Uncertainty Quantification and Sensitivity Analysis of Crosstalk in PCB Lines via the Combination of Machine Learning and Polynomial Chaos Expansion

Uncertainty Quantification and Sensitivity Analysis of Crosstalk in PCB Lines via the Combination of Machine Learning and Polynomial Chaos Expansion

1st Paolo Manfredi
Department of Electronics and Telecommunications
Politecnico di Torino
10129 Torino, Italy
paolo.manfredi@polito.it

2nd Riccardo Trinchero
Department of Electronics and Telecommunications
Politecnico di Torino
10129 Torino, Italy
riccardo.trinchero@polito.it

***Abstract*—Manufacturing tolerances are becoming one of the major limitations to achieving high data rates in modern integrated circuits. Polynomial chaos expansion (PCE) recently became a popular alternative to classical Monte Carlo for efficient uncertainty quantification (UQ). Statistical moments and sensitivity information are analytically derived from the PCE coefficients. However, computing the model coefficients typically requires a large amount of data samples from possibly expensive design simulations. On the other hand, kernel-based machine learning methods were recently employed to create efficient surrogate models with a limited amount of training data. In this paper, we combine the kernel Gaussian process regression (GPR) method and the PCE. The former is used to surrogate the expensive simulator in the calculation of PCE coefficients via numerical integration. The use of GPR also allows obtaining confidence levels of the estimated PCE coefficients and the relative statistical information. The proposed hybrid PCE-GPR method is applied to UQ of maximum crosstalk in a PCB interconnect, for which higher accuracy with a very limited amount of data is obtained compared to state-of-the-art approaches.**

***Index Terms*—Crosstalk, Gaussian process, Kriging, machine learning, polynomial chaos, regression, signal integrity, surrogate modeling, training, transmission lines, uncertainty quantification.**

I. Introduction

As technology is pushing towards increased miniaturization and higher bitrates, manufacturing tolerances are becoming a critical limitation to high-speed signaling over PCB lines. In this scenario, signal integrity analysis must unavoidably take process variability into account, which is usually accomplished by means of Monte Carlo methods.

More efficient alternatives based on the polynomial chaos expansion (PCE) framework were recently proposed [1]. The PCE leverages expansions of special orthogonal polynomials that are specifically tailored towards uncertainty quantification (UQ). Statistical information such as moments and sensitivity indices are analytically derived from the model coefficients [2]. However, the calculation of such coefficients may be inefficient in terms of data samples required. In this regard, kernel machine learning methods, such as those based on Gaussian process regression (GPR), are more parsimonious in terms of data [3], yet they lack interpretability in an UQ scenario.

In this paper, we propose a hybrid PCE-GPR method that combines the advantages of both PCE (i.e., model interpretability) and GPR (i.e., efficient training). The advocated method is applied to the UQ of crosstalk in a PCB line.

II. PCE

Let us introduce a target quantity of interest (QoI) y that depends on a set of d uncertain design parameters $\boldsymbol{x} \in \mathbb{R}^d$ as

$$y = \mathcal{M}(\boldsymbol{x}), \tag{1}$$

where $\mathcal{M}$ generically denotes the computational model or code that is run in order to obtain y for a given configuration of the design parameters $\boldsymbol{x}$.

The PCE approximates the QoI in (1) by means of the expansion of suitable orthogonal polynomials, i.e., [2]

$$y \approx \hat{y} = \mathcal{M}_{\mathrm{PCE}}(\boldsymbol{x}) = \sum_{\boldsymbol{k} \in \mathcal{K}} c_{\boldsymbol{k}} \psi_{\boldsymbol{k}}(\boldsymbol{x}), \tag{2}$$

where $\psi_{\boldsymbol{k}}$ are multivariate basis functions constructed as the product of univariate polynomials, $c_{\boldsymbol{k}}$ are the corresponding coefficients, and $\mathcal{K} \subset \mathbb{N}^d$ is a set of multi-indices $\boldsymbol{k}$ defining the degree of the basis function in each dimension. Special orthogonal polynomials exist depending on the distribution of the parameters $\boldsymbol{x}$, e.g., Hermite polynomials for Gaussian variability. A popular definition for $\mathcal{K}$ is based on the so-called total degree truncation, meaning that the 1-norm of the multi-indices is bounded by a maximum order p (i.e., $\|\boldsymbol{k}\|_1 \leq p, \forall \boldsymbol{k} \in \mathcal{K}$), leading to $|\mathcal{K}| = (p+d)!/(p!d!)$ terms.

A peculiar property of PCEs is that statistical moments and sensitivity indices are analytically derived from the model coefficients. Indeed, the expectation and the variance of the QoI are given by

$$\mu_y = \mathrm{E}(y) \approx \mathrm{E}(\hat{y}) = c_{\boldsymbol{0}} \tag{3}$$

and

$$\sigma_y^2 = \mathrm{Var}(y) \approx \mathrm{Var}(\hat{y}) = \sum_{\boldsymbol{k} \in \mathcal{K} \setminus \boldsymbol{0}} c_{\boldsymbol{k}}^2, \tag{4}$$

respectively. The total Sobol' sensitivity indices, describing the individual impact of each random input on the variance of the QoI, are instead given by [2]

$$S_j = \frac{1}{\sigma_y^2} \sum_{\boldsymbol{k} \in \mathcal{K}_j} c_{\boldsymbol{k}}^2, \tag{5}$$

for $j = 1, \dots, d$, where $\mathcal{K}_j \subset \mathcal{K} = \{\boldsymbol{k} : k_j \neq 0\}$ is the subset of multi-indices with a non-zero j-th component.

Two important classes of methods are available to estimate the PCE coefficients and are discussed in the next sections.

A. Quadrature-Based Approaches

In these approaches, the projection integral that defines the rigorous calculation of the coefficients is approximated by means of a suitable quadrature rule, leading to

$$c_{\boldsymbol{k}} = \int_{\mathbb{R}^d} \mathcal{M}(\boldsymbol{x}) \psi_{\boldsymbol{k}}(\boldsymbol{x}) w(\boldsymbol{x}) d\boldsymbol{x} \approx \sum_{q=1}^{N_Q} y_q \psi_{\boldsymbol{k}}(\boldsymbol{x}_q) w_q, \tag{6}$$

where $\boldsymbol{x}_q$ are the quadrature nodes and w_q the corresponding weights, whereas $y_q = \mathcal{M}(\boldsymbol{x}_q)$ are the samples of the QoI (i.e., the model responses) evaluated at the quadrature nodes.

A tensor-product Gauss quadrature rule requires $N_Q = (p+1)^d$ nodes. Sparse grid quadratures, such as the one based on Smolyak' rule, are more parsimonious at the price of a reduced accuracy. In either case, the number of simulations required is much greater than the number of coefficients to be estimated, i.e., $N_Q \gg |\mathcal{K}|$, which makes these approaches inefficient when the underlying model (1) is expensive to simulate.

B. Regression-Based Approaches

These methods fit the model (2) in the least-square sense, i.e., they solve

$$\{c_{\boldsymbol{k}}\}_{\boldsymbol{k} \in \mathcal{K}} = \underset{c_{\boldsymbol{k}} \in \mathcal{K}}{\arg\min}\, \mathrm{E}\left(\left(\mathcal{M}(\boldsymbol{x}) - \sum_{\boldsymbol{k} \in \mathcal{K}} c_{\boldsymbol{k}} \psi_{\boldsymbol{k}}(\boldsymbol{x}) \right)^2 \right) \tag{7}$$

based on a set of data samples $\{(\boldsymbol{x}_l, y_l)\}_{l=1}^{N_L}$. For the ordinary regression, it must hold $N_L > |\mathcal{K}|$ in order for the regression problem to be overdetermined. This approach is typically more parsimonious compared to quadrature-based ones, but also less rigorous and more prone to larger errors due to its sensitivity to the specific choice of regression samples.

Least-angle regression (LAR) is a more efficient alternative to the ordinary least-square regression [2]. It identifies a sparse subset $\mathcal{A} \subseteq \mathcal{K}$ of basis functions, based on which it solves a modified version of the minimization problem (7) with a penalty proportional to the 1-norm of the coefficients. The penalty acts as a regularization term that favors low-rank solutions. This relaxes the requirement of having more regression samples than the (total) number of PCE coefficients. However, the number of estimated coefficients is still bounded by the number of regression samples (i.e., $|\mathcal{A}| < N_L$).

III. GPR

The fundamental assumption of GPR, also known as Kriging, is that the QoI y is a realization of a given *prior* Gaussian process with mean function (or trend) $\mu(\boldsymbol{x})$ and covariance function (or kernel) $k(\boldsymbol{x}, \boldsymbol{x}')$ [4]. Without loss of generality, we adopt the common assumption of a constant trend, i.e., $\mu(\boldsymbol{x}) = \beta_0$.

Based on a given set of N_L training data, the GPR identifies the prior realization that is the most consistent with the observation. The result is a *posterior* process, whose mean function provides the prediction at an arbitrary point:

$$\begin{aligned} y \approx \hat{y} = \mathcal{M}_{\mathrm{GPR}}(\boldsymbol{x}) &= \beta_0 + \boldsymbol{k}(\boldsymbol{x})^{\mathsf{T}} \boldsymbol{K}^{-1} (\boldsymbol{y} - \beta_0) \\ &= \beta_0 + \sum_{l=1}^{N_L} \alpha_l k(\boldsymbol{x}, \boldsymbol{x}_l) \end{aligned} \tag{8}$$

where

- $\boldsymbol{y} = (y_1, \dots, y_{N_L})^{\mathsf{T}}$ is the vector of the observations;
- $\boldsymbol{k}(\boldsymbol{x}) = (k(\boldsymbol{x}, \boldsymbol{x}_1), \dots, k(\boldsymbol{x}, \boldsymbol{x}_{N_L}))^{\mathsf{T}}$ is a vector of kernel functions centered at the training samples;
- $\boldsymbol{K}$ is the covariance matrix of the training samples, with entries $K_{lm} = k(\boldsymbol{x}_l, \boldsymbol{x}_m)$, for $l, m = 1, \dots, N_L$.

The vector of the α-coefficients is readily found as $\boldsymbol{\alpha} = \boldsymbol{K}^{-1}(\boldsymbol{y} - \beta_0)$.

Typically, GPR is able to achieve accurate predictions with a limited amount of training data [3]. Moreover, one of the peculiar and attractive features of GPR is that an estimate of the prediction uncertainty, due to the limited amount of observed data, is assigned to model predictions. Specifically, the covariance between two predictions $\hat{y}_i$ and $\hat{y}_j$, at points $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ respectively, is given by

$$\mathrm{cov}\,(\hat{y}_i, \hat{y}_j) = k(\boldsymbol{x}_i, \boldsymbol{x}_j) - \boldsymbol{k}(\boldsymbol{x}_i)^{\mathsf{T}} \boldsymbol{K}^{-1} \boldsymbol{k}(\boldsymbol{x}_j). \tag{9}$$

For a single point $\boldsymbol{x}_i = \boldsymbol{x}_j$, (9) yields the variance of the prediction. Hence, the GPR prediction is in fact a Gaussian random variable with mean and variance provided by (8) and (9), respectively, which allows estimating its confidence level. This contrasts with the classical PCE method, which provides a deterministic model with no associated uncertainty.

IV. Combined GPR and PCE Framework

The main limitation of the PCE method is that it requires a large number of data samples to (accurately) estimate the model coefficients. On the other hand, the GPR method is typically more parsimonious, yet it lacks interpretability in an UQ scenario. Statistical moments and their confidence levels can be obtained numerically by using GPR as a surrogate of the actual model (1) in a MC-like analysis [3]. However, this requires to evaluate the covariance (9) at a potentially large number of MC samples, which may become computationally prohibitive.

In this section, we propose an effective combination of the PCE and GPR methods. Basically, we use the GPR model to surrogate observations in the quadrature-based approach (6), i.e., we replace y_q with predictions $\hat{y}_q$ at the quadrature

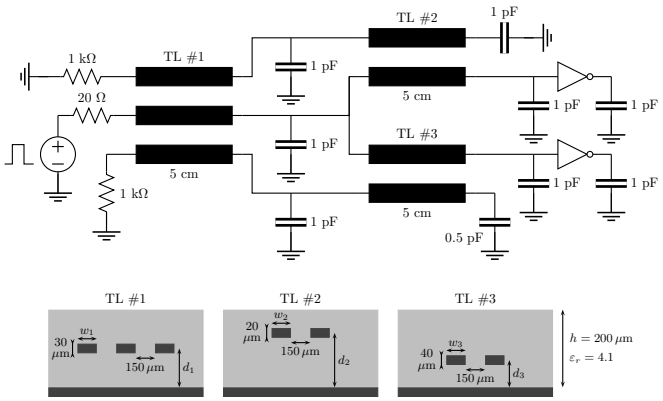

Fig. 1. Illustration of the PCB interconnect for the proposed application.

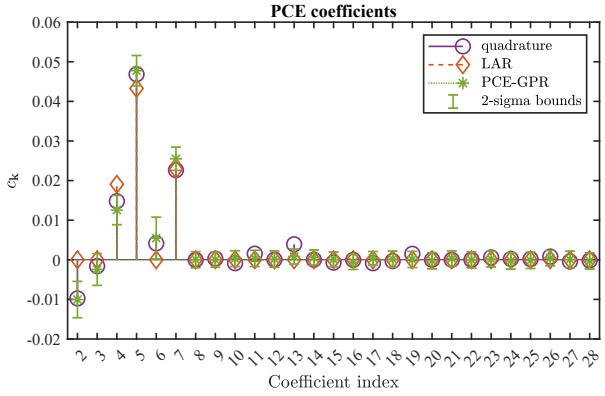

Fig. 2. PCE coefficients estimated with a quadrature-based approach (purple circles), LAR (red diamonds), and the proposed PCE-GPR method (green asterisks with confidence bars).

nodes obtained from (8). Since the GPR typically achieves an accurate model with a limited number of training data, this replacement incurs negligible accuracy loss, while evaluating (8) is inexpensive even for a large number of quadrature nodes. This leads to

$$c_{\mathbf{0}} = \beta_0 + \sum_{q=1}^{N_Q} \sum_{l=1}^{N_L} \alpha_l k(\boldsymbol{x}_q, \boldsymbol{x}_l) w_q \tag{10}$$

for the first coefficient (since $\psi_{\mathbf{0}} = 1$), and

$$c_{\boldsymbol{k}} = \sum_{q=1}^{N_Q} \sum_{l=1}^{N_L} \alpha_l k(\boldsymbol{x}_q, \boldsymbol{x}_l) \psi_{\boldsymbol{k}}(\boldsymbol{x}_q) w_q \tag{11}$$

for $\boldsymbol{k} \succ \mathbf{0}$.

Moreover, since (6) is in fact a linear combination of model predictions, the covariance (9) is readily propagated to obtain the covariance of the estimated PCE coefficients. This allows obtaining confidence levels for the estimated coefficients, and in turn also for the moments and sensitivity indices. Details are deferred to a future report.

Indeed, an important outcome of the outlined framework is that the abovementioned confidence information provides an indication of the accuracy of the estimated PCE coefficients due to the lack of data, which is missing in state-of-the-art PCE implementations. It should be noted the confidence level does not account for the error introduced by the quadrature. It is however reasonable to assume that this error is small compared to the overall model error.

V. Application Example and Numerical Results

The proposed framework is applied to the statistical analysis of the far-end crosstalk in the PCB interconnect of Fig. 1. The uncertainty is provided by $d = 6$ geometrical parameters, namely the width of the traces and their height above the ground plane in each of the three transmission line sections. The distribution is Gaussian with the nominal values $w_1 = 150\ \mu$m, $w_2 = 130\ \mu$m, $w_3 = 170\ \mu$m, $d_1 = 100\ \mu$m, $d_2 = 140\ \mu$m, $d_3 = 70\ \mu$m and a relative standard deviation of 10%. In particular, the QoI is taken as the maximum crosstalk occurring over time for a voltage pulse excitation with an amplitude of 5 V, rise/fall times of 100 ps, and a duration of 1 ns. The structure is implemented in HSPICE and a transient simulation is performed to obtain the maximum of the absolute value of the far-end crosstalk over time, i.e.,

$$y = \mathcal{M}(\boldsymbol{x}) = \max_t \left(|v_{FEXT}(t; \boldsymbol{x})| \right). \tag{12}$$

We compare four different methods:

1) A second-order PCE with coefficients computed by means of an accurate tensor-product Gauss quadrature, requiring simulations at $N_Q = 729$ nodes. The model has $|\mathcal{K}| = 28$ terms based on a total degree truncation.
2) An analogous PCE model with coefficients computed via the LAR. For this method, we use $N_L = 10$ training data, sampled from a Latin hypercube scheme.
3) A GPR model with a squared-exponential kernel function, trained with the same dataset as used for LAR.
4) The proposed hybrid PCE-GPR method, in which the GPR model of method 3) is combined with the quadrature rule of method 1) to obtain the PCE coefficients.

We use the UQLab toolbox [5] for methods 1) and 2) and the Statistics and Machine Learning Toolbox™ [6] in MATLAB® to train the GPR models in methods 3) and 4).

Figure 2 shows the PCE coefficients obtained via the accurate Gauss quadrature (purple circles) as well as with the LAR and the hybrid PCE-GPR method (red diamonds and green asterisks, respectively), both trained with the same 10 samples. The first coefficient (corresponding to the crosstalk mean) is dominant and is therefore omitted, but it is estimated with good accuracy by all methods. The 2-sigma confidence interval of the coefficients estimated with PCE-GPR is indicated by the green error bars.

This first comparison shows that the PCE-GPR method achieves good accuracy with 10 observations only. In particular, the estimate is generally better than LAR (cfr., e.g., the estimate of coefficients 2 to 6) using the same amount of data, and in good agreement with the quadrature-based approach, which however requires way more data samples. It is also interesting to note that the reference quadrature coefficients

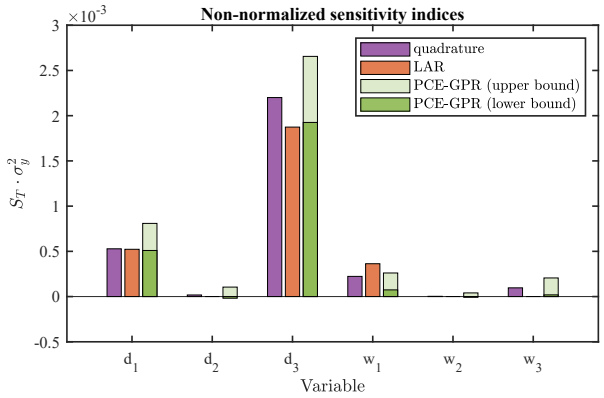

Fig. 3. Non-normalized Sobol' sensitivity indices describing the relative impact of each of the six uncertain parameters.

are mostly within the confidence interval of the PCE-GPR prediction, which therefore provides a good indication of the estimate uncertainty due to the limited amount of data.

The PCE-GPR method predicts a mean crosstalk of $\mu_y = 0.479$ V with a 2-sigma confidence interval of $[0.476, 0.481]$ V and a standard deviation of $\sigma_y = 0.0572$ V with a confidence interval of $[0.0535, 0.0607]$ V. The reference results from a Monte Carlo analysis with 5000 runs are $\mu_y = 0.476$ V and $\sigma_y = 0.0549$ V. These values lie within the confidence intervals of the PCE-GPR prediction.

Figure 3 shows the Sobol' sensitivity indices describing the relative impact of the six uncertain inputs on the variability of crosstalk. The purple and red bars refer to the results of the quadrature- and LAR-based methods in UQLab. The light and dark shades of green indicate instead the upper and lower 2-sigma bounds of the PCE-GPR prediction, respectively. For an easier propagation of the prediction uncertainty, the indices are not normalized by the variance of y. This does not alter the relative magnitude of the indices, since σ_y^2 is a common denominator for all.

The sensitivity analysis indicates that the trace-to-ground distance of the third transmission line section has the largest impact on the crosstalk, followed by the trace-to-ground distance of the first section. It is also reasonable that the parameters of the second section (w_2 and d_2), which lies in the opposite (upper) propagation branch, have negligible impact on the crosstalk in the lower branch.

Concerning the accuracy of the various methods, it is observed that the PCE-GPR method achieves both a good accuracy and a good estimation of the prediction confidence. Indeed, the reference result from the quadrature method always lies within the confidence interval, i.e., between the lighter and darker green bar. The confidence interval shrinks by increasing the number of training samples, yet the result is omitted due to the lack of space.

Finally, Table I reports the accuracy of the considered methods in terms of root-mean-square error (RMSE) between the model predictions and the reference results from the Monte Carlo analysis. We report results also for the GPR method alone. Furthermore, for the LAR and GPR-based methods, we consider training datasets of increasing size, i.e., with $N_L = \{10, 20, 30\}$ samples, and three independent runs for each dataset size.

TABLE I
RMSE FOR THE VARIOUS CONSIDERED METHODS.

Method	N	run #1	run #2	run #3
PCE-quad	729		2.533×10^{-3}	
PCE-LAR	10	1.260×10^{-2}	2.985×10^{-2}	2.346×10^{-2}
GPR alone	10	7.055×10^{-3}	7.768×10^{-3}	6.583×10^{-3}
PCE-GPR	10	7.026×10^{-3}	7.696×10^{-3}	6.474×10^{-3}
PCE-LAR	20	6.317×10^{-3}	6.628×10^{-3}	4.574×10^{-3}
GPR alone	20	5.756×10^{-3}	5.515×10^{-3}	3.870×10^{-3}
PCE-GPR	20	5.723×10^{-3}	5.499×10^{-3}	3.796×10^{-3}
PCE-LAR	30	4.769×10^{-3}	4.317×10^{-3}	4.003×10^{-3}
GPR alone	30	4.088×10^{-3}	4.493×10^{-3}	3.157×10^{-3}
PCE-GPR	30	4.063×10^{-3}	4.462×10^{-3}	3.101×10^{-3}

The above analysis allows drawing some interesting conclusions. As expected, the quadrature-based approach is the most accurate and achieves the lowest RMSE. However, it requires a number of data samples that is almost two orders of magnitude larger. The hybrid GPR-PCE method achieves almost always a lower RMSE compared to LAR (the only exception being run #2 for the dataset with $N_L = 30$ samples).

In terms of computational times, the surrogate models are trained within a few seconds, while most time is consumed during the acquisition of the data samples. This step takes 44 s for 30 samples (LAR and GPR), 799 s for 729 samples (quadrature), and 7690 s for 5000 samples (Monte Carlo).

VI. Conclusions

This paper proposed a hybrid PCE-GPR method. While the PCE formulation analytically provides statistical information such as moments and sensitivity indices, the GPR framework allows for an efficient training with limited data and provides confidence levels for the predicted statistics. The proposed method was successfully applied to the statistical investigation of crosstalk in a PCB interconnect, for which accurate results were achieved with 10 training simulations only.

References

[1] P. Manfredi and D. Vande Ginste, *Uncertainty Quantification of Electromagnetic Devices, Circuits, and Systems*. Stevenage, U.K.: IET, 2021, ch. Polynomial chaos based uncertainty quantification in electrical engineering: Theory, pp. 21–48.

[2] G. Blatman and B. Sudret, "Adaptive sparse polynomial chaos expansion based on least angle regression," *J. Comput. Phys.*, vol. 230, no. 6, pp. 2345–2367, Mar. 2011.

[3] P. Manfredi and R. Trinchero, "A probabilistic machine learning approach for the uncertainty quantification of electronic circuits based on Gaussian process regression," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 41, no. 8, pp. 2638–2651, Aug. 2021.

[4] C. E. Rasmussen and C. K. Williams, *Gaussian Processes for Machine Learning*. Springer, 2006, vol. 1.

[5] S. Marelli, N. Lüthen, and B. Sudret, "UQLab user manual – Polynomial chaos expansions," Chair of Risk, Safety and Uncertainty Quantification, ETH Zurich, Switzerland, Tech. Rep., 2022, report UQLab-V2.0-104.

[6] *Statistics and Machine Learning Toolbox, Version 12.1*. Natick, MA, USA: The MathWorks Inc., 2023. [Online]. Available: https://www.mathworks.com