

Stoch PDE: Anal Comp manuscript No.
(will be inserted by the editor)

Strong Approximation of Monotone Stochastic Partial Differential Equations Driven by Multiplicative Noise

Zhihui Liu · Zhonghua Qiao

Received: / Accepted: date

Abstract We establish a general theory of optimal strong error estimation for numerical approximations of a second-order parabolic stochastic partial differential equation with monotone drift driven by a multiplicative infinite-dimensional Wiener process. The equation is spatially discretized by Galerkin methods and temporally discretized by drift-implicit Euler and Milstein schemes. By the monotone and Lyapunov assumptions, we use both the variational and semigroup approaches to derive a spatial Sobolev regularity under the $L_\omega^p L_t^\infty \dot{H}^{1+\gamma}$ -norm and a temporal Hölder regularity under the $L_\omega^p L_x^2$ -norm for the solution of the proposed equation with an $\dot{H}^{1+\gamma}$ -valued initial datum for $\gamma \in [0, 1]$. Then we make full use of the monotonicity of the equation and tools from stochastic calculus to derive the sharp strong convergence rates $\mathcal{O}(h^{1+\gamma} + \tau^{1/2})$ and $\mathcal{O}(h^{1+\gamma} + \tau^{(1+\gamma)/2})$ for the Galerkin-based Euler and Milstein schemes, respectively.

Keywords monotone stochastic partial differential equation · stochastic Allen–Cahn equation · Galerkin finite element method · Euler scheme · Milstein scheme

Mathematics Subject Classification (2010) 65M60 · 60H15 · 60H35

Zhihui Liu
Department of Mathematics,
The Hong Kong University of Science and Technology,
Clear Water Bay, Kowloon, Hong Kong
E-mail: zhliu@ust.hk

Zhonghua Qiao
Department of Applied Mathematics,
The Hong Kong Polytechnic University,
Hung Hom, Kowloon, Hong Kong
E-mail: zhonghua.qiao@polyu.edu.hk

1 Introduction

There exists a general theory of strong error estimation for numerical approximations of stochastic partial differential equations (SPDEs) with Lipschitz coefficients; see, e.g., [5], [17], [20], [21] and references therein. For SPDEs with non-Lipschitz coefficients driven by multiplicative noises, there are only a few results on strong approximations. Since these SPDEs cannot be solved explicitly, one might need to develop efficient numerical techniques to study them. Depending on particular physical models, it may be necessary to design numerical schemes for the underlying SPDEs with strong convergence rates, i.e., rates in $L^p(\Omega)$ concerning the sample variables for some $p \geq 1$, see, e.g., [1] and references quoted therein.

To deal with the nonlinearity of SPDEs, one usually applies the truncation technique, which only produces convergence rate in certain sense such as in probability or pathwise which is weaker than strong sense, see, e.g., [18] for the stochastic Allen–Cahn equation and [23] for the stochastic nonlinear Schrödinger equation. A few recent contributions on strong error estimations for numerical approximations of particular types of SPDEs with non-Lipschitz coefficients by using the so-called exponential integrability for both the exact and numerical solutions. However, such exponential integrability is only shown to hold for several special SPDEs. In fact, up to our knowledge, only the two-dimensional stochastic Navier–Stokes equation [9], the one-dimensional stochastic Burgers equation and Cahn–Hilliard–Cook equation [15], and the one-dimensional stochastic nonlinear Schrödinger equation [7], [8] were proved to possess this type of exponential integrability.

1.1 Stochastic Allen–Cahn Equation

Currently, several works are performed to give strong error estimations for semidiscrete or fully discrete schemes of the stochastic Allen–Cahn equation (or stochastic Ginzburg–Landau equation) driven by an additive white or colored noises:

$$du = (\Delta u + u - u^3)dt + dW(t), \quad u(0) = u_0. \quad (1.1)$$

A major feature of the nonlinear drift term in Eq. (1.1) is that the following monotone (more precisely, one-sided Lipschitz) condition holds:

$$[(x - x^3) - (y - y^3)](x - y) \leq C(x - y)^2, \quad x, y \in \mathbb{R}. \quad (1.2)$$

This type of stochastic equation, arising from phase transition in materials science by stochastic perturbation such as impurities of the materials, has been extensively studied in the literature; see, e.g., [2], [3], [26], [27], [35] for one-dimensional equation with space-time white noise and [4], [6], [19], [34] for $d = 1, 2, 3$ -dimensional equation with colored noises. The main ingredient of these approaches in all of the aforementioned papers is the full use of the additive nature of the noise in Eq. (1.1).

The case of multiplicative noise is more subtle and challenging. For the $d = 1, 2, 3$ -dimensional stochastic Allen–Cahn equation under periodic boundary condi-

tion driven by a multiplicative $K \in \mathbb{N}_+$ -dimensional Wiener process $\beta = (\beta_1, \dots, \beta_K)$:

$$du = (\Delta u + u - u^3)dt + \sum_{k=1}^K g_k(u) d\beta_k(t), \quad u(0) = u_0, \quad (1.3)$$

the author in [33] proved that the drift-implicit Euler–Galerkin finite element scheme (1.9) with $F(u_h^{m+1}) = u_h^{m+1} - (u_h^{m+1})^3$ possesses the strong convergence rate

$$\sup_{0 \leq m \leq M} \|u(t_m) - u_h^m\|_{L_\omega^2 L_x^2} = \mathcal{O}(h^{\frac{1}{3}-\varepsilon} + \tau^{\frac{1}{2}-\varepsilon}),$$

and sharpened to $h^{1-\varepsilon}$ for $d \leq 2$ with an infinitesimal factor ε , under certain smooth and bounded assumptions on (g_k) and the assumption that $u_0 \in H^2$. When $g = g_1 \in \mathcal{C}_b^2$ (in the case $K = 1$), this convergence rate was improved to

$$\sup_{0 \leq m \leq M} \|u(t_m) - u_h^m\|_{L_\omega^2 L_x^2} = \mathcal{O}(h + \tau^{\frac{1}{2}-\varepsilon}), \quad (1.4)$$

by [28] for a modified scheme of (1.9) where the nonlinear drift term $F(u_h^{m+1})$ was replaced by $\frac{1}{2}(u_h^{m+1} + u_h^m)(1 - (u_h^{m+1})^2)$. We also refer to [11] for a theory of strong approximations of a class of quasilinear, monotone SPDEs driven by finite-dimensional Wiener processes with certain regularity conditions assumed for the solution.

It is an interesting and difficult problem to generalize the above strong error estimations for numerical approximations of Eq. (1.3) driven by an infinite-dimensional Wiener process. On the other hand, as is well-known that the finite element method possesses an optimal convergence rate for deterministic parabolic PDEs with Lipschitz coefficients, the following conjecture arises naturally.

Conjecture 1.1 *The drift-implicit Euler–Galerkin scheme (1.9) of the stochastic Allen–Cahn equation driven by a multiplicative infinite-dimensional \mathbf{Q} -Wiener process (Eq. (1.3) with $K = \infty$) with mild assumptions on the diffusion coefficients satisfies*

$$\sup_{0 \leq m \leq M} \|u(t_m) - u_h^m\|_{L_\omega^2 L_x^2} = \mathcal{O}(h^s + \tau^{\frac{1}{2}}), \text{ provided } u_0 \in \dot{H}^s \text{ for } s = 1, 2.$$

The above conjecture imposes the homogeneous Dirichlet boundary condition on the equation; such question can be readily modified to cover other self-adjoint boundary conditions, such as periodic and homogeneous Neumann boundary conditions.

1.2 Main Motivations

The proposed Conjecture 1.1 is one of our main motivations for this study. In the finite-dimensional case, provided that the drift function f satisfies one-sided Lipschitz continuity and polynomial growth conditions and the diffusion function g satisfies global Lipschitz condition (see Remark 2.4), it was shown in [12] that the drift-implicit Euler scheme (i.e., the so-called backward Euler scheme) applied to the stochastic ordinary differential equation (SODE)

$$du(t) = f(u(t))dt + g(u(t))d\beta(t), \quad (1.5)$$

is stable and possesses the standard strong order $1/2$. A general result for the infinite-dimensional case under the aforementioned conditions is still unknown and remains an open problem. This problem forms our second motivation for this study.

Our main purpose in the present paper is giving a general theory of optimal strong error estimations for numerical approximations of semilinear SPDEs with a monotone drift which grows at most polynomially such that similarity of the one-sided Lipschitz condition (1.2) holds. We focus on the following second-order parabolic SPDE:

$$du(t, x) = (\Delta u(t, x) + f(u(t, x)))dt + g(u(t, x))dW(t, x), \quad (1.6)$$

on the physical time-space domain $(t, x) \in \mathbb{R}_+ \times \mathcal{O}$, where $\mathcal{O} \subset \mathbb{R}^d$ ($d = 1, 2, 3$) is a bounded domain with smooth boundary. Eq. (1.6) is subject to the following initial value and homogeneous Dirichlet boundary condition:

$$u(t, x) = 0, \quad (t, x) \in \mathbb{R}_+ \times \partial\mathcal{O}; \quad u(0, x) = u_0(x), \quad x \in \mathcal{O}. \quad (1.7)$$

Here the drift function f is only assumed to be of monotone-type with polynomial growth which includes the case $f(x) = x - x^3$ for $x \in \mathbb{R}$, g satisfies the usual Lipschitz condition in infinite-dimensional setting (see Assumptions 2.1-2.2), and $\{W(t) : t \geq 0\}$ is an infinite-dimensional \mathbf{Q} -Wiener process in a stochastic basis $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$.

To study Eq. (1.6)–(1.7), we consider an equivalent infinite-dimensional stochastic evolution equation (SEE)

$$du(t) = (Au(t) + F(u(t)))dt + G(u(t))dW(t); \quad u(0) = u_0, \quad (1.8)$$

where A is the Dirichlet Laplacian, F and G are Nemytskii operators associated to f and g , respectively; see Section 2.2 for details. We will use different settings for the solutions of Eq. (1.8) and show that they are essentially equivalent (see Lemma 2.3). Our first main purpose is to derive the optimal strong convergence rate of the drift-implicit Euler–Galerkin finite element scheme (see the scheme (DIEG) in Section 5.1)

$$u_h^{m+1} = u_h^m + \tau A_h u_h^{m+1} + \tau \mathcal{P}_h F(u_h^{m+1}) + \mathcal{P}_h G(u_h^m)(W(t_{m+1}) - W(t_m)), \quad (1.9)$$

with the initial value $u_h^0 = \mathcal{P}_h u_0$ applied to Eq. (1.8).

Besides the optimality of spatial Galerkin approximations, the last motivation is to construct a temporal higher-order scheme for Eq. (1.8). To construct a higher-order scheme for an SODE (1.5), a popular and impressively effective numerical scheme is the Milstein scheme based on Itô–Taylor expansion on the diffusion term. Recently, such infinite-dimensional analog of Milstein scheme had been investigated by [17], [20] for SPDEs with Lipschitz coefficients which fulfill a certain commutativity type condition. Our second main purpose is to derive the sharp strong convergence rate of the Milstein–Galerkin finite element scheme (see the scheme (DIEMG) in Section 5.2)

$$\begin{aligned} U_h^{m+1} = & U_h^m + \tau A_h U_h^{m+1} + \tau \mathcal{P}_h F(U_h^{m+1}) + \mathcal{P}_h G(U_h^m)(W(t_{m+1}) - W(t_m)) \\ & + \mathcal{P}_h DG(U_h^m)G(U_h^m) \left[\int_{t_m}^{t_{m+1}} (W(r) - W(t_m))dW(r) \right], \end{aligned} \quad (1.10)$$

with the initial value $U_h^0 = \mathcal{P}_h u_0$ applied to Eq. (1.8).

1.3 Main Ideas and Results

The main aim in the present paper is to derive the optimal strong convergence rates of the Galerkin-based Euler or Milstein schemes (1.9) and (1.10), respectively, and thus give a positive answer to Conjecture 1.1 for a more general Eq. (1.6)–(1.7) with merely monotone drift which has polynomial growth and more general initial datum which belongs to $\dot{H}^{1+\gamma}$ for any $\gamma \in [0, 1]$. To answer Conjecture 1.1, there are two major difficulties. The first one is the sharp trajectory Sobolev as well as Hölder regularity for the solution of Eq. (1.8). Another one is to deal with the non-Lipschitz drift term as well as the multiplicative noise term.

To overcome the first difficulty, we combine semigroup framework and factorization method with variational framework and monotone condition. Precisely, we utilize the well-posedness theory for monotone SEEs from [24] to conclude the \dot{H}^1 -well-posedness and moments' estimation of Eq. (1.8) under Assumptions 2.1–2.2 (see Theorem 3.1). Such variational method could not handle the \dot{H}^2 -well-posedness of Eq. (1.8) since the second derivative of f has no upper bound under Assumption 2.1; see Remark 3.1 for more details. To get over this obstacle, we use a bootstrapping argument and factorization method to lift the $\dot{H}^{1+\gamma}$ -regularity for any $\gamma \in [0, 1]$ (see Proposition 3.1 and Theorem 3.3).

To overcome another difficulty for the drift-implicit Euler–Galerkin finite element scheme (1.9), we introduce an auxiliary process (4.10) and combine the factorization method with the theory of error estimation for Euler–Galerkin scheme (see Lemma 4.2). For the Milstein–Galerkin finite element scheme (1.10), we apply the Itô–Taylor expansion to lift the temporal convergence rate between an analogous auxiliary process (5.17) and the exact solution of Eq. (1.8) (see Lemma 5.1). By making full use of the variational method and the one-sided Lipschitz condition (2.3), we reduce the convergence rate between the aforementioned two auxiliary processes and the Galerkin-based fully discrete Euler or Milstein schemes to the convergence rate between the auxiliary processes and the exact solution of Eq. (1.8).

Our main results are the following strong convergence rates for both the drift-implicit Euler–Galerkin finite element scheme (1.9) and the Milstein–Galerkin finite element scheme (1.10) of Eq. (1.8).

Theorem 1.1 *Let $\gamma \in [0, 1)$, $u_0 \in \dot{H}^{1+\gamma}$, and Assumptions 2.1–2.2 hold. Let u and u_h^m be the solutions of Eq. (1.8) and (1.9), respectively. Then*

$$\sup_{m \in \mathbb{Z}_M} \|u(t_m) - u_h^m\|_{L_\omega^2 L_x^2} = \mathcal{O}(h^{1+\gamma} + \tau^{\frac{1}{2}}). \quad (1.11)$$

Assume furthermore that $u_0 \in \dot{H}^2$ and Assumption 2.3 holds, then (1.11) is valid with $\gamma = 1$.

Theorem 1.2 *Let $\gamma \in [0, 1)$, $u_0 \in \dot{H}^{1+\gamma}$, and Assumptions 2.1–2.2 and 5.1–5.2 hold. Let u and U_h^m be the solutions of Eq. (1.8) and (1.10), respectively. Then*

$$\sup_{m \in \mathbb{Z}_M} \|u(t_m) - U_h^m\|_{L_\omega^2 L_x^2} = \mathcal{O}(h^{1+\gamma} + \tau^{\frac{1+\gamma}{2}}). \quad (1.12)$$

Assume furthermore that $u_0 \in H^2$ and Assumption 2.3 holds, then (1.12) is valid with $\gamma = 1$.

We note that similar convergence results in Theorems 1.1-1.2 hold true where h is replaced by $N^{-1/d}$ for spectral Galerkin-based fully discrete Euler and Milstein schemes (see Theorems 4.1 and 5.1). The obtained spatial convergence rate in Theorems 1.1-1.2 for finite element method is optimal, since it exactly coincides with the optimal exponent of Sobolev regularity of the solution in Theorem 3.1, Proposition 3.1, and Theorem 3.3. The rate for the fully discrete scheme (1.9) is also sharp which removes the infinitesimal factor appearing in (1.4), agreeing with the standard theory of strong error estimation for the drift-implicit Euler scheme in the globally Lipschitz case; see [21] for $\gamma \in [0, 1)$. In fact, in the simplest case of the stochastic heat equation (1.6) where $f \equiv 0$ and $g = \text{Id}$, [30, Theorem 1] gave a lower error bound $\mathcal{O}(\tau^{1/2})$ for temporal numerical approximation once one only uses a total of $\lceil \tau^{-1} \rceil$ evaluations of one-dimensional components for the driving \mathbf{Q} -Wiener process W .

Associated with the aforementioned problems, the temporal approximation result can be considered as a generalization of the finite-dimensional case in [12] to the infinite-dimensional case without adding any smooth or bounded assumptions on the coefficients of Eq. (1.6) (see Remark 2.4). The spatial approximation result can be seen as an improvement of the convergence rates in [28] and [33]; see Remark 2.2. Theorem 1.1 gives more than expected in Conjecture 1.1; indeed, it shows that the drift-implicit Euler–Galerkin finite element scheme (1.9) applied to the general SPDE (1.8) which includes the stochastic Allen–Cahn equation possesses the sharp strong convergence rate under general regularity assumptions on the initial datum. Finally, the convergence rate for the Milsten–Galerkin finite element scheme (1.10) in Theorem 1.2 can be seen as a generalization of [17] and [20] for SPDEs with Lipschitz coefficients to SPDEs with non-Lipschitz but monotone coefficients.

The rest of the paper is organized as follows. We list some preliminaries including frequently used notations, assumptions, stochastic tools, and formulations for solutions of Eq. (1.8) in the next section. In Section 3, we proposed a detailed well-posedness and regularity analysis for Eq. (1.8). Rigorous strong error estimations for the aforementioned two fully discrete schemes (1.9) and (1.10) are performed in Sections 4-5, respectively. The proofs of Theorems 1.1-1.2 and similar strong error estimations for spectral Galerkin-based fully discrete Euler and Milstein schemes are given in these two sections.

2 Preliminaries

In this section, we first define functional spaces, norms, and notations that will be used throughout the paper. Then we give the main assumptions and different formulations of solutions for Eq. (1.8).

2.1 Notations

Let $T \in \mathbb{R}_+^* = (0, \infty)$ be a fixed terminal time. For an integer $M \in \mathbb{N}_+ = \{1, 2, \dots\}$, we denote by $\mathbb{Z}_M := \{0, 1, \dots, M\}$ and $\mathbb{Z}_M^* = \{1, \dots, M\}$. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space with a normal filtration $\mathbb{F} := (\mathcal{F}_t)_{t \in [0, T]}$ which forms a stochastic basis.

Throughout, we use the notations $(L_x^q := L^q(\mathcal{O}; \mathbb{R}), \|\cdot\|_{L_x^q})$ for $q \in [2, \infty]$ and $(H^s = H^s(\mathcal{O}; \mathbb{R}), \|\cdot\|_{H^s})$ for $s \in \mathbb{N}_+$ to denote the usual Lebesgue and Sobolev spaces, respectively; when $q = 2$, L_x^2 is denoted by H . We will use Id to denote the identity operator on various finite-dimensional or infinite-dimensional Hilbert spaces such as $\mathbb{R}^{K \times K}$ for some $K \in \mathbb{N}_+$ and H if there is no confusion. For convenience, sometimes we use the temporal, sample path, and spatial mixed norm $\|\cdot\|_{L_\omega^p L_t^r L_x^q}$ in different orders, such as

$$\|X\|_{L_\omega^p L_t^r L_x^q} := \left(\int_\Omega \left(\int_0^T \left(\int_0^1 |X(t, x, \omega)|^q dx \right)^{\frac{r}{q}} dt \right)^{\frac{p}{r}} \mathbb{P}(d\omega) \right)^{\frac{1}{p}}$$

for $X \in L_\omega^p L_t^r L_x^q$, with the usual modification for $r = \infty$ or $q = \infty$. For an integer $s \in \mathbb{N}_+$, $(\mathcal{C}_b^s, \|\cdot\|_{\mathcal{C}_b^s})$ is used to denote the Banach space of bounded functions together with its derivatives up to order s .

Denote by $A : \text{Dom}(A) \subset H \rightarrow H$ the Dirichlet Laplacian on H . Let $\{(\lambda_k, e_k)\}_{k=1}^\infty$ be an eigensystem of $-A$ with $\{\lambda_k\}_{k=1}^\infty$ being an increasing order. Let $N \in \mathbb{N}_+$ and $V_N := \text{Span}\{e_1, \dots, e_N\}$. Then A is the infinitesimal generator of an analytic C_0 -semigroup $S(\cdot) = e^{A\cdot}$ on H , and thus one can define the fractional powers $(-A)^\theta$ for $\theta \in \mathbb{R}$ of the self-adjoint and positive operator $-A$. Let \dot{H}^θ be the domain of $(-A)^{\theta/2}$ equipped with the norm $\|\cdot\|_\theta$ (related inner product is denoted by $\langle \cdot, \cdot \rangle_\theta$):

$$\|u\|_\theta := \|(-A)^{\frac{\theta}{2}} u\|, \quad u \in \dot{H}^\theta. \quad (2.1)$$

In particular, one has $\dot{H}^0 = H$, $\dot{H}^1 = V := H_0^1(\mathcal{O}; \mathbb{R})$ and $\dot{H}^2 = V \cap H^2$. The inner products (and related norms) of H and V are denoted by $\|\cdot\|$ (and $\langle \cdot, \cdot \rangle$) and $\|\cdot\|_1$ (and $\langle \cdot, \cdot \rangle_1$), respectively. Let $V^* = \dot{H}^{-1}$. Then V^* is the dual space of V (with respect to $\langle \cdot, \cdot \rangle$); the dualization between V and V^* is denoted by $_{-1}\langle \cdot, \cdot \rangle_1$. We will need the following ultracontractive and smoothing properties of the analytic \mathcal{C}_0 -semigroup S for any $t \in (0, T]$, $\mu \geq 0$ and $\rho \in [0, 1]$ (see, e.g., [32, Theorem 6.13 in Chapter 2]):

$$\|(-A)^\mu S(t)\|_{\mathcal{L}(H)} \leq C t^{-\mu}, \quad \|(-A)^{-\rho} (S(t) - \text{Id}_H)\|_{\mathcal{L}(H)} \leq C t^\rho, \quad (2.2)$$

where $(\mathcal{L}(H; \dot{H}^\theta), \|\cdot\|_{\mathcal{L}(H; \dot{H}^\theta)})$ denotes the space of bounded linear operators from H to \dot{H}^θ for $\theta \in \mathbb{R}$ and $\mathcal{L}(H) := \mathcal{L}(H; H)$. Here and what follows we use C, C_1, C_2, \dots , to denote generic constants independent of various discrete parameters that may be different in each appearance. For simplicity, we use the notation $a \simeq_p b$ for $a, b \in \mathbb{R}$ if there exists two constants $C_1 = C_1(p)$ and $C_2 = C_2(p)$ such that $C_1 b \leq a \leq C_2 b$.

Let U be another separable Hilbert space and $\mathbf{Q} \in \mathcal{L}(U)$ be a self-adjoint and nonnegative definite operator on U . Denote by $U_0 := \mathbf{Q}^{\frac{1}{2}} U$ and $(\mathcal{L}_2^\theta := HS(U_0; \dot{H}^\theta), \|\cdot\|_{\mathcal{L}_2^\theta})$ the space of Hilbert–Schmidt operators from U_0 to \dot{H}^θ for $\theta \in \mathbb{R}_+$. The spaces U, U_0 , and \mathcal{L}_2^θ are equipped with Borel σ -algebras $\mathcal{B}(U), \mathcal{B}(U_0)$, and $\mathcal{B}(\mathcal{L}_2^\theta)$, respectively. Let $W := \{W(t) : t \in [0, T]\}$ be a U -valued \mathbf{Q} -Wiener process in the

stochastic basis $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$, i.e., there exists an orthonormal basis $\{g_k\}_{k=1}^\infty$ of U which forms the eigenvectors of \mathbf{Q} subject to the eigenvalues $\{\lambda_k^{\mathbf{Q}}\}_{k=1}^\infty$ and a sequence of mutually independent Brownian motions $\{\beta_k\}_{k=1}^\infty$ such that (see [25, Proposition 2.1.10])

$$W(t) = \sum_{k \in \mathbb{N}_+} \mathbf{Q}^{\frac{1}{2}} g_k \beta_k(t) = \sum_{k \in \mathbb{N}_+} \sqrt{\lambda_k^{\mathbf{Q}}} g_k \beta_k(t), \quad t \in [0, T].$$

2.2 Main Assumptions

Our main conditions on the coefficients of Eq. (1.6)–(1.7) or the equivalent Eq. (1.8) are the following two assumptions.

Assumption 2.1 *$f : \mathbb{R} \rightarrow \mathbb{R}$ is differentiable and there exist constants $L_f \in \mathbb{R}$, $L'_f \in \mathbb{R}_+$ and $q \geq 2$ such that*

$$(f(x) - f(y))(x - y) \leq L_f(x - y)^2, \quad x, y \in \mathbb{R}, \quad (2.3)$$

$$|f'(x)| \leq L'_f(1 + |x|^{q-2}), \quad x \in \mathbb{R}. \quad (2.4)$$

Throughout, we assume that

$$q \in [2, \infty), \quad d = 1, 2; \quad q \in [2, 4], \quad d = 3, \quad (2.5)$$

and thus the following frequently used Sobolev embedding holds (see Remark 3.4 for the source of this technique restriction):

$$\dot{H}^1 \hookrightarrow L_x^{2(q-1)} \hookrightarrow L_x^q, \quad d = 1, 2, 3. \quad (2.6)$$

From the one-sided Lipschitz condition (2.3) and the differentiability of f , we get

$$f(x)x = (f(x) - f(0))(x - 0) + f(0)x \leq C(1 + x^2), \quad x \in \mathbb{R}, \quad (2.7)$$

and

$$f'(x) \leq L_f, \quad x \in \mathbb{R}. \quad (2.8)$$

By mean value theorem and the growth condition (2.4), it is clear that f grows as most polynomially of order $(q - 1)$, i.e.,

$$|f(x)| \leq C(1 + |x|^{q-1}), \quad x \in \mathbb{R}. \quad (2.9)$$

Example 2.1 *A concrete example such that Assumption 2.1 holds is the odd polynomials with negative leading coefficients. More precisely, let $q - 1 \in \mathbb{N}_+$ be an odd number and*

$$f(x) = -a_{q-1}x^{q-1} + \sum_{k=0}^{q-2} a_k x^k, \quad a_{q-1} \in \mathbb{R}_+^*, \quad a_k \in \mathbb{R}, \quad k \in \mathbb{Z}_{q-2}, \quad x \in \mathbb{R}, \quad (2.10)$$

then one can easily check that the conditions (2.3)–(2.4) hold true.

Denote by q' the conjugation of q (i.e., $1/q' + 1/q = 1$), $F : L_x^q \rightarrow L_x^{q'}$ the Nemytskii operator associated with f , i.e.,

$$F(u)(x) := f(u(x)), \quad u \in L_x^{q'}, \quad x \in \mathcal{O},$$

and $\langle \cdot, \cdot \rangle_{L_x^{q'}} \langle \cdot, \cdot \rangle_{L_x^q}$ the dual between $L_x^{q'}$ and L_x^q . Then it follows from Assumption 2.1 that

$$F \text{ has a continuous extension from } L_x^{q'} \text{ to } L_x^q, \quad (2.11)$$

$$\langle F(u) - F(v), u - v \rangle_{L_x^q} \leq L_f \|u - v\|_{L_x^q}^2, \quad u, v \in L_x^q, \quad \text{and} \quad (2.12)$$

$$\langle F(u), u \rangle_{L_x^q} \leq C(1 + \|u\|_{L_x^q}^2), \quad u \in L_x^q. \quad (2.13)$$

Remark 2.1 Compared with the assumption on f in the case of space-time white noise where instead of (2.3) the authors in [26] and [27] assumed that

$$(f(x) - f(y))(x - y) \leq L_f |x - y|^2 - c|x - y|^q, \quad x, y \in \mathbb{R}, \quad (2.14)$$

for some $L_f \in \mathbb{R}$ and $c \in \mathbb{R}_+^*$, Assumption 2.1 weakens this monotone-type condition due to the smoothness of the noise.

Denote by $G : H \rightarrow \mathcal{L}_2^0$ the Nemytskii operator associated with g :

$$G(u)v(x) := g(u(x))v(x), \quad u \in H, \quad v \in U_0, \quad x \in \mathcal{O}.$$

Assumption 2.2 The operator $G : H \rightarrow \mathcal{L}_2^0$ is Lipschitz continuous, i.e., there exists a constant $L_g \in \mathbb{R}_+$ such that

$$\|G(u) - G(v)\|_{\mathcal{L}_2^0} \leq L_g \|u - v\|, \quad u, v \in H. \quad (2.15)$$

Moreover, $G(\dot{H}_x^1) \subset \mathcal{L}_2^1$ and

$$\|G(z)\|_{\mathcal{L}_2^1} \leq L_g(1 + \|z\|_1), \quad z \in \dot{H}_x^1. \quad (2.16)$$

To derive the \dot{H}^2 -well-posedness and moments' estimation of Eq. (1.8) in Section 3.2, we need the following growth condition of G in $\dot{H}_x^{1+\theta}$ for some positive θ .

Assumption 2.3 There exist constants $\theta \in (0, 1)$ and $\rho \geq 1$ such that $G(\dot{H}_x^{1+\theta}) \subset \mathcal{L}_2^{1+\theta}$ and

$$\|G(z)\|_{\mathcal{L}_2^{1+\theta}} \leq C(1 + \|z\|_{1+\theta}^\rho), \quad z \in \dot{H}_x^{1+\theta}. \quad (2.17)$$

Remark 2.2 Under Majee–Prohl's condition on g of Eq. (1.3) with $K = 1$ that $g \in \mathcal{C}_b^2(\mathbb{R}; \mathbb{R})$ in [28, Assumption A.1], Assumption 2.2 holds true. In fact, in this case we have $U = \mathbb{R}$ and $\mathbf{Q} = \text{Id}$ and thus

$$\|G(u) - G(v)\|_{\mathcal{L}_2^0} = \|g(u) - g(v)\| \leq C\|u - v\|, \quad u, v \in L_x^2,$$

$$\|G(z)\|_{\mathcal{L}_2^1} \leq \sqrt{\|g(z)\|^2 + \|g'(z)\nabla z\|^2} \leq C(1 + \|z\|_1), \quad z \in \dot{H}_x^1,$$

for some constant C depending only on $\|g\|_{\mathcal{C}_b^1}$ but independent of $\|g\|_{\mathcal{C}_b^2}$.

Remark 2.3 In the additive noise case $G \equiv \text{Id}$ the conditions (2.15)-(2.16) are equivalent to the assumption $\|(-A)^{1/2} \mathbf{Q}^{1/2}\|_{HS(H;H)} < \infty$ which was imposed in [34, Section 2.1]; a multiplicative example fulfilling Assumption 2.2 is given in Example 5.1.

Remark 2.4 Provided that $f : \mathbb{R}^r \rightarrow \mathbb{R}^r$ and $g : \mathbb{R}^r \rightarrow \mathbb{R}^{r \times K}$ are \mathcal{C}^1 functions such that the following conditions hold for some constants $L_f \in \mathbb{R}$, $L_g, L'_f \in \mathbb{R}_+$, $l \in \mathbb{N}$ and for any $x, y \in \mathbb{R}^r$ (see [12, Assumptions 3.1 and 4.1]):

$$(f(x) - f(y))^\top (x - y) \leq L_f |x - y|^2, \quad (2.18)$$

$$\|g(x) - g(y)\|_{HS(\mathbb{R}^K; \mathbb{R}^r)} \leq L_g |x - y|, \quad (2.19)$$

$$|f(0)| < \infty, \quad |f(x) - f(y)|^2 \leq L'_f (1 + |x|^l + |y|^l) |x - y|^2, \quad (2.20)$$

the authors in [12, Theorem 5.3] proved that the backward Euler scheme

$$u^{m+1} = u^m + \tau f(u^{m+1}) + g(u^m)(\beta(t_{m+1}) - \beta(t_m)), \quad m \in \mathbb{Z}_{M-1}; \quad u^0 = u_0,$$

for the r -dimensional SODE (1.5) driven by a K -dimensional Wiener process β is convergent with mean-square order $1/2$.

The first two conditions (2.18) and (2.19) are exactly (2.3) and (2.15), respectively, in the finite-dimensional case $U = \mathbb{R}^K$, $\mathbf{Q} = \text{Id}$, and $H = \mathbb{R}^r$. It is clear that once the last condition (2.20) is valid, our condition (2.4) holds true with $q = 2 + l/2 \geq 2$.

2.3 Formulations of Solutions

Before formulating the solutions of Eq. (1.8) in variational and semigroup frameworks, we need to introduce several types of Burkholder–Davis–Gundy inequalities, which are frequently used in the regularity analysis and strong error estimations to control the appearing stochastic integrals in both continuous and discrete settings.

We begin with a Burkholder–Davis–Gundy inequality for Hilbert space valued continuous martingales; see, [29, Theorem 1.1].

Lemma 2.1 Let $(\mathbb{H}, \|\cdot\|_{\mathbb{H}})$ be a real separable Hilbert space. Then for any $p > 0$ and any \mathbb{H} -valued continuous \mathbb{F} -martingale $\{N(t) : t \in [0, T]\}$ with $N(0) = 0$,

$$\mathbb{E} \left[\sup_{t \in [0, T]} \|N(t)\|_{\mathbb{H}}^p \right] \simeq_p \mathbb{E} \left[\langle N \rangle_T^{\frac{p}{2}} \right]. \quad (2.21)$$

If we take $\mathbb{H} = H$ and $N(\cdot) := \int_0^\cdot \Phi(r) dW(r)$ with an \mathbb{F} -adapted process $\Phi \in L_\omega^p L_t^2 \mathcal{L}_2^0$ in Lemma 2.1, then we get $\langle N \rangle_T = \int_0^T \|\Phi(r)\|_{\mathcal{L}_2^0}^2 dr$ and the following type of Burkholder–Davis–Gundy inequality for stochastic integral:

$$\mathbb{E} \left[\sup_{t \in [0, T]} \left\| \int_0^t \Phi(r) dW(r) \right\|^p \right] \simeq_p \mathbb{E} \left[\left(\int_0^T \|\Phi(r)\|_{\mathcal{L}_2^0}^2 dr \right)^{\frac{p}{2}} \right]. \quad (2.22)$$

We also need the following Burkholder–Davis–Gundy inequality for a sequence of H -valued discrete martingales, see [16, Lemma 4.1] for the scalar case.

Lemma 2.2 Let $p \geq 2$ and $\{Z_m\}_{m \in \mathbb{Z}_M}$ be a sequence of H -valued random variables with bounded p -moments such that $\mathbb{E}[Z_{m+1}|Z_0, \dots, Z_m] = 0$, $m \in \mathbb{Z}_{M-1}$. Then there exists a constant $C = C(p)$ such that

$$\left(\mathbb{E} \left[\left\| \sum_{i=0}^m Z_i \right\|^p \right] \right)^{\frac{1}{p}} \leq C \left(\sum_{i=0}^m \left(\mathbb{E} \left[\|Z_i\|^p \right] \right)^{\frac{2}{p}} \right)^{\frac{1}{2}}. \quad (2.23)$$

Next, we recall the following definitions of the variational and mild solutions of Eq. (1.8). The relation of these two types of solutions to general semilinear SPDEs is given in [25, Appendix G]. We mainly focus on a solution which is an \dot{H}^1 -valued, \mathbb{F} -adapted process.

Definition 2.1 Let $u = \{u(t) : t \in [0, T]\}$ be an \dot{H}^1 -valued, \mathbb{F} -adapted process.

(1) u is called a variational solution of Eq. (1.8) if

$$u(t) = u_0 + \int_0^t (Au(r) + F(u(r)))dr + \int_0^t G(u(r))dW(r), \quad t \in [0, T]. \quad (2.24)$$

Eq. (2.24) takes values in \dot{H}^{-1} and explained by

$$\langle u(t), v \rangle = \langle u_0, v \rangle + \int_0^t \langle Au(r) + F(u(r)), v \rangle_1 dr + \int_0^t \langle v, G(u(r))dW(r) \rangle,$$

for any $t \in [0, T]$ and $v \in \dot{H}^1$.

(2) u is called a mild solution of Eq. (1.8) if

$$u(t) = S(t)u_0 + S * F(u)(t) + S \diamond G(u)(t), \quad t \in [0, T], \quad (2.25)$$

where $S * F(u)(\cdot) := \int_0^\cdot S(\cdot - r)F(u(r))dr$ and $S \diamond G(u)(\cdot) := \int_0^\cdot S(\cdot - r)G(u(r))dW(r)$ denote the deterministic and stochastic convolutions, respectively.

For these two types of formulations, the uniqueness of solutions are both understood in the sense of stochastic equivalence.

All of the appeared equations in this paper are valid in almost surely (a.s.) sense and we omit the a.s. in these equations for simplicity.

The main ingredient to derive the Sobolev and Hölder regularity for the solution u of Eq. (1.8) in the next section (see Section 3.2 for details) is identifying the variational and mild solutions and combining them with a bootstrap approach. We have the following equivalence between the variational and mild solutions of Eq. (1.8).

Lemma 2.3 Let $p \geq q$, u_0 is \mathcal{F}_0 -measurable such that $u_0 \in L^p(\Omega; \dot{H}^1)$, and Assumptions 2.1-2.2 hold. Then a variational solution $\{u(t) : t \in [0, T]\}$ in $L^p(\Omega; \mathcal{C}([0, T]; \dot{H}^1))$ of Eq. (1.8) is also its mild solution in the sense of Definition 2.1.

Proof By [25, Remark G.0.6], the variational solution u is also a weak solution of Eq. (1.8) in the sense that

$$\langle u(t), v \rangle = \langle u_0, v \rangle + \int_0^t \langle u(r), A^* v \rangle dr + \int_0^t \langle F(u(r)), v \rangle dr + \int_0^t \langle G(u(r))dW(r), v \rangle,$$

for any $v \in \dot{H}^2$ and $t \in [0, T]$. On the other hand, by the embedding (2.6), Young inequality, and the Burkholder–Davis–Gundy inequality (2.22) we get

$$\mathbb{E} \left[\int_0^T \|u(t)\| dt + \int_0^T \|F(u(t))\| dt + \int_0^T \|G(u(t))\|_{\mathcal{L}_2^0}^2 dt \right] \leq C \left(1 + \|u\|_{L_T^\infty L_{\dot{H}^1}^q}^q \right),$$

which is finite by the assumption $u \in L^p(\Omega; \mathcal{C}([0, T]; \dot{H}^1))$ with $p \geq q$. We conclude by [25, Proposition G.0.5(i)] that this weak solution is also a mild solution.

3 Well-posedness and Regularity

Our main aim in this part is to derive the global well-posedness and sharp trajectory Sobolev and Hölder regularity results for Eq. (1.8) under Assumptions 2.1-2.3.

3.1 \dot{H}^1 -well-posedness and Moments' Estimation

In this part, we will show the well-posedness and moments' estimation of Eq. (1.8) in V , following a known result that Eq. (1.8) whose coefficients satisfy certain monotone and Lyapunov conditions has a unique variational solution $\{u(t) : t \in [0, T]\}$ in the sense of Definition 2.1 which is essentially bounded a.s. in V .

To apply the variational method, let us introduce the Gelfand triple $\dot{H}^1 = V \hookrightarrow H = L_x^2 \hookrightarrow V^* = \dot{H}^{-1}$. By the classical stochastic variational theory, see, e.g., [24], one can prove that the equation

$$du(t) = \mathbf{A}(u(t))dt + B(u(t))dW(t), \quad t \in [0, T]; \quad u(0) = u_0, \quad (3.1)$$

whose coefficients satisfy certain monotone and Lyapunov conditions is well-posedness in V . The following result shows that the conditions given in [24, Theorem 1.1] hold for Eq. (1.8) under Assumptions 2.1-2.2. We also show that the solution of Eq. (1.8) is indeed continuous a.s. and satisfies the moments' estimation (3.2) in V .

Theorem 3.1 *Let $p \geq q$, u_0 is \mathcal{F}_0 -measurable such that $u_0 \in L^p(\Omega; \dot{H}^1)$, and Assumptions 2.1-2.2 hold. Then Eq. (1.8) exists a unique variational solution $\{u(t) : t \in [0, T]\}$ in $L^p(\Omega; \mathcal{C}([0, T]; \dot{H}^1))$ such that*

$$\mathbb{E} \left[\sup_{t \in [0, T]} \|u(t)\|_1^p \right] + \mathbb{E} \left[\int_0^T \|u(t)\|_1^{p-2} \|Au(t)\|^2 dt \right] \leq C \left(1 + \mathbb{E} \left[\|u_0\|_1^p \right] \right). \quad (3.2)$$

Proof We first verify the following conditions for Eq. (3.1) with $\mathbf{A}(u) = Au + F(u)$ and $B(u) = G(u)$ for all $u, v, z \in \dot{H}^1$ and $u_N \in V_N$:

$$\mathbb{R} \ni \lambda \mapsto -1 \langle \mathbf{A}(u + \lambda v), z \rangle_1 \text{ is continuous}; \quad (3.3)$$

$$-1 \langle \mathbf{A}(u) - \mathbf{A}(v), u - v \rangle_1 + \|B(u) - B(v)\|_{\mathcal{L}_2^0}^2 \leq C \|u - v\|^2; \quad (3.4)$$

$$\langle \mathbf{A}(u_N), u_N \rangle_1 \leq C(1 + \|u_N\|_1^2); \quad (3.5)$$

$$\|\mathbf{A}(u)\|_{-1} \leq C(1 + \|u\|_1^{p-1}); \quad (3.6)$$

$$\|B(u)\|_{\mathcal{L}_2^1} \leq C(1 + \|u\|_1). \quad (3.7)$$

The first statement (3.3) follows from (2.11) and the embedding (2.6); (3.4) and (3.7) follow from the monotonicity (2.12) and the Lipschitz continuity (2.15). Indeed,

$$-1 \langle (Au + F(u)) - (Av + F(v)), u - v \rangle_1 + \|G(u) - G(v)\|_{\mathcal{L}_2^0}^2 \leq (L_f + L_g) \|u - v\|^2.$$

Similarly, (3.5) follows from (2.13) and (2.8):

$$\langle Au_N + F(u_N), u_N \rangle_1 \leq C(1 + \|u_N\|_1^2) - \|Au_N\|^2. \quad (3.8)$$

To show (3.6) with $p \geq q$, we use the dual argument and the embedding (2.6) to get

$$\begin{aligned} \|F(u)\|_{-1} &= \sup_{v \in \dot{H}^1} \frac{-1 \langle F(u), v \rangle_1}{\|v\|_1} \leq \sup_{v \in \dot{H}^1} \|F(u)\|_{L_x^{q'}} \frac{\|v\|_{L_x^q}}{\|v\|_1} \\ &\leq C(1 + \|u\|_{L_x^q}^{q-1}) \leq C(1 + \|u\|_1^{q-1}), \end{aligned} \quad (3.9)$$

and thus $\|Au + F(u)\|_{-1} \leq \|Au\|_{-1} + \|F(u)\|_{-1} \leq C(1 + \|u\|_1^{q-1})$ for any $u \in \dot{H}^1$. Applying [24, Theorem 1.1] and the assumption that $u_0 \in L^p(\Omega; \dot{H}^1)$, we conclude that Eq. (1.8) exists a unique variational solution u which belongs to $L^p(\Omega; L^\infty(0, T; \dot{H}^1))$ such that $\mathbb{E} \left[\int_0^T \|u(t)\|_1^{p-2} \|Au(t)\|^2 dt \right] < \infty$.

It remains to show that the solution u is indeed continuous in \dot{H}^1 a.s. such that (3.2) holds. To lighten the notations, here and after we omit the temporal variable when an integral appears. Applying Itô formula (see, e.g., [25, Theorem 4.2.5]) to the functional $F : H \rightarrow \mathbb{R}$ defined by $F(v) := \frac{1}{p} \|v\|^p$ with $v = \nabla u$ leads to

$$\begin{aligned} \frac{1}{p} \|u(t)\|_1^p &= \frac{1}{p} \|u_0\|_1^p + \frac{p-2}{2} \int_0^t \|u\|_1^{p-4} \sum_{k \in \mathbb{N}_+} \langle u, G(u) \mathbf{Q}^{\frac{1}{2}} g_k \rangle_1^2 dr \\ &\quad + \int_0^t \|u\|_1^{p-2} \langle u, G(u) dW(r) \rangle_1 \\ &\quad + \int_0^t \|u\|_1^{p-2} \left(\langle Au + F(u), u \rangle_1 + \frac{1}{2} \|G(u)\|_{\mathcal{L}_2^1}^2 \right) dr. \end{aligned} \quad (3.10)$$

By Assumptions 2.1-2.2, we have

$$\langle Au + F(u), u \rangle_1 \leq C(1 + \|u\|_1^2) - \|Au\|^2, \quad u \in \dot{H}^2,$$

and

$$\|u\|_1^{p-4} \sum_{k \in \mathbb{N}_+} \langle u, G(u) \mathbf{Q}^{\frac{1}{2}} g_k \rangle_1^2 + \|u\|_1^{p-2} \|G(u)\|_{\mathcal{L}_2^1}^2 \leq C(1 + \|u\|_1^p), \quad u \in \dot{H}^1,$$

from which we obtain

$$\begin{aligned} &\|u(t)\|_1^p + p \int_0^t \|u\|_1^{p-2} \|Au\|^2 dr \\ &\leq \|u_0\|_1^p + \int_0^t C(1 + \|u\|_1^p) dr + p \int_0^t \|u\|_1^{p-2} \langle u, G(u) dW(r) \rangle_1. \end{aligned} \quad (3.11)$$

Since $\int_0^\cdot \|u\|_1^{p-2} \langle u, G(u) dW(r) \rangle_1$ is a real-valued continuous martingale, by the Burkholder–Davis–Gundy inequality (2.21), Young and Hölder inequalities, and the condition (2.16), we get an estimation of the stochastic integral in (3.11):

$$\begin{aligned} & \mathbb{E} \left[\sup_{t \in [0, T]} \left| \int_0^t \|u\|_1^{p-2} \langle u, G(u) dW(r) \rangle_1 \right| \right] \\ & \leq C \mathbb{E} \left[\left(\int_0^T \|u\|_1^{2p-2} \|G(u)\|_{\mathcal{L}_2^1}^2 dt \right)^{\frac{1}{2}} \right] \\ & \leq \frac{1}{2p} \mathbb{E} \left[\sup_{t \in [0, T]} \|u(t)\|_1^p \right] + C \int_0^T \left(1 + \mathbb{E} \left[\|u(t)\|_1^p \right] \right) dt. \end{aligned} \quad (3.12)$$

Now taking $L_\omega^1 L_t^\infty$ -norm on both sides of (3.11), we obtain

$$\begin{aligned} & \mathbb{E} \left[\sup_{t \in [0, T]} \|u(t)\|_1^p \right] + p \mathbb{E} \left[\int_0^T \|u(t)\|_1^{p-2} \|\Delta u(t)\|^2 dt \right] \\ & \leq \mathbb{E} \left[\|u_0\|_1^p \right] + \int_0^T C \left(1 + \mathbb{E} \left[\|u(t)\|_1^p \right] \right) dt, \end{aligned}$$

from which we get (3.2) by Grönwall inequality. It is clear from Eq. (3.10) that u is indeed continuous in V a.s. by the absolute continuity of Lebesgue and Itô integrals under the estimation (3.2).

Remark 3.1 Unlike (3.8) in Theorem 3.1, the following uniform coercive condition does not hold in the \dot{H}^2 -setting:

$$\langle \mathcal{P}_N F(u_N), u_N \rangle_2 \leq C(1 + \|u_N\|_2^2), \quad u_N \in \dot{H}^2,$$

since the second derivative of f has no upper bound under Assumption 2.1. Thus one could not use the arguments in Theorem 3.1 to get the \dot{H}^2 -well-posedness of Eq. (1.8). We refer to [31, Theorem 2] for an $\dot{H}^2(\mathcal{O}')$ -well-posedness result for a similar equation as Eq. (1.8) driven by an affine noise for any compact domain $\mathcal{O}' \subset \mathcal{O}$ and analogous results in weighted Sobolev spaces on the whole spatial domain \mathcal{O} using results from [22] along with several L^p -estimates.

3.2 Sobolev and Hölder Regularity

In the present part, we will derive the Sobolev and Hölder regularity for the solution u of Eq. (1.8). The first type of regularity is an estimation of u under the $L_\omega^p L_t^\infty \dot{H}^{1+\gamma}$ -norm for general $\gamma \in [0, 1]$. Another type of regularity is a temporal Hölder regularity of u under the $L_\omega^p L_x^2$ -norm. These two types of regularity are both useful, in the rest Sections 4–5, to derive the optimal strong convergence rate of fully discrete numerical approximations for Eq. (1.8).

The main ingredients are Lemma 2.3, the identification of the variational and mild solutions to Eq. (1.8), in combination with the factorization method. We first derive a sharp Hölder regularity of u under the $L_\omega^p L_x^2$ -norm.

Theorem 3.2 *Let $p \geq q$, $u_0 \in L^{p(q-1)}(\Omega; \dot{H}^1)$, and Assumptions 2.1-2.2 hold. There exists a constant C such that*

$$\sup_{0 \leq s < t \leq T} \frac{\|u(t) - u(s)\|_{L_{\omega}^p L_x^2}}{(t-s)^{1/2}} \leq C \left(1 + \|u_0\|_{L_{\omega}^{p(q-1)} \dot{H}^1}^{q-1}\right). \quad (3.13)$$

Proof By the mild formulation (2.25) and Minkovskii inequality, we get

$$\|u(t) - u(s)\|_{L_{\omega}^p L_x^2} \leq I + II + III,$$

where

$$\begin{aligned} I &= \|(S(t-s) - \text{Id})u(s)\|_{L_{\omega}^p L_x^2}, \\ II &= \left\| \int_s^t S(t-r)F(u(r))dr \right\|_{L_{\omega}^p L_x^2}, \\ III &= \left\| \int_s^t S(t-r)G(u(r))dW(r) \right\|_{L_{\omega}^p L_x^2}. \end{aligned}$$

The smoothing property (2.2) and the estimation (3.2) yield that

$$I \leq \|S(t-s) - \text{Id}\|_{\mathcal{L}(H; \dot{H}^{-1})} \|u(s)\|_{L_{\omega}^p \dot{H}^1} \leq C(t-s)^{\frac{1}{2}} (1 + \|u_0\|_{L_{\omega}^p \dot{H}^1}). \quad (3.14)$$

For the second term II , by Minkovskii inequality, the ultracontractive and smoothing properties (2.2), the embedding (2.6), and the estimation (3.2), we have

$$\begin{aligned} II &\leq \int_s^t \|S(t-r)\|_{\mathcal{L}(H)} \|F(u(r))\|_{L_{\omega}^p L_x^2} dr \\ &\leq C \|F(u)\|_{L_t^{\infty} L_{\omega}^p L_x^2} (t-s) \leq C (1 + \|u_0\|_{L_{\omega}^{p(q-1)} \dot{H}^1}^{q-1}) (t-s). \end{aligned} \quad (3.15)$$

For the last term III , by the Burkholder–Davis–Gundy inequality (2.22), the conditions (2.15)–(2.16), the smoothing property (2.2), and the estimation (3.2), we get

$$\begin{aligned} III &\leq \left(\int_s^t \|S(t-r)\|_{\mathcal{L}(H)}^2 \|G(u(r))\|_{L_{\omega}^p \mathcal{L}_2^0}^2 dr \right)^{\frac{1}{2}} \\ &\leq C \|G(u)\|_{L_t^{\infty} L_{\omega}^p \mathcal{L}_2^0} (t-s)^{\frac{1}{2}} \leq C (1 + \|u_0\|_{L_{\omega}^p \dot{H}^1}) (t-s)^{\frac{1}{2}}. \end{aligned} \quad (3.16)$$

Combining the above estimations (3.14)–(3.16), we conclude (3.13).

Remark 3.2 *One can use the factorization method as in Proposition 3.1 and then apply Lemma 3.1 to derive $u \in L_{\omega}^p \mathcal{C}^{\delta'} L_x^2$ for a $\delta' > 0$. However, our main interest here is to derive the sharp Hölder exponent δ of u such that $u \in \mathcal{C}^{\delta} L_{\omega}^p L_x^2$. In this case, the derived temporal strong convergence rate of numerical approximations for Eq. (1.8) is optimal (and without any infinitesimal factor).*

Next, we study the Sobolev regularity for the solution u of Eq. (1.8). Note that it seems impossible to use Theorem 3.1 by variational approach to derive the \dot{H}^2 -well-posedness of Eq. (1.8); see Remark 3.1. For simplicity, in the rest we always assume that the initial datum u_0 is a nonrandom function; a similar argument can handle the case of random initial data which possess finite, sufficiently large algebraic order moments.

The main tool is the following factorization formulas for deterministic and stochastic convolutions:

$$S * F(u)(t) = \frac{\sin(\pi\alpha)}{\pi} \int_0^t (t-r)^{\alpha-1} S(t-r) F_\alpha(r) dr, \quad (3.17)$$

$$S \diamond G(u)(t) = \frac{\sin(\pi\alpha)}{\pi} \int_0^t (t-r)^{\alpha-1} S(t-r) G_\alpha(r) dr, \quad (3.18)$$

where $\alpha \in (0, 1)$ and

$$F_\alpha(t) := \int_0^t (t-r)^{-\alpha} S(t-r) F(u(r)) dr,$$

$$G_\alpha(t) := \int_0^t (t-r)^{-\alpha} S(t-r) G(u(r)) dW(r),$$

for $t \in [0, T]$, see, e.g., [13], [14], and references cited therein.

To this purpose, we need the following characterization, see, e.g., [14, Proposition 4.1], about the convolution operator R_α defined by

$$R_\alpha F(t) := \int_0^t (t-r)^{\alpha-1} S(t-r) F(r) dr, \quad t \in [0, T]. \quad (3.19)$$

Lemma 3.1 *Let $p > 1$, $1/p < \alpha < 1$ and $\rho, \theta, \delta \geq 0$. Then R_α defined by (3.19) is a bounded linear operator from $L^p(0, T; \dot{H}^\rho)$ to $\mathcal{C}^{\alpha-1/p-(\theta-\rho)/2}([0, T]; \dot{H}^\theta)$ for $\theta > \rho$ and $\alpha > (\theta - \rho)/2 + 1/p$.*

Using Lemma 3.1 and a bootstrapping argument, we have the following $\dot{H}^{1+\gamma}$ -regularity of the solution to Eq. (1.8) for any $\gamma \in (0, 1)$.

Proposition 3.1 *Let $\gamma \in (0, 1)$ and $u_0 \in \dot{H}^{1+\gamma}$. Under Assumptions 2.1-2.2, the solution u of Eq. (1.8) belongs to $L^p(\Omega; \mathcal{C}([0, T]; \dot{H}^{1+\gamma}))$ for any $p \geq 1$ such that*

$$\mathbb{E} \left[\sup_{t \in [0, T]} \|u(t)\|_{1+\gamma}^p \right] \leq C \left(1 + \|u_0\|_{1+\gamma}^{p(q-1)} \right). \quad (3.20)$$

Proof Let $p > 2$ and $\alpha \in (1/p, 1/2)$. For any $\gamma \in [0, 1)$, Minkovskii and Young inequalities yield that

$$\begin{aligned} \|S * F(u)\|_{L_\omega^p L_t^\infty \dot{H}^{1+\gamma}} &\leq \left\| \int_0^t \|(-A)^{\frac{1+\gamma}{2}} S(t-r) F(u(r))\| dr \right\|_{L_\omega^p L_t^\infty} \\ &\leq C \|t^{-\frac{1+\gamma}{2}} * \|F(u)\| \|_{L_\omega^p L_t^\infty} \leq CT^{\frac{1-\gamma}{2}} \|F(u)\|_{L_\omega^p L_t^\infty L_x^2}. \end{aligned}$$

Then we get by the growth condition (2.4), the embedding (2.6), and the estimation (3.2) that

$$\|S * F(u)\|_{L_{\omega}^p L_t^{\infty} \dot{H}^{1+\gamma}} \leq C(1 + \|u\|_{L_{\omega}^{p(q-1)} L_t^{\infty} \dot{H}^1}^{q-1}) \leq C(1 + \|u_0\|_1^{q-1}). \quad (3.21)$$

For the stochastic convolution, we use the factorization formula (3.18). By the growth condition (2.16), we have

$$\|G_{\alpha}\|_{L_{\omega}^p L_t^p \dot{H}^1}^p \leq C \left[\int_0^T \left(\int_0^t (t-r)^{-2\alpha} dr \right)^{\frac{p}{2}} dt \right] \left(1 + \|u\|_{L_{\omega}^p L_t^{\infty} \dot{H}^1}^p \right) \leq C \left(1 + \|u_0\|_1^p \right).$$

Thus Lemma 3.1 with $\rho = 1$ and $\theta = 1 + \gamma$ for $\gamma \in [0, 1 - 2/p]$ yields that $S \diamond G(u) \in L^p(\Omega; \mathcal{C}([0, T]; \dot{H}^{1+\gamma}))$ and

$$\|S \diamond G(u)\|_{L_{\omega}^p L_t^{\infty} \dot{H}^{1+\gamma}} \leq C \|G_{\alpha}\|_{L_{\omega}^p L_t^p \dot{H}^1} \leq C(1 + \|u_0\|_1). \quad (3.22)$$

Combining the estimations (3.21)-(3.22) with the standard estimation for $S(\cdot)u_0$ that

$$\|S(\cdot)u_0\|_{L_{\omega}^p L_t^{\infty} \dot{H}^{1+\gamma}} \leq C \|u_0\|_{1+\gamma}, \quad (3.23)$$

we conclude (3.20) for $\gamma \in [0, 1 - 2/p]$. Taking p large enough, we prove the result for general $\gamma \in [0, 1)$.

Corollary 3.1 *Let $u_0 \in \dot{H}^1 \cap L_x^{\infty}$. Under Assumptions 2.1-2.2, the solution u of Eq. (1.8) belongs to $L^p(\Omega; \mathcal{C}([0, T]; \dot{H}^1)) \cap L^p(\Omega; \mathcal{C}([0, T]; L_x^{\infty}))$ for any $p \geq 1$ such that*

$$\mathbb{E} \left[\sup_{t \in [0, T]} \|u(t)\|_1^p \right] + \mathbb{E} \left[\sup_{t \in [0, T]} \|u(t)\|_{L_x^{\infty}}^p \right] \leq C \left(1 + \|u_0\|_1^{p(q-1)} + \|u_0\|_{L_x^{\infty}}^p \right). \quad (3.24)$$

Proof The proof of Proposition 3.1 and the embedding $\dot{H}^{1+\beta} \hookrightarrow L_x^{\infty}$ for any $\beta \in (1/2, 1]$ imply that

$$\begin{aligned} & \|S * F(u)\|_{L_{\omega}^p L_t^{\infty} L_x^{\infty}} + \|S \diamond G(u)\|_{L_{\omega}^p L_t^{\infty} L_x^{\infty}} \\ & \leq \|S * F(u)\|_{L_{\omega}^p L_t^{\infty} \dot{H}^{1+\beta}} + \|S \diamond G(u)\|_{L_{\omega}^p L_t^{\infty} \dot{H}^{1+\beta}} \\ & \leq C(1 + \|u\|_{L_{\omega}^{p(q-1)} L_t^{\infty} \dot{H}^1}^{q-1}) + C(1 + \|u\|_{L_{\omega}^p L_t^{\infty} \dot{H}^1}) \leq C(1 + \|u_0\|_1^{q-1}). \end{aligned}$$

Similarly to (3.23), we have $\|S(\cdot)u_0\|_{L_{\omega}^p L_t^{\infty} L_x^{\infty}} \leq C \|u_0\|_{L_x^{\infty}}$. Thus we get

$$\begin{aligned} \|u\|_{L_{\omega}^p L_t^{\infty} L_x^{\infty}} & \leq \|S(\cdot)u_0\|_{L_{\omega}^p L_t^{\infty} L_x^{\infty}} + \|S * F(u)\|_{L_{\omega}^p L_t^{\infty} L_x^{\infty}} + \|S \diamond G(u)\|_{L_{\omega}^p L_t^{\infty} L_x^{\infty}} \\ & \leq C(1 + \|u_0\|_1^{q-1} + \|u_0\|_{L_x^{\infty}}). \end{aligned} \quad (3.25)$$

Combining the estimation (3.25) with the estimation (3.2), we derive (3.24).

Our next step is to show the \dot{H}^2 -regularity for the solution of Eq. (1.8). To this end, we need the following result.

Lemma 3.2 *Let $\beta \in [0, 1/2)$ and (2.4) hold. For any $u \in \dot{H}^{\beta} \cap L_x^{\infty}$, there exists a constant C such that*

$$\|F(u)\|_{\beta} \leq C(1 + \|u\|_{L_x^{\infty}}^{q-1} + \|u\|_{\beta}^{q-1}). \quad (3.26)$$

Proof We assume that $\beta \in (0, 1/2)$, while the inequality (3.26) for $\beta = 0$ is trivial. Recall that for $\beta \in (0, 1/2)$ the space \dot{H}^β coincides with the Sobolev–Slobodeckij space $W^{\beta,2}$ whose norm is defined by

$$\|u\|_{W^{\beta,2}}^2 := \|u\|^2 + \int_{\mathcal{O}} \int_{\mathcal{O}} \frac{|u(x) - u(y)|^2}{|x - y|^{d+2\beta}} dx dy, \quad u \in W^{\beta,2}.$$

Then by mean value theorem, the growth condition (2.4), the embedding (2.6), and Young inequality, we get

$$\begin{aligned} \|F(u)\|_\beta^2 &\leq C \left(\|F(u)\|^2 + \int_{\mathcal{O}} \int_{\mathcal{O}} \frac{|f(u(x)) - f(u(y))|^2}{|x - y|^{d+2\beta}} dx dy \right) \\ &\leq C \left(1 + \|u\|_{L_x^{2(q-1)}}^{2(q-1)} \right) + C \left(1 + \|u\|_{L_x^\infty}^{2(q-2)} \right) \int_{\mathcal{O}} \int_{\mathcal{O}} \frac{|u(x) - u(y)|^2}{|x - y|^{d+2\beta}} dx dy \\ &\leq C \left(1 + \|u\|_{L_x^\infty}^{2(q-1)} + \|u\|_\beta^{2(q-1)} \right). \end{aligned}$$

This completes the proof of (3.26).

Theorem 3.3 *Let $u_0 \in \dot{H}^2$ and Assumptions 2.1-2.3 hold. Then the solution u of Eq. (1.8) is in $L^p(\Omega; \mathcal{C}([0, T]; \dot{H}^2))$ for any $p \geq 1$ such that*

$$\mathbb{E} \left[\sup_{t \in [0, T]} \|u(t)\|_2^p \right] \leq C \left(1 + \|u_0\|_2^{pp(q-1)} + \|u_0\|_2^{p(q-1)^2} \right). \quad (3.27)$$

Proof By Young inequality, the factorization formula (3.17), and Lemma 3.2, we get for $p > 1$, $1/p < \alpha < 1$, and $\beta \in (0, 1/2)$ that

$$\begin{aligned} \|F_\alpha\|_{L_{\omega,t}^p \dot{H}^\beta} &\leq \|t^{-\alpha} * F(u)\|_{\dot{H}^\beta} \|L_{\omega,t}^p\| \leq CT^{1-\alpha} \|F(u)\|_{L_{\omega,t}^p \dot{H}^\beta} \\ &\leq CT^{1-\alpha} \left(1 + \|u\|_{L_{\omega,t}^{p(q-1)} L_t^\infty L_x^\infty}^{q-1} + \|u\|_{L_{\omega,t}^{p(q-1)} L_t^\infty \dot{H}^\beta}^{q-1} \right). \end{aligned}$$

Since $u_0 \in \dot{H}^2$, Proposition 3.1 yields that $u \in L^p(\Omega; \mathcal{C}([0, T]; \dot{H}^{1+\gamma}))$ for any $p \geq 2$ and $\gamma \in [0, 1)$ such that (3.20) holds. In particular, we take $\gamma \in (1/2, 1)$ such that $\dot{H}^{1+\gamma} \hookrightarrow L_x^\infty$. Consequently,

$$\|F_\alpha\|_{L_{\omega,t}^p \dot{H}^\beta} \leq C \left(1 + \|u\|_{L_{\omega,t}^{p(q-1)} L_t^\infty \dot{H}^{1+\gamma}}^{q-1} \right) \leq C \left(1 + \|u_0\|_2^{(q-1)^2} \right).$$

As a result of Lemma 3.1 with $\rho = \beta = 5/p$ (take $p > 10$) and $\alpha = 1 - 1/p$, we obtain $S * F(u) \in L^p(\Omega; \mathcal{C}([0, T]; \dot{H}^2))$ such that

$$\|S * F(u)\|_{L_{\omega,t}^p L_t^\infty \dot{H}^2} \leq C \|F_\alpha\|_{L_{\omega,t}^p \dot{H}^\beta} \leq C \left(1 + \|u\|_2^{(q-1)^2} \right). \quad (3.28)$$

For the stochastic convolution $S \diamond G(u)$, the condition (2.17) and the estimation (3.20) yield that

$$\|G_\alpha\|_{L_{\omega,t}^p \dot{H}^{1+\theta}} \leq C \left(1 + \|u\|_{L_{\omega,t}^{pp} L_t^\infty \dot{H}^{1+\theta}}^p \right) \leq C \left(1 + \|u_0\|_{1+\theta}^{p(q-1)} \right) \leq C \left(1 + \|u_0\|_2^{p(q-1)} \right),$$

with $\alpha \in (1/p, 1/2)$. Taking for example $p > 4/\theta > 4$ and $\alpha = 1/2 - 1/p$ such that $\alpha \in (1/p, 1/2)$ and $2\alpha - 2/p + 1 + \theta > 2$, we obtain by Lemma 3.1 that $S \diamond G(u) \in L^p(\Omega; \mathcal{C}([0, T]; \dot{H}^2))$ such that

$$\|S \diamond G(u)\|_{L_{\omega}^p L_t^\infty \dot{H}^2} \leq C \|G_\alpha\|_{L_{\omega, t}^p \dot{H}^{1+\theta}} \leq C(1 + \|u_0\|_2^{\rho(q-1)}).$$

261 The above inequality, in combination with (3.28) and (3.23) with $\gamma = 1$, shows (3.27).

Remark 3.3 One could not expect that $u \in \mathcal{C}([0, T]; L^p(\Omega; \dot{H}^{2+\varepsilon}))$ provided $u_0 \in \dot{H}^{2+\varepsilon}$ for some $\varepsilon \in \mathbb{R}_+^*$ under Assumption 2.2. This is due to the restriction of the regularity for the stochastic convolution. For the additive noise case, the equivalent assumption $\|(-A)^{1/2} \mathbf{Q}^{1/2}\|_{HS(H, H)} < \infty$ in Remark 2.3 is not sufficient to ensure that

$$\mathbb{E} \left[\left\| \int_0^t S(t-r) dW(r) \right\|_{2+\varepsilon}^2 \right] < \infty, \quad t \in (0, T].$$

In fact, one can find a counterexample that $\mathbf{Q} = (-A)^{-\alpha}$ for some $\alpha \in \mathbb{R}$. Then

$$\mathbb{E} \left[\left\| \int_0^t S(t-r) dW(r) \right\|_{2+\varepsilon}^2 \right] = \int_0^t \sum_{k \in \mathbb{N}_+} \lambda_k^{2+\varepsilon-\alpha} e^{-2\lambda_k r} dr = \sum_{k \in \mathbb{N}_+} \frac{\lambda_k^{1+\varepsilon-\alpha} (1 - e^{-2\lambda_k t})}{2},$$

262 which is convergent for $t \in (0, T]$ if and only if $\sum_{k \in \mathbb{N}_+} \lambda_k^{1+\varepsilon-\alpha} < \infty$, i.e., $\alpha > 1 + \varepsilon + \frac{d}{2}$.

263 However, the assumption $\|(-A)^{1/2} \mathbf{Q}^{1/2}\|_{HS(H, H)} < \infty$ is equivalent to the convergence of the previous series with $\varepsilon = 0$. We also refer to [22, Examples 1.1 and 1.2] for an interesting counterexample.

266 Finally, we give another sharp Hölder regularity under \dot{H}^β -norm with $\beta \in [0, 1]$ for the solution of Eq. (1.8) with $\dot{H}^{1+\gamma}$ -valued initial datum for some $\gamma \in [0, 1]$. It is also frequently used in the derivation of the strong convergence rates of numerical approximations in Sections 4-5.

Corollary 3.2 Let $\gamma \in [0, 1]$, $u_0 \in \dot{H}^{1+\gamma}$, and Assumptions 2.1-2.2 hold. Assume that Assumption 2.3 holds if $\gamma = 1$. Then for any $p \geq 1$ and $\beta \in [0, 1]$, there exists a constant C such that

$$\sup_{0 \leq s < t \leq T} \frac{\|u(t) - u(s)\|_{L_{\omega}^p \dot{H}^\beta}}{|t-s|^{\frac{1+\gamma-\beta}{2} \wedge \frac{1}{2}}} \leq \begin{cases} C(1 + \|u_0\|_{1+\gamma}^{q-1}), & \gamma \in [0, 1); \\ C(1 + \|u_0\|_2^{\rho(q-1)} + \|u_0\|_2^{(q-1)^2}), & \gamma = 1. \end{cases} \quad (3.29)$$

Proof Without loss of generality, let $0 \leq s < t \leq T$. We use similar idea from Theorem 3.2 to show (3.29):

$$\|u(t) - u(s)\|_{L_{\omega}^p \dot{H}^\beta} \leq I_1 + II_1 + III_1,$$

where

$$\begin{aligned} I_1 &= \|(S(t) - \text{Id})u(s)\|_{L_{\omega}^p \dot{H}^\beta}, \\ II_1 &= \left\| \int_s^t S(t-r) F(u(r)) dr \right\|_{L_{\omega}^p \dot{H}^\beta}, \end{aligned}$$

$$III_1 = \left\| \int_s^t S(t-r)G(u(r))dW(r) \right\|_{L_\omega^p \dot{H}^\beta}.$$

The smoothing property (2.2) and the estimations (3.20) and (3.27) imply that

$$\begin{aligned} I_1 &\leq \|S(t-s) - \text{Id}\|_{\mathcal{L}(H; \dot{H}^{-1-\gamma+\beta})} \|u(s)\|_{L_\omega^p \dot{H}^{1+\gamma}} \leq C(t-s)^{\frac{1+\gamma-\beta}{2}} \|u\|_{L_t^\infty L_\omega^p \dot{H}^{1+\gamma}} \\ &\leq \begin{cases} C(t-s)^{\frac{1+\gamma-\beta}{2}} (1 + \|u_0\|_{1+\gamma}^{q-1}), & \gamma \in [0, 1); \\ C(t-s)^{\frac{1+\gamma-\beta}{2}} (1 + \|u_0\|_2^{p(q-1)} + \|u_0\|_2^{(q-1)^2}), & \gamma = 1. \end{cases} \end{aligned} \quad (3.30)$$

For the second term II_1 , by the ultracontractive property (2.2), the embedding (2.6), and the estimation (3.2) we have

$$II_1 \leq \|F(u)\|_{L_t^\infty L_\omega^p L_x^2} \left[\int_s^t (t-r)^{-\frac{\beta}{2}} dr \right] \leq C(1 + \|u_0\|_1^{q-1})(t-s)^{1-\frac{\beta}{2}}. \quad (3.31)$$

For the last term III_1 , by the Burkholder–Davis–Gundy inequality (2.22), the conditions (2.15), (2.16), and (2.17), the smoothing property (2.2), and the estimations (3.2) and (3.20), we get

$$III_1 \leq \|G(u)\|_{L_t^\infty L_\omega^p \mathcal{L}_2^1} (t-s)^{\frac{1}{2}} \leq C(1 + \|u_0\|_1)(t-s)^{\frac{1}{2}}. \quad (3.32)$$

Combining the above estimations (3.30)–(3.32), we conclude (3.29).

Remark 3.4 The technical requirement that $q \leq 4$ when $d = 3$ in (2.5) comes from the embedding $H^1 \hookrightarrow L_x^{2(q-1)}$ to control $\|F(u)\|$ by $\|u\|_1$, which was used in Lemma 2.3, (3.15) in Theorem 3.2, (3.21) in Proposition 3.1, and (3.31) in Corollary 3.2.

4 Euler–Galerkin Scheme

Our main aim in this section is to give a Galerkin-based fully discrete scheme and derive its optimal strong convergence rate. The proof of Theorem 1.1 is given at the end of Section 4.2.

4.1 Drift-implicit Euler–Galerkin Finite Element Scheme

Let $h \in (0, 1)$, \mathcal{T}_h be a regular family of partitions of \mathcal{O} with maximal length h , and $V_h \subset V$ be the space of continuous functions on $\bar{\mathcal{O}}$ which are piecewise linear over \mathcal{T}_h and vanish on the boundary $\partial\mathcal{O}$. Let $A_h : V_h \rightarrow V_h$ and $\mathcal{P}_h : V^* \rightarrow V_h$ be the discrete Laplacian and generalized orthogonal projection operator, respectively, defined by

$$\langle A_h u_h, v_h \rangle = -\langle \nabla u_h, \nabla v_h \rangle, \quad u_h, v_h \in V_h, \quad (4.1)$$

$$\langle \mathcal{P}_h u, v_h \rangle = -1 \langle u, v_h \rangle_1, \quad u \in V^*, v_h \in V_h. \quad (4.2)$$

Taking $u \in H$ and $v_h = \mathcal{P}_h u$ in (4.2) and using Cauchy–Schwarz inequality implies the following H -contraction property of \mathcal{P}_h :

$$\|\mathcal{P}_h u\| \leq \|u\|, \quad u \in H. \quad (4.3)$$

The finite element approximation for Eq. (1.8) is to find an \mathbb{F} -adapted V_h -valued process $\{u_h(t) : t \in [0, T]\}$ such that the following variational equality holds for any $t \in [0, T]$ and $v_h \in V_h$:

$$\begin{aligned} \langle u_h(t), v_h \rangle + \int_0^t \langle \nabla u_h(r), \nabla v_h \rangle dr \\ = \langle u_h(0), v_h \rangle + \int_0^t \langle F(u_h(r)), v_h \rangle dr + \int_0^t \langle v_h, G(u_h(r)) dW(r) \rangle. \end{aligned} \quad (4.4)$$

To complement (4.4), we set the initial datum to be $u_h(0) = \mathcal{P}_h u_0$. Then the finite element approximation (4.4) is equivalent to

$$\begin{aligned} du_h(t) &= (A_h u_h(t) + \mathcal{P}_h F(u_h(t))) dt + \mathcal{P}_h G(u_h(t)) dW(t), \quad t \in [0, T]; \\ u_h(0) &= \mathcal{P}_h u_0. \end{aligned} \quad (4.5)$$

Let $M \in \mathbb{N}_+$ and $\{(t_m, t_{m+1}] : m \in \mathbb{Z}_{M-1}\}$ be an equal length subdivision of $(0, T]$ with temporal step-size $\tau = t_{m+1} - t_m$ for each $m \in \mathbb{Z}_{M-1}$. The drift-implicit Euler (DIE) scheme of the finite element approximation (4.5), that we call the DIE Galerkin (DIEG) scheme, is to find a V_h -valued discrete process $\{u_h^m : m \in \mathbb{Z}_M\}$ such that

$$u_h^{m+1} = u_h^m + \tau A_h u_h^{m+1} + \tau \mathcal{P}_h F(u_h^{m+1}) + \mathcal{P}_h G(u_h^m) \delta_m W, \quad u_h^0 = \mathcal{P}_h u_0, \quad (\text{DIEG})$$

where $\delta_m W = W(t_{m+1}) - W(t_m)$, $m \in \mathbb{Z}_{M-1}$. This DIEG scheme had been widely studied; see, e.g., [10], [33], and [34].

It is clear that the DIEG scheme (DIEG) is equivalent to the scheme

$$u_h^{m+1} = S_{h,\tau} u_h^m + \tau S_{h,\tau} \mathcal{P}_h F(u_h^{m+1}) + S_{h,\tau} \mathcal{P}_h G(u_h^m) \delta_m W, \quad m \in \mathbb{Z}_{M-1}, \quad (4.6)$$

with initial datum $u_h^0 = \mathcal{P}_h u_0$, where $S_{h,\tau} := (\text{Id} - \tau A_h)^{-1}$ is a space-time approximation of the continuous semigroup S in one step. Iterating (4.6) for m -times, we obtain

$$u_h^{m+1} = S_{h,\tau}^{m+1} u_h^0 + \tau \sum_{i=0}^m S_{h,\tau}^{m+1-i} \mathcal{P}_h F(u_h^{i+1}) + \sum_{i=0}^m S_{h,\tau}^{m+1-i} \mathcal{P}_h G(u_h^i) \delta_i W, \quad m \in \mathbb{Z}_{M-1}. \quad (4.7)$$

Throughout we take $\tau \in (0, 1)$ when $L_f \leq 0$ and $\tau < 1/(4L_f)$ when $L_f \in \mathbb{R}_+^*$. Under Assumptions 2.1-2.2, it is not difficult to show that the DIEG scheme (DIEG) is solvable and stable.

4.2 Strong Convergence Rate of DIEG Scheme

In this part, we give the optimal strong convergence rate of the DIEG scheme (DIEG).

Denote by $E_{h,\tau}(t) = S(t) - S_{h,\tau}^{m+1} \mathcal{P}_h$ for $t \in (t_m, t_{m+1}]$ with $m \in \mathbb{Z}_{M-1}$. The following estimations of $E_{h,\tau}$ play a pivotal role in the error estimation of the DIEG scheme (DIEG); see, e.g., [21, Lemmas 4.3(i) and 4.4(ii)].

Lemma 4.1 Let $t \in (0, T]$.

1. For $0 \leq v \leq \mu \leq 2$ and $x \in \dot{H}^v$,

$$\|E_{h,\tau}(t)x\| \leq C(h^\mu + \tau^{\frac{\mu}{2}})t^{-\frac{\mu-v}{2}}\|x\|_v. \quad (4.8)$$

2. For $0 \leq \mu \leq 1$ and $x \in \dot{H}^\mu$,

$$\left(\int_0^t \|E_{h,\tau}(r)x\|^2 dr\right)^{\frac{1}{2}} \leq C(h^{1+\mu} + \tau^{\frac{1+\mu}{2}})\|x\|_\mu. \quad (4.9)$$

To estimate the solution u of Eq. (1.8) and u_h^m , we introduce the auxiliary process

$$\tilde{u}_h^{m+1} = S_{h,\tau}^{m+1}u_h^0 + \tau \sum_{i=0}^m S_{h,\tau}^{m+1-i} \mathcal{P}_h F(u(t_{i+1})) + \sum_{i=0}^m S_{h,\tau}^{m+1-i} \mathcal{P}_h G(u(t_i)) \delta_i W, \quad (4.10)$$

for $m \in \mathbb{Z}_{M-1}$, where the terms u_h^{i+1} and u_h^i in the discrete deterministic and stochastic convolutions of (4.7) are replaced by $u(t_{i+1})$ and $u(t_i)$, respectively. We start with the following uniform boundedness of \tilde{u}_h^m in $L_\omega^p L_t^\infty \dot{H}_x^1$:

$$\sup_{m \in \mathbb{Z}_M} \mathbb{E} \left[\|\tilde{u}_h^m\|_1^p \right] \leq C \left(1 + \|u_0\|_1^{p(q-1)} \right), \quad (4.11)$$

provided $u_0 \in \dot{H}_x^1$ and Assumptions 2.1-2.2 hold, which can be shown in view of the boundedness (3.2) and the well-known uniform boundedness

$$\sup_{m \in \mathbb{Z}_M} \|S_{h,\tau}^m \mathcal{P}_h x\| \leq C\|x\|, \quad x \in H. \quad (4.12)$$

Next, we show the strong error estimation between the exact solution u of Eq. (1.8) and the auxiliary process $\{\tilde{u}_h^m\}_{m \in \mathbb{Z}_M}$ defined by (4.10).

Lemma 4.2 *Let $\gamma \in [0, 1)$, $u_0 \in \dot{H}^{1+\gamma}$, and Assumptions 2.1-2.2 hold. Then for any $p \geq 1$, there exist a constant C such that*

$$\sup_{m \in \mathbb{Z}_M} \|u(t_m) - \tilde{u}_h^m\|_{L_\omega^p L_x^2} \leq C(h^{1+\gamma} + \tau^{\frac{1}{2}}). \quad (4.13)$$

Assume furthermore that $u_0 \in \dot{H}^2$ and Assumption 2.3 holds, then (4.13) is valid with $\gamma = 1$.

Proof Let $m \in \mathbb{Z}_{M-1}$. Subtracting the auxiliary process \tilde{u}_h^{m+1} defined by (4.10) from the mild formulation (2.25) with $t = t_{m+1}$, i.e.,

$$u(t_{m+1}) = S(t_{m+1})u_0 + \int_0^{t_{m+1}} S(t_{m+1}-r)F(u)dr + \int_0^{t_{m+1}} S(t_{m+1}-r)G(u)dW(r),$$

we get

$$J^{m+1} := \|u(t_{m+1}) - \tilde{u}_h^{m+1}\|_{L_\omega^p L_x^2} \leq \sum_{i=1}^3 J_i^{m+1}, \quad (4.14)$$

where

$$J_1^{m+1} = \|E_{h,\tau}(t_{m+1})u_0\|_{L_\omega^p L_x^2},$$

$$J_2^{m+1} = \left\| \int_0^{t_{m+1}} S(t_{m+1} - r) F(u) dr - \tau \sum_{i=0}^m S_{h,\tau}^{m+1-i} \mathcal{P}_h F(u(t_{i+1})) \right\|_{L_\omega^p L_x^2},$$

$$J_3^{m+1} = \left\| \int_0^{t_{m+1}} S(t_{m+1} - r) G(u) dW(r) - \sum_{i=0}^m S_{h,\tau}^{m+1-i} \mathcal{P}_h G(u(t_i)) \delta_i W \right\|_{L_\omega^p L_x^2}.$$

294 In the sequel, we treat the above three terms one by one.
The estimation (4.8) with $\mu = \nu = 1 + \gamma$ yields that

$$J_1^{m+1} \leq C(h^{1+\gamma} + \tau^{\frac{1+\gamma}{2}}) \|u_0\|_{1+\gamma}, \quad \gamma \in [0, 1]. \quad (4.15)$$

To deal with the second term, we decompose it into the following two terms:

$$J_2^{m+1} \leq \left\| \sum_{i=0}^m \int_{t_i}^{t_{i+1}} S(t_{m+1} - r) [F(u) - F(u(t_{i+1}))] dr \right\|_{L_\omega^p L_x^2}$$

$$+ \left\| \sum_{i=0}^m \int_{t_i}^{t_{i+1}} E_{h,\tau}(t_{m+1} - r) F(u(t_{i+1})) dr \right\|_{L_\omega^p L_x^2} := \sum_{i=1}^2 J_{2i}^{m+1}.$$

Similarly to (3.9), we have by the embedding (2.6) that

$$\|(-A)^{-\frac{1}{2}}(F(u) - F(v))\| = \sup_{z \in \dot{H}^1} \frac{-1 \langle F(u) - F(v), z \rangle_1}{\|z\|_1}$$

$$\leq C(1 + \|u\|_{L_x^{2(q-1)}}^{q-2} + \|v\|_{L_x^{2(q-1)}}^{q-2}) \|u - v\| \cdot \sup_{z \in \dot{H}^1} \frac{\|z\|_{L_x^{2(q-1)}}}{\|z\|_1}$$

$$\leq C(1 + \|u\|_1^{q-2} + \|v\|_1^{q-2}) \|u - v\|. \quad (4.16)$$

By Minkovskii inequality, the uniform boundedness (2.2), the above dual estimation (4.16) and the Hölder estimate (3.13), we get

$$J_{21}^{m+1} \leq C \left(1 + \|u\|_{L_t^\infty L_\omega^{2p(q-2)} \dot{H}^1}^{q-2} \right) \left[\sup_{0 \leq s < t \leq T} \frac{\|u(t) - u(s)\|_{L_\omega^{2p} L_x^2}}{|t - s|^{1/2}} \right]$$

$$\times \left[\sum_{i=0}^m \int_{t_i}^{t_{i+1}} \|S(t_{m+1} - r)\|_{\mathcal{L}(H; \dot{H}^1)} (t_{i+1} - r)^{\frac{1}{2}} dr \right]$$

$$\leq C(1 + \|u_0\|_1^{(q-1)(q-2)}) \tau^{\frac{1}{2}}.$$

Applying Minkovskii inequality and using (4.8) with $\mu = 1 + \gamma$ with $\gamma \in [0, 1)$ and $\nu = 0$, the embedding (2.6) and the estimations (3.2), we derive

$$J_{22}^{m+1} \leq \sum_{j=0}^m \int_{t_j}^{t_{j+1}} \|E_{h,\tau}(\sigma) F(u(t_{m+1-j}))\|_{L_\omega^p L_x^2} d\sigma$$

$$\leq C(h^{1+\gamma} + \tau^{\frac{1+\gamma}{2}}) \|F(u)\|_{L_t^\infty L_\omega^p L_x^2} \left[\sum_{j=0}^m \int_{t_j}^{t_{j+1}} \sigma^{-\frac{1+\gamma}{2}} d\sigma \right]$$

$$\leq C(h^{1+\gamma} + \tau^{\frac{1+\gamma}{2}}) (1 + \|u_0\|_1^{q-1}), \quad \gamma \in [0, 1). \quad (4.17)$$

Combining the above two estimations implies that

$$J_2^{m+1} \leq C(h^{1+\gamma} + \tau^{\frac{1}{2}})(1 + \|u_0\|_1^{q-1} + \|u_0\|_1^{(q-1)(q-2)}), \quad \gamma \in [0, 1). \quad (4.18)$$

The last term J_3^{m+1} can be decomposed as

$$\begin{aligned} J_3^{m+1} &\leq \left\| \sum_{i=0}^m \int_{t_i}^{t_{i+1}} S_{h,\tau}^{m+1-i} \mathcal{P}_h[G(u) - G(u(t_i))] dW(r) \right\|_{L_\omega^p L_x^2} \\ &\quad + \left\| \int_0^{t_{m+1}} E_{h,\tau}(t_{m+1} - r) G(u) dW(r) \right\|_{L_\omega^p L_x^2} := \sum_{i=1}^2 J_{3i}^{m+1}. \end{aligned}$$

Applying both the discrete and continuous Burkholder–Davis–Gundy inequalities (2.23) and (2.22), respectively, and using the uniform stability property (4.12), the conditions (2.15)–(2.16), and the estimations (3.2), (3.13), and (4.8) with $(\mu, \nu) = (1 + \gamma, 1)$ for $\gamma \in [0, 1)$, we get

$$\begin{aligned} J_3^{m+1} &\leq C \left(\sum_{i=0}^m \int_{t_i}^{t_{i+1}} \|S_{h,\tau}^{m+1-i} \mathcal{P}_h[G(u) - G(u(t_i))]\|_{L_\omega^p \mathcal{L}_2^0}^2 dr \right)^{\frac{1}{2}} \\ &\quad + C \left(\int_0^{t_{m+1}} \|E_{h,\tau}(t_{m+1} - r) G(u)\|_{L_\omega^p \mathcal{L}_2^0}^2 dr \right)^{\frac{1}{2}} \\ &\leq C \left[\sup_{0 \leq s < t \leq T} \frac{\|u(t) - u(s)\|_{L_\omega^p L_x^2}}{(t-s)^{1/2}} \right] \left(\sum_{i=0}^m \int_{t_i}^{t_{i+1}} (r - t_i) dr \right)^{\frac{1}{2}} \\ &\quad + C(h^{1+\gamma} + \tau^{\frac{1+\gamma}{2}}) \left(1 + \|u\|_{L_t^\infty L_\omega^p \dot{H}^1} \right) \left(\int_0^{t_{m+1}} r^{-\gamma} dr \right)^{\frac{1}{2}} \\ &\leq C(1 + \|u_0\|_1^{q-1})(h^{1+\gamma} + \tau^{\frac{1}{2}}), \quad \gamma \in [0, 1). \end{aligned} \quad (4.19)$$

Putting the estimations (4.15), (4.18), and (4.19) together results in

$$J^{m+1} \leq C(1 + \|u_0\|_1^{q-1} + \|u_0\|_1^{(q-1)(q-2)})(h^{1+\gamma} + \tau^{\frac{1}{2}}), \quad \gamma \in [0, 1).$$

This inequality in combination with estimation on J^0 that

$$J^0 = \|u_0 - \mathcal{P}_h u_0\|_{L_\omega^p L_x^2} \leq Ch^{1+\gamma} \|u_0\|_{1+\gamma}, \quad \gamma \in [0, 1],$$

completes the proof of (4.13) with $\gamma \in [0, 1)$.

To show (4.13) for $\gamma = 1$, we just need to give refined estimations for the J_{22}^{m+1} and J_{32}^{m+1} provided $u_0 \in \dot{H}^2$ and (2.17) holds. Applying Minkovskii inequality and using (4.8) with $\mu = 2$ and $\nu = \beta \in (0, 1)$, the embedding $\dot{H}^2 \hookrightarrow \dot{H}^\beta \cap L_x^\infty$, and the estimations (3.2), (3.24), and (3.26), we derive

$$\begin{aligned} J_{22}^{m+1} &\leq C(h^2 + \tau) \left(1 + \|u\|_{L_t^\infty L_\omega^p \dot{H}^\beta}^{q-1} + \|u\|_{L_t^\infty L_\omega^p L_x^\infty}^{q-1} \right) \left[\sum_{j=0}^m \int_{t_j}^{t_{j+1}} \sigma^{-\frac{2-\beta}{2}} d\sigma \right] \\ &\leq C(h^2 + \tau) (1 + \|u_0\|_2^{(q-1)^2}). \end{aligned} \quad (4.20)$$

On the other hand, (4.8) with $\mu = 2$ and $\nu = 1 + \theta$, the condition (2.17) and the estimation (3.20) imply that

$$\begin{aligned} J_{32}^{m+1} &\leq C(h^2 + \tau) \|G(u)\|_{L_t^\infty L_\omega^p \dot{H}^{1+\theta}} \left(\int_0^T \sigma^{-(1-\theta)} d\sigma \right)^{\frac{1}{2}} \\ &\leq C(1 + \|u_0\|_2^{q-1}) (h^2 + \tau). \end{aligned} \quad (4.21)$$

Remark 4.1 Assumption 2.3 is not needed in the additive noise case $G(u) \equiv G \in \mathcal{L}_2^1$ for any $u \in H$ when $\gamma = 1$. In fact, the Burkholder–Davis–Gundy inequality (2.22) and the estimation (4.9) with $\mu = 1$ show the following sharp estimation:

$$\left\| \int_0^{t_{m+1}} E_{h,\tau}(t_{m+1} - r) dW(r) \right\|_{L_\omega^p L_x^2} \leq C \left(\int_0^T \|E_{h,\tau}(r)\|_{\mathcal{L}_2^0}^2 dr \right)^{\frac{1}{2}} \leq C(h^2 + \tau) \|G\|_{\mathcal{L}_2^1}.$$

Combining Lemma 4.2 with a variational approach under the monotone Assumption (2.1), we can prove Theorem 1.1 on the strong convergence rate between the solution u of Eq. (1.8) and the numerical solution u_h^m of the DIEG scheme (DIEG), which gives a positive answer to Conjecture 1.1.

Proof of Theorem 1.1 Let $m \in \mathbb{Z}_{M-1}$. By Minkovskii inequality, we get

$$\|u(t_{m+1}) - u_h^{m+1}\|_{L_\omega^p L_x^2} \leq \|u(t_{m+1}) - \tilde{u}_h^{m+1}\|_{L_\omega^p L_x^2} + \|\tilde{u}_h^{m+1} - u_h^{m+1}\|_{L_\omega^p L_x^2}.$$

In terms of (4.13), it suffices to prove that

$$\sup_{m \in \mathbb{Z}_M} \|\tilde{u}_h^m - u_h^m\|_{L_\omega^p L_x^2} \leq C(h^{1+\gamma} + \tau^{\frac{1}{2}}), \quad \gamma \in [0, 1]. \quad (4.22)$$

Define $\tilde{e}_h^{m+1} := \tilde{u}_h^{m+1} - u_h^{m+1}$. Then $\tilde{e}_h^{m+1} \in V_h$ with vanishing initial datum $\tilde{e}_h^0 = 0$. In terms of Eq. (4.6) and (4.10), it is not difficult to show that

$$\begin{aligned} u_h^{m+1} &= u_h^m + \tau A_h u_h^{m+1} + \tau \mathcal{P}_h F(u_h^{m+1}) + \mathcal{P}_h G(u_h^m) \delta_m W, \\ \tilde{u}_h^{m+1} &= \tilde{u}_h^m + \tau A_h \tilde{u}_h^{m+1} + \tau \mathcal{P}_h F(u(t_{m+1})) + \mathcal{P}_h G(u(t_m)) \delta_m W. \end{aligned}$$

Consequently,

$$\begin{aligned} \tilde{e}_h^{m+1} - \tilde{e}_h^m &= \tau A_h \tilde{e}_h^{m+1} + \tau \mathcal{P}_h (F(u(t_{m+1})) - F(u_h^{m+1})) \\ &\quad + \mathcal{P}_h (G(u(t_m)) - G(u_h^m)) \delta_m W. \end{aligned}$$

Multiplying \tilde{e}_h^{m+1} on both sides of the above equation, we obtain

$$\begin{aligned} &\langle \tilde{e}_h^{m+1} - \tilde{e}_h^m, \tilde{e}_h^{m+1} \rangle + \tau \|\nabla \tilde{e}_h^{m+1}\|^2 \\ &= \tau \langle F(u(t_{m+1})) - F(\tilde{u}_h^{m+1}), \tilde{e}_h^{m+1} \rangle + \tau \langle F(\tilde{u}_h^{m+1}) - F(u_h^{m+1}), \tilde{e}_h^{m+1} \rangle \\ &\quad + \int_{t_m}^{t_{m+1}} \langle \tilde{e}_h^{m+1}, (G(u(t_m)) - G(u_h^m)) dW(r) \rangle. \end{aligned} \quad (4.23)$$

By the dual estimation (4.16), the one-sided Lipschitz condition (2.12), Cauchy–Schwarz inequality and the embedding (2.6), the right-hand side term of the above equation can be controlled by

$$\tau \|(-A)^{-\frac{1}{2}} (F(u(t_{m+1})) - F(\tilde{u}_h^{m+1}))\| \times \|\nabla \tilde{e}_h^{m+1}\| + L_f \tau \|\tilde{e}_h^{m+1}\|^2$$

$$\begin{aligned}
& + \int_{t_m}^{t_{m+1}} \langle \tilde{e}_h^{m+1}, (G(u(t_m)) - G(u_h^m)) dW(r) \rangle \\
& \leq \frac{\tau}{2} \|\nabla \tilde{e}_h^{m+1}\|^2 + L_f \tau \|\tilde{e}_h^{m+1}\|^2 \\
& \quad + C \tau \|u(t_{m+1}) - \tilde{u}_h^{m+1}\|^2 (1 + \|u(t_{m+1})\|_1^{q-2} + \|\tilde{u}_h^{m+1}\|_1^{q-2}) \\
& \quad + \int_{t_m}^{t_{m+1}} \langle \tilde{e}_h^{m+1}, (G(u(t_m)) - G(u_h^m)) dW(r) \rangle. \tag{4.24}
\end{aligned}$$

The elementary identity $(a-b)a = \frac{1}{2}(a^2 - b^2) + \frac{1}{2}(a-b)^2$ for $a, b \in \mathbb{R}$ yields that

$$\langle \tilde{e}_h^{m+1} - \tilde{e}_h^m, \tilde{e}_h^{m+1} \rangle = \frac{1}{2} (\|\tilde{e}_h^{m+1}\|^2 - \|\tilde{e}_h^m\|^2) + \frac{1}{2} \|\tilde{e}_h^{m+1} - \tilde{e}_h^m\|^2. \tag{4.25}$$

To deal with the stochastic term in (4.23), we use the martingale property of the stochastic integral, Cauchy–Schwarz inequality, and Itô isometry to deduce that

$$\begin{aligned}
& \mathbb{E} \left[\int_{t_m}^{t_{m+1}} \langle \tilde{e}_h^{m+1}, (G(u(t_m)) - G(u_h^m)) dW(r) \rangle \right] \\
& = \mathbb{E} \left[\int_{t_m}^{t_{m+1}} \langle \tilde{e}_h^{m+1} - \tilde{e}_h^m, (G(u(t_m)) - G(u_h^m)) dW(r) \rangle \right] \\
& \leq \frac{1}{2} \mathbb{E} [\|\tilde{e}_h^{m+1} - \tilde{e}_h^m\|^2] + L_g^2 \mathbb{E} [\|u(t_m) - \tilde{u}_h^m\|^2] \tau + L_g^2 \mathbb{E} [\|\tilde{e}_h^m\|^2] \tau. \tag{4.26}
\end{aligned}$$

Now taking expectation in (4.23), we use (4.24)-(4.26) to get

$$\begin{aligned}
& \frac{1}{2} \left(\mathbb{E} [\|\tilde{e}_h^{m+1}\|^2] - \mathbb{E} [\|\tilde{e}_h^m\|^2] \right) \\
& \leq L_f \mathbb{E} [\|\tilde{e}_h^{m+1}\|^2] \tau + C \left(\mathbb{E} [\|u(t_{m+1}) - \tilde{u}_h^{m+1}\|^4] \right)^{\frac{1}{2}} \tau \\
& \quad \times \left(1 + \left(\mathbb{E} [\|u(t_{m+1})\|_1^{2(q-2)}] \right)^{\frac{1}{2}} + \mathbb{E} [\|\tilde{u}_h^{m+1}\|_1^{2(q-2)}] \right)^{\frac{1}{2}} \\
& \quad + L_g^2 \mathbb{E} [\|u(t_m) - \tilde{u}_h^m\|^2] \tau + L_g^2 \mathbb{E} [\|\tilde{e}_h^m\|^2] \tau.
\end{aligned}$$

Then by the estimations (3.2), (4.11), and (4.13), we obtain

$$(1 - 2L_f \tau) \mathbb{E} [\|\tilde{e}_h^{m+1}\|^2] \leq C \tau (h^{1+\gamma} + \tau^{\frac{1}{2}})^2 + (1 + 2L_g^2 \tau) \mathbb{E} [\|\tilde{e}_h^m\|^2]. \tag{4.27}$$

Summing over $m = 0, 1, \dots, l-1$ with $1 \leq l \leq M$, we obtain

$$(1 - 2L_f \tau) \mathbb{E} [\|\tilde{e}_h^l\|^2] \leq C (h^{1+\gamma} + \tau^{\frac{1}{2}})^2 + 2(L_f + L_g^2) \tau \sum_{m=0}^{l-1} \mathbb{E} [\|\tilde{e}_h^m\|^2],$$

from which we conclude (4.22) by the classical discrete Grönwall inequality and thus completes the proof.

Remark 4.2 Our analysis can be extended to a slight general nonlinear drift $\tilde{F} = \tilde{F}(u; \nabla u) = F(u) + C\nabla u$, where F is the Nemytskii operator associated with f that satisfies Assumption (2.1) and C is a constant. Indeed, for the estimate of $u(t_m) - \tilde{u}_h^m$ in Lemma 4.2, the corresponding term J_2^{m+1} involving $F(u)$ has the same estimate as (4.18); for the estimate of \tilde{e}_h^{m+1} in the proof of Theorem 1.1, one only needs to estimate the term $C\tau \langle \nabla(u(t_{m+1}) - \tilde{u}_h^{m+1}), \tilde{e}_h^{m+1} \rangle + C\tau \langle \nabla \tilde{e}_h^{m+1}, \tilde{e}_h^{m+1} \rangle$ added in the right-hand side of (4.23). It can be bounded by $C\tau \|u(t_{m+1}) - \tilde{u}_h^{m+1}\|^2 + C\tau \|\tilde{e}_h^{m+1}\|^2 + \frac{\tau}{2} \|\nabla \tilde{e}_h^{m+1}\|^2$, that appeared also in (4.24), so the proof of Theorem 1.1 suffices.

4.3 Drift-implicit Euler–Spectral Galerkin Scheme

The parallel result, Theorem 1.1, for DIEG scheme (DIEG) in Section 4.1 is also valid for the spectral Galerkin-based Euler scheme of Eq. (1.8).

Let us first recall and introduce more related notations in Section 3.1. Given an $N \in \mathbb{N}_+$, V_N is the linear space spanned by the first N eigenvectors of Dirichlet Laplacian. Let $\mathcal{P}_N : H \rightarrow V_N$ be the orthogonal projection operator in H and A_N be the Laplacian restricted in V_N .

The spectral Galerkin approximation for Eq. (1.8) is to find a sequence of \mathbb{F} -adapted V_N -valued processes $\{u_N(t) : t \in [0, T]\}$ such that

$$du_N(t) = (A_N u_N(t) + \mathcal{P}_N F(u_N(t)))dt + \mathcal{P}_N G(u_N(t))dW(t), \quad t \in [0, T], \quad (4.28)$$

with initial datum $u_N(0) = \mathcal{P}_N u_0$. By the arguments in the proof of Theorem 1.1, one can show the similar solvability and convergence results for the DIE spectral Galerkin (DIESG) scheme

$$u_N^{m+1} = u_N^m + \tau A_N u_N^{m+1} + \tau \mathcal{P}_N F(u_N^{m+1}) + \mathcal{P}_N G(u_N^m) \delta_m W, \quad u_N^0 = \mathcal{P}_N u_0. \quad (\text{DIESG})$$

Theorem 4.1 Let $\gamma \in [0, 1)$, $u_0 \in \dot{H}^{1+\gamma}$, and Assumptions 2.1–2.2 hold. For any $p \geq 1$, there exist a constant C such that

$$\sup_{m \in \mathbb{Z}_M} \|u(t_m) - u_N^m\|_{L_\omega^2 L_x^2} \leq C(N^{-\frac{1+\gamma}{d}} + \tau^{\frac{1}{2}}). \quad (4.29)$$

Assume furthermore that $u_0 \in \dot{H}^2$ and Assumption 2.3 holds, then (4.29) is valid with $\gamma = 1$.

5 Milstein–Galerkin Scheme

Our main aim in this last section is to construct a temporal higher-order Galerkin-based fully discrete Milstein scheme for Eq. (1.8) based on Itô–Taylor expansion on the diffusion term. We will give the proof of Theorem 1.2 at the end of Section 5.2.

The Milstein scheme of the finite element approximation (4.5) is to find a V_h -valued discrete process $\{U_h^m : m \in \mathbb{Z}_M\}$ such that

$$U_h^{m+1} = U_h^m + \tau A_h U_h^{m+1} + \tau \mathcal{P}_h F(U_h^{m+1}) + \mathcal{P}_h G(U_h^m) \delta_m W + \mathcal{P}_h DG(U_h^m) G(U_h^m) \left[\int_{t_m}^{t_{m+1}} (W(r) - W(t_m)) dW(r) \right], \quad (\text{DIEMG})$$

with initial datum $U_h^0 = \mathcal{P}_h u_0$. This scheme is derived by adding the last term in (DIEMG) to the DIEG scheme (DIEG) and followed from an Itô formula to $G(u)$. We call (DIEMG) the DIE Milstein Galerkin (DIEMG) scheme. This type of DIEMG scheme had been studied and reviewed in e.g. [17], [20], and references therein.

It is clear that the DIEMG scheme (DIEMG) is equivalent to the following compact scheme for $m \in \mathbb{Z}_{M-1}$:

$$U_h^{m+1} = S_{h,\tau} U_h^m + \tau S_{h,\tau} \mathcal{P}_h F(U_h^{m+1}) + S_{h,\tau} \mathcal{P}_h G(U_h^m) \delta_m W + S_{h,\tau} \mathcal{P}_h DG(U_h^m) G(U_h^m) \left[\int_{t_m}^{t_{m+1}} (W(r) - W(t_m)) dW(r) \right], \quad (5.1)$$

with initial datum $U_h^0 = \mathcal{P}_h u_0$, where $S_{h,\tau}$ is given in the previous section. Iterating (5.1) for m -times, we obtain

$$U_h^{m+1} = S_{h,\tau}^{m+1} u_h^0 + \tau \sum_{i=0}^m S_{h,\tau}^{m+1-i} \mathcal{P}_h F(U_h^{i+1}) + \sum_{i=0}^m S_{h,\tau}^{m+1-i} \mathcal{P}_h G(U_h^i) \delta_i W + \sum_{i=0}^m S_{h,\tau}^{m+1-i} \mathcal{P}_h DG(U_h^i) G(U_h^i) \left[\int_{t_i}^{t_{i+1}} (W(r) - W(t_i)) dW(r) \right]. \quad (5.2)$$

It should be noted that in the case of additive noise, the last term of the DIEMG scheme (DIEMG) vanishes and thus it reduces to the DIEG scheme (DIEG) (with $G(u) \equiv G$ for any $u \in H$). In the following, we first show the DIEMG scheme (DIEMG) or equivalently, DIEG scheme (DIEG) applied to Eq. (1.8) with additive noise possesses a temporal higher-order of convergence. Then we pass to the general multiplicative noise which satisfies the additional Assumption 5.2.

5.1 Additive Noise

In the additive noise case, for simplicity, G is a time independent constant operator, the conditions (2.15)-(2.16) are equivalent to the assumption $G \in \mathcal{L}_2^1$ which was imposed in [34] as $\|(-A)^{1/2} \mathbf{Q}^{1/2}\|_{HS(H;H)} < \infty$ with $G \equiv \text{Id}$. We can show more temporal convergence order for the fully discrete scheme (DIEG), provided that the following growth condition for the second-order derivative of f holds.

Assumption 5.1 $f : \mathbb{R} \rightarrow \mathbb{R}$ is second time differentiable and there exist constants $L_f'' \in \mathbb{R}_+^*$ and $\tilde{q} \geq 3$ with $\tilde{q} \in [3, \infty)$ for $d = 1, 2$ and $\tilde{q} \in [3, 5]$ for $d = 3$ such that

$$|f''(x)| \leq L_f''(1 + |x|^{\tilde{q}-3}), \quad x \in \mathbb{R}. \quad (5.3)$$

The condition (5.3) is consistent with (2.4). In fact, for an odd $(q-1)$ -order polynomial (2.10) with q satisfying (2.5), the conditions (2.4) and (5.3) are valid with $\tilde{q} = q$ when $q \geq 3$ and $\tilde{q} = 3$ when $q < 3$ (i.e., $q = 2$).

Corollary 5.1 *Let $\gamma \in [0, 1]$, $u_0 \in \dot{H}^{1+\gamma}$, $G \in \mathcal{L}_2^1$, and Assumptions 2.1 and 5.1 hold. Let u and u_h^m be the solutions of Eq. (1.8) and (DIEG) with $G(u) \equiv G$ for any $u \in H$, respectively. Then there exists a constant C such that*

$$\sup_{m \in \mathbb{Z}_M} \|u(t_m) - u_h^m\|_{L_\omega^2 L_x^2} \leq C(h^{1+\gamma} + \tau^{\frac{1+\gamma}{2}}). \quad (5.4)$$

Proof Let $\gamma \in [0, 1]$ and $\gamma = 1$ when (2.17) holds. According to the proof in Theorem 1.1, it suffices to show the following refined estimation for any $p \geq 2$ between the exact solution u and the auxliary process \tilde{u}_h^m defined by (4.10):

$$\sup_{m \in \mathbb{Z}_M} \|u(t_m) - \tilde{u}_h^m\|_{L_\omega^p L_x^2} \leq C(h^{1+\gamma} + \tau^{\frac{1+\gamma}{2}}). \quad (5.5)$$

By means of the proof in Lemma 4.2, we only need to give the following two refined estimations for J_{21}^{m+1} and J_3^{m+1} under the additional condition (5.3):

$$J_{21}^{m+1} = \left\| \sum_{i=0}^m \int_{t_i}^{t_{i+1}} S(t_{m+1}-r) [F(u(r)) - F(u(t_{i+1}))] dr \right\|_{L_\omega^p L_x^2} \leq C\tau^{\frac{1+\gamma}{2}}, \quad (5.6)$$

$$J_3^{m+1} = \left\| \int_0^{t_{m+1}} E_{h,\tau}(t_{m+1}-r) G dW(r) \right\|_{L_\omega^p L_x^2} \leq C(h^{1+\gamma} + \tau^{\frac{1+\gamma}{2}}). \quad (5.7)$$

The last estimation (5.7) can be handled by the Burkholder–Davis–Gundy inequality (2.22) and (4.9) with $\mu = \gamma$:

$$J_3^{m+1} \leq C \left(\int_0^{t_{m+1}} \|E_{h,\tau}(\sigma) G\|_{\mathcal{L}_2^0}^2 d\sigma \right)^{\frac{1}{2}} \leq C \|G\|_{\mathcal{L}_2^1} (h^{1+\gamma} + \tau^{\frac{1+\gamma}{2}}).$$

To show the first estimation (5.6), note that

$$\begin{aligned} u(t_{i+1}) &= S(t_{i+1}-r)u(r) + \int_r^{t_{i+1}} S(t_{i+1}-\sigma)F(u(\sigma))d\sigma \\ &\quad + \int_r^{t_{i+1}} S(t_{i+1}-\sigma)GdW(\sigma), \quad r \in [t_i, t_{i+1}). \end{aligned}$$

Then using the Taylor formula leads to the splitting of J_{21}^{m+1} into

$$\begin{aligned} J_{21}^{m+1} &\leq \left\| \sum_{i=0}^m \int_{t_i}^{t_{i+1}} S(t_{m+1}-r) DF(u(r)) (S(t_{i+1}-r) - \text{Id}) u(r) dr \right\|_{L_\omega^p L_x^2} \\ &\quad + \left\| \sum_{i=0}^m \int_{t_i}^{t_{i+1}} S(t_{m+1}-r) DF(u(r)) \left[\int_r^{t_{i+1}} S(t_{i+1}-\sigma) F(u(\sigma)) d\sigma \right] dr \right\|_{L_\omega^p L_x^2} \\ &\quad + \left\| \sum_{i=0}^m \int_{t_i}^{t_{i+1}} S(t_{m+1}-r) DF(u(r)) \left[\int_r^{t_{i+1}} S(t_{i+1}-\sigma) G dW(\sigma) \right] dr \right\|_{L_\omega^p L_x^2} \end{aligned}$$

$$+ \left\| \sum_{i=0}^m \int_{t_i}^{t_{i+1}} S(t_{m+1}-r) R_F(u(r), u(t_{i+1})) dr \right\|_{L_\omega^p L_x^2} =: \sum_{i=1}^4 J_{21i}^{m+1},$$

where R_F denotes the remainder term

$$\begin{aligned} & R_F(u(r), u(t_{i+1})) \\ &:= \int_0^1 D^2 F(u(r) + \lambda(u(t_{i+1}) - u(r))) (u(t_{i+1}) - u(r), u(t_{i+1}) - u(r)) (1 - \lambda) d\lambda \\ &= \int_0^1 f''(u(r) + \lambda(u(t_{i+1}) - u(r))) (u(t_{i+1}) - u(r))^2 (1 - \lambda) d\lambda. \end{aligned}$$

342 We shall estimate J_{21i}^{m+1} , $i = 1, 2, 3, 4$, successively.

Since for $\delta \in (3/2, 2)$, $\dot{H}^\delta \hookrightarrow \mathcal{C}$, it follows by the dual argument that

$$\|x\|_{-\delta} \leq C \|x\|_{L_x^1}, \quad x \in L_x^1. \quad (5.8)$$

This inequality, in conjunction with Minkovskii and Cauchy–Schwarz inequalities, the condition (2.4), the embedding (2.6), and the estimations (2.2), (3.2), (3.20), and (3.27), yields that

$$\begin{aligned} J_{211}^{m+1} &\leq \sum_{i=0}^m \int_{t_i}^{t_{i+1}} \|(-A)^{\frac{\delta}{2}} S(t_{m+1}-r) \|_{\mathcal{L}(H)} \|DF(u(r))(S(t_{i+1}-r) - \text{Id})u(r)\|_{L_\omega^p \dot{H}^{-\delta}} dr \\ &\leq C \sum_{i=0}^m \int_{t_i}^{t_{i+1}} (t_{m+1}-r)^{-\frac{\delta}{2}} \|DF(u(r))(S(t_{i+1}-r) - \text{Id})u(r)\|_{L_\omega^p L_x^1} dr \\ &\leq C \sum_{i=0}^m \int_{t_i}^{t_{i+1}} (t_{m+1}-r)^{-\frac{\delta}{2}} \|f'(u(r))\|_{L_\omega^{2p} L_x^2} \|(S(t_{i+1}-r) - \text{Id})u(r)\|_{L_\omega^{2p} L_x^2} dr \\ &\leq C \tau^{\frac{1+\gamma}{2}} \|f'(u)\|_{L_t^\infty L_\omega^{2p} L_x^2} \|u\|_{L_t^\infty L_\omega^{2p} \dot{H}^{1+\gamma}} \left(\int_0^{t_{m+1}} r^{-\frac{\delta}{2}} dr \right) \\ &\leq C \tau^{\frac{1+\gamma}{2}} \times \begin{cases} (1 + \|u_0\|_1^{q-2})(1 + \|u_0\|_1), & \gamma = 0; \\ (1 + \|u_0\|_1^{q-2})(1 + \|u_0\|_{1+\gamma}^{q-1}), & \gamma \in (0, 1); \\ (1 + \|u_0\|_1^{q-2})(1 + \|u_0\|_2^{\rho(q-1)} + \|u_0\|_2^{(q-1)^2}), & \gamma = 1. \end{cases} \end{aligned} \quad (5.9)$$

Similar argument implies that

$$\begin{aligned} J_{212}^{m+1} &\leq C \sum_{i=0}^m \int_{t_i}^{t_{i+1}} \int_r^{t_{i+1}} (t_{m+1}-r)^{-\frac{\delta}{2}} \|f'(u(r))\|_{L_\omega^{2p} L_x^2} \|f(u(\sigma))\|_{L_\omega^{2p} L_x^2} d\sigma dr \\ &\leq C \tau \|f'(u)\|_{L_t^\infty L_\omega^{2p} L_x^2} \|f(u)\|_{L_t^\infty L_\omega^{2p} L_x^2} \left(\int_0^{t_{m+1}} r^{-\frac{\delta}{2}} dr \right) \\ &\leq C \tau (1 + \|u_0\|_1^{q-2})(1 + \|u_0\|_1^{q-1}). \end{aligned} \quad (5.10)$$

To estimate the third term J_{213}^{m+1} , we apply stochastic Fubini theorem and the discrete and continuous Burkholder–Davis–Gundy inequalities (2.23) and (2.22), respectively, to derive

$$J_{213}^{m+1} = \left\| \sum_{i=0}^m \int_{t_i}^{t_{i+1}} \left[\int_{t_i}^{t_{i+1}} \chi_{[r, t_{i+1})}(\sigma) S(t_{m+1}-r) DF(u(r)) S(t_{i+1}-\sigma) G dr \right] dW(\sigma) \right\|_{L_\omega^p L_x^2}$$

$$\leq C \left(\sum_{i=0}^m \int_{t_i}^{t_{i+1}} \left\| \int_{t_i}^{t_{i+1}} S(t_{m+1}-r) DF(u(r)) S(t_{i+1}-\sigma) G dr \right\|_{L_{\omega}^p \mathcal{L}_2^0}^2 d\sigma \right)^{\frac{1}{2}}.$$

Then by Cauchy–Schwarz inequality, the condition (2.4), and the estimation (3.2), we obtain

$$\begin{aligned} J_{213}^{m+1} &\leq C \tau^{\frac{1}{2}} \left(\sum_{i=0}^m \int_{t_i}^{t_{i+1}} \int_{t_i}^{t_{i+1}} \|S(t_{m+1}-r) DF(u(r)) S(t_{i+1}-\sigma) G\|_{L_{\omega}^p \mathcal{L}_2^0}^2 dr d\sigma \right)^{\frac{1}{2}} \\ &\leq C \tau^{\frac{1}{2}} \left[\sum_{i=0}^m \left(\int_{t_i}^{t_{i+1}} \|f'(u(r))\|_{L_{\omega}^p \dot{H}^{-1}}^2 dr \right) \left(\int_{t_i}^{t_{i+1}} \|S(\sigma) G\|_{\mathcal{L}_2^1}^2 d\sigma \right) \right]^{\frac{1}{2}} \\ &\leq C \tau \|f'(u)\|_{L_t^{\infty} L_{\omega}^{2p} \dot{H}^{-1}} \|G\|_{\mathcal{L}_2^1} \leq C(1 + \|u_0\|_1^{q-2}) \tau. \end{aligned} \quad (5.11)$$

Finally, in terms of the estimations (2.2) and (5.8) and Minkovskii and Hölder inequalities, we estimate the last term J_{214}^{m+1} similarly to J_{211}^{m+1} by

$$\begin{aligned} J_{214}^{m+1} &\leq C \sum_{i=0}^m \int_{t_i}^{t_{i+1}} (t_{m+1}-r)^{-\frac{\delta}{2}} \|R_F(u(r), u(t_{i+1}))\|_{L_{\omega}^p L_x^1} dr \\ &\leq C \sum_{i=0}^m \int_{t_i}^{t_{i+1}} (t_{m+1}-r)^{-\frac{\delta}{2}} \times \int_0^1 \|f''(u(r) + \lambda(u(t_{i+1}) - u(r)))\|_{L_{\omega}^{3p} L_x^3} \\ &\quad \|u(t_{i+1}) - u(r)\|_{L_{\omega}^{3p} L_x^3}^2 (1-\lambda) d\lambda dr. \end{aligned}$$

It follows from the condition (5.3), the embeddings (2.6) and $\dot{H}^{1/2} \hookrightarrow L_x^3$, the moments' estimation (3.2), and the Hölder estimation (3.29) that

$$\begin{aligned} J_{214}^{m+1} &\leq C \tau^{(\frac{1}{2}+\gamma) \wedge 1} \left(1 + \|u\|_{L_t^{\infty} L_{\omega}^{3p(\tilde{q}-3)} L_x^{3(\tilde{q}-3)}}^{\tilde{q}-3} \right) \left(\int_0^{t_{m+1}} r^{-\frac{\delta}{2}} dr \right) \\ &\quad \times \left[\sup_{0 \leq s < t \leq T} \frac{\|u(t) - u(s)\|_{L_{\omega}^{3p} \dot{H}^{1/2}}^2}{|t-s|^{(1/2+\gamma) \wedge 1}} \right] \\ &\leq C \tau^{(\frac{1}{2}+\gamma) \wedge 1} \times \begin{cases} (1 + \|u_0\|_1^{\tilde{q}+q-4}) (1 + \|u_0\|_1^{q-1}), & \gamma \in [0, 1); \\ (1 + \|u_0\|_1^{\tilde{q}+q-4}) (1 + \|u_0\|_2^{\rho(q-1)} + \|u_0\|_2^{(q-1)^2}), & \gamma = 1. \end{cases} \end{aligned} \quad (5.12)$$

Collecting the above four estimations (5.9)–(5.12) together, we obtain (5.6) and complete the proof.

5.2 Strong Convergence Rate of DIEMG Scheme

Next, we consider the DIEMG scheme (DIEMG) for Eq. (1.8) with general multiplicative noise under an infinite-dimensional analog of the commutativity type condition in the SODE setting. Such commutativity condition was first introduced in [17, Assumption 3] where the authors observed in [17, Section 4] that a certain class of semilinear SPDEs with multiplicative trace class noise naturally fulfills such commutativity condition. The following assumption holds true automatically in the case of the additive noise case $G(u) \equiv G \in \mathcal{L}_2^0$ for any $u \in H$.

Assumption 5.2 $G : H \rightarrow \mathcal{L}_2^0$ be a twice continuously Fréchet differentiable mapping such that for every $z \in \dot{H}^1$, $DG(z) \in \mathcal{L}(H; \mathcal{L}_2^0)$ and $D^2G(z) \in \mathcal{L}^{\otimes 2}(\dot{H}^\beta; \mathcal{L}_2^0)$ for some $\beta \in [0, \frac{1+\gamma}{2}]$. Moreover, there exists a constant $L'_g \in \mathbb{R}_+$ such that

$$\|DG(z)\|_{\mathcal{L}(L_x^2; \mathcal{L}_2^0)} + \|D^2G(z)\|_{\mathcal{L}^{\otimes 2}(\dot{H}^\beta; \mathcal{L}_2^0)} \leq L'_g, \quad z \in \dot{H}^1, \quad (5.13)$$

$$\|DG(u)G(u) - DG(v)G(v)\|_{HS(U_0; \mathcal{L}_2^0)} \leq L'_g \|u - v\|, \quad u, v \in H. \quad (5.14)$$

Example 5.1 Assume that $g : \mathbb{R} \rightarrow \mathbb{R}$ a twice continuously differentiable function such that

$$\sup_{x \in \mathbb{R}} |g'(x)| + \sup_{x \in \mathbb{R}} |g''(x)| + \sup_{x, y \in \mathbb{R}, x \neq y} \frac{\|g'(x)g(x) - g'(y)g(y)\|}{|x - y|} < \infty, \quad (5.15)$$

$g_k \in W^{1, \infty}$ for each $k \in \mathbb{N}_+$ such that

$$\sum_{k \in \mathbb{N}_+} \lambda_k^Q \|g_k\|_{W^{1, \infty}}^2 < \infty, \quad (5.16)$$

and that either $g(0) = 0$ or g_k vanishes on the boundary $\partial \mathcal{O}$, then Assumptions 2.2 and 5.2 hold true. In fact, for any $u, v \in L_x^2$ and $z \in \dot{H}^1$, we have

$$\begin{aligned} \|G(u) - G(v)\|_{\mathcal{L}_2^0} &= \left(\sum_{k \in \mathbb{N}_+} \lambda_k^Q \|(g(u) - g(v))g_k\|^2 \right)^{\frac{1}{2}} \\ &\leq \sup_{x \in \mathbb{R}} |g'(x)| \left(\sum_{k \in \mathbb{N}_+} \lambda_k^Q \|g_k\|_{L_x^\infty}^2 \right)^{\frac{1}{2}} \|u - v\|, \end{aligned}$$

and

$$\begin{aligned} \|G(z)\|_{\mathcal{L}_2^0}^2 &\leq C \sum_{k \in \mathbb{N}_+} \lambda_k^Q \left(\|g(z)g_k\|^2 + \|g'(z)\nabla z g_k + g(z)g'_k\|^2 \right) \\ &\leq C \left(1 + \sup_{x \in \mathbb{R}} |g'(x)|^2 \right) \sum_{k \in \mathbb{N}_+} \lambda_k^Q \|g_k\|_{W_x^{1, \infty}}^2 (1 + \|z\|_1)^2, \end{aligned}$$

which show Assumption 2.2. On the other hand, for any $z \in \dot{H}^1$,

$$\begin{aligned} &\|DG(z)\|_{\mathcal{L}(L_x^2; \mathcal{L}_2^0)} + \|D^2G(z)\|_{\mathcal{L}^{\otimes 2}(\dot{H}^\beta; \mathcal{L}_2^0)} \\ &= \sup_{v \in L_x^2} \frac{\left(\sum_{k \in \mathbb{N}_+} \lambda_k^Q \|g'(z)vg_k\|^2 \right)^{\frac{1}{2}}}{\|v\|} + \sup_{u, v \in \dot{H}^\beta} \frac{\left(\sum_{k \in \mathbb{N}_+} \lambda_k^Q \|(g''(z)uv)g_k\|^2 \right)^{\frac{1}{2}}}{\|u\|_\beta \|v\|_\beta} \\ &\leq C \left(\sum_{k \in \mathbb{N}_+} \lambda_k^Q \|g_k\|_{L_x^\infty}^2 \right)^{\frac{1}{2}} \left[\sup_{x \in \mathbb{R}} |g'(x)| + \sup_{x \in \mathbb{R}} |g''(x)| \cdot \sup_{u, v \in \dot{H}^\beta} \frac{\|u\|_{L_x^4} \|v\|_{L_x^4}}{\|v\|_\beta \|z\|_\beta} \right]. \end{aligned}$$

Due to the assumptions (5.15)-(5.16) and the embedding $\dot{H}^\beta \hookrightarrow L_x^4$ with $\beta = 1/2$ when $d = 1, 2$ and $\beta = 3/4$ when $d = 3$, we have

$$\|DG(z)\|_{\mathcal{L}(L_x^2; \mathcal{L}_2^0)} + \|D^2G(z)\|_{\mathcal{L}^{\otimes 2}(\dot{H}^\beta; \mathcal{L}_2^0)}$$

$$\leq C \sup_{v \in L_x^2} \left(\sum_{k \in \mathbb{N}_+} \lambda_k^Q \|g_k\|_{L_x^\infty}^2 \right)^{\frac{1}{2}} \left[\sup_{x \in \mathbb{R}} |g'(x)| + \sup_{x \in \mathbb{R}} |g''(x)| \right].$$

Moreover, for any $u, v \in L_x^2$,

$$\begin{aligned} & \|DG(u)G(u) - DG(v)G(v)\|_{HS(U_0; \mathcal{L}_2^0)} \\ &= \left(\sum_{k \in \mathbb{N}_+} \sum_{l \in \mathbb{N}_+} \lambda_k^Q \lambda_l^Q \|DG(u)(G(u)g_k)g_l - DG(v)(G(v)g_l)g_k\|^2 \right)^{\frac{1}{2}} \\ &\leq \left(\sum_{k \in \mathbb{N}_+} \lambda_k^Q \|g_k\|_{L_x^\infty}^2 \right) \|g'(u)g(u) - g'(v)g(v)\| \leq \left(\sum_{k \in \mathbb{N}_+} \lambda_k^Q \|g_k\|_{L_x^\infty}^2 \right) \|u - v\|. \end{aligned}$$

The above two estimations show Assumption 5.2 with $\gamma \in [0, 1]$ and $\beta \in [0, 1]$ when $d = 1, 2$ and $\gamma \in [1/2, 1]$ and $\beta \in [0, (1 + \gamma)/2]$ when $d = 3$.

We introduce a version \tilde{U}_h^{m+1} of the auxiliary process \tilde{u}_h^{m+1} given by (4.10) in consistent with the DIEMG scheme (4.6), which is defined as

$$\begin{aligned} \tilde{U}_h^{m+1} &= S_{h,\tau}^{m+1} u_h^0 + \tau \sum_{i=0}^m S_{h,\tau}^{m+1-i} \mathcal{P}_h F(u(t_{i+1})) + \sum_{i=0}^m S_{h,\tau}^{m+1-i} \mathcal{P}_h G(u(t_i)) \delta_i W \\ &\quad + \sum_{i=0}^m S_{h,\tau}^{m+1-i} \mathcal{P}_h DG(u(t_i)) G(u(t_i)) \left[\int_{t_i}^{t_{i+1}} (W(r) - W(t_i)) dW(r) \right], \quad (5.17) \end{aligned}$$

for $m \in \mathbb{Z}_{M-1}$, where the terms U_h^{i+1} and U_h^i in the discrete deterministic and stochastic convolutions of (5.2) are replaced by $u(t_{i+1})$ and $u(t_i)$, respectively. The boundedness (3.2) and (4.12) imply the solvability and uniform boundedness of \tilde{U}_h^m :

$$\sup_{m \in \mathbb{Z}_M} \mathbb{E} \left[\|\tilde{U}_h^m\|_1^p \right] \leq C \left(1 + \|u_0\|_1^{p(q-1)} \right), \quad (5.18)$$

provided $u_0 \in \dot{H}^1$ and Assumptions 2.1-2.2 and 5.2 hold. We begin with the strong error estimation between u and $\{\tilde{U}_h^m\}_{m \in \mathbb{Z}_M}$.

Lemma 5.1 Let $\gamma \in [0, 1]$, $u_0 \in \dot{H}^{1+\gamma}$, and Assumptions 2.1-2.2 and 5.1-5.2 hold. Then for any $p \geq 1$, there exists a constant C such that

$$\sup_{m \in \mathbb{Z}_M} \|u(t_m) - \tilde{U}_h^m\|_{L_\omega^p L_x^2} \leq C(h^{1+\gamma} + \tau^{\frac{1+\gamma}{2}}). \quad (5.19)$$

Assume furthermore that $u_0 \in \dot{H}^2$ and Assumption 2.3 holds, then (5.19) is valid with $\gamma = 1$.

Proof For $m \in \mathbb{Z}_{M-1}$, define $\tilde{J}^{m+1} := \|u(t_{m+1}) - \tilde{U}_h^{m+1}\|_{L_\omega^p L_x^2}$. By the proofs of Lemma 4.2 and Corollary 5.1, we have

$$\tilde{J}^{m+1} \leq \sum_{i=1}^2 J_i^{m+1} + \tilde{J}_3^{m+1},$$

where J_i^{m+1} , $i = 1, 2$, are defined from (4.14) and satisfy the estimate

$$J_1^{m+1} + J_2^{m+1} \leq C(h^{1+\gamma} + \tau^{\frac{1+\gamma}{2}}), \quad (5.20)$$

by (4.15), (4.17), (4.20), and (5.6), and the last term is defined as

$$\begin{aligned} \tilde{J}_3^{m+1} := & \left\| \int_0^{t_{m+1}} S(t_{m+1} - r) G(u(r)) dW(r) - \sum_{i=0}^m S_{h,\tau}^{m+1-i} \mathcal{P}_h G(u(t_i)) \delta_i W \right. \\ & \left. - \sum_{i=0}^m S_{h,\tau}^{m+1-i} \mathcal{P}_h DG(u(t_i)) G(u(t_i)) \left[\int_{t_i}^{t_{i+1}} (W(r) - W(t_i)) dW(r) \right] \right\|_{L_\omega^p L_x^2}. \end{aligned}$$

Thus to prove (5.19), it suffices to show that

$$\tilde{J}_3^{m+1} \leq C(h^{1+\gamma} + \tau^{\frac{1+\gamma}{2}}). \quad (5.21)$$

The term \tilde{J}_3^{m+1} can be decomposed as

$$\begin{aligned} \tilde{J}_3^{m+1} \leq & \left\| \sum_{i=0}^m \int_{t_i}^{t_{i+1}} H_i^{m+1}(r) dW(r) \right\|_{L_\omega^p L_x^2} \\ & + \left\| \int_0^{t_{m+1}} E_{h,\tau}(t_{m+1} - r) G(u(r)) dW(r) \right\|_{L_\omega^p L_x^2} := \tilde{J}_{31}^{m+1} + J_{32}^{m+1}, \end{aligned}$$

where

$$H_i^{m+1}(r) = S_{h,\tau}^{m+1-i} \mathcal{P}_h [G(u(r)) - G(u(t_i)) - DG(u(t_i)) G(u(t_i)) (W(r) - W(t_i))].$$

By (4.19) and (4.21), the second term J_{32}^{m+1} has the estimation

$$J_{32}^{m+1} \leq C(h^{1+\gamma} + \tau^{\frac{1+\gamma}{2}}) \times \begin{cases} (1 + \|u_0\|_1^{q-1}), & \gamma \in [0, 1), \\ (1 + \|u_0\|_2^{q-1}), & \gamma = 1. \end{cases} \quad (5.22)$$

It remains to show a similar estimation for \tilde{J}_{31}^{m+1} . It is clear that

$$\begin{aligned} H_i^{m+1}(r) = & S_{h,\tau}^{m+1-i} \mathcal{P}_h DG(u(t_i)) [u(r) - u(t_i) - G(u(t_i)) (W(r) - W(t_i))] \\ & + S_{h,\tau}^{m+1-i} \mathcal{P}_h \int_0^1 D^2 G(u(t_i) + \lambda(u(r) - u(t_i))) \\ & (u(r) - u(t_i), u(r) - u(t_i)) (1 - \lambda) d\lambda, \quad r \in [t_i, t_{i+1}). \end{aligned}$$

As a consequence of the discrete and continuous Burkholder–Davis–Gundy inequalities (2.23) and (2.22), respectively, we have

$$\tilde{J}_{31}^{m+1} \leq \sum_{i=1}^2 \sqrt{\tilde{J}_{31i}^{m+1}}, \quad (5.23)$$

where

$$\tilde{J}_{31i}^{m+1} := \sum_{i=0}^m \int_{t_i}^{t_{i+1}} \left\| S_{h,\tau}^{m+1-i} \mathcal{P}_h DG(u(t_i)) \left[u(r) - u(t_i) - \int_{t_i}^r G(u(t_i)) dW(s) \right] \right\|_{L_\omega^p \mathcal{L}_2^0}^2 dr,$$

$$\begin{aligned} \tilde{J}_{312}^{m+1} := & \sum_{i=0}^m \int_{t_i}^{t_{i+1}} \left\| S_{h,\tau}^{m+1-i} \mathcal{P}_h \int_0^1 D^2 G(u(t_i) + \lambda(u(r) - u(t_i))) \right. \\ & \left. (u(r) - u(t_i), u(r) - u(t_i)) (1 - \lambda) d\lambda \right\|_{L_\omega^p \mathcal{L}_2^0}^2 dr. \end{aligned}$$

We begin with the first term \tilde{J}_{311}^{m+1} . By the boundedness (2.2), the condition (5.13), the mild formulation (2.25), and Hölder inequality, we have

$$\begin{aligned} \tilde{J}_{311}^{m+1} & \leq C \sum_{i=0}^m \int_{t_i}^{t_{i+1}} \|u(r) - u(t_i) - G(u(t_i))(W(r) - W(t_i))\|_{L_\omega^p L_x^2}^2 dr \\ & \leq C \sum_{i=0}^m \int_{t_i}^{t_{i+1}} \|(S(r - t_i) - \text{Id})u(t_i)\|_{L_\omega^p L_x^2}^2 dr \\ & \quad + C \sum_{i=0}^m \int_{t_i}^{t_{i+1}} \left\| \int_{t_i}^r S(r - t_i) F(u(s)) ds \right\|_{L_\omega^p L_x^2}^2 dr \\ & \quad + C \sum_{i=0}^m \int_{t_i}^{t_{i+1}} \left\| \int_{t_i}^r (S(r - s) - \text{Id}) G(u(s)) dW(s) \right\|_{L_\omega^p L_x^2}^2 dr \\ & \quad + C \sum_{i=0}^m \int_{t_i}^{t_{i+1}} \left\| \int_{t_i}^r [G(u(s)) - G(u(t_i))] dW(s) \right\|_{L_\omega^p L_x^2}^2 dr. \end{aligned}$$

The conditions (2.9) and (2.15), the embedding (2.6), and the estimations (2.2), (3.2), (3.13), (3.20), and (3.27) imply that

$$\begin{aligned} \tilde{J}_{311}^{m+1} & \leq C \tau^{1+\gamma} \|u\|_{L_t^\infty L_\omega^p \dot{H}^{1+\gamma}}^2 + C \tau^2 \left(1 + \|u\|_{L_t^\infty L_\omega^{p(q-1)} \dot{H}^1}^{q-1} \right)^2 \\ & \quad + C \tau^2 \left(1 + \|u\|_{L_t^\infty L_\omega^p \mathcal{L}_2^1} \right)^2 + C \tau^2 \sup_{0 \leq s < t \leq T} \frac{\|u(t) - u(s)\|_{L_\omega^p L_x^2}^2}{(t - s)} \\ & \leq C \tau^{1+\gamma} \times \begin{cases} (1 + \|u_0\|_1^{q-1})^2, & \gamma \in [0, 1), \\ (1 + \|u_0\|_2^{p(q-1)} + \|u_0\|_2^{(q-1)^2})^2, & \gamma = 1. \end{cases} \end{aligned} \quad (5.24)$$

Next, we estimate the second term \tilde{J}_{312}^{m+1} by the boundedness (4.12), (5.13) with $\beta \in [0, \frac{1+\gamma}{2}]$, Cauchy–Schwarz inequality, and the Hölder estimation (3.13):

$$\begin{aligned} \tilde{J}_{312}^{m+1} & \leq C \sup_{z \in \dot{H}^1} \|D^2 G(z)\|_{\mathcal{L}^{\beta,2}(\dot{H}^\beta; \mathcal{L}_2^0)}^2 \left[\sum_{i=0}^m \int_{t_i}^{t_{i+1}} \|u(r) - u(t_i)\|_{L_\omega^{2p} \dot{H}^\beta}^4 dr \right] \\ & \leq C \sup_{0 \leq s < t \leq T} \frac{\|u(t) - u(s)\|_{L_\omega^{2p} L_x^2}^4}{|t - s|^{2(1+\gamma-\beta)\wedge 2}} \left[\sum_{i=0}^m \int_{t_i}^{t_{i+1}} (r - t_i)^{2(1+\gamma-\beta)\wedge 2} dr \right] \\ & \leq C \tau^{2(1+\gamma-\beta)\wedge 2} (1 + \|u_0\|_1^{q-1})^4 \leq C \tau^{1+\gamma} (1 + \|u_0\|_1^{q-1})^4. \end{aligned} \quad (5.25)$$

Collecting the estimations (5.24)–(5.25), we derive

$$\tilde{J}_{31}^{m+1} \leq C \tau^{\frac{1+\gamma}{2}}. \quad (5.26)$$

The above estimation (5.26), in combination with (5.22), shows (5.21) and thus completes the proof.

Remark 5.1 Under the conditions of Lemma 5.1 with $\gamma \in [0, 1]$ and $\beta \in [0, 1]$, the proof of Lemma 5.1 yields that, for any $p \geq 1$, there exists a constant C such that

$$\sup_{m \in \mathbb{Z}_M} \|u(t_m) - \tilde{U}_h^m\|_{L_\omega^p L_x^2} \leq C(h^{1+\gamma} + \tau^{\frac{1+\gamma}{2} \wedge (1+\gamma-\beta)}).$$

Now we can prove Theorem 1.2 by combining Lemma 5.1 with a variational approach as in the proof of Theorem 1.1.

Proof (Proof of Theorem 1.2) Let $m \in \mathbb{Z}_{M-1}$. In terms of Minkovskii inequality and the estimation (5.19), it suffices to prove that

$$\sup_{m \in \mathbb{Z}_M} \|\tilde{U}_h^m - U_h^m\|_{L_\omega^p L_x^2} \leq C(h^{1+\gamma} + \tau^{\frac{1+\gamma}{2}}), \quad \gamma \in [0, 1]. \quad (5.27)$$

Define $\hat{e}_h^{m+1} := \tilde{U}_h^{m+1} - U_h^{m+1}$. Then $\hat{e}_h^{m+1} \in V_h$ with vanishing initial datum $\hat{e}_h^0 = 0$. In terms of Eq. (DIEMG) and (5.17), it is clear that

$$\begin{aligned} \tilde{U}_h^{m+1} &= \tilde{U}_h^m + \tau A_h \tilde{U}_h^{m+1} + \tau \mathcal{P}_h F(u(t_{m+1})) + \mathcal{P}_h G(u(t_m)) \delta_m W \\ &\quad + \mathcal{P}_h D G(u(t_m)) G(u(t_m)) \left[\int_{t_m}^{t_{m+1}} (W(r) - W(t_m)) dW(r) \right], \\ U_h^{m+1} &= U_h^m + \tau A_h U_h^{m+1} + \tau \mathcal{P}_h F(U_h^{m+1}) + \mathcal{P}_h G(U_h^m) \delta_m W \\ &\quad + \mathcal{P}_h D G(U_h^m) G(U_h^m) \left[\int_{t_m}^{t_{m+1}} (W(r) - W(t_m)) dW(r) \right]. \end{aligned}$$

Consequently,

$$\begin{aligned} \hat{e}_h^{m+1} - \hat{e}_h^m &= \tau A_h \hat{e}_h^{m+1} + \tau \mathcal{P}_h (F(u(t_{m+1})) - F(U_h^{m+1})) \\ &\quad + \mathcal{P}_h (G(u(t_m)) - G(U_h^m)) \delta_m W + \int_{t_m}^{t_{m+1}} \mathcal{P}_h H^{m+1} dW(r), \end{aligned}$$

where

$$H^{m+1}(r) := \int_{t_m}^r [D G(u(t_m)) G(u(t_m)) - D G(U_h^m) G(U_h^m)] dW(s), \quad r \in [t_m, t_{m+1}).$$

Multiplying \hat{e}_h^{m+1} on both sides of the above equation as in (4.23), we obtain

$$\begin{aligned} \langle \hat{e}_h^{m+1} - \hat{e}_h^m, \hat{e}_h^{m+1} \rangle + \tau \|\nabla \hat{e}_h^{m+1}\|^2 &= \tau \langle F(u(t_{m+1})) - F(U_h^{m+1}), \hat{e}_h^{m+1} \rangle \\ &\quad + \langle (G(u(t_m)) - G(U_h^m)) \delta_m W, \hat{e}_h^{m+1} \rangle + \int_{t_m}^{t_{m+1}} \langle H^{m+1} dW(r), \hat{e}_h^{m+1} \rangle. \end{aligned} \quad (5.28)$$

Similarly to (4.25), we have

$$\mathbb{E} [\langle \hat{e}_h^{m+1} - \hat{e}_h^m, \hat{e}_h^{m+1} \rangle] = \frac{1}{2} \left(\mathbb{E} [\|\hat{e}_h^{m+1}\|^2] - \mathbb{E} [\|\hat{e}_h^m\|^2] \right) + \frac{1}{2} \mathbb{E} [\|\hat{e}_h^{m+1} - \hat{e}_h^m\|^2]. \quad (5.29)$$

By analogous arguments in the proof of Theorem 1.1, we have

$$\tau \langle F(u(t_{m+1})) - F(U_h^{m+1}), \hat{e}_h^{m+1} \rangle \leq \frac{1}{2} \tau \|\nabla \hat{e}_h^{m+1}\|^2 + L_f \tau \|\hat{e}_h^{m+1}\|^2$$

$$+ C\tau \|u(t_{m+1}) - \tilde{U}_h^{m+1}\|^2 (1 + \|u(t_{m+1})\|_1^{q-2} + \|\tilde{U}_h^{m+1}\|_1^{q-2}), \quad (5.30)$$

and

$$\begin{aligned} & \langle (G(u(t_m)) - G(U_h^m)) \delta_m W, \tilde{e}_h^{m+1} \rangle \\ & \leq \frac{1}{4} \mathbb{E} [\|\tilde{e}_h^{m+1} - \tilde{e}_h^m\|^2] + 2L_g^2 \mathbb{E} [\|u(t_m) - \tilde{U}_h^m\|^2] \tau + 2L_g^2 \mathbb{E} [\|\tilde{e}_h^m\|^2] \tau. \end{aligned} \quad (5.31)$$

To deal with the last stochastic term in (5.28), we use the martingale property of the stochastic integral, Cauchy–Schwarz inequality, and Itô isometry as in (5.24):

$$\mathbb{E} \left[\int_{t_m}^{t_{m+1}} \langle H^{m+1} dW(r), \tilde{e}_h^{m+1} \rangle \right] \leq \frac{1}{4} \mathbb{E} [\|\tilde{e}_h^{m+1} - \tilde{e}_h^m\|^2] + C \int_{t_m}^{t_{m+1}} \mathbb{E} [\|H^{m+1}\|_{\mathcal{L}_2^0}^2] dr.$$

By Itô isometry, the condition (5.14), and Cauchy–Schwarz inequality, we get

$$\begin{aligned} & \int_{t_m}^{t_{m+1}} \mathbb{E} [\|H^{m+1}\|_{\mathcal{L}_2^0}^2] dr \\ & = \int_{t_m}^{t_{m+1}} \mathbb{E} \left[\left\| \int_{t_m}^r [DG(U_h^m)G(U_h^m) - DG(u(t_m))G(u(t_m))] dW(s) \right\|_{\mathcal{L}_2^0}^2 \right] dr \\ & = \frac{1}{2} \tau^2 \mathbb{E} [\|DG(U_h^m)G(U_h^m) - DG(u(t_m))G(u(t_m))\|_{HS(U_0; \mathcal{L}_2^0)}^2] \\ & = \frac{(L'_g)^2}{2} \tau^2 \mathbb{E} [\|u(t_m) - U_h^m\|^2] \leq L_g'^2 \tau^2 \mathbb{E} [\|u(t_m) - \tilde{U}_h^m\|^2] + L_g'^2 \tau^2 \mathbb{E} [\|\tilde{e}_h^m\|^2], \end{aligned}$$

and thus

$$\begin{aligned} & \mathbb{E} \left[\int_{t_m}^{t_{m+1}} \langle H^{m+1} dW(r), \tilde{e}_h^{m+1} \rangle \right] \\ & \leq \frac{1}{4} \mathbb{E} [\|\tilde{e}_h^{m+1} - \tilde{e}_h^m\|^2] + L_g'^2 \tau^2 \mathbb{E} [\|u(t_m) - \tilde{U}_h^m\|^2] + L_g'^2 \tau^2 \mathbb{E} [\|\tilde{e}_h^m\|^2]. \end{aligned} \quad (5.32)$$

Now taking expectation in (5.28), we use (5.29)–(5.32) and Hölder inequality to get

$$\begin{aligned} & \frac{1}{2} \left(\mathbb{E} [\|\tilde{e}_h^{m+1}\|^2] - \mathbb{E} [\|\tilde{e}_h^m\|^2] \right) \\ & \leq L_f \mathbb{E} [\|\tilde{e}_h^{m+1}\|^2] \tau + C \left(\mathbb{E} [\|u(t_{m+1}) - \tilde{U}_h^{m+1}\|^4] \right)^{\frac{1}{2}} \tau + L_g'^2 \tau^2 \mathbb{E} [\|u(t_m) - \tilde{U}_h^m\|^2] \\ & \quad \times \left(1 + \left(\mathbb{E} [\|u(t_{m+1})\|_1^{2(q-2)}] \right)^{\frac{1}{2}} + \mathbb{E} [\|\tilde{U}_h^{m+1}\|_1^{2(q-2)}] \right)^{\frac{1}{2}} \\ & \quad + 2L_g^2 \mathbb{E} [\|u(t_m) - \tilde{U}_h^m\|^2] \tau + (2L_g^2 \tau + L_g'^2 \tau^2) \mathbb{E} [\|\tilde{e}_h^m\|^2]. \end{aligned}$$

Then by the estimations (3.2), (5.18), and (5.19), we obtain

$$(1 - 2L_f \tau) \mathbb{E} [\|\tilde{e}_h^{m+1}\|^2] \leq C\tau (h^{1+\gamma} + \tau^{\frac{1+\gamma}{2}})^2 + (1 + C\tau) \mathbb{E} [\|\tilde{e}_h^m\|^2].$$

We conclude (5.27) by the same arguments as in the proof of Theorem 1.1.

5.3 Milstein–Spectral Galerkin Scheme

Similarly to Section 4.3, the parallel result in Theorem 1.2 for DIEMG scheme (DIEMG) is also valid for the spectral Galerkin-based Milstein scheme of Eq. (1.8).

The DIE Milstein spectral Galerkin (DIEMSG) scheme of Eq. (1.8) is to find a V_h -valued discrete process $\{U_N^m : N \in \mathbb{N}_+, m \in \mathbb{Z}_M\}$ such that

$$U_N^{m+1} = U_N^m + \tau A_N U_N^{m+1} + \tau \mathcal{P}_N F(U_N^{m+1}) + \mathcal{P}_N G(U_N^m) \delta_m W + \mathcal{P}_N D G(U_N^m) G(U_N^m) \left[\int_{t_m}^{t_{m+1}} (W(r) - W(t_m)) dW(r) \right], \quad (\text{DIEMSG})$$

with initial datum $U_N^0 = \mathcal{P}_N u_0$. By the arguments in the proof of Theorem 1.2, one can show the following strong convergence results for this scheme with sharp rates.

Theorem 5.1 *Let $\gamma \in [0, 1)$, $u_0 \in \dot{H}^{1+\gamma}$, and Assumptions 2.1-2.2 and 5.1-5.2 hold. There exist a constant C such that*

$$\sup_{m \in \mathbb{Z}_M} \|u(t_m) - U_N^m\|_{L_\omega^p L_x^2} \leq C(N^{-\frac{1+\gamma}{d}} + \tau^{\frac{1+\gamma}{2}}). \quad (5.33)$$

Assume furthermore that $u_0 \in \dot{H}^2$ and Assumption 2.3 holds, then (5.33) is valid with $\gamma = 1$.

Acknowledgements

We thank the anonymous referee for very helpful remarks and suggestions. The first author is partially supported by Hong Kong RGC General Research Fund, No. 16307319, and the UGC - Research Infrastructure Grant, No. IRS20SC39. The second author is partially supported by Hong Kong RGC General Research Fund, No. 15325816, and the Hong Kong Polytechnic University Start-up Fund for New Recruits, No. 1-ZE33.

References

1. C. Bayer, P. Friz, S. Riedel, and J. Schoenmakers, *From rough path estimates to multilevel Monte Carlo*, SIAM J. Numer. Anal. **54** (2016), no. 3, 1449–1483. MR 3501366
2. S. Becker, B. Gess, A. Jentzen, and P. Kloeden, *Strong convergence rates for explicit space-time discrete numerical approximations of stochastic Allen–Cahn equations*, arXiv:1711.02423.
3. S. Becker and A. Jentzen, *Strong convergence rates for nonlinearity-truncated Euler-type approximations of stochastic Ginzburg–Landau equations*, Stochastic Process. Appl. **129** (2019), no. 1, 28–69. MR 3906990
4. C.-E. Bréhier, J. Cui, and J. Hong, *Strong convergence rates of semidiscrete splitting approximations for the stochastic Allen–Cahn equation*, IMA J. Numer. Anal. **39** (2019), no. 4, 2096–2134. MR 4019051
5. Y. Cao, J. Hong, and Z. Liu, *Approximating stochastic evolution equations with additive white and rough noises*, SIAM J. Numer. Anal. **55** (2017), no. 4, 1958–1981. MR 3686802
6. J. Cui and J. Hong, *Strong and weak convergence rates of a spatial approximation for stochastic partial differential equation with one-sided Lipschitz coefficient*, SIAM J. Numer. Anal. **57** (2019), no. 4, 1815–1841. MR 3984308

- 394 7. J. Cui, J. Hong, and Z. Liu, *Strong convergence rate of finite difference approximations for stochastic cubic Schrödinger equations*, J. Differential Equations **263** (2017), 3687–3713. MR 3670034
- 395 8. J. Cui, J. Hong, Z. Liu, and W. Zhou, *Strong convergence rate of splitting schemes for stochastic nonlinear Schrödinger equations*, J. Differential Equations **266** (2019), no. 9, 5625–5663. MR 3912762
- 396 9. P. Dörsek, *Semigroup splitting and cubature approximations for the stochastic Navier-Stokes equations*, SIAM J. Numer. Anal. **50** (2012), no. 2, 729–746. MR 2914284
- 399 10. X. Feng, Y. Li, and Y. Zhang, *Finite element methods for the stochastic Allen–Cahn equation with gradient-type multiplicative noises*, SIAM J. Numer. Anal. **55** (2017), no. 1, 194–216. MR 3600370
- 400 11. I. Gyöngy and A. Millet, *Rate of convergence of space time approximations for stochastic evolution equations*, Potential Anal. **30** (2009), no. 1, 29–64. MR 2465711
- 401 12. D. Higham, X. Mao, and A. Stuart, *Strong convergence of Euler-type methods for nonlinear stochastic differential equation*, SIAM J. Numer. Anal. **40** (2002), no. 3, 1041–1063. MR 1949404
- 402 13. J. Hong, C. Huang, and Z. Liu, *Optimal regularity of stochastic evolution equations in M -type 2 Banach spaces*, J. Differential Equations **267** (2019), no. 3, 1955–1971. MR 3945622
- 403 14. J. Hong and Z. Liu, *Well-posedness and optimal regularity of stochastic evolution equations with multiplicative noises*, J. Differential Equations **266** (2019), no. 8, 4712–4745. MR 3912731
- 404 15. M. Hutzenthaler and A. Jentzen, *On a perturbation theory and on strong convergence rates for stochastic ordinary and partial differential equations with non-globally monotone coefficients*, to appear at Ann. Probab. (arXiv:1401.0295).
- 405 16. ———, *Convergence of the stochastic Euler scheme for locally Lipschitz coefficients*, Found. Comput. Math. **11** (2011), no. 6, 657–706. MR 2859952
- 406 17. A. Jentzen and M. Röckner, *A Milstein scheme for SPDEs*, Found. Comput. Math. **15** (2015), no. 2, 313–362. MR 3320928
- 407 18. M. Kovács, S. Larsson, and F. Lindgren, *On the backward Euler approximation of the stochastic Allen–Cahn equation*, J. Appl. Probab. **52** (2015), no. 2, 323–338. MR 3372078
- 408 19. ———, *On the discretisation in time of the stochastic Allen–Cahn equation*, Math. Nachr. **291** (2018), no. 5–6, 966–995. MR 3795566
- 409 20. R. Kruse, *Consistency and stability of a Milstein–Galerkin finite element scheme for semilinear SPDE*, Stoch. Partial Differ. Equ. Anal. Comput. **2** (2014), no. 4, 471–516. MR 3274889
- 410 21. ———, *Optimal error estimates of Galerkin finite element methods for stochastic partial differential equations with multiplicative noise*, IMA J. Numer. Anal. **34** (2014), no. 1, 217–251. MR 3168284
- 411 22. N. Krylov, *A W_2^n -theory of the Dirichlet problem for SPDEs in general smooth domains*, Probab. Theory Related Fields **98** (1994), no. 3, 389–421. MR 1262972
- 412 23. J. Liu, *Order of convergence of splitting schemes for both deterministic and stochastic nonlinear Schrödinger equations*, SIAM J. Numer. Anal. **51** (2013), no. 4, 1911–1932. MR 3072234
- 413 24. W. Liu, *Well-posedness of stochastic partial differential equations with Lyapunov condition*, J. Differential Equations **255** (2013), no. 3, 572–592. MR 3053478
- 414 25. W. Liu and M. Röckner, *Stochastic partial differential equations: an introduction*, Universitext, Springer, Cham, 2015. MR 3410409
- 415 26. Z. Liu and Z. Qiao, *Strong approximation of stochastic Allen–Cahn equation with white noise*, to appear at IMA J. Num. Anal. (<https://doi.org/10.1093/imanum/dry088>).
- 416 27. ———, *Wong–Zakai approximation of stochastic Allen–Cahn equation*, Int. J. Numer. Anal. Model. **16** (2019), no. 5, 681–694. MR 3946652
- 417 28. A. Majee and A. Prohl, *Optimal strong rates of convergence for a space-time discretization of the stochastic Allen–Cahn equation with multiplicative noise*, Comput. Methods Appl. Math. **18** (2018), no. 2, 297–311. MR 3776047
- 418 29. C. Marinelli and M. Röckner, *On the maximal inequalities of Burkholder, Davis and Gundy*, Expo. Math. **34** (2016), no. 1, 1–26. MR 3463679
- 419 30. T. Müller-Gronbach and K. Ritter, *Lower bounds and nonuniform time discretization for approximation of stochastic heat equations*, Found. Comput. Math. **7** (2007), no. 2, 135–181. MR 2324415
- 420 31. Neelima and D. Šiška, *l^p -estimates and regularity for SPDEs with monotone semilinearity*, to appear at Stoch. Partial Differ. Equ. Anal. Comput. (<https://doi.org/10.1007/s40072-019-00150-w>).
- 421 32. A. Pazy, *Semigroups of linear operators and applications to partial differential equations*, Applied Mathematical Sciences, vol. 44, Springer-Verlag, New York, 1983. MR 710486
- 422 33. A. Prohl, *Strong rates of convergence for a space-time discretization of the stochastic Allen–Cahn equation with multiplicative noise*, preprint, <https://na.uni-tuebingen.de/preprints> (2014).
- 423 34. R. Qi and X. Wang, *Optimal error estimates of Galerkin finite element methods for stochastic Allen–Cahn equation with additive noise*, J. Sci. Comput. **80** (2019), no. 2, 1171–1194. MR 3977202
- 424 35. X. Wang, *An efficient explicit full discrete scheme for strong approximation of stochastic Allen–Cahn equation*, arXiv:1802.09413.