*Electronics Letters*

LETTER OPEN ACCESS

# Design of a Fabrication-Tolerant Directional Coupler via Robust Bayesian Optimization

Thijs Ullrick[1,2] | Yichen Liu[2] | Jixiang Qingy[3] | Wim Bogaerts[2] | Tom Dhaene[1]

[1]IDLab, Department of Information Technology, Ghent University - imec, Ghent, Belgium | [2]Photonics Research Group, Department of Information Technology, Ghent University - imec, Ghent, Belgium | [3]Department of Computing, Imperial College London, London, UK

**Correspondence:** Thijs Ullrick (thijs.ullrick@ugent.be)

**Received:** 12 October 2025 | **Revised:** 9 February 2026 | **Accepted:** 19 February 2026

## ABSTRACT

The performance of silicon photonic devices is highly sensitive to fabrication-induced deviations in waveguide core dimensions from their design values. This motivates the development of optimization algorithms that explicitly account for such uncertainties and can identify robust designs, that is, designs whose performance is insensitive to fabrication-induced variations. In this work, a novel scheme based on robust Bayesian optimization (RBO) is proposed, which leverages well-established robustness measures and aims to minimize deviations from nominal performance under input uncertainty. The study focuses on an S-bend directional coupler (DC), but the proposed methodology is general and can be applied to a wide range of passive and active photonic components.

## 1 | Introduction

Silicon photonics has emerged as the leading platform for large-scale, high-complexity, and cost-effective photonic integrated circuits. A unique feature of this technology platform is the high refractive index contrast of silicon-on-insulator (SOI) substrates, which facilitates sub-micrometer waveguide geometries and allows for high-density integration of optical functionalities on a single chip. Though a key enabler for dense integration, the high index contrast in silicon photonics also means that nanometer-scale deviations in waveguide dimensions can significantly alter the effective index, impacting phase and coupling performance in photonic circuits. This means that variability introduced by the fabrication process can have a significant impact on the overall performance of a circuit. To mitigate these effects, this paper presents a data-efficient optimization scheme based on RBO for the design of fabrication-tolerant devices.

A fundamental building block in many photonic circuits is the DC. However, its coupling characteristics exhibit a strong dependence on waveguide dimensions; even minor deviations from design specifications can result in significant discrepancies between the intended and actual coupling ratio, potentially degrading overall system performance. In this work, the proposed method is applied to identify a design configuration of the DC with reduced sensitivity to linewidth variations. Furthermore, the results obtained are validated against those reported in a prior study on a similar device [1].

## 2 | Preliminary

Passive building blocks, such as DCs, are typically designed to achieve a specific power-splitting ratio at some target wavelength. The objective of this work is to identify design configurations that not only satisfy these specifications but also minimize the expected deviation from nominal performance caused by fabrication-induced variations. This problem is addressed through robust optimization, where uncertainties in the design parameters are explicitly accounted for during the search for an optimal configuration. The robust design problem can be formulated as the minimization of the expected squared error

under input uncertainty

$$\min_{\mathbf{x}\in\Omega} \quad L(\mathbf{x}) = \mathbb{E}_{\xi}\left[\left(f(\mathbf{x}+\boldsymbol{\xi}) - y_0\right)^2\right] \tag{1}$$

where $\mathbf{x}$ is the vector holding the design variables, $\Omega$ denotes the prespecified design space, $\boldsymbol{\xi}$ represents the perturbation on the design parameters, $f(\mathbf{x})$ represents the simulated device response, and $y_0$ is the target value describing the desired device performance. The elements of $\boldsymbol{\xi}$ are assumed to be independently and normally distributed, that is, $\boldsymbol{\xi} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma})$, where the covariance matrix $\boldsymbol{\Sigma}$ is diagonal, indicating that perturbations to the design variables are uncorrelated.

Bayesian optimization (BO) is a sample-efficient global optimization framework particularly well-suited for photonic design tasks, where each function evaluation involves computationally expensive electromagnetic simulations. BO begins with a limited set of initial training samples and constructs a probabilistic surrogate model—typically a Gaussian process (GP)—to approximate the objective function [2]. The GP is defined by its covariance function, which encodes the correlation between pairs of input variables. Given a set of observed data, the GP model provides a probabilistic estimate at any input $\mathbf{x}$

$$\mathcal{GP}(\mathbf{x}) \sim \mathcal{N}(\mu(\mathbf{x}), \sigma^2(\mathbf{x})), \tag{2}$$

where $\mu(\mathbf{x})$ and $\sigma^2(\mathbf{x})$ denote the predictive mean and variance of the posterior distribution, respectively. As new samples are acquired, the GP is iteratively updated. Leveraging a computationally inexpensive acquisition function derived from the GP's posterior, BO effectively balances the exploration of unknown regions and exploitation of known optima in the design space.

Minimizing the objective function defined in (1) poses a significant challenge, as it constitutes a non-observable quantity that cannot be directly modeled using a GP. A related effort is the recent work by Boning et al. [3], which formulates the design of fabrication-tolerant devices as a Pareto front optimization between the squared error and a robustness metric derived from the second derivative of the GP. While demonstrated as a scalable and effective framework, this work takes a different approach based on established Bayesian robustness measures from the machine learning literature. By directly integrating these robustness measures within the acquisition function, a highly data-efficient and automated design procedure is proposed, explicitly tailored to minimize the objective defined in (1).

## 3 | Proposed Method

Two widely adopted robustness measures for uncertainty quantification are the mean and variance [4], which correspond to the first and second central moments of the objective function distribution induced by input uncertainty, respectively. They are defined as

$$\begin{aligned} \mathbb{J}_{\xi}(f) &= \mathbb{E}_{\xi}\left[f(\mathbf{x}+\boldsymbol{\xi})\right] \\ \mathbb{V}_{\xi}(f) &= \mathbb{E}_{\xi}\left[f^2(\mathbf{x}+\boldsymbol{\xi})\right] - \left[\mathbb{E}_{\xi}\left[f(\mathbf{x}+\boldsymbol{\xi})\right]\right]^2 \end{aligned} \tag{3}$$

To facilitate the use of these robustness measures to solve (1), the squared error objective function $L(\mathbf{x})$ in (1) is reformulated in terms of the mean and variance. Leveraging the linearity property of the expectation operator, $L(\mathbf{x})$ can be rewritten as

$$L(\mathbf{x}) = \mathbb{E}_{\xi}\left[f(\mathbf{x}+\boldsymbol{\xi})^2\right] - 2y_0\mathbb{E}_{\xi}\left[f(\mathbf{x}+\boldsymbol{\xi})\right] + y_0^2 \tag{4}$$

Then, using (3), it is easy to show that

$$L(\mathbf{x}) = \mathbb{V}_{\xi}(f) + \mathbb{J}_{\xi}(f)^2 - 2y_0\mathbb{J}_{\xi}(f) + y_0^2 \tag{5}$$

Within the context of robust optimization, accurately inferring the latent robustness measures is critical. While the mean robustness measure is analytically tractable under the GP framework [5], the variance is much more challenging to infer. A common workaround in practice involves approximating the variance by applying Monte Carlo sampling on the GP posterior mean. Although this approach provides a convenient point estimate, it does not capture the full distributional properties of the variance and, as demonstrated in [6], results in a biased estimate of its expected value. To overcome this limitation, an effective inference scheme for the mean and variance, leveraging a spectral representation of the GP via Fourier feature-based kernel approximation, was presented in [6]. Accordingly, these robustness measures are adopted in the present study to enable accurate and tractable optimization of (5) under input uncertainty.

A GP can be viewed as a generalization of a parametric Bayesian linear model in weight-space, where the latent function is expressed as $f(\mathbf{x}) = \boldsymbol{\phi}(\mathbf{x})^T\boldsymbol{\theta}$, with $\boldsymbol{\phi}(\mathbf{x})$ representing a feature mapping function and $\boldsymbol{\theta}$ denoting a vector of weights. Applying the kernel trick, an infinite number of features can be implicitly replaced by a positive-definite kernel function $k(\mathbf{x}, \mathbf{x}') = \boldsymbol{\phi}(\mathbf{x})^{\top}\boldsymbol{\phi}(\mathbf{x}')$, allowing to compute the posterior distribution without explicitly constructing the feature mapping.

In [6], this duality is used to approximate the squared exponential kernel in weight-space. Specifically, by invoking Bochner's theorem (e.g., [2]), which asserts that any stationary kernel $k$ and its spectral density $S(\omega)$ are Fourier duals, the kernel is approximated by a finite number of Fourier features $k(\mathbf{x}, \mathbf{x}') \approx \boldsymbol{\phi}^{\top}(\mathbf{x})\boldsymbol{\phi}(\mathbf{x}')$, where $\boldsymbol{\phi}(\mathbf{x}) := [\phi_1(\mathbf{x}), \phi_2(\mathbf{x}), \ldots, \phi_{N_f}(\mathbf{x})]^{\top}$ denotes a vector of $N_f$ parametric feature functions. Leveraging this formulation, the GP posterior can be asymptotically approximated as a Bayesian linear model [7]. This approximation is also referred to as the spectral representation of GPs

$$\mathrm{GP}(f) \dot{\sim} \boldsymbol{\phi}(\mathbf{x})^T\boldsymbol{\theta} \tag{6}$$

where $\dot{\sim}$ denotes an approximation of distributions.

In [6], two different types of Fourier features are introduced: Random Fourier features (RFF) and quadrature fourier features (QFF). RFFs are known to suffer from variance starvation, leading to potentially inaccurate approximations of the GP posterior if the training data size is large. Consequently, QFFs—designed to mitigate the starvation problem—are adopted in the present work. For stationary kernels with a decomposable spectral density $S(\boldsymbol{\omega})$, QFFs yield significantly more accurate posterior approximations with the same number of features $N_f$ [8].

By leveraging the QFF-based approximation of the SE kernel, the posterior distributions of the robustness measures in (3) can be expressed in terms of the expectations $E_{\boldsymbol{\xi}}[\boldsymbol{\phi}(\mathbf{x}+\boldsymbol{\xi})_m^T]$ and $E_{\boldsymbol{\xi}}[\boldsymbol{\phi}(\mathbf{x}+\boldsymbol{\xi})_m \boldsymbol{\phi}(\mathbf{x}+\boldsymbol{\xi})_n^T]$. In [6], closed-form analytic expressions for these expectations—computed assuming a normally distributed input uncertainty—are derived. As a result, the robustness measures themselves admit analytical forms, eliminating the need for a degenerated variance approximation based on Monte Carlo sampling.

Because the computational cost scales cubically with the number of Fourier features $N_f$, and $N_f$ itself grows exponentially with the input dimensionality, the QFF- and RFF-based robustness measures are currently practical only for low- to medium-dimensional problems [6]. However, since the proposed acquisition function is agnostic to the specific inference mechanism used for robustness estimation, Monte Carlo sampling can still be adopted when QFF- or RFF-based inference becomes computationally infeasible.

Using the QFF-based closed-form analytical expressions for the mean and variance, the objective function in (5) can be formulated under the GP framework. Starting from this formulation, the novel Fourier feature based, mean variance considered, squared error, expected improvement acquisition function (FF-MV-SE-EI) is proposed

$$\begin{aligned}\alpha_{\text{FF-MV-SE-EI}} &= \int_{\psi^*}\int_{\mathbb{J},\mathbb{V}} l(L,\psi^*)p(\mathbb{J},\mathbb{V}\mid\psi^*)p(\psi^*\mid\mathcal{D})d\psi^* d\mathbb{J}d\mathbb{V} \\ &\approx \frac{1}{N}\sum_{i=1}^{N} l(\tilde{L},\tilde{\psi}_i^*\mid\boldsymbol{\theta}_i,\mathcal{D})\end{aligned} \tag{7}$$

where $\tilde{\psi}_i^*$ is extracted from the training data and $l(\tilde{L},\tilde{\psi}^*,\boldsymbol{\theta}_i) := \max(\tilde{L}-\tilde{\psi}^*,0)$. Since the joint probability $p(\mathbb{J},\mathbb{V}\mid\psi^*)$ is not directly available, the numerical integration in (7) is avoided. Instead, Monte Carlo sampling is applied to the posterior distributions of $\mathbb{J}$ and $\mathbb{V}$. This approximation can be efficiently implemented within the Fourier feature-based robustness framework, as joint samples of $\tilde{\mathbb{J}}_i$, $\tilde{\mathbb{V}}_i$ can be obtained directly by sampling $\boldsymbol{\theta}_i$.

The posterior variance of the robustness measures accounts for two sources of uncertainty: epistemic uncertainty, which arises due to limited training data, and input uncertainty, which results from variability in the design variables. This means that the inferred Bayes risk at a test point $\mathbf{x}_t$ may remain uncertain, even if the GP model has seen data at $\mathbf{x}_t$. Since the acquisition function cannot intrinsically handle this inconsistency, two possible issues may arise when integrating it into a standard BO scheme: First, the algorithm may waste resources by sampling duplicate candidates $\mathbf{x}_t$ in a futile attempt to reduce the non-zero posterior variance. Second, when sampling is constrained within the design space, the posterior variance cannot be lowered to zero, often resulting in redundant sampling near the boundaries. To resolve these issues, the two-stage acquisition process from [9] is adopted. At iteration $t$, the function $f(\mathbf{x}_t+\boldsymbol{\xi}_t)$ is evaluated according

$$\begin{aligned}\mathbf{x}_t &= \arg\max_{\mathbf{x}\in\mathcal{X}} \alpha_t(\mathbf{x}) \\ \tilde{\boldsymbol{\xi}}_t &= \arg\max_{\boldsymbol{\xi}\in\Delta} \sigma_t(\tilde{\mathbf{x}}_t+\boldsymbol{\xi})\end{aligned} \tag{8}$$

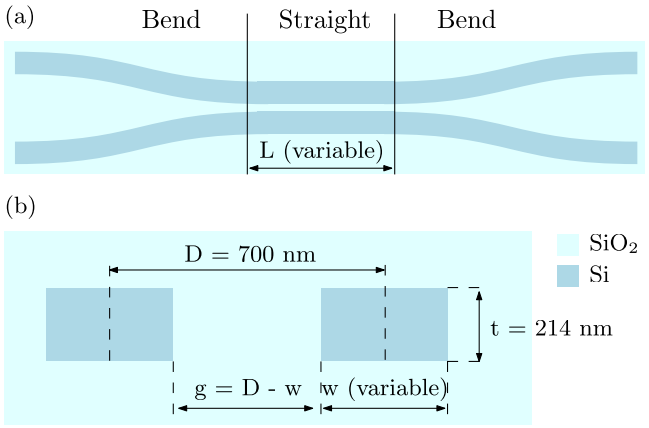


**FIGURE 1** | (a) Schematic diagram of a directional coupler showing the straight and the curved sections. (b) Diagram showing the cross-section through the straight section of the directional coupler.

where $\alpha_t(\cdot)$ and $\sigma_t(\cdot)$ denote the acquisition function defined in (7) and posterior standard deviation, respectively, both evaluated at iteration $t$. The active learning (AL) search space is given by $\Delta = [-\text{CI}_{0.025}, \text{CI}_{0.975}]$, $\text{CI}_{0.975} = Q(0.975\mid\boldsymbol{\xi})$ where Q is the quantile function built upon marginal distribution $\boldsymbol{\xi}$. The rationale behind this procedure is to avoid direct sampling at $\mathbf{x}_t$ by identifying a candidate that maximally reduces the model's uncertainty within its neighborhood. As argued in [10], it is only sensible to perform AL if one of these conditions is met: (i) $\mathbf{x}_t$ lies near the boundary of the design space, that is, $|\mathbf{x}_t-\mathbf{x}_{min}|<\epsilon$ or $|\mathbf{x}_t-\mathbf{x}_{max}|<\epsilon$; or (ii) optimization of $\alpha_t$ leads to duplicate sampling within the design space, that is, $|\mathbf{x}_t-\mathbf{X}|<\epsilon$. Here, $\mathbf{x}_{min}$ and $\mathbf{x}_{max}$ represent the lower and upper bounds of the design space, respectively and $\mathbf{X}$ represents the set of all previously evaluated design points.

## 4 | Numerical Results

The proposed RBO framework is applied to find the tolerant operating region of an S-bend DC. The schematic diagram of the device, featuring a $SiO_2$ substrate and cladding, is shown in Figure 1a. A cross-sectional view of its straight section is illustrated in Figure 1b. In this study, the device is designed for a 50:50 power coupling at a target wavelength of 1.55$\mu$m. This is achieved by varying the length $L$ of the straight coupling section within the range of [10.0, 20.0]$\mu$m and the waveguide width $w$ within [350, 600]nm. Due to spatially correlated variations in the waveguide width, inherent to the CMOS manufacturing process, two closely spaced waveguides tend to experience nearly identical dimensional changes. As a result, it is reasonable to model the width of the symmetrical waveguides using a single parameter $w$. Furthermore, this also means that the center line distance remains invariant under the fabrication tolerances. As demonstrated in [1], both finite-difference time-domain (FDTD) and EigenMode expansion (EME) simulations exhibit excellent convergence, yielding similar results. However, since power transmission needs to be evaluated at only a single wavelength, EME offers a significant reduction in CPU time per design configuration and is therefore adopted in this work.

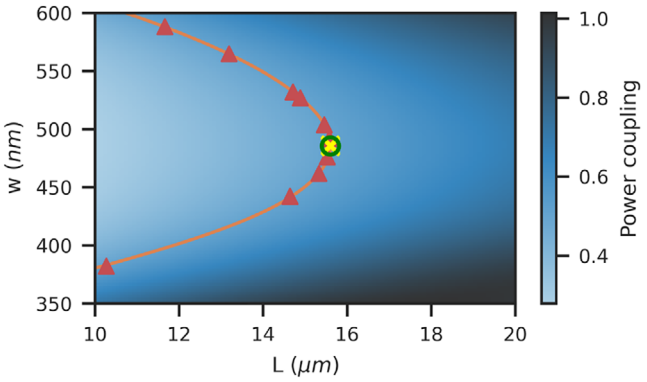


**FIGURE 2** | Regression surface generated using the reference GP. The orange contour indicates the set of global optimal solutions achieving the target 50:50 coupling. Red triangles denote solutions found by standard BO, yellow cross markers show robust configurations identified by RBO, and the green circle marks the robust optimum found via sampling of the reference GP.

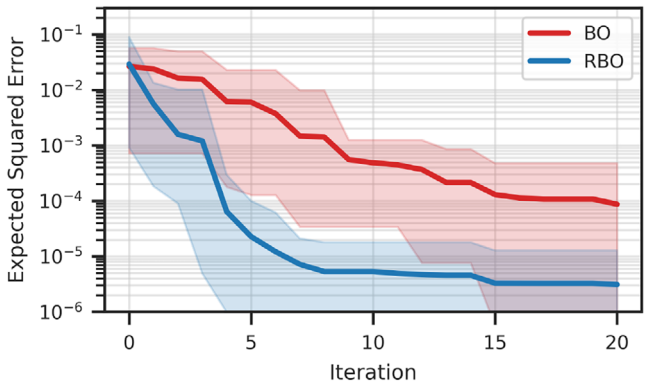


**FIGURE 3** | Progression of the cumulative best expected squared error for the RBO and standard BO schemes as a function of iteration number.

**TABLE 1** | Computational runtime associated with the different optimization schemes.

| Method | EME time (s) | Algorithm time (s) | Total time (s) |
|---|---|---|---|
| BO | 4003 | 151 | 4154 |
| Ref. GP | 20014 | 1313 | 21327 |
| RBO | 4003 | 719 | 4722 |

A study of process variations in imec's iSiPP50G technology platform reports linewidth variations with a standard deviation of $\sigma_w = 4.6$nm [11]. Based on these findings, the input uncertainty considered during the robust Bayesian optimization (RBO) is modeled as

$$(\Delta w, \Delta L) \sim \mathcal{N}\left(0, \begin{bmatrix} \sigma_w^2 & 0.0 \\ 0.0 & 0.0 \end{bmatrix}\right) \tag{9}$$

The AL threshold of the two-step acquisition scheme is set to $\epsilon = 0.5$ nm. This threshold also defines an upper bound on the resolution of the optimization process: During function evaluation at the robust configuration, AL will be triggered if a previously sampled point lies within an $\epsilon$-radius. Ideally, the number of Fourier features should be set as high as possible. However, experiments on benchmark functions [6] indicate that $N_f \approx 32^d$, where $d$ is the input dimension, serves as a practical guideline. Following this guideline, the number of QFF features is set to $N_f = 1024$.

To quantitatively benchmark the robustness and efficiency of the proposed algorithm, it is compared against two baseline approaches. First, a GP model is trained on a dataset consisting of 100 samples generated via Latin hypercube sampling. Monte Carlo sampling is then applied to this reference model to estimate the expected squared error over a dense grid in the design space, from which the robust optimum is identified. Second, the proposed method is compared against standard BO employing an expected improvement (EI) acquisition function. The RBO framework is implemented using Trieste, an open-source Python library for Bayesian optimization built on TensorFlow. Both the BO and RBO schemes were executed 10 times, each starting from five samples and running for 20 iterations. To assess the sensitivity of the optimization schemes to the choice of the initial dataset, a different seed was used for each run.

The optimal parameter configurations identified by the RBO and standard BO methods are overlaid on the regression surface in Figure 2, which was generated using the reference GP. As discussed in [1], multiple $(w, L)$ configurations can yield the target 50:50 power coupling; however, only one corresponds to the robust optimum. Standard BO may converge to any configuration within the set of nominally optimal solutions, indicated by the orange contour in the surface plot. In contrast, RBO consistently converges to the robust optimum at $w_* = 486$ nm and $L_* = 15.6\,\mu$m, as identified through Monte Carlo sampling of the reference GP.

Figure 3 presents the cumulative best expected squared error for both the BO and RBO methods as a function of the iteration number. The curves report the mean over the ten runs, with the shaded regions indicating the best and worst outcomes. As the expected squared error is a latent quantity that cannot be directly observed, it is inferred for both methods using the reference GP. The results show that the proposed RBO method consistently converges to the robust optimum within ten iterations, independent of the choice of the initial dataset. In contrast, standard BO tends to converge to an arbitrary solution within the set of global optima.

The computational runtime for the three methods, including the time required to generate the corresponding EM training data, is summarized in Table 1. While inference of the QFF robustness measures incurs some overhead, the total runtime remains significantly lower than that required to train a GP surrogate on a brute-force parameter sweep. Notably, as simulation costs increase, the relative advantage of the proposed scheme becomes more pronounced.

It was shown in [1] that the DC power coupling exhibits a parabolic dependence on the waveguide width, as established through both numerical simulations and experimental measurements. The robust optimum corresponds to the minimum of this parabolic response, where the first derivative with respect to width vanishes, yielding a locally flat response. Leveraging

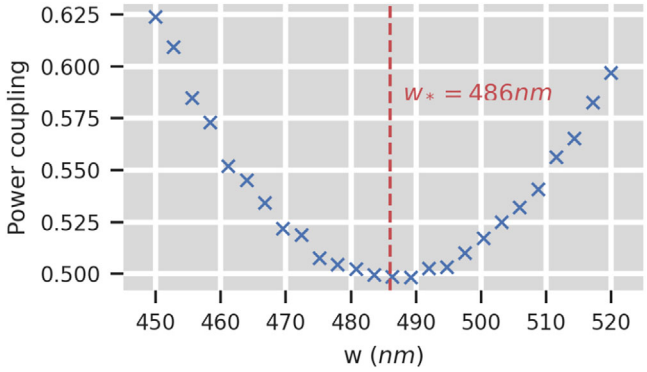


**FIGURE 4** | EME simulation of the cross transmission of the DC at $L = 15.6\ \mu$m.

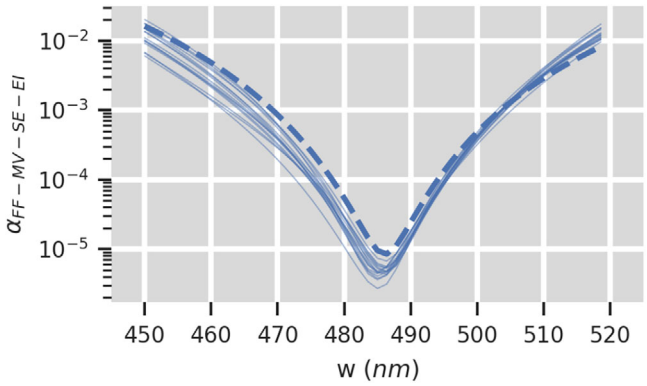


**FIGURE 5** | Plot of the acquisition function at $L_* = 15.6\ \mu$m. The solid blue lines represent the squared error acquisition function derived by sampling the QFF robustness measures. The blue dashed line indicates the acquisition function estimate obtained via MC sampling applied to the reference GP.

this insight, our robust solution is validated by performing a parameter sweep of the waveguide width $w$ over a narrow range around the identified optimum $w_*$, while keeping the coupler length fixed at $L_*$. The results in Figure 4 confirm that the robust optimum identified by the RBO scheme coincides with the minimum of the parabolic trend, demonstrating agreement between our simulation-based optimization and the experimental findings reported in [1]. A GP model is then trained on this dataset and used to estimate the expected squared error via Monte Carlo sampling. As shown in Figure 5, the acquisition function derived from the QFF robustness measures in the first RBO run closely matches the expected squared error obtained from the GP.

## 5 | Conclusion

The proposed acquisition function, derived from spectral robustness metrics, demonstrates high effectiveness in the design of robust devices and consistently guides RBO to the robust optimum within a limited number of iterations. The active learning policy prioritizes sampling in the vicinity of the optimum, resulting in more efficient use of computational resources and improving model accuracy near the robust design point. Moreover, because the spectral robustness measures admit closed-form expressions within the GP framework, they furnish a principled statistical basis for post-optimization yield estimation and variability analysis.


### Funding

This work was supported by Flemish Research Foundation (FWO-Vlaanderen G031421N) and the Flanders AI Research Program.


### Author Contributions

**Thijs Ullrick:** conceptualization, methodology, software, writing – original draft. **Yichen Liu:** investigation (photonic design), writing – review. **Jixiang Qing:** methodology, software, writing – review. **Wim Bogaerts:** supervision, writing – review. **Tom Dhaene:** supervision, writing – review.

### Conflicts of Interest

The authors declare no conflicts of interest

### Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

### References


1. Y. Liu, U. Khan, and W. Bogaerts, "Fabrication Tolerant Directional Coupler," in *24th European Conference on Integrated Optics (ECIO)* (Springer, 2023), https://www.ecio-conference.org/2023-proceedings.

2. C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*, ser. Adaptive Computation and Machine Learning (MIT Press, 2006), ocm61285753.

3. Z. Gao, Z. Zhang, Z. He, J. Gu, D. Z. Pan, and D. S. Boning, "Selecting Robust Silicon Photonic Designs After Bayesian Optimization Without Extra Simulations," *Optics Express* 32, no. 21 (Oct. 2024): 37585–37598, https://opg.optica.org/oe/abstract.cfm?uri=oe-32-21-37585.

4. P. Manfredi and R. Trinchero, "A Probabilistic Machine Learning Approach for the Uncertainty Quantification of Electronic Circuits Based on Gaussian Process Regression," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems* 41, no. 8 (Aug. 2022): 2638–2651, https://ieeexplore.ieee.org/abstract/document/9536589.

5. A. O'Hagan, "Bayes–Hermite Quadrature," *Journal of Statistical Planning and Inference* 29, no. 3 (Nov. 1991): 245–260, https://www.sciencedirect.com/science/article/pii/037837589190002V.

6. J. Qing, T. Dhaene, and I. Couckuyt, "Spectral Representation of Robustness Measures for Optimization Under Input Uncertainty," in *Proceedings of the 39th International Conference on Machine Learning* (PMLR, June 2022), 18096–18121, https://proceedings.mlr.press/v162/qing22a.html.

7. J. Wilson, V. Borovitskiy, A. Terenin, P. Mostowsky, and M. Deisenroth, "Efficiently Sampling Functions From Gaussian Process Posteriors," in *Proceedings of the 37th International Conference on Machine Learning* (PMLR, Nov. 2020), 10292–10302, https://proceedings.mlr.press/v119/wilson20a.html.

8. M. Mutny and A. Krause, "Efficient High Dimensional Bayesian Optimization With Additivity and Quadrature Fourier Features," in *Advances in Neural Information Processing Systems* vol. 31 (Curran Associates, Inc., 2018), https://proceedings.neurips.cc/paper/2018/hash/4e5046fc8d6a97d18a5f54beaed54dea-Abstract.html.

9. S. Iwazaki, Y. Inatsu, and I. Takeuchi, "Mean-Variance Analysis in Bayesian Optimization under Uncertainty," in *Proceedings of The 24th International Conference on Artificial Intelligence and Statistics* (PMLR, Mar. 2021), 973–981, https://proceedings.mlr.press/v130/iwazaki21a.html.

10. J. Qing, N. Knudde, D. Spina, I. Couckuyt, and T. Dhaene, "Bayesian Active Learning for Electromagnetic Structure Design," in *European Conference on Antennas and Propagation (EuCAP)* (2020).

11. Y. Xing, M. Wang, A. Ruocco, J. Geessels, U. Khan, and W. Bogaerts, "Compact Silicon Photonics Circuit to Extract Multiple Parameters for Process Control Monitoring," *OSA Continuum* 3, no. 2 (Feb. 2020): 379–390.