A Posteriori Error and Optimal Reduced Basis for Stochastic Processes Defined by a Finite Set of Realizations

Guillaume Perrin, Christian Soize, Denis Duhamel, Christine Funfschilling

A POSTERIORI ERROR AND OPTIMAL REDUCED BASIS FOR STOCHASTIC PROCESSES DEFINED BY A FINITE SET OF REALIZATIONS.

G. PERRIN*†‡, C. SOIZE*, D. DUHAMEL†, AND C. FUNFSCHILLING‡

Abstract.
The use of reduced basis has spread to many scientific fields for the last fifty years to condense the statistical properties of stochastic processes. Among these basis, the classical Karhunen-Loève basis corresponds to the Hilbertian basis that is constructed as the eigenfunctions of the covariance operator of the stochastic process of interest. The importance of this basis stems from its optimality in the sense that it minimizes the total mean square error. When the available information about this stochastic process is characterized by a limited set of independent realizations, the covariance operator is not perfectly known. In this case, there is no reason for the Karhunen-Loève basis associated with any estimator of the covariance that are not converged to be still optimal. This paper presents therefore an adaptation of the Karhunen-Loève expansion in order to characterize optimal basis for projection of stochastic processes that are only characterized by a relatively small set of independent realizations.

Key words. Karhunen-Loève expansion, stochastic processes, a posteriori error, optimal basis.

AMS subject classifications. 60H35, 60H15, 60H25, 60H40, 65C50

1. Introduction. For the fifty years, stochastic process analysis has been used in an increasing number of scientific fields, such as uncertainty quantification, material science, seismology, geophysics, quantitative finance, signal processing, control engineering etc. It is indeed an interesting tool for stochastic modeling, forecasting, classification, signal detection and estimation. In most of these applications, the knowledge of these stochastic processes, that we write $X = \{X(s),\ s \in \Omega \subset \mathbb{R}\}$, is however limited. Indeed, their statistical properties are generally known through a set of independent realizations that stem from experimental measurements. From these available measurements, several statistical techniques have therefore been developed to identify the distribution P_X of stochastic process X. These methods then allow the generation of additional realizations of X that are realistic and representative of the measurements.

When $\Omega = \mathbb{R}$, AutoRegressive-Moving-Average (ARMA) models, that were first introduced by Whittle for time series [38, 39] and popularized by Box and Jenkins [7], allow the description of Gaussian stationary stochastic processes as a parameterized integral of a Gaussian white noise stochastic process. Based on limited knowledge of stochastic process X, these models can therefore be used to emphasize particular properties of X and to extrapolate its value. More recently, in the case $\Omega = [0, S]$, with $S < +\infty$, methods based on a two-step approach have given promising results to identify the distribution of *a priori* non-Gaussian and non-stationary stochastic processes. The first step of these methods is generally the approximation of the stochastic process, X, by its projection, $\widehat{X}^{\mathcal{B}^{(M)}}$, on a set $\mathcal{B}^{(M)}$ of M deterministic functions, which are supposed to be square integrable on Ω and orthonormal. The vector gathering the projection coefficients of X on each element of $\mathcal{B}^{(M)}$, denoted by

*Université Paris-Est, Modélisation et Simulation Multi-Échelle (MSME UMR 8208 CNRS), 5 Bd. Descartes, 77454 Marne-la-Vallée, France (christian.soize@univ-paris-est.fr).

†Université Paris-Est, Laboratoire Navier (UMR 8205), CNRS, ENPC, IFSTTAR, F-77455 Marne-la-Vallée (guillaume.perrin@enpc.fr, duhamel@lami.enpc.fr).

‡Innovation and Research Department, SNCF, Paris, France (christine.funfschilling@sncf.fr).

$\boldsymbol{C}$, is thus a M-dimension random vector whose components are *a priori* dependent. Then, the second step is the identification of the multidimensional distribution of $\boldsymbol{C}$. There are several techniques to perform such an identification, such as the Priori Algebraic Stochastic Modeling (PASM) methods, methods based on the Information Theory and the Maximum Entropy Principle (MEP) [15, 31], and methods based on a direct projection of random vector $\boldsymbol{C}$ on a polynomial basis of its probability space, such as Polynomial Chaos Expansion (PCE) methods (see [2, 10, 11, 12, 13, 20, 24, 32, 33]). In such two-step approaches, the more relevant regarding X the projection basis $\mathcal{B}^{(M)}$ is, the lower the dimension M has to be, to guarantee an amplitude of the residue $X - \widehat{X}^{\mathcal{B}^{(M)}}$ lower than a given threshold, and so the easier and the more precise the identification of the distribution of $\boldsymbol{C}$ will be.

As an example of such a projection basis, the Karhunen-Loève (KL) basis has played for the last decades a major role and has been applied in many works [1, 3, 5, 8, 9, 14, 26, 17, 16, 18, 19, 21, 23, 24, 25, 27, 28, 30, 33, 35, 34, 37, 40, 41]. Indeed, for any integer M, it can be extracted from the KL basis associated with X the M-dimensional family that is optimal in the sense that it minimizes among all the M-dimensional families the L^2-error associated with X. Mathematically, the KL expansion corresponds to the orthogonal projection theorem in separable Hilbert spaces, for which the Hilbertian basis, $\{k_m,\ m \geq 1\}$, is constructed as the eigenfunctions of the covariance operator of X, defined by the covariance function, R_{XX}, which is assumed to be square integrable on $\Omega \times \Omega$. When the available information about this stochastic process is characterized by a limited set of ν independent realizations, this covariance function is not perfectly known but can only be estimated. If we define $\widehat{R}(\nu)$ as the empirical estimator of R_{XX}, there is however no reason for the eigenfunctions of $\widehat{R}$ to be still optimal.

In reply to this concern, this paper presents a method to identify projection families that are as relevant as possible for X, even if the number of available realizations is relatively small. This method is based on the coupling of an alternative approximation of the covariance operator and an original use of the "leave-one-out" method to *a posteriori* evaluate the projection errors for X. Section 2 introduces therefore the method we propose to identify optimal projection basis from a set of independent realizations, and Section 3 then illustrates the possibilities of such a method on two examples.

2. Identification of optimal basis from a finite set of independent realizations.

2.1. Theoretical frame. Let $(\Theta, \mathcal{C}, P)$ be a probability space. For all $M \geq 1$, let $L^2(\Theta, \mathbb{R}^M)$ be the space of all the second-order random vectors defined on $(\Theta, \mathcal{C}, P)$ with values in $\mathbb{R}^M$, equipped with the inner product $\langle \cdot, \cdot \rangle$, such that for all $\boldsymbol{U}$ and $\boldsymbol{V}$ in $L^2(\Theta, \mathbb{R}^M)$,

$$\langle \boldsymbol{U}, \boldsymbol{V} \rangle = \int_\Theta \boldsymbol{U}^T(\theta) \boldsymbol{V}(\theta) dP(\theta) = E\left[\boldsymbol{U}^T \boldsymbol{V}\right], \tag{2.1}$$

where $E\left[\cdot\right]$ is the mathematical expectation. Let $\mathcal{P}(\Omega)$ be the space of all the second-order $\mathbb{R}$-valued stochastic processes, indexed by the compact interval $\Omega = [0, S]$, where $S < +\infty$. Let $\mathbb{H} = L^2(\Omega, \mathbb{R})$ be the space of square integrable functions on Ω, with values in $\mathbb{R}$, equipped with the inner product $(\cdot, \cdot)$, such that for all u and v in $\mathbb{H}$,

$$(u,v) = \int_\Omega u(s)v(s)ds. \tag{2.2}$$

In addition, the notation $\|\cdot\|$ will be used to denote the L^2 norm in $\mathcal{P}(\Omega)$, such that for all Z in $\mathcal{P}(\Omega)$,

$$\|Z\|^2 = E\left[(Z,Z)\right]. \tag{2.3}$$

Let $X = \{X(s),\ s \in \Omega\}$ be an element of $\mathcal{P}(\Omega)$, for which ν independent realizations, $\{X(\theta_1), \ldots, X(\theta_\nu)\}$, are supposed to be known. For the sake of simplicity, and without any loss of generality, only centered stochastic processes X are considered in this work, which reads:

$$E\left[X(s)\right] = 0, \quad s \in \Omega. \tag{2.4}$$

It is moreover assumed that the covariance function, R_{XX}, of X is square integrable on $\Omega \times \Omega$, such that:

$$R_{XX}(s,s') \overset{\text{def}}{=} E\left[X(s)X(s')\right], \quad \int_\Omega \int_\Omega R_{XX}(s,s')^2 ds ds' < +\infty. \tag{2.5}$$

Let $\mathcal{B} = \{b_m(s),\ s \in \Omega\}_{m \geq 1}$, be a Hilbertian basis of $\mathbb{H}$, such that:

$$X = \sum_{m \geq 1} C_m b_m, \quad (b_m, b_p) = \delta_{mp}, \quad C_m = (X, b_m), \tag{2.6}$$

where the projection coefficients, $\{C_m,\ m \geq 1\}$, are centered random variables that are *a priori* correlated and δ_{mp} is the kronecker symbol that is equal to 1 if $m = p$ and 0 otherwise. For practical purposes, this basis has to be truncated. For all $M \geq 1$, $\widehat{X}^{\mathcal{B}^{(M)}}$ is thus introduced as the projection of X on any M-dimension subset $\mathcal{B}^{(M)}$ of $\mathcal{B}$. Finally, the relevance of $\mathcal{B}^{(M)}$ to characterize X is analyzed with respect to the L^2-error, that is denoted by ε^2, such that:

$$\varepsilon^2(\mathcal{B}^{(M)}) \overset{\text{def}}{=} \left\| X - \widehat{X}^{\mathcal{B}^{(M)}} \right\|^2. \tag{2.7}$$

2.2. Difficulties concerning the identification of the Karhunen-Loève expansion from independent realizations. If $\mathcal{S}(\mathbb{R})$ is the set of all square integrable kernel functions defined on $\Omega \times \Omega$, such that

$$\mathcal{S}(\mathbb{R}) = \left\{ A \in L^2(\Omega \times \Omega, \mathbb{R}), \ \mid \ A(s,s') = A(s',s) \in \mathbb{R}, \quad (s,s') \in \Omega \times \Omega \right\}, \tag{2.8}$$

the solving of the eigenvalue problem associated with any function A in $\mathcal{S}(\mathbb{R})$ can be seen as a generator of a particular Hilbertian basis of $\mathcal{P}(\Omega)$, $\left\{b_m^A,\ m \geq 1\right\}$, such that:

$$\int_\Omega A(s,s') b_m^A(s') ds' = \lambda_m^A b_m^A(s), \quad \lambda_1^A \geq \lambda_2^A \geq \ldots \to 0, \quad \left(b_m^A, b_p^A\right) = \delta_{mp}. \tag{2.9}$$

Due to the orthogonal projection theorem in Hilbert spaces, the KL expansion associated with X, which is denoted by $\mathcal{K} = \{k_m,\ m \geq 1\}$ and which can be directly computed by solving the eigenvalue problem defined by Eq. (2.8) with $A = R_{XX}$, is optimal in the sense that, for all $M \geq 1$, $\mathcal{K}^{(M)} = \{k_m,\ 1 \leq m \leq M\}$ minimizes error ε^2 among all the M-dimensional families of $\mathbb{H}$:

$$\mathcal{K}^{(M)} = \arg \min_{\mathcal{B}^{(M)} \in \mathbb{H}^M} \left\{ \varepsilon^2(\mathcal{B}^{(M)}) \right\}. \tag{2.10}$$

When dealing with correlated stochastic processes that are only known through a set of ν independent realizations, $\{X(\theta_1), \ldots, X(\theta_\nu)\}$, covariance function R_{XX} is however unknown. The eigenvalue problem defined by Eq. (2.8) cannot thus be directly solved. Indirect methods can however be proposed to identify approximations of $\mathcal{K}$ from this maximal available information for X. Indeed, noticing that operator R_{XX} is also optimal for the minimization of ε^2 on $\mathcal{S}(\mathbb{R})$,

$$R_{XX} = \arg \min_{A \in \mathcal{S}(\mathbb{R})} \left\{ \varepsilon^2(\mathcal{B}_A^{(M)}) \right\}, \tag{2.11}$$

the best approximation for R_{XX} from these realizations, the most relevant for X the corresponding M-dimensional family. In this prospect, the *a priori* best evaluation of the covariance function of X is given by the following empirical estimator, $\widehat{R}(\nu)$, such that:

$$R_{XX}(s, s') \approx \widehat{R}(\nu, s, s') \stackrel{\text{def}}{=} \frac{1}{\nu} \sum_{n=1}^{\nu} X(s, \theta_n) X(s', \theta_n). \tag{2.12}$$

The solving of the eigenvalue problem defined by Eq. (2.8) with $A = \widehat{R}(\nu)$ allows therefore the definition of a basis of $\mathcal{P}(\Omega)$, $\mathcal{B}_{\widehat{R}(\nu)} = \{\widehat{k}_m, m \leq 1\}$, whose elements are likely to be well-adapted to X. However, as the rank of $\widehat{R}(\nu)$ is by construction lower than or equal to ν, only ν functions are needed to characterize this kernel:

$$\widehat{R}(\nu, s, s') = \sum_{m=1}^{\nu} \widehat{\lambda}_m \widehat{k}_m(s) \widehat{k}_m(s'), \quad (s, s') \in \Omega \times \Omega, \tag{2.13}$$

such that, from Eqs. (2.12) and (2.13), we get:

$$\left(X(\theta_n), \widehat{k}_m \right) = 0, \quad 1 \leq n \leq \nu, \quad m > \nu. \tag{2.14}$$

Hence, set $\{X(\theta_1), \ldots, X(\theta_\nu)\}$ of available realizations of X is orthogonal to the subset $\{\widehat{k}_m,\ m > \nu\}$, and the relevance of $\mathcal{B}_{\widehat{R}(\nu)}^{(M)}$ to characterize X will be limited to values of M lower than ν.

When covariance function R_{XX} is not perfectly known, the idea of the following section is therefore to propose an original method to identify M-dimensional projection families that will be adapted to X in the particular case $M > \nu$.

2.3. Proposition of an original subset of $\mathcal{S}(\mathbb{R})$. From Eq. (2.11), optimal family $\mathcal{K}^{(M)}$ can be seen as the solution of an optimization problem in $\mathcal{S}(\mathbb{R})$. This set being vast, this section aims at presenting a particular parametrized subspace of $\mathcal{S}(\mathbb{R})$ to identify accurate approximations of $\mathcal{K}^{(M)}$ when stochastic process X is only characterized by a finite set of ν independent realizations. In this prospect, another estimator for R_{XX}, that is denoted by $\widetilde{R}(\nu)$, is introduced, such that for all (s,s') in $\Omega \times \Omega$:

$$\widetilde{R}(\nu,s,s') = \begin{cases} \frac{1}{S-(s'-s)} \int_0^{S-(s'-s)} \widehat{R}(\nu,x,x+(s'-s))dx & \text{if } S > s'-s \geq 0, \\ \frac{1}{S-(s-s')} \int_0^{S-(s-s')} \widehat{R}(\nu,x+(s-s'),x)dx & \text{if } S > s-s' > 0, \\ \widehat{R}(\nu,s,s') \text{ otherwise.} \end{cases} \tag{2.15}$$

Hence, if stochastic process X is the restriction to Ω of a mean-square stationary stochastic process indexed by s in $\mathbb{R}$, that is to say if $R_{XX}(s,s')$ only depends on the difference $|s-s'|$, function $\widetilde{R}(\nu)$ is by construction the best stationary estimator of R_{XX}. The interest of such an estimator is that, although it is computed from only ν realizations of X, its rank is by construction much higher than ν. As a consequence, the number of eigenfunctions associated with $\widetilde{R}(\nu)$, which will be adapted to X, will be much higher than ν. For any error threshold T, we then define by $\widehat{M}(T)$ and $\widetilde{M}(T)$ the minimal number of eigenfunctions associated with $\widehat{R}(\nu)$ and $\widetilde{R}(\nu)$ respectively to guarantee a projection error lower than T:

$$\max\left\{\varepsilon^2(\mathcal{B}_{\widehat{R}(\nu)}^{(\widehat{M})}), \varepsilon^2(\mathcal{B}_{\widetilde{R}(\nu)}^{(\widetilde{M})})\right\} \leq T \leq \min\left\{\varepsilon^2(\mathcal{B}_{\widehat{R}(\nu)}^{(\widehat{M}+1)}), \varepsilon^2(\mathcal{B}_{\widetilde{R}(\nu)}^{(\widetilde{M}+1)})\right\}. \tag{2.16}$$

From the point of view of the minimization of Eq. (2.11), even if X is actually not mean-square stationary, $\widetilde{R}(\nu)$ can thus be considered as a better function than $\widehat{R}(\nu)$ to characterize X if and only if $\widetilde{M} \leq \widehat{M}$. From a more general point of view, for α in $[0,1]$, if we denote by $\mathbb{A}(\alpha)$ the function of $\mathcal{S}(\mathbb{R})$, such that:

$$\mathbb{A}(\alpha) = \alpha \widehat{R}(\nu) + (1-\alpha)\widetilde{R}(\nu) \in \mathcal{S}(\mathbb{R}), \tag{2.17}$$

interesting projection basis for X can be searched as the solution, $\mathcal{B}_{\mathbb{A}(\alpha^*)}^{(M)}$, of the following optimization problem:

$$\begin{cases} \mathcal{K}^{(M)} \approx \mathcal{B}_{\mathbb{A}(\alpha^*)}^{(M)}, \\ \alpha^* = \arg \min\limits_{\alpha \in [0,1]} \left\{\varepsilon^2(\mathcal{B}_{\mathbb{A}(\alpha)}^{(M)})\right\}. \end{cases} \tag{2.18}$$

Such an approach appears indeed to be efficient when the number of available realizations, ν, is small. First, by construction of any function $\mathbb{A}(\alpha)$, the relevance of optimal family $\mathcal{B}_{\mathbb{A}(\alpha^*)}^{(M)}$ is at least equal to the relevance of classical families $\mathcal{B}_{\widehat{R}(\nu)}^{(M)}$ or $\mathcal{B}_{\widetilde{R}(\nu)}^{(M)}$. Secondly, for $\alpha \neq 1$, the rank of $\mathbb{A}(\alpha)$ is by construction higher than ν, such that the number of eigenfunctions associated with $\mathbb{A}(\alpha)$, which won't be orthogonal to each available realization of X, will be greater than ν. Finally, the optimization problem defined by Eq. (2.18) being parametrized by a single scalar α, whose definition domain is a bounded interval, the computational costs associated with its minimization can be easily controlled.

Although the convexity of such a problem has not been proved in the general case, we propose in this work to use a dichotomy-based algorithm (for which formulation is presented in detail in Appendix A) for the iterative identification of such an optimal value for α, with T a given error threshold and $M \geq 1$. Under this hypothesis and using the proposed algorithm, approximations of α^* with a precision of T can be computed from the solving of at most $N_{\text{cost}} = 2 + p$ eigenvalue problems defined by Eq. (2.9), where p is the exponent for the smallest power of two that satisfies $(1/2)^p \leq T$.

2.4. A posteriori evaluation of the representativeness error. In order to solve the problem defined by Eq. (2.18), a method to *a posteriori* evaluate $\varepsilon^2(\mathcal{B}^{(M)}_{\mathbb{A}(\alpha)})$ from only ν available independent realizations, $\{X(\theta_n),\ 1 \leq n \leq \nu\}$, for all α in $[0, 1]$, then is required. Indeed, from the limited set $\{X(\theta_n),\ 1 \leq n \leq \nu\}$, the L^2-error, $\varepsilon^2(\mathcal{B}^{(M)})$, corresponding to any M-dimensional family $\mathcal{B}^{(M)}$ of $\mathbb{H}^M$, cannot be exactly calculated, but has to be evaluated as precisely as possible. Two cases can be distinguished:

- case 1: $\mathcal{B}^{(M)}$ is defined without any reference to $\{X(\theta_1), \ldots, X(\theta_\nu)\}$.
- case 2: the knowledge of $\{X(\theta_1), \ldots, X(\theta_\nu)\}$ is used to optimize the representativeness of $\mathcal{B}^{(M)}$. In this case, $\mathcal{B}^{(M)}$ depends on the available realizations of X.

2.4.1. Case 1: realizations and projection basis are independent. If $\mathcal{B}^{(M)}$ has been computed without any reference to the set $\{X(\theta_1), \ldots, X(\theta_\nu)\}$, error $\varepsilon^2(\mathcal{B}^{(M)})$ can be evaluated from its empirical estimation, $\widehat{\varepsilon}^2_\nu(\mathcal{B}^{(M)})$, such that for sufficiently high values of ν:

$$\begin{aligned}\varepsilon^2(\mathcal{B}^{(M)}) &= E\left\{\left(X - \widehat{X}^{\mathcal{B}^{(M)}}, X - \widehat{X}^{\mathcal{B}^{(M)}}\right)\right\} \\ &\approx \widehat{\varepsilon}^2_\nu(\mathcal{B}^{(M)}) \stackrel{\text{def}}{=} \frac{1}{\nu}\sum_{n=1}^{\nu}\left(X(\theta_n) - \widehat{X}^{\mathcal{B}^{(M)}}(\theta_n), X(\theta_n) - \widehat{X}^{\mathcal{B}^{(M)}}(\theta_n)\right).\end{aligned} \tag{2.19}$$

2.4.2. Case 2: the projection basis depends on the available realizations. In order to make projection family $\mathcal{B}^{(M)}$ be particularly adapted to stochastic process X, it can be interesting to exploit as much as possible the information about X that is gathered in independent realizations $\{X(\theta_1), \ldots, X(\theta_\nu)\}$. In this case, $\mathcal{B}^{(M)}$ is dependent on $\{X(\theta_1), \ldots, X(\theta_\nu)\}$, and error $\widehat{\varepsilon}^2_\nu$ can strongly underestimate ε^2. This phenomenon is generally called *over-fitting*. For instance, if we define $\mathcal{B}^{(M)} = \{b_m,\ 1 \leq m \leq M\}$ as the Gram-Schmidt orthogonalization of the deterministic family of available independent realizations $\{X(\theta_1), \ldots, X(\theta_\nu)\}$ (see Appendix B for further details about the computation of $\mathcal{B}^{(M)}$), the projection, $\widehat{X}^{\mathcal{B}^{(M)}}$, of X on $\mathcal{B}^{(M)}$ then verifies, for $1 \leq n \leq \nu$:

$$\widehat{X}^{\mathcal{B}^{(M)}}(\theta_n) = \sum_{m=1}^{M}(X(\theta_n), b_m)\, b_m = X(\theta_n). \tag{2.20}$$

By construction, error $\widehat{\varepsilon}^2_\nu(\mathcal{B}^{(M)})$ thus is always equal to zero, whereas $\varepsilon^2(\mathcal{B}^{(M)})$ should be in general strictly greater than 0, as the number of available realization, ν, and the dimension of the projection basis, M, are limited. Hence, in order to correctly evaluate error ε^2, a separation in two sets of the available realizations is generally performed:

- the first set, $\{X(\theta_1), \dots, X(\theta_{\nu^*})\}$, is a learning set, on which the definition of $\mathcal{B}^{(M)}$ is based,
- the second set, $\{X(\theta_{\nu^*+1}), \dots, X(\theta_\nu)\}$, is a validation set, on which the computation of $\widehat{\varepsilon}^2_{\nu-\nu^*}$ is achieved to evaluate ε^2.

With such a method, it can be noticed that the higher ν^*, the less precise the evaluation of ε^2. This limits strongly the scope of such approaches when the number of available realizations ν is small, compared to the number of functions that are needed to characterize X.

In this context, the application of the jackknife theory [22, 29, 6] to the evaluation of projection errors appears to be a good method to evaluate precise approximations for $\varepsilon^2(\mathcal{B}^{(M)})$ when all realizations of X have been used to compute $\varepsilon^2(\mathcal{B}^{(M)})$. To this end, for all $1 \leq i \leq \nu$, we denote by $\mathbb{X}^{(-i)} = \{X(\theta_n),\ 1 \leq n \leq \nu,\ n \neq i\}$ the set gathering all the available realizations of X but the i^{th} one, by $\mathcal{B}^{(M)}(\mathbb{X}^{(-i)})$ the M-dimensional family that has been computed from this set $\mathbb{X}^{(-i)}$ (the family that stems from the Gram-Schmidt orthogonalization defined by Eq. (B.1) is an example of such a family), and by $\widehat{X}^{(-i)}$ the projection of X on $\mathcal{B}^{(M)}(\mathbb{X}^{(-i)})$. Then, the two following hypotheses are assumed:

1. First, it is supposed that error $\varepsilon^2\left(\mathcal{B}^{(M)}(\mathbb{X}^{(-i)})\right)$ decreases when ν increases.
2. Then, given two sets $\mathbb{X}^{(-i)}$ and $\mathbb{X}^{(-j)}$ with $1 \leq i \neq j \leq \nu$, it is assumed that:

$$P_{\widehat{X}^{(-i)}} \approx P_{\widehat{X}^{(-j)}}, \tag{2.21}$$

where $P_{\widehat{X}^{(-i)}}$ and $P_{\widehat{X}^{(-j)}}$ are the distributions of $\widehat{X}^{(-i)}$ and $\widehat{X}^{(-j)}$ respectively.

In other words, the first hypothesis means that the modeling errors are expected to decrease when the information is increasing, whereas the second hypothesis asks for the application that compute the projection basis from a limited set of realizations to give robust results. It can therefore be deduced from these two hypotheses that, for ν sufficiently high:

$$\varepsilon^2(\mathcal{B}^{(M)}) \approx \varepsilon^2_{LOO}(\mathcal{B}^{(M)}), \tag{2.22}$$

$$\varepsilon^2_{LOO}(\mathcal{B}^{(M)}) \overset{\text{def}}{=} \frac{1}{\nu}\sum_{n=1}^{\nu}\left(X(\theta_n) - \widehat{X}^{(-n)}(\theta_n), X(\theta_n) - \widehat{X}^{(-n)}(\theta_n)\right). \tag{2.23}$$

According to the central limit theorem, estimator ε^2_{LOO}, which is called Leave-one-out (LOO) error, converges to ε^2 at the convergence rate of $1/\sqrt{\nu}$. Finally, the L^2-error, ε^2, in the optimization problem defined by Eq. (2.18), can be replaced by its LOO estimator, such that the M-dimensional optimal projection family, $\mathcal{K}^{(M)}$, can be approximated by the following optimization problem:

$$\begin{cases} \mathcal{K}^{(M)} \approx \mathcal{B}^{(M)}_{\mathbb{A}(\alpha^*)}, \\ \alpha^* = \arg \min_{\alpha \in [0,1]} \left\{ \varepsilon^2_{LOO}(\mathcal{B}^{(M)}_{\mathbb{A}(\alpha)}) \right\}. \end{cases} \tag{2.24}$$

As in Section 2.3, under the hypothesis that the optimization problem defined by Eq. (2.18) is convex, such that the dichotomy-based algorithm defined by Eq. (A.1) can be used, the solving of $N_{\text{cost}}(= 2 + p) \times \nu$ eigenvalue problems defined by Eq. (2.9), is required to solve the optimization problem defined by Eq. (2.24) with a precision of T for α^*. However, keeping in mind that computing the LOO error associated with any projection family amounts to solving a series of slightly modified eigenvalue problems, it was noticed that iterative methods, such as the subspace iteration methods (see [4] for further details about these methods), can be used to reduce drastically the computational time.

3. Applications. In order to illustrate the benefits that stem from the optimization problem defined by Eq. (2.24), two applications based on simulated data are introduced in this section. The first application deals with a non-Gaussian stochastic process characterized by a rather smooth correlation function. This application aims thus at justifying the relevance of the Leave-one-out error and at emphasizing the difficulties of the classical KL expansion to identify optimal basis when the number of available realizations is low, while the generalized KL expansion, characterized by Eq. (2.24), gives promising results. The second application then shows to what extent such an approach can also be applied to processes characterized by non-smooth correlation functions.

3.1. Relevance of the generalized KL expansion to approximate stochastic processes characterized by smooth covariance functions. The objective of this section is to illustrate the results shown in Section 3. To this end, we introduce the following quantities:

- $\Omega = [0, 1]$.
- R_{XX} is a particular element of $\mathcal{S}(\mathbb{R})$, such that for all κ in $[0, 1]$ and all (s, s') in $\Omega \times \Omega$:

$$\begin{cases} R_{XX}(s, s') = P(\kappa, s, s') \exp\left(-5|s - s'|\right) \cos\left(10\pi|s - s'|\right), \\ P(\kappa, s, s') = \cos\left(2\pi\kappa s\right) \cos\left(2\pi\kappa s'\right), \end{cases} \tag{3.1}$$

- $\{k_m,\ m \geq 1\}$ and $\{\lambda_m,\ m \geq 1\}$ are the respective sets of eigenfunctions and eigenvalues associated with R_{XX}. In this work, the eigenvalue problem defined by Eq. (2.9) with $A = R_{XX}$ is solved using a Galerkin-type approximation at the spatial step $h = 2.5 \times 10^{-3} \ll 1$. For the sake of simplicity, no distinction will be made between R_{XX} and its Galerkin approximation at the spatial step h, which amounts to assuming that for (s, s') in $\Omega \times \Omega$:

$$R_{XX}(s, s') = \sum_{m=1}^{N} \lambda_m k_m(s) k_m(s'), \quad N = 1/h + 1 = 401. \tag{3.2}$$

- $\boldsymbol{C} = (C_1, \ldots, C_N)$ is a N-dimension non-Gaussian random vector, whose components are uncorrelated but highly dependent, such that:

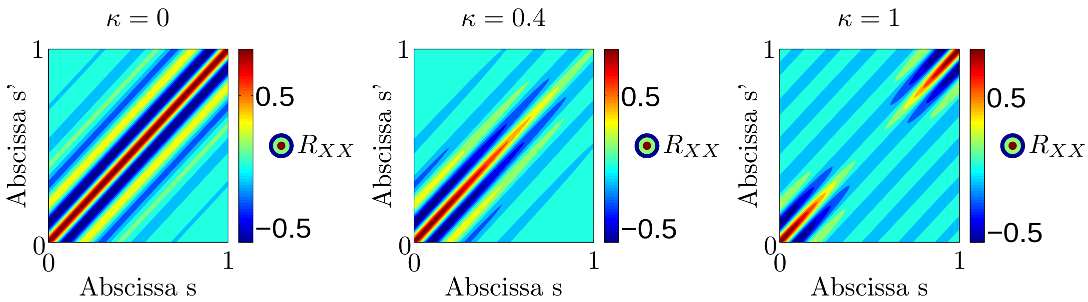

Figure 3.1: Representation of the covariance function $(s, s') \mapsto R_{XX}(s, s')$ for three particular values of κ in $[0, 1]$.

$$\boldsymbol{C} = \sum_{0 \leq \alpha_1 + \cdots + \alpha_{10} \leq 4} \boldsymbol{c}_{\alpha_1, \cdots, \alpha_{10}} \xi_1^{\alpha_1} \times \cdots \times \xi_{10}^{\alpha_{10}}, \tag{3.3}$$

where $\boldsymbol{\xi} = (\xi_1, \dots, \xi_{10})$ is a 10-dimensional random vector, whose components are independent and uniformly distributed on $[-1, 1]$, and $\boldsymbol{c}_{\alpha_1, \cdots, \alpha_{10}}$ are deterministic coefficients, whose components are uniformly and independently chosen between -1 and 1 before being normalized in order to guarantee that $E\left[C_i C_j\right] = \delta_{ij}$.

Finally, under such a formalism, we define X as the stochastic process defined by:

$$X(s) = \sum_{m=1}^{N} C_m k_m, \quad s \in \Omega. \tag{3.4}$$

By construction, $X \in \mathcal{P}(\Omega)$ is a centered, non-Gaussian stochastic process, whose covariance function is equal to R_{XX}. From each independent realization, $\boldsymbol{\xi}(\theta)$, of $\boldsymbol{\xi}$, we have therefore access to an independent realization, $X(\theta)$, of X. As an illustration, for $\kappa = 1$, four independent realizations of X are represented in Figure 3.2. Function P has moreover been introduced as a non-stationary perturbation such that the higher κ, the less mean-square stationary X. The influence of P can indeed be seen in Figure 3.1, where function R_{XX} is represented for three values of κ. Two sets, $\mathcal{X}^{\mathrm{exp}} = \{X(\theta_1), \dots, X(\theta_\nu)\}$ and $\mathcal{X}^{\mathrm{valid}} = \{X\Theta_1), \dots, X(\Theta_{\nu^{\mathrm{valid}}})\}$, of independent realizations of X then are generated. Set $\mathcal{X}^{\mathrm{exp}}$ represents the available information for X, whereas $\mathcal{X}^{\mathrm{valid}}$ is the validation set. According to Section 2.4.1, this set will only be used to evaluate the projection error corresponding to any projection family $\mathcal{B}^{(M)}$ in $\mathbb{H}^M$. In order to lighten the notations, no distinction will be made in the following between error $\varepsilon^2(\mathcal{B}^{(M)})$ and its empirical estimator, $\widehat{\varepsilon}^2_{\nu^{\mathrm{valid}}}(\mathcal{B}^{(M)})$, calculated from the realizations gathered in $\mathcal{X}^{\mathrm{valid}}$.

3.1.1. Improvement of the projection basis with respect to the available information. The number of available realizations, ν, is now supposed to be in the set $\{0, 20, 50, 100, 400\}$. The case $\nu = 0$ corresponds to a limit case when no realization

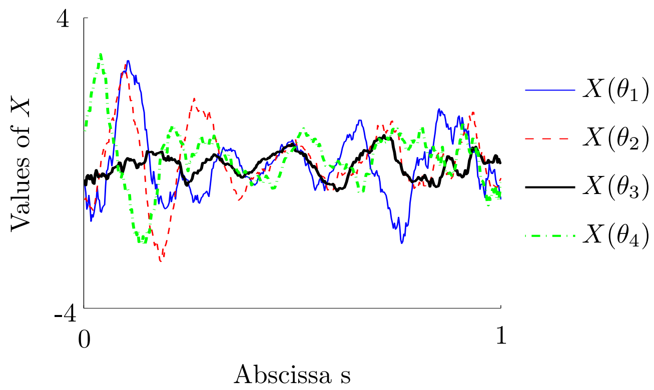

Figure 3.2: Representation of four independent realizations of X, for $\kappa = 1$.

of X is available. For the other cases, the empirical estimator of covariance function R_{XX}, which is denoted by $\widehat{R}(\nu)$, such that:

$$\widehat{R}(\nu, s, s') = \frac{1}{\nu} \sum_{n=1}^{\nu} X(\theta_n, s) X(\theta_n, s'), \quad (s, s') \in \Omega \times \Omega, \tag{3.5}$$

is compared in Figure 3.3, in the case $\kappa = 1$. In these figures, it can be verified that the higher ν, the more relevant $\widehat{R}(\nu)$. For all $\nu > 0$, $\widehat{R}(\nu)$ is in $\mathcal{S}(\mathbb{R})$ and we denote by $\mathcal{B}^{(M)}_{\widehat{R}(\nu)}$ the set gathering the M eigenfunctions of $\widehat{R}(\nu)$ of highest eigenvalues. In particular, in the case $\nu = 0$, $\mathcal{B}^{(M)}_{\widehat{R}(\nu=0)}$ corresponds to any M-dimensional set of orthonormal functions of $\mathbb{H}$. In Figure 3.4 are thus compared the evolutions of the error functions, $\varepsilon^2(\mathcal{B}^{(M)}_{\widehat{R}(\nu)})$, with respect to M, for different values of ν and κ. First, it can be noticed in these figures that $\varepsilon^2(\mathcal{B}^{(M)}_{\widehat{R}(\nu=0)})$ decreases linearly with respect to M, which means that the relevance of each element of $\mathcal{B}^{(M)}_{\widehat{R}(\nu=0)}$ to describe X is approximatively the same. This is a direct and natural consequence of the fact that all these elements have been defined without information on X. Then, two phases can clearly be identified in the evolution of $\varepsilon^2(\mathcal{B}^{(M)}_{\widehat{R}(\nu)})$ with respect to M, for $\nu = 10, 20, 50, 200$: the decrease of $\varepsilon^2(\mathcal{B}^{(M)}_{\widehat{R}(\nu)})$ is indeed much faster for $M \leq \nu$ than for $M > \nu$, where a quasi-linear decrease is found again. As explained in Section 2.2, this behavior can be justified by the fact that the ν first elements of $\mathcal{B}^{(M)}_{\widehat{R}(\nu)}$ are based on the available realizations of X, whereas the $M - \nu$ last elements are not.

3.1.2. Relevance of the Leave-one-out error. As presented in Section 2.4, when the assessment set, $\mathcal{X}^{\text{valid}}$, is not available, which is the general case, the Leave-one-out error allows us to evaluate the projection error from the only set $\mathcal{X}^{\text{exp}}$. For $\nu = 10, 20, 50, 200$ and $\kappa = 0, 0.4, 1$, the relevance of the Leave-one-out error can also be noticed in Figure 3.4, where, even for low values of ν, Leave-one-out error

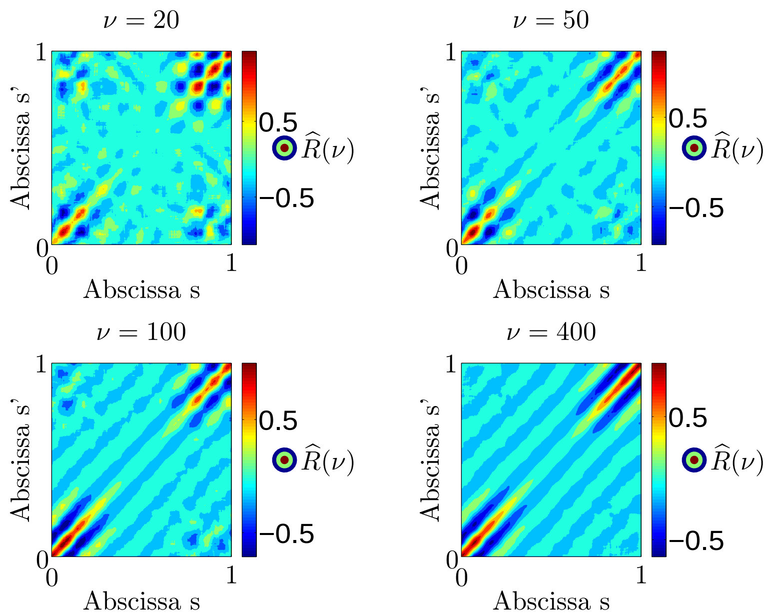

Figure 3.3: Empirical estimators $\widehat{R}(\nu)$ for four values of ν.

$\varepsilon^2_{LOO}(\mathcal{B}^{(M)}_{\widehat{R}(\nu)})$, which is only based on the ν available realizations of X, is very close to the validation error $\widehat{\varepsilon}^2_{\nu^{\text{valid}}}(\mathcal{B}^{(M)}_{\widehat{R}(\nu)})$, defined by Eq. (2.19), which is based on the $\nu^{\text{valid}} \gg \nu$ realizations gathered in $\mathcal{X}^{\text{valid}}$.

3.1.3. Optimal basis when few realizations are available. In the former section, it has been shown that family $\mathcal{B}^{(M)}_{\widehat{R}(\nu)}$ is particularly relevant to characterize X when the number of available realizations, ν, is higher than M. This section aims at illustrating the benefits of the approach introduced in Section 2.3, in cases when $M \gg \nu$ elements are needed for the projection error to be sufficiently small. To this end, the evolution of error $\varepsilon^2(\mathcal{B}^{(M)}_{\mathbb{A}(\alpha)})$ with respect to α is therefore represented in Figure 3.5, for several values of κ and M, when only $\nu = 50$ independent realizations of X are supposed to be available. On these figures, the optimal value of α to characterize X, α^*, which is solution of the optimization problem defined by Eq. (2.18), has been added. On the first hand, if X is mean-square stationary, that is to say if $\kappa = 0$, we can thus verify in this figure that the choice $\alpha = 0$ leads us particularly adapted projection basis for X. On the other hand, if X is not mean-square stationary, significant reductions for $\varepsilon^2(\mathcal{B}^{(M)}_{\mathbb{A}(\alpha)})$ can also be obtained by searching the optimal value of α in $[0, 1]$, especially when $M = 100 > \nu = 50$.

More precisely, the relevance of $\mathcal{B}^{(M)}_{\widehat{R}(\nu)}$ and $\mathcal{B}^{(M)}_{\mathbb{A}(\alpha^*)}$ is compared in Figure 3.6, for representative values of κ and ν (improvements of the same order of magnitude were obtained for the other values of (κ, ν) in $\{0, 0.4, 1\} \times \{20, 50, 100, 400\}$). The optimal value of the projection error, $\varepsilon^2(\mathcal{K}^{(M)})$, has been added in these figures as a limit state. In each case, it can thus be seen that $\varepsilon^2(\mathcal{B}^{(M)}_{\mathbb{A}(\alpha^*)}) \leq \varepsilon^2(\mathcal{B}^{(M)}_{\widehat{R}(\nu)})$, especially when

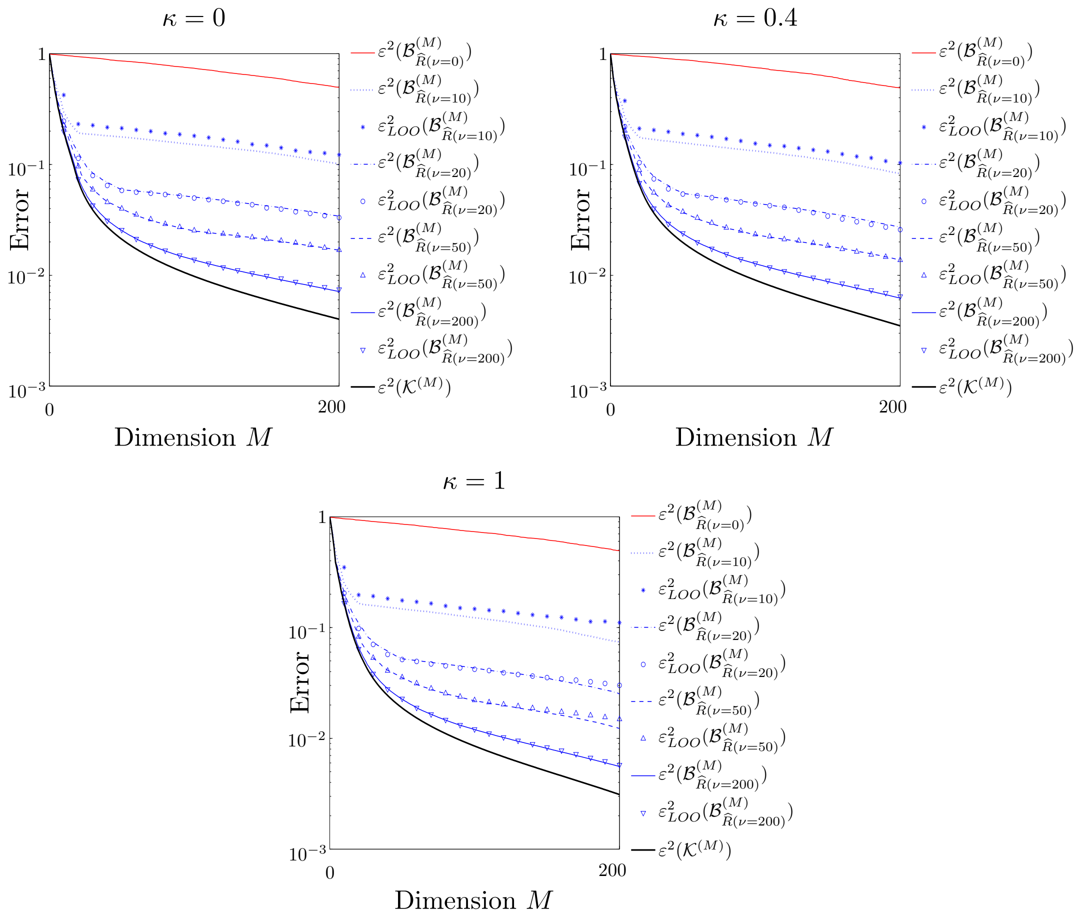

Figure 3.4: Improvement of the projection basis with respect to the number of available realizations, ν.

$\nu \ll M$. Hence, for these four choices for (κ, ν), this figure underlines the great benefits brought by the formulation defined by Eq. (2.18). For instance, for $\nu = 20$, $\kappa = 0.4$ and $M = 50$, we get:

$$\varepsilon^2(\mathcal{B}^{(50)}_{\widehat{R}(\nu)}) = 1.3\% \gg \varepsilon^2(\mathcal{B}^{(50)}_{\mathbb{A}(\alpha^*)}) = 0.16\% > \varepsilon^2(\mathcal{K}^{(50)}) = 0.12\%. \tag{3.6}$$

Indeed, whereas the rank of $\widehat{R}(\nu = 20)$ is $20 < M = 50$, the rank of $\mathbb{A}(\alpha^*)$ is by construction much higher than ν. Therefore, more elements of $\mathcal{B}^{(50)}_{\mathbb{A}(\alpha^*)}$ are based on the knowledge of X than the elements of $\mathcal{B}^{(50)}_{\widehat{R}(\nu)}$, which explains such an improvement of the projection basis, even if X is not mean-square stationary.

For all $1 \leq M \leq 200$, and all (κ, ν) in $\{0, 0.4, 1\} \times \{20, 50, 100, 400\}$, the computation of optimal value α^* was based on the dichotomy-based algorithm presented in Section 2.3, with a chosen precision $T = 5\%$. Hence, whereas $\mathcal{B}^{(50)}_{\widehat{R}(\nu)}$ is computed from

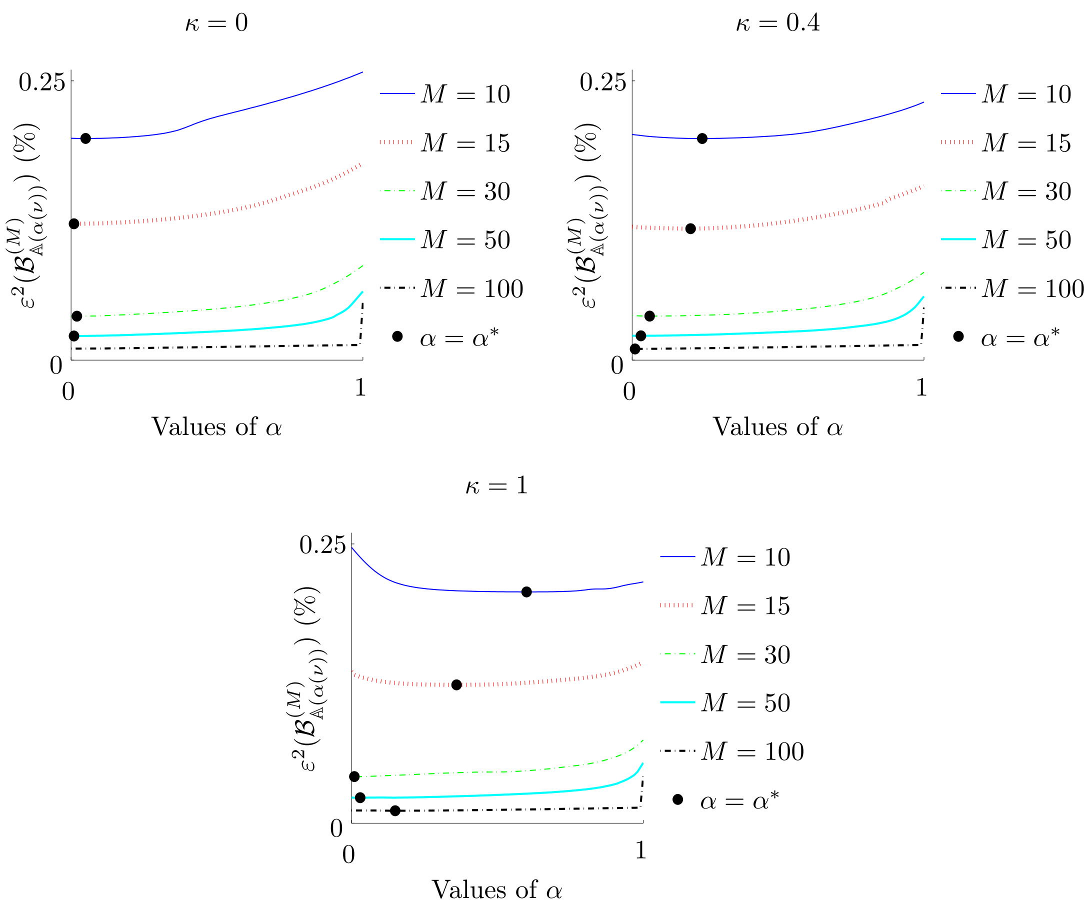

Figure 3.5: Evolution of the projection errors, $\varepsilon^2(\mathcal{B}^{(M)}_{\mathbb{A}(\alpha)})$, with respect to α, for M in $\{10, 15, 30, 50, 100\}$ and κ in $\{0, 0.4, 1\}$, when the available information is characterized by a set of $\nu = 50$ independent realizations of X.

the direct solving of the eigenvalue problem defined by Eq. (2.9), with $A = \widehat{R}(\nu)$, the computation of $\mathcal{B}^{(50)}_{\mathbb{A}(\alpha^*)}$ requires the solving of the same eigenvalue problem, with $A = \mathbb{A}(\alpha)$, for $N_{\text{cost}} = 7$ different values of α.

3.2. Relevance of the generalized KL expansion to characterize stochastic processes characterized by non-smooth correlation functions. In order to underline that the proposed approach can be applied to any type of correlation function, the Ornstein-Uhlenbeck case [36], which is characterized by a non-smooth correlation function, is tested in this section. To this end, let $X = \{X(s),\ s \in [0, 1]\}$ be the Gaussian stochastic process, whose mean function μ and covariance function R_{XX} are given by, for (s, s') in $\Omega \times \Omega$:

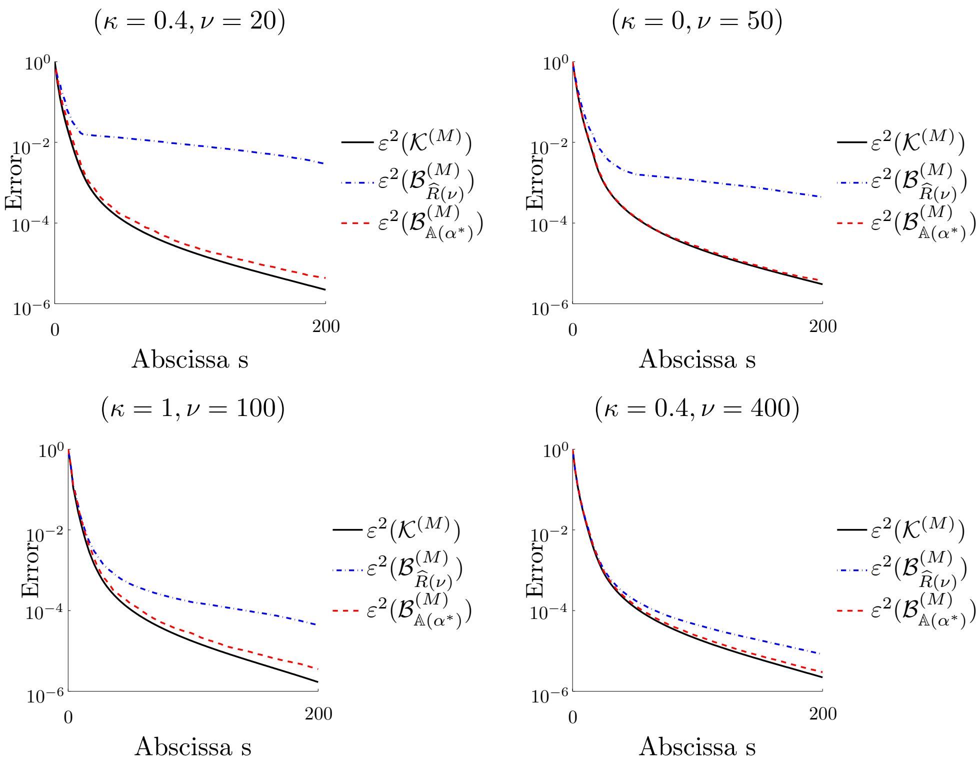

Figure 3.6: Improvement of the projection basis.

$$\begin{cases} \mu(s) = E\left[X(s)\right] = a\left(1 - \mathrm{e}^{-bs}\right), \\ R_{XX}(s,s') = E\left[X(s)X(s')\right] = \dfrac{c^2\ \mathrm{e}^{-b(s+s')}}{2b}\left(\mathrm{e}^{2b\min(s,s')} - 1\right). \end{cases} \tag{3.7}$$

where $a = 1$, $b = 10^{-2}$, and $c = 0.4$ are deterministic parameters.

In the same manner as in Section 3.1, it is now supposed that the maximal information about X is characterized by a set of only $\nu = 50$ independent realizations, which are once again denoted by $\{X(\theta_1), \ldots, X(\theta_\nu)\}$. As an illustration of such a stochastic process, Figure 3.7 compares the evolutions of the covariance function of X, R_{XX}, and of its empirical estimator, $\widehat{R}(\nu)$, whereas Figure 3.8 shows four particular realizations of X. Using the same notation as in Section 3.1, we define:

- $\mathcal{K}^{(M)}$ and $\mathcal{B}^{(M)}_{\widehat{R}(\nu)}$ as the sets gathering the M eigenfunctions of highest eigenvalues associated with R_{XX} and $\widehat{R}(\nu)$ respectively;
- α^* and $\mathbb{A}(\alpha^*)$ as the solution of the optimization problem defined by Eq. (2.18) and the corresponding function in $\mathcal{S}(\mathbb{R})$ respectively.

In this section, it is supposed that no validation set is available to evaluate the relevance of any projection family to characterize X. The Leave-one-out procedure, described in Section 2.4 has thus been applied. For all $1 \leq M \leq 100$, the evaluation of

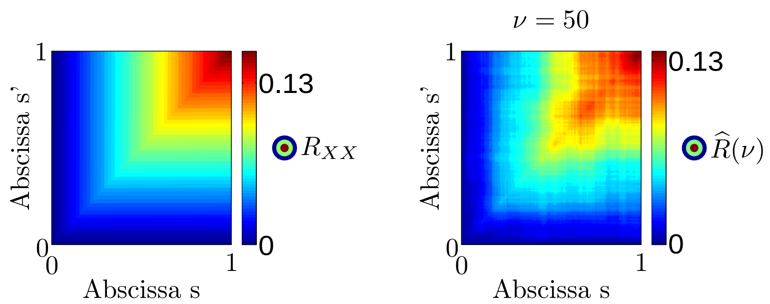

Figure 3.7: Representation of the chosen Ornstein-Uhlenbeck covariance function, R_{XX}, and of its empirical approximation, $\widehat{R}(\nu)$, for $\nu = 50$.

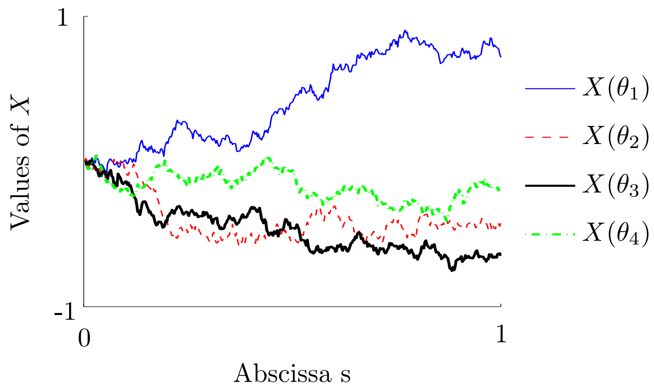

Figure 3.8: Representation of four independent realizations of Ornstein-Uhlenbeck process X.

the projection error associated with a particular projection family is therefore based on the computation of $\nu = 50$ slightly modified eigenvalue problems. Moreover, as in Section 3.1.3, the dichotomy-based algorithm presented in Section 2.3, with a precision $T = 5\%$, was once again used to identify accurate approximations of optimal value α^*. Hence, for all $1 \leq M \leq 100$, $\nu + 1 = 51$ and $N_{\text{cost}} = 7(\nu + 1) = 357$ solvings of the eigenvalue problem defined by Eq. (2.9) were needed to evaluate $\varepsilon^2_{LOO}(\mathcal{B}^{(M)}_{\widehat{R}(\nu)})$ and $\varepsilon^2_{LOO}(\mathcal{B}^{(M)}_{\mathbb{A}(\alpha^*)})$ respectively.

Finally, when the available information about X is characterized by a finite set of independent realizations, the relevance of the proposed method is illustrated in Figure 3.9. Indeed, by comparing the evolutions of errors $\varepsilon^2_{LOO}(\mathcal{B}^{(M)}_{\widehat{R}(\nu)})$, $\varepsilon^2_{LOO}(\mathcal{B}^{(M)}_{\mathbb{A}(\alpha^*)})$ and $\varepsilon^2(\mathcal{K}^{(M)})$ with respect to M, it appears that interesting improvements can be brought by the coupling of a Leave-one-out procedure and the solving of an original approximation of the covariance operator to identify accurate reduced basis for X.

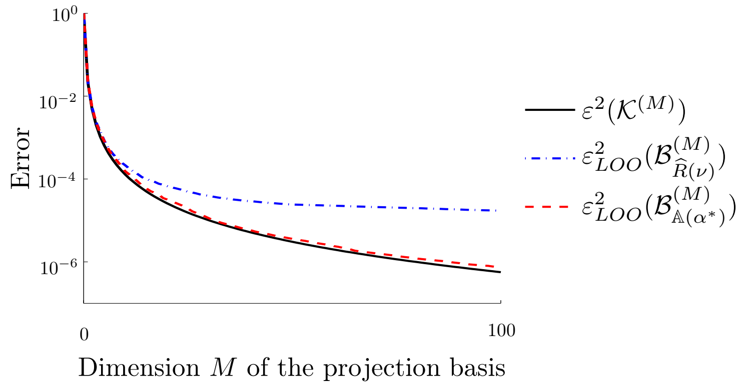

Figure 3.9: Improvement of the projection basis for the characterization of Ornstein-Uhlenbeck process X.

4. Conclusions. For the last fifty years, the increasing computational power has encouraged many scientific fields to take into account stochastic processes in their modeling. The development of reduced basis that could condense at best the statistical properties of these stochastic processes is therefore of great interest. In most of these applications, the knowledge of these stochastic processes is however limited to a finite set of independent realizations. In this context, using two examples, this paper emphasized the efficiency of a method based on an adaptation of the Karhunen-Loève expansion, in order to construct optimal basis from a relatively small set of independent realizations. This method defined first an original optimization problem, and secondly, required a way to *a posteriori* evaluate projection errors. Finally, when interested in studying complex systems that are excited by stochastic processes that are only known through a set of limited independent realizations, the proposed method opens new opportunities to optimize the projection basis with respect to the available information.

Acknowledgments. This work was supported by SNCF (Innovation and Research Department).

Appendix A. Dichotomy-based identification algorithm.

$$\left[\begin{array}{l}
\textbf{Initialize: } \alpha_1 = 0,\ \alpha_2 = 1. \\
\textbf{while } |\alpha_1 - \alpha_2| > T : \\
\quad \textbf{Compute:} \\
\quad\quad \alpha = (\alpha_1 + \alpha_2)/2, \\
\quad\quad \alpha_1^* = \arg \min_{\beta \in \{\alpha_1, \alpha, \alpha_2\}} \left\{ \varepsilon^2 (\mathcal{B}^{(M)}_{\mathbb{A}(\beta)}) \right\}, \\
\quad\quad \alpha_2^* = \arg \min_{\beta \in \{\alpha_1, \alpha, \alpha_2\},\ \beta \neq \alpha_1^*} \left\{ \varepsilon^2 (\mathcal{B}^{(M)}_{\mathbb{A}(\beta)}) \right\}. \\
\quad \textbf{Update:} \\
\quad\quad \textbf{if } (\alpha_1 = \alpha_1^*, \alpha_2 = \alpha_2^*) : \\
\quad\quad\quad \textbf{exit while}. \\
\quad\quad \textbf{else}: \\
\quad\quad\quad \alpha_1 = \alpha_1^*,\ \alpha_2 = \alpha_2^*. \\
\quad\quad \textbf{end if} \\
\textbf{end while} \\
\alpha^* = \alpha.
\end{array}\right. \tag{A.1}$$

Appendix B. Gram-Schmidt orthogonalization algorithm.

If the set $\{X(\theta_1), \ldots, X(\theta_\nu)\}$ gathers ν independent realizations of X, the following algorithm allows the identification of a projection family for X, $\mathcal{B}^{(M)} = \{b_m,\ 1 \leq m \leq M\}$, which corresponds to the Gram-Schmidt orthogonalization of $\{X(\theta_1), \ldots, X(\theta_\nu)\}$.

$$\left[\begin{array}{l}
b_1 = X(\theta_1) / \left(X(\theta_1), X(\theta_1) \right),\ K = 1. \\
\textbf{for } 2 \leq m \leq M : \\
\quad b_m^* = X(\theta_i) - \sum_{k=1}^{m-1} \left(X(\theta_m), b_k \right) b_k \\
\quad \textbf{if } (b_m^*, b_m^*) > 0 : \\
\quad\quad K = K + 1, \quad b_K = b_m^* / \sqrt{(b_m^*, b_m^*)}. \\
\quad \textbf{end if} \\
\textbf{end for} \\
M = K.
\end{array}\right. \tag{B.1}$$

REFERENCES

[1] C. Allery, A. Hambouni, D. Ryckelynck, and N. Verdon, *A priori reduction method for solving the two-dimensional burgers' equations*, Applied Mathematics and Computation, 217 (2011), pp. 6671–6679.

[2] M. Arnst, R. Ghanem, and C. Soize, *Identification of bayesian posteriors for coefficients of chaos expansions*, Journal of Computational Physics, 229 (9) (2010), pp. 3134–3154.

[3] J.A. Atwell and B.B. King, *Proper orthogonal decomposition for reduced basis feedback controllers for parabolic equations*, Math. Comput. Modell., 33 (1-3) (2001), pp. 1–19.

[4] K. J. Bathe and E.L. Wilson, *Numerical Methods in Finite Element Analysis*, Prentice-Hall, 1976.

[5] G. Berkooz, P. Holmes, and J.L. Lumley, *The proper orthogonal decomposition in the analysis of turbulent flows*, Annu. Rev. Fluid Mech., 25 (1993), pp. 539–575.

[6] G. Blatman and B. Sudret, *Adaptive sparse polynomial chaos expansion based on least angle regression*, J. Comput. Phys., 230 pages=2345-2367, (2011).

[7] G. Box and G.M. Jenkinsl, *Time Series Analysis: Forecasting and Control*, Holden-Day, San Francisco, 1970.
[8] G. P. Brooks and J. M. Powers, *A karhunen-loève least-squares technique for optimization of geometry of a blunt body in supersonic flow*, Journal of Computational Physics, 195 (2004), pp. 387–412.
[9] E.A. Christensen, M. Brons, and J.M. Sorensen, *Evaluation of proper orthogonal decomposition-based decomposition techniques applied to parameter dependent nonturbulent flows*, SIAM J. Sci. Comput, 21 (4) (2000), pp. 1419–1434.
[10] S. Das, R. Ghanem, and S. Finette, *Polynomial chaos representation of spatio-temporal random field from experimental measurements*, J. Comput. Phys., 228 (2009), pp. 8726–8751.
[11] C. Desceliers, R. Ghanem, and C. Soize, *Maximum likelihood estimation of stochastic chaos representations from experimental data*, Internat. J. Numer. Methods Engrg., 66 (2006), pp. 978–1001.
[12] C. Desceliers, C. Soize, and R. Ghanem, *Identification of chaos representations of elastic properties of random media using experimental vibration tests*, Comput. Mech., 39 (2007), pp. 831–838.
[13] R.G. Ghanem and A. Doostan, *On the construction and analysis of stochastic models: Characterization and propagation of the errors associated with limited data*, J. Comput. Phys., 217 (2006), pp. 63–81.
[14] S.P. Huang, S.T. Quek, and K.K. Phoon, *Convergence study of the truncated karhunen-loève expansion for simulation of stochastic processes*, Int J Num Meth Engng, 52 (9) (2001), pp. 1029–43.
[15] E. T. Jaynes, *Information theory and statistical mechanics*, The Physical Review, 106 (4) (1963), pp. 620–630.
[16] K. Kunisch and S. Volkwein, *Galerkin proper orthogonal decomposition methods for parabolic problems*, Numer. Math., 90 (1) (2001), pp. 117–148.
[17] O.P. Le Maître and O.M. Knio, *Spectral Methods for Uncertainty Quantification*, Springer, 2010.
[18] L. Li, K. Phoon, and S. Quek, *Comparison between karhunen-loève expansion and translation-based simulation of non-gaussian processes*, Computers and Structures, 85 (2007), pp. 264–76.
[19] X. Ma and N. Zabaras, *Kernel principal component analysis for stochastic input model generation*, Comptes rendus de l'Académie des sciences de Paris, 220 (1945).
[20] Y. M. Marzouk, H. N. Najm, and L. A. Rahn, *spectral methods for efficient bayesian solution of inverse problems*, J. Comput. Phys., 224 (2007), pp. 560–586.
[21] H.G. Matthies, *Stochastic finite elements: Computational approaches to stochastic partial differential equations*, Zamm-Zeitschrift für Angewandte Mathematik und Mechanik, 88 (11) (2008), pp. 849–873.
[22] R. G. Miller, *The jackknife - a review*, Biometrika, 61 (1974), pp. 1–15.
[23] A. Nouy and O.P. Le Maître, *Generalized spectral decomposition method for stochastic non-linear problems*, J. Comput. Phys., 228 (1) (2009), pp. 202–235.
[24] G. Perrin, C. Soize, D. Duhamel, and C. Funfschilling, *Identification of polynomial chaos representations in high dimension from a set of realizations*, SIAM J. Sci. Comput., 34(6) (2012), pp. 2917–2945.
[25] ———, *Karhunen-loève expansion revisited for vector-valued random fields: scaling, errors and optimal basis*, Journal of Computational Physics, 242 (2013), pp. 607–622.
[26] ———, *Track irregularities stochastic modeling*, Probabilistic Engineering Mechanics, 34 (2013), pp. 123–130.
[27] K.K. Phoon, S.P. Huang, and S.T. Quek, *Implementation of karhunen-loeve expansion for simulation using a wavelet-galerkin scheme*, Probabilistic Engineering Mechanics, 17 (2002), pp. 293–303.
[28] K.K. Phoon, S.P. Huang, and S.T. Quek, *Simulation of strongly non-gaussian processes using karhunen-loeve expansion*, Probabilistic Engineering Mechanics, 20 (2005), pp. 188–198.
[29] N.T. Quan, *The prediction sum of squares as a general measure for regression diagnostics*, Journal of Business and Economic Statistics, 6 (4) (1988), pp. 501–504.
[30] C. Schwab and R. A. Todor, *Karhunen-loeve approximation of random fields by generalized fast multipole methods*, Journal of Computational Physics, 217 (2006), pp. 100–122.
[31] C. Soize, *Construction of probability distributions in high dimension using the maximum entropy principle. applications to stochastic processes, random fields and random matrices.*, International Journal for Numerical Methods in Engineering, 76(10) (2008), pp. 1583–1611.

[32] C. Soize, *Generalized probabilistic approach of uncertainties in computational dynamics using random matrices and polynomial chaos decompositions.*, Internat. J. Numer. Methods Engrg., 81 (2010), pp. 939–970.

[33] C. Soize, *Identification of high-dimension polynomial chaos expansions with random coefficients for non-gausian tensor-valued random fields using partial and limited experimental data.*, Computer Methods in Applied Mechanics and Engineering, 199 (2010), pp. 2150–2164.

[34] P.D. Spanos, M. Beer, and J. Red-Horse, *Karhunen -loève expansion of stochastic processes with a modified exponential covariance kernel*, Journal of Engineering Mechanics, 133 (7) (2007), pp. 773–779.

[35] P.D. Spanos and B.A. Zeldin, *Galerkin sampling method for stochastic mechanics problems*, Journal of Engineering Mechanics, 120 (5) (1994), pp. 1091–1106.

[36] G. E. Uhlenbeck and L. S. Ornstein, *On the theory of the brownian motion*, Phys. Rev., 36 (1930), pp. 823–841.

[37] B. Wen and N. Zabaras, *A multiscale approach for model reduction of random microstructures*, Computational Materials Science, 63 (2012), pp. 269–285.

[38] P. Whittle, *Hypothesis testing in time series, PhD thesis.*, University of Uppsala, 1951.

[39] P. Whittle, *Prediction and Regulation by Linear Least-Square Methods.*, University of Minnesota Press, 1983.

[40] M.M.R. Williams, *The eigenfunctions of the karhunen-loeve integral equation for a spherical system*, Propabilistic Engineering Mechanics, 26 (2011), pp. 202–207.

[41] S. Q. Wu and S. S. Law, *Statistical moving load identification including uncertainty*, Probabilistic Engineering Mechanics, 29 (2012), pp. 70–78.