A Parameter-Space-Reduction-Technique-Assisted Optimization Method for Characterizing Recirculating Waveguide Meshes. Supplementary Material

Ran Tao, Jifang Qiu,* Yuchen Chen, Yan Li, Hongxiang Guo and Jian Wu
State Key Laboratory of Information Photonics and Optical Communications, School of Electronic Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China
**jifangqiu@bupt.edu.cn*

Section S1: state of art in recirculating waveguide meshes characterization

Table S1 present the state of art in recirculating waveguide meshes characterization, summarizing the range of parameters that can be characterized by various methods. ✓ indicates that the parameter can be characterized, while × indicates it cannot be obtained. Our method can broaden the range of imperfection parameters that can be characterized, making it sufficient for use in simulation methods and analytical gradient calculation methods, thus paving the way towards a much faster off-chip configuration.

Table S1: State of the art of recirculating waveguide meshes characterization.

		Whether characterized			
		[1]	[2]	[3]	This work
Parameter	Voltage required to set TBU to cross/bar state	✓	✓	✓	✓
	Coupling factor versus applied voltage	×	✓	✓	✓
	Passive phase difference between TBU upper and lower arm phase shifters $(d\vartheta = \vartheta_l - \vartheta_u)$	×	×	✓	✓
	Passive phase of each phase shifter $(\vartheta_u \text{ and } \vartheta_l)$	×	×	×	✓
	Phase increment versus voltage of phase shifter $(\varphi_u(V) \text{ and } \varphi_l(V))$	×	×	×	✓
	Group index of waveguide (n_g)	×	×	×	✓
	Splitting ratio of beam splitters $(K_{BS,\alpha}, K_{BS,\beta})$	×	×	×	✓
	Whether sufficient for simulation methods [4-7] or analytical gradient calculation methods [5, 6]	×	×	×	✓

The working principle of simulation methods and gradient calculation methods require the specific imperfection parameters of each component of TBU, such as passive phase of its phase shifter, beam splitting ratio of its beam splitter, rather than just the resultant TBU coupling factor. Thus, existing methods from earlier works are not sufficient for these methods.

Section S2: Modeling of recirculating waveguide meshes considering fabrication imperfections

As shown in Fig. S1(a), recirculating waveguide meshes rely on the interconnection of TBUs (Fig. S1(b)), which is complemented with a symmetric MZI composed of 50:50 beam splitters (BSs) and phase shifters

(PSs) attached to both arms. As labeled in Fig. S1(b), due to fabrication errors, the passive phase (without voltage applied) (ϑ_u and ϑ_l) and phase-voltage relationship curves ($\varphi_u(V)$ and $\varphi_l(V)$) of the upper and lower PS, splitting ratio of BSs (α and β), group index of waveguide (n_g) may deviate from the ideal one. Here we build a model of recirculating waveguide meshes considering these fabrication imperfections.

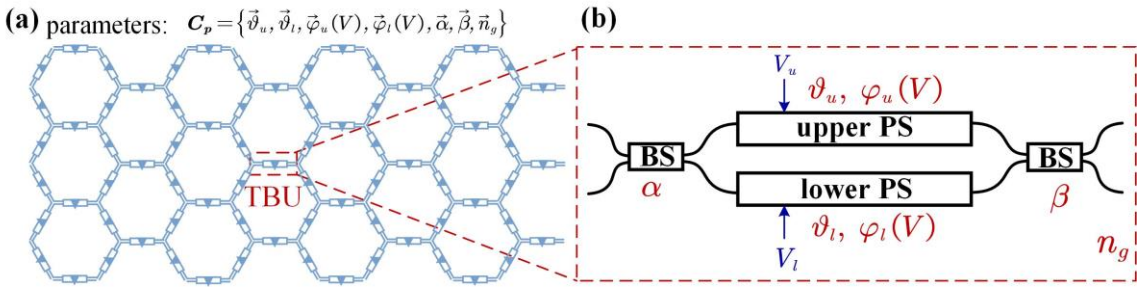

Fig. S1. (a) Recirculating waveguide mesh consisting of interconnected TBUs (b) TBU employing Mach-Zehnder interferometer (MZI) structure, composed of beam splitters (BSs) and phase shifters (PSs). ϑ_u and ϑ_l are the passive phase of the PS on the upper and lower arm of MZI. $\varphi_u(V)$ and $\varphi_l(V)$ are the phase-voltage relationship curves of the upper and lower PS. α and β indicate the splitting ratio of BSs. n_g indicated the group index of waveguide. V_u and V_l are the voltages applied on the upper and lower PS.

The scattering matrix of an imperfect TBU can be expressed as:

$$S_{TBU}=\begin{bmatrix}\cos(\beta) & j\sin(\beta)\\ j\sin(\beta) & \cos(\beta)\end{bmatrix}\begin{bmatrix}e^{j\theta_u} & 0\\ 0 & e^{j\theta_l}\end{bmatrix}\begin{bmatrix}\cos(\alpha) & j\sin(\alpha)\\ j\sin(\alpha) & \cos(\alpha)\end{bmatrix} \tag{S1}$$

where, θ_u/θ_l is the phase shift of the upper/lower PS when apply voltage V_u/V_l on it. Considering recirculating waveguide meshes are often used to configuring filters, thus their performance across the entire frequency band needs to be evaluated. Therefore, we perform a first-order Taylor expansion of θ_u and θ_l with respect to optical frequency f, to provide a multi-frequency model. θ_u and θ_l can be expressed as:

$$\theta_u(f,V_u)=\vartheta_u+\frac{2\pi}{c}Ln_g(f-f_0)+\varphi_u(V_u) \tag{S2}$$

$$\theta_l(f,V_l)=\vartheta_l+\frac{2\pi}{c}Ln_g(f-f_0)+\varphi_l(V_l) \tag{S3}$$

where ϑ_u and ϑ_l are the passive phase (without voltage applied) of the upper and lower PS at reference frequency f_0, $\varphi_u(V)$ and $\varphi_l(V)$ are phase increment versus voltage curves of the upper and lower PS, L is the length of the PS, n_g is the group index of waveguide.

With $\boldsymbol{C_p}=\{\vec{\vartheta}_u,\vec{\vartheta}_l,\vec{\varphi}_u(V),\vec{\varphi}_l(V),\vec{\alpha},\vec{\beta},\vec{n}_g\}$ (the corresponding parameters of each TBU in the mesh) known, the scattering matrix of each TBU can be deduces using Eq. (S1)(S2)(S3), and the scattering matrix of the whole mesh can then be deduced using method proposed in [4-7].

Section S3: Characterization of $d\vec{\vartheta}=\vec{\vartheta}_l-\vec{\vartheta}_u$, $\vec{\varphi}_u(V)$ and $\vec{\varphi}_l(V)$

Using method proposed in [3], for each TBU, its coupling factor relative to the applied voltages (V) can be characterized. In more detail, the coupling factor of TBU are affect by both the upper and lower phase shifter of the TBU. We first keep the lower PS unactuated (without voltage applied), and sweep the voltage applied on the upper PS to characterize the evolution curve of coupling factor (such coupling factor is gained by actuating the upper PS, thus called $K_u(V)$). Then keep the upper PS unactuated, and sweep the voltage applied on the lower PS to characterize coupling factor (such coupling factor is gained by actuating the lower

PS, thus called $K_l(V)$). From these two curves $K_u(V)$ and $K_l(V)$, we can then retrieve $d\vartheta = \vartheta_l - \vartheta_u$, $\varphi_u(V)$ and $\varphi_l(V)$. The detailed principle is given as follows:

We can first retrieve $d\vartheta = \vartheta_l - \vartheta_u$. $d\vartheta$ can be extracted from $K_u(V)$ using Eq.(S4) depending on the $K_u(V)$ curve's upward/downward trend at $V = 0$.

$$d\vartheta = \begin{cases} \pi - \arccos(K_u(0)) & K_u(\delta V) < K_u(0) \\ \arccos(K_u(0)) + \pi & K_u(\delta V) \geq K_u(0) \end{cases} \tag{S4}$$

where $K_u(0)$ and $K_u(dV)$ respectively corresponding to coupling factor when apply voltage $V = 0$ and $V = \delta V$ on the upper PS. δV is a small increment of applied voltage V, used to determine the upward/downward trend of $K_u(V)$ at $V = 0$.

Secondly, with $d\vartheta$ known, we can then extract $\varphi_u(V)$ from $K_u(V)$ using Eq.(S5).

$$\varphi_u(V) = \begin{cases} mod(\arccos(K_u(V)) - \pi + d\vartheta) & K_u(V + \delta V) < K_u(V) \\ mod(-\arccos(K_u(V)) - \pi + d\vartheta) & K_u(V + \delta V) \geq K_u(V) \end{cases} \tag{S5}$$

similarly, the small increment δV is used to determine the upward/downward trend of $K_u(V)$ at V. mod means taking the modulus in radians.

Lastly, we can extract $\varphi_l(V)$ from $K_l(V)$ using Eq.(S6).

$$\varphi_l(V) = \begin{cases} mod(\arccos(K_l(V)) + \pi - d\vartheta) & K_l(V + \delta V) < K_l(V) \\ mod(-\arccos(K_l(V)) + \pi - d\vartheta) & K_l(V + \delta V) \geq K_l(V) \end{cases} \tag{S6}$$

Section S4: Characterization of $\vec{n}_g$

To characterize n_g, we synthesize a MZI resonator on the mesh by setting certain TBUs to cross, bar, or 50:50 coupling state (the voltage values needed for TBUs to reach their target coupling state is known after Step 1). Then we extract Free Spectral Range (FSR) from the resonance spectrum and calculate n_g using Eq. (S7).

$$n_g = \frac{c}{10 \cdot L \cdot FSR} \tag{S7}$$

Note that, we assume a same value of n_g for all TBUs, since the difference of group refractive index among waveguides within the same region is not significant [8, 9].

Section S5: Narrowing down value range of $\vec{\vartheta}_u$ and $\vec{\vartheta}_l$

This method ingeniously constructs two delay lines, allowing for the extraction of sum of ϑ_u and ϑ_l from the phase difference between the two delay lines. Combine with the information of $d\vartheta = \vartheta_l - \vartheta_u$, which we have already obtained in Step 1 of our characterization procedure, the value of ϑ_u and ϑ_l can then be solved. The detailed principle of the method is provided as follows.

Assuming perfect splitting ratio $K_{BS} = 50\%$ for BSs, thus Eq. (S1) can be written as:

$$S_{TBU} = e^{j\phi} \begin{bmatrix} \sin\Delta & \cos\Delta \\ \cos\Delta & -\sin\Delta \end{bmatrix} \tag{S8}$$

in which,

$$\Delta = \frac{\theta_u - \theta_l}{2} = \frac{\vartheta_u - \vartheta_l + \varphi_u(V_u) - \varphi_l(V_l)}{2} \tag{S9}$$

$$\phi = \frac{\pi}{2} + \frac{\theta_u + \theta_l}{2} = \frac{\pi}{2} + \frac{\vartheta_u + \vartheta_l + \varphi_u(V_u) + \varphi_l(V_l)}{2} \tag{S10}$$

As we can see, ϕ contains the information of $\vartheta_u + \vartheta_l$. If we know ϕ, plus that we have already characterize $d\vartheta = \vartheta_l - \vartheta_u$, we can solve for the specific values of ϑ_u and ϑ_l.

We use ϕ^{cross}, corresponding to ϕ when TBU is set to cross state, to calculate $\vartheta_u + \vartheta_l$. When keeping the voltage on the lower PS to be zero, and apply voltage V_u^{cross} on the upper PS to set TBU to cross state, Eq. (S9) and Eq. (S10) can be written as:

$$\Delta^{cross} = \frac{\vartheta_u - \vartheta_l + \varphi_u(V_u^{cross})}{2} \tag{S11}$$

$$\phi^{cross} = \frac{\pi}{2} + \frac{\vartheta_u + \vartheta_l + \varphi_u(V_u^{cross})}{2} \tag{S12}$$

As a matter of fact, obtaining the actual value of ϕ^{cross} is challenging, but the modulus of $2\phi^{cross}$ in radians can be easily obtained. Thus, we reorganize Eq. (S12) as:

$$\vartheta_u + \vartheta_l \ = \ mod(2\phi^{corss}) - \pi + \ 2m\pi \ - \varphi_u(V_u^{cross}) \tag{S13}$$

$m \in \boldsymbol{Z}$. mod means taking the modulus in radians. When TBU is at cross state, Δ^{cross} should satisfy $\Delta^{cross} = n\pi$, substitute into Eq. (S11), we get:

$$\vartheta_u - \vartheta_l + \varphi_u(V_u^{cross}) = 2n\pi \tag{S14}$$

$n \in \boldsymbol{Z}$. Combining Eq. (S13) and Eq. (S14), as well as $d\vartheta = \vartheta_l - \vartheta_u$, we get:

$$\vartheta_u + \vartheta_l = mod(2\phi^{cross}) - d\vartheta - \pi + 2(m-n)\pi \tag{S15}$$

$$\vartheta_u - \vartheta_l = \ -d\vartheta \tag{S16}$$

Thus ϑ_u and ϑ_l can be solved as:

$$\vartheta_u = \frac{1}{2} mod(2\phi^{cross}) - d\vartheta - \frac{\pi}{2} + (m-n)\pi \tag{S17}$$

$$\vartheta_l = \ \vartheta_u + d\vartheta \tag{S18}$$

where m and n are chosen so that ϑ_u are in the range of $[0, 2\pi)$. In fact, ϑ_u has two solutions, thus, we have narrowed down the possible range of ϑ_u to two possibilities.

From Eq. (S17) and Eq. (S18), we can see that, ϑ_u and ϑ_l can be solved once we obtain $mod(2\phi^{cross})$. To obtain the value of $mod(2\phi^{cross})$, we designed a MZI resonator, the upper arm of which is an external phase shifter, and the lower arm is a delay line configured on the waveguide mesh, as shown in Fig. S2(a). With different settings of the waveguide mesh, the delay line light path shapes either “8” or “0”, as shown in Fig. S2(b) and Fig. S2(c).

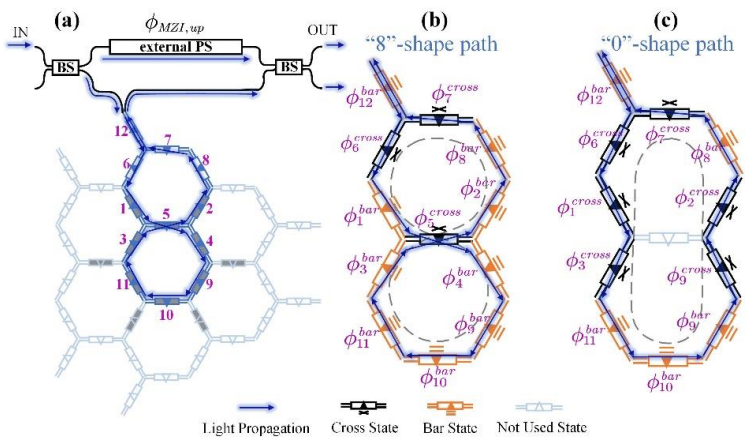

Fig. S2. (a) Constructed MZI resonator, the upper arm of which is an external phase shifter, and the lower arm is a delay line configured on the waveguide mesh. The TBU identifiers are marked with purple numbers. A delay line configured on the waveguide mesh, shaping

like (b) "8", and (c) "0". ϕ_i^{st} is the corresponding ϕ of the i-th TBU in st state ($st \in \{cross, bar\}$); $\phi_{MZI,up}$ represents the phase of the upper arm of the constructed MZI.

$\phi_{"8"}$ and $\phi_{"0"}$, represent the total phase shift of the "8"-shape and "0"-shape delay lines, respectively, can be expressed as sum of ϕ of their constituted TBUs:

$$\phi_{"8"} = \sum_{i=1}^{4} \phi_i^{bar} + \sum_{i=6}^{7} \phi_i^{cross} + \sum_{i=8}^{12} \phi_i^{bar} + 2\phi_5^{cross} \tag{S19}$$

$$\phi_{"0"} = \sum_{i=1}^{4} \phi_i^{cross} + \sum_{i=6}^{7} \phi_i^{cross} + \sum_{i=8}^{12} \phi_i^{bar} \tag{S20}$$

where ϕ_i^{st} is the corresponding ϕ of the i-th TBU in st state ($st \in \{cross, bar\}$). From Eq. (S11) and Eq. (S12) we can deduce that $\phi_i^{bar} - \phi_i^{cross} = \pi/2$.

Subtract Eq. (S19) and Eq. (S20), and substitute $\phi_i^{bar} - \phi_i^{cross} = \pi/2$, we get:

$$mod(2\phi_5^{cross}) = mod(\phi_{"8"} - \phi_{"0"}) \tag{S21}$$

It can be observed in Eq. (S21) that we can have $mod(2\phi_5^{cross})$ once we know $\phi_{"8"}$ and $\phi_{"0"}$. However, obtaining the absolute values of $\phi_{"8"}$ and $\phi_{"0"}$ is still difficult. But we can have $mod(\phi_{"8"} - \phi_{"0"})$ by acquiring its equal value $mod(\delta\phi_{"8"} - \delta\phi_{"0"})$, $\delta\phi_{"8"}$ and $\delta\phi_{"0"}$ are the phase differences between the two arms of the constructed MZI (Fig. S2(a)), respectively corresponding to using the "8"-shape and the "0"-shape delay line as its lower arm. $\phi_{"8"}$ and $\phi_{"0"}$ are related to $\delta\phi_{"8"}$ and $\delta\phi_{"0"}$ as in Eq. (S22) and Eq. (S23).

$$\delta\phi_{"8"} = mod(\phi_{"8"} - \phi_{MZI,up}) \tag{S22}$$

$$\delta\phi_{"0"} = mod(\phi_{"0"} - \phi_{MZI,up}) \tag{S23}$$

$\phi_{MZI,up}$ represents the phase of the upper arm of the constructed MZI.

Subtracting Eq. (S22) and Eq. (S23), $mod(\phi_{"8"} - \phi_{"0"})$ can be expressed as:

$$mod(\phi_{"8"} - \phi_{"0"}) = mod(\delta\phi_{"8"} - \delta\phi_{"0"}) \tag{S24}$$

$\delta\phi_{"8"}$ and $\delta\phi_{"0"}$ can be easily obtained. Sweeping the voltage applied on the external PS of the constructed MZI, while testing the output power of its output port, $\delta\phi_{"8"}$ and $\delta\phi_{"0"}$ can be easily extracted from the power-voltage curve. Thus, according to Eq. (S24) and Eq. (S21), we now have $mod(2\phi^{cross})$, ϑ_u and of ϑ_l TBU 5 can then be solved using Eq. (S17) and Eq. (S18). Similarly, for all TBUs that are not on the edge, painted gray as in Fig. S2(a), we can perform the above operation to solve their ϑ_u and ϑ_l.

SectionS6: characterizing utilizing optimization

In the final step of our characterization procedure, we utilize an optimization method, taking $\boldsymbol{C_p}$ as the variables for optimization, find the optimal value of $\boldsymbol{C_p}$ based on the criterion of making the mesh behavior closely resembles that of the fabricated one.

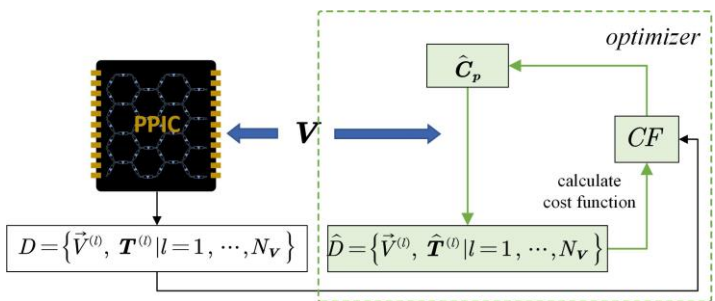

Fig. S3. Optimization system diagram.

The workflow of the optimization is illustrated in Fig. S3. Generate N_V sets of random voltage settings $\vec{V}$ (vector with n_{PS} entries, n_{PS} is the number of PS in the mesh). For each voltage setting $\vec{V}^{(l)}$, apply it on the fabricated mesh, and measure its corresponding transmission matrix $\boldsymbol{T}^{(l)}$, where superscript l denote the index for a particular voltage setting, forming a dataset $D=\left\{\vec{V}^{(l)},\ \boldsymbol{T}^{(l)}|l=1,\cdots,N_{\boldsymbol{V}}\right\}$. Such dataset describes the behavior of the fabricated mesh under various voltage settings.

For each voltage setting $\vec{V}^{(l)}$, we calculate the corresponding transmission matrix $\hat{\boldsymbol{T}}^{(l)}$ would manifest under current predicted $\hat{\boldsymbol{C}}_{\boldsymbol{p}}$, forming a dataset $\hat{D}=\left\{\vec{V}^{(l)},\ \hat{\boldsymbol{T}}^{(l)}|l=1,\cdots,N_{\boldsymbol{V}}\right\}$. Such dataset describes the behavior of a virtual mesh with current predicted $\hat{\boldsymbol{C}}_{\boldsymbol{p}}$.

We define a cost, as in Eq.(S25), to evaluate the difference between the behaviors of the virtual mesh featuring the predicted $\hat{\boldsymbol{C}}_{\boldsymbol{p}}$ and the actual fabricated mesh.

$$cost=\sum_{l=1}^{N_V}\sum_{k=1}^{N_f}\sum_{j=1}^{N_{port}}\sum_{i=1}^{N_{port}}\left(\sqrt{\hat{t}_{i,j,k}^{(l)}}-\sqrt{t_{i,j,k}^{(l)}}\right)^2 \tag{S25}$$

where $t_{i,j,k}^{(l)}$ and $\hat{t}_{i,j,k}^{(l)}$ are the weights of $\boldsymbol{T}^{(l)}$ and $\hat{\boldsymbol{T}}^{(l)}$. The superscript l denote the index for a particular voltage setting, $N_{\boldsymbol{V}}$ is the total number of voltage settings. The subscripts i, j and k represent the row, column, and spectral slice index, respectively. N_{port} is the number of ports of the waveguide mesh, determining the row and column count of $\boldsymbol{T}^{(l)}$, N_f refers to the spectral slice count of $\boldsymbol{T}^{(l)}$, i.e., the number of frequency points at which the mesh transmission matrices are measured.

As shown in Fig. S3, with each generation predicted $\hat{\boldsymbol{C}}_{\boldsymbol{p}}$, calculate cost using Eq.(S25) to determine whether the current predicted $\hat{\boldsymbol{C}}_{\boldsymbol{p}}$ is good or bad. Adjust the next generation $\hat{\boldsymbol{C}}_{\boldsymbol{p}}$ accordingly. The optimal value of $\boldsymbol{C}_{\boldsymbol{p}}$ will be found when cost is minimized, and $\boldsymbol{C}_{\boldsymbol{p}}$ is thus characterized.

This is the final step of our characterization procedure, after the previous three steps, some parameters in $\boldsymbol{C}_{\boldsymbol{p}}$ have already been characterized or had their potential value ranges narrowed down. Thus, in this step, these parameters are set to their pre-characterized values or have their searching range reduced, which helps to reduce the parameter space for optimization.

Given that we are dealing with an optimization problem characterized by continuous space, high dimensionality, non-convexity, and a solution space featuring numerous hills and valleys, we choose to employ Particle Swarm Optimization (PSO) algorithm [10, 11] to conduct optimization, which is well-suited for addressing such challenges.

Section S7: Random $\boldsymbol{C}_{\boldsymbol{p}}$ generation

In validation of our proposed method, imperfection parameters in $\boldsymbol{C}_{\boldsymbol{p}}=\left\{\vec{\vartheta}_u,\vec{\vartheta}_l,\vec{\varphi}_u(V),\vec{\varphi}_l(V),\vec{\alpha},\vec{\beta}\right\}$ are randomly generated, and then retrieved using the proposed the characterization method. When randomly generating $\boldsymbol{C}_{\boldsymbol{p}}$, probabilistic models that different parameters in $\boldsymbol{C}_{\boldsymbol{p}}$ follow and their corresponding model parameters are summarized in Table S2.

Table S2: Probabilistic models that different parameters in $\boldsymbol{C}_{\boldsymbol{p}}$ follow and their corresponding model parameters.

$\boldsymbol{C}_{\boldsymbol{p}}$	ϑ_u,ϑ_l	$\vartheta_u,\vartheta_l\sim U[0,2\pi)$ [a]
	α,β	$K_{BS,\alpha},\ K_{BS,\beta}\sim N(50\%,2.5\%^2)$ [b][12] Beam splitter splitting ratios $K_{BS,\alpha},K_{BS,\beta}$ are first generated, then transfer to α,β using $K_{BS,\alpha}=\cos^2(\alpha),\ K_{BS,\beta}=\cos^2(\beta)$

	$\varphi_u(V), \varphi_l(V)$	Assuming $\varphi_u(V)$ and $\varphi_l(V)$ are proportional to square of applied voltage V, i.e., $\varphi_u(V) = k_{h,u} \cdot V^2, \varphi_l(V) = k_{h,l} \cdot V^2$, where $k_{h,u}, k_{h,l} \sim N(\pi/5^2,\ (\pi/4.7^2 - \pi/5^2)^2)$.
	n_g	$n_g \sim N(4.3,\ 0.01^2)$ [8, 9]

[a] $U[a,b)$ indicates a uniform distribution within $[a,b)$.

[b] $N(\mu,\ \sigma^2)$ indicates a normal distribution with mean μ and standard deviation σ.

Section S8: Stability and robustness analysis

We carried out our method under various scenarios considering different beam splitter splitting ratio variance, inaccurate measurements of mesh, inaccurate TBU insertion loss characterization, and inaccurate voltage control to analyze the stability and robustness of our method.

We first evaluate the stability and robustness of our characterization method under different beam splitter splitting ratio (K_{BS}) variation scenario. Assuming K_{BS} follows a Gaussian distribution with a standard deviation of σ_{BS} of 2.5%, 5%, 7.5% (a typical wafer-level K_{BS} variation is $\sigma_{BS} = 2\%$ [12]), respectively. Under each scenario, conduct 10 experiments by randomly generating 10 sets of different $\boldsymbol{C_p}$ values, and use our characterization method to respectively retrieve them. The resulting RMSEs are depicted in Fig. S4(a). Notably, our characterization method proves effective even under extreme splitting ratios, which is way worse than the typical wafer-level variance (beam splitter variation as small as 2% [12]), and is very stable, cross all experimental runs for various splitting ratio conditions, the RMSEs stay below 1.0 dB.

Secondly, we consider the measurement inaccuracy when measuring the transmission matrices $\boldsymbol{T}$ of the fabricated mesh. The measurement inaccuracy is simulated by introducing random fluctuations to the accurate $\boldsymbol{T}$, where the magnitude of fluctuations follows a Gaussian distribution with standard deviation $\sigma_{\boldsymbol{T}}$ of 1%, 2%, and 3%, respectively. Under each scenario, conducted 10 experiments. The resulting RMSEs are depicted in Fig. S4(b). The performance of the characterization method indeed decreases with the rise in measurement inaccuracy, but the decrease is small. Our method maintains good performance even under vary bad conditions.

Thirdly, we consider the inaccuracies in TBU insertion loss (IL) characterization. The IL of TBUs can be characterized using method outlined in [2], when characterize correctly, it does not affect the proposed method performance in characterization of $\boldsymbol{C_p}$. However, acknowledging potential inaccuracies in IL characterization, as described in [2], where the average IL characterization error is 0.18 dB, we also tested the proposed method in presence of inaccurate IL characterization. Assume the actual ILs of TBUs are uniformly distributed within the range of (0.5, 0.7) dB [13], and the IL characterization inaccuracy is simulated by introducing errors to the actual IL, where the errors follow a uniform distribution within ranges of (-0.2,0.2) dB, (-0.4,0.4) dB, and (-0.6, 0.6) dB, corresponding to average absolute error $\mu_{IL,e}$ of 0.1dB, 0.2dB and 0.3dB, respectively. Under each scenario, conduct 10 experiments. The resulting RMSEs are depicted in Fig. S4(c). It appears that the RMSE indeed increases with higher IL characterization errors. However, the main reason is that the incorrect IL itself leads to errors in the predicted transmission matrix $\boldsymbol{T}$ (since the calculation of $\boldsymbol{T}$ involves IL, higher IL errors naturally result in less accurate $\boldsymbol{T}$, leading to a larger RMSE), rather than the IL characterization error hindering the proposed method from correctly characterizing $\boldsymbol{C_p}$.

Lastly, we evaluated the impact of inaccurate voltage control. Assuming Digital-to-Analog Converter (DAC) bit resolution of 4 bits, 6 bits, and 8 bits, respectively, under each condition, our characterization method was tested. The resulting RMSEs are depicted in Fig. S4(d). In addition to these conditions, we also considered an ideal scenario where any target voltage can be accurately applied to the phase shifter without being affected by DAC accuracy, labeled as “ideal” on the x-axis of Fig. S4(d). From Fig. S4(d), we can see that the DAC control accuracy indeed affects the performance of the proposed method, but when DAC resolution reached 6 bits, the performance of our method is close to that of the ideal control condition. Therefore, although higher-bit is preferred, 6-bit DAC is sufficient. The main impact brought by inaccurate

voltage control is that it prevents the TBU from achieving accurate cross/bar states, such poor extinction ratios would result in optical leakage to other TBUs outside the intended optical path, causing uncontrollable interference, which leads the phase estimated by step 3 of our characterization procedure to deviate from the actual value, resulting in poor characterization performance. So, if the hardware DAC bits are indeed quite low, the impact of it can be mitigated by expanding the phase search range during the optimization process.

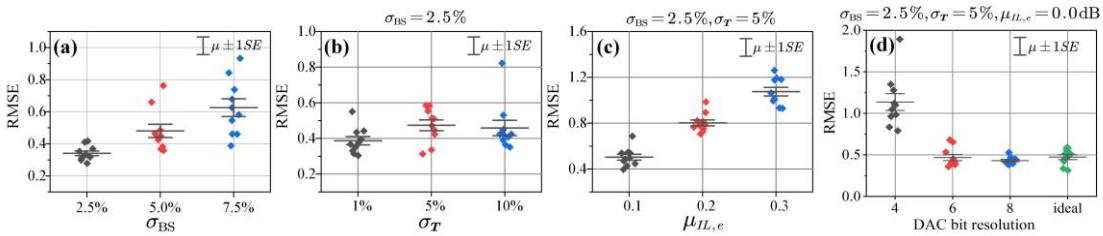

Fig. S4. (a) RMSEs obtained under various σ_{BS} conditions. (b) RMSEs obtained under different levels of measurement inaccuracy of $\boldsymbol{T}$, under the condition of $\sigma_{BS} = 2.5\%$. (c) RMSEs obtained under different IL characterization error levels, under the condition of $\sigma_{BS} = 2.5\%$, $\sigma_{\boldsymbol{T}} = 5\%$. (d) RMSEs obtained under different DAC bit resolutions and an ideal voltage control scenario, under the condition of $\sigma_{BS} = 2.5\%$, $\sigma_{\boldsymbol{T}} = 5\%$ and $\mu_{IL,e} = 0.0\text{dB}$. In each scenario, conduct 10 experiments. Plot all obtained RMSE values, and mark the Mean plus Standard Deviation ($\mu + 1SE$, μ indicates mean value, SE indicates standard deviation).

Section S9: Verification of the proposed method by implementing infinite impulse response (IIR) filter applications applications

We also demonstrated the implementations of infinite impulse response (IIR) filter applications, including an optical ring resonator (ORR), a triple ORR coupled resonator waveguide (CROW) filter and a double ORR ring-loaded MZI, as shown in Fig. S5(a)(b)(c).

We first configured these IIR filters using voltages chosen based on an ideal assumption, as a result, the normalized spectral responses of an ideal mesh, the actual mesh, and the characterized mesh were plotted in Fig. S5(d)(e)(f)). As we can see, the spectral responses of the actual mesh deviate significantly from those of the ideal mesh, indicating that, configuring the mesh based on an ideal assumption would lead the actual defective mesh to deviate from the targeted functionality. Also, the spectral responses of the characterized mesh align with those of the actual mesh, validating the effectiveness of our characterization method in accurately predicting the actual mesh behavior. Then we configure the filters using voltages chosen based on the characterized $\hat{\boldsymbol{C}}_{\boldsymbol{p}}$ ($\hat{\boldsymbol{C}}_{\boldsymbol{p}}$ used here is characterized under the circumstance of $\sigma_{BS} = 2.5\%$, $\sigma_{\boldsymbol{T}} = 5\%$, exhibiting an RMSE of 0.58 dB), the normalized spectral responses of the actual and the characterized mesh are plotted in Fig. S5(g)(h)(i), as we can see, they are in perfect alignment, and both have achieved the intended functionality, highlighting that our characterization method can ensure configurations suitable for the actual mesh.

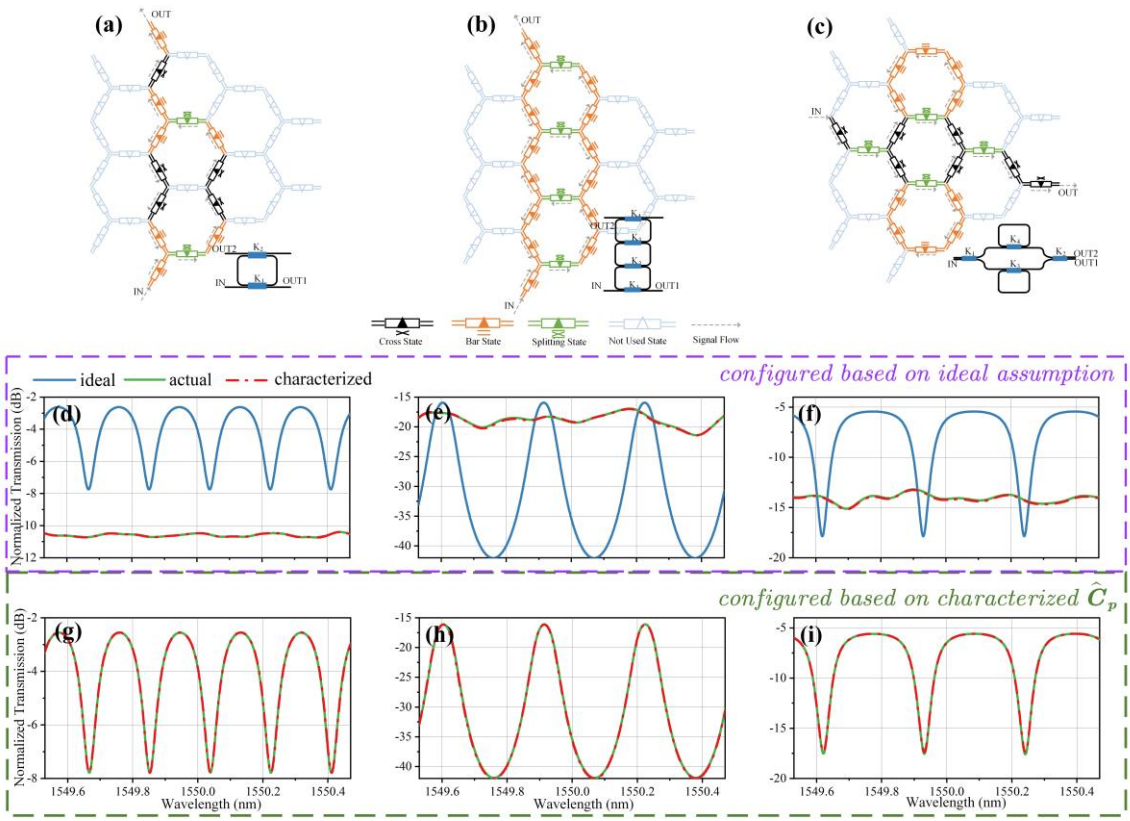

Fig. S5. IIR applications: Circuit layout diagram, waveguide mesh arrangements for three different IIR applications. (a) an optical ring resonator (ORR). (b) a triple ORR coupled resonator waveguide (CROW) filter. (c) a double ORR ring-loaded MZI. (d)(e)(f) spectral responses of the ideal, actual, and characterized mesh, when configured with voltages chosen based on ideal assumption, respectively, for the three applications. (g)(h)(i) spectral response of the actual and characterized mesh, when configured with voltages chosen base on characterized $\hat{C}_p$, respectively, for the three applications.

References

1. D. Perez Lopez, A. Lopez Hernandez, P. DasMahapatra, J. Capmany, and Ieee, "Field-Programmable Photonic Array for multipurpose microwave photonic applications," in *IEEE International Topical Meeting on Microwave Photonics (MWP)*, (2019), 147.
2. D. Pérez-López, A. López, P. DasMahapatra, and J. Capmany, "Multipurpose self-configuration of programmable photonic circuits: supplementary material," Nature Communications **11,** 6359 (2020).
3. A. López-Hernández, M. Gutiérrez-Zubillaga, and D. Pérez-López, "Automatic Self-calibration of Programmable Photonic Processors," in *IEEE Photonics Conference (IPC)*, (2022), 1.
4. E. Sánchez, A. López, and D. Pérez-López, "Simulation of Highly Coupled Programmable Photonic Circuits," Journal of Lightwave Technology **40,** 6423 (2022).
5. Z. Gao, X. Chen, Z. Zhang, U. Chakraborty, W. Bogaerts, and D. S. Boning, "Automatic Realization of Light Processing Functions for Programmable Photonics," in *IEEE Photonics Conference (IPC)*, (2022), 1.
6. Z. Gao, X. Chen, Z. Zhang, U. Chakraborty, W. Bogaerts, and D. S. Boning, "Automatic synthesis of light-processing functions for programmable photonics: theory and realization," Photonics Research **11,** 643 (2023).
7. D. Pérez and J. Capmany, "Scalable analysis for arbitrary photonic integrated waveguide meshes," Optica **6,** 19 (2019).
8. Y. Xing, J. Dong, S. Dwivedi, U. Khan, and W. Bogaerts, "Accurate extraction of fabricated geometry using optical measurement," Photonics Research **6,** 1008 (2018).
9. Y. Xing, J. Dong, U. Khan, and W. Bogaerts, "Capturing the Effects of Spatial Process Variations in Silicon Photonic Circuits," ACS Photonics **10,** 928 (2023).
10. R. Eberhart and J. Kennedy, "A new optimizer using particle swarm theory," in *MHS'95. Proceedings of the Sixth International Symposium on Micro Machine and Human Science*, (1995), 39.
11. M. Clerc and J. Kennedy, "The particle swarm - explosion, stability, and convergence in a multidimensional complex space," IEEE Transactions on Evolutionary Computation **6,** 58 (2002).

12. J. C. Mikkelsen, W. D. Sacher, and J. K. S. Poon, "Dimensional variation tolerant silicon-on-insulator directional couplers," Optics Express **22,** 3145 (2014).
13. D. Pérez, I. Gasulla, L. Crudgington, D. J. Thomson, A. Z. Khokhar, K. Li, W. Cao, G. Z. Mashanovich, and J. Capmany, "Multipurpose silicon photonics signal processor core: supplementary material," Nature Communications **8,** 636 (2017).