A POSTERIORI ESTIMATES FOR THE STOCHASTIC TOTAL VARIATION FLOW

LUBOMÍR BAŇAS AND ANDRÉ WILKE

Abstract. We derive a posteriori error estimates for a fully discrete time-implicit finite element approximation of the stochastic total variation flow (STVF) with additive space time noise. The estimates are first derived for an implementable fully discrete approximation of a regularized stochastic total variation flow. We then show that the derived a posteriori estimates remain valid for the unregularized flow up to a perturbation term that can be controlled by the regularization parameter. Based on the derived a posteriori estimates we propose a pathwise algorithm for the adaptive space-time refinement and perform numerical simulation for the regularized STVF to demonstrate the behavior of the proposed algorithm.

1. Introduction

In this paper we derive a posteriori error estimates for a fully discrete finite element approximation of the stochastic total variation flow (STVF)

\[ dX = \text{div} \left( \frac{\nabla X}{|\nabla X|} \right) dt - \lambda (X - g) dt + \sigma dW, \quad \text{in } (0,T) \times \Omega, \]
\[ X = 0 \quad \text{on } (0,T) \times \partial \Omega, \]
\[ X(0) = x_0 \quad \text{in } \Omega, \]

where \( \Omega \subset \mathbb{R}^d \), \( d \geq 1 \) is a bounded, convex polyhedral domain, \( \lambda \geq 0 \), \( T > 0 \) are constants and \( x_0, g \in H^1_0 \) are given functions. The term \( W \) is a \( \mathbb{R} \)-valued Wiener process (Brownian motion) on a given filtered probability space \((\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{0 \leq t \leq T}, \mathbb{P})\), and \( \sigma \in L^2(\Omega; L^\infty([0,T]; H^1_0)) \) is a \( \{\mathcal{F}_t\}_{0 \leq t \leq T} \)-adapted stochastic process on \((\Omega, \mathcal{F}, \mathbb{P})\). Equation (1) can formally be interpreted as a stochastically perturbed gradient flow of the penalized total variation energy functional

\[ \mathcal{J}_\lambda(u) := \int_\Omega |\nabla u| \, dx + \frac{\lambda}{2} \int_\Omega |u - g|^2 \, dx. \]

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – SFB 1283/2 2021 – 317210226.
Due to the singular character of total variation flow (1), it is convenient to perform numerical simulations using a regularized problem

$$dX^\varepsilon = \text{div} \left( \frac{\nabla X^\varepsilon}{|\nabla X^\varepsilon|_\varepsilon} \right) \, dt - \lambda (X^\varepsilon - g) \, dt + \sigma \, dW \quad \text{in } (0, T) \times \mathcal{O},$$

(3)  

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

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

where $|\cdot|_\varepsilon := \sqrt{|\cdot|^2 + \varepsilon^2}$ 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

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

(4)

It is well-known that the minimizers of the above regularized energy functional converge to the minimizers of (2) for $\varepsilon \to 0$, cf. [11] and the references therein. Owing 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. [14], [15], is not applicable to this problem. To study singular gradient flow problems it is convenient to apply the solution framework developed in [2] which characterizes the solutions of (1) as stochastic variational inequalities (SVIs). Convergent fully discrete finite element approximation of (1) and (3) was proposed in [5]. Throughout the paper, we refer to the solutions which satisfy a stochastic variational inequality as SVI solutions, and to the classical SPDE solutions as variational solutions.

We note that 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. We refer, for instance, to [13] where a stochastic level-set method is applied in image segmentation, and [18] which uses stochastic gradient flow of a modified (non-convex) total variation energy functional for binary tomography.

In the deterministic setting only few work exist on the adaptive finite element solution of the total variation problem; we mention [12] and [3] which study a posteriori estimates and adaptive finite element schemes for the (time independent and deterministic) minimization problems (2), (4). As far as we are aware, even in the deterministic setting, the present work is the first one which addresses a posteriori estimates and adaptivity for the finite element approximation of the time dependent total variation flow. Very few results exist on adaptive finite element methods for stochastic partial differential equations (SPDEs). Algorithmic aspects of adaptive finite element approximation of SPDEs have been addressed in [4],[17]. The first work to derive rigorous a posteriori estimates for (linear) SPDEs is [16]; for related overview of time-adaptive methods for stochastic differential equations we refer the reader to the references in this work. As far as we are aware, apart from the present paper, the only other result that studies a posteriori estimates for nonlinear SPDEs is [6].
A POSTERIORI ESTIMATES FOR THE STOCHASTIC TOTAL VARIATION FLOW

The present work is inspired by [16], where a posteriori estimates are derived for the stochastic heat equation with additive noise. The approach of [16] uses a transformation of the SPDE to a random PDE (RPDE) which enjoys improved time regularity properties that enable to postulate an error equation in a form that can be used to derive the a posteriori estimates. Due to the nonlinear character of the diffusion term in (3) the (linear) transformation used in [16] is not directly applicable in the present setting. Nevertheless, we are able to adopt the approach by only applying the transformation to the linear part of the problem which is related to the time derivative in the resulting RPDE; an advantage of this approach is that we only require $H^1$-regularity of the noise term, as opposed to the $H^2$-regularity required in [16]. Hence, we obtain the existence of a time derivative in the resulting RPDE, which is crucial ingredient for the derivation of the error equation for the numerical approximation. Consequently, we derive an a posteriori error bound for the finite element approximation of the regularized problem (3). Furthermore, by using the convergence of (3) to the SVI solution of (1) for $\varepsilon \to 0$, we obtain a posteriori estimates for the finite element approximation of the SVI solution of (3). As far as we are aware, the present work is the first to show a a posteriori estimate of implementable numerical approximation for singular stochastic gradient flows in the framework of stochastic variational inequalities as well as in the deterministic setting.

The paper is organized as follows. In Section 2 we introduce the notation, summarize useful analytical results on the solutions of STVF and introduce the transformation of the SPDE (3) into a RPDE. A practical fully discrete finite element scheme for the approximation of the regularized problem (3) is introduced in Section 3. The a posteriori error estimates for the proposed finite element approximation are derived in Section 4. In Section 5 we formulate the adaptive finite element algorithm and perform numerical simulation to demonstrate its properties.

2. Notation and Preliminaries

For $1 \leq p \leq \infty$, we denote by $(L^p, \|\cdot\|_{L^p})$ the standard spaces of $p$-th order integrable functions on $\mathcal{O}$, and use $\|\cdot\| := \|\cdot\|_{L^2}$ and $(\cdot, \cdot) := (\cdot, \cdot)_{L^2}$ for the $L^2$-inner product. For $k \in \mathbb{N}$ we denote the usual Sobolev space on $\mathcal{O}$ as $(H^k, \|\cdot\|_{H^k})$, and $(H^k_0, \|\cdot\|_{H^k_0})$ stands for the $H^k$ space with zero trace on $\partial \mathcal{O}$ with its dual space $(H^{-1}, \|\cdot\|_{H^{-1}})$. Furthermore, we set $(\cdot, \cdot) := (\cdot, \cdot)_{H^{-1} \times H^1_0}$, where $(\cdot, \cdot)_{H^{-1} \times H^1_0}$ is the duality pairing between $H^1_0$ and $H^{-1}$. We say that a function $X \in L^1(\Omega \times (0,T); L^2)$ is $\mathcal{F}_t$-progressively measurable if $X \mathbb{1}_{[0,t]}$ is $\mathcal{F}_t \otimes \mathcal{B}([0,t])$-measurable for all $t \in [0,T]$.

**Definition 2.1.** A function $u \in L^1$ is called a function of bounded variation, if its total variation

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

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


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

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

We define the functionals

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

and

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

where $\gamma_0(x)$ is the trace of $x$ on the boundary and $d\mathcal{H}^{n-1}$ is the Hausdorff measure on $\partial \mathcal{O}$. $J_{\varepsilon, \lambda}$ and $J_{\lambda}$ are both convex and lower semicontinuous over $\mathbb{L}^2$ and the lower semicontinuous hulls of $\tilde{J}_{\varepsilon, \lambda}|_{\mathbb{N}^1}$ or $\tilde{J}_{\lambda}|_{\mathbb{N}^1}$ respectively, cf. [1, Proposition 11.3.2]. A SVI solution of (3) and (1) was defined in [5] (see also [2]) as a stochastic variational inequality.

**Definition 2.2.** 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 $L^2$-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

$$
dZ(t) = -G(t) \, dt + Z(t) \, dW(t), \quad t \in [0, T],
$$

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

$$
\frac{1}{2} \mathbb{E} \left[ \|X^\varepsilon(t) - Z(t)\|^2 \right] + \mathbb{E} \left[ \int_0^t \tilde{J}_{\varepsilon, \lambda}(X^\varepsilon(s)) \, ds \right] \leq \frac{1}{2} \mathbb{E} \left[ \|x_0 - Z(0)\|^2 \right] + \mathbb{E} \left[ \int_0^t \tilde{J}_{\varepsilon, \lambda}(Z(s)) \, ds \right]
$$

$$
+ \frac{1}{2} \mathbb{E} \left[ \int_0^t \|\sigma(s) - Z(s)\|^2 \, ds \right] + \mathbb{E} \left[ \int_0^t (X^\varepsilon(s) - Z(s), G) \, ds \right],
$$

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

$$
\frac{1}{2} \mathbb{E} \left[ \|X(t) - Z(t)\|^2 \right] + \mathbb{E} \left[ \int_0^t \tilde{J}_{\lambda}(X(s)) \, ds \right] \leq \frac{1}{2} \mathbb{E} \left[ \|x_0 - Z(0)\|^2 \right] + \mathbb{E} \left[ \int_0^t \tilde{J}_{\lambda}(Z(s)) \, ds \right]
$$

$$
+ \frac{1}{2} \mathbb{E} \left[ \int_0^t \|\sigma(s) - Z(s)\|^2 \, ds \right] + \mathbb{E} \left[ \int_0^t (X(s) - Z(s), G) \, ds \right].
$$
Remark 2.1. In [5], [2] SVI solutions of (1) and (3) are defined as SVI solutions in the sense of Definition 2.2 for $\sigma \equiv X$ and $\sigma \equiv \varepsilon X$, respectively. This corresponds to the multiplicative noise case, which was considered in [3]. We note that the well-posedness result from [5] directly carries over to the present case of (1), (3) with additive noise. Furthermore, we note that the results of the present paper are directly applicable the case of $H^1$-regular trace class Wiener process, cf. [16].

Owing to the $H^1_0$-regularity of the data $x_0 g$ we may follow the line of arguments in the proof of [5, Lemma 3.2 and 3.3] to conclude the existence and uniqueness for any $\varepsilon > 0$ of a square integrable $H^1_0$-valued $\{F_t\}_{0 \leq t \leq T}$-predictable stochastic process $X^\varepsilon \in L^2(\Omega; C([0, T]; L^2)) \cap L^2(\Omega; L^\infty([0, T]; H^1_0))$ such that the following variational formulation holds: for all $t \in [0, T]$ and $\mathbb{P}$-a.s,

$$
(X^\varepsilon(t), \Phi) = (x_0, \Phi) - \int_0^t \left( \frac{\nabla X^\varepsilon(s)}{\|\nabla X^\varepsilon(s)\|_{L^\infty}}, \nabla \Phi \right) \, ds - \lambda \int_0^t (X^\varepsilon(s) - g, \Phi) \, ds + \int_0^t (\sigma(t), \Phi) \, dW(s) \quad \forall \Phi \in H^1_0.
$$

(10)

Furthermore, there exists a $C \equiv C(T) > 0$ such that the following estimates holds

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

(11)

and

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

(12)

**Transformation into a RPDE.** The lack of time-differentiability of the solutions of SPDEs (which is due to the low time-regularity of the driving Wiener process $W$) is a major obstacle in the derivation of a suitable error equation for the numerical approximation which then allows to deduce the a posteriori estimates.

To overcome this issue for the solution $X^\varepsilon$ of (3), we consider the transformation $\mathbb{P}$-a.s, for a.a $t \in [0, T]$ and a.a. $x \in \mathcal{O}$:

$$
Y^\varepsilon(t) := X^\varepsilon(t) - \int_0^t \sigma(s) \, dW(s).
$$

(13)

On noting the regularity properties (11), (12) the triangle inequality and the Burkholder-Davis-Gundy inequality imply that

$$
\mathbb{E} \left[ \sup_{t \in [0, T]} \|Y^\varepsilon(t)\|^2_{H^1_0} \right] \leq C \mathbb{E} \left[ \sup_{t \in [0, T]} \|X^\varepsilon(t)\|^2_{H^1_0} \right] + C \mathbb{E} \left[ \int_0^T \|\sigma(t)\|^2_{H^1_0} \, dt \right] \leq C,
$$

where $C$ is a constant.
i.e., the square integrable, \( \{ \mathcal{F}_t \}_{0 \leq t \leq T} \)-adapted process \( Y^\varepsilon \in L^2(\Omega; L^\infty([0, T]; \mathbb{H}_0^1)) \). Furthermore, \( Y^\varepsilon \) satisfies \( \mathbb{P} \)-a.s. for a.a. \( t \in [0, T] \)

\[
\langle \partial_t Y^\varepsilon, \varphi \rangle = -\left( \frac{\nabla X^\varepsilon}{|\nabla X^\varepsilon|_\varepsilon}, \nabla \varphi \right) - \lambda(X^\varepsilon - g, \varphi) \quad \forall \varphi \in \mathbb{H}_0^1,
\]

where \( \partial_t Y^\varepsilon \in L^2([0, T]; \mathbb{H}^{-1}) \), \( Y^\varepsilon \in C([0, T]; L^2) \) \( \mathbb{P} \)-a.s., follows by standard arguments, see for instance [10, Theorem 5.9.3] (cf., also [8, Section 2.1]). For instance, the \( \mathbb{P} \)-a.s bound \( \partial_t Y^\varepsilon \in L^2([0, T]; \mathbb{H}^{-1}) \) can be deduced (formally) from (14) by the Cauchy-Schwarz and triangle inequalities

\[
\langle \partial_t Y^\varepsilon, \varphi \rangle \leq C \left( \left\| \frac{\nabla X^\varepsilon}{|\nabla X^\varepsilon|_\varepsilon} \right\|_\varepsilon + \lambda (\|X^\varepsilon\| + \|g\|) \right) \|\varphi\|_{\mathbb{H}_0^1} \quad \forall \varphi \in \mathbb{H}_0^1,
\]

on noting (11), (12) and \( \left| \frac{\nabla}{|\nabla|_\varepsilon} \right| < 1 \); the result can be derived rigorously by Galerkin approximation.

**Remark 2.2.** In [16] the transformation (13) is used to reformulate the considered linear SPDE into an equivalent linear RPDE. An analogous transformation of (3) into a RPDE that only involves the transformed solution \( Y^\varepsilon \) is not applicable in the present setting due to the nonlinearity in (3). Hence, we just transform the linear part of equation (3) and, as discussed above, the existence of \( \partial_t Y^\varepsilon \) follows from the uniqueness and regularity which are available for the variational solution \( X^\varepsilon \) of the SPDE (3). We also remark that the present "partial" transformation has the advantage that it only requires \( \mathbb{H}^1 \)-regularity of the noise as opposed to the \( \mathbb{H}^2 \)-regularity required in [16].

### 3. Numerical approximation

For simplicity we assume throughout this section that all discretization parameters are deterministic, a generalization to random setting is straightforward, also, cf., Section 5.

We consider a sequence of time step sized \( \{ \tau_n \}_{n \geq 1} \) and set \( t_n = \sum_{i=1}^n \tau_i \). At time level \( t_n \) we consider a quasi-uniform partition \( T_h^n \) of \( \mathcal{O} \) into simplices which is obtained from \( T_h^{n-1} \) by a suitable refinement/coarsening procedures. For an element \( T \in T_h^n \) we denote by \( \mathcal{E}_T \) the set of all faces (or edges for \( d = 2 \)) of \( \partial T \). The set of all faces of the elements of the mesh \( T_h^n \) is denoted as \( \mathcal{E}_h^n := \bigcup_{T \in T_h^n} \mathcal{E}_T \); the diameter of \( T \in T_h^n \) and \( E \in \mathcal{E}_h^n \) is denoted as \( h_T \) and \( h_E \), respectively. We split \( \mathcal{E}_h^n \) into the set of all interior edges and boundary edges \( \mathcal{E}_h^\partial := \mathcal{E}_h^\partial \mathcal{O} \cup \mathcal{E}_h^\partial \mathcal{O} \) where \( \mathcal{E}_h^\partial := \{ E \in \mathcal{E}_h^n : E \subset \partial \mathcal{O} \} \). Given a \( E \in \mathcal{E}_h^n \) we denote by \( \mathcal{N}(E) \) the set of its nodes and for \( T \in T_h^n \), \( E \in \mathcal{E}_h^n \) we define the local patches \( \omega_T := \bigcup_{E \in \mathcal{E}(T) \cap \mathcal{E}(T) \neq \emptyset} T \), \( \omega_E := \bigcup_{E \in \mathcal{E}(T)} T \). The space \( \mathbb{V}_h^n \) is the usual \( \mathbb{H}^1 \)-conforming finite element space of continuous piecewise linear functions on \( T_h^n \). For \( n = 0, \ldots, N \) we denote the \( L^2 \)-orthogonal projection operator onto \( \mathbb{V}_h^n \) as \( P_n : \mathbb{L}^2 \to \mathbb{V}_h^n \). We consider the Clément-Scott-Zhang interpolation operator \( \Pi_n : \mathbb{H}_0^1 \to \mathbb{V}_h^n \) with the following local approximation properties for \( \psi \in \mathbb{H}_0^1 \):

\[
\| \psi - \Pi_n \psi \|_T + h_T \| \nabla [\psi - \Pi_n \psi] \|_T \leq C_h T \| \nabla \psi \|_{\omega_T} \quad \forall T \in T_h^n,
\]
We define continuous piecewise linear time-interpolants of \( \{Y_{\varepsilon,h}^n\}_{n \in \mathbb{N}} \), \( \{X_{\varepsilon,h}^n\}_{n \in \mathbb{N}} \) and \( \{\Sigma^n\}_{n \in \mathbb{N}} \) on \([0,T]\) as follows: for \( t \in [t_{n-1}, t_n] \) we set

\[
Y_{\varepsilon,h}^T(t) := \frac{t - t_{n-1}}{\tau_n} Y_{\varepsilon,h}^n + \frac{t_n - t}{\tau_n} Y_{\varepsilon,h}^{n-1},
\]

\[
X_{\varepsilon,h}^T(t) := \frac{t - t_{n-1}}{\tau_n} X_{\varepsilon,h}^n + \frac{t_n - t}{\tau_n} X_{\varepsilon,h}^{n-1},
\]

and

\[
\|\psi - \Pi_n \psi\|_E \leq C^* h_E^{1/2} \|\nabla \psi\|_{\omega_E} \quad \forall E \in \mathcal{E}_h^n,
\]

where the constant \( C^* > 0 \) only depends on the minimum angle of the mesh \( T^n \), see \([7]\).

We consider the following space-time discretization of \((14)\): for \(Y_{\varepsilon,h}^{n-1} \) a \( \mathbb{V}_h^{n-1}\)-valued random variable and \( n = 1, 2, \ldots \), find the \( \mathcal{F}_{t_n}\)-measurable random variable \( Y_{\varepsilon,h}^n \) taking values in \( \mathbb{V}_h^n \), such that \( \mathbb{P}\)-almost surely

\[
\left( \frac{Y_{\varepsilon,h}^n - Y_{\varepsilon,h}^{n-1}}{\tau_n}, \varphi_h \right) = - \left( \frac{\nabla X_{\varepsilon,h}^n}{|\nabla X_{\varepsilon,h}^n|}, \nabla \varphi_h \right) - \lambda (X_{\varepsilon,h}^n - g_h, \varphi_h) \quad \forall \varphi_h \in \mathbb{V}_h^n.
\]

The function \( X_{\varepsilon,h}^n \) is the solution of the implicit fully-discrete approximation of \((3)\) which is defined as follows: set \( X_{\varepsilon,h}^0 := \mathcal{P}_0 x^0 \in \mathbb{V}_h^0 \) and determine \( X_{\varepsilon,h}^n \in \mathbb{V}_h^n \), as the solution of

\[
(X_{\varepsilon,h}^n, \varphi) = (X_{\varepsilon,h}^{n-1}, \varphi) - \tau_n \left( \frac{\nabla X_{\varepsilon,h}^n}{|\nabla X_{\varepsilon,h}^n|}, \nabla \varphi_h \right) - \tau_n \lambda (X_{\varepsilon,h}^n - g_h, \varphi_h)
\]

\[
+ (\sigma(t_{n-1}) \Delta W_n, \varphi_h) \quad \forall \varphi_h \in \mathbb{V}_h^n,
\]

where \( \Delta W_n := W(t_n) - W(t_{n-1}) \).

From \((17)\) and \((18)\) we deduce the discrete counterpart of the transformation \((13)\):

\[
Y_{\varepsilon,h}^n = X_{\varepsilon,h}^n - \mathcal{P}_n \left( \sum_{i=1}^n (\mathcal{P}_{n-1} \circ \cdots \circ \mathcal{P}_i)(\sigma(t_{i-1})) \Delta W_i \right).
\]

On denoting \( \tilde{\sigma}_{n-1}^i := (\mathcal{P}_{n-1} \circ \cdots \circ \mathcal{P}_i)(\sigma(t_{i-1})) \) the above transformation can be formulated as

\[
Y_{\varepsilon,h}^n = X_{\varepsilon,h}^n - \mathcal{P}_n \left( \sum_{i=1}^n \tilde{\sigma}_{n-1}^{i-1} \Delta W_i \right).
\]

Further, we denote the stochastic integral as

\[
\Sigma(t) := \int_0^t \sigma(s) \, dW(s),
\]

and its discrete counterpart as

\[
\Sigma^n := \mathcal{P}_n \left( \sum_{i=1}^n \tilde{\sigma}_{n-1}^{i-1} \Delta W_i \right).
\]

We define continuous piecewise linear time-interpolants of \( \{Y_{\varepsilon,h}^n\}_{n \in \mathbb{N}} \), \( \{X_{\varepsilon,h}^n\}_{n \in \mathbb{N}} \) and \( \{\Sigma^n\}_{n \in \mathbb{N}} \) on \([0,T]\) as follows: for \( t \in [t_{n-1}, t_n] \) we set

\[
Y_{\varepsilon,h}^T(t) := \frac{t - t_{n-1}}{\tau_n} Y_{\varepsilon,h}^n + \frac{t_n - t}{\tau_n} Y_{\varepsilon,h}^{n-1},
\]

\[
X_{\varepsilon,h}^T(t) := \frac{t - t_{n-1}}{\tau_n} X_{\varepsilon,h}^n + \frac{t_n - t}{\tau_n} X_{\varepsilon,h}^{n-1},
\]
and analogically, recalling (20),
$$
\Sigma^\tau(t) := \frac{t - t_{n-1}}{\tau_n} \Sigma^n + \frac{t_n - t}{\tau_n} \Sigma^{n-1}
= \frac{t - t_{n-1}}{\tau_n} p_n \left( \sum_{i=1}^{n} \bar{\sigma}_{i-1} \Delta W_i \right) + \frac{t_n - t}{\tau_n} p_{n-1} \left( \sum_{i=1}^{n-1} \bar{\sigma}_{i-2} \Delta W_i \right).
$$

On noting (20) and (23), (24), (25), we deduce that
$$
Y_{\tau,\epsilon,h}(t) = X_{\tau,\epsilon,h}(t) - \Sigma^\tau(t).
$$

4. A posteriori estimates

In this section we consider the numerical solution \( \{ X_{\epsilon,h}^n \}_{n=0}^N \) of (18) obtained using possibly non-equidistant time grid with variable time steps \( \{ \tau_n \}_{n=0}^N \) and locally adapted spatial meshes \( \{ T_h^n \}_{n=0}^N \). We derive residual a posteriori estimates for the numerical solution \( \{ X_{\epsilon,h}^n \}_{n=0}^N \) which allow to control the approximation error to the variational solution \( X_\epsilon \) of (3) as well as to the SVI solution \( X_1 \) of (1). The constants in the derived a posteriori estimates only depend on the given data \( x_0, g, O, T \) and the shape of the meshes \( \{ T_h^n \}_{n=0}^N \).

We define the (random) interior residual \( \{ R_h^n \}_{n=0}^N \) as
$$
R_h^n := \lambda (g_h - X_{\epsilon,h}^n) - \frac{X_{\epsilon,h}^n - X_{\epsilon,h}^{n-1}}{\tau_n} p_n (t_{n-1}) \Delta W_i.
$$
The jump residual \( J_E^n \) across an interior face \( E = \partial K_1 \cap \partial K_2 \in \mathcal{E}_{h,O}^n \) is defined as
$$
J_E^n := \frac{1}{2} \left( \frac{\nabla X_{\epsilon,h}^n}{|\nabla X_{\epsilon,h}^n|} |_{K_1} - \frac{\nabla X_{\epsilon,h}^n}{|\nabla X_{\epsilon,h}^n|} |_{K_2} \right) \cdot \nu_E,
$$
where we use the convention that the unit normal vector \( \nu_E \) to \( E \) points from \( K_1 \) to \( K_2 \).

By applying the integration by parts formula on each element \( K \in T_h^n \), we deduce the following equality
$$
\left( \frac{\nabla X_{\epsilon,h}^n}{|\nabla X_{\epsilon,h}^n|} \cdot \nabla \varphi \right) = \sum_{E \in \mathcal{E}_{h,O}^n} \int_E J_E^n \varphi \, dS \quad \forall \varphi \in H_0^1,
$$
$$
\nabla \cdot \frac{\nabla X_{\epsilon,h}^n}{|\nabla X_{\epsilon,h}^n|} |_{K} = 0, \text{ by the linearity of } X_{\epsilon,h}^n |_K.
$$
We define the time error indicators \( \eta_{\text{time},1}^n, \eta_{\text{time},2}^n \) to be the \( \mathbb{R} \)-valued random variables
$$
\eta_{\text{time},1}^n := \| X_{\epsilon,h}^n - X_{\epsilon,h}^{n-1} \|,
$$
and
$$
\eta_{\text{time},2}^n := \| \nabla (X_{\epsilon,h}^n - X_{\epsilon,h}^{n-1}) \|.
$$
The correspondingly space error indicators \( \eta_{\text{space},1}^n, \eta_{\text{space},2}^n \) are defined as
$$
\eta_{\text{space},1}^n := h_T \| R_h^n \| T.
$$
and

\begin{equation}
\eta_{\text{space},2}^n := h_E^{\frac{1}{2}} \| J_E^n \| E.
\end{equation}

The terms \( \eta_{\text{space},2}^n, \eta_{\text{space},1}^n \) are often referred to as the local space error indicators. We define the time noise error indicators \( \eta_{\text{noise},1}^n, \eta_{\text{noise},2}^n, \eta_{\text{noise},3}^n \),

\begin{align}
\eta_{\text{noise},1}^n &:= \sum_{n=1}^{m} \tau_n \sum_{i=1}^{n} \int_{t_{i-1}}^{t_i} \| \nabla (s) - \sigma(t_{i-1}) \| \, ds + \sum_{n=1}^{m} \int_{t_{n-1}}^{t_n} \int_{t}^{t_n} \| \sigma(s) \| \, ds \, dt \\
&+ \sum_{n=1}^{m} \tau_n^2 \| \sigma(t_{n-1}) \| \, ds + \sum_{n=1}^{m} \tau_n \sum_{i=1}^{n} \tau_i \| \sigma(t_{n-1}) - \tilde{\sigma}^{n-1}_i \| \, ds \\
&+ \sum_{n=1}^{m} \tau_n \sum_{i=1}^{n} \tau_i \| \mathcal{P}_n(\tilde{\sigma}^{i-1}_{n-1} - \tilde{\sigma}^{i-1}_{n-2}) \| \, ds + \sum_{n=1}^{m} \tau_n \sum_{i=1}^{n} \tau_i \| (\mathcal{P}_n - \mathcal{P}_{n-1}) \tilde{\sigma}^{i-1}_{n-2} \| \, ds \\
\end{align}

\begin{align}
\eta_{\text{noise},2}^n &:= \sum_{n=1}^{m} \tau_n \sum_{i=1}^{n} \int_{t_{i-1}}^{t_i} \| \nabla (s) - \sigma(t_{i-1}) \| \, ds + \sum_{n=1}^{m} \int_{t_{n-1}}^{t_n} \int_{t}^{t_n} \| \nabla (s) \| \, ds \, dt \\
&+ \sum_{n=1}^{m} \tau_n^2 \| \nabla \sigma(t_{n-1}) \| \, ds + \sum_{n=1}^{m} \tau_n \sum_{i=1}^{n} \tau_i \| \nabla (\sigma(t_{i-1}) - \tilde{\sigma}^{i-1}_n) \| \, ds \\
&+ \sum_{n=1}^{m} \tau_n \sum_{i=1}^{n} \tau_i \| \nabla \mathcal{P}_n(\tilde{\sigma}^{i-1}_{n-1} - \tilde{\sigma}^{i-1}_{n-2}) \| \, ds + \sum_{n=1}^{m} \tau_n \sum_{i=1}^{n} \tau_i \| \nabla (\mathcal{P}_n - \mathcal{P}_{n-1}) \tilde{\sigma}^{i-1}_{n-2} \| \, ds
\end{align}

and

\begin{align}
\eta_{\text{noise},3}^n &:= \sum_{n=1}^{m} \int_{t_{n-1}}^{t_n} \| \sigma(s) - \sigma(t_{n-1}) \| \, ds + \sum_{n=1}^{m} \tau_n \| \sigma(t_{n-1}) - \tilde{\sigma}^{n-1}_m \| \, ds \\
&+ \sum_{n=1}^{m} \tau_n \| \sigma(t_{n-1}) - \tilde{\sigma}^{n-1}_m \| \, ds.
\end{align}

**Theorem 4.1.** Let \( X^\varepsilon \) be the variational solution of (10) and let \( \{ X^n_{\varepsilon,h} \}_{n=0}^N \) be the solution of (18) and \( X^\tau \) be the corresponding interpolation defined in (24). There exists a constant \( C > 0 \) depending on \( T, |\mathcal{O}|, \sigma, x_0 \) and a constant \( \tilde{C} > 0 \) which additionally depends on the shape of the meshes \( \{ T^n_h \}_{n=0}^N \) and \( \lambda \) such that the following a posteriori estimates holds for
any integer \(1 \leq m \leq N\):

\[
E[\|X(\tau_m) - X(\tau_{\epsilon,h})(\tau_m)\|^2] + E\left[\sum_{n=1}^{m} \int_{t_{n-1}}^{t_n} \int_{\Omega} \left(1 - \frac{|\nabla X(\epsilon,h)(s)|}{|\nabla X(\epsilon,h)(s)|_{\epsilon}}\right) \frac{|\nabla(X(\epsilon) - X(\epsilon,h)(s))|_{\epsilon}}{|\nabla X(\epsilon)|_{\epsilon}} \, dx \, ds\right]
\]

\[
\leq C \left(E\left[\|Y(0) - Y(\epsilon,h)\|^2\right] + \lambda \|g - g_h\|^2\right) + \mathcal{C} \left\{ \tau_n \sum_{n=1}^{m} E\left[\left(\eta_{\text{time,}1}^n\right)^2 + \eta_{\text{time,}2}^n\right]\right. \\
+ \sum_{n=1}^{m} \tau_n \left(E\left[\left(\sum_{T \in T_{\epsilon,h}} \eta_{\text{space,}1}^n\right)^2\right]\right) \right. \\
+ \sum_{n=1}^{m} \tau_n \left(E\left[\left(\sum_{E \in E_{\epsilon,h,0}} \left(\eta_{\text{space,}2}^n\right)^2\right)\right]\right) \\
+ \sum_{n=1}^{m} \tau_n E\left[\left(\sum_{T \in T_{\epsilon,h}} \left(\eta_{\text{space,}1}^n\right)^2\right) + \sum_{E \in E_{\epsilon,h,0}} \left(\eta_{\text{space,}2}^n\right)^2\right]\right. \\
+ \sum_{n=1}^{m} \tau_n E\left[\eta_{\text{noise,}1}^n + \eta_{\text{noise,}2}^n + \eta_{\text{noise,}3}^n\right] \left. + \left(E\left[\eta_{\text{noise,}2}^n\right]\right)^{1/2}\right) \right\}.
\]

**Proof.** On noting (14), (17) and (27) we deduce that for a.a. \(t \in (t_{n-1}, t_n]\), \(P\)-a.s., and for any \(\varphi \in H_0^1\) and \(\varphi_h \in V_h^n\)

\[
\langle \partial_t(Y(\epsilon) - Y(\epsilon,h)), \varphi \rangle + \left(\frac{\nabla X(\epsilon)}{|\nabla X(\epsilon)|_{\epsilon}} - \frac{\nabla X(\epsilon,h)}{|\nabla X(\epsilon,h)|_{\epsilon}}, \nabla \varphi \right) + \lambda(\varphi - X(\epsilon,h), \varphi) - \lambda(g - g_h, \varphi)
\]

\[
= (R_h^n, \varphi - \varphi_h) - \left(\frac{\nabla X(\epsilon,h)}{|\nabla X(\epsilon,h)|_{\epsilon}}, \nabla (\varphi - \varphi_h)\right).
\]

We set \(\varphi = [Y(\epsilon) - Y(\epsilon,h)](\omega, t) \in H_0^1\) and \(\varphi_h = \Pi_n [Y(\epsilon) - Y(\epsilon,h)](\omega, t) \in V_h^n\) in (37), to obtain after integration by parts and taking the expectation, substitute (13) and (26) and integrate
in time from 0 to $t_m$

$$\frac{1}{2} \mathbb{E} \left[ \left\| Y^\varepsilon(t_m) - Y^\varepsilon_{\varepsilon,h}(t_m) \right\|^2 \right] + \mathbb{E} \left[ \sum_{n=1}^{m} \int_{t_{n-1}}^{t_n} \left( \frac{\nabla X^\varepsilon}{\| \nabla X^\varepsilon \|}, \nabla (X^\varepsilon - X^\varepsilon_{\varepsilon,h}) \right) ds \right]$$

$$+ \mathbb{E} \left[ \sum_{n=1}^{m} \int_{t_{n-1}}^{t_n} \lambda (X^\varepsilon - X^\varepsilon_{\varepsilon,h}, X^\varepsilon - X^\varepsilon_{\varepsilon,h}) - \lambda (g - g_h, X - X^\varepsilon_{\varepsilon,h}) ds \right]$$

$$= \frac{1}{2} \mathbb{E} \left[ \left\| Y(0) - Y^\varepsilon_{\varepsilon,h}(0) \right\|^2 \right] + \mathbb{E} \left[ \sum_{n=1}^{m} \int_{t_{n-1}}^{t_n} \left( \frac{\nabla X^\varepsilon}{\| \nabla X^\varepsilon \|}, \nabla (\Sigma - \Sigma^\tau) \right) ds \right]$$

$$+ \mathbb{E} \left[ \sum_{n=1}^{m} \int_{t_{n-1}}^{t_n} \lambda (X^\varepsilon - X^\varepsilon_{\varepsilon,h}, \Sigma - \Sigma^\tau) - \lambda (g - g_h, \Sigma - \Sigma^\tau) ds \right]$$

$$+ \mathbb{E} \left[ \sum_{n=1}^{m} \int_{t_{n-1}}^{t_n} \left( R^n_h, X^\varepsilon - X^\varepsilon_{\varepsilon,h} - \Pi_n [X^\varepsilon - X^\varepsilon_{\varepsilon,h}] \right) ds \right]$$

$$+ \mathbb{E} \left[ \sum_{n=1}^{m} \int_{t_{n-1}}^{t_n} \sum_{E \in \mathcal{E}_h^0} \int_E J^n_E (X^\varepsilon - X^\varepsilon_{\varepsilon,h} - \Pi_n [X^\varepsilon - X^\varepsilon_{\varepsilon,h}] ) ds ds \right]$$

$$- \mathbb{E} \left[ \sum_{n=1}^{m} \int_{t_{n-1}}^{t_n} \left( R^n_h, \Sigma - \Sigma^\tau - \Pi_n [\Sigma - \Sigma^\tau] \right) ds \right]$$

$$- \mathbb{E} \left[ \sum_{n=1}^{m} \int_{t_{n-1}}^{t_n} \sum_{E \in \mathcal{E}_h^0} \int_E J^n_E (\Sigma - \Sigma^\tau - \Pi_n [\Sigma - \Sigma^\tau] ) ds ds \right]$$

(38)

We rewrite the scalar product in the last sum at the left hand side in (38)

$$\lambda (X^\varepsilon - X^\varepsilon_{\varepsilon,h}, X^\varepsilon - X^\varepsilon_{\varepsilon,h}) + \lambda (g - g_h, X^\varepsilon - X^\varepsilon_{\varepsilon,h})$$

(39) $$= \lambda \left\| X^\varepsilon - X^\varepsilon_{\varepsilon,h} \right\|^2 + \lambda (X^\varepsilon_{\varepsilon,h} - X^\varepsilon_{\varepsilon,h}, X^\varepsilon - X^\varepsilon_{\varepsilon,h}) + \lambda (g - g_h, X^\varepsilon - X^\varepsilon_{\varepsilon,h}).$$

We put the second and the third term of the equality (39) on the right hand side in (38) and estimate it as follows

$$- \lambda (X^\varepsilon_{\varepsilon,h} - X^\varepsilon_{\varepsilon,h}, X^\varepsilon - X^\varepsilon_{\varepsilon,h}) - \lambda (g - g_h, X^\varepsilon - X^\varepsilon_{\varepsilon,h})$$

$$\leq \frac{\lambda}{2} \left\| X^\varepsilon - X^\varepsilon_{\varepsilon,h} \right\|^2 + \frac{\lambda}{2} \left\| X^\varepsilon_{\varepsilon,h} - X^\varepsilon_{\varepsilon,h} \right\|^2 + 4\lambda \left\| g - g_h \right\|^2 + \frac{\lambda}{16} \left\| X^\varepsilon - X^\varepsilon_{\varepsilon,h} \right\|^2.$$
We estimate the terms in the first sum on the right hand side in (38). At first, we split the scalar product

\[
\begin{aligned}
& \left( \frac{\nabla X^\varepsilon}{|\nabla X^\varepsilon|_\varepsilon} - \frac{\nabla X^n_{\varepsilon,h}}{|\nabla X^n_{\varepsilon,h}|_\varepsilon}, \nabla X^\varepsilon - \nabla X^\tau_{\varepsilon,h} \right) \\
= & \left( \frac{\nabla X^\varepsilon}{|\nabla X^\varepsilon|_\varepsilon} - \frac{\nabla X^n_{\varepsilon,h}}{|\nabla X^n_{\varepsilon,h}|_\varepsilon}, \nabla X^\varepsilon - \nabla X^n_{\varepsilon,h} \right) + \left( \frac{\nabla X^\varepsilon}{|\nabla X^\varepsilon|_\varepsilon} - \frac{\nabla X^n_{\varepsilon,h}}{|\nabla X^n_{\varepsilon,h}|_\varepsilon}, \nabla X^n_{\varepsilon,h} - \nabla X^\tau_{\varepsilon,h} \right)
\end{aligned}
\]

Then rewrite and estimate the second term on the r.h.s in (40) with the elementary inequality \( ||a| - |b|| \leq |a - b| \)

\[
\begin{aligned}
& \int_{\partial} \left| \frac{\nabla(X^\varepsilon - X^n_{\varepsilon,h})}{|\nabla X^\varepsilon|_\varepsilon} \right|^2 \left| \frac{\nabla(X^n_{\varepsilon,h})}{|\nabla X^n_{\varepsilon,h}|_\varepsilon} \right|^2 + \left( \frac{1}{|\nabla X^\varepsilon|_\varepsilon} - \frac{1}{|\nabla X^n_{\varepsilon,h}|_\varepsilon} \right) \nabla X^n_{\varepsilon,h} \nabla(X^\varepsilon - X^n_{\varepsilon,h}) \ dx \\
\geq & \int_{\partial} \frac{|\nabla(X^\varepsilon - X^n_{\varepsilon,h})|^2}{|\nabla X^\varepsilon|_\varepsilon^2} - \frac{|\nabla X^n_{\varepsilon,h}|}{|\nabla X^\varepsilon|_\varepsilon |\nabla X^n_{\varepsilon,h}|_\varepsilon} |\nabla(X^\varepsilon - X^n_{\varepsilon,h})|^2 \ dx \\
\geq & \int_{\partial} \left( 1 - \frac{|\nabla X^n_{\varepsilon,h}|}{|\nabla X^\varepsilon|_\varepsilon} \right) \frac{|\nabla(X^\varepsilon - X^n_{\varepsilon,h})|^2}{|\nabla X^\varepsilon|_\varepsilon} \ dx.
\end{aligned}
\]
We substitute (41) into (38) and obtain
\[
\frac{1}{2} \mathbb{E} \left[ \left\| Y(t_m) - Y^{\tau}_{\varepsilon,h}(t_m) \right\|^2 \right] + \mathbb{E} \left[ \sum_{n=1}^{m} \int_{t_{n-1}}^{t_n} \left( 1 - \frac{\nabla X^n_{\varepsilon,h}}{|\nabla X^n_{\varepsilon,h}|} \right) \frac{\nabla (X^\varepsilon - X^n_{\varepsilon,h})^2}{|\nabla X^\varepsilon|} \, dx \, ds \right] \\
+ \mathbb{E} \left[ \sum_{n=1}^{m} \int_{t_{n-1}}^{t_n} \lambda \left\| X^\varepsilon - X^n_{\varepsilon,h} \right\|^2 \, ds \right] \\
\leq \frac{1}{2} \mathbb{E} \left[ \left\| Y(0) - Y^{\tau}_{\varepsilon,h}(0) \right\|^2 \right] + \mathbb{E} \left[ \sum_{n=1}^{m} \int_{t_{n-1}}^{t_n} \left( \frac{\nabla X^\varepsilon}{|\nabla X^\varepsilon|} - \frac{\nabla X^n_{\varepsilon,h}}{|\nabla X^n_{\varepsilon,h}|} \right) \nabla (X^\varepsilon - X^n_{\varepsilon,h}) \, ds \right] \\
+ \mathbb{E} \left[ \sum_{n=1}^{m} \int_{t_{n-1}}^{t_n} \frac{\lambda}{2} \left\| X^\varepsilon - X^n_{\varepsilon,h} \right\|^2 \, ds \right] + \mathbb{E} \left[ \sum_{n=1}^{m} \int_{t_{n-1}}^{t_n} \frac{\lambda}{16} \left\| X^\varepsilon - X^n_{\varepsilon,h} \right\|^2 \, ds \right] \\
+ \mathbb{E} \left[ \sum_{n=1}^{m} \int_{t_{n-1}}^{t_n} \left( \frac{\nabla X^\varepsilon}{|\nabla X^\varepsilon|} - \frac{\nabla X^n_{\varepsilon,h}}{|\nabla X^n_{\varepsilon,h}|} \right) \nabla (\Sigma - \Sigma^{\tau}) \, ds \right] \\
+ \mathbb{E} \left[ \sum_{n=1}^{m} \int_{t_{n-1}}^{t_n} \lambda (X^\varepsilon - X^n_{\varepsilon,h}, \Sigma - \Sigma^{\tau}) \, ds \right] + \mathbb{E} \left[ \sum_{n=1}^{m} \int_{t_{n-1}}^{t_n} \lambda (g - g_h, \Sigma - \Sigma^{\tau}) \, ds \right] \\
+ \mathbb{E} \left[ \sum_{n=1}^{m} \int_{t_{n-1}}^{t_n} \left( P^n_{h}, X^\varepsilon - X^n_{\varepsilon,h} - \Pi_n \left[ X^\varepsilon - X^n_{\varepsilon,h} \right] \right) \, ds \right] \\
+ \mathbb{E} \left[ \sum_{n=1}^{m} \int_{t_{n-1}}^{t_n} \sum_{E \in \mathcal{E}^{n}_{h,0}} \int_{E} \left. \mathbb{J}_E^n \left( \nabla (X^\varepsilon - X^n_{\varepsilon,h} - \Pi_n \left[ X^\varepsilon - X^n_{\varepsilon,h} \right]) \right) \right| \, dx \, ds \right] \\
- \mathbb{E} \left[ \sum_{n=1}^{m} \int_{t_{n-1}}^{t_n} \left( P^n_{h}, \Sigma - \Sigma^{\tau} - \Pi_n \left[ \Sigma - \Sigma^{\tau} \right] \right) \, ds \right] \\
- \mathbb{E} \left[ \sum_{n=1}^{m} \int_{t_{n-1}}^{t_n} \sum_{E \in \mathcal{E}^{n}_{h,0}} \int_{E} \left. \mathbb{J}_E^n \left( \Sigma - \Sigma^{\tau} - \Pi_n \left[ \Sigma - \Sigma^{\tau} \right] \right) \right| \, dx \, ds \right] \\
= \frac{1}{2} \left\| Y(0) - Y^{\tau}_{\varepsilon,h}(0) \right\|^2 + \sum_{i=1}^{12} A_i.
\]

We use the bound \( \left| \frac{\nabla}{|\nabla X^\varepsilon|} \right| < 1 \) to estimate \( A_1 \) with the Cauchy-Schwarz inequality
\[
A_1 = \mathbb{E} \left[ \sum_{n=1}^{m} \int_{t_{n-1}}^{t_n} \left( \frac{\nabla X^\varepsilon}{|\nabla X^\varepsilon|} - \frac{\nabla X^n_{\varepsilon,h}}{|\nabla X^n_{\varepsilon,h}|} \right) \nabla X^n_{\varepsilon,h} - \nabla X^n_{\varepsilon,h} \right) \, ds \right] \leq 2 \mathbb{E} \left[ \sum_{n=1}^{m} \mathbb{T}_n \mathbb{H}_{\text{time},2} \right].
On recalling the definition (24) of $X^*_{ε,h}$, we see that

$$A_3 \leq \frac{λ}{2} \mathbb{E} \left[ \sum_{n=1}^{m} \tau_n (\eta_{time,1}^n)^2 \, ds \right].$$

Similarly, we estimate $A_5$ after applying Cauchy-Schwarz and Youngs inequality

$$A_5 \leq \mathbb{E} \left[ \frac{λ}{4} \sum_{n=1}^{m} \int_{t_{n-1}}^{t_n} \left\| X^ε - X^*_{ε,h} \right\|^2 \, ds + \frac{3λ}{8} \sum_{n=1}^{m} \tau_n (\eta_{time,1}^n)^2 \, ds \right].$$

On recalling the definitions (21) and (25) of $Σ$ and $Σ^*$ we deduce

$$\mathbb{E} \left[ \sum_{n=1}^{m} \int_{t_{n-1}}^{t_n} \left\| Σ(s) - Σ^*(s) \right\|^2 \, ds \right] \leq C \mathbb{E} \left[ \sum_{n=1}^{m} \int_{t_{n-1}}^{t_n} \left\| \int_{t_{n-1}}^{s} \sigma(r) \, dW(r) - \frac{s - t_n}{τ_n} \sigma(t_{n-1}) \Delta W_n - \sum_{i=1}^{n} \sigma(t_{i-1}) \Delta W_i \right\|^2 \, ds \right] + C \mathbb{E} \left[ \sum_{n=1}^{m} \int_{t_{n-1}}^{t_n} \left\| \frac{s - t_n}{τ_n} \left[ \sigma(t_{n-1}) - \bar{\sigma}^n_{n-1} \right] \Delta W_n \right\|^2 \, ds \right] + C \mathbb{E} \left[ \sum_{n=1}^{m} \tau_n \left\| \sum_{i=1}^{n} \sigma(t_{i-1}) \Delta W_i - \mathcal{P}_n \left( \sum_{i=1}^{n} \bar{\sigma}^{i-1}_{n-1} \Delta W_i \right) \right\|^2 \right] + C \mathbb{E} \left[ \sum_{n=1}^{m} \int_{t_{n-1}}^{t_n} \left\| \frac{t_n - s}{τ_n} \mathcal{P}_n \left( \sum_{i=1}^{n-1} \bar{\sigma}^{i-1}_{n-2} - \bar{\sigma}^{i-1}_{n-2} \right) \Delta W_i \right\|^2 \, ds \right] + C \mathbb{E} \left[ \sum_{n=1}^{m} \int_{t_{n-1}}^{t_n} \left\| \frac{t_n - s}{τ_n} \left( \mathcal{P}_n - \mathcal{P}_{n-1} \right) \left( \sum_{i=1}^{n-1} \bar{\sigma}^{i-1}_{n-2} \Delta W_i \right) \right\|^2 \, ds \right] =: B_1 + B_2 + B_3 + B_4 + B_5.
We estimate each term above separately. After applying Cauchy-Schwarz and Young’s inequalities and Itô’s isometry we obtain

\[ B_1 \leq \mathbb{E} \left[ \sum_{n=1}^{m} \tau_n \left\| \sum_{i=1}^{n} \int_{t_{i-1}}^{t_i} \sigma(r) - \sigma(t_{i-1}) \, dW(r) \right\|^2 \right] \]
\[ + \mathbb{E} \left[ \sum_{n=1}^{m} \int_{t_{n-1}}^{t_n} \left\| \int_{s}^{t_n} \sigma(r) \, dW(r) - \frac{s-t_n}{\tau_n} \sigma(t_{n-1}) \Delta W_n \right\|^2 \, ds \right] \]
\[ \leq C \sum_{n=1}^{m} \int_{t_{n-1}}^{t_n} \mathbb{E} \left[ \sum_{i=1}^{n} \int_{t_{i-1}}^{t_i} \|\sigma(r) - \sigma(t_{i-1})\|^2 \, dr \right] \, ds \]
\[ + C \sum_{n=1}^{m} \int_{t_{n-1}}^{t_n} \mathbb{E} \left[ \int_{s}^{t_n} \|\sigma(r)\|^2 \, dr \right] \, ds + C \sum_{n=1}^{m} \tau_n^2 \mathbb{E} \left[ \|\sigma(t_{n-1})\|^2 \right]. \]

Similarly, we estimate \( B_2, \ldots, B_4 \) as

\[ B_2 \leq C \mathbb{E} \left[ \sum_{n=1}^{m} \tau_n^2 \|\sigma(t_{n-1}) - \tilde{\sigma}^{n-1}_n\|^2 \, ds \right], \]
\[ B_3 \leq C \mathbb{E} \left[ \sum_{n=1}^{m} \tau_n \sum_{i=1}^{n-1} \tau_i \|\sigma(t_{i-1}, x) - \tilde{\sigma}_n^{i-1}\|^2 \right], \]
and

\[ B_4 \leq C \mathbb{E} \left[ \sum_{n=1}^{m} \tau_n \sum_{i=1}^{n-1} \tau_i \|\mathcal{P}_n(\tilde{\sigma}^{i-1}_n - \tilde{\sigma}^{i-1}_{n-2})\|^2 \, ds \right]. \]

The last term is estimated as

\[ B_5 \leq C \mathbb{E} \left[ \sum_{n=1}^{m} \tau_n \sum_{i=1}^{n-1} \tau_i \|\mathcal{P}_n - \mathcal{P}_{n-1}\|_{\tilde{\sigma}^{i-1}_{n-2}}^2 \right]. \]

After substituting the above estimates for \( B_1, \ldots, B_5 \) into (43) we arrive at

\[ \mathbb{E} \left[ \sum_{n=1}^{m} \int_{t_{n-1}}^{t_n} \|\Sigma(s) - \Sigma^\tau(s)\|^2 \, ds \right] \leq \mathbb{E} \left[ \eta_{\text{noise}, 1}^n \right]. \]

Analogically one can show

\[ \mathbb{E} \left[ \sum_{n=1}^{m} \int_{t_{n-1}}^{t_n} \|\nabla(\Sigma(s) - \Sigma^\tau(s))\|^2 \, ds \right] \leq \mathbb{E} \left[ \eta_{\text{noise}, 2}^n \right]. \]
Next, the Cauchy-Schwarz inequality, the bound $\left| \frac{\nabla}{|\nabla\epsilon|} \right| < 1$ yield

$$A_6 \leq E \left[ \sum_{n=1}^{m} \int_{t_{n-1}}^{t_n} \left( \frac{\nabla X^\epsilon}{|\nabla X^\epsilon|} - \frac{\nabla X_{\epsilon,h}^n}{|\nabla X_{\epsilon,h}^n|}, \nabla(\Sigma - \Sigma') \right) \right] \leq 2T |\mathcal{O}| \left( E \left[ \eta_{\text{noise,2}}^n \right] \right)^{\frac{1}{2}}.$$  

We estimate $A_7, A_8$ with the Cauchy-Schwarz and Young inequality

$$A_7 \leq \frac{\lambda}{4} E \left[ \sum_{n=1}^{m} \int_{t_{n-1}}^{t_n} \left( \| X^\epsilon - X_{\epsilon,h}^n \|^2 \right) \right] + \lambda E \left[ \eta_{\text{noise,1}}^n \right],$$

$$A_8 \leq \frac{\lambda}{2} T E \left[ \| g - gh \|^2 \right] + \frac{\lambda}{2} E \left[ \eta_{\text{noise,1}}^n \right].$$

We use the interpolation estimate (15) and (12) to estimate $A_9$

$$A_9 \leq E \left[ \sum_{n=1}^{m} \int_{t_{n-1}}^{t_n} \left( \| R_{h}^n \|_{T} \left\| X^\epsilon - X_{\epsilon}^\tau - \Pi_n [X^\epsilon - X_{\epsilon}^\tau] \right\|_{T} \right) \right]
\leq C \left( E \left[ \|\nabla x_0\|^2 \right] + \|\nabla g\|^2 + E \left[ \sup_{t \in [0,T]} \|\nabla \sigma(t)\|^2 \right] \right)^{\frac{1}{2}}
\times \sum_{n=1}^{m} (\tau_n)^{\frac{1}{2}} \left( E \left[ \int_{t_{n-1}}^{t_n} \left( \sum_{T \in T_h^n} C^* h_T \| R_{h}^n \|_{T} \right)^2 \right] \right)^{\frac{1}{2}}
\leq CC^* \sum_{n=1}^{m} \tau_n \left( E \left[ \left( \sum_{T \in T_h^n} \eta_{\text{space,1}}^n \right)^{2} \right] \right)^{\frac{1}{2}},$$

where we also used that $\Pi_n X_{\epsilon}^\tau = X_{\epsilon}^\tau$ by the projection property of $\Pi_n$, cf., [7].

Analogically we estimate $A_{10}$ as

$$A_{10} \leq CC^* \sum_{n=1}^{m} \tau_n \left( E \left[ \left( \sum_{E \in \mathcal{E}_h^2} \eta_{\text{space,2}}^n \right)^{2} \right] \right)^{\frac{1}{2}}.$$
Next, we use (15) and (16) and (44) and obtain after applying the Cauchy-Schwarz and Young’s inequalities and Itô’s isometry

\[
A_{11} \leq \frac{1}{2} C^* \mathbb{E} \left[ \sum_{n=1}^{m} \tau_n \sum_{\epsilon \in \mathcal{E}_{\nu,0}} (\eta_{\text{space},1}^n)^2 \right] + \mathbb{E} \left[ \eta_{\text{noise},2}^n \right],
\]

\[
A_{12} \leq \frac{1}{2} C^* \mathbb{E} \left[ \sum_{n=1}^{m} \tau_n \sum_{\epsilon \in \mathcal{E}_{\nu,0}} (\eta_{\text{space},2}^n)^2 \right] + \mathbb{E} \left[ \eta_{\text{noise},2}^n \right].
\]

After substituting the above estimates for \( A_1, \ldots, A_{12} \) into (42) we arrive at

\[
\frac{1}{2} \mathbb{E} \left[ \| Y(t_m) - Y^\tau_{\epsilon,h}(t_m) \|^2 \right] + \mathbb{E} \left[ \sum_{n=1}^{m} \int_{t_{n-1}}^{t_n} \int_{\mathcal{O}} \left( 1 - \frac{\nabla X^\nu_{\epsilon,h}^n}{\nabla X^\epsilon_{\epsilon,h}^n} \right) \frac{|\nabla(X^\epsilon - X^\nu_{\epsilon,h})|^2}{|\nabla X^\epsilon|^2} \, dx \, ds \right] \\
\leq \frac{1}{2} \mathbb{E} \left[ \| Y(0) - Y^\tau_{\epsilon,h}(0) \|^2 \right] + \frac{7\lambda}{8} \mathbb{E} \left[ \sum_{n=1}^{m} \tau_n (\eta_{\text{time},1}^n)^2 \, ds \right] + 2 \mathbb{E} \left[ \sum_{n=1}^{m} \tau_n \eta_{\text{noise},2}^n \right] \\
+ CC^* \sum_{n=1}^{m} \tau_n \left( \mathbb{E} \left[ \left( \sum_{\epsilon \in \mathcal{E}_{\nu,0}} \eta_{\text{space},1}^n \right)^2 \right] \right)^{\frac{1}{2}} + \frac{1}{2} C^* \mathbb{E} \left[ \sum_{n=1}^{m} \tau_n \sum_{\epsilon \in \mathcal{E}_{\nu,0}} (\eta_{\text{space},1}^n)^2 \right] \\
+ CC^* \sum_{n=1}^{m} \tau_n \left( \mathbb{E} \left[ \left( \sum_{\epsilon \in \mathcal{E}_{\nu,0}} \eta_{\text{space},2}^n \right)^2 \right] \right)^{\frac{1}{2}} + \frac{1}{2} C^* \mathbb{E} \left[ \sum_{n=1}^{m} \tau_n \sum_{\epsilon \in \mathcal{E}_{\nu,0}} (\eta_{\text{space},2}^n)^2 \right] \\
+ \frac{3\lambda}{2} \mathbb{E} \left[ \eta_{\text{noise},1}^n \right] + 2 \mathbb{E} \left[ \eta_{\text{noise},2}^n \right] + 2 T |\mathcal{O}| \left( \mathbb{E} \left[ \eta_{\text{noise},2}^n \right] \right)^\frac{1}{2} \\
+ \frac{\lambda}{2} T \mathbb{E} \left[ \| g - g_n \|^2 \right].
\]

On recalling (13), (20) we obtain by the triangle and Young’s inequalities

\[
\mathbb{E} \left[ \| Y^\epsilon(t_m) - Y^\tau_{\epsilon,h}(t_m) \|^2 \right] \geq \frac{1}{2} \mathbb{E} \left[ \| X^\epsilon(t_m) - X^\tau_{\epsilon,h}(t_m) \|^2 \right] - \mathbb{E} \left[ \| \Sigma(t_m) - \Sigma^\tau(t_m) \|^2 \right].
\]

Furthermore, on recalling (21), (25) we get using the Cauchy-Schwarz inequality, Young inequality and Itô’s isometry that

\[
\mathbb{E} \left[ \| \Sigma(t_m) - \Sigma^\tau(t_m) \|^2 \right] \leq C \mathbb{E} \left[ \sum_{n=1}^{m} \int_{t_{n-1}}^{t_n} \| \sigma(s, x) - \sigma(t_{n-1}, x) \|^2 \, ds \right] \\
+ C \sum_{n=1}^{m} \tau_n \mathbb{E} \left[ \| \sigma(t_{n-1}, x) - \sigma_m^{-1} \|^2 \right] + C \sum_{n=1}^{m} \tau_n \mathbb{E} \left[ \| \sigma_m^{-1} - \sigma_m^{-1} \|^2 \right] = C \mathbb{E} \left[ \eta_{\text{noise},3}^n \right].
\]

Hence, the assertion of the theorem follows after substituting the above two inequalities into (45). \(\square\)
The next lemma controls the (global) difference between the SVI solution $X$ of (1) and the regularized solution $X^\varepsilon$ of (3) in terms of the regularization parameter $\varepsilon$.

**Lemma 4.1.** Let $0 < T < \infty$ and $x_0 \in L^2(\Omega, \mathcal{F}_0; \mathbb{H}^1_0)$, $g \in \mathbb{H}^1_0$ be fixed. Let $X^\varepsilon$ be the variational solutions of (3) for $\varepsilon \in (0, 1]$. and $X$ be the SVI solution of (1). Then the following estimate holds true

$$
\sup_{t \in [0, T]} \mathbb{E} \left[ \|X(t) - X^\varepsilon(t)\|^2 \right] \leq \varepsilon |\mathcal{O}| T.
$$

**Proof.**

We set $Z = X^\varepsilon, G = A^\varepsilon(X^\varepsilon)$ in (9) and observe that

$$
\frac{1}{2} \mathbb{E} \left[ \|X(t) - X^\varepsilon(t)\|^2 \right] + \mathbb{E} \left[ \int_0^t \tilde{J}_\lambda(X(s)) \, ds \right] \leq \mathbb{E} \left[ \int_0^t \tilde{J}_\lambda(X^\varepsilon(s)) \, ds \right] - \mathbb{E} \left[ \int_0^t \left( X(s) - X^\varepsilon(s), \text{div} \frac{\nabla X^\varepsilon(s)}{\nabla X^\varepsilon(s)} \right) \, ds \right] + \mathbb{E} \left[ \int_0^t \lambda(X(s) - X^\varepsilon(s), X^\varepsilon(s) - g) \, ds \right]
$$

$$
(47) \quad := I + II + III.
$$

We show the following estimate

$$
II = \left( X - X^\varepsilon, -\text{div} \frac{\nabla X^\varepsilon}{|\nabla X^\varepsilon|} \right) \leq \tilde{J}_{\varepsilon, 0}(X) - \tilde{J}_{\varepsilon, 0}(X^\varepsilon).
$$

We consider an approximating sequence $x_k \in C^\infty(\mathcal{O}) \cap BV(\mathcal{O})$, s.t., $x_k \to X$ strongly in $\mathbb{L}^1$ and $\tilde{J}_{\varepsilon, 0}(x_k) \to \tilde{J}_{\varepsilon, 0}(X)$ for $k \to \infty$, cf. [1, Theorem 13.4.1]. Integration by parts then yields

$$
(49) \quad \left( x_k - X^\varepsilon, -\text{div} \frac{\nabla X^\varepsilon}{|\nabla X^\varepsilon|} \right) = \left( \nabla (x_k - X^\varepsilon), \frac{\nabla X^\varepsilon}{|\nabla X^\varepsilon|} \right) + \int_{\partial \mathcal{O}} \gamma_0(x_k) \frac{\nabla X^\varepsilon}{|\nabla X^\varepsilon|} \nu \, d\mathcal{H}^{n-1} - \int_{\partial \mathcal{O}} \gamma_0(X^\varepsilon) \frac{\nabla X^\varepsilon}{|\nabla X^\varepsilon|} \nu \, d\mathcal{H}^{n-1},
$$

where $\nu$ is the unit outer normal vector to the boundary $\partial \mathcal{O}$ and $\gamma_0$ is the trace operator. Since $X^\varepsilon(\omega, t) \in \mathbb{H}^1_0$ for almost all $(\omega, t) \in \Omega \times (0, T]$, the second boundary integral vanishes. We estimate the first boundary integral as

$$
\int_{\partial \mathcal{O}} \gamma_0(x_k) \frac{\nabla X^\varepsilon}{|\nabla X^\varepsilon|} \nu \, d\mathcal{H}^{n-1} \leq \int_{\partial \mathcal{O}} |\gamma_0(x_k)| \, d\mathcal{H}^{n-1}.
$$

On noting that $\|X(\omega, t)\| \leq C$ for a.a. $(\omega, t) \in \Omega \times (0, T)$, the convergence $x_k \to X$ for $k \to \infty$ in $\mathbb{L}^1$ implies $x_k \to X$ in $\mathbb{L}^2$ a.e. in $\Omega \times (0, T)$. We further assert that the trace of each approximating function $x_k \in C^\infty(\mathcal{O}) \cap BV(\mathcal{O})$, coincides with the trace of $X$ on
the boundary of $\mathcal{O}$, see [1, Remark 10.2.1]. Furthermore, we note that by the convexity of $J_{\varepsilon, 0}$ it follows that

$$\left( \nabla (x_k - X^\varepsilon), \frac{\nabla X^\varepsilon}{|\nabla X^\varepsilon|} \right) \leq J_{\varepsilon, 0}(x_k) - J_{\varepsilon, 0}(X^\varepsilon).$$

Hence, on noting the definition (6) we observe that (48) follows from (49) after taking the limit for $k \to \infty$.

Next, we obtain

$$III \leq \frac{\lambda}{2} \mathbb{E} \left[ \int_0^t \|X(s) - g\|^2 - \|X^\varepsilon(s) - g\|^2 \, ds \right].$$

After substituting $II-III$ into (47) we arrive at

$$\frac{1}{2} \mathbb{E} \left[ \|X(t) - X^\varepsilon(t)\|^2 \right] + \mathbb{E} \left[ \int_0^t \tilde{J}_0(X) \, ds \right] - \mathbb{E} \left[ \int_0^t \tilde{J}_{\varepsilon, 0}(X) \, ds \right] \leq \mathbb{E} \left[ \int_0^t \tilde{J}_0(X^\varepsilon) \, ds \right] - \mathbb{E} \left[ \int_0^t \tilde{J}_{\varepsilon, 0}(X^\varepsilon) \, ds \right].$$

We consider an approximating sequence $x_k \in C^\infty(\mathcal{O}) \cap BV(\mathcal{O})$, s.t., $x_k \to X$ strongly in $L^1$ and $\tilde{J}_0(x_k) \to \tilde{J}_0(X)$ for $k \to \infty$, cf. [1, Theorem 13.4.1]. Hence, using the lower semi-continuity of $\tilde{J}_{\varepsilon, 0}$ with respect to $L^1$ convergence, we arrive at

$$\tilde{J}_{\varepsilon, 0}(X) - \tilde{J}_0(X) \leq \liminf_{k \to \infty} \tilde{J}_{\varepsilon, 0}(x_k) - \lim_{k \to \infty} \tilde{J}_0(x_k) \leq \limsup_{k \to \infty} \tilde{J}_{\varepsilon, 0}(x_k) - \limsup_{k \to \infty} \tilde{J}_0(x_k) \leq \limsup_{k \to \infty} (\tilde{J}_{\varepsilon, 0}(x_k) - \tilde{J}_0(x_k)) \leq \|\mathcal{O}\| \varepsilon$$

where we used $|\cdot|_{\varepsilon} - |\cdot| \leq \varepsilon$. Following a similar argument for the difference of $\tilde{J}_0(X^\varepsilon) - \tilde{J}_{\varepsilon, 0}(X^\varepsilon)$ we conclude the proof. \hfill $\Box$

The following a posteriori error estimate for the numerical approximation of the SVI solution of (1) is the direct consequence of Lemma 4.1 and Theorem 4.1.

**Corollary 4.1.** Let $X$ be the SVI solution of (1), let $\{X_{\varepsilon,h}^n\}_{n=0}^N$ be the solution of (18), $X_{\varepsilon,h}^\tau$ the corresponding interpolation (24) and let the conditions of Theorem 4.1 hold. Then
the following estimates holds for $1 \leq m \leq N$:

\[
\frac{1}{2} \mathbb{E} \left[ \| X(t_m) - X_{\epsilon,h}^\tau(t_m) \|^2 \right]
\leq C \left( \varepsilon + \mathbb{E} \left[ \| Y(0) - Y_{\epsilon,h}^0 \|^2 \right] + \lambda \| g - g_h \|^2 \right) + \hat{C} \left\{ \tau_n \sum_{n=1}^{m} \mathbb{E} \left[ (n_{\text{time},1}^n)^2 + n_{\text{time},2}^n \right] 
+ \sum_{n=1}^{m} \tau_n \left( \mathbb{E} \left[ \left( \sum_{T \in T_h^n} n_{\text{space},1}^n \right)^2 \right] \right)^{\frac{1}{2}} 
+ \sum_{n=1}^{m} \tau_n \mathbb{E} \left[ \left( \sum_{T \in T_h^n} (\eta_{\text{space},1}^n)^2 + \sum_{E \in \mathcal{E}_{h,0}} (\eta_{\text{space},2}^n)^2 \right) \right] \right\}.
\]

5. Numerical experiments

We perform the numerical experiments using the fully discrete finite element scheme (18) on the unit square $\mathcal{O} = (0,1)^2$ with a slightly more general noise term. The scheme for $n = 1, \ldots, N$ then reads as

\[
(X_{\epsilon,h}^n, \varphi_h) = (X_{\epsilon,h}^{n-1}, \varphi_h) - \tau_n \left( \frac{\nabla X_{\epsilon,h}^n}{|\nabla X_{\epsilon,h}^n|}, \nabla \varphi_h \right) 
- \tau_n \lambda \left( X_{\epsilon,h}^n - g_h^n, \varphi_h \right) + \sigma \left( \Delta_h W, \varphi_h \right) 
\quad \forall \varphi_h \in V_h^n,
\]

where $g_h^n \in V_h^n$ is a suitable approximations of the function $g$ and $\sigma$ is a constant. Furthermore, we set $X_{\epsilon,h}^0 = 0$ (i.e., we use a homogeneous initial condition $x_0 \equiv 0$) and $T = 0.05$, $\lambda = 200$ in all experiments.

The space-time noise in (50) has the form

\[
\Delta_h W(x,y) = \sin(4\pi x) \sin(4\pi y) \Delta_h \beta_1 + \sin(5\pi x) \sin(5\pi y) \Delta_h \beta_2,
\]

where $\beta_1$, $\beta_2$ are independent scalar-valued Wiener processes.

The initial time step is chosen as $\tau_0 = 10^{-5}$, and $\varepsilon = 2^{-5}$, $\sigma = 0.25$, if not mentioned otherwise. The initial triangulation $T_h^0$ is constructed by dividing the domain $\mathcal{O}$ into squares with side $h = 2^{-5}$, each square is sub-divided into four triangles with equal size.

We take the function $g$ to be the characteristic function of a circle with radius 0.25, and set $g_h^n = T_h^n g + T_h^n \xi_h^{n-1} \in V_h^n$ with $\xi_h^n(x) = 0.1 \sum_{\ell=1}^{L} \varphi_\ell(x) \xi_\ell, x \in \mathcal{O}$ where $\varphi_\ell, \ell = 1, \ldots, L$ are realizations of independent $\mathcal{U}(-1,1)$-distributed random variables and $\{\varphi_\ell\}_{\ell=1}^{L}$ is the standard Lagrange basis is the space $V_h^n$. The corresponding realization of $\xi_h^n$ is displayed in Figure 1 (right).
For simplicity we employ a pathwise refinement algorithm which is explicit in time. That is, we use a Monte-Carlo approach and for each $\omega \in \Omega$ we determine $\tau_n = \tau_n(\omega)$ and $T_h^n \equiv T_h^n(\omega)$ using the solution from the previous time step as follows. Given a realization $X_{\varepsilon,h}^{n-1}(\omega_k)$ of the random variable $X_{\varepsilon,h}^{n-1}$ and the tolerances $TOL_\tau$, $TOL_h$ we compute the realization $X_{\varepsilon,h}^n(\omega_k)$ using the scheme (50) with $\tau_n = \tau_n(\omega_k)$, $V_h^n \equiv V_h^n(T_h^n(\omega_k))$. The triangulation $T_h^n(\omega_k)$ is obtained from $T_h^{n-1}(\omega_k)$ by local refinement and coarsening using an error equidistribution strategy, cf., [9]:

(i) For all $T \in T_h^{n-1}(\omega_k)$ compute the local error estimate

$$\eta^n_T = \eta_{\text{space},1}^{n-1}(T) + \sum_{E \in \partial T} \eta_{\text{space},1}^{n-1}(E),$$

using the values of $X_{\varepsilon,h}^{n-1}(\omega_k)$,

(a) if $\eta^n_T > 0.9 TOL_h/(\# T_h^{n-1})^{1/2}$ mark $T$ for refinement,

(b) if $\eta^n_T < 0.1 TOL_h/(\# T_h^{n-1})^{1/2}$ mark $T$ for coarsening;

(ii) construct $T_h^n(\omega_k)$ by local refinement/coarsening of elements in $T_h^{n-1}(\omega_k)$.

Once the triangulation $T_h^n(\omega_k)$ has been constructed we compute $X_{\varepsilon,h}^n(\omega_k) \in V_h^n(T_h^n(\omega_k))$ using a simple fixed-point nonlinear solver and determine the new time step $\tau_{n+1}(\omega_k)$ as follows

(i) Compute the time-error indicator $\eta_{\text{time},2}^n(\omega_k)$;

(a) if $\eta_{\text{time},2}^n(\omega_k) > TOL_\tau$ (or the fixed point iterations require $> 30$ iterations) set $\tau^{n+1}(\omega_k) = 0.5 \tau^n(\omega_k)$;

(b) if $\eta_{\text{time},2}^n(\omega_k) < 0.3 TOL_\tau$ (and the fixed point iterations require $< 15$ iterations) set $\tau^{n+1}(\omega_k) = 1.5 \tau^n(\omega_k)$;

(ii) set $n \leftarrow n + 1$ and proceed to next time level.

We note that the fixed-point iteration is terminated once the difference of two subsequent iterates in the $L^\infty$-norm drop below the tolerance $10^{-4}$.

In Figure 2 we display solution for $TOL_h = 0.25$, $TOL_\tau = 0.02$, the corresponding meshes are displayed in Figure 3. We observe that initial the mesh is refined within the circle due to the effect of the penalization term $\lambda$ which dominates the error estimate in the
early stages of the computation, since we start with an initial condition that is far from the function $g$. In the later stages the mesh refinement concentrates along the edge of the circle with radius 0.25, i.e., along the discontinuity of the function $g$, cf., [3]. The evolution of the time step for $TOL_\tau = 0.02, 0.01$ (and $TOL_h = 0.25$) is displayed in Figure 4; generally the time steps increases towards the end of the computation as the solution approaches the steady state. In the stochastic case the time step size is limited by the effects of the noise for smaller tolerances, whereas in the deterministic case the time step grows linearly once the steady state is reached, see Figure 4 (right).

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{solution_t0.png}
\caption{Solution at time $t = 0, 0.001, 0.004, 0.05$.}
\end{figure}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{mesh_t0.png}
\caption{Finite element mesh at time $t = 0, 0.001, 0.004, 0.05$.}
\end{figure}

One realization of the solution produced by the adaptive finite element algorithm with $2 \times$ finer tolerance $TOL_h = 0.125, TOL_\tau = 0.01$ is displayed in Figure 5, the corresponding meshes are displayed in Figure 6. The adaptive algorithm behaves analogically as in the previous experiment. Eventually, the meshes are refined along the discontinuity of the solution, with a mesh size that is $4 \times$ smaller than in the previous experiment. In Figure 6 we also observe periodic patterns away from the discontinuity which reflect the effect of the noise on the numerical solution. In contrast to the previous computation, the evolution of the time step (not sidplayed) is dictated predominantly by the convergence of the fixed-point algorithm, this is due to the fact that the fixed-point algorithm is typically expected to converge under the time step restriction $\tau \approx h^2$.

For comparison we display in Figure 7 the solution and the corresponding adaptive meshes for the deterministic problem (i.e., for $\sigma = 0$) at the final time level $t = T$ computed
with tolerance $TOL_h = 0.125$, $TOL_{\tau} = 0.01$. We observe that, along the discontinuity, the mesh is roughly $2 \times$ coarser as in the stochastic case, furthermore (before the steady state is reached, cf. Figure 4) the required time steps are roughly $2 \times$ larger than the corresponding time steps in the stochastic setting.

References

Figure 7. Finite element solution and the corresponding mesh for the deterministic problem at time $t = 0.05$.


**DEPARTMENT OF MATHEMATICS, BIELEFELD UNIVERSITY, 33501 BIELEFELD, GERMANY**

*Email address*: bananas@math.uni-bielefeld.de

**DEPARTMENT OF MATHEMATICS, BIELEFELD UNIVERSITY, 33501 BIELEFELD, GERMANY**

*Email address*: awilke@math.uni-bielefeld.de