Combining LS-SVM and GP Regression for the Uncertainty Quantification of the EMI of Power Converters Affected by Several Uncertain Parameters

Combining LS-SVM and GP Regression for the Uncertainty Quantification of the EMI of Power Converters Affected by Several Uncertain Parameters

Riccardo Trinchero, *Member, IEEE,* and Flavio G. Canavero, *Fellow, IEEE*

***Abstract*—This paper deals with the development of a probabilistic surrogate model for the uncertainty quantification of the voltage output spectral envelope of a power converter with several stochastic parameters. The proposed approach relies on the combination of the least-squares support vector machine (LS-SVM) regression with the Gaussian process regression (GPR), but it can suitably be applied to any deterministic regression techniques. As a first step, the LS-SVM regression is used to build an accurate and fast-to-evaluate deterministic model of the system responses starting from a limited set of training samples provided by the full-computational model. Then, the GPR is used to provide a probabilistic model of the regression error. The resulting LS-SVM+GPR probabilistic model not only approximates the system responses for any configuration of its input parameters, but it also provides an estimation of its prediction uncertainty, such as the confidence intervals (CI). The above technique has been applied to qualify the uncertainty of the spectral envelope of the output voltage of a buck converter with 17 independent Gaussian parameters. The feasibility and the accuracy of the resulting model has been investigated by comparing its predictions and CI with the ones obtained by five different surrogate models based on state-of-the-art techniques and by the reference Monte Carlo results.**

***Index Terms*—Machine Learning, least-squares support vector machine, Gaussian process, confidence interval, conductive emission, switching converter.**

I. Introduction

Switching power converters play a key role in modern devices, since they provide an efficient and compact solution for power conversion. However, due to their time-varying activity, the current and voltage waveforms at the input and output stage of the converter are usually characterized by high-frequency noisy components behaving as conducted emissions (CE). The effect of possible component tolerances on the spectral content of the electromagnetic interferences (EMI) generated by the switching converter must be carefully investigated, especially during the early design phase, through statistical tools and methodologies for the uncertainty quantification in order to avoid possible electromagnetic compatibility (EMC) issues and thus expensive redesign [1].

The work has been partly supported by the Joint Project for Internalization of Research 2018: "Machine Learning to Improve the Reliability of Complex Systems".

R. Trinchero and F. G. Canavero are with the EMC Group, Department of Electronics and Telecommunications, Politecnico di Torino, 10129 Torino, Italy (e-mail: riccardo.trinchero@polito.it, flavio.canavero@polito.it).

Monte Carlo (MC) simulation can be considered as a traditional technique for such kind of analysis. The underlying idea is to run thousands of deterministic circuital simulations in which the components parameters are varied according to their probabilistic distribution, in order to capture the actual statistical behavior of the quantity of interest [2]. Despite its simplicity, this direct approach turns out to be computational expensive, since it requires a large number of simulations with the full-computational model. Also, the MC approach turns out to be a blind method which does not provide any relationship between the input parameters and the simulation output. In addition, it is important to remark that, due to the nonlinearity and time-varying activity of the switching converters, their simulations must be carried on in time-domain and the CE spectrum is then calculated off-line from the steady-state waveforms via the Fourier transform, leading to a non-negligible computational overhead [3]–[5].

In the last decades, advanced techniques such as Polynomial Chaos (PC) expansion [6]–[8] and its advanced variants, such as the least-angle regressions (LARS) PC expansion [9]–[11] have been proposed as alternatives to the traditional MC analysis for the uncertainty quantification in complex systems. The above techniques allow building accurate and fast-to-evaluate surrogate models for the statistical analysis of the outputs of a generic nonlinear system affected by stochastic parameters. Recently, several advanced general purpose regression techniques belonging to the Machine Learning (ML) framework [12] have been adopted for the surrogate modeling and the uncertainty quantification in many research fields. In particular, the support vector machine (SVM) [13], [14] and the least-squares support vector machine (LS-SVM) [15] regression can be seen as a viable and accurate solution for the surrogate modeling in high-dimensional parameter space, sometime providing an improved accuracy with respect to well-established PC-based expansions [16], [17].

All the above techniques provide as result a *deterministic* model. This means that the resulting model can be interpreted as a function, which allows predicting the system response for any configuration of its input parameters, without any information on the degree of confidence and the uncertainty of its predictions. In fact, the degree of confidence of the model prediction is usually known only for the set of training samples used to build it, but it is completely unknown when the model is evaluated for a generic point in the parameters space. Indeed,

when we talk about deterministic models, one of the most challenge question that can arise is: *"How can we predict the accuracy of a model without running an equivalent simulation with the full-computational model?"*

Gaussian process regression (GPR) [18]–[25], also known as Kriging model, represents a possible solution to the above challenging problem. The GPR belongs to the ML techniques and it allows building a *probabilistic* model of the nonlinear response of a complex system starting from a limited set of training samples. The resulting model not only provides a prediction of the model output, but it also allows estimating the uncertainty of its prediction for any configuration of its input parameters, such as the confidence intervals (CI) [26]. Also, the GPR is so general that it can be used to enrich any kind of deterministic model resulting from a generic regression with the CI of its predictions, thus providing the user with a *probabilistic* model [18], [25]. It is important to point out that there are several reasons why one might wish to combine the GPR with an available explicit regression model, instead of using the set of mean functions available within the GPR, including: interpretability of the model, convenience of expressing prior information and improved accuracy [18].

This paper presents an unconventional technique called LS-SVM+GPR (preliminary results have been recently presented in [27]) for the generation of a probabilistic model based on a two-step procedure: (i) generate a deterministic model based on a LS-SVM regression; (ii) use the GPR to build a probabilistic model of the LS-SVM regression error function.

The proposed modeling technique has been applied to the uncertainty quantification of the spectral envelope of the output voltage of a realistic buck converter as a function of 17 independent Gaussian distributed parameters. The model predictions are then compared with the ones provided by five different surrogate models, i.e., the deterministic LS-SVM regression with linear and RBF kernel and the GPR with constant, linear and polynomial (order 2) trends, respectively. The accuracy of all considered models will be investigated by comparing their predictions with the results of a MC simulation in LTspice [28].

The remainder of this paper is organized as follows. Section II presents the proposed LS-SVM+GPR techniques. Section III compares the accuracy of the proposed LS-SVM+GPR probabilistic modeling with respect to the accuracy of five different surrogate models based on the LS-SVM and the GPR for the prediction of output voltage spectral envelope of the output voltage of a buck converter as a function of 17 uncertain parameters. Section IV concludes the paper.

II. Probabilistic Model based on LS-SVM+GPR

This section focuses on the development of a probabilistic surrogate model of the responses of a generic nonlinear function in a high-dimensional parameter space based on the combination of the LS-SVM regression with the GPR.

A. Step 1: Deterministic model via LS-SVM Regression

Let us start considering the problem of fitting a given set of training pairs $\mathcal{D}_{1:L} = \{(\mathbf{x}_i, y_i)\}_{i=1}^{L}$, provided by a full computational model $\mathcal{M}$ (i.e., $y_i = \mathcal{M}(\mathbf{x}_i)$) where $y_i \in \mathbb{R}$ and $\mathbf{x}_i \in \mathcal{P}$ with $\mathcal{P} \subset \mathbb{R}^d$ (d represents the problem dimensionality, i.e., the number of uncertain parameters) via the following LS-SVM regression $\mathcal{M}_{LS\text{-}SVM}(\mathbf{x})$ in the dual space [15]:

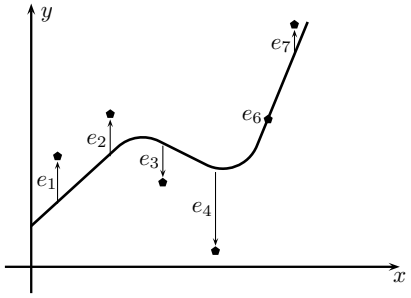

Fig. 1. Graphical interpretation of the errors ε_i used within the LS-SVM regression optimization. For illustration purposes, a 1-D parameter space is considered only.

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

where $\alpha_i \in \mathbb{R}$ are scalar coefficients, $K(\cdot,\cdot) : \mathbb{R}^{d\times d} \to \mathbb{R}$ is the *kernel function* and $b \in \mathbb{R}$ is the bias term.

In summary, the LS-SVM regression allows building accurate and fast-to-evaluate *deterministic* models of the response of a generic high-dimensional nonlinear function $\mathcal{M}$ starting from a number of L training samples [17]. This regression provides an alternative interpretation to the standard SVM regression [13]–[16] based on a more intuitive least-squares formulation [15], as further explained at the end of this subsection. The most common kernels used in both the SVM and the LS-SVM regression are listed below [13], [14], [15]:

- 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): $K(\mathbf{x}_i, \mathbf{x}) = \exp\left(-\|\mathbf{x}_i - \mathbf{x}\|^2 / 2\sigma^2\right)$.

It is worth to remark that, different from the standard regression techniques (e.g., the plain least-squares regression), thanks to the use of kernels, any SVM-based regression in the dual form provides a model for which the number of unknowns to be estimated during the training phase (i.e., the number of coefficients α_i in (1)) turns out to be independent from the dimensionality d of the input parameter space. Indeed, the number of unknowns for the regression is equal to the number of training samples L used to train the model [14], [15].

The goal of the LS-SVM regression in (1) is to minimize the squared of the error e_i between the model output and the training samples, where $e_i = \mathcal{M}(\mathbf{x}_i) - \mathcal{M}_{LS\text{-}SVM}(\mathbf{x}_i)$ (see Fig. 1 for a pictorial illustration in a 1-D parameter space). The least-squares problem leads to the following linear system, allowing to estimate the parameters α_i and b:

$$\begin{bmatrix} 0 & \mathbf{1}^T \\ \mathbf{1} & \boldsymbol{\Omega} + \mathbf{I}/\gamma \end{bmatrix} \begin{bmatrix} b \\ \boldsymbol{\alpha} \end{bmatrix} = \begin{bmatrix} 0 \\ \mathbf{y} \end{bmatrix} \qquad (2)$$

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 $\boldsymbol{\Omega} \in$

$\mathbb{R}^{L\times L}$ is the kernel matrix, whose elements represents the kernel computed for all combinations of training points, i.e., $\Omega_{ij} = K(\mathbf{x}_i, \mathbf{x}_j)$ for any $i, j = 1, \ldots, L$. The LS-SVM regression is already available in MATLAB within LS-SVMLab Toolbox version 1.8 [29].

B. Step 2: Probabilistic model of the regression error via GPR

Let us consider the error function $e(\mathbf{x})$ between the deterministic model $\mathcal{M}_{LS\text{-}SVM}(\mathbf{x})$ and the full-computational model $\mathcal{M}(\mathbf{x})$, which simply writes:

$$e(\mathbf{x}) = \mathcal{M}(\mathbf{x}) - \mathcal{M}_{LS\text{-}SVM}(\mathbf{x}). \tag{3}$$

For a generic surrogate model built through a deterministic regression, such as the LS-SVM regression, the error function $e(\mathbf{x})$ is only known for a discrete set of configurations of input parameters $\mathbf{x}_i$ for which the corresponding responses of the full-computational model $y_i = \mathcal{M}(\mathbf{x}_i)$ are available (i.e., these are the values used to train the deterministic regression or to validate the model). This means that we are unable to quantify the precision of the model prediction for a generic input configuration $\mathbf{x}_*$, without running an equivalent simulation of the full-computational model (i.e., without knowing $y_* = \mathcal{M}(\mathbf{x}_*)$).

At this stage, we make use of GPR as a viable solution which allows overcoming the above-mentioned limitation. Indeed, the GPR can be adopted to enrich any deterministic models built via a regression technique [18]. Without loss of generality, in the remaining of this Section, the proposed modeling scheme is applied to the LS-SVM regression shown in eq. (1), leading to the following formulation:

$$\mathcal{M}(\mathbf{x}) \approx \overbrace{\mathcal{M}_{LS\text{-}SVM}(\mathbf{x}) + \tilde{e}(\mathbf{x})}^{\mathcal{M}_{LS\text{-}SVM+GPR}(\mathbf{x})}, \tag{4}$$

where $\mathcal{M}_{LS\text{-}SVM}$ is the deterministic regression estimated via the LS-SVM regression presented in Sec. II.1, and $\tilde{e}(\mathbf{x})$ is an unknown function accounting for the regression residuals. This means that we are assuming the existence of a nonlinear function $\tilde{e}(\mathbf{x})$ approximating the error function $e(\mathbf{x})$ in (3) of the deterministic model, i.e., $e(\mathbf{x}) \approx \tilde{e}(\mathbf{x})$.

As it is very very unlikely that, in practice, the error $\tilde{e}(\mathbf{x})$ is an uncorrelated random error (like a white noise signal), a commonly used approach is to describe the error as a Gaussian process, i.e., $\tilde{e}(\mathbf{x}) \sim GP(0, k(\mathbf{x}, \mathbf{x}'))$ with zero mean and variance function $k(\cdot, \cdot)$ [19]. In elementary terms, a GP is analogous to a function, but instead of returning a scalar value of $\tilde{e}(\mathbf{x})$ for an arbitrary $\mathbf{x}$, it returns an ensemble of values drawn from a Gaussian distribution, subject to some smoothness condition imposing a given correlation function $k(\mathbf{x}, \mathbf{x}')$ between any two inputs $\mathbf{x}$ and $\mathbf{x}'$, [20]. Hence, the mean and variance over the possible values of $\tilde{e}$ at $\mathbf{x}$ are readily evaluated (see Appendix A for additional information). For the sake of terminological precision, the formulation in (4) is a particular case of the GPR with a *fixed* mean function [18].

More technically, the above-mentioned correlation implies that the residuals calculated on the training samples $e(\mathbf{x}_i) = \mathcal{M}(\mathbf{x}_i) - \mathcal{M}_{LS\text{-}SVM}(\mathbf{x}_i) \neq 0$[1] for $i = 1, \ldots, L$ can be modeled via a L-dimensional multivariate distribution, given by:

$$\begin{bmatrix} e(\mathbf{x}_1) \\ \vdots \\ e(\mathbf{x}_L) \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} k(\mathbf{x}_1, \mathbf{x}_1) & \ldots & k(\mathbf{x}_1, \mathbf{x}_L) \\ \vdots & \ddots & \vdots \\ k(\mathbf{x}_L, \mathbf{x}_1) & \ldots & k(\mathbf{x}_L, \mathbf{x}_L) \end{bmatrix} \right), \tag{5}$$

where $\mathcal{N}$ indicates the Normal distribution.

It is important to remark that the choice of covariance function $k(\cdot, \cdot)$ is extremely important for our modeling purposes, since it specifies the correlation among the values of the error function in (3) for any value of $\mathbf{x} \in \mathcal{P}$ [21]. The underlying idea is, that points with similar predictor values are aspected to have close response values, therefore we are implicitly assuming that $e(\mathbf{x})$ is *smooth* [18]. Various correlation functions are available in literature [20], [18]. Without loss of generality, we focus on the Matern 5/2 covariance function with an automatic relevance determination (ARD) hyper-parameters (i.e., $\boldsymbol{\theta}$) [18], [20], [30], which writes:

$$k(\mathbf{x}, \mathbf{x}'|\boldsymbol{\theta}) = \sigma_f^2 \left(1 + \sqrt{5}r + \frac{5}{3}r^2\right) \exp(-\sqrt{5}r), \tag{6}$$

with,

$$r = \sqrt{\sum_{m=1}^{d} \frac{(x_m - x'_m)^2}{\sigma_m^2}}, \tag{7}$$

where σ_f and σ_m for $m = 1, \ldots, d$ are the so-called hyper-parameters collected in the vector $\boldsymbol{\theta}$. The hyper-parameters $\boldsymbol{\theta}$ can be estimated through optimization from the available information on the error training samples $e(\mathbf{x}_i)$ for $i = 1, \ldots, L$. As an example, the GPR tool of MATLAB allows estimating the above quantities by maximization of the log likelihood [18].

Thanks to the properties of GP and to the *prior* information provided by the training error samples, for any new value of the input parameter $\mathbf{x}_* \in \mathcal{P}$, such that $\mathbf{x}_* \neq \mathbf{x}_i$ for $i = 1, \ldots, L$, the samples of $\{e(\mathbf{x}_1), \ldots, e(\mathbf{x}_L), e(\mathbf{x}_*)\}$ follow an $L+1$-dimensional joint Gaussian distribution [19], which writes:

$$\begin{bmatrix} \mathbf{e} \\ \tilde{e}_* \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \mathbf{0} \\ 0 \end{bmatrix}, \begin{bmatrix} \mathbf{K} & \mathbf{k}_*^T \\ \mathbf{k}_* & k_{**} \end{bmatrix} \right), \tag{8}$$

where $\mathbf{e} = [e(\mathbf{x}_1), \ldots, e(\mathbf{x}_L)]^T$, $\mathbf{K} \in \mathbb{R}^{L\times L}$ is the correlation matrix given by $K_{ij} = k(\mathbf{x}_i, \mathbf{x}_j)$, $\mathbf{k}_* = [k(\mathbf{x}_*, \mathbf{x}_1), \ldots, k(\mathbf{x}_*, \mathbf{x}_L)] \in \mathbb{R}^{1\times L}$, $k_{**} = k(\mathbf{x}_*, \mathbf{x}_*)$ and $\tilde{e}_* = \tilde{e}(\mathbf{x}_*)$ is a prediction of the error function at $\mathbf{x}_*$.

The probability of predicting $\mathcal{M}(\mathbf{x}_*)$, called posterior distribution, given the prior information on the training samples $\mathcal{D}_{1:L}$, corresponds to conditioning the joint distribution in (8) on the observations (i.e., training samples):

$$p(\mathcal{M}_{LS\text{-}SVM+GPR}|\mathbf{x}_*, \mathcal{D}_{1:L}) \sim N(\mu_{\mathbf{x}_*}, \sigma^2_{\mathbf{x}_*}) \tag{9}$$

[1] It is ought to remark that, for a regression, the values of the error function on the training samples, i.e., $e_i = e(\mathbf{x}_i)$ are usually different from zero.

where $\mu_{\mathbf{x}_*}$ and $\sigma^2_{\mathbf{x}_*}$ are defined as follows:

$$\mu_{\mathbf{x}_*} = \mathcal{M}_{LS\text{-}SVM}(\mathbf{x}_*) + \mathbf{k}_* \mathbf{K}^{-1} \mathbf{e} \tag{10a}$$

$$\sigma^2_{\mathbf{x}_*} = k_{**} - \mathbf{k}_* \mathbf{K}^{-1} \mathbf{k}_*^T. \tag{10b}$$

We are then going to use the prediction mean $\mu_{\mathbf{x}_*}$ instead of the deterministic LS-SVM regression $\mathcal{M}_{LS\text{-}SVM}(\mathbf{x}_*)$ in (1), whereas the variance $\sigma^2_{\mathbf{x}_*}$ gives a local error indicator about the precision of the estimate. The above probabilistic interpretation allows estimating the confidence interval at the $100(1-\alpha)\%$ level, such that the full computational model $\mathcal{M}(\mathbf{x}_*)$ at any point $\mathbf{x}_* \in \mathcal{P}$:

$$\left(\mu_{\mathbf{x}_*} - z_{1-\frac{\alpha}{2}} \sigma_{\mathbf{x}_*}\right) \leq \mathcal{M}(\mathbf{x}_*) \leq \left(\mu_{\mathbf{x}_*} + z_{1-\frac{\alpha}{2}} \sigma_{\mathbf{x}_*}\right), \tag{11}$$

with a probability of $(1-\alpha)$, [26] where z denotes the inverse of the Gaussian cumulative distribution function evaluated at $1 - \frac{\alpha}{2}$ and $\mu_{\mathbf{x}_*} \pm z_{1\text{-}\frac{\alpha}{2}} \sigma^2_{\mathbf{x}_*}$ represent the upper and lower confidence bound, respectively.

It is ought to be remarked that the above formulation holds only for a GPR with a fixed mean function. In fact, the standard GPR does not use the LS-SVM model as a trend and writes [18], [25]:

$$\mathcal{M}(\mathbf{x}) \approx GP(\boldsymbol{\beta}\mathbf{f}(\mathbf{x})^T, k(\mathbf{x}, \mathbf{x}')), \tag{12}$$

where $k(.,.)$ is the covariance function (see as an example the ADR Matern 5/2 covariance function in (6)) and $\boldsymbol{\beta}\mathbf{f}(\mathbf{x})^T$ is the GP trend in which $\mathbf{f}(\mathbf{x}) = [f_1(\mathbf{x}), \ldots, f_P(\mathbf{x})]$ indicates a set of bases functions and $\boldsymbol{\beta} = [\beta_1, \ldots, \beta_P]$ are the regression parameters to be estimated during the training of the GPR. As an example the MATLAB tool for the GPR works directly with constant, linear and polynomial bases functions.

III. Application Example

The LS-SVM+GPR modeling technique presented in Sec. II has been applied to build a probabilistic surrogate model for the uncertainty quantification of the output voltage spectral envelope of the switching converter in Fig. 2, as a function of 17 stochastic parameters. The converter is a 12V:5V switching buck converter with its feedback network (see [31] for additional details) operating at a switching frequency of 100 kHz through a sawtooth signal defined between $0\,V$ and $5\,V$.

In the following analysis, the values of all the components specified in the schematic of Fig. 2 have been considered as Gaussian stochastic variables centered at their nominal value and with a standard deviation of 10% around their mean value, leading to 17 uncorrelated Gaussian parameters (i.e., $\mathbf{x} \in \mathbb{R}^{17}$). The full-computational model used to evaluate the global effect of the uncertainty parameters on the stochastic behavior of the spectral envelope of the voltage output of the converter is based on a parametric transient simulation in LTspice. Specifically, the spectral envelope of the output voltage v_{out}, namely $V_{out,E}(f; \mathbf{x}_*)$, for a generic configuration of the circuit parameters $\mathbf{x}_*$ is calculated through the full-computational model via the following procedure: (i) use the input parameter configuration to run the corresponding transient simulation in LTspice; (ii) compute the spectrum $\hat{V}_{out}(f; \mathbf{x}_*)$ by applying the Fast Fourier transform (FFT) on the steady-state portion of the voltage waveform $v_{out}(t; \mathbf{x}_*)$; (iii) compute the magnitude of the peak spectral envelope $V_{out,E}(f; \mathbf{x}_*)$ via the MATLAB function `envelop` and convert the resulting spectrum in dB.

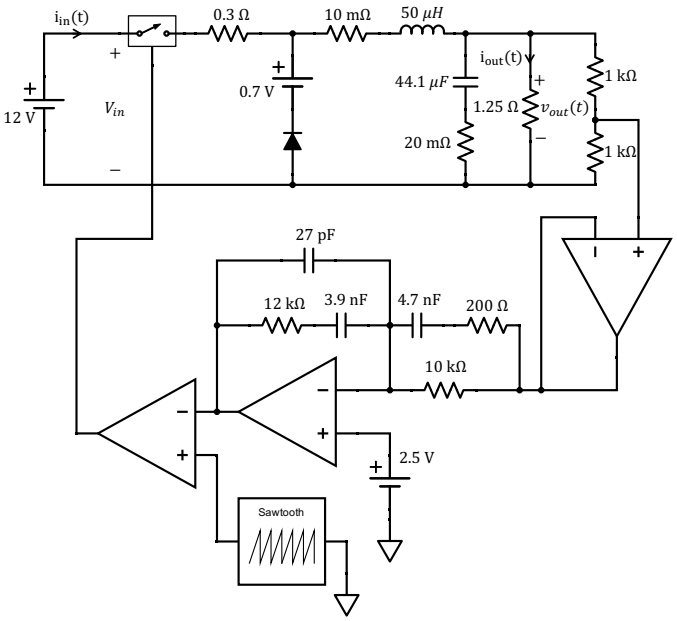

Fig. 2. The buck converter schematic considered in Sec. III [31]. For each component, nominal values are indicated.

For any configuration of the input parameters, the transient simulation has been run in the time window $[0, 3]$ ms with a time-step of 10 ns. In order to ensure that all the waveforms have reached the steady-state, the FFT has been applied only to the last 3 switching periods of the voltage waveform $v_{out}(t; \mathbf{x}_*)$ [3], [4]. The resulting spectral envelope $V_{out,E}(f_k; \mathbf{x}_*)$ with $k = 1, \ldots, N_f$ covers a frequency bandwidth from DC to 30 MHz via a set of $N_f = 91$ linearly spaced frequency samples.

For each of the frequency component f_k, the above simulation scheme (i.e., the full-computational model) has been used to generate a set of L training samples $\{(\mathbf{x}_i, y_i(f_k)\}_{i=1}^L$, where the input parameter configurations $[\mathbf{x}_1, \ldots, \mathbf{x}_L]$ have been drawn based on the latin hypercube sampling scheme [32] and $y_i(f_k) = V_{out,E}(f_k; \mathbf{x}_i))$. The training samples have been used to train the proposed LS-SVM+GPR surrogate model which writes:

$$\begin{aligned} V_{out,E}(f_k; \mathbf{x}) &\approx \mathcal{M}_{LS\text{-}SVM+GPR}(f_k; \mathbf{x}) \\ &= \mathcal{M}_{(LS\text{-}SVM+GPR),k}(\mathbf{x}) \end{aligned} \tag{13}$$

for any $\mathbf{x} \in \mathcal{P}$, and $k = 1, \ldots, N_f$.

Specifically, two different LS-SVM+GPR models have been trained by considering two different LS-SVM regressions with either linear or RBF kernel. For the sake of completeness, the same training samples have been also used to build other five different surrogate models based on state-of-the-art techniques such the deterministic LS-SVM regression with linear and RBF kernel and the standard probabilistic GPR with constant,

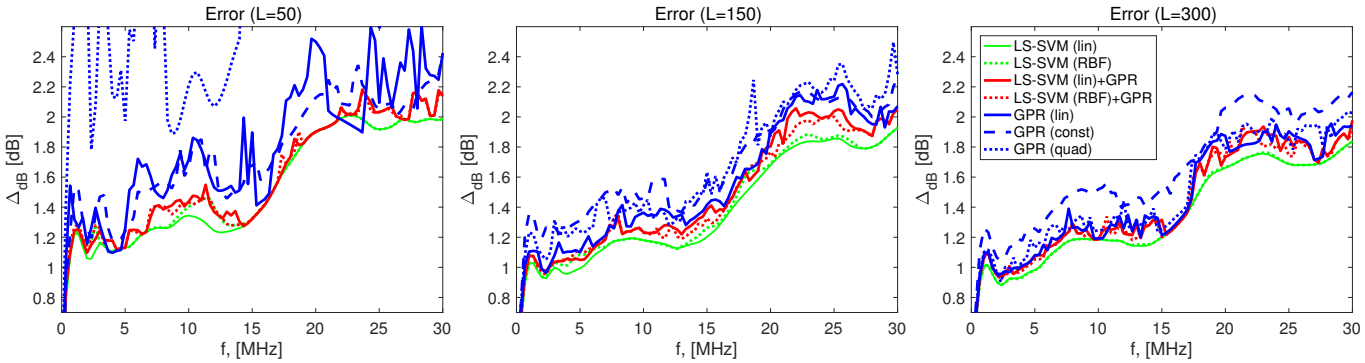

Fig. 3. Average absolute error $\Delta_{dB}(f)$ calculated by comparing the predictions of the surrogate models based on the LS-SVM (solid and dashed green curves), the LS-SVM+GPR (solid and dashed red curves) and the plain GPR (blue solid, dashed and dotted curves) with the corresponding ones obtained via a MC simulation with 10,000 samples for an increasing number of training samples $L = 50, 150$ and 300.

linear and polynomial (order 2) trend [18], respectively. All the proposed probabilistic models (i.e., the ones based on the LS-SVM+GPR and the GPR) use the ARD Matern 5/2 covariance function in (6).

The accuracy of each model is then illustrated in Fig. 3. The plots provide a comparison among the accuracy provided by each of the considered surrogate models in terms of the average absolute error spectrum $\Delta_{dB}(f_k)$ defined for $k = 1, \ldots, N_f$, as follows

$$\Delta_{dB}(f_k) = \sum_{i=1}^{N_{MC}} \frac{\left| V_{out,E}(f_k; \mathbf{x}_i) - \tilde{V}_{out,E}(f_k; \mathbf{x}_i) \right|}{N_{MC}}, \quad (14)$$

where $V_{out,E}(f_k; \mathbf{x}_i)$ corresponds to the envelope amplitude in dB obtained via the LTspice simulations for each configuration $\mathbf{x}_i$ of the uncertain parameters considered in a MC simulation with $N_{MC} = 10,000$ samples, whilst $\tilde{V}_{out,E}(f_k; \mathbf{x}_i)$ is the corresponding value estimate by a given surrogate. The above error is computed by using each of the considered surrogate built with an increasing number of training samples $L = 50, 150$ and 300.

From the curves of Fig. 3, the most accurate models are the ones based on the deterministic LS-SVM regression with linear and RBF kernel (solid and dashed green lines), however, the proposed modeling scheme based on the LS-SVM with linear and RBF kernel+GPR (see the solid and dashed red curves) provides the most accurate probabilistic surrogate models for all the considered set of training samples (i.e., $L = 50$, 150 and 300). Indeed, the curves related to the LS-SVM+GPR are usually below the ones related to standard GPR (solid, dashed and dotted blue lines). The results also show the improved convergence of the LS-SVM-based deterministic and probability models when a small set of training samples are available, i.e., $L = 50$, [16], [17].

As a further validation, Fig. 4 provides a graphical comparison between the scattering plots obtained by comparing the results of a MC simulation with 10,000 samples for all the $N_f = 91$ frequency points, with the corresponding ones provided by the deterministic surrogate models based on LS-SVM regression with linear kernel (green dots) and the mean values predicted by the proposed LS-SVM (RBF)+GPR

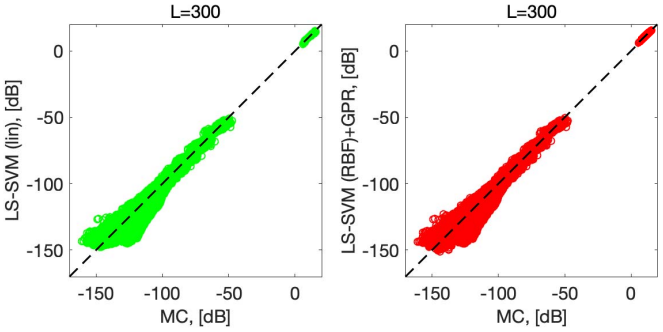

Fig. 4. Scatter plot (10,000 samples) comparing, for all the considered frequency points, the voltage spectral envelope $V_{out,E}$ predicted by the deterministic model based on the LS-SVM with linear kernel (green dots; left panel) and by the statistical surrogate model based on LS-SVM (RBF)+GPR (red dots; right panel) against the MC samples generated by the full-computational model.

statistical model (red dots). The plots highlight the capability of the two models to accurately predict the actual value of the MC simulation, since the samples are very close to the dashed line, which represents the perfect agreement between the model and the reference samples.

Also, Fig. 5 compares the probability density functions (PDFs) of the spectral envelope magnitude at $f_0 = 100\,\text{kHz}$ provided by the deterministic LS-SVM regression with linear kernel (solid green line) and the mean values of the LS-SVM (RBF)+GPR (solid red line) in (10a) with the histogram resulting from 10,000 MC samples (gray bins). The results highlight once again the excellent capability of the two models to capture the main feature of the reference PDF resulting from the MC simulations.

As a final comparison between the two models, Fig. 6 shows two realizations of the spectral envelope randomly selected among the results of the MC simulation (black curve) along with the corresponding predictions provided by both the deterministic surrogate model based on the LS-SVM with linear kernel (dashed green curve) and results of the proposed probabilistic model (i.e., the mean values and the 99% CI) based on the LS-SVM (RBF)+GPR model (red vertical bars) for $L = 300$. The results clearly highlight the advantages of the proposed modeling techniques with respect to the

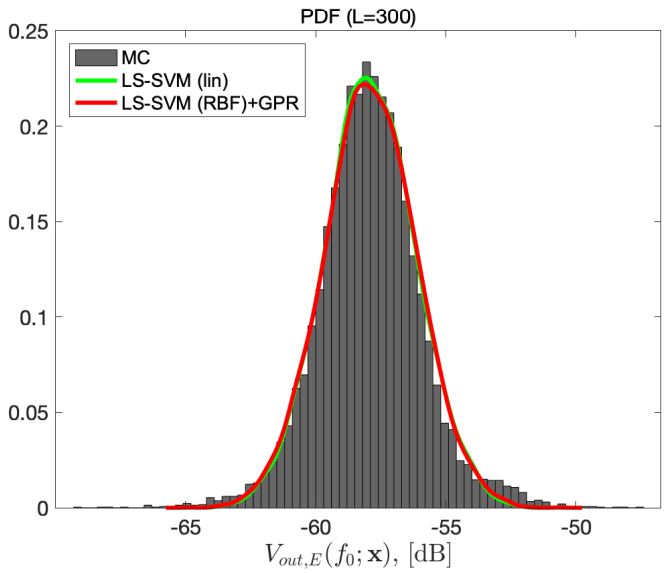

Fig. 5. Comparison among the PDFs computed for the realizations of the spectral envelop $V_{out,E}(f_0;\mathbf{x})$ at $f_0 = 100$ kHz obtained from the mean values of the probabilist model based on the LS-SVM (RBF)+GPR (solid red line) and the LS-SVM regression with linear kernel (solid green line) with the histogram of 10,000 MC samples (black bins).

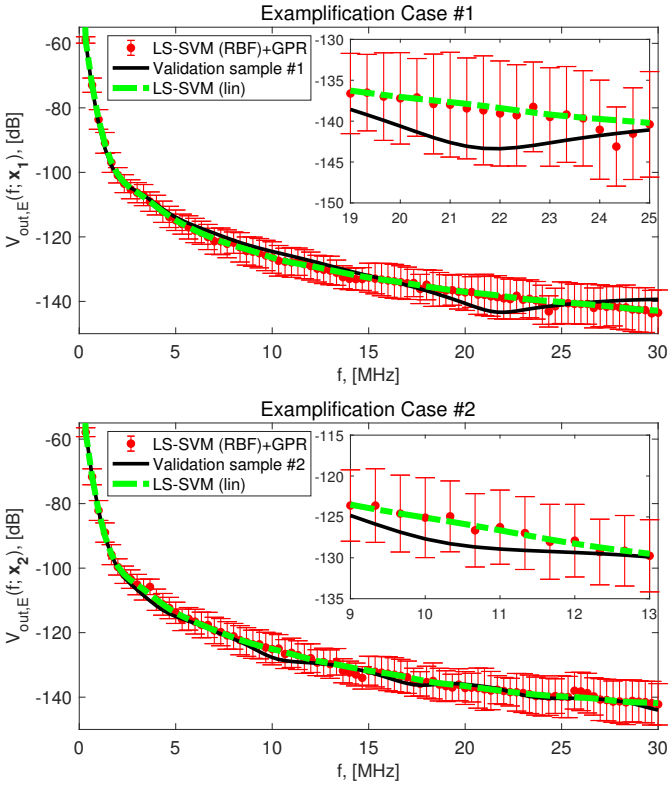

Fig. 6. Comparison between the envelope spectra for two different configurations of the converter parameters (black curve) randomly chosen among the 10,000 realizations of the MC simulation with the corresponding predictions of the deterministic model based on the LS-SVM with RBF kernel (dashed green curve) and the mean values and 99% CI estimated by the probabilistic models built via the LS-SVM (linear)+GPR (red vertical bars).

classical deterministic regression. In fact, different from the deterministic model based on the LS-SVM regression, for any given configuration of the input parameters, the proposed probabilistic one built via the LS-SVM+GPR does not only provide an approximation of the envelope spectra (red dots), but it also provides the users with a statistical information on the model error and reliability by means of the 99% CI (red error bars). The accuracy of such CI can be easily appreciated by noticing that the actual spectral envelope provided by the full-computational model (black curve) lays between the CI estimated by the proposed models. It is important to remark that the CI shown in Fig. 6 cannot be computed from the results of the MC simulation, since they are not related to any statistical quantity (e.g., statistical moments, quartiles, confidence limits, etc..) associated to the uncertain responses of the system under modeling. Indeed, the CI provide the user with a statistical information on the model error, only.

Table I provides a detailed summary of the computational cost required to build each of the considered surrogate models t_{model} and the computational cost t_{pred} required by each model to predict 10,000 realizations of the envelope spectra. All the simulations have been performed on a MacBook Pro with an Intel Core i5 CPU running at 3.1GHz and 16GB of RAM. The individual simulation with the full-computational model takes 1.7s, therefore the computational time required to generate the $L = 50, 150$ and 300 training samples is about 85 s, 255 s and 510 s. However, even for the maximum number of training samples (i.e., $L = 300$), the most complex surrogate model can be generated in less than 465 s (i.e., less than 8 min). The results highlight the advantage of the proposed technique based on the surrogate models with respect to the standard MC simulation, since each of the considered models allow predicting the envelope spectra for 10,000 realizations of the uncertain parameters in less than 30 s, while the corresponding MC simulation requires about 4 h 43 min.

IV. Conclusions

This paper deals with the development of a probabilistic model for the prediction of the spectral envelope of the output voltage waveform of a switching converter with a feedback network; 17 uncertain parameters related to the circuit components values are considered. The proposed technique relies on a two-step modeling scheme which combines the LS-SVM regression with the GPR. The accuracy of the resulting probabilistic model is then assessed by comparing its predictions with 10,000 MC simulations. For the sake of completeness, the results of the proposed model are then compared with the ones provided by five different surrogate models, such as the ones based on the deterministic LS-SVM regression with both linear and RBF kernel and the probabilistic models obtained via the standard GPR with constant, linear and polynomial (order 2) trend, respectively. From the results presented in this work, the proposed probabilistic models built with the LS-SVM+GPR can be considered as viable approaches for the development of accurate and fast-to-evaluate probabilist models for the prediction of the uncertain response of complex non-linear system in a high-dimensional parameter space.

TABLE I
COMPARISON AMONG THE COMPUTATIONAL COSTS NEEDED TO GENERATE THE MODELS AND TO EVALUATE 10,000 SAMPLES. REFERENCE IS THE MC SIMULATION.

Method	$L = 50$		$L = 150$		$L = 300$	
	t_{model}	t_{pred}	t_{model}	t_{pred}	t_{model}	t_{pred}
MC	–	4 h 43 min	–	4 h 43 min	–	4 h 43 min
LS-SVM (linear)	42 s	<1 s	45 s	<1 s	115 s	<1 s
LS-SVM (RBF)	51 s	<1 s	61 s	2 s	165 s	4 s
LS-SVM (linear)+GPR	52 s	2.5 s	129 s	10 s	414 s	23 s
LS-SVM (RBF)+GPR	59 s	3 s	131 s	12 s	465 s	26 s
GPR (constant trend)	13 s	2.2 s	73 s	10 s	304 s	19 s
GPR (linear trend)	10 s	2.3 s	73 s	10 s	294 s	19 s
GPR (poly trend order 2)	11 s	2.3 s	73 s	10 s	307 s	19 s

APPENDIX
GAUSSIAN PROCESS

A Gaussian process (GP) is a potentially infinite collection of random variables such that any finite subset of it has a joint multivariate Gaussian distribution. It can be considered as an extension of the concept of multivariate Gaussian distributions to infinite dimensionality [18].

A generic GP writes:

$$f(\mathbf{x}) \sim GP(\mu(\mathbf{x}), k(\mathbf{x}, \mathbf{x}')), \tag{15}$$

where $\mu(\mathbf{x}) : \mathbb{R}^d \to \mathbb{R}$ is a function defining the mean value (trend) of the GP and $k(\mathbf{x}, \mathbf{x}') : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ is the covariance function. The GP is completely characterized by the above quantities which are defined as:

$$m(\mathbf{x}) = E[f(\mathbf{x})], \tag{16a}$$

$$k(\mathbf{x}, \mathbf{x}') = E[(f(\mathbf{x}) - \mu(\mathbf{x}))(f(\mathbf{x}') - \mu(\mathbf{x}'))]. \tag{16b}$$

As an example let us considering the following random function [22]:

$$y(x) = b_0 + b_1 x + b_2 x^2, \tag{17}$$

where the coefficients b_1, b_2 and b_3 mutually independent Gaussian variable with $b_i \sim \mathcal{N}(0, \sigma_i^2)$ for $i = 0, 1, 2$. For any $x \in [-1, +1]$ the draws have zero mean, i.e.,

$$\begin{aligned} E[y(x))] &= E[b_0 + b_1 x + b_2 x^2] \\ &= E[b_0]x + E[b_1]x + E[b_2]x^2 \\ &= 0 + 0x + 0x^2 = 0, \end{aligned} \tag{18}$$

and covariance function:

$$\begin{aligned} k(y(x_1), y(x_2)) &= E[(b_0 + b_1 x_1 + b_2 x_1^2)(b_0 + b_1 x_2 + b_2 x_2^2)] \\ &= \sigma_0 + \sigma_1 x_1 x_2 + \sigma_2^2 x_1^2 x_2^2 = k(x_1, x_2). \end{aligned} \tag{19}$$

Because linear combinations of a fixed set of independent normal random variables have a multivariate normal distribution, given a set of values of the input parameter $x_1 \ldots, x_L$, the probability of getting the responses $[y(x_1), \ldots, y(x_L)]$ has a multivariate normally distribution, even if it is degenerate when $L > 4$ (i.e., the covariance matrix $\mathbf{K}$, computed form the covariance function $k(\cdot, \cdot)$ in (19), is not full-rank anymore). Without requiring any additional assumptions, the random function $y(x)$ in (17) can be modeled in more generic way as a Gaussian process, such as $y(x) \sim GP(0, k(x, x'))$.

REFERENCES

[1] Y. Song and B. Wang, "Survey on Reliability of Power Electronic Systems," in *IEEE Trans. Power Electron.*, vol. 28, no. 1, pp. 591–604, Jan. 2013.

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

[3] R. Trinchero, I. S. Stievano and F. G. Canavero, "Steady-state response of periodically switched linear circuits via augmented time-invariant nodal analysis," *J. Elect. Comput. Eng.*, vol. 2014, article ID 198273, 2014.

[4] R. Trinchero, I. S. Stievano and F. G. Canavero, "EMI Prediction of Switching Converters," *IEEE Trans. Electromagn. Compat.*, vol. 57, no. 5, pp. 1270–1273, Oct. 2015.

[5] R. Trinchero, I. S. Stievano and F. G. Canavero, "Steady-state analysis of switching power converters via augmented time-invariant equivalents," *IEEE Trans. Power Electron.*, vol. 29, no. 11, pp. 5657–5661, Nov. 2014.

[6] D. Xiu and G. E. Karniadakis, "The Wiener-Askey polynomial chaos for stochastic differential equations," *SIAM, Journal of Sci. Computation*, vol. 24, no. 2, pp. 619–622, 2002.

[7] P. Manfredi, I. S. Stievano and F. G. Canavero, "Stochastic Analysis of Switching Power Converters via Deterministic SPICE Equivalents," *IEEE Trans. Power Electron.*, vol. 29, no. 9, pp. 4475–4478, Sept. 2014.

[8] P. Manfredi, D. V. Ginste, I. S. Stievano, D. De Zutter, F. G. Canavero, "Stochastic transmission line analysis via polynomial chaos methods: an overview," *IEEE Electromagn. Compat. Mag.*, vol. 6, no. 3, pp. 77–84, Third Quarter 2017.

[9] G. Blatman and B. Sudret, "Adaptive sparse polynomial chaos expansion based on least angle regression," *Journal of Computational Physics*, vol. 230, no 6, pp. 2345–2367, 2011.

[10] M. Berveiller, B. Sudret and M. Lemaire, "Stochastic finite element: a non intrusive approach by regression," *European Journal of Computational Mechanics*, vol. 15, no. 3, pp. 81–92, 2006.

[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. Shalev-Shwartz and S. Ben-David, *Understanding Machine Learning: From Theory to Algorithms*, Cambridge University Press., Cambridge, 2014.

[13] V. Vapnik, *The Nature of Statistical Learning Theory*, 2nd edition, Springer, 1999.

[14] V. Vapnik, *Statistical Learning Theory*, John Wiley and Sons, NY, 1998.

[15] J. A. K. Suykens, *at al.*, "Least Squares Support Vector Machines", World Scientific Pub Co Inc, 2002.

[16] 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.
[17] R. Trinchero, M. Larbi, H. M. Torun, F. G. Canavero and M. Swaminathan, "Machine Learning and Uncertainty Quantification for Surrogate Models of Integrated Devices With a Large Number of Parameters," in *IEEE Access*, vol. 7, pp. 4056-4066, 2019.
[18] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*, MIT Press. Cambridge, Massachusetts, 2006.
[19] K.T. Fang, R.Z. Li, A. Sudjianto, *Design and Modeling for Computer Experiments*, Chapman and Hall/CRC, New York, 2006.
[20] E. Brochu, V. M. Cora, and N. de Freitas. (Dec. 2010). *A tutorial on bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning.* [Online]. Available: https://arxiv.org/abs/1012.2599.
[21] F. Lindsten, T. B. Schon, A. Svensson, and N. Wahlstrom, (2017). *Probabilistic modeling: linear regression & Gaussian processes*. Uppsala University, Department of Information Technology. [Online]. Available: http://www.it.uu.se/edu/course/homepage/sml/literature/-probabilistic_modeling_compendium.pdf
[22] T. J. Santner; B. J. Williams; W. I. Notz. *The design and analysis of computer experiments* , Springer, New York, N.Y., 2003.
[23] F. Perez-Cruz, S. Van Vaerenbergh, J. J. Murillo-Fuentes, M. Lazaro-Gredilla and I. Santamaria, "Gaussian Processes for Nonlinear Signal Processing: An Overview of Recent Advances," *IEEE Signal Processing Magazine*, vol. 30, no. 4, pp. 40–50, July 2013.
[24] H. M. Torun, M. Swaminathan, A. Kavungal Davis and M. L. F. Bellaredj, "A Global Bayesian Optimization Algorithm and Its Application to Integrated System Design," in *IEEE Trans. on Very Large Scale Integration (VLSI) Systems*, vol. 26, no. 4, pp. 792–802, April 2018.
[25] R. Schobi, B. Sudret and J. Wiart, "Polynomial-chaos-based Kriging," *Int. J. Uncertainty Quantification*, vol. 5, pp. 171–193, 2015.
[26] A. N. Shiryaev, *Probability*, 2nd ed. New York: Springer, 1996.
[27] R. Trinchero, M. Larbi, M. Swaminathan and F. G. Canavero, "Statistical Analysis of the Efficiency of an Integrated Voltage Regulator by means of a Machine Learning Model Coupled with Kriging Regression," *Proc. IEEE 23rd Workshop on Signal and Power Integrity (SPI)*, Chambry, France, 2019, pp. 1-4.
[28] LTspice ver. XVII. Accessed: March, 2019. Available: https://www.analog.com/en/design-center/design-tools-and-calculators/ltspice-simulator.html.
[29] LS-SVMlab, version 1.8; Department of Electrical Engineering (ESAT), Katholieke Universiteit Leuven: Leuven, Belgium, 2011. Available online: http://www.esat.kuleuven.be/sista/lssvmlab/.
[30] R. M. Neal, *Bayesian Learning for Neural Networks*. Springer, New York. Lecture Notes in Statistics, 118, 996.
[31] R. Muyshondt and P. T. Krein, "20 W benchmark converters for simulation and control comparisons," in *Record 6th Workshop on Computer in Power Electronics (Cat. No.98TH8358)*, pp. 201–212, Cernobbio, Italy, 1998,
[32] M. McKay, R. Beckman and W. 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.

Riccardo Trinchero (M'16) was born in Casale Monferrato, Italy in 1987. He received the M.Sc. and the Ph.D. degrees in Electronics and Communication Engineering from Politecnico di Torino, Torino, Italy, in 2011 and 2015, respectively.

He is currently an Assistant Professor within the EMC Group with the Department of Electronics and Telecommunications at the 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.

Flavio G. Canavero (SM'99-F'07) received his electronic engineering degree from Politecnico (Technical University) of Torino, Italy, and the PhD degree from the Georgia Institute of Technology, Atlanta, USA, in 1986. Currently he is a Professor of Circuit Theory with the Department of Electronics and Telecommunications, Politecnico di Torino. He is an IEEE Fellow. He has been the Editor-in-Chief of IEEE Transactions on Electromagnetic Compatibility, V.P. for Communication Services of the EMC Society and Chair of 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 EMC Society. His research interests include signal integrity and EMC design issues, interconnect modeling, black-box characterization of digital integrated circuits, EMI and statistics in EMC.