Tandem Neural Networks for the Inverse Programming of Linear Photonic Processors

José Roberto Rausell-Campo
Photonics Research Labs, iTEAM
Universitat Politecnica de València

the MZI structures is isolated from the rest of the system ensuring the input light exclusively goes through one of the input waveguides of the MZI, and the output light is guided to one of the outputs of the mesh. The applied current on the MZI is then swept and the optical power at the output is recorded. The dependence of the optical output power with the applied phase shift is:

$$P_{cross} = P_{in}\cos^2(\ ) \qquad (2)$$
$$P_{bar} = P_{in}\sin^2(\ )$$

where P_{in} represent the input optical power, and P_{cross} and P_{bar} represent the optical output power at the bar and cross port, respectively. The obtained curve is fitted using (1) and (2), and the free parameters $_0$ and are estimated. The isolating procedure is architecture dependent. For the case of the rectangular mesh a calibration procedure has been proposed [7]. Once the calibration is finished, a decomposition process is necessary to relate the target matrix with the current that needs to be applied to each phase shifter in the mesh. This method usually assumes ideal components on each of the MZI, which is not the case in practice due to fabrication imperfections. These deviations from the ideal behaviour introduce discrepancies between the target matrix and the real matrix implemented on the chip. A calibration and decomposition algorithm to incorporate the effects of fabrication errors has been demonstrated in [8]. Nevertheless, other sources of errors such as thermal drift, thermal cross-talk and quantization errors cannot be characterized using this procedure.

Recently, a data-driven model has been proposed to include

theration a lnedr aphotonic-2381(ntrrRecenated)-2582caircuit-2381(ith)utp]TJ 0 -11.955 Td [(proio)-4492as thewer25(v)15(ert40(,)-

talibrationcla current

dpplied

ixcp8(compoR0(isplemented)-4R0(iatrix)-4R0(in)-430(the)-340(chaip,-350(she)-340(cFig.-340(c2-350(ap051,)]TJ 9.963 -12.081 Td [(RI)

inplemented the tnd the of

ofe phss-ibl

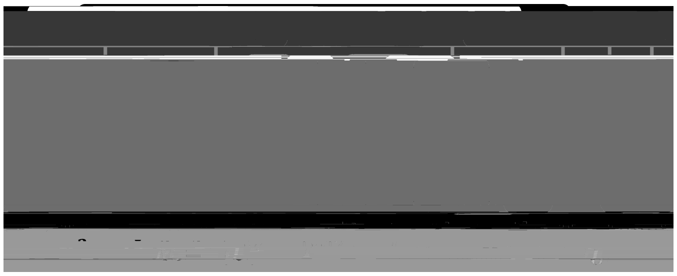

Fig. 3. Picture of the chip used to gather the data to train the tandem network. It is a 3x3 mesh consisting on 3 building blocks using the Clements architecture. The building block (black square) consists of one MZI and one external PS.

To generate the dataset for the tandem network, first we generate a set of 6 currents from a uniform random distribution

a) b) c)

Fig. 6. Tandem neural network results: a) difference between the validation error during the training of the inverse model when using a single feedforward network (red) and the proposed tandem network (blue), b) comparison between the predicted and target weights using the tandem network and c) distribution of the difference between the target weights and the predicted weights (W) for each of the 9 weights of the 3x3 matrix.

TABLE I
MODEL HYPERPARAMETERS