SILICON PHOTONICS

Meshing optics with applications

Two concurrent demonstrations of programmable photonic processors based on large meshes of interconnected waveguides on a silicon chip provide new hope for optical information processing.

David A. B. Miller

Silicon photonics enables the realization of precise integrated optical circuits with complexities well beyond the arrangements of classical bench-top optics of bulk lenses, mirrors and beamsplitters. Two papers in this issue of *Nature Photonics*, by Shen *et al.*[1] and Harris *et al.*[2], demonstrate the considerable progress made in the area with the report of programmable photonic circuits with hundreds of optical components in chips of millimetre sizes. The chips are put to use to demonstrate two innovative applications in a preliminary form: linear-optics matrix multiplication for neural networks[1] and simulation of quantum transport[2].

These papers are impressive examples of a growing field of programmable linear optical processors[3–11]. Such processors use meshes of single-mode waveguides laid out on the flat silicon substrate in patterns like chain-link fences, typically in the form of an interconnected array of Mach–Zehnder interferometers (MZIs; Fig. 1). Phase shifters in the arms of the waveguides are used to control the interference of beams where the waveguides intersect, thus allowing the relative amplitude and phase of the beams in the output waveguides to be controlled — in effect, programming the transmission response of the mesh. This simple recipe allows a surprisingly large range of optical functions to be performed (see ref. 12 for a short introduction) with some layouts allowing arbitrary linear transforms, giving universal linear optical components[3,5].

Such interferometer meshes are expected to be useful for applications as diverse as spatial mode (de)multiplexing in telecommunications; tracking multiple targets in imaging; finding the best transmission channels through scattering media; linear optical quantum computing; lossless beam-power combining for multiple coherent beams with any relative amplitudes and phases; and unscrambling scattered light[12]. Some architectures can set themselves up automatically with progressive algorithms based on simple local feedback loops[5–7,11]. Such architectures allow systems that can perform tasks such as self-aligning the coupling of a beam to a waveguide; sorting and separating beams based on training using the beams themselves, without calculations or calibrations; correcting themselves for fabrication errors; and automatically undoing scattering between beams, even if the scattering changes in time. Extended mesh architectures can implement many of the functions and circuits needed in photonic processing of microwave signals[8–10].

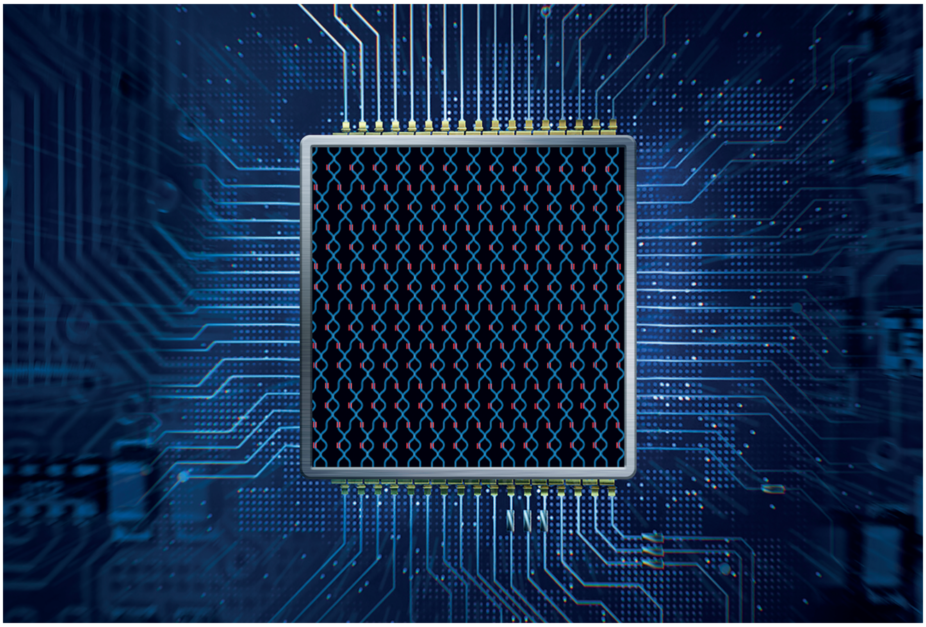

Figure 1 | Artist's impression of a programmable photonic processor chip such as that demonstrated by Shen *et al.*[1] consisting of a waveguide mesh of interconnected Mach–Zehnder interferometers. The phase shifters are indicated by the red blocks. Image: RedCube Inc.

Several of these mesh architectures differ qualitatively from most conventional optics in one important way; each output waveguide amplitude can be an arbitrary linear combination of all of the input waveguide amplitudes. Crucially, the underlying mathematics of such an arrangement enables the mesh architectures to implement an arbitrary matrix, and with little or no unnecessary loss. Mesh architectures for unitary matrices have been known for some time[3], but more recent universal architectures, based on the singular value decomposition (SVD) of the matrix[5], can implement arbitrary (and hence also non-unitary) matrices in an optimally compact form.

Shen *et al.*[1] and Harris *et al.*[2] each implement linear classical analog processors, operating on the input light amplitudes to generate outputs. Both exploit large mesh networks, with 56 (ref. 1) and 88 (ref. 2) MZIs, respectively; with two phase shifters and two waveguide beamsplitters for each interferometer, the total number of subcomponents is measured in the hundreds, well beyond the complexity of bench-top optics. The performance of the individual interferometers can be very high; Harris *et al.*[2] show fabricated interferometers with rejection ratios of ~66 dB, a record value implying precise beamsplitter fabrication, even compared with devices with automatic post-fabrication correction[11].

Shen *et al.*[1] use their mesh to demonstrate key aspects of an optical neural network processor. Such processors, which have been considered for many decades, exploit aspects of the architectures of the brain. They involve highly connected networks of nonlinear elements or 'neurons' that are 'trained' or 'self-learn' to perform useful decision-making tasks. Interest in such networks has grown recently, in part because of improved understanding of the power of networks with multiple buried layers of neurons, and in part because of growing interest in a range of opportunities in image processing, language understanding and translation, decision support, and other areas. Several major companies are currently making large electronic systems for such 'deep learning' applications.

Once the various 'weights' or strengths of interconnections between layers of a neural network are calculated in a 'training' step (which Shen *et al.*[1] perform separately on a conventional computer), the physical connection strengths in a network can implement those weights. Operation of the whole network then consists of a matrix multiplication, followed by a set of nonlinear 'neuron' operations on the outputs, followed by another matrix multiplication and neuron 'layer', and so on for all the layers in the network. Data to be processed are fed in as an appropriate vector of amplitudes to the first matrix multiplier.

Shen and colleagues' optical work focuses on matrix multiplication, using the SVD architecture[5] as an optical interference unit. Their MZI mesh is large enough to allow two of the three matrices in the SVD of a 4×4 matrix to be implemented at once. They propose approaches for the required nonlinear optical elements to form the neurons. Using a combination of actual optical multiplications and emulations of other required aspects, they construct a multilayer optical neural network and test its performance. On a vowel recognition task, its performance is not much worse than that of a 64-bit electronic processor, even though the optical system may be operating at an effectively much lower resolution.

Their main argument for developing such an optical approach is that linear optics can potentially perform the multiplication of the input by the weights with little or no energy consumption. This has been a feature of linear optical processing for many decades — in the 1980s, other linear optical matrix-vector multipliers were used for neural networks. Ultimately, such work did not lead to optical systems that were attractive enough to replace electronics, especially as the performance of the latter continued its Moore's Law march of exponential improvement.

However, several things have changed since then: the use of the MZI mesh avoids the '$1/N$' optical loss of previous schemes (for an N element vector); the complexity of optical circuits that can be realized in an integrated platform and the interest in neural architectures have both grown substantially; and the bounds on improvements in the performance of electronic systems are more evident, both in the slowing down of Moore's Law and the limitations of electrical interconnects.

Several questions remain open for the photonics approach and need to be resolved for it to gain traction. First, previous optical schemes foundered on the difficulty of driving optical data fast enough into the processors. Can modern approaches — such as multiple high-bit-rate streams as proposed by Shen *et al.*[1] — solve this? The necessary time-multiplexing itself consumes significant energy, for example. Second, although nonlinear optical elements are possible, can those practically achieve low enough energies and background loss? Third, can we make optical approaches at the scale necessary to compete with electronics? Fourth, can optics compete with advancing or alternative electronic approaches, such as low precision or integer multiplication, or electronic linear analog matrix multiplication? Shen *et al.*[1] do, however, propose a serious model for comparison, correctly emphasizing the dominating importance of energy in information processing.

With a completely different application in mind, Harris *et al.*[2] set up a MZI mesh to simulate the transport of different paths in a one-dimensional scattering medium. With this mesh, they perform simulation experiments for a finite number of scatterers with 'nearest-neighbour' scattering and for several simulation 'time steps'. In different experiments, they vary the parameters that characterize the disorder in the medium, especially in the phase of different paths in the scattering. Rather than simulating in actual time steps, they use successive layers in the mesh to represent successive time steps. With this approach, they emulate the time evolution for two different classes of disorder — 'static', where the phases in the scattering paths do not change with each successive 'time' step, and 'dynamic', where the phases change with successive time steps. For static disorder, they use sets of phases that are the same in each successive layer of their mesh; for dynamic disorder, they vary the phases between successive layers. With their processor, they were able to perform over 64,000 different experiments with sets of randomly chosen phases, which allows statistically meaningful characterization.

They can emulate ballistic transport, diffusive incoherent transport and Anderson localization (resulting from static phase variations). They can see the phenomenon of environment-assisted quantum transport in which a particle initially localized at a point can escape, ironically, through the introduction of dynamic disorder. As they introduce more dynamic disorder, rather than further enhancing the transport, they see instead inhibition — a 'quantum Goldilocks' regime of an optimum range of dynamic disorder to enhance transport. These experiments show the first evidence for environment-assisted quantum transport and a quantum Goldilocks regime in such discrete time scattering.

Such emulations are not fundamentally beyond the capabilities of an electronic digital computer, but the ability to perform complex calculations just by changing a few physical phase shifters is intriguing. Where such processors move substantially beyond any conventional machines is when they are fed, not with classical light, but with identical single photons[3,4] to allow quantum computation. To explore such a regime, the mesh processor can remain essentially the same, so the success of this large-scale 'classical' demonstration is encouraging for future quantum machines.

The papers by Shen *et al.*[1] and Harris *et al.*[2] show rapid progress in the physical technology of complex mesh processors, and give serious explorations of new application areas. Technical challenges remain; for example, phase shifters without the large power dissipation and lengths of thermal devices are required, as are even larger sizes of meshes with simple and low-loss coupling of very large numbers of beams. But, hopefully, work such as this gives increasing confidence and motivation to advance this promising emerging area of optics. ❐

David A. B. Miller is at the Edward L. Ginzton Laboratory, Stanford University, 348 Via Pueblo Mall, Stanford, California 94305, USA.
e-mail: dabm@stanford.edu

References

1. Shen, Y. *et al. Nat. Photon.* **11,** 441–446 (2017).
2. Harris, N. C. *et al. Nat. Photon.* **11,** 447–452 (2017).
3. Carolan, J. *et al. Science* **349,** 711–716 (2015).
4. Tillmann, M. *et al. Nat. Photon.* **7,** 540–544 (2013).
5. Miller, D. A. B. *Photon. Res.* **1,** 1–15 (2013).
6. Annoni, A. *et al.* Preprint at https://arxiv.org/abs/1512.06762 (2017).
7. Ribeiro, A., Ruocco, A., Vanacker, L. & Bogaerts, W. *Optica* **3,** 1348–1357 (2016).
8. Zhuang, L., Roeloffzen, C. G. H., Hoekman, M., Boller, K. J. & Lowery, A. J. *Optica* **2,** 854–859 (2015).
9. Pérez, D., Gasulla, I., Capmany, J. & Soref, R. A. *Opt. Express* **24,** 12093–12106 (2016).
10. Hall, T. J. & Hasan, M. *Opt. Express* **24,** 7600–7610 (2016).
11. Wilkes, C. M. *et al. Opt. Lett.* **41,** 5318–5321 (2016).
12. Miller, D. A. B. *Science* **347,** 1423–1424 (2015).