

# Design of Asymptotically Perfect Linear Feedforward Photonic Circuits

Ryan Hamerly<sup>(1,2)</sup>, Saumil Bandyopadhyay<sup>(1)</sup>, Alexander Sludds<sup>(1)</sup>, Dirk Englund<sup>(1)</sup>

<sup>(1)</sup> MIT RLE, 50 Vassar Street, Cambridge, MA 02139, USA

<sup>(2)</sup> NTT Research Inc., PHI Laboratories, 940 Stewart Drive, Sunnyvale, CA 94085, USA

rhamerly@mit.edu

**Abstract:** We propose a new architecture for feedforward photonic circuits based on a 3-splitter MZI. This architecture is more error tolerant than the standard mesh, supports self-configuration, and yields asymptotically perfect circuits for large mesh sizes. © 2022 The Author(s)

## 1. Introduction

Large-scale programmable photonic circuits have opened up transformative new applications for optics [1–3]. Central to many applications is the universal multiport interferometer: an  $N \times N$  programmable linear circuit (Fig. 1(a-b)) usually constructed from a dense mesh of Mach-Zehnder interferometers (MZIs) [4]. Component errors arising from fabrication process variations (Fig. 1(c)) pose a major practical challenge to scaling multiport interferometers, as the error in matrix fidelity increases with the circuit depth,  $O(\sqrt{N})$  for uncorrelated errors.

Two solutions exist to the problem of hardware errors: (1) eliminate them by constructing MZI-doubled “perfect” photonic circuits, at the cost a 1.5–2× increase in chip area and pad count, or (2) employ a “zero-change” error-correction algorithm based on global optimization [7], self-configuration [8,9], or local correction [10]. Such algorithms do not increase the hardware complexity, but the error reduction is limited to a quadratic factor [9], so errors still pose a strict limit to scaling of the circuit.

In this paper, we propose a hybrid approach that achieves the best of both options. Instead of a standard MZI, we use a 3-splitter design (3-MZI, Fig. 1(d)) [11], programmed using progressive self-configuration [9]. Our approach is motivated by the intimate correspondence between  $SU(2)$  and the Riemann sphere, where hardware errors lead to “forbidden regions” near the poles of this sphere, and the effect of the 3-MZI is to rotate these regions away from the peak probability density. Numerical simulations confirm that, with our approach, errors are reduced several orders of magnitude compared to the standard MZI mesh. Remarkably, the overall circuit error *decreases* with size, allowing asymptotically perfect circuits in the limit of large mesh sizes.

## 2. Self-Configuration and the 3-MZI

To correctly configure an MZI mesh, we use the *physical nulling* method based on a combination of diagonalization and physical measurements [9]. Fig. 1(e-f) illustrates the procedure for the triangular mesh [4]. Starting from

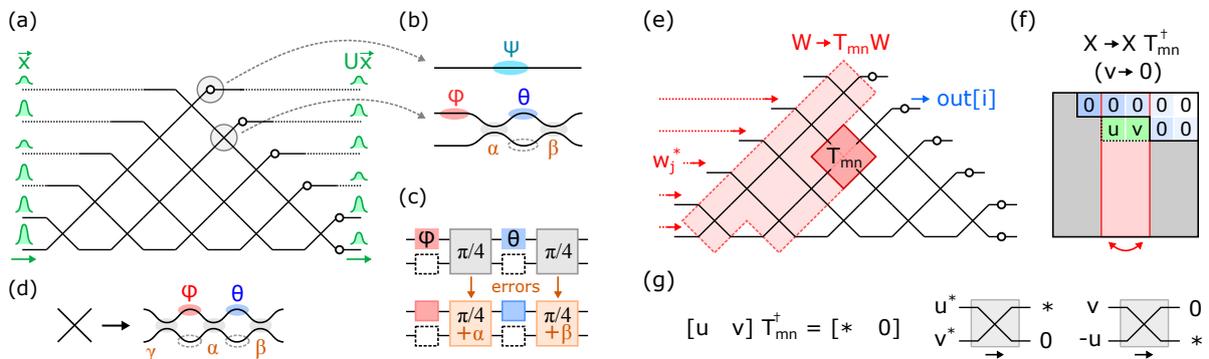


Fig. 1. (a)  $6 \times 6$  triangular photonic mesh, composed of (b) a phase screen  $\psi$  and tunable MZI crossings  $\theta, \phi$ . (c) Component errors  $\alpha, \beta$  arising from fabrication. (d) Proposed alternative 3-splitter MZI design. (e-g) Self-configuration procedure for correcting errors in the mesh (Sec. 2).

the transfer matrix, represented as a product of  $2 \times 2$  blocks  $U = DW$  ( $W = \prod_{mn} T_{mn}$ , where  $T_{mn}$  is the  $n^{\text{th}}$  MZI of the  $m^{\text{th}}$  diagonal), we configure the mesh by “building up”  $W$  in a sequence of steps that diagonalize a target matrix  $X = UW^\dagger$ . Each step adds a crossing to  $W \rightarrow T_{mn}W$  (Fig. 1(e)), performing the update  $X \rightarrow XT_{mn}^\dagger$  (Fig. 1(f)). The phase shifts  $(\theta, \phi)$  are chosen to zero a particular matrix element  $v \rightarrow 0$  (green in figure), satisfying the equation:

$$[u \ v]T^\dagger = [* \ 0] \Leftrightarrow T_{11}/T_{12} = u/v \quad (1)$$

This is illustrated in Fig. 1(g). Physically, nulling corresponds to injecting  $w_j^*$  (the  $j^{\text{th}}$  column of  $W^\dagger$ ) and zeroing the power at the  $i^{\text{th}}$  output [9]. Mathematically, it corresponds to matching the complex splitting ratio  $s = T_{11}/T_{12}$  to a target value  $\hat{s} \equiv u/v$ . For an MZI, this is not always possible, as the allowed range of splitting ratios  $|s| \in [\tan|\alpha + \beta|, \cot|\alpha - \beta|]$  is constrained by the component errors  $\alpha, \beta$  (Fig. 1(c)) [7]. Therefore, errors lead to *forbidden regions* (Fig. 2(a)) in the space of splitting ratios  $s$ , where nulling cannot be achieved perfectly. Imperfect nulling leads to off-diagonal residuals on the target matrix  $X$ , whose magnitude is proportional to the Euclidean distance  $d(s, \hat{s})$  between the target  $\hat{s}$  and its closest realizable value  $s$  (Fig. 2(a)). Imperfect diagonalization of  $X$  leads to imperfect realization of the target matrix  $U$ .

Viewed in this light, there are two paths to improve the accuracy of the mesh. The first is to eliminate the forbidden regions altogether, restoring perfect fidelity. As previously mentioned, this can be done by MZI doubling [5, 6], but at the cost of a significant increase in hardware complexity. Alternatively, one can rotate (Fig. 2(b)) the forbidden regions away from the peak of probability density (Fig. 2(c)). Rotation is achieved by appending a third splitter to the MZI, which performs a Möbius transformation  $s \rightarrow (s + i)/(1 + is)$  on the splitting ratio  $s$ : 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 [11]. However, the presence of forbidden regions means that the

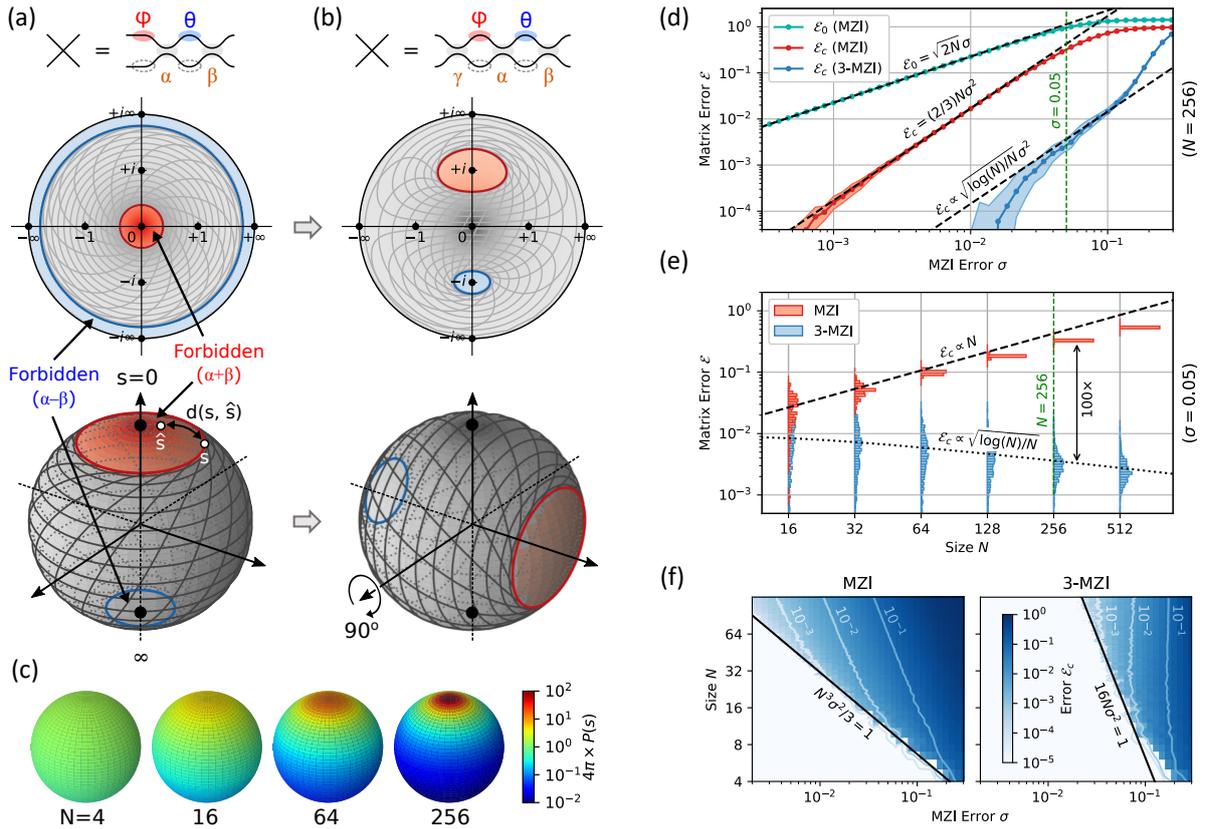


Fig. 2. (a) Allowed range in MZI for  $s = T_{11}/T_{12}$ ; regions near  $s = 0$  and  $s = \infty$  are forbidden due to imperfections. Contours are lines of constant  $(\theta, \phi)$ , with  $\alpha = 0.23$ ,  $\beta = 0.07$ . (b) Probability density  $P(s)$  as a function of mesh size. (c) Adding a third splitter shifts the forbidden regions. (d) Matrix error  $\epsilon_0$ ,  $\epsilon_c$  as a function of splitter variation  $\sigma$  (fixed  $N = 256$ ), comparing the standard and 3-splitter MZI designs. (e) Scaling of  $\epsilon_c$  with mesh size  $N$  (fixed  $\sigma = 0.05$ ). (f) Matrix error as a function of both  $N$  and  $\sigma$ , showing the sharp onset of “perfect” error correction in regions where the coverage Eq. (3) is of order unity.

relative *phase* of this splitter cannot be fully controlled, which means that errors can still occur when programming the mesh (unlike the “perfect” MZIs [5,6]). However, in this case compared to the standard MZI, couplers fall into the forbidden regions only infrequently; therefore, the error for a 3-MZI will be much lower than for a standard mesh, and perhaps more surprisingly, this error decreases with mesh size.

### 3. Accuracy and Scaling

To confirm the advantage of the 3-MZI design, we performed numerical experiments on triangular meshes [4] of sizes  $16 \leq N \leq 512$ , with uncorrelated Gaussian errors  $\alpha, \beta \sim N(0, \sigma)$  (correlated errors show similar statistics [12]), with Haar-random target matrices [13]. Our algorithms are available as part of the MESHES package [14].

The matrix error  $\mathcal{E} = \langle \|\Delta U\|_{\text{rms}} \rangle / \sqrt{N}$ , which quantifies the average relative error per matrix element, is the most important figure of merit for most applications. Fig. 2(d) plots  $\mathcal{E}$  as a function of  $\sigma$  for a fixed mesh size  $N = 256$ . If an MZI mesh is straightforwardly programmed without taking any account of the imperfections (“uncorrected” error), the error under an uncorrelated perturbation model would be  $\mathcal{E}_0 = \sqrt{2N}\sigma$ . Correction leads to a quadratic suppression of errors [9, 12]:  $(\mathcal{E}_c)_{\text{MZI}} = (2/3)N\sigma^2 = \mathcal{E}_0^2/3$ . However, error correction only reaps benefits when the errors are sufficiently small to begin with; moreover, the scaling  $\mathcal{E}_c \propto N$  means that errors still pose a hard upper limit on the mesh size.

For the 3-MZI architecture, following the distribution of splitting angles for Haar-random unitaries [13], we calculate:

$$(\mathcal{E}_c)_{3\text{-MZI}} = 8\sigma^2 \left[ 2 \frac{\log(N) + \gamma_e - \frac{5}{4} - \log(2)}{N} \right]^{1/2} \quad (2)$$

where  $\gamma_e \approx 0.5772$  is the Euler-Mascheroni constant. For most mesh sizes,  $\mathcal{E}_c$  is 1–2 orders of magnitude smaller for the 3-MZI design. Remarkably, the error actually *decreases* with increasing  $N$ , scaling as  $\mathcal{E}_c \propto \sqrt{\log(N)/N}$  (Fig. 2(e)). In the asymptotic limit  $N \rightarrow \infty$ , the mesh is perfect.

Another advantage of the 3-splitter MZI is a higher coverage (fraction of  $U(N)$  that is perfectly realized by the mesh). Coverage is computed by summing the probabilities that individual crossings fall within the forbidden regions [13]; this is exponentially larger for the 3-MZI for large meshes:

$$\text{cov}(N) = \begin{cases} e^{-N^3\sigma^2/3} & (\text{MZI}) \\ e^{-16N\sigma^2} & (3\text{-MZI}) \end{cases} \quad (3)$$

Consequently, the threshold for perfect error correction is higher, scaling as  $\sigma_{\text{th}} \propto N^{1/2}$ , as opposed to the  $N^{3/2}$  scaling observed for a standard MZI. Perfect correction is available under a wider range of conditions (Fig. 2(f)).

### 4. Conclusion

Hardware error correction is critical for many-mode multiport interferometers, but error correction schemes suffer from a tradeoff between accuracy and complexity. The 3-MZI architecture proposed in this paper offers a solution to this tradeoff, yielding orders-of-magnitude error reduction over the standard MZI mesh at the small cost of a single additional splitter per MZI. Moreover, with progressive self-configuration, the 3-MZI mesh is asymptotically perfect, as the matrix error decreases with mesh size. By breaking the scaling bound set by hardware errors, the 3-MZI eliminates a major obstacle to the development of very-large-scale linear photonic circuits.

### References

1. J. Carolan, C. Harrold, C. Sparrow, E. Marín-López et al., “Universal linear optics,” *Science*, **349**, 711–716 (2015).
2. Y. Shen et al., “Deep learning with coherent nanophotonic circuits,” *Nature Photonics* **11**, 441 (2017).
3. L. Zhuang et al., “Programmable photonic signal processor chip for radiofrequency applications,” *Optica* **2**, 854 (2015).
4. M. Reck et al., “Experimental realization of any discrete unitary operator,” *Phys. Rev. Lett.* **73**, 58 (1994).
5. D. A. B. Miller, “Perfect optics with imperfect components,” *Optica* **2**, 747–750 (2015).
6. K. Suzuki et al., “Ultra-high-extinction-ratio  $2 \times 2$  silicon optical switch with variable splitter,” *Opt. Exp.* **23**, 9086 (2015).
7. R. Burgwal et al., “Using an imperfect photonic network to implement random unitaries,” *Opt. Exp.* **25**, 28236 (2017).
8. D. A. B. Miller, “Self-configuring universal linear optical component,” *Phot. Res.* **1**, 1 (2013).
9. R. Hamerly et al., “Accurate self-configuration of rectangular multiport interferometers,” *arXiv:2106.03249* (2021).
10. S. Bandyopadhyay et al., “Hardware error correction for programmable photonics,” *Optica* **8**, 1247–1255 (2021).
11. M. Wang et al., “Tolerant, broadband tunable  $2 \times 2$  coupler circuit,” *Opt. Exp.* **28**, 5555 (2020).
12. R. Hamerly et al., “Stability of Self-Configuring Large Multiport Interferometers,” *arXiv:2106.04363* (2021).
13. N. J. Russell et al., “Direct dialling of Haar random unitary matrices,” *New J. Phys* **19**, 033007 (2017).
14. R. Hamerly, *Meshes: Tools for modeling photonic beamsplitter mesh networks*, Online at: <https://github.com/QPG-MIT/meshes> (2021).