OPTIMAL STRONG RATES OF CONVERGENCE FOR A SPACE-TIME DISCRETIZATION OF THE STOCHASTIC ALLEN-CAHN EQUATION WITH MULTIPLICATIVE NOISE

ANANTA K. MAJEE AND ANDREAS PROHL

ABSTRACT. The stochastic Allen-Cahn equation with multiplicative noise involves the nonlinear drift operator $\mathscr{A}(x) = \Delta x - (|x|^2 - 1)x$. We use the fact that $\mathscr{A}(x) = -\mathcal{J}'(x)$ satisfies a weak monotonicity property to deduce uniform bounds in strong norms for solutions of the temporal, as well as of the spatio-temporal discretization of the problem. This weak monotonicity property then allows for the estimate $\sup_{1 \le j \le J} \mathbb{E}\big[\|X_{t_j} - Y^j\|_{\mathbb{L}^2}^2\big] \le C_\delta(k^{1-\delta} + h^2)$ for all small $\delta > 0$, where X is the strong variational solution of the stochastic Allen-Cahn equation, while $\{Y^j : 0 \le j \le J\}$ solves a structure preserving finite element based space-time discretization of the problem on a temporal mesh $\{t_j; 1 \le j \le J\}$ of size $k > 0$ which covers $[0, T]$.

1. INTRODUCTION

Let $\big(\mathbb{H}, (\cdot, \cdot)_{\mathbb{H}}\big)$ be a separable Hilbert space, and $\mathbb{V}$ be a reflexive Banach space, such that $\mathbb{V} \hookrightarrow \mathbb{H} \hookrightarrow \mathbb{V}'$ constitutes a Gelfand triple. The main motivation for this work is to identify the structural properties for the drift operator of the nonlinear SPDE

$$\mathrm{d}X_t = \mathscr{A}(X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}W_t \quad (t > 0), \qquad X(0) = x \in \mathbb{H}, \tag{1.1}$$

which allow to construct a space-time discretization of (1.1) for which optimal strong rates of convergence may be shown. Relevant works in this direction are [5, 6], where both, σ and $\mathscr{A}$ are required to be Lipschitz. The Lipschitz assumption for the drift operator $\mathscr{A} : \mathbb{V} \to \mathbb{V}'$ does not hold for many nonlinear SPDEs including the stochastic Navier-Stokes equation, or the stochastic version of general phase field models (including (1.2)) below for example. A usual strategy for a related numerical analysis is then to truncate nonlinearities (see e.g. [9]), or to quantify the mean square error on large subsets $\Omega_{k,h} := \Omega_k \cap \Omega_h \subset \Omega$. As an example, the following estimate for a (time-implicit, finite element based) space-time discretization of the 2D stochastic Navier-Stokes equation with solution $\{\mathbf{U}^m; m \ge 0\}$ was obtained in [2],

$$\mathbb{E}\Big[\chi_{\Omega_{k,h}} \max_{1 \le m \le M} \|\mathbf{u}(t_m) - \mathbf{U}^m\|_{\mathbb{L}^2}^2\Big] \le C\big(k^{\eta - \varepsilon} + kh^{-\varepsilon} + h^{2-\varepsilon}\big) \qquad (\varepsilon > 0)$$

for all $\eta \in (0, \frac{1}{2})$, where $\Omega_k \subset \Omega$ (resp. $\Omega_h \subset \Omega$) is such that $\mathbb{P}[\Omega \setminus \Omega_k] \to 0$ for $k \to 0$ (resp. $\mathbb{P}[\Omega \setminus \Omega_h] \to 0$ for $h \to 0$). We also mention the work [8] which studies a spatial discretization of the stochastic Cahn-Hilliard equation.

Let $\mathscr{O} \subset \mathbb{R}^d$, $d \in \{1, 2, 3\}$ be a bounded Lipschitz domain. We consider the stochastic Allen-Cahn equation with multiplicative noise, where the process $X : \Omega \times [0, T] \times \overline{\mathscr{O}} \to \mathbb{R}$ solves

$$\mathrm{d}X_t - \Big(\Delta X_t - \big(|X_t|^2 - 1\big)X_t\Big)\mathrm{d}t = \sigma(X_t)\mathrm{d}W_t \quad (t > 0), \qquad X_0 = x, \tag{1.2}$$

where $W \equiv \{W_t; 0 \le t \le T\}$ is an $\mathbb{R}$-valued Wiener process which is defined on the given filtered probability space $\mathfrak{P} \equiv (\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$; however it is easily possible to generalize the analysis below

to a trace class Q- Wiener process. Obviously, the drift operator $\mathscr{A}(y) = \Delta y - (|y|^2 - 1)y$ is only locally Lipschitz, but is the negative Gâteaux differential of $\mathcal{J}(y) = \frac{1}{2}\|\nabla y\|_{\mathbb{L}^2}^2 + \frac{1}{4}\||y|^2 - 1\|_{\mathbb{L}^2}^2$ and satisfies the weak monotonicity property

$$\left\langle \mathscr{A}(y_1) - \mathscr{A}(y_2), y_1 - y_2 \right\rangle_{(\mathbb{W}_{\mathrm{per}}^{1,2})^* \times \mathbb{W}_{\mathrm{per}}^{1,2}} \leq K \|y_1 - y_2\|_{\mathbb{L}^2}^2 - \|\nabla(y_1 - y_2)\|_{\mathbb{L}^2}^2 \ \ \forall y_1, y_2 \in \mathbb{W}_{\mathrm{per}}^{1,2} \quad (1.3)$$

for some $K > 0$; see Section 2 for the notation. Our goal is a (variational) error analysis for the structure preserving finite element based space-time discretization (2.5) which accounts for this structural property, avoiding arguments that exploit only the locally Lipschitz property of $\mathscr{A}$ to arrive at optimal strong error estimate.

The existing literature (see e.g. [8]) for estimating the numerical strong error on the problem (1.2) mainly uses the involved linear semigroup theory; the authors have considered the additive colored noise case, in which they have benefited from it by using the stochastic convolution, and then used a truncation of the nonlinear drift operator $\mathscr{A}$ to prove a *rate of convergence* for a (spatial) semi-discretization on sets of probability close to 1 without exploiting the weak monotonicity of $\mathscr{A}$. In contrast, property (1.3) and variational arguments were used in the recent work [10], where strong error estimates for both, semi-discrete (in time) and fully-discrete schemes for (1.2) were obtained, which are of sub-optimal order $\mathcal{O}(\sqrt{k} + h^{\frac{2-\delta}{6}})$ for the fully discrete scheme in the case $d = 3$. In [10], a standard implicit discretization of (1.2) was considered for which it is not clear to obtain uniform bounds for arbitrary higher moments of the solution of the fully discrete scheme, thus leading to sub-optimal convergence rates above. In this work, we consider the modified scheme (2.5) for (1.2), and derive optimal strong numerical error estimate.

The subsequent analysis for the scheme (2.5) is split into two steps to independently address errors due to the temporal and spatial discretization. First we exploit the variational solution concept for (1.2) and the semi-linear structure of $\mathscr{A}(y) = -\mathcal{J}'(y)$ to derive uniform bounds for the arbitrarily higher moments of the solution of (1.2) in strong norms; these bounds may then be used to bound increments of the solution of (1.2) in Lemma 3.2. The second ingradient to achieve optimal error bounds is a temporal discretization which inherits the structural properties of (1.2); the scheme (4.1) is constructed to allow for bounds of arbitrary moments of $\{\mathcal{J}(X^j); 0 \leq j \leq J\}$ in Lemma 4.1, which then settles the error bounds in Theorem 4.2 by using property (1.3) to effectively handle the nonlinear terms. We recover the asymptotic rate $\frac{1}{2}$ which is known for SPDEs of the form (1.1) when $\mathscr{A}$ is linear elliptic. It is interesting to compare the present error analysis for the SPDE (1.2) with the one in [7] for a general SODE with polynomial drift (see [7, Ass.s 3.1, 4.1, 4.2]) which also exploits the weak monotonicity of the drift.

The temporal semi-discretization was studied as a first step rather than spatial discretization to inherit bounds in strong norms which are needed for a complete error analysis of the problem. The second part of the error analysis is then on the structure preserving finite element based fully discrete scheme (2.5), for which we first verify the uniform bounds of arbitrary moments of $\{\mathcal{J}(Y^j); 0 \leq j \leq J\}$ (cf. Lemma 5.1). It is worth mentioning that, if $\{Y^j : 0 \leq j \leq J\}$ is a solution to a *standard* space-time discretization which involves the nonlinearity $\mathscr{A}(Y^j) = -\mathcal{J}'(Y^j)$, then only basic uniform bounds may be obtained (see [10, Lemma 2.5]), as opposed to those in Lemma 5.1. Next to it, we use again (1.3) for the drift, in combination with well-known approximation results for a finite element discretization to show that the error part due to spatial discretization is of order $\mathcal{O}(\sqrt{k} + h)$ where $k > 0$ is the time discretization parameter and $h > 0$ is the space discretization parameter (see Theorem 5.2). In this context, we mention the numerical analysis in [11] for an extended model of (1.2), where the uniform bounds for the exponential moments next to arbitrary moments in stronger norms are obtained for the solution of a semi-discretization in space in the case $d = 1$ (see [11, Prop.s 4.2, 4.3]); those bounds, together with a monotonicity argument are then used to properly address the nonlinear effects

in the error analysis and arrive at the (lower) strong rate $\frac{1}{2}$ for the p-th mean convergence of the numerical solution.

2. Technical framework and main result

Throughout this paper, we use the letter $C > 0$ to denote various generic constants. Let $\mathscr{O} \equiv (0, R)^d$, $1 \leq d \leq 3$, with $R \in (0, \infty)$ be a cube in $\mathbb{R}^d$. Let us denote $\Gamma_j = \partial\mathscr{O} \cap \{x_j = 0\}$ and $\Gamma_{j+d} = \partial\mathscr{O} \cap \{x_j = R\}$ for $j = 1, \ldots, d$. The problem (1.2) is then supplemented by the space-periodic boundary condition

$$X\big|_{\Gamma_j} = X\big|_{\Gamma_{j+d}} \qquad (1 \leq j \leq d).$$

Let $\left(\mathbb{L}_{\mathrm{per}}^p, \|\cdot\|_{\mathbb{L}^p}\right)$ resp. $\left(\mathbb{W}_{\mathrm{per}}^{m,p}, \|\cdot\|_{\mathbb{W}^{m,p}}\right)$ denote the Lebesgue resp. Sobolev space of R-periodic functions $\varphi \in \mathbb{W}_{\mathrm{loc}}^{m,p}(\mathbb{R}^d)$. Recall that functions in $\mathbb{W}_{\mathrm{per}}^{m,2}$ may be characterized by their Fourier series expansion, i.e.,

$$
\mathbb{W}_{\mathrm{per}}^{m,2}(\mathscr{O}) = \left\{ \varphi : \mathbb{R}^d \to \mathbb{R} : \varphi(x) = \sum_{k \in \mathbb{Z}^d} c_k \exp\left(2i\pi \frac{\langle k, x \rangle}{R}\right), \right.
$$
$$
\left. \bar{c}_k = c_{-k}, \ \sum_{k \in \mathbb{Z}^d} |k|^{2m} |c_k|^2 < \infty \right\}.
$$

Below, we set $\psi(x) = \frac{1}{4}\||x|^2 - 1\|_{\mathbb{L}^2}^2$ for $x \in \mathbb{L}^2$. Throughout this article, we make the following assumption on $\sigma : \mathbb{R} \to \mathbb{R}$.

A.1 σ, σ', and σ'' are bounded. Moreover, σ is Lipschitz continuous, i.e., there exists a constant $K_1 > 0$ such that

$$\|\sigma(u) - \sigma(v)\|_{\mathbb{L}^2}^2 \leq K_1 \|u - v\|_{\mathbb{L}^2}^2 \quad \forall u, v \in \mathbb{L}_{\mathrm{per}}^2. \tag{2.1}$$

Definition 2.1. (Strong variational solution) Fix $T \in (0, \infty)$, and $x \in \mathbb{L}_{\mathrm{per}}^2$. A $\mathbb{W}_{\mathrm{per}}^{1,2}$-valued $\mathbb{F}$-adapted stochastic process $X \equiv \{X_t; \ t \in [0, T]\}$ is called a strong variational solution of (1.2) if $X \in L^2\left(\Omega; C([0, T]; \mathbb{L}_{\mathrm{per}}^2)\right)$ satisfies $\mathbb{P}$-a.s. for all $t \in [0, T]$ that

$$
\left(X_t, \phi\right)_{\mathbb{L}^2} + \int_0^t \left\{ \left(\nabla X_s, \nabla\phi\right)_{\mathbb{L}^2} + \left(D\psi(X_s), \phi\right)_{\mathbb{L}^2} \right\} \mathrm{d}s
$$
$$
= \left(x, \phi\right)_{\mathbb{L}^2} + \int_0^t \left(\sigma(X_s), \phi\right)_{\mathbb{L}^2} \mathrm{d}W_s \quad \forall \phi \in \mathbb{W}_{\mathrm{per}}^{1,2}. \tag{2.2}
$$

The following estimate for the strong solution is well-known ($p \geq 1$),

$$
\sup_{t \in [0, T]} \mathbb{E}\left[\frac{1}{p}\left(\|X_t\|_{\mathbb{L}^2}^p - \|x\|_{\mathbb{L}^2}^p\right) + \int_0^T \|X_s\|_{\mathbb{L}^2}^{p-2}\left(\|\nabla X_s\|_{\mathbb{L}^2}^2 + \|X_s\|_{\mathbb{L}^4}^4\right) \mathrm{d}s\right] \leq C. \tag{2.3}
$$

2.1. **Fully discrete scheme.** Let us introduce some notation needed to define the structure preserving finite element based fully discrete scheme. Let $0 = t_0 < t_1 < \ldots < t_J$ be a uniform partition of $[0, T]$ of size $k = \frac{T}{J}$. Let $\mathscr{T}_h$ be a quasi-uniform triangulation of the domain $\mathscr{O}$. We consider the $\mathbb{W}_{\mathrm{per}}^{1,2}$-conforming finite element space (cf. [1]) $\mathbb{V}_h \subset \mathbb{W}_{\mathrm{per}}^{1,2}$ such that

$$\mathbb{V}_h = \left\{ \phi \in C(\overline{\mathscr{O}}; \mathbb{R}); \ \phi\big|_K \in \mathscr{P}_1(K) \quad \forall K \in \mathscr{T}_h \right\},$$

where $\mathscr{P}_1(K)$ is the space of $\mathbb{R}$-valued functions on K which are polynomials of degree less or equal to 1. We may then consider the space-time discretization of (1.2): Let $Y^0 = \mathscr{P}_{\mathbb{L}^2} x \in \mathbb{V}_h$, where $\mathscr{P}_{\mathbb{L}^2} : \mathbb{L}_{\mathrm{per}}^2 \to \mathbb{V}_h$ denotes the $\mathbb{L}_{\mathrm{per}}^2$-orthogonal projection, i.e., for all $g \in \mathbb{L}_{\mathrm{per}}^2$

$$\left(g - \mathscr{P}_{\mathbb{L}^2} g, \phi\right)_{\mathbb{L}^2} = 0 \quad \forall \phi \in \mathbb{V}_h. \tag{2.4}$$

Find the $\{\mathcal{F}_{t_j}; 0 \le j \le J\}$-adapted $\mathbb{V}_h$-valued process $\{Y^j; 0 \le j \le J\}$ such that $\mathbb{P}$-almost surely

$$\left(Y^j - Y^{j-1}, \phi\right)_{\mathbb{L}^2} + k\Big[\left(\nabla Y^j, \nabla\phi\right)_{\mathbb{L}^2} + \left(f(Y^j, Y^{j-1}), \phi\right)_{\mathbb{L}^2}\Big] = \Delta_j W\big(\sigma(Y^{j-1}), \phi\big)_{\mathbb{L}^2} \quad \forall \phi \in \mathbb{V}_h, \tag{2.5}$$

where

$$\Delta_j W := W(t_j) - W(t_{j-1}) \sim \mathcal{N}(0, k), \ \text{ and } \ f(y, z) = \big(|y|^2 - 1\big)\frac{y + z}{2}. \tag{2.6}$$

Solvability for $k < 1$ is again immediate via Brouwer fixed point theorem.

We are now in a position to state the main result of this article.

Main Theorem. Let the assumption **A**.1 hold and $x \in \mathbb{W}^{2,2}_{\text{per}}$. For every $\delta > 0$, there exist constants $0 < C_\delta < \infty$, independent of the discretized parameters $k, h > 0$, and $k_0 = k_0(T, x) > 0$ such that for all $k \le k_0$ sufficiently small, there holds

$$\sup_{1 \le j \le J} \mathbb{E}\Big[\|X_{t_j} - Y^j\|^2_{\mathbb{L}^2}\Big] + k\sum_{j=1}^J \mathbb{E}\Big[\|\nabla(X_{t_j} - Y^j)\|^2_{\mathbb{L}^2}\Big] \le C_\delta\big(k^{1-\delta} + h^2\big),$$

where $\{X_t; t \in [0, T]\}$ solves (1.2) while $\{Y^j; 0 \le j \le J\}$ solves (2.5).

The proof is detailed in Sections 3, 4 and 5 and uses the semi-linear structure of $\mathscr{A}$ along with the weak monotonicity property (1.3). We first consider a semi-discrete (in time) scheme (4.1) of the problem (1.2) and derive the error estimate between the strong solution X of (1.2) and the discretized solution $\{X^j; 0 \le j \le J\}$ of (4.1), see Theorem 4.2. Again, using uniform bounds for higher moments of the solutions $\{X^j\}$ and $\{Y^j\}$, we derive the error estimate of X^j and Y^j in strong norm, cf. Theorem 5.2. Putting things together then settles the main theorem.

3. Stochastic Allen-Cahn equation: The continuous case

In this section, we derive uniform bounds of arbitrary moments for the strong solution $X \equiv \{X_t : t \in [0, T]\}$ of (1.2) and using these uniform bounds, we estimate the expectation of the increment $\|X_t - X_s\|^2_{\mathbb{L}^2}$ in terms of $|t - s|$.

The following estimate may be shown by a standard Galerkin method which employs a (finite) sequence of ($\mathbb{W}^{1,2}_{\text{per}}$-orthonormal) eigenfunctions $\{w_j; 1 \le j \le N\}$ of the inversely compact, self-adjoint isomorphic operator $\mathrm{I} - \Delta : \mathbb{W}^{2,2}_{\text{per}} \to \mathbb{L}^2_{\text{per}}$, the use of Itô's formula to the functional $y \mapsto \mathcal{J}(y) := \frac{1}{2}\|\nabla y\|^2_{\mathbb{L}^2} + \psi(y)$, and the final passage to the limit (see e.g. [4]),

$$\sup_{t \in [0,T]} \mathbb{E}\Big[\mathcal{J}(X_t) + \int_0^t \|\Delta X_s - D\psi(X_s)\|^2_{\mathbb{L}^2}\,\mathrm{d}s\Big] \le C\Big(1 + \mathbb{E}\big[\mathcal{J}(x)\big]\Big), \tag{3.1}$$

for which we require the improved regularity property $x \in \mathbb{W}^{1,2}_{\text{per}}$. Thanks to the assumption **A**.1, we see that σ satisfies the following estimates:

a) There exists a constant $C > 0$ such that

$$\|\nabla\sigma(\xi)\|^2_{\mathbb{L}^2} + \Big(D^2\psi(\xi)\sigma(\xi), \sigma(\xi)\Big)_{\mathbb{L}^2} \le C\big(1 + \mathcal{J}(\xi)\big) \quad \forall \xi \in \mathbb{W}^{1,2}_{\text{per}}. \tag{3.2}$$

b) There exist constants $K_2, K_3, K_4 > 0$ and $L_2, L_3, L_4 > 0$ such that for all $\xi \in \mathbb{W}^{2,2}_{\text{per}}$

$$\|\Delta\sigma(\xi)\|^2_{\mathbb{L}^2} \le \begin{cases} K_2\|\Delta\xi\|^2_{\mathbb{L}^2} + K_3\|\nabla\xi\|^2_{\mathbb{L}^2}\|\Delta\xi\|^2_{\mathbb{L}^2} + K_4\|\nabla\xi\|^4_{\mathbb{L}^2} & \text{if } d = 2 \\ L_2\|\Delta\xi\|^2_{\mathbb{L}^2} + L_3\|\nabla\Delta\xi\|^{\frac{3}{2}}_{\mathbb{L}^2}\|\nabla\xi\|^{\frac{5}{2}}_{\mathbb{L}^2} + L_4\|\nabla\xi\|^4_{\mathbb{L}^2} & \text{if } d = 3. \end{cases} \tag{3.3}$$

Lemma 3.1. *Suppose that the assumption **A**.1 holds and $p \in \mathbb{N}$. Then, there exists a constant $C \equiv C(\|x\|_{\mathbb{W}^{1,2}}, p, T) > 0$ such that*

(i) $\quad \sup_{t \in [0,T]} \mathbb{E}\left[\left(\mathcal{J}(X_t)\right)^p\right] + \mathbb{E}\left[\int_0^T \|\nabla X_s\|_{\mathbb{L}^2}^{2(p-1)} \|\Delta X_s\|_{\mathbb{L}^2}^2 \, ds\right] \leq C \,.$

Suppose in addition $x \in \mathbb{W}_{per}^{2,2}$. Then, there exists a constant $C \equiv C(\|x\|_{\mathbb{W}^{2,2}}, T) > 0$ such that

(ii) $\quad \sup_{t \in [0,T]} \mathbb{E}\left[\|\Delta X_t\|_{\mathbb{L}^2}^2\right] + \mathbb{E}\left[\int_0^T \|\nabla \Delta X_s\|_{\mathbb{L}^2}^2 \, ds\right] \leq C \,.$

Proof. (i) We proceed formally. Note that

$$\mathcal{J}(\xi) := \frac{1}{2}\|\nabla \xi\|_{\mathbb{L}^2}^2 + \psi(\xi) \qquad \text{and} \qquad -\mathscr{A}(\xi) \equiv \mathcal{J}'(\xi) = -\Delta \xi + D\psi(\xi).$$

We use Itô's formula for $\xi \mapsto g(\xi) := \left(\mathcal{J}(\xi)\right)^p$.

$$Dg(\xi) = -p\left(\mathcal{J}(\xi)\right)^{p-1} \mathscr{A}(\xi)$$
$$D^2 g(\xi) = p(p-1)\left(\mathcal{J}(\xi)\right)^{p-2} \mathscr{A}(\xi) \otimes \mathscr{A}(\xi) + p\left(\mathcal{J}(\xi)\right)^{p-1}\left(-\Delta + D^2\psi(\xi)\right),$$

where $a \otimes b \cdot c = a(b,c)_{\mathbb{L}^2}$ for all $a, b, c \in \mathbb{L}^2$. By Cauchy-Schwartz inequality, and (3.2), we have

$$\mathbb{E}\left[\left(\mathcal{J}(X_t)\right)^p - \left(\mathcal{J}(x)\right)^p + p\int_0^t \left(\mathcal{J}(X_s)\right)^{p-1}\|\mathscr{A}(X_s)\|_{\mathbb{L}^2}^2 \, ds\right]$$

$$= \frac{p}{2}\mathbb{E}\left[\int_0^t \left\{(p-1)\left(\mathcal{J}(X_s)\right)^{p-2}\left(\mathscr{A}(X_s), \sigma(X_s)\right)_{\mathbb{L}^2}^2 \right.\right.$$

$$\left.\left. + \left(\mathcal{J}(X_s)\right)^{p-1}\left([-\Delta + D^2\psi(X_s)]\sigma(X_s), \sigma(X_s)\right)_{\mathbb{L}^2}\right\} ds\right]$$

$$\leq \frac{p}{2}\mathbb{E}\left[\int_0^t (p-1)\left(\mathcal{J}(X_s)\right)^{p-2}\left(\mathscr{A}(X_s), \sigma(X_s)\right)_{\mathbb{L}^2}^2 \, ds\right] + C(p)\int_0^t \mathbb{E}\left[\left(\mathcal{J}(X_s)\right)^p\right] ds + C.$$

Since σ is bounded, by using Young's inequality, we have

$$\left(\mathcal{J}(X_s)\right)^{p-2}\left(\mathscr{A}(X_s), \sigma(X_s)\right)_{\mathbb{L}^2}^2 \leq C\left(\mathcal{J}(X_s)\right)^{p-2}\|\mathscr{A}(X_s)\|_{\mathbb{L}^2}^2$$

$$\leq \theta\left(\mathcal{J}(X_s)\right)^{p-1}\|\mathscr{A}(X_s)\|_{\mathbb{L}^2}^2 + C(\theta)\|\mathscr{A}(X_s)\|_{\mathbb{L}^2}^2.$$

We choose $\theta > 0$ such that $p - \frac{p}{2}(p-1)\theta > 0$. With this choice of θ, by (3.1), we have for some constant $C_1 = C_1(p) > 0$,

$$\mathbb{E}\left[\left(\mathcal{J}(X_t)\right)^p\right] + C_1(p)\mathbb{E}\left[\int_0^t \left(\mathcal{J}(X_s)\right)^{p-1}\|\mathscr{A}(X_s)\|_{\mathbb{L}^2}^2 \, ds\right]$$

$$\leq \mathbb{E}\left[\left(\mathcal{J}(x)\right)^p\right] + C(p)\int_0^t \mathbb{E}\left[\left(\mathcal{J}(X_s)\right)^p\right] ds + C.$$

We use Gronwall's lemma to conclude that

$$\sup_{t \in [0,T]} \mathbb{E}\left[\left(\mathcal{J}(X_t)\right)^p\right] + \mathbb{E}\left[\int_0^T \left(\mathcal{J}(X_s)\right)^{p-1}\|\mathscr{A}(X_s)\|_{\mathbb{L}^2}^2 \, ds\right] \leq C. \tag{3.4}$$

In view of (2.3), and (3.4) it follows that

$$\sup_{t \in [0,T]} \mathbb{E}\left[\|X_t\|_{\mathbb{W}^{1,2}}^p\right] \leq C, \quad \forall p \geq 1. \tag{3.5}$$

Note that,

$$\|\Delta X_s\|_{\mathbb{L}^2}^2 \leq \|\mathscr{A}(X_s)\|_{\mathbb{L}^2}^2 + \|X_s\|_{\mathbb{L}^6}^6 + \|X_s\|_{\mathbb{L}^2}^2 \leq \|\mathscr{A}(X_s)\|_{\mathbb{L}^2}^2 + C\|X_s\|_{\mathbb{W}^{1,2}}^6 + \|X_s\|_{\mathbb{W}^{1,2}}^2,$$

where we use the embedding $\mathbb{W}^{1,2} \hookrightarrow \mathbb{L}^6$ for $d \leq 3$. Thanks to (3.4) and (3.5), together with Cauchy-Schwartz inequality and the above estimate, we see that

$$\mathbb{E}\Big[\int_0^T \big(\mathcal{J}(X_s)\big)^{p-1} \|\Delta X_s\|_{\mathbb{L}^2}^2 \, ds \Big]$$

$$\leq C + \mathbb{E}\Big[\int_0^T \big(\mathcal{J}(X_s)\big)^{p-1} \big(\|X_s\|_{\mathbb{W}^{1,2}}^6 + \|X_s\|_{\mathbb{W}^{1,2}}^2 \big) ds \Big]$$

$$\leq C + T \sup_{t \in [0,T]} \mathbb{E}\Big[\big(\mathcal{J}(X_t)\big)^p \Big] + T \sup_{t \in [0,T]} \mathbb{E}\Big[\|X_t\|_{\mathbb{W}^{1,2}}^{6p} + \|X_t\|_{\mathbb{W}^{1,2}}^{2p} \Big] \leq C. \tag{3.6}$$

One can combine (3.4) and (3.6) to conclude the assertion.

(ii) Use Itô's formula for $\xi \mapsto g(\xi) = \frac{1}{2}\|\Delta \xi\|_{\mathbb{L}^2}^2$. We compute its derivatives.

$$Dg(\xi) = \Delta^2 \xi, \qquad D^2 g(\xi) = \Delta^2 .$$

Note that by integration by parts

$$\Big(\Delta^2 X_s, -\Delta X_s + |X_s|^2 X_s \Big)_{\mathbb{L}^2} \geq \frac{1}{2} \Big[\|\nabla \Delta X_s\|_{\mathbb{L}^2}^2 - 3 \| |X_s|^2 \nabla X_s \|_{\mathbb{L}^2}^2 \Big]. \tag{3.7}$$

Because of $\mathbb{W}^{1,2} \hookrightarrow \mathbb{L}^6$ for $d \leq 3$ we estimate the last term through

$$\| |X_s|^2 \nabla X_s \|_{\mathbb{L}^2}^2 \leq \|X_s\|_{\mathbb{L}^6}^4 \|\nabla X_s\|_{\mathbb{L}^6}^2 \leq C \|X_s\|_{\mathbb{W}^{1,2}}^4 \big(\|\nabla X_s\|_{\mathbb{L}^2}^2 + \|\Delta X_s\|_{\mathbb{L}^2}^2 \big)$$

$$\leq C \|X_s\|_{\mathbb{W}^{1,2}}^6 + C \big(\|X_s\|_{\mathbb{L}^2}^2 + \|\nabla X_s\|_{\mathbb{L}^2}^2 \big)^2 \|\Delta X_s\|_{\mathbb{L}^2}^2$$

$$\equiv C \|X_s\|_{\mathbb{W}^{1,2}}^6 + \mathcal{M}. \tag{3.8}$$

Note that $\|X_s\|_{\mathbb{L}^2}^2 \leq C(1 + \psi(X_s))$, and therefore we see that

$$\mathcal{M} \leq C \Big(1 + \big(\mathcal{J}(X_s)\big)^2 \Big) \|\Delta X_s\|_{\mathbb{L}^2}^2. \tag{3.9}$$

Inserting (3.8) and (3.9) in (3.7), we obtain

$$\mathbb{E}\Big[\|\Delta X_t\|_{\mathbb{L}^2}^2 \Big] + \mathbb{E}\Big[\int_0^t \|\nabla \Delta X_s\|_{\mathbb{L}^2}^2 \, ds \Big]$$

$$\leq \mathbb{E}\big[\|\Delta x\|_{\mathbb{L}^2}^2 \big] + C \mathbb{E}\Big[\int_0^t \|X_s\|_{\mathbb{W}^{1,2}}^6 \, ds \Big] + C \mathbb{E}\Big[\int_0^T \big(\mathcal{J}(X_s)\big)^2 \|\Delta X_s\|_{\mathbb{L}^2}^2 \, ds \Big]$$

$$+ C \int_0^t \mathbb{E}\big[\|\Delta X_s\|_{\mathbb{L}^2}^2 \big] \, ds + C \mathbb{E}\Big[\int_0^t \|\Delta \sigma(X_s)\|_{\mathbb{L}^2}^2 \, ds \Big]. \tag{3.10}$$

Let $d = 2$. Then by (3.3), we see that

$$\mathbf{G} := \mathbb{E}\Big[\int_0^t \|\Delta \sigma(X_s)\|_{\mathbb{L}^2}^2 \, ds \Big]$$

$$\leq K_2 \int_0^t \mathbb{E}\big[\|\Delta X_s\|_{\mathbb{L}^2}^2 \big] \, ds + K_3 \mathbb{E}\Big[\int_0^T \|\nabla X_s\|_{\mathbb{L}^2}^2 \|\Delta X_s\|_{\mathbb{L}^2}^2 \, ds \Big] + K_4 \int_0^T \mathbb{E}\big[\|\nabla X_s\|_{\mathbb{L}^2}^4 \big] \, ds.$$

One can combine the above estimate in (3.10) and use (3.5), (3.6) along with Gronwall's lemma to conclude the assertion for $d \leq 2$.

Let $d = 3$. Then, thanks to (3.3) and Cauchy-Schwartz inequality, we have

$$\mathbf{G} \leq \frac{1}{2} \mathbb{E}\Big[\int_0^t \|\nabla \Delta X_s\|_{\mathbb{L}^2}^2 \, ds \Big] + C \mathbb{E}\Big[\int_0^T \|\nabla X_s\|_{\mathbb{L}^2}^{10} \, ds \Big] + L_4 \int_0^T \mathbb{E}\big[\|\nabla X_s\|_{\mathbb{L}^2}^4 \big] \, ds$$

$$+ L_2 \int_0^t \mathbb{E}\big[\|\Delta X_s\|_{\mathbb{L}^2}^2 \big] \, ds. \tag{3.11}$$

We combine (3.10) and (3.11) and use (3.5), (3.6) along with Gronwall's lemma to conclude the assertion for $d = 3$. This completes the proof. $\square$

The following result is to bound the increments $X_t - X_s$ of the solutions of (1.2) in terms of $|t - s|^\alpha$ for some $\alpha > 0$; its proof uses Lemma 3.1 in particular.

Lemma 3.2. *Let the assumption* **A**.1 *holds and* $x \in \mathbb{W}^{2,2}_{\text{per}}$. *Then for every* $0 \leq s \leq t \leq T$, *there exists a constant* $C \equiv C(p, T) > 0$ *such that*

$$\text{(i)} \quad \mathbb{E}\big[\|X_t - X_s\|^p_{\mathbb{L}^2}\big] \leq C|t - s| \qquad (p \geq 2),$$
$$\text{(ii)} \quad \mathbb{E}\big[\|X_t - X_s\|^2_{\mathbb{W}^{1,2}}\big] \leq C|t - s| \,.$$

Proof. (i) Fix $s \geq 0$. An application of Itô's formula for $u \mapsto \frac{1}{p}\|u - \beta\|^p_{\mathbb{L}^2}$ with $\beta = X_s(\cdot, \omega) \in \mathbb{R}$ to (1.2) yields, after taking expectation

$$\mathbb{E}\Big[\frac{1}{p}\|X_t - X_s\|^p_{\mathbb{L}^2} + \int_s^t \|X_\zeta - X_s\|^{p-2}_{\mathbb{L}^2}\Big(\mathscr{A}(X_\zeta) - \mathscr{A}(X_s), X_\zeta - X_s\Big)_{\mathbb{L}^2} \mathrm{d}\zeta\Big]$$
$$\leq \mathbb{E}\Big[\int_s^t \|X_\zeta - X_s\|^{p-2}_{\mathbb{L}^2}\Big(-\mathscr{A}(X_s), X_\zeta - X_s\Big)_{\mathbb{L}^2}\mathrm{d}\zeta\Big] + C_p\mathbb{E}\Big[\int_s^t \|X_\zeta - X_s\|^{p-2}_{\mathbb{L}^2}\|\sigma(X_\zeta)\|^2_{\mathbb{L}^2} \mathrm{d}\zeta\Big]$$
$$\equiv A_1 + A_2.$$

We use the weak monotonicity property (1.3) to bound from below the second term on the left-hand side,

$$\geq \mathbb{E}\Big[\int_s^t \Big(\|X_\zeta - X_s\|^{p-2}_{\mathbb{L}^2}\|\nabla(X_\zeta - X_s)\|^2_{\mathbb{L}^2} - C\,\|X_\zeta - X_s\|^p_{\mathbb{L}^2}\Big)\mathrm{d}\zeta\Big].$$

The integration by parts formula and Young's inequality reveal that

$$\Big(-\mathscr{A}(X_s), X_\zeta - X_s\Big)_{\mathbb{L}^2} \leq \|\nabla X_s\|_{\mathbb{L}^2}\|\nabla(X_\zeta - X_s)\|_{\mathbb{L}^2} + \|D\psi(X_s)\|_{\mathbb{L}^2}\|X_\zeta - X_s\|_{\mathbb{L}^2}$$
$$\leq \Big(\|X_s\|^3_{\mathbb{L}^6} + \|X_s\|_{\mathbb{W}^{1,2}}\Big)\|X_\zeta - X_s\|_{\mathbb{W}^{1,2}}$$

Since $\mathbb{W}^{1,2} \hookrightarrow \mathbb{L}^6$ for $d \leq 3$, by using Young's inequality, we see that

$$A_1 \leq C\,\mathbb{E}\Big[\int_s^t \|X_\zeta - X_s\|^{p-2}_{\mathbb{L}^2}\Big(\|X_s\|_{\mathbb{W}^{1,2}} + \|X_s\|^3_{\mathbb{L}^6}\Big)\|X_\zeta - X_s\|_{\mathbb{W}^{1,2}} \mathrm{d}\zeta\Big]$$
$$\leq \mathbb{E}\Big[\frac{1}{2}\int_s^t \|X_\zeta - X_s\|^{p-2}_{\mathbb{L}^2}\Big(\|X_\zeta - X_s\|^2_{\mathbb{L}^2} + \|\nabla(X_\zeta - X_s)\|^2_{\mathbb{L}^2}\Big)\mathrm{d}\zeta\Big]$$
$$\quad + \frac{1}{2}\mathbb{E}\Big[\int_s^t \|X_\zeta - X_s\|^{p-2}_{\mathbb{L}^2}\Big(\|X_s\|^2_{\mathbb{W}^{1,2}} + \|X_s\|^6_{\mathbb{W}^{1,2}}\Big)\mathrm{d}\zeta\Big]$$
$$\leq \frac{1}{2}\mathbb{E}\Big[\int_s^t \|X_\zeta - X_s\|^{p-2}_{\mathbb{L}^2}\|\nabla(X_\zeta - X_s)\|^2_{\mathbb{L}^2} \mathrm{d}\zeta\Big] + C_p\int_s^t \mathbb{E}\Big[\|X_\zeta - X_s\|^p_{\mathbb{L}^2} \mathrm{d}\zeta\Big]$$
$$\quad + C|t - s|\sup_{\zeta \in [s,t]}\mathbb{E}\Big[\|X_\zeta\|^p_{\mathbb{W}^{1,2}} + \|X_\zeta\|^{3p}_{\mathbb{W}^{1,2}}\Big].$$

Again, thanks to the boundedness of σ and Young's inequality, we see that

$$A_2 \leq C_p\int_s^t \mathbb{E}\Big[\|X_\zeta - X_s\|^p_{\mathbb{L}^2} \mathrm{d}\zeta + C|t - s|.$$

We combine all the above estimates and use (2.3) and Lemma 3.1, (i) along with Gronwall's inequality to get the result.

(ii) We apply Itô's formula to the function $\frac{1}{2}|\nabla X_t - \beta|^2$ for any $\beta \in \mathbb{R}$ to (1.2), and then use $\beta = \nabla X_s$ for fixed $0 < s \leq t$ and integrate with respect to spatial variable. Thanks to Young's inequality, and the boundedness of σ',

$$
\frac{1}{2}\mathbb{E}\Big[\|\nabla(X_t - X_s)\|_{\mathbb{L}^2}^2\Big] \leq \Big|\int_s^t \mathbb{E}\Big[\big(-\Delta[X_\zeta - X_s], \mathscr{A}(X_\zeta)\big)_{\mathbb{L}^2}\Big]\,\mathrm{d}\zeta\Big| + C\int_s^t \mathbb{E}\Big[\|\nabla\sigma(X_\zeta)\|_{\mathbb{L}^2}^2\Big]\,\mathrm{d}\zeta
$$

$$
\leq C|t - s|\sup_{\zeta \in [s,t]} \mathbb{E}\Big[\|\Delta X_\zeta\|_{\mathbb{L}^2}^2\Big] + C\int_s^t \mathbb{E}\Big[\|\mathscr{A}(X_\zeta)\|_{\mathbb{L}^2}^2 + \|\nabla X_\zeta\|_{\mathbb{L}^2}^2\Big]\,\mathrm{d}\zeta.
$$

Notice that $\|\mathscr{A}(X_\zeta)\|_{\mathbb{L}^2}^2 \leq \|\Delta X_\zeta\|_{\mathbb{L}^2}^2 + C\big(\|X_\zeta\|_{\mathbb{L}^6}^6 + \|X_\zeta\|_{\mathbb{L}^2}^2\big)$. From the above estimate, and Lemma 3.1, (ii), (3.5), and the embedding $\mathbb{W}^{1,2} \hookrightarrow \mathbb{L}^6$ for $d \leq 3$, we conclude

$$
\mathbb{E}\Big[\|\nabla(X_t - X_s)\|_{\mathbb{L}^2}^2\Big] \leq C|t - s|\sup_{\zeta \in [s,t]} \mathbb{E}\Big[\|\Delta X_\zeta\|_{\mathbb{L}^2}^2 + \|X_\zeta\|_{\mathbb{W}^{1,2}}^6 + \|X_\zeta\|_{\mathbb{W}^{1,2}}^2\Big] \leq C|t - s|. \quad (3.12)
$$

One can use (i) of Lemma 3.2 for $p = 2$, and (3.12) to arrive at (ii). This finishes the proof. $\quad\square$

4. Semi-discrete scheme (in time) and its bound

Let $0 = t_0 < t_1 < \ldots < t_J$ be an equi-distant partition of $[0, T]$ of size $k = \frac{T}{J}$. The structure preserving time discrete version of (1.2) defines an $\{\mathcal{F}_{t_j};\ 0 \leq j \leq J\}$-adapted $\mathbb{W}_{\mathrm{per}}^{1,2}$-valued process $\{X^j;\ 0 \leq j \leq J\}$ such that $\mathbb{P}$-almost surely and for all $\phi \in \mathbb{W}_{\mathrm{per}}^{1,2}$

$$
\begin{cases}
\big(X^j - X^{j-1}, \phi\big)_{\mathbb{L}^2} + k\Big[\big(\nabla X^j, \nabla\phi\big)_{\mathbb{L}^2} + \big(f(X^j, X^{j-1}), \phi\big)_{\mathbb{L}^2}\Big] = \Delta_j W\big(\sigma(X^{j-1}), \phi\big)_{\mathbb{L}^2} \\
X^0 = x \in \mathbb{L}_{\mathrm{per}}^2,
\end{cases} \quad (4.1)
$$

where $\Delta_j W$ and f are defined in (2.6). Solvability for $k < 1$ easily follows from a coercivity property of the drift operator, and the Lipschitz continuity property (2.1) for the diffusion operator. Below, we denote again

$$
\mathcal{J}(X^j) = \frac{1}{2}\|\nabla X^j\|_{\mathbb{L}^2}^2 + \psi(X^j).
$$

The proof of the following lemma evidences why $D\psi(X^j)$ is substituted by $f(X^j, X^{j-1})$ in (4.1) to recover uniform bounds for arbitrary higher moments of $\mathcal{J}(X^j)$.

Lemma 4.1. *Suppose* $x \in \mathbb{W}_{\mathrm{per}}^{1,2}$, *and that assumption* **A**.1 *holds. For every* $p = 2^r$, $r \in \mathbb{N}^*$, *there exists a constant* $C \equiv C(p, T) > 0$ *such that*

$$
\max_{1 \leq j \leq J} \mathbb{E}\Big[\big|\mathcal{J}(X^j)\big|^p\Big] + \sum_{j=1}^J \mathbb{E}\Bigg[\prod_{\ell=1}^r \Big[[\mathcal{J}(X^j)]^{2^{\ell-1}} + [\mathcal{J}(X^{j-1})]^{2^{\ell-1}}\Big] \times \Big(\|\nabla(X^j - X^{j-1})\|_{\mathbb{L}^2}^2
$$

$$
+ \big\||X^j|^2 - |X^{j-1}|^2\big\|_{\mathbb{L}^2}^2 + k\big\|-\Delta X^j + f(X^j, X^{j-1})\big\|_{\mathbb{L}^2}^2\Big)\Bigg] \leq C.
$$

Proof. **1.** Consider (4.1) for a fixed $\omega \in \Omega$ and choose $\phi = -\Delta X^j(\omega) + f(X^j, X^{j-1})(\omega)$. Then one has $\mathbb{P}$-a.s.,

$$
\big(X^j - X^{j-1}, -\Delta X^j + f(X^j, X^{j-1})\big)_{\mathbb{L}^2} + k\big\|-\Delta X^j + f(X^j, X^{j-1})\big\|_{\mathbb{L}^2}^2
$$

$$
= \Delta_j W\big(\nabla\sigma(X^{j-1}), \nabla X^j\big)_{\mathbb{L}^2} + \Delta_j W\big(\sigma(X^{j-1}), f(X^j, X^{j-1})\big)_{\mathbb{L}^2} =: \mathcal{A}_1 + \mathcal{A}_2. \quad (4.2)
$$

By using the identity $(a - b)a = \frac{1}{2}\big(|a|^2 - |b|^2 + |a - b|^2\big)\ \forall a, b \in \mathbb{R}$ along with integration by parts formula, we calculate

$$
\big(X^j - X^{j-1}, -\Delta X^j + f(X^j, X^{j-1})\big)_{\mathbb{L}^2}
$$

$$= \left(\nabla(X^j - X^{j-1}), \nabla X^j\right)_{\mathbb{L}^2} + \frac{1}{2}\left(|X^j|^2 - 1, |X^j|^2 - 1 - (|X^{j-1}|^2 - 1)\right)_{\mathbb{L}^2}$$

$$= \frac{1}{2}\left(\|\nabla X^j\|_{\mathbb{L}^2}^2 - \|\nabla X^{j-1}\|_{\mathbb{L}^2}^2 + \|\nabla(X^j - X^{j-1})\|_{\mathbb{L}^2}^2\right)$$

$$+ \frac{1}{4}\left(\||X^j|^2 - 1\|_{\mathbb{L}^2}^2 - \||X^{j-1}|^2 - 1\|_{\mathbb{L}^2}^2 + \||X^j|^2 - |X^{j-1}|^2\|_{\mathbb{L}^2}^2\right)$$

$$= \mathcal{J}(X^j) - \mathcal{J}(X^{j-1}) + \frac{1}{2}\|\nabla(X^j - X^{j-1})\|_{\mathbb{L}^2}^2 + \frac{1}{4}\||X^j|^2 - |X^{j-1}|^2\|_{\mathbb{L}^2}^2. \tag{4.3}$$

Since σ' is bounded, we observe that

$$\mathcal{A}_1 \le \frac{1}{4}\|\nabla(X^j - X^{j-1})\|_{\mathbb{L}^2}^2 + C\|\nabla X^{j-1}\|_{\mathbb{L}^2}^2|\Delta_j W|^2 + \left(\nabla\sigma(X^{j-1}), \nabla X^{j-1}\right)_{\mathbb{L}^2}\Delta_j W$$

$$\le \frac{1}{4}\|\nabla(X^j - X^{j-1})\|_{\mathbb{L}^2}^2 + C\mathcal{J}(X^{j-1})|\Delta_j W|^2 + \left(\nabla\sigma(X^{j-1}), \nabla X^{j-1}\right)_{\mathbb{L}^2}\Delta_j W.$$

We decompose $\mathcal{A}_2$ into the sum of two terms $\mathcal{A}_{2,1}$ and $\mathcal{A}_{2,2}$ where

$$\begin{cases} \mathcal{A}_{2,1} = \left(\sigma(X^{j-1}), (|X^j|^2 - |X^{j-1}|^2)\frac{X^j + X^{j-1}}{2}\right)_{\mathbb{L}^2}\Delta_j W \\ \mathcal{A}_{2,2} = \left(\sigma(X^{j-1}), (|X^{j-1}|^2 - 1)\frac{X^j + X^{j-1}}{2}\right)_{\mathbb{L}^2}\Delta_j W. \end{cases}$$

In view of Young's inequality and the boundedness of σ, we have

$$\mathcal{A}_{2,1} \le \frac{1}{8}\||X^j|^2 - |X^{j-1}|^2\|_{\mathbb{L}^2}^2 + C\left(\|X^j - X^{j-1}\|_{\mathbb{L}^2}^2 + \|X^{j-1}\|_{\mathbb{L}^2}^2\right)|\Delta_j W|^2,$$

$$\mathcal{A}_{2,2} = \left(\sigma(X^{j-1}), (|X^{j-1}|^2 - 1)\frac{X^j - X^{j-1}}{2}\right)_{\mathbb{L}^2}\Delta_j W + \left(\sigma(X^{j-1}), (|X^{j-1}|^2 - 1)X^{j-1}\right)_{\mathbb{L}^2}\Delta_j W$$

$$\le \|X^j - X^{j-1}\|_{\mathbb{L}^2}^2 + \||X^{j-1}|^2 - 1\|_{\mathbb{L}^2}^2|\Delta_j W|^2 + \left(\sigma(X^{j-1}), (|X^{j-1}|^2 - 1)X^{j-1}\right)_{\mathbb{L}^2}\Delta_j W$$

$$\le \|X^j - X^{j-1}\|_{\mathbb{L}^2}^2 + C\mathcal{J}(X^{j-1})|\Delta_j W|^2 + \left(\sigma(X^{j-1}), (|X^{j-1}|^2 - 1)X^{j-1}\right)_{\mathbb{L}^2}\Delta_j W.$$

Next we estimate $\|X^j - X^{j-1}\|_{\mathbb{L}^2}^2$ independently to bound $\mathcal{A}_{2,2}$. To do so, we choose as test function $\phi = (X^j - X^{j-1})(\omega)$ in (4.1) and obtain

$$\|X^j - X^{j-1}\|_{\mathbb{L}^2}^2 + \frac{k}{2}\left(\|\nabla X^j\|_{\mathbb{L}^2}^2 - \|\nabla X^{j-1}\|_{\mathbb{L}^2}^2 + \|\nabla(X^j - X^{j-1})\|_{\mathbb{L}^2}^2\right)$$

$$+ \frac{k}{2}\left(|X^j|^2 - 1, |X^j|^2 - |X^{j-1}|^2\right)_{\mathbb{L}^2} = \left(\sigma(X^{j-1}), X^j - X^{j-1}\right)_{\mathbb{L}^2}\Delta_j W. \tag{4.4}$$

Note that

$$\frac{k}{2}\left(|X^j|^2 - 1, |X^j|^2 - |X^{j-1}|^2\right)_{\mathbb{L}^2} = k\left(\psi(X^j) - \psi(X^{j-1}) + \frac{1}{4}\||X^j|^2 - |X^{j-1}|^2\|_{\mathbb{L}^2}^2\right),$$

$$\left(\sigma(X^{j-1}), X^j - X^{j-1}\right)_{\mathbb{L}^2}\Delta_j W \le \frac{1}{2}\|X^j - X^{j-1}\|_{\mathbb{L}^2}^2 + C|\Delta_j W|^2, \tag{4.5}$$

where in the last inequality we have used the boundedness property of σ. We use (4.5) in (4.4) to get

$$\|X^j - X^{j-1}\|_{\mathbb{L}^2}^2 \le Ck\mathcal{J}(X^{j-1}) + C|\Delta_j W|^2. \tag{4.6}$$

Again, since $\mathscr{O}$ is a bounded domain, one has

$$\|X^{j-1}\|_{\mathbb{L}^2}^2 = \int_{\mathscr{O}}\left(|X^{j-1}|^2 - 1\right)\mathrm{d}x + |\mathscr{O}| \le C\left(1 + \frac{1}{4}\||X^{j-1}|^2 - 1\|_{\mathbb{L}^2}^2\right) \le C\left(1 + \mathcal{J}(X^{j-1})\right), \tag{4.7}$$

where $|\mathscr{O}|$ denotes the Lebesgue measure of $\mathscr{O}$. Combining the above estimates and then those for $\mathcal{A}_1$ and $\mathcal{A}_2$ in (4.2), and then (4.3), we obtain after taking expectation

$$\mathbb{E}\left[\mathcal{J}(X^j) - \mathcal{J}(X^{j-1})\right] + \frac{1}{4}\mathbb{E}\left[\|\nabla(X^j - X^{j-1})\|_{\mathbb{L}^2}^2\right] + \frac{1}{8}\mathbb{E}\left[\||X^j|^2 - |X^{j-1}|^2\|_{\mathbb{L}^2}^2\right]$$

$$+ k\mathbb{E}\Big[\|-\Delta X^j + f(X^j, X^{j-1})\|_{\mathbb{L}^2}^2\Big] \leq Ck\Big(1 + \mathbb{E}[\mathcal{J}(X^{j-1})]\Big).$$

Summation over all time steps, and the discrete Gronwall's lemma then establish the assertion for $r = 0$.

2. In order to validate the assertion for $p = 2^r, r \in \mathbb{N}^*$, we proceed inductively and illustrate the argument for $r = 1$. Recall that we have from before

$$\mathcal{J}(X^j) - \mathcal{J}(X^{j-1}) + \frac{1}{4}\|\nabla(X^j - X^{j-1})\|_{\mathbb{L}^2}^2 + \frac{1}{8}\||X^j|^2 - |X^{j-1}|^2\|_{\mathbb{L}^2}^2$$
$$+ k\|-\Delta X^j + f(X^j, X^{j-1})\|_{\mathbb{L}^2}^2$$
$$\leq C\mathcal{J}(X^{j-1})\Big\{k(1 + |\Delta_j W|^2) + |\Delta_j W|^2\Big\} + C|\Delta_j W|^2\big(1 + |\Delta_j W|^2\big)$$
$$+ \big(\nabla\sigma(X^{j-1}), \nabla X^{j-1}\big)_{\mathbb{L}^2}\Delta_j W + \big(\sigma(X^{j-1}), (|X^{j-1}|^2 - 1)X^{j-1}\big)_{\mathbb{L}^2}\Delta_j W. \quad (4.8)$$

To prove the assertion for $r = 1$, one needs to multiply (4.8) by some quantity to produce a term like $\mathcal{J}^2(X^j) - \mathcal{J}^2(X^{j-1}) + \alpha|\mathcal{J}(X^j) - \mathcal{J}(X^{j-1})|^2$ with $\alpha > 0$ on the left hand side of the inequality in order to absorb related terms coming from the right-hand side of the inequality before discrete Gronwall's lemma. Therefore, we multiply (4.8) with $\mathcal{J}(X^j) + \frac{1}{2}\mathcal{J}(X^{j-1})$ to get by binomial formula

$$\frac{3}{4}\Big(\mathcal{J}^2(X^j) - \mathcal{J}^2(X^{j-1})\Big) + \frac{1}{4}|\mathcal{J}(X^j) - \mathcal{J}(X^{j-1})|^2 + \frac{1}{2}\big(\mathcal{J}(X^j) + \mathcal{J}(X^{j-1})\big) \times$$
$$\Big\{\frac{1}{4}\|\nabla(X^j - X^{j-1})\|_{\mathbb{L}^2}^2 + \frac{1}{8}\||X^j|^2 - |X^{j-1}|^2\|_{\mathbb{L}^2}^2 + k\|-\Delta X^j + f(X^j, X^{j-1})\|_{\mathbb{L}^2}^2\Big\}$$
$$\leq C\mathcal{J}(X^{j-1})\big(\mathcal{J}(X^j) + \frac{1}{2}\mathcal{J}(X^{j-1})\big)\Big\{k(1 + |\Delta_j W|^2) + |\Delta_j W|^2\Big\}$$
$$+ C\big(\mathcal{J}(X^j) + \frac{1}{2}\mathcal{J}(X^{j-1})\big)|\Delta_j W|^2\big(1 + |\Delta_j W|^2\big)$$
$$+ \big(\mathcal{J}(X^j) + \frac{1}{2}\mathcal{J}(X^{j-1})\big)\big(\sigma(X^{j-1}), (|X^{j-1}|^2 - 1)X^{j-1}\big)_{\mathbb{L}^2}\Delta_j W$$
$$+ \big(\mathcal{J}(X^j) + \frac{1}{2}\mathcal{J}(X^{j-1})\big)\big(\nabla\sigma(X^{j-1}), \nabla X^{j-1}\big)_{\mathbb{L}^2}\Delta_j W := \mathcal{A}_3 + \mathcal{A}_4 + \mathcal{A}_5 + \mathcal{A}_6. \quad (4.9)$$

By Young's inequality, we have $(\theta_1, \theta_2 > 0)$

$$\mathcal{A}_3 \leq \theta_1|\mathcal{J}(X^j) - \mathcal{J}(X^{j-1})|^2 + C(\theta_1)\mathcal{J}^2(X^{j-1})\Big\{k(1 + |\Delta_j W|^2) + |\Delta_j W|^2\Big\}^2$$
$$+ C\mathcal{J}^2(X^{j-1})\Big\{k(1 + |\Delta_j W|^2) + |\Delta_j W|^2\Big\},$$
$$\mathcal{A}_4 \leq \theta_2|\mathcal{J}(X^j) - \mathcal{J}(X^{j-1})|^2 + C(\theta_2)|\Delta_j W|^4(1 + |\Delta_j W|^4) + C\mathcal{J}^2(X^{j-1})|\Delta_j W|^4$$
$$+ C(1 + |\Delta_j W|^4).$$

We can decompose $\mathcal{A}_6$ as

$$\mathcal{A}_6 = \big(\mathcal{J}(X^j) - \mathcal{J}(X^{j-1})\big)\big(\nabla\sigma(X^{j-1}), \nabla X^{j-1}\big)_{\mathbb{L}^2}\Delta_j W$$
$$+ \frac{3}{2}\mathcal{J}(X^{j-1})\big)\big(\nabla\sigma(X^{j-1}), \nabla X^{j-1}\big)_{\mathbb{L}^2}\Delta_j W := \mathcal{A}_{6,1} + \mathcal{A}_{6,2}.$$

Note that $\mathbb{E}\big[\mathcal{A}_{6,2}\big] = 0$. By using Young's inequality and the boundedness of σ', we estimate $\mathcal{A}_{6,1}$,

$$\mathcal{A}_{6,1} \leq \theta_3|\mathcal{J}(X^j) - \mathcal{J}(X^{j-1})|^2 + C(\theta_3)\|\nabla X^{j-1}\|_{\mathbb{L}^2}^4|\Delta_j W|^2$$
$$\leq \theta_3|\mathcal{J}(X^j) - \mathcal{J}(X^{j-1})|^2 + C(\theta_3)|\Delta_j W|^2\big(1 + \mathcal{J}^2(X^{j-1})\big).$$

Again, $\mathcal{A}_5$ can be written as $\mathcal{A}_{5,1} + \mathcal{A}_{5,2}$ with $\mathbb{E}\big[\mathcal{A}_{5,2}\big] = 0$, where

$$\mathcal{A}_{5,1} = \big(\mathcal{J}(X^j) - \mathcal{J}(X^{j-1})\big)\big(\sigma(X^{j-1}), (|X^{j-1}|^2 - 1)X^{j-1}\big)_{\mathbb{L}^2}\Delta_j W.$$

Thanks to Young's inequality, the boundedness of σ and (4.7) we get for $\theta_4 > 0$

$$
\begin{aligned}
\mathcal{A}_{5,1} &\le \theta_4 \big|\mathcal{J}(X^j) - \mathcal{J}(X^{j-1})\big|^2 + C(\theta_4)\big\||X^{j-1}|^2 - 1\big\|_{\mathbb{L}^2}^2 \|X^{j-1}\|_{\mathbb{L}^2}^2 |\Delta_j W|^2 \\
&\le \theta_4 \big|\mathcal{J}(X^j) - \mathcal{J}(X^{j-1})\big|^2 + C(\theta_4)\big(\mathcal{J}^2(X^{j-1}) + \|X^{j-1}\|_{\mathbb{L}^2}^4\big)|\Delta_j W|^2 \\
&\le \theta_4 \big|\mathcal{J}(X^j) - \mathcal{J}(X^{j-1})\big|^2 + C(\theta_4)\big(1 + \mathcal{J}^2(X^{j-1})\big)|\Delta_j W|^2.
\end{aligned}
$$

We combine all the above estimates in (4.9), and choose $\theta_1, \cdots, \theta_4 > 0$ with $\sum_{i=1}^4 \theta_i < \frac{1}{4}$ to have, after taking expectation

$$
\begin{aligned}
\mathbb{E}\Big[\mathcal{J}^2(X^j) &- \mathcal{J}^2(X^{j-1}) + C_1\big|\mathcal{J}(X^j) - \mathcal{J}(X^{j-1})\big|^2\Big] + C_2 \mathbb{E}\Big[\big(\mathcal{J}(X^j) + \mathcal{J}(X^{j-1})\big) \\
&\times \Big\{\|\nabla(X^j - X^{j-1})\|_{\mathbb{L}^2}^2 + \big\||X^j|^2 - |X^{j-1}|^2\big\|_{\mathbb{L}^2}^2 + k\| - \Delta X^j + f(X^j, X^{j-1})\|_{\mathbb{L}^2}^2\Big\}\Big] \\
&\le C_3(1 + k) + C_4 k \mathbb{E}\Big[\mathcal{J}^2(X^{j-1})\Big].
\end{aligned}
\tag{4.10}
$$

Summation over all time steps $0 \le j \le J$ in (4.10), together with the discrete Gronwall's lemma then validates the assertion of the theorem for $r = 1$. This completes the proof. $\qquad\square$

We employ the bounds for arbitrary moments of X in the strong norms in Lemma 3.1, (i), and a weak monotonicity argument to prove the following error estimate for the solution $\{X^j;\, 0 \le j \le J\}$ of (4.1).

Theorem 4.2. *Assume that $x \in \mathbb{W}_{\mathrm{per}}^{2,2}$, and the assumption $\mathbf{A}$.1 holds true. Then, for every $\delta > 0$, there exist constants $0 \le C_\delta < \infty$ and $k_1 = k_1(x, T) > 0$ such that for all $k \le k_1$ sufficiently small*

$$\sup_{0 \le j \le J} \mathbb{E}\big[\|X_{t_j} - X^j\|_{\mathbb{L}^2}^2\big] + k\sum_{j=0}^J \mathbb{E}\big[\|\nabla(X_{t_j} - X^j)\|_{\mathbb{L}^2}^2\big] \le C_\delta k^{1-\delta},$$

where $\{X_t;\, t \in [0,T]\}$ solves (2.2) while $\{X^j;\, 0 \le j \le J\}$ solves (4.1).

The parameter $\delta > 0$ which appears in Theorem 4.2 is due to the non-Lipschitz drift in the problem and is caused by the estimate (4.12) below.

Proof. Consider (2.2) for the time interval $[t_{j-1}, t_j]$, and denote $e^j := X_{t_j} - X^j$. There holds $\mathbb{P}$-a.s. for all $\phi \in \mathbb{W}_{\mathrm{per}}^{1,2}$

$$
\begin{aligned}
\big(e^j - e^{j-1}, \phi\big)_{\mathbb{L}^2} &+ \int_{t_{j-1}}^{t_j} \Big(\big(\nabla[X_{t_j} - X^j], \nabla\phi\big)_{\mathbb{L}^2} + \big(D\psi(X_{t_j}) - D\psi(X^j), \phi\big)_{\mathbb{L}^2}\Big)\mathrm{d}s \\
&= -\int_{t_{j-1}}^{t_j} \big(\nabla[X_s - X_{t_j}], \nabla\phi\big)_{\mathbb{L}^2}\,\mathrm{d}s - \int_{t_{j-1}}^{t_j} \big(D\psi(X_s) - D\psi(X_{t_j}), \phi\big)_{\mathbb{L}^2}\mathrm{d}s \\
&\quad -\frac{1}{2}\int_{t_{j-1}}^{t_j} \big((|X^j|^2 - 1)(X^j - X^{j-1}), \phi\big)_{\mathbb{L}^2}\,\mathrm{d}s + \int_{t_{j-1}}^{t_j} \big(\sigma(X_s) - \sigma(X_{t_{j-1}}), \phi\big)_{\mathbb{L}^2}\,\mathrm{d}W_s \\
&\qquad -\int_{t_{j-1}}^{t_j} \big(\sigma(X_{t_{j-1}}) - \sigma(X^{j-1}), \phi\big)_{\mathbb{L}^2}\,\mathrm{d}W_s \\
&=: I_j + II_j + III_j + IV_j + V_j.
\end{aligned}
\tag{4.11}
$$

The third term on the right-hand side attributes to the use of $f(X^j, X^{j-1})$ instead of $D\psi(X^j)$ in (4.1). Choose $\phi = e^j(\omega)$, and apply expectation. By the weak monotonicity property (1.3) of the drift, the left-hand side of (4.11) is then bounded from below by

$$\frac{1}{2}\mathbb{E}\Big[\|e^j\|_{\mathbb{L}^2}^2 - \|e^{j-1}\|_{\mathbb{L}^2}^2 + \|e^j - e^{j-1}\|_{\mathbb{L}^2}^2 + 2k\big(\|\nabla e^j\|_{\mathbb{L}^2}^2 - \|e^j\|_{\mathbb{L}^2}^2\big)\Big].$$

Because of Young's inequality and Lemma 3.2, (ii) we conclude

$$\mathbb{E}[I_j] \le Ck^2 + \frac{k}{8}\mathbb{E}\Big[\|\nabla e^j\|_{\mathbb{L}^2}^2\Big].$$

Next we bound $\mathbb{E}[II_j]$. For this purpose, we use the embedding $\mathbb{W}^{1,2} \hookrightarrow \mathbb{L}^6$ for $d \le 3$, the algebraic identity $a^3 - b^3 = \frac{1}{2}(a-b)\big((a+b)^2 + a^2 + b^2\big)$, and Young's and Hölder's inequalities in combination with Lemma 3.2 to estimate ($\delta > 0$)

$$
\begin{aligned}
\mathbb{E}[II_j] &\le \frac{1}{2}\int_{t_{j-1}}^{t_j} \mathbb{E}\Big[\|X_s - X_{t_j}\|_{\mathbb{L}^2}\|(X_s + X_{t_j})^2 + X_s^2 + X_{t_j}^2\|_{\mathbb{L}^3}\|e^j\|_{\mathbb{L}^6}\Big]\,\mathrm{d}s \\
&\quad + \int_{t_{j-1}}^{t_j} \mathbb{E}\Big[\|X_s - X_{t_j}\|_{\mathbb{L}^2}\|e^j\|_{\mathbb{L}^2}\Big]\,\mathrm{d}s \\
&\le Ck \sup_{s \in [t_{j-1}, t_j]} \mathbb{E}\Big[\|X_s - X_{t_j}\|_{\mathbb{L}^2}^2\big(\|X_{t_j}\|_{\mathbb{L}^6}^4 + \|X_s\|_{\mathbb{L}^6}^4\big)\Big] + \frac{k}{8}\mathbb{E}\Big[\|e^j\|_{\mathbb{W}^{1,2}}^2\Big] + Ck^2 \\
&\le Ck \sup_{s \in [t_{j-1}, t_j]} \Big(\mathbb{E}\Big[\|X_s - X_{t_j}\|_{\mathbb{L}^2}^{2(1+\delta)}\Big]\Big)^{\frac{1}{1+\delta}} \Big(\mathbb{E}\Big[\big(\|X_{t_j}\|_{\mathbb{W}^{1,2}}^4 + \|X_s\|_{\mathbb{W}^{1,2}}^4\big)^{\frac{1+\delta}{\delta}}\Big]\Big)^{\frac{\delta}{1+\delta}} \\
&\quad + \frac{k}{8}\mathbb{E}\Big[\|\nabla e^j\|_{\mathbb{L}^2}^2\Big] + Ck\mathbb{E}\Big[\|e^j\|_{\mathbb{L}^2}^2\Big] + Ck^2. \tag{4.12}
\end{aligned}
$$

The leading factor is bounded by $Ck^{\frac{1}{1+\delta}}$ by Lemma 3.2, (i), while the second factor may be bounded by C_δ due to (3.5). Thus we have

$$\mathbb{E}[II_j] \le C_\delta k^{\frac{2+\delta}{1+\delta}} + Ck^2 + \frac{k}{8}\mathbb{E}\Big[\|\nabla e^j\|_{\mathbb{L}^2}^2\Big] + Ck\mathbb{E}\Big[\|e^j\|_{\mathbb{L}^2}^2\Big].$$

It is immediate to validate

$$\big|\mathbb{E}[IV_j]\big| + \big|\mathbb{E}[V_j]\big| \le Ck^2 + \frac{1}{8}\|e^j - e^{j-1}\|_{\mathbb{L}^2}^2 + Ck\,\|e^{j-1}\|_{\mathbb{L}^2}^2$$

by adding and subtracting e^{j-1} in the second argument and proceeding as before, and Itô's isometry in combination with (2.1) and Lemma 3.2, (i). Next we focus on the term III_j. In view of generalized Hölder's inequality, and the embedding $\mathbb{W}^{1,2} \hookrightarrow \mathbb{L}^6$ for $d \le 3$,

$$
\begin{aligned}
\mathbb{E}\big[III_j\big] &\le \frac{1}{2}\mathbb{E}\Big[\int_{t_{j-1}}^{t_j} \|e^j\|_{\mathbb{L}^6}\|X^j - X^{j-1}\|_{\mathbb{L}^2}\||X^j|^2 - 1\|_{\mathbb{L}^3}\,\mathrm{d}s\Big] \\
&\le \frac{k}{8}\mathbb{E}\big[\|e^j\|_{\mathbb{L}^6}^2\big] + Ck\mathbb{E}\Big[\|X^j - X^{j-1}\|_{\mathbb{L}^2}^2\||X^j|^2 - 1\|_{\mathbb{L}^3}^2\Big] \\
&\le \frac{k}{8}\mathbb{E}\big[\|e^j\|_{\mathbb{W}^{1,2}}^2\big] + Ck\mathbb{E}\Big[\|X^j - X^{j-1}\|_{\mathbb{L}^2}^2\big(1 + \|X^j\|_{\mathbb{W}^{1,2}}^4\big)\Big] \\
&\equiv \frac{k}{8}\mathbb{E}\big[\|e^j\|_{\mathbb{W}^{1,2}}^2\big] + III_{j,1}.
\end{aligned}
$$

In view of Lemma 4.1, we see that

$$\sup_j \mathbb{E}\big[\|X^j\|_{\mathbb{W}^{1,2}}^p\big] \le C \quad \text{for any } p \ge 2, \tag{4.13}$$

We use (4.6), Lemma 4.1 and (4.13) to estimate $III_{j,1}$,

$$III_{j,1} \leq Ck\mathbb{E}\Big[\big(k\mathcal{J}(X^{j-1}) + |\Delta_j W|^2\big)\big(1 + \|X^j\|_{\mathbb{W}^{1,2}}^4\big)\Big]$$

$$\leq Ck^2\mathbb{E}\Big[1 + \mathcal{J}(X^{j-1}) + \big(\mathcal{J}(X^{j-1})\big)^2 + \|X^j\|_{\mathbb{W}^{1,2}}^8\Big] + C\mathbb{E}\Big[k|\Delta_j W|^2\|X^j\|_{\mathbb{W}^{1,2}}^4\Big]$$

$$\leq Ck^2 + Ck^2\mathbb{E}\Big[\|X^j\|_{\mathbb{W}^{1,2}}^8\Big] + C\mathbb{E}\Big[|\Delta_j W|^4\Big] \leq Ck^2,$$

and therefore we obtain

$$\mathbb{E}\big[III_j\big] \leq \frac{k}{8}\mathbb{E}\Big[\|\nabla e^j\|_{\mathbb{L}^2}^2\Big] + Ck\mathbb{E}\big[\|e^j\|_{\mathbb{L}^2}^2\big] + Ck^2.$$

We combine all the above estimates to have

$$\mathbb{E}\Big[\|e^j\|_{\mathbb{L}^2}^2 - \|e^{j-1}\|_{\mathbb{L}^2}^2 + k\|\nabla e^j\|_{\mathbb{L}^2}^2\Big] \leq Ck^2 + C_\delta k^{\frac{2+\delta}{1+\delta}} + Ck\Big(\mathbb{E}\Big[\|e^j\|_{\mathbb{L}^2}^2 + \|e^{j-1}\|_{\mathbb{L}^2}^2\Big]\Big). \qquad (4.14)$$

Summation over all time steps $0 \leq j \leq J$ in (4.14), together with the discrete (implicit form) Gronwall's lemma then validates the assertion of the theorem. $\qquad \square$

5. Space-time discretization and strong error estimate

In this section, we first derive the uniform moment estimate for the discretized solution $\{Y^j : 0 \leq j \leq J\}$ of the structure preserving finite element based fully discrete scheme (2.5). Then by using these uniform bounds along with Lemma 4.1 we bound the error $E^j := X^j - Y^j$, where $\{X^j; 0 \leq j \leq J\}$ solves (4.1).

We define the discrete Laplacian $\Delta_h : \mathbb{V}_h \to \mathbb{V}_h$ by the variational identity

$$-\big(\Delta_h \phi_h, \psi_h\big)_{\mathbb{L}^2} = \big(\nabla \phi_h, \nabla \psi_h\big)_{\mathbb{L}^2} \quad \forall \phi_h, \psi_h \in \mathbb{V}_h.$$

One can use the test function $\phi = -\Delta_h Y^j + \mathscr{P}_{\mathbb{L}^2} f(Y^j, Y^{j-1}) \in \mathbb{V}_h$ in (2.5) and proceed as in the proof of Lemma 4.1 along with (2.4), the $\mathbb{W}^{1,2}$ and $\mathbb{L}^q$ $(1 \leq q \leq \infty)$-stabilities of the projection operator $\mathscr{P}_{\mathbb{L}^2}$ (cf. [3]) to arrive at the following uniform moment estimates for $\{Y^j; 0 \leq j \leq J\}$.

Lemma 5.1. *For every $p = 2^r$, $r \in \mathbb{N}^*$, there exists a constant $C \equiv C(p, T) > 0$ such that*

$$\max_{1 \leq j \leq J}\mathbb{E}\Big[\big|\mathcal{J}(Y^j)\big|^p\Big] + \sum_{j=1}^{J}\mathbb{E}\Bigg[\prod_{\ell=1}^{r}\big[[\mathcal{J}(Y^j)]^{2^{\ell-1}} + [\mathcal{J}(Y^{j-1})]^{2^{\ell-1}}\big] \times \Big(\|\nabla(Y^j - Y^{j-1})\|_{\mathbb{L}^2}^2$$

$$+ \big\||Y^j|^2 - |Y^{j-1}|^2\big\|_{\mathbb{L}^2}^2 + k\big\|-\Delta_h Y^j + \mathscr{P}_{\mathbb{L}^2} f(Y^j, Y^{j-1})\big\|_{\mathbb{L}^2}^2\Big)\Bigg] \leq C,$$

provided $\mathbb{E}\big[|\mathcal{J}(Y^0)|^p\big] \leq C$.

In view of Lemma 5.1, it follows that

$$\sup_{0 \leq j \leq J}\mathbb{E}\Big[\|Y^j\|_{\mathbb{W}^{1,2}}^p\Big] \leq C, \quad \forall p \geq 2. \qquad (5.1)$$

We have the following theorem regarding the error E^j in strong norm.

Theorem 5.2. *Assume that $x \in \mathbb{W}_{\mathrm{per}}^{2,2}$. Then, under the assumption **A**.1, there exist constants $C > 0$, independent of the discretization parameters $h, k > 0$ and $k_2 \equiv k_2(T, x) > 0$ such that for all $k \leq k_2$ sufficiently small, there holds*

$$\sup_{0 \leq j \leq J}\mathbb{E}\Big[\|X^j - Y^j\|_{\mathbb{L}^2}^2\Big] + k\sum_{j=0}^{J}\mathbb{E}\Big[\|\nabla(X^j - Y^j)\|_{\mathbb{L}^2}^2\Big] \leq C\big(k + h^2\big),$$

where $\{X^j; 0 \leq j \leq J\}$ solves (4.1) while $\{Y^j; 0 \leq j \leq J\}$ solves (2.5).

Proof. We subtract (2.5) from (4.1), and restrict to the test functions $\phi \in \mathbb{V}_h$. Choosing $\phi = \mathscr{P}_{\mathbb{L}^2} E^j(\omega)$, and using (2.4), we obtain

$$\frac{1}{2}\mathbb{E}\Big[\|\mathscr{P}_{\mathbb{L}^2} E^j\|_{\mathbb{L}^2}^2 - \|\mathscr{P}_{\mathbb{L}^2} E^{j-1}\|_{\mathbb{L}^2}^2 + \|\mathscr{P}_{\mathbb{L}^2}[E^j - E^{j-1}]\|_{\mathbb{L}^2}^2\Big]$$

$$+ k\mathbb{E}\Big[\big(\nabla E^j, \nabla E^j\big)_{\mathbb{L}^2} + \big(D\psi(X^j) - D\psi(Y^j), E^j\big)_{\mathbb{L}^2}\Big]$$

$$= k\mathbb{E}\Big[\big(\nabla E^j, \nabla(E^j - \mathscr{P}_{\mathbb{L}^2} E^j)\big)_{\mathbb{L}^2} + \big(D\psi(X^j) - D\psi(Y^j), E^j - \mathscr{P}_{\mathbb{L}^2} E^j\big)_{\mathbb{L}^2}\Big]$$

$$+ k\mathbb{E}\Big[\Big((|X^j|^2 - 1)\frac{X^j - X^{j-1}}{2} - (|Y^j|^2 - 1)\frac{Y^j - Y^{j-1}}{2}, \mathscr{P}_{\mathbb{L}^2} E^j\Big)_{\mathbb{L}^2}\Big]$$

$$+ \mathbb{E}\Big[\big(\sigma(X^{j-1}) - \sigma(Y^{j-1}), \mathscr{P}_{\mathbb{L}^2}[E^j - E^{j-1}]\big)_{\mathbb{L}^2}\Delta_j W\Big]$$

$$\leq \frac{k}{2}\mathbb{E}\Big[\|\nabla E^j\|_{\mathbb{L}^2}^2\Big] + \frac{k}{2}\mathbb{E}\Big[\|\nabla(X^j - \mathscr{P}_{\mathbb{L}^2} X^j)\|_{\mathbb{L}^2}^2\Big] - k\mathbb{E}\Big[\|E^j\|_{\mathbb{L}^2}^2\Big]$$

$$+ k\mathbb{E}\Big[\|\mathscr{P}_{\mathbb{L}^2} E^j\|_{\mathbb{L}^2}^2\Big] + \Big|\mathbb{E}\Big[\big(|X^j|^2 X^j - |Y^j|^2 Y^j, X^j - \mathscr{P}_{\mathbb{L}^2} X^j\big)_{\mathbb{L}^2}\Big]\Big|$$

$$+ k\mathbb{E}\Big[\Big((|X^j|^2 - 1)\frac{X^j - X^{j-1}}{2} - (|Y^j|^2 - 1)\frac{Y^j - Y^{j-1}}{2}, \mathscr{P}_{\mathbb{L}^2} E^j\Big)_{\mathbb{L}^2}\Big]$$

$$+ \mathbb{E}\Big[\big(\sigma(X^{j-1}) - \sigma(Y^{j-1}), \mathscr{P}_{\mathbb{L}^2}[E^j - E^{j-1}]\big)_{\mathbb{L}^2}\Delta_j W\Big].$$

Note that the third term on the right-hand side of the first equality reflects that $f(X^j, X^{j-1})$ is a perturbation of $D\psi(X^j)$. By the weak monotonicity property (1.3), we see that

$$\mathbb{E}\Big[\|\nabla E^j\|_{\mathbb{L}^2}^2 - \|E^j\|_{\mathbb{L}^2}^2\Big] \leq \mathbb{E}\Big[\big(\nabla E^j, \nabla E^j\big)_{\mathbb{L}^2} + \big(D\psi(X^j) - D\psi(Y^j), E^j\big)_{\mathbb{L}^2}\Big],$$

and therefore we arrive at the following inequality

$$\frac{1}{2}\mathbb{E}\Big[\big(\|\mathscr{P}_{\mathbb{L}^2} E^j\|_{\mathbb{L}^2}^2 - \|\mathscr{P}_{\mathbb{L}^2} E^{j-1}\|_{\mathbb{L}^2}^2\big) + \|\mathscr{P}_{\mathbb{L}^2}[E^j - E^{j-1}]\|_{\mathbb{L}^2}^2 + k\|\nabla E^j\|_{\mathbb{L}^2}^2\Big]$$

$$\leq Ck\mathbb{E}\big[\|\mathscr{P}_{\mathbb{L}^2} E^j\|_{\mathbb{L}^2}^2\big] + Ck\mathbb{E}\Big[\|\nabla(X^j - \mathscr{P}_{\mathbb{L}^2} X^j)\|_{\mathbb{L}^2}^2\Big]$$

$$+ Ck\Big|\mathbb{E}\Big[\big(|X^j|^2 X^j - |Y^j|^2 Y^j, X^j - \mathscr{P}_{\mathbb{L}^2} X^j\big)_{\mathbb{L}^2}\Big]\Big|$$

$$+ k\mathbb{E}\Big[\Big((|X^j|^2 - 1)\frac{X^j - X^{j-1}}{2} - (|Y^j|^2 - 1)\frac{Y^j - Y^{j-1}}{2}, \mathscr{P}_{\mathbb{L}^2} E^j\Big)_{\mathbb{L}^2}\Big]$$

$$+ \mathbb{E}\Big[\big(\sigma(X^{j-1}) - \sigma(Y^{j-1}), \mathscr{P}_{\mathbb{L}^2}[E^j - E^{j-1}]\big)_{\mathbb{L}^2}\Delta_j W\Big]$$

$$=: Ck\mathbb{E}\big[\|\mathscr{P}_{\mathbb{L}^2} E^j\|_{\mathbb{L}^2}^2\big] + \mathbf{B}_{1,j} + \mathbf{B}_{2,j} + \mathbf{B}_{3,j} + \mathbf{B}_{4,j}. \tag{5.2}$$

Note that, in view of Lemma 4.1, Young's inequality and the embedding $\mathbb{W}^{1,2} \hookrightarrow \mathbb{L}^6$ for $d \leq 3$

$$\mathbb{E}\Big[\|f(X^j, X^{j-1})\|_{\mathbb{L}^2}^2\Big] \leq C\mathbb{E}\Big[\int_{\mathscr{O}}(|X^j|^4 + 1)(|X^j|^2 + |X^{j-1}|^2)\,dx\Big]$$

$$\leq C\mathbb{E}\Big[\int_{\mathscr{O}}\big(|X^j|^6 + |X^{j-1}|^6 + |X^j|^2 + |X^{j-1}|^2\big)\,dx\Big]$$

$$\leq C\Big(1 + \sup_j \mathbb{E}\big[|\mathcal{J}(X^j)|^8\big]\Big).$$

Thus using Lemma 4.1 and the estimate above, we see that

$$k\sum_{j=1}^{J}\mathbb{E}\big[\|\Delta X^j\|_{\mathbb{L}^2}^2\big]$$

$$\leq k \sum_{j=1}^{J} \mathbb{E}\big[\| - \Delta X^j + f(X^j, X^{j-1})\|_{\mathbb{L}^2}^2\big] + k \sum_{j=1}^{J} \mathbb{E}\big[\|f(X^j, X^{j-1})\|_{\mathbb{L}^2}^2\big] \leq C. \qquad (5.3)$$

Let us recall the following well-known properties of $\mathscr{P}_{\mathbb{L}^2}$, see [1]

$$\begin{cases} \|g - \mathscr{P}_{\mathbb{L}^2}g\|_{\mathbb{L}^2} \leq Ch\|g\|_{\mathbb{W}^{1,2}} \quad \forall g \in \mathbb{W}^{1,2}, \\ \|g - \mathscr{P}_{\mathbb{L}^2}g\|_{\mathbb{L}^2} + h\|\nabla[g - \mathscr{P}_{\mathbb{L}^2}g]\|_{\mathbb{L}^2} \leq Ch^2\|\Delta g\|_{\mathbb{L}^2} \quad \forall g \in \mathbb{W}^{2,2}. \end{cases} \qquad (5.4)$$

We use (5.3) and (5.4) to infer that

$$\sum_{j=1}^{J} \mathbf{B}_{1,j} \leq Ch^2 \sum_{j=1}^{J} k \mathbb{E}\big[\|\Delta X^j\|_{\mathbb{L}^2}^2\big] \leq Ch^2.$$

Next we estimate $\sum_{j=1}^{J} \mathbf{B}_{4,j}$. A simple approximation argument, (2.1), and (5.3) together with Young's inequality lead to

$$\sum_{j=1}^{J} \mathbf{B}_{4,j} \leq \sum_{j=1}^{J} \mathbb{E}\Big[\|\sigma(X^{j-1}) - \sigma(Y^{j-1})\|_{\mathbb{L}^2}\|\mathscr{P}_{\mathbb{L}^2}[E^j - E^{j-1}]\|_{\mathbb{L}^2}|\Delta_j W|\Big]$$

$$\leq \frac{1}{4} \sum_{j=1}^{J} \mathbb{E}\Big[\|\mathscr{P}_{\mathbb{L}^2}[E^j - E^{j-1}]\|_{\mathbb{L}^2}^2\Big] + C \sum_{j=1}^{J} k \mathbb{E}\Big[\|E^{j-1}\|_{\mathbb{L}^2}^2\Big]$$

$$\leq \frac{1}{4} \sum_{j=1}^{J} \mathbb{E}\Big[\|\mathscr{P}_{\mathbb{L}^2}[E^j - E^{j-1}]\|_{\mathbb{L}^2}^2\Big] + Ck \sum_{j=1}^{J} \mathbb{E}\Big[\|\mathscr{P}_{\mathbb{L}^2}E^{j-1}\|_{\mathbb{L}^2}^2 + \|X^{j-1} - \mathscr{P}_{\mathbb{L}^2}X^{j-1}\|_{\mathbb{L}^2}^2\Big]$$

$$\leq \frac{1}{4} \sum_{j=1}^{J} \mathbb{E}\big[\|\mathscr{P}_{\mathbb{L}^2}[E^j - E^{j-1}]\|_{\mathbb{L}^2}^2\big] + C\Big(h^4 + k \sum_{j=1}^{J} \mathbb{E}\big[\|\mathscr{P}_{\mathbb{L}^2}E^{j-1}\|_{\mathbb{L}^2}^2\big]\Big).$$

We now bound the term $\mathbf{B}_{2,j}$. We use the algebraic formula given before (4.12), the embedding $\mathbb{W}^{1,2} \hookrightarrow \mathbb{L}^6$ for $d \leq 3$, and a generalized Young's inequality to have

$$\mathbf{J}_{2,j} \leq Ck\mathbb{E}\Big[\|E^j\|_{\mathbb{L}^6}\big(\|X^j\|_{\mathbb{L}^6}^2 + \|Y^j\|_{\mathbb{L}^6}^2\big)\|X^j - \mathscr{P}_{\mathbb{L}^2}X^j\|_{\mathbb{L}^2}\Big]$$

$$\leq Ck\mathbb{E}\Big[\|E^j\|_{\mathbb{W}^{1,2}}\big(\|X^j\|_{\mathbb{W}^{1,2}}^2 + \|Y^j\|_{\mathbb{W}^{1,2}}^2\big)\|X^j - \mathscr{P}_{\mathbb{L}^2}X^j\|_{\mathbb{L}^2}\Big]$$

$$\leq \frac{k}{8}\mathbb{E}\Big[\|\nabla E^j\|_{\mathbb{L}^2}^2\Big] + \frac{k}{8}\mathbb{E}\Big[\|E^j\|_{\mathbb{L}^2}^2\Big] + Ck\mathbb{E}\Big[\big(\|X^j\|_{\mathbb{W}^{1,2}}^4 + \|Y^j\|_{\mathbb{W}^{1,2}}^4\big)\|X^j - \mathscr{P}_{\mathbb{L}^2}X^j\|_{\mathbb{L}^2}^2\Big]$$

$$=: \frac{k}{8}\mathbb{E}\Big[\|\nabla E^j\|_{\mathbb{L}^2}^2\Big] + \mathbf{B}_{2,j}^1 + \mathbf{B}_{2,j}^2.$$

Thanks to (4.13) and (5.4), we note that

$$\sum_{j=0}^{J} \mathbf{B}_{2,j}^1 \leq \sum_{j=0}^{J} k\mathbb{E}\Big[\|\mathscr{P}_{\mathbb{L}^2}E^j\|_{\mathbb{L}^2}^2\Big] + Ch^2 k \sum_{j=0}^{J} \mathbb{E}\Big[\|X^j\|_{\mathbb{W}^{1,2}}^2\Big] \leq Ch^2 + \sum_{j=0}^{J} k\mathbb{E}\Big[\|\mathscr{P}_{\mathbb{L}^2}E^j\|_{\mathbb{L}^2}^2\Big].$$

We use (4.13), (5.1) and (5.4), together with Young's inequality to get

$$\sum_{j=0}^{J} \mathbf{B}_{2,j}^2 \leq Ch^2 k \sum_{j=0}^{J} \mathbb{E}\Big[\big(\|X^j\|_{\mathbb{W}^{1,2}}^4 + \|Y^j\|_{\mathbb{W}^{1,2}}^4\big)\|X^j\|_{\mathbb{W}^{1,2}}^2\Big]$$

$$\leq Ch^2 k \sum_{j=0}^{J} \mathbb{E}\Big[\|X^j\|_{\mathbb{W}^{1,2}}^8 + \|Y^j\|_{\mathbb{W}^{1,2}}^8 + \|X^j\|_{\mathbb{W}^{1,2}}^4\Big] \leq Ch^2.$$

It remains to bound $\mathbf{B}_{3,j}$. We decompose $\mathbf{B}_{3,j}$ as follows.

$$\mathbf{B}_{3,j} = \frac{k}{2}\mathbb{E}\Big[\Big((|X^j|^2 - |Y^j|^2)(X^j - X^{j-1}), \mathscr{P}_{\mathbb{L}^2}E^j\Big)_{\mathbb{L}^2}\Big]$$
$$+ \frac{k}{2}\mathbb{E}\Big[\Big((|Y^j|^2 - 1)(E^j - E^{j-1}), \mathscr{P}_{\mathbb{L}^2}E^j\Big)_{\mathbb{L}^2}\Big] =: \mathbf{B}_{3,j}^1 + \mathbf{B}_{3,j}^2.$$

Thanks to generalized Hölder's inequality, the $L^q(1 \le q \le \infty)$-stability of $\mathscr{P}_{\mathbb{L}^2}$, the embedding $\mathbb{W}^{1,2} \hookrightarrow \mathbb{L}^6$ for $d \le 3$, the estimates (5.4), (5.1) and (4.13), we have

$$\mathbf{B}_{3,j}^2 \le Ck\mathbb{E}\Big[\|E^j - E^{j-1}\|_{\mathbb{L}^2}\|E^j\|_{\mathbb{W}^{1,2}}\||Y^j|^2 - 1\|_{\mathbb{L}^3}\Big]$$
$$\le Ck\mathbb{E}\Big[\Big(\|\mathscr{P}_{\mathbb{L}^2}(E^j - E^{j-1})\|_{\mathbb{L}^2} + \|X^j - \mathscr{P}_{\mathbb{L}^2}X^j - (X^{j-1} - \mathscr{P}_{\mathbb{L}^2}X^{j-1})\|_{\mathbb{L}^2}\Big)$$
$$\times \|E^j\|_{\mathbb{W}^{1,2}}\||Y^j|^2 - 1\|_{\mathbb{L}^3}\Big]$$
$$\le \frac{k}{16}\mathbb{E}\big[\|E^j\|_{\mathbb{W}^{1,2}}^2\big] + Ck\mathbb{E}\Big[\|\mathscr{P}_{\mathbb{L}^2}(E^j - E^{j-1})\|_{\mathbb{L}^2}^{1+1}\||Y^j|^2 - 1\|_{\mathbb{L}^3}^2\Big]$$
$$+ Ck\mathbb{E}\Big[\|X^j - \mathscr{P}_{\mathbb{L}^2}X^j - (X^{j-1} - \mathscr{P}_{\mathbb{L}^2}X^{j-1})\|_{\mathbb{L}^2}^2\||Y^j|^2 - 1\|_{\mathbb{L}^3}^2\Big]$$
$$\le \frac{k}{16}\mathbb{E}\big[\|\nabla E^j\|_{\mathbb{L}^2}^2\big] + Ck\mathbb{E}\big[\|\mathscr{P}_{\mathbb{L}^2}E^j\|_{\mathbb{L}^2}^2\big] + \frac{1}{8}\mathbb{E}\big[\|\mathscr{P}_{\mathbb{L}^2}(E^j - E^{j-1})\|_{\mathbb{L}^2}^2\big]$$
$$+ Ck^2\mathbb{E}\Big[\|\mathscr{P}_{\mathbb{L}^2}(E^j - E^{j-1})\|_{\mathbb{L}^2}^2\||Y^j|^2 - 1\|_{\mathbb{L}^3}^4\Big] + Ckh^2\mathbb{E}\big[\|X^j\|_{\mathbb{W}^{1,2}}^2\big]$$
$$+ Ckh^2\mathbb{E}\Big[\Big(\|X^j\|_{\mathbb{W}^{1,2}}^2 + \|X^{j-1}\|_{\mathbb{W}^{1,2}}^2\Big)\||Y^j|^2 - 1\|_{\mathbb{L}^3}^2\Big]$$
$$\le \frac{k}{16}\mathbb{E}\big[\|\nabla E^j\|_{\mathbb{L}^2}^2\big] + \frac{1}{8}\mathbb{E}\Big[\|\mathscr{P}_{\mathbb{L}^2}(E^j - E^{j-1})\|_{\mathbb{L}^2}^2\Big] + Ck\mathbb{E}\big[\|\mathscr{P}_{\mathbb{L}^2}E^j\|_{\mathbb{L}^2}^2\big] + Ckh^2$$
$$+ Ck^2\mathbb{E}\Big[\|E^j - E^{j-1}\|_{\mathbb{L}^2}^2(1 + \|Y^j\|_{\mathbb{W}^{1,2}}^8)\Big]$$
$$+ Ckh^2\mathbb{E}\Big[1 + \|X^j\|_{\mathbb{W}^{1,2}}^4 + \|X^{j-1}\|_{\mathbb{W}^{1,2}}^4 + \|Y^j\|_{\mathbb{W}^{1,2}}^8\Big]$$
$$\le \frac{k}{16}\mathbb{E}\big[\|\nabla E^j\|_{\mathbb{L}^2}^2\big] + \frac{1}{8}\mathbb{E}\Big[\|\mathscr{P}_{\mathbb{L}^2}(E^j - E^{j-1})\|_{\mathbb{L}^2}^2\Big] + Ck\mathbb{E}\big[\|\mathscr{P}_{\mathbb{L}^2}E^j\|_{\mathbb{L}^2}^2\big] + Ck(h^2 + k).$$

Next we estimate $\mathbf{B}_{3,j}^1$. We use the generalized Hölder's inequality, the $L^q(1 \le q \le \infty)$-stability of $\mathscr{P}_{\mathbb{L}^2}$, the embedding $\mathbb{W}^{1,2} \hookrightarrow \mathbb{L}^6$ for $d \le 3$, Young's inequality, the estimates (5.1), (5.4) and (4.6), (4.7) and (4.13), along with Lemma 4.1 to get

$$\mathbf{B}_{3,j}^1 \le Ck\mathbb{E}\Big[\|\mathscr{P}_{\mathbb{L}^2}E^j\|_{\mathbb{L}^6}\|X^j - X^{j-1}\|_{\mathbb{L}^2}\||X^j|^2 - |Y^j|^2\|_{\mathbb{L}^3}\Big]$$
$$\le Ck\mathbb{E}\Big[\|E^j\|_{\mathbb{W}^{1,2}}\|X^j - X^{j-1}\|_{\mathbb{L}^2}\big(\|X^j\|_{\mathbb{L}^6}^2 + \|Y^j\|_{\mathbb{L}^6}^2\big)\Big]$$
$$\le \frac{k}{16}\mathbb{E}\big[\|\nabla E^j\|_{\mathbb{L}^2}^2\big] + \frac{k}{16}\mathbb{E}\big[\|E^j\|_{\mathbb{L}^2}^2\big] + Ck\mathbb{E}\Big[\|X^j - X^{j-1}\|_{\mathbb{L}^2}^2\big(\|X^j\|_{\mathbb{L}^6}^2 + \|Y^j\|_{\mathbb{L}^6}^2\big)^2\Big]$$
$$\le \frac{k}{16}\mathbb{E}\big[\|\nabla E^j\|_{\mathbb{L}^2}^2\big] + Ck\mathbb{E}\big[\|\mathscr{P}_{\mathbb{L}^2}E^j\|_{\mathbb{L}^2}^2\big] + Ckh^2\mathbb{E}\big[\|X^j\|_{\mathbb{W}^{1,2}}^2\big]$$
$$+ Ck^2\mathbb{E}\Big[\|X^j\|_{\mathbb{W}^{1,2}}^8 + \|Y^j\|_{\mathbb{W}^{1,2}}^8\Big] + C\mathbb{E}\Big[\|X^j - X^{j-1}\|_{\mathbb{L}^2}^4\Big]$$
$$\le \frac{k}{16}\mathbb{E}\big[\|\nabla E^j\|_{\mathbb{L}^2}^2\big] + Ck\mathbb{E}\big[\|\mathscr{P}_{\mathbb{L}^2}E^j\|_{\mathbb{L}^2}^2\big] + Ck(h^2 + k) + C\mathbb{E}\Big[k^2\mathcal{J}^2(X^{j-1}) + |\Delta_j W|^4\Big]$$
$$\le \frac{k}{16}\mathbb{E}\big[\|\nabla E^j\|_{\mathbb{L}^2}^2\big] + Ck\mathbb{E}\big[\|\mathscr{P}_{\mathbb{L}^2}E^j\|_{\mathbb{L}^2}^2\big] + Ck(h^2 + k) + Ck^2\big(1 + \mathbb{E}\big[\mathcal{J}^2(X^{j-1})\big]\big)$$
$$\le \frac{k}{16}\mathbb{E}\big[\|\nabla E^j\|_{\mathbb{L}^2}^2\big] + Ck\mathbb{E}\big[\|\mathscr{P}_{\mathbb{L}^2}E^j\|_{\mathbb{L}^2}^2\big] + Ck(h^2 + k).$$

Putting things together in (5.2) and using the discrete Gronwall's lemma (implicit form) then yields

$$\sup_{0 \le j \le J} \mathbb{E}\big[\|\mathscr{P}_{\mathbb{L}^2} E^j\|_{\mathbb{L}^2}^2\big] + k \sum_{j=0}^{J} \mathbb{E}\big[\|\nabla(X^j - Y^j)\|_{\mathbb{L}^2}^2\big] \le C(k + h^2). \tag{5.5}$$

Thus, thanks to (4.13), (5.4) and (5.5), we conclude that

$$\sup_{0 \le j \le J} \mathbb{E}\Big[\|X^j - Y^j\|_{\mathbb{L}^2}^2\Big] + k \sum_{j=0}^{J} \mathbb{E}\big[\|\nabla(X^j - Y^j)\|_{\mathbb{L}^2}^2\big]$$

$$\le \sup_{0 \le j \le J} \mathbb{E}\big[\|\mathscr{P}_{\mathbb{L}^2} E^j\|_{\mathbb{L}^2}^2\big] + k \sum_{j=0}^{J} \mathbb{E}\big[\|\nabla(X^j - Y^j)\|_{\mathbb{L}^2}^2\big] + Ch^2 \sup_{0 \le j \le J} \mathbb{E}\big[\|X^j\|_{\mathbb{W}^{1,2}}^2\big]$$

$$\le C(k + h^2).$$

This finishes the proof. $\qquad\square$

5.1. **Proof of Main Theorem.** Let the assumption **A**.1 hold and $x \in \mathbb{W}_{\mathrm{per}}^{2,2}$. Then thanks to Theorem 4.2, for every $\delta > 0$, there exist constants $0 \le C_\delta < \infty$ and $k_1 \equiv k_1(T, x) > 0$ such that for all $k \le k_1$ sufficiently small

$$\sup_{0 \le j \le J} \mathbb{E}\big[\|X_{t_j} - X^j\|_{\mathbb{L}^2}^2\big] + k \sum_{j=0}^{J} \mathbb{E}\big[\|\nabla(X_{t_j} - X^j)\|_{\mathbb{L}^2}^2\big] \le C_\delta k^{1-\delta}, \tag{5.6}$$

where $\{X_t; t \in [0, T]\}$ solves (2.2) while $\{X^j; 0 \le j \le J\}$ solves (4.1). Again, Theorem 5.2 asserts that there exist constants $C > 0$, independent of the discretization parameters $h, k > 0$ and $k_2 \equiv k_2(T, x) > 0$ such that for all $k \le k_2$ sufficiently small

$$\sup_{0 \le j \le J} \mathbb{E}\Big[\|X^j - Y^j\|_{\mathbb{L}^2}^2\Big] + k \sum_{j=0}^{J} \mathbb{E}\Big[\|\nabla(X^j - Y^j)\|_{\mathbb{L}^2}^2\Big] \le C\big(k + h^2\big). \tag{5.7}$$

Let $k_0 = \min\{k_1, k_2\}$. Then (5.6) and (5.7) hold true for all $k \le k_0$ sufficiently small. We combine (5.6) and (5.7) to conclude the proof of the main theorem.

References

[1] S.C. Brenner, L.R. Scott. The mathematical theory of finite element methods. *2nd edition, Springer* (2002).

[2] E. Carelli, A. Prohl. Rates of convergence for discretizations of the stochastic incompressible Navier-Stokes equations. *SIAM J. Numer. Anal.* **50**, pp. 2467-2496 (2012).

[3] J. Douglas, Jr. T. Dupont, L. Wahlbin. The stability in L^q of the L^2-projection into finite element function spaces. *Numer. Math.* 23, 193-197 (1975).

[4] B. Gess. Strong solutions for stochastic partial differential equations of gradient type. *J. Funct. Anal.* **263**, pp. 2355-2383 (2012).

[5] I. Gyöngy, A. Millet. Rate of convergence of implicit approximations for stochastic evolution equations. In: Baxendale, P., Lototsky, S. (eds.) Stochastic Differential Equations: Theory and Applications (A volume in honor of Boris L. Rosovskii), vol. **2**, pp. 281-310. World Scientific Interdisciplinary Mathematical Sciences. World Scientific, Singapore (2007).

[6] I. Gyöngy, A. Millet. Rate of convergence of space time approximations for stochastic evolution equations. *Potential Anal.* **30**, pp. 29–64 (2009).

[7] D.J. Higham, X. Mao, A.M. Stuart. Strong convergence of Euler-type methods for nonlinear stochastic differential equations. *SIAM J. Numer. Anal.* **40**, pp. 1041-1063 (2002).

[8] M. Kovács, S. Larsson, A. Mesforush. Finite element approximation of the Cahn-Hilliard-Cook equation. *SIAM J. Numer. Anal.* **49**, pp. 2407-2429 (2011).

[9] J. Printems. On the discretization in time of parabolic stochastic partial differential equations. M2AN **35**, pp. 1055-1078 (2001).

[10] 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).

[11] M. Sauer, W. Stannat. Lattice Approximation for Stochastic Reaction Diffusion Equations with One-Sided Lipschitz Condition. *Math. Comp.* 84(2015), no. 292, 743-766.

(Ananta K. Majee)

MATHEMATISCHES INSTITUT UNIVERSITÄT TÜBINGEN AUF DER MORGENSTELLE 10 D-72076 TÜBINGEN, GERMANY

E-mail address: `majee@na.uni-tuebingen.de`

(Andreas Prohl)

MATHEMATISCHES INSTITUT UNIVERSITÄT TÜBINGEN AUF DER MORGENSTELLE 10 D-72076 TÜBINGEN, GERMANY

E-mail address: `prohl@na.uni-tuebingen.de`