

Characterization and Applications of Spatial Variation Models for Silicon Microring-Based Optical Transceivers

Yuyang Wang^{*†}, Jared Hulme[‡], Peng Sun[‡], Mudit Jain[‡], M. Ashkan Seyed[‡],
Marco Fiorentino[‡], Raymond G. Beausoleil[‡], and Kwang-Ting Cheng[§]

[†]*Department of Electrical and Computer Engineering, University of California, Santa Barbara, California, U.S.A.*

[‡]*Hewlett Packard Labs, Hewlett Packard Enterprise, Palo Alto, California, U.S.A.*

[§]*School of Engineering, Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong*

^{*}wyw@ece.ucsb.edu

Abstract—Photonic integrated circuits suffer from large process variations. Effective and accurate characterization of the variation patterns is a critical task for enabling the development of novel techniques to alleviate the variation challenges. In this study, we propose a hierarchical approach that effectively decomposes the spatial variations of silicon microring-based optical transceivers into wafer-level, intra-die, and inter-die components. We then demonstrate that the characterized variation models can be used to generate trustworthy synthetic data for architecture- and system-level solutions for variation alleviation. We further demonstrate the utility of our variation characterization method for accurate yield prediction based on partial measurement data.

I. INTRODUCTION

Integrated silicon photonics has emerged as a promising alternative to traditional CMOS electronics [1], [2]. To substantially improve the bandwidth capacity and energy efficiency of future high-performance computing systems (HPC), optical interconnects have been proposed to host the increasingly intensive communication traffic [3], [4]. However, due to the intrinsically high sensitivity of silicon photonic devices to fabrication imperfection, optical interconnects have been suffering from large process variations and relying on post-fabrication compensation techniques to alleviate the variation challenges [5]–[11]. Therefore, effectively and accurately characterizing the variation patterns is a critical task for enabling the development of architectural and system solutions for variation alleviation [12].

Process variations of semiconductor manufacturing can be examined hierarchically. Spatial variations often exist in forms of wafer-level, intra-die, and inter-die patterns, while temporal variations may present at wafer-to-wafer and lot-to-lot levels [13]. Limited by the fabrication quantity, previous studies on variation characterization of silicon photonic devices have mainly focused on sub-wafer levels. The authors of [14]–[16] extracted wafer-level variation patterns by modeling the variations as a function of the device location on the wafer. However, such methods only capture a smooth trend of the variations across the wafer and leave high-frequency components in the residuals which may not be entirely random. The authors of [17]–[19] observed that the variation between two devices on the same wafer is largely proportional to their physical distance. Accordingly, the authors of [9], [20] proposed to characterize the intra-die and inter-die variations separately. However, the limited number of dies and devices per die in these studies prohibited further extraction of location dependencies of the variations. Instead, independent Gaussian distributions were assumed for both intra-die and inter-die variations [9], which led to oversimplified variation models. In [21], [22], variation characterization methods encompassing both wafer-level and intra-die components were presented, but also limited by the small number of fabricated devices.

On the other hand, variation characterization of electronic manufacturing has been extensively studied by leveraging abundant fabrication and measurement data. Spatial variation models can be learned

using decomposition-based methods at various levels [23]–[29] or estimation-based methods characterizing the entire wafer [30]–[35]. However, such techniques usually assume a large number of devices per wafer. As optical devices tend to be bulkier and thus less can be fabricated on a single wafer, such techniques are usually less effective if directly applied to silicon photonic wafers.

In this study, we aim to develop an effective method for accurately characterizing the spatial variations of silicon microring-based optical transceivers, and to explore its applications to architecture- and system-level design optimization. We propose to take a hierarchical approach incorporating domain-specific knowledge, spatial-frequency analyses, and low-rank tensor factorization methods to decompose the variations of optical test items into wafer-level, intra-die, and inter-die components. We also provide analyses on the residuals from each step to evaluate their randomness. Wafer-scale measurement data of 66 transceivers containing 3000+ microrings were used to validate our variation characterization method.

We then demonstrate that the characterized variation models can be used to generate trustworthy synthetic data of microring-based optical transceivers for the development of novel variation alleviation techniques. The synthetic data generated by our approach closely resemble the actual data and can estimate the system performance with substantially higher accuracy than existing methods. We further explore the utility of our variation characterization method for yield prediction, and demonstrate that a reasonably accurate prediction of the transceiver yield can be achieved with only 60% of measurement data of a wafer.

The rest of the paper will be organized as follows. In Section II, we provide an overview of the fabricated microring-based optical transceivers, and describe the preprocessing steps to extract the variations of individual test items from raw measurement. In Section III, we provide implementation details of our variation characterization method and residual analyses. In Section IV, we demonstrate the applications of our variation characterization method for trustworthy synthetic data generation and accurate yield prediction. And finally, in Section V, we discuss the conclusion and future work.

II. OVERVIEW AND DATA PREPROCESSING

A. Background of Microring-Based Optical Transceivers

A silicon microring resonator (Fig. 1a) is a highly wavelength-selective device that can modulate or filter optical signals at its specific resonance wavelength [36]. The transmission spectrum of a microring resonator can be modeled as a Lorentzian function parameterized by three key properties, namely λ_r , ER , and Q :

$$T(\lambda) = 1 - \frac{1 - 1/ER}{1 + (2Q \cdot (\lambda - \lambda_r)/\lambda_r)^2} \quad (1)$$

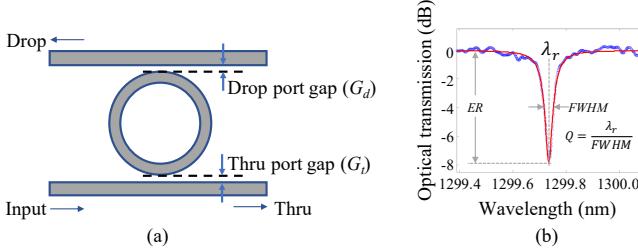


Fig. 1. Illustrations of typical (a) geometry design, and (b) thru-port transmission spectrum of a microring resonator, where $FWHM$: full width at half maximum.

where λ_r , ER , and Q are the resonance wavelength, extinction ratio, and quality factor of the spectrum (Fig. 1b). A microring-based optical transceiver (TRx) achieves dense wavelength-division multiplexing (DWDM) by deploying cascaded microring resonators along a common waveguide [37]–[39]. At the transmitter (Tx) side, voltages applied to the microring modulators can slightly shift their resonance wavelengths to perform on-off keying (OOK) modulation of the optical signal. At the receiver (Rx) side, corresponding microring filters can couple the signal out for detection. The overall transmission spectrum of a Tx/Rx can be modeled as the product of all transmission functions of individual microrings:

$$T_n(\lambda) = \prod_{i=1}^n \left(1 - \frac{1 - 1/ER_i}{1 + (2Q_i \cdot (\lambda - \lambda_{r,i})/\lambda_{r,i})^2} \right) \quad (2)$$

where n is the number of DWDM channels, which is also the number of microrings in the Tx/Rx.

Variations of λ_r , ER , and Q have a significant impact on the energy efficiency of the transceiver. First of all, the resonance wavelengths of the Tx/Rx channels must be actively tuned to align with the carrier wavelengths of the laser source, and this tuning power is non-trivial [40]. Besides, variations of both ER and Q affect the loss and crosstalk noise of the optical link, and thus the optical power required to attain a specific data rate will vary. Therefore, we will focus on these three key parameters in this study for variation characterization.

B. Overview of Fabricated Devices

Comprehensive measurement data of the transmission spectra of 24-channel microring-based transceivers were collected from a fabricated wafer containing 66 of them (Fig. 2a). The fabrication was carried out by a commercial foundry (STMicroelectronics) on a 300 mm silicon-on-insulator (SOI) wafer. Each die consists of a transceiver block (marked in Fig. 2b). In each transceiver block (Fig. 2c), a 24-channel Tx and a 24-channel Rx were measured

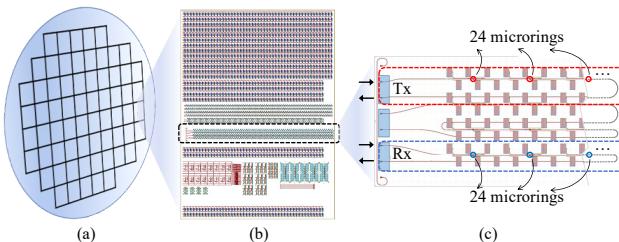


Fig. 2. Organization of fabricated devices: (a) a wafer of 66 dies; (b) each die consisting of one TRx; and (c) each Tx/Rx consisting of 24 microrings.

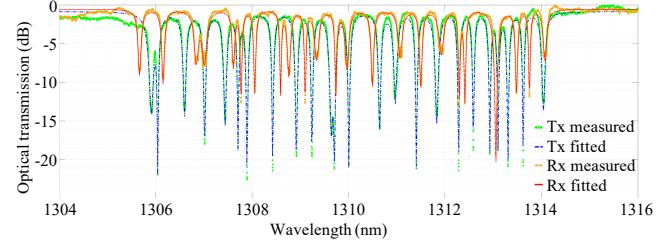


Fig. 3. Measured and fitted transmission spectra of a 24-channel transceiver.

for optical transmission spectra. Thus, the total number of microrings involved is $66 \times 2 \times 24 = 3168$. Each group of 24 microrings starts with a $5\text{ }\mu\text{m}$ radius and ramps up to $5.046\text{ }\mu\text{m}$ radius with a step size of 2 nm. All microrings share the same drop-port gap width (G_d) of 200 nm. The thru-port gap widths (G_t) differ for Tx and Rx microrings and are 175 nm and 200 nm, respectively.

C. Model-Based Data Preprocessing

We preprocessed the measured transmission spectra to extract the values and variations of λ_r , ER , and Q for each microring.

1) *Extraction of λ_r , ER , and Q :* The resonance wavelengths were extracted by detecting significant peaks of the measured spectra. In practice, a resonance peak may split into two adjacent ones due to back-reflection in the microring [41]. Techniques suggested in [42] were employed to effectively recognize the peak splitting as a single resonance. Then, ER and Q were extracted for all microrings by fitting Eq. 2 to the measured spectra via non-linear least squares (an example shown in Fig. 3). We plot the extracted values of λ_r , ER , and Q in three wafer maps (Figs. 4a-c). Each wafer map is 16×240 , in which each 2×24 block corresponds to a die. Odd rows are for Tx and even rows are for Rx. The ranges of the color bars were obscured to protect the manufacturer. Note that some microrings were found without significant resonance and represented by blank pixels in the wafer maps.

2) *Computation of nominal values:* The variations of λ_r , ER , and Q can be computed as the actual values subtracted by the nominal values. However, the nominal values are related to the microring geometry, and thus are different for microrings with various radii and gap widths. Figs. 4a-c clearly depicts the incremental resonance wavelengths within each transceiver, and the differences between Tx/Rx rows for ER and Q .

Models for microring resonators [43] specify that:

$$\lambda_r = \frac{2\pi n_{\text{eff}}}{m} \cdot r = c_0 \cdot r \quad (3)$$

$$ER = \left(\frac{\delta_t + \delta_d}{\delta_t - \delta_d} \right)^2 \quad (4)$$

$$Q = \frac{2\pi\lambda_r}{FSR \cdot (\delta_t + \delta_d)} \quad (5)$$

where $FSR = 13.09\text{ nm}$ is the average free spectrum range for these fabricated microrings, and

$$\delta_t = c_1 \cdot \exp(-c_2 \cdot G_t) + c_3 \quad (6)$$

$$\delta_d = c_4 \cdot \exp(-c_5 \cdot G_d) + c_6 \quad (7)$$

are the coupling ratios of the thru port and the drop port, respectively. We define an auxiliary function:

$$\mathbf{y} = \begin{bmatrix} \lambda_r^* - c_0 \cdot r \\ ER^* - \left(\frac{\delta_t + \delta_d}{\delta_t - \delta_d} \right)^2 \\ Q^* - \frac{2\pi\lambda_r^*}{FSR \cdot (\delta_t + \delta_d)} \end{bmatrix}, \text{ where } \begin{cases} \delta_t = c_1 \cdot \exp(-c_2 \cdot G_t) + c_3 \\ \delta_d = c_4 \cdot \exp(-c_5 \cdot G_d) + c_6 \end{cases} \quad (8)$$

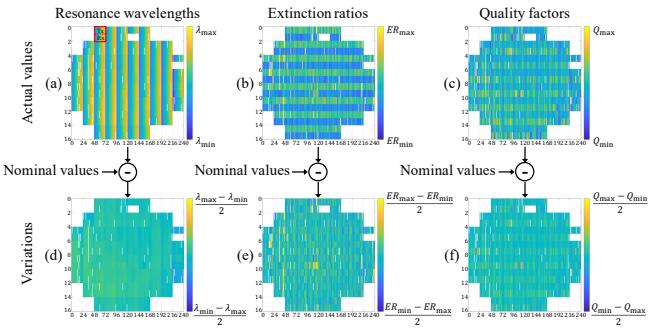


Fig. 4. Wafer maps for the (a-c) actual values, and (d-f) variations of λ_r , ER , and Q .

Here, λ_r^* , ER^* , and Q^* are the actual values extracted from the measured spectra, and r , G_t , and G_d are the design values of the microring radii and thru/drop-port gap widths. The coefficients $c_0 \sim c_6$ can then be solved as $\arg\min \|y\|_2^2$ using existing non-linear minimization algorithms [44]. The nominal values of λ_r , ER , and Q for each microring can then be computed based on Eqs. 3-7.

Figs. 4d-f plots the variations of λ_r , ER , and Q in three wafer maps after subtracting the nominal values from the actual values. The color bars are also shifted to have zero mean. In the next section, we describe our hierarchical characterization method for the spatial variations of λ_r , ER , and Q .

III. VARIATION CHARACTERIZATION

We present our method for effectively decomposing the spatial variations of microring-based optical transceivers into wafer-level, intra-die, and inter-die components. As categorized in [25], wafer-level variation components capture a low-frequency trend of the variations across the wafer. They are usually assumed to be related to the fabrication process and independent of the layout. Intra-die components capture the device variations within a die and manifest themselves as repetitive patterns across the wafer. Inter-die components further capture the die-to-die differences that still appear more systematic than random.

A. Wafer-level variation components

The authors of [14] decomposed the spatial variations of optical devices into leveling and radial components. The two basis functions are written as:

$$f_{\text{leveling}}(x, y) = a_1 x + a_2 y \quad (9)$$

$$f_{\text{radial}}(x, y) = a_3 \cdot \cos(a_4 \sqrt{x^2 + y^2} + a_5) \quad (10)$$

where x and y are the coordinates of the device on the wafer. We further introduced two basis functions, namely wafer-edge and wafer-center, to capture the different behavior of dies located near the edge/center of the wafer [28]:

$$f_{\text{edge}}(x, y) = \begin{cases} a_6, & (x, y) \in E \\ 0, & \text{otherwise} \end{cases} \quad (11)$$

$$f_{\text{center}}(x, y) = \begin{cases} a_7, & (x, y) \in C \\ 0, & \text{otherwise} \end{cases} \quad (12)$$

where E and C are predefined sets of edge dies and center dies, respectively. The coefficients were derived by robust regression [45] that fits the sum of the basis functions to the spatial variations. It is noteworthy that some alternating patterns between the Tx and Rx rows can be observed in the variation wafer maps for ER and Q (Figs. 4e&f), and the authors of [29] suggested extra basis functions dedicated to those. However, we assume that the alternating

patterns largely relate to the different thru-port gap widths (G_t) of Tx/Rx microrings, but that the wafer-level variation components are layout independent. Later we show that these alternating patterns can be effectively extracted as part of the intra-die variation components.

In our analysis, we first removed some outliers from the wafer maps by identifying those more than 1.5 interquartile ranges [46] above the 75% quartile or below the 25% quartile. This method works generally well when the data are not normally distributed. We then scaled the variations into the range of $[-0.5, 0.5]$ before decomposition. Fig. 5 shows the scaled variations for λ_r with the outliers removed and the wafer-level decomposition results. For simplicity, the three components corresponding to Eqs. 10-12 are merged into a single plot for the radial component. The significance of each component was computed as the variance of the component divided by that of the total variations. It can be observed that over 40% of the variations were left in the residuals where systematic patterns still exist, despite the extracted wafer-level components.

We provide two analyses of the wafer-level residuals for λ_r . First, we compared the distribution of the residuals to a standard normal distribution using a quantile-quantile plot (QQ-plot). As shown in Fig. 6a, the quantiles of the wafer-level residuals fall below a normal distribution at both high and low ends. This can also be verified by the histogram of the residuals which has a heavier tail to the left (Fig. 6b). We further investigated the autocorrelation of each row/column of the wafer-level residuals. Fig. 6c plots the autocorrelation of a center row and the 95%-confidence intervals for a white noise process. Major peaks can be observed at every period of 24. For columns, a period of 2 was also observed. These findings indicate that a repetitive pattern with the size of a die (2×24) exists across the wafer.

Wafer-level decomposition was conducted for variations of ER and Q as well, and the existence of repetitive variation patterns was also confirmed among the wafer-level residuals. We summarize the wafer-level decomposition results for λ_r , ER , and Q in Table I for later revisit. Next, we demonstrate the decomposition of intra-die variation components.

B. Intra-die variation components

In this study, we employed a spatial-frequency-domain analysis proposed in [23] to extract the intra-die variation components. In contrast to [21] which simply takes an average of all dies as the intra-die variation pattern, this approach leverages the periodicity observed

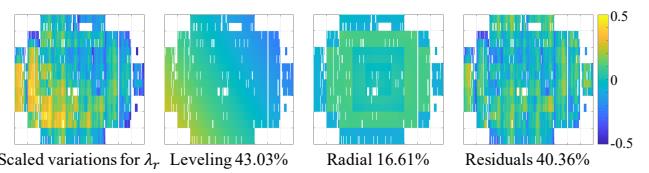


Fig. 5. Wafer-level variation components for λ_r .

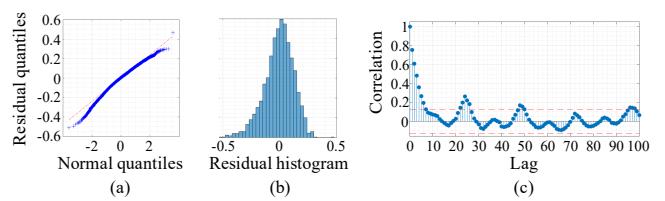


Fig. 6. Analyses of wafer-level residuals for λ_r .

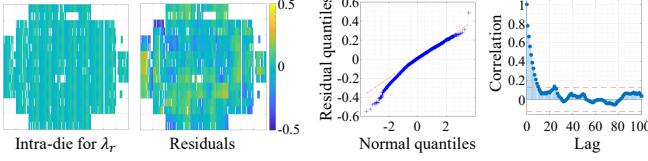


Fig. 7. Intra-die variation patterns and residual analyses for λ_r .

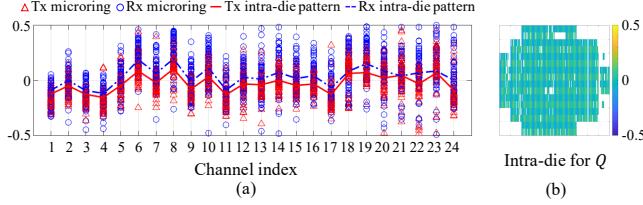


Fig. 8. Intra-die variation patterns for Q .

from the wafer-level residuals. Specifically, we converted the wafer-level residuals into the frequency domain using a 2-dimensional fast Fourier transform (FFT). The transformed matrix has the same size as the input matrix, i.e., 16×240 . Theoretically, a periodic pattern in the spatial domain with the size of 2×24 will correspond to frequency-domain elements at every $16/2$ and $240/24$ intervals in the y- and x-directions. Therefore, the intra-die variation pattern can be extracted by first downsampling the frequency-domain matrix at these intervals, followed by an inverse FFT.

In Fig. 7, we show the intra-die variation patterns extracted for λ_r and the analyses of the residuals (hereinafter the *intra-die residuals*). Visually, the high end of the QQ-plot is closer to a standard normal distribution than that of the wafer-level residuals. The peaks in the autocorrelation plot are also reduced and mostly fall within the 95%-confidence intervals for a white noise process.

In Fig. 8a, we plot the wafer-level residuals of Q for all microrings vs. their locations within a die. The two rows in a die correspond to a Tx and an Rx. We also plot the extracted intra-die variation patterns in thick lines. We observe that the extracted patterns closely follow the general trends of the data groups. The differences between the Tx and Rx rows were also captured by the intra-die variation components (Fig. 8b), accounting for the alternating patterns mentioned in Section III-A. The significance of the intra-die variation components to the total variations are also summarized in Table I for λ_r , ER , and Q .

C. Inter-die variation components

Discontinuities at die borders can still be observed from the intra-die residuals, which is more likely due to systematic than random cause. We propose to further extract inter-die variation components from the intra-die residuals. Previous methods such as the one proposed in [25] are based on the availability of multiple wafers. In this study, however, we present a method based on low-rank tensor factorization for inter-die variation characterization.

Tensor factorization is a generalization of matrix factorization, which decomposes a tensor into another tensor of a smaller size and several factor matrices [47]. Existing applications of tensor factorization, however, have been focusing on data compression or accurate reconstruction of data with missing entries. Algorithms developed for these purposes usually leverage the low-rank properties of the data [48]–[51], but cannot guarantee their effectiveness if the data are inherently not low-rank. To extract the inter-die variation

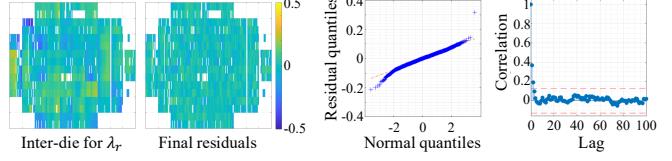


Fig. 9. Inter-die variation characterization and final residual analyses for λ_r .

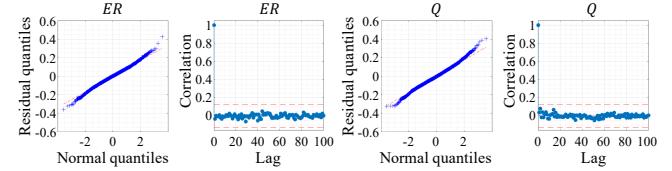


Fig. 10. Final residual analyses for ER and Q .

patterns, however, we aim to learn the hidden low-rank factors of the data disregarding the noise. In this context, we force a very low maximum rank to be explored by the algorithm in order not to capture the noise into the systematic variation components.

We first reshape the intra-die residuals into a tensor by splitting the wafer map at die borders and stacking all dies in a third dimension. Then, we apply tensor factorization to approximate the tensor with a low-rank version. This low-rank tensor is reshaped back to a matrix as the inter-die variation pattern, and the approximation errors are the final residuals of our whole decomposition workflow. In practice, we found that several existing tensor factorization algorithms all work considerably well with a maximum rank as low as 5. Fig. 9 shows the inter-die variation component for λ_r extracted by smooth PARAFAC decomposition [50] and the final residual analyses. The linearity of the QQ-plot indicates the resemblance of the residuals to a standard normal distribution within the range of ± 2 standard deviations. The flat autocorrelation curve also confirms the successful extraction of periodic variation components. The residual analyses for ER and Q (Fig. 10) reach the same conclusion.

In Table I, we summarize the hierarchical variation components extracted for λ_r , ER , and Q . We observed that the resonance wavelengths of the microrings are more prone to wafer-level variations, while ER and Q present more significant intra-die (layout-dependent) patterns. Considerable amounts of inter-die variations exist for all three test items. Overall, the variations of ER and Q appear more random than that of λ_r .

IV. APPLICATIONS OF THE VARIATION MODELS

In addition to the implications of variation sources, we further explore the applications of our characterized variation models for trustworthy synthetic data generation and accurate yield prediction for microring-based optical transceivers.

TABLE I
SUMMARY OF SPATIAL VARIATION DECOMPOSITION.

	Wafer-level		Intra-die	Inter-die	Residual	Residual distribution
	Leveling	Radial				
λ_r	43.03%	16.61%	5.12%	27.45%	7.79%	$\mathcal{N}(0, 0.0407)$
ER	0.04%	1.60%	32.67%	18.61%	47.08%	$\mathcal{N}(0, 0.0891)$
Q	1.59%	2.31%	29.47%	21.72%	44.91%	$\mathcal{N}(0, 0.0940)$

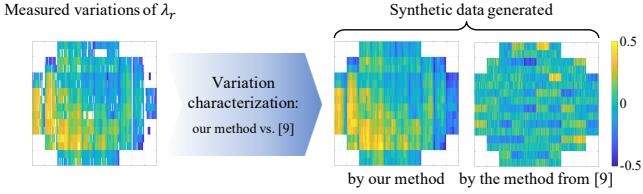


Fig. 11. Comparison of the synthetic data generated by our method and the method from [9].

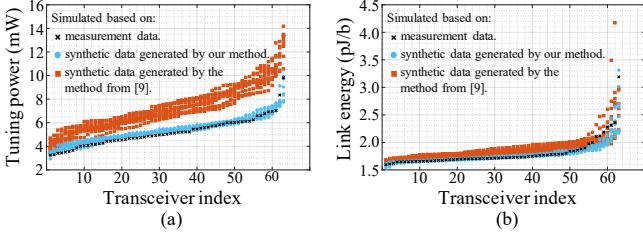


Fig. 12. Simulations of tuning power and link energy show better quality of our synthetic data.

A. Synthetic Data Generation

Synthetic data are often used to complement measurement data for the development and validation of architecture- and system-level variation alleviation techniques [9]–[11]. However, inaccurate synthetic data generated by oversimplified variation models may result in the misestimation of the system performance metrics. We generate synthetic wafer maps for the variations of λ_r , ER , and Q by adding a random term to the characterized systematic components. The random terms follow the normal distributions specified in Table I. As a comparison, we characterize the same measurement data using the method featured in [9] and generate another set of synthetic wafer maps. In [9], the variations of the microrings are decomposed into global, local, and Tx-Rx offset terms. All three terms are directly approximated by independent normal distributions without characterizing their location dependencies. Fig. 11 compares the synthetic data for λ_r generated by the two methods. Visually, our synthetic data can better capture the spatial patterns of the measurement data.

We designed two experiments to further evaluate the quality of our synthetic data, namely to simulate the microring tuning power and the transceiver energy per bit. The microring tuning power is the power required to thermally shift the resonance wavelengths of the Tx and Rx toward a common set of laser wavelengths, and is related to the variations of λ_r . We assume a tuning efficiency of 0.15 nm/mW [52] in the experiments. Fig. 12a plots the average tuning power per channel in ascending order computed for all measured and synthetic transceivers. Ten synthetic wafers were generated using either our method or the method from [9]. The results clearly indicate that our synthetic data closely resemble the measurement data and can estimate the tuning power with substantially higher accuracy than that of [9].

The simulation of the transceiver energy per bit involves variations of λ_r , ER , and Q . It is computed as the total power consumption divided by the data rate. Apart from the microring tuning power, the power consumption of electrical driving circuitry has been modeled in [40] as a function of the data rate. As for the laser power, it has to be sufficiently high to overcome the various power losses within the transceiver link and eventually satisfy the receiver sensitivity

requirement:

$$P_{\text{laser}} \cdot \text{WPE} \cdot \prod_i \text{PL}_i \geq P_{\text{sensitivity}} \quad (13)$$

Here, WPE (wall-plug efficiency) is the efficiency with which the laser converts the electrical power into optical power. The various power losses (PL_i in Eq. 13) are related to the variations of ER and Q . Finally, by specifying a target data rate, the communication energy per bit can be computed for each transceiver with distinct variation profiles. Fig. 12b shows the simulated energy per bit for measured and synthetic transceivers at a target data rate of 20 Gb/s per channel and a laser WPE of 20% . It is further confirmed that our variation characterization method generates more realistic synthetic data than that of [9]. Novel techniques for variation alleviation can thus be evaluated by our synthetic data with improved accuracy.

B. Yield Prediction

Spatial variation models can also be characterized from incomplete wafer maps. We further explore the opportunity for transceiver yield prediction based on a partially measured wafer. We define a *working* transceiver as one that can achieve a bit error rate (BER) of 10^{-12} at a given data rate. We then define the yield as the number of working transceivers on a wafer divided by the total number of transceivers. The authors of [53] modeled the receiver sensitivity requirement for a BER of 10^{-12} as a function of the target data rate. Due to the optical nonlinearities of the microrings and the silicon waveguides, there exists a maximum optical power that can be injected into the transceiver [54]. Based on Eq. 13, a maximum optical power at the receiver end can be computed for each transceiver with distinct variation profiles. If this maximum power at the receiver end is still lower than the sensitivity required for a BER of 10^{-12} , the transceiver is considered *not working* at this data rate.

In the experiments, we characterize the spatial variation models based on the measurement data of randomly sampled transceivers and predict a complete wafer map (Fig. 13a). Yield is computed based on this predicted wafer map and compared to that of a fully measured wafer. In Fig. 13b, we plot the predicted yield for various sampling ratios and target data rates. The dashed lines correspond to the yield computed from full measurement data. As can be seen, a sampling ratio too low tends to produce over-optimistic predictions of the yield, which indicates that the severity of the variations is underestimated. However, a sampling ratio of 60% is high enough for the characterized variation models to accurately predict the yield of the whole wafer. Our variation characterization method can thus be used to investigate the measurement order of the transceivers in order to provide an early estimation of the wafer yield.

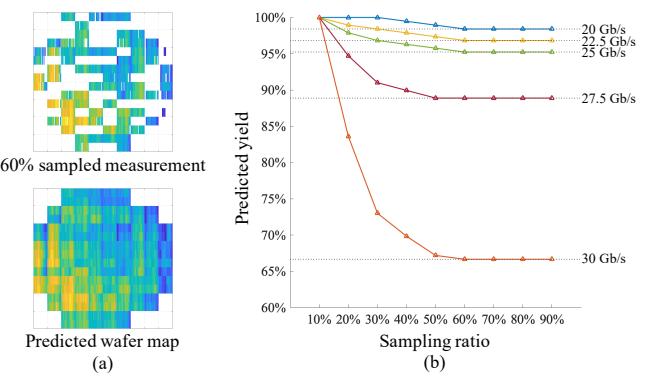


Fig. 13. Predicted yield at various sampling ratio.

V. CONCLUSION AND FUTURE WORK

In this study, we present a hierarchical method for characterizing the spatial variations of microring-based optical transceivers. The extracted wafer-level, intra-die, and inter-die variation components effectively capture the location dependencies of the variations. We demonstrated the applications of the characterized variation models for trustworthy synthetic data generation and accurate yield prediction. If more fabricated wafers become available in the future, topics that are worth exploring include calibration of the spatial variation models, characterization of wafer-to-wafer temporal variations, and broad applications to system-level variation-aware design optimization for optical interconnects.

ACKNOWLEDGMENT

This work was partially supported by a GRF grant from RGC of the Hong Kong SAR, China (HKUST 16207917).

REFERENCES

- [1] B. Jalali *et al.*, "Silicon photonics," *J. Light. Technol.*, 2006.
- [2] R. Soref, "The past, present, and future of silicon photonics," *IEEE J. Sel. Top. Quantum Electron.*, 2006.
- [3] R. G. Beausoleil, "Large-scale integrated photonics for high-performance interconnects," *ACM J. Emerg. Technol. Comput. Syst.*, 2011.
- [4] R. G. Beausoleil *et al.*, "Photonic architectures for high-performance data centers," *IEEE J. Sel. Top. Quantum Electron.*, 2013.
- [5] A. V. Krishnamoorthy *et al.*, "Exploiting cmos manufacturing to reduce tuning requirements for resonant optical devices," *IEEE Photonics J.*, 2011.
- [6] M. Georgas *et al.*, "Addressing link-level design tradeoffs for integrated photonic interconnects," in *2011 IEEE Cust. Integr. Circuits Conf.*, 2011.
- [7] Y. Zheng *et al.*, "Power-efficient calibration and reconfiguration for optical network-on-chip," *J. Opt. Commun. Netw.*, 2012.
- [8] R. Wu *et al.*, "Variation-aware adaptive tuning for nanophotonic interconnects," in *2015 IEEE/ACM Int. Conf. Comput. Des.*, 2015.
- [9] R. Wu *et al.*, "Pairing of microring-based silicon photonic transceivers for tuning power optimization," in *2018 23rd Asia South Pacific Des. Autom. Conf.*, 2018.
- [10] Y. Wang *et al.*, "Energy-efficient channel alignment of dwdm silicon photonic transceivers," *Des. Autom. Test Eur. Conf. Exhib. (DATE)*, 2018, 2018.
- [11] Y. Wang *et al.*, "Bidirectional tuning of microring-based silicon photonic transceivers for optimal energy efficiency," *Asia South Pacific Des. Autom. Conf.*, 2019.
- [12] Y. Wang *et al.*, "Taming emerging devices' variation and reliability challenges with architectural and system solutions [invited]," in *2019 IEEE 32nd Int. Conf. Microelectron. Test Struct.*, 2019.
- [13] K. Qian *et al.*, "Hierarchical modeling of spatial variability with a 45nm example," *Des. Manuf. through Des. Integr. III*, 2009.
- [14] R. Wu *et al.*, "Spatial pattern analysis of process variations in silicon microring modulators," in *2016 IEEE Opt. Interconnects Conf.*, 2016.
- [15] N. Boynton *et al.*, "Characterization of systematic process variation in a silicon photonic platform," *6th IEEE Photonics Soc. Opt. Interconnects Conf. OI 2017*, 2017.
- [16] P. Sun *et al.*, "Statistical Behavioral Models of Silicon Ring Resonators at a Commercial CMOS Foundry," in *2019 IEEE Opt. Interconnects Conf.*, 2019.
- [17] X. Chen *et al.*, "Process variation in silicon photonic devices," *Appl. Opt.*, 2013.
- [18] L. Chrostowski *et al.*, "Impact of fabrication non-uniformity on chip-scale silicon photonic integrated circuits," *Opt. Fiber Commun. Conf. OFC 2014*, 2014.
- [19] Z. Lu *et al.*, "Performance prediction for silicon photonics integrated circuits with layout-dependent correlated manufacturing variability," *Opt. Express*, 2017.
- [20] S. K. Selvaraja *et al.*, "Subnanometer linewidth uniformity in silicon nanophotonic waveguide devices using cmos fabrication technology," *IEEE J. Sel. Top. Quantum Electron.*, 2010.
- [21] Y. Xing *et al.*, "Hierarchical model for spatial variations of integrated photonics," in *2018 IEEE 15th Int. Conf. Gr. IV Photonics*, 2018.
- [22] W. A. Zortman *et al.*, "Silicon photonics manufacturing," *Opt. Express*, 2010.
- [23] C. Yu *et al.*, "Use of short-loop electrical measurements for yield improvement," *IEEE Trans. Semicond. Manuf.*, 1995.
- [24] B. Stine *et al.*, "Inter- and intra-die polysilicon critical dimension variation," in *Proc. SPIE - Int. Soc. Opt. Eng.*, 1996.
- [25] B. Stine *et al.*, "Analysis and decomposition of spatial variation in integrated circuit processes and devices," *IEEE Trans. Semicond. Manuf.*, 1997.
- [26] S. J. Bae *et al.*, "Yield prediction via spatial modeling of clustered defect counts across a wafer map," *IIE Trans.*, 2007.
- [27] W. Zhang *et al.*, "Toward efficient spatial variation decomposition via sparse regression," in *2011 IEEE/ACM Int. Conf. Comput. Des.*, 2011.
- [28] W. Zhang *et al.*, "Efficient spatial pattern analysis for variation decomposition via robust sparse regression," *IEEE Trans. Comput. Des. Integr. Circuits Syst.*, 2013.
- [29] K. Huang *et al.*, "Process monitoring through wafer-level spatial variation decomposition," in *2013 IEEE Int. Test Conf.*, 2013.
- [30] W. Zhang *et al.*, "Bayesian virtual probe: Minimizing variation characterization cost for nanoscale ic technologies via bayesian inference," *Proc. - Des. Autom. Conf.*, 2010.
- [31] W. Zhang *et al.*, "Virtual probe : A statistical framework for low-cost silicon characterization of nanoscale integrated circuits," *IEEE Trans. Comput. Des. Integr. Circuits Syst.*, 2011.
- [32] N. Kupp *et al.*, "Spatial estimation of wafer measurement parameters using gaussian process models," in *2012 IEEE Int. Test Conf.*, 2012.
- [33] C. K. Hsu *et al.*, "Test data analytics - exploring spatial and test-item correlations in production test data," in *Proc. - Int. Test Conf.*, 2013.
- [34] S. Zhang *et al.*, "Joint virtual probe: Joint exploration of multiple test items' spatial patterns for efficient silicon characterization and test prediction," in *Des. Autom. Test Eur. Conf. Exhib. (DATE)*, 2014, 2014.
- [35] J. Luan *et al.*, "Prediction of multi-dimensional spatial variation data via bayesian tensor completion," *IEEE Trans. Comput. Des. Integr. Circuits Syst.*, 2019.
- [36] Q. Xu *et al.*, "Micrometre-scale silicon electro-optic modulator," *Nature*, 2005.
- [37] Q. Xu *et al.*, "Cascaded silicon micro-ring modulators for wdm optical interconnection," *Opt. Express*, 2006.
- [38] C. H. Chen *et al.*, "Concurrent multi-channel transmission of a dwdm silicon photonic transmitter based on a comb laser and microring modulators," in *2015 International Conference on Photonics in Switching (PS)*, 2015.
- [39] M. A. Seyed *et al.*, "Concurrent dwdm transmission with ring modulators driven by a comb laser with 50ghz channel spacing," in *2016 21st OptoElectronics and Communications Conference (OECC) held jointly with 2016 International Conference on Photonics in Switching (PS)*, 2016.
- [40] R. Polster *et al.*, "Efficiency optimization of silicon photonic links in 65-nm cmos and 28-nm fdsoi technology nodes," *IEEE Trans. Very Large Scale Integr. Syst.*, 2016.
- [41] B. E. Little *et al.*, "Surface-roughness-induced contradirectional coupling in ring and disk resonators," *Opt. Lett.*, 1997.
- [42] A. Li *et al.*, "Backscattering in silicon microring resonators: a quantitative analysis," *Laser Photon. Rev.*, 2016.
- [43] R. Wu, "Variation-aware modeling and design of nanophotonic interconnects," Ph.D. dissertation, 2017.
- [44] T. F. Coleman *et al.*, "An interior trust region approach for nonlinear minimization subject to bounds," *SIAM J. Optim.*, 1996.
- [45] P. J. Huber, *Robust Statistics*, 2011.
- [46] G. Upton *et al.*, *Understanding statistics*, 1996.
- [47] S. Rabanser *et al.*, "Introduction to tensor decompositions and their applications in machine learning," 2017.
- [48] E. Acar *et al.*, "Scalable tensor factorizations for incomplete data," *Chemom. Intell. Lab. Syst.*, 2011.
- [49] Q. Zhao *et al.*, "Bayesian robust tensor factorization for incomplete multiway data," *IEEE Trans. Neural Networks Learn. Syst.*, 2016.
- [50] T. Yokota *et al.*, "Smooth parafac decomposition for tensor completion," *IEEE Trans. Signal Process.*, 2016.
- [51] L. Yuan *et al.*, "Tensor ring decomposition with rank minimization on latent space: An efficient approach for tensor completion," *Proc. AAAI Conf. Artif. Intell.*, 2019.
- [52] K. Yu *et al.*, "A 25 gb/s hybrid-integrated silicon photonic source-synchronous receiver with microring wavelength stabilization," *IEEE J. Solid-State Circuits*, 2016.
- [53] M. Bahadori *et al.*, "Energy-performance optimized design of silicon photonic interconnection networks for high-performance computing," in *Des. Autom. Test Eur. Conf. Exhib. (DATE)*, 2017, 2017.
- [54] Q. Li *et al.*, "Experimental characterization of the optical-power upper bound in a silicon microring modulator," *2012 Opt. Interconnects Conf. OIC 2012*, 2012.