Uncertainty quantification of silicon photonic devices with correlated and non-Gaussian random parameters
preprint version

Tsui-Wei Weng, Zheng Zhang, Zhan Su,
Youssef Marzouk, Andrea Melloni, and Luca Daniel

Massachusetts Institute of Technology

February 2015

Abstract

Process variations can significantly degrade device performance and chip yield in silicon photonics. In order to reduce the design and production costs, it is highly desirable to predict the statistical behavior of a device before the final fabrication. Monte Carlo is the mainstream computational technique used to estimate the uncertainties caused by process variations. However, it is very often too expensive due to its slow convergence rate. Recently, stochastic spectral methods based on polynomial chaos expansions have emerged as a promising alternative, and they have shown significant speedup over Monte Carlo in many engineering problems. The existing literature mostly assumes that the random parameters are mutually independent. However, in practical applications such assumption may not be necessarily accurate. In this paper, we develop an efficient numerical technique based on stochastic collocation to simulate silicon photonics with correlated and non-Gaussian random parameters. The effectiveness of our proposed technique is demonstrated by the simulation results of a silicon-on-insulator based directional coupler example. Since the mathematic formulation in this paper is very generic, our proposed algorithm can be applied to a large class of photonic design cases as well as to many other engineering problems.

1 Introduction

The silicon photonic technology has emerged as a promising alternative to electrical interconnects due to its ability to achieve higher bandwidth and lower power dissipation [1]. Thanks to its ease of integration with the CMOS process, silicon photonics can remarkably reduce fabrication costs, increase the integration scale and improve the overall system performance [2]. However, silicon-based optical devices are very sensitive to fabrication process variations, leading to potentially significant device performance degradations and potential system failures [3–8].

There is a lack of uncertainty quantification techniques for silicon photonics, however some results have been reported recently for nanometer integrated circuits (IC) [9–12] and microelectromechanical systems (MEMS) [13, 14]. The mainstream stochastic simulation technique in commercial design software is Monte Carlo [15]. Despite some recent improvements [16], Monte Carlo

simulation is still inefficient for most design problems. This limitation is clearly evident in silicon photonics where reliable Monte Carlo results can be obtained only by performing a large number of simulations, resulting in a very high computational cost.

Generalized polynomial chaos (gPC) expansion techniques have been proposed [17] to efficiently approximate a stochastic solution dependent on a set of random parameters. Furthermore, they have recently been exploited in the analysis of photonic circuits [18]. Based on such expansions, fast stochastic spectral methods have been developed [19–21]. Standard spectral methods include an intrusive method we refer to as stochastic Galerkin (SG) [19], a non-intrusive method called stochastic collocation (SC) [20, 21], and their hybrid variant called stochastic testing [9–11]. In most existing publications [18, 22–25], the input parameters are assumed to be mutually independent, or Gaussian-correlated, in which case one can convert the random variables to a set of independent ones by a linear transform. However, in practical applications such assumption may not be necessarily accurate.

In this paper, we propose a computational framework to quantify the uncertainties in silicon photonic devices caused by non-Gaussian correlated random parameters. Our algorithm is based on the theoretical results of [26] which has been rarely utilized in practice due to relevant issues faced in its numerical implementation. In our work, we assume that the process variations are described by a Gaussian Mixture (GM). Gaussian mixture is quite a reasonable modeling technique for process variations, since it can capture both correlated and uncorrelated process variations. Based on this assumption, we first develop a numerical scheme to construct a set of orthogonal basis functions to represent the stochastic solution. We further propose an implementation of stochastic collocation to compute each of the coefficient associated with each of the basis functions. Differently from our preliminary conference paper [27], this paper provides also all the theoretical and numerical details for computing the basis functions and for implementing stochastic collocation with correlated random parameters. As an example, a silicon-on-insulator (SOI) based directional coupler is considered and the impact of tolerances on geometric dimensions is evaluated. Moreover, the accuracy of our proposed algorithm is compared with Monte Carlo in our numerical experiments.

This paper is structured as follows. In Section 2, we give a brief review of orthogonal basis functions with respect to a joint probability density measure, as well as some basics about Gaussian mixture models. Section 3 consists of two parts describing our proposed computational techniques. The first part demonstrates how to construct the orthogonal bases given a Gaussian-mixture density function; the second part presents how to implement efficiently a stochastic collocation scheme to obtain the stochastic solution. In Section 4, we report the simulation results of a SOI based directional coupler under process variations. Finally, Section 5 concludes this paper.

2 Background review

This section briefly reviews some background related to orthogonal basis functions [26] and Gaussian mixture models.

2.1 Orthogonal basis function construction

Let $\vec{\xi} = \{\xi_1, \cdots, \xi_N\} \in \mathbb{R}^N$ be N random variables with a given joint probability density function $p_{\vec{\xi}}$ on the support Ξ:

$$p_{\vec{\xi}}(\vec{\xi}) = 0, \quad \text{if} \quad \vec{\xi} \notin \Xi. \tag{1}$$

We further denote the marginal probability density function of the i-th random variable by $p_i(\xi_i)$:

$$p_i(\xi_i) = \int_{\Xi} p_{\vec{\xi}}(\xi_1, \cdots, \xi_N) d\xi_1 \cdots d\xi_{i-1} d\xi_{i+1} \cdots d\xi_N. \tag{2}$$

Assume that $\{\phi_k^{(i)}(\xi_i)\}$ are a set of generalized polynomial-chaos bases for ξ_i [17], where k denotes the degree of a polynomial. Such basis functions are orthogonal with respect to the marginal density function p_i:

$$\langle \phi_k^{(i)}, \phi_{k'}^{(i)} \rangle_{p_i} = \int_{\mathbb{R}} \phi_k^{(i)}(\xi_i) \phi_{k'}^{(i)}(\xi_i) p_i(\xi_i) d\xi_i = \|\phi_k^{(i)^2}\| \delta_{kk'} \tag{3}$$

where $\langle \rangle_{p_i}$ denotes an inner product defined by the marginal probability density function $p_i(\xi_i)$. In [26], it is shown that a set of N-dimensional orthogonal functions for the joint density $p_{\vec{\xi}}(\vec{\xi})$ can be constructed as

$$\Psi_{\vec{\alpha}}(\vec{\xi}) = \left[\frac{p_1(\xi_1) \cdots p_N(\xi_N)}{p_{\vec{\xi}}(\vec{\xi})} \right]^{1/2} \phi_{\alpha_1}^{(1)}(\xi_1) \cdots \phi_{\alpha_N}^{(N)}(\xi_N) \tag{4}$$

where the functionals $\Psi_{\vec{\alpha}}$ are defined for $\vec{\xi} \in \Xi$ and $\vec{\alpha} = (\alpha_1, \cdots, \alpha_N) \in \mathbb{N}^N$ is a multi-index. The proof of orthogonality of $\{\Psi_{\vec{\alpha}}(\vec{\xi})\}$ is shown in Appendix.

2.2 Gaussian mixture

In this paper, we assume that a Gaussian-mixture model has already been constructed based on some measurement data, using for instance the Expectation-Maximization (EM) algorithm [28]. We focus here instead on how to predict the effect on the performance of a silicon photonic device influenced by process variations. In practice, there are correlated and uncorrelated process variations. In this paper, we aim at developing numerical techniques that can deal with correlated process variations, since there are already many techniques available for simulating design problems with uncorrelated process variations.

A Gaussian mixture with M mixed terms is the weighted sum of M multivariate Gaussian densities. Purely for simplicity of exposition of mathematical derivations, we set here $N = M = 2$. However, all results presented in this paper are valid and can be extended in a straightforward way to the more general cases where N and M are not equal to 2 and N is not necessarily equal to M.

With a Gaussian mixture model, the process variations can be described by the following distribution

$$\begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix} \sim a \cdot N(\vec{\mu}_A, \Sigma_A) + b \cdot N(\vec{\mu}_B, \Sigma_B) \tag{5}$$

with

$$a + b = 1, \quad 0 \leq a, b \leq 1$$

and

$$\mu_A = \begin{bmatrix} \mu_{A_1} \\ \mu_{A_2} \end{bmatrix}, \quad \Sigma_A = \begin{bmatrix} \sigma_{A_1}^2 & \rho_A \sigma_{A_1} \sigma_{A_2} \\ \rho_A \sigma_{A_1} \sigma_{A_2} & \sigma_{A_2}^2 \end{bmatrix}$$

$$\mu_B = \begin{bmatrix} \mu_{B_1} \\ \mu_{B_2} \end{bmatrix}, \quad \Sigma_B = \begin{bmatrix} \sigma_{B_1}^2 & \rho_B \sigma_{B_1} \sigma_{B_2} \\ \rho_B \sigma_{B_1} \sigma_{B_2} & \sigma_{B_2}^2 \end{bmatrix}$$

where $\vec{\xi}$ can be the geometrical or material parameters of a silicon photonic device. It is well-known that the marginal probability density function of a multivariate Gaussian is a univariate Gaussian. Therefore, the i-th random parameter follows a univariate Gaussian-mixture distribution

$$\xi_i \sim a \cdot N(\mu_{A_i}, \sigma_{A_i}^2) + b \cdot N(\mu_{B_i}, \sigma_{B_i}^2), \tag{6}$$

and its marginal probability density function is

$$p_i(\xi_i) = \frac{a \cdot \exp(\frac{(\xi_\mathrm{i} - \mu_{\mathrm{A_i}})^2}{-2\sigma_{\mathrm{A_i}}^2})}{\sqrt{2\pi\sigma_{A_i}^2}} + \frac{b \cdot \exp(\frac{(\xi_\mathrm{i} - \mu_{\mathrm{B_i}})^2}{-2\sigma_{\mathrm{B_i}}^2})}{\sqrt{2\pi\sigma_{B_i}^2}}. \tag{7}$$

3 Uncertainty quantification with non-Gaussian correlated parameters

In this section, we first develop some numerical techniques to build the orthogonal basis functions introduced in Section 2. We then describe how to implement efficiently a stochastic collocation technique to calculate a stochastic quantity of interest that depends on some correlated non-Gaussian process variation parameters. This section provides all mathematical details to construct a stochastic solution based on stochastic collocation, and the flows are summarized in Algorithm 1 and 2. For readers who are more interested in the applications of the proposed computational technique, please refer to Section 4 for a directional coupler example.

3.1 Basis function computation

Let $u(x, \vec{\xi})$ be the quantity of interest, smoothly dependent on the process variation parameters. Here u can be for instance the effective phase index, the resonant wavelength or the power dissipation of a silicon photonic device. Vector $\vec{\xi}$ could contain for instance the dimensions and/or the material properties of a device, and x can be the frequency or wavelength. For example, in order to investigate the uncertainties caused by geometric variations of a directional coupler, we can set u to be the power coupling coefficient, and $\vec{\xi}$ to be the exterior and interior waveguide sidewall position. Meanwhile, x can be set as the operating frequency or wavelength that also influences the power coupling coefficient. Given the basis functions $\Psi_{\vec{\alpha}}(\vec{\xi})$ described in Section 2, we can approximate u as

$$u(x, \vec{\xi}) \approx \sum_{\|\vec{\alpha}\|_\infty \leq \vec{t}} C_{\vec{\alpha}}(x) \Psi_{\vec{\alpha}}(\xi_1, \cdots, \xi_N) \tag{8}$$

when u has a bounded variance. Here $C_{\vec{\alpha}}(x)$ is the coefficient of $\Psi_{\vec{\alpha}}(\xi_1, \cdots, \xi_N)$, $\vec{\alpha}$ is a multi-index which consists of N-tuples, and a full-tensor truncation is used for gPCe. The basis functions are still not known, and thus our first task is to compute them given the Gaussian-mixture joint density function.

In order to construct $\Psi_{\vec{\alpha}}$, we first need to compute the univariate basis functions $\phi_k^{(i)}(\xi_i)$'s defined in Eq. (3) for each random parameter ξ_i. The univariate basis functions can be built using the three term recurrence relations [29], and the construction of such univariate orthogonal polynomials is shown in the following. Since our numerical scheme is applicable to all random parameters ξ_i, the parameter index i in $\phi_k^{(i)}(\xi_i)$ and $p_i(\xi_i)$ will be omitted in the following derivation.

The three term recurrence relations is

$$\phi_{k+1}(\xi) = (\xi - a_k)\phi_k(\xi) + b_k\phi_{k-1}(\xi), \quad k \geq 1 \tag{9}$$

with $\phi_0 \equiv 1$ and $\phi_1 \equiv \xi - \frac{\langle 1,\xi\rangle_p}{\langle 1,1\rangle_p}$. In order to ensure that $\{\phi_k\}_{k\geq 0}$ is an orthogonal set with respect to measure p, the coefficients a_k and b_k are defined as

$$a_k = \frac{\langle \xi\phi_k, \phi_k\rangle_p}{\langle \phi_k, \phi_k\rangle_p} \tag{10}$$

$$b_k = \frac{\langle \phi_k, \phi_k\rangle_p}{\langle \phi_{k-1}, \phi_{k-1}\rangle_p}. \tag{11}$$

In order to compute a_k and b_k, we need to evaluate $\langle \xi\phi_k, \phi_k\rangle_p$ and $\langle \phi_k, \phi_k\rangle_p$, $\langle \phi_{k-1}, \phi_{k-1}\rangle_p$. Notice that

$$\langle \xi\phi_k, \phi_k\rangle_p = \int_{\mathbb{R}} \xi\phi_k^2(\xi)p(\xi)d\xi \tag{12}$$

with $p(\xi)$ defined in Eq. (7). It is non-trivial to obtain Eq. (12) in an analytical way. However, by letting $y_i = \frac{\xi-\mu_{A_i}}{\sigma_{A_i}}$ and $z_i = \frac{\xi-\mu_{B_i}}{\sigma_{B_i}}$, the integral in Eq. (12) can be rewritten as

$$\begin{aligned} &a\int_{\mathbb{R}} (\mu_{A_i} + \sigma_{A_i}y_i)\phi_k^2(\mu_{A_i} + \sigma_{A_i}y_i)\frac{\exp(\frac{y_i^2}{-2})}{\sqrt{2\pi}}dy_i \\ &+b\int_{\mathbb{R}} (\mu_{B_i} + \sigma_{B_i}z_i)\phi_k^2(\mu_{B_i} + \sigma_{B_i}z_i)\frac{\exp(\frac{z_i^2}{-2})}{\sqrt{2\pi}}dz_i. \end{aligned} \tag{13}$$

As a result, a Gauss-Hermite quadrature rule can be used to numerically compute Eq. (13) with high accuracy. For instance, the two terms in Eq. (13) can be computed as

$$a\sum_{j=1}^{q} w_j(\mu_{A_i} + \sigma_{A_i}x_j)\phi_k^2(\mu_{A_i} + \sigma_{A_i}x_j) + b\sum_{j=1}^{q} w_j(\mu_{B_i} + \sigma_{B_i}x_j)\phi_k^2(\mu_{B_i} + \sigma_{B_i}x_j) \tag{14}$$

where x_j and w_j are the j-th well-known Gauss-Hermite abscissa and weight respectively [30] , and q is the number of quadrature points. The inner products $\langle \phi_k, \phi_k\rangle_p$ and $\langle \phi_{k-1}, \phi_{k-1}\rangle_p$ can be calculated in a similar way:

$$\langle \phi_k, \phi_k\rangle_p = \int_{\mathbb{R}} \phi_k^2(\xi)p(\xi)d\xi \approx a\sum_{j=1}^{q} w_j\phi_k^2(\mu_{A_i} + \sigma_{A_i}x_j) + b\sum_{j=1}^{q} w_j\phi_k^2(\mu_{B_i} + \sigma_{B_i}x_j) \tag{15}$$

With a_k and b_k computed with the above procedures, the formula in Eq. (9) can be used to calculate the basis function for each parameter ξ_i. Finally, the multivariate basis function $\Psi_{\vec{\alpha}}$ can be obtained according to Eq. (4).

3.2 Implementation of stochastic collocation

The coefficients $C_{\vec{\alpha}}(x)$ in Eq. (8) can be computed based on stochastic collocation, and the implementation is detailed below.

Algorithm 1 Univariate basis function construction and Quadrature points computation

1: Initialization: $\phi_0 = 1$;
2: Compute $\langle \xi\phi_0, \phi_0 \rangle_p$ using Eq. (12) and $\langle \phi_0, \phi_0 \rangle_p$ using Eq. (15) to obtain ϕ_1;
3: **for** $k = 1, \cdots, n-1$ **do**
4: Compute $\langle \xi\phi_k, \phi_k \rangle_p$ using Eq. (12) and compute $\langle \phi_k, \phi_k \rangle_p$ using Eq. (15);
5: Compute a_k using Eq. (10) and compute b_k using Eq. (11);
6: Construct ϕ_{k+1} using Eq. (9);
7: **end for**
8: Construct a n by n matrix J using Eq. (21);
9: Compute the abscissae $\xi_j^{(i_j)}$ by solving the eigenvalues of J, and compute the corresponding weights $w_j^{(i_j)}$ using Eq. (22).

Algorithm 2 Compute gPCe coefficients with Stochastic collocation

1: **for** each dimension $j = 1, \cdots, N$ **do**
2: Compute $\xi_j^{(i_j)}$ and $w_j^{(i_j)}$ using Algorithm 1;
3: **end for**
4: Compute $\langle u(x, \vec{\xi}), \Psi_{\vec{\alpha}}(\vec{\xi}) \rangle_{p_{\vec{\xi}}}$ using Eq. (20) and compute $||\Psi_{\vec{\alpha}}(\vec{\xi})||^2$ using Eq. (27);
5: Compute gPCe coefficient $C_{\vec{\alpha}}$ using Eq. (16).

Since $\{\Psi_{\vec{\alpha}}\}$ is an orthogonal set for the measure $p_{\vec{\xi}}$, we have

$$C_{\vec{\alpha}}(x) = \frac{\langle u(x, \vec{\xi}), \Psi_{\vec{\alpha}}(\vec{\xi}) \rangle_{p_{\vec{\xi}}}}{\langle \Psi_{\vec{\alpha}}(\vec{\xi}), \Psi_{\vec{\alpha}}(\vec{\xi}) \rangle_{p_{\vec{\xi}}}} = \frac{\langle u(x, \vec{\xi}), \Psi_{\vec{\alpha}}(\vec{\xi}) \rangle_{p_{\vec{\xi}}}}{||\Psi_{\vec{\alpha}}(\vec{\xi})||^2} \tag{16}$$

The denominator is given in Eq. (27), and it can be obtained as a by-product of the three-term recurrence relation discussed in Section 3.1. The numerator of $C_{\vec{\alpha}}(x)$ can be approximated using some numerical quadrature rules. We first rewrite the numerator in Eq. (16) in the integral form:

$$\int_{\Xi} u(x, \vec{\xi}) \left[\frac{p_1(\xi_1) \cdots p_N(\xi_N)}{p_{\vec{\xi}}(\vec{\xi})} \right]^{1/2} \prod_{i=1}^{N} \phi_{\alpha_i}^{(i)}(\xi_i) p_{\vec{\xi}}(\vec{\xi}) d\vec{\xi} \tag{17}$$

and define

$$g(x, \vec{\xi}) = u(x, \vec{\xi}) \left[\frac{p_{\vec{\xi}}(\vec{\xi})}{p_1(\xi_1) \cdots p_N(\xi_N)} \right]^{1/2} \prod_{i=1}^{N} \phi_{\alpha_i}^{(i)}(\xi_i). \tag{18}$$

Then, Eq. (17) can be rewritten as

$$\int_{\Xi} g(x, \xi_1, \cdots, \xi_N) p_1(\xi_1) \cdots p_N(\xi_N) d\xi_1 \cdots d\xi_N \tag{19}$$

and it can be computed by a Gauss quadrature rule with the weighting functions being the tensor product of all marginal probability density functions. Specifically, $\langle u(x, \vec{\xi}), \Psi_{\vec{\alpha}}(\vec{\xi}) \rangle_{p_{\vec{\xi}}}$ is approximated by

$$\sum_{i_1=1}^{q_1} \sum_{i_2=1}^{q_2} \cdots \sum_{i_N=1}^{q_N} g(x, \xi_1^{(i_1)}, \cdots, \xi_N^{(i_N)}) w_1^{(i_1)} \cdots w_N^{(i_N)} \tag{20}$$

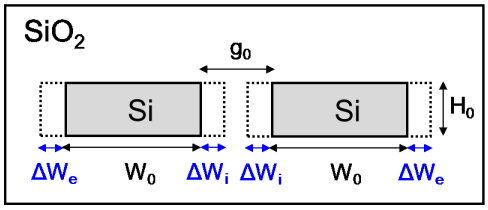


Figure 1. The cross section of a SOI-based directional coupler with nominal width W_0, nominal gap g_0, height H_0, and refractive indices $n_{Si} = 3.48, n_{SiO_2} = 1.445$.

where $\xi_j^{(i_j)}$'s and $w_j^{(i_j)}$'s are the abscissae and corresponding weights for the quadrature rule in the jth dimension (i.e., for parameter ξ_j), and q_j is the number of quadrature points used for dimension j. To obtain the quadrature points and weights for each marginal distribution p_i, we use Golub and Welsch algorithm [29], where $\{\xi_j^{(i_j)}\}$ are the eigenvalues of the symmetric tridiagonal matrix. According to the Golub and Welsch algorithm, $\{\xi_j^{(i_j)}\}_{i_j=1,\ldots,n}$ are obtained as the eigenvalues of the symmetric tridiagonal matrix

$$J_{mk} = \begin{cases} a_{m-1}, & \text{if} \quad k = m, \quad m = 1, \ldots, n \\ \sqrt{b_{m-1}}, & \text{if} \quad k = m-1, \quad m = 2, \ldots, n \\ \sqrt{b_m}, & \text{if} \quad k = m+1, \quad m = 1, \ldots, n-1 \end{cases} \tag{21}$$

and the corresponding weight is computed as

$$w_j^{(i_j)} = \mu_j \left(v_j^{(i_j,1)} \right)^2 \tag{22}$$

with $\mu_j = \int_{\mathbb{R}} p_j(\xi_j) d\xi_j$ and $v_j^{(i_j,1)}$ being the first component of the normalized eigenvector corresponding to eigenvalue $\xi_j^{(i_j)}$ [31].

4 Directional coupler example

4.1 Benchmark description

Due to the high refractive index contrast between silicon and silica, silicon photonic devices are extremely sensitive to variations in the geometry, for instance in waveguide width. Such geometrical variations can cause significant fluctuations in the effective phase index (n_{eff}) and can further lead to performance degradations in many photonic devices such as directional couplers and ring resonators.

In practical fabrications, due to tolerances of the fabrication process, the fabricated waveguide width (denoted as W) and the gap (denoted as g) between two waveguides of a directional coupler (as shown in Fig. 1) are different from chip to chip and wafer to wafer. Here, we assume that exterior and interior waveguide sidewall position, ΔW_e and ΔW_i, are correlated random parameters. The two coupled waveguides are assumed identical for simplicity and ΔW_i is assumed to be different from ΔW_e because of the typical proximity effects in the etching process. The fabricated W and g are related to the originally designed width (W_0) and (g_0) as

$$W = W_0 + \Delta W_e + \Delta W_i, \quad g = g_0 - 2\Delta W_i \tag{23}$$

Therefore, the total variation of width ΔW is $\Delta W_e + \Delta W_i$ and the total variation of gap Δg is $2\Delta W_i$. The quantity of interest for a directional coupler is its power coupling coefficient and output phase difference. The power coupling coefficient $K(z)$ is a sinusoidal function of the coupler's length z

$$K(z) = sin^2(\delta z) \tag{24}$$

where δ is the field coupling coefficient and it is equal to the effective phase index difference between the symmetric mode and the asymmetric mode of a coupler times a factor π/λ, where λ is the wavelength of concern. On the other hand, the output phase difference is always $\pi/2$ due to the assumed symmetry. Note that exact density functions for ΔW_e and ΔW_i can be extracted by measuring a large set of practical chips. However, in this paper, for the purpose of algorithm verification, we assume ΔW_e and ΔW_i are correlated, since there are already extensive techniques for the non-correlated process variations. Under the above assumption, our computational methods developed in Section 3 can be employed to analyze how such geometric uncertainties affect the power coupling coefficient. According to Eq. (24), it will be more convenient to approximate δ using gPC, since we can compute $K(z)$ for any length z with δ at hand. Specifically, we approximate $\delta(\lambda, \Delta W_e, \Delta W_i)$ by a linear combination of a set of orthogonal functions $\{\Psi_{\vec{\alpha}}(\Delta W_e, \Delta W_i)\}$, where λ is the wavelength. When computing $C_{\vec{\alpha}}(\lambda)$, we need to evaluate δ for each multidimensional quadrature point in Eq. (20). For each sample, a finite difference mode solver is called to solve for δ with the given value $(\Delta W_e, \Delta W_i, \lambda)$ using Eq. (23).

In our numerical experiments, we focus on synchronous couplers where the nominal width W_0 is 0.4 μm and the nominal gap g_0 is 0.2 μm. The waveguides height are fixed to 0.22 μm, the refractive index of silicon and silica are 3.48 and 1.445 respectively, and the wavelength of interest is 1.55 μm. The simulated δ with nominal width and gap is 0.11. All simulations are performed using a personal computer with a intel-xeon CPU X5650 and 24G of RAM.

4.2 Numerical results

In order to validate our computational methods, we use the following Gaussian mixture for simulation:

$$\begin{bmatrix} \Delta W_e \\ \Delta W_i \end{bmatrix} \sim 0.6 \cdot N(\vec{\mu}_A, \Sigma_A) + 0.4 \cdot N(\vec{\mu}_B, \Sigma_B)$$

with

$$\mu_A = \begin{bmatrix} 9 \\ 6 \end{bmatrix} \text{nm}, \quad \Sigma_\text{A} = \begin{bmatrix} 6 & 0 \\ 0 & 3 \end{bmatrix} \text{nm}^2$$

and

$$\mu_B = \begin{bmatrix} 8 \\ 7 \end{bmatrix} \text{nm}, \quad \Sigma_\text{B} = \begin{bmatrix} 5 & 1 \\ 1 & 4 \end{bmatrix} \text{nm}^2.$$

With a proper Gaussian mixture, we can approximate any measured distribution. For simplicity we have included only two terms in the Gaussian mixture here. The numbers in the mean and covariance matrix are picked to have the total variations of coupler's width (ΔW) and gap (Δg) range from around 0 to 25 nm, as shown in Fig. 2. The chosen numbers are reasonable numbers in current fabrication process tolerance. To show our algorithm's capability of dealing with correlated parameters, the weights a and b are chosen close to 0.5, and the exact values are 0.6 and 0.4 respectively. Note also that the choice of different distributions for ΔW_e and ΔW_i here allows taking into account the proximity effect that occurs in the gap during the fabrication.

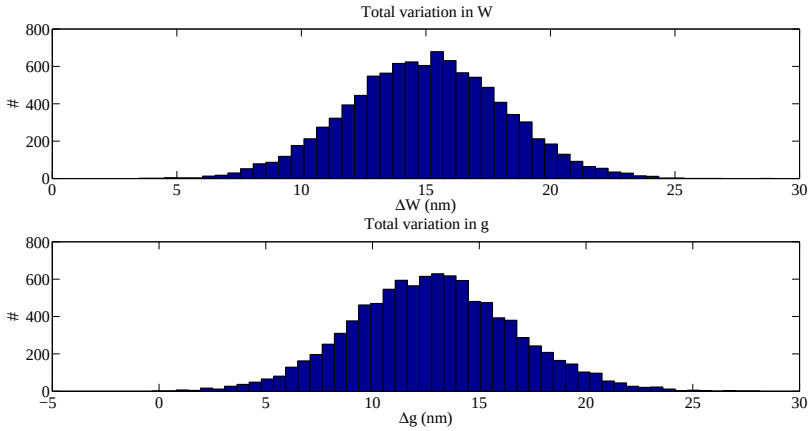


Figure 2. Top plot: Histograms of ΔW in the example. Bottom plot: Histogram of Δg in the example.

Our algorithm constructs an approximation of the parameter-dependent δ in the form of Eq. (8), and then computes $K(z)$ using Eq. (24). By doing so, we can easily evaluate its density function without calling the expensive mode solver every time in Monte-Carlo simulation. Figure 3 shows the probability density function of δ at $\lambda = 1.55$ μm. The solid line represents δ obtained from our stochastic collocation with level $\vec{m} = (5, 5)$. Here $\vec{m} = (5, 5)$ means that a 5-level Gauss-Hermite quadrature rule is used for both parameter dimensions to evaluate Eq. (20). For parameter ξ_i, the level m_i is determined as

$$q_i = 2m_i - 1 \tag{25}$$

where q_i is the number of quadrature points used for ξ_i. Therefore, 9 quadrature points are used in each dimension to obtain the results in Fig. 3, i.e. $q_1 = q_2 = 9$ in Eq. (20). Note that to avoid aliasing error in the gPC coefficient, the relation between $\vec{m}$ and $\vec{t}$ in Eq. (8) is $\vec{t} = 2\vec{m} - 2$. In order to verify the accuracy of our approach, the simulated probability density function from standard Monte Carlo with 10^4 samples is also plotted as a dashed line in Fig. 3. The result from our stochastic collocation is consistent with that from Monte Carlo method. The nominal value of δ for $W_0 = 400$ nm and $g_0 = 200$ nm is 0.11, while the analysis shows an expected value asymmetrically distributed around $\delta = 0.102$. With the pdf of δ at hand, we can compute $K(z)$ using Eq. (24), and the result is shown in Fig. 4. Figures 4(a) and 4(b) show the simulated probability density functions of the power coupling coefficient at $z = 7.1376\mu$m and $z = 2.9250\mu$m, which are the lengths that correspond to $K_0 = 0.5$ and $K_0 = 0.1$ respectively. K_0 is the nominal power coupling coefficient for the nominal width $W_0 = 400$ nm and gap $g_0 = 200$ nm. From Fig. 4, it is clearly seen that K deviates from its original desired values under process variations. Notice that with the choice of the parameters in our example ($W_0 = 400$ nm and $g_0 = 200$ nm), the coupler coefficients are on average smaller than designed ($K_0 = 0.5$, and $K_0 = 0.1$). This can be explained by looking at the average waveguide width and gap in our example: the waveguide width are always larger than the nominal values (around 15 nm in average) while the gap is always smaller (around 12.5 nm in average). For example, the computed K($z = 7.1376\mu$ m) with $W = 415$ nm and $g = 187.5$ nm is around 0.44, which is smaller than the designed value 0.5 and complies the simulated results in Fig. 4(a). Similar results can be obtained for the case of K($z = 2.9250\mu$ m) with $W = 415$ nm and $g = 187.5$ nm.

In order to show the convergence of our algorithm, we gradually increase the level numbers and compare the truncation error of $\delta(\lambda, \Delta W_e, \Delta W_i)$. We use the 10^4 samples Monte Carlo simulation

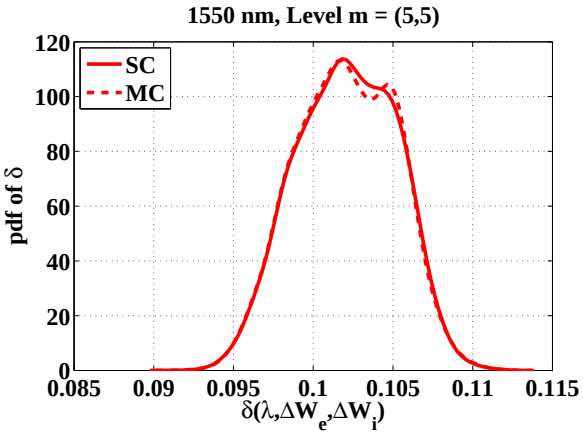
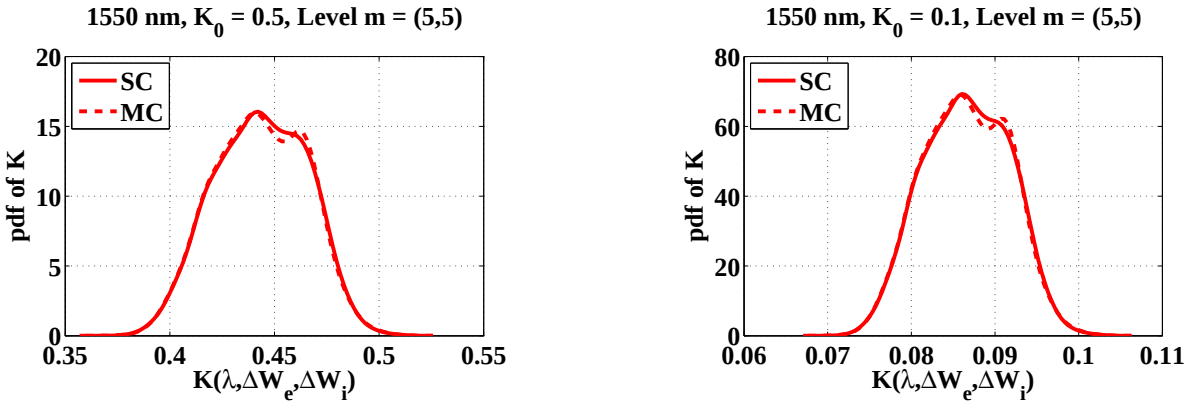


Figure 3. The simulated PDF of $\delta(\lambda, \Delta W_e, \Delta W_i)$ with $\lambda = 1.55$ μm. The solid line is the Stochastic Collocation (SC) result, whereas the dash line represents Monte Carlo (MC) result. The nominal value of δ is 0.11.



(a) $z = 7.1376$ μm (b) $z = 2.9250$ μm

Figure 4. The simulated pdf of power coupling coefficient $K(z)$. The solid line is the Stochastic Collocation (SC) result, whereas the dash line represents Monte Carlo (MC) result.

result as the reference solution and plot the L_2-norm of the error as a function of the level parameter. The result in Fig. 5 shows that the error decreases when more basis functions and more Gauss quadrature points are used. In addition, to quantitatively show the accuracy of our solver, Table 1 compares the mean value and standard deviation of the simulated δ from Monte Carlo and from our technique. It clearly shows that Monte Carlo result converges towards that from our proposed solver as the number of samples increases. Moreover, to achieve a similar level of accuracy in the mean and standard deviation of δ, Monte Carlo uses 10^4 samples and 97 minutes of CPU time, whereas our method consumes about 1 minute and 45 seconds when 9 quadrature points are used in each dimension. This clearly shows the superior efficiency of our technique over Monte Carlo simulation, and the speedup factor is about 55.

Remark: In our implementations, we have observed slower convergence rates for some cases even when the δ is smoothly dependent on the process variations parameters ΔW_e and ΔW_i. The slow convergence rate is caused by the theoretical limitations of the basis functions proposed in [26]. Observing Eq. (4), we note that the basis functions $\Psi_{\vec{\alpha}}(\vec{\xi})$ might be highly non-linear when the

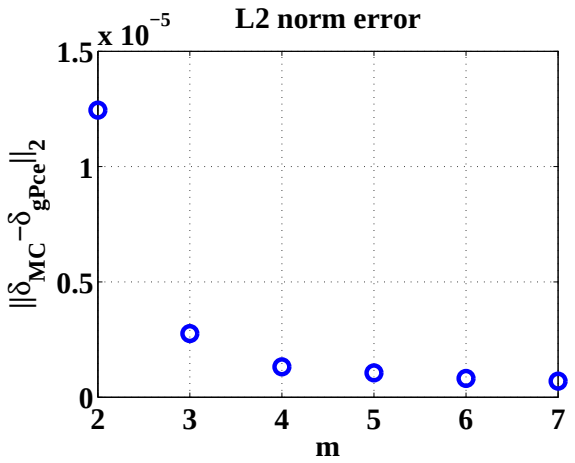


Figure 5. The truncation error in terms of the level m.

Table 1: Performance summary of Stochastic Collocation and Monte Carlo simulation

	Monte Carlo			**Stochastic Collocation**
Number of Samples	100	1000	10000	81 Quadrature points
Mean value	0.102156	0.102074	0.102055	0.102083
S.t.d. value	0.003314	0.003152	0.003097	0.003096
CPU time (sec.)	58	580	5800	105

parameters are correlated. Thus, for some cases many basis functions are needed to approximate a smooth quantity of interest (which is δ in our example), resulting in a slower convergence rate for the solution. In order to improve the convergence rate, we are working on a whole new set of basis functions which can potentially achieve high accuracy with much fewer basis functions.

5 Conclusions

In this paper, we have developed a stochastic collocation scheme to simulate silicon photonic devices with non-Gaussian correlated parameters described by a Gaussian-mixture joint probability density function. The numerical techniques for constructing a set of orthogonal basis functions and for implementing stochastic collocation are presented. The proposed numerical method is applied to analyze the process variations in a directional coupler. The simulation results of the δ and K are consistent with Monte Carlo. Our technique has achieved a 55× speedup factor over standard Monte Carlo.

Appendix

The orthogonality of $\{\Psi_{\vec{\alpha}}(\vec{\xi})\}$ can be shown by computing the inner product of $\Psi_{\vec{\alpha}}(\vec{\xi})$ and $\Psi_{\vec{\gamma}}(\vec{\xi})$:

$$\langle \Psi_{\vec{\alpha}}, \Psi_{\vec{\gamma}} \rangle_{p_{\vec{\xi}}} = \int_{\Xi} \Psi_{\vec{\alpha}}(\vec{\xi}) \Psi_{\vec{\gamma}}(\vec{\xi}) p_{\vec{\xi}}(\vec{\xi}) d\vec{\xi}. \tag{26}$$

Substituting Eq. (4) into Eq. (26), one can obtain

$$\begin{aligned}\left\langle \Psi_{\vec{\alpha}}(\vec{\xi}), \Psi_{\vec{\gamma}}(\vec{\xi}) \right\rangle_{p_{\vec{\xi}}} &= \prod_{i=1}^{N} \int_{\mathbb{R}} \phi_{\alpha_i}^{(i)}(\xi_i) \phi_{\gamma_i}^{(i)}(\xi_i) p_i(\xi_i) d\xi_i \\ &= \prod_{i=1}^{N} \left\| \phi_{\alpha_i}^{(i)}(\xi_i) \right\|^2 \delta_{\alpha_i \gamma_i} \\ &= \left\| \Psi_{\vec{\alpha}}(\vec{\xi}) \right\|^2 \delta_{\vec{\alpha}\vec{\gamma}}.\end{aligned} \tag{27}$$

Here, $\delta_{\vec{\alpha}\vec{\gamma}}$ is a Delta function defined as below:

$$\delta_{\vec{\alpha}\vec{\gamma}} = \begin{cases} 1, \text{ if } \vec{\alpha} = \vec{\gamma}, \\ 0, \text{ otherwise.} \end{cases} \tag{28}$$

Acknowledgments

T.-W. Weng and Z. Zhang contributed equally to this work. This work was supported by the NSF/SRC NEEDS program and by the MIT-SkolTech program. The authors would like to thank Dr. Daniele Melati for all valuable suggestions and help.

References

[1] A. Biberman and K. Bergman, "Optical interconnection networks for high-performance computing systems," *Reports on Progress in Physics*, vol. 75, no. 4, p. 046402, 2012.

[2] J. Sun, E. Timurdogan, A. Yaacobi, E. S. Hosseini, and M. R. Watts, "Large-scale nanophotonic phased array," *Nature*, vol. 493, no. 7431, pp. 195–199, 2013.

[3] W. A. Zortman, D. C. Trotter, and M. R. Watts, "Silicon photonics manufacturing," *Opt. Express*, vol. 18, no. 23, pp. 23 598–23 607, Nov 2010.

[4] A. Krishnamoorthy, X. Zheng, G. Li, J. Yao, T. Pinguet, A. Mekis, H. Thacker, I. Shubin, Y. Luo, K. Raj, and J. Cunningham, "Exploiting CMOS manufacturing to reduce tuning requirements for resonant optical devices," *IEEE Photon. J.*, vol. 3, no. 3, pp. 567–579, June 2011.

[5] X. Chen, M. Mohamed, Z. Li, L. Shang, and A. R. Mickelson, "Process variation in silicon photonic devices," *Appl. Opt.*, vol. 52, no. 31, pp. 7638–7647, Nov 2013.

[6] S. Selvaraja, E. Rosseel, L. Fernandez, M. Tabat, W. Bogaerts, J. Hautala, and P. Absil, "SOI thickness uniformity improvement using corrective etching for silicon nano-photonic device," in *2011 8th IEEE International Conference on Group IV Photonics (GFP)*, Sept 2011, pp. 71–73.

[7] X. Wang, W. Shi, H. Yun, S. Grist, N. A. Jaeger, and L. Chrostowski, "Narrow-band waveguide bragg gratings on SOI wafers with CMOS-compatible fabrication process," *Opt. Express*, vol. 20, no. 14, pp. 15 547–15 558, 2012.

[8] L. Chrostowski, X. Wang, J. Flueckiger, Y. Wu, Y. Wang, and S. Talebi Fard, "Impact of fabrication non-uniformity on chip-scale silicon photonic integrated circuits," in *Optical Fiber Communication Conference*. OSA, 2014, pp. Th2A–37.

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

[10] Z. Zhang, T. El-Moselhy, P. Maffezzoni, I. Elfadel, and L. Daniel, "Efficient uncertainty quantification for the periodic steady state of forced and autonomous circuits," *IEEE Trans. Circuits Syst. II: Exp. Briefs*, vol. 60, no. 10, pp. 687–691, Oct 2013.

[11] Z. Zhang, I. Elfadel, and L. Daniel, "Uncertainty quantification for integrated circuits: Stochastic spectral methods," in *IEEE/ACM International Conference on Computer-Aided Design*, Nov 2013, pp. 803–810.

[12] K. Strunz and Q. Su, "Stochastic formulation of spice-type electronic circuit simulation with polynomial chaos," *ACM Trans. Model Comput. Simul.*, vol. 18, no. 4, p. 15, 2008.

[13] N. Agarwal and N. Aluru, "Stochastic analysis of electrostatic MEMS subjected to parameter variations," *J. Microelectromech. Syst.*, vol. 18, no. 6, pp. 1454–1468, Dec 2009.

[14] Z. Zhang, X. Yang, G. Marucci, P. Maffezzoni, I. M. Elfadel, G. Karnidadakis, and L. Danie, "Stochastic testing simulator for integrated circuits and MEMS: Hierarchical and sparse techniques," in *IEEE Custom Integrated Circuits Conference*, Sept 2014.

[15] S. Weinzierl, "Introduction to Monte Carlo methods," *arXiv preprint hep-ph/0006269*, 2000.

[16] A. Singhee and R. A. Rutenbar, "Why quasi-Monte Carlo is better than Monte Carlo or latin hypercube sampling for statistical circuit analysis," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 29, no. 11, pp. 1763–1776, 2010.

[17] D. Xiu and G. E. Karniadakis, "The Wiener–Askey polynomial chaos for stochastic differential equations," *SIAM J. Sci. Comput.*, vol. 24, no. 2, pp. 619–644, 2002.

[18] D. Cassano, F. Morichetti, and A. Melloni, "Statistical analysis of photonic integrated circuits via polynomial-chaos expansion," in *Signal Processing in Photonic Communications*. OSA, 2013, pp. JT3A–8.

[19] R. G. Ghanem and P. D. Spanos, *Stochastic Finite Elements: A Spectral Approach*. Springer, 1991, vol. 41.

[20] D. Xiu and J. S. Hesthaven, "High-order collocation methods for differential equations with random inputs," *SIAM J. Sci. Comput.*, vol. 27, no. 3, pp. 1118–1139, 2005.

[21] F. Nobile, R. Tempone, and C. G. Webster, "A sparse grid stochastic collocation method for partial differential equations with random input data," *SIAM J. Numer. Anal.*, vol. 46, no. 5, pp. 2309–2345, 2008.

[22] T. El-Moselhy and L. Daniel, "Stochastic integral equation solver for efficient variation-aware interconnect extraction," in *Proceedings of Design Automation Conference*. ACM, 2008, pp. 415–420.

[23] ——, "Variation-aware stochastic extraction with large parameter dimensionality: Review and comparison of state of the art intrusive and non-intrusive techniques," in *International Symposium on Quality Electronic Design (ISQED)*. IEEE, 2011, pp. 1–10.

[24] D. Vande Ginste, D. De Zutter, D. Deschrijver, T. Dhaene, P. Manfredi, and F. Canavero, "Stochastic modeling-based variability analysis of on-chip interconnects," *IEEE Trans. Compon. Packag. Manuf. Technol.*, vol. 2, no. 7, pp. 1182–1192, 2012.

[25] I. S. Stievano, P. Manfredi, and F. G. Canavero, "Carbon nanotube interconnects: Process variation via polynomial chaos," *IEEE Trans. Electromagn. Compat.*, vol. 54, no. 1, pp. 140–148, 2012.

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

[27] T.-W. Weng, Z. Zhang, Z. Su, and L. Daniel, "Fast stochastic simulation of silicon waveguide with non-gaussian correlated process variations," in *Asia Communications and Photonics Conference*. OSA, 2014, p. AF3B.7.

[28] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the em algorithm," *J. R. Stat. Soc. Series B*, pp. 1–38, 1977.

[29] G. H. Golub and J. H. Welsch, "Calculation of Gauss quadrature rules," *Math. Comput.*, vol. 23, no. 106, pp. 221–230, 1969.

[30] Q. Liu and D. A. Pierce, "A note on gausshermite quadrature," *Biometrika*, vol. 81, no. 3, pp. 624–629, 1994.

[31] H. Wilf, *Mathematics for the Physical Sciences*. Wiley, 1962, no. Problem 9.