Infinitely Scalable Multiport Interferometers

Ryan Hamerly\textsuperscript{1,2}, Saumil Bandyopadhyay\textsuperscript{1}, Dirk Englund\textsuperscript{1}

September 14, 2021

\textsuperscript{1} Research Laboratory of Electronics, MIT, 50 Vassar Street, Cambridge, MA 02139, USA
\textsuperscript{2} NTT Research Inc., Physics and Informatics Laboratories, 940 Stewart Drive, Sunnyvale, CA 94085, USA

Abstract—Component errors limit the scaling of multiport interferometers based on MZI meshes. These errors arise because imperfect MZIs cannot be perfectly programmed to the cross state. Here, we introduce two modified mesh architectures that overcome this limitation: (1) a 3-splitter MZI for generic errors, and (2) a broadband MZI+Crossing design for correlated errors. Because these designs allow for perfect realization of the cross state, the matrix fidelity no longer decreases with mesh size, allowing scaling to arbitrarily large meshes. The proposed architectures support progressive self-configuration, are more compact than previous MZI-doubling schemes, and do not require additional phase shifters. This eliminates a major obstacle to the development of very-large-scale linear photonic circuits.

1 Introduction

Large-scale programmable photonic circuits are opening up radical new possibilities for optics. Of central importance in many devices is the universal multiport interferometer, which functions as an $N \times N$ programmable linear circuit (Fig. 1(a-b)). This device, usually constructed from a dense mesh of Mach-Zehnder interferometers (MZIs) [1, 2], is widely employed in applications ranging from spatially multiplexed optical communications to machine learning and quantum computing [3–7].

Sadly, component errors (Fig. 1(c)) are a critical factor limiting the size of such circuits. Since the circuit depth of MZI meshes scales as $O(N)$, the effect of errors grows with mesh size, meaning that, in practice, even modestly sized circuits cannot be programmed to high accuracy. Motivated by this challenge, a large body of recent work has focused on “correcting” hardware errors by global optimization [8–10], self-configuration [11–17], or local correction [18, 19]. For conventional MZI meshes, correction reduces errors by a quadratic factor [16, 18]; however, the effect of errors still grows with mesh size and poses a fundamental limit to the scaling of these circuits.

To overcome this limit, various alternative mesh architectures have been proposed. Non-compact structures such as binary trees avoid the extreme splitting-ratio requirements [20, 21], but suffer from large chip area and the need for many crossings. A complementary approach is to stick to conventional geometries [11, 12], but insert redundant MZIs to realize the full range of splitting ratios even in imperfect hardware [22, 23]. This solves the scaling problem, but at the cost of a $\sim 2\times$ increase in the number of splitters and phase shifters. The resulting effect on chip area and pad count makes this option unappealing.

In this paper, we propose two new mesh architectures that achieve the same perfect scaling without significant added complexity: a 3-splitter MZI that corrects all hardware errors (Fig. 1(d)) and an MZI+Crossing design that only corrects correlated errors, but has the added advantage of broader bandwidth (Fig. 1(e)). These designs take up significantly less chip area than the “perfect” redundant MZIs [22, 23], and do not require additional phase shifters. Moreover, the proposed architectures support progressive self-configuration [16, 17], allowing for error correction even when the hardware errors are unknown. This work will enable the development of freely scalable, broadband, and compact linear photonic circuits.

This paper is structured as follows: in Sec. 2 we introduce the formalism of error correction in MZI meshes, focusing on the self-configuration approach. Splitting ratios are visualized as points on the Riemann sphere, where hardware imperfections lead to forbidden regions around the poles (bar- and cross-state), where the probability density is at a maximum. To avoid this unfortunate coincidence, our architectures “rotate” the Riemann sphere to move the forbidden regions away from this peak, so that a larger fraction of MZIs are perfectly realized. Sec. 3 in-

![Figure 1](https://example.com/figure1.png)

Figure 1: (a) $6 \times 6$ triangular mesh, composed of (b) a phase screen $\psi$ and tunable MZI crossings $\theta, \phi$. (c) Fabrication imperfections lead to splitting-ratio errors $\alpha, \beta$. (d) Alternative 3-splitter MZI. (e) MZI+Crossing design.
roduces the 3-splitter MZI, which can correct arbitrary errors by rotating the forbidden regions to the equator. Sec. 4 introduces the MZI+Crossing, which flips the poles of the Riemann sphere. While this design is only robust against correlated errors, it has the added advantage of broader bandwidth. For both architectures, we compare the matrix fidelity to the standard MZI to demonstrate the scaling advantage of both schemes. Finally, Sec. 5 states our conclusions.

2 Error Correction Formalism

To correctly configure an MZI mesh in the presence of errors, one uses a nulling method based on physical measurements [16, 17]. Fig. 2(a) illustrates the case of the triangular mesh [1], where the procedure is more straightforward. The transfer matrix for this system is a product of a phase screen and a sequence of $2 \times 2$ unitaries:

$$U = D \left( \left( T_{N-1,1} \right) \ldots \left( T_{2, N-2} \ldots T_{21} \right) (T_{1, N-1} \ldots T_{11}) \right)$$

where $T_{mn}$ is the $n^{th}$ MZI of the $m^{th}$ rising diagonal. We configure the mesh by building up matrix $W$ in a sequence of steps designed to diagonalize a target matrix $X = UW^\dagger$. In each step, we add one crossing to $W$, performing the update $W \rightarrow T_{mn}W$, which right-multiplies the target matrix $X \rightarrow XT_{mn}^\dagger$ (Fig. 2(b)). The phase shifts $(\theta, \phi)$ are chosen to zero a particular matrix element $v \rightarrow 0$ (green in figure), satisfying the equation (indices $m, n$ suppressed for notational simplicity):

$$[u \ v]^\dagger \left[ \begin{array}{c} T_{mn} \end{array} \right] = [ \ast \ 0] \iff T_{11}/T_{12} = u/v$$

(2)

This is illustrated in Fig. 2(c). Nulling physically corresponds to injecting $w_j^\dagger$ (the $j^{th}$ column of $W^\dagger$) and zeroing the power at the $i^{th}$ output [17]. If all nulling steps are performed exactly, the mesh will perfectly realize the target matrix $U$.

Figure 2: Nulling method of self configuration. (a) Configuring MZI $T_{mn}$ updates matrix $W$. (b) Corresponding nulling update to $X = UW^\dagger$, which is (c) equivalent to zeroing an output of $T_{mn}$ given a fixed input.

Mathematically, nulling corresponds to matching the complex splitting ratio $s \equiv T_{11}/T_{12} = -(T_{21}/T_{22})^*$ to a target value $\hat{s} \equiv u/v$. This is not always possible in practice, as the range of splitting ratios is constrained by hardware imperfections (Fig. 2(c)):

$$\tan |\alpha + \beta| \leq |s| \leq \cot |\alpha - \beta|$$

(3)

Here $\alpha, \beta$ are the splitting angle errors for the splitters in a standard MZI. These imperfections lead to forbidden regions (Fig. 3(a)) for small and large $s$, where nulling cannot be achieved perfectly. It is also instructive to view this chart on the Riemann sphere, which shows that these forbidden regions are centered around the poles (Fig. 3(b)).

If nulling cannot be achieved perfectly, the “zero” region of matrix $X$ is left with a residual of magnitude:

$$r = |T_{11}v - T_{12}u| = \sqrt{|u|^2 + |v|^2 d(s, \hat{s})^2/2}$$

(4)

where $\hat{s}$ is the target splitting ratio, $s$ is the closest physically realizable value, and $d(s, \hat{s})$ is the Euclidean distance on the Riemann sphere with stereographic projection $s = (x + iy)/(1 + z)$. As a fidelity metric, we consider the normalized matrix error $E_c = \langle \|\Delta U\|_{\text{rms}}\rangle / \sqrt{N}$, which is approximately the quadrature sum of these residuals:

$$(E_c)^2 = \frac{\langle \|\Delta U\|^2 \rangle}{N} = \frac{2}{N} \sum_{mn} r_{mn}^2$$

(5)

Here, $\langle \ldots \rangle$ refers to the ensemble average over Haar-distributed unitaries $U [24, 25]$. Eq. (5) is highly sensitive to this distribution, since $r_{mn}$ is nonzero only within

Figure 3: (a) Allowed range of $s = T_{11}/T_{12}$; regions near $s = 0$ and $s = \infty$ are forbidden due to imperfections. Contours are lines of constant $(\theta, \phi)$, with $\alpha = 0.23, \beta = 0.07$. (b) Equivalent Riemann sphere plot. (c) Probability density $P(s)$ as a function of mesh size.
the forbidden regions. It is well known that, for large meshes, this distribution clusters tightly near the cross state \( s \neq 10,26 \), or the upper pole of the Riemann sphere (Fig. 3(c)). The distribution depends on the MZI’s location in the mesh; for a given \( T_{mn} \) it takes the form \[ P_{mn}(s) = \frac{n}{4\pi} \left( \frac{z + 1}{2} \right)^{-1} \] (6)
This is uniform for the lowest row of crossings, and becomes increasingly concentrated as one approaches the triangle’s apex.

We calculate the mean residual \( \langle r^2 \rangle \) by averaging Eq. (4) over the distribution \( P(s) \). This is simplified in the case of small hardware errors, because the forbidden region is correspondingly small and where we can assume \( P(s) \) is approximately constant:

\[ \langle r^2_{mn} \rangle = \left( \langle |u|^2 + |v|^2 \rangle \right) P_{mn}(s_0) \frac{\pi R_0^4}{24} \] (7)
where \( s_0 \) and \( R_0 \) are the center and radius of the forbidden region. Consider the forbidden region near the upper pole (the lower region plays a negligible role for large meshes), where \( s_0 = 0 \) and \( R_0 = 2|\alpha + \beta| \). The residual is also proportional to the quantity \( q_{mn} = \langle |u|^2 + |v|^2 \rangle \) (green squares in Fig. 2(b)). Following the Gaussian elimination procedure of a Haar matrix, this evaluates to \( q_{mn} = (n + 1)/(N + 1 - m) \), and we have:

\[ \langle r^2_{mn} \rangle \rightarrow \frac{n(n + 1)}{N + 1 - m} \frac{\langle |\alpha + | |\beta|\rangle^4}{6} \] (8)
In the large-\( N \) limit, we can replace the discrete sum in Eq. (4) with an integral. Finally, assuming an uncorrelated Gaussian perturbation model with \( \langle \alpha \rangle_{rms} = \langle |\beta|\rangle_{rms} = \sigma \), the resulting matrix error is:

\[ \mathcal{E}_c = \left( \frac{N^2}{24} \langle |\alpha + | |\beta|\rangle^4 \right)^{1/2} \rightarrow \frac{2}{3} N \sigma^2 \] (9)
In contrast, if the mesh were straightforwardly programmed without taking any account of the imperfections (“uncorrected” error), the error under an uncorrelated perturbation model would be \( \mathcal{E}_0 = \sqrt{2N} \sigma \). Self-configuration therefore leads to a quadratic suppression of errors: \( \mathcal{E}_c = (\mathcal{E}_0)^2/3 \). However, \( \mathcal{E}_c \) still increases with mesh size, so fabrication imperfections will still set a limit on the scaling of meshes based on this architecture.

Some matrices can be exactly realized even with an imperfect mesh. The coverage is the fraction of such unitaries under the Haar measure. This is computed by summing up the probabilities that individual MZIs fall within the forbidden region:

\[ \text{cov}(N) = \exp \left( - \sum_{mn} P_{mn}(0) \pi R_0^4 \right) = e^{-N^3 \sigma^2/3} \] (10)
In practice, even moderately sized meshes will have vanishingly small coverage, so the error metric \( \mathcal{E}_c \) is usually more relevant.

3 3-Splitter MZI

The main challenge to error correction in multiport interferometers is that the forbidden regions overlap with the peak of the probability distribution (cross state \( s = 0 \)). Adding redundant components (MZI doubling) solves this problem by eliminating the forbidden regions altogether \[ 22, 23 \], but at the cost of added optical and electrical complexity. Here, we take the alternative approach of displacing the forbidden regions away from the cross state rather than eliminating them altogether. This can be performed by placing a third splitter at the input of the MZI, as shown in Fig. 4(a). The extra splitter performs a Möbius transformation on \( s \)

\[ s \rightarrow s + i \tan(\eta) \frac{1 + i s \tan(\eta)}{1 + i s \tan(\eta)} \] (11)
which for a 50:50 splitting ratio (\( \eta = \pi/4 \)), maps the bar and cross states to \( s = \pm i \) (Fig. 4(b)). This can be visualized as a 90° rotation on the Riemann sphere, which pushes the forbidden regions to the equator, while the probability density is still concentrated at the poles (Fig. 4(c)).

The 3-splitter MZI can realize the full range of (absolute value) splitting ratios \( |s| \in [0, \infty) \), and for certain parameter choices, this ratio is wavelength-independent \[ 27 \]. However, the presence of forbidden regions means that the relative phase of this splitter cannot be fully controlled; which means that errors can still occur when programming the mesh (unlike the “perfect” MZIs of Refs. 22, 23, which cure this defect with redundant phase shifters). However, as we will show next, with the 3-splitter structure, MZIs fall into the forbidden regions only rarely; consequently, the error metric of the 3-splitter MZI is much lower than the standard MZI, and perhaps more strikingly, does not increase with mesh size \( N \).

Figure 4: (a) 3-splitter MZI design. (b) Splitter Möbius transformation on \( s \in \mathbb{C} \), which pushes the forbidden regions away from \( s = \{0, \infty\} \). (c) Equivalent Riemann sphere rotation.
Applying Eq. (7), the mean residual left by MZI $T_{mn}$ is:

$$\langle r_{mn}^2 \rangle = 2 \times \frac{n+1}{N+1-m} \frac{n}{q_{mn}} \frac{\pi}{24} \frac{(192\sigma^4)}{(r_0^4)} \tag{12}$$

The factor of two in Eq. (12) arises because both forbidden regions contribute equally. Summing over all MZIs,

$$E_c = \frac{(128\sigma^4)}{N} \left[ \sum_{n=1}^{N} \frac{1}{n} - \frac{5}{4} - \log(2) \right] \right]^{1/2}$$

$$\approx 8\sigma^2 \left[ 2\log(N) + \gamma_e - \frac{5}{4} - \log(2) \right]^{1/2} \tag{13}$$

where $\gamma_e \approx 0.5772$ is the Euler-Mascheroni constant. To test Eqs. (9), we numerically simulate self-configuration on imperfect meshes using the MESHES package [28]; the agreement with theory is shown in Fig. 5(b).

For most mesh sizes, the matrix error $E_c$ is 1–2 orders of magnitude smaller for the 3-MZI design. Remarkably, the factor of two in Eq. (12) arises because both forbidden regions contribute equally. Summing over all MZIs, the threshold scales as $E_c \propto \sqrt{\log(N)/N}$. In the asymptotic limit $N \to \infty$, matrices can be programmed perfectly.

This non-intuitive effect arises from the fact that, under the Haar measure, only a small fraction of MZIs have significant probability density near $s = \pm i$, where the forbidden regions are centered. This probability decreases exponentially with the distance from the triangle’s base (or from the edge in the case of the rectangular mesh) due to the $2^{-n}$ dependence in Eq. (6). Therefore, although the mesh has $N(N - 1)/2$ MZIs, only $O(N)$ contribute significantly to the matrix error under self-configuration.

A naïve estimate assuming uncorrelated errors would give $||\Delta U|| \propto \sqrt{N}\sigma^2$, which would lead to a constant $E_c$. However, during the self-configuration process, subsequent MZIs can partially correct for errors in earlier MZIs that cannot be properly configured. Only the residual terms Eq. (4) remain, as these cannot be cancelled out by subsequent Givens rotations. The end result is to reduce the error of each MZI by a factor $q_{mn} = (|u|^2 + |v|^2)$, which is often $\ll 1$. Properly accounting for this factor explains the further reduction to $E_c \propto \sqrt{\log(N)/N}$ observed in Eq. (13).

Another advantage of the 3-splitter MZI is that the threshold for perfect error correction is higher. One obtains this threshold is found by computing the coverage; analogous to Eq. (10), we find:

$$\text{cov}(N) = \exp\left(-2\sum_{mn} P_{mn}(\pm i) \pi R_0^2\right) = e^{-16N\sigma^2} \tag{14}$$

The threshold scales as $\sigma_{th} \propto N^{-1/2}$, as opposed to the $N^{-3/2}$ scaling observed for a standard MZI. Consequently, perfect error correction is available under a much larger range of conditions (Fig. 6).

4 MZI+Crossing

Sec. 3 considered the case of generic, uncorrelated component errors, for which the 3-splitter MZI is well-suited. However, since the correlation lengths of process variations tend to be larger than a single MZI, errors are often correlated in practice. This is especially true for broadband couplers based on multimode interference (MMI) [29, 30], subwavelength gratings [31, 32], and asymmetric designs [33, 34], all of which are highly dependent on
Figure 7: (a) MZI+Crossing architecture. (b) Effect of the crossing is to flip the s = 0 and s = ∞ forbidden regions. For correlated errors, the forbidden region around s = 0 disappears. (c) Riemann sphere projection.

the device geometry, which can vary slightly from run to run. Moreover, even with perfect 50:50 couplers, the splitting ratios are still wavelength-dependent. Operating the mesh away from its design wavelength will lead to correlated device errors, so sensitivity to these errors is closely tied to the operational bandwidth of the device.

Consider the case of a constant offset µ for all splitting ratios: α = β = µ. In a standard MZI, the forbidden region near s = ∞ disappears since |α − β| = 0, while the region near s = 0 (the peak of the probability distribution) remains in place (Fig. 3). This is consistent with the common observation that the extinction ratio in an MZI is much higher in the cross port than in the bar port. The optimal error reduction strategy, illustrated in Fig. 7(a), has been previously proposed in the context of broadband optical switching: place a waveguide crossing before the MZI [35]. The added crossing performs the Mőbius transformation s → 1/s, rotating the Riemann sphere by 180° to move the forbidden region to the minimum of the probability distribution (Fig. 7(b-c)).

As before, we use the residual formula Eq. (7) to calculate the matrix error. In this case, there is only one forbidden region, centered at s₀ = ∞, with R₀ = 4µ. Only the MZIs in the bottom row contribute to the sum, because the probability distribution Eq. (6) vanishes at s = ∞ for the upper rows. We find:

\[ E_c = 4\mu^2 \left[ \frac{2 \log(N) + \gamma - 1}{N} \right]^{1/2} \]

This is plotted in Fig. 8. Like the 3-MZI design, this metric scales as \( E_c \propto \sqrt{\log(N)/N\mu^2} \), in contrast to the trend \( E_c = (4/3\sqrt{2})N\mu^2 \) observed for the standard MZI under correlated errors. The coverage also increases:

\[ \text{cov}(N) = \begin{cases} e^{-(2/3)N\mu^2} & \text{(MZI)} \\ e^{-4N\mu^2} & \text{(MZI+X)} \end{cases} \]

As an added bonus, the MZI+Crossing design also reduces the effect of errors in the absence of correction. To see how, consider the transfer matrix of the standard MZI:

\[ T_{\alpha,\beta}(\theta, \phi) = S(\frac{\pi}{4} + \beta) \begin{bmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\phi} \end{bmatrix} S(\frac{\pi}{4} + \alpha) \begin{bmatrix} 0 & 1 \\ e^{i\phi} & 0 \end{bmatrix} \]

where \( S(\eta) = e^{i\sigma_x \eta} \) is the transfer matrix for a symmetric crossing, and \( \sigma_x = [0, 1], [1, 0] \) is a Pauli matrix. To first order in (α, β), the norm of the matrix error is:

\[ \| \Delta T \|_{\text{MZI}}^2 = 2[\cos^2(\theta/2)(\alpha + \beta) + \sin^2(\theta/2)(\alpha - \beta)^2] \]

The sensitivity to correlated errors α = β is maximized when the MZI is in the cross state \( \theta = 0 \). Since most MZIs in large meshes are close to the cross state, the matrix error \( \epsilon_0 = 2\sqrt{N\mu} \) for the overall mesh is larger by a factor of \( \sqrt{2} \) compared to the uncorrelated case. For the MZI+Crossing, we find:

\[ T_{\alpha,-\beta}(\theta, \phi) = S(\frac{\pi}{4} + \beta) \begin{bmatrix} e^{-i\theta} & 0 \\ 0 & -1 \end{bmatrix} S(\frac{\pi}{4} + \alpha) \begin{bmatrix} e^{-i\phi} & 0 \\ 0 & 1 \end{bmatrix} S(\frac{\pi}{4}) \]

(Here we assume perfect crossings; the effect of crossing imperfections is discussed later in this section). Up to
some irrelevant output phases, the effect of the crossing is to flip the relative sign of $\alpha$ and $\beta$, so the component errors appear anticorrelated. As a result, $|\Delta T|_{\text{MZI+X}} \propto \sin(\theta/2)\mu$, which is zero for the cross state. Of course, the MZIs in a realistic mesh are not exactly in the cross state, to the actual error will be nonzero. Adding the errors $|\Delta T_{\text{MZI}}|$ in quadrature, and averaging over the probability distribution $P_m(\theta) = n \sin(\theta/2) \cos(\theta/2)^{2n-1}$, we find:

$$E_0 = 2\sqrt{2(\log N + \gamma_\mu - 2)}\mu$$  \hspace{1cm} (20)

Correlated errors (both corrected and uncorrected) are important because they are tightly connected to the operational bandwidth of the mesh. All directional couplers are dispersive, and this dispersion leads to a correlated wavelength-dependent splitter error, which can usually be expanded to first order $\mu \approx (d\mu/d\lambda)\Delta\lambda$. Two important wavelength-dependent figures of merit are:

1. The **bandwidth** is related to the number of wavelength channels that can be **simultaneously** processed by the mesh. Since the mesh will be programmed at a specific wavelength (the center channel), the other channels will inevitably have distorted matrices. Given a target error $E_{\text{max}}$, the bandwidth is the range over which $E_0(\lambda) < E_{\text{max}}$:

$$\lambda_{\text{BW}} = \frac{E_{\text{max}}}{|d\lambda/d\mu|} \left\{ \frac{\sqrt{N}}{\sqrt{2(\log N - 1.42)}} \right\}$$  \hspace{1cm} (21)

2. The **tuning range** refers to the range of $\lambda$ over which the mesh can be programmed to a given accuracy. This is governed by the condition $E_c(\lambda) < E_{\text{max}}$, which yields:

$$\lambda_{\text{TR}} = \frac{E_{\text{max}}}{|d\lambda/d\mu|} \left\{ \frac{\sqrt{N}}{\sqrt{2(\log N - 0.42)}} \right\}$$  \hspace{1cm} (22)

As Fig. 9 illustrates, the MZI+Crossing architecture enjoys a significantly larger tuning range, in addition to modestly greater bandwidth. The respective "enhancement factors", which scale as $F_{\text{BW}} \propto \sqrt{N/\log N}$ and $F_{\text{TR}} \propto (N^{3/4}/\log N)^{1/4}$, are tabulated in Table 1.

| $N$   | 16   | 32   | 64   | 128  | 256  |
|-------|------|------|------|------|------|
| $F_{\text{BW}}$ | $2.4\times$ | $2.8\times$ | $3.4\times$ | $4.3\times$ | $5.6\times$ |
| $F_{\text{TR}}$  | $5.6\times$ | $10\times$  | $18\times$  | $33\times$  | $61\times$  |

Table 1: Bandwidth and tuning range enhancement factors for mesh sizes up to $N = 256$, Eqs. (21)\&(22).

Real crossings have a small amount of nonzero crosstalk, quantified by the S-matrix element $S_{21}$; scattering into the forward-facing port leads to a perturbation $S(\theta/2) \rightarrow S(\theta/2 + \gamma)$ in the transfer matrix, where

$$\gamma = 10^{-S_{21}[\text{dB}]/20}$$  \hspace{1cm} (23)

This does not degrade the effectiveness of self-configuration, since the additional scattering angle merely rotates the Riemann sphere Fig. 7(c) by an additional angle $\gamma \ll 1$, and the forbidden region is still far from $s = 0$. In-plane crossings in silicon can achieve sub-40 dB crosstalk suppression ($\gamma < 0.01$) with insertion losses well below 0.1 dB [36-40]. Unlike directional couplers, crossings are inherently broadband; the insertion loss and crosstalk depend only very weakly on $\lambda$, so any crossing imperfections can be treated as (correctable) wavelength-independent errors that do not affect the bandwidth enhancements of the MZI+Crossing scheme.

In addition to the forward-scattered light, a $90^\circ$ crossing will scatter light into the backward-facing port. Back-reflected light can be subsequently reflected in other crossings, leading to a spurious signal that interferes with the forward-propagating light. Provided that the phases of reflected beams are random, these add in quadrature: with amplitude $\gamma^2$ and $O(N^2)$ scattering paths, we expect this to induce an $O(N\gamma^2)$ error, which may be uncorrectable and set a limit on scaling. However, if this effect is small, gradient-based methods or iterative self-configuration may enable correction of these errors.

## 5 Conclusion

As photonic circuits grow larger, error tolerance becomes increasingly important. Many techniques exist to manage
hardware errors, but all involve a tradeoff between accuracy and complexity. At opposite poles lie “zero-change” error correction, which has limited scalability \[10-15\], and “perfect” photonic circuits, which require a larger number of photonic and electronic components \[22-23\]. This paper has introduced two new designs for programmable circuits that strike a tradeoff between these extremes, as shown in Fig. 10 and Table 2 achieving performance that is almost as good as the perfect designs, but with less added complexity.

The main insight from this paper is that, by adding a single passive component (either a splitter or a waveguide crossing) to the MZI, we can recover behavior that is asymptotically perfect—that is, the average normalized matrix error decreases with size. Our design choices are motivated by the elegant theory of self-configuration by matrix diagonalization \[17\], where splitting ratios are set to successively zero the off-diagonal elements of the target unitary. By visualizing the MZI state on the Riemann sphere, we can intuitively understand the increased error robustness of our designs in terms of “rotating” the forbidden regions away from the peak probability density. This leads to a several-orders-of-magnitude reduction in post-correction errors compared to the standard MZI mesh. The ability to achieve near-perfect and freely scalable MZI meshes with less complexity than the MZI-doubled designs \[22-23\] (especially with respect to the number of active components and pads) removes a major obstacle to the realization of very-large-scale photonic circuits.

An interesting direction for future work is to explore to what extent multiport interferometers can be made robust to imperfections in the absence of error correction. The near-cancellation of correlated errors in the MZI+Crossing architecture explains the \(O(\sqrt{N}/\log N)\) reduction in the uncorrected error, and corresponding increase in bandwidth. Further design modifications based on the theory of composite pulse sequences \[12-13\] may allow this imperfect cancellation to be made exact, further improving the bandwidth (and multiplexing capabilities) of linear photonics.

S.B. is supported by an NSF Graduate Research Fellowship. D.E. acknowledges funding from AFOSR (no. FA9550-20-1-0113, FA9550-16-1-0391). The authors thank Prof. David A. B. Miller and Sunil Pai for helpful discussions.

References

[1] M. Reck, A. Zeilinger, H. J. Bernstein, and P. Bertani, “Experimental realization of any discrete unitary operator,” Physical Review Letters, vol. 73, no. 1, p. 58, 1994.
[2] W. R. Clements, P. C. Humphreys, B. J. Metcalf, W. S. Kolthammer, and I. A. Walmsley, “Optimal design for universal multiport interferometers,” Optica, vol. 3, no. 12, pp. 1460–1465, 2016.
[3] J. Carolan, C. Harrold, C. Sparrow, E. Martín-López, N. J. Russell, J. W. Silverstone, P. J. Shadbolt, N. Matsuda, M. Oguma, M. Itoh et al., “Universal linear optics,” Science, vol. 349, no. 6249, pp. 711–716, 2015.
[4] H.-S. Zhong, H. Wang, Y.-H. Deng, M.-C. Chen, L.-C. Peng, Y.-H. Luo, J. Qin, D. Wu, X. Ding, Y. Hu, P. Hu, X.-Y. Yang, W.-J. Zhang, H. Li, Y. Li, X. Jiang, L. Gan, G. Yang, L. You, Z. Wang, L. Li, N.-L. Liu, C.-Y. Lu, and J.-W. Pan, “Quantum computational advantage using photons,” Science, no. abe8770, 2020.
[5] Y. Shen, N. C. Harris, S. Skirlo, M. Prabhu, T. Baehr-Jones, M. Hochberg, X. Sun, S. Zhao, H. Larochelle, D. Englund et al., “Deep learning with coherent nanophotonic circuits,” Nature Photonics, vol. 11, no. 7, p. 441, 2017.
[6] D. Marpaung, C. Roeloffzen, R. Heideman, A. Leinse, S. Sales, and J. Capmany, “Integrated microwave photonics,” Laser & Photonics Reviews, vol. 7, no. 4, pp. 506–538, 2013.
[7] L. Zhuang, C. G. Roeloffzen, M. Hoekman, K.-J. Boller, and A. J. Lowery, “Programmable photonic signal processor chip for radiofrequency applications,” Optica, vol. 2, no. 10, pp. 854–859, 2015.
[8] R. Burgwal, W. R. Clements, D. H. Smith, J. C. Gates, W. S. Kolthammer, J. J. Renema, and I. A. Walmsley, “Using an imperfect photonic network to implement random unitaries,” *Optics Express*, vol. 25, no. 23, pp. 28236–28245, 2017.

[9] J. Mower, N. C. Harris, G. R. Steinbrecher, Y. Lahini, and D. Englund, “High-fidelity quantum state evolution in imperfect photonic integrated circuits,” *Physical Review A*, vol. 92, no. 3, p. 032322, 2015.

[10] S. Pai, B. Bartlett, O. Solgaard, and D. A. Miller, “Matrix optimization on universal unitary photonic devices,” *Physical Review Applied*, vol. 11, no. 6, p. 064044, 2019.

[11] S. Pai, I. Williamson, T. W. Hughes, M. Minkov, O. Solgaard, S. Fan, and D. A. Miller, “Parallel programming of an arbitrary feedforward photonic network,” *IEEE Journal of Selected Topics in Quantum Electronics*, 2020.

[12] T. W. Hughes, M. Minkov, Y. Shi, and S. Fan, “Training of photonic neural networks through in situ backpropagation and gradient measurement,” *Optica*, vol. 5, no. 7, pp. 864–871, 2018.

[13] D. A. Miller, “Self-aligning universal beam coupler,” *Optica*, vol. 2, no. 8, pp. 6360–6370, 2015.

[14] ——, “Self-configuring universal linear optical component,” *Photons Research*, vol. 1, no. 1, pp. 1–15, 2013.

[15] ——, “Setting up meshes of interferometers–reversed local light interference method,” *Optics Express*, vol. 25, no. 23, pp. 29233–29248, 2017.

[16] R. Hamerly, S. Bandyopadhyay, and D. Englund, “Stability of self-configuring large multiport interferometers,” *arXiv preprint arXiv:2106.04363*, 2021.

[17] ——, “Accurate self-configuration of rectangular multipart interferometers,” *arXiv preprint arXiv:2106.05249*, 2021.

[18] S. Bandyopadhyay, R. Hamerly, and D. Englund, “Hardware error correction for programmable photonic properties,” *arXiv preprint arXiv:2103.04993*, 2021.

[19] S. P. Kumar, L. Neuhaus, L. G. Helt, H. Qi, B. Morrison, D. H. Mahler, and I. Dhand, “Mitigating linear optics imperfections via port allocation and compilation,” *arXiv preprint arXiv:2103.03183*, 2021.

[20] V. J. López-Pastor, J. S. Lundeen, and F. Marquardt, “Arbitrary optical wave evolution with fourier transforms and phase masks,” *arXiv preprint arXiv:1912.04721*, 2019.

[21] S. Pai et al., *in preparation*, 2021.

[22] D. A. Miller, “Perfect optics with imperfect components,” *Optica*, vol. 2, no. 8, pp. 747–750, 2015.

[23] K. Suzuki, G. Cong, K. Tanizawa, S.-H. Kim, K. Ikeda, S. Namiki, and H. Kawashima, “Ultra-high-extinction-ratio 2×2 silicon optical switch with variable splitter,” *Optics Express*, vol. 23, no. 7, pp. 9086–9092, 2015.

[24] A. Haar, “Der massbegriff in der theorie der kontinuierlichen gruppen,” *Annals of Mathematics*, pp. 147–169, 1933.

[25] W.-K. Tung, *Group theory in physics: an introduction to symmetry principles, group representations, and special functions in classical and quantum physics*. World Scientific Publishing Company, 1985.

[26] N. J. Russell, L. Chakhmakhchyan, J. L. O’Brien, and A. Laing, “Direct dialling of Haar random unitary matrices,” *New Journal of Physics*, vol. 19, no. 3, p. 033007, 2017.

[27] M. Wang, A. Ribero, Y. Xing, and W. Bogaerts, “Tolerant, broadband tunable 2×2 coupler circuit,” *Optics Express*, vol. 28, no. 4, pp. 5555–5566, 2020.

[28] R. Hamerly, “Meshes: tools for modeling photonic beamsplitter mesh networks,” Online at: https://github.com/QPG-MIT/meshes, 2021.

[29] L. B. Soldano and E. C. Pennings, “Optical multimode interference devices based on self-imaging: principles and applications,” *Journal of Lightwave Technology*, vol. 13, no. 4, pp. 615–627, 1995.

[30] A. Maese-Novo, R. Halir, S. Romero-García, D. Pérez-Galacho, L. Zavargo-Peche, A. Ortega-Moñux, I. Molina-Fernández, J. Wangüemert-Pérez, and P. Cheben, “Wavelength independent multimode interference coupler,” *Optica Express*, vol. 21, no. 6, pp. 7033–7040, 2013.

[31] Y. Wang, Z. Lu, M. Ma, H. Yun, F. Zhang, N. A. Jaeger, and L. Chrostowski, “Compact broadband directional couplers using subwavelength gratings,” *IEEE Photonics Journal*, vol. 8, no. 3, pp. 1–8, 2016.

[32] C. Ye and D. Dai, “Ultra-compact broadband 2×2 3 dB power splitter using a subwavelength-grating-assisted asymmetric directional coupler,” *Journal of Lightwave Technology*, vol. 38, no. 8, pp. 2370–2375, 2020.

[33] H. Morino, T. Maruyama, and K. Iiyama, “Reduction of wavelength dependence of coupling characteristics using Si optical waveguide curved directional coupler,” *Journal of Lightwave Technology*, vol. 32, no. 12, pp. 2188–2192, 2014.
[34] Z. Lu, H. Yun, Y. Wang, Z. Chen, F. Zhang, N. A. Jaeger, and L. Chrostowski, “Broadband silicon photonic directional coupler using asymmetric-waveguide based phase control,” Optics Express, vol. 23, no. 3, pp. 3795–3808, 2015.

[35] K. Suzuki, R. Konoike, J. Hasegawa, S. Suda, H. Matsuura, K. Ikeda, S. Namiki, and H. Kawashima, “Low-insertion-loss and power-efficient 32×32 silicon photonics switch with extremely high-δ silica PLC connector,” Journal of Lightwave Technology, vol. 37, no. 1, pp. 116–122, 2018.

[36] T. Fukazawa, T. Hirano, F. Ohno, and T. Baba, “Low loss intersection of Si photonic wire waveguides,” Japanese Journal of Applied Physics, vol. 43, no. 2R, p. 646, 2004.

[37] H. Chen and A. W. Poon, “Low-loss multimode-interference-based crossings for silicon wire waveguides,” IEEE Photonics Technology Letters, vol. 18, no. 21, pp. 2260–2262, 2006.

[38] Y. Ma, Y. Zhang, S. Yang, A. Novack, R. Ding, A. E.-J. Lim, G.-Q. Lo, T. Baehr-Jones, and M. Hochberg, “Ultralow loss single layer submicron silicon waveguide crossing for SOI optical interconnect,” Optics express, vol. 21, no. 24, pp. 29374–29382, 2013.

[39] P. Dumais, D. Goodwill, D. Celo, J. Jiang, and E. Bernier, “Three-mode synthesis of slab gaussian beam in ultra-low-loss in-plane nanophotonic silicon waveguide crossing,” in 2017 IEEE 14th International Conference on Group IV Photonics (GFP). IEEE, 2017, pp. 97–98.

[40] S. Wu, X. Mu, L. Cheng, S. Mao, and H. Fu, “State-of-the-art and perspectives on silicon waveguide crossings: a review,” Micromachines, vol. 11, no. 3, p. 326, 2020.

[41] B. A. Bell and I. A. Walmsley, “Further compactifying linear optical unitaries,” arXiv preprint arXiv:2104.07561, 2021.

[42] K. R. Brown, A. W. Harrow, and I. L. Chuang, “Arbitrarily accurate composite pulse sequences,” Physical Review A, vol. 70, no. 5, p. 052318, 2004.

[43] J. Bulmer, J. Jones, and I. Walmsley, “Drive-noise tolerant optical switching inspired by composite pulses,” Optics Express, vol. 28, no. 6, pp. 8646–8657, 2020.