \mapsto n^{\text{resp}}(\mathbf{r})$  with  $\chi(\mathbf{r}, \mathbf{r}')$  equal to the molecular density-density response function, there is a suitable way to compute and store the most significant subset of eigenvalues and eigenfunctions. We were able to calculate the eigensystem representation and could already show that the calculation of converged response densities would require several thousand eigenvalues and eigenstates.<sup>[1,2]</sup> We want to emphasize that we furthermore have to calculate several thousand scalar products for the evaluation of each single perturbing potential. This is a computationally demanding protocol.

In this article, we present a transformation of this subset of eigenstates, which condenses the physically relevant information into a new set of states of considerably reduced dimensionality (moment expansion). Our aim in this manuscript is to provide a thorough mathematical basis for this transformation by means of two theorems.

For the specific case of the static linear density-density response function and a specific basis of the perturbing potential, we recently published a recipe for the calculation of the eigensystem representation<sup>[1,2]</sup> as well as the moment expansion.<sup>[3,4]</sup> Here, we generalize these results to an entire class of linear operators and arbitrary basis sets.

## Conditions/Restrictions for the Application of the Theorems

In this article, we focus on linear operators defined by integral transforms according to eq. (1). Thus, both domain and image consist of square integrable functions. A closer inspection of eq. (2) reveals that only domain vectors  $f(\mathbf{r})$  with a nonvanishing overlap with the eigenstates ( $\exists i \in \mathbb{N}: \langle \chi_i, f \rangle \neq 0$ ) will contribute to the image vector. Unfortunately, we do not have any information about the shape of the eigenstates in general.

However, we assume the linear operator  $\hat{T}$  will have compact support, that is,  $\forall i \in \mathbb{N}: \chi_i(\mathbf{r}) \in L^2[X]$  with  $X \subset \mathbb{R}^3$ . Thus, we are only interested in a local expansion of the domain vectors on a compact subspace  $X \subset \mathbb{R}^3$ .

For concrete evaluations of the operator  $\hat{T}$ , we have to choose a basis  $\{P_n(\mathbf{r}) | n \in \mathbb{N}\}$  for the representation of a domain vector  $f(\mathbf{r})$ :

$$f(\mathbf{r}) = \sum_{n=1}^{\infty} c_n P_n(\mathbf{r}) \quad (3)$$

with  $c_n$  being the  $n$ -th basis expansion coefficient.

A precise approximation of a function  $f(\mathbf{r})$  in a local environment (at a compact subspace) is often possible within a finite

number of basis functions. For example, an analytic function can always be expanded within a Taylor series (using the monomials as basis of the vector space). The terms of the Taylor expansion are ordered with respect to descending importance and it is even possible to approximate the introduced error by a truncated expansion.

We can validate, if the conditions for the domain vectors are fulfilled for a specific choice of  $\hat{\mathbf{T}}$ , for example, for the operator  $\hat{\mathbf{T}}: V_{\text{pert}}(\mathbf{r}) \mapsto n^{\text{resp}}(\mathbf{r})$  obtained by setting the kernel  $\chi(\mathbf{r}, \mathbf{r}')$  to the molecular density-density response function. This operator links an arbitrary perturbing potential to the resulting molecular response density. Due to the locality of the quantum mechanical Hamilton operator, these response densities occur only in the vicinity of a molecule, which corresponds to the condition of compact support. Thus, we are only interested in an (approximate) representation of the perturbing potential in an area around the responding molecule, that is, in particular not around the physical origin of the perturbation.

A truncated multipole expansion (e.g., in terms of Racah normalized irregular solid harmonics [15,16]) allows such a finite representation of the perturbing potential.

A demonstration of the truncated expression of the perturbing potential of a water molecule is presented in the last section of this article.

## Non-Orthogonal Basis Functions

The Dirac notation is a standard tool in quantum chemistry in order to express manipulations of linear operators and wave functions in an abstract formalism. In particular, it allows the efficient transformations of the representation with respect to different bases. The theoretical justification of the Dirac notation is the Riesz representation theorem, which states that all separable infinite Hilbert spaces are isometric isomorph.

We defined the operator  $\hat{\mathbf{T}}$  as linear map between two infinite separable Hilbert spaces, which justifies the use of the Dirac notation even through the physical meaning of a bra/ket vector differs from the conventional one (potentials and densities as opposed to wave functions). Here we will briefly review the Dirac notation, in particular highlighting deviations from standard notation for non-orthogonal basis sets.

The resolution of identity  $Id_{\{p\}}$  can be expressed for an orthogonal basis set  $\{|p_n\rangle \mid n \in \mathbb{N}\}$  as:

$$Id_{\{p\}} = \sum_n |p_n\rangle\langle p_n| \quad (4)$$

and for a non-orthogonal basis  $\{|q_n\rangle \mid n \in \mathbb{N}\}$  we have to write:

$$Id_{\{q\}} = \sum_n |q_n\rangle \sum_m \left(\mathbf{S}_{\{q\}}^{-1}\right)_{nm} \langle q_m| \quad (5)$$

with  $\mathbf{S}_{\{q\}}$  being the overlap matrix with the elements  $(\mathbf{S}_{\{q\}})_{nm} = \langle q_n | q_m \rangle$ . The matrix  $\mathbf{S}_{\{q\}}$  is known as Gramian matrix of the vectors  $\{|q_n\rangle \mid n \in \mathbb{N}\}$  and it is a non-singular matrix, if the

$\{|q_n\rangle \mid n \in \mathbb{N}\}$  are linearly independent. Thus, the existence of the inverse  $\mathbf{S}_{\{q\}}^{-1}$  is always ensured. A derivation of eq. (5) and an explanation of the treatment of non-orthogonal basis sets is given in the Supporting Information.

Within an orthogonal basis  $\{|p_n\rangle \mid n \in \mathbb{N}\}$ , the linear operator  $\hat{\mathbf{T}}$  can be expressed as

$$\hat{\mathbf{T}} = Id_{\{p\}} \hat{\mathbf{T}} Id_{\{p\}} = \sum_n \sum_m |p_n\rangle \langle p_n | \hat{\mathbf{T}} |p_m\rangle \langle p_m| \quad (6)$$

with the element of the transformation matrix:

$$\left(\mathbf{T}_{\{p_k\}}^{\{p_k\}}\right)_{nm} = \langle p_n | \hat{\mathbf{T}} |p_m\rangle. \quad (7)$$

Herein,  $\left(\mathbf{T}_{\{p_k\}}^{\{p_k\}}\right)_{nm}$  is the (n,m)-th element of the transformation matrix. The elements of the transformation matrix depend on the choice of the basis of the domain as well as of the basis of the image of the linear operator  $\hat{\mathbf{T}}$ . In order to distinguish transformation matrices expressed in different bases, we assign the super/subscript indices to the transformation matrix  $\mathbf{T}_{\{p_k\}}^{\{p_k\}}$  to define the basis  $\{p_k\} := \{|p_k\rangle \mid k \in \mathbb{N}\}$  of the image/domain. The index k in  $\{p_k\}$  will only be used to refer to the entity of basis vectors  $\{|p_k\rangle \mid k \in \mathbb{N}\}$ , but is not for a specific indexing.

Within a non-orthogonal basis  $\{|q_n\rangle \mid n \in \mathbb{N}\}$  of the image and an orthogonal basis  $\{|p_n\rangle \mid n \in \mathbb{N}\}$  of the domain, the linear operator  $\hat{\mathbf{T}}$  can be expressed as

$$\hat{\mathbf{T}} = Id_{\{q\}} \hat{\mathbf{T}} Id_{\{p\}} = \sum_o \sum_n \sum_m |q_o\rangle \left(\mathbf{S}_{\{q\}}^{-1}\right)_{on} \langle q_n | \hat{\mathbf{T}} |p_m\rangle \langle p_m| \quad (8)$$

with the element of the transformation matrix:

$$\left(\mathbf{T}_{\{p_k\}}^{\{q_k\}}\right)_{om} = \sum_n \left(\mathbf{S}_{\{q\}}^{-1}\right)_{on} \langle q_n | \hat{\mathbf{T}} |p_m\rangle. \quad (9)$$

## Derivation of the Moment Expansion

### Theorem 1: Moment Expansion.

Let  $\hat{\mathbf{T}}: L^2(X) \rightarrow L^2(X)$  with  $X \subset \mathbb{R}^3$  be a compact, positive-definite and self-adjoint linear operator with eigensystem representation from eq. (2). Further, let  $\{P_n \mid n \in \mathbb{N}\}$  be an orthonormal basis of the domain of  $\hat{\mathbf{T}}$ . Then:

1)  $\forall \{P_n \mid n \in \mathbb{N}\} \exists$  a unique transformation  $\mathbf{Q}$  with  $|\xi_n\rangle := \sum_{i=1}^{\infty} Q_{in} |\sqrt{\lambda_i} \chi_i\rangle$ , such that

$$\langle \xi_i | P_n \rangle = 0 \quad \forall i > n \quad (10)$$

and

$$\left(\mathbf{T}_{\{p_k\}}^{\{\xi_k\}}\right)_{in} = \langle \xi_i | P_n \rangle. \quad (11)$$

2) The operator  $\hat{\mathbf{T}}$  can be expressed as

$$\hat{\mathbf{T}} = \sum_{i=1}^{\infty} |\xi_i\rangle\langle\xi_i|. \quad (12)$$

The restriction  $\hat{\mathbf{T}}|_{\mathcal{P}[\mathbf{N}]}$  of  $\hat{\mathbf{T}}$  to  $\mathcal{P}[\mathbf{N}] := \text{Span}(P_1, P_2, \dots, P_N)$  is given by

$$\hat{\mathbf{T}}|_{\mathcal{P}[\mathbf{N}]} |f\rangle = \sum_{i=1}^N |\xi_i\rangle\langle\xi_i, f| \quad \forall f \in \mathcal{P}[\mathbf{N}]. \quad (13)$$

**Proof:**

**Proof of sub-item 1:**

The overlap matrix  $\mathbf{S}_{\{\sqrt{\lambda_k}\chi_k\}}$  defined by  $(\mathbf{S}_{\{\sqrt{\lambda_k}\chi_k\}})_{ij} = \langle\sqrt{\lambda_i}\chi_i, \sqrt{\lambda_j}\chi_j\rangle = \lambda_i\delta_{ij}$  is a diagonal matrix with inverse  $\mathbf{S}_{\{\sqrt{\lambda_k}\chi_k\}}^{-1}$ :

$$(\mathbf{S}_{\{\sqrt{\lambda_k}\chi_k\}}^{-1})_{ij} = \frac{1}{\lambda_i}\delta_{ij}. \quad (14)$$

We choose the non-normalized  $\{\sqrt{\lambda_i}\chi_i \mid i \in \mathbb{N}\}$  as basis of the image and the orthonormal  $\{P_i \mid i \in \mathbb{N}\}$  as basis of the domain of  $\hat{\mathbf{T}}$ . According to eq. (9) we can express the elements of the infinite transformation matrix as following:

$$\begin{aligned} (\mathbf{T}_{\{\sqrt{\lambda_k}\chi_k\}}^{\{\sqrt{\lambda_k}\chi_k\}})_{jn} &= \sum_{i=1}^{\infty} (\mathbf{S}_{\{\sqrt{\lambda_k}\chi_k\}}^{-1})_{ij} \langle\sqrt{\lambda_j}\chi_j, \hat{\mathbf{T}}P_n\rangle = \lambda_j^{-1}\lambda_j \langle\sqrt{\lambda_j}\chi_j, P_n\rangle \\ &= \langle\sqrt{\lambda_j}\chi_j, P_n\rangle \end{aligned} \quad (15)$$

As the transformation matrix of  $\hat{\mathbf{T}}$  represents a bounded linear operator on the separable Hilbert space  $\ell^2$ , there is a QR decomposition of  $\mathbf{T}_{\{\sqrt{\lambda_k}\chi_k\}}^{\{\sqrt{\lambda_k}\chi_k\}}$  into the product of an orthogonal matrix  $\mathbf{Q}$  and an upper triangular matrix  $\mathbf{R}$ .<sup>[17,18]</sup>

$$\mathbf{T}_{\{\sqrt{\lambda_k}\chi_k\}}^{\{\sqrt{\lambda_k}\chi_k\}} = \mathbf{Q}\mathbf{R} \Leftrightarrow \mathbf{R} = \mathbf{Q}^{-1}\mathbf{T}_{\{\sqrt{\lambda_k}\chi_k\}}^{\{\sqrt{\lambda_k}\chi_k\}}. \quad (16)$$

Equation (16) reads element-wise:

$$\mathbf{R}_{ij} = \sum_{l=1}^{\infty} (\mathbf{Q}^{-1})_{il} (\mathbf{T}_{\{\sqrt{\lambda_k}\chi_k\}}^{\{\sqrt{\lambda_k}\chi_k\}})_{lj}. \quad (17)$$

With  $(\mathbf{Q}^{-1})_{il} = \mathbf{Q}_{li}$  and  $(\mathbf{T}_{\{\sqrt{\lambda_k}\chi_k\}}^{\{\sqrt{\lambda_k}\chi_k\}})_{lj} = \langle\sqrt{\lambda_l}\chi_l, P_j\rangle$  we obtain:

$$\mathbf{R}_{ij} = \sum_{l=1}^{\infty} \mathbf{Q}_{li} \langle\sqrt{\lambda_l}\chi_l, P_j\rangle = \left\langle \sum_{l=1}^{\infty} \mathbf{Q}_{li} \sqrt{\lambda_l}\chi_l, P_j \right\rangle. \quad (18)$$

We define

$$\xi_i = \sum_{l=1}^{\infty} \mathbf{Q}_{li} \sqrt{\lambda_l}\chi_l, \quad (19)$$

so that  $\mathbf{R}_{ij} = \langle\xi_i, P_j\rangle$ . The fact that  $\mathbf{R}$  is an upper triangular matrix implies that:  $\langle\xi_i, P_j\rangle = 0 \forall i > j$ . The last expression proves equation (10) from sub-item 1.

In the next step, we prove equation (11) from sub-item 1. Equation (19) shows that  $\mathbf{Q}$  represents a change-of-basis matrix which we will denote by  $\mathbf{W}_{\{\xi_k\}}^{\{\sqrt{\lambda_k}\chi_k\}} := \mathbf{Q}$ . This matrix transforms the  $\{\xi_i \mid i \in \mathbb{N}\}$  into the  $\{\sqrt{\lambda_i}\chi_i \mid i \in \mathbb{N}\}$ . We can write for the inverse basis transformation:

$$\mathbf{W}_{\{\xi_k\}}^{\{\sqrt{\lambda_k}\chi_k\}} = \left(\mathbf{W}_{\{\xi_k\}}^{\{\sqrt{\lambda_k}\chi_k\}}\right)^{-1} \quad (20)$$

On the one hand, applying the basis transformation to  $\mathbf{T}_{\{\sqrt{\lambda_k}\chi_k\}}^{\{\sqrt{\lambda_k}\chi_k\}}$  yields by definition (compare commutative diagram of standard basis transformations in the Supporting Information):

$$\mathbf{W}_{\{\xi_k\}}^{\{\sqrt{\lambda_k}\chi_k\}} \mathbf{T}_{\{\sqrt{\lambda_k}\chi_k\}}^{\{\sqrt{\lambda_k}\chi_k\}} = \mathbf{T}_{\{\xi_k\}}^{\{\xi_k\}} \quad (21)$$

and on the other hand, the result of the transformation can be obtained from eqs. (16) and (20):

$$\mathbf{W}_{\{\xi_k\}}^{\{\sqrt{\lambda_k}\chi_k\}} \mathbf{T}_{\{\sqrt{\lambda_k}\chi_k\}}^{\{\sqrt{\lambda_k}\chi_k\}} = \left(\mathbf{W}_{\{\xi_k\}}^{\{\sqrt{\lambda_k}\chi_k\}}\right)^{-1} \mathbf{T}_{\{\sqrt{\lambda_k}\chi_k\}}^{\{\sqrt{\lambda_k}\chi_k\}} = \mathbf{Q}^{-1} \mathbf{T}_{\{\sqrt{\lambda_k}\chi_k\}}^{\{\sqrt{\lambda_k}\chi_k\}} = \mathbf{R}. \quad (22)$$

Combining eqs. (21) and (22) we get:  $\mathbf{T}_{\{\xi_k\}}^{\{\xi_k\}} = \mathbf{R}$  with

$$(\mathbf{T}_{\{\xi_k\}}^{\{\xi_k\}})_{ij} = \langle\xi_i, P_j\rangle. \quad (23)$$

**Proof of sub-item 2:**

We choose the non-orthonormal  $\{\xi_i \mid i \in \mathbb{N}\}$  as basis of the image and the orthonormal  $\{P_i \mid i \in \mathbb{N}\}$  as basis of the domain of  $\hat{\mathbf{T}}$ . According to eq. (11) we can express the operator  $\hat{\mathbf{T}}$  as:

$$\hat{\mathbf{T}} = \sum_{i=1}^{\infty} \sum_{n=1}^{\infty} |\xi_i\rangle (\mathbf{T}_{\{\xi_k\}}^{\{\xi_k\}})_{in} \langle P_n| = \sum_{i=1}^{\infty} \sum_{n=1}^{\infty} |\xi_i\rangle\langle\xi_i, P_n\rangle \langle P_n| \quad (24)$$

$$= \sum_{i=1}^{\infty} |\xi_i\rangle\langle\xi_i| \underbrace{\left(\sum_{n=1}^{\infty} |P_n\rangle\langle P_n|\right)}_{I_{d(\rho_n)}} = \sum_{i=1}^{\infty} |\xi_i\rangle\langle\xi_i| \quad (25)$$

For arbitrary elements  $|f\rangle$  of the domain of  $\hat{\mathbf{T}}$ , we can write:

$$\hat{\mathbf{T}}|f\rangle = \sum_{i=1}^{\infty} |\xi_i\rangle\langle\xi_i, f\rangle \quad (26)$$

Finally, we have to show: the restriction  $\hat{\mathbf{T}}|_{\mathcal{P}[\mathbf{N}]}$  is obtained by truncating eq. (26) after  $N$  terms. For an arbitrary vector  $|f\rangle \in \mathcal{P}[\mathbf{N}]$  with  $|f\rangle = \sum_{n=1}^N |P_n\rangle\langle P_n, f\rangle$ , we have:

$$\begin{aligned} \left[ \hat{\mathbf{T}}|_{\mathcal{P}[\mathbf{N}]} \right] |f\rangle &= \sum_{i=1}^N |\xi_i\rangle \langle \xi_i, f\rangle \\ &= \sum_{i=1}^N |\xi_i\rangle \sum_{n=1}^N \langle \xi_i, P_n \rangle \langle P_n, f\rangle \stackrel{\text{eq. (10)}}{=} \sum_{i=1}^N |\xi_i\rangle \sum_{n=1}^N \langle \xi_i, P_n \rangle \langle P_n, f\rangle \end{aligned} \quad (27)$$

$$|f\rangle \stackrel{\text{eq. (10)}}{=} \sum_{i=1}^N |\xi_i\rangle \langle \xi_i | \underbrace{\left( \sum_{n=1}^N |P_n\rangle \langle P_n| \right)}_{Id_{\mathcal{P}[\mathbf{N}]}} |f\rangle = \sum_{i=1}^N |\xi_i\rangle \langle \xi_i, f\rangle = \hat{\mathbf{T}} |f\rangle. \quad (28)$$

□

**General Remarks:**

- Sub-item 2 of this theorem corresponds to the following statement: The complete image due to domain vectors from the restriction  $f \in \mathcal{P}[\mathbf{N}]$  can be calculated using eq. (13) within  $N$  terms.
- The results of Theorem 1 do also hold for non-orthogonal bases of the domain. Starting from an expansion of the domain vectors in non-orthogonal basis functions (involving the Gramian matrix for the expansion coefficients) the same proof can be used to derive the desired properties of the moment expansion. Formal expressions involving non-orthogonal basis sets for image as well as domain are given in the Supporting Information.

**Remarks for the special case of  $\hat{\mathbf{T}} : V_{\text{pert}} \mapsto n^{\text{resp}}$ :**

- For the specific choice of Racah normalized irregular solid harmonics as basis functions for the expression of the perturbing potential, Scherrer et al.<sup>[3]</sup> showed in 2016, the existence of the moment expansion. In this article, we put the moment expansion on a more solid mathematical fundament and generalize it to arbitrary self-adjoint, positive, and compact operators (and arbitrary bases of the domain of these operators).
- The set  $\{P_n | n \in \mathbb{N}\}$  denotes the basis functions for the domain of  $\hat{\mathbf{T}}$ , which is different for every molecule. For practical application we aim to use standard basis sets of quantum chemistry such as solid harmonics or polynomials. The support of these standard basis sets is in general different from the compact support of the linear operator  $\hat{\mathbf{T}}$ . Thus inner products  $\langle P_i, V_{\text{pert}} \rangle$  are taken on the wrong vector space, which introduces some problems for actual calculations of the response density. The problem can be circumvented by expressing  $V_{\text{pert}}$  within the set of  $\{P_1, P_2, \dots, P_n\}$  and employing eq. (13), because the  $\xi_i$  vanish outside the domain of  $\hat{\mathbf{T}}$ . The overlap  $\langle \xi_i, V_{\text{pert}} \rangle \forall V_{\text{pert}} \in \mathcal{P}[\mathbf{N}]$  is automatically taken at the correct subspace. A detailed discussion of this problem is given in the Supporting Information.

**Direct Moment Expansion**

Ahler et al.<sup>[4]</sup> developed already an idea how to obtain the moment expanded states without calculation of thousands of eigenstates. In the next theorem we put this approach to the next level and show how to obtain the first  $N$  moment expanded states within  $N$  ab initio calculations (by means of the a cholesky decomposition of a  $N \times N$  matrix):

**Theorem 2:** Let  $\hat{\mathbf{T}}$  be the linear operator,  $\{|P_n\rangle | n \in \mathbb{N}\}$  the orthonormal basis of the domain of  $\hat{\mathbf{T}}$  and  $\{|\xi_n\rangle | n \in \mathbb{N}\}$  the moment expanded states from Theorem 1. We define

$$|\tilde{n}_i\rangle := \hat{\mathbf{T}} |P_i\rangle \quad (29)$$

and two matrices  $(\Theta^{\xi})$  and  $(\Theta^{\tilde{n}})$  with the elements:

$$(\Theta^{\xi})_{ij} := \langle \xi_i, P_j \rangle \quad (30)$$

$$(\Theta^{\tilde{n}})_{ij} := \langle \tilde{n}_i, P_j \rangle. \quad (31)$$

Then:

$$1) \quad |\xi_i\rangle = \sum_{j=1}^i \left( (\Theta^{\xi})^{-1} \right)_{ji} |\tilde{n}_j\rangle$$

$$2) \quad \Theta^{\tilde{n}} = (\Theta^{\xi})^T (\Theta^{\xi})$$

**Proof:**

**Proof of the first sub-item:**

We start with the definition from eq. (29):

$$\begin{aligned} |\tilde{n}_i\rangle &= \hat{\mathbf{T}} |P_i\rangle \stackrel{\text{eq. (26)}}{=} \sum_{j=1}^{\infty} |\xi_j\rangle \langle \xi_j, P_i\rangle \stackrel{\text{eq. (10)}}{=} \sum_{j=1}^i |\xi_j\rangle \langle \xi_j, P_i\rangle \\ &\stackrel{\text{eq. (30)}}{=} \sum_{j=1}^i |\xi_j\rangle (\Theta^{\xi})_{ji}. \end{aligned} \quad (32)$$

Equation (32) shows that  $\Theta^{\xi}$  is a change-of-basis matrix. The inverse of the change-of-basis matrix is the change-of-basis matrix of the inverse basis transformation:

$$|\xi_i\rangle = \sum_{j=1}^i |\tilde{n}_j\rangle \left( (\Theta^{\xi})^{-1} \right)_{ji}. \quad (33)$$

$\Theta^{\xi}$  is an upper triangular matrix (compare eq. (10) and eq. (30)), which in turn ensures the full rank of the matrix and hence the existence of the inverse  $(\Theta^{\xi})^{-1}$ .

**Proof of the second sub-item:**

The symmetry of  $\Theta^{\tilde{n}}$  is obtained from:

$$\begin{aligned} (\Theta^{\tilde{n}})_{ji} &\stackrel{\text{eq. (31)}}{=} \langle P_j, \tilde{n}_i \rangle \stackrel{\text{eq. (29)}}{=} \langle P_j, \hat{\mathbf{T}} P_i \rangle = \langle \hat{\mathbf{T}} P_j, P_i \rangle \stackrel{\text{eq. (29)}}{=} \langle \tilde{n}_j, P_i \rangle \\ &\stackrel{\text{eq. (31)}}{=} (\Theta^{\tilde{n}})_{ij} \end{aligned} \quad (34)$$

Multiplying eq. (32) from the left with  $\langle P_j |$  we obtain:

$$\underbrace{\langle P_j, \tilde{n}_i \rangle}_{(\Theta^{\tilde{n}})_{ji}} = \left\langle P_j, \sum_{n=1}^i |\xi_n\rangle \langle \xi_n, P_i \rangle \right\rangle = \sum_{n=1}^i \langle P_j, \xi_n \rangle \langle \xi_n, P_i \rangle$$

eq. (10)  $\sum_{n=1}^{\min\{i,j\}} \langle P_j, \xi_n \rangle \langle \xi_n, P_i \rangle$  (35)

eq. (30)  $\sum_{n=1}^{\min\{i,j\}} (\Theta^{\xi})_{nj} (\Theta^{\xi})_{ni} = \sum_{n=1}^{\min\{i,j\}} (\Theta^{\xi})_{jn}^T (\Theta^{\xi})_{ni}$  (36)

This is equivalent to the following matrix equation:

$$\Theta^{\tilde{n}} = (\Theta^{\xi})^T (\Theta^{\xi}) \quad (37)$$

Equation (37) actually defines a Cholesky decomposition of the symmetric matrix  $(\Theta^{\xi})$  into a product of an upper triangular matrix  $(\Theta^{\tilde{n}})$  and its transposed  $(\Theta^{\xi})^T$ . This upper triangular matrix is nothing but the moment matrix of the moment expanded states.

□

#### Remarks:

1. A **Cholesky decomposition** is a decomposition of a symmetric, positive-definite matrix into the product of a lower triangular matrix and its transpose. We assumed  $\hat{\mathbf{T}}$  as positive operator. Thus we can obtain the symmetry as well as the positive definiteness of  $\Theta^{\tilde{n}}$  from  $(\Theta^{\tilde{n}})_{ij} = \langle \tilde{n}_i, P_j \rangle = \langle P_i, \hat{\mathbf{T}}P_j \rangle$ .
2. In general  $\Theta^{\tilde{n}}$  and  $\Theta^{\xi}$  are infinite matrices. In the next remark, we show that only the  $N \times N$  restrictions of these matrices are necessary for the calculation of the first  $N$  moment expanded states  $\{|\xi_1\rangle, |\xi_2\rangle, \dots, |\xi_N\rangle\}$ .
3. According to the last theorem, we want to summarize the effort for the calculation of the first  $N$  moment expanded states  $\{|\xi_1\rangle, |\xi_2\rangle, \dots, |\xi_N\rangle\}$ . We have to calculate:

- $N$  evaluations of  $\hat{\mathbf{T}}$  according to  $\left\{ \begin{array}{c} |\hat{\mathbf{T}}P_1\rangle, |\hat{\mathbf{T}}P_2\rangle, \dots, |\hat{\mathbf{T}}P_N\rangle \\ |\tilde{n}_1\rangle, |\tilde{n}_2\rangle, \dots, |\tilde{n}_N\rangle \end{array} \right\}$
- the  $N \times N$  restriction of the inverse of  $(\Theta^{\xi})$ . This  $N \times N$  restriction can be calculated by the Cholesky decomposition of the  $N \times N$  restriction of  $\Theta^{\tilde{n}}$ . The formation of  $\Theta^{\tilde{n}} \in \mathbb{R}^{N \times N}$  requires only the  $\{|\tilde{n}_1\rangle, |\tilde{n}_2\rangle, \dots, |\tilde{n}_N\rangle\}$  (mentioned in the first item).  
The numerical determination of the first  $N$  moment expanded states requires of the order of  $N$  applications of  $\hat{\mathbf{T}}$  (and linear algebra operations with negligible effort).

#### Algorithm of the direct moment expansion

Based on the second theorem, we want to report an explicit protocol for the calculation of the moment expanded states:

- Choose a basis  $\{ |P_k\rangle \mid k \in \mathbb{N} \}$  for the domain of the linear operator  $\hat{\mathbf{T}}$ .

- Compute the set of functions  $\{ |\tilde{n}_i\rangle := \hat{\mathbf{T}} |P_i\rangle \mid i \in \{1, \dots, N\} \}$ .
- Calculation of an  $N \times N$  matrix with elements  $(\Theta^{\tilde{n}})_{ij} = \langle \tilde{n}_i, R_j \rangle$ .

For the actual transformation two steps are necessary:

**First step:** Cholesky decomposition (via library algorithms e.g., lapack) of  $\Theta^{\tilde{n}} = (\Theta^{\xi})^T (\Theta^{\xi})$ .

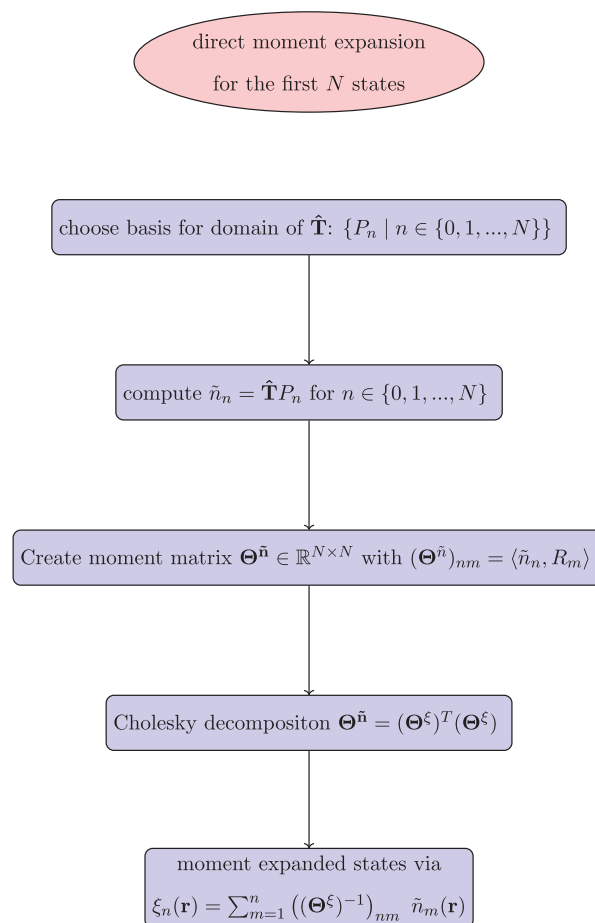
**Second step:** The inverse of the  $N \times N$  moment expanded moment matrix  $(\Theta^{\tilde{n}})^{-1}$  transforms the set  $\{|\tilde{n}_i| i \leq N\}$  into the moment expanded states  $\{|\xi_i| i \leq N\}$  according to eq. (33).

A schematic representation of the direct moment expansion is given in Figure 1.

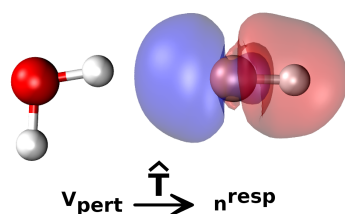
## Application of Theorem 1 and 2 to the Linear Density-Density Response Function

### General description

The calculation of the electronic response due to a perturbation potential is common to numerous important applications from spectroscopy to intermolecular interactions (Figure 2). One example is fragmentation methods, where a large system is



**Figure 1.** Schematic representation of the calculation of the first  $N$  moment expanded states using the Direct Moment Expansion. [Color figure can be viewed at wileyonlinelibrary.com]



**Figure 2.** Principal illustration of the response density of the water molecule (right) due to a perturbing water molecule (left). [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

partitioned into smaller fragments that are more efficient to compute. The mutual interaction between the fragments can be realized via a perturbation theory approach, where each fragment perturbs the other ones via its electrostatic potential (acting as a perturbation potential).<sup>[19–26]</sup> Following these ideas, many embedding or subsystem density functional theory (DFT) methods<sup>[27]</sup> like frozen density embedding<sup>[28–34]</sup> or embedded correlated wave functions methods<sup>[34–37]</sup> were developed in the literature.

For the specific case of interacting molecules, a fragmentation approach, which considers the individual molecules as elementary subsystems is suited. Then, the intermolecular interaction energy has to be obtained via post processing of the results for the isolated molecules. In particular for the electrostatic interaction energy, methods like multipole expansions,<sup>[38–44]</sup> density fitting<sup>[45–48]</sup> or perturbation theories<sup>[19–26]</sup> are utilized.

Accurate electrostatic interaction energies should take polarization effects into account. For the specific choice of the molecular linear density-density response function as kernel  $\chi(\mathbf{r}, \mathbf{r}')$  the integral transform in eq. (1) constitutes an operator  $\hat{\mathbf{T}}$ :  $V_{\text{pert}}(\mathbf{r}) \mapsto n^{\text{resp}}(\mathbf{r})$  which maps perturbing potentials  $V_{\text{pert}}(\mathbf{r})$  to response densities  $n^{\text{resp}}(\mathbf{r})$ . These response densities fully incorporate polarization effects.

In Theorem 1 and 2 we prove the validity of an efficient protocol for the calculation of the density response due to approximated/truncated perturbing potentials  $V_{\text{pert}}(\mathbf{r}) \in \mathcal{P}[\mathbf{N}] := \text{Span}(P_1(\mathbf{r}), P_2(\mathbf{r}), \dots, P_N(\mathbf{r}))$  using the operator  $\hat{\mathbf{T}}$ . We present a generalized representation (i.e., independent of the specific shape of the perturbing potential) of the response density, which can be computed and stored with a reasonable computational effort by means of the moment expansion. The latter is designed to optimally reflect the shape of physically meaningful (yet arbitrary) perturbation potentials which will be generated by nearby molecules.

### Efficiency gain due to moment expansion

The quantum chemical calculation of the (linear) response of a molecule to an external perturbation potential can be done (1) via a conventional self-consistent perturbation theory calculation<sup>[6–12]</sup> or (2) via application of the generalized density-density response function  $\chi(\mathbf{r}, \mathbf{r}')$ . For the determination of a single response density for one given perturbation potential (or at least a limited number of such perturbations), the direct calculation (1) is more efficient than our approach (2) due to

the comparably high computational cost of obtaining  $\chi(\mathbf{r}, \mathbf{r}')$  explicitly. The actual application of  $\chi(\mathbf{r}, \mathbf{r}')$  to any specific perturbation potential in turn is quite small, because only a limited number of scalar products of the specific perturbation potential function with the basis functions for  $\chi(\mathbf{r}, \mathbf{r}')$  is necessary (in particular, no self-consistency cycle is involved any more). Thus, the advantage of using an explicit representation of the density-density response lies in *repeated* applications of response calculations for numerous *different* perturbation potentials for the same molecule, that is, when only one initial determination of  $\chi(\mathbf{r}, \mathbf{r}')$  is required. For the original eigensystem formulation of  $\chi(\mathbf{r}, \mathbf{r}')$ , the break-even point depends on the number of eigenstates that are needed for an accurate representation, typically around 1000–5000 (e.g., for a water molecule). The number of required self-consistent perturbation theory cycles for the initial determination of  $\chi(\mathbf{r}, \mathbf{r}')$  roughly equals the number of desired eigenstates.

The efficiency gain of the moment expansion stems from the reduction of that number (the number of eigenstates required for an accurate representation of  $\chi(\mathbf{r}, \mathbf{r}')$ ) by about two orders of magnitude. In our new representation, the focus is shifted from optimally representing the local properties of  $\chi(\mathbf{r}, \mathbf{r}')$  toward optimally adapting the representation to the shape of the perturbation potential  $V_{\text{pert}}(\mathbf{r})$ . Hence, our moment expansion shifts the break-even point compared to conventional self-consistent perturbation theory calculations from 1000 to 5000 toward a number of about 10–100 conventional calculations. While this figure of course still means that for a small number of response calculations, the conventional path is more advantageous, the perspective changes when a very large number of response calculations for the same molecule is required. A case in point for this situation is molecular dynamics simulations of a molecular liquid. From the perspective of a given molecular in the liquid, each new molecular dynamics frame constitutes a new chemical environment, to which the molecular density will respond. This response is currently computed self-consistently (Born-Oppenheimer molecular dynamics) or iteratively (Car-Parrinello scheme). Our approach could be used to switch the determination of the density response to a new molecular dynamics step to a non-self-consistent way at considerably reduced computational cost without sacrificing quantum chemical accuracy.

What is the physical interpretation of this dimensionality reduction? The spatial shape of the eigenstates of the density-density response function resembles a kind of symmetrized set of molecular orbitals. The eigenstates often appear like atom-centered (or bond-centered) functions of localized nature and strong nodal structure, that is, strong local oscillations. Furthermore, subsequent eigenstates often cover the same spatial area but just show a higher nodal structure. From a more general perspective, an analogy to fitting a function via a polynomial interpolation comes into mind: Fitting  $N$  data points yields an interpolation polynomial of order  $N$ , which however is rarely the optimal choice for a relatively smooth function. This problem of the original eigensystem representation of the density-density response function is cured by the moment expansion formulation.



The specific impact of the two theorems for the linear density-density response function is discussed in the following.

### Relevance of Theorem 1

The eigensystem representation offers a straightforward but inefficient possibility for the calculation of the density response. The infinite expansion in eq. (2) can be truncated after  $M$  eigenstates to the  $M$  largest eigenvalues ( $M \sim 10^3 - 10^4$ ),<sup>[1,2]</sup> as the sequence of the eigenvalues decays to zero. In order to obtain the response density due to a perturbing potential  $V_{\text{pert}}(\mathbf{r}) \in \text{Span}(P_1(\mathbf{r}), P_2(\mathbf{r}), \dots, P_N(\mathbf{r}))$ , the calculation of  $M$  dot products/projections between the eigenstates and the specific perturbing potential  $V_{\text{pert}}(\mathbf{r})$  is necessary in eq. (38).

$$n^{\text{resp}}(\mathbf{r}) = \sum_{i=1}^M \chi_i(\mathbf{r}) \lambda_i \langle \chi_i, V_{\text{pert}} \rangle \quad (38)$$

Utilizing the moment expanded states resulting from Theorem 1, a drastic reduction of the dimensionality of the expression of the response density is achieved. The full density response due to  $V_{\text{pert}}(\mathbf{r}) \in \text{Span}(P_1(\mathbf{r}), P_2(\mathbf{r}), \dots, P_N(\mathbf{r}))$  can be calculated within  $N$  moment expanded states with  $N \ll M$  (subitem 2 of Theorem 1):

$$n^{\text{resp}}(\mathbf{r}) = \sum_{n=1}^N \xi_n(\mathbf{r}) \langle \xi_n, V_{\text{pert}} \rangle. \quad (39)$$

### Relevance of Theorem 2

The second theorem gives rise to a much more efficient algorithm for the calculation of the moment-expanded states (referred to as direct moment expansion). For the brute force construction of the transformation  $\mathbf{Q}$  derived in theorem 1 (which corresponds to the algorithm of the moment expansion published in 2016<sup>[3]</sup>) a large number of  $M$  eigenstates are calculated, followed by  $M \times N$  Givens rotations to condense the physically information.

Via the direct moment expansion, we are able to calculate the first  $N$  moment expanded states from the first  $N$  direct density expanded states (by a Cholesky decomposition of a  $N \times N$  matrix). Each direct density expanded state can be computed via a single quantum chemical perturbation theory calculation. A concrete, step-by-step protocol for the calculation of the direct moment expansion is given in the previous section.

## Numerical Example: Density Response of a Water Molecule

In this section, we will demonstrate that the linear operator defined by the linear density-density response function fulfills the conditions for the application of the moment expansion, that is, we will show that the relevant domain vectors (perturbing potentials arising from neighboring molecules) can be expressed within a finite number of basis functions.

We choose the perturbative effect of one water molecule on an adjacent one as an elementary example (Figure 2). For

simplicity, we use the Hartree potential using the partial charges of one of the most common force field water models (TIP3P) as perturbation, while the perturbed water molecule is represented within standard density functional theory, complemented with the density-density response function computed using density functional perturbation theory calculations (as practical realizations of the operator  $\hat{\mathbf{T}}$ ). The Hartree potential of the TIP3P partial charges reads:<sup>[49]</sup>

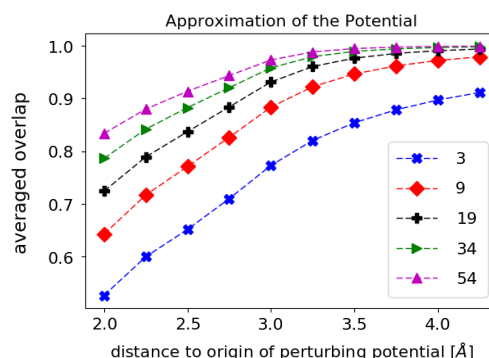
$$V_{\text{TIP3P}}^{\text{frag}}(\mathbf{r}) = \sum_{\lambda=1}^{N_n} \frac{Q_{\lambda}^{\text{partial}}}{|\mathbf{r}-\mathbf{R}_{\lambda}|} \text{erf}\left(\frac{|\mathbf{r}-\mathbf{R}_{\lambda}|}{\sqrt{2}\sigma}\right). \quad (40)$$

For 10 different distances to the origin of the TIP3P water molecule, we choose six representative points (spanning an octahedron) and expanded the perturbing potential around these positions using up to 54 monomials. Subsequently, we calculated the overlap of the polynomial approximated potential and the Hartree potential on a  $4 \text{ \AA} \times 4 \text{ \AA} \times 4 \text{ \AA}$  cube centered at the origin of the potential expansion—note, that the origin of the perturbing potential (center of mass of a TIP3P water molecule) and the origin of the potential expansion (center of the responding water molecule) do not coincide—the overlap  $O_{|f\rangle}^{|\tilde{f}\rangle}$  of a function  $|\tilde{f}\rangle$  and  $|f\rangle$  was taken according to:

$$O_{|f\rangle}^{|\tilde{f}\rangle} = \frac{\langle \tilde{f}, f \rangle}{\langle f, f \rangle} \quad (41)$$

Finally, we averaged the six overlaps of the Hartree potential and the expanded potential for each distance to the origin and presented the results with respect to different numbers of basis functions in Figure 3.

Furthermore, we demonstrated that we can use the moment expansion in order to express the full density response of water



**Figure 3.** Averaged and normalized overlap of the Hartree potential and the potential obtained from the expansion in monomials. The different colors refer to different number of basis functions. Please note, that the origin of the perturbing potential (center of mass of a TIP3P water molecule) and the origin of the potential expansion do not coincide! The overlap was taken at a  $4 \text{ \AA} \times 4 \text{ \AA} \times 4 \text{ \AA}$  cube centered at the origin of the expansion of the perturbing potential. For each distance to the spatial origin of the potential expansion, we averaged over six overlaps (corresponding to six different spatial origins of the expansion of the perturbing potential, which span an octahedron). [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

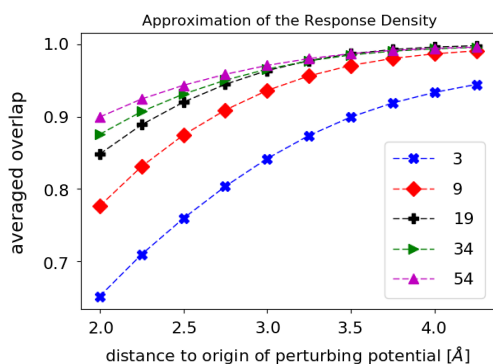
molecule within a few moment expanded states. For each distance to the perturbing TIP3P water molecule, we reused the octahedral alignment of the origins of the expansion of the perturbing potentials from the previous paragraph as center of masses of a responding water molecule.

Due to the perturbing potential, we calculated the response density of these water molecules via the moment-expanded states and via Density Functional Perturbation Theory (DFPT) (as reference calculations).

We calculated the overlap of the resulting densities and averaged all overlaps corresponding to the same distance to the perturbing TIP3P water molecule. The obtained values are depicted in Figure 4. From Figure 3 and Figure 4, we can deduce that already within nine states we can obtain reasonable good approximations of response densities or perturbing potentials. For small intermolecular distances, we can systematically improve the approximation by addition of further moment expanded states/basis functions. Within 54 states we obtain more than 90% of the response density/ the perturbing potential, even in the regime of the first maximum of the water-water radial distribution function ( $\sim 2.5$  Å). This is a remarkable reduction compared the employment of several of thousands eigenstates of the linear density-density response function.

## Conclusion

In this article, we generalize the concept of the moment expansion to arbitrary compact, positive and self-adjoint linear operators  $\hat{T}$ . This class of linear operators includes all integral transforms with a positive, symmetric and continuous kernel  $\chi(\mathbf{r}, \mathbf{r}')$ . We show that only comparably weak conditions have to be fulfilled (such as functions of the domain has to be Taylor expandable in an local environment) in order to allow for an efficient explicit representation (i.e., calculation and memory storage) of the linear operator  $\hat{T}$ , which enables the evaluation



**Figure 4.** Averaged normalized overlap of the response density (of a water molecule) from the moment expansion and a reference DFPT calculation (due to the perturbing potential generated by the TIP3P water). The different colors refer to different number of moment-expanded states. We averaged over six density overlaps for each distance between the perturbing TIP3P water molecule and the responding water molecule. The six positions of the responding water molecules span an octahedron for each distance. Please note, that we compare the density differences ( $\cong$  response densities) due to a perturbation and not the resulting overall densities. [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

of  $\hat{T}$  with a drastically reduced costs compared to its eigen-system representation.

We have demonstrated that the moment expansion can be considered as a **QR** decomposition of the transformation matrix. The straightforward calculation of the infinite orthogonal matrix **Q** would in particular include the calculation of thousands of eigenvalues and eigenstates. We presented an efficient version of the moment expansion (termed as direct moment expansion), which allows the calculation of the first  $N$  moment expanded states at the computational costs of about  $N$  evaluations of  $\hat{T}$ .

For the specific case of the linear density-density response functions as kernel  $\chi(\mathbf{r}, \mathbf{r}')$ , we presented a numerical example for the efficient calculations of converged molecular response densities via the moment expansion. We demonstrated that perturbing potentials from neighboring molecules could be truncated within a few ( $\cong 34$ ) basis functions.

The actual scope of our work exceeds far beyond this specific example. All Green's-function based formalisms in physics and chemistry have the problem that the Green's function itself can rarely be expressed/computed explicitly because of its dimensionality. In most cases, it is merely considered as a handy nomenclature for formally expressing that a linear differential equation can be solved (inverted) for a specific right-hand-side function. Our work shows that this problem can possibly be overcome in many situations. While we do not believe that we have a comprehensive overview of all possible applications ourselves, an indicative list of quantum chemical problems would include (a) fragmentation/embedding approaches, (b) van-der-Waals interaction energies via a Lundqvist expression<sup>[50]</sup> and correlation energies within the random phase approximation, and (c) resolution of identity approaches. Possible fields of application beyond this directly related area are (d) polarizable force fields, (e) optimal control theory, and (f) signal processing (impulse response function). However, we believe that our Ansatz may also find applications in the wider area of "fundamental solutions" (in mathematical terms) where spectral theory/Fredholm integral equations are applicable.

**Keywords:** density-density response function · molecular interaction · density functional perturbation theory · linear compact operator

How to cite this article: C. Dreßler, A. Scherrer, P. Ahlert, D. Sebastiani. *J. Comput. Chem* **2019**, *40*, 2712–2721. DOI: 10.1002/jcc.26046

Additional Supporting Information may be found in the online version of this article.

## Appendix

### Computational details

The moment expansion has been implemented in our development version of the CPMD<sup>[51]</sup> electronic structure package. The calculations have been performed using Density Functional



- Perturbation Theory<sup>[8,11,12,52,53]</sup> with Troullier-Martins<sup>[54]</sup> pseudo potentials in the Becke<sup>[55]</sup> and Lee-Yang-Parr<sup>[56]</sup> approximation for the exchange correlation kernel. The perturbing potential was expressed by a polynomial expansion. The basis was formed by the set of polynomials  $x^l y^m z^n$  with  $l, m, n \in \mathbb{N}$ , which fulfill the condition  $l + m + n \leq i_{max}$ . The largest basis set consisted out of 54 polynomials ( $i_{max} = 5$ ). The calculation of the reference density response with CPMD was done for relaxed water geometry. We used a plane wave basis for all involved quantities, be it the electronic density, the perturbing potentials or the moment-expanded states. The real space representation naturally gives rise to a regular grid. For our chosen parameters, the grid increment is about 0.08 Angstroms, whereas the spread of the Gaussians in eq. (40) is chosen as  $\sigma = 0.5a_0$ .
- [1] A. Scherrer, V. Verschinin, D. Sebastiani, *J. Chem. Theo. Comput.* **2012**, pMID: 26592873, 8, 106. <https://doi.org/10.1021/ct200695y>.
- [2] A. C. Ihrig, A. Scherrer, D. Sebastiani, *J. Chem. Phys.* **2013**, 139, 094102. <https://doi.org/10.1063/1.4819070>.
- [3] A. Scherrer, D. Sebastiani, *J. Comput. Chem.* **2016**, 37, 665.
- [4] P. Ahlert, A. Scherrer, C. Dressler, D. Sebastiani, *Europ. Phys. J. B* **2018**, 91, 94.
- [5] W. Dirk, *Funktionalanalysis*; Springer Berlin Heidelberg: Berlin, Heidelberg, **2011**. [https://doi.org/10.1007/978-3-642-21017-4\\_6](https://doi.org/10.1007/978-3-642-21017-4_6).
- [6] P. Giannozzi, S. de Gironcoli, P. Pavone, S. Baroni, *Phys. Rev. B* **1991**, 43, 7231.
- [7] X. Gonze, *Phys. Rev. A* **1995**, 52, 1086.
- [8] X. Gonze, *Phys. Rev. A* **1995**, 52, 1096.
- [9] X. Gonze, D. C. Allan, M. P. Teter, *Phys. Rev. Lett.* **1992**, 68, 3603.
- [10] X. Gonze, J.-P. Vigneron, *Phys. Rev. B* **1989**, 39, 13120.
- [11] S. Baroni, S. de Gironcoli, A. del Corso, P. Giannozzi, *Rev. Mod. Phys.* **2001**, 73, 515.
- [12] A. Putrino, D. Sebastiani, M. Parrinello, *J. Chem. Phys.* **2000**, 113, 7102.
- [13] H. F. Wilson, F. m. c. Gygi, G. Galli, *Phys. Rev. B* **2008**, 78, 113303.
- [14] H. F. Wilson, D. Lu, F. m. c. Gygi, G. Galli, *Phys. Rev. B* **2009**, 79, 245106.
- [15] J. D. Jackson, *Classical electrodynamics*, 3rd ed., Wiley, New York, NY, **1999**.
- [16] S. Chakrabarti, D. P. Dewangan, *J. Phys. B Atom. Mole. Opt. Phys.* **1995**, 28, L769.
- [17] A. C. Hansen, *J. Funct. Anal.* **2008**, 254, 2092.
- [18] P. Deift, L. Li, C. Tomei, *J. Funct. Anal.* **1985**, 64, 358.
- [19] B. Jeziorski, R. Moszynski, K. Szalewicz, *Chem. Rev.* **1994**, 94, 1887. <https://doi.org/10.1021/cr00031a008>.
- [20] A. Heßelmann, *J. Chem. Theo. Comput.* **2018**, pMID: 29566325, 14, 1943. <https://doi.org/10.1021/acs.jctc.7b01233>.
- [21] D. M. Benoit, D. Sebastiani, M. Parrinello, *Phys. Rev. Lett.* **2001**, 87, 226401.
- [22] A. Heßelmann, G. Jansen, *Chem. Phys. Lett.* **2003**, 367, 778.
- [23] D. Rocca, D. Lu, G. Galli, *J. Chem. Phys.* **2010**, 133, 164109. <https://doi.org/10.1063/1.3494540>.
- [24] A. J. Misquitta, B. Jeziorski, K. Szalewicz, *Phys. Rev. Lett.* **2003**, 91, 033201.
- [25] A. Hesselmann, *Chem. A Eur. J.* **2011**, pMID: 21806071, 115, 11321. <https://doi.org/10.1021/jp205031e>.
- [26] D. Lu, H.-V. Nguyen, G. Galli, *J. Chem. Phys.* **2010**, 133, 154110. <https://doi.org/10.1063/1.3494541>.
- [27] A. Goez, J. Neugebauer, Embedding methods in quantum chemistry. In *Frontiers of Quantum Chemistry*; M. J. Wójcik, H. Nakatsuji, B. Kirtman, Y. Ozaki, Eds., Springer, Singapore, Singapore, **2018**, p. 139.
- [28] C. R. Jacob, J. Neugebauer, L. Visscher, *J. Comput. Chem.* **2008**, 29, 1011.
- [29] T. A. Wesolowski, A. Warshel, *J. Phys. Chem.* **1993**, 97, 8050.
- [30] D. G. Artiukhin, C. R. Jacob, J. Neugebauer, *J. Chem. Phys.* **2015**, 142, 234101. <https://doi.org/10.1063/1.4922429>.
- [31] T. A. Wesolowski, One-electron equations for embedded electron density: Challenge for theory and practical payoffs in multi-level modelling of complex polyatomic systems. In *Computational Chemistry: Reviews of Current Trends*; WORLD SCIENTIFIC: Singapore, **2006**; p. 1.
- [32] C. R. Jacob, J. Neugebauer, *Comput. Mole. Sci.* **2014**, 4, 325.
- [33] P. Cortona, *Phys. Rev. B* **1991**, 44, 8454.
- [34] F. R. Manby, M. Stella, J. D. Goodpaster, T. F. Miller, *J. Chem. Theo. Comput.* **2012**, pMID: 22904692, 8, 2564. <https://doi.org/10.1021/ct300544e>.
- [35] F. Libisch, C. Huang, E. A. Carter, *Acc. Chem. Res.* **2014**, pMID: 24873211, 47, 2768. <https://doi.org/10.1021/ar500086h>.
- [36] X. Zhang, E. A. Carter, *J. Chem. Theo. Comput.* **2019**, 15, 949. <https://doi.org/10.1021/acs.jctc.8b00990>, <https://doi.org/10.1021/acs.jctc.8b00990>.
- [37] S. Sharifzadeh, P. Huang, E. A. Carter, *Chem. Phys. Lett.* **2009**, 470, 347.
- [38] A. D. Buckingham, Permanent and induced molecular moments and long-range intermolecular forces. In *Advances in Chemical Physics*; John Wiley and Sons, Ltd: Hoboken, New Jersey, **2007**; p. 107.
- [39] A. Stone, *Chem. Phys. Lett.* **1981**, 83, 233.
- [40] J.-P. Piquemal, N. Gresh, C. Giessner-Prettre, *Chem. A Eur. J.* **2003**, pMID: 26313624, 107, 10353. <https://doi.org/10.1021/jp035748t>.
- [41] G. Naray-Szabo, G. G. Ferenczy, *Chem. Rev.* **1995**, 95, 829. <https://doi.org/10.1021/cr00036a002>.
- [42] J. G. Ángyán, C. Chipot, F. Dehez, C. Hättig, G. Jansen, C. Millot, *J. Comput. Chem.* **2003**, 24, 997.
- [43] D. Elking, T. Darden, R. J. Woods, *J. Comput. Chem.* **2007**, 28, 1261.
- [44] R. J. Wheatley, *Mole. Phys.* **1993**, 79, 597. <https://doi.org/10.1080/00268979300101481>.
- [45] G. G. Hall, C. M. Smith, *Int. J. Quant. Chem.* **1984**, 25, 881.
- [46] D. M. Elking, G. A. Cisneros, J.-P. Piquemal, T. A. Darden, L. G. Pedersen, *J. Chem. Theo. Comput.* **2010**, 6, 190. <https://doi.org/10.1021/ct900348b>.
- [47] K. Eichkorn, O. Treutler, H. Öhm, M. Häser, R. Ahlrichs, *Chem. Phys. Lett.* **1995**, 242, 652.
- [48] G. A. Cisneros, J.-P. Piquemal, T. A. Darden, *J. Chem. Phys.* **2005**, 123, 044109. <https://doi.org/10.1063/1.1947192>.
- [49] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, *J. Chem. Phys.* **1983**, 79, 926. <https://doi.org/10.1063/1.445869>.
- [50] K. Berland, V. R. Cooper, K. Lee, E. Schröder, T. Thonhauser, P. Hyldgaard, B. I. Lundqvist, *Rep. Prog. Phys.* **2015**, 78, 066501.
- [51] CPMD 3.15.3, Copyright IBM Corp 1990–2008, Copyright MPI für Festkörperforschung Stuttgart 1997–2001 Available at: <http://www.cpmid.org/>.
- [52] W. Kohn, L. J. Sham, *Phys. Rev.* **1965**, 140, A1133.
- [53] P. Hohenberg, W. Kohn, *Phys. Rev.* **1964**, 136, B864.
- [54] N. Troullier, J. L. Martins, *Phys. Rev. B* **1991**, 43, 1993.
- [55] A. D. Becke, *Phys. Rev. A* **1988**, 38, 3098.
- [56] C. Lee, W. Yang, R. G. Parr, *Phys. Rev. B* **1988**, 37, 785.

Received: 17 May 2019

Revised: 24 July 2019

Accepted: 25 July 2019

Published online on 29 August 2019

### 3.2.3 Article VIII: Iterative Approach for the Moment Representation of the LDDRF.

Paul Ahlert, Arne Scherrer, Christian Dreßler, and Daniel Sebastiani.

Iterative approach for the moment representation of the density-density response function.

*The European Physical Journal B*, 91(6):94, **2018**.

A. Scherrer developed the idea of using symmetry-adapted starting functions for the iterative response calculation scheme and implemented the code. P. Ahlert and I were responsible for benchmarking and application calculations. D. Sebastiani initiated and supervised the project. While P. Ahlert wrote the main part of the manuscript, I was responsible for section 3 of the article.

# Iterative approach for the moment representation of the density-density response function<sup>\*</sup>

Paul Ahlert<sup>a</sup>, Arne Scherrer, Christian Dressler, and Daniel Sebastiani

Institute of Chemistry, Martin-Luther-University Halle-Wittenberg, Von-Danckelmann-Platz 4, 06120 Halle (Saale), Germany

Received 23 January 2018 / Received in final form 16 March 2018

Published online 4 June 2018 – © EDP Sciences, Società Italiana di Fisica, Springer-Verlag 2018

**Abstract.** The linear density-density response function  $\chi(\mathbf{r}, \mathbf{r}')$  can be transformed from its Eigensystem representation into a computationally more efficient moment expansion representation using a suitable unitary transformation. Here, we propose an iterative approach for the direct calculation of this moment representation without resorting to either the direct-space or the conventional Eigensystem representation.

## 1 Introduction

When atoms and molecules are exposed to various perturbations from the environment, e.g. electric fields from surrounding molecules, they react to this external perturbation with a response of their electronic quantum state. For many small perturbations, perturbation theory can be applied to determine the linear response to the perturbation. This linearity also applies to the electron density in density functional theory (DFT) [1–3], giving rise to density functional perturbation theory (DFPT) [4–9]. The central quantity within linear response theory is the linear density-density response function  $\chi$ , which connects the (arbitrary) perturbing potential to the induced electronic response density:

$$n_{\text{resp}}(\mathbf{r}) = \int \chi(\mathbf{r}, \mathbf{r}') V_{\text{pert}}(\mathbf{r}') d^3 r', \quad (1)$$

where  $n_{\text{resp}}(\mathbf{r})$  denotes the response density and  $V_{\text{pert}}(\mathbf{r}')$  denotes the perturbing potential. One possible representation of  $\chi(\mathbf{r}, \mathbf{r}')$  is its spectral decomposition in eigenvectors  $\chi_i(\mathbf{r})$  and eigenvalues  $\lambda_i$  [10]:

$$\chi(\mathbf{r}, \mathbf{r}') = \sum_i^{\infty} \chi_i(\mathbf{r}) \lambda_i \chi_i(\mathbf{r}') \quad (2)$$

The spectral decomposition can be obtained numerically by a Lanczos diagonalisation technique [10–15].

<sup>\*</sup> Contribution to the Topical Issue “Special issue in honor of Hardy Gross”, edited by C.A. Ullrich, F.M.S. Nogueira, A. Rubio, and M.A.L. Marques.

<sup>a</sup> e-mail: [daniel.sebastiani@chemie.uni-halle.de](mailto:daniel.sebastiani@chemie.uni-halle.de)

The sum in equation (2) can be expressed in the form of a matrix product according to

$$\chi(\mathbf{r}, \mathbf{r}') = \underline{Q} \underline{A} \underline{Q}^T \quad (3)$$

with a vector of functions  $\underline{Q} = [\chi_1(\mathbf{r}), \chi_2(\mathbf{r}), \dots]$  and a diagonal real matrix  $\underline{A} = [\lambda_1, \lambda_2, \dots]$ .

Although the sum in equation (2) is in principle infinite, the spectrum decays, so that it can be truncated after a sufficient number of states. Nevertheless, several thousand states are necessary in order to obtain a converged response density even for small molecules, as we and others have reported [10–16]. This large number of eigenstates presents two problems: it makes, firstly, the computation of the eigenstates very demanding and, secondly, the evaluation of equation (1) rather cumbersome. The latter problem can be solved by a change of representation of the linear density-density response function, generating new states that allow to separate the contributions from different multipole moments of the perturbing potential to the response density. As a consequence, the physical information of the linear density-density response function is condensed in few states [16]. This moment expanded representation provides also a means to solve the former problem, which shall be shown in this paper. A condensed, computationally efficient representation of the density-density susceptibility tensor is crucial for potential applications of intermolecular interactions in the form of polarizable force fields [17–22], but could be of equal interest for the calculation of intermolecular dispersion energies [23–25]. Complementary relevance stems from the area of sum-frequency spectroscopy [26–32], and recently also from “alchemical” morphing between elements [33,34].

The change of representation is achieved by a unitary transformation of the eigenstates that yields irreducible representations of the linear density-density response

function with respect to rotations in SO(3). In the following, we first give a brief summary of the moment expanded representation of the linear density-density response function and show afterwards how this representation drastically reduces the computational effort necessary for the calculation of the response function.

## 2 Moment expansion

The central idea of the transformation of the density density response function into its moment expansion is that it is possible to determine a unitary transformation for the spectral decomposition of the response function (“Eigensystem representation”) which yields a representation that condenses its physical information in only a few states [16].

In the first step, we expand the perturbing potential in a Laplace series using solid harmonic functions [35]. While other choices such as non-polynomial radial functions (Gaussian, exponential, ...) or a direct cartesian Taylor expansion) are equally possible, we have found solid harmonics to be the most suitable option thanks to their mathematical properties, in particular the existence of a fast and stable method for rotating them in space [35] and the fact that their multipole moments are traceless.

$$V_{\text{pert}}(\mathbf{r}') = \sum_{l=0}^{\infty} \sum_{m=-l}^l V_l^m R_l^m(\mathbf{r}'), \quad (4)$$

where

$$V_l^m = \langle V_{\text{pert}} | R_l^m \rangle. \quad (5)$$

Insertion of equations (2) and (4) in (1) yields:

$$n_{\text{resp}}(\mathbf{r}) = \sum_{i=1}^{\infty} \chi_i(\mathbf{r}) \left( \sqrt{\lambda_i} \sum_{l=1}^{\infty} \sum_{m=-l}^l V_l^m \Theta_i^{lm} \right), \quad (6)$$

where we have defined multipole moments

$$\Theta_i^{lm} = \sqrt{\lambda_i} \langle \chi_i | R_l^m \rangle. \quad (7)$$

In the next step, we introduce a new representation of  $\chi(\mathbf{r}, \mathbf{r}')$  by a unitary transformation of the eigenstates in equation (3):

$$\chi(\mathbf{r}, \mathbf{r}') = \underline{Q} \underline{A}^{\frac{1}{2}} \underline{U} \underline{U}^T \underline{A}^{\frac{1}{2}} \underline{Q}^T =: \underline{R} \underline{R}^T \quad (8)$$

where  $\underline{R} = [(\xi_1(\mathbf{r}), \xi_2(\mathbf{r}), \dots)]$  with transformed states

$$\xi_i(\mathbf{r}) := \sum_j^{\infty} U_{ji} \sqrt{\lambda_j} \chi_j(\mathbf{r}). \quad (9)$$

The elements of the transformation matrix  $\underline{U}$  are yet to be determined. The aim of this step is to obtain a new set of states  $\{\xi_i(\mathbf{r})\}$  which exhibit symmetries of angular momenta with adequate quantum numbers  $(l, m)$ . As we

have shown previously [16], it is possible to assign a suitable couple  $i \mapsto (l, m)$  to each  $\xi_i(\mathbf{r})$  so that a new set of functions  $\{\xi_l^m(\mathbf{r})\}$  emerges which satisfies

$$\Xi_{ll'}^{mm'} := \langle \xi_l^m | R_{l'}^{m'} \rangle = 0 \quad \text{if } l > l' \text{ or } m > m'. \quad (10)$$

This property means that a given  $\{\xi_l^m(\mathbf{r})\}$  has only contributions from solid harmonics with quantum numbers *larger than* its own ones  $(l, m)$ . The new representation of  $\chi(\mathbf{r}, \mathbf{r}')$  in terms of the  $\xi_l^m(\mathbf{r})$ , analogous to equation (2), follows from equation (8):

$$\chi(\mathbf{r}, \mathbf{r}') = \sum_l^{\infty} \sum_{m=-l}^l \xi_l^m(\mathbf{r}) \xi_l^m(\mathbf{r}'). \quad (11)$$

Insertion of this new representation for  $\chi(\mathbf{r}, \mathbf{r}')$  into equation (1) and using the Laplace expansion of  $V_{\text{pert}}$  according to equation (4) yields the response density

$$n_{\text{resp}}(\mathbf{r}) = \sum_{l'=1}^{\infty} \sum_{l=1}^{l'} \sum_{m'=-l'}^{l'} \sum_{m=-l}^l \xi_l^m(\mathbf{r}) \Xi_{ll'}^{mm'} V_{l'}^{m'}, \quad (12)$$

where  $\Xi_{ll'}^{mm'}$  from equation (10) has been substituted. The condition expressed in equation (10) has been exploited to reduce the summation limits for  $l$  in equation (12) from  $(1 \rightarrow \infty)$  to  $(1 \rightarrow l')$ . Hence, the algorithmic efficiency of the representation of (12) compared to the original one (combining Eqs. (1) and (2)) now depends on a different criterium: it does no more depend on the convergence of the susceptibility spectrum, i.e. the decay of its eigenvalue series  $(\lambda_i)$ , but instead on the quantum number  $l'_{\text{max}}$  necessary to expand the perturbing potential in terms of solid harmonics functions.

Physically, the coefficients  $V_l^m$  represent the  $(l, m)$ -components of an external potential expanded at the position of the perturbed molecule/fragment, which is generally several Ångströms away from the position of the perturbing object. Hence, the spatial variation of  $V_{\text{pert}}(\mathbf{r})$  at this distance is likely to be weaker, and the angular moment expansion might be truncated at moderate values of  $(l, m)$ :

$$n_{\text{resp}}(\mathbf{r}) = \sum_{l'=1}^{l'_{\text{max}}} \sum_{l=1}^{l'} \sum_{m'=-l'}^{l'} \sum_{m=-l}^l \xi_l^m(\mathbf{r}) \Xi_{ll'}^{mm'} V_{l'}^{m'}. \quad (13)$$

We have demonstrated this for the hydrogen bond interaction in a water dimer[16]: it turns out that  $l'_{\text{max}} = 4$  delivers an accurate response density of one water molecule to the perturbing potential of the other one.

The spatial shape of these moment expanded states  $\xi_l^m(\mathbf{r})$  for water are discussed in detail in [16] as well. The number of moment expanded states for a maximal angular momentum channel  $l'_{\text{max}}$  grows with  $\mathcal{O}(l_{\text{max}}^2)$ . Thus, a converged response density can be obtained by a number of moment expanded states which is very small compared to the several thousand states of the spectral decomposition.

The algorithm for *obtaining* them according to (8), however, still requires the initial determination of virtually the full spectrum of  $\hat{\chi}$ , i.e. several thousands of couples  $(\lambda_i, \chi_i(\mathbf{r}))$ , plus the unitary transformation matrix  $U_{ij}$ . In the following section, we will derive an algorithm which allows to compute the moment expanded states directly.

### 3 Computation of the $\xi_l^m(\mathbf{r})$

In this section, we will present an algorithm which allows to compute the moment expanded states  $\xi_l^m(\mathbf{r})$  iteratively, without the need to obtain the spectral decomposition beforehand. In the first step, we determine the response density to a perturbing potential which is equal to the first solid harmonic, i.e.

$$V_{\text{pert}}(\mathbf{r}') = R_1^{-1}(\mathbf{r}'). \quad (14)$$

This is equivalent to setting  $V_1^{-1} = 1$  in the Laplace expansion and all other coefficients  $V_{l \neq 1}^{m \neq -1} = 0$ .

The response density to such a potential is given by equation (13):

$$n_{\text{resp}}(\mathbf{r}) = \sum_{m=-l}^l \xi_l^m(\mathbf{r}) \Xi_{11}^{m-1} \quad (15)$$

$$= n_1^{-1}(\mathbf{r}) = \xi_1^{-1}(\mathbf{r}) \Xi_{11}^{-1-1}. \quad (16)$$

Projection on the solid harmonic  $R_1^{-1}(\mathbf{r}')$  yields

$$\langle R_1^{-1} | n_1^{-1} \rangle = |\Xi_{11}^{-1-1}|^2. \quad (17)$$

Eventually, the first moment expanded state  $\xi_1^{-1}(\mathbf{r})$  can be obtained by

$$\xi_1^{-1}(\mathbf{r}) = \frac{n_1^{-1}(\mathbf{r})}{\sqrt{\langle R_1^{-1} | n_1^{-1} \rangle}}. \quad (18)$$

The response density to the perturbing potential (14) can be calculated directly by means of an explicit DFPT run, yielding  $\xi_1^{-1}(\mathbf{r})$  by equation (18). In order to obtain the state of next higher symmetry  $\xi_1^0(\mathbf{r})$ , the response density to a perturbing potential equal to the second solid harmonic has to be determined:

$$V_{\text{pert}}(\mathbf{r}') = \xi_1^0(\mathbf{r}'). \quad (19)$$

The response density is then given by equation (13) as

$$n_{\text{resp}}(\mathbf{r}) = \xi_1^{-1}(\mathbf{r}) \Xi_{11}^{-10} + \xi_1^0(\mathbf{r}) \Xi_{11}^{00}, \quad (20)$$

so that follows

$$\begin{aligned} \xi_1^0(\mathbf{r}) \Xi_{11}^{00} &= n_{\text{resp}}(\mathbf{r}) - \xi_1^{-1}(\mathbf{r}) \Xi_{11}^{-10} \\ &=: n_1^0(\mathbf{r}). \end{aligned} \quad (21)$$

$\Xi_{11}^{00}$  can be obtained analogous to equation (17), eventually leading to

$$\xi_1^0(\mathbf{r}) = \frac{n_1^0(\mathbf{r})}{\sqrt{\langle R_1^0 | n_1^0 \rangle}}. \quad (22)$$

The total response density to a perturbing potential

$$V_{\text{pert}}(\mathbf{r}') = R_1^1(\mathbf{r}') \quad (23)$$

is given as:

$$n_{\text{resp}}(\mathbf{r}) = \xi_1^{-1}(\mathbf{r}) \Xi_{11}^{-11} + \xi_1^0(\mathbf{r}) \Xi_{11}^{01} + \xi_1^1(\mathbf{r}) \Xi_{11}^{11}. \quad (24)$$

It follows

$$\xi_1^1(\mathbf{r}) \Xi_{11}^{11} = n_{\text{resp}}(\mathbf{r}) - \xi_1^{-1}(\mathbf{r}) \Xi_{11}^{-11} - \xi_1^0(\mathbf{r}) \Xi_{11}^{01} =: n_1^1(\mathbf{r}) \quad (25)$$

and

$$\xi_1^1(\mathbf{r}) = \frac{n_1^1(\mathbf{r})}{\sqrt{\langle R_1^1 | n_1^1 \rangle}}. \quad (26)$$

The above algorithm can be generalized for the calculation of any moment expanded state  $\xi_l^m(\mathbf{r})$  with **specific**  $(l, m)$ , which shall be shown in the following. The total response density to a perturbation potential equal to the solid harmonic  $R_l^m(\mathbf{r}')$  is

$$n_{\text{resp}}(\mathbf{r}) = \int \chi(\mathbf{r}, \mathbf{r}') R_l^m(\mathbf{r}') d^3 r'. \quad (27)$$

It follows from equations (10) and (13) that

$$n_{\text{resp}}(\mathbf{r}) = \sum_{l'=1}^{l-1} \sum_{m'=-l'}^{l'} \xi_{l'}^{m'}(\mathbf{r}) \Xi_{l'l}^{m'm} + \sum_{m'=-l}^m \xi_l^{m'}(\mathbf{r}) \Xi_{ll}^{m'm}. \quad (28)$$

Furthermore, it follows

$$\begin{aligned} \xi_l^m(\mathbf{r}) \Xi_{ll}^{mm} &= n_{\text{resp}}(\mathbf{r}) - \sum_{l'=1}^{l-1} \sum_{m'=-l'}^{l'} \xi_{l'}^{m'}(\mathbf{r}) \Xi_{l'l}^{m'm} \\ &- \sum_{m'=-l}^{m-1} \xi_l^{m'}(\mathbf{r}) \Xi_{ll}^{m'm} =: n_l^m(\mathbf{r}). \end{aligned} \quad (29)$$

If all the moment expanded states of lower symmetry  $(l', m')$  are known,  $\xi_l^m(\mathbf{r})$  can be easily calculated:

$$\xi_l^m(\mathbf{r}) = \frac{n_l^m(\mathbf{r})}{\sqrt{\langle R_l^m | n_l^m \rangle}}. \quad (30)$$

Since the explicit calculation of a single state using DFPT requires only one additional perturbation calculation, only  $l^2 + l + m$  perturbation calculations are necessary to compute a moment expanded state of order  $(l, m)$ . This drastically reduces the computational effort and the required storage in comparison to the computation of several thousand eigenstates for the spectral decomposition

according to equation (2). In the case of a water dimer, the corresponding numbers are about 5000 DFPT iterations (one for each Lanczos step) and 3000 retained Eigenfunctions  $\chi_i(\mathbf{r})$  for the original scheme using equation (2), compared to 25 DFPT calculations for all new states up to  $l_{\max} = 4$  and the same number of 25 retained functions  $\xi_l^m(\mathbf{r})$ .

## 4 Conclusion

We have derived an iterative method for the efficient and direct computation of the linear density-density response function in its moment expanded representation. Our novel method does not rely on the convergence of the Lanczos-based diagonalisation of the response tensor  $\chi(\mathbf{r}, \mathbf{r}')$  and the determination of the subsequent transformation of the spectral decomposition into the moment expanded states. Instead, the moment expanded states  $\xi_l^m(\mathbf{r})$  are obtained directly from the elementary iterations of the response function.

The moment expanded representation of the density-density response function allows new insight into the nature of the reaction of a complex set of molecular orbitals to external perturbations such as intermolecular interactions.

The considerable increase in algorithmic efficiency relies on the possibility to expand the perturbing potential in a comparably small number of solid harmonics functions, i.e. to truncate the sum in equation (4) at low values of  $l$ . While in principle, this truncation does *not* depend on the size of the molecule (rather on the spatial variation of the perturbing potential), it is clear that for a large molecule, this variation will most likely be larger. Hence, the algorithmic advantage of our new scheme may somewhat shrink. However, for a large molecule, the decay of the regular Eigenvalue spectrum of the susceptibility tensor will be much slower than for a small molecule. Hence, the new scheme will again be more advantageous compared to the direct Eigensystem representation according to equation (2).

It should be noted that neither the increase in efficiency nor the accuracy do *not* rely on the properties (size, molecular symmetry, ...) of the molecule under consideration, but instead on the variability of the external potential in the region of that molecule. In turn, this means that for an external potential that has a very anisotropic shape in the region of the considered molecule, the expansion of the potential in terms of solid harmonics may be difficult to converge and the response calculation via equation (13) may become inaccurate for small values of  $l_{\max}$ . We believe that many common intermolecular interaction types such as hydrogen bonds and polarization effects due to polar and ionic groups can be represented with a very modest quantum number of about  $l_{\max} = 4$ .

## Author contribution statement

A. Scherrer developed the idea of using symmetry-adapted starting functions for the iterative response calculation scheme and implemented the code; P. Ahlert

and C. Dressler were responsible for benchmarking and application calculations; D. Sebastiani initiated and supervised the project. All authors were involved in the preparation of the manuscript and have approved its final version.

## References

1. P. Hohenberg, W. Kohn, Phys. Rev. **136**, B864 (1964)
2. W. Kohn, L.J. Sham, Phys. Rev. **140**, A1133 (1965)
3. R.O. Jones, O. Gunnarsson, Rev. Mod. Phys. **61**, 689 (1989)
4. X. Gonze, J.P. Vigneron, Phys. Rev. B. **39**, 13120 (1989)
5. P. Giannozzi, S. de Gironcoli, P. Pavone, S. Baroni, Phys. Rev. B **43**, 7231 (1991)
6. X. Gonze, D.C. Allan, M.P. Teter, Phys. Rev. Lett. **68**, 3603 (1992)
7. X. Gonze, Phys. Rev. A **52**, 1096 (1995)
8. S. Baroni, S. de Gironcoli, A. dal Corso, P. Giannozzi, Rev. Mod. Phys. **73**, 515 (2001)
9. A. Putrino, D. Sebastiani, M. Parrinello, J. Chem. Phys. **113**, 7102 (2000)
10. A. Scherrer, V. Verschinin, D. Sebastiani, J. Chem. Theory Comput. **8**, 106 (2012)
11. T. Pham, T. Li, S. Shankar, F. Gygi, G. Galli, Phys. Rev. B **84**, 45308 (2011)
12. D. Lu, F. Gygi, G. Galli, Phys. Rev. Lett. **100**, 147601 (2008)
13. H.F. Wilson, D. Lu, F. Gygi, G. Galli, Phys. Rev. B **79**, 245106 (2009)
14. A.C. Ihrig, A. Scherrer, D. Sebastiani, J. Chem. Phys. **139**, 094102 (2013)
15. A. Scherrer, C. Dreßler, P. Ahlert, D. Sebastiani, J. Chem. Phys. **144**, 144111 (2016)
16. A. Scherrer, D. Sebastiani, J. Comput. Chem. **37**, 665 (2016)
17. D. Loco, L. Lagardère, S. Caprasecca, F. Lipparini, B. Mennucci, J.P. Piquemal, J. Chem. Theory Comput. **13**, 4025 (2017)
18. G. Naray-Szabo, G.G. Ferenczy, Chem. Rev. **95**, 829 (1995)
19. W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey, M.L. Klein, J. Chem. Phys. **79**, 926 (1983)
20. V.V. Gobre, A. Tkatchenko, Nat. Commun. **4**, 2341 (2013)
21. F. Rob, K. Szalewicz, Mol. Phys. **111**, 1430 (2013)
22. T. Watermann, A. Scherrer, D. Sebastiani, in *Many-electron approaches in physics, chemistry and mathematics*, edited by V. Bach, L. Delle Site (Springer International Publishing, 2014), Mathematical Physics Studies, pp. 97–110
23. A. Heßelmann, G. Jansen, Chem. Phys. Lett. **367**, 778 (2003)
24. A.J. Misquitta, B. Jeziorski, K. Szalewicz, Phys. Rev. Lett. **91**, 033201 (2003)
25. A. Heßelmann, G. Jansen, M. Schütz, J. Chem. Phys. **122**, 014103 (2005)
26. Y.R. Shen, V. Ostroverkhov, Chem. Rev. **106**, 1140 (2006)
27. J.L. Skinner, P.A. Pieniazek, S.M. Gruenbaum, Acc. Chem. Res. **45**, 93 (2012)



28. T. Ishiyama, T. Imamura, A. Morita, Chem. Rev. **114**, 8447 (2014)
29. E. Anim-Danso, Y. Zhang, A. Alizadeh, A. Dhinojwala, J. Am. Chem. Soc. **135**, 2734 (2013)
30. M. Sulpizi, M. Salanne, M. Sprik, M.P. Gaigeot, J. Phys. Chem. Lett. **4**, 83 (2013)
31. K. Matsuzaki, S. Nihonyanagi, S. Yamaguchi, T. Nagata, T. Tahara, J. Phys. Chem. Lett. **4**, 1654 (2013)
32. M. Sovago, R.K. Campen, G.W.H. Wurpel, M. Müller, H.J. Bakker, M. Bonn, Phys. Rev. Lett. **100**, 173901 (2008)
33. A. Solovyeva, O.A. von Lilienfeld, Phys. Chem. Chem. Phys. **18**, 31078 (2016)
34. K.Y.S. Chang, O.A. von Lilienfeld, Chimia **68**, 602 (2014)
35. Z. Gimbutas, L. Greengard, J. Comput. Phys. **228**, 5621 (2009)

### 3.2.4 Article IX: Generalization of the Electronic Susceptibility for Arbitrary Molecular Geometries.

Arne Scherrer, Christian Dreßler, Paul Ahlert, and Daniel Sebastiani.

Generalization of the electronic susceptibility for arbitrary molecular geometries.

*The Journal of Chemical Physics*, 144(14):144111, **2016**.

In this article, I performed the numerical calculations of the geometry dependence of the linear density-density response function. A. Scherrer derived the theory. A. Scherrer, P. Ahlert and I implemented the theory. A. Scherrer and I wrote the manuscript. D. Sebastiani supervised the project and provided conceptual advice.



# Generalization of the electronic susceptibility for arbitrary molecular geometries

 Arne Scherrer, Christian Dreßler, Paul Ahlert, and Daniel Sebastiani<sup>a)</sup>
*Institute of Chemistry, Martin-Luther-University Halle-Wittenberg, Von-Danckelmann-Platz 4, 06120 Halle (Saale), Germany*

(Received 12 January 2016; accepted 7 March 2016; published online 14 April 2016)

We generalize the explicit representation of the electronic susceptibility  $\chi_{[\mathbf{R}]}(\mathbf{r}, \mathbf{r}')$  for arbitrary molecular geometries  $\mathbf{R}$ . The electronic susceptibility is a response function that yields the response of the molecular electronic charge density at linear order to an arbitrary external perturbation. We address the dependence of this response function on the molecular geometry. The explicit representation of the molecular geometry dependence is achieved by means of a Taylor expansion in the nuclear coordinates. Our approach relies on a recently developed low-rank representation of the response function  $\chi_{[\mathbf{R}]}(\mathbf{r}, \mathbf{r}')$  which allows a highly condensed storage of the expansion and an efficient application within dynamical chemical environments. We illustrate the performance and accuracy of our scheme by computing the vibrationally induced variations of the response function of a water molecule and its resulting Raman spectrum. © 2016 AIP Publishing LLC. [<http://dx.doi.org/10.1063/1.4945372>]

## I. INTRODUCTION

The accurate evaluation of intermolecular electrostatic interactions is of great importance for large scale molecular dynamics simulations. In many cases a fragmentation approach can be used to increase the efficiency of such calculations. Common methods employed are based on multipole expansions,<sup>1–6</sup> density fitting,<sup>7–9</sup> or perturbation theories.<sup>10–13</sup> Provided that polarization effects are taken into account, this involves the calculation of *response* properties of the fragments, e.g., in the simplest case the dipole-dipole-polarizability  $\alpha$ .

Wave function and density functional theory based perturbation theories are very successful in providing chemically accurate intermolecular interaction energies yielding instructive energy decomposition schemes.<sup>10–13</sup> However, their power comes at the price of a high computational complexity. In view of their application to molecular dynamics simulations, the dependence of the involved quantities on the atomic configuration  $\mathbf{R} = \{\mathbf{R}_i\}$  has to be considered explicitly.

The orbital based calculation of the linear electronic response density via density functional perturbation theory<sup>14–18</sup> (DFPT) requires a self-consistent solution of the Sternheimer equation

$$\begin{aligned}
 & -\left[\hat{H}_{KS}^{(0)}(\mathbf{R}) - \epsilon_o^{(0)}(\mathbf{R})\right] \hat{P}_{[\mathbf{R}]e} \left| \phi_{[\mathbf{R}]o}^{(1)} \right\rangle \\
 & = \hat{P}_{[\mathbf{R}]e} \hat{H}_{KS}^{(1)}[\{\phi_o\}] \left| \phi_{[\mathbf{R}]o}^{(0)} \right\rangle, \quad \text{for } o = 1 \dots N, \quad (1)
 \end{aligned}$$

$$n_{[\mathbf{R}]}^{\text{resp}}(\mathbf{r}) = 2 \sum_{o=1}^{N_e} \text{Re} \left[ \phi_{[\mathbf{R}]o}^{(0)*}(\mathbf{r}) \phi_{[\mathbf{R}]o}^{(1)}(\mathbf{r}) \right], \quad (2)$$

where  $\hat{P}_{[\mathbf{R}]e} = 1 - \sum_{o=1}^{N_e} |\phi_{[\mathbf{R}]o}\rangle \langle \phi_{[\mathbf{R}]o}|$  is a projector on the manifold of unoccupied states.

We have explicitly included the parametric dependence of the involved operators and orbitals on the molecular geometry

$\mathbf{R}$ . The perturbation Hamiltonian on the right hand side  $\hat{H}_{KS}^{(1)}[\{\phi_o\}]$  depends on the electronic density response and hence implicitly on the perturbed orbitals on the left hand side, which necessitates a self-consistent solution of the equations.

During a molecular dynamics simulation, an orbital-based evaluation of the electronic response corresponds to solving the unperturbed ground state electronic structure  $H_{KS}^{(0)}(\mathbf{R})$ ,  $\epsilon_o^{(0)}(\mathbf{R})$ ,  $\phi_{[\mathbf{R}]o}^{(0)}(\mathbf{r})$  for each fragment and configuration  $\mathbf{R}$  and subsequently a series of self-consistent solutions of the perturbation equations  $\phi_{[\mathbf{R}]o}^{(1)}(\mathbf{r})$ ,  $n_{[\mathbf{R}]}^{\text{resp}}(\mathbf{r})$  for each pairwise interaction  $V_{\text{pert}}(\mathbf{r})$  of the fragments. This is evidently a very costly undertaking, limiting its applicability to comparatively small system sizes.

Density based methods provide a considerable reduction of dimensionality. The long-range regime of the electrostatic interaction is elegantly described by the (distributed) point multipole approximation, which drastically simplifies the interaction.<sup>1–4,19,20</sup> For larger molecules, distributed point multipoles have proven to work even if the single point multipole approximation diverges.<sup>19,21</sup> A further generalization allowing for polarization effects is the distributed polarizability method which attributes multipole polarizabilities to different sites of the molecules.<sup>22–25</sup> By construction, point multipoles give a poor description of the short range regime of the interaction.<sup>26</sup> Strategies to overcome this problem employ a spatial representation of the electronic density via Gaussian charge distributions<sup>27,28</sup> or Gaussian multipoles.<sup>6,29–31</sup>

Our work aims to push this approach to the next level, combining the advantages of a reduced computational complexity for the evaluation of the response density while keeping the response density in its full non-local spatial dependence and not only its multipole moments. In analogy to Gaussian multipole moments<sup>6,31</sup> (as a generalization of point multipoles), our representation of the electronic susceptibility  $\chi(\mathbf{r}, \mathbf{r}')$  provides a generalization of the multipole-multipole-polarizabilities.<sup>22–25,32,33</sup>

<sup>a)</sup>Electronic mail: daniel.sebastiani@chemie.uni-halle.de

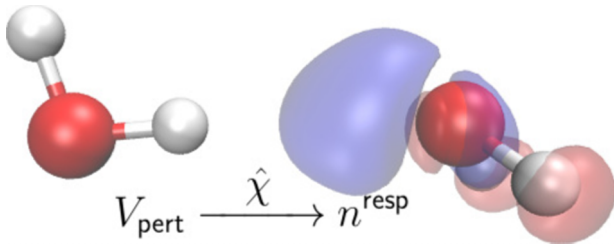


FIG. 1. Density response  $n_{[\mathbf{R}]}^{\text{resp}}(\mathbf{r})$  of a water molecule (right) due to the perturbation potential  $V_{\text{pert}}(\mathbf{r})$  generated by the electrostatic potential of a hydrogen bond donor (left water molecule).

Here, we make a step towards a higher usability of the approach without the need of solving Eqs. (1) and (2) self-consistently for each

- (a) novel external perturbation potential and
- (b) new molecular conformation  $\mathbf{R}$ .

The first aspect (a) is achieved by our recently developed explicit representation of the electronic susceptibility  $\chi_{[\mathbf{R}]}(\mathbf{r}, \mathbf{r}')$  within a low-rank approximation.<sup>32–37</sup> The electronic linear response density  $n_{[\mathbf{R}]}^{\text{resp}}(\mathbf{r})$  due to a perturbation potential  $V_{\text{pert}}(\mathbf{r}')$  then is determined by the linear electronic susceptibility  $\chi_{[\mathbf{R}]}(\mathbf{r}, \mathbf{r}')$  via

$$n_{[\mathbf{R}]}^{\text{resp}}(\mathbf{r}) = \int \chi_{[\mathbf{R}]}(\mathbf{r}, \mathbf{r}') V_{\text{pert}}(\mathbf{r}') d^3 r'. \quad (3)$$

The specific example of the density response  $n_{[\mathbf{R}]}^{\text{resp}}(\mathbf{r})$  within a water molecule due to a hydrogen bond is illustrated in Fig. 1. Here, the specific perturbation potential  $V_{\text{pert}}(\mathbf{r}) = \int \frac{n_{\text{left}}^{\text{H}_2\text{O}}(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d^3 r'$  was used to compute  $n_{[\mathbf{R}]}^{\text{resp}}(\mathbf{r})$  via Eq. (3). In principle, however, Eq. (3) defines a response function which is formally *independent* of the perturbation potential. The simultaneous incorporation of the implicit geometry dependence (b) requires in principle an explicit representation of the susceptibility  $\chi_{[\mathbf{R}]}(\mathbf{r}, \mathbf{r}')$  on the nuclear coordinates  $\mathbf{R}$  which is nontrivial. In this article we show that this explicit representation of the full non-local response function on the molecular geometry can be achieved by means of a Taylor expansion in the nuclear coordinates. Our approach relies on the very condensed representations<sup>32,33</sup> of the full response function which provides an efficient way to address this problem.

## II. MOLECULAR GEOMETRY EXPANSION

Formally, there is no obstacle to explicitly calculate the response function  $\chi_{[\mathbf{R}]}(\mathbf{r}, \mathbf{r}')$  for different configurations  $\mathbf{R}$ . This allows to analyze the dependencies of the electronic susceptibility on the molecular geometry and it will be used for comparison in the following. However, its brute force tabulation for all relevant configurations is no practicable way to approach problem (b).

An obvious approach to explicitly express the geometry dependence of the response function is a multi-dimensional Taylor expansion around the geometry of interest, e.g., the

equilibrium geometry  $\mathbf{R}_0$

$$\chi_{[\mathbf{R}]}(\mathbf{r}, \mathbf{r}') = \sum_{|a| \geq 0} \frac{(\mathbf{R} - \mathbf{R}_0)^a}{a!} \partial^a \chi_{[\mathbf{R}_0]}(\mathbf{r}, \mathbf{r}'), \quad (4)$$

where we have used multi-index notation. This involves the calculation of derivatives of the high-dimensional response function, i.e., for the displacement of nucleus  $\nu$  in Cartesian direction  $k$  we obtain

$$\chi_{[\mathbf{R}_0]}^{(\nu k)}(\mathbf{r}, \mathbf{r}') = \left. \frac{\partial \chi_{[\mathbf{R}]}(\mathbf{r}, \mathbf{r}')}{\partial R_k^\nu} \right|_{\mathbf{R}=\mathbf{R}_0}, \quad (5)$$

where the superscripts in parenthesis ( $\nu k$ ) denote the derivative. Its direct calculation, i.e., an explicit diagonalization of the analytical derivative  $\chi_{[\mathbf{R}_0]}^{(\nu k)}(\mathbf{r}, \mathbf{r}')$  might be possible. Prior work on the non-local polarizability density has expressed its analytical derivative in terms of the non-local hyperpolarizability density.<sup>38,39</sup> This nicely illustrates the physics underlying the derivative of the response function but barely provides a numerically practical scheme for its computation.

An evident way to calculate the derivatives in Eq. (4) is the numerical derivative of the response function via finite differences, e.g., via forward difference for the first order. For a geometry  $\mathbf{R}_k^\nu$  with nucleus  $\nu$  displaced in direction  $k$ , i.e.,  $R_k^\nu = R_{k,0}^\nu + \Delta \delta_{kl} \delta_{\nu\nu'}$ , this gives

$$\chi_{[\mathbf{R}_0]}^{(\nu k)}(\mathbf{r}, \mathbf{r}') = \frac{\chi_{[\mathbf{R}_k^\nu]}(\mathbf{r}, \mathbf{r}') - \chi_{[\mathbf{R}_0]}(\mathbf{r}, \mathbf{r}')}{\Delta}. \quad (6)$$

However, this operation is computationally very inconvenient due to the extremely high dimensionality of the involved objects. Therefore, the direct difference of the response function will not be used in this work. In order to obtain a feasible expression for the derivative of the total response function, its spectral decomposition is used<sup>32,34</sup>

$$\chi_{[\mathbf{R}]}(\mathbf{r}, \mathbf{r}') = \sum_{i=1}^{\infty} \chi_{[\mathbf{R}]i}(\mathbf{r}) \chi_{[\mathbf{R}]i}(\mathbf{r}'). \quad (7)$$

The full response function is represented as a weighted outer product over eigenvalues  $\chi_{[\mathbf{R}]i}$  and eigenfunctions  $\chi_{[\mathbf{R}]i}(\mathbf{r})$ .

The direct derivation of Eq. (7) gives via the product rule

$$\begin{aligned} \chi_{[\mathbf{R}_0]}^{(\nu k)}(\mathbf{r}, \mathbf{r}') &= \sum_i \left( \chi_{[\mathbf{R}_0]i}^{(\nu k)}(\mathbf{r}) \chi_{[\mathbf{R}_0]i}(\mathbf{r}') \right. \\ &\quad + \chi_{[\mathbf{R}_0]i}(\mathbf{r}) \chi_{[\mathbf{R}_0]i}^{(\nu k)}(\mathbf{r}') \\ &\quad \left. + \chi_{[\mathbf{R}_0]i}(\mathbf{r}) \chi_{[\mathbf{R}_0]i}(\mathbf{r}') \chi_{[\mathbf{R}_0]i}^{(\nu k)}(\mathbf{r}') \right), \end{aligned} \quad (8)$$

where derivatives are taken at  $\mathbf{R} = \mathbf{R}_0$ .

This suggests that the numerical derivative can be performed on the eigenstates  $\chi_{[\mathbf{R}_0]i}(\mathbf{r})$  and eigenvalues  $\chi_{[\mathbf{R}_0]i}$  of the spectral decomposition

$$\chi_{[\mathbf{R}_0]i}^{(\nu k)}(\mathbf{r}) = \frac{\chi_{[\mathbf{R}_k^\nu]i}(\mathbf{r}) - \chi_{[\mathbf{R}_0]i}(\mathbf{r})}{\Delta}, \quad (9)$$

$$\chi_{[\mathbf{R}_0]i}^{(\nu k)} = \frac{\chi_{[\mathbf{R}_k^\nu]i} - \chi_{[\mathbf{R}_0]i}}{\Delta}. \quad (10)$$

At first glance, this appears to solve the problem. In particular, the representation and the subtraction of the states are technically feasible.

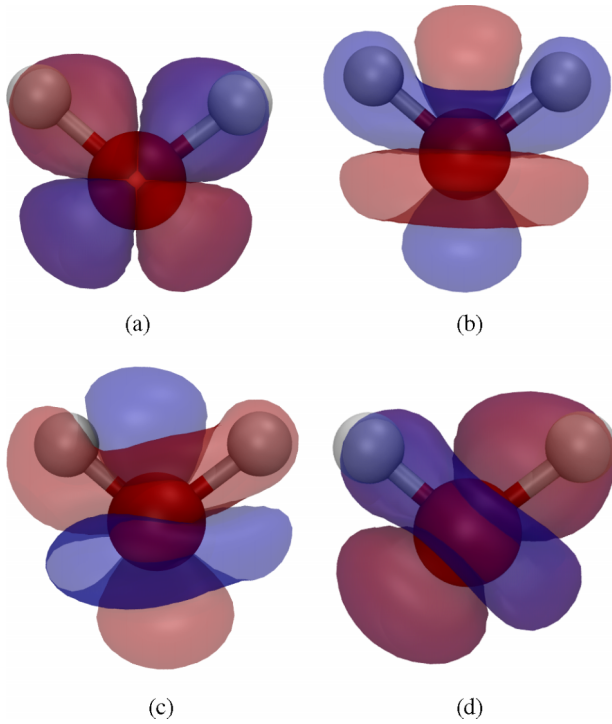


FIG. 2. Eigenstates 13 and 14 for H<sub>2</sub>O in equilibrium geometry ((a) and (b)) and in slightly displaced geometry ((c) and (d)). The sign and order of the states have changed. This illustrates that small changes of the system can lead to completely different eigenstates. (a)  $\chi_{[\mathbf{R}_0]13}(\mathbf{r})$ , (b)  $\chi_{[\mathbf{R}_0]14}(\mathbf{r})$ , (c)  $\chi_{[\mathbf{R}_k^y]13}(\mathbf{r})$ , (d)  $\chi_{[\mathbf{R}_k^y]14}(\mathbf{r})$ .

However, a closer look at the eigenstates for different geometries reveals an additional complication. In Fig. 2 we show the eigenstates  $\chi_{[\mathbf{R}]13}(\mathbf{r})$  and  $\chi_{[\mathbf{R}]14}(\mathbf{r})$  of an isolated water molecule for different geometries.

The change of symmetry of the eigenstates stems from the intermixing of near degenerate eigenstates. If the system changes gradually, e.g., due to a different configuration, some eigenstates shift their eigenvalues. This reveals itself via a change of the ordering eigenvalues and eigenstates. If two states have near degenerate eigenvalues, this change can lead to an intermixing or swap of the eigenstates. The eigenstates with the same index  $i$  for different configurations may, even if the displacement is very small, have completely different symmetry. This applies in particular if the displacement breaks a molecular symmetry. Furthermore, only the absolute values of the eigenstates are well defined, their signs have no physical relevance since they cancel themselves in the symmetric outer product in Eq. (7). This means that eigenstates are in general not continuously differentiable and Eq. (8) provides no feasible way for the calculation of the derivative. Hence, an additional theoretical step is necessary.

### III. DIFFERENTIABILITY THANKS TO SYMMETRY DECOMPOSITION

We look for a decomposition of the response function in states that are generally continuously differentiable and hence suited for finite differences. That is, the states should be continuously differentiable in the nuclear coordinates.

For electronic systems far from conical intersections, the overall response function should fulfill this property. For the eigenstates, this condition is not fulfilled. Therefore, a change of representation is needed that yields transformed states with the desired properties.

We have very recently derived a new representation of the full electronic susceptibility via a moment expansion.<sup>33</sup> In this very condensed representation, the total response function is given as a sum of the outer products of transformed states, partitioned in orders of the angular momentum expansion

$$\chi_{[\mathbf{R}]}(\mathbf{r}, \mathbf{r}') = \sum_{l=1}^{\infty} \chi_{[\mathbf{R}]}^l(\mathbf{r}, \mathbf{r}'), \quad (11)$$

$$\chi_{[\mathbf{R}]}^l(\mathbf{r}, \mathbf{r}') = \sum_{m=-l}^l \xi_{[\mathbf{R}]}^m(\mathbf{r}) \xi_{[\mathbf{R}]}^m(\mathbf{r}'). \quad (12)$$

The transformed states  $\xi_{[\mathbf{R}]}^m(\mathbf{r})$  are labeled by their corresponding multipole moment and fulfill the property

$$\langle \xi_{[\mathbf{R}]}^m | R_{l'}^{m'} \rangle = \begin{cases} \Xi_{[\mathbf{R}]}^{mm'} & \text{if } l \leq l' \\ 0 & \text{if } l > l' \end{cases}, \quad (13)$$

where  $R_{l'}^{m'}(\mathbf{r})$  are real Racah normalized regular solid harmonic functions and  $\Xi_{[\mathbf{R}]}^{mm'}$  are multipole moments of the states and in general non-zero. In other words, the new representation  $\{\xi_{[\mathbf{R}]}^m(\mathbf{r})\}$  yields states for which the subset  $\{\xi_{[\mathbf{R}]}^m(\mathbf{r}) | l \leq l_{\max} \wedge m \in \{-l, \dots, l\}\}$  comprises the *complete*  $\chi$  response up to an angular momentum channel of  $l_{\max}$ .

The derivation of the change of representation as well as its properties is not in the scope of this article and is discussed extensively elsewhere.<sup>33</sup> Important for this work are the new properties of the moment expanded states  $\xi_{[\mathbf{R}]}^m(\mathbf{r})$ . For each angular momentum channel  $l$  of the moment expansion  $\chi_{[\mathbf{R}]}^l(\mathbf{r}, \mathbf{r}')$ , the new decomposition in Eq. (11) is uniquely defined, i.e., the derivative of the whole response function can be split into a sum of derivatives of angular momentum contributions

$$\chi_{[\mathbf{R}_0]}^{(vk)}(\mathbf{r}, \mathbf{r}') = \sum_{l=1}^{\infty} \chi_{[\mathbf{R}_0]}^{l(vk)}(\mathbf{r}, \mathbf{r}'). \quad (14)$$

For a given choice of the coordinate system, also the decomposition of each angular momentum contribution  $\chi_{[\mathbf{R}]}^l(\mathbf{r}, \mathbf{r}')$  in moment expanded states  $\xi_{[\mathbf{R}]}^m(\mathbf{r})$  according to Eq. (12) is well defined. Therefore, the physically observable contribution of one transformed state to the total response is uniquely determined by its symmetry

$$n_{[\mathbf{R}]}^{\text{resp}}(\mathbf{r}) = \sum_{l=1}^{\infty} n_{[\mathbf{R}]}^l(\mathbf{r}), \quad (15)$$

$$n_{[\mathbf{R}]}^l(\mathbf{r}) = \int \chi_{[\mathbf{R}]}^l(\mathbf{r}, \mathbf{r}') V_{\text{pert}}(\mathbf{r}') d^3 r' \quad (16)$$

$$= \sum_{m=-l}^l \xi_{[\mathbf{R}]}^m(\mathbf{r}) \langle \xi_{[\mathbf{R}]}^m | V_{\text{pert}} \rangle. \quad (17)$$

This property is the key difference to the eigenstates: the partition of the total response into physically observable contributions of distinct symmetry guarantees that the underlying states are also continuously differentiable. It is this property that prohibits a discontinuous intermixing of

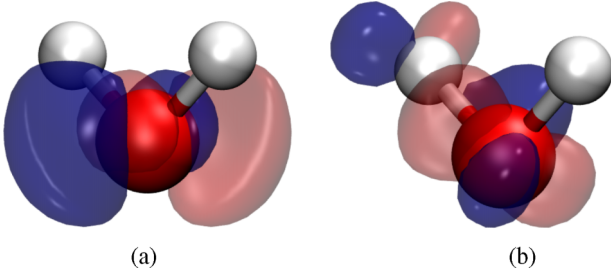


FIG. 3. (a) First moment-expanded state  $\xi_{[\mathbf{R}_0]l}^{-1}(\mathbf{r})$  for H<sub>2</sub>O in equilibrium geometry. (b) Derivative state  $\xi_{[\mathbf{R}_0]l}^{-1(vk)}(\mathbf{r})$  for OH-bond elongation.

different states and hence guarantees that the finite-difference method gives well-defined derivatives

$$\chi_{[\mathbf{R}_0]}^{l(vk)}(\mathbf{r}, \mathbf{r}') = \sum_{m=-l}^l \left( \xi_{[\mathbf{R}_0]l}^{m(vk)}(\mathbf{r}) \xi_{[\mathbf{R}_0]l}^m(\mathbf{r}') + \xi_{[\mathbf{R}_0]l}^m(\mathbf{r}) \xi_{[\mathbf{R}_0]l}^{m(vk)}(\mathbf{r}') \right). \quad (18)$$

In contrast to Eq. (8) where only the total sum is continuously differentiable, in Eq. (18) all single terms in the sum are continuously differentiable on their own.

Therefore, the numerical finite difference calculation of the derivatives of moment expanded states is possible

$$\xi_{[\mathbf{R}_0]l}^{m(vk)}(\mathbf{r}) = \frac{\xi_{[\mathbf{R}_k]l}^m(\mathbf{r}) - \xi_{[\mathbf{R}_0]l}^m(\mathbf{r})}{\Delta}. \quad (19)$$

In Fig. 3 we show a moment expanded state  $\xi_{[\mathbf{R}_0]l}^m(\mathbf{r})$  and its corresponding bond-elongation derivative state  $\xi_{[\mathbf{R}_0]l}^{m(vk)}(\mathbf{r})$ .

#### IV. NUMERICAL VERIFICATION

The theoretical implication of continuous differentiability of the moment expanded states can be verified numerically. We have calculated the spectral decomposition and the corresponding moment expanded representation for different nuclear configurations  $\mathbf{R}$ . As a benchmark system we use an isolated water molecule and vary the nuclear configuration. For the time being we truncate the Taylor expansion in Eq. (4) after the second order and restrict ourselves to the first two angular momentum channels of the moment expansion  $l = 1, 2$ ,

$$\begin{aligned} \chi_{[q]}(\mathbf{r}, \mathbf{r}') \approx & \sum_{l=1}^2 \left( \chi_{[q_0]l}^l(\mathbf{r}, \mathbf{r}') + \sum_{n=1}^{N_{int}} \chi_{[q_0]l}^{l(n)}(\mathbf{r}, \mathbf{r}') \Delta q_n \right. \\ & \left. + \frac{1}{2} \sum_{n,n'=1}^{N_{int}} \chi_{[q_0]l}^{l(nn')}(\mathbf{r}, \mathbf{r}') \Delta q_n \Delta q_{n'} \right), \quad (20) \end{aligned}$$

with  $\Delta q_n = q_n - q_{n0}$ . The superscript ( $n$ ) denotes the derivatives with respect to the internal coordinate  $q_n$  and  $N_{int}$  is the number of internal coordinates,  $N_{int} = 3N-6$  for a nonlinear isolated molecule. It is important to note that the systematic extension of this approach to higher orders in the Taylor expansion or the moment expansion poses no conceptual difficulties. For the second order traced moment expansion altogether 9 states have to be considered, 3 for the linear order and 5 + 1 for the second order.<sup>33</sup>

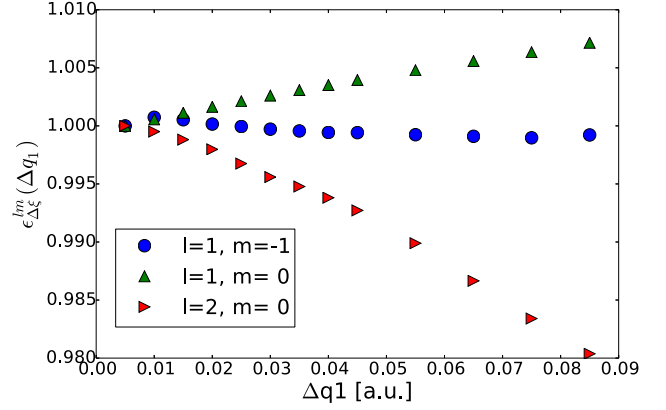


FIG. 4. Overlap between the explicitly calculated and the extrapolated states  $\xi_l^m$  for a single water molecule according to Eq. (22). The displacements roughly correspond to temperatures of 2000 K. Also the second order states show a well defined dependence on the nuclear displacement (a.u. refers to atomic units).

To quantify the errors of the expansion for finite displacements  $q_{n'} = q_{n'0} + \delta_{nn'} \Delta q_n$  we can compare the extrapolated states with the exact explicitly calculated state at the displaced geometry  $\xi_{[q]l}^m$  by means of their overlap

$$\epsilon_{\xi}^{lm}(\Delta q_n) = \frac{\langle \xi_{[q]l}^m | \xi_{[q_0]l}^m + \xi_{[q_0]l}^{m(n)} \Delta q_n \rangle}{\langle \xi_{[q]l}^m | \xi_{[q]l}^m \rangle}, \quad (21)$$

where ideally  $\epsilon_{\xi}^{lm}(\Delta q_n) = 1$  for an exact finite order expansion. As a more significant measure, we look at the corresponding relative errors for the displacement induced changes of the states

$$\epsilon_{\Delta\xi}^{lm}(\Delta q_n) = \frac{\langle \xi_{[q]l}^m - \xi_{[q_0]l}^m | \xi_{[q_0]l}^{m(n)} \Delta q_n \rangle}{\langle \xi_{[q]l}^m - \xi_{[q_0]l}^m | \xi_{[q]l}^m - \xi_{[q_0]l}^m \rangle}. \quad (22)$$

The relative errors of the moment expanded states for different nuclear displacements are shown in Fig. 4. The deviation of the extrapolated states from the explicitly diagonalized states shows a smooth and continuous behavior. Its relative error is of the order of a few percent in the displacement range typical for nuclear vibrations. (Also the relative error of the change of the state is at most two percent over the relevant range.) These results confirm the validity of the first order expansion of the molecular geometry dependence of the electronic susceptibility for geometries expected within an ambient temperature MD trajectory ( $\Delta q = 0.045$  Å).

#### V. FINITE TEMPERATURE RAMAN SIGNATURE OF A WATER MOLECULE

A first important application of the geometry dependence of the molecular polarizability tensor  $\alpha_{[\mathbf{R}]}$  is the calculation of an anharmonic Raman spectra.<sup>40-46</sup> By construction of the moment expanded representation, the full polarizability can be determined from the first order  $l = 1$  angular momentum contribution  $\chi_{[q]l}^{l=1}(\mathbf{r}, \mathbf{r}')$  of the susceptibility. This requires only three moment expanded states for  $m \in -1, 0, 1$



$$\alpha_{[q]ij} = \sum_{m=-1}^1 M_{[q]i}^m M_{[q]j}^m, \quad (23)$$

where  $M_{[q]i}^m = \int r_i \xi_{[q]i}^m(\mathbf{r}) d^3r$  is the linear order Cartesian multipole moment in direction  $i$  of state  $l = 1, m$ .<sup>32,33</sup> In analogy to Eq. (20) the polarizability tensor can be expanded in terms of derivatives of the first moments

$$\alpha_{[q]ij} \approx \alpha_{[q_0]ij} + \sum_{n=1}^{N_{\text{int}}} \alpha_{[q_0]ij}^{(n)} \Delta q_n + \frac{1}{2} \sum_{n,n'=1}^{N_{\text{int}}} \alpha_{[q_0]ij}^{(nn')} \Delta q_n \Delta q_{n'}. \quad (24)$$

Plugging Eq. (23) into Eq. (24) yields for the first and second order,

$$\alpha_{[q_0]ij}^{(n)} = \sum_{m=-1}^1 (M_{[q_0]i}^{m(n)} M_{[q_0]j}^m + M_{[q_0]i}^m M_{[q_0]j}^{m(n)}), \quad (25)$$

$$\alpha_{[q_0]ij}^{(nn')} = \sum_{m=-1}^1 (M_{[q_0]i}^{m(nn')} M_{[q_0]j}^m + M_{[q_0]i}^{m(n)} M_{[q_0]j}^{m(n')} + M_{[q_0]i}^{m(n')} M_{[q_0]j}^{m(n)} + M_{[q_0]i}^m M_{[q_0]j}^{m(nn')}). \quad (26)$$

The evaluation of the moment derivatives can be done either via calculation of the moments of the derivative state or equivalently via finite difference derivative of the moments for the different geometries.

In Fig. 5 we show the correlation of the first and second order moments for nuclear displacements along an internal coordinate (O–H distance). The first and second order moments, derived from extrapolated values and explicit calculations, are in excellent agreement.

We can calculate the change of the polarizability tensor from the change of the linear moments of the moment expanded states according to Eq. (25) or Eq. (26). This change of the polarizability for nuclear displacements along different internal coordinates is shown in Fig. 6. We would like to stress that these results are a verification and benchmark of our method, that is, the explicit geometry dependence of the general response function. Of course, a simple interpolation

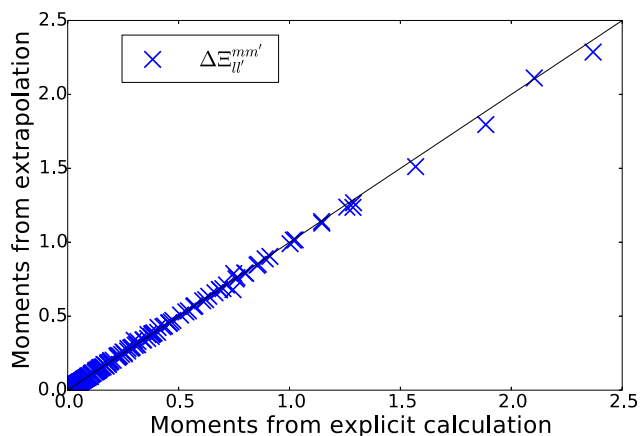


FIG. 5. Correlation of the vibrationally induced variations of  $\Delta \Xi_{ll'}^{mm'}$  obtained from explicit calculation and extrapolation. The  $\Delta \Xi_{ll'}^{mm'}$  are depicted for all moments up to  $\Xi_{2,2}^{2,2}$  over a representative range. The displacements correspond to temperatures of 2000 K.

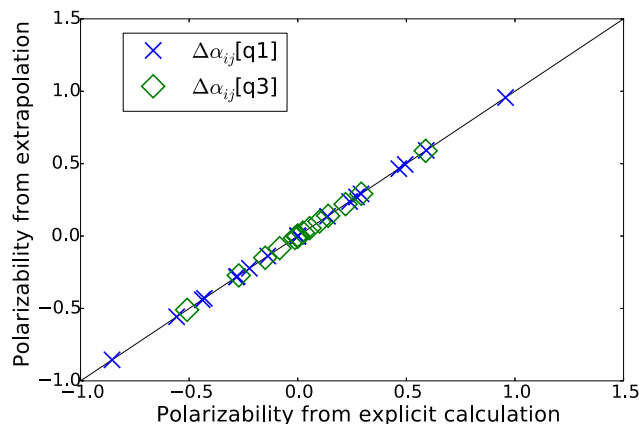


FIG. 6. Correlation of the geometry dependence of the polarizability for nuclear displacements along the O–H-bond  $q_1$  and the H–O–H-angle  $q_3$ . Explicit calculations for the corresponding geometry are correlated with interpolated values via second order Taylor expansion Eq. (26).

and tabulation of the polarizability tensor is much easier to achieve, this is however not our primary goal.

In order to obtain an estimate for the deviation of our approach from the exact calculation we have so far resorted to the comparison with the explicitly calculated susceptibility at the displaced geometry of interest. By construction, the deviation is zero for the equilibrium geometry and increases for larger displacements, especially if individual Cartesian coordinates are far from equilibrium.

For the application we are aiming at, i.e., spectroscopic sampling along molecular dynamics simulations or molecular dynamics simulations via fragmentation, the actual error of our method has to be weighted with the probability of the corresponding configuration during the simulation under ambient conditions. An explicit calculation of the polarizability via diagonalization of the non-local electronic susceptibility for many configurations along a molecular dynamics is far to expensive, which is the motivation for this work after all. We therefore compare the polarizability changes in our approach to direct DFPT calculations of the polarizability.

We have determined an average deviation of our method from direct DFPT calculations for the change of the polarizability along a molecular dynamic simulation. As a benchmark system we choose again a water molecule at 350 K. The average relative error of the induced changes of the polarizability using a first order Taylor expansion is 4.43% for the trace and 7.47% for the full tensor. Correcting with second order terms including mixed derivatives we obtain a mean relative deviation of 1.69% for the trace and 3.74% for the full tensor. These results nicely confirm our idea that the explicit geometry dependence for relevant vibrational displacements can be obtained via first- or second order Taylor expansion of the electronic susceptibility.

A direct illustration of the applicability of these results is the calculation of a Raman spectrum, which is obtained by Fourier-transform autocorrelation of the vibrationally induced polarizability changes.<sup>47,48</sup> In Fig. 7 we show the Raman spectrum of an isolated water molecule at 350 K. Our explicit geometry dependence of the electronic susceptibility can

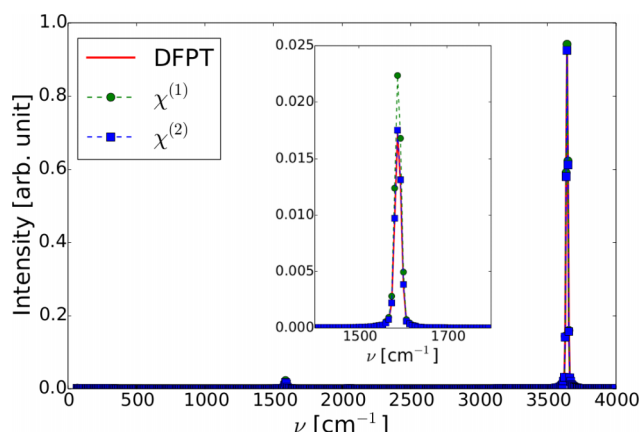


FIG. 7. Raman scattering intensities from vibrationally induced polarizability changes. The red line shows the explicit DFPT calculation as a reference. The first order (green) and second order (blue) Taylor expansion of the electronic susceptibility can reproduce all the features of the spectrum with systematic convergence to the reference.

reproduce all the features of the total spectrum and show only minor deviations of the intensity at the bending peak.

## VI. CONCLUSION

In this work we validate the accuracy of a first and second order Taylor expansion of the molecular geometry dependence of the full non-local electronic susceptibility. Our long-term motivation is to work towards a new generation of fragmentation based molecular dynamics, which requires an efficient calculation of electronic response properties. The calculation of electronic response properties requires a much higher initial (once for all) effort but has a considerable better scaling for repeated application (e.g., along an AIMD trajectory) compared to trajectory sampling with an explicit DFPT approach. The key step enabling a Taylor expansion of the geometry dependence of the susceptibility is to resort to an explicit representation via a spectral decomposition of the response function. Combined with our recently developed moment expanded representation, this yields an efficiently differentiable representation of the response function. We validated our method numerically for an isolated water molecule, yielding a quantitative agreement for the Raman spectrum of water in the gas phase. We could show that a first order Taylor expansion is sufficient to reproduce moments and states within an error of two percent for molecular displacements corresponding to ambient temperatures.

## VII. COMPUTATIONAL DETAILS

The presented theory was implemented in our development version of the CPMD<sup>49</sup> electronic structure package. The calculations have been performed using density functional perturbation theory<sup>15–18,50,51</sup> with Troullier-Martins<sup>52</sup> pseudo potentials in the Becke<sup>53</sup> Lee-Yang-Parr<sup>54</sup> approximation for the exchange correlation kernel. We have employed a plane wave cutoff of 70 Ry and used the optimized geometry of an isolated water at this level of theory for all our calculations. In practice, the infinite space of the eigenfunctions is restricted

to a subspace of 5000 converged eigenstates. The molecular dynamics was generated with the CP2K<sup>55,56</sup> program package using the TZV2P-MOLOPT-GTH basis<sup>57</sup> and GTH pseudo potentials<sup>58–60</sup> with a 0.5 fs time step. The temperature was set to 350 K by a Nosé-Hoover chain thermostat.<sup>61,62</sup> For the comparison of polarizability tensors we employ the Frobenius norm, i.e.,  $\epsilon = |\alpha_{\text{ref}} - \alpha|_F / |\alpha_{\text{ref}}|_F$ .

- <sup>1</sup>A. D. Buckingham, *Adv. Chem. Phys.* **12**, 107 (1967).
- <sup>2</sup>A. J. Stone, *Chem. Phys. Lett.* **83**, 233 (1981).
- <sup>3</sup>G. Naray-Szabo and G. G. Ferenczy, *Chem. Rev.* **95**, 829 (1995).
- <sup>4</sup>J. G. Angyán, C. Chipot, F. Dehez, C. Hättig, G. Jansen, and C. Millot, *J. Comput. Chem.* **24**, 997 (2003).
- <sup>5</sup>D. Elking, T. Darden, and R. J. Woods, *J. Comput. Chem.* **28**, 1261 (2007).
- <sup>6</sup>R. J. Wheatley, *Mol. Phys.* **79**, 597 (1993).
- <sup>7</sup>G. G. Hall and C. M. Smith, *Int. J. Quantum. Chem.* **25**, 881 (1984).
- <sup>8</sup>K. Eichkorn, O. Treutler, H. Öhm, M. Häser, and R. Ahlrichs, *Chem. Phys. Lett.* **240**, 283 (1995).
- <sup>9</sup>G. A. Cisneros, J.-P. Piquemal, and T. A. Darden, *J. Chem. Phys.* **123**, 044109 (2005).
- <sup>10</sup>B. Jeziorski, R. Moszynski, and K. Szalewicz, *Chem. Rev.* **94**, 1887 (1994).
- <sup>11</sup>D. M. Benoit, D. Sebastiani, and M. Parrinello, *Phys. Rev. Lett.* **87**, 226401 (2001).
- <sup>12</sup>A. Hebelmann and G. Jansen, *Chem. Phys. Lett.* **367**, 778 (2003).
- <sup>13</sup>A. J. Misquitta, B. Jeziorski, and K. Szalewicz, *Phys. Rev. Lett.* **91**, 033201 (2003).
- <sup>14</sup>X. Gonze, *Phys. Rev. A* **52**, 1086 (1995).
- <sup>15</sup>X. Gonze, *Phys. Rev. A* **52**, 1096 (1995).
- <sup>16</sup>A. Putrino, D. Sebastiani, and M. Parrinello, *J. Chem. Phys.* **113**, 7102 (2000).
- <sup>17</sup>S. Baroni, S. de Gironcoli, A. dal Corso, and P. Giannozzi, *Rev. Mod. Phys.* **73**, 515 (2001).
- <sup>18</sup>T. Watermann, A. Scherrer, and D. Sebastiani, in *Many-Electron Approaches in Physics, Chemistry and Mathematics*, edited by V. Bach and L. Delle Site, Mathematical Physics Studies (Springer International Publishing, 2014), pp. 97–110.
- <sup>19</sup>A. J. Stone and M. Alderton, *Mol. Phys.* **56**, 1047 (1985).
- <sup>20</sup>J. D. Jackson, *Classical Electrodynamics*, 3rd ed. (Wiley, 1998), pp. 145–151.
- <sup>21</sup>A. J. Misquitta, A. J. Stone, and F. Fazeli, *J. Chem. Theory Comput.* **10**, 5405 (2014).
- <sup>22</sup>A. J. Stone, *Mol. Phys.* **56**, 1065 (1985).
- <sup>23</sup>C. R. L. Sueur and A. J. Stone, *Mol. Phys.* **78**, 1267 (1993).
- <sup>24</sup>C. R. L. Sueur and A. J. Stone, *Mol. Phys.* **83**, 293 (1994).
- <sup>25</sup>A. J. Misquitta and A. J. Stone, *J. Chem. Phys.* **124**, 024111 (2006).
- <sup>26</sup>A. Stone, *The Theory of Intermolecular Forces*, 2nd ed. (Oxford University Press, 2013).
- <sup>27</sup>R. Chelli, R. Righini, S. Califano, and P. Procacci, *J. Mol. Liq.* **96–97**, 87 (2002).
- <sup>28</sup>P. Paricaud, M. Predota, A. A. Chialvo, and P. T. Cummings, *J. Chem. Phys.* **122**, 244511 (2005).
- <sup>29</sup>D. Martin and G. G. Hall, *Theor. Chem. Acc.* **59**, 281 (1981).
- <sup>30</sup>R. J. Wheatley and J. B. O. Mitchell, *J. Comput. Chem.* **15**, 1187 (1994).
- <sup>31</sup>D. M. Elking, G. A. Cisneros, J.-P. Piquemal, T. A. Darden, and L. G. Pedersen, *J. Chem. Theory Comput.* **6**, 190 (2010).
- <sup>32</sup>A. Scherrer, V. Verschinin, and D. Sebastiani, *J. Chem. Theory Comput.* **8**, 106 (2012).
- <sup>33</sup>A. Scherrer and D. Sebastiani, *J. Comput. Chem.* **37**, 665 (2016).
- <sup>34</sup>S. Hamel, A. J. Williamson, H. F. Wilson, F. Gygi, G. Galli, E. Ratner, and D. Wack, *Appl. Phys. Lett.* **92**, 3115 (2008).
- <sup>35</sup>D. Lu, F. Gygi, and G. Galli, *Phys. Rev. Lett.* **100**, 147601 (2008).
- <sup>36</sup>H. F. Wilson, D. Lu, F. Gygi, and G. Galli, *Phys. Rev. B* **79**, 245106 (2009).
- <sup>37</sup>A. C. Ihrig, A. Scherrer, and D. Sebastiani, *J. Chem. Phys.* **139**, 094102 (2013).
- <sup>38</sup>K. L. C. Hunt, *J. Chem. Phys.* **90**, 4909 (1989).
- <sup>39</sup>K. L. C. Hunt, Y. Q. Liang, R. Nimalakirithi, and R. A. Harris, *J. Chem. Phys.* **91**, 5251 (1989).
- <sup>40</sup>J. S. Bader and B. J. Berne, *J. Chem. Phys.* **100**, 8359 (1994).
- <sup>41</sup>J. Borysow, M. Moraldi, and L. Frommhold, *Mol. Phys.* **56**, 913 (1985).
- <sup>42</sup>P. L. Silvestrelli, M. Bernasconi, and M. Parrinello, *Chem. Phys. Lett.* **277**, 478 (1997).
- <sup>43</sup>R. Iftimie and M. E. Tuckerman, *J. Chem. Phys.* **122**, 214508 (2005).
- <sup>44</sup>M.-P. Gaigeot and M. Sprik, *J. Phys. Chem. B* **107**, 10344 (2003).

- <sup>45</sup>J. Neugebauer, M. Reiher, C. Kind, and B. A. Hess, *J. Comput. Chem.* **23**, 895 (2002).
- <sup>46</sup>D. A. A. McQuarrie, *Statistical Mechanics* (University Science Books, 2000).
- <sup>47</sup>M. Thomas, M. Brehm, R. Fligg, P. Vohringer, and B. Kirchner, *Phys. Chem. Chem. Phys.* **15**, 6608 (2013).
- <sup>48</sup>R. Futrelle and D. McGinty, *Chem. Phys. Lett.* **12**, 285 (1971).
- <sup>49</sup>CPMD, Computer code, <http://www.cpmc.org/>, 2016.
- <sup>50</sup>P. Hohenberg and W. Kohn, *Phys. Rev.* **136**, B864 (1964).
- <sup>51</sup>W. Kohn and L. J. Sham, *Phys. Rev.* **140**, A1133 (1965).
- <sup>52</sup>N. Troullier and J. L. Martins, *Phys. Rev. B* **43**, 1993 (1991).
- <sup>53</sup>A. D. Becke, *Phys. Rev. A* **38**, 3098 (1988).
- <sup>54</sup>C. Lee, W. Yang, and R. G. Parr, *Phys. Rev. B* **37**, 785 (1988).
- <sup>55</sup>G. Lippert, J. Hutter, and M. Parrinello, *Mol. Phys.* **92**, 477 (1997).
- <sup>56</sup>J. VandeVondele, M. Krack, F. Mohamed, M. Parrinello, T. Chassaing, and J. Hutter, *Comput. Phys. Commun.* **167**, 103 (2005).
- <sup>57</sup>J. VandeVondele and J. Hutter, *J. Chem. Phys.* **127**, 114105 (2007).
- <sup>58</sup>S. Goedecker, M. Teter, and J. Hutter, *Phys. Rev. B* **54**, 1703 (1996).
- <sup>59</sup>C. Hartwigsen, S. Goedecker, and J. Hutter, *Phys. Rev. B* **58**, 3641 (1998).
- <sup>60</sup>M. Krack, *Theor. Chem. Acc.* **114**, 145 (2005).
- <sup>61</sup>S. Nosé, *J. Chem. Phys.* **81**, 511 (1984).
- <sup>62</sup>W. G. Hoover, *Phys. Rev. A* **31**, 1695 (1985).

### 3.3 Proton Conduction in Solid Acids

In the article [Dreßler, Sebastiani, *Phys. Chem. Chem. Phys.*, **2020**.] in this thesis, the low-temperature phase (LTP) and high-temperature phase (HTP) of  $\text{CsH}_2\text{PO}_4$ ,  $\text{CsHSO}_4$  and  $\text{CsHSeO}_4$  are investigated by means of *ab initio* molecular dynamics simulations (AIMD) in order to obtain a general picture of the proton transfer mechanism. The following questions are addressed by the article:

- How can we explain the big difference in proton conductivity between the high- and low-temperature phases?
- How can we describe the general mechanism for proton conduction in the  $\text{CsH}_y\text{XO}_4$  (X= S, P, Se,  $y = 1, 2$ ) family by means of AIMD simulations?
- Can we qualitatively and quantitatively explain the trend of high proton conductivity in the high-temperature phases of  $\text{CsHSeO}_4$ ,  $\text{CsHSO}_4$  and  $\text{CsH}_2\text{PO}_4$ ?

Our treatment of the first two questions will contribute three specific examples to overall knowledge of proton motion in the solid acids family, while our work on question number three will shed light on new aspects of the long-range proton transfer mechanism.

In the article [Dreßler, Sebastiani, *Phys. Chem. Chem. Phys.*, **2020**.] in this thesis, the Grotthuss-type proton conduction mechanism in these compounds is analyzed in terms of the proton transfer rate and the relaxation rate of the molecular environment (rotation rate of the anions). It is demonstrated that only the balanced interplay of both contributing processes enables effective long-range proton transfer. While (massive) proton rattling is observed for the low-temperature phases investigated within this work, the oxo-anions do not rotate in these compounds at the time scale of the AIMD simulations. Due to this highly ordered hydrogen bond network, no long-range proton transfer can be observed for the LTPs of the solid acids  $\text{CsH}_2\text{PO}_4$ ,  $\text{CsHSO}_4$  and  $\text{CsHSeO}_4$ .

A high diffusion coefficient can be reported for all HTPs of the solid acids ( $1.4 \cdot 10^{-3}$ – $1.5 \cdot 10^{-2} \text{ \AA}^2/\text{ps}$ ). While the differences in the diffusion coefficients of  $\text{CsH}_2\text{PO}_4$  and  $\text{CsHSO}_4$  are within the uncertainty of the results of the AIMD simulations,  $\text{CsHSeO}_4$  possesses the highest proton conductivity. Among the three HTPs, the highest proton transfer rate and the lowest anion reorientation frequency is observed for  $\text{CsH}_2\text{PO}_4$ . For  $\text{CsHSO}_4$ , the relation between the two rates is opposed. The highest proton conductivity in  $\text{CsHSeO}_4$  can be explained by the combination of intermediate proton transfer and anion reorientation rates, enabling effective long-range proton transfer.

In addition to *ab initio* molecular dynamics, the importance of the proton transfer frequency and relaxation rate of the molecular environment was also studied within our scale-bridging combined Molecular Dynamics/Lattice Monte Carlo approach (cMD/LMC) for the simulation of long-range proton transfer (the cMD/LMC method is described in detail in sections 2.8 and 3.1). In this approach, the system is reduced to the oxygen sites. Each of these sites can be occupied by a maximum of one proton. While the evolution of the oxygen topology is obtained from the underlying molecular dynamics simulation, the long-range proton motion is modeled in the Monte Carlo part by proton jumps on the oxygen grid according to a distance-dependent jump rate. This jump rate yields the oxygen-oxygen distance-dependent probability for a proton transfer event within a certain time interval. The jump rate can be directly determined from an AIMD trajectory (for details cf. sections 2.8 and 3.1).



Following these explanations, the proton transfer frequency and relaxation rate of the molecular environment are reflected within the cMD/LMC method by the two fundamental parameters of the approach: the proton jump rate and the update of the oxygen topology from an MD trajectory. The two parameters are only linked as they are determined from the same AIMD trajectory of an underlying compound. The combination of proton jump rates and oxygen trajectories from different compounds appears to be arbitrary, but opens up new possibilities for the investigation of the proton conduction mechanism.

In the article [Dreßler, Kabbe, Sebastiani, *J. Phys. Chem. C*, **2016**.] in this thesis, the limiting factor of the proton conduction mechanism in CsHSO<sub>4</sub> and CsH<sub>2</sub>PO<sub>4</sub> is investigated by exchanging the jump rates and trajectories of the oxygen atoms between these two compounds. The results of these cMD/LMC studies confirm that the proton transfer frequency is decisive for long-range proton transfer in CsHSO<sub>4</sub>, while the efficiency of proton transfer is limited in CsH<sub>2</sub>PO<sub>4</sub> by the relaxation rate of the molecular environment.

Furthermore, it is demonstrated in the article [Dreßler, Kabbe, Sebastiani, *J. Phys. Chem. C*, **2016**.] that the cMD/LMC approach is able to reproduce (in a black box manner) the very high proton diffusion coefficient in the high-temperature phases of CsHSO<sub>4</sub> and CsH<sub>2</sub>PO<sub>4</sub> and the very low proton diffusion coefficient in the low-temperature phases of CsH<sub>2</sub>PO<sub>4</sub>. In contrast with other lattice site jump models,<sup>250</sup> the key factor achieving a correct simulation of the diffusion coefficients is the periodic updating of the oxygen sites from a molecular dynamics trajectory, which reflects the evolution of the underlying supra-molecular structure. The diffusion coefficients were calculated with the cMD/LMC scheme using the dynamically updated oxygen grid and from various fixed oxygen grids (oxygen positions from crystal structure, geometry optimization or averaged positions from the MD). All calculations based on fixed oxygen topologies yielded diffusion coefficients that were several magnitudes too small.

In the article [Dreßler, Kabbe, Sebastiani, *Fuel Cells*, **2016**.] in this thesis, theoretical findings with regard to proton transfer in aqueous media and water-free proton conductors are reviewed.

### 3.3.1 Article X: Effect of Anion Reorientation on Proton Mobility in the Solid Acids Family CsH<sub>y</sub>XO<sub>4</sub> (X = S, P, Se; y = 1, 2) from AIMD

Christian Dreßler and Daniel Sebastiani.

Effect of chemical environment on proton mobility in the solid acids family CsH<sub>y</sub>XO<sub>4</sub> (x = S, P, Se; y = 1, 2) from *ab initio* molecular dynamics simulations.

*Physical Chemistry Chemical Physics*, 22:10738–10752, **2020**.

In this article, I performed all AIMD simulations. I designed and conducted all analysis of the AIMD trajectories. D. Sebastiani supervised the project and provided me with valuable advice during the project work and while I was writing the manuscript.



Cite this: DOI: 10.1039/c9cp06473g

# Effect of anion reorientation on proton mobility in the solid acids family $\text{CsH}_y\text{XO}_4$ ( $\text{X} = \text{S}, \text{P}, \text{Se}, y = 1, 2$ ) from *ab initio* molecular dynamics simulations†

Christian Dreßler and Daniel Sebastiani \*

The high temperature phases of the solid acids  $\text{CsHSeO}_4$ ,  $\text{CsHSO}_4$  and  $\text{CsH}_2\text{PO}_4$  show extraordinary high proton conductivities, while the low temperature phases do not conduct protons at all. We systematically investigate proton dynamics in the low and high temperature phases of these compounds by means of *ab initio* molecular dynamics simulations in order to develop a general picture of the proton transfer mechanism. For all of these compounds, proton conduction follows a Grotthuss mechanism via a combined proton transfer and subsequent structural reorientation of the environment. We demonstrate that the drastically reduced conductivity of the low temperature phases is caused by a highly ordered, rigid hydrogen bond network, while efficient long range proton transfer in the high temperature phases is enabled by the interplay of high proton transfer rates and frequent anion reorientation. Furthermore, we present a simple descriptor for the quantitative prediction of the diffusion coefficient within the solid acids family. As a side result, we show that the rate of the most elementary proton hopping reaction depends on the heavy-atom configuration of the nearest atoms in a ubiquitous manner, and is in turn almost independent from the global nature of the compound, *i.e.* whether it is organic or inorganic, ordered or disordered.

Received 29th November 2019,  
Accepted 26th February 2020

DOI: 10.1039/c9cp06473g

[rsc.li/pccp](http://rsc.li/pccp)

## 1 Introduction

Molecular dynamics (MD) simulations are the most common tool for the investigation of the dynamics of atomistic processes. The temporal evolution of a system of particles is obtained by numerical integration of the Newtonian equations of motions constituted by the forces acting on each particle. The quality of the results of an MD simulation is highly dependent on the quality of the calculated forces. The most accurate approach for the calculation of the forces is the approximate solution of the Schrödinger equation. This computational demanding ansatz is known as first principle or *ab initio* molecular dynamics (AIMD). Using the truly predictive power of the *ab initio* techniques, unknown systems can be studied and theoretical challenging effects starting from polarization up to the breaking of covalent bonds are intrinsically included within the simulation. Thus electrochemical phenomena<sup>1,2</sup> and surface interactions,<sup>3–8</sup> ion transport in condensed phases,<sup>9–14</sup> chemical reactions<sup>15</sup> as well as solvation in novel complex liquids such as ionic liquids<sup>16–18</sup> are investigated by AIMD simulations.

In this work, we will use AIMD to study proton conduction in solid acids. Solid acids have been an object of high interest for chemists and physicists in the last years, as these compounds show extraordinarily high proton conductivities.<sup>19–28</sup> Compared to other state of the art proton conductors like Nafion, residual water is not a prerequisite for the proton conduction.<sup>29</sup> In particular members of the  $\text{CsH}_y\text{XO}_4$  ( $\text{X} = \text{S}, \text{P}, \text{Se}, y = 1, 2$ ) family are promising candidates for application as proton exchange membrane materials (PEM) in fuel cells.<sup>27,30–32</sup> For  $\text{CsH}_2\text{PO}_4$  the successful usage in a fuel cell was already demonstrated.<sup>33</sup> The solid acids of the  $\text{CsH}_y\text{XO}_4$  ( $\text{X} = \text{S}, \text{P}, \text{Se}, y = 1, 2$ ) family undergo at least one phase transition at higher temperature (often between 400 and 500 K), corresponding to a drastic increase of the proton conductivity.<sup>34</sup>

In this article, we will focus on  $\text{CsH}_2\text{PO}_4$ ,  $\text{CsHSO}_4$  and  $\text{CsHSeO}_4$ . Table 1 provides the temperature of the phase transition as well as the space groups of the low temperature phases (LTP) and high temperature phases (HTP) of these compounds.

### 1.1 Crystal structures of the HTP and LTP of $\text{CsH}_2\text{PO}_4$

The crystal structure of the monoclinic phase (LTP) of  $\text{CsH}_2\text{PO}_4$  is shown in Fig. 1.<sup>35</sup> The hydrogen bond network within this phase is constituted by two different types of hydrogen bonds.<sup>35</sup> Symmetric double-minima hydrogen bonds (2.48 Å) form one dimensional chains of oxo-anions in *b* direction. These chains

Institute of Chemistry, Martin Luther University Halle-Wittenberg,  
Von-Danckelmann-Platz 4, 06120 Halle (Saale), Germany.  
E-mail: daniel.sebastiani@chemie.uni-halle.de

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c9cp06473g

**Table 1** Characteristic crystal structure parameters for the high and low temperature phases (HTP/LTP) of CsH<sub>2</sub>PO<sub>4</sub>, CsHSeO<sub>4</sub> and CsHSO<sub>4</sub><sup>34</sup>

	CsH <sub>2</sub> PO <sub>4</sub>	CsHSO <sub>4</sub>	CsHSeO <sub>4</sub>
HTP	<i>Pm</i> $\bar{3}$ <i>m</i> Cubic	<i>I</i> 4 <sub>1</sub> / <i>amd</i> Tetragonal	<i>I</i> 4 <sub>1</sub> / <i>amd</i> Tetragonal
LTP	<i>P</i> 2 <sub>1</sub> / <i>m</i> Monoclinic	<i>P</i> 2 <sub>1</sub> / <i>c</i> Monoclinic	<i>P</i> 2 <sub>1</sub> / <i>c</i> Monoclinic
$\lambda_c$	503 K	414 K	384 K

are linked by asymmetric single-minimum hydrogen bonds (2.54 Å) leading to two dimensional layers of hydrogen bonded tetrahedra. Hydrogen atoms involved in the symmetric hydrogen bond with the double potential well undergo rapid exchange between two different oxygen sites.<sup>36–38</sup>

One unit cell of the crystal structure of the cubic phase (HTP) of CsH<sub>2</sub>PO<sub>4</sub> is presented in Fig. 2.<sup>39</sup> The large number of partially colored atoms (partial occupation of the crystal sites) indicates a strong disorder of the oxygen atoms. Twenty-four partial occupied oxygen sites are reported resulting in six different orientation of the phosphate tetrahedra. The PO<sub>4</sub> groups reside in the center of the cubic unit cell. The reorientational disorder is originated by the mismatch of the tetrahedral oxo-anion unit and the octahedral symmetry of the cubic center. An alignment of the P–O bonds with the (111) body diagonal would retain the overall cubic symmetry while removing the octahedral symmetry from the center. In fact, this is not possible due to the small phosphorus–caesium distance of 4.30 Å in this direction. Displacing the oxygen atoms from this diagonal, retains the cubic symmetry and results in six different orientations of the phosphate anions depicted in Fig. 2.<sup>40</sup> Besides a strong (single minimum) hydrogen bond of 2.46 Å other hydrogen bonds of 2.76 Å, 2.85 Å and 3.03 Å length are possible from the crystal structure.

## 1.2 Crystal structures of the HTP and LTP of CsHSO<sub>4</sub> and CsHSeO<sub>4</sub>

The low as well as the high temperature phase of CsHSO<sub>4</sub> and CsHSeO<sub>4</sub> are isostructural. Thus, we are only discussing the crystal structure of CsHSO<sub>4</sub>.

Three different phases of CsHSO<sub>4</sub> (phase I, II, and III) are reported within the temperature range between room temperature and approximate 500 K.<sup>41–43</sup> Within this article, we are

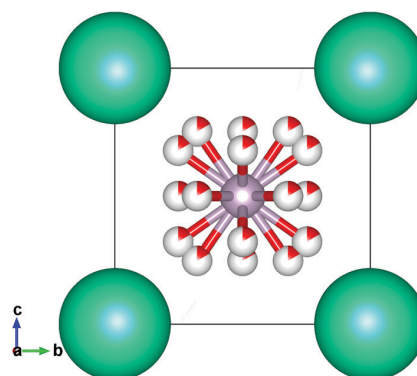


Fig. 2 One unit cell of the crystal structure of the cubic phase (HTP) of CsH<sub>2</sub>PO<sub>4</sub> is shown. Partial occupancies are presented by partial coloring of the atoms. Red: oxygen, green: caesium, lilac: phosphorus.<sup>39,101</sup>

interested in the proton conduction mechanism just above and just below the super protonic phase transition. Therefore, we restrict our investigations to phase II (LTP) and phase III (HTP). The crystal structures of the monoclinic (LTP) and the tetragonal (HTP) phase of CsHSO<sub>4</sub> are presented in Fig. 3 and 4. While the monoclinic phase forms infinite chains of hydrogen bonded sulfate tetrahedra, cyclic dimers linked by hydrogen bonds are reported for the tetragonal phase. This structural motifs are also known from other solid acids and the hydrogen bonds within zigzag chains tend to be stronger than hydrogen bonds formed by the cyclic dimers.<sup>38</sup> For CsHSO<sub>4</sub> this also holds for the monoclinic and the tetragonal phase and hydrogen bond lengths of 2.63 Å (monoclinic phase) and 2.79 Å (tetragonal phase) can be observed from the crystal structures. Slightly different crystal structures of the tetragonal phase were proposed by Jiráak *et al.*,<sup>44</sup> Merinov *et al.*<sup>45</sup> and Belushkin *et al.*<sup>46</sup> These structures differ by the number and the position of the partially occupied oxygen positions. In Fig. 4, we present the crystal structure with two possible orientations of the oxygen tetrahedron reported by Jiráak *et al.*

## 1.3 High proton conductivity in the high temperature phases of the solid acids

Solid acids were studied extensively and sometimes controversially by experiment and theory.<sup>40,47–55</sup> The structure and the

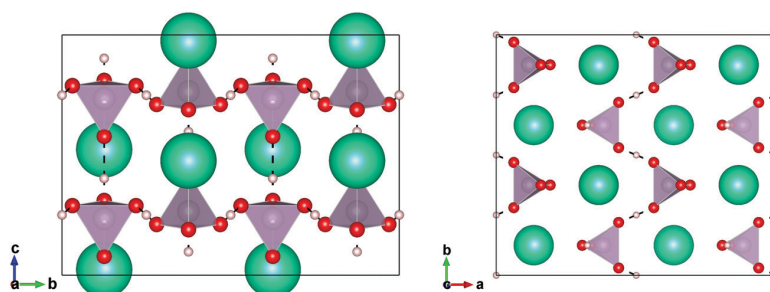


Fig. 1 Eight unit cells of the crystal structure of the monoclinic phase (LTP) of CsH<sub>2</sub>PO<sub>4</sub> are shown with different lattice orientations. Red: oxygen, green: caesium, gray: hydrogen.<sup>101</sup>

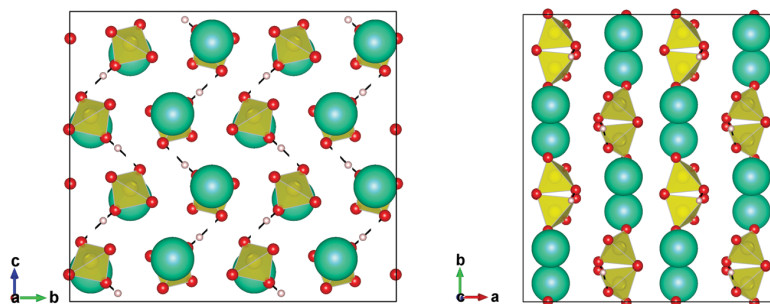


Fig. 3 Eight unit cells of the crystal structure of the monoclinic phase (LTP) of  $\text{CsHSO}_4$  are shown with different lattice orientations. Red: oxygen, green: caesium, gray: hydrogen.<sup>101</sup>

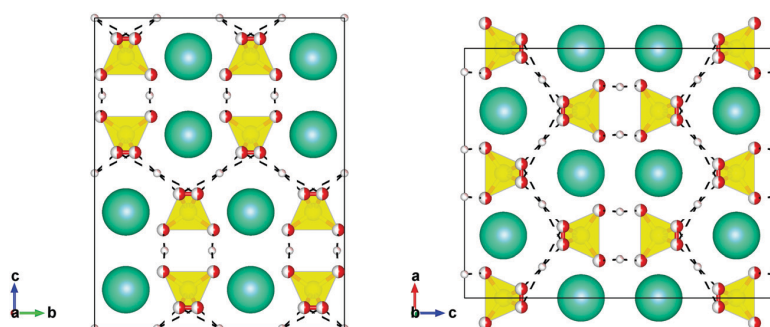


Fig. 4 Four unit cells of the crystal structure of the tetragonal phase (HTP) of  $\text{CsHSO}_4$  are shown with different lattice orientations. Red: oxygen, green: caesium, gray: hydrogen.<sup>101</sup>

phase transition were characterized by X-ray diffraction and calorimetry. The proton dynamics were investigated with the help of PFG-NMR techniques, neutron scattering and impedance spectroscopy.<sup>34,56–61</sup> The general proton conducting mechanism for all high temperature phases of the solid acids is already known and follows a Grotthuss mechanism: The combination of intra hydrogen bond proton transfer and subsequent relaxation of the environment (anion rotation) enables the high conductivity of these compounds.<sup>62,63</sup> While the phase transition from the low (monoclinic) to the high temperature phase (tetragonal or cubic) leads to an increase in the symmetry, it results in disorder for the oxygen sites. A case in point is  $\text{CsH}_2\text{PO}_4$ . Here, the orientational disorder of the oxygen tetrahedra arises from the incompatibility

of the tetrahedral oxo-anion with the octahedral symmetry of the cubic center. Thus, the oxygen atoms in the high temperature phases are free to vibrate and librate between crystallographic identical positions. The hydrogen bond network corresponding to such fast tetrahedral dynamics can be characterized as dynamically disordered and the dynamics of the oxo-anions is often described as “liquid like”. The disordered structure of the high temperature phases (in comparison to the low temperature phase) is illustrated in Fig. 5 for the specific example of  $\text{CsHSeO}_4$ .

Despite this consensus regarding the general picture of proton conduction in the high temperature phase, the relative importance of the different processes which are contributing to the Grotthuss mechanism is not clear. For  $\text{CsH}_2\text{PO}_4$ , theoretical

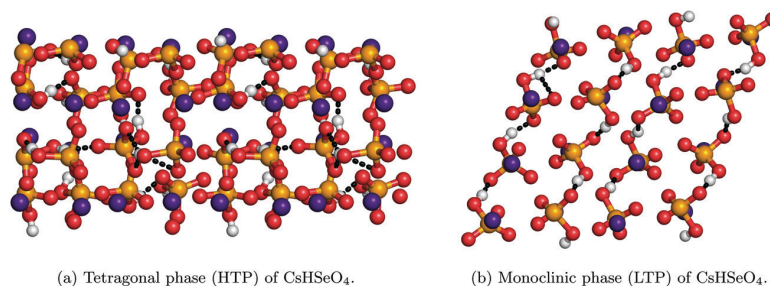


Fig. 5 Snapshots from the simulation of the HTP and LTP of  $\text{CsHSeO}_4$ . Hydrogen atoms are shown in white, oxygen in red, selenium in yellow and caesium in blue. Hydrogen bonds are represented by dashed lines.

studies suggested that the rotation of the anions (rather than the proton transfer rate between the anions) is the limiting step for the proton diffusion.<sup>59,64,65</sup> On the one hand, the same relationship of tetrahedral dynamics and proton transfer along hydrogen bonds was proposed for CsHSO<sub>4</sub>.<sup>66</sup> On the other hand, computational simulations<sup>65,67–70</sup> as well as experiments<sup>46,57,60,71,72</sup> suggested instead that the proton transfer rate (rather than the rotational anion dynamics) is limiting the proton diffusion.

From the theoretical point of view, static calculations of the electronic structure, classical and *ab initio* molecular dynamics were used to study the solid acids.<sup>64,66,68,69,73–76</sup> The focus of these investigations was on CsHSO<sub>4</sub>. Ke and Tanaka calculated activation energies for proton jumps in CsHSO<sub>4</sub> and showed that inter-anion proton transfer was forbidden.<sup>75,77</sup> Several classical MD studies of CsHSO<sub>4</sub> were conducted.<sup>67–69</sup> We want to point to the study of Chisholm *et al.* who investigated the phase transition from monoclinic to tetragonal CsHSO<sub>4</sub>.<sup>69</sup> Remarkably, they could predict the correct temperature of the phase transition. While the classical MD studies lacks the ability to simulate bond breaking, leading to fixed covalent O–H bonds/connection topology through the entire simulation, *ab initio* molecular dynamics (AIMD) allow to overcome these drawbacks and enable the simulation of real proton transfer events. Due to their high computational costs, AIMD methods can only be used to simulate several hundred atoms for several picoseconds. This timescale is comparable to the timescale of proton transfer itself. Thus, proton transfers are rare events in AIMD simulations, which becomes evident in the AIMD studies of CsH<sub>2</sub>PO<sub>4</sub> by Lee and Tuckerman and of CsHSO<sub>4</sub> by Wood and Marzari.<sup>64,66</sup> Tuckerman could reproduce the proton diffusion coefficient of CsH<sub>2</sub>PO<sub>4</sub> by a factor of two, whereas Marzari obtained a ten times larger diffusion coefficient compared with the experimental value.

Until now, the focus of the AIMD studies was on the high temperature phases.<sup>64–66,78</sup> The proton conductivity of the low temperature phases is drastically reduced compared with the proton conductivity of the high temperature phases. Nevertheless, the mechanism of proton conductivity in the paraelectric (or low temperature) phase of CsH<sub>2</sub>PO<sub>4</sub> were studied from the experimental side by <sup>17</sup>O, <sup>1</sup>H and <sup>31</sup>P NMR measurements.<sup>36,37,39,79</sup> This studies characterized the hydrogen bond network and proposed that H<sub>2</sub>PO<sub>4</sub><sup>–</sup> anion rotation plays a crucial role for the proton conduction. Kim *et al.* reported two different time constants for the hydrogen phosphate rotation corresponding to different strengths of the hydrogen bonds along the conduction path.<sup>36</sup> While the faster rotational process is enabled by local fluctuations of the protons within a hydrogen bond, the slower rotational process enforces the long range transport of a proton. In this work, we will shed light on the question of the timescale of the tetrahedral dynamics of the oxo-anions in the low temperature phases of the solid acid compound family.

The previous isolated pure AIMD studies of solid acids investigated only a single phase of a single compound (CsH<sub>2</sub>PO<sub>4</sub> and CsHSO<sub>4</sub>), neglecting the comparison between the different compounds. It is in particular important for the quantitative comparison of the contributing processes to the

overall conduction mechanism (short range proton transfer *vs.* tetrahedral dynamics) to define these processes in an identical manner. A case in point is the determination of the proton transfer rates between the anions, because there is no exact definition of a proton jump (“real” long range proton transfer *vs.* proton rattling). Thus, different protocols for the determination of the number of proton jumps can lead to different results. Within this work, we tackle this problem by the parallel simulation and analysis of different phases of different solid acids in order to draw a coherent picture of proton motion within this compounds.

We present the first AIMD study covering the LTP and the HTP phases of three different members of the CsH<sub>y</sub>XO<sub>4</sub> (X = S, P, Se, y = 1, 2) family (CsHSeO<sub>4</sub>, CsHSO<sub>4</sub> and CsH<sub>2</sub>PO<sub>4</sub>). We want to highlight, that neither the HTP of CsHSeO<sub>4</sub> nor the LTP of CsHSeO<sub>4</sub> and CsHSO<sub>4</sub> were investigated by AIMD simulations up to now. Only the combination of the entirety of the AIMD trajectories will enable us to draw and intuitive as well as extended picture of proton diffusion in these compounds, highlighting the overall similarities as well as the smaller differences for different compounds. The comparison of different compounds within one study is necessary, because many years of research on solid acids lead to a large amount of isolated articles with partial contrary results (*cf.* discussion of rate determining step of proton conduction in CsHSO<sub>4</sub>). In this article, we will answer the following questions:

- How can we explain the big difference in proton conductivity between high and low temperature phases?
- How can we describe the general mechanism for proton conduction in the CsH<sub>y</sub>XO<sub>4</sub> (X = S, P, Se, y = 1, 2) family from AIMD simulations?
- Can we explain (qualitatively and quantitatively) the trend of the high proton conductivity of the high temperature phases of CsHSeO<sub>4</sub>, CsHSO<sub>4</sub> and CsH<sub>2</sub>PO<sub>4</sub>?

While the answers to the first two questions will add three specific examples to the overall known picture of proton motion in the solid acids family, the answer to question number three will shed light on new aspects of the long-range proton transfer mechanism.

## 2 Computational details

We applied Born–Oppenheimer Molecular Dynamics (BOMD) using the CP2K<sup>80</sup> program package to simulate the proton transfer in a series of solid acids. We utilized the Quickstep module<sup>81</sup> and orbital transformation<sup>82</sup> for faster convergence. The electronic structure was calculated with density functional theory utilizing the PBE<sup>83–85</sup> functional. A basis set of the type DZVP-MOLOPT-SR-GTH<sup>86</sup> and GTH pseudopotentials<sup>87,88</sup> were applied. Furthermore, we used the empirical dispersion correction (D3) from Grimme.<sup>89</sup> The temperature was set by a Nosé–Hoover chain thermostat<sup>90–92</sup> (NVT ensemble). The timestep was chosen as 0.5 fs. The dimensions of the simulation box and the starting configurations of the systems were obtained from crystal structure data from the literature<sup>35,39,44,93,94</sup>



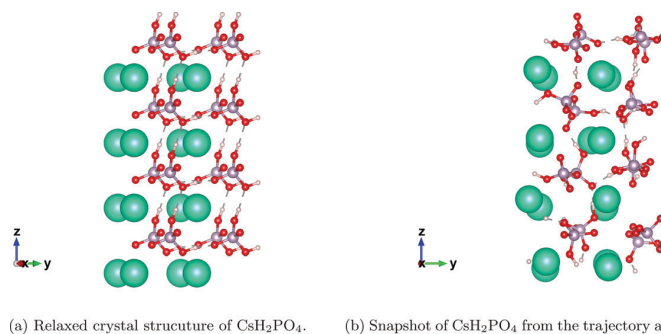


Fig. 6 Geometry optimized structure of  $\text{CsH}_2\text{PO}_4$  and a snapshot from the trajectory after equilibration. Hydrogen atoms are shown in white, oxygen in red, phosphorus in lilac and caesium in green. Hydrogen bonds are represented by dashed lines.

For the high temperature structures, partial occupancies of the oxygen atoms were observed in the literature. From the set of partial occupancies, we chose one possible set of coordinates and used them as initial structure for the geometry optimization. Afterwards, we performed a 10 ps equilibration run using massive and global thermostating. It turned out, that the choice of the initial set of coordinates from the partial occupancies did not affect the resulting trajectory due to the large amount of orientational disorder of the oxygen tetrahedra. The crystal structure of the HTP of  $\text{CsH}_2\text{PO}_4$  is presented in Fig. 2. In Fig. 6, we show the resulting geometry optimized structure of  $\text{CsH}_2\text{PO}_4$  and a snapshot from the trajectory after the equilibration procedure.

Systems of 16 formula units of  $\text{CsH}_2\text{PO}_4/\text{CsHSeO}_4/\text{CsHSO}_4$  were investigated above (510/500/500 K) and below (490/385/385 K) the phase transition. All production runs were performed for 250 ps.

### 2.1 Processing of the trajectory

A snapshot of the simulation of the HTP and LTP of  $\text{CsHSeO}_4$  is presented in Fig. 5. We calculated the mean square displacement of the protons in order to obtain the diffusion coefficients. The diffusion coefficient is accessible from the linear part of the mean square displacement (MSD) *via* the Einstein equation:

$$D_s = \frac{1}{6} \frac{d}{dt} [\text{MSD}(t)] \quad (1)$$

Diffusion coefficients, radial distribution functions (RDFs), vector-autocorrelation functions and combined distributed functions (CDFs) were created by Travis.<sup>95</sup> Histograms of the proton jumps and the covalent OH-bond autocorrelation function were calculated by own scripts. A detailed explanation of

the functions or scripts is given within the Result section, accompanying the corresponding plots.

## 3 Results

### 3.1 AIMD vs. experimental diffusion coefficients

In Table 2, experimental diffusion coefficients and diffusion coefficients obtained by AIMD simulations are depicted. From the experimental side, many different methods can be used to determine the diffusion coefficient (PFG-NMR, neutron scattering, impedance spectroscopy). None of these techniques alone is able to cover the whole range of protonic motion. Thus, we obtain for each method a different range of values for the diffusion coefficient corresponding to a specific aspect of long range proton transfer. A point in case is  $\text{CsH}_2\text{PO}_4$ . A diffusion coefficient of  $25 \times 10^{-3} \text{ \AA}^2 \text{ ps}^{-1}$  obtained by quasi elastic neutron scattering as well as a diffusion coefficient of  $2.9 \times 10^{-3} \text{ \AA}^2 \text{ ps}^{-1}$  obtained by PFG-NMR and impedance measurements were reported within the same study.<sup>59</sup> The authors of this study assigned the first diffusion coefficient to the (short range) proton motion within a hydrogen bond (proton rattling) and the second diffusion coefficient to the long range proton transfer. In addition, several other studies of  $\text{CsH}_2\text{PO}_4$  reported its conductivity.<sup>39,49,53,56,96</sup> Using the Nernst–Einstein equation,<sup>34</sup> we can also derive the diffusion coefficient from these values. From the variety of articles, we obtain diffusion coefficients within the range of  $0.5\text{--}6.5 \times 10^{-3} \text{ \AA}^2 \text{ ps}^{-1}$ .

For the sake of comparability, all simulations of the HTPs were conducted at the same temperature (500–510 K). No experimental data is available for  $\text{CsHSO}_4$  and  $\text{CsHSeO}_4$  at this temperature. Thus, the diffusion coefficients were extrapolated

Table 2 Diffusion coefficients for several solid acids

	$\text{CsH}_2\text{PO}_4$ HTP	$\text{CsH}_2\text{PO}_4$ LTP	$\text{CsHSeO}_4$ HTP	$\text{CsHSeO}_4$ LTP	$\text{CsHSO}_4$ HTP	$\text{CsHSO}_4$ LTP
Temperature [K]	510	490	500	385	500	385
Diffusion coefficient from AIMD [ $10^{-3} \text{ \AA}^2 \text{ ps}^{-1}$ ]	2.3	$\ll 0.01$	13.6	$\ll 0.01$	1.8	$\ll 0.01$
Diffusion coefficient from experiment [ $10^{-3} \text{ \AA}^2 \text{ ps}^{-1}$ ]	$2.9\text{--}25^a$	$\ll 0.01$	$\sim 9.6^b$	$\ll 0.01$	$\sim 5.2^b$	$\ll 0.01$
	$0.5\text{--}6.5^c$					

<sup>a</sup> Ishikawa, Maekawa, Yamamura, Kawakita, Shibata, Kawai, *Solid State Ionics*, 2008, **179**, 2345–2349. <sup>b</sup> Blinc, Dolinsek, Lahajnar, Zupancic, Shuvalov, Baranov, *Phys. Status Solidi B*, 1984, **123**, K83–K87. <sup>c</sup> See ref. 39, 49, 53, 56 and 96.

from data at lower temperatures.<sup>55</sup> For CsHSeO<sub>4</sub> diffusion coefficients from 400 to 444 K and for CsHSO<sub>4</sub> diffusion coefficients from 420 to 455 K were used to extrapolate the diffusion coefficients at 500 K. Our equation for extrapolation derived from data of ref. 55 is shown in the ESI†

All AIMD diffusion coefficients in Table 2 agree with the experimental diffusion coefficients within a factor of 3. The low temperature phases show no proton conductivity at all. The drastic increase of the proton diffusion coefficient for the phase transition from the low to the high temperature phase is clearly demonstrated for every compound (e.g. CsHSeO<sub>4</sub> from less than  $0.01 \times 10^{-3} \text{ \AA}^2 \text{ ps}^{-1}$  (LTP) to  $13.6 \times 10^{-3} \text{ \AA}^2 \text{ ps}^{-1}$  (HTP)). The HTP of the CsHSeO<sub>4</sub> shows the largest diffusion coefficient ( $13.6 \times 10^{-3} \text{ \AA}^2 \text{ ps}^{-1}$ ) compared with the HTPs of CsH<sub>2</sub>PO<sub>4</sub> ( $2.3 \times 10^{-3} \text{ \AA}^2 \text{ ps}^{-1}$ ) and CsHSO<sub>4</sub> ( $1.8 \times 10^{-3} \text{ \AA}^2 \text{ ps}^{-1}$ ). The diffusion coefficients of the HTPs of CsH<sub>2</sub>PO<sub>4</sub> and CsHSO<sub>4</sub> deviate by 25%, which is within the range of uncertainty for the diffusion coefficients obtained by AIMD simulations.

### 3.2 Characterization of hydrogen bonds and protonic motion

A careful inspection of the distribution of oxygen–oxygen and oxygen–hydrogen distances is important for the characterization of the hydrogen bonds in these systems.

**Intermolecular oxygen–oxygen RDFs.** The intermolecular oxygen–oxygen radial distribution functions (RDF) for the solid acids are depicted in Fig. 7. The RDFs show only small differences between the HTP and LTP of a given compound. Comparing the different compounds, it turns out that the RDFs of CsHSO<sub>4</sub> and CsHSeO<sub>4</sub> resemble each other but differ from the RDF of CsH<sub>2</sub>PO<sub>4</sub>. The first maximum of the O–O RDF of CsH<sub>2</sub>PO<sub>4</sub> is observed at 240 pm, while the first maximum of the O–O RDF of CsHSO<sub>4</sub> and CsHSeO<sub>4</sub> is observed at 260 pm. Furthermore, the first peak of the O–O RDF of CsH<sub>2</sub>PO<sub>4</sub> is much broader compared with the first peak of the O–O RDF of CsHSO<sub>4</sub> and CsHSeO<sub>4</sub>.

The grouping of the RDFs into two different types (CsH<sub>2</sub>PO<sub>4</sub> vs. CsHSO<sub>4</sub>/CsHSeO<sub>4</sub>) is expected from the crystal structures of these compounds. While the low as well as the high temperature phases

of CsHSO<sub>4</sub> and CsHSeO<sub>4</sub> are isostructural, the crystal structures of the corresponding phases of CsH<sub>2</sub>PO<sub>4</sub> are different (cf. discussion of crystal structures in the Introduction). First of all, the hydrogen bonds lengths detected from the crystal structures are shorter in CsH<sub>2</sub>PO<sub>4</sub> (indicating stronger hydrogen bonds) compared to CsHSO<sub>4</sub>/CsHSeO<sub>4</sub>. Second, the distribution of the hydrogen bond lengths (mainly originated by the number of possible orientations of the oxo-anions in the crystal structures) is broader for CsH<sub>2</sub>PO<sub>4</sub> compared to CsHSO<sub>4</sub>/CsHSeO<sub>4</sub>, which explains the much broader first peak of the O–O RDF of CsH<sub>2</sub>PO<sub>4</sub>.

A discussion of the intermolecular oxygen–hydrogen RDFs is given in the ESI† yielding the same conclusion regarding the proton diffusion mechanisms in the LTPs and HTPs: concerning the RDFs, there are no (significant) differences of the HTPs and LTPs. HTPs and LTPs form hydrogen bonds in a similar amount and strength.

**Distinction between proton conductors and insulators: combined distribution functions (CDF).** Considering the similarity of the RDFs between the HTP and LTP of a given compound, we could not distinguish whether long range proton transfer takes place within a compound or not. We will show in the following that combined distribution functions (CDFs) of the solid acids fill this gap, enabling a clear distinction between a proton conductor or insulator.

In a 2d-CDF plot, the values of two distribution functions are combined *via*  $\text{CDF}(d_1, d_2) = \text{RDF}(d_1) \cdot \text{RDF}(d_2)$ .<sup>95</sup> It turns out that the H–O<sub>initial donor</sub> ( $:=d_1$ ) and the H–O<sub>initial acceptor</sub> ( $:=d_2$ ) distances (cf. Fig. 8) are ideal parameters for  $(d_1, d_2)$  within the CDF( $d_1, d_2$ ). The CDFs shown in Fig. 9 are calculated by the application of the following protocol: at the initial frame of the trajectory, we created a list of distances of covalently bonded hydrogen–oxygen pairs. The oxygen atoms, involved within this covalent bond were termed as O<sub>initial donor</sub> (cf. Fig. 8) and the distances as H–O<sub>initial donor</sub> distance. An RDF of these distance is used as y-axis in Fig. 9. The remaining oxygen atoms (not covalently bound to the hydrogen atom at the initial frame of the trajectory) are considered as hydrogen bond acceptors, denoted by O<sub>initial acceptor</sub>. The RDF obtained from the distance of O<sub>initial acceptor</sub> and the hydrogen atom is used as x-axis in Fig. 9 (H–O<sub>initial acceptor</sub> distance).

Please keep in mind two important features of this protocol: first, we determine the initial donor oxygen atom O<sub>initial donor</sub> for a hydrogen atom within a hydrogen bond at the initial frame of the trajectory (and do not update this assignment, even in case of a proton jump). This definition will lead to the appearance of considerably larger O<sub>initial donor</sub>–H distances as soon as the initially chosen proton diffuses away from its originally donor oxygen. Second, while there is only one unique initial O<sub>initial donor</sub> oxygen atom, all the other oxygen atoms of the oxo-anions are considered as potential hydrogen accepting oxygen atoms O<sub>initial acceptor</sub>. Thus, very large distances for O<sub>initial acceptor</sub> and the hydrogen atom are already observe at the initial time frame.

In Fig. 8 three different sub-processes (A–C) of a proton transfer step are defined. Each of this sub-processes corresponds to a specific area in a CDF. We highlight these areas

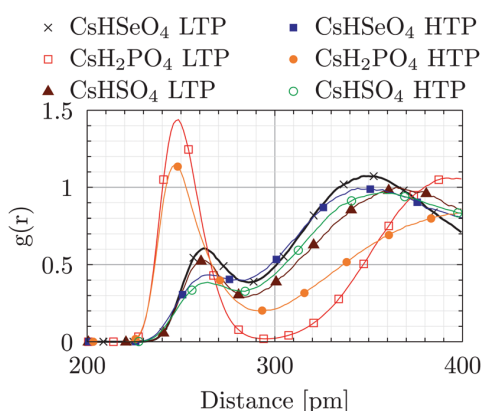


Fig. 7 RDF of the intermolecular oxygen–oxygen distances for the solid acids.

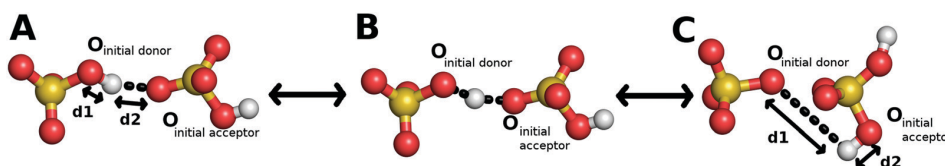


Fig. 8 Explanation of three different sub-processes (A–C) of a long-range proton transfer step. We denote the H–O<sub>initial donor</sub> distance by  $d_1$  and the H–O<sub>initial acceptor</sub> distance by  $d_2$ . The special nomenclature using the terms O<sub>initial acceptor</sub> and O<sub>initial donor</sub> is only important for the understanding of the CDFs in Fig. 9.

by black ellipses in Fig. 9 (top, left). In the following, we give a brief description of the three sub-processes:

- Configuration A refers to a hydrogen bond. No proton rattling or long range proton transfer can be observed. This state is represented by the dashed lines in Fig. 8.

- Configuration B refers to a proton, which is oscillating between two oxygen atoms (proton rattling). Proton rattling is characterized by equal H–O<sub>initial donor</sub> and H–O<sub>initial acceptor</sub> distances (*cf.* area B in Fig. 9).

- If long range proton transfer is observed, some intensity can be found in region C of Fig. 9 (top, left). This area corresponds to a H–O<sub>initial donor</sub> distance greater than 200 pm and a H–O<sub>initial acceptor</sub> distance in the range of a covalent H–O bond. From such a distance tuple, the proton is transferred from the donor-oxygen to an acceptor-oxygen and moved afterward at least 100 pm.

In Fig. 9 the CDF of the HTPs (left) and LTPs (right) of CsH<sub>2</sub>PO<sub>4</sub>, CsHSeO<sub>4</sub> and CsHSO<sub>4</sub> are shown. The HTP and LTP of all compounds possess intensity in area A and B, while only the HTPs show intensity in area C. Thus, proton rattling can be reported for all compounds and phases but long range proton transfer takes only place in the HTPs.

For the comparison of the LTPs of the different compounds, it turns out that the “rattling” part of the CDF is almost symmetric for CsH<sub>2</sub>PO<sub>4</sub> whereas the H–O<sub>initial donor</sub> distances are only shorter or equal to the H–O<sub>initial acceptor</sub> distances (and not *vice versa*) for CsHSeO<sub>4</sub> and CsHSO<sub>4</sub>. Thus, proton rattling seems to be more extensive in CsH<sub>2</sub>PO<sub>4</sub> compared with CsHSeO<sub>4</sub> and CsHSO<sub>4</sub>. This finding can be explained by referring to the crystal structures (which we discussed in the Introduction). The LTP of CsH<sub>2</sub>PO<sub>4</sub> possesses two crystallographic different types of protons: the first type of protons forms localized asymmetric hydrogen bonds and the second type of protons undergoes rapid exchange between two oxygen atoms in a hydrogen bond with a symmetric double potential well. This fundamental difference arising from the hydrogen bond topology is able to explain the more frequent “proton rattling” within the LTP of CsH<sub>2</sub>PO<sub>4</sub>.

For the interpretation of our CDFs we have to keep in mind, that they represent only a non-converged snapshot of proton motion. Following this line, only qualitative (signal positions and signal shapes) instead of quantitative features (*i.e.* the intensity) should be derived from the CDF. Thus, the intensity of the peaks from area C can not be used to explain differences of diffusion coefficients in the HTPs.

Summarizing the results obtained from the CDFs, we can state that long range proton transfer could only be observed in the HTPs in agreement to their high proton conductivities. For the LTPs proton rattling can be stated from the CDFs, explaining

the similarities in the single H–O and O–O radial distance distribution functions. CDFs are a powerful tool to decide *via* the characteristic pattern of the peaks, if long range proton transfer takes place *i.e.* if the compound is a (potentially good) proton conductor.

### 3.3 Proton transfer rates

Proton conduction in solid acids follows a general Grotthuss mechanism.<sup>62,63</sup> Thus, the overall proton conductivity is a function of the proton jump rate between oxo-anions and the rotational relaxation of the oxygen tetrahedra.

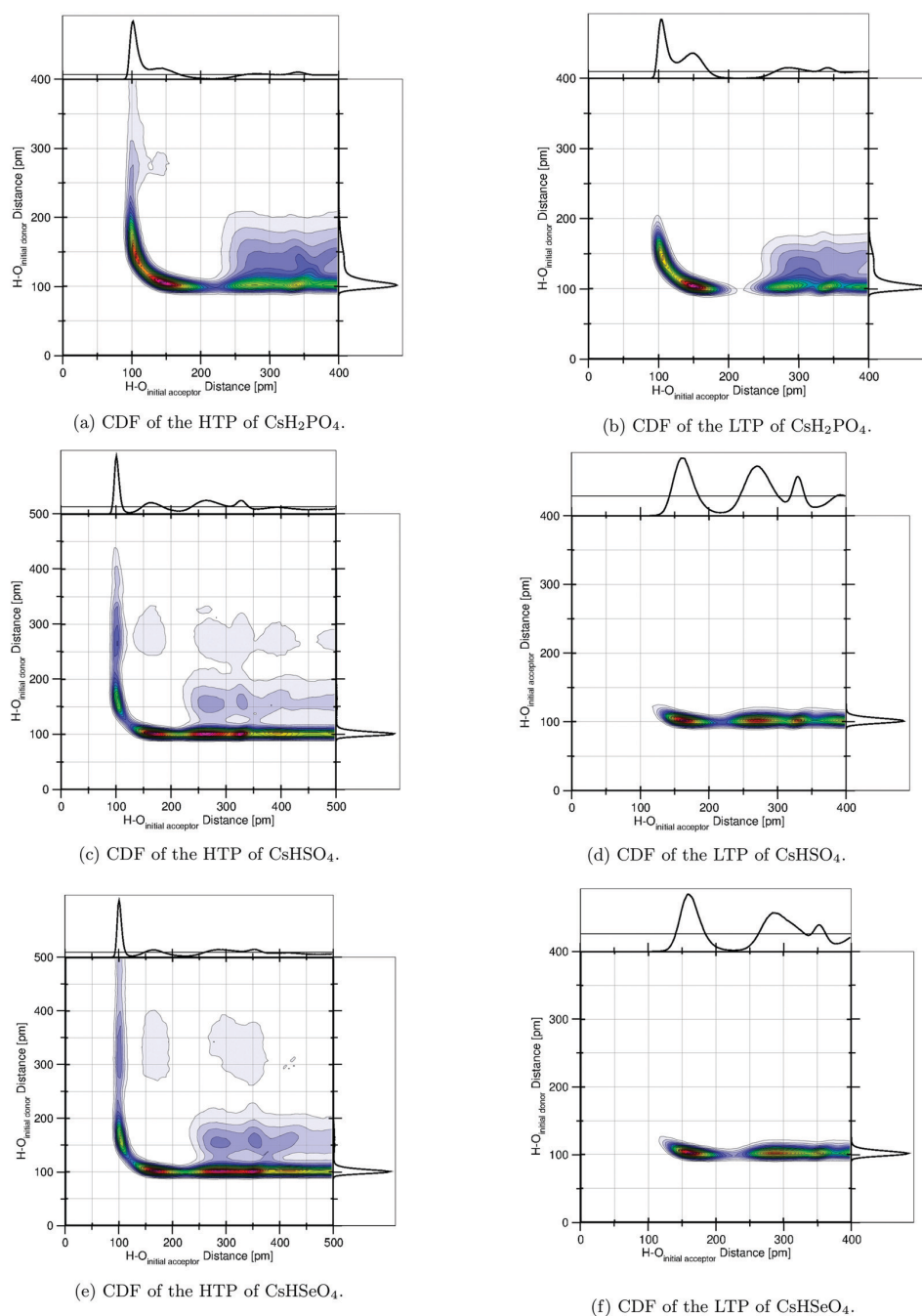
In order to explain the different proton conductivities, we start to investigate the frequency of proton jumps in the solid acids. The number of proton jumps per 1 ps and 1 proton for the solid acids CsH<sub>2</sub>PO<sub>4</sub>, CsHSO<sub>4</sub> and CsHSeO<sub>4</sub> are depicted in Table 3. We define a proton transfer as change of the next oxygen neighbor of a hydrogen atom (in two subsequent AIMD steps). Thus, also “proton rattling” between two oxygen atoms of a hydrogen bond does contribute to the number of proton transfers.

For CsH<sub>2</sub>PO<sub>4</sub>, the HTP and the LTP possess an extraordinary high proton transfer rate. The proton transfer rate of the HTP of CsHSeO<sub>4</sub> respectively CsHSO<sub>4</sub> is smaller by a factor of 7 and 15, respectively, to the proton transfer rate of CsH<sub>2</sub>PO<sub>4</sub>. The proton transfer rate of the LTPs of CsHSeO<sub>4</sub> and CsHSO<sub>4</sub> is even smaller by one magnitude compared with the transfer rate of the corresponding HTPs. The larger number of proton jumps in the LTP of CsH<sub>2</sub>PO<sub>4</sub> compared with the LTPs of the other compounds can be explained by a symmetric hydrogen bond with a double potential well (*cf.* description of crystal structures in the Introduction). Protons within this hydrogen bond undergo rapid exchange between two oxygen atoms. Hence, we conclude that the frequency of proton jumps as a solely parameter is not suited in order to describe the trend of the diffusion coefficients. Concerning only the proton transfer rate, CsH<sub>2</sub>PO<sub>4</sub> should possess the largest diffusion coefficient among all investigated solid acids and the conductivity of the LTP of CsH<sub>2</sub>PO<sub>4</sub> should be similar to the conductivity of the HTP. Both statements are wrong. The HTP of CsHSeO<sub>4</sub> exhibits the largest diffusion coefficient and the LTP of CsH<sub>2</sub>PO<sub>4</sub> does not show any proton conductivity at all.

### 3.4 Rotational dynamics of the anions

In order to investigate the rotational dynamics of the anions (HSO<sub>4</sub><sup>−</sup>, HSeO<sub>4</sub><sup>−</sup>, H<sub>2</sub>PO<sub>4</sub><sup>−</sup>) of the solid acids, we prepared vector autocorrelation functions (VCF) of the X–O (X = P, S, Se) vector. The X–O vector is defined as the difference between the position vector of the atom X and the position vector of one of its covalently bonded oxygen atoms. The autocorrelation of a vector  $\vec{a}_i$  is defined as the normalized sum over the dot product





**Fig. 9** Combined distribution function (CDF) of the HTPs and LTPs of CsH<sub>2</sub>PO<sub>4</sub>, CsHSeO<sub>4</sub> and CsHSO<sub>4</sub>. Explanations of the H–O<sub>initial acceptor</sub> distance and the H–O<sub>initial donor</sub> distance as well as the interpretation of the regions A–C are given in Fig. 8. We are only interested in relative intensities of the CDFs. For the sake of completeness, we show the color bars associated with the CDFs in the ESI.†

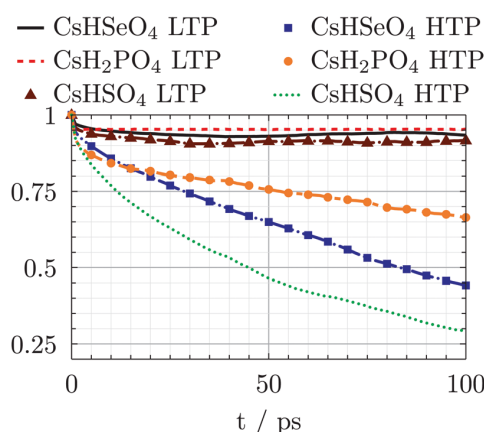
between the vector at some time  $t$  and the same vector at some later time  $t + \tau$  for all starting times  $t$ :<sup>95</sup>

$$\text{VCF}[\tau] = N \left\langle \sum_{i=0}^{T-\tau} \vec{a}_i(t) \cdot \vec{a}_i(t + \tau) \right\rangle_i \quad (2)$$

In eqn (2),  $T$  denotes the total simulation time,  $i$  the index of a specific X–O vector and  $N$  a normalization constant. The final autocorrelation function is obtained by taking the average over all X–O vectors within the system. The vector autocorrelation function starts at one and shows a decay with increasing  $t$ . The decay time of the autocorrelation function corresponds to the

**Table 3** Number of proton transfers per 1 ps and 1 proton

	CsH <sub>2</sub> PO <sub>4</sub>	CsHSeO <sub>4</sub>	CsHSO <sub>4</sub>
High temperature phases			
Proton jumps per ps and proton	11.6	1.61	0.75
Low temperature phases			
Proton jumps per ps and proton	12.0	0.20	0.09

**Fig. 10** Autocorrelation function of the P–O, S–O or Se–O vector.

rotational frequency of the anion. The resulting autocorrelation functions are shown in Fig. 10. The autocorrelation functions of all LTPs do not show a decay at all. This translates to the conclusion that anion rotation is not observed within the LTPs at the time scale of our AIMD simulations. In contrast, the autocorrelation functions of the HTPs possess a significant decay corresponding to the fast orientational dynamics of the oxygen tetrahedra in these compounds. The decay of the autocorrelation functions of the HTPs increases in the order CsH<sub>2</sub>PO<sub>4</sub>, CsHSeO<sub>4</sub>, CsHSO<sub>4</sub>.

## 4 General mechanism of proton conduction in the solid acid family

Combining the proton transfer rates (*cf.* Table 3) and rotational rates (*cf.* Fig. 10), we are now able to explain the qualitative trend of the diffusion coefficients (*cf.* Table 2).

The anions in the LTPs do not rotate at all. This prevents long range proton transfer (and hence proton conduction), because even if protons are transferred between the anions (*cf.* high proton transfer rate in the LTP of CsH<sub>2</sub>PO<sub>4</sub>) they will only “rattle” between two oxygen atoms. Opposed to that, the

HTP of CsH<sub>2</sub>PO<sub>4</sub> combines a very high proton transfer rate with a low rotational rate of the anion, whereas the HTP of CsHSO<sub>4</sub> combines a low proton transfer rate with a very high rotational rate of the anion. Both opposed combinations of rates lead to the same (intermediate) diffusion coefficient. The HTP of CsHSeO<sub>4</sub> possess neither the highest proton transfer rate (CsH<sub>2</sub>PO<sub>4</sub>) nor the highest rotational rate of the anion (CsHSO<sub>4</sub>). Nevertheless it is the best proton conductor among the three investigated compounds, because it combines an intermediate rotational rate of the anion and an intermediate proton transfer rate. The results of this discussion are summarized in Table 4. It turns out that we can explain the trend of the diffusion coefficient in terms of the proton transfer rates and rates of anion rotation. Each of these rates can limit the diffusion coefficient. Only a combination of a high proton transfer rate and a high rotational rate of the anion will lead to the optimal proton conductivity. This conclusion is in accordance with the overall chemical picture that proton conductivity in the solid acid family CsH<sub>y</sub>XO<sub>4</sub> (X = S, P, Se, y = 1, 2) follows a Grotthuss-type mechanism, which requires both proton transfer between different anions and subsequent relaxation of the environment (rotation of the anions).<sup>34,40,62,63</sup> A high proton transfer rate without relaxation of the environment only leads to proton rattling between two oxygen atoms and will not enable long range proton transfer.

In a last step, we want to propose a chemical explanation for the observed trend of the proton conductivity. As already indicated in this article, a strong as well as fluctuating hydrogen bond network is a prerequisite for a good proton conductor. The hydrogen bond acceptor strength increases in the series sulphate, selenate, phosphate.<sup>97</sup> Stronger hydrogen bonds in CsH<sub>2</sub>PO<sub>4</sub> compared to CsHSeO<sub>4</sub> and CsHSO<sub>4</sub> are also observed from the crystal structure (*cf.* Introduction) as well as the RDFs (*cf.* Fig. 7). For CsHSO<sub>4</sub> the proton jump frequency is the limiting factor, indicating too weak hydrogen bonds and too high activation energies for the proton jumps for efficient long range proton transfer. For CsH<sub>2</sub>PO<sub>4</sub> the frequency of anion reorientation limits the conductivity. While hydrogen bonds have to be broken in order to enable anion rotations, this indicates too strong hydrogen bonds in CsH<sub>2</sub>PO<sub>4</sub> for efficient long range proton transfer. The hydrogen bond strength of HSeO<sub>4</sub><sup>−</sup> seems to be in between the strength of HSO<sub>4</sub><sup>−</sup> and H<sub>2</sub>PO<sub>4</sub><sup>−</sup>, which balances best the opposing requirements for proton transfer between anions and anion rotation.

### 4.1 A quantitative predictor of the diffusion coefficient trend

**O–H-bond autocorrelation functions.** In this section, we want to develop a quantitative descriptor of proton dynamics which enables the explanation of the diffusion coefficients trends in the solid acids family. First, we state a criterion for

**Table 4** Qualitative explanation of diffusion coefficient trends

	CsH <sub>2</sub> PO <sub>4</sub> HTP	CsHSeO <sub>4</sub> HTP	CsHSO <sub>4</sub> HTP	CsH <sub>2</sub> PO <sub>4</sub> LTP	CsHSeO <sub>4</sub> LTP	CsHSO <sub>4</sub> LTP
Proton transfer rate	Very high	Medium	Low	Very high	Very low	Very low
Anion rotation frequency	Low	Medium	Very high	No rotation	No rotation	No rotation
Diffusion coefficient	Medium	Very high	Medium	No diffusion	No diffusion	No diffusion

the formation of a covalent bond between hydrogen and oxygen atoms: we assume a covalent bond is formed between a hydrogen atom and its next oxygen neighbor. Starting from this definition, we obtain the autocorrelation function of the covalent HO-bonds from an AIMD trajectory according to the following protocol: at the initial time step ( $t = 0$ ), we determine the next oxygen neighbor of every hydrogen atom and assume a covalent bond between these pairs of atoms. At time  $t \neq 0$ , an initial covalent OH-bond remained intact if the next oxygen neighbor of a hydrogen atom did not change (compared with  $t = 0$ ). Each intact initial covalent bond at time  $t$  contributes a value of one to the autocorrelation function of the covalent HO-bonds, while every hydrogen atom with a new next oxygen neighbor contribute a value of zero. Finally, we divide the number of intact covalent H-O bonds by the overall number of initial O-H covalent bonds in order to obtain a normalized function. Thus the covalent HO-autocorrelation function describes the fraction of protons which are still bonded to their initial covalent bond partner. From a pictorial/qualitative point of view it is apparent that the decay of the covalent HO-autocorrelation function and the diffusivity are related.

The autocorrelation functions of the covalent HO-bond autocorrelation function for the solid acids are shown in Fig. 11. We performed a bi-exponential fit of the covalent HO-bond autocorrelation function according to:

$$f(t) = c_1 \cdot e^{-k_1 \cdot t} + (1 - c_1) \cdot e^{-k_2 \cdot t} \quad (3)$$

We choose  $k_1 > k_2$  and define  $\tau_1 = 1/k_1$  as short-term and  $\tau_2 = 1/k_2$  as long-term relaxation time. The relaxation times for the investigated compounds are listed in Table 5. We have verified that a bi-exponential target function from eqn (3) yields a good fit for our simulation data, indicating that the underlying physical process can indeed be considered a two-step mechanism.

We will now demonstrate two different correlations concerning the short- as well as long-term relaxation times of the covalent HO-bond autocorrelation function.

First, we will correlate the long-term relaxation times and proton conductivity: it turns out (from Fig. 11) that covalent HO-bond autocorrelation functions of all LTPs remain almost

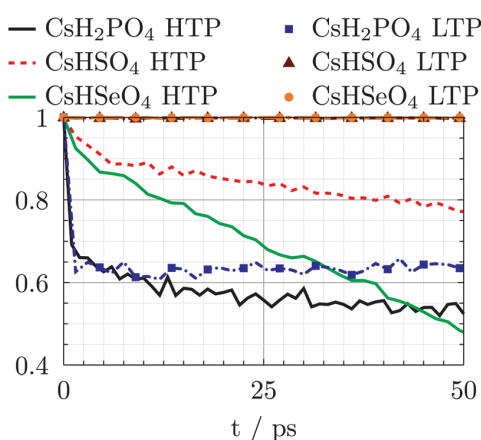


Fig. 11 Autocorrelation function of the covalent O-H bonds.

Table 5 Short-term relaxation time  $\tau_1$  and long-term relaxation time  $\tau_2$  of the covalent HO-bond autocorrelation function obtained from eqn (3) with  $k_2 < k_1$  and  $\tau_i = 1/k_i$ . The infinity symbol means that the covalent HO-bond autocorrelation function is virtually constant. The LTP of  $\text{CsH}_2\text{PO}_4$  is a special case. Its covalent HO-bond autocorrelation function possesses a fast decay at the beginning ( $\tau_1 < \infty$ ) and remains afterwards constant ( $\tau_2 = \infty$ ). Therefore, we performed for this system a mono-exponential fit

	$\text{CsH}_2\text{PO}_4$	$\text{CsHSeO}_4$	$\text{CsHSO}_4$
High temperature phases			
Relaxation time $\tau_1$ in ps	0.31	0.57	2.5
Relaxation time $\tau_2$ in ps	248	82	322
Low temperature phases			
Relaxation time $\tau_1$ in ps	0.19	$\infty$	$\infty$
Relaxation time $\tau_2$ in ps	$\infty$	$\infty$	$\infty$

constant (beside a very short and fast decay of the HO-bond autocorrelation function of the LTP of  $\text{CsH}_2\text{PO}_4$ ). Thus the number of intact initial covalent O-H bonds is constant and no long range proton transfer among different oxygens atoms is observed. An infinite long-term relaxation time of the covalent HO-bond autocorrelation functions (*cf.* Table 5) fits well to the negligible proton conductivity of the LTPs. The long-term relaxation times of the HTPs of  $\text{CsH}_2\text{PO}_4$  and  $\text{CsHSO}_4$  are almost equal. Opposed to that, the long-term relaxation times of the HTP of  $\text{CsHSeO}_4$  is drastically decreased compared with the long-term relaxation times of the HTPs of  $\text{CsH}_2\text{PO}_4$  and  $\text{CsHSO}_4$ . Keeping in mind, that the difference of the diffusion coefficients of  $\text{CsH}_2\text{PO}_4$  and  $\text{CsHSO}_4$  was within the error of the AIMD simulations, the long-term relaxation time of the covalent HO-bond autocorrelation function is able to predict the quantitative trend of the diffusion coefficients (compare Tables 2 and 5).

In a second step, we want to correlate the short-term relaxation times and the proton transfer rates from Table 3: compared to the LTPs of  $\text{CsHSeO}_4$  and  $\text{CsHSO}_4$  only the LTP of  $\text{CsH}_2\text{PO}_4$  possesses a short and fast decay from 1 to 0.64. This different behavior can be explained from the crystal structure (*cf.* crystal structure discussion in the Introduction): two types of hydrogen bonds are observed within the LTP of  $\text{CsH}_2\text{PO}_4$ . While the first type of protons within an asymmetric hydrogen bond do not change their next oxygen neighbor, the other protons “rattle” extensively within the double minimum potential well of the symmetric hydrogen bond.<sup>36</sup>

We reported the highest proton jump rates for the HTP and LTP of  $\text{CsH}_2\text{PO}_4$  which fits well to the smallest short-term relaxations times ( $< 0.5$  ps). Also the reduced number of jumps within the series  $\text{CsH}_2\text{PO}_4$ ,  $\text{CsHSeO}_4$  and  $\text{CsHSO}_4$  is reflected by the increasing trend of the short-term relaxation times for these compounds ( $0.31 \text{ ps} < 0.57 \text{ ps} < 2.5 \text{ ps}$ ).

In conclusion, the covalent HO-bond autocorrelation function combines the information of the proton transfer rate and the overall proton conductivity into a single function. For short timescales, the number of initially formed and still intact covalent O-H bonds is dominated by the rattling/hopping frequency of the protons, while the number of remaining covalent bonded H-O pairs for longer timescales is connected to the overall proton conductivity. For solid acids, we have

demonstrated that the anion reorientation dynamics is the missing link between the proton transfer frequency and the proton conductivity, and hence can be indirectly deduced from the covalent HO-bond autocorrelation function. Compared to the diffusion coefficient, the covalent HO-bond autocorrelation function as a single descriptor allows a direct atomistic interpretation and its relation to the proton transfer rate and anion rotation rate is immediately apparent for the case of the solid acids.

## 5 A fundamental invariant of proton transfer

The overall aim of an AIMD simulation is to gain atomistic insights into the structure and the dynamics of an investigated system. This information can be used to construct models of atomistic processes and the comparison (of these models) to existing experimental data. On this occasion, we can use the predictive power of *ab initio* simulations not only for evaluation but also for an explicit statement of an invariant of proton transfer in  $O \cdots H-O$  hydrogen bonded systems, which is relevant for the question how materials science can benefit from theoretical investigations.

In our group, proton transfer is an object of interest since many years and beside the inorganic solid acid family in this article, we have also investigated proton transfer in an organic polymer (poly(vinylphosphonic acid) (PVPA)) and columnar discotic compound (hexa(*p*-phosphonatophenyl)benzene (HPB)).<sup>98–100</sup> We have performed a careful analysis of the trajectories of these compounds ( $CsHSO_4$ ,  $CsHSeO_4$ ,  $CsH_2PO_4$ , HPB, PVPA), in order to discover similarities and differences of proton transfer. In Fig. 12, we present the histogram of the proton jumps with respect to the oxygen–oxygen distance within an AIMD trajectory (normalized to 100 ps and 16 protons for reasons of comparability). We define a proton jump as change of the next oxygen neighbor of a hydrogen atom (in two subsequent AIMD steps).

The histogram illustrates that all systems exhibit the same shape of the dependence of the jump number on the O–O

distance: in the logarithmic plot, the function resembles a negative parabola, which corresponds to a Gaussian dependence of the jump function on O–O distance. This Gaussian function has the same center ( $d_{OO} \approx 2.4 \text{ \AA}$ ) for all compounds. However, the jump functions vary considerably in absolute amplitude: every system has a specific prefactor (corresponding to a vertical shift of the function in the logarithmic plot).

This similarity between the systems is quite notable, especially when considering that the jump function is a product of the probability for a given O–O distance (*i.e.* the radial distribution function, see Fig. 7) and the conditional probability for a jump at that given O–O distance. None of these contributions, however, has a real Gaussian shape, and the radial distribution functions clearly have different centers (see Fig. 7).

In conclusion, we obtain a common Gaussian shape of the actual number of proton jumps as function of the acceptor–donor distance, with a common center at  $2.4 \text{ \AA}$  and a system-specific (but distance-independent) prefactor which depends on the exact chemistry of the constituents and can only be reported numerically.

Thus, we conclude from Fig. 12 that a necessary but not sufficient prerequisite of efficient proton transfer is the high probability of O–O distances of about  $2.4 \text{ \AA}$ .

## 6 Conclusion

In this article, we investigated systematically proton dynamics in the low and high temperature phases of  $CsHSeO_4$ ,  $CsHSO_4$  and  $CsH_2PO_4$  by AIMD simulations. While the high temperature phases are excellent proton conductors, the low temperature phases do not conduct protons at all. We demonstrate that the drastic reduced conductivity of the low temperature phases is caused by a highly ordered, rigid hydrogen bond network, while efficient long range proton transfer in the high temperature phases is enabled by the interplay of high proton transfer rates and frequent anion reorientation. We have shown that the relaxation time of the OH-bond autocorrelation function is a simple, one-dimensional quantitative descriptor for the prediction of the diffusion coefficient trend within the solid acids family.

Finally, we reported a fundamental invariant of proton transfer in  $O \cdots H-O$  hydrogen bonded systems obtained by analysis of the trajectories of several proton conducting materials, namely the acceptor–donor distance at which the proton hopping rate is maximum. From this invariant, we deduced a design rule for the development of new proton conductors: a prerequisite of efficient proton transfer is the presence of O–O distances of about  $2.4 \text{ \AA}$ .

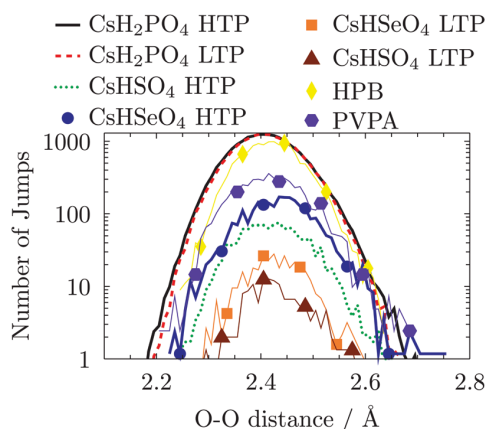


Fig. 12 Number of proton jumps with respect to the O–O distance observed in a 100 ps simulation of a system containing 16 protons.

## Conflicts of interest

There are no conflicts to declare.

## Acknowledgements

This work was supported by the Deutsche Forschungsgemeinschaft (Project-ID 435886714).



## References

- J. Cheng, M. Sulpizi and M. Sprik, Redox potentials and  $pK_a$  for benzoquinone from density functional theory based molecular dynamics, *J. Chem. Phys.*, 2009, **131**(15), 154504.
- J. Cheng, X. Liu, J. VandeVondele, M. Sulpizi and M. Sprik, Redox potentials and acidity constants from density functional theory based molecular dynamics, *Acc. Chem. Res.*, 2014, **47**(12), 3522–3529. PMID: 25365148.
- J. Cheng and M. Sprik, Acidity of the aqueous rutile  $TiO_2(110)$  surface from density functional theory based molecular dynamics, *J. Chem. Theory Comput.*, 2010, **6**(3), 880–889. PMID: 26613315.
- C. Zhang, J. Hutter and M. Sprik, Coupling of surface chemistry and electric double layer at  $TiO_2$  electrochemical interfaces, *J. Phys. Chem. Lett.*, 2019, **10**(14), 3871–3876. PMID: 31241948.
- I. L. Geada, I. Petit, M. Sulpizi and F. Tielens, Unravelling the Gly-Pro-Glu tripeptide induced reconstruction of the Au(110) surface at the molecular scale, *Surf. Sci.*, 2018, **677**, 271–277.
- M. Gierada, F. De Proft, M. Sulpizi and F. Tielens, Understanding the acidic properties of the amorphous hydroxylated silica surface, *J. Phys. Chem. C*, 2019, **123**(28), 17343–17352.
- M.-P. Gageot, M. Sprik and M. Sulpizi, Oxide/water interfaces: how the surface chemistry modifies interfacial water properties, *J. Phys.: Condens. Matter*, 2012, **24**(12), 124106.
- M. Mangold, L. Rolland, F. Costanzo, M. Sprik, M. Sulpizi and J. Blumberger, Absolute  $pK_a$  values and solvation structure of amino acids from density functional based molecular dynamics simulation, *J. Chem. Theory Comput.*, 2011, **7**(6), 1951–1961. PMID: 26596456.
- M. Hellström, V. Quaranta and J. Behler, One-dimensional vs. two-dimensional proton transport processes at solid-liquid zinc-oxide-water interfaces, *Chem. Sci.*, 2019, **10**, 1232–1243.
- M. Hellström and J. Behler, Concentration-dependent proton transfer mechanisms in aqueous NaOH solutions: From acceptor-driven to donor-driven and back, *J. Phys. Chem. Lett.*, 2016, **7**(17), 3302–3306. PMID: 27504986.
- V. Quaranta, M. Hellström and J. Behler, Proton-transfer mechanisms at the water-ZnO interface: The role of pre-solvation, *J. Phys. Chem. Lett.*, 2017, **8**(7), 1476–1483. PMID: 28296415.
- M. Hellström and J. Behler, Proton-transfer-driven water exchange mechanism in the  $Na^+$  solvation shell, *J. Phys. Chem. B*, 2017, **121**(16), 4184–4190. PMID: 28375608.
- M. Brehm, H. Weber, A. S. Pensado, A. Stark and B. Kirchner, Proton transfer and polarity changes in ionic liquid-water mixtures: a perspective on hydrogen bonds from ab initio molecular dynamics at the example of 1-ethyl-3-methylimidazolium acetate-water mixtures-part 1, *Phys. Chem. Chem. Phys.*, 2012, **14**, 5030–5044.
- A. Stirling, M. Bernasconi and M. Parrinello, Ab initio simulation of water interaction with the (100) surface of pyrite, *J. Chem. Phys.*, 2003, **118**(19), 8917–8926.
- A. Stirling, T. Rozgonyi, M. Krack and M. Bernasconi, Prebiotic  $NH_3$  formation: Insights from simulations, *Inorg. Chem.*, 2016, **55**(4), 1934–1939. PMID: 26831570.
- J. Ingenmey, S. Gehrke and B. Kirchner, How to harvest Grotthuss diffusion in protic ionic liquid electrolyte systems, *ChemSusChem*, 2018, **11**(12), 1900–1910.
- E. Perlt, P. Ray, A. Hansen, F. Malberg, S. Grimme and B. Kirchner, Finding the best density functional approximation to describe interaction energies and structures of ionic liquids in molecular dynamics studies, *J. Chem. Phys.*, 2018, **148**(19), 193835.
- B. Kirchner, F. Malberg, D. S. Firaha and O. Hollóczki, Ion pairing in ionic liquids, *J. Phys.: Condens. Matter*, 2015, **27**(46), 463002.
- G. Y. Foran, D. H. Brouwer and G. R. Goward, Quantifying site-specific proton dynamics in phosphate solid acids by  $^1H$  double quantum NMR spectroscopy, *J. Phys. Chem. C*, 2017, **121**(46), 25641–25650.
- A. Goñ-Urtiaga, D. Presvytes and K. Scott, Solid acids as electrolyte materials for proton exchange membrane (PEM) electrolysis: Review, *Int. J. Hydrogen Energy*, 2012, **37**(4), 3358–3372. International Conference on Renewable Energy (ICRE 2011).
- V. G. Ponomareva, K. A. Kovalenko, A. P. Chupakhin, E. S. Shutova and V. P. Fedin,  $CsHSO_4$  – proton conduction in a crystalline metal-organic framework, *Solid State Ionics*, 2012, **225**, 420–423. Solid State Ionics 18 Proceedings of the 18th International Conference on Solid State Ionics Warsaw, Poland, July 3–8, 2011.
- A. B. Papandrew, R. W. Atkinson III, R. R. Unocic and A. Thomas, Zawodzinski. Ruthenium as a co-tolerant hydrogen oxidation catalyst for solid acid fuel cells, *J. Mater. Chem. A*, 2015, **3**, 3984–3987.
- L. Liu, H. Li, X. Chen and X. Lei, Electrolyte membranes based on molten  $KH_5(PO_4)_2$  for intermediate temperature fuel cells, *Fuel Cells*, 2019, **19**(3), 280–288.
- J. H. Leal, H. Martinez, I. Martinez, A. D. Price, A. G. Goos and C. E. Botez, Stability of the superprotonic conduction of  $(1-x)CsH_2PO_4/xSiO_2$  ( $0 \leq x \leq 0.3$ ) composites under dry and humid environments, *Mater. Today Commun.*, 2018, **15**, 11–17.
- I. N. Bagryantseva, V. G. Ponomareva and N. P. Lazareva, Proton-conductive membranes based on  $CsH_2PO_4$  and ultra-dispersed polytetrafluoroethylene, *Solid State Ionics*, 2019, **329**, 61–66.
- H. Takahashi, Y. Suzuki and T. Sakuma, New phase transition in superprotonic phase of inorganic solid acid  $Cs_2(HSO_4)(H_2PO_4)$ , *Solid State Ionics*, 2016, **285**, 155–159. The 40th Symposium on Solid State Ionics in Japan.
- N. Mohammad, A. B. Mohamad, A. A. H. Kadhum and K. S. Loh, A review on synthesis and characterization of solid acid materials for fuel cell applications, *J. Power Sources*, 2016, **322**, 77–92.
- F. P. Lohmann-Richters, B. Abel and Á. Varga, In situ determination of the electrochemically active platinum

- surface area: key to improvement of solid acid fuel cells, *J. Mater. Chem. A*, 2018, **6**, 2700–2707.
- 29 Q. Li, R. He, J. O. Jensen and N. J. Bjerrum, Approaches and recent development of polymer electrolyte membranes for fuel cells operating above 100°, *Chem. Mater.*, 2003, **15**(26), 4896–4915.
- 30 S. M. Haile, D. A. Boysen, C. R. Chisholm and R. B. Merle, Solid acids as fuel cell electrolytes, *Nature*, 2001, **410**, 910–913.
- 31 D. A. Boysen, T. Uda, C. R. I. Chisholm and S. M. Haile, High-performance solid acid fuel cells through humidity stabilization, *Science*, 2004, **303**, 68–70.
- 32 C. R. I. Chisholm, D. A. Boysen, A. B. Papandrew, S. Zecevic, S. Y. Cha, K. A. Sasaki, A. Varga, K. P. Giapis and S. M. Haile, From laboratory breakthrough to technological realization: The development path for solid acid fuel cells, *J. Electrochem. Soc. Interface*, 2009, **18**, 53–59.
- 33 <http://www.safcell.com/>.
- 34 A. I. Baranov, Crystals with disordered hydrogen-bond networks and superprotonic conductivity. review, *Crystallogr. Rep.*, 2003, **48**, 1012–1037.
- 35 A. Preisinger, K. Mereiter and W. Bronowska. The phase transition of CsH<sub>2</sub>PO<sub>4</sub> (CDP) at 505 K. *European Powder Diffraction 3, volume 166 of Materials Science Forum*. Trans Tech Publications Ltd, 1994, vol. 7, pp. 511–516.
- 36 G. Kim, J. M. Griffin, F. Blanc, S. M. Haile and C. P. Grey, Characterization of the dynamics in the protonic conductor CsH<sub>2</sub>PO<sub>4</sub> by <sup>17</sup>O solid-state NMR spectroscopy and first-principles calculations: Correlating phosphate and protonic motion, *J. Am. Chem. Soc.*, 2015, **137**(11), 3867–3876. PMID: 25732257.
- 37 G. Kim, F. Blanc, Y.-Y. Hu and C. P. Grey, Understanding the conduction mechanism of the protonic conductor CsH<sub>2</sub>PO<sub>4</sub> by solid-state NMR spectroscopy, *J. Phys. Chem. C*, 2013, **117**(13), 6504–6515.
- 38 P. Colomban and C. Philippe, *Proton Conductors: Solids, membranes and gels-materials and devices*, Cambridge University Press, 1992, vol. 2.
- 39 K. Yamada, T. Sagara, Y. Yamane, H. Ohki and T. Okuda, Superprotonic conductor CsH<sub>2</sub>PO<sub>4</sub> studied by <sup>1</sup>H, <sup>31</sup>P NMR and X-ray diffraction, *Solid State Ionics*, 2004, **175**(1), 557–562. Fourteenth International Conference on Solid State Ionics.
- 40 S. M. Haile, C. R. I. Chisholm, K. Sasaki, D. A. Boysen and T. Uda, Solid acid proton conductors: from laboratory curiosities to fuel cell electrolytes, *Faraday Discuss.*, 2007, **134**, 17–39.
- 41 A. M. Balagurov, A. V. Belushkin, I. D. Dutt, I. Natkaniec, N. M. Plakida, B. N. Savenko, L. A. Shuvalov and J. Wasicki, Neutron scattering studies on structural phase transitions of superionic conductor CsHSO<sub>4</sub>, *Ferroelectrics*, 1985, **63**(1), 59–67.
- 42 A. V. Belushkin, I. Natkaniec, N. M. Pakida, L. A. Shuvalov and J. Wasicki, Neutron scattering studies of vibrational spectra and structural transformations in the superionic conductors CsHSO<sub>4</sub> and CsHSeO<sub>4</sub>, *J. Phys. C: Solid State Phys.*, 1987, **20**(5), 671–687.
- 43 P. Colomban, M. Pham-Thi and A. Novak, Thermal history and phase transitions in the superionic protonic conductors CsHSO<sub>4</sub> and CsHSeO<sub>4</sub>, *Solid State Ionics*, 1986, **20**(2), 125–134.
- 44 Z. Jiráak, M. Dlouhá, S. Vratislav, A. M. Balagurov, A. I. Beskrovnyi, V. I. Gordelii, I. D. Datt and L. A. Shwalov, A neutron diffraction study of the superionic phase in CsHSO<sub>4</sub>, *Phys. Status Solidi A*, 1987, **100**(2), K117–K122.
- 45 B. V. Merinov, A. I. Baranov, L. A. Shuvalov and B. A. Maksimov, Crystal structure of superionic phase CsDSO<sub>4</sub> and phase transitions in cesium hydro- and deuteriosulfates (selenates), *Kristallografiya*, 1987, **32**(1), 86–92.
- 46 A. V. Belushkin, W. I. F. David, R. M. Ibberson and L. A. Shuvalov, High-resolution neutron powder diffraction studies of the structure of CsDSO<sub>4</sub>, *Acta Crystallogr., Sect. B: Struct. Sci.*, 1991, **47**(2), 161–166.
- 47 J.-H. Park, C.-S. Kim, C. Byung-Chun, B.-K. Moon and H.-J. Seo, Physical properties of CsH<sub>2</sub>PO<sub>4</sub> crystal at high temperatures, *J. Phys. Soc. Jpn.*, 2003, **72**(6), 1592–1593.
- 48 E. Ortiz, R. A. Vargas and B.-E. Mellander, On the high-temperature phase transitions of some KDP-family compounds: a structural phase transition? a transition to a bulk-high proton conducting phase?, *Solid State Ionics*, 1999, **125**(1), 177–185.
- 49 D. A. Boysen, S. M. Haile, H. Liu and R. A. Secco, High-temperature behavior of CsH<sub>2</sub>PO<sub>4</sub> under both ambient and high pressure conditions, *Chem. Mater.*, 2003, **15**, 727–736.
- 50 K. Yamada, T. Sagara, Y. Yamane, H. Ohki and T. Okuda, Superprotonic conductor CsH<sub>2</sub>PO<sub>4</sub> studied by <sup>1</sup>H, <sup>31</sup>P NMR and X-ray diffraction, *Solid State Ionics*, 2004, **175**(1), 557–562. Fourteenth International Conference on Solid State Ionics.
- 51 J. Otomo, T. Tamaki, S. Nishida, S. Wang, M. Ogura, T. Kobayashi, C.-j. Wen, H. Nagamoto and H. Takahashi, Effect of water vapor on proton conduction of cesium dihydrogen phosphate and application to intermediate temperature fuel cells, *J. Appl. Electrochem.*, 2005, **35**(9), 865–870.
- 52 A. I. Baranov, V. V. Grebenev, A. N. Khodan, V. V. Dolbinina and E. P. Efremova, Optimization of superprotonic acid salts for fuel cell applications, *Solid State Ionics*, 2005, **176**(39), 2871–2874. Selected Papers from the 12th International Conference on Solid State Proton Conductors (SSPC-12).
- 53 J. Otomo, N. Minagawa, C. j. Wen, K. Eguchi and H. Takahashi, Protonic conduction of CsH<sub>2</sub>PO<sub>4</sub> and its composite with silica in dry and humid atmospheres, *Solid State Ionics*, 2003, **156**(3), 357–369.
- 54 B. V. Troeye, M. J. van Setten, M. Giantomassi, M. Torrent, G.-M. Rignanese and X. Gonze, First-principles study of paraelectric and ferroelectric CsH<sub>2</sub>PO<sub>4</sub> including dispersion forces: Stability and related vibrational, dielectric, and elastic properties, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2017, **95**, 024112.
- 55 R. Blinc, J. Dolinsek, G. Lahajnar, I. Zupancic, L. A. Shuvalov and A. I. Baranov, Spin-lattice relaxation and self-diffusion study of the protonic superionic conductors CsHSeO<sub>4</sub> and CsHSO<sub>4</sub>, *Phys. Status Solidi B*, 1984, **123**(1), K83–K87.

- 56 A. I. Baranov, V. P. Khiznichenko and L. A. Shuvalov, High temperature phase transitions and proton conductivity in some KDP-family crystals, *Ferroelectrics*, 1989, **100**, 135–141.
- 57 J. Badot, RF-microwave dielectric relaxations and phase transitions in superionic protonic acid sulphates (selenates), *Solid State Ionics*, 1989, **35**, 143–149.
- 58 R. Blinc, J. Dolinsek, G. Lahajnar, I. Zupancic, L. A. Shuvalov and A. I. Baranov, Spin-lattice relaxation and self-diffusion study of the protonic superionic conductors CsHSeO<sub>4</sub> and CsHSO<sub>4</sub>, *Phys. Status Solidi B*, 1984, **123**, K83–K87.
- 59 A. Ishikawa, H. Maekawa, T. Yamamura, Y. Kawakita, K. Shibata and M. Kawai, Proton dynamics of CsH<sub>2</sub>PO<sub>4</sub> studied by quasi-elastic neutron scattering and PFG-NMR, *Solid State Ionics*, 2008, **179**, 2345–2349.
- 60 A. V. Belushkin, C. J. Carlile and L. A. Shuvalov, The diffusion of protons in the superionic conductor CsHSO<sub>4</sub> by quasielastic neutron scattering, *J. Phys.: Condens. Matter*, 1992, **4**, 389–398.
- 61 G. Kim, F. Blanc, Y.-Y. Hu and C. P. Grey, Understanding the conduction mechanism of the protonic conductor CsH<sub>2</sub>PO<sub>4</sub> by solid-state NMR spectroscopy, *J. Phys. Chem. C*, 2013, **117**, 6504–6515.
- 62 A. I. Baranov, B. V. Merinov, A. V. Tregubchenko, V. P. Khiznichenko, L. A. Shuvalov and N. M. Schagina, Fast proton transport in crystals with a dynamically disordered hydrogen bond network, *Solid State Ionics*, 1989, **36**(3), 279–282.
- 63 K. D. Kreuer, Fast proton conductivity: A phenomenon between the solid and the liquid state?, *Solid State Ionics*, 1997, **94**(1), 55–62. Papers from the International Workshop.
- 64 H.-S. Lee and M. E. Tuckerman, The structure and proton transport mechanisms in the superprotonic phase of CsH<sub>2</sub>PO<sub>4</sub>: An *ab initio* molecular dynamics study, *J. Phys. Chem. C*, 2008, **112**, 9917–9930.
- 65 C. Drefßler, G. Kabbe and D. Sebastiani, Proton conductivity in hydrogen phosphate/sulfates from a coupled molecular dynamics/lattice Monte Carlo (CMD/LMC) approach, *J. Phys. Chem. C*, 2016, **120**(36), 19913–19922.
- 66 B. C. Wood and N. Marzari, Proton dynamics in superprotonic CsHSO<sub>4</sub>, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2007, **76**, 134301.
- 67 W. Munch, K. D. Kreuer, U. Traub and J. Maier, Proton transfer in the three-dimensional hydrogen bond network of the high temperature phase of CsHSO<sub>4</sub>: a molecular dynamics study, *J. Mol. Struct.*, 1996, **381**, 1–8.
- 68 W. Munch and K. D. Kreuer, A molecular dynamics study of the high proton conducting phase of CsHSO<sub>4</sub>, *Solid State Ionics*, 1995, **77**, 10–14.
- 69 C. R. I. Chisholm, Y. H. Jang, S. M. Haile and W. A. Goddard, Superprotonic phase transition of CsHSO<sub>4</sub>: A molecular dynamics simulation study, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2005, **72**, 134103.
- 70 P. Zetterström, A. V. Belushkin, R. L. McGreevy and L. A. Shuvalov, Structure and proton conduction in CsDSO<sub>4</sub>, *Solid State Ionics*, 1999, **116**(3), 321–329.
- 71 A. Damyanovich, M. M. Pintar, R. Blinc and J. Slak, Proton pseudoglass-to-fast-ion-conductor phase transition in CsHSO<sub>4</sub>, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1997, **56**, 7942–7946.
- 72 J. C. Badot and P. Colomban, RF-microwave dielectric relaxations and phase transitions in superionic protonic acid sulphates (selenates), *Solid State Ionics*, 1989, **35**(1), 143–149.
- 73 B. Merinov, Proton transport mechanism and pathways in the superprotonic phase of CsHSO<sub>4</sub> from experiment and theory, *Solid State Ionics*, 2012, **213**, 72–75.
- 74 X. Ke and I. Tanaka, Atomistic mechanism of proton conduction in solid CsHSO<sub>4</sub> by a first-principles study, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2004, **69**, 165114.
- 75 X. Ke and I. Tanaka, Proton transfer mechanism in solid CsHSO<sub>4</sub> by first-principles study, *Solid State Ionics*, 2004, **172**, 145–148.
- 76 C. Drefßler, G. Kabbe and D. Sebastiani, Insight from atomistic simulations of protonation dynamics at the nanoscale, *Fuel Cells*, 2016, **16**(6), 682–694.
- 77 X. Ke and I. Tanaka, Atomistic mechanism of proton conduction in solid CsHSO<sub>4</sub> by a first-principles study, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2004, **69**, 165114.
- 78 G. Kabbe, C. Drefßler and D. Sebastiani, Toward realistic transfer rates within the coupled molecular dynamics/lattice Monte Carlo approach, *J. Phys. Chem. C*, 2016, **120**(36), 19905–19912.
- 79 J. W. Traer, K. J. Soo, M. Vijayakumar and G. R. Goward, Elucidating the time scale and geometry of phosphate and phosphonate rotation in solid acid electrolytes using multinuclear NMR, *J. Phys. Chem. C*, 2011, **115**(13), 6064–6072.
- 80 J. Hutter, M. Iannuzzi, F. Schiffmann and J. VandeVondele, CP2K: atomistic simulations of condensed matter systems, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2014, **4**, 15–25.
- 81 J. VandeVondele, M. Krack, F. Mohamed, M. Parrinello, T. Chassaing and J. Hutter, Quickstep: Fast and accurate density functional calculations using a mixed Gaussian and plane waves approach, *Comput. Phys. Commun.*, 2005, **167**, 103–128.
- 82 J. VandeVondele and J. Hutter, An efficient orbital transformation method for electronic structure calculations, *J. Chem. Phys.*, 2003, **118**, 4365–4369.
- 83 Y. Zhang and W. Yang, Comment on “generalized gradient approximation made simple”, *Phys. Rev. Lett.*, 1998, **80**, 890.
- 84 J. P. Perdew, A. Ruzsinszky, G. I. Csonka, O. A. Vydrov, G. E. Scuseria, L. A. Constantin, X. Zhou and K. Burke, Restoring the density-gradient expansion for exchange in solids and surfaces, *Phys. Rev. Lett.*, 2008, **100**, 136406.
- 85 J. P. Perdew, K. Burke and M. Ernzerhof, Generalized gradient approximation made simple, *Phys. Rev. Lett.*, 1996, **77**, 3865–3868.
- 86 J. VandeVondele and J. Hutter, Gaussian basis sets for accurate calculations on molecular systems in gas and condensed phases, *J. Chem. Phys.*, 2007, **127**, 114105.
- 87 C. Hartwigsen, S. Goedecker and J. Hutter, Relativistic separable dual-space Gaussian pseudopotentials from H to Rn, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1998, **58**, 3641–3662.



- 88 M. Krack, Pseudopotentials for H to Kr optimized for gradient-corrected exchange-correlation functionals, *Theor. Chem. Acc.*, 2005, **114**, 145–152.
- 89 S. Grimme, J. Antony, S. Ehrlich and H. Krieg, A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu, *J. Chem. Phys.*, 2010, **132**, 154104.
- 90 S. Nosé, A unified formulation of the constant temperature molecular dynamics methods, *J. Chem. Phys.*, 1984, **81**, 511–519.
- 91 G. J. Martyna, M. L. Klein and M. Tuckerman, Nosé-Hoover chains: The canonical ensemble via continuous dynamics, *J. Chem. Phys.*, 1992, **97**, 2635–2643.
- 92 S. Nosé, A molecular dynamics method for simulations in the canonical ensemble, *Mol. Phys.*, 1970, **52**, 255–268.
- 93 C. R. I. Chisholm and S. M. Haile, X-ray structure refinement of CsHSO<sub>4</sub> in phase ii, *Mater. Res. Bull.*, 2000, **35**(7), 999–1005.
- 94 M. A. Zakharov, S. I. Troyanov and E. Kemnitz, Superprotonic high temperature phase and refinement of the low temperature structure of CsHSeO<sub>4</sub>, *Z. Kristallogr. - Cryst. Mater.*, 2001, **216**(3), 172–175.
- 95 M. Brehm and B. Kirchner, Travis – a free analyzer and visualizer for Monte Carlo and molecular dynamics trajectories, *J. Chem. Inf. Model.*, 2011, **51**(8), 2007–2023. PMID: 21761915.
- 96 E. Ortiz, R. A. Vargas and B.-E. Mellander, On the high-temperature phase transitions of CsH<sub>2</sub>PO<sub>4</sub>: A polymorphic transition? a transition to a superprotonic conducting phase?, *J. Chem. Phys.*, 1999, **110**, 4847–4853.
- 97 H. D. Lutz, Structure and strength of hydrogen bonds in inorganic solids, *J. Mol. Struct.*, 2003, **646**(1), 227–236.
- 98 C. Wehmeyer, M. Schrader, D. Andrienko and D. Sebastiani, Water-free proton conduction in hexakis(*p*-phosphonato-phenyl)benzene nanochannels, *J. Phys. Chem. C*, 2013, **117**, 12366–12372.
- 99 G. Kabbe, C. Wehmeyer and D. Sebastiani, A coupled molecular dynamics/kinetic Monte Carlo approach for protonation dynamics in extended systems, *J. Chem. Theory Comput.*, 2014, **10**, 4221–4228.
- 100 G. A. Luduena, T. D. Kühne and D. Sebastiani, Mixed Grotthuss and vehicle transport mechanism in proton conducting polymers from *ab initio* molecular dynamics simulations, *Chem. Mater.*, 2011, **23**, 1424–1429.
- 101 K. Momma and F. Izumi, VESTA3 for three-dimensional visualization of crystal, volumetric and morphology data, *J. Appl. Crystallogr.*, 2011, **44**(6), 1272–1276, DOI: 10.1107/S0021889811038970.

### 3.3.2 Article XI: Proton Conductivity in Hydrogen Phosphate/Sulfates from a Coupled Molecular Dynamics/Lattice Monte Carlo (cMD/LMC) Approach

Christian Dreßler, Gabriel Kabbe, and Daniel Sebastiani.

Proton conductivity in hydrogen phosphate/sulfates from a coupled Molecular Dynamics/Lattice Monte Carlo (cMD/LMC) approach.

*The Journal of Physical Chemistry C*, 120(36):19913–19922, 2016.

In this article, I performed all cMD/LMC calculations and I prepared the underlying AIMD simulations. I designed and conducted all analysis of the cMD/LMC approach and the AIMD trajectories. D. Sebastiani supervised the project and provided me with valuable advice during the project work and while I was writing the manuscript. G. Kabbe adapted the kinetic Monte Carlo scheme to the new molecular systems.

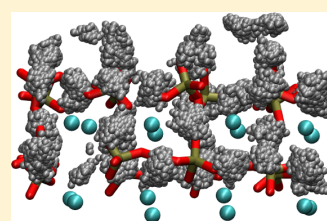
# Proton Conductivity in Hydrogen Phosphate/Sulfates from a Coupled Molecular Dynamics/Lattice Monte Carlo (cMD/LMC) Approach

C. Dreßler, G. Kabbe, and D. Sebastiani\*

Institute of Chemistry, Martin-Luther-University Halle-Wittenberg, Von-Danckelmann-Platz 4, 06120 Halle (Saale), Germany

## Supporting Information

**ABSTRACT:** The ionic conductivity of solid acids of  $\text{CsH}_n\text{XO}_4$  type ( $X = \text{P}, \text{S}; n = 2, 1$ ) varies upon chemical substitution  $\text{P} \leftrightarrow \text{S}$  and between different crystal structures ( $T_c = 503 \text{ K}$  for  $\text{CsH}_2\text{PO}_4$ ). We apply a recently developed coupled molecular dynamics/lattice Monte Carlo simulation approach (cMD/LMC<sup>1,2</sup>) to explain both the phosphate/sulfate and temperature/phase-related variations of the proton conductivity on a molecular level. Our simulation method elucidates the relative importance of the two key components of the Grotthuss-type proton conduction mechanism, proton hopping and structural reorientation, as a function of the chemical/thermodynamical conditions. We find that the chemical substitution leads to a substantial change in the proton hopping rate, which however results only in a modest variation of the proton diffusivity. The variation of the temperature of  $\text{CsH}_2\text{PO}_4$  results in a significant response of the anion rotation frequency, which turns out to be the rate-limiting process for proton conduction. In particular, the dramatic conductivity response to the phase transition can be explained by a large change of the rotation frequency. In contrast to this, our simulations show that for  $\text{CsHSO}_4$ , the local proton hopping rate is the decisive mechanism which controls long-range proton transport. These findings illustrate that the actually rate limiting factor of proton conduction in such solid acids is clearly system-dependent. Our simulated results for the proton conductivities agree almost quantitatively with experimental values, providing further evidence for the high predictive capabilities of our scale-bridging cMD/LMC simulation approach.



## 1. INTRODUCTION

Proton conduction is a process of fundamental scientific interest and has been extensively studied by chemists and physicists.<sup>3–14</sup> If we focus on the field of energy conversion, the core component of a hydrogen fuel cell is the proton exchange membrane. This membrane has to be chemically stable, durable and proton conducting.<sup>15</sup>

In the specific field of proton exchange membranes, aqueous proton transport is widely used but has an important drawback: the operating temperature is limited to 100 °C.

Solid state proton conductors (or solid acids) of the type  $\text{CsH}_y\text{XO}_4$  ( $X = \text{S}, \text{P}, \text{Se}, \text{As}; y = 1, 2$ ) have attracted considerable interest in recent years as alternative high-temperature, water-free proton exchange fuel cell membrane materials.<sup>16–20</sup> They have been shown to exhibit high ionic conductivity [ $\approx 10^{-2} \Omega^{-1} \text{ cm}^{-1}$ ].<sup>17</sup> Fuel-cell operation using electrolytes based on  $\text{CsHSO}_4$  and  $\text{CsH}_2\text{PO}_4$  has already been successfully demonstrated in the laboratory.<sup>19–21</sup>

Unlike electrolyte solutions, where proton carriers show a high degree of conformational freedom, and thus strongly promote proton conductivity, hydrogen-bonded crystalline systems exhibit a much more ordered structure, especially at low temperatures, and therefore feature much lower proton conductivity. These low temperature phases (LTP) are characterized by highly ordered hydrogen bond networks, which in some cases form 1-D chains. With increasing temperature a phase transition toward a high temperature

phase (HTP) can be observed, at which the disorder increases and the oxyanion groups are able to rotate almost freely. This leads to a disordered hydrogen bond network and a strong increase in proton conductivity. At temperatures above the transition (400–500 K), the conductivity is about  $10^{-2} \text{ S/cm}$ , approaching the value of aqueous systems.<sup>9</sup> For increased temperatures, a moderately increased conductivity of the HTP is reported.<sup>22,23</sup> The compounds  $\text{CsH}_2\text{PO}_4$  and  $\text{CsHSO}_4$  are subject to a superprotonic phase transition at 503 and 414 K leading to an increase of proton conductivity by several orders of magnitude. The involved phases are depicted in Table 1. The geometry optimized crystal structures of the HTP of  $\text{CsH}_2\text{PO}_4$  and  $\text{CsHSO}_4$  are depicted in Figure 1.

**Table 1. Characteristic Crystal Structure Parameters for the High and Low Temperature Phases (HTP/LTP) of  $\text{CsH}_2\text{PO}_4$  and  $\text{CsHSO}_4$** <sup>22</sup>

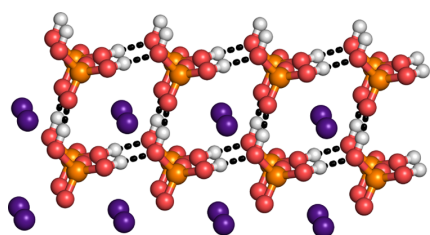
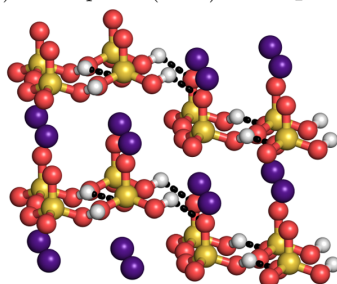
	$\text{CsH}_2\text{PO}_4$	$\text{CsHSO}_4$
HTP	$Pm\bar{3}m$ cubic	$I4_1/amd$ tetragonal
LTP	$P2_1/c$ monoclinic	$P2_1/c$ monoclinic
$\lambda_c$	503 K	414 K

Received: June 9, 2016

Revised: August 19, 2016

Published: August 22, 2016



(a) Cubic phase (HTP) of  $\text{CsH}_2\text{PO}_4$ .(b) Tetragonal (HTP) phase of  $\text{CsHSO}_4$ .

**Figure 1.** Geometry-optimized crystal structures of the HTP of  $\text{CsH}_2\text{PO}_4$  and  $\text{CsHSO}_4$ . Hydrogen atoms are shown in white, oxygen in red, sulfur in yellow, phosphorus in orange, and cesium in blue. Hydrogen bonds are represented by dashed lines.

Several experimental<sup>22–28</sup> and theoretical<sup>16,17,29–33</sup> studies of proton conduction suggested a proton conducting mechanism in  $\text{CsH}_2\text{PO}_4$  and  $\text{CsHSO}_4$  relying on the combination of two fundamental processes: proton transfer between anions and rotation of the anions, in complete analogy to the Grotthuss proton conduction in liquid water.<sup>28,34</sup> Only the interplay of both processes enables the long-range proton transfer. Considering frequent proton transfer processes without anion rotation leads to a rattling motion of the proton between two anions and no sustained proton conduction. Tuckerman stated for  $\text{CsH}_2\text{PO}_4$  a 10-fold higher frequency of proton jumps compared to the frequency of anion rotations<sup>16,28</sup>. This indicates that the anion rotation represents the rate limiting step for proton conductivity. For the sulfate compound ( $\text{CsHSO}_4$ ), Marzari also reported that proton jumps are more frequent than anion rotations in  $\text{CsHSO}_4$ . On the other hand, many experimentalists<sup>25,27,35,36</sup> and other theoretical studies<sup>32,33,37</sup> claim a reduced frequency of proton transfers between anions for  $\text{CsHSO}_4$  compared to  $\text{CsH}_2\text{PO}_4$ ,<sup>25,27</sup> derived from quasielastic neutron scattering and rf-microwave dielectric relaxation.

The determination of the diffusion coefficient is actually a challenging task for experimentalists. The diffusion coefficient can be derived by several methods, e.g., NMR relaxation times or conductivity measurements. A quick survey of the corresponding publications shows a surprisingly large variation of the reported diffusion coefficients. Depending on the experimental method deviations of up to a factor of 10 can be observed.<sup>23,28,38</sup> This large experimental uncertainty may at least be partially attributed to the variety of complex microscopic processes responsible for the diffusion phenomena at different length and time scales. In this work, we attempt to

elucidate the mechanistic picture of diffusion with a scale bridging simulation approach, which allows us to simulate proton transport on large time scales while also incorporating the atomistic details of the hydrogen bond network.

The comparatively high computational costs of purely *ab initio* molecular dynamics (AIMD) based studies of the solid acids typically restrict them to time scales of about 20–100 ps. Regarding ion conduction, however, only few elementary atomic hopping processes occur within several picoseconds. As a consequence the statistical sampling of the proton translation process can hardly be considered converged, leading to considerable deviations from experiment. So far the predicted theoretical values of  $\text{CsHSO}_4$  are too large by a factor of 10 compared to the experimental value.<sup>17,26,27</sup>

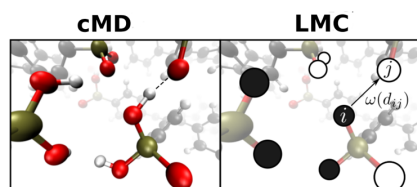
A scale bridging approach can overcome the limitations of pure electronic structure methods. In this context we apply a combination of first-principles molecular dynamics simulations and a lattice Monte Carlo scheme (cMD/LMC approach).

The algorithmic details of this combined cMD/LMC approach and its validation in terms of accuracy for the reference systems have been published elsewhere.<sup>1</sup> In this work, we use a refined version of the cMD/LMC approach<sup>2</sup> to explain the changes in proton conductivity in a series of solid acids, upon temperature variations, phase transitions, and chemical variations.

## 2. METHODS

**2.1. cMD/LMC Approach.** We have designed a simulation approach on the basis of a combination of first-principles molecular dynamics simulations and a lattice Monte Carlo scheme (cMD/LMC approach).<sup>1</sup> While the molecular dynamics simulation provides a representative conformational ensemble of the supramolecular structure at the level of density functional theory (DFT) under periodic boundary conditions, the Monte Carlo part models the long-range mobility of the acidic protons taking into account the dynamically evolving molecular structure of the system.

In the LMC part, the system is reduced to a lattice of sites which can be occupied or unoccupied, and correspond chemically to the oxygen atoms in the system. These oxygen atoms are treated as nodes and can hold one proton. Proton jumps between the oxygen atoms occur, according to a distance dependent jump rate. The basic idea behind the cMD/LMC method is illustrated in Figure 2. The jump rate can be obtained from an AIMD trajectory. It is a physical rate for the frequency of a proton transfer between two oxygen atoms. Therefore, it is dependent on the O–O distance. The unit of the proton jump rate is probability per time. According to this rate function a physical time can be assigned to the LMC steps.



**Figure 2.** Illustration of the basic cMD/LMC idea. Between the oxygen atoms  $\text{O}_i$  and  $\text{O}_j$  proton jumps occur according to a distance dependent jump rate  $\omega(d_{ij})$ .

Following these explanations, this distance dependent jump rates constitutes one of the fundamental input parameter of the LMC scheme. The actual network of transition rates between all LMC lattice sites, is obtained from the specific molecular geometry of the system at a given snapshot of the MD trajectory. The propagation of the LMC state is then followed by an update of the transition/jump rate matrix corresponding to the evolution of the MD trajectory. A crucial aspect of our model is the coupling between MD and the LMC algorithm, i.e., the way the LMC model obtains its topology from the MD trajectory. The LMC loads the atomistic structure from the MD trajectory at a time  $T$ . In each LMC time step, every proton held by an oxygen atom is considered for a jump to all neighboring unprotonated oxygen atoms with a probability given by the jump rate. The length of a LMC step is determined by the temporal unit obtained from the jump rates. For the subsequent LMC step, the coordinates of the next AIMD-trajectory-frame are loaded into the program. It is particular noteworthy that LMC steps do not alter the topology of the system.

The atomic nuclei are treated as classical particles in our AIMD simulations (Born–Oppenheimer molecular dynamics). Therefore, no quantum effects are included in our cMD/LMC scheme, which neglect the contribution of the tunneling effect to the proton transfer. For increased temperatures the influence of proton tunneling is rather small.

There are essentially only two fundamental input parameters required for a LMC/MD simulation: the distance dependent proton jump rate and the supramolecular structure (heavy atom structure from MD snapshots). Both parameters can be obtained from a AIMD simulation. By applying the LMC approach to a system, as an advantage, the proton conduction mechanism is automatically decomposed into contributions from the proton jump rate and the heavy atom structure. This allows us to explain the proton conducting mechanism in term of the two important cMD/LMC parameters and to compare the cMD/LMC results to mechanistic studies from atomistic simulations. As a result from a single cMD/LMC run of a system, the mean square displacement and the H–O–bond–autocorrelation function for several nanoseconds can be obtained on a single CPU.

**2.2. Diffusion Coefficients.** The diffusion coefficient is accessible from the linear part of the mean square displacement via the Einstein equation (eq 1).

$$D_s = \frac{1}{6} \frac{d}{dt}(\text{MSD}(t)) \quad (1)$$

Long trajectories are necessary to observe the linear regime of the MSD. These time scales can only hardly be reached by AIMD simulations. In particular for a short given AIMD trajectory, the diffusion coefficient can not be uniquely determined, because none or more than one linear regimes of the MSD can be identified.

We obtained the MSD from the AIMD by applying a multiinterval processing of the trajectory. For the determination of the MSD in the cMD/LMC method, we choose a time interval from 50 to 250 ps. The diffusion coefficient is computed by taking the slope of the linear function obtained by fitting the MSD in this interval. The error for the diffusion coefficient was obtained by regarding the calculated standard deviation of the MSD during the fitting process.

**2.3. Computational Details. AIMD.** We applied Born–Oppenheimer Molecular Dynamics (BOMD) using the

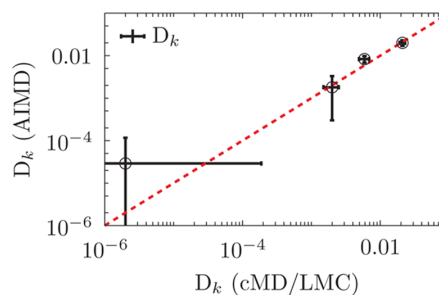
CP2K<sup>39</sup> program package to simulate the proton transfer in a series of solid acids. We utilized the Quickstep module<sup>40</sup> and orbital transformation<sup>41</sup> for faster convergence. The electronic structure was calculated with density functional theory utilizing the PBE<sup>42–44</sup> functional. A basis set of the type DZVP-MOLOPT-SR-GTH<sup>45</sup> and GTH pseudopotentials<sup>46,47</sup> were applied. Furthermore, we used the empirical dispersion correction (D3) from Grimme<sup>48</sup>. The temperature was set by a Nosé–Hoover chain thermostat<sup>49–51</sup> (NVT ensemble). The time step was chosen as 0.5 fs. The dimensions of the simulation box and the starting configurations of the systems were obtained from crystal structure data from the literature.<sup>52–54</sup>

For the initial equilibration we performed a geometry optimization of the systems followed by 10 ps AIMD using massive and global thermostating. The SCF convergence parameter was set to  $10^{-7}$ . The SCF convergence parameter is calculated as the matrix norm of the difference of the density matrices between two SCF steps. Systems of 16 formula units of  $\text{CsH}_2\text{PO}_4$  were investigated above (610, 510 K) and below (490 K) the phase transition. A system of 16 formula units of tetragonal  $\text{CsHSO}_4$  was investigated at 420 K (6 K above the phase transition). All production runs were performed for 120 ps.

**cMD/LMC.** Each Monte Carlo run was performed for the equivalent of 250 ps following the recently improved version of our cMD/LMC scheme.<sup>2</sup> The heavy atom structure was updated in the continuous mode.<sup>1</sup> An angular cut off criterion for the P–O–O and S–O–O angle of  $90^\circ$  was used for the proton jumps (for explanation see ref 2). Before starting the production run, every MC simulation was equilibrated for 125 ps.

### 3. RESULTS

**3.1. Diffusion Coefficients from cMD/LMC Simulations vs AIMD and Experimental References.** We have computed the proton diffusion coefficients for a series of solid acids (different compounds, phases and temperatures) with the combined lattice Monte Carlo (cMD/LMC) and the *ab initio* molecular dynamics (AIMD) method (see Figure 3). The four investigated systems represent the full range of proton conductivities including very strong proton conductors ( $10^{-2} \text{ \AA}^2/\text{ps}$ ) as well as insulators. Figure 3 depicts the correlation between the diffusion coefficients obtained by the two different



**Figure 3.** Correlation plot of the diffusion coefficients  $D_k$  [ $\text{\AA}^2/\text{ps}$ ] obtained from the AIMD method and the cMD/LMC approach for the systems listed in Table 2. The ideal correlation is indicated by the dotted line.



Table 2. Diffusion Coefficients [ $\text{\AA}^2/\text{ps}$ ] from Different Compounds<sup>a</sup>

	CsH <sub>2</sub> PO <sub>4</sub> cubic (HTP) 510 K	CsH <sub>2</sub> PO <sub>4</sub> monoclinic (LTP) 490 K	CsH <sub>2</sub> PO <sub>4</sub> cubic (HTP) 610 K	CsHSO <sub>4</sub> tetragonal (HTP) 420 K
D <sub>k</sub> cMD/LMC	$(5.9 \pm 1.0) \times 10^{-3}$	$\sim 0.0^b$	$(2.1 \pm 0.2) \times 10^{-2}$	$(2.0 \pm 0.5) \times 10^{-3}$
D <sub>k</sub> AIMD	$(8.3 \pm 1.5) \times 10^{-3}$	$\sim 0.0^b$	$(2.0 \pm 0.3) \times 10^{-2}$	$(1.8 \pm 1.5) \times 10^{-3}$
experimental	$(0.5\text{--}6.5) \times 10^{-3c}$	$\sim 0.0^c$	$\sim 5.5 \times 10^{-2d}$	$\sim 1 \times 10^{-3d}$
D <sub>k</sub>	$(2.9\text{--}25) \times 10^{-3c}$		(560 K)	

<sup>a</sup>AIMD diffusion coefficient and cMD/LMC diffusion coefficients rely on a 60 ps AIMD trajectory. <sup>b</sup>D<sub>k</sub> very small ( $\leq 3 \times 10^{-5}$ ) and error larger than the value itself <sup>c</sup>Reference 24. <sup>d</sup>Reference 27. <sup>e</sup>Reference 23.

methods. The diffusion coefficients do not differ more than by a factor of 2.

We validate our simulations by comparing the theoretical diffusion coefficient to experiment (Table 2). Discussing only one representative example from Table 2, the diffusion coefficient of CsH<sub>2</sub>PO<sub>4</sub> (HTP) at 510 K obtained from an AIMD trajectory is  $8.3 \times 10^{-3} \text{\AA}^2/\text{ps}$ . The diffusion coefficient determined from the cMD/LMC method was  $5.9 \times 10^{-3} \text{\AA}^2/\text{ps}$ . This value matches well the broad range of diffusion coefficients ( $0.5 \times 10^{-3}$  to  $6.5 \times 10^{-2} \text{\AA}^2/\text{ps}$ ) published by experimentalists. Also all other diffusion coefficients calculated with the cMD/LMC method are in good agreement with the diffusion coefficient from experimental investigations. The cMD/LMC diffusion coefficients do not differ more than by a factor of 2 compared to AIMD and by a factor of 3 compared to experimental diffusion coefficients.

The difficulty of benchmarking our method lies in the fact that reported values of diffusion coefficients vary by an order of magnitude. Tuckerman compared in his ab initio study of CsH<sub>2</sub>PO<sub>4</sub> the theoretically obtained proton conductivity to other methods.<sup>16</sup> He found proton conductivities ranging from 0.002 to 0.025 S/cm for the experimental values (corresponding to diffusion coefficients from  $(0.5\text{--}6.5) \times 10^{-3} \text{\AA}^2/\text{ps}$ ).<sup>24,38,55–57</sup> These “indirect” studies for the determination of the diffusion coefficient use the Nernst–Einstein equation (eq 2) to extrapolate the diffusion coefficients from conductivity measurements. In eq 2,  $N_p$  is the number of protons in the unit volume and  $D_k$  is the proton diffusion constant.

$$\sigma = \frac{e^2 N_p D_k}{k_b T} \quad (2)$$

A recent study from Kawai published in 2008, investigating the diffusion coefficient of CsH<sub>2</sub>PO<sub>4</sub> with quasi-elastic neutron scattering (QENS) and pulse field gradient nuclear magnetic resonance (PFG-NMR), stated a broad range of the experimental diffusion coefficient of 2.9 to  $25 \times 10^{-3} \text{\AA}^2/\text{ps}$  (temperature: 511–513 K).<sup>23</sup> The study of Kawai is very reliable, because they combined and compared direct and indirect methods within the same publication. The translational diffusion coefficient of  $2.9 \times 10^{-3} \text{\AA}^2/\text{ps}$  was confirmed by direct (PFG-NMR) and indirect (ac impedance) methods and is in good agreement with our theoretical value obtained by AIMD or the cMD/LMC method.

Proton diffusion is an activated process and therefore it is a challenging task to determine diffusion coefficients with AIMD. Earlier studies, using standard trajectory processing techniques, could often only compute the diffusion coefficient with a 10-fold deviation.<sup>17,26,27</sup> The large calculated errors for the diffusion coefficient in Table 2 underline the uncertainty for the determination of diffusion coefficients. Regarding proton conduction, the large error of the AIMD method can be explained, because only few elementary proton hopping

processes occur within a whole simulation. In particular simulations of materials with smaller conductivities lead to statistically not converged MSDs and large errors for the diffusion coefficient (compare Table 2, CsHSO<sub>4</sub>). Other factors, which could hinder the comparability of theoretically diffusion coefficients, are effective pressure differences between the calculations.

Keeping in mind the widespread range of the diffusion coefficients obtained from experimental and ab initio methods, the agreement of the cMD/LMC results with the other methods is noteworthy.

**3.2. Importance of Structure Update in the cMD/LMC Method.** The evolution of the underlying supramolecular structure has a major influence on the proton dynamics. In contrast to other scale-bridging methods,<sup>58</sup> the cMD/LMC approach takes this into account via an update of the oxygen positions.

In order to clarify the importance of the update of the oxygen position from the AIMD, we set up four LMC runs, which differ only by the underlying oxygen structure (see Figure 4). In

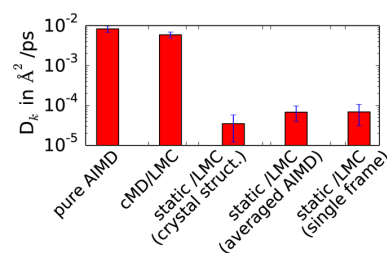


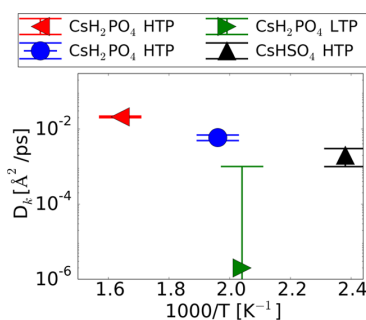
Figure 4. cMD/LMC diffusion coefficients calculated with different structure update methods. Using a fixed supramolecular structure during the whole LMC run, a drastically decreased diffusion coefficient is obtained. In such a fixed setup, the excellent proton conductor CsH<sub>2</sub>PO<sub>4</sub> exhibits conductivities in the range of an insulator.

the first LMC run, the standard cMD/LMC approach is used, updating the oxygen positions continuously (second bar Figure 4). In the second run, we used fixed oxygen positions, averaged from the MD run (fourth bar Figure 4) whereas in the third run fixed oxygen positions, obtained from the first geometry optimized initial frame of the MD (third bar Figure 4) are used. Finally we set up an LMC run with fixed oxygen positions, obtained from a randomly picked AIMD trajectory time frame (fifth bar in Figure 4).

From Figure 4, we can conclude that the diffusion coefficients for LMC runs with static oxygen positions are drastically decreased, compared to the LMC runs relying on a continuously updated AIMD trajectory. For the fixed oxygen-position-based LMC simulations, the diffusion coefficients collapse to values on the order of an insulator.

**3.3. Atomistic Explanations of Diffusion Coefficient Trends.** A very interesting aspect is the dependence of the proton diffusion coefficients on the chemical nature of the investigated systems. Here, we explain the trend of the diffusion coefficients in terms of the central cMD/LMC input parameters: jump rate and supramolecular structure.

Figure 5 depicts the diffusion coefficient for different temperatures and phases of  $\text{CsH}_2\text{PO}_4$  and for the HTP of



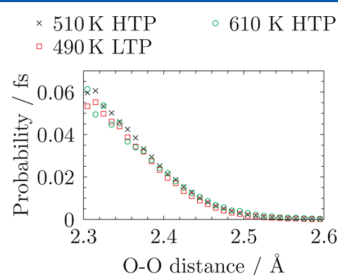
**Figure 5.** CMD/LMC diffusion coefficients for  $\text{CsH}_2\text{PO}_4$  and  $\text{CsHSO}_4$  at different temperatures are shown in this figure. The numerical values of the diffusion coefficients were already shown in Table 2.

$\text{CsHSO}_4$  obtained by the cMD/LMC approach. Two questions arise from this figure:

- 1 Why does a temperature shift influence only moderately the proton conductivity, whereas a phase shift causes drastic changes of the diffusion coefficient?
- 2  $\text{CsHSO}_4$  and  $\text{CsH}_2\text{PO}_4$  are different compounds but the diffusion coefficient of the HTP of  $\text{CsHSO}_4$  at 420 K could be extrapolated from the temperature trend of the HTP of  $\text{CsH}_2\text{PO}_4$ . Can we therefore conclude the same proton conduction mechanism?

**$\text{CsH}_2\text{PO}_4$ : Impact of Crystal Structure and Temperature.** A challenging question is the response of proton diffusion to phase transitions and temperature changes (compare Figure 5). We investigate the proton transfer mechanism in terms of the two basic cMD/LMC parameters: jump rate and heavy atom structure (supramolecular structure).

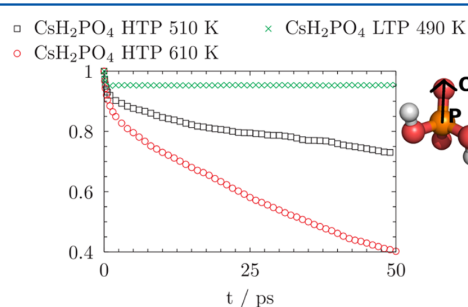
The jump rates for the simulation of the monoclinic (390 K) and the two cubic (510 and 610 K)  $\text{CsH}_2\text{PO}_4$  systems are depicted in Figure 6. The graphs of the jump rates for these three compounds are nearly the same. To exclude a jump rate dependency of the diffusion coefficients for temperature shifts



**Figure 6.** Distance dependent jump rates for  $\text{CsH}_2\text{PO}_4$  determined from the AIMD trajectory.

and phase transitions, we simulated three diffusion coefficients with the cMD/LMC approach obtained from one supramolecular structure and three different jump rates. For the calculation of these diffusion coefficient we used the supramolecular structure of the cubic  $\text{CsH}_2\text{PO}_4$  at 510 K combined with the jump rates obtained from the other simulations. The resulting diffusion coefficients differ most by 1%. Hence the jump rates of the different compounds are not able to explain the trend of the proton conductivity.

The second important parameter in terms of the cMD/LMC approach is the supramolecular structure obtained from a comparatively short AIMD trajectory. For investigation of the supramolecular structure, vector autocorrelation functions (vec-acfs) of the P–O vector for the monoclinic phase (490 K) and the cubic phase (510 and 610 K) were computed. The vector autocorrelation functions of  $\text{CsH}_2\text{PO}_4$  are depicted in Figure 7.



**Figure 7.** Vector autocorrelation function of the P–O vector determined from an AIMD trajectory.

The vec-acf of the cubic phases show an exponential decay. This decay is faster for the simulation at increased temperature. The vec-acf of the monoclinic phase shows a short decay from 1.0 to 0.95 in the first picosecond. After this short decay the vec-acf remain constant.

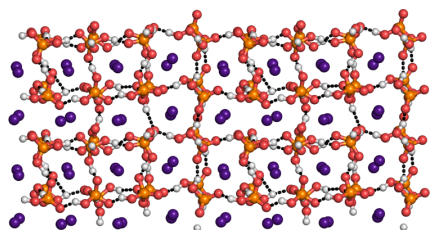
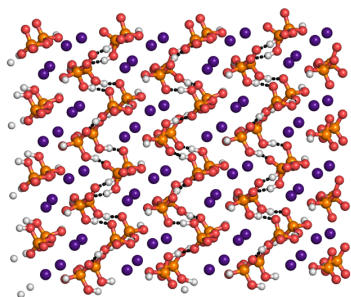
The relaxation time of the vec-acf of the P–O vectors characterizes the frequency of the phosphate rotation. Figure 7 shows, that phosphate rotation in the monoclinic phase is strongly reduced compared to the cubic phase. This is correlated with a much more ordered hydrogen bond network (HBN): The monoclinic phase has a tight 1D network and the cubic phase has a fluctuating HBN due to a variety of stable oxygen positions and corresponding hydrogen bond orientations. This has previously already been reported by Tuckerman.<sup>16</sup>

Characteristic frames from the AIMD trajectory of the monoclinic and the cubic phase are depicted in Figure 8. The figure underlines in an intuitive way the different properties of the hydrogen bond network.

Now we can combine the results from the investigations of the jump rate and the supramolecular structure. On the one hand, we observe for the monoclinic phase a high proton jump rate and no reorientation of the phosphate anions, leading to a very low proton diffusion coefficient. On the other hand, we observe for the cubic phase a high proton jump rate and a comparably fast reorientation of the phosphate anions, leading to a large proton diffusion coefficient.

While both phases exhibit frequent proton jumps thanks to the high hopping rates, the slow phosphate rotation in the monoclinic phase leads only to a rattling motion of the proton between the same two neighboring oxygen atoms. In contrast



(a) Cubic phase (HTP) of CsH<sub>2</sub>PO<sub>4</sub>.(b) Monoclinic (LTP) phase of CsH<sub>2</sub>PO<sub>4</sub>.

**Figure 8.** Snapshots of HTP and LTP CsH<sub>2</sub>PO<sub>4</sub> simulation boxes during the simulation run. Hydrogen bonds are represented by dashed lines. Note that the HBN in the cubic phase (a) is three-dimensional but less ordered compared to the one-dimensional HBN of the monoclinic phase (b).

to this, the fast phosphate rotation in the cubic phase enables a considerably larger final diffusion coefficient. This effectively constitutes a Grotthuss-type motion, in complete analogy to proton conduction in liquid water.

In summary, we compared the proton dynamics of CsH<sub>2</sub>PO<sub>4</sub> in its cubic and its monoclinic phase. By decomposing the proton jump mechanism into contributions of a distance dependent jump rate and contributions of the supramolecular structure with our cMD/LMC approach, we could show that the latter limits the overall proton conduction. In particular, a variation of the jump rate parameters could not significantly alter the overall MSD. The model of proton transfer based on these two processes—proton jumps and phosphate rotations—is widely discussed in the literature (Grotthuss mechanism). In particular for the CsH<sub>2</sub>PO<sub>4</sub>-based system we can state, in coincidence with the ab initio study of Tuckerman<sup>16</sup> and some experimental results,<sup>22–24,28,38</sup> that the rotation of the phosphate group is the rate-determining step of proton diffusion for the CsH<sub>2</sub>PO<sub>4</sub> compound. The experimental study of Kawai<sup>23</sup> stated two different diffusion coefficients, that can be interpreted as a librational ( $25 \times 10^{-3} \text{ \AA}^2/\text{ps}$ ) and translational ( $2.9 \times 10^{-3} \text{ \AA}^2/\text{ps}$ ) diffusion coefficients. In our picture of the proton conducting mechanism, we can identify the large librational diffusion coefficient with the fast proton jumps between the phosphate anions, whereas the smaller translational diffusion coefficient corresponds to the resulting overall diffusion coefficient, which is limited by the slow anion rotation.

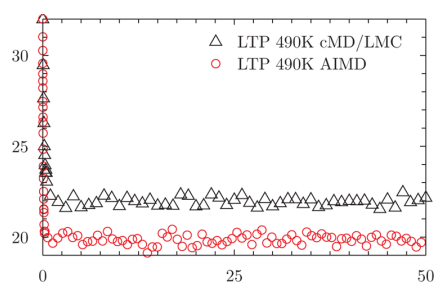
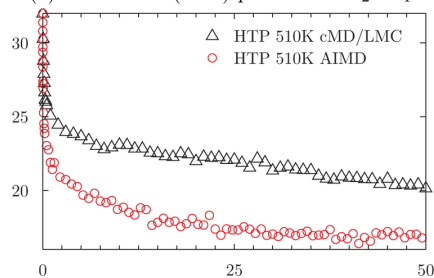
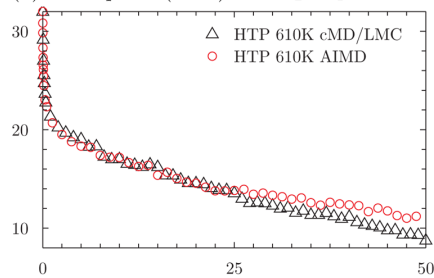
**CsH<sub>2</sub>PO<sub>4</sub>: HO-Bond Autocorrelation Function.** We used the cMD/LMC approach to investigate the covalent bond

dynamics of the HO bonds. For this we define a vector  $\mathbf{H}(t)$ , which holds the index of its covalent bonding partner for each acidic proton in the system at time  $t$ . In the case of the AIMD trajectories, we store the indices of each proton's nearest oxygen neighbor, whereas for the cMD/LMC model, the indices of the oxygen sites at which the protons are residing, are saved.

$$\eta_{\text{cov}}(t) = \left\langle \sum_i \delta_{H_i^{\text{cov}}(t_0), H_i^{\text{cov}}(t+t_0)} \right\rangle_{t_0} \quad (3)$$

The quantity reduces the multidimensional vector  $\mathbf{H}(t)$  to a scalar value by counting the covalent bondings that have not changed since time  $t_0$ .

The upper graph in Figure 9 describes the HO-bond-autocorrelation function of the monoclinic phase of CsH<sub>2</sub>PO<sub>4</sub>.

(a) Monoclinic (LTP) phase of CsH<sub>2</sub>PO<sub>4</sub>.(b) Cubic phase (HTP) of CsH<sub>2</sub>PO<sub>4</sub> at 510 K.(c) Cubic phase (HTP) of CsH<sub>2</sub>PO<sub>4</sub> at 610 K.

**Figure 9.** O–H bond autocorrelation functions for CsH<sub>2</sub>PO<sub>4</sub>. The maximum number of H-bonds in the system is 32.

After a short decay, the autocorrelation function remains constant for this compound. The short decay corresponds to the proton rattling between two oxygen atoms. Also in the monoclinic phase, frequent proton transfers are observed, but the lack of phosphate rotation leads to a constant HO-bond-autocorrelation function for longer time scales. The constant

HO-bond-autocorrelation functions for the monoclinic phase fits well to the low diffusion coefficient of this material. In contrast to that, an exponential decay of the HO-bond-autocorrelation function is observed for the super protonic phase (HTP) of  $\text{CsH}_2\text{PO}_4$  (Figure 9, middle). The decay of the HO-bond-autocorrelation function at 610 K (Figure 9, bottom) is faster compared to 510 K. The stronger exponential decay at higher temperatures corresponds with the increased diffusion coefficient. For the comparison of the AIMD and cMD/LMC method, the HO-bond-autocorrelation functions for different phases and temperatures for  $\text{CsH}_2\text{PO}_4$  are in good agreement (see Figure 9). The underlying jump rate is nearly the same for all three  $\text{CsH}_2\text{PO}_4$  simulations. That means that a similar jump rate can lead to a wide range of HO-bond-autocorrelation functions.

**Anion Driven Changes for the Proton Conduction Mechanism.** Up to this point we only investigated trends of the proton diffusion coefficient caused by phase and temperature changes. We are also interested in changes of the diffusion coefficient due to chemical modifications, e.g., upon the exchange of the anion. For this purpose, we prepared an AIMD trajectory of the  $\text{CsHSO}_4$  (HTP) system, which was subsequently fed into the cMD/LMC approach. The AIMD and cMD/LMC diffusion coefficients were in good agreement (compare Figure 3 and Table 2). Therefore, we can conclude that the proton conduction mechanism can be described quite accurately in terms of the cMD/LMC method: proton hopping and subsequent relaxation of the involved supramolecular structure (Grotthuss-type process).

For  $\text{CsH}_2\text{PO}_4$ , we already showed that the anion rotational dynamics is the rate limiting step for proton conduction. From previous AIMD studies, it is known that  $\text{CsHSO}_4$  exhibits a significantly faster anion rotational dynamics compared to  $\text{CsH}_2\text{PO}_4$ .<sup>16,17</sup> Nevertheless, the final diffusion coefficients of  $\text{CsHSO}_4$  and  $\text{CsH}_2\text{PO}_4$  are almost equal, if we assume ideal behavior with respect to the temperature (compare Figure 5).

To explain these findings, we compare the jump rates of the different high temperature phases. The magnitude of the jump rate of  $\text{CsHSO}_4$  is 2 orders smaller compared to  $\text{CsH}_2\text{PO}_4$ . The detailed graph for the distance dependent jump rate is shown in the Supporting Information. This result is remarkable, because the jump rate of the other systems investigated in this work and also of hexakis(*p*-phosphonatophenyl)benzene from an earlier publication<sup>59</sup> were in the same range. The big deviation of the jump rates of  $\text{CsH}_2\text{PO}_4$ (HTP) and  $\text{CsHSO}_4$ (HTP) is a first hint toward a jump rate dependency of the diffusion coefficient.

These observations give rise to the question, whether the diffusion coefficients of the compounds  $\text{CsHSO}_4$ (HTP) and  $\text{CsH}_2\text{PO}_4$ (HTP) are more sensitive to differences in the supramolecular structure, or rather to the differing jump rates. To address this aspect we set up a fictitious cMD/LMC simulation using the supramolecular structure of  $\text{CsH}_2\text{PO}_4$  with the jump rate parameters of  $\text{CsHSO}_4$  and vice versa. The results of these cMD/LMC runs are depicted in Table 3.

**Table 3. Diffusion Coefficients [ $\text{\AA}^2/\text{ps}$ ] from Different Combinations of Supramolecular Structure and Jump Rates**

	$\text{CsHSO}_4$ structure	$\text{CsH}_2\text{PO}_4$ structure
$\text{CsHSO}_4$ jump rate	$(2.6 \pm 0.5) \times 10^{-3}$	$(1.6 \pm 0.5) \times 10^{-3}$
$\text{CsH}_2\text{PO}_4$ jump rate	$(2.4 \pm 0.3) \times 10^{-2}$	$(5.4 \pm 1.0) \times 10^{-3}$

Utilizing the large jump rate of  $\text{CsH}_2\text{PO}_4$  in combination with the heavy atom structure of  $\text{CsHSO}_4$ , we obtained a diffusion coefficient of  $2.4 \times 10^{-2} \text{ \AA}^2/\text{ps}$ . This diffusion coefficient is five times larger compared to the diffusion coefficient obtained by the standard cMD/LMC run for  $\text{CsH}_2\text{PO}_4$  ( $5.4 \times 10^{-3} \text{ \AA}^2/\text{ps}$ ). If we use the small jump rate of  $\text{CsHSO}_4$  in combination with the supramolecular structure of  $\text{CsH}_2\text{PO}_4$  for a cMD/LMC run, a diffusion coefficient of  $1.6 \times 10^{-3} \text{ \AA}^2/\text{ps}$  is observed. This diffusion coefficient is close to the small diffusion coefficient obtained by the standard cMD/LMC run for  $\text{CsHSO}_4$  ( $2.6 \times 10^{-3} \text{ \AA}^2/\text{ps}$ ). From these observations, we conclude that the diffusion coefficient of  $\text{CsHSO}_4$  is more sensitive with respect to the jump rate compared to the diffusion coefficient of  $\text{CsH}_2\text{PO}_4$ . For variations of the anion ( $\text{HSO}_4^-$  instead of  $\text{H}_2\text{PO}_4^-$ ), the magnitude of the diffusion coefficient is mostly driven by the jump rate.

Compared to standard simulation methods only the cMD/LMC approach enabled us to figure out this jump rate dependency of the diffusion coefficients via the simulation with exchanged jump rates. Our findings explain the apparent contradiction between previous experimental and computational studies on  $\text{CsHSO}_4$ , where the computational side argued that the number of proton transfers in  $\text{CsHSO}_4$  is comparable to  $\text{CsH}_2\text{PO}_4$ ,<sup>16,17</sup> while experimental studies indicated a reduced jump rate.<sup>25</sup>

#### 4. DISCUSSION

The main feature of the cMD/LMC scheme is the coupling of microscopic structural and dynamical data from accurate molecular dynamics simulations with long time scales within the lattice Monte Carlo part. A particular advantage of our implementation is that almost the entire chemistry is contained in the first part (the MD run), which in turn renders the LMC part system-independent. This feature allows to apply our cMD/LMC algorithm to a variety of systems without the need to perform a manual analysis and decomposition of the protonation dynamics mechanisms beforehand, as it is required for other approaches. In our scheme, only the molecular dynamics simulation contains system-specific parameters (i.e., the force-field parameters), which already exist for a large variety of systems.

A particularly encouraging success of the cMD/LMC scheme is the accurate prediction of the dramatic conductivity response of the  $\text{CsH}_2\text{PO}_4$  crystal upon the transition from its cubic to the monoclinic phase. Similarly, the (considerably weaker) thermally induced changes in protonation dynamics are well reproduced. In both cases, no input other than the crystal structure (and temperature) was required for the cMD/LMC approach. In this context, it is interesting to note that the relative importance of heavy-atom structure and proton rattling between oxygen acceptors can vary considerably between different systems.

Finally, we consistently observe a significantly better convergence behavior of the cMD/LMC scheme on the duration of the underlying molecular dynamics trajectory, compared to the direct extraction of dynamical parameters from the trajectory itself. This means that the combination of a short MD trajectory (in our case: about 30 ps) with the LMC part yields diffusion properties which are much better converged than the diffusion coefficient obtained directly from that (short) trajectory. A longer molecular dynamics trajectory will improve the directly calculated diffusion

coefficient, while the results from the cMD/LMC scheme are virtually unaffected (for more details see [Supporting Information](#)). This phenomenon is caused by the fact that the actual proton transfer processes in the MC run are rare events. During a 30 ps run, only very few processes (other than proton rattling) occur within the picosecond time scale. In turn, the heavy atom structure is statistically converged to a much better degree. Since only the latter is used in the LMC part (together with accurate geometry-dependent hopping rates), the cMD/LMC scheme does not suffer from the insufficient proton transfer statistics in the molecular dynamics trajectory.

In case of the solid acids the rotation of the anions (supramolecular structure) is slow and the bottleneck for the convergence of the diffusion coefficient obtained by the cMD/LMC approach. In the future our overarching goal is to utilize classical MD for the update of the supramolecular structure. Under the assumption that the positions of the oxygen atoms are well reproduced compared to AIMD, we can easily simulate the supramolecular structure for several ns. Using this new supramolecular structure update we can assume to obtain a converged diffusion coefficient with drastically reduced computational effort.

Beside the incorporation of classical molecular dynamics trajectories, we are planning to investigate the pressure dependency of the proton transfer mechanism in the future. Toward a realistic description of long-range proton transfer, we are also planning to include the contribution of proton tunneling via post processing of the jump rates in our cMD/LMC scheme.

## 5. CONCLUSION

We have investigated the proton dynamics in several solid acid structures with very similar chemistry and very different proton conductivity. Using a recently improved simulation approach (combining MD and MC techniques) we elucidate the microscopic origin of the conductivity changes between different phases/structures. The observed mechanism of proton transfer in  $\text{CsH}_2\text{PO}_4$  and  $\text{CsHSO}_4$  can be described in terms of a general Grotthuss-type process: proton hopping and subsequent reorientation of the involved particles (anion rotation). We illustrate that the drastic change in proton conductivity between the monoclinic and cubic phase of  $\text{CsH}_2\text{PO}_4$  (at the same temperature) is caused by the different rotation ability of the phosphate groups due to a different hydrogen bonding network. In turn, gradual temperature variations for a given crystal phase yield moderate shifts of the proton conductivity due to increasing/decreasing rotation rates of the phosphates. In opposite to that the proton conductivity for the variation of the anion (hydrogen phosphate to hydrogen sulfate) is limited by the jump rate. We could show that, depending on the simulated systems, the proton jump rate (exchange of anions) or the time evolution of the supramolecular structure (thermodynamic changes) can limit the proton conductivity. Considering only one of these parameters could lead to contrary conclusions. The interplay of a high proton jump rate and a suited heavy atoms structure are required to enable efficient long-range proton transfer.

## ■ ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: [10.1021/acs.jpcc.6b05822](https://doi.org/10.1021/acs.jpcc.6b05822).

Diffusion coefficients obtained from cMD/LMC and AIMD method with respect to different AIMD trajectory lengths and comparison of the proton jump rates of  $\text{CsHSO}_4$  and  $\text{CsH}_2\text{PO}_4$  ([PDF](#))

## ■ AUTHOR INFORMATION

### Corresponding Author

\*(D.S.) E-mail: [daniel.sebastiani@chemie.uni-halle.de](mailto:daniel.sebastiani@chemie.uni-halle.de).

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

The authors gratefully acknowledge the computing time granted by the John von Neumann Institute for Computing (NIC) and provided on the supercomputer JUROPA and JURECA at the Jülich Supercomputing Centre (JSC). This work has been supported by the German Research Foundation (DFG) under Grant No. SE 1008/6-2.

## ■ REFERENCES

- (1) Kabbe, G.; Wehmeyer, C.; Sebastiani, D. A Coupled Molecular Dynamics/Kinetic Monte Carlo Approach for Protonation Dynamics in Extended Systems. *J. Chem. Theory Comput.* **2014**, *10*, 4221–4228.
- (2) Kabbe, G.; Dreßler, C.; Sebastiani, D. Toward Realistic Transfer Rates Within the Coupled Molecular Dynamics/Lattice Monte Carlo Approach. *J. Phys. Chem. C* **2016**.
- (3) Chen, K.; Hirst, J.; Camba, R.; Bonagura, C. A.; Stout, C. D.; Burgess, B. K.; Armstrong, F. A. Atomically Defined Mechanism for Proton Transfer to a Buried Redox Centre in a Protein. *Nature* **2000**, *405*, 814–817.
- (4) Tuckerman, M.; Laasonen, K.; Sprik, M.; Parrinello, M. Ab Initio Molecular Dynamics Simulation of the Solvation and Transport of Hydronium and Hydroxyl Ions in Water. *J. Chem. Phys.* **1995**, *103*, 150–161.
- (5) Fang, C.; Frontiera, R. R.; Tran, R.; Mathies, R. A. Mapping GFP Structure Evolution During Proton Transfer with Femtosecond Raman Spectroscopy. *Nature* **2009**, *462*, 200–204.
- (6) Vilciauskas, L.; Tuckerman, M. E.; Bester, G.; Paddison, S. J.; Kreuer, K.-D. The Mechanism of Proton Conduction in Phosphoric Acid. *Nat. Chem.* **2012**, *4*, 461–466.
- (7) Marx, D.; Tuckerman, M. E.; Hutter, J.; Parrinello, M. The Nature of the Hydrated Excess Proton in Water. *Nature* **1999**, *397*, 601–604.
- (8) Prod'homme, B.; Pietrobon, D.; Hess, P. Direct Measurement of Proton Transfer Rates to a Group Controlling the Dihydropyridine-Sensitive  $\text{Ca}^{2+}$  Channel. *Nature* **1987**, *329*, 243–246.
- (9) Kreuer, K.-D. Proton Conductivity: Materials and Applications. *Chem. Mater.* **1996**, *8*, 610–641.
- (10) Vuilleumier, R.; Borgis, D. Proton Conduction: Hopping along Hydrogen Bonds. *Nat. Chem.* **2012**, *4*, 432–433.
- (11) Tuckerman, M. E.; Marx, D.; Parrinello, M. The Nature and Transport Mechanism of Hydrated Hydroxide Ions in Aqueous Solution. *Nature* **2002**, *417*, 925–929.
- (12) Wraight, C. A. Chance and Design-Proton Transfer in Water, Channels and Bioenergetic proteins. *Biochim. Biophys. Acta, Bioenerg.* **2006**, *1757*, 886–912.
- (13) Bekcioglu, G.; Hoffmann, F.; Sebastiani, D. Solvation-Dependent Latency of Photoacid Dissociation and Transient IR Signatures of Protonation Dynamics. *J. Phys. Chem. A* **2015**, *119*, 9244–9251.
- (14) Chandra, A.; Tuckerman, M. E.; Marx, D. Connecting Solvation Shell Structure to Proton Transport Kinetics in Hydrogen-Bonded Networks via Population Correlation Functions. *Phys. Rev. Lett.* **2007**, *99*, 145901.
- (15) Crabtree, G. W.; Dresselhaus, M. S.; Buchanan, M. V. The Hydrogen Economy. *Phys. Today* **2004**, *57*, 39–44.



- (16) Lee, H.-S.; Tuckerman, M. E. The Structure and Proton Transport Mechanisms in the Superprotonic Phase of CsH<sub>2</sub>PO<sub>4</sub>: An Ab Initio Molecular Dynamics Study. *J. Phys. Chem. C* **2008**, *112*, 9917–9930.
- (17) Wood, B. C.; Marzari, N. Proton Dynamics in Superprotonic CsHSO<sub>4</sub>. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2007**, *76*, 134301.
- (18) Merle, R. B.; Chisholm, C. R. I.; Boysen, D. A.; Haile, S. M. Instability of Sulfate and Selenate Solid Acids in Fuel Cell Environments. *Energy Fuels* **2003**, *17*, 210–215.
- (19) Haile, S. M.; Boysen, D. A.; Chisholm, C. R.; Merle, R. B. Solid Acids as Fuel Cell Electrolytes. *Nature* **2001**, *410*, 910–913.
- (20) Chisholm, C. R. I.; Boysen, D. A.; Papandrew, A. B.; Zecevic, S.; Cha, S.; Sasaki, K. A.; Varga, A.; Giapis, K. P.; Haile, S. M. From Laboratory Breakthrough to Technological Realization: The Development Path for Solid Acid Fuel Cells. *J. Electrochem. Soc. Interface* **2009**, *18* (3), 53–59.
- (21) Boysen, D. A.; Uda, T.; Chisholm, C. R. I.; Haile, S. M. High-Performance Solid Acid Fuel Cells through Humidity Stabilization. *Science* **2004**, *303*, 68–70.
- (22) Baranov, A. I. Crystals with Disordered Hydrogen-Bond Networks and Superprotonic Conductivity. Review. *Crystallogr. Rep.* **2003**, *48*, 1012–1037.
- (23) Ishikawa, A.; Maekawa, H.; Yamamura, T.; Kawakita, Y.; Shibata, K.; Kawai, M. Proton Dynamics of CsH<sub>2</sub>PO<sub>4</sub> Studied by Quasi-Elastic Neutron Scattering and PFG-NMR. *Solid State Ionics* **2008**, *179*, 2345–2349.
- (24) Baranov, A. I.; Khiznichenko, V. P.; Shuvalov, L. A. High Temperature Phase Transitions and Proton Conductivity in Some Kdp-Family Crystals. *Ferroelectrics* **1989**, *100*, 135–141.
- (25) Badot, J. RF-Microwave Dielectric Relaxations and Phase Transitions in Superionic Protonic Acid Sulphates (Selenates). *Solid State Ionics* **1989**, *35*, 143–149.
- (26) Blinc, R.; Dolinsek, J.; Lahajnar, G.; Zupancic, I.; Shuvalov, L. A.; Baranov, A. I. Spin-Lattice Relaxation and Self-Diffusion Study of the Protonic Superionic Conductors CsHSeO<sub>4</sub> and CsHSO<sub>4</sub>. *Phys. Status Solidi B* **1984**, *123*, K83–K87.
- (27) Belushkin, A. V.; Carlile, C. J.; Shuvalov, L. A. The Diffusion of Protons in the Superionic Conductor CsHSO<sub>4</sub> by Quasielastic Neutron Scattering. *J. Phys.: Condens. Matter* **1992**, *4*, 389–398.
- (28) Kim, G.; Blanc, F.; Hu, Y.-Y.; Grey, C. P. Understanding the Conduction Mechanism of the Protonic Conductor CsH<sub>2</sub>PO<sub>4</sub> by Solid-State NMR Spectroscopy. *J. Phys. Chem. C* **2013**, *117*, 6504–6515.
- (29) Merinov, B. Proton Transport Mechanism and Pathways in the Superprotonic Phase of CsHSO<sub>4</sub> from Experiment and Theory. *Solid State Ionics* **2012**, *213*, 72–75.
- (30) Ke, X.; Tanaka, I. Atomistic Mechanism of Proton Conduction in Solid CsHSO<sub>4</sub> by a First-Principles Study. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2004**, *69*, 165114.
- (31) Ke, X.; Tanaka, I. Proton Transfer Mechanism in Solid CsHSO<sub>4</sub> by First-Principles Study. *Solid State Ionics* **2004**, *172*, 145–148.
- (32) Munch, W.; Kreuer, K. D. A Molecular Dynamics Study of the High Proton Conducting Phase of CsHSO<sub>4</sub>. *Solid State Ionics* **1995**, *77*, 10–14.
- (33) Chisholm, C. R. I.; Jang, Y. H.; Haile, S. M.; Goddard, W. A. Superprotonic Phase Transition of CsHSO<sub>4</sub>: A Molecular Dynamics Simulation Study. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2005**, *72*, 134103.
- (34) Goni-Urtiaga, A.; Presvytes, D.; Scott, K. Solid Acids as Electrolyte Materials for Proton Exchange Membrane (PEM) Electrolysis: Review. *Int. J. Hydrogen Energy* **2012**, *37*, 3358–3372.
- (35) Belushkin, A. V.; David, W. I. F.; Ibberson, R. M.; Shuvalov, L. A. High-Resolution Neutron Powder Diffraction Studies of the Structure of CsDSO<sub>4</sub>. *Acta Crystallogr., Sect. B: Struct. Sci.* **1991**, *47*, 161–166.
- (36) Damyanovich, A.; Pintar, M. M.; Blinc, R.; Slak, J. Proton Pseudoglass-to-fast-Ion-Conductor Phase Transition in CsHSO<sub>4</sub>. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1997**, *S6*, 7942–7946.
- (37) Munch, W.; Kreuer, K. D.; Traub, U.; Maier, J. Proton Transfer in the Three-Dimensional Hydrogen Bond Network of the High Temperature Phase of CsHSO<sub>4</sub>: A Molecular Dynamics Study. *J. Mol. Struct.* **1996**, *381*, 1–8.
- (38) Yamada, K. Superprotonic Conductor CsH<sub>2</sub>PO<sub>4</sub> Studied by <sup>1</sup>H, <sup>31</sup>P NMR and X-Ray Diffraction. *Solid State Ionics* **2004**, *175*, 557–562.
- (39) Hutter, J.; Iannuzzi, M.; Schiffmann, F.; VandeVondele, J. *cp2k: Atomistic Simulations of Condensed Matter Systems* **2014**, *4*, 15–25.
- (40) VandeVondele, J.; Krack, M.; Mohamed, F.; Parrinello, M.; Chassaing, T.; Hutter, J. QUICKSTEP: Fast and Accurate Density Functional Calculations Using a Mixed Gaussian and Plane Waves Approach. *Comput. Phys. Commun.* **2005**, *167*, 103–128.
- (41) VandeVondele, J.; Hutter, J. An Efficient Orbital Transformation Method for Electronic Structure Calculations. *J. Chem. Phys.* **2003**, *118*, 4365–4369.
- (42) Zhang, Y.; Yang, W. Comment on “Generalized Gradient Approximation Made Simple”. *Phys. Rev. Lett.* **1998**, *80*, 890.
- (43) Perdew, J. P.; Ruzsinszky, A.; Csonka, G. L.; Vydrov, O. A.; Scuseria, G. E.; Constantin, L. A.; Zhou, X.; Burke, K. Restoring the Density-Gradient Expansion for Exchange in Solids and Surfaces. *Phys. Rev. Lett.* **2008**, *100*, 136406.
- (44) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.
- (45) VandeVondele, J.; Hutter, J. Gaussian Basis Sets for Accurate Calculations on Molecular Systems in Gas and Condensed Phases. *J. Chem. Phys.* **2007**, *127*, 114105.
- (46) Hartwigsen, C.; Goedecker, S.; Hutter, J. Relativistic Separable Dual-Space Gaussian Pseudopotentials from H to Rn. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1998**, *58*, 3641–3662.
- (47) Krack, M. Pseudopotentials for H to Kr Optimized for Gradient-Corrected Exchange-Correlation Functionals. *Theor. Chem. Acc.* **2005**, *114*, 145–152.
- (48) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A Consistent and Accurate Ab Initio Parametrization of Density Functional Dispersion Correction (DFT-D) for the 94 Elements H-Pu. *J. Chem. Phys.* **2010**, *132*, 154104.
- (49) Nose, S. A. Unified Formulation of the Constant Temperature Molecular Dynamics Methods. *J. Chem. Phys.* **1984**, *81*, 511–519.
- (50) Martyna, G. J.; Klein, M. L.; Tuckerman, M. Nose-Hoover chains: The Canonical Ensemble via Continuous Dynamics. *J. Chem. Phys.* **1992**, *97*, 2635–2643.
- (51) Nose, S. A. Molecular Dynamics Method for Simulations in the Canonical Ensemble. *Mol. Phys.* **1970**, *52*, 255–268.
- (52) Bronowska, W. Comment on “Does the Structural Superionic Phase Transition at 231 °C in CsH<sub>2</sub>PO<sub>4</sub> Really not Exist?” [*J. Chem. Phys.* **110**, 4847 (1999)]. *J. Chem. Phys.* **2001**, *114*, 611–612.
- (53) Preisinger, A.; Mereiter, K.; Bronowska, W. The Phase Transition of CsH<sub>2</sub>PO<sub>4</sub> (CDP) at 505 K. *Mater. Sci. Forum* **1994**, *166–169*, 511–516.
- (54) Jiraak, Z.; Dlouha, M.; Vratislav, S.; Balagurov, A. M.; Beskrovnyi, A. I.; Gordelii, V. I.; Datt, I. D.; Shuvalov, L. A. A Neutron Diffraction Study of the Superionic Phase in CsHSO<sub>4</sub>. *Phys. Status Solidi A* **1987**, *100*, K117–K122.
- (55) Haile, S. M.; Liu, H.; Secco, R. A. High-Temperature Behavior of CsH<sub>2</sub>PO<sub>4</sub> Under Both Ambient and High Pressure Conditions. *Chem. Mater.* **2003**, *15*, 727–736.
- (56) Ortiz, E.; Vargas, R. A.; Mellander, B.-E. On the High-Temperature Phase Transitions of CsH<sub>2</sub>PO<sub>4</sub>: A Polymorphic Transition? A Transition to a Superprotonic Conducting Phase? *J. Chem. Phys.* **1999**, *110*, 4847–4853.
- (57) Otomo, J. Protonic Conduction of CsH<sub>2</sub>PO<sub>4</sub> and its Composite with Silica in Dry and Humid Atmospheres. *Solid State Ionics* **2003**, *156*, 357–369.
- (58) Maslowski, T.; Drzewinski, A.; Ulner, J.; Wojtkiewicz, J.; Zdanowska-Fraczek, M.; Nordlund, K.; Kuronen, A. Kinetic Monte Carlo Simulations of Proton Conductivity. *Phys. Rev. E Stat Nonlin. Soft Matter Phys.* **2014**, *90*, 012135.

(59) Wehmeyer, C.; Schrader, M.; Andrienko, D.; Sebastiani, D. Water-Free Proton Conduction in Hexakis(p-Phosphonatophenyl)-benzene Nanochannels. *J. Phys. Chem. C* **2013**, *117*, 12366–12372.

### 3.3.3 Review XII: Insight from Atomistic Simulations of Protonation Dynamics at the Nanoscale

Christian Dreßler, Gabriel Kabbe, and Daniel Sebastiani.

Insight from atomistic simulations of protonation dynamics at the nanoscale.

*Fuel Cells*, 16(6):682–694, **2016**.

This publication constitutes a review of theoretical findings with regard to proton transfer in aqueous media and water-free proton conductors. I reviewed solid acids and molecular liquids. G. Kabbe focused on the review of articles about proton transfer in Nafion-like membranes and within nano-sized pores. D. Sebastiani supervised the project and proofread the manuscript.



# Insight from Atomistic Simulations of Protonation Dynamics at the Nanoscale<sup>▲</sup>

C. Dreßler<sup>1†</sup>, G. Kabbe<sup>1†</sup>, D. Sebastiani<sup>1\*</sup>

<sup>1</sup> Martin-Luther-University Halle-Wittenberg, Institute of Chemistry, von Danckelmann-Platz 4, 06120 Halle, Germany

Received December 26, 2015; accepted October 04, 2016; published online November 11, 2016

## Abstract

In this paper, we give an overview of the role of molecular dynamics (MD) simulations in the field of proton exchange membranes. We focus on structural and dynamical findings regarding the topology of hydrogen bond networks and proton diffusion. On the one hand, findings about water-containing polymer electrolyte membrane fuel cell materials, such as

Nafion and liquid containing pore materials are discussed. On the other hand, proton conduction in water-free systems is elucidated. Here, the focus lies on phosphonic acids, which possess a rigid structure, and polymers based on phosphonic acids.

**Keywords:** Model Calculations, Molecular Dynamics Simulations, Numerical Simulation, Proton Exchange Membrane, Theoretical Model

## 1 Introduction

Proton exchange membrane or polymer electrolyte membrane (PEM) fuel cells have become an important topic as their light weight and compactness make them a good choice for portable applications, in particular in the automobile sector. In view of this range of application, certain criteria must be met by the proton exchange membrane in order to become a viable alternative to current technologies. Especially chemical stability, durability and proton conductivity need to be ensured [1]. Two central mechanisms enable proton conduction: proton hopping (also known as Grotthuss or structural diffusion) among solvent molecules such as water, and vehicular transport, where an excess proton can travel through the solvent with the help of a host molecule, to which it attaches. The Grotthuss-mechanism is characterized at the atomistic scale by two coupled processes: proton hopping and subsequent relaxation of the involved supramolecular structure. While both transport principles have been extensively studied, there is still an ongoing discussion regarding the nature of the underlying atomistic processes, even in relatively simple and uniform systems like liquid water [2–6].

In the specific field of proton exchange membranes, aqueous proton transport is very common but has an important drawback: the operating temperature is limited to 100 °C. Consequently, there is a quest for finding water-free high-temperature proton conductors [7], in order to increase the working

temperature and reduce the amount of catalysts needed. State-of-the-art material is Nafion, which relies on water channels for the proton conduction. Due to their high proton conductivity, complementary compounds like solid, phosphonic and sulfonic acids or imidazole based compounds seem to be suitable candidates for water-free proton conduction in fuel cells [8–13].

Promising candidates of solid acids for fuel cell applications are those formed by sulfate, phosphate and selenate anions. The anions are linked together by hydrogen bonds and are charge balanced by large alkali cation species [13, 14]. Unlike electrolyte solutions, where proton carriers show a high degree of freedom, and thus strongly promote proton conductivity, hydrogen-bonded crystalline systems exhibit a much more ordered structure, especially at low temperatures, and therefore, feature much lower proton conductivity. With increasing temperature, however, the disorder increases and the oxyanion groups are able to reorder almost freely. This leads to a strong increase in proton conductivity (up to four orders of magnitude for superprotonic CsH<sub>2</sub>PO<sub>4</sub>). At temperatures above the phase transition, the conductivity is about 10<sup>-2</sup> S cm<sup>-1</sup>, close to the value of aqueous systems [10, 15].

Phosphonic acids represent another class of high-temperature water-free proton conductors. In these proton conductors, the concentration of the phosphonic acid groups should, however, be high enough to ensure efficient proton transport

<sup>▲</sup> Publication is part of the Topical Issue "Theory and Modeling of Fuel Cells" 2016.

[\*] Corresponding author, daniel.sebastiani@chemie.uni-halle.de  
† Both authors contributed equally to this work.



[16, 2–4]. Finally, the spatial arrangement of these functional groups should form percolating pathways for protons moving through the membrane. Both issues can be addressed by using self-organizing supramolecular assemblies, such as recently proposed organic phosphonated molecules (see section 4.3.2) or by using new polymeric phosphonic acid materials with intrinsic proton conductivity [8, 9]. Theoretical methods can help to gain a deeper understanding of the proton conduction mechanisms in these new materials. The results of such simulations can be used to improve present materials and to develop new proton conducting materials. However, one must differentiate between the different kinds of theoretical approaches, which are used to describe fuel cell performances. A good overview can be found in [17]. The key phenomenon that a multitude of methods in this field try to elucidate, is the mechanism of proton conduction, which has been studied extensively [18]. While many heuristic approaches based on fluid dynamics exist, we focus here on the atomistic picture of proton transport.

## 2 Method Overview

In theoretical chemistry, a multitude of methods exist for the simulation of atomistic motion and the determination of molecular structures. Here, we focus in particular on molecular dynamics approaches, that is the explicit propagation of the atom cores in time caused by the forces of their surrounding environment. While the propagation is simply calculated *via* the integration of Newton's second law, the determination of the forces is the crucial point, and depends on the method used.

### 2.1 Classical Force Field Molecular Dynamics

In force field based molecular dynamics, the topology of the molecular system to be simulated needs to be specified in advance. This means, bond distances, angles between triples of atoms, and dihedral angles need to be known beforehand. For bonds and angles, a quadratic, harmonic oscillator like potential is assumed, whereas the torsional potential is modeled by a sum of cosine functions.

Several parametrized force fields exist [19–21] which are suited for different molecule types. The low computational cost of calculating the forces allows classical molecular dynamics simulations to explore time scales up to the order of microseconds with ten thousands of atoms [22–25]. For the simulation of chemically active materials, however, classical MD is unsuited because of the predefined bond topology between atoms, which does not allow bond breaking or forming. Even though reactive force fields exist, which allow bond breaking, they still need knowledge of the possible reaction paths beforehand.

### 2.2 *Ab Initio* Molecular Dynamics

For the simulation of chemical reactions, quantum chemical methods need to be employed, which are able to simulate bond breaking and formation.

In most quantum methods presented here, only the electronic structure is treated quantum mechanically, whereas the atom cores are treated like classical particles. By solving the Schrödinger equation of the electrons,  $E_{\text{el}}\Psi_{\text{el}}(\mathbf{r}) = E\Psi_{\text{el}}(\mathbf{r})$ , the electron density of a molecular system, and consequently the forces acting on each atom core, can be determined.

In contrast to classical MD, *ab initio* MD (AIMD) needs no predetermined bond specifications, as the bonds are determined from the calculated electronic densities. The drawback are the high computational costs associated with the determination of the electronic structure. Typical AIMD simulations are usually in the picosecond range with up to several hundred atoms.

### 2.3 Quantum-Mechanics/Molecular-Mechanics Methods

At the density functional theory (DFT) level, simulations of large systems are computationally several orders of magnitude more expensive than classical MD simulation, which limits their accessible time scales severely. But in many cases, the parts of a system, in which the relevant chemical reactions occur, are confined only to a partial volume of the system. Therefore, combined quantum-mechanics/molecular-mechanics (QM/MM) methods [26–29] have been developed that speed up the simulation considerably while still maintaining the benefits of the high level method. While one part of the system, often a surrounding solvent, is modeled with a classical force field, the other part is modeled with a quantum mechanical level of theory (often in a relatively small and confined region). One of the difficulties of this hybrid approach, however, is the interface region between the two different regions. If one of the atoms is located in the quantum (QM) region and the other one in the classical (MM) part, then a chemical bond is “broken” as a consequence. Similar problems arise when MM atoms are located near a QM region, because the QM and MM descriptions are not genuinely compatible.

### 2.4 Empirical Valence Bond Approach

Another method that can be considered a mixture of classical and quantum mechanical methods is the empirical valence bond (EVB) approach [30, 31]. In this approach, the wave function of a chemical system is written as a linear combination of a few selected valence bonds, which are considered essential states during some reaction. From this basis set of states the EVB Hamiltonian can be constructed. Diagonal elements are constructed using classical force fields, whereas the off-diagonal elements, which describe the transition between states, are obtained from quantum chemical calculations [32]. Special care must be taken for the initial selection of elementary valence bond states in order to fully describe the possible reac-

tions that can occur. Several extensions of the original EVB model were developed (two-state EVB [33–35], multistate-EVB [36]), which improve the method further. The multistate empirical valence bond model (MS-EVB) for example constantly redefines the EVB states according to the close environment of the reactive species. This makes it feasible to combine MS-EVB with MD (MS-MD) for the simulation of chemically active systems. In order to do an MD simulation using the (MS-)EVB approach, the EVB Hamiltonian needs to be diagonalized in each time step. Using the Hellmann-Feynman theorem for the lowest eigenstate, the resulting forces can be calculated, and thus the trajectory can be propagated to the next time step. A disadvantage of MS-EVB models of the first generations was that the computational effort increased exponentially with number of excess protons [37]. By now, methods have been developed that show only linear increase of the computation time with the number of excess protons [38].

### 3 Proton Transfer Along Water Channels

#### 3.1 Sulfonated Hydrophobic Materials: Nafion

The simulation of a Nafion membrane requires a large system size because of the complex structure of the chains, and the large amount of water molecules included in a hydrated membrane. Therefore, classical MD is often the preferred method. While it cannot simulate the actual proton transport, it can give insight into structural properties of materials. Accordingly, the following papers mostly focus on structural features and their dependence on different parameters.

Paddison et al. investigated the structure of Nafion by means of static *ab initio* calculations and classical MD simulations [41]. The systems examined were isolated triflic acid ( $\text{CF}_3\text{SO}_3\text{H}$ ), a triflic acid molecule with an additional water molecule ( $\text{CF}_3\text{SO}_3\text{H} + \text{H}_2\text{O}$ ), and di-trifluoromethane ether with a water molecule ( $\text{CF}_3\text{OCF}_3 + \text{H}_2\text{O}$ ) and the complete side chain. Furthermore, a classical MD simulation of the side chain was conducted for 20 ns at a temperature of 300 K. From their results they concluded that the ether part of the side chain is hydrophobic and stiff, whereas the  $\text{SO}_3^-$  is strongly hydrophilic and more flexible. Both the static calculations and the MD simulation of the side chain resulted in a folded conformation.

A stronger focus on the effects of the environment upon the Nafion structure is found in a paper by Vishnyakov et al. [42]. They conducted geometry optimizations of a ten-unit per-

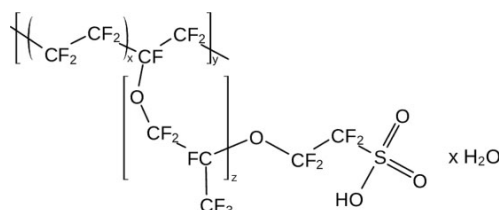


Fig. 1 Chemical structure of Nafion [39].

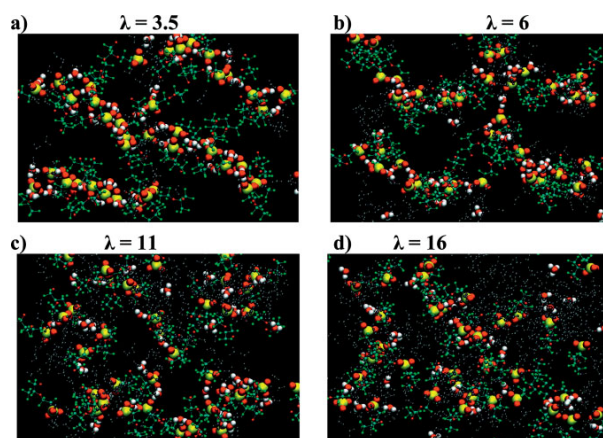


Fig. 2 Structure of a Nafion pore for varying water content. Reprinted (adapted) with permission from [40]. Copyright (2015) American Chemical Society.

fluorosulfate oligomer in vacuum and did MD simulations of a four-unit oligomer in water and methanol at  $T = 298 \text{ K}$ . In both solvents, the CCCC dihedrals always preferred *trans* positions, but solvation in water led to a more strongly folded structure than in methanol. For the determination of hydrogen bonds in the system, a geometric criterion was used: the distance between two oxygens must not be larger than  $3.4 \text{ \AA}$ , and the O-H-O angle must not exceed  $120^\circ$ . Both solvents formed strong hydrogen bonds with the sulfonic acid groups. The Nafion side chain was shown to be rather stiff, and both solvents formed stable hydrogen bonds with the  $\text{SO}_3^-$  groups. The rest of the side chain showed hydrophobic behavior. In water, an  $\text{SO}_3^-$  group formed five hydrogen bonds with the surrounding solvent, whereas in methanol, only four hydrogen bonds were formed.

In a similar vein, the same authors studied the effect of a water-methanol mixture on Nafion [43]. They found out that in the mixture, the skeleton folding was similarly strong to that in pure water. At the same time, the conformational transitions of the skeleton in the mixture were three times as high as in pure water. However, the side chains showed a high stiffness, as they did not change their conformation during the 500 ps simulation. The formation of a first solvation shell around the  $\text{SO}_3^-$  groups of the side chain could be observed, but no second solvation shell. The number of water and methanol molecules in the first solvation shell was found to be almost equal.

In the same year, Vishnyakov et al. investigated the behavior of Nafion membranes at different hydration levels [44]. The simulations showed microphase segregation in the hydrated membranes. Density and water diffusion coefficient were in good agreement with experimental results. The simulations also revealed that at high water content, isolated water clusters formed containing around 100 water molecules. During the simulation, bridges between clusters emerged for a short time with a frequency of around  $0.01 \text{ ps}^{-1}$ .

Similarly to the aforementioned paper, Urata et al. investigated the effect of the hydration level on the dynamical and static properties of water and the side chain in a swollen membrane that consists of poly(tetrafluoroethylene) backbones and perfluorosulfonic pendant side chains [45]. They did MD simulations for a percentage of water between 5 wt.% and 40 wt.% at 360 K. The results showed that water is mainly bound to the sulfonic acid. Therefore, the neighboring side chains arrange perpendicular to the hydrophilic and hydrophobic interfaces. From 20 wt.% upwards, stable water clusters form, whereas in lower hydrated settings, water paths form and break dynamically. Free water can be observed from a humidity of 20 wt.% or more, but only in 40 wt.% membranes, it forms continuous paths. In less hydrated membranes, water interacts strongly with the acidic sites, which leads to a restricted mobility.

Cui et al. investigated in a classical MD simulation, how the water content in Nafion determines its structural and dynamical properties [46]. Varying the humidity from 5 to 20 wt.%, they could show that Nafion forms hydrophobic regions consisting of the polymer backbone, and hydrophilic regions where water clusters formed. With larger amounts of H<sub>2</sub>O, more hydrophilic regions connected, forming a large water cluster. At a humidity of 5 wt.%, most hydronium ions were found to be hydrated by only up to two water molecules. With increasing water content, the hydration of the hydronium ions also increased. They conclude, that the experimentally observed drop of conductivity at around 5 wt.% can be due to the insufficiently hydrated hydronium ions, as well as the disrupted water channel network.

Focussing more on the role of the polymer backbone, Jang et al. studied *via* classical MD the nanophase segregation and diffusion in Nafion using two different polymer structures consisting of 70 nonpolar tetrafluoroethylen (TFE) segments and 10 polar perfluorosulfonic vinyl ether (PSVE) segments [47]. In the dispersed structure, the PSVE segments are evenly distributed along the backbone with one PSVE segment after every 7 TFE segments. In the 'blocky' structure all 10 PSVE segments are clustered at one end of the backbone. The simulation showed that the 'blocky' structure gives rise to larger channels and therefore, a higher diffusion, compared to the dispersed structure, where the PSVE groups are evenly distributed. From these results they suggest that the actual structure of Nafion resembles the 'blocky' structure.

Paddison et al. investigated how the number of difluoromethylene groups affect the connectivity of the SO<sub>3</sub><sup>-</sup> groups in a short-side-chain perfluorosulfonic acid membrane [48]. Geometric optimizations showed that under dry conditions with more than four difluoromethylene (CF<sub>2</sub>) groups, no hydrogen bonds were formed between neighboring SO<sub>3</sub><sup>-</sup> groups on the same backbone. In order to form a continuous hydrogen bond network between the SO<sub>3</sub><sup>-</sup> groups, one (two, three) water molecules were needed if the chains were separated by five (seven, nine) CF<sub>2</sub> groups. The amount of water needed to observe proton dissociation to the first hydration shell also depends on the separation of the side chains: In the case of five CF<sub>2</sub> groups separating the side chains, after adding five water

molecules, proton dissociation occurred on both SO<sub>3</sub><sup>-</sup> groups. On the other hand, in the case of seven (nine) CF<sub>2</sub> groups separating the side chains, only one proton dissociation could be observed with five (six) added water molecules. Paddison et al. see in these results the importance of the Zundel ion in a minimally hydrated setting, as it may suggest that proton transport *via* structural diffusion can still occur, even though the water molecules are hydrogen bonded to the SO<sub>3</sub><sup>-</sup> groups.

Several studies focus on the water diffusion within the Nafion membrane by means of classical MD. The authors often try to bypass the lack of bond breaking within classical MD, by manually including H<sub>3</sub>O<sup>+</sup> ions within the simulation box.

In an attempt to elucidate the mechanisms of water diffusion, Venkatnathan et al. conducted classical MD simulations of a Nafion membrane including hydronium ions [40]. Figure 1 shows the structure of the system in the MD simulation. They showed that at low hydration levels, hydronium ions form hydrogen bonds with water molecules between the SO<sub>3</sub><sup>-</sup> groups. The diffusion coefficient for water agrees well with experimental data, but that of the hydronium ions is underestimated. The reason for this is probably that the Grotthuss mechanism cannot be modeled by classical MD. Furthermore, at low hydration levels, Nafion loses its ionic form when protons attach chemically to the SO<sub>3</sub><sup>-</sup> groups. This feature cannot be modeled by classical MD, either.

Brandell et al. conducted classical MD simulations on water diffusion in Nafion, as well as Dow and Aciplex systems (which have different side chain lengths compared to Nafion) for two levels of hydration (5 and 15 water molecules per HSO<sub>3</sub> group) [49]. Their results show that Nafion has the highest side-chain mobility compared with the two other compounds. They give two possible reasons for this: one reason is that Nafion has the optimal ratio of hydrophobic and hydrophilic regions, which results in an ideal water channel network. On the other hand, the side chain motion may give rise to higher ion mobility *via* a "paddle-wheel" mechanism. At the same time, however, the highest water diffusion can be observed in the middle of the water channels, where the influence of the side chains is negligible. Therefore, they also point out that the Grotthuss mechanism is not modeled in their simulation, which might alter the presented results.

Devanathan et al. investigated the change of the Nafion nanostructure *via* MD simulations [50]. Their simulations revealed that low hydration of the Nafion pore leads to hydronium ions being strongly bound to the sulfonate groups, which in turn prevents the vehicular proton transport. Furthermore, at low hydration levels, a single hydronium ion is surrounded by multiple SO<sub>3</sub><sup>-</sup> groups, which opposes structural diffusion of protons. Variation of the hydration level showed that the number of hydronium ions being in a bridging configuration is in excellent agreement with results from neutron scattering experiments.

Devanathan et al. also determined the time H<sub>2</sub>O molecules and H<sub>3</sub>O<sup>+</sup> ions reside within the first solvation shell of the SO<sub>3</sub><sup>-</sup> groups [51]. Variation of the number of water molecules

per sulfonic group (denoted as  $\lambda$ ) showed that the mean residence time of water is at about 1 ns at  $\lambda = 1$ , and decreases to around 75 ps at  $\lambda = 20$ . The residence times of  $\text{H}_3\text{O}^+$  are larger by a factor of 2.5 to 4.5.

Even though the investigations of water diffusion *via* classical MD does yield information about diffusion processes in Nafion, quantitative comparisons with experimental data often reveal deviations, which are due to the lack of the Grotthuss type proton hopping within those simulations.

Hofmann et al. addressed this issue by developing a reactive force field for the modeling of the Grotthuss mechanism in classical MD [52]. From the MD simulation they conclude that proton transfer in membranes under low hydration is a second order reaction, whose time step is limited by the time a pair of one hydronium ion and one water molecule needs to form. They report that the self-diffusion coefficients of water are underestimated from 30% to 40% compared to experimental data. The temperature dependence of the self-diffusion, on the other hand, shows good agreement. Nevertheless, they state that their method does not reach *ab initio* quality.

Besides classical MD, several publications investigate the nature of proton conduction by means of the EVB approach.

Petersen et al. investigated *via* an EVB model the role of vehicular and Grotthuss type proton transport within the Nafion membrane [53]. By decomposing the mean squared displacement into a contribution from the vehicular mechanism and from the Grotthuss type charge transport, they could observe that there is actually a negative correlation of these two effects, resulting in an overall diffusion, which is smaller than the sum of the contributions from Grotthuss and vehicular mechanism. Interestingly, this effect is not observed in bulk water, where almost no negative correlation of the two mechanisms occurs.

Feng et al. [54] investigated proton transport in hydrated Nafion by means of an MS-EVB model. They could show that the distance to a sulfonate group plays an important role for the formation of Zundel- and Eigen-cations. On transport times over 1.27 ps, proton hopping was shown to be an important factor.

Seeliger et al. simulated proton transport in Nafion using an EVB model and purely classical MD [55]. They investigated both structural changes of Nafion and proton transport properties for a dry ( $\lambda = 5$ ) and a wet ( $\lambda = 10$ ) membrane. The dominant proton transfer mechanism observed for dry and wet membrane is the Grotthuss type mechanism. From this the authors conclude that the increased activation energy of the proton transfer at low hydration levels is not due to surface conduction mechanisms. Instead they suggest a fluctuative bridging mechanism as described by Vishnyakov and Neimark [44]. They add, however, that because of the simulation lengths (up to 30 ns) the existence of such bridges cannot unambiguously be proven.

Actual *ab initio* simulations of Nafion were done by Devanathan et al. [56]. They modeled a Nafion membrane for hydration levels 3, 9 and 15 (water molecules per acid group). The simulated Nafion system consists of 680 atoms, and addi-

tional 20 to 40 water molecules (depending on the hydration level). The simulations were run between 75 ps ( $\lambda = 9, 15$ ) and 130 ps ( $\lambda = 3$ ). Whereas in the dry setting, the hydrated proton is mostly bound to the  $\text{SO}_3^-$  group, a percolating water cluster is formed at higher hydration levels making it possible for the proton to move through the water network. For the hydration levels  $\lambda = 9$  and  $\lambda = 15$ ,  $\text{H}_5\text{O}_2^+$ ,  $\text{H}_7\text{O}_3^+$  and  $\text{H}_9\text{O}_4^+$  cations can be observed, which agrees with experimental observations. At the highest hydration level, the proton diffusion coefficient is  $0.9 \cdot 10^{-5} \text{ cm}^2 \text{ s}^{-1}$ . This agrees with experimental results and shows that in Nafion, proton hopping is the more important mechanism for proton transport compared to vehicular transport.

Several studies have investigated the importance and mechanisms of surface proton transport. In view of the performance of PEM fuel cells, this aspect is especially important for low hydration levels, where the role of proton transport in bulk water diminishes, and the contribution of the proton transport along the Nafion-water-interface increases.

Roudgar et al. simulated proton transport at an array of sulfonic acid terminated surface groups [57]. They determined the minimum energy paths of a proton transfer and found that there is an accompanying rotation and tilt of accepting and donating surface groups during the proton transfer.

Vartak et al. investigated interfacial proton transport *via ab initio* MD at sulfonic acid terminated surface groups [58]. On a hexagonal grid of surface groups of type  $\text{CF}_3\text{SO}_3\text{H}$  with one water molecule per surface group, proton transport was observed at minimal hydration. The authors distinguish between two types of transitions: a local defect-type, where the donor sulfonic group loses a hydrogen bond and the acceptor sulfonic group gains a hydrogen bond; the other type is a concerted hydronium ion transition, in which the number of interfacial hydrogen bonds stays constant. The resulting activation/reaction Helmholtz energies are 0.6 eV/0.5 eV for the local defect-type and 0.3 eV/0.0 eV for the concerted type (with a computational error of 0.1 eV). Compared to the energy barrier in bulk water, the activation energy for the concerted movement is two to three times higher.

In conclusion, proton transport in sulfonated hydrophobic materials has been elucidated from different angles including both parameterized methods (MD, EVB) as well as *ab initio* calculations. Even though many questions regarding the mechanism of proton transport have been answered, a unified description featuring both atomistic quantum mechanical processes and processes on mesoscopic time scales still poses a challenge. Nevertheless, the understanding of the different atomistic mechanisms involved in the proton transport process within Nafion can serve as a basis for more abstract models, such as coarse-grained approaches or statistic models. For example a phenomenological description of the transport processes involved within a PEM fuel cell was developed by Eikerling et al. [59].



### 3.2 Stiff Nano-sized Pores

Similar to Nafion, pore materials trap the solvent in a confined region, in which bulk water-like conditions prevail.

Habenicht et al. investigated the proton transfer and dissociation of water confined in carbon nanotubes (CNT), which were functionalized with perfluorosulfonic acid (PFSA) groups *via ab initio* MD [5]. They showed that a decrease of the distance between sulfonic acid groups leads to a strong increase of the proton dissociation. Furthermore, fluorination of the CNT walls led to an increased occurrence of hydronium ions compared to Zundel ions.

In another paper, Habenicht et al. investigated the importance of a hydrophobic environment in CNT functionalized with PFSA groups [61]. It was found that an increase of the sulfonic acid groups led to an increased dissociation of the acidic protons. However, this can sometimes lead to trap states, where a hydronium ion forms hydrogen bonds to two sulfonic acid groups, which then hinders the full solvation of the acidic proton. Addition of fluorine to the CNT walls led to a stabilization of the hydrogen bonds, and an enhanced dissociation of acidic protons.

Allolio et al. investigated the characteristics of bulk water confined in an amorphous silica pore (see Figure 3) [60]. They found that the tetrahedral structure of H<sub>2</sub>O was distorted up to the second solvation shell because of the partially broken hydrogen bond network at the walls. While the diffusion behavior slowed down near the walls, it showed bulk characteristics in the middle of the pore.

Adeagbo et al. simulated confined bulk water between layers of  $\alpha$ -quartz to observe transport phenomena and chemical processes at the interface by means of *ab initio*  $\alpha$ -quartz MD [62]. The surfaces were Si-terminated and O-terminated. Within two picoseconds, dissociation of the water molecules on the Si-terminated quartz surface could be observed, which finally leads to a complete hydroxylation of both surfaces.

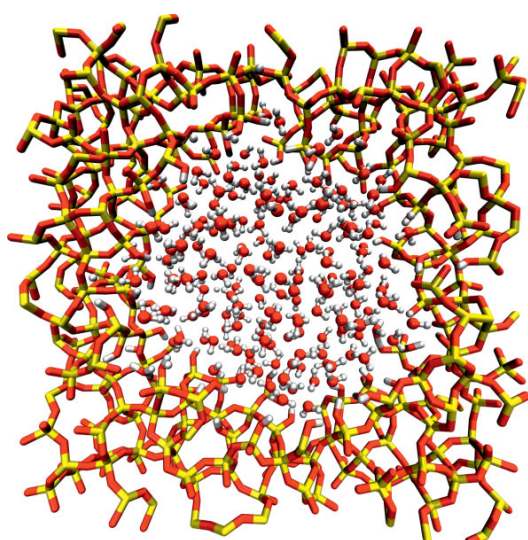


Fig. 3 Bulk water confined within a silica pore [60].

Compared to bulk water, the confinement reduces the water diffusion by one third.

Marschall et al. simulated the behavior of water confined within a mesoporous silica material functionalized with SO<sub>3</sub>H (Si-MCM-41) [63]. Experiments showed a continuous increase of the proton conductivity with temperature. Their simulations indicate that the interplay between water content and group density is decisive for the resulting proton conductivity: whereas for small densities of the SO<sub>3</sub>H groups, the proton mobility will strongly depend on the amount of water molecules inside the pore, pores with a higher SO<sub>3</sub>H density show less dependence on the water amount.

### 3.3 Bulk Liquid Water

Proton transfer in bulk water constitutes an important part of the overall conductivity of Nafion-like PEM fuel cells. While proton transport along the surface regions of a pore depends on the composition of the fuel cell material, the bulk water mechanism can be investigated separately. Due to the involved bond breaking between proton and oxygen atoms, quantum chemical methods need to be employed.

Laasonen et al. did an *ab initio* simulation of water using DFT [64]. They were able to reproduce several quantities such as RDF, dipole moment, ionization energy, and the diffusion coefficient of water in good agreement with experimental results,

Silvestrelli et al. investigated structural and electronic properties of water liquid water by means of *ab initio* MD [65]. Compared to the gas phase, the liquid phase shows a more spherically distributed electronic charge density around the oxygen atoms, and an increased intramolecular O–H distance, which leads to a 60% larger dipole moment than in the gas phase.

Going one step further, Marx et al. observed the behavior of a hydrated excess proton in water, treating all nuclei as quantum mechanical particles [66]. They could show that the intermediate steps during a proton transfer cannot solely be described by the formation of HO<sub>3</sub><sup>+</sup> or H<sub>9</sub>O<sub>4</sub><sup>+</sup> complexes. Instead, complexes occur, where the proton is shared equally between two water molecules.

Tuckerman et al. then investigated the behavior of the shared proton in hydrogen bonds [67]. They found out, that the OH<sup>+</sup> complex can be described classically, whereas for the Eigen-/Zundel-cation hydrogen bond, quantum zero-point motion needs to be considered even at room temperature. They also investigated the characterization of OH<sup>-</sup> and H<sup>+</sup> ion movement in water by means of *ab initio* simulation [68]. Their results show that the behavior of OH<sup>-</sup> and H<sup>+</sup> differs strongly. Even though protons are not treated quantum mechanically, and nonadiabatic electron dynamics are neglected, they find good agreement with experimental data, which leads them to believe that these effects play only a negligible role.

Besides the aforementioned purely quantum chemical calculations of bulk liquid water, several papers have been pub-

lished that determine bulk water properties using the EVB approach.

James et al. used an EVB model to determine global minima geometries of water clusters [32], Vuilleumier et al. developed an EVB model of an excess proton in water, which they used to compute the forces in an MD run [69]. Their results show good agreement with pure *ab initio* results.

Sagnella et al. used a two-state EVB model to describe proton transfer in a  $\text{H}_5\text{O}_2^+$  cluster [35]. They fitted the parameters of the EVB potential to *ab initio* data (at the Hartree Fock and MP2 level) and used it for both classical and path integral MD.

Several papers exist which have incrementally refined the EVB method. For example, polarizable water models were investigated in conjunction with EVB by Lefohn et al. [70] and Walbaran et al. [71]. Wick used a polarizable EVM model to investigate the hydronium behavior at the air-water interface [72]. Proton transfer in bulk water by means of an anharmonic water force field was employed by Park et al. [73].

## 4 Water-free Proton Conduction

### 4.1 Solid Acids

Solid acids represent a complementary membrane material to polymers. Their ion lattice is rigid and cannot relax during a proton transfer. Therefore, fluctuations of the local structure are even more important. As an immediate consequence, different phases of a given system can show a totally different behavior. Particular examples are  $\text{CsHSO}_4$  and  $\text{CsH}_2\text{PO}_4$ .

Ke et al. [74, 75] studied the electronic structure and the proton-transfer mechanism in monoclinic (CHS-II, no proton conductor) and cubic (CHS-I, super protonic)  $\text{CsHSO}_4$  phases. They determined proton transfer paths and energy barriers using first-principles methods. They observed that in both phases proton transfer within the sulfate tetrahedron is not possible because of the high energy barrier, and the protons can jump forwards and backwards within the hydrogen bond. Proton transfer to the next neighboring sulfate tetrahedron is feasible for CHS-I, whereas it is almost forbidden for CHS-II. The arrangement of the hydrogen bond is responsible for this difference. In CHS-I, the hydrogen bond network is disordered, which enables the formation of a new hydrogen bond by rotating one tetrahedron without breaking other hydrogen-bonds. In contrast to this observation, the hydrogen bond network of CHS-II is highly ordered and therefore, anion rotation is not possible. Consequently, protons cannot be transferred to the next neighboring anion in this case. Ke et al. extrapolated from their calculations, that reorientations of the sulfate tetrahedron in CHS-I can take place very frequently, which is in accordance with the experimental observation.

The reorientation and proton transfer rate in  $\text{CsHSO}_4$  (super protonic phase) was studied explicitly by Münch et al. [76] using classical molecular dynamics (MD) simulations. The proton transfer rate was calculated indirectly from the distribution of the O-O distances and a semi-empirical model potential. They reported a 100 times faster reorientation rate

( $10^{11}\text{ s}^{-1}$ ) compared to the proton transfer rate ( $10^9\text{ s}^{-1}$ ). These values are in the same order of magnitude as experimental results. The activation energy for the proton transfer was determined as 0.35 eV.

In a second classical MD study, Münch et al. [77] simulated proton transfer by treating the proton as an independent entity. This way, they could describe the dynamics by means of an adequate intermolecular potential. The authors characterized the proton motion in  $\text{CsHSO}_4$  (super protonic phase) as largely uncorrelated, which is found to be in reasonable agreement with randomly distributed protons. The protonic motion was, therefore, considered to be rather liquid-like. The macroscopic transport properties, such as the proton diffusion coefficient and the protonic conductivity, were also determined from the simulation. Both, proton diffusion coefficient and protonic conductivity, were found to be in agreement with experimental data.

Haile et al. [78] utilized classical MD to take a closer look at the phase transition between CHS-II and CHS-I. They studied the structure as a function of temperature from 298 to 723 K, and reproduced the phase transition to the expected tetragonal phase, even though the simulated temperature of 598 K deviated from the experimental temperature (414 K). It is particularly noteworthy that they confirmed a drastic change of the rotational dynamics of the anions at the transition. The authors suggested that the increased conductivity is caused by the rapid reorientation occurring in the high temperature phase. Phase transitions in similar systems are known to be associated with an order-disorder transition of the proton positions in the O-H-O bonds. The successful observation of the phase transition in this system led Haile et al. to the conclusion that in this phase transition only a heavy atom transfer is involved.

Wood et al. [79] explicitly investigated the proton dynamics of  $\text{CsHSO}_4$  in the high temperature phase at three different temperatures (500, 650, 750 K) *via ab initio* molecular dynamics (AIMD). They confirmed that the Grotthuss mechanism is the dominant picture, characterized by very frequent proton jumps with a high reversal rate. For the sulfate tetrahedra, they stated a reorientation behavior in terms of small angular jumps. A proton jump leads to lattice stress, which is followed by a fast relaxation in terms of the reorientation of the anion. Previous experimental and theoretical investigations suggested a higher rotational frequency of the anions compared to the frequency of the proton jumps. Surprisingly, Wood et al. reported, that the timescale of hydrogen bond dynamics is the same for proton transfer. The AIMD proton diffusion coefficient deviates by a factor of ten from the experimental value. A common explanation is that AIMD diffusion coefficients are suffering from the short trajectories and therefore, limited statistics.

In 2008, Lee et al. [80] published an *ab initio* MD study of the superprotonic phase of  $\text{CsH}_2\text{PO}_4$ . They benchmarked their simulations by comparison of calculated infra-red spectra to experimental ones. They stated that proton diffusion in this compound follows a general Grotthuss mechanism. After a

slow reorientation of the phosphate anion a fast proton hopping to a neighboring phosphate ion was observed. The smaller frequency of phosphate reorientation appeared to be the bottleneck for proton diffusion, compared to the greater proton hopping rate. In agreement with the investigation of the high temperature phase of  $\text{CsHSO}_4$  (CHS-I), the origin of high proton conductivity seems to be the dynamical disorder of the cubic system. In Figure 4, a snapshot from the disordered cubic phase is shown in comparison to the ordered monoclinic phase of  $\text{CsH}_2\text{PO}_4$ . Due to the short timescales accessible by AIMD (20 ps), a full phosphate rotation was not observed. Nevertheless, the diffusion coefficient was reproduced in good agreement with experimental values [80]. The reduced phosphate rotation frequency of  $\text{CsH}_2\text{PO}_4$ , compared to the sulfate rotation frequency of  $\text{CsHSO}_4$ , can be explained in terms of the increased number of hydrogen bonds that have to be broken to enable anion rotation.

The AIMD studies of Lee et al. [80] and Wood et al. [79] revealed a detailed direct view on the proton dynamics. Besides the classical Grotthuss diffusion motion, including proton jumps and relaxation of the anions, both studies reported a second diffusion mechanism, which does not require phosphate reorientation. This Grotthuss type structural diffusion mechanism occurred on a magnitude faster time scale.

## 4.2 Molecular Liquids

### 4.2.1 Phosphoric Acid

Neat liquid phosphoric acid has the highest intrinsic proton conductivity of any known substance [11]. Therefore, it is a

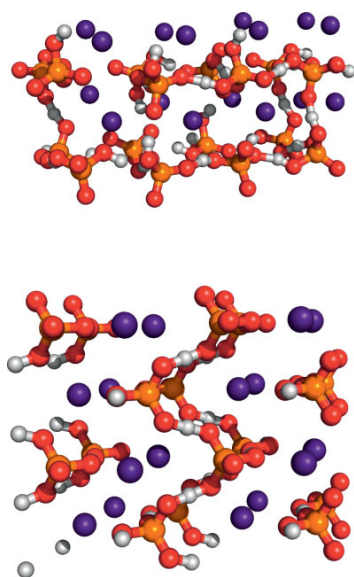


Fig. 4 Snapshots of cubic and monoclinic  $\text{CsH}_2\text{PO}_4$  simulation boxes during the simulation run. In the cubic phase, the hydrogen bond network shows dynamical disorder, whereas in the monoclinic phase a highly ordered hydrogen bond network is observed.

suited compound as a starting point towards the investigation of more complex proton conductors containing functional groups based on phosphorus. Vilčiauskas et al. [81] investigated binding energies of small clusters (two to six molecules of phosphoric acid) as a function of the cluster size. The authors calculated *via* static quantum mechanic methods increasing binding energies with respect to the increasing number and strength of hydrogen bonds formed for larger cluster sizes. Consequently, they estimated that the largest cluster size was still too small to describe the binding in the condensed phase of the phosphonic acid. For larger cluster sizes they stated a higher degree of charge delocalization and therefore, a facilitated proton transfer.

Vilčiauskas et al. [11] prepared an AIMD simulation to fully encompass the proton dynamics in bulk phosphoric acid. They reported strong, polarizable hydrogen bonds in phosphoric acid, leading to the formation of extended hydrogen-bonded chains. The formation of ordered structures is in contrast to the expected situation in an excellent proton conductor. Nevertheless, the interplay of these chains (solvent interactions) and a sufficient degree of disorder (caused by a frustrated hydrogen-bond network), enable the high concentration of diffusive protons.

### 4.2.2 Phosphonic Acid

Phosphonic acid derivatives are used widely as low humidity proton conducting electrolytes in fuel cells. For instance, the polyphosphonic compounds Hexakis-(p-phosphonatophenyl)benzene and Poly(vinylphosphoric acid) (PVPA) are discussed later in this article. Joswig et al. [82] utilized semi-empirical density functional based method (DFT-B) to investigate proton transfer in liquid phosphonic acid *via* molecular dynamics. Proton transfer was only observed in the case of an excess/defect proton. They reported that local geometric constraints, corresponding to a strong hydrogen bond (O-H-O angle of  $180^\circ$ , O-O distances lower than  $2.5 \text{ \AA}$ ), are necessary to enable proton transfer. In many cases, a Zundel-like species (an excess proton is shared between two phosphonic acid molecules) prevents efficient proton conduction. For the Zundelions, the same geometric criterion was observed as for the proton transfer events. So the authors concluded a nonlocal contribution to the mechanism of proton transfer. The self-diffusion coefficient obtained from the trajectories was one magnitude smaller compared to experimental values. The authors suggested finite size effects as an explanation.

### 4.2.3 Sulfonic Acid and Imidazole Based Compounds

Sulfonic acid and imidazoles, representing sulfur and nitrogen containing proton conductors, were also investigated as model compounds *via* theoretical methods. Paddison et al. [12] studied imidazole, sulfonic and phosphonic functionalized heptane with first principles based electronic structure calculations. They calculated torsional barriers and reported that imidazole is clearly the most labile when used as a substituent in an alkyl chain. By simulating minimum energy conformations



of the methyl dimers of imidazole, sulfonic and phosphonic acid, they figured out that binding of the pairs is greatest in the phosphonic acids and lowest for the imidazoles. Also, interactions of the methyl compounds with a single water molecule were studied. They reported that the magnitude of the interaction of methyl phosphonic and sulfonic acid were in the same range. The interaction energy for the imidazole-water dimer was reduced. Their study suggested, that the oxoacids will probably retain water better under very low humidity conditions and the water molecules of the first hydration shell will be stronger bound to phosphonic and sulfonic acid.

Via the combination of  $^1\text{H-NMR}$  measurements and theoretical calculations, Sebastiani et al. [83] and Goward et al. [84] investigated proton conducting crystals formed from ethylene oxide tethered imidazole heterocycles (see Figure 5). They calculated  $^1\text{H-NMR}$  shifts in excellent agreement with experimental investigations and characterized the microscopic hydrogen bonding structure of this class of materials. The authors reported the existence of highly ordered and disordered domains in the crystal. The ordered domains do not contribute to the proton conduction, due to their rigid hydrogen bond network.

To get a deeper insight into the proton dynamics in an imidazole molecule chain, Münch et al. [85] prepared an AIMD at 390 K. The authors reported that proton conduction follows a Grotthuss mechanism, involving proton transfer and a local rather than a long-range cooperative reorientation of the imidazole chain. The reorientation step was rate determining, as proton transfer occurred within a time scale of 0.3 ps and the reorientation time of the imidazole was extrapolated to values of 30 ps. Due to the short total length of the AIMD trajectory (6 ps), a full successful reorientation could not be observed.

Hanning et al. used an MS-EVB model to simulate proton transfer in liquid imidazole [86]. Using *ab initio* PES from a protonated imidazole in the gas phase, the authors found that

their results were in good agreement with experiments of liquid imidazole. They observed that the first solvation shell of protonated imidazolium molecules is highly ordered, due to formation of hydrogen bonds. In contrast, the second solvation shell is much more disordered. From this, they conclude that proton hopping is highly localized in this compound, and only molecules close to the excess proton play a role during the hopping process.

To overcome the restrictions of electronic structure theory, Maslowski et al. [87] used a simplified one-dimensional microscopic model of the proton transport in benzimidazole  $\text{C}_7\text{H}_6\text{N}_2$ . They identified two important elementary steps (benzimidazole rotation and proton jumps between the benzimidazoles) for the long range proton transfer and created a two-stage Grotthuss proton migration mechanism. They reported that the relative frequency of reorientation and diffusion processes is crucial for the proton conductivity. The theoretical proton currents for benzimidazole were found to be in very good agreement with the experimental data. The authors pointed out that the thermal lattice vibrations, which modify the H-bond potential, play an essential role in the conduction process.

To get insight into more realistic proton conducting materials, Cavalcanti et al. [88] and Harvey et al. [89] utilized classical MD to study immobilized imidazole molecules tethered to alkyl backbones. Cavalcanti et al. reported that a spacer length of three  $-\text{CH}_2-$  groups between imidazole and a solid surface is most favorable to a proton-transfer reaction. In accordance with these findings, Harvey et al. stated for the exchange of pentyl chains against methylene groups the tendency to form percolating hydrogen-bond networks in combination with dramatically decreasing imidazole ring reorientation times.

In order to study another kind of immobilization effects in a complex chemical environment, Eisbein et al. [90] investigated confined imidazole in a metal organic framework

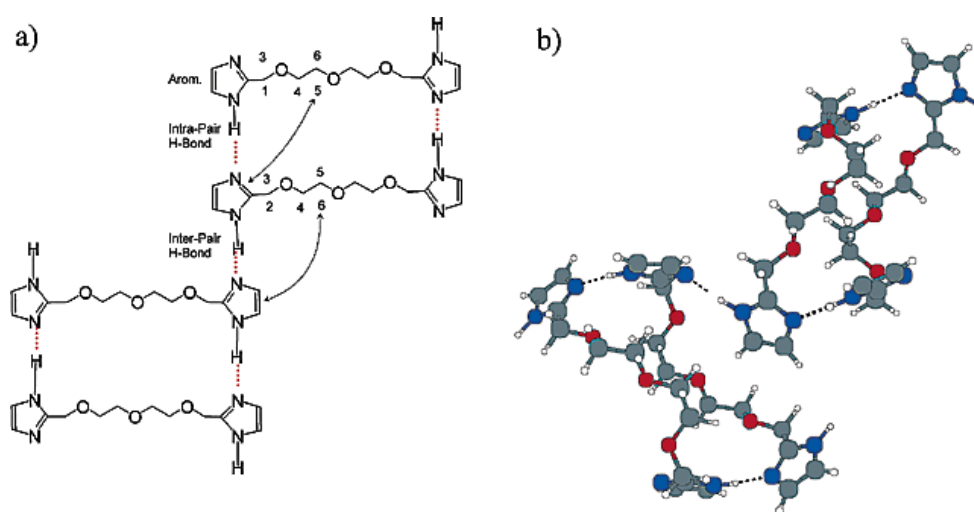


Fig. 5 (a) Molecular structure of paired Imi-2EO molecules and (b) segment of crystal structure of Imi-2EO. Labels are placed in crystal structure to depict the positions of the protons according to the assignments determined from the calculations of the  $^1\text{H}$  resonance frequencies.

(MOF) compound. Protonated imidazole molecules exhibit stronger interaction with the carboxylate groups of the MOF and showed therefore, a reduced diffusivity. In this system, the most important contribution to the proton conductivity is provided by the vehicle mechanism. In addition, hopping events increase the overall transport.

### 4.3 Polymeric and Liquid Crystalline Compounds

State-of-the-art materials used in proton conducting fuel cell membranes are mostly complex polymeric compounds. Often, the polymeric matrix is functionalized by some imidazoles, triazoles, phosphonic or sulfonic acids. Before starting studying these complex materials it is helpful to develop an understanding of the basic proton conduction mechanism in simpler systems.

#### 4.3.1 Poly(vinylphosphonic acid)

Compared to other polymeric proton conducting materials, poly(vinylphosphonic acid) (PVPA) is a rather simple compound. But as a nontrivial polymer, simulation of its proton dynamics is still a challenging task. Many of the proton conducting materials, relying on phosphonic acid groups, suffer from phosphonic acid condensation leading to anhydrides. This anhydrid formation, which decreases the proton conductivity, was investigated by Lee et al. [91, 92]. They studied the structure and the local proton mobility of poly(vinylphosphonic acid) in terms of a polymer model via first principle calculations. They prepared very short AIMD simulations in the range of 1 ps to obtain relaxed structures for the calculation of  $^1\text{H}$ - and  $^{31}\text{P}$ -NMR shifts. Based on these calculations, two different types of hydrogen bonded P-OH resonances were observed (acidic protons and phosphonic acid anhydride). It was reported, that formation of anhydride leads to a decrease of proton conductivity via two different mechanisms: At the one hand, charge transport pathways were blocked through immobilization of charge carriers. On the other hand, rotation of the anhydride group, as a prerequisite of Grotthuss proton diffusion, was hindered.

Utilizing static quantum mechanic calculations, Heggen et al. [93] studied the condensation behavior of organic phosphonic acid groups. They used phosphonic acid and methylphosphonic acid as model compounds and calculated the reaction energies for the formation of the dimer, the trimer, and the cyclic trimer. They stated that the formation of dimers and trimers is energetically possible both for unsubstituted phosphonic acid and for organic phosphonic acids. Due to the ring strain, the formation of the cyclic trimer could be ruled out. A direct mechanism for the dimerization of two molecules of phosphonic acid or methylphosphonic acid was not observed, due to high transition-state energies. Consequently, this leads to the assumption, that anhydride formation in condensed phase is a more complex process, at least involving proton transfer with other proton donating and accepting groups.

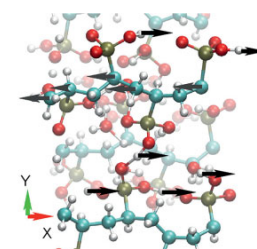
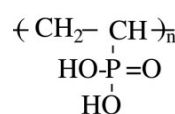


Fig. 6 Molecular structure of poly(vinylphosphonic acid) (PVPA) and our model system for PVPA. The arrows indicate the external force bias.

From various experimental studies it is known that the water molecules cannot be totally removed from the channels of PVPA and that the residual water strongly influences the proton conductivity. So, the simulations of Ludueña et al. [9] addressed the question to which extent the proton conductivity is a function of the amount of residual water. The study also shed light on the general conduction mechanism in PVPA. They prepared AIMD trajectories of PVPA in a temperature interval ranging from 100 to 600 K, suited to simulate the proton dynamics over several ps. As a result of the simulations, the authors were able to show, that the residual water has a local vehicle/carrier function for the excess protons. Proton hopping events between adjacent acids were very frequently observed. Nevertheless, they did not contribute to long range proton conduction. Ludueña et al. stated, that the Grotthuss style hopping mechanism must be supported by short distance transport of hydronium ions to the neighboring acids. So the main role of water in the conduction mechanism is to bridge the H-bond percolation path complementing Grotthuss proton hopping between phosphonic acid groups. This proton conduction mechanism in PVPA was termed carrier-mediated Grotthuss mechanism.

#### 4.3.2 Hexakis(*p*-phosphonatophenyl)benzene

Jiménez-García et al. investigated the proton conductivity of organic molecules containing phosphonic acid groups. They could show that the supramolecular ordering of the molecules plays a large role for proton conduction in PEM fuel cells.

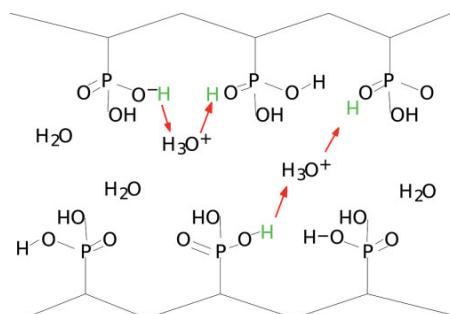


Fig. 7 Scheme of the suggested mechanism, where a Grotthuss-style hopping mechanism is supported by short-distance transport of hydronium ions to neighboring acids. We call this a carrier mediated Grotthuss mechanism.

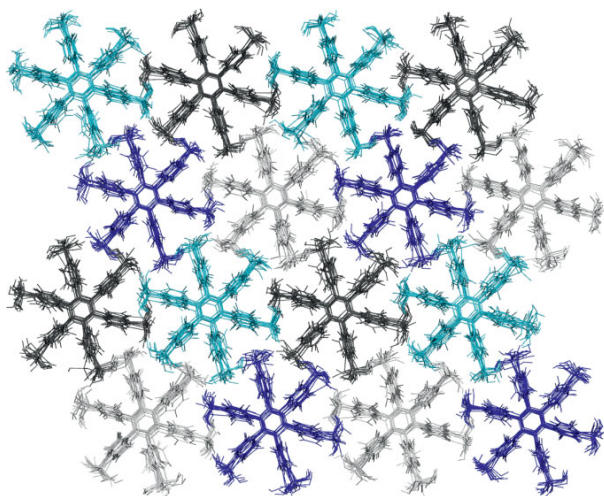


Fig. 8 Supramolecular arrangement of the *p*-6PA-HPB disks in MD simulations. At each interface of three columns, a proton channel forms, along which proton transport is possible.

They concluded that molecules that self-assemble into columns are the most promising candidates for proton conducting materials [94]. One of the materials examined, hexakis(*p*-phosphonatophenyl)benzene (*p*-6PA-HPB), showed especially promising properties: a conductivity of  $3.2 \cdot 10^{-3} \text{ S cm}^{-1}$ , which stays constant over a temperature range from 100 to 180 °C [95]. In X-ray experiments, a columnar stacking of the individual *p*-6PA-HPB molecules could be observed.

Wehmeyer et al. simulated the proton conduction mechanisms in *p*-6PA-HPB via *ab initio* MD under dehydrated conditions [8]. The AIMD showed that in each center of three neighboring columnar stacks, a quasi-one-dimensional channel is formed, along which proton transfer is possible. In contrast to membrane materials such as Nafion or PVPA, *p*-6PA-HPB shows, an anisotropic behavior of the proton mobility, as movement in *z*-direction, i.e. parallel to the columnar alignment, is preferred. Furthermore, spontaneous autodissociation of the phosphonic acid groups could be observed, even though no water molecules are present.

## 5 Conclusion

This overview has presented selected topics in the field of proton transport in PEM fuel cells by means of atomistic simulations. On the material side, simulations of proton transport in water-based membrane materials and water-free materials have been reviewed. The majority of the studies presented here find that the Grotthuss mechanism is the prevailing proton transport process. In general, two prerequisites are necessary for Grotthuss-like proton conduction: proton transfer and afterwards structural reorientation. Both prerequisites can be fulfilled by a weak and fluctuating hydrogen bond network. Water-free systems on the other hand, often exhibit a more rigid topology.

MD methods can help in this regard to analyze the arrangement of the hydrogen bond networks, and to determine, how structural and thermodynamic changes influences can alter them. For the observation of proton hopping, *ab initio* methods are necessary, however, as they are able to simulate bond breaking without any preliminary assumptions about the final proton path.

## References

- [1] G. W. Crabtree, M. S. Dresselhaus, M. V. Buchanan, *Phys. Today* **2004**, *57*, 39.
- [2] K.-D. Kreuer, A. Wohlfarth, *Angew. Chem. Int. Ed. Engl.* **2012**, *51*, 10454.
- [3] R. Vuilleumier, D. Borgis, *Nat. Chem.* **2012**, *4*, 432.
- [4] K. Kobayashi, T. Shirasaka, A. Sato, E. Horn, N. Furukawa, *Angew. Chem. Int. Ed.* **1999**, *38*, 3483.
- [5] B. F. Habenicht, S. J. Paddison, M. E. Tuckerman, *Phys. Chem. Chem. Phys.* **2010**, *12*, 8728.
- [6] C. Wang, S. J. Paddison, *Phys. Chem. Chem. Phys.* **2010**, *12*, 970.
- [7] S. J. Paddison, *Annu. Rev. Mater. Res.* **2003**, *33*, 289.
- [8] C. Wehmeyer, M. Schrader, D. Andrienko, D. Sebastiani, *J. Phys. Chem. C.* **2013**, *117*, 12366.
- [9] G. A. Ludueña, T. D. Kühne, D. Sebastiani, *Chem. Mat.* **2011**, *23*, 1424.
- [10] K.-D. Kreuer, *Chem. Mat.* **1996**, *8*, 610.
- [11] L. Vilčiauskas, M. E. Tuckerman, G. Bester, S. J. Paddison, K.-D. Kreuer, *Nat. Chem.* **2012**, *4*, 461.
- [12] S. J. Paddison, K.-D. Kreuer, J. Maier, *Phys. Chem. Chem. Phys.* **2006**, *8*, 4530.
- [13] S. M. Haile, C. R. I. Chisholm, K. Sasaki, D. A. Boysen, T. Uda, *Farad Discuss* **1970**, *134*, 17.
- [14] S. M. Haile, D. A. Boysen, C. R. Chisholm, R. B. Merle, *Nature* **2001**, *410*, 910.
- [15] C. R. I. Chisholm, D. A. Boysen, A. B. Papandrew, S. Zecvic, S. Cha, K. A. Sasaki, A. Varga, K. P. Giapis, S. M. Haile, *J. Electrochem. Soc. Interface* **2009**, *18*, 53.
- [16] H. Steininger, M. Schuster, K. D. Kreuer, A. Kaltbeitzel, B. Bingöl, W. H. Meyer, S. Schauff, G. Brunklaus, J. Maier, H. W. Spiess, *Phys. Chem. Chem. Phys.* **2007**, *9*, 1764.
- [17] A. Z. Weber, J. Newman, *Chem. Rev.* **2004**, *104*, 4679.
- [18] M. Eikerling, *J. Phys. Condens. Matter.* **2011**, *23*, 230301.
- [19] W. Damm, A. Frontera, J. Tirado-Rives, W. L. Jorgensen, *J. Comput. Chem.* **1997**, *18*, 1955.
- [20] J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, D. A. Case, *J. Comput. Chem.* **2004**, *25*, 1157.
- [21] S. L. Mayo, B. D. Olafson, W. A. Goddard, *J. Phys. Chem.* **1990**, *94*, 8897.
- [22] L. Monticelli, E. J. Sorin, D. P. Tieleman, V. S. Pande, G. Colombo, *J. Comput. Chem.* **2008**, *29*, 1740.
- [23] S. V. Lyulin, A. A. Gurtovenko, S. V. Larin, V. M. Nazarychev, A. V. Lyulin, *Macromolecules* **2013**, *46*, 6357.
- [24] P. L. Freddolino, F. Liu, M. Gruebele, K. Schulten, *Biophys. J.* **2008**, *94*, L75.

- [25] D. Bai, G. Chen, X. Zhang, W. Wang, *Langmuir* **2011**, *27*, 5961.
- [26] M. Eichinger, P. Tavan, J. Hutter, M. Parrinello, *J. Chem. Phys.* **1999**, *110*, 10452.
- [27] M. J. Field, P. A. Bash, M. Karplus, *J. Comput. Chem.* **1990**, *11*, 700.
- [28] A. Laio, J. VandeVondele, U. Rothlisberger, *J. Chem. Phys.* **2002**, *116*, 6941.
- [29] P. D. Lyne, M. Hodoscek, M. Karplus, *J. Phys. Chem. A.* **1999**, *103*, 3462.
- [30] A. Warshel, R. M. Weiss, *J. Am. Chem. Soc.* **1980**, *102*, 6218.
- [31] A. Warshel, R. M. Weiss, *Annals New York Academy of Sciences* **1981**, *367*, 370.
- [32] T. James, D. J. Wales, *J. Chem. Phys.* **2005**, *122*, 134306.
- [33] T. Mabuchi, A. Fukushima, T. Tokumasu, *J. Chem. Phys.* **2015**, *143*, 014501.
- [34] J. Lobaugh, G. A. Voth, *J. Chem. Phys.* **1996**, *104*, 2056.
- [35] D. E. Sagnella, M. E. Tuckerman, *J. Chem. Phys.* **1998**, *108*, 2073.
- [36] U. W. Schmitt, G. A. Voth, *J. Chem. Phys.* **1998**, *102*, 5547.
- [37] G. A. Voth, *Acc. Chem. Res.* **2006**, *39*, 143.
- [38] F. Wang, G. A. Voth, *J. Chem. Phys.* **2005**, *122*, 144105.
- [39] Wikipedia, the free encyclopedia Nafion-polymer **2009**. [Online, accessed December 18, 2015].
- [40] A. Venkatnathan, R. Devanathan, M. Dupuis, *J. Phys. Chem. B.* **2007**, *111*, 7234.
- [41] S. Paddison, *Solid State Ionics* **1998**, *113*, 333.
- [42] A. Vishnyakov, A. V. Neimark, *J. Phys. Chem. B* **2000**, *104*, 4471.
- [43] A. Vishnyakov, A. V. Neimark, *J. Phys. Chem. B* **2001**, *105*, 7830.
- [44] A. Vishnyakov, A. V. Neimark, *J. Phys. Chem. B* **2001**, *105*, 9586.
- [45] S. Urata, J. Irisawa, A. Takada, W. Shinoda, S. Tsuzuki, M. Mikami, *J. Phys. Chem. B.* **2005**, *109*, 4269.
- [46] S. Cui, J. Liu, M. E. Selvan, D. J. Keffer, B. J. Edwards, W. V. Steele, *J. Phys. Chem. B.* **2007**, *111*, 2208.
- [47] S. S. Jang, V. Molinero, T. Çagin, W. A. Goddard, *J. Phys. Chem. B.* **2004**, *108*, 3149.
- [48] S. J. Paddison, J. A. Elliott, *J. Phys. Chem. A* **2005**, *109*, 7583.
- [49] D. Brandell, J. Karo, A. Liivat, J. O. Thomas, *J. Mol. Model* **2007**, *13*, 1039.
- [50] R. Devanathan, A. Venkatnathan, M. Dupuis, *J. Phys. Chem. B.* **2007**, *111*, 8069.
- [51] R. Devanathan, A. Venkatnathan, M. Dupuis, *J. Phys. Chem. B.* **2007**, *111*, 13006.
- [52] D. W. M. Hofmann, L. Kuleshova, B. D'Aguzzo, *J. Mol. Model* **2008**, *14*, 225.
- [53] M. K. Petersen, G. A. Voth, *J. Phys. Chem. B.* **2006**, *110*, 18594.
- [54] S. Feng, G. A. Voth, *J. Phys. Chem. B.* **2011**, *115*, 5903.
- [55] D. Seeliger, C. Hartnig, E. Spohr, *Electrochimica Acta* **2005**, *50*, 4234.
- [56] R. Devanathan, N. Idupulapati, M. D. Baer, C. J. Mundy, M. Dupuis, *J. Phys. Chem. B.* **2013**, *117*, 16522.
- [57] A. Roudgar, S. P. Narasimachary, M. Eikerling, *M. Chemical Physics Letters* **2008**, *457*, 337.
- [58] S. Vartak, A. Roudgar, A. Golovnev, M. Eikerling, *J. Phys. Chem. B.* **2013**, *117*, 583.
- [59] M. Eikerling, A. A. Kornyshev, A. M. Kuznetsov, J. Ulstrup, S. Walbran, *J. Phys. Chem. B.* **2001**, *105*, 3646.
- [60] C. Allolio, F. Klameth, M. Vogel, D. Sebastiani, *Chemphyschem.* **2014**, *15*, 3955.
- [61] B. F. Habenicht, S. J. Paddison, M. E. Tuckerman, *J. Mater. Chem.* **2010**, *20*, 6342.
- [62] W. A. Adeagbo, N. L. Doltsinis, K. Klevakina, J. Renner, *Chemphyschem.* **2008**, *9*, 994.
- [63] R. Marschall, P. Tölle, W. L. Cavalcanti, M. Wilhelm, C. Köhler, T. Frauenheim, M. Wark, *J. Phys. Chem. C.* **2009**, *113*, 19218.
- [64] K. Laasonen, M. Sprik, M. Parrinello, R. Car, *J. Chem. Phys.* **1993**, *99*, 9080.
- [65] P. L. Silvestrelli, M. Parrinello, *J. Chem. Phys.* **1999**, *111*, 3572.
- [66] D. Marx, M. E. Tuckerman, J. Hutter, M. Parrinello, *Nature* **1999**, *397*, 601.
- [67] M. E. Tuckerman, *Science* **1997**, *275*, 817.
- [68] M. E. Tuckerman, D. Marx, M. Parrinello, *Nature* **2002**, *417*, 925.
- [69] R. Vuilleumier, D. Borgis, *Chemical Physics Letters* **1998**, *284*, 71.
- [70] A. E. Lefohn, M. Ovchinnikov, G. A. Voth, *J. Phys. Chem. B.* **2001**, *105*, 6628.
- [71] S. Walbran, A. A. Kornyshev, *J. Chem. Phys.* **2001**, *114*, 10039.
- [72] C. D. Wick, *J. Phys. Chem.* **2012**, *116*, 4026.
- [73] K. Park, W. Lin, F. Paesani, *J. Phys. Chem. B.* **2012**, *116*, 343.
- [74] X. Ke, I. Tanaka, *Solid State Ionics* **2004**, *172*, 145.
- [75] X. Ke, I. Tanaka, *Phys. Rev. B.* **2004**, *69*, 165114.
- [76] W. Münch, K. D. Kreuer, *Solid State Ionics* **1995**, *77*, 10.
- [77] W. Münch, K. D. Kreuer, U. Traub, J. Maier, *J. Mol. Struct.* **1996**, *381*, 1.
- [78] C. R. I. Chisholm, Y. H. Jang, S. M. Haile, W. A. Goddard, *Phys. Rev. B.* **2005**, *72*, 134103.
- [79] B. C. Wood, N. Marzari, *Phys Rev B* **2007**, *76*, 134301.
- [80] H.-S. Lee, M. E. Tuckerman, *J. Phys. Chem. C.* **2008**, *112*, 9917.
- [81] L. Vilčiauskas, S. J. Paddison, K.-D. Kreuer, *J. Phys. Chem. A.* **2009**, *113*, 9193.
- [82] J.-O. Joswig, G. Seifert, *J Phys Chem B* **2009**, *113*, 8475.
- [83] D. Sebastiani, G. Goward, I. Schnell, H. W. Spiess, *Comp. Theor. Chem.* **2003**, *625*, 283.
- [84] G. R. Goward, M. F. H. Schuster, D. Sebastiani, I.vSchnell, H. W. Spiess, *J. Phys. Chem. B.* **2002**, *106*, 9322.
- [85] W. Münch, K.-D. Kreuer, W. Silvestri, J. Maier, G. Seifert, *Solid State Ionics* **2001**, *145*, 437.



- [86] H. Chen, T. Yan, G. A. Voth, *J. Phys. Chem. A* **2009**, *113*, 4507.
- [87] T. Masłowski, A. Drzewiński, J. Ulner, J. Wojtkiewicz, M. Zdanowska-Frączek, K. Nordlund, A. Kuronen, *Phys. Rev. E. Stat. Nonlin Soft Matter. Phys.* **2014**, *90*, 012135.
- [88] W. L. Cavalcanti, D. F. Portaluppi, J.-O. Joswig, *J. Chem. Phys.* **2010**, *133*, 104703.
- [89] J. A. Harvey, S. M. Auerbach, *J. Phys. Chem. B* **2014**, *118*, 7609.
- [90] E. Eisbein, J.-O. Joswig, G. Seifert, *J. Phys. Chem. C* **2014**, *118*, 13035.
- [91] Y. J. Lee, B. Bingöl, T. Murakhtina, D. Sebastiani, W. H. Meyer, G. Wegner, H. W. Spiess, *J. Phys. Chem. B* **2007**, *111*, 9711.
- [92] Y. J. Lee, T. Murakhtina, D. Sebastiani, H. W. Spiess, *J. Am. Chem. Soc.* **2007**, *129*, 12406.
- [93] B. Heggen, S. Roy, F. Müller-Plathe, *J. Phys. Chem. C* **2008**, *112*, 14209.
- [94] L. Jiménez-García, A. Kaltbeitzel, V. Enkelmann, J. S. Gutmann, M. Klapper, K. Müllen, *Adv. Funct. Mater.* **2011**, *21*, 2216.
- [95] L. Jiménez-García, A. Kaltbeitzel, W. Pisula, J. S. Gutmann, M. Klapper, K. Müllen, *Angew. Chem. Int. Ed. Engl.* **2009**, *48*, 9951.



## Chapter 4

# Conclusion

This thesis deals with the development of scale-bridging methods for the simulation of proton conduction and aims to establish an efficient approach for the calculation of intermolecular electrostatic interactions. The investigation of proton conduction started from *ab initio* molecular dynamics simulations of the solid acid family  $\text{CsH}_y\text{XO}_4$  ( $X = \text{S, P, Se}, y = 1, 2$ ) to investigate the quantitative behavior of the diffusion coefficients in these compounds. Proton diffusion in this family of solid acids can be explained in terms of the interplay of the proton transfer frequency between the anions and the rotational reorientation rate of the anions. The *ab initio* molecular dynamics trajectories of the  $\text{CsH}_y\text{XO}_4$  ( $X = \text{S, P, Se}, y = 1, 2$ ) family were also used to benchmark the combined Molecular Dynamics/Lattice Monte Carlo (cMD/LMC) approach for the simulation of long-range proton transfer.

For this compound class, the cMD/LMC method was able to reproduce the enormous change in proton mobility due to phase transitions as well as the moderate increase in the conductivity due to simple variations of the temperature. It is clearly demonstrated that only the dynamic updating of the lattice allows for the simulation of the high proton conductivities within the solid acid family. In addition, the cMD/LMC approach is refined in this thesis by introducing a cut-off angle for the calculation of proton transfer probabilities in order to avoid unphysical proton jumps within the oxygen grid.

By neglecting the explicit proton correlation in the cMD/LMC approach, a further drastic reduction in the dimensionality becomes possible. The resulting Molecular Dynamics/Matrix Propagation (MDM) method condenses the dynamic information on proton transport within an entire molecular dynamics simulation into an  $M \times M$  matrix where  $M$  is the number of oxygen atoms. The system is further reduced to the positions of the  $M$  oxygen atoms and the proton distribution is described by an  $M$ -dimensional state space. This enormous reduction in dimensionality allows for analytical (instead of numerical) analysis of the model. A thorough mathematical discussion resulted in the verification of the uniqueness of the solutions of the MDM approach and provided the proof of its correct asymptotic behavior, i.e. all protons are equally distributed across the (chemically equivalent) oxygen atoms for large time scales. With appropriate modifications of the transition matrix, proton correlation can be implicitly incorporated into the MDM method. It is proved in this thesis that these changes conserve the Markov character of the MDM method.

In order to highlight the increase in length and time scales, the MDM approach was used to explicitly compute the non-equilibrium molecular dynamics of protons in the solid acid  $\text{CsH}_2\text{PO}_4$  on the micrometer length scale. It is shown that an excess proton distribution initially localized within a few nanometers diffuses through a  $8 \mu\text{m}$  sized system within 5 ms, in full agreement with the common diffusion laws. Alongside this proof-of-principle example, the scale-bridging approach for the simulation of long-range proton transfer is also used to explain the experimentally measured conductivity



behavior of a nanoporous network of  $\text{CsH}_2\text{PO}_4$  in a composite solid acid fuel cell electrode.

The second part of this thesis is dedicated to the investigation of efficient representations of the linear density-density response function (LDDRF). The LDDRF provides a tool for the calculation of the fully self-consistent density response of a molecule to arbitrary potentials (in practice: the electrostatic field arising from nearby molecules), corresponding to the exact polarization at all multipolar orders. The moment expansion introduced by Scherrer and Sebastiani achieves a drastic reduction in dimensionality compared to the eigensystem representation of the LDDRF. This thesis provides a thorough derivation of the mathematical foundations of the moment expansion and allow for its generalization to arbitrary compact, positive and self-adjoint linear operators. It is demonstrated that the moment expansion can be understood in terms of a QR decomposition of an appropriate matrix representation of the LDDRF. Unfortunately, the calculation of this initial matrix representation requires thousands of *ab initio* calculations. To address this problem, a much more efficient algorithm - the direct moment expansion - for the calculation of the moment expansion itself was developed. It is shown that the direct moment expansion can be explained in terms of a Cholesky decomposition of a small matrix. In a final development step, a modified representation - the reduced eigensystem representation - is derived that allows us to define a trivial criterion for the convergence of the approximation to the density response. Thanks to its novel eigensystem-like structure, the significant reduction in the dimensionality becomes apparent for the calculation of the density-density response function.

When a very large number of response calculations for the same molecule is required, the efficiency gain obtained from the application of the moment expansion or the reduced eigensystem representation has a maximum. A case in point for this situation is molecular dynamics simulation of a molecular liquid. The employment of the efficient representations of the LDDRF in classical molecular dynamics would enable the incorporation of polarization at all multipolar orders. To facilitate this, the geometry dependence of the moment expanded states by means of a Taylor expansion is also investigated in this thesis.

## Appendix A

# General Linear Response Theory

In this appendix, the Lehmann representation of a frequency-dependent response function is derived. The discussion will closely follow reference 38.

We start from the expectation value of a quantum mechanical operator  $\hat{\alpha}$ :

$$\alpha_0 = \langle \Psi_0, \hat{\alpha} \Psi_0 \rangle, \quad (\text{A.1})$$

where  $\Psi_0$  is the ground-state many-body wavefunction, which fulfills  $\hat{H}_0 \Psi_0 = E_0 \Psi_0$ , where  $\hat{H}_0$  is the static Hamiltonian and  $E_0$  the lowest eigenvalue.

Let  $F(t)$  be an external field that is coupled with an observable  $\hat{\beta}$ . At time  $t_0$ , we switch on a time-dependent perturbation according to

$$\hat{H}_1(t) = F(t)\hat{\beta}, \quad t \geq t_0, \quad (\text{A.2})$$

which leads to a perturbed system and a time-dependent expectation value of  $\hat{\alpha}$ :

$$\alpha(t) = \langle \Psi(t), \hat{\alpha} \Psi(t) \rangle, \quad t \geq t_0. \quad (\text{A.3})$$

The response  $\alpha(t) - \alpha_0$  of  $\hat{\alpha}$  to the perturbation from equation (A.2) can be expanded in terms of powers of the field  $F(t)$ :

$$\alpha(t) - \alpha_0 = \alpha_1(t) + \alpha_2(t) + \dots, \quad (\text{A.4})$$

where  $\alpha_1(t)$  is the linear response,  $\alpha_2(t)$  is the quadratic (second-order) response,  $\alpha_3(t)$  is the third-order response and so on.

In order to derive an explicit expression for the time-dependent linear response  $\alpha_1(t)$ , we have to start from the following two equations (A.5) and (A.9), which arise from time-dependent many-particle theory. For a more detailed description, see reference 38.

First, the time evolution of the operator  $\hat{\beta}$  is given in the interaction picture by

$$\hat{\beta}(\tilde{t}) = e^{i\hat{H}_0\tilde{t}} \hat{\beta} e^{-i\hat{H}_0\tilde{t}}. \quad (\text{A.5})$$

Second, the solution of the time-dependent Schrödinger equation

$$i \frac{\partial}{\partial t} \Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, t) = \hat{H}(t) \Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, t), \quad (\text{A.6})$$

can be written in terms of the time evolution operator

$$\Psi(t) = \hat{U}(t, t_0) \Psi(t_0). \quad (\text{A.7})$$

In time-dependent perturbation theory, the Hamiltonian can be expressed as:

$$\hat{H}(t) = \hat{H}_0 + \hat{H}_1(t), \quad (\text{A.8})$$

which is used in reference 38 to obtain the first-order approximation to the time evolution operator:

$$\hat{U}(t, t_0) \approx e^{-i\hat{H}_0(t-t_0)} \left\{ 1 - i \int_{t_0}^t dt' F(t') \hat{\beta}(t' - t_0) \right\}. \quad (\text{A.9})$$

Using equation (A.9) for the first-order approximation to the time evolution operator and equation (A.5) for the interaction picture representation of the operators  $\hat{\alpha}$  and  $\hat{\beta}$ , we obtain the linear response as:

$$\alpha_1(t) = -i \int_{t_0}^t dt' F(t') \langle \Psi_0, [\hat{\alpha}(t), \hat{\beta}(t')] \Psi_0 \rangle. \quad (\text{A.10})$$

The initial-state Hamiltonian  $\hat{H}_0$  is time-independent. Thus, the commutator  $[\hat{\alpha}(t), \hat{\beta}(t')]$  can be replaced by  $[\hat{\alpha}(t-t'), \hat{\beta}]$  in equation (A.10). By introducing the step function  $\Theta(t-t')$  (which is 1 for  $t \geq t'$  and 0 otherwise), a retarded response function is defined as follows:

$$\chi_{\alpha\beta}(t-t') = -i\theta(t-t') \langle \Psi_0, [\hat{\alpha}(t-t'), \hat{\beta}] \Psi_0 \rangle. \quad (\text{A.11})$$

In this context, the term ‘‘retarded’’ is used to express the fact that the response at time  $t$  is due to a perturbation at an earlier time  $t' \leq t$ . The linear response  $\alpha_1(t)$  is therefore given by

$$\alpha_1(t) = \int_{-\infty}^{\infty} dt' \chi_{\alpha\beta}(t-t') F(t'), \quad (\text{A.12})$$

where we are allowed to replace the lower integration limit  $t_0$  by  $-\infty$  since the external field  $F(t)$  is zero for all times before  $t_0$ . Please note that the response function does not depend on the time  $t_0$  at which the perturbation is switched on.

### Frequency dependent response

A Fourier transformation of a given function with respect to time  $t$  introduces frequency as a basic variable. We will see that this modification will allow us to obtain excitation energies of a system from its linear response.

We define the Fourier transform of the perturbing field, together with its inverse (for simplicity we do not distinguish between the Fourier transform of a function and the function itself, since the distinction can be inferred from the functional arguments  $t$  and  $\omega$ ):

$$F(t) = \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} F(\omega) e^{-i\omega t}, \quad F(\omega) = \int_{-\infty}^{\infty} dt F(t) e^{i\omega t}, \quad (\text{A.13})$$

and similarly for all other time-dependent quantities. In the next step, we will make use of the relation  $\int_{-\infty}^{\infty} \exp^{it(\omega-\omega')} = 2\pi\delta(\omega-\omega')$ . We insert equation (A.13)

into equation (A.12) and obtain

$$\int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \alpha_1(\omega) e^{-i\omega t} = \int_{-\infty}^{\infty} dt' \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \chi_{\alpha\beta}(\omega) e^{-i\omega(t-t')} \int_{-\infty}^{\infty} \frac{d\omega'}{2\pi} F(\omega') e^{-i\omega' t'} \quad (\text{A.14})$$

$$= \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \chi_{\alpha\beta}(\omega) F(\omega) e^{-i\omega t}. \quad (\text{A.15})$$

A careful comparison of the right- and left-hand sides of the latter equation leads to the following linear response equation in frequency space (since all exponentials  $\exp^{-it\omega}$  are linearly independent):

$$\alpha_1(\omega) = \chi_{\alpha\beta}(\omega) F(\omega) \quad (\text{A.16})$$

A Fourier transformation of equation (A.11) formally yields an expression for the frequency-dependent response function

$$\chi_{\alpha\beta}(\omega) = -i \int_{-\infty}^{\infty} d\tau \Theta(\tau) \langle \Psi_0, [\hat{\alpha}(\tau), \hat{\beta}] \Psi_0 \rangle e^{i\omega\tau} \quad (\text{A.17})$$

The set of (orthonormal) eigenfunctions  $\{\Psi_n | n = 0, 1, 2, \dots\}$ , where  $\Psi_0$  is the many-body ground state, with energy  $E_0$ ,  $\Psi_1$  is the first excited state, with energy  $E_1$ , and so on can be used to obtain the resolution of the identity:

$$\hat{I}d = \sum_{i=1}^{\infty} |\Psi_i\rangle \langle \Psi_i|. \quad (\text{A.18})$$

Furthermore, we denote the  $n$ -th excitation energy of the system as  $\Omega_n := E_n - E_0$ . By inserting the completeness relation (eq. (A.18)) into equation (A.17) and using the explicit form of the interaction representation from equation (A.5) of the operator  $\hat{\alpha}$ , we obtain

$$\chi_{\alpha\beta}(\omega) = -i \sum_{n=1}^{\infty} \int_{-\infty}^{\infty} d\tau \Theta(\tau) e^{i\omega\tau} \left\{ \langle \Psi_0, \hat{\alpha} \Psi_n \rangle \langle \Psi_n, \hat{\beta} \Psi_0 \rangle e^{-i\Omega_n \tau} - \langle \Psi_0, \hat{\beta} \Psi_n \rangle \langle \Psi_n, \hat{\alpha} \Psi_0 \rangle e^{i\Omega_n \tau} \right\}. \quad (\text{A.19})$$

The step function  $\Theta(\tau)$  can be expressed by the following integral representation

$$\Theta(\tau) = \lim_{\eta \rightarrow 0^+} \frac{i}{2\pi} \int_{-\infty}^{\infty} d\omega' \frac{e^{-i\omega' \tau}}{\omega' + i\eta}. \quad (\text{A.20})$$

This relation is obtained by contour integration in the complex frequency plane.

### Lehmann representation

Inserting equation (A.20) into equation (A.19) yields the Lehmann representation of the linear response function

$$\chi_{\alpha\beta}(\omega) = \lim_{\eta \rightarrow 0^+} \sum_{n=1}^{\infty} \left\{ \frac{\langle \Psi_0, \hat{\alpha} \Psi_n \rangle \langle \Psi_n, \hat{\beta} \Psi_0 \rangle}{\omega - \Omega_n + i\eta} - \frac{\langle \Psi_0, \hat{\beta} \Psi_n \rangle \langle \Psi_n, \hat{\alpha} \Psi_0 \rangle}{\omega - \Omega_n + i\eta} \right\}. \quad (\text{A.21})$$

Equation (A.21) is of great importance for the field of response theory because it couples a frequency-dependent perturbation to the excitation spectrum of a system.



# Bibliography

- <sup>1</sup> G. Kabbe. Development of a coupled Molecular Dynamics / Lattice Monte Carlo scheme - [kumulative dissertation], 2018.
- <sup>2</sup> K. Yamada, T. Sagara, Y. Yamane, H. Ohki, and T. Okuda. Superprotonic conductor  $\text{CsH}_2\text{PO}_4$  studied by  $^1\text{H}$ ,  $^{31}\text{P}$  NMR and X-ray diffraction. *Solid State Ion.*, 175(1):557 – 562, 2004.
- <sup>3</sup> A. I. Baranov. Crystals with disordered hydrogen-bond networks and superprotonic conductivity. *Crystallogr. Rep.*, 48:1012–1037, 2003.
- <sup>4</sup> P. A. M. Dirac. Quantum mechanics of many-electron systems. *Proc. R. Soc. Lond. A*, 123(792):714–733, 1929.
- <sup>5</sup> C. Dreßler and D. Sebastiani. Reduced eigensystem representation of the linear density-density response function. *Int. J. Quantum Chem.*, 120(3):e26085, 2020.
- <sup>6</sup> C. Dreßler, A. Scherrer, P. Ahlert, and D. Sebastiani. Efficient representation of the linear density-density response function. *J. Comput. Chem.*, 40(31):2712–2721, 2019.
- <sup>7</sup> P. Ahlert, A. Scherrer, C. Dreßler, and D. Sebastiani. Iterative approach for the moment representation of the density-density response function. *Eur. Phys. J. B*, 91(6):94, 2018.
- <sup>8</sup> A. Scherrer, C. Dreßler, P. Ahlert, and D. Sebastiani. Generalization of the electronic susceptibility for arbitrary molecular geometries. *J. Chem. Phys.*, 144(14):144111, 2016.
- <sup>9</sup> C. Dreßler, G. Kabbe, M. Brehm, and D. Sebastiani. Exploring non-equilibrium molecular dynamics of mobile protons in the solid acid  $\text{CsH}_2\text{PO}_4$  at the micrometer and microsecond scale. *J. Chem. Phys.*, 152(16):164110, 2020.
- <sup>10</sup> C. Dreßler, G. Kabbe, M. Brehm, and D. Sebastiani. Dynamical matrix propagator scheme for large-scale proton dynamics simulations. *J. Chem. Phys.*, 152(11):114114, 2020.
- <sup>11</sup> M. Wagner, C. Dreßler, F. P. Lohmann-Richters, K. Hanus, D. Sebastiani, A. Varga, and B. Abel. Mechanism of ion conductivity through polymer-stabilized  $\text{CsH}_2\text{PO}_4$  nanoparticulate layers from experiment and theory. *J. Mater. Chem. A*, 7:27367–27376, 2019.
- <sup>12</sup> G. Kabbe, C. Dreßler, and D. Sebastiani. Proton mobility in aqueous systems: combining *ab initio* accuracy with millisecond timescales. *Phys. Chem. Chem. Phys.*, 19:28604–28609, 2017.
- <sup>13</sup> G. Kabbe, C. Dreßler, and D. Sebastiani. Toward realistic transfer rates within the coupled Molecular Dynamics/Lattice Monte Carlo approach. *J. Phys. Chem. C*, 120(36):19905–19912, 2016.

- <sup>14</sup> C. Dreßler and D. Sebastiani. Effect of anion reorientation on proton mobility in the solid acids family  $\text{CsH}_y\text{XO}_4$  ( $X = \text{S, P, Se}$ ;  $y = 1, 2$ ) from *ab initio* molecular dynamics simulations. *Phys. Chem. Chem. Phys.*, 22:10738–10752, 2020.
- <sup>15</sup> C. Dreßler, G. Kabbe, and D. Sebastiani. Insight from atomistic simulations of protonation dynamics at the nanoscale. *Fuel Cells*, 16(6):682–694.
- <sup>16</sup> C. Dreßler, G. Kabbe, and D. Sebastiani. Proton conductivity in hydrogen phosphate/sulfates from a coupled Molecular Dynamics/Lattice Monte Carlo (cMD/LMC) approach. *J. Phys. Chem. C*, 120(36):19913–19922, 2016.
- <sup>17</sup> A. D. Buckingham. *Permanent and Induced Molecular Moments and Long-Range Intermolecular Forces*, pages 107–142. John Wiley and Sons, Ltd, 2007.
- <sup>18</sup> A. Stone. Distributed multipole analysis, or how to describe a molecular charge distribution. *Chem. Phys. Lett.*, 83(2):233 – 239, 1981.
- <sup>19</sup> J. Piquemal, N. Gresh, and C. Giessner-Prettre. Improved formulas for the calculation of the electrostatic contribution to the intermolecular interaction energy from multipolar expansion of the electronic distribution. *J. Phys. Chem. A*, 107(48):10353–10359, 2003.
- <sup>20</sup> G. Naray-Szabo and G. G. Ferenczy. Molecular electrostatics. *Chem. Rev.*, 95(4):829–847, 1995.
- <sup>21</sup> J. G. Ángyán, C. Chipot, F. Dehez, C. Hättig, G. Jansen, and C. Millot. Opep: A tool for the optimal partitioning of electric properties. *J. Comput. Chem.*, 24(8):997–1008, 2003.
- <sup>22</sup> D. Elking, T. Darden, and R. J. Woods. Gaussian induced dipole polarization model. *J. Comput. Chem.*, 28(7):1261–1274, 2007.
- <sup>23</sup> R. J. Wheatley. Gaussian multipole functions for describing molecular charge distributions. *Molecular Physics*, 79(3):597–610, 1993.
- <sup>24</sup> G. G. Hall and C. M. Smith. Fitting electron densities of molecules. *Int. J. Quantum Chem.*, 25(5):881–890, 1984.
- <sup>25</sup> D. M. Elking, G. A. Cisneros, J. Piquemal, T. A. Darden, and L. G. Pedersen. Gaussian multipole model (GMM). *J. Chem. Theory Comput.*, 6(1):190–202, 2010.
- <sup>26</sup> K. Eichkorn, O. Treutler, H. Öhm, M. Häser, and R. Ahlrichs. Auxiliary basis sets to approximate Coulomb potentials (Chem. Phys. Letters 240 (1995) 283-290). *Chem. Phys. Lett.*, 242(6):652 – 660, 1995.
- <sup>27</sup> G. A. Cisneros, J. Piquemal, and T. A. Darden. Intermolecular electrostatic energies using density fitting. *J. Chem. Phys.*, 123(4):044109, 2005.
- <sup>28</sup> B. Jeziorski, R. Moszynski, and K. Szalewicz. Perturbation theory approach to intermolecular potential energy surfaces of van der Waals complexes. *Chem. Rev.*, 94(7):1887–1930, 1994.
- <sup>29</sup> A. Heßelmann. DFT-SAPT intermolecular interaction energies employing exact-exchange Kohn–Sham response methods. *J. Chem. Theory Comput.*, 14(4):1943–1959, 2018.



- <sup>30</sup> D. M. Benoit, D. Sebastiani, and M. Parrinello. Accurate total energies without self-consistency. *Phys. Rev. Lett.*, 87:226401, 2001.
- <sup>31</sup> A. Hesselmann and G. Jansen. Intermolecular dispersion energies from time-dependent density functional theory. *Chem. Phys. Lett.*, 367(5):778 – 784, 2003.
- <sup>32</sup> A. J. Misquitta, B. Jeziorski, and K. Szalewicz. Dispersion energy from density-functional theory description of monomers. *Phys. Rev. Lett.*, 91:033201, 2003.
- <sup>33</sup> A. Hesselmann. Comparison of intermolecular interaction energies from SAPT and DFT including empirical dispersion contributions. *J. Phys. Chem. A*, 115(41):11321–11330, 2011.
- <sup>34</sup> A. C. Ihrig, A. Scherrer, and D. Sebastiani. Electronic density response to molecular geometric changes from explicit electronic susceptibility calculations. *J. Chem. Phys.*, 139(9):094102, 2013.
- <sup>35</sup> A. Scherrer, V. Verschinin, and D. Sebastiani. Eigensystem representation of the electronic susceptibility tensor for intermolecular interactions within density functional theory. *J. Chem. Theory Comput.*, 8(1):106–111, 2012.
- <sup>36</sup> H. F. Wilson, F. Gygi, and G. Galli. Efficient iterative method for calculations of dielectric matrices. *Phys. Rev. B*, 78:113303, 2008.
- <sup>37</sup> X. Gonze, D. C. Allan, and M. P. Teter. Dielectric tensor, effective charges, and phonons in  $\alpha$ -quartz by variational density-functional perturbation theory. *Phys. Rev. Lett.*, 68:3603–3606, 1992.
- <sup>38</sup> C. A. Ullrich. *Time-dependent density-functional theory: concepts and applications*. OUP Oxford, 2011.
- <sup>39</sup> M. Marques, A. Rubio, E. K. Gross, K. Burke, F. Nogueira, and C. A. Ullrich. *Time-dependent density functional theory*, volume 706. Springer Science & Business Media, 2006.
- <sup>40</sup> M. A. Marques, N. T. Maitra, F. Nogueira, E. K. Gross, and A. Rubio. *Fundamentals of time-dependent density functional theory*, volume 837. Springer Science & Business Media, 2012.
- <sup>41</sup> K. Chen, J. Hirst, R. Camba, C. A. Bonagura, C. D. Stout, B. K. Burgess, and F. A. Armstrong. Atomically defined mechanism for proton transfer to a buried redox centre in a protein. *Nature*, 405:814–817, 2000.
- <sup>42</sup> M. Tuckerman, K. Laasonen, M. Sprik, and M. Parrinello. *Ab Initio* molecular dynamics simulation of the solvation and transport of hydronium and hydroxyl ions in water. *J. Chem. Phys.*, 103:150–161, 1995.
- <sup>43</sup> C. Fang, R. R. Frontiera, R. Tran, and R. A. Mathies. Mapping GFP structure evolution during proton transfer with femtosecond Raman spectroscopy. *Nature*, 462:200–204, 2009.
- <sup>44</sup> L. Vilčiauskas, M. E. Tuckerman, G. Bester, S. J. Paddison, and K. Kreuer. The mechanism of proton conduction in phosphoric acid. *Nat. Chem.*, 4:461–466, 2012.
- <sup>45</sup> D. Marx, M. E. Tuckerman, J. Hutter, and M. Parrinello. The nature of the hydrated excess proton in water. *Nature*, 397:601–604, 1999.

- <sup>46</sup> B. Prod'hom, D. Pietrobon, and P. Hess. Direct measurement of proton transfer rates to a group controlling the dihydropyridine-sensitive  $\text{Ca}^{2+}$  channel. *Nature*, 329:243–246, 1987.
- <sup>47</sup> K. Kreuer. Proton conductivity: Materials and applications. *Chem. Mat.*, 8:610–641, 1996.
- <sup>48</sup> R. Vuilleumier and D. Borgis. Proton conduction: Hopping along hydrogen bonds. *Nat. Chem.*, 4:432–433, 2012.
- <sup>49</sup> M. E. Tuckerman, D. Marx, and M. Parrinello. The nature and transport mechanism of hydrated hydroxide ions in aqueous solution. *Nature*, 417:925–929, 2002.
- <sup>50</sup> C. A. Wraight. Chance and design—proton transfer in water, channels and bioenergetic proteins. *Biochim. Biophys. Acta*, 1757:886–912, 2006.
- <sup>51</sup> A. Chandra, M. E. Tuckerman, and D. Marx. Connecting solvation shell structure to proton transport kinetics in hydrogen-bonded networks via population correlation functions. *Phys. Rev. Lett.*, 99:145901, 2007.
- <sup>52</sup> J. A. Morrone and M. E. Tuckerman. *Ab initio* molecular dynamics study of proton mobility in liquid methanol. *J. Chem. Phys.*, 117(9):4403–4413, 2002.
- <sup>53</sup> R. Vuilleumier and D. Borgis. Proton conduction: Hopping along hydrogen bonds. *Nat. Chem.*, 4(6):432, 2012.
- <sup>54</sup> J. A. Morrone, K. E. Haslinger, and M. E. Tuckerman. *Ab Initio* molecular dynamics simulation of the structure and proton transport dynamics of methanol–water solutions. *J. Phys. Chem. B*, 110(8):3712–3720, 2006.
- <sup>55</sup> T. C. Berkelbach, H. Lee, and M. E. Tuckerman. Concerted hydrogen-bond dynamics in the transport mechanism of the hydrated proton: A first-principles molecular dynamics study. *Phys. Rev. Lett.*, 103:238302, 2009.
- <sup>56</sup> L. Vilčiauskas, M. E. Tuckerman, G. Bester, S. J. Paddison, and K. Kreuer. The mechanism of proton conduction in phosphoric acid. *Nat. Chem.*, 4(6):461, 2012.
- <sup>57</sup> R. Devanathan, N. Idupulapati, M. D. Baer, C. J. Mundy, and M. Dupuis. *Ab Initio* molecular dynamics simulation of proton hopping in a model polymer membrane. *J. Phys. Chem. B*, 117(51):16522–16529, 2013.
- <sup>58</sup> V. Glezakou, M. Dupuis, and C. J. Mundy. Acid/base equilibria in clusters and their role in proton exchange membranes: computational insight. *Phys. Chem. Chem. Phys.*, 9:5752–5760, 2007.
- <sup>59</sup> M. D. Baer, J. L. Fulton, M. Balasubramanian, G. K. Schenter, and C. J. Mundy. Persistent ion pairing in aqueous hydrochloric acid. *J. Phys. Chem. B*, 118(26):7211–7220, 2014.
- <sup>60</sup> M. D. Baer, D. J. Tobias, and C. J. Mundy. Investigation of interfacial and bulk dissociation of HBr, HCl, and  $\text{HNO}_3$  using density functional theory-based molecular dynamics simulations. *J. Phys. Chem. C*, 118(50):29412–29420, 2014.
- <sup>61</sup> J. Schmidt, J. VandeVondele, I. W. Kuo, D. Sebastiani, J. I. Siepmann, J. Hutter, and C. J. Mundy. Isobaric–isothermal molecular dynamics simulations utilizing density functional theory: An assessment of the structure and density of water at near-ambient conditions. *J. Phys. Chem. B*, 113(35):11959–11964, 2009.

- <sup>62</sup> J. C. Grossman, E. Schwegler, E. W. Draeger, F. Gygi, and G. Galli. Towards an assessment of the accuracy of density functional theory for first principles simulations of water. *J. Chem. Phys.*, 120(1):300–311, 2004.
- <sup>63</sup> E. Schwegler, J. C. Grossman, F. Gygi, and G. Galli. Towards an assessment of the accuracy of density functional theory for first principles simulations of water II. *J. Chem. Phys.*, 121(11):5400–5409, 2004.
- <sup>64</sup> L. Delle Site. What is a multiscale problem in molecular dynamics? *Entropy*, 16(1):23–40, 2014.
- <sup>65</sup> M. Praprotnik, L. Delle Site, and K. Kremer. Multiscale simulation of soft matter: From scale bridging to adaptive resolution. *Annu. Rev. Phys. Chem.*, 59(1):545–571, 2008.
- <sup>66</sup> K. Wendler, F. Dommert, Y. Y. Zhao, R. Berger, C. Holm, and L. Delle Site. Ionic liquids studied across different scales: A computational perspective. *Faraday Discuss.*, 154:111–132, 2012.
- <sup>67</sup> J. Behler. Constructing high-dimensional neural network potentials: A tutorial review. *Int. J. Quantum Chem.*, 115(16):1032–1050, 2015.
- <sup>68</sup> M. Hellström and J. Behler. Proton-transfer-driven water exchange mechanism in the Na<sup>+</sup> solvation shell. *J. Phys. Chem. B*, 121(16):4184–4190, 2017.
- <sup>69</sup> B. Nebgen, N. Lubbers, J. S. Smith, A. E. Sifain, A. Lokhov, O. Isayev, A. E. Roitberg, K. Barros, and S. Tretiak. Transferable dynamic molecular charge assignment using deep neural networks. *J. Chem. Theory Comput.*, 14(9):4687–4698, 2018.
- <sup>70</sup> K. T. Butler, D. W. Davies, H. Cartwright, O. Isayev, and A. Walsh. Machine learning for molecular and materials science. *Nature*, 559(7715):547, 2018.
- <sup>71</sup> M. Hellström and J. Behler. Concentration-dependent proton transfer mechanisms in aqueous NaOH solutions: From acceptor-driven to donor-driven and back. *J. Phys. Chem. Lett.*, 7(17):3302–3306, 2016.
- <sup>72</sup> M. Gastegger, C. Kauffmann, J. Behler, and P. Marquetand. Comparing the accuracy of high-dimensional neural network potentials and the systematic molecular fragmentation method: A benchmark study for all-*trans* alkanes. *J. Chem. Phys.*, 144(19):194110, 2016.
- <sup>73</sup> V. Quaranta, M. Hellström, and J. Behler. Proton-transfer mechanisms at the Water-ZnO interface: The role of presolvation. *J. Phys. Chem. Lett.*, 8(7):1476–1483, 2017.
- <sup>74</sup> J. Chodera, G. Bowman, V. Pande, and Noé F. An introduction to Markov state models and their application to long timescale molecular simulation, vol. 797 of *Advances in Experimental Medicine and Biology*, 2014.
- <sup>75</sup> G. R. Bowman, E. R. Bolin, K. M. Hart, B. C. Maguire, and S. Marqusee. Discovery of multiple hidden allosteric sites by combining Markov state models and experiments. *Proc. Natl. Acad. Sci. U.S.A.*, 112(9):2734–2739, 2015.
- <sup>76</sup> B. E. Husic and V. S. Pande. Markov state models: From an art to a science. *J. Am. Chem. Soc.*, 140(7):2386–2396, 2018.

- <sup>77</sup> V. S. Pande, K. Beauchamp, and G. R. Bowman. Everything you wanted to know about Markov state models but were afraid to ask. *Methods*, 52(1):99 – 105, 2010.
- <sup>78</sup> C. Schütte and M. Sarich. A critical appraisal of Markov state models. *Eur. Phys. J. ST*, 224(12):2445–2462, 2015.
- <sup>79</sup> C. R. Schwantes, R. T. McGibbon, and V. S. Pande. Perspective: Markov models for long-timescale biomolecular dynamics. *J. Chem. Phys.*, 141(9):090901, 2014.
- <sup>80</sup> L. J. LaBerge and J. C. Tully. A rigorous procedure for combining molecular dynamics and Monte Carlo simulation algorithms. *Chem. Phys.*, 260:183–191, 2000.
- <sup>81</sup> E. C. Neyts and A. Bogaerts. Combining molecular dynamics with Monte Carlo simulations: implementations and applications. *Theor. Chem. Acc.*, 132(2):1–12, 2012.
- <sup>82</sup> B. M. Forrest and U. W. Suter. Hybrid Monte Carlo simulations of dense polymer systems. *J. Chem. Phys.*, 101:2616–2629, 1994.
- <sup>83</sup> D. G. Gromov and J. J. de Pablo. Structure of binary polymer blends: Multiple time step hybrid Monte Carlo simulations and self-consistent integral-equation theory. *J. Chem. Phys.*, 103:8247–8256, 1995.
- <sup>84</sup> A. Irback. Hybrid Monte Carlo simulation of polymer chains. *J. Chem. Phys.*, 101:1661–1667, 1994.
- <sup>85</sup> D. W. Heermann and L. Yixue. A global-update simulation method for polymer systems. *Macromol. Chem. Phys.*, 2(3):299–308, 1993.
- <sup>86</sup> I. Martin-Bragado, R. Borges, J. P. Balbuena, and M. Jaraiz. Kinetic Monte Carlo simulation for semiconductor processing: A review. *Prog. Mater. Sci.*, 92:1 – 32, 2018.
- <sup>87</sup> G. Betz and W. Husinsky. A combined molecular dynamics and kinetic Monte Carlo calculation to study sputter erosion and beam assisted deposition. *Nucl. Instr. Meth. Phys. Res. B*, 193:352–358, 2002.
- <sup>88</sup> A. Ghoufi and G. Maurin. Hybrid Monte Carlo simulations combined with a phase mixture model to predict the structural transitions of a porous metal-organic framework material upon adsorption of guest molecules. *J. Phys. Chem. C*, 114(14):6496–6502, 2010.
- <sup>89</sup> A. A. Knizhnik, A. A. Bagaturyants, I. V. Belov, B. V. Potapkin, and A. A. Korokin. An integrated kinetic Monte Carlo molecular dynamics approach for film growth modeling and simulation: ZrO<sub>2</sub> deposition on Si(100) surface. *Comput. Mater. Sci.*, 24(1-2):128–132, 2002.
- <sup>90</sup> U. H. E. Hansmann and Y. Okamoto. New Monte Carlo algorithms for protein folding. *Curr. Opin. Struct. Biol.*, 9(2):177–183, 1999.
- <sup>91</sup> E. K. Peter and J. Shea. A hybrid MD-kMC algorithm for folding proteins in explicit solvent. *Phys. Chem. Chem. Phys.*, 16(14):6430–6440, 2014.
- <sup>92</sup> H. Zhang. A new hybrid Monte Carlo algorithm for protein potential function test and structure refinement. *Proteins: Struct., Funct., Genet.*, 34(4):464–471, 1999.

- <sup>93</sup> F. Noé and E. Rosta. Markov models of molecular kinetics. *J. Chem. Phys.*, 151(19):190401, 2019.
- <sup>94</sup> M. Dibak, M. J. del Razo, D. De Sancho, C. Schütte, and F. Noé. MSM/RD: Coupling Markov state models of molecular kinetics with reaction-diffusion simulations. *J. Chem. Phys.*, 148(21):214107, 2018.
- <sup>95</sup> B. G. Keller, J. Prinz, and F. Noé. Markov models and dynamical fingerprints: Unraveling the complexity of molecular kinetics. *Chem. Phys.*, 396:92 – 107, 2012.
- <sup>96</sup> B. G. Keller, A. Kobitski, A. Jäschke, G. U. Nienhaus, and F. Noé. Complex RNA folding kinetics revealed by single-molecule FRET and hidden Markov models. *J. Am. Chem. Soc.*, 136(12):4534–4543, 2014.
- <sup>97</sup> F. Nüske, B. G. Keller, G. Pérez-Hernández, A. S. J. S. Mey, and F. Noé. Variational approach to molecular kinetics. *J. Chem. Theory Comput.*, 10(4):1739–1752, 2014.
- <sup>98</sup> F. Nüske, H. Wu, J. Prinz, C. Wehmeyer, C. Clementi, and F. Noé. Markov state models from short non-equilibrium simulations – analysis and correction of estimation bias. *J. Chem. Phys.*, 146(9):094104, 2017.
- <sup>99</sup> J. Prinz, H. Wu, M. Sarich, B. Keller, M. Senne, M. Held, J. D. Chodera, C. Schütte, and Noé F. Markov models of molecular kinetics: Generation and validation. *J. Chem. Phys.*, 134(17):174105, 2011.
- <sup>100</sup> C. Schütte, F. Noé, J. Lu, M. Sarich, and E. Vanden-Eijnden. Markov state models based on milestoning. *J. Chem. Phys.*, 134(20):204105, 2011.
- <sup>101</sup> B. Trendelkamp-Schroer, H. Wu, F. Paul, and F. Noé. Estimation and uncertainty of reversible Markov models. *J. Chem. Phys.*, 143(17):174101, 2015.
- <sup>102</sup> F. Vitalini, A. S. J. S. Mey, F. Noé, and B. G. Keller. Dynamic properties of force fields. *J. Chem. Phys.*, 142(8):084101, 2015.
- <sup>103</sup> H. Wu, F. Paul, C. Wehmeyer, and F. Noé. Multiensemble Markov models of molecular thermodynamics and kinetics. *Proc. Natl. Acad. Sci. U.S.A.*, 2016.
- <sup>104</sup> E. Hruska, J. R. Abella, F. Nüske, L. E. Kaviraki, and C. Clementi. Quantitative comparison of adaptive sampling methods for protein dynamics. *J. Chem. Phys.*, 149(24):244119, 2018.
- <sup>105</sup> U. Sengupta, M. Carballo-Pacheco, and B. Strodel. Automated Markov state models for molecular dynamics simulations of aggregation and self-assembly. *J. Chem. Phys.*, 150(11):115101, 2019.
- <sup>106</sup> F. Noé, C. Schütte, E. Vanden-Eijnden, L. Reich, and T. R. Weigl. Constructing the equilibrium ensemble of folding pathways from short off-equilibrium simulations. *Proc. Natl. Acad. Sci. U.S.A.*, 106(45):19011–19016, 2009.
- <sup>107</sup> T. J. Lane, G. R. Bowman, K. Beauchamp, V. A. Voelz, and V. S. Pande. Markov state model reveals folding and functional dynamics in ultra-long md trajectories. *J. Am. Chem. Soc.*, 133(45):18413–18419, 2011.
- <sup>108</sup> G. Pérez-Hernández, F. Paul, T. Giorgino, G. De Fabritiis, and F. Noé. Identification of slow molecular order parameters for Markov model construction. *J. Chem. Phys.*, 139(1):015102, 2013.

- <sup>109</sup> N. Stanley and Esteban-Martí. Kinetic modulation of a disordered protein domain by phosphorylation.
- <sup>110</sup> I. Buch, T. Giorgino, and G. De Fabritiis. Complete reconstruction of an enzyme-inhibitor binding process by molecular dynamics simulations. *Proc. Natl. Acad. Sci. U.S.A.*, 108(25):10184–10189, 2011.
- <sup>111</sup> M. Held, P. Metzner, J. Prinz, and F. Noé. Mechanisms of protein-ligand association and its modulation by protein mutations. *Biophys. J.*, 100(3):701 – 710, 2011.
- <sup>112</sup> D. Silva, G. R. Bowman, A. Sosa-Peinado, and X. Huang. A role for both conformational selection and induced fit in ligand binding by the LAO protein. *PLOS Comput. Biol.*, 7(5):1–11, 2011.
- <sup>113</sup> G. R. Bowman and P. L. Geissler. Equilibrium fluctuations of a single folded protein reveal a multitude of potential cryptic allosteric sites. *Proc. Natl. Acad. Sci. U.S.A.*, 109(29):11681–11686, 2012.
- <sup>114</sup> N. Plattner and F. Noé. Protein conformational plasticity and complex ligand-binding kinetics explored by atomistic simulations and Markov models. *Nat. comm.*, 6:7653, 2015.
- <sup>115</sup> D. Chakraborty and D. J. Wales. Dynamics of an adenine-adenine RNA conformational switch from discrete path sampling. *J. Chem. Phys.*, 150(12):125101, 2019.
- <sup>116</sup> G. Pinamonti, F. Paul, F. Noé, A. Rodriguez, and G. Bussi. The mechanism of RNA base fraying: Molecular dynamics simulations analyzed with core-set Markov state models. *J. Chem. Phys.*, 150(15):154123, 2019.
- <sup>117</sup> H. Lee and M. E. Tuckerman. The structure and proton transport mechanisms in the superprotonic phase of CsH<sub>2</sub>PO<sub>4</sub>: An *ab initio* Molecular Dynamics study. *J. Phys. Chem. C*, 112:9917–9930, 2008.
- <sup>118</sup> B. C. Wood and N. Marzari. Proton dynamics in superprotonic CsHSO<sub>4</sub>. *Phys. Rev. B*, 76:134301, 2007.
- <sup>119</sup> R. B. Merle, C. R. I. Chisholm, D. A. Boysen, and S. M. Haile. Instability of sulfate and selenate solid acids in fuel cell environments. *Energy Fuels*, 17:210–215, 2003.
- <sup>120</sup> S. M. Haile, D. A. Boysen, C. R. Chisholm, and R. B. Merle. Solid acids as fuel cell electrolytes. *Nature*, 410:910–913, 2001.
- <sup>121</sup> C. R. I. Chisholm, D. A. Boysen, A. B. Papandrew, S. Zecevic, S. Cha, K. A. Sasaki, A. Varga, K. P. Giapis, and S. M. Haile. From laboratory breakthrough to technological realization: The development path for solid acid fuel cells. *J. Electrochem. Soc. Interface*, 18:53–59, 2009.
- <sup>122</sup> D. A. Boysen, T. Uda, C. R. I. Chisholm, and S. M. Haile. High-performance solid acid fuel cells through humidity stabilization. *Science*, 303:68–70, 2004.
- <sup>123</sup> <http://www.safcell.com/>, (accessed: 14.07.2020).
- <sup>124</sup> T. Schlick. *Molecular modeling and simulation: an interdisciplinary guide*, volume 21. Springer Science & Business Media, 2010.

- <sup>125</sup> M. De Vivo, M. Masetti, G. Bottegoni, and A. Cavalli. Role of molecular dynamics and related methods in drug discovery. *J. Med. Chem.*, 59(9):4035–4061, 2016.
- <sup>126</sup> L. Verlet. Computer "experiments" on classical fluids. II. equilibrium correlation functions. *Phys. Rev.*, 165:201–214, 1968.
- <sup>127</sup> G. J. Martyna, M. L. Klein, and M. Tuckerman. Nosé–hoover chains: The canonical ensemble via continuous dynamics. *J. Chem. Phys.*, 97(4):2635–2643, 1992.
- <sup>128</sup> A. Szabo and N. S. Ostlund. *Modern quantum chemistry: introduction to advanced electronic structure theory*. Courier Corporation, 2012.
- <sup>129</sup> P. Hohenberg and W. Kohn. Inhomogeneous electron gas. *Phys. Rev.*, 136:B864–B871, 1964.
- <sup>130</sup> W. Ritz. Über eine neue Methode zur Lösung gewisser Variationsprobleme der mathematischen Physik. *J. Reine Angew. Math.*, 1909(135):1–61, 1909.
- <sup>131</sup> S. Gasiorowicz. *Quantum physics*. John Wiley & Sons, 2007.
- <sup>132</sup> M. Levy. Universal variational functionals of electron densities, first-order density matrices, and natural spin-orbitals and solution of the v-representability problem. *Proc. Natl. Acad. Sci. U.S.A.*, 76(12):6062–6065, 1979.
- <sup>133</sup> E. H. Lieb. Density functionals for coulomb systems. 24(3):243–277, 1983.
- <sup>134</sup> R. M. Martin. *Electronic structure: basic theory and practical methods*. Cambridge university press, 2004.
- <sup>135</sup> J. P. Perdew and K. Schmidt. Jacob’s ladder of density functional approximations for the exchange-correlation energy. *AIP Conference Proceedings*, 577(1):1–20, 2001.
- <sup>136</sup> S. Baroni, S. de Gironcoli, A. Dal Corso, and P. Giannozzi. Phonons and related crystal properties from density-functional perturbation theory. *Rev. Mod. Phys.*, 73:515–562, 2001.
- <sup>137</sup> X. Gonze. Perturbation expansion of variational principles at arbitrary order. *Phys. Rev. A*, 52:1086–1095, 1995.
- <sup>138</sup> X. Gonze. Adiabatic density-functional perturbation theory. *Phys. Rev. A*, 52:1096–1114, 1995.
- <sup>139</sup> A. Putrino, D. Sebastiani, and M. Parrinello. Generalized variational density functional perturbation theory. *J. Chem. Phys.*, 113(17):7102–7109, 2000.
- <sup>140</sup> A. C. Ihrig, A. Scherrer, and D. Sebastiani. Electronic density response to molecular geometric changes from explicit electronic susceptibility calculations. *J. Chem. Phys.*, 139(9):094102, 2013.
- <sup>141</sup> A. Scherrer, V. Verschinin, and D. Sebastiani. Eigensystem representation of the electronic susceptibility tensor for intermolecular interactions within density functional theory. *J. Chem. Theory Comput.*, 8(1):106–111, 2012.
- <sup>142</sup> H. F. Wilson, D. Lu, F. Gygi, and G. Galli. Iterative calculations of dielectric eigenvalue spectra. *Phys. Rev. B*, 79:245106, 2009.



- <sup>143</sup> J. D. Jackson. *Classical electrodynamics*. John Wiley & Sons, 2007.
- <sup>144</sup> A. Scherrer and D. Sebastiani. Moment expansion of the linear density-density response function. *J. Comput. Chem.*, 37(7):665–674, 2016.
- <sup>145</sup> A. Heßelmann and G. Jansen. Intermolecular dispersion energies from time-dependent density functional theory. *Chem. Phys. Lett.*, 367(5–6):778–784, 2003.
- <sup>146</sup> K. Eichkorn, O. Treutler, H. Öhm, M. Häser, and R. Ahlrichs. Auxiliary basis sets to approximate Coulomb potentials. *Chem. Phys. Lett.*, 240:283–290, 1995.
- <sup>147</sup> A. J. Misquitta and A. J. Stone. Distributed polarizabilities obtained using a constrained density-fitting algorithm. *J. Chem. Phys.*, 124(2):024111, 2006.
- <sup>148</sup> A. Heßelmann, G. Jansen, and M. Schütz. Density-functional theory-symmetry-adapted intermolecular perturbation theory with density fitting: A new efficient method to study intermolecular interaction energies. *J. Chem. Phys.*, 122(1):014103, 2005.
- <sup>149</sup> R. Bukowski, R. Podeszwa, and K. Szalewicz. Efficient calculation of coupled Kohn–Sham dynamic susceptibility functions and dispersion energies with density fitting. *Chem. Phys. Lett.*, 414(1–3):111 – 116, 2005.
- <sup>150</sup> H. Nguyen, T. A. Pham, D. Rocca, and G. Galli. Improving accuracy and efficiency of calculations of photoemission spectra within the many-body perturbation theory. *Phys. Rev. B*, 85:81101, 2012.
- <sup>151</sup> T. A. Pham, H. Nguyen, D. Rocca, and G. Galli. GW calculations using the spectral decomposition of the dielectric matrix: Verification, validation, and comparison of methods. *Phys. Rev. B*, 87:155148, 2013.
- <sup>152</sup> F. Caruso, P. Rinke, X. Ren, A. Rubio, and M. Scheffler. Self-consistent GW: All-electron implementation with localized basis functions. *Phys. Rev. B*, 88:075105, 2013.
- <sup>153</sup> M. Fuchs and X. Gonze. Accurate density functionals: Approaches using the adiabatic-connection fluctuation-dissipation theorem. *Phys. Rev. B*, 65:235109, 2002.
- <sup>154</sup> F. Furche and T. Van Voorhis. Fluctuation-dissipation theorem density-functional theory. *J. Chem. Phys.*, 122(16):164106, 2005.
- <sup>155</sup> J. Toulouse, I. C. Gerber, G. Jansen, A. Savin, and J. G. Ángyán. Adiabatic-connection fluctuation-dissipation density-functional theory based on range separation. *Phys. Rev. Lett.*, 102:096404, 2009.
- <sup>156</sup> J. F. Dobson and J. Wang. Successful test of a seamless van der Waals density functional. *Phys. Rev. Lett.*, 82:2123–2126, 1999.
- <sup>157</sup> M. Dion, H. Rydberg, E. Schröder, D. C. Langreth, and B. I. Lundqvist. Van der Waals density functional for general geometries. *Phys. Rev. Lett.*, 92:246401, 2004.
- <sup>158</sup> D. Lu, Y. Li, D. Rocca, and G. Galli. *Ab initio* calculation of van der Waals bonded molecular crystals. *Phys. Rev. Lett.*, 102:206411, 2009.
- <sup>159</sup> J. F. Dobson and T. Gould. Calculation of dispersion energies. *J. Phys. Condens. Matter*, 24(7):073201, 2012.

- <sup>160</sup> F. Furche. Molecular tests of the random phase approximation to the exchange-correlation energy functional. *Phys. Rev. B*, 64:195120, 2001.
- <sup>161</sup> F. Furche. Developing the random phase approximation into a practical post-Kohn-Sham correlation model. *J. Chem. Phys.*, 129(11):114105, 2008.
- <sup>162</sup> X. Ren, P. Rinke, C. Joas, and M. Scheffler. Random-phase approximation and its applications in computational chemistry and materials science. *J. Mater. Sci.*, 47(21):7447–7471, 2012.
- <sup>163</sup> D. Rocca. Random-phase approximation correlation energies from Lanczos chains and an optimal basis set: theory and applications to the benzene dimer. *J. Chem. Phys.*, 140(18):18A501, 2014.
- <sup>164</sup> X. Ren, P. Rinke, G. E. Scuseria, and M. Scheffler. Renormalized second-order perturbation theory for the electron correlation energy: Concept, implementation, and benchmarks. *Phys. Rev. B*, 88:035120, 2013.
- <sup>165</sup> J. E. Bates and F. Furche. Communication: Random phase approximation renormalized many-body perturbation theory. *J. Chem. Phys.*, 139(17):171103, 2013.
- <sup>166</sup> N. Colonna, M. Hellgren, and S. de Gironcoli. Correlation energy within exact-exchange adiabatic connection fluctuation-dissipation theory: Systematic development and simple approximations. *Phys. Rev. B*, 90(12), 2014.
- <sup>167</sup> B. Mussard, D. Rocca, G. Jansen, and J. G. Ángyán. Dielectric matrix formulation of correlation energies in the random phase approximation: Inclusion of exchange effects. *J. Chem. Theory Comput.*, 12(5):2191–2202, 2016.
- <sup>168</sup> H. Hu and W. Yang. Free energies of chemical reactions in solution and in enzymes with *ab initio* quantum mechanics/molecular mechanics methods. *Ann. Rev. Phys. Chem.*, 59(1):573–601, 2008.
- <sup>169</sup> P. Pulay and T. Janowski. Efficient calculation of the energy of a molecule in an arbitrary electric field. *Int. J. Quant. Chem.*, 109(10):2113–2120, 2009.
- <sup>170</sup> T. Janowski, K. Wolinski, and P. Pulay. Ultrafast quantum mechanics/molecular mechanics Monte Carlo simulations using generalized multipole polarizabilities. *Chem. Phys. Lett.*, 530:1–9, 2012.
- <sup>171</sup> N. Sablon, F. De Proft, M. Sola, and P. Geerlings. The linear response kernel of conceptual DFT as a measure of aromaticity. *Phys. Chem. Chem. Phys.*, 14:3960–3967, 2012.
- <sup>172</sup> N. Sablon, F. D. Proft, and P. Geerlings. The linear response kernel: Inductive and resonance effects quantified. *J. Phys. Chem. Lett.*, 1(8):1228–1234, 2010.
- <sup>173</sup> N. Sablon, F. D. Proft, and P. Geerlings. The linear response kernel of conceptual DFT as a measure of electron delocalisation. *Chem. Phys. Lett.*, 498(1–3):192–197, 2010.
- <sup>174</sup> P. Geerlings and F. De Proft. Conceptual DFT: the chemical relevance of higher response functions. *Phys. Chem. Chem. Phys.*, 10:3028–3042, 2008.
- <sup>175</sup> P. Geerlings, S. Fias, Z. Boisdenghien, and F. De Proft. Conceptual DFT: chemistry from the linear response function. *Chem. Soc. Rev.*, 43:4989–5008, 2014.

- <sup>176</sup> M. to Baben, J. O. Achenbach, and O. A. von Lilienfeld. Guiding *ab initio* calculations by alchemical derivatives. *J. Chem. Phys.*, 144(10):104103, 2016.
- <sup>177</sup> R. Balawender, M. A. Welearegay, M. Lesiuk, F. D. Proft, and P. Geerlings. Exploring chemical space with the alchemical derivatives. *J. Chem. Theory Comput.*, 9(12):5327–5340, 2013.
- <sup>178</sup> K. Y. S. Chang, S. Fias, R. Ramakrishnan, and O. A. von Lilienfeld. Fast and accurate predictions of covalent bonds in chemical space. *J. Chem. Phys.*, 144(17):174110, 2016.
- <sup>179</sup> S. Hamel, A. J. Williamson, H. F. Wilson, F. Gygi, G. Galli, E. Ratner, and D. Wack. First-principles calculations of the dielectric properties of silicon nanostructures. *Appl. Phys. Lett.*, 92:3115, 2008.
- <sup>180</sup> A. Savin, F. Colonna, and M. Allavena. Analysis of the linear response function along the adiabatic connection from the Kohn-Sham to the correlated system. *J. Chem. Phys.*, 115(15):6827–6833, 2001.
- <sup>181</sup> Z. Boisdenghien, C. V. Alsenoy, F. D. Proft, and P. Geerlings. Evaluating and interpreting the chemical relevance of the linear response kernel for atoms. *J. Chem. Theory Comput.*, 9(2):1007–1015, 2013.
- <sup>182</sup> D. Werner. *Funktionalanalysis*. Springer, 2006.
- <sup>183</sup> W. Kabbalo. *Grundkurs Funktionalanalysis*. Springer-Verlag, 2018.
- <sup>184</sup> D. Rocca, R. Gebauer, Y. Saad, and S. Baroni. Turbo charging time-dependent density-functional theory with Lanczos chains. *J. Chem. Phys.*, 128(15):154105, 2008.
- <sup>185</sup> F. Kaoui and D. Rocca. Random phase approximation correlation energy using a compact representation for linear response functions: application to solids. *J. Phys. Condens. Matter*, 28(3):035201, 2016.
- <sup>186</sup> S. P. Meyn and R. L. Tweedie. *Markov chains and stochastic stability*. Springer Science & Business Media, 2012.
- <sup>187</sup> F. Baudoin. Stochastic processes. In Penelope Peterson, Eva Baker, and Barry McGaw, editors, *International Encyclopedia of Education (Third Edition)*, pages 451 – 452. Elsevier, Oxford, 2010.
- <sup>188</sup> T. Müller-Gronbach, E. Novak, and K. Ritter. *Monte Carlo-Algorithmen*. Springer-Verlag, 2012.
- <sup>189</sup> B. Derrida. An exactly soluble non-equilibrium system: The asymmetric simple exclusion process. *Phys. Rep.*, 301(1):65 – 83, 1998.
- <sup>190</sup> H. Yau. Law of the two dimensional asymmetric simple exclusion process. *Ann. Math.*, pages 377–405, 2004.
- <sup>191</sup> C. T. MacDonald, J. H. Gibbs, and A. C. Pipkin. Kinetics of biopolymerization on nucleic acid templates. *Biopolymers*, 6(1):1–25, 1968.
- <sup>192</sup> F. Spitzer. Interaction of Markov processes. *Adv. Math.*, 5(2):246 – 290, 1970.
- <sup>193</sup> F. Jensen. *Introduction to computational chemistry*. John Wiley & Sons, 2017.

- <sup>194</sup> A. Goñi-Urtiaga, D. Presvytes, and K. Scott. Solid acids as electrolyte materials for proton exchange membrane (PEM) electrolysis: Review. *Int. J. Hydrog. Energy*, 37(4):3358 – 3372, 2012.
- <sup>195</sup> G. Y. Foran, D. H. Brouwer, and G. R. Goward. Quantifying site-specific proton dynamics in phosphate solid acids by <sup>1</sup>H double quantum NMR spectroscopy. *J. Phys. Chem. C*, 121(46):25641–25650, 2017.
- <sup>196</sup> V. Ponomareva, K. Kovalenko, A. Chupakhin, E. Shutova, and V. Fedin. CsHSO<sub>4</sub> – proton conduction in a crystalline metal–organic framework. *Solid State Ion.*, 225:420 – 423, 2012.
- <sup>197</sup> A. B. Papandrew, R. W. Atkinson, R. R. Unocic, and T. A. Zawodzinski. Ruthenium as a co-tolerant hydrogen oxidation catalyst for solid acid fuel cells. *J. Mater. Chem. A*, 3:3984–3987, 2015.
- <sup>198</sup> L. Liu, H. Li, X. Chen, and X. Lei. Electrolyte membranes based on molten KH<sub>5</sub>(PO<sub>4</sub>)<sub>2</sub> for intermediate temperature fuel cells. *Fuel Cells*, 19(3):280–288, 2019.
- <sup>199</sup> J. Leal, H. Martinez, I. Martinez, A. Price, A. Goos, and C. Botez. Stability of the superprotonic conduction of (1-x) CsH<sub>2</sub>PO<sub>4</sub>/x SiO<sub>2</sub> (0 ≤ x ≤ 0.3) composites under dry and humid environments. *Mater. Today*, 15:11 – 17, 2018.
- <sup>200</sup> I. N. Bagryantseva, V. G. Ponomareva, and N. P. Lazareva. Proton-conductive membranes based on CsH<sub>2</sub>PO<sub>4</sub> and ultra-dispersed polytetrafluoroethylene. *Solid State Ion.*, 329:61 – 66, 2019.
- <sup>201</sup> H. Takahashi, Y. Suzuki, and T. Sakuma. New phase transition in superprotonic phase of inorganic solid acid Cs<sub>2</sub>(HSO<sub>4</sub>)(H<sub>2</sub>PO<sub>4</sub>). *Solid State Ion.*, 285:155 – 159, 2016.
- <sup>202</sup> N. Mohammad, A. B. Mohamad, A. A. H. Kadhun, and K. S. Loh. A review on synthesis and characterization of solid acid materials for fuel cell applications. *J. Power Sources*, 322:77 – 92, 2016.
- <sup>203</sup> F. P. Lohmann-Richters, B. Abel, and A. Varga. In situ determination of the electrochemically active platinum surface area: key to improvement of solid acid fuel cells. *J. Mater. Chem. A*, 6:2700–2707, 2018.
- <sup>204</sup> A. Baranov, L. Shuvalov, and N. Shchagina. Superior conductivity and phase transitions in CsHSO<sub>4</sub> and CsHSeO<sub>4</sub> crystals. *JETP lett.*, 36(11):459–462, 1982.
- <sup>205</sup> S. M. Haile, C. R. I. Chisholm, K. Sasaki, D. A. Boysen, and T. Uda. Solid acid proton conductors: from laboratory curiosities to fuel cell electrolytes. *Faraday Discuss.*, 134:17–39, 2007.
- <sup>206</sup> Q. Li, R. He, J. O. Jensen, and N. J. Bjerrum. Approaches and recent development of polymer electrolyte membranes for fuel cells operating above 100 °C. *Chem. Mater.*, 15(26):4896–4915, 2003.
- <sup>207</sup> J. Park, C. Kim, B. Choi, B. K. Moon, and H. Seo. Physical properties of CsH<sub>2</sub>PO<sub>4</sub> crystal at high temperatures. *J. Phys. Soc. Jpn.*, 72(6):1592–1593, 2003.
- <sup>208</sup> E. Ortiz, R. Vargas, and B. Mellander. On the high-temperature phase transitions of some KDP-family compounds: a structural phase transition? a transition to a bulk-high proton conducting phase? *Solid State Ion.*, 125(1):177 – 185, 1999.

- <sup>209</sup> E. Ortiz, R. A. Vargas, and B. Mellander. On the high-temperature phase transitions of  $\text{CsH}_2\text{PO}_4$ : A polymorphic transition? a transition to a superprotonic conducting phase? *J. Chem. Phys.*, 110(10):4847–4853, 1999.
- <sup>210</sup> K. Yamada, T. Sagara, Y. Yamane, H. Ohki, and T. Okuda. Superprotonic conductor  $\text{CsH}_2\text{PO}_4$  studied by  $^1\text{H}$ ,  $^{31}\text{P}$  NMR and X-ray diffraction. *Solid State Ion.*, 175(1):557 – 562, 2004.
- <sup>211</sup> J. Otomo, T. Tamaki, S. Nishida, S. Wang, M. Ogura, T. Kobayashi, C. Wen, H. Nagamoto, and H. Takahashi. Effect of water vapor on proton conduction of cesium dihydrogen phosphate and application to intermediate temperature fuel cells. *J. Appl. Electrochem.*, 35(9):865–870, 2005.
- <sup>212</sup> A. Baranov, V. Grebenev, A. Khodan, V. Dolbinina, and E. Efremova. Optimization of superprotonic acid salts for fuel cell applications. *Solid State Ion.*, 176(39):2871 – 2874, 2005.
- <sup>213</sup> J. Otomo, N. Minagawa, C. Wen, K. Eguchi, and H. Takahashi. Protonic conduction of  $\text{CsH}_2\text{PO}_4$  and its composite with silica in dry and humid atmospheres. *Solid State Ion.*, 156(3):357 – 369, 2003.
- <sup>214</sup> S. M. Haile, H. Liu, and R. A. Secco. High-temperature behavior of  $\text{CsH}_2\text{PO}_4$  under both ambient and high pressure conditions. *Chem. Mater.*, 15(3):727–736, 2003.
- <sup>215</sup> A. Baranov, B. Merinov, A. Tregubchenko, V. Khiznichenko, L. Shuvalov, and N. Schagina. Fast proton transport in crystals with a dynamically disordered hydrogen bond network. *Solid State Ion.*, 36(3):279 – 282, 1989.
- <sup>216</sup> K. Kreuer. Fast proton conductivity: A phenomenon between the solid and the liquid state? *Solid State Ion.*, 94(1):55 – 62, 1997.
- <sup>217</sup> P. Colomban and C. Philippe. *Proton Conductors: Solids, membranes and gels-materials and devices*, volume 2. Cambridge University Press, 1992.
- <sup>218</sup> A. Preisinger, K. Mereiter, and W. Bronowska. The phase transition of  $\text{CsH}_2\text{PO}_4$  (CDP) at 505 K. In *European Powder Diffraction 3*, volume 166 of *Materials Science Forum*, pages 511–516. Trans Tech Publications Ltd, 1994.
- <sup>219</sup> G. Kim, J. M. Griffin, F. Blanc, S. M. Haile, and C. P. Grey. Characterization of the dynamics in the protonic conductor  $\text{CsH}_2\text{PO}_4$  by  $^{17}\text{O}$  solid-state NMR spectroscopy and first-principles calculations: Correlating phosphate and protonic motion. *J. Am. Chem. Soc.*, 137(11):3867–3876, 2015.
- <sup>220</sup> G. Kim, F. Blanc, Y. Hu, and C. P. Grey. Understanding the conduction mechanism of the protonic conductor  $\text{CsH}_2\text{PO}_4$  by solid-state NMR spectroscopy. *J. Phys. Chem. C*, 117(13):6504–6515, 2013.
- <sup>221</sup> A. M. Balagurov, A. V. Belushkin, I. D. Dutt, I. Natkaniec, N. M. Plakida, B. N. Savenko, L. A. Shuvalov, and J. Wasicki. Neutron scattering studies on structural phase transitions of superionic conductor  $\text{CsHSO}_4$ . *Ferroelectrics*, 63(1):59–67, 1985.
- <sup>222</sup> A. V. Belushkin, I. Natkaniec, N. M. Pakida, L. A. Shuvalov, and J. Wasicki. Neutron scattering studies of vibrational spectra and structural transformations in the superionic conductors  $\text{CsHSO}_4$  and  $\text{CsHSeO}_4$ . *J. Phys. C Solid State Phys*, 20(5):671–687, 1987.

- <sup>223</sup> P. Colomban, M. Pham-Thi, and A. Novak. Thermal history and phase transitions in the superionic protonic conductors CsHSO<sub>4</sub> and CsHSeO<sub>4</sub>. *Solid State Ion.*, 20(2):125 – 134, 1986.
- <sup>224</sup> Z. Jiráak, M. Dlouhá, S. Vratislav, A. M. Balagurov, A. I. Beskrovnyi, V. I. Gordelii, I. D. Datt, and L. A. Shwalov. A neutron diffraction study of the superionic phase in CsHSO<sub>4</sub>. *Phys. Status Solidi A*, 100(2):K117–K122, 1987.
- <sup>225</sup> B. Merinov, A. Baranov, L. Shuvalov, and B. Maksimov. Crystal structure of superionic phase CsDSO<sub>4</sub> and phase transitions in cesium hydro- and deuteriosulfates (selenates). *Kristallografiya*, 32(1):86–92, 1987.
- <sup>226</sup> A. V. Belushkin, W. I. F. David, R. M. Ibberson, and L. A. Shuvalov. High-resolution neutron powder diffraction studies of the structure of CsDSO<sub>4</sub>. *Acta Crystallogr. B*, 47(2):161–166, 1991.
- <sup>227</sup> D. A. Boysen, S. M. Haile, H. Liu, and R. A. Secco. High-temperature behavior of CsH<sub>2</sub>PO<sub>4</sub> under both ambient and high pressure conditions. *Chem. Mat.*, 15:727–736, 2003.
- <sup>228</sup> B. Van Troeye, M. J. van Setten, M. Giantomassi, M. Torrent, G. Rignanese, and X. Gonze. First-principles study of paraelectric and ferroelectric CsH<sub>2</sub>PO<sub>4</sub> including dispersion forces: Stability and related vibrational, dielectric, and elastic properties. *Phys. Rev. B*, 95:024112, 2017.
- <sup>229</sup> R. Blinc, J. Dolinsek, G. Lahajnar, I. Zupancic, L. A. Shuvalov, and A. I. Baranov. Spin-lattice relaxation and self-diffusion study of the protonic superionic conductors CsHSeO<sub>4</sub> and CsHSO<sub>4</sub>. *Phys. Status Solidi B*, 123(1):K83–K87, 1984.
- <sup>230</sup> W. Munch, K. D. Kreuer, U. Traub, and J. Maier. Proton transfer in the three-dimensional hydrogen bond network of the high temperature phase of CsHSO<sub>4</sub>: a molecular dynamics study. *J. Mol. Struct.*, 381:1–8, 1996.
- <sup>231</sup> W. Munch and K. D. Kreuer. A molecular dynamics study of the high proton conducting phase of CsHSO<sub>4</sub>. *Solid State Ion.*, 77:10–14, 1995.
- <sup>232</sup> C. R. I. Chisholm, Y. H. Jang, S. M. Haile, and W. A. Goddard. Superprotonic phase transition of CsHSO<sub>4</sub>: A molecular dynamics simulation study. *Phys. Rev. B*, 72:134103, 2005.
- <sup>233</sup> P. Zetterström, A. Belushkin, R. McGreevy, and L. Shuvalov. Structure and proton conduction in CsDSO<sub>4</sub>. *Solid State Ion.*, 116(3):321 – 329, 1999.
- <sup>234</sup> M. Krzystyniak, K. Druzbicki, and F. Fernandez-Alonso. Nuclear dynamics in the metastable phase of the solid acid caesium hydrogen sulfate. *Phys. Chem. Chem. Phys.*, 17:31287–31296, 2015.
- <sup>235</sup> B. Merinov. Proton transport mechanism and pathways in the superprotonic phase of CsHSO<sub>4</sub> from experiment and theory. *Solid State Ion.*, 213:72–75, 2012.
- <sup>236</sup> X. Ke and I. Tanaka. Atomistic mechanism of proton conduction in solid CsHSO<sub>4</sub> by a first-principles study. *Phys. Rev. B*, 69:165114, 2004.
- <sup>237</sup> X. Ke and I. Tanaka. Proton transfer mechanism in solid CsHSO<sub>4</sub> by first-principles study. *Solid State Ion.*, 172:145–148, 2004.

- <sup>238</sup> A. Ishikawa, H. Maekawa, T. Yamamura, Y. Kawakita, K. Shibata, and M. Kawai. Proton dynamics of  $\text{CsH}_2\text{PO}_4$  studied by quasi-elastic neutron scattering and PFG-NMR. *Solid State Ion.*, 179:2345–2349, 2008.
- <sup>239</sup> A. V. Belushkin, C. J. Carlile, and L. A. Shuvalov. The diffusion of protons in the superionic conductor  $\text{CsHSO}_4$  by quasielastic neutron scattering. *J. Phys.: Condens. Matter*, 4:389–398, 1992.
- <sup>240</sup> A. Damyanovich, M. M. Pintar, R. Blinc, and J. Slak. Proton pseudoglass-to-fast-ion-conductor phase transition in  $\text{CsHSO}_4$ . *Phys. Rev. B*, 56:7942–7946, 1997.
- <sup>241</sup> J. Badot. RF-microwave dielectric relaxations and phase transitions in superionic protonic acid sulphates (selenates). *Solid State Ion.*, 35:143–149, 1989.
- <sup>242</sup> J. Badot and P. Colomban. RF-microwave dielectric relaxations and phase transitions in superionic protonic acid sulphates (selenates). *Solid State Ion.*, 35(1):143–149, 1989.
- <sup>243</sup> S. Hayashi and M. Mizuno. Proton diffusion in the superprotonic phase of  $\text{CsHSO}_4$  studied by  $^1\text{H}$  NMR relaxation. *Solid State Ion.*, 171(3):289–293, 2004.
- <sup>244</sup> J. W. Traer, K. J. Soo, M. Vijayakumar, and G. R. Goward. Elucidating the time scale and geometry of phosphate and phosphonate rotation in solid acid electrolytes using multinuclear NMR. *J. Phys. Chem. C*, 115(13):6064–6072, 2011.
- <sup>245</sup> R. C. Desai and M. Nelkin. Atomic motions in a rigid sphere gas as a problem in neutron transport. *Nucl. Sci. Eng.*, 24(2):142–152, 1966.
- <sup>246</sup> Otto J. Eder. The velocity autocorrelation function and the diffusion coefficient for a dilute hard sphere gas. *J. Chem. Phys.*, 66(9):3866–3870, 1977.
- <sup>247</sup> G. Kabbe, C. Wehmeyer, and D. Sebastiani. A coupled Molecular Dynamics/Kinetic Monte Carlo approach for protonation dynamics in extended systems. *J. Chem. Theory. Comput.*, 10:4221–4228, 2014.
- <sup>248</sup> D. Marx, A. Chandra, and M. E. Tuckerman. Aqueous basic solutions: Hydroxide solvation, structural diffusion, and comparison to the hydrated proton. *Chem. Rev.*, 110(4):2174–2216, 2010.
- <sup>249</sup> D. Marx. Proton transfer 200 years after von Grotthuss: Insights from *ab initio* simulations. *Chem. Phys. Chem.*, 7(9):1848–1870, 2006.
- <sup>250</sup> N. Hatada, K. Toyoura, T. Onishi, Y. Adachi, and T. Uda. Fast and anisotropic proton conduction in a crystalline polyphosphate. *J. Phys. Chem. C*, 118(51):29629–29635, 2014.



# ACADEMIC CURRICULUM VITAE

---

## PERSÖNLICHE DATEN

Name: Christian Dreßler  
Geburtsdatum: 11.10.1989  
Geburtsort: Löbau

---

## SCHULISCHE AUSBILDUNG/STUDIUM

2008 Abschluss des Abiturs mit der Note 1,0  
2008- 2011 Bachelorstudium der Chemie an der Universität Leipzig (Abschlussnote: 1,5)  
2011- 2013 Masterstudium der Chemie an der Universität Leipzig (Abschlussnote: 1,3)  
seit 2013 Student im Diplomstudiengang Mathematik an der Universität Leipzig; Scheinfreiheit für das Vordiplom seit Februar 2020; zwei von drei Vordiplomsprüfungen mit Note 1,3 bestanden; außerdem Teilnahme an weiteren Vorlesungen aus dem Hauptstudium  
2013-2014 Wissenschaftlicher Mitarbeiter in der Theoretischen Chemie an der Universität Leipzig  
seit 2014 Wissenschaftlicher Mitarbeiter in der Theoretischen Chemie an der MLU Halle-Wittenberg und Beginn der Promotion im Arbeitskreis von Prof. Sebastiani

---

## STIPENDIEN

2009 - 2013 Stipendiat der Studienstiftung des deutschen Volkes

---

## BETREUUNG VON ABSCHLUSSARBEITEN

2015 Masterarbeit Christopher Peschel: "Computational Investigation of Properties of 1,2,3-Triazoles"  
2018 Bachelorarbeit Christoph Kirsch: "*Ab initio* Molekulardynamik-Simulationen der Mobilität von Lithium-Ionen in Festkörpern"  
2020 Masterarbeit Thomas Kunze: "Molecular Dynamics Simulation of Hybrid Protein Systems"

- 2020 Masterarbeit Christoph Kirsch: "Molekulardynamik-Simulationen der Mobilität von Lithium-Ionen in Lithiumsiliciden"
- 2020 Bachelorarbeit Johnny Alexander Jimenez Siegert: "Investigation of proton conduction in methanesulfonic acid using the combined Molecular Dynamics/Lattice Monte Carlo (cMD/LMC) approach"
- 

#### LEHRTÄTIGKEIT

- 2013 Seminarleiter für die Vorlesung "Einführung in die Theoretische Chemie" (in Leipzig)
- 2014 Erstellung der Aufgabenserien für die Vorlesung "Physikalische Chemie I" (in Halle)
- 2014 - 2016 Betreuer im Masterpraktikum für Physikalische Chemie für den Versuch "Klassische Molekulardynamik" (in Halle)
- 2015 - 2020 Seminarleiter für die Vorlesung "Theoretische Chemie" (in Halle) mit Unterbrechung 2018 wegen Elternzeit
- 2016 - 2017 Seminarleiter für die Vorlesung "Physikalische Chemie II für Biochemiker" (in Halle)
- 2016 - 2020 Seminarleiter für die Vorlesung "Quantenchemie" (in Halle)
- 2017 Tutor für das Modul "Mathematik für Chemiker" im Fernstudiengang Chemie des Springer-Verlages
- 

#### PUBLIKATIONEN

- 12 Christian Dreßler and Daniel Sebastiani.  
Effect of chemical environment on proton mobility in the solid acids family  $\text{CsH}_y\text{XO}_4$  ( $x = \text{S, P, Se}$ ;  $y = 1, 2$ ) from *ab initio* molecular dynamics simulations.  
*Physical Chemistry Chemical Physics*, 22:10738–10752, **2020**.
- 11 Christian Dreßler, Gabriel Kabbe, Martin Brehm, and Daniel Sebastiani.  
Exploring non-equilibrium molecular dynamics of mobile protons in the solid acid  $\text{CsH}_2\text{PO}_4$  at the micrometer and microsecond scale.  
*The Journal of Chemical Physics*, 152(16):164110, **2020**.
- 10 Christian Dreßler, Gabriel Kabbe, Martin Brehm, and Daniel Sebastiani.  
Dynamical matrix propagator scheme for large-scale proton dynamics simulations.  
*The Journal of Chemical Physics*, 152(11):114114, **2020**.

- 9 Christian Dreßler and Daniel Sebastiani  
Reduced eigensystem representation of the linear density-  
density response function.  
*International Journal of Quantum Chemistry*, 120(3):e26085,  
**2020**.
- 8 Maximilian Wagner, Christian Dreßler, Felix P. Lohmann-  
Richters, Kevin Hanus, Daniel Sebastiani, Aron Varga, and  
Bernd Abel.  
Mechanism of ion conductivity through polymer-stabilized  
CsH<sub>2</sub>PO<sub>4</sub> nanoparticulate layers from experiment and theory.  
*Journal of Materials Chemistry A*, 7:27367–27376, **2019**.
- 7 Christian Dreßler, Arne Scherrer, Paul Ahlert, and Daniel  
Sebastiani.  
Efficient representation of the linear density-density response  
function.  
*Journal of Computational Chemistry*, 40(31):2712–2721, **2019**.
- 6 Paul Ahlert, Arne Scherrer, Christian Dreßler, and Daniel  
Sebastiani.  
Iterative approach for the moment representation of the density-  
density response function.  
*The European Physical Journal B*, 91(6):94, **2018**.
- 5 Gabriel Kabbe, Christian Dreßler, and Daniel Sebastiani.  
Proton mobility in aqueous systems: combining ab initio  
accuracy with millisecond timescales.  
*Physical Chemistry Chemical Physics*, 19(42):28604–28609,  
**2017**.
- 4 Arne Scherrer, Christian Dreßler, Paul Ahlert, and Daniel  
Sebastiani.  
Generalization of the electronic susceptibility for arbitrary  
molecular geometries.  
*The Journal of Chemical Physics*, 144(14):144111, **2016**.
- 3 Christian Dreßler, Gabriel Kabbe, and Daniel Sebastiani.  
Insight from atomistic simulations of protonation dynamics  
at the nanoscale.  
*Fuel Cells*, 16(6):682–694, **2016**.
- 2 Christian Dreßler, Gabriel Kabbe, and Daniel Sebastiani.  
Proton conductivity in hydrogen phosphate/sulfates from a  
coupled Molecular Dynamics/Lattice Monte Carlo (cMD/LMC)  
approach.  
*The Journal of Physical Chemistry C*, 120(36):19913–19922,  
**2016**.
- 1 Gabriel Kabbe, Christian Dreßler, and Daniel Sebastiani.  
Toward realistic transfer rates within the coupled Molecular  
Dynamics/Lattice Monte Carlo approach.

*The Journal of Physical Chemistry C*, 120(36):19905–19912,  
**2016.**

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbstständig und nur unter Verwendung der angegebenen Hilfsmittel und Quellen angefertigt habe. Alle Stellen, die wörtlich oder sinngemäß aus Veröffentlichungen entnommen sind, habe ich als solche kenntlich gemacht. Des Weiteren versichere ich, dass ich diese Arbeit an keiner anderen Institution eingereicht habe. Mir ist bekannt, dass bei Angabe falscher Aussagen die Prüfung als nicht bestanden gilt. Mit meiner Unterschrift versichere ich die Richtigkeit dieser Angaben und erkenne die rechtlichen Grundlagen an.

---

Datum

---

Unterschrift