

© 2020 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

Machine Learning based Hybrid Random-Fuzzy Uncertainty Quantification for EMC and SI Assessment

Simon De Ridder *Student Member, IEEE*, Domenico Spina *Senior Member, IEEE*, Nicola Toscani,
Flavia Grassi *Senior Member, IEEE*, Dries Vande Ginste *Senior Member, IEEE* and Tom Dhaene *Senior
Member, IEEE*

Abstract—Modeling the effects of uncertainty is of crucial importance in the Signal Integrity (SI) and Electromagnetic Compatibility (EMC) assessment of electronic products. In this paper, a novel machine-learning based approach for uncertainty quantification problems involving both random and epistemic variables is presented. The proposed methodology leverages evidence theory to represent probabilistic and epistemic uncertainties in a common framework. Then, Bayesian optimization is used to efficiently propagate this hybrid uncertainty on the performance of the system under study. Two suitable application examples validate the accuracy and efficiency of the proposed method.

Index Terms—Bayesian Optimization, epistemic uncertainty, random-fuzzy problems.

I. INTRODUCTION

Resorting to statistical techniques for uncertainty quantification (UQ) and propagation (UP) within EMC and SI problems is common practice nowadays. In particular, RF designers have to cope with several sources of uncertainty, such as parasitic effects, non-ideal behavior of components, tolerances due to the manufacturing process, and so on. Even in relatively simple test setups several parameters are inherently unknown and/or hard to control. For these reasons, advanced statistical techniques have recently been applied to EMC and SI problems [1], [2], [3], [4], [5], [6], [7] with the objective to outperform the standard brute-force approach, based on Monte Carlo (MC) repeated simulations, in terms of computational efficiency, while retaining comparable accuracy in predicting the variability of the output variables.

All these techniques are firmly based on probability theory, i.e. on the *a priori* assumption that to every variable in the problem a probability distribution describing its variability over a given interval of values and/or around a nominal value can be assigned. This assumption is well-grounded when the uncertainty is caused by random effects. If the distribution of these random effects is known or can be estimated, a probabilistic treatment is a sound choice. However, when the uncertainty stems from a lack of knowledge about the value

of a parameter or its distribution, a probabilistic view falls short. Indeed, the standard approach to resort to a uniform distribution when a parameter is known to fall in a certain interval implies the arbitrary assumption that all values in that interval are equally probable, though without any evidence supporting this specific claim.

To overcome this intrinsic limitation of probability theory in providing an adequate representation of parameters affected by epistemic uncertainty, the more general framework of evidence theory or Dempster-Shafer theory was introduced [8]. In this framework, probability and possibility theories represent special cases. On the one hand, aleatory uncertainty due to random variability of parameters finds adequate description through the use of random variables (RVs) to which *probability* distribution functions (PDFs) are assigned. On the other hand, epistemic uncertainty due to lack of information is tackled through the use of fuzzy variables (FVs) to which *possibility* distribution functions are assigned.

Since typical problems usually involve both parameters affected by random and by epistemic uncertainty, hybrid probabilistic-possibilistic UP approaches have recently been developed and successfully applied in different engineering fields, spanning from risk management [9], [10] to measurement [11]. In an EMC context, the MC-based hybrid algorithm of [9], [10] was considered in a previous work [12], with the objective to estimate the radiated susceptibility of a non-ideally twisted wire pair (above ground) illuminated by a partially-unknown impinging electromagnetic (EM) field. Comparison with a fully probabilistic solution of the problem (where all uncertain parameters were assigned specific PDFs) revealed the potential of the aforementioned hybrid approach to provide a more general and exhaustive representation of the variability of the output variables. Nonetheless, this also unveiled the computational inefficiency of such an algorithm, which requires repeated-run MC simulations to assess the influence of RVs, and where for each MC run time-consuming grid-search sub-problems need to be solved to quantify the effect of the FVs. Hence, depending on the computational cost of the underlying deterministic model, which can be evaluated through either analytical expressions or full-wave EM simulations (as in the examples provided in this manuscript), this algorithm can become computationally prohibitive.

To tackle this limitation, this work proposes to use Bayesian optimization (BO) to deal with FVs. In this way, far fewer

S. De Ridder, D. Spina, T. Dhaene and D. Vande Ginste are with IDLab, Department of Information Technology, Ghent University-imec, 9000 Gent, Belgium. F. Grassi is with the Department of Electronics, Information and Bioengineering, Politecnico di Milano, 20133 Milano, Italy. N. Toscani is with the Department of Mechanical Engineering, Politecnico di Milano, 20156 Milano, Italy.

E-mail: simon.deridder@ugent.be

model evaluations are required to accurately propagate the epistemic uncertainty. The BO framework is hybridized with a state-of-the-art polynomial chaos (PC) approach to efficiently include the random variability. The computational efficiency and accuracy of the novel hybrid method is demonstrated by application to two realistic design examples. In particular, it is demonstrated that the proposed approach allows the quantification of combined possibilistic and probabilistic uncertainty of, among others, interconnects on printed circuit boards in spite of the time-consuming full-wave EM simulations required to describe their behavior. It is instructive to note that BO can also be used to optimize the design of electronic devices [13], [14], [15].

The manuscript organization is as follows. In Section II, basic elements of evidence theory are recalled with focus on possibility theory and fuzzy sets. Also, main features of the original hybrid probabilistic-possibilistic MC algorithm are described in this section. Section III introduces the novel BO framework, and explains how BO combined with Polynomial Chaos (PC) expansion is used to speed up the original hybrid algorithm. Numerical examples are presented in Section IV. Section V draws concluding remarks.

II. UNCERTAINTY PROPAGATION IN RANDOM-FUZZY PROBLEMS

This section introduces basic elements of evidence theory as a general framework for the representation of aleatory (probability theory) and epistemic (possibility theory) uncertainty. Furthermore, the main features of the MC-based algorithm [9], [10] are recalled, which allows for UP in hybrid problems involving both random variables (RVs) and epistemic fuzzy variables (FVs).

A. Basic Elements of Evidence Theory

Consider a set Σ that contains all possible values x that a variable X may assume. Its power set Ω consists of all possible subsets B_i of Σ , each of which can be assigned a probability weight $w(B_i)$, whose overall sum is unit over Ω . B_i represents a focal element of Ω if its weight $w(B_i)$ is non-zero. Consider now a subset $A \subseteq \Sigma$, called ‘event’. The evidence of the statement “ x may belong to A ” (i.e., to any of its subsets B_i) is expressed by (i) the Belief function (Bel)

$$\text{Bel}(A) = \sum_{B_i \subseteq A} w(B_i) \quad (1)$$

which measures the confidence we have in the statement based on the available information, i.e. the weights $w(B_i)$; and by (ii) the Plausibility function (Pl):

$$\text{Pl}(A) = \sum_{B_i \cap A \neq \emptyset} w(B_i) = 1 - \text{Bel}(\bar{A}) \quad (2)$$

which measures the conceivability of the statement based on the confidence in the occurrence of events that contradict the statement. It can be proven that Bel and Pl represent a limiting upper and lower cumulative distribution function

(CDF), encompassing all families of probability distributions $P(A)$ [16], i.e.:

$$\text{Pl}(A) \leq P(A) \leq \text{Bel}(A) \quad (3)$$

If the focal elements B_i are singletons, it is possible to assign weights, i.e., the traditional probabilities, to every element of Σ in a unique and precise way, and Bel and Pl degenerate into a single CDF P , that is $\text{Bel} = \text{Pl} = P$. As such, probability theory can be seen as a particular case of evidence theory.

Another particular case is obtained if, instead, the focal elements B_i are nested intervals, i.e.,

$$B_1 \subseteq B_2 \subseteq \dots \subseteq B_N, \quad (4)$$

where N denotes the number of focal elements. Then, a possibility distribution (PD) function $\pi(x) : \Sigma \rightarrow [0, 1]$ can be introduced as follows:

$$\forall x \in \Sigma | \pi(x) = \sum_{B_k \ni x} w(B_k). \quad (5)$$

$\pi(x) = 0$ corresponds to a definitely impossible value and $\pi(x) = 1$ to a totally possible value. In the case of (4), the Belief function (1) and Plausibility function (2) are called *necessity* (N) and *possibility* (Π) functions respectively, and they reduce to

$$N(A) = 1 - \sup_{x \in \bar{A}} \pi(x) \quad (6)$$

$$\Pi(A) = \sup_{x \in A} \pi(x). \quad (7)$$

In Figs. 1(a) and 1(b) two examples of typical possibility (PD) distribution functions, i.e. a rectangular distribution π_1 and a triangular distribution π_2 respectively, are shown for the case where $\Sigma = \mathbb{R}$. In the case of Fig. 1(a), all values x in the interval $A = [x_1, x_2]$ are perfectly possible, whereas all other values, in \bar{A} , are impossible. This information is, on the one hand, more vague than the information provided by a uniform probability density function, as it represents a total ignorance about x in that interval $A = [x_1, x_2]$. It is only known that these values are all totally possible. On the other hand, the information provided by π_1 , is more complete, as it does not make any assumptions; in contrast, the uniform probability density function assumes all values to be equally probable in $[x_1, x_2]$. Consequently, all possible cumulative density functions $P(A)$, assigned to the interval $A = [x_1, x_2]$ (and including the uniform CDF), are bounded by the possibility $\Pi_1(A)$ and necessity $N_1(A)$ functions, i.e. $N_1(A) \leq P(A) \leq \Pi_1(A)$. These functions are shown in Fig. 1(c). As opposed to the total ignorance case of Fig. 1(a) and (c), the triangular PD $\pi_2(x)$ of Fig. 1(b), with corresponding N_2 and Π_2 (see Fig. 1(d)), indicates a higher degree of confidence in one value of x , where $\pi_2(x) = 1$, and a gradual decrease for all other values in A .

The mathematical framework to deal with PDs is the theory of fuzzy sets [17]. Within this framework, the focal elements B_i are nested intervals obtained by cutting the PDs at different levels α in $[0, 1]$. For example, in Fig. 1(b), $\alpha = 0.3$ identifies the interval $[c_1, c_2]$, while $\alpha = 0.8$ leads to $[d_1, d_2]$. Each such

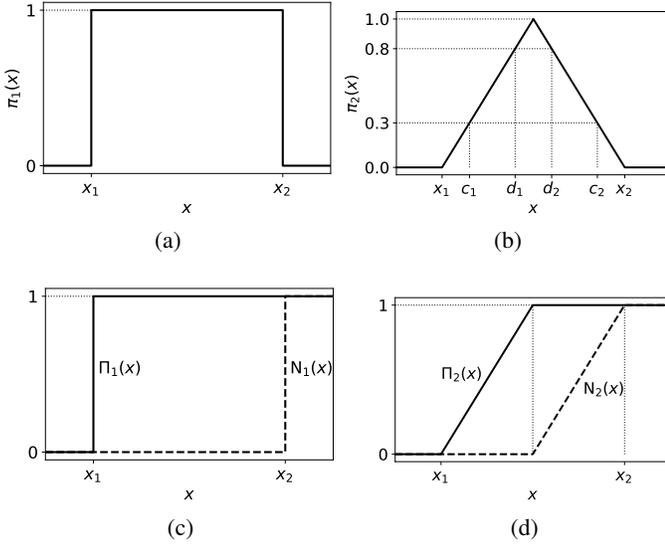


Figure 1: Rectangular (a) and triangular (b) possibility distributions with their corresponding possibility $\Pi(x)$ (solid) and necessity $N(x)$ (dashed) measures in (c) and (d), respectively. Here we used the shorthand $\Pi(x)$ and $N(x)$ to denote $\Pi([-\infty, x])$ and $N([-\infty, x])$.

interval $A_\alpha = [\inf_\alpha, \sup_\alpha]$ is called an α -cut and the entire PD is determined by the knowledge of the extreme values \inf_α and \sup_α of a sufficient amount of α -cuts. Moreover, each α -cut corresponds to a confidence interval with necessity $N(A_\alpha)$ as follows:

$$N(A_\alpha) = 1 - \sup_{x \notin A_\alpha} \pi(x) = 1 - \alpha \quad (8)$$

B. Hybrid Random-Fuzzy MC-based Algorithm

Owing to their different meaning and mathematical representation, UP of parameters affected by epistemic (FVs) and aleatory (RVs) uncertainty follows different rules. Hence, several approaches were developed for UP within hybrid problems involving both RVs and FVs. Among these, the MC-based hybrid algorithm of [10] along with the aggregation method of [9], [18] is first described here. Next, in Section III, a novel hybrid BO-PC framework is introduced, strongly optimizing the MC-based algorithm, as such allowing for the EMC and SI assessment of full-wave EM problems.

MC-based hybrid algorithm: Assume z to be defined as a function of K random variables x_1, \dots, x_K and of J epistemic variables y_1, \dots, y_J , i.e., $z = f(x_1, \dots, x_K, y_1, \dots, y_J)$. As can be seen in Fig. 2, the uncertainty of z is evaluated by generating N_x random realizations of the K random variables. For each i -th realization ($i = 1, \dots, N_x$), the i -th possibility distribution function π_i^z of z with respect to the J epistemic variables is constructed. Thereto, a preset number of values for α , ranging from 1 to 0, is chosen. These determine the α -cuts of the J epistemic variables. For each α -value, z is evaluated on a dense grid in the space of the epistemic variables. The extreme values (i.e., \inf_α and \sup_α) of z are then identified by taking the minimum and maximum of these evaluations, and they are assigned as lower and upper limits of the corresponding α -cut of z . As such, these α -cuts yield the

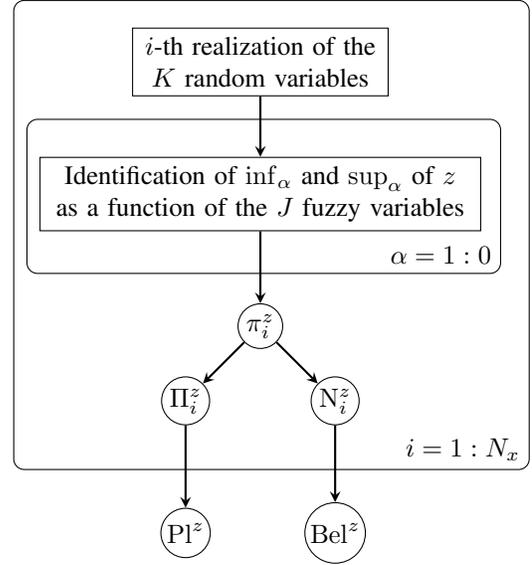


Figure 2: Schematic representation of the hybrid random-fuzzy approach in both this paper and [10]. The indexed rectangles indicate loops: the outer one over the N_x realizations of the RVs, the inner one over the α -cuts for the FVs.

possibility distribution π_i^z of z for the i -th realization of the K RVs. Then, repeating the above procedure for the N_x RV realizations, a set of N_x possibility distributions $\pi_1^z, \dots, \pi_{N_x}^z$ is obtained. N_x corresponding possibility Π_i^z and necessity N_i^z functions are constructed. Finally, to aggregate the results, Pl^z and Bel^z of z are computed as the weighted averages:

$$Pl^z = \frac{1}{N_x} \sum_{i=1}^{N_x} \Pi_i^z, \quad (9)$$

$$Bel^z = \frac{1}{N_x} \sum_{i=1}^{N_x} N_i^z. \quad (10)$$

These Pl^z and Bel^z can again be interpreted as limiting upper and lower bounds, encompassing all possible CDFs of z , i.e.

$$Bel^z \leq P(z) \leq Pl^z \quad (11)$$

III. BAYESIAN OPTIMIZATION FRAMEWORK

The brute force implementation of the hybrid algorithm presented in Section II-B easily becomes computationally prohibitive even in the presence of few uncertain parameters, due the large amount of dense grid evaluations required to find \inf_α and \sup_α for each α -value. This is especially true for EM problems where numerical full-wave simulations are required in order to achieve solution of the underlying deterministic model.

To alleviate the computational burden, BO is introduced in this section as a way to obtain these \inf_α and \sup_α for each α -value in a significantly more efficient way. When combined with PC for the RVs, the number of simulations is reduced even further.

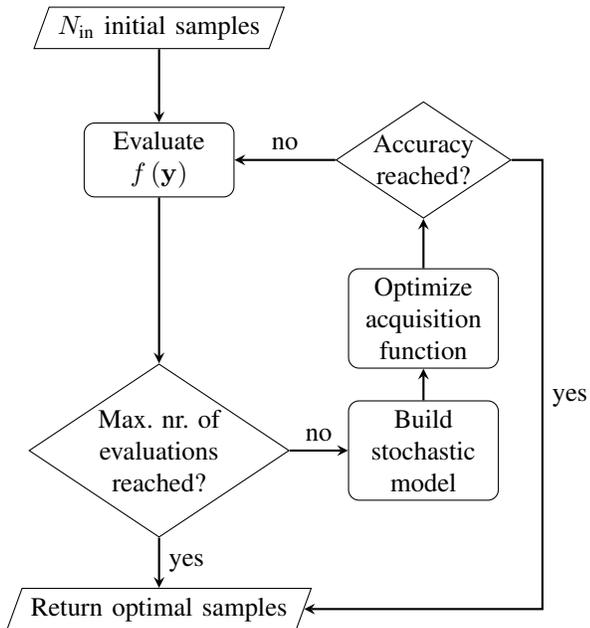


Figure 3: Flowchart of a BO algorithm.

A. General characteristics of Bayesian Optimization

BO aims to solve global optimization problems of the form

$$f_{\min} = \min_{\mathbf{y}} [f(\mathbf{y})], \quad (12)$$

where $f(\mathbf{y})$ is the quantity to be minimized (or maximized), also called cost or objective function in the BO framework, and depends on a set of parameters collected in the vector \mathbf{y} . In particular, $f(\mathbf{y})$ can be considered a “black box” function: it does not have to be known *a priori*, but can be evaluated for any combination of its parameters \mathbf{y} through simulations or measurements. BO is particularly useful when the objective function is expensive to evaluate and dynamic with respect to its parameters, causing $f(\mathbf{y})$ to have several local optima [19].

The general methodology of BO is summarized in Fig. 3. First, the cost function $f(\mathbf{y})$ is evaluated through simulations or measurements for an initial set of N_{in} samples \mathbf{y}_i , $i = 1, \dots, N_{\text{in}}$, chosen e.g. according to a Latin hypercube.

With these initial samples, a suitable stochastic surrogate model is then built. This model serves as a cheap to evaluate substitute for $f(\mathbf{y})$, which is computationally easy to optimize. The predictions of this model are stochastic, and because of this, the uncertainty of the model can be used to select the next point to evaluate. Several mathematical techniques can be adopted to build the desired surrogate model, among which are Bayesian neural networks [20] and Gaussian Processes (GPs) [21], [22].

The selection of the point to evaluate next is generally done by optimizing an acquisition function. Such an acquisition function transforms the model predictions into a measure that identifies where the next evaluation should be in order to obtain a global optimum (12) efficiently. Among the most widely adopted acquisition functions are Expected Improvement (EI) [23], [24] and Probability of Improvement (PoI) [25].

If the selected point is very close to the previously evaluated point, convergence is assumed and the optimization is finished. If not, the suggested point is evaluated through a new (expensive) simulation. As long as the computational budget is not spent, the stochastic surrogate model is then updated, another point is selected, and so on.

For a more complete treatment of BO, the interested reader is referred to [19], [24], [26].

B. Bayesian Optimization for the possibilistic part of the hybrid algorithm

In light of the hybrid algorithm detailed in Section II-B, BO is particularly suited to replace the grid-search and to optimize the solution of the possibilistic sub-problems, where the \inf_{α} and \sup_{α} of all α -cuts must be identified for each RV realization. To this end, the widely popular GP regression model [21], [22] is chosen as the stochastic surrogate model, due to its accuracy and modeling power. A GP can be considered a generalization of the multivariate Gaussian distribution; the latter is a distribution over a finite set of random variables, completely specified by a mean vector and covariance matrix, whereas a GP considers an infinite, parameterized set of RVs. As such it can be interpreted as a distribution over functions that is fully characterized by its mean function and covariance function [21], [22]. Often, and also in this paper, the mean is chosen to be zero, as the data can be normalized, and no trend is assumed to be known beforehand. As the covariance function (also called kernel), a Matérn (5/2) kernel is used, which has the form:

$$K_{M,5/2}(\mathbf{y}, \mathbf{y}') = \sigma^2 \left(1 + \frac{\sqrt{5}d}{\rho} + \frac{5d^2}{3\rho^2} \right) \exp \left(-\frac{\sqrt{5}d}{\rho} \right), \quad (13)$$

where \mathbf{y} and \mathbf{y}' are two vectors of length J , representing two points in the space of FVs, and $d = \|\mathbf{y} - \mathbf{y}'\|$. This kernel is adopted due to its capability to model a wide class of functions (including non-differentiable ones) [21], [22]. The hyper-parameters σ^2 and ρ are optimized using maximum likelihood estimation, as implemented by the GPyOpt package [27], or the GPFlowOpt package [28].

Regarding the sampling strategy, the acquisition function chosen in this paper is EI [23], [24]. For pure minimization problems, EI is defined as follows. The current evaluated point with minimum value is called f_{\min} . Improvement at a point \mathbf{y} , denoted $I(\mathbf{y})$, is equal to $f_{\min} - f(\mathbf{y})$ if $f(\mathbf{y})$ is smaller than f_{\min} , and 0 otherwise. Since a GP is a stochastic surrogate model, its prediction t at location \mathbf{y} , is not deterministic, but is a Gaussian RV, whose mean represents the estimated value of $f(\mathbf{y})$. Therefore, the expected value of the improvement can be written as:

$$\mathbb{E}[I(\mathbf{y})] = \mathbb{E}[\max(f_{\min} - t, 0)]. \quad (14)$$

The expectation in (14) and the corresponding gradient can be calculated analytically, making optimization of EI computationally cheap.

Since for each α -cut of the J FVs (see Fig. 2) both minimum and maximum need to be found, the following

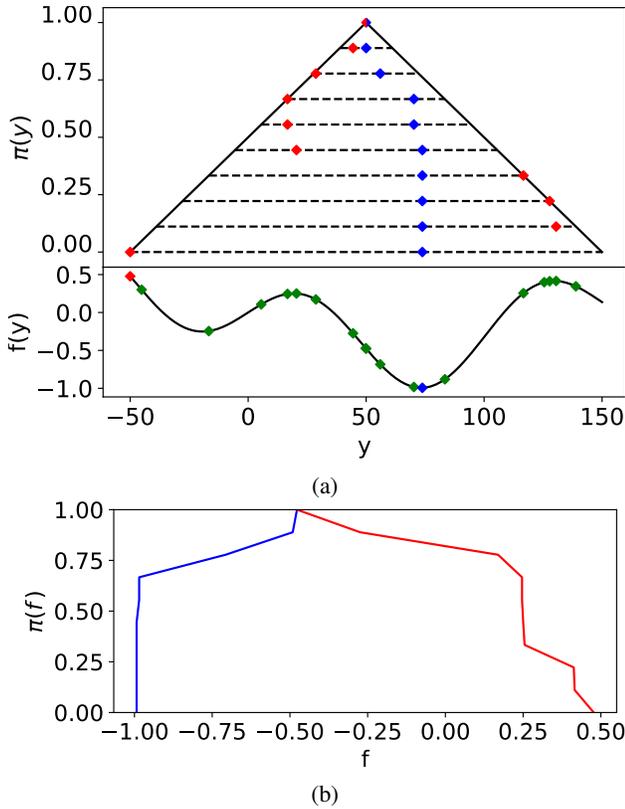


Figure 4: Illustration of the “top-down” optimization strategy for a one-dimensional analytical example with 10 α -cuts, and 19 samples in total as a computational budget. In (a), the red diamonds are the maxima found in each α -cut (or the global maximum on the curve), while the blue diamonds represent the found minima in each α -cut. The green diamonds on the curve are evaluated samples that are not optimal. The dashed lines represent each of the ten α -cuts, while the black lines delimiting them describe the possibility distribution. (b) shows the obtained possibility distribution, using all evaluated samples. The two halves of this curve are plotted in blue and red, to show that they are obtained from the minimum and maximum, respectively, at each α -cut.

modified version of EI is adopted as an acquisition function here:

$$EI_{\text{mm}}(\mathbf{y}) = \max(\mathbb{E}[\max(f_{\min} - t, 0)], \mathbb{E}[\max(t - f_{\max}, 0)]). \quad (15)$$

This is the maximum of: EI for minimization; and EI for maximization. Hence, the point where, according to the model, most potential for a better optimum of either kind can be found, will be evaluated. Thus, the proposed optimization strategy finds both optima with a minimal amount of evaluations.

It is important to remark that both the minimum and the maximum of the cost function must be computed for *each* α -cut. This is a trivial task for FVs with a rectangular possibility distribution as shown in Fig. 1(a), as then all α -cuts are the same. Still, a suitable strategy must be defined when dealing with different possibility distributions, e.g., the one of Fig. 1(b). To this end, considering that α -cuts are nested by definition, a “top-down” approach is proposed here. Fig. 4(a) displays the proposed “top-down” strategy for a toy example.

The analytical function under study is

$$f(y) = \cos\left(\frac{x}{23}\right) \sin\left(\frac{x}{50}\right). \quad (16)$$

This function depends on one FV y with a triangular distribution $\pi(y)$. BO is first carried out for the top α -cut, corresponding to $\alpha = 1$, with a small portion of the total computational budget (though in the one-dimensional triangular case of Fig. 4, no BO is needed as the domain consists of only one point). Then, one by one, the objective function is also optimized over the α -cuts corresponding to lower possibility values, making use of the samples already evaluated in the α -cuts above them, while allowing a few additional samples to be evaluated for each new α -cut. The PD of $f(y)$, denoted $\pi(f)$, constructed in this way, is shown in Fig. 4(b). Note that, in post-processing, a better optimum found at lower α can replace an optimum found at higher α , as long as it falls within the bounds defined by that α -cut.

C. Hybridization of BO with PC

BO successfully expedites the solution of the epistemic sub-problems. Nonetheless, the slow convergence of the MC sampling, to account for the aleatory variability of the RVs, renders the hybrid algorithm still slow. Therefore, in this paper, PC expansions [29] are adopted, given their accuracy and efficiency in characterizing stochastic variations [2], [3], [4], [30], [31], [32], [33]. Using the PC expansion, a suitable model is computed for both the minimum and maximum depending on the RVs and the α -cut, in the form:

$$F_{\min}(\alpha, \mathbf{x}) = \sum_{k=0}^K \beta_{\min,k}(\alpha) \phi_k(\mathbf{x}),$$

$$F_{\max}(\alpha, \mathbf{x}) = \sum_{k=0}^K \beta_{\max,k}(\alpha) \phi_k(\mathbf{x}), \quad (17)$$

where $\phi_k(\mathbf{x})$, for $k = 1, \dots, K$, are suitable orthogonal polynomials depending on the randomly distributed variables collected in the vector \mathbf{x} , while $\beta_{\min,k}(\alpha)$ and $\beta_{\max,k}(\alpha)$ are the corresponding PC coefficients, which depend on the α -cut [29].

It is important to remark that the PC basis functions in (17) depend only on the distribution of the RVs \mathbf{x} , and can be determined upfront for different distributions [33]. Thus, only the coefficients $\beta_{\min,k}$ must be estimated, as described in the following. First, a number of samples $[\mathbf{x}_i]_{i=1}^{N_{PC}}$ (also called collocation points) are chosen in the random variables’ space according to the method described in [31]. At each of these points, BO is performed in order to find the minima and maxima of the cost function for the different α -cuts. Next, the desired PC coefficients can be computed by solving a suitable linear system [31].

F_{\min} and F_{\max} at each of the α -cuts and for a specific instance of the random variables \mathbf{x}_j form the α -cuts for z itself, and thus define its possibility distribution π_j^z . As such, an MC sampling of N_x instances of the random variables, evaluated using F_{\min} and F_{\max} (17), is used to construct many such possibility distributions. These can then be combined using (9) and (10), to obtain Pl^z and Bel^z (see Fig. 2). Since

the models in (17) are analytical, they can be evaluated with great efficiency compared to a direct sampling of the cost function.

IV. NUMERICAL EXAMPLES

The proposed methodology is applied to two distinct numerical examples, of increasing complexity. The first example is an analytical one, constitutes a typical EMC setup and is used to illustrate and validate the proposed methodology. The second case study, instead, represents a relevant SI application example where full-wave EM simulations are required to estimate the chosen cost function.

A. Twisted wire pair

In this first example, the proposed methodology is applied to the radiated susceptibility problem illustrated in Fig. 5 and also adopted in [12]. Two lossless wires are wound into a double helix, $h = 5$ cm above a perfectly conducting ground plane. The line length is $l = 200$ cm. The terminal loads are the same both at the left and at the right terminations, and are balanced to prevent mode conversion. A uniform plane wave (1 V m^{-1}) with polarization angle η , azimuth angle ψ and elevation angle θ illuminates the twisted wire pair.

In this example, the goal is to estimate the variation of the induced differential mode (DM) noise current with respect to polarization η and azimuth angle ψ of the incident field. Since positioning of the irradiating antenna is affected by random uncertainty, they are considered Gaussian RVs: $\eta \sim \mathcal{N}(90^\circ, 2^\circ)$ and $\psi \sim \mathcal{N}(90^\circ, 2^\circ)$ (where $\mathcal{N}(\mu, \sigma)$ is the Gaussian distribution with mean μ and standard deviation σ).

The elevation angle θ , as suggested in aerospace standards, is usually set to 73° [34]. However, the test operator is allowed some flexibility in setting this angle, and thus a triangular distribution [$68^\circ, 78^\circ$] [12] is chosen to represent the epistemic uncertainty for this parameter.

The number of twists N_t can be treated as a real number to also indicate a fraction of an incomplete twist at the end of the structure. It has been shown that such an incomplete twist produces a similar variability to that of a non-uniformity in the twists [12], [35]. This justifies approximating the effect of an unknown non-uniformity by imposing a uniform possibility distribution on N_t , more specifically in the range [40, 41].

The DM current (which is the $z = f(\mathbf{x}, \mathbf{y})$ for this example, with $\mathbf{x} = [\eta, \phi]$ and $\mathbf{y} = [\theta, N_t]$) can be analytically computed for any frequency value of interest by means of the deterministic model described in [36].

In order to compute the PC models (17), the proposed BO is carried out for $K = 10$ samples of the random variables (η, ψ) . In particular, 51 α -cuts, ranging from possibility level 1 to 0, are considered for each collocation point. The BO starts at $\alpha = 1$ with 3 initial points in the range of $N_t \in [40, 41]$ (since the distribution of θ is triangular) to construct a GP model. A maximum of (in total) 10 samples is allowed for the BO at this α -cut. In each following α -cut, the bounds of the optimization are widened according to the possibility distributions of N_t and θ and an additional 1.8 samples are added to the maximum total budget for BO (at each step the

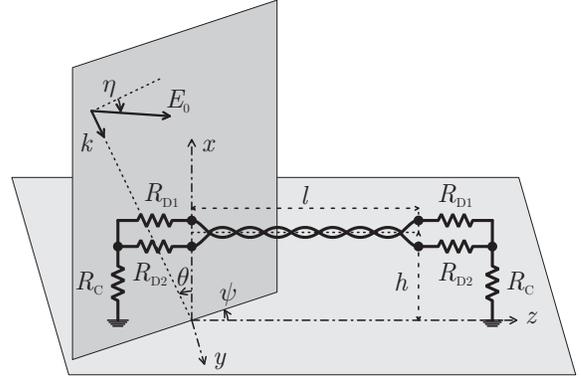


Figure 5: Twisted wire pair. $R_C = 100 \Omega$, $R_{D1} = R_{D2} = 50 \Omega$.

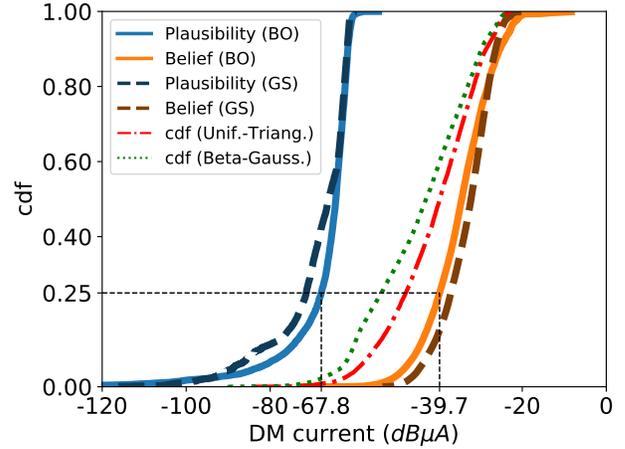


Figure 6: Twisted wire pair: Plausibility and Belief functions of the DM current computed through PC (random variables) and BO (epistemic variables) are shown in thick blue and orange lines. The same functions computed with the approach presented in [12] are displayed in thick dashed dark blue and brown lines. Examples of CDFs obtained by treating all variables as probabilistic are shown in thinner lines. With N_t distributed as a uniform distribution between 40 and 41 and θ following a triangular distribution from 68 to 78 with mode at 73, the CDF given by the dash-dotted red line is obtained. If N_t follows a Beta distribution with Beta $(\frac{1}{2}, \frac{1}{2})$ between 40 and 41, and if θ follows a Gaussian distribution $\mathcal{N}(73, 2)$, truncated between 68 and 78, the green dotted CDF is calculated.

budget is rounded down), summing to a total maximum of 100 samples over all α -cuts.

Next, a PC model (with polynomials up to third order) as a function of the random variables is calculated for both the minimum and maximum of the DM current, estimated via BO for each α -cut. Finally, a large number of MC samples are drawn from the RVs distributions, and the PC models are evaluated for all of them. The obtained results are then aggregated using (9) and (10) in terms of P1 and Bel functions. The uncertainty quantification results for the frequency of 30.54 MHz are displayed in Fig. 6.

For comparison, the same example was also solved with the original implementation of the algorithm [12], that is by making use of MC sampling for the RVs and grid search (GS) (i.e., sampling on a pre-determined grid formed by 101×101

samples) for the possibilistic sub-problems. The corresponding Pl and Bel functions (aggregated by (9) and (10)) are also shown in Fig. 6. Clearly, the proposed method offers a comparable accuracy with respect to the reference one. However, fewer than 1000 samples were simulated for the proposed method, corresponding to fewer than 100 (θ, N_t) samples for each of the 10 (η, ψ) collocation points, while the reference approach required to simulate 10201000 samples, given by 10201 (θ, N_t) samples for each of the 1000 (η, ψ) MC sample points. Due to the availability of an analytical solution and vectorization speedups, the grid search approach takes 5 min and 35 s, while the proposed approach takes 39 min. Ignoring the vectorization speedups and sequential function call overhead, the difference amounts to an average of about 2 s per model update and acquisition-function optimization, as an indication. Note that in the next subsection, these vectorization speedups will no longer be possible, as the solution of each deterministic problem will require a full-wave simulation.

In general, whenever simulation of the device under study is very fast (e.g. with an analytical model), a vectorized MC-based approach can be faster than the proposed sequential approach, due to the overhead of building and updating the GP models. However, when the simulations are the bottleneck, as is often the case for full-wave simulations, the proposed method can deliver significant speedups, as evidenced in Section IV-B.

Note that Fig. 6 also features two CDFs, obtained by treating N_t and θ as RVs. This demonstrates that the obtained Pl and Bel functions can serve as lower and upper bounds to any possible CDF of the DM current. This can be interpreted as follows. At any value z^* of the DM current, the belief function is a lower bound to the percentage of instances of the setup that are smaller than z^* . For example, at least 25% of the setups lead to a DM current of less than -39.7 dB μ A. Similarly, the plausibility function is an upper bound to the percentage of instances of the setup that are smaller than z^* , or alternatively, $1 - \text{Pl}$ is a lower bound to the percentage of instances of the setup that are larger than z^* . For example, maximally 25% of the setups lead to a DM current of less than -67.8 dB μ A or, differently put, minimally 75% of the setups lead to a DM current of more than -67.8 dB μ A.

When implemented on a server node with a Quad-Core AMD Opteron Processor (2 GHz) and 31.4 GiB DDR2 RAM, the

B. Bent microstrip lines

As a second example, the proposed framework is applied to two bent coupled microstrip lines making a 90° bend (see Fig. 7). This layout was introduced in [37].

For this example, four independent RVs are considered: the lengths of both of the line segments, $l_1 \sim \mathcal{N}(50 \text{ mm}, 0.1 \text{ mm})$ and $l_2 \sim \mathcal{N}(50 \text{ mm}, 0.1 \text{ mm})$, and the width of each line, $w_1 \sim \mathcal{N}(1.8 \text{ mm}, 0.1 \text{ mm})$ and $w_2 \sim \mathcal{N}(1.8 \text{ mm}, 0.1 \text{ mm})$. Since they are only specified by manufacturers within certain tolerances, the relative permittivity ϵ_r and the height h of the substrate are treated as FVs with rectangular possibility distributions with supports $[3.36, 3.96]$ and $[1.374 \text{ mm}, 1.674 \text{ mm}]$,

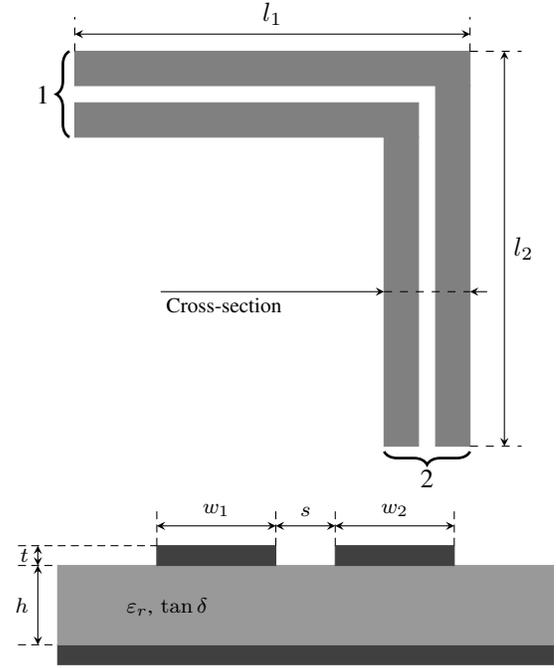


Figure 7: Bent microstrip lines: The top schematic shows the layout of the bent microstrip lines, demarcating the line lengths l_1 and l_2 . The differential signaling ports are also indicated using braces. The cross-section of the bent microstrip lines is shown in the bottom schematic, demarcating remaining parameters. The fixed parameters are $s = 0.7$ mm, $t = 35$ μ m, $\tan \delta = 0.003$ and the line conductivity is 4.1×10^7 S m^{-1} .

respectively. Given that mode conversion is especially critical for bent interconnects, we construct an objective function that is a measure of the total DM to common mode (CM) conversion [38], as follows:

$$C = \left[\int_{0 \text{ GHz}}^{6 \text{ GHz}} (|S_{cd11}(f)|^2 + |S_{cd21}(f)|^2) df \right]^{1/2}, \quad (18)$$

where S_{cd11} and S_{cd21} are the relevant elements of the modal S -parameters matrix. In contrast to the previous example, no analytical model for the S -parameters nor the cost function is available. For a specific sample of the random and epistemic variables, the cost function is computed by first acquiring the S -parameters in the frequency range of interest [0-6 GHz] by means of full wave simulations, performed using Advanced Design System (ADS) [39]. Then, the modal scattering parameters are computed [37], and the integral in (18) is calculated using standard numerical techniques [38].

PC models with $K = 35$ (with polynomials up to third order) of the minimum and maximum of the cost function were constructed from the optima found by BO in the collocation points. In particular, since the FVs follow a rectangular possibility distribution, BO only needs to be carried out once for each of the collocation points. A maximum of 50 samples in the space of the epistemic variables was allowed.

Evaluating these PC models in 1000 Monte Carlo sampled points leads to an aggregated distribution for the cost function C , as indicated in Fig. 8. As an example, at least 21%

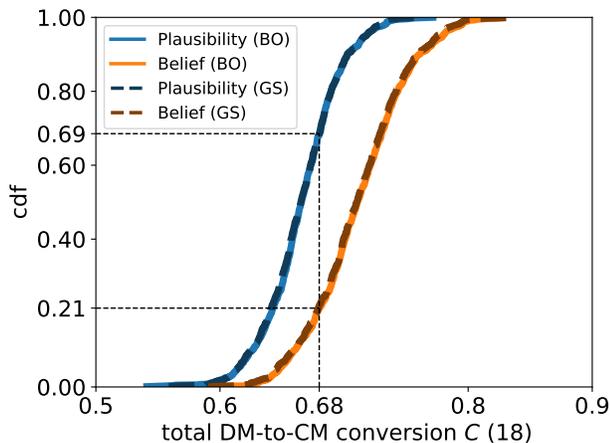


Figure 8: Bent microstrip lines: Plausibility and Belief functions for the cost function C (18), obtained through PC (random variables) and BO (epistemic variables) are shown in blue and orange lines. The same functions computed with the approach presented in [12] are displayed in dashed dark blue and brown lines.

Table I: Bent microstrip lines: Computational cost of the Uncertainty Quantification.

	PC + BO	MC + grid search
Total number of simulations	696	100 000
Total computational time	17 h, 57 min	112 d, 16 h, 3 min

of the samples has a total DM-to-CM conversion of less than 0.68, but no more than 69%.

For comparison, a grid search (GS) with 10×10 points for the epistemic variables was also carried out in each of these 1000 Monte Carlo points, and the corresponding aggregated results are also shown in Fig. 8. The Pl and Bel functions obtained with both methods are in excellent agreement, while the proposed BO-based approach, as displayed in Table I, is much more efficient. This speedup makes a hybrid possibilistic-probabilistic approach as feasible as a fully probabilistic one to electronics designers, while retaining the advantage of treating epistemic variables in a possibilistic manner. Both methodologies were implemented on a server node with a Quad-Core AMD Opteron Processor (2 GHz) and 31.4 GiB DDR2 RAM.

The proposed method ostensibly opens the door to perform a hybrid possibilistic-probabilistic analysis on real-world data. Because both PC and especially BO scale quite well when increasing the number of dimensions, problems with a higher number of stochastic dimensions can be tackled as well. However, we have not included such examples in this manuscript as a comparison with the method outlined in Section II.B, based on MC and GS, would become infeasible.

V. CONCLUSIONS

A hybrid novel machine-learning based approach for the UQ of the EMC and SI behavior of electronic problems prone to both probabilistic and epistemic variability is presented in this paper. The method leverages evidence theory as a general framework to deal with both probabilistic and possibilistic

definitions of uncertainty. Bayesian optimization based on Gaussian process regression and a modified version of the Expected Improvement acquisition function are used to solve uncertainty propagation problems with accuracy and efficiency, whenever expensive full-wave simulations are required. Two relevant application examples, a twisted wire pair and a bent differential microstrip, have been used to validate the proposed approach, demonstrating its usefulness in providing bounds for the probability of events.

REFERENCES

- [1] F. Paladian, P. Bonnet, and S. Lall ch re, "Modeling complex systems for EMC applications by considering uncertainties," in *2011 XXXth URSI General Assembly and Scientific Symposium*, Aug 2011, pp. 1–4.
- [2] T. Pham, E. Gad, M. S. Nakhla, and R. Achar, "Decoupled polynomial chaos and its applications to statistical analysis of high-speed interconnects," *IEEE Transactions on Components, Packaging and Manufacturing Technology*, vol. 4, no. 10, pp. 1634–1647, Oct 2014.
- [3] D. Spina, F. Ferranti, T. Dhaene, L. Knockaert, G. Antonini, and D. Vande Ginste, "Variability analysis of multiport systems via polynomial-chaos expansion," *IEEE Transactions on Microwave Theory and Techniques*, vol. 60, no. 8, pp. 2329–2338, Aug 2012.
- [4] Y. Ye, D. Spina, P. Manfredi, D. Vande Ginste, and T. Dhaene, "A comprehensive and modular stochastic modeling framework for the variability-aware assessment of signal integrity in high-speed links," *IEEE Transactions on Electromagnetic Compatibility*, vol. 60, no. 2, pp. 459–467, April 2018.
- [5] Z. Fei, Y. Huang, J. Zhou, and Q. Xu, "Uncertainty quantification of crosstalk using stochastic reduced order models," *IEEE Transactions on Electromagnetic Compatibility*, vol. 59, no. 1, pp. 228–239, Feb 2017.
- [6] Y. Wang, S. Jin, S. Penugonda, J. Chen, and J. Fan, "Variability analysis of crosstalk among differential vias using polynomial-chaos and response surface methods," *IEEE Transactions on Electromagnetic Compatibility*, vol. 59, no. 4, pp. 1368–1378, Aug 2017.
- [7] B. Chen, M. Ouyang, S. Yong, Y. Wang, J. Wang, S. Jin, Y. Bai, Y. Zhou, and J. Fan, "Differential integrated crosstalk noise (icn) reduction among multiple differential bga and via pairs by using design of experiments (doe) method," in *2017 IEEE International Symposium on Electromagnetic Compatibility Signal/Power Integrity (EMCSI)*, Aug 2017, pp. 112–117.
- [8] G. Shafer, *A Mathematical Theory of Evidence*. Princeton, NJ, USA: Princeton University Press, 1976, vol. 42.
- [9] C. Baudrit, D. Dubois, and D. Guyonnet, "Joint propagation and exploitation of probabilistic and possibilistic information in risk assessment," *IEEE Transactions on Fuzzy Systems*, vol. 14, no. 5, pp. 593–608, Oct. 2006.
- [10] P. Baraldi, I. C. Popescu, and E. Zio, "Predicting the time to failure of a randomly degrading component by a hybrid monte carlo and possibilistic method," in *2008 International Conference on Prognostics and Health Management*, Oct 2008, pp. 1–8.
- [11] A. Ferrero and S. Salicone, "The random-fuzzy variables: A new approach for the expression of uncertainty in measurement," in *Proceedings of the 20th IEEE Instrumentation Technology Conference*, vol. 2, May 2003, pp. 1502–1507.
- [12] L. Badini, N. Toscani, G. Spadacini, F. Grassi, and S. A. Pignari, "A possibilistic approach to radiated susceptibility of twisted-wire pairs," in *2017 IEEE International Symposium on Electromagnetic Compatibility & Signal/Power Integrity (EMCSI)*, Aug 2017, pp. 96–101.
- [13] P. Chen, B. M. Merrick, and T. J. Brazil, "Bayesian optimization for broadband high-efficiency power amplifier designs," *IEEE Transactions on Microwave Theory and Techniques*, vol. 63, no. 12, pp. 4263–4272, Dec 2015.
- [14] N. Knudde, I. Couckuyt, D. Spina, K. Łukasik, P. Barmuta, D. Schreurs, and T. Dhaene, "Data-efficient bayesian optimization with constraints for power amplifier design," in *2018 IEEE MTT-S International Conference on Numerical Electromagnetic and Multiphysics Modeling and Optimization (NEMO)*, Aug 2018, pp. 1–3.
- [15] H. M. Torun and M. Swaminathan, "High-dimensional global optimization method for high-frequency electronic design," *IEEE Transactions on Microwave Theory and Techniques*, vol. 67, no. 6, pp. 2128–2142, June 2019.
- [16] D. Dubois and H. Prade, "When upper probabilities are possibility measures," *Fuzzy Sets and Systems*, vol. 49, no. 1, pp. 65–74, Jul 1992.

- [17] L. A. Zadeh, "Fuzzy sets as a basis for a theory of possibility," *Fuzzy Sets and Systems*, vol. 1, pp. 3–28, 1978.
- [18] C. Baudrit, D. Guyonnet, and D. Dubois, "Post-processing the hybrid method for addressing uncertainty in risk assessment," *Journal of Environmental Engineering*, vol. 131, no. 2, pp. 1750–1754, Dec 2005.
- [19] B. Shahriari, K. Swersky, Z. Wang, R. P. Adams, and N. de Freitas, "Taking the human out of the loop: A review of bayesian optimization," *Proceedings of the IEEE*, vol. 104, no. 1, pp. 148–175, Jan 2016.
- [20] J. T. Springenberg, A. Klein, S. Falkner, and F. Hutter, "Bayesian optimization with robust Bayesian neural networks," in *Advances in Neural Information Processing Systems*, D. D. Lee, M. Sugiyama, U. V. Luxburg, I. Guyon, and R. Garnett, Eds., vol. 29. Curran Associates, Inc., 2016, pp. 4134–4142.
- [21] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*. MIT Press, 2006. [Online]. Available: <http://www.gaussianprocess.org/gpml>
- [22] D. J. C. MacKay, *Information Theory, Inference, and Learning Algorithms*. Cambridge University Press, 2003.
- [23] J. Moćkus, "On Bayesian methods for seeking the extremum," in *Optimization Techniques IFIP Technical Conference: Novosibirsk, July 1–7, 1974*, G. I. Marchuk, Ed. Berlin, Heidelberg: Springer Berlin Heidelberg, 1975, pp. 400–404.
- [24] D. R. Jones, M. Schonlau, and W. J. Welch, "Efficient global optimization of expensive black-box functions," *Journal of Global optimization*, vol. 13, no. 4, pp. 455–492, Dec 1998.
- [25] H. J. Kushner, "A new method of locating the maximum point of an arbitrary multipeak curve in the presence of noise," *Journal of Basic Engineering*, vol. 86, no. 1, pp. 97–106, 1964.
- [26] J. Snoek, H. L. Hugo, and R. P. Adams, "Practical Bayesian optimization of machine learning algorithms," in *Advances in Neural Information Processing Systems*, F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, Eds., vol. 25. Curran Associates, Inc., Jun 2012, pp. 2951–2959.
- [27] The GPyOpt authors, "GPyOpt: A Bayesian optimization framework in python," <http://github.com/SheffieldML/GPyOpt>, 2016.
- [28] N. Knudde, J. van der Herten, T. Dhaene, and I. Couckuyt, "GPFlowOpt: A Bayesian optimization library using TensorFlow," in *Neural Information Processing Systems 2017-Workshop on Bayesian Optimization*, 2017, pp. 1–5.
- [29] D. Xiu and G. E. Karniadakis, "The Wiener-Askey polynomial chaos for stochastic differential equations," *SIAM journal on scientific computing*, vol. 24, no. 2, pp. 619–644, Apr 2002.
- [30] D. Spina, T. Dhaene, L. Knockaert, and G. Antonini, "Polynomial chaos-based macromodeling of general linear multiport systems for time-domain analysis," *IEEE Transactions on Microwave Theory and Techniques*, vol. 65, no. 5, pp. 1422–1433, May 2017.
- [31] 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 Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 32, no. 10, pp. 1533–1545, Oct 2013.
- [32] M. R. Ruftuie, E. Gad, M. Nakhla, and R. Achar, "Generalized Hermite polynomial chaos for variability analysis of macromodels embedded in nonlinear circuits," *IEEE Transactions on Components, Packaging and Manufacturing Technology*, vol. 4, no. 4, pp. 673–684, April 2014.
- [33] C. Cui and Z. Zhang, "Stochastic collocation with non-Gaussian correlated process variations: Theory, algorithms and applications," *IEEE Transactions on Components, Packaging and Manufacturing Technology*, pp. 1–1, 2018.
- [34] L. Badini, F. Grassi, S. A. Pignari, G. Spadacini, P. Bisognin, P. Pelissou, and S. Marra, "Conducted-susceptibility testing as an alternative approach to unit-level radiated-susceptibility verifications," in *2016 ESA Workshop on Aerospace EMC (Aerospace EMC)*, May 2016, pp. 1–5.
- [35] G. Spadacini and S. A. Pignari, "Numerical assessment of radiated susceptibility of twisted-wire pairs with random nonuniform twisting," *IEEE Transactions on Electromagnetic Compatibility*, vol. 55, no. 5, pp. 956–964, Oct 2013.
- [36] S. A. Pignari and G. Spadacini, "Plane-wave coupling to a twisted-wire pair above ground," *IEEE Transactions on Electromagnetic Compatibility*, vol. 53, no. 2, pp. 508–523, May 2011.
- [37] C. Gazda, D. Vande Ginste, H. Rogier, R.-B. Wu, and D. De Zutter, "A wideband common-mode suppression filter for bend discontinuities in differential signaling using tightly coupled microstrips," *IEEE Transactions on Advanced Packaging*, vol. 33, no. 4, pp. 969–978, Nov 2010.
- [38] C. Gazda, I. Couckuyt, H. Rogier, D. Vande Ginste, and T. Dhaene, "Constrained multiobjective optimization of a common-mode suppression filter," *IEEE Transactions on Electromagnetic Compatibility*, vol. 54, no. 3, pp. 704–707, June 2012.
- [39] Keysight EEs of EDA, "Advanced design system," 06 2018, 471.update1.0. [Online]. Available: <http://www.keysight.com/find/eesof-ads>