Journal of Computational and Applied Mathematics 419 (2023) 114634

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

journal homepage: www.elsevier.com/locate/cam

The BDF2-Maruyama method for the stochastic Allen–Cahn equation with multiplicative noise

Raphael Kruse, Rico Weiske *

Martin-Luther-Universität Halle-Wittenberg, Institut für Mathematik, 06099 Halle (Saale), Germany

A R T I C L E I N F O

Article history:
Received 31 January 2022
Received in revised form 3 July 2022

2010 MSC:
60H15
60H35
65C30
65M15

Keywords:
Stochastic Allen–Cahn equation
BDF2-Maruyama method
Multistep scheme
Strong convergence rate

A B S T R A C T

We investigate the numerical approximation of the stochastic Allen–Cahn equation on a bounded domain $\mathcal{D}$ under Dirichlet boundary conditions and with multiplicative noise. The considered numerical method combines the two-step backward differentiation formula (BDF2) for the temporal discretization in conjunction with an abstract Galerkin scheme for the spatial approximation. In dependence on the regularity of the exact solution we derive a rate of convergence for the BDF2-Maruyama method with respect to the root-mean-square error in discrete analogues of the spaces $L^\infty([0,T];L^2(\mathcal{D}))$ and $L^2([0,T];H^1_0(\mathcal{D}))$. Our error analysis is based on the variational approach for stochastic evolution equations. Finally, several numerical experiments illustrate our theoretical results, where a finite element method is used as an example for a Galerkin scheme.

1. Introduction

In this paper we study the spatio-temporal discretization of the stochastic Allen–Cahn (SAC) equation with multiplicative noise. To introduce the underlying problem, fix $T \in (0,\infty)$ and let $\mathcal{D} \subset \mathbb{R}^d$, $d \in \{1,2,3\}$, be a bounded domain with sufficiently smooth boundary. We consider the stochastic partial differential equation (SPDE)

$$\begin{aligned} \mathrm{d}u - \left(\Delta u + u - u^3\right)\mathrm{d}t &= b(u)\,\mathrm{d}W(t) && \text{in } (0,T]\times\mathcal{D},\\ u(t,x) &= 0, && (t,x)\in(0,T]\times\partial\mathcal{D},\\ u(0,x) &= u_0(x), && x\in\mathcal{D}, \end{aligned} \tag{1}$$

where $b\colon \mathbb{R} \to \mathbb{R}$ is Lipschitz continuous and satisfies a growth condition (cf. Assumption 2.3), W is an infinite-dimensional trace-class Q-Wiener process defined on a stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$, and the deterministic initial value $u_0\colon \mathcal{D} \to \mathbb{R}$ is assumed to be sufficiently smooth.

The SAC equation (1) is an often studied benchmark problem in numerical analysis of semi-linear SPDEs. The particular difficulty in the error analysis stems from the super-linearly growing non-linearity in the drift. In contrast to the deterministic case, the super-linear growth can cause instabilities and, consequently, the divergence of standard solvers which treat the semi-linear part, for instance, with an explicit step of the Euler method. An in-depth description of this phenomenon is given in [1].

* Corresponding author.
E-mail addresses: raphael.kruse@mathematik.uni-halle.de (R. Kruse), rico.weiske@mathematik.uni-halle.de (R. Weiske).

https://doi.org/10.1016/j.cam.2022.114634

The well-posedness of the SAC equation itself is, however, well-established. We refer to [2], where the variational approach [3] as well as the semi-group approach [4] to stochastic evolution equations are discussed. In detail, we rewrite the problem (1) equivalently as a stochastic evolution equation of the form

$$\mathrm{d}X(t) + A(X(t))\,\mathrm{d}t = B(X(t))\,\mathrm{d}W(t) \quad \text{on } (0,T], \quad X(0) = X_0. \tag{2}$$

Here the stochastic process X is identified as an abstract function of u. Further, the nonlinear operator A consists of the Laplace operator with Dirichlet boundary conditions and the Nemytskii operator associated to the polynomial part of the drift. Similarly, B denotes the Nemytskii operator associated to the mapping b. We provide more details on this framework in Section 2.2.

The analytical treatment of (2) then relies on some monotonicity property and the local Lipschitz continuity of the drift operator A. These properties have also been used to design and analyze various numerical methods for the approximation of the SAC equation. We refer to [1] for an exhaustive list of references and mention, in addition, [2,5–8] as a brief selection of more recent results.

In this paper we apply and analyze the two-step backward difference formula (BDF2) for the numerical approximation of (1). For deterministic problems, the BDF2 method is well-known for its good stability properties [9] and it is often used for the discretization of stiff problems. Compared to the backward Euler method, its order of converge is twice as high while it has essentially the same computational cost. Moreover, the BDF2 method can easily be combined with Galerkin schemes such as the finite element method, see [10, Chapter 10]. In the context of stochastic ordinary differential equations, the BDF2-Maruyama method was first introduced in [11] and studied further in [12].

In this paper we combine the BDF2-Maruyama method with an abstract Galerkin scheme and obtain a fully discrete numerical method for the SAC equation (1) with multiplicative noise. The resulting numerical method is introduced in full detail in Section 3.1. Our goal is then to derive convergence rates for the strong error. Our error analysis of the BDF2 method is primarily based on analytical tools from the variational approach for (stochastic) evolution equations from [13,14]. It has the advantage that we obtain error estimates with respect to the root-mean-square norms in discrete analogues of the spaces $L^\infty([0,T];L^2(\mathcal{D}))$ and $L^2([0,T];H_0^1(\mathcal{D}))$.

We remark that the error estimation in Theorem 3.4 is a modification of [14, Theorem 4.7]. While the underlying analytical framework is mostly identical in both papers, the latter result is not directly applicable to stochastic evolution equations (2) with a locally Lipschitz continuous drift operator A. In the present paper we focus solely on the SAC equation (1) which enables us to make use of the specific structure of this semi-linear problem. In contrast, the error analysis in [14, Theorem 4.7] also applies to a class of quasi-linear problems but under more restrictive conditions on the non-linearity. In addition, due to well-established results on the temporal regularity of the SAC equation (1), e.g., [2], the error estimation in Theorem 3.4 gives a self-contained prediction of the temporal order of convergence. Comparable regularity results are not readily available for the abstract stochastic evolution equation considered in [14].

The paper is structured as follows. In Section 2, we introduce the notation that we use throughout the paper. Moreover, we derive the variational formulation (2) and discuss sufficient conditions under which the problem is well-posed. Section 3 is devoted to the formulation of the BDF2-Maruyama method and to establish a strong error estimate. Furthermore, we also derive error estimates for the spatial discretization if the finite element method is used in place of the abstract Galerkin scheme. In Section 4 we illustrate the theoretical results through numerical experiments.

2. Preliminaries

First, we introduce functional spaces, norms, as well as further notations which will be utilized in the following. Then, we formulate the main assumptions under which we derive well-posedness of the variational formulation (2).

2.1. Notation

Throughout this paper, we consider a fixed terminal time $T \in (0,\infty)$ and a bounded domain $\mathcal{D} \subset \mathbb{R}^d$, $d \in \{1,2,3\}$, with smooth boundary. Further, let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$ define a stochastic basis. We denote by $C > 0$ a generic constant which is independent of any discretization parameters. Moreover, all equalities and inequalities involving random variables are assumed to hold $\mathbb{P}$-almost surely if not stated otherwise.

We denote the Lebesgue spaces by $L^p(\mathcal{D}) := L^p(\mathcal{D};\mathbb{R})$ with norm $\|\cdot\|_{L^p}$ for $p \in [1,\infty]$ and the Sobolev spaces by $H^s := H^s(\mathcal{D};\mathbb{R})$ with norm $\|\cdot\|_{H^s}$ for $s \in \mathbb{N}$. In particular, we identify $(H_0^1, (\cdot,\cdot)_{H_0^1})$ as the Hilbert space consisting of functions $u \in H^1$ with zero trace. We denote by H^{-1} the dual space of H_0^1 and use $\langle\cdot,\cdot\rangle_{H^{-1}\times H_0^1}$ for the dual pairing between these spaces. It holds $H_0^1 \hookrightarrow L^2(\mathcal{D}) \hookrightarrow H^{-1}$ where $\hookrightarrow$ denotes dense and continuous embeddings.

Let $\Delta\colon \mathrm{dom}(\Delta) \subset L^2(\mathcal{D}) \to L^2(\mathcal{D})$ denote the Laplace operator with zero Dirichlet boundary conditions. For $\gamma \in (0,\infty)$, we set $\dot{H}^\gamma := \mathrm{dom}\big((-\Delta)^{\frac{\gamma}{2}}\big)$ and endow it with the norm $\|u\|_{\dot{H}^\gamma} := \|(-\Delta)^{\frac{\gamma}{2}}u\|_{L^2}$, $u \in \dot{H}^\gamma$. We remark that the identifications $\dot{H}^1 = H_0^1$ and $\dot{H}^2 = H^2 \cap H_0^1$ hold.

Let V be a Banach space endowed with the Borel σ-algebra $\mathcal{B}(V)$. For $p \in [1,\infty)$ we denote by $L^p(\Omega;V) := L^p(\Omega,\mathcal{F},\mathbb{P};V)$ and $L^p([0,T]\times\Omega;V) := L^p([0,T]\times\Omega, \mathcal{B}([0,T])\otimes\mathcal{F}, \mathrm{d}t\otimes\mathbb{P};V)$ the Bochner–Lebesgue spaces equipped, respectively, with the norms

$$\|X\|_{L^p(\Omega;V)} := \left(\mathbb{E}\left[\|X\|_V^p\right]\right)^{\frac{1}{p}}, \quad \|X\|_{L^p([0,T]\times\Omega;V)} := \left(\mathbb{E}\left[\int_0^T \|X(t)\|_V^p \, \mathrm{d}t\right]\right)^{\frac{1}{p}}.$$

An introduction to Bochner–Lebesgue spaces is given, e.g., in [15,16].

Let $(U, (\cdot,\cdot)_U)$ and $(H, (\cdot,\cdot)_H)$ be two separable Hilbert spaces. We denote by $\mathcal{L}(U,H)$ the Banach space of all linear, bounded operators from U to H. Further, let $Q \in \mathcal{L}(U,U)$ be a non-negative, symmetric operator with finite trace. Recall from [3] that there exists a unique operator $Q^{\frac{1}{2}} \in \mathcal{L}(U)$ which satisfies $Q = Q^{\frac{1}{2}} \circ Q^{\frac{1}{2}}$ and induces a Hilbert space $U_0 := Q^{\frac{1}{2}}(U)$ endowed with the inner product $(u,v)_{U_0} := (Q^{-\frac{1}{2}}u, Q^{-\frac{1}{2}}v)_U$ for all $u, v \in U_0$, where $Q^{-\frac{1}{2}}$ denotes the pseudo-inverse of $Q^{\frac{1}{2}}$. Then we denote by $\mathcal{L}_2(U_0,H)$ the Hilbert space of all operators $B \in \mathcal{L}(U_0,H)$ with finite Hilbert–Schmidt norm $\|B\|^2_{\mathcal{L}_2(U_0,H)} := \mathrm{Tr}(B^*B)$.

Let $W = (W(t))_{t\in[0,T]}$ denote a U-valued Q-Wiener process with respect to the stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$ as defined in [3]. Such a Wiener process W has the representation

$$W(t) = \sum_{j=1}^{\infty} \sqrt{q_j}\chi_j\beta_j(t), \quad t \in [0,T], \tag{3}$$

where $\{\chi_j\}_{j\in\mathbb{N}}$ is an orthonormal basis of U consisting of eigenfunctions of Q with summable eigenvalues $q_j \geq 0$ and $\{\beta_j\}_{j\in\mathbb{N}}$ is a family of independent scalar Brownian motions.

We use an abstract Galerkin method for the spatial approximation of the exact solution. A family $(V_h)_{h\in(0,1)}$ of finite dimensional subspaces of H_0^1 is called a *Galerkin scheme* if for every $v \in L^2(\mathcal{D})$ it holds $\inf_{v_h\in V_h} \|v_h - v\|_{L^2} \to 0$ as $h \to 0$. We denote by $N_h \in \mathbb{N}$ the dimension of the subspace V_h. Notice that $N_h = \dim(V_h) \to \infty$ as $h \to 0$ since the space $L^2(\mathcal{D})$ is infinite dimensional.

Moreover, let $P_h : L^2(\mathcal{D}) \to V_h$ and $R_h : H_0^1 \to V_h$ denote the orthogonal projection mappings onto V_h with respect to the inner products $(\cdot,\cdot)_{L^2}$ and $(\cdot,\cdot)_{H_0^1}$, respectively. Both projectors satisfy best approximation properties in their underlying norms, i.e.

$$\begin{aligned} \|P_h u - u\|_{L^2} &= \inf_{v_h\in V_h} \|v_h - u\|_{L^2} \quad \text{for all } u \in L^2(\mathcal{D}), \\ \|R_h v - v\|_{H_0^1} &= \inf_{v_h\in V_h} \|v_h - v\|_{H_0^1} \quad \text{for all } v \in H_0^1. \end{aligned} \tag{4}$$

These properties follow from the Hilbert space structure of $L^2(\mathcal{D})$ and H_0^1, see, e.g., [17, Theorem 5.2].

2.2. Variational framework and assumptions

The aim of this subsection is to formulate the problem (1) as a stochastic evolution equation of the form (2) in the variational framework and to provide the definition of a variational solution. In addition, we discuss the assumptions under which the problem (2) is well-posed.

For the variational formulation, we identify $u : [0,T] \times \Omega \times \overline{\mathcal{D}} \to \mathbb{R}$ as the solution to (1) with a stochastic process taking values in H_0^1. More precisely, we consider

$$X : [0,T] \times \Omega \to H_0^1, \quad (t,\omega) \mapsto u(t,\cdot,\omega).$$

Furthermore, let us introduce the operators A and B associated to the drift and the function b in (1). These operators are defined, respectively, by

$$\begin{aligned} &A : H_0^1 \to H^{-1}, \quad A(v)(x) := -\Delta v(x) - v(x) + v^3(x), \\ &B : L^2(\mathcal{D}) \to \mathcal{L}_2(U_0, L^2(\mathcal{D})), \quad (B(v)w)(x) := b(v(x))w(x), \end{aligned} \tag{5}$$

for $v \in H_0^1 \hookrightarrow L^2(\mathcal{D})$, $w \in U_0$ and $x \in \mathcal{D}$. We remark that $A(v) \in H^{-1}$ defines a linear functional which is given by

$$\langle A(v), w\rangle_{H^{-1}\times H_0^1} = \int_{\mathcal{D}} \nabla v(x)\nabla w(x)\,\mathrm{d}x - \int_{\mathcal{D}} \big(v(x) - v^3(x)\big) w(x)\,\mathrm{d}x$$

for all $u, w \in H_0^1$. Notice that the operator A is well-defined due to the Sobolev embedding $H_0^1 \hookrightarrow L^6(\mathcal{D})$ which holds on the domain $\mathcal{D} \subset \mathbb{R}^d$ of dimension $d \in \{1,2,3\}$, see, e.g., [18, Section 5.6].

In order to derive (2) as the variational formulation of (1), we multiply the stochastic partial differential equation (1) with a test function in H_0^1, integrate over the time-space domain and apply integration by parts with respect to the spatial variable. This approach motivates the following definition based on [3, Definition 5.1.2].

Definition 2.1. Assume that $X_0 \in L^p(\Omega, \mathcal{F}_0, \mathbb{P}; L^2(\mathcal{D}))$. A continuous, $L^2(\mathcal{D})$-valued and $(\mathcal{F}_t)_{t\in[0,T]}$-adapted stochastic process $X \in L^2([0,T] \times \Omega; H_0^1)$ is called a variational solution of (2) if

$$(X(t), v)_{L^2} + \int_0^t \langle A(\bar{X}(s)), v\rangle_{H^{-1}\times H_0^1}\,\mathrm{d}s = (X_0, v)_{L^2} + \int_0^t (v, B(\bar{X}(s))\,\mathrm{d}W(s))_{L^2} \tag{6}$$

holds $\mathbb{P}$-almost surely for all $t \in [0,T]$ and $v \in H_0^1$, where $\bar{X}$ is a H_0^1-valued, progressively measurable modification of X.

In view of Definition 2.1, the stochastic evolution equation (2) is formally understood as an equation in H^{-1}. In the following, we will not distinguish notationally between the variational solution X and its modification.

Before we discuss the well-posedness of the problem (2), we gather some properties of the drift operator A and state the assumptions on the operator B. The proof of is included for completeness and uses similar techniques as in [2,7].

Lemma 2.2. *The operator A defined in (5) satisfies the monotonicity property*

$$\langle A(u) - A(v), u - v\rangle_{H^{-1}\times H_0^1} \geq \|u-v\|_{H_0^1}^2 - \|u-v\|_{L^2}^2 \tag{7}$$

and the local Lipschitz condition

$$\|A(u) - A(v)\|_{H^{-1}} \leq \|v-u\|_{H_0^1} + (1 + \|u\|_{H_0^1}^2 + \|v\|_{H_0^1}^2)\|v-u\|_{L^2} \tag{8}$$

for all $u, v \in H_0^1$.

We refer (7) as the monotonicity property of the operator A since it is equivalent to the strong monotonicity of the operator $A + \mathrm{id}\colon H_0^1 \to H^{-1}$, where id denotes the identity on H_0^1.

Proof. Let $w \in H_0^1$. By integration by parts we have

$$\langle A(u) - A(v), w\rangle_{H^{-1}\times H_0^1} = (u-v, w)_{H_0^1} - (u-v, w)_{L^2} + (u^3 - v^3, w)_{L^2}. \tag{9}$$

Notice that

$$u^3 - v^3 = \frac{1}{2}(u-v)((u+v)^2 + u^2 + v^2).$$

Hence testing (9) with $w = u - v$ yields the monotonicity property, since the last summand is non-negative. Furthermore, we deduce from (9) by applying Hölder's inequality, the inequality $(a+b)^2 \leq 2a^2 + 2b^2$ and the embedding $H_0^1 \hookrightarrow L^6(\mathcal{D})$ that

$$\begin{aligned}\langle A(u) - A(v), w\rangle_{H^{-1}\times H_0^1} &\leq \|u-v\|_{H_0^1}\|w\|_{H_0^1} + \|u-v\|_{L^2}\|w\|_{L^2} + \frac{3}{2}\|u-v\|_{L^2}\|u^2+v^2\|_{L^3}\|w\|_{L^6} \\ &\leq C\Big(\|u-v\|_{H_0^1} + (1 + \|u\|_{H_0^1}^2 + \|v\|_{H_0^1}^2)\|u-v\|_{L^2}\Big)\|w\|_{H_0^1}.\end{aligned}$$

Dividing this estimate by $\|w\|_{H_0^1}$ and taking the supremum over all $w \in H_0^1$ yields the local Lipschitz condition. □

Assumption 2.3. The operator $B\colon L^2(\mathcal{D}) \to \mathcal{L}_2(U_0, L^2(\mathcal{D}))$ defined in (5) satisfies the Lipschitz condition

$$\|B(u) - B(v)\|_{\mathcal{L}_2(U_0, L^2(\mathcal{D}))} \leq L_B\|v-u\|_{L^2}, \quad \text{for all } u, v \in L^2(\mathcal{D}), \tag{10}$$

with Lipschitz constant $L_B \in [0,\infty)$. Furthermore, it holds $B(H_0^1) \subset \mathcal{L}_2(U_0, H_0^1)$ and

$$\|B(v)\|_{\mathcal{L}_2(U_0, H_0^1)} \leq L_B(1 + \|v\|_{H_0^1}), \quad \text{for all } v \in H_0^1. \tag{11}$$

The existence of a unique variational solution of (2) was established in [2]. Here we understand uniqueness of a solution as the pathwise uniqueness or, equivalently, as the indistinguishability of the stochastic process. For a proof of the following result we refer to [2, Theorem 3.1].

Theorem 2.4. *Let $p \geq 4$. Assume that $X_0 \in L^p(\Omega, \mathcal{F}_0, \mathbb{P}; H_0^1)$ and Assumption 2.3 holds. Then there exists a unique variational solution X to (2) such that $X \in L^p(\Omega; C([0,T]; H_0^1))$.*

To derive $H^2 \cap H_0^1$-regularity of the variational solution of (2), a stronger growth condition on B needs to be imposed.

Assumption 2.5. The operator $B\colon L^2(\mathcal{D}) \to \mathcal{L}_2(U_0, L^2(\mathcal{D}))$ defined in (5) satisfies $B(\dot{H}^{1+\theta}) \subset \mathcal{L}_2(U_0, \dot{H}^{1+\theta})$ for some $\theta \in (0,\infty)$. In addition, there exists $\rho \geq 1$ such that

$$\|B(v)\|_{\mathcal{L}_2(U_0, \dot{H}^{1+\theta})} \leq C(1 + \|v\|_{\dot{H}^{1+\theta}}^{\rho}), \quad \text{for all } v \in \dot{H}^{1+\theta}.$$

Under an additional regularity assumption on the initial value, the following bounds of the moments and the Hölder regularity of the solution X were proven in [2]. In particular, the Hölder continuity in the $L^2(\Omega; H_0^1)$-norm depends on the smoothness of the deterministic initial value and will be crucial in the error analysis in Section 3.2.

Theorem 2.6. *Assume that the initial value satisfies $X_0 \in \dot{H}^{1+\gamma}$ for $\gamma \in (0, 1]$. Let Assumption 2.3 hold and, in case $\gamma = 1$, also Assumption 2.5 be satisfied. Then the variational solution X to (2) belongs to $L^p(\Omega; C([0, T]; \dot{H}^{1+\gamma}))$ for every $p \in [1, \infty)$ and there exists a constant $C > 0$ depending on $\|X_0\|_{\dot{H}^{1+\gamma}}$ such that*

$$\mathbb{E}\Big[\sup_{t\in[0,T]} \|X(t)\|^p_{\dot{H}^{1+\gamma}}\Big]^{\frac{1}{p}} \le C. \tag{12}$$

Moreover, it holds

$$\begin{aligned} &\mathbb{E}\big[\|X(s) - X(t)\|^p_{L^2}\big]^{\frac{1}{p}} \le C|s - t|^{\frac{1}{2}},\\ &\mathbb{E}\big[\|X(s) - X(t)\|^2_{H^1_0}\big]^{\frac{1}{2}} \le C|s - t|^{\frac{\gamma}{2}}, \end{aligned} \tag{13}$$

for every $s, t \in [0, T]$ and $p \in [2, \infty)$.

We remark that the Hölder estimates (13) are a slight generalization of the estimates of [7, Lemma 3.2] which are based on the variational approach. Theorem 2.6 has been derived by studying the mild solution to (2) in the semigroup approach and by noticing that the mild solution and the variational solution coincide. For further details we refer to [2, Proposition 3.1, Theorem 3.3, Corollary 3.2].

Remark 2.7. A generalized setting of the stochastic Allen–Cahn equation (1) can be considered, where the non-linearity of the drift is induced by an odd polynomial of at most degree three with negative leading coefficient such that the properties (7) and (8) hold, see [2, Example 2.1]. In this case, Theorem 2.6 still holds such that the aforementioned properties of the drift operator A are sufficient for the derivation of the main result, see Theorem 3.4.

A direct consequence of Theorem 2.6 and Lemma 2.2 is the following result.

Lemma 2.8. *Under the same assumptions as Theorem 2.6 with $\gamma \in (0, 1]$ there exists $C \in (0, \infty)$ such that the solution X to (2) satisfies for all $s, t \in [0, T]$*

$$\mathbb{E}\big[\|A(X(s)) - A(X(t))\|^2_{H^{-1}}\big] \le C|s - t|^{\gamma}.$$

Proof. Fix $s, t \in [0, T]$ arbitrarily. From the local Lipschitz condition (8) we deduce

$$\|A(X(s)) - A(X(t))\|_{H^{-1}} \le \|X(s) - X(t)\|_{H^1_0} + \big(1 + \|X(s)\|^2_{H^1_0} + \|X(t)\|^2_{H^1_0}\big)\|X(s) - X(t)\|_{L^2}. \tag{14}$$

Due to the Hölder continuity (13) of X with respect to the H^1_0-norm we have

$$\big(\mathbb{E}\big[\|X(s) - X(t)\|^2_{H^1_0}\big]\big)^{\frac{1}{2}} \le C|s - t|^{\frac{\gamma}{2}}.$$

Hence, after applying the $L^2(\Omega; \mathbb{R})$-norm to both sides of (14) we obtain

$$\big(\mathbb{E}\big[\|A(X(s)) - A(X(t))\|^2_{H^{-1}}\big]\big)^{\frac{1}{2}} \le C|s - t|^{\frac{\gamma}{2}} + \Big(\mathbb{E}\Big[\big(1 + \|X(s)\|^2_{H^1_0} + \|X(t)\|^2_{H^1_0}\big)^2\|X(s) - X(t)\|^2_{L^2}\Big]\Big)^{\frac{1}{2}}.$$

For the estimation of the second term on the right-hand side we apply Hölder's inequality with exponents $\rho = \frac{3}{2}$ and $\rho' = 3$. Then, it follows from the moments' bound (12), the triangle inequality, and the Hölder continuity (13) of the solution X with respect to the L^2-norm that

$$\begin{aligned} &\Big(\mathbb{E}\Big[\big(1 + \|X(s)\|^2_{H^1_0} + \|X(t)\|^2_{H^1_0}\big)^2\|X(s) - X(t)\|^2_{L^2}\Big]\Big)^{\frac{1}{2}}\\ &\quad\le \Big(\mathbb{E}\Big[\big(1 + \|X(s)\|^2_{H^1_0} + \|X(t)\|^2_{H^1_0}\big)^3\Big]\Big)^{\frac{1}{3}}\big(\mathbb{E}\big[\|X(s) - X(t)\|^6_{L^2}\big]\big)^{\frac{1}{6}}\\ &\quad\le C\Big(1 + 2\big(\mathbb{E}\big[\sup_{r\in[0,T]} \|X(r)\|^6_{H^1_0}\big]\big)^{\frac{1}{3}}\Big)|s - t|^{\frac{1}{2}} \le CT^{\frac{1-\gamma}{2}}|s - t|^{\frac{\gamma}{2}} \end{aligned}$$

since $\gamma \in (0, 1]$. Altogether, this completes the proof of the assertion. □

3. The BDF2-Maruyama method for the SAC equation

The goal of this section is to formulate the BDF2-Maruyama method with an underlying abstract Galerkin scheme for the spatial approximation and to establish a strong convergence rate using the variational approach introduced in Section 2.2. In addition, we discuss how the abstract convergence rate result transfers to the BDF2-Maruyama finite element method.

3.1. Discretization scheme

To formulate the numerical method, let us fix a step size $k = \frac{T}{N_k}$, $N_k \in \mathbb{N}$, and consider an equidistant temporal grid on $[0, T]$ with points $t_n := nk$ for $n \in \{0, \dots, N_k\}$. We denote Wiener increments by $\Delta_k W^n := W(t_n) - W(t_{n-1})$ for $n \in \{1, \dots, N_k\}$. Further, let $V_h \subset H_0^1$ be a finite dimensional subspace depending on some parameter $h \in (0, 1)$. For given initial values $(X_{k,h}^n)_{n=0,1}$ we define the *BDF2-Maruyama method* by

$$\Big(\frac{3}{2}X_{k,h}^n - 2X_{k,h}^{n-1} + \frac{1}{2}X_{k,h}^{n-2}, v\Big)_{L^2} + k\langle A(X_{k,h}^n), v\rangle_{H^{-1}\times H_0^1} = \Big(\frac{3}{2}B(X_{k,h}^{n-1})\Delta_k W^n - \frac{1}{2}B(X_{k,h}^{n-2})\Delta_k W^{n-1}, v\Big)_{L^2} \quad \mathbb{P}\text{-a.s.} \tag{15}$$

for all $v \in V_h$ and $n \in \{2, \dots, N_k\}$. This method is well-defined if there exists a unique discrete stochastic process $(X_{k,h}^n)_{n=0}^{N_k}$, which is $(\mathcal{F}_{t_n})_{n=0}^{N_k}$-adapted, $\mathbb{P}$-almost surely V_h-valued and solves (15). Here we understand uniqueness of a solution again in the sense of indistinguishable processes.

In order to initialize the two-step scheme (15), two suitable initial values are required. The following assumption on the initial values ensures that the numerical approximation is adapted and square-integrable.

Assumption 3.1. The initial values $(X_{k,h}^n)_{n=0,1}$ satisfy for each $n \in \{0, 1\}$

$$X_{k,h}^n \in L^2(\Omega, \mathcal{F}_{t_n}, \mathbb{P}; L^2(\mathcal{D})) \quad \text{and} \quad \mathbb{P}(\{\omega \in \Omega : X_{k,h}^n(\omega) \in V_h\}) = 1.$$

Theorem 3.2. *Let Assumption 2.3 hold and assume that some initial values $(X_{k,h}^n)_{n=0,1}$ satisfy Assumption 3.1. Fix $h \in (0, 1)$ and $k = \frac{T}{N_k}$, $N_k \in \mathbb{N}$. For every sufficiently small temporal step size k there exists a unique solution $(X_{k,h}^n)_{n=0}^{N_k}$ to (15) such that $X_{k,h}^n \in L^2(\Omega; H_0^1)$ holds for each $n \in \{2, \dots, N_k\}$.*

Proof. From the local Lipschitz condition (8) we deduce that the operator A is hemicontinuous, i.e. the mapping $z : [0, 1] \to \mathbb{R}$, $\lambda \mapsto \langle A(u + \lambda v), w\rangle_{H^{-1}\times H_0^1}$ is continuous for all $u, v, w \in H_0^1$. In addition, the monotonicity property (7) and Assumption 2.3 imply that a coupled monotonicity condition and coercivity condition given, respectively, by

$$2\langle A(u) - A(v), u - v\rangle_{H^{-1}\times H_0^1} + C\|u - v\|_{L^2}^2 \geq \|B(u) - B(v)\|^2_{\mathcal{L}_2(U_0, L^2(\mathcal{D}))},$$
$$2\langle A(v), v\rangle_{H^{-1}\times H_0^1} + C\|v\|_{L^2}^2 \geq 2\|B(v)\|^2_{\mathcal{L}_2(U_0, L^2(\mathcal{D}))} + 2\|v\|^2_{H_0^1},$$

hold for all $u, v \in H_0^1$. Furthermore, Assumption 2.3 yields that $B(X_{k,h}^n) \in L^2(\Omega, \mathcal{F}_{t_n}, \mathbb{P}; \mathcal{L}_2(U_0, L^2(\mathcal{D})))$ for $n \in \{0, 1\}$. Hence all assumptions for [14, Theorem 3.5] are satisfied with $p = 2$ except for a linear growth condition on the operator A. However, a detailed check of the corresponding proof shows that this property is not required. Therefore, the result follows. □

Remark 3.3. A suitable procedure to generate the two required initial values is to choose $X_{k,h}^0$ as the projection of X_0 onto V_h and to compute $X_{k,h}^1$ by one iteration of the backward Euler–Maruyama method. In particular, under Assumption 2.3 and $X_0 \in H_0^1$, the strong convergence of the backward Euler–Maruyama method for the stochastic Allen–Cahn equation was established in [2, Theorem 1.1]. There it is shown that

$$\mathbb{E}\big[\|X_{k,h}^1 - X(t_1)\|_{L^2}^2\big] = \mathcal{O}(h^2 + k). \tag{16}$$

3.2. Strong convergence rate of the BDF2-Maruyama method

The main result of this paper is presented in the following theorem.

Theorem 3.4. *Assume that the initial value satisfies $X_0 \in \dot{H}^{1+\gamma}$ for $\gamma \in (0, 1]$. Let Assumption 2.3 hold and, if $\gamma = 1$, let also Assumption 2.5 be satisfied. Further, let the initial values $(X_{k,h}^n)_{n=0,1}$ satisfy Assumption 3.1. Fix $h \in (0, 1)$ and $k = \frac{T}{N_k}$, $N_k \in \mathbb{N}$, with $k < \frac{1}{4}$. Then for every $p \in (2, \infty)$ there exists $C \in (0, \infty)$ independent of the discretization parameters h and k such that*

$$
\begin{aligned}
&\max_{n\in\{2,\ldots,N_k\}} \mathbb{E}\big[\|X_{k,h}^n - X(t_n)\|_{L^2}^2\big] + k\sum_{n=2}^{N_k}\mathbb{E}\big[\|X_{k,h}^n - X(t_n)\|_{H_0^1}^2\big]\\
&\le C\Big(k^\gamma + \sum_{n=0}^{1}\mathbb{E}\big[\|X_{k,h}^n - X(t_n)\|_{L^2}^2\big] + k\sum_{n=2}^{N_k}\mathbb{E}\big[\|(\mathrm{id}-P_h)X(t_n)\|_{H_0^1}^2\big]\Big)\\
&\quad + C\big(1+\|P_h\|_{\mathcal{L}(H_0^1)}^4\big)\max_{n\in\{2,\ldots,N_k\}}\big(\mathbb{E}\big[\|(\mathrm{id}-P_h)X(t_n)\|_{L^2}^p\big]\big)^{\frac{2}{p}}.
\end{aligned}
$$

Observe that the order of convergence of the BDF2-Maruyama methods depends on the Hölder regularity $\gamma \in (0,1]$ of the exact solution X with respect to the H_0^1-norm, cf. Theorem 2.6, which in turn is determined by the regularity of the initial value X_0. In addition, the error bound also depends on a consistent choice of the initial values $(X_{h,k}^n)_{n=0,1}$ and on estimates of the spatial discretization error of the abstract Galerkin scheme. To be more precise, we require error bounds on the L^2-orthogonal projector $\mathrm{id} - P_h$. We discuss the spatial discretization errors for the standard finite element method in more detail in Section 3.3.

Lemma 3.5. *Assume that the initial value satisfies $X_0 \in \dot{H}^{1+\gamma}$ for $\gamma \in (0,1]$. Let Assumption 2.3 hold and, in case $\gamma = 1$, also Assumption 2.5 be satisfied. For every $p \in (2,\infty)$ there exists $C \in (0,\infty)$ such that for all $h \in (0,1)$ and $t \in [0,T]$ it holds*

$$
\big(\mathbb{E}\big[\|A(X(t)) - A(P_hX(t))\|_{H^{-1}}^2\big]\big)^{\frac{1}{2}} \le \big(\mathbb{E}\big[\|(\mathrm{id}-P_h)X(t)\|_{H_0^1}^2\big]\big)^{\frac{1}{2}} + C\big(1+\|P_h\|_{\mathcal{L}(H_0^1)}^2\big)\big(\mathbb{E}\big[\|(\mathrm{id}-P_h)X(t)\|_{L^2}^p\big]\big)^{\frac{1}{p}},
$$

where X denotes the variational solution to (2).

Proof. Let $t \in [0,T]$, $p \in (2,\infty)$, and $h \in (0,1)$ be arbitrary. From the local Lipschitz condition (8) we deduce

$$
\|A(X(t)) - A(P_hX(t))\|_{H^{-1}} \le \|X(t) - P_hX(t)\|_{H_0^1} + \big(1 + \|X(t)\|_{H_0^1}^2 + \|P_hX(t)\|_{H_0^1}^2\big)\|X(t) - P_hX(t)\|_{L^2}. \tag{17}
$$

Hence, after applying the $L^2(\Omega;\mathbb{R})$-norm to both sides of (17) we obtain

$$
\begin{aligned}
&\big(\mathbb{E}\big[\|A(X(t)) - A(P_hX(t))\|_{H^{-1}}^2\big]\big)^{\frac{1}{2}}\\
&\le \big(\mathbb{E}\big[\|(\mathrm{id}-P_h)X(t)\|_{H_0^1}^2\big]\big)^{\frac{1}{2}} + \Big(\mathbb{E}\Big[\big(1+\|X(t)\|_{H_0^1}^2 + \|P_hX(t)\|_{H_0^1}^2\big)^2\|(\mathrm{id}-P_h)X(t)\|_{L^2}^2\Big]\Big)^{\frac{1}{2}}.
\end{aligned} \tag{18}
$$

For the estimation of the last summand in (18) we make use of the assumption that $\|P_h\|_{\mathcal{L}(H_0^1)} < \infty$. Hence, we arrive at

$$
\begin{aligned}
&\Big(\mathbb{E}\Big[\big(1+\|X(t)\|_{H_0^1}^2 + \|P_hX(t)\|_{H_0^1}^2\big)^2\|(\mathrm{id}-P_h)X(t)\|_{L^2}^2\Big]\Big)^{\frac{1}{2}}\\
&\le \big(1+\|P_h\|_{\mathcal{L}(H_0^1)}^2\big)\Big(\mathbb{E}\Big[\big(1+\|X(t)\|_{H_0^1}^2\big)^2\|(\mathrm{id}-P_h)X(t)\|_{L^2}^2\Big]\Big)^{\frac{1}{2}}.
\end{aligned}
$$

Finally, after applying the moments' bound (12) and the Hölder inequality with exponents $\rho = \frac{p}{p-2}$ and $\rho' = \frac{p}{2}$ we conclude that

$$
\begin{aligned}
&\Big(\mathbb{E}\Big[\big(1+\|X(t)\|_{H_0^1}^2 + \|P_hX(t)\|_{H_0^1}^2\big)^2\|(\mathrm{id}-P_h)X(t)\|_{L^2}^2\Big]\Big)^{\frac{1}{2}}\\
&\le \big(1+\|P_h\|_{\mathcal{L}(H_0^1)}^2\big)\Big(\mathbb{E}\Big[\big(1+\|X(t)\|_{H_0^1}^2\big)^{2\rho}\Big]\Big)^{\frac{1}{2\rho}}\Big(\mathbb{E}\big[\|(\mathrm{id}-P_h)X(t)\|_{L^2}^p\big]\Big)^{\frac{1}{p}}\\
&\le C\big(1+\|P_h\|_{\mathcal{L}(H_0^1)}^2\big)\Big(\mathbb{E}\big[\|(\mathrm{id}-P_h)X(t)\|_{L^2}^p\big]\Big)^{\frac{1}{p}}.
\end{aligned}
$$

Altogether, this completes the proof of the assertion. □

Proof of Theorem 3.4. For an improved readability, we omit the dependence of the discrete solution on the parameters k and h by writing $X^n := X_{k,h}^n$ throughout the proof. Further, by $E^n := X^n - X(t_n)$ we denote the error of the numerical method (15) at time t_n for $n \in \{0,\ldots,N_k\}$. In the following, we make use of the orthogonal decomposition

$$
E^n = P_hE^n + (\mathrm{id}-P_h)E^n = \big(X^n - P_hX(t_n)\big) + (P_h - \mathrm{id})X(t_n) =: \Theta^n + \Xi^n.
$$

Notice that Θ^i and Ξ^j are indeed orthogonal with respect to the L^2-inner product for all $i,j \in \{0,\ldots,N_k\}$ $\mathbb{P}$-almost surely.

A straightforward algebraic calculation verifies the identity

$$
\begin{aligned}
&\|E^n\|_{L^2}^2 - \|E^{n-1}\|_{L^2}^2 + \|2E^n - E^{n-1}\|_{L^2}^2 - \|2E^{n-1} - E^{n-2}\|_{L^2}^2 + \|E^n - 2E^{n-1} + E^{n-2}\|_{L^2}^2\\
&= 4\Big(\frac{3}{2}E^n - 2E^{n-1} + \frac{1}{2}E^{n-2}, E^n\Big)_{L^2} =: \Gamma^n.
\end{aligned} \tag{19}
$$

Fix $j \in \{2, \dots, N_k\}$. Due to the telescopic structure on the left-hand side we obtain by summing over $n \in \{2, \dots, j\}$ that

$$\|E^j\|_{L^2}^2 - \|E^1\|_{L^2}^2 + \|2E^j - E^{j-1}\|_{L^2}^2 - \|2E^1 - E^0\|_{L^2}^2 = \sum_{n=2}^{j}\Big(\Gamma^n - \|E^n - 2E^{n-1} + E^{n-2}\|_{L^2}^2\Big). \tag{20}$$

Next, we make use of the L^2-orthogonal decomposition of the error E^n on the right-hand side of (19) to obtain

$$\Gamma^n = 4\Big(\frac{3}{2}E^n - 2E^{n-1} + \frac{1}{2}E^{n-2}, \Theta^n\Big)_{L^2} + 2\big(3\Xi^n - 4\Xi^{n-1} + \Xi^{n-2}, \Xi^n\big)_{L^2}.$$

After inserting the definitions of the numerical method (15) and the variational solution (6) we arrive at

$$\begin{aligned}\Gamma^n &= -4k\big\langle A(X^n), \Theta^n\big\rangle_{H^{-1}\times H_0^1} + 2\Big\langle 3\int_{t_{n-1}}^{t_n} A(X(s))\,\mathrm{d}s - \int_{t_{n-2}}^{t_{n-1}} A(X(s))\,\mathrm{d}s, \Theta^n\Big\rangle_{H^{-1}\times H_0^1}\\ &\quad + 2\Big(3\int_{t_{n-1}}^{t_n} B(X^{n-1}) - B(X(s))\,\mathrm{d}W(s), \Theta^n\Big)_{L^2} - 2\Big(\int_{t_{n-2}}^{t_{n-1}} B(X^{n-2}) - B(X(s))\,\mathrm{d}W(s), \Theta^n\Big)_{L^2}\\ &\quad + 2\big(3\Xi^n - 4\Xi^{n-1} + \Xi^{n-2}, \Xi^n\big)_{L^2}\\ &= \Gamma_A^n + \Gamma_B^n + \Gamma_\Xi^n.\end{aligned} \tag{21}$$

We derive error bounds for the terms Γ_A^n, Γ_B^n, and Γ_Ξ^n, separately. To estimate the term Γ_A^n, we consider the decomposition

$$\begin{aligned}\Gamma_A^n &= -4k\langle A(X^n) - A(P_hX(t_n)), \Theta^n\rangle_{H^{-1}\times H_0^1} + 4\int_{t_{n-1}}^{t_n}\langle A(X(s)) - A(P_hX(t_n)), \Theta^n\rangle_{H^{-1}\times H_0^1}\,\mathrm{d}s\\ &\quad + 2\int_{t_{n-1}}^{t_n}\langle A(X(s)) - A(P_hX(t_{n-1})), \Theta^n\rangle_{H^{-1}\times H_0^1}\,\mathrm{d}s - 2\int_{t_{n-2}}^{t_{n-1}}\langle A(X(s)) - A(P_hX(t_{n-1})), \Theta^n\rangle_{H^{-1}\times H_0^1}\,\mathrm{d}s,\end{aligned}$$

where we also used that the grid points t_n are equidistant with step size k. Recall that the operator A satisfies the monotonicity property (7). Together with several applications of Young's inequality we deduce

$$\Gamma_A^n \le -k\|\Theta^n\|_{H_0^1}^2 + 4k\|\Theta^n\|_{L^2}^2 + 4\int_{t_{n-1}}^{t_n}\|A(X(s)) - A(P_hX(t_n))\|_{H^{-1}}^2\,\mathrm{d}s + \int_{t_{n-2}}^{t_n}\|A(X(s)) - A(P_hX(t_{n-1}))\|_{H^{-1}}^2\,\mathrm{d}s.$$

After taking expectations and applying Fubini's theorem, we are able to make use of Lemma 2.8 and Lemma 3.5 with $p \in (2, \infty)$ arbitrary. This yields

$$\int_{t_{n-1}}^{t_n}\|A(X(s)) - A(P_hX(t_n))\|_{H^{-1}}^2\,\mathrm{d}s \le Ck^{1+\gamma} + 2k\mathbb{E}\big[\|\Xi^n\|_{H_0^1}^2\big] + C(1 + \|P_h\|_{\mathcal{L}(H_0^1)}^4)k\big(\mathbb{E}\big[\|\Xi^n\|_{L^2}^p\big]\big)^{\frac{2}{p}}$$

as well as

$$\int_{t_{n-2}}^{t_n}\|A(X(s)) - A(P_hX(t_{n-1}))\|_{H^{-1}}^2\,\mathrm{d}s \le Ck^{1+\gamma} + 4k\mathbb{E}\big[\|\Xi^{n-1}\|_{H_0^1}^2\big] + C(1 + \|P_h\|_{\mathcal{L}(H_0^1)}^4)k\big(\mathbb{E}\big[\|\Xi^{n-1}\|_{L^2}^p\big]\big)^{\frac{2}{p}}.$$

To sum up, we have for every $p \in (2, \infty)$ that

$$\begin{aligned}\mathbb{E}\big[\Gamma_A^n\big] &\le -k\mathbb{E}\big[\|\Theta^n\|_{H_0^1}^2\big] + 4k\,\mathbb{E}\big[\|\Theta^n\|_{L^2}^2\big] + Ck^{1+\gamma} + 8k\mathbb{E}\big[\|\Xi^n\|_{H_0^1}^2\big] + 4k\mathbb{E}\big[\|\Xi^{n-1}\|_{H_0^1}^2\big]\\ &\quad + C\big(1 + \|P_h\|_{\mathcal{L}(H_0^1)}^4\big)k\Big(\big(\mathbb{E}\big[\|\Xi^n\|_{L^2}^p\big]\big)^{\frac{2}{p}} + \big(\mathbb{E}\big[\|\Xi^{n-1}\|_{L^2}^p\big]\big)^{\frac{2}{p}}\Big).\end{aligned} \tag{22}$$

For the estimation of Γ_B^n, we set $\mathrm{I}^n := \int_{t_{n-1}}^{t_n}\big(B(X^{n-1}) - B(X(s))\big)\,\mathrm{d}W(s)$ for $n \in \{1, \dots, N_k\}$. We decompose Γ_B^n further by

$$\begin{aligned}\Gamma_B^n &= 2\big(3\mathrm{I}^n - \mathrm{I}^{n-1}, \Theta^n\big)_{L^2}\\ &= 2\big(\mathrm{I}^n - \mathrm{I}^{n-1}, \Theta^n - 2\Theta^{n-1} + \Theta^{n-2}\big)_{L^2} + 2\big(\mathrm{I}^n, 2\Theta^n - \Theta^{n-1}\big)_{L^2} - 2\big(\mathrm{I}^{n-1}, 2\Theta^{n-1} - \Theta^{n-2}\big)_{L^2}\\ &\quad + 2\big(\mathrm{I}^n, 3\Theta^{n-1} - \Theta^{n-2}\big)_{L^2}.\end{aligned} \tag{23}$$

It holds $\mathbb{E}\big[\mathrm{I}^n \mid \mathcal{F}_{t_{n-1}}\big] = 0$ due to the martingale property of the stochastic integral. Since the random variable $3\Theta^{n-1} - \Theta^{n-2}$ is $\mathcal{F}_{t_{n-1}}$-measurable, the expectation of the last term in the decomposition (23) of Γ_B^n equals zero. Hence, after applying the Cauchy–Schwarz inequality, the Young inequality and taking expectation, we have

$$\begin{aligned}\mathbb{E}\big[\Gamma_B^n\big] &\le \mathbb{E}\big[\|\mathrm{I}^n - \mathrm{I}^{n-1}\|_{L^2}^2\big] + \mathbb{E}\big[\|\Theta^n - 2\Theta^{n-1} + \Theta^{n-2}\|_{L^2}^2\big]\\ &\quad + 2\mathbb{E}\big[\big(\mathrm{I}^n, 2\Theta^n - \Theta^{n-1}\big)_{L^2}\big] - 2\mathbb{E}\big[\big(\mathrm{I}^{n-1}, 2\Theta^{n-1} - \Theta^{n-2}\big)_{L^2}\big].\end{aligned} \tag{24}$$

Notice that I^n and I^{n-1} are orthogonal with respect to the inner product in $L^2(\Omega; L^2)$ such that $\mathbb{E}\big[\|\mathrm{I}^n - \mathrm{I}^{n-1}\|_{L^2}^2\big] = \mathbb{E}\big[\|\mathrm{I}^n\|_{L^2}^2\big] + \mathbb{E}\big[\|\mathrm{I}^{n-1}\|_{L^2}^2\big]$ holds for $n \in \{2, \dots, N_k\}$. Therefore, the summation in (24) over n from 2 to $j \in \{2, \dots, N_k\}$

with a further application of the Cauchy–Schwarz inequality yields

$$\sum_{n=2}^{j}\mathbb{E}\big[\Gamma_B^n\big] \le 2\sum_{n=1}^{j}\mathbb{E}\big[\|\mathrm{I}^n\|_{L^2}^2\big] + \sum_{n=2}^{j}\mathbb{E}\big[\|\Theta^n - 2\Theta^{n-1} + \Theta^{n-2}\|_{L^2}^2\big] + \mathbb{E}\big[\|2\Theta^j - \Theta^{j-1}\|_{L^2}^2\big] + \mathbb{E}\big[\|2\Theta^1 - \Theta^0\|_{L^2}^2\big]. \tag{25}$$

Because of the Itō isometry and the Lipschitz continuity (10) of the operator B, we have

$$\mathbb{E}\big[\|\mathrm{I}^n\|_{L^2}^2\big] = \mathbb{E}\Big[\int_{t_{n-1}}^{t_n}\|B(X^{n-1}) - B(X(s))\|_{\mathcal{L}_2(U_0,L^2(\mathcal{D}))}^2\,\mathrm{d}s\Big] \le 2L_Bk\mathbb{E}\big[\|E^{n-1}\|_{L^2}^2\big] + 2L_B\mathbb{E}\Big[\int_{t_{n-1}}^{t_n}\|X(t_{n-1}) - X(s)\|_{L^2}^2\,\mathrm{d}s\Big]$$

for every $n \in \{1,\ldots,N_k\}$. Together with the Hölder regularity (13) of the variational solution with respect to the L^2-norm, we conclude from (25) that

$$\sum_{n=2}^{j}\mathbb{E}\big[\Gamma_B^n\big] \le 2L_Bk\sum_{n=1}^{j}\mathbb{E}\big[\|E^{n-1}\|_{L^2}^2\big] + \sum_{n=2}^{j}\mathbb{E}\big[\|\Theta^n - 2\Theta^{n-1} + \Theta^{n-2}\|_{L^2}^2\big] + \mathbb{E}\big[\|2\Theta^j - \Theta^{j-1}\|_{L^2}^2\big] + \mathbb{E}\big[\|2\Theta^1 - \Theta^0\|_{L^2}^2\big] + Ck. \tag{26}$$

We similarly rewrite the expectation of the term Γ_{Ξ}^n as in (19) by

$$\mathbb{E}\big[\Gamma_{\Xi}^n\big] = \mathbb{E}\big[\|\Xi^n\|_{L^2}^2\big] - \mathbb{E}\big[\|\Xi^{n-1}\|_{L^2}^2\big] + \mathbb{E}\big[\|2\Xi^n - \Xi^{n-1}\|_{L^2}^2\big] - \mathbb{E}\big[\|2\Xi^{n-1} - \Xi^{n-2}\|_{L^2}^2\big] + \mathbb{E}\big[\|\Xi^n - 2\Xi^{n-1} + \Xi^{n-2}\|_{L^2}^2\big]. \tag{27}$$

After taking expectation in (20), we are able to insert the estimates (22), (26), and (27). After also recalling that the error decomposition $E^n = \Theta^n + \Xi^n$ is L^2-orthogonal for every $n \in \{2,\ldots,j\}$ we infer

$$\begin{aligned}\mathbb{E}\big[\|E^j\|_{L^2}^2\big] &\le -k\sum_{n=2}^{j}\mathbb{E}\big[\|\Theta^n\|_{H_0^1}^2\big] + 4k\sum_{n=2}^{j}\mathbb{E}\big[\|\Theta^n\|_{L^2}^2\big] + 2L_Bk\sum_{n=2}^{j}\mathbb{E}\big[\|E^{n-1}\|_{L^2}^2\big]\\ &\quad + Ck^{\gamma} + 2\mathbb{E}\big[\|2\Theta^1 - \Theta^0\|_{L^2}^2\big] + (1+2k)\mathbb{E}\big[\|\Theta^1\|_{L^2}^2\big]\\ &\quad + 8k\mathbb{E}\big[\|\Xi^j\|_{H_0^1}^2\big] + 12k\sum_{n=2}^{j}\mathbb{E}\big[\|\Xi^{n-1}\|_{H_0^1}^2\big] + C\big(1 + \|P_h\|_{\mathcal{L}(H_0^1)}^4\big)\Big(\max_{n\in\{1,\ldots,j\}}\mathbb{E}\big[\|\Xi^n\|_{L^2}^p\big]^{\frac{2}{p}}\Big).\end{aligned}$$

Here we used that $\mathbb{E}\big[\|\Xi^n\|_{L^2}^p\big] \le \mathbb{E}\big[\|X(t_n)\|_{L^2}^p\big] \le C$ holds uniformly in n which follows from the projection property of $\mathrm{id} - P_h$ and the moments' estimate (12). From the L^2-orthogonality of the error decomposition and the triangle inequality we further deduce

$$\begin{aligned}\mathbb{E}\big[\|\Theta^n\|_{L^2}^2\big] &\le \mathbb{E}\big[\|E^n\|_{L^2}^2\big],\\ -\mathbb{E}\big[\|\Theta^n\|_{H_0^1}^2\big] &\le -\frac{1}{2}\mathbb{E}\big[\|E^n\|_{H_0^1}^2\big] + \mathbb{E}\big[\|\Xi^n\|_{H_0^1}^2\big].\end{aligned}$$

Hence for every $k \in (0, \frac{1}{4})$ and $j \in \{2,\ldots,N_k\}$ it follows that

$$\begin{aligned}(1-4k)\mathbb{E}\big[\|E^j\|_{L^2}^2\big] + \frac{1}{2}k\sum_{n=2}^{j}\mathbb{E}\big[\|E^n\|_{H_0^1}^2\big] &\le 2(2+L_B)k\sum_{n=2}^{j-1}\mathbb{E}\big[\|E^n\|_{L^2}^2\big] + Ck^{\gamma} + C\sum_{n=0}^{1}\mathbb{E}\big[\|E^n\|_{L^2}^2\big]\\ &\quad + C\Big(k\sum_{n=2}^{j}\mathbb{E}\big[\|\Xi^n\|_{H_0^1}^2\big] + \big(1 + \|P_h\|_{\mathcal{L}(H_0^1)}^4\big)\max_{n\in\{2,\ldots,j\}}\big(\mathbb{E}\big[\|\Xi^n\|_{L^2}^p\big]\big)^{\frac{2}{p}}\Big).\end{aligned} \tag{28}$$

Finally, we divide by $1 - 4k$, apply a discrete Gronwall inequality, see, e.g., [19], and arrive for every $j \in \{2,\ldots,N_k\}$ at the error bound

$$\begin{aligned}\mathbb{E}\big[\|E^j\|_{L^2}^2\big] + k\sum_{n=2}^{j}\mathbb{E}\big[\|E^n\|_{H_0^1}^2\big] &\le C\Big(k^{\gamma} + \sum_{n=0}^{1}\mathbb{E}\big[\|E^n\|_{L^2}^2\big]\Big)\\ &\quad + C\Big(k\sum_{n=2}^{N_k}\mathbb{E}\big[\|\Xi^n\|_{H_0^1}^2\big] + \big(1 + \|P_h\|_{\mathcal{L}(H_0^1)}^4\big)\max_{n\in\{2,\ldots,N_k\}}\big(\mathbb{E}\big[\|\Xi^n\|_{L^2}^p\big]\big)^{\frac{2}{p}}\Big).\end{aligned}$$

Notice that the right-hand side yields a bound for each summand on the left-hand side independently of $j \in \{2,\ldots,N_k\}$. Hence the proof is complete. □

3.3. Strong convergence rate of the BDF2-Maruyama finite element method

In this subsection we deduce from Theorem 3.4 a strong convergence rate for a version of the BDF2-Maruyama scheme (15) that is combined with the standard finite element method for the spatial discretization.

As in Section 3.1, we consider the BDF2-Maruyama scheme (15) on an equidistant partition of $[0,T]$ with step size $k = \frac{T}{N_k}$, $N_k \in \mathbb{N}$. Regarding the spatial discretization, let $\{\mathcal{T}_h\}_{h\in(0,1)}$ be a quasi-uniform family of triangulations of $\mathcal{D}$ with maximal diameter h. We define the space V_h consisting of piecewise linear finite elements by

$$V_h = \big\{ v \in C(\overline{\mathcal{D}};\mathbb{R}) \colon v|_K \in \mathcal{P}_1(K) \text{ for all } K \in \mathcal{T}_h \text{ and } v|_{\partial\mathcal{D}} = 0 \big\}, \tag{29}$$

where $\mathcal{P}_1(K)$ denotes the space of $\mathbb{R}$-valued polynomials on K up to degree 1. Further, we denote by $\{\phi_i\}_{i=1}^{N_h} \subset V_h$ the Lagrange basis functions of V_h and by $\{x_j\}_{j=1}^{N_h} \subset \overline{\mathcal{D}}$ the nodes of the triangulation which are uniquely determined by $\phi_i(x_j) = \delta_{ij}$ for all $i,j = 1,\dots,N_h$. Let us recall from [20, Section 4.4] that the family of spaces $\{V_h\}_{h\in(0,1)}$ defines a Galerkin scheme for the Sobolev space H_0^1. Moreover, the corresponding orthogonal projector P_h is H^1-stable as the following lemma shows. We refer for a corresponding proof to [21,22].

Lemma 3.6. *Let the space V_h be defined by (29) and let P_h be the orthogonal L^2-projector onto V_h. Then it holds $\|P_h\|_{\mathcal{L}(H_0^1)} < \infty$ uniformly in $h \in (0,1)$.*

We denote by

$$I_h \colon C(\overline{\mathcal{D}};\mathbb{R}) \to V_h, \quad v \mapsto I_h(v) = \sum_{i=1}^{N_h} v(x_i)\phi_i \tag{30}$$

the interpolation operator. Notice that the operator I_h is well-defined on H^2 due to the Sobolev embedding $H^2 \hookrightarrow C(\overline{\mathcal{D}};\mathbb{R})$ holding on domains $\mathcal{D} \subset \mathbb{R}^d$, $d \in \{1,2,3\}$, with C^1-boundary, see, e.g. [18, Section 5.6]. The following estimates of the interpolation error

$$\|(I_h - \mathrm{id})v\|_{L^2} \le Ch^2\|v\|_{H^2}, \quad \text{as well as} \quad \|(I_h - \mathrm{id})v\|_{H_0^1} \le Ch\|v\|_{H^2},$$

hold for every $v \in H^2 \cap H_0^1$ and a constant $C > 0$ independent of $h \in (0,1)$, see, e.g., [20, Theorem 4.4.20]. The next lemma is a direct consequence from these error estimates and the best approximation properties (4) of the orthogonal projectors P_h and R_h.

Lemma 3.7. *Let the space V_h be defined by (29). Further, let P_h and R_h be the orthogonal L^2-projector and H_0^1-projector onto V_h, respectively. Then, there exists $C \in (0,\infty)$ such that for every $h \in (0,1)$ and every $v \in H^2 \cap H_0^1$ it holds*

$$\|(\mathrm{id} - P_h)v\|_{L^2} \le Ch^2\|v\|_{H^2}, \quad \text{as well as} \quad \|(\mathrm{id} - R_h)v\|_{H_0^1} \le Ch\|v\|_{H^2}.$$

We generate the initial values for the BDF2-Maruyama finite element method according to Remark 3.3. More precisely, we set $X_{k,h}^0 = I_h(X_0)$ and compute $X_{k,h}^1$ by one iteration of the backward Euler–Maruyama finite element method. If $X_0 \in H^2 \cap H_0^1$, then it follows from the interpolation error estimates and the error estimate (16) that

$$\sum_{n=0}^{1} \mathbb{E}\big[\|X_{k,h}^n - X(t_n)\|_{L^2}^2\big] = \mathcal{O}(h^2 + k). \tag{31}$$

Combining these results with Theorem 3.4 we arrive at the following error estimates.

Corollary 3.8. *Assume $X_0 \in H^2 \cap H_0^1$. Let Assumptions 2.3 and 2.5 be satisfied. Then there exists a constant $C \in (0,\infty)$ such that for every $h \in (0,1)$ and for every $k = \frac{T}{N_k} < \frac{1}{4}$, $N_k \in \mathbb{N}$, the BDF2-Maruyama finite element method satisfies*

$$\max_{n\in\{2,\dots,N_k\}} \mathbb{E}\big[\|X_{k,h}^n - X(t_n)\|_{L^2}^2\big] + k\sum_{n=2}^{N_k} \mathbb{E}\big[\|X_{k,h}^n - X(t_n)\|_{H_0^1}^2\big] \le C(k + h^2).$$

4. Numerical experiments

In this section, we simulate the stochastic Allen–Cahn equation (1) by using the BDF2-Maruyama finite element method from Section 3.3 and compare its performance to the backward Euler–Maruyama (BEM) finite element method.

In our simulations, we set $T = 1$ and consider the domain $\mathcal{D} = [0,1]$. For the initial value we choose $u_0 = \sin(\pi\,\cdot) \in H^2 \cap H_0^1$. Further, for the Q-Wiener process W we use a truncated version of the Karhunen–Loève expansion (3) with the first $J = 2^{12}$ summands. To be more precise, we choose the eigenbasis $\chi_j(x) = \sqrt{2}\sin(j\pi x)$ and $q_j = j^{-(5+\varepsilon)}$ with $\varepsilon = 10^{-3}$ as the corresponding eigenvalues of the covariance operator Q. As discussed in [23, Example 10.9], this yields a Wiener process which takes values in $H^2 \cap H_0^1$ almost surely.

For the intensity of the multiplicative noise we take the affine-linear mapping $b(u) = \sigma(u+1)$ with parameter values $\sigma \in \{0, 0.4, 1.0\}$. The Hilbert–Schmidt norm of the associated Nemytskii-type operator B is then given by

$$\|B(v)\|^2_{\mathcal{L}(U_0,H^2)} = \sum_{j=1}^{\infty} q_j \|B(v)\chi_j\|^2_{H^2} = 2\sigma^2 \sum_{j=1}^{\infty} j^{-(5+\varepsilon)} \big\|(1+v(\cdot))\sin(j\pi\,\cdot)\big\|^2_{H^2},$$

where we inserted the definition of b and the eigenbasis $(q_j, \chi_j)_{j\in\mathbb{N}}$ of the operator Q. Since the eigenfunctions χ_j of Q and their first and second order derivatives are bounded on $\mathcal{D}$ one finds a constant $C \in (0,\infty)$ such that

$$\big|\big|(1+v(\cdot))\sin(j\pi\,\cdot)\big|\big|_{H^2} \le Cj^2\big(1+\|v\|_{H^2}\big), \quad \text{for all } v \in H^2.$$

From this it follows that B satisfies Assumption 2.5. Analogously, one also verifies the Lipschitz continuity of B in Assumption 2.3.

For the temporal discretization we consider an equidistant grid with step size $k = \frac{T}{N_k}$, where $N_k = 2^l$ for $l = 5, \dots, 11$. For the spatial discretization we employ the standard finite element method with piecewise linear basis functions over an equidistant partition of $\mathcal{D}$ with $N_h = 2^{12}$ interior nodes and spatial step size $h = \frac{1}{N_h+1}$.

In our numerical experiments, we compute the strong error between the approximate solution of the respective scheme and the exact solution of the stochastic Allen–Cahn equation (1) with respect to time discrete versions of the $L^\infty([0,T]; L^2(\Omega; L^2(0,1)))$-norm and the $L^2([0,T]; L^2(\Omega; H^1_0(0,1)))$-norm using a Monte Carlo simulation with $M = 1000$ independent samples. In detail, we consider the strong error estimators

$$\begin{aligned} L^2\text{-error}_{k,h} &:= \max_{n\in\{2,\dots,N_k\}} \Big(\frac{1}{M}\sum_{m=1}^{M} \|X^{n,(m)}_{k,h} - X^{(m)}(t_n)\|^2_{L^2}\Big)^{\frac{1}{2}}, \\ H^1_0\text{-error}_{k,h} &:= \Big(\frac{k}{M}\sum_{n=2}^{N_k}\sum_{m=1}^{M} \|X^{n,(m)}_{k,h} - X^{(m)}(t_n)\|^2_{H^1_0}\Big)^{\frac{1}{2}}, \end{aligned} \tag{32}$$

where $\{X^{n,(m)}_{k,h} - X^{(m)}(t_n)\}_{m=1,\dots,M}$ are independently generated samples of the error $X^n_{k,h} - X(t_n)$. Hereby, we use a numerical reference solution computed by the BEM finite element method with $N_k = 2^{17}$ steps and $N_h = 2^{12}$ degrees of freedom as a substitute for the exact solution in (32). It was shown in [2] that the method used for the reference solution is strongly convergent with respect to the time discrete $L^\infty([0,T]; L^2(\Omega; L^2(0,1)))$-norm with order $\mathcal{O}(k^{\frac{1}{2}})$. Since we initialize the BEM finite element method also with $X^0_{k,h} = I_h(X_0)$, the approximate solutions of both schemes coincide at the first two temporal grid points. Hence we exclude the points $\{t_0, t_1\}$ in the error computation (32).

To estimate the temporal convergence rates, we compute the *experimental order of convergence* (EOC) which we define for successive temporal step sizes k_{i-1}, k_i and fixed spatial step size h by

$$\text{EOC} = \frac{\log(\text{error}_{k_i,h}) - \log(\text{error}_{k_{i-1},h})}{\log(k_i) - \log(k_{i-1})}.$$

To simplify the implementation of the numerical experiments, we project the nonlinearity of the drift and the noise term on the finite element space V_h by applying the piecewise linear interpolation operator I_h defined in (30), i.e. we consider the approximations

$$\begin{aligned} \big(X^3 - X, v\big)_H &\approx \big(I_h(X^3 - X), v\big)_{L^2}, \\ \big(B(X)\Delta_k W^n, v\big)_H &\approx \sigma\Big(I_h\big(B(X)\Delta_k W^n\big), v\Big)_H, \end{aligned}$$

for some $X \in H^1$, $v \in V_h$ and $n \in \{1,\dots,N_k\}$. Since $f\colon \mathbb{R} \to \mathbb{R}, x \mapsto x^3 - x$ is a C^1-mapping, the composition $f(X)$ is H^1-valued for every $X \in H^1$, cf. [17, Corollary 8.11]. Moreover, the Wiener increment $\Delta_k W^n$ is smooth by the chosen construction such that the term $B(X)\Delta_k W^n$ is also H^1-valued. Due to the embedding $H^1(0,1) \hookrightarrow C([0,1];\mathbb{R})$, this approximation is well-defined. Furthermore, the interpolation error is of order $\mathcal{O}(h)$, see, e.g., [20, Theorem 4.4.20].

The results from the error analysis are presented for each choice of the parameter σ in Table 1 to Table 3, respectively. The simulations showed that there is no noticeable difference in the computational time between the two considered methods. Furthermore, the computations for each experiment took on average 3.67 min for each method and sample, i.e. about 7.3 min for the first deterministic experiment in Table 1 and up to 121.6 h for each stochastic experiment in Tables 2 and 3 (if the computations are executed successively for every sample).

In the first experiment with $\sigma = 0$, we observe that the BDF2-Maruyama method performs significantly better than the BEM method with respect to both considered error norms. In addition, the two-step method converges in comparison to the one-step method with almost twice the rate for fine grids with $N_k \in \{512, 1024, 2024\}$ steps. This is in line with the well studied case of the deterministic Allen–Cahn equation.

The second experiment, where the small noise intensity $\sigma = 0.4$ is used, indicates that both methods initially exceed the expected convergence rate with respect to both error norms. To be more precise, the experimental order of convergence is generally larger than $\frac{1}{2}$ but decreases towards the expected rate for both numerical methods as N_k increases.

Table 1
Error for deterministic Allen–Cahn equation (1) with $\sigma = 0$.

N_k	BEM				BDF2			
	L^2-err	EOC	H_0^1-err	EOC	L^2-err	EOC	H_0^1-err	EOC
32	0.034408		0.049235		0.020286		0.015438	
64	0.018339	0.91	0.026210	0.91	0.007643	1.41	0.005285	1.55
128	0.009455	0.96	0.013451	0.96	0.002402	1.67	0.001569	1.75
256	0.004802	0.98	0.006805	0.98	0.000695	1.79	0.000430	1.87
512	0.002418	0.99	0.003419	0.99	0.000191	1.86	0.000111	1.95
1024	0.001212	1.00	0.001711	1.00	0.000051	1.92	0.000027	2.05
2048	0.000605	1.00	0.000854	1.00	0.000013	1.96	0.000007	1.89

Table 2
Error for stochastic Allen–Cahn equation (1) with $\sigma = 0.4$.

N_k	BEM				BDF2			
	L^2-err	EOC	H_0^1-err	EOC	L^2-err	EOC	H_0^1-err	EOC
32	0.037165		0.067985		0.023253		0.043852	
64	0.020047	0.89	0.038858	0.81	0.010299	1.17	0.025518	0.78
128	0.010639	0.91	0.021886	0.83	0.005183	0.99	0.015143	0.75
256	0.005609	0.92	0.012403	0.82	0.003111	0.74	0.009243	0.71
512	0.003158	0.83	0.007410	0.74	0.002128	0.55	0.005984	0.63
1024	0.001897	0.74	0.004563	0.70	0.001458	0.54	0.003930	0.61
2048	0.001163	0.71	0.002914	0.65	0.000987	0.56	0.002655	0.57

Table 3
Error for stochastic Allen–Cahn equation (1) with $\sigma = 1.0$.

N_k	BEM				BDF2			
	L^2-err	EOC	H_0^1-err	EOC	L^2-err	EOC	H_0^1-err	EOC
32	0.065305		0.174214		0.059198		0.160974	
64	0.042929	0.61	0.116304	0.58	0.039144	0.60	0.108204	0.57
128	0.027631	0.64	0.075971	0.61	0.025711	0.61	0.071529	0.60
256	0.018670	0.57	0.051158	0.57	0.017916	0.52	0.048964	0.55
512	0.012964	0.53	0.034586	0.56	0.012651	0.50	0.033645	0.54
1024	0.009007	0.53	0.023955	0.53	0.008844	0.52	0.023543	0.52
2048	0.006407	0.49	0.016685	0.52	0.006330	0.48	0.016512	0.51

This can be explained as follows: Due to the small noise intensity, the error stemming from the approximation of the drift integral contributes more dominantly to the overall error than the error from the approximation of the stochastic integral. This becomes evident by comparing the absolute values of the errors for $N_k \in \{32, 64, 128\}$ in the Table 2 with those in Table 1. If the temporal grid becomes finer, the relative contribution of the error of the approximation of the drift part decreases more rapidly than the error of approximating the stochastic integrals. Hence, at some point the orders of convergence approach those of standard Maruyama-type approximations on sufficiently fine time grids. The analysis of this phenomenon for linear multistep methods was already investigated in [11] for SODEs. Moreover, since the BDF2-Maruyama method resolves the drift part more accurately, the error from the stochastic part dominates the overall error already for $N_k = 512$ explaining why its EOC values deteriorate earlier than those of the BEM method. However, the absolute height of the computed errors still indicates a better performance the full range of considered step sizes.

Finally, for high noise intensity $\sigma = 1.0$, we observe for both methods in Table 3 that the error converges independently of the chosen norm with the expected rate of $\frac{1}{2}$. In this case, the advantage of the BDF2-Maruyama method over the BEM method seems negligible.

References

[1] M. Beccari, M. Hutzenthaler, A. Jentzen, R. Kurniawan, F. Lindner, D. Salimova, Strong and weak divergence of exponential and linear-implicit Euler approximations for stochastic partial differential equations with superlinearly growing nonlinearities, 2019, arXiv:1903.06066.

[2] Z. Liu, Z. Qiao, Strong approximation of monotone stochastic partial differential equations driven by multiplicative noise, Stoch. Partial Differ. Equ. Anal. Comput. 9 (3) (2021) 559–602, http://dx.doi.org/10.1007/s40072-020-00179-2.

[3] W. Liu, M. Röckner, Stochastic Partial Differential Equations: An Introduction, Universitext, Springer, Cham, 2015, http://dx.doi.org/10.1007/978-3-319-22354-4.

[4] G. Da Prato, J. Zabczyk, Stochastic equations in infinite dimensions, in: Encyclopedia of Mathematics and Its Applications, Vol. 44, Cambridge University Press, Cambridge, 1992.

[5] C.-E. Bréhier, J. Cui, J. Hong, Strong convergence rates of semidiscrete splitting approximations for the stochastic Allen-Cahn equation, IMA J. Numer. Anal. 39 (4) (2019) 2096–2134, http://dx.doi.org/10.1093/imanum/dry052.

[6] X. Feng, Y. Li, Y. Zhang, Strong convergence of a fully discrete finite element method for a class of semilinear stochastic partial differential equations with multiplicative noise, J. Comput. Math. 39 (4) (2021) 574–598, http://dx.doi.org/10.4208/jcm.2003-m2019-0250.

[7] A.K. Majee, A. Prohl, Optimal strong rates of convergence for a space–time discretization of the stochastic allen-cahn equation with multiplicative noise, Comput. Methods Appl. Math. 18 (2) (2018) 297–311, http://dx.doi.org/10.1515/cmam-2017-0023.
[8] X. Wang, An efficient explicit full-discrete scheme for strong approximation of stochastic Allen-Cahn equation, Stochastic Process. Appl. 130 (10) (2020) 6271–6299, http://dx.doi.org/10.1016/j.spa.2020.05.011.
[9] E. Hairer, G. Wanner, Solving Ordinary Differential Equations. II, second ed., in: Springer Series in Computational Mathematics, vol. 14, Springer-Verlag, Berlin, 1996, http://dx.doi.org/10.1007/978-3-642-05221-7, stiff and differential–algebraic problems.
[10] V. Thomée, Galerkin Finite Element Methods for Parabolic Problems, second ed., in: Springer Series in Computational Mathematics, vol. 25, Springer-Verlag, Berlin, 2006.
[11] E. Buckwar, R. Winkler, Multistep methods for SDEs and their application to problems with small noise, SIAM J. Numer. Anal. 44 (2) (2006) 779–803, http://dx.doi.org/10.1137/040602857.
[12] A. Andersson, R. Kruse, Mean-square convergence of the BDF2-maruyama and backward Euler schemes for SDE satisfying a global monotonicity condition, BIT 57 (1) (2017) 21–53, http://dx.doi.org/10.1007/s10543-016-0624-y.
[13] E. Emmrich, Two-step BDF time discretisation of nonlinear evolution problems governed by monotone operators with strongly continuous perturbations, Comput. Methods Appl. Math. 9 (1) (2009) 37–62, http://dx.doi.org/10.2478/cmam-2009-0003.
[14] R. Kruse, R. Weiske, The BDF2-Maruyama scheme for stochastic evolution equations with monotone drift, 2021, arXiv:2105.08767.
[15] D.L. Cohn, Measure theory, in: BirkhäUser Advanced Texts: Basler Lehrbücher, second ed., Birkhäuser/ Springer, New York, 2013, http://dx.doi.org/10.1007/978-1-4614-6956-8.
[16] N.S. Papageorgiou, P. Winkert, Applied Nonlinear Functional Analysis: An Introduction, De Gruyter Graduate, De Gruyter, Berlin, 2018.
[17] H. Brezis, Functional Analysis, Sobolev Spaces and Partial Differential Equations, Universitext, Springer, New York, 2011.
[18] L.C. Evans, Partial Differential Equations, in: Graduate Studies in Mathematics, vol. 19, American Mathematical Society, 1998.
[19] D.S. Clark, Short proof of a discrete Grönwall inequality, Discrete Appl. Math. 16 (3) (1987) 279–281, http://dx.doi.org/10.1016/0166-218X(87)90064-3.
[20] S.C. Brenner, L.R. Scott, The Mathematical Theory of Finite Element Methods, third ed., in: Texts in Applied Mathematics, vol. 15, Springer, New York, 2008, http://dx.doi.org/10.1007/978-0-387-75934-0.
[21] M. Crouzeix, V. Thomée, The stability in L_p and W_p^1 of the L_2-projection onto finite element function spaces, Math. Comp. 48 (178) (1987) 521–532, http://dx.doi.org/10.2307/2007825.
[22] C. Carstensen, Merging the Bramble-Pasciak-Steinbach and the Crouzeix-Thomée criterion for H^1-stability of the L^2-projection onto finite element spaces, Math. Comp. 71 (237) (2002) 157–163, http://dx.doi.org/10.1090/S0025-5718-01-01316-3.
[23] G.J. Lord, C.E. Powell, T. Shardlow, An Introduction to Computational Stochastic PDEs, first ed., in: Cambridge Texts in Applied Mathematics, 50, Cambridge University Press, Cambridge, 2014.