A Robust Bayesian Optimization Framework for Microwave Circuit Design under Uncertainty

Duygu De Witte * [ID]**, Jixiang Qing** [ID]**, Ivo Couckuyt** [ID]**, Tom Dhaene** [ID]**, Dries Vande Ginste** [ID]
and Domenico Spina [ID]

IDLab, Department of Information Technology, Ghent University-imec, Technologiepark-Zwijnaarde 126,
9052 Ghent, Belgium; jixiang.qing@ugent.be (J.Q.); ivo.couckuyt@ugent.be (I.C.); tom.dhaene@ugent.be (T.D.);
dries.vandeginste@ugent.be (D.V.G.); domenico.spina@ugent.be (D.S.)
*** Correspondence: duygu.kan@ugent.be

Abstract: In modern electronics, there are many inevitable uncertainties and variations of design parameters that have a profound effect on the performance of a device. These are, among others, induced by manufacturing tolerances, assembling inaccuracies, material diversities, machining errors, etc. This prompts wide interests in enhanced optimization algorithms that take the effect of these uncertainty sources into account and that are able to find robust designs, i.e., designs that are insensitive to the uncertainties early in the design cycle. In this work, a novel machine learning-based optimization framework that accounts for uncertainty of the design parameters is presented. This is achieved by using a modified version of the expected improvement criterion. Moreover, a data-efficient Bayesian Optimization framework is leveraged to limit the number of simulations required to find a robust design solution. Two suitable application examples validate that the robustness is significantly improved compared to standard design methods.

Keywords: Bayesian Optimization; robust optimization; Gaussian processes; microwave design

1. Introduction

The emergence of new technologies and the constant miniaturization of integrated circuits challenge engineers to obtain designs that satisfy stringent functional specifications and signal integrity (SI) requirements, as well as electromagnetic compatibility (EMC) constraints. Various optimization techniques are used to find optimal designs and improve the device performance at the early design stage [1–3]. While in other engineering domains it is sometimes possible to solve complex numerical equations with efficient-stable numerical methods [4–6]; given the high computational cost of simulating modern high-frequency circuits, surrogate modeling techniques are typically utilized to efficiently perform design optimization [7–11]. In particular, since the complexity of design optimization problems is constantly increasing, machine learning-based algorithms have become a popular choice to cope with the multiscale issues in radio frequency (RF) and microwave designs [12–15].

While, theoretically speaking, the performance of an electromagnetic device can be improved by applying adequate optimization techniques, in reality, there are many unavoidable geometrical and material uncertainties that degrade the device's performance. In particular, the real-life performance of a device is significantly affected by the manufacturing technology employed, the assembling inaccuracy, uncertainties due to material diversities, operation environment, etc. From the perspective of industrial design and manufacturing, there are two types of optimization models: deterministic and robust optimization models [16]. In deterministic optimization methods, all design parameters and system variables assume known values, which are freely specified by designers. Since no random variability and data uncertainty are investigated, deterministic optimization models fall short to find robust designs. In contrast, the latter are considered (to a certain extent) insensitive to variations of design parameters.

As a relatively simple case of an electromagnetic system, consider, for example, a pair of coupled microstrip lines for high-speed digital circuits. There are several manufacturing uncertainties for these types of structures. These include an asymmetric ground pin configuration, signal trace length mismatch, and asymmetric discontinuities in the circuit layout, such as at a bend of coupled microstrip lines [17]. These manufacturing variations lead to severe signal degradation, thus decreasing performance in terms of its SI. To increase the reliability of electromagnetic circuits, a robust design optimization is thus a necessity.

Several formulations for robust design have been proposed in the literature [16,18–23]. The main idea of these approaches is to capture and account for the process variations and manufacturing tolerances early in the design cycle. In [24], Kriging surrogate models assist a worst-case robustness scenario for the optimization of electromagnetic devices. Simulation-based robust design strategies were developed by employing the Monte Carlo (MC) [25] sampling method to investigate the impact of parameter variations on product performance [26,27]. Most of these approaches, however, rely on methods that are not data-efficient. This motivates us to introduce a data-efficient optimization methodology while accounting for the uncertainties described earlier in this section.

The term data-efficient is used here to indicate techniques that minimize the amount of data necessary to reach the desired objective. Data-efficiency is particularly relevant for modern microwave design problems since (full-wave) simulations are computationally very expensive. Therefore, in this paper, we present a new data-efficient methodology for robust optimization of microwave circuits and systems based on Bayesian inference, which allows designers to take manufacturing imperfections during the optimization process into account, while minimizing the related computational cost. More precisely, our methodology adopts Gaussian Processes (GPs) [28] and propagates the input uncertainty on the design parameters by moment matching the predictive distribution of the underlying GP. Furthermore, we propose a modified version of the Expected Improvement (EI) criterion that accounts for the input uncertainty, called stochastic EI (sEI), which is used to guide the optimization process [29]. To the best of the authors' knowledge, in this paper, sEI is used for the first time for the data-efficient, robust optimization of microwave circuits and systems.

The paper is organized as follows. Section 2 first discusses the concept of robust optimum by means of an illustrative example. Next, it continues with a brief discussion on robust design optimization in engineering and introduces some of the most recent developments. Since our methodology is based on Bayesian inference, the relevant features of BO, GPs and EI are introduced in Sections 2.1–2.3, respectively. In Section 3, the complete novel methodology for solving robust BO problems in electrical engineering is proposed. Two representative numerical examples and their results are presented in Section 4. Finally, conclusions are drawn and the future research avenues are described in Section 5.

2. Problem Formulation

In electronics, engineers are interested in a design solution that satisfies specifications and, at the same time, is stable under uncertainty. Such a robust solution may represent a local rather than a global optimum, as long as it is significantly less sensitive to small production perturbations compared to the global optimum. A schematic example of a minimization problem is shown in Figure 1. In this figure, the horizontal axis shows a design parameter x, while $f(x)$ indicates the function that we want to minimize (often referred to as the objective function in an optimization framework). For microwave circuits, $f(x)$ typically corresponds to a figure of merit describing the performance of the circuit, such as bandwidth, gain, attenuation, etc. Deterministic methods tend to find the global optimum of a function $f(x)$, indicated as point A in Figure 1, without taking any parameter variations into account [16]. This may result in big fluctuations (Δf_1) of the objective with respect to small variations Δx of a design parameter (the length of a microstrip, for example). Looking at Figure 1, the solution B is considered as robust, as the objective

function $f(x)$ does not change much with respect to the same perturbation Δx. Therefore, despite being a local optimum, in practice design B is more favorable for engineers.
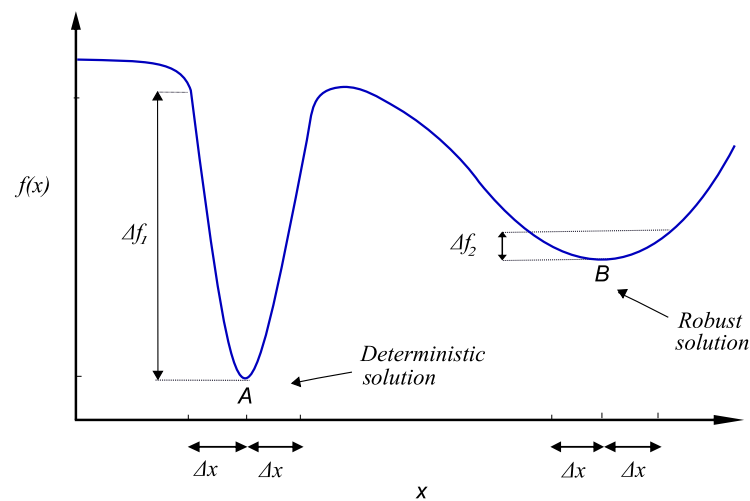

Figure 1. Illustration of global versus robust solutions.

Recently, several procedures have been presented to find such robust solutions in different engineering domains. While in [30] a sensitivity analysis was exploited to build a framework for robust design optimization of Mechatronics Systems, a global two-layer meta-model approximation was presented in [31] to cope with the computational challenges of robust design optimization. Moreover, in [32] a multidisciplinary robust design optimization framework, which takes both parameter and metamodeling uncertainties into account, was introduced.

In general, evaluating a relatively simple robustness measure (e.g., expectancy or worst value) of the cost function is a common practice in literature. However, worst-case analysis methods typically result in an overly pessimistic estimation of the tolerance effects [33,34]. Nowadays, Uncertainty Quantification (UQ) of electronic circuits is a popular alternative where surrogate models are adopted for efficient statistical analysis [35–38]. UQ methods can assist engineers to achieve robust designs. In [39], an UQ-based optimization strategy, leveraging on knowledge-based feature surrogates, is introduced for the robust design of microwave components. For a broad overview of several robust optimization approaches, the reader is referred to [2,16,19,40].

However, the current state-of-the-art techniques are either not probabilistic or not data-efficient. To overcome these shortcomings, we present a novel data-efficient approach that relies on a robustness measure and stochastic EI, detailed in Section 3. Still, first, the preliminary concepts BO, GP, and EI are succinctly introduced in Sections 2.1–2.3, respectively.

2.1. Bayesian Optimization

Consider an objective function $f \colon X \to \mathbb{R}$, also called cost function in the BO framework, defined on a compact subset $X \subseteq \mathbb{R}^D$. BO is a model-based approach to solve a global minimization (or maximization) problem defined as follows:

$$\min_{\mathbf{x} \in X \subset \mathbb{R}^D} f(\mathbf{x}), \tag{1}$$

where $\mathbf{x}$ is the vector of length D containing the D design parameters. In the context of an electromagnetics problem, one might want to minimize the return loss of a filter at its center frequency or maximize the gain of an antenna over a certain frequency range. Here, the objective function f is often expensive to evaluate, as it requires full-wave modelling implemented in simulation tools such as Advanced Design System (ADS) [41].

BO is particularly effective to solve expensive design optimization problems, as it aims at minimizing the number of expensive function evaluations.

BO relies on two main elements: a surrogate model that mimics the objective function, and a sequential sampling strategy that selects the next sample intelligently in order to reach the global optimum as fast as possible. Let us consider the general BO methodology depicted in Figure 2. First, the objective function $f(\mathbf{x})$ is evaluated for a limited set of design parameters $[\mathbf{x}_k]_{k=1}^{K} \in X$. Initial samples are typically generated according to suitable space filling techniques, such as a Latin Hypercube Design [42]. Based on this initial data, a surrogate model of $f(\mathbf{x})$ is computed. This model is very cheap to evaluate compared to performing (full-wave) simulations, and it is used to calculate the location of the candidate optimum. It is important to remark that surrogate models used in BO are stochastic: the model not only predicts the value of the objective function with regards to the parameters $\mathbf{x}$, but also the degree of confidence in its prediction.
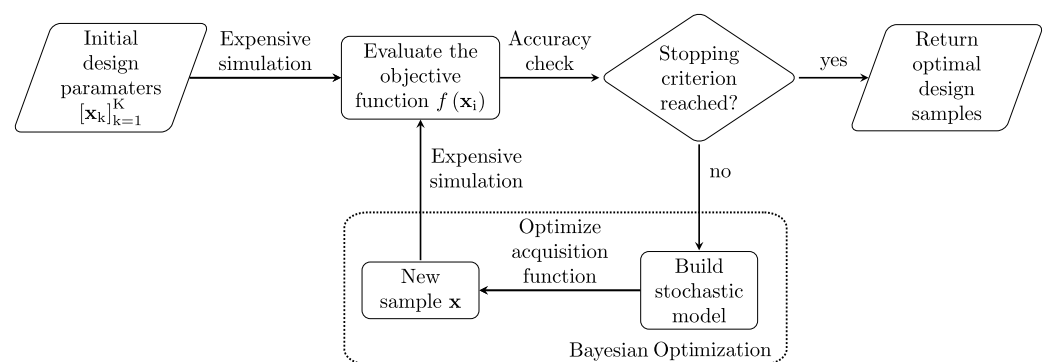

Figure 2. Flowchart of the BO framework.

The sampling strategy in BO relies on a so-called acquisition function. More specifically, the acquisition function is used to determine the location of the new sample to be evaluated, based on the stochastic model's predictions and its confidence bounds. The sample selected is then evaluated by a new (expensive) simulation of the objective function and the surrogate model is updated. This procedure continues until suitable stopping criteria are met, and each simulation refines the surrogate, increasing the probability of finding the global optima of the problem (1). A complete description of the BO properties is given in [43,44].

2.2. Gaussian Processes

In this work, we adopt a BO framework that leverages on GP as a surrogate model. Different surrogate models can be used in BO, such as Bayesian neural networks [45] and GPs [28,46]. The latter choice (GPs) is common in a BO context and also used in this paper. This is because it is analytically tractable and provides a predictive distribution given new input data. More specifically, a GP $f \sim \mathcal{GP}(m, k)$ represents a distribution over functions $f: X \to \mathbb{R}$, which is completely characterized by its mean function $m: X \to \mathbb{R}$ and a positive-definite kernel, or covariance function, $k: X \times X \to \mathbb{R}$. As such, a finite set of function values $[y_1, y_2, \ldots, y_N] = [f(x_1), f(x_2), \ldots, f(x_N)]$ are distributed according to a multivariate Gaussian distribution with mean $\mathbf{m}$ and covariance matrix $K_{\mathbf{xx}}$, where $\mathbf{m}_i = m(\mathbf{x}_i)$ and $(K_{\mathbf{xx}})_{ij} = k(\mathbf{x}_i, \mathbf{x}_j)$. The mean is typically chosen to be zero. Among the different covariance functions presented in literature, the popular squared exponential (SE) kernel is used in this work, which has the form:

$$k_{SE}(\mathbf{x}, \mathbf{x}') = \sigma^2 \exp\left(-\sum_{d=1}^{D} \frac{(x_d - x'_d)^2}{2\, l_d^2} \right),$$

where the hyperparameters θ are the collection of the kernel variance σ and the length-scales l_d, $d = 1, \ldots, D$. The hyperparameters θ are tuned for a given training data set using the Maximum Likelihood Estimation (MLE):

$$\hat{\theta} = \arg\max_{\theta} \log p(\mathbf{f}|\theta)$$
$$= \arg\max_{\theta} -\frac{1}{2}\left(\log|2\pi K_{\mathbf{xx}}| + \mathbf{f}^T K_{\mathbf{xx}}^{-1}\mathbf{f}\right).$$

Let $\mathcal{D}_n = \{\mathbf{x}_n, y_n\}$, $n = 1, \ldots, N$, denote the set of observations of the design under study. The predictive distribution of GPs for a new input $\mathbf{x}_\star$ based on $\mathcal{D}_n$ (also called posterior distribution) is denoted $p(f_\star|\mathbf{x}_\star, \mathcal{D}_n)$ and it can be analytically calculated, resulting in a Gaussian distribution with the following moments [28]:

$$\mu(\mathbf{x}_\star) = \mathbb{E}(f_\star|\mathbf{x}_\star, \mathcal{D}_n) = \mathbf{k}_{\star\mathbf{x}}^T K_{\mathbf{xx}}^{-1}\mathbf{y}_n$$
$$\sigma^2(\mathbf{x}_\star) = \mathrm{Var}(f_\star|\mathbf{x}_\star, \mathcal{D}_n) = k_{\star\star} - \mathbf{k}_{\mathbf{x}\star}^T K_{\mathbf{xx}}^{-1}\mathbf{k}_{\mathbf{x}\star}$$

where $(K_{\mathbf{xx}})_{ij} = k(\mathbf{x}_i, \mathbf{x}_j)$, $(\mathbf{k}_{\mathbf{x}\star})_i = k(\mathbf{x}_\star, \mathbf{x}_i)$, $k_{\star\star} = k(\mathbf{x}_\star, \mathbf{x}_\star)$. In our optimization problem, the posterior mean $\mu(\mathbf{x}_\star)$ represents the GP prediction, while its variance $\sigma^2(\mathbf{x}_\star)$ indicates the model's confidence in its predictions.

2.3. Expected Improvement

Among the most commonly used acquisition functions are the Expected Improvement (EI) [47,48] and Probability of Improvement (PoI) [49]. In particular, a stochastic version of EI is adopted in this manuscript (see Section 3). Traditional EI is defined as:

$$\alpha_{\mathrm{EI},n}(\mathbf{x}_\star) = \int_{-\infty}^{y_{min}} (y_{min} - f_\star)\, p(f_\star|\mathbf{x}_\star, \mathcal{D}_n)\, \mathrm{d}f_\star. \tag{2}$$

Here, y_{min} is the minimum value observed thus far, $y_{min} = \min_{\mathbf{x} \in \mathcal{D}_n} \mathbb{E}(f|\mathbf{x}, \mathcal{D}_n)$ and $|y_{min} - f_\star|$ represents the improvement. Hence, under the GP model, EI can be rewritten in closed form as:

$$\alpha_{\mathrm{EI},n}(\mathbf{x}_\star) = (y_{min} - \mu(\mathbf{x}_\star))\Phi\left(\frac{y_{min} - \mu(\mathbf{x}_\star)}{\sigma(\mathbf{x}_\star)}\right)$$
$$+ \sigma(\mathbf{x}_\star)\phi\left(\frac{y_{min} - \mu(\mathbf{x}_\star)}{\sigma(\mathbf{x}_\star)}\right),$$

when $\sigma > 0$ and vanishes otherwise. Note that $\phi(\cdot)$ and $\Phi(\cdot)$ denote the probability density function and cumulative distribution function of the standard normal distribution, respectively.

The goal is to find the point that maximizes the EI and add it to the data set $\mathcal{D}_n$. Using this data, the predictive distribution $p(f_\star|\mathbf{x}_\star, \mathcal{D}_n)$ is updated and EI is recalculated to determine the next point to evaluate. This corresponds to one iteration of the BO loop and continues until the global optimum is found.

Note that finding the value of the parameters $\mathbf{x}$ that maximizes the acquisition function is an optimization problem per se. However, solving this problem is a relatively easy task: the EI is fast to compute since it is calculated using the current posterior distribution. It is also differentiable and can therefore be maximized with a standard gradient-based optimizer [50].

3. Robust Bayesian Optimization in Engineering Design

In circuits, design parameters are usually affected by uncertainties which degrade the overall performance. Among different types of uncertainty, manufacturing tolerances are one of the primary uncertainty sources, and therefore the focus of this work.

The optimization strategy described in Section 2 is deterministic. Thus, it does not take the uncertainty on the design parameters into account when searching for the optimal design configuration. Hence, the goal of robust design optimization is finding the minimum of $f(\mathbf{x} + \delta)$, where $\mathbf{x}$ is the set of design parameters and δ represents the perturbation on the design parameters. Note that δ is a vector of stochastic (random) variables and not deterministic ones. In this paper, the elements of δ are assumed to be normally distributed, that is $\delta \sim \mathcal{N}(0, \Sigma)$, where the correlation matrix Σ is chosen as a diagonal matrix. Thus, the elements of δ are considered uncorrelated (and since they are Gaussian random variables, they are also independent). Additionally, we assume that the elements of Σ are constant, so their values are independent on $\mathbf{x}$.

Finding a robust optimum instead of a global optimum is a challenging mathematical problem. Moreover, there is no unified mathematical definition of robustness [16,18]. One main idea portrayed in literature is to evaluate an expectation measure of the objective function. This was first introduced in literature as Type 1 robustness by Deb and Gupta [51].

In this manuscript, inspired by [51,52], the used expectation measure E_r is:

$$E_r(\mathbf{x}) = \int_{\mathbb{R}^D} f(\mathbf{z}) \mathcal{N}(\mathbf{z}|\mathbf{x}, \Sigma) \, d\mathbf{z}, \tag{3}$$

where $p(\mathbf{z}) = \mathcal{N}(\mathbf{z}|\mathbf{x}, \Sigma)$ is the probability density function of a multivariate normal distribution with mean $\mathbf{x}$ and covariance Σ. This expectation measure E_r is, in essence, the expectation value of the cost function under perturbation [29]. Figure 3 illustrates the optimal solution resulting from minimizing an objective function (point B) versus its corresponding expectation measure E_r (point A). Solution A is considered robust, as for a small amount of perturbation δ of the design parameters, the objective function value of the solution does not change significantly. Solution B offers a better minimum, but it is sensitive to perturbations: it may not be suited for the application at hand.
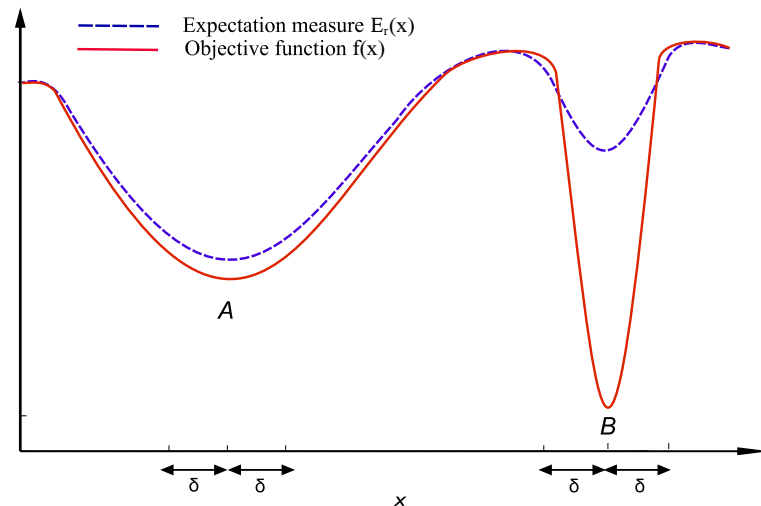

Figure 3. Illustration of the expectation measure E_r in Equation (3) versus the traditional cost function.

The main idea of the proposed robust optimization strategy is to adopt the BO algorithm described in Section 2.1 and summarized in Figure 2, where the uncertainty on the design parameters is propagated through the GP model and a new acquisition function is defined to estimate a robust optimum following the measure in Equation (3). Hence, the EI presented in Section 2.3 is replaced by a *stochastic* EI as follows: first, the expectation of EI in Equation (2) is taken with regards to $p(\mathbf{z})$, which is the equivalent of the regular EI, but using the metric defined by Equation (3). Next, using Fubini's theorem [53], the following acquisition function is obtained:

$$\alpha_{\text{sEI},n}(\mathbf{x}_\star) = \mathbb{E}_{p(\mathbf{z}|\mathbf{x}_\star,\Sigma)}[\alpha_{\text{EI},n}(\mathbf{z})]$$

$$= \int_{-\infty}^{y_{min}} (y_{min} - f_\star) \left[\int_{\mathbb{R}^D} p(f_\star|\mathbf{z}, \mathcal{D}_n) p(\mathbf{z}|\mathbf{x}, \Sigma) \, \mathrm{d}\mathbf{z} \right] \mathrm{d}f_\star. \tag{4}$$

The last integral corresponds to the marginalization of $\mathbf{z}$ and has the predictive distribution $p(f_\star|\mathbf{x}_\star, \mathcal{D}_n)$ as a result.

The chosen acquisition function $\alpha_{\text{sEI},n}$ requires the marginalization of the input space of the GP. However, this is analytically intractable. Relying on the literature, the most accurate way to calculated the predictive distribution is to apply Markov Chain Monte Carlo (MCMC) methods. Since the moments of the predictive distribution are analytically tractable if the kernel expectations are tractable [54], a moment matching method is used to approximate the predictive distribution $p(f_\star|\mathbf{x}_\star, \mathcal{D}_n)$. The kernel expectations are noted as follows:

$$\xi_\star = \mathbb{E}(k_{\star\star})$$
$$(\psi_{\star\mathbf{x}})_i = \mathbb{E}((\mathbf{k}_{\star\mathbf{x}})_i) \tag{5}$$
$$(\Phi_{\star\mathbf{x}})_{ij} = \mathbb{E}((\mathbf{k}_{\star\mathbf{x}})_i(\mathbf{k}_{\star\mathbf{x}})_j)$$

In Section 2.2, it is described that the SE is used as a kernel of the GP model: the expectations in Equation (5) can be calculated *analytically* for the SE kernel. To obtain a sufficiently good approximation of the real distribution, in practice, we suggest to adopt only the mean and variance and calculate them through the law of total cumulance [55]:

$$\mathbb{E}(f_\star) = \mathbb{E}_{p(\mathbf{x}_\star)}(\mathbb{E}(f_\star|\mathbf{x}_\star, \mathcal{D}_n)) = \psi_{\star\mathbf{x}}^T K_{\mathbf{x}\mathbf{x}}^{-1} \mathbf{y}_n$$
$$\text{Var}(f_\star) = \text{Var}_{p(\mathbf{x}_\star)}(\mathbb{E}(f_\star|\mathbf{x}_\star, \mathcal{D}_n)) + \mathbb{E}_{p(\mathbf{x}_\star)}(\text{Var}(f_\star|\mathbf{x}_\star, \mathcal{D}_n))$$
$$= \xi_\star + \text{Tr}(K_{\mathbf{x}\mathbf{x}}^{-1}(\mathbf{y}_n\mathbf{y}_n^T - K_{\mathbf{x}\mathbf{x}})K_{\mathbf{x}\mathbf{x}}^{-1}\Phi_{\star\mathbf{x}}) - \mathbb{E}(f_\star)^2.$$

These stochastic moments are matched with a Gaussian distribution to approximate $p(f_\star|\mathbf{x}_\star, \mathcal{D}_n)$. In this case, the expression of the sEI is similar to the deterministic case:

$$\alpha_{\text{sEI},n} = (y_{min} - \mathbb{E}(f_\star))\Phi\left(\frac{y_{min} - \mathbb{E}(f_\star)}{\sqrt{\text{Var}(f_\star)}}\right)$$
$$+ \sqrt{\text{Var}(f_\star)}\phi\left(\frac{y_{min} - \mathbb{E}(f_\star)}{\sqrt{\text{Var}(f_\star)}}\right).$$

Finally, the global minimum y_{min} (which is unknown) is replaced by the lowest expectation under uncertainty: $y_{min} \approx \min_{\mathbf{x} \in X_n} \mathbb{E}_{p(\mathbf{z}|\mathbf{x},\Sigma)}(\mathbb{E}(f_\star|\mathbf{z}, \mathcal{D}_n))$. The derived formula for sEI is used in the BO framework defined in Figure 2 to find the robust optimum.

A 1D example that illustrates the sEI versus deterministic EI is shown in Figure 4. In this example, the sEI is much smoother than the regular EI because of the averaging over the input distribution. The robust optimum is a local minimum. Clearly, the EI samples at the global optimum (i.e., it has higher values around the global optimum), while the sEI samples near the robust optimum.

In the traditional BO setting, the current best optimum is easily obtained by comparing the function values evaluated so far. However, when using sEI, it is impossible to *know* what the current best optimum is. Indeed, to calculate the current best optimum, the GP model is used to approximate the expectation measure E_r. Consequently, unlike in the regular BO setting, the convergence graph of the current best optimum may not exhibit a monotonously decreasing trend.
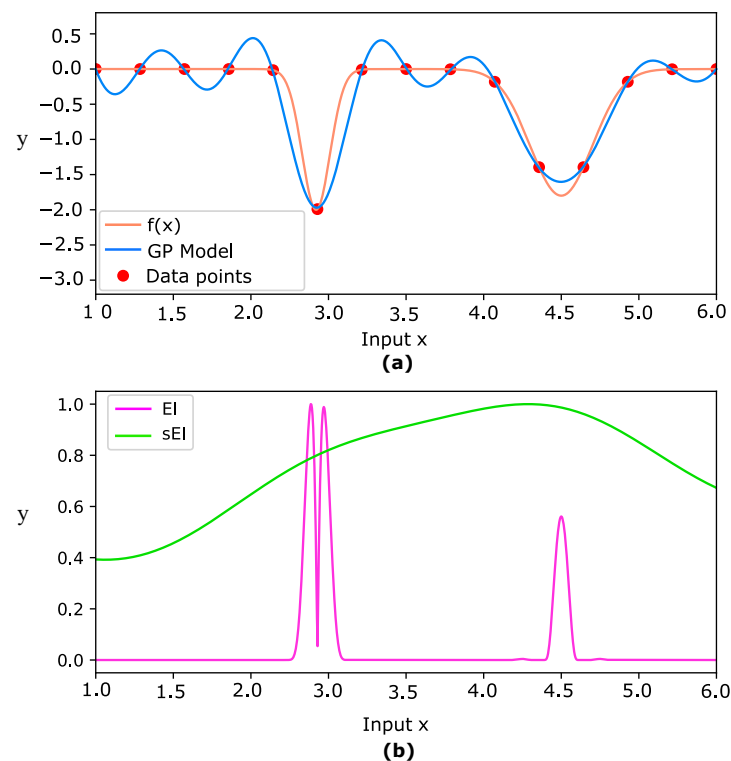
Figure 4. (**a**) The GP model and (**b**) comparison of acquisition functions EI and sEI for an objective function $f(x)$.

4. Application Examples

In this section, the proposed method is applied to optimize two microwave filter designs, and their performances are compared to standard BO. The proposed method was implemented using the GPFlowOpt Python package [56].

4.1. Microstrip Lowpass Filter

The two-port lowpass stepped impedance microstrip filter presented in [57] is studied in this section. The filter is formed by six microstrip line sections with different lengths l_i, $i = 1, \ldots, 6$, while the following relation holds for their widths [57]: $w_1 = w_3 = w_5$ and $w_2 = w_4 = w_6$. The corresponding filter layout is presented in Figure 5. This filter operates in the $[1, 4]$ GHz band, and resides on a substrate with a relative permittivity $\epsilon_r = 4.2$ and thickness $h = 1.58$ mm. The nominal values of its geometrical parameters are shown in Table 1. The simulator used to estimate the filter's performance is the MATLAB RF Toolbox (Mathworks Inc., Natick, MA, USA).

Table 1. Microstrip low-pass filter geometrical parameters.

Parameter	Value
Microstrip lengths	$l_1 = 2.05$ mm, $l_2 = 6.63$ mm, $l_3 = 7.69$ mm, $l_4 = 9.04$ mm, $l_5 = 5.63$ mm, $l_6 = 2.41$ mm
Microstrip widths	$w_1 = w_3 = w_5 = 11.3$ mm $w_2 = w_4 = w_6 = 0.428$ mm

For this example, we will consider the length of each microstrip section as a relevant design parameter, defined in the supports $l_1, l_6 \in [1, 5]$ mm, $l_2, l_5 \in [4, 8]$ mm and, $l_2, l_3 \in [6.5, 10.5]$ mm. The design goal is a filter with a 3 dB cut-off frequency at 2.4 GHz. This is achieved by minimizing the objective function formulated as: $f_c(\mathbf{x}) = |f_c - 2.4|$, where $f_c(\mathbf{x})$ is the cut-off frequency expressed in GHz for a specific design configuration.
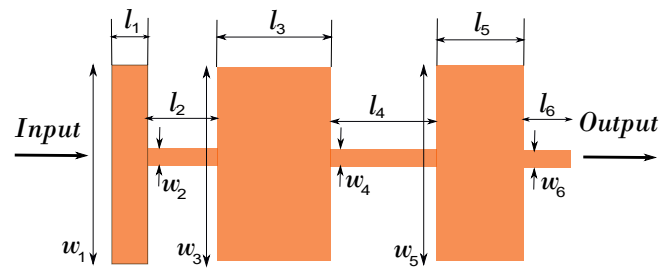
Figure 5. Top view of the layout of the lowpass filter under study.

In this example, we assume that the perturbation δ on the six design parameters has the following covariance matrix $\Sigma = \mathrm{diag}(0.6^2, 1.2^2, 1.7^2, 1.7^2, 1.2^2, 0.6^2)$ mm^2. Comparing the elements of the covariance matrix with the nominal values defined in Table 1, it is clear that a relatively large perturbation of the design parameters is assumed. This choice is made to clearly illustrate the performance of the proposed method compared to standard BO. The optimization is performed by assuming a total computational budget of 100 (l_1, l_2, l_3, l_4, l_5, l_6) samples for both standard BO (with EI) and our novel robust BO methodology (with sEI). Ten initial samples are chosen via a Latin Hypercube Design, while the remaining samples are sequentially chosen by the optimization algorithm, comparing EI with sEI.

In order to assess the robustness to the value of the initial samples, 12 sets of initial samples are chosen via a Latin Hypercube Design and the optimization process is repeated for each sample set. This allows one to estimate a confidence interval for both standard BO and the advocated methodology.

Results are shown in Figure 6, which illustrates the progression of the robust measure E_r with respect to the number of simulations performed for both BO and the proposed robust algorithm. The proposed optimization method clearly outperforms BO to find a robust optimum, and the difference already becomes clear after a few iterations.
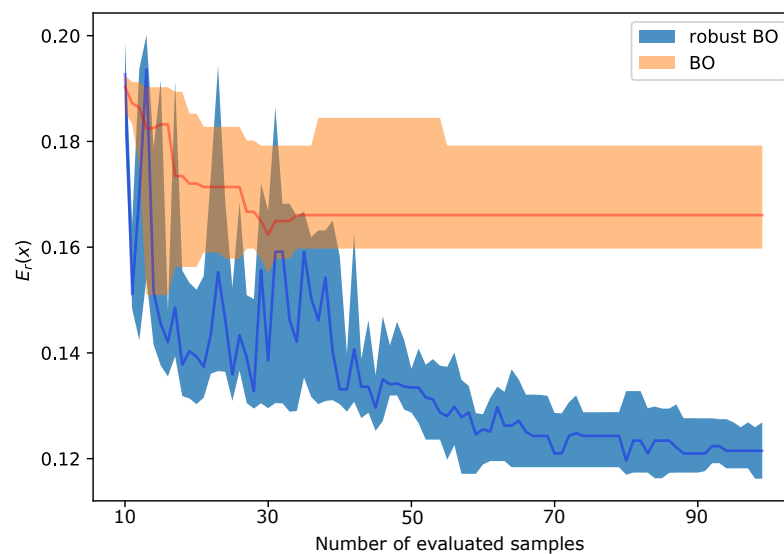

Figure 6. The expectation measure E_r as a function of samples. We show the median (lines) and the 25/75th percentiles (shaded area) of the E_r measure calculated for 12 different sets of initial samples.

To evaluate how different E_r values describe the robustness of the filter performance with respect to the perturbation δ on the design parameters, additional results concerning the filter's transmission characteristic (element S_{21} of the filter's scattering matrix) are

shown in Figure 7. First, consider the optimization results for one set of the initial samples: standard BO individuates a global optimum $\mathbf{x}_{BO} = (5.00, 5.25, 9.89, 8.38, 8.00, 1.33)$ corresponding to $f(\mathbf{x}_{BO}) = 3.4e - 16$, while the proposed robust optimization strategy selects a robust local optimum at $\mathbf{x}_{RBO} = (1.00, 7.81, 10.50, 6.50, 8.00, 5.00)$, corresponding to $f(\mathbf{x}_{RBO}) = 0.025$. Figure 7 shows the results of the MC analysis performed around $\mathbf{x}_{BO}$ and $\mathbf{x}_{RBO}$: the design solution individuated by the novel algorithm is significantly more robust to the perturbation δ on the design parameters. Similar results hold for all 12 initial sample sets.
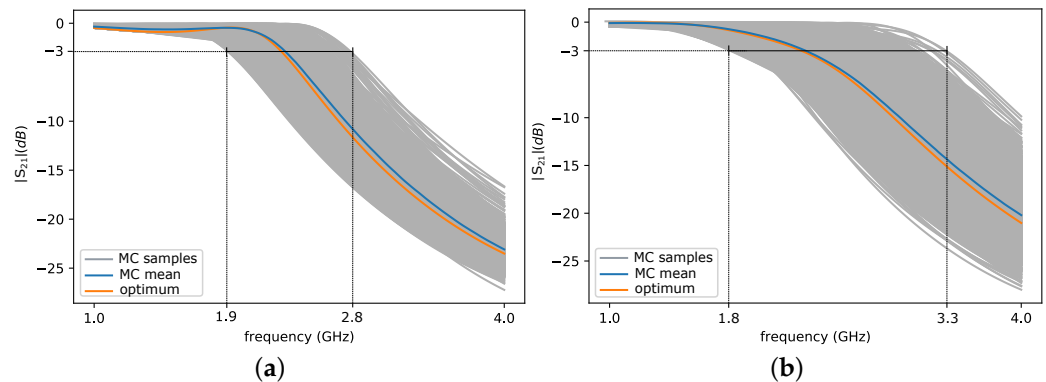

(a)   **(b)**

Figure 7. (**a**) Transmission characteristic $|S_{21}|$ of the lowpass filter using the robust methodology and (**b**) standard BO. (**a**) The optimal 3 dB cut-off frequency found by the robust BO is 2.38 GHz, and MC analysis reveals that it varies in the [1.9, 2.8] GHz range; (**b**) The optimal 3 dB cut-off frequency found by the standard BO is 2.40 GHz, and MC analysis reveals that it varies in the [1.8, 3.3] GHz range.

4.2. Zigzag Narrow Bandpass Filter

The second microwave example is a zigzag narrow bandpass filter (see Figure 8) based on the design proposed in [58]. The filter is defined over the frequency range [2, 3] GHz on a substrate with height $h = 0.5$ mm, relative permittivity $\epsilon_r = 2.2$ and loss tangent $tan\delta = 0.003$. The width of both the horizontal and vertical conductors is 0.4 mm. This filter has a very narrow pass-bandwidth around [2.4, 2.6] GHz and a center frequency at 2.5 GHz.
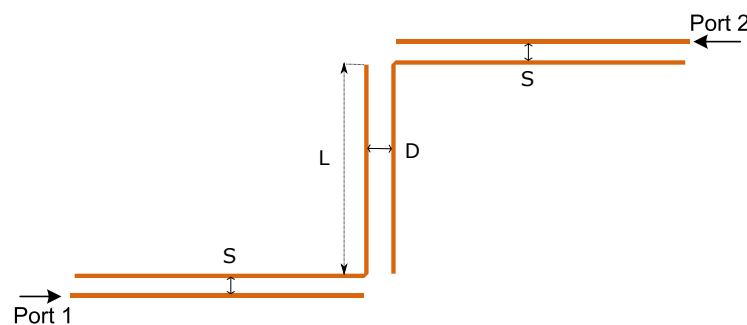

Figure 8. Top-view of the zigzag bandpass filter.

Since the bandwidth is very sensitive to the filter's geometrical parameters, three design parameters are considered: the distance D, the length L of the two coupling parts and the gap S between the horizontal conductors (see Figure 8). These parameters are defined in the supports [0.2, 2] mm, [16, 20] mm and [0.1, 1] mm, respectively. The goal is to design a filter with passband in the range [2.4, 2.6] GHz. For simplicity, we only consider performance specifications imposed on the reflection coefficient (element S_{11} of the filter scattering matrix). In order to achieve our optimization goal, the objective is formulated as follows:

$$f(\mathbf{x}) = |f_L - 2.4| + |f_H - 2.6| .$$

where f_L and f_H are the frequencies determining the intended operating bandwidth expressed in GHz. The perturbation on the design parameters has the following covariance matrix $\Sigma = \text{diag}(0.03^2, 0.06^2, 0.9^2)$, which corresponds to around 5% of their nominal values.

The optimization starts with 15 initial points, and continues for 100 iterations adding up to 115 (D,L,S) samples in total. As for the previous example, 16 sets of different initial samples are chosen via Latin Hypercube Design and the optimization process is repeated for each sample set. Differently from the example in Section 4.1, where the scattering parameters of the filter are computed using a quasi-analytical model in MATLAB, here, full-wave simulations, using Advanced Design System (Momentum EEsof EDA, Keysight Technologies) [41], are adopted to calculate the reflection coefficient. Due to the high associated computational cost, it was not feasible to calculate the E_r measure over all 100 iterations. Therefore, we present the results only for the last iteration of the optimization loop. The results are shown in Table 2.

Table 2. Median across 16 replications.

Method	Corresponding E_r Value
BO	0.145
Robust BO	0.137

As observed, the proposed methodology indeed achieves a lower E_r value compared to the BO method with respect to the small perturbations imposed. The global minimum found by BO is $\mathbf{x}_{BO} = (0.3, 0.7, 18)$, leading to $f(\mathbf{x}_{BO}) = 4.4e - 16$. The robust minimum defined by metric E_r lies at $\mathbf{x}_{RBO} = (0.5, 1.1, 17)$, corresponding to $f(\mathbf{x}_{RBO}) = 0.089$. The results of the MC analysis performed around $\mathbf{x}_{BO}$ and $\mathbf{x}_{RBO}$ are shown in Figure 9: despite being a local optimum, the design solution individuated by the novel algorithm is more robust to the perturbation δ on the design parameters.
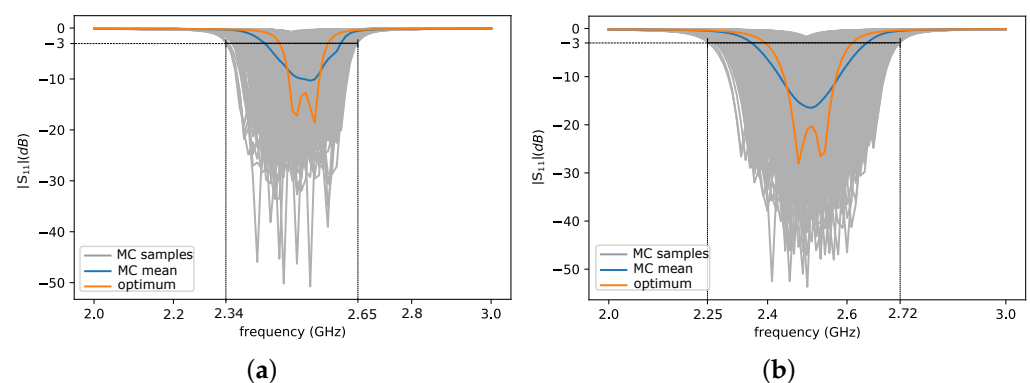

(a)   **(b)**

Figure 9. (**a**) Return loss $|S_{11}|$ of the zigzag bandpass filter using robust BO and (**b**) standard BO. (**a**) The optimal 3 dB bandwidth by the robust BO is [2.47, 2.58] GHz. MC analysis reveals, however, that it may extend up to [2.34, 2.65] GHz. (**b**) The optimal 3 dB bandwidth by the standard BO is [2.40, 2.61] GHz. MC analysis reveals, however, that it may extend up to [2.25, 2.72] GHz.

5. Conclusions

We presented a framework for robust optimization within the standard Bayesian optimization framework, which considers input uncertainty. The standard BO framework is extended to the novel robust methodology by using the sEI, a modified version of the EI acquisition function, as the sampling strategy. sEI takes a measure of robustness (E_r) into account to converge to a robust optimum rather than a global one. Such a robust solution may be a local optimum, however, it is more stable under perturbations compared to the global optimum obtained by standard BO. GPs are employed as the surrogate model and the uncertainty on the input parameters is propagated by moment matching the predictive

distribution. Moreover, our advocated robust optimization is performed in a data-efficient manner: the number of the expensive function evaluations is limited owing to the BO scheme. The effectiveness of the optimization framework is demonstrated on two filter design examples. Clearly in both cases, a lower median of the measure E_r, as well as a lower variance across replications of the optimization method, is observed.

Several extensions and improvements to the current approach are possible. In particular, we aim at extending the current approach to high dimensional inputs. Another improvement would be considering an input uncertainty distribution other than Gaussian. Moreover, for engineering design, a multi-objective version of a robust optimization method is very relevant.

Author Contributions: Conceptualization, D.S.; Data curation, D.D.W.; Formal analysis, D.D.W., J.Q. and D.V.G.; Funding acquisition, D.V.G.; Investigation, D.D.W.; Methodology, I.C., T.D. and D.S.; Resources, D.V.G.; Software, I.C. and T.D.; Supervision, D.V.G. and D.S.; Validation, D.D.W.; Writing—original draft, D.D.W.; Writing—review & editing, J.Q., I.C., T.D., D.V.G. and D.S. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. El Misilmani, H.; Naous, T.; Al Khatib, S. A Review on the Design and Optimization of Antennas Using Machine Learning Algorithms and Techniques. *Int. J. Microw.-Comput.-Aided Eng.* **2020**, *30*, e22356. [CrossRef]
2. Cingoska, M.V.; Sarac, V.J.; Gelev, S.A.; Cingoski, V.T. Efficiency optimization of electrical devices. In Proceedings of the 2018 23rd International Scientific-Professional Conference on Information Technology (IT), Zabljak, Montenegro, 19–24 February 2018; pp. 1–4. [CrossRef]
3. Li, Y.; Lei, G.; Bramerdorfer, G.; Peng, S.; Sun, X.; Zhu, J. Machine Learning for Design Optimization of Electromagnetic Devices: Recent Developments and Future Directions. *Appl. Sci.* **2021**, *11*, 1627. [CrossRef]
4. Mahdy, A. A numerical method for solving the nonlinear equations of Emden-Fowler models. *J. Ocea. Eng. Sci.* **2022**. [CrossRef]
5. Akinyemi, L.; Akpan, U.; Veeresha, P.; Rezazadeh, H.; Inc, M. Computational techniques to study the dynamics of generalized unstable nonlinear Schrödinger equation. *J. Ocean Eng. Sci.* **2022**. [CrossRef]
6. Cool, V.; Jonckheere, S.; Deckers, E.; Desmet, W. Black box stability preserving reduction techniques in the Loewner framework for the efficient time domain simulation of dynamical systems with damping treatments. *J. Sound Vib.* **2022**, *529*, 116922. [CrossRef]
7. Pietrenko-Dabrowska, A.; Koziel, S. Cost-Efficient EM-Driven Size Reduction of Antenna Structures by Multi-Fidelity Simulation Models. *Electronics* **2021**, *10*, 1536. [CrossRef]
8. Wang, G.G.; Shan, S. Review of Metamodeling Techniques in Support of Engineering Design Optimization. *J. Mech. Des.* **2006**, *129*, 370–380. [CrossRef]
9. Garbaya, A.; Kotti, M.; Fakhfakh, M.; Tlelo-Cuautle, E. Surrogate Assisted Optimization for Low-Voltage Low-Power Circuit Design. *J. Low Power Electron. Appl.* **2020**, *10*, 20. [CrossRef]
10. Younis, S.; Saleem, M.M.; Zubair, M.; Zaidi, S.M.T. Multiphysics design optimization of RF-MEMS switch using response surface methodology. *Microelectron. J.* **2018**, *71*, 47–60. [CrossRef]
11. Garg, L. Variability Aware Transistor Stack Based Regression Surrogate models for Accurate and Efficient Statistical Leakage Estimation. *Microelectron. J.* **2017**, *69*, 1–19. [CrossRef]
12. Knudde, N.; Couckuyt, I.; Spina, D.; Łukasik, K.; Barmuta, P.; Schreurs, D.; Dhaene, T. Data-Efficient Bayesian Optimization with Constraints for Power Amplifier Design. In Proceedings of the 2018 IEEE MTT-S International Conference on Numerical Electromagnetic and Multiphysics Modeling and Optimization (NEMO), Reykjavik, Iceland, 8–10 August 2018; pp. 1–3.
13. Gao, Y.; Yang, T.; Bozhko, S.; Wheeler, P.; Dragičević, T. Filter Design and Optimization of Electromechanical Actuation Systems Using Search and Surrogate Algorithms for More-Electric Aircraft Applications. *IEEE Trans. Transp. Electrif.* **2020**, *6*, 1434–1447. [CrossRef]
14. Mahouti, T.; Yildirim, T.; Kuşkonmaz, N. Artificial intelligence–based design optimization of nonuniform microstrip line band pass filter. *Int. J. Numer. Model. Electron. Netw. Devices Fields* **2021**, *34*, e2888. [CrossRef]

15. Afacan, E.; Lourenço, N.; Martins, R.; Dündar, G. Review: Machine learning techniques in analog/RF integrated circuit design, synthesis, layout, and test. *Integration* **2021**, *77*, 113–130. [CrossRef]

16. Lei, G.; Zhu, J.; Guo, Y.; Liu, C.; Ma, B. A Review of Design Optimization Methods for Electrical Machines. *Energies* **2017**, *10*, 1962. [CrossRef]

17. Gazda, C.; Vande Ginste, D.; Rogier, H.; Wu, R.B.; De Zutter, D. A Wideband Common-Mode Suppression Filter for Bend Discontinuities in Differential Signaling Using Tightly Coupled Microstrips. *IEEE Trans. Adv. Packag.* **2010**, *33*, 969–978. [CrossRef]

18. Barrico, C.; Antunes, C.H. Robustness Analysis in Multi-Objective Optimization Using a Degree of Robustness Concept. In Proceedings of the 2006 IEEE International Conference on Evolutionary Computation, Vancouver, BC, Canada, 16–21 July 2006; pp. 1887–1892.

19. Beyer, H.G.; Sendhoff, B. Robust optimization – A comprehensive survey. *Comput. Methods Appl. Mech. Eng.* **2007**, *196*, 3190–3218. [CrossRef]

20. Lei, G.; Zhu, J.; Liu, C.; Ma, B. Robust design optimization of electrical machines and drive systems for high quality mass production. In Proceedings of the 2016 6th International Electric Drives Production Conference (EDPC), Nuremberg, Germany, 30 November–1 December 2016; pp. 217–223. [CrossRef]

21. Lei, G.; Zhu, J.G.; Guo, Y.G.; Hu, J.F.; Xu, W.; Shao, K.R. Robust Design Optimization of PM-SMC Motors for Six Sigma Quality Manufacturing. *IEEE Trans. Magn.* **2013**, *49*, 3953–3956. [CrossRef]

22. Orosz, T.; Rassõlkin, A.; Kallaste, A.; Arsénio, P.; Pánek, D.; Kaska, J.; Karban, P. Robust Design Optimization and Emerging Technologies for Electrical Machines: Challenges and Open Problems. *Appl. Sci.* **2020**, *10*, 6653. [CrossRef]

23. Qing, J.; Couckuyt, I.; Dhaene, T. A Robust Multi-Objective Bayesian Optimization Framework Considering Input Uncertainty. *arXiv* **2022**, arXiv:2202.12848.

24. Xia, B.; Ren, Z.; Koh, C.S. Utilizing Kriging Surrogate Models for Multi-Objective Robust Optimization of Electromagnetic Devices. *IEEE Trans. Magn.* **2014**, *50*, 693–696. [CrossRef]

25. Fishman, G. *Monte Carlo: Concepts, Algorithms, and Applications*; Springer Series in Operations Research and Financial Engineering; Springer: New York, NY, USA, 2013.

26. Fonseca, J.R.; Friswell, M.I.; Lees, A.W. Efficient robust design via Monte Carlo sample reweighting. *Int. J. Numer. Methods Eng.* **2007**, *69*, 2279–2301. [CrossRef]

27. Wang, H.; Gong, Z.; Huang, H.Z.; Zhang, X.; Lv, Z. System Reliability Based Design Optimization with Monte Carlo simulation. In Proceedings of the 2012 International Conference on Quality, Reliability, Risk, Maintenance, and Safety Engineering, Chengdu, China, 15–18 June 2012; pp. 1143–1147.

28. Rasmussen, C.E.; Williams, C.K.I. *Gaussian Processes for Machine Learning*; MIT Press: Cambridge, MA, USA, 2006.

29. Knudde, N. Gaussian Processes for Modelling and Optimization of Engineering Systems. Ph.D. Thesis, Ghent University, Ghent, Belgium, 2020.

30. Rosich, A.; López, C.; Dewangan, P.; Abedrabbo, G. Robust Design Optimization of Mechatronics Systems: Parallel Electric Drivetrain Application. *Proc. Des. Soc.* **2022**, *2*, 1727–1736. [CrossRef]

31. Chatterjee, T.; Friswell, M.I.; Adhikari, S.; Chowdhury, R. A global two-layer meta-model for response statistics in robust design optimization. *Eng. Optim.* **2022**, *54*, 153–169. [CrossRef]

32. Li, W.; Gao, L.; Garg, A.R.; Xiao, M. Multidisciplinary robust design optimization considering parameter and metamodeling uncertainties. *Eng. Comput.* **2020**, *38*, 191–208. [CrossRef]

33. Sengupta, M.; Saxena, S.; Daldoss, L.; Kramer, G.; Minehane, S.; Cheng, J. Application specific worst case corners using response surfaces and statistical models. In Proceedings of the International Symposium on Signals, Circuits and Systems. Proceedings, SCS 2003. (Cat. No.03EX720), San Jose, CA, USA, 22–24 March 2004; pp. 351–356.

34. Zhang, B.; Rahmat-Samii, Y. Robust Optimization With Worst Case Sensitivity Analysis Applied to Array Synthesis and Antenna Designs. *IEEE Trans. Antennas Propag.* **2018**, *66*, 160–171. [CrossRef]

35. Manfredi, P.; Trinchero, R. A Probabilistic Machine Learning Approach for the Uncertainty Quantification of Electronic Circuits Based on Gaussian Process Regression. *IEEE Trans.-Comput.-Aided Des. Integr. Circuits Syst.* **2021**. [CrossRef]

36. De Ridder, S.; Spina, D.; Toscani, N.; Grassi, F.; Vande Ginste, D.; Dhaene, T. Machine-Learning-Based Hybrid Random-Fuzzy Uncertainty Quantification for EMC and SI Assessment. *IEEE Trans. Electromagn. Compat.* **2020**, *62*, 2538–2546. [CrossRef]

37. Kan, D.; Ridder, S.D.; Spina, D.; Couckuyt, I.; Grassi, F.; Dhaene, T.; Rogier, H.; Vande Ginste, D. Machine Learning-Based Hybrid Random-Fuzzy Modeling Framework For Antenna Design. In Proceedings of the 2020 14th European Conference on Antennas and Propagation (EuCAP), Copenhagen, Denmark, 15–20 March 2020; pp. 1–5.

38. Prasad, A.K.; Roy, S. Reduced Dimensional Chebyshev-Polynomial Chaos Approach for Fast Mixed Epistemic-Aleatory Uncertainty Quantification of Transmission Line Networks. *IEEE Trans. Compon. Packag. Manuf. Technol.* **2019**, *9*, 1119–1132. [CrossRef]

39. Pietrenko-Dabrowska, A.; Koziel, S. Optimization-based robustness enhancement of compact microwave component designs with response feature regression surrogates. *Knowl.-Based Syst.* **2022**, *240*, 108161. [CrossRef]

40. Murphy, T.; Tsui, K.L.; Allen, J. A Review of Robust Design Methods for Multiple Responses. *Res. Eng. Des.* **2005**, *15*, 201–215. [CrossRef]

41. Keysight EESof EDA. Advanced Design System. Available online: https://www.keysight.com/be/en/assets/7018-01027/brochures/5988-3326.pdf (accesed on 19 July 2022).

42. Viana, F.A.C.; Venter, G.; Balabanov, V. An algorithm for fast optimal Latin hypercube design of experiments. *Int. J. Numer. Methods Eng.* **2010**, *82*, 135–156. [CrossRef]
43. I. Frazier, P. A Tutorial on Bayesian Optimization. *arXiv* **2018**, arXiv:1807.02811.
44. Shahriari, B.; Swersky, K.; Wang, Z.; Adams, R.P.; de Freitas, N. Taking the Human Out of the Loop: A Review of Bayesian Optimization. *Proc. IEEE* **2016**, *104*, 148–175. [CrossRef]
45. Springenberg, J.T.; Klein, A.; Falkner, S.; Hutter, F. Bayesian Optimization with Robust Bayesian Neural Networks. In Proceedings of the Advances in Neural Information Processing Systems, Barcelona, Spain, 5–10 December 2016; Lee, D.D., Sugiyama, M., Luxburg, U.V., Guyon, I., Garnett, R., Eds.; Curran Associates, Inc.: Red Hook, NY, USA, 2016; Volume 29, pp. 4134–4142.
46. MacKay, D.J.C. *Information Theory, Inference, and Learning Algorithms*; Cambridge University Press: Cambridge, MA, USA, 2003; Chapter 45.
47. Močkus, J. On Bayesian methods for seeking the extremum. In Proceedings of the Optimization Techniques IFIP Technical Conference, Novosibirsk, Russia, 1–7 July 1974; Marchuk, G.I., Ed.; Springer: Berlin/Heidelberg, Germany, 1975.
48. Jones, D.R.; Schonlau, M.; Welch, W.J. Efficient global optimization of expensive black-box functions. *J. Glob. Optim.* **1998**, *13*, 455–492. [CrossRef]
49. Kushner, H.J. A new method of locating the maximum point of an arbitrary multipeak curve in the presence of noise. *J. Basic Eng.* **1964**, *86*, 97–106. [CrossRef]
50. Gelbart, M.; Snoek, J.; Adams, R. Bayesian optimization with unknown constraints. In Proceedings of the 30th Conference on Uncertainty in Artificial Intelligence, Quebec City, QC, Canada, 23–27 July 2014; AUAI Press: Arlington, TX, USA, 2014; pp. 250–259.
51. Deb, K.; Gupta, H. Introducing Robustness in Multi-Objective Optimization. *Evol. Comput.* **2006**, *14*, 463–494. [CrossRef]
52. Mirjalili, S.; Lewis, A. Obstacles and difficulties for robust benchmark problems: A novel penalty-based robust optimisation method. *Inf. Sci.* **2016**, *328*, 485–509. [CrossRef]
53. Fubini, G. Sugli integrali multipli. *Rom. Acc. L. Rend.* **1907**, *16*, 608–614.
54. Girard, A.; Rasmussen, C.E.; Murray-Smith, R. *Gaussian Process Priors with Uncertain Inputs: Multiple-Step-Ahead Prediction*; University of Glasgow: Glasgow, UK, 2002.
55. Brillinger, D. The calculation of cumulants via conditioning. *Ann. Inst. Stat. Math.* **1969**, *21*, 215–218. [CrossRef]
56. Knudde, N.; van der Herten, J.; Dhaene, T.; Couckuyt, I. GPFlowOpt: A Bayesian optimization library using TensorFlow. In Proceedings of the Neural Information Processing Systems 2017-Workshop on Bayesian Optimization, Long Beach, CA, USA, 4–9 December 2017; pp. 1–5.
57. Borazjani, O.; Rezaee, A. Design, Simulation and Construction a Low Pass Microwave Filters on the Micro Strip Transmission Line. *Int. J. Comput. Theory Eng.* **2012**, *4*, 784–787. [CrossRef]
58. Puttadilok, D.; Eungdamrong, D.; Tanacharoenwat, W. A study of narrow-band and compact size microstrip bandpass filters for wireless communications. In Proceedings of the SICE Annual Conference 2007, Takamatsu, Japan, 17–20 September 2007; pp. 1418–1421.