IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received November 28, 2018, accepted December 8, 2018, date of publication December 20, 2018,
date of current version January 11, 2019.

Digital Object Identifier 10.1109/ACCESS.2018.2888903

Machine Learning and Uncertainty Quantification for Surrogate Models of Integrated Devices With a Large Number of Parameters

RICCARDO TRINCHERO[1], (Member, IEEE), MOURAD LARBI[2], HAKKI M. TORUN[2], (Student Member, IEEE), FLAVIO G. CANAVERO[1], (Fellow, IEEE), AND MADHAVAN SWAMINATHAN[2], (Fellow, IEEE)

[1]Department of Electronics and Telecommunications, Politecnico di Torino, Turin 10129, Italy
[2]School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA

Corresponding author: Riccardo Trinchero (riccardo.trinchero@polito.it)

This work was supported in part by the Joint Project for Internalization of Research 2018: Machine Learning to Improve the Reliability of Complex Systems and in part by the National Science Foundation through the Center for Advanced Electronics through Machine Learning (CAEML) under Grant CNS 16-24810.

ABSTRACT This paper deals with the application of the support vector machine (SVM) and the least-squares SVM regressions to the uncertainty quantification of complex systems with a high-dimensional parameter space. The above regression techniques are used to build accurate and compact surrogate models of the system responses from a limited set of training samples. The accuracy and the feasibility of the proposed modeling techniques are then investigated by comparing their results with the ones predicted by a sparse polynomial chaos expansion by considering two real-life problems with 8 and 30 random variables, respectively.

INDEX TERMS Machine learning, uncertainty quantification, parameterized modeling, surrogate models, SVM regression, LS-SVM regression, sparse PC expansion, integrated voltage regulator (IVR), wireless power transfer (WPT).

I. INTRODUCTION

Uncontrolled variation of system parameters due to manufacturing processes, tolerances and uncertain device characteristics can heavily affect the response of electrical and electronic circuits and systems. In order to guarantee the correct assessment of the system performance and to avoid expensive redesign, the impact of the above effects on the circuit responses is usually investigated during the early design phase via computationally expensive Monte Carlo (MC) simulations [1].

In modern electronic circuits and systems, the level of integration, the system complexity and the number of uncertain parameters is so high that the standard MC analysis is becoming challenging in terms of computer resources and simulation time. To this aim, in the last decades, a number of powerful techniques for the parametric and statistical analysis in complex nonlinear problems have been developed as alternatives to the *brute force* approach presented by MC simulation. Well-known examples are provided by parametric macromodeling [2]–[5], uncertainty analysis [6]–[12] and worst-case approaches [13]–[16]. Unfortunately, none of the available techniques provide an ultimate solution for the problem at hand, since they maximize different objectives w.r.t. the specific engineering application, e.g., the number of variables, their variability and their probability distributions.

Recently, the keyword *machine learning* has gained widespread reputation in different research areas for both classification and regression purposes [17]–[24]. Among the machine learning techniques, the support vector machine (SVM) regression [26]–[28] and its variant, namely, the least-squares support vector machine (LS-SVM) regression [29]–[31] can be considered as promising alternatives, which allow building compact parametric surrogates of the nonlinear system responses with several uncertain parameters [17], [32]–[36]. However, the benefits of the application of such techniques to realistic structures in a high-dimensional parameter space still has to be proven. A preliminary feasibility study has been recently proposed in [36]. Such work focuses on the application of the SVM regression with polynomial kernels to the uncertainty quantification and the parametric modeling of structures with a small number of

uniformly distributed random variables (i.e., 4–5 variables) characterized by a large variability. Also, [36] provides a comparison in terms of accuracy, convergence and robustness to noise between the SVM regression with polynomial kernel and the polynomial chaos (PC) expansion [37]–[43] with least squares regression [8].

The goal of this work is to investigate the application of both the standard SVM and the LS-SVM regressions with polynomial and Gaussian radial base function (RBF) kernels for the statistical analysis of realistic structures in high-dimensional input parameter space (i.e., up to 30 uncorrelated stochastic variables). The above regression techniques have been applied to generate several surrogate models trained with a limited set of simulation results. For the sake of completeness, the results of the above surrogate models are compared with the ones predicted by an advanced technique for uncertainty quantification, such as the sparse PC expansion [38], [39] for two different test-cases i.e., an integrated voltage regulator with 8 uniformly distributed parameters and a wireless power transfer application with 30 Gaussian parameters.

The remainder of this paper is organized as follows. Section II presents the mathematical background behind the SVM and the LS-SVM regressions. Section III provides a brief overview on the sparse PC expansion. Section IV investigates the accuracy of the LSV, the LS-SVM regression with polynomial and RBF kernels and the sparse PC expansion by considering two different realistic test-cases. Section V concludes the paper.

II. SVM & LS-SVM REGRESSION

This Section provides a complete overview of the mathematical framework behind the SVM and the LS-SVM regressions.

A. SVM REGRESSION

Let us consider the problem of approximating a set of L training data $\{(\mathbf{x}_i, y_i)\}_{i=1}^{L}$ provided by a generic nonlinear system $y = \mathcal{M}(\mathbf{x})$ with input parameters $\mathbf{x} = [x_1, \ldots, x_d] \in \mathbb{R}^d$ via the following nonlinear SVM regression [25], [26]:

$$\mathcal{M}_{SVM}(\mathbf{x}) = \langle \mathbf{w}, \Phi(\mathbf{x}) \rangle + b, \tag{1}$$

where $\Phi(\mathbf{x}) = [\phi_1(\mathbf{x}), \ldots, \phi_D(\mathbf{x})]$ is a nonlinear map $\Phi(\cdot) : \mathbb{R}^d \rightarrow \mathbb{R}^D$ which maps the parameter space of dimension d into the corresponding *feature space* of dimension D; $\mathbf{w} \in \mathbb{R}^D$ is a vector collecting the unknown coefficients of the nonlinear regression; $b \in \mathbb{R}$ is the bias term; $\langle \mathbf{w}, \Phi(\mathbf{x}) \rangle$ is defined as the inner product in $\mathbb{R}^D$.

It is important to remark that the dimensionality of the feature space (i.e., D) is defined by the nonlinear map $\Phi(\mathbf{x})$, and, therefore, it turns out to be independent from both the number of training pairs L and the number of system parameters d. Also, it should be noted that (1) is linear w.r.t the nonlinear transformation $\Phi(\mathbf{x})$.

The goal of the SVM regression is to find the best combination of the parameters $(\mathbf{w}, b)$ in (1) that minimizes the following risk function:

$$R_{emp}(\mathbf{w}, b) = \frac{1}{L} \sum_{i=1}^{L} |y_i - \mathcal{M}_{SVM}(\mathbf{x}_i)|_\varepsilon, \tag{2}$$

where $|\cdot|_\varepsilon$ is the so-called linear ε-intensive loss function [26] defined as:

$$|y_i - \mathcal{M}_{SVM}(\mathbf{x}_i)|_\varepsilon = \begin{cases} 0, & \text{if } |y_i - \mathcal{M}_{SVM}(\mathbf{x}_i)| \leq \varepsilon \\ |y_i - \mathcal{M}_{SVM}(\mathbf{x}_i)| - \varepsilon, & \text{otherwise.} \end{cases} \tag{3}$$

Minimizing the risk function (2) is equivalent to finding the best combination of the parameters $(\mathbf{w}, b)$ that minimizes the deviation of the model predictions from the training samples outside the ε-intensive zone. This can be done via the following optimization problem in the primal space, which can be written as [26]:

$$\begin{aligned} &\text{minimize } \frac{1}{2}\|\mathbf{w}\|^2 + C \sum_{i=1}^{L} (\xi_i + \xi_i^*) \\ &\text{subject to } \begin{cases} y_i - \langle \mathbf{w}, \Phi(\mathbf{x}_i) \rangle - b \leq \varepsilon + \xi_i \\ \langle \mathbf{w}, \Phi(\mathbf{x}_i) \rangle + b - y_i \leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0, \quad \text{for } i = 1, \ldots, L \end{cases} \end{aligned} \tag{4}$$

where ξ_i, ξ_i^* are the slack variables and C is a parameter, chosen by the user, which provides a trade-off between the accuracy of the model and its flatness [27]. Figure 1(a) provides a graphical interpretation of the above optimization problem. The underlying idea is to minimize the positive ξ_i and negative ξ_i^* deviations of the training samples which lay outside the ε-intensive zone (gray area), but at the same time maximizing the model flatness to avoid the overfitting problem.

The nonlinear optimization problem in (4) can be solved by minimizing the following Lagrangian function:

$$\begin{aligned} &\mathcal{L}(\mathbf{w}, b, \boldsymbol{\xi}, \boldsymbol{\xi}; \boldsymbol{\alpha}, \boldsymbol{\alpha}^*, \boldsymbol{\eta}, \boldsymbol{\eta}^*) \\ &= \frac{1}{2}\|\mathbf{w}\|^2 + C \sum_{i=1}^{L} (\xi_i + \xi_i^*) + \end{aligned}$$

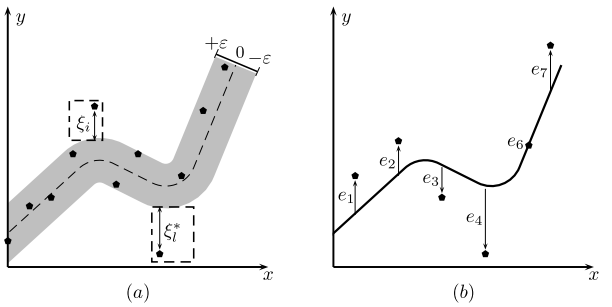

FIGURE 1. Panel (a): graphical interpretation of the SVM regression optimization problem in (4) (inspired by [27] and [33]). Panel (b): illustration of the corresponding least-square formulation in (9) for the LS-SVM regression (inspired by [29]).

$$- \sum_{i=1}^{L} \alpha_i(\varepsilon + \xi_i - y_i + \langle \mathbf{w}, \Phi(\mathbf{x}_i)\rangle + b) + \\ - \sum_{i=1}^{L} \alpha_i^*(\varepsilon + \xi_i^* + y_i - \langle \mathbf{w}, \Phi(\mathbf{x}_i)\rangle - b) + \\ - \sum_{i=1}^{L} (\eta_i \xi_i + \eta_i^* \xi_i^*), \tag{5}$$

where the parameters $\boldsymbol{\alpha}, \boldsymbol{\alpha}^*, \boldsymbol{\eta}, \boldsymbol{\eta}^* \geq 0$ are the Lagrangian multipliers (dual parameters) related to the constrained optimization problem. The saddle point of the Lagrangian (5) is obtained by computing the partial derivatives w.r.t. the primal variables:

$$\frac{\partial \mathcal{L}}{\partial \mathbf{w}} = 0 \rightarrow \mathbf{w} = \sum_{i=1}^{L} (\alpha_i - \alpha_i^*)\Phi(\mathbf{x}_i) \tag{6a}$$

$$\frac{\partial \mathcal{L}}{\partial b} = 0 \rightarrow \sum_{i=1}^{L} (\alpha_i - \alpha_i^*) = 0 \tag{6b}$$

$$\frac{\partial \mathcal{L}}{\partial \xi_i} = 0 \rightarrow C - \alpha_i - \eta_i = 0 \tag{6c}$$

$$\frac{\partial \mathcal{L}}{\partial \xi_i^*} = 0 \rightarrow C - \alpha_i^* - \eta_i^* = 0. \tag{6d}$$

Substituting (6a), (6b) and (6c) into (4) leads to the following dual optimization problem:

$$\text{maximize} \begin{cases} -\frac{1}{2} \sum_{i,j=1}^{L} (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*) K(\mathbf{x}_i, \mathbf{x}_j) + \\ -\varepsilon \sum_{i=1}^{L} (\alpha_i + \alpha_i^*) + \sum_{i=1}^{L} y_i(\alpha_i + \alpha_i^*) \end{cases}$$

$$\text{subject to} \begin{cases} \sum_{i=1}^{L} (\alpha_i - \alpha_i^*) = 0 \\ \alpha_i, \alpha_i^* \in [0, C], \end{cases} \tag{7}$$

where $K(\mathbf{x}_i, \mathbf{x}_j) = \langle \Phi(\mathbf{x}_i), \Phi(\mathbf{x}_j)\rangle$ is the kernel function defined as the inner product in the feature space between the function $\Phi(\mathbf{x})$ evaluated at the training samples $\mathbf{x}_i$ and $\mathbf{x}_j$.

Therefore, the nonlinear regression in (1) can be rewritten in the dual space as follows:

$$\mathcal{M}_{SVM}(\mathbf{x}) = \sum_{i=1}^{L} (\alpha_i - \alpha_i^*)\langle \Phi(\mathbf{x}_i), \Phi(\mathbf{x})\rangle + b \\ = \sum_{i=1}^{L} (\alpha_i - \alpha_i^*) K(\mathbf{x}_i, \mathbf{x}) + b. \tag{8}$$

The solution of the optimization problem in (7) allows estimating the parameters α_i and α_i^* in the above nonlinear regression. The bias term b is computed by exploiting the Karush-Kuhn-Tucker conditions [28].

It is important to remark that the direct computation of the inner product $\langle \Phi(\mathbf{x}_i), \Phi(\mathbf{x})\rangle$ can be extremely inefficient for a high-dimensional feature space (e.g., for the specific case of the Gaussian RBF kernel, D grows to infinity). However since both the dual optimization problem in (7) and the nonlinear regression in (8) do not require an explicit calculation of the inner product in the D-space, the latter can be efficiently estimated by operating directly in the parameter space with finite dimensionality (i.e., $\mathbb{R}^d$) via the kernel functions $K(\cdot, \cdot) : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$. This means that the SVM nonlinear regression is fully defined by the kernel function K without requiring an explicit definition of the nonlinear transformation Φ. This is the so-called *kernel trick*.

The most common kernels are listed below:

- linear: $K(\mathbf{x}_i, \mathbf{x}) = \mathbf{x}_i^T \mathbf{x}$;
- polynomial of order q: $K(\mathbf{x}_i, \mathbf{x}) = (1 + \mathbf{x}_i^T \mathbf{x})^q$;
- Gaussian radial basis function (RBF): $K(\mathbf{x}_i, \mathbf{x}) = \exp\left(-\|\mathbf{x}_i - \mathbf{x}\|^2 / 2\sigma^2\right)$.

The SVM regression algorithm is already available in MATLAB. A generic SVM regression can be suitably trained via the command `fitrsvm` and the resulting surrogate model can be evaluated for an arbitrary value of the input parameter $\mathbf{x}$ via the command `predict` [36].

B. LEAST-SQUARES SVM REGRESSION (LS-SVM)

The LS-SVM technique provides an alternative solution to the standard SVM regression which allows recasting the convex nonlinear optimization problem for the SVM regression in (1), in terms of a more standard least-squares formulation, without losing the advantages of the standard SVM regression [29]–[31].

First of all, the optimization problem in the primal space in (4) developed for the standard SVM regression has to be rewritten in the following form:

$$\text{minimize } \frac{1}{2}\|\mathbf{w}\|^2 + \gamma \frac{1}{2} \sum_{i=1}^{L} e_i^2 \\ \text{subject to } y_i = \langle \mathbf{w}, \Phi(\mathbf{x}_i)\rangle + b + e_i, \quad \text{for } i = 1, \ldots, L, \tag{9}$$

where $e_i \in \mathbb{R}$ are error variables and γ is an empirical parameter which provides a trade-off between the accuracy of the model and its flatness, playing the same role of the parameter C in the SVM primal optimization of (4).

The underlying idea is to minimize the squares of the regression errors e_i, which are defined as the deviation of the model prediction from the corresponding training samples as shown in Fig. 1(b), but at the same time by keeping the model as flat as possible, thus avoiding the oversampling problem. Also in this case the above optimization problem can be solved by minimizing the following Lagrangian:

$$\mathcal{L}(\mathbf{w}, b, \mathbf{e}; \boldsymbol{\alpha}) = \frac{1}{2}\|\mathbf{w}\|^2 + \gamma \frac{1}{2} \sum_{i=1}^{L} e_i^2 + \\ - \sum_{i=1}^{L} \alpha_i \{\langle \mathbf{w}, \Phi(\mathbf{x}_i)\rangle + b + e_i - y_i\} \tag{10}$$

where $\alpha_k \geq 0$ are the Lagrange multipliers.

By computing the partial derivatives of the above Lagrangian leads to:

$$\frac{\partial \mathcal{L}}{\partial \mathbf{w}} = 0 \rightarrow \mathbf{w} = \sum_{i=1}^{L} \alpha_i \Phi(\mathbf{x}_i) \tag{11a}$$

$$\frac{\partial \mathcal{L}}{\partial b} = 0 \rightarrow \sum_{i=1}^{L} \alpha_i = 0 \tag{11b}$$

$$\frac{\partial \mathcal{L}}{\partial e_i} = 0 \rightarrow \alpha_i = \gamma e_i \tag{11c}$$

$$\frac{\partial \mathcal{L}}{\partial \alpha_k} = 0 \rightarrow \langle \mathbf{w}, \Phi(\mathbf{x}_i) \rangle + b + e_i - y_i = 0. \tag{11d}$$

By substituting (11a), (11b) and (11c) into (11d) we can write following system of equations in the dual space:

$$\begin{cases} \sum_{i=1}^{L} \alpha_i = 0 \\ \sum_{i=1}^{L} \alpha_i K(\mathbf{x}_i, \mathbf{x}_1) + b + \gamma \alpha_1 - y_1 = 0 \\ \vdots \\ \sum_{i=1}^{L} \alpha_i K(\mathbf{x}_i, \mathbf{x}_L) + b + \gamma \alpha_L - y_L = 0 \end{cases} \tag{12}$$

where again $K(\mathbf{x}_i, \mathbf{x}_j) = \langle \Phi(\mathbf{x}_i), \Phi(\mathbf{x}_j) \rangle$ is the kernel function previously defined for the case of the SVM regression.

The linear system of equations in (12) can be recasted via the following matrix formulation:

$$\begin{bmatrix} 0 & \mathbf{1}^T \\ \mathbf{1} & \Omega + \mathbf{I}/\gamma \end{bmatrix} \begin{bmatrix} b \\ \boldsymbol{\alpha} \end{bmatrix} = \begin{bmatrix} 0 \\ \mathbf{y} \end{bmatrix} \tag{13}$$

where $\boldsymbol{\alpha} = [\alpha_1, \ldots, \alpha_L]^T$, $\mathbf{y} = [y_1, \ldots, y_L]^T$, $\mathbf{1}^T = [1, \ldots, 1] \in \mathbb{R}^{1 \times L}$, $\boldsymbol{I} \in \mathbb{R}^{L \times L}$ is the identity matrix and $\Omega \in \mathbb{R}^{L \times L}$ is the kernel matrix for which the element $\Omega_{ij} = K(\mathbf{x}_i, \mathbf{x}_j)$ for any $i, j = 1, \ldots, L$.

The solution of the square linear system of equations in (13) for $\boldsymbol{\alpha}$ and b leads to the following nonlinear regression in the dual space:

$$\mathcal{M}_{LS\text{-}SVM}(\mathbf{x}) = \textstyle\sum_{i=1}^{L} \alpha_i K(\mathbf{x}_i, \mathbf{x}) + b. \tag{14}$$

The LS-SVM regression is already implemented within LS-SVMLab Toolbox version 1.8 [44], which is fully compatible with the MATLAB environment.

III. POLYNOMIAL CHAOS EXPANSION

This section presents a quick overview on the mathematical framework behind two state-of-the-art techniques, the standard PC expansion and the advanced sparse PC expansion, respectively. The latter will be used hereafter in this paper as a reference approach for uncertainty quantification in high-dimensional problems.

A. STANDARD PC APPROACH

The discussion starts by considering the vector $\mathbf{x} \in \mathbb{R}^d$ collecting d independent random variables $(x_1, \ldots, x_d)$ with a joint probability density function (PDF) $f_{\mathbf{x}}(\boldsymbol{x})$, representing the uncertain input parameters of the physical problem. Assume that the problem is described by a numerical model $\mathcal{M}$, i.e., $y = \mathcal{M}(\mathbf{x})$, where y is the model response supposed to be a scalar quantity with a finite variance.

The PC expansion of the model response may be written as [37]:

$$y = \mathcal{M}_{PC}(\mathbf{x}) = \sum_{\boldsymbol{\lambda} \in \mathbb{N}^d} a_{\boldsymbol{\lambda}} \Psi_{\boldsymbol{\lambda}}(\mathbf{x}) \tag{15}$$

where $a_{\boldsymbol{\lambda}}$ are the unknown deterministic coefficients and $\Psi_{\boldsymbol{\lambda}}$ are the multivariate polynomials which are orthonormal w.r.t. the joint PDF $f_{\mathbf{x}}(\boldsymbol{x})$, i.e., $\mathbb{E}\left[\Psi_{\boldsymbol{\lambda}}(\mathbf{x})\Psi_{\boldsymbol{\beta}}(\mathbf{x})\right] = \delta_{\boldsymbol{\lambda}\boldsymbol{\beta}}$ and $\delta_{\boldsymbol{\lambda}\boldsymbol{\beta}} = 1$, if $\boldsymbol{\lambda} = \boldsymbol{\beta}$ and 0 otherwise.

Let $\mathcal{X} = \{\mathbf{x}_1, \ldots, \mathbf{x}_L\}$ be an experimental design (ED) of $\mathbf{x}$, and $\mathcal{Y} = \{\mathcal{M}(\mathbf{x}_1), \ldots, \mathcal{M}(\mathbf{x}_L)\}$ be the associated set of model response quantities. The PC coefficients $a_{\boldsymbol{\lambda}}$ may be estimated by using a non-intrusive technique such as the ordinary least square regression [38] from a set of model evaluations $\mathcal{Y}$. This technique relies on the choice of a *truncation set* denoted $\mathcal{A} = \{\boldsymbol{\lambda}_0, \ldots, \boldsymbol{\lambda}_{h-1}\} \subset \mathbb{N}^d$ defining the multi-indices of the selected basis polynomials $\{\Psi_{\boldsymbol{\lambda}_0}, \ldots, \Psi_{\boldsymbol{\lambda}_{h-1}}\}$. The PC expansion is usually truncated in order to preserve the polynomials of the basis whose total degree is less than or equal h, defined by the standard truncation set $\mathcal{A}^{d,h} = \{\boldsymbol{\lambda} \in \mathbb{N}^d : \|\boldsymbol{\lambda}\|_1 = \sum_{i=1}^{d} \lambda_i \leq h\}$. The number of coefficients retained by this strategy is $H = \frac{(d+h)!}{d!h!}$. We then notice that the number of terms $a_{\boldsymbol{\lambda}}$ to estimate blows up for large number of input random variables d and high degree h. In order to overcome this limitation, an advanced technique for reducing the number of polynomial bases is now introduced.

B. REDUCTION OF THE PC BASIS THROUGH A SPARSE APPROACH

As mentioned previously, the size of the PC expansion retained in the truncation set $\mathcal{A}^{d,h}$ can be too large when dealing with high-dimensional problems. In order to overcome the above issue, a hyperbolic truncation strategy based on a parameter k, with $0 < k \leq 1$, allowing to reduce the size of the PC basis is then introduced as follows:

$$\mathcal{A}_k^{d,h} = \{\boldsymbol{\lambda} \in \mathbb{N}^d : \|\boldsymbol{\lambda}\|_k = \left(\sum_{i=1}^{d} \lambda_i^k\right)^{1/k} \leq h\}. \tag{16}$$

This hyperbolic truncation strategy favors the most relevant effects and low-order interactions, which are known to have the largest impact on the variability of the response according to the *sparsity-of-effects principle* [49]. It is important to point out that lower values of k imply a larger number of neglected high-rank interactions. In addition, when $k = 1$, this scheme is equivalent to the standard PC expansion, which is defined by the truncation set $\mathcal{A}^{d,h}$. When $k < 1$,

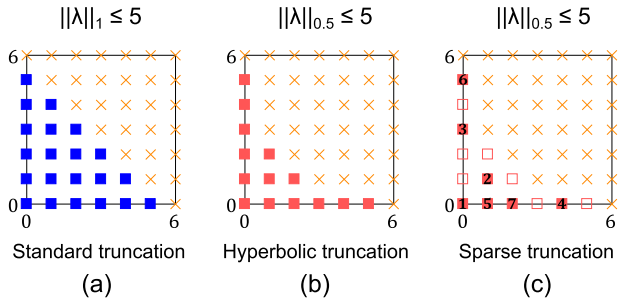

FIGURE 2. Number of terms of the polynomial basis of degree less or equal to the degree $h = 5$ retained by the hyperbolic truncation strategy when (a) k = 1 (squares) and (b) k = 0.5 (squares). (c) Numbered squares are the polynomial basis terms selected by the LARS algorithm from the 1^{st} to the 7^{th} iteration [39].

the retained terms of the polynomial basis can be substantially reduced as compared to H [39]. This improved truncation scheme is represented for two input random variables ($d = 2$) in Fig. 2(a) and 2(b), with the squares illustrating all terms of the polynomial basis of degree less than or equal to $l = 5$, included in the set (16) for $k = 1$ and $k = 0.5$, respectively. As a result, Fig. 2(b) shows that, for $k = 0.5$, this scheme chooses a number of polynomials (squares) smaller than those selected from a standard truncation set $\mathcal{A}^{M,l}$ (squares) as in Fig. 2(a).

Given a truncation set $\mathcal{A}^{M,h,k}$ of cardinal K, the hyperbolic truncation strategy enables to reduce the number of coefficients to be estimated in the PC expansion. However, this may remain too expensive in terms of model evaluations when the number of input random variables is large. For this reason, the number of elements of the polynomial basis may still be decreased by using a variable selection algorithm, such as the LARS algorithm [41].

The concept of the LARS algorithm is summarized in the following paragraph, while the reader may refer to [39] and [41] for additional details. Based on an iterative approach, LARS builds up a sparse approximation including from 1 to K terms of the polynomial basis (from one to all the squares in Fig. 2(b)), according to their decreasing impact. It starts by selecting the polynomial basis $\Psi_{\boldsymbol{\lambda}_1}$ which is better correlated with the model response y. In practice, the correlation is evaluated from a set of realizations (i.e., an ED) of the response $\mathcal{Y}$. After the identification of the first polynomial $\Psi_{\boldsymbol{\lambda}_1}$, the associated coefficient is estimated so that the residual $y - a^{(1)}_{\boldsymbol{\lambda}_1}\Psi_{\boldsymbol{\lambda}_1}(\mathbf{x})$ becomes equi-correlated with another polynomial basis, defined as being $\Psi_{\boldsymbol{\lambda}_2}$. Afterwards, the selection of a third polynomial basis will be performed by moving forward the direction $(\Psi_{\boldsymbol{\lambda}_1} + \Psi_{\boldsymbol{\lambda}_2})$ up to the new residual becomes equi-correlated with a third polynomial basis $\Psi_{\boldsymbol{\lambda}_3}$, and so on. An illustration of the selected polynomials by LARS after seven iterations is given by squares in Fig. 2(c).

The estimation of suitable terms of the polynomial basis by means of LARS is carried out for each degree $h = 1, 2, \ldots, h_{\max}$. In the end, LARS produces a set of sparse expansions including an increasing number of polynomial elements. The quality of each expansion of degree h is then evaluated according to a leave-one-out cross validation error ϵ_{LOO} as follows:

$$\epsilon_{LOO} = \frac{\sum_{i=1}^{N}\left(\mathcal{M}(\mathbf{x}_i) - \mathcal{M}^{PC}_{-i}(\mathbf{x}_i)\right)^2}{\sum_{i=1}^{N}\left(\mathcal{M}(\mathbf{x}_i) - \frac{1}{N}\sum_{i=1}^{N}\mathcal{M}(\mathbf{x}_i)\right)^2}, \tag{17}$$

where $\mathcal{M}^{PC}_{-i}(\mathbf{x}_i)$ are N surrogate models built up on the ED $\boldsymbol{\mathcal{X}}$ such that $\mathbf{x}_i = \{\mathbf{x}_q, q = 1, \ldots, N,\ q \neq i\}$. The retained degree h for the sparse PC expansion is the one minimizing the leave-one-out error ϵ_{LOO}.

In the following, the quality of the surrogate model is computed via the Q^2 coefficient defined by $Q^2 = 1 - \epsilon_{LOO}$, $0 \leq Q^2 \leq 1$. It is important to note that the larger Q^2 is, the better is the prediction of the surrogate model.

In order to identify quantities of interest of the model response, a post-processing of the coefficients of the basis can be performed at a relatively low computational cost. Indeed, the orthonormality property of the polynomial basis allows to estimate the expectation and the variance [43] of the output y as:

$$\mathbb{E}[y] = a_{\mathbf{0}}, \tag{18}$$

and

$$\mathbb{V}[y] = \sum_{\boldsymbol{\lambda}\in\mathcal{A}\setminus\{\mathbf{0}\}} a^2_{\boldsymbol{\lambda}}. \tag{19}$$

IV. APPLICATION TEST-CASES

In this section the accuracy and the feasibility of the techniques presented in Sec. II and Sec. III have been investigated by considering two realistic test-cases: an integrated voltage regulator (IVR) and a wireless power transfer (WPT) application, respectively. All the simulations are performed on a Dell PC with an Intel(R) Core(TM) i7-7700M, CPU running at 3.60 GHz and 16 GB of RAM.

A. EXAMPLE 1: EMBEDDED INDUCTOR AND INTEGRATED VOLTAGE REGULATOR

As a first test-case, the SVM, the LS-SVM with RBF and polynomial kernels and the sparse PC expansion have been adopted to quantify the impact of 8 uncertain parameters of an embedded inductor on the IVR efficiency while the converter operates in pulse-width modulation (PWM) mode at 100 MHz.

The considered architecture is shown in Fig. 3. It consists of a system-in-package (SiP) solution including two chips (buck converter and low-dropout (LDO)/load) with an integrated inductor on an organic package [45]. The converter architecture shown in Fig. 3 is designed with stacked topology, using 130 nm GLOBALFOUNDRIES (GF) process and consists of four-phases (one master-three slaves). The structure of the inductor is a solenoid with Nickel-Zinc (NiZn) ferrite magnetic core as shown in Fig. 4(a) and (b).

The converter efficiency has been accurately estimated via an extensive model that accounts for switching and conduction losses of power switches, DC and AC losses of inductor, power delivery network (PDN) and output capacitance.

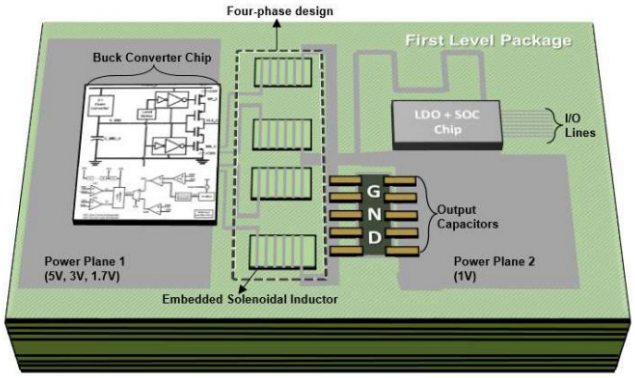

FIGURE 3. Illustration of the two-chip SiP IVR architecture considered in Sec. IV-A.

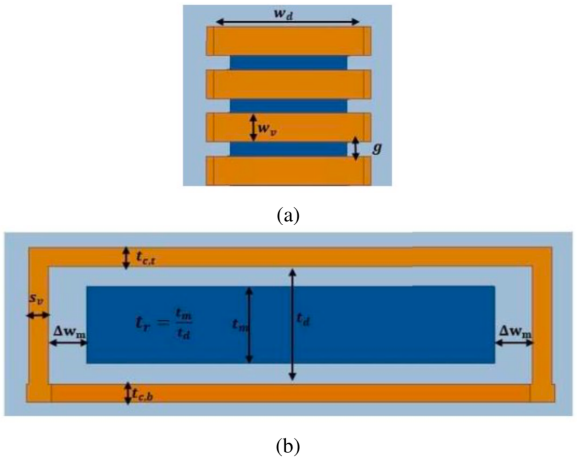

FIGURE 4. Top view (panel (a)) and side view (panel (b)) of the geometrical parameters of solenoidal inductor with magnetic core [23].

A more detailed description of the buck converter topology along with the model verification are provided in [45]. In order to quantify the effect of possible geometrical tolerances on the inductor electromagnetic behavior and thus on the IVR efficiency, 8 geometrical parameters have been considered as uniform random variables as shown in Table 1. The number of windings and the magnetic core thickness ratio have been fixed to $N_w = 6$ and $t_m = 0.9$, respectively. To account for proximity and skin effect as well as demagnetization effect caused by the magnetic core, the embedded inductor is characterized using the full-wave solver of Ansys HFSS [47]. Additional details on the inductor simulation framework adopted in this paper can be found in [23].

The SVM and the LS-SVM regressions with polynomial kernels of order from 1 to 3 and RBF kernel are adopted to generate 8 different surrogate models. In addition, a sparse PC expansion is constructed with an adaptive degree varying from 1 to 10. The hyperbolic truncation scheme in (16) is set to $k = 0.75$ to reduce the size of the polynomial basis [39], [40].

All surrogate models have been trained from $L = 200$ samples based on latin hypercube sampling (LHS) [48] resulting

TABLE 1. Uncertain Geometrical Parameters of the Solenoidal Inductor used for the IVR in Sec. IV-A.

Uniform random variables		Unit	$\mathcal{U}\,[Min; Max]$
Gap between windings	g	mil	$\mathcal{U}\,[4; 6]$
Size of via	s_v	μm	$\mathcal{U}\,[80; 120]$
Copper Trace Width	w_v	mil	$\mathcal{U}\,[9; 11]$
Copper Thickness Bottom	$t_{c,b}$	μm	$\mathcal{U}\,[64; 96]$
Copper Thickness Top	$t_{c,t}$	μm	$\mathcal{U}\,[64; 96]$
Dielectric Thickness	t_d	μm	$\mathcal{U}\,[180; 220]$
Dielectric Width	w_d	mil	$\mathcal{U}\,[59.4; 60.6]$
Magnetic Core Width offset	Δw_m	mil	$\mathcal{U}\,[9; 11]$

TABLE 2. Comparison on the accuracy and the computational cost of the SVM, LS-SVM and PC surrogates computed for the IVR in Sec. IV-A

Method	Kernel Regression	RMSE	$\hat{\mu}$	$\hat{\sigma}$	t_{model}	t_{cost}
MC	–	–	67.01	0.31	–	7 days
SVM	Linear	0.159	67.02	0.28	<1s	2.4s
	Poly Order 2	0.179	67.04	0.30	<1s	2.9s
	Poly Order 3	0.428	67.00	0.52	1.2s	4.5s
	RBF	0.166	67.04	0.28	<1s	1.8s
LS-SVM	Linear	0.158	67.03	0.28	<1s	<1s
	Poly Order 2	0.166	67.04	0.29	<1s	<1s
	Poly Order 3	0.443	67.00	0.54	<1s	<1s
	RBF	**0.155**	67.02	0.28	1.4s	<1s
Sparse PC	Poly Order 9	0.170	67.02	0.28	<1s	<1s

from the HFSS simulations with a computational cost of 3 h 27 min (one simulation with the full-wave solver of Ansys HFSS model takes about 1 min). From the sparse PC, the Q^2 coefficient is equal to 0.83, emphasizing a surrogate model with a high level of accuracy. In order to investigate the performance of the obtained surrogate models, their predictions are then compared with the results of a MC simulation with 10000 samples. Table 2 provides a detailed comparison on the accuracy and the computational cost of the proposed modeling techniques by collecting the root mean square error (RMSE), the mean value $\hat{\mu}$ and the standard deviation $\hat{\sigma}$ estimated by the proposed surrogate models, along with the corresponding computational times t_{model} and t_{cost} required to build each surrogate models and to evaluate them to compute 10000 predictions of the output, respectively.

Table 2 shows that the LS-SVM regression with the RBF kernel turns out to be the most accurate surrogate model with a RMSE of 0.1552, which is slightly better than the accuracy obtained by the linear regression with the LS-SVM (i.e., RMSE = 0.1580) and SVM (i.e., RMSE = 0.1585) regressions. However, a good accuracy is also achieved by the sparse PC expansion, i.e., RMSE = 0.1696. The results of the SVM and the LS-SVM regressions with polynomial kernel highlight that for the specific example at hand, a linear

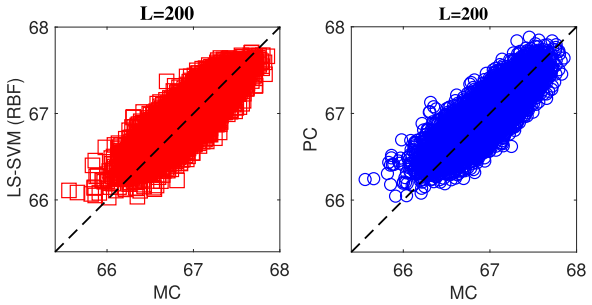

FIGURE 5. Scatter plots of the IVR efficiency providing a comparison between the predictions of the LS-SVM regression with RBF kernel (red marker) and the sparse PC expansion (blue marker) and the results of a MC simulation with 10000 samples.

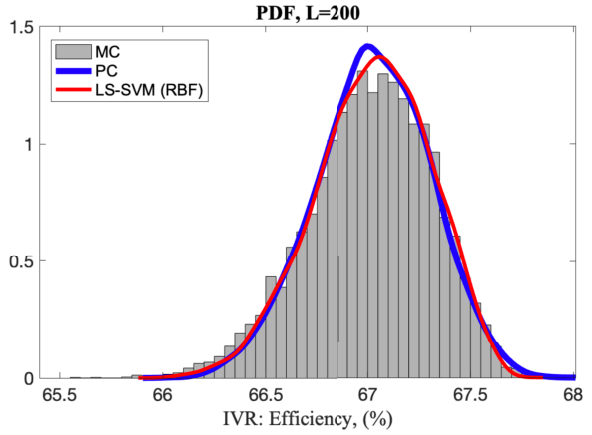

FIGURE 6. PDFs of the IVR efficiency obtained from the LS-SVM regression with RBF kernel (solid red curve) and sparse PC expansion (solid blue curve) compared with the histogram of 10000 MC samples (gray bins).

expansion is enough to accurately reproduce the actual behavior of the IVR efficiency as a function of the considered input variables. In fact, the RMSE of the surrogate models is not improved by increasing the order of the polynomial kernel.

As an illustration, Fig. 5 provides the scatter plots for the LS-SVM with RBF kernel and sparse PC surrogate models. The plots emphasize the good accuracy between the two surrogate models and the MC simulation as the samples are very close to the dashed lines, which represent a perfect agreement between the surrogate models and the reference samples. The impact of the variability on the IVR efficiency is illustrated in Fig. 6, where the PDFs estimated via the LS-SVM based surrogate model with RBF kernel and the sparse PC expansion are compared with the histogram of 10000 MC samples. We see that the main variability of the IVR efficiency is well captured by both surrogate models for which the mean values and the standard deviations are quite close to those calculated by MC simulation. This confirms a good estimation of the PDF of the IVR efficiency with both surrogate models, highlighting a similar level of accuracy. In terms of computational cost, 10000 MC simulations required about 7 days, while the LS-SVM with RBF kernel and the sparse PC needed less than 1 s, as shown in Table 2.

TABLE 3. Uncertain Parameters considered for the WPT in Sec. IV-B.

Gaussian Random Variables		Unit	μ	σ
Feeding Gap for TX coil	$g_{f,TX}$	mm	2.46	0.1230
Feeding Gap for RX coil	$g_{f,RX}$	mm	2.3	0.1150
Width & Height of TX coil	$g_{\{x,y\},TX}$	mm	2.1	0.1050
Width & Height of RX coil	$g_{\{x,y\},RX}$	mm	3	0.1500
GND Cut-Out Ratio TX	$slot_{\{x,y\},TX}$	–	1.19	0.0595
GND Cut-Out Ratio RX	$slot_{\{x,y\},RX}$	–	1.16	0.0580
Match Capacitor TX	C_1	pF	4.09	0.2045
Resonance Capacitor TX	C_2	pF	4.59	0.2295
Resonance Capacitor RX	C_3	pF	2.16	0.1080
Match Capacitor RX	C_4	pF	0.3	0.015
Match inductor TX	L_1	nH	1.25	0.0625
Match inductor RX	L_2	nH	7.82	0.3910
Line Width TX Coil	$l_{w,TX}$	mm	1.89	0.0945
Line Width RX Coil	$l_{w,RX}$	mm	0.71	0.0355
Width TL1	w_{TL1}	mm	0.381	0.0191
Width TL2	w_{TL2}	mm	0.407	0.0204
Width TL3	w_{TL3}	mm	0.386	0.0193
Width TL4	w_{TL4}	mm	0.528	0.0264
Width TL5	w_{TL5}	mm	0.521	0.0191
Width TL6	w_{TL6}	mm	0.397	0.0198
Width TL7	w_{TL7}	mm	0.607	0.0303
Width TL8	w_{TL8}	mm	0.520	0.0260
Length TL1	l_{TL1}	mm	1.656	0.0828
Length TL2	l_{TL2}	mm	1.424	0.0712
Length TL3	l_{TL3}	mm	0.723	0.0362
Length TL4	l_{TL4}	mm	2.153	0.0828
Length TL5	l_{TL5}	mm	0.612	0.1076
Length TL6	l_{TL6}	mm	1.666	0.0833
Length TL7	l_{TL7}	mm	0.502	0.0251
Length TL8	l_{TL8}	mm	0.555	0.0278

It is worth noting that this computational cost does not include the training time required to generate the $L = 200$ samples from LHS needed for constructing the surrogate models.

B. EXAMPLE 2: WIRELESS POWER TRANSFER

In a second and more complex example, the modeling approaches presented in Sec. III and Sec. II have been applied to predict the impact of 30 uncertain parameters on the maximum efficiency of a wireless power transfer (WPT) application in the bandwidth 500 MHz to 1.5 GHz.

The WPT schematic is shown in Fig. 7. It consists of an integrated board architecture with embedded rectangular RF

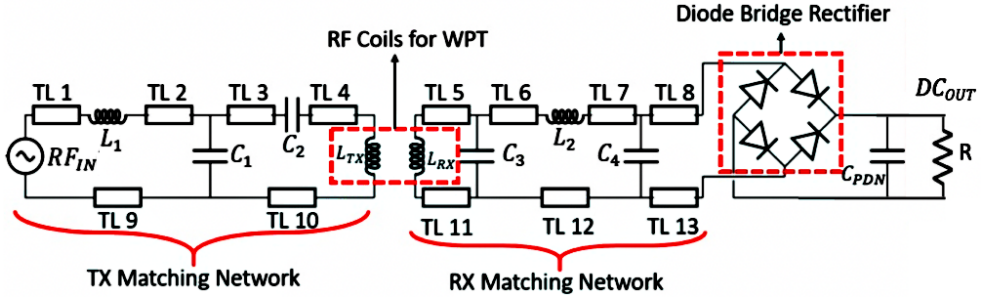

FIGURE 7. Schematic of the WPT based power delivery architecture considered in Sec. IV-B.

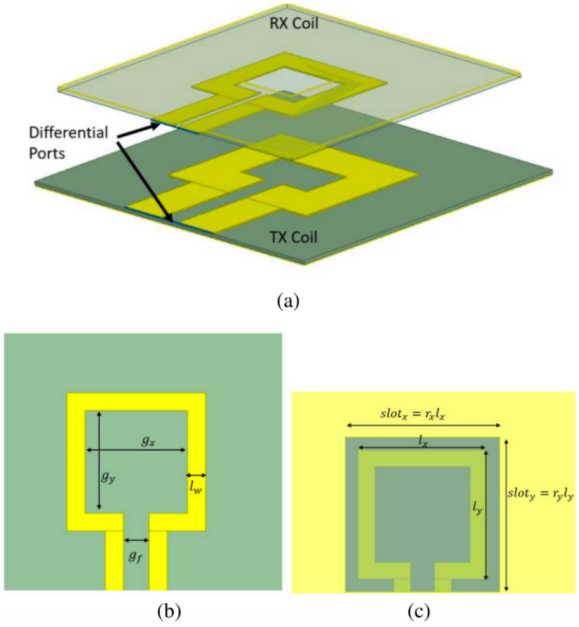

FIGURE 8. TX and RX coils for the WPT based power delivery architecture in the schematic of Fig. 7, with their main geometrical parameters [21]. Panel (a): WPT structure; Panel (b): top view; Panel (c): bottom view.

coils, shown in Fig 8, connected to a TX and RX matching networks and a full bridge diode rectifier loaded by the parallel between a capacitor $C_{PDN} = 1$ nF and a resistor $R = 868\,\Omega$. The transmission distance between the TX and RX coil is fixed to 1 mm. The matching networks on both TX and RX modules consist of a LC network and a resonating capacitor connected with the RF coils. In order to maximize the magnetic flux generated by the TX coil and the power delivered to the load, the resonant capacitors are connected in series and in shunt with the TX and RX coils as suggested in [46]. Additional details on the parameters selection are available in [21].

The effect of the fabrication tolerances and uncertainties on the efficiency of the WPT is examined through 30 Gaussian random variables, given in Table 3. Each parameter has a nominal value and a standard deviation of 5%. Similar to the previous example, different surrogate models have been generated by SVM and LS-SVM regressions with polynomial kernels of degree from 1 to 4 and RBF kernel. A sparse PC surrogate model has also been generated based on an adaptive degree varying from 1 to 10 and a hyperbolic truncation scheme fixed by $k = 0.75$.

The mentioned surrogates have been trained with an increasing number of training samples L generated from LHS scheme. Table 4 provides, for 10000 samples, a comparison of the performances of the resulting 11 surrogate models by considering the RMSE, the mean values $\hat{\mu}$ and the standard deviations $\hat{\sigma}$ for $L = 200, 400$ and 600 computed w.r.t. the results of a MC simulation. Moreover, Table 4 provides, w.r.t the several sizes of training data (i.e. $L = 200, 400, 600$ samples), the computational times t_{model} and t_{cost} required to construct each surrogate model and to evaluate them for the predictions of 10000 realizations of the output, respectively. It is important to note that the generation of the training data with $L = 200, 400, 600$ samples took about 5 h, 10 h and 15 h, respectively (one single simulation with the full-wave solver of Ansys HFSS model takes about 1 min 30 s).

From the results collected in Table 4, the LS-SVM regression with RBF kernel is confirmed to be the most accurate model with a $RMSE = 0.4307$ for $L = 600$ samples, which is slightly better than the error obtained by the sparse PC expansion (i.e., $RMSE = 0.4385$). A remarkable accuracy is also achieved by the LS-SVM regression with polynomial kernel of order 4, for which $RMSE = 0.4798$. The results of Table 4 also highlight the improved performance of the SVM-based approach in terms of RMSE, for $L = 200$ and 400 samples. This is in line with the observations reported in [36].

As a further validation, Fig. 9 shows the scatter plots computed, for 10000 samples, by comparing the predictions of the LS-SVM regression with RBF kernel and the sparse PC expansion surrogates, generated for an increasing number of training samples (i.e., $L = 200, 400, 600$), with the results of a MC simulation. The scatter plots confirm the results collected in Table 4. Indeed, due to the strong nonlinearity and high-dimensionality of the problem at hand, the accuracy of the predictions obtained from the models generated via $L = 200$ samples is rather poor and turns out to be inadequate for our modeling purposes. For the specific case of the surrogate model based on the LS-SVM regression, the resulting accuracy is slightly improved for $L = 400$ samples. Finally, an acceptable accuracy has been achieved

TABLE 4. Comparison on the accuracy and the computational cost of the SVM, LS-SVM and PC surrogates computed for an increasing number of training samples L for the WPT in Sec. IV-B.

Method	Kernel Regression	$L = 200$ (cost=5h)					$L = 400$ (cost=10h)					$L = 600$ (cost=15h)				
		RMSE	$\hat{\mu}$	$\hat{\sigma}$	t_{model}	t_{cost}	RMSE	$\hat{\mu}$	$\hat{\sigma}$	t_{model}	t_{cost}	RMSE	$\hat{\mu}$	$\hat{\sigma}$	t_{model}	t_{cost}
MC	–	–	74.19	0.68	–	11 days	–	74.19	0.68	–	11 days	–	74.19	0.68	–	11 days
SVM	Linear	0.543	74.29	0.55	<1s	3.0s	0.523	74.27	0.45	<1s	2.8s	0.526	74.27	0.50	1.2s	3.2s
	Poly Order 2	0.608	74.25	0.49	<1s	3.3s	0.876	74.26	0.92	<1s	3.4s	0.990	74.25	1.09	18.9s	3.6s
	Poly Order 3	0.573	74.21	0.28	6.1s	4.1s	0.561	74.27	0.35	13.2s	4.8s	0.541	74.24	0.43	20.0s	5.4s
	Poly Order 4	>1	73.32	>1	7.1s	4.0s	0.633	74.26	0.15	12.4s	8.0s	0.605	74.22	0.20	19.8s	8.7s
	RBF	0.524	74.28	0.40	<1s	2.3s	0.506	74.25	0.39	<1s	2.7s	0.495	74.25	0.42	<1s	3.4s
LS-SVM	Linear	0.537	74.29	0.52	<1s	<1s	0.520	74.27	0.45	<1s	<1s	0.512	74.27	0.48	<1s	<1s
	Poly Order 2	0.536	74.25	0.59	<1s	<1s	0.861	74.26	0.94	<1s	<1s	0.840	74.25	0.98	<1s	<1s
	Poly Order 3	0.509	74.21	0.51	<1s	<1s	0.513	74.27	0.56	<1s	<1s	0.501	74.27	0.61	<1s	<1s
	Poly Order 4	0.509	73.32	0.45	<1s	<1s	0.497	74.26	0.49	<1s	<1s	0.480	74.22	0.54	<1s	<1s
	RBF	**0.488**	74.23	0.54	1.4s	<1s	**0.456**	74.24	0.51	4.1s	<1s	**0.431**	74.21	0.57	9.8s	<1s
Sparse PC	Poly Order 3	0.565	74.25	0.52	<1s	<1s	0.527	74.27	0.45	<1s	<1s	0.439	74.22	0.55	1.1s	<1s

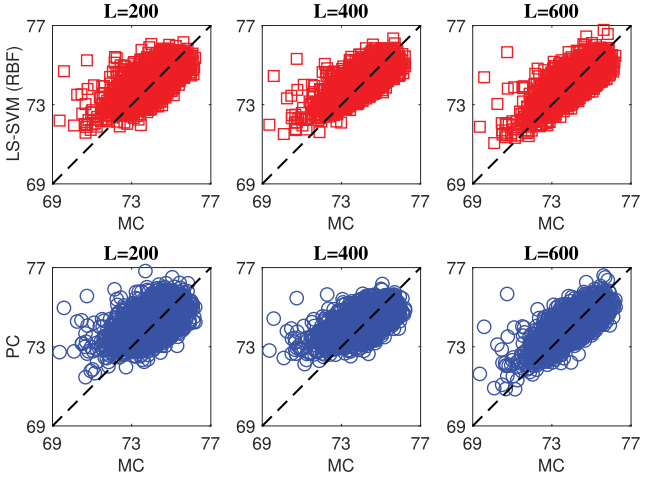

FIGURE 9. Scatter plots of the WPT efficiency obtained by comparing the predictions of the LS-SVM with RBF (red squares) kernel and the PC (blue circles) surrogates with the results of a MC simulation with 10000 samples for an increasing number of training samples i.e., $L = 200$, 400 and 600.

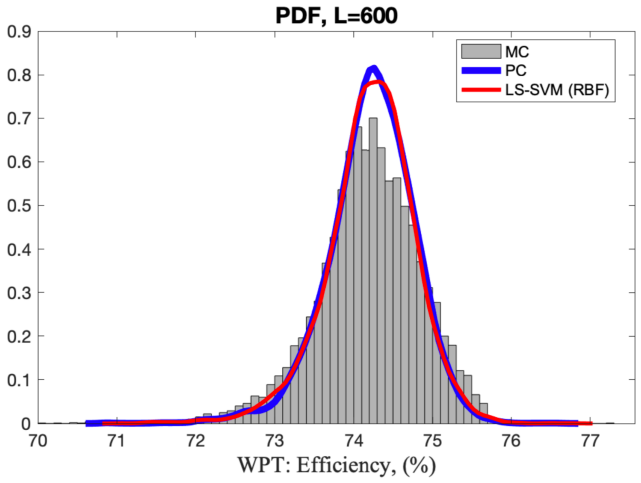

FIGURE 10. PDFs of the WPT efficiency obtained from the LS-SVM regression with RBF kernel (solid red curve) and sparse PC expansion (solid blue curve) compared with the histogram of 10000 MC samples (gray bins).

by increasing the number of training samples up to $L = 600$. The improvement in terms of accuracy can be observed in the corresponding scatter plots in Fig. 9, where the samples (squares and circles) are closer to the dashed line representing a perfect surrogate model. The model accuracy is also confirmed by the coefficient $Q^2 = 0.63$ computed for the sparse PC expansion.

As regard to the response variability, Fig. 10 presents an illustrative comparison between the PDFs predicted by the surrogate models built with the LS-SVM regression with RBF kernel and the sparse PC expansion, trained with $L = 600$ samples, with the results of a MC simulation with 10000 samples. The curves highlight once again the capability of the two surrogate models to capture the actual shape of the PDF provided by the MC simulation. As far as the computational cost is concerned, 10000 MC simulations required about 11 days whereas the LS-SVM with RBF kernel and the sparse PC needed less than 1 s. This comparison does not include the computational cost required to generate the training data ($L = 600$ samples), that has taken about 15 h as shown in Table 4.

V. CONCLUSIONS

This paper investigates the strength and the accuracy of three different approaches for the uncertainty quantification in high-dimensional space. Specifically, the SVM and the LS-SVM regressions with either polynomial or RBF Gaussian kernels and the sparse PC expansion, have been applied to generate a set of surrogate models built from a limited set of training samples for the responses of two realistic test-cases i.e., the IVR with 8 uniformly distributed uncertain parameters and the WPT with 30 uncertain parameters with Gaussian distribution, respectively. The accuracy provided by each of the proposed surrogate models has been investigated by comparing the models prediction with the results obtained by means of MC simulations. According to the results obtained for the proposed test-cases, the SVM and the LS-SVM regressions can be considered as an effective solution for uncertainty quantification in high-dimensional nonlinear problems, with an accuracy which turns out to be comparable or even better than an advanced and more consolidated technique such as the sparse PC expansion.

REFERENCES

[1] R. Spence and R. S. Soin, *Tolerance Design of Electronic Circuits*. London, U.K.: Imperial College Press, 1997.

[2] P. Triverio, S. Grivet-Talocia, and M. S. Nakhla, "A parameterized macromodeling strategy with uniform stability test," *IEEE Trans. Adv. Packag.*, vol. 32, no. 1, pp. 205–215, Feb. 2009.

[3] D. Deschrijver and T. Dhaene, "Fully parameterized macromodeling of S-parameter data by interpolation of numerator & denominator," *IEEE Microw. Wireless Compon. Lett.*, vol. 22, no. 6, pp. 309–311, Jun. 2012.

[4] E. R. Samuel, L. Knockaert, F. Ferranti, and T. Dhaene, "Guaranteed passive parameterized macromodeling by using Sylvester state-space realizations," *IEEE Trans. Microw. Theory Techn.*, vol. 61, no. 4, pp. 1444–1454, Apr. 2013.

[5] S. Grivet-Talocia and R. Trinchero, "Behavioral, parameterized, and broadband modeling of wired interconnects with internal discontinuities," *IEEE Trans. Electromagn. Compat.*, vol. 60, no. 1, pp. 77–85, Feb. 2018.

[6] Z. Zhang, T. A. El-Moselhy, I. M. Elfadel, and L. Daniel, "Stochastic testing method for transistor-level uncertainty quantification based on generalized polynomial chaos," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 32, no. 10, pp. 1533–1545, Oct. 2013.

[7] M. R. Rufuie, E. Gad, M. Nakhla, and R. Achar, "Generalized Hermite polynomial chaos for variability analysis of macromodels embedded in nonlinear circuits," *IEEE Trans. Compon., Packag., Manuf. Technol.*, vol. 4, no. 4, pp. 673–684, Apr. 2014.

[8] D. Spina, D. De Jonghe, D. Deschrijver, G. Gielen, L. Knockaert, and T. Dhaene, "Stochastic macromodeling of nonlinear systems via polynomial chaos expansion and transfer function trajectories," *IEEE Trans. Microw. Theory Techn.*, vol. 62, no. 7, pp. 1454–1460, Jul. 2014.

[9] P. Manfredi, D. Vande Ginste, D. De Zutter, and F. G. Canavero, "Stochastic modeling of nonlinear circuits via spice-compatible spectral equivalents," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 61, no. 7, pp. 2057–2065, Jul. 2014.

[10] M. Ahadi and S. Roy, "Sparse linear regression (SPLINER) approach for efficient multidimensional uncertainty quantification of high-speed circuits," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 35, no. 10, pp. 1640–1652, Oct. 2016.

[11] M. Larbi, I. S. Stievano, F. G. Canavero, and P. Besnier, "Variability impact of many design parameters: The case of a realistic electronic link," *IEEE Trans. Electromagn. Compat.*, vol. 60, no. 1, pp. 34–41, Feb. 2018.

[12] S. De Ridder *et al.*, "A generative modeling framework for statistical link analysis based on sparse data," *IEEE Trans. Compon., Packag. Manuf. Techol.*, vol. 8, no. 1, pp. 21–31, Jan. 2018.

[13] N. Femia and G. Spagnuolo, "True worst-case circuit tolerance analysis using genetic algorithms and affine arithmetic," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 47, no. 9, pp. 1285–1296, Sep. 2000.

[14] M. Wu, D. G. Beetner, T. H. Hubing, H. Ke, and S. Sun, "Statistical prediction of 'reasonable worst-case' crosstalk in cable bundles," *IEEE Trans. Electromagn. Compat.*, vol. 51, no. 3, pp. 842–851, Aug. 2009.

[15] R. Trinchero, P. Manfredi, T. Ding, and I. S. Stievano, "Combined parametric and worst case circuit analysis via Taylor models," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 63, no. 7, pp. 1067–1078, Jul. 2016.

[16] R. Trinchero, P. Manfredi, and I. S. Stievano, "TMsim: An algorithmic tool for the parametric and worst-case simulation of systems with uncertainties," *Math. Problems Eng.*, vol. 2017, Mar. 2017, Art. no. 6739857.

[17] Y.-B. Yuan, W.-Q. Qiu, Y.-J. Wang, J. Gao, and P. He, "Classification of heart failure with polynomial smooth support vector machine," in *Proc. Int. Conf. Mach. Learn. Cybern.*, Ningbo, China, Jul. 2017, pp. 48–54.

[18] M. Sabzekar, H. S. Yazdi, and M. Naghibzadeh, "Relaxed constraints support vector machines for noisy data," *Neural Comput. Appl.*, vol. 20, no. 5, pp. 671–685, 2011.

[19] J. M. Moguerza and A. Muñoz, "Support vector machines with applications," *Stat. Sci.*, vol. 21, no. 3, pp. 322–336, 2006.

[20] D. Plets *et al.*, "Surrogate modeling based cognitive decision engine for optimization of WLAN performance," *Wireless Netw.*, vol. 23, no. 8, pp. 2347–2359, May 2017.

[21] H. M. Torun, C. Pardue, M. L. F. Belleradj, A. K. Davis, and M. Swaminathan, "Machine learning driven advanced packaging and miniaturization of IoT for wireless power transfer solutions," in *Proc. IEEE 68th Electron. Compon. Technol. Conf. (ECTC)*, San Diego, CA, USA, May 2018, pp. 2374–2381.

[22] B. Tang, T. Song, F. Li, and L. Deng, "Fault diagnosis for a wind turbine transmission system based on manifold learning and Shannon wavelet support vector machine," *Renew. Energy*, vol. 62, pp. 1–9, Feb. 2014.

[23] H. M. Torun, M. Swaminathan, A. K. Davis, and M. L. F. Bellaredj, "A global Bayesian optimization algorithm and its application to integrated system design," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 26, no. 4, pp. 792–802, Apr. 2018.

[24] H. Yu, M. Swaminathan, C. Ji, and D. White, "A method for creating behavioral models of oscillators using augmented neural networks," in *Proc. IEEE 26th Conf. Elect. Perform. Electron. Packag. Syst. (EPEPS)*, San Jose, CA, USA, Oct. 2017, pp. 1–3.

[25] V. Vapnik, *Statistical Learning Theory*. Hoboken, NJ, USA: Wiley, 1998.

[26] V. Vapnik, *The Nature of Statistical Learning Theory*, 2nd ed. New York, NY, USA: Springer, 1999.

[27] V. Cherkassky and Y. Ma, "Practical selection of SVM parameters and noise estimation for SVM regression," *Neural Netw.*, vol. 17, no. 1, pp. 113–126, 2004.

[28] A. J. Smola and B. Schölkopf, "A tutorial on support vector regression," *Statist. Comput.*, vol. 14, no. 3, pp. 199–222, 2004.

[29] H. Wang and D. Hu, "Comparison of SVM and LS-SVM for regression," in *Proc. Int. Conf. Neural Netw. Brain*, Beijing, China, Oct. 2005, pp. 279–283.

[30] M. Espinoza, J. A. K. Suykens, R. Belmans, and B. D. Moor, "Electric load forecasting," *IEEE Control Syst.*, vol. 27, no. 5, pp. 43–57, Oct. 2007.

[31] J. A. K. Suykens, T. Van Gestel, J. De Brabanter, B. De Moor, and J. Vandewalle, *Least Squares Support Vector Machines*. Singapore: World Scientific, 2002.

[32] M. Arefi and B. Chowdhury, "Ensemble adaptive neuro fuzzy support vector machine for prediction of transient stability," in *Proc. North Amer. Power Symp.*, Morgantown, WV, USA, Sep. 2017, pp. 1–6.

[33] V. R. Kohestani and M. Hassanlourad, "Modeling the mechanical behavior of carbonate sands using artificial neural networks and support vector machines," *Int. J. Geomech.*, vol. 16, no. 1, pp. 1–9, 2016.

[34] R. Trinchero and F. G. Canavero, "Modeling of eye diagram height in high-speed links via support vector machine," in *Proc. IEEE 22nd Workshop Signal Power Integr. (SPI)*, Brest, France, May 2018, pp. 1–4.

[35] R. Trinchero and F. G. Canavero, "Design of passive equalizer for space wire links via support vector machine," in *Proc. Int. Symp. Electromagn. Compat. (EMC EUROPE)*, Amsterdam, The Netherlands, Aug. 2018, pp. 53–56.

[36] R. Trinchero, P. Manfredi, I. S. Stievano, and F. G. Canavero, "Machine learning for the performance assessment of high-speed links," *IEEE Trans. Electromagn. Compat.*, vol. 60, no. 6, pp. 1627–1634, Dec. 2018.

[37] C. Soize and R. Ghanem, "Physical systems with random uncertainties: Chaos representations with arbitrary probability measure," *SIAM J. Sci. Comput.*, vol. 26, no. 2, pp. 395–410, 2004.

[38] M. Berveiller, B. Sudret, and M. Lemaire, "Stochastic finite element: A non intrusive approach by regression," *Eur. J. Comput. Mech.*, vol. 15, nos. 1–3, pp. 81–92, 2006.

[39] 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, 2011.

[40] M. Larbi, H. M. Torun, M. Swaminathan, I. S. Stievano, F. G. Cañavero, and P. Besnier, "Uncertainty quantification of SiP based integrated voltage regulator," in *Proc. IEEE 22nd Workshop Signal Power Integr. (SPI)*, Brest, France, May 2018, pp. 1–4.

[41] B. Efron, T. Hastie, I. Johnstone, and R. Tibshirani, "Least angle regression," *Ann. Statist.*, vol. 32, no. 2, pp. 407–499, 2004.

[42] I. M. Sobol, "Sensitivity estimates for nonlinear mathematical models," *Math. Model. Comput. Exp.*, vol. 1, no. 4, pp. 407–414, 1990.

[43] B. Sudret, "Global sensitivity analysis using polynomial chaos expansions," *Rel. Eng. Syst. Saf.*, vol. 93, no. 7, pp. 964–979, Jul. 2008.

[44] Department of Electrical Engineering (ESAT), Katholieke Universiteit Leuven: Leuven, Belgium. (2011). *LS-SVMlab, Version 1.8*. [Online]. Available: http://www.esat.kuleuven.be/sista/lssvmlab/

[45] S. Mueller *et al.*, "Design of high efficiency integrated voltage regulators with embedded magnetic core inductors," in *Proc. IEEE 66th Electron. Compon. Technol. Conf. (ECTC)*, May 2016, pp. 566–573.

[46] R. Jay and S. Palermo, "Resonant coupling analysis for a two-coil wireless power transfer system," in *Proc. IEEE Dallas Circuits Syst. Conf. (DCAS)*, Oct. 2014, pp. 1–4.

[47] *Ansys HFSS Ver. 2015.2*. Accessed: Dec. 2018. [Online]. Available: http://www.ansys.com

[48] 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 from a computer code," *Technometrics*, vol. 42, no. 1, pp. 55–61, 2000.

[49] D. C. Montgomery, *Design and Analysis of Experiments*. New York, NY, USA: Wiley, 2004.

RICCARDO TRINCHERO (M'16) received the M.Sc. and the Ph.D. degrees in electronics and communication engineering from the Politecnico di Torino, Torino, Italy, in 2011 and 2015, respectively.

He is currently a Researcher with the EMC Group, Department of Electronics and Telecommunications, Politecnico di Torino. His research interests include the analysis of linear time-varying systems, modeling and simulation of switching converters, and statistical simulation of circuits and systems.

MOURAD LARBI received the M.S. degree in applied statistics from the University of Nice Sophia-Antipolis, Nice, France, in 2011, and the Ph.D. degree in electronics and telecommunications from the Institute of Electronics and Telecommunications of Rennes, Rennes, France, in 2016. From 2016 to 2017, he was a Post-Doctoral Researcher with the Electromagnetic Compatibility Group, Department of Electronics and Telecommunications, Politecnico di Torino, Turin, Italy. He is currently a Post-Doctoral Researcher with the Center for Co-design, Chip, Package, System, School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA, USA. His current research interests include behavioral modeling and the risk analysis of signal propagation on interconnects in high-dimensional uncertainty quantification problems. He received the Richard B. Schulz Best EMC Transactions Paper Award—Honorable Mention, in 2016, and the Best Paper Award at the IEEE International Symposium and Exhibition on Electromagnetic Compatibility (EMC Europe), in 2017.

HAKKI M. TORUN (S'15) received the B.Sc. degree in electrical and electronics engineering from Bilkent University, Turkey, in 2016. He is currently pursuing the Ph.D. degree with the School of Electrical and Computer Engineering, Georgia Institute of Technology, under the supervision of Dr. M. Swaminathan. His research interests include developing machine learning algorithms for system level optimization and modeling with applications in signal and power integrity in high-speed channels, microwave electronics, and VLSI systems.

Mr. Torun was a recipient of the Best Student Paper Award of the IEEE 27th Conference on Electrical Performance of Electronic Packaging and Systems, in 2018.

FLAVIO G. CANAVERO (SM'99–F'07) received the Electronic Engineering degree from the Politecnico di Torino (Technical University), Italy, and the Ph.D. degree from the Georgia Institute of Technology, Atlanta, GA, USA, in 1986. He is currently a Professor of Circuit Theory with the Department of Electronics and Telecommunications, Politecnico di Torino, where he also serves as the Director of the Doctoral School. His research interests include signal integrity and EMC design issues, interconnect modeling, black-box characterization of digital integrated circuits, EMI, and statistics in EMC.

He is an IEEE Fellow. He has been the Editor-in-Chief of the IEEE TRANSACTIONS ON ELECTROMAGNETIC COMPATIBILITY, the Vice President for Communication Services of the EMC Society, and the Chair of the URSI Commission E. He received several industry and IEEE awards, including the prestigious Richard R. Stoddard Award for Outstanding Performance, which is the EMC society's highest technical award, and the Honored Member Award of the EMC Society.

MADHAVAN SWAMINATHAN (M'95–SM'98–F'06) received the M.S. and Ph.D. degrees in electrical engineering from Syracuse University, in 1989 and 1991, respectively. He was with IBM, where he was involved in packaging for supercomputers. He was a Joseph M. Pettit Professor in electronics with the School of Electrical and Computer Engineering and the Deputy Director of the NSF Microsystems Packaging Research Center, Georgia Institute of Technology (Georgia Tech), Atlanta, GA, USA. He is currently a John Pippin Chair Professor in microsystems packaging and electromagnetics with the School of Electrical and Computer Engineering and the Director of the Center for Co-Design of Chip, Packages, System, Georgia Tech. He is also the Founder or the Co-Founder of two start-up companies, E-System Design and Jacket Micro Devices. He has authored over 450 refereed technical publications and holds 29 patents. He is a primary author and a co-editor of three books. He is the Founder of the IEEE Conference Electrical Design of Advanced Packaging and Systems, a premier conference sponsored by the CPMT Society on Signal Integrity in the Asian region. He has served as the Distinguished Lecturer for the IEEE EMC Society.