Wafer level characterization of silicon nitride CWDM (de)multiplexers using Bayesian inference

Jun Rong Ong, Tina X. Guo, Thomas Y. L. Ang, Soon Thor Lim, Hong Wang, Ching Eng Png

***Abstract*—A cascaded Mach-Zehnder interferometer based filter for coarse wavelength (de)multiplexing (CWDM) at the O-band is fabricated and tested on a silicon nitride on SOI platform. We characterize the chip-to-chip performance variability of the filter devices on a wafer. Using the optical measurement data, we apply Bayesian inference methods to estimate the waveguide geometery parameters and also quantify the uncertainty of the estimates.**

***Index Terms*—Parameter estimation, Wavelength division multiplexing, Photonic integrated circuits, Silicon photonics**

I. Introduction

Wavelength division multiplexing is an essential technology needed to meet the ever increasing demand for bandwidth in long haul and data center communications. In particular, 4-channel CWDM at O-band wavelengths with 20 nm spacing (1271, 1291, 1311 and 1331 nm) is a common specification for 100G data center applications. A (de)multiplexing filter is needed to split and combine the 4 different wavelengths. Such a filter should have the following characteristics: accurate passband spacing centered at channel wavelengths, low loss and loss uniformity within the channel, sufficiently wide and flat passband to support high speed signals, low inter-channel crosstalk and low temperature sensitivity. Several different kinds of WDM filters have been proposed and demonstrated, e.g. arrayed waveguide gratings [1], echelle gratings [2] and contra-directional couplers [3]. Of these, cascaded Mach-Zehnder interferometer (MZI) filters are of particular interest due to their low insertion loss, low cross-talk and flat passband charateristics [4]–[11].

In this letter, we design and demonstrate a 4-channel CWDM cascaded MZI filter at O-band using silicon nitride (SiN) waveguides on SOI platform. SiN has a lower index contrast and lower thermo-optic coefficient than SOI and hence has the advantages of lower suceptibility to fabrication error and thermal drift [7]. Nevertheless, the spectral response of the filters, which depend on interference effects, are still affected by process variations. Performance prediction of devices under process variation is an area of research which has received greater attention recently, especially for high index contrast SOI photonics [12], [13]. This involves obtaining device performance statistics, extracting process variations on a wafer level, and performing robustness optimization. As such, we have measured cascaded MZI filter devices on a sample of 22 chips to obtain wafer level statistics. To further make sense of the data, we constructed models of the filters through polynomial fitting of the waveguide and directional coupler effective indices. Using these models and together with statistical inference methods, we are able to extract waveguide geometric parameters. Furthermore, by using Bayesian inference we also obtain the uncertainty of the estimates. Accurate extraction of waveguide geometry usually requires specialized test structures at scattered locations on the wafer [14]–[16]. Our way of parameter extraction can be complementary to these methods by making use of data collected directly from the device-under-test itself to give additional insights.

Manuscript received Month D, YYYY; revised Month DD, YYYY. This work was supported by A*STAR-NTU-SUTD AI Partnership Grant (RGANS1901).

J.R. Ong, T.Y.L. Ang, S.T. Lim and C.E. Png are with the Institute of High Performance Computing, Agency for Science, Technology and Research (A*STAR), 138632, Singapore. (e-mail: ongjr@ihpc.a-star.edu.sg; pngce@ihpc.a-star.edu.sg)

T.X. Guo and H. Wang are with Silicon Technologies – Centre of Excellence, Nanyang Technological University 50 Nanyang Avenue, S1-B2b-15, 639798, Singapore. (email: ewanghong@ntu.edu.sg)

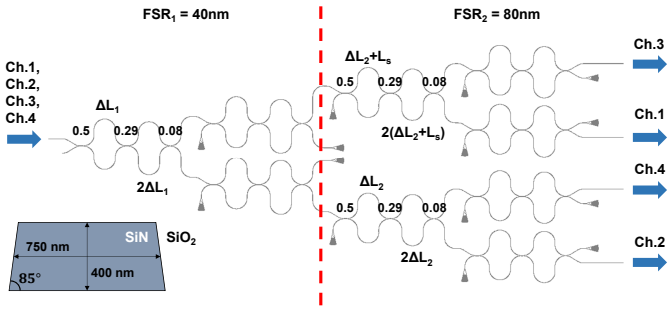

Fig. 1. 4-channel cascaded MZI filter design and layout. The directional coupler coefficients κ^2 and differential lengths ΔL are labeled. L_S is an additional length to center the odd channels. Inset: silicon nitride (SiN) on SOI waveguide cross-section.

II. Design and Modeling

Cascaded MZI filters are also called lattice Fourier filters, since the transfer function is the sum of all the distinct optical paths in the interferometer, which together make up a truncated Fourier series [4]. Our device design is similar to [6] and consists of two stages (see Fig. 1): the first stage has a free-spectral range (FSR) of 40 nm acting as an interleaver to separate odd and even wavelength channels, the second stage has a FSR of 80 nm and filters out the individual channels. Each stage of the lattice filter consists of two cascaded building block filters to reduce cross-talk and each building block filter consists of three directional couplers and two differential lengths. An additional length is needed at the second stage to shift the center of some of the channels to the correct wavelengths [5]. We used a common waveguide

cross-section throughout the device, which is nominally of $w_0 = 750$ nm width and $h_0 = 400$ nm height, within the single mode condition. The bending radius is 40 μm and the coupler gap is fixed at 400 nm to meet the design rule specifications [17]. Since the lattice filter only consists of differential waveguide lengths and directional couplers, the filter transmission spectrum can be calculated efficiently using transfer matrices. We apply polynomial regression fitting to build accurate models of the effective indices which are used to calculate the phase shifts and coupling coefficients in the transfer matrix model. Below, we describe more details of our lattice filter model.

A. Polynomial regression fit of effective indices

We first generated a simulation dataset by calculating the effective indices of the SiN waveguide quasi-TE modes n_{eff} and directional coupler symmetric and anti-symmetric supermodes n_s, n_a using a commercial finite difference solver (Lumerical). The refractive index used for the LPCVD SiN was 1.98 [18]. The swept range (in μm) of the dependent parameters are as follows: waveguide width $w \in$ (0.68, 0.82), waveguide height $h \in$ (0.34, 0.46), wavelength $\lambda \in$ (1.26, 1.36) and (for coupler supermodes) the coupler gap $g \in$ (0.3, 0.5). These parameter ranges were chosen according to the maximum geometry variations, $\Delta w = \pm 50$nm and $\Delta h = \pm 40$nm [17]. The trapezoidal waveguide sidewall angle is fixed at 85° and the width w is the middle width (see Fig. 1). We assume that the directional coupler is symmetric, i.e. both waveguides are of the same width, and the gap g is the middle gap. We use 4-th order polynomial regression fits to build accurate models of the effective indices [15], with the maximum error obtained at 0.01% which equals to a phase error of $< 10^{-3}$ rad/μm.

B. Directional coupler model

The directional couplers can be modeled as unitary 2×2 matrices $\left[\begin{smallmatrix} t & i\kappa \\ i\kappa & t \end{smallmatrix}\right]$, with $\kappa^2 + t^2 = 1$. We assume identical coupled waveguides, so $\kappa^2 = \sin^2\left(\frac{\pi}{2}\frac{l_c}{L_c}\right)$, with $L_c = \frac{\lambda}{2(n_s - n_a)}$. The 0.01% error of our polynomial fit of the supermode effective indices n_s, n_a corresponds to 0.5% error in predicted L_c. Additionally, we have to include the coupling effect at the bends where the coupled waveguides are brought close together such that the total coupling length is the sum of the straight and bend contributions, $l_c = l_s + l_b$. Typically, we can calculate l_b by an integral over the coupler gap as a function of position [19]. However, we noticed this method slightly underestimated l_b for our chosen radius of 40 μm, particularly for thinner and narrower waveguides that have poorer optical mode confinement. As such, we generated an additional dataset of the coupling length contributions from bends using 3D FDTD simulations (Lumerical). Since 3D FDTD simulations are computationally expensive, we assumed that the positions of the waveguides are fixed (i.e. $g + w =$ 1150 nm) to reduce the size of the parameter space to cover. We used a 4-th order polynomial fit to model l_b, giving a maximum error of 0.1%.

III. Measurements

To obtain wafer level statistics, we measured the 4 channel transmission spectrum of the cascaded MZI filters on 22 different chips across a wafer. We used a tunable laser source, scanning from 1260 nm to 1360 nm, in steps of 0.2 nm. A waveguide adjacent to the filter, with fiber-to-silicon-waveguide edge couplers and silicon-to-silicon-nitride interlayer couplers [20], was used to normalize out the coupling losses. As mentioned previously, a good performing filter can be determined by measurement of several key metrics: passband center wavelength, insertion loss, 1dB passband bandwidth, inter-channel crosstalk. We show in Table I summary statistics of these metrics. We see that the (median) passband wavelengths are shifted from their nominal design values by about -5 nm, with an the inter-quartile range (IQR) of 1.4 nm. Comparing the insertion loss and crosstalk at the nominal channel wavelength and the passband center wavelength, we see a degradation in performance due to the wavelength shift. Fortunately, because of the large 1dB bandwidth, the insertion loss and crosstalk was not too severely degraded. Channel 2 was most affected due to its narrowest passband.

TABLE I
Statistics of measured performance specifications

		Ch. 1	Ch. 2	Ch. 3	Ch. 4
Nom. λ (nm)		1271	1291	1311	1331
Center λ (nm)	Median	1266.2	1285.6	1305.3	1326.5
	IQR	1.4	1.4	1.4	1.4
	Max.	1271.5	1290.9	1310.7	1332.0
	Min.	1261.3	1280.6	1300.3	1321.4
1dB bandwidth (nm)	Median	9.6	8.9	12.8	13.9
	IQR	0.7	0.8	0.1	0.4
	Max.	10.1	9.8	12.9	14.3
	Min.	8.8	7.8	12.6	13.6
Insertion Loss (dB)	Median	4.5	5.3	4.1	2.9
at nom. λ	IQR	1.5	2.2	1.4	1.6
	Max.	16.9	20.2	17.6	13.9
	Min.	2.7	3.4	2.2	1.3
Crosstalk (dB)	Median	-20.9	-16.1	-22.5	-23.4
at nom. λ	IQR	4.6	4.6	2.6	5.5
	Max.	-12.3	-12.3	-13.9	-17.2
	Min.	-26.6	-28.1	-30.5	-34.6
Insertion Loss (dB)	Median	3.9	3.0	3.6	2.8
at center λ	IQR	1.4	2.6	2.3	1.4
	Max.	10.9	10.6	10.4	9.3
	Min.	2.1	1.8	1.7	1.4
Crosstalk (dB)	Median	-24.4	-24.0	-23.2	-27.3
at center λ	IQR	3.4	7.2	6.1	4.7
	Max.	-17.6	-17.7	-14.5	-19.7
	Min.	-31.8	-33.7	-37.6	-34.0

To gain more insight from the data, we perform a least squares data fitting using our transfer matrix model to obtain estimates of the waveguide geometry w and h. We minimize the sum of squared errors

$$S = \sum_{n=1}^{4} \sum_{i=1}^{N} \left(y_{n,i} - f(\lambda_{n,i}, w, h)\right)^2 \tag{1}$$

where $y_{n,i}$ is the measured normalized transmission at the n-th channel and the i-th wavelength, giving a total of $4N$ data points. Our transfer matrix model prediction $f(\lambda_{n,i}, w, h) = \hat{y}_{n,i}$ assumes $g + w = 1150$ nm, so g is not an adjustable

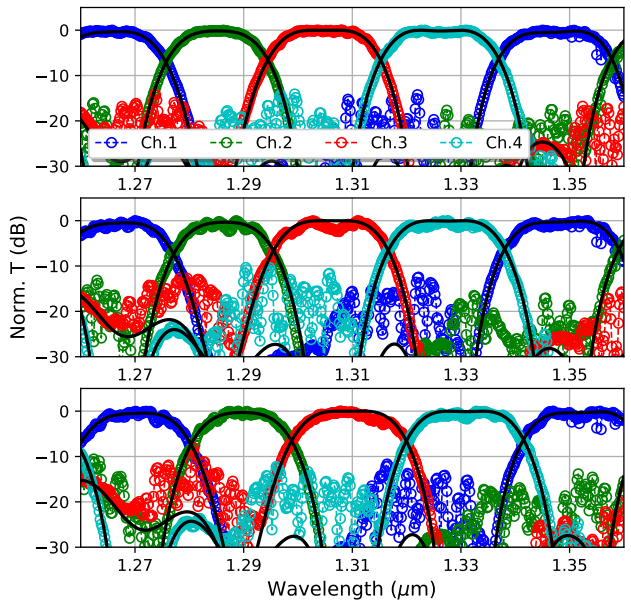

Fig. 2. Normalized transmission spectra of 4-channel (de)multiplexer obtained from measurement of three different samples. Black lines indicate the predictions obtained from least squares fitting model.

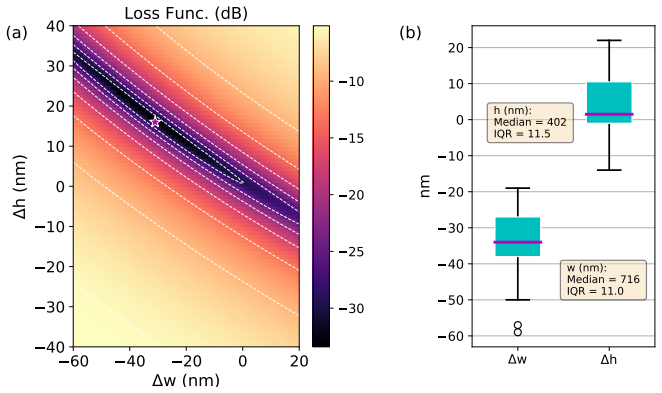

Fig. 3. (a) An example of the loss function contours in dB as a function of the geometry deviations Δw and Δh. The star indicates the minimum point. (b) Boxplots of geometry deviations extracted from the measured samples, using least squares fitting. The text boxes show the median and IQR of extracted w and h.

parameter. Figure 2 shows a few examples of the fitted data, with good agreement between model and experiment. We note also that there is an additional source of crosstalk in the experiment data that is not present in the model, which we believe originates from the limited polarization extinction when coupling from the laser source. The crosstalk can be reduced by inserting a high extinction polarization beam splitter either before coupling to the chip or on the chip itself [21].

Figure 3(a) shows an example of the loss function contours as a function of the geometry deviations, $\Delta w = w - w_0$ and $\Delta h = h - h_0$. The minimum point (star) gives the estimate of Δw and Δh. We perform this data fitting on the 22 measured samples to obtain the wafer level summary statistics for the waveguide geometry estimates (see Fig. 3(b)). The distribution of extracted heights has a median of 402 nm and IQR of 11.5

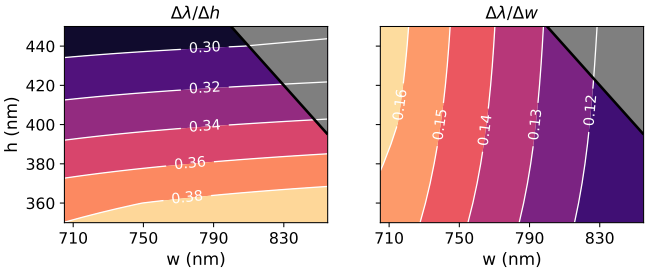

Fig. 4. Contour plots of the sensitivities $\frac{\Delta\lambda}{\Delta w}$, $\frac{\Delta\lambda}{\Delta h}$ with units of nm $\cdot$ nm^{-1}, as a function of waveguide geometry w and h. The gray regions indicate multi-mode waveguides.

nm and is well centered near the nominal height h_0. However, the distribution of extracted widths has a median of 716 nm and IQR of 11.0 nm, which is different from the nominal width w_0 by -34 nm. The channel wavelength shift of about -5 nm can be attributed to this difference.

The amount of shift in the peaks of MZI interference fringes caused by geometry changes is $\frac{\Delta\lambda}{\Delta x} = \frac{dn_{\text{eff}}}{dx}\frac{\lambda}{n_g}$ to a first order approximation, where x represents w or h [22]. In Fig. 4, we plot the peak wavelength sensitivities $\frac{\Delta\lambda}{\Delta w}$, $\frac{\Delta\lambda}{\Delta h}$. We can see that greater insensivity to geometry changes is obtained by increasing the width and height of the waveguides, while being careful to remain outside the multi-mode region (gray). Although we have designed our filters for relatively flat and wide passbands, we still observed performance degradation due to the passband wavelength shift. Future designs should make use of this geometry insensitivity, since too large a wavelength shift can cause degradation in performance due to the steep roll-off of the passbands.

IV. Bayesian Inference

Least squares fitting, which is equivalent to maximum likelihood estimation in the probabilistic interpretation, gives us a point estimate. Next, we perform Bayesian estimation which gives an uncertainty of the estimate and also lets us incorporate prior information regarding the parameters. We model the measurement outcome y as a normally distributed random variable $Y \sim \mathcal{N}(\mu = \hat{y}, \sigma^2)$. We apply normal priors on the parameters w, h and a half-normal prior on the standard deviation σ of the error term $\epsilon_{n,i} = y_{n,i} - \hat{y}_{n,i}$

$$w \sim \mathcal{N}(\mu_w = w_0, \sigma_w^2) \tag{2a}$$
$$h \sim \mathcal{N}(\mu_h = h_0, \sigma_h^2) \tag{2b}$$
$$\sigma \sim |\mathcal{N}(\mu_\epsilon = 0, \sigma_\epsilon^2)| \tag{2c}$$

choosing $\sigma_w = \sigma_h = 60$ nm and $\sigma_\epsilon = 0.1$ to give weak priors, so as to not overly constrain the inferred values. The error ϵ accounts for random deviations due to measurement noise and fabrication imperfections that are not included in the model. We do not consider a joint distribution for w, h as we did not observe a strong correlation from the least squares estimates. Using Bayes' theorem, with the likelihood and priors, we sample from the posterior distribution

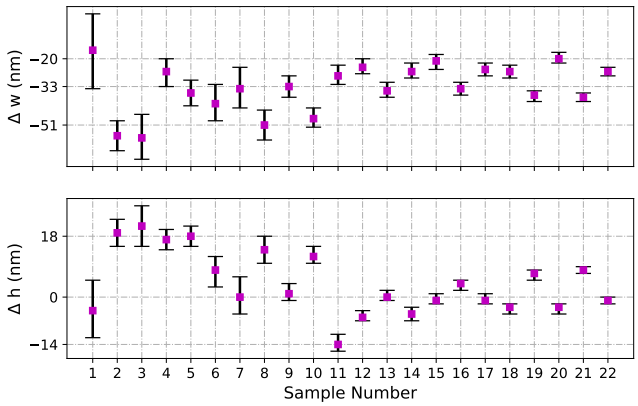

Fig. 5. Estimates for geometry deviations Δw and Δh, arranged in descending order of σ. Large σ corresponds to larger errors and hence more uncertainty in geometry estimates.

$P(w, h, \sigma|y) \propto P(y|w, h, \sigma)P(w, h, \sigma)$. The log-likelihood is given as

$$\log L = -\frac{1}{2}\sum_{n,i}\left(\frac{(y_{n,i} - \hat{y}_{n,i})^2}{\sigma^2} + \log(2\pi\sigma^2)\right). \quad (3)$$

We use probabilistic programming to perform inference with the Markov chain Monte Carlo method [23]. Figure 5 shows the parameter estimates, sorted in descending order of σ, with the markers indicating the mean and the whiskers showing the 94% credible intervals. A larger σ indicates larger errors ϵ in the measurement data, and thus at the same time more uncertainty in the geometry estimates. Dropping the samples with the 3 largest σ gives a revised estimate for width and height $(\text{Median}, \text{IQR})$ of $(717, 10.5)$ nm and $(400, 10.0)$ nm. The inferred distribution for σ is $(0.041, 0.021)$.

V. Conclusion

We have designed and fabricated a cascaded MZI filter for O-band CWDM applications and performed wafer level testing. The 4 channel transmission spectrum data from a sample of 22 chips on a single wafer was analyzed, giving the wafer level statistics of the device performance. A fast and accurate transfer matrix model, using polynomial regression fit of the effective indices, was used to characterize the measurement data. By least squares data fitting, we obtained estimates of the waveguide geometry deviations Δw and Δh. Finally, by Bayesian estimation, we quantified the uncertainty of the Δw and Δh estimation. We obtained a wafer median width and height estimate of 717 nm and 400 nm respectively. The parameter estimation results points to the waveguide width deviation Δw as the source of the observed passband wavelength shift. The methods presented will be useful for process monitoring and performance variability analysis which is crucial for scalable production of silicon photonic circuits.

References

[1] S. Guerber, C. Alonso-Ramos, D. Perez-Galacho, X. Le Roux, N. Vulliet, S. Crémer, D. Marris-Morini, F. Boeuf, L. Vivien, and C. Baudot, "Design and integration of an o-band silicon nitride awg for cwdm applications," in *2017 IEEE 14th International Conference on Group IV Photonics (GFP)*. IEEE, 2017, pp. 133–134.

[2] C. Sciancalepore, R. J. Lycett, J. A. Dallery, S. Pauliac, K. Hassan, J. Harduin, H. Duprez, U. Weidenmueller, D. F. Gallagher, S. Menezo *et al.*, "Low-crosstalk fabrication-insensitive echelle grating demultiplexers on silicon-on-insulator," *IEEE Photon. Technol. Lett.*, vol. 27, no. 5, pp. 494–497, 2014.

[3] W. Shi, H. Yun, C. Lin, M. Greenberg, X. Wang, Y. Wang, S. T. Fard, J. Flueckiger, N. A. Jaeger, and L. Chrostowski, "Ultra-compact, flat-top demultiplexer using anti-reflection contra-directional couplers for cwdm networks on silicon," *Opt. Express*, vol. 21, no. 6, pp. 6733–6738, 2013.

[4] Y. Li, C. Henry, E. Laskowski, H. Yaffe, and R. Sweatt, "Monolithic optical waveguide 1.31/1.55 μm wdm with -50 db crosstalk over 100 nm bandwidth," *Electron. Lett.*, vol. 31, no. 24, pp. 2100–2101, 1995.

[5] F. Horst, W. M. Green, S. Assefa, S. M. Shank, Y. A. Vlasov, and B. J. Offrein, "Cascaded mach-zehnder wavelength filters in silicon photonics for low loss and flat pass-band wdm (de-) multiplexing," *Opt. Express*, vol. 21, no. 10, pp. 11 652–11 658, 2013.

[6] S. Dwivedi, P. De Heyn, P. Absil, J. Van Campenhout, and W. Bogaerts, "Coarse wavelength division multiplexer on silicon-on-insulator for 100 gbe," in *2015 IEEE 12th International Conference on Group IV Photonics (GFP)*. IEEE, 2015, pp. 9–10.

[7] G. Gao, D. Chen, S. Tao, Y. Zhang, S. Zhu, X. Xiao, and J. Xia, "Silicon nitride o-band (de) multiplexers with low thermal sensitivity," *Opt. Express*, vol. 25, no. 11, pp. 12 260–12 267, 2017.

[8] S. Tao, Q. Huang, L. Zhu, J. Liu, Y. Zhang, Y. Huang, Y. Wang, and J. Xia, "Athermal 4-channel (de-) multiplexer in silicon nitride fabricated at low temperature," *Photonics Res.*, vol. 6, no. 7, pp. 686–691, 2018.

[9] J. Mikkelsen, A. Bois, T. Lordello, D. Mahgerefteh, S. Menezo, and J. Poon, "Polarization-insensitive silicon nitride mach-zehnder lattice wavelength demultiplexers for cwdm in the o-band," *Opt. Express*, vol. 26, no. 23, pp. 30 076–30 084, 2018.

[10] S. Cheung and M. R. Tan, "Silicon nitride (si3n4) (de-) multiplexers for 1-μm cwdm optical interconnects," *Journal of Lightwave Technology*, 2019.

[11] S.-H. Jeong, "Broadband $1\times$ 8 channel silicon-nanowire-waveguide wdm filter based on point-symmetric mach-zehnder interferometric optical couplers in the o-band spectral regime," *OSA Continuum*, vol. 2, no. 12, pp. 3564–3575, 2019.

[12] Y. Xing, D. Spina, A. Li, T. Dhaene, and W. Bogaerts, "Stochastic collocation for device-level variability analysis in integrated photonics," *Photonics Res.*, vol. 4, no. 2, pp. 93–100, 2016.

[13] A. Waqas, D. Melati, P. Manfredi, and A. Melloni, "Stochastic process design kits for photonic circuits based on polynomial chaos augmented macro-modelling," *Opt. Express*, vol. 26, no. 5, pp. 5894–5907, 2018.

[14] Z. Lu, J. Jhoja, J. Klein, X. Wang, A. Liu, J. Flueckiger, J. Pond, and L. Chrostowski, "Performance prediction for silicon photonics integrated circuits with layout-dependent correlated manufacturing variability," *Opt. Express*, vol. 25, no. 9, pp. 9712–9733, 2017.

[15] Y. Xing, J. Dong, S. Dwivedi, U. Khan, and W. Bogaerts, "Accurate extraction of fabricated geometry using optical measurement," *Photonics Res.*, vol. 6, no. 11, pp. 1008–1020, 2018.

[16] 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*, vol. 3, no. 2, pp. 379–390, 2020.

[17] "Compoundtek pte ltd." https://compoundtek.com/, accessed: 3-April-2020.

[18] C. J. Krückel, A. Fülöp, Z. Ye, P. A. Andrekson *et al.*, "Optical bandgap engineering in nonlinear silicon nitride waveguides," *Opt. Express*, vol. 25, no. 13, pp. 15 370–15 380, 2017.

[19] C. K. Madsen and J. H. Zhao, *Optical filter design and analysis*. Wiley New York, 1999.

[20] Y. Huang, J. Song, X. Luo, T.-Y. Liow, and G.-Q. Lo, "Cmos compatible monolithic multi-layer si 3 n 4-on-soi platform for low-loss high performance silicon photonics dense integration," *Opt. Express*, vol. 22, no. 18, pp. 21 859–21 865, 2014.

[21] J. R. Ong, T. Y. Ang, E. Sahin, B. Pawlina, G. Chen, D. Tan, S. T. Lim, and C. E. Png, "Broadband silicon polarization beam splitter with a high extinction ratio using a triple-bent-waveguide directional coupler," *Opt. Lett.*, vol. 42, no. 21, pp. 4450–4453, 2017.

[22] S. Dwivedi, H. D'heer, and W. Bogaerts, "Maximizing fabrication and thermal tolerances of all-silicon fir wavelength filters," *IEEE Photon. Technol. Lett.*, vol. 27, no. 8, pp. 871–874, 2015.

[23] J. Salvatier, T. V. Wiecki, and C. Fonnesbeck, "Probabilistic programming in python using pymc3," *PeerJ Comput. Sci.*, vol. 2, p. e55, 2016.