Global Sensitivity Analysis for Models Described by Stochastic Differential Equations

Pierre Étoré[1] · Clémentine Prieur[1] · Dang Khoi Pham[1] · Long Li[1]

Abstract

Many mathematical models involve input parameters, which are not precisely known. Global sensitivity analysis aims to identify the parameters whose uncertainty has the largest impact on the variability of a quantity of interest. One of the statistical tools used to quantify the influence of each input variable on the quantity of interest are the Sobol' sensitivity indices. In this paper, we consider stochastic models described by stochastic differential equations (SDE). We focus the study on mean quantities, defined as the expectation with respect to the Wiener measure of a quantity of interest related to the solution of the SDE itself. Our approach is based on a Feynman-Kac representation of the quantity of interest, from which we get a parametrized partial differential equation (PDE) representation of our initial problem. We then handle the uncertainty on the parametrized PDE using polynomial chaos expansion and a stochastic Galerkin projection.

Keywords Stochastic differential equations · Sobol' indices · Feynman-Kac representation · Polynomial chaos expansion · Stochastic Galerkin projection

Mathematics Subject Classification (2010) 65C30 · 49Q12 · 65L60

1 Introduction

Many mathematical models (or numerical simulators) encountered in applied sciences involve numerous poorly-known input parameters. Moreover, stochastic models are often necessary for a realistic description of the physical phenomena. In deterministic models, the model response (output) is fully determined by the values of the input parameters. A specificity of nondeterministic simulators is that the same set of parameter values will lead to an

✉ Pierre Étoré
pierre.etore@univ-grenoble-alpes.fr

Clémentine Prieur
clementine.prieur@univ-grenoble-alpes.fr

[1] Université Grenoble Alpes, CNRS, Inria, Grenoble INP (Institute of Engineering, Université Grenoble Alpes), LJK, 38000 Grenoble, France

ensemble of different model responses. It is critical for practitioners to evaluate the impact of the parameter lack of knowledge onto the model response. This is the aim of sensitivity analysis. We consider in the present work the framework of global sensitivity analysis (GSA).

In the framework of GSA for deterministic models (see, e.g., Saltelli et al. 2000, and references therein), the uncertain vector of input parameters is modeled by some random vector, whose probability distribution accounts for the practitioner's belief about the input parameters uncertainty. This turns the model response into a random response, whose total variance can be split down into different parts of variance under the assumption that the input parameters are independent (this is the so-called Hoeffding decomposition, see van der Vaart 1998). Each part of variance corresponds to the contribution of each set of input parameters on the variance of the model response. By considering the ratio of each part of variance to the variance of the model response, we obtain a measure of importance for each set of input parameters that is called the Sobol' index or sensitivity index Sobol (1993); the most influential set of input parameters can then be identified and ranked as the set of input parameters with the largest Sobol' indices.

Considering a similar stochastic formalism for modeling the uncertainty on the parameters of a nondeterministic model, we obtain a random response whose randomness has two sources: the randomness from the parameters, and the one due to the stochasticity inherent to the model itself. Then there are two different settings: either one is interested in the full probability distribution of the output, or we are only concerned with the output averaged over the inherent randomness of the physical system. Note that in the following, we consider that the stochasticity of the simulator can be modeled by an additional parameter, independent of the initial uncertain parameters, also considered as a random seed variable, which can be settled by the user. In other more complex frameworks, the randomness of the process is uncontrollable in the sense that it is managed by the simulator itself, and classical sensitivity analysis tools, like Monte Carlo algorithms or metamodeling, cannot be used.

In case we are only interested in the mean value relative to the intrinsic randomness of the model, a quite common procedure consists in replacing the mean quantity by the empirical mean, and then in performing usual GSA.

If one is interested in the full probability distribution of the output, both papers Iooss and Ribatet (2009) and Marrel et al. (2012) propose a strategy based on a joint modeling of the mean and dispersion of stochastic model outputs. From these joint models, it is possible to compute, for each uncertain parameter, the first-order Sobol' index and any Sobol' index associated to an interaction with a subset of uncertain parameters. It is also possible to compute a total Sobol' index which contains the part of the output variance explained by the intrinsic noise of the model, by itself or in interaction with the uncertain parameters. However, such an approach does not allow to separate the effects enclosed in this total index. In Hart et al. (2017), the authors propose another point of view. They define Sobol' indices as random variables and study their statistical properties.

In the present paper, we focus our analysis on stochastic models described by stochastic differential equations (SDE). Such equations are frequently used for the simulation of complex systems, such as chemical kinetics (see, e.g., Gillespie and Petzold 2003, and related works) or ocean models relying on stochastic parametrizations of unresolved scales (see, e.g., Cooper 2017, and references therein), to cite only a few. Global sensitivity analysis for parametrized SDEs has been considered in Le Maître and Knio (2015) and Jimenez et al. (2017). In these two papers, the authors assume that the Wiener noise and the uncertain parameters in the parametrized SDE are independent. We will also consider this assumption. They are then interested in the full probability distribution of the solution of the

parametrized SDE. Their study relies on polynomial chaos (PC) expansion Wiener (1938), Cameron and Martin (1947), Ghanem and Spanos (1991), and Le Maître and Knio (2010) to represent the uncertain parameters, leading through a Galerkin formalism to a hierarchy of stochastic differential equations governing the evolution of the modes in the PC expansion of the overall solution. The main advantage of the PC expansion of their uncertain stochastic process is that it allows a readable expression of the conditional expectations and variances.

Contrarily to the setting of both aforementioned papers, we are not considering the full probability distribution of the solution. We are rather considering mean quantities, such as the expectation with respect to the Wiener measure of a quantity of interest related to the solution itself. We focus the study on two quantities: the first one is the expectation of the exit time τ_D from a regular bounded domain $D \subset \mathbb{R}^d$, the second one is the expectation of a functional of the solution at a time t on the event $t \leq \tau_D$. In this framework, we introduce a new methodology. We first use a Feynman-Kac representation of the quantity of interest (see, e.g., Karatzas and Shreve 1991; Gobet 2016). This leads to a parametrized partial differential equation (PDE) representation of our problem. We then handle the uncertainty on the parametrized PDE using polynomial chaos expansion and a stochastic Galerkin projection. The use of PC expansion for sensitivity analysis, or more generally uncertainty quantification, of parametrized PDE is common as it leads to analytical representations of the conditional expectations and variances, as already mentioned.

Our paper not only proposes a practical implementation of our new methodology. It also studies its theoretical justification. It is organised as follows: in Section 2 we fix some notation, state our problem and give some first results on the integrability of our quantities of interest. We also give a brief overview of existing methods for addressing our problem (Section 2.3). In Section 3 we present our methodology based on Feynman-Kac formulae. Both theoretical and numerical issues are discussed. Section 4 presents some numerical experiments on a toy model. An Appendix gathers the proofs of several technical results, mostly related to stochastic partial differential equations.

2 Notation, Problem Statement and First Results

2.1 First Notation and Assumptions

Let $\Xi \subset \mathbb{R}^m$, equipped with the Borel σ-field $\mathcal{B}_\Xi$. Let $\mathbb{P}_\xi$ be a probability measure on $(\Xi, \mathcal{B}_\Xi)$. Note that by construction the random variable $\xi(z) = z$, $z \in \Xi$, defined on $(\Xi, \mathcal{B}_\Xi, \mathbb{P}_\xi)$ has law $\mathbb{P}_\xi$. In the sequel we will denote in short $L^2(\Xi, \mathbb{P}_\xi) = L^2(\Xi, \mathcal{B}_\Xi, \mathbb{P}_\xi)$. We introduce the following assumption.

Assumption 1 (A1) Assume that the components of $\xi = (\xi_1, \ldots, \xi_m)^T$ are independent so that Ξ and $\mathbb{P}_\xi$ have product structures

$$\Xi = \prod_{j=1}^m \Xi_j \,, \ \mathbb{P}_\xi(dz) = \prod_{j=1}^m \mathbb{P}_{\xi_j}(dz_j) \,.$$

Assume moreover that for any $j = 1, \ldots, m$, the probability measure $\mathbb{P}_{\xi_j}$ is characterized by its moments. One sufficient condition for this last assertion is that the moment generating function of ξ_j has positive radius of convergence. Assumption (A1) ensures the existence of an orthonormal polynomial basis of $L^2(\Xi, \mathbb{P}_\xi)$.

The random variable ξ will represent the uncertain parameter in the SDEs we will consider.

Now let $(C, \mathcal{C}, \mathbb{W})$ denote the usual d-dimensional Wiener space. That is to say C is the space of continuous functions from $[0, \infty)$ to $\mathbb{R}^d$, equipped with the σ-field $\mathcal{C}$ endowed by the open sets for the metric defined in Equation (2.4.1) p60 of Karatzas and Shreve (1991). We recall that the Wiener measure $\mathbb{W}$ is such that the canonical process $W(\theta) = \theta, \theta \in C$, is a d-Brownian motion on $(C, \mathcal{C}, \mathbb{W})$.

We now set $\Omega = C \times \Xi, \mathcal{F} = \mathcal{C} \otimes \mathcal{B}_\Xi$ and $\mathbb{P} = \mathbb{W} \otimes \mathbb{P}_\xi$. A natural consequence of the product structure of the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is Assumption 2 below:

Assumption 2 (A2) On $(\Omega, \mathcal{F}, \mathbb{P})$ the random variables W and ξ are independent.

Consider now $b : \mathbb{R}^d \times \Xi \to \mathbb{R}^d$ and $\sigma : \mathbb{R}^d \times \Xi \to \mathbb{R}^{d \times d}$. Denoting $| \cdot |$ both the euclidean norm on $\mathbb{R}^d$ and on $\mathbb{R}^{d \times d}$ (i.e. $|a|^2 = \sum_{k,j=1}^d a_{kj}^2$ for any $a \in \mathbb{R}^{d \times d}$) we introduce the following assumptions.

Assumption 3 (A3) There exists a constant $0 \le M < \infty$ such that for all $x, y \in \mathbb{R}^d$ and all $z \in \Xi$,

$$|b(x, z) - b(y, z)| + |\sigma(x, z) - \sigma(y, z)| \le M|x - y|$$

and

$$|b(x, z)|^2 + |\sigma(x, z)|^2 \le M^2(1 + |x|^2). \tag{1}$$

We now consider the SDE with uncertain parameter ξ,

$$dX_t = b(X_t, \xi)dt + \sigma(X_t, \xi)dW_t, \quad t \ge 0, \quad X_0 = x. \tag{2}$$

Thanks to (A3), and using Theorem 5.2.9 in Karatzas and Shreve (1991), we can claim that for every value of ξ the SDE (2) has a unique strong solution X. That is to say there exists $X = (X_t)_{t \ge 0}$ defined on $(\Omega, \mathcal{F}, \mathbb{P})$ such that for a.e. $\omega = (\theta, z) \in \Omega = C \times \Xi$, we have: $\forall t \ge 0, \forall 1 \le i \le d$,

$$X_t^i(\omega) = x^i + \int_0^t b_i(X_s(\omega), \xi(z)) \, ds + \sum_{j=1}^d \int_0^t \sigma_{ij}(X_s(\omega), \xi(z)) \, dW_s^j(\theta).$$

In fact, under (A3), such a process has all its moments as we state in the following lemma, whose proof is postponed to the Appendix.

Lemma 1 *Assume (A2) and (A3). Let X be the solution of (2) and $q \in \mathbb{N}^*$. For any $t \ge 0$ we have $\mathbb{E}|X_t|^q < \infty$ (in other words $X_t \in L^q(\Omega, \mathbb{P})$).*

Remark 1 Assumption 3, in order to get the existence of the solution X and its square integrability, could be weakened. But it provides a reasonable setting to study the partial differential equation associated to our original problem (see Section 3).

Remark 2 Let $F : C \to \mathbb{R}$ measurable such that $F(X)$ belongs to $L^1(\Omega, \mathbb{P})$. Note that thanks to (A2),

$$\mathbb{E}[\, F(X) \,|\, \xi \,] = \int_C F(X(\theta, \xi)) \mathbb{W}(d\theta),$$

that is to say $\mathbb{E}[\, \cdot \,|\, \xi \,]$ can be seen as the average with respect to the Brownian motion W driving (2).

2.2 Problem Statement

As already mentioned in the introduction, the d-Brownian motion in our study models the physical system randomness. The quantities we are interested in are then mean values, with respect to the d-Brownian motion, of outputs related to the SDE. These quantities depend on uncertain parameters, and we aim at determining which parameters are influential on these quantities, by performing a global sensitivity analysis. This section aims at defining properly two quantities we are interested in (see Eqs. 3 and 6).

We first need to introduce a set of assumptions and notation, as far as preliminary results.

Let σ denote the diffusion term in Eq. 2, we then define $a = \sigma \sigma^T$ and introduce the following uniform ellipticity assumption:

Assumption 4 (A4). There exists $\lambda > 0$ such that

$$\forall x, y \in \mathbb{R}^d, \ \forall z \in \varXi, \quad y^T a(x, z) y \geq \lambda |y|^2.$$

Let D be an open bounded subset of $\mathbb{R}^d$. Recall that X is solution of Eq. 2. We define the first exit time of the process X from D as:

$$\tau_D = \inf\{t \geq 0 : X_t \notin D\}.$$

We have the following result (Lemma 2) whose proof is postponed to the Appendix.

Lemma 2 *Assume (A2)–(A4). For any starting point $x \in D$ of* Eq. 2 *we have $\tau_D \in L^1(\varOmega, \mathbb{P})$.*

In the sequel the starting point $x \in \mathbb{R}^d$ of X solution of Eq. 2 may vary. We classically denote by $\mathbb{E}^x(\cdot|\xi)$ the expectation with respect to W computed under the initial condition $X_0 = x$.

The first quantity we are interested in is the averaged (with respect to W) exit time of the process X from the domain D (we assume $x \in \bar{D}$, the closure of D). It is defined as:

$$\mathfrak{U}(x, \xi) = \mathbb{E}^x[\, \tau_D \,|\, \xi \,]. \tag{3}$$

To ensure that $\mathfrak{U}(x, \xi)$ is in $L^2(\varXi, \mathbb{P}_\xi)$, we add a regularity assumption on the domain D:

Assumption 5 (A5) The boundary ∂D is of class $C^{2,\alpha}$, $\alpha = 1$ (see Gilbarg and Trudinger 1983, p94 for a definition).

Then, it is possible to prove the following lemma:

Lemma 3 *Assume (A2)–(A5). For any $x \in D$ and almost every value of ξ we have $|\mathfrak{U}(x, \xi)| \leq C$ where C is a finite constant.*

Proof Note that Lemma 3 is a corollary of Theorem 1 stated in Section 3. Its proof is postponed to the Appendix. □

Remark 3 Even if the result of Theorem 1 is required for the proof of Lemma 3, we decided to postpone its statement to Section 3, devoted to the introduction of our new methodology, based on Feynman-Kac representations of the mean quantities we are interested in.

Theorem 1 provides indeed a clear interpretation of $\mathfrak{U}(x, \xi)$ as the solution of an elliptic stochastic partial differential equation.

Let now $0 \le t < \infty$. Let f be a function from $\bar{D}$ to $\mathbb{R}$. We introduce the following set of assumptions on f:

Assumption 6 (A6) The function f is of class $C^2(\bar{D})$ and satisfies the compatibility condition

$$f \equiv 0 \quad \text{and} \quad \frac{1}{2}\sum_{k,j}^{d} a_{kj}\partial^2_{x_k x_j} f + \sum_{j=1}^{d} b_j \partial_{x_j} f \equiv 0 \text{ on } \partial D. \tag{4}$$

Besides, for all $x, y \in \bar{D}$,

$$|f(x) - f(y)| + \sum_{k=1}^{d} |\partial_{x_k} f(x) - \partial_{x_k} f(y)| + \sum_{k,j=1}^{d} |\partial^2_{x_j x_k} f(x) - \partial^2_{x_j x_k} f(y)| \le M|x - y|$$

and

$$|f(x)| \le M(1 + |x|^\mu), \tag{5}$$

with constants $0 \le M, \mu < \infty$.

From Lemma 1 we immediately get Lemma 4 just below.

Lemma 4 *Assume (A2)–(A3) and Eq. 5. Then for any $q \in \mathbb{N}^*$ we have $f(X_t) \in L^q(\Omega, \mathbb{P})$, for any $t \ge 0$.*

It is then possible to define the second quantity we are interested in. The quantity $\mathfrak{V}(t, x, \xi)$ is defined as the expectation (with respect to W) of $f(X_t)$, computed on the event $\{t \le \tau_D\}$:

$$\mathfrak{V}(t, x, \xi) = \mathbb{E}^x[f(X_t)\mathbf{1}_{t \le \tau_D} \,|\, \xi]. \tag{6}$$

Remark 4 Lemma 4 ensures that, under (A2)–(A3) and (A6), the quantity $\mathfrak{V}(t, x, \xi)$ is in $L^2(\varXi, \mathbb{P}_\xi)$.

In the sequel, in order to lighten notation, we will sometimes drop any reference to x or t in the quantities of interest $\mathfrak{U}$ or $\mathfrak{V}$.

Recall that we aim at determining which parameters, or set of parameters, are influential for $\mathfrak{U}$, *resp.* $\mathfrak{V}$. We propose to compute Sobol' sensitivity indices, with respect to the components of $\xi = (\xi_1, \ldots, \xi_m)^T$ assumed to be independent. These indices are defined (see, e.g., Sobol 1993) as:

$$S_I(\mathfrak{U}) = \frac{\mathbb{V}[\mathbb{E}(\mathfrak{U}(\xi)|\xi_I)]}{\mathbb{V}(\mathfrak{U}(\xi))}, \quad \forall I \subset \{1, \ldots, m\}, \quad \xi_I = \{\xi_\ell, \ell \in I\} \tag{7}$$

and

$$S_I(\mathfrak{V}) = \frac{\mathbb{V}[\mathbb{E}(\mathfrak{V}(\xi)|\xi_I)]}{\mathbb{V}(\mathfrak{V}(\xi))}, \quad \forall I \subset \{1, \ldots, m\}, \quad \xi_I = \{\xi_\ell, \ell \in I\}. \tag{8}$$

Remark 5 In the following, we assume that $x \in D$ and $t > 0$ to ensure $\mathbb{V}(\mathfrak{U}(\xi)) \neq 0$ and $\mathbb{V}(\mathfrak{V}(\xi)) \neq 0$.

Remark 6 Assumption (A6) on f is a technical assumption which may appear restrictive. In practice, if f does not satisfy (A6), we approximate it by some f_ε satisfying (A6) and propose a bound to control the approximation error. E.g., for $f \equiv 1$ on $\bar{D}$, and for any $\varepsilon > 0$, it is possible via convolution arguments to construct a function $f_\varepsilon \in C^2(\bar{D})$ satisfying (A6), such that $0 \leq f_\varepsilon \leq 1$ and $f_\varepsilon \equiv 1$ on K_ε, with K_ε a compact subset of D with $|D \setminus K_\varepsilon| \leq \varepsilon$ (note that the existence of K_ε is ensured by Assumption (A5) on the regularity of the boundary). Then we have:

$$\left| \mathbb{E}^x[\, f(X_t)\mathbf{1}_{t \leq \tau_D} \,|\, \xi\,] - \mathbb{E}^x[\, f_\varepsilon(X_t)\mathbf{1}_{t \leq \tau_D} \,|\, \xi\,] \right| \leq \mathbb{P}^x\,(X_t \in D \setminus K_\varepsilon \,|\, \xi)\,. \tag{9}$$

In fact, as in the left hand side of Eq. 9 we are only concerned with the paths of X before it exits D, we can without loss of generality assume that σ and b are globally bounded and Lipschitz on $\mathbb{R}^d$ (with constants uniform in ξ; in particular the coefficients are bounded by some $0 < \bar{M} < \infty$). Then using Theorems 6.4.5, 6.4.7 and 6.5.4 in Friedman (1975), we get the following estimate for the transition probability density of $(X_t)t \geq 0$, under assumptions (A3), (A4):

$$p(t, x, y, \xi) \leq \frac{M_T}{t^{d/2}} \exp\left(-\frac{|y - x|^2}{M_T t} \right), \quad \forall x, y \in D \tag{10}$$

with M_T depending on $T, \bar{M}, \lambda$ (but not on ξ). We thus conclude that the right hand side in Eq. 9 behaves as $O\,(\varepsilon)$.

2.3 State of the Art

In this section, we first recall the crude Monte Carlo procedure for the estimation of $S_I(\mathfrak{U})$ and $S_I(\mathfrak{V})$. We then present the approach introduced recently in Jimenez et al. (2017) and Le Maître and Knio (2015) for the estimation of the $S_I(\mathfrak{V}_i)$, with $\mathfrak{V}_i(t, x, \xi) = \mathbb{E}^x[X_t^i|\xi]$, for $1 \leq i \leq d$.

2.3.1 Crude Monte Carlo Estimation Procedure

The crude Monte Carlo estimation procedure is probably the most intuitive. However, its main issue is its computational cost. Let us detail below what we mean by crude Monte Carlo estimation procedure.

Let $\mathcal{F}_{\mathfrak{U}} : C \times \varXi \to \mathbb{R}$ (*resp.* $\mathcal{F}_{\mathfrak{V}} : C \times \varXi \to \mathbb{R}$) be the application such that $\mathcal{F}_{\mathfrak{U}}(W, \xi) = \tau_D$ (*resp.* $\mathcal{F}_{\mathfrak{V}}(W, \xi) = f(X_t)\mathbf{1}_{t \leq \tau_D}$). Note that the deterministic maps $\mathcal{F}_{\mathfrak{U}}$ and $\mathcal{F}_{\mathfrak{V}}$ exist because for fixed ξ one passes from a path of W to a path of X in a deterministic way, thanks to the Yamada-Watanabe causality principle (see Karatzas and Shreve 1991).

In practice our Monte Carlo procedures will involve some Euler scheme to approximate the paths of X, so that we will in fact work with approximations $\widehat{\mathcal{F}}_{\mathfrak{U}}$ and $\widehat{\mathcal{F}}_{\mathfrak{V}}$ of $\mathcal{F}_{\mathfrak{U}}$ and $\mathcal{F}_{\mathfrak{V}}$ (e.g., in Section 4).

In the sequel we focus on $\mathfrak{U}(\xi)$, as $\mathfrak{V}(\xi)$ can be treated in a similar manner. The quantity $\mathfrak{U}(\xi)$ is first approximated by

$$\mathbb{E}[\,\widehat{\mathcal{F}}_{\mathfrak{U}}(W, \xi)\,|\, \xi\,].$$

Second, let $M \in \mathbb{N}^*$ (M is supposed to be large) and $W^{(k)}, k = 1, \ldots, M$, independent samplings of the Brownian motion W. We approximate $\mathbb{E}[\,\widehat{\mathcal{F}}_{\mathfrak{U}}(W, \xi)\,|\, \xi\,]$ by the Monte Carlo mean

$$\overline{E}^M\left[\widehat{\mathcal{F}}_{\mathfrak{U}}\right](\xi) := \frac{1}{M} \sum_{k=1}^M \widehat{\mathcal{F}}_{\mathfrak{U}}(W^{(k)}, \xi).$$

We then perform a sensitivity analysis of $\overline{E}^M\left[\widehat{\mathcal{F}}_{\mathfrak{U}}\right](\xi)$ with respect to the components of $\xi = (\xi_1, \ldots, \xi_m)^T$. Sobol' indices can be estimated with a classical *pick freeze* procedure (see, e.g., Sobol' 1993, 2001). Let $I \subset \{1, \ldots, m\}$. Let $N \in \mathbb{N}^*$ (N is supposed to be large). Let $\xi^{A,(l)}, \xi^{B,(l)}, l = 1, \ldots, N$ independent samplings of ξ. We then construct samples $\xi^{I,(l)}, l = 1, \ldots, N$ in the following manner:

$$\forall 1 \le l \le N, \forall 1 \le \ell \le m, \quad \xi_\ell^{I,(l)} = \begin{cases} \xi_\ell^{B,(l)} & \text{if } \ell \in I \\ \xi_\ell^{A,(l)} & \text{otherwise.} \end{cases}$$

Then the Sobol' index $S_I(\mathfrak{U})$ is approximated by

$$S_I(\overline{E}^M\left[\widehat{\mathcal{F}}_{\mathfrak{U}}\right](\xi)) = \frac{\mathbb{V}\left[\mathbb{E}\left(\overline{E}^M\left[\widehat{\mathcal{F}}_{\mathfrak{U}}\right](\xi) \mid \xi_I\right)\right]}{\mathbb{V}\left[\overline{E}^M\left[\widehat{\mathcal{F}}_{\mathfrak{U}}\right](\xi)\right]} = \frac{\mathbb{V}\left[\overline{E}^M\left[\widehat{\mathcal{F}}_{\mathfrak{U}}\right](\xi_I)\right]}{\mathbb{V}\left[\overline{E}^M\left[\widehat{\mathcal{F}}_{\mathfrak{U}}\right](\xi)\right]}$$

which itself is estimated by

$$\frac{\frac{1}{N}\sum_{l=1}^N \overline{E}^M\left[\widehat{\mathcal{F}}_{\mathfrak{U}}\right](\xi^{I,(l)}) \times \overline{E}^M\left[\widehat{\mathcal{F}}_{\mathfrak{U}}\right](\xi^{B,(l)}) - \left(\overline{C}\right)^2}{\frac{1}{N}\sum_{l=1}^N \frac{\left\{\overline{E}^M\left[\widehat{\mathcal{F}}_{\mathfrak{U}}\right](\xi^{I,(l)})\right\}^2 + \left\{\overline{E}^M\left[\widehat{\mathcal{F}}_{\mathfrak{U}}\right](\xi^{B,(l)})\right\}^2}{2} - \left(\overline{C}\right)^2} \tag{11}$$

with

$$\overline{C} = \frac{1}{N}\sum_{l=1}^N \frac{\overline{E}^M\left[\widehat{\mathcal{F}}_{\mathfrak{U}}\right](\xi^{I,(l)}) + \overline{E}^M\left[\widehat{\mathcal{F}}_{\mathfrak{U}}\right](\xi^{B,(l)})}{2}.$$

This estimator has been introduced first in Monod et al. (2006) and its asymptotic properties have been studied in Janon et al. (2014). The main advantage of *pick freeze* estimators is that it only requires the square integrability for the model. However, the number of model evaluations it requires can be huge. It is equal to $MN(m+1)$, with m the number of uncertain parameters.

2.3.2 Hybrid Galerkin-Monte Carlo Procedure

In the setting of parametrized SDE, we can exploit regularity properties of the underlying model for proposing alternatives to the crude Monte Carlo procedure. In this section, we present the procedure introduced in Le Maître and Knio (2015) and Jimenez et al. (2017). It is based on a polynomial chaos analysis with stochastic expansion coefficients. In that paper, the uncertainty due to the Wiener noise and the one due to the uncertain parameters ξ are handled at the same level for sensitivity analysis. Recall that the components of $\xi = (\xi_1, \ldots, \xi_m)^T$ are assumed to be independent so that $\varXi$ and $\mathbb{P}_\xi$ have product structures

$$\varXi = \prod_{j=1}^m \varXi_j, \quad \mathbb{P}_\xi(dz) = \prod_{j=1}^m \mathbb{P}_{\xi_j}(dz_j).$$

We endow $L^2(\varXi, \mathbb{P}_\xi)$ with the inner product and associated norm denoted $< \cdot, \cdot >$ and $\|\cdot\|_2$ respectively

$$\forall U, V \in L^2(\varXi, \mathbb{P}_\xi), \quad < U, V > = \int_\varXi U(z)V(z)\mathbb{P}_\xi(dz);$$

$$\|U\|_2 = < U, U >^{1/2} < \infty \iff U \in L^2(\varXi, \mathbb{P}_\xi).$$

Then, introducing an orthonormal basis $\{\Psi_q(\cdot)\,,\ q \in \mathbb{N}\}$ for $L^2(\Xi, \mathbb{P}_\xi)$ (e.g., any tensorized basis), any second-order random variable $U(\xi)$ can be expanded as

$$U(\xi) = \sum_{q \in \mathbb{N}} [U_q] \Psi_q(\xi)\,.$$

The authors in Le Maître and Knio (2015) and Jimenez et al. (2017) consider then the following tensor representation: for a.e. $\omega = (\theta, z) \in \Omega = C \times \Xi$,

$$X_t^i(\theta, z) = \sum_{q \in \mathbb{N}} [(X_t^i)_q](\theta) \Psi_q(z)\,,\ i = 1, \ldots, d. \tag{12}$$

Then, they consider a stochastic Galerkin projection (see, e.g., Ghanem and Spanos (1991)) to derive equations that enable the determination of the stochastic processes $[(X_t^i)_q](\theta)$, $i = 1, \ldots, d$. E.g., let $i \in \{1, \ldots, d\}$, we get

$$d[(X_t^i)_q](\theta) = \ <b \left(\sum_{r \in \mathbb{N}} [(X_t^i)_r](\theta) \Psi_r(\cdot), \cdot \right) dt, \Psi_q(\cdot) >$$

$$+\sigma \left(\sum_{r \in \mathbb{N}} [(X_t^i)_r](\theta) \Psi_r(\cdot), \cdot \right) dW_t(\theta), \Psi_q(\cdot) > .$$

From Assumption (A2), we know that ξ and W are independent thus

$$d[(X_t^i)_q](\theta) = \ <b \left(\sum_{r \in \mathbb{N}} [(X_t^i)_r](\theta) \Psi_r(\cdot), \cdot \right), \Psi_q(\cdot) > dt$$

$$+\sigma \left(\sum_{r \in \mathbb{N}} [(X_t^i)_r](\theta) \Psi_r(\cdot), \cdot \right), \Psi_q(\cdot) > dW_t(\theta).$$

From computational purposes, the authors truncate this infinite sequence of coupled problems to an order P: $0 \leq q \leq P$. They thus get a system of $P+1$ coupled stochastic differential equations, with initial conditions obtained by projecting the initial data on the stochastic basis $[(X_0^i)_q](\theta) = <X_0, \Psi_q>$. The system is then solved using standard Monte Carlo simulation, introducing a time scheme and generating trajectories for the d-Brownian motion.

Consider now the quantities of interest analyzed in Jimenez et al. (2017) and Le Maître and Knio (2015):

$$\forall 1 \leq i \leq d, \quad \mathfrak{V}_i(t, x, \xi) = \mathbb{E}^x[X_t^i \,|\, \xi],$$

denoted in short $\mathfrak{V}_i(\xi)$, $1 \leq i \leq d$, and recall (12).

Then, due to the orthonormality of the stochastic basis $\{\Psi_q\,,\ q \in \mathbb{N}\}$, we get

$$\mathbb{V}[\mathfrak{V}_i(\xi)] \approx \sum_{q=1}^{P} \{\mathbb{E}[(X_t^i)_q]\}^2 \tag{13}$$

and

$$\mathbb{V}(\mathbb{E}[\mathfrak{V}_i(\xi)|\xi_I]) \approx \sum_{q \in K_I} \mathbb{E}\{[(X_t^i)_q]^2\} \tag{14}$$

with K_I defined as

$$K_I = \{p \in \{1, \ldots, P\} : \psi_p(\xi) = \psi_p(\xi_I)\}. \tag{15}$$

The application of such a procedure requires technical assumptions on the initial parametrized SDE, which we do not develop here for sake of concision.

In the next section, we propose a new methodology for the estimation of Sobol' indices defined by Eq. 7 and Eq. 8, combining Feynman-Kac representation formulas with polynomial orthonormal decompositions on the basis $\{\Psi_q, q \in \mathbb{N}\}$.

3 A New Methodology Based on Feynman-Kac Representation Formulas

This section is devoted to the presentation of the new methodology we propose for performing sensitivity analysis of parametrized stochastic differential equations. Let us recall that we consider in our approach that our model is a stochastic model depending on unknown parameters $\xi = (\xi_1, \ldots, \xi_m)^T$, the stochasticity being induced by the d-Brownian motion. We then are only interested by the mean value of a functional of the model output, where the mean is taken with respect to the d-Brownian motion which drives the inherent randomness of the model. It is different in nature from the point of view adopted in Le Maître and Knio (2015) and Jimenez et al. (2017), in which the authors are interested in the sensitivity of the model output with respect to both the uncertain parameters and the noise inherent to the system. Our approach is based on Feynman-Kac representation formulas which establish a link between parabolic or elliptic partial differential equations and stochastic processes. More precisely, Feynman-Kac formulas allow interpreting the quantities $\mathfrak{U}$ and $\mathfrak{V}$ as the solutions of some partial differential equations. The literature proposes a broad range of methods for the sensitivity analysis of parametrized partial differential equations (see, e.g., Nouy 2017, and references therein). Hereafter, for each of both quantities of interest under study in this paper, we focus on a method based on a stochastic Galerkin polynomial chaos approximation of the solution of the parametrized PDE we obtain from the Feynman-Kac representation.

In Section 3.1 below, we state in Theorems 1 and 2 the Feynman-Kac representations of our quantities of interest $\mathfrak{U}$ and $\mathfrak{V}$. Then in Section 3.2 (resp. Section 3.3), we introduce our methodology based on stochastic Galerkin approximation for the estimation of Sobol' indices associated to $\mathfrak{U}$ (resp. $\mathfrak{V}$).

3.1 Link Between $\mathfrak{U}$ and $\mathfrak{V}$ and Some Stochastic Partial Differential Equations

Theorem 1 below provides an interpretation of $\mathfrak{U}$ as the solution of an elliptic problem.

Theorem 1 *Let us consider the following elliptic stochastic partial differential equation*

$$\begin{cases} \frac{1}{2} \sum_{k,j=1}^{d} a_{kj}(x, \xi) \partial^2_{x_k x_j} u(x, \xi) + \sum_{k=1}^{d} b_k(x, \xi) \partial_{x_k} u(x, \xi) = -1 \ \forall x \in D \\ \\ u(x, \xi) = 0 \quad \forall x \in \partial D. \end{cases} \tag{16}$$

i) *Assume (A3)–(A5). Then, for almost every value of ξ, there exists a unique solution to* Eq. 16 *in* $C^2(\overline{D})$.
ii) *If moreover (A2) is satisfied, then for a.e. ξ, the quantity $\mathfrak{U}(x, \xi)$ is solution of the partial differential equation defined by* Eq. 16.

Proof (*Proof Theorem 1*) Point i) is a direct consequence of Theorem 6.14 in Gilbarg and Trudinger (1983). Taking into account Remark 2 and Lemma 2, ii) follows simply from the elliptic Feynman-Kac formula given in Theorem 5.7.2 in Karatzas and Shreve (1991). $\square$

We now turn to the interpretation of $\mathfrak{V}$.

Theorem 2 *Let us consider the following parabolic stochastic partial differential equation*

$$
\begin{cases}
\partial_t v(t, x, \xi) = \frac{1}{2} \sum_{k,j=1}^d a_{kj}(x, \xi) \partial^2_{x_k x_j} v(t, x, \xi) + \sum_{k=1}^d b_k(x, \xi) \partial_{x_k} v(t, x, \xi), \\
\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \forall (t, x) \in (0, T] \times D \\
v(t, x, \xi) = 0, \ \forall (t, x) \in (0, T] \times \partial D \\
v(0, x, \xi) = f(x), \ \forall x \in \bar{D}.
\end{cases}
\tag{17}
$$

i)   *Assume (A3)–(A6). Then for almost every value of ξ, there exists a unique solution to Eq. 17 in $C^{1,2}([0, T] \times \bar{D})$.*

ii) *Assume in addition (A2). Then the quantity $\mathfrak{V}(t, x, \xi)$, $0 < t \leq T$, is given by the solution at time t of the parabolic stochastic partial differential Eq. 17.*

Proof (Proof of Theorem 2) As f satisfies the compatibility condition (4) one may extend it for any ξ in a function $\bar{f}$ of class $C^{1,2}([0, T] \times \bar{D})$, with uniformly Lipschitz derivatives (to order 1 in time and up to order 2 in space) and satisfying $\bar{f} \equiv 0$ on $(0, T] \times \partial D$, and $\bar{f} \equiv f$ on $\{0\} \times \bar{D}$ and $\frac{1}{2} \sum_{k,j}^d a_{kj} \partial^2_{x_k x_j} \bar{f} + \sum_{j=1}^d b_j \partial_{x_j} \bar{f} \equiv 0$ on $\{0\} \times \partial D$ (e.g., let us consider $\bar{f}(t, x) = f(x) \ \forall (t, x) \in [0, T] \times \bar{D}$). Point i) then follows from Theorem 5.14 in Lieberman (1996). Taking into account Remark 2, the result ii) follows simply from the parabolic Feynman-Kac formula given in Theorem 4.4.5 in Gobet (2016). Note however that this Feynman-Kac formula is given for parabolic problems in backward form and with terminal condition instead of initial one, as the authors deal with the general case of time-inhomogeneous coefficients. But in the case of time-homogeneous coefficients, one may use a reverting time argument, and the time-homogeneous Markov property of X, in order to get the Feynman-Kac formula for the parabolic problem in the forward form and with initial condition. □

At this stage of the paper, we are concerned with *classical solutions* of stochastic PDEs, in the sense of Definition 1 below. In Sections 3.2 and 3.3, we will introduce the notion of *weak solutions at the stochastic level*.

Definition 1 A function $u : \bar{D} \times \varXi \to \mathbb{R}$ (resp. $v : [0, T] \times \mathbb{R}^d \times \varXi \to \mathbb{R}$) is a classical solution of Eq. 16 (resp. (17)) if for a.e. value of ξ the function $x \mapsto u(x, \xi)$ (resp. $(t, x) \mapsto v(t, x, \xi)$) is a solution to Eq. 16 (resp. (17)) in $C^2(\bar{D})$ (resp. $C^{1,2}([0, T] \times \bar{D})$).

Corollary 1 *Assume (A3)–(A5) (resp. (A3)–(A6)). There exists a unique classical solution of the stochastic PDE (16) (resp. (17)). If in addition (A2) is satisfied, this classical solution is provided by $\mathfrak{U}(x, \xi)$ (resp. $\mathfrak{V}(t, x, \xi)$).*

From the Feynman-Kac representations of our quantities of interest stated in Corollary 1, we propose in Sections 3.2 and 3.3 to apply a methodology based on stochastic Galerkin polynomial chaos approximations to approximate Sobol' indices. For the clarity of exposure we have found convenient to present first the elliptic case (Section 3.2), and then to turn to the parabolic one which is technically more cumbersome (Section 3.3).

3.2 Approximation of $S_I(\mathfrak{U})$ Using the Elliptic Stochastic PDE

We want now to construct a discretized approximation of $\mathfrak{U}(\xi)$. So far (Corollary 1) the function $\mathfrak{U}$ is seen as a classical solution of Eq. 16 but it can be shown to be a *weak solution at the stochastic level* of some equivalent PDE in divergence form, living in the space $H_0^1(D) \otimes L^2(\varXi, \mathbb{P}_\xi)$ (this will be addressed in Theorem 3 below; the definition of $H_0^1(D)$ will be recalled in Section 3.2.1). Therefore it is possible to perform a Galerkin projection on some finite dimensional subspace $V^N \otimes \mathcal{S}^P$ where $\mathcal{S}^P = \text{Span}(\{\Psi_q\}_{q=0}^P)$ with $\{\Psi_q\}_{q=0}^P$ introduced in Section 2.3.2 and $V^N = \text{Span}(\{\phi_i^N\}_{i=1}^{N-1})$ with $\{\phi_i^N\}_{i=1}^{N-1}$ some finite element basis of functions in $H_0^1(D)$ (see Section 3.2.1 for details). That is to say $\mathfrak{U}(x, \xi)$ will be discretized both with respect to the space variable x and the uncertain parameter ξ. Concerning weak solutions and Galerkin projection we are inspired by the framework in Nouy (2009, 2017). In this article we provide a rigorous proof of the convergence of our Galerkin approximation toward the weak solution when the discretization parameters in x and ξ simultaneously converge (see in particular the proof of Lemma 5, postponed to the Appendix). Once we have got an approximation (also called reduced model or metamodel) $\mathfrak{U}^{N,P}$ of the weak solution, therefore of $\mathfrak{U}(\xi)$, we deduce an analytical formula for approximating Sobol' indices $S_I(\mathfrak{U})$, $I \subset \{1, \ldots, m\}$. The arguments to derive such an analytical formula are similar to the ones used to derive Eqs. 13 and 14 in Section 2.3.2, and mainly rely on the orthonormality of the basis $\{\Psi_q\}_{q=0}^P$.

The subsection is organised ad follows: in Section 3.2.1 we introduce some notation on the functional spaces we will use both in this subsection and in Section 3.3 about the parabolic case. In Section 3.2.2 we reformulate the elliptic stochastic PDE (16) in divergence form. This is needed in order to introduce in Section 3.2.3 the definition of weak solutions at the stochastic level of elliptic stochastic PDEs, and of their Galerkin approximation. The aforementioned Theorem 3 is stated in Section 3.2.4 and establishes the convergence of the Galerkin approximation $\mathfrak{U}^{N,P}$ toward the quantity of interest $\mathfrak{U}$. In Section 3.2.5 it is explained how to use $\mathfrak{U}^{N,P}$ in order to compute approximated Sobol' indices of $\mathfrak{U}$.

3.2.1 Functional Spaces Notation

We denote by $H_0^1(D)$ the usual Sobolev space of functions in $H^1(D)$ vanishing on ∂D (note that in the sequel the derivatives may be understood in the weak sense). The space $H_0^1(D)$ is equipped with the norm $u \mapsto ||u||_{H^1(D)} = \int_D u^2 + \int_D |\nabla u|^2$. We denote $H^{-1}(D)$ the topological dual space of $H_0^1(D)$. Recall that we have the Gelfand triple $H_0^1(D) \hookrightarrow L^2(D) \hookrightarrow H^{-1}(D)$ (embeddings are continuous and dense).

We then introduce the tensor Hilbert spaces $\mathcal{H} = L^2(D) \otimes L^2(\varXi, \mathbb{P}_\xi)$ and $\mathcal{V} = H_0^1(D) \otimes L^2(\varXi, \mathbb{P}_\xi)$. The space $\mathcal{H}$ is isomorphic to $L^2(\varXi, L^2(D), \mathbb{P}_\xi)$ and $L^2(D, L^2(\varXi, \mathbb{P}_\xi))$. The space $\mathcal{V}$ is isomorphic to $L^2(\varXi, H_0^1(D), \mathbb{P}_\xi)$ and $H_0^1(D, L^2(\varXi, \mathbb{P}_\xi))$. The space $\mathcal{H}$ is equipped with the scalar product $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ defined by

$$\forall v, w \in \mathcal{H}, \quad \langle v, w \rangle_{\mathcal{H}} = \mathbb{E}_\xi \left(\int_D vw \right),$$

and with the norm $u \mapsto ||u||_{\mathcal{H}} := \sqrt{\langle u, u \rangle_{\mathcal{H}}}$. The space $\mathcal{V}$ is equipped with the norm $|| \cdot ||_{\mathcal{V}}$ defined as:

$$u \mapsto \mathbb{E}_\xi ||u(\cdot, \xi)||_{H^1(D)}^2 =: ||u||_{\mathcal{V}}^2.$$

Denoting $\mathcal{V}'$ the topological dual of $\mathcal{V}$ we have the new Gelfand triple $\mathcal{V} \hookrightarrow \mathcal{H} \hookrightarrow \mathcal{V}'$. Note that $\mathcal{V}' = H^{-1}(D) \otimes L^2(\mathbb{P}_\xi)$. Note that $\langle \cdot, \cdot \rangle_\mathcal{H}$ will also denote the extension by continuity of the previous scalar product to the dual pairing $\mathcal{V}' \times \mathcal{V}$.

We introduce the finite dimensional approximation space $\mathcal{V}^{N,P} \subset \mathcal{V}$ defined as the tensor space

$$\mathcal{V}^{N,P} = V^N \otimes \mathcal{S}^P. \tag{18}$$

Recall that the space $\mathcal{S}^P$ in Eq. 18 is defined as $\mathrm{Span}(\{\Psi_q\}_{q=0}^P)$ with $\{\Psi_q\}_{q=0}^P$ introduced in Section 2.3.2. Recall also that the space V^N is defined as $\mathrm{Span}(\{\phi_i^N\}_{i=1}^{N-1})$ with $\{\phi_i^N\}_{i=1}^{N-1}$ some finite element basis of functions in $H_0^1(D)$. More precisely, we choose ϕ_i^N, $N \geq 2$, $1 \leq N - 1$ such that $V_2 \subset \ldots \subset V_{N-1} \subset V_N$. This can be done for example by using linear B-spline basis functions. Using interpolation properties, we get that $\cup_{N \geq 2} V_N$ is dense in $H_0^1(D)$.

We then have $\mathcal{V}^{N,P} \subset \mathcal{V}^{N+1,P+1}$ for any N, P, and $\bigcup_{N,P} \mathcal{V}^{N,P}$ is dense in $\mathcal{V}$.

3.2.2 Elliptic PDE in Divergence Form

For any vector field $U \in C^1(D; \mathbb{R}^d)$ we denote $\nabla \cdot U = \sum_{k=1}^d \partial_{x_k} U^k$ the divergence of U and, for any $U \in C^1(D; \mathbb{R})$ we denote $\nabla U = (\partial_{x_1} U, \ldots, \partial_{x_d} U)^T$ the gradient of U.

Assume that for a.e. value of ξ the function $a(\cdot, \xi)$ is of class $C^1(D; \mathbb{R}^{d \times d})$ and define the coefficients $\tilde{a}$, $\tilde{b}$, the vector field $\partial \tilde{a}$, and the function $\tilde{f}$ on $D \times \Xi$ by

$$\tilde{a} = a/2, \qquad (\partial \tilde{a})_j = \nabla \cdot (\tilde{a}_{1j}, \ldots, \tilde{a}_{dj})^T, \quad \forall 1 \leq j \leq d$$
$$\tilde{b} = \partial \tilde{a} - b \quad \text{and} \quad \tilde{f} = 1 \tag{19}$$

It is then clear that Eq. 16 is equivalent to

$$\begin{cases} -\nabla \cdot (\tilde{a}(x, \xi) \nabla u(x, \xi)) + (\nabla u(x, \xi))^T \tilde{b}(x, \xi) = \tilde{f}(x, \xi) \ \forall x \in D \\ \\ u(x, \xi) = 0 \qquad \forall x \in \partial D. \end{cases} \tag{20}$$

3.2.3 Weak Solutions at the Stochastic Level to Eq. 20 and Their Galerkin Projection

We now introduce the bilinear form

$$\forall u, v \in \mathcal{V}, \ \mathcal{A}(u, v) = \mathbb{E}_\xi \left(\int_D (\nabla u(\cdot, \xi))^T \tilde{a}(\cdot, \xi) \nabla v(\cdot, \xi) + \int_D (\nabla u(\cdot, \xi))^T \tilde{b}(\cdot, \xi) v(\cdot, \xi) \right) \tag{21}$$

and the linear form

$$\forall v \in \mathcal{V}, \quad F(v) = \mathbb{E}_\xi \left(\int_D \tilde{f}(\cdot, \xi) v(\cdot, \xi) \right) = \langle \tilde{f}, v \rangle_\mathcal{H}.$$

We introduce now assumptions which ensure the continuity and the coercivity of the bilinear form $\mathcal{A}$.

Assumption 7 (A7) There is a constant $\tilde{\Lambda}$ s.t.

$$\forall 1 \leq i, j \leq d, \ \forall x \in \bar{D}, \ \forall z \in \Xi, \quad |\tilde{a}_{ij}(x, z)| \leq \tilde{\Lambda}.$$

Assumption 8 (A8) There is a constant $\tilde{\lambda} > 0$ s.t.

$$\forall y \in \mathbb{R}^d, \ \forall x \in \bar{D}, \ \forall z \in \Xi, \quad y^T \tilde{a}(x, z) y \geq \tilde{\lambda} |y|^2.$$

Let us now recall that thanks to Poincaré's inequalities the norm $v \mapsto \left(\int_D |\nabla v|^2\right)^{1/2}$ is equivalent to $|| \cdot ||_{H^1(D)}$ on $H^1_0(D)$. More precisely

$$\forall v \in H^1_0(D), \quad \left(\int_D |\nabla v|^2\right)^{1/2} \leq ||u||_{H^1(D)} \leq C(d, |D|) \left(\int_D |\nabla v|^2\right)^{1/2}, \quad (22)$$

with

$$C(d, |D|) = \sqrt{1 + \left(\frac{d\,\Gamma(d/2)}{2\pi^{d/2}}|D|\right)^{1/d}} \quad (23)$$

where $\Gamma(\cdot)$ denotes the Gamma function and $|D|$ is the volume of D (cf Gilbarg and Trudinger 1983, p164). With this in mind we introduce the following assumption.

Assumption 9 (A9) Assume (A8) and

$$\forall x \in \bar{D}, \quad \forall z \in \varXi, \quad |\tilde{b}(x, z)| \leq \tilde{M},$$

with $\tilde{M} < \dfrac{\tilde{\lambda}}{\sqrt{2C(d, |D|)}}$ where $C(d, |D|)$ is the constant defined in Eq. 23.

We have the following result:

Lemma 5 *Assume (A7)–(A9). There exists a unique weak solution u at the stochastic level of* Eq. 20, *in the sense that $u \in \mathcal{V}$ and satisfies*

$$\forall v \in \mathcal{V}, \quad \mathcal{A}(u, v) = F(v).$$

This solution u can be approximated by its Galerkin projection $u^{N,P}$, which is the unique element in $\mathcal{V}^{N,P}$ that satisfies

$$\forall v \in \mathcal{V}^{N,P}, \quad \mathcal{A}(u^{N,P}, v) = F(v). \quad (24)$$

More precisely we have

$$||u^{N,P} - u||_{\mathcal{V}} \xrightarrow[N\to\infty, P\to\infty]{} 0.$$

Proof See the Appendix. □

3.2.4 Approximation of the Quantity of Interest $\mathfrak{U}$

The metamodel for $\mathfrak{U}$ is provided by the next theorem:

Theorem 3 *Assume $a(\cdot, \xi)$ belongs to $C^1(D; \mathbb{R}^{d \times d})$ for a.e. value of ξ and that $\tilde{a}$, $\tilde{b}$ and $\tilde{f}$ are defined by* Eq. 19. *Assume (A2)-(A5). Assume (A9). Then the quantity of interest $\mathfrak{U}(x, \xi)$ is a weak solution to* Eq. 20, *and can be approximated by $\mathfrak{U}^{N,P}$ the solution of* Eq. 24, *in the sense that*

$$||\mathfrak{U}^{N,P} - \mathfrak{U}||_{\mathcal{V}} \xrightarrow[N\to\infty, P\to\infty]{} 0.$$

Proof See the Appendix. □

3.2.5 Approximation of Sobol' Indices

We now discuss how to solve Eq. 24, for fixed N, $P \in \mathbb{N}^*$ (we omit some superscripts N, P in what follows). A function

$$\mathfrak{U}^{N,P} = \sum_{j=1}^{N-1} \sum_{q=0}^{P} U_j^q \psi_q \phi_j \in \mathcal{V}^{N,P} \tag{25}$$

satisfies Eq. 24 if and only if

$$\forall 1 \leq i \leq N - 1, \quad \forall 0 \leq p \leq P, \quad \mathcal{A}(\mathfrak{U}^{N,P}, \psi_p \phi_i) = F(\psi_p \phi_i).$$

Therefore if and only if the block vector

$$U = (U^0, \ldots, U^P)^T, \text{ with } U^q = (U_1^q, \ldots, U_{N-1}^q) \in \mathbb{R}^{1 \times (N-1)}, \, 0 \leq q \leq P$$

solves

$$\mathbf{A}U = \mathbf{F} \tag{26}$$

where $\mathbf{A} = (\mathbf{A}_{pq})_{0 \leq p, q \leq P}$ is the block matrix defined by

$$\mathbf{A}_{pq} = \left(\mathcal{A}(\psi_q \phi_j, \psi_p \phi_i)\right)_{1 \leq i, j \leq N-1} \text{ for any } 0 \leq p, q \leq P$$

and $\mathbf{F} = (\mathbf{F}_0, \ldots, \mathbf{F}_P)^T$ is the block vector defined by

$$\mathbf{F}_p = \left(F(\psi_p \phi_1), \ldots, F(\psi_p \phi_{N-1})\right) \in \mathbb{R}^{1 \times (N-1)} \text{ for any } 0 \leq p \leq P.$$

We thus have to solve Eq. 26 and to recover $\mathfrak{U}^{N,P}$ by Eq. 25, as it is usually done in finite elements methods.

Assume then we have got an approximation $\mathfrak{U}^{N,P}(x, \xi) = \sum_{q=0}^{P} \mathfrak{U}_q^N(x) \psi_q(\xi)$ of the random variable $\mathfrak{U}(x, \xi)$, for $x \in D$ (here we have denoted $\mathfrak{U}_q^N(x) = \sum_{j=1}^{N-1} U_j^q \phi_j(x)$, for any $0 \leq q \leq P$). Then, thanks to the orthonormality of the basis $\{\psi_p\}_{p=0}^{P}$, the Sobol' index $S_I(\mathfrak{U})$ can be approximated using Parseval identity by

$$\frac{\sum_{q \in K_I} [\mathfrak{U}_q^N(x)]^2}{\sum_{q=1}^{P} [\mathfrak{U}_q^N(x)]^2} \tag{27}$$

where $K_I = \{p \in \{1, \ldots, P\} : \psi_p(\xi) = \psi_p(\xi_I)\}$ has been defined in Eq. 15.

3.3 Computation of $S_I(\mathfrak{V})$ Using the Parabolic Stochastic PDE

The structure of this subsection is similar to the structure of Section 3.2 about the elliptic case. Note that throughout the subsection we use the notation already introduced in Section 3.2.1. In Section 3.3.1 we reformulate (17) in divergence form. In Section 3.3.2 we introduce the notion of weak solution at the stochastic level of the parabolic PDE in divergence form, and explain how to get an approximation $(\mathfrak{V}^{m,N,P})_m$ of this weak solution. Here we have to perform a discretization in the space variable x, the uncertainty parameter ξ, but also the time variable t. As again our quantity of interest $\mathfrak{V}$ turns out to be a weak solution of the stochastic PDE of interest, we thus get an approximation of it that allows the computation on approximated Sobol' indices.

3.3.1 Parabolic PDE in Divergence Form

As in Section 3.2 we first rewrite (17) in divergence form. We assume again that for a.e. value of ξ the function $a(\cdot, \xi)$ is of class $C^1(D; \mathbb{R}^{d \times d})$ and define the coefficients $\tilde{a}, \tilde{b}$ as in Eq. 19. It is then clear that Eq. 17 is equivalent to

$$
\begin{cases}
\partial_t v(t, x, \xi) = \nabla \cdot (\tilde{a}(x, \xi) \nabla v(t, x, \xi)) - (\nabla v(t, x, \xi))^T \tilde{b}(x, \xi), \ \forall (t, x) \in (0, T] \times D \\[2mm]
v(t, x, \xi) = 0, \quad \forall (t, x) \in (0, T] \times \partial D \\[2mm]
v(0, x, \xi) = f(x), \quad \forall x \in \bar{D}.
\end{cases}
\tag{28}
$$

3.3.2 Weak Solutions at the Stochastic Level of Eq. 28 and Their Approximation; Link with the Quantity of Interest $\mathfrak{V}$

In the same spirit as in Section 3.2.3 we will present the notion of weak solution of Eq. 28 at the stochastic level and of its approximation. The weak solution will be searched for in the space

$$
\left(L^2(0, T; H_0^1(D)) \cap H^1(0, T; H^{-1}(D)) \right) \otimes L^2(\Xi, \mathbb{P}_\xi)
$$

Here the space $L^2(0, T; H_0^1(D)) \cap H^1(0, T; H^{-1}(D))$ denotes the space of those functions v such that v is in $L^2(0, T; L^2(D))$ (i.e. $\int_0^T \int_D v^2 < \infty$), the $\partial_{x_i} v$'s, for $1 \le i \le d$, are in $L^2(0, T; L^2(D))$ too, and $\partial_t v$ is in $L^2(0, T; H^{-1}(D))$ (the derivatives are understood in the distribution sense). Note that we have

$$
\left(L^2(0, T; H_0^1(D)) \cap H^1(0, T; H^{-1}(D)) \right) \otimes L^2(\Xi, \mathbb{P}_\xi) = L^2(0, T; \mathcal{V}) \cap H^1(0; T, \mathcal{V}').
$$

We start with the two following results (in Lemma 6 the form $\mathcal{A}$ is the one defined in Section 3.2).

Lemma 6 *Assume (A7)-(A8) and that $\tilde{b}$ is bounded. Assume $f \in L^2(D)$. There exists a unique weak solution v at the stochastic level to Eq. 28, in the sense that v belongs to $L^2(0, T; \mathcal{V}) \cap H^1(0; T, \mathcal{V}') \cap C(0, T; \mathcal{H})$ and satisfies*

$$
\forall w \in \mathcal{V}, \ \langle \partial_t v(t, \cdot, \cdot), w \rangle_{\mathcal{H}} + \mathcal{A}(v(t, \cdot, \cdot), w) = 0 \ \text{for a.e. } t \in [0, T] \text{ and } v(0, \cdot) = f.
\tag{29}
$$

Proof See the Appendix. $\qquad\square$

Lemma 7 *Assume $a(\cdot, \xi)$ belongs to $C^1(D; \mathbb{R}^{d \times d})$ for a.e. value of ξ and $\tilde{a}$ and $\tilde{b}$ are defined by Eq. 19. Assume (A2)-(A6). Then the quantity of interest $\mathfrak{V}(t, x, \xi)$ is a weak solution to Eq. 28.*

Proof See the Appendix. $\qquad\square$

Remark 7 Note that in the above results we have dropped Assumption (A9) in itself and we only have assumed that $\tilde{b}$ is bounded. This because in the parabolic case we do not need the form $\mathcal{A}$ to be coercive in itself; it suffices to get the coercivity of $(u, v) \mapsto \mathcal{A}(u, v) + \mu \langle u, v \rangle_{\mathcal{H}}$ for μ large enough (see the proof of Lemma 6).

Galerkin Discretisation We now aim at approximating the weak solution of Eq. 28. We first consider a Galerkin type discretisation of Eq. 29 with respect to the space variable and the uncertain parameter, similar to the one we have used for the elliptic case (the time-discretisation will be treated in a second time). We consider the problem of finding $v^{N,P}$ in $C^1(0, T; \mathcal{V}^{N,P})$ satisfying

$$\begin{cases} \forall w \in \mathcal{V}^{N,P}, \ \partial_t v^{N,P}(t, \cdot, \cdot), \ w\rangle_{\mathcal{H}} + \mathcal{A}\left(v^{N,P}(t, \cdot, \cdot), \ w\right) = 0 \text{ for a.e. } t \in [0, T], \\ v^{N,P}(0, \cdot) = f^{N,P}. \end{cases} \tag{30}$$

Here $f^{N,P}$ is an element of $\mathcal{V}^{N,P}$, and the sequence $(f^{N,P})$ satisfies

$$||f^{N,P} - f||_{\mathcal{H}} \xrightarrow[N \to \infty, P \to \infty]{} 0. \tag{31}$$

Let us now rewrite Eq. 30 in matrix form. We have

$$v^{N,P}(t, x, \xi) = \sum_{j=1}^{N-1} \sum_{q=0}^{P} V_j^q(t)\psi_q(\xi)\phi_j(x)$$

and denote $V(t)$ the block vector

$$(V^0(t), \ldots, V^P(t))^T, \text{ with } V^q(t) = \left(V_1^q(t), \ldots, V_{N-1}^q(t)\right) \in \mathbb{R}^{1 \times (N-1)}, \ 0 \le q \le P.$$

We denote $\mathbf{f} = (\mathbf{f}_0, \ldots, \mathbf{f}_P)^T$ the block vector defined by

$$\mathbf{f}_p = (f_{1,p}^{N,P}, \ldots, f_{N-1,p}^{N,P}), \ 0 \le p \le P, \text{ with } f^{N,P} = \sum_{i=1}^{N-1} \sum_{p=0}^{P} f_{i,p}^{N,P} \psi_p \phi_i.$$

Using then the block matrix $\mathbf{A}$ defined in Section 3.2.5 and introducing $\mathbf{M} = (\mathbf{M}_{pq})_{0 \le p,q \le P}$ the block matrix defined by $\mathbf{M}_{pq} = \left(\langle \psi_q \phi_j, \psi_p \phi_i \rangle_{\mathcal{H}}\right)_{1 \le i,j \le N-1}$ for any $0 \le p, q \le P$, it is clear that Eq. 30 is equivalent to

$$\mathbf{M}V'(t) + \mathbf{A}V(t) = 0, \quad V(0) = \mathbf{f}. \tag{32}$$

Equation 32 is a system of Ordinary Differential Equations (ODE). Note that we know by the Cauchy-Lipschitz theorem that this system has a solution in $C^1(0, T; \mathbb{R}^{(N-1) \times (P+1)})$ (see Theorem VII.3 in Brezis 1983); we use here the fact that $\mathbf{M}$ is invertible and constant). It ensures the existence of $v^{N,P}$ in $C^1(0, T; \mathcal{V}^{N,P})$ that solves Eq. 30.

Under for example (A7)–(A8) and $\tilde{b} \equiv 0$ it can be proved (see Corollary 7.4-1 in Raviart and Thomas 2004) that $||v^{N,P}(t, \cdot, \cdot) - v(t, \cdot, \cdot)||_{\mathcal{H}} \to 0$, as $N \to \infty, P \to \infty$, for any $t \in [0, T]$, where v is the weak solution to Eq. 28. Thus $v^{N,P}$ is an approximation of v. However we will not access to $v^{N,P}$ itself but to its approximation through a time finite difference scheme for Eq. 32.

Indeed, in a second time, we introduce a θ-scheme for Eq. 32 ($\theta \in [0, 1]$). We choose $M \in \mathbb{N}^*$ and denote $\Delta t = \frac{T}{M}$ and $t_m = m\Delta t$ for any $0 \le m \le M$. For any $0 \le m \le M$, define $V[m] \in \mathbb{R}^{(N-1) \times (P+1)}$ as the solution of:

$$\frac{1}{\Delta t}\mathbf{M}(V[m+1] - V[m]) + \mathbf{A}(\theta V[m+1] + (1-\theta)V[m]) = 0, \quad V[0] = \mathbf{f},$$

which can be written as:

$$(\mathbf{M} + \theta \Delta t \mathbf{A})V[m+1] = (\mathbf{M} - (1-\theta)\Delta t \mathbf{A})V[m], \quad V[0] = \mathbf{f}. \tag{33}$$

We then have the following result:

Lemma 8 *Assume (A7)–(A8) and that $\tilde{b}$ is bounded. Assume $f \in H^1_0(D)$ and that we have a sequence of functions $f^{N,P} \in \mathcal{V}^{N,P}$ satisfying $||f^{N,P} - f||_\mathcal{H} \to 0$ as $N, P \to \infty$.*

Consider then v the weak solution of Eq. 28 and $v^{m,N,P} = \sum_{j=1}^{N-1} \sum_{q=0}^{P} V_j^q[m]\psi_q\phi_j$, where the $V_j^q[m]$'s are computed by Eq. 33.

Assume $\theta = 1/2$ and $u \in C^1(0, T; \mathcal{V}) \cap C^3(0, T; \mathcal{H})$. Then

$$\sup_{0 \le m \le M} ||v^{m,N,P} - v(t_m, \cdot, \cdot)||_\mathcal{H} \xrightarrow[M\to\infty, N\to\infty, P\to\infty]{} 0.$$

Proof See the Appendix. $\square$

Remark 8 In fact the scheme (33) is unconditionally stable for $\theta \in [\frac{1}{2}, 1]$. If $\theta = 1/2$ and $u \in C^1(0, T; \mathcal{V}) \cap C^3(0, T; \mathcal{H})$ it is consistent and thus convergent (Lax principle), with order $(\Delta t)^2$ in time. See the proof of Lemma 8 in the Appendix for details. For $\theta = 1/2$, this scheme is known as the Crank-Nicholson scheme.

Gathering the above results we are led to the following theorem, which provides the metamodel for $\mathfrak{V}$.

Theorem 4 *Assume $a(\cdot, \xi)$ belongs to $C^1(D; \mathbb{R}^{d \times d})$ for a.e. value of ξ and that $\tilde{a}$ and $\tilde{b}$ are defined by Eq. 19. Assume (A2)–(A5) and that f satisfies (A6).*

Let the sequence $f^{N,P}$ be defined as in Lemma 8. Choose $M \in \mathbb{N}^*$ and compute $\left(\mathfrak{V}^{m,N,P} = \sum_{j=1}^{N-1} \sum_{q=0}^{P} V_j^q[m]\psi_q\phi_j\right)_{0 \le m \le M}$ by solving Eq. 33 with $\theta = 1/2$.

If the quantity of interest $\mathfrak{V}(t, x, \xi)$ is in $C^1(0, T; \mathcal{V}) \cap C^3(0, T; \mathcal{H})$ it can be approximated by $(\mathfrak{V}^{m,N,P}(x, \xi))_{0 \le m \le M}$ in the sense that

$$\sup_{0 \le m \le M} ||\mathfrak{V}^{m,N,P} - \mathfrak{V}(t_m, \cdot, \cdot)||_\mathcal{H} \xrightarrow[M\to\infty, N\to\infty, P\to\infty]{} 0.$$

Remark 9 For any $N \in \mathbb{N}^*$, $P \in \mathbb{N}$, $f_{N,P}$ can be, e.g., built as an interpolation function or as a projection on $\mathcal{V}^{N,P}$.

Remark 10 Note that the assumption that $\mathfrak{V}(t, x, \xi)$ is in $C^1(0, T; \mathcal{V}) \cap C^3(0, T; \mathcal{H})$ is sufficient but not necessary. If f is in C^∞, with f and all its derivatives vanishing at the boundary ∂D, then the aforementioned assumption is satisfied (see, e.g., Theorem VII.5. Brezis 1983). This last assumption is not restrictive in the sense that similar arguments to the ones in Remark 6 can be applied in case it is not satisfied.

Then, as in Eq. 34, the Sobol' index $S_I(\mathfrak{V}(T, x, \xi))$ can be approximated by

$$\frac{\sum_{q \in K_I}[\mathfrak{V}_q^{M,N}(x)]^2}{\sum_{q=1}^{P}[\mathfrak{V}_q^{M,N}(x)]^2} \tag{34}$$

where $\mathfrak{V}_q^{M,N}(x) = \sum_{j=1}^{N-1} V_j^q[M]\phi_j(x)$ for $0 \le q \le P$.

4 Validation and Illustration: Numerical Experiments on a One-Dimensional Ornstein-Uhlenbeck Process

In this section we choose a simple example, namely a one-dimensional Ornstein-Uhlenbeck (OU) process whose coefficients depend on uncertain parameters, in order to validate and illustrate the PDE based approach we propose. The results we obtain are promising and motivate further work to extend the method to more complex settings.

More precisely our toy model is given by the SDE

$$dX_t = -\alpha(\xi)X_t dt + \sigma(\xi)dW_t, \quad t \geq 0, \quad X_0 = x, \tag{35}$$

where W is a one-dimensional standard Brownian motion and $x \in \mathbb{R}$.

We choose $\xi = (\xi_1, \xi_2)^T$ with ξ_1, ξ_2 independent, distributed as $\mathcal{U}([0, 1])$, and independent of W, and define

$$\alpha(\xi) = \mu_1 + \sqrt{3}\sigma_1(2\xi_1 - 1) \quad \text{and} \quad \sigma(\xi) = \mu_2 + \sqrt{3}\sigma_2(2\xi_2 - 1)$$

with $\sigma_i > 0$, $\mu_i > \sqrt{3}\sigma_i$, $i = 1, 2$. Note that the probability distributions of $\alpha(\xi) > 0$ and $\beta(\xi) > 0$ are then given by

$$\alpha(\xi) \sim \mathcal{U}\left([\mu_1 - \sqrt{3}\sigma_1, \mu_1 + \sqrt{3}\sigma_1]\right) \quad \text{and} \quad \sigma(\xi) \sim \mathcal{U}\left([\mu_2 - \sqrt{3}\sigma_2, \mu_2 + \sqrt{3}\sigma_2]\right). \tag{36}$$

Note that the truncated orthonormal basis $\{\Psi_q\}_{q=0}^P$ is then chosen as the tensorized Legendre polynomials basis.

To start with, we have the following straightforward result.

Proposition 1 *Consider the stochastic differential equation described by* Eq. 35. *Assume ξ and $\alpha(\xi)$, $\sigma(\xi)$ are defined as above.*

Assumptions A1 and A2 are then trivially satisfied. In addition the coefficients in Eq. 35 satisfy A3 and we thus know that Eq. 35 *admits a solution.*

We set $D = (0, 10)$ and start with studying $S_I(\mathfrak{U})$ in Section 4.1. We set $\mathfrak{V}(T, x, \xi) = \mathbb{P}^x(T \leq \tau_D|\xi)$ (this corresponds to $f \equiv 1$ in Eq. 6) and will study $S_I(\mathfrak{V})$ in Section 4.2.

The OU process is a well marked test case. For example, for any $t \geq 0$, the solution to Eq. 35 is given by $X_t = xe^{-\alpha(\xi)t} + \sigma(\xi)\int_0^t e^{-\alpha(\xi)(t-s)}dW_s$ (see Karatzas and Shreve 1991, Example 5.6.8).

4.1 Study of $S_I(\mathfrak{U})$

Proposition 2 *Consider the stochastic differential equation described by* Eq. 35. *Assume that $D = (0, 10)$. Assume ξ and $\alpha(\xi)$, $\sigma(\xi)$ are defined as above.*

In addition to (A1)–(A3), the Assumptions 4 and 5 are also satisfied. Thus, for any $x \in D$, the quantity $\mathfrak{U}(x, \xi)$ is given by the solution of the stochastic elliptic PDE

$$\begin{cases} \frac{1}{2}\sigma^2(\xi)\partial_{xx}^2 u(x, \xi) - \alpha(\xi)x\,\partial_x u(x, \xi) = -1 \; \forall x \in D \\ \\ u(0, \xi) = u(1, \xi) = 0. \end{cases} \tag{37}$$

Proof (A4) is satisfied because of Eq. 36. As we are in dimension one in space ($d = 1$), Assumption 5 will be trivially met by any segment $(a, b) \subset \mathbb{R}$. Using Theorem 1 we get the last statement of the proposition. □

In view of Proposition 2 we wish now to use Theorem 3 in order to compute an approximation of the quantity of interest $\mathfrak{U}(x, \xi)$, and then compute the Sobol' indices by Eq. 34.

Note that Eq. 37 may trivially be rewritten in the form Eq. 20 with $\tilde{a}(x, z) = \sigma^2(z)/2$, $\tilde{b}(x, z) = \alpha(z) x$ and $\tilde{f}(x, z) = 1$, for $x \in \bar{D}$, $z \in \Xi$. In order to check (A9) we have to impose that

$$10(\mu_1 + \sqrt{3}\sigma_1) < \frac{(\mu_2 - \sqrt{3}\sigma_2)^2}{2(24)^{1/4}}. \tag{38}$$

Then we have the following result.

Proposition 3 *In the context of Proposition 2, assume that σ_1, σ_2, μ_1 and μ_2 satisfy* Eq. 38. *Then $\mathfrak{U}(x, \xi)$ can be approximated by the solution $\mathfrak{U}^{N,P}$ of Eq.* 24 *with, $\forall u, v \in \mathcal{V}$,*

$$\mathcal{A}(u, v) = \mathbb{E}_\xi \left(\int_0^{10} \frac{\sigma^2(\xi)}{2} \partial_x u(x, \xi) \partial_x v(x, \xi) dx + \int_0^{10} \alpha(\xi) x \, \partial_x u(x, \xi)) v(x, \xi) dx \right)$$

and $F(v) = \mathbb{E}_\xi \left(\int_0^{10} v(x, \xi) \right) dx$, $\forall v \in \mathcal{V}$.

Let $N, P \in \mathbb{N}^*$ fixed. Denote $x_i = 10\frac{i}{N}$, $1 \le i \le N$, and $h = 10/N$. We use as finite element basis functions the ϕ_i, $1 \le i \le N - 1$, defined by $\phi_i(x) = \phi(\frac{x - x_i}{h})$ with $\phi(x) = \mathbf{1}_{|x| \le 1}(1 - |x|)$. Note that $\phi_1(0) = \phi_{N-1}(10) = 0$.

In order to follow the program proposed in Proposition 3 we have then to solve $\mathbf{A}U = \mathbf{F}$, with $\mathbf{A}$ and $\mathbf{B}$ defined as after Eq. 26. Here simple computations show that ($\otimes$ stands for the Kronecker product)

$$\mathbf{A} = \Sigma \otimes A + \beta \otimes B$$

with

$$\Sigma = \left(\mathbb{E}_\xi \left[\frac{\sigma^2(\xi)}{2} \psi_q(\xi) \psi_p(\xi) \right] \right)_{0 \le p, q \le P} \quad \text{and} \quad \beta = \left(\mathbb{E}_\xi \left[\alpha(\xi) \psi_q(\xi) \psi_p(\xi) \right] \right)_{0 \le p, q \le P},$$

$$A = \left(\int_0^{10} \partial_x \phi_j \partial_x \phi_i \right)_{1 \le i, j \le N-1} = \frac{1}{h} \begin{pmatrix} 2 & -1 & & & 0 \\ -1 & 2 & & & \\ & & \ddots & & \\ & & & 2 & -1 \\ 0 & & & -1 & 2 \end{pmatrix},$$

$$B = \left(\int_0^{10} x \partial_x \phi_j(x) \phi_i(x) dx \right)_{1 \le i, j \le N-1} = \begin{pmatrix} -\frac{h}{3} & \frac{x_1}{2} + \frac{h}{6} & & & 0 \\ -\frac{x_2}{2} + \frac{h}{6} & -\frac{h}{3} & & & \\ & & \ddots & & \\ & & & -\frac{h}{3} & \frac{x_{N-2}}{2} + \frac{h}{6} \\ 0 & & & -\frac{x_{N-1}}{2} + \frac{h}{6} & -\frac{h}{3} \end{pmatrix}.$$

Note that Σ, β and A are symmetric but not B. Besides $\mathbf{F} = (\mathbf{F}_p)_{0 \le p \le P}$ is simply given by $\mathbf{F}_0 = h(1, \ldots, 1)^T \in \mathbb{R}^{N-1}$ and $\mathbf{F}_p = 0$ for $1 \le p \le P$.

We will compare our method with crude Monte Carlo estimator. In our Monte Carlo estimation of the exit time τ_D we use a Euler scheme and a boundary shifting method to be described in Gobet (2016).

More precisely the function $\widehat{\mathcal{F}}_{\mathfrak{U}}$ we use (see Section 2.3.1) can be described as:

INPUT: the brownian path W, the uncertain parameter ξ and the time step δt.
COMPUTE: $\sigma(\xi) = \mu_2 + \sqrt{3}\sigma_2(2\xi_2 - 1)$, $\alpha(\xi) = \mu_1 + \sqrt{3}\sigma_1(2\xi_1 - 1)$ and $c(\xi) = 0.5826\,\sigma(\xi)\,\sqrt{\delta t}$.
INITIALIZE: $X = x$, $t = 0$ and $d = 100$ (in fact any value that is greater than $c(\xi)$, for any ξ).

      While { $X \in (0, 10)$ and $c(\xi) < d$ }
          Do: Set $X = X + \sqrt{\sigma(\xi)}\sqrt{\delta t}(W_{t+\delta t} - W_t) - \alpha(\xi)\, X\, \delta t$
              $t = t + \delta t$
              $d = \text{dist}(X, \partial D)$
      Endwhile

OUTPUT: the time t that is reached when exiting the while loop.

In Table 1 we illustrate the efficiency of our approach with $\mu_1 = 1$, $\mu_2 = 9$ and $\sigma_1 = \sigma_2 = 0.2$, and with the starting point $x = 5$.

4.2 Study of $S_l(\mathfrak{V})$

Recall that we consider $\mathfrak{V}(T, x, \xi) = \mathbb{P}^x(T \leq \tau_D | \xi)$ (this corresponds to $f \equiv 1$ in Eq. 6). Note that $f \equiv 1$ do not satisfy all the requirements of (A6). In particular it does not satisfy the uniform Dirichlet boundary condition. However it is possible to approach f by f_ε satisfying (A6). Using the construction based on convolution arguments suggested in Remark 6, we get $f_\varepsilon \equiv f$ on K_ε with K_ε a compact subset of D such that $|D \setminus K_\varepsilon| \leq \varepsilon$. A consequence is that, numerically, if we choose a discretization step small enough, f_ε will coincide with f.

We take the same finite element basis ϕ_i, $1 \leq i \leq N - 1$ as in Section 4.1. We will use M time steps. We take $\epsilon < h$ so that f_ϵ can be simply interpolated by

$$f^{N,P} = \sum_{i=1}^{N-1} f_\epsilon(x_i)\phi_i = \sum_{i=1}^{N-1} \phi_i$$

and thus the block vector $\mathbf{f} = (\mathbf{f}_p)_{0 \leq p \leq P}$ is given by $\mathbf{f}_0 = (1, \ldots, 1)^T \in \mathbb{R}^{N-1}$ and $\mathbf{f}_p = 0$ for $1 \leq p \leq P$. The block matrix $\mathbf{A}$ remains as in Section 4.1. In order to solve Eq. 33 (with $\theta = \frac{1}{2}$) it remains to compute $\mathbf{M}$. But simple computations show that

$$\mathbf{M} = \mathcal{M} \otimes M_1$$

Table 1 Sobol' indices for $\mathfrak{U}(x, \xi)$, $x = 5$, in the Ornstein-Uhlenbeck example, with $\mu_1 = 1$, $\mu_2 = 9$ and $\sigma_1 = \sigma_2 = 0.2$

	Galerkin	Double MC
$S_1(\mathfrak{U})$	$(N = 1000, P = 10)$ 0.0253	$(N = M = 10^4, \delta t = 10^{-3})$ 0.024927
$S_2(\mathfrak{U})$	$(N = 1000, P = 10)$ 0.9747	$(N = M = 2 \times 10^4, \delta t = 10^{-4})$ 0.971277

Table 2 Sobol' indices for $\mathfrak{V}(T, x, \xi)$, $x = 1$, $T = 0.3$ in the Ornstein-Uhlenbeck example, with $\mu_1 = 1$, $\mu_2 = 2$ and $\sigma_1 = \sigma_2 = 0.2$

	Galerkin / Crank-Nicholson	Double MC
$S_1(\mathfrak{V})$	$(M = 300, N = 1000, P = 10)$ 0.0961	$(N = M = 5 \times 10^4, \delta t = 6 \times 10^{-4})$ 0.095812
$S_2(\mathfrak{V})$	$(M = 300, N = 1000, P = 10)$ 0.9039	$(N = M = 5 \times 10^4, \delta t = 6 \times 10^{-4})$ 0.903182

with $\mathcal{M} = \left(\mathbb{E}_\xi \left[\psi_q(\xi) \psi_p(\xi) \right] \right)_{0 \le p, q \le P}$ and

$$M_1 = \left(\int_0^{10} \phi_j(x) \phi_i(x) \, dx \right)_{1 \le i, j \le N-1} = \begin{pmatrix} 2h/3 & h/6 & & & 0 \\ h/6 & 2h/3 & & & \\ & & \ddots & & \\ & & & 2h/3 & h/6 \\ 0 & & & h/6 & 2h/3 \end{pmatrix}.$$

We compare our method with crude Monte Carlo estimator. That is to say we use in Eq. 11 a function $\widehat{\mathcal{F}}_{\mathfrak{V}}$ that is constructed in the same spirit as $\widehat{\mathcal{F}}_{\mathfrak{U}}$ in Section 4.1 (in particular we use again a boundary shifting method). In Table 2 we illustrate the efficiency of our approach with $\mu_1 = 1$, $\mu_2 = 2$ and $\sigma_1 = \sigma_2 = 0.2$, and with the starting point $x = 1$ and the time horizon $T = 0.3$.

Appendix

A.1 Integrability of the Stochastic Quantities of Interest

Proof (*Proof of Lemma 1*) Let $q \in \mathbb{N}^*$ and $t \ge 0$. Using Problem 5.2.15 in Karatzas and Shreve (1991), we can claim that, under (A3), for any $z \in \mathcal{E}$

$$\int_C |X_t(\theta, z)|^q \mathbb{W}(d\theta) \le C(1 + |x|^q) \exp(Ct), \tag{39}$$

where $C = C(t, M)$. As the constant M in Eq. 1 is uniform in z, the constant C does not depend on z either. Integrating Eq. 39 over $\mathcal{E}$ and against $\mathbb{P}_\xi$ we get the desired result. □

Proof (*Proof of Lemma 2*) Here we revisit the proof of Lemma 5.7.4 in Karatzas and Shreve (1991) and give details for the sake of completeness. Thanks to (A4), we have for example $a_{11}(x, z) \ge \lambda$ for any $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$, $z \in \mathcal{E}$. We note $q = \min_{x \in \bar{D}} x_1$ and b a constant s.t. $|b(x, z)| \le b < \infty$, $\forall x \in D, \forall z \in \mathcal{E}$.

Such a finite constant b exists thanks to Eq. 1. Set now $\nu > 2b/\lambda$ and consider the function $h(x) = -\mu \exp(\nu x^1)$, $x \in D$, where the constant μ will be determined later. This function is of class $C^\infty(D)$ and satisfies for any $x \in D$ and $z \in \mathcal{E}$,

$$-\frac{1}{2} \sum_{i,j=1}^d a_{ij}(x, z) \partial^2_{x_i x_j} h(x) - \sum_{j=1}^d b_j(x, z) \partial_{x_i} h(x)$$
$$= \mu e^{\nu x^1} \left(\frac{1}{2} \nu^2 a_{11}(x, z) + \nu b_1(x, z) \right) \ge \frac{1}{2} \mu \nu \lambda e^{\nu q} \left(\nu - \frac{2b}{\lambda} \right).$$

Choosing μ sufficiently large, we have for any $x \in D$, any $z \in \varXi$:

$$\frac{1}{2} \sum_{i,j=1}^{d} a_{ij}(x,z)\partial^2_{x_i x_j} h(x) + \sum_{j=1}^{d} b_j(x,z)\partial_{x_i} h(x) \leq -1.$$

Let us write $\tau_D(\omega) = \tau_D(\theta, z)$. As $X_0 = x \in D$, using Itô Formula and taking the expectation against $\mathbb{W}$ we get:

$$\forall z \in \varXi \int_C (t \wedge \tau_D(\theta, z)\mathbb{W}(d\theta) \leq h(x) - \int_C h(X_{t \wedge \tau_D}(\theta, z))\mathbb{W}(d\theta) \leq 2 \max_{y \in \bar{D}} |h(y)|.$$

Note that in the above computation, the expectation of the stochastic integral vanishes, as h is bounded with bounded derivatives. Then, integrating over $\varXi$ and against $\mathbb{P}_\xi$ the above inequality we get the result by monotone convergence. □

Proof (*Proof of Lemma 3*) According to Theorem 1, the function $\mathfrak{U}(\cdot, \xi)$ solves Eq. 16 (for a.e. fixed value of ξ). We now use a maximum principle argument to get an a priori bound on $\mathfrak{U}(\cdot, \xi)$. Without loss of generality we assume that D lies in the slab $0 < x^1 < \delta$ for a certain $0 < \delta < \infty$ (the general case can be recovered by translation arguments). According to Theorem 3.7 in Gilbarg and Trudinger (1983) and its proof we have

$$\sup_{x \in \bar{D}} |\mathfrak{U}(x, \xi)| \leq \frac{C'(\xi)}{\lambda} \tag{40}$$

where $C'(\xi) = e^{\alpha(\xi)\delta} - 1$ with $\alpha(\xi)$ a quantity chosen s.t. $\alpha(\xi) \geq 1 + \frac{1}{\lambda}\sup_{x \in D} |b(x,\xi)|$ (note that the uniform ellipticity constant λ in (A4) does not depend on z). Then, thanks to the uniformity in z of the constant M in Eq. 1 and to the boundedness of D, the quantity $\alpha(\xi)$ may be chosen independent on ξ. Therefore $C'(\xi) = C'$ does not depend on ξ, and thus we get the result with $C = C'/\lambda$. □

A.2 Results on Stochastic Partial Differential Equations

Proof (*Proof of Lemma 5*) As $\tilde{f}$ is uniformly bounded in x and z, it is easy to prove that the form F is continuous on $\mathcal{V}$ (with the help of Cauchy-Schwarz's inequality). Thanks to (A7) and (A9), the bilinear form $\mathcal{A}$ is continuous on $\mathcal{V}$. Thanks to (A8)–(A9) one may check the coercivity of $\mathcal{A}$. Indeed, using Lemma 8.4 in Gilbarg and Trudinger (1983), one gets for any $z \in \varXi$,

$$\int_D (\nabla v(\cdot, z))^T \tilde{a}(\cdot, z)\nabla v(\cdot, z) + \int_D (\nabla v(\cdot, z))^T \tilde{b}(\cdot, z)v(\cdot, z)$$
$$\geq \frac{\tilde{\lambda}}{2} \int_D |\nabla v(\cdot, z)|^2 - \tilde{\lambda}\nu^2 \int_D v^2(\cdot, z) \tag{41}$$

where $\nu = \frac{\tilde{M}}{\tilde{\lambda}}$. Using now the definition of $||\cdot||_{H^1(D)}$, Inequality Eq. 22, and integrating Eq. 41 over $\varXi$ against $\mathbb{P}_\xi$ we get

$$\mathcal{A}(v, v) \geq \left(\frac{\tilde{\lambda}}{2C(d, |D|)} - \tilde{\lambda}\nu^2\right) ||v||^2_{\mathcal{V}}$$

(we recall that $C(d, |D|)$ is defined by Eq. 23). As $\tilde{M} < \dfrac{\tilde{\lambda}}{\sqrt{2C(d, |D|)}}$, we have that

$\dfrac{\tilde{\lambda}}{2C(d, |D|)} - \tilde{\lambda}\nu^2 > 0$. The existence of a unique weak solution at the stochastic level then follows from the Lax-Milgram theorem.

Now, as F and $\mathcal{A}$ are continuous, as $\mathcal{A}$ is coercive and as $\mathcal{V}^{N,P} \subset \mathcal{V}$, the existence of a unique solution to Eq. 24 follows from Lax-Milgram theorem again, but applied this time on $\mathcal{V}^{N,P}$. In order to prove the convergence result, we first notice that thanks to the Céa lemma we have for any $N, P \in \mathbb{N}^*$

$$||u - u^{N,P}||_{\mathcal{V}} \leq \tilde{C} \min_{v \in \mathcal{V}^{N,P}} ||u - v||_{\mathcal{V}},$$

where the constant $\tilde{C}$ depends on $\tilde{\lambda}$, $\tilde{\Lambda}$, $\tilde{M}$ and $C(d, |D|)$. Second, we recall that $\mathcal{V}^{N,P} \subset \mathcal{V}^{N+1,P+1}$ for any N, P, and that $\bigcup_{N,P} \mathcal{V}^{N,P}$ is dense in $\mathcal{V}$. This is sufficient in order to prove that $\min_{v \in \mathcal{V}^{N,P}} ||u - v||_{\mathcal{V}} \to 0$, as $N, P \to \infty$, and the result follows. $\qquad \square$

Proof (*Proof of Theorem 3*) By Corollary 1, $\mathfrak{U}(x, \xi)$ is a classical solution of Eq. 16. Then, using Theorem 6.6 in Gilbarg and Trudinger (1983) we get that, for a.e. value of ξ

$$\forall 1 \leq i \leq d, \quad \sup_{x \in \bar{D}} |\partial_{x_i} \mathfrak{U}(x, \xi)| \leq C''(\sup_{x \in \bar{D}} |\mathfrak{U}(x, \xi)| + 1),$$

with C'' depending on d, M, $|D|$ and λ, but not of ξ. But, as already noticed in the proof of Lemma 3, we have $\sup_{x \in \bar{D}} |\mathfrak{U}(x, \xi)| \leq C$ with C depending on M, $|D|$ and λ. Thus $\mathfrak{U}(x, \xi)$ belongs to the space $\mathcal{V}$.

As noticed in Section 3.2 any classical solution of Eq. 16 is a classical solution of Eq. 20). Let $v \in C_c^1(D) \otimes L^2(\Xi, \mathbb{P}_\xi)$. Multiplying the first line of Eq. 20 by v, integrating over $D \times \Xi$ against $dx \otimes \mathbb{P}_\xi(dz)$, and performing integration by parts w.r.t the variable x, we get that $\mathcal{A}(\mathfrak{U}, v) = F(v)$. Thanks to Eq. 1 in (A3), the coefficients $\tilde{a}$ and $\tilde{b}$ are obviously bounded uniformly in x and z, because D is bounded, and we have therefore (A7) and the continuity of the form $\mathcal{A}$. Using then density arguments we get $\mathcal{A}(\mathfrak{U}, v) = F(v)$ for all $v \in \mathcal{V}$. (A8) is a consequence of (A4), and we are under (A9). Thus we may get the approximation result by applying Lemma 5. $\qquad \square$

Proof (*Proof of Lemma 6*) In the proof of Lemma 5 we have seen that (A7) and the boundedness of $\tilde{b}$ imply that the form $\mathcal{A}$ is continuous on $\mathcal{V}$.

We now prove that we can find $\mu \geq 0$ large enough such that the form $\mathcal{A}_\mu : (u, v) \mapsto \mathcal{A}(u, v) + \mu \langle u, v \rangle_{\mathcal{H}}$ is coercive on $\mathcal{V}$. Let us denote $\tilde{M}$ the upper bound for $\tilde{b}$ (i.e. $|\tilde{b}(x, z)| \leq \tilde{M}$ for any $x \in \bar{D}$, $z \in \Xi$). Proceeding as for Eq. 41 (that we integrate over Ξ against $\mathbb{P}_\xi$) we get that for any $v \in \mathcal{V}$

$$\mathcal{A}_\mu(v, v) \geq \frac{\tilde{\lambda}}{2} \mathbb{E}_\xi \left(\int_D |\nabla v(\cdot, \xi)|^2 \right) + \left(\mu - \frac{\tilde{M}^2}{\tilde{\lambda}} \right) \mathbb{E}_\xi \left(\int_D v^2(\cdot, \xi) \right).$$

Note that if $\tilde{M} < \dfrac{\tilde{\lambda}}{\sqrt{2C(d, |D|)}}$ it suffices to take $\mu = 0$ (see the proof of Lemma 5). If this is not the case we take $\mu > \dfrac{\tilde{M}^2}{\tilde{\lambda}}$ and set $c = \min\left(\dfrac{\tilde{\lambda}}{2}, \mu - \dfrac{\tilde{M}^2}{\tilde{\lambda}} \right)$. We then have $\mathcal{A}_\mu(v, v) \geq c||v||_{\mathcal{V}}^2$, for any $v \in \mathcal{V}$.

This is sufficient to prove the result (see Lemma 7.1.2 and Theorem 7.1.4 in Ern and Guermond 2002; see also the earlier reference Lions and Magenes 1972). $\qquad\square$

Proof (*Proof of Lemma 7*) By Corollary 1, $\mathfrak{V}(t, x, \xi)$ is a classical solution of Eq. 17. Using Theorem 5.14 in Lieberman (1996) we get that for a.e. value of ξ

$$\sup_{(t,x)\in[0,T]\times\bar{D}} |\mathfrak{V}(t, x, \xi)| + \sup_{(t,x)\in[0,T]\times\bar{D}} |\partial_t \mathfrak{V}(t, x, \xi)| + \sum_{i=1}^{d} \sup_{(t,x)\in[0,T]\times\bar{D}} |\partial_{x_i} \mathfrak{V}(t, x, \xi)|$$

is bounded above by some constant C that depends on M, $|D|$, λ and d but not on ξ. We use that the coefficients a and b are Lipschitz continuous and that the initial condition f is in the H
"older space $H_{2+\alpha}$ for $\alpha = 1$; see pp 46–47 of Lieberman (1996) for more details on Hölder spaces and norms in the parabolic setting. The point is that here we can write these properties of a, b and f with quantities that are uniform w.r.t. the uncertain parameter ξ. Thus the function $\mathfrak{V}(t, x, \xi)$ belongs to the space $L^2(0, T; \mathcal{V}) \cap H^1(0; T, \mathcal{V}')$. As it is continuous in time and bounded (in particular uniformly w.r.t. ξ) it is also in $C([0, T]; \mathcal{H})$.

As noticed in Section 3.3.1 the function $\mathfrak{V}(t, x, \xi)$ is a classical solution of Eq. 28. Multiplying the first line of Eq. 28 by a test function in $C_c^1(D) \otimes L^2(\varXi, \mathbb{P}_\xi)$, integrating over $D \times \varXi$ against $dx \otimes \mathbb{P}_\xi(dz)$, and using density arguments we get

$$\forall w \in \mathcal{V}, \langle \partial_t \mathfrak{V}(t, \cdot, \cdot), w \rangle_{\mathcal{H}} + \mathcal{A}(\mathfrak{V}(t, \cdot, \cdot), w) = 0, \forall t \in [0, T] \text{ and } \mathfrak{V}(0, \cdot, \cdot) = f.$$

The result is proved. $\qquad\square$

Proof (*Proof of Lemma 8*) Remember that in the proof of Lemma 6 we have seen that we can find $\mu \geq 0$ large enough such that the form $\mathcal{A}_\mu$ is coercive. Here for simplicity we assume that $\mu = 0$, i.e. $\mathcal{A}$ is coercive (with constant c). We claim that this is without loss of generality. Indeed if $\mu > 0$ we consider v_μ the solution of

$$\forall w \in \mathcal{V}, \langle \partial_t v_\mu(t, \cdot), w \rangle_{\mathcal{H}} + \mathcal{A}_\mu\left(v_\mu(t, \cdot), w\right) = 0 \text{ for a.e. } t \in [0, T] \text{ and } v_\mu(0, \cdot) = f \tag{42}$$

(the solution exists because the form $\mathcal{A}_\mu$ is continuous and coercive). Then it is obvious that the function $v(t, \cdot) = e^{\mu t} v_\mu(t, \cdot)$ solves Eq. 29 (in fact Eqs. 29 and 42 are equivalent via the change of variable). Thus one may approach v_μ by the θ-scheme and get an approximation of v applying again the change of variable.

Then for any N, P the form $\mathcal{A}$ defines a scalar product on $\mathcal{V}^{N,P}$, so that for any $u \in \mathcal{V}$ there is by the Riesz theorem an element $\Pi^{N,P} u$ of $\mathcal{V}^{N,P}$ such that

$$\forall w \in \mathcal{V}^{N,P}, \quad \mathcal{A}(\Pi^{N,P} u, w) = \mathcal{A}(u, w).$$

The application $\Pi^{N,P}$ is linear and continuous from $\mathcal{V}$ to $\mathcal{V}^{N,P}$ and is called the elliptic projection operator. It satisfies

$$\forall u \in \mathcal{V}, \quad \forall w \in \mathcal{V}^{N,P}, \quad \mathcal{A}(u - \Pi^{N,P} u, w) = 0.$$

Then, following the proof of Céa's Lemma (Theorem 3.1-2 in Raviart and Thomas 2004) one can prove that

$$||u - \Pi^{N,P} u||_{\mathcal{V}} \leq \tilde{C} \min_{w \in \mathcal{V}^{N,P}} ||u - w||_{\mathcal{V}},$$

where again $\tilde{C}$ depends on the continuity and coercivity constants of $\mathcal{A}$. Using the monotonicity and density assumptions on the $\mathcal{V}^{N,P}$'s we thus get that

$$\forall u \in \mathcal{V}, \quad ||u - \Pi^{N,P} u||_{\mathcal{V}} \xrightarrow[N\to\infty, P\to\infty]{} 0. \tag{43}$$

In the sequel we note Π for $\Pi^{N,P}$ in order to lighten notations. For any $y \in C(0, T; \mathcal{H})$, $t \in [0, T]$ we denote $y(t) := y(t, \cdot, \cdot)$.

Then for any $0 \le m \le M$ we set

$$e_{N,P}^m = v^{m,N,P} - \Pi v(t_m)$$

(here we follow for example Raviart and Thomas (2004) §77.5). Note that Eq. 33 is equivalent to: $\forall 0 \le m \le M - 1, \forall w \in \mathcal{V}^{N,P}$,

$$\frac{1}{\Delta t} \langle v^{m+1,N,P} - v^{m,N,P}, w \rangle_{\mathcal{H}} + \mathcal{A}\left(\theta v^{m+1,N,P} + (1-\theta)v^{m,N,P}, w\right) = 0,$$

so that by some algebraic computations we see that: $\forall 0 \le m \le M - 1, \forall w \in \mathcal{V}^{N,P}$,

$$\frac{1}{\Delta t} \langle e_{N,P}^{m+1} - e_{N,P}^m, w \rangle_{\mathcal{H}} + \mathcal{A}\left(\theta e_{N,P}^{m+1} + (1-\theta)e_{N,P}^m, w\right) = \langle \varepsilon_{N,P}^m, w \rangle_{\mathcal{H}} \tag{44}$$

where $\varepsilon_{N,P}^m \in \mathcal{V}'$ is defined for any $0 \le m \le M - 1$ by: $\forall w \in \mathcal{V}$,

$$\langle \varepsilon_{N,P}^m, w \rangle_{\mathcal{H}} = -\frac{1}{\Delta t} \langle \Pi v(t_{m+1}) - \Pi v(t_m), w \rangle_{\mathcal{H}} - \mathcal{A}\left(\theta v(t_{m+1}) + (1-\theta)v(t_m), w\right). \tag{45}$$

In the sequel we denote e^m (resp. ε) for $e_{N,P}^m$ (resp. $\varepsilon_{N,P}^m$).

We now aim at showing that for any $\theta \in [\frac{1}{2}, 1]$ we have the stability result

$$\forall 1 \le m \le M, \quad ||e^m||_{\mathcal{H}}^2 \le ||e^0||_{\mathcal{H}}^2 + \frac{\Delta t}{c} \sum_{k=1}^{m-1} ||\varepsilon^k||_{\mathcal{V}'}^2. \tag{46}$$

Then we will aim at controlling the $||\varepsilon^k||_{\mathcal{V}'}^2$'s (consistency result; we will only treat the case $\theta = \frac{1}{2}$). This will allow to get convergence.

Stability. Here we adapt the energy estimate method to be found pp66–67 in Dautray and Lions (1993). Let $\theta \in [\frac{1}{2}, 1]$ and fix $0 \le k \le M - 1$. Taking $w = \theta e^{k+1} + (1-\theta)e^k =: \bar{e}^k$ in Eq. 44 we get

$$\frac{1}{\Delta t} \langle e^{k+1} - e^k, \bar{e}^k \rangle_{\mathcal{H}} + \mathcal{A}(\bar{e}_k, \bar{e}_k) = \langle \varepsilon^k, \bar{e}^k \rangle_{\mathcal{H}}. \tag{47}$$

Using then the algebraic equality

$$\langle e^{k+1} - e^k, \theta e^{k+1} + (1-\theta)e^k \rangle_{\mathcal{H}} = \frac{1}{2}||e^{k+1}||_{\mathcal{H}}^2 - \frac{1}{2}||e^k||_{\mathcal{H}}^2 + (\theta - \frac{1}{2})||e^{k+1} - e^k||_{\mathcal{H}}^2$$

in Eq. 47 we get

$$\frac{1}{2}||e^{k+1}||_{\mathcal{H}}^2 - \frac{1}{2}||e^k||_{\mathcal{H}}^2 + \left(\theta - \frac{1}{2}\right)||e^{k+1} - e^k||_{\mathcal{H}}^2 + \Delta t \mathcal{A}(\bar{e}_k, \bar{e}_k) = \Delta t \langle \varepsilon^k, \bar{e}^k \rangle_{\mathcal{H}}$$

and then

$$\frac{1}{2}||e^{k+1}||_{\mathcal{H}}^2 - \frac{1}{2}||e^k||_{\mathcal{H}}^2 + \Delta t \mathcal{A}(\bar{e}_k, \bar{e}_k) \le \Delta t \langle \varepsilon^k, \bar{e}^k \rangle_{\mathcal{H}}. \tag{48}$$

But for any $f \in \mathcal{V}'$ and any $w \in \mathcal{V}$ we have, using Young's inequality and the coercivity of $\mathcal{A}$,

$$\langle f, w \rangle_{\mathcal{H}} \le ||f||_{\mathcal{V}'} ||w||_{\mathcal{V}} \le \frac{c}{2}||w||_{\mathcal{V}}^2 + \frac{1}{2c}||f||_{\mathcal{V}'}^2 \le \frac{1}{2}\mathcal{A}(w, w) + \frac{1}{2c}||f||_{\mathcal{V}'}^2.$$

Using this in Eq. 48 we get

$$\frac{1}{2}||e^{k+1}||^2_{\mathcal{H}} - \frac{1}{2}||e^k||^2_{\mathcal{H}} + \Delta t \mathcal{A}(\bar{e}_k, \bar{e}_k) \le \Delta t \left(\frac{1}{2} \mathcal{A}(\bar{e}^k, \bar{e}^k) + \frac{1}{2c}||e^k||^2_{\mathcal{V}'} \right)$$

and then $||e^{k+1}||^2_{\mathcal{H}} - ||e^k||^2_{\mathcal{H}} \le \frac{\Delta t}{c}||e^k||^2_{\mathcal{V}'}$. Summing over $0 \le k \le m-1$ for any $1 \le m \le M$ we get Eq. 46.

Consistency. Here we follow the lines of the proof of Lemma 7.5-1 in Raviart and Thomas (2004). Using Eq. 29 and the fact that $v \in C^1(0, T; \mathcal{V})$ we can rewrite Eq. 45 into: $\forall w \in \mathcal{V}$,

$$\langle \varepsilon^m, w \rangle_{\mathcal{H}} = \langle \theta \partial_t v(t_{m+1}) + (1-\theta)\partial_t v(t_m) \rangle_{\mathcal{H}} - \frac{1}{\Delta t} \langle \Pi v(t_{m+1}) - \Pi v(t_m), w \rangle_{\mathcal{H}}$$

and further: $\forall w \in \mathcal{V}$,

$$\langle \varepsilon^m, w \rangle_{\mathcal{H}} = \langle \theta \partial_t v(t_{m+1}) + (1-\theta)\partial_t v(t_m) \rangle_{\mathcal{H}} - \frac{1}{\Delta t} \langle v(t_{m+1}) - v(t_m), w \rangle_{\mathcal{H}}$$

$$+ \frac{1}{\Delta t} \int_{t_m}^{t_{m+1}} \langle (I - \Pi)\partial_t v(s), w \rangle_{\mathcal{H}} \, ds.$$

From this we get that for any $0 \le m \le M-1$

$$||\varepsilon^m||_{\mathcal{V}'} = ||\varepsilon^m||_{\mathcal{H}} \le ||\eta(t_m)||_{\mathcal{H}} + \frac{1}{\Delta t} \int_{t_m}^{t_{m+1}} ||(I - \Pi)\partial_t v(s)||_{\mathcal{H}} \, ds$$

with

$$\eta(t_m) = \frac{1}{\Delta t}(v(t_{m+1}) - v(t_m)) - \theta \partial_t v(t_{m+1}) - (1-\theta)\partial_t v(t_m).$$

We take now $\theta = \frac{1}{2}$. Using a Taylor expansion with rest in integral form and the fact that $v \in C^3(0, T; \mathcal{H})$ we get:

$$\eta(t_m) = \frac{1}{\Delta t}\left(\partial_t v(t_m)\Delta t + \frac{1}{2}\partial_{t^2}^2 v(t_m)(\Delta t)^2 + \frac{1}{2}\int_{t_m}^{t_{m+1}} \partial_{t^3}^3 v(s)(t_{m+1} - s)^2 ds \right)$$

$$-\frac{1}{2}\partial_t v(t_{m+1}) - \frac{1}{2}\partial_t v(t_m)$$

$$= \frac{1}{2}\partial_{t^2}^2 v(t_m)\Delta t + \frac{1}{2\Delta t}\int_{t_m}^{t_{m+1}} \partial_{t^3}^3 v(s)(t_{m+1} - s)^2 ds$$

$$-\frac{1}{2}\left(\partial_{t^2}^2 v(t_m)\Delta t + \int_{t_m}^{t_{m+1}} \partial_{t^3}^3 v(s)(t_{m+1} - s)ds \right)$$

$$= \int_{t_m}^{t_{m+1}} \partial_{t^3}^3 v(s)(t_{m+1} - s)\left[-\frac{1}{2} + \frac{1}{2\Delta t}(t_{m+1} - s) \right] ds$$

$$= \frac{1}{2\Delta t}\int_{t_m}^{t_{m+1}}(t_{m+1} - s)(t_n - s)\partial_{t^3}^3 v(s)ds.$$

Then we have $||\eta(t_m)||_{\mathcal{H}} \le C\Delta t \int_{t_m}^{t_{m+1}} \left|\left| \partial_{t^3}^3 v(s) \right|\right|_{\mathcal{H}} ds$ where C is some universal constant. To sum up we have for any $0 \le m \le M-1$

$$||\varepsilon^m||_{\mathcal{V}'} \le \frac{1}{\Delta t}\int_{t_m}^{t_{m+1}} ||(I - \Pi)\partial_t v(s)||_{\mathcal{H}} \, ds + C\Delta t \int_{t_m}^{t_{m+1}} \left|\left| \partial_{t^3}^3 v(s) \right|\right|_{\mathcal{H}} ds$$

and, using Cauchy-Schwarz inequality,

$$||\varepsilon^m||_{\mathcal{V}'} \leq \frac{1}{\sqrt{\Delta t}} \left(\int_{t_m}^{t_{m+1}} ||(I - \Pi)\partial_t v(s)||_{\mathcal{H}}^2 \, ds \right)^{\frac{1}{2}} + C(\Delta t)^{\frac{3}{2}} \left(\int_{t_m}^{t_{m+1}} \left\| \partial_{t^3}^3 v(s) \right\|_{\mathcal{H}}^2 \, ds \right)^{\frac{1}{2}}.$$

(49)

Convergence. Let $0 \leq m \leq M$. Using now $||v^{m,N,P} - v(t_m)||_{\mathcal{H}} \leq ||e^m||_{\mathcal{H}} + ||v(t_m) - \Pi v(t_m)||_{\mathcal{H}}$ and Eqs. 46 and 49 we get

$$\begin{aligned}
||v^{m,N,P} - v(t_m)||_{\mathcal{H}} &\leq ||v(t_m) - \Pi v(t_m)||_{\mathcal{H}} + \Bigg\{ ||f^{N,P} - \Pi f||_{\mathcal{H}}^2 \\
&\quad + \frac{1}{c} \int_0^{t_m} ||(I - \Pi)\partial_t v(s)||_{\mathcal{H}}^2 \, ds + \frac{C}{c}(\Delta t)^4 \int_0^{t_m} \left\| \partial_{t^3}^3 v(s) \right\|_{\mathcal{H}}^2 \, ds \\
&\quad + \frac{C}{c}(\Delta t)^2 \sum_{k=1}^{m-1}\sum_{j=1}^{m-1} \left(\int_{t_k}^{t_{k+1}} ||(I - \Pi)\partial_t v(s)||_{\mathcal{H}}^2 \, ds \times \int_{t_j}^{t_{j+1}} \left\| \partial_{t^3}^3 v(s) \right\|_{\mathcal{H}}^2 \, ds \right)^{\frac{1}{2}} \Bigg\}^{\frac{1}{2}} \\
&\leq \sup_{s\in[0,T]} ||(I - \Pi)v(s)||_{\mathcal{H}} + \Bigg\{ \left(||f^{N,P} - f||_{\mathcal{H}} + ||f - \Pi f||_{\mathcal{H}} \right)^2 \\
&\quad + \frac{T}{c} \sup_{s\in[0,T]} ||(I - \Pi)\partial_t v(s)||_{\mathcal{H}}^2 + \frac{C}{c} ||\partial_{t^3}^3 v||_{L^2(0,T;\mathcal{H})}^2 (\Delta t)^4 \\
&\quad + \frac{C}{c} T^2 \sup_{s\in[0,T]} ||\partial_{t^3}^3 v(s)||_{\mathcal{H}} \Delta t \sup_{s\in[0,T]} ||(I - \Pi)\partial_t v(s)||_{\mathcal{H}} \Bigg\}^{\frac{1}{2}}
\end{aligned}$$

(50)

Recall now that we have $||f^{N,P} - f||_{\mathcal{H}} \to 0$ as $N, P \to \infty$ and Eq. 43. If $y \in C(0,T;\mathcal{V})$ the family $(I - \Pi)y(s)$, $s \in [0,T]$, is equicontinuous, and we know thanks to Eq. 43 that for any $s \in [0,T]$ we have $||(I - \Pi)y(s)||_{\mathcal{H}} \to 0$ as $N, P \to \infty$. With the help of Ascoli theorem we can see that we have then

$$\sup_{s\in[0,T]} ||(I - \Pi)y(s)||_{\mathcal{H}} \xrightarrow[N\to\infty, P\to\infty]{} 0.$$

Using this with $y = v$, $\partial_t v$, in Eq. 50 we get the announced convergence (note that it is in order $(\Delta t)^2$ is time; cf Remark 8). $\qquad \square$

References

Brezis H (1983) Analyse fonctionnelle. Collection Mathématiques Appliquées pour la Maîtrise, Masson, Paris, théorie et applications

Cameron RH, Martin WT (1947) The orthogonal development of non-linear functionals in series of fourier-hermite functionals. Ann Math 48(2):385–392

Cooper FC (2017) Optimisation of an idealised primitive equation ocean model using stochastic parameterization. Ocean Model 113:187–200

Dautray R, Lions J (1993) Mathematical analysis and numerical methods for science and technology. No. vol. 6 in Mathematical Analysis and Numerical Methods for Science and Technology. Springer

Ern A, Guermond J (2002) Éléments finis: théorie, applications, mise en oeuvre. Mathématiques et Applications. Springer, Berlin

Friedman A (1975) Stochastic differential equations and applications. No. vol. 1 in Probability and mathematical statistics. Academic Press, New York

Ghanem RG, Spanos PD (1991) Stochastic finite elements: a spectral approach. Springer, Berlin

Gilbarg D, Trudinger NS (1983) Elliptic partial differential equations of second order. Grundlehren der mathematischen Wissenschaften. Springer, Berlin

Gillespie DT, Petzold LR (2003) Improved leap-size selection for accelerated stochastic simulation. J Chem Phys 119(16):8229–8234

Gobet E (2016) Monte-Carlo methods and stochastic processes: from linear to non-linear. Chapman & Hall/CRC

Hart JL, Alexanderian A, Gremaud PA (2017) Efficient computation of sobol'indices for stochastic models. SIAM J Sci Comput 39(4):A1514–A1530

Iooss B, Ribatet M (2009) Global sensitivity analysis of computer models with functional inputs. Reliab Eng Syst Saf 94(7):1194–1204

Janon A, Klein T, Lagnoux A, Nodet M, Prieur C (2014) Asymptotic normality and efficiency of two sobol index estimators. ESAIM Probab Statist 18:342–364

Jimenez MN, Le Maître O, Knio O (2017) Nonintrusive polynomial chaos expansions for sensitivity analysis in stochastic differential equations. SIAM/ASA J Uncertain Quantif 5(1):378–402

Karatzas I, Shreve S (1991) Brownian motion and stochastic calculus, 2nd edn. Graduate Texts in Mathematics, 113. Springer, New York. xxiii

Le Maître O, Knio OM (2010) Spectral methods for uncertainty quantification: with applications to computational fluid dynamics. Springer Science & Business Media

Le Maître O, Knio OM (2015) Pc analysis of stochastic differential equations driven by wiener noise. Reliab Eng Syst Saf 135:107–124

Lieberman G (1996) Second order parabolic differential equations. World Scientific, Singapore

Lions JL, Magenes E (1972) Non-homogeneous boundary value problems and applications, vol I. Springer, New York-Heidelberg. translated from the French by P. Kenneth, Die Grundlehren der mathematischen Wissenschaften, Band 181

Marrel A, Iooss B, Da Veiga S, Ribatet M (2012) Global sensitivity analysis of stochastic computer models with joint metamodels. Stat Comput 22(3):833–847

Monod H, Naud C, Makowski D (2006) Uncertainty and sensitivity analysis for crop models. Working with dynamic crop models: evaluation, analysis, parameterization, and applications, vol 4, pp 55–100

Nouy A (2009) Recent developments in spectral stochastic methods for the numerical solution of stochastic partial differential equations. Arch Comput Methods Eng 16(3):251–285

Nouy A (2017) Low-rank methods for high-dimensional approximation and model order reduction. Model Reduction and Approximation: Theory and Algorithms 15:171

Raviart P, Thomas J (2004) Introduction à l'analyse numérique des équations aux dérivées partielles. Mathématiques appliquées pour la maîtrise, Dunod

Saltelli A, Chan K, Scott EM et al (2000) Sensitivity analysis, vol 1. Wiley, New York

Sobol IM (1993) Sensitivity estimates for nonlinear mathematical models. Math Modeling Comput Experiment 1(4):407–414

Sobol IM (2001) Global sensitivity indices for nonlinear mathematical models and their monte carlo estimates. Math Comput Simul 55(1):271–280

van der Vaart AW (1998) Asymptotic statistics, Cambridge Series in Statistical and Probabilistic Mathematics, vol 3. Cambridge University Press, Cambridge

Wiener N (1938) The homogeneous chaos. Am J Math 60(4):897–936