CONVERGENT NUMERICAL APPROXIMATION OF THE STOCHASTIC TOTAL VARIATION FLOW WITH LINEAR MULTIPLICATIVE NOISE: THE HIGHER DIMENSIONAL CASE

ĽUBOMÍR BAŇAS, MICHAEL RÖCKNER, AND ANDRÉ WILKE

ABSTRACT. We consider fully discrete finite element discretization of the stochastic total variation flow equation (STVF) with linear multiplicative noise which was previously proposed in [4]. Due to lack of a discrete counterpart of stronger a priori estimates in higher spatial dimensions the original convergence analysis of the numerical scheme was limited to one spatial dimension, cf. [5]. In this paper we generalize the convergence proof to higher dimensions.

1. INTRODUCTION

We study the convergence of numerical approximation of the stochastic total variation flow (STVF) equation

$$
\begin{aligned}
\mathrm{d}X &= \mathrm{div}\left(\frac{\nabla X}{|\nabla X|}\right)\mathrm{d}t - \lambda(X-g)\,\mathrm{d}t + X\,\mathrm{d}W, && \text{in } (0,T)\times\mathcal{O},\\
X &= 0 && \text{on } (0,T)\times\partial\mathcal{O},\\
X(0) &= x_0 && \text{in } \mathcal{O},
\end{aligned}
\tag{1}
$$

where $\mathcal{O} \in \mathbb{R}^d$, $d \geq 1$, is a bounded, convex polyhedral domain, $\lambda \geq 0$, $T > 0$ are constants and x_0, $g \in \mathbb{L}^2$. For simplicity we take W to be a one dimensional real-valued Wiener process.

We adopt the approach from [4] and construct a fully discrete approximation scheme (cf. (40) below) of (1) using a regularization approach. Given a regularization parameter $\varepsilon > 0$ we consider the following regularized problem

$$
\begin{aligned}
\mathrm{d}X^\varepsilon &= \mathrm{div}\left(\frac{\nabla X^\varepsilon}{\sqrt{|\nabla X^\varepsilon|^2+\varepsilon^2}}\right)\mathrm{d}t - \lambda(X^\varepsilon-g)\,\mathrm{d}t + X^\varepsilon\,\mathrm{d}W && \text{in } (0,T)\times\mathcal{O},\\
X^\varepsilon &= 0 && \text{on } (0,T)\times\partial\mathcal{O},\\
X^\varepsilon(0) &= x_0 && \text{in } \mathcal{O}\,.
\end{aligned}
\tag{2}
$$

Equations (1), (2), respectively, admit unique solutions in the sense of stochastic variational inequalities, see [2], [4], [3]. Throughout the paper we refer to the solutions of (1), (2) as SVI solutions, see Definition 3.1 below. The first numerical approximation of (1) was

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – SFB 1283/2 2021 – 317210226.

constructed in [4] and its convergence was shown by considering the full discretization of the regularized problem (2) as an intermediate step. The convergence proof of the numerical approximation in [4] relies on the discrete counterpart of a priori estimates in stronger norm (cf. Lemma 3.2 below), which are so-far restricted to spatial dimension $d = 1$, cf. [5]. The recent work [3] shows convergence of numerical approximation with a random walk representation of the noise to probabilistically weak SVI solutions of (1). The numerical analysis in [3] is valid in higher spatial dimensions $d \geq 1$, but does not cover the case of linear multiplicative noise, except for $d = 1$. In this work we show convergence of the numerical approximation of the stochastic total variation flow (1) with linear multiplicative noise in spatial dimension $d > 1$.

The paper is organized as follows. In Section 2 we introduce the notation and state some auxiliary results. The existence of a unique SVI solution of the regularized stochastic TV flow (2) and its convergence towards a unique SVI solution of (1) is discussed in Section 3. In Section 4 we introduce a time semi-discrete numerical scheme for the regularizared problem (3) below and show its convergence to the variational solution of (3) for initial data with higher regularity. Finally, in Section 5 we show the convergence of the fully discrete finite element scheme for the regularizared problem (2) and show its convergence to the SVI solution of (1).

2. Notation and preliminaries

Throughout the paper by C we denote a generic positive constant that may change from line to line. By $\mathbb{L}^p := L^p(\mathcal{O})$ for $1 \leq p \leq \infty$ we denote the standard spaces of p-th order integrable functions on $\mathcal{O}$; we use $\|\cdot\| = \|\cdot\|_{\mathbb{L}^2}$ for the $\mathbb{L}^2$-norm and $(\cdot,\cdot) = (\cdot,\cdot)_{\mathbb{L}^2}$ for the $\mathbb{L}^2$-inner product. For $k, p \in \mathbb{N}$ we denote the usual Sobolev space on $\mathcal{O}$ by $(\mathbb{W}^{p,k}, \|\cdot\|_{\mathbb{W}^{p,k}})$; for $p = 2$ we use $\mathbb{H}^k := \mathbb{W}^{2,k}$. Furthermore $\mathbb{H}^1_0$ stands for the $\mathbb{H}^1$ space with zero trace on $\partial\mathcal{O}$ with its dual denoted as $\mathbb{H}^{-1}$ and we set $\langle\cdot,\cdot\rangle = \langle\cdot,\cdot\rangle_{\mathbb{H}^{-1}\times\mathbb{H}^1_0}$, where $\langle\cdot,\cdot\rangle_{\mathbb{H}^{-1}\times\mathbb{H}^1_0}$ is the duality pairing between $\mathbb{H}^1_0$ and $\mathbb{H}^{-1}$.

For $u \in \mathbb{H}^1_0$ we consider the energy functional

$$\mathcal{J}_{\varepsilon,\lambda}(u) = \int_{\mathcal{O}} \sqrt{|\nabla u|^2 + \varepsilon^2}\,\mathrm{d}\mathbf{x} + \frac{\lambda}{2}\int_{\mathcal{O}} |u - g|^2\,\mathrm{d}\mathbf{x}\,,$$

With a slight abuse of notation we set $\mathcal{J}_\varepsilon := \mathcal{J}_{\varepsilon,0}$ if $\lambda = 0$ and $\mathcal{J}_\lambda := \mathcal{J}_{0,\lambda}$ if $\varepsilon = 0$.

Next, we state basic definitions related to the functions of bounded variation.

Definition 2.1. *A function $u \in L^1(\mathcal{O})$ is called a function of bounded variation, if its total variation*

$$\int_{\mathcal{O}} |\nabla u|\,\mathrm{d}\mathbf{x} := \sup\left\{-\int_{\mathcal{O}} u \operatorname{div}\mathbf{v}\,\mathrm{d}\mathbf{x};\ \mathbf{v} \in C_0^\infty(\mathcal{O},\mathbb{R}^d),\ \|\mathbf{v}\|_{L^\infty} \leq 1\right\},$$

is finite. The space of functions of bounded variation is denoted by $BV(\mathcal{O})$.

Furthermore, for $u \in BV(\mathcal{O})$ we set

$$\int_{\mathcal{O}} \sqrt{|\nabla u|^2 + \varepsilon^2}\,\mathrm{d}\mathbf{x} := \sup\left\{\int_{\mathcal{O}} \Big(-u \operatorname{div}\mathbf{v} + \varepsilon\sqrt{1 - |\mathbf{v}|^2}\Big)\,\mathrm{d}\mathbf{x};\ \mathbf{v} \in C_0^\infty(\mathcal{O},\mathbb{R}^d),\ \|\mathbf{v}\|_{L^\infty} \leq 1\right\}.$$

3. The continuous problem

In this section we construct a unique SVI solution of (1) (see Definition 3.1 below) via a two-level regularization procedure. Given the data $x_0 \in L^2(\Omega, \mathcal{F}_0; \mathbb{L}^2)$, $g \in \mathbb{L}^2$ we consider an $\mathbb{H}^1_0$-approximating sequences $\{x_0^n\}_{n\in\mathbb{N}} \subset L^2(\Omega, \mathcal{F}_0; \mathbb{H}^1_0)$, $\{g^n\}_{n\in\mathbb{N}} \subset \mathbb{H}^1_0$, s.t. $x_0^n \to x_0$, $g^n \to g$ in $L^2(\Omega, \mathcal{F}_0; \mathbb{L}^2)$ for $n \to \infty$, respectively. For $\delta > 0$ we introduce a regularization of (2) as

$$\begin{aligned} \mathrm{d}X_n^\delta =& \delta\Delta X_n^\delta + \mathrm{div}\left(\frac{\nabla X_n^\delta}{\sqrt{|\nabla X_n^\delta|^2+\varepsilon^2}}\right)\mathrm{d}t - \lambda(X_n^\delta - g_n)\,\mathrm{d}t + X_n^\delta\,\mathrm{d}W(t), \\ X_n^\delta(0) =& x_0^n\,. \end{aligned} \tag{3}$$

We define the operator $A^\delta : \mathbb{H}^1_0 \to \mathbb{H}^{-1}$ as

$$\langle A^\delta u, v\rangle_{\mathbb{H}^{-1}\times\mathbb{H}^1_0} = \delta\,(\nabla u, \nabla v) + \left(\frac{\nabla u}{\sqrt{|\nabla u|^2+\varepsilon^2}}, \nabla v\right) + \lambda\,(u - g_n, v) \quad \forall u, v \in \mathbb{H}^1_0, \tag{4}$$

and note that (3) can be equivalently formulated as

$$\begin{aligned} \mathrm{d}X_n^\delta + A^\delta X_n^\delta\,\mathrm{d}t &= X_n^\delta\,\mathrm{d}W(t)\,, \\ X_n^\delta(0) &= x_0^n. \end{aligned} \tag{5}$$

The operator $A^\delta : \mathbb{H}^1_0 \to \mathbb{H}^{-1}$ is demicontinuos and satisfies (cf. [8, Remark 4.1.1])

$$\langle A^\delta(u) - A^\delta(v), u - v\rangle_{\mathbb{H}^{-1}\times\mathbb{H}^1_0} \geq \delta\|u-v\|^2_{\mathbb{H}^1_0} + \lambda\|u-v\|^2, \qquad \forall u, v \in \mathbb{H}^1_0, \tag{6}$$

$$\|A^\delta(u)\|_{\mathbb{H}^{-1}} \leq C(\delta, \lambda, \|g_n\|)(\|u\|_{\mathbb{H}^1_0} + 1), \qquad \forall u \in \mathbb{H}^1_0. \tag{7}$$

We recall that the convexity of the function $\sqrt{|\cdot|^2+\varepsilon^2}$ implies the monotonicity property

$$\begin{aligned} &\left(\frac{\nabla X}{\sqrt{|\nabla X|^2+\varepsilon^2}} - \frac{\nabla Y}{\sqrt{|\nabla Y|^2+\varepsilon^2}}, \nabla(X-Y)\right) \\ &\quad = \left(\frac{\nabla X}{\sqrt{|\nabla X|^2+\varepsilon^2}}, \nabla(X-Y)\right) + \left(\frac{\nabla Y}{\sqrt{|\nabla Y|^2+\varepsilon^2}}, \nabla(Y-X)\right) \\ &\quad \geq \mathcal{J}_\varepsilon(X) - \mathcal{J}_\varepsilon(Y) + \mathcal{J}_\varepsilon(Y) - \mathcal{J}_\varepsilon(X) = 0. \end{aligned} \tag{8}$$

The well-posedness of the regularized problem (3) follows from standard theory of monotone SPDEs, see for instance [8, Chapter 4] and [4].

Lemma 3.1. *For any $\varepsilon, \delta > 0$ and $x_0^n \in L^2(\Omega, \mathcal{F}_0; \mathbb{H}^1_0)$, $g_n \in \mathbb{H}^1_0$ there exists a unique variational solution $X_n^\delta \in L^2(\Omega; C([0,T]; \mathbb{L}^2))$ of (3). Furthermore, there exists a constant $C \equiv C(T) > 0$ such that the following estimate holds*

$$\mathbb{E}\left[\sup_{t\in[0,T]}\|X_n^\delta(t)\|^2\right] \leq C(\mathbb{E}\left[\|x_0^n\|^2\right] + \|g_n\|^2).$$

We recall that in addition to the above $\mathbb{L}^2$-estimate, the solution of the regularized equation (3) satisfies the following stronger a priori estimate, see [4, Lemma 3.2].

Lemma 3.2. *Let $x_0^n \in L^2(\Omega,\mathcal{F}_0;\mathbb{H}_0^1)$, $g_n \in \mathbb{H}_0^1$. There exists a constant $C \equiv C(T)$ such that for any $\varepsilon,\delta > 0$ the corresponding variational solution X_n^δ of (3) satisfies*

$$\mathbb{E}\left[\sup_{t\in[0,T]} \|\nabla X_n^\delta(t)\|^2 + \delta\int_0^T \|\Delta X_n^\delta(t)\|^2\,\mathrm{d}t\right] \le C\left(\mathbb{E}\left[\|x_0^n\|_{\mathbb{H}_0^1}^2\right] + \|g_n\|_{\mathbb{H}_0^1}^2\right). \tag{9}$$

We consider the following functionals

$$\bar{\mathcal{J}}_{\varepsilon,\lambda}(u) = \begin{cases} \mathcal{J}_{\varepsilon,\lambda}(u) + \int_{\partial\mathcal{O}} |\gamma_0(u)|\,\mathrm{d}\mathcal{H}^{n-1} & \text{for } u \in BV(\mathcal{O})\cap L^2(\mathcal{O}), \\ +\infty & \text{for } u \in BV(\mathcal{O})\setminus L^2(\mathcal{O}), \end{cases}$$

and (for $\varepsilon = 0$)

$$\bar{\mathcal{J}}_{\lambda}(u) = \begin{cases} \mathcal{J}_{\lambda}(u) + \int_{\partial\mathcal{O}} |\gamma_0(u)|\,\mathrm{d}\mathcal{H}^{n-1} & \text{for } u \in BV(\mathcal{O})\cap L^2(\mathcal{O}), \\ +\infty & \text{for } u \in BV(\mathcal{O})\setminus L^2(\mathcal{O}), \end{cases}$$

where $\gamma_0(u)$ is the trace of u on the boundary and $\mathrm{d}\mathcal{H}^{n-1}$ is the Hausdorff measure. The functionals $\bar{\mathcal{J}}_{\varepsilon,\lambda}$ and $\bar{\mathcal{J}}_{\lambda}$ are both convex and lower semicontinuous on $\mathbb{L}^2$ and the lower semicontinuous hulls of $\bar{\mathcal{J}}_{\varepsilon,\lambda}|_{\mathbb{H}_0^1}$ and $\bar{\mathcal{J}}_{\lambda}|_{\mathbb{H}_0^1}$, respectively, cf. [1, Proposition 11.3.2].

As in [4] we interpret (1), (2) as stochastic variational inequalities.

Definition 3.1. *Let $0 < T < \infty$, $\varepsilon \in [0,1]$ and $x_0 \in L^2(\Omega,\mathcal{F}_0;\mathbb{L}^2)$ and $g \in \mathbb{L}^2$. Then an $\mathcal{F}_t$-adapted stochastic process $X^\varepsilon \in L^2(\Omega;C([0,T];\mathbb{L}^2))\cap L^1(\Omega;L^1((0,T);BV(\mathcal{O})))$ (denoted by $X \in L^2(\Omega;C([0,T];\mathbb{L}^2))\cap L^1(\Omega;L^1((0,T);BV(\mathcal{O})))$ for $\varepsilon = 0$) is called a SVI solution of (2) (or (1) if $\varepsilon = 0$) if $X^\varepsilon(0) = x_0$ ($X(0) = x_0$), and for each $(\mathcal{F}_t)$-progressively measurable process $G \in L^2(\Omega\times(0,T),\mathbb{L}^2)$ and for each $(\mathcal{F}_t)$-adapted $\mathbb{L}^2$-valued process Z with $\mathbb{P}$-a.s. continuous sample paths, s.t, $Z \in L^2(\Omega\times(0,T);\mathbb{H}_0^1)$, which satisfy the equation*

$$\mathrm{d}Z(t) = -G(t)\,\mathrm{d}t + Z(t)\,\mathrm{d}W(t),\ \ t\in[0,T], \tag{10}$$

it holds for $\varepsilon \in (0,1]$ that

$$\begin{aligned} &\frac{1}{2}\mathbb{E}\left[\|X^\varepsilon(t) - Z(t)\|^2\right] + \mathbb{E}\left[\int_0^t \bar{\mathcal{J}}_{\varepsilon,\lambda}(X^\varepsilon(s))\,\mathrm{d}s\right] \\ &\quad\le \frac{1}{2}\mathbb{E}\left[\|x_0 - Z(0)\|^2\right] + \mathbb{E}\left[\int_0^t \bar{\mathcal{J}}_{\varepsilon,\lambda}(Z(s))\,\mathrm{d}s\right] \\ &\quad + \frac{1}{2}\mathbb{E}\left[\int_0^t \|X^\varepsilon(s) - Z(s)\|^2\,\mathrm{d}s\right] + \frac{1}{2}\mathbb{E}\left[\int_0^t (X^\varepsilon(s) - Z(s), G)\,\mathrm{d}s\right], \end{aligned} \tag{11}$$

and analogously for $\varepsilon = 0$ it holds that

$$
\begin{aligned}
&\frac{1}{2}\mathbb{E}\left[\|X(t)-Z(t)\|^2\right] + \mathbb{E}\left[\int_0^t \bar{\mathcal{J}}_\lambda(X(s))\,\mathrm{d}s\right] \\
&\leq \frac{1}{2}\mathbb{E}\left[\|x_0-Z(0)\|^2\right] + \mathbb{E}\left[\int_0^t \bar{\mathcal{J}}_\lambda(Z(s))\,\mathrm{d}s\right] \\
&+ \frac{1}{2}\mathbb{E}\left[\int_0^t \|X(s)-Z(s)\|^2\,\mathrm{d}s\right] + \frac{1}{2}\mathbb{E}\left[\int_0^t (X(s)-Z(s),G)\,\mathrm{d}s\right].
\end{aligned} \tag{12}
$$

The next theorem shows that the solutions of the regularized problem (3) converge to the SVI solution of (1) for $\varepsilon, n \to \infty, \delta \to 0$; the proof of the theorem follows as [4, Theorem 3.2].

Theorem 3.1. *Let $0 < T < \infty$ and $x_0 \in L^2(\Omega, \mathcal{F}_0; \mathbb{L}^2)$, $g \in \mathbb{L}^2$ be fixed and consider $\mathbb{H}^1_0$-approximating sequences $\{x_0^n\}_{n\in\mathbb{N}} \subset L^2(\Omega, \mathcal{F}_0; \mathbb{H}^1_0)$, $\{g^n\}_{n\in\mathbb{N}} \subset \mathbb{H}^1_0$, s.t. $x_0^n \to x_0$, $g^n \to g$ in $L^2(\Omega, \mathcal{F}_0; \mathbb{L}^2)$ for $n \to \infty$. Let $\{X_n^\delta\}_{\delta>0}$ be the variational solutions of (3) associated with x_0^n, g^n, $\varepsilon \in (0,1]$ and $\delta > 0$. Then X_n^δ converges to the unique SVI variational solution X of (1) in $L^2(\Omega; C([0,T]; \mathbb{L}^2))$ for $\varepsilon \to 0, n \to \infty, \delta \to 0$, i.e.,*

$$
\lim_{\varepsilon\to 0}\lim_{n\to\infty}\lim_{\delta\to 0}\mathbb{E}\left[\sup_{t\in[0,T]}\|X_n^\delta(t)-X(t)\|^2\right] = 0. \tag{13}
$$

4. Semi-discretization in time

For $N \in \mathbb{N}$ we consider a partition of the time interval $t_i = i\tau$ for $i = 0, \ldots, N$ with the time-step $\tau = T/N$, and denote the discrete Wiener increments as $\Delta_i W = W(t_i) - W(t_{i-1})$.

The implicit time-discrete approximation of (3) is defined as follows: set $X^0_{\delta,n} = x_0^n$ and determine $X^i_{\delta,n} \in \mathbb{H}^1_0$, $i = 1, \ldots, N$ as the solution of

$$
\begin{aligned}
\left(X^i_{\delta,n}, \Phi\right) = \left(X^{i-1}_{\delta,n}, \Phi\right) - \tau\delta\left(\nabla X^i_{\delta,n}, \nabla\Phi\right) - \tau\left(\frac{\nabla X^i_{\delta,n}}{\sqrt{|\nabla X^i_{\delta,n}|^2 + \varepsilon^2}}, \nabla\Phi\right) \\
- \tau\lambda\left(X^i_{\delta,n} - g_n, \Phi\right) + \left(X^{i-1}_{\delta,n}, \Phi\right)\Delta_i W \qquad \forall \Phi \in \mathbb{H}^1_0.
\end{aligned} \tag{14}
$$

The existence, uniqueness and measurability of $\{X^i_{\delta,n}\}_{i=1}^N$ can be shown via finite dimensional Galerkin approximation; we summarize the main steps below:

- consider a finite dimensional subspace $\mathbb{V}_m$ and the corresponding Galerkin approximation $X^i_{\delta,n,m} \in \mathbb{V}_m$ of the solution $X^i_{\delta,n}$ of (14);
- proceed by induction: assuming that an $\mathcal{F}_{t_{i-1}}$-measurable solution $X^{i-1}_{\delta,n,m} \in \mathbb{V}_m$ exists, the existence of an $\mathcal{F}_{t_i}$-measurable solution $X^i_{\delta,n,m}$ follows by Brouwer's fixed point theorem and the uniqueness by the monotonicity property (8), cf. [4, Lemma 4.3];
- for any $m \in \mathbb{N}$ the Galerkin approximation $\{X^i_{\delta,n,m}\}_{i=1}^n$ satisfies the same a priori estimates as in Lemma 4.1 below;

- by the (uniform in m) a priori estimates it holds that $X^i_{\delta,n,m} \rightharpoonup X^i_{\delta,n}$ for $m \to \infty$. Furthermore, by the monotonicity (8) it follows that the limit $X^i_{\delta,n}$ is unique and satisfies (14), cf., Lemma 4.2 below.

In the next lemma we state the stability properties of the time-discrete solution of the scheme (14) which are discrete analogues of estimates in Lemma 3.1 and Lemma 3.2. Later on, we will consider sequences $\{x_0^n\}_{n\in\mathbb{N}}$, $\{g_n\}_{n\in\mathbb{N}}$ which are uniformly bounded in $\mathbb{L}^2$ but not in $\mathbb{H}^1_0$. Hence, in the following we suppress the dependence of the constants on the data in (15) but not in (16).

Lemma 4.1. *Let $x_0^n \in L^2(\Omega, \mathcal{F}_0; \mathbb{H}^1_0)$ and $g_n \in \mathbb{H}^1_0$ be given. Then there exists a constant $C \equiv C(\mathbb{E}\left[\|x_0^n\|_{\mathbb{L}^2}\right], \|g_n\|_{\mathbb{L}^2}) > 0$ such that for any $\tau > 0$ the solution of scheme (14) satisfies*

$$\max_{i=1,\dots,N} \mathbb{E}\left[\|X^i_{\delta,n}\|^2\right] + \frac{1}{4}\mathbb{E}\left[\sum_{k=1}^N \|X^k_{\delta,n} - X^{k-1}_{\delta,n}\|^2\right] + \tau\mathbb{E}\left[\sum_{k=1}^N \mathcal{J}_\varepsilon(X^k_{\delta,n})\right] + \frac{\tau\lambda}{2}\mathbb{E}\left[\sum_{k=1}^N \|X^k_{\delta,n}\|^2\right] \leq C\,, \tag{15}$$

and a constant $C_n \equiv C(\mathbb{E}[\|x_0^n\|_{\mathbb{H}^1_0}], \|g_n\|_{\mathbb{H}^1_0}) > 0$ such that for any $\tau > 0$

$$\max_{i=1,\dots,N} \mathbb{E}\left[\|\nabla X^i_{\delta,n}\|^2\right] + \mathbb{E}\left[\sum_{k=1}^N \|\nabla(X^k_{\delta,n} - X^{k-1}_{\delta,n})\|^2\right] + \tau\delta\mathbb{E}\left[\sum_{k=1}^N \|\Delta X^k_{\delta,n}\|^2\right] \leq C_n. \tag{16}$$

Proof. We set $\Phi = X^i_{\delta,n}$ (14) and use the identity $2(a-b)a = a^2 - b^2 + (a-b)^2$ to get for $i = 1, \dots, N$

$$\frac{1}{2}\|X^i_{\delta,n}\|^2 + \frac{1}{2}\|X^i_{\delta,n} - X^{i-1}_{\delta,n}\|^2 + \tau\delta\|\nabla X^i_{\delta,n}\|^2 + \tau\left(\frac{\nabla X^i_{\delta,n}}{\sqrt{|\nabla X^i_{\delta,n}|^2 + \varepsilon^2}}, \nabla X^i_{\delta,n}\right) = \frac{1}{2}\|X^{i-1}_{\delta,n}\|^2 - \tau\lambda\left(\|X^i_{\delta,n}\|^2 - \left(g^n, X^i_{\delta,n}\right)\right) + \left(X^{i-1}_{\delta,n}, X^i_{\delta,n}\right)\Delta_i W. \tag{17}$$

We take expectation in (17) and use the properties of the Wiener increments $\mathbb{E}\left[\Delta_i W\right] = 0$, $\mathbb{E}\left[|\Delta_i W|^2\right] = \tau$ and the independence of $\Delta_i W$ and $X^{i-1}_{\delta,n}$ to estimate the stochastic term as

$$\begin{aligned}
\mathbb{E}\left[\left(X^{i-1}_{\delta,n}, X^i_{\delta,n}\right)\Delta_i W\right] &= \mathbb{E}\left[\left(X^{i-1}_{\delta,n}, X^i_{\delta,n} - X^{i-1}_{\delta,n}\right)\Delta_i W\right] + \mathbb{E}\left[\left(X^{i-1}_{\delta,n}, X^{i-1}_{\delta,n}\right)\Delta_i W\right] \\
&\leq \mathbb{E}\left[\frac{1}{4}\|X^{i-1}_{\delta,n} - X^i_{\delta,n}\|^2 + \|X^{i-1}_{\delta,n}\|^2|\Delta_i W|^2\right] + \mathbb{E}\left[\|X^{i-1}_{\delta,n}\|^2\right]\mathbb{E}\left[\Delta_i W\right] \\
&= \frac{1}{4}\mathbb{E}\left[\|X^i_{\delta,n} - X^{i-1}_{\delta,n}\|^2\right] + \tau\mathbb{E}\left[\|X^{i-1}_{\delta,n}\|^2\right].
\end{aligned}$$

From (17) by the convexity of $\mathcal{J}_\varepsilon$ and using $\mathcal{J}_\varepsilon(0) = \varepsilon|\mathcal{O}|$ it follows that

$$\begin{aligned}\frac{1}{2}\mathbb{E}\left[\|X^i_{\delta,n}\|^2\right] + \frac{1}{4}\mathbb{E}\left[\|X^i_{\delta,n} - X^{i-1}_{\delta,n}\|^2\right] + \frac{\tau\lambda}{2}\mathbb{E}\left[\|X^i_{\delta,n}\|^2\right] + \tau\delta\|\nabla X^i_{\delta,n}\|^2 + \tau\mathbb{E}\left[\mathcal{J}_\varepsilon(X^i_{\delta,n})\right] \\ \leq \tau\varepsilon|\mathcal{O}| + \frac{1}{2}\mathbb{E}\left[\|X^{i-1}_{\delta,n}\|^2\right] + \tau\mathbb{E}\left[\|X^{i-1}_{\delta,n}\|^2\right] + \tau\lambda\|g^n\|^2 .\end{aligned}$$

We sum up the above inequality for $k = 1, \ldots, i$ and obtain

$$\begin{aligned}&\frac{1}{2}\mathbb{E}\left[\|X^i_{\delta,n}\|^2\right] + \frac{1}{4}\mathbb{E}\left[\sum_{k=1}^{i}\|X^k_{\delta,n} - X^{k-1}_{\delta,n}\|^2\right] + \frac{\tau\lambda}{2}\mathbb{E}\left[\sum_{k=1}^{i}\|X^k_{\delta,n}\|^2\right] \\ &+\tau\delta\mathbb{E}\left[\sum_{k=1}^{i}\|\nabla X^k_{\delta,n}\|\right] + \tau\mathbb{E}\left[\sum_{k=1}^{i}\mathcal{J}_\varepsilon(X^k_{\delta,n})\right] \\ &\leq T\varepsilon|\mathcal{O}| + \frac{1}{2}\mathbb{E}\left[\|x^n_0\|^2\right] + T\lambda\|g^n\|^2 + \tau\mathbb{E}\left[\sum_{k=1}^{i}\|X^{k-1}_{\delta,n}\|^2\right] .\end{aligned} \tag{18}$$

Then (15) follows from (18) after an application of the discrete Gronwall lemma.

To show the estimate (16) we proceed formally, the calculations can be made rigorous via finite dimensional Galerkin approximation, cf. [4, Lemma 3.2]. We set $\Phi = -\Delta X^i_{\delta,n}$ in (14), use integration by parts and proceed analogously to the first part of the proof.

As in the proof of [4, Lemma 3.2] we deduce that

$$\left(\frac{\nabla X^i_{\delta,n}}{\sqrt{|\nabla X^i_{\delta,n}|^2 + \varepsilon^2}}, \nabla(-\Delta X^i_{\delta,n})\right) \geq 0. \tag{19}$$

Hence, we neglect the above term and conclude that

$$\begin{aligned}&\frac{1}{2}\mathbb{E}\left[\|\nabla X^i_{\delta,n}\|^2\right] + \frac{1}{4}\mathbb{E}\left[\sum_{k=1}^{i}\|\nabla(X^k_{\varepsilon,n} - X^{k-1}_{\varepsilon,n})\|^2\right] + \frac{\tau\lambda}{2}\mathbb{E}\left[\sum_{k=1}^{i}\|\nabla X^k_{\varepsilon,n}\|^2\right] \\ &+ \tau\delta\mathbb{E}\left[\sum_{k=1}^{i}\|\Delta X^k_{\delta,n}\|\right] \leq \frac{1}{2}\mathbb{E}\left[\|\nabla x^n_0\|^2\right] + T\lambda\|\nabla g_n\|^2 + \tau\mathbb{E}\left[\sum_{k=1}^{i}\|\nabla X^{k-1}_{\varepsilon,n}\|^2\right] .\end{aligned}$$

Estimate (16) then follows after an application of the discrete Gronwall lemma. □

Remark 4.1. *The proof of the convergence of the numerical approximation given in* [4] *relies on the stronger a priori estimate* (16)*. The above proof of the estimate* (16) *requires property* (19) *to hold. So far, the proof of the spatially discrete counterpart of the estimate* (19) *is restricted to spatial dimension* $d = 1$ [5, Lemma 3.1]*. In the proof of the convergence of the fully discrete numerical approximation below we circumvent the lack of a (rigorous) discrete counterpart of* (19) *for* $d > 1$ *by considering the time-discrete problem (14) as an intermediate step.*

We define piecewise constant time-interpolants of the numerical solution $\{X^i_{\delta,n}\}_{i=0}^N$ of (14) for $t \in [0,T]$ as

$$\overline{X}^{\delta,n}_{\tau}(t) = X^i_{\delta,n} \quad \text{if} \quad t \in (t_{i-1}, t_i] \tag{20}$$

and

$$\overline{X}^{\delta,n}_{\tau_-}(t) = X^{i-1}_{\delta,n} \quad \text{if} \quad t \in [t_{i-1}, t_i)\,. \tag{21}$$

We note that (14) can be reformulated as

$$\begin{aligned} &\Big(\overline{X}^{\delta,n}_{\tau}(t), \Phi\Big) + \left\langle \int_0^{\theta_+(t)} A^\delta \overline{X}^{\delta,n}_{\tau}(s)\,\mathrm{d}s, \Phi \right\rangle \\ &= \big(X^0_{\varepsilon,n}, \Phi\big) + \left(\int_0^{\theta_+(t)} \overline{X}^{\delta,n}_{\tau_-}(s)\,\mathrm{d}W(s), \Phi\right) \qquad \text{for } t \in [0,T],\ \Phi \in \mathbb{H}^1_0, \end{aligned} \tag{22}$$

where $\theta_+(0) = 0$ and $\theta_+(t) = t_i$ if $t \in (t_{i-1}, t_i]$.

Estimates (15), (16) imply the bounds

$$\begin{aligned} &\sup_{t\in[0,T]} \mathbb{E}\left[\|\overline{X}^{\delta,n}_{\tau}(t)\|^2\right] \le C, \qquad \sup_{t\in[0,T]} \mathbb{E}\left[\|\overline{X}^{\delta,n}_{\tau_-}(t)\|^2\right] \le C, \\ &\delta\mathbb{E}\left[\int_0^T \|\nabla \overline{X}^{\delta,n}_{\tau}(s)\|^2\,\mathrm{d}s\right] \le C. \end{aligned} \tag{23}$$

Furthermore, (23) and (7) imply

$$\mathbb{E}\left[\int_0^T \|A^\delta \overline{X}^{\delta,n}_{\tau}(s)\|^2_{\mathbb{H}^{-1}}\,\mathrm{d}s\right] \le C. \tag{24}$$

The estimates in (23) for fixed $n \in \mathbb{N}$, $\varepsilon, \delta > 0$ imply the existence of a subsequence, still denoted by $\{\overline{X}^{\delta,n}_{\tau}\}_{\tau>0}$, and a $Y \in L^2(\Omega\times(0,T);\mathbb{L}^2)\cap L^2(\Omega\times(0,T);\mathbb{H}^1_0)\cap L^\infty((0,T);L^2(\Omega;\mathbb{L}^2)$, s.t., for $\tau \to 0$

$$\begin{aligned} \overline{X}^{\delta,n}_{\tau} &\rightharpoonup Y \text{ in } L^2(\Omega\times(0,T);\mathbb{L}^2), \\ \overline{X}^{\delta,n}_{\tau} &\rightharpoonup Y \text{ in } L^2(\Omega\times(0,T);\mathbb{H}^1_0), \\ \overline{X}^{\delta,n}_{\tau} &\rightharpoonup^* Y \text{ in } L^\infty((0,T);L^2(\Omega;\mathbb{L}^2)). \end{aligned} \tag{25}$$

In addition, there exists $\nu \in L^2(\Omega;\mathbb{L}^2)$ such that $\overline{X}^{\delta,n}_{\tau}(T) \rightharpoonup \nu$ in $L^2(\Omega;\mathbb{L}^2)$ as $\tau \to 0$ and estimate (24) implies the existence of $a^\delta \in L^2(\Omega\times(0,T);\mathbb{H}^{-1})$, s.t.,

$$A^\delta \overline{X}^{\delta,n}_{\tau} \rightharpoonup a^\delta \text{ in } L^2(\Omega\times(0,T);\mathbb{H}^{-1}) \quad \text{for } \tau \to 0. \tag{26}$$

Furthermore, the estimates in (23) for fixed $n \in \mathbb{N}$, $\varepsilon, \delta > 0$ imply the existence of a subsequence, still denoted by $\{\overline{X}^{\delta,n}_{\tau_-}\}_{\tau>0}$, and of $Y^- \in L^2(\Omega\times(0,T);\mathbb{L}^2)$, s.t.,

$$\overline{X}^{\delta,n}_{\tau_-} \rightharpoonup Y^- \text{ in } L^2(\Omega\times(0,T);\mathbb{L}^2) \quad \text{for } \tau \to 0.$$

Finally, inequality (18) implies

$$\begin{aligned}\lim_{\tau\to 0}\mathbb{E}\left[\int_0^T \|\overline{X}_\tau^{\delta,n}(s)-\overline{X}_{\tau_-}^{\delta,n}(s)\|^2\,\mathrm{d}s\right] &= \lim_{\tau\to 0}\tau\mathbb{E}\left[\sum_{k=1}^N \|X_{\delta,n}^k - X_{\delta,n}^{k-1}\|^2\right]\\ &\leq \lim_{\tau\to 0} C\tau = 0\,.\end{aligned}$$

which shows that the weak limits of Y and Y^- coincide.

From the above convergence properties we deduce by standard arguments, cf. [4, Lemma 4.6], that the solutions of the semi-discrete scheme (14) converge to the unique variational solution of (3) for $\tau \to 0$.

Lemma 4.2. *Let $x_0^n \in L^2(\Omega,\mathcal{F}_0;\mathbb{H}_0^1)$ and $g_n \in \mathbb{H}_0^1$ be given, let $\varepsilon,\delta,\lambda > 0$, $n \in \mathbb{N}$ be fixed. Further, let X_n^δ be the unique variational solution of* (3) *and $\overline{X}_\tau^{\delta,n}$, $\overline{X}_{\tau_-}^{\delta,n}$ be the respective time-interpolant (20), (21) of the numerical solution $\{X_{\delta,n}^i\}_{i=1}^N$ of* (14)*. Then $\overline{X}_\tau^{\delta,n}$, $\overline{X}_{\tau_-}^{\delta,n}$ converge to X_n^δ for $\tau \to 0$ in the sense that the weak limits from (25), (26) satisfy $Y \equiv X_n^\delta$, $a^\delta \equiv A^\delta Y \equiv A^\delta X_n^\delta$ and $\nu = Y(T) \equiv X_n^\delta(T)$. In addition, it holds for almost all $(\omega,t) \in \Omega \times (0,T)$ that*

$$Y(t) = Y(0) - \int_0^t A^\delta Y(s)\,\mathrm{d}s + \int_0^t Y(s)\,\mathrm{d}W(s),$$

and there is an $\mathbb{L}^2$-valued continuous modification of Y (denoted again by Y) such that for all $t \in [0,T]$

$$\begin{aligned}\frac{1}{2}\|Y(t)\|^2 = &\frac{1}{2}\|Y(0)\|^2 - \int_0^t \langle A^\delta Y(s), Y(s)\rangle + \frac{1}{2}\|Y(s)\|^2\,\mathrm{d}s\\ &+ \int_0^t (Y(s),Y(s))\,\mathrm{d}W(s).\end{aligned} \tag{27}$$

The strong monotonicity property (6) of the operator A^δ implies strong convergence of the time-discrete approximation in $L^2(\Omega \times (0,T);\mathbb{L}^2)$, cf. [4, Lemma 4.7].

Lemma 4.3. *Let $x_0^n \in L^2(\Omega,\mathcal{F}_0;\mathbb{H}_0^1)$ and $g_n \in \mathbb{H}_0^1$ be given, let $\varepsilon,\delta,\lambda > 0$, $n \in \mathbb{N}$ be fixed. Furthermore, let X_n^δ be the variational solution of* (3) *and $\overline{X}_\tau^{\delta,n}$ be the time-interpolants (20) of the time-discrete solution $\{X_{\delta,n}^i\}_{i=1}^N$ of* (14)*. Then*

$$\lim_{\tau\to 0}\|X_n^\delta - \overline{X}_\tau^{\delta,n}\|^2_{L^2(\Omega\times(0,T);\mathbb{L}^2)} \to 0. \tag{28}$$

5. Full Discretization

Given a quasi-uniform triangulation $\mathcal{T}_h$ of $\mathcal{O}$ we consider the $\mathbb{H}_0^1$-conforming finite element space of globally continuous piecewise linear functions over $\mathcal{T}_h$ given as

$$\mathbb{V}_h = \left\{w_h \in C^0(\mathcal{O}) : w_h|_T \in \mathcal{P}_1(T)\ \forall T \in \mathcal{T}_h\right\} \subset \mathbb{H}_0^1.$$

The orthogonal $\mathbb{L}^2$-projection $\Pi_h : \mathbb{H}^1 \to \mathbb{V}_h$ is defined as

$$(v - \Pi_h v, w_h) = 0 \qquad \forall w_h \in \mathbb{V}_h \,.$$

It is well-known, see e.g., [6], [7], that the projection operator satisfies the following interpolation and stability properties for $\psi \in \mathbb{H}^1$:

$$\|\psi - \Pi_h \psi\| \leq Ch\|\nabla\psi\| \quad \text{and} \quad \|\Pi_h\psi\|_{\mathbb{H}^1} \leq C\|\psi\|_{\mathbb{H}^1} \,. \tag{29}$$

For $\psi \in \mathbb{H}^2$ one has the following estimate

$$\|\psi - \Pi_h\psi\| + h\|\nabla[\psi - \Pi_h\psi]\| \leq Ch^2\|\nabla^2\psi\|. \tag{30}$$

By the estimate (16) we deduce from (30) that

$$\tau \sum_{i=1}^{N} \mathbb{E}\left[\|\nabla\big(X^i_{\delta,n} - \Pi_h X^i_{\delta,n}\big)\|^2\right] \leq C_n \delta^{-1} h^2 \,, \tag{31}$$

uniformly for all $\tau > 0$.

Given $\mathbb{H}^1$-regular data x_0^n, g_n we consider the following auxiliary fully discrete numerical scheme. Set $X^0_{\varepsilon.n,h} = \Pi_h x_0^n$, $g_{n,h} = \Pi_h g_n$, $\tau = T/N$ and determine $X^i_{\varepsilon,n,h} \in \mathbb{V}_h$, $i = 1, \ldots, N$ as the solution of

$$\begin{aligned} \big(X^i_{\varepsilon,n,h}, \Phi_h\big) = \big(X^{i-1}_{\varepsilon,n,h}, \Phi_h\big) - \tau \left(\frac{\nabla X^i_{\varepsilon,n,h}}{\sqrt{|\nabla X^i_{\varepsilon,n,h}|^2 + \varepsilon^2}}, \nabla\Phi_h\right) \\ - \tau\lambda\left(X^i_{\varepsilon,n,h} - g_{n,h}, \Phi_h\right) + \left(X^{i-1}_{\varepsilon,n,h}, \Phi_h\right)\Delta_i W \qquad \forall \Phi_h \in \mathbb{V}_h \;. \end{aligned} \tag{32}$$

The existence, uniqueness and measurability of the numerical solution $\{X^i_{\varepsilon,n,h}\}_{i=1}^N$ follows as in [4, Lemma 5.3].

In the next lemma we state the stability properties of the auxiliary numerical scheme (32). The proof of the estimate is a direct counterpart of the proof of (15) and is therefore omitted.

Lemma 5.1. *Let $x_0^n, g_n \in \mathbb{H}^1_0$ and $T > 0$. Then there exists a constant $C \equiv C(T)$ such that the solutions of scheme* (32) *satisfy for any $\varepsilon, h \in (0,1]$, $N \in \mathbb{N}$*

$$\begin{aligned} \max_{i=1,\ldots,N} \mathbb{E}\left[\|X^i_{\varepsilon,n,h}\|^2\right] + \frac{1}{4}\mathbb{E}\left[\sum_{k=1}^{N}\|X^k_{\varepsilon,n,h} - X^{k-1}_{\varepsilon,n,h}\|^2\right] \\ + \tau\mathbb{E}\left[\sum_{k=1}^{N}\mathcal{J}_\varepsilon(X^k_{\varepsilon,n,h})\right] + \frac{\tau\lambda}{2}\mathbb{E}\left[\sum_{k=1}^{N}\|X^k_{\varepsilon,n,h}\|^2\right] \leq C\,. \end{aligned} \tag{33}$$

The next lemma provides an estimate for the difference between the solutions of the auxiliary fully discrete numerical scheme (32) and the solutions of its semi-discrete counterpart (14).

Lemma 5.2. *Let $\varepsilon > 0$, $\delta > 0$, $n \in \mathbb{N}$ be fixed. Let $X^i_{\delta,n}$ be the solution of the semi-discrete scheme* (14) *and let $X^i_{\varepsilon,n,h}$ be the numerical solution of the fully-discrete scheme* (32)*. Then the following estimate holds for $0 < \tau \leq \frac{1}{2}$:*

$$\max_{i=1,\dots,N} \mathbb{E}\left[\|X^i_{\delta,n} - X^i_{\varepsilon,n,h}\|^2\right] \leq C\left(C_n h + C_n^{1/2}\delta^{-\frac{1}{2}} h + C_n\delta + \lambda\|g_n - g_{n,h}\|^2\right).$$

Proof. We set $Z^i = X^i_{\delta,n} - X^i_{\varepsilon,n,h}$ and observe the following equality

(34)
$$\left(Z^i - Z^{i-1}, \Pi_h Z^i\right) = \left(\Pi_h(Z^i - Z^{i-1}), \Pi_h Z^i\right) = \frac{1}{2}\|\Pi_h Z^i\|^2 + \frac{1}{2}\|\Pi_h(Z^i - Z^{i-1})\|^2 - \frac{1}{2}\|\Pi_h Z^{i-1}\|^2,$$

where we used the elementary property of the orthogonal projection that $(v, \Pi_h v) = (\Pi_h v, \Pi_h v)$.

We set $\Phi = \Phi_h = \Pi_h(X^i_{\delta,n} - X^i_{\varepsilon,n,h})$ in (32), (14) (note $\Pi_h X^i_{\varepsilon,n,h} = X^i_{\varepsilon,n,h}$) and obtain after subtracting the respective equations and using (34)

$$\begin{aligned}
&\frac{1}{2}\|\Pi_h Z^i\|^2 + \frac{1}{2}\|\Pi_h(Z^i - Z^{i-1})\|^2 - \frac{1}{2}\|\Pi_h Z^{i-1}\|^2 \\
&\qquad + \tau\left(\frac{\nabla X^i_{\delta,n}}{\sqrt{|\nabla X^i_{\delta,n}|^2 + \varepsilon^2}} - \frac{\nabla X^i_{\varepsilon,n,h}}{\sqrt{|\nabla X^i_{\varepsilon,n,h}|^2 + \varepsilon^2}}, \nabla(X^i_{\delta,n} - X^i_{\varepsilon,n,h})\right) \\
&\qquad - \tau\left(\frac{\nabla X^i_{\delta,n}}{\sqrt{|\nabla X^i_{\delta,n}|^2 + \varepsilon^2}} - \frac{\nabla X^i_{\varepsilon,n,h}}{\sqrt{|\nabla X^i_{\varepsilon,n,h}|^2 + \varepsilon^2}}, \nabla(X^i_{\delta,n} - \Pi_h X^i_{\delta,n})\right) \\
&\qquad + \tau\lambda\left(\left(X^i_{\delta,n} - g_n, \Pi_h Z^i\right) - \left(X^i_{\varepsilon,n,h} - g_{n,h}, \Pi_h Z^i\right)\right) \\
&\quad = \left(Z^{i-1}, \Pi_h Z^i\right)\Delta_i W + \tau\delta\left(\Delta X^i_{\delta,n}, Z^i\right).
\end{aligned} \tag{35}$$

By (8) the fourth term on the left hand side is positive and can be neglected. We estimate the fifth term on the left hand side in (35) using $|\frac{\nabla \cdot}{\sqrt{|\nabla \cdot|^2 + \varepsilon^2}}| \leq 1$ and the Cauchy-Schwarz inequality as

$$\begin{aligned}
&\left(\frac{\nabla X^i_{\delta,n}}{\sqrt{|\nabla X^i_{\delta,n}|^2 + \varepsilon^2}} - \frac{\nabla X^i_{\varepsilon,n,h}}{\sqrt{|\nabla X^i_{\varepsilon,n,h}|^2 + \varepsilon^2}}, \nabla(X^i_{\delta,n} - \Pi_h X^i_{\delta,n})\right) \\
&\qquad \leq 2|\mathcal{O}|^{\frac{1}{2}}\|\nabla(X^i_{\delta,n} - \Pi_h X^i_{\delta,n})\|.
\end{aligned} \tag{36}$$

Using the Cauchy-Schwarz and Young inequalities the last term on the left-hand in (35) can be estimated as

$$\begin{aligned}
\left(X^i_{\delta,n} - g_n, \Pi_h Z^i\right) - \left(X^i_{\varepsilon,n,h} - g_{n,h}, \Pi_h Z^i\right) &= \left(g_{n,h} - g_n, \Pi_h Z^i\right) + \left(X^i_{\delta,n} - X^i_{\varepsilon,n,h}, \Pi_h Z^i\right) \\
&\geq \frac{1}{2}\|\Pi_h Z^i\|^2 - \frac{1}{2}\|g_n - g_{n,h}\|^2,
\end{aligned}$$

and the last term on the right-hand side as

$$\delta\left(-\Delta X^i_{\delta,n},\Pi_h Z^i\right) \le \frac{\delta^2}{2}\|\Delta X^i_{\delta,n}\|^2 + \frac{1}{2}\|\Pi_h Z^i\|^2 .$$

After substituting the above inequalities into (35) we obtain

$$\begin{aligned}(37)\qquad \frac{1}{2}\|\Pi_h Z^i\|^2 &+ \frac{1}{2}\|\Pi_h (Z^i - Z^{i-1})\|^2 - \frac{1}{2}\|\Pi_h Z^{i-1}\|^2 \\ &\le \left(Z^{i-1},\Pi_h Z^i\right)\Delta_i W + \tau|\mathcal{O}|^{\frac{1}{2}}\|\nabla(X^i_{\delta,n} - \Pi_h X^i_{\delta,n})\| \\ &\quad + \frac{\lambda\tau}{2}\|g_n - g_{n,h}\|^2 + \frac{\tau\delta^2}{2}\|\Delta X^i_{\delta,n}\|^2 + \frac{\tau}{2}\|\Pi_h Z^i\|^2.\end{aligned}$$

We estimate the stochastic term as

$$\begin{aligned}\mathbb{E}\left[\left(Z^{i-1},\Pi_h Z^i\right)\Delta_i W\right] &= \mathbb{E}\left[\left(Z^{i-1},\Pi_h Z^i - \Pi_h Z^{i-1}\right)\Delta_i W + \left(Z^{i-1},\Pi_h Z^{i-1}\right)\Delta_i W\right] \\ &= \mathbb{E}\left[\left(\Pi_h Z^{i-1},\Pi_h Z^i - \Pi_h Z^{i-1}\right)\Delta_i W + \left(\Pi_h Z^{i-1},\Pi_h Z^{i-1}\right)\Delta_i W\right] \\ &\le \mathbb{E}\left[\frac{1}{2}\|\Pi_h Z^{i-1}\|^2|\Delta_i W|^2 + \frac{1}{2}\|\Pi_h (Z^i - Z^{i-1})\|^2\right] \\ &= \frac{\tau}{2}\mathbb{E}\left[\|\Pi_h Z^{i-1}\|^2\right] + \frac{1}{2}\mathbb{E}\left[\|\Pi_h (Z^i - Z^{i-1})\|^2\right].\end{aligned}$$

Hence, we obtain after taking expectation in (37) and summing over i that

$$\begin{aligned}\frac{1}{2}\mathbb{E}\left[\|\Pi_h Z^i\|^2\right] &\le \frac{1}{2}\mathbb{E}\left[\|\Pi_h Z^0\|^2\right] + \frac{\tau}{2}\mathbb{E}\left[\sum_{k=1}^{i}\|\Pi_h Z^{k-1}\|^2\right] \\ (38)\qquad &\quad + 2\tau|\mathcal{O}|^{\frac{1}{2}}\mathbb{E}\left[\sum_{k=1}^{i}\|\nabla[X^k_{\delta,n} - \Pi_h X^k_{\delta,n}]\|\right] + \frac{\tau}{2}\mathbb{E}\left[\sum_{k=1}^{i}\|\Pi_h Z^k\|^2\right] \\ &\quad + \frac{T\lambda}{2}\|g_n - g_{n,h}\|^2 + \frac{\tau\delta^2}{2}\mathbb{E}\left[\sum_{k=1}^{i}\|\Delta X^i_{\delta,n}\|^2\right] \\ &= \mathrm{I} + \mathrm{II} + \mathrm{III} + \mathrm{IV} + \mathrm{V} + \mathrm{VI}.\end{aligned}$$

By the Cauchy-Schwartz inequality and (31) we obtain

$$(39)\qquad \mathrm{III} \le (T|\mathcal{O}|)^{\frac{1}{2}}\left(\sum_{k=0}^{i}\tau\mathbb{E}\left[\|\nabla[X^k_{\delta,n} - \Pi_h X^k_{\delta,n}]\|^2\right]\right)^{\frac{1}{2}} \le C\left(C_n\delta^{-1}h^2\right)^{\frac{1}{2}} .$$

Estimate (16) implies

$$\mathrm{VI} \le C_n\delta,$$

and since $X^0_{\varepsilon,n,h} = \Pi_h x^n_0$ and $X^0_{\delta,n} = x^n_0$, we deduce

$$I = \frac{1}{2}\|\Pi_h X^0_{\delta,n} - X^0_{\varepsilon,n,h}\|^2 = 0 .$$

After substituting the above estimates for I, III, VI into (38), we obtain by the discrete Gronwall lemma for sufficiently small τ (e.g. $\tau \leq \frac{1}{2}$) that

$$\max_{i=1,\ldots,N} \mathbb{E}\left[\|\Pi_h Z^i\|^2\right] \leq C\left(\lambda\|g_n - g_{n,h}\| + C_n^{1/2}\delta^{-\frac{1}{2}}h + C_n\delta\right).$$

The statement of the lemma then follows from the above estimate by (29) and (16), since

$$\begin{aligned}\mathbb{E}\left[\|Z^i\|^2\right] &\leq 2\mathbb{E}\left[\|Z^i - \Pi_h Z^i\|^2\right] + 2\mathbb{E}\left[\|\Pi_h Z^i\|^2\right]\\ &= 2\mathbb{E}\left[\|X^i_{\delta,n} - \Pi_h X^i_{\delta,n}\|^2\right] + 2\mathbb{E}\left[\|\Pi_h Z^i\|^2\right]\\ &\leq C_n h + 2\mathbb{E}\left[\|\Pi_h Z^i\|^2\right].\end{aligned}$$

□

The fully discrete numerical approximation of (2) is constructed as follows. For $x_0, g \in \mathbb{L}^2$ we set $X^0_{\varepsilon,h} = \Pi_h x_0$ and $g_h = \Pi_h g$ and determine $X^i_{\varepsilon,h}$, $i = 1, \ldots, N$ as the solution of:

$$\begin{aligned}\left(X^i_{\varepsilon,h}, \Phi_h\right) = \left(X^{i-1}_{\varepsilon,h}, \Phi_h\right) - \tau\left(\frac{\nabla X^i_{\varepsilon,h}}{\sqrt{|\nabla X^i_{\varepsilon,h}|^2 + \varepsilon^2}}, \nabla\Phi_h\right)\\ - \tau\lambda\left(X^i_{\varepsilon,h} - g_h, \Phi_h\right) + \left(X^{i-1}_{\varepsilon,h}, \Phi_h\right)\Delta_i W \qquad \forall \Phi_h \in \mathbb{V}_h\,.\end{aligned} \tag{40}$$

The existence, uniqueness and measurability properties of the solutions of (40) follow analogously as for the solutions of (32).

In the next lemma we estimate the difference between the solutions of the fully discrete numerical scheme (40) and the auxiliary scheme (32).

Lemma 5.3. *Let $x_0 \in L^2(\Omega, \mathcal{F}_0; \mathbb{L}^2)$ and $g \in \mathbb{L}^2$ be given. Then for each $n \in \mathbb{N}$ there exists a constant $C \equiv C(T) > 0$, such that for any $N \in \mathbb{N}$, $n \in \mathbb{N}$, $h, \varepsilon \in (0, 1]$ the following estimate holds for the difference of the numerical solutions of (32) and (40):*

$$\max_{i=1,\ldots,N} \mathbb{E}\left[\|X^i_{\varepsilon,n,h} - X^i_{\varepsilon,h}\|^2\right] \leq C\left(\mathbb{E}\left[\|x_0 - x_0^n\|^2\right] + \lambda\|g - g_n\|^2\right).$$

Proof. We define $Z^i_\varepsilon = X^i_{\varepsilon,n,h} - X^i_{\varepsilon,h}$. After subtracting (32) and (40) we get

$$\begin{aligned}\left(Z^i_\varepsilon, \Phi_h\right) = \left(Z^{i-1}_\varepsilon, \Phi_h\right) - \tau\left(\frac{\nabla X^i_{\varepsilon,n,h}}{\sqrt{|\nabla X^i_{\varepsilon,n,h}|^2 + \varepsilon^2}} - \frac{\nabla X^i_{\varepsilon,h}}{\sqrt{|\nabla X^i_{\varepsilon,h}|^2 + \varepsilon^2}}, \nabla\Phi_h\right)\\ - \tau\lambda\left(Z^i_\varepsilon, \Phi_h\right) - \tau\lambda\left(g_h - g_{n,h}, \Phi_h\right) + \left(Z^{i-1}_\varepsilon, \Phi_h\right)\Delta_i W.\end{aligned}$$

We set $\Phi_h = Z^i_\varepsilon$ and obtain

$$\begin{aligned}\left(Z^i_\varepsilon - Z^{i-1}_\varepsilon, Z^i_\varepsilon\right) = -\tau\left(\frac{\nabla X^i_{\varepsilon,n,h}}{\sqrt{|\nabla X^i_{\varepsilon,n,h}|^2 + \varepsilon^2}} - \frac{\nabla X^i_{\varepsilon,h}}{\sqrt{|\nabla X^i_{\varepsilon,h}|^2 + \varepsilon^2}}, \nabla Z^i_\varepsilon\right)\\ - \tau\lambda\|Z^i_\varepsilon\|^2 - \tau\lambda\left(g_h - g_{n,h}, Z^i_\varepsilon\right) + \left(Z^{i-1}_\varepsilon, Z^i_\varepsilon\right)\Delta_i W.\end{aligned} \tag{41}$$

We rewrite the left-hand side in (41) as

$$\left(Z_\varepsilon^i - Z_\varepsilon^{i-1}, Z_\varepsilon^{i-1}\right) = \frac{1}{2}\|Z_\varepsilon^i\|^2 - \frac{1}{2}\|Z_\varepsilon^{i-1}\|^2 + \frac{1}{2}\|Z_\varepsilon^i - Z_\varepsilon^{i-1}\|^2 ,$$

and by the Cauchy-Schwarz and Young inequalities we estimate

$$\tau\lambda\left(g_h - g_{n,h}, Z_\varepsilon^i\right) \leq \frac{\tau\lambda}{2}\|g_h - g_{n,h}\|^2 + \frac{\tau\lambda}{2}\|Z_\varepsilon^i\|^2.$$

Furthermore, the convexity (8) implies that

$$-\tau\left(\frac{\nabla X_{\varepsilon,n,h}^i}{\sqrt{|\nabla X_{\varepsilon,n,h}^i|^2+\varepsilon^2}} - \frac{\nabla X_{\varepsilon,h}^i}{\sqrt{|\nabla X_{\varepsilon,h}^i|^2+\varepsilon^2}}, \nabla(X_{\varepsilon,n,h}^i - X_{\varepsilon,h}^i)\right) \leq 0.$$

Using the above estimates we deduce from (41) that

$$\begin{aligned}\frac{1}{2}\|Z_\varepsilon^i\|^2 + \frac{1}{2}\|Z_\varepsilon^i - Z_\varepsilon^{i-1}\|^2 - \frac{1}{2}\|Z_\varepsilon^{i-1}\|^2 + \frac{\tau\lambda}{2}\|Z_\varepsilon^{i-1}\|^2 \\ \leq \frac{\tau\lambda}{2}\|g_h - g_{n,h}\|^2 + \left(Z_\varepsilon^{i-1}, Z_\varepsilon^i\right)\Delta_i W\,.\end{aligned} \tag{42}$$

We estimate the last term on the right-hand side above as

$$\begin{aligned}\left(Z_\varepsilon^{i-1}, Z_\varepsilon^i\right)\Delta_i W &= \left(Z_\varepsilon^{i-1}, Z_\varepsilon^i - Z_\varepsilon^{i-1}\right)\Delta_i W + \|Z_\varepsilon^{i-1}\|^2\Delta_i W \\ &\leq \frac{1}{2}\|Z_\varepsilon^i - Z_\varepsilon^{i-1}\|^2 + \frac{1}{2}\|Z_\varepsilon^{i-1}\|^2|\Delta_i W|^2 + \|Z_\varepsilon^{i-1}\|^2\Delta_i W.\end{aligned}$$

We substitute the above identity into (42), neglecting the positive term multiplied by λ on the left-hand side and arrive at

$$\frac{1}{2}\|Z_\varepsilon^i\|^2 - \frac{1}{2}\|Z_\varepsilon^{i-1}\|^2 \leq \frac{\tau\lambda}{2}\|g_h - g_{n,h}\|^2 + \frac{1}{2}\|Z_\varepsilon^{i-1}\|^2|\Delta_i W|^2 + \|Z_\varepsilon^{i-1}\|^2\Delta_i W.$$

Hence, we sum the above inequality over i, take expectation and obtain

$$\frac{1}{2}\mathbb{E}\left[\|Z_\varepsilon^i\|^2\right] \leq \frac{1}{2}\mathbb{E}\left[\|Z_\varepsilon^0\|^2\right] + \frac{\tau}{2}\sum_{k=0}^{i-1}\mathbb{E}\left[\|Z_{\varepsilon,h}^k\|^2\right] + \frac{T\lambda}{2}\|g_h - g_{n,h}\|^2.$$

Finally, an application of the discrete Gronwall lemma yields that

$$\max_{i=1,\ldots,N}\mathbb{E}\left[\|Z_\varepsilon^i\|^2\right] \leq C\left(\mathbb{E}\left[\|\Pi_h(x_0 - x_0^n)\|^2\right] + \lambda\|\Pi_h(g - g_n)\|^2\right),$$

and the statement of the lemma follows by the stability of the $\mathbb{L}^2$-projection (29). □

We define piecewise constant time-interpolants of the discrete solutions $\{X_{\delta,n}^i\}_{i=0}^N$ of (14), $\{X_{\varepsilon,n,h}^i\}_{i=0}^N$ of (32) and $\{X_{\varepsilon,h}^i\}_{i=0}^N$ of (40) for $t \in [0,T)$ as

$$\overline{X}_\tau^{\delta,n} = X_{\delta,n}^i, \quad \overline{X}_{\tau,h}^{\varepsilon,n} = X_{\varepsilon,n,h}^i, \quad \overline{X}_{\tau,h}^{\varepsilon} = X_{\varepsilon,h}^i \quad \text{if} \quad t \in (t_{i-1}, t_i]. \tag{43}$$

In the next theorem we conclude the paper by showing the convergence of the fully discrete numerical approximation (40) to the unique SVI solution of the total variation flow (1) (cf. Definition 3.1).

Theorem 5.1. *Let X be the SVI solution of* (1) *and let $\overline{X}^{\varepsilon}_{\tau,h}$ be the time-interpolant (43) of the solutions of the fully-discrete scheme* (40)*. Then*

$$\lim_{\varepsilon\to 0}\lim_{\tau,h\to 0}\|X-\overline{X}^{\varepsilon}_{\tau,h}\|^2_{L^2(\Omega\times(0,T);\mathbb{L}^2)}\to 0. \tag{44}$$

Proof. For $x_0\in L^2(\Omega,\mathcal{F}_0;\mathbb{L}^2)$ and $g\in\mathbb{L}^2$ for $n\in\mathbb{N}$ we set $x_0^n=\mathcal{P}_n x_0$, $g_n=\mathcal{P}_n g$ where $\mathcal{P}_n:\mathbb{L}^2\to\mathbb{V}_n$ is the orthogonal $\mathbb{L}^2$-projection onto the finite dimensional eigenspace $\mathbb{V}_n=\operatorname{span}\{e_0,\dots,e_n\}\subset\mathbb{H}^1_0$. By construction the sequences $\{x_0^n\}_{n\in\mathbb{N}}\subset\mathbb{H}^1_0$, $\{g_n\}_{n\in\mathbb{N}}\subset\mathbb{H}^1_0$ satisfy $x_0^n\to x_0\in L^2(\Omega,\mathcal{F}_0;\mathbb{L}^2)$, $n\in\mathbb{N}$, $g_n\to g\in\mathbb{L}^2$. Below, we consider (3), (32) with the data x_0^n, g_n defined above.

By the triangle inequality we get

$$\begin{aligned}\frac{1}{4}\|X-\overline{X}^{\varepsilon}_{\tau,h}\|^2_{L^2(\Omega\times(0,T);\mathbb{L}^2)}\leq&\|X-X^{\delta}_n\|^2_{L^2(\Omega\times(0,T);\mathbb{L}^2)}+\|X^{\delta}_n-\overline{X}^{\delta,n}_{\tau}\|^2_{L^2(\Omega\times(0,T);\mathbb{L}^2)}\\&+\|\overline{X}^{\delta,n}_{\tau}-\overline{X}^{\varepsilon,n}_{\tau,h}\|^2_{L^2(\Omega\times(0,T);\mathbb{L}^2)}+\|\overline{X}^{\varepsilon,n}_{\tau,h}-\overline{X}^{\varepsilon}_{\tau,h}\|^2_{L^2(\Omega\times(0,T);\mathbb{L}^2)}\\&=:\mathrm{I}+\mathrm{II}+\mathrm{III}+\mathrm{IV}.\end{aligned}\tag{45}$$

From Theorem 3.1 it follows that

$$\lim_{\varepsilon\to 0}\lim_{n\to\infty}\lim_{\delta\to 0}\mathrm{I}=\lim_{\varepsilon\to 0}\lim_{n\to\infty}\lim_{\delta\to 0}\mathbb{E}\left[\|X-X^{\delta}_n\|^2\right]=0.$$

By Lemma 4.3 we deduce for the second term that

$$\lim_{\tau\to 0}\mathrm{II}=\lim_{\tau\to 0}\mathbb{E}\left[\|X^{\delta}_n-\overline{X}^{\delta,n}_{\tau}\|^2\right]=0.$$

For the third term we get by Lemma 5.2 and (29) that

$$\lim_{\delta\to 0}\lim_{\tau,h\to 0}\mathrm{III}\leq\lim_{\delta\to 0}\lim_{\tau,h\to 0}C\left(C_nh+C_n^{1/2}\delta^{-1/2}h+C_n\delta+\|g_n-\Pi_hg_n\|^2\right)=0.$$

By Lemma 5.3 the fourth term satisfies

$$\lim_{n\to\infty}\mathrm{IV}\leq\lim_{n\to\infty}\max_{i=1,\dots,N}\mathbb{E}\left[\|X^i_{\varepsilon,n,h}-X^i_{\varepsilon,h}\|^2\right]\leq\lim_{n\to\infty}C\left(\mathbb{E}\left[\|x_0-x_0^n\|^2\right]+\|g-g_n\|^2\right)=0.$$

Finally, we consecutively take $\tau,h\to 0$, $\delta\to 0$, $n\to\infty$ and $\varepsilon\to 0$ in (45) and use the above convergence of $\mathrm{I}-\mathrm{IV}$ to obtain (44). □

References

[1] H. Attouch, G. Buttazzo, and G. Michaille. *Variational analysis in Sobolev and BV spaces*, volume 6 of *MPS/SIAM Series on Optimization*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA; Mathematical Programming Society (MPS), Philadelphia, PA, 2006. Applications to PDEs and optimization.

[2] V. Barbu and M. Röckner. Stochastic variational inequalities and applications to the total variation flow perturbed by linear multiplicative noise. *Arch. Ration. Mech. Anal.*, 209(3):797–834, 2013.

[3] Ľ. Baňas and M. Ondreját. Numerical approximation of probabilistically weak and strong solutions of the stochastic total variation flow. *M2AN Math. Model. Numer. Anal.*, 2022. accepted.

[4] Ľ. Baňas, M. Röckner, and A. Wilke. Convergent numerical approximation of the stochastic total variation flow. *Stoch. Partial Differ. Equ. Anal. Comput.*, 9(2):437–471, 2021.

[5] Ľ. Baňas, M. Röckner, and A. Wilke. Correction to: Convergent numerical approximation of the stochastic total variation flow. *Stoch. Partial Differ. Equ. Anal. Comput.*, 2022. accepted.

[6] S. C. Brenner and L. R. Scott. *The Mathematical Theory of Finite Element Methods (second edition).* Springer-Verlag, New York, 2002.

[7] X. Feng and A. Prohl. Analysis of total variation flow and its finite element approximations. *M2AN Math. Model. Numer. Anal.*, 37(3):533–556, 2003.

[8] W. Liu and M. Röckner. *Stochastic partial differential equations: an introduction*. Universitext. Springer, Cham, 2015.

Department of Mathematics, Bielefeld University, 33501 Bielefeld, Germany
Email address: banas@math.uni-bielefeld.de

Department of Mathematics, Bielefeld University, 33501 Bielefeld, Germany and Academy of Mathematics and Systems Science, CAS, Beijing
Email address: roeckner@math.uni-bielefeld.de

Department of Mathematics, Bielefeld University, 33501 Bielefeld, Germany
Email address: awilke@math.uni-bielefeld.de