CONVERGENT NUMERICAL APPROXIMATION OF THE STOCHASTIC TOTAL VARIATION FLOW

ĽUBOMÍR BAŇAS, MICHAEL RÖCKNER, AND ANDRÉ WILKE

ABSTRACT. We study the stochastic total variation flow (STVF) equation with linear multiplicative noise. By considering a limit of a sequence of regularized stochastic gradient flows with respect to a regularization parameter ε we obtain the existence of a unique variational solution of the STVF equation which satisfies a stochastic variational inequality. Furthermore, we propose a fully discrete implicit numerical scheme for the regularized gradient flow equation and show that the numerical solution converges to the solution of the unregularized STVF equation. We perform numerical experiments to demonstrate the practicability of the numerical approximation.

1. INTRODUCTION

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

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

where $\mathcal{O}\subset\mathbb{R}^d$, $d\geq 1$ is a bounded, convex domain with a piecewise C^2-smooth boundary $\partial\mathcal{O}$, and $\lambda\geq 0$, $T>0$ are constants. We assume that $x_0,\,g\in\mathbb{L}^2$ and consider a one dimensional real-valued Wiener process W, for simplicity; generalization for a sufficiently regular trace-class noise is straightforward.

Equation (1) can be interpreted as a stochastically perturbed gradient flow of the penalized total variation energy functional

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

The minimization of above functional, so-called ROF-method, is a prototypical approach for image denoising, cf. [13]; in this context the function g represents a given noisy image and λ serves as a penalization parameter. Further applications of the functional include, for instance, elastoplasticity and the modeling of damage and fracture, for more details see for instance [4] and the references therein.

The use of stochastically perturbed gradient flows has proven useful in image processing. Stochastic numerical methods for models with nonconvex energy functionals are able to avoid local energy minima and thus achieve faster convergence and/or more accurate results

than their deterministic counterparts; see [10] which applies stochastic level-set method in image segmentation, and [14] which uses stochastic gradient flow of a modified (non-convex) total variation energy functional for binary tomography.

Due to the singular character of total variation flow (1), it is convenient to perform numerical simulations using a regularized problem

$$\mathrm{d}X^\varepsilon = \mathrm{div}\left(\frac{\nabla X^\varepsilon}{\sqrt{|\nabla X^\varepsilon|^2 + \varepsilon^2}}\right)\mathrm{d}t - \lambda(X^\varepsilon - g)\,\mathrm{d}t + X^\varepsilon\,\mathrm{d}W \qquad \text{in } (0, T) \times \mathcal{O},$$

$$(3) \qquad X^\varepsilon = 0 \qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{on } (0, T) \times \partial\mathcal{O},$$

$$X^\varepsilon(0) = x_0 \qquad\qquad\qquad\qquad\qquad\qquad\quad \text{in } \mathcal{O},$$

with a regularization parameter $\varepsilon > 0$. In the deterministic setting ($W \equiv 0$) equation (3) corresponds to the gradient flow of the regularized energy functional

$$(4) \qquad \mathcal{J}_{\varepsilon,\lambda}(u) := \int_{\mathcal{O}} \sqrt{|\nabla u|^2 + \varepsilon^2}\,\mathrm{d}x + \frac{\lambda}{2}\int_{\mathcal{O}} |u - g|^2\,\mathrm{d}x.$$

It is well-known that the minimizers of the above regularized energy functional converge to the minimizers of (2) for $\varepsilon \to 0$, cf. [7] and the references therein.

Due to the singular character of the diffusion term in (1) the classical variational approach for the analysis of stochastic partial differential equations (SPDEs), see e.g. [11], [12], is not applicable to this problem. To study well-posedeness of singular gradient flow problems it is convenient to apply the solution framework developed in [3] which characterizes the solutions of (1) as stochastic variational inequalities (SVIs). In this paper, we show the well posedness of variational solutions using the practically relevant regularization procedure (3) which, in the regularization limit, yields a SVI solution in the sense of [3].

Convergence of numerical approximation of (3) in the deterministic setting ($W \equiv 0$) has been shown in [7]. Analogically to the deterministic setting we construct a numerical approximation of (1) via a full discretization of the regularized problem (3). The deterministic variational inequality framework used in the the numerical analysis of [7] is not directly transferable to the stochastic setting. Instead, we show the convergence of the proposed numerical approximation of (3) to the SVI solution of (1) via an additional regularization step on the discrete level. The convergence analysis of the discrete approximation is inspired by the analytical approach of [8] where the SVI solution concept was applied to the stochastic p-Laplace equation. As far as we are aware, the present work is the first to show convergence of implementable numerical approximation for singular stochastic gradient flows in the framework of stochastic variational inequalities. Throughout the paper, we will refer to the solutions which satisfy a stochastic variational inequality as SVI solutions, and to the classical SPDE solutions as variational solutions.

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 (3) and its convergence towards a unique SVI solution of (1) for $\varepsilon \to 0$ is shown in Section 3. In Section 4 we introduce numerical approximation of (3) and show the

convergence of the proposed numerical approximation to the SVI solution of (1). Numerical experiments are presented in Section 5.

2. NOTATION AND PRELIMINARIES

Throughout the paper we denote by C a generic positive constant that may change from line to line. For $1 \leq p \leq \infty$, we denote by $(\mathbb{L}^p, \|\cdot\|_{\mathbb{L}^p})$ the standard spaces of p-th order integrable functions on $\mathcal{O}$, and use $\|\cdot\| := \|\cdot\|_{\mathbb{L}^2}$ and $(\cdot, \cdot) := (\cdot, \cdot)_{\mathbb{L}^2}$ for the $\mathbb{L}^2$-inner product. For $k \in \mathbb{N}$ we denote the usual Sobolev space on $\mathcal{O}$ as $(\mathbb{H}^k, \|\cdot\|_{\mathbb{H}^k})$, and $(\mathbb{H}_0^1, \|\cdot\|_{\mathbb{H}_0^1})$ stands for the $\mathbb{H}^1$ space with zero trace on $\partial\mathcal{O}$ with its dual space $(\mathbb{H}^{-1}, \|\cdot\|_{\mathbb{H}^{-1}})$. Furthermore, we set $\langle\cdot, \cdot\rangle := \langle\cdot, \cdot\rangle_{\mathbb{H}^{-1} \times \mathbb{H}_0^1}$, where $\langle\cdot, \cdot\rangle_{\mathbb{H}^{-1} \times \mathbb{H}_0^1}$ is the duality pairing between $\mathbb{H}_0^1$ and $\mathbb{H}^{-1}$. The functional (4) with $\lambda = 0$ will be denoted as $\mathcal{J}_\varepsilon := \mathcal{J}_{\varepsilon,0}$. We say that a function $X \in L^1(\Omega \times (0, T); \mathbb{L}^2)$ is $\mathcal{F}_t$-progressively measurable if $X\mathbb{1}_{[0,t]}$ is $\mathcal{F}_t \otimes \mathcal{B}([0,t])$-measurable for all $t \in [0, T]$.

For the convenience of the reader we state some basic definitions below.

Definition 2.1. *Let $\mathbb{H}$ be a real Banach space, $A : D(A) \to \mathbb{H}$ a linear operator and $\rho(A)$ its resolvent set. For a real number $\xi \in \rho(A)$ we define the resolvent $R_\xi : \mathbb{H} \to \mathbb{H}$ of A as*

$$R_\xi(x) := (I - \xi A)^{-1}x.$$

Furthermore we define the Yosida approximation of A as

$$T_\xi(x) := AR_\xi = \frac{1}{\xi}(I - R_\xi)x.$$

Definition 2.2. *The mapping $\mathcal{P}_m : \mathbb{L}^2 \to \mathbb{V}_m \subset \mathbb{L}^2$ which satisfies*

$$(w - \mathcal{P}_m w, v_m) = 0 \quad \forall v_m \in \mathbb{V}_m.$$

defines the $\mathbb{L}^2$-orthogonal projection onto $\mathbb{V}_m$.

Definition 2.3. *A function $u \in L^1(\mathcal{O})$ is called a function of bounded variation, if its total variation*

$$(5) \qquad \int_{\mathcal{O}} |\nabla u|\, \mathrm{d}x := \sup\left\{-\int_{\mathcal{O}} u \operatorname{div} \mathbf{v}\, \mathrm{d}x;\ \mathbf{v} \in C_0^\infty(\mathcal{O}, \mathbb{R}^d),\ \|\mathbf{v}\|_{L^\infty} \leq 1\right\},$$

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

For $u \in BV(\mathcal{O})$ we denote

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

The following proposition plays an important role in the analysis below; the proposition holds for convex domains with piecewise smooth boundary, which includes the case of practically relevant polygonal domains, cf. [3, Proposition 8.2 and Remark 8.1].

Proposition 2.1. *Let $\mathcal{O} \subset \mathbb{R}^d$, $d \geq 1$ be a bounded domain with a piecewise C^2-smooth and convex boundary. Let $g : [0, \infty) \to [0, \infty)$ be a continuous and convex function of at most quadratic growth such that $g(0) = 0$, then it holds*

$$
(6) \qquad \int_{\mathcal{O}} g(|\nabla R_\xi(y)|) \, dx \leq \int_{\mathcal{O}} g(|\nabla y|) \, dx, \quad \forall y \in \mathbb{H}_0^1.
$$

3. Well posedness of STVF

In this section we show existence and uniques of the SVI solution of (1) (see below for a precise definition) via a two-level regularization procedure. To be able to treat problems with $\mathbb{L}^2$-regular data, i.e., $x_0 \in L^2(\Omega, \mathcal{F}_0; \mathbb{L}^2)$, $g \in \mathbb{L}^2$ we consider a $\mathbb{H}_0^1$-approximating sequence $\{x_0^n\}_{n \in \mathbb{N}} \subset L^2(\Omega, \mathcal{F}_0; \mathbb{H}_0^1)$ s.t. $x_0^n \to x_0$ in $L^2(\Omega, \mathcal{F}_0; \mathbb{L}^2)$ for $n \to \infty$ and $\{g^n\}_{n \in \mathbb{N}} \subset \mathbb{H}_0^1$ s.t. $g^n \to g$ in $\mathbb{L}^2$ for $n \to \infty$. We introduce a regularization of (3) as

$$
dX_n^{\varepsilon,\delta} = \delta \Delta X_n^{\varepsilon,\delta} \, dt + \mathrm{div} \left(\frac{\nabla X_n^{\varepsilon,\delta}}{\sqrt{|\nabla X_n^{\varepsilon,\delta}|^2 + \varepsilon^2}} \right) \, dt
$$

$$
(7) \qquad\qquad - \lambda(X_n^{\varepsilon,\delta} - g^n) \, dt + X_n^{\varepsilon,\delta} \, dW(t) \qquad\qquad \text{in } (0, T) \times \mathcal{O},
$$

$$
X_n^{\varepsilon,\delta}(0) = x_0^n \qquad\qquad\qquad\qquad \text{in } \mathcal{O},
$$

where $\delta > 0$ is an additional regularization parameter.

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

$$
(8) \quad \langle A^{\varepsilon,\delta} u, v \rangle_{\mathbb{H}^{-1} \times \mathbb{H}_0^1} = \int_{\mathcal{O}} \delta \nabla u \nabla v + \frac{\nabla u}{\sqrt{|\nabla u|^2 + \varepsilon^2}} \nabla v + \lambda(u - g^n) v \, dx \quad \forall u, v \in \mathbb{H}_0^1,
$$

and note that (7) is equivalent to

$$
(9) \qquad\qquad dX_n^{\varepsilon,\delta} + A^{\varepsilon,\delta} X_n^{\varepsilon,\delta} \, dt = X_n^{\varepsilon,\delta} \, dW(t),
$$

$$
X_n^{\varepsilon,\delta}(0) = x_0^n.
$$

The operator $A^{\varepsilon,\delta}$ is coercive, demicontinuos and satisfies (cf. [12, Remark 4.1.1])

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

$$
(11) \quad \|A^{\varepsilon,\delta}(u)\|_{\mathbb{H}^{-1}} \leq C(\delta, \lambda, \|g^n\|)(\|u\|_{\mathbb{H}_0^1} + 1), \qquad\qquad \forall u \in \mathbb{H}_0^1.
$$

The following monotonicity property, which follows from the convexity of the function $\sqrt{|\cdot|^2 + \varepsilon^2}$, will be used frequently in the subsequent arguments

$$
\left(\frac{\nabla X}{\sqrt{|\nabla X|^2 + \varepsilon^2}} - \frac{\nabla Y}{\sqrt{|\nabla Y|^2 + \varepsilon^2}}, \nabla(X - Y) \right)
$$

$$
(12) \qquad = \left(\frac{\nabla X}{\sqrt{|\nabla X|^2 + \varepsilon^2}}, \nabla(X - Y) \right) + \left(\frac{\nabla Y}{\sqrt{|\nabla Y|^2 + \varepsilon^2}}, \nabla(Y - X) \right)
$$

$$
\geq \mathcal{J}_\varepsilon(X) - \mathcal{J}_\varepsilon(Y) + \mathcal{J}_\varepsilon(Y) - \mathcal{J}_\varepsilon(X) = 0.
$$

The existence and uniqueness of a variational solution of (7) is established in the next lemma; we note that the result only requires $\mathbb{L}^2$-regularity of data.

Lemma 3.1. *For any $\varepsilon, \delta > 0$ and $x_0^n \in L^2(\Omega, \mathcal{F}_0; \mathbb{H}_0^1)$, $g^n \in \mathbb{H}_0^1$ there exists a unique variational solution $X_n^{\varepsilon,\delta} \in L^2(\Omega; C([0,T]; \mathbb{L}^2))$ of (7). Furthermore, there exists a $C \equiv C(T) > 0$ such that the following estimate holds*

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

Proof of Lemma 3.1. On noting the properties (10)-(11) of the operator $A^{\varepsilon,\delta}$ for $\varepsilon, \delta > 0$ the classical theory, cf. [12], implies that for any given data $x_0^n \in L^2(\Omega, \mathcal{F}_0; \mathbb{H}_0^1)$, $g^n \in \mathbb{H}_0^1$ there exists a unique variational solution $X_n^{\varepsilon,\delta} \in L^2(\Omega; C([0,T]; \mathbb{L}^2))$ of (7) which satisfies the stability estimate.   □

In next step, we show a priori estimate for the solution of (7) in stronger norms; the estimate requires $\mathbb{H}_0^1$-regularity of the data.

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 $X_n^{\varepsilon,\delta}$ of (7) satisfies*

$$(13) \qquad \mathbb{E}\left[\sup_{t \in [0,T]} \|\nabla X_n^{\varepsilon,\delta}(t)\|^2\right] + \delta \mathbb{E}\left[\int_0^t \|\Delta X_n^{\varepsilon,\delta}(s)\|^2 \, \mathrm{d}s\right] \leq C\left(\mathbb{E}\left[\|x_0^n\|_{\mathbb{H}_0^1}^2\right] + \|g^n\|_{\mathbb{H}_0^1}^2\right).$$

Proof of Lemma 3.2. Let $\{e_i\}_{i=0}^{\infty}$ be an orthonormal basis of eigenfunctions of the Dirichlet Laplacian $-\Delta$ on $\mathbb{L}^2$ and $\mathbb{V}_m := \mathrm{span}\{e_0, \ldots, e_m\}$. Let $\mathcal{P}_m : \mathbb{L}^2 \to \mathbb{V}_m$ be the $\mathbb{L}^2$-orthogonal projection onto $\mathbb{V}_m$.

For fixed ε, δ, n the Galerkin approximation $X_{n,m}^{\varepsilon,\delta} \in \mathbb{V}_m$ of $X_n^{\varepsilon,\delta}$ satisfies

$$\mathrm{d}X_{n,m}^{\varepsilon,\delta} = \delta \Delta X_{n,m}^{\varepsilon,\delta} \, \mathrm{d}t + \mathcal{P}_m \mathrm{div}\left(\frac{\nabla X_{n,m}^{\varepsilon,\delta}}{\sqrt{|\nabla X_{n,m}^{\varepsilon,\delta}|^2 + \varepsilon^2}}\right) \mathrm{d}t$$

$$(14) \qquad\qquad - \lambda(X_{n,m}^{\varepsilon,\delta} - \mathcal{P}_m g^n) \, \mathrm{d}t + X_{n,m}^{\varepsilon,\delta} \, \mathrm{d}W(t),$$

$$X_{n,m}^{\varepsilon,\delta}(0) = \mathcal{P}_m x_0^n.$$

By standard arguments, cf. [12, Theorem 5.2.6], there exists a $X_n^{\varepsilon,\delta} \in L^2(\Omega; C([0,T]; \mathbb{L}^2))$ such that $X_{n,m}^{\varepsilon,\delta} \rightharpoonup X_n^{\varepsilon,\delta}$ in $L^2(\Omega \times (0,T); \mathbb{L}^2)$ for $m \to \infty$. We use Itô's formula for

$\|\nabla X_{n,m}^{\varepsilon,\delta}(t)\|^2$ to obtain

$$
\begin{aligned}
\frac{1}{2}\|\nabla X_{n,m}^{\varepsilon,\delta}(t)\|^2 = & \frac{1}{2}\|\nabla \mathcal{P}_m x_0^n\|^2 - \delta \int_0^t \|\Delta X_{n,m}^{\varepsilon,\delta}(s)\|^2 \, \mathrm{d}s \\
& - \int_0^t \left(\operatorname{div} \frac{\nabla X_{n,m}^{\varepsilon,\delta}(s)}{\sqrt{|\nabla X_{n,m}^{\varepsilon,\delta}(s)|^2 + \varepsilon^2}}, \Delta X_{n,m}^{\varepsilon,\delta}(s) \right) \, \mathrm{d}s \\
& - \lambda \int_0^t \left((X_{n,m}^{\varepsilon,\delta}(s) - g^n), \Delta X_{n,m}^{\varepsilon,\delta}(s) \right) \, \mathrm{d}s \\
& - \int_0^t \left(\Delta X_{n,m}^{\varepsilon,\delta}(s), X_{n,m}^{\varepsilon,\delta}(s) \, \mathrm{d}W(s) \right) \, \mathrm{d}s \\
& + \frac{1}{2} \int_0^t \|X_{n,m}^{\varepsilon,\delta}(s)\|_{\mathbb{H}_0^1}^2 \, \mathrm{d}s.
\end{aligned}
\tag{15}
$$

Let T_ξ be the Yosida-approximation and R_ξ the resolvent of the Dirichlet Laplacian $-\Delta$ on L^2, respectively; see Definition 2.1. By the convexity, cf. (12), we get

$$
\begin{aligned}
& \left(-\Delta X_{n,m}^{\varepsilon,\delta}(s), \operatorname{div} \frac{\nabla X_{n,m}^{\varepsilon,\delta}(s)}{\sqrt{|\nabla X_{n,m}^{\varepsilon,\delta}(s)|^2 + \varepsilon^2}} \right) \\
& = \lim_{\xi \to \infty} \left(T_\xi X_{n,m}^{\varepsilon,\delta}(s), \operatorname{div} \frac{\nabla X_{n,m}^{\varepsilon,\delta}(s)}{\sqrt{|\nabla X_{n,m}^{\varepsilon,\delta}(s)|^2 + \varepsilon^2}} \right) \\
& = \lim_{\xi \to \infty} \frac{1}{\xi} \left(X_{n,m}^{\varepsilon,\delta}(s) - R_\xi X_{n,m}^{\varepsilon,\delta}(s), \operatorname{div} \frac{\nabla X_{n,m}^{\varepsilon,\delta}(s)}{\sqrt{|\nabla X_{n,m}^{\varepsilon,\delta}(s)|^2 + \varepsilon^2}} \right) \\
& = \lim_{\xi \to \infty} \frac{1}{\xi} \left(\nabla R_\xi X_{n,m}^{\varepsilon,\delta}(s) - \nabla X_{n,m}^{\varepsilon,\delta}(s), \frac{\nabla X_{n,m}^{\varepsilon,\delta}(s)}{\sqrt{|\nabla X_{n,m}^{\varepsilon,\delta}(s)|^2 + \varepsilon^2}} \right) \\
& \leq \lim_{\xi \to \infty} \frac{1}{\xi} \left(\mathcal{J}_\varepsilon(\nabla R_\xi X_{n,m}^{\varepsilon,\delta}(s)) - \mathcal{J}_\varepsilon(\nabla X_{n,m}^{\varepsilon,\delta}(s)) \right) \\
& \leq 0,
\end{aligned}
$$

where we used Proposition 2.1 in the last step above. The Burkholder-Davis-Gundy inequaltiy for $p = 1$ implies that

$$
\begin{aligned}
\mathbb{E}\left[\sup_{t\in[0,T]}\int_0^t \|\nabla X_{n,m}^{\varepsilon,\delta}(s)\|^2\,\mathrm{d}W(s)\right] &\leq C\mathbb{E}\left[\left(\int_0^T \|\nabla X_{n,m}^{\varepsilon,\delta}(s)\|^4\,\mathrm{d}s\right)^{\frac{1}{2}}\right] \\
&\leq C\mathbb{E}\left[\sup_{t\in[0,T]}\|\nabla X_{n,m}^{\varepsilon,\delta}(t)\|\left(\int_0^T \|\nabla X_{n,m}^{\varepsilon,\delta}(s)\|^2\,\mathrm{d}s\right)^{\frac{1}{2}}\right] \\
&\leq \frac{1}{4}\mathbb{E}\left[\sup_{t\in[0,T]}\|\nabla X_{n,m}^{\varepsilon,\delta}(t)\|^2\right] + C\mathbb{E}\left[\int_0^T \|\nabla X_{n,m}^{\varepsilon,\delta}(s)\|^2\,\mathrm{d}s\right].
\end{aligned}
\tag{16}
$$

After taking supremum over t and expectation in (15), using (16) along with the Tonelli and Gronwall lemmas we obtain

$$
\mathbb{E}\left[\sup_{t\in[0,T]}\|\nabla X_{n,m}^{\varepsilon,\delta}(t)\|^2 + \delta\int_0^T \|\Delta X_{n,m}^{\varepsilon,\delta}(s)\|^2\,\mathrm{d}s\right] \leq C(\mathbb{E}\left[\|x_0^n\|_{\mathbb{H}_0^1}^2\right] + \|g^n\|_{\mathbb{H}_0^1}^2).
$$

Hence, from the sequence $\{X_{n,m}^{\varepsilon,\delta}\}_{m\in\mathbb{N}}$ we can extract a subsequence (not relabeled), s.t. for $m \to \infty$

$$
\begin{aligned}
X_{n,m}^{\varepsilon,\delta} &\rightharpoonup X_n^{\varepsilon,\delta} \text{ in } L^2(\Omega; L^2((0,T);\mathbb{H}^2) \\
X_{n,m}^{\varepsilon,\delta} &\rightharpoonup^* X_n^{\varepsilon,\delta} \text{ in } L^2(\Omega; L^\infty((0,T);\mathbb{H}_0^1)).
\end{aligned}
$$

By lower-semicontinuity of the norms, we get

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

$\square$

We define the SVI solution of (3) and (1) analogically to [3, Definition 3.1] as a stochastic variational inequality.

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 a $\mathcal{F}_t$-progressively measurable map $X^\varepsilon \in L^2(\Omega; C([0,T]; \mathbb{L}^2)) \cap L^2(\Omega; L^1((0,T); BV(\mathcal{O})))$ (denoted by $X \in L^2(\Omega; C([0,T]; \mathbb{L}^2)) \cap L^2(\Omega; L^1((0,T); BV(\mathcal{O})))$ for $\varepsilon = 0$) is called a SVI solution of (3) (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 $Z \in L^2(\Omega \times (0,T); \mathbb{H}_0^1)$ with $\mathbb{P}$-a.s. continuous sample paths which satisfy the equation*

$$
\mathrm{d}Z(t) = -G(t)\,\mathrm{d}t + Z(t)\,\mathrm{d}W(t), \ t \in [0,T],
\tag{17}
$$

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

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

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

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

In the next theorem we show the existence and uniqueness of a SVI solution to (3) for $\varepsilon > 0$ in the sense of the Definition 3.1.

Theorem 3.1. *Let $0 < T < \infty$ and $x_0 \in L^2(\Omega, \mathcal{F}_0; \mathbb{L}^2)$, $g \in \mathbb{L}^2$. For each $\varepsilon \in (0,1]$ there exists a unique SVI solution X^ε of (3). Moreover, any two SVI solutions $X_1^\varepsilon, X_2^\varepsilon$ with $x_0 \equiv x_0^1$, $g \equiv g^1$ and $x_0 \equiv x_0^2$, $g \equiv g^2$ satisfy*

$$
(20) \qquad \mathbb{E}\left[\|X_1^\varepsilon(t) - X_2^\varepsilon(t)\|^2\right] \leq C\left(\mathbb{E}\left[\|x_0^1 - x_0^2\|^2\right] + \|g^1 - g^2\|^2\right),
$$

for all $t \in [0, T]$.

Proof of Theorem 3.1.
We show that for fixed $\varepsilon > 0$ the sequence $\{X_n^{\varepsilon,\delta}\}_{\delta,n}$ of variational solutions of (7) is a Cauchy-sequence w.r.t. δ for any fixed $n \in \mathbb{N}$, and then show that it is a Cauchy-sequence w.r.t. n for $\delta \equiv 0$.

We denote by $X_{n_1}^{\varepsilon,\delta_1}, X_{n_2}^{\varepsilon,\delta_2}$ the solutions of (7) for $\delta \equiv \delta_1$, $\delta \equiv \delta_2$ and $x_0 \equiv x_0^{n_1} \in L^2(\Omega, \mathcal{F}_0; \mathbb{H}_0^1)$, $x_0 \equiv x_0^{n_2} \in L^2(\Omega, \mathcal{F}_0; \mathbb{H}_0^1)$, respectively, where $x_0^{n_1}$, $x_0^{n_2}$ belong to the $\mathbb{H}_0^1$-approximating sequence of $x_0 \in L^2(\Omega, \mathcal{F}_0; \mathbb{L}^2)$. By Itô's formula it follows that

$$
\frac{1}{2}\|X_{n_1}^{\varepsilon,\delta_1}(t) - X_{n_2}^{\varepsilon,\delta_2}(t)\|^2
$$

$$
= \frac{1}{2}\|x_0^{n_1} - x_0^{n_2}\|^2 + \int_0^t \left(\delta_1 \Delta X_{n_1}^{\varepsilon,\delta_1}(s) - \delta_2 \Delta X_{n_2}^{\varepsilon,\delta_2}(s), X_{n_1}^{\varepsilon,\delta_1}(s) - X_{n_2}^{\varepsilon,\delta_2}(s)\right)\,\mathrm{d}s
$$

$$
- \int_0^t \left(\frac{\nabla X_{n_1}^{\varepsilon,\delta_1}(s)}{\sqrt{|\nabla X_{n_2}^{\varepsilon,\delta_2}(s)|^2 + \varepsilon^2}} - \frac{\nabla X_{n_2}^{\varepsilon,\delta_2}(s)}{\sqrt{|\nabla X_{n_2}^{\varepsilon,\delta_2}(s)|^2 + \varepsilon^2}}, \nabla(X_{n_1}^{\varepsilon,\delta_1}(s) - X_{n_2}^{\varepsilon,\delta_2}(s))\right)\,\mathrm{d}s
$$

$$
- \lambda \int_0^t \|(X_{n_1}^{\varepsilon,\delta_1}(s) - X_{n_2}^{\varepsilon,\delta_2}(s)\|^2\,\mathrm{d}s + \int_0^t \|X_{n_1}^{\varepsilon,\delta_1}(s) - X_{n_2}^{\varepsilon,\delta_2}(s)\|^2\,\mathrm{d}W(s)
$$

$$
+ \int_0^t \|X_{n_1}^{\varepsilon,\delta_1}(s) - X_{n_2}^{\varepsilon,\delta_2}(s)\|^2\,\mathrm{d}s.
$$

We note that

$$
\left(\delta_1 \Delta X_{n_1}^{\varepsilon,\delta_1}(s) - \delta_2 \Delta X_{n_2}^{\varepsilon,\delta_2}(s), X_{n_1}^{\varepsilon,\delta_1}(s) - X_{n_2}^{\varepsilon,\delta_2}(s)\right)
$$

$$
= -\left(\delta_1 \nabla X_{n_1}^{\varepsilon,\delta_1}(s) - \delta_2 \nabla X_{n_2}^{\varepsilon,\delta_2}(s), \nabla X_{n_1}^{\varepsilon,\delta_1}(s) - \nabla X_{n_2}^{\varepsilon,\delta_2}(s)\right)
$$

$$
\leq C(\delta_1 + \delta_2)(\|\nabla X_{n_1}^{\varepsilon,\delta_1}(s)\|^2 + \|\nabla X_{n_2}^{\varepsilon,\delta_2}(s)\|^2).
$$

Hence by using the convexity (12), Lemma 3.2, the Burkholder-Davis-Gundy inequality for $p = 1$, the Tonelli and Gronwall lemmas we obtain

$$
\mathbb{E}\left[\sup_{t\in[0,T]} \|X_{n_1}^{\varepsilon,\delta_1}(t) - X_{n_2}^{\varepsilon,\delta_2}(t)\|^2\right] \leq C\mathbb{E}\left[\|x_0^{n_1} - x_0^{n_2}\|^2\right]
$$

(21)
$$
+ C\left(\mathbb{E}\left[\|x_0^{n_1}\|_{\mathbb{H}_0^1}^2\right], \mathbb{E}\left[\|x_0^{n_2}\|_{\mathbb{H}_0^1}^2\right], \|g^n\|_{\mathbb{H}_0^1}^2\right)(\delta_1 + \delta_2).
$$

Inequality (21) implies for $x_0^{n_1} \equiv x_0^{n_2} \equiv x_0^n$ that

$$
\mathbb{E}\left[\sup_{t\in[0,T]} \|X_n^{\varepsilon,\delta_1}(t) - X_n^{\varepsilon,\delta_2}(t)\|^2\right] \leq C\left(\mathbb{E}\left[\|x_0^n\|_{\mathbb{H}_0^1}^2\right], \|g^n\|_{\mathbb{H}_0^1}^2\right)(\delta_1 + \delta_2).
$$

Hence for any fixed n, ε there exists a $\{\mathcal{F}_t\}$-adapted process $X_n^\varepsilon \in L^2(\Omega, C([0,T]; \mathbb{L}^2))$, s.t.

(22)
$$
\lim_{\delta \to 0} \mathbb{E}\left[\sup_{t\in[0,T]} \|X_n^{\varepsilon,\delta}(t) - X_n^\varepsilon(t)\|^2\right] \to 0.
$$

For fixed n_1, n_2, ε we get from (21) using (22) by the lower-semicontinuity of norms that

$$
\begin{aligned}
\text{(23)} \qquad \mathbb{E}\left[\sup_{t\in[0,T]} \|X_{n_1}^{\varepsilon}(t) - X_{n_2}^{\varepsilon}\|^2\right] &\leq \liminf_{\delta\to 0} \mathbb{E}\left[\sup_{t\in[0,T]} \|X_{n_1}^{\varepsilon,\delta}(t) - X_{n_2}^{\varepsilon,\delta}\|^2\right] \\
&\leq \frac{1}{2}\mathbb{E}\left[\|x_0^{n_1} - x_0^{n_2}\|^2\right].
\end{aligned}
$$

Since $x_0^{n_1}$, $x_0^{n_2} \to x_0$ for $n_1, n_2 \to \infty$ we deduce from (23) that for any fixed ε there exists an $\{\mathcal{F}_t\}$-adapted process $X^{\varepsilon} \in L^2(\Omega; C([0,T]; \mathbb{L}^2))$ such that

$$
\text{(24)} \qquad \lim_{n\to\infty} \mathbb{E}\left[\sup_{t\in[0,T]} \|X^{\varepsilon}(t) - X_n^{\varepsilon}(t)\|^2\right] \to 0.
$$

In the next step, we show that the limiting process X^{ε} is a SVI solution of (3). We subtract the process

$$
\mathrm{d}Z(t) = -G(t)\,\mathrm{d}t + Z(t)\,\mathrm{d}W(t),
$$

with $Z(t) = z_0$ from (7) and obtain

$$
\mathrm{d}\left(X_n^{\varepsilon,\delta}(t) - Z(t)\right) = \left(-A^{\varepsilon,\delta} X_n^{\varepsilon,\delta}(t) + G(t)\right)\,\mathrm{d}t + \left(X_n^{\varepsilon,\delta}(t) - Z(t)\right)\,\mathrm{d}W(t).
$$

The Itô formula implies

$$
\begin{aligned}
\text{(25)} \qquad &\frac{1}{2}\mathbb{E}\left[\|X_n^{\varepsilon,\delta}(t) - Z(t)\|^2\right] \\
&= \frac{1}{2}\mathbb{E}\left[\|X_n^{\varepsilon,\delta}(0) - z_0\|^2\right] - \mathbb{E}\left[\int_0^t \langle A^{\varepsilon,\delta} X_n^{\varepsilon,\delta}(s), X_n^{\varepsilon,\delta}(s) - Z(s)\rangle\,\mathrm{d}s\right] \\
&\quad + \mathbb{E}\left[\int_0^t \left(G(s), X_n^{\varepsilon,\delta}(s) - Z(s)\right)\,\mathrm{d}s\right] + \mathbb{E}\left[\int_0^t \|X_n^{\varepsilon,\delta}(s) - Z(s)\|^2\,\mathrm{d}s\right].
\end{aligned}
$$

We rewrite the second term on the right-hand side in above inequality as

$$
\begin{aligned}
&\mathbb{E}\left[\int_0^t \langle A^{\varepsilon,\delta} X_n^{\varepsilon,\delta}(s), X_n^{\varepsilon,\delta}(s) - Z(s)\rangle\,\mathrm{d}s\right] \\
&= \mathbb{E}\left[\int_0^t \delta(\nabla X_n^{\varepsilon,\delta}(s), \nabla(X_n^{\varepsilon,\delta}(s) - Z(s)))\,\mathrm{d}s\right] \\
&\quad + \mathbb{E}\left[\int_0^t \left(\frac{\nabla X_n^{\varepsilon,\delta}(s)}{\sqrt{|\nabla X_n^{\varepsilon,\delta}(s)|^2 + \varepsilon^2}}, \nabla\left(X_n^{\varepsilon,\delta}(s) - Z(s)\right)\right)\,\mathrm{d}s\right] \\
&\quad + \mathbb{E}\left[\int_0^t \lambda(X_n^{\varepsilon,\delta}(s) - g^n, X_n^{\varepsilon,\delta}(s) - Z(s))\,\mathrm{d}s\right].
\end{aligned}
$$

The convexity of $\mathcal{J}_\varepsilon$ along with the Cauchy-Schwarz and Young's inequalities imply that

$$\mathbb{E}\left[\int_0^t (\frac{\nabla X_n^{\varepsilon,\delta}(s)}{\sqrt{|\nabla X_n^{\varepsilon,\delta}(s)|^2 + \varepsilon^2}}, \nabla(X_n^{\varepsilon,\delta}(s) - Z(s)))\,\mathrm{d}s\right]$$

$$+ \mathbb{E}\left[\int_0^t \lambda(X_n^{\varepsilon,\delta}(s) - g^n, X_n^{\varepsilon,\delta}(s) - Z(s))\,\mathrm{d}s\right]$$

$$\geq \mathbb{E}\left[\int_0^t \mathcal{J}_\varepsilon(X_n^{\varepsilon,\delta}(s)) - \mathcal{J}_\varepsilon(Z(s))\,\mathrm{d}s\right]$$

$$+ \mathbb{E}\left[\int_0^t \frac{\lambda}{2}\|X_n^{\varepsilon,\delta}(s) - g^n\|^2 - \frac{\lambda}{2}\|Z(s) - g^n\|^2\,\mathrm{d}s\right].$$

By combining two inequalities above with (25) we get

$$\frac{1}{2}\mathbb{E}\left[\|X_n^{\varepsilon,\delta}(t) - Z(t)\|^2\right] + \mathbb{E}\left[\int_0^t \mathcal{J}_\varepsilon(X_n^{\varepsilon,\delta}(s))\,\mathrm{d}s + \frac{\lambda}{2}\|X_n^{\varepsilon,\delta}(s) - g^n\|^2\,\mathrm{d}s\right]$$

$$+ \frac{\delta}{2}\mathbb{E}\left[\int_0^t \|\nabla X_n^{\varepsilon,\delta}(s)\|)\,\mathrm{d}s\right]$$

(26) $$\leq \frac{1}{2}\mathbb{E}\left[\|X_n^{\varepsilon,\delta}(0) - Z(0)\|^2\right] + \mathbb{E}\left[\int_0^t \mathcal{J}_\varepsilon(Z(s))\,\mathrm{d}s + \frac{\lambda}{2}\|Z(s) - g^n\|^2\,\mathrm{d}s\right]$$

$$+ \frac{\delta}{2}\mathbb{E}\left[\int_0^t \|\nabla Z(s)\|)\,\mathrm{d}s\right] + \mathbb{E}\left[\int_0^t \left(G(s), X_n^{\varepsilon,\delta}(s) - Z(s)\right)\,\mathrm{d}s\right]$$

$$+ \mathbb{E}\left[\int_0^t \|X_n^{\varepsilon,\delta}(s) - Z(s)\|^2\,\mathrm{d}s\right].$$

The lower-semicontinuity of $\mathcal{J}_\varepsilon$ in $BV(\mathcal{O})$ with respect to convergence in $\mathbb{L}^1$, cf. [1], and (22), (24) and the strong convergence $g^n \to g$ in $\mathbb{L}^2$ imply that for $\delta \to 0$ and $n \to \infty$ the limiting process $X^\varepsilon \in L^2(\Omega; C([0,T]; \mathbb{L}^2))$ satisfies (18).

To conclude that X^ε is a SVI solution of (3) it remains to show that $X^\varepsilon \in L^2(\Omega; L^1((0,T); BV(\mathcal{O})))$. Setting $G \equiv 0$ in (17) (which implies $Z \equiv 0$ by (17)) yields

(27)
$$\frac{1}{2}\mathbb{E}\left[\|X^\varepsilon(t)\|^2\right] + \mathbb{E}\left[\int_0^t \mathcal{J}_{\varepsilon,\lambda}(X^\varepsilon(s))\,\mathrm{d}s\right]$$

$$\leq \frac{1}{2}\mathbb{E}\left[\|x_0\|^2\right] + \mathbb{E}\left[\int_0^t \mathcal{J}_{\varepsilon,\lambda}(0)\,\mathrm{d}s\right] + \mathbb{E}\left[\int_0^t \|X^\varepsilon(s)\|^2\,\mathrm{d}s\right].$$

On noting that (cf. Definition 2.3 or [7, proof of Theorem 1.3])

$$\mathcal{J}_{\varepsilon,\lambda}(X^\varepsilon) \geq \mathcal{J}_\lambda(X^\varepsilon),$$

and $\mathcal{J}_{\varepsilon,\lambda}(0) = \varepsilon|\mathcal{O}| + \frac{\lambda}{2}\|g\|^2$, we deduce from (27) that

$$\frac{1}{2}\mathbb{E}\left[\|X^\varepsilon(t)\|^2\right] + \mathbb{E}\left[\int_0^t \mathcal{J}_\lambda(X^\varepsilon(s))\,\mathrm{d}s\right]$$
$$\leq \frac{1}{2}\mathbb{E}\left[\|x_0\|^2\right] + \mathbb{E}\left[\int_0^t \varepsilon\left(|\mathcal{O}| + \frac{\lambda}{2}\|g\|^2\right)\mathrm{d}s\right] + \mathbb{E}\left[\int_0^t \|X^\varepsilon(s)\|^2\,\mathrm{d}s\right].$$

Hence, by the Tonelli and Gronwall lemmas it follows that

$$\frac{1}{2}\mathbb{E}\left[\|X^\varepsilon(t)\|^2\right] + \mathbb{E}\left[\int_0^t \mathcal{J}_\lambda(X^\varepsilon(s))\,\mathrm{d}s\right]$$
(28)
$$\leq C_T \exp(T)\left(\mathbb{E}\left[\|x_0\|^2\right] + |\mathcal{O}| + \lambda\|g\|^2\right).$$

Hence $X^\varepsilon \in L^2(\Omega; C([0,T];\mathbb{L}^2)) \cap L^2(\Omega; L^1((0,T); BV(\mathcal{O})))$ is a SVI solution of (3) for $\varepsilon \in (0,1]$.

In the next step we show the uniqueness of the SVI solution. Let $X_1^\varepsilon, X_2^\varepsilon$ be two SVI solutions to (3) for a fixed $\varepsilon \in (0,1]$ with initial values $x_0 \equiv x_0^1, x_0^2$ and $g \equiv g^1, g^2$, respectively. Let $\{x_0^{2,n}\}_{n\in\mathbb{N}} \subset L^2(\Omega, \mathcal{F}_0; \mathbb{H}_0^1)$ be a sequence, s.t. $x_0^{2,n} \to x_0^2$ in $L^2(\Omega, \mathcal{F}_0; \mathbb{L}^2)$ and $\{g^{2,n}\}_{n\in\mathbb{N}} \subset \mathbb{H}_0^1$ be a sequence, s.t. $g_0^{2,n} \to g^2$ in $\mathbb{L}^2$ for $n \to \infty$ and let $\{X_{2,n}^{\varepsilon,\delta}\}_{n\in\mathbb{N},\delta>0}$ be a sequence of variational solutions of (7) (for fixed $\varepsilon > 0$) with $x_0 \equiv x_0^{2,n}$, $g \equiv g^{2,n}$. We note that the first part of the proof implies that $X_{2,n}^{\varepsilon,\delta} \to X_2^\varepsilon$ in $L^2(\Omega; C([0,T];\mathbb{L}^2))$ for $\delta \to 0$, $n \to \infty$. We set $Z = X_{2,n}^{\varepsilon,\delta}, G = A^{\varepsilon,\delta}(X_{2,n}^{\varepsilon,\delta})$ in (18) and observe that

$$\frac{1}{2}\mathbb{E}\left[\|X_1^\varepsilon(t) - X_{2,n}^{\varepsilon,\delta}(t)\|^2\right] + \mathbb{E}\left[\int_0^t \mathcal{J}_{\varepsilon,\lambda}(X_1^\varepsilon(s))\,\mathrm{d}s\right]$$
$$\leq \frac{1}{2}\mathbb{E}\left[\|x^1 - x_0^{2,n}\|^2\right] + \mathbb{E}\left[\int_0^t \mathcal{J}_{\varepsilon,\lambda}(X_{2,n}^{\varepsilon,\delta}(s))\,\mathrm{d}s\right]$$
(29)
$$+ \mathbb{E}\left[\delta \int_0^t \left(\nabla X_1^\varepsilon(s) - \nabla X_{2,n}^{\varepsilon,\delta}(s), \nabla X_{2,n}^{\varepsilon,\delta}(s)\right)\mathrm{d}s\right]$$
$$+ \mathbb{E}\left[\int_0^t \left(\nabla X_1^\varepsilon(s) - \nabla X_{2,n}^{\varepsilon,\delta}(s), \frac{\nabla X_{2,n}^{\varepsilon,\delta}(s)}{\sqrt{|\nabla X_{2,n}^{\varepsilon,\delta}(s)|^2 + \varepsilon^2}}\right)\mathrm{d}s\right]$$
$$+ \mathbb{E}\left[\int_0^t \lambda\left(X_1^\varepsilon(s) - X_{2,n}^{\varepsilon,\delta}(s), X_{2,n}^{\varepsilon,\delta}(s) - g^{2,n}\right)\mathrm{d}s\right]$$
$$+ \mathbb{E}\left[\int_0^t \|X_1^\varepsilon(s) - X_{2,n}^{\varepsilon,\delta}(s)\|^2\,\mathrm{d}s\right]$$
$$:= I + II + III + IV + V + VI.$$

The term III is estimated using Young's inequality as

$$III = \delta \mathbb{E}\left[\int_0^t \left(\nabla X_1^\varepsilon(s) - \nabla X_{2,n}^{\varepsilon,\delta}(s), \nabla X_{2,n}^{\varepsilon,\delta}(s))\right) \, \mathrm{d}s\right]$$

$$\leq C \mathbb{E}\left[\int_0^t \delta^{\frac{2}{3}} \|X_1^\varepsilon(s) - X_{2,n}^{\varepsilon,\delta}(s)\|^2 + \delta^{\frac{4}{3}} \|\Delta X_{2,n}^{\varepsilon,\delta}(s))\|^2 \, \mathrm{d}s\right].$$

By the convexity (12) we estimate

$$IV = \mathbb{E}\left[\int_0^t \left(\nabla X_1^\varepsilon(s) - \nabla X_{2,n}^{\varepsilon,\delta}(s), \frac{\nabla X_{2,n}^{\varepsilon,\delta}(s)}{\sqrt{|\nabla X_{2,n}^{\varepsilon,\delta}(s)|^2 + \varepsilon^2}}\right) \, \mathrm{d}s\right]$$

$$\leq \mathbb{E}\left[\int_0^t \sqrt{|\nabla X_1^\varepsilon(s)|^2 + \varepsilon^2} - \sqrt{|\nabla X_{2,n}^{\varepsilon,\delta}(s)|^2 + \varepsilon^2} \, \mathrm{d}s\right].$$

Next, we obtain

$$V = \lambda \mathbb{E}\left[\int_0^t \left(X_1^\varepsilon(s) - X_{2,n}^{\varepsilon,\delta}(s), X_{2,n}^{\varepsilon,\delta}(s) - g^{2,n}\right) \, \mathrm{d}s\right]$$

$$\leq \frac{\lambda}{2} \mathbb{E}\left[\int_0^t \|X_1^\varepsilon(s) - g^{2,n}\|^2 - \|X_{2,n}^{\varepsilon,\delta}(s) - g^{2,n}\|^2 \, \mathrm{d}s\right].$$

After substituting III-V into (29) we arrive at

$$\frac{1}{2}\mathbb{E}\left[\|X_1^\varepsilon(t) - X_{2,n}^{\varepsilon,\delta}(t)\|^2\right] + \frac{\lambda}{2}\mathbb{E}\left[\int_0^t \|X_1^\varepsilon(s) - g^1\|^2 \, \mathrm{d}s\right]$$

$$\leq \frac{1}{2}\mathbb{E}\left[\|x_0^1 - x_0^{2,n}\|^2\right] + \frac{\lambda}{2}\mathbb{E}\left[\int_0^t \|X_{2,n}^{\varepsilon,\delta}(s) - g^1\|^2 \, \mathrm{d}s\right]$$

$$+ C\mathbb{E}\left[\int_0^t \delta^{\frac{2}{3}} \|X_1^\varepsilon(s) - X_{2,n}^{\varepsilon,\delta}(s)\|^2 + \delta^{\frac{4}{3}} \|\Delta X_{2,n}^{\varepsilon,\delta}(s))\|^2 \, \mathrm{d}s\right]$$

(30)
$$\frac{\lambda}{2}\mathbb{E}\left[\int_0^t \|X_1^\varepsilon(s) - g^{2,n}\|^2 \, \mathrm{d}s\right] - \frac{\lambda}{2}\mathbb{E}\left[\int_0^t \|X_{2,n}^{\varepsilon,\delta}(s) - g^{2,n}\|^2 \, \mathrm{d}s\right]$$

$$+ \mathbb{E}\left[\int_0^t \|X_1^\varepsilon(s) - X_{2,n}^{\varepsilon,\delta}(s)\|^2 \, \mathrm{d}s\right].$$

The convergences (22), (24) imply the convergence $X_{2,n}^{\varepsilon,\delta} \to X_2^\varepsilon$ in $L^2(\Omega; C([0,T]; \mathbb{L}^2))$ for $\delta \to 0$, $n \to \infty$. We note that for $\delta \to 0$ the fourth term on the right-hand side of (30) vanishes due to Lemma 3.2. Hence, by taking the limits for $\delta \to 0$, $n \to \infty$ in (30), using the strong convergence $g^{2,n} \to g^2$ in $\mathbb{L}^2$ for $n \to \infty$, the lower-semicontinuity of norms and

(22), (24) we obtain

$$
\begin{aligned}
\mathbb{E}\left[\|X_1^\varepsilon(t) - X_2^\varepsilon(t)\|^2\right] &\leq C\mathbb{E}\left[\|x_0^1 - x_0^2\|^2\right] \\
&+ \frac{\lambda}{2}\mathbb{E}\left[\int_0^t \|X_1^\varepsilon(s) - g^2\|^2 + \|X_2^\varepsilon(s) - g^1\|^2 - \|X_1^\varepsilon(s) - g^1\|^2 - \|X_2^\varepsilon(s) - g^2\|^2 \, \mathrm{d}s\right] \\
&+ \mathbb{E}\left[\int_0^t \|X_1^\varepsilon(s) - X_2^\varepsilon(s)\|^2 \, \mathrm{d}s\right] \\
&\leq C\left(\mathbb{E}\left[\|x_0^1 - x_0^2\|^2\right] + \mathbb{E}\left[\int_0^t \|X_1^\varepsilon(s) - X_2^\varepsilon(s)\|^2 \, \mathrm{d}s\right] + \|g^1 - g^2\|^2\right),
\end{aligned}
$$

for all $t \in [0, T]$. After applying the Tonelli and Gronwall lemmas we obtain (20). $\square$

Our second main theorem establishes existence and uniqueness of a SVI solution to (1) in the sense of Definition 3.1. The solution is obtained as a limit of solutions of the regularized gradient flow (3) for $\varepsilon \to 0$.

Theorem 3.2. *Let $0 < T < \infty$ and $x_0 \in L^2(\Omega, \mathcal{F}_0; \mathbb{L}^2)$, $g \in \mathbb{L}^2$ be fixed. Let $\{X^\varepsilon\}_{\varepsilon>0}$ be the SVI solutions of (3) for $\varepsilon \in (0, 1]$. Then X^ε converges to the unique SVI variational solution X of (1) in $L^2(\Omega; C([0, T]; \mathbb{L}^2))$ for $\varepsilon \to 0$, i.e., there holds*

$$
\tag{31}
\lim_{\varepsilon \to 0} \mathbb{E}\left[\sup_{t \in [0,T]} \|X^\varepsilon(t) - X(t)\|^2\right] = 0.
$$

Furthermore, the following estimate holds

$$
\tag{32}
\mathbb{E}\left[\|X_1(t) - X_2(t)\|^2\right] \leq C\left(\mathbb{E}\left[\|x_0^1 - x_0^2\|^2\right] + \|g^1 - g^2\|^2\right) \quad \text{for all } t \in [0, T],
$$

where X_1 and X_2 are SVI solutions of (1) with $x_0 \equiv x_0^1$, $g \equiv g^1$ and $x_0 \equiv x_0^2$, $g \equiv g^2$, respectively.

Proof of Theorem 3.2.
We consider $\mathbb{L}^2$-approximating sequences $\{x_0^n\}_{n\in\mathbb{N}} \subset L^2(\Omega, \mathcal{F}_0; \mathbb{H}_0^1)$ and $\{g^n\}_{n\in\mathbb{N}} \subset \mathbb{H}_0^1$ of the initial condition $x_0 \in L^2(\Omega, \mathcal{F}_0; \mathbb{L}^2)$ and $g \in \mathbb{L}^2$, respectively. For $n \in \mathbb{N}$, $\delta > 0$ we denote by $X_n^{\varepsilon_1,\delta}, X_n^{\varepsilon_2,\delta}$ the variational solutions of (7) with $\varepsilon \equiv \varepsilon_1$, $\varepsilon \equiv \varepsilon_2$, respectively. By

Itô's formula the difference satisfies

$$
\frac{1}{2}\|X_n^{\varepsilon_1,\delta}(t) - X_n^{\varepsilon_2,\delta}(t)\|^2
$$

$$
= -\delta \int_0^t \|\nabla(X_n^{\varepsilon_1,\delta}(s) - X_n^{\varepsilon_2,\delta}(s))\|^2 \, \mathrm{d}s
$$

(33)
$$
- \int_0^t \left(\frac{\nabla X_n^{\varepsilon_1,\delta}(s)}{\sqrt{|\nabla X_n^{\varepsilon_1,\delta}(s)|^2 + \varepsilon_1^2}} - \frac{\nabla X_n^{\varepsilon_2,\delta}(s)}{\sqrt{|\nabla X_n^{\varepsilon_2,\delta}(s)|^2 + \varepsilon_2^2}}, \nabla(X_n^{\varepsilon_1,\delta}(s) - X_n^{\varepsilon_2,\delta}(s)) \right) \, \mathrm{d}s
$$

$$
- \lambda \int_0^t \|(X_n^{\varepsilon_1,\delta}(s) - X_n^{\varepsilon_2,\delta}(s))\|^2 \, \mathrm{d}s + \int_0^t \|X_n^{\varepsilon_1,\delta}(s) - X_n^{\varepsilon_2,\delta}(s)\|^2 \, \mathrm{d}W(s)
$$

$$
+ \int_0^t \|X_n^{\varepsilon_1,\delta}(s) - X_n^{\varepsilon_2,\delta}(s)\|^2 \, \mathrm{d}s.
$$

We estimate the second term on the right-hand side of (33) using the convexity (12)

$$
\left(\frac{\nabla X_n^{\varepsilon_1,\delta}(s)}{\sqrt{|\nabla X_n^{\varepsilon_1,\delta}(s)|^2 + \varepsilon_1^2}} - \frac{\nabla X_n^{\varepsilon_2,\delta}(s)}{\sqrt{|\nabla X_n^{\varepsilon_2,\delta}(s)|^2 + \varepsilon_2^2}}, \nabla(X_n^{\varepsilon_1,\delta}(s) - X_n^{\varepsilon_2,\delta}(s)) \right)
$$

$$
= \left(\frac{\nabla X_n^{\varepsilon_1,\delta}(s)}{\sqrt{|\nabla X_n^{\varepsilon_1,\delta}(s)|^2 + \varepsilon_1^2}}, \nabla(X_n^{\varepsilon_1,\delta}(s) - X_n^{\varepsilon_2,\delta}(s)) \right)
$$

(34)
$$
+ \left(\frac{\nabla X_n^{\varepsilon_2,\delta}(s)}{\sqrt{|\nabla X_n^{\varepsilon_2,\delta}(s)|^2 + \varepsilon_2^2}}, \nabla(X_n^{\varepsilon_2,\delta}(s) - X_n^{\varepsilon_1,\delta}(s)) \right)
$$

$$
\geq \int_{\mathcal{O}} \sqrt{|\nabla X_n^{\varepsilon_1,\delta}|^2 + \varepsilon_1^2} - \sqrt{|\nabla X_n^{\varepsilon_2,\delta}|^2 + \varepsilon_1^2} \, \mathrm{d}x
$$

$$
+ \int_{\mathcal{O}} \sqrt{|\nabla X_n^{\varepsilon_2,\delta}|^2 + \varepsilon_2^2} - \sqrt{|\nabla X_n^{\varepsilon_1,\delta}|^2 + \varepsilon_2^2} \, \mathrm{d}x.
$$

Next, we observe that

$$
\int_{\mathcal{O}} \left(\sqrt{|\nabla X_n^{\varepsilon_1,\delta}|^2 + \varepsilon_1^2} - \sqrt{|\nabla X_n^{\varepsilon_1,\delta}|^2 + \varepsilon_2^2} \right) \mathrm{d}x
$$

$$
= \int_{\mathcal{O}} \frac{\left(\sqrt{|\nabla X_n^{\varepsilon_1,\delta}|^2 + \varepsilon_1^2} - \sqrt{|\nabla X_n^{\varepsilon_1,\delta}|^2 + \varepsilon_2^2} \right) \left(\sqrt{|\nabla X_n^{\varepsilon_1,\delta}|^2 + \varepsilon_1^2} + \sqrt{|\nabla X_n^{\varepsilon_1,\delta}|^2 + \varepsilon_2^2} \right)}{\sqrt{|\nabla X_n^{\varepsilon_1,\delta}|^2 + \varepsilon_1^2} + \sqrt{|\nabla X_n^{\varepsilon_1,\delta}|^2 + \varepsilon_2^2}} \, \mathrm{d}x
$$

$$
= \int_{\mathcal{O}} \frac{|\nabla X_n^{\varepsilon_1,\delta}|^2 + \varepsilon_1^2 - |\nabla X_n^{\varepsilon_1,\delta}|^2 - \varepsilon_2^2}{\sqrt{|\nabla X_n^{\varepsilon_1,\delta}|^2 + \varepsilon_1^2} + \sqrt{|\nabla X_n^{\varepsilon_1,\delta}|^2 + \varepsilon_2^2}} \, \mathrm{d}x
$$

$$
= \int_{\mathcal{O}} \frac{(\varepsilon_1 + \varepsilon_2)(\varepsilon_1 - \varepsilon_2)}{\sqrt{|\nabla X_n^{\varepsilon_1,\delta}|^2 + \varepsilon_1^2} + \sqrt{|\nabla X_n^{\varepsilon_1,\delta}|^2 + \varepsilon_2^2}} \, \mathrm{d}x
$$

$$
\leq \int_{\mathcal{O}} |\varepsilon_1 - \varepsilon_2| \left(\frac{\varepsilon_1}{\sqrt{|\nabla X_n^{\varepsilon_1,\delta}|^2 + \varepsilon_1^2}} + \frac{\varepsilon_2}{\sqrt{|\nabla X_n^{\varepsilon_1,\delta}|^2 + \varepsilon_2^2}} \right) \mathrm{d}x \leq C(\varepsilon_1 + \varepsilon_2).
$$

Using the inequality above, we get

$$
\int_{\mathcal{O}} \sqrt{|\nabla X_n^{\varepsilon_1,\delta}|^2 + \varepsilon_1^2} - \sqrt{|\nabla X_n^{\varepsilon_2,\delta}|^2 + \varepsilon_2^2} \, \mathrm{d}x + \int_{\mathcal{O}} \sqrt{|\nabla X_n^{\varepsilon_2,\delta}|^2 + \varepsilon_2^2} - \sqrt{|\nabla X_n^{\varepsilon_1,\delta}|^2 + \varepsilon_2^2} \, \mathrm{d}x
$$

$$
\geq - \left| \int_{\mathcal{O}} \sqrt{|\nabla X_n^{\varepsilon_1,\delta}|^2 + \varepsilon_1^2} - \sqrt{|\nabla X_n^{\varepsilon_1,\delta}|^2 + \varepsilon_2^2} \, \mathrm{d}x \right|
$$

$$
- \left| \int_{\mathcal{O}} \sqrt{|\nabla X_n^{\varepsilon_2,\delta}|^2 + \varepsilon_1^2} - \sqrt{|\nabla X_n^{\varepsilon_2,\delta}|^2 + \varepsilon_2^2} \, \mathrm{d}x \right|
$$

$$
\geq - C(\varepsilon_1 + \varepsilon_2).
$$

Substituting (34) along with the last inequality into (33) yields

$$
\begin{aligned}
\frac{1}{2} \| X_n^{\varepsilon_1,\delta}(t) - X_n^{\varepsilon_2,\delta}(t) \|^2 \leq\, & C(\varepsilon_1 + \varepsilon_2) \\
& + \int_0^t \| X_n^{\varepsilon_1,\delta}(s) - X_n^{\varepsilon_2,\delta}(s) \|^2 \, \mathrm{d}W(s) \\
& + \int_0^t \| X_n^{\varepsilon_1,\delta}(s) - X_n^{\varepsilon_2,\delta}(s) \|^2 \, \mathrm{d}s.
\end{aligned}
$$
(35)

After using the Burkholder-Davis-Gundy inequality for $p = 1$, the Tonelli and Gronwall lemmas we obtain that

$$(36) \qquad \mathbb{E}\left[\sup_{t\in[0,T]} \|X_n^{\varepsilon_1,\delta}(t) - X_n^{\varepsilon_2,\delta}(t)\|^2\right] \leq C(\varepsilon_1 + \varepsilon_2).$$

We take the limit for $\delta \to 0$ in (36) for fixed n and $\varepsilon_1, \varepsilon_2$, and obtain using (22) by the lower-semicontinuity of norms that

$$\mathbb{E}\left[\sup_{t\in[0,T]} \|X_n^{\varepsilon_1}(t) - X_n^{\varepsilon_2}(t)\|^2\right] \leq \liminf_{\delta\to 0} \mathbb{E}\left[\sup_{t\in[0,T]} \|X_n^{\varepsilon_1,\delta}(t) - X_n^{\varepsilon_2,\delta}(t)\|^2\right]$$

$$(37) \qquad\qquad\qquad\qquad \leq C(\varepsilon_1 + \varepsilon_2).$$

Hence, by (24) and the lower-semicontinuity of norms, after taking the limit $n \to \infty$ in (37) for fixed $\varepsilon_1, \varepsilon_2$ we get

$$(38) \qquad \mathbb{E}\left[\sup_{t\in[0,T]} \|X^{\varepsilon_1}(t) - X^{\varepsilon_2}(t)\|^2\right] \leq \liminf_{n\to\infty} \mathbb{E}\left[\sup_{t\in[0,T]} \|X_n^{\varepsilon_1}(t) - X_n^{\varepsilon_2}(t)\|^2\right]$$

$$\leq C(\varepsilon_1 + \varepsilon_2).$$

The above inequality implies that $\{X^\varepsilon\}_{\varepsilon>0}$ is a Cauchy Sequence in ε. Consequently there exists a unique $\{\mathcal{F}_t\}$-adapted process $X \in L^2(\Omega; C([0,T]; \mathbb{L}^2))$ with $X(0) = x_0$ such that

$$(39) \qquad \lim_{\varepsilon\to 0} \mathbb{E}\left[\sup_{t\in[0,T]} \|X^\varepsilon(t) - X(t)\|^2\right] = 0.$$

This concludes the proof of (31).

Next, we show that the limiting process X is the SVI solution of (1), i.e., we show that (19) holds. We note that (28) implies that

$$(40) \qquad \sup_{\varepsilon\in(0,1]} \mathbb{E}\left[\int_0^t \mathcal{J}_\lambda(X^\varepsilon(s)\,\mathrm{d}s\right] \leq C.$$

Hence using (39), (40) we get by Fatou's lemma and [2, Prop. 10.1.1] that

$$\liminf_{\varepsilon\to 0} \mathbb{E}\left[\int_0^t \mathcal{J}_\lambda(X^\varepsilon(s))\,\mathrm{d}s\right] \geq \mathbb{E}\left[\int_0^t \mathcal{J}_\lambda(X(s))\,\mathrm{d}s\right].$$

By Theorem 3.1 we know that X^ε satisfies (18) for any $\varepsilon \in (0,1]$. By taking the limit for $\varepsilon \to 0$ in (18), using the above inequality and (39) it follows that X satisfies (19). Finally, inequality (32) follows after taking the limit for $\varepsilon \to 0$ in (39), by (20) and the lower semicontinuity of norms. $\qquad\square$

4. NUMERICAL APPROXIMATION

We will construct numerical scheme for the STVF equation (1) via a time-implicit full discretization of the regularized STVF equation (3). We fix $N \in \mathbb{N}$ and define the time-step $\tau := T/N$, set $t_i := i\tau$ for $i = 0, \ldots, N$ and denote the discrete Wiener increments as $\Delta_i W := W(t_i) - W(t_{i-1})$. For the spatial discretization we employ a spectral Galerkin Scheme: for a fixed m we consider a finite-dimensional subspace $\mathbb{V}_m \subset \mathbb{H}_0^1$ which consists of the first m eigenfunctions of the Dirichlet Laplacian and set set $x_0^m := \mathcal{P}_m x_0$, $g^m := \mathcal{P}_m g$, where $\mathcal{P}_m$ is the $\mathbb{L}^2$-projection onto $\mathbb{V}_m$.

We note that the convergence analysis below directly covers the case of $\mathbb{H}^2$-conforming finite element based spatial discretization, i.e., $\mathbb{V}_m$ may be replaced by a finite element space $\mathbb{V}_h \subset \mathbb{H}^2 \cap \mathbb{H}_0^1$ and $\mathcal{P}_m$ by the corresponding orthogonal $\mathbb{L}^2$-projection $\mathcal{P}_h : \mathbb{L}^2 \to \mathbb{V}_h$; this includes, for instance, the $C^1(\mathcal{O})$-conforming Argyris element, cf. [5].

The fully discrete approximation of (3) is defined as follows: fix $N, m \in \mathbb{N}$ set $X_\varepsilon^0 = x_0^m \in \mathbb{V}_m$ and determine $X_\varepsilon^i \in \mathbb{V}_m$, $i = 1, \ldots, N$ as the solution of

$$(41) \qquad \left(X_\varepsilon^i, v_h\right) = \left(X_\varepsilon^{i-1}, v_h\right) - \tau \left(\frac{\nabla X_\varepsilon^i}{\sqrt{|\nabla X_\varepsilon^i|^2 + \varepsilon^2}}, \nabla v_h\right)$$

$$- \tau\lambda\left(X_\varepsilon^i - g^m, v_h\right) + \left(X_\varepsilon^{i-1}, v_h\right)\Delta_i W \qquad \forall v_h \in \mathbb{V}_m.$$

Remark 4.1. *Throughout this section, to simplify the notation, whenever we take the limit w.r.t. to the time-step τ we implicitly assume that the limit is simultaneously taken with respect to the spatial discretization parameter m, i.e., we will denote by $\lim_{\tau \to 0} \equiv \lim_{\tau \to 0, m \to \infty}$. Nevertheless we stress, that due to the implicit time-discretization in (41) there is no explicit coupling between the time-step τ and the spatial discretization m.*

To show convergence of the solution of the numerical scheme (41) we need to consider a fully discrete approximation of the corresponding regularized problem (7). Given $x_0 \in L^2(\Omega, \mathcal{F}_0; \mathbb{L}^2)$, $g \in \mathbb{L}^2$ and $n \in \mathbb{N}$ we choose $x_0^n := \mathcal{P}_n x_0 \in \mathbb{V}_n$, $g^n := \mathcal{P}_n g \in \mathbb{V}_n$ in (7). Since $\mathbb{V}_n \subset \mathbb{H}_0^1$ the sequences $\{x_0^n\}_{n \in \mathbb{N}} \subset L^2(\Omega, \mathcal{F}_0; \mathbb{H}_0^1)$, $\{g_n\}_{n \in \mathbb{N}} \in \mathbb{H}_0^1$ constitute $\mathbb{H}_0^1$-approximating sequences of $x_0 \in L^2(\Omega, \mathcal{F}_0; \mathbb{L}^2)$, $g \in \mathbb{L}^2$, respectively. To simplify the notation, we assume without the loss of generality that $m \geq n$, which implies that $\mathcal{P}_m x_0^n \equiv x_0^n$ and $\mathcal{P}_m g^n \equiv g^n$. The numerical approximation of (7) for fixed $n \in \mathbb{N}$ is then defined as follows: fix $m, N \in N$, $m \geq n$ set $X_{\varepsilon,\delta,n}^0 = x_0^n$ and determine $X_{\varepsilon,\delta,n}^i \in \mathbb{V}_m$, $i = 1, \ldots, N$ as the solution of

$$\left(X_{\varepsilon,\delta,n}^i, v_h\right) = \left(X_{\varepsilon,\delta,n}^{i-1}, v_h\right) - \tau\delta\left(\nabla X_{\varepsilon,\delta,n}^i, \nabla v_h\right) - \tau\left(\frac{\nabla X_{\varepsilon,\delta,n}^i}{\sqrt{|\nabla X_{\varepsilon,\delta,n}^i|^2 + \varepsilon^2}}, \nabla v_h\right)$$

$$(42) \qquad - \tau\lambda\left(X_{\varepsilon,\delta,n}^i - g^n, v_h\right) + \left(X_{\varepsilon,\delta,n}^{i-1}, v_h\right)\Delta_i W \qquad \forall v_h \in \mathbb{V}_m.$$

The next lemma, cf. [15, Lemma II.1.4] is used to show $\mathbb{P}$-a.s. existence of discrete solutions $\{X_\varepsilon^i\}_{i=1}^N$, $\{X_{\varepsilon,\delta,n}^i\}_{i=1}^N$ of numerical schemes (41), (42), respectively.

Lemma 4.1. *Let $h : \mathbb{R}^m \to \mathbb{R}^m$ be continuous. If there is $R > 0$ such that $h(v)v \geq 0$ whenever $\|v\|_{\mathbb{R}^m} = R$ then there exist $\bar{v}$ satisfying $\|\bar{v}\|_{\mathbb{R}^m} \leq R$ and $h(\bar{v}) = 0$.*

In order to show $\{\mathcal{F}_{t_i}\}_{i=1}^N$-measurability of the random variables $\{X_\varepsilon^i\}_{i=1}^N$, $\{X_{\varepsilon,\delta,n}^i\}_{i=1}^N$ we make use of the following lemma, cf. [9, 6].

Lemma 4.2. *Let (S, Σ) be a measure space. Let $f : S \times \mathbb{R}^m \to \mathbb{R}^m$ be a function that is Σ-measurable in its first argument for every $x \in \mathbb{R}^m$, that is continuous in its second argument for every $\alpha \in S$ and moreover such that for every $\alpha \in S$ the equation $f(\alpha, x) = 0$ has an unique solution $x = g(\alpha)$. Then $g : S \to \mathbb{R}^m$ is Σ-measurable.*

Below we show the existence, uniqueness and measurability of numerical solutions of (41), (42). We state the result for the scheme (42) only, since the proof also holds for $\delta = 0$ (i.e. for (41)) without any modifications.

Lemma 4.3. *Let $x_0 \in L^2(\Omega, \mathcal{F}_0; \mathbb{L}^2)$, $g \in \mathbb{L}^2$ and let $m, n, N \in \mathbb{N}$ be fixed. The for any $\delta \geq 0$, $\varepsilon > 0$, $i = 1, \ldots, N$, there exist $\mathcal{F}_{t_i}$-measurable $\mathbb{P}$-a.s. unique random variables $X_{\varepsilon,\delta,n}^i \in \mathbb{V}_m$ which solves (42).*

Proof of Lemma 4.3. Assume that the $\mathbb{V}_m$-valued random variables $X_{\varepsilon,\delta,n}^0, \ldots, X_{\varepsilon,\delta,n}^{i-1}$ satisfy (42) and that $X_{\varepsilon,\delta,n}^k$ is $\mathcal{F}_{t_k}$-measurable for $k = 1, \ldots, i - 1$. We show that there is a $\mathcal{F}_{t_i}$ measurable random variable $X_{\varepsilon,\delta,n}^i$, that satisfies (41). Let $\{\varphi_\ell\}_{\ell=1}^m$ be the basis of $\mathbb{V}_m$. We identify every $v \in \mathbb{V}_m$ with a vector $\bar{v} \in \mathbb{R}^m$ with $v = \sum_{\ell=1}^m \bar{v}_\ell \varphi_\ell$ and define a norm on $\mathbb{R}^m$ as $\|\bar{v}\|_{\mathbb{R}^m} := \|v\|_{\mathbb{H}_0^1}$. For an arbitrary $\omega \in \Omega$ we represent $X_\omega \in \mathbb{V}_m$ as a vector $\bar{X}_\omega \in \mathbb{R}^m$ and define a function $h : \Omega \times \mathbb{R}^m \to \mathbb{R}^m$ component-wise for $\ell = 1, \ldots, m$ as

$$h(\omega, \bar{X}_\omega)_\ell := (X_\omega - X_{\varepsilon,\delta,n}^{i-1}(\omega), \varphi_\ell) + \tau\delta(\nabla X_\omega, \nabla\varphi_\ell) + \tau\left(\frac{\nabla X_\omega}{\sqrt{|\nabla X_\omega|^2 + \varepsilon^2}}, \nabla\varphi_\ell\right)$$
$$+ \tau\lambda(X_\omega - g^n, \varphi_\ell) - (X_{\varepsilon,\delta,n}^{i-1}(\omega), \varphi_\ell)\Delta_i W(\omega).$$

We show, that for each $\omega \in \Omega$ there exists an $\bar{X}_\omega$ such that $h(\omega, \bar{X}_\omega) = 0$. We note the following inequality

$$h(\omega, \bar{X}_\omega) \cdot \bar{X}_\omega = (X_\omega - X_{\varepsilon,\delta,n}^{i-1}(\omega), X) + \tau\delta\|\Delta X_\omega\|^2 + \tau\left(\frac{\nabla X_\omega}{\sqrt{|\nabla X_\omega|^2 + \varepsilon^2}}, \nabla X_\omega\right)$$
$$+ \tau\lambda(X_\omega - g^n, X_\omega) - (X_{\varepsilon,\delta,n}^{i-1}(\omega), X_\omega)\Delta_i W(\omega)$$
$$\geq \|X_\omega\|^2 - (X_{\varepsilon,\delta,n}^{i-1}(\omega), X_\omega) + \tau\left(\frac{\nabla X_\omega}{\sqrt{|\nabla X_\omega|^2 + \varepsilon^2}}, \nabla X_\omega\right)$$
$$- (X_{\varepsilon,\delta,n}^{i-1}(\omega), X_\omega)\Delta_i W + \tau\lambda\|X_\omega\| - \tau\lambda(g^n, X_\omega)$$
$$\geq \|X_\omega\| \left(\|X_\omega\| - \|X_{\varepsilon,\delta,n}^{i-1}(\omega)\| - \|X_{\varepsilon,\delta,n}^{i-1}(\omega)\||\Delta_i W(\omega)| - \|g^n\|\right).$$

On choosing $\|X_\omega\| = R_\omega$ large enough, the existence of $X_{\varepsilon,\delta,n}^i(\omega) \in \mathbb{V}_m$ for each $\omega \in \Omega$ then follows by Lemma 4.1, since $h(\omega, \cdot)$ is continuous by the demicontinuity of the operator $A^{\varepsilon,\delta}$, which follows from hemicontinuity and and monotonicty of $A^{\varepsilon,\delta}$ for $\delta \geq 0$, $\varepsilon > 0$, see [12, Remark 4.1.1]. The $\mathcal{F}_{t_i}$-measurabilty follows by Lemma 4.2 for unique $X_{\varepsilon,\delta,n}^i$.

Hence, it remains to show that $X^i_{\varepsilon,\delta,n}$ is $\mathbb{P}$-a.s. unique. Assume there are two different solution X_1, X_2, s.t. $h(\omega, \overline{X}_1(\omega)) = 0 = h(\omega, \overline{X}_2(\omega))$ for $\omega \in \Omega$. Then by the convexity (12) we observe that

$$
\begin{aligned}
0 =& \big(h(\omega, \overline{X}_1(\omega)) - h(\omega, \overline{X}_2(\omega))\big) \cdot \big(\overline{X}_1(\omega) - \overline{X}_2(\omega)\big) \\
=& (1 + \tau\lambda)\|X_1(\omega) - X_2(\omega)\|^2 + \tau\delta\|\nabla(X_1 - X_2)(\omega)\|^2 \\
& + \tau \left(\frac{\nabla X_1(\omega)}{\sqrt{|\nabla X_1(\omega)|^2 + \varepsilon^2}} - \frac{\nabla X_2(\omega)}{\sqrt{|\nabla X_2(\omega)|^2 + \varepsilon^2}}, \nabla X_1(\omega) - \nabla X_2(\omega) \right) \\
\geq& (1 + \tau\lambda)\|X_1(\omega) - X_2(\omega)\|^2 + \tau\delta\|\nabla(X_1 - X_2)(\omega)\|^2 .
\end{aligned}
$$

Hence $X_1 \equiv X_2$ $\mathbb{P}$-a.s. $\qquad\square$

In the next lemma we state the stability properties of the numerical solution of the scheme (42) which are discrete analogues of estimates in Lemma 3.1 and Lemma 3.2.

Lemma 4.4. *Let $x_0 \in L^2(\Omega, \mathcal{F}_0; \mathbb{L}^2)$ and $g \in \mathbb{L}^2$ be given. Then there exists a constant $C \equiv C(\mathbb{E}\left[\|x_0\|_{\mathbb{L}^2}\right], \|g\|_{\mathbb{L}^2}) > 0$ such that for any $m \geq n \in \mathbb{N}$, $\tau > 0$ the solution of scheme (42) satisfies*

$$
\begin{aligned}
\sup_{i=1,\dots,N} \mathbb{E}\left[\|X^i_{\varepsilon,\delta,n}\|^2\right] + \frac{1}{4}\mathbb{E}&\left[\sum_{k=1}^N \|X^k_{\varepsilon,\delta,n} - X^{k-1}_{\varepsilon,\delta,n}\|^2\right] \\
& +\tau\delta\mathbb{E}\left[\sum_{k=1}^N \|\nabla X^k_{\varepsilon,\delta,n}\|^2\right] + \frac{\tau\lambda}{2}\mathbb{E}\left[\sum_{k=1}^N \|X^k_{\varepsilon,\delta,n}\|^2\right] \leq C,
\end{aligned}
\tag{43}
$$

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

$$
\sup_{i=1,\dots,N} \mathbb{E}\left[\|\nabla X^i_{\varepsilon,\delta,n}\|^2\right] + \frac{1}{4}\mathbb{E}\left[\sum_{k=1}^N \|\nabla(X^k_{\varepsilon,\delta,n} - X^{k-1}_{\varepsilon,\delta,n})\|^2\right] + \tau\delta\mathbb{E}\left[\sum_{k=1}^N \|\Delta X^k_{\varepsilon,\delta,n}\|^2\right] \leq C_n.
\tag{44}
$$

Proof of Lemma 4.4. We set $v_h = X^i_{\varepsilon,\delta,n}$ in (42), use the identity $2(a-b)a = a^2 - b^2 + (a-b)^2$ and get for $i = 1, \dots, N$

$$
\begin{aligned}
\frac{1}{2}\|X^i_{\varepsilon,\delta,n}\|^2 + \frac{1}{2}\|X^i_{\varepsilon,\delta,n} - X^{i-1}_{\varepsilon,\delta,n}\|^2 + \tau\delta\|\nabla X^i_{\varepsilon,\delta,n}\|^2 &+ \tau \left(\frac{\nabla X^i_{\varepsilon,\delta,n}}{\sqrt{|\nabla X^i_{\varepsilon,\delta,n}|^2 + \varepsilon^2}}, \nabla X^i_{\varepsilon,\delta,n} \right) \\
= \frac{1}{2}\|X^{i-1}_{\varepsilon,\delta,n}\|^2 - \tau\lambda \left(\|X^i_{\varepsilon,\delta,n}\|^2 - \left(g^n, X^i_{\varepsilon,\delta,n}\right)\right) &+ \left(X^{i-1}_{\varepsilon,\delta,n}, X^i_{\varepsilon,\delta,n}\right) \Delta_i W.
\end{aligned}
\tag{45}
$$

We take expected value in (45) and on noting the properties of 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 $X^{i-1}_{\varepsilon,\delta,n}$ we estimate the stochastic term

as

$$\mathbb{E}\left[\left(X_{\varepsilon,\delta,n}^{i-1}, X_{\varepsilon,\delta,n}^i\right)\Delta_i W\right] = \mathbb{E}\left[\left(X_{\varepsilon,\delta,n}^{i-1}, X_{\varepsilon,\delta,n}^i - X_{\varepsilon,\delta,n}^{i-1}\right)\Delta_i W\right] + \mathbb{E}\left[\left(X_{\varepsilon,\delta,n}^{i-1}, X_{\varepsilon,\delta,n}^{i-1}\right)\Delta_i W\right]$$

$$\leq \mathbb{E}\left[\frac{1}{4}\|X_{\varepsilon,\delta,n}^{i-1} - X_{\varepsilon,\delta,n}^i\|^2 + \|X_{\varepsilon,\delta,n}^{i-1}\|^2|\Delta_i W|^2\right] + \mathbb{E}\left[\|X_{\varepsilon,\delta,n}^{i-1}\|^2\right]\mathbb{E}\left[\Delta_i W\right]$$

$$= \frac{1}{4}\mathbb{E}\left[\|X_{\varepsilon,\delta,n}^i - X_{\varepsilon,\delta,n}^{i-1}\|^2\right] + \tau\mathbb{E}\left[\|X_{\varepsilon,\delta,n}^{i-1}\|^2\right].$$

We neglect the positive term

$$\left(\frac{\nabla X_{\varepsilon,\delta,n}^i}{\sqrt{|\nabla X_{\varepsilon,\delta,n}^i|^2 + \varepsilon^2}}, \nabla X_{\varepsilon,\delta,n}^i\right) \geq 0,$$

and get from (45) that

$$\frac{1}{2}\mathbb{E}\left[\|X_{\varepsilon,\delta,n}^i\|^2\right] + \frac{1}{4}\mathbb{E}\left[\|X_{\varepsilon,\delta,n}^i - X_{\varepsilon,\delta,n}^{i-1}\|^2\right] + \tau\delta\mathbb{E}\left[\|\nabla X_{\varepsilon,\delta,n}^i\|^2\right] + \frac{\tau\lambda}{2}\mathbb{E}\left[\|X_{\varepsilon,\delta,n}^i\|^2\right]$$

$$\leq \frac{1}{2}\mathbb{E}\left[\|X_{\varepsilon,\delta,n}^{i-1}\|^2\right] + \tau\mathbb{E}\left[\|X_{\varepsilon,\delta,n}^{i-1}\|^2\right] + \tau\lambda\|g^n\|^2.$$

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

$$\frac{1}{2}\mathbb{E}\left[\|X_{\varepsilon,\delta,n}^i\|^2\right] + \frac{1}{4}\mathbb{E}\left[\sum_{k=1}^i\|X_{\varepsilon,\delta,n}^k - X_{\varepsilon,\delta,n}^{k-1}\|^2\right] + \tau\delta\mathbb{E}\left[\sum_{k=1}^i\|\nabla X_{\varepsilon,\delta,n}^k\|^2\right] + \frac{\tau\lambda}{2}\mathbb{E}\left[\sum_{k=1}^i\|X_{\varepsilon,\delta,n}^k\|^2\right]$$

$$(46) \qquad \leq \frac{1}{2}\mathbb{E}\left[\|x_0^n\|^2\right] + \tau\mathbb{E}\left[\sum_{k=1}^i\|X_{\varepsilon,\delta,n}^{k-1}\|^2\right] + T\lambda\|g^n\|^2.$$

By the discrete Gronwall lemma it follows from (46) that

$$\sup_{i=1,\ldots,N}\mathbb{E}\left[\|X_{\varepsilon,\delta,n}^i\|^2\right] \leq \exp(2T)\Big(\mathbb{E}\left[\|x_0\|^2\right] + 2T\lambda\|g\|^2\Big).$$

We substitute the above estimate into the right-hand side of (46) to conclude (43). To show the estimate (44) we set $v_h = \Delta X_{\varepsilon,\delta,n}^i$ in (42) use integration by parts and proceed analogically to the first part of the proof. We note that by Proposition 2.1, as in the proof of Lemma 3.2, it holds that

$$(47) \qquad \tau\left(\text{div}\frac{\nabla X_{\varepsilon,\delta,n}^i}{\sqrt{|\nabla X_{\varepsilon,\delta,n}^i|^2 + \varepsilon^2}}, \Delta X_{\varepsilon,\delta,n}^i\right) \geq 0.$$

Hence we may neglect the positive term and get that

$$\frac{1}{2}\mathbb{E}\left[\|\nabla X_{\varepsilon,\delta,n}^i\|^2\right] + \frac{1}{4}\mathbb{E}\left[\sum_{k=1}^{i}\|\nabla(X_{\varepsilon,\delta,n}^k - X_{\varepsilon,\delta,n}^{k-1})\|^2\right] + \tau\delta\mathbb{E}\left[\sum_{k=1}^{i}\|\Delta X_{\varepsilon,\delta,n}^k\|^2\right]$$

$$+ \frac{\tau\lambda}{2}\mathbb{E}\left[\sum_{k=1}^{i}\|\nabla X_{\varepsilon,\delta,n}^k\|^2\right] \leq \frac{1}{2}\mathbb{E}\left[\|\nabla x_0^n\|^2\right] + \tau\mathbb{E}\left[\sum_{k=1}^{i}\|\nabla X_{\varepsilon,\delta,n}^{k-1}\|^2\right] + T\lambda\|\nabla g^n\|^2.$$

and obtain (44) after an application of the discrete Gronwall lemma. $\qquad\square$

We define piecewise constant time-interpolants of the fully discrete solution $\{X_{\varepsilon,\delta,n}^i\}_{i=0}^N$ of (42) for $t \in [0,T]$ as

(48) $$\overline{X}_\tau^{\varepsilon,\delta,n}(t) := X_{\varepsilon,\delta,n}^i \quad \text{if} \quad t \in (t_{i-1}, t_i]$$

and

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

We note that (42) can be reformulated as

$$\left(\overline{X}_\tau^{\varepsilon,\delta,n}(t), v_h\right) + \left\langle \int_0^{\theta_+(t)} A^{\varepsilon,\delta}\overline{X}_\tau^{\varepsilon,\delta,n}(s)\,\mathrm{d}s, v_h \right\rangle$$

(50) $$= \left(X_{\varepsilon,\delta,n}^0, v_h\right) + \left(\int_0^{\theta_+(t)} \overline{X}_{\tau_-}^{\varepsilon,\delta,n}(s)\,\mathrm{d}W(s), v_h\right) \qquad \text{for } t \in [0,T],$$

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

Estimate (43) yields the bounds

(51) $$\sup_{t\in[0,T]} \mathbb{E}\left[\|\overline{X}_\tau^{\varepsilon,\delta,n}(t)\|^2\right] \leq C, \qquad\qquad \sup_{t\in[0,T]} \mathbb{E}\left[\|\overline{X}_{\tau_-}^{\varepsilon,\delta,n}(t)\|^2\right] \leq C,$$

$$\delta\mathbb{E}\left[\int_0^T \|\nabla\overline{X}_\tau^{\varepsilon,\delta,n}(s)\|^2\,\mathrm{d}s\right] \leq C.$$

Furthermore, (51) and (11) imply

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

The estimates in (51) imply for fixed $n \in \mathbb{N}$, ε, $\delta > 0$ the existence of a subsequence, still denoted by $\{\overline{X}_\tau^{\varepsilon,\delta,n}\}_{\tau>0}$, and a $Y \in L^2(\Omega\times(0,T);\mathbb{L}^2)\cap L^2(\Omega\times(0,T);\mathbb{H}_0^1)\cap L^\infty((0,T);L^2(\Omega;\mathbb{L}^2))$, s.t., for $\tau \to 0$

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

In addition, there exists a $\nu \in L^2(\Omega; \mathbb{L}^2)$ such that $\overline{X}_\tau^{\varepsilon,\delta,n}(T) \rightharpoonup \nu$ in $L^2(\Omega; \mathbb{L}^2)$ as $\tau \to 0$ and the estimate (52) implies the existence of a $a^{\varepsilon,\delta} \in L^2(\Omega \times (0,T); \mathbb{H}^{-1})$, s.t.,

$$(54) \qquad A^{\varepsilon,\delta} \overline{X}_\tau^{\varepsilon,\delta,n} \rightharpoonup a^{\varepsilon,\delta} \text{ in } L^2(\Omega \times (0,T); \mathbb{H}^{-1}) \quad \text{for } \tau \to 0.$$

The estimates in (51) also implies for fixed $n \in \mathbb{N}$, ε, $\delta > 0$ the existence of a subsequence, still denoted by $\{\overline{X}_{\tau_-}^{\varepsilon,\delta,n}\}_{\tau > 0}$, and a $Y^- \in L^2(\Omega \times (0,T); \mathbb{L}^2)$, s.t.,

$$(55) \qquad \overline{X}_{\tau_-}^{\varepsilon,\delta,n} \rightharpoonup Y^- \text{ in } L^2(\Omega \times (0,T); \mathbb{L}^2) \quad \text{for } \tau \to 0.$$

Finally, the inequality (46) implies

$$\lim_{\tau \to \infty} \mathbb{E}\left[\int_0^T \|\overline{X}_\tau^{\varepsilon,\delta,n}(s) - \overline{X}_{\tau_-}^{\varepsilon,\delta,n}(s)\|^2 \, \mathrm{d}s\right] = \lim_{\tau \to \infty} \tau \mathbb{E}\left[\sum_{k=1}^N \|X_{\varepsilon,\delta,n}^k - X_{\varepsilon,\delta,n}^{k-1}\|^2\right]$$

$$(56) \qquad\qquad\qquad\qquad\qquad\qquad\qquad \leq \lim_{\tau \to \infty} C\tau$$

$$= 0.$$

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

The following result shows that the limit $Y \equiv X_n^{\varepsilon,\delta}$, i.e., that the numerical solution of scheme (42) converges to the unique variational solution of (7) for $\tau \to 0$. Owing to the properties (10), (11) the convergence proof follows standard arguments for the convergence of numerical approximations of monotone equations, see for instance [9], [6], and is therefore omitted. We note that the convergence of the whole sequence $\{\overline{X}_\tau^{\varepsilon,\delta,n}\}_{\tau > 0}$ follows by the uniqueness of the variational solution.

Lemma 4.5. *Let $x_0 \in L^2(\Omega, \mathcal{F}_0; \mathbb{L}^2)$ and $g \in \mathbb{L}^2$ be given, let $\varepsilon, \delta, \lambda > 0$, $n \in \mathbb{N}$ be fixed. Further, let $X_n^{\varepsilon,\delta}$ be the unique variational solution of (7) for $x_0^n = \mathcal{P}_n x_0$, $g^n = \mathcal{P}_n g$ and $\overline{X}_\tau^{\varepsilon,\delta,n}$, $\overline{X}_{\tau_-}^{\varepsilon,\delta,n}$ be the respective time-interpolant (48), (49) of the numerical solution $\{X_{\varepsilon,\delta,n}^i\}_{i=1}^N$ of (42). Then $\overline{X}_\tau^{\varepsilon,\delta,n}$, $\overline{X}_{\tau_-}^{\varepsilon,\delta,n}$ converge to $X_n^{\varepsilon,\delta}$ for $\tau \to 0$ in the sense that the weak limits from (53), (54) satisfy $Y \equiv X_n^{\varepsilon,\delta}$, $a^{\varepsilon,\delta} \equiv A^{\varepsilon,\delta} Y \equiv A^{\varepsilon,\delta} X_n^{\varepsilon,\delta}$ and $\nu = Y(T) \equiv X_n^{\varepsilon,\delta}(T)$. In addition it holds for almost all $(\omega, t) \in \Omega \times (0,T)$ that*

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

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

$$(57) \qquad \frac{1}{2}\|Y(t)\|^2 = \frac{1}{2}\|Y(0)\|^2 + \int_0^t \langle A^{\varepsilon,\delta} Y(s), Y(s)\rangle + \frac{1}{2}\|Y(s)\|^2 \, \mathrm{d}s$$

$$+ \int_0^t (Y(s), Y(s)) \, \mathrm{d}W(s).$$

The strong monotonicity property (10) of the operator $A^{\varepsilon,\delta}$ implies strong convergence of the numerical approximation in $L^2(\Omega \times (0,T); \mathbb{L}^2)$.

Lemma 4.6. *Let $x_0 \in L^2(\Omega, \mathcal{F}_0; \mathbb{L}^2)$ and $g \in \mathbb{L}^2$ be given, let $\varepsilon, \delta, \lambda > 0$, $n \in \mathbb{N}$ be fixed. Further, let $X_n^{\varepsilon,\delta}$ be the variational solution of (7) for $x_0^n = \mathcal{P}_n x_0$, $g^n = \mathcal{P}_n g$ and $\overline{X}_\tau^{\varepsilon,\delta,n}$ be the time-interpolant (48) of the numerical solution $\{X_{\varepsilon,\delta,n}^i\}_{i=1}^N$ of (42). Then the following convergence holds true*

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

Proof of Lemma 4.6. The proof follows along the lines of [9], [6]. We sketch the main steps of the proof for the convenience of the reader.

We note that $\overline{X}_\tau^{\varepsilon,\delta,n}$ satisfies (cf. proof of Lemma 4.4)

$$
\begin{aligned}
e^{-\kappa T}\mathbb{E}\left[\|\overline{X}_\tau^{\varepsilon,\delta,n}(T)\|^2\right] \leq & \mathbb{E}\left[\|x_0^n\|^2\right] - \kappa \int_0^T e^{-\kappa s}\mathbb{E}\left[\|\overline{X}_\tau^{\varepsilon,\delta,n}(s)\|^2\right]\,\mathrm{d}s \\
& -2\mathbb{E}\left[\int_0^T e^{-\kappa s}\langle A^{\varepsilon,\delta}\overline{X}_\tau^{\varepsilon,\delta,n}(s), \overline{X}_\tau^{\varepsilon,\delta,n}(s)\rangle\,\mathrm{d}s\right] \\
& +\mathbb{E}\left[\int_0^T e^{-\kappa s}\|\overline{X}_\tau^{\varepsilon,\delta,n}(s)\|^2\,\mathrm{d}s\right] + \kappa \int_0^T e^{-\kappa s}|R_\tau(s)|\,\mathrm{d}s,
\end{aligned}
$$

(59)

where $R_\tau(t) := \mathbb{E}\left[\int_t^{\theta_+(t)} 2\langle A^{\varepsilon,\delta}\overline{X}_\tau^{\varepsilon,\delta,n}(s), \overline{X}_\tau^{\varepsilon,\delta,n}(s)\rangle - \|\overline{X}_\tau^{\varepsilon,\delta,n}(s)\|^2\,\mathrm{d}s\right]$.

We reformulate the third term on the right-hand side in (59) as

$$
\begin{aligned}
& \mathbb{E}\left[\int_0^T e^{-\kappa s}\langle A^{\varepsilon,\delta}\overline{X}_\tau^{\varepsilon,\delta,n}(s), \overline{X}_\tau^{\varepsilon,\delta,n}(s)\rangle\,\mathrm{d}s\right] \\
& = \mathbb{E}\left[\int_0^T e^{-\kappa s}\langle A^{\varepsilon,\delta}\overline{X}_\tau^{\varepsilon,\delta,n}(s) - A^{\varepsilon,\delta}X_n^{\varepsilon,\delta}(s), \overline{X}_\tau^{\varepsilon,\delta,n}(s) - X_n^{\varepsilon,\delta}(s)\rangle\,\mathrm{d}s\right] \\
& + \mathbb{E}\left[\int_0^T e^{-\kappa s}\langle A^{\varepsilon,\delta}X_n^{\varepsilon,\delta}(s), \overline{X}_\tau^{\varepsilon,\delta,n}(s) - X_n^{\varepsilon,\delta}(s)\rangle + \langle A^{\varepsilon,\delta}\overline{X}_\tau^{\varepsilon,\delta,n}(s), X_n^{\varepsilon,\delta}(s)\rangle\,\mathrm{d}s\right].
\end{aligned}
$$

We substitute the equality above into (59) and obtain for $\kappa \geq 1$ that

$$
\begin{aligned}
& e^{-\kappa T}\mathbb{E}\left[\|\overline{X}_\tau^{\varepsilon,\delta,n}(T)\|^2\right] + 2\mathbb{E}\left[\int_0^T e^{-\kappa s}\langle A^{\varepsilon,\delta}\overline{X}_\tau^{\varepsilon,\delta,n}(s) - A^{\varepsilon,\delta}X_n^{\varepsilon,\delta}(s), \overline{X}_\tau^{\varepsilon,\delta,n}(s) - X^\varepsilon(s)\rangle\,\mathrm{d}s\right] \\
& \leq \mathbb{E}\left[\|x_0^n\|^2\right] - 2\mathbb{E}\left[\int_0^T e^{-\kappa s}\langle A^{\varepsilon,\delta}X_n^{\varepsilon,\delta}(s), \overline{X}_\tau^{\varepsilon,\delta,n}(s) - X_n^{\varepsilon,\delta}(s)\rangle + \langle A^{\varepsilon,\delta}\overline{X}_\tau^{\varepsilon,\delta,n}(s), X_n^{\varepsilon,\delta}(s)\rangle\,\mathrm{d}s\right] \\
& + \kappa \int_0^T e^{-\kappa s}|R_\tau(s)|\,\mathrm{d}s.
\end{aligned}
$$

We observe that $\int_0^T e^{-\kappa s}|R_\tau(s)|\,\mathrm{d}s \to 0$ for τ. Hence, by the lower-semicontinuity of norms using the convergence properties from Lemma 4.5 and the monotonicity property (10) we

get for $\tau \to 0$ that

$$
e^{-\kappa T}\mathbb{E}\left[\|X_n^{\varepsilon,\delta}(T)\|^2\right] + 2\lambda \lim_{\tau \to 0}\mathbb{E}\left[\int_0^T e^{-\kappa s}\|\overline{X}_\tau^{\varepsilon,\delta,n}(s) - X_n^{\varepsilon,\delta}(s)\|^2\,\mathrm{d}s\right]
$$

$$
(60) \qquad + 2\delta \lim_{\tau \to 0}\mathbb{E}\left[\int_0^T e^{-\kappa s}\|\nabla\big(\overline{X}_\tau^{\varepsilon,\delta,n}(s) - X_n^{\varepsilon,\delta}(s)\big)\|^2\,\mathrm{d}s\right]
$$

$$
\leq \mathbb{E}\left[\|x_0^n\|^2\right] - 2\mathbb{E}\left[\int_0^T e^{-\kappa s}\langle A^{\varepsilon,\delta}X_n^{\varepsilon,\delta}(s), X_n^{\varepsilon,\delta}(s)\rangle\,\mathrm{d}s\right].
$$

It is not difficult to see that (57) for $Y \equiv X_n^{\varepsilon,\delta}$ implies

$$
(61) \qquad e^{-\kappa T}\mathbb{E}\left[\|X_n^{\varepsilon,\delta}(T)\|^2\right] = \mathbb{E}\left[\|x_0^n\|^2\right] - 2\mathbb{E}\left[\int_0^T e^{-\kappa s}\langle A^{\varepsilon,\delta}(s)X_n^{\varepsilon,\delta}, X_n^{\varepsilon,\delta}(s)\rangle\,\mathrm{d}s\right]
$$

$$
- \kappa \int_0^T e^{-\kappa s}\mathbb{E}\left[\|X_n^{\varepsilon,\delta}(s)\|^2\right]\,\mathrm{d}s + \mathbb{E}\left[\int_0^T e^{-\kappa s}\|X_n^{\varepsilon,\delta}(s)\|^2\,\mathrm{d}s\right].
$$

We subtract the equality (61) from (60) and obtain for $\kappa \geq 1$

$$
\lambda \lim_{\tau \to \infty}\mathbb{E}\left[\int_0^T e^{-\kappa s}\|\overline{X}_\tau^{\varepsilon,\delta,n}(s) - X_n^{\varepsilon,\delta}(s)\|^2\,\mathrm{d}s\right] \leq 0.
$$

Hence, we conclude that $\overline{X}_\tau^{\varepsilon,\delta,n} \to X_n^{\varepsilon,\delta}$ in $L^2(\Omega; L^2((0,T); \mathbb{L}^2))$. $\qquad\square$

Remark 4.2. *It is obvious from the proof of Lemma 4.6 that the strong convergence in $L^2(\Omega \times (0,T); \mathbb{L}^2)$ remains valid for $\lambda = 0$ due to (10) by the Poincaré inequality.*

Next lemma guarantees the convergence of the numerical solution of scheme (42) to the numerical solution of scheme (41) for $\delta \to 0$.

Lemma 4.7. *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 $C \equiv C(T) > 0$, $C_n \equiv C(\mathbb{E}[\|x_0^n\|_{\mathbb{H}_0^1}], \|g^n\|_{\mathbb{H}_0^1}) > 0$ such that for any $N \in \mathbb{N}$, $\delta > 0$, $m \geq n \in \mathbb{N}$, $\varepsilon \in (0,1]$ the following estimate holds for the difference of numerical solutions of (41) and (42):*

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

We note that the n-dependent constant C_n in the estimate above is due to the a priori estimate (44), for $\mathbb{H}_0^1$-regular data x_0, g it holds that $C_n \equiv C(\mathbb{E}[\|x_0\|_{\mathbb{H}_0^1}], \|g\|_{\mathbb{H}_0^1})$ by the stability of the $\mathbb{L}^2$ projection in $\mathbb{H}_0^1$.

Proof of Lemma 4.7. We define $Z_\varepsilon^i := X_\varepsilon^i - X_{\varepsilon,\delta,n}^i$. From (41) and (42) we get

$$\left(Z_\varepsilon^i, v_h\right) = \left(Z_\varepsilon^{i-1}.v_h\right) - \tau\delta\left(\Delta X_{\varepsilon,\delta,n}^i, v_h\right)$$
$$- \tau\left(\frac{\nabla X_\varepsilon^i}{\sqrt{|\nabla X_\varepsilon^i|^2 + \varepsilon^2}}, v_h\right) - \tau\left(\frac{\nabla X_{\varepsilon,\delta,n}^i}{\sqrt{|\nabla X_{\varepsilon,\delta,n}^i|^2 + \varepsilon^2}}, \nabla v_h\right)$$
$$- \tau\lambda\left(Z_\varepsilon^i, v_h\right) - \tau\lambda\left(g^m - g^n, v_h\right)$$
$$+ \left(Z_\varepsilon^{i-1}, v_h\right)\Delta_i W.$$

We set $v_h = Z_\varepsilon^i$ and obtain

$$\left(Z_\varepsilon^i - Z_\varepsilon^{i-1}, Z_\varepsilon^i\right) = -\tau\delta\left(\Delta X_{\varepsilon,\delta,n}^i, Z_\varepsilon^i\right)$$
$$- \tau\left(\frac{\nabla X_\varepsilon^i}{\sqrt{|\nabla X_\varepsilon^i|^2 + \varepsilon^2}} - \frac{\nabla X_{\varepsilon,\delta,n}^i}{\sqrt{|\nabla X_{\varepsilon,\delta,n}^i|^2 + \varepsilon^2}}, \nabla Z_\varepsilon^i\right)$$
$$- \tau\lambda\|Z_\varepsilon^i\|^2 - \tau\lambda\left(g^m - g^n, Z_\varepsilon^i\right)$$
$$+ \left(Z_\varepsilon^{i-1}, Z_\varepsilon^i\right)\Delta_i W.$$

We note that

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

and by the Cauchy-Schwarz and Young's inequalities

$$\tau\delta\left(\Delta X_{\varepsilon,\delta,n}^i, Z_\varepsilon^i\right) \leq \frac{\tau\delta^2}{2\lambda}\|\Delta X_{\varepsilon,\delta,n}^i\|^2 + \frac{\tau\lambda}{2}\|Z_\varepsilon^i\|^2,$$
$$\tau\lambda\left(g^m - g^n, Z_\varepsilon^i\right) \leq \frac{\tau\lambda}{2}\|g^m - g^n\|^2 + \frac{\tau\lambda}{2}\|Z_\varepsilon^i\|^2.$$

From the convexity (12) it follows that

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

Hence, we obtain that

(62) $$\frac{1}{2}\|Z_\varepsilon^i\|^2 + \frac{1}{2}\|Z_\varepsilon^i - Z_\varepsilon^{i-1}\|^2$$
$$\leq \frac{1}{2}\|Z_\varepsilon^{i-1}\|^2 + \frac{\tau\delta^2}{2\lambda}\|\Delta X_{\varepsilon,\delta,n}^i\|^2 + \frac{\tau\lambda}{2}\|g^m - g^n\|^2 + \left(Z_\varepsilon^{i-1}, Z_\varepsilon^i\right)\Delta_i W.$$

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

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

and substitute the above identity into (62)

$$\frac{1}{2} \| Z_\varepsilon^i \|^2 + \frac{1}{2} \| Z_\varepsilon^i - Z_\varepsilon^{i-1} \|^2 \leq \frac{1}{2} \| Z_\varepsilon^{i-1} \|^2 + \frac{\tau \delta^2}{2\lambda} \| \Delta X_{\varepsilon,\delta,n}^i \|^2 + \frac{\tau \lambda}{2} \| g^m - g^n \|^2 + \frac{1}{2} \| Z_\varepsilon^i - Z_\varepsilon^{i-1} \|^2$$
$$+ \frac{1}{2} \| Z_\varepsilon^{i-1} \|^2 |\Delta_i W|^2 + \| Z_\varepsilon^{i-1} \|^2 \Delta_i W \,.$$

Next, we sum up the above inequality up to $i \leq N$ and obtain

$$\frac{1}{2} \| Z_\varepsilon^i \|^2 \leq \frac{1}{2} \| Z_\varepsilon^0 \|^2 + \frac{\tau \delta^2}{2\lambda} \sum_{k=1}^i \| \Delta X_{\varepsilon,\delta}^k \|^2 + \frac{1}{2} \sum_{k=1}^i \| Z_\varepsilon^{k-1} \|^2 |\Delta_k W|^2 + \sum_{k=1}^i \| Z_\varepsilon^{k-1} \|^2 \Delta_k W$$
$$+ \frac{T\lambda}{2} \| g^m - g^n \|^2.$$

After taking expectation in the above and using the independence properties of Wiener increments and the estimate (44) we arrive at

$$\frac{1}{2} \mathbb{E} \left[\| Z_\varepsilon^i \|^2 \right] \leq \frac{1}{2} \| Z_\varepsilon^0 \|^2 + \frac{\tau \delta^2}{2\lambda} \mathbb{E} \left[\sum_{k=1}^i \| \Delta X_{\varepsilon,\delta}^k \|^2 \right] + \frac{\tau}{2} \sum_{k=1}^i \mathbb{E} \left[\| Z_\varepsilon^{k-1} \|^2 \right]$$
$$\leq C_n \delta + \frac{1}{2} \mathbb{E} \left[\| Z_\varepsilon^0 \|^2 \right] + \frac{\tau}{2} \sum_{k=0}^{i-1} \mathbb{E} \left[\| Z_\varepsilon^k \|^2 \right] + \frac{T\lambda}{2} \| g^m - g^n \|^2.$$

with $C_n \equiv C(\| x_0^n \|_{\mathbb{H}_0^1}, \| g^n \|_{\mathbb{H}_0^1})$. Finally, the Discrete Gronwall lemma yields for $i = 1, \ldots, N$ that

$$(63) \qquad \mathbb{E} \left[\| Z_\varepsilon^i \|^2 \right] \leq \exp(T)(C_n \delta + \frac{1}{2} \mathbb{E} \left[\| x_0^m - \mathcal{P}_m x_0^n \|^2 \right] + \frac{T\lambda}{2} \| g^m - g^n \|^2).$$

which concludes the proof . $\qquad \square$

We define piecewise constant time-interpolant of the fully discrete solution $\{ X_\varepsilon^i \}_{i=0}^N$ of (41) for $t \in [0, T)$ as

$$(64) \qquad \overline{X}_\tau^\varepsilon(t) = X_\varepsilon^i \quad \text{if} \quad t \in (t_{i-1}, t_i].$$

We are now ready to state the second main result of this paper which is the convergence of the fully discrete numerical approximation (41) to the unique SVI solution of the total variation flow (1) (cf. Definition 3.1).

Theorem 4.1. *Let X be the SVI solution of* (1) *and let $\overline{X}_\tau^\varepsilon$ be the time-interpolant (64) of the numerical solution of the scheme* (41). *Then the following convergence holds true*

$$(65) \qquad \lim_{\varepsilon \to 0} \lim_{\tau \to 0} \| X - \overline{X}_\tau^\varepsilon \|_{L^2(\Omega \times (0,T); \mathbb{L}^2)}^2 \to 0.$$

Proof of Theorem 4.1. For $x_0 \in L^2(\Omega, \mathcal{F}_0; \mathbb{L}^2)$ and $g \in \mathbb{L}^2$ we define the $\mathbb{H}_0^1$-approximating sequences $\{x_0^n\}_{n \in \mathbb{N}} \subset \mathbb{H}_0^1$, $x_0^n \to x_0 \in L^2(\Omega, \mathcal{F}_0; \mathbb{L}^2)$, $\{g^n\}_{n \in \mathbb{N}} \subset \mathbb{H}_0^1$, $n \in \mathbb{N}$, $g^n \to g \in \mathbb{L}^2$ via the $\mathbb{L}^2$-projection onto $\mathbb{V}_n \subset \mathbb{H}_0^1$. We consider the solutions $X^\varepsilon, X_n^{\varepsilon,\delta}$ of (3), (7), respectively, and denote by X_n^ε the SVI solution of (3) for $x_0 \equiv x_0^n$, $g \equiv g^n$. Furthermore, we recall that the interpolant $\overline{X}_\tau^{\varepsilon,\delta,n}$ of the numerical solution of (42) was defined in (48).

We split the numerical error as

$$
\begin{aligned}
\|X - \overline{X}_\tau^\varepsilon\|_{L^2(\Omega \times (0,T); \mathbb{L}^2)}^2 \leq & \|X - X^\varepsilon\|_{L^2(\Omega \times (0,T); \mathbb{L}^2)}^2 + \|X^\varepsilon - X_n^\varepsilon\|_{L^2(\Omega \times (0,T); \mathbb{L}^2)}^2 \\
(66) \qquad & + \|X_n^\varepsilon - X_n^{\varepsilon,\delta}\|_{L^2(\Omega \times (0,T); \mathbb{L}^2)}^2 + \|X_n^{\varepsilon,\delta} - \overline{X}_\tau^{\varepsilon,\delta,n}\|_{L^2(\Omega \times (0,T); \mathbb{L}^2)}^2 \\
& + \|\overline{X}_\tau^{\varepsilon,\delta,n} - \overline{X}_\tau^\varepsilon\|_{L^2(\Omega \times (0,T); \mathbb{L}^2)}^2 \\
=: & \ I + II + III + IV + V.
\end{aligned}
$$

By Theorem 3.1 it follows that

$$
\lim_{\varepsilon \to 0} I = \lim_{\varepsilon \to 0} \|X - X^\varepsilon\|_{L^2(\Omega \times (0,T); \mathbb{L}^2)}^2 = 0.
$$

To estimate the second term we consider the solutions X_n^ε of (3) with $x_0 \equiv x_0^n$ and $g \equiv g^n$. From (20) we deduce that

$$
\lim_{n \to \infty} II = \lim_{n \to \infty} \|X^\varepsilon - X_n^\varepsilon\|_{L^2(\Omega \times (0,T); \mathbb{L}^2)}^2 = 0.
$$

We use (22) to estimate the third term as

$$
\lim_{\delta \to 0} III = \lim_{\delta \to 0} \|X_n^\varepsilon - X_n^{\varepsilon,\delta}\|_{L^2(\Omega \times (0,T); \mathbb{L}^2)}^2 = 0.
$$

The fourth term is estimated by Lemma 4.6

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

For the last term we use Lemma 4.7

$$
\lim_{n \to \infty} \limsup_{\delta \to 0} \limsup_{m \to \infty} V = \lim_{n \to \infty} \limsup_{\delta \to 0} \limsup_{m \to \infty} \|\overline{X}_\tau^{\varepsilon,\delta,n} - \overline{X}_\tau^\varepsilon\|_{L^2(\Omega \times (0,T); \mathbb{L}^2)}^2 = 0.
$$

Finally, we consecutively take $\tau \to 0$ ($m \to \infty$), $\delta \to 0$, $n \to \infty$ and $\varepsilon \to 0$ in (66) and use the above convergence of $I - V$ to obtain (65). $\qquad \square$

We conclude this section by showing unconditional stability of scheme (41), i.e., we show that the numerical solution satisfies a discrete energy law which is an analogue of the energy estimate (27).

Lemma 4.8. *Let $x_0, g \in \mathbb{L}^2$ and $T > 0$. Then there exist a constant $C \equiv C(T)$ such that the solutions of scheme (41) for $\varepsilon \in (0, 1]$ satisfy for any $m, N \in \mathbb{N}$*

$$
\begin{aligned}
(67) \qquad \sup_{i=1,\dots,N} \frac{1}{2} \mathbb{E}\left[\|X_\varepsilon^i\|^2\right] + \tau \mathbb{E}\left[\sum_{k=1}^N \mathcal{J}_\varepsilon(X_\varepsilon^i) + \frac{\tau\lambda}{2}\|X_\varepsilon^i - g^m\|^2\right] \\
\leq C\left(\frac{1}{2}\mathbb{E}\left[\|x_0\|^2\right] + T|\mathcal{O}| + \frac{\lambda}{2}\|g\|^2\right).
\end{aligned}
$$

Proof of Lemma 4.8. We set $v_h \equiv X_\varepsilon^i$ in (41) and obtain

$$
\frac{1}{2}\|X_\varepsilon^i\|^2 + \frac{1}{2}\|X_\varepsilon^i - X_\varepsilon^{i-1}\|^2 + \tau\left(\frac{\nabla X_\varepsilon^i}{\sqrt{|\nabla X_\varepsilon^i|^2 + \varepsilon^2}}, \nabla X_\varepsilon^i\right) + \tau\lambda(X_\varepsilon^i - g^m, X_\varepsilon^i)
$$

(68) $$
= \frac{1}{2}\|X_\varepsilon^{i-1}\|^2 + (X_\varepsilon^{i-1}, X_\varepsilon^i)\Delta_i W.
$$

Using the the convexity of $\mathcal{J}_\varepsilon$ along with the identity

$$
(X_\varepsilon^{i-1}, X_\varepsilon^i)\Delta_i W = (X_\varepsilon^{i-1}, X_\varepsilon^i - X_\varepsilon^{i-1})\Delta_i W + \|X_\varepsilon^{i-1}\|^2 \Delta_i W,
$$

we get from (68) that

$$
\frac{1}{2}\|X_\varepsilon^i\|^2 + \frac{1}{2}\|X_\varepsilon^i - X_\varepsilon^{i-1}\|^2 + \tau\mathcal{J}_\varepsilon(X_\varepsilon^i) + \frac{\tau\lambda}{2}\|X_\varepsilon^i - g^m\|^2
$$

(69) $$
\leq \tau\mathcal{J}_\varepsilon(0) + \frac{\tau\lambda}{2}\|g^m\|^2 + \frac{1}{2}\|X_\varepsilon^{i-1}\|^2 + \frac{1}{2}\|X_\varepsilon^i - X_\varepsilon^{i-1}\|^2
$$

$$
+ \frac{1}{2}\|X_\varepsilon^{i-1}\|^2\|\Delta_i W\|^2 + \|X_\varepsilon^{i-1}\|^2\Delta_i W.
$$

After taking the expectation and summing up over i in (69), and noting that $\mathcal{J}_\varepsilon(0) = \varepsilon|\mathcal{O}|$ we obtain

$$
\frac{1}{2}\mathbb{E}\left[\|X_\varepsilon^i\|^2\right] + \tau\mathbb{E}\left[\sum_{k=1}^{i}\mathcal{J}_\varepsilon(X_\varepsilon^i) + \frac{\lambda}{2}\|X_\varepsilon^i - g^m\|^2\right]
$$

$$
\leq \frac{1}{2}\mathbb{E}\left[\|x_0\|^2\right] + T\left(\varepsilon|\mathcal{O}| + \frac{\lambda}{2}\|g\|^2\right) + \frac{\tau}{2}\mathbb{E}\left[\sum_{k=0}^{i-1}\|X_\varepsilon^k\|^2\right].
$$

Hence (67) follows after an application of the discrete Gronwall lemma. $\qquad\square$

5. NUMERICAL EXPERIMENTS

We perform numerical experiments using a variant of the spectral scheme (41) on the unit square $\mathcal{O} = (0,1)^2$. For convenience we employ spatial discretization based on the $\mathbb{H}_0^1$-conforming finite element method; for a given regular triangulation $\mathcal{T}_h$ of $\mathcal{O}$ with mesh size $h = \max_{K\in\mathcal{T}_h}\{\mathrm{diam}(K)\}$ we consider the associated space of piecewise linear continuous functions $\mathbb{V}_h = \{v_h \in C^0(\overline{\mathcal{O}}); v_h|_K \in \mathcal{P}^1(K) \; \forall K \in \mathcal{T}_h\} \subset \mathbb{H}_0^1$, for more detail see, e.g., [5], [7], [4]. The fully discrete finite element scheme for $i = 1, \ldots, N$ then reads as

$$
\left(X_\varepsilon^i, v_h\right) = \left(X_\varepsilon^{i-1}, v_h\right) - \tau\left(\frac{\nabla X_\varepsilon^i}{\sqrt{|\nabla X_\varepsilon^i|^2 + \varepsilon^2}}, \nabla v_h\right)
$$

(70) $$
- \tau\lambda\left(X_\varepsilon^i - g_h, v_h\right) + \mu\left(\sigma(X_\varepsilon^{i-1})\Delta_i W^h, v_h\right) \qquad \forall v_h \in \mathbb{V}_h,
$$

$$
X_\varepsilon^0 = x_h^0,
$$

where $g_h, x_h^0 \in \mathbb{V}_h$ are suitable approximations of g, x_0 (e.g., orthogonal projections onto $\mathbb{V}_h$), respectively, and $\mu > 0$ is a constant. The multiplicative space-time noise

$\sigma(X_\varepsilon^{i-1})\Delta_i W^h$ is constructed as follows. The term W^h is taken to be a $\mathbb{V}_h$-valued space-time noise of the form

$$\Delta_i W^h(x) = \sum_{\ell=1}^{L} \varphi_\ell(x)\Delta_i\beta_\ell \qquad \forall\, x \in \overline{\mathcal{D}},$$

where β_ℓ, $\ell = 1, \ldots, L$ are independent scalar-valued Wiener processes and $\{\varphi_\ell\}_{\ell=1}^{L}$ is the standard 'nodal' finite element basis of $\mathbb{V}_h$. In the simulations below we employ three practically relevant choices of σ: a tracking-type noise $\sigma(X) \equiv \sigma_1(X) = |X - g_h|$, a gradient type noise $\sigma(X) \equiv \sigma_2(X) = |\nabla X|$ and the additive noise $\sigma(X) \equiv \sigma_3 = 1$; in the first case the noise is small when the solution is close to the 'noisy image' g_h, in the gradient noise case the noise is localized along the edges of the image.

We note that the fully discrete finite element scheme (70) satisfies a corresponding counterpart of the discrete energy law (67) for $\sigma \equiv \sigma_1, \sigma_2, \sigma_3$; the proof is analogical to the proof of Lemma 4.8.

In all experiments we set $T = 0.05$, $\lambda = 200$, $x_0 \equiv x_0^h \equiv 0$. If not mentioned otherwise we use the time step $\tau = 10^{-5}$, the mesh size $h = 2^{-5}$ and set $\varepsilon = h = 2^{-5}$, $\mu = 1$. We define $g \in \mathbb{V}_h$ as a piecewise linear interpolation of the characteristic function of a circle with radius 0.25 on the finite element mesh, see Figure 1 (left), and set $g_h = g + \xi_h \in \mathbb{V}_h$ with $\xi_h(x) = \nu \sum_{\ell=1}^{L} \varphi_\ell(x)\xi_\ell$, $x \in \mathcal{O}$ where ξ_ℓ, $\ell = 1, \ldots, L$ are realizations of independent $U(-1, 1)$-distributed random variables. If not indicated otherwise we use $\nu = 0.1$; the corresponding realization of ξ_h is displayed in Figure 1 (right).
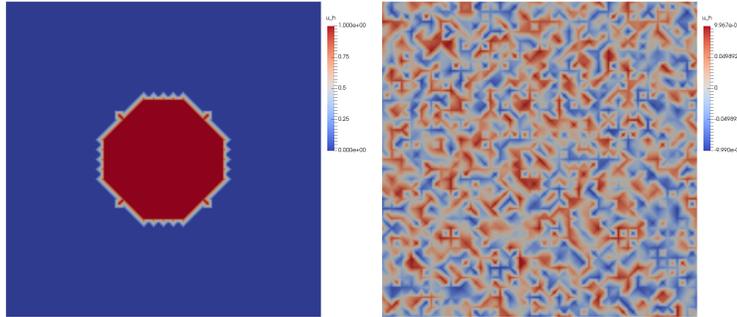
FIGURE 1. The function g (left) and the noise ξ_h (right).

We choose $\varepsilon = h = 2^{-5}$, $\mu = 1$, $\sigma \equiv \sigma_1$ as parameters for the 'baseline' experiment; the individual parameters are then varied in order to demonstrate their influence on the evolution. The time-evolution of the discrete energy functional $\mathcal{J}_{\varepsilon,\lambda}(X_\varepsilon^i)$, $i = 1, \ldots, N$ for a typical realization of the space-time noise W^h is displayed in Figure 2; in the legend of the graph we state parameters which differ from the parameters of the baseline experiment, e.g., the legend '$sigma_2$, $mu = 0.125$' corresponds to the parameters $\sigma \equiv \sigma_2$, $\mu = 0.125$ and the remaining parameters are left unchanged, i.e., $\varepsilon = h = 2^{-5}$. For all considered parameter setups, except for the case of noisier image $\nu = 0.2$, the evolution remained close

to the discrete energy of the deterministic problem (i.e., (70) with $\mu = 0$). The energy decreases over time until the solution is close to the (discrete) minimum of $\mathcal{J}_{\varepsilon,\lambda}$; to highlight the differences we display a zoom at the graphs. We observe that in the early stages (not displayed) the energy of stochastic evolutions with sufficiently small noise typically remained below the energy of the deterministic problems and the situation reversed as the solution approached the stationary state.
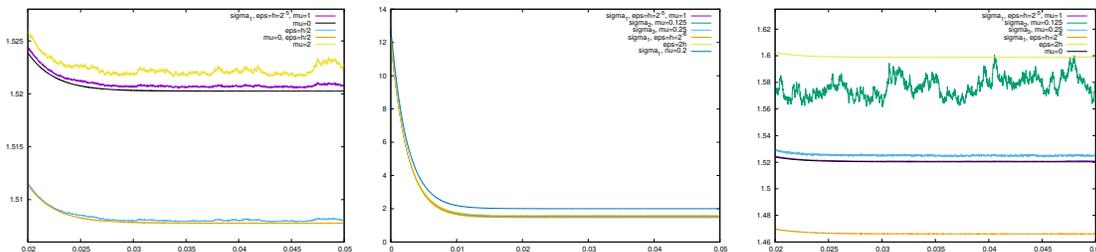
FIGURE 2. Evolution of the discrete energy: $\sigma \equiv \sigma_1$, $h = 2^{-5}$, $\varepsilon = h, \frac{h}{2}$, $\mu = 1, 2$ (left); $\sigma \equiv \sigma_1, \sigma_2, \sigma_3$, $\sigma \equiv \sigma_1$, $h = 2^{-5}$, $\varepsilon = 2h$, $\sigma = \sigma_1$, $\varepsilon = h = 2^{-6}$ and $\nu = 0.2$ (middle and right).

In Figure 3 we display the solution at the final time computed with $\sigma \equiv \sigma_1$, $\varepsilon = h$ for $h = 2^{-5}, 2^{-6}$, respectively, and $\sigma \equiv \sigma_2$, $\varepsilon = h = 2^{-5}$; graphically the results of the remaining simulations did not significantly differ from the first case. The displayed results may indicate that the noise σ_2 yields worse results than the noise σ_1 and σ_2; however, for sufficiently small value of μ the results would remain close to the deterministic simulation as well. We have magnified noise intensity μ to highlight the differences to the other noise types (i.e., the noise is concentrated along the edges of the image). We note that the gradient type noise σ_2 might be a preferred choice for practical computations, cf. [14].
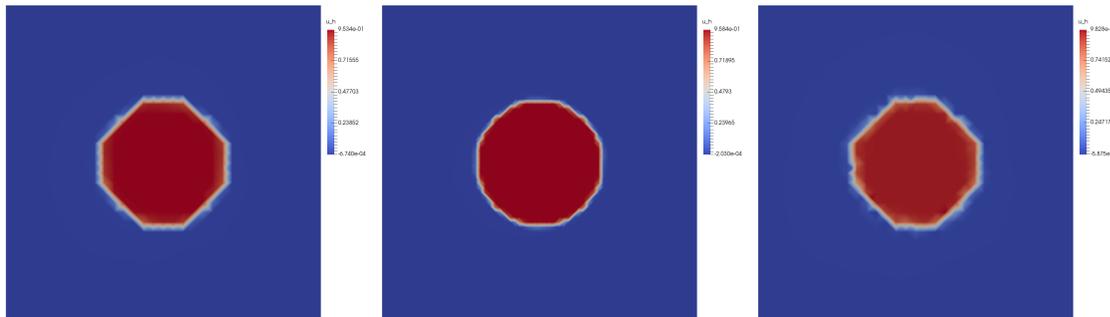
FIGURE 3. From let to right: the solution for $\sigma \equiv \sigma_1$ with $\varepsilon = h = 2^{-5}$, $\sigma \equiv \sigma_1$ with $\varepsilon = h = 2^{-6}$ and $\sigma \equiv \sigma_2$ with $\varepsilon = 2^{-5}$.

References

[1] L. Ambrosio, N. Fusco, and D. Pallara. *Functions of bounded variation and free discontinuity problems.* Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, New York, 2000.

[2] 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.

[3] 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.

[4] S. Bartels and M. Milicevic. Stability and experimental comparison of prototypical iterative schemes for total variation regularized problems. *Comput. Methods Appl. Math.*, 16(3):361–388, 2016.

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

[6] E. Emmrich and D. Šiška. Nonlinear stochastic evolution equations of second order with damping. *Stoch. Partial Differ. Equ. Anal. Comput.*, 5(1):81–112, 2017.

[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] B. Gess and J. Tölle. Stability of solutions to stochastic partial differential equations. *J. Differential Equations*, 260(6):4973–5025, 2016.

[9] I. Gyöngy and A. Millet. On discretization schemes for stochastic evolution equations. *Potential Anal.*, 23(2):99–134, 2005.

[10] O. Juan, R. Keriven, and G. Postelnicu. Stochastic motion and the level set method in computer vision: Stochastic active contours. *International Journal of Computer Vision*, 69:7–25, 2006.

[11] N.V. Krylov and B.L. Rozovskii. Stochastic evolution equations. In *Stochastic differential equations: theory and applications*, volume 2 of *Interdiscip. Math. Sci.*, pages 1–69. World Sci. Publ., Hackensack, NJ, 2007.

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

[13] L.I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Phys. D*, 60(1-4):259–268, 1992.

[14] B. Sixou, L. Wang, and F. Peyrin. Stochastic diffusion equation with singular diffusivity and gradient-dependent noise in binary tomography. *J. Phys.: Conf. Ser.*, 69:012001, 542.

[15] R. Temam. *Navier-Stokes equations. Theory and numerical analysis.* North-Holland Publishing Co., Amsterdam-New York-Oxford, 1977. Studies in Mathematics and its Applications, Vol. 2.

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

Department of Mathematics, Bielefeld University, 33501 Bielefeld, Germany
E-mail address: `roeckner@math.uni-bielefeld.de`

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