Model Order Reduction for Stochastic Differential Equations driven by Standard and Fractional Brownian motion

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

Frau Nahid Jamshidi

Gutachter:

Herr Prof. Dr. Martin Redmann (Martin-Luther-Universität Halle-Wittenberg)

Herr Prof. Dr. Wilfried Grecksch (Martin-Luther-Universität Halle-Wittenberg)

Herr Prof. Dr. Ralf Wunderlich (Brandenburgische Technische Universität Cottbus-Senftenberg)

Tag der Verteidigung: 07.11.2025

To my father

Acknowledgments

First and foremost, I would like to express my deepest gratitude to **Prof. Dr. Martin Redmann** for his invaluable guidance, constant support, and insightful mentorship throughout my PhD journey.

My sincere appreciation also goes to **Prof. Dr. Wilfried Grecksch** for his continuous support and advice, without which this dissertation would not have been possible. His contributions were crucial to my progress, and I am truly grateful for his help.

I am equally grateful to **Prof. Dr. Ralf Wunderlich** for kindly agreeing to serve as a reviewer of this dissertation. I sincerely thank him for his valuable discussions and thoughtful comments, which have enriched the scientific depth of my work.

Special thanks are extended to the members of the *Numeric Group* for their collaboration, stimulating discussions, and technical support throughout the various stages of my research. Their contributions have made this journey both productive and enjoyable.

I am deeply thankful to my colleagues for their friendship and companionship, and to my family for their unwavering love and encouragement. My heartfelt thanks go to my spouse, whose patience, support, and belief in me have been a constant source of strength during the most challenging times of this journey.

Finally, I wish to acknowledge the broader academic community and all those who have, directly or indirectly, contributed to the realization of this dissertation. This work would not have been possible without your collective support.

Abstract

This dissertation addresses the challenges of solving high-dimensional stochastic differential equations (SDEs) often spatially discretized stochastic partial differential equations (SPDEs) driven by both standard Brownian motion (sBm) and fractional Brownian motion (fBm). The primary focus is on model order reduction (MOR) techniques, essential for simplifying large-scale systems into lower-dimensional, computationally efficient models while preserving critical dynamics.

We study numerical solutions for large-scale linear stochastic systems, such as spatially discretized SPDEs driven by standard and fractional Brownian motion. Particular emphasis is placed on the stochastic heat equation, where a spectral Galerkin scheme is employed for spatial discretization, extending existing results to SPDEs influenced by Wiener and fractional noise. Due to the lack of asymptotic stability in semi-discretized SPDEs under large noise, we develop MOR schemes specifically designed for unstable systems with Wiener noise. A novel Gramian-based approach is proposed to identify dominant subspaces, with Gramians constructed through Lyapunov equations. Since covariance information is not directly available, efficient sampling-based methods incorporating variance reduction techniques and deterministic approximations of covariance functions are introduced. An error bound is established, providing a priory criteria for selecting the reduced system dimension, and ensuring the applicability of the method even in deterministic settings. Comprehensive numerical experiments validate the proposed MOR schemes, demonstrating their computational efficiency and effectiveness in high-dimensional stochastic systems.

We also investigate systems driven by fBm with the Hurst parameter $H \in [1/2, 1)$, using the Young and Stratonovich interpretations. Fractional Young differential equations capture memory effects, and we analyze fundamental solutions to introduce empirical reduced order methods based on snapshots or approximated Gramians. For $H > 1/2$, empirical Gramians from the simulation data are proposed. We present projection-based ROMs, noting the need for improvements in Stratonovich settings. Numerical experiments validate our techniques, offering insights into effective MOR for stochastic systems with fractional noise, and aiding efficient computational strategies for practical applications.

Contents

Introduction **1**

1 Preliminaries in Stochastic **8**

1.1 The Basic Concepts of Probability Theory 8
1.2 Stochastic Processes 9
1.2.1 Markov Processes 10
1.2.2 Martingales 11
1.2.3 Standard Brownian Motion 11
1.3 Stochastic Integrals 12
1.3.1 Stochastic Integral for Simple Functions 13
1.3.2 Stochastic Integrals for General Processes 15
1.4 Itô Formula 18
1.4.1 One Dimensional Itô Formula 18
1.4.2 Multidimensional Itô Formula 18
1.5 Existence and Uniqueness of Solutions for SDEs 19
1.6 Convergence of Random Sequences 20
1.6.1 Strong Convergence 20
1.6.2 Weak Convergence 21
1.7 Stability Analysis of SDEs 21
1.7.1 Stochastic Stability 21
1.7.2 Mean Square Stability 22
1.8 Euler-Maruyama Method for SDEs 22

2 Fractional Brownian Motion **24**

2.1 Properties of Fractional Brownian Motion 24
2.2 Different Stochastic Representation of fBm 26
2.3 Simulation of Fractional Brownian Motion 28
2.4 Semimartingale and Markov Property 30
2.5 Stochastic Calculus with respect to fBm 31
2.5.1 Stochastic Integration of Deterministic Processes 31

2.5.2 Stochastic Integration of Random Processes . . . 33
2.5.3 Malliavin Calculus for fBm . . . 35
2.5.4 Fractional Wick Itô Skorohod Integrals . . . 37
2.6 Fractional Itô Formula . . . 40
2.7 Existence and Uniqueness of Solution of SDE with fBm . . . 41
2.8 Numerical Methods for SDEs with fBm . . . 43
2.8.1 Euler Method . . . 43
2.8.2 Implicit Midpoint Method . . . 44

3 Model Order Reduction **47**
3.1 Linear Deterministic Systems . . . 47
3.1.1 Lyapunov Operators . . . 49
3.1.2 Stability . . . 50
3.1.3 Model Order Reduction for Deterministic Systems . . . 51
3.2 linear Stochastic Dynamical Systems . . . 56
3.2.1 Lyapunov Equations in Stochastic Case . . . 59
3.2.2 Stability . . . 62
3.2.3 Characterization of Dominant Subspaces Using Gramian Matrices . . . 62
3.2.4 Model Order Reduction for Stochastic Systems . . . 66

4 SPDEs Driven by Fractional Brownian Motion **72**
4.1 Preliminary . . . 72
4.1.1 C_0-semigroups . . . 72
4.1.2 Operator Spaces . . . 74
4.2 Evolution Equations with fBm . . . 75
4.2.1 Existence and Uniqueness of Mild Solution . . . 76
4.3 Spectral Galerkin Method . . . 82
4.4 Example . . . 87
4.4.1 Stochastic Heat Equation . . . 87

5 Gramian-Based Model Reduction for Unstable Stochastic Systems **91**
5.1 Gramian-Based MOR . . . 91
5.1.1 Gramians and Characterization of Dominant Subspaces . . . 91
5.1.2 Reduced Order Modeling by Transformation of Gramians . . . 96
5.2 Output Error Bound . . . 97
5.3 Computation of Gramians . . . 104
5.3.1 Exact Methods . . . 104

5.3.2 Sampling Based Approaches . . . 105
5.3.3 Gramians Based on Deterministic Approximations of $F(T)$ and $G(T)$ . . . 108
5.4 Numerical Experiments . . . 109
5.4.1 Simulations for $n = 100$ and $T = 1$ . . . 110
5.4.2 Simulations for $n = 1000$ and $T = 1$ . . . 111
5.4.3 Relevance of T and K . . . 112

6 Model Reduction for Stochastic Systems Driven by fBm **114**
6.1 Setting and (Projection-Based) Reduced System . . . 114
6.2 Fundamental Solutions and Gramians . . . 116
6.2.1 Fundamental Solutions and Their Properties . . . 116
6.2.2 Exact and Empirical Gramians . . . 119
6.3 Model Reduction of Young/Stratonovich Differential Equations . . . 125
6.3.1 State Space Transformation and Balancing . . . 125
6.3.2 Reduced Order Models Based on Projection . . . 127
6.3.3 An Alternative Approach for the Stratonovich Setting ($H = 1/2$) 127
6.4 Numerical Results . . . 133
6.4.1 Time Integration . . . 133
6.4.2 POD-Based Method . . . 133
6.4.3 Dimension Reduction for a Stochastic Heat Equation . . . 134

Conclusions **138**

Summary of Contributions **140**

A Norms **142**
A.1 Norms of Vectors and Matrices in Finite Dimensions . . . 142
A.1.1 Singular Value Decomposition . . . 143
A.1.2 Schatten p-norm . . . 143
A.2 Banach and Hilbert Spaces . . . 144
A.2.1 Nuclear Operator . . . 144
A.2.2 Hilbert-Schmidt Operator . . . 145

B Kronecker Product **146**

List of Abbreviations

SPDE	Stochastic Partial Differential Equation
fBm	Fractional Brownian Motion
sBm	Standard Brownian Motion
SDE	Stochastic Differential Equation
DS	Dynamical System
ROM	Reduced Order Model
MOR	Model Order Reduction
POD	Proper Orthogonal Decomposition
PCA	Principal Component Analysis
SVD	Singular Value Decomposition
LBR	Lyapunov Balanced Reduction
HSVs	Hankel Singular Values
BT	Balanced Truncation
fWIS integral	Fractional Wick Itô Skorohod integral

List of Symbols

$\mathbb{N}$	Natural numbers, i.e., $\{1, 2, 3, \ldots\}$,
$\mathbb{Z}_+$	$\mathbb{N} \cup \{0\}$,
$\mathbb{R}$	Real numbers,
$\mathbb{R}_+$	Nonnegative real numbers,
$\mathbb{R}^{m\times n}$	Vector space of real matrices with m rows and n columns,
$\mathbb{R}^n$	$\mathbb{R}^{n\times 1}$,
$\mathbb{C}$	Complex numbers,
$\mathbb{C}_-$	Open left half of the complex plane,
A^T	Transpose of a matrix A,
A^*	Adjoint operator of A,
A^{-1}	Inverse of a matrix or an operator A,
W^H	Fractional Brownian motion (fBm) with Hurst parameter H,
$\mathbb{E}$	Expected value,
$(\Omega, \mathcal{F}, \mathbb{P})$	Probability space,measure space with $\mathcal{P}(\Omega) = 1$,
$(\mathcal{F}_t)_{t\geq 0}$	filtration; increasing family of sub-σ-algebras of $\mathcal{F}$; $\mathcal{F}_{t_1} \subseteq \mathcal{F}_{t_2}$ for $t_1 \leq t_2$,
$\mathcal{H}$	Hilbert space,
$\circ$	Indicates Young and Stratonovich integral,
δ	Malliavin divergence operator,
$\mathcal{C}\left(D; \mathbb{R}^d\right)$	The family of continuous $\mathbb{R}^d$-valued functions defined on D,
$\mathcal{C}^k\left(D; \mathbb{R}^d\right)$	The family of continuously k-times differentiable $\mathbb{R}^d$-valued functions defined on D,
$\mathcal{C}^\alpha\left(D; \mathbb{R}^d\right)$	The family of $\mathbb{R}^d$-valued Hölder continuous with exponent α functions defined on D
$\mathcal{C}_b^k\left(D; R^d\right)$	The family of functions in $\mathcal{C}^k\left(D; \mathbb{R}^d\right)$ with bounded derivatives,

$\mathcal{B}^d$	$d-$Dimension Borel $\sigma-$algebra,
D_t^H	Malliavin derivative,
D_s^ϕ	Malliavin ϕ-derivative,
$\diamond$	Wick product,
$\mathbb{D}^{1,p}$	Sobolev space of functions,
$\sigma(A)$	Spectrum of a matrix or an operator A,
$\mathrm{tr}(\cdot)$	Trace of a matrix or an operator; sum of the diagonal entries (matrix case),
$\mathrm{vec}(\cdot)$	Vectorization of a matrix,
$\|.\|_u, \langle\cdot,\cdot\rangle_U$	Norm and inner product in a Hilbert space U,
$\|\cdot\|_2, \langle\cdot,\cdot\rangle_2$	Euclidean norm and inner product,
$\|\cdot\|_F, \langle\cdot,\cdot\rangle_F$	Frobenius norm and inner product,
$L_H(0,T)$	Space of processes suitable for fractional integration,
$L(U,V)$	Space of Linear and bounded operators from U to V
$L(U)$	$L(U,V)$ with $V=U$
$L_1(U)$	Space of nuclear operators on U
$L_2(V,U)$	Space of Hilbert-Schmidt operators from V to U
$L^p(\Omega;U)$	The family of U-valued random variables ξ with $\mathbb{E}\|\xi\|_U^p < \infty$,
$L^p\left(\Omega;\mathbb{R}^d\right)$	The family of $\mathbb{R}^d$-valued random variables ξ with $\mathbb{E}\|\xi\|_2^p < \infty$,
$L^p\left((a,b);\mathbb{R}^d\right)$	The family of Borel measurable functions $h:[a,b]\to\mathbb{R}^d$ such that $\int_a^b \|h(t)\|_2^p dt < \infty$,
$\mathcal{L}^p\left((a,b);\mathbb{R}^d\right)$	The family of $\mathcal{R}^d$-valued $\mathcal{F}_t$-adapted processes $\{f(t)\}_{a\le t\le b}$ such that $\int_a^b \|f(t)\|_2^p dt < \infty$ a.s. ,
$\mathcal{M}^p\left((a,b);\mathbb{R}^d\right)$	The family of processes $\{f(t)\}_{a\le t\le b}$ in $\mathcal{L}^p\left([a,b];\mathbb{R}^d\right)$ such that $\mathbb{E}\int_a^b \|f(t)\|_2^p dt < \infty$,
$\mathcal{L}^p\left(\mathbb{R}_+;\mathbb{R}^d\right)$	The family of processes $\{f(t)\}_{t\ge0}$ such that for every $T>0$, $\{f(t)\}_{0\le t\le T}\in\mathcal{L}^p\left([0,T];\mathbb{R}^d\right)$,
$\mathcal{M}^p\left(\mathbb{R}_+;\mathbb{R}^d\right)$	The family of processes $\{f(t)\}_{t\ge0}$ such that for every $T>0$, $\{f(t)\}_{0\le t\le T}\in\mathcal{M}^p\left([0,T];\mathbb{R}^d\right)$,
$\Gamma(\cdot)$	Gamma function,
$x(t;x_0,u)$	State of the system depending on initial state x_0 and control u,

$\|u\|_{L_T^2}$	Norm of control function u in L^2 space over time interval T,
$\|\cdot\|_2$	Euclidean norm,
$A^{\frac{1}{2}}$	Square root of a matrix or an operator A,
$A > 0$	Symmetric positive definite matrix A,
$\mathrm{diag}(a_1, a_2, \ldots, a_n)$	Diagonal matrix with the diagonal entries $a_1, a_2, \ldots, a_n$,
$D(\cdot)$	Domain of an operator,
$\mathrm{im}(\cdot)$	Image of a matrix or an operator; space spanned by the columns (matrix case),
I_n, I	n-dimensional identity matrix or identity matrix of suitable dimension,
$\ker(\cdot)$	Kernel of a matrix or an operator; inverse image of zero,
$A \otimes B$	Kronecker product of matrices A and B,
$\mathrm{span}\{x_1, x_2, \ldots, x_n\}$	Space spanned by the vectors $x_1, x_2, \ldots, x_n$,
$\mathcal{N}(0, h)$	Normal distribution with mean zero and variance h,

Introduction

Computational science has emerged as a critical third pillar in science and industry, complementing theoretical and experimental approaches. It enables deep insights into systems in various fields like physics, chemistry, biology, economics, and engineering. The growing need for real-time simulation, control, and prediction of complex systems presents significant challenges. These challenges often lead to optimal control problems involving dynamical systems (DS) governed by (stochastic) partial differential equations ((S)PDEs).

Advancements over the past decades, driven by increased computational power and improved numerical algorithms, have made it possible to solve many complex problems that were previously intractable. However, standard spatial discretization methods for high-resolution DS often result in large-scale, high-dimensional systems of stochastic or ordinary differential equations. To address these challenges, MOR

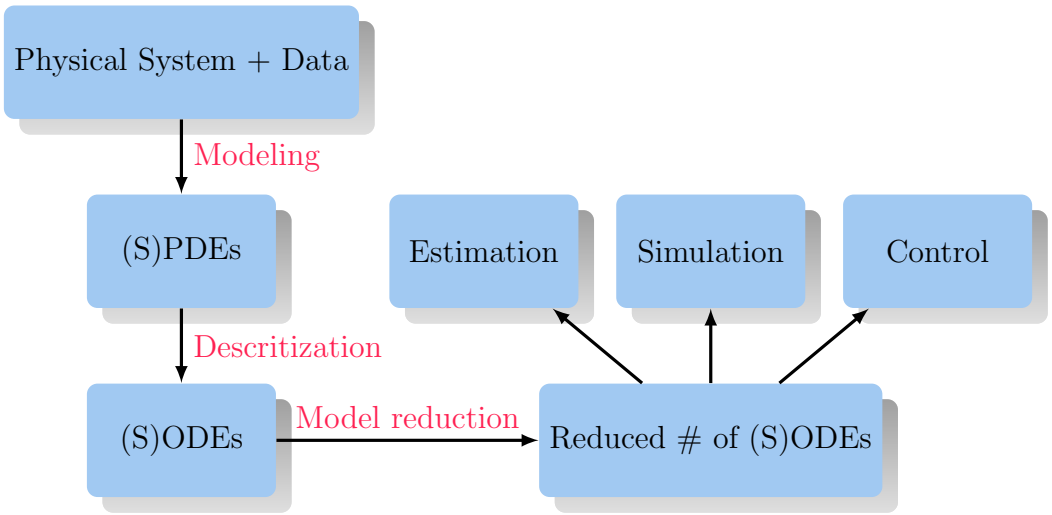

Figure 0.1: Process to a Reduced-Order Modeling

techniques provide a promising approach. These techniques simplify large-scale systems into low-dimensional Reduced-Order Models (ROMs), facilitating robust simulation and active control (Figure 0.1). MOR aims to find low-dimensional approximations of high-dimensional DS by focusing on dominant modes, thus reducing computational complexity while preserving essential input-output behaviors (Figure 0.2). The reduction process must be reliable, computationally efficient, and result in minimal approximation errors. There has been an enormous interest in MOR techniques for deterministic equations. Let us refer to [5, 13], where an overview of different approaches is given and further references can be found. MOR for Itô

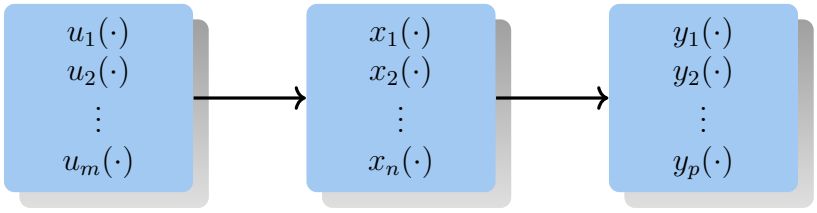

Figure 0.2: Input-output systems

stochastic differential equations (SDEs) is also very natural thinking of computationally very involved techniques like Mont-Carlo methods. There has been vast progress in the development of MOR schemes in the Itô setting. Let us refer to [14, 100, 116] in order to point out three different approaches in this context. Chapter 3 is devoted to the concept of MOR and defines various MOR methods, highlighting the most critical ones: the Balanced Truncation (BT) method and the Proper Orthogonal Decomposition (POD) method, which are pivotal in reducing system complexity while preserving essential dynamics.

Balanced Truncation is a prominent method for MOR in both deterministic and stochastic systems. It uses Singular Value Decomposition (SVD) to approximate a matrix to its lower-rank form, as detailed by Moore [76] and further developed by Mullis and Roberts [74]. In deterministic systems, Lyapunov Balanced Reduction (LBR), which involves solving reachability and observability Lyapunov equations, ensures stability and provides error bounds, as established by Glover [89] and Enns [33]. For stochastic systems, balanced truncation incorporates stochastic Gramians to account for noise impact, allowing for accurate simplification of complex models with clear error bounds, crucial for fields like aerospace and financial engineering. This adaptation was notably advanced in the late 1980s and 1990s [41, 25], and recent discussions on Itô type SDEs can be found in [10, 14].

POD method is another crucial MOR technique developed through contributions by Karhunen [56], Loéve [67], and others [63, 92, 15]. It has been extensively applied in various scientific and engineering domains, including fluid dynamics [42, 49, 107, 112], electric circuit analysis [88], and structural dynamics [4]. Comprehensive reviews on the history and applications of POD can be found in [20, 22, 58, 65, 118]. While POD has been effectively used for deterministic systems derived from PDEs, its extension to SDEs influenced by Wiener processes has been less explored. Notable applications in stochastic settings include the stochastic Burgers equation [48, 122] and stochastic Hamiltonian systems, where POD techniques were shown to improve solution accuracy and stability [116]. This dissertation explores advanced MOR techniques in stochastic systems driven by standard Brownian motion (sBM) and fractional Brownian motion (fBm).

Fractional Brownian motion, introduced by Kolmogorov [62] and further explored by Mandelbrot and Van Ness [69], differs from classical Brownian motion by its Hurst parameter H, which ranges from 0 to 1 and adjusts its self-similarity index. While sBm has a self-similarity index of $1/2$ and stationary increments, fBm's varying H allows it to model a wider array of phenomena but also prevents it from being a

semimartingale or Markov process when $H \neq 1/2$. This distinction complicates the analytical treatment of fBm but opens new research avenues. Integrating processes with respect to fBm, particularly for $H > 1/2$, can be approached using pathwise methods based on Young's integration [123] or fractional calculus [34]. Malliavin calculus has also been pivotal, with foundational work by Decreusefond and Üstünel [27], and further contributions by Carmona and Coutin [21], Alós, Mazet, and Nualart [1, 2], and the monographs by Hu and Biagini [44, 18]. A comprehensive discussion of fBm, detailing its properties and various integration techniques is provided in Chapter 2.

Due to its self-similarity and long-range dependency, fBm is an excellent candidate for simulating various phenomena in practice. The significance of long-range dependence is underscored by the extensive volume of literature that features this concept in their titles. Such publications span diverse fields including finance [66], econometrics [105], internet modeling [55], hydrology [87], climate studies [117], linguistics [3], DNA sequencing [57], and physics [38, 39, 93].

In this dissertation, we study a $\mathbb{R}^q$-valued fBm process $W^H = \left(W_1^H, \dots, W_q^H\right)^\top$, characterized by a mean of zero, Hurst parameter $H \in [1/2, 1)$, and the covariance matrix $\mathbf{K} = (k_{ij})_{i,j=1,\dots,q}$. The process is characterized by

$$\mathbb{E}[W^H(t)(W^H(t))^\top] = \mathbf{K}t^{2H}, \quad \text{for} \quad t \in [0, T],$$

where $T > 0$ denotes the terminal time. These processes are defined on a filtered probability space $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P}\right)$, where $(\mathcal{F}_t)_{t\in[0,T]}$ is assumed to be right-continuous and complete.

Initially, we focus on Wiener noise, corresponding to $H = \frac{1}{2}$, and assume $W^{\frac{1}{2}} := W$ is $(\mathcal{F}_t)_{t\in[0,T]}$-adapted, with increments $W(t+h) - W(t)$ being independent of $\mathcal{F}_t$ for $t, h \geq 0$. Comprehensive details and concepts about SDEs and Itô calculus are provided in Chapter 1. In the latter part of this dissertation, we extend our analysis to fBm with Hurst parameter $H \geq \frac{1}{2}$. This unified approach allows us to analyze the properties and behaviors of both Wiener processes and fBm, using their unique characteristics in various stochastic modeling scenarios.

We consider the following large-scale controlled linear SDE:

$$\mathrm{d}x(t) = [Ax(t) + Bu(t)]\mathrm{d}t + \sum_{i=1}^{q} N_i x(t)\mathrm{d}W_i(t), \qquad x(0) = x_0, \tag{0.1a}$$

$$y(t) = Cx(t), \qquad t \in [0, T], \tag{0.1b}$$

where $A, N_i \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$ and $C \in \mathbb{R}^{p\times n}$. The state dimension n is assumed to be large and the quantity of interest y is often low-dimensional, i.e., $p \ll n$, but we also discuss the case of a large p. By $x(t; x_0, u)$, we denote the state in dependence on the initial state x_0 and the control u, for which we assume that it is $(\mathcal{F}_t)_{t\in[0,T]}$-adapted and $\|u\|_T^2 := \mathbb{E}\int_0^T \|u(s)\|_2^2\, ds < \infty$ with $\|\cdot\|_2$ representing the Euclidean norm.

The first goal of this dissertation is to construct a system with state x_r and

quantity of interest y_r having the same structure as (0.1) but a much smaller state dimension $r \ll n$. At the same time, it is aimed to ensure $y \approx y_r$. Such a reduced order model (ROM) is particularly beneficial if many evaluations (0.1) for several controls u are required (e.g. in an optimal control problem) combined with the need to generate many samples of y for each u. Now, a ROM shall be achieved under very general conditions such as the absence of mean square asymptotic stability, i.e., $\mathbb{E}\,\|x(t;x_0,0)\|_2^2 \to 0$ (as $t \to \infty$) is not given. Methods involving such a stability condition are intensively studied in the literature [9, 14, 96, 103] since it is often guaranteed if (0.1a) results from a spatial discretization of a SPDE such as

$$\frac{\partial \mathcal{X}(t,\zeta)}{\partial t} = \Delta \mathcal{X}(t,\zeta) + \mathcal{B}u(t) + \sum_{i=1}^{q} \mathcal{N}_i \mathcal{X}(t,\zeta) \frac{\partial W_i(t)}{\partial t}. \tag{0.2}$$

The solution $\mathcal{X}(t,\cdot)$ to the heat equation (0.2) is viewed as a stochastic process taking values in a Hilbert space and shall be approximated by x. In this context, A can be seen as a discretized version of the Laplacian Δ and B, N_i represent discretizations of the linear bounded operators $\mathcal{B}$, $\mathcal{N}_i$. Moreover, W_i can be interpreted as Fourier coefficients corresponding to a truncated series of space-time noise. Further explanations on different schemes for spatial discretization can, e.g., be found in [8, 24, 40]. However, even in a setting like in (0.2), mean square asymptotic stability can be violated since the noise can easily cause instabilities (e.g. if it is sufficiently large).

Such a scenario is of interest in this thesis. We establish generalizations of balancing related MOR schemes in order to make them applicable to general systems (0.1). These MOR methods rely on matrices called Gramians that can be used to identify the dominant subspaces of (0.1). Based on this characterization of the relevance of different state directions, less important information in the dynamics is removed leading to the desired ROM. This step can be interpreted as an optimization procedure applied to spatially discretized SPDE. In an unstable setting, Gramians need to be defined that generally exist in contrast to previous approaches. We consider generalized time-limited Gramians in Chapter 5. Such type of Gramians have been used in deterministic frameworks [37, 64, 74, 102]. Although such an ansatz is beneficial for the setting we want to cover, the analysis of MOR methods based on generalized time-limited Gramians is much more challenging. Furthermore, the question of how to compute these Gramians in practice is very difficult but vital since they are required in order to derive the ROM.

In Chapter 5, we introduce the time-limited Gramian in the stochastic setting studied here. This chapter presents the work published in [101], where the theoretical findings are elaborated and applied. We point out the relation between these Gramians and the dominant subspaces of (0.1) and show their relation to matrix (differential) equations. Subsequently, we discuss two different MOR techniques based on these Gramians and analyze the respective error. In particular, an error bound is established that allows us to point out situations in which the approaches work well. It is important to point out that this bound is more than just a gener-

alization of the deterministic case [102]. The new type of representation links the truncated Hankel singular values of the system or the truncated eigenvalues of the reachability Gramian, respectively, to the error of the approximation without needing asymptotic stability and is hence beneficial also in unstable settings. Moreover, we discuss different strategies that can be used to compute the proposed Gramians. They are solutions to Lyapunov equations. However, in a time-limited scenario, covariance information at the terminal time enters these Lyapunov equations which is not immediately available. Since direct methods only work in moderately high dimensions, we focus on sampling based approaches to estimate the required covariances. In order to increase the efficiency of such procedures we apply variance reduction methods in this context leading to an efficient way of solving for the time-limited Gramians. Apart from this empirical procedure, a second strategy to approximate covariance functions and hence the Gramians is investigated, where potentially expensive sampling is not required. This chapter concludes with several numerical experiments showing the efficiency of the MOR methods.

The second objective of this thesis is to develop MOR methods for stochastic systems driven by fBm with non-zero initial conditions. The system under consideration is described by:

$$\begin{aligned} \mathrm{d}x(t) &= [Ax(t) + Bu(t)]\mathrm{d}t + \sum_{i=1}^{q} N_i x(t) \circ \mathrm{d}W_i^H(t), \quad x(0) = x_0 = X_0 z, \\ y(t) &= Cx(t), \quad t \in [0,T], \end{aligned} \tag{0.3}$$

where $W^H = \{W_1^H, \dots, W_q^H\}$ are independent fBm with Hurst index $H \in [1/2, 1)$. The matrices are defined similarly to those in system (0.1), with $X_0 \in \mathbb{R}^{n\times v}$, $z \in \mathbb{R}^v$ and $T > 0$ being the terminal time. System (0.3) is defined as an integral equation using Young integration ($H > 1/2$) and Strtonovich integration ($H = 1/2$) to make sense of $\int_0^t N_i x(s) \circ \mathrm{d}W_i^H(s)$.

As mentioned before, when $H \neq \frac{1}{2}$, the process W^H is neither a semimartingale nor a Markov process. These are the main obstacles when MOR techniques are designed for such systems. The dimension reduction we focus on is conducted by identifying the dominant subspaces using quadratic forms of the solution to the stochastic equation, specifically Gramian matrices. By characterizing the relevance of different state directions using Gramians, less important information can be removed to achieve the desired ROM. Our work considers various types of Gramians depending on their availability in different settings. The exact Gramians are studied on compact intervals $[0, T]$ as well as on infinite time horizons. As stated earlier, these have previously been used in deterministic frameworks or Itô stochastic differential equations (see Chapter 5 and [9, 10, 14]). Given the Young case of $H > 1/2$, the fractional driver does not have independent increments making it hard to extend the concept of Gramians to this setting. One of our contributions is the analysis of fundamental solutions of Young differential equations. We prove a weak form of semigroup property in Lemma 6.4 which is the basis for a proper definition of Gramians for $H > 1/2$ and new even if $H = 1/2$. This lemma is the key for the entire

theory and opens up opportunities to study MOR for equations with drivers solely having stationary increments. The lemma 6.4 is, for example, exploited to show that certain eigenspaces of these Gramians are associated with dominant subspaces of the system and therefore confirms that the choice of Gramians is meaningful. However, this approach is still very challenging from the computational point of view for fBms with $H > 1/2$. This is due to a missing link of the proposed exact Gramians to Lyapunov equations, a connection that is the foundation for the previous theory of MOR with $H = 1/2$ (Itô case). The link to matrix equations only exists when $H = 1/2$, because the increments of the driver are independent in that case. Therefore, empirical Gramians based on simulation data are introduced. Computing this approximation of the exact Gramians is still challenging yet vital since they are needed for deriving the ROM. We further point out, how exact Gramians can be computed for Stratonovich stochastic differential equations. Here, the equivalence to Itô equations is exploited. Although we show that these Gramians identify redundant information in Stratonovich settings, MOR turns out to be not as natural as in the Itô case. In fact, we illustrate that projection-based dimension reduction for Stratonovich equations leads to ROMs that lack important properties. For instance, stability might not be preserved in the ROM and the error does not solely depend on the truncated eigenvalues of the Gramians. This indicates that there are situations in which the projection-based ROM performs poorly. For that reason, we propose a modification of the ROM having all these nice properties known for Itô equations (stability and meaningful error bounds).

As previously mentioned, such a system (0.3) results from the spatial discretization of a SPDE driven by fBm, such as the heat equation (0.2) with fBm instead of Wiener noise. SPDEs driven by fBm have attracted considerable research attention in mathematics. Numerous studies have examined various aspects of SPDEs influenced by fBm with different values of the Hurst parameter. In recent years, significant progress has been made in the theory of SPDEs driven by fBm, especially for $H \in \left(\frac{1}{2}, 1\right)$. Noteworthy advancements include the extensive study of SPDEs in Hilbert spaces with infinite-dimensional fBm, as discussed in [31, 53, 71]. Further details on the theory of such equations can be found in Chapter 4.

To address the spatial discretization of SPDEs, the Galerkin method is frequently employed. This method approximates the solution of the SPDE by using a finite-dimensional subspace of trial functions. In Chapter 4, we also explore a spectral Galerkin scheme for specific SPDEs driven by fBm, building upon the methods studied in [29] for certain SPDEs with fractional noise. This scheme is applied specifically to equation (0.2) with fBm. The convergence of the spectral Galerkin solution to the mild solution of the corresponding SPDE was demonstrated. This discretization resulted in high-dimensional linear SDEs, motivating the extension of balancing-based model order reduction to mean square asymptotically stable controlled stochastic systems. Several investigations have focused on the implementation of the Galerkin method for spatial discretization in the context of SPDEs driven by fBm. The papers in [53, 120] apply the spectral Galerkin method for the spatial discretization of SPDEs driven by fBm with a Hurst parameter $H > \frac{1}{2}$.

In the end, Chapter 6, which is based on our works published in [50, 51], begins with a brief discussion of the setting and the general structure of the reduced system by projection. This includes an initial insight into how projection-based reduced systems need to be modified to ensure better approximation quality in the Stratonovich setting. Following this, we study the properties of the fundamental solution to the underlying stochastic system which is vital for any kind of theoretical concerns and new for each choice of the Hurst index H. A weak type of semigroup property leads to a natural notion of Gramians, which we show to characterize the dominant subspaces of the system and form the basis for our dimension reduction. This subspace identification is an essential theoretical contribution as it explains the role of the Gramian introduced here. Since exact Gramians are not available for every choice of H, several modifications and approximations are discussed. We also provide strategies for computing Gramians for Stratonovich equations. Subsequently, we describe the concept of balancing for all variations of the proposed Gramians. This is followed by a truncation procedure to yield a ROM. We further prove that the truncation method is not optimal in the Stratonovich case (no stability preservation and a potentially large error) and suggest an alternative that is based on transformation into the equivalent Ito framework. This is another key contribution of this thesis. It is noted that the truncation method is not optimal in the Stratonovich case ($H = 1/2$), and an alternative based on transformation into the equivalent Itô framework is suggested. Finally, we apply the methods described to solve the stochastic heat equation with fractional noise. This section presents the results of our simulations that demonstrate the effectiveness of the proposed methods in solving these equations under various noise conditions.

1 Preliminaries in Stochastic

This chapter explores fundamental concepts and theorems essential to the theory of SDEs. The content provided is derived from the foundational works of Kloeden and Platen [60], Øksendal [86] and Mao [70].

1.1 The Basic Concepts of Probability Theory

Definition 1.1 Suppose that we have a non-empty set denoted as Ω, representing the sample space. We define a σ-algebra, also known as a σ-field $\mathcal{F}$, as a collection of sets $\{A_i\}_{i\geq 1}$, where each A_i is a subset of Ω. This collection must satisfy the following conditions:

(i) $\emptyset \in \mathcal{F}$

(ii) $A \in \mathcal{F} \Rightarrow A^{\complement} \in \mathcal{F}$ where $A^{\complement} = \Omega \setminus A$ is complement of A in Ω,

(iii) $\{A_i\}_{i\geq 1} \subset \mathcal{F} \Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$.

The pair $(\Omega, \mathcal{F})$ is known as a measurable space, where the sets belonging to $\mathcal{F}$ are termed as $\mathcal{F}$-measurable sets. Given $\mathcal{C} \subseteq 2^{\Omega}$, the smallest σ-algebra containing $\mathcal{C}$ is denoted $\sigma(\mathcal{C})$. On $\mathbb{R}^d$, the *Borel σ-algebra* is $\mathcal{B}(\mathbb{R}^d) := \sigma(\mathcal{O})$ where $\mathcal{O}$ is the family of open sets. Its elements are called Borel sets.

Afterward, we introduce the probability measure denoted as $\mathbb{P}$ and illustrate its associated properties.

Definition 1.2 A probability measure $\mathbb{P}$, which is defined over the measurable space $(\Omega, \mathcal{F})$, is a function $\mathbb{P} : \mathcal{F} \to [0, 1]$ such that

(i) $\mathbb{P}(\Omega) = 1$,

(ii) for any disjoint sequence $\{A_i\}_{i\geq 1} \subset \mathcal{F}$ (i.e. $A_i \cap A_j = \emptyset$ if $i \neq j$)

$$\mathbb{P}(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mathbb{P}(A_i).$$

The triple $(\Omega, \mathcal{F}, \mathbb{P})$ is called a probability space.
In this context, we assume that the probability space is complete, which means that for every $B \in \mathcal{F}$ with $\mathbb{P}(B) = 0$ and every $A \subset B$ we also have $A \in \mathcal{F}$.

Definition 1.3 If $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space, then $x : (\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ is a random variable if, for every $a \in \mathbb{R}$,

$$\{\omega \in \Omega : x(\omega) \leq a\} \in \mathcal{F}.$$

This function x is also called real-valued $\mathcal{F}$-measurable random variable. A $\mathbb{R}^d$-valued function $x(\omega) = (x_1(\omega), \ldots, x_d(\omega))^\top : (\Omega, \mathcal{F}) \to (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ is said to be a random vector if all elements x_i are random variables.

We consider x as an integrable real-valued random variable with respect to the probability measure $\mathbb{P}$. The expression

$$\mathbb{E}[x] = \int_\Omega x(\omega)\, d\mathbb{P}(\omega),$$

denotes the expected value of x under the measure $\mathbb{P}$. If the law of $x(\omega)$ is absolutely continuous with respect to Lebesgue measure on $\mathbb{R}$, then there exists a probability density function f such that the cumulative distribution function satisfies

$$F(x) = \int_{-\infty}^{x} f(t)\mathrm{d}t$$

and in that case

$$\mathbb{E}[x] = \int_{-\infty}^{\infty} x f(x)\mathrm{d}x.$$

Furthermore, the expression

$$\mathrm{Var}(x) = \mathbb{E}\left[(x - \mathbb{E}[x])^2\right] = \int_{-\infty}^{\infty} (x - \mathbb{E}[x])^2 f(x)\mathrm{d}x$$

represents the variance of x, assuming the existence of all relevant integrals within this context. The p-th moment of x, for $p > 0$, is defined as $\mathbb{E}\left[|x|^p\right]$ if $\mathbb{E}\left[|x|^p\right] < \infty$. Given another real-valued random variable y, the covariance between x and y is given by

$$\mathrm{Cov}(x, y) = \mathbb{E}\left[(x - \mathbb{E}[x])\,(y - \mathbb{E}[y])\right],$$

x and y are said to be uncorrelated if $\mathrm{Cov}(x, y) = 0$. Note that uncorrelated random variables need not be independent. However, for jointly Gaussian random variables, zero covariance does imply independence. This fact will be used later in the characterization of Brownian motion

1.2 Stochastic Processes

Definition 1.4 A collection $\{x_t\}_{t \in I}$ of random variables taking values in $\mathbb{R}^d$ is referred to as a stochastic process, where I represents the index set and $\mathbb{R}^d$ denotes the state space. In most cases, the index set I is taken to be the non-negative real line, $\mathbb{R}_+ = [0, \infty)$, though it could also be an interval $[a, b]$, the nonnegative integers,

or even subsets of $\mathbb{R}^d$. For each fixed $t \in I$, the process defines a random variable

$$\Omega \ni \omega \mapsto x_t(\omega) \in \mathbb{R}^d.$$

Alternatively, for each fixed $\omega \in \Omega$, one obtains a function

$$I \ni t \mapsto x_t(\omega) \in \mathbb{R}^d,$$

which is termed a sample path of the process, for fixed ω often denoted as $x.(\omega)$. In some cases, it is more practical to express $x(t, \omega)$ in place of $x_t(\omega)$, viewing the stochastic process as a mapping from $I \times \Omega$ into $\mathbb{R}^d$.

Remark 1.5 Instead of $x(t)$, it might happen during this thesis that we also write x_t.

Definition 1.6 Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. A filtration (on $(\Omega, \mathcal{F}, \mathbb{P})$) is a family $\{\mathcal{F}_t\}_{t\in[0,T]}$ of σ-algebras $\mathcal{F}_t \subset \mathcal{F}$ such that

$$0 \leq s < t \Rightarrow \mathcal{F}_s \subset \mathcal{F}_t \quad \text{for} \quad s, t \in [0, T]$$

which means that $\{\mathcal{F}_t\}$ is an increasing family. The filtration is said to be right continuous if $\mathcal{F}_t = \bigcap_{s>t} \mathcal{F}_s$ for all $t \geq 0$.

Definition 1.7 The process $x(t)$, where t belongs to the interval $[0, T]$, is considered non-anticipating (or adapted to $\mathcal{F}_t$) if and only if $x(t)$ is measurable with respect to $\mathcal{F}_t$. Often we use the natural filtration, that is, $\mathcal{F}_t = \sigma\{x(s) : s \leq t\}$.

1.2.1 Markov Processes

A stochastic process can be classified as a Markov process if the evolution of an event is solely determined by its current state, without considering its past history. Hence, knowledge of the event's past is unnecessary. This feature is applicable in other types of processes, including the Wiener process, which is the stochastic process used in this study. The Markov process is now explicitly defined.

Definition 1.8 Consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a stochastic process $x(t), t \in [0, T]$. The process $x(t)$ is defined as a Markov process if, for any sequence of times $t_1 < t_2 < \ldots < t_{n+1}$ where each $t_i \in [0, T]$ and for any sequence of real numbers $x_1, x_2, \ldots, x_{n+1}$, the specified conditional probability satisfies

$$\begin{aligned} &\mathbb{P}\Big(x(t_{n+1}) < x_{n+1} \Big| x(t_n) = x_n, x(t_{n-1}) = x_{n-1}, \ldots, x(t_1) = x_1\Big) \\ &\quad = \mathbb{P}\Big(x(t_{n+1}) < x_{n+1} \Big| x(t_n) = x_n\Big) \end{aligned} \tag{1.1}$$

In the continuous case, we can also consider the Markov property with respect to a

filtration $\{\mathcal{F}_t\}_{t\in[0,T]}$:

$$\mathbb{P}(x(t) \in B|\mathcal{F}_s) = \mathbb{P}(x(t) \in B|x(s)), \tag{1.2}$$

where $\mathcal{F}_t := \sigma\{x(s)|0 \leq s \leq t\}$ for all $t \in [0,T]$ and B is a Borel set, then (1.2) is equivalent to (1.1). In fact, a Markov process has the property that the past and future are independent when the present is known.

1.2.2 Martingales

Martingales represent a particular class of stochastic processes, often described as "fair games" due to their defining property that the conditional expected future value equals the current value.

Definition 1.9 A stochastic process $x = \{x(t)\}_{t\in[0,T]}$ on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with respect to the filtration $\{\mathcal{F}_t\}_{t>0}$ is a martingale if it satisfies the following conditions:

1. For every $t \in [0,T]$, $\mathbb{E}(|x(t)|) < \infty$.
2. For every $t \in [0,T]$, $x(t)$ is $\mathcal{F}_t$-measurable.
3. For every $0 \leq s < t$, the following property holds with probability one

$$\mathbb{E}(x(t)|\mathcal{F}_s) = x(s), \tag{1.3}$$

In the case of a discrete-time martingale process $x_0, x_1, x_2, \ldots$, the identity (1.3) reduced to

$$E(x_n \mid x_{n-1}) = x_{n-1}, \quad \text{w.p.1},$$

for $n = 1, 2, 3, \ldots$ Interpreted in a gambling context, this means that the expected winnings of the next game, conditioned on the knowledge of the winnings from games up to the present, equals the winnings of the current game.

1.2.3 Standard Brownian Motion

A stochastic process that captures significant interest is the sBm, commonly referred to as the Wiener process in much of the stochastic processes literature. This process plays a crucial role in the theory of SDEs and Itô calculus. We then define this process and outline its characteristics.

Definition 1.10 Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space with filtration $\{\mathcal{F}_t\}_{t\geq 0}$. A (standard) one-dimensional Brownian motion (or Wiener process) is a real-valued continuous $\{\mathcal{F}_t\}$ adapted process $\{W(t)\}_{t\geq 0}$ with the following properties:

1. $W(0) = 0 \quad a.s.$,
2. The process $W(t),\ t \geq 0$ has independent and stationary increments,

3. The increment $W(t)-W(s)$ follows a normal distribution with mean zero and variance $t-s$, in summary,

$$W(t) - W(s) \sim \mathcal{N}(0, t-s).$$

The term "independent increments" means that for any selection of non-negative real numbers:

$$0 \leq s_1 \leq t_1 \leq s_2 \leq t_2 \leq \ldots \leq s_n \leq t_n < \infty,$$

the random increments:

$$W(t_1) - W(s_1), W(t_2) - W(s_2), \ldots, W(t_n) - W(s_n)$$

are independent and "stationary increments" imply that for any $0 < s, t < \infty$, the distribution of the increment $W(t+s) - W(s)$ is the same as $W(t) - W(0)$. For further use (see Section 2.2), we also introduce the two-sided sBm $W = (W(t))_{t\in\mathbb{R}}$ as

$$W(t) = \begin{cases} W_1(t) & \text{if } t \geq 0, \\ W_2(-t) & \text{if } t < 0, \end{cases} \tag{1.4}$$

where $W_1(t)$ and $W_2(-t)$ are two independent (one-sided) standard Brownian motions.

In the context of sBm, from now on we shall consider the following filtration.

Definition 1.11 Suppose that $W(t,\omega)$ is a n-dimensional Brownian motion. Then, we define the natural filtration $\mathcal{F}_t := \mathcal{F}_t^{(n)}$ as the σ-algebra generated by the random variables $\{W_i(s)\}_{\substack{1\leq i\leq n \\ 0\leq s\leq t}}$. In other words, $\mathcal{F}_t$ is the smallest σ-algebra containing all sets of the form:

$$\{\omega : W(t_1,\omega) \in \mathcal{B}_1, \ldots, W(t_k,\omega) \in \mathcal{B}_k\},$$

where $t_j \leq t$ for $j \leq k = 1, 2, \ldots$ and $\mathcal{B}_j \subset \mathbb{R}^n$ are Borel sets.

1.3 Stochastic Integrals

In this section we introduce stochastic integrals with respect to Brownian motion and use them to formalize the notion of a solution to a stochastic differential equation (SDE). Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[0,T]}, \mathbb{P})$ be a filtered probability space, where the filtration is given by

$$\mathcal{F}_t := \sigma\{W(s) : 0 \leq s \leq t\},$$

the natural filtration of a one-dimensional Wiener process $W = \{W(t)\}_{t\in[0,T]}$, augmented in the usual way to satisfy the standard hypotheses (completeness and right-continuity).

Let $f : [0,T] \times \mathbb{R} \to \mathbb{R}$ and $g : [0,T] \times \mathbb{R} \to \mathbb{R}$ be Borel-measurable functions. We

consider the following SDE

$$\mathrm{d}x(t) = f(t, x(t))\,\mathrm{d}t + g(t, x(t))\,\mathrm{d}W(t), \qquad t \in [0, T], \tag{1.5}$$

where $x = \{x(t)\}_{t\in[0,T]}$ is an $\mathbb{R}$-valued stochastic process with almost surely continuous sample paths.

A process $x = \{x(t)\}_{t\in[0,T]}$ is said to be a (strong) solution of the SDE (1.5) with initial condition $x(0) = x_0$ if

$$x(t) = x_0 + \int_0^t f(s, x(s))\,\mathrm{d}s + \int_0^t g(s, x(s))\,\mathrm{d}W(s), \qquad t \in [0, T], \tag{1.6}$$

where, function f is referred to as the *drift*, and g is referred to as the *diffusion coefficient*.

While the first integral in equation (1.6) is a standard Lebesgue integral, the second integral presents a significant challenge due to the fact that Wiener process paths are nowhere of bounded variation. Consequently, the usual Riemann-Stieltjes or Lebesgue integration techniques are not directly applicable, necessitating the development of stochastic integration theory. It is important to note that the Wiener process can be replaced by other processes satisfying the martingale property to define generalized stochastic integrals.

In order to define the above stochastic integrals, we first need to introduce some necessary concepts.

1.3.1 Stochastic Integral for Simple Functions

We begin by considering indicator functions, which serve as the building blocks for constructing stochastic integrals. Define the indicator function of an interval $[0, T]$ as follows:

$$\chi_{[0,T]}(t) = \begin{cases} 1 & \text{if } 0 \le t \le T, \\ 0 & \text{otherwise.} \end{cases} \tag{1.7}$$

Then, the stochastic integral of $\chi_{[0,T]}(t)$ with respect to $W(t)$ is given by

$$\int_0^\infty \chi_{[0,T]}(t)\mathrm{d}W(t) = \int_0^T \chi_{[0,T]}(t)\mathrm{d}W(t) := W(T) - W(0).$$

Now, consider the partition $0 = t_0 < t_1 < \ldots < t_n = T$ of the interval $[0, T]$, with the partition width $\Delta t_i = t_{i+1} - t_i$. We define a deterministic function $f(t)$ as a step function if

$$f(t) = \sum_{i=0}^{n-1} f(t_i)\chi_{[t_i,t_{i+1}]}(t),$$

and, if the integral exists, we define the integral of $f(t)$ with respect to $W(t)$ as the Itô integral.

Definition 1.12 Let $\delta = \max_i \Delta t_i > 0$ be the largest step size in the partition. The Itô integral of the step function $f(t)$ is defined as:

$$\int_0^T f(t)\mathrm{d}W(t) := \lim_{\delta \to 0} \sum_{i=0}^{n-1} f(t_i) \left[W(t_{i+1}) - W(t_i)\right],$$

where the function f is evaluated at the left endpoint, specifically at t_i, for each subinterval $[t_i, t_{i+1}]$. The resulting stochastic integral $\int_0^T f(t)\mathrm{d}W(t)$ is $\mathcal{F}_T$-measurable and belongs to the space $L^2(\Omega; \mathbb{R})$.

Remark 1.13 For the Stratonovich integral, which is denoted by $\int_0^T f(t) \circ \mathrm{d}W(t)$, the conditions remain similar, with the only difference being that the function $f(t)$ is evaluated at the midpoint of the interval, i.e.

$$\int_0^T f(t) \circ \mathrm{d}W(t) := \lim_{\delta \to 0} \sum_{i=0}^{n-1} f\left(\frac{t_i + t_{i+1}}{2}\right) \left[W(t_{i+1}) - W(t_i)\right].$$

Before discussing the properties of the Itô integral, it is necessary first to define second-moment stochastic functions.

Definition 1.14 Consider $\mathcal{M}^2([0,T];\mathbb{R})$ as the space of all real-valued $\{\mathcal{F}_t\}$-adapted processes $f = \{f(t)\}_{0 \le t \le T}$ such that

$$\|f\|_T^2 = \mathbb{E} \int_0^T |f(t)|^2 dt < \infty.$$

Definition 1.15 A real-valued stochastic process $g = \{g(t)\}_{0 \le t \le T}$ is called a simple (or step) process if there exists a partition $0 = t_0 < t_1 < \ldots < t_k = T$ of the interval $[0,T]$ and a sequence of bounded, $\mathcal{F}_{t_i}-$measurable random variables ξ_i for $0 \le i \le k-1$, such that

$$g(t) = \xi_0 \chi_{[t_0,t_1]}(t) + \sum_{i=1}^{k-1} \xi_i \chi_{(t_i,t_{i+1}]}(t). \tag{1.8}$$

Let $\mathcal{M}_0([0,T];\mathbb{R})$ represent the collection of all such processes. It is clear that $\mathcal{M}_0([0,T];\mathbb{R}) \subset \mathcal{M}^2([0,T];\mathbb{R})$. We proceed to define the Itô integral for these simple processes.

Definition 1.16 For a simple process $g \in \mathcal{M}_0([0,T];\mathbb{R})$, the Itô integral is defined as:

$$\int_a^b g(t)\mathrm{d}W(t) = \sum_{i=0}^{k-1} \xi_i \left(W(t_{i+1}) - W(t_i)\right).$$

This integral is $\mathcal{F}_T$-measurable and belongs to $L^2(\Omega;\mathbb{R})$.

Lemma 1.17 [70] For any simple process $g \in \mathcal{M}_0([0,T];\mathbb{R})$

$$\mathbb{E}\left(\int_0^T g(t)\mathrm{d}W(t)\right) = 0,$$
$$\mathbb{E}\left|\int_0^T g(t)\mathrm{d}W(t)\right|^2 = \mathbb{E}\int_0^T |g(t)|^2\mathrm{d}t.$$

Using this property, we can extend the definition of Itô integral from simple processes to all processes in $\mathcal{M}^2([0,T];\mathbb{R})$.

1.3.2 Stochastic Integrals for General Processes

Definition 1.18 For $f \in \mathcal{M}^2([0,T];\mathbb{R})$, define the Itô integral as

$$\int_0^T f(t)\mathrm{d}W(t) = \lim_{n\to\infty}\int_0^T g_n(t)\mathrm{d}W(t) \quad \text{in } L^2(\Omega;\mathbb{R}),$$

where $\{g_n\}$ is a sequence of simple process such that

$$\lim_{n\to\infty}\mathbb{E}\int_a^b |f(t)-g_n(t)|^2\mathrm{d}t = 0.$$

In the similar manner, we can define the Stratonovich integral for $f \in \mathcal{M}^2([0,T];\mathbb{R})$.

The following lemma establishes key properties of the Itô integral, linking it to the Lebesgue integral.

Lemma 1.19 [70](Properties of the Itô Integral) Let $f, g \in \mathcal{M}^2([0,T];\mathbb{R})$ and α, β be two real numbers. Then, the Itô integral satisfies the following properties:

(i) The operator is linear, meaning that

$$\int_0^T [\alpha f(t) + \beta g(t)]\mathrm{d}W(t) = \alpha\int_0^T f(t)\mathrm{d}W(t) + \beta\int_0^T g(t)\mathrm{d}W(t).$$

(ii) The expectation of the Itô integral is zero:

$$\mathbb{E}\int_0^T f(t)\mathrm{d}W(t) = 0.$$

(iii) The Itô integral $\int_0^T f(t)\mathrm{d}W(t)$ is $\mathcal{F}_T$-measurable.

(iv) If both $\int_0^T f(t)\mathrm{d}W(t)$ and $\int_0^T g(t)\mathrm{d}W(t)$ exist, then

$$\mathbb{E}\left(\int_0^T f(s)\mathrm{d}W(s)\int_0^T g(\tau)\mathrm{d}W(\tau)\right) = \int_0^T \mathbb{E}\left(f(t)g(t)\right)\mathrm{d}t.$$

In particular, the Itô isometry holds:

$$\left\| \int_0^T f(t) \mathrm{d}W(t) \right\|_T^2 = \mathbb{E}\left(\int_0^T f(t) \mathrm{d}W(t) \right)^2 = \int_0^T \mathbb{E}\left(|f(t)|^2 \right) \mathrm{d}t$$

(v) If $s, t \in [0, T]$ with $s \leq t$, we have

$$\mathbb{E}\left(\int_0^s f(\tau) \mathrm{d}W(\tau) \int_0^t g(\tau) \mathrm{d}W(\tau) \right) = \int_0^s \mathbb{E}\left(f(\tau) g(\tau) \right) d\tau,$$

In particular,

$$\mathbb{E}\left(\int_0^s f(\tau) \mathrm{d}W(\tau) \int_0^t f(\tau) \mathrm{d}W(\tau) \right)^2 = \int_0^s \mathbb{E}\left(f^2(\tau) \right) d\tau.$$

For further information, see [60, 70, 86].

Definition 1.20 Let $\mathcal{L}^p\left(\mathbb{R}_+; \mathbb{R}^d\right)$ for $p \geq 1$ represent the set of all $\mathcal{F}_t$-adapted, $\mathbb{R}^d$-valued measurable processes $f = \{f(t)\}_{t \geq 0}$ such that

$$\mathbb{E}\left(\int_0^T \|f(t)\|_2^p \mathrm{d}t \right) < \infty \quad \text{for every } T > 0,$$

where $\| \cdot \|_2$ is the Euclidean norm.

Definition 1.21 A one-dimensional Itô (Stratonovich) process is a continuous, adapted process $x(t)$ for $t \geq 0$ of the form

$$x(t) = x(0) + \int_0^t f(s) \mathrm{d}s + \int_0^t g(s) (\circ) \mathrm{d}W(s), \tag{1.9}$$

where $f \in \mathcal{L}^1\left(\mathbb{R}_+; \mathbb{R}\right)$ and $g \in \mathcal{L}^2\left(\mathbb{R}_+; \mathbb{R}\right)$. The stochastic differential equation is

$$dx(t) = f(t) \mathrm{d}t + g(t) (\circ) \mathrm{d}W(t). \tag{1.10}$$

Definition 1.22 The above definition could be extended to the multi-dimensional case. Let $W(t) = (W_1(t), \ldots, W_q(t))^\top, t \geq 0$ be a q-dimensional Brownian motion defined on the complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ adapted to the filtration $\{\mathcal{F}_t\}_{t \geq 0}$. Then, a d-dimensional Itô (Stratonovich) process is an $\mathbb{R}^d$-valued continuous adapted process $x(t) = (x_1(t), \ldots, x_d(t))^T$ on $t \geq 0$, where $f = (f_1, \ldots, f_d)^\top \in \mathcal{L}^1\left(\mathbb{R}_+; \mathbb{R}^d\right)$ and $g = (g_{ij})_{d \times q} \in \mathcal{L}^2\left(\mathbb{R}_+; \mathbb{R}^{d \times q}\right)$.

Remark 1.23 A straightforward relationship exists between the solutions of Itô and Stratonovich equations. When $x = (x(t))_{t \in [0, \infty)}$ is a solution to the Itô equation

given below:

$$x(t) = x_0 + \int_0^t f(s, x(s))\mathrm{d}s + \int_0^t g(s, x(s))\mathrm{d}W(s), \tag{1.11}$$

where $(W(t))_{t\in[0,\infty)}$ is a one-dimensional Wiener process, it follows that $(x(t))_{t\in[0,\infty)}$ is a solution to the Stratonovich equation presented below (see [121]):

$$x(t) = x_0 + \int_0^t \underline{f}(s, x(s))\mathrm{d}s + \int_0^t g(s, x(s)) \circ \mathrm{d}W(s), \tag{1.12}$$

where $\underline{f}(t,x) = f(t,x) - \frac{1}{2}g(t,x)\frac{\partial g}{\partial x}(t,x)$. Hence, Itô and Stratonovich stochastic equations can be easily transformed to each other. So we can say that the relation between Itô and Stratonovich Integral is defined as follows

$$\int_0^t g(s, x(s)) \circ \mathrm{d}W(s) = \int_0^t g(s, x(s))\mathrm{d}W(s) + \frac{1}{2}\int_0^t g(s, x(s))\frac{\partial g}{\partial x}(s, x(s))\mathrm{d}s \tag{1.13}$$

As shown in Lemma 1.19, a key advantage of Itô integrals over Stratonovich integrals is their compatibility with the basic properties of the Wiener process. The Itô model is often preferred because it does not anticipate future events, making it useful in various fields such as biology [115], option pricing, risk management and modeling financial derivatives [47, 52], as well as in signal processing and noise reduction in stochastic systems [19]. It is also important to note that the equations (1.11) and (1.12) are equivalent when $g(t,x)$ does not depend on x.

The Stratonovich integral is widely used in physical sciences and engineering because it follows the traditional chain rule, making it easier for those in these fields to understand. It is essential in feedback systems and cases where the noise of the system has multiplicative characteristics. This type of integration is commonly applied in areas such as modeling thermal fluctuations, statistical mechanics [35], control theory with non-linear noise effects [70] and population dynamics under environmental changes [75].

Since there is a direct relationship between the Itô and Stratonovich integrals (equation (1.13)), calculations can often be done using either type. A key feature of the Stratonovich integral is that it follows the standard chain rule for transformations, unlike the Itô formula, which includes second-order terms, as shown in Theorems 1.24 and 1.25. This makes the Stratonovich integral more convenient in certain applications. However, unlike Itô integrals, Stratonovich integrals do not form martingales. Despite being less convenient for transformations, the Itô integral has advantages in computation.

1.4 Itô Formula

1.4.1 One Dimensional Itô Formula

One of the fundamental concepts in stochastic analysis, particularly in Itô calculus, is the Itô formula.

Theorem 1.24 [60] Assume that $x(t)$ is an one-dimensional Itô process (see Definition 1.21) generated by:

$$\mathrm{d}x(t) = f(t)\mathrm{d}t + g(t)\mathrm{d}W(t),$$

where $f \in \mathcal{L}^1(\mathbb{R}_+;\mathbb{R})$ and $g \in \mathcal{L}^2(\mathbb{R}_+;\mathbb{R})$. Let $h \in \mathcal{C}^{1,2}(\mathbb{R}_+ \times \mathbb{R};\mathbb{R})$, which means that h is a continuously differentiable function with respect to the first variable and twice continuously differentiable function with respect to the second variable on $\mathbb{R}_+ \times \mathbb{R}$. Then, $h(t,x(t))$ is also an Itô process and the following relation holds almost sure:

$$\begin{aligned} dh(t,x(t)) &= \frac{\partial}{\partial t}h(t,x(t))\,\mathrm{d}t + \frac{\partial}{\partial x}h(t,x(t))\,\mathrm{d}x(t) + \frac{1}{2}\frac{\partial^2}{\partial x^2}h(t,x(t))\,(\mathrm{d}x(t))^2 \\ &= \left(\frac{\partial}{\partial t}h(t,x(t)) + f(t)\frac{\partial}{\partial x}h(t,x(t)) + \frac{1}{2}g^2(t)\frac{\partial^2}{\partial x^2}h(t,x(t))\right)\mathrm{d}t \\ &\quad + g(t)\frac{\partial}{\partial x}h(t,x(t))\,\mathrm{d}W(t). \end{aligned}$$

1.4.2 Multidimensional Itô Formula

Now, we extend the Itô formula to the d-dimensional case.

Theorem 1.25 [60] Let $x(t)$ be a d-dimensional Itô process on $t \geq 0$ (see Definition 1.22) with the stochastic differential

$$dx(t) = f(t)dt + g(t)dW(t)$$

where $f \in \mathcal{L}^1(\mathbb{R}_+;\mathbb{R}^d)$ and $g \in \mathcal{L}^2(\mathbb{R}_+;\mathbb{R}^{d\times q})$. Let $h(t,x) \in \mathcal{C}^{1,2}(\mathbb{R}_+ \times \mathbb{R}^d;\mathbb{R}^p)$. Then, $h(t,x(t))$ is also an Itô process with the stochastic differential given by

$$\begin{aligned} dh(t,x(t)) = \Big[&\frac{\partial}{\partial t}h(t,x(t)) + \frac{\partial}{\partial x}h(t,x(t))f(t) \\ &+ \frac{1}{2}trace\left(g^T(t)\frac{\partial^2}{\partial x^2}h(t,x(t))g(t)\right)\Big]dt + \frac{\partial}{\partial x}h(t,x(t))g(t)dW(t) \quad \text{a.s.} \end{aligned}$$

where

$$\begin{aligned} \mathrm{d}t\mathrm{d}t = 0 \quad \mathrm{d}t\mathrm{d}W_i(t) = 0 \\ \mathrm{d}W_i(t)\mathrm{d}W_i(t) = dt \quad \mathrm{d}W_i(t)\mathrm{d}W_j(t) = 0 \quad \text{if} \quad i \neq j. \end{aligned}$$

Remark 1.26 The q-dimensional Wiener process can be extended as follows:

Consider $W(t) = (W_1(t), \ldots, W_q(t))^\top$, a $\mathbb{R}^q$-valued Wiener process defined on a complete filtered probability space $\big(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P}\big)$. This process has mean zero and a covariance matrix $\mathbf{K} = (k_{ij})_{i,j=1,\ldots,q}$. It satisfies the property $\mathbb{E}[W(t)W(t)^\top] = \mathbf{K}t$ for $t \in [0, T]$, where $T > 0$ represents the terminal time. It is worth mentioning that, often we consider $K = I_q$, where I_q is the identity matrix.

Example 1.27 [86] Let $x(t)$ and $y(t)$ be Itô processes on $\mathbb{R}$. Applying the Itô formula to the function $g(x, y) = x \cdot y$, we obtain the following:

$$\begin{aligned} d(x(t)y(t)) = \; & dg(x(t), y(t)) = \frac{\partial g}{\partial x}(x(t), y(t))dx(t) + \frac{\partial g}{\partial y}(x(t), y(t))dy(t) \\ & + \frac{1}{2}\frac{\partial^2 g}{\partial x^2}(x(t), y(t))(dx(t))^2 + \frac{\partial^2 g}{\partial x \partial y}(x(t), y(t))dx(t)dy(t) \\ & + \frac{1}{2}\frac{\partial^2 g}{\partial y^2}(x(t), y(t))(dy(t))^2 \\ = \; & y(t)dx(t) + x(t)dy(t) + dx(t)dy(t). \end{aligned}$$

From this derivation, we obtain the integrated form:

$$x(t)y(t) = x(0)y(0) + \int_0^t y(s)\, dx(s) + \int_0^t x(s)\, dy(s) + \int_0^t dx(s)\, dy(s),$$

which is also known as the Itô product rule.

1.5 Existence and Uniqueness of Solutions for SDEs

Consider $W(t) = (W_1(t), ..., W_q(t))^\top$ as an q-dimensional Wiener process defined on a filtered probability space $\big(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P}\big)$[1]. Let $0 \le t_0 < T < \infty$ and x_0 be a random variable in $\mathbb{R}^d$, which is $\mathcal{F}_{t_0}$-measurable and satisfies $\mathbb{E}\left[x_0\right]^2 < \infty$. Let $f : [t_0, T] \times \mathbb{R}^d \to \mathbb{R}^d$ and $g : [t_0, T] \times \mathbb{R}^d \to \mathbb{R}^{d\times q}$ be two Borel measurable functions. We consider the following d-dimensional Itô stochastic differential equation:

$$\mathrm{d}x(t) = f(t, x(t))\mathrm{d}t + g(t, x(t))\mathrm{d}W(t) \quad \text{for} \quad t_0 \le t \le T \tag{1.14}$$

with the initial condition $x(t_0) = x_0$. Using Definition 1.22, this equation can be rewritten in its integral form as:

$$x(t) = x_0 + \int_{t_0}^t f(s, x(s))\mathrm{d}s + \int_{t_0}^t g(s, x(s))\mathrm{d}W(s) \quad \text{for} \quad t_0 \le t \le T \tag{1.15}$$

Now, we define the solution of the stochastic integral equation (1.15).

[1] $(\mathcal{F}_t)_{t\in[0,T]}$ shall be right continuous and complete.

Definition 1.28 A stochastic process $\{x(t)\}_{t_0 \le t \le T}$ taking values in $\mathbb{R}^d$ is said to be a solution of equation (1.15) if it satisfies the following conditions:

- $\{x(t)\}$ is continuous and adapted to $\mathcal{F}_t$,
- $\{f(t, x(t))\} \in \mathcal{L}^1([t_0, T]; \mathbb{R}^d)$ and $\{g(t, x(t))\} \in \mathcal{L}^2([t_0, T]; \mathbb{R}^{d\times q})$,
- for each $t \in [t_0, T]$, equation (1.15) holds with probability 1.

Theorem 1.29 [70] Assume that there exist two positive constants C_1 and C_2 such that

(i) (Lipschitz Condition): for all $x, y \in \mathbb{R}^d$ and $t \in [0, T]$

$$\|f(t, x) - f(t, y)\|_2 + \|g(t, x) - g(t, y)\|_2 \le C_1\|x - y\|_2.$$

(ii) (Linear Growth Bound): for all $x \in \mathbb{R}^d$ and $t \in [0, T]$,

$$\|f(t, x)\|_2 + \|g(t, x)\|_2 \le C_2(1 + \|x\|_2).$$

where $\|\cdot\|_2$ is Euclidean norm. Then there exists a unique solution $x(t)$ to equation (1.14) and the solution belongs to $\mathcal{M}^2\left([t_0, T]; \mathbb{R}^d\right)$.

1.6 Convergence of Random Sequences

Consider the scenario where we have a sequence of $\mathbb{R}^d-$valued random variables $\{X_i\}_{i=1}^{\infty}$ and our interest lies in their long-term behavior, specifically whether a $\mathbb{R}^d-$valued random variable X exists as a limit of the sequence X_n in a certain sense. The concept of convergence for such sequences can be classified in various ways, essentially dividing into two categories: a stronger form requiring the sample paths of X_n to approximate those of X closely and a weaker form demanding only that their probability distributions converge.

1.6.1 Strong Convergence

Assuming that all random variables are defined on a shared probability space, we identify three primary modes of convergence within the stronger category:

(i) **Convergence with probability one (w.p.1):**

$$\mathbb{P}\left(\left\{\omega \in \Omega : \lim_{n\to\infty} \|X_n(\omega) - X(\omega)\|_2 = 0\right\}\right) = 1.$$

This mode is also called almost sure convergence.

(ii) **Mean-square convergence:** Let $E\left(\|X_n\|_2^2\right) < \infty$ for $n = 1, 2, \ldots$ and also $E\left(\|X\|_2^2\right) < \infty$,

$$\lim_{n\to\infty} E\left(\|X_n - X\|_2^2\right) = 0.$$

(iii) **Convergence in probability:**

$$\lim_{n\to\infty} \mathbb{P}\left(\{\omega \in \Omega : \|X_n(\omega) - X(\omega)\|_2 \geq \epsilon\}\right) = 0 \text{ for all } \epsilon > 0.$$

Remark 1.30 Convergence in probability is also referred to as stochastic convergence. It is important to note that both almost sure convergence (*i*) and mean-square convergence (*ii*) imply convergence in probability (*iii*).

1.6.2 Weak Convergence

For less strict forms of convergence, it is not necessary to have detailed knowledge of the specific random variables or their respective probability spaces. Instead, understanding their distribution functions suffices. Highlighted below are key types of convergence within this category.

(iv) **Convergence in Distribution:** The criterion for convergence in distribution, also termed as convergence in law, is outlined as follows:

$$\lim_{n\to\infty} F_{X_n}(x) = F_X(x) \quad \text{at every continuity point of } F_X.$$

This concept refers to the convergence of the distribution functions associated with the sequence of random variables X_n to the distribution function of X.

(v) **Weak Convergence:** The definition of weak convergence is given by:

$$\lim_{n\to\infty} \int_{-\infty}^{\infty} f(x) dF_{X_n}(x) = \int_{-\infty}^{\infty} f(x) dF_X(x),$$

applicable for all test functions $f : \mathbb{R} \to \mathbb{R}$, typically continuous functions that become zero outside a certain bounded interval, which might vary based on the specific function.

1.7 Stability Analysis of SDEs

Stability represents a crucial aspect in the analysis of dynamical systems. Specifically, for systems governed by SDEs, various approaches exist to examine their stability, commonly based on probabilistic criteria as outlined in [25, 113]. These include concepts like stochastic stability and moment stability. In the following, we introduce several of these criteria, starting with the concept of stochastic stability.

1.7.1 Stochastic Stability

Also referred to as stability in probability, the definition of stochastic stability is provided as follows:

Definition 1.31 A system described by the SDE (1.14) is called stochastically stable if, for any initial condition x_0, there exists a $\delta > 0$ such that $\|x_0\| < \delta$, ensuring that for every $\epsilon > 0$,

$$\lim_{x_0 \to 0} \mathbb{P}\left(\sup_{t \in [0,\infty)} \|x(t, x_0)\|_2 \geq \epsilon \right) = 0. \tag{1.16}$$

Definition 1.32 The system (1.14) is considered stochastically asymptotically stable, or stable with probability one if it is stable and satisfies

$$\lim_{x_0 \to 0} \mathbb{P}\left(\lim_{t \to \infty} \|x(t, x_0)\|_2 = 0 \right) = 1. \tag{1.17}$$

A more strict criterion than stochastic stability for the SDE (1.14) is also discussed in the following subsection.

1.7.2 Mean Square Stability

Definition 1.33 A system characterized by the SDE (1.14) is called mean square stable if, for any $x_0 \in L^2\left(\Omega, \mathbb{R}^d\right)$, there exists a $\delta > 0$ such that $\|x_0\|_{L^2(\Omega,\mathbb{R}^d)} < \delta$ and for all $\epsilon > 0$ and all $t > 0$, it holds that

$$\|x(t, x_0)\|_{\mathcal{L}^2([t_0,T];\mathbb{R}^d)} \leq \epsilon. \tag{1.18}$$

Remark 1.34 The mean square stability implies stochastic stability.

Definition 1.35 If the system in (1.14) is mean square stable and fulfills

$$\lim_{t \to \infty} \|x(t, x_0)\|^2_{\mathcal{L}^2([t_0,T];\mathbb{R}^d)} = 0, \tag{1.19}$$

then it is termed asymptotically mean square stable.

Definition 1.36 We define equation (1.14) as exponentially mean-square stable, possessing a decay rate of at least $\alpha \geq 0$, provided that there exists a positive constant M such that for all $t > 0$, the following inequality holds

$$\|x(t, x_0)\|_{\mathcal{L}^2([t_0,T];\mathbb{R}^d)} \leq M e^{-\alpha t}. \tag{1.20}$$

1.8 Euler-Maruyama Method for SDEs

In this section, we address the numerical solution of the $d-$dimensional Itô equation (1.14) for the stochastic process $x(t)$, defined over the interval $t \in [0, T]$ with the initial condition $x(0) = x_0$. To solve equation (1.14) numerically, we employ the Euler-Maruyama method. This method represents the first-order approximation in the

Taylor series expansion for stochastic processes. Specifically, the Euler-Maruyama scheme is formulated as:

$$Y_{i+1} = Y_i + f(t_i, Y_i)\Delta t_i + g(t_i, Y_i)\Delta W_i,$$

for $i = 1, \ldots, N-1$, where Y_i denotes the numerical approximation of $x(t)$ in discrete time steps and $\Delta W_i = W(t_{i+1}) - W(t_i)$.

Definition 1.37 Given time steps $\Delta_i = \Delta t_i$ with $\delta = \max_i \Delta_i > 0$ representing the maximum step size, the numerical approximation $Y^{(\delta)}$ of $x(t)$ exhibits strong convergence of order $\gamma > 0$ at time T if there exists a positive constant C, independent of δ and a $\delta_0 > 0$ such that

$$\epsilon(\delta) = \left\| x(T) - Y_T^{(\delta)} \right\|_{L^2(\Omega)} \leq C\delta^\gamma$$

for each $\delta \in (0, \delta_0)$.

In the following theorem, we establish the order of strong convergence, denoted by $\gamma = \frac{1}{2}$, for the Euler-Maruyama method.

Theorem 1.38 ([61, Theorem 10.2.2]) Suppose the conditions

$$\begin{aligned} \|x_0\|_{L^2(\Omega)} &< \infty, \\ \left\| x_0 - Y_0^{(\delta)} \right\|_{L^2(\Omega)} &\leq C_1\delta^{1/2}, \\ \|f(t,x) - f(t,y)\|_2 + \|g(t,x) - g(t,y)\|_2 &\leq C_2\|x-y\|_2, \\ \|f(t,x)\|_2 + \|g(t,x)\|_2 &\leq C_3(1 + \|x\|_2) \end{aligned}$$

and

$$\|f(s,x) - f(t,x)\|_2 + \|g(s,x) - g(t,x)\|_2 \leq C_4(1 + \|x\|_2)|s-t|^{1/2}$$

hold for all $s, t \in [0, T]$ and $x, y \in \mathbb{R}^n$, where constants $C_1, \ldots, C_4$ do not depend on δ. Then, for the Euler-Maruyama approximation $Y^{(\delta)}$,

$$\left\| x(T) - Y_T^{(\delta)} \right\|_{L^2(\Omega)} \leq C_5\delta^{1/2}$$

holds, where the constant C_5 does not depend on δ.

2 Fractional Brownian Motion

It is well known that classical Brownian motion is a random process with stationary increments and self-similarity with an index of 1/2. This continuous Gaussian process exhibits these properties, which are evident in natural phenomena such as the movement of particles in a fluid (Brownian motion) and the fluctuation in financial asset prices.

Given these characteristics, an interesting question arises: can we identify a stochastic process that also exhibits a Gaussian distribution, stationary increments and self-similarity, but with a self-similarity index different from 1/2? Indeed, such a process exists. Originally proposed by Kolmogorov [62] in the early 1940s to describe fluid turbulence. This process, now widely known as fractional Brownian motion (fBm), was later popularized by Mandelbrot and Van Ness [69].

The characterization of fBm is based on the Hurst parameter H, or self-similarity index, which ranges from 0 to 1. This flexibility makes fBm particularly valuable for modeling various phenomena, as the Hurst parameter H can be adjusted to closely align with the empirical data. However, it is important to note the limitations of fBm, especially when the self-similarity index H diverges from 1/2. Under such conditions, fBm does not satisfy the criteria of a semimartingale or a Markov process, which significantly restricts the analytical tools available for problem-solving. Despite these challenges, the exploration of fBm introduces new opportunities for experts in stochastic calculus to develop innovative solutions.

The following chapter will provide an analysis of the properties associated with fBm, drawing on references [18, 30, 83, 84].

2.1 Properties of Fractional Brownian Motion

In this chapter, we consider all random variables within the filtered probability space $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P}\right)$[1] (see Chapter 1 for more details).

A zero-mean Gaussian random process $W^H = \{W^H(t) : t \geq 0\}$ is called a fractional Brownian motion with Hurst parameter $H \in (0, 1)$, if all sample paths are continuous and its covariance function is given by:

$$\mathbb{E}\left[W^H(t)W^H(s)\right] = R_H(t, s) = \frac{1}{2}(s^{2H} + t^{2H} - |t - s|^{2H}). \tag{2.1}$$

FBm exhibits the following properties:

[1] $(\mathcal{F}_t)_{t\in[0,T]}$ shall be right continuous and complete.

- Self-Similarity: The processes $\{W^H(t), t \geq 0\}$ and $\{a^{-H}W^H(at), t \geq 0\}$ for any constant $a > 0$ have the same probability distributions, because

$$\begin{aligned}\mathbb{E}[a^{-H}W^H(at)a^{-H}W^H(as)] &= a^{-2H}\mathbb{E}[W^H(at)W^H(as)] \\ &= \frac{a^{-2H}}{2}\left\{(at)^{2H} + (as)^{2H} - |at-as|^{2H}\right\} \\ &= \frac{1}{2}\left\{t^{2H} + s^{2H} - |t-s|^{2H}\right\} = \mathbb{E}[W^H(t)W^H(s)].\end{aligned} \tag{2.2}$$

- Stationary increments: Equation (2.1) shows that the moments of fBm in the interval $[s, t]$ follow a normal distribution with mean zero and variance is given by

$$\mathbb{E}\left[(W^H(t) - W^H(s))^2\right] = |t-s|^{2H},$$

and for any $k \geq 1$,

$$\mathbb{E}\left[(W^H(t) - W^H(s))^{2k}\right] = \frac{(2k)!}{k!2^k}|t-s|^{2Hk}. \tag{2.3}$$

Definition 2.1 Let $\{x(t)\}_{t\geq 0}$ and $\{y(t)\}_{t\geq 0}$ be two stochastic processes defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We say that $\{y(t)\}_{t\geq 0}$ is a modification of $\{x(t)\}_{t\geq 0}$ if for every $t \geq 0$,

$$\mathbb{P}(\{\omega \in \Omega : x(t,\omega) = y(t,\omega)\}) = 1,$$

meaning that for every $t \geq 0$, $y(t) = x(t)$ almost surely.

Theorem 2.2 (Kolmogorov's Continuity Criterion) [18] Let $\{x(t)\}_{t\in[0,T]}$ be a stochastic process. If for every $T > 0$ and positive constants D, α and β, there exist such values that for all $0 \leq s, t \leq T$,

$$\mathbb{E}(|x(t) - x(s)|^\alpha) \leq D|t-s|^{1+\beta},$$

then $x(t)$ has a continuous modification.

If we do not assume the continuity of all sample paths, we select k in equation (2.3) such that $2Hk > 1$. Consequently, Kolmogorov's Continuity Criterion ensures that the fBm has a modification with continuous paths. Furthermore, the following lemma, adapted from [36], establishes the Hölder continuity of the fBm paths:

Lemma 2.3 [36] For any $\epsilon > 0$ and $T > 0$, there exists a non-negative random variable $G_{\epsilon,T}$ such that, for all $p \geq 1$, $\mathbb{E}(|G_{\epsilon,T}|^p) < \infty$ and for $s, t \in [0, T]$, the following inequality holds:

$$|W^H(t) - W^H(s)| \leq G_{\epsilon,T}|t-s|^{H-\epsilon}. \tag{2.4}$$

In other words, the paths of fBm are Hölder continuous of order $H - \epsilon$ for $\epsilon > 0$

with $H - \epsilon > 0$.

For $H = 1/2$, the covariance function (2.1) reduces to $R_{1/2}(t,s) = \min\{t,s\}$ and the process $W^{1/2}$ is equivalent to sBm. In this case, the increments of the process in disjoint intervals are independent. However, for $H \neq 1/2$, the increments depend on each other.

If we define $x(n) = W^H(n) - W^H(n-1)$ for $n \geq 1$, then $\{x(n) : n \geq 1\}$ forms a stationary Gaussian sequence with variance 1 and the covariance function is given by:

$$\begin{aligned}
\rho_H(n) &:= Cov(x(k), x(k+n)) \\
&= \mathbb{E}[(W^H(k) - W^H(k-1))(W^H(n+k) - W^H(n+k-1))] \\
&= \mathbb{E}[W^H(k)W^H(n+k)] - \mathbb{E}[W^H(k)W^H(n+k-1)] \\
&\quad - \mathbb{E}[W^H(k-1)W^H(n+k)] + \mathbb{E}[W^H(k-1)W^H(n+k-1)] \\
&= \frac{1}{2}\left\{(n+1)^{2H} + (n-1)^{2H} - 2n^{2H}\right\} \\
&= \frac{1}{2}\{\binom{2H}{0}n^{2H} + \binom{2H}{1}n^{2H-1} + \binom{2H}{2}n^{2H-2} + \dots \\
&\quad + \binom{2H}{0}n^{2H} - \binom{2H}{1}n^{2H-1} + \binom{2H}{2}n^{2H-2} + \dots - 2n^{2H}\} \\
&= \frac{1}{2}\left\{2\binom{2H}{2}n^{2H-2} + \dots\right\} \approx \frac{(2H)!}{2!(2H-2)!}n^{2H-2} \\
&= H(2H-1)n^{2H-2}. \qquad (2.5)
\end{aligned}$$

Therefore, as n approaches infinity, $\rho_H(n) \to 0$. Consequently, for all $n > 0$, if $H > \frac{1}{2}$, then $\rho_H(n) > 0$ and $\sum_{n=1}^{\infty} \rho_H(n) = \infty$. In this case, we say that the sequence $\{x(n) : n \geq 1\}$ exhibits long-range dependence. Due to this long-range dependency property, fBm with a Hurst parameter $H > \frac{1}{2}$ is particularly well-suited for simulating a variety of phenomena. If $H < \frac{1}{2}$, then for all $n > 0$, $\rho_H(n) < 0$ and $\sum_{n=1}^{\infty} |\rho_H(n)| < \infty$.

2.2 Different Stochastic Representation of fBm

In this section, we show that fBm can be represented as a Wiener integral (Itô integral, see Section 1.3) in different ways.

The first representation of fBm (as introduced in [69]) is known as the time representation.

Proposition 2.4 Let $H \in (0, \frac{1}{2}) \cup (\frac{1}{2}, 1)$, set

$$c_H = \sqrt{\frac{1}{2H} + \int_0^{\infty} \left((1+u)^{H-\frac{1}{2}} - u^{H-\frac{1}{2}}\right)^2 du} < \infty,$$

and let $W = (W(t))_{t\in\mathbb{R}}$ be a two-sided sBm (see (1.4)). Then, (any continuous modification of) the process $W^H = (W^H(t))_{t>0}$, defined as

$$W^H(t) = \frac{1}{c_H}\left(\int_{-\infty}^{0}\left((t-u)^{H-\frac{1}{2}} - (-u)^{H-\frac{1}{2}}\right)\mathrm{d}W(u) + \int_0^t (t-u)^{H-\frac{1}{2}}\mathrm{d}W(u)\right), \tag{2.6}$$

is a fBm with Hurst parameter H.

The second representation of fBm is the spectral representation (also called harmonizable representation [83]).

Proposition 2.5 Let $H \in (0, \frac{1}{2}) \cup (\frac{1}{2}, 1)$, set

$$d_H = \sqrt{2\int_0^\infty \frac{1-\cos u}{u^{2H+1}}\mathrm{d}u} < \infty,$$

and let $W = (W(t))_{t\in\mathbb{R}}$ be a two-sided sBm (see (1.4)). Then, (any continuous modification of) the process $W^H = (W^H(t))_{t>0}$, defined as

$$W^H(t) = \frac{1}{d_H}\left(\int_{-\infty}^{0}\frac{1-\cos(ut)}{|u|^{H+\frac{1}{2}}}\mathrm{d}W(u) + \int_0^\infty \frac{\sin(ut)}{|u|^{H+\frac{1}{2}}}\mathrm{d}W(u)\right), \tag{2.7}$$

is a fBm with Hurst parameter H.

In conclusion, the following proposition referenced in [27, 82], presents an alternative representation of fBm. This representation shows that fBm can be expressed as a Volterra process, which means that it can be depicted as $W^H(t) = \int_0^t K_H(t,s)\mathrm{d}W(s)$, with $W = \{W(t)\}_{t>0}$ denoting a sBm and K_H being a well-defined square-integrable kernel. (This representation will be used in Chapter 4.)

Proposition 2.6 Let $H \in (0, \frac{1}{2}) \cup (\frac{1}{2}, 1)$ and, for $t > s > 0$, set

$$K_H(t,s) = \begin{cases} \sqrt{\frac{H(2H-1)}{\int_0^1 (1-x)^{1-2H}x^{H-\frac{3}{2}}\mathrm{d}x}}\, s^{\frac{1}{2}-H}\int_s^t (u-s)^{H-\frac{3}{2}}u^{H-\frac{1}{2}}\mathrm{d}u, & \text{if } H > \frac{1}{2} \\ \sqrt{\frac{2H}{(1-2H)\int_0^1 (1-x)^{-2H}x^{H-\frac{1}{2}}\mathrm{d}x}} & \\ \times\left[\left(\frac{t}{s}\right)^{H-\frac{1}{2}}(t-s)^{H-\frac{1}{2}} - (H-\frac{1}{2})s^{\frac{1}{2}-H}\int_s^t u^{H-\frac{3}{2}}(u-s)^{H-\frac{1}{2}}\mathrm{d}u\right] & \text{if } H < \frac{1}{2} \end{cases}$$

Let $W = (W(t))_{t\geq 0}$ be a sBm and define $W^H = (W^H(t))_{t\geq 0}$ by

$$W^H(t) = \int_0^t K_H(t,s)\mathrm{d}W(s). \tag{2.8}$$

Then, (any continuous modification of) W^H is a fBm with Hurst parameter H.

2.3 Simulation of Fractional Brownian Motion

This section is dedicated to simulating fBms. In this section, we will explain the Cholesky decomposition method to simulate fBms. As we know, in the case of $H = 1/2$, fBm is identical to sBm; hence, this process has independent increments. However, for $H \neq 1/2$, the increments of the process are dependent. As shown in the previous section, if we define $x(k) = W^H(k) - W^H(k-1)$, then the covariance between $x(k)$ and $x(k+n)$ is given by

$$Cov(x(k), x(k+n)) = \frac{1}{2}\big[(n+1)^{2H} - 2n^{2H} + (n-1)^{2H}\big]. \tag{2.9}$$

Therefore, the primary task in simulating fBms is to generate the increments of this process. To achieve this, consider the vector

$$Z = (W^H(1), W^H(2) - W^H(1), W^H(3) - W^H(2), \ldots, W^H(N) - W^H(N-1)).$$

Vector Z follows a normal distribution with mean zero and covariance matrix Σ, i.e., $Z \sim \mathcal{N}(0, \Sigma)$, where

$$\begin{aligned} \Sigma_{i,j} &= \mathbb{E}\left(W^H(i+1) - W^H(i), W^H(j+1) - W^H(j)\right) \\ &= \frac{1}{2}\big[(j-i-1)^{2H} - 2(j-i)^{2H} + (j-i+1)^{2H}\big], \quad i,j = 0, \ldots, N-1. \end{aligned} \tag{2.10}$$

Now, to simulate Z, let C be an $N \times N$ matrix and vector $V = (v_1, \ldots, v_N)^\top$ with $V_i \sim \mathcal{N}(0,1)$ for $i = 1, \ldots, N$. If we find matrix C such that $C^\top C = \Sigma$, then it is clear that $C^\top V \sim \mathcal{N}(0, C^\top C)$ since $C^\top V = \left(\sum_{j=1}^N c_{ji} v_j\right)_{i=1,\ldots,N}$, then

$$\mathbb{E}\left[C^\top V, C^\top V\right] = \begin{pmatrix} \sum_{j=1}^N c_{j1}^2 & \sum_{j=1}^N c_{j1}c_{j2} & \cdots & \sum_{j=1}^N c_{j1}c_{jN} \\ \vdots & \vdots & & \vdots \\ \sum_{j=1}^N c_{j1}c_{jN} & \sum_{j=1}^N c_{j2}c_{jN} & \cdots & \sum_{j=1}^N c_{jN}^2 \end{pmatrix} = C^\top C,$$

by setting $Z = C^\top V$, we successfully simulate Z.

Given that the matrix Σ is a positive definite symmetric matrix, it admits a Cholesky decomposition. This decomposition implies that Σ can be expressed as $\Sigma = LDL^\top$, where L is a lower triangular matrix and D is a diagonal matrix. Consequently, Z can be easily obtained using this decomposition. Figure 2.1 displays the simulated sample paths for three different values of the Hurst parameter H. In this figure, the cumulative sums of the fBm samples are illustrated. Negative correlations are evident for $H = 0.2$, resulting in more irregular sample paths, whereas the paths are smoother for $H = 0.8$ due to positive correlations. From equation (2.5), we can infer that the covariances are negative when $H < 1/2$ and positive when $H > 1/2$. This behavior is also depicted in Figure 2.2, where the increments $x(n) = W^H(n) - W^H(n-1)$ are shown for the same values of H as in Figure 2.1. For $H = 0.2$, the negative correlation results in high variability, while

for $H = 0.8$, the sample path exhibits periods of increase and decrease.

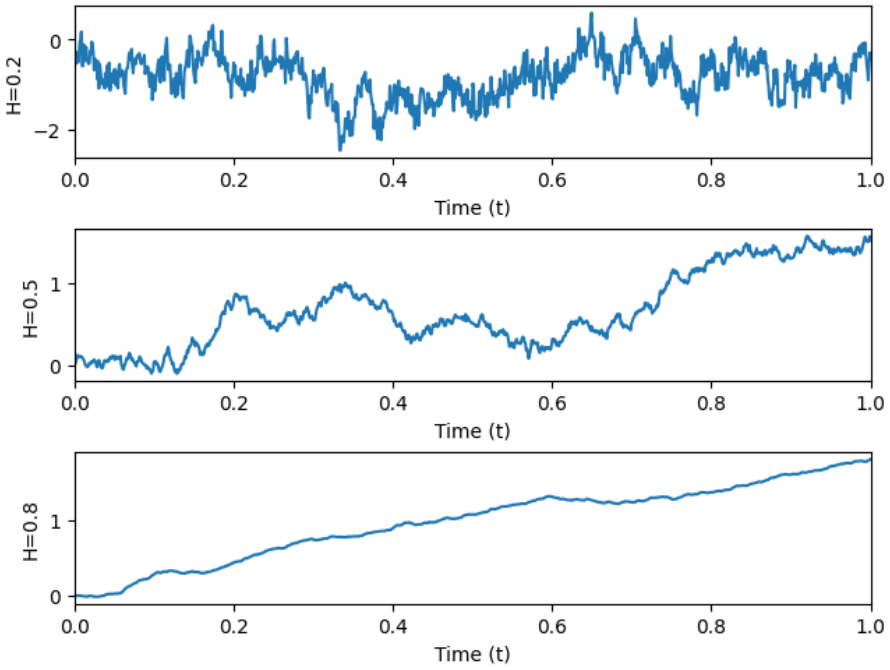

Figure 2.1: Samples of fBm for different Hurst parameters

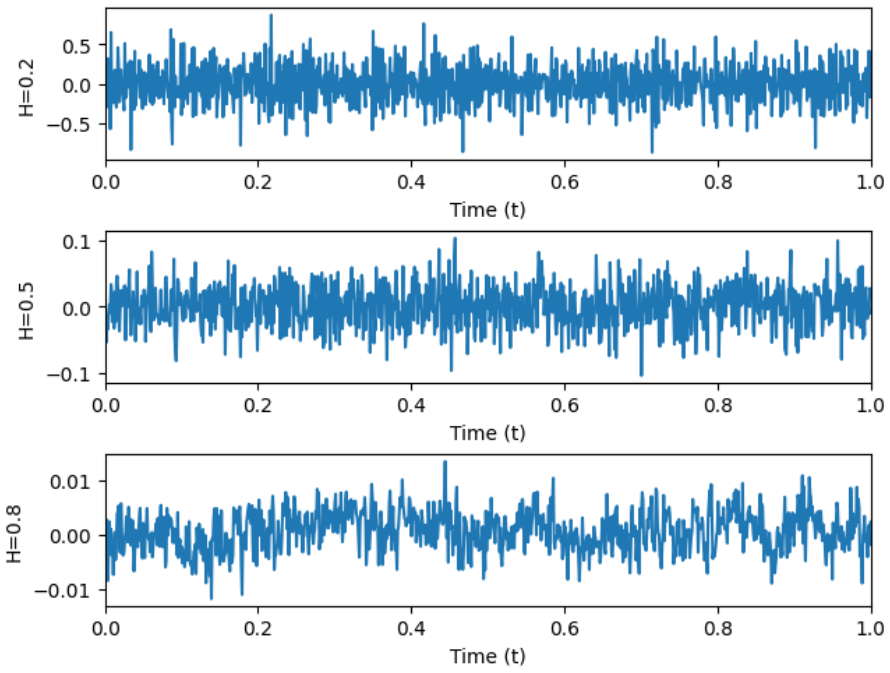

Figure 2.2: Samples of fBm increments for different Hurst parameters.

Remark 2.7 Using the self-similarity property of fBm (2.2), we can simulate the sequence $\left(W^H(t_0), W^H(t_1) - W^H(t_0), \ldots, W^H(t_n) - W^H(t_{n-1})\right)$, where $t_i \in [a, b]$ for $i = 0, 1, \ldots, n$ and $b > a \geq 0$, by applying the simulation method given in equation (2.10).

2.4 Semimartingale and Markov Property

This section explores the asymptotic behavior p-variation of fBm. Furthermore, it will be shown that fBm does not belong to the semimartingale class unless it coincides with sBm, which is the case for $H = \frac{1}{2}$. A process is classified as a semimartingale if it can be represented as a combination of a local martingale and a càdlàg adapted process with locally finite variation (for an authoritative text on semimartingales, see [91]). However, it is worth mentioning that, if $W^H(t)$ is a fBm with Hurst parameter $H \in \left(\frac{3}{4}, 1\right)$ and $W(t)$ is an independent sBm, then the process defined by $M_t = W^H(t) + W(t)$ for $t > 0$ qualifies as a semimartingale. The comprehensive exploration of this phenomenon is available in [23].

Consider a stochastic process with continuous paths denoted as $x = \{x(t), t \geq 0\}$ and let $p > 0$ be a constant. The p-variations of x over the interval $[0, T]$ are defined as the following limit:

$$\mathbb{P} - \lim_{n\to\infty} \sum_{i=0}^{n-1} |x(\frac{(i+1)T}{n}) - x(\frac{iT}{n})|^p,$$

where the expression "$\mathbb{P} - \lim$" indicates the limit in probability.

Remark 2.8 If the p-variation exists and is almost surely nonzero, then for any $q > p$, the q-variation is zero and for $q < p$, the q-variation is infinite.

According to [83], we deduce the following result about the p-variations of fBm.

Corollary 2.9 Let W^H be a fBm of the Hurst parameter $H \in (0, 1)$ and let $p \in [1, +\infty)$. Then, in $L^2(\Omega)$ and as $n \to \infty$, one has

$$\sum_{i=0}^{n-1} \left| W^H(\frac{(i+1)T}{n}) - W^H(\frac{iT}{n}) \right|^p \to \begin{cases} 0, & \text{if } p > \frac{1}{H}, \\ \mathbb{E}\,|G|^p, & \text{if } p = \frac{1}{H}, \\ +\infty, & \text{if } p < \frac{1}{H}. \end{cases} \quad \text{with } G \sim \mathcal{N}(0, 1),$$

Based on Corollary 2.9 and the observations noted in Remark 2.8, it can be concluded that, if $H < \frac{1}{2}$, the quadratic variation is unbounded, while for $H > \frac{1}{2}$, the quadratic variation becomes zero and the 1-variation becomes unbounded.

We have now established that fBm typically does not fall into the category of semimartingales, with the only exception being when the Hurst parameter is equal to $\frac{1}{2}$. This particularity makes integration in relation to it a challenging and significant issue.

Theorem 2.10 [106] Let W^H be a fBm of Hurst index $H \in (0, 1/2) \cup (1/2, 1)$.

Then W^H is not a semimartingale.

For fBm, we have the following result about the Markov property.

Theorem 2.11 [83] Let W^H be a fBm of Hurst index $H \in (0, 1/2) \cup (1/2, 1)$. Then W^H is not a Markov process (see Definition 1.8).

2.5 Stochastic Calculus with respect to fBm

One of the key challenges in working with fBm is the definition and interpretation of stochastic integrals with respect to fBm. Traditional Itô calculus, which is well-suited for integration with respect to sBm, does not extend naturally to fBm due to its non-Markovian nature and long-range dependencies. The objective of stochastic calculus lies in the establishment of stochastic integrals in the form of

$$\int_0^T u(t) \mathrm{d}W^H(t),$$

where $u = \{u(t), t \in [0, T]\}$ represents a stochastic process. From this point on, our focus will be on fBm in the scenario where $H > \frac{1}{2}$. This is because, in Chapter 6, our analysis and discussions will be specifically related to this case.

2.5.1 Stochastic Integration of Deterministic Processes

If u is a deterministic function, a general procedure is available to define the stochastic integral of u concerning a Gaussian process, employing convergence within $L^2(\Omega)$. We shall commence by examining this general approach, specifically in the context of fBm.

Let us consider an one-dimensional fBm $W^H = \{W^H(t), t \geq 0\}$ where the Hurst parameter $H \in (\frac{1}{2}, 1)$. Consider a fixed time interval $[0, T]$ and denote the set of step functions defined on this interval as ε. The integration of a step function, denoted as $\phi(t) = \sum_{i=0}^{n-1} a_j \chi_{[t_i, t_{i+1}]}(t)$, can be expressed in a straightforward way as follows:

$$\int_0^T \phi(t) \mathrm{d}W^H(t) = \sum_{i=0}^{n-1} a_j \left(W^H(t_{i+1}) - W^H(t_i) \right),$$

where $\chi_{[t_i, t_{i+1}]}(t)$ is the indicator function defined in equation (1.7). Our objective is to expand the definition of this integral by including a wider range of functions, using convergence within the space $L^2(\Omega)$. In order to achieve this goal, we proceed by introducing the Hilbert space $\mathcal{H}$ (see Appendix A.2), which is defined as the closure of ε with respect to the scalar product as:

$$\langle \chi_{[0,t]}, \chi_{[0,s]} \rangle_{\mathcal{H}} = R_H(t, s),$$

where $R_H(t, s)$ defined in equation (2.1). The expression for the second partial

derivative of the covariance function (2.1), represented as

$$\frac{\partial^2 R_H}{\partial t \partial s} = \alpha_H |t-s|^{2H-2},$$

where α_H is defined as $H(2H-1)$ and it should be emphasized that this equation is integrable. Therefore,

$$R_H(t,s) = \alpha_H \int_0^t \int_0^s |r-u|^{2H-2} \mathrm{d}u\mathrm{d}r. \tag{2.11}$$

Formula (2.11) suggests that the scalar product within the Hilbert space $\mathcal{H}$ can be written as:

$$\langle \phi, \psi \rangle_{\mathcal{H}} = \alpha_H \int_0^T \int_0^T |r-u|^{2H-2} \phi(r)\, \psi(u) \mathrm{d}u\mathrm{d}r, \tag{2.12}$$

this holds true for any pair of step functions ϕ and ψ belonging to ε.

Consequently, it is possible to establish a definition for a linear subspace of functions that are included within the Hilbert space $\mathcal{H}$ in the following way. Let $|\mathcal{H}|$ be the space of measurable functions $\phi : [0,T] \to \mathbb{R}$ satisfying the condition

$$\|\phi\|^2_{|\mathcal{H}|} = \alpha_H \int_0^T \int_0^T |r-u|^{2H-2} |\phi(r)||\phi(u)| \mathrm{d}u\mathrm{d}r < \infty.$$

The incompleteness of the space $|\mathcal{H}|$ with the inner product $\langle ., . \rangle_{\mathcal{H}}$, together with its isometric embedding into a subspace of $\mathcal{H}$, has been established and documented in reference [90].

The following lemma, which was proposed in [72], offers an additional estimate.

Lemma 2.12 Let $H > \frac{1}{2}$ and $\phi \in L^{\frac{1}{H}}([0,T])$, it holds that

$$\|\phi\|_{|\mathcal{H}|} \leq b_H \|\phi\|_{L^{\frac{1}{H}}([0,T])}$$

where b_H is a constant.

As a result, we may witness the embeddings

$$L^2([0,T]) \subset L^{\frac{1}{H}}([0,T]) \subset |\mathcal{H}| \subset \mathcal{H}.$$

Moreover, it is possible to establish the Wiener-type integral $\int_0^T \phi(t) \mathrm{d}W^H(t)$ for functions ϕ that belong to the space $|\mathcal{H}|$. It is important to note that an additional range of functions can be incorporated in contrast to the scenario of sBm. Within this particular framework, the isometry property of the Itô stochastic integral is substituted by the use of the formula:

$$\mathbb{E}\left[\left(\int_0^T \phi(t) \mathrm{d}W^H(t)\right)^2\right] = \alpha_H \int_0^T \int_0^T |r-u|^{2H-2} \phi(r)\, \phi(u) \mathrm{d}u\mathrm{d}r = \|\phi\|^2_{\mathcal{H}}.$$

2.5.2 Stochastic Integration of Random Processes

The literature has presented numerous ways for integrating random processes with respect to fBm. In the case where $H > \frac{1}{2}$, the stochastic integral $\int_0^T u(t)\mathrm{d}W^H(t)$ can be defined using a pathwise approach, taking advantage of the results established by Young [123]. He defined it as a limit in L^2 of the Riemann sums. Another approach involves the application of Malliavin calculus techniques to construct stochastic calculus for fBm. The beginnings of this technique can be traced back to the influential research conducted by Decreusefond and Üstünel [27]. Other relevant sources supporting this methodology include [1, 2, 18, 21, 44], along with numerous other scientific references. In the following discussion, we will introduce a pathwise approach to define Young integration.

2.5.2.1 Young Integration

The development of the definition of $\int_0^T u(t)\mathrm{d}W^H(t)$ can be achieved by employing a pathwise approach involving Riemann-Stieltjes integrals, considering the insights provided by the Young study [123].

Let us consider $T > 0$ to be the endpoint of our time period. Henceforth, we assume that all functions under consideration are defined over the interval $[0, T]$. For a given $\alpha \in [0, 1]$, $\mathcal{C}^\alpha$ denotes the class of functions $f : [0, T] \to \mathbb{R}$ that are Hölder continuous with index α, i.e., functions f that meet the criterion

$$|f|_\alpha = \sup_{0\le s<t\le T} \frac{|f(t) - f(s)|}{(t-s)^\alpha} < \infty.$$

Furthermore, we define $\|f\|_\infty := \sup_{t\in[0,T]} |f(t)|$ and endow $\mathcal{C}^\alpha$ with the norm

$$\|f\|_\alpha = \|f\|_\infty + |f|_\alpha.$$

Definition 2.13 Suppose that $f \in \mathcal{C}^\alpha$ and $g \in \mathcal{C}^\beta$, where $\alpha + \beta > 1$. Given a sequence $(t_i^n)_{i=0}^{k_n}$ of partitions of $[0, T]$ with $\lim_{n\to\infty} \max_{i=0}^{k_n-1}\{t_{i+1}^n - t_i^n\} = 0$. Then, the Young integral $\int_0^T f(s)dg(s)$ is then defined as

$$\int_0^T f(s)dg(s) := \lim_{n\to\infty} \sum_{i=0}^{k_n-1} f(t_i^n)\big[g(t_{i+1}^n) - g(t_i^n)\big].$$

It should be noted that the existence of the Young integral was established in [123]. Therefore, if $u = \{u(t), t \in [0, T]\}$ represents a stochastic process with trajectories that exhibit $\gamma-$Hölder continuity, where $\gamma > 1 - H$, then the Riemann-Stieltjes integral $\int_0^T u(t)\mathrm{d}W^H(t)$ exists pathwise. This implies that for any specific element $\omega \in \Omega$, the integral $\int_0^T u(t, \omega)\mathrm{d}W^H(t, \omega)$ exists as the path-wise limit of the Riemann sums.

In particular, when $H > \frac{1}{2}$, as the paths of W^H are a.s. Hölder continuous with $\alpha = H - \epsilon$ and considering $F \in \mathcal{C}^2$, then the following change of variables formula

is valid:

$$F(W^H(t)) = F(0) + \int_0^t F'(W^H(s))\mathrm{d}W^H(s) \tag{2.13}$$

Furthermore, The Young integral is consistent with the following chain rule.

Theorem 2.14 [83] Suppose $\phi : \mathbb{R}^2 \to \mathbb{R}$ is a $\mathcal{C}^2$ function and consider f, g belonging to $\mathcal{C}^\alpha$ for some $\alpha \in \left(\frac{1}{2}, 1\right]$. It is established that the integrals $\int_0^\cdot \frac{\partial\phi}{\partial f}(f(u), g(u))\, df(u)$ and $\int_0^\cdot \frac{\partial\phi}{\partial g}(f(u), g(u))\, dg(u)$ are well-defined Young integrals. Furthermore, for any $t \in [0, T]$, it holds that

$$\phi(f(t), g(t)) = \phi(f(0), g(0)) + \int_0^t \frac{\partial\phi}{\partial f}(f(u), g(u))\, df(u) + \int_0^t \frac{\partial\phi}{\partial g}(f(u), g(u))\, dg(u). \tag{2.14}$$

$H = 1/2$ represents the boundary case, in which the Young integration no longer works. For that reason, the probabilistic approach of Stratonovich is chosen in the following way.

Definition 2.15 Let $H = 1/2$ and $(t_i^n)_{i=0}^{k_n}$ a partition like in Definition 2.13. Given a continuous semimartingale Y, we set

$$\int_0^T Y(s) \circ \mathrm{d}W^H(s) := \int_0^T Y(s)\mathrm{d}W^H(s) + \frac{1}{2}[Y, W^H]_T,$$

where the first term is the Itô integral

$$\int_0^T Y(s)\mathrm{d}W^H(s) := \mathbb{P} - \lim_{n\to\infty} \sum_{i=0}^{k_n - 1} Y(t_i^n)\left(W^H(t_{i+1}) - W^H(t_i)\right)$$

and

$$[Y, W^H]_T := \mathbb{P} - \lim_{n\to\infty} \sum_{i=0}^{k_n - 1} \left(Y(t_{i+1}^n) - Y(t_i^n)\right)\left(W^H(t_{i+1}) - W^H(t_i)\right)$$

is the quadratic covariation.

Let us refer to Chapter 1 for more details concerning the stochastic calculus given $H = 1/2$. The Stratonovich integral can be viewed as the natural extension of Young since the Stratonovich setting still ensures that it has a "classical" chain rule. Moreover, W^H, $H = 1/2$, can be approximated by "smooth" processes $W^{H,\epsilon}$ with bounded variation paths when Stratonovich stochastic differential equations are considered, e.g., $W^{H,\epsilon}$ can be piecewise linear (Wong-Zakai) [54, 110, 121]. Due to these connections and in order to distinguish from the Itô setting, we use the circle notation $\circ\mathrm{d}W^H$ for both the Young and Stratonovich cases. Nevertheless, the Young and Stratonovich differential equations driven by a fBm have important

applications in various fields.

Remark 2.16 [83] Consider W^H as a fBm with Hurst parameter $H \in (0, \frac{1}{2})$. Given the unbounded nature of the quadratic variation of W^H, which follows directly from Corollary 2.9 for $p = 2$, it can be easily verified that the trajectories of W^H do not belong to the space $\bigcup_{\alpha\in(\frac{1}{2},1]} \mathcal{C}^\alpha$. This observation implies that Definition 2.13 is inapplicable to $H < \frac{1}{2}$. Consequently, the Young integral turns out to be an unsuitable option for performing integration concerning a fBm of Hurst parameter less than $\frac{1}{2}$.

In contrast to the Itô stochastic integral with respect to sBm, the Young integral $\int_0^T F(W^H(t)) \circ \mathrm{d}W^H(t)$ does not exhibit zero mean and its variance does not possess a straightforward formula. Now, we aim to clarify how the methods employed in Malliavin calculus facilitate the calculation of the mean and the variance of the integral mentioned above.

2.5.3 Malliavin Calculus for fBm

Consider the fBm $W^H = \{W^H(t), t \geq 0\}$ characterized by the Hurst parameter $H \in (\frac{1}{2}, 1)$. This process, W^H, is Gaussian, allowing us to establish the associated stochastic calculus of variations, often referred to as the Malliavin calculus. Malliavin calculus is an infinite-dimensional differential calculus initially introduced by Malliavin in [68]. The fundamental operators within the Malliavin calculus consist of the derivative operator D^H and its adjoint, the divergence operator δ. For a comprehensive exploration of the Malliavin calculus and its applications within the framework of fBm, we recommend referring to [73, 85, 114].

Let us consider a fixed time interval denoted as $[0, T]$. The set S is defined as a collection of elementary random variables, every single one of them represented as:

$$F = f(W^H(\phi_1), \ldots, W^H(\phi_n)), \tag{2.15}$$

where $n \geq 1$, $f \in \mathcal{C}_p^\infty(\mathbb{R}^n)$ (meaning f and all its partial derivatives are continuous and exhibit polynomial growth order) and $\phi_i \in \mathcal{H}$. The derivative operator D^H associated with an elementary random variable F in the form of (2.15) is defined as the $\mathcal{H}$-valued random variable:

$$D^H F = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(W^H(\phi_1), \ldots, W^H(\phi_n))\phi_i.$$

The following integration-by-parts identity is established as follows.

Lemma 2.17 [84] Let F denote an elementary random variable, as represented in equation (2.15). Then, for every arbitrary element ϕ belonging to the Hilbert space $\mathcal{H}$, the following relationship is valid:

$$\mathbb{E}(\langle D^H F, \phi\rangle_{\mathcal{H}}) = \mathbb{E}(F W^H(\phi)). \tag{2.16}$$

As a result, if F and G are elementary random variables and $h \in \mathcal{H}$, the following relationship is established:

$$\mathbb{E}(G\langle D^H F, h\rangle_{\mathcal{H}}) = \mathbb{E}(-F\langle D^H G, h\rangle_{\mathcal{H}} + FGW^H(h)) \tag{2.17}$$

Formula (2.17) leads to the conclusion that the derivative operator D^H is a closable operator, transforming functions from $L^p(\Omega)$ to $L^p(\Omega;\mathcal{H})$, for any $p \geq 1$. The Sobolev space $\mathbb{D}^{1,p}$ is defined as the closure of S under the norm

$$\|F\|_{1,p} = \left[\mathbb{E}(|F|^p) + \mathbb{E}\left(\|D^H F\|_{\mathcal{H}}^p\right)\right]^{\frac{1}{p}},$$

where $p \geq 1$. $\mathbb{D}^{1,p}$ can be regarded as an infinite-dimensional weighted Sobolev space.

The divergence operator δ serves as the adjoint of the derivative operator. More specifically, we consider a random variable u in $L^2(\Omega;\mathcal{H})$ belonging to the domain of the divergence operator, denoted as $\operatorname{Dom}\delta$, if

$$|\mathbb{E}(\langle D^H F, u\rangle_{\mathcal{H}})| \leq c_u \|F\|_{L^2(\Omega)}$$

for any $F \in S$. In such cases, $\delta(u)$ is defined through the duality relationship:

$$\mathbb{E}(F\delta(u)) = \mathbb{E}(\langle D^H F, u\rangle_{\mathcal{H}}), \tag{2.18}$$

valid for any $F \in \mathbb{D}^{1,2}$. To provide an illustration, we will investigate an elementary $\mathcal{H}$-valued random variable in the form of $u = \sum_{k=1}^m F_k\phi_k$, where $F_k \in \mathbb{D}^{1,2}$ and $\phi_k \in \mathcal{H}$. In the given scenario, the variable u is considered to be within the domain of the divergence operator. By utilizing equation (2.17), it is possible to deduce that the expression for $\delta(u)$ can be derived as:

$$\delta(u) = \sum_{k=1}^m \left(F_k W^H(\phi_k) - \langle D^H F_k, \phi_k\rangle_{\mathcal{H}}\right). \tag{2.19}$$

The term $F_k W^H(\phi_k) - \langle D^H F_k, \phi_k\rangle_{\mathcal{H}}$ is denoted as the Wick product of the random variables F_k and $W^H(\phi_k)$ and is represented as

$$F_k \diamond W^H(\phi_k) = F_k W^H(\phi_k) - \langle D^H F_k, \phi_k\rangle_{\mathcal{H}} \tag{2.20}$$

Using this notation, equation (2.19) can be modified as:

$$\delta(u) = \sum_{k=1}^m F_k \diamond W^H(\phi_k).$$

The notation $\delta(u) = \int_0^T u(t)\diamond \mathrm{d}W^H(t)$ will be utilized in cases when u is a stochastic process that falls within the domain of the divergence operator.

2.5.4 Fractional Wick Itô Skorohod Integrals

Now we aim to define the stochastic integral $\delta(F) = \int_0^T F(t) \diamond \mathrm{d}W^H(t)$ as an element within $L^2(\mathcal{H})$, provided that $F(t)$ is a suitable integrand. It is important to note that we are assuming $H > \frac{1}{2}$. For simplicity in notation, we denote the fractional Wick Itô Skorohod (fWIS) integral by

$$\int_0^T F(t) \diamond \mathrm{d}W^H(t) := \int_0^T F(t) \delta W^H(t).$$

To define this integral, we consider an arbitrary partition of the interval $[0, T]$, denoted by $\{t_i^n\}_{i=0}^{k_n}$, where $\lim_{n\to\infty} \max_{0\leq i\leq k_n-1}\{t_{i+1}^n - t_i^n\} = 0$ and introduce the Riemann sum:

$$S(F, \pi) = \sum_{i=0}^{k_n-1} F(t_i) \diamond \left(W^H(t_{i+1}^n) - W^H(t_i^n)\right).$$

Definition 2.18 Let $L_H(0, T)$ be a family consisting of stochastic processes F on $[0, T]$ possessing the following properties:

- $F \in L_H(0, T)$ if and only if

$$\mathbb{E}[\|F\|_{\mathcal{H}}^2] = \alpha_H \mathbb{E}\left[\int_0^T \int_0^T |r-u|^{2H-2} F(r) F(u) \mathrm{d}u\mathrm{d}r\right] < \infty$$

- F is Malliavin differentiable,
- For $0 \leq t \leq T$, there exists $D_t^H F(s)$ such that:

$$\mathbb{E}\left[\int_{[0,T]^4} D_{t_1}^H F(s_1) D_{t_2}^H F(s_2) \phi(t_1, t_2) \phi(s_1, s_2) \mathrm{d}t_1 \mathrm{d}s_1 \mathrm{d}t_2 \mathrm{d}s_2\right] < \infty.$$

The following theorem, which is the most significant in this section, introduces the fWIS integral based on Theorem 3.6.1 in [18].

Theorem 2.19 Let $(F(t), t \in [0, T])$ be a stochastic process such that $F \in L_H(0, T)$. Then, the following limit exists in $L^2(\Omega, \mathcal{F}, \mathbb{P})$:

$$\lim_{n\to 0} \sum_{i=0}^{k_n-1} F(t_i) \diamond \left(W^H(t_{i+1}) - W^H(t_i)\right),$$

and this limit is denoted by $\int_0^T F(s) \delta W^H(s)$, known as the fWIS integral. Furthermore, this integral satisfies the following relations:

$$\mathbb{E}\left[\int_0^T F(s) \delta W^H(s)\right] = 0,$$

and

$$\left\|\int_0^T F(s)\delta W^H(s)\right\|_{L_H(0,T)} := \mathbb{E}\left[|\int_0^T F(s)\delta W^H(s)|^2\right]$$
$$= \mathbb{E}\left[\int_{[0,T]^4} D^H_{t_1}F(s_1)D^H_{t_2}F(s_2)\phi(t_1,t_2)\phi(s_1,s_2)\mathrm{d}t_1\mathrm{d}s_1\mathrm{d}t_2\mathrm{d}s_2\right]$$
$$+ \alpha_H\mathbb{E}\left[\int_{[0,T]^2}|r-u|^{2H-2}F(r)\,F(u)\mathrm{d}u\mathrm{d}r\right].$$

The following properties are directly deduced from the above theorem:

- If $F, G \in L_H(0,T)$, then for any arbitrary constants a and b, the following relation holds:
$$\int_0^t (aF(s)+bG(s))\delta W^H(s) = a\int_0^t F(s)\delta W^H(s) + b\int_0^t G(s)\delta W^H(s).$$
- If $F \in L_H(0,T)$, $\mathbb{E}\left[\sup_{0\le s\le t}F(s)\right]^2 < \infty$ and also $\sup_{0\le s\le t}\mathbb{E}|D^H_t F(s)|^2 < \infty$, then $\int_0^t F(s)\delta W^H(s)$ has a continuous version.

Remark 2.20 [30] Suppose $\{t_i^n\}_{i=0}^{k_n}$ is a sequence of partitions of the interval $[0,T]$, such that $\lim_{n\to\infty}\max_{0\le i\le k_n-1}\{t^n_{i+1}-t^n_i\} = 0$. If $\sum_{i=0}^{k_n-1}F(t_i)\left(W^H(t^n_{i+1})-W^H(t^n_i)\right)$ converges in the space of $L^2(\Omega,\mathcal{F},\mathbb{P})$, then this limit corresponds to the Young integral (see Subsections 2.5.2.1).

The following theorem expresses the relationship between the fWIS integral and the Young integral.

Theorem 2.21 [18] If $F \in L_H(0,T)$, then the following relation holds:

$$\int_0^t F(s)\circ\mathrm{d}W^H(s) = \int_0^t F(s)\delta W^H(s) + \int_0^t\int_0^s D^H_v F\phi(s,v)\mathrm{d}v\mathrm{d}s.$$

In a special case when F is a function of $W^H(t)$, we can write

$$\int_0^t F(W^H(s))\circ\mathrm{d}W^H(s) = \int_0^t F(W^H(s))\delta W^H(s)$$
$$+ H(2H-1)\int_0^t\int_0^s F'(W^H(s))(s-v)^{2H-2}\mathrm{d}v\mathrm{d}s$$
$$= \int_0^t F(W^H(s))\delta W^H(s) + H\int_0^t F'(W^H(s))s^{2H-1}\mathrm{d}s. \tag{2.21}$$

These two stochastic integrals have the following important properties:

- $\mathbb{E}\left[\int_0^t F(s)\delta W^H(s)\right] = 0$, while, in general, $\mathbb{E}\left[\int_0^t F(s)\circ\mathrm{d}W^H(s)\right] \neq 0$.

- The chain rule for the fWIS integral is more complex compared to the Young integral.

Since the rules for the Young integral are similar to the Riemann-Stieltjes integral, in this thesis we consider the fractional stochastic integral in the Young integral form.

Example 2.22 According to the definition of Young integration, we have:

$$\int_0^t W^H(s) \circ \mathrm{d}W^H(s) = \frac{1}{2}(W^H(t))^2,$$

hence (2.21) implies that:

$$\int_0^t W^H(s)\delta W^H(s) = \frac{1}{2}(W^H(t))^2 - H\int_0^t s^{2H-1}\mathrm{d}s = \frac{1}{2}\left[(W^H(t))^2 - t^{2H}\right].$$

The following example shows that $\mathbb{E}\left[\int_0^t F(s) \circ \mathrm{d}W^H(s)\right]$ is not necessarily 0.

Example 2.23 [30] It is well-known that if x is a standard normal random variable, $x \sim N(0,1)$, then

$$\mathbb{E}[x^n] = \begin{cases} \frac{n!}{(\sqrt{2})^n\left(\frac{n}{2}\right)!} & \text{if } n \text{ is even,} \\ 0 & \text{if } n \text{ is odd.} \end{cases}$$

Consider $f(x) = x^n$ and referring to equation (2.21), we obtain

$$\begin{aligned}
\mathbb{E}\left[\int_0^t f\left(W^H(s)\right) \circ \mathrm{d}W^H(s)\right] &= \mathbb{E}\left[H\int_0^t s^{2H-1} f'\left(W^H(s)\right)\mathrm{d}s\right] \\
&= nH\int_0^t s^{2H-1}\mathbb{E}\left[f'\left(W^H(s)\right)\right]\mathrm{d}s \\
&= nH\int_0^t s^{2H-1}\mathbb{E}\left[\left(W^H(s)\right)^{n-1}\right]\mathrm{d}s \\
&= nH\int_0^t s^{(n+1)H-1}\mathbb{E}\left[\left(\frac{W^H(s)}{s^H}\right)^{n-1}\right]\mathrm{d}s \\
&= \begin{cases} \frac{n!t^{(n+1)H}}{(\sqrt{2})^{n-1}(n+1)\left(\frac{n-1}{2}\right)!} & \text{if } n \text{ is odd,} \\ 0 & \text{if } n \text{ is even.} \end{cases}
\end{aligned}$$

Remark 2.24 [30] Consider an arbitrary partition $\{t_i^n\}_{i=0}^{k_n}$ of the interval $[0,T]$, such that $\lim_{n\to\infty}\max_{0\le i\le k_n-1}\{t_{i+1}^n - t_i^n\} = 0$. Also, suppose that $F \in L_H(0,T)$ is a stochastic process. As mentioned above, for a standard Brownian motion process $(W(t), t \ge 0)$, the Itô integral is defined as the limit of Riemann sums, that is, $\sum_{i=0}^{k_n-1} F(t_i)\left(W(t_{i+1}) - W(t_i)\right)$, as $n \to \infty$. Similarly, the Stratonovich integral is defined as $\sum_{i=0}^{k_n-1} F(\frac{t_i+t_{i+1}}{2})\left(W(t_{i+1}) - W(t_i)\right)$. In the case of fBm, it has been shown

that both limits,

$$\lim_{n\to\infty}\sum_{i=0}^{k_n-1} F(t_i)\left(W^H(t_{i+1})-W^H(t_i)\right),$$

and

$$\lim_{n\to\infty}\sum_{i=0}^{k_n-1} F(\frac{t_i+t_{i+1}}{2})\left(W^H(t_{i+1})-W^H(t_i)\right),$$

converge to the pathwise Riemann Stieltjes integral (Young integral).

2.6 Fractional Itô Formula

Now, we proceed to express the fractional Itô formula. To illustrate this formula, we first present the following theorem, which characterizes the Malliavin ϕ-derivative of fWIS integral. If $F:\Omega\to\mathbb{R}$ is a given function, we define the Malliavin ϕ-derivative of F at s as follows:

$$D_s^{\phi}F=\int_{\mathbb{R}} D_v^H F\,\phi(s,v)\mathrm{d}v,$$

where $D_v^H F$ denotes the Malliavin derivative.

Theorem 2.25 [18] Consider the stochastic process $(F(t),t\in[0,T])$ in the space $L_H(0,T)$ such that $\sup_{0\le s\le T}\mathbb{E}[|D_s^{\phi}F(s)|^2]<\infty$. Define $\eta_t=\int_0^t F(u)\,\delta W^H(u)$, for $t\in[0,T]$. Then, for $s,t\in[0,T]$, the following equation holds almost surely:

$$D_s^{\phi}\eta_t=\int_0^t D_s^{\phi}F(u)\,\delta W_u^H+\int_0^t F(u)\phi(s,u)\mathrm{d}u.$$

Now, let us express the fractional Itô formula in the general case.

Theorem 2.26 [18] Suppose that $\eta_t=\xi+\int_0^t G(u)\mathrm{d}u+\int_0^t F(u)\,\delta W^H(u)$, where $(F(u),0\le u\le T)$ is a stochastic process in $L_H(0,T)$. Assume that there is an $\alpha>1-H$ such that

$$\mathbb{E}\left[|F(u)-F(v)|^2\right]\le C|u-v|^{2\alpha},$$

where $|u-v|\le\delta$ for some $\delta>0$ and

$$\lim_{0\le u,v\le t,|u-v|\to 0}\mathbb{E}\left[\left|D_u^{\phi}\left(F(u)-F(v)\right)\right|^2\right]=0.$$

Also, let $f\in\mathcal{C}_b^{1,2}(\mathbb{R}_+\times\mathbb{R};\mathbb{R})$. Furthermore, assume that

$$\mathbb{E}\left[\int_0^t |F(s)D_s^{\phi}\eta_s|\mathrm{d}s\right]<\infty,\quad \mathbb{E}\left[\sup_{0\le s\le T}|G(s)|\right]<\infty,$$

and $(\partial f(s,\eta_s)/\partial x)\,F(s)$ for $s\in[0,T]$ belong to the space $L_H(0,T)$. Then, for each

$0 \leq t \leq T$, the following equation holds almost surely:

$$f(t,\eta_t) = f(0,\xi) + \int_0^t \frac{\partial f}{\partial s}(s,\eta_s)\mathrm{d}s + \int_0^t \frac{\partial f}{\partial x}(s,\eta_s)G(s)\mathrm{d}s + \int_0^t \frac{\partial f}{\partial x}(s,\eta_s)F(s)\,\delta W^H(s) + \int_0^t \frac{\partial^2 f}{\partial x^2}(s,\eta_s)F(s)D_s^{\phi}\eta_s\mathrm{d}s.$$

We end this section by presenting an Itô formula for fBm in the specific case where $\eta_t = W^H(t)$, which holds for all values of H in the interval $(0,1)$.

Theorem 2.27 [18] Let $H \in (0,1)$ and the function $f(t,x) : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$ is of class $\mathcal{C}^{1,2}(\mathbb{R}\times\mathbb{R};\mathbb{R})$ and also the following random variables are all elements of $L^2(\Omega,\mathcal{F},\mathbb{P})$:

$$f\left(t, W^{(H)}(t)\right), \quad \int_0^t \frac{\partial f}{\partial s}\left(s, W^{(H)}(s)\right)\mathrm{d}s, \quad \text{and} \quad \int_0^t \frac{\partial^2 f}{\partial x^2}\left(s, W^{(H)}(s)\right)s^{2H-1}\mathrm{d}s.$$

Then,

$$f(t,W^H(t)) = f(0,0) + \int_0^t \frac{\partial f}{\partial s}(s,W^H(s))\mathrm{d}s + \int_0^t \frac{\partial f}{\partial x}(s,W^H(s))\,\delta W^H(s) + H\int_0^t \frac{\partial^2 f}{\partial x^2}(s,W^H(s))s^{2H-1}\mathrm{d}s.$$

Example 2.28 Let $f(t,W^H(t)) = (W^H(t))^2$. Consequently, we have:

$$d(W^H(t))^2 = 2W^H(t)\,\delta W^H(t) + 2Ht^{2H-1}\mathrm{d}t.$$

2.7 Existence and Uniqueness of Solution of SDE with fBm

In this section, we aim to establish the theorem for the global existence and uniqueness of solutions for multidimensional, time-dependent SDEs driven by fBm with a Hurst parameter $H > \frac{1}{2}$. We consider the following multidimensional SDE:

$$\begin{aligned} x_i(t) &= x_{0,i} + \int_0^t a_i(s,x(s))\mathrm{d}s + \sum_{j=1}^{q}\int_0^t \sigma_{ij}(s,x(s)) \circ \mathrm{d}W_j^H(s) \quad \text{for} \quad i = 1,\ldots,d \\ &= x_0 + \int_0^t a(s,x(s))\mathrm{d}s + \int_0^t \sigma(s,x(s)) \circ \mathrm{d}W^H(s), \quad t \in [0,T]. \end{aligned} \tag{2.22}$$

In this scenario, we consider $W^H = (W_1^H,\ldots,W_q^H)$ as a q-dimensional fBm with the same Hurst parameter $H \in (\frac{1}{2},1)$ defined on complete filtered probability space

$\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P}\right)$. x_0 is a $d-$dimensional random variable and the coefficients $a_i, \sigma_{ij} : \Omega\times[0,T]\times\mathbb{R}^d \to \mathbb{R}$ are measurable functions and for a matrix $M = (m_{i,j})_{d\times q}$ denotes $\|M\|_2^2 = \sum_{i,j} |m_{i,j}|^2$. Equation (2.22) is considered as the Young integral equation (Section 2.5.2.1).

Let H be such that $\frac{1}{2} < H < 1$ and α satisfy $1 - H < \alpha < \frac{1}{2}$. Define the space $\mathcal{V}_0^{\alpha,\infty}\left([0,T];\mathbb{R}^d\right)$ to consist of all measurable functions $f : [0,T] \to \mathbb{R}^d$ which satisfy the following condition:

$$\|f\|_{\alpha,\infty} := \sup_{t\in[0,T]} \left(|f(t)| + \int_0^t \frac{|f(t) - f(s)|}{(t-s)^{\alpha+1}} \mathrm{d}s \right) < \infty,$$

and we have, for all $0 < \varepsilon < \alpha$

$$\mathcal{C}^{\alpha+\varepsilon}\left([0,T];\mathbb{R}^d\right) \subset \mathcal{V}_0^{\alpha,\infty}\left([0,T];\mathbb{R}^d\right) \subset \mathcal{C}^{\alpha-\varepsilon}\left([0,T];\mathbb{R}^d\right)$$

where $\mathcal{C}^\alpha$ denotes the class of Hölder continuous functions.

Let us consider the following assumptions on the coefficients, which are supposed to hold for $\mathbb{P}$-almost all $\omega \in \Omega$. The constants M_N, L_N and the function a_0 may depend on ω.

(H_σ) : $\sigma(t,x)$ is continuously differentiable with respect to x and there exist constants $0 < \beta, \delta \leq 1$. Additionally, for every $N \geq 0$, there is a constant $M_N > 0$ such that the following conditions are satisfied:

(i) Lipschitz continuity

$$\|\sigma(t,x) - \sigma(t,y)\|_2 \leq M_0\|x - y\|_2, \quad \forall x \in \mathbb{R}^d, \forall t \in [0,T]$$

(ii) Hölder continuity in time:

$$\|\sigma(t,x) - \sigma(s,x)\|_2 + \|\partial_{x_i}\sigma(t,x) - \partial_{x_i}\sigma(s,x)\|_2 \leq M_0|t-s|^\beta$$

for all $x \in \mathbb{R}^d$ and $t, s \in [0,T]$.

(iii) Local Hölder continuity of $\partial_{x_i}\sigma$

$$\|\partial_{x_i}\sigma(t,x) - \partial_{y_i}\sigma(t,y)\|_2 \leq M_N\|x-y\|_2^\delta \quad \forall \|x\|_2, \|y\|_2 \leq R, \quad \forall t \in [0,T]$$

for $1 \leq i \leq d$,

(iv) There exist $\gamma \in [0,1]$ and $K_0 > 0$ such that

$$\|\sigma(t,x)\|_2 \leq K_0\left(1 + \|x\|_2^\gamma\right), \quad \forall x \in \mathbb{R}^d, \quad \forall t \in [0,T].$$

(H_a) : There exists a function $a_0 \in L^p\left([0,T];\mathbb{R}^d\right)$, where $p \geq 2$ and for each $N \geq 0$, there exists a constant $L_N > 0$ such that the following conditions hold:

(i) Local Lipschitz continuity:

$$\|a(t,x)-a(t,y)\|_2 \leq L_N\|x-y\|_2, \forall \|x\|_2, \|y\|_2 \leq N, \forall t \in [0,T]$$

(ii) Boundedness:

$$\|a(t,x)\|_2 \leq L_0\|x\|_2 + a_0(t), \quad \forall x \in \mathbb{R}^d, \forall t \in [0,T]$$

Let

$$\alpha_0 = \min\left\{\frac{1}{2}, \beta, \frac{\delta}{1+\delta}\right\}.$$

The existence and uniqueness theorem of the solution to equation (2.22) is established in Theorem 2.1 in [95].

Theorem 2.29 [95] Let x_0 be an $\mathbb{R}^d-$valued random variable, the coefficients $\sigma(t,x)$ and $a(t,x)$ satisfy assumptions H_σ and H_a with $\beta > 1-H$ and $\delta > \frac{1}{H}-1$. If $\alpha \in (1-H, \alpha_0)$ and $p \geq 1/\alpha$, then there exists a unique stochastic process $x \in L^0\left((\Omega,\mathcal{F},\mathbb{P}); \mathcal{V}_0^{\alpha,\infty}\left([0,T];\mathbb{R}^d\right)\right)$ that solves the stochastic equation (2.22) and, moreover, for $\mathbb{P}$-almost all $\omega \in \Omega$

$$x(\omega,\cdot) = (x_i(\omega,\cdot))_{d\times 1} \in C^{1-\alpha}\left([0,T];\mathbb{R}^d\right).$$

2.8 Numerical Methods for SDEs with fBm

The pursuit of explicit solutions for stochastic differential equations driven by fBm is sometimes unachievable for a significant number of equations. Consequently, the development of numerical methods becomes crucial in addressing this class of equations. The stochastic differential equations (2.22) can be numerically solved using various general-purpose stochastic numerical techniques (refer to, for example, [60, 77] and related literature). [28] delve into the 2-step Euler scheme simplified through Wong-Zakai approximations, achieving a pathwise convergence rate of close to $\left(H-\frac{1}{3}\right)$ in the Hölder norm for H within the interval $\left(\frac{1}{3},1\right)$. The authors in [78, 80] suggested that the optimal convergence rate for the supremum norm could be $2H-\frac{1}{2}$. Furthermore, [45, 46] establish the Crank-Nicolson scheme and the modified Euler scheme's optimal strong convergence rates for $H \in \left(\frac{1}{2},1\right)$, employing a combination of Malliavin calculus and fractional calculus techniques. In this section, we discuss two numerical methods for SDEs driven by fBm, which will be utilized in Chapter 6.

2.8.1 Euler Method

The results presented in this subsection are based on the findings of [79], where the authors established the Euler method in the case of fBm. We consider the following

SDE:

$$x(t) = x_0 + \int_0^t a(x(s))\mathrm{d}s + \int_0^t \sigma(x(s)) \circ \mathrm{d}W^H(s), \quad t \in [0,T]. \tag{2.23}$$

The interval $[0,T]$ is partitioned by the set of $\{0 = t_0 < t_1 < \ldots < t_n = T\}$, where $t_k = \frac{kT}{n}$ for $0 \leq k \leq n$ and $n \in \mathbb{N}$. The rates of convergence will be discussed relative to this partition scheme. For simplicity, we denote ΔW_k^H instead of $W^H(t_{k+1}) - W^H(t_k)$ and $\Delta_k = t_{k+1} - t_k$. Assuming $H > \frac{1}{2}$, we consider the integral with respect to W^H as defined by the Young integral (see Definition 2.13). The Euler scheme y^n for equation (2.23) is given by:

$$\begin{cases} y_0^n = x_0, \\ y_{k+1}^n = y_k^n + a(y_k^n)\Delta_k + \sigma(y_k^n)\Delta W_k^H, \quad k \in \{0, \ldots, n-1\}, \end{cases} \tag{2.24}$$

where standard assumptions are made on a and σ and $y_t^n = y_{\frac{[nt]}{n}}^n$ for $t \in [0,T]$.

Theorem 2.30 [79] Let $a \in \mathcal{C}_b^2$ and $\sigma \in \mathcal{C}_b^3$, indicating that the functions a and σ are continuously differentiable twice and three times with bounded derivatives, respectively. Furthermore, let us assume that they also fulfill the conditions outlined in Theorem 2.29. Then, the Euler scheme y^n for equation (2.23), as defined by (2.24) and as $n \to \infty$, satisfies:

$$n^{2H-1}[y_T^n - x(T)] \xrightarrow{\text{a.s.}} -\frac{1}{2}\int_0^T \sigma'(x(s)) D_s^H x(T)\mathrm{d}s,$$

where $D_s^H x(t)$, for $s,t \in [0,T]$, denotes the Malliavin derivative at time s of $x(t)$ with respect to fBm W^H defined in Section 2.5.3. This result is notable because it lacks an analog in the classical theory of stochastic processes driven by sBm.

Remark 2.31 In [79], the author introduced a different form of the pathwise integral with respect to fBm, specifically the forward integral $\int_0^t Y(s)dX^-(s)$ (refer to [108] for further information on this integral). As established in [109], when X and Y are two real-valued processes with paths that are almost surely Hölder continuous of orders $\alpha > 0$ and $\beta > 0$ respectively, such that $\alpha + \beta > 1$, the forward integral $\int_0^t Y(s)d^-X(s)$ is equivalent to the Young integral $\int_0^t Y(s)dX(s)$.

2.8.2 Implicit Midpoint Method

In practical applications, stiff differential equations often pose significant challenges for numerical simulations in both deterministic and stochastic systems. Implicit methods are generally more effective than explicit methods when dealing with stiff problems. This section aims to explore an implicit numerical method suitable for addressing stiff SDEs. The focus of the numerical section for fBm will be on the stochastic implicit midpoint method. The methodology is based on the framework presented in [43] and therefore the following discussion and results conform to the

findings of this reference. We consider the SDE as follows:

$$dy(t) = V\left(y(t)\right) \mathrm{d}x(t) = \sum_{l=1}^{d} V_l\left(y(t)\right) \mathrm{d}x_l(t), \quad y(0) = y_0 \in \mathbb{R}^m, t \in [0,T], \tag{2.25}$$

where $x(t) = (x_1(t), \ldots, x_d(t))^\top \in \mathbb{R}^d$ with $x_1(t) = t$ and $x_2(t), \ldots, x_d(t)$ being independent fBms with Hurst parameter $H \in \left(\frac{1}{2}, 1\right)$. The well-posedness of (2.25) is interpreted pathwise through the Young integral (see Section 2.5.2.1), which is valid for fractional Brownian motion with Hurst index $H > 1/2$. This follows the analytic framework established in [43], where the existence and uniqueness of solutions are obtained via fractional calculus estimates for Hölder continuous driving signals.

In the context of a numerical algorithm, we utilize a uniform division of the interval $[0,T]$ with a step size defined as $\Delta t = \frac{T}{n}$, where $n \in \mathbb{N}$. We represent t_k as $t_k = k\Delta t$, for $k = 0, \ldots, n$. Consider an s-stage Runge-Kutta method applied to (2.25):

$$y^n_{t_{k+1},i} = y^n_{t_k} + \sum_{j=1}^{\mathrm{s}} a_{ij} V\left(y^n_{t_{k+1},j}\right) \Delta x_k, \tag{2.26}$$

$$y^n_{t_{k+1}} = y^n_{t_k} + \sum_{i=1}^{\mathrm{s}} b_i V\left(y^n_{t_{k+1},i}\right) \Delta x_k, \tag{2.27}$$

with $i, j = 1, \ldots, s, k = 0, \ldots, n-1, \Delta x_k = x(t_{k+1}) - x(t_k) \in \mathbb{R}^d$ and $y^n_{t_0} = y_0 \in \mathbb{R}^m$. Here, $y^n_{t_{k+1},i}$, $i = 1, \ldots, \mathbf{s}$, are called stage values.

Following [43], we construct continuous extensions of the discrete scheme. This allows one to perform a precise pathwise error analysis by comparing the exact solution $y(t)$ and the numerical interpolate y^n_t through an error decomposition into a local truncation term and a quadrature-type remainder. The continuous form is defined as follows:

To derive order conditions on coefficients of Runge-Kutta methods with the strong convergence rate $2H - \frac{1}{2}$, the authors in [43] first construct the continuous versions (2.28) and (2.29) for the Runge-Kutta methods (2.26) and (2.27), taking advantages of the stage values $y^n_{t_k,i}$. Denote $\lceil t \rceil^n := t_{k+1}$ for $t \in (t_k, t_{k+1}]$. In particular, $t = t_k$ if and only if $t = \lceil t \rceil^n$ for some $k = 0, \cdots, n$. The continuous version reads

$$y^n_{t,i} := y^n_{(t-h)\vee 0} + \sum_{j=1}^{\mathrm{s}} \int_{(t-h)\vee 0}^{t} a_{ij} V\left(y^n_{[s]^n,j}\right) \mathrm{d}X_s, \quad i = 1, \cdots, \mathrm{s}, \tag{2.28}$$

$$y^n_t := y_0 + \sum_{i=1}^{\mathrm{s}} \int_0^t b_i V\left(y^n_{[s]^n,i}\right) \mathrm{d}X_s, \tag{2.29}$$

where $s \vee t = \max\{s, t\}$.

Theorem 2.32 [43] Let us assume that $V \in \mathcal{C}^3_b(\mathbb{R}^m; \mathbb{R}^{m\times d})$, which means that V is three times continuously differentiable with bounded derivatives, with $H > \frac{1}{2}$.

Define c_i as $c_i = \sum_{j=1}^{\mathbf{s}} a_{ij}$. Given the conditions

$$\sum_{i=1}^{\mathbf{s}} b_i = 1 \quad \text{and} \quad \sum_{i=1}^{\mathbf{s}} b_i c_i = \frac{1}{2}, \tag{2.30}$$

the Runge-Kutta method applied to equation (2.25) exhibits a strong convergence rate of $2H-\frac{1}{2}$. Specifically, for a constant C that does not depend on n, the following inequality holds:

$$\left\| \sup_{t\in[0,T]} |y(t) - y_t^n| \right\|_{L^p(\Omega)} \leq C(\Delta t)^{2H-\frac{1}{2}}, \quad p \geq 1,$$

where $\Delta t = \frac{T}{n}$ and y_t^n are defined in equation (2.29).

For solving SDE (2.22), we consider the 1-stage Runge-Kutta method with coefficients defined by the following Butcher tableau:

$$\begin{array}{c|c} \frac{1}{2} & \frac{1}{2} \\ \hline & 1 \end{array}.$$

This Runge-Kutta method is the implicit midpoint scheme. Hence, it takes the following form when applied to (2.22):

$$x_n(t_{k+1}) = x_n(t_k) + \Big[a\Big(\frac{x_n(t_k) + x_n(t_{k+1})}{2}\Big)\Big]\Delta t + \sum_{i=1}^{q} \sigma^i\Big(\frac{x_n(t_k) + x_n(t_{k+1})}{2}\Big)\Delta W_{i,k}^H, \tag{2.31}$$

Here, we define $\Delta W_{i,k}^H = W_i^H(t_{k+1}) - W_i^H(t_k)$. This method satisfies condition (2.30) and Theorem 2.32, indicate that its mean-square convergence rate (see Section 1.6) is $2H - 1/2$,i.e.,

$$\left\| \max_{1\leq k\leq n} |x(t) - x_n(t_k)| \right\|_{L^2(\Omega)} \leq C(\Delta t)^{2H-\frac{1}{2}}.$$

3 Model Order Reduction

Model reduction techniques have been developed to reduce the computational cost associated with high-dimensional dynamical systems. These methods aim to create simplified models with lower dimensions that require less computational resources to simulate while preserving the essential dynamics of the full system (refer to [6, 12, 13, 94] and related references). This chapter focuses on model order reduction for deterministic and stochastic linear dynamical systems. It begins with an exploration of Lyapunov operators and stability in deterministic systems, extending these concepts to stochastic systems. It introduces Gramian matrices as tools for characterizing dominant subspaces, forming the basis for MOR techniques. Special attention is given to the Balanced Truncation method and the POD method, which is applied to reduce the computational complexity of high-dimensional systems while retaining their essential dynamics.

3.1 Linear Deterministic Systems

In the context of linear state-space systems, we analyze a system G characterized by differential equations governing states $x(t) \in \mathbb{R}^n$, input $u(t) \in \mathbb{R}^m$ and output $y(t) \in \mathbb{R}^p$, described as:

$$\begin{cases} \dot{x}(t) = Ax(t) + Bu(t), \quad x(0) = x_0 \\ y(t) = Cx(t) \end{cases} \tag{3.1}$$

where $A \in \mathbb{R}^{n\times n}$ with $\sigma(A) \subset \mathbb{C}_-$, defined as the spectrum of A within the open left half complex plane $\mathbb{C}_-$, along with $B \in \mathbb{R}^{n\times m}$ and $C \in \mathbb{R}^{p\times n}$, represent components of the system.

The study of system (3.1), which can be alternatively identified as G, (A, B, C) and (A, B), requires a revisit of fundamental concepts such as stability, controllability, reachability and observability. These concepts are explained in the following.

Definition 3.1 The system (A, B) is considered controllable if there exists a control $u(t)$ capable of steering the system from an initial state $x(0) = x_0$ to the state $x(t) = 0$ within a time t. Alternatively, it should enable steering from the initial state x_0 to any desired state x_1 within a time t_1.

Conversely, the concept of reachability is defined when the zero state is steered toward a desired state, presenting a converse statement as given in Definition 3.1.

Definition 3.2 The system G is reachable if a control $u(t)$ exists that can drive the

system either from the initial state $x(0) = 0$ to the final state $x(t_1) = x_1$ or from the initial state $x(0) = x_0 \neq 0$ to a final state $x(t_1) = x_1$.

The controllability (reachability) matrix, termed the Kalman matrix of the system G, is denoted as

$$\mathcal{R} = [B \; AB \; A^2B \; \dots \; A^{n-1}B].$$

The system is identified as controllable (reachable) if this matrix is obtained in full rank. Additionally, the notation $\mathcal{R}(A, B)$ denotes the range of $\mathcal{R}$, corresponding to the reachability space of the deterministic system.

Definition 3.3 The system G is observable if the initial state $x(0)$ can be determined from the outputs of the system over time and the applied inputs. The observability matrix, which is key in determining this property, is presented as follows:

$$\mathcal{O} = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{n-1} \end{bmatrix}.$$

In a similar manner to the controllability matrix $\mathcal{R}$, the system is observable if the observability matrix $\mathcal{O}$ has full rank. In the study of the system G, controllability and observability serve as mathematical complements, unaffected by changes in basis representation. As a result, the reachability of G corresponds directly to the observability of the dual system G^*. For additional techniques to determine the controllability and observability of the system, see [5].

Here, the function $x(t; x_0, u)$ denotes the solution to the state equation (3.1). Specifically, the solution takes the form:

$$x(t; x_0, u) = e^{At}x_0 + \int_0^t e^{A(t-\tau)}Bu(\tau)\, d\tau, \quad \text{for } t \geq 0,$$

so, we have

$$y(t) = Cx(t; x_0, u) \quad \text{for } t \geq 0.$$

In the next section, we derive the link between controllability and observability matrices and the algebraic Lyapunov equations.

3.1.1 Lyapunov Operators

Associated with (3.1), the Lyapunov equations are given by

$$AP + PA^\top = -BB^\top, \tag{3.2a}$$
$$A^\top Q + QA = -C^\top C, \tag{3.2b}$$

where P and Q are symmetric matrices in $\mathbb{R}^{n\times n}$ representing the controllability and observability Gramians of the system, respectively. Considering the stability assumptions of the system, P and Q can be expressed as integrals:

$$P = \int_0^\infty e^{At} BB^\top e^{A^\top t} \mathrm{d}t, \tag{3.3a}$$
$$Q = \int_0^\infty e^{A^\top t} C^\top C e^{At} \mathrm{d}t. \tag{3.3b}$$

Utilizing Theorem 3.4 to verify the controllability and observability of the system, we find that the Gramians P and Q must be positive definite. Furthermore, for the controllability and observability matrices $\mathcal{R}$ and $\mathcal{O}$, it is required that they possess full rank.

Theorem 3.4 [5][Reachability and Observability conditions] The following statements are equivalent:

(i) The pair (A, B), $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, is reachable.

(ii) The rank of the reachability matrix is full: $\operatorname{rank} \mathcal{R}(A, B) = n$.

(iii) The reachability Gramian is positive definite $P > 0$.

and also the following statements are equivalent:

(i) The pair (C, A), $C \in \mathbb{R}^{p\times n}$, $A \in \mathbb{R}^{n\times n}$, is observable.

(ii) The rank of the observability matrix is full: $\operatorname{rank} \mathcal{O}(C, A) = n$.

(iii) The observability Gramian is positive definite: $Q > 0$.

3.1.1.1 Vectorization and Kronecker Product

The concept of vectorization and the Kronecker product are introduced as computational tools to facilitate the calculation of the controllability and observability of Gramians, P and Q, derived from (3.2). The results are as follows:

$$(A \otimes I + I \otimes A)\operatorname{vec}(P) = -\operatorname{vec}(BB^\top), \tag{3.4a}$$
$$\left(A^\top \otimes I + I \otimes A^\top\right)\operatorname{vec}(Q) = -\operatorname{vec}(C^\top C). \tag{3.4b}$$

where $\cdot \otimes \cdot$ is the Kronecker product between two matrices and vec$(\cdot)$ be the vectorization of a matrix (see Appendix B for more details).

We then introduce the operator $\mathcal{L}_A$ from equation (3.2) as $\mathcal{L}_A : \mathcal{S}_n \to \mathcal{S}_n$ with $\mathcal{S}_n := \{X \mid X \in \mathbb{R}^{n\times n}$ symmetric matrix $\}$, defined by:

$$\mathcal{L}_A : X \longrightarrow AX + XA^\top. \tag{3.5}$$

The adjoint operator $\mathcal{L}_A^*$, with respect to the Frobenius inner product (which means that $\langle X, Y\rangle = trace(X^\top Y)$ where X and Y are two matrices), is given as:

$$\mathcal{L}_A^* : X \longrightarrow A^\top X + XA. \tag{3.6}$$

With the introduction of $\mathcal{L}_A$, it becomes compatible to work with the Kronecker product notation as well, defining:

$$\mathcal{L}_A = A \otimes I + I \otimes A, \tag{3.7}$$
$$\mathcal{L}_A^* = A^\top \otimes I + I \otimes A^\top. \tag{3.8}$$

Expanding on equation (3.2), we can establish that:

$$\mathcal{L}_A(P) \leq -BB^\top, \quad \mathcal{L}_A^*(Q) \leq -C^\top C. \tag{3.9}$$

These characterizations are vital in the analysis of stability, controllability and observability of the system (3.1).

3.1.2 Stability

To ascertain the stability of a deterministic system, we utilize the following theorem as a set of equivalent conditions:

Theorem 3.5 [Stability of Linear Systems] For a linear system with deterministic, time-invariant and homogeneous characteristics, described by (3.1), the following statements are equivalent:

(i) The system is asymptotically stable,

(ii) $\sigma(\mathcal{L}_A) \subset \mathbb{C}_-$, i.e. $\sigma(A) \subset \mathbb{C}_-$,

(iii) $\exists X > 0 : \mathcal{L}_A(X) < 0$,

(iv) $\forall Y > 0, \exists X > 0 : \mathcal{L}_A(X) = -Y$,

where $\sigma(A)$ is the spectrum of the matrix A and $\mathbb{C}_-$ denotes the set of complex numbers with negative real parts, ensuring stability. The stability criteria presented here can be extended to linear stochastic systems. Subsequently, we introduce model order reduction strategies for deterministic systems. The extension of these strategies to stochastic systems is explored in Section 3.2.2.

3.1.3 Model Order Reduction for Deterministic Systems

The primary concept of model reduction involves approximating a high-dimensional dynamical system with a lower-dimensional model that preserves its essential dynamic features. In this section, we present several model order reduction (MOR) techniques. We begin by defining the reduced-order system.

3.1.3.1 Reduced order system

The system G_r, representing a lower-order model, approximates the full state-space system G of order n. The equations governing the dynamics of G_r are given by:

$$\begin{cases} \dot{x}_r(t) = A_r x_r(t) + B_r u(t), \quad x_r(0) = x_0, \\ y_r(t) = C_r x_r(t), \end{cases} \tag{3.10}$$

where $x_r(t) \in \mathbb{R}^r$ and $r \ll n$. Our goal is to minimize the norm difference between y and y_r to evaluate the maximum deviation between the output of the original system y and its approximation y_r. To generate G_r, we utilize orthogonal truncation, involving a coordinate transformation $x_S(t) = Sx(t)$ via a transformation matrix $S \in \mathbb{R}^{n\times n}$ that is invertible. Putting this coordinate transformation into (3.1) yields a modified state-space representation:

$$\begin{aligned} \dot{x}_S(t) &= A_S x_S(t) + B_S u(t), \quad x_S(0) = Sx_0, \\ y(t) &= C_S x_S(t), \quad t \geq 0, \end{aligned} \tag{3.11}$$

where

$$\begin{aligned} A_S &= SAS^{-1} = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}, && A_{11} \in \mathbb{R}^{r\times r}, \\ B_S &= SB = \begin{pmatrix} B_1 \\ B_2 \end{pmatrix}, && B_1 \in \mathbb{R}^{r\times m}, \\ C_S &= CS^{-1} = \begin{pmatrix} C_1 & C_2 \end{pmatrix}, && C_1 \in \mathbb{R}^{p\times r}. \end{aligned}$$

Hence, the model for G_r is configured as:

$$\begin{bmatrix} A_r & B_r & C_r \end{bmatrix} = \begin{bmatrix} A_{11} & B_1 & C_1 \end{bmatrix}.$$

This technique of orthogonal truncation serves as a projection method, enabling the reduction from the original system G in $\mathbb{R}^n$ to a more manageable system G_r in $\mathbb{R}^r$.

3.1.3.2 Petrov Projection and Galerkin Projection

Within the framework of non-orthogonal projection methodologies, transformations labeled as M and V are designed as

$$M^\top = \begin{pmatrix} I_r & 0_{r\times(n-r)} \end{pmatrix} S, \quad \mathbb{R}^n \to \mathbb{R}^r \tag{3.12}$$

$$V = S^{-1} \begin{pmatrix} I_r \\ 0^\top_{r\times(n-r)} \end{pmatrix}, \quad \mathbb{R}^r \to \mathbb{R}^n \tag{3.13}$$

with the purpose of enabling the projection from the vector space $\mathbb{R}^n$ to the vector space $\mathbb{R}^r$, as well as its corresponding inverse operation. It is worth noting that these transformations satisfy

$$M^\top V = I_r, \quad VM^\top = (VM^\top)(VM^\top) \tag{3.14}$$

Such projection methodologies fall under the classification of a Petrov-Galerkin projection. The Galerkin projection, characterized by the condition $M = V$, which means that S denotes an orthogonal matrix, is used to transform the model from G in $\mathbb{R}^n$ to G_r in $\mathbb{R}^r$. For the Petrov Galerkin projection, we have

$$\begin{bmatrix} A_r & B_r & C_r \end{bmatrix} = \begin{bmatrix} M^\top AV & M^\top B & CV \end{bmatrix}.$$

In order to understand the Petrov-Galerkin projection, let us consider the following analysis. Suppose we aim to represent the solution $x(t) \in \mathbb{R}^n$ of the original model G using only r variables in the ROM G_r. We can write

$$x(t) = Vx_r(t) \in \text{Range}(V), \quad \text{where} \quad x(t) \in \mathbb{R}^n, \quad x_r(t) \in \mathbb{R}^r.$$

As a consequence, the residual R is expressed as:

$$R = V\dot{x}_r(t) - [AVx_r(t) + Bu(t)].$$

The Petrov-Galerkin projection requires that the projection of the residual R into the range (W) is zero. This condition can be expressed as $M^\top R = 0$. This equation yields

$$M^\top(V\dot{x}_r(t) - AVx_r(t) - Bu(t)) = \dot{x}_r(t) - M^\top AVx_r(t) - M^\top Bu(t) = 0,$$

resulting in

$$\dot{x}_r(t) = M^\top AVx_r(t) + M^\top Bu(t). \tag{3.15}$$

In the upcoming section, we present the Balanced Truncation method. However, before doing so, we will clarify the notion of "balance" as used in this context.

Remark 3.6 A stable system (3.1), represented by G, is said to be balanced if the Gramians, i.e., the solutions P and Q to the Lyapunov equations (3.2a) and (3.2b),

satisfy the following condition:

$$P = Q = \text{diag}\,(\sigma_1, \ldots, \sigma_n)$$

where the singular values are ordered as $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n > 0$.

3.1.3.3 Balanced Truncation

This section focuses on the utilization of Balanced Truncation (BT) as a method to approximate a system. Using Singular Value Decomposition (SVD), see Appendix A.1.1, a matrix can be approximated to its lower-rank form, optimized via the L^2 norm. This technique forms the basis for methodologies in model reduction for linear dynamical systems, such as BT method. BT aligns the reachability and observability of states, initially outlined by [76] and further detailed by [74]. A specific BT variant, Lyapunov Balanced Reduction (LBR), is implemented by solving the reachability and observability Lyapunov equations in tandem. The stability of this method was established by [89], with error bounds provided by [33]. This method is one among several balancing strategies.

We rewrite the algebraic Lyapunov equations (3.2) as follow:

$$AP + PA^\top + BB^\top = 0,$$
$$A^\top Q + QA + C^\top C = 0.$$

The eigenvalues derived from the product of the controllability and observability Grammians are known as Hankel singular values (HSVs), which are denoted by $\sigma_i = (\lambda_i(PQ))^{\frac{1}{2}}$.The significance of these singular values, which remain unchanged under coordinate transformations, can be seen in the areas of system theory and control.

The strategy of reducing the state space by deleting the dimensions that are least observable or controllable, resulting in a lower-order approximate model, can be achieved through a five-step procedure:

(i) Compute the reachability Grammian P and observability Grammian Q,

(ii) Compute the Cholesky factor R of Q and L of P such that $Q = R^\top R$ and $P = LL^\top$, so we have

$$PQ = (LL^\top)(R^\top R) = L(RL)^\top R,$$

(iii) Compute the SVD $RL = V\Sigma U^\top$, where U and V are unitary matrices. Then, multiplying from the left by $V^\top$ and from the right by U we have

$$V^\top RLU = \Sigma,$$

and finally multiplying from the left and from the right by $\Sigma^{-1/2}$, we obtain

$$\Sigma^{-1/2} V^\top RLU \Sigma^{-1/2} = I.$$

(iv) Use the coordinate transformation $x_S(t) = Sx(t)$, where

$$S = \Sigma^{-1/2} V^\top R, \quad S^{-1} = LU\Sigma^{-1/2}.$$

(v) Compute the truncation system G_r (3.10) after transforming the state-space system G (3.1).

In the transformed coordinate system, denoted by

$$\begin{aligned} A_S &= SAS^{-1} = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}, && A_{11} \in \mathbb{R}^{r\times r} \\ B_S &= SB = \begin{pmatrix} B_1 \\ B_2 \end{pmatrix}, && B_1 \in \mathbb{R}^{r\times m} \\ C_S &= CS^{-1} = \begin{pmatrix} C_1 & C_2 \end{pmatrix}, && C_1 \in \mathbb{R}^{p\times r} \end{aligned} \tag{3.16}$$

so we have

$$[A_r \quad B_r \quad C_r] = [A_{11} \quad B_1 \quad C_1],$$

the transformed Grammians $P_S = SPS^\top$ and $Q_S = S^{-\top} Q S^{-1}$ can be obtained.

$$\begin{aligned} P_S &= SPS^\top = \left(\Sigma^{-\frac{1}{2}} V^\top R\right) P \left(R^\top V \Sigma^{-\frac{1}{2}}\right) = \left(\Sigma^{-\frac{1}{2}} V^\top\right) RL(RL)^\top \left(V\Sigma^{-\frac{1}{2}}\right) \\ &= \left(\Sigma^{-\frac{1}{2}} V^\top\right) V\Sigma^2 V^\top \left(V\Sigma^{-\frac{1}{2}}\right) = \left(\Sigma^{-\frac{1}{2}}\right) \Sigma^2 \left(\Sigma^{-\frac{1}{2}}\right) = \Sigma, \\ Q_S &= S^{-\top} Q S^{-1} = \left(LU\Sigma^{-\frac{1}{2}}\right)^\top Q \left(LU\Sigma^{-\frac{1}{2}}\right) = \left(\Sigma^{-\frac{1}{2}} U^\top\right) (RL)^\top RL \left(U\Sigma^{-\frac{1}{2}}\right) \\ &= \left(\Sigma^{-\frac{1}{2}} U^\top\right) U\Sigma^2 U^\top \left(U\Sigma^{-\frac{1}{2}}\right) = \left(\Sigma^{-\frac{1}{2}}\right) \Sigma^2 \left(\Sigma^{-\frac{1}{2}}\right) = \Sigma. \end{aligned}$$

Therefore, according to Remark 3.6, the Gramians P_S and Q_S characterize the system in a balanced form. The approximation by BT method ensures stability preservation by the following theorem (see[5, 74])

Theorem 3.7 (Stability preservation) Given a homogeneous linear time-invariant system (3.1), if $\sigma(\mathcal{L}_A) \subseteq \mathbb{C}_-$ and a block diagonal matrix $\Sigma = \operatorname{diag}(\Sigma_1, \Sigma_2) > 0$ with $\sigma(\Sigma_1) \cap \sigma(\Sigma_2) = \emptyset$ exists, satisfying

$$\mathcal{L}_A(\Sigma) \le 0 \quad \text{and} \quad \mathcal{L}_A^*(\Sigma) \le 0,$$

then for the balanced truncated system, we have $\sigma(\mathcal{L}_{A_r}) \subseteq \mathbb{C}_-$ where

$$\mathcal{L}_{A_r} = A_r \otimes I + I \otimes A_r. \tag{3.17}$$

Additionally, the H_∞ norm (representing the maximum frequency response) of the error system remains bounded by the following theorem.

Theorem 3.8 Let G be a system as in (3.1), where $\mathcal{L}_A \subseteq \mathbb{C}_-$. When a balanced truncation is performed of r-th order, the resultant system is denoted as G_r. The

bound on the error due to truncation is given by:

$$||y - y_r||_\infty \leq 2(\sigma_{r+1} + \ldots + \sigma_n).$$

3.1.3.4 Proper Orthogonal Decomposition

The development of POD was a collective effort by various scientists. In particular, Karhunen [56] and Loéve [67] contributed significantly, giving rise to what is known as the Karhunen-Loéve decomposition. Other important contributions to the field include [63, 92, 15]. The POD method has been successfully used in a wide variety of scientific and engineering problems, such as fluid dynamics [42, 49, 107, 112], electric circuit analysis [88], or structural dynamics [4]. For an expansive review of the history and diverse applications of the POD, [20, 22, 58, 65, 118] are recommended.

In addressing the challenges posed by the analysis and simulation of large-scale nonlinear dynamical systems, MOR techniques serve as essential tools. This section elaborates on the application of a prominent method: the POD method also termed Principal Component Analysis (PCA). These methods are pivotal in addressing the complexities of nonlinear systems derived from PDEs, especially after their discretization into finite-dimensional dynamical systems.

We consider the dynamics of nonlinear systems represented by the general form:

$$\begin{aligned} \dot{x}(t) &= Ax(t) + f(x(t)) + Bu(t), \quad x(0) = x_0, \\ y(t) &= Cx(t), \end{aligned} \tag{3.18}$$

where matrices A, B and C respectively define the linear dynamics, control inputs effects and observation matrix, as described in system (3.1). $x(t) \in \mathbb{R}^n$ denotes the state vector at time t and $f : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ represents the non-linear interactions within the system. The primary goal is to construct a ROM that accurately captures the essential dynamics of (3.18) while significantly reducing computational complexity.

The strategy for reducing the state space through the application of the POD method, which yields a lower-order approximate model, can be systematically achieved by the following three-step procedure:

(i) **Data Collection and Snapshot Matrix Construction:** Conduct simulations using specified initial conditions and inputs to collect data. These data are structured into a snapshot matrix Z, capturing dynamics of the system in discrete time instances:

$$Z = [x(t_0)\ x(t_1)\ \ldots\ x(t_N)].$$

(ii) **Extraction of the Dominant Subspace via SVD:** Apply SVD to the snapshot matrix Z to decompose it and extract the most influential modes of system behavior. This decomposition,

$$Z = [V_r\ V_{\bar{r}}] \begin{bmatrix} \Sigma_r & 0 \\ 0 & \Sigma_{\bar{r}} \end{bmatrix} \begin{bmatrix} U_r^\top \\ U_{\bar{r}}^\top \end{bmatrix} \quad \rightarrow \quad \text{POD basis } V_r = [v_1, \ldots, v_r]$$

where $V_r^\top V_r = I_r$, isolates the significant singular values (and their corresponding vectors) which are then used to construct the POD basis V_r. The basis vectors v_i in V_r span the dominant subspace capturing the core dynamics.

(iii) **Galerkin Projection and Reduced Model Derivation:** Utilize the POD basis for Galerkin projection (See Subsection 3.1.3.2) on the original system, leading to a reduced model. This model includes the system dynamics within a reduced-dimensional space, maintaining a balance between simplicity and accuracy. It is characterized by a reduced linear matrix, nonlinearity and input component, as shown in

$$\begin{aligned} \dot{\bar{x}}(t) &= V_r^\top A V_r \bar{x}(t) + V_r^\top f(V_r \bar{x}(t)) + V_r^\top B u(t), \quad \Rightarrow \quad V_r \bar{x}(t) \approx x(t), \\ y_r(t) &= C V_r \bar{x}(t). \end{aligned}$$

The essence of MOR lies in the selection of an optimal subspace, represented by V_r, that ensures the reduced model's output $y_r(t)$ closely approximates the original system output $y(t)$ with minimal error, which means that

$$\|y(t) - y_r(t)\| \ll \varepsilon,$$

effectively capturing the essential dynamics of the system within a lower-dimensional framework.

This MOR strategy employs snapshot-based approaches to efficiently handle nonlinear systems where traditional analytical solutions are infeasible. Using the dynamic data stored in the snapshot matrix and carefully selecting the projection subspace, POD enables the creation of efficient and representative ROMs. For a more detailed exploration of POD in combination with SVD, refer to [118].

Remark 3.9 Selecting the appropriate dimension r for the POD basis is essential and is typically determined by analyzing the decay rate of singular values to assess their significance. Additionally, the effectiveness of the POD method heavily relies on the conditions of the inputs used in initial simulations, emphasizing the necessity for precise and comprehensive data collection.

3.2 linear Stochastic Dynamical Systems

Consider a q-dimensional Wiener process $W(t) = (W_1(t), \ldots, W_q(t))^\top$ with zero mean and a covariance matrix $\mathbf{K} = (k_{ij})_{i,j=1,\ldots,q}$ on the interval $[0, T]$, where $T > 0$. This process is part of a stochastic system on a complete, right-continuous filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$. We assume that W is adapted to the filtration $(\mathcal{F}_t)_{t\in[0,T]}$ and the increments $W(t+h) - W(t)$ are independent of $\mathcal{F}_t$ for $t, h \geq 0$. Consider a controlled, high-dimensional linear SDE system of Itô type(see Section

1.3) is given by

$$\mathrm{d}x(t) = [Ax(t) + Bu(t)]\mathrm{d}t + \sum_{i=1}^{q} N_i x(t)\mathrm{d}W_i(t), \qquad x(0) = x_0, \tag{3.19a}$$

$$y(t) = Cx(t), \qquad t \geq 0, \tag{3.19b}$$

with $A, N_i \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $C \in \mathbb{R}^{p\times n}$ and the state $x(t; x_0, u)$ depending on the initial state x_0 and control u. We consider $\mathcal{M}^2([0,T];\mathbb{R}^m)$ as the set of all $(\mathcal{F}_t)_{t\geq 0}$-adapted stochastic processes v with values in $\mathbb{R}^m$ that are square-integrable with respect to the measure $\mathbb{P}\otimes \mathrm{d}t$ (see Definition 1.14). The energy norm in this space is expressed as

$$\|v\|_T^2 := \mathbb{E}\left[\int_0^T v^\top(t)v(t)\,\mathrm{d}t\right] = \mathbb{E}\left[\int_0^T \|v(t)\|_2^2\,\mathrm{d}t\right], \tag{3.20}$$

where $\|\cdot\|_2$ represents the Euclidean norm and two processes v_1 and v_2 belonging to $\mathcal{M}^2([0,T];\mathbb{R}^m)$ are deemed equivalent if they are almost surely identical with respect to $\mathbb{P}\otimes \mathrm{d}t$. For the infinite horizon, the space is denoted by $\mathcal{M}^2(\mathbb{R}_+;\mathbb{R}^m)$.

The stochastic system (3.19) can be identified as $G_N, (A, N, B, C)$ or (A, N).

Definition 3.10 An $\mathbb{R}^n$-valued process $\{x(t)\}_{t\geq 0}$ is said to be the solution to the SDE (3.19) if it satisfies:

$$x(t; x_0, u) = x_0 + \int_0^t [Ax(s) + Bu(s)]\mathrm{d}s + \sum_{i=1}^{q}\int_0^t N_i x(s)\mathrm{d}W_i(s)$$

almost surely for all $t \geq 0$, where $x_0 \in \mathbb{R}^n$ represents the initial condition and $u \in \mathcal{M}^2([0,T];\mathbb{R}^m)$ is control term. For the solution of (3.19) in the uncontrolled case, i.e,

$$\mathrm{d}x(t) = Ax(t)\mathrm{d}t + \sum_{i=1}^{q} N_i x(t)\mathrm{d}W_i(t), \quad y(t) = Cx(t), \quad t \geq 0, \quad x(0) = x_0, \tag{3.21}$$

We briefly write $x_{x_0} := x\,(t, x_0, 0)$, where x_{x_0} is called homogeneous solution.

Theorem 3.11 Let x_{x_0} be the solution of system (3.21) with any initial value $x_0 \in \mathbb{R}^n$, then $\mathrm{Y}(t) = \mathbb{E}\left[x_{x_0}(t)x_{x_0}^\top(t)\right]$ is the solution of the matrix integral equation

$$\mathrm{Y}(t) = x_0x_0^\top + \int_0^t Y(s)\mathrm{d}sA^\top + A\int_0^t Y(s)\mathrm{d}s + \sum_{i,j=1}^{q} N_i\int_0^t Y(s)\mathrm{d}sN_j^\top k_{i,j}, \tag{3.22}$$

for $t \geq 0$.

Proof. See Theorem 1.4.3 in [25]. □

The solutions of the SDE (3.19) and (3.21) satisfy the conditions of Theorem 1.29, which guarantees the existence and uniqueness of their solutions. The homogeneous system (3.21) corresponds to the fundamental matrix solution

$$\Phi(t,\tau) = I_n + \int_\tau^t A\Phi(s,\tau)\mathrm{d}s + \sum_{i=1}^{q}\int_\tau^t N_i\Phi(s,\tau)\mathrm{d}W_i(s), \quad t \geq 0, \tag{3.23}$$

for $t \geq \tau \geq 0$. We have

$$\Phi(t,0) = \Phi(t) = I_n + \int_0^t A\Phi(s)\mathrm{d}s + \sum_{i=1}^{q}\int_0^t N_i\Phi(s)\mathrm{d}W_i(s), \quad t \geq 0, \tag{3.24}$$

and I_n is the identity matrix. Hence, the general solution x_{x_0} of uncontrolled SDE (3.21) with initial condition $x(0) = x_0$ can be expressed as $x_{x_0} = \Phi(t)x_0$.

As a result, considering $t \geq \tau \geq 0$, we can express $\Phi(t)$, defined in equation (3.24), as follows:

$$\begin{aligned}\Phi(t) &= I + \int_0^\tau A\Phi(s)\,\mathrm{d}s + \int_\tau^t A\Phi(s)\,\mathrm{d}s \\ &\quad + \sum_{i=1}^{q}\left[\int_0^\tau N_i\Phi(s)\,\mathrm{d}W_i(s) + \int_\tau^t N_i\Phi(s)\,\mathrm{d}W_i(s)\right] \\ &= \Phi(\tau) + \int_\tau^t A\Phi(s)\,\mathrm{d}s + \sum_{i=1}^{q}\int_\tau^t N_i\Phi(s)\,\mathrm{d}W_i(s).\end{aligned}$$

Upon multiplying both sides of the above equation by $\Phi^{-1}(\tau)$ from the right, we obtain:

$$\Phi(t)\Phi^{-1}(\tau) = I_n + \int_\tau^t A\Phi(s)\Phi^{-1}(\tau)\mathrm{d}s + \sum_{i=1}^{q}\int_\tau^t N_i\Phi(s)\Phi^{-1}(\tau)\mathrm{d}W_i(s). \tag{3.25}$$

As evident from equation (3.25), it aligns with the original formulation given in equation (3.23). We summarize this result in the following remark:

Remark 3.12 For $t \geq \tau \geq 0$, we have $\Phi(t,\tau) = \Phi(t)\Phi^{-1}(\tau)$, since $\Phi(t)\Phi^{-1}(\tau)$ satisfies equation (3.23).

We define the following proposition as an interesting result that we use in following chapters.

Proposition 3.13 [97] Assume that all matrices $A, N_1, \ldots, N_q$ commute. Hence, these matrices commute with the fundamental solution Φ, i.e.

$$A\Phi(t) = \Phi(t)A \quad \text{and} \quad N_i\Phi(t) = \Phi(t)N_i$$

for all $t \in [0,T]$ and $i = 1, \ldots, q$.

A drawback of the stochastic system is the inability to provide a general explicit description of the fundamental matrix solution. However, a specific case of an explicit fundamental matrix solution $\Phi(t)$ occurs when all matrices A and N_i for $i = 1, \ldots, q$ in equation (3.21) commute, as seen below.

Proposition 3.14 Assume that all A, N_i commute. Then the fundamental solution $\Phi(t)$ of (3.21) is

$$\Phi(t) = \exp\left(\left(A - \frac{1}{2}\sum_{i,j=1}^{q}(N_iN_j)\,k_{ij}\right)t + \sum_{i=1}^{q}N_iW_i(t)\right).$$

Proof. We write $Z(t)$ for the exponent in (3.21), such that

$$\mathrm{d}Z(t) = \left(A - \frac{1}{2}\sum_{i,j=1}^{q}(N_iN_j)\,k_{ij}\right)\mathrm{d}t + \sum_{i=1}^{q}N_i\mathrm{d}W_i(t).$$

Then we have

$$\begin{aligned} d\Phi(t) &= e^{Z(t)}\mathrm{d}Z(t) + \frac{1}{2}e^{Z(t)}(\mathrm{d}Z(t))^2 \\ &= \Phi(t)\left(\left(A - \frac{1}{2}\sum_{i,j=1}^{q}(N_iN_j)\,k_{ij}\right)\mathrm{d}t + \sum_{i=1}^{q}N_i\mathrm{d}W_i(t) + \frac{1}{2}\sum_{i,j=1}^{q}(N_iN_j)\,k_{ij}\mathrm{d}t\right) \\ &= A\Phi(t)\mathrm{d}t + \sum_{i=1}^{q}N_i\Phi(t)\mathrm{d}W_i(t), \end{aligned}$$

since $\Phi(t)$ commutes by assumption with A and all N_i for $i \in \{1, \ldots, q\}$. □

We now proceed to the nonhomogeneous SDE presented in (3.19), employing the fundamental matrix solution (3.24), as described in the following theorem.

Theorem 3.15 Given the fundamental matrix solution $\Phi(t)$, we obtain as the solution of (3.19)

$$x(t) = \Phi(t)\left(x_0 + \int_0^T \Phi^{-1}(s)Bu(s)ds\right). \tag{3.26}$$

Proof. For a more detailed exposition of the proof, which is provided in a broader context, reference [25] is suggested. □

We proceed by expanding the generalized Lyapunov equations stated in (3.2) to cover the stochastic system described in (3.19).

3.2.1 Lyapunov Equations in Stochastic Case

Lyapunov equations play a critical role in characterizing the stability and control properties of linear stochastic systems. For the system described by (3.19), the

corresponding Lyapunov equations can be obtained in a generalized form as:

$$AP + PA^\top + \sum_{i,j=1}^{q} N_i P N_j^\top k_{ij} = -BB^\top, \tag{3.27a}$$

$$A^\top Q + QA + \sum_{i,j=1}^{q} N_i^\top Q N_j k_{ij} = -C^\top C. \tag{3.27b}$$

The symmetric matrices P and Q from $\mathbb{R}^{n\times n}$, which satisfy the semi-positive definite conditions $P \geq 0$ and $Q \geq 0$, uniquely solve these Lyapunov equations under specific system conditions. These conditions are closely connected to the system's observability and reachability, key concepts that can be defined as follows (adapted from [10]):

Definition 3.16 (Unobservability and Unreachability): Consider a system as described in (3.19). A vector $v \in \mathbb{R}^n$ is termed:

- *Unobservable*, if for the initial condition $x(0) = v$ with $u \equiv 0$, the output $y \equiv 0$.
- *Unreachable*, if for any input u and for all $t > 0$, $x(t) \neq v$ when starting from the initial condition $x(0) = 0$.

Based on the stability conditions of the system, particularly the asymptotic mean-square stability of the system (3.19), the subspaces associated with unobservability and unreachability can be described through the kernels of Q and P, respectively.

Theorem 3.17 [10] Additionally outlines

(a) A state v is *unobservable* if and only if $Qv = 0$.

(b) A state v is *unreachable* if and only if $Pv = 0$.

Crucially, the system is completely observable and reachable if and only if $Q > 0$ and $P > 0$ respectively. These conditions ensure that all states significantly affect the system's outputs and can be influenced by the inputs, aligning with the foundational requirements for effective control and monitoring. The stochastic version of Gramian matrices are defined as follows

$$P = \mathbb{E}\left(\int_0^\infty \Phi(\tau) B B^\top \Phi(\tau)\, d\tau\right), \tag{3.28a}$$

$$Q = \mathbb{E}\left(\int_0^\infty \Phi(\tau) C^\top C \Phi(\tau)\, d\tau\right). \tag{3.28b}$$

The existence of the infinite integrals in (3.28) is guaranteed by the mean square asymptotic stability of the system (Theorem 3.19). Now, we present the generalized Lyapunov equations (3.27) for the linear stochastic system using the Kronecker product notation, while preserving the analogy with the linear deterministic system.

3.2.1.1 Vectorization and Kronecker product in stochastic case

By utilizing the Kronecker product notation, the generalized Lyapunov equations from (3.27) can be reformulated as follows:

$$\left(A \otimes I + I \otimes A + \sum_{i,j=1}^{q} N_i \otimes N_j k_{ij}\right) \operatorname{vec}(P) = -\operatorname{vec}\left(BB^\top\right), \tag{3.29a}$$

$$\left(A^\top \otimes I + I \otimes A^\top + \sum_{i,j=1}^{q} N_i^\top \otimes N_j^\top k_{ij}\right) \operatorname{vec}(Q) = -\operatorname{vec}\left(C^\top C\right). \tag{3.29b}$$

where, as mentioned before, $\cdot \otimes \cdot$ is the Kronecker product between two matrices and $\operatorname{vec}(\cdot)$ is the vectorization of a matrix (see Appendix B for more details). Let us review the definition of the operator $\mathcal{L}_A$ as stated in equations (3.5) and (3.7) and subsequently introduce the operator $\Pi_N : \mathcal{S}_n \to \mathcal{S}_n$, where $\mathcal{S}_n$ is the space of $n \times n$ symmetric matrices, given by

$$\Pi_N(X) := \sum_{i,j=1}^{q} N_i X N_j^\top k_{ij}, \quad \text{or} \quad \Pi_N := \sum_{i,j=1}^{q} N_i \otimes N_j k_{ij}, \tag{3.30}$$

with adjoints concerning the Frobenius inner product given by

$$\Pi_N^*(X) := \sum_{i,j=1}^{q} N_i^\top X N_j k_{ij}, \quad \text{or} \quad \Pi_N^* := \sum_{i,j=1}^{q} N_i^\top \otimes N_j^\top k_{ij}. \tag{3.31}$$

Based on these definitions, we may establish the following inequalities for the stochastic system, as compared to the deterministic inequalities in (3.9).

$$\mathcal{L}_A(P) + \Pi_N(P) \leq -BB^\top, \tag{3.32}$$

$$\mathcal{L}_A^*(Q) + \Pi_N^*(Q) \leq -C^\top C. \tag{3.33}$$

As we have outlined the roles of the operators $\mathcal{L}_A$ and Π_N in the generalized Lyapunov equations, we now aim to establish a more direct mathematical equivalence using the Kronecker product formulation. This approach not only simplifies the representation, but also enhances our understanding of the interactions between these operators and the matrix structures within the stochastic system framework. To this end, we introduce the following remark that contains the equivalence of these formulations under the Kronecker product notation.

Remark 3.18 Defining

$$\mathcal{K} = A \otimes I + I \otimes A + \sum_{i,j=1}^{q} N_i \otimes N_j k_{ij}, \tag{3.34}$$

$$\mathcal{K}^\top = A^\top \otimes I + I \otimes A^\top + \sum_{i,j=1}^{q} N_i^\top \otimes N_j^\top k_{ij}. \tag{3.35}$$

This establishes the equivalence of equation (3.29) by employing equations (3.5), (3.6), (3.30) and (3.31) as follows:

$$\mathcal{K}\,\mathrm{vec}(P) = (\mathcal{L}_A + \Pi_N)\,\mathrm{vec}(P) = -\,\mathrm{vec}\left(BB^\top\right), \tag{3.36a}$$

$$\mathcal{K}^\top\,\mathrm{vec}(Q) = (\mathcal{L}_A^* + \Pi_N^*)\,\mathrm{vec}(Q) = -\,\mathrm{vec}\left(C^\top C\right). \tag{3.36b}$$

Next, we introduce the stability of the stochastic system to the generalized Lyapunov equations and their solutions.

3.2.2 Stability

The following theorem presents equivalent criteria for ensuring the asymptotic mean-square stability of equation (3.21), similar to Theorem 3.5. For more details about mean-square stability, see Section 1.7.2.

Theorem 3.19 [25](Stability of Stochastic Linear System) The criteria below are equivalent:

(i) Equation (3.21) achieves asymptotic mean-square stability (see Definition 1.35).

(ii) Equation (3.21) exhibits exponential mean-square stability (see Definition 1.36).

(iii) $\sigma\left(\mathcal{L}_A + \Pi_N\right) \subset \mathbb{C}_-$ or equivalently $\sigma\left(\mathcal{K}\right) \subset \mathbb{C}_-$.

(iv) $\exists X > 0$ such that $\mathcal{L}_A(X) + \Pi_N(X) < 0$.

(v) $\forall Y < 0 : \exists X > 0$ such that $\mathcal{L}_A(X) + \Pi_N(X) = Y$.

3.2.3 Characterization of Dominant Subspaces Using Gramian Matrices

Let us explore the relation between P and the dominant subspaces in (3.19a) for the case of zero initial data. We recall an argument from [96, Section 3] blow. To obtain this relationship, consider the matrix partition of B:

$$B = \begin{bmatrix} b_1 & \dots & b_m \end{bmatrix},$$

which leads to

$$\Phi(t)B = \begin{bmatrix} x(t,b_1,0) & \dots & x(t,b_m,0) \end{bmatrix} = \begin{bmatrix} x_{b_1} & \dots & x_{b_m} \end{bmatrix},$$

thus, we have the following identity:

$$\mathbb{E}\left[\Phi(t)BB^\top\Phi^\top(t)\right] = \sum_{k=1}^{m} \mathbb{E}\left[x_{b_k}x_{b_k}^\top\right].$$

Applying Theorem 3.11 to each summand leads to:

$$\begin{aligned}\mathbb{E}\left[\Phi(t)BB^\top\Phi^\top(t)\right] = BB^\top &+ A\int_0^t \mathbb{E}\left[\Phi(s)BB^\top\Phi^\top(s)\right]\mathrm{d}s \\ &+ \int_0^t \mathbb{E}\left[\Phi(s)BB^\top\Phi^\top(s)\right]\mathrm{d}sA^\top \\ &+ \sum_{i,j=1}^{q} N_i \int_0^t \mathbb{E}\left[\Phi(s)BB^\top\Phi^\top(s)\right] N_j^\top k_{ij}. \end{aligned} \tag{3.37}$$

By considering the limit as $t \to \infty$ and assuming asymptotic mean square stability (Theorem 3.19 (i)), the left-hand side tends to zero, thus confirming equation (3.27a).

We now proceed to analyze the process $\langle x(t;0,u),\rho\rangle_2$, where $\rho \in \mathbb{R}^n$, setting $\Phi(t,s) = \Phi(t)\Phi^{-1}(s)$ for $t \geq s \geq 0$ (Remark 3.12). Inserting equation (3.26) yields the following bound:

$$\begin{aligned}\mathbb{E}\left|\langle x(t;0,u),\rho\rangle_2\right| &= \mathbb{E}\left|\int_0^t \langle \rho, \Phi(t,s)Bu(s)\rangle_2 \mathrm{d}s\right| \\ &= \mathbb{E}\left|\int_0^t \langle B^\top\Phi^\top(t,s)\rho, u(s)\rangle_2 \mathrm{d}s\right| \\ &\leq \mathbb{E}\int_0^t \left\|B^\top\Phi^\top(t,s)\rho\right\|_2 \|u(s)\|_2 \mathrm{d}s.\end{aligned}$$

By Cauchy's inequality, it follows that

$$\mathbb{E}\left|\langle x(t;0,u),\rho\rangle_2\right| \leq \left(\mathbb{E}\int_0^t \left\|B^\top\Phi^\top(t,s)\rho\right\|_2^2 \mathrm{d}s\right)^{\frac{1}{2}} \left(\mathbb{E}\int_0^t \|u(s)\|_2^2 \mathrm{d}s\right)^{\frac{1}{2}}.$$

Remark 3.20 [14] In stochastic systems, the fundamental matrix $\Phi(t,\tau)$ is defined differently from its deterministic counterparts. Unlike deterministic systems where the semigroup property $\Phi(t,\tau) = \Phi(t-\tau)$ typically holds, in stochastic contexts, this property does not apply $\mathbb{P}$-almost surely due to variability in noise trajectories. This highlights a fundamental difference in behavior between stochastic and deterministic systems. Furthermore, although a direct comparison of $\Phi(t,\tau)$ and $\Phi(t-\tau)$ is not possible on a pointwise basis due to different noise influences over intervals $[0, t-\tau]$

and $[\tau, t]$, their expected values related to system behavior are equivalent, which means that

$$\mathbb{E}\left[\Phi(t,\tau)BB^\top\Phi^\top(t,\tau)\right] = \mathbb{E}\left[\Phi(t-\tau)BB^\top\Phi^\top(t-\tau)\right].$$

This equivalence is derived since both terms solve the equation (3.22). Therefore, both sides conform to the integral equation (3.37) from the initial time $\tau \leq t \leq T$, which has a unique solution.

Consequently, based on Remark 3.20, we derive:

$$\begin{aligned}\mathbb{E}\int_0^t \left\|B^\top\Phi^\top(t,\tau)\rho\right\|_2^2 \mathrm{d}s &= \rho^\top\mathbb{E}\int_0^t \Phi(t,\tau)BB^\top\Phi^\top(t,\tau)\mathrm{d}s\,\rho \\ &= \rho^\top\mathbb{E}\int_0^t \Phi(t-\tau)BB^\top\Phi^\top(t-\tau)\mathrm{d}s\,\rho \\ &= \rho^\top\mathbb{E}\int_0^t \Phi(s)BB^\top\Phi^\top(s)\mathrm{d}s\,\rho \leq \rho^\top P\rho,\end{aligned}$$

and consequently,

$$\sup_{t\in[0,T]} \mathbb{E}\left|\langle x(t;0,u),\rho\rangle_2\right| \leq \left(\rho^\top P\rho\right)^{1/2}\|u\|_T. \tag{3.38}$$

If $\rho \in \ker P$, then the right-hand side of (3.38) is zero, which implies that $\langle x(t;0,u),\rho\rangle_2 = 0$, for all $t \in [0,T]$, $\mathbb{P}$-almost surely, regardless of the control u used. This indicates that the trajectories of x are orthogonal to $\ker P$ and thus,

$$\mathbb{P}\left\{x(t;0,u) \in \operatorname{im} P, t \in [0,T]\right\} = 1,$$

for every $u \in \mathcal{M}^2([0,T];\mathbb{R}^m)$, indicating that no state outside $\operatorname{im} P$ is reachable (from zero). Let $\{\rho_k\}_{k=1,\dots,n}$ be an orthonormal basis of $\mathbb{R}^n$, consisting of eigenvectors of P. The representation

$$x(t;0,u) = \sum_{k=1}^{n} \langle x(t;0,u),\rho_k\rangle_2\,\rho_k,$$

the Fourier coefficient can be bound from (3.38) as follows

$$\sup_{t\in[0,T]} \mathbb{E}\left|\langle x(t;0,u),\rho_k\rangle_2\right| \leq \lambda_k^{\frac{1}{2}}\|u\|_T, \tag{3.39}$$

where λ_k is the eigenvalue corresponding to ρ_k. If λ_k is small, the same is true for $\langle x(\cdot,0,u),\rho_k\rangle_2$ and therefore ρ_k is a less relevant direction that can be neglected.

To quantify the energy contained by the initial states x_0 via observations, we consider the observability Gramian matrix Q as the unique solution to the matrix

differential equation (3.27b):

$$A^\top Q + QA + \sum_{i,j=1}^{q} N_i^\top Q N_j k_{ij} = -C^\top C.$$

This formulation of Q is justified under the first part of Theorem 3.19. Using the relation

$$\mathbb{E}\left[x_{x_0}^\top Q x_{x_0}\right] = \operatorname{tr}\left(Q\mathbb{E}\left[x_{x_0}x_{x_0}^\top\right]\right),$$

we can integrate the result from Theorem 3.11. Observing that the trace operator allows for permutation within a matrix product, we derive:

$$\mathbb{E}\left[x_{x_0}^\top Q x_{x_0}\right] = x_0^\top Q x_0 + \mathbb{E}\int_0^t x(s,x_0,0)^\top \left(A^\top Q + QA + \sum_{i,j=1}^{q} N_i^\top Q N_j k_{ij}\right) x(s,x_0,0)\mathrm{d}s,$$

Substituting (3.27b) into this equation, we get:

$$\mathbb{E}\left[x_{x_0}^\top Q x_{x_0}\right] = x_0^\top Q x_0 - \mathbb{E}\int_0^t x(s,x_0,0)^\top C^\top C x(s,x_0,0)\mathrm{d}s, \tag{3.40}$$

given the mean square asymptotic stability of system (3.19a), the left side of equation (3.40) vanishes as $t \to \infty$. Thus, the energy observed is expressed by:

$$\mathbb{E}\int_0^\infty \|y(s,x_0,0)\|_2^2\mathrm{d}s = \mathbb{E}\int_0^\infty \|Cx(s,x_0,0)\|_2^2\mathrm{d}s = x_0^\top Q x_0. \tag{3.41}$$

From equation (3.41), we derive that the states which are difficult to observe, those that contribute minimally to the overall observation energy, are characterized by producing lower values of $x_0^\top Q x_0$. These states are hence considered unimportant in the context of the system's output observability because they have minimal influence on the output data collected over an infinite time horizon.

Essentially, equation (3.41) clarifies that such "difficult-to-observe" states are principally aligned with the eigenspaces of the observability Gramian Q corresponding to its smaller eigenvalues. The eigenvectors associated with these smaller eigenvalues define directions in the state space that are inherently less detectable through output measurements. Thus, the magnitude of the eigenvalues of Q serves as an indicator of the degree of observability of the corresponding state directions: smaller eigenvalues signify lower observability.

Additionally, the practical expression for Q (equation (3.28b)) can be derived by substituting the solution form $Cx_{x_0} = C\Phi(t)x_0$ into equation (3.41), resulting in the following:

$$x_0^\top \mathbb{E}\int_0^\infty \Phi(t)^\top C^\top C\Phi(t)\mathrm{d}t x_0 = x_0^\top Q x_0,$$

since this is valid for any $x_0 \in \mathbb{R}^n$, we conclude:

$$Q = \mathbb{E}\int_0^\infty \Phi(t)^\top C^\top C\Phi(t)\mathrm{d}t.$$

In summary, our analysis shows that identifying and removing less important state directions from equations (3.19a) and (3.19b), along with findings from (3.39) and (3.41),

simplifies the model. This simplification makes the model easier to manage and faster to run without losing accuracy. By focusing only on the essential parts of the system, we use computational resources more effectively, leading to better performance.

3.2.4 Model Order Reduction for Stochastic Systems

Suppose we aim to solve (3.19). When the dimension n of the stochastic process $x(t)$ is exceptionally large, numerically solving the system (3.19) becomes computationally expensive. As previously mentioned, the core idea behind model reduction is to approximate this high-dimensional stochastic dynamical system with a lower-dimensional model that preserves its essential dynamic characteristics. In this section, we introduce several MOR techniques for stochastic systems.

3.2.4.1 Reduced order system

Similar to the linear deterministic system discussed in Section 3.1.3.1, we denote the reduced system by $G_{N,r}$, which is defined by:

$$\mathrm{d}x_r(t) = [A_r x_r(t) + B_r u(t)]\mathrm{d}t + \sum_{i=1}^{q} N_{i,r}\mathrm{d}W_i(t), \quad x_r(0) = x_{0,r} \tag{3.42a}$$

$$y_r(t) = C_r x_r(t) \quad t \geq 0 \tag{3.42b}$$

where $x_r(t) \in \mathbb{R}^r$ and $r \ll n$. We introduce a new variable $x_S(t) = Sx(t)$. This can be interpreted as a coordinate transform that is chosen in order to transform the original system using a suitable regular matrix $S \in \mathbb{R}^{n\times n}$. This transformation is the basis for the dimension reduction. Now, inserting $x_S(t) = Sx(t)$ into (3.19), we obtain

$$dx_S(t) = [A_S x_S(t) + B_S u(t)]\mathrm{d}t + \sum_{i=1}^{q} N_{S,i} x_S(t)\mathrm{d}W_i(t), \quad x_S(0) = x_{0,S}, \tag{3.43a}$$

$$y(t) = C_S x_S(t), \quad t \geq 0, \tag{3.43b}$$

where

$$\begin{aligned}
A_S &= SAS^{-1} = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}, && A_{11} \in \mathbb{R}^{r\times r} \\
N_{S,i} &= SN_iS^{-1} = \begin{pmatrix} N_{i,11} & N_{i,12} \\ N_{i,21} & N_{i,22} \end{pmatrix}, && N_{i,11} \in \mathbb{R}^{r\times r} \\
B_S &= SB = \begin{pmatrix} B_1 \\ B_2 \end{pmatrix}, && B_1 \in \mathbb{R}^{r\times m} \\
C_S &= CS^{-1} = \begin{pmatrix} C_1 & C_2 \end{pmatrix}, && C_1 \in \mathbb{R}^{p\times r} \\
x_S(t) &= Sx(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix}, && x_1(t) \in \mathbb{R}^{r}
\end{aligned}$$

A reduced system is derived by eliminating the equations associated with x_2 in (3.43). Subsequently, $x_2 \equiv 0$ within the equations about x_1, resulting in a reduced system (3.42)

with following components:

$$[x_r(t) \quad A_r \quad N_{i,r} \quad B_r \quad C_r] = [x_1(t) \quad A_{11} \quad N_{i,11} \quad B_1 \quad C_1]\,. \tag{3.44}$$

3.2.4.2 Petrov-Galerkin Projection in the Stochastic Framework

Referencing Section 3.1.3.2, the Petrov-Galerkin projection uses matrices $M, V \in \mathbb{R}^{n\times r}$, as defined in equations (3.12) and (3.13), respectively, with properties described in (3.14). In the context of the Petrov-Galerkin projection, the model transformation is expressed as:

$$\begin{bmatrix} A_r & N_{i,r} & B_r & C_r \end{bmatrix} = \begin{bmatrix} M^\top AV & M^\top N_i V & M^\top B & CV \end{bmatrix}. \tag{3.45}$$

To establish the stochastic relationship related to (3.15), the residual R is defined in the stochastic setting as:

$$R = V\mathrm{d}x_r(t) - AVx_r(t)\mathrm{d}t - Bu(t)\mathrm{d}t - \sum_{i=1}^{q} N_i x_r(t)\mathrm{d}W_i(t).$$

The Petrov-Galerkin projection requires that the residual R to be orthogonal to the subspace spanned by M. This condition, given by $M^\top R = 0$, results in the following equation for the reduced-order dynamics:

$$\begin{aligned} &M^\top(V\mathrm{d}x_r(t) - AVx_r(t)\mathrm{d}t - Bu(t)\mathrm{d}t - \sum_{i=1}^{q} N_i x_r(t)\mathrm{d}W_i(t)) \\ &\quad = \mathrm{d}x_r(t) - M^\top AVx_r(t)\mathrm{d}t - M^\top Bu(t)\mathrm{d}t - \sum_{i=1}^{q} M^\top N_i V x_r(t)\mathrm{d}W_i(t) = 0, \end{aligned}$$

ultimately yielding the compact form:

$$\mathrm{d}x_r(t) = [M^\top AVx_r(t) + M^\top Bu(t)]\mathrm{d}t + \sum_{i=1}^{q} M^\top N_i V x_r(t)\mathrm{d}W_i(t).$$

If $M = V$ has orthonormal columns, we obtain what is known as a Galerkin approximation, this is also mentioned in Section 3.1.3.2. This approximation technique will be defined in detail in the following section.

3.2.4.3 Reduced-order Model by Galerkin Projection

The following subsection is based on the framework and discussions presented in [99]. To simplify the analysis, we introduce the eigenvalue decomposition of the reachability matrix

$$P = S^\top \Sigma S,$$

where $S^{-1} = S^\top$ and $\Sigma = \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} = \mathrm{diag}\,(\lambda_1, \ldots, \lambda_n)$ is the matrix of eigenvalues of P. Suppose that the spectrum of P is ordered, that is, $\lambda_1 \geq \ldots \geq \lambda_n \geq 0$. This ordering facilitates defining the transformed state vector $x_S(t) = Sx(t)$. The fundamental solution

of the balanced realization (3.43), Φ_S, is given by

$$\Phi_S = S\Phi S^\top,$$

which can be derived by multiplying equation (3.24) with S from the left and with $S^\top$ from the right. Consequently, the reachability Gramian for equation (3.43) is expressed as

$$P_S := \mathbb{E}\int_0^\infty \Phi_S(s)B_S B_S^\top \Phi_S^\top(s)\mathrm{d}s = SPS^\top = \Sigma.$$

We partition x_S as

$$x_S = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix},$$

where x_1 and x_2 correspond to Σ_1 and Σ_2, respectively. Drawing from the insights in Section 3.2.3, it is evident that x_2 has a minimal impact on the system dynamics. We achieve the MOR by truncating the equations associated with x_2 in equation (3.43) and we set the remaining components of x_2 to zero. Using the Galerkin projection method discussed in Subsection 3.2.4.2 and applying equation (3.45), we derive the reduced system characterized by equation (3.42), with the associated matrices detailed as follows:

$$A_r = A_{11} = V^\top AV, \quad N_{i,r} = N_{i,11} = V^\top N_i V, \quad B_r = B_1 = V^\top B, \quad C_r = C_1 = CV,$$

where V represents the first r columns of $S^\top = [V \quad \star]$ (see equation (3.13)).

3.2.4.4 Stochastic Balanced Truncation

The balanced truncation method, when adapted for stochastic systems, marks a crucial step forward in simplifying complex models that are affected by randomness and uncertainty. This adaptation, developed in the late 1980s and 1990s, aimed to ensure that essential aspects of systems influenced by unpredictable disturbances could be captured more effectively. It did so by including stochastic Gramians, which consider how noise impacts the behavior of system dynamics. This change allows for a more accurate simplification of complicated stochastic models. A key benefit of using this method for stochastic systems is its ability to reduce models in a way that provides clear error bounds. This means that we can be confident that the simplified models still accurately reflect the original system's behavior in uncertain conditions. As a result, the balanced truncation method has become essential in fields like aerospace, climate modeling and financial engineering, where it is critical to simplify complex systems without losing important details. This evolution highlights the method's significant role in control theory and systems engineering today, showing how it helps create simpler, yet still reliable, models [41, 25]. More recently, for SDEs of Itô type, you can find relevant discussions in [10, 14].

We rewrite the algebraic Lyapunov equations (3.27) as follows:

$$AP + PA^\top + \sum_{i,j=1}^{q} N_i P N_j^\top k_{ij} = -BB^\top,$$
$$A^\top Q + QA + \sum_{i,j=1}^{q} N_i^\top Q N_j k_{ij} = -C^\top C.$$

As we mentioned before, by eliminating the dimensions with minimal observability or controllability, we can obtain a ROM approximation. This process is detailed in a five-step method described in Subsection 3.1.3.3. Within the new coordinate framework, represented as

$$[A_r \quad N_{r,i} \quad B_r \quad C_r] = [A_{11} \quad N_{i,11} \quad B_1 \quad C_1],$$

where

$$N_{S,i} = SN_iS^{-1} = \begin{pmatrix} N_{i,11} & N_{i,12} \\ N_{i,21} & N_{i,22} \end{pmatrix}, \quad N_{i,11} \in \mathbb{R}^{r\times r},$$

and A_{11}, B_1, and C_1 are defined as (3.16). The corresponding transformed Gramians, $P_S = SPS^\top$ and $Q_S = S^{-\top}QS^{-1}$, with the following property, are then computed.

$$P_S = Q_S = SPS^\top = S^{-\top}QS^{-1} = \Sigma$$

where

$$\Sigma = \begin{pmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_n \end{pmatrix},$$

here, $\sigma_i = \sqrt{\lambda_i(PQ)}$ are HSVs of the system, where they are organized in descending order, such that $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n \geq 0$. We finalize our discussion on MOR through the balanced truncation method, applied to linear stochastic systems (3.19), with the following theorem.

Theorem 3.21 [25] Consider the full system G_N (3.19) to be asymptotically stochastically stable, which implies

$$\sigma(\mathcal{L}_A + \Pi_N) \subset \mathbb{C}_-.$$

Assume $\Sigma = \operatorname{diag}(\Sigma_1, \Sigma_2) > 0$ is a block-diagonal matrix with $\Sigma_1 \in \mathbb{R}^{r\times r}$ and also $\sigma(\Sigma_1) \cap \sigma(\Sigma_2) = \emptyset$ and the conditions

$$\mathcal{L}_A(\Sigma) + \Pi_N(\Sigma) \leq 0, \quad \text{and} \quad \mathcal{L}_A^*(\Sigma) + \Pi_N^*(\Sigma) \leq 0,$$

are met. Then, for the reduced system $G_{N,r}$ (3.42), it is established that

$$\sigma(\mathcal{L}_{A_r} + \Pi_{N_r}) \subset \mathbb{C}_-,$$

where $\mathcal{L}_{A_r}$ defined in (3.17) and

$$\Pi_{N_r} := \sum_{i,j=1}^{q} N_{i,r} \otimes N_{j,r} k_{ij}.$$

This theorem highlights the preservation of stability when applying the balanced truncation method to an asymptotically stable system.

3.2.4.5 Stochastic POD Mehod

To our knowledge, the extension of the POD technique to SDEs influenced by Wiener processes has received limited attention. An application of the POD method is mentioned in [48] and [122], focusing exclusively on the stochastic Burgers equation. However, these studies do not tackle the empirical approximation of the nonlinear term and rely solely on low-order time integration methods. [116] extended POD techniques to stochastic Hamiltonian systems, emphasizing structure preservation for improved solution accuracy and stability, as demonstrated in experiments with the stochastic nonlinear Schrödinger equation.

Let Z_j be an $n \times N$ matrix representing empirical data on the system (3.19) for the jth realization of the Wiener process, as follows:

$$Z_j = \left[x\left(t_1, W^j\right), x\left(t_2, W^j\right), \ldots, x\left(t_N, W^j\right)\right], \quad \text{for} \quad W^j \in \Omega, \quad j = 1, \ldots, N_s,$$

where $N, N_s > 0$ are the number of time points and samples of the Wiener process, respectively. We want this method to be computationally inexpensive. Therefore, N_s should be small. We consider

$$Z := [x_0, Z_1, Z_2, \ldots, Z_{N_s}].$$

In fact, Z can be a collection of snapshots of a solution of this system for different realizations of the Wiener process $W(t)$. These snapshots are calculated for a particular set of initial conditions x_0 and control term $u(t)$. A low-rank approximation of Z can be done by performing the SVD of Z and truncating it after the first r largest singular values, that is,

$$Z = \hat{V}\hat{\Sigma}\hat{U}^\top = \begin{pmatrix} V & * \end{pmatrix} \begin{pmatrix} \Sigma & \\ & * \end{pmatrix} \begin{pmatrix} U^\top \\ * \end{pmatrix} \approx V\Sigma U^\top,$$

where $\hat{\Sigma} = \operatorname{diag}(\sigma_1, \sigma_2, \ldots)$ is the diagonal matrix of the singular values, $\hat{V}$ and $\hat{U}$ are orthogonal matrices, Σ is the diagonal matrix of the first r largest singular values and V and U are orthogonal matrices constructed by taking the first r columns of $\hat{V}$ and $\hat{U}$, respectively. Let $\hat{x}$ denote a vector in $\mathbb{R}^r$. Substituting $x = V\hat{x}$ in (3.19) yields a reduced SDE for $\hat{x}$ as

$$d\hat{x}(t) = [V^\top AV\hat{x}(t) + V^\top Bu(t)]\mathrm{d}t + \sum_{i=1}^{q} V^\top N_i V\hat{x}(t)\mathrm{d}W_i(t), \quad \hat{x}(0) = V^\top x_0, \tag{3.46a}$$

$$y(t) = CV\hat{x}(t), \qquad t \geq 0. \tag{3.46b}$$

If the singular values of Z exhibit rapid decay, a high-quality approximation of Z can be achieved for r, with $r \ll n$. Consequently, (3.46) serves as a low-dimensional approximation of (3.19), allowing for more efficient computational solutions. The solution approximation to (3.19) is reconstructed using $x(t) = V\hat{x}(t)$. This procedure, which involves deriving the matrix V from the empirical or simulation data set Z, is often described

as the offline phase in the context of model reduction. Although this stage can be computationally expensive, it is performed only once. On the other hand, computing solutions within the reduced-dimensional framework of (3.46) referred to the online stage of model reduction and is supposed to be faster and more efficient than solving the full system (3.19).

4 SPDEs Driven by Fractional Brownian Motion

In this chapter, we explore SPDEs driven by fBm with an additional control component. We employ an abstract evolution equation approach to represent these SPDEs. In recent years, significant developments have been made in the theory of SPDEs driven by fBm, particularly for $H \in \left(\frac{1}{2}, 1\right)$. In particular, SPDEs in a Hilbert space with infinite-dimensional fBm have been extensively studied, as referenced in [31, 32, 71]. To begin, we define C_0-semigroups to characterize mild solutions. Next, we introduce and validate the model of focus. The chapter continues with an example of a stochastic heat equation driven by fBm. In the following sections, we apply the Galerkin scheme to our example, similar to the one used in [29] for certain SPDEs with fractional noise. For SPDEs driven by fBm, references such as [53, 120] discuss the use of the spectral Galerkin method for spatial discretization when the Hurst parameter $H > \frac{1}{2}$.

4.1 Preliminary

4.1.1 C_0-semigroups

The concept of C_0-semigroups and their corresponding generators is crucial to introduce evolution equations. The following definitions and theorems are adapted from the book of Vrabie [119]. While Vrabie discusses C_0-semigroups on Banach spaces, our focus will be on the Hilbert space setting, which is sufficient for our purposes. Let U denote a separable Hilbert space (see Appendix A.2) and let $L(U)$ be the set of all linear bounded operators from U to U.

Definition 4.1 Afamily $\{S(t); t \geq 0\}$ of bounded linear operators $S(t) \in L(U)$ is called a C_0-semigroup on U if:

(i) $S(0) = I$, where I is the identity operator on U,

(ii) $S(t+s) = S(t)S(s)$ for all $t, s \geq 0$,

(iii) $\lim_{t \to 0^+} S(t)x = x$ for all $x \in U$.

Theorem 4.2 [119] Let $\{S(t); t \geq 0\}$ be a C_0-semigroup. Then there exist constants $\beta \in \mathbb{R}$ and $M \geq 1$ such that for all $t \geq 0$,

$$\|S(t)\|_{L(U)} \leq M e^{\beta t}. \tag{4.1}$$

Definition 4.3 The infinitesimal generator or simply the generator of the C_0-semigroup $\{S(t); t \geq 0\}$ is the operator $\mathcal{A} : D(\mathcal{A}) \subset U \to U$ defined by:

$$D(\mathcal{A}) = \left\{ x \in U \mid \lim_{t \to 0^+} \frac{1}{t}(S(t)x - x) \text{ exists} \right\},$$
$$\mathcal{A}x = \lim_{t \to 0^+} \frac{1}{t}(S(t)x - x), \quad x \in D(\mathcal{A}).$$

Equivalently, we say that $\mathcal{A}$ generates $\{S(t); t \geq 0\}$.

Remark 4.4

- The generator of a C_0-semigroup is a linear operator but not necessarily bounded.
- A C_0-semigroup is called a contraction semigroup if $M = 1$ and $\beta = 0$ in (4.1),i.e., for all $t \geq 0$, we have
$$\|S(t)\|_{L(U)} \leq 1.$$

Next, we state the basic properties of C_0-semigroups.

Theorem 4.5 [119] Let $\mathcal{A} : D(\mathcal{A}) \subset U \to U$ be the generator of a C_0-semigroup $\{S(t); t \geq 0\}$. Then:

(i) For all $x \in U$ and $t \geq 0$, we have
$$\lim_{h \to 0^+} \frac{1}{h} \int_t^{t+h} S(s)x\mathrm{d}s = S(t)x.$$

(ii) For all $x \in U$ and $t > 0$, we have
$$\int_0^t S(s)x\mathrm{d}s \in D(\mathcal{A}) \quad \text{and} \quad \mathcal{A}\left(\int_0^t S(s)x\mathrm{d}s\right) = S(t)x - x.$$

(iii) For all $x \in D(\mathcal{A})$ and $t \geq 0$, we have $S(t)x \in D(\mathcal{A})$. In addition, the mapping $t \mapsto S(t)x$ is continuously differentiable on $[0, +\infty)$ and satisfies
$$\frac{d}{dt}(S(t)x) = \mathcal{A}S(t)x = S(t)\mathcal{A}x.$$

(iv) For all $x \in D(\mathcal{A})$ and $0 \leq t_0 \leq t < \infty$, we have
$$\int_{t_0}^t \mathcal{A}S(s)x\mathrm{d}s = \int_{t_0}^t S(s)\mathcal{A}x\mathrm{d}s = S(t)x - S(t_0)x.$$

We summarize further important properties of the generator of a C_0-semigroup in the following theorems.

Theorem 4.6 [119] Let $\mathcal{A} : D(\mathcal{A}) \subset U \to U$ be the generator of a C_0-semigroup $\{S(t); t \geq 0\}$. Then $D(\mathcal{A})$ is dense in U and $\mathcal{A}$ is a closed operator[1].

[1]The graph of operator $\mathcal{A}$ is closed in $U \times U$

Theorem 4.7 [119] If $\mathcal{A} : D(\mathcal{A}) \subset U \to U$ is the generator of two C_0-semigroups $\{S(t); t \geq 0\}$ and $\{T(t); t \geq 0\}$, then $S(t) = T(t)$ for all $t \geq 0$.

4.1.2 Operator Spaces

Consider two separable Hilbert spaces denoted as $(U, \|\cdot\|_U, \langle\cdot,\cdot\rangle_U)$ and $(V, \|\cdot\|_V, \langle\cdot,\cdot\rangle_V)$. As it mentioned before, the space of all bounded linear operators from U to V is represented by $L(U;V)$. For simplicity, we abbreviate $L(U) = L(U;U)$. The set $L_1(V)$ represents the space of nuclear operators (see Appendix A.2.1), while $L_2(V,U)$ denotes the space of Hilbert-Schmidt operators from V to U (refer to Appendix A.2.2). The inner product for these operators, $\langle \mathcal{S}, \mathcal{T}\rangle_{HS}$, is given by $\langle \mathcal{S}, \mathcal{T}\rangle_{HS} = \sum_{k=1}^{\infty} \langle \mathcal{S}v_k, \mathcal{T}v_k\rangle_U$, where $\{v_k\}_{k=1}^{\infty}$ is an orthonormal basis in V. The corresponding norm is denoted by $\|\cdot\|_{HS}$.

Throughout this chapter, we assume that the operator Q meets the following condition.

Assumption 4.8 Let $Q \in L_1(V)$ be a self-adjoint, non-negative definite and bounded linear operator.

Define $V_0 := Q^{\frac{1}{2}}(V)$. We denote by $\mathcal{L}_2^0$ the Hilbert space of the Hilbert-Schmidt operators from V_0 to U, with the inner product defined as

$$\langle \Phi_1, \Phi_2\rangle_{\mathcal{L}_2^0} = \sum_{k=1}^{\infty} \left\langle \Phi_1 Q^{\frac{1}{2}} v_k, \Phi_2 Q^{\frac{1}{2}} v_k \right\rangle_U,$$

where $\{v_k\}_{k=1}^{\infty}$ is an orthonormal basis in V. The norm on this space is then defined as $\|\mathcal{T}\|_{\mathcal{L}_2^0} = \left\|\mathcal{T}Q^{\frac{1}{2}}\right\|_{HS}$.

In this chapter, we introduce the following condition on the operator $\mathcal{A}$:

Assumption 4.9 Let $\mathcal{A} : D(\mathcal{A}) \subset U \longrightarrow U$ be a linear, unbounded, self-adjoint and positive definite operator defined on a Hilbert space U, with a compact inverse.

Based on Assumption 4.9, the fractional power $\mathcal{A}^{-\alpha}$ can be expressed as:

$$\mathcal{A}^{-\alpha} = \frac{1}{\Gamma(\alpha)} \int_0^t t^{\alpha-1} e^{-\mathcal{A}t}\, \mathrm{d}t, \quad \alpha > 0,$$

where $e^{-\mathcal{A}t}$ is the analytic semigroup generated by $-\mathcal{A}$, given by

$$e^{-\mathcal{A}t} = \frac{1}{2\pi i} \int_\Gamma e^{-zt} R(z;\mathcal{A})\, \mathrm{d}z, \quad t \geq 0, \tag{4.2}$$

with Γ being a contour in the resolvent set $\rho(-\mathcal{A})$, such that $\arg z \to \pm\theta$ as $|z| \to \infty$, for some $\theta \in \left(\frac{\pi}{2}, \pi\right)$. Here, $R(z;\mathcal{A}) = (zI - \mathcal{A})^{-1}$ is the resolvent of $\mathcal{A}$.

Under Assumption 4.9, $\mathcal{A}^{-\alpha}$ has a unique inverse denoted by $\mathcal{A}^{\alpha}$ and the domain $\dot{U}^{\alpha} := \mathrm{dom}(\mathcal{A}^{\alpha})$ is a Banach space equipped with the norm

$$\|u\|_{\dot{U}^{\alpha}}^2 := \left\|\mathcal{A}^{\frac{\alpha}{2}} u\right\|_U^2 = \sum_{k=1}^{\infty} \lambda_k^{\alpha} \langle u, u_k\rangle_U^2,$$

where $\alpha \in \mathbb{R}$, $u \in \dot{U}^{\alpha}$ and $\{u_k\}_{k\in\mathbb{N}}$ is an orthonormal basis of operator $\mathcal{A}$ in the separable

Hilbert space U, consisting of its eigenvectors, satisfying:

$$\mathcal{A}u_k = \lambda_k u_k, \tag{4.3}$$

where $0 < \lambda_1 \leq \lambda_2 \leq \ldots...$ Specifically, for $\alpha < 0$, $\dot{U}^\alpha$ is defined as the closure of U with the norm $\|\cdot\|_{\dot{U}^\alpha}$. For $\alpha > 0$, the spaces $\dot{U}^\alpha$ are Banach spaces and $\dot{U}^{-\alpha}$ is isometrically isomorphic to the dual of $\dot{U}^\alpha$. The spaces $\dot{U}^\alpha \subset U \subset \dot{U}^{-\alpha}$ form a Gelfand triple when $\alpha > 0$.

Introducing the fractional order space, we denote by $\mathcal{L}^0_{2,\gamma}$, with $\gamma \geq 0$, the Hilbert space of Hilbert-Schmidt operators from V_0 to $\dot{U}^\gamma$, with the inner product

$$\langle \Phi_1, \Phi_2 \rangle_{\mathcal{L}^0_{2,\gamma}} = \sum_{k=1}^{\infty} \left\langle \Phi_1 Q^{\frac{1}{2}} v_k, \Phi_2 Q^{\frac{1}{2}} v_k \right\rangle_{\dot{U}^\gamma},$$

where $\{v_k\}_{k=1}^\infty$ is an orthonormal basis in V. Furthermore, the norm of $\mathcal{T}$ in $\mathcal{L}^0_{2,\gamma}$ is given by

$$\|\mathcal{T}\|_{\mathcal{L}^0_{2,\gamma}} = \left\| \mathcal{A}^{\frac{\gamma}{2}} \mathcal{T} Q^{\frac{1}{2}} \right\|_{HS}.$$

In particular, when $\gamma = 0$, $\mathcal{L}^0_{2,\gamma}$ reduces to $\mathcal{L}^0_2$.

4.2 Evolution Equations with fBm

In this section, we address an infinite-dimensional system characterized by the noise process W^H. The process W^H takes values in a separable Hilbert space V and is defined on a complete right-continuous filtered probability space $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P}\right)$. Furthermore, we assume that W^H is a fBm process with respect to the filtration $(\mathcal{F}_t)_{t\geq 0}$. The key properties of this process and the definition of an integral with respect to W^H are discussed in Chapter 2. In this context, we consider Young integration with respect to fBm (refer to Definition 2.13).

Let Q satisfies Assumption 4.8, then the V-valued fBm W^H is defined by the formal sum:

$$W^H(t) = \sum_{i=1}^{\infty} \sqrt{\mu_i} v_i \omega_i^H(t). \tag{4.4}$$

where $\{\omega_i^H\}_{i\in\mathbb{N}}$ is a sequence of stochastically independent scalar fBms with the same Hurst parameter $H > \frac{1}{2}$, $\{v_i\}_{i\in\mathbb{N}}$ is an orthonormal basis of V and additionally $\mu_1 \geq \mu_2 \geq \cdots \geq 0$ is a bounded sequence of nonnegative numbers such that $Qv_i = \mu_i v_i$ with $\lim_{n\to\infty} \mu_n = 0$ and $\sum_{i=1}^\infty \mu_i < \infty$.

Let $G \in \mathcal{L}^0_2$, then the stochastic integral with respect to Q-fBm W^H can be defined as

$$\begin{aligned} \int_0^T G(s) \mathrm{d}W^H(s) &= \sum_{i=1}^{\infty} \int_0^T G(s) Q^{1/2} v_i d\omega_i^H(s) \\ &= \sum_{i=1}^{\infty} \sqrt{\mu_i} \int_0^T G(s) v_i d\omega_i^H(s), \end{aligned} \tag{4.5}$$

where the convergence of the sums is understood in the mean square sense in V. Throughout the remainder of this chapter, we will use C, K and L to represent generic positive constants, which may vary in value with each appearance.

4.2.1 Existence and Uniqueness of Mild Solution

In this section, our main aim is to establish the existence and uniqueness of mild solutions to the following stochastic evolution equation with multiplicative fBms:

$$\begin{aligned} d\mathcal{X}(t) &= \left(\tilde{\mathcal{A}}\mathcal{X}(t) + \mathcal{B}u(t)\right) \mathrm{d}t + \mathcal{N}(\mathcal{X}(t)) \circ \mathrm{d}W^H(t), \quad t \in [0,T], \\ \mathcal{X}(0) &= X_0 \in U. \end{aligned} \tag{4.6}$$

Where $W^H(t)$ denotes the fBm with the covariance operator Q, as specified in Assumption 4.8 and the operator $\tilde{\mathcal{A}} = -\mathcal{A}$, where the operator $\mathcal{A}$ satisfies Assumption 4.9.

The process $u : \mathbb{R}_+ \times \Omega \to \mathbb{R}^m$ is adapted to $(\mathcal{F}_t)_{t\geq 0}$, satisfying

$$\sup_{t\in[0,T]} \mathbb{E}\|u(t)\|_2^2 \leq C \tag{4.7}$$

for every $T > 0$, where $\|\cdot\|_2$ represents the Euclidean norm in $\mathbb{R}^m$. We also impose the following assumptions on the initial value X_0 the operator $\mathcal{B}$ and the operator $\mathcal{N}$.

Assumption 4.10 Let $\beta \in (0,1]$ and $\gamma \in [0,\beta)$. We assume that the initial value X_0 is $\mathcal{F}_0$-measurable and takes values in $\dot{U}^\gamma$, with $X_0 \in L^2(\Omega;\dot{U}^\gamma)$. This implies the existence of a constant $K > 0$ such that

$$\|X_0\|_{L^2(\Omega;\dot{U}^\gamma)} \leq K.$$

Assumption 4.11 Let $\beta \in (0,1]$ and $\gamma \in [0,\beta)$. Suppose that $\mathcal{B}$ is a bounded linear operator on $\mathbb{R}^m$ with values in $\dot{U}^{\gamma+\beta-1}$ and satisfies

$$\|\mathcal{B}\|_{L(\mathbb{R}^m,\dot{U}^{\gamma+\beta-1})} < C.$$

Assumption 4.12 Let $\beta \in (0,1]$ and $\gamma \in [0,\beta)$. Consider a mapping $\mathcal{N} : \dot{U}^\gamma \to \mathcal{L}_2^0$ which is measurable and satisfies $\mathcal{N}(\dot{U}^\gamma) \subset \mathcal{L}_{2,\gamma+\beta-1}^0$. Assume that there exists a constant $L > 0$ such that the following conditions are met:

(i) $\left\|\mathcal{A}^{\frac{\gamma+\beta-1}{2}}\mathcal{N}(u)\right\|_{\mathcal{L}_2^0} \leq L\left(1 + \|u\|_{\dot{U}^\gamma}\right), \quad \text{for } u \in \dot{U}^\gamma,$

(ii) $\left\|\mathcal{A}^{\frac{\gamma+\beta-1}{2}}\left(\mathcal{N}(u) - \mathcal{N}(v)\right)\right\|_{\mathcal{L}_2^0} \leq L\|u-v\|_{\dot{U}^\gamma}, \quad \text{for } u,v \in \dot{U}^\gamma,$

(iii) $\left\|D_s^H\mathcal{A}^{\frac{\gamma+\beta-1}{2}}\mathcal{N}(u)\right\|_{\mathcal{L}_2^0} \leq L\left(1 + \|u\|_{\dot{U}^\gamma}\right), \quad \text{for } u \in \dot{U}^\gamma,$

(iv) $\left\|D_s^H\mathcal{A}^{\frac{\gamma+\beta-1}{2}}\left(\mathcal{N}(u) - \mathcal{N}(v)\right)\right\|_{\mathcal{L}_2^0} \leq L\|u-v\|_{\dot{U}^\gamma}, \quad \text{for } u,v \in \dot{U}^\gamma.$

Here, D_s^H denotes the Malliavin derivative defined in Section 2.5.3 in Chapter 2.

It is established that, under Assumption 4.9, the following initial value problem

$$d\mathcal{X}(t) + \mathcal{A}\mathcal{X}(t)\,\mathrm{d}t = 0, \quad \mathcal{X}(0) = X_0, \quad t \in [0, T],$$

has a unique solution given by $\mathcal{X}(t) = S(t)X_0$, where $S(t) = e^{-\mathcal{A}t}$ for $t \geq 0$, as defined in (4.2). It is widely recognized that the semigroup $S(t)$ exhibits the following properties:

$$\|\mathcal{A}^s S(t)\|_{L(U)} \leq Ct^{-s}, \quad t, s \geq 0, \tag{4.8}$$
$$\|\mathcal{A}^s \dot{S}(t)\|_{L(U)} \leq Ct^{-s-1}, \quad t, s \geq 0, \tag{4.9}$$
$$\|\mathcal{A}^{-\rho}(I - S(t))\|_{L(U)} \leq Ct^{\rho}, \quad t \geq 0,\ \rho \in [0, 1]. \tag{4.10}$$

These properties are crucial for the regularity analysis of the solution.

Remark 4.13 In this chapter, we refer to the results in [29] to derive the following outcome. As mentioned earlier, in our case, we consider the fractional Brownian motion W^H and the solution $\mathcal{X}(t)$ to be defined on different Hilbert spaces V and U, respectively, while in [29] the authors treated both on the same Hilbert space.

Now we give the definition of mild solution to equation (4.6).

Definition 4.14 Let $\{\mathcal{X}(t)\}_{t\in[0,T]}$ be a predictable U-valued stochastic process. We refer to $\{\mathcal{X}(t)\}_{t\in[0,T]}$ as a mild solution of the equation (4.6) if the operator

$$\Psi_{X_0}(\mathcal{X})(t) := S(t)X_0 + \int_0^t S(t-s)\mathcal{B}u(s)\mathrm{d}s + \int_0^t S(t-s)\mathcal{N}(\mathcal{X}(s)) \circ \mathrm{d}W^H(s), \tag{4.11}$$

is well-defined and for almost every $t \in [0, T]$, it satisfies $\Psi_{X_0}(\mathcal{X})(t) = \mathcal{X}(t)$ almost surely. But we will see that the mild solution has more regularity with the Assumption 4.12 (see Theorem 4.17).

To demonstrate that the concept of a mild solution is applicable in $L^2(\Omega; U)$, it is necessary to verify that the stochastic convolution

$$\mathcal{R}_t := \int_0^t S(t-s)\mathcal{N}(\mathcal{X}(s)) \circ \mathrm{d}W^H(s), \quad t \in [0, T],$$

is well-defined within $L^2(\Omega; U)$. According to Theorem 2.21 presented in Chapter 2, we can express

$$\mathcal{R}_t = \int_0^t S(t-s)\mathcal{N}(\mathcal{X}(s))\delta W^H(s) + \int_0^t \int_0^s D_s^H(S(t-\tau)\mathcal{N}(\mathcal{X}(\tau)))\phi(s,\tau)\,d\tau\mathrm{d}s, \tag{4.12}$$

where $\phi(s, \tau) = H(2H-1)|s-\tau|^{2H-2}$ and the first term on the right-hand side represents the fWIS integral, as defined in Section 2.5.4 of Chapter 2.

In [29], the author consider the following Wiener integral representation of W^H:

$$W^H(t) = \int_0^t K_H(t, s)\mathrm{d}W(s)$$

where the kernel function $K_H(t,s)$ is defined in Proposition 2.6 as follows

$$K_H(t,s) = \begin{cases} c_H \int_s^t \left(\frac{u}{s}\right)^{H-\frac{1}{2}} (u-s)^{H-\frac{3}{2}} \mathrm{d}u, & \text{if } H > \frac{1}{2} \\ b_H \left[(\frac{t}{s})^{H-\frac{1}{2}} (t-s)^{H-\frac{1}{2}} - (H-\frac{1}{2}) s^{\frac{1}{2}-H} \int_s^t u^{H-\frac{3}{2}} (u-s)^{H-\frac{1}{2}} \mathrm{d}u \right] & \text{if } H < \frac{1}{2} \end{cases} \tag{4.13}$$

where b_H and c_H are constants dependent on the Hurst parameter H. By employing this kernel function, the integral with respect to fBm is defined as

$$\int_0^t f(s) \mathrm{d}W^H(s) = \int_0^t K^*_{H,t}(f)(s) \mathrm{d}W(s),$$

which indicates that, instead of directly integrating with respect to fBm, we can express the integral as the Itô integral (see Section 1.3 for more details). In particular, when $\frac{1}{2} < H < 1$, the operator $K^*_{H,t}$ can be expressed as

$$K^*_{H,t}(f)(s) = \int_s^t f(\tau) \frac{\partial K_H(\tau, s)}{\partial \tau} d\tau,$$

where, according to (4.13),

$$\frac{\partial K_H(\tau,s)}{\partial \tau} = c_H \left(\frac{\tau}{s}\right)^{H-\frac{1}{2}} (\tau - s)^{H-\frac{3}{2}}.$$

In the literature, it has been established that $K^*_{H,t}$ is an isometry from ε to $L^2([0,T])$, where ε represents the space of step functions, denoted by $\phi_t = \sum_{j=1}^m a_j 1_{(t_{j-1}, t_j]}(t)$ (for more details, see Section 2.5.1).

As stated in Proposition 6.12 of [44], the first integral on the right-hand side of (4.12) can be expressed as

$$\int_0^t K^*_{H,t}(S(t-s)\mathcal{N}(\mathcal{X}(s))) \mathrm{d}W_s.$$

In other words, by applying equation (4.12) alongside the above expression, we can rewrite

$$\mathcal{R}_t = \int_0^t K^*_{H,t}(S(t-s)\mathcal{N}(\mathcal{X}(s))) \mathrm{d}W_s + \int_0^t \int_0^t D^H_s(S(t-\tau)\mathcal{N}(\mathcal{X}(\tau)))\phi(s,\tau)\, d\tau \mathrm{d}s. \tag{4.14}$$

To demonstrate that $\mathcal{R}_t$ is well-defined, we need to introduce the following function space.

For $p \geq 2$ and $\gamma \geq 0$, let $\mathcal{O}^2_\gamma$ denote the space of all $\dot{U}^\gamma$-valued predictable processes $\{\mathcal{X}(t) : t \in [0,T]\}$ satisfying

$$\|\mathcal{X}\|_{\mathcal{O}^2_\gamma} := \sup_{t\in[0,T]} \|\mathcal{X}(t)\|_{L^2(\Omega;\dot{U}^\gamma)} = \sup_{t\in[0,T]} \left(\mathbb{E}\|\mathcal{X}(t)\|^2_{\dot{U}^\gamma}\right)^{\frac{1}{2}} < \infty.$$

For the special case where $\gamma = 0$, the space $\mathcal{O}^2_0$ becomes a Banach space equipped with the norm

$$\|\mathcal{X}\|_{\mathcal{O}^2_0} := \sup_{t\in[0,T]} \left(\mathbb{E}\|\mathcal{X}(t)\|^2_U\right)^{\frac{1}{2}} < \infty.$$

In this scenario, $\mathcal{O}^2_0$ is simply denoted by $\mathcal{O}^2$. However, when $\gamma > 0$, the space

$\left(\mathcal{O}_\gamma^2, \|\cdot\|_{\mathcal{O}_\gamma^2}\right)$ is not a Banach space. However, it is straightforward to verify that for any $M \in (0,\infty)$, the subset

$$\mathcal{O}_\gamma^2(M) = \left\{x \in \mathcal{O}_\gamma^2 : \|\mathcal{X}\|_{\mathcal{O}_\gamma^2} \leq M\right\}$$

equipped with the norm $\|\cdot\|_{\mathcal{O}^2}$ forms a complete metric space.

Lemma 4.15 [29] Assume that Assumptions 4.8, 4.9, (i) and (iii) from Assumption 4.12 are satisfied. Let $\beta \in (0,1]$, $\gamma \in [0,\beta)$, $\max\{0, 3-4H\} < \beta - \gamma$ and consider $\mathcal{X}(t)$ as a predictable process such that $\mathcal{X}(t) \in \mathcal{O}_\gamma^2$ for any $t \in [0,T]$. Then, there exists a constant $C > 0$ such that

$$\|\mathcal{R}_t\|_{L^2(\Omega;\dot{U}^\gamma)} \leq C\left(1 + \sup_{s\in[0,T]} \|\mathcal{X}(s)\|_{L^2(\Omega;\dot{U}^\gamma)}\right) t^{\frac{4H+\beta-3}{2}}, \quad t \in [0,T].$$

Proof. The proof follows the same steps as the proof of Lemma 3.2 in [29], as follow

$$\begin{aligned}
\|\mathcal{R}_t\|_{L^2(\Omega;\dot{U}^\gamma)} &\leq \left\|\int_0^t K_{H,t}^*(S(t-s)\mathcal{N}(\mathcal{X}(s)))\mathrm{d}W_s\right\|_{L^2(\Omega;\dot{U}^\gamma)} \\
&\quad + \left\|\int_0^t \int_0^t D_s^H(S(t-\tau)\mathcal{N}(\mathcal{X}(\tau)))\phi(s,\tau)\mathrm{d}s\,d\tau\right\|_{L^2(\Omega;\dot{U}^\gamma)} \\
&\leq C\left\|\left(\int_0^t \left\|K_{H,t}^*\mathcal{A}^{\frac{\gamma}{2}}S(t-s)\mathcal{N}(\mathcal{X}(s))Q^{\frac{1}{2}}\right\|_{\mathrm{HS}}^2 \mathrm{d}s\right)^{\frac{1}{2}}\right\|_{L^2(\Omega;\mathbb{R})} \\
&\quad + C\left\|\left(\int_0^t \int_0^t \left\|D_s^H\mathcal{A}^{\frac{\gamma}{2}}S(t-\tau)\mathcal{N}(\mathcal{X}(\tau))\right\|_{\mathcal{L}_2^0}^2 \mathrm{d}s\,d\tau\right)^{\frac{1}{2}}\right\|_{L^2(\Omega;\mathbb{R})} \\
&\quad + \left\|\left(\int_0^t \int_0^t \left\|D_s^H\mathcal{A}^{\frac{\gamma}{2}}S(t-\tau)\mathcal{N}(\mathcal{X}(\tau))\right\|_{\mathcal{L}_2^0}^2 |\phi(s,\tau)|^2\mathrm{d}s\,d\tau\right)^{\frac{1}{2}}\right\|_{L^2(\Omega;\mathbb{R})} \\
&\leq C\left(1 + \sup_{s\in[0,T]} \|\mathcal{X}(s)\|_{L^2(\Omega;\dot{U}^\gamma)}\right)\left(\int_0^t \left\|\mathcal{K}_{H,t}^*\mathcal{A}^{\frac{-\beta+1}{2}}S(t-s)\right\|_U^2 \mathrm{d}s\right)^{\frac{1}{2}} \\
&\quad + C\left(1 + \sup_{s\in[0,T]} \|\mathcal{X}(s)\|_{L^2(\Omega;\dot{U}^\gamma)}\right) t \\
&\quad + C\left(1 + \sup_{s\in[0,T]} \|\mathcal{X}(s)\|_{L^2(\Omega;\dot{U}^\gamma)}\right)\left(\int_0^t \int_0^t \left\|\mathcal{A}^{\frac{-\beta+1}{2}}S(t-\tau)\right\|_U^2 |\phi(s,\tau)|^2\mathrm{d}s\,d\tau\right)^{\frac{1}{2}}.
\end{aligned}$$

So we can write

$$\|\mathcal{R}_t\|_{L^2(\Omega;\dot{U}^\gamma)} \leq C(1 + \sup_{s\in[0,T]} \|\mathcal{X}(s)\|_{L^2(\Omega;\dot{U}^\gamma)})\,(I_1 + t + I_2)\,. \tag{4.15}$$

According to Lemma 3.2 from [29], for I we have

$$I_1 \leq Ct^{\frac{\beta+2H-1}{2}} \tag{4.16}$$

Now, for I_2 we have

$$\begin{aligned} I_2^2 &\leq CH^2(2H-1)^2 \int_0^t \int_0^t (t-\tau)^{\beta-1}|s-\tau|^{4H-4}\mathrm{d}s\, d\tau \\ &= 2CH^2(2H-1)^2 \int_0^t (t-\tau)^{\beta-1} \int_0^\tau (\tau-s)^{4H-4}\mathrm{d}s\, d\tau \\ &= \frac{2CH^2(2H-1)^2}{4H-3} \int_0^t (t-\tau)^{\beta-1}\tau^{4H-3}\, d\tau \end{aligned}$$

By substituting $u = \frac{\tau}{t}$, applying the Beta function definition and using the relation $B(x,y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$, we obtain

$$\begin{aligned} I_2^2 &\leq \frac{2CH^2(2H-1)^2}{4H-3} t^{4H+\beta-3} \int_0^1 (1-u)^{\beta-1}u^{4H-3}\mathrm{d}u \\ &= \frac{2CH^2(2H-1)^2}{4H-3} B(4H-2,\beta)t^{4H+\beta-3} \\ &= \frac{2CH^2(2H-1)^2\Gamma(4H-2)\Gamma(\beta)}{(4H-3)\Gamma(4H+\beta-2)} t^{4H+\beta-3}, \end{aligned} \tag{4.17}$$

where the integral is finite, since $\beta \in (3-4H, 1]$. Combining (4.15), (4.16) and (4.17), we can obtain that there exists a constant $C > 0$ such that

$$\|\mathcal{R}_t\|_{L^2(\Omega;\dot{U}^\gamma)} \leq C\left(1 + \sup_{s\in[0,T]} \|\mathcal{X}(s)\|_{L^2(\Omega;\dot{U}^\gamma)}\right) t^{\frac{4H+\beta-3}{2}}, \quad t \in [0,T].$$

□

Next, we present a lemma that serves as a crucial component in establishing the global existence and uniqueness of the mild solution for equation (4.6).

Lemma 4.16 [29] Let $0 < \zeta \leq 1$, $0 < \alpha < 1$, $\delta > -\alpha_i$ and let b, c and d be non-negative constants. Additionally, let $a : [0,T] \to \mathbb{R}^+$ be a non-decreasing, bounded function and let $x(t)$ be a non-negative, bounded function on $[0,T)$ satisfying the inequality

$$x(t) \leq a(t) + b\int_0^t (t-s)^{\zeta-1}x(s)\,\mathrm{d}s + c\int_0^t (t-s)^{\alpha-1}s^\delta x(s)\,\mathrm{d}s.$$

Under these conditions, there exists a constant $\mu > 0$ such that

$$x(t) \leq 2e^{\theta\mu t}a(t), \quad t \in [0,T),$$

where $0 < \theta < \min\{\alpha, \zeta\}$ and $\alpha + \theta < 1$.

We are now ready to present the existence and uniqueness theorem, following the approach in [29].

Theorem 4.17 [29] Assume that Assumptions 4.8, 4.9, 4.10, 4.11 and 4.12 are satisfied. Let $\beta \in (0,1]$, $\gamma \in [0,\beta)$ and $\max\{0, 3-4H\} < \beta - \gamma$. Then, equation (4.6) has a unique

mild solution and there exists a constant $C > 0$ such that

$$\|\mathcal{X}(t)\|_{\mathcal{O}^2_\gamma} = \sup_{t\in[0,T]} \left(\mathbf{E}\|\mathcal{X}(t)\|^2_{\dot{U}^\gamma}\right)^{\frac{1}{2}} \leq C. \tag{4.18}$$

Proof. In [29], the author employed a standard technique to demonstrate that the operator Ψ_{X_0}, as defined in equation (4.11), maps $\mathcal{O}^2_\gamma(M)$ into itself. Furthermore, it was shown that there exists a sufficiently small time τ such that Ψ_{X_0} acts as a contraction mapping under the norm $\|\cdot\|_{\mathcal{O}^2}$, where $M \in (0,\infty)$. This result implies the existence of a finite time $\tau \in (0,T)$ for which the equation (4.6) admits a unique local mild solution for $t \in [0,\tau)$, which is a predictable process satisfying

$$\sup_{t\in[0,\tau)} \left(\mathbf{E}\|\mathcal{X}(t)\|^2_{\dot{U}^\gamma}\right)^{\frac{1}{2}} \leq M. \tag{4.19}$$

Following the approach in [29], we now establish that the uniform estimate (4.18) guarantees the existence of a unique global mild solution to equation (4.6) for $t \in [0,T]$.

Let $t \in [0,\tau)$. Applying equation (4.12) and Assumption 4.10 and utilizing Hölder inequality, we obtain

$$\begin{aligned}
\|\mathcal{X}(t)\|_{L^2(\Omega;\dot{U}^\gamma)} \leq & CK + \left\|\int_0^t S(t-s)\mathcal{B}u(s)\,ds\right\|_{L^2(\Omega;\dot{U}^\gamma)} \\
& + \left\|\int_0^t S(t-\tau)\mathcal{N}(\mathcal{X}(\tau))\delta W^H(\tau)\right\|_{L^2(\Omega;\dot{U}^\gamma)} \\
& + \left\|\int_0^t D^H_s S(t-\tau)\mathcal{N}(\mathcal{X}(\tau))\phi(s,\tau)\mathrm{d}s\,d\tau\right\|_{L^2(\Omega;\dot{U}^\gamma)} \\
\leq & CK + J + I_1 + I_2.
\end{aligned} \tag{4.20}$$

Utilizing equation (4.8), along with Assumption 4.11, we obtain the following result:

$$\begin{aligned}
\left\|\int_0^t S(t-s)\mathcal{B}u(s)\,ds\right\|^2_{L^2(\Omega;\dot{U}^\gamma)} &\leq \left\|\int_0^t \left\|\mathcal{A}^{\frac{\gamma}{2}}S(t-s)\mathcal{B}\right\|^2_{\mathcal{L}^0_2}\|u(s)\|^2_2\mathrm{d}s\right\|_{L(\Omega;\mathbb{R})} \\
&\leq \|\mathcal{B}\|^2_{L(\mathbb{R}^m;\dot{U}^{\gamma+\beta-1})}\left\|\int_0^t \|\mathcal{A}^{\frac{1-\beta}{2}}S(t-s)\|^2_U\|u(s)\|^2_2\mathrm{d}s\right\|_{L(\Omega;\mathbb{R})} \\
&\leq C\|\mathcal{B}\|^2_{L(\mathbb{R}^m;\dot{U}^{\gamma+\beta-1})}\int_0^t (t-s)^{\beta-1}\mathbb{E}\|u(s)\|^2_2\mathrm{d}s \\
&\leq C\|\mathcal{B}\|^2_{L(\mathbb{R}^m;\dot{U}^{\gamma+\beta-1})}\sup_{s\in[0,T]}\mathbb{E}\|u(s)\|^2_2 t^\beta \leq CT^\beta
\end{aligned}$$

which is derived from the relation (4.7) for the control term $u(t)$. So it yields

$$J \leq CT^{\frac{\beta}{2}}. \tag{4.21}$$

Based on the proof of Theorem 3.5 in [29], the following inequality holds for I_1:

$$\begin{aligned} I_1 \leq & C\left(\int_0^t s^{1-2H}(t-s)^{2H+\beta-2}\left(1+\|\mathcal{X}(s)\|_{L^2(\Omega;\dot{U}^\gamma)}\right)^2 \mathrm{d}s\right)^{\frac{1}{2}} \\ & + C\left(\int_0^t (t-s)^{\beta-1}\left(1+\|\mathcal{X}(s)\|_{L^2(\Omega;\dot{U}^\gamma)}\right)^2 \mathrm{d}s\right)^{\frac{1}{2}}. \end{aligned} \tag{4.22}$$

According to Assumption 4.12 (iii) and by following the reasoning outlined in the proof of Lemma 4.15, we obtain

$$\begin{aligned} I_2 &\leq C\left(\int_0^t \int_0^t \left\|\mathcal{A}^{\frac{-\beta+1}{2}} S(t-\tau)\right\|_U^2 |\phi(s,\tau)|^2 \left(1+\|\mathcal{X}(s)\|_{L^2(\Omega;\dot{U}^\gamma)}\right)^2 ds\, d\tau\right)^{\frac{1}{2}} \\ &\leq C\left(\int_0^t (t-s)^{\beta-1} s^{4H-3}(1+\|\mathcal{X}(s)\|_{L^2(\Omega;\dot{U}^\gamma)})^2 \mathrm{d}s\right)^{\frac{1}{2}}. \end{aligned} \tag{4.23}$$

By substituting (4.21), (4.22) and (4.23) into (4.20) and applying both the Cauchy and Hölder inequalities, the following result is derived

$$\begin{aligned} \|\mathcal{X}(t)\|^2_{L^2(\Omega;\dot{U}^\gamma)} \leq & CT^\beta + C\int_0^t (t-s)^{\beta-1}\|\mathcal{X}(s)\|^2_{L^2(\Omega;\dot{U}^\gamma)}\mathrm{d}s \\ & + C\int_0^t s^{1-2H}(t-s)^{2H+\beta-2}\|\mathcal{X}(s)\|^2_{L^2(\Omega;\dot{U}^\gamma)}\mathrm{d}s \\ & + C\int_0^t s^{4H-3}(t-s)^{\beta-1}\|\mathcal{X}(s)\|^2_{L^2(\Omega;\dot{U}^\gamma)}\mathrm{d}s. \end{aligned} \tag{4.24}$$

By utilizing (4.19) in conjunction with Lemma 4.16 and incorporating it into (4.24), we derive that

$$\|\mathcal{X}(t)\|_{L^2(\Omega;\dot{U}^\gamma)} \leq C,$$

for $t \in [0,\tau]$, the same estimate holds for $t \in [\tau, 2\tau]$ and so on. So, the proof is completed. □

4.3 Spectral Galerkin Method

In this section, we apply a spectral Galerkin method for the spatial discretization of equation (4.6). Let $\tilde{\mathcal{A}} = (-\mathcal{A})$ with $\mathcal{A} : D(\mathcal{A}) \to U$ be a densely defined, linear operator that is self-adjoint and positive definite. The operator $\mathcal{A}$ has an orthonormal basis $\{u_k\}_{k\in\mathbb{N}}$ in the separable Hilbert space U, consisting of its eigenvectors, satisfying:

$$\mathcal{A}u_k = \lambda_k u_k, \tag{4.25}$$

where $0 \leq \lambda_1 \leq \lambda_2 \leq \ldots$ and $\lim_{n\to\infty} \lambda_n = \infty$. The operator $\tilde{\mathcal{A}}$ generates a contraction C_0-semigroup $\{S(t); t \geq 0\}$ defined by

$$S(t)x = \sum_{k=1}^{\infty} \mathrm{e}^{-\lambda_k t}\,(x, u_k)\,u_k, \tag{4.26}$$

for $x \in U$. The system is exponentially stable ($\beta < 0$ in (4.1)) when $\lambda_1 > 0$. For a given $n \in \mathbb{N}$, We introduce the finite-dimensional subspaces $U_n \subset U$ and $V_n \subset V$, defined as follows:

$$U_n = \operatorname{span}\{u_1, u_2, \ldots, u_n\}, \quad V_n = \operatorname{span}\{v_1, v_2, \ldots, v_n\}.$$

Therefore, we introduce projection operators $\mathcal{P}_n : U \to U_n$ and $\tilde{\mathcal{P}}_n : V \to V_n$ given by:

$$\mathcal{P}_n v = \sum_{k=1}^{n} \langle u, u_k \rangle_U u_k, \quad u \in U,$$
$$\tilde{\mathcal{P}}_n v = \sum_{k=1}^{n} \langle v, v_k \rangle_V v_k, \quad v \in V.$$

This implies that

$$\|\mathcal{P}_n u\|_U^2 = \left\| \sum_{k=1}^{n} \langle u, u_k \rangle_U u_k \right\|_U^2 = \sum_{k=1}^{n} |\langle u, u_k \rangle_U|^2 \leq \sum_{k=1}^{\infty} |\langle u, u_k \rangle_U|^2 = \|u\|_U^2, \quad u \in U,$$
$$\left\| \tilde{\mathcal{P}}_n v \right\|_V^2 = \left\| \sum_{k=1}^{n} \langle v, v_k \rangle_V v_k \right\|_V^2 = \sum_{k=1}^{n} |\langle v, v_k \rangle_V|^2 \leq \sum_{k=1}^{\infty} |\langle v, v_k \rangle_V|^2 = \|v\|_V^2, \quad v \in V.$$

We define the operator $\mathcal{A}_n : U \to U_n$ as $\mathcal{A}_n = \mathcal{A}\mathcal{P}_n$, so that $\tilde{\mathcal{A}}_n = -\mathcal{A}_n$. Since $\tilde{\mathcal{A}}_n$ is bounded for each $n \in \mathbb{N}$, it is known that $\tilde{\mathcal{A}}_n$ generates a C_0-semigroup on U_n, denoted by $S_n(t) = S(t)\mathcal{P}_n = e^{\tilde{\mathcal{A}}_n t}$ for $t \in [0, T]$. For any $x \in U_n$, we have the representation $S_n(t)x = \sum_{k=1}^{n} \mathrm{e}^{-\lambda_k t} \langle x, u_k \rangle_U u_k$. It is straightforward to verify that

$$S_n(t)\mathcal{P}_n = S(t)\mathcal{P}_n, \tag{4.27}$$

and

$$\left\| \mathcal{A}^{-\nu} (I - \mathcal{P}_n) u \right\|_U \leq \lambda_{n+1}^{-\nu} \|u\|_U, \quad v \geq 0, \ u \in U. \tag{4.28}$$

Furthermore, utilizing (4.27) and (4.28), we obtain

$$\|(S(t) - S_n(t)\mathcal{P}_n) u\|_U = \|S(t) (I - \mathcal{P}_n) u\|_U \leq M t^{-\nu} \lambda_{n+1}^{-\nu} \|u\|_U, \quad u \geq 0, \ u \in U. \tag{4.29}$$

The spectral Galerkin method for equation (4.6) can thus be formulated as

$$\mathrm{d}X_n(t) = \tilde{\mathcal{A}}_n X_n(t)\, dt + \mathcal{P}_n \mathcal{B} u(t)\, dt + \mathcal{P}_n \mathcal{N}(X_n(t))\, \tilde{\mathcal{P}}_n \circ dW^H(t), \quad t \in [0, T], \tag{4.30}$$

with the initial condition $X_n(0) = \mathcal{P}_n X_0$ and $X_n(t) \in U_n$ for all t.

The corresponding mild solution to equation (4.30) is expressed as

$$\begin{aligned} X_n(t) =& S_n(t) X_n(0) + \int_0^t S_n(t-s) \mathcal{P}_n \mathcal{B} u(s)\, ds \\ &+ \int_0^t S_n(t-s) \mathcal{P}_n \mathcal{N}(X_n(s))\, \tilde{\mathcal{P}}_n \circ dW^H(s), \end{aligned} \tag{4.31}$$

where $X_n(0) = \mathcal{P}_n X_0$ and $t \in [0, T]$. Our primary objective now is to establish an error

bound for the spatial semidiscretization. To achieve this, we begin by presenting a key lemma.

Lemma 4.18 Assume that the conditions of Theorem 4.17 are satisfied. Then, there exists a constant $C > 0$ such that

$$\|X_n(t)\|_{\mathcal{O}_\gamma^2} = \sup_{t\in[0,T]} \left(\mathbb{E}\|X_n(t)\|_{\dot{U}^\gamma}^2\right)^{\frac{1}{2}} \leq C.$$

Proof. The proof follows similarly to that of Theorem 4.17 and is therefore omitted for brevity. □

Theorem 4.19 Under the assumptions of Theorem 4.17, there exists a constant $C > 0$ such that

$$\|\mathcal{X}(t) - X_n(t)\|_{L^2(\Omega;U)} \leq C\left(\mu_{n+1}^{\frac{1}{2}} + \lambda_{n+1}^{-\frac{\gamma}{2}}\right), \quad t \in [0,T]. \tag{4.32}$$

Proof. The following proof is based on the fundamental reasoning used in Theorem 4.2 of [29]. By subtracting equation (4.31) from equation (4.11) and evaluating the norms, we obtain

$$\begin{aligned}
\|\mathcal{X}(t) - X_n(t)\|_{L^2(\Omega;U)} &\leq \|(S(t) - S_n(t)\mathcal{P}_n)\, X_0\|_{L^2(\Omega;U)} \\
&+ \left\| \int_0^t (S(t-s)\mathcal{B}u(s) - S_n(t-s)\mathcal{P}_n\mathcal{B}u(s))\,\mathrm{d}s \right\|_{L^2(\Omega;U)} \\
&+ \left\| \int_0^t \left(S(t-s)\mathcal{N}(\mathcal{X}(s)) - S_n(t-s)\mathcal{P}_n\mathcal{N}(X_n(s))\tilde{\mathcal{P}}_n\right) \circ \mathrm{d}W^H(s) \right\|_{L^2(\Omega;U)} \\
&=: J_1 + J_2 + J_3,
\end{aligned} \tag{4.33}$$

Clearly, by considering equation (4.8), Assumption 4.10 and equation (4.28), we can derive the following estimate for J_1:

$$J_1 = \left\| \mathcal{A}^{\frac{\gamma}{2}} S(t) \mathcal{A}^{-\frac{\gamma}{2}} \left(I - \mathcal{P}_n\right) X_0 \right\|_{L^2(\Omega;U)} \leq C\lambda_{n+1}^{-\frac{\gamma}{2}}. \tag{4.34}$$

For the term J_2, applying Assumption 4.11, along with equations (4.7), (4.8) and (4.28), we obtain

$$\begin{aligned}
J_2 &= \left\| \int_0^t \mathcal{A}^{\frac{1-\beta}{2}} S(t-s) \mathcal{A}^{-\frac{\gamma}{2}} \left(I - \mathcal{P}_n\right) A^{\frac{\gamma+\beta-1}{2}} \mathcal{B}u(s) \mathrm{d}s \right\|_{L^2(\Omega;U)} \\
&\leq C \left(\sup_{s\in[0,T]} \mathbb{E}\|u(s)\|_2^2 \right)^{\frac{1}{2}} \|\mathcal{B}\|_{L(\mathbb{R}^m;\dot{U}^{\gamma+\beta-1})} T^{\frac{\beta}{2}} \lambda_{n+1}^{-\frac{\gamma}{2}} \\
&\leq C\lambda_{n+1}^{-\frac{\gamma}{2}}
\end{aligned} \tag{4.35}$$

To estimate J_3, we begin by splitting it into three components:

$$\begin{aligned} J_3 \leq & \left\| \int_0^t \left(S(t-s)\mathcal{N}(\mathcal{X}(s)) - S_n(t-s)\mathcal{P}_n\mathcal{N}(\mathcal{X}(s)) \right) \circ \mathrm{d}W^H(s) \right\|_{L^2(\Omega;U)} \\ & + \left\| \int_0^t \left(S_n(t-s)\mathcal{P}_n\mathcal{N}(\mathcal{X}(s)) - S_n(t-s)\mathcal{P}_n\mathcal{N}(X_n(s)) \right) \circ \mathrm{d}W^H(s) \right\|_{L^2(\Omega;U)} \\ & + \left\| \int_0^t \left(S_n(t-s)\mathcal{P}_n\mathcal{N}(X_n(s)) - S_n(t-s)\mathcal{P}_n\mathcal{N}(X_n(s))\tilde{\mathcal{P}}_n \right) \circ \mathrm{d}W^H(s) \right\|_{L^2(\Omega;U)} \\ := & J_{3,1} + J_{3,2} + J_{3,3}. \end{aligned} \tag{4.36}$$

Using equation (4.12) along with the Minkowski inequality, we obtain

$$\begin{aligned} J_{3,1} \leq & \left\| \int_0^t S(t-s)(I-\mathcal{P}_n)\mathcal{N}(\mathcal{X}(s))\delta W^H(s) \right\|_{L^2(\Omega;U)} \\ & + \left\| \int_0^t \int_0^s D_s^H \left(S(t-\tau)(I-\mathcal{P}_n)\mathcal{N}(\mathcal{X}(\tau)) \right) \phi(s,\tau)\, d\tau \mathrm{d}s \right\|_{L^2(\Omega;U)} \\ =: & J'_{3,1} + J''_{3,1}, \end{aligned} \tag{4.37}$$

Utilizing Theorem 4.2 from [29], we obtain the following:

$$J'_{3,1} \leq C\lambda_{n+1}^{-\frac{\gamma}{2}} t^{\frac{2H+\beta-\gamma-1}{2}} + C\lambda_{n+1}^{-\frac{\gamma}{2}}. \tag{4.38}$$

In the case of $J''_{3,1}$, utilizing Assumption 4.12 (iii) and applying the steps outlined in Lemma 4.15, we arrive at the following:

$$\begin{aligned} J''_{3,1} \leq & \left\| \left(\int_0^t \int_0^t \left\| D_s^H \mathcal{A}^{\frac{1-\beta}{2}} S(t-\tau)\mathcal{A}^{\frac{-\gamma}{2}} (I-\mathcal{P}_n)\mathcal{A}^{\frac{\gamma-\beta+1}{2}} \mathcal{N}(\mathcal{X}(\tau)) \right\|_{\mathcal{L}_2^0}^2 |\phi(s,\tau)|^2 \mathrm{d}s\, d\tau \right)^{\frac{1}{2}} \right\|_{L^2(\Omega;\mathbb{R})} \\ \leq & C\lambda_{n+1}^{-\frac{\gamma}{2}} \left(\int_0^t \int_0^t \left\| \mathcal{A}^{\frac{1-\beta}{2}} S(t-\tau) \right\|^2 |\phi(s,\tau)|^2 \mathrm{d}s\, d\tau \right)^{\frac{1}{2}} \\ \leq & C\lambda_{n+1}^{-\frac{\gamma}{2}} t^{\frac{4H+\beta-3}{2}}. \end{aligned} \tag{4.39}$$

Substituting (4.38) and (4.39) into (4.37), we obtain

$$J_{3,1} \leq C\lambda_{n+1}^{-\frac{\gamma}{2}} + C\lambda_{n+1}^{-\frac{\gamma}{2}} t^{\frac{2H+\beta-\gamma-1}{2}} + C\lambda_{n+1}^{-\frac{\gamma}{2}} t^{\frac{4H+\beta-3}{2}} \tag{4.40}$$

In the case of $J_{3,3}$, we can write

$$J_{3,3} = \left\| \int_0^t \left(S_n(t-s)\mathcal{P}_n\mathcal{N}(X_n(s)) \right) (I-\tilde{\mathcal{P}}_n) \circ \mathrm{d}W^H(s) \right\|_{L^2(\Omega;U)},$$

now, by applying (4.4) in place of $W^H(t)$, (4.12) and Assumption 4.8, then the Theorem

7.10 of [44], we obtain the following expression for $J_{3,3}$:

$$\begin{aligned}
J_{3,3} &= \left\| \int_0^t \left(S_n(t-s)\mathcal{P}_n\mathcal{N}(X_n(s))\right)\left(I-\tilde{\mathcal{P}}_n\right)\circ \mathrm{d}W^H(s)\right\|_{L^2(\Omega;U)} \\
&\leq C\mu_{n+1}^{\frac{1}{2}} \left\| \left(\int_0^t \|K^*_{H,t}\mathcal{A}^{\frac{-\beta+1}{2}} S_n(t-s)\mathcal{P}_n\mathcal{A}^{\frac{\beta-1}{2}}\mathcal{N}(X_n(s))\|^2_{\mathcal{L}_2^0}\mathrm{d}s\right)^{\frac{1}{2}}\right\|_{L^2(\Omega;\mathbb{R})} \\
&\quad + C\mu_{n+1}^{\frac{1}{2}} \left\| \left(\int_0^t \int_0^t \|D_s^H\mathcal{A}^{\frac{-\beta+1}{2}} S_n(t-\tau)\mathcal{P}_n\mathcal{A}^{\frac{\beta-1}{2}}\mathcal{N}(X_n(\tau))\|_{\mathcal{L}_2^0}\, d\tau\mathrm{d}s\right)^{\frac{1}{2}}\right\|_{L^2(\Omega;\mathbb{R})} \\
&\quad + C\mu_{n+1}^{\frac{1}{2}} \left\| \left(\int_0^t \int_0^t \|D_s^H\mathcal{A}^{\frac{-\beta+1}{2}} S_n(t-\tau)\mathcal{P}_n\mathcal{A}^{\frac{\beta-1}{2}}\mathcal{N}(X_n(\tau))\|^2_{\mathcal{L}_2^0}|\phi(s,\tau)|^2\, d\tau\mathrm{d}s\right)^{\frac{1}{2}}\right\|_{L^2(\Omega;\mathbb{R})}
\end{aligned}$$

Following the same argument as in Lemma 4.15, we obtain

$$J_{3,3} \leq C\mu_{n+1}^{\frac{1}{2}} t^{\frac{2H+\beta-1}{2}} + C\mu_{n+1}^{\frac{1}{2}} + C\mu_{n+1}^{\frac{1}{2}} t^{\frac{4H+\beta-3}{2}} \tag{4.41}$$

We now turn our attention to the estimation $J_{3,2}$. Utilizing Assumption 4.12 (ii) along with equation (4.12), we obtain

$$\begin{aligned}
J_{3,2} &\leq \left\| \int_0^t \mathcal{A}^{\frac{1-\beta}{2}} S_n(t-s)\mathcal{P}_n\mathcal{A}^{\frac{\beta-1}{2}} \left(\mathcal{N}(\mathcal{X}(s)) - \mathcal{N}(X_n(s))\right)\delta W^H(s)\right\|_{L^2(\Omega;U)} \\
&\quad + \left\| \int_0^t \int_0^t D_s^H\mathcal{A}^{\frac{1-\beta}{2}} S_n(t-\tau)\mathcal{P}_n\mathcal{A}^{\frac{\beta-1}{2}} \left(\mathcal{N}(\mathcal{X}(\tau)) - \mathcal{N}(X_n(\tau))\right)\phi(s,\tau)\mathrm{d}s\, d\tau\right\|_{L^2(\Omega;U)} \\
&:= J'_{3,2} + J''_{3,2}
\end{aligned} \tag{4.42}$$

Based on Theorem 4.2 from [29], we can write

$$\begin{aligned}
J'_{3,2} &\leq C\left(\int_0^t s^{1-2H}(t-s)^{2H+\beta-2} \|\mathcal{X}(s) - X_n(s)\|^2_{L^2(\Omega;U)}\,\mathrm{d}s\right)^{\frac{1}{2}} \\
&\quad + C\left(\int_0^t (t-s)^{\beta-1} \|\mathcal{X}(s) - X_n(s)\|^2_{L^2(\Omega;U)}\,\mathrm{d}s\right)^{\frac{1}{2}}
\end{aligned} \tag{4.43}$$

By using (ii) from Assumption 4.12 and (4.27), we obtain

$$\begin{aligned}
J''_{3,2} &\leq \left(\int_0^t \int_0^t \left\| D_s^H\mathcal{A}^{\frac{1-\beta}{2}} S(t-\tau)\mathcal{P}_n\mathcal{A}^{\frac{\beta-1}{2}} \left(\mathcal{N}(\mathcal{X}(\tau)) - \mathcal{N}(X_n(\tau))\right)\right\|^2_{\mathcal{L}_2^0} |\phi(s,\tau)|^2 \mathrm{d}s\, d\tau\right)^{\frac{1}{2}} \\
&\leq C\left(\int_0^t \int_0^t \left\| \mathcal{A}^{\frac{-\beta+1}{2}} S(t-\tau)\right\|^2 |\phi(s,\tau)|^2 \|\mathcal{X}(s) - X_n(s)\|^2_{L^2(\Omega;U)}\,\mathrm{d}s\, d\tau\right)^{\frac{1}{2}} \\
&\leq C\left(\int_0^t (t-s)^{\beta-1} s^{4H-3} \|\mathcal{X}(s) - X_n(s)\|^2_{L^2(\Omega;U)}\,\mathrm{d}s\right)^{\frac{1}{2}}
\end{aligned} \tag{4.44}$$

Substituting (4.43) and (4.44) into (4.42), we obtain

$$\begin{aligned} J_{3,2} \leq C &\left(\int_0^t s^{1-2H}(t-s)^{2H+\beta-2} \left\| \mathcal{X}(s) - X_n(s) \right\|_{L^2(\Omega;U)}^2 \mathrm{d}s \right)^{\frac{1}{2}} \\ &+ C \left(\int_0^t (t-s)^{\beta-1} \left\| \mathcal{X}(s) - X_n(s) \right\|_{L^2(\Omega;U)}^2 \mathrm{d}s \right)^{\frac{1}{2}} \\ &+ C \left(\int_0^t (t-s)^{\beta-1} s^{4H-3} \left\| \mathcal{X}(s) - X_n(s) \right\|_{L^2(\Omega;U)}^2 \mathrm{d}s \right)^{\frac{1}{2}} \end{aligned} \tag{4.45}$$

By substituting equations (4.40), (4.41) and (4.45) into (4.36), we obtain the following estimate for J_3:

$$\begin{aligned} J_3 \leq & C\mu_{n+1}^{\frac{1}{2}} + C\lambda_{n+1}^{-\frac{\gamma}{2}} + C \left(\int_0^t s^{1-2H}(t-s)^{2H+\beta-2} \left\| \mathcal{X}(s) - X_n(s) \right\|_{L^2(\Omega;U)}^2 \mathrm{d}s \right)^{\frac{1}{2}} \\ &+ C \left(\int_0^t (t-s)^{\beta-1} \left\| \mathcal{X}(s) - X_n(s) \right\|_{L^2(\Omega;U)}^2 \mathrm{d}s \right)^{\frac{1}{2}} \\ &+ C \left(\int_0^t (t-s)^{\beta-1} s^{4H-3} \left\| \mathcal{X}(s) - X_n(s) \right\|_{L^2(\Omega;U)}^2 \mathrm{d}s \right)^{\frac{1}{2}} \end{aligned} \tag{4.46}$$

Furthermore, inserting (4.34), (4.35) and (4.46) into (4.33), we have

$$\begin{aligned} \left\| \mathcal{X}(t) - X_n(t) \right\|_{L^2(\Omega;U)} \leq C &\left(\mu_{n+1}^{\frac{1}{2}} + \lambda_{n+1}^{-\frac{\gamma}{2}} \right) \\ &+ C \left(\int_0^t s^{1-2H}(t-s)^{2H+\beta-2} \left\| \mathcal{X}(s) - X_n(s) \right\|_{L^2(\Omega;U)}^2 \mathrm{d}s \right)^{\frac{1}{2}} \\ &+ C \left(\int_0^t (t-s)^{\beta-1} \left\| \mathcal{X}(s) - X_n(s) \right\|_{L^2(\Omega;U)}^2 \mathrm{d}s \right)^{\frac{1}{2}} \\ &+ C \left(\int_0^t (t-s)^{\beta-1} s^{4H-3} \left\| \mathcal{X}(s) - X_n(s) \right\|_{L^2(\Omega;U)}^2 \mathrm{d}s \right)^{\frac{1}{2}} \end{aligned}$$

In this situation, the conditions of Lemma 4.16 are satisfied. Hence, applying the Cauchy inequality, we similarly derive (4.32). This concludes the proof. □

4.4 Example

4.4.1 Stochastic Heat Equation

A key example that meets the generator assumptions is $\tilde{\mathcal{A}} = \Delta$, which corresponds to the heat equation. In the following, we present an example to illustrate the application of the abstract framework discussed in Section 4.2. We begin with a modified version of an example studied in [14]. In particular, not an Itô equation driven by a Brownian motion is studied. Instead, we consider the following Young/Stratonovich stochastic partial differential equation driven by a (scalar) fractional Brownian motion W^H with the Hurst

parameter $H \in [1/2, 1)$.

Example 4.20 Consider a two-dimensional surface with perfect insulation along its boundaries and a heat source located at its center. This situation can be described using the following controlled stochastic partial differential equation, where $t \geq 0$ and $\zeta \in [0, \pi]^2$:

$$\begin{aligned} \frac{\partial \mathcal{X}(t,\zeta)}{\partial t} &= \Delta \mathcal{X}(t,\zeta) + 1_{[\frac{\pi}{4},\frac{3\pi}{4}]^2}(\zeta)u(t) + e^{-|\zeta_1 - \frac{\pi}{2}| - \zeta_2}\mathcal{X}(t,\zeta) \circ \frac{\partial W^H(t)}{\partial t}, \\ \mathcal{X}(t,\zeta) &= 0, \quad t \geq 0, \quad \zeta \in \partial[0,\pi]^2, \quad \text{and} \quad \mathcal{X}(0,\zeta) = 0, \end{aligned} \tag{4.47}$$

where W^H is a scalar fBm that can model a random heat source or the impact of wind.

Furthermore, we define

(i) $U = L^2\left([0,\pi]^2\right), \quad V = \mathbb{R}, \quad m = 1,$

(ii) $\tilde{\mathcal{A}}$ is the Laplace operator, $\mathcal{B} = 1_{[\frac{\pi}{4},\frac{3\pi}{4}]^2}(\cdot)$,

(iii) Let $g(\zeta_1, \zeta_2) = e^{-|\zeta_1 - \frac{\pi}{2}| - \zeta_2}$, then $\mathcal{N}(x) = g(\cdot)x = e^{-|\cdot - \frac{\pi}{2}| - \cdot}x$ for $x \in L^2([0,\pi]^2)$.

The eigenvalues of the Laplacian on $[0,\pi]^2$ are expressed as $-\lambda_{ij} = -(i^2 + j^2)$ for $i, j \in \mathbb{Z}^+$. The corresponding eigenvectors, which form an orthonormal basis, are defined as $u_{ij} = \frac{f_{ij}}{\|f_{ij}\|_U}$, where $f_{ij} = \sin(i\cdot)\sin(j\cdot)$. For simplicity and to match the form presented in (4.25), we denote the k-th largest eigenvalue as $-\lambda_k$, $k \in \mathbb{N}$ and the associated eigenvector as u_k.

The scalar output of the system is the average temperature over the non-heated area as follows

$$\mathcal{Y}(t) = \frac{4}{3\pi^2}\int_{[0,\pi]^2 \setminus [\frac{\pi}{4},\frac{3\pi}{4}]^2} \mathcal{X}(t,\zeta)d\zeta, \tag{4.48}$$

where $\mathcal{C}x = \frac{4}{3\pi^2}\int_{[0,\pi]^2\setminus[\frac{\pi}{4},\frac{3\pi}{4}]^2} x(\zeta)d\zeta$ for $x \in L^2([0,\pi]^2)$.

Now, we apply the spectral Galerkin method introduced in Section 4.3 to the type of equation introduced in Example 4.20. We assume the presence of an orthonormal basis $(u_k)_{k\in\mathbb{N}}$ of U, which is included in $D(\mathcal{A})$. Using (4.30), we can achieve the spectral Galerkin method as follows.

$$\begin{aligned} dX_n(t) &= \left[\tilde{\mathcal{A}}_n X_n(t) + \mathcal{B}_n u(t)\right]dt + \mathcal{N}_n\left(X_n(t)\right) \circ \mathrm{d}W^H(t), \quad t \in [0,1], \\ X_n(0) &= X_{0,n}, \end{aligned} \tag{4.49}$$

where

- $\tilde{\mathcal{A}}_n x = \mathcal{P}_n(-\mathcal{A})x = \sum_{k=1}^n \langle -\mathcal{A}x, u_k\rangle_U u_k \in U_n$ for all $x \in D(\mathcal{A})$,
- $\mathcal{B}_n x = \mathcal{P}_n\mathcal{B}x = \sum_{k=1}^n \langle \mathcal{B}x, u_k\rangle_U u_k \in U_n$ for all $x \in \mathbb{R}^m$,
- $\mathcal{N}_n(x) = \mathcal{P}_n(g(\cdot)x) = \sum_{k=1}^n \langle g(\cdot)x, u_k\rangle_U u_k \in U_n$ for all $x \in U$,
- $X_{0,n} = \mathcal{P}_n X_0 = \sum_{k=1}^n \langle X_0, u_k\rangle_U u_k \in U_n$.

The mild solution of equation (4.49) is expressed as

$$X_n(t) = S_n(t)X_{0,n} + \int_0^t S_n(t-s)\mathcal{B}_n u(s)\mathrm{d}s + \int_0^t S_n(t-s)\mathcal{N}_n\left(X_n(s)\right) \circ \mathrm{d}W^H(s), \quad (4.50)$$

for $t \geq 0$. Furthermore, we consider the p-dimensional approximating output.

$$y_n(t) = CX_n(t), \ t \geq 0.$$

We express $X_n(t)$ as:

$$X_n(t) = \sum_{k=1}^{n} \langle \mathcal{X}(t), u_k \rangle_U u_k,$$

where $\{u_k\}_{k\in\mathbb{N}}$ is an orthonormal basis of the Hilbert space U. We now describe the vector of Fourier coefficients for the Galerkin solution X_n and define

$$x(t) = \left(\langle X_n(t), u_1\rangle_U, \ldots, \langle X_n(t), u_n\rangle_U\right)^\top.$$

The components of x satisfy the following:

$$\begin{aligned}\langle X_n(t), u_k\rangle_U = \langle S_n(t)X_{0,n}, u_k\rangle_U &+ \int_0^t \langle S_n(t-s)\mathcal{B}_n u(s), u_k\rangle_U \,\mathrm{d}s \\ &+ \left\langle \int_0^t S_n(t-s)\mathcal{N}_n\left(X_n(s)\right) \circ \mathrm{d}W^H(s), u_k \right\rangle_U.\end{aligned}$$

Using the representation $S_n(t)x = \sum_{i=1}^n e^{-\lambda_i t}\langle x, u_i\rangle_U u_i \, (x \in U_n)$, we obtain

$$\langle S_n(t)X_{0,n}, u_k\rangle_U = e^{-\lambda_k t}\langle X_{0,n}, u_k\rangle_U = e^{-\lambda_k t}\langle x_0, u_k\rangle_U,$$

and

$$\langle S_n(t-s)\mathcal{B}_n u(s), u_k\rangle_U = e^{-\lambda_k(t-s)}\langle \mathcal{B}_n u(s), u_k\rangle_U = e^{-\lambda_k(t-s)}\langle \mathcal{B}, u_k\rangle_U u(s),$$

this holds for $k = 1, \ldots, n$. Furthermore, we have

$$\begin{aligned}\langle \int_0^t S_n(t-s)\mathcal{N}_n\left(X_n(s)\right) \circ \mathrm{d}W^H(s), u_k\rangle_U &\\ = \int_0^t \langle S_n(t-s)\mathcal{N}_n\left(X_n(s)\right), u_k\rangle_U \circ \mathrm{d}W^H(s) &\\ = \sum_{i=1}^n \int_0^t \langle S_n(t-s)g(\cdot)u_i, u_k\rangle_U \langle X_n(s), u_i\rangle_U \circ \mathrm{d}W^H(s) &\\ = \sum_{i=1}^n \int_0^t e^{-\lambda_k(t-s)}\langle g(\cdot)u_i, u_k\rangle_U \langle X_n(s), u_i\rangle_U \circ \mathrm{d}W^H(s) &\end{aligned}$$

Hence, in compact form, x is given by

$$x(t) = e^{At}x_0 + \int_0^t e^{A(t-s)}Bu(s)\mathrm{d}s + \int_0^t e^{A(t-s)}Nx(s) \circ \mathrm{d}W^H(s) \quad (4.51)$$

where

- $A = \operatorname{diag}(-\lambda_1, \ldots, -\lambda_n)$
- $B = (\langle \mathcal{B}, u_k \rangle_U)_{k=1,\ldots,n}$
- $N = (\langle g(\cdot)u_i, u_k \rangle_U)_{k,i=1,\ldots,n}$
- $x_0 = (\langle x_0, u_1 \rangle_U, \ldots, \langle x_0, u_n \rangle_U)^\top$.

Below, we demonstrate that the solution to equation (4.51) also satisfies the strong solution equation. Define

$$g(t) := x_0 + \int_0^t \mathrm{e}^{-As} B u(s) \mathrm{d}s + \int_0^t \mathrm{e}^{-As} N x(s) \circ \mathrm{d}W^H(s), \quad t \geq 0.$$

According to Theorem 2.14, for $x(t) = \mathrm{e}^{At} g(t)$, we have

$$\begin{aligned} e_i^\top x(t) = e_i^\top \mathrm{e}^{At} g(t) &= e_i^\top g(0) + \int_0^t d\left(e_i^\top \mathrm{e}^{As}\right) g(s) + \int_0^t e_i^\top \mathrm{e}^{As}\, dg(s) \\ &= e_i^\top \left(x_0 + \int_0^t A \mathrm{e}^{As} g(s) \mathrm{d}s + \int_0^t B u(s) \mathrm{d}s + \int_0^t N x(s) \circ \mathrm{d}W^H(s) \right), \end{aligned}$$

where e_i is the i-th unit vector of $\mathbb{R}^n$. Therefore, we have

$$x(t) = x_0 + \int_0^t [A x(s) + B u(s)] \mathrm{d}s + \int_0^t N x(s) \circ \mathrm{d}W^H(s), \quad t \geq 0. \tag{4.52}$$

The scalar output corresponding to the Galerkin solution is given by

$$y_n(t) = C x(t), \quad t \geq 0,$$

where $C^\top = (\mathcal{C} u_k)_{k=1,\ldots,n}$. This follows from the relation

$$y_n(t) = C x(t) = \sum_{k=1}^n \langle X_n(t), u_k \rangle_U \mathcal{C} u_k. \tag{4.53}$$

The Fourier coefficients of the Galerkin solution for SPDE in Example (4.20) are given by (4.52) and are expressed in terms of the following components:

- $A = \operatorname{diag}(0, -1, -1, -2, \ldots)$,
- $B = \left(\left\langle 1_{\left[\frac{\pi}{4}, \frac{3\pi}{4}\right]^2}(\cdot), u_k \right\rangle_U \right)_{k=1,\ldots,n}$,
- $N = \left(\left\langle e^{-|\cdot - \frac{\pi}{2}| - \cdot} u_i, u_k \right\rangle_U \right)_{k,i=1,\ldots,n}$.

5 Gramian-Based Model Reduction for Unstable Stochastic Systems

As we revisit the system dynamics previously discussed in Chapter 3, we present the system (3.19) again for continuity and ease of reference. The matrices within the system retain their defined properties as follows:

$$dx(t) = [Ax(t) + Bu(t)]\mathrm{d}t + \sum_{i=1}^{q} N_i x(t)\mathrm{d}W_i(t), \qquad x(0) = x_0, \tag{5.1a}$$

$$y(t) = Cx(t), \qquad t \in [0, T], \tag{5.1b}$$

where (5.1a) represents the state dynamics and (5.1b) defines the output equation of the system. The state dimension n is assumed to be large and the quantity of interest y is often low-dimensional, i.e., $p \ll n$, but we also discuss the case of a large p. Using Definition 3.10, $x(t, x_0, u)$, defined the state in dependence on the initial state x_0 and the control $u \in \mathcal{M}^2([0, T]; \mathbb{R}^m)$ (see (3.20)).

The goal is to construct a system with state x_r and quantity of interest y_r having the same structure as (5.1) but a much smaller state dimension $r \ll n$ (see Section 3.2.4.1). At the same time, the aim is to ensure $y \approx y_r$. The detailed results of this study have been published in [101].

5.1 Gramian-Based MOR

5.1.1 Gramians and Characterization of Dominant Subspaces

Identifying the effective dimensionality of the system (5.1) requires the study of the fundamental solution to the homogeneous stochastic state equation defined in (3.24). As previously developed, the matrix-valued stochastic process Φ, which characterizes these fundamental solutions, satisfies

$$\Phi(t) = I_n + \int_0^t A\Phi(s)\mathrm{d}s + \sum_{i=1}^{q} \int_0^t N_i \Phi(s) dW_i(s), \quad t \in [0, T], \tag{5.2}$$

where I_n denotes the identity matrix. By multiplying (5.2) by the initial state vector x_0 from the right and assuming no control input ($u \equiv 0$), we obtain the solution to the stochastic state equation (5.1a).

Now, we extend the concept of stochastic Gramians discussed in Chapter 3 by introducing their time-limited counterparts. Based on the stochastic fundamental matrix Φ defined previously, we establish the time-limited Gramians as follows:

$$P_T := \mathbb{E} \int_0^T \Phi(s) B B^\top \Phi^\top(s) \mathrm{d}s \tag{5.3}$$

$$Q_T := \mathbb{E} \int_0^T \Phi^\top(s) C^\top C \Phi(s) \mathrm{d}s, \tag{5.4}$$

where P_T and Q_T are supposed to identify the less relevant states in (5.1a) and (5.1b), respectively. The time-limited Gramians P_T and Q_T serve as stochastic analogs to the deterministic time-limited Gramians, which are obtained in the absence of noise influences ($N_i = 0$ for all i), leading to $\Phi(t) = e^{At}$. MOR scheme based on such Gramians in a deterministic framework is investigated, e.g., in [37, 64, 102]. P_T and Q_T generally exist in contrast to their limits $\lim_{T\to\infty} P_T$ and $\lim_{T\to\infty} Q_T$ which require mean square asymptotic stability. MOR methods based on these limits are, for example, considered in [9, 14, 96, 103] and are already analyzed in detail. However, in practice, the necessary stability condition is often not satisfied.

Building on the foundational analysis introduced in Chapter 3, this section delves deeper into the relationship between the covariance matrix P_T and the dominant subspaces of the system described by (5.1a), specifically under the condition of zero initial data. As previously outlined, we consider an orthonormal basis $(p_k)_{k=1,\dots,n}$ of $\mathbb{R}^n$, consisting of eigenvectors of P_T, to express the state evolution:

$$x(t;0,u) = \sum_{k=1}^{n} \langle x(t;0,u), p_k \rangle_2 \, p_k.$$

As delineated in Section 3.2.3, with the initial state set to zero, the upper bound of the Fourier coefficients is governed by:

$$\sup_{t\in[0,T]} \mathbb{E} \left| \langle x(t,0,u), p_k \rangle_2 \right| \leq \lambda_k^{\frac{1}{2}} \left\| u \right\|_T, \tag{5.5}$$

where λ_k denotes the eigenvalue associated with p_k. A smaller λ_k suggests a negligible influence of the corresponding eigenvector p_k on the behavior of the system, allowing the reduction of its presence in the system model.

Further, to assess the impact of the state directions on the quantity of interest y, we revisit the expansion of the initial state x_0 as:

$$x_0 = \sum_{k=1}^{n} \langle x_0, q_k \rangle_2 \, q_k,$$

where $(q_k)_{k=1,\dots,n}$ is an orthonormal basis of eigenvectors of Q_T, each associated with an

eigenvalue μ_k. The solution representation of the state variable provides that:

$$\begin{aligned} y(t;x_0,u) &= C\Phi(t)x_0 + C\int_0^t \Phi(t,s)Bu(s)\mathrm{d}s \\ &= \sum_{k=1}^{n} \langle x_0, q_k\rangle_2 \, C\Phi(t)q_k + C\int_0^t \Phi(t,s)Bu(s)\mathrm{d}s, \end{aligned}$$

with $t \in [0,T]$. The eigenspaces corresponding to minor eigenvalues μ_k are determined to be of minimal consequence to y, supported by:

$$\mathbb{E}\int_0^T \|C\Phi(t)q_k\|_2^2 \,\mathrm{d}t = q_k^\top Q_T q_k = \mu_k. \tag{5.6}$$

This indicates that the eigenspaces of Q_T corresponding to small eigenvalues μ_k are not crucial for the performance of the system. Given that the less significant directions in the state space, as identified in equations (5.1a) and (5.1b) through (5.5) and (5.6), are intended to be eliminated. This can be done by diagonalizing P_T such that less important variables in (5.1a) can be easily identified and truncated. Another, but computationally more expensive, approach is based on simultaneously diagonalizing P_T and Q_T which allows the removal of more redundant information from the system. Both strategies are discussed in Section 5.1.2.

Below, we point out the relation between the Gramians and linear matrix differential equations, crucial for computing these Gramians P_T and Q_T to facilitate the derivation of a reduced system. Specifically, by employing the differential formulation from equation (3.37), it can be demonstrated that the matrix function

$$F(t) = \mathbb{E}[\Phi(t)BB^\top\Phi^\top(t)], \quad \text{for} \quad t \in [0,T], \tag{5.7}$$

satisfies the following differential equation:

$$\dot{F}(t) = AF(t) + F(t)A^\top + \sum_{i,j=1}^{q} N_i F(t) N_j^\top K_{ij}, \quad F(0) = BB^\top.$$

Utilizing the operators $\mathcal{L}_A$ and Π_N, as defined in equations (3.5) and (3.30) respectively, we can express this relationship more precisely:

$$\dot{F}(t) = \mathcal{L}_A\left(F(t)\right) + \Pi_N\left(F(t)\right), \quad F(0) = BB^\top. \tag{5.8}$$

Integrating both sides of (5.8) yields

$$F(T) - BB^\top = \mathcal{L}_A\left(P_T\right) + \Pi_N\left(P_T\right). \tag{5.9}$$

The link between Q_T and the corresponding matrix equation is established in a different way. Before formulating this result in the following proposition, we first clarify the definitions of the Lyapunov operators $\mathcal{L}_A^*$ and Π_N^*, as delineated in (3.6) and (3.31)

respectively:

$$\mathcal{L}_A^* = A^\top X + XA, \quad \Pi_N^*(X) := \sum_{i,j=1}^{q} N_i^\top X N_j k_{ij}.$$

Proposition 5.1 Let $C^\top C$ be contained in the eigenspace of the Lyapunov operator $\mathcal{L}_A^* + \Pi_N^*$. Then, $G(t) = \mathbb{E}[\Phi^\top(t) C^\top C \Phi(t)]$, $t \in [0, T]$, satisfies

$$\dot{G}(t) = \mathcal{L}_A^*(G(t)) + \Pi_N^*(G(t)), \quad G(0) = C^\top C. \tag{5.10}$$

Proof. Since $C^\top C$ is contained in the eigenspace of the Lyapunov operator, there exists $\alpha_1, \dots, \alpha_{n^2} \in \mathbb{C}$ such that $C^\top C = \sum_{k=1}^{n^2} \alpha_k \mathcal{V}_k$, where $(\mathcal{V}_k)$ are eigenvectors of $\mathcal{L}_A^* + \Pi_N^*$ corresponding to the eigenvalues (β_k). Then, we have

$$\mathbb{E}[\Phi^\top(t) C^\top C \Phi(t)] = \sum_{k=1}^{n^2} \alpha_k \mathbb{E}[\Phi^\top(t) \mathcal{V}_k \Phi(t)].$$

Let us apply Ito's product rule, see Example 1.27, to $\Phi^\top(t)\mathcal{V}_k\Phi(t)$ resulting in

$$d\left(\Phi^\top(t)\mathcal{V}_k\Phi(t)\right) = d\left(\Phi^\top(t)\right)\mathcal{V}_k\Phi(t) + \Phi^\top(t)\mathcal{V}_k d\left(\Phi(t)\right) + d\left(\Phi^\top(t)\right)\mathcal{V}_k d\left(\Phi(t)\right).$$

We insert the stochastic differential of Φ above, compare with (5.2), leading to

$$\begin{aligned}
d\left(\Phi^\top(t)\mathcal{V}_k\Phi(t)\right) &= \left(\Phi^\top(t)A^\top \mathrm{d}t + \sum_{i=1}^{q} \Phi^\top(t) N_i^\top \mathrm{d}W_i(t)\right)\mathcal{V}_k\Phi(t) \\
&\quad + \Phi^\top(t)\mathcal{V}_k\left(A\Phi(t)\mathrm{d}t + \sum_{i=1}^{q} N_i\Phi(t)\mathrm{d}W_i(t)\right) \\
&\quad + \Phi^\top(t)\sum_{i,j=1}^{q} N_i^\top \mathcal{V}_k N_j k_{ij}\Phi(t)\mathrm{d}t \\
&= \Phi^\top(t)\left(A^\top\mathcal{V}_k + \mathcal{V}_k A + \sum_{i,j=1}^{q} N_i^\top\mathcal{V}_k N_j k_{ij}\right)\Phi(t)\mathrm{d}t \\
&\quad + \sum_{i=1}^{q} \Phi^\top(t)\left(N_i^\top\mathcal{V}_k N_i\right)\Phi(t)\mathrm{d}W_i(t)
\end{aligned}$$

We apply the expected value to both sides of the above identity and exploit that the integrals have mean zero (see Lemma 1.19). Hence, we obtain

$$\frac{d}{dt}\mathbb{E}[\Phi^\top(t)\mathcal{V}_k\Phi(t)] = \mathbb{E}[\Phi^\top(t)(\mathcal{L}_A^* + \Pi_N^*)(\mathcal{V}_k)\Phi(t)] = \beta_k\mathbb{E}[\Phi^\top(t)\mathcal{V}_k\Phi(t)].$$

This implies that $\mathbb{E}[\Phi^\top(t)\mathcal{V}_k\Phi(t)] = \mathrm{e}^{\beta_k t}\,\mathcal{V}_k$ providing $\mathbb{E}[\Phi^\top(t)C^\top C\Phi(t)] = \sum_{k=1}^{n^2}\alpha_k\,\mathrm{e}^{\beta_k t}\,\mathcal{V}_k$.

Consequently, we obtain

$$\frac{d}{dt}\mathbb{E}[\Phi^\top(t)C^\top C\Phi(t)] = \sum_{k=1}^{n^2}\alpha_k\,\mathrm{e}^{\beta_k t}\,\beta_k\mathcal{V}_k = \sum_{i=1}^{n^2}\alpha_k\,\mathrm{e}^{\beta_k t}(\mathcal{L}_A^* + \Pi_N^*)(\mathcal{V}_k)$$
$$= (\mathcal{L}_A^* + \Pi_N^*)(\mathbb{E}[\Phi^\top(t)C^\top C\Phi(t)])$$

using the linearity of $\mathcal{L}_A^* + \Pi_N^*$. This concludes the proof. □

Remark 5.2 The assumptions of Proposition 5.1 are invariably satisfied if the operator $\mathcal{K}$, defined in Remark 3.18 as

$$\mathcal{K} = A \otimes I + I \otimes A + \sum_{i,j=1}^{q} N_i \otimes N_j k_{ij}, \tag{5.11}$$

is diagonalizable over $\mathbb{C}$ because in that case there is a basis of $\mathbb{C}^{n^2}$ consisting of eigenvectors of $\mathcal{K}^\top$. Hence, $\mathrm{vec}(C^\top C)$ can be spanned by these eigenvectors which are of the form $\mathrm{vec}(\mathcal{V}_k)$ with $\mathcal{V}_k$ being an eigenvector of $\mathcal{L}_A^* + \Pi_N^*$ providing that $C^\top C$ is in the eigenspaces of this operator. Therefore, from the computational point of view, the assumption of Proposition 5.1 does not restrict the generality since the set of diagonalizable $n^2 \times n^2$ matrices is dense in $\mathbb{C}^{n^2\times n^2}$.
In fact, we can find a stochastic representation of the solution to (5.10) different from $\mathbb{E}[\Phi^\top(t)C^\top C\Phi(t)]$, $t \in [0,T]$. Introducing the fundamental solution Φ_d by the equation

$$\Phi_d(t) = I + \int_0^t A^\top\Phi_d(s)\mathrm{d}s + \sum_{i=1}^{q}\int_0^t N_i^\top\Phi_d(s)\mathrm{d}W_i(s),$$

we see that $G(t) = \mathbb{E}[\Phi_d(t)C^\top C\Phi_d^\top(t)]$. This is a direct consequence of the relation between $\mathbb{E}[\Phi(t)BB^\top\Phi^\top(t)]$ and the solution of (5.8) when (A,B,N_i) is replaced by $(A^\top,C^\top,N_i^\top)$. Therefore, $\mathbb{E}[\Phi_d(t)C^\top C\Phi_d^\top(t)]$, $t\in[0,T]$, solves the equation (5.10) and hence coincides with $\mathbb{E}[\Phi^\top(t)C^\top C\Phi(t)]$, $t\in[0,T]$, given the assumption of Proposition 5.1.
Generally, we have $\Phi_d(t) \neq \Phi^\top(t)$. In case all matrices $A, N_1, \ldots, N_q$ commute, we know that A and N_i commute with Φ[1]. Hence, $\Phi_d(t) = \Phi^\top(t)$ which can be seen as transposing (5.2) and subsequently exploiting the commutative property. This is particularly given in the deterministic case where $N_i = 0$ for all $i = 1,\ldots,q$.

Under the assumption of Proposition 5.1, it holds that

$$G(T) - C^\top C = \mathcal{L}_A^*(Q_T) + \Pi_N^*(Q_T), \tag{5.12}$$

exploiting (5.10). In fact, we need to compute P_T and Q_T within the MOR procedure described later. Lyapunov equations (5.9) and (5.12) are used to do so. However, one needs to have access to $F(T)$ and $G(T)$ which are the terminal values of the matrix-differential equations (5.8) and (5.10). This is indeed very challenging in a framework, where $n \gg 100$. We will address possible approaches for computing P_T and Q_T for such settings in Section 5.3.

[1] see Proposition 3.13

5.1.2 Reduced Order Modeling by Transformation of Gramians

In this chapter, we delve deeper into MOR techniques that are previously introduced in Section 3.2.4. A crucial step in this process involves selecting an appropriate regular transformation matrix S, which redefines the state variable as $x_S(t) = Sx(t)$. This transformation, discussed in detail in Section 3.1.3.1, facilitates the dimensionality reduction of the system by focusing on its most significant dynamic behaviors.

Inserting this transformed state variable into the original system equation (5.1), as derived in Chapter 3, we obtain the transformed stochastic differential equations:

$$dx_S(t) = [A_S x_S(t) + B_S u(t)]\mathrm{d}t + \sum_{i=1}^{q} N_{i,S} x_S(t)\mathrm{d}W_i(t), \quad y(t) = C_S x_S(t), \quad t \in [0,T], \tag{5.13}$$

where the transformed system matrices are defined as follows:

$$(A_S, B_S, C_S, N_{i,S}) = (SAS^{-1}, SB, CS^{-1}, SN_iS^{-1}).$$

This setup ensures that the transformed system (5.13) has the same input-output behavior as (5.1) but the fundamental solution and hence the Gramians are different. The fundamental solution of (5.13) is $\Phi_S(t) = S\Phi(t)S^{-1}$ which can be observed by multiplying (5.2) with S from the left and with S^{-1} from the right. Consequently, the new Gramians are

$$P_{T,S} = \mathbb{E}\int_0^T \Phi_S(s)B_S B_S^\top \Phi_S^\top(s)\,ds = SP_TS^\top$$
$$Q_{T,S} = \mathbb{E}\int_0^T \Phi_S^\top(s)C_S^\top C_S\Phi_S(s)\mathrm{d}s = S^{-\top}Q_TS^{-1}$$

The idea is to diagonalize at least one of these Gramians, since in a system with diagonal Gramians, the orthonormal bases (p_k) and (q_k) are canonical unit vectors (columns of the identity matrix). Thus, unimportant directions can be identified easily by (5.5) and (5.6) and are associated to the small diagonal entries of the new Gramians. For the first approach, we set $S = S_1$, where S_1 is part of the eigenvalue decomposition $P_T = S_1^\top \Sigma_T^{(1)} S_1$. This leads to $P_{T,S} = \Sigma_T^{(1)}$ with $\Sigma_T^{(1)}$ being the diagonal matrix of eigenvalues of P_T. Notice that $S^\top = S^{-1}$ holds in this case. If (5.1a) is mean square asymptotically stable, P_T can be replaced by $\lim_{T\to\infty} P_T$. This method based on the limit is investigated in Section 3.2.4.3.

The second approach uses $S = S_2$ as the stochastic Balance Truncation method[2], which leads to $P_T = Q_T = \Sigma_T^{(2)}$, where $\Sigma_T^{(2)}$ is the diagonal matrix of the square roots of eigenvalues of P_TQ_T or HSVs of the system. Given $P_T, Q_T > 0$, the transformation S_2 and its inverse are obtained by

$$S_2 = {\Sigma_T^{(2)}}^{-\frac{1}{2}} U^\top R, \quad S_2^{-1} = LU{\Sigma_T^{(2)}}^{-\frac{1}{2}}. \tag{5.14}$$

where the ingredients of (5.14) are computed by the factorizations $P_T = LL^\top$, $Q_T = R^\top R$

[2] See Section 3.2.4.4

and the singular value decomposition of $RL = U\Sigma_T^{(2)}U^\top$. The same procedure can be conducted for the limits of the Gramians (as $T \to \infty$) if mean square asymptotic stability is given (see Section 3.2.4.4). However, such a stability condition is generally too restrictive in practice. We introduce the matrix

$$\Sigma_T = \operatorname{diag}(\sigma_{T,1}, \dots, \sigma_{T,n}) = \Sigma_T^{(1)}, \Sigma_T^{(2)}$$

as the diagonal matrix of either eigenvalues of P_T or of HSVs of system (5.1). For $S = S_1, S_2$ the coefficients of (5.13) are partitioned as follows

$$\begin{aligned} A_S &= \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}, \quad B_S = \begin{pmatrix} B_1 \\ B_2 \end{pmatrix}, \quad C_S = \begin{pmatrix} C_1 & C_2 \end{pmatrix}, \\ N_{i,S} &= \begin{pmatrix} N_{i,11} & N_{i,12} \\ N_{i,21} & N_{i,22} \end{pmatrix}, \quad x_S(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix}, \quad \Sigma_T = \begin{pmatrix} \Sigma_{T,1} & \\ & \Sigma_{T,2} \end{pmatrix}, \end{aligned} \tag{5.15}$$

where $x_1(t) \in \mathbb{R}^r$, $A_{11} \in \mathbb{R}^{r\times r}$, $B_1 \in \mathbb{R}^{r\times m}$, $C_1 \in \mathbb{R}^{p\times r}$, $N_{i,11} \in \mathbb{R}^{r\times r}$ and $\Sigma_{T,1} \in \mathbb{R}^{r\times r}$ etc. The variables x_2 are associated to the matrix $\Sigma_{T,2}$ of small diagonal entries of Σ_T and are the less relevant ones. A reduced system is now obtained by truncating the equations of x_2 in (5.13). Additionally, we set $x_2 \equiv 0$ in the equations for x_1 leading to a reduced system

$$\mathrm{d}x_r(t) = [A_r x_r(t) + B_r u(t)]\mathrm{d}t + \sum_{i=1}^{q} N_{i,r}\mathrm{d}W_i(t), \quad x_r(0) = x_{0,r} \tag{5.16a}$$

$$y_r(t) = C_r x_r(t), \quad t \in [0, T], \tag{5.16b}$$

As in equation (3.44), we observe that

$$\begin{bmatrix} x_r(t) & A_r & N_{i,r} & B_r & C_r \end{bmatrix} = \begin{bmatrix} x_1(t) & A_{11} & N_{i,r} & B_1 & C_1 \end{bmatrix},$$

approximating (5.1). Below, we give another interpretation for (5.16). Let us decompose the transformation, another representation for (3.12) and (3.13)

$$S = \begin{pmatrix} M^\top \\ \star \end{pmatrix}, \quad S^{-1} = \begin{pmatrix} V & \star \end{pmatrix} \tag{5.17}$$

where $M^\top$ and V are the first r rows and columns of S and S^{-1}, respectively (see Section 3.2.4.2). Notice that $M^\top V = I$ and hence $VM^\top$ is a projection (see equation (3.14)). Furthermore, we have $M = V$ if $S = S_1$. Consequently, (5.16) can be seen as a Petrov-Galerkin projection model with $A_r = M^\top AV$, $B_r = M^\top B$, $C_r = CV$ and $N_{i,r} = M^\top N_i V$ which is obtained by the state approximation $x(t) \approx Vx_r(t)$, as derived in equation (3.45) of Chapter 3. Inserting this approximation into (5.1) and subsequently multiplying the state equation with $M^\top$ to enforce the remainder term to be zero then results in (5.16).

5.2 Output Error Bound

In this section, we prove a bound for the error between (5.1) and (5.16). Below, we assume zero initial conditions, i.e., $x_0 = 0$ and $x_{0,r} = 0$. We begin with a general bound following the steps of [14, 96]. The solutions $x(t)$ and $x_r(t)$, $t \in [0, T]$, to (5.1) and (5.16)

can be expressed using their fundamental matrices $\Phi(t)$ and $\Phi_r(t)$, respectively, see [96]. Therefore, we have

$$x(t;0,u) = \int_0^t \Phi(t,s)Bu(s)\mathrm{d}s, \quad x_r(t;0,u) = \int_0^t \Phi_r(t,s)B_r u(s)\mathrm{d}s.$$

As discussed in Remark 3.12, we have $\Phi(t,s) = \Phi(t)\Phi^{-1}(s)$ and $\Phi_r(t,s) = \Phi_r(t)\Phi_r^{-1}(s)$. Consequently, representations for the outputs are

$$\begin{aligned} y(t) &= Cx(t;0,u) = C\int_0^t \Phi(t,s)Bu(s)\mathrm{d}s, \\ y_r(t) &= C_r x_r(t;0,u) = C_r\int_0^t \Phi_r(t,s)B_r u(s)\mathrm{d}s, \end{aligned} \tag{5.18}$$

where $t \in [0,T]$. Then, we find

$$\begin{aligned} \mathbb{E}\|y(t)-y_r(t)\|_2 &= \mathbb{E}\left\|C\int_0^t \Phi(t,s)Bu(s)\mathrm{d}s - C_r\int_0^t \Phi_r(t,s)B_r u(s)\mathrm{d}s\right\|_2 \\ &\le \mathbb{E}\int_0^t \left\| \left(C\Phi(t,s)B - C_r\Phi_r(t,s)B_r\right)u(s)\right\|_2 \mathrm{d}s \\ &\le \mathbb{E}\int_0^t \left\|C\Phi(t,s)B - C_r\Phi_r(t,s)B_r\right\|_F \|u(s)\|_2\mathrm{d}s. \end{aligned} \tag{5.19}$$

Here, $\|\cdot\|_F$ denotes the Frobenius norm (see Appendix A.1.2). Using Cauchy's inequality, it holds that

$$\begin{aligned} \mathbb{E}\|y(t)-y_r(t)\|_2 &\le \left(\mathbb{E}\int_0^t \left\|C\Phi(t,s)B - C_r\Phi_r(t,s)B_r\right\|_F^2\mathrm{d}s\right)^{\frac{1}{2}} \left(\mathbb{E}\int_0^t \|u(s)\|_2^2\mathrm{d}s\right)^{\frac{1}{2}} \\ &= \left(\mathbb{E}\int_0^t \left\|C^e\Phi^e(t,s)B^e\right\|_F^2\mathrm{d}s\right)^{\frac{1}{2}} \left(\mathbb{E}\int_0^t \|u(s)\|_2^2\mathrm{d}s\right)^{\frac{1}{2}}, \end{aligned}$$

where $\Phi^e = \left[\begin{smallmatrix}\Phi & 0\\ 0 & \Phi_r\end{smallmatrix}\right]$ is the fundamental solution to the system with coefficients:

$$A^e = \left[\begin{smallmatrix}A & 0\\ 0 & A_r\end{smallmatrix}\right], \quad N_i^e = \left[\begin{smallmatrix}N_i & 0\\ 0 & N_{i,r}\end{smallmatrix}\right], \quad B^e = \left[\begin{smallmatrix}B\\ B_r\end{smallmatrix}\right], \quad C^e = \left[\begin{smallmatrix}C & -C_r\end{smallmatrix}\right].$$

Applying the arguments that are used in Remark 3.20, we know that

$$\mathbb{E}[\Phi^e(t,s)B^eB^{e\top}\Phi^{e\top}(t,s)] = \mathbb{E}[\Phi^e(t-s)B^eB^{e\top}\Phi^{e\top}(t-s)]. \tag{5.20}$$

For $t \in [0,T]$, the identity in (5.20) yields

$$\begin{aligned} \mathbb{E}\int_0^t \left\|C^e\Phi^e(t,s)B^e\right\|_F^2\mathrm{d}s &= \mathbb{E}\int_0^t \mathrm{tr}(C^e\Phi^e(t,s)B^eB^{e\top}\Phi^{e\top}(t,s)C^{e\top})\mathrm{d}s \\ &= \mathbb{E}\int_0^t \mathrm{tr}(C^e\Phi^e(s)B^eB^{e\top}\Phi^{e\top}(s)C^{e\top})\mathrm{d}s \\ &\le \mathrm{tr}\left(C^e\int_0^T F^e(s)\mathrm{d}s\, C^{e\top}\right) \end{aligned} \tag{5.21}$$

with $F^e(t) = \mathbb{E}\left[\Phi^e(t)B^eB^{e\top}\Phi^{e\top}(t)\right]$ exploiting Fubini's theorem as well as the fact that the trace and C^e are linear operators. Since $F(t) = \mathbb{E}\left[\Phi(t)BB^\top\Phi^\top(t)\right]$ is a stochastic representation for equation (5.8), see Section 5.1.1, F^e satisfies

$$\dot{F}^e(t) = A^eF^e(t) + F^e(t)A^{e\top} + \sum_{i,j=1}^{q} N_i^eF^e(t)N_j^{e\top}k_{ij}, \quad F^e(0) = B^eB^{e\top}, \tag{5.22}$$

using the same arguments. From (5.22), it can be seen that the left upper $n \times n$ block of F^e is F which solves (5.8). On the other hand, the right lower $r \times r$ block F_r and the right upper $n \times r$ block $\tilde{F}$ of F^e satisfy

$$\dot{F}_r(t) = A_rF_r(t) + F_r(t)A_r^\top + \sum_{i,j=1}^{q} N_{i,r}F_r(t)N_{j,r}^\top k_{ij}, \quad F_r(0) = B_rB_r^\top, \tag{5.23}$$

$$\dot{\tilde{F}}(t) = A\tilde{F}(t) + \tilde{F}(t)A_r^\top + \sum_{i,j=1}^{q} N_i\tilde{F}(t)N_{j,r}^\top k_{ij}, \quad \tilde{F}(0) = BB_r^\top, \tag{5.24}$$

with stochastic representations

$$F_r(t) = \mathbb{E}[\Phi_r(t)B_rB_r^\top\Phi_r^\top(t)], \quad \tilde{F}(t) = \mathbb{E}[\Phi(t)BB_r^\top\Phi_r^\top(t)]. \tag{5.25}$$

Consequently, using (5.21) with the partition $F^e = \left[\begin{smallmatrix} F & \tilde{F} \\ \tilde{F}^\top & F_r \end{smallmatrix}\right]$, we find

$$\mathbb{E}\int_0^t \left\|C^e\Phi^e(t,s)B^e\right\|_F^2 \mathrm{d}s \le \operatorname{tr}\left(CP_TC^\top\right) + \operatorname{tr}\left(C_r\bar{P}_TC_r^\top\right) - 2\operatorname{tr}\left(C\tilde{P}_TC_r^\top\right),$$

where $P_{T,r} = \int_0^T F_r(t)\mathrm{d}t$ and $\tilde{P}_T = \int_0^T \tilde{F}(t)\mathrm{d}t$ solve

$$F_r(T) - B_rB_r^\top = A_rP_{T,r} + P_{T,r}A_r^\top + \sum_{i,j=1}^{q} N_{i,r}P_{T,r}N_{j,r}^\top k_{ij}, \tag{5.26}$$

$$\tilde{F}(T) - BB_r^\top = A\tilde{P}_T + \tilde{P}_TA_r^\top + \sum_{i,j=1}^{q} N_i\tilde{P}_TN_{j,r}^\top k_{ij}. \tag{5.27}$$

Summing up, we obtain that

$$\sup_{t\in[0,T]} \mathbb{E}\|y(t) - y_r(t)\|_2 \le \left(\operatorname{tr}(CP_TC^\top) + \operatorname{tr}(C_rP_{T,r}C_r^\top) - 2\operatorname{tr}(C\tilde{P}_TC_r^\top)\right)^{\frac{1}{2}} \|u\|_T. \tag{5.28}$$

The bound in (5.28) is very useful in order to check for the quality of a reduced system. Since P_T has to be computed to obtain (5.16), the actual cost to determine the bound lies in solving the low-dimensional matrix equations (5.26) and (5.27). However, (5.28) is only an a-posteriori estimate which is computed after the reduced order model is derived. Therefore, we discuss the role of $\Sigma_{T,2} = \operatorname{diag}(\sigma_{T,r+1}, \ldots, \sigma_{T,n})$ which is either the matrix of neglected eigenvalues of P_T or HSVs of the system. $\Sigma_{T,2}$ is associated to the truncated state variables x_2 of (5.13), compare with (5.15). By (5.5) and (5.6), it is already known that such variables x_2 are less relevant if $\sigma_{T,r+1}, \ldots, \sigma_{T,n}$ are small. This makes the values

σ_i a good a-priori criterion for the choice of r. In the following, we want to investigate how the truncated values $\sigma_{T,r+1}, \dots, \sigma_{T,n}$ characterize the error of the approximation. For that reason, we prove an error bound depending on $\Sigma_{T,2}$. As we will see, $\Sigma_{T,2}$ is not the only factor having an impact on the bound that is structurally independent of whether we choose $S = S_1$ or $S = S_2$.

Theorem 5.3 Let y be the output of (5.1) and y_r be the one of (5.16). Suppose that $S = S_1, S_2$, where S_1 is the factor of the eigenvalue decomposition of the Gramian P_T and S_2 is the balancing transformation defined in (5.14). Using partition (5.15) of the realization $(A_S, B_S, C_S, N_{i,S})$, we have

$$\begin{aligned}\sup_{t\in[0,T]} \mathbb{E}\|y(t) - y_r(t)\|_2 \\ \leq \Bigg(\operatorname{tr}\left(\Sigma_{T,2}\left[C_2^\top C_2 + 2A_{12}^\top \tilde{Q}_2 + \sum_{i,j=1}^{q} N_{i,12}^\top \left(2\tilde{Q}\begin{pmatrix} N_{j,12} \\ N_{j,22}\end{pmatrix} - Q_r N_{j,12}\right) k_{ij}\right]\right) \\ + 2\operatorname{tr}\left(\tilde{Q}\begin{pmatrix}\tilde{F}_1 - F_{11} \\ \tilde{F}_2 - F_{21}\end{pmatrix}\right) + \operatorname{tr}\left(Q_r(F_{11} - F_r)\right)\Bigg)^{\frac{1}{2}} \|u\|_T,\end{aligned}$$

where Q_r and $\tilde{Q} = \begin{pmatrix}\tilde{Q}_1 & \tilde{Q}_2\end{pmatrix}$ and are the unique solutions to

$$A_r^\top Q_r + Q_r A_r + \sum_{i,j=1}^{q} N_{i,r}^\top Q_r N_{j,r} k_{ij} = -C_r^\top C_r, \tag{5.29}$$

$$A_r^\top \tilde{Q} + \tilde{Q} A_S + \sum_{i,j=1}^{q} N_{i,r}^\top \tilde{Q} N_{j,S} k_{ij} = -C_r^\top C_S. \tag{5.30}$$

Moreover, the above bound involves

$$F_S(T) := SF(T)S^\top = \begin{bmatrix} F_{11} & F_{12} \\ F_{21} & F_{22}\end{bmatrix} \quad \textit{and} \quad \tilde{F}_S(T) := S\tilde{F}(T) = \begin{bmatrix}\tilde{F}_1 \\ \tilde{F}_2\end{bmatrix},$$

where $F(T)$, $F_r = F_r(T)$ and $\tilde{F}(T)$ are the terminal values of (5.8), (5.23) and (5.24), respectively.

The terms in the bound of Theorem 5.3 that do not directly depend on $\Sigma_{T,2}$ are related to the covariance error of the dimension reduction at the terminal time T (with $u \equiv 0$). To see this, let V be the matrix introduced in (5.17). As explained below (5.17), the state of the reduced system (5.16) can be interpreted as an approximation of the original state in the subspace spanned by the columns of V. By the stochastic representations of $F(T)$, $\tilde{F}(T)$ and $F_r(T)$ (see (5.7) and (5.25)), we can view $F(T)$ and $F_r(T)$ as covariances of the original and reduced model at time T, whereas $\tilde{F}(T)$ describes the correlations between both systems. Let us now assume that

$$F(T) \approx \tilde{F}(T)V^\top, \tag{5.31}$$

$$F(T) \approx VF_r(T)V^\top, \tag{5.32}$$

i.e., the covariance at T is well-approximated in the reduced system. This is, e.g., given if the uncontrolled state is well-approximated in the range of V at time T, i.e.,

$\Phi(T)B \approx V\Phi_r(T)B_r$. Now, multiplying (5.31) with S from the left and with M (defined in (5.17)) from the right, we obtain that $\begin{bmatrix} \tilde{F}_1 - F_{11} \\ \tilde{F}_2 - F_{21} \end{bmatrix}$ is small. Multiplying (5.32) with $M^\top$ from the left and with M the right provides a low deviation between $F_{11} = M^\top F(T) M$ and F_r. Although we additionally have these terms related to the covariance error, looking at $\Sigma_{T,2}$ is still suitable for getting an intuition concerning the error and hence a first idea for the choice of r. This is because a small $\Sigma_{T,2}$ goes along with a small error between $\Phi(T)B$ and its approximation $V\Phi_r(T)B_r$ in the range of V. This observation can be made due to

$$\mathbb{E}\int_0^T \| \left(\Phi(t)B\right)^\top z_T\|_2^2 \mathrm{d}t = z_T^\top P_T z_T = 0,$$

where $z_T \in \ker P_T$. Since $t \mapsto \Phi(t)$ is $\mathbb{P}$-almost surely continuous, we have $\left(\Phi(t)B\right)^\top z_T = 0$ $\mathbb{P}$-almost surely for all $t \in [0,T]$. Choosing $t = T$, we therefore know that the columns of $\Phi(T)B$ are orthogonal to $\ker P_T$. Given that P_T is symmetric, its image and kernel are orthogonal complements of each other in the vector space. Therefore, the orthogonality of $\Phi(T)B$ to the kernel implies that $\Phi(T)B \in \operatorname{im} P_T$. Hence, there is a matrix Z_T such that

$$\Phi(T)B = P_T Z_T = S^{-1}\Sigma_T S^{-\top} Z_T = \begin{pmatrix} V & \star \end{pmatrix} \begin{pmatrix} \Sigma_{T,1} & \\ & \Sigma_{T,2} \end{pmatrix} \begin{pmatrix} V^\top \\ \star \end{pmatrix} Z_T \approx V\Sigma_{T,1}V^\top Z_T,$$

i.e., the columns of $\Phi(T)B$ lie almost in the span of V if $\Sigma_{T,2}$ is small. Therefore, a good approximation can be expected if one truncates states with associated small values $\sigma_{T,r+1}, \dots, \sigma_{T,n}$. This can be confirmed by computing the representation in (5.28) after a reduced order dimension r was chosen based on the values $\sigma_{T,i}$. We prove the error bound in the following:

Proof of Theorem 5.3. Since $S = S_1, S_2$ diagonalizes P_T, we have

$$A_S\Sigma_T + \Sigma_T A_S^\top + \sum_{i,j=1}^{q} N_{i,S}\Sigma_T N_{j,S}^\top k_{ij} = -B_S B_S^\top + F_S(T). \tag{5.33}$$

We set $\tilde{Y}_T := S\tilde{P}_T$ and obtain the corresponding equation by multiplying (5.27) with S from the left resulting in

$$A_S\tilde{Y}_T + \tilde{Y}_T A_r^\top + \sum_{i,j=1}^{q} N_{i,S}\tilde{Y}_T N_{j,r}^\top k_{ij} = -B_S B_r^\top + \tilde{F}_S(T). \tag{5.34}$$

Now, we analyze the trace expression $\epsilon^2 := \left(\operatorname{tr}(CP_TC^\top) + \operatorname{tr}(C_rP_{T,r}C_r^\top) - 2\operatorname{tr}(C\tilde{P}_TC_r^\top)\right)$ in (5.28). We see that

$$\begin{aligned} \epsilon^2 &= \left(\operatorname{tr}(C_S\Sigma_T C_S^\top) + \operatorname{tr}(C_rP_{T,r}C_r^\top) - 2\operatorname{tr}(C_S\tilde{Y}_TC_r^\top)\right) \\ &= \left(\operatorname{tr}(C_r\Sigma_{T,1}C_r^\top) + \operatorname{tr}(C_2\Sigma_{T,2}C_2^\top) + \operatorname{tr}(C_rP_{T,r}C_r^\top) - 2\operatorname{tr}(C_S\tilde{Y}_TC_r^\top)\right). \end{aligned} \tag{5.35}$$

Exploiting (5.30) yields

$$\begin{aligned}-\operatorname{tr}(C_S\tilde{Y}_T C_r^\top) &= -\operatorname{tr}(\tilde{Y}_T C_r^\top C_S) = \operatorname{tr}\left(\tilde{Y}_T\left[A_r^\top\tilde{Q} + \tilde{Q}A_S + \sum_{i,j=1}^{q} N_{i,r}^\top\tilde{Q}N_{j,S}k_{ij}\right]\right)\\ &= \operatorname{tr}\left(\tilde{Q}\left[A_S\tilde{Y}_T + \tilde{Y}_T A_r^\top + \sum_{i,j=1}^{q} N_{i,S}\tilde{Y}_T N_{j,r}^\top k_{ij}\right]\right).\end{aligned}$$

Comparing (5.30) and (5.34), we find that

$$-\operatorname{tr}(C_S\tilde{Y}_T C_r^\top) = -\operatorname{tr}(\tilde{Q}B_S B_r^\top) + \operatorname{tr}(\tilde{Q}\tilde{F}_S(T)). \tag{5.36}$$

Using the partition in (5.15), the first r columns of (5.33) are

$$\begin{aligned}&\begin{pmatrix}A_r\\A_{21}\end{pmatrix}\Sigma_{T,1} + \begin{pmatrix}\Sigma_{T,1}A_r^\top\\\Sigma_{T,2}A_{12}^\top\end{pmatrix} + \sum_{i,j=1}^{q}\left(\begin{pmatrix}N_{i,r}\\N_{i,21}\end{pmatrix}\Sigma_{T,1}N_{j,r}^\top + \begin{pmatrix}N_{i,12}\\N_{i,22}\end{pmatrix}\Sigma_{T,2}N_{j,12}^\top\right)k_{ij}\\ &= -B_S B_r^\top + \begin{pmatrix}F_{11}\\F_{21}\end{pmatrix}.\end{aligned} \tag{5.37}$$

We insert (5.37) into (5.36) and obtain

$$\begin{aligned}-\operatorname{tr}(C_S\tilde{Y}_T C_r^\top) &= \operatorname{tr}\left(\tilde{Q}\begin{pmatrix}\tilde{F}_1 - F_{11}\\\tilde{F}_2 - F_{21}\end{pmatrix}\right)\\ &+ \operatorname{tr}\left(\tilde{Q}\left[\begin{pmatrix}A_r\\A_{21}\end{pmatrix}\Sigma_{T,1} + \begin{pmatrix}\Sigma_{T,1}A_r^\top\\\Sigma_{T,2}A_{12}^\top\end{pmatrix}\right.\right.\\ &\left.\left.+ \sum_{i,j=1}^{q} k_{ij}\left(\begin{pmatrix}N_{i,r}\\N_{i,21}\end{pmatrix}\Sigma_{T,1}N_{j,r}^\top + \begin{pmatrix}N_{i,12}\\N_{i,22}\end{pmatrix}\Sigma_{T,2}N_{j,12}^\top\right)\right]\right)\\ &= \operatorname{tr}\left(\tilde{Q}\begin{pmatrix}\tilde{F}_1 - F_{11}\\\tilde{F}_2 - F_{21}\end{pmatrix}\right) + \operatorname{tr}\left(\Sigma_{T,2}\left[A_{12}^\top\tilde{Q}_2 + \sum_{i,j=1}^{q} N_{i,12}^\top\tilde{Q}\begin{pmatrix}N_{j,12}\\N_{j,22}\end{pmatrix}k_{ij}\right]\right)\\ &+ \operatorname{tr}\left(\Sigma_{T,1}\left[\tilde{Q}\begin{pmatrix}A_r\\A_{21}\end{pmatrix} + A_r^\top\tilde{Q}_1 + \sum_{i,j=1}^{q} N_{i,r}^\top\tilde{Q}\begin{pmatrix}N_{j,r}\\N_{j,21}\end{pmatrix}k_{ij}\right]\right).\end{aligned}$$

Using the partition of the balanced realization in (5.15), we observe that the last term of the above equation is the first r columns of (5.30). So, we can say

$$\begin{aligned}-\operatorname{tr}(C_S\tilde{Y}_T C_r^\top) &= \operatorname{tr}\left(\tilde{Q}\begin{pmatrix}\tilde{F}_1 - F_{11}\\\tilde{F}_2 - F_{21}\end{pmatrix}\right) + \operatorname{tr}\left(\Sigma_{T,2}\left[A_{12}^\top\tilde{Q}_2 + \sum_{i,j=1}^{q} N_{i,12}^\top\tilde{Q}\begin{pmatrix}N_{j,12}\\N_{j,22}\end{pmatrix}k_{ij}\right]\right)\\ &- \operatorname{tr}(\Sigma_{T,1}C_r^\top C_r).\end{aligned} \tag{5.38}$$

Inserting (5.38) into (5.35), we have

$$\begin{aligned}\epsilon^2 = \operatorname{tr}\left(\Sigma_{T,2}\left[C_2^\top C_2 + 2A_{12}^\top \tilde{Q}_2 + 2\sum_{i,j=1}^{q} N_{i,12}^\top \tilde{Q}\begin{pmatrix} N_{j,12} \\ N_{j,22}\end{pmatrix} k_{ij}\right]\right) \\ + 2\operatorname{tr}\left(\tilde{Q}\begin{pmatrix}\tilde{F}_1 - F_{11} \\ \tilde{F}_2 - F_{21}\end{pmatrix}\right) + \operatorname{tr}\left((P_{T,r} - \Sigma_{T,1})C_r^\top C_r\right).\end{aligned} \tag{5.39}$$

Equation (5.29) now yields

$$\begin{aligned}&\operatorname{tr}\left((P_{T,r} - \Sigma_{T,1})C_r^\top C_r\right) \\ &= -\operatorname{tr}\left(Q_r\left[A_r(P_{T,r} - \Sigma_{T,1}) + (P_{T,r} - \Sigma_{T,1})A_r^\top + \sum_{i,j=1}^{q} N_{i,r}(P_{T,r} - \Sigma_{T,1})N_{j,r}^\top k_{ij}\right)\right]\end{aligned}$$

The combination of (5.26) and the left upper block of (5.33) gives

$$\begin{aligned}&A_r(P_{T,r} - \Sigma_{T,1}) + (P_{T,r} - \Sigma_{T,1})A_r^\top + \sum_{i,j=1}^{q} N_{i,r}(P_{T,r} - \Sigma_{T,1})N_{j,r}^\top k_{ij} \\ &= \sum_{i,j=1}^{q} N_{i,12}\Sigma_{T,2}N_{j,12}^\top k_{ij} + (F_r - F_{11}).\end{aligned}$$

Consequently, we have

$$\operatorname{tr}\left((P_{T,r} - \Sigma_{T,1})C_r^\top C_r\right) = -\operatorname{tr}\left(\Sigma_{T,2}\left[\sum_{i,j=1}^{q} N_{i,12}^\top Q_r N_{j,12} k_{ij}\right]\right) + \operatorname{tr}\left(Q_r(F_{11} - F_r)\right).$$

So, we obtain that

$$\begin{aligned}\epsilon^2 = \operatorname{tr}\left(\Sigma_{T,2}\left[C_2^\top C_2 + 2A_{12}^\top \tilde{Q}_2 + \sum_{i,j=1}^{q} N_{i,12}^\top\left(2\tilde{Q}\begin{pmatrix} N_{j,12} \\ N_{j,22}\end{pmatrix} - Q_r N_{j,12}\right)k_{ij}\right]\right) \\ + 2\operatorname{tr}\left(\tilde{Q}\begin{pmatrix}\tilde{F}_1 - F_{11} \\ \tilde{F}_2 - F_{21}\end{pmatrix}\right) + \operatorname{tr}\left(Q_r(F_{11} - F_r)\right),\end{aligned}$$

which concludes the proof of this theorem. □

Notice that the estimate in Theorem 5.3 is also beneficial if $N_i = 0$ for all $i = 1, \ldots, q$, since it improves the deterministic bound [102] in the sense that we can generally deduce the relation between the truncated HSVs and the actual approximation error here. It is important to note that, in the deterministic case, "improvement" is not meant in terms of accuracy. The error bound representation in [102] just has the drawback that it allows making similar conclusions only if the underlying system is asymptotically stable. Moreover, the result of Theorem 5.3 is a generalization of the bounds for mean square asymptotically stable stochastic systems [14, 103], where the covariance related terms vanish as $T \to \infty$.

5.3 Computation of Gramians

In this section, we discuss how to compute P_T and Q_T which allow us to identify redundant information in the system. These matrices are solutions to Lyapunov equations (5.9) and (5.12) with left hand sides depending on $F(T)$ and $G(T)$, respectively. Given $F(T)$ and $G(T)$ it is therefore required to solve generalized Lyapunov equations

$$L(X) = \mathcal{L}_A(X) + \Pi_N(X), \tag{5.40}$$

efficiently, where L is some matrix of suitable dimension. According to Remark 3.18 in Chapter 3 this can be done by vectorization, i.e., one can try to solve $\operatorname{vec}(L) = \mathcal{K}\operatorname{vec}(X)$ with the Kronecker matrix $\mathcal{K}$ defined in (5.11). Since $\mathcal{K}$ is of order n^2, the complexity of deriving $\operatorname{vec}(X)$ from this linear system of equations is $\mathcal{O}(n^6)$ making this procedure infeasible for $n \gg 100$.

However, more efficient techniques have been developed in order to solve (5.40), see, e.g. [26], where a sequence of standard Lyapunov equations ($\Pi_N = 0$) are solved to find X. Such standard Lyapunov equations can either be tackled by direct methods, such as Bartels-Stewart [7], which cost $\mathcal{O}(n^3)$ operations, or by iterative methods such as ADI or Krylov subspace methods [111], which have a much smaller complexity than the Bartels-Stewart algorithm, in particular, when the left hand side is of low rank or structured (complexity of $\mathcal{O}(n^2)$ or less).

Solving for P_T and Q_T now relies on having access to $F(T)$ and $G(T)$ which are the terminal values of the matrix-differential equations (5.8) and (5.10). The remainder of this section will deal with strategies to compute these terminal values.

5.3.1 Exact Methods

One solution to overcome the issue of unknown $F(T)$ and $G(T)$ is to use vectorizations of (5.8) and (5.10) for dimensions n of a few hundreds. If we define $f(t) := \operatorname{vec}(F(t))$ and $g(t) = \operatorname{vec}(G(t))$, then

$$\dot{f}(t) = \mathcal{K}f(t), \quad f(0) = \operatorname{vec}(BB^\top), \quad \dot{g}(t) = \mathcal{K}^\top g(t), \quad g(0) = \operatorname{vec}(C^\top C),$$

where $\mathcal{K}$ is defined in (5.11). Therefore, obtaining $F(T)$ and $G(T)$ rely on the efficient computation of a matrix exponential, since

$$f(T) = \mathrm{e}^{\mathcal{K}T}\operatorname{vec}(BB^\top), \quad g(T) = \mathrm{e}^{\mathcal{K}^\top T}\operatorname{vec}(C^\top C).$$

One can find a discussion on how to determine a matrix exponential efficiently in [64] and references therein. Alternatively, one might think of discretizing the matrix differential equations (5.8) and (5.10) to find an approximation of $F(T)$ and $G(T)$. However, as stated above, these equations are equivalent to ordinary differential equations of order n^2. Solving such extremely large scale systems is usually not feasible. In addition, only implicit schemes would allow for a reasonable step size in the discretization making the problem even more complex. For that reason, we discuss more suitable numerical approximations in the following.

5.3.2 Sampling Based Approaches

We aim to derive an approximation of the terminal value $F(T) = \mathbb{E}[\Phi(T)BB^\top\Phi^\top(T)]$ of (5.8) by different stochastic representations. This alternative approach is required since computing $\mathrm{e}^{\mathcal{K}T}$ is not feasible if $n \gg 100$ knowing that $\mathcal{K} \in \mathbb{R}^{n^2\times n^2}$. Therefore, we discuss sampling-based approaches in the following. Let $\Phi^i(T)$, $i \in \{1,\dots,M\}$, be i.i.d. copies of $\Phi(T)$. Then, we have $\frac{1}{M}\sum_{i=1}^{M}\Phi^i(T)BB^\top\Phi^i(T)^\top \approx F(T)$ if M is sufficiently large. This requires to sample the random variable $\Phi(T)B$ possibly many times. $\Phi(T)B$ is the terminal value of the stochastic differential equation

$$\mathrm{d}x_B(t) = Ax_B(t)\mathrm{d}t + \sum_{i=1}^{q} N_i x_B(t)\mathrm{d}W_i(t), \quad x_B(0) = B, \tag{5.41}$$

where $x_b(t) \in \mathbb{R}^{n\times m}$. System (5.41) can be seen as a homogeneous version of (5.1a) since the control dependence is gone[3]. If (5.1) needs to be evaluated for many different controls u and additionally a large number of samples are required for each fixed u, it even pays off to generate many samples of the solution to (5.41). In particular, this is true if the number of columns of B is low. However, we want to avoid evaluating (5.41) too often. The number of samples M required for a good estimate of $F(T)$ depends on the variance of $\Phi(T)BB^\top\Phi^\top(T)$. Therefore, we want to reduce the variance by finding a better stochastic representation than $\mathbb{E}[\Phi(T)BB^\top\Phi^\top(T)]$. In the spirit of variance reduction techniques, we first find the zero-variance unbiased estimator. To do so, we apply Ito's product rule (see Example 1.27) in order to obtain

$$\begin{aligned}
d\left(x_B(t)x_B^\top(t)\right) &= d\left(x_B(t)\right)x_B^\top(t) + x_B(t)d\left(x_B^\top(t)\right) + d\left(x_B(t)\right)d\left(x_B^\top(t)\right)\\
&= \left(Ax_B(t)\mathrm{d}t + \sum_{i=1}^{q} N_i x_B(t)\mathrm{d}W_i(t)\right)x_B^\top(t)\\
&\quad + x_B(t)\left(x_B^\top(t)A^\top\mathrm{d}t + \sum_{i=1}^{q} x_B^\top(t)N_i^\top\mathrm{d}W_i(t)\right)\\
&\quad + \sum_{i,j=1}^{q} N_i x_B(t)x_B^\top(t)N_j^\top k_{ij}\\
&= (\mathcal{L}_A + \Pi_N)\left(x_B(t)x_B^\top(t)\right)dt + \sum_{i=1}^{q}\mathcal{L}_{N_i}\left(x_B(t)x_B^\top(t)\right)\mathrm{d}W_i(t).
\end{aligned}$$

This stochastic differential is now exploited to find

$$\begin{aligned}
d\left(\mathrm{e}^{\mathcal{K}(T-t)}\operatorname{vec}(x_B(t)x_B^\top(t))\right) &= -\mathrm{e}^{\mathcal{K}(T-t)}\,\mathcal{K}\operatorname{vec}(x_B(t)x_B^\top(t))\mathrm{d}t\\
&\quad + \mathrm{e}^{\mathcal{K}(T-t)}\,d\left(\operatorname{vec}(x_B(t)x_B^\top(t))\right)\\
&= \sum_{i=1}^{q}\mathrm{e}^{\mathcal{K}(T-t)}\operatorname{vec}\left(\mathcal{L}_{N_i}\left(x_B(t)x_B^\top(t)\right)\right)\mathrm{d}W_i(t)
\end{aligned}$$

[3] See equation (3.21)

using that $\operatorname{vec}\left(\left(\mathcal{L}_A + \Pi_N\right)\left(x_B(t)x_B^\top(t)\right)\right) = \mathcal{K}\operatorname{vec}(x_B(t)x_B^\top(t))$. Hence, we have

$$\operatorname{vec}\left(x_B(T)x_B^\top(T)\right) = \mathrm{e}^{\mathcal{K}T}\operatorname{vec}(BB^\top) + \sum_{i=1}^q \int_0^T \mathrm{e}^{\mathcal{K}(T-t)}\operatorname{vec}\left(\mathcal{L}_{N_i}\left(x_B(t)x_B^\top(t)\right)\right)\mathrm{d}W_i(t).$$

Devectorizing this equation yields

$$F(T) = x_B(T)x_B^\top(T) - \sum_{i=1}^q \int_0^T F\left(T-t, \mathcal{L}_{N_i}\left(x_B(t)x_B^\top(t)\right)\right)\mathrm{d}W_i(t), \tag{5.42}$$

where the second argument in F represents the initial condition of (5.8). The right hand side of (5.42) now is an unbiased zero variance estimator of $F(T)$. However, this estimator depends on F which is not available. Therefore, given a symmetric matrix X_0, we approximate $F(t, X_0)$ by a computable matrix function $\mathcal{F}(t, X_0)$ that we specify later. This leads to the unbiased estimator

$$E_{\mathcal{F}}(T) := x_B(T)x_B^\top(T) - \sum_{i=1}^q \int_0^T \mathcal{F}\left(T-t, \mathcal{L}_{N_i}\left(x_B(t)x_B^\top(t)\right)\right)\mathrm{d}W_i(t), \tag{5.43}$$

for $F(T)$. The hope is that a few samples of $E_{\mathcal{F}}(T)$ can give an accurate approximation of $F(T)$. Of course, $E_{\mathcal{F}}(T)$ can only be simulated by further discretizing the above Ito integrals, e.g., by a Riemann-Stieltjes sum approximation. The variance of $E_{\mathcal{F}}(T)$ is

$$\begin{aligned}
&\mathbb{E}\left\|E_{\mathcal{F}}(T) - F(T)\right\|_F^2 = \mathbb{E}\left\|\sum_{i=1}^q \int_0^T F\left(T-t, X_i(t)\right) - \mathcal{F}\left(T-t, X_i(t)\right)\mathrm{d}W_i(t)\right\|_F^2 \\
&= \sum_{i,j=1}^q \mathbb{E}\int_0^T \Big\langle F\left(T-t, X_i(t)\right) - \mathcal{F}\left(T-t, X_i(t)\right), \\
&\qquad\qquad F\left(T-t, X_j(t)\right) - \mathcal{F}\left(T-t, X_j(t)\right)\Big\rangle_F k_{ij}\mathrm{d}t
\end{aligned}$$

setting $X_i(t) = N_i x_B(t)x_B^\top(t) + x_B(t)x_B^\top(t)N_i^\top$ and exploiting Ito's isometry (see Lemma 1.19). Consequently, the benefit of the variance reduction depends on the difference $F(t, X_0) - \mathcal{F}(t, X_0)$.

We conclude this section by discussing suitable approximations $\mathcal{F}(t, X_0)$ of $F(t, X_0)$. For that reason, we establish the following theorem.

Theorem 5.4 Let $F(t, X_0)$, $t \in [0, T]$, be the solution to

$$\dot{F}(t) = \mathcal{L}_A\left(F(t)\right) + \Pi_N\left(F(t)\right), \quad F(0) = X_0,$$

where the initial data X_0 is a symmetric matrix. Then, there exist constants $\underline{c}$ and $\overline{c}$ such that

$$\mathrm{e}^{At}X_0\,\mathrm{e}^{A^\top t} + \underline{c}\int_0^t \mathrm{e}^{As}\,\Pi_N(I)\,\mathrm{e}^{A^\top s}\,\mathrm{d}s \le F(t) \le \mathrm{e}^{At}X_0\,\mathrm{e}^{A^\top t} + \overline{c}\int_0^t \mathrm{e}^{As}\,\Pi_N(I)\,\mathrm{e}^{A^\top s}\,\mathrm{d}s.$$

Proof. Exploiting the product rule, it can be seen that F is implicitly given by

$$F(t) = \mathrm{e}^{At} X_0 \mathrm{e}^{A^\top t} + \int_0^t \mathrm{e}^{A(t-s)} \Pi_N(F(s)) \mathrm{e}^{A^\top (t-s)} \, \mathrm{d}s. \tag{5.44}$$

The solution $t \mapsto F(t)$ is continuous and $F(t)$ is a symmetric matrix for all $t \in [0,T]$. Consequently, exploiting (see Corollary B.8), there exist continuous and real functions $\lambda_1, \dots, \lambda_n$ such that $\lambda_1(t), \dots, \lambda_n(t)$ represent the eigenvalues of $F(t)$ for each fixed t. We now define continuous functions by $\underline{\lambda} := \min\{\lambda_1, \dots, \lambda_n\}$ and $\overline{\lambda} := \max\{\lambda_1, \dots, \lambda_n\}$. Symmetric matrices can be estimated from below and above by their smallest and largest eigenvalue, respectively, leading to $\underline{\lambda}(t) I \leq F(t) \leq \overline{\lambda}(t) I$. Therefore, given an arbitrary vector in $v \in \mathbb{R}^n$, we have

$$\begin{aligned}
v^\top \Pi_N(F(t)) v &= \sum_{i,j=1}^q (N_i v)^\top F(t) N_j v k_{ij} = \sum_{i,j=1}^q (N_i v)^\top F(t) N_j v e_i^\top \mathbf{K}^{\frac{1}{2}} \mathbf{K}^{\frac{1}{2}} e_j \\
&= \sum_{i,j=1}^q (N_i v)^\top F(t) N_j v \sum_{k=1}^q \langle \mathbf{K}^{\frac{1}{2}} e_i, e_k \rangle_2 \langle \mathbf{K}^{\frac{1}{2}} e_j, e_k \rangle_2 \\
&= \sum_{k=1}^q \left(\sum_{i=1}^q N_i v \langle \mathbf{K}^{\frac{1}{2}} e_i, e_k \rangle_2 \right)^\top F(t) \Bigg(\underbrace{\sum_{j=1}^q N_j v \langle \mathbf{K}^{\frac{1}{2}} e_j, e_k \rangle_2}_{=: v_k} \Bigg) \\
&\begin{cases} \leq \overline{\lambda}(t) \sum_{k=1}^q v_k^\top I v_k \\ \geq \underline{\lambda}(t) \sum_{k=1}^q v_k^\top I v_k \end{cases}
\end{aligned}$$

resulting in $\underline{\lambda}(t) \Pi_N(I) \leq \Pi_N(F(t)) \leq \overline{\lambda}(t) \Pi_N(I)$, where e_i is the canonical basis of $\mathbb{R}^q$. Since $\underline{\lambda}, \overline{\lambda}$ are continuous on $[0,T]$, they can be bounded from below and above by some suitable constants. Applying this to (5.44), we obtain the result by substitution. □

Of course, the constants in Theorem 5.4 are generally unknown. However, this result gives us the intuition that $F(t, X_0)$ can be approximated by

$$\mathcal{F}(t, X_0) = \mathrm{e}^{At} X_0 \mathrm{e}^{A^\top t} + c \int_0^t \mathrm{e}^{As} \Pi_N(I) \mathrm{e}^{A^\top s} \, \mathrm{d}s, \tag{5.45}$$

where $c \in [\underline{c}, \bar{c}]$ is a real number. From the proof of Theorem 5.4, we further know that $\underline{c}, \bar{c} \geq 0$ if X_0 is positive semidefinite. We cannot generally expect a reduction of the variance for all choices of c. However, a good candidate will reduce the computational complexity. A general strategy how to find such a candidate is an interesting question for future research.

Remark 5.5 Besides generating (a few) samples of x_B from (5.41), we require the matrix exponentials e^{At_i} an a grid $0 = t_0 < t_1 < \ldots < t_{n_k} = T$ to determine the estimator (5.43) with $\mathcal{F}$ as in (5.45). Here, n_g is the number of grid points when discretizing the Itô integral in (5.43). If the points t_i are equidistant with step size h, one first computes e^{Ah}. The other exponentials are then powers of e^{Ah} such that a certain number of matrix multiplications (depending on n_g) have to be conducted.

The Gramian Q_T can be computed from (5.12) requiring to determine $G(T)$. According to Remark 5.2, we know that $G(T) = \mathbb{E}[x_C(T)x_C^\top(T)]$, where

$$dx_C(t) = A^\top x_C(t)\mathrm{d}t + \sum_{i=1}^{q} N_i^\top x_C(t)\mathrm{d}W_i(t), \quad x_C(0) = C^\top,$$

with $x_C(t) \in \mathbb{R}^{n\times p}$. Exploiting the above consideration regarding $F(T)$, we can see that

$$E_{\mathcal{G}}(T) := x_C(T)x_C^\top(T) - \sum_{i=1}^{q}\int_0^T \mathcal{G}\left(T-t, \mathcal{L}_{N_i}^*\left(x_C(t)x_C^\top(t)\right)\right)\mathrm{d}W_i(t) \tag{5.46}$$

is a possible unbiased estimator for $G(T)$. The approximation $\mathcal{G}$ of G can be chosen as in (5.45) replacing $(A, N_i) \mapsto (A^\top, N_i^\top)$.

5.3.3 Gramians Based on Deterministic Approximations of $F(T)$ and $G(T)$

Based on Theorem 5.4, an estimation of $F(T)$ (and also $G(T)$) is given in (5.45). Instead of using these approximations in a variance reduction procedure like in Section 5.3.2, we exploit it directly in (5.9) and (5.12). This leads to matrices $\mathcal{P}_T$ and $\mathcal{Q}_T$ solving

$$\begin{aligned}\mathcal{F}(T, BB^\top) - BB^\top &= \mathcal{L}_A(\mathcal{P}_T) + \Pi_N(\mathcal{P}_T),\\ \mathcal{G}(T, C^\top C) - C^\top C &= \mathcal{L}_A^*(\mathcal{Q}_T) + \Pi_N^*(\mathcal{Q}_T),\end{aligned}$$

where the left-hand sides are defined by

$$\mathcal{F}(T, BB^\top) = \mathrm{e}^{AT}BB^\top\mathrm{e}^{A^\top T} + c_F\int_0^T \mathrm{e}^{As}\Pi_N(I)\mathrm{e}^{A^\top s}\,\mathrm{d}s, \quad c_F \in \mathbb{R}, \tag{5.47}$$

$$\mathcal{G}(T, C^\top C) = \mathrm{e}^{A^\top T}C^\top C\,\mathrm{e}^{AT} + c_G\int_0^T \mathrm{e}^{A^\top s}\Pi_N^*(I)\mathrm{e}^{As}\,\mathrm{d}s, \quad c_G \in \mathbb{R}. \tag{5.48}$$

Certainly, the choice of the constants c_F and c_G determine how well P_T and Q_T are approximated by $\mathcal{P}_T$ and $\mathcal{Q}_T$, e.g., in terms of the characterization of the respective dominant subspaces of system (5.1). Notice that for $N_i = 0$, $\mathcal{F}(T, BB^\top)$ and $\mathcal{G}(T, C^\top C)$ yield the exact values for $F(T, BB^\top)$ and $G(T, C^\top C)$. At this point, it is important to mention that the Gramian approximation of this section is computationally less complex than the one in Section 5.3.2. First of all, we do not need to sample from (5.41) and secondly, no Itô integral as in (5.43) has to be discretized. Calculating $\mathcal{F}$ and $\mathcal{G}$ might also require to compote matrix exponentials on a partition of $[0, T]$, company with Remark 5.5. However, fewer grid points than for the sampled Gramians of Section 5.3.2have to be considered since an ordinary integral can be discretized with a larger step size compared to an Ito integral. Alternatively, the integrals in (5.47) and (5.48) can also be determined without a discretization since it holds that

$$\mathcal{L}_A\left(\int_0^T \mathrm{e}^{As}\Pi_N(I)\mathrm{e}^{A^\top s}ds\right) = -\Pi_N(I) + \mathrm{e}^{AT}\Pi_N(I)\mathrm{e}^{A^\top T}$$
$$\mathcal{L}_A^*\left(\int_0^T \mathrm{e}^{A^\top s}\Pi_N^*(I)\mathrm{e}^{As}ds\right) = -\Pi_N^*(I) + \mathrm{e}^{A^\top T}\Pi_N^*(I)\mathrm{e}^{AT}$$

This approach has the advantage that only the matrix exponential e^{AT} at the terminal time is needed.

5.4 Numerical Experiments

In order to indicate the benefit of the model reduction method presented in Section 5.1, we consider a linear controlled SPDE as follow

$$\frac{\partial \mathcal{X}(t,\zeta)}{\partial t} = \Delta\mathcal{X}(t,\zeta) + \mathcal{B}u(t) + \sum_{i=1}^{q}\mathcal{N}_i\mathcal{X}(t,\zeta)\frac{\partial W_i(t)}{\partial t}. \tag{5.49}$$

In addition, we emphasize the applicability to unstable systems by rescaling and shifting the Laplacian. The concrete example of interest is (for $t \in [0,1]$ and $\zeta \in [0,\pi]^2$)

$$\begin{aligned}\frac{\partial \mathcal{X}(t,\zeta)}{\partial t} &= (\alpha\Delta + \beta I)\,\mathcal{X}(t,\zeta) + 1_{[\frac{\pi}{4},\frac{3\pi}{4}]^2}(\zeta)u(t) + \gamma\,\mathrm{e}^{-|\zeta_1-\frac{\pi}{2}|-\zeta_2}\,\mathcal{X}(t,\zeta)\frac{\partial W(t)}{\partial t},\\ \mathcal{X}(t,\zeta) &= 0, \quad t\in[0,1], \quad \zeta\in\partial[0,\pi]^2, \quad \text{and} \quad \mathcal{X}(0,\zeta)\equiv 0,\end{aligned}$$

where $\alpha,\beta > 0$, $\gamma \in \mathbb{R}$ and W is an one-dimensional Wiener process. $\mathcal{X}(t,\cdot)$, $t\in[0,T]$, is interpreted as a process taking values in $H = L^2([0,\pi]^2)$. The input operator $\mathcal{B}$ in (5.49) is characterized by $1_{[\frac{\pi}{4},\frac{3\pi}{4}]^2}(\cdot)$ and the noise operator $\mathcal{N}_1 = \mathcal{N}$ is defined trough $\mathcal{N}\mathcal{X} = \mathrm{e}^{-|\cdot-\frac{\pi}{2}|-\cdot}\,\mathcal{X}$ for $\mathcal{X}\in L^2([0,\pi]^2)$. Since the Dirichlet Laplacian generates a C_0-semigroup and its eigenfunctions $(h_k)_{k\in\mathbb{N}}$ represent a basis of H, the same is true for $\alpha\Delta+\beta I$. Therefore, we interpret the solution of the above SPDE in the mild sense. For more information on SPDEs and the concept of mild solution, we refer to [24]. The quantity of interest is the average temperature in the noncontrolled area, i.e.,

$$\mathcal{Y}(t) = \mathcal{C}\mathcal{X}(t,\cdot) := \frac{4}{3\pi^2}\int_{[0,\pi]^2\setminus[\frac{\pi}{4},\frac{3\pi}{4}]^2}\mathcal{X}(t,\zeta)d\zeta.$$

In order to solve this SPDE numerically, a spatial discretization can be considered as a first step. Here, we choose a spectral Galerkin method that is based on the global basis of eigenfunctions $(h_k)_{k\in\mathbb{N}}$. The idea is to construct an approximation $\mathcal{X}_n$ to $\mathcal{X}$ taking values in the subspace $H_n = \operatorname{span}\{h_1,\dots,h_n\}$ and which converges to the SPDE solution with $n\to\infty$. For more detailed information on this discretization scheme, we refer to [40]. The vector of Fourier coefficients $x(t) = (\langle\mathcal{X}_n(t),h_1\rangle_H,\dots,\langle\mathcal{X}_n(t),h_n\rangle_H)^\top$ is a solution of a system like (5.1) with $q=1$ and discretized operators

- $A = \alpha\operatorname{diag}(-\lambda_1,\dots,-\lambda_n) + \beta I$, $B = (\langle\mathcal{B},h_k\rangle_H)_{k=1\dots n}$, $C = (\mathcal{C}h_k)_{k=1\dots n}$,
- $N_1 = (\langle\mathcal{N}h_i,h_k\rangle_H)_{k,i=1\dots n}$ and $x_0 = 0$,

where $(-\lambda_k)_{k\in\mathbb{N}}$ are the ordered eigenvalues of Δ. We refer to Chapter 4, where a similar example with fBm noise was studied, providing further details on the derivation of this system and its associated matrices. For the case of Wiener noise, we refer to [14]. Now, a small α and a larger β yield an unstable A, i.e., $\sigma(A) \not\subset \mathbb{C}_-$ which already violates asymptotic mean square stability of (5.1), i.e., $\mathbb{E}\,\|x(t;x_0,0)\|_2^2 \nrightarrow 0$ as $t \to \infty$. Moreover, a larger γ (larger noise) causes further instabilities. For that reason, we pick $\alpha = 0.4$, $\beta = 3$ and $\gamma = 2$ in order to demonstrate the MOR procedure for a relatively unstable system. Notice that enlarging β or γ (or making α smaller) leads to a higher degree of instability. This affects the approximation quality in the reduced system given T is fixed. The intuition is that the less stable a system is the stronger the dominant subspaces are expanding in time. This is because some variables in unstable systems are strongly growing such that initially redundant directions become more relevant from a certain point of time. This can also be observed in numerical experiments.

In the following, we fix a normalized control $u(t) = c_u\,\mathrm{e}^{-0.1t}$, $t \in [0,T]$, (the constant c_u ensures $\|u\|_T = 1$) and apply the MOR method to the spatially discretized SPDE that is based on the balancing transformation $S = S_2$ described in Section 5.1.2. In Section 5.4.1, we compare the approximation quality of the ROMs using either the exact Gramian or inexact Gramians introduced in Section 5.3. Subsequently, Section 5.4.2 shows the reduced model accuracy in higher state space dimension, where solely inexact Gramians are available. We conclude the numerical experiments by discussing the impact of the terminal time T and the covariance matrix K in Section 5.4.3.

5.4.1 Simulations for $n = 100$ and $T = 1$

We compare the associated ROM (5.16) with the original system in dimension $n = 100$ first since this choice allows to determine $F(T), G(T)$ and hence the Gramians P_T, Q_T exactly according to Section 5.3.1. As a consequence, we can compare the MOR scheme involving the exact Gramians with the same type of scheme relying on the approximated Gramians that are computed exploiting the approaches in Sections 5.3.2 and 5.3.3. In particular, we first approximate $F(T)$ and $G(T)$ based on a Monte-Carlo simulation using 10 realizations of the estimators (5.43) and (5.46), respectively. The functions $\mathcal{F}$ and $\mathcal{G}$ entering these estimators are chosen as in (5.45) with $c = 0$. We refer to the resulting matrices as the Section 5.3.2 Gramians. At this point, we want to emphasize that these sampling based Gramians do not necessarily have to be accurate approximations of the exact Gramians in a component-wise sense. It is more important that the dominant subspaces of the system (eigenspaces of the Gramians) are captured in the approximation. Notice that the dominant subspace characterization is not improved if the number of samples is enlarged to 1000. Secondly, we determine the approximations $\mathcal{P}_T$ and $\mathcal{Q}_T$ according to Section 5.3.3 and call them Section 5.3.3 Gramians. The associated constants are chosen to be $c_F = c_G = 0$.

In Figure 5.1, the HSVs $\sigma_{T,i}$, $i = \{1,\dots,50\}$, of system (5.1) are displayed. By Theorem 5.3 and the explanations below this theorem, it is known that small truncated $\sigma_{T,i}$ go along with a small reduction error of the MOR scheme. Due to the rapid decay of these values, we can therefore conclude that a small error can already be achieved for small reduced dimensions r. For instance, we observe that $\sigma_{T,i} < 3.5\text{e}{-}06$ for $i \geq 8$ indicating very high accuracy in the ROM for $r \geq 7$. This is confirmed by the error plot in Figure 5.2 and the second column of Table 5.1. Moreover, Figure 5.2 shows the tightness of the error bound

in (5.28) that was specified in Theorem 5.3. The bound differs from the exact error only by a factor between 2.5 and 4.6 for the reduced dimensions considered in Figure 5.2 and is hence a good indicator of expected performance. Notice that the error is only exact up to deviations occurring due to the semi-implicit Euler-Maruyama discretization of (5.1) and (5.16) as well as the Monte-Carlo approximation of the expected value using 10 000 paths. Besides the MOR error based on P_T and Q_T, Table 5.1 states the errors in case the approximating Gramians of Sections 5.3.2 and 5.3.3 are used. It can be seen that both approximations perform roughly the same and that one loses an order of accuracy compared to the exact Gramian approach. However, one can lower the reduction error by an optimization with respect to the constants c, c_F, c_G. Moreover, we see that the accuracy is very good for the estimators of the covariances $F(T)$ and $G(T)$ used here.

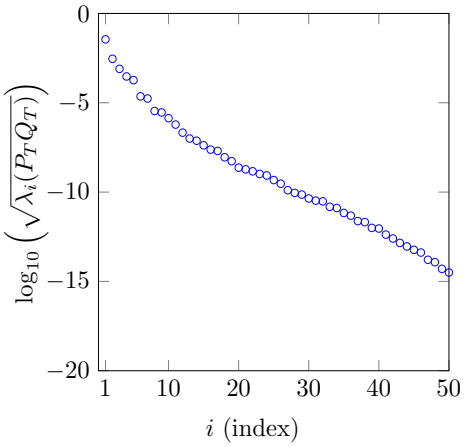

Figure 5.1: Decay of first 50 logarithmic HSVs of system (5.1) based on time-limited Gramians P_T and Q_T.

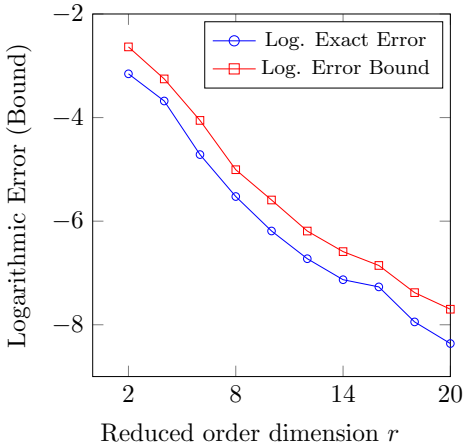

Figure 5.2: $\log_{10}\left(\sup_{t\in[0,1]}\mathbb{E}\|y(t) - y_r(t)\|_2\right)$ and logarithmic bound in (5.28) for different values of r.

	Error $\sup_{t\in[0,1]}\mathbb{E}\|y(t)-y_r(t)\|_2$ of MOR using		
Reduced dimension r	exact Gramians P_T, Q_T	Section 5.3.2 Gramians	Section 5.3.3 Gramians
2	7.00e−04	2.61e−03	1.75e−03
4	2.09e−04	1.82e−03	8.61e−04
8	2.99e−06	2.63e−05	4.51e−05
16	5.38e−08	1.31e−06	1.55e−06

Table 5.1: Error between the output y of (5.1) with $n = 100$ and the reduced output y_r of (5.16) using different Gramians to compute the balancing transformation $S = S_2$.

5.4.2 Simulations for $n = 1000$ and $T = 1$

We repeat the simulations of Subsection 5.4.1 for $n = 1000$. This is a scenario, where the exact Gramians are not available anymore. Therefore, we conduct the balancing MOR scheme using the Sections 5.3.2 and 5.3.3 Gramians only. In the context of the Section

5.3.2 Gramians, it is important to mention that in higher dimensions it is required to use very efficient discretizations of the Ito integrals in (5.43) and (5.46). Otherwise, a very small step size is needed such that from the computational point of view it is better to omit these Ito integrals within the estimators, i.e., just x_B and x_C are supposed to be sampled to approximate $F(T)$ and $G(T)$. Table 5.2 shows that the balancing related MOR technique based on the approximated Gramians of Sections 5.3.2 and 5.3.3 is beneficial in high dimensions. A very small reduction error can be observed and in the majority of the cases the sampling based approach seems slightly more accurate than the approach of Section 5.3.3 given the same type of approximations for $F(T)$ and $G(T)$ for each ansatz.

	Error $\sup_{t\in[0,1]} \mathbb{E}\|y(t) - y_r(t)\|_2$ of MOR using	
Reduced dimension r	Section 5.3.2 Gramians	Section 5.3.3 Gramians
2	1.43e−03	1.72e−03
4	2.07e−03	8.57e−04
8	5.18e−05	9.26e−05
16	2.13e−06	4.88e−06

Table 5.2: Error between the output y of (5.1) with $n = 1000$ and the reduced output y_r of (5.16) using Sections 5.3.2 and 5.3.3 Gramians to compute the balancing transformation $S = S_2$.

5.4.3 Relevance of T and K

As in Section 5.4.1, let us fix $n = 100$ to be able to compute the Gramians exactly. We begin with deriving reduced systems on different intervals $[0, T]$. Secondly, we extend our model to a stochastic differential equation with noise dimension $q = 2$ and investigate the effect of different correlations between the two Wiener processes.

Relevance of the Terminal Time Let us study the scenario of Section 5.4.1 with $T = 0.5, 1, 2, 3$ using the exact Gramians to illustrate that dominant subspaces are changing in time. Indeed, we observe in Table 5.3 that for a fixed reduced dimension r the error gets bigger the larger the interval $[0, T]$ is. This means that with increasing T the reduced dimension has to be enlarged to ensure a certain desired approximation error. This is also intuitive in the sense that it is generally harder to find a good approximation on a larger interval in comparison to a smaller one.

	Error $\sup_{t\in[0,T]} \mathbb{E}\|y(t) - y_r(t)\|_2$ of MOR for			
Reduced dimension r	$T = 0.5$	$T = 1$	$T = 2$	$T = 3$
2	3.98e−04	7.00e−04	2.17e−02	3.13e−02
4	1.46e−05	2.09e−04	2.86e−04	6.86e−04
8	2.82e−07	2.99e−06	7.80e−06	2.23e−05
16	5.46e−09	5.38e−08	1.12e−07	2.90e−07

Table 5.3: Error between the output y of (5.1) and the reduced output y_r of (5.16) using the exact Gramians: $n = 100$, $S = S_2$ and $T = 0.5, 1, 2, 3$.

Relevance the the Covariance Structure Let us extend the SPDE discretization by introducing $N_2 := N_1^{\frac{6}{5}}$ so that we have a system of the form (5.1) with $q = 2$ and standard Wiener processes W_1 and W_2. The goal is to investigate how the correlation between W_1 and W_2 influences the MOR error. For that reason, we choose the following three scenarios: $\mathbb{E}[W_1(t)W_2(t)] = \rho t$ with $\rho = 0, 0.5, 1$. Table 5.4 states the MOR errors for these correlations. In this example, we can observe that a higher correlation between the processes yields a larger error. A different observation was made in numerical examples studied in [98], where systems with high correlations in the noise processes gave a smaller reduction error. However, [98] studies different types of stochastic differential equations in the context of asset price models that do not have control inputs.

	Error $\sup_{t\in[0,1]} \mathbb{E}\|y(t) - y_r(t)\|_2$ of MOR for		
Reduced dimension r	$\rho = 0$	$\rho = 0.5$	$\rho = 1$
2	1.10e−03	1.43e−03	1.79e−03
4	2.44e−04	2.34e−04	3.24e−04
8	5.71e−06	8.95e−06	1.34e−05
16	1.64e−07	2.37e−07	3.36e−07

Table 5.4: Error between the output y of (5.1) and the reduced output y_r of (5.16) using the exact Gramians: $n = 100$, $S = S_2$, $T = 1$, $q = 2$ and different correlations $\rho = 0, 0.5, 1$.

6 Model Reduction for Stochastic Systems Driven by fBm

In this chapter, we study large-scale linear fractional stochastic systems representing, for example, spatially discretized stochastic partial differential equations (SPDEs) driven by fBm with a Hurst parameter $H \in [\frac{1}{2}, 1)$. Such equations in the case of $H < \frac{1}{2}$ are more realistic in modeling real-world phenomena compared to frameworks that do not capture memory effects. To the best of our knowledge, dimension reduction schemes for fBm settings have not been studied so far.

In this chapter, we investigate empirical reduced-order methods that are either based on snapshots (e.g., POD method) or on approximated Gramians. In each case, dominant subspaces are learned from data. These model reduction techniques are introduced and analyzed for stochastic systems with fractional and Wiener noise and later applied to spatially discretized SPDEs driven by fBm to reduce the computational cost arising from both the high dimension of the considered stochastic system and the large number of required Monte Carlo runs.

We validate our proposed techniques with numerical experiments for some large-scale stochastic differential equations driven by fBm. These results are published in [50, 51].

6.1 Setting and (Projection-Based) Reduced System

We consider the following Young/Stratonovich stochastic differential equation controlled by $u \in \mathcal{M}^2([0,T];\mathbb{R}^m)$ satisfying $\|u\|_T^2 := \mathbb{E}\int_0^T \|u(t)\|_2^2 \mathrm{d}t < \infty$ (see (3.20)).

$$\begin{aligned} \mathrm{d}x(t) &= [Ax(t) + Bu(t)]\mathrm{d}t + \sum_{i=1}^{q} N_i x(t) \circ \mathrm{d}W_i^H(t), \quad x(0) = x_0 = X_0 z, \\ y(t) &= Cx(t), \quad t \in [0,T], \end{aligned} \tag{6.1}$$

where $A, N_i \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $C \in \mathbb{R}^{p\times n}$, $X_0 \in \mathbb{R}^{n\times v}$, $z \in \mathbb{R}^v$ and $T > 0$ is the terminal time. $W_1^H, \ldots, W_q^H$ are independent fBm with Hurst index $H \in [1/2, 1)$. System (6.1) is defined as an integral equation using Definitions 2.13 ($H > 1/2$) and 2.15 ($H = 1/2$) to make sense of $\int_0^t N_i x(s) \circ \mathrm{d}W_i^H(s)$.

For the latter reduction procedure, rewriting the Stratonovich setting in the Itô form can be beneficial. Given $H = 1/2$, the state equation in (6.1) is equivalent to the Itô equation

$$\mathrm{d}x(t) = [(A + \frac{1}{2}\sum_{i=1}^{q} N_i^2)x(t) + Bu(t)]\mathrm{d}t + \sum_{i=1}^{q} N_i x(t)\mathrm{d}W_i^H(t) \tag{6.2}$$

exploiting that the quadratic covariation process is $\sum_{i=1}^{q}\int_0^t N_i^2 x(s)ds$, $t \in [0,T]$ (See (1.13)

in Chapter 1).

The goal of this chapter is to find a system of reduced order. This reduction is achieved through the application of the Petrov-Galerkin projection method, which is detailed in Section 3.2.4.2 of Chapter 3. For clarity, the Petrov-Galerkin method involves identifying a subspace that is spanned by the columns of a matrix $V \in \mathbb{R}^{n\times r}$. This subspace facilitates the approximation $x(t) \approx Vx_r(t)$, where $x_r(t)$ represents the reduced state vector. Inserting this into (6.1) yields

$$\begin{aligned} Vx_r(t) &= X_0 z + \int_0^t [AVx_r(s) + Bu(s)]\mathrm{d}s + \sum_{i=1}^{q} \int_0^t N_i V x_r(s) \circ \mathrm{d}W_i^H(s) + e(t), \\ y_r(t) &= CVx_r(t), \end{aligned} \tag{6.3}$$

We enforce the error $e(t)$ to be orthogonal to some space spanned by columns of $M \in \mathbb{R}^{n\times r}$, for which we assume that $M^\top V = I$. Multiplying (6.3) with $M^\top$ from the left yields

$$\begin{aligned} dx_r(t) &= [A_r x_r(t) + B_r u(t)]\mathrm{d}t + \sum_{i=1}^{q} N_{i,r} x_r(t) \circ \mathrm{d}W_i^H(t), \quad x_r(0) = x_{0,r} = X_{0,r} z, \\ y_r(t) &= C_r x_r(t), \quad t \in [0,T], \end{aligned} \tag{6.4}$$

where $X_{0,r} = M^\top X_0$ and

$$A_r = M^\top AV, \quad B_r = M^\top B, \quad N_{i,r} = M^\top N_i V, \quad C_r = CV,$$

refer to equation (3.45) in Chapter 3. If $M = V$ has orthonormal columns, we obtain a Galerkin approximation, as discussed in Section 3.2.4.3. On the other hand, we want to point out that reduced order systems can also be of a different form when $H = 1/2$. Inserting $x(t) \approx Vx_r(t)$ into (6.2) instead of (6.1) and conducting the same Petrov-Galerkin procedure, we obtain a reduced Itô system with drift coefficient $A_r + \frac{1}{2}\sum_{i=1}^{q} M^\top N_i^2 V$. Transforming this back into a Stratonovich equation yields

$$d\bar{x}_r(t) = [\big(A_r + \frac{1}{2}\sum_{i=1}^{q}(M^\top N_i^2 V - N_{i,r}^2)\big)\bar{x}_r(t) + B_r u(t)]\mathrm{d}t + \sum_{i=1}^{q} N_{i,r}\bar{x}_r(t) \circ \mathrm{d}W_i^H(t), \tag{6.5}$$

which is clearly different from the state equation in (6.4). This is due to the Itô - Stratonovich correction not being a linear transformation, as derived in (1.13). Another goal of this chapter is to analyze whether x_r or $\bar{x}_r$ performs better for $H = 1/2$. This question arises due to the fact that as soon as classical integration (Young/Stieltjes-sense) does not work anymore (e.g., $H \leq 1/2$), potential extensions of Young/Stieltjes-integrals are no longer unique. Therefore, it is interesting to ask for an optimal setting in which the dimension reduction is conducted. So, it can make sense to consider reduced systems (6.5) rather than the direct and maybe more intuitive choice (6.4).

6.2 Fundamental Solutions and Gramians

6.2.1 Fundamental Solutions and Their Properties

Before we are able to compute suitable reduced systems, we require fundamental solutions Φ.

Definition 6.1 A process $x(t)$ valued in $\mathbb{R}^n$, defined for $t \in [0,T]$, is recognized as a solution to the stochastic differential equation (SDE) specified in (6.1) if it almost surely satisfies the equation:

$$x(t;x_0,u) = x_0 + \int_0^t [Ax(s) + Bu(s)]\mathrm{d}s + \sum_{i=1}^{q} \int_0^t N_i x(s) \circ \mathrm{d}W_i^H(s) \tag{6.6}$$

where x_0 in $\mathbb{R}^n$ is the initial state and u belongs to the space $\mathcal{M}^2([0,T];\mathbb{R}^m)$, serving as the control function. In the specific case where no control is applied, the SDE simplifies to:

$$dx(t) = Ax(t)\mathrm{d}t + \sum_{i=1}^{q} N_i x(t) \circ \mathrm{d}W_i^H(t), \tag{6.7}$$

$$y(t) = Cx(t), \quad \text{for } t \in [0,T], \quad x(0) = x_0, \tag{6.8}$$

which defines the homogeneous solution x_{x_0} as $x_{x_0} := x(t;x_0,0)$.

These Φ will later lead to the concept of Gramians that identify dominant subspaces. The fundamental solution associated to (6.7) is a two-parameter matrix-valued stochastic process Φ solving

$$\Phi(t,s) = I + \int_s^t A\Phi(\tau,s)d\tau + \sum_{i=1}^{q} \int_s^t N_i \Phi(\tau,s) \circ \mathrm{d}W_i^H(\tau) \tag{6.9}$$

for $t \geq s \geq 0$. For simplicity, we set $\Phi(t) := \Phi(t,0)$ meaning that we omit the second argument if it is zero. We can separate the variables, since we have $\Phi(t,s) = \Phi(t)\Phi(s)^{-1}$ for $t \geq s \geq 0$, as discussed in Remark 3.12 in the Wiener case. This result can be extended to the context of fractional noise. Now, we derive the solution of the state equation (6.1) in the following proposition, which is a known result based on the product rule.

Proposition 6.2 The solution of the state equation (6.1) for $H \in [1/2,1)$ is given by

$$x(t;x_0,u) = \Phi(t)x_0 + \int_0^t \Phi(t,s)Bu(s)\mathrm{d}s, \quad t \in [0,T]. \tag{6.10}$$

Proof. Defining $k(t) = x_0 + \int_0^t \Phi(s)^{-1}Bu(s)\mathrm{d}s$, the result follows directly from the classical

product rule (available in the Young/Stratonovich case) to $\Phi(t)k(t), t \in [0,T]$. Thus

$$\begin{aligned}\Phi(t)k(t) &= x_0 + \int_0^t \Phi(s)\, dk(s) + \int_0^t (d\Phi(s))k(s) \\ &= x_0 + \int_0^t Bu(s)\mathrm{d}s + \int_0^t A\Phi(s)k(s)\mathrm{d}s + \sum_{i=1}^q \int_0^t N_i\Phi(s)k(s) \circ \mathrm{d}W_i^H(s),\end{aligned}$$

meaning that $\Phi(t)k(t)$, $t \in [0,T]$, is the solution to (6.1). The desired result follows directly from the identity $\Phi(t,s) = \Phi(t)\Phi^{-1}(s)$. □

As noted previously, a limitation of the stochastic system is its inability to offer a universally explicit expression for the fundamental matrix solution. Nevertheless, an explicit representation of the fundamental matrix solution, $\Phi(t)$, is achievable in instances where the matrices A and N_i for $i = 1, \ldots, q$ in Equation (6.7) are commutative, as demonstrated subsequently.

Proposition 6.3 Assume that all A, N_i commute. Then the fundamental solution $\Phi(t)$ of (6.7) is

$$\Phi(t) = \exp\left(At + \sum_{i=1}^q N_i W_i^H(t)\right).$$

Proof. We write $Z(t)$ for the exponent in (6.7), such that

$$dZ(t) = A\mathrm{d}t + \sum_{i=1}^q N_i \circ \mathrm{d}W_i^H(t).$$

Then we have

$$\begin{aligned}d\Phi(t) = e^{Z(t)}\, dZ(t) &= \Phi(t)\left(A\mathrm{d}t + \sum_{i=1}^q N_i \circ \mathrm{d}W_i^H(t)\right) \\ &= A\Phi(t)dt + \sum_{i=1}^q N_i \circ \mathrm{d}W_i^H(t),\end{aligned}$$

since $\Phi(t)$ commutes by assumption with A and all N_i for $i \in \{1, \ldots, q\}$. □

The fundamental solution lacks the strong semigroup feature compared to the deterministic case ($N_i = 0$). This means that $\Phi(t,s) = \Phi(t-s)$ does not hold $\mathbb{P}$-almost surely, as the trajectories of W^H on $[0, t-s]$ and $[s,t]$ are distinct. As discussed in Remark 3.20 of Chapter 3, the same property holds in the Wiener case. In the following lemma, we can demonstrate that the semigroup property holds in distribution exploiting the stationary increments of W^H. In fact, this lemma is the key for studying MOR for stochastic systems driven by processes with stationary (not necessarily independent) increments and can therefore be applied to settings beyond the case studied in this chapter.

Lemma 6.4 It holds that the fundamental solution of (6.1) satisfies

$$\Phi(t,s) \overset{d}{=} \Phi(t-s), \quad t \geq s \geq 0.$$

Proof. We consider $\Phi(\cdot)$ on the interval $[0, t-s]$ and $\Phi(\cdot, s)$ on $[s, t]$. Introducing the step size $\Delta t = \frac{t-s}{N}$, we find the partitions $t_k = k\Delta t$ and $t_k^{(s)} = s + t_k$, $k \in \{0, 1, \dots, N\}$, of $[0, t-s]$ and $[s, t]$. We employ the Euler discretization of Equation (6.9), as introduced in Section 2.8.1, as follows:

$$\begin{aligned} \Phi_{k+1} &= \Phi_k + A\Phi_k \Delta t + \sum_{j=1}^{q} N_j \Phi_k \Delta W_{j,k}^H, \\ \Phi_{k+1}^{(s)} &= \Phi_k^{(s)} + A\Phi_k^{(s)} \Delta t + \sum_{j=1}^{q} N_j \Phi_k^{(s)} \Delta W_{j,k}^{H,(s)}, \end{aligned} \tag{6.11}$$

where we define $\Delta W_{j,k}^H = W_j^H(t_{k+1}) - W_j^H(t_k)$ and $\Delta W_{j,k}^{H,(s)} = W_j^H(t_{k+1}^{(s)}) - W_j^H(t_k^{(s)})$. According to Theorem 2.30, the Euler scheme converges $\mathbb{P}$-almost surely for $H > 1/2$ yielding in particular convergence in distribution, that is

$$\Phi_N \xrightarrow{d} \Phi(t-s), \quad \Phi_N^{(s)} \xrightarrow{d} \Phi(t, s), \tag{6.12}$$

as $N \to \infty$. The Euler method does not converge almost surely in the Stratonovich setting. However, for $H = 1/2$, we can rewrite (6.9) as the Itô equation

$$\Phi(t, s) = I + \int_s^t (A + \frac{1}{2} \sum_{i=1}^{q} N_i^2) \Phi(\tau, s) d\tau + \sum_{i=1}^{q} \int_s^t N_i \Phi(\tau, s) \mathrm{d}W_i^H(\tau).$$

This equation can be discretized by a scheme like in (6.11) (Euler-Maruyama). The corresponding convergence is in $L^1(\Omega, \mathcal{F}, \mathbb{P})$ (see Theorem 1.38), so that we also have (6.12) for $H = 1/2$ as well. By simple calculation, we can get from (6.11) that

$$\begin{aligned} \Phi_N &= \prod_{k=0}^{N-1} \left(I + A\Delta t + \sum_{j=1}^{q} N_j \Delta W_{j,k}^H \right) =: F(Z), \\ \Phi_N^{(s)} &= \prod_{k=0}^{N-1} \left(I + A\Delta t + \sum_{j=1}^{q} N_j \Delta W_{j,k}^{H,(s)} \right) = F(Z^{(s)}), \end{aligned}$$

where $Z := (\Delta W_{j,k}^H)$ and $Z^{(s)} := (\Delta W_{j,k}^{H,(s)})$ ($j = 1, \dots, q$ and $k = 0, \dots, N-1$) are Gaussian vectors with mean zero. Notice that the function F is just slightly different for $H = 1/2$, i.e., A is replaced by $A + \frac{1}{2}\sum_{i=1}^{q} N_i^2$. It remains to show that the covariance matrices of Z and $Z^{(s)}$ coincide leading to $\Phi_N(t, s) \overset{d}{=} \Phi_N(t-s)$. Subsequently, the result is followed by (6.12). Using the independence of W_i^H and W_j^H for $i \neq j$, the non zero entries of the covariances of Z and $Z^{(s)}$ are $\mathbb{E}[\Delta W_{j,k}^H \Delta W_{j,\ell}^H]$ and $\mathbb{E}[\Delta W_{j,k}^{H,(s)} \Delta W_{j,\ell}^{H,(s)}]$ ($k, \ell = 0, 1, \dots, N-1$), respectively. These expressions are the same, since exploiting (2.1), we obtain that

$$\begin{aligned} \mathbb{E}[\Delta W_{j,k}^{H,(s)} \Delta W_{j,\ell}^{H,(s)}] &= \mathbb{E}[\left(W_j^H(s + t_{k+1}) - W_j^H(s + t_k)\right)\left(W_j^H(s + t_{\ell+1}) - W_j^H(s + t_\ell)\right)] \\ &= \frac{1}{2}\left(|t_{k+1} - t_\ell|^{2H} + |t_k - t_{\ell+1}|^{2H} - |t_{k+1} - t_{\ell+1}|^{2H} - |t_k - t_\ell|^{2H}\right) \end{aligned}$$

is independent of s. This concludes the proof. □

Let us mention that the result of Lemma 6.4 is new even for the well-studied case of $H = 1/2$. In fact, we later exploit that Lemma 6.4 yields

$$\mathbb{E}\left[\Phi(t,s)M\Phi(t,s)^\top\right] = \mathbb{E}\left[\Phi(t-s)M\Phi(t-s)^\top\right],$$

for a matrix M of suitable dimension. This property was proved for $H = 1/2$ using relations to matrix ODEs. This is not possible for general H, so that the much stronger result of Lemma 6.4 is required.

6.2.2 Exact and Empirical Gramians

6.2.2.1 Exact Gramians and Dominant Subspaces

Similar to the approach presented in the POD-based method outlined in Section 6.4.2, our methodology involves partitioning the primary system described in equation (6.1) into distinct subsystems in the following manner:

$$dx_u(t) = [Ax_u(t) + Bu(t)]\mathrm{d}t + \sum_{i=1}^{q} N_i x_u(t) \circ \mathrm{d}W_i^H(t), \quad x_u(0) = 0, \quad y_u(t) = Cx_u(t), \tag{6.13}$$

$$dx_{x_0}(t) = Ax_{x_0}(t)\mathrm{d}t + \sum_{i=1}^{q} N_i x_{x_0}(t) \circ \mathrm{d}W_i^H(t), \quad x_{x_0}(0) = x_0 = X_0 z, \quad y_{x_0}(t) = Cx_{x_0}(t). \tag{6.14}$$

Proposition 6.2 shows that we have the representations

$$x_{x_0}(t) = \Phi(t)x_0, \quad \text{and} \quad x_u(t) = \int_0^t \Phi(t,s)Bu(s)\mathrm{d}s,$$

so that $y(t) = y_{x_0}(t) + y_u(t)$ follows. Lemma 6.4 is now vital for a suitable definition of Gramians. Due to the weak semigroup property of the fundamental solution in Lemma 6.4, it turns out that (see Section 3.2.3 for more details):

$$P_{u,T} := \mathbb{E}\left[\int_0^T \Phi(s)BB^\top\Phi(s)^\top \mathrm{d}s\right], \quad P_{x_0,T} := \mathbb{E}\left[\int_0^T \Phi(s)X_0X_0^\top\Phi(s)^\top \mathrm{d}s\right]. \tag{6.15}$$

are the right notion of Gramians for (6.13) and (6.14). With (6.15) we then define a Gramian $P_T := P_{u,T} + P_{x_0,T}$ for the original state equation (6.1). In case of the output equation in (6.1), a Gramian can be introduced directly by

$$Q_T := \mathbb{E}\int_0^T \Phi(s)^\top C^\top C\Phi(s)ds. \tag{6.16}$$

The following proposition contains estimates that tell us in which sense the above Gramians characterize dominant subspaces of the system. It heavily relies on Lemma 6.4 indicating the importance of this key lemma in the theory of dimension reduction for stochastic

systems driven by fBm. In particular, the next proposition addresses the theoretical shortcomings of [51].

Proposition 6.5 Given $v \in \mathbb{R}^n$, an initial state of the form $x_0 = X_0 z$, a parameter $p = 1, 2$ and a control $u \in \mathcal{M}^2([0,T];\mathbb{R}^m)$ that is assumed to be deterministic if $p = 2$. Then, we have that

$$\int_0^T \mathbb{E}\langle x_{x_0}(t), v\rangle_2^2 \mathrm{d}t \leq v^\top P_{x_0,T} v \|z\|_2^2, \quad \sup_{t\in[0,T]} \mathbb{E}|\langle x_u(t), v\rangle_2|^p \leq (v^\top P_{u,T} v)^{\frac{p}{2}} \|u\|_T^p \tag{6.17}$$

for $p = 1, 2$. Consequently, we have

$$\int_0^T \mathbb{E}|\langle x(t), v\rangle_2| \mathrm{d}t \leq \sqrt{2}\sqrt{v^\top P_T v} \max\{\sqrt{T}\|z\|_2, T\|u\|_T\}. \tag{6.18}$$

for general $u \in \mathcal{M}^2([0,T];\mathbb{R}^m)$ and

$$\int_0^T \mathbb{E}|\langle x(t), v\rangle_2|^2 \mathrm{d}t \leq 2 v^\top P_T v \max\{\|z\|_2^2, T\|u\|_T^2\}. \tag{6.19}$$

if u is further deterministic. Moreover, it holds that

$$\int_0^T \mathbb{E}\|C\Phi(t)v\|_2^2 \mathrm{d}t = v^\top Q_T v. \tag{6.20}$$

Proof. The first relation is a simple consequence of the inequality of Cauchy-Schwarz and the representation of x_{x_0} in Proposition 6.2. Thus,

$$\begin{aligned}
\int_0^T \mathbb{E}\langle x_{x_0}(t), v\rangle_2^2 \mathrm{d}t &= \mathbb{E}\int_0^T \langle \Phi(t) X_0 z, v\rangle_2^2 \mathrm{d}t = \mathbb{E}\int_0^T \langle z, X_0^\top \Phi(t)^\top v\rangle_2^2 \mathrm{d}t \\
&\leq \|z\|_2^2 \mathbb{E}\int_0^T \|X_0^\top \Phi(t)^\top v\|_2^2 \mathrm{d}t \\
&= \|z\|_2^2 v^\top \int_0^T \mathbb{E}\left[\Phi(t) X_0 X_0^\top \Phi(t)^\top\right] \mathrm{d}t v \\
&= v^\top P_{x_0,T} v \|z\|_2^2.
\end{aligned}$$

Utilizing equation (6.10) and the Cauchy-Schwarz inequality once more, we have

$$\begin{aligned}
\mathbb{E}|\langle x_u(t), v\rangle_2|^p &= \mathbb{E}\left|\langle \int_0^t \Phi(t,s) B u(s) \mathrm{d}s, v\rangle_2\right|^p \\
&\leq \mathbb{E}\left[\left(\int_0^t |\langle \Phi(t,s) B u(s), v\rangle_2| \,\mathrm{d}s\right)^p\right] \\
&= \mathbb{E}\left[\left(\int_0^t \left|\langle u(s), B^\top \Phi(t,s)^\top v\rangle_2\right| \mathrm{d}s\right)^p\right] \\
&\leq \mathbb{E}\left[\left(\int_0^t \|u(s)\|_2 \|B^\top \Phi(t,s)^\top v\|_2 \mathrm{d}s\right)^p\right] \\
&\leq \left(v^\top \mathbb{E}\int_0^t \Phi(t,s) B B^\top \Phi(t,s)^\top \mathrm{d}s\, v\right)^{\frac{p}{2}} \|u\|_T^p.
\end{aligned}$$

for $t \in [0,T]$. Based on Lemma 6.4, we obtain that

$$\mathbb{E}\left[\Phi(t,s)BB^\top\Phi(t,s)^\top\right] = \mathbb{E}\left[\Phi(t-s)BB^\top\Phi(t-s)^\top\right].$$

Hence,

$$\mathbb{E}|\langle x_u(t),v\rangle_2|^p \le (v^\top\mathbb{E}\int_0^t \Phi(t-s)BB^\top\Phi(t-s)^\top \mathrm{d}s\, v)^{\frac{p}{2}}\,\|u\|_T^p \le (v^\top P_{u,T}\,v)^{\frac{p}{2}}\,\|u\|_T^p$$

by variable substitution and the increasing nature of $P_{u,T}$ and $\|u\|_T^2$ in T. This shows the second part of (6.17). Exploiting Proposition 6.2, we know that $x = x_{x_0} + x_u$. Therefore, we have

$$\begin{aligned}\int_0^T \mathbb{E}\langle x(t),v\rangle_2^2\mathrm{d}t &\le 2\Big(\int_0^T \mathbb{E}\langle x_{x_0}(t),v\rangle_2^2\mathrm{d}t + \int_0^T \mathbb{E}\langle x_u(t),v\rangle_2^2\mathrm{d}t\Big)\\ &\le 2\Big(\int_0^T \mathbb{E}\langle x_{x_0}(t),v\rangle_2^2\mathrm{d}t + T\sup_{t\in[0,T]}\mathbb{E}\langle x_u(t),v\rangle_2^2\Big)\end{aligned}$$

by the linearity of the inner product in the first argument. Applying (6.17) to this inequality yields (6.18) using that $P_T = P_{x_0,T} + P_{u,T}$. On the other hand, we obtain

$$\begin{aligned}\int_0^T \mathbb{E}|\langle x(t),v\rangle_2|\mathrm{d}t &\le \int_0^T \mathbb{E}|\langle x_{x_0}(t),v\rangle_2|\mathrm{d}t + \int_0^T \mathbb{E}|\langle x_u(t),v\rangle_2|\mathrm{d}t\\ &\le \sqrt{T}\sqrt{\int_0^T \mathbb{E}\langle x_{x_0}(t),v\rangle_2^2\mathrm{d}t} + T\sup_{t\in[0,T]}\mathbb{E}|\langle x_u(t),v\rangle_2|.\end{aligned}$$

Applying (6.17) with $p=1$ to this inequality yields (6.18) using that

$$\sqrt{v^\top P_{x_0,T}v} + \sqrt{v^\top P_{u,T}v} \le \sqrt{2}\sqrt{v^\top P_T v}.$$

By the definitions of Q_T in (6.16) and the Euclidean norm, we have

$$\int_0^T \mathbb{E}\|C\Phi(t)v\|_2^2\mathrm{d}t = v^\top\int_0^T \mathbb{E}\left[\Phi(t)^\top C^\top C\Phi(t)\right]\mathrm{d}t v = v^\top Q_T v.$$

So this proof is concluded. □

Remark 6.6 If the limits $P_{x_0} = \lim_{T\to\infty} P_{x_0,T}$, $P_u = \lim_{T\to\infty} P_{u,T}$, $P = \lim_{T\to\infty} P_T$ and $Q = \lim_{T\to\infty} Q_T$ exist, the Gramians in Proposition 6.5 can be replaced by their limit as we have $v^\top P_T v \le v^\top P v$, $v^\top Q_T v \le v^\top Q v$ etc for all $v \in \mathbb{R}^n$, as discussed in Section 3.2.3 of Chapter 3.

The following remark explains the role of the results in Proposition 6.5 in more detail. In fact, Proposition 6.5 delivers the theoretical motivation for the dimension reduction procedure studied in this chapter.

Remark 6.7 We can read Proposition 6.5 as follows. If v is an eigenvector of $P_{x_0,T}$ and $P_{u,T}$, respectively, associated to a small eigenvalue, then x_{x_0} and x_u are small in the

direction of v. Such state directions can therefore be neglected. The same interpretation holds for x using (6.18) when v is a respective eigenvector of P_T. Now, given $t_0 \in [0,T)$, we expand the state variable as

$$x(t_0) = \sum_{k=1}^{n} \langle x(t_0), q_k \rangle_2 q_k,$$

where $(q_k)_{k=1,\dots,n}$ represents an orthonormal set of eigenvectors of Q_T. We aim to answer the question which directions q_k in $x(t_0)$ barely contribute to y on $[t_0,T]$. We can represent the state in (6.1) by

$$x(t) = x(t_0) + \int_{t_0}^{t} [Ax(s) + Bu(s)]\mathrm{d}s + \sum_{i=1}^{q} \int_{t_0}^{t} N_i x(s) \circ \mathrm{d}W_i^H(s), \quad t \in [t_0,T]. \tag{6.21}$$

and introduce $\tilde{x}$ as the solution of (6.21) when replacing $x(t_0)$ by

$$\tilde{x}(t_0) := x(t_0) - \langle x(t_0), q_k \rangle_2 q_k,$$

i.e., the direction q_k is neglected. The process $x - \tilde{x}$ then solves (6.21) with $u \equiv 0$ starting in $\langle x(t_0), q_k \rangle_2 q_k$ at t_0. Therefore, the difference in the associated outputs is

$$y(t) - C\tilde{x}(t) = \langle x(t_0), q_k \rangle_2 C\Phi(t,t_0)q_k, \quad t \in [t_0,T],$$

using the solution representation in (6.10). For that reason, we solely focus on the term $C\Phi(t,t_0)q_k$ and observe that

$$\int_{t_0}^{T} \mathbb{E}\|C\Phi(t,t_0)q_k\|_2^2 \mathrm{d}t = \int_{t_0}^{T} \mathbb{E}\|C\Phi(t-t_0)q_k\|_2^2 \mathrm{d}t \leq \int_{0}^{T} \mathbb{E}\|C\Phi(t)q_k\|_2^2 \mathrm{d}t$$

using Lemma 6.4. Identity (6.20) therefore tells us that the direction $v = q_k$ in $x(t_0)$ has a low impact on $y(t)$, $t \in [t_0,T]$, if the corresponding eigenvalue is small. Such q_k can be removed from the each state $x(t_0)$ without causing a large error in between the exact output y and its approximation $C\tilde{x}$.

6.2.2.2 Approximation and Computation of Gramians

In theory, Proposition 6.5 together with Remark 6.7 is the key when aiming to identify dominant subspaces of (6.1) that lead to ROMs. However, for practical purposes, strategies to compute the associated Gramians are vital.

Empirical Gramians for $H \geq 1/2$ The Gramians that we defined above are hard to compute. In fact, no established connection exists between these Gramians and algebraic Lyapunov equations or matrix differential equations when $H > \frac{1}{2}$, in contrast to the Wiener case, which is detailed in Section 5.1.1 of Chapter 5. For that reason, we suggest an empirical approach in the following in which approximate Gramians based on sampling are calculated. In particular, we consider a discretization of the integral representations by a Monte Carlo method. Let us introduce an equidistant time grid $0 = s_0 < s_1 < \cdots < s_N = T$ and let $\mathcal{N}_s$ further be the number of Monte-Carlo samples.

Given that $\mathcal{N}$ and $\mathcal{N}_s$ are sufficiently large, we obtain

$$\begin{aligned} P_{u,T} \approx \bar{P}_{u,T} &= \frac{T}{\mathcal{N}\cdot\mathcal{N}_s}\sum_{i=1}^{\mathcal{N}}\sum_{j=1}^{\mathcal{N}_s}\Phi(s_i,\omega_j)BB^\top\Phi(s_i,\omega_j)^\top, \\ P_{x_0,T} \approx \bar{P}_{x_0,T} &= \frac{T}{\mathcal{N}\cdot\mathcal{N}_s}\sum_{i=1}^{\mathcal{N}}\sum_{j=1}^{\mathcal{N}_s}\Phi(s_i,\omega_j)X_0X_0^\top\Phi(s_i,\omega_j)^\top, \end{aligned} \tag{6.22}$$

where $\omega_j \in \Omega$. Now, the advantage is that $\Phi(\cdot)B$ and $\Phi(\cdot)X_0$ are easy to sample as they are the solutions of the control independent variable x_{x_0} in (6.14) with initial states $x_0 \mapsto B$ and $x_0 \mapsto X_0$, respectively. This is particularly feasible if B and X_0 only have a few columns. Based on (6.22), we can then define $\bar{P}_T := \bar{P}_{x_0,T} + \bar{P}_{u,T}$ approximating P_T. Here, the goal is to choose $\mathcal{N}$ and $\mathcal{N}_s$ so that the estimates in Proposition 6.5 still hold (approximately) ensuring the dominant subspace characterization by the empirical Gramians. Notice that if the limits of the Gramians as $T \to \infty$ shall be considered, then the terminal time needs to be chosen sufficiently large. In fact, it is also not an issue to write down the empirical version of Q_T which is

$$\bar{Q}_T = \frac{T}{\mathcal{N}\cdot\mathcal{N}_s}\sum_{i=1}^{\mathcal{N}}\sum_{j=1}^{\mathcal{N}_s}\Phi(s_i,\omega_j)^\top C^\top C\Phi(s_i,\omega_j).$$

However, this object is computationally much more involved. This is because $C\Phi(\cdot)$ is not a solution to an equations like (6.14) that can be sampled easily in case only a few initial states are of interest. In fact, we might have to sample from (6.9) to determine $\bar{Q}_T$. This is equivalent to computing samples of x_{x_0} in (6.14) for n different initial states, i.e., $x_0 \mapsto I$. The issue is that n is very large, whereas the number of columns of B and X_0 is generally low. This leaves the open question of whether $\bar{Q}_T$ is numerically tractable.

Exact computation of Gramians for $H = 1/2$ Let us briefly discuss that the computation of P_T, Q_T or their limits as $T \to \infty$ is easier when we are in the Stratonovich setting of $H = 1/2$. Once more let us point out the relation between Itô and Stratonovich differential equation. So, the fundamental solution of the state equation in (6.1) defined in (6.9) is also the fundamental solution of (6.2), i.e., it satisfies

$$\Phi(t) = I + \int_0^t A_N\Phi(s)ds + \sum_{i=1}^{q}\int_0^t N_i\Phi(s)\mathrm{d}W_i^H(s),$$

where $A_N := A + \frac{1}{2}\sum_{i=1}^{q} N_i^2$. Let us consider the linear operators

$$\mathcal{L}_{A_N}(X) = A_NX + XA_N^\top, \quad \text{and} \quad \Pi_N(X) = \sum_{i=1}^{q} N_iXN_i^\top,$$

then it is a well-established fact, as a consequence of Itô's product rule demonstrated in Example 1.27, that $Z(t) = \mathbb{E}\left[\Phi(t)M\Phi(t)^\top\right]$ solves

$$\frac{d}{dt}Z(t) = \mathcal{L}_{A_N}\big[Z(t)\big] + \Pi_N\big[Z(t)\big], \quad Z(0) = M, \quad t \geq 0, \tag{6.23}$$

where M is a matrix of suitable dimension. Setting $M = BB^\top + X_0X_0^\top$ and integrating (6.23) yields

$$Z(T) - BB^\top - X_0X_0^\top = \mathcal{L}_{A_N}[P_T] + \Pi_N[P_T] \tag{6.24}$$

using that $P_T = \mathbb{E}\Big[\int_0^T \Phi(s)\Big(BB^\top + X_0X_0^\top\Big)\Phi(s)^\top \mathrm{d}s\Big]$. If system (6.1) is mean square asymptotically stable, that is, $\mathbb{E}\|\Phi(t)\|^2$ decays exponentially to zero, then we even find

$$-BB^\top - X_0X_0^\top = \mathcal{L}_{A_N}[P] + \Pi_N[P]$$

for the limit P of P_T. There is still a small gap in the theory left in Theorem 5.1 on how to compute Q_T in the case of $H = 1/2$. Therefore, the following proposition was stated under the additional assumption that $C^\top C$ is contained in the eigenspace of $\mathcal{L}^*_{A_N} + \Pi^*_N$, where $\mathcal{L}^*_{A_N}(X) = A_N^\top X + XA_N$, $\Pi^*_N(X) = \sum_{i=1}^q N_i^\top X N_i$. We prove this result in full generality below.

Proposition 6.8 Given that we are in the Stratonovich setting of $H = 1/2$. Then, the function $Z_*(t) = \mathbb{E}\big[\Phi(t)^\top C^\top C\Phi(t)\big]$ solves

$$\frac{d}{\mathrm{d}t}Z_*(t) = \mathcal{L}^*_{A_N}[Z_*(t)] + \Pi^*_N[Z_*(t)], \quad Z_*(0) = C^\top C, \quad t \geq 0. \tag{6.25}$$

Proof. Let us vectorize the matrix differential equation (6.23) leading to

$$\frac{d}{\mathrm{d}t}\operatorname{vec}[Z(t)] = \mathcal{K}\operatorname{vec}[Z(t)], \quad \operatorname{vec}[Z(0)] = \operatorname{vec}[M],$$

where $\mathcal{K}$ defined as follows

$$\mathcal{K} = A_N \otimes I + I \otimes A_N + \sum_{i=1}^q N_i \otimes N_i.$$

Therefore, we know that

$$\mathrm{e}^{\mathcal{K}t}\operatorname{vec}[M] = \operatorname{vec}[Z(t)] = \operatorname{vec}\Big[\mathbb{E}\big[\Phi(t)M\Phi(t)^\top\big]\Big] = \mathbb{E}\big[\Phi(t)\otimes\Phi(t)\big]\operatorname{vec}[M],$$

again exploiting the relation between the vectorization and the Kronecker product (see Proposition B.5). Since this holds for all matrices M, it follows that $\mathbb{E}\big[\Phi(t)\otimes\Phi(t)\big] = \mathrm{e}^{\mathcal{K}t}$. This is now applied to

$$\operatorname{vec}\big[Z_*(t)\big] = \operatorname{vec}\Big[\mathbb{E}\big[\Phi(t)^\top C^\top C\Phi(t)\big]\Big] = \mathbb{E}\big[\Phi(t)^\top\otimes\Phi(t)^\top\big]\operatorname{vec}[C^\top C] = \mathrm{e}^{\mathcal{K}^\top t}\operatorname{vec}[C^\top C]$$

since $\mathbb{E}\big[\Phi(t)^\top\otimes\Phi(t)^\top\big] = \big(\mathbb{E}\big[\Phi(t)\otimes\Phi(t)\big]\big)^\top$. Therefore, it holds that

$$\frac{d}{\mathrm{d}t}\operatorname{vec}[Z_*(t)] = \mathcal{K}^\top\operatorname{vec}[Z_*(t)], \quad \operatorname{vec}[Z_*(0)] = \operatorname{vec}[C^\top C].$$

Devectorizing this equation and exploiting that $\mathcal{K}^\top$ serves as the matrix representation of $\mathcal{L}^*_{A_N} + \Pi^*_N$ leads to the claim of this proposition. □

Integrating (6.25) and using that $Q_T = \mathbb{E}\big[\int_0^T \Phi(t)^\top C^\top C \Phi(t) \mathrm{d}t\big]$ leads to

$$Z_*(T) - C^\top C = \mathcal{L}^*_{A_N}\big[Q_T\big] + \Pi^*_N\big[Q_T\big]. \tag{6.26}$$

Once more, mean square asymptotic stability yields the well-known relation

$$-C^\top C = \mathcal{L}^*_{A_N}\big[Q\big] + \Pi^*_N\big[Q\big]$$

by taking the limit as $T \to \infty$ in (6.26). Although we found algebraic equation (6.24) and (6.26) from which P_T and Q_T could be computed, it is still very challenging to solve these equations. This is mainly due to the unknowns $Z(T)$ and $Z_*(T)$. In fact, Section 5.3 of Chapter 5 discusses strategies based on sampling and variance reduction to address the problems defined in equations (6.24) and (6.26). For further details, the reader is referred to this chapter.

6.3 Model Reduction of Young/Stratonovich Differential Equations

In this section, we introduce ROMs that are based on the (empirical) Gramians of Section 6.2.2 as they (approximately) identify the dominant subspaces of (6.1). In order to accomplish this, we discuss state space transformations first that diagonalize these Gramians. This diagonalization facilitates the assignment of unimportant directions in the dynamics to specific state components, as outlined in Proposition 6.5. Subsequently, the issue is split up into two parts. A truncation procedure is briefly explained for the general case of $H \in [1/2, 1)$, in which unimportant state variables are removed. This strategy is associated with (Petrov-)Galerkin schemes sketched in Section 6.1. Later, we focus on the case of $H = 1/2$ and point out an alternative ansatz that is supposed to perform better than the previously discussed projection method. Let us notice once more that since a fractional Brownian motion with $H > 1/2$ does not have independent increments, no Lyapunov equations associated with the Gramians can be derived. Therefore, we frequently refer to the empirical versions of these Gramians and the corresponding reduced dimension techniques.

6.3.1 State Space Transformation and Balancing

We introduce a new variable $x_S(t) = Sx(t)$, where S is a regular matrix. This can be interpreted as a coordinate transform that is chosen in order to diagonalize the Gramians of Section 6.2.2. This transformation is the basis for the dimension reduction discussed in Sections 6.3.2 and 6.3.3. Using the same procedure as discussed in Section 5.1.2, we substitute $x_S(t) = Sx(t)$ into Equation (6.1) to obtain:

$$\begin{aligned} dx_S(t) &= [A_S x_S(t) + B_S u(t)]\mathrm{d}t + \sum_{i=1}^{q} N_{i,S} x_S(t) \circ \mathrm{d}W_i^H(t), \quad x_S(0) = x_{0,S} = X_{0,S} z, \\ y(t) &= C_S x_S(t), \quad t \in [0, T], \end{aligned} \tag{6.27}$$

where $A_S = SAS^{-1}$, $B_S = SB$, $N_{i,S} = SN_iS^{-1}$, $X_{0,S} = SX_0$ and $C_S = CS^{-1}$. As previously noted throughout this thesis and as evidenced by equation (6.27), the output remains unchanged under the transformation. However, the fundamental solution of the state equation in (6.27) is

$$\Phi_S(t) = S\Phi(t)S^{-1}. \tag{6.28}$$

Relation (6.28) immediately transfers to the Gramians which are

$$P_{T,S} := \mathbb{E}\int_0^T \Phi_S(s)(B_SB_S^\top + X_{0,S}X_{0,S}^\top)\Phi_S(s)^\top \mathrm{d}s = SP_TS^\top \tag{6.29}$$

$$Q_{T,S} := \mathbb{E}\int_0^T \Phi_S(s)^\top C_S^\top C_S\Phi_S(s)\mathrm{d}s = S^{-\top}Q_TS^{-1}. \tag{6.30}$$

Exploiting (6.28) again, the same relations like in (6.29) and (6.30) hold true if P_T and Q_T are replaced by their limits P, Q or their empirical versions $\bar{P}_T, \bar{Q}_T$. In the next definition, different diagonalizing transformations S are introduced.

Definition 6.9 (*i*) Let the state space transformation S be given by the eigenvalue decomposition $P_T = S^\top \Sigma S$, where Σ is the diagonal matrix of eigenvalues of P_T. Then, the procedure is called P_T-balancing.

(*ii*) Let P_T and Q_T be positive definite matrices. If S is of the form $S = \Sigma^{\frac{1}{2}}U^\top L^{-1}$ with the factorization $P_T = LL^\top$ and the spectral decomposition $L^\top Q_T L = U\Sigma^2 U^\top$, where Σ^2 is the diagonal matrix of eigenvalues of P_TQ_T. Then, the transformation is called P_T/Q_T-balancing.

(*iii*) Replacing P_T and Q_T by their limits (as $T \to \infty$) in (*i*) and (*ii*), then the schemes are called P-balancing (see Section 3.2.4.3) or P/Q-balancing (see Section 3.2.4.4), respectively, where in these cases Σ is either the matrix of eigenvalues of P or Σ^2 contains the eigenvalues of PQ.

(*iv*) Using the empirical versions of P_T and Q_T instead, the methods in (*i*) and (*ii*) are called $\bar{P}_T$-balancing and $\bar{P}_T/\bar{Q}_T$-balancing. Here, Σ can be viewed as a random diagonal matrix of the respective eigenvalues.

Notice that balancing based on Gramians P_T, P or $\bar{P}_T$ refers to the aim of an approximation of the full state x instead of y. Diagonalizing only one Gramian is, of course, also computationally cheaper but certainly leads to a worse approximation of y if information in Q_T, Q or $\bar{Q}_T$ is not involved in the model reduction procedure. It is not difficult to check that the transformations introduced in Definition 6.9 diagonalize the underlying Gramians. Nevertheless, we formulate the following proposition.

Proposition 6.10
- Using the matrix S in Definition 6.9 (*i*), we find that the state variable Gramian of system (6.27) is $P_{T,S} = \Sigma$.
- If instead S is of the form given in Definition 6.9 (*ii*), we have $P_{T,S} = Q_{T,S} = \Sigma$.
- The same type of diagonalization is established if the underlying Gramians are either P, Q or $\bar{P}_T, \bar{Q}_T$.

Proof. The result follows by inserting the respective S into (6.29) and (6.30). Since these relations also hold true for the pairs P, Q and $\bar{P}_T, \bar{Q}_T$, the same argument applies in these cases as well. □

Having diagonal Gramians Σ, Proposition 6.5 (choose v to be the ith unit vector in $\mathbb{R}^n$) together with Remark 6.7 tells us that we can neglect state components in (6.27) that correspond to small diagonal entries σ_i of Σ. Those have to be truncated to obtain a reduced system.

6.3.2 Reduced Order Models Based on Projection

In that spirit, we decompose the diagonal Gramian based on one of the balancing procedures in Definition 6.9. We write

$$\Sigma = \begin{bmatrix} \Sigma_1 & \\ & \Sigma_2 \end{bmatrix}, \tag{6.31}$$

where $\Sigma_1 \in \mathbb{R}^{r\times r}$ contains the r large diagonal entries of Σ and Σ_2 the remaining small ones. We further partition the balanced coefficient of (6.27) as follows

$$A_S = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad B_S = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}, \quad N_{i,S} = \begin{bmatrix} N_{i,11} & N_{i,12} \\ N_{i,21} & N_{i,22} \end{bmatrix} \quad X_{0,S} = \begin{bmatrix} X_{0,1} \\ X_{0,2} \end{bmatrix} \quad C_S = \begin{bmatrix} C_1 & C_2 \end{bmatrix}. \tag{6.32}$$

The balanced state of (6.27) is decomposed as $x_S = \left[\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\right]$, where x_1 and x_2 are associated to Σ_1 and Σ_2, respectively. Now, exploiting the insights of Proposition 6.5, x_2 barely contributes to (6.27). We remove the equation for x_2 from the dynamics and set it equal to zero in the remaining parts. This yields a reduced system

$$\begin{aligned} dx_r(t) &= [A_{11}x_r(t) + B_1u(t)]\mathrm{d}t + \sum_{i=1}^{q} N_{i,11}x_r(t) \circ \mathrm{d}W_i^H(t), \quad x_r(0) = x_{0,r} = X_{0,1}z, \\ y_r(t) &= C_1x_r(t), \quad t \in [0,T], \end{aligned} \tag{6.33}$$

which is of the form like in (6.4). If balancing according to Definition 6.9 is used, then V are the first r columns of S^{-1}, whereas W represents the first r columns of $S^\top$, as derived in equation (5.17) of Chapter 5. Notice that if solely P_T, P or $\bar{P}_T$ are diagonalized (instead of a pair of Gramians), we have $S^{-1} = S^\top$ and hence $W = V$. This method is discussed in detail under the Galerkin approximation in Section 3.2.4.3.

6.3.3 An Alternative Approach for the Stratonovich Setting ($H = 1/2$)

6.3.3.1 The Alternative

As sketched in Section 6.1, the truncation/projection procedure is not unique for $H = 1/2$ meaning that (6.5) can be considered instead of (6.33) (being of the form (6.4)). Such a reduced system is obtained if we rewrite the state of (6.27) as a solution to an Itô equation

meaning that A_S becomes $A_{SN} = A_S + \frac{1}{2}\sum_{i=1}^{q} N_{i,S}^2$ in the Itô setting. Now, removing x_2 from this system like we explained in Section 6.3.2, we obtain a reduced Itô system

$$\begin{aligned} dx_r(t) &= [A_{N,11}x_r(t) + B_1u(t)]\mathrm{d}t + \sum_{i=1}^{q} N_{i,11}x_r(t)\mathrm{d}W_i^H(t), \quad x_r(0) = x_{0,r} = X_{0,1}z, \\ y_r(t) &= C_1x_r(t), \quad t \in [0,T], \end{aligned} \tag{6.34}$$

where $A_{N,11} = A_{11} + \frac{1}{2}\sum_{i=1}^{q}(N_{i,11}^2 + N_{i,12}N_{i,21})$ is the left upper $r \times r$ block of A_{SN}. In Stratonovich form, the system is

$$\begin{aligned} dx_r(t) &= [(A_{11} + \frac{1}{2}\sum_{i=1}^{q} N_{i,12}N_{i,21})x_r(t) + B_1u(t)]\mathrm{d}t + \sum_{i=1}^{q} N_{i,11}x_r(t) \circ \mathrm{d}W_i^H(t), \\ y_r(t) &= C_1x_r(t), \quad t \in [0,T], \quad x_r(0) = x_{0,r} = X_{0,1}z, \end{aligned} \tag{6.35}$$

which has a state equation of the structure given in (6.5).

6.3.3.2 Comparison of (6.33) and (6.35) for $H = 1/2$

Let us continue setting $H = 1/2$. Moreover, we assume $x_0 = 0$ in this subsection for simplicity. We only focus on P- as well as P/Q-balancing (explained in Definition 6.9 (iii)) in order to emphasize our arguments. In addition, we always suppose that P and Q are positive definite. Let us point out that relations between (6.1) and (6.35) are well-studied due to the model reduction theory of Itô equations exploiting that these Stratonovich equations are equivalent to (6.2) and (6.34). In fact, the (uncontrolled) state equation is mean square asymptotically stable ($\mathbb{E}\|\Phi(t)\|^2 \to 0$ as $t \to \infty$) if and only if the same is true for (6.2) (see Theorem 3.21). This type of stability is well-investigated in Itô settings, see, e.g., [25, 59]. It is again equivalent to the existence of a positive definite matrix X, so that the operator $\mathcal{L}_{A_N} + \Pi_N$ evaluated at X is a negative definite matrix (as discussed in part (iv) of Theorem 3.19), i.e.,

$$\mathcal{L}_{A_N}[X] + \Pi_N[X] < 0. \tag{6.36}$$

Now, applying P/Q-balancing to (6.1) under the assumptions we made in this subsection, the reduced system (6.35) preserves this property, i.e., there exists a positive definite matrix X_r, so that

$$A_{N,11}X_r + X_rA_{N,11}^\top + \sum_{i=1}^{q} N_{i,11}X_rN_{i,11}^\top < 0. \tag{6.37}$$

This result was established in [11] and Theorem 3.21 given that $\sigma_r \neq \sigma_{r+1}$, where σ_i is the ith diagonal entry of Σ. If P-balancing is used instead, (6.37) basically holds as described in [99]. However, generally a further Galerkin projection of the reduced system (not causing an error) is required in order to ensure stability preservation. We illustrated with the following example that stability is not necessarily preserved in (6.33) given the Stratonovich case.

Example 6.11 Let us fix $x_0 = 0$, $q = 1$ and consider (6.1) with

$$A = \begin{bmatrix} -\frac{13}{8} & \frac{5}{4} \\ -\frac{5}{4} & -2 \end{bmatrix}, \quad B = C^\top = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad N_1 = \begin{bmatrix} \frac{3}{2} & -1 \\ 1 & 1 \end{bmatrix}$$

and hence $A_N = \left[\begin{smallmatrix} -1 & 0 \\ 0 & -2 \end{smallmatrix}\right]$. This system is asymptotically mean square stable, since (6.36) is satisfied. We apply P/Q-balancing in order to compute ROMs (6.33) and (6.35) for $r = 1$ and $H = 1/2$. Now, we find that $2A_{N,11} + N_{1,11}^2 = -0.85926 < 0$ which is equivalent to (6.37) in the scalar case. On the other hand, (6.33) is not stable, because $2(A_{11} + 0.5N_{1,11}^2) + N_{1,11}^2 = 0.13825 > 0$.

Example 6.11 shows us that we cannot generally expect a good approximation of (6.1) by (6.33) in the Stratonovich setting as the asymptotic behavior can be contrary. This is an important theoretical finding as it indicates that direct dimension reduction in the Stratonovich framework is not optimal.

We emphasize this argument further by looking at the error of the approximations if the full model (6.1) and the reduced system (6.33) have the same asymptotic behavior. First, let us note the following. If (6.1) is mean square asymptotically stable, then applying P- or P/Q-balancing to this equation ensures the existence of a matrix $\mathcal{W}$ (depending on the method), so that

$$\sup_{t\in[0,T]} \mathbb{E}\,\|y(t) - y_r(t)\|_2 \leq \left(\operatorname{tr}\left(\Sigma_2 \mathcal{W}\right)\right)^{\frac{1}{2}} \|u\|_T\,, \tag{6.38}$$

where y_r is the output of (6.35). This was proved in [14, 99]. Notice that $\mathcal{W}$ is independent of the diagonalized Gramian Σ and Σ_2 contains the truncated eigenvalues only, see (6.31). It is important to mention that [99] just looked at the P-balancing case if $C = I$ but (6.38) holds for general C, too. Let us now look at ROM (6.33) and check for a bound like (6.38). First of all, we need to assume stability preservation in (6.33) for the existence of a bound. This preservation is not naturally given according to Example 6.11 in contrast to (6.35).

Theorem 6.12 Given that we consider the Stratonovich setting of $H = 1/2$. Let system (6.1) with output y and $x_0 = 0$ be mean square asymptotically stable. Moreover, suppose that (6.33) with output y_r and $x_{0,r} = 0$ preserves this stability. In case (6.33) is based on either P-balancing or P/Q-balancing according to Definition 6.9 (iii), we have

$$\sup_{t\in[0,T]} \mathbb{E}\,\|y(t) - y_r(t)\|_2 \leq \left(\operatorname{tr}\left(\Sigma_1(\hat{Q}_1^\top - Q_r)\Delta_{N,11}\right) + \operatorname{tr}\left(\Sigma_2 \mathcal{W}\right)\right)^{\frac{1}{2}} \|u\|_T\,, \tag{6.39}$$

where

$$\mathcal{W} := C_2^\top C_2 + 2A_{N,12}^\top \hat{Q}_2 + \sum_{i=1}^{q} N_{i,12}^\top \left(2\hat{Q}\begin{bmatrix} N_{i,12} \\ N_{i,22} \end{bmatrix} - Q_r N_{i,12}\right).$$

The above matrices result from the partition (6.32) of the balanced realization (6.27) of (6.1) and $A_{SN} = \left[\begin{smallmatrix} A_{N,11} & A_{N,12} \\ A_{N,21} & A_{N,22} \end{smallmatrix}\right]$, where $A_{SN} = A_S + \frac{1}{2}\sum_{i=1}^{q} N_{i,S}^2$. Furthermore, we set

$\Delta_{N,11} = \sum_{i=1}^{q} N_{i,12} N_{i,21}$ and assume that $\hat{Q} = [\,\hat{Q}_1 \;\hat{Q}_2\,]$ and Q_r are the unique solutions to

$$(A_{N,11} - \frac{1}{2}\Delta_{N,11})^\top \hat{Q} + \hat{Q} A_{SN} + \sum_{i=1}^{q} N_{i,11}^\top \hat{Q} N_{i,S} = -C_1^\top C_S, \tag{6.40}$$

$$(A_{N,11} - \frac{1}{2}\Delta_{N,11})^\top Q_r + Q_r (A_{N,11} - \frac{1}{2}\Delta_{N,11}) + \sum_{i=1}^{q} N_{i,11}^\top Q_r N_{i,11} = -C_1^\top C_1. \tag{6.41}$$

The bound in (6.39) further involves the matrix $\Sigma = \begin{bmatrix} \Sigma_1 & \\ & \Sigma_2 \end{bmatrix}$ of either eigenvalues of P (P-balancing) or square roots of eigenvalues of PQ (P/Q-balancing). In particular, Σ_2 represents the truncated eigenvalues of the system.

Proof. We have to compare the outputs of (6.27) and (6.33). This is the same like calculating the error between the corresponding Itô versions of these systems. In the Itô equation of (6.27), A_S is replaced by A_{SN} and the Itô form of (6.33) involves

$$A_{11} + \frac{1}{2}\sum_{i=1}^{q} N_{i,11}^2 = A_{N,11} - \frac{1}{2}\Delta_{N,11},$$

instead of A_{11}. Since either P-balancing or P/Q-balancing is used, we know that at least one of the Gramians is diagonal, i.e., $P = \Sigma$ (see Proposition 6.10). Since we are in the case of $H = 1/2$, we also know the relation to Lyapunov equations by Section 6.2.2.2, so that we obtain

$$A_{SN}\Sigma + \Sigma A_{SN}^\top + \sum_{i=1}^{q} N_{i,S}\Sigma N_{i,S}^\top = -B_S B_S^\top. \tag{6.42}$$

In the Itô setting, an error bound has been established in [14]. Applying this result yields

$$\sup_{t\in[0,T]} \mathbb{E}\,\|y(t) - y_r(t)\|_2 \le \left(\operatorname{tr}(C_S \Sigma C_S^\top) + \operatorname{tr}(C_1 P_r C_1^\top) - 2\operatorname{tr}(C_S \hat{P} C_1^\top)\right)^{\frac{1}{2}} \|u\|_T\,. \tag{6.43}$$

The reduced system Gramian P_r as well as the mixed Gramian $\hat{P}$ exist due to the assumption that stability is preserved in the reduced system. They can be defined as the unique solutions of

$$(A_{N,11} - \frac{1}{2}\Delta_{N,11}) P_r + P_r (A_{N,11} - \frac{1}{2}\Delta_{N,11})^\top + \sum_{i=1}^{q} N_{i,11} P_r N_{i,11}^\top = -B_1 B_1^\top, \tag{6.44}$$

$$A_{SN}\hat{P} + \hat{P}(A_{N,11} - \frac{1}{2}\Delta_{N,11})^\top + \sum_{i=1}^{q} N_{i,S}\hat{P} N_{i,11}^\top = -B_S B_1^\top. \tag{6.45}$$

Using the partitions of A_{SN} and the other matrices in (6.32), we evaluate the first r

columns of (6.42) to obtain

$$-B_S B_1^\top = A_{SN} \begin{bmatrix} \Sigma_1 \\ 0 \end{bmatrix} + \Sigma \begin{bmatrix} A_{N,11}^\top \\ A_{N,12}^\top \end{bmatrix} + \sum_{i=1}^{q} N_{i,S} \Sigma \begin{bmatrix} N_{i,11}^\top \\ N_{i,12}^\top \end{bmatrix} \tag{6.46}$$
$$= \begin{bmatrix} A_{N,11} \\ A_{N,21} \end{bmatrix} \Sigma_1 + \begin{bmatrix} \Sigma_1 A_{N,11}^\top \\ \Sigma_2 A_{N,12}^\top \end{bmatrix} + \sum_{i=1}^{q} \left(\begin{bmatrix} N_{i,11} \\ N_{i,21} \end{bmatrix} \Sigma_1 N_{i,11}^\top + \begin{bmatrix} N_{i,12} \\ N_{i,22} \end{bmatrix} \Sigma_2 N_{i,12}^\top \right).$$

Using the properties of the trace, we find the relation $\operatorname{tr}(C\hat{P}C_1^\top) = \operatorname{tr}(\hat{Q}B_S B_1^\top)$ between the mixed Gramians satisfying (6.40) and (6.45). In more detail, one can find this relation by inserting (6.45) into $\operatorname{tr}(\hat{Q}\tilde{B}B_1^\top)$ and exploiting that two matrices can be switched in the trace of a product of both without changing the result. We insert (6.46) into this relation giving us

$$\begin{aligned} -\operatorname{tr}(C_S \hat{P} C_1^\top) &= \operatorname{tr}\left(\hat{Q} \left[\begin{bmatrix} A_{N,11} \\ A_{N,21} \end{bmatrix} \Sigma_1 + \begin{bmatrix} \Sigma_1 A_{N,11}^\top \\ \Sigma_2 A_{N,12}^\top \end{bmatrix} \right.\right. \\ &\quad \left.\left. + \sum_{i=1}^{q} \begin{bmatrix} N_{i,11} \\ N_{i,21} \end{bmatrix} \Sigma_1 N_{i,11}^\top + \begin{bmatrix} N_{i,12} \\ N_{i,22} \end{bmatrix} \Sigma_2 N_{i,12}^\top \right] \right). \\ &= \operatorname{tr}\left(\Sigma_1 \left[\hat{Q} \begin{bmatrix} A_{N,11} \\ A_{N,21} \end{bmatrix} + (A_{N,11} - \frac{1}{2}\Delta_{N,11})^\top \hat{Q}_1 + \sum_{i=1}^{q} N_{i,11}^\top \hat{Q} \begin{bmatrix} N_{i,11} \\ N_{i,21} \end{bmatrix} \right] \right) \\ &\quad + \frac{1}{2} \operatorname{tr}\left(\Sigma_1 \Delta_{N,11}^\top \hat{Q}_1 \right) + \operatorname{tr}\left(\Sigma_2 \left[A_{N,12}^\top \hat{Q}_2 + \sum_{i=1}^{q} N_{i,12}^\top \hat{Q} \begin{bmatrix} N_{i,12} \\ N_{i,22} \end{bmatrix} \right] \right). \end{aligned}$$

The first r columns of (6.40) give us

$$\hat{Q} \begin{bmatrix} A_{N,11} \\ A_{N,21} \end{bmatrix} + (A_{N,11} - \frac{1}{2}\Delta_{N,11})^\top \hat{Q}_1 + \sum_{i=1}^{q} N_{i,11}^\top \hat{Q} \begin{bmatrix} N_{i,11} \\ N_{i,21} \end{bmatrix} = -C_1^\top C_1$$

and hence

$$\begin{aligned} -\operatorname{tr}(C_S \hat{P} C_1^\top) &= -\operatorname{tr}(C_1 \Sigma_1 C_1^\top) + \frac{1}{2} \operatorname{tr}\left(\Sigma_1 \Delta_{N,11}^\top \hat{Q}_1 \right) \\ &\quad + \operatorname{tr}\left(\Sigma_2 \left[A_{N,12}^\top \hat{Q}_2 + \sum_{i=1}^{q} N_{i,12}^\top \hat{Q} \begin{bmatrix} N_{i,12} \\ N_{i,22} \end{bmatrix} \right] \right). \end{aligned}$$

We exploit this for the bound in (6.43) and further find that

$$\operatorname{tr}(C_S \Sigma C_S^\top) = \operatorname{tr}(C_1 \Sigma_1 C_1^\top) + \operatorname{tr}(C_2 \Sigma_2 C_2^\top).$$

Thus, we have

$$\begin{aligned}\operatorname{tr}(C_S\Sigma C_S^\top)+\operatorname{tr}(C_1P_rC_1^\top)-2\operatorname{tr}(C_S\hat{P}C_1^\top)\\ =\operatorname{tr}(C_1(P_r-\Sigma_1)C_1^\top)+\operatorname{tr}\left(\Sigma_1\Delta_{N,11}^\top\hat{Q}_1\right)\\ +\operatorname{tr}\left(\Sigma_2\left[C_2^\top C_2+2A_{N,12}^\top\hat{Q}_2+2\sum_{i=1}^q N_{i,12}^\top\hat{Q}\begin{bmatrix}N_{i,12}\\N_{i,22}\end{bmatrix}\right]\right).\end{aligned}\tag{6.47}$$

Now, we analyze $P_r-\Sigma_1$. The left upper $r\times r$ block of (6.42) fulfills

$$\begin{aligned}&(A_{N,11}-\frac{1}{2}\Delta_{N,11})\Sigma_1+\Sigma_1(A_{N,11}-\frac{1}{2}\Delta_{N,11})^\top+\sum_{i=1}^q N_{i,11}\Sigma_1N_{i,11}^\top\\ &=-B_1B_1^\top-\sum_{i=1}^q N_{i,12}\Sigma_2N_{i,12}^\top-\frac{1}{2}\Delta_{N,11}\Sigma_1-\Sigma_1\frac{1}{2}\Delta_{N,11}^\top.\end{aligned}$$

Comparing this with (6.44) yields

$$\begin{aligned}&(A_{N,11}-\frac{1}{2}\Delta_{N,11})(P_r-\Sigma_1)+(P_r-\Sigma_1)(A_{N,11}-\frac{1}{2}\Delta_{N,11})^\top+\sum_{i=1}^q N_{i,11}(P_r-\Sigma_1)N_{i,11}^\top\\ &=\sum_{i=1}^q N_{i,12}\Sigma_2N_{i,12}^\top+\frac{1}{2}\Delta_{N,11}\Sigma_1+\Sigma_1\frac{1}{2}\Delta_{N,11}^\top.\end{aligned}$$

Therefore, using (6.41), we obtain that

$$\begin{aligned}&\operatorname{tr}(C_1(P_r-\Sigma_1)C_1^\top)=\operatorname{tr}((P_r-\Sigma_1)C_1^\top C_1)\\ &=-\operatorname{tr}\left((P_r-\Sigma_1)[(A_{N,11}-\frac{1}{2}\Delta_{N,11})^\top Q_r+Q_r(A_{N,11}-\frac{1}{2}\Delta_{N,11})+\sum_{i=1}^q N_{i,11}^\top Q_rN_{i,11}]\right)\\ &=-\operatorname{tr}\Big([(A_{N,11}-\frac{1}{2}\Delta_{N,11})(P_r-\Sigma_1)+(P_r-\Sigma_1)(A_{N,11}-\frac{1}{2}\Delta_{N,11})^\top\\ &+\sum_{i=1}^q N_{i,11}(P_r-\Sigma_1)N_{i,11}^\top]Q_r\Big)=-\operatorname{tr}\left([\sum_{i=1}^q N_{i,12}\Sigma_2N_{i,12}^\top+\Delta_{N,11}\Sigma_1]Q_r\right)\\ &=-\operatorname{tr}\left([\Sigma_2\sum_{i=1}^q N_{i,12}^\top Q_rN_{i,12}+\Sigma_1Q_r\Delta_{N,11}]\right).\end{aligned}$$

Inserting this into (6.47) concludes the proof. □

Even if stability is preserved in (6.33), we cannot ensure a small error if we only know that Σ_2 has small diagonal entries. This is the main conclusion from Theorem 6.12 as the bound depends on a matrix Σ_1 with potentially very large diagonal entries reflecting the dominant eigenvalues associated with the key modes of the system. This is an indicator that there are cases in which (6.33) might perform poorly. The correction term $\frac{1}{2}\Delta_{N,11}=\frac{1}{2}\sum_{i=1}^q N_{i,12}N_{i,21}$ in (6.35) ensures that the expression in (6.39) that depends on $\Delta_{N,11}$ is canceled out. This leads to the bound in (6.38). At this point, let us also refer to the error analysis for P_T/Q_T-balancing for $H=1/2$ in the Ito setting in Section 5.2 of Chapter 5.

Let us conclude this Chapter by conducting a numerical experiment.

6.4 Numerical Results

In this section, the reduced order techniques that are based on balancing and lead to a system like in (6.33) or (6.35) are applied to an example. In detail, stochastic heat equation driven by fractional Brownian motions with different Hurst parameters H are considered and formally discretized in space. This discretization yields a system of the form (6.1) which we reduce concerning the state space dimension. Before we provide details on the model reduction procedure, let us briefly describe the time-discretization that is required here as well. We use an implicit scheme, because spatial discretizations of the underlying stochastic partial differential equations are stiff.

6.4.1 Time Integration

The stochastic differential equations (6.1), (6.33) and (6.35) can be numerically solved using various general-purpose stochastic numerical schemes. As previously mentioned, stiff differential equations pose significant challenges for numerical simulation in both deterministic and stochastic systems. Implicit methods are generally more effective than explicit methods for solving stiff problems. This work aims to utilize an implicit numerical method well-suited for addressing stiff stochastic differential equations. The stochastic implicit midpoint method will be the focus throughout the numerical section. For a more detailed discussion of Runge-Kutta methods based on increments of the driver, refer to Section 2.8.2 for $(H > 1/2)$ and [104] for $(H = 1/2)$. In particular, we rewrite the stochastic implicit midpoint method, as defined in (6.48), as follows:

$$x_{k+1} = x_k + \left[A\left(\frac{x_k + x_{k+1}}{2}\right) + Bu\left(t_k + \frac{\Delta t}{2}\right)\right]\Delta t + \sum_{i=1}^{q} N_i \left(\frac{x_k + x_{k+1}}{2}\right) \Delta W_{i,k}^H \quad (6.48)$$

when applying it to (6.1), where Δt denotes the time step related to equidistant grid points t_k. Moreover, we define $\Delta W_{i,k}^H = W_i^H(t_k + 1) - W_i^H(t_k)$. The midpoint method converges with almost sure/L^p-rate (arbitrary close to) $2H - 1/2$ for $H \in [1/2, 1)$. Before proceeding to the numerical experiments, let us briefly sketch the POD scheme that we use as a reference method within the numerics.

6.4.2 POD-Based Method

The proper orthogonal decomposition (POD) method is a data-driven strategy for the reduction of large-scale models that is based on the singular value decomposition (SVD) of snapshot matrices. However, POD techniques for stochastic differential equations driven by fBm have not been studied yet. For the convenience of the readers, a brief explanation of the POD method is provided here. For a more detailed discussion, please refer to Section 3.2.4.5.

The idea is to sample the solution for fixed u and x_0 to obtain matrices

$$Z_j = \left[x\left(t_1, \omega_j\right), x\left(t_2, \omega_j\right), \ldots, x\left(t_{\mathcal{N}}, \omega_j\right)\right], \quad \text{for} \quad \omega_j \in \Omega, \quad j = 1, \ldots, \mathcal{N}_s,$$

where $\mathcal{N}, \mathcal{N}_s > 0$ are the number of considered time points and samples. We introduce a data matrix $Z := [X_0, Z_1, Z_2, \ldots, Z_{\mathcal{N}_s}]$ and calculate its SVD:

$$Z = \begin{pmatrix} V & \star \end{pmatrix} \begin{pmatrix} \Sigma_Z & \\ & \star \end{pmatrix} \begin{pmatrix} U^\top \\ \star \end{pmatrix}.$$

The dominant subspace is identified by considering only singular vectors associated to the singular values in Σ_Z above a certain threshold. We end up with a POD-based reduced system (6.4), where the projection matrix $V = W$ consists of vectors associated to large singular values of the snapshot matrix. Instead of using POD for (6.1) directly, we can also apply it to subsystems (6.13) and (6.14). Subsequently, we find an approximation for (6.1) by the sum of the reduced subsystems.

6.4.3 Dimension Reduction for a Stochastic Heat Equation

We recall the stochastic heat equation from Example 4.20 $\left(t \in [0,1], \zeta \in [0,\pi]^2\right)$:

$$\begin{aligned} \frac{\partial \mathcal{X}(t,\zeta)}{\partial t} &= a\Delta \mathcal{X}(t,\zeta) + 1_{[\frac{\pi}{4},\frac{3\pi}{4}]^2}(\zeta)u(t) + \gamma e^{-|\zeta_1 - \frac{\pi}{2}| - \zeta_2}\mathcal{X}(t,\zeta) \circ \frac{\partial W^H(t)}{\partial t}, \\ \mathcal{X}(t,\zeta) &= 0, \quad t \in [0,1], \quad \zeta \in \partial[0,\pi]^2, \quad \text{and} \quad \mathcal{X}(0,\zeta) = b\cos(\zeta), \end{aligned} \tag{6.49}$$

where $a, b > 0$, $\gamma \in \mathbb{R}$ and a single input meaning that $m = 1$. Instead of considering the entire state, we focus on a finite number of observations, specifically the average temperature in the non-heated region, given by:

$$\mathcal{Y}(t) = \frac{4}{3\pi^2} \int_{[0,\pi]^2 \setminus [\frac{\pi}{4},\frac{3\pi}{4}]^2} \mathcal{X}(t,\zeta) d\zeta. \tag{6.50}$$

We approximate $\mathcal{Y}(t)$ using the output of the Galerkin solution discussed in Chapter 4. The Galerkin approximation for this specific case is detailed in Example 4.20 and is given by:

$$y_n(t) = Cx(t)$$

where $C^T = (\mathcal{C}u_k)_{k=1,\ldots,n}$. Here, C is the integral operator defined on the right-hand side of equation (6.50) and $(u_k)_{k=1,\ldots,n}$ are the eigenvectors of the Laplace operator with Dirichlet boundary conditions. The corresponding state x is expressed as (refer to Example 4.20):

$$x(t) = \int_0^t Ax(s) + Bu(s)\mathrm{d}s + \int_0^t Nx(s) \circ \mathrm{d}W^H(s),$$

where:

- $A = \mathrm{diag}(0, -1, -1, -2, \ldots)$ and $U = L^2([0,\pi]^2)$,
- $N = \left(\left\langle e^{-|\cdot - \frac{\pi}{2}| - \cdot} u_i, u_k \right\rangle_U\right)_{k,i=1,\ldots,n}$,
- $B = \left(\left\langle 1_{[\frac{\pi}{4},\frac{3\pi}{4}]^2}(\cdot), u_k \right\rangle_U\right)_{k=1,\ldots,n}$.

In the following, we fix $a = 0.2$, $b = 1$ and set $n = 1024$. We investigate two cases. These are $H = 0.5$ and $H = 0.75$. In the following, we explain the particular dimension reduction techniques for each scenario.

Case H = 0.75 : We have pointed out in Section 6.2.2.2 that Gramians P_T and Q_T (or their limits P and Q) are hard to compute for $H > 1/2$, since a link of these matrices to ordinary differential or algebraic equations is unknown. Therefore, we solely consider empirical Gramians discussed in Section 6.2.2.2 for $H = 0.75$. In fact, $\bar{P}_T$ is available by sampling the solution of (6.14), whereas $\bar{Q}_T$ seems computationally much more involved. For that reason, we apply $\bar{P}_T$-balancing (see Definition 6.9 (iv)) to system (6.1) that obtained from the above heat equation. This results in (6.27) which is truncated in order to find the reduced equation (6.33). Two other related approaches are conducted in this section as well.

- We apply the same $\bar{P}_T$-balancing procedure to subsystems (6.13) and (6.14), i.e., $\bar{P}_{u,T}$-balancing is used for (6.13) and $\bar{P}_{x_0,T}$-balancing for (6.14), compared to (6.22). The sum of the resulting reduced order systems is then used to approximate (6.1). Refer to this second ansatz as splitting-based $\bar{P}_T$-balancing.

- Another empirical dimension reduction technique, as discussed above, is the POD method. In this method, the solution space of (6.1) is learned using samples, which are potentially based on various initial states x_0 and controls u. Note that the snapshot matrices are computed from a small set of x_0 and u to provide a POD-based reduced system (6.4) that perform well for a larger number of x_0 and u. In this approach, we apply the POD scheme to the subsystems (6.13) and (6.14) and approximate (6.1) by the sum of the reduced subsystems. We refer to this method as splitting-based POD.

Case H = 0.5: Similar techniques are exploited for the Stratonovich setting. However, we have the advantage that P_T and Q_T can be computed from matrix equations; see (6.24) and (6.26). Still, these equations are difficult to solve. Therefore, we use the sampling and variance reduction-based schemes proposed in Section 5.3 of Chapter 5 in order to solve them. Due to the availability of both Gramians, we apply P_T/Q_T-balancing, see Definition 6.9 (ii), instead of the procedure based on diagonalizing $\bar{P}_T$. However, we truncate differently, i.e., the reduced system (6.35) is used instead due to the drawbacks of (6.33) pointed out in Section 6.3.3.2 when $H = 0.5$. The splitting-based P_T/Q_T-balancing is defined the same way. It is the technique, where $P_{u,T}/Q_T$-balancing is conducted for (6.13) and $P_{x_0,T}/Q_T$-balancing is exploited for (6.14) to obtain reduced systems of the form (6.35) for each subsystem. Again, we use a splitting-based POD scheme according to Section 6.4.2 for $H = 0.5$.

For the discretization in time, the stochastic midpoint method (6.48), stated in Section 6.4.1, is employed here, where the number of time steps is $\mathcal{N} = 100$. Moreover, all empirical objects are calculated based on $\mathcal{N}_s = 10^3$ samples. The error between the reduced systems and the original model is computed for the control $u(t) = \sqrt{\frac{2}{\pi}}\sin(t)$, where the reduction error is measured by the quantity $\mathcal{R}_E = \dfrac{\sup_{t\in[0,1]} \mathbb{E}\|y(t) - y_r(t)\|_2}{\sup_{t\in[0,1]} \mathbb{E}\|y(t)\|_2}$.

In the case of $H = 0.5$, Figure 6.1 illustrates that splitting-based P_T/Q_T-balancing (2. Gramian), which was described just above and P_T/Q_T-balancing (1. Gramian), which

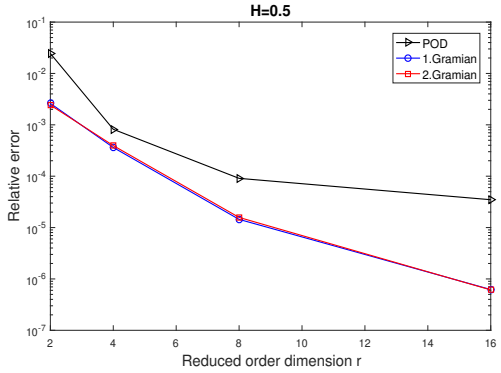

Figure 6.1: $\mathcal{R}_E$ for three approaches with Hurst parameters $H = 0.5$.

Table 6.1: $\mathcal{R}_E$ for $r \in \{2, 4, 8, 16\}$ and $H = 0.5$.

r	POD	1. Gramian	2. Gramian
2	$2.4471e-02$	$2.6131e-03$	$2.4251e-03$
4	$8.1898e-04$	$3.6254e-04$	$3.9410e-04$
8	$9.0777e-05$	$1.4427e-05$	$1.5756e-5$
16	$3.4842e-05$	$6.2128e-07$	$6.1161e-07$

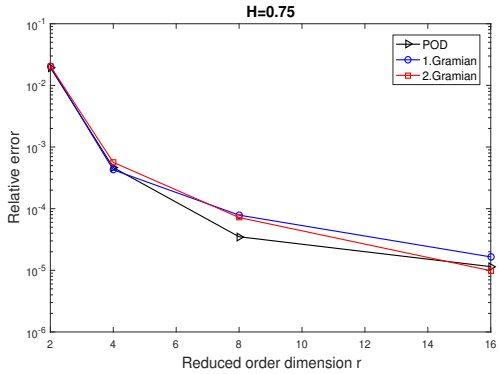

Figure 6.2: $\mathcal{R}_E$ for three approaches with Hurst parameters $H = 0.75$.

Table 6.2: $\mathcal{R}_E$ for $r \in \{2, 4, 8, 16\}$ and $H = 0.75$.

r	POD	1. Gramian	2. Gramian
2	$1.9428e-02$	$2.0531e-02$	$2.0543e-02$
4	$4.6419e-04$	$4.2626e-04$	$5.6448e-04$
8	$3.5032e-05$	$7.8586e-05$	$7.1846e-05$
16	$1.1479e-05$	$1.652e-05$	$9.8581e-06$

generate very similar results, produces notably better outcomes compared to the splitting-based POD method. The worst case errors of the plot are also state in the associated Table 6.1.

On the other hand, the Young setting in which we have $H = 0.75$ presents a different scenario. Figure 6.2 demonstrates that splitting-based POD exhibits a better performance compared to splitting-based $\bar{P}_T$-balancing (2. Gramian) and the usual $\bar{P}_T$-balancing (1. Gramian) methods, except when the reduced dimension is 16. Surprisingly, for $r = 16$, the 2. Gramian method yields better results compared to the POD method. It is worth noting that both empirical Gramian methods provide similar outcomes, which is an indicator for a nearly identical reduction potential for both subsystems (6.13) and (6.14). Note that the error of the plot can be found in Table 6.2.

For both, $H = 0.5$ and $H = 0.75$ an enormous reduction potential can be observed, meaning that small dimensions r lead to accurate approximations. According to Remark 6.7 this is known a-priori by the strong decay of certain eigenvalues associated with the system Gramians, since small eigenvalues indicate variables of low relevance. Given $H = 0.75$, Figure 6.3 shows the eigenvalues of $\bar{P}_T$ (1. Gramian), the sum eigenvalues of $\bar{P}_{u,T}$ and $\bar{P}_{x_0,T}$ (2. Gramian) as well as the sum of the singular values corresponding to the POD snapshot matrices of subsystems (6.13) and (6.14). Similar types of algebraic values

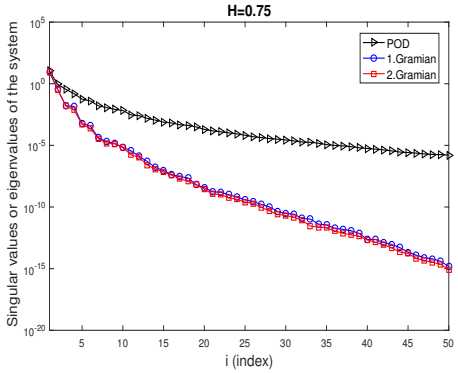

Figure 6.3: First 50 POD singular values or eigenvalues associated to $\bar{P}_T$ for $H = 0.75$.

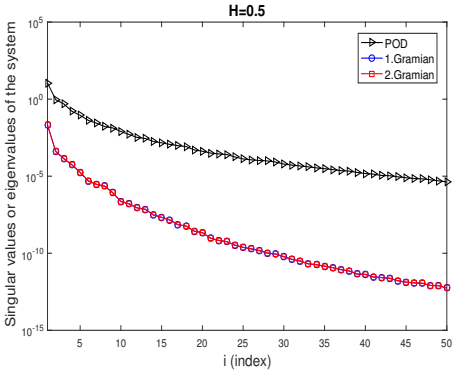

Figure 6.4: First 50 POD singular values or eigenvalues associated to P_T/Q_T for $H = 0.5$.

are considered for $H = 0.5$ in Figure 6.4. Here, square roots of eigenvalues of $P_T Q_T$ (1. Gramian) or the sum of square roots of eigenvalues of $P_{u,T} Q_T$ and $P_{x_0,T} Q_T$ (2. Gramian) are depicted. The large number of small eigenvalues (or singular values) explains why small errors could be achieved in our simulations.

Conclusions

This dissertation develops advanced techniques in model order reduction for stochastic systems driven by both standard and fractional Brownian motion. The focus is on addressing theoretical challenges and proposing practical computational frameworks to reduce the complexity of high-dimensional stochastic systems. The following is a summary of the key contributions presented in each chapter.

The introductory chapter, Chapter 1, establishes the mathematical foundation of the dissertation. It covers essential concepts in stochastic calculus, including stochastic processes, stochastic integrals, and the Itô formula. These preliminaries set the stage for subsequent chapters by providing a robust framework for analyzing stochastic systems.

Chapter 2 delves into fractional Brownian motion, highlighting its distinct properties such as self-similarity and long-range dependence, governed by the Hurst parameter H. This chapter discusses various integration techniques suitable for fBm, including Young integration and Malliavin calculus, and explores numerical methods for solving stochastic differential equations driven by fBm.

Chapter 3 focuses on model order reduction for deterministic and stochastic linear dynamical systems. The chapter begins with an exploration of Lyapunov operators and stability in deterministic systems, extending these concepts to stochastic systems. It introduces Gramian matrices as tools for characterizing dominant subspaces, which form the basis for MOR techniques. Special attention is given to the Balanced Truncation method and the Proper Orthogonal Decomposition method, which are applied to reduce the computational complexity of high-dimensional systems while retaining their essential dynamics.

Chapter 4 introduces linear stochastic partial differential equations (SPDEs) in an abstract evolution equation framework. We focused on stochastic heat equations with fractional noise, approximated using a spectral Galerkin scheme. The convergence of the spectral Galerkin solution to the mild solution of the corresponding SPDE was demonstrated. This discretization resulted in high-dimensional linear SDEs, motivating the extension of balancing-based model order reduction to mean square asymptotically stable controlled stochastic systems.

Chapter 5 addresses model order reduction for large-scale linear stochastic systems, such as spatially discretized stochastic partial differential equations, where asymptotic stability is often not guaranteed due to noise. The focus is on developing Gramian-based MOR schemes suitable for unstable systems. These Gramians are constructed to identify dominant subspaces and can be computed using Lyapunov equations, which require covariance information. To address this, efficient sampling-based methods with variance reduction are proposed, along with deterministic approximations of covariance functions. An error bound is derived, providing a criterion for selecting the reduced dimension of the system. Numerical experiments demonstrate the effectiveness of the proposed MOR techniques.

Finally, in Chapter 6, we study large-scale linear systems driven by fBm with Hurst parameter $H \in [1/2, 1)$. These equations are interpreted in the sense of Young ($H > 1/2$)

or Stratonovich ($H = 1/2$), with Young equations particularly suited to capture memory effects in real-world phenomena. Addressing the computational challenges of high-dimensional systems, we investigate model reduction techniques for both settings. We analyze fundamental solutions of the systems, introducing empirical reduced-order methods based on snapshots (e.g., POD method) or approximated Gramians, which identify dominant subspaces. For $H > 1/2$, the absence of links between Gramians and algebraic equations complicates the computation, so we propose empirical Gramians derived from simulation data. Projection-based models are constructed using dominant subspaces, though such projections may not preserve stability in Stratonovich settings. To address this, we propose an improved reduced-order model for $H = 1/2$. The proposed techniques are validated through numerical experiments on large-scale stochastic differential equations derived from spatially discretized fractional stochastic PDEs. This study provides valuable insights into reduced-order methods for stochastic systems with fractional noise, enabling more efficient computational strategies for practical applications.

In conclusion, this dissertation makes significant contributions to the field of MOR by developing innovative techniques tailored to stochastic systems driven by sBM and fBm. These methods provide a unified approach to dimension reduction, applicable to a broad range of stochastic modeling scenarios.

Summary of Contributions

This thesis is partially based on the papers that are listed below.

Preprint

N. Jamshidi and M. Redmann, *(Empirical) Gramian-based dimension reduction for stochastic differential equations driven by fractional Brownian motion*, Applied Probability Journals, submitted.

Sections 6.2, 6.3 and 6.4 are based on this paper. However, in this thesis the results are explained more detailed.

Publications

M. Redmann and N. Jamshidi, *Gramian-based model reduction for unstable stochastic systems*, Mathematics of Control, Signals, and Systems, vol. 34, 2022, pp. 855–881.

Results of this paper enter in Chapter 5.

N. Jamshidi and M. Redmann, *Sampling-based model order reduction for stochastic differential equations driven by fractional Brownian motion*, Proceedings in Applied Mathematics and Mechanics, vol. 23, 2023, pp. 1–6.

Section 6.4.2 are based on this paper.

Breakdown of Contributions

My individual contributions to each of the papers are as follows:

- **"Gramian-based model reduction for unstable stochastic systems"**:
  The core idea for this paper was proposed by Prof. Redmann, and we collaboratively developed the theoretical framework. However, the majority of the theoretical contributions were led by Prof. Redmann. I was responsible for writing the experimental code.
- **"Sampling-based model order reduction for stochastic differential equations driven by fractional Brownian motion"**:
  This paper focused on developing model order reduction schemes for large-scale systems driven by fractional Brownian motions. The concept was initiated by Prof. Redmann. I contributed by applying the Proper Orthogonal Decomposition (POD) method in this context and comparing it with the Gramian-based approach. I was

responsible for implementing the experimental code, and Prof. Redmann and I co-authored the paper together.

- **"(Empirical) Gramian-based dimension reduction for stochastic differential equations driven by fractional Brownian motion"**:
  This work aimed to extend the results from the previous paper on sampling-based MOR for stochastic differential equations driven by fBms to a more general case. Prof. Redmann incorporated the case where $H = \frac{1}{2}$ (standard Brownian motion), while I conducted the numerical experiments. I also wrote the section on fractional Brownian motion, while Prof. Redmann contributed the Stratonovich calculus for the standard Brownian motion case and revised the entire paper.

A Norms

A.1 Norms of Vectors and Matrices in Finite Dimensions

Consider $V \to \mathbb{R}$ as a linear space over the real or complex number field. A norm on V is a function $v : V \to \mathbb{R}$ such that the following properties are satisfied:

- Strictly positive: $v(x) \geq 0$ for all $x \in V$ and $v(x) = 0$ if and only if $x = 0$.
- Satisfies the triangle inequality: $v(x + y) \leq v(x) + v(y)$ for all $x, y \in V$.
- Exhibits positive homogeneity: $v(\alpha x) = |\alpha| v(x)$ for any scalar $\alpha \in \mathbb{C}$ and $x \in V$.

For a vector $x = (x_1, \ldots, x_n)^\top \in \mathbb{C}^n$, the norm, often referred to as the Hölder or p-norm, is articulated as:

$$\|x\|_p = \begin{cases} \left(\sum_{i=1}^n |x_i|^p\right)^{\frac{1}{p}}, & 1 \leq p < \infty, \\ \max\limits_{1 \leq i \leq n} |x_i|, & p = \infty. \end{cases}$$

Matrix norms of significance are those formulated from the vector p-norm specified above. Specifically, for a matrix $A = (a_{ij}) \in \mathbb{C}^{n \times m}$, the corresponding induced p-norm is:

$$\|A\|_p = \sup_{x \neq 0} \frac{\|Ax\|_p}{\|x\|_p},$$

For $p = 1, 2, \infty$, the matrix norms are expressed as follows:

$$\begin{aligned} \|A\|_1 &= \max_{1 \leq j \leq m} \sum_{i=1}^n |a_{ij}|, \\ \|A\|_2 &= \sqrt{\max(\lambda(A^* A))}, \\ \|A\|_\infty &= \max_{1 \leq i \leq n} \sum_{j=1}^m |a_{ij}|, \end{aligned}$$

where $\lambda(A^* A)$ signifies the maximal eigenvalue of the positive-semidefinite matrix $A^* A$, and A^* is the conjugate transpose of the matrix A.

Besides the induced matrix norms, there are alternative matrix norms, such as the Schatten p-norms, which are invariant under unitary transformations. For the formulation of these norms, we reference the singular value decomposition.

A.1.1 Singular Value Decomposition

Singular value decomposition (SVD), as referenced in [81], applied to a matrix $A \in \mathbb{C}^{n\times m}$, ensures the existence of non-negative real numbers $\{\sigma_i\}_{i=1}^r$ and unitary matrices

$$\begin{aligned} U &= [u_1 \quad \dots \quad u_n] \in \mathbb{C}^{n\times n}, \quad UU^* = I_n \\ V &= [v_1 \quad \dots \quad v_m] \in \mathbb{C}^{m\times m}; \quad VV^* = I_m \end{aligned}$$

where I_n and I_m denote the identity matrices of dimensions $n\times n$ and $m\times m$, respectively, such that:

$$A = U\Sigma V^*,$$

where

$$\Sigma = \begin{pmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{pmatrix}$$

where Σ_1 is a diagonal matrix with non-negative real numbers on the diagonal:

$$\Sigma_1 = \begin{pmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_r \end{pmatrix},$$

here the singular values σ_i of A are arranged in a non-increasing order,i.e $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r > 0$ and $r = \operatorname{rank}(A)$. The columns of U and V are known as the left and right singular vectors of A, respectively. Moreover, the vectors $\{u_i\}_{i=1}^r$ and $\{v_i\}_{i=1}^r$ satisfy

$$Av_i = \sigma_i u_i \quad \text{and} \quad A^*u_i = \sigma_i v_i \quad \text{for} \quad i = 1, \dots, r.$$

They are eigenvectors of AA^* and A^*A, respectively, with eigenvalues $\lambda_i = \sigma_i^2$ for $i = 1, \dots, r$. The vectors $\{u_i\}_{i=r+1}^n$ and $\{v_i\}_{i=r+1}^m$ (if $r < n$ and $r < m$, respectively) are eigenvectors of AA^* and A^*A with eigenvalue 0.

A.1.2 Schatten p-norm

Consequently, for a matrix $A = (a_{ij})$ within the complex space $\mathbb{C}^{n\times m}$ with $m \leq n$, we define the Schatten p-norm, which is not derived from other norms, in the following manner:

$$\|A\|_{\mathrm{S},p} = \begin{cases} \left(\sum_{i=1}^m (\sigma_i(A))^p\right)^{\frac{1}{p}}, & 1 \leq p < \infty, \\ \sigma_{\max}(A), & p = \infty, \end{cases}$$

where $\sigma_i(A)$ are referred to as the singular values of the matrix A, or equivalently, the square roots of the i-th largest eigenvalues of the product AA^*.

For particular cases where $p = 1, 2, \infty$, the Schatten p-norms are specified as follows:

$$\|A\|_{\mathrm{S},p} = \begin{cases} \sum_{i=1}^m \sigma_i(A), & p = 1, \quad \text{(trace norm)}, \\ \left(\sum_{i=1}^m (\sigma_i(A))^2\right)^{\frac{1}{2}} = trace(A^*A), & p = 2, \quad \text{(Frobenius norm)}, \\ (\lambda_{\max}(AA^*))^{\frac{1}{2}}, & p = \infty, \quad \text{(spectral norm)}. \end{cases}$$

A.2 Banach and Hilbert Spaces

A Banach space is identified as a vector space V over the field $\mathbb{R}$ for real numbers or $\mathbb{C}$ for complex numbers equipped with a norm $\|\cdot\|$ that turns V into a complete metric space. This means that every Cauchy sequence in V has a limit in V. Hilbert spaces are spacial cases of Banach spaces with a richer structure that arises from the presence of an inner product, a mapping from $V \times V$ to $\mathbb{R}$ defined by:

$$\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R} \tag{A.1}$$
$$(x, y) \to \langle x, y \rangle \in \mathbb{R} \tag{A.2}$$

This function is characterized by the following properties:

- Strict positiveness: For all $x \in V$, $\langle x, x \rangle \geq 0$ with equality if and only if $x = 0$.
- Linearity in the first argument: For all $x, y \in V$ and scalars $\alpha, \beta \in \mathbb{R}$, it holds that

$$\langle \alpha x + \beta y, z \rangle = \alpha \langle x, z \rangle + \beta \langle y, z \rangle$$

  .
- Conjugate symmetry: For all $x, y \in V$, $\langle x, y \rangle^* = \langle y, x \rangle$.

The norm induced by this inner product on V is given by $\|x\| = \sqrt{\langle x, x \rangle}$.

Definition A.1 A Hilbert space is separable provided it contains a dense countable subset.

Theorem A.2 A Hilbert space $\mathcal{H}$ is separable if and only if it has one countable orthonormal basis, and this is equivalent to every orthonormal basis for $\mathcal{H}$ being countable.

A.2.1 Nuclear Operator

Let $\mathcal{H}$ be a separable Hilbert space with $\{e_k\}$ as a complete orthonormal basis in $\mathcal{H}$. If $T \in L_1(\mathcal{H}, \mathcal{H})$, the trace of T is defined as

$$\operatorname{Tr} T = \sum_{j=1}^{\infty} \langle T e_j, e_j \rangle_{\mathcal{H}} .$$

Proposition A.3 [24] A nonnegative operator $T \in L(\mathcal{H})$ is classified as nuclear (or trace class) operator if and only if, for any orthonormal basis $\{e_k\}$ of $\mathcal{H}$, we have

$$\sum_{j=1}^{\infty} \langle T e_j, e_j \rangle_{\mathcal{H}} < +\infty.$$

Furthermore, in this situation, the trace of T satisfies $\operatorname{Tr} T = \|T\|_1$.

A.2.2 Hilbert-Schmidt Operator

A linear continuous operator $T : V \to U$ between two Hilbert spaces V and U is called a Hilbert-Schmidt operator if

$$\|T\|_{HS}^2 := \sum_{k=1}^{\infty} \|Te_k\|_U^2 < \infty,$$

where $\{e_k\}_{k=1}^{\infty}$ is an orthonormal basis for the Hilbert space V, and $\|\cdot\|_U$ is the norm on U. The quantity $\|T\|_{HS}$ is called the Hilbert-Schmidt norm of T, and it is independent of the choice of the orthonormal basis $\{e_k\}$. The space of all Hilbert-Schmidt operators is a Hilbert space itself, denoted by $L_2(V, U)$, with the inner product

$$\langle T_1, T_2 \rangle_{HS} = \sum_{k=1}^{\infty} \langle T_1 e_k, T_2 e_k \rangle_U.$$

For a comprehensive treatment of Hilbert-Schmidt operators, see [24] .

B Kronecker Product

In this section, we explore the process of vectorizing a matrix alongside an exploration of the Kronecker product, including an examination of its attributes. This particular matrix operation simplifies various computational procedures. As such, it is utilized in deducing the controllability and observability Gramians from the Lyapunov equations, as indicated in (3.2) or (3.27), and in determining the matrices $\mathcal{K}$ and $\mathcal{K}^\top$ of a system as discussed in Remark 3.18. For an in-depth discussion on the characteristics of the Kronecker product and proof of these properties, the reader is referred to [16].

Definition B.1 Consider a matrix $A \in \mathbb{R}^{n\times m}$. The vectorization operator is defined as

$$\operatorname{vec} A = \begin{bmatrix} \operatorname{col}_1(A) \\ \vdots \\ \operatorname{col}_m(A) \end{bmatrix} \in \mathbb{R}^{nm\times 1},$$

resulting in a $nm \times 1$ column vector by consolidating the columns of A. To reconstruct A from $\operatorname{vec} A$, we utilize the inverse vectorization process:

$$A = \operatorname{vec}^{-1}(\operatorname{vec} A).$$

Further, we demonstrate the following for $C \in \mathbb{R}^{m\times p}$:

$$trace(AC) = (\operatorname{vec} C^\top) \operatorname{vec} A.$$

Definition B.2 Let $A \in \mathbb{R}^{n\times m}$ and $B \in \mathbb{R}^{p\times k}$. Then, the Kronecker product $A \otimes B$ of $A \in \mathbb{R}^{np\times mk}$ is the partitioned matrix

$$A \otimes B = \begin{bmatrix} A_{(1,1)}B & A_{(1,2)}B & \dots & A_{(1,m)}B \\ A_{(2,1)}B & A_{(2,2)}B & \dots & A_{(2,m)}B \\ \vdots & \vdots & \ddots & \vdots \\ A_{(n,1)}B & A_{(n,2)}B & \dots & A_{(n,m)}B \end{bmatrix}. \tag{B.1}$$

Unlike matrix multiplication, the Kronecker product $A \otimes B$ does not entail a restriction on either the size of A or the size of B.

The following results are immediate consequences of the definition of the Kronecker product.

Proposition B.3 [16] Let $\alpha \in \mathbb{R}$, $A \in \mathbb{R}^{n\times m}$, $B \in \mathbb{R}^{p\times k}$ and $C \in \mathbb{R}^{m\times p}$. Then,

$$A \otimes (\alpha B) = (\alpha A) \otimes B = \alpha(A \otimes B), \tag{B.2}$$
$$A \otimes B = B \otimes A, \tag{B.3}$$
$$(A \otimes B)^\top = A^\top \otimes B^\top, \tag{B.4}$$

Proposition B.4 [16] Let $A, B \in \mathbb{R}^{n\times m}$ and $C \in \mathbb{R}^{p\times k}$. Then,

$$(A + B) \otimes C = A \otimes C + B \otimes C, \tag{B.5}$$
$$C \otimes (A + B) = C \otimes A + C \otimes B. \tag{B.6}$$

Proposition B.5 [16] Let $A \in \mathbb{R}^{n\times m}$, $B \in \mathbb{R}^{p\times k}$, and $C \in \mathbb{R}^{q\times l}$. Then,

$$A \otimes (B \otimes C) = (A \otimes B) \otimes C.$$

Hence, we write $A \otimes B \otimes C$ for $A \otimes (B \otimes C)$ and $(A \otimes B) \otimes C$.

Proposition B.6 [16] Let $A \in \mathbb{R}^{n\times m}$, $B \in \mathbb{R}^{p\times k}$ and $C \in \mathbb{R}^{m\times p}$. Then,

$$\operatorname{vec}(ACB) = (B^\top \otimes A)\operatorname{vec}(C)$$

The following result demonstrates a practical alignment between traditional matrix multiplication and the Kronecker product operation.

Proposition B.7 [16] Let $A \in \mathbb{R}^{n\times m}$, $B \in \mathbb{R}^{l\times k}$, $C \in \mathbb{R}^{m\times q}$, and $D \in \mathbb{R}^{k\times p}$. Then,

$$(A \otimes B)(C \otimes D) = AC \otimes BD.$$

Corollary B.8 [17] Let $t \to A(t)$ be a continuous map from an interval I into the space of $n \times n$ matrices. Then there exist continuous functions $\lambda_1(t) \dots \lambda_n(t)$ that, for each $t \in I$, are the eigenvalues of $A(t)$.

Bibliography

[1] E. Alòs, O. Mazet, and D. Nualart, *Stochastic calculus with respect to fractional brownian motion with hurst parameter lesser than* $\frac{1}{2}$, Stochastic processes and their applications, 86 (2000), pp. 121–139.

[2] E. Alos, O. Mazet, and D. Nualart, *Stochastic calculus with respect to gaussian processes*, The Annals of Probability, 29 (2001), pp. 766–801.

[3] E. Alvarez-Lacalle, B. Dorow, J.-P. Eckmann, and E. Moses, *Hierarchical structures induce long-range dynamical correlations in written texts*, Proceedings of the National Academy of Sciences, 103 (2006), pp. 7956–7961.

[4] M. Amabili, A. Sarkar, and M. Paidoussis, *Reduced-order models for nonlinear vibrations of cylindrical shells via the proper orthogonal decomposition method*, Journal of Fluids and Structures, 18 (2003), pp. 227–250.

[5] A. C. Antoulas, *Approximation of large-scale dynamical systems*, vol. 6 of Advances in Design and Control, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2005. With a foreword by Jan C. Willems.

[6] A. C. Antoulas, D. C. Sorensen, and S. Gugercin, *A survey of model reduction methods for large-scale systems*, in Structured matrices in mathematics, computer science, and engineering, I (Boulder, CO, 1999), vol. 280 of Contemp. Math., Amer. Math. Soc., Providence, RI, 2001, pp. 193–219.

[7] R. H. Bartels and G. W. Stewart, *Solution of the matrix equation* $AX + XB = C$, Communications of the ACM, 15 (1972), pp. 820–826.

[8] A. Barth, *A finite element method for martingale-driven stochastic partial differential equations*, Commun. Stoch. Anal., 4 (2010), pp. 355–375.

[9] S. Becker and C. Hartmann, *Infinite-dimensional bilinear and stochastic balanced truncation with error bounds*, Math. Control. Signals, Syst., 31 (2019), pp. 1–37.

[10] P. Benner and T. Damm, *Lyapunov equations, energy functionals, and model order reduction of bilinear and stochastic systems*, SIAM J. Control Optim., 49 (2011), pp. 686–711.

[11] P. Benner, T. Damm, M. Redmann, and Y. R. Rodriguez Cruz, *Positive operators and stable truncation*, Linear Algebra Appl., 498 (2016), pp. 74–87.

[12] P. Benner, S. Gugercin, and K. Willcox, *A survey of projection-based model reduction methods for parametric dynamical systems*, SIAM Rev., 57 (2015), pp. 483–531.

[13] P. BENNER, M. OHLBERGER, A. COHEN, AND K. WILLCOX, *Model reduction and approximation: theory and algorithms*, SIAM, 2017.

[14] P. BENNER AND M. REDMANN, *Model Reduction for Stochastic Systems.*, Stoch PDE: Anal Comp, 3 (2015), pp. 291–338.

[15] G. BERKOOZ, P. HOLMES, AND J. L. LUMLEY, *The proper orthogonal decomposition in the analysis of turbulent flows*, Annual review of fluid mechanics, 25 (1993), pp. 539–575.

[16] D. S. BERNSTEIN, *Matrix mathematics*, Princeton University Press, Princeton, NJ, second ed., 2009. Theory, facts, and formulas.

[17] R. BHATIA, *Matrix analysis*, vol. 169 of Graduate Texts in Mathematics, Springer-Verlag, New York, 1997.

[18] F. BIAGINI, Y. HU, B. ØKSENDAL, AND T. ZHANG, *Stochastic calculus for fractional Brownian motion and applications*, Springer Science & Business Media, 2008.

[19] A. BRANDT, *Noise and vibration analysis: signal analysis and experimental procedures*, John Wiley & Sons, 2023.

[20] P. BUCHFINK, A. BHATT, AND B. HAASDONK, *Symplectic model order reduction with non-orthonormal bases*, Mathematical and Computational Applications, 24 (2019), p. 43.

[21] P. CARMONA, L. COUTIN, AND G. MONTSENY, *Stochastic integration with respect to fractional brownian motion*, in Annales de l'IHP Probabilités et statistiques, vol. 39, 2003, pp. 27–68.

[22] S. CHATURANTABUT AND D. C. SORENSEN, *Nonlinear model reduction via discrete empirical interpolation*, SIAM J. Sci. Comput., 32 (2010), pp. 2737–2764.

[23] P. CHERIDITO, *Mixed fractional Brownian motion*, Bernoulli, 7 (2001), pp. 913–934.

[24] G. DA PRATO AND J. ZABCZYK, *Stochastic equations in infinite dimensions*, vol. 44 of Encyclopedia of Mathematics and its Applications, Cambridge University Press, Cambridge, 1992.

[25] T. DAMM, *Rational matrix equations in stochastic control*, vol. 297 of Lecture Notes in Control and Information Sciences, Springer-Verlag, Berlin, 2004.

[26] T. DAMM, *Direct methods and ADI-preconditioned Krylov subspace methods for generalized Lyapunov equations.*, Numer. Linear Algebra Appl., 15 (2008), pp. 853–871.

[27] L. DECREUSEFOND AND A. S. ÜSTÜNEL, *Stochastic analysis of the fractional Brownian motion*, Potential Anal., 10 (1999), pp. 177–214.

[28] A. DEYA, A. NEUENKIRCH, AND S. TINDEL, *A Milstein-type scheme without Lévy area terms for SDEs driven by fractional Brownian motion*, Ann. Inst. Henri Poincaré Probab. Stat., 48 (2012), pp. 518–550.

[29] X.-L. Ding and D. Wang, *Regularity analysis for SEEs with multiplicative fBms and strong convergence for a fully discrete scheme*, IMA J. Numer. Anal., 44 (2024), pp. 1435–1463.

[30] T. E. Duncan, Y. Hu, and B. Pasik-Duncan, *Stochastic calculus for fractional brownian motion i. theory*, SIAM Journal on Control and Optimization, 38 (2000), pp. 582–612.

[31] T. E. Duncan, B. Maslowski, and B. Pasik-Duncan, *Semilinear stochastic equations in a hilbert space with a fractional brownian motion*, SIAM Journal on Mathematical Analysis, 40 (2009), pp. 2286–2315.

[32] T. E. Duncan, B. Pasik-Duncan, and B. Maslowski, *Fractional Brownian motion and stochastic equations in Hilbert spaces*, Stoch. Dyn., 2 (2002), pp. 225–250.

[33] D. F. Enns, *Model reduction with balanced realizations: An error bound and a frequency weighted generalization*, in The 23rd IEEE conference on decision and control, IEEE, 1984, pp. 127–132.

[34] D. Feyel and A. de La Pradelle, *On fractional Brownian processes*, Potential Anal., 10 (1999), pp. 273–288.

[35] C. Gardiner, *Stochastic methods*, vol. 4, Springer Berlin, 2009.

[36] A. M. Garsia, E. Rodemich, H. Rumsey, and M. Rosenblatt, *A real variable lemma and the continuity of paths of some gaussian processes*, Indiana University Mathematics Journal, 20 (1970), pp. 565–578.

[37] W. Gawronski and J. Juang, *Model reduction in limited time and frequency intervals*, Int. J. Syst. Sci., 21 (1990), pp. 349–376.

[38] W. Grecksch and H. Lisei, eds., *Infinite dimensional and finite dimensional stochastic equations and applications in physics*, World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, [2020] ©2020.

[39] T. Guggenberger, A. Chechkin, and R. Metzler, *Fractional brownian motion in superharmonic potentials and non-boltzmann stationary distributions*, Journal of Physics A: Mathematical and Theoretical, 54 (2021), p. 29LT01.

[40] E. Hausenblas, *Approximation for Semilinear Stochastic Evolution Equations.*, Potential Anal., 18 (2003), pp. 141–186.

[41] D. Hinrichsen and A. J. Pritchard, *Stochastic H^∞*, SIAM J. Control Optim., 36 (1998), pp. 1504–1538.

[42] P. Holmes, J. L. Lumley, G. Berkooz, and C. W. Rowley, *Turbulence, coherent structures, dynamical systems and symmetry*, Cambridge Monographs on Mechanics, Cambridge University Press, Cambridge, second ed., 2012.

[43] J. HONG, C. HUANG, AND X. WANG, *Optimal rate of convergence for two classes of schemes to stochastic differential equations driven by fractional Brownian motions*, IMA J. Numer. Anal., 41 (2021), pp. 1608–1638.

[44] Y. HU, *Integral transformations and anticipative calculus for fractional Brownian motions*, Mem. Amer. Math. Soc., 175 (2005), pp. viii+127.

[45] Y. HU, Y. LIU, AND D. NUALART, *Rate of convergence and asymptotic error distribution of Euler approximation schemes for fractional diffusions*, Ann. Appl. Probab., 26 (2016), pp. 1147–1207.

[46] ——, *Crank-Nicolson scheme for stochastic differential equations driven by fractional Brownian motions*, Ann. Appl. Probab., 31 (2021), pp. 39–83.

[47] J. C. HULL AND S. BASU, *Options, futures, and other derivatives*, Pearson Education India, 2016.

[48] T. ILIESCU, H. LIU, AND X. XIE, *Regularized reduced order models for a stochastic Burgers equation*, Int. J. Numer. Anal. Model., 15 (2018), pp. 594–607.

[49] K. ITO AND S. S. RAVINDRAN, *A reduced-order method for simulation and control of fluid flows*, J. Comput. Phys., 143 (1998), pp. 403–425.

[50] N. JAMSHIDI AND M. REDMANN, *(empirical) gramian-based dimension reduction for stochastic differential equations driven by fractional brownian motion*, arXiv preprint arXiv:2307.04614, (2023).

[51] ——, *Sampling-based model order reduction for stochastic differential equations driven by fractional brownian motion*, PAMM, 23 (2023), p. e202200109.

[52] M. S. JOSHI, *The concepts and practice of mathematical finance*, Mathematics, Finance and Risk, Cambridge University Press, Cambridge, 2003.

[53] M. KAMRANI AND N. JAMSHIDI, *Implicit Euler approximation of stochastic evolution equations with fractional Brownian motion*, Commun. Nonlinear Sci. Numer. Simul., 44 (2017), pp. 1–10.

[54] ——, *Implicit Milstein method for stochastic differential equations via the Wong-Zakai approximation*, Numer. Algorithms, 79 (2018), pp. 357–374.

[55] T. KARAGIANNIS, M. MOLLE, AND M. FALOUTSOS, *Long-range dependence ten years of internet traffic modeling*, IEEE internet computing, 8 (2004), pp. 57–64.

[56] K. KARHUNEN, *Über lineare Methoden in der Wahrscheinlichkeitsrechnung*, Ann. Acad. Sci. Fennicae Ser. A. I. Math.-Phys., 1947 (1947), p. 79.

[57] KARMESHU AND A. KRISHNAMACHARI, *Sequence variability and long-range dependence in dna: An information theoretic perspective*, in International Conference on Neural Information Processing, Springer, 2004, pp. 1354–1361.

[58] G. KERSCHEN, J.-C. GOLINVAL, A. F. VAKAKIS, AND L. A. BERGMAN, *The method of proper orthogonal decomposition for dynamical characterization and order reduction of mechanical systems: an overview*, Nonlinear Dynam., 41 (2005), pp. 147–169.

[59] R. Z. KHASMINSKII, *Stochastic stability of differential equations.* Monographs and Textbooks on Mechanics of Solids and Fluids. Mechanics: Analysis, 7. Alphen aan den Rijn, The Netherlands; Rockville, Maryland, USA: Sijthoff & Noordhoff., 1980.

[60] P. E. KLOEDEN AND E. PLATEN, *Numerical solution of stochastic differential equations*, vol. 23 of Applications of Mathematics (New York), Springer-Verlag, Berlin, 1992.

[61] P. E. KLOEDEN AND E. PLATEN, *Stochastic differential equations*, in Numerical Solution of Stochastic Differential Equations, Springer, 1992, pp. 103–160.

[62] A. N. KOLMOGOROV, *Wienersche spiralen und einige andere interessante kurven in hilbertscen raum, cr (doklady)*, Acad. Sci. URSS (NS), 26 (1940), pp. 115–118.

[63] D. D. KOSAMBI, *Statistics in function space*, J. Indian Math. Soc. (N.S.), 7 (1943), pp. 76–88.

[64] P. KÜRSCHNER, *Balanced truncation model order reduction in limited time intervals for large systems*, Advances in Computational Mathematics, 44 (2018), pp. 1821–1844.

[65] Y. C. LIANG, H. P. LEE, S. P. LIM, W. Z. LIN, K. H. LEE, AND C. G. WU, *Proper orthogonal decomposition and its applications. I. Theory*, J. Sound Vibration, 252 (2002), pp. 527–544.

[66] A. LO, *Fat tails, long memory, and the stock market since the 1960's', economic notes, 26 (2), 213-46*, International Library of Critical Writings in Economics, 146 (2002), pp. 221–254.

[67] M. LOÈVE, *Fonctions aléatoires de second ordre*, Revue Sci., 84 (1946), pp. 195–206.

[68] P. MALLIAVIN, *Stochastic calculus of variation and hypoelliptic operators*, in Proc. Intern. Symp. SDE Kyoto 1976, Kinokuniya, 1978, pp. 195–263.

[69] B. B. MANDELBROT AND J. W. VAN NESS, *Fractional brownian motions, fractional noises and applications*, SIAM review, 10 (1968), pp. 422–437.

[70] X. MAO, *Stochastic differential equations and applications*, Elsevier, 2007.

[71] B. MASLOWSKI AND D. NUALART, *Evolution equations driven by a fractional Brownian motion*, J. Funct. Anal., 202 (2003), pp. 277–305.

[72] J. MÉMIN, Y. MISHURA, AND E. VALKEILA, *Inequalities for the moments of wiener integrals with respect to a fractional brownian motion*, Statistics & Probability Letters, 51 (2001), pp. 197–206.

[73] Y. S. MISHURA, *Stochastic calculus for fractional Brownian motion and related processes*, vol. 1929 of Lecture Notes in Mathematics, Springer-Verlag, Berlin, 2008.

[74] B. C. MOORE, *Principal component analysis in linear systems: controllability, observability, and model reduction*, IEEE Trans. Automat. Control, 26 (1981), pp. 17–32.

[75] F. MOSS AND P. V. MCCLINTOCK, *Noise in nonlinear dynamical systems*, vol. 2, Cambridge University Press, 1989.

[76] C. T. MULLIS AND R. A. ROBERTS, *Synthesis of minimum roundoff noise fixed point digital filters*, IEEE Trans. Circuits and Systems, CAS-2 (1976), pp. 551–562.

[77] A. NEUENKIRCH, *Optimal approximation of sde's with additive fractional noise*, Journal of Complexity, 22 (2006), pp. 459–474.

[78] ——, *Optimal pointwise approximation of stochastic differential equations driven by fractional Brownian motion*, Stochastic Process. Appl., 118 (2008), pp. 2294–2333.

[79] A. NEUENKIRCH AND I. NOURDIN, *Exact rate of convergence of some approximation schemes associated to SDEs driven by a fractional Brownian motion*, J. Theoret. Probab., 20 (2007), pp. 871–899.

[80] A. NEUENKIRCH AND T. SHALAIKO, *The maximum rate of convergence for the approximation of the fractional Lévy area at a single point*, J. Complexity, 33 (2016), pp. 107–117.

[81] B. NOBLE AND J. W. DANIEL, *Applied linear algebra*, Prentice-Hall, Inc., Englewood Cliffs, NJ, second ed., 1977.

[82] I. NORROS, E. VALKEILA, AND J. VIRTAMO, *An elementary approach to a Girsanov formula and other analytical results on fractional Brownian motions*, Bernoulli, 5 (1999), pp. 571–587.

[83] I. NOURDIN, *Selected aspects of fractional Brownian motion*, vol. 4 of Bocconi & Springer Series, Springer, Milan; Bocconi University Press, Milan, 2012.

[84] D. NUALART, *Fractional brownian motion: stochastic calculus and applications*, in International Congress of Mathematicians, vol. 3, Citeseer, 2006, pp. 1541–1562.

[85] ——, *Malliavin calculus and its applications*, vol. 110 of CBMS Regional Conference Series in Mathematics, Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, 2009.

[86] B. OKSENDAL, *Stochastic differential equations: an introduction with applications*, Springer Science & Business Media, 2013.

[87] S. PAINTER, *Long-range dependence in the subsurface: Empirical evidence and simulation methods*, in Invited paper at the American Geophysical Union 1998 Fall Meeting, 1998.

[88] P. A. Parrilo, S. Lall, F. Paganini, G. C. Verghese, B. C. Lesieutre, and J. E. Marsden, *Model reduction for analysis of cascading failures in power systems*, in Proceedings of the 1999 American Control Conference (Cat. No. 99CH36251), vol. 6, IEEE, 1999, pp. 4208–4212.

[89] L. Pernebo and L. M. Silverman, *Model reduction via balanced state space representations*, IEEE Trans. Automat. Control, 27 (1982), pp. 382–387.

[90] V. Pipiras and M. S. Taqqu, *Are classes of deterministic integrands for fractional brownian motion on an interval complete?*, Bernoulli, (2001), pp. 873–897.

[91] P. E. Protter, *Stochastic integration and differential equations*, vol. 21 of Stochastic Modelling and Applied Probability, Springer-Verlag, Berlin, second ed., 2005. Corrected third printing.

[92] V. S. Pugačev, *The general theory of correlation of random functions*, Izv. Akad. Nauk SSSR Ser. Mat., 17 (1953), pp. 401–420.

[93] H. Qian, *Fractional brownian motion and fractional gaussian noise*, in Processes with Long-Range Correlations: Theory and Applications, Springer, 2003, pp. 22–33.

[94] A. Quarteroni, A. Manzoni, and F. Negri, *Reduced basis methods for partial differential equations: an introduction*, vol. 92, Springer, 2015.

[95] A. Rascanu et al., *Differential equations driven by fractional brownian motion*, Collectanea Mathematica, (2002), pp. 55–81.

[96] M. Redmann, *Type II singular perturbation approximation for linear systems with Lévy noise*, SIAM J. Control Optim., 56 (2018), pp. 2120–2158.

[97] M. Redmann, C. Bayer, and P. Goyal, *Low-dimensional approximations of high-dimensional asset price models*, SIAM J. Financial Math., 12 (2021), pp. 1–28.

[98] M. Redmann, C. Bayer, and P. Goyal, *Low-dimensional approximations of high-dimensional asset price models*, SIAM J. Financial Math., 12 (2021), pp. 1–28.

[99] M. Redmann and I. P. Duff, *Full state approximation by Galerkin projection reduced order models for stochastic and bilinear systems*, Appl. Math. Comput., 420 (2022), pp. Paper No. 126561, 21.

[100] M. Redmann and M. A. Freitag, *Optimization based model order reduction for stochastic systems*, Appl. Math. Comput., 398 (2021), pp. Paper No. 125783, 18.

[101] M. Redmann and N. Jamshidi, *Gramian-based model reduction for unstable stochastic systems*, Math. Control Signals Systems, 34 (2022), pp. 855–881.

[102] M. Redmann and P. Kürschner, *An output error bound for time-limited balanced truncation*, Syst. Control. Lett., 121 (2018), pp. 1–6.

[103] M. Redmann and I. Pontes Duff, *Full state approximation by Galerkin projection reduced order models for stochastic and bilinear systems*, Appl. Math. Comput., (accepted 2021).

[104] M. REDMANN AND S. RIEDEL, *Runge-kutta methods for rough differential equations*, arXiv preprint arXiv:2003.12626, (2020).

[105] P. M. ROBINSON, *Time series with long memory*, Advanced Texts in Econometrics, 2003.

[106] L. C. G. ROGERS, *Arbitrage with fractional brownian motion*, Mathematical Finance, 7 (1997), pp. 95–105.

[107] C. W. ROWLEY, T. COLONIUS, AND R. M. MURRAY, *Model reduction for compressible flows using POD and Galerkin projection*, Phys. D, 189 (2004), pp. 115–129.

[108] F. RUSSO AND P. VALLOIS, *Forward, backward and symmetric stochastic integration*, Probab. Theory Related Fields, 97 (1993), pp. 403–421.

[109] ——, *Elements of stochastic calculus via regularisation*, Preprint Paris, 13 (2005).

[110] A. SHMATKOV, *Rate of convergence of wong-zakai approximations for sdes and spdes*, (2006).

[111] V. SIMONCINI, *Computational Methods for Linear Matrix Equations*, SIAM Rev., 58 (2016), pp. 377–441.

[112] L. SIROVICH, *Turbulence and the dynamics of coherent structures. I. Coherent structures*, Quart. Appl. Math., 45 (1987), pp. 561–571.

[113] K. SOBCZYK, *Stochastic differential equations*, vol. 40 of Mathematics and its Applications (East European Series), Kluwer Academic Publishers Group, Dordrecht, 1991. With applications to physics and engineering.

[114] M. S. SOLÉ, *Malliavin calculus: with applications to stochastic partial differential equations*, EPFL press, 2005.

[115] M. TURELLI, *Random environments and stochastic calculus*, Theoret. Population Biol., 12 (1977), pp. 140–178.

[116] T. M. TYRANOWSKI, *Data-driven structure-preserving model reduction for stochastic Hamiltonian systems*, J. Comput. Dyn., 11 (2024), pp. 220–255.

[117] C. VAROTSOS AND D. KIRK-DAVIDOFF, *Long-memory processes in global ozone and temperature variations*, Atmospheric Chemistry and Physics Discussions, 6 (2006), pp. 4325–4340.

[118] S. VOLKWEIN, *Proper orthogonal decomposition: Theory and reduced-order modelling*, Lecture Notes, University of Konstanz, 4 (2013), pp. 1–29.

[119] I. I. VRABIE, *C_0-semigroups and applications*, vol. 191 of North-Holland Mathematics Studies, North-Holland Publishing Co., Amsterdam, 2003.

[120] X. WANG, R. QI, AND F. JIANG, *Sharp mean-square regularity results for SPDEs with fractional noise and optimal convergence rates for the numerical approximations*, BIT, 57 (2017), pp. 557–585.

[121] E. Wong and M. Zakai, *On the relation between ordinary and stochastic differential equations*, Internat. J. Engrg. Sci., 3 (1965), pp. 213–229.

[122] X. Xie, F. Bao, and C. G. Webster, *Evolve filter stabilization reduced-order model for stochastic burgers equation*, Fluids, 3 (2018), p. 84.

[123] L. C. Young, *An inequality of the hölder type, connected with stieltjes integration*, (1936).

List of Figures

0.1 Process to a Reduced-Order Modeling 1
0.2 Input-output systems 2

2.1 Samples of fBm for different Hurst parameters 29
2.2 Samples of fBm increments for different Hurst parameters. 29

5.1 Decay of first 50 logarithmic HSVs of system (5.1) based on time-limited Gramians P_T and Q_T. 111
5.2 $\log_{10}\left(\sup_{t\in[0,1]}\mathbb{E}\|y(t)-y_r(t)\|_2\right)$ and logarithmic bound in (5.28) for different values of r. 111

6.1 $\mathcal{R}_E$ for three approaches with Hurst parameters $H = 0.5$. 136
6.2 $\mathcal{R}_E$ for three approaches with Hurst parameters $H = 0.75$. 136
6.3 First 50 POD singular values or eigenvalues associated to $\bar{P}_T$ for $H = 0.75$. 137
6.4 First 50 POD singular values or eigenvalues associated to P_T/Q_T for $H = 0.5$. 137

Lebenslauf

Personal Information

Name:	Nahid Jamshidi
Email:	nahid.jamshidi@mathematik.uni-halle.de

Work Experience

01/2021- Present **Lecturer & Researcher**
Institute for Mathematics, Martin-Luther-University Halle-Wittenberg, Halle (Saale), Germany

09/2013 - 06/2020 **Lecturer & Researcher**
Kermanshah University of Technology and Razi University, Kermanshah, Iran

Education

01/2021- 11/2025 **PhD in Applied Stochastic**
Martin-Luther-University Halle-Wittenberg, Halle (Saale), Germany
Thesis: Model Order Reduction for Stochastic Differential Equations driven by Standard and Fractional Brownian motion.
Advisor: Prof. Dr. Martin Redmann.

09/2010-09/2012 **MSc in Applied Mathematics, Ordinary Differential Equations**
Sahand University of Technology, Tabriz, East Azerbaijan, Iran
Thesis: Positive solution of singular boundary value problem for nonlinear fractional differential equations with nonlinearity that changes sign.
Advisor: Prof. Dr. Kazem Ghanbari.

09/2006-07/2010 **BSc in Applied Mathematics**
Razi University, Kermanshah, Iran
Advisor: Dr. Mehrdad AhmadZadeh Raji.

Academia appointment

03/2017 - 12/2017 **Visiting Scholar**
Southern Denmark University, Odense, Denmark
Department of Mathematics and Computer Science (IMADA)
Under the direction of Prof. Dr. Kristian Debrabant.

Publications

2025 N. Jamshidi, M. Redmann, *(Empirical) Gramian-based dimension reduction for stochastic differential equations driven by fractional Brownian motion*, Applied Probability Journals
DOI: 10.1017/jpr.2025.10039

2024 M. Kamrani and K. Debrabant, N. Jamshidi, *Exponential Euler method for stiff stochastic differential equations with additive fractional Brownian noise*, International Journal of Computer Mathematics
DOI: 10.1080/00207160.2024.2327615

2023 N. Jamshidi, M. Redmann, *Sampling-based model order reduction for stochastic differential equations driven by fractional Brownian motion*, Proceedings in Applied Mathematics and Mechanics
DOI: 10.1002/pamm.202200109

2022 M. Redmann, N. Jamshidi, *Gramian-based model reduction for unstable stochastic systems*, Mathematics of Control, Signals, and Systems
DOI: 10.1007/s00498-022-00328-z

2021 N. Jamshidi, M. Kamrani,*Convergence of a numerical scheme associated to stochastic differential equations with fractional Brownian motion*, Applied Numerical Mathematics
DOI: 10.1016/j.apnum.2021.05.001

2018 M. Kamrani, N. Jamshidi,*Implicit Milstein method for stochastic differential equations via the Wong-Zakai approximation*, Numerical Algorithm
DOI: 10.1007/s11075-017-0440-8

2017 M. Kamrani, N. Jamshidi,*Implicit Euler approximation of stochastic evaluation equation with fractional Brownian motion*, Communications in Nonlinear Science and Numerical Simulation
DOI: 10.1016/j.cnsns.2016.07.023

Halle (Saale), 11. March 2025
Nahid Jamshidi

Eidesstattliche Erklärung

Ich erkläre an Eides statt, dass ich die vorliegende Arbeit

Model Order Reduction for Stochastic Differential Equations driven by Standard and Fractional Brownian motion

selbstständig und ohne fremde Hilfe verfasst, keine anderen als die von mir angegebenen Quellen und Hilfsmittel benutzt und die den benutzten Werken wörtlich oder inhaltlich entnommenen Stellen als solche kenntlich gemacht habe.

Halle (Saale), 11. March 2025

Nahid Jamshidi