IMA Journal of Numerical Analysis (2025) **45**, 404–450
https://doi.org/10.1093/imanum/drae009
Advance Access publication on 4 May 2024

Strong convergence of adaptive time-stepping schemes for the stochastic Allen–Cahn equation

Chuchu Chen, Tonghe Dang* and Jialin Hong
LSEC, ICMSEC, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, 100190, China and School of Mathematical Sciences, University of Chinese Academy of Sciences, Beijing, 100049, China
*Corresponding author: dangth@lsec.cc.ac.cn

[Received on 9 April 2022; revised on 9 September 2023]

It is known from Beccari *et al.* (2019) that the standard explicit Euler-type scheme (such as the exponential Euler and the linear-implicit Euler schemes) with a uniform timestep, though computationally efficient, may diverge for the stochastic Allen–Cahn equation. To overcome the divergence, this paper proposes and analyzes adaptive time-stepping schemes, which adapt the timestep at each iteration to control numerical solutions from instability. The *a priori* estimates in $\mathscr{C}(\mathscr{O})$-norm and $\dot{H}^{\beta}(\mathscr{O})$-norm of numerical solutions are established provided the adaptive timestep function is suitably bounded, which plays a key role in the convergence analysis. We show that the adaptive time-stepping schemes converge strongly with order $\frac{\beta}{2}$ in time and $\frac{\beta}{d}$ in space with d ($d = 1, 2, 3$) being the dimension and $\beta \in (0, 2]$. Numerical experiments show that the adaptive time-stepping schemes are simple to implement and at a lower computational cost than a scheme with the uniform timestep.

Keywords: Stochastic Allen–Cahn equation; Adaptive time-stepping scheme; Strong convergence.

1. Introduction

Numerical approximations for stochastic partial differential equations (SPDEs) with globally Lipschitz coefficients have been studied in recent decades (see e.g., the monograph, Lord *et al.*, 2014). In contrast, numerical analysis of SPDEs with nonglobally Lipschitz coefficients, for example the stochastic Allen–Cahn equation, has been considered (see e.g., Cui & Hong, 2019; Liu & Qiao, 2020; Cui *et al.*, 2021, and references therein) and is still not fully understood. It is pointed out in Beccari *et al.* (2019) that the explicit Euler, the exponential Euler and the linear-implicit Euler schemes with the uniform timestep fail to converge in the strong sense for SPDEs with superlinearly growing coefficients; see also Jentzen & Pušnik (2020). Implicit schemes like fully drift-implicit scheme (see e.g., Kovács *et al.*, 2018; Majee & Prohl, 2018; Qi & Wang, 2019; Liu & Qiao, 2020, 2021, and references therein) can be strongly convergent in this setting. It is known that the implementation of the implicit scheme requires solving an algebraic equation at each iteration step, which needs additional computational effort. These reasons have led to the research on the construction of explicit schemes that can ensure convergence under the nonglobally Lipschitz condition. For instance, Bréhier & Goudenège (2019) proposes the splitting scheme and studies the convergence in strong, weak and probability senses. It is shown that the mean-square convergence order is almost 1/4, localized on an event of arbitrarily large probability, and that the convergence order in probability is almost 1/4 for the space-time white noise case; the authors in Bréhier *et al.* (2019) study the strong convergence order of the explicit temporal splitting numerical scheme for the case of different noises with varying degrees of smoothness; Cai *et al.* (2021) proves that the weak convergence order of the tamed exponential Euler scheme is almost β for the generalized Q-Wiener process case, and Wang (2020) shows that the strong convergence order of the nonlinearity-tamed accelerated exponential Euler scheme is almost $1/2$ for the space-time white noise case; Becker & Jentzen (2019) studies the strong convergence order of the nonlinearity-truncated Euler-type schemes, which is almost $1/4$ for the cylindrical Wiener process case. The present work makes further contributions on the numerical study of the explicit, adaptive time-stepping schemes for the stochastic Allen–Cahn equation.

Adaptive time-stepping schemes, which adapt the timestep at each iteration to control the numerical solution from divergence, have been deeply studied for stochastic ordinary differential equations (SODEs) with nonglobally Lipschitz drift. As for the selection of adaptive timesteps, we refer to e.g., Kelly & Lord (2018); Fang & Giles (2020); Kelly & Lord (2022) for the admissible strategy, Lamba (2003); Lemaire (2007) for the strategy based on the local error control and Merle & Prohl (2021) for the strategy based on the *a posteriori* weak error estimate. Numerically, this adaptive scheme is simple to implement and the complexity is similar to that of an Euler scheme, which is a big advantage for high dimensional problems (see Lamba, 2003). It is also pointed out in Hoel *et al.* (2012); Fang & Giles (2020) that such an adaptive scheme can lead to better computational performance for multi-level Monte–Carlo simulations. To our knowledge, there are few works on the study of the adaptive time-stepping scheme for SPDEs. The first attempt to apply the adaptive time-stepping scheme to the simulation of SPDEs is Campbell & Lord (2018), where the strong convergence rate is obtained under the assumption that the Fréchet derivative of the drift coefficient is bounded polynomially in $L^2(\mathbb{D})$-norm (see Campbell & Lord, 2018, Assumption 2.4), where $\mathbb{D} \subset \mathbb{R}^d$.

Consider another important class of nonlinear SPDEs, including the stochastic Allen–Cahn equation driven by additive noise

$$\begin{cases} \mathrm{d}X(t) + AX(t)\,\mathrm{d}t = F(X(t))\,\mathrm{d}t + \mathrm{d}W(t), \ t \in (0,T], \\ X(0) = X_0, \end{cases} \tag{1.1}$$

where $-A := \Delta : \mathrm{Dom}(A) \subset H \to H$ is the Laplacian operator with homogeneous Dirichlet boundary condition with $H := L^2(\mathscr{O}), \mathscr{O} := [0,1]^d, d = 1,2,3$ endowed with the usual inner product $\langle \cdot, \cdot \rangle$ and the norm $\| \cdot \|$, and the stochastic process $\{W(t)\}_{t\in[0,T]}$ is a generalized Q-Wiener process on a filtered probability space $(\Omega, \mathscr{F}, \{\mathscr{F}_t\}_{0\le t\le T}, \mathbb{P})$, subject to $\|A^{\frac{\beta-1}{2}} Q^{\frac{1}{2}}\|_{\mathscr{L}_2(H)} < \infty$, $\beta \in (0,2]$. The nonlinear drift F is a Nemytskii operator defined by $F(X)(x) = f(X(x))$ for $X \in H$ and $x \in \mathscr{O}$, where f is a polynomial and satisfies $f(\xi) = \sum_{i=0}^{3} a_i \xi^i$, $a_3 < 0$, $\xi \in \mathbb{R}$. In this case, the Fréchet derivative of F is bounded polynomially in E-norm ($E := \mathscr{C}(\mathscr{O})$). The main contribution of this work is to present the *a priori* estimates and rigorous strong convergence analysis of the adaptive time-stepping schemes for (1.1).

To be specific, the adaptive time-stepping scheme for (1.1), whose spatial discretization is using the spectral Galerkin method, and temporal direction is based on the adaptive exponential integrator, is an explicit numerical scheme with adaptive timesteps. The prerequisite of the convergence analysis is the *a priori* estimates in E-norm and $\dot{H}^{\beta}$-norm of the numerical solution, which are derived by a bootstrap argument. We refer to Wang (2020) for the use of this argument for the nonlinearity-tamed scheme with the uniform timestep. Based on the above *a priori* estimates, and combining the smoothing effect of the analytic semigroup and regularity properties of the generalized Q-Wiener process, the strong convergence order of this fully discrete scheme for (1.1) is finally carefully analyzed, which is the same as usual, i.e., order $\frac{\beta}{2}$ in time and $\frac{\beta}{d}$ in space. Moreover, we give the numerical analysis of the adaptive time-stepping scheme for the multiplicative noise case, and show that the convergence order is $1/2$ in time and 1 in space when $d = 1$.

For the feasibility of an adaptive time-stepping scheme, some bounds for the adaptive timestep function are proposed. We would like to mention that in practice, instead of verifying whether an adaptive timestep function satisfies the given lower bound, one generally introduces a backstop scheme with a uniform timestep and couples it with the adaptive time-stepping scheme to ensure that a simulation over the interval $[0, T]$ can be completed in a finite number of timesteps (see Kelly & Lord, 2018, for the case of SODEs). More precisely, when the lower bound is invalid for the adaptive timestep function, for example, we perform a single step with the tamed exponential integrator with a uniform timestep instead. It can be shown that the corresponding coupled scheme is strongly convergent with the order being the same as the adaptive time-stepping scheme. Further, it can be observed from numerical experiments in Section 7 that the coupled schemes are at a lower computational cost, measured in terms of the CPU time.

The outline of this paper is as follows. In the next section, some preliminaries are listed. In Section 3, we propose the adaptive time-stepping schemes, and present the main convergence theorem of this paper. Section 4 presents the *a priori* estimates in E-norm and $\dot{H}^{\beta}$-norm of numerical solutions. In Section 5, we give the proof of the main convergence theorem of the schemes. In Section 6, we give the discussion of the numerical analysis for the multiplicative noise case. Section 7 is devoted to the numerical experiments, which verify our theoretical results.

2. Preliminaries

In this section, we give assumptions on A, F, $W(t)$ and the initial datum, as well as the well-posedness of (1.1), see e.g., Cerrai (2001); Cui *et al.* (2021) for details. Throughout this paper, C is a constant that may change from one line to another, and sometimes we write $C(a, b, c, \ldots)$ to emphasize the dependence on the parameters $a, b, c, \ldots$

Let $H^s := H^s(\mathscr{O})$ be the usual Sobolev space. Then the domain of the operator A is $\mathrm{Dom}(A) := H^2 \cap H_0^1$, and there is a sequence of real numbers $\lambda_i \sim i^{\frac{2}{d}}, i \in \mathbb{N}_+$ (see Chen *et al.*, 2022, Section 1), and an orthonormal basis $\{e_i(x)\}_{i\in\mathbb{N}_+}$ such that $Ae_i = \lambda_i e_i$. It is known that A is positive, self-adjoint and densely defined operator on H, and that $-A$ generates an analytic semigroup $\{S(t) := e^{-tA}, t \geq 0\}$ on H. Define the Hilbert space $\dot{H}^{\gamma} := \mathrm{Dom}(A^{\frac{\gamma}{2}})$, $\gamma \in \mathbb{R}$, equipped with the inner product $\langle \cdot, \cdot \rangle_{\gamma} := \langle A^{\frac{\gamma}{2}} \cdot, A^{\frac{\gamma}{2}} \cdot \rangle$ and the norm $\| \cdot \|_{\gamma} := \langle \cdot, \cdot \rangle_{\gamma}^{\frac{1}{2}}$. Furthermore, $\mathscr{L}(H, U)$ and $\mathscr{L}_2(H, U)$ denote spaces of the usual bounded linear operators and Hilbert–Schmidt operators from a Hilbert space H to another Hilbert space U, respectively. When $H = U$, we use notations $\mathscr{L}(H)$ and $\mathscr{L}_2(H)$ for simplicity.

It is well-known (see e.g., Kruse, 2014, Lemma B.9) that there is a positive constant C such that

$$\|A^{\gamma} S(t)\|_{\mathscr{L}(H)} \leq Ct^{-\gamma}, \quad t > 0, \gamma \geq 0, \tag{2.1}$$

$$\|A^{-\gamma}(Id - S(t))\|_{\mathscr{L}(H)} \leq Ct^{\gamma}, \quad t > 0, \gamma \in [0, 1], \tag{2.2}$$

$$\int_s^t \left\| A^{\frac{\gamma}{2}} S(t-r)u \right\|^2 \mathrm{d}r \le C(t-s)^{1-\gamma} \|u\|^2, \quad u \in H,\ 0 \le s \le t,\ \gamma \in [0,1], \tag{2.3}$$

and

$$\left\| A^{\gamma} \int_s^t S(t-r)u \, \mathrm{d}r \right\| \le C(t-s)^{1-\gamma} \|u\|, \quad u \in H,\ 0 \le s \le t,\ \gamma \in [0,1]. \tag{2.4}$$

ASSUMPTION 2.1 Let $F : L^6(\mathscr{O}) \to H$ be the Nemytskii operator defined by

$$F(X)(x) = f(X(x)) := a_3 X^3(x) + a_2 X^2(x) + a_1 X(x) + a_0,\ a_3 < 0,\ x \in \mathscr{O},\ \text{a.e.}$$

It can be verified that there exist positive constants L_0 and L_1 such that for $X, Y \in E$,

$$\langle X - Y, F(X) - F(Y) \rangle \le L_0 \|X - Y\|^2, \tag{2.5}$$

$$\|F(X) - F(Y)\| \le L_1 \left(1 + \|X\|_E^2 + \|Y\|_E^2\right) \|X - Y\|. \tag{2.6}$$

And for $X, \psi, \psi_1, \psi_2 \in L^6(\mathscr{O})$,

$$\left(DF(X)(\psi)\right)(x) = \left(3a_3 X^2(x) + 2a_2 X(x) + a_1\right) \psi(x), \quad x \in \mathscr{O}, \tag{2.7}$$

$$\left(D^2 F(X)(\psi_1, \psi_2)\right)(x) = (6a_3 X(x) + 2a_2)\psi_1(x)\psi_2(x), \quad x \in \mathscr{O}. \tag{2.8}$$

Moreover, there is a positive constant C such that for $X \in \dot{H}^2$,

$$\|F(X)\|_1^2 = \left\| A^{\frac{1}{2}} F(X) \right\|^2 = \int_{\mathscr{O}} \left| \left(3a_3 X^2(x) + 2a_2 X(x) + a_1\right) \nabla X(x) \right|^2 \mathrm{d}x \le C \left(\|X\|_E^4 + 1\right) \|X\|_1^2, \tag{2.9}$$

$$\begin{aligned} \|F(X)\|_2^2 &= \int_{\mathscr{O}} \left| (6a_3 X(x) + 2a_2) |\nabla X(x)|^2 + \left(3a_3 X^2(x) + 2a_2 X(x) + a_1\right) \Delta X(x) \right|^2 \mathrm{d}x \\ &\le C \left(\|X\|_E^4 + 1\right) \left(\|X\|_2^4 + 1\right), \end{aligned} \tag{2.10}$$

where we have used the Gagliardo–Nirenberg inequality $\|u\|_{L^4(\mathscr{O})} \le C \|\nabla u\|^{\theta} \|u\|^{1-\theta}$ for $u \in H^1$ with $\theta = \frac{d}{4} \in (0,1]$, see e.g., Liu & Röckner (2015, Eq. (5.71)).

We make the following assumption on the stochastic process $\{W(t,\cdot)\}_{t \in [0,T]}$.

Assumption 2.2 Let $\{W(t,\cdot)\}_{t\in[0,T]}$ be a generalized Q-Wiener process on a filtered probability space $(\Omega, \mathscr{F}, \{\mathscr{F}_t\}_{0\le t\le T}, \mathbb{P})$, which can be represented as $W(t,x) := \sum_{n=1}^{\infty} Q^{\frac{1}{2}} e_n(x)\beta_n(t)$, where $\{\beta_n(t)\}_{n\in\mathbb{N}_+}$ is a sequence of independent real valued standard Brownian motions. Assume that for some $\beta \in (0,2]$,

$$\left\| A^{\frac{\beta-1}{2}} Q^{\frac{1}{2}} \right\|_{\mathscr{L}_2(H)} < \infty. \tag{2.11}$$

In the case $\beta \le \frac{d}{2}$, we in addition assume that Q commutes with A.

There are two important cases included in (2.11): the trace-class noise case (i.e., $tr(Q) < \infty$) for $\beta \in [1,2]$, and the space-time white noise case (i.e., $Q = \mathrm{Id}$) for $d = 1$ and $\beta < \frac{1}{2}$. We remark that the condition that Q commutes with A for $\beta \le \frac{d}{2}$ is used to ensure

$$\sup_{t\in[0,T]} \mathbb{E}\left[\left\| \int_0^t S(t-s)\,\mathrm{d}W(s) \right\|_E^p \right] < \infty \tag{2.12}$$

for $p > \frac{2d}{\beta\wedge 1}$, while for the case of $\beta > \frac{d}{2}$, (2.12) can be proved directly by the Sobolev embedding $\dot{H}^{\beta} \hookrightarrow E$; see also Cui *et al.* (2021, Lemma 2).

Assumption 2.3 The initial datum satisfies $\mathbb{E}\big[\exp\{\|X_0\|_{\dot{H}^{\beta}\cap E}\}\big] < \infty$, where β is given in Assumption 2.2. In addition, X_0 is $\mathscr{F}_0/\mathscr{B}(H\cap E)$-measurable.

With the above assumptions, we can obtain the existence, uniqueness and regularity estimates of the mild solution of (1.1). The proofs can be found in e.g., Cerrai (2001, Proposition 6.2.2), Cai *et al.* (2021, Theorem 2.1) and Wang (2020, Theorem 2.6).

Theorem 2.4 Under Assumptions 2.1–2.3, the stochastic Allen–Cahn equation (1.1) has a unique mild solution given by

$$X(t) = S(t)X_0 + \int_0^t S(t-s)F(X(s))\,\mathrm{d}s + \int_0^t S(t-s)\,\mathrm{d}W(s) \quad \text{a.s.}$$

For $p \ge 2$, we have

$$\sup_{t\in[0,T]} \|X(t)\|_{L^p(\Omega;E)} + \sup_{t\in[0,T]} \|X(t)\|_{L^p(\Omega;\dot{H}^{\beta})} \le C_1, \tag{2.13}$$

$$\|X(t)-X(s)\|_{L^p(\Omega;H)} \le C_2 (t-s)^{\frac{\beta\wedge 1}{2}}, \quad 0 \le s < t \le T, \tag{2.14}$$

where constants $C_1, C_2 > 0$ depend on $X_0, p, T, \|A^{\frac{\beta-1}{2}} Q^{\frac{1}{2}}\|_{\mathscr{L}_2(H)}$.

3. Adaptive schemes

In this section, we first introduce the adaptive time-stepping scheme, and assumptions to ensure that the final time T can be attained in finite many steps. Then we present the coupled scheme for the practical use. Finally, we show the strong convergence orders of these numerical schemes.

3.1 *Schemes*

Introduce the adaptive timestep function $\tau : H \to \mathbb{R}_+$. We consider the following adaptive time-stepping scheme, whose spatial discretization is based on the spectral Galerkin method, and temporal direction is the adaptive exponential integrator:

$$X^N_{t_{m+1}} = S^N(\tau_m)X^N_{t_m} + S^N(\tau_m)F^N\left(X^N_{t_m}\right)\tau_m + S^N(\tau_m)P^N\Delta W_m, \quad X^N_0 = P^N X_0, \tag{AE}$$

where $\tau_m := \tau\left(X^N_{t_m}\right)$, $t_0 = 0$, $t_{m+1} = t_m + \tau_m$, $S^N(t) := P^N S(t) = e^{-tA^N}$ with $P^N : H \to H_N$ (:= $span\{e_1, \ldots, e_N\}$) being the spectral projection operator and $A^N := P^N A$, $F^N := P^N F$ and the increment $\Delta W_m := W(t_{m+1}) - W(t_m)$. If the existing time span is longer than T after adding the last timestep, then we take a smaller timestep such that the existing time span just attains T after adding it. Namely, letting M_T be the number of timesteps for a given timestep function τ, if $t_{M_T-1} + \tau_{M_T-1} > T$, then we enforce the last timestep $\tau_{M_T-1} := T - t_{M_T-1}$. In the sequel, we will give some assumptions on the timestep function so that the numerical solution can attain T with finite many timesteps.

The continuous version of (AE) is given by

$$X^N_t = S^N(t)P^N X_0 + \int_0^t S^N\left(t - t_{m_s}\right)F^N\left(X^N_{t_{m_s}}\right)\mathrm{d}s + \int_0^t S^N\left(t - t_{m_s}\right)P^N\,\mathrm{d}W(s), \tag{3.1}$$

where $m_s := \max\{m : t_m \le s\}$.

In order to bound the number of timesteps, we give the following assumption on the adaptive timestep function with the uniform lower bound.

Assumption 3.1 The adaptive timestep function $\tau : H \to \mathbb{R}_+$ is continuous and satisfies that for $X(\omega) \in H$,

$$\tau(X(\omega))^{1-\frac{d}{4}}\|F(X(\omega))\| \le L_2, \quad \text{a.s.}, \tag{3.2}$$

$$\tau(X(\omega)) \ge \tau_{min}, \quad \text{a.s.} \tag{3.3}$$

with positive constants L_2 and τ_{min} independent of ω.

Under the assumption (3.3), we have $M_T \le \frac{T}{\tau_{min}} < \infty$, a.s. That is to say, T is a.s. attainable in finite many timesteps. The power $1 - \frac{d}{4}$ in (3.2) is for technical reason to derive the *a priori* estimate in E-norm of the solution of (AE). Examples for adaptive timestep functions that satisty Assumption 3.1 are given in Section 7.

Remark 3.2 If the expected supremum of the pth moment of the numerical solution is finite, i.e., $\mathbb{E}\left[\sup_{0\le t\le T}\left\|X^N_t\right\|^p\right] < \infty$ for some large $p \ge 2$, then the bounds of adaptive timestep function in Assumption 3.1 can be weaken to the adaptive ones:

$$\tau(X(\omega))^{1-\frac{d}{4}}\|F(X(\omega))\| \le L_2(\omega), \quad \text{a.s.}, \tag{3.4}$$

$$\tau(X(\omega)) \ge \left(\zeta_1\|X(\omega)\|^{q_0} + \zeta_2\right)^{-1}, \quad \text{a.s.} \tag{3.5}$$

with positive constants ζ_1, ζ_2 and q_0 independent of ω, and $q_0 \le p$, $\|L_2(\cdot)\|_{L^{2p}(\Omega)} < \infty$. In this setting, T is still a.s. attainable, i.e.,

$$\mathbb{E}[M_T] \le T\mathbb{E}\left[\sup_{0\le t_m\le T} \frac{1}{\tau\left(X_{t_m}^N\right)}\right] \le T\mathbb{E}\left[\sup_{0\le t_m\le T}\left(\zeta_1\left\|X_{t_m}^N\right\|^{q_0}+\zeta_2\right)\right] < \infty.$$

Notice that for the trace-class noise case (i.e., $tr(Q) < \infty$ with $\beta \in [1,2]$ in Assumption 2.2), under the assumption

$$\langle X(\omega), F(X(\omega))\rangle + \frac{1}{2}\tau(X(\omega))\|F(X(\omega))\|^2 \le \bar{L}_2\|X(\omega)\|^2 + L_3, \quad \text{a.s.} \tag{3.6}$$

with positive constants $\bar{L}_2$ and L_3 independent of ω, following the approach of Fang & Giles (2020, Theorem 1) and combining with the contractivity of the semigroup $\{S(t), t \ge 0\}$ in H, we can get the finiteness of the expected supremum of the pth moment of the numerical solution. However, for $\beta \in (0,1)$, we haven't derived the finiteness of the expected supremum of the pth moment of the numerical solution. Hence, the main result in this paper is still hold under assumptions (3.4)–(3.6) when $\beta \in [1,2]$.

In practice, instead of verifying whether a timestep function satisfies the lower bound in (3.3) or (3.5), people usually introduce a backstop scheme with a uniform timestep and couple it with (AE) to ensure that a simulation over the interval $[0,T]$ can be completed in a finite number of timesteps (see Kelly & Lord (2018) for the case of SODEs). More precisely, if $\tau_m < \tau_{min}$ at time t_m, then we apply a single step of some convergent scheme $\Psi : H_N \times \mathbb{R}_+ \times H \to H_N$, which is called the backstop scheme over a timestep of length τ_{min} instead, i.e.,

$$X_{t_{m+1}}^{N,C} = \Phi\left(X_{t_m}^{N,C}, \tau_m, \Delta W_m\right)\chi_{\{\tau_m\ge\tau_{min}\}} + \Psi\left(X_{t_m}^{N,C}, \tau_{min}, \Delta W_m\right)\chi_{\{\tau_m<\tau_{min}\}}, \tag{3.7}$$

where the map $\Phi : H_N \times \mathbb{R}_+ \times H \to H_N,\ (x,h,y) \mapsto S^N(h)(x + F^N(x)h + y)$ denotes the scheme (AE).

The backstop scheme is usually chosen to be an explicit and convergent scheme, for instance, the tamed exponential integrator (see Wang, 2020), the nonlinearity-truncated exponential integrator and the linear-implicit nonlinearity-truncated scheme (see Becker & Jentzen, 2019). In the following, we take the backstop scheme Ψ as the tamed exponential integrator:

$$\Psi(x,h,y) = S^N(h)\left(x + \frac{F^N(x)h}{1+\|F^N(x)\|h} + y\right), \tag{3.8}$$

and give estimates of the corresponding coupled scheme

$$X_{t_{m+1}}^{N,(1)} = S^N(\tau_m)\left[X_{t_m}^{N,(1)} + F^N\left(X_{t_m}^{N,(1)}\right)\tau_m + \Delta W_m\right]\chi_{\{\tau_m\ge\tau_{min}\}} \tag{CAU 1}$$
$$+ S^N(\tau_{min})\left[X_{t_m}^{N,(1)} + \frac{F^N\left(X_{t_m}^{N,(1)}\right)\tau_{min}}{1+\left\|F^N\left(X_{t_m}^{N,(1)}\right)\right\|\tau_{min}} + \Delta W_m\right]\chi_{\{\tau_m<\tau_{min}\}}.$$

Moreover, the adaptive timestep function that satisfies (3.2) can be chosen as, e.g., $\tau_m = \tau\left(X_{t_m}^N\right) = \left(\frac{L_2}{\|F(X_{t_m}^N)\|+c}\right)^{\frac{4}{4-d}}$ with some $c > 0$. We can write the continuous versions of (CAU 1) into the compact integral form by defining a new timestep function denoted by τ_{new} with

$$\tau_{new} := \begin{cases} \tau_m, & \text{for } \tau_m \geq \tau_{min}, \\ \tau_{min}, & \text{for } \tau_m < \tau_{min}. \end{cases}$$

Then $t_{m+1} = t_m + \tau_{new}$, and the continuous versions of (CAU 1) is

$$X_t^{N,(1)} = S^N(t)P^N X_0 + \int_0^t S^N\left(t - t_{m_s}\right)\left(F^N\left(X_{t_{m_s}}^{N,(1)}\right)\chi_{\{\tau_{m_s}\geq\tau_{min}\}} + \frac{F^N\left(X_{t_{m_s}}^{N,(1)}\right)}{1+\left\|F^N\left(X_{t_{m_s}}^{N,(1)}\right)\right\|\tau_{min}}\chi_{\{\tau_{m_s}<\tau_{min}\}}\right)\mathrm{d}s + \int_0^t S^N\left(t-t_{m_s}\right)P^N\,\mathrm{d}W(s). \tag{3.9}$$

3.2 *Main result*

In this subsection, we give strong convergence orders for the scheme (AE) as well as the coupled scheme (CAU 1), whose proofs are postponed to Section 5 and Appendix B, respectively.

Since the timestep function τ is determined by the numerical solution, we need to make a modification when considering the convergence of adaptive time-stepping schemes (see Fang & Giles, 2020). Namely, for a given timestep function τ that satisfies Assumption 3.1, we introduce the refined timestep function τ^δ controlled by a scalar parameter $\delta \in (0,1)$ and consider the convergence when $\delta \to 0$ as well as the order with respect to δ.

Assumption 3.3 The refined timestep function τ^δ satisfies that for $X(\omega) \in H$,

$$\delta \min\{T, \tau(X(\omega))\} \leq \tau^\delta(X(\omega)) \leq \min\{T\delta, \tau(X(\omega))\}, \quad \text{a.s.}$$

Examples of τ and τ^δ are given in Section 7. We also remark that in this setting, the lower bound in (3.3) is defined as $\tau_{min}^\delta := \delta\tau_{min}$. With this assumption in hand, in the following, we present strong convergence orders of schemes (AE) and (CAU 1) with the timestep function τ^δ. Before that, we put an additional assumption on the initial datum, which is used to get the *a priori* estimates of numerical solutions in E-norm, see Section 4 for details.

Assumption 3.4 The initial datum of (1.1) satisfies that $\sup_{N\in\mathbb{N}_+}\mathbb{E}\left[\exp\left\{\left\|P^N X_0\right\|_E\right\}\right] < \infty$.

Based on the Sobolev embedding theorem, Assumption 3.4 is fulfilled if for $\beta > \frac{d}{2}$, $\sup_{N\in\mathbb{N}_+}\mathbb{E}\left[\exp\{\|X_0\|_{\dot{H}^\beta}\}\right] < \infty$.

Theorem 3.5 Under Assumptions 2.1–2.3, 3.1, 3.3, and 3.4, for $p \geq 2$,

$$\sup_{0\leq t\leq T}\left\|X(t) - \mathbb{X}_t^N\right\|_{L^p(\Omega;H)} \leq C\left(\lambda_N^{-\frac{\beta}{2}} + \delta^{\frac{\beta}{2}}\right),$$

where $\mathbb{X}_t^N$ is the numerical solution X_t^N (or $X_t^{N,(1)}$) of the scheme (AE) (or (CAU 1)), and $C > 0$ depends on $p, T, \tau_{min}, L_0, L_1, L_2, X_0$ and $\|A^{\frac{\beta-1}{2}} Q^{\frac{1}{2}}\|_{\mathscr{L}_2(H)}$.

Noting that $\lambda_N \sim N^{-\frac{2}{d}}$, the convergence result in Theorem 3.5 can be rewritten as $\sup_{0\le t\le T} \|X(t) - \mathbb{X}_t^N\|_{L^p(\Omega;H)} \le C(N^{-\frac{\beta}{d}} + \delta^{\frac{\beta}{2}})$. Then we say that the spatial convergence order is $\frac{\beta}{d}$ with respect to N.

Remark 3.6 As stated in Remark 3.2, for the trace-class noise case (i.e., $\beta \in [1,2]$), the bounds of the adaptive timestep function in Assumption 3.1 can be weaken to (3.4) and (3.5). Similarly to the definition of (3.7), when the critical parameter for the adaptive timestep size is $\left(\zeta_1 \|\widetilde{X_{t_m}^{N,C}}\|^{q_0} + \zeta_2\right)^{-1}$, the coupled scheme can also be defined as

$$\widetilde{X_{t_{m+1}}^{N,C}} = \Phi\left(\widetilde{X_{t_m}^{N,C}}, \tau_m, \Delta W_m\right) \chi_{\left\{\tau_m \ge \left(\zeta_1 \left\|\widetilde{X_{t_m}^{N,C}}\right\|^{q_0} + \zeta_2\right)^{-1}\right\}} + \Psi\left(\widetilde{X_{t_m}^{N,C}}, \tau_m, \Delta W_m\right) \chi_{\left\{\tau_m < \left(\zeta_1 \left\|\widetilde{X_{t_m}^{N,C}}\right\|^{q_0} + \zeta_2\right)^{-1}\right\}}. \tag{3.10}$$

If we still choose Ψ to be (3.8), and denote the solution of the continuous version of the corresponding coupled scheme by $\widetilde{X_t^{N,(1)}}$, then the similar proof as that of Theorem 3.5 yields that: under Assumptions 2.1–2.3, Assumptions 3.3–3.4 and Eq. 3.4–3.6, for $p \ge 2$,

$$\sup_{0\le t\le T} \left\| X(t) - \widetilde{X_t^{N,(1)}} \right\|_{L^p(\Omega;H)} \le C\left(\lambda_N^{-\frac{\beta}{2}} + \delta^{\frac{\beta}{2}}\right) \le C\left(\lambda_N^{-\frac{\beta}{2}} + (\mathbb{E}[M_{T,\delta}])^{-\frac{\beta}{2}}\right),$$

where $M_{T,\delta}$ is the number of timesteps for the given timestep function τ^δ, and $C > 0$ depends on $p, T, \zeta_1, \zeta_2, L_i (i = 0, \dots, 3), \bar{L}_2, X_0$ and $\|A^{\frac{\beta-1}{2}} Q^{\frac{1}{2}}\|_{\mathscr{L}_2(H)}$.

Remark 3.7 We remark that it is interesting to investigate if one can obtain the strong convergence order for the coupled scheme directly from some error estimates of schemes Φ and Ψ. For this problem, a fundamental convergence theorem that characterizes the relation between the local error and the global error might be helpful. For the study of such a theorem, we refer to Milstein & Tretyakov (2021) for the case of SODEs with either the globally or locally Lipschitz drift, and to Chen & Hong (2016) for the case of SPDEs with the globally Lipschitz drift. However, for SPDEs with the non-Lipschitz drift, there has been no work on such theorem. We leave this as the future work.

4. Estimates of numerical solutions

In this section, we analyze the *a priori* estimates of numerical solutions (AE) and (CAU 1) in E-norm and $\dot{H}^\beta$-norm, respectively. Proofs of all the results in this section are given in Appendix A for readers' convenience.

We give the following lemma on the properties of the semigroup $\{S(t), t \ge 0\}$, which is important in the *a priori* estimates of numerical solutions.

Lemma 4.1 We have (*i*) $\sup_{N\in\mathbb{N}_+} \|A^\rho P^N S(t)u\|_E \le C(\rho, d)\left(t^{-\rho} + t^{-\rho-\frac{d}{4}}\right)\|u\|$, for $\rho \ge 0,\ t > 0,\ u \in H$; (*ii*) $\sup_{N\in\mathbb{N}_+} \|P^N S(t)u\|_E \le C(\rho, d) t^{\frac{2\rho-d}{4}} \|u\|_\rho$, for $\rho \in \left[0, \frac{d}{2}\right),\ t > 0,\ u \in \dot{H}^\rho$.

Let $v^N, z^N : [0,T] \to H_N$ satisfy the perturbed differential equation

$$\begin{cases} \mathrm{d}v_t^N = \left(-A^N v_t^N + F^N\left(v_t^N + z_t^N\right)\right)\,\mathrm{d}t, & t \in (0,T], \\ v_0^N = 0. \end{cases} \tag{4.1}$$

In what follows, we aim to show that the E-norm of the solution v_t^N of (4.1) can be controlled by the $L^\eta([0,t];E)$-norm of the perturbation $z_\cdot^N$ with some $\eta > 0$, which plays a crucial role in deriving the moment bounds in E-norm for the scheme (AE).

LEMMA 4.2 The solution of (4.1) satisfies

$$\left\|v_t^N\right\|_E \le C\left[1 + \int_0^t \left((t-s)^{-\frac{d}{4}}\left\|z_s^N\right\|_E^3 + \left\|z_s^N\right\|_E^{72}\right)\mathrm{d}s\right], \quad t \in (0,T]$$

with $C > 0$.

With these preparations, we can establish the *a priori* estimates of the numerical solution of (AE) in E-norm by the standard bootstrap argument; see Wang (2020) for the description of this approach for the nonlinearity-tamed scheme with the uniform timestep.

PROPOSITION 4.3 Under conditions in Theorem 3.5, we have for $p \ge 2$,

$$\sup_{N\in\mathbb{N}_+}\sup_{0\le t\le T}\left\|X_t^N\right\|_{L^p(\Omega;E)} \le C, \tag{4.2}$$

where X_t^N is the numerical solution of (AE) with timestep function τ^δ for some $\delta \in (0,1)$, and $C > 0$ depends on $p, T, L_1, L_2, \tau_{min}, X_0$ and $\|A^{\frac{\beta-1}{2}}Q^{\frac{1}{2}}\|_{\mathscr{L}_2(H)}$.

With the *a priori* estimate of the numerical solution of (AE) in E-norm in hand, we can obtain the following *a priori* estimate in $\dot{H}^\beta$-norm directly by means of the mild form of the solution. A standard argument gives the Hölder continuity of the numerical solution; see also Becker & Jentzen (2019, Lemma 4.3).

PROPOSITION 4.4 Under conditions in Proposition 4.3, for $p \ge 2$, $\beta \in (0,2]$ and $\gamma \in (0,\beta]$, we have

$$\sup_{N\in\mathbb{N}_+}\sup_{0\le t\le T}\left\|X_t^N\right\|_{L^p(\Omega;\dot{H}^\beta)} \le C_1, \tag{4.3}$$

$$\sup_{N\in\mathbb{N}_+}\left\|X_t^N - X_s^N\right\|_{L^p(\Omega;\dot{H}^\gamma)} \le C_2(t-s)^{\frac{(\beta-\gamma)\wedge 1}{2}}, \quad 0 \le s < t \le T, \tag{4.4}$$

where constants $C_1, C_2 > 0$ depend on $p, T, L_1, L_2, \tau_{min}, X_0$ and $\|A^{\frac{\beta-1}{2}}Q^{\frac{1}{2}}\|_{\mathscr{L}_2(H)}$.

With the above two propositions, combining regularity estimates of the tamed exponential integrator, which can be proved similarly as Propositions 4.3–4.4 and Wang (2020), we get the following the *a priori* estimates in $\dot{H}^\beta$ and the Hölder continuity for the numerical solution of (CAU 1).

COROLLARY 4.5 Under conditions in Proposition 4.3, for $p \ge 2$, $\beta \in (0,2]$ and $\gamma \in (0,\beta]$, we have

$$\sup_{N\in\mathbb{N}_+}\sup_{0\le t\le T}\left\|X_t^{N,(1)}\right\|_{L^p(\Omega;E\cap\dot{H}^\beta)} \le C_1,$$

$$\sup_{N\in\mathbb{N}_+}\left\|X_t^{N,(1)} - X_s^{N,(1)}\right\|_{L^p(\Omega;\dot{H}^\gamma)} \le C_2(t-s)^{\frac{(\beta-\gamma)\wedge 1}{2}}, \quad 0 \le s < t \le T,$$

where constants $C_1, C_2 > 0$ depend on $p, T, L_1, L_2, \tau_{min}, X_0$ and $\|A^{\frac{\beta-1}{2}}Q^{\frac{1}{2}}\|_{\mathscr{L}_2(H)}$.

5. Proof of the main result

In this section, based on the *a priori* estimates of numerical solutions presented in Section 4, we show the proof of the convergence order of the scheme (AE) in Theorem 3.5, and leave that of the coupled scheme (CAU 1) to Appendix B.

Proof of Theorem 3.5. By introducing an auxiliary process,

$$Y_t^N = S^N(t)P^N X_0 + \int_0^t S^N(t-s)F^N(X(s))\,\mathrm{d}s + \int_0^t S^N\left(t - t_{m_s}\right) P^N\,\mathrm{d}W(s),$$

the error can be divided into the following terms:

$$\left\|X(t) - X_t^N\right\|_{L^p(\Omega;H)} \le \|X(t) - P^N X(t)\|_{L^p(\Omega;H)} + \left\|P^N X(t) - Y_t^N\right\|_{L^p(\Omega;H)} + \left\|Y_t^N - X_t^N\right\|_{L^p(\Omega;H)}.$$

The term $\|X(t) - P^N X(t)\|_{L^p(\Omega;H)}$ can be estimated as

$$\|X(t) - P^N X(t)\|_{L^p(\Omega;H)} = \left\|A^{-\frac{\beta}{2}}(\mathrm{Id} - P^N)A^{\frac{\beta}{2}}X(t)\right\|_{L^p(\Omega;H)} \le C\lambda_N^{-\frac{\beta}{2}}\|X(t)\|_{L^p(\Omega;\dot{H}^\beta)}. \tag{5.1}$$

For the term $\|P^N X(t) - Y_t^N\|_{L^p(\Omega;H)}$, using the Burkholder–Davis–Gundy inequality (see e.g., Kruse, 2014, Proposition 2.12), (2.2)–(2.3), the Hölder inequality and Assumption 3.3 gives that

$$\begin{aligned}
&\left\|P^N X(t) - Y_t^N\right\|_{L^p(\Omega;H)}^p = \left\|\int_0^t S^N(t-s)\left(\mathrm{Id} - S^N\left(s - t_{m_s}\right)\right)\mathrm{d}W(s)\right\|_{L^p(\Omega;H)}^p \\
&\le C(p)\mathbb{E}\left[\left(\int_0^t \left\|S^N(t-s)\left(\mathrm{Id} - S^N\left(s - t_{m_s}\right)\right)Q^{\frac{1}{2}}\right\|_{\mathscr{L}_2(H)}^2 \mathrm{d}s\right)^{\frac{p}{2}}\right] \\
&= C(p)\mathbb{E}\left[\left(\int_0^t \left\|A^{-\frac{\beta}{2}}\left(\mathrm{Id} - S^N\left(s - t_{m_s}\right)\right)A^{\frac{1}{2}}S^N(t-s)A^{\frac{\beta-1}{2}}Q^{\frac{1}{2}}\right\|_{\mathscr{L}_2(H)}^2 \mathrm{d}s\right)^{\frac{p}{2}}\right] \\
&\le C(p)\mathbb{E}\left[\left(\int_0^t \sum_{j=1}^\infty \left(s - t_{m_s}\right)^{\frac{\beta}{2}}\left\|A^{\frac{1}{2}}S^N(t-s)A^{\frac{\beta-1}{2}}Q^{\frac{1}{2}}e_j\right\|^2 \mathrm{d}s\right)^{\frac{p}{2}}\right] \\
&\le C(p,T)\left\|A^{\frac{\beta-1}{2}}Q^{\frac{1}{2}}\right\|_{\mathscr{L}_2(H)}^p \delta^{\frac{\beta p}{2}} \le C\left(p,T,\left\|A^{\frac{\beta-1}{2}}Q^{\frac{1}{2}}\right\|_{\mathscr{L}_2(H)}\right)\delta^{\frac{\beta p}{2}}.
\end{aligned} \tag{5.2}$$

For the estimate of the term $\|Y_t^N - X_t^N\|_{L^p(\Omega;H)}$, combining the differential form

$$\mathrm{d}\left(X_t^N - Y_t^N\right) = -A^N\left(X_t^N - Y_t^N\right)\mathrm{d}t + \left(S^N\left(t - t_{m_t}\right)F^N\left(X_{t_{m_t}}^N\right) - F^N(X(t))\right)\mathrm{d}t,$$

and applying the Taylor formula yield

$$\begin{aligned}\left\|X_t^N - Y_t^N\right\|^2 &= 2\int_0^t \left\langle X_s^N - Y_s^N, -A^N\left(X_s^N - Y_s^N\right)\right\rangle \mathrm{d}s \\ &\quad + 2\int_0^t \left\langle X_s^N - Y_s^N, S^N\left(s - t_{m_s}\right)F^N\left(X_{t_{m_s}}^N\right) - F^N(X(s))\right\rangle \mathrm{d}s \\ &= 2\int_0^t \left\langle X_s^N - Y_s^N, -A^N\left(X_s^N - Y_s^N\right)\right\rangle \mathrm{d}s + 2\int_0^t \left\langle X_s^N - Y_s^N, S^N\left(s - t_{m_s}\right)F^N\left(X_{t_{m_s}}^N\right) - F^N\left(X_s^N\right)\right\rangle \mathrm{d}s \\ &\quad + 2\int_0^t \left\langle X_s^N - Y_s^N, F^N\left(X_s^N\right) - F^N\left(Y_s^N\right)\right\rangle \mathrm{d}s + 2\int_0^t \left\langle X_s^N - Y_s^N, F^N\left(Y_s^N\right) - F^N(X(s))\right\rangle \mathrm{d}s \\ &\le (2L_0 + 1)\int_0^t \left\|X_s^N - Y_s^N\right\|^2 \mathrm{d}s + \int_0^t \left\|F^N\left(Y_s^N\right) - F^N(X(s))\right\|^2 \mathrm{d}s + 2J_{1,1}(t) + 2J_{1,2}(t), \qquad (5.3)\end{aligned}$$

where we have used the condition (2.5) and the Young inequality, and

$$J_{1,1}(t) := \int_0^t \left\langle X_s^N - Y_s^N, \left(S^N\left(s - t_{m_s}\right) - \mathrm{Id}\right)F^N\left(X_{t_{m_s}}^N\right)\right\rangle \mathrm{d}s, \quad J_{1,2}(t) := \int_0^t \left\langle X_s^N - Y_s^N, F^N\left(X_{t_{m_s}}^N\right) - F^N\left(X_s^N\right)\right\rangle \mathrm{d}s.$$

Estimate of $J_{1,1}$. For the case $\beta = 1$ or $\beta = 2$, it follows from (2.2), (2.9)–(2.10) and the Young inequality that

$$\begin{aligned}J_{1,1}(t) &\le \frac{1}{2}\int_0^t \left\|X_s^N - Y_s^N\right\|^2 \mathrm{d}s + \frac{1}{2}\int_0^t \left\|A^{-\frac{\beta}{2}}\left(S^N\left(s - t_{m_s}\right) - \mathrm{Id}\right)A^{\frac{\beta}{2}}F\left(X_{t_{m_s}}^N\right)\right\|^2 \mathrm{d}s \\ &\le \frac{1}{2}\int_0^t \left\|X_s^N - Y_s^N\right\|^2 \mathrm{d}s + C\int_0^t \left(s - t_{m_s}\right)^\beta \left(1 + \left\|X_{t_{m_s}}^N\right\|_E^4\right)\left(\left\|X_{t_{m_s}}^N\right\|_\beta^4 + 1\right).\end{aligned}$$

This, combining Assumption 3.3 and Propositions 4.3 and 4.4 leads to

$$\|J_{1,1}(t)\|_{L^p(\Omega)} \le \frac{1}{2}\int_0^t \left\|X_s^N - Y_s^N\right\|_{L^{2p}(\Omega;H)}^2 \mathrm{d}s + C\delta^\beta, \quad t \in [0,T].$$

For the case $\beta \in (0,1) \cup (1,2)$, we need to further split the term $J_{1,1}$ into three parts, which are denoted by $J_{1,1}^i, i = 1,2,3$, based on

$$\begin{aligned}X_s^N - Y_s^N &= \int_0^s \left(S^N\left(s - t_{m_r}\right) - S^N(s - r)\right)F^N\left(X_{t_{m_r}}^N\right)\mathrm{d}r + \int_0^s S^N(s - r)\left(F^N\left(X_{t_{m_r}}^N\right) - F^N\left(X_r^N\right)\right)\mathrm{d}r \\ &\quad + \int_0^s S^N(s - r)\left(F^N\left(X_r^N\right) - F^N(X(r))\right)\mathrm{d}r. \qquad (5.4)\end{aligned}$$

Namely,

$$J_{1,1}^1(t) := \int_0^t \left\langle \int_0^s S^N(s - r)\left(S^N\left(r - t_{m_r}\right) - \mathrm{Id}\right)F^N\left(X_{t_{m_r}}^N\right)\mathrm{d}r,\ \left(S^N\left(s - t_{m_s}\right) - \mathrm{Id}\right)F^N\left(X_{t_{m_s}}^N\right)\right\rangle \mathrm{d}s,$$

$$J_{1,1}^2(t) := \int_0^t \left\langle \int_0^s S^N(s - r)\left(F^N\left(X_{t_{m_r}}^N\right) - F^N\left(X_r^N\right)\right)\mathrm{d}r,\ \left(S^N\left(s - t_{m_s}\right) - \mathrm{Id}\right)F^N\left(X_{t_{m_s}}^N\right)\right\rangle \mathrm{d}s,$$

$$J_{1,1}^3(t) := \int_0^t \left\langle \int_0^s S^N(s - r)\left(F^N\left(X_r^N\right) - F^N(X(r))\right)\mathrm{d}r,\ \left(S^N\left(s - t_{m_s}\right) - \mathrm{Id}\right)F^N\left(X_{t_{m_s}}^N\right)\right\rangle \mathrm{d}s.$$

For the term $J_{1,1}^1$, by the Young inequality, we obtain that for $\beta \in (0,1)$,

$$J_{1,1}^1(t) = \int_0^t \left\langle \int_0^s A^{\beta} S^N(s-r) A^{-\frac{\beta}{2}} \left(S^N\left(r-t_{m_r}\right) - \mathrm{Id}\right) F^N\left(X_{t_{m_r}}^N\right) \mathrm{d}r,\ A^{-\frac{\beta}{2}}\left(S^N\left(s-t_{m_s}\right)-\mathrm{Id}\right) F^N\left(X_{t_{m_s}}^N\right)\right\rangle \mathrm{d}s$$
$$\leq C \int_0^t \int_0^s (s-r)^{-\beta}\left[\left(r-t_{m_r}\right)^{\beta}\left\|F\left(X_{t_{m_r}}^N\right)\right\|^2 + \left(s-t_{m_s}\right)^{\beta}\left\|F\left(X_{t_{m_s}}^N\right)\right\|^2\right] \mathrm{d}r\,\mathrm{d}s.$$

When $\beta \in (1,2)$, applying (2.1)–(2.2) and (2.9), the term $J_{1,1}^1$ can be estimated as

$$J_{1,1}^1(t) = \int_0^t \left\langle \int_0^s A^{\beta-1} S^N(s-r) A^{-\frac{\beta}{2}} \left(S^N\left(r-t_{m_r}\right)-\mathrm{Id}\right) A^{\frac{1}{2}} F^N\left(X_{t_{m_r}}^N\right) \mathrm{d}r,\ A^{-\frac{\beta}{2}}\left(S^N\left(s-t_{m_s}\right)-\mathrm{Id}\right) A^{\frac{1}{2}} F^N\left(X_{t_{m_s}}^N\right)\right\rangle \mathrm{d}s$$
$$\leq C \int_0^t \int_0^s (s-r)^{1-\beta}\left[\left(r-t_{m_r}\right)^{\beta}\left(1+\left\|X_{t_{m_r}}^N\right\|_E^4\right)\left\|X_{t_{m_r}}^N\right\|_1^2 + \left(s-t_{m_s}\right)^{\beta}\left(1+\left\|X_{t_{m_s}}^N\right\|_E^4\right)\left\|X_{t_{m_s}}^N\right\|_1^2\right] \mathrm{d}r\,\mathrm{d}s.$$

Hence, we derive from Propositions 4.3 and 4.4 that for $\beta \in (0,1) \cup (1,2)$,

$$\left\|J_{1,1}^1(t)\right\|_{L^p(\Omega)} \leq C\delta^{\beta}, \quad t \in [0,T].$$

For the term $J_{1,1}^2$, when $\beta \in (0,1)$,

$$J_{1,1}^2(t) = \int_0^t \left\langle \int_0^s A^{\frac{\beta}{2}} S^N(s-r)\left(F^N\left(X_{t_{m_r}}^N\right) - F^N\left(X_r^N\right)\right) \mathrm{d}r,\ A^{-\frac{\beta}{2}}\left(S^N\left(s-t_{m_s}\right)-\mathrm{Id}\right) F^N\left(X_{t_{m_s}}^N\right)\right\rangle \mathrm{d}s$$
$$\leq C \int_0^t \int_0^s (s-r)^{-\frac{\beta}{2}}\left[\left(1+\left\|X_{t_{m_r}}^N\right\|_E^4 + \left\|X_r^N\right\|_E^4\right)\left\|X_r^N - X_{t_{m_r}}^N\right\|^2 + \left(s-t_{m_s}\right)^{\beta}\left\|F\left(X_{t_{m_s}}^N\right)\right\|^2\right] \mathrm{d}r\,\mathrm{d}s,$$

which gives $\|J_{1,1}^2(t)\|_{L^p(\Omega)} \leq C\delta^{\beta}$ based on Proposition 4.4. And when $\beta \in (1,2)$, since the order of the Hölder continuity of X_r^N is $\frac{1}{2}$ in H-norm, hence the term $J_{1,1}^2$ need to be further split, based on the Taylor formula to F and the fact that

$$X_r^N - X_{t_{m_r}}^N = \left(S^N\left(r-t_{m_r}\right)-\mathrm{Id}\right) X_{t_{m_r}}^N + \int_{t_{m_r}}^r S^N\left(r-t_{m_r}\right) F^N\left(X_{t_{m_r}}^N\right) \mathrm{d}u + \int_{t_{m_r}}^r S^N\left(r-t_{m_r}\right) \mathrm{d}W(u).$$

Namely, we arrive at

$$J_{1,1}^2(t) = -\int_0^t \left\langle \int_0^s A^{\frac{\beta-1}{2}} S^N(s-r) DF\left(X_{t_{m_r}}^N\right)\left(S^N\left(r-t_{m_r}\right)-\mathrm{Id}\right) X_{t_{m_r}}^N \mathrm{d}r,\ A^{-\frac{\beta}{2}}\left(S^N\left(s-t_{m_s}\right)-\mathrm{Id}\right) A^{\frac{1}{2}} F^N\left(X_{t_{m_s}}^N\right)\right\rangle \mathrm{d}s$$
$$- \int_0^t \left\langle \int_0^s A^{\frac{\beta-1}{2}} S^N(s-r) DF\left(X_{t_{m_r}}^N\right) \int_{t_{m_r}}^r S^N\left(r-t_{m_r}\right) F^N\left(X_{t_{m_r}}^N\right) \mathrm{d}u\,\mathrm{d}r,\ A^{-\frac{\beta}{2}}\left(S^N\left(s-t_{m_s}\right)-\mathrm{Id}\right) A^{\frac{1}{2}} F^N\left(X_{t_{m_s}}^N\right)\right\rangle \mathrm{d}s$$
$$- \int_0^t \left\langle \int_0^s A^{\frac{\beta-1}{2}} S^N(s-r) DF\left(X_{t_{m_r}}^N\right) \int_{t_{m_r}}^r S^N\left(r-t_{m_r}\right) \mathrm{d}W(u)\,\mathrm{d}r,\ A^{-\frac{\beta}{2}}\left(S^N\left(s-t_{m_s}\right)-\mathrm{Id}\right) A^{\frac{1}{2}} F^N\left(X_{t_{m_s}}^N\right)\right\rangle \mathrm{d}s$$
$$- \int_0^t \left\langle \int_0^s S^N(s-r) R_F\left(X_{t_{m_r}}^N, X_r^N\right) \mathrm{d}r,\ \left(S^N\left(s-t_{m_s}\right)-\mathrm{Id}\right) F^N\left(X_{t_{m_s}}^N\right)\right\rangle \mathrm{d}s =: II_1(t) + II_2(t) + II_3(t) + II_4(t)$$

with the remainder

$$R_F\left(X^N_{t_{m_r}}, X^N_r\right) := \int_0^1 D^2F\left(X^N_{t_{m_r}} + (1-\theta)\left(X^N_r - X^N_{t_{m_r}}\right)\right)\left(\left(X^N_r - X^N_{t_{m_r}}\right), \left(X^N_r - X^N_{t_{m_r}}\right)\right)(1-\theta)\,\mathrm{d}\theta.$$

Then we treat the four terms $II_i, i = 1,2,3,4$ one by one. It follows from (2.1)–(2.2) and (2.7) that

$$II_1(t) \le C\int_0^t\int_0^s (s-r)^{-\frac{\beta-1}{2}}\left(1+\left\|X^N_{t_{m_r}}\right\|^2_E\right)(r-t_{m_r})^{\frac{\beta}{2}}\left\|X^N_{t_{m_r}}\right\|_\beta \mathrm{d}r\,(s-t_{m_s})^{\frac{\beta}{2}}\left\|F\left(X^N_{t_{m_s}}\right)\right\|_1 \mathrm{d}s,$$
$$II_2(t) \le C\int_0^t\int_0^s (s-r)^{-\frac{\beta-1}{2}}\left(1+\left\|X^N_{t_{m_r}}\right\|^2_E\right)(r-t_{m_r})\left\|F\left(X^N_{t_{m_r}}\right)\right\| \mathrm{d}r\,(s-t_{m_s})^{\frac{\beta}{2}}\left\|F\left(X^N_{t_{m_s}}\right)\right\|_1 \mathrm{d}s,$$

which combining Propositions 4.3 and 4.4 gives $\|II_1(t)\|_{L^p(\Omega)} + \|II_2(t)\|_{L^p(\Omega)} \le C\delta^\beta,\ t \in [0,T]$. For the term II_3, applying the stochastic Fubini theorem (see e.g., Kruse, 2014, Theorem 4.18) and the Burkholder–Davis–Gundy inequality (see e.g., Kruse, 2014, Proposition 2.12), and combining Propositions 4.3 and 4.4 yield

$$\begin{aligned}
&\|II_3(t)\|_{L^p(\Omega)}\\
&\le \int_0^t \left\|\sum_{i=0}^{m_s}\int_{t_i}^{t_{i+1}}\int_{t_i}^{t_{i+1}} \chi_{[t_i,r)}(u)\chi_{[t_i,s)}(r)A^{\frac{\beta-1}{2}}S^N(s-r)DF\left(X^N_{t_i}\right)S^N(r-t_i)\,\mathrm{d}W(u)\,\mathrm{d}r\right\|_{L^{2p}(\Omega;H)}\\
&\quad\times\left\|(s-t_{m_s})^{\frac{\beta}{2}}\left\|F\left(X^N_{t_{m_s}}\right)\right\|_1\right\|_{L^{2p}(\Omega)}\mathrm{d}s\\
&\le C\delta^{\frac{\beta}{2}}\int_0^t\left\|\sum_{i=0}^{m_s}\int_{t_i}^{t_{i+1}}\int_{t_i}^{t_{i+1}} \chi_{[t_i,r)}(u)\chi_{[t_i,s)}(r)A^{\frac{\beta-1}{2}}S^N(s-r)DF\left(X^N_{t_i}\right)S^N(r-t_i)\,\mathrm{d}r\,\mathrm{d}W(u)\right\|_{L^{2p}(\Omega;H)}\mathrm{d}s\\
&\le C\delta^{\frac{\beta}{2}}\int_0^t\left(\int_0^s\left\|\int_{t_{m_u}}^{t_{m_u}+\tau^\delta_{m_u}} \chi_{[t_{m_u},r)}(u)\chi_{[t_{m_u},s)}(r)A^{\frac{\beta-1}{2}}S^N(s-r)DF\left(X^N_{t_{m_u}}\right)S^N(r-t_{m_u})Q^{\frac12}\,\mathrm{d}r\right\|^2_{L^{2p}(\Omega;\mathscr{L}_2(H))}\mathrm{d}u\right)^{\frac12}\mathrm{d}s\\
&\le C\delta^{\frac{1+\beta}{2}}\int_0^t\Bigg[\left(\int_0^{(s-T\delta)\vee 0}\left\|\sum_{j=1}^\infty\int_{t_{m_u}}^{t_{m_u}+\tau^\delta_{m_u}}\left\|A^{\frac{\beta-1}{2}}S^N(s-r)DF\left(X^N_{t_{m_u}}\right)S^N(r-t_{m_u})Q^{\frac12}e_j\right\|^2\mathrm{d}r\right\|_{L^p(\Omega)}\mathrm{d}u\right)^{\frac12}\\
&\quad+\left(\int_{(s-T\delta)\vee 0}^s\left\|\sum_{j=1}^\infty\int_{(s-2T\delta)\vee 0}^s\left\|A^{\frac{\beta-1}{2}}S^N(s-r)DF\left(X^N_{t_{m_u}}\right)S^N(r-t_{m_u})Q^{\frac12}e_j\right\|^2\mathrm{d}r\right\|_{L^p(\Omega)}\mathrm{d}u\right)^{\frac12}\Bigg]\mathrm{d}s\\
&\le C\delta^{\frac{1+\beta}{2}}\int_0^t\left[\delta^{\frac12}\left(\int_0^{(s-T\delta)\vee 0}(s-u-T\delta)^{-\beta+1}\,\mathrm{d}u\right)^{\frac12}+\left(\int_{(s-T\delta)\vee 0}^s\int_{(s-2T\delta)\vee 0}^s(s-r)^{-\beta+1}\,\mathrm{d}r\,\mathrm{d}u\right)^{\frac12}\right]\mathrm{d}s\\
&\quad\times\sup_{s\in[0,T]}\left\|DF\left(X^N_s\right)A^{-\frac{\beta-1}{2}}\right\|_{L^{2p}(\Omega;\mathscr{L}(H))}\left\|A^{\frac{\beta-1}{2}}Q^{\frac12}\right\|_{\mathscr{L}_2(H)}.
\end{aligned}\tag{5.5}$$

It can be calculated that

$$\int_0^t \left[\delta^{\frac{1}{2}} \left(\int_0^{(s-T\delta)\vee 0} (s-u-T\delta)^{-\beta+1} \,\mathrm{d}u \right)^{\frac{1}{2}} + \left(\int_{(s-T\delta)\vee 0}^{s} \int_{(s-2T\delta)\vee 0}^{s} (s-r)^{-\beta+1} \,\mathrm{d}r \,\mathrm{d}u \right)^{\frac{1}{2}} \right] \mathrm{d}s$$

$$\leq \int_{T\delta}^{t} \delta^{\frac{1}{2}} \left(\int_0^{s-T\delta} (s-u-T\delta)^{-\beta+1} \,\mathrm{d}u \right)^{\frac{1}{2}} + \int_0^{T\delta} \left(\int_0^s \int_0^s (s-r)^{-\beta+1} \,\mathrm{d}s \,\mathrm{d}u \right)^{\frac{1}{2}} \mathrm{d}s$$

$$+ \int_{T\delta}^{2T\delta} \left(\int_{s-T\delta}^{s} \int_0^s (s-r)^{-\beta+1} \,\mathrm{d}r \,\mathrm{d}u \right)^{\frac{1}{2}} \mathrm{d}s + \int_{2T\delta}^{t} \left(\int_{s-T\delta}^{s} \int_{s-2T\delta}^{s} (s-r)^{-\beta+1} \,\mathrm{d}r \,\mathrm{d}u \right)^{\frac{1}{2}} \mathrm{d}s$$

$$\leq C \left(\delta^{\frac{1}{2}} + \delta^{\frac{5-\beta}{2}} + \delta^{\frac{3-\beta}{2}} \right) \leq C\delta^{\frac{1}{2}},$$

which leads to $\|II_3(t)\|_{L^p(\Omega)} \leq C\delta^{1+\frac{\beta}{2}},\ t \in [0,T]$. The term II_4 is treated separately for cases $d = 1,2$ and $d = 3$. When $d = 1,2$, by the Sobolev embedding $L^1(\mathscr{O}) \hookrightarrow \dot{H}^{-\frac{d+\epsilon_0}{2}}$ with $\epsilon_0 = \frac{2-\beta}{2}$ ($\beta \in (1,2)$), we deduce from (2.1)–(2.2) and (2.8) that

$$\begin{aligned}
\|II_4(t)\|_{L^p(\Omega)} &\leq \int_0^t \int_0^s \left\| A^{\frac{\beta-1}{2}+\frac{d+\epsilon_0}{4}} S^N(s-r) \right\|_{\mathscr{L}(H)} \left\| A^{-\frac{d+\epsilon_0}{4}} R_F\left(X^N_{t_{m_r}}, X^N_r\right) \right\|_{L^{2p}(\Omega;H)} \mathrm{d}r \\
&\quad \times \left\| A^{-\frac{\beta}{2}} \left(S^N\left(s-t_{m_s}\right) - \mathrm{Id} \right) A^{\frac{1}{2}} F^N\left(X^N_{t_{m_s}}\right) \right\|_{L^{2p}(\Omega;H)} \mathrm{d}s \\
&\leq C\delta^{\frac{\beta}{2}} \int_0^t \int_0^s (s-r)^{-\frac{\beta-1}{2}-\frac{d+\epsilon_0}{4}} \left\| R_F\left(X^N_{t_{m_r}}, X^N_r\right) \right\|_{L^{2p}(\Omega;L^1(\mathscr{O}))} \mathrm{d}r \left\| F\left(X^N_{t_{m_s}}\right) \right\|_{L^{2p}(\Omega;\dot{H}^1)} \mathrm{d}s \\
&\leq C\delta^{\frac{\beta}{2}} \int_0^t \int_0^s (s-r)^{-\frac{\beta-1}{2}-\frac{d+\epsilon_0}{4}} \left(1 + \sup_{0\leq r\leq T} \left\| X^N_r \right\|_{L^{4p}(\Omega;E)} \right) \left\| X^N_r - X^N_{t_{m_r}} \right\|^2_{L^{8p}(\Omega;H)} \mathrm{d}r \\
&\quad \times \left\| F\left(X^N_{t_{m_s}}\right) \right\|_{L^{2p}(\Omega;\dot{H}^1)} \mathrm{d}s \leq C\delta^{1+\frac{\beta}{2}},
\end{aligned}$$

where in the last step we use the regularity and the Hölder continuity of X^N_t (see Propositions 4.3 and 4.4). And when $d = 3$, applying (2.1), (2.8), the Gagliardo–Nirenberg inequality $\|u\|_{L^4(\mathscr{O})} \leq C\|\nabla u\|^{\frac{3}{4}} \|u\|^{\frac{1}{4}},\ u \in H^1$, Propositions 4.3 and 4.4 yields

$$\begin{aligned}
&\|II_4(t)\|_{L^p(\Omega)} \\
&= \left\| \int_0^t \left\langle \int_0^s A^{\frac{1}{2}} S^N(s-r) R_F\left(X^N_{t_{m_r}}, X^N_r\right) \mathrm{d}r,\ A^{-1}\left(S^N\left(s-t_{m_s}\right) - \mathrm{Id} \right) A^{\frac{1}{2}} F^N\left(X^N_{t_{m_s}}\right) \right\rangle \mathrm{d}s \right\|_{L^p(\Omega)} \\
&\leq C \int_0^t \int_0^s (s-r)^{-\frac{1}{2}} \left(1 + \sup_{0\leq r\leq T} \|X^N_r\|_{L^{4p}(\Omega;E)} \right) \left\| X^N_r - X^N_{t_{m_r}} \right\|^2_{L^{8p}(\Omega;L^4(\mathscr{O}))} \mathrm{d}r \left\| \left(s-t_{m_s}\right) \left\| F\left(X^N_{t_{m_s}}\right) \right\|_1 \right\|_{L^{2p}(\Omega)} \mathrm{d}s \\
&\leq C\delta \int_0^t \int_0^s (s-r)^{-\frac{1}{2}} \left(1 + \sup_{0\leq r\leq T} \left\| X^N_r \right\|_{L^{4p}(\Omega;E)} \right) \left\| \left\| X^N_r - X^N_{t_{m_r}} \right\|^{\frac{1}{4}} \left\| X^N_r - X^N_{t_{m_r}} \right\|_1^{\frac{3}{4}} \right\|^2_{L^{8p}(\Omega)} \mathrm{d}r \,\mathrm{d}s \\
&\leq C\delta \int_0^t \int_0^s (s-r)^{-\frac{1}{2}} \left\| X^N_r - X^N_{t_{m_r}} \right\|^{\frac{1}{2}}_{L^{4p}(\Omega;H)} \left\| X^N_r - X^N_{t_{m_r}} \right\|^{\frac{3}{2}}_{L^{12p}(\Omega;\dot{H}^1)} \mathrm{d}r \,\mathrm{d}s \\
&\leq C\delta \int_0^t \int_0^s (s-r)^{-\frac{1}{2}} \left(r - t_{m_r} \right)^{\frac{1}{4}+\frac{3(\beta-1)}{4}} \mathrm{d}r \,\mathrm{d}s \leq C\delta^{\beta},
\end{aligned}$$

where in the last step we have used the fact that $\frac{1}{4} + \frac{3(\beta-1)}{4} > \beta - 1$ for $\beta \in (1,2)$.

Hence, we deduce that for $\beta \in (0,1) \cup (1,2)$,

$$\left\| J_{1,1}^2(t) \right\|_{L^p(\Omega)} \le C\delta^{\beta}, \quad t \in [0,T].$$

For the term $J_{1,1}^3$, when $\beta \in (0,2)$,

$$\begin{aligned}
J_{1,1}^3(t) &= \int_0^t \left\langle \int_0^s A^{\frac{\beta}{2}} S^N(s-r) \left(F^N\left(X_r^N\right) - F^N(X(r))\right) \mathrm{d}r,\ A^{-\frac{\beta}{2}} \left(S^N\left(s - t_{m_s}\right) - \mathrm{Id}\right) F^N\left(X_{t_{m_s}}^N\right)\right\rangle \mathrm{d}s \\
&\le C \int_0^t \int_0^s (s-r)^{-\frac{\beta}{2}} \left[\left\| X(r) - X_r^N \right\|^2 + \left(s - t_{m_s}\right)^{\beta} \left(1 + \|X(r)\|_E^4 + \left\| X_r^N \right\|_E^4 \right) \left\| F\left(X_{t_{m_s}}^N\right)\right\|^2 \right] \mathrm{d}r\, \mathrm{d}s \\
&\le C \int_0^t \left\| X_s^N - Y_s^N \right\|^2 \mathrm{d}s + C \int_0^t \left\| Y_s^N - X(s) \right\|^2 \mathrm{d}s \\
&\quad + C \int_0^t \int_0^s (s-r)^{-\frac{\beta}{2}} \left(s - t_{m_s}\right)^{\beta} \left(1 + \|X(r)\|_E^4 + \left\| X_r^N \right\|_E^4 \right) \left\| F\left(X_{t_{m_s}}^N\right)\right\|^2 \mathrm{d}r\, \mathrm{d}s,
\end{aligned}$$

where in the last step we transform the integral domain and use the Minkowski inequality. This, together with (5.1)–(5.2), implies that for $\beta \in (0,2)$,

$$\left\| J_{1,1}^3(t) \right\|_{L^p(\Omega)} \le C \int_0^t \left\| X_s^N - Y_s^N \right\|_{L^{2p}(\Omega;H)}^2 \mathrm{d}s + C\lambda_N^{-\beta} + C\delta^{\beta}.$$

Altogether, we obtain that for $\beta \in (0,2]$,

$$\|J_{1,1}(t)\|_{L^p(\Omega)} \le C \int_0^t \left\| X_s^N - Y_s^N \right\|_{L^{2p}(\Omega;H)}^2 \mathrm{d}s + C\lambda_N^{-\beta} + C\delta^{\beta}, \quad t \in [0,T]. \tag{5.6}$$

Estimate of $J_{1,2}$. When $\beta \in (0,1]$, the Hölder continuity of X_t^N and Assumption 3.3 give

$$\begin{aligned}
\|J_{1,2}(t)\|_{L^p(\Omega)} &\le \frac{1}{2} \int_0^t \left\| \left\| X_s^N - Y_s^N \right\|^2 \right\|_{L^p(\Omega)} \mathrm{d}s + C \int_0^t \left\| X_s^N - X_{t_{m_s}}^N \right\|_{L^{4p}(\Omega;H)}^2 \left(1 + \sup_{s\in[0,T]} \left\| X_s^N \right\|_{L^{8p}(\Omega;E)}^4 \right) \mathrm{d}s \\
&\le \frac{1}{2} \int_0^t \left\| X_s^N - Y_s^N \right\|_{L^{2p}(\Omega;H)}^2 \mathrm{d}s + C\delta^{\beta}.
\end{aligned}$$

When $\beta \in (1,2]$, by (5.4), we split $J_{1,2}$ further as $J_{1,2} = J_{1,2}^1 + J_{1,2}^2 + J_{1,2}^3$ with

$$\begin{aligned}
J_{1,2}^1(t) &:= \int_0^t \left\langle \int_0^s S^N(s-r) \left(S^N\left(r - t_{m_r}\right) - \mathrm{Id}\right) F^N\left(X_{t_{m_r}}^N\right) \mathrm{d}r,\ F^N\left(X_{t_{m_s}}^N\right) - F^N\left(X_s^N\right)\right\rangle \mathrm{d}s, \\
J_{1,2}^2(t) &:= \int_0^t \left\langle \int_0^s S^N(s-r) \left(F^N\left(X_{t_{m_r}}^N\right) - F^N\left(X_r^N\right)\right) \mathrm{d}r,\ F^N\left(X_{t_{m_s}}^N\right) - F^N\left(X_s^N\right)\right\rangle \mathrm{d}s, \\
J_{1,2}^3(t) &:= \int_0^t \left\langle \int_0^s S^N(s-r) \left(F^N\left(X_r^N\right) - F^N(X(r))\right) \mathrm{d}r,\ F^N\left(X_{t_{m_s}}^N\right) - F^N\left(X_s^N\right)\right\rangle \mathrm{d}s.
\end{aligned}$$

The estimate of the term $J^1_{1,2}$ is similar to that of $J^2_{1,1}$, whose proof is based on the Taylor formula to F. And one can obtain $\|J^1_{1,2}(t)\|_{L^p(\Omega)} \le C\delta^\beta$. Now we show the estimate of the term $J^2_{1,2} = \sum_{i=1}^3 J^{2,i}_{1,2}$, where

$$
\begin{aligned}
J^{2,1}_{1,2}(t) &:= \int_0^t \left\langle \int_0^s S^N(s-r)\mathscr{A}_r \,\mathrm{d}r,\ \mathscr{A}_s \right\rangle \mathrm{d}s,\\
J^{2,2}_{1,2}(t) &:= \int_0^t \left\langle \int_0^s S^N(s-r)\mathscr{A}_r \,\mathrm{d}r,\ DF\left(X^N_{t_{m_s}}\right)\int_{t_{m_s}}^s S^N\left(s-t_{m_s}\right)\,\mathrm{d}W(u) \right\rangle \mathrm{d}s\\
&\quad + \int_0^t \left\langle \int_0^s S^N(s-r)DF\left(X^N_{t_{m_r}}\right)\int_{t_{m_r}}^r S^N\left(r-t_{m_r}\right)\,\mathrm{d}W(u)\,\mathrm{d}r,\ \mathscr{A}_s \right\rangle \mathrm{d}s,\\
J^{2,3}_{1,2}(t) &:= \int_0^t \left\langle \int_0^s S^N(s-r)DF\left(X^N_{t_{m_r}}\right)\int_{t_{m_r}}^r S^N\left(r-t_{m_r}\right)\,\mathrm{d}W(u_2)\,\mathrm{d}r,\ DF\left(X^N_{t_{m_s}}\right)\int_{t_{m_s}}^s S^N\left(s-t_{m_s}\right)\,\mathrm{d}W(u_1) \right\rangle \mathrm{d}s
\end{aligned}
$$

with $\mathscr{A}_r := DF\left(X^N_{t_{m_r}}\right)\left(\left(S^N\left(r-t_{m_r}\right)-\mathrm{Id}\right)X^N_{t_{m_r}} + \int_{t_{m_r}}^r S^N\left(r-t_{m_r}\right)F^N\left(X^N_{t_{m_r}}\right)\mathrm{d}u\right) + R_F\left(X^N_{t_{m_r}}, X^N_r\right)$.

For the term $J^{2,1}_{1,2}$, noting that for $d=1$,

$$
\begin{aligned}
&\left\| \int_0^t \left\langle \int_0^s S^N(s-r)R_F\left(X^N_{t_{m_r}}, X^N_r\right)\,\mathrm{d}r, R_F(X^N_{t_{m_s}}, X^N_s) \right\rangle \mathrm{d}s \right\|_{L^p(\Omega)}\\
&\le C\int_0^t\int_0^s (s-r)^{-\frac{d+\epsilon_0}{2}} \left\| \left\|R_F\left(X^N_{t_{m_r}}, X^N_r\right)\right\|_{-\frac{d+\epsilon_0}{2}} \left\|R_F\left(X^N_{t_{m_s}}, X^N_s\right)\right\|_{-\frac{d+\epsilon_0}{2}} \right\|_{L^p(\Omega)} \mathrm{d}r\,\mathrm{d}s \le C\delta^2,
\end{aligned}
$$

and for $d=2,3$,

$$
\begin{aligned}
&\left\| \int_0^t \left\langle \int_0^s S^N(s-r)R_F\left(X^N_{t_{m_r}}, X^N_r\right)\,\mathrm{d}r, R_F\left(X^N_{t_{m_s}}, X^N_s\right) \right\rangle \mathrm{d}s \right\|_{L^p(\Omega)}\\
&\le C\int_0^t\int_0^s (s-r)^{-\frac{d+\epsilon_0}{4}} \left\| \left\|R_F\left(X^N_{t_{m_r}}, X^N_r\right)\right\|_{-\frac{d+\epsilon_0}{2}} \left\|R_F\left(X^N_{t_{m_s}}, X^N_s\right)\right\| \right\|_{L^p(\Omega)} \mathrm{d}r\,\mathrm{d}s\\
&\le C\int_0^t\int_0^s (s-r)^{-\frac{d+\epsilon_0}{4}} \left\|R_F\left(X^N_{t_{m_r}}, X^N_r\right)\right\|_{L^{2p}(\Omega;\dot H^{-\frac{d+\epsilon_0}{2}})} \sup_{s\in[0,T]} \|X^N_s\|_{L^{4p}(\Omega;E)} \left\|X^N_s - X^N_{t_{m_s}}\right\|^2_{L^{8p}(\Omega;L^4(\mathscr{O}))} \mathrm{d}r\,\mathrm{d}s\\
&\le C\delta\int_0^t\int_0^s (s-r)^{-\frac{d+\epsilon_0}{4}} \left\| \left\|X^N_s - X^N_{t_{m_s}}\right\|_1^{\frac{d}{4}} \left\|X^N_s - X^N_{t_{m_s}}\right\|^{1-\frac{d}{4}} \right\|^2_{L^{8p}(\Omega)} \mathrm{d}r\,\mathrm{d}s \le C\delta^\beta,
\end{aligned}
$$

we have

$$
\begin{aligned}
\left\|J^{2,1}_{1,2}(t)\right\|_{L^p(\Omega)} &\le \left\| \int_0^t \left\langle \int_0^s S^N(s-r)R_F\left(X^N_{t_{m_r}}, X^N_r\right)\,\mathrm{d}r, R_F\left(X^N_{t_{m_s}}, X^N_s\right) \right\rangle \mathrm{d}s \right\|_{L^p(\Omega)}\\
&\quad + C\Bigg[\delta^\beta \sup_{r\in[0,T]} \left\| \left\|DF\left(X^N_{t_{m_r}}\right)\right\|_{\mathscr{L}(H)} \left\|X^N_{t_{m_r}}\right\|_\beta \right\|^2_{L^{2p}(\Omega)}\\
&\quad + \delta^2 \sup_{r\in[0,T]} \left\| \left\|DF\left(X^N_{t_{m_r}}\right)\right\|_{\mathscr{L}(H)} \left\|F\left(X^N_{t_{m_r}}\right)\right\| \right\|^2_{L^{2p}(\Omega)}
\end{aligned}
$$

$$+\,\delta^{\frac{\beta}{2}}\int_0^t\int_0^s (s-r)^{-\frac{d+\epsilon_0}{4}}\left\|\left\|R_F\left(X^N_{t_{m_r}},X^N_r\right)\right\|_{-\frac{d+\epsilon_0}{2}}\left\|DF\left(X^N_{t_{m_s}}\right)\right\|_{\mathscr{L}(H)}\right.$$
$$\left.\left(\left\|X^N_{t_{m_s}}\right\|_\beta+\left\|F\left(X^N_{t_{m_s}}\right)\right\|\right)\right\|_{L^p(\Omega)}\mathrm{d}r\,\mathrm{d}s\Bigg]\le C\delta^{\beta}.$$

For the term $J^{2,2}_{1,2}$, applying the stochastic Fubini theorem yields

$$\left\|J^{2,2}_{1,2}(t)\right\|_{L^p(\Omega)}$$
$$\le\left\|\sum_{i=0}^{m_t}\int_{t_i}^{t_{i+1}}\chi_{[t_i,s)}(u)\chi_{[t_i,t)}(s)\left\langle\int_0^s S^N(s-r)\mathscr{A}_r\,\mathrm{d}r,\ DF\left(X^N_{t_i}\right)S^N(s-t_i)\,\mathrm{d}s\,\mathrm{d}W(u)\right\rangle\right\|_{L^p(\Omega)}$$
$$+\int_0^t\left\|\sum_{i=0}^{m_s}\int_{t_i}^{t_{i+1}}\int_{t_i}^{t_{i+1}}\chi_{[t_i,r)}(u)\chi_{[t_i,s)}(r)\left\langle\mathscr{A}_s,\ S^N(s-r)DF\left(X^N_{t_i}\right)S^N(r-t_i)\,\mathrm{d}r\,\mathrm{d}W(u)\right\rangle\right\|_{L^p(\Omega)}\mathrm{d}s.$$

And by a similar proof to that of (5.5), one can show that $\|J^{2,2}_{1,2}(t)\|_{L^p(\Omega)}\le C\delta^{\frac{\beta}{2}+1}$. For the term $J^{2,3}_{1,2}$, by using the stochastic Fubini theorem twice and letting $G(s):=\int_0^s S^N(s-r)DF\left(X^N_{t_{m_r}}\right)\int_{t_{m_r}}^r S^N\left(r-t_{m_r}\right)\mathrm{d}W(u_2)\,\mathrm{d}r$, we derive

$$\left\|J^{2,3}_{1,2}(t)\right\|_{L^p(\Omega)}$$
$$=\left\|\sum_{i=0}^{m_t}\int_{t_i}^{t_{i+1}}\int_{t_i}^{t_{i+1}}\left\langle G(s),\ DF\left(X^N_{t_i}\right)\chi_{[t_i,s)}(u_1)\chi_{[t_i,t)}(s)S^N(s-t_i)\,\mathrm{d}s\,\mathrm{d}W(u_1)\right\rangle\right\|_{L^p(\Omega)}$$
$$\le C\left(\int_0^t\left\|\int_{t_{m_{u_1}}}^{t_{m_{u_1}}+\tau^\delta_{m_{u_1}}}Q^{\frac12}S^N\left(s-t_{m_s}\right)\chi_{[t_{m_{u_1}},s)}(u_1)\chi_{[t_{m_{u_1}},t)}(s)DF\left(X^N_{t_{m_s}}\right)G(s)\,\mathrm{d}s\right\|^2_{L^p(\Omega;H)}\mathrm{d}u_1\right)^{\frac12}$$
$$\le C\delta^{\frac12}\left(\int_0^t\int_{(u_1-T\delta)\vee 0}^{u_1+T\delta}\sup_{s\in[0,T]}\left\|Q^{\frac12}S^N\left(s-t_{m_s}\right)\chi_{[t_{m_{u_1}},s)}(u_1)\chi_{[t_{m_{u_1}},t)}(s)DF\left(X^N_{t_{m_s}}\right)G(s)\right\|^2_{L^p(\Omega;H)}\mathrm{d}s\,\mathrm{d}u_1\right)^{\frac12}$$
$$\le C\delta\sup_{s\in[0,T]}\left\|\sum_{j=0}^{m_s}\int_{t_j}^{t_{j+1}}\int_{t_j}^{t_{j+1}}Q^{\frac12}S^N\left(s-t_{m_s}\right)DF\left(X^N_{t_{m_s}}\right)S^N(s-r)DF\left(X^N_{t_j}\right)\chi_{[t_j,r)}(u_2)\chi_{[t_j,s)}(r)\right.$$
$$\left.\times S^N(r-t_j)\,\mathrm{d}r\,\mathrm{d}W(u_2)\right\|_{L^p(\Omega;H)}$$
$$\le C\delta\sup_{s\in[0,T]}\left(\int_0^s\left\|\int_{t_{m_{u_2}}}^{t_{m_{u_2}}+\tau^\delta_{m_{u_2}}}Q^{\frac12}S^N\left(s-t_{m_s}\right)DF\left(X^N_{t_{m_s}}\right)S^N(s-r)DF(X^N_{t_{m_{u_2}}})\chi_{[t_{m_{u_2}},r)}(u_2)\chi_{[t_{m_{u_2}},s)}(r)\right.\right.$$
$$\left.\left.\times S^N\left(r-t_{m_{u_2}}\right)Q^{\frac12}\,\mathrm{d}r\right\|^2_{L^p(\Omega;\mathscr{L}_2(H))}\mathrm{d}u_2\right)^{\frac12}\le C\delta^2.$$

For the term $J_{1,2}^3$, we split it as $J_{1,2}^3 = J_{1,2}^{3,1} + J_{1,2}^{3,2}$ with

$$J_{1,2}^{3,1}(t) := \int_0^t \left\langle \int_0^s S^N(s-r)\Big(F^N\left(X_r^N\right) - F^N(X(r))\Big)\,\mathrm{d}r,\ \mathscr{A}_s \right\rangle \mathrm{d}s,$$

$$J_{1,2}^{3,2}(t) := \int_0^t \left\langle \int_0^s S^N(s-r)\Big(F^N\left(X_r^N\right) - F^N(X(r))\Big)\,\mathrm{d}r,\ DF\left(X_{t_{m_s}}^N\right)\int_{t_{m_s}}^s S^N\left(s-t_{m_s}\right)\,\mathrm{d}W(u) \right\rangle \mathrm{d}s.$$

Combining (5.1)–(5.2), the term $J_{1,2}^{3,1}$ can be estimated as

$$\begin{aligned}
\left\| J_{1,2}^{3,1}(t) \right\|_{L^p(\Omega)} &\le C\int_0^t\int_0^s (s-r)^{-\frac{d+\epsilon_0}{4}} \Bigg[\left\|X_r^N - X(r)\right\|_{L^{2p}(\Omega;H)}^2 + \left(1 + \left\|X_r^N\right\|_{L^{8p}(\Omega;E)}^4 + \|X(r)\|_{L^{8p}(\Omega;E)}^4\right) \\
&\quad \times \left(\delta^\beta \left\| \left\|DF\left(X_{t_{m_s}}^N\right)\right\|_{\mathscr{L}(H)}\left(\left\|X_{t_{m_s}}^N\right\|_\beta + \left\|F\left(X_{t_{m_s}}^N\right)\right\|\right)\right\|_{L^{4p}(\Omega)}^2 + \left\| \left\|R_F\left(X_{t_{m_s}}^N, X_s^N\right)\right\|_{-\frac{d+\epsilon_0}{2}}\right\|_{L^{4p}(\Omega)}^2\right)\Bigg]\,\mathrm{d}r\,\mathrm{d}s \\
&\le C\int_0^t \left\|X_s^N - Y_s^N\right\|_{L^{2p}(\Omega;H)}^2 \mathrm{d}s + C\delta^\beta + C\lambda_N^{-\beta}.
\end{aligned}$$

By the stochastic Fubini theorem and the Young inequality, the term $J_{1,2}^{3,2}$ can be estimated as

$$\begin{aligned}
\left\|J_{1,2}^{3,2}(t)\right\|_{L^p(\Omega)} &= \Bigg\| \sum_{i=0}^{m_t} \int_{t_i}^{t_{i+1}} \int_{t_i}^{t_{i+1}} \Big\langle \int_0^s S^N(s-r)(F^N(X_r^N) - F^N(X(r)))\,\mathrm{d}r, \\
&\qquad \times DF\left(X_{t_i}^N\right) S^N(s-t_i)\chi_{[t_i,s)}(u)\chi_{[t_i,t)}(s)\,\mathrm{d}s\,\mathrm{d}W(u)\Big\rangle \Bigg\|_{L^p(\Omega)} \\
&\le C\Bigg(\int_0^t \Bigg\| \int_{t_{m_u}}^{t_{m_u}+\tau_{m_u}^\delta} Q^{\frac12} S^N\left(t-t_{m_u}\right)\chi_{[t_{m_u},s)}(u)\chi_{[t_{m_u},t)}(s) DF\left(X_{t_{m_u}}^N\right) \\
&\qquad \int_0^s S^N(s-r)\left(F^N\left(X_r^N\right) - F^N(X(r))\right)\mathrm{d}r\,\mathrm{d}s\Bigg\|_{L^p(\Omega;H)}^2 \mathrm{d}u\Bigg)^{\frac12} \\
&\le C\Bigg(\int_0^t \Bigg\| \int_{t_{m_u}}^{t_{m_u}+\tau_{m_u}^\delta} \left\|Q^{\frac12} S^N\left(t-t_{m_u}\right)\chi_{[t_{m_u},s)}(u)\chi_{[t_{m_u},t)}(s) DF\left(X_{t_{m_u}}^N\right)\right\|_{\mathscr{L}(H)} \\
&\qquad \times \int_0^t \left(1 + \left\|X_r^N\right\|_E^2 + \|X(r)\|_E^2\right)\left\|X_r^N - X(r)\right\|\,\mathrm{d}r\,\mathrm{d}s\Bigg\|_{L^p(\Omega)}^2 \mathrm{d}u\Bigg)^{\frac12} \\
&\le C\Bigg(\int_0^t \Bigg[\int_0^t \Bigg\| \int_{t_{m_u}}^{t_{m_u}+\tau_{m_u}^\delta} \left\|Q^{\frac12} S^N\left(t-t_{m_u}\right)\chi_{[t_{m_u},s)}(u)\chi_{[t_{m_u},t)}(s) DF\left(X_{t_{m_u}}^N\right)\right\|_{\mathscr{L}(H)} \\
&\qquad \times \left(1 + \left\|X_r^N\right\|_E^2 + \|X(r)\|_E^2\right)\mathrm{d}s\Bigg\|_{L^{2p}(\Omega)} \left\|X_r^N - X(r)\right\|_{L^{2p}(\Omega;H)}\mathrm{d}r\Bigg]^2 \mathrm{d}u\Bigg)^{\frac12} \\
&\le C\int_0^t \left\|X_r^N - X(r)\right\|_{L^{2p}(\Omega;H)}^2 \mathrm{d}r + C\Bigg(\int_0^t \Bigg[\int_0^t \Bigg\| \int_{t_{m_u}}^{t_{m_u}+\tau_{m_u}^\delta} \chi_{[t_{m_u},s)}(u)\chi_{[t_{m_u},t)}(s) \\
&\qquad \times \left\|Q^{\frac12} S^N\left(t-t_{m_u}\right) DF\left(X_{t_{m_u}}^N\right)\right\|_{\mathscr{L}(H)}\left(1 + \left\|X_r^N\right\|_E^2 + \|X(r)\|_E^2\right)\mathrm{d}s\Bigg\|_{L^{2p}(\Omega)}^2 \mathrm{d}r\Bigg]^2 \mathrm{d}u\Bigg)^{\frac12} \\
&\le C\int_0^t \left\|X_r^N - Y_r^N\right\|_{L^{2p}(\Omega;H)}^2 \mathrm{d}r + C\delta^\beta + C\lambda_N^{-\beta},
\end{aligned}$$

where in the third step we exchange the order of the integral and use the Hölder inequality, and in the last step we use (5.1)–(5.2).

Hence, we derive that for $\beta \in (0, 2]$,

$$\|J_{1,2}(t)\|_{L^p(\Omega)} \leq C \int_0^t \left\| X_r^N - Y_r^N \right\|^2_{L^{2p}(\Omega;H)} \mathrm{d}r + C\delta^\beta + C\lambda_N^{-\beta}, \quad t \in [0, T]. \tag{5.7}$$

Therefore, combining estimates of terms $J_{1,1}$ and $J_{1,2}$ (i.e., (5.6) and (5.7)), we derive

$$\left\| X_t^N - Y_t^N \right\|^2_{L^{2p}(\Omega;H)} \leq C \left(\int_0^t \left\| X_s^N - Y_s^N \right\|^2_{L^{2p}(\Omega;H)} \mathrm{d}s + \lambda_N^{-\beta} + \delta^\beta \right),$$

where $p \geq 1$, and the constant $C > 0$ depends on $p, T, \tau_{min}, L_0, L_1, L_2, X_0$ and $\|A^{\frac{\beta-1}{2}} Q^{\frac{1}{2}}\|_{\mathscr{L}_2(H)}$. Applying the Grönwall inequality, we obtain

$$\sup_{0 \leq t \leq T} \left\| X_t^N - Y_t^N \right\|_{L^{2p}(\Omega;H)} \leq C \left(\lambda_N^{-\frac{\beta}{2}} + \delta^{\frac{\beta}{2}} \right), \; p \geq 1,$$

which together with (5.1)–(5.2) finishes the proof of the scheme (AE). □

6. Discussions on the multiplicative noise case

For the stochastic Allen–Cahn equation driven by a multiplicative noise, there have been some works on the numerical study, see e.g., Majee & Prohl (2018); Jentzen & Pušnik (2020); Feng *et al.* (2021); Liu & Qiao (2021) and references therein. To be specific, for the stochastic Allen–Cahn equation driven by the Brownian motion, the authors in Feng *et al.* (2021) propose the fully discrete finite element method and prove the strong convergence with nearly optimal rates; the authors in Majee & Prohl (2018) give variational error analysis for the structure preserving finite element based space-time discretization of the strong variational solution. For the stochastic Allen–Cahn equation driven by the Q-Wiener process, Jentzen & Pušnik (2020) derives the strong convergence rate for the nonlinearity-truncated exponential Euler approximation scheme; Liu & Qiao (2021) proves strong convergence rates for both the drift-implicit Euler–Galerkin finite element scheme and the Milstein–Galerkin finite element scheme.

In this section, we present the numerical analysis of the adaptive time-stepping scheme for the multiplicative noise case, i.e., the stochastic Allen–Cahn equation with the multiplicative noise

$$\begin{cases} \mathrm{d}X(t) + AX(t)\,\mathrm{d}t = F(X(t))\,\mathrm{d}t + G(X(t))\,\mathrm{d}W(t), & t \in (0, T], \\ X(0) = X_0, \end{cases} \tag{6.1}$$

where A and F are defined as in (1.1). In this section, $\{W(t)\}_{t\in[0,T]}$ is a generalized Q-Wiener process valued on another separable Hilbert space U. Here, $G : H \to \mathscr{L}_2^0 := \mathscr{L}_2(U_0, H)$ is Lipschitz continuous with $U_0 := Q^{\frac{1}{2}}(U)$, i.e.,

$$\|G(X) - G(Y)\|_{\mathscr{L}_2^0} \leq C\|X - Y\|, \quad X, Y \in H$$

for some constant $C > 0$. The corresponding adaptive time-stepping scheme for (6.1) is

$$X_{t_{m+1}}^N = S^N(\tau_m)\left(X_{t_m}^N + F^N\left(X_{t_m}^N\right)\tau_m + P^N G\left(X_{t_m}^N\right)\Delta W_m\right), \quad X_0^N = P^N X_0, \qquad \text{(AE-M)}$$

where τ_m is defined similarly as in (AE). The differential form of the continuous version of (AE-M) is given by

$$\mathrm{d}X_t^N = -A^N X_t^N\,\mathrm{d}t + S^N\left(t - t_{m_t}\right)F^N\left(X_{t_{m_t}}^N\right)\mathrm{d}t + S^N\left(t - t_{m_t}\right)P^N G\left(X_{t_{m_t}}^N\right)\mathrm{d}W(t). \qquad (6.2)$$

We remark that in this section we still use notations $X(t), X_{t_m}^N$ and X_t^N in the multiplicative noise case for the simplicity of symbols. And we will emphasize on them when these symbols are referred to solutions in the additive noise case.

The following lemma shows the boundedness of the expected supremum of the pth moment for the solution of (AE-M).

LEMMA 6.1 Under Assumptions 2.1, 2.3 and (3.5)(3.6), we have $\mathbb{E}\big[\sup_{0\le t\le T}\left\|X_t^N\right\|^p\big] \le C$ for $p \ge 2$ with some constant $C := C(p, T, X_0) > 0$.

Proof. We only consider the case of $p \ge 4$ since the case of $2 \le p < 4$ can be obtained by the use of the Hölder inequality. The proof is based on the truncation technique. For $K > \left\|P^N X_0\right\|$, define the truncated function $\Theta_K \in \mathscr{C}_0^\infty : H \to [0,1]$ satisfying $\Theta_K(x) = 1$ for $\|x\| \le K$ and $\Theta_K(x) = 0$ for $\|x\| \ge 2K$. Introduce the truncated stochastic process of the numerical solution $\{X_{t_m}^N\}_{m\ge 0}$ as

$$X_{t_{m+1}}^{N,K} = S^N(\tau_m)\left(X_{t_m}^{N,K} + \Theta_K\left(X_{t_m}^{N,K}\right)F^N\left(X_{t_m}^{N,K}\right)\tau_m + P^N G\left(X_{t_m}^{N,K}\right)\Delta W_m\right).$$

The corresponding continuous version is defined as

$$X_t^{N,K} = S^N(t)P^N X_0 + \int_0^t S^N\left(t - t_{m_s}\right)\Theta_K\left(X_{t_{m_s}}^{N,K}\right)F^N\left(X_{t_{m_s}}^{N,K}\right)\mathrm{d}s + \int_0^t S^N\left(t - t_{m_s}\right)P^N G\left(X_{t_{m_s}}^{N,K}\right)\mathrm{d}W(s).$$

Then the proof is separated into two steps.

Step 1: Expected supremum of the pth moment of the truncated numerical solution.

By the contraction property of P^N and $S(t)$ in H, i.e., $\|P^N u\| \le \|u\|$, $\|S(t)u\| \le \|u\|$, and using (3.6), we have

$$\begin{aligned}
\left\|X_{t_{m+1}}^{N,K}\right\|^2 &\le \left\|X_{t_m}^{N,K} + \Theta_K\left(X_{t_m}^{N,K}\right)F^N\left(X_{t_m}^{N,K}\right)\tau_m + P^N G\left(X_{t_m}^{N,K}\right)\Delta W_m\right\|^2 \\
&= \left\|X_{t_m}^{N,K}\right\|^2 + 2\tau_m\Theta_K\left(X_{t_m}^{N,K}\right)\left(\left\langle X_{t_m}^{N,K}, F^N\left(X_{t_m}^{N,K}\right)\right\rangle + \frac{1}{2}\tau_m\Theta_K\left(X_{t_m}^{N,K}\right)\left\|F^N\left(X_{t_m}^{N,K}\right)\right\|^2\right) \\
&\quad + 2\left\langle X_{t_m}^{N,K} + \Theta_K\left(X_{t_m}^{N,K}\right)F^N\left(X_{t_m}^{N,K}\right)\tau_m,\, P^N G\left(X_{t_m}^{N,K}\right)\Delta W_m\right\rangle + \left\|P^N G\left(X_{t_m}^{N,K}\right)\Delta W_m\right\|^2 \\
&\le \left\|X_{t_m}^{N,K}\right\|^2 + 2\tau_m\left(\bar{L}_2\left\|X_{t_m}^{N,K}\right\|^2 + L_3\right) + 2\left\langle X_{t_m}^{N,K} + \Theta_K\left(X_{t_m}^{N,K}\right)F^N\left(X_{t_m}^{N,K}\right)\tau_m, G\left(X_{t_m}^{N,K}\right)\Delta W_m\right\rangle + \left\|G\left(X_{t_m}^{N,K}\right)\Delta W_m\right\|^2.
\end{aligned}$$

Similarly, for $t \in (t_m, t_{m+1})$,

$$\left\|X_t^{N,K}\right\|^2 \leq \left\|X_{t_m}^{N,K} + \Theta_K\left(X_{t_m}^{N,K}\right)F^N\left(X_{t_m}^{N,K}\right)(t-t_m) + P^N G\left(X_{t_m}^{N,K}\right)(W(t)-W(t_m))\right\|^2$$
$$\leq \left\|X_{t_m}^{N,K}\right\|^2 + 2(t-t_m)\left(\bar{L}_2\left\|X_{t_m}^{N,K}\right\|^2 + L_3\right) + 2\left\langle X_{t_m}^{N,K} + \Theta_K\left(X_{t_m}^{N,K}\right)F^N\left(X_{t_m}^{N,K}\right)(t-t_m), G\left(X_{t_m}^{N,K}\right)(W(t)-W(t_m))\right\rangle$$
$$+ \left\|G\left(X_{t_m}^{N,K}\right)(W(t)-W(t_m))\right\|^2.$$

By induction, we have

$$\left\|X_t^{N,K}\right\|^2 \leq \left\|P^N X_0\right\|^2 + 2\bar{L}_2\int_0^t \left\|X_{t_{m_s}}^{N,K}\right\|^2 \mathrm{d}s + 2L_3 t + 2\int_0^t \left\langle X_{t_{m_s}}^{N,K} + \Theta_K\left(X_{t_{m_s}}^{N,K}\right)F^N\left(X_{t_{m_s}}^{N,K}\right)\left(\left(t_{m_s}+\tau_{m_s}\right)\wedge t - t_{m_s}\right),\right.$$
$$\left. G\left(X_{t_{m_s}}^{N,K}\right)\mathrm{d}W(s)\right\rangle + \sum_{i=0}^{m_t-1}\left\|G\left(X_{t_i}^{N,K}\right)\Delta W_i\right\|^2 + \left\|G\left(X_{t_{m_t}}^{N,K}\right)\left(W(t) - W\left(t_{m_t}\right)\right)\right\|^2.$$

Taking the $\frac{p}{2}$th moment with $p \geq 4$ and applying the Burkholder–Davis–Gundy inequality yield

$$\mathbb{E}\left[\left\|X_t^{N,K}\right\|^p\right] \leq C(p,T,\bar{L}_2)\left(\mathbb{E}\left[\|X_0\|^p\right] + T^{\frac{p}{2}} + \int_0^t \mathbb{E}\left[\left\|X_{t_{m_s}}^{N,K}\right\|^p\right]\mathrm{d}s\right.$$
$$+ \mathbb{E}\left[\left(\int_0^t \left\|X_{t_{m_s}}^{N,K} + \Theta_K\left(X_{t_{m_s}}^{N,K}\right)F^N\left(X_{t_{m_s}}^{N,K}\right)\left(\left(t_{m_s}+\tau_{m_s}\right)\wedge t - t_{m_s}\right)\right\|^2 \left\|G\left(X_{t_{m_s}}^{N,K}\right)\right\|_{\mathscr{L}_2^0}^2 \mathrm{d}s\right)^{\frac{p}{4}}\right]$$
$$\left.+\mathbb{E}\left[\left(\sum_{i=0}^{m_t-1}\left\|G\left(X_{t_i}^{N,K}\right)\Delta W_i\right\|^2 + \left\|G\left(X_{t_{m_t}}^{N,K}\right)\left(W(t)-W(t_{m_t})\right)\right\|^2\right)^{\frac{p}{2}}\right]\right).$$

Notice that

$$\left\|X_{t_{m_s}}^{N,K} + \Theta_K\left(X_{t_{m_s}}^{N,K}\right)F^N\left(X_{t_{m_s}}^{N,K}\right)\left(\left(t_{m_s}+\tau_{m_s}\right)\wedge t - t_{m_s}\right)\right\|^2$$
$$= \left\|X_{t_{m_s}}^{N,K}\right\|^2 + 2\left(\left(t_{m_s}+\tau_{m_s}\right)\wedge t - t_{m_s}\right)\Theta_K\left(X_{t_{m_s}}^{N,K}\right)$$
$$\times\left(\left\langle X_{t_{m_s}}^{N,K}, F\left(X_{t_{m_s}}^{N,K}\right)\right\rangle + \frac{1}{2}\left(\left(t_{m_s}+\tau_{m_s}\right)\wedge t - t_{m_s}\right)\Theta_K\left(X_{t_{m_s}}^{N,K}\right)\left\|F\left(X_{t_{m_s}}^{N,K}\right)\right\|^2\right)$$
$$\leq \left\|X_{t_{m_s}}^{N,K}\right\|^2 + 2\left(\left(t_{m_s}+\tau_{m_s}\right)\wedge t - t_{m_s}\right)\left(\bar{L}_2\left\|X_{t_{m_s}}^{N,K}\right\|^2 + L_3\right)$$

and

$$
\begin{aligned}
&\mathbb{E}\left[\left(\sum_{i=0}^{m_t-1}\left\|G\left(X_{t_i}^{N,K}\right)\Delta W_i\right\|^2\right)^{\frac{p}{2}}\right]=\mathbb{E}\left[\left(\sum_{i=0}^{m_t-1}\tau_i\frac{\left\|G\left(X_{t_i}^{N,K}\right)\Delta W_i\right\|^2}{\tau_i}\right)^{\frac{p}{2}}\right]\\
&\le\mathbb{E}\left[\left(\sum_{i=0}^{m_t-1}\tau_i\right)^{\frac{p}{2}-1}\sum_{i=0}^{m_t-1}\tau_i\frac{\left\|G\left(X_{t_i}^{N,K}\right)\Delta W_i\right\|^p}{\tau_i^{\frac{p}{2}}}\right]\\
&\le T^{\frac{p}{2}-1}\mathbb{E}\left[\sum_{i=0}^{m_t-1}\tau_i\frac{\left\|G\left(X_{t_i}^{N,K}\right)\Delta W_i\right\|^p}{\tau_i^{\frac{p}{2}}}\right]\le T^{\frac{p}{2}-1}\mathbb{E}\left[\int_0^t\frac{\left\|G\left(X_{t_{m_s}}^{N,K}\right)\Delta W_{m_s}\right\|^p}{\tau_{m_s}^{\frac{p}{2}}}\,\mathrm{d}s\right]\\
&=T^{\frac{p}{2}-1}\int_0^t\mathbb{E}\left[\mathbb{E}\left[\tau_{m_s}^{-\frac{p}{2}}\left\|G\left(X_{t_{m_s}}^{N,K}\right)\Delta W_{m_s}\right\|^p|\mathscr{F}_{t_{m_s}}\right]\right]\mathrm{d}s=T^{\frac{p}{2}-1}\int_0^t\mathbb{E}\left[\tau_{m_s}^{-\frac{p}{2}}\mathbb{E}\left[\left\|G\left(X_{t_{m_s}}^{N,K}\right)\Delta W_{m_s}\right\|^p|\mathscr{F}_{t_{m_s}}\right]\right]\mathrm{d}s\\
&\le C\int_0^t\mathbb{E}\left[1+\left\|X_{t_{m_s}}^{N,K}\right\|^p\right]\mathrm{d}s,
\end{aligned}
$$

where we have used the inequality $\left(\sum_i\tau_iu_i\right)^p\le\left(\sum_i\tau_i\right)^{p-1}\sum_i\tau_iu_i^p$, the property of the conditional expectation, the Burkholder–Davis–Gundy inequality and the Lipschitz continuity of G. Therefore, combining the Hölder inequality, we get for $p\ge 4$,

$$
\begin{aligned}
\mathbb{E}\left[\left\|X_t^{N,K}\right\|^p\right]&\le C\left(p,T,\|Q\|_{\mathscr{L}_2(H)},\bar{L}_2\right)\left(1+\mathbb{E}\left[\|X_0\|^p\right]+\int_0^t\mathbb{E}\left[\left\|X_{t_{m_s}}^{N,K}\right\|^p\right]\mathrm{d}s\right)\\
&\le C\left(p,T,\|Q\|_{\mathscr{L}_2(H)},\bar{L}_2\right)\left(1+\mathbb{E}\left[\|X_0\|^p\right]+\int_0^t\sup_{0\le u\le s}\mathbb{E}\left[\left\|X_u^{N,K}\right\|^p\right]\mathrm{d}s\right),
\end{aligned}
$$

which implies

$$
\sup_{0\le r\le t}\mathbb{E}\left[\|X_r^{N,K}\|^p\right]\le C\left(p,T,\|Q\|_{\mathscr{L}_2(H)},\bar{L}_2\right)\left(1+\mathbb{E}\left[\|X_0\|^p\right]+\int_0^t\sup_{0\le u\le s}\mathbb{E}\left[\left\|X_u^{N,K}\right\|^p\right]\mathrm{d}s\right).
$$

Applying the Grönwall inequality leads to that for $p\ge 4$,

$$
\sup_{0\le t\le T}\mathbb{E}\left[\left\|X_t^{N,K}\right\|^p\right]\le C\left(p,T,\|Q\|_{\mathscr{L}_2(H)},\bar{L}_2\right)\left(1+\mathbb{E}\left[\|X_0\|^p\right]\right). \tag{6.3}
$$

To obtain the expected supremum of the pth moment of $X_t^{N,K}$, we need to apply the following Burkholder–Davis–Gundy inequality:

$$\mathbb{E}\left[\sup_{0\le t\le T}\left|\int_0^t \left\langle X_{t_{m_s}}^{N,K} + \Theta_K\left(X_{t_{m_s}}^{N,K}\right) F\left(X_{t_{m_s}}^{N,K}\right)\tau_{m_s}, G\left(X_{t_{m_s}}^{N,K}\right)\,\mathrm{d}W(s)\right\rangle\right|^p\right]$$
$$\le C\,\mathbb{E}\left[\left(\int_0^T \left\|X_{t_{m_s}}^{N,K} + \Theta_K\left(X_{t_{m_s}}^{N,K}\right) F^N\left(X_{t_{m_s}}^{N,K}\right)\tau_{m_s}\right\|^2 \left\|G\left(X_{t_{m_s}}^{N,K}\right)\right\|_{\mathscr{L}_2^0}^2\,\mathrm{d}s\right)^{\frac{p}{2}}\right]$$
$$\le C\int_0^T \sup_{s\in[0,T]}\mathbb{E}\left[\left\|X_{t_{m_s}}^{N,K}\right\|^p\right]\mathrm{d}s \le C,$$

where we use (6.3). And the remaining proof of $\mathbb{E}\left[\sup_{0\le t\le T}\left\|X_t^{N,K}\right\|^p\right] \le C(1+\mathbb{E}\left[\|X_0\|^p\right])$ can be given similarly. Moreover, the assumption (3.5) gives that $\mathbb{E}[M_T] \le T\mathbb{E}\left[\sup_{0\le t\le T}\tau(X_t^{N,K})^{-1}\right] \le T\mathbb{E}\left[\sup_{0\le t\le T}\left(\zeta_1\left\|X_t^{N,K}\right\|^{q_0}+\zeta_2\right)\right] \le C$ with C independent of K, which implies that T is a.s. attainable.

Step 2: Expected supremum of the pth moment of X_t^N.

Define the stopping time by $\tilde{\tau}_K = \inf\{t\in[0,T] : \left\|X_t^{N,K}\right\| \ge K\}$. Then for $t\in[0,\tilde{\tau}_{K_1}\wedge\tilde{\tau}_{K_2})$, we have $\Theta_{K_1}(X_t^{N,K_1}) = \Theta_{K_2}(X_t^{N,K_2}) = 1$ and thus $X_t^{N,K_1} = X_t^{N,K_2}$ a.s. due to the existence and uniqueness of the solution. Hence, $\tilde{\tau}_K$ is nondecreasing with respect to K. Denote $\lim_{K\to\infty}\tilde{\tau}_K = T_*$ a.s. The Chebyshev inequality gives

$$\mathbb{P}\{\tilde{\tau}_K = T\} = \mathbb{P}\left\{\sup_{0\le t\le T}\left\|X_t^{N,K}\right\| \le K\right\}$$
$$\ge 1 - K^{-4}\mathbb{E}\left[\sup_{0\le t\le T}\left\|X_t^{N,K}\right\|^4\right] \to 1 \quad \text{as } K\to\infty,$$

which implies $T = T_*$. Define X_t^N on $[0,T]$ by $X_t^N = X_t^{N,K}$ on $[0,\tilde{\tau}_K)$. Then X_t^N is the solution of (6.2) with $\lim_{K\to\infty}\sup_{t\in[0,\tilde{\tau}_K)}\left\|X_t^{N,K}\right\| = \sup_{t\in[0,T]}\left\|X_t^N\right\|$ a.s. Hence, applying the Fatou lemma and *Step 1* leads to

$$\mathbb{E}\left[\sup_{0\le t\le T}\left\|X_t^N\right\|^p\right] \le \lim_{K\to\infty}\mathbb{E}\left[\sup_{0\le t\le T}\left\|X_t^{N,K}\right\|^p\right] \le C\left(1+\mathbb{E}\left[\|X_0\|^p\right]\right).$$

The proof is finished. □

Below, we give the *a priori* estimate in E-norm of the solution of (AE-M) for the case of $d=1$ based on the Sobolev embedding $H^{\frac{1}{2}+} \hookrightarrow E$.

PROPOSITION 6.2 Let $d = 1$. Under Assumptions 2.1, 2.3, 3.3–3.4 and (3.4)–(3.6), we have for $p \geq 2$,

$$\sup_{N \in \mathbb{N}_+} \sup_{0 \leq t \leq T} \left\| X_t^N \right\|_{L^p(\Omega;E)} \leq C$$

for some $C := C(p, T, X_0) > 0$, where X_t^N is the solution of (AE-M) with timestep function τ^δ for $\delta \in (0, 1)$.

Proof. We split the numerical solution as $X_t^N = Y_{1,t}^N + Y_{2,t}^N + \mathbb{Z}_t^N$, where $Y_{1,t}^N, Y_{2,t}^N$ and $\mathbb{Z}_t^N$ are solutions of

$$\mathrm{d}Y_{1,t}^N = -A^N Y_{1,t}^N \,\mathrm{d}t + S^N\left(t - t_{m_t}\right) F^N\left(X_{t_{m_t}}^N\right) \mathrm{d}t - F^N\left(X_t^N\right) \mathrm{d}t, \quad Y_{1,0}^N = P^N X_0, \tag{6.4}$$

$$\mathrm{d}Y_{2,t}^N = F^N\left(X_t^N\right) \mathrm{d}t, \quad Y_{2,0}^N = 0,$$

and

$$\mathrm{d}\mathbb{Z}_t^N = -A^N \mathbb{Z}_t^N \,\mathrm{d}t + S^N\left(t - t_{m_t}\right) P^N G\left(X_{t_{m_t}}^N\right) \mathrm{d}W_t, \quad \mathbb{Z}_0^N = 0, \tag{6.5}$$

respectively.

Applying Lemma 4.2 to $Y_{2,t}^N$ yields that

$$\left\| Y_{2,t}^N \right\|_E \leq C\left(1 + \int_0^t (t-s)^{-\frac{d}{4}} \left\| Y_{1,s}^N + \mathbb{Z}_s^N \right\|_E^3 + \left\| Y_{1,s}^N + \mathbb{Z}_s^N \right\|_E^{72} \mathrm{d}s\right).$$

Recall the definition Z_t^N (see (A5)) in the proof of Proposition 4.3. It can be observed that $Y_{1,t}^N + \mathbb{Z}_t^N$ has a similar expression to that of Z_t^N, in which the counterpart of W_A^N is $\mathbb{Z}_t^N$ in this setting. Thus, it suffices to prove that $\sup_{0 \leq t \leq T} \mathbb{E}\big[\|\mathbb{Z}_t^N\|_E^q\big] < \infty$ for $q \geq 2$ in the multiplicative noise case. In fact, the Sobolev embedding $H^{\frac{1}{2}+\epsilon} \hookrightarrow E$ with $\epsilon > 0$, the Burkholder–Davis–Gundy inequality and the Lipschitz continuity of G give that

$$\begin{aligned}
\sup_{0 \leq t \leq T} \mathbb{E}\left[\left\| \mathbb{Z}_t^N \right\|_E^q\right] &\leq C \sup_{0 \leq t \leq T} \mathbb{E}\left[\left\| \int_0^t S^N\left(t - t_{m_s}\right) P^N G\left(X_{t_{m_s}}^N\right) \mathrm{d}W(s) \right\|_{\frac{1}{2}+\epsilon}^q\right] \\
&\leq C \sup_{0 \leq t \leq T} \mathbb{E}\left[\left(\int_0^t \left\| A^{\frac{1+2\epsilon}{4}} S^N\left(t - t_{m_s}\right) P^N G\left(X_{t_{m_s}}^N\right) \right\|_{\mathscr{L}_2^0}^2 \mathrm{d}s\right)^{\frac{q}{2}}\right] \\
&\leq C \sup_{0 \leq t \leq T} \mathbb{E}\left[\left(\int_0^t \left(t - t_{m_s}\right)^{-\frac{1+2\epsilon}{2}} \left\| P^N G\left(X_{t_{m_s}}^N\right) \right\|_{\mathscr{L}_2^0}^2 \mathrm{d}s\right)^{\frac{q}{2}}\right] \\
&\leq C \sup_{0 \leq t \leq T} \mathbb{E}\left[\left(\int_0^t (t-s)^{-\frac{1+2\epsilon}{2}} \left(1 + \left\| X_{t_{m_s}}^N \right\|^2\right) \mathrm{d}s\right)^{\frac{q}{2}}\right] \\
&\leq C \sup_{0 \leq t \leq T} \left(\int_0^t (t-s)^{-\frac{1+2\epsilon}{2}} \left(1 + \sup_{0 \leq s \leq T} \left\| X_{t_{m_s}}^N \right\|_{L^q(\Omega;H)}^2\right) \mathrm{d}s\right)^{\frac{q}{2}} \leq C.
\end{aligned}$$

The remaining proof is similar to that of Proposition 4.3 under conditions (3.4)–(3.5) and hence is omitted. □

With the above regularity estimates, one can derive the *a priori* estimate in $\dot{H}^1$-norm of the solution of (AE-M) by a similar approach to Proposition 4.4. Then similar to the proof of Theorem 3.5, one can obtain the following convergence order of the adaptive time-stepping scheme (AE-M).

PROPOSITION 6.3 Let $d = 1$. Under conditions in Proposition 6.2, for $p \geq 4$,

$$\sup_{t\in[0,T]} \left\| X(t) - X_t^N \right\|_{L^p(\Omega;H)} \leq C\left(\lambda_N^{-\frac{1}{2}} + \delta^{\frac{1}{2}}\right),$$

where X_t^N is the solution of (AE-M) and $C := C(p, T, X_0) > 0$.

Proof. Similar to the proof of Theorem 3.5, we introduce the auxiliary process $\mathbb{Y}_t^N$,

$$\mathbb{Y}_t^N = S^N(t)P^N X_0 + \int_0^t S^N(t-s)F^N(X(s))\,\mathrm{d}s + \int_0^t S^N\left(t - t_{m_s}\right) P^N G\left(\mathbb{Y}_s^N\right) \mathrm{d}W(s).$$

It can be shown that $\|\mathbb{Y}_t^N\|_{L^p(\Omega;H)} \leq C$ and $\|\mathbb{Y}_t^N - \mathbb{Y}_s^N\|_{L^p(\Omega;H)} \leq C(t-s)^{\frac{1}{2}}$. The error can be divided into

$$\left\| X(t) - X_t^N \right\|_{L^p(\Omega;H)} \leq \|X(t) - P^N X(t)\|_{L^p(\Omega;H)} + \left\| P^N X(t) - \mathbb{Y}_t^N \right\|_{L^p(\Omega;H)} + \left\| \mathbb{Y}_t^N - X_t^N \right\|_{L^p(\Omega;H)}.$$

With the regularity of the solution X, we have $\|X(t) - P^N X(t)\|_{L^p(\Omega;H)} \leq C\lambda_N^{-\frac{1}{2}} \|X(t)\|_{L^p(\Omega;\dot{H}^1)}$.

For the term $\|P^N X(t) - \mathbb{Y}_t^N\|_{L^p(\Omega;H)}$, it follows from the Burkholder–Davis–Gundy inequality that for $p \geq 2$,

$$\begin{aligned}
&\left\| P^N X(t) - \mathbb{Y}_t^N \right\|_{L^p(\Omega;H)} \\
&\leq \left\| \int_0^t S^N(t-s)\left(G(X(s)) - G\left(\mathbb{Y}_s^N\right)\right) \mathrm{d}W(s) \right\|_{L^p(\Omega;H)} \\
&\quad + \left\| \int_0^t S^N(t-s)\left(\mathrm{Id} - S^N\left(s - t_{m_s}\right)\right) G\left(\mathbb{Y}_s^N\right) \mathrm{d}W(s) \right\|_{L^p(\Omega;H)} \\
&\leq C\Bigg[\left\| \int_0^t \left\| G(X(s)) - G(P^N X(s)) \right\|^2_{\mathscr{L}_2^0} \mathrm{d}s \right\|^{\frac{1}{2}}_{L^{\frac{p}{2}}(\Omega)} + \left\| \int_0^t \left\| G(P^N X(s)) - G\left(\mathbb{Y}_s^N\right) \right\|^2_{\mathscr{L}_2^0} \mathrm{d}s \right\|^{\frac{1}{2}}_{L^{\frac{p}{2}}(\Omega)} \\
&\quad + \left\| \int_0^t \left\| S^N(t-s)\left(\mathrm{Id} - S^N\left(s - t_{m_s}\right)\right) G\left(\mathbb{Y}_s^N\right) \right\|^2_{\mathscr{L}_2^0} \mathrm{d}s \right\|^{\frac{1}{2}}_{L^{\frac{p}{2}}(\Omega)} \Bigg] \\
&\leq C\Bigg[\lambda_N^{-\frac{1}{2}} + \left(\int_0^t \left\| P^N X(s) - \mathbb{Y}_s^N \right\|^2_{L^p(\Omega;H)} \mathrm{d}s \right)^{\frac{1}{2}} + \left\| \int_0^t \left\| A^{\frac{1}{2}} S^N(t-s) A^{-\frac{1}{2}} \left(\mathrm{Id} - S^N\left(s - t_{m_s}\right)\right) G\left(\mathbb{Y}_t^N\right) \right\|^2_{\mathscr{L}_2^0} \mathrm{d}s \right\|^{\frac{1}{2}}_{L^{\frac{p}{2}}(\Omega)} \\
&\quad + \left\| \int_0^t \left\| A^{\frac{1}{2}} S^N(t-s) A^{-\frac{1}{2}} \left(\mathrm{Id} - S^N\left(s - t_{m_s}\right)\right) \left(G\left(\mathbb{Y}_t^N\right) - G\left(\mathbb{Y}_s^N\right)\right) \right\|^2_{\mathscr{L}_2^0} \mathrm{d}s \right\|^{\frac{1}{2}}_{L^{\frac{p}{2}}(\Omega)} \Bigg],
\end{aligned}$$

which together with (2.1)–(2.3), the Hölder continuity of $\mathbb{Y}_t^N$ and the Grönwall inequality yields that $\|P^N X(t) - \mathbb{Y}_t^N\|_{L^p(\Omega;H)} \leq C(\lambda_N^{-\frac{1}{2}} + \delta^{\frac{1}{2}})$.

For the term $\|\mathbb{Y}_t^N - X_t^N\|_{L^p(\Omega;H)}$, applying the Itô formula gives

$$\begin{aligned}\left\|X_t^N - \mathbb{Y}_t^N\right\|^2 &= 2\int_0^t \left\langle X_s^N - \mathbb{Y}_s^N, -A^N\left(X_s^N - \mathbb{Y}_s^N\right)\right\rangle \mathrm{d}t + 2\int_0^t \left\langle X_s^N - \mathbb{Y}_s^N, S^N\left(s-t_{m_s}\right)F^N\left(X_{t_{m_s}}^N\right) - F^N(X(s))\right\rangle \mathrm{d}s \\ &+ 2\int_0^t \left\langle X_s^N - \mathbb{Y}_s^N, S^N\left(s-t_{m_s}\right)P^N\left(G\left(X_{t_{m_s}}^N\right) - G\left(\mathbb{Y}_s^N\right)\right)\mathrm{d}W(s)\right\rangle + \int_0^t \left\|S^N\left(s-t_{m_s}\right)P^N\left(G\left(X_{t_{m_s}}^N\right) - G\left(\mathbb{Y}_s^N\right)\right)\right\|_{\mathcal{L}_2^0}^2 \mathrm{d}s.\end{aligned} \tag{6.6}$$

Then compared with the additive noise case (see (5.3)), the main difference of the estimation of $\|X_t^N - \mathbb{Y}_t^N\|_{L^p(\Omega;H)}$ lies in the estimation of the last two terms in (6.6), which can be estimated as for $p \geq 4$,

$$\begin{aligned}&\mathbb{E}\left[\left|\int_0^t \left\langle X_s^N - \mathbb{Y}_s^N, S^N\left(s-t_{m_s}\right)P^N\left(G\left(X_{t_{m_s}}^N\right) - G\left(\mathbb{Y}_s^N\right)\right)\mathrm{d}W(s)\right\rangle\right|^{\frac{p}{2}}\right] \\ &\leq C\mathbb{E}\left[\left(\int_0^t \left\|X_s^N - \mathbb{Y}_s^N\right\|^2 \left\|G\left(X_{t_{m_s}}^N\right) - G\left(\mathbb{Y}_s^N\right)\right\|_{\mathcal{L}_2^0}^2 \mathrm{d}s\right)^{\frac{p}{4}}\right] \\ &\leq C\int_0^t \mathbb{E}\left[\left\|X_s^N - \mathbb{Y}_s^N\right\|^p\right]\mathrm{d}s + C\delta^{\frac{p}{2}},\end{aligned}$$

and

$$\begin{aligned}&\mathbb{E}\left[\left|\int_0^t \left\|S^N\left(t-t_{m_s}\right)P^N\left(G\left(X_{t_{m_s}}^N\right) - G\left(\mathbb{Y}_s^N\right)\right)\right\|_{\mathcal{L}_2^0}^2 \mathrm{d}s\right|^{\frac{p}{2}}\right] \\ &\leq C\int_0^t \mathbb{E}\left[\left\|X_s^N - \mathbb{Y}_s^N\right\|^p\right]\mathrm{d}s + C\delta^{\frac{p}{2}},\end{aligned}$$

respectively, where the Hölder continuity of X_t^N, the Hölder inequality, the Young inequality and the Lipschitz condition of G are used. Hence, similar to the proof of (5.3), one can obtain that $\|X_t^N - \mathbb{Y}_t^N\|_{L^p(\Omega;H)} \leq C\lambda_N^{-\frac{1}{2}} + \delta^{\frac{1}{2}}$. This finishes the proof. □

Note that the proof of Proposition 6.2 is not applicable for the case of $d > 1$ due to that the Sobolev embedding $H^{\frac{d}{2}+} \hookrightarrow E$ requires a higher space regularity. Hence other skills are needed to obtain the regularity estimate of (AE-M) in E-norm. We leave this as the future work and attempt to study it in the future.

7. Numerical experiments

In this section, we present numerical experiments for the trace-class noise case and the space-time white noise case to verify the previous theoretical results, respectively. Meanwhile, some alternative choices of the adaptive timestep function are given.

Consider the following stochastic Allen–Cahn equation with the generalized Q-Wiener process:

$$\begin{cases} \frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + u - u^3 + \dot{W}(t,x), \quad t \in (0,1],\ x \in (0,1), \\ u(0,x) = \sqrt{2}\sin(\pi x), \quad x \in (0,1), \\ u(t,0) = u(t,1) = 0, \quad t \in (0,1]. \end{cases} \tag{7.1}$$

For the trace-class noise case, we choose Q such that $Qe_i = \frac{1}{i^2}e_i,\ i \geq 1$, which implies that $tr(Q) < \infty$ and Assumption 2.2 holds for some $\beta \in [1,2]$. For the space-time white noise case, i.e., $Q = \mathrm{Id}$, Assumption 2.2 holds for $\beta \in (0, \frac{1}{2})$.

We are going to compare the performance of adaptive schemes with the scheme with a uniform timestep. The scheme with a uniform timestep is chosen as the tamed exponential integrator (TE), see e.g., Wang (2020). In the following, error bounds are measured in root-mean-square (RMS) sense at the end point $T = 1$, caused by spatio-temporal discretization. And the expectations are approximated by computing averages over 1000 samples. The infinite-dimensional Hilbert space $L^2(0,1)$ is approximated through the finite-dimensional subspace spanned by the first 2^{10} eigenfunctions of the Laplacian, i.e., in the spectral Galerkin method, we take $N = 2^{10}$. Since the exact solution cannot be given explicitly, we take the solution generated by the same method with a timestep that is three times smaller as the reference solution. Numerical experiments are tested by Matlab R2017a in MacBook Pro (13-inch, 2019, Two Thunderbolt 3 ports).

Besides (3.8), there are other choices for the backstop scheme Ψ, for example, the nonlinearity-truncated exponential integrator (see Becker & Jentzen, 2019, Eq. (6))

$$\Psi(x,h,y) = S^N(h)\left(x + hF^N(x)\chi_{\left\{\|x\|^6_{L^{18}(\mathscr{O})} \leq h^{-1}\right\}} + y\right), \tag{7.2}$$

and the linear-implicit nonlinearity-truncated scheme (see Becker & Jentzen, 2019, Eq. (7))

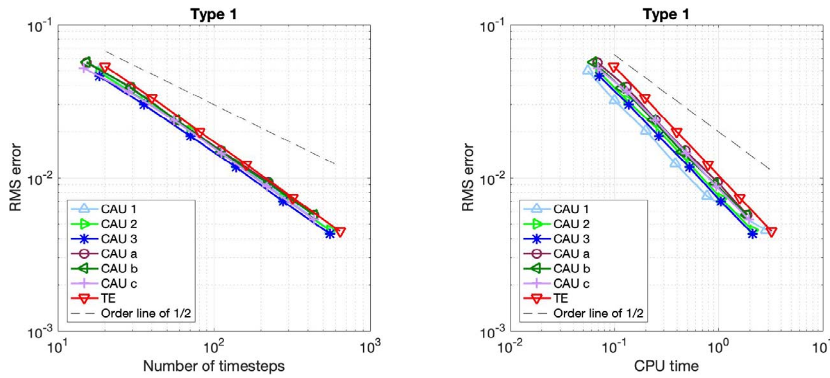
$$\Psi(x,h,y) = (\mathrm{Id} - A^N h)^{-1}\left(x + hF^N(x)\chi_{\left\{\|x\|^6_{L^{18}(\mathscr{O})} \leq h^{-1}\right\}} + y\right). \tag{7.3}$$

Hence, for the coupled scheme (3.7), when the backstop scheme Ψ is chosen to be (7.2) or (7.3), we obtain two coupled schemes denoted by (CAU 2) and (CAU 3), respectively. Similarly, for the coupled scheme (3.10), when the backstop scheme Ψ is chosen to be (3.8), (7.2) or (7.3), we obtain three coupled schemes denoted by (CAU a), (CAU b) and (CAU c), respectively.

7.1 *Trace-class noise*

We show the experiment results for the trace-class noise case in this subsection. For the trace-class noise case, the seven methods tested are: the scheme TE that is with the uniform timestep, schemes (CAU 1), (CAU 2), (CAU 3), (CAU a), (CAU b) and (CAU c) that are with the adaptive timestep. Recall that τ^δ is the refined timestep function controlled by the scalar parameter $\delta \in (0,1)$ and satisfies Assumption 3.3. We choose the following two types of adaptive timestep functions: *type 1:*

$$\tau^\delta_{11}(X) = \delta\tau_{11}(X) \text{ with } \tau_{11}(X) = \left(\frac{\|X\| + 0.5}{2\|F(X)\| + 1.8}\right)^{\frac{4}{3}},$$



(a) timestep type 1 (b) timestep type 1

FIG. 1. RMS error for numerical schemes with trace-class noise for type 1.

$$\tau_{12}^{\delta}(X) = \delta\tau_{12}(X) \text{ with } \tau_{12}(X) = \min\left\{\frac{2\|X\|_{L^4}^4 + 0.01}{\|X\|_{L^6}^6 + 0.01}, \left(\frac{\|X\| + 0.2}{\|F(X)\| + 0.8}\right)^{\frac{4}{3}}\right\},$$

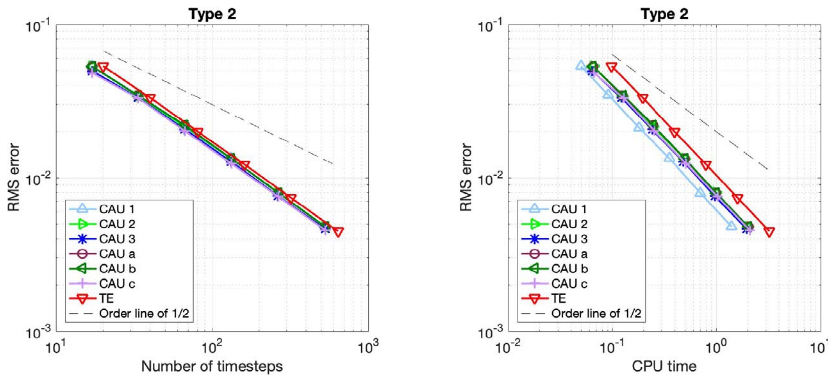
type 2:

$$\tau_{21}^{\delta}(X) = \delta\tau_{21}(X) \text{ with } \tau_{21}(X) = \left(\frac{1.1}{\|F(X)\| + 3.2}\right)^{\frac{4}{3}},$$

$$\tau_{22}^{\delta}(X) = \delta\tau_{22}(X) \text{ with } \tau_{22}(X) = \min\left\{\frac{\|X\|^2 + 0.01}{\|X\|_{L^6}^6 + 0.04}, \left(\frac{1.1}{\|F(X)\| + 3.2}\right)^{\frac{4}{3}}\right\},$$

where $\delta = 2^{-l}, l = 2, \ldots, 7$. Here, τ_{11}, τ_{21} are adaptive timestep functions for (CAU 1), (CAU 2) and (CAU 3), and τ_{12}, τ_{22} are adaptive timestep functions for (CAU a), (CAU b) and (CAU c). In *type 1*, we let $L_2(\omega) = \|X(\omega)\| + c$ with X being the reference solution or numerical solution, $c = 0.5$ in $\tau_{1,1}^{\delta}$ and $c = 0.2$ in $\tau_{1,2}^{\delta}$, and $\bar{L}_2 = \frac{3}{2}$. In *type 2*, we let $L_2 = 1.1$ and $\bar{L}_2 = 2$ independent of ω. One can verify that the *type 1* and *type 2* timestep functions satisfy Assumption 3.3.

The comparisons of the seven schemes with *type 1* and *type 2* timestep functions are presented in Fig. 1 and Fig. 2, respectively, where the left ones are about the RMS error against the average number of timesteps, and the right ones are about the RMS error against the CPU time. Recall (3.3) and (3.5), and set $\tau_{min} = 0.2\delta$, $\zeta_1 = 3, q_0 = 1$ and $\zeta_2 = 4$ for *type 1* and *type 2*. The length of a timestep for TE is set by τ_{min}. From Figs 1 and 2, we observe that the convergence orders in the temporal direction of these seven schemes are slightly bigger than $\frac{1}{2}$. As for the *type 1* timestep function, for a given RMS error, we see from Fig. 1 (*a*) that the adaptive schemes cost slightly fewer numbers of timesteps than TE that has the uniform timestep. And we observe from Fig. 1 (*b*) that the adaptive schemes cost less CPU time than TE for a given RMS error. A similar phenomenon is observed from Fig. 2 for the *type 2* timestep function. These mean that for the *type 1* and *type 2* timestep functions, the adaptive schemes perform slightly better than TE that uses the uniform timestep.



(a) timestep type 2 (b) timestep type 2

FIG. 2. RMS error for numerical schemes with trace-class noise for type 2.

TABLE 1 *Time in seconds* ($\delta = 2^{-6}$)

	CAU 1	CAU 2	CAU 3	CAU a	CAU b	CAU c	TE
Type 1	81.78	82.80	82.83	77.66	77.00	86.08	93.52
Type 2	76.84	82.74	79.54	87.29	88.75	86.55	93.52

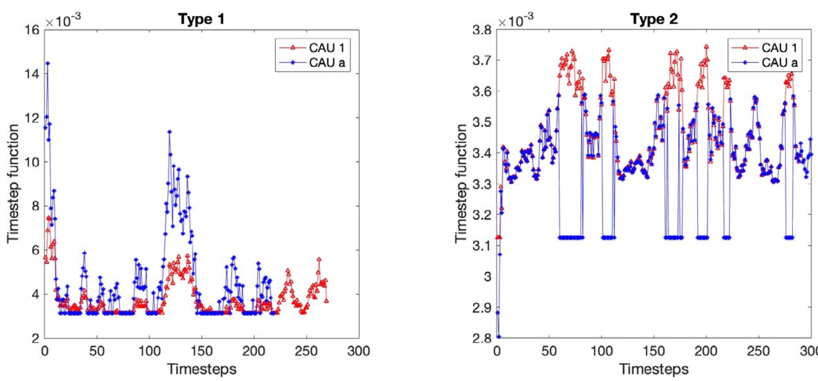
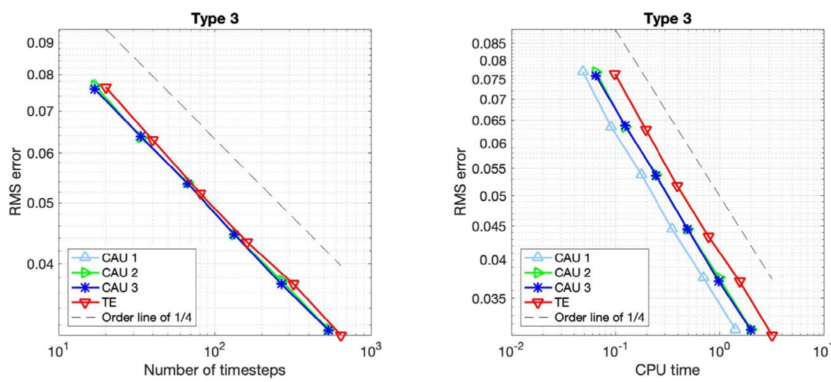


FIG. 3. Timestep function for numerical schemes with trass-class noise ($\delta = 2^{-6}$).

We list the CPU time costed by these schemes with the *type i* ($i = 1, 2$) timestep functions in Table 1. Here, the uniform timestep is $\tau_{min} = 0.2\delta$ with $\delta = 2^{-6}$. The critical parameter for (CAU 1)-(CAU 3) is τ_{min}, and that for (CAU a)-(CAU c) is $(\zeta_1 \|X\|^{q_0} + \zeta_2)^{-1}$ with $\zeta_1 = 3, \zeta_2 = 4, q_0 = 1$. And 1000 realizations are calculated. It can be observed that the coupled schemes are in a lower computational cost in terms of the CPU time compared with TE that uses the uniform timestep. This can be explained by Fig. 3, since coupled schemes like (CAU 1) and (CAU c) use many large timesteps compared with the uniform timestep, which may reduce the computational cost of the schemes.

TABLE 2 *Time in seconds* ($\delta = 2^{-6}$)

	CAU 1	CAU 2	CAU 3	TE
Type 3	82.80	81.98	77.85	102.62



(a) timestep type 3 (b) timestep type 3

FIG. 4. RMS error for numerical schemes with space-time white noise.

7.2 *Space-time white noise*

We show the experiment results for the space-time white noise case in this subsection. For the space-time white noise case, the four schemes tested are: the scheme TE that is with the uniform timestep, schemes (CAU 1), (CAU 2) and (CAU 3) that are with the adaptive timestep. We choose the following adaptive timestep function for the considered three adaptive schemes: *type 3:*

$$\tau_{31}^{\delta}(X) = \delta\tau_{31}(X) \text{ with } \tau_{31}(X) = \left(\frac{1.1}{\|F(X)\| + 3}\right)^{\frac{4}{3}},$$

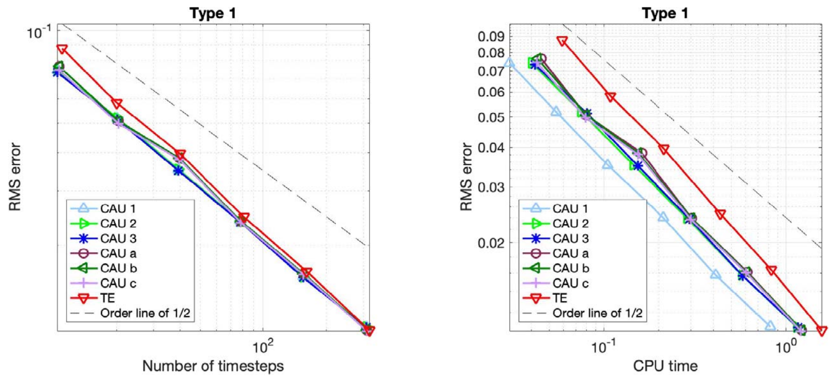
where δ is the same as before. In this case, the *type 3* timestep function can also be verified to satisfy Assumption 3.3.

For space-time white noise case, Fig. 4 shows the comparison of RMS error against the number of timesteps and CPU time. It can be gained from Fig. 4 that the convergence orders in the temporal direction of the four schemes are $\frac{1}{4}$. For a given RMS error, we see from Fig. 4 (*a*) that the adaptive schemes cost slightly fewer numbers of timesteps than TE that uses the uniform timestep. And we observe from Fig. 4 (*b*) that the adaptive schemes cost less CPU time than TE for a given RMS error. The total time for these schemes with *type 3* timestep function is listed in Table 2, from which we can see that the coupled schemes are in a lower computational cost in terms of the CPU time. These mean that for the *type 3* timestep function, the adaptive schemes still perform slightly better than TE that uses the uniform timestep.

At the end of this section, we present the experiment result of the multiplicative noise case in one dimensional case with the diffusion coefficient $G(X)(x) = X(x) + 1, x \in \mathcal{O}$ and the trace-class operator Q satisfying $Qe_i = \frac{1}{i^2}e_i, i \geq 1$. Here, we only take *type 1* timestep functions as an example, which are

TABLE 3 *Time in seconds* ($\delta = 2^{-6}$)

	CAU 1	CAU 2	CAU 3	CAU a	CAU b	CAU c	TE
Type 1	48.88	49.68	55.21	54.61	55.55	54.75	97.26



(a) timestep type 1 (b) timestep type 1

FIG. 5. RMS error for numerical schemes of multiplicative noise case.

taken as

$$\tau_{11}^{\delta}(X) = \delta\left(\frac{\|X\| + 0.6}{\|F(X)\| + 1.4}\right)^{\frac{4}{3}}, \quad \tau_{12}^{\delta}(X) = \delta \min\left\{\frac{2\|X\|_{L^4}^4 + 0.01}{\|X\|_{L^6}^6 + 0.01}, \left(\frac{\|X\| + 0.3}{\|F(X)\| + 0.8}\right)^{\frac{4}{3}}\right\}.$$

Figure 5 shows the comparison of RMS error against the number of timesteps and CPU time, which implies the convergence orders in the temporal direction of schemes are $\frac{1}{2}$. For a given RMS error, it can be observed that in Fig. 5 (*a*), the adaptive schemes cost slightly fewer number of timesteps than TE with the uniform timestep, and that in Fig. 5 (*b*), the adaptive schemes cost less CPU time than TE. The total CPU time is listed in Table 3, from which we can see that the coupled schemes are in a lower computational cost in terms of the CPU time. These show the better performance of adaptive schemes for the multiplicative noise case.

Funding

National Key R&D Program of China (No. 2020YFA0713701); National Natural Science Foundation of China (No. 12022118, No. 12031020, No. 11971470, No. 11871068); Youth Innovation Promotion Association CAS.

REFERENCES

BECCARI, M., HUTZENTHALER, M., JENTZEN, A., KURNIAWAN, R., LINDNER, F. & SALIMOVA, D. (2019) Strong and weak divergence of exponential and linear-implicit Euler approximations for stochastic partial differential equations with superlinearly growing nonlinearities. arXiv: 1903.06066.

Becker, S. & Jentzen, A. (2019) Strong convergence rates for nonlinearity-truncated Euler-type approximations of stochastic Ginzburg–Landau equations. *Stochastic Process. Appl.*, **129**, 28–69.

Bréhier, C., Cui, J. & Hong, J. (2019) Strong convergence rates of semidiscrete splitting approximations for the stochastic Allen–Cahn equation. *IMA J. Numer. Anal.*, **39**, 2096–2134.

Bréhier, C. & Goudenège, L. (2019) Analysis of some splitting schemes for the stochastic Allen–Cahn equation. *Discrete Contin. Dyn. Syst. Ser. B*, **24**, 4169–4190.

Cai, M., Gan, S., Wang, X. (2021) Weak convergence rates for an explicit full-discretization of stochastic Allen-Cahn equation with additive noise. *J. Sci. Comput.*, **86**, Paper No. 34, 30.

Campbell, S. & Lord, G. (2018) Adaptive time-stepping for stochastic partial differential equations with non-lipschitz drift. arXiv: 1812.09036.

Cerrai, S. (2001) *Second Order PDE's in Finite and Infinite Dimension. A Probabilistic Approach*. Lecture Notes in Mathematics, vol. 1762. Berlin: Springer.

Chen, C. & Hong, J. (2016) Symplectic Runge–Kutta semidiscretization for stochastic Schrödinger equation. *SIAM J. Numer. Anal.*, **54**, 2569–2593.

Chen, H., Bhakta, M. & Hajaiej, H. (2022) On the bounds of the sum of eigenvalues for a Dirichlet problem involving mixed fractional Laplacians. *J. Differential Equations*, **317**, 1–31.

Cui, J. & Hong, J. (2019) Strong and weak convergence rates of a spatial approximation for stochastic partial differential equation with one-sided Lipschitz coefficient. *SIAM J. Numer. Anal.*, **57**, 1815–1841.

Cui, J., Hong, J. & Sun, L. (2021) Weak convergence and invariant measure of a full discretization for parabolic SPDEs with non-globally Lipschitz coefficients. *Stochastic Process. Appl.*, **134**, 55–93.

Fang, W. & Giles, M. (2020) Adaptive Euler–Maruyama method for SDEs with nonglobally Lipschitz drift. *Ann. Appl. Probab.*, **30**, 526–560.

Feng, X., Li, Y. & Zhang, Y. (2021) 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**, 574–598.

Hoel, H., von Schwerin, E., Szepessy, A. & Tempone, R. (2012) Adaptive multilevel Monte Carlo simulation. *Numerical Analysis of Multiscale Computations*. Lect. Notes Comput. Sci. Eng., vol. 82. Heidelberg: Springer, pp. 217–234.

Jentzen, A. & Pušnik, P. (2020) Strong convergence rates for an explicit numerical approximation method for stochastic evolution equations with non-globally Lipschitz continuous nonlinearities. *IMA J. Numer. Anal.*, **40**, 1005–1050.

Kelly, C. & Lord, G. J. (2018) Adaptive time-stepping strategies for nonlinear stochastic systems. *IMA J. Numer. Anal.*, **38**, 1523–1549.

Kelly, C. & Lord, G. J. (2022) Adaptive Euler methods for stochastic systems with non-globally Lipschitz coefficients. *Numer. Algorithms*, **89**, 721–747.

Kovács, M., Larsson, S. & Lindgren, F. (2018) On the discretisation in time of the stochastic Allen–Cahn equation. *Math. Nachr.*, **291**, 966–995.

Kruse, R. (2014) *Strong and Weak Approximation of Semilinear Stochastic Evolution Equations*. Lecture Notes in Mathematics, vol. 2093. Cham: Springer.

Lamba, H. (2003) An adaptive timestepping algorithm for stochastic differential equations. *J. Comput. Appl. Math.*, **161**, 417–430.

Lemaire, V. (2007) An adaptive scheme for the approximation of dissipative systems. *Stochastic Process. Appl.*, **117**, 1491–1518.

Liu, W. & Röckner, M. (2015) *Stochastic Partial Differential Equations: An Introduction*. Universitext. Cham: Springer.

Liu, Z. & Qiao, Z. (2020) Strong approximation of monotone stochastic partial differential equations driven by white noise. *IMA J. Numer. Anal.*, **40**, 1074–1093.

Liu, Z. & Qiao, Z. (2021) Strong approximation of monotone stochastic partial differential equations driven by multiplicative noise. *Stoch. Partial Differ. Equ. Anal. Comput.*, **9**, 559–602.

Lord, G. J., Powell, C. E. & Shardlow, T. (2014) *An Introduction to Computational Stochastic PDEs*. Cambridge Texts in Applied Mathematics. New York: Cambridge University Press.

Majee, A. K. & Prohl, A. (2018) Optimal strong rates of convergence for a space-time discretization of the stochastic Allen–Cahn equation with multiplicative noise. *Comput. Methods Appl. Math.*, **18**, 297–311.

Merle, F. & Prohl, A. (2021) An adaptive time-stepping method based on a posteriori weak error analysis for large SDE systems. *Numer. Math.*, **149**, 417–462.

Milstein, G. N. & Tretyakov, M. V. (2021) *Stochastic Numerics for Mathematical Physics*, 2nd edn. Scientific Computation. Cham: Springer [of 2069903].

Qi, R. & Wang, X. (2019) Optimal error estimates of Galerkin finite element methods for stochastic Allen–Cahn equation with additive noise. *J. Sci. Comput.*, **80**, 1171–1194.

Wang, X. (2020) An efficient explicit full-discrete scheme for strong approximation of stochastic Allen–Cahn equation. *Stochastic Process. Appl.*, **130**, 6271–6299.

Appendix A.

Proof of Lemma 4.1. (*i*) Let $u=\sum_{i=1}^{\infty}\langle u,e_i\rangle e_i$. Then the Hölder inequality gives that

$$\|A^{\rho}P^N S(t)u\|_E=\sup_{x\in\mathscr{O}}\left|\sum_{i=1}^{N}e^{-\lambda_i t}\lambda_i^{\rho}\langle u,e_i\rangle e_i(x)\right|\le C\sum_{i=1}^{N}e^{-\lambda_i t}\lambda_i^{\rho}|\langle u,e_i\rangle|$$

$$\le C\left(\sum_{i=1}^{N}\lambda_i^{2\rho}e^{-2\lambda_i t}\right)^{\frac12}\left(\sum_{i=1}^{N}|\langle u,e_i\rangle|^2\right)^{\frac12}\le C\left(\sum_{i=1}^{N}(i^{2/d})^{2\rho}e^{-cti^{2/d}}\right)^{\frac12}\|u\|.$$

By the monotonicity of the function $y^{4\rho/d}\,e^{-cty^{2/d}}$ w.r.t. $y\in(0,\infty)$, which takes the maximum value at the point $y=(\frac{2\rho}{ct})^{d/2}$, and letting $i_0:=\lfloor(\frac{2\rho}{ct})^{d/2}\rfloor$ with $\lfloor\cdot\rfloor$ being the floor function, we have

$$\sum_{i=1}^{\infty}i^{4\rho/d}e^{-cti^{2/d}}=i_0^{4\rho/d}e^{-cti_0^{2/d}}+\sum_{i=1}^{i_0-1}i^{4\rho/d}e^{-cti^{2/d}}+\sum_{i=i_0+1}^{\infty}i^{4\rho/d}e^{-cti^{2/d}}$$

$$\le\left(\frac{2\rho}{ct}\right)^{2\rho}e^{-2\rho}+\int_0^{\infty}y^{4\rho/d}e^{-cty^{2/d}}\,\mathrm{d}y$$

$$\le C(\rho,d)\left(\frac{1}{t^{2\rho}}+\frac{1}{t^{2\rho+\frac d2}}\int_0^{\infty}z^{2\rho+\frac d2-1}e^{-cz}\,\mathrm{d}z\right)\le C(\rho,d)\left(\frac{1}{t^{2\rho}}+\frac{1}{t^{2\rho+\frac d2}}\right),$$

where in the last step we have used the fact that $\int_0^{\infty}z^{2\rho+\frac d2-1}e^{-cz}\,\mathrm{d}z=c^{-2\rho-\frac d2}\Gamma(2\rho+\frac d2)$ with Γ being the Gamma function. Applying the inequality $(a+b)^{\frac12}\le a^{\frac12}+b^{\frac12}$ leads to the desired result.

(*ii*) By the Hölder inequality, we obtain

$$\|P^N S(t)u\|_E=\sup_{x\in\mathscr{O}}\left|\sum_{i=1}^{N}e^{-\lambda_i t}\langle u,e_i\rangle e_i(x)\right|\le C\sum_{i=1}^{N}e^{-\lambda_i t}|\langle u,e_i\rangle|$$

$$\le C\left(\sum_{i=1}^{N}e^{-2\lambda_i t}\lambda_i^{-\rho}\right)^{\frac12}\left(\sum_{i=1}^{N}\lambda_i^{\rho}|\langle u,e_i\rangle|^2\right)^{\frac12}\le C\left(\int_0^{\infty}y^{-2\rho/d}e^{-cy^{2/d}t}\,\mathrm{d}y\right)^{\frac12}\|u\|_{\rho}$$

$$\le C(\rho,d)\left(t^{\rho-\frac d2}\int_0^{\infty}z^{-\rho+d/2-1}e^{-cz}\,\mathrm{d}z\right)^{\frac12}\|u\|_{\rho}\le C(\rho,d)t^{\frac{\rho}{2}-\frac d4}\|u\|_{\rho},$$

where in the last step we have used the fact that $\int_0^\infty z^{-\rho+\frac{d}{2}-1}e^{-cz}\,\mathrm{d}z = c^{\rho-\frac{d}{2}}\Gamma(-\rho+\frac{d}{2})$ for $\rho \in [0, \frac{d}{2})$. The proof is finished. □

Proof of Lemma 4.2. By Lemma 4.1 (ii) with $\rho = 0$ and the Sobolev embedding $\dot{H}^1 \hookrightarrow L^6(\mathscr{O})$, we get

$$\left\|v_t^N\right\|_E \le C\int_0^t (t-s)^{-\frac{d}{4}}\left(1+\left\|\nabla v_s^N\right\|^3 + \left\|z_s^N\right\|^3_{L^6(\mathscr{O})}\right)\mathrm{d}s. \tag{A.1}$$

The Taylor formula, the Young inequality, the Gagliardo–Nirenberg inequality $\|u\|_{L^{\eta_1}(\mathscr{O})} \le C\left(\|\Delta u\|^\alpha\|u\|^{1-\alpha} + \|u\|_{L^{\eta_2}(\mathscr{O})}\right)$ for $u \in H^2$ with $\alpha = \frac{(\eta_1-2)d}{4\eta_1} \in (0,1]$, $\eta_1 > 2$, $\eta_2 > 0$ imply that

$$\begin{aligned}
\left\|\nabla v_t^N\right\|^2 &= \int_0^t \left(-2\left\|Av_s^N\right\|^2 + 2\left\langle \nabla v_s^N, \nabla F^N\left(v_s^N+z_s^N\right)\right\rangle\right)\mathrm{d}s \\
&\le -(2-\epsilon)\int_0^t \left\|Av_s^N\right\|^2\mathrm{d}s + C(\epsilon)\int_0^t\left(1+\left\|v_s^N\right\|^4_{L^4(\mathscr{O})}\left(\left\|z_s^N\right\|_E^2+1\right)+\left\|v_s^N\right\|^2\left\|z_s^N\right\|_E^4+\left\|z_s^N\right\|_E^6+\left\|v_s^N\right\|^2\right)\mathrm{d}s \\
&\le -(2-\epsilon)\int_0^t \left\|Av_s^N\right\|^2\mathrm{d}s + C(\epsilon)\int_0^t\left(1+\left(\left\|Av_s^N\right\|^{\frac{d}{2}}\left\|v_s^N\right\|^{4-\frac{d}{2}}+\left\|v_s^N\right\|^4\right)\left(\left\|z_s^N\right\|_E^2+1\right)+\left\|v_s^N\right\|^2\left\|z_s^N\right\|_E^4\right. \\
&\quad \left.+\left\|z_s^N\right\|_E^6+\left\|v_s^N\right\|^2\right)\mathrm{d}s \\
&\le C(\epsilon)\int_0^t\left(1+\left\|v_s^N\right\|^{\frac{2(8-d)}{4-d}}\left(\left\|z_s^N\right\|_E^{\frac{8}{4-d}}+1\right)+\left\|v_s^N\right\|^6+\left\|z_s^N\right\|_E^6\right)\mathrm{d}s,
\end{aligned} \tag{A.2}$$

where in the second step we use the fact that $\langle\nabla v_s^N, 3a_3(v_s^N)^2\nabla v_s^N\rangle \le 0$, and in the third step we take $\eta_1 = 4, \eta_2 = 2$. Using the Taylor formula again, and combining (2.5) and the Young inequality give that

$$\begin{aligned}
\left\|v_t^N\right\|^2 &\le -2\int_0^t\left\|\nabla v_s^N\right\|^2\mathrm{d}s + 2L_0\int_0^t\left\|v_s^N\right\|^2\mathrm{d}s + 2\int_0^t\left\langle v_s^N, F^N\left(z_s^N\right)\right\rangle\mathrm{d}s \\
&\le 2L_0\int_0^t\left\|v_s^N\right\|^2\mathrm{d}s + C\int_0^t\left(1+\left\|v_s^N\right\|^2+\left\|z_s^N\right\|_E^6\right)\mathrm{d}s,
\end{aligned}$$

which yields that $\|v_t^N\|^2 \le Ce^{(2L_0+C)T}\int_0^t(1+\|z_s^N\|_E^6)\,\mathrm{d}s$ due to the Grönwall inequality. This, together with (A.2) leads to

$$\begin{aligned}
\sup_{0\le s\le t}\left\|\nabla v_s^N\right\|^2 &\le C\left[\sup_{0\le s\le t}\left\|v_s^N\right\|^{\frac{2(8-d)}{4-d}}\int_0^t\left(\|z_s^N\|_E^{\frac{8}{4-d}}+1\right)\mathrm{d}s + \int_0^t\left(1+\left\|z_s^N\right\|_E^6\right)\mathrm{d}s + \sup_{0\le s\le t}\left\|v_s^N\right\|^6\right] \\
&\le C\left[1+\left(\int_0^t\left(1+\left\|z_s^N\right\|_E^8\right)\mathrm{d}s\right)^{\frac{8-d}{4-d}+1}\right].
\end{aligned}$$

Plugging the above inequality into the right-hand side of (A.1), one has

$$\begin{aligned}\left\|v_t^N\right\|_E &\le C\int_0^t (t-s)^{-\frac{d}{4}}\left(1+\left\|z_s^N\right\|_E^3\right)\,\mathrm{d}s + C\sup_{0\le s\le t}\left\|\nabla v_s^N\right\|^3\\ &\le C\int_0^t (t-s)^{-\frac{d}{4}}\left(1+\left\|z_s^N\right\|_E^3\right)\,\mathrm{d}s + C\left[1+\left(\int_0^t\left(1+\left\|z_s^N\right\|_E^8\right)\,\mathrm{d}s\right)^{\frac{8-d}{4-d}+1}\right]^{\frac{3}{2}}\\ &\le C\left[1+\int_0^t\left((t-s)^{-\frac{d}{4}}\left\|z_s^N\right\|_E^3+\left\|z_s^N\right\|_E^{72}\right)\,\mathrm{d}s\right],\end{aligned}$$

where we have used the fact that $\frac{3}{2}(\frac{8-d}{4-d}+1)\le 9$ for $d\le 3$. The proof is finished. □

Proof of Proposition 4.3. First, define a sequence of nonincreasing events as follows:

$$\Omega_t := \left\{\omega\in\Omega : \sup_{j\in\{0,1,\cdots,m_t(\omega)\}}\left(\tau_j^\delta\right)^\alpha\left\|X_{t_j}^N(\omega)\right\|_E\le 1\right\},\quad 0\le t\le T$$

with $\alpha>0$ being determined later. It is clear that $\chi_{\Omega_t}\in\mathscr{F}_t$. Note that one can always choose $\delta>0$ sufficiently small so that $\mathbb{P}\left(\{(\tau_0^\delta)^\alpha\|X_0^N\|_E>1\}\right)\le(T\delta)^\alpha\mathbb{E}[\|X_0^N\|_E]<1$, and hence $\mathbb{P}(\Omega_0)>0$, which implies $\Omega_0\ne\emptyset$. Intuitively, Ω_t is the event that the E-norm of the numerical solution can be bounded by timestep sizes with a certain order before time t (including t).

We claim that

$$\mathbb{E}\left[\chi_{\Omega_t}\left\|X_t^N\right\|_E^p\right]\le C\left(p,T,L_1,L_2,\left\|A^{\frac{\beta-1}{2}}Q^{\frac{1}{2}}\right\|_{\mathscr{L}_2(H)},\|X_0\|_{L^{q_1}(\Omega;\dot{H}^\beta)},\left\|P^NX_0\right\|_{L^{q_1}(\Omega;E)}\right)\tag{A.3}$$

and

$$\mathbb{E}\left[\chi_{\Omega_t^c}\left\|X_t^N\right\|_E^p\right]\le C\left(p,T,\tau_{min},L_1,L_2,\left\|A^{\frac{\beta-1}{2}}Q^{\frac{1}{2}}\right\|_{\mathscr{L}_2(H)},\|X_0\|_{L^{q_2}(\Omega;\dot{H}^\beta)},\left\|P^NX_0\right\|_{L^{q_2}(\Omega;E)}\right)\tag{A.4}$$

with $q_1,q_2>0$ sufficiently large, whose proofs are given in *Step 1* and *Step 2*, respectively. Once we prove (A.3) and (A.4), the proof of (4.2) is finished due to the Minkowski inequality.

Step 1: Proof of (A.3).

Let

$$Z_t^N := S^N(t)P^NX_0+\int_0^t S^N\left(t-t_{m_s}\right)F^N\left(X_{t_{m_s}}^N\right)\,\mathrm{d}s-\int_0^t S^N(t-s)F^N\left(X_s^N\right)\,\mathrm{d}s+W_A^N(t),\tag{A.5}$$

where $W_A^N(t):=P^NW_A(t)$ with $W_A(t)=\int_0^t S\left(t-t_{m_s}\right)\mathrm{d}W(s)$. Then $X_t^N=\int_0^t S^N(t-s)F^N\left(X_s^N\right)\mathrm{d}s+Z_t^N$. Let $\bar{X}_t^N:=X_t^N-Z_t^N$. It satisfies

$$\begin{cases}\mathrm{d}\bar{X}_t^N=\left(-A^N\bar{X}_t^N+F^N\left(\bar{X}_t^N+Z_t^N\right)\right)\,\mathrm{d}t, & t\in(0,T],\\ \bar{X}_0^N=0.\end{cases}$$

Then applying Lemma 4.2 yields that $\|\bar{X}_t^N\|_E\le C\int_0^t\left(1+(t-s)^{-\frac{d}{4}}\left\|Z_s^N\right\|_E^3+\left\|Z_s^N\right\|_E^{72}\right)\mathrm{d}s$. This implies

$$\left\|\chi_{\Omega_t}\left\|\bar{X}_t^N\right\|_E\right\|_{L^p(\Omega)}\le C\int_0^t\left(1+(t-s)^{-\frac{d}{4}}\left\|\chi_{\Omega_t}\left\|Z_s^N\right\|_E\right\|_{L^{3p}(\Omega)}^3+\left\|\chi_{\Omega_t}\left\|Z_s^N\right\|_E\right\|_{L^{72p}(\Omega)}^{72}\right)\,\mathrm{d}s.\tag{A.6}$$

It follows from the definition of Z_s^N that for $0 \le s \le t$,

$$\begin{aligned}\chi_{\Omega_t}\left\|Z_s^N\right\|_E &\le \left\|S^N(s)P^NX_0\right\|_E + \chi_{\Omega_t}\left\|\int_0^s S^N(s-r)\left(F^N\left(X_{t_{m_r}}^N\right)-F^N\left(X_r^N\right)\right)\mathrm{d}r\right\|_E \\ &\quad + \chi_{\Omega_t}\left\|\int_0^s\left(S^N\left(s-t_{m_r}\right)-S^N(s-r)\right)F^N\left(X_{t_{m_r}}^N\right)\mathrm{d}r\right\|_E + \left\|W_A^N(s)\right\|_E \\ &=: \left\|S^N(s)P^NX_0\right\|_E + I_0(s) + I_1(s) + \left\|W_A^N(s)\right\|_E. \end{aligned} \tag{A.7}$$

Estimate of I_0. Since

$$X_r^N = S^N\left(r-t_{m_r}\right)X_{t_{m_r}}^N + \int_{t_{m_r}}^r S^N\left(r-t_{m_u}\right)F^N\left(X_{t_{m_u}}^N\right)\mathrm{d}u + W_A^N(r) - S^N\left(r-t_{m_r}\right)W_A^N(t_{m_r}), \tag{A.8}$$

and by the contractivity of the semigroup in E, i.e., $\|S(t)u\|_E \le \|u\|_E$, (3.2), Assumption 3.3 and Lemma 4.1 (*ii*) with $\rho = 0$, we obtain

$$\begin{aligned}\chi_{\Omega_t}\|X_r^N\|_E &\le \chi_{\Omega_t}\left(\left\|X_{t_{m_r}}^N\right\|_E + C\int_{t_{m_r}}^r\left(r-t_{m_u}\right)^{-\frac{d}{4}}\left\|F^N\left(X_{t_{m_u}}^N\right)\right\|\mathrm{d}u + \left\|W_A^N(r)\right\|_E + \left\|W_A^N\left(t_{m_r}\right)\right\|_E\right) \\ &\le \chi_{\Omega_t}\left(\left\|X_{t_{m_r}}^N\right\|_E + C\left(\tau_{m_r}^\delta\right)^{1-\frac{d}{4}}\left\|F\left(X_{t_{m_r}}^N\right)\right\| + \left\|W_A^N(r)\right\|_E + \left\|W_A^N\left(t_{m_r}\right)\right\|_E\right) \\ &\le \chi_{\Omega_t}\left(\left\|X_{t_{m_r}}^N\right\|_E + CL_2 + \left\|W_A^N(r)\right\|_E + \left\|W_A^N\left(t_{m_r}\right)\right\|_E\right). \end{aligned} \tag{A.9}$$

We derive from (2.1)–(2.2) and (3.1)–(3.2) that for $\beta \in (0,2]$, $0 \le r \le t$,

$$\begin{aligned}\chi_{\Omega_t}\left\|X_r^N - X_{t_{m_r}}^N\right\| &\le \left\|\left(S(r)-S\left(t_{m_r}\right)\right)P^NX_0\right\| + \chi_{\Omega_t}\left\|\int_0^r S\left(r-t_{m_u}\right)F^N\left(X_{t_{m_u}}^N\right)\mathrm{d}u\right. \\ &\quad \left. - \int_0^{t_{m_r}} S\left(t_{m_r}-t_{m_u}\right)F^N\left(X_{t_{m_u}}^N\right)\mathrm{d}u\right\| + \left\|P^N\left(W_A(r)-W_A\left(t_{m_r}\right)\right)\right\| \\ &\le \left(\tau_{m_r}^\delta\right)^{\frac{\beta}{2}}\left\|P^NX_0\right\|_\beta + \chi_{\Omega_t}\left\|\int_0^{t_{m_r}} S\left(t_{m_r}-t_{m_u}\right)\left(S\left(r-t_{m_r}\right)-\mathrm{Id}\right)F^N\left(X_{t_{m_u}}^N\right)\mathrm{d}u\right\| \\ &\quad + \chi_{\Omega_t}\left\|\int_{t_{m_r}}^r S\left(r-t_{m_r}\right)F^N\left(X_{t_{m_r}}^N\right)\mathrm{d}u\right\| + \left\|P^N\left(W_A(r)-W_A\left(t_{m_r}\right)\right)\right\| \\ &\le \left(\tau_{m_r}^\delta\right)^{\frac{\beta}{2}}\left\|P^NX_0\right\|_\beta + C\chi_{\Omega_t}\int_0^{t_{m_r}}\left(t_{m_r}-t_{m_u}\right)^{-1+\epsilon}\left(\tau_{m_r}^\delta\right)^{1-\epsilon}\left\|F\left(X_{t_{m_u}}^N\right)\right\|\mathrm{d}u \\ &\quad + C\chi_{\Omega_t}\tau_{m_r}^\delta\left\|F\left(X_{t_{m_r}}^N\right)\right\| + \left\|W_A(r)-W_A\left(t_{m_r}\right)\right\| \\ &\le C(L_2)\left(\tau_{m_r}^\delta\right)^{\frac{\beta\wedge1\wedge\frac{d}{2}}{2}-\epsilon}\left(1 + \|X_0\|_\beta + \frac{\left\|W_A(r)-W_A\left(t_{m_r}\right)\right\|}{\left(\tau_{m_r}^\delta\right)^{\frac{\beta\wedge1}{2}}}\right), \end{aligned} \tag{A.10}$$

where $0 < \epsilon \ll 1$, and in the last step we have used $\frac{\tau^\delta_{m_r}}{\tau^\delta_{m_u}} \leq \frac{T\delta}{\delta \min\{T,\tau_{m_u}\}} \leq T(T^{-1} + \tau_{min}^{-1})$ due to (3.3) and Assumption 3.3. We deduce from Lemma 4.1 (*ii*) with $\rho = 0$, (2.6) and (A9) that

$$
\begin{aligned}
I_0(s) &\leq \chi_{\Omega_t} \int_0^s (s-r)^{-\frac{d}{4}} \left\| F\left(X^N_r\right) - F\left(X^N_{t_{m_r}}\right) \right\| \mathrm{d}r \\
&\leq C(L_1)\chi_{\Omega_t} \int_0^s (s-r)^{-\frac{d}{4}} \left(1 + \left\|X^N_r\right\|_E^2 + \left\|X^N_{t_{m_r}}\right\|_E^2\right) \left\|X^N_r - X^N_{t_{m_r}}\right\| \mathrm{d}r \\
&\leq C(L_1,L_2)\chi_{\Omega_t} \int_0^s (s-r)^{-\frac{d}{4}} \left(1 + \left\|X^N_{t_{m_r}}\right\|_E^2 + \left\|W^N_A(r)\right\|_E^2 + \left\|W^N_A\left(t_{m_r}\right)\right\|_E^2\right) \left\|X^N_r - X^N_{t_{m_r}}\right\| \mathrm{d}r,
\end{aligned}
$$

which together with (A.10) leads to

$$
\begin{aligned}
I_0(s) &\leq C(L_1,L_2)\chi_{\Omega_t} \int_0^s (s-r)^{-\frac{d}{4}} \left(1 + \left\|X^N_{t_{m_r}}\right\|_E^2 + \left\|W^N_A(r)\right\|_E^2 + \left\|W^N_A\left(t_{m_r}\right)\right\|_E^2\right) \\
&\quad \times \left(\tau^\delta_{m_r}\right)^{\frac{\beta\wedge 1\wedge \frac{d}{2}}{2}} \left(1 + \|X_0\|_\beta + \frac{\left\|W_A(r) - W_A\left(t_{m_r}\right)\right\|}{\left(\tau^\delta_{m_r}\right)^{\frac{\beta\wedge 1}{2}}}\right) \mathrm{d}r \\
&\leq C(L_1,L_2) \int_0^s (s-r)^{-\frac{d}{4}} \left(1 + \left\|W^N_A(r)\right\|_E^2 + \left\|W^N_A\left(t_{m_r}\right)\right\|_E^2\right) \left(1 + \|X_0\|_\beta + \frac{\left\|W_A(r) - W_A\left(t_{m_r}\right)\right\|}{\left(\tau^\delta_{m_r}\right)^{\frac{\beta\wedge 1}{2}}}\right) \mathrm{d}r,
\end{aligned}
$$

where we have used the definition of Ω_t and taken $\alpha \in \left(0, \frac{\beta\wedge 1\wedge \frac{d}{2}}{4}\right)$. Taking $L^{\tilde{p}}(\Omega)$-norm ($\tilde{p} \geq 2$) on both sides of the above equation, and applying the Minkowski inequality and the Hölder inequality, we derive

$$
\begin{aligned}
\|I_0(s)\|_{L^{\tilde{p}}(\Omega)} &\leq C(L_1,L_2) \int_0^s (s-r)^{-\frac{d}{4}} \left\| 1 + \left\|W^N_A(r)\right\|_E^2 + \left\|W^N_A\left(t_{m_r}\right)\right\|_E^2 \right\|_{L^{2\tilde{p}}(\Omega)} \\
&\quad \times \left\| 1 + \|X_0\|_\beta + \frac{\left\|W_A(r) - W_A\left(t_{m_r}\right)\right\|}{\left(\tau^\delta_{m_r}\right)^{\frac{\beta\wedge 1}{2}}} \right\|_{L^{2\tilde{p}}(\Omega)} \mathrm{d}r.
\end{aligned}
$$

Note that $\sup_{t\in[0,T]} \|A^{\frac{\beta\wedge 1}{2}} W_A(t)\|_{L^{2\tilde{p}}(\Omega;H)} \leq C$ and (2.2) imply the Hölder continuity of W_A

$$
\begin{aligned}
\left\|W_A(r) - W_A\left(t_{m_r}\right)\right\|_{L^{2\tilde{p}}(\Omega;H)} &\leq \left\|A^{\frac{\beta\wedge 1}{2}}\left(S\left(r - t_{m_r}\right) - \mathrm{Id}\right)\right\|_{\mathscr{L}(H)} \left\|W_A\left(t_{m_r}\right)\right\|_{L^{2\tilde{p}}(\Omega;\dot{H}^{\beta\wedge 1})} \\
&\quad + \left\| \int_{t_{m_r}}^r S\left(r - t_{m_r}\right) \mathrm{d}W(u) \right\|_{L^{2\tilde{p}}(\Omega;H)} \leq C\left(\tilde{p}, \left\|A^{\frac{\beta-1}{2}} Q^{\frac{1}{2}}\right\|_{\mathscr{L}_2(H)}\right) \left(\tau^\delta_{m_r}\right)^{\frac{\beta\wedge 1}{2}},
\end{aligned}
$$

which combining properties of the conditional expectation gives

$$\mathbb{E}\left[\left(\frac{\|W_A(r)-W_A(t_{m_r})\|}{\left(\tau^{\delta}_{m_r}\right)^{\frac{\beta\wedge 1}{2}}}\right)^{2\tilde{p}}\right]=\mathbb{E}\left[\mathbb{E}\left(\frac{\|W_A(r)-W_A(t_{m_r})\|^{2\tilde{p}}}{\left(\tau^{\delta}_{m_r}\right)^{\frac{\beta\wedge 1}{2}\times 2\tilde{p}}}\Big|\mathscr{F}_{t_{m_r}}\right)\right]$$

$$=\mathbb{E}\left[\left(\tau^{\delta}_{m_r}\right)^{-\frac{\beta\wedge 1}{2}\times 2\tilde{p}}\mathbb{E}\left(\left\|W_A(r)-W_A\left(t_{m_r}\right)\right\|^{2\tilde{p}}\Big|\mathscr{F}_{t_{m_r}}\right)\right]\le C\left(\tilde{p},\left\|A^{\frac{\beta-1}{2}}Q^{\frac{1}{2}}\right\|_{\mathscr{L}_2(H)}\right).$$

Similar as the proof of (2.12), we have

$$\sup_{0\le t\le T}\mathbb{E}\left[\left\|W^N_A(t)\right\|^{4\tilde{p}}_E\right]<\infty. \tag{A.11}$$

Therefore, we arrive at

$$\sup_{0\le s\le t}\|I_0(s)\|_{L^{\tilde{p}}(\Omega)}\le C\left(\tilde{p},T,L_1,L_2,\left\|A^{\frac{\beta-1}{2}}Q^{\frac{1}{2}}\right\|_{\mathscr{L}_2(H)}\right)\left(1+\|X_0\|_{L^{2\tilde{p}}(\Omega;\dot{H}^{\beta})}\right).$$

Estimate of I_1. Applying Lemma 4.1 (i) with $\rho=1-\frac{d}{4}-\epsilon=:\rho_0$ and (2.2) yields that for $0\le s\le t$,

$$\begin{aligned}I_1(s)&=\chi_{\Omega_t}\left\|\int_0^s A^{\rho_0}S^N(s-r)A^{-\rho_0}\left(S^N\left(r-t_{m_r}\right)-\mathrm{Id}\right)F^N\left(X^N_{t_{m_r}}\right)\mathrm{d}r\right\|_E\\&\le C(L_1)\chi_{\Omega_t}\int_0^s\left((s-r)^{-\rho_0}+(s-r)^{-\rho_0-\frac{d}{4}}\right)\left(\tau^{\delta}_{m_r}\right)^{\rho_0}\left(1+\left\|X^N_{t_{mr}}\right\|^3_E\right)\mathrm{d}r\\&\le C(L_1)\int_0^s\left((s-r)^{-\rho_0}+(s-r)^{-\rho_0-\frac{1}{4}}\right)\mathrm{d}r,\end{aligned}$$

where we have used the definition of Ω_t and taken $\alpha\in(0,\frac{\rho_0}{3}]$. Therefore, we get $\sup_{0\le s\le t}\|I_1(s)\|_{L^{\tilde{p}}(\Omega)}\le C(\tilde{p},T,L_1)$.

Combining estimates of I_0 and I_1, and taking parameter α in Ω_t as $\alpha=\alpha_0:=\left(\frac{\beta\wedge 1\wedge\frac{d}{2}}{4}\wedge\frac{4-d}{12}\right)-\epsilon$, we derive that

$$\begin{aligned}&\sup_{s\in[0,T]}\left[\left\|\chi_{\Omega_t}\left\|Z^N_s\right\|_E\right\|^3_{L^{3p}(\Omega)}+\left\|\chi_{\Omega_t}\left\|Z^N_s\right\|_E\right\|^{72}_{L^{72p}(\Omega)}\right]\\&\quad\le C\left(p,T,L_1,L_2,\left\|A^{\frac{\beta-1}{2}}Q^{\frac{1}{2}}\right\|_{\mathscr{L}_2(H)},\left\|P^NX_0\right\|_{L^{72p}(\Omega;E)},\|X_0\|_{L^{144p}(\Omega;\dot{H}^{\beta})}\right).\end{aligned}$$

Then (A.6) can be estimated as

$$\left\|\chi_{\Omega_t}\left\|\bar{X}^N_t\right\|_E\right\|_{L^p(\Omega)}\le C\left(p,T,L_1,L_2,\left\|A^{\frac{\beta-1}{2}}Q^{\frac{1}{2}}\right\|_{\mathscr{L}_2(H)},\left\|P^NX_0\right\|_{L^{72p}(\Omega;E)},\|X_0\|_{L^{144p}(\Omega;\dot{H}^{\beta})}\right), \tag{A.12}$$

which leads to

$$\begin{aligned}\left\|\chi_{\Omega_t}\left\|X^N_t\right\|_E\right\|_{L^p(\Omega)}&\le\left\|\chi_{\Omega_t}\left\|\bar{X}^N_t\right\|_E\right\|_{L^p(\Omega)}+\left\|\chi_{\Omega_t}\left\|Z^N_t\right\|_E\right\|_{L^p(\Omega)}\\&\le C\left(p,T,L_1,L_2,\left\|A^{\frac{\beta-1}{2}}Q^{\frac{1}{2}}\right\|_{\mathscr{L}_2(H)},\left\|P^NX_0\right\|_{L^{72p}(\Omega;E)},\|X_0\|_{L^{144p}(\Omega;\dot{H}^{\beta})}\right).\end{aligned}$$

Step 2: Proof of (A.4). Recall that in *Step 1* we take $\alpha = \alpha_0$. Note that for $0 \le t \le T$,

$$\Omega_t^c = \Omega_{t_{m_t}}^c \cup \left(\Omega_{t_{m_t}} \cap \left\{ \omega \in \Omega : \left(\tau_{m_t}^\delta\right)^{\alpha_0} \left\| X_{t_{m_t}}^N(\omega) \right\|_E > 1 \right\} \right).$$

Hence, by iteration,

$$\begin{aligned}
\chi_{\Omega_t^c}(\omega) &= \chi_{\Omega_{t_{m_t}}^c}(\omega) + \chi_{\Omega_{t_{m_t}}}(\omega) \cdot \chi_{\left\{ \left(\tau_{m_t}^\delta\right)^{\alpha_0} \left\| X_{t_{m_t}}^N \right\|_E > 1 \right\}}(\omega) \\
&= \sum_{j=0}^{m_t(\omega)} \chi_{\Omega_{t_{j-1}}}(\omega) \cdot \chi_{\left\{ \left(\tau_j^\delta\right)^{\alpha_0} \left\| X_{t_j}^N \right\|_E > 1 \right\}}(\omega),
\end{aligned}$$

where $\chi_{\Omega_{t_{-1}}} := 1$. Then for $0 \le t \le T$, combining the Hölder inequality yields

$$\begin{aligned}
\mathbb{E}\left[\chi_{\Omega_t^c} \left\| X_t^N \right\|_E^p \right] &= \mathbb{E}\left[\sum_{j=0}^{m_t} \left\| X_t^N \right\|_E^p \cdot \chi_{\Omega_{t_{j-1}}} \cdot \chi_{\left\{ \left(\tau_j^\delta\right)^{\alpha_0} \left\| X_{t_j}^N \right\|_E > 1 \right\}} \right] \\
&= \mathbb{E}\left[\sum_{j=0}^{m_t} \left(\chi_{\Omega_{t_{j-1}}} \cdot \chi_{\left\{ \left(\tau_j^\delta\right)^{\alpha_0} \left\| X_{t_j}^N \right\|_E > 1 \right\}} \tau_j^\delta \right) \frac{1}{\tau_j^\delta} \left\| X_t^N \right\|_E^p \right] \\
&\le \mathbb{E}\left[\sup_{0 \le t_j \le T} \left(\frac{1}{\tau_j^\delta} \right) \int_0^t \chi_{\Omega_{t_{m_s-1}}} \cdot \chi_{\left\{ \left(\tau_{m_s}^\delta\right)^{\alpha_0} \left\| X_{t_{m_s}}^N \right\|_E > 1 \right\}} \mathrm{d}s \left\| X_t^N \right\|_E^p \right] \\
&\le \left(\mathbb{E}\left[\sup_{0 \le t_j \le T} \left(\frac{1}{\tau_j^\delta} \right)^3 \right] \right)^{\frac{1}{3}} \left(\mathbb{E}\left[\int_0^t \chi_{\Omega_{t_{m_s-1}}} \cdot \chi_{\left\{ \left(\tau_{m_s}^\delta\right)^{\alpha_0} \left\| X_{t_{m_s}}^N \right\|_E > 1 \right\}} \mathrm{d}s \right] \right)^{\frac{1}{3}} \left(\mathbb{E}\left[\left\| X_t^N \right\|_E^{3p} \right] \right)^{\frac{1}{3}}.
\end{aligned}$$

It can be shown that

$$\left(\mathbb{E}\left[\sup_{0 \le t_j \le T} \left(\frac{1}{\tau_j^\delta} \right)^3 \right] \right)^{\frac{1}{3}} \le C(T, \tau_{min}) \delta^{-1}. \tag{A.13}$$

Notice that

$$\begin{aligned}
\mathbb{E}\left[\chi_{\Omega_{t_{m_s-1}}} \cdot \chi_{\left\{ \left(\tau_{m_s}^\delta\right)^{\alpha_0} \left\| X_{t_{m_s}}^N \right\|_E > 1 \right\}} \right] &= \mathbb{P}\left(\left\{ \omega \in \Omega_{t_{m_s-1}} : \left(\tau_{m_s}^\delta\right)^{\alpha_0} \left\| X_{t_{m_s}}^N \right\|_E > 1 \right\} \right) \\
&= \mathbb{P}\left(\left\{ \omega \in \Omega : \chi_{\Omega_{t_{m_s-1}}} \cdot \left(\tau_{m_s}^\delta\right)^{\alpha_0} \left\| X_{t_{m_s}}^N \right\|_E > 1 \right\} \right).
\end{aligned}$$

Combining the Chebyshev inequality, the Hölder inequality and Assumption 3.3 gives

$$\begin{aligned}
&\mathbb{P}\left(\left\{\omega\in\Omega:\chi_{\Omega_{t_{m_s-1}}}\cdot\left(\tau_{m_s}^{\delta}\right)^{\alpha_0}\left\|X_{t_{m_s}}^N\right\|_E>1\right\}\right) \\
&\leq \mathbb{E}\left[\left(\chi_{\Omega_{t_{m_s-1}}}\cdot\left(\tau_{m_s}^{\delta}\right)^{\alpha_0}\left\|X_{t_{m_s}}^N\right\|_E\right)^{\frac{3(p+1)}{\alpha_0}}\right] \\
&\leq \left(\mathbb{E}\left[\left(\tau_{m_s}^{\delta}\right)^{6(p+1)}\right]\right)^{\frac{1}{2}}\left(\mathbb{E}\left[\chi_{\Omega_{t_{m_s-1}}}\left\|X_{t_{m_s}}^N\right\|_E^{\frac{6(p+1)}{\alpha_0}}\right]\right)^{\frac{1}{2}} \\
&\leq C\left(p,T,\tau_{min},L_1,\left\|A^{\frac{\beta-1}{2}}Q^{\frac{1}{2}}\right\|_{\mathscr{L}_2(H)},\left\|P^NX_0\right\|_{L^q(\Omega;E)},\|X_0\|_{L^{2q}(\Omega;\dot{H}^\beta)}\right)\delta^{3(p+1)},
\end{aligned} \tag{A.14}$$

where we have utilized the fact that

$$\mathbb{E}\left[\chi_{\Omega_{t_{m_s-1}}}\left\|X_{t_{m_s}}^N\right\|_E^{\frac{6(p+1)}{\alpha_0}}\right]\leq C\left(p,T,\tau_{min},L_1,\left\|A^{\frac{\beta-1}{2}}Q^{\frac{1}{2}}\right\|_{\mathscr{L}_2(H)},\left\|P^NX_0\right\|_{L^q(\Omega;E)},\|X_0\|_{L^{2q}(\Omega;\dot{H}^\beta)}\right) \tag{A.15}$$

with $q=72\times\frac{6(p+1)}{\alpha_0}$. The proof of (A.15) is a combination of (A.3) and

$$\begin{aligned}
\chi_{\Omega_{t_{m_s-1}}}\left\|X_{t_{m_s}}^N\right\|_E &= \chi_{\Omega_{t_{m_s-1}}}\left\|S^N\left(\tau_{m_s-1}\right)\left(X_{t_{m_s-1}}^N+F^N\left(X_{t_{m_s-1}}^N\right)\tau_{m_s-1}+P^N\Delta W_{m_s-1}\right)\right\|_E \\
&\leq C\chi_{\Omega_{t_{m_s-1}}}\left(\left\|X_{t_{m_s-1}}^N\right\|_E+\tau_{m_s-1}^{1-\frac{d}{4}}\left\|F\left(X_{t_{m_s-1}}^N\right)\right\|+\left\|P^N\Delta W_{m_s-1}\right\|_E\right) \\
&\leq C\chi_{\Omega_{t_{m_s-1}}}\left(\left\|X_{t_{m_s-1}}^N\right\|_E+L_2+\left\|P^N\Delta W_{m_s-1}\right\|_E\right).
\end{aligned}$$

It remains to estimate $\left(\mathbb{E}\left[\,\left\|X_t^N\right\|_E^{3p}\,\right]\right)^{\frac{1}{3}}$. For $0\leq t_{m+1}\leq T$ and the fixed $\omega\in\Omega$, it follows from Lemma 4.1 (*ii*) with $\rho=0$ and the contractivity of the semigroup $\{S(t),t\geq 0\}$ in E that

$$\begin{aligned}
\|X_{t_{m+1}}^N\|_E &\leq \|S(\tau_m^\delta)X_{t_m}^N\|_E+\|S(\tau_m^\delta)F^N\left(X_{t_m}^N\right)\tau_m^\delta\|_E+\|S(\tau_m^\delta)P^N\Delta W_m\|_E \\
&\leq \|X_{t_m}^N\|_E+C(\tau_m^\delta)^{1-\frac{d}{4}}\left\|F\left(X_{t_m}^N\right)\right\|+\|W_A^N(t_{m+1})\|_E+\|W_A^N(t_m)\|_E \\
&\leq \left\|P^NX_0\right\|_E+C\sum_{i=0}^{m}\left((\tau_i^\delta)^{1-\frac{d}{4}}\left\|F\left(X_{t_i}^N\right)\right\|+\|W_A^N(t_{i+1})\|_E+\|W_A^N(t_i)\|_E\right).
\end{aligned}$$

Similarly,

$$\left\|X_t^N\right\|_E\leq\|X_{t_{m_t}}^N\|_E+C\left(t-t_{m_t}\right)^{1-\frac{d}{4}}\left\|F\left(X_{t_{m_t}}^N\right)\right\|+\|W_A^N(t)\|_E+\|W_A^N(t_{m_t})\|_E.$$

Hence, by Assumption 3.1, we obtain

$$\left\|X_t^N\right\|_E\leq\left\|P^NX_0\right\|_E+C\sum_{i=0}^{m_t}\left((\tau_i^\delta)^{1-\frac{d}{4}}\left\|F\left(X_{t_i}^N\right)\right\|+\|W_A^N(t)\|_E+\|W_A^N(t_i)\|_E\right).$$

This, together with (A.11) implies that for $0 \le t \le T$,

$$\left\|X_t^N\right\|_{L^p(\Omega;E)} \le \left\|P^N X_0\right\|_{L^p(\Omega;E)} + C\int_0^T \left\|\left(\tau_{m_s}^\delta\right)^{-1}\right\|_{L^{2p}(\Omega)} (L_2 + \|W_A^N(s)\|_{L^{2p}(\Omega;E)} + \|W_A^N(t_{m_s})\|_{L^{2p}(\Omega;E)})\,\mathrm{d}s.$$

It can be deduced from $(\tau^\delta)^{-1} \le \delta^{-1}(\tau_{min} + T)^{-1}$ and (A.11) that

$$\mathbb{E}\left[\left\|X_t^N\right\|_E^{3p}\right] \le C\left(p, T, \tau_{min}, L_2, \|A^{\frac{\beta-1}{2}} Q^{\frac{1}{2}}\|_{\mathscr{L}_2(H)}\right)\left(\left\|P^N X_0\right\|_{L^{3p}(\Omega;E)}^{3p} + \delta^{-3p}\right). \tag{A.16}$$

It follows from (A.13), (A.14) and (A.16) that

$$\mathbb{E}\left[\chi_{\Omega_t^c}\left\|X_t^N\right\|_E^p\right] \le C\delta^{-1}\delta^{p+1}\left(1 + \left\|P^N X_0\right\|_{L^{3p}(\Omega;E)}^p + \delta^{-p}\right) \le C$$

with $C > 0$ depending on $p, T, L_1, L_2, \tau_{min}, \|A^{\frac{\beta-1}{2}} Q^{\frac{1}{2}}\|_{\mathscr{L}_2(H)}, \left\|P^N X_0\right\|_{L^q(\Omega;E)}, \|X_0\|_{L^{2q}(\Omega;\dot{H}^\beta)}$.

Therefore, combining *Step 1* and *Step 2*, we get

$$\sup_{0\le t\le T} \mathbb{E}\left[\left\|X_t^N\right\|_E^p\right] \le C\left(p, T, L_1, L_2, \tau_{min}, \|A^{\frac{\beta-1}{2}} Q^{\frac{1}{2}}\|_{\mathscr{L}_2(H)}, \|X_0\|_{L^{2q}(\Omega;\dot{H}^\beta)}^q, \|P^N X_0\|_{L^q(\Omega;E)}^q\right)$$

for some large $q := q(p, \beta, d) > 0$, which finishes the proof. □

Proof of Proposition 4.4. For the proof of the regularity (4.3), when $\beta \in (0, 2)$, applying (2.1) and (2.3) yields that

$$\sup_{0\le t\le T} \|X_t^N\|_{L^p(\Omega;\dot{H}^\beta)} \le C(1 + \|X_0\|_{L^p(\Omega;\dot{H}^\beta)} + \sup_{0\le s\le T} \|X_s^N\|_{L^{3p}(\Omega;E)}^3 + \|A^{\frac{\beta-1}{2}} Q^{\frac{1}{2}}\|_{\mathscr{L}_2(H)}).$$

And when $\beta = 2$, it follows from (2.1), (2.9) and the Hölder inequality that

$$\begin{aligned}\left\|X_t^N\right\|_{L^p(\Omega;\dot{H}^2)} &\le \|X_0\|_{L^p(\Omega;\dot{H}^2)} + C\int_0^t (t-s)^{-\frac{1}{2}}\left\|\left\|X_{t_{m_s}}^N\right\|_1\left(1 + \left\|X_{t_{m_s}}^N\right\|_E^2\right)\right\|_{L^p(\Omega)} \mathrm{d}s + C\|A^{\frac{1}{2}} Q^{\frac{1}{2}}\|_{\mathscr{L}_2(H)}\\ &\le C(1 + \|X_0\|_{L^p(\Omega;\dot{H}^2)} + \sup_{0\le s\le T} \|X_s^N\|_{L^{2p}(\Omega;\dot{H}^1)} + \sup_{0\le s\le T} \|X_s^N\|_{L^{4p}(\Omega;E)}^2).\end{aligned}$$

These, combining Proposition 4.3 finish the proof of (4.3).

For the proof of the Hölder continuity (4.4), when $\gamma \in (0, \beta]$ with $\beta \in (0, 2)$, or $\gamma \in (0, \beta)$ with $\beta = 2$, we can deduce from (2.1)–(2.3), Proposition 4.3 and (4.3) that

$$\begin{aligned}\|X_t^N - X_s^N\|_{L^p(\Omega;\dot{H}^\gamma)} &\le \|A^{\frac{\gamma-\beta}{2}}(S^N(t-s) - \mathrm{Id})A^{\frac{\beta}{2}} X_s^N\|_{L^p(\Omega;H)} + \left\|\int_s^t A^{\frac{\gamma}{2}} S^N\left(t - t_{m_s}\right) F^N\left(X_{t_{m_s}}^N\right) \mathrm{d}s\right\|_{L^p(\Omega;H)}\\ &\quad + \left\|\int_s^t A^{\frac{1-\beta+\gamma}{2}} S^N\left(t - t_{m_s}\right) A^{\frac{\beta-1}{2}}\, \mathrm{d}W(s)\right\|_{L^p(\Omega;H)}\\ &\le C(t-s)^{\frac{\beta-\gamma}{2}}\|X_s^N\|_{L^p(\Omega;\dot{H}^\beta)} + C\int_s^t \left(t - t_{m_s}\right)^{-\frac{\gamma}{2}}\left(1 + \left\|X_{t_{m_s}}^N\right\|_{L^{3p}(\Omega;E)}^3\right) \mathrm{d}s\\ &\quad + C\|A^{\frac{\beta-1}{2}} Q^{\frac{1}{2}}\|_{\mathscr{L}_2(H)}(t-s)^{\frac{(\beta-\gamma)\wedge 1}{2}} \le C(t-s)^{\frac{(\beta-\gamma)\wedge 1}{2}}.\end{aligned}$$

And when $\gamma = \beta = 2$, it follows from (4.3) that $\|X_t^N - X_s^N\|_{L^p(\Omega;\dot{H}^2)} \le C$. The proof is finished. □

Appendix B.

Proof. Proof of the convergence of (CAU 1).Based on the auxiliary process Y_t^N, the error can be divided into the following terms,

$$\|X(t)-X_t^{N,(1)}\|_{L^p(\Omega;H)} \le \|X(t)-P^N X(t)\|_{L^p(\Omega;H)} + \|P^N X(t)-Y_t^N\|_{L^p(\Omega;H)} + \|Y_t^N - X_t^{N,(1)}\|_{L^p(\Omega;H)},$$

where terms $\|X(t)-P^N X(t)\|_{L^p(\Omega;H)}$ and $\|P^N X(t)-Y_t^N\|_{L^p(\Omega;H)}$ have been estimated; see (5.1)–(5.2).

For the estimate of the term $\|Y_t^N - X_t^{N,(1)}\|_{L^p(\Omega;H)}$, combining the differential form

$$\begin{aligned} \mathrm{d}\left(X^{N,(1)} - Y_t^N\right) = -A^N\left(X^{N,(1)} - Y_t^N\right)\mathrm{d}t + S^N\left(t-t_{m_t}\right)\Bigg(F^N\left(X_{t_{m_t}}^{N,(1)}\right)\chi_{\{\tau_{m_t}^\delta \ge \tau_{min}^\delta\}} \\ + \frac{F^N\left(X_{t_{m_t}}^{N,(1)}\right)}{1+\left\|F^N\left(X_{t_{m_t}}^{N,(1)}\right)\right\|\tau_{min}^\delta}\chi_{\{\tau_{m_t}^\delta < \tau_{min}^\delta\}}\Bigg)\mathrm{d}t - F^N(X(t))\,\mathrm{d}t \end{aligned}$$

with $\tau_{min}^\delta = \delta\tau_{min}$, and applying the Taylor formula give

$$\begin{aligned} \left\|X_t^{N,(1)} - Y_t^N\right\|^2 &= 2\int_0^t \left\langle X_s^{N,(1)} - Y_s^N,\ -A^N\left(X_s^{N,(1)} - Y_s^N\right)\right\rangle \mathrm{d}s \\ &\quad + 2\int_0^t \left\langle X_s^{N,(1)} - Y_s^N,\ S^N\left(s-t_{m_s}\right)F^N\left(X_{t_{m_s}}^{N,(1)}\right) - F^N(X(s))\right\rangle \chi_{\{\tau_{m_s}^\delta \ge \tau_{min}^\delta\}}\,\mathrm{d}s \\ &\quad + 2\int_0^t \left\langle X_s^{N,(1)} - Y_s^N,\ S^N\left(s-t_{m_s}\right)\frac{F^N\left(X_{t_{m_s}}^{N,(1)}\right)}{1+\left\|F^N\left(X_{t_{m_s}}^{N,(1)}\right)\right\|\tau_{min}^\delta} - F^N(X(s))\right\rangle \chi_{\{\tau_{m_s}^\delta < \tau_{min}^\delta\}}\,\mathrm{d}s \\ &=:\ 2\int_0^t \langle X_s^{N,(1)} - Y_s^N,\ -A^N(X_s^{N,(1)} - Y_s^N)\rangle\,\mathrm{d}s + 2K_1(t) + 2K_2(t). \end{aligned}$$

For the estimate of the term K_1, by (2.5) and the Young inequality, we have

$$\begin{aligned} K_1(t) &= \int_0^t \left\langle X_s^{N,(1)} - Y_s^N,\ \left(S^N\left(s-t_{m_s}\right) - \mathrm{Id}\right)F^N\left(X_{t_{m_s}}^{N,(1)}\right)\right\rangle \chi_{\{\tau_{m_s}^\delta \ge \tau_{min}^\delta\}}\,\mathrm{d}s \\ &\quad + \int_0^t \left\langle X_s^{N,(1)} - Y_s^N,\ F^N\left(X_{t_{m_s}}^{N,(1)}\right) - F^N(Y_s^N)\right\rangle \chi_{\{\tau_{m_s}^\delta \ge \tau_{min}^\delta\}}\,\mathrm{d}s \\ &\quad + \int_0^t \langle X_s^{N,(1)} - Y_s^N,\ F^N(Y_s^N) - F^N(X(s))\rangle \chi_{\{\tau_{m_s}^\delta \ge \tau_{min}^\delta\}}\,\mathrm{d}s \\ &\le \left(L_0 + \frac{1}{2}\right)\int_0^t \|X_s^{N,(1)} - Y_s^N\|^2\,\mathrm{d}s + \frac{1}{2}\int_0^t \|F^N(Y_s^N) - F^N(X(s))\|^2\,\mathrm{d}s + K_{1,1}(t) + K_{1,2}(t), \end{aligned}$$

where

$$K_{1,1}(t) := \int_0^t \langle X_s^{N,(1)} - Y_s^N,\ \left(S^N\left(s-t_{m_s}\right) - \mathrm{Id}\right) F^N\left(X_{t_{m_s}}^{N,(1)}\right)\rangle \chi_{\{\tau_{m_s}^\delta \geq \tau_{min}^\delta\}}\,\mathrm{d}s,$$

$$K_{1,2}(t) := \int_0^t \langle X_s^{N,(1)} - Y_s^N,\ F^N\left(X_{t_{m_s}}^{N,(1)}\right) - F^N(X_s^{N,(1)})\rangle \chi_{\{\tau_{m_s}^\delta \geq \tau_{min}^\delta\}}\,\mathrm{d}s.$$

We first show the estimate of the term $K_{1,1}$. For the case $\beta = 1$ or $\beta = 2$, similar to the estimate of $J_{1,1}$, the term $K_{1,1}$ can be estimated as $\|K_{1,1}(t)\|_{L^p(\Omega)} \leq C\int_0^t \|X_r^{N,(1)} - Y_r^N\|^2_{L^{2p}(\Omega;H)}\,\mathrm{d}r + C\delta^\beta,\ t\in[0,T]$. The proof is omitted.

For the case $\beta \in (0,1)\cup(1,2)$, we need to further split the term $K_{1,1}$ into two parts, which are denoted by $K_{1,1,i}, i = 1,2$, based on

$$X_s^{N,(1)} - Y_s^N = \int_0^s \left(S^N\left(s-t_{m_r}\right) F^N\left(X_{t_{m_r}}^{N,(1)}\right) - S^N(s-r)F^N(X(r))\right) \chi_{\{\tau_{m_r}^\delta \geq \tau_{min}^\delta\}}\,\mathrm{d}r$$
$$+ \int_0^s \left(S^N\left(s-t_{m_r}\right) \frac{F^N\left(X_{t_{m_r}}^{N,(1)}\right)}{1+\left\|F^N\left(X_{t_{m_r}}^{N,(1)}\right)\right\|\tau_{min}^\delta} - S^N(s-r)F^N(X(r))\right) \chi_{\{\tau_{m_r}^\delta < \tau_{min}^\delta\}}\,\mathrm{d}r.$$

Namely,

$$K_{1,1,1}(t) := \int_0^t \left\langle \int_0^s \left(S^N\left(s-t_{m_r}\right) F^N\left(X_{t_{m_r}}^{N,(1)}\right) - S^N(s-r)F^N(X(r))\right) \chi_{\{\tau_{m_r}^\delta \geq \tau_{min}^\delta\}}\,\mathrm{d}r,\right.$$
$$\left.\left(S^N\left(s-t_{m_s}\right) - \mathrm{Id}\right) F^N\left(X_{t_{m_s}}^{N,(1)}\right)\right\rangle \chi_{\{\tau_{m_s}^\delta \geq \tau_{min}^\delta\}}\,\mathrm{d}s,$$

$$K_{1,1,2}(t) := \int_0^t \left\langle \int_0^s \left(S^N\left(s-t_{m_r}\right) \frac{F^N\left(X_{t_{m_r}}^{N,(1)}\right)}{1+\left\|F^N\left(X_{t_{m_r}}^{N,(1)}\right)\right\|\tau_{min}^\delta} - S^N(s-r)F^N(X(r))\right) \chi_{\{\tau_{m_r}^\delta < \tau_{min}^\delta\}}\,\mathrm{d}r,\right.$$
$$\left.\left(S^N\left(s-t_{m_s}\right) - \mathrm{Id}\right) F^N\left(X_{t_{m_s}}^{N,(1)}\right)\right\rangle \chi_{\{\tau_{m_s}^\delta \geq \tau_{min}^\delta\}}\,\mathrm{d}s.$$

For the term $K_{1,1,1}$, it can be further split as

$$K_{1,1,1}(t) = \int_0^t \left\langle \int_0^s S^N(s-r)(S^N\left(r-t_{m_r}\right) - \mathrm{Id})F^N\left(X_{t_{m_r}}^{N,(1)}\right) \chi_{\{\tau_{m_r}^\delta \geq \tau_{min}^\delta\}}\,\mathrm{d}r,\right.$$
$$\left.\left(S^N\left(s-t_{m_s}\right) - \mathrm{Id}\right) F^N\left(X_{t_{m_s}}^{N,(1)}\right)\right\rangle \chi_{\{\tau_{m_s}^\delta \geq \tau_{min}^\delta\}}\,\mathrm{d}s$$
$$+ \int_0^t \left\langle \int_0^s S^N(s-r)\left(F^N\left(X_{t_{m_r}}^{N,(1)}\right) - F^N\left(X_r^{N,(1)}\right)\right) \chi_{\{\tau_{m_r}^\delta \geq \tau_{min}^\delta\}}\,\mathrm{d}r,\right.$$
$$\left.\left(S^N\left(s-t_{m_s}\right) - \mathrm{Id}\right) F^N\left(X_{t_{m_s}}^{N,(1)}\right)\right\rangle \chi_{\{\tau_{m_s}^\delta \geq \tau_{min}^\delta\}}\,\mathrm{d}s$$
$$+ \int_0^t \left\langle \int_0^s S^N(s-r)\left(F^N\left(X_r^{N,(1)}\right) - F^N(X(r))\right) \chi_{\{\tau_{m_r}^\delta \geq \tau_{min}^\delta\}}\,\mathrm{d}r,\right.$$
$$\left.\left(S^N\left(s-t_{m_s}\right) - \mathrm{Id}\right) F^N\left(X_{t_{m_s}}^{N,(1)}\right)\right\rangle \chi_{\{\tau_{m_s}^\delta \geq \tau_{min}^\delta\}}\,\mathrm{d}s =: K_{1,1,1}^1(t) + K_{1,1,1}^2(t) + K_{1,1,1}^3(t).$$

Similar to estimates of $J_{1,1}^1$ and $J_{1,1}^3$, terms $K_{1,1,1}^1$ and $K_{1,1,1}^3$ can be estimated as $\|K_{1,1,1}^1(t)\|_{L^p(\Omega)} + \|K_{1,1,1}^3(t)\|_{L^p(\Omega)} \leq C\int_0^t \|X_r^{N,(1)} - Y_r^N\|_{L^{2p}(\Omega;H)}^2 \,\mathrm{d}r + C\lambda_N^{-\beta} + C\delta^\beta$ for $\beta \in (0,1)\cup(1,2)$ and $t \in [0,T]$. The proof is omitted.

For the term $K_{1,1,1}^2$, when $\beta \in (0,1)$, the estimate is still similar as $J_{1,1}^2$ and $\|K_{1,1,1}^2(t)\|_{L^p(\Omega)} \leq C\delta^\beta$. And when $\beta \in (1,2)$, by the Taylor formula, and the fact that

$$
\begin{aligned}
&K_{1,1,1}^2(t)\\
&= -\int_0^t \left\langle \int_0^s A^{\frac{\beta-1}{2}} S^N(s-r) DF\left(X_{t_{m_r}}^{N,(1)}\right)\left(S^N\left(r-t_{m_r}\right)-\mathrm{Id}\right) X_{t_{m_r}}^{N,(1)} \chi_{\{\tau_{m_r}^\delta \geq \tau_{min}^\delta\}}\,\mathrm{d}r,\right.\\
&\quad \left. A^{-\frac{\beta}{2}}\left(S^N\left(s-t_{m_s}\right)-\mathrm{Id}\right) A^{\frac{1}{2}} F^N\left(X_{t_{m_s}}^{N,(1)}\right)\right\rangle \chi_{\{\tau_{m_s}^\delta \geq \tau_{min}^\delta\}}\,\mathrm{d}s\\
&\quad -\int_0^t \left\langle \int_0^s A^{\frac{\beta-1}{2}} S^N(s-r) DF\left(X_{t_{m_r}}^{N,(1)}\right)\int_{t_{m_r}}^r S^N\left(r-t_{m_r}\right)\left(F^N\left(X_{t_{m_u}}^{N,(1)}\right)\chi_{\{\tau_{m_u}^\delta \geq \tau_{min}^\delta\}}\right.\right.\\
&\quad \left.\left. +\frac{F^N\left(X_{t_{m_u}}^{N,(1)}\right)}{1+\left\|F^N\left(X_{t_{m_u}}^{N,(1)}\right)\right\|\tau_{min}^\delta}\chi_{\{\tau_{m_u}^\delta < \tau_{min}^\delta\}}\right)\mathrm{d}u\chi_{\{\tau_{m_r}^\delta \geq \tau_{min}^\delta\}}\,\mathrm{d}r,\ A^{-\frac{\beta}{2}}\left(S^N\left(s-t_{m_s}\right)-\mathrm{Id}\right)A^{\frac{1}{2}}F^N\left(X_{t_{m_s}}^{N,(1)}\right)\right\rangle \chi_{\{\tau_{m_s}^\delta \geq \tau_{min}^\delta\}}\,\mathrm{d}s\\
&\quad -\int_0^t \left\langle \int_0^s A^{\frac{\beta-1}{2}} S^N(s-r) DF\left(X_{t_{m_r}}^{N,(1)}\right)\int_{t_{m_r}}^r S^N\left(r-t_{m_r}\right)\mathrm{d}W(u)\chi_{\{\tau_{m_r}^\delta \geq \tau_{min}^\delta\}}\,\mathrm{d}r,\right.\\
&\quad \left. A^{-\frac{\beta}{2}}\left(S^N\left(s-t_{m_s}\right)-\mathrm{Id}\right) A^{\frac{1}{2}} F^N\left(X_{t_{m_s}}^{N,(1)}\right)\right\rangle \chi_{\{\tau_{m_s}^\delta \geq \tau_{min}^\delta\}}\,\mathrm{d}s\\
&\quad -\int_0^t \left\langle \int_0^s S^N(s-r) R_F\left(X_{t_{m_r}}^{N,(1)}, X_r^{N,(1)}\right)\chi_{\{\tau_{m_r}^\delta \geq \tau_{min}^\delta\}}\,\mathrm{d}r,\ \left(S^N\left(s-t_{m_s}\right)-\mathrm{Id}\right)F^N\left(X_{t_{m_s}}^{N,(1)}\right)\right\rangle \chi_{\{\tau_{m_s}^\delta \geq \tau_{min}^\delta\}}\,\mathrm{d}s\\
&=: \widetilde{II}_1(t)+\widetilde{II}_2(t)+\widetilde{II}_3(t)+\widetilde{II}_4(t).
\end{aligned}
$$

Proofs of terms $\widetilde{II}_1, \widetilde{II}_3, \widetilde{II}_4$ are similar to those of II_1, II_3, II_4 and $\|\widetilde{II}_1(t)\|_{L^p(\Omega)} + \|\widetilde{II}_3(t)\|_{L^p(\Omega)} + \|\widetilde{II}_4(t)\|_{L^p(\Omega)} \leq C\delta^\beta$ for $\beta \in (1,2)$. For the term $\widetilde{II}_2$, it follows from (2.1)–(2.2), (2.7) and Corollary 4.5 that

$$
\begin{aligned}
\|\widetilde{II}_2(t)\|_{L^p(\Omega)} &\leq C\delta^{1+\frac{\beta}{2}}\int_0^t\int_0^s (s-r)^{-\frac{\beta-1}{2}}\left\|(1+\|X_{t_{m_r}}^{N,(1)}\|_E^2)\|F\left(X_{t_{m_r}}^{N,(1)}\right)\|\chi_{\{\tau_{m_r}^\delta \geq \tau_{min}^\delta\}}\right\|_{L^{2p}(\Omega)}\\
&\quad \times \|F\left(X_{t_{m_s}}^{N,(1)}\right)\|_{L^{2p}(\Omega;\dot{H}^1)}\,\mathrm{d}r\,\mathrm{d}s \leq C\delta^\beta.
\end{aligned}
$$

Hence, we get that for $\beta \in (0,1)\cup(1,2)$, $\|K_{1,1,1}^2(t)\|_{L^p(\Omega)} \leq C\delta^\beta,\ t \in [0,T]$.

Similarly, the term $K_{1,1,2}$ can be split as

$$K_{1,1,2}(t) = \int_0^t \left\langle \int_0^s S^N(s-r)\left(S^N\left(r-t_{m_r}\right)-\mathrm{Id}\right)\frac{F^N\left(X^{N,(1)}_{t_{m_r}}\right)}{1+\|F^N\left(X^{N,(1)}_{t_{m_r}}\right)\|\tau^\delta_{min}}\chi_{\{\tau^\delta_{m_r}<\tau^\delta_{min}\}}\,\mathrm{d}r,\right.$$

$$\left.\left(S^N\left(s-t_{m_s}\right)-\mathrm{Id}\right)F^N\left(X^{N,(1)}_{t_{m_s}}\right)\right\rangle\chi_{\{\tau^\delta_{m_s}\geq\tau^\delta_{min}\}}\,\mathrm{d}s$$

$$+\int_0^t\left\langle\int_0^s S^N(s-r)\left(F^N\left(X^{N,(1)}_{t_{m_r}}\right)-F^N(X(r))\right)\chi_{\{\tau^\delta_{m_r}<\tau^\delta_{min}\}}\,\mathrm{d}r,\right.$$

$$\left.\left(S^N\left(s-t_{m_s}\right)-\mathrm{Id}\right)F^N\left(X^{N,(1)}_{t_{m_s}}\right)\right\rangle\chi_{\{\tau^\delta_{m_s}\geq\tau^\delta_{min}\}}\,\mathrm{d}s$$

$$-\int_0^t\left\langle\int_0^s S^N(s-r)\frac{\left\|F^N\left(X^{N,(1)}_{t_{m_r}}\right)\right\|\tau^\delta_{min}F^N\left(X^{N,(1)}_{t_{m_r}}\right)}{1+\left\|F^N\left(X^{N,(1)}_{t_{m_r}}\right)\right\|\tau^\delta_{min}}\chi_{\{\tau^\delta_{m_r}<\tau_{min}\}}\,\mathrm{d}r,\right.$$

$$\left.\left(S^N\left(s-t_{m_s}\right)-\mathrm{Id}\right)F^N\left(X^{N,(1)}_{t_{m_s}}\right)\right\rangle\chi_{\{\tau^\delta_{m_s}\geq\tau^\delta_{min}\}}\,\mathrm{d}s =: K^1_{1,1,2}(t)+K^2_{1,1,2}(t)+K^3_{1,1,2}(t).$$

Terms $K^1_{1,1,2}, K^2_{1,1,2}$ can be estimated similarly to terms $K^1_{1,1,1}$ and $K^2_{1,1,1}+K^3_{1,1,1}$, respectively, and we can get $\|K^1_{1,1,2}(t)\|_{L^p(\Omega)}+\|K^2_{1,1,2}(t)\|_{L^p(\Omega)}\leq C\int_0^t\|X^{N,(1)}_s-Y^N_s\|^2_{L^{2p}(\Omega;H)}\,\mathrm{d}s+C\lambda_N^{-\beta}+C\delta^\beta$ for $\beta\in(0,1)\cup(1,2)$. Then we show the estimate of the term $K^3_{1,1,2}$. By (2.1)–(2.2) and the Young inequality, we arrive at for $\beta\in(0,2)$,

$$K^3_{1,1,2}(t) = -\int_0^t\left\langle\int_0^s A^{\frac{\beta}{2}}S^N(s-r)\frac{\left\|F^N\left(X^{N,(1)}_{t_{m_r}}\right)\right\|\tau^\delta_{min}F^N\left(X^{N,(1)}_{t_{m_r}}\right)}{1+\left\|F^N\left(X^{N,(1)}_{t_{m_r}}\right)\right\|\tau^\delta_{min}}\chi_{\{\tau^\delta_{m_r}<\tau^\delta_{min}\}}\,\mathrm{d}r,\right.$$

$$\left.A^{-\frac{\beta}{2}}\left(S^N\left(s-t_{m_s}\right)-\mathrm{Id}\right)F^N\left(X^{N,(1)}_{t_{m_s}}\right)\right\rangle\chi_{\{\tau^\delta_{m_s}\geq\tau^\delta_{min}\}}\,\mathrm{d}s$$

$$\leq C\int_0^t\int_0^s(s-r)^{-\frac{\beta}{2}}\left[\left\|F\left(X^{N,(1)}_{t_{m_r}}\right)\right\|^4(\tau^\delta_{min})^2\chi_{\{\tau^\delta_{m_r}<\tau^\delta_{min}\}}+\left(s-t_{m_s}\right)^\beta\left\|F\left(X^{N,(1)}_{t_{m_s}}\right)\right\|^2\chi_{\{\tau^\delta_{m_s}\geq\tau^\delta_{min}\}}\right]\mathrm{d}r\,\mathrm{d}s,$$

which gives $\|K^3_{1,1,2}(t)\|_{L^p(\Omega)}\leq C\delta^\beta,\ t\in[0,T]$.

Hence, we derive $\|K_{1,1}(t)\|_{L^p(\Omega)}\leq C\int_0^t\|X^{N,(1)}_r-Y^N_r\|^2_{L^{2p}(\Omega;H)}\,\mathrm{d}r+C\lambda_N^{-\beta}+C\delta^\beta,\ t\in[0,T]$ for $\beta\in(0,2]$.

The estimate of $K_{1,2}$ is similar to that of $J_{1,2}$, and we can obtain $\|K_{1,2}(t)\|_{L^p(\Omega)}\leq C\int_0^t\|X^{N,(1)}_s-Y^N_s\|^2_{L^{2p}(\Omega;H)}\,\mathrm{d}s+C\lambda_N^{-\beta}+C\delta^\delta$ for $\beta\in(0,2]$ and $t\in[0,T]$. The proof is omitted.

Based on Corollary 4.5, similarly to the estimate of the term $K_{1,1,2}$ and Wang (2020, Theorem 4.12), the term K_2 can be estimated as $\|K_2(t)\|_{L^p(\Omega)}\leq C\int_0^t\|X^{N,(1)}_s-Y^N_s\|^2_{L^{2p}(\Omega;H)}\,\mathrm{d}s+C\lambda_N^{-\beta}+C\delta^\delta$ for $\beta\in(0,2]$ and $t\in[0,T]$. The proof is omitted.

Therefore, combining estimates of terms K_1 and K_2 yields

$$\left\| X_t^{N,(1)} - Y_t^N \right\|_{L^{2p}(\Omega;H)}^2 \leq C \left(\int_0^t \left\| X_t^{N,(1)} - Y_s^N \right\|_{L^{2p}(\Omega;H)}^2 \mathrm{d}s + \lambda_N^{-\beta} + \delta^\beta \right),$$

where $p \geq 1$ and the constant $C > 0$ depends on $p, T, \tau_{min}, L_0, L_1, L_2, X_0$ and $\|A^{\frac{\beta-1}{2}} Q^{\frac{1}{2}}\|_{\mathscr{L}_2(H)}$. Recalling that $\lambda_N \sim N^{-\frac{2}{d}}$, and applying the Grönwall inequality, we obtain

$$\sup_{0 \leq t \leq T} \left\| X_t^{N,(1)} - Y_t^N \right\|_{L^{2p}(\Omega;H)} \leq C \left(N^{-\frac{\beta}{d}} + \delta^{\frac{\beta}{2}} \right), \ p \geq 1,$$

which combining (5.1)–(5.2) finishes the proof of (CAU 1). □