Mode Division Multiplexing (MDM) Weight Bank Design For Use in Photonic Neural Networks

Ethan K. Gordon, Alexander N. Tait, Paul R. Prucnal

Princeton University, Department of Electrical Engineering, EQuad, Olden St., Princeton, NJ 08544 USA

egordon@princeton.edu

PRINCETON Electrical Engineering

Motivation and Previous Work

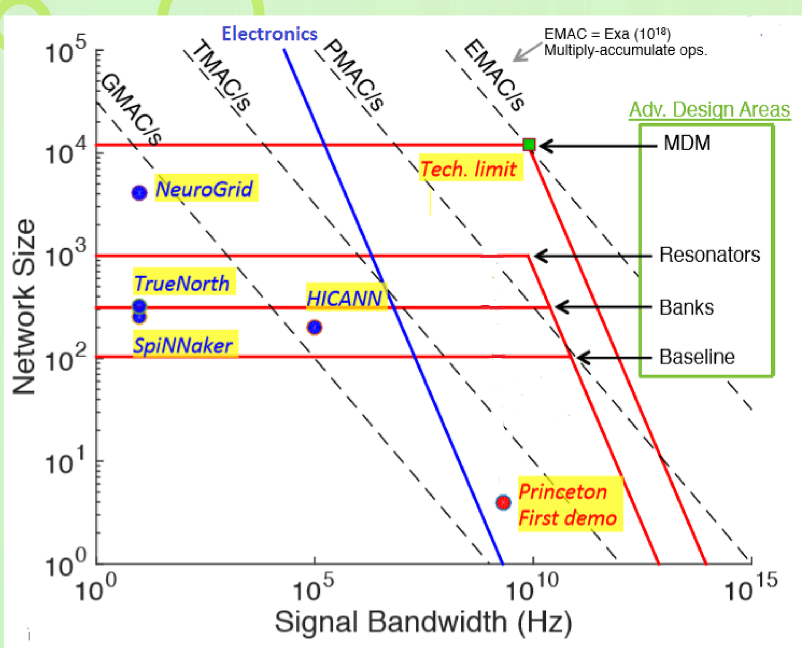
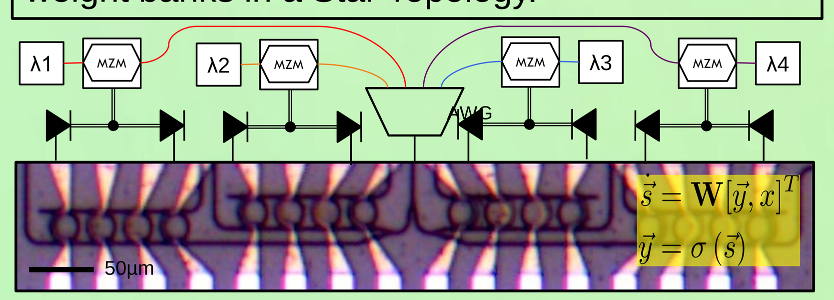


Figure 1 (above): Network Size / Bandwidth Tradeoff. MDM can increase network size by an order of magnitude.
Figure 2 (below): Princeton's First Network, four weight banks in a Star Topology.



The Neuron-Bandwidth Tradeoff

One fundamental limit on neural networks is the tradeoff between network size and the total signal bandwidth. As more neurons go on the network, the network slows down. As shown in **Figure 1**, electronic networks have been thousands of neurons in size, but the nature of electronics in general prohibits bandwidths in the Terahertz range. Optical networks circumvent this bandwidth limitation, but introduce a size limitation. Only so many wavelength channels can be effectively used in a single waveguide. Princeton's first network demonstration only had four wavelengths, corresponding to only four neurons.

Experiments 1 and 2 investigate the addition of mode-division multiplexing to create two dimensions of channels, increasing potential network size by an order of magnitude without affecting bandwidth.

Princeton's First Photonic Neural Network

- **Star Topology:** The combined axon outputs would originate at a single point and branch off to all the different neurons in the network.
- **Off-chip Elements:** The neurons and photodetectors are off-chip.
- **Four Neurons:** Limited wavelengths available in a waveguide.

Experiment 3 investigates the use of a folding bus, or "hairpin," topology, allowing for a more flexible and scalable chip design.

Experiment 1: Optimizing Mode Coupling

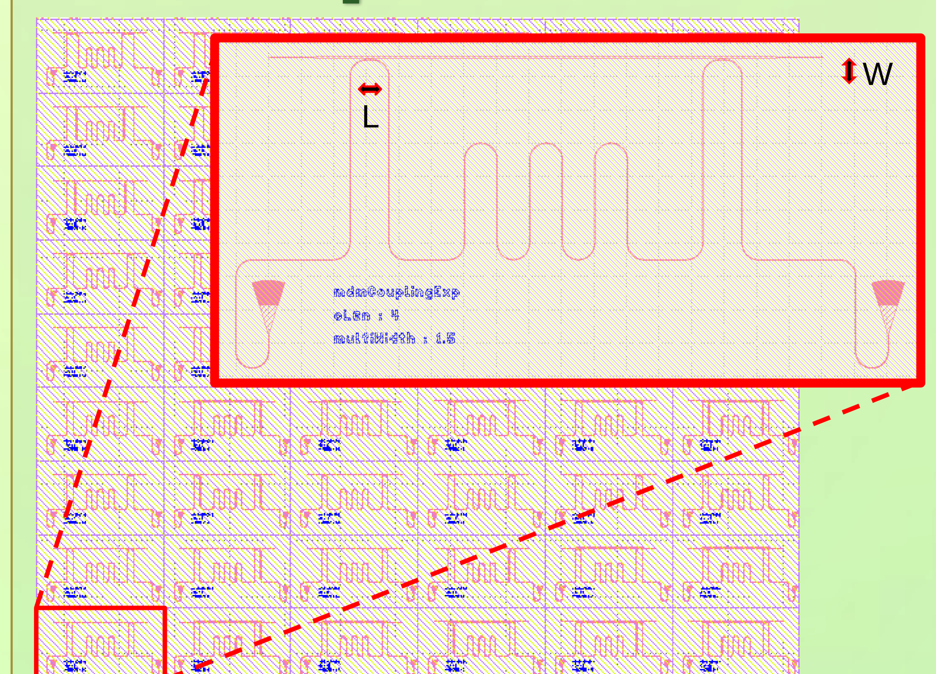


Figure 3: Asymmetric Mach-Zehnder Interferometer Design. Waveguid Width (W) and Coupling Length (L) need to be optimized for maximum coupling.

Experimental Design

- **Goal:** Determine the optical multi-mode waveguide width and coupling length to achieve 100% coupling.
- **Device:** An asymmetric Mach-Zehnder interferometer, as in **Figure 3**. Depending on the wavelength, the transfer coefficient will differ.
- **Expected Data:** A sinusoidal spectrogram, as shown in **Figure 4,** where the amplitude is called the *Extinction Ratio*.
- **Method:** The Extinction Ratio maximizes at the optimal width for any given mode. However, a maximum extinction ratio is also only 50% coupling. We adjust the coupling length to bring the extinction ratio to 0 (a null), which corresponds with 100% coupling.

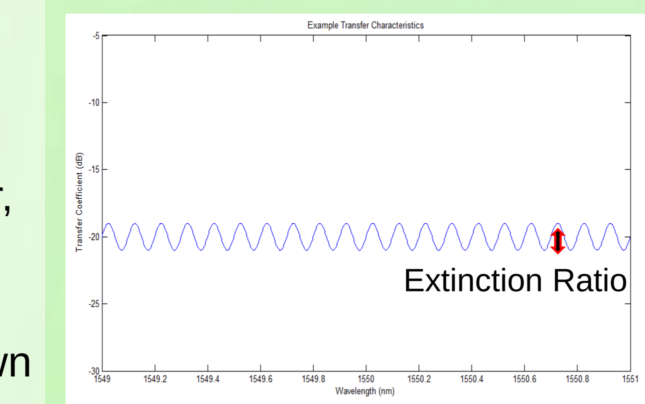


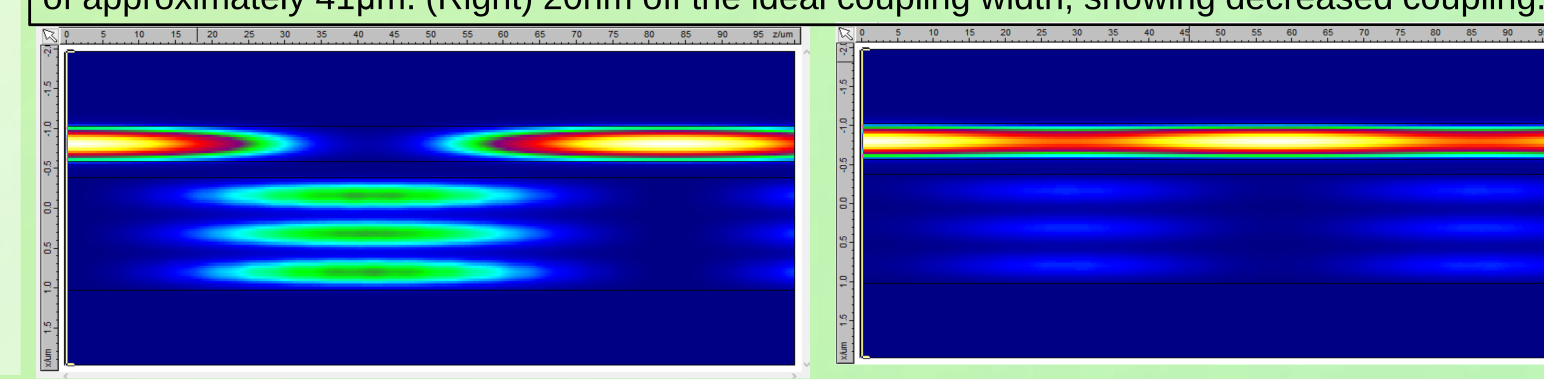
Figure 4 (above): Sample Spectrogram Data for transfer characteristics of each modular block. The extinction ratio (shown) will be maximized when the coupling width is optimal.

Simulated Results

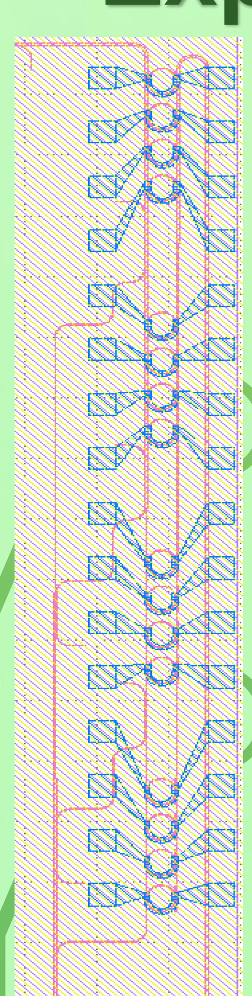
Since a null extinction ratio could actually correspond with 100% or 0% coupling, we can use simulation data to determine which is more likely.

(All μm)	TE_0	TE_1	TE_2	TE_3
W	0.5	0.95	1.4	1.9
L	25	35	41	47

Figure 5: Simulated Coupling Results. (Left) The ideal coupling width, showing a half-beat length of approximately 41μm. (Right) 20nm off the ideal coupling width, showing decreased coupling.



Experiment 2: Weight Bank



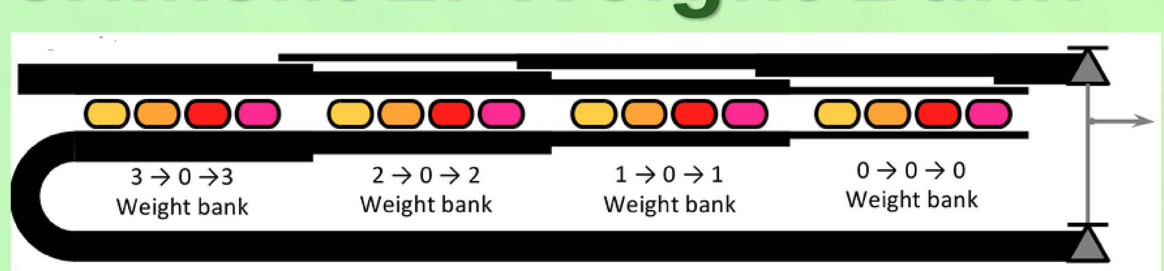


Figure 6 (above): Layout of 4-mode, 4-wavelength Weight Bank
Figure 7 (left): Final CAD of Weight Bank on chip.

Experimental Design

- **Goal:** Demonstrate a device that can perform a weighted sum on a channels at different modes and wavelengths.
- **Device:** As shown in **Figure 6 and** 7: N stages for N modes, with each stage consisting of M microring resonators (MRRs) for M wavelengths. Only the highest order mode will couple in each stage.
- **Expected Data:** Measure the transfer coefficient for each wavelength and mode while electrically modifying the coupling percentage of each MRR. Using off-chip balanced photodiodes, weights can range from -1 (100% coupling to the MRR), and 1 (0% coupling).

Experiment 3: Network Topology

Experimental Design

- **Goal:** Demonstrate a network in a hairpin topology with arbitrary waveguide bends and multi-mode neurons.
- **Device:** Two neurons on different mode channels, including axons, one *cascaded weight bank* to drop power to the first neuron, and on non-cascaded weight bank for the second neuron.

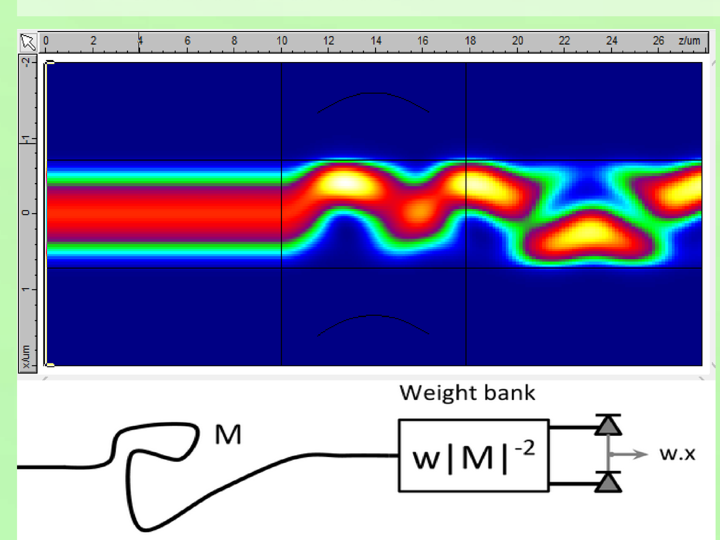


Figure 8: Example of Intermodal Mixing in a bent fiber and the use of a weight bank to invert the mix matrix and apply a custom weight "w."

Intermodal Mixing

While the hairpin topology lends itself well to multi-mode systems, as there is no easy way to split multi-mode optical power into different waveguides, the arbitrary bending of the multi-mode waveguides lead to intermodal mixing, as shown in **Figure 8 (top)**. The bending effectively applies a mixing matrix to the incoming modes. Fortunately, as the weight bank is taking an effective inner product anyway, the inverse of this matrix can be factored into the weight bank.

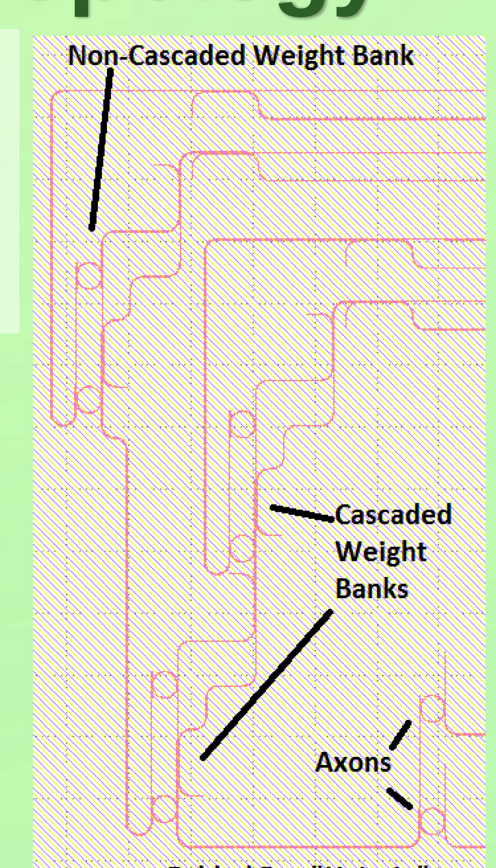


Figure 9: Final CAD of two neurons in a hairpin topology.

Future Work

MDM WDM Full Network

The next logical step is to put all of these experiments together into a full mode-division-multiplexed and wave-division-multiplexed network, which can be applied where a non-optical network would be too slow, such as classification at RF frequencies.

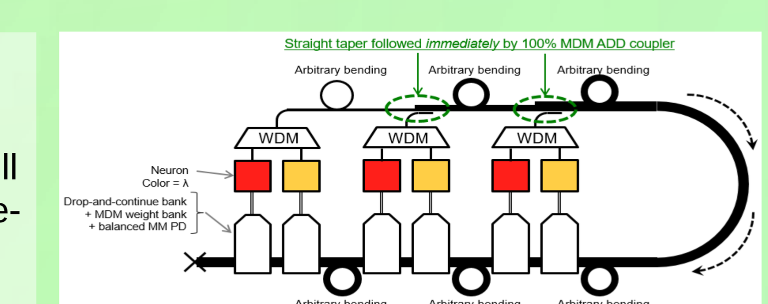


Figure 10: High level representation of a 3-mode 2-wavelength hairpin network.

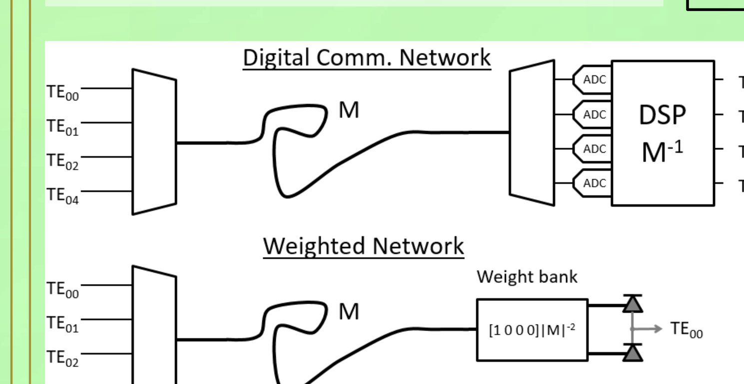


Figure 11: How a multimode weight bank can select a single channel without DSP. Four weight banks can demix all channels.

Application: Channel Demixing

The weight bank itself can also become an important component of integrated digital applications, which cannot not use arbitrarily bent multi-mode waveguides on chip due to the DSP that was required to undo intermodal mixing. For N modes, N weight banks are sufficient to completely invert the intermodal mixing matrix.

I would like to thank everybody in the Lightwave Communications Laboratory for their support and advice. This work was made possible by funding from the School for Engineering and Applied Science and the Department of Electrical Engineering.