

Non-Intrusive Polynomial Chaos Expansions for Sensitivity Analysis in Stochastic Differential Equations*

M. Navarro Jimenez[†], O.P. Le Maître[‡], and O.M. Knio[§]

Abstract. A Galerkin polynomial chaos (PC) method was recently proposed to perform variance decomposition and sensitivity analysis in stochastic differential equations (SDEs), driven by Wiener noise and involving uncertain parameters. The present paper extends the PC method to non-intrusive approaches enabling its application to more complex systems hardly amenable to stochastic Galerkin projection methods. We also discuss parallel implementations and the variance decomposition of derived quantity of interest within the framework of non-intrusive approaches. In particular, a novel hybrid PC sampling based strategy is proposed in the case of non-smooth quantities of interest (QoIs) but smooth SDE solution. Numerical examples are provided that illustrate the decomposition of the variance of QoIs into contributions arising from the uncertain parameters, the inherent stochastic forcing, and joint effects. The simulations are also used to support a brief analysis of the computational complexity of the method, providing insight on type of problems that would benefit from the present developments.

Key words. Variance decomposition, Stochastic differential equation, Polynomial Chaos

AMS subject classifications. 60H10, 65C30, 65D15, 65D32

1. Introduction. Stochastic differential equations are used to model complex systems and phenomena in many different domains, such as biology [36], epidemiology [3], economics and finance [2, 26], weather prediction [15], and building energy consumption [4]. In addition to their inherent random dynamics, these stochastic models also involve parameters that are often uncertain. As a result, the output of the random model is uncertain and depends on both the inherent variability caused by the driving noise and the uncertain parameters.

The simplest approach type to analyze such stochastic systems, with or without uncertain parameters, is the Monte Carlo method [1, 23]. Typically, MC methods involve a random sampling step, to generate a sample set of model predictions, followed by a statistical analyses to characterize the model output variability. The characterizations can include, among others, the estimation of the expected values and statistical moments of QoIs, as well as variance decomposition to rank the importance of different uncertain model parameters [43, 10]. Although the implementation of MC methods is usually straightforward, it provides random estimates with usually limited accuracy. Specifically, the order of convergence of MC estimates is $\mathcal{O}(N^{1/2})$, where N is the number of samples. Therefore, MC methods can be computationally expensive if accurate estimates are required.

To decrease the computational effort of MC methods, several improvements of the sampling procedure have been introduced. These include Latin Hypercube Sampling (LHS) [24, 28], the Quasi-Monte Carlo (QMC) method [5, 12], the Markov Chain Monte Carlo method

*Support from the SRI UQ center of the King Abdullah University of Science and Technology is acknowledged.

[†]CEMSE Division, King Abdullah University of Science and Technology (KAUST), Thuwal, Saudi Arabia.

[‡]LIMSI-CNRS, UPR-3251, Orsay, France.

[§]Department of Mechanical Engineering and Materials Science, Duke University, Durham, NC, USA.

(MCMC) [25] and probability density evolution methods [7, 47]. None of these improved MC methods exploits the possible smoothness in the model output with respect to the random inputs to accelerate the estimation. The use of functional approximations in the context of inherently stochastic systems with uncertain parameters was introduced more recently, in particular using Polynomial Chaos (PC) expansions. Classically, the functional approximations exploit some regularity with respect to the random parameters to estimate sensitivity coefficients. Since smoothness with respect to inherent noise is generally not to be expected, functional approximations have been mostly limited to the investigation of the effects of uncertain parameters on the model output statistics [32, 33, 34, 35, 39, 40]. Typically, the dependence with the uncertain parameters of the two first moments of the stochastic model output are characterized by means of functional approximations.

In a previous work [21], a new methodology was introduced to simulate and to analyze SDEs in the presence of uncertain inputs. Under the assumption that the Wiener noise and uncertain parameters could be treated as independent random variables, a PC [6, 20, 48] representation of the stochastic process was performed. The resulting spectral representation was exploited to perform an orthogonal decomposition of the variance of the process. As a result, contributions arising from the uncertainty in parameters, the stochastic forcing, and a couple term were identified. This methodology was focused on *intrusive* or Galerkin methods to compute the stochastic modes. However, there are complex settings, industrial applications and situations where only a black-box code is available. Therefore the implementation of Galerkin is not always feasible. For this reason, in this work, we propose an extension of the methodology in [21] to *non-intrusive* or sampling based methods, and to develop efficient techniques that enable a global sensitivity analysis of the stochastic process.

The implementation of non-intrusive methods to analyze and simulate stochastic differential equations driven by Wiener noise is not a straightforward problem, since the robustness of the estimations could be lost [39]. In particular, we rely on non-intrusive pseudo-spectral projection method, and show that the implementation retains all the attractive features of non-intrusive methods when the problem is smooth. We also exploit the fact that non-intrusive methods enable us to focus on specific QoIs. For both smooth and non-smooth QoIs, we present techniques for performing an orthogonal decomposition of their variance.

The outline of this paper is as follows. In Section 2, we fix the notation and briefly discuss the background concepts used throughout the paper. We introduce the non-intrusive projection method for the case of stochastic differential equations driven by Wiener noise with parametric uncertainty in Section 2.2, detail its application to the computation of the sensitivity indices in Section 2.3, and its implementation in Section 2.4. Then, in Section 3, we carry out a validation of the NISP method with respect to Galerkin to show that all the attractive features of the non-intrusive approach still hold in the presence of irreducible noise when the problem is smooth. We further illustrate in Section 3.2 the application of the approach to the case of smooth QoIs, amenable to direct non-intrusive projection. In section 4, we describe how to perform indirectly the global sensitivity analysis of QoIs being non-smooth functional of the smooth SDE solution, and discuss the potential improvement brought by this indirect approach. Concluding remarks are provided in Section 5.

2. Non-intrusive pseudo-spectral projection for SDEs with parametric uncertainty. In this section, we first introduce the uncertain SDE problem and briefly introduce the background materials used throughout the paper. The non-intrusive projection method and the variance decomposition approach used to characterize the variability of the model output are respectively discussed in Sections 2.1, 2.2 and 2.3.

Let (Ω, Σ, P) be a probability space where Ω is the event space, Σ denotes a σ -algebra on Ω , and P is a probability measure. In this space, we consider the following SDE

$$dX(t, \omega) = C(X(t, \omega))dt + D(X(t, \omega))dW(t, \omega), \quad X(t = 0, \omega) = X_{(0)}, \quad (2.1)$$

where $X : (t, \omega) \in T \times \Omega \mapsto \mathbb{R}$ is a real-valued stochastic process, defined in the time interval $T = [0, T_f]$ with $T_f > 0$, and $W(t, \omega)$ is a Wiener process. The function $C : \mathbb{R} \mapsto \mathbb{R}$ is the *drift* coefficient and $D : \mathbb{R} \mapsto \mathbb{R}$ is the *diffusion* coefficient. The Wiener process is the source of their inherent stochasticity, inducing variability (randomness) in the the solution X .

In the present paper, we also consider uncertain drift and diffusion coefficients through the introduction of random parameters $Q(\omega)$ in the definitions of these two functions. The SDE (2.1) is consequently rewritten as

$$dX(t, \omega) = C(X(t, \omega), Q(\omega))dt + D(X(t, \omega), Q(\omega))dW(t, \omega), \quad X(t = 0, Q(\omega)) = X_{(0)}(Q(\omega)). \quad (2.2)$$

The SDE solution X has now two sources of randomness: the inherent noise induced by Wiener process and the uncertain parameters Q . Our objective is to perform a global sensitivity analysis to quantify the respective impact of these two sources of randomness on the overall variability.

2.1. Polynomial chaos expansion. In the following, it is assumed that the random parameters Q are parametrized using a finite-dimensional random germ, $\boldsymbol{\xi} = \{\xi_1, \dots, \xi_N\}$, whose components are independent, canonical, real-valued independent random variables. Hence $Q(\omega) = Q(\boldsymbol{\xi}(\omega))$. Let $\Xi \subseteq \mathbb{R}^N$ denote the range of the random vector $\boldsymbol{\xi}(\omega)$, and p_i the probability density function (pdf) of ξ_i , for $i = 1, \dots, N$. By virtue of the independence of the ξ_i 's, the joint pdf of the random vector $\boldsymbol{\xi}$ is

$$p_{\boldsymbol{\xi}}(\xi_1, \dots, \xi_N) = \prod_{i=1}^N p_i(\xi_i). \quad (2.3)$$

Let $L_2(\Xi, p_{\boldsymbol{\xi}})$ be the space of square-integrable functionals in $\boldsymbol{\xi}$, endowed with the following inner product and associated norm

$$\forall U, V \in L_2(\Xi, p_{\boldsymbol{\xi}}), \quad \langle U, V \rangle = \int_{\Xi} U(\boldsymbol{\xi})V(\boldsymbol{\xi})p_{\boldsymbol{\xi}}(\boldsymbol{\xi})d\boldsymbol{\xi}, \quad \|U\|_{L_2} = \langle U, U \rangle^{1/2}. \quad (2.4)$$

Any functional $U \in L_2(\Xi, p_{\boldsymbol{\xi}})$ has a PC expansion of the form [6]:

$$U(\boldsymbol{\xi}) = \sum_{\boldsymbol{\alpha} \in \mathbb{N}_0^N} [U_{\boldsymbol{\alpha}}] \Psi_{\boldsymbol{\alpha}}(\boldsymbol{\xi}), \quad (2.5)$$

where the coefficients $[U_\alpha]$ are called modes or PC coefficients of U , $\alpha = (\alpha_1, \dots, \alpha_N)$ is a multi-index, and the $\Psi_\alpha(\boldsymbol{\xi})$ are multivariate orthonormal polynomials. The polynomials are mutually orthogonal with respect to the inner product:

$$\langle \Psi_\alpha, \Psi_\beta \rangle = \delta_{\alpha\beta} = \begin{cases} 1 & \alpha = \beta, \\ 0 & \alpha \neq \beta. \end{cases} \quad (2.6)$$

Classically, the polynomial Ψ_α is the product of one-dimensional polynomials defined through

$$\Psi_\alpha(\boldsymbol{\xi}) \doteq \prod_{i=1}^N \psi_{\alpha_i}^i(\xi_i), \quad (2.7)$$

where $\{\psi_\alpha^i, \alpha \in \mathbb{N}_0\}$ is a complete orthonormal set (with respect to the density p_i), and ψ_α^i is a polynomial of degree α . In practice, the expansion of U in (2.5) is truncated to a finite set \mathcal{A} of multi-indices,

$$U(\boldsymbol{\xi}) \approx \hat{U}(\boldsymbol{\xi}) \doteq \sum_{\alpha \in \mathcal{A}} [U_\alpha] \Psi_\alpha(\boldsymbol{\xi}). \quad (2.8)$$

We shall denote $P = |\mathcal{A}| \doteq \text{card}(\mathcal{A})$, and call P the dimension of the PC basis.

2.2. Non-intrusive method. Consider $U \in L_2(\Xi, p_\xi)$. We seek to approximate U by a function in the subspace $\mathbb{S}_\mathcal{A}$ spanned by the polynomial basis $\{\Psi_\alpha, \alpha \in \mathcal{A}\}$. For a non-intrusive method, the expansion coefficients $[U_\alpha]$ in (2.8) are determined from a set of evaluations of $U(\boldsymbol{\xi})$ at selected points $\boldsymbol{\xi}_j \in \Xi$. The sample set of evaluation points $\boldsymbol{\xi}_j$, as well as the expansion coefficients, can be determined in different ways. These include the so-called collocation methods [27, 30], which consist in multivariate interpolations, and quadrature-based methods [22], which roughly consist in approximating the L_2 -inner product by discrete quadrature rules. In the present work, we rely on non-intrusive projections which are quadrature based approaches, but the developments proposed below can be easily extended to collocation methods.

In the non-intrusive Spectral Projection (NISP) method [22, 31] one defines \hat{U} as the orthogonal projection of U in $\mathbb{S}_\mathcal{A}$, leading to the P conditions

$$\left\langle U - \sum_{\beta \in \mathcal{A}} [U_\beta] \Psi_\beta, \Psi_\alpha \right\rangle = 0, \quad \forall \alpha \in \mathcal{A}. \quad (2.9)$$

Equivalently, using the orthogonality of the polynomials, these conditions can be expressed as

$$[U_\alpha] = \langle U, \Psi_\alpha \rangle, \quad \forall \alpha \in \mathcal{A}. \quad (2.10)$$

Upon the introduction of a suitable quadrature rule to approximate the inner product, with points $\boldsymbol{\xi}_j$ and weights w_j , one obtains

$$[U_\alpha] = \langle U, \Psi_\alpha \rangle \approx \sum_{j=1}^{N_Q} w_j U(\boldsymbol{\xi}_j) \Psi_\alpha(\boldsymbol{\xi}_j), \quad (2.11)$$

where N_Q is the number of quadrature points. Denoting $\mathbf{U} = (U(\boldsymbol{\xi}_j))_{j=1,\dots,N_Q}$ the vector containing the values of U at the quadrature nodes, $[\mathbf{U}] = ([U_\alpha])_{\alpha \in \mathcal{A}}$ the vector of PC coefficients of U , the NISP method can be recast as

$$[\mathbf{U}] = \mathcal{P}^{\text{NISP}} \mathbf{U}, \quad \mathcal{P}^{\text{NISP}} \in \mathbb{R}^{(P+1) \times N_Q}. \quad (2.12)$$

In (2.12) the NISP operator $\mathcal{P}^{\text{NISP}}$ has for entries

$$\mathcal{P}_{\alpha j}^{\text{NISP}} = w_j \Psi_\alpha(\boldsymbol{\xi}_j), \quad \alpha \in \mathcal{A}, \quad 1 \leq j \leq N_Q.$$

A minimal requirement for $\mathcal{P}^{\text{NISP}}$ is that it is exact for any $U \in \mathbb{S}_{\mathcal{A}}$. This requirement translates in the following conditions on the quadrature rule underlying the NISP method:

$$\sum_{j=1}^{N_Q} w_j \Psi_\alpha(\boldsymbol{\xi}_j) \Psi_{\alpha'}(\boldsymbol{\xi}_j) = \delta_{\alpha\alpha'}, \quad \forall \alpha, \alpha' \in \mathcal{A}. \quad (2.13)$$

For such a quadrature rule and polynomial basis, $\mathcal{P}^{\text{NISP}}$ is free of internal aliasing and the projection error due to the substitution of the continuous inner product with the quadrature rule reduces to external aliasing (that is aliasing error with the component of the function orthogonal to $\mathbb{S}_{\mathcal{A}}$). In practice, the quadrature rule is first constructed and \mathcal{A} is subsequently defined as the largest set of polynomial multi-indices satisfying (2.13). Operating this way provides a great flexibility in the construction of the quadrature rule, possibly withing an adaptive framework to come-up with anisotropic rules (and polynomial bases). Sparse grid methods [13, 16], based on generalized Smolyak formula [41], are in particular crucial to enable the application of NISP methods in high dimensions.

In the present work we rely on an improved variant of the sparse grid projection, called the Pseudo-Spectral Projection (PSP), that was proposed in [8, 9]. The PSP method uses sparse tensorization of one-dimensional projection operators at different levels to construct a sparse (multi-dimensional) projection operator. The crucial difference between the PSP and SP methods, is that the projection at different levels uses different quadrature rules *and* projection basis. Doing so, PSP achieve exactness for generally larger multi-index sets \mathcal{A} than for the NISP method, with higher order polynomial spaces for the same set of points. This results in a lower projection error for similar computational complexity. Further, the PSP can be written as in (2.12), but with an alternative operator \mathcal{P}^{PSP} .

In the following we shall denote \mathcal{P}^{NI} a generic non intrusive projection operator, without referring explicitly to the actual non-intrusive method considered. The linear operator \mathcal{P}^{NI} then maps the model solutions at the N_Q grid points to the P coefficients of the PC approximation.

We now return to the determination of the expansion coefficients of $X(t, \omega)$, solution of (2.2). Note that in the remainder of the section, by a slight abuse of notation, we will write $X(t, W(\omega), \boldsymbol{\xi}(\omega))$, to explicitly highlight the dependence of X on the Wiener noise and the uncertain parameters. As mentioned previously, it was proposed in [21] to use PC expansions to account for parametric dependence of X . Along this line, the expansion coefficients are random processes (functions of W) and the approximation is expressed as

$$X(t, W, \boldsymbol{\xi}) \approx \sum_{\alpha \in \mathcal{A}} [X_\alpha](t, W) \Psi_\alpha(\boldsymbol{\xi}). \quad (2.14)$$

For the non-intrusive approach, the stochastic processes $[X_\alpha](t, W)$ for $\alpha \in \mathcal{A}$ are in fact determined by applying the discrete projection operator \mathcal{P}^{NI} to a set of N_Q solutions of the SDE corresponding to the parameter values $Q(\xi_j)$:

$$[X_\alpha](t, W) = \sum_{j=1}^{N_Q} \mathcal{P}_{\alpha j}^{\text{NI}} X(t, W, \xi_j). \quad (2.15)$$

In other words, to perform the NI projection, one has to solve a SDE at each of the N_Q sparse grid points and for the *same* Wiener process W . To this end, let us denote $\mathbf{Y}(t, W)$ the random solution of the *system* of N_Q stochastic differential equations

$$d\mathbf{Y} = \mathbf{C}(\mathbf{Y})dt + \mathbf{D}(\mathbf{Y})dW, \quad (2.16)$$

where the drift and diffusion vectors are defined by

$$\mathbf{C}(\mathbf{Y}) \doteq (C(Y_1, Q(\xi_1)) \dots C(Y_{N_Q}, Q(\xi_{N_Q})))^T, \quad \mathbf{D}(\mathbf{Y}) \doteq (D(Y_1, Q(\xi_1)) \dots D(Y_{N_Q}, Q(\xi_{N_Q})))^T.$$

In words, the j -th component of the drift and diffusion vectors are evaluated at the parameter value $Q(\xi_j)$ and the corresponding solution point of the sparse grid (Y_j). The stochastic system (2.16) is completed with the initial conditions,

$$(\mathbf{Y}(t=0))_j = X_{(0)}(\xi_j).$$

One key property of the system above is that the resolution of the components Y_j are in fact independent from one to another. This will enable efficient parallel procedure as discussed in the next section.

The resolution of the stochastic system then proceeds classically by means of Monte Carlo method to sample the Wiener process. Let us denote $W^{(i)} = W(\omega_i)$ a particular trajectory of the Wiener process and $\mathbf{Y}^{(i)}(t)$ the corresponding trajectory of the random vector $\mathbf{Y}(t, W^{(i)})$, solving (2.16). The expansion coefficients of the corresponding trajectory of $X(t)$ are finally obtained through

$$[X_\alpha]^{(i)}(t) = \sum_{j=1}^{N_Q} \mathcal{P}_{\alpha j}^{\text{NI}} Y_j^{(i)}(t). \quad (2.17)$$

Obviously, the resolution of (2.16) conditioned on $W = W^{(i)}$ requires the introduction of an appropriate time discretization; in the present work we use for simplicity the classical Euler-Maruyana method. Note that any more advanced integration method could be considered for the generation of the trajectories $\mathbf{Y}^{(i)}(t)$.

2.3. Variance decomposition. From the approximation of $X(t, W, \xi)$ in (2.14), the objective is to apply a global sensitivity analysis, in particular to quantify the respective importance of the noise W and parameters $Q(\xi)$ on the variability of X at some time T . For notational convenience, the reference to time is dropped in the remainder of this section, and we detail only the expressions used in the analysis. A complete derivation of these expression can be found in [21]. A review of classical sensitivity analysis can be found in [38].

The variance-based global sensitivity analysis of X relies on its Sobol-Hoeffding (SH) decomposition [42]:

$$X = \bar{X} + X_{\text{par}}(\boldsymbol{\xi}) + X_{\text{noise}}(W) + X_{\text{mix}}(\boldsymbol{\xi}, W), \quad (2.18)$$

where $\bar{X} = \mathbb{E}\{X\}$ and

$$\begin{aligned} X_{\text{par}}(\boldsymbol{\xi}) &= \mathbb{E}\{X|\boldsymbol{\xi}\} - \bar{X}, \\ X_{\text{noise}}(W) &= \mathbb{E}\{X|W\} - \bar{X}, \\ X_{\text{mix}}(\boldsymbol{\xi}, W) &= X - \bar{X} - \mathbb{E}\{X|\boldsymbol{\xi}\} - \mathbb{E}\{X|W\}. \end{aligned}$$

It can be easily shown that the SH functions appearing on the right-hand-side of (2.18) are mutually orthogonal. As such, the variance of X can be decomposed into the sum of the three contributions,

$$\mathbb{V}\{X\} = V_{\text{par}} + V_{\text{noise}} + V_{\text{mix}}, \quad (2.19)$$

which are respectively identified as the (partial) variances due to the parametric uncertainty (V_{par}), to the Wiener noise (V_{noise}) and their interactions (V_{mix}). Classically, the importance of the different contributions is characterized using the so-called sensitivity indices, which are simply the corresponding partial variances normalized by the total variance of X :

$$S_{\text{par}} = \frac{V_{\text{par}}}{\mathbb{V}\{X\}}, \quad S_{\text{noise}} = \frac{V_{\text{noise}}}{\mathbb{V}\{X\}}, \quad S_{\text{mix}} = \frac{V_{\text{mix}}}{\mathbb{V}\{X\}}. \quad (2.20)$$

Having introduced previously the PC approximation of X , we now want to reuse this approximation to obtain approximations of the sensitivity indices. We shall adopt here the classical convention, defining $\Psi_{\mathbf{0}}(\boldsymbol{\xi}) \doteq 1$ for some $\mathbf{0} \in \mathcal{A}$, such that all other polynomials have vanishing expectation (by the orthogonality of the $\Psi_{\boldsymbol{\alpha}}$):

$$\mathbb{E}\{\Psi_{\boldsymbol{\alpha}}\} = \begin{cases} 1 & \boldsymbol{\alpha} = \mathbf{0}, \\ 0 & \boldsymbol{\alpha} \neq \mathbf{0}. \end{cases} \quad (2.21)$$

We first restate a major assumption of this work, which is that the Wiener noise and the SDE parameters are two independent random quantities. Therefore, the solution mean $\bar{X} \doteq \mathbb{E}\{X\}$ is approximated from its PC via

$$\bar{X} \approx \sum_{\boldsymbol{\alpha} \in \mathcal{A}} \mathbb{E}\{[X_{\boldsymbol{\alpha}}](W)\} \mathbb{E}\{\Psi_{\boldsymbol{\alpha}}\} = \mathbb{E}\{[X_{\mathbf{0}}](W)\}.$$

Similarly the variance of X is approximated by

$$\mathbb{V}\{X\} = \mathbb{E}\{X^2\} - \bar{X}^2 \approx \sum_{\boldsymbol{\alpha} \in \mathcal{A}} \mathbb{E}\{[X_{\boldsymbol{\alpha}}]^2\} - \mathbb{E}\{[X_{\mathbf{0}}]\}^2. \quad (2.22)$$

PC-based expressions can also be obtained for the parametric conditional expectations and variances. Specifically, the conditional expectation and variance of X given $\boldsymbol{\xi} = \boldsymbol{\eta}$ are expressed as:

$$\begin{aligned} \mathbb{E}\{X|\boldsymbol{\xi} = \boldsymbol{\eta}\} &\approx \mathbb{E}\left\{\sum_{\boldsymbol{\alpha} \in \mathcal{A}} [X_{\boldsymbol{\alpha}}]\Psi_{\boldsymbol{\alpha}}|\boldsymbol{\xi} = \boldsymbol{\eta}\right\} = \sum_{\boldsymbol{\alpha} \in \mathcal{A}} \mathbb{E}\{[X_{\boldsymbol{\alpha}}]\} \Psi_{\boldsymbol{\alpha}}(\boldsymbol{\eta}), \\ \mathbb{V}\{X|\boldsymbol{\xi} = \boldsymbol{\eta}\} &= \mathbb{E}\{(X - \mathbb{E}\{X|\boldsymbol{\xi} = \boldsymbol{\eta}\})^2|\boldsymbol{\xi} = \boldsymbol{\eta}\} \approx \sum_{\boldsymbol{\alpha}, \boldsymbol{\alpha}' \in \mathcal{A}} [C_{X,X}]_{\boldsymbol{\alpha}, \boldsymbol{\alpha}'} \Psi_{\boldsymbol{\alpha}}(\boldsymbol{\eta}) \Psi_{\boldsymbol{\alpha}'}(\boldsymbol{\eta}), \end{aligned}$$

where $[\mathbb{C}_{\mathbf{x},\mathbf{x}}] \in \mathbb{R}^{P \times P}$ is the covariance matrix of the expansion coefficients of X . Conditioning X on a particular trajectory of W , say $W^{(i)}$, the following approximations can be similarly derived:

$$\begin{aligned} \mathbb{E} \left\{ X | W = W^{(i)} \right\} &\approx \sum_{\alpha \in \mathcal{A}} \mathbb{E} \left\{ [X_{\alpha}] | W = W^{(i)} \right\} \mathbb{E} \{ \Psi_{\alpha} \} = [X_{\mathbf{0}}]^{(i)}, \\ \mathbb{V} \left\{ X | W = W^{(i)} \right\} &= \mathbb{E} \left\{ \left(X - \mathbb{E} \left\{ X | W = W^{(i)} \right\} \right)^2 | W = W^{(i)} \right\} \approx \sum_{\alpha \in \mathcal{A} \setminus \mathbf{0}} \left([X_{\alpha}]^{(i)} \right)^2. \end{aligned}$$

The approximated expressions of the partial variance V_{par} and V_{noise} can then be obtained using the PC coefficients. This leads to

$$V_{\text{par}} = \sum_{\alpha \in \mathcal{A} \setminus \mathbf{0}} \mathbb{E} \{ [X_{\alpha}] \}^2, \quad V_{\text{noise}} = \mathbb{V} \{ [X_{\mathbf{0}}] \}, \quad (2.23)$$

from which it can be deduced using (2.22) and (2.23) that:

$$V_{\text{mix}} = \sum_{\alpha \in \mathcal{A} \setminus \mathbf{0}} \mathbb{V} \{ [X_{\alpha}] \}. \quad (2.24)$$

2.4. Discussion and implementation. The developments above assume that the process X can actually be expanded in the form (2.14). This means that X belongs to the space of square integrable random variables in Q tensored with the space of squared integrable functionals of W , that is using standard notation $L_2(\Xi, p_{\xi}) \otimes L_2(\mathbb{R}^{\infty}, p_W)$. For instance, this requirement rules out the case of SDE solutions blowing up in time with finite probabilities. It is also remarked that the expansion in (2.14) approximates the dependence of the solution with respect to the finite set of uncertain parameters, with expansion coefficients still depending on the infinite dimensional noise. Their dependence on W is, however, never computed; instead the proposed approach (jointly) samples the expansion coefficients $[X_{\alpha}](t, W)$ for realizations $W^{(i)}$ of the noise. In fact, the approach is applicable as soon as the coefficients $[X_{\alpha}](W^{(i)})$ have second moment with respect to the sampled paths $W^{(i)}$, as detailed below.

Observe that the expressions in (2.23) and (2.24) involve only the expectations and variances of the expansion coefficients of X . Their computations involve averaging over the Wiener measure only, and can therefore be estimated by means of MC sampling. Specifically, the expectation and variance of the PC coefficients $[X_{\alpha}](W)$ can be estimated through the classical formulas:

$$\mathbb{E} \{ [X_{\alpha}] \} \approx \mathbb{E} \{ \widehat{[X_{\alpha}]} \} \doteq \frac{1}{N_W} \sum_{i=1}^{N_W} [X_{\alpha}]^{(i)}, \quad \mathbb{V} \{ [X_{\alpha}] \} = \frac{1}{N_W - 1} \sum_{i=1}^{N_W} \left([X_{\alpha}]^{(i)} \right)^2 - \mathbb{E} \{ \widehat{[X_{\alpha}]} \}^2, \quad (2.25)$$

where we have denoted N_W the number of MC samples $W^{(i)}$ of the Wiener process. Recalling that $[X]^{(i)}$ is given by (2.17), where $\mathbf{Y}^{(i)}$ is the solution of (2.16) using $W = W^{(i)}$. The resolution of (2.16) can be achieved by any solver suitable to the solution of stochastic differential Equations [17]. In the present work we used the Euler-Maruyama method, but

higher order methods (*e.g.* Milstein [29] or Runge-Kutta [44]) or a correction method (*e.g.* Wong-Zakai's [14, 46]) could be considered as well. In fact, future works should consider the problem of selecting a solution method for $\mathbf{Y}^{(i)}$ that induces an error consistent with the projection error on the PC basis. Assuming that the cost of solving the SDE for $\mathbf{Y}^{(i)}$ is independent of the realization $W^{(i)}$, for instance when using a fixed time-discretization, one can estimate the computational complexity of estimating the partial variances as $\mathcal{O}(N_W \times N_Q)$. This complexity estimate only assumes a computational cost independent of the parameter values for solving the SDE given $W = W^{(i)}$. However, one can further exploit the possibility of computing independently the N_Q components of $\mathbf{Y}^{(i)}$. In other words, the computation of the set of trajectories $\{Y_j^{(i)}(t); j = 1, \dots, N_Q; i = 1, \dots, N_W\}$ can be made in parallel as long as the trajectories $W^{(i)}$ of the Wiener process are consistently generated for different non-intrusive points ξ_j . This is in particular the case for pseudo-random sequences entirely determined by a seed and when using the same time-discretization for all the non-intrusive points.

In the present work, we have tested two parallel strategies. In the first strategy, different MC realizations of the PC coefficients of X were generated in parallel, solving (sequentially) system (2.16) and applying (2.17) to get the whole set of coefficients $[X_\alpha]^{(i)}$ at once. Note that the resolution of (2.16) could be performed in parallel as well, corresponding to a fine-grained parallelization. In the second strategy, the parallelization was implemented over the non-intrusive points, generating in parallel sample sets $\{Y_j^{(i)}, i = 1, \dots, N_W\}$. This was made possible by reusing consistently the same set of N_W seeds for the random number generator at each of the non-intrusive points. Note, again, that the problems for $Y_j^{(i)}$ could also be solved in parallel for different i (fine-grained parallelization). Both parallel implementations are completely equivalent, in the sense that they produce the same results. The second strategy is more demanding in terms of memory since it requires us to store in memory all the solutions $Y_j^{(i)}$ to subsequently compute the PC coefficients $[X_\alpha]^{(i)}$ and estimate their first two moments needed for the variance decomposition. On the contrary, the first parallelization on $W^{(i)}$ seems better suited because the estimators of the partial variances can be computed on the fly, without having to store $[X_\alpha]^{(i)}$. However, the second approach may be computationally more efficient depending on the hardware architecture and infrastructure available.

Finally, it is emphasized that the PC decomposition can conceptually be performed for any functional of X (provided it has W -almost surely finite second moments). This feature is further discussed and illustrated in the following sections.

3. Validation. In this section we present a validation of the proposed NI approach. The accuracy of the solution is also contrasted with the Galerkin solution constructed using the methodology proposed in [21].

3.1. Test problem. Let us consider the process governed by the following SDE

$$dX(W, \xi) = (Q_1(\xi) - X(W, \xi))dt + (\nu X(W, \xi) + 1)Q_2(\xi)dW, \quad (3.1)$$

with uncertain parameter Q and the initial condition $X(t = 0) = 0$ almost surely. Note that for $\nu = 0$, X is the Ornstein-Uhlenbeck (OU) process. We set in this section $\nu = 0.2$ so the problem has multiplicative noise. The random functions Q_1 and Q_2 are assumed

to be independent and uniformly distributed, according to $Q_1 \sim \mathcal{U}([0.95, 1.15])$ and $Q_2 \sim \mathcal{U}([0.02, 0.22])$. These two functions are then parametrized using two independent random variables, ξ_1 and ξ_2 , having uniform distributions on the unit interval. Then $\boldsymbol{\xi} = (\xi_1, \xi_2) \sim \mathcal{U}([0, 1])^2$.

To monitor the convergence of the approximation, we shall report the normalized L_2 error norm defined by

$$\varepsilon(t) = \frac{\|\hat{X}(t) - X(t)\|}{\|X(t)\|}, \quad (3.2)$$

where X and \hat{X} are the exact and PC approximations of the SDE solution, and $\|\cdot\|$ is the L2-norm with respect to W and $\boldsymbol{\xi}$ given by:

$$\|U(W, \boldsymbol{\xi})\|^2 = \mathbb{E} \{U^2\}^{1/2}. \quad (3.3)$$

We focus on the error at $t = 10$ and drop the reference to the time in the notation of the error. The error norm is evaluated by means of a MC sampling as follows. First we draw a trajectory $W^{(i)}$ and construct the PC approximation $\hat{X}^{(i)} = \sum_{\alpha \in \mathcal{A}} [X_\alpha]^{(i)} \Psi_\alpha(\boldsymbol{\xi})$ through the PSP method as described above. Then we draw at random $\boldsymbol{\xi}^{(i)}$ in the unit square and solve for $X^{(i)}$ the stochastic differential equations for the noise trajectory $W^{(i)}$ and drift and diffusion function given by $Q_1(\xi_1^{(i)})$ and $Q_2(\xi_2^{(i)})$. The error norm is finally estimated as

$$\varepsilon^2 \approx \frac{\sum_{i=1}^M |\hat{X}^{(i)}(\boldsymbol{\xi}^{(i)}) - X^{(i)}|^2}{\sum_{i=1}^M |X^{(i)}|^2}, \quad (3.4)$$

for sufficiently large M (in practice few millions of samples are used to estimate the error). In addition to the MC sampling, the resolution of the SDE for the sampled values of the parameters and the Wiener noise requires a time-discretization. To obtain consistent error estimates, all computations use the same time step $\Delta t = 0.01$ for the integration of the differential equations.

Below, we investigate the evolution of the error in the PSP approximation when using different levels of the sparse grid. The different levels correspond to quadratures consisting in (sparse) tensorizations of the one-dimensional F ej er rule. The resulting projection bases have maximal degree $(2^\ell - 1)$, where $\ell \geq 1$ denotes the PSP level. Table 3.1 reports the dependence on the PSP level ℓ of the number of sparse grid points N_Q and the dimension of the PC basis P in the computations presented below.

ℓ	1	2	3	4	5
N_Q	5	17	49	129	321
P	3	8	20	48	112

Table 3.1: Number of sparse grid points and dimension of the PC basis for the 2-dimensional sparse PSP based on F ej er’s rule.

The left plot of Figure 3.1 shows the dependence of the L_2 error defined in (3.4) on the PSP level. It is seen that the approximation error has an exponential decay with the PSP level

and reaches machine precision for $\ell \geq 4$. This fast convergence highlights the smoothness of the dependence of the trajectories of X with respect to the uncertain parameters appearing in the definition of the process. The behavior of the non-intrusive approach is further illustrated in the right plot of Figure 3.1 which quantifies the difference between the PSP and Galerkin solutions of (3.1). The Galerkin solution is computed using the method presented in [21]. Specifically, the right plot depicts the norm of the difference between the PSP and Galerkin solution, normalized by the norm of the exact solution. For each estimate, the Galerkin solutions use the same PSP basis (\mathcal{A}) as in the corresponding PSP solution. These plots indicate that the PSP and Galerkin methods yield comparable convergence rates and error levels.

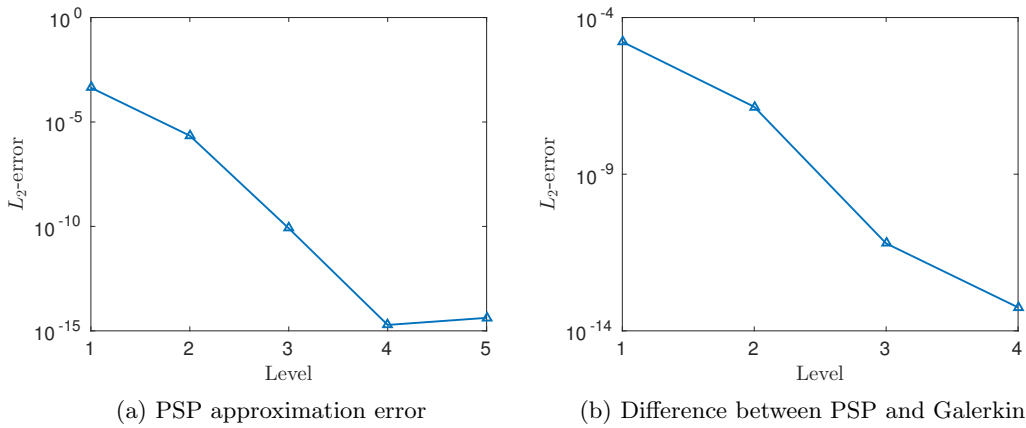


Figure 3.1: Left: error norm ϵ versus PSP level ℓ for the approximation of X governed by (3.1) at time $t = 10$. Right: norm of the difference between the Galerkin and PSP approximations at $t = 10$ versus the PSP level ℓ . The estimates are normalized by the norm of the Galerkin solution.

3.2. Direct NI projection of QoIs. Compared to the Galerkin method used in [21], the present NI approach has the advantage of allowing for the direct approximation of some functionals of X . As an example, assume that we are not interested in X , but in some derived quantity such as the integral of X over $t \in [0, 10]$. Figure 3.2a shows several trajectories of X , solution of (3.1), corresponding to different parameter values of $Q(\xi)$ but a single realization $W^{(i)}$ of W . Clearly, the integral

$$A(W, \xi) \doteq \int_0^{10} X(t, W, \xi) dt. \tag{3.5}$$

has variability due to W and ξ . (For fixed $W^{(i)}$, the dependence of A on ξ is illustrated in Figure 3.2b.) This integral can be estimated by a large variety of methods; for simplicity, we use here a simple trapezoidal rule with fixed time step for the SDE integration. Each trajectory of the noise, $W^{(i)}$, has an associated vector of SDE solutions at the sparse grid

points $\mathbf{Y}^{(i)}(t)$, from which one can *directly* obtain the PC coefficients of $A^{(i)}(\boldsymbol{\xi})$:

$$A(W^{(i)}, \boldsymbol{\xi}) \approx \hat{A}(W^{(i)}, \boldsymbol{\xi}) = \sum_{\boldsymbol{\alpha} \in \mathcal{A}} [A_{\boldsymbol{\alpha}}]^{(i)} \Psi_{\boldsymbol{\alpha}}(\boldsymbol{\xi}), \quad [A_{\boldsymbol{\alpha}}]^{(i)} = \sum_{j=1}^{N_Q} \mathcal{P}_{\boldsymbol{\alpha}j}^{\text{NI}} A_j^{(i)}, \quad A_j^{(i)} = \int_0^{10} Y_j^{(i)}(t) dt.$$

It is seen that this direct projection of A consists in first estimating the QoI at each sparse grid point $\boldsymbol{\xi}_j$, followed by a projection of these values. Such an approach is not possible in the Galerkin approach as no governing equation would be available for this QoI. This feature is one of the attractive aspects of NI methods.

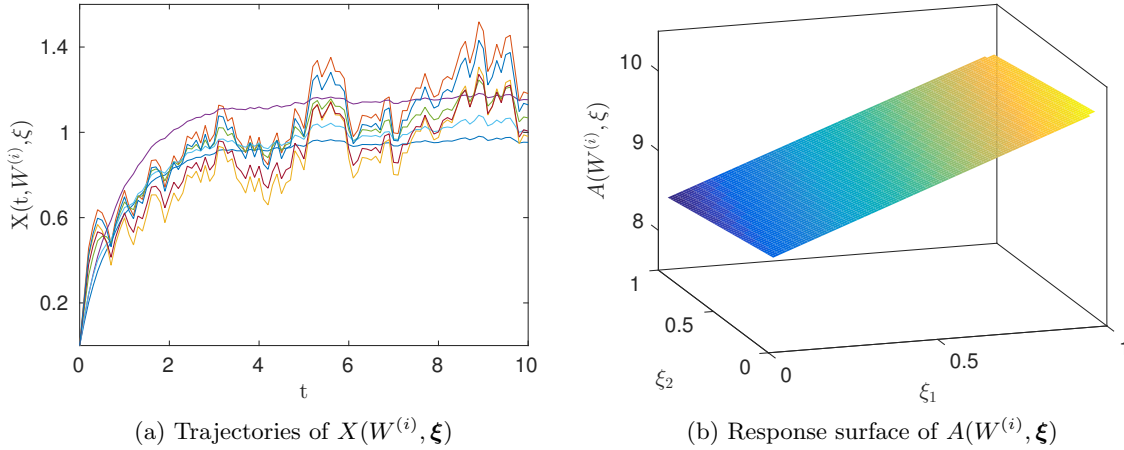
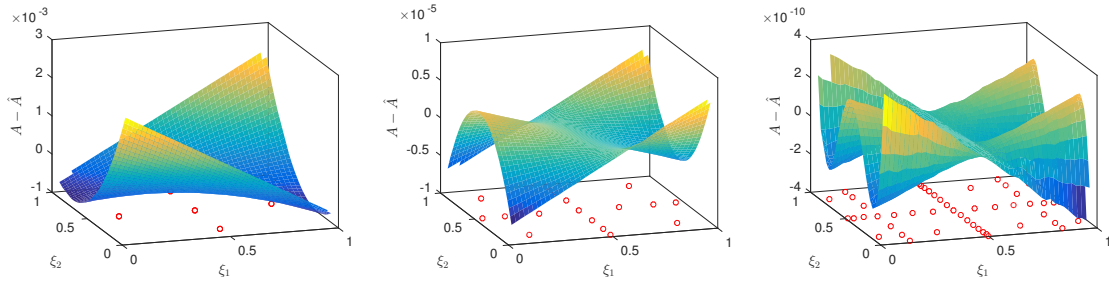


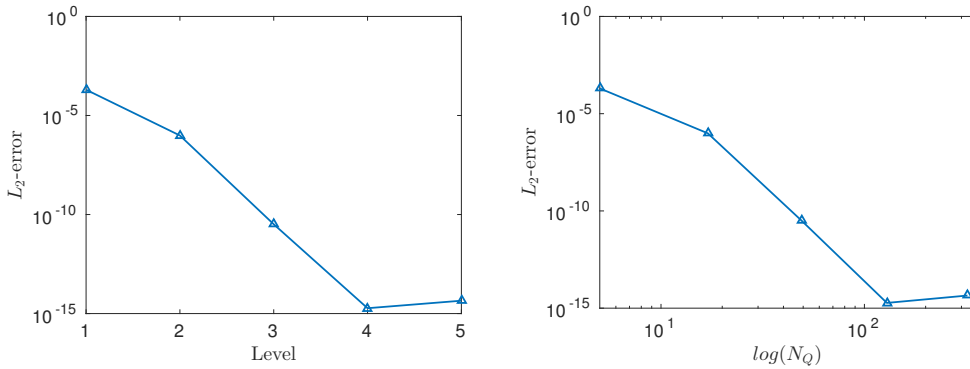
Figure 3.2: (a) Different trajectories of X , for a fixed noise $W^{(i)}$ and different realizations of $Q(\boldsymbol{\xi})$. (b) $A^{(i)}(\boldsymbol{\xi})$ versus ξ_1 and ξ_2 for fixed $W^{(i)}$.

We first check the convergence of the direct PSP projection of A , presenting in Figure 3.3a the difference $A(W^{(i)}, \boldsymbol{\xi}) - \hat{A}(W^{(i)}, \boldsymbol{\xi})$ as function of $\boldsymbol{\xi}$ for different PSP levels. Figure 3.3b shows the normalized approximation error $\|A - \hat{A}\|$, as expressed in (3.2) with A instead of X , as a function of PSP level (left plot) and number of sparse grid points (right plot). The results demonstrate the fast convergence of the PSP approximation, and indicate that even with $\ell = 1$ the PSP may be sufficiently well-converged for the purpose of performing the variance decomposition.

The sensitivity indices of A are then estimated from a sample set of N_W realizations of the PC coefficients $[Q]_{\boldsymbol{\alpha}}^{(i)}$ (see Sections 2.3 and 2.4). Figure 3.4 illustrates the dependence of S_{par} , S_{noise} and S_{mix} on N_W . For validation and comparison purposes, also shown are estimates obtained using a pure MC sampling strategy [37], based on sampling jointly W and $\boldsymbol{\xi}$ and without relying on PC projection. The plots show that the results of the PSP and direct MC methods asymptotically agree when N_W increases. It can also be seen that the effect of the PSP level for $\ell > 1$ is barely visible, as one would have expected from the fast convergence of the PSP approximation for this problem. Also note that the level of fluctuations in the random estimates of the sensitivity indices is much smaller for PSP than with MC. This may



(a) Approximation error in \hat{A} .



(b) L_2 -norm of approximation error.

Figure 3.3: (a) $A - \hat{A}$ versus $\boldsymbol{\xi}$ for fixed $W^{(i)}$. Plotted are 2D surfaces for $\ell = 1, 2$ and 3 arranged from left to right. (b) Normalized L_2 error versus ℓ (left) and $\log(N_Q)$ (right).

be explained by the fact that the PC expansion yields an “exact” $\boldsymbol{\xi}$ sampling for each $W^{(i)}$ with an improved convergence as a result for the estimators of S_{par} , S_{noise} and S_{mix} . In fact, if A was independent of W , using $N_W = 1$ would be enough to capture all the variability in A . This can be appreciated from Fig. 3.4, where dashed lines are used to depict error curves for the same problem, but with a reduced diffusion coefficient ($Q_2(\boldsymbol{\xi}) \sim [0.02, 0.03]$) and consequently a lower noise influence.

To analyze the effect of “exact” $\boldsymbol{\xi}$ on the sensitivity estimates, we plot in Figure 3.5a the standard errors (SE) of the PSP and direct MC methods against N_W . The SE are estimated by means of bootstrapping [11]. The results indicate that PSP has lower error than MC for the same number of samples N_W . In particular, we observe a difference of four orders of magnitude between the two methods. However, the reduction of the SE will not necessarily produce computational savings since the PSP requires the resolution of N_Q SDEs for each sample $W^{(i)}$. The PSP cost is estimated as $N_Q \times N_W$, that is total number of solves for the SDE to construct the PSP approximation, when it is just $3N_W$ for MC for the three sensitivity indices as explained in Appendix A. Hence, the reduction in the SE in the PSP estimators may not necessarily translate in an improvement of the overall computational efficiency if N_Q is large. Figure 3.5b shows the evolution of the SE in the MC and PSP approaches as function

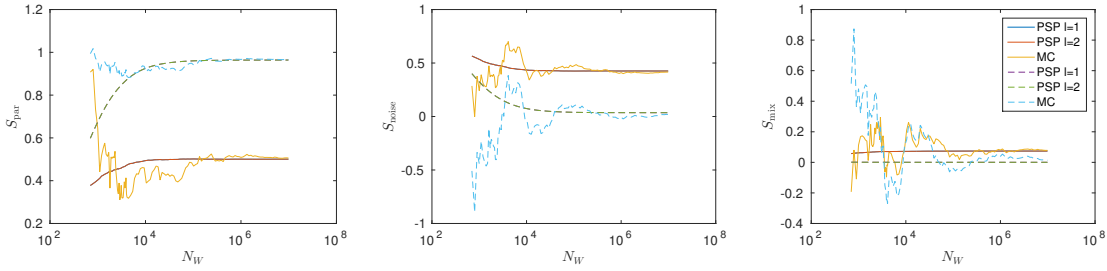


Figure 3.4: S_{par} , S_{noise} and S_{mix} versus N_W of W for different PSP levels. Also shown for comparison are the estimators for a direct MC approach (without PSP approximation, labeled MC). The solid lines correspond to the settings of Section 3.1, whereas dashed lines correspond to results obtained using PSP with a reduced diffusion coefficient $Q_2(\boldsymbol{\xi}) \sim U[0.02, 0.03]$.

of their respective cost. Plotted are results obtained for the two previous settings, i.e. with high and low noise levels. In both cases and for all three indices, it is seen that the PSP approach out performs the classical MC estimation.

4. Case of Non-Smooth QoI. In contrast to the example of the time integral in Section 3.2, the smoothness (with respect to uncertain parameters) of the SDE solution does not necessarily translate to smooth quantities of interest. As an example, we continue to consider the solution, X , of the SDE in (3.1), but focus on the variance decomposition of the exit time corresponding to the exit boundary $X = c$. The quantity of interest, denoted T is now defined by

$$T(W, \boldsymbol{\xi}) = \min_{t>0} \{X(t, W, \boldsymbol{\xi}) > c\}. \tag{4.1}$$

In the following we set $c = 1$. Figure 4.1 shows trajectories X up to their respective exit times T . Again, these trajectories correspond to the same realization $W^{(i)}$ of the noise but different random realizations of the uncertain parameters $Q(\boldsymbol{\xi})$. We observe that even for a fixed realization of the noise, the variance of $T(\cdot, \boldsymbol{\xi})$ can be very large as for some parameter values X quickly exits, while for other cases it takes a very long time to hit the exit boundary. In fact, the dependence of the exit time on the random parameters is not expected to be continuous, as an infinitesimal change in the parameter values can lead to finite delay in T .

4.1. Indirect PSP approximation. The non-smooth dependence of $T(W^{(i)}, \boldsymbol{\xi})$ with respect to both of its arguments results in a compromised convergence of its PC approximation. Specifically, the direct expansion of T over bases of globally smooth polynomial Ψ_α exhibits a slow convergence rate with the polynomial degree of the expansion, and therefore would require a prohibitively large number of sparse grid points. This is illustrated in Figure 4.2 which shows the direct PSP approximation of $T(W^{(i)}, \boldsymbol{\xi})$, computed as in Section 3.2. As the PSP level ℓ increases, the directed PSP projection of $T(W^{(i)}, \boldsymbol{\xi})$ does not converge, at least not as fast as reported in the experiments of the previous section. Instead, the approximation is seen to be plagued by higher and higher frequency oscillations as ℓ is increased, denoting the emergence of Gibbs phenomena. For the highest PSP level reported ($\ell = 5$ in the bottom

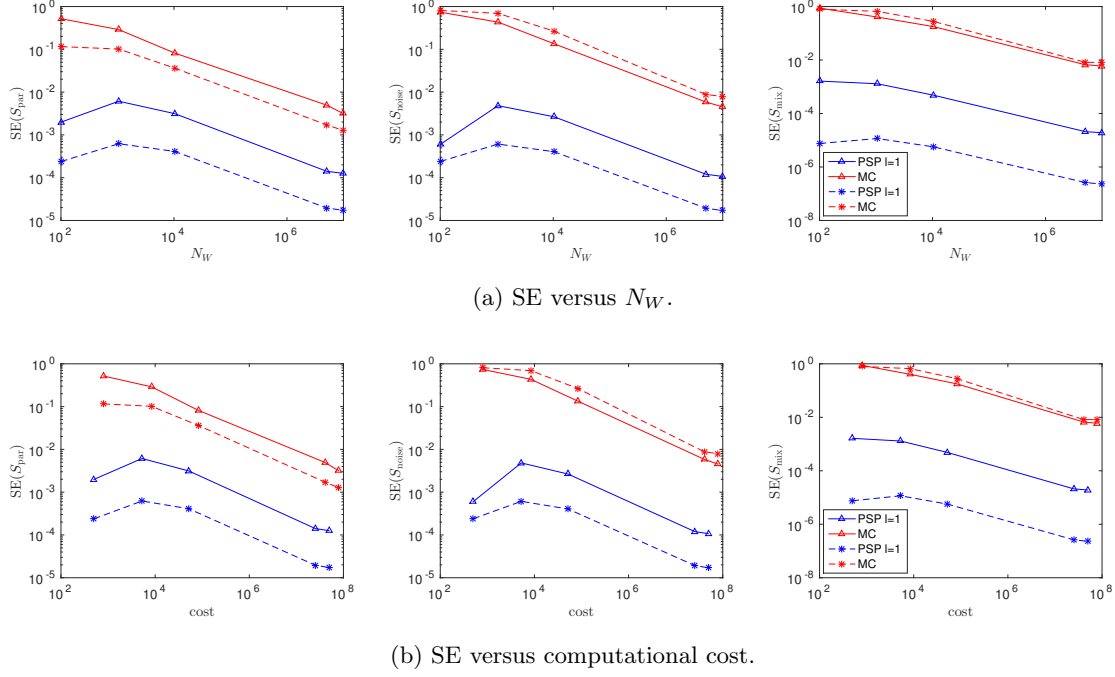


Figure 3.5: Standard Errors in the sensitivity indices of A , using the direct PSP with $N_Q = 5$ and standard MC methods. The solid lines correspond to the settings of Section 3.1, whereas the dashed lines correspond to the reduced diffusion coefficient $Q_2(\boldsymbol{\xi}) \sim U[0.02, 0.03]$. The top rows present the SEs as function of N_W , whereas the bottom rows show their dependences on the computational cost.

right plot of Figure 4.2), we remark that although the number of sparse grid points is significant ($N_Q = 321$), the approximation error on the exit time is still very large, in particular for small values of ξ_1 and ξ_2 . In fact, these situations correspond to a low asymptotic expected value for X (less than the exit level) and small diffusion coefficient. For this combination, the probability of $W^{(i)}$ to push X up to the exit boundary is small, with large expected exit time T as a result.

Even though the mapping $X \mapsto T(X)$ is not smooth, with the detrimental effects just underlined, we can still exploit the smoothness of X and the fast convergence of \hat{X} to accurately estimate the sensitivity indices of the exit time T . The key idea is to substitute the exact exit time T (or its direct non-intrusive approximation), with the surrogate $\tilde{T} = T(\hat{X})$, *indirectly* constructed through

$$T(W, \boldsymbol{\xi}) \approx \tilde{T}(W, \boldsymbol{\xi}) = \min_{t>0} \{\hat{X}(t, W, \boldsymbol{\xi}) > c = 1\}, \quad \hat{X}(t, W, \boldsymbol{\xi}) = \sum_{\alpha \in \mathcal{A}} [X_\alpha](t, W) \Psi_\alpha(\boldsymbol{\xi}). \quad (4.2)$$

In essence, given a noise trajectory $W^{(i)}$ we first approximate the PSP approximation of $X(W^{(i)}, \boldsymbol{\xi})$ and use this PSP approximation to approximate the exit time $T(W^{(i)}, \boldsymbol{\xi})$. Therefore, this approach is termed indirect. To demonstrate the validity of the indirect approach,

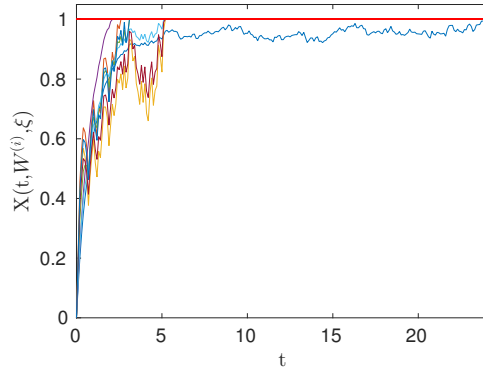


Figure 4.1: Several trajectories of the solution of the SDE, stopped when hitting the exit boundary $X = 1$. Shown are trajectories for the same noise realization $W^{(i)}$, and different parameter values $Q(\xi)$.

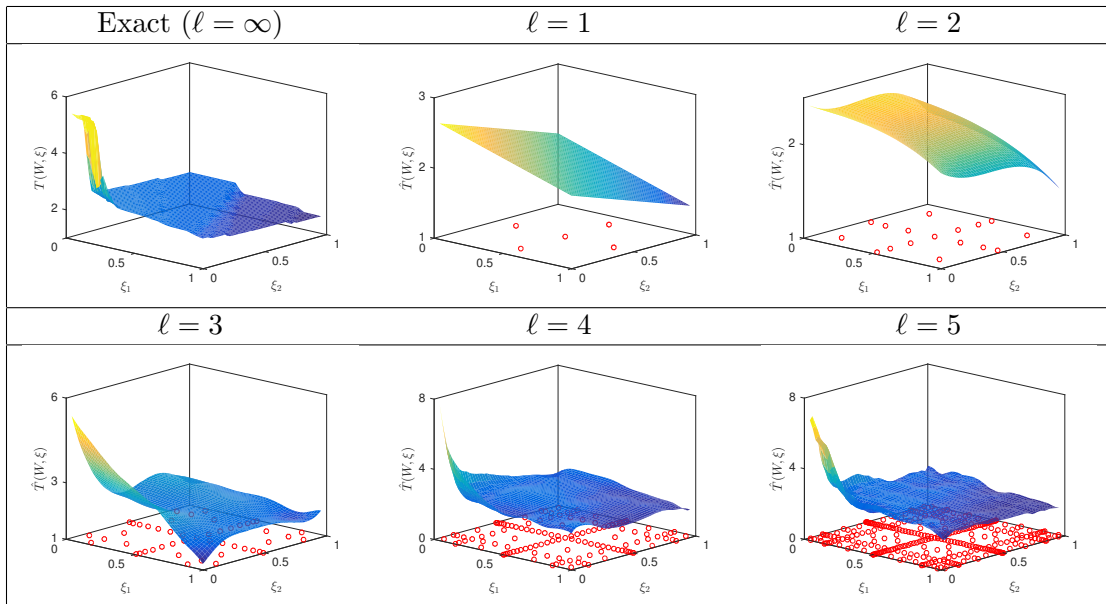


Figure 4.2: Direct PSP approximation of $T(W^{(i)}, \xi)$ for different levels ℓ as indicated. Also shown as circles are the sparse grid points used in the PSP constructions.

we present in Figure 4.3 the approximated exit time \tilde{T} defined from (4.2). The plots shows the dependence on ξ of $\tilde{T}(W, \xi)$, for a single realization of W and different levels ℓ in the approximation of X (top row), as well as the corresponding indirect approximation errors $T - \tilde{T}$. It is seen that for this realization W of the noise, a PSP expansion of X at level $\ell = 2$ is enough to obtain an absolute indirect error in the exit time of less than 10^{-5} .

We remark that the indirect approximation of $T(\cdot, \xi)$, for any $\xi \in \Xi$, requires the con-

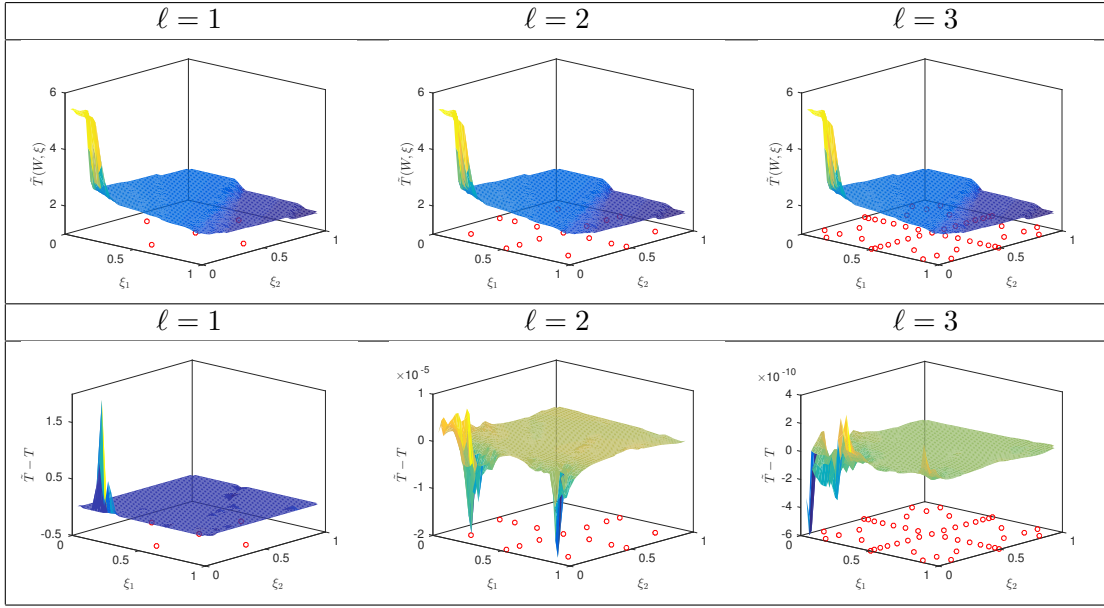


Figure 4.3: Indirect approximation \tilde{T} of the exit time through (4.2) (top row) and absolute indirect approximation error $T - \tilde{T}$ (bottom row). The plots correspond to a fixed trajectory W of the noise and different PSP levels ℓ as indicated.

struction of $\hat{X}(t, \cdot, \xi)$ up to a time t_f , when the system has exited for all values of ξ . It implies that the whole set of N_Q PSP points are integrated up to (at least) time t_f .

To better appreciate the convergence of the indirect approximation \tilde{T} based on (4.2), Figure 4.4 shows the convergence of the L_2 error norm (following (3.2) with T instead of X). The left plot shows the dependence of the error norm on the number of levels ℓ in the PSP approximation of T ; the right plot depicts the same error norm but as a function of the number of PSP points N_Q (note the log-scale). The two plots look similar as N_Q increases roughly exponentially with ℓ . They demonstrate the exponential convergence rate of the indirect approximation for this non-smooth QoI; specifically, approximation within machine precision is achieved for just $\ell = 4$ levels.

4.2. Estimation of the Sensitivity indices. We now turn to the estimation of the partial variances and sensitivity indices of the exit time T . Contrary to the direct approximation, where PC coefficients of T would be available, the expressions of the partial variances in (2.23) and (2.24) can not be used for the indirect approximation, and an alternative approach is needed to estimate V_{par} , V_{noise} and V_{mix} . The indirect approximation of the exit time provides natural means of proceeding by Monte Carlo sampling, where instead of sampling T one samples \tilde{T} . The Monte Carlo sampling for the estimation of the sensitivity indices was proposed in [37], and is briefly summarized in the Appendix. For simplicity, the estimation of the partial variance V_{noise} is considered here. As shown in Section A, the MC estimation of V_{noise} requires the estimation of a correlation of the form $\mathbb{E}\{T(W, \xi)T(W, \xi')\}$, where ξ and ξ' are two independent replicas of the uncertain parameters. The MC estimate can be expressed

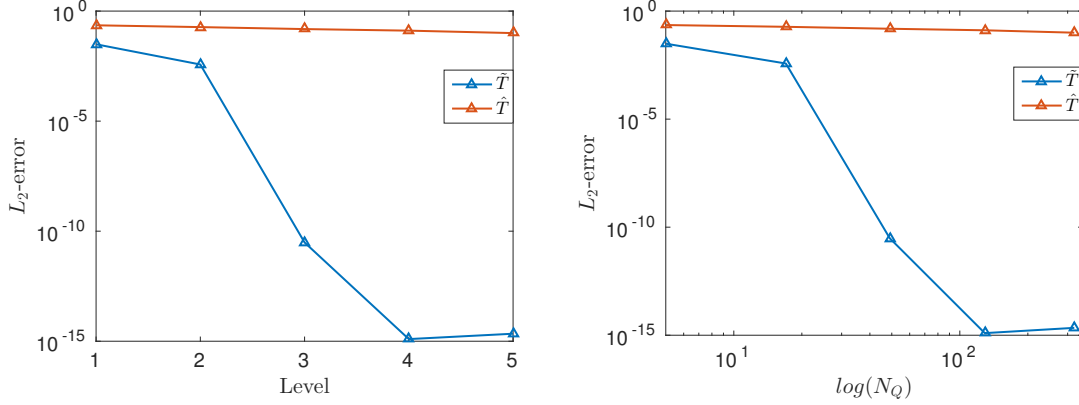


Figure 4.4: L_2 error norms of the direct and indirect PSP approximations of T .

as

$$\mathbb{E} \{T(W, \boldsymbol{\xi})T(W, \boldsymbol{\xi}')\} \approx \frac{1}{N} \sum_{i=1}^N T(W^{(i)}, \boldsymbol{\xi}^{(i)})T(W^{(i)}, \boldsymbol{\xi}'^{(i)}). \quad (4.3)$$

In this approach, one jointly samples (twice) both the Wiener and parameter spaces. Thus for each element of the sum in (4.3), the SDE needs be solved twice using the same noise $W^{(i)}$ but two parameters values $Q(\boldsymbol{\xi}^{(i)})$ and $Q(\boldsymbol{\xi}'^{(i)})$. We propose in the following to rely on the indirect exit time approximation \tilde{T} in place of T . Replacing T will introduce a bias that can be controlled by increasing the accuracy of the approximation (increasing ℓ in the PSP approach). In addition, the statistical independence of the noise and parameters is here exploited to recast the MC estimate (4.3) into the following form

$$\mathbb{E} \{T(W, \boldsymbol{\xi})T(W, \boldsymbol{\xi}')\} \approx \frac{1}{N_W} \sum_{i=1}^{N_W} \left[\frac{1}{N_\xi} \sum_{j=1}^{N_\xi} \tilde{T}(W^{(i)}, \boldsymbol{\xi}^{(i,j)})\tilde{T}(W^{(i)}, \boldsymbol{\xi}'^{(i,j)}) \right]. \quad (4.4)$$

In (4.3), $\boldsymbol{\xi}^{(i,j)}$ and $\boldsymbol{\xi}'^{(i,j)}$ are two independent random samples. The main difference between (4.3) and (4.4) is that for each sample $W^{(i)}$ of the noise, the later formula uses multiple samples (N_ξ) of the parameters. A similar expression can be derived for the correlation appearing in V_{par} , with a sum over N_ξ samples $\boldsymbol{\xi}^{(i,j)}$.

The two approaches are contrasted in Figure 4.5, which depicts the sensitivity indices computed using (4.3) (labeled MC in the plot) and using (4.4) for different PSP levels. Two computational cases are reported. First, the case of the process with significant noise level with $Q_1 \sim \mathcal{U}[0.95, 1.05]$, $Q_2 \sim \mathcal{U}[0.02, 0.22]$ and $\nu = 0.2$ in the top row and, second, the case with $Q_1 \sim \mathcal{U}[0.95, 1.05]$, $Q_2 \sim \mathcal{U}[0.02, 0.03]$ and $\nu = 0.0$, corresponding to a low noise OU process, in the bottom row. The plots show the convergence with ℓ of the indirect approach. In fact for the additive noise case, the PSP approximation of X with $\ell = 1$ is exact such that the estimates of the sensitivity indices are unbiased. We also observe the much faster convergence of the indirect approach, which exhibits much smaller statistical fluctuations

compared to the classical MC estimates. The improvement in the random estimates of the sensitivity indices is due to the effect of the ξ -averaging in the indirect approach, which used here $N_\xi = 2000$. Obviously, when using $N_\xi = 1$ the two approaches are equivalent and no gain in the convergence with N_W is then reported (not shown). Clearly, the improvement of the indirect approach is the more significant in the case of the OU process with low noise impact as one would have expected. Indeed, most of the variance is then explained by the parametric variability and, $V_{\text{noise}} + V_{\text{mix}}$ being small, the sampling error is less important.

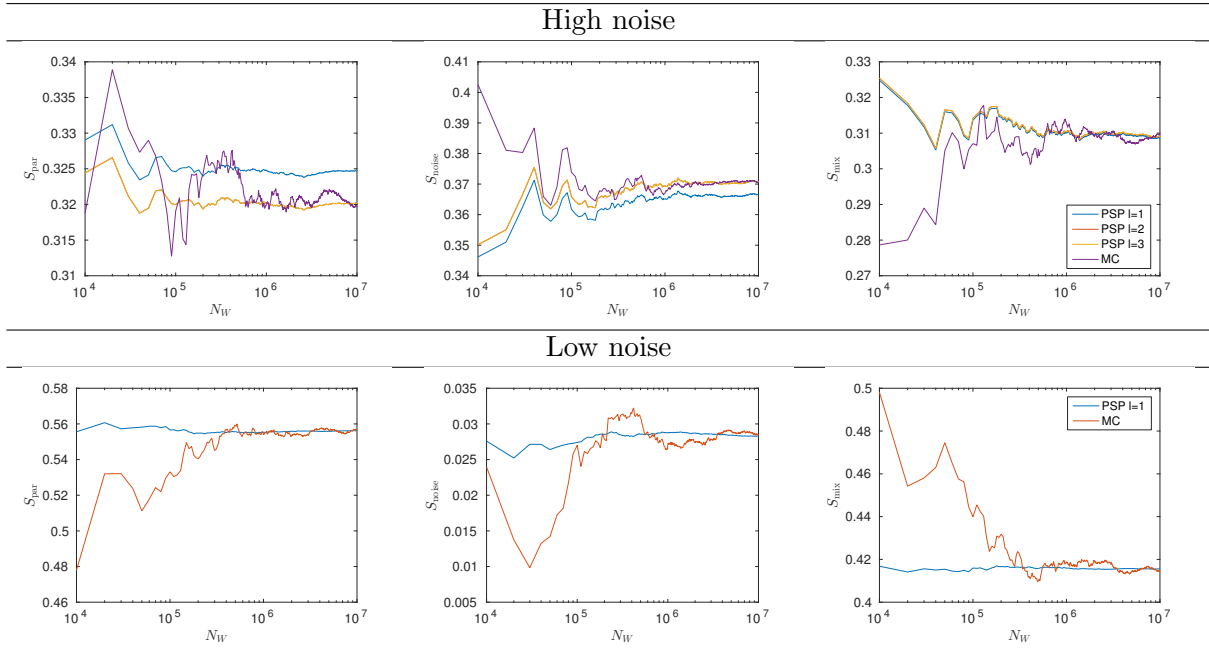


Figure 4.5: Sensitivity indices S_{par} , S_{noise} and S_{mix} versus the number, N_W , of noise samples for different PSP levels. Also shown for comparison are the standard MC estimators (labeled MC). The two rows corresponds to different systems with parameters given in the text.

To better appreciate the gain achieved by the proposed indirect approach, we present in Figure 4.6 the evolutions with the number of noise samples N_W of the standard error in the sensitivity indices (*i.e.* the standard deviation of the random estimates of S_{par} , S_{noise} and S_{mix}). The curves reported in Figure 4.6 confirm and quantify the previous claim that the indirect approach has systematically a lower standard error than the classical MC estimator for the same number of noise samples. Specifically, to obtain comparable standard error, the estimation of S_{par} using the indirect approach requires roughly a hundred time less samples than the classical MC estimator. This is the immediate benefit of the ξ -averaging. For the two other indices, S_{noise} and S_{mix} , the magnitude in the reduction of their standard errors depends on the effect of the noise on the exit time, and is less pronounced when this effect is large.

This trend can be explained by observing that the stronger the noise effects, the larger the number of W samples to reduce the sampling error, and the beneficial impact of ξ -averaging

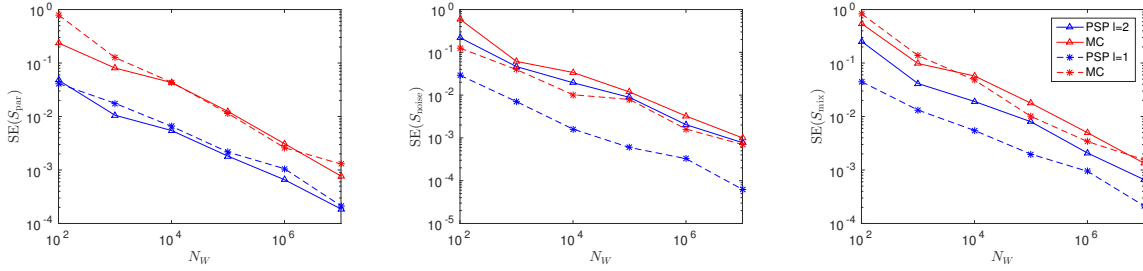


Figure 4.6: Standard errors in the sensitivity indices of the exit time, as a function of the number of samples N_W , for the indirect PSP and standard MC approaches. The indirect PSP method uses $N_\xi = 2000$. The solid lines correspond to the multiplicative noise process with large noise, whereas the dashed lines correspond to the low additive noise case (OU process).

(using N_ξ samples for each $W^{(i)}$). Figure 4.7 illustrates this trend, namely by plotting the standard errors as function of N_W for different values of N_ξ . For $N_\xi = 2$, the error has similar behavior as for classical MC (dashed line). Increasing the number of samples N_ξ , the standard error is seen to first improve, and then to level-off when the W -sampling error becomes dominant. It is also seen that the benefit of the PSP approach is more significant for S_{par} than for the two other indices.

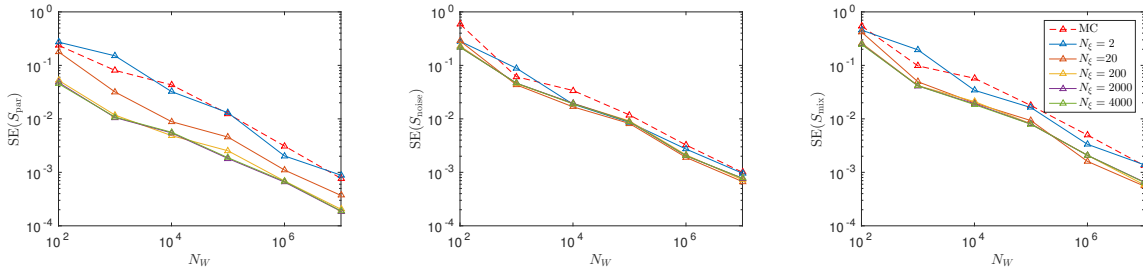


Figure 4.7: Standard errors in the sensitivity indices of the exit time versus the number of samples N_W , for the indirect PSP with different values of N_ξ , and standard MC approach. The results correspond to the process with large level multiplicative noise.

To complete the results on the treatment of non-smooth QoIs using indirect PC approximations, we discuss the efficiency of the approach. In fact, the reduction of the standard error in the random estimate of the sensitivity indices will not necessarily yield computational savings. Indeed, if the possibility of reducing the number of noise sample is essential to improve the efficiency of the estimates, one needs to account for the overhead inherent in the determination of the PC approximation of X given $W^{(i)}$. In many situations, this overhead will be proportional to the number of sparse grid points N_Q involved in the PSP approximation. This overhead must then be compared with the improvement in the standard error resulting from the ξ -averaging. For the present tests on the exit time, the PSP overhead is even

more important than just the number of sparse grid points, because the approach requires us to integrate system 2.16 up to t_f , time at which the trajectories corresponding to sample points $\xi^{(i,j)}$ have exited. For a fixed noise realization $W^{(i)}$, the last exit time t_f is random (because $\xi^{(i,j)}$ is random) with an expected values that increases with N_ξ . Indeed, denoting $T^{(i,j)} = T(W^{(i)}, \xi^{(i,j)})$ we have

$$t_f^{(i)} = \max_{j=1, \dots, N_\xi} T^{(i,j)}, \quad \mathbb{E} \left\{ t_f^{(i)} \right\} = T_f(N_\xi) \geq \mathbb{E} \left\{ T^{(i,j)} \right\}.$$

This shows the existence of a trade-off between the reduction of the standard error with increasing N_ξ and the increase of the expected computational cost per N_W sample with N_ξ . Deciding *a priori* of the optimal choice for N_ξ highly depends on the problem considered, specifically on the relative importance of the contributions to the variance (which we actually aim at characterizing) and the computational cost of building the non-intrusive approximation. We illustrate this trade-off in Figure 4.8, which shows the standard errors of Figure 4.6, plotted against the actual computational cost. The computational cost was estimated as the sum of the number of sparse grid points times the number of time steps performed till *all* the N_ξ samples have exited. In these numerical experiments we used $N_\xi = 2000$ and the computational cost is compared to the direct MC cost which reflects only the number of time steps to exit. It is seen that the efficiency of the indirect approach is less pronounced than when monitoring the standard error as a function of N_W . In particular, in the case with the largest noise effects, the indirect approach can be less effective than the classical MC estimation for V_{noise} . Note, however, that having constructed the PSP of $X(t)$ for a given $W^{(i)}$ gives access to many more information than just the sensitivity indices considered here. For instance, one could compute all sensitivity indices (*e.g.* separating the contributions of the drift and diffusion) at no additional cost, whereas the MC approach would require many more computations.

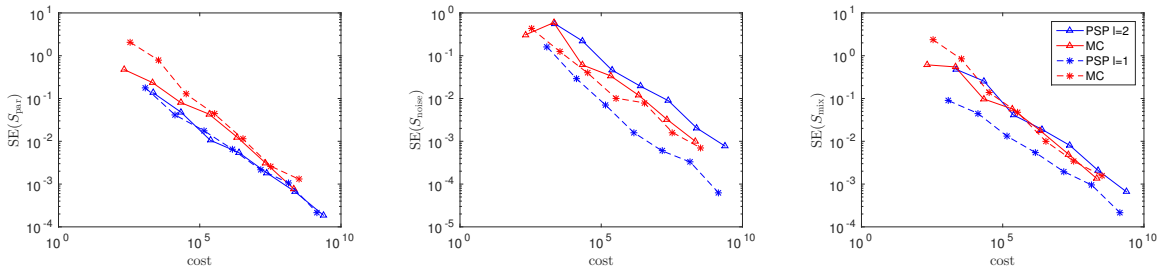


Figure 4.8: SE errors in the sensitivity indices of the exit time versus the global cost of the indirect PSP and standard MC methods. The indirect PSP method uses $N_\xi = 2000$. The solid lines correspond to the multiplicative noise process with large noise, whereas the dashed lines correspond to the low additive noise case (OU process).

5. Conclusions. In this work we have proposed an adaptation of the spectral Galerkin approach presented in [21] to non-intrusive or sampling-type methods to perform variance decompositions in uncertain SDEs. Here, we have relied on the non-intrusive pseudo-spectral

projection (PSP) method to compute the Polynomial Chaos expansion coefficients of the uncertain SDE solution. We have shown using an elementary example that the PSP with Galerkin methods yield comparable errors and have similar convergence rate, and that PSP exhibits the advantages of the non-intrusive methods making it more suitable for more complex problems. In particular, the non-intrusive approaches can be more easily implemented in parallel, compared to the Galerkin method, and we have discussed possible parallel alternatives depending on the available computational infrastructure.

We have also exploited to non-intrusive character of the approach to perform global sensitivity analysis of QoIs that correspond to functionals of the stochastic solution. The cases of smooth and non-smooth QoIs were distinguished. For the case of smooth QoI (with respect to the uncertain parameters of the SDE), a direct NI projection can be performed, and the sensitivity indices can be computed from their expansion coefficients directly. We have compared our results with a MC sampling method, and report that our PSP approach exhibits a significantly lower standard error in the estimated sensitivity coefficients for same number, N_W , of SDE noise samples. We pointed out that this improvement is due to the essentially “exact” ξ -averaging, which comes at the cost of having to solve a set of N_Q SDEs for each sample $W^{(i)}$ of the noise. This suggests that the relative efficiency of the proposed non-intrusive method depends on the problem considered, balancing the reduction in the variance of the sensitivity indices estimators, with the number of N_Q additional solves needed to construct the PSP approximation. For the case of a non-smooth QoI, we have shown that the direct non-intrusive projection on a spectral bases results in a very slow convergence. This fact was avoided by introducing a NI strategy, termed indirect, which consists in first constructing the spectral approximation of the SDE solution, and secondly sampling this approximation to generate realizations of the QoI. These samples of the QoI can then be used in a classical MC sampling method to estimate the sensitivity indices. This methodology was illustrated for the case of the exit time as QoI. We found that as expected the proposed approach has lower standard error than a classical MC estimator, for the same number N_W of noise samples. This improvement is due to the possibility of estimating efficiently conditional averages (with respect to the noise realization). However, we have shown that for the exit time example, the computational cost of building the SDE solution approximation could not only be a function of the number of non-intrusive points, N_Q , but can also be affected by the number of samples used to estimated the conditional averages of the QoIs. Situations of low noise effects, relatively to parameter uncertainty, are the most favorable to the proposed method. Further, assuming that the computational cost is dominated by the non-intrusive approximation of the SDE solution, the proposed approach has a cost essentially independent of the number of sensitivity indices being estimated, in contrast to the Monte Carlo method whose cost is proportional to the number of indices sought.

To complete the discussion, let us stress that computational savings are obtained only if the determination of accurate PC expansions is feasible at a cost (number of simulations) significantly less than that is needed by a direct MC method to sample the parameter space with the same accuracy. This may not be the case for all models and parametric variabilities; in particular, parametric bifurcations and bimodal solution models are generally expected to necessitate higher dimensional projection spaces, namely to accommodate discontinuities (see for instance Section 6 of [21]). Projection on enriched spaces (*e.g.* multiwavelet spaces [18, 19])

demands a significantly higher computational effort and would require the introduction of adaptivity and multi-resolution procedures [45] to eventually achieve computational savings in a non-intrusive framework.

Appendix. MC estimation of the sensitivity indices. We briefly recall the MC sampling to estimate the sensitivity indices, using the Homma and Saltelli method [37]. We adapt the notations to the problem considered here, namely the computation of the partial variances associated to noise (W), parametric (ξ), and mixed contributions. To this end, consider a generic quantity of interest $F(W, \xi)$. The estimators use two independent random sample sets of realizations of (W, ξ) , each with size N . We denote \mathcal{S} and $\tilde{\mathcal{S}}$ these two sample sets.

$$\mathcal{S} = \{(W^{(i)}, \xi^{(i)}), i = 1, \dots, N\}, \quad \tilde{\mathcal{S}} = \{(\tilde{W}^{(i)}, \tilde{\xi}^{(i)}), i = 1, \dots, N\}. \quad (\text{A.1})$$

The partial variances V_{par} and V_{noise} can be estimated by combining the two sample sets as follows:

$$V_{\text{par}} \approx \hat{V}_{\text{par}} = \frac{1}{N-1} \sum_{i=1}^N F(W^{(i)}, \xi^{(i)})F(\tilde{W}^{(i)}, \xi^{(i)}) - \frac{1}{N} \sum_{i=1}^N F(W^{(i)}, \xi^{(i)})^2,$$

$$V_{\text{noise}} \approx \hat{V}_{\text{noise}} = \frac{1}{N-1} \sum_{i=1}^N F(W^{(i)}, \xi^{(i)})F(W^{(i)}, \tilde{\xi}^{(i)}) - \frac{1}{N} \sum_{i=1}^N F(W^{(i)}, \xi^{(i)})^2.$$

The first-order sensitivity indices are then estimated using

$$S_{\text{par}} \approx \frac{\hat{V}_{\text{par}}}{\widehat{\mathbb{V}}\{F\}}, \quad S_{\text{noise}} \approx \frac{\hat{V}_{\text{noise}}}{\widehat{\mathbb{V}}\{F\}}, \quad S_{\text{mix}} \approx 1 - S_{\text{par}} - S_{\text{noise}}. \quad (\text{A.2})$$

In the previous equations, the estimation of the variance of F , $\widehat{\mathbb{V}}\{F\}$, follows the classical MC estimator:

$$\widehat{\mathbb{V}}\{F\} \doteq \frac{1}{N-1} \sum_{i=1}^N F(W^{(i)}, \xi^{(i)})^2 - \widehat{\mathbb{E}}\{F\}^2, \quad \widehat{\mathbb{E}}\{F\} \doteq \frac{1}{N} \sum_{i=1}^N F(W^{(i)}, \xi^{(i)}). \quad (\text{A.3})$$

The MC estimation above can be extended to other sensitivity indices considering different groupings and combinations of the sampled variables in the two initial sample sets \mathcal{S} and $\tilde{\mathcal{S}}$.

REFERENCES

- [1] H. L. ANDERSON, *Metropolis, Monte Carlo and the MANIAC*, Los Alamos Science, 1986.
- [2] O. E. BARNDORFF-NIELSEN AND N. SHEPHARD, *Non-gaussian ornstein uhlenbeck-based models and some of their uses in financial economics*, Journal of the Royal Statistical Society: Series B (Statistical Methodology), 63 (2001), pp. 167–241.
- [3] F. BRAUER, P. VAN DEN DRIESSCHE, AND J. WU, *Mathematical epidemiology*, Springer, 2008.
- [4] H. BROHUS, C. FRIER, P. HEISELBERG, AND F. HAGHIGHAT, *Quantification of uncertainty in predicting building energy consumption: A stochastic approach*, Energy and Buildings, 55 (2012), pp. 127–140.
- [5] R. E. CAFLISCH, *Monte carlo and quasi-monte carlo methods*, Acta Numerica, (1998), pp. 1–49.
- [6] R. H. CAMERON AND W. T. MARTIN, *The orthogonal development of nonlinear functionals in series of Fourier-Hermite functionals*, Annals of Mathematics, 48 (1947), pp. 385–392.

- [7] J.-B. CHEN, R. GHANEM, AND J. LI, *Partition of the probability-assigned space in probability density evolution analysis of nonlinear stochastic structures*, Probabilistic Engineering Mechanics, 24 (2009), pp. 27–42.
- [8] P. R. CONRAD AND Y. M. MARZOUK, *Adaptive smolyak pseudospectral approximations*, SIAM Journal on Scientific Computing, 35 (2013), pp. A2643–A2670.
- [9] P. CONSTANTINE, M. ELDRED, AND E. PHIPPS, *Sparse pseudospectral approximation method*, Computer Methods in Applied Mechanics and Engineering, 229–232 (2012), pp. 1–12.
- [10] I. DIMOV AND R. GEORGIEVA, *Monte carlo algorithms for evaluating sobol’ sensitivity indices*, Mathematics and Computers in Simulation, 81 (2010), pp. 506–514.
- [11] B. EFRON AND J. TIBSHIRANI, *An introduction to the bootstrap*, Chapman and Hall, 1993.
- [12] B. L. FOX, *Strategies for Quasi-Monte Carlo*, Kluwer academic Pub., 1999.
- [13] T. GERSTNER AND M. GRIEBEL, *Dimension-adaptive tensor-product quadrature*, Computing, 71 (2003), pp. 65–87.
- [14] N. IKEDA AND S. WATANABE, *Stochastic Differential Equations and Diffusion Processes*, North Holland, 1981.
- [15] E. B. IVERSEN, J. M. MORALES, J. K. MOLLER, AND H. MADSEN, *Short-term probabilistic forecasting of wind speed using stochastic differential equations*, International Journal of Forecasting, (2015).
- [16] A. KEESE AND H. G. MATTHIES, *Numerical methods and Smolyak quadrature for nonlinear stochastic partial differential equations*, tech. report, Institute of Scientific Computing TU Braunschweig Brunswick, 2003.
- [17] P. E. KLOEDEN AND E. PLATEN, *Numerical Solution of Stochastic Differential Equations*, Springer-Verlag, 1999.
- [18] O. LE MAÎTRE, R. GHANEM, O. KNIO, AND H. NAJM, *Uncertainty propagation using Wiener-Haar expansions*, J. Comput. Phys., 197 (2004), pp. 28–57.
- [19] O. LE MAÎTRE, H. NAJM, R. GHANEM, AND O. KNIO, *Multi-resolution analysis of Wiener-type uncertainty propagation schemes*, J. Comput. Phys., 197 (2004), pp. 502–531.
- [20] O. P. LE MAÎTRE AND O. M. KNIO, *Spectral Methods for Uncertainty Quantification*, Springer, New York, NY, 2010.
- [21] O. P. LE MAÎTRE AND O. M. KNIO, *Pc analysis of stochastic differential equations driven by wiener noise*, Reliability Engineering & System Safety, 135 (2015), pp. 107–124.
- [22] O. P. LE MAÎTRE, M. T. REAGAN, H. N. NAJM, R. G. GHANEM, AND O. M. KNIO, *A stochastic projection method for fluid flow. II. Random process*, Journal of Computational Physics, 181 (2002), pp. 9–44.
- [23] J. S. LIU, *Monte Carlo strategies in scientific computing*, Springer Series in Statistics, 2004.
- [24] W.-L. LOH, *On latin hypercube sampling*, Annals of Statistics, 24 (1996), pp. 2058–2080.
- [25] N. MADRAS, *Lectures on Monte Carlo methods*, American Mathematical Society, Providence, RI, 2001.
- [26] M. C. MARIANI AND O. K. TWENEBOAH, *Stochastic differential equations applied to the study of geophysical and financial time series*, Physica A: Statistical Mechanics and its Applications, 443 (2016), pp. 170–178.
- [27] L. MATHELIN AND M. Y. HUSSAINI, *A Stochastic Collocation Algorithm for Uncertainty Analysis*, tech. report, NASA/CR-2003-212153, February 2003.
- [28] M. D. MCKAY, R. J. BECKMAN, AND W. J. CONOVER, *A comparison of three methods for selecting values of input variables in the analysis of output for a computer code*, Technometrics, 21 (1979), pp. 239–245.
- [29] G. N. MILL’SHTEJN, *Approximate integration of stochastic differential equations*, Theory Probab. Appl., 19 (1975), pp. 557–562.
- [30] F. NOBILE, R. TEMPONE, AND C. G. WEBSTER, *A sparse grid stochastic collocation method for partial differential equations with random input data*, SIAM J. Num. Anal., 46 (2008), pp. 2309–2345.
- [31] M. T. REAGAN, H. N. NAJM, R. G. GHANEM, AND O. M. KNIO, *Uncertainty quantification in reacting flow simulations through non-intrusive spectral projection*, Combustion and Flame, 132 (2003), pp. 545–555.
- [32] F. RIZZI, R. JONES, B. J. DEBUSSCHERE, AND O. M. KNIO, *Uncertainty quantification in MD simulations of concentration driven ionic flow through a silica nanopore. Part I: sensitivity to physical parameters of the pore*, J. Chem. Phys., 138 (2013), p. 194104.

- [33] F. RIZZI, R. JONES, B. J. DEBUSSCHERE, AND O. M. KNIO, *Uncertainty quantification in MD simulations of concentration driven ionic flow through a silica nanopore. Part II: uncertain potential parameters*, J. Chem. Phys., 138 (2013), p. 194105.
- [34] F. RIZZI, H. N. NAJM, B. J. DEBUSSCHERE, K. SARGSYAN, M. SALLOUM, H. ADALSTEINSSON, AND O. M. KNIO, *Uncertainty quantification in MD simulations. Part I: Forward propagation*, Multiscale Model. Simul., 10 (2012), p. 14281459.
- [35] F. RIZZI, H. N. NAJM, B. J. DEBUSSCHERE, K. SARGSYAN, M. SALLOUM, H. ADALSTEINSSON, AND O. M. KNIO, *Uncertainty quantification in MD simulations. Part II: Bayesian inference of force-field parameters*, Multiscale Model. Simul., 10 (2012), p. 14601492.
- [36] A. SAARINEN, M. L. LINNE, AND O. YLI-HARJA, *Stochastic differential equation model for cerebellar granule cell excitability*, PLoS Computational Biology, 4 (2008), p. e1000004.
- [37] A. SALTELLI, *Making best use of model evaluations to compute sensitivity indices*, Computer Physics Communications, 145 (2002), pp. 280–297.
- [38] A. SALTELLI, S. TARANTOLA, F. CAMPOLONGO, AND M. RATTO, *Sensitivity analysis in practice: a guide to assessing scientific models*, John Wiley & Sons, 2004.
- [39] K. SARGSYAN, B. DEBUSSCHERE, H. N. NAJM, AND Y. MARZOUK, *Bayesian inference of spectral expansions for predictability assessment in stochastic reaction networks*, Journal of Computational and Theoretical Nanoscience, 6 (2009), pp. 2283–2297.
- [40] K. SARGSYAN, C. SAFTA, B. DEBUSSCHERE, AND H. NAJM, *Multiparameter spectral representation of noise-induced competence in bacillus subtilis*, IEEE/ACM Transactions on Computational Biology and Bioinformatics, 6 (2012), pp. 1709–1723.
- [41] S. A. SMOLYAK, *Quadrature and interpolation formulas for tensor products of certain classes of functions*, Dokl. Akad. Nauk SSSR, 4 (1963), pp. 240–243.
- [42] I. M. SOBOL', *Sensitivity estimates for nonlinear mathematical models*, Math. Modeling & Comput. Exp., 1 (1993), pp. 407–414.
- [43] I. M. SOBOL', *Global sensitivity indices for nonlinear mathematical models and their monte carlo estimates*, Math. Comput. Sim., 55 (2001), pp. 271–280.
- [44] A. TOCINO AND R. ARDANUY, *Runge-Kutta methods for numerical solution of stochastic differential equations*, J. Comp. App. Math, 138 (2002), pp. 219–241.
- [45] J. TRYOEN, O. LE MAÎTRE, AND A. ERN, *Adaptive anisotropic spectral stochastic methods for uncertain scalar conservation laws*, SIAM J. Sci. Comp., 34 (2012), pp. 2459–2481.
- [46] K. TWARDOSKA, *Wong-Zakai approximations for stochastic differential equations*, Acta Applic. Math., 43 (1996), pp. 317–359.
- [47] D. VENTURI, T. P. SAPSIS, H. CHO, AND G. E. KARNIADAKIS, *A computable evolution equation for the joint response-excitation probability density function of stochastic dynamical systems*, Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, 468 (2012), pp. 759–783, doi:10.1098/rspa.2011.0186, <http://rspa.royalsocietypublishing.org/content/468/2139/759>, arXiv:<http://rspa.royalsocietypublishing.org/content/468/2139/759.full.pdf>.
- [48] N. WIENER, *Sensitivity estimates for nonlinear mathematical models*, American Journal of Mathematics, 60 (1938), pp. 897–936.