UNIVERSITÀ
DI PAVIA

cea leti

Dipartimento di Fisica

Corso di Laurea Magistrale in Scienze Fisiche

Calibration of an integrated Clement's interferometer

From mathematical modelling to experimental characterization

Tesi per la Laurea di:
Luca Enea Nugnes

Relatore:
Prof. Matteo Galli, Università di Pavia

Correlatore:
Dr. Emanuel Peinke, CEA Leti

A.A. 2025-2026

Abstract

Since transistors are approaching the size of atoms and given an ever increasing necessity of energy-efficient chips, a search for new technological paradigms to complement electronic chips has begun, and integrated photonics is emerging as a competing platform in multiple electronic-dominated technological fields. With the ability to scale optical systems to chip sized devices with high fabrication consistency, integrated photonics already finds widespread application in sensing and communication, and it is also now surging as an alternative computing platform, both for classical and quantum applications. Large integrated interferometric meshes promise to become a fast and low power computing platform, which naturally behaves as a matrix multiplicator. With the rapid emergence of AI agents, of which the main bottleneck as of today is the energy required for their operation and training, devices which intrinsically act as neural network create a genuine interest as potential candidates for a new generation of specialized computing chips for AI.
The same devices made of large integrated interferometric meshes, e.g. Clement's interferometers, are also a new promise in the rush to achieve a quantum computer [2, 15, 20]. While integrated quantum photonics is a relatively new approach when compared to more mainstream solutions, e.g. superconducting qubits or trapped ions [26], the possibility to realize all the required components on a single millimetric device, with the sub-wavelength precision enabled by the decades long industrial experience of CMOS technology, has made it an emerging candidate in the field of quantum information processing.

In the context of photonic computing, the Clement's interferometers plays a central role as a reconfigurable universal interferometer. Its ability to be programmed to any unitary transformation and its intrinsic resistance to propagation losses makes it the main candidate for various photonic computing platforms. Though, fabrication imperfection, together with its complex structure, pose a significant challenge in its operation. Therefore, accurate and efficient calibration algorithms become a necessity to operate an integrated Clement's interferometer with a high fidelity.

This work investigates the limits of the performance achievable by a Clement's interferometer in relation to the entity of the non-idealities of its components. A simulation was developed to understand the effect of fabrication imperfection on the performance of the chip. Also, different calibration algorithms were tested to find one robust to non-idealities. The best performing calibration algorithm was subsequently tested on a physical chip realized at CEA Leti, which allowed a comparison between theoretical and experimental results.

Contents

Chapter 1

Introduction

Between the various platforms competing to become the leading approach for quantum information processing, photon based technologies present some attractive properties. The low interaction between photons and matter translate to high decoherence times and room temperature operation. Also, the propagating nature of photons makes them the main candidate to transport quantum information across large distances to implement quantum communication protocols. On the other hand, photons do not interact with each other, an essential characteristic to implement multiple qubits gates [14].

In photonic quantum information processing, quantum states are encoded in the degrees of freedom of light, including polarization, spatial modes, time bins, and frequency modes. Linear optical elements, such as beam splitters and phase shifters, implement unitary transformations allowing the realization of arbitrary operations [22]. To circumevent the limitations posed by the non interaction between photons, different non deterministic computation protocols have been developed, e.g. the KLM scheme [13], where two qbits operation are realized via measurements.

Integrated photonics has emerged as the leading platform for the implementation of photonic quantum information processing. By transferring optical circuits from free-space optical setups to lithographically fabricated chips, integrated photonics enables stable, compact, and scalable manipulation of light [24, 26]. Modern fabrication techniques allow sub-wavelength precision in waveguide definition, enabling tight optical confinement and dense circuit integration. High index-contrast platforms such as silicon-on-insulator (SOI) allow compact device footprints, while low-loss materials such as silicon nitride offer lower propagation losses for interferometric circuits [27]. Lithium niobate on insulator offers strong electro-optic coefficients, enabling high-speed and efficient phase modulation [26]. These material platforms provide different advantages in terms of optical confinement, loss performance, and modulation capabilities.
A major technological advantage of integrated photonics is its compatibility with complementary metal–oxide–semiconductor (CMOS) fabrication processes. Silicon photonics, in particular, benefits from decades of industrial development originally aimed at microelectronics. CMOS-compatible fabrication enables high-throughput production and tight dimensional control. Also, CMOS technology supports the integration of both photonic and electronic components on the same platform.
Another advantage of integrated photonics is intrinsic stability. In free-space optical setups, interferometric experiments are sensitive to mechanical vibrations. By contrast, waveguide circuits significantly reduce environmental sensitivity. Such stability is essential for reliable operation, especially in circuits requiring precise phase control. In addition to stability, scalability represents a defining feature of integrated platforms. Lithographic fabrication allows the replication of optical components with high reproducibility, making it possible to scale from few-mode interferometers to large meshes.

Programmable interferometric meshes also find a natural application as photonic neural networks (PNN). With the perspective of Moore's law coming to an end and the need for energy-efficient chips to allow larger and larger systems, the performances of classical electronic chip is falling behind the necessities of the ever growing AI industry, and new designs for computing devices specialized in performing tensor multiplication are already coming to the market. Under this context PNNs are also an emerging candidate as a new generation of high-performance computing chips for AI-specific applications [1]. In fact, their structures intrinsically behaves as a multiplication between multiple layers of reprogrammable matrices, de facto mimicking a neural network. These devices promise to achieve faster and more energy efficient computation, allowing increased performances and scalability [28].

The realization of large programmable interferometric networks introduces significant challenges related to control and calibration. Each tunable element must be accurately controlled to reproduce the desired transformation. Fabrication imperfections lead to deviations from ideal splitting ratios, while thermal crosstalk between adjacent phase shifters introduces phase errors. In this context, as the number of controllable parameters usually grows quadratically with the number of modes, accurate and efficient calibration strategies become essential for achieving precise operation.

The work presented in this thesis focuses on the calibration of a Clement's interferometer, a widespread architecture to realize a universal programmable interferometer. A simulation of the chip has been developed to account for non-ideal beam splitter ratios, static phase offsets, and thermal crosstalk. This simulation allowed to perform an in depth study of the performance of the device in function of the entity of its imperfections, and to test multiple calibration algorithms. The best performing calibration algorithm has subsequently been applied to an integrated Clement's interferometer realized at CEA Leti, of which a full characterization has been attempted. The calibration and characterization of the fabricated device allowed to compare the theoretical predictions and the experimental results.

This thesis will first briefly cover the mathematical description of the photonic and electrical components found on the studied device and give a comprehensive description of a Clement's interferometer (Ch 2 and 3). It will then provide an overview on interferometric-pattern-analysis based calibration algorithms (Ch 4), which will then be benchmarked on a numerical simulation of the device (Ch 5). Finally it will report the experimental results obtained from the application of such algorithms to two physical devices and other measurements performed on them (Ch 6).

Chapter 2

The device and its components

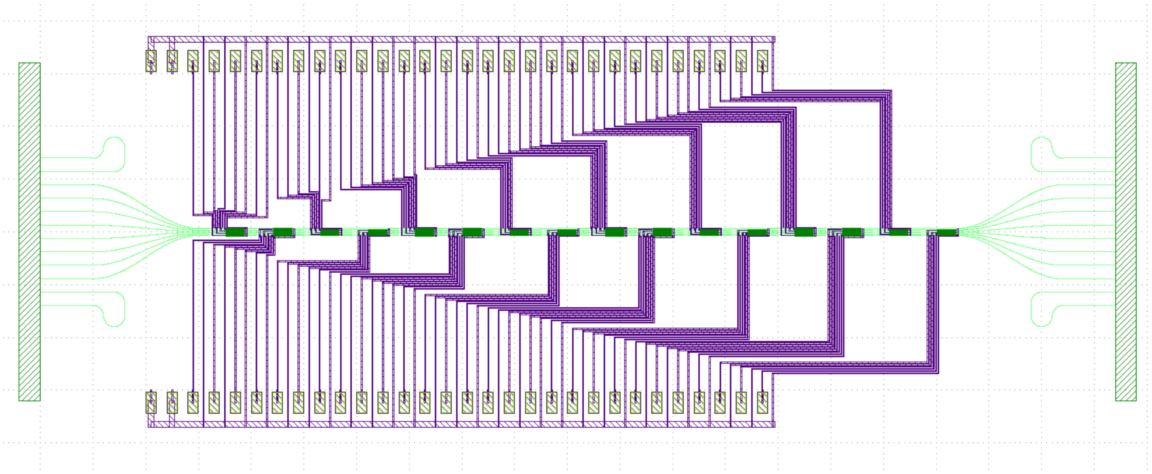
The aim of this work is the characterization, calibration and programming of a photonic computing device: the Clement's interferometer [5], a universal multi-port interferometer (UMI).
UMIs are instruments, composed of multiple interconnected Mach-Zehnder interferometers, able to implement any arbitrary $N \times N$ unitary transformation on a set of N optical modes. They find applications in matrix multiplication for AI or neuromorphics and to implement logical gates for quantum computing.
The idea of the UMI was introduced by Reck et al. in [22], where they showed how to realize one as a triangular grid of MZIs. This first design though had one major flaw that made challenging its experimental realization: optical paths in the device leading to the same MZI had different lengths, which, in presence of propagation losses, degraded the interferometric patterns [5].
Clement's design was proposed to overcome this problem: it abandoned the triangular geometry for a more symmetrical rectangular one. The increased regularity of this new mesh solved the original problem, but it introduced its own downside: a more complex structure makes it more challenging to characterize the MZIs inside it, requiring complex algorithms to calibrate and program the device.
Multiple new architectures have been proposed since to obtain a UMI with the advantages of the Clement's interferometer and a simpler calibration procedure, like the diamond architecture [23] or the Bokun architecture [17]. These designs embed the Reck or Clement's interferometers in larger geometries to get around the problems of their original counterparts, which evidently comes at the cost of a higher number of components and size of the device. All these factors have ultimately lead the Clement's design to be the most widespread way to physically realize a UMI, making the development of robust and precise calibration algorithms for it of particular interest.

The test device used in this study, a 8×8 UMI of which a photo can be seen in fig. 2.1c, was realized as a semiconductor chip in SiN on SiO_2/Si substrate with CMOS technology at CEA Leti, were both the design, shown in fig. 2.1a, and the fabrication took place.
It presents photonics structures (waveguides, edge couplers and directional couplers) and electronics components (heaters, electrical wiring and pads), its meant to work on $1550nm$ light and it acts on 8 input modes.

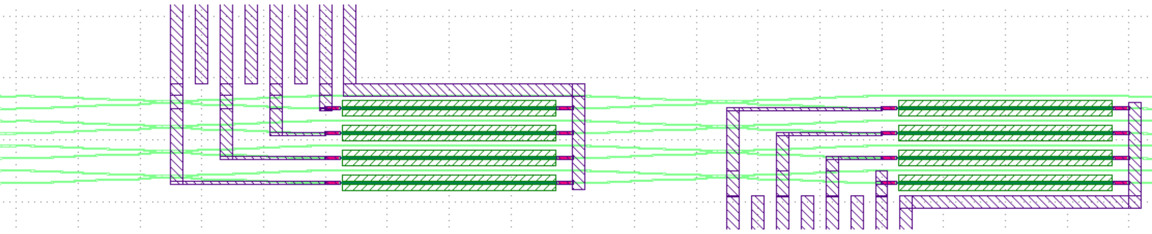
This chapter will briefly cover the working mechanism of the basic components found on the device and the impact of their fabrication imperfections, and it will give a way to describe them mathematically. It will then explain what's a Mach-Zehnder interferometer and how these components form one.

2.1 Waveguides and edge couplers

Waveguides are the fundamental building block of photonic integrated circuits. They exploit the difference of refractive index between materials to confine light, which travels inside them in discrete optical modes, and are used to move light inside the chip.
In the chip under consideration strip waveguides are used, an image of one can be seen in fig. 2.2.



(a) Design of the device (by Emanuel Peinke). Waveguides in green and metal wires in purple.
On the left we see the 8 inputs and 2 alignment loops used for packaging, similarly we can see the 8 ouputs and two alignment loops on the right.
On the top: 30 metal pins to control the thermo-optical phase shifters, the first two from the left connect to the common ground, same on the bottom. In the center: the 8×8 interferometer, in dark green we see the insulating trenches positioned between waveguides were the thermo-optical phase shifter are present, with the purpose of thermally isolating adjacent waveguides



(b) Detail of the design: a column of MZI. We can clearly observe the directional couplers obtained by reducing the gap between waveguides, the heaters (in dark green) on top of alternating waveguides and the isolating trenches (dashed green) between them.

(c) A photo of the as-fabricated chip

Figure 2.1

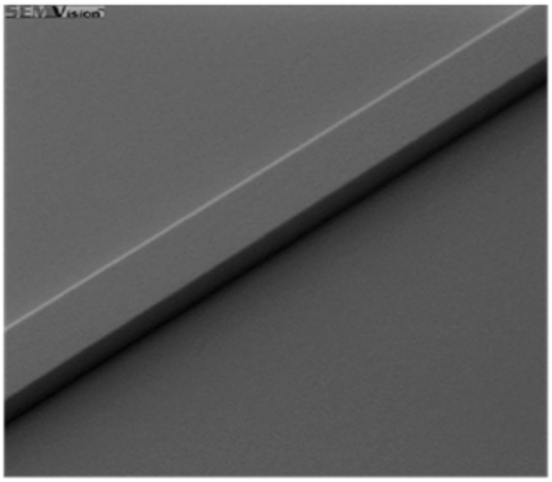


Figure 2.2: A SEM image of a rectangular SiN waveguide on top of SiO_2

The electric field of a monochromatic wave confined in a waveguide can be written as

$$\mathbf{E}(x,y,z,t) = \mathbf{E}(x,y)e^{i(\beta z-\omega t)} \tag{2.1}$$

where β is the propagation constant of the waveguide and z is the propagation axis.
As we can see from Eq. 2.1, light traveling through a waveguide of length L gains a phase

$$\phi = Re[\beta]L \tag{2.2}$$

and faces propagation losses described by $Im[\beta]$, which decreases its amplitude exponentially following

$$\mathbf{E}(x,y,z) = \mathbf{E}(x,y)e^{-Im[\beta]L} \tag{2.3}$$

The SiN platform was chosen for its very low propagation losses at 1550 nm, that were measured at $0.14dB/cm$ for a mono mode waveguide.
An ideal fabrication would lead to perfectly defined rectangular waveguides, but fabrication process imperfections always induce variations in thickness and surface roughness. This makes β non constant over their length, causing a difference of phase to accumulate between different waveguides with the same length [9]. Disregarding losses, the phase introduced by an array of waveguides can be described by a diagonal unitary matrix, which takes the the form

$$\begin{pmatrix} e^{i\phi_0} & 0 & \cdots & 0 \\ 0 & e^{i\phi_1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & e^{i\phi_n} \end{pmatrix} \tag{2.4}$$

Edge couplers are used to couple light in and out of the circuit (i.e. from an optical fiber to an-on chip waveguide and vice versa).They can be described by a coupling coefficient defined as the ratio between the intensity of light entering the edge coupler and the intensity of light coming out of it, called insertion losses.
They depend strongly on the alignment between the input(output) optical fiber and the edge coupler and on the shape and size of the modes entering(exiting) and, as such, they may vary slightly between ports on the chip. The individually different insertion losses will have to be taken into account in the calibration and programming of the real circuit.

2.2 Directional couplers

Directional couplers (DC) are used to couple light between different waveguides and are employed on the chip as beamsplitters.

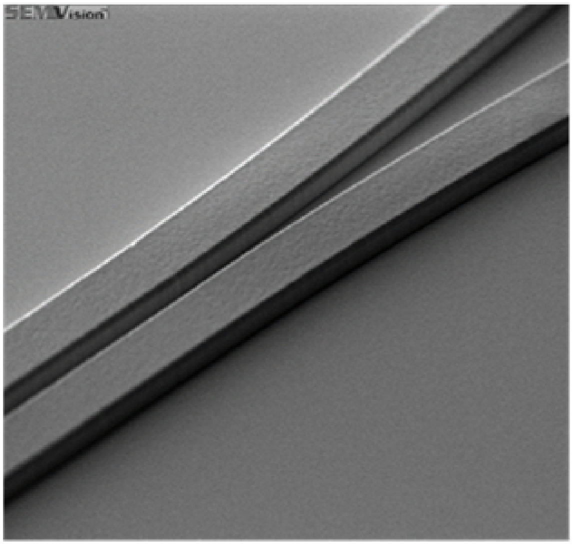


Figure 2.3: A zoom on the beginning of the coupling region with small gap of a directional coupler

They are made by reducing the gap between two waveguides, so light from one can couple to the other. The percentage of light transferred depends on their length, the size of the gap, the wavelength and the polarization of light. A photo of one can be seen in fig. 2.3.

They can be described by a matrix of the form [16]

$$U_{DC} = \begin{pmatrix} \sqrt{1-k} & i\sqrt{k} \\ i\sqrt{k} & \sqrt{1-k} \end{pmatrix} = \begin{pmatrix} \sqrt{\frac{1}{2}-\delta k} & i\sqrt{\frac{1}{2}+\delta k} \\ i\sqrt{\frac{1}{2}+\delta k} & \sqrt{\frac{1}{2}-\delta k} \end{pmatrix} \tag{2.5}$$

where k is the splitting coefficient of the DC and $\delta k = k - \frac{1}{2}$. The DC in this work are designed as to have $k = \frac{1}{2}, \delta k = 0$ but, as it will be discussed later, in reality this value may change.
Since

$$(\sqrt{1-k})^2 + (\sqrt{k})^2 = 1 \tag{2.6}$$

it is possible to rewrite the matrix in terms of the sine and cosine of an angle as [25]

$$U_{DC} = \begin{pmatrix} \cos(\gamma) & i\sin(\gamma) \\ i\cos(\gamma) & \sin(\gamma) \end{pmatrix} = \begin{pmatrix} \cos(\frac{\pi}{4}+\delta\gamma) & i\sin(\frac{\pi}{4}+\delta\gamma) \\ i\sin(\frac{\pi}{4}+\delta\gamma) & \cos(\frac{\pi}{4}+\delta\gamma) \end{pmatrix} \tag{2.7}$$

where $\delta\gamma = \gamma - \frac{\pi}{4}$. By comparing the two matrices we find that the two descriptions are related by

$$2\delta k = \sin(2\delta\gamma) \tag{2.8}$$

so, for small variations from the ideal value, $\delta k \simeq \delta\gamma$. To stay coherent with [6] we'll keep working with the δk representation, and we'll refer to the shifts δk as unbalancements.

2.3 Thermo-Optical phase shifters

A thermo-optical phase shifter (TOPS) relies on the thermo-optic effect (the dependence of the refractive index of a material on its temperature) to induce a phase shift. It consists of a resistor placed on top of the waveguide: when current flows through it, it heats up due to Joule effect, also increasing the temperature of the waveguide below changing its refractive index.

The introduced phase change for a small change in temperature follows [11]

$$\Delta\phi = \frac{2\pi L}{\lambda_0}\frac{dn}{dT}\Delta T \tag{2.9}$$

where L is the TOPS length, n is the refractive index of the waveguide and λ_0 is the wavelength. If we suppose that the change in temperature of the waveguide is directly proportional to the energy dissipated from the heater ($\Delta T = kP$, k being an arbitrary constant) we can rewrite this relation as

$$\Delta\phi = k'P \tag{2.10}$$

where $k' = k\frac{2\pi L}{\lambda_0}\frac{dn}{dT}$. Keeping in mind that the power dissipated trough Joule effect is

$$P = IV = \frac{V^2}{R} \tag{2.11}$$

allows to write the phase voltage relation of a TOPS as

$$\Delta\phi = k''V^2 \tag{2.12}$$

This relation doesn't take into account the change of resistance of the TOPS when it heats up, in fact the resistance of a resistor changes with its temperature following

$$R = R_0(1 + \alpha\Delta T) \tag{2.13}$$

where α for a conductor is a positive coefficient and R_0 is the resistance at the reference temperature. The increase of the resistance with temperature reduces the current flowing through the device at a given voltage and the power dissipated with it.
At the equilibrium the power dissipated by the device must be equal to the power dissipated through Joule effect, if we ignore radiation cooling this translates to:

$$\frac{V^2}{R_0(1 + \alpha\Delta T)} = cS\Delta T \tag{2.14}$$

where S is the surface of the device and c is the thermal dissipation coefficient. Solving for ΔT leads to

$$\Delta T = \frac{\sqrt{1 + \left(\frac{V}{V_0}\right)^2} - 1}{2\alpha} \tag{2.15}$$

with $V_0 = \sqrt{\frac{cSR_0}{4\alpha}}$.

If we substitute it in eq. 2.9, assuming the shift in temperature of the waveguide to be proportional to the one of the heater, we end with

$$\Delta\phi = k'''\left(\sqrt{1 + \left(\frac{V}{V_0}\right)^2} - 1\right) \tag{2.16}$$

while it leads to a current voltage relation of the form

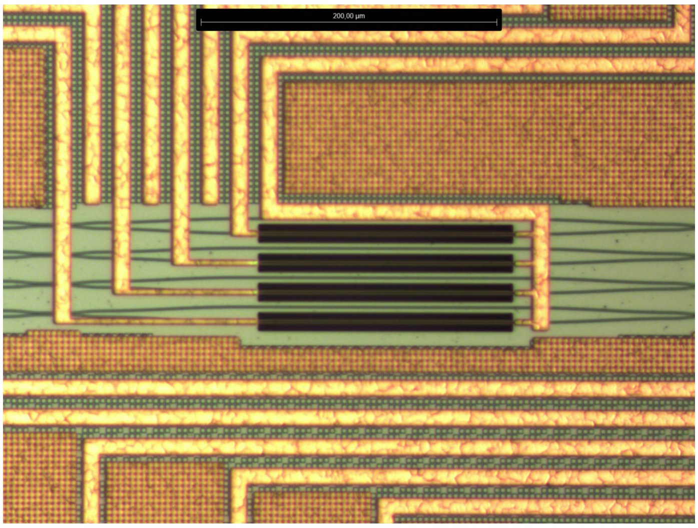
$$I = \frac{V}{R} = \frac{V}{R_0(1 + \alpha\Delta T)} = \frac{V}{\frac{R_0}{2}\left(1 + \sqrt{1 + \left(\frac{V}{V_0}\right)^2}\right)} \tag{2.17}$$

The heating of the resistor changes the expected quadratic phase voltage relation to a more complex one. We observe empirically that is possible to approximate the behavior of the TOPS with a curve of the form

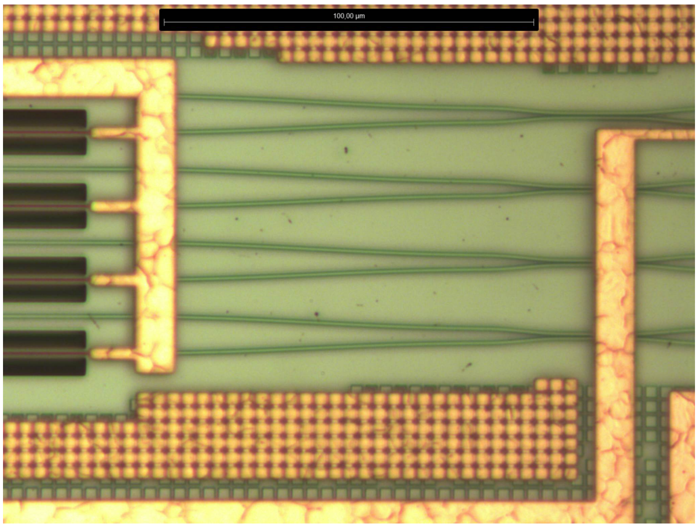
$$\Delta\phi = kV^n \tag{2.18}$$

where $n <= 2$. In addition, as we will explain later in Ch. 3, we can 'move' the phase offset due to imperfections in the waveguides into the phase response of the phase shifters as a constant offset, which leads to relations of the form

$$\Delta\phi = \phi_0 + kV^n \tag{2.19}$$



(a) A column of thermo-optical phase shifter (TOPS) . Waveguides in dark green, metal contact in bright yellow and the insulating trenches in black.



(b) Detail of a column of TOPS, we can see in red the heaters positioned on top of the waveguides. To the right we also see 4 DC.

Figure 2.4

for the empirical curve and

$$\Delta\phi = \phi_0 + k''' \left(\sqrt{1 + \left(\frac{V}{V_0} \right)^2} - 1 \right) \tag{2.20}$$

for the theoretical one. In both cases 3 parameters are needed to describe the phase-voltage relation of a TOPS. An experimental test of the validity of Eq. 2.17 and Eq. 2.19 can be found in Ch. 7.1 .

When a TOPS is working, in addition to the waveguide below it, it also heats surrounding ones introducing unwanted phase offsets in them, a phenomenon known as thermal crosstalk.
To mitigate it deep trenches are dug between waveguides where the heaters are present to reduce the heat conduction in the space between them. Photos of a column of TOPS can be seen in fig. 2.4.

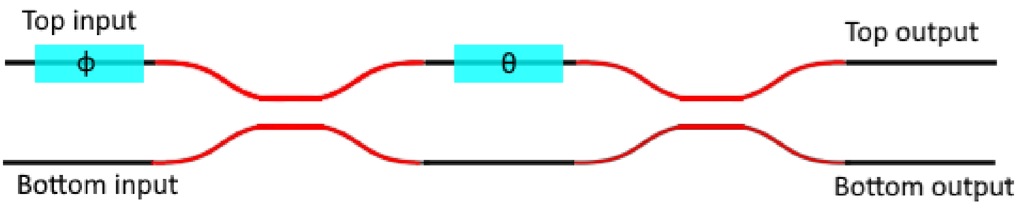


Figure 2.5: A schematic picture of an MZI as realized on the chip. In black waveguides, directional couplers in red and phase shifters in light blue.

2.4 The Mach-Zehnder Interferometer

A Mach-Zehnder Interferometer (MZI) is an optical interferometer formed by two 50:50 beamsplitters (BS) and a phase shifter between them, a schematic representation of it can be see in Fig. 2.5.
The phase shifter controls the type of interference happening at the second beamsplitter, allowing to manipulate the repartition of light between the outputs. We will call it the internal phase shifter and we'll refer to the phase it introduces as θ.
The matrix representing the unitary transformation T of an MZI is [16]

$$T = e^{i\frac{\theta}{2}} \begin{pmatrix} \sin\frac{\theta}{2} & \cos\frac{\theta}{2} \\ \cos\frac{\theta}{2} & -\sin\frac{\theta}{2} \end{pmatrix} \tag{2.21}$$

In the context of the universal multiport interferometer another phase shifter is added to each MZI on one of the input or output arms. We will refer to it as the external phase shifter and to the phase it introduces as ϕ. The placement of the external phase shifter(on one input or on one output) is a matter of preference, as we will see. In the chip under consideration it's on the top output, but to stay coherent with [5, 6, 22] we will consider it to be on the top input arm for the theoretical discussions.
Adding this phase shift changes the transfer matrix of the MZI to

$$T = e^{i\frac{\theta}{2}} \begin{pmatrix} e^{i\phi}\sin\frac{\theta}{2} & \cos\frac{\theta}{2} \\ e^{i\phi}\cos\frac{\theta}{2} & -\sin\frac{\theta}{2} \end{pmatrix} \tag{2.22}$$

Two configuration are especially important on an MZI: the bar state, when $\theta = \pi$, where the device acts as the identity(minus a phase factor) and the cross state, when $\theta = 0$, in which it acts as the swap operator. The two states can be clearly seen in fig. 2.6.

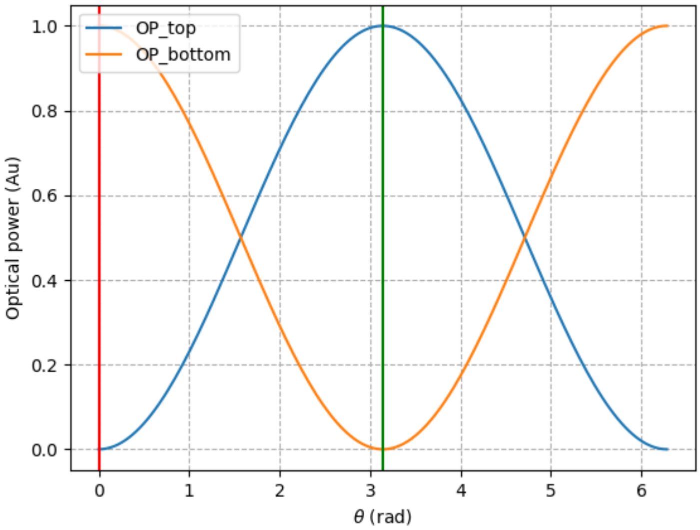


Figure 2.6: Output optical powers of an MZI in function of θ when light is shone in the top input. The vertical line in red highlights the cross state, while the one in green the bar state

Lets also consider two particular situations:

- When light is shone in the top input the output powers(normalized) follow

$$P_1 = \frac{1}{2} - \frac{1}{2}\cos(\theta) \tag{2.23}$$

- When both inputs are illuminated with same intensity and phase and $\theta = \frac{\pi}{2}$

$$P_1 = \frac{1}{2} - \frac{1}{2}\cos(\phi) \tag{2.24}$$

P_2 can be simply found as $1 - P1$ for conservation of energy, where P_1 and P_2 are respectively the top and bottom output optical powers. These two special cases have a simple functional form of the type

$$f(x) = A + B\cos(x) \tag{2.25}$$

that, as will be shown later, can be easily inverted to obtain the phase voltage relation of the phase shifters.

2.5 The imperfect MZI

In the chip, MZIs are realized through DC, TOPS and waveguides. The fabrication imperfections of these components will change the behavior of the real MZIs from the ideal one. As discussed above, these imperfection come in two types: phase differences accumulating between waveguides and deviations from the ideal 50:50 splitting coefficient of the DC. It will be shown in ch. 3 that in a Clement's interferometer we can always consider these phase offsets as if they were introduced by the phase shifters of the MZIs, and, as such, this section will be dedicated to the study of the effect of unbalancement on the MZI.

The matrix T representing an unbalanced MZI can be found by multiplying together the transfer matrices of its components as

$$T = \begin{pmatrix} \sqrt{1-k_2} & i\sqrt{k_2} \\ i\sqrt{k_2} & \sqrt{1-k_2} \end{pmatrix} \begin{pmatrix} e^{i\theta} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \sqrt{1-k_1} & i\sqrt{k_1} \\ i\sqrt{k_1} & \sqrt{1-k_1} \end{pmatrix} \begin{pmatrix} e^{i\phi} & 0 \\ 0 & 1 \end{pmatrix}$$

from which it follows

$$\begin{aligned} T_{11} &= e^{i\phi}(Ae^{i\theta} - B) \\ T_{12} &= i(Ce^{i\theta} + D) \\ T_{21} &= ie^{i\phi}(C + De^{i\theta}) \\ T_{11} &= (A - Be^{i\theta}) \end{aligned} \tag{2.26}$$

where

$$\begin{aligned} A &= \sqrt{(1-k_1)(1-k_2)} = \sqrt{(\frac{1}{2} - \delta k_1)(\frac{1}{2} - \delta k_2)} \\ B &= \sqrt{(k_1)(k_2)} = \sqrt{(\frac{1}{2} + \delta k_1)(\frac{1}{2} + \delta k_2)} \\ C &= \sqrt{(k_1)(1-k_2)} = \sqrt{(\frac{1}{2} + \delta k_1)(\frac{1}{2} - \delta k_2)} \\ D &= \sqrt{(1-k_1)(k_2)} = \sqrt{(\frac{1}{2} - \delta k_1)(\frac{1}{2} + \delta k_2)} \end{aligned} \tag{2.27}$$

By substituting $\theta = 0$ and $\theta = \pi$, it can be observed that T_{12} and T_{21} won't vanish. In general, an imperfect MZI will not act as the identity or the swap operator in the bar and cross state respectively. Two values are often considered as figures of merit for the 'ideality' of an MZI: the extinction ratio in the bar and cross states, defined as:

$$\begin{aligned} ER_{bar} &= \frac{P_2(\theta = \pi)}{P_1(\theta = \pi)} \\ ER_{cross} &= \frac{P_1(\theta = 0)}{P_2(\theta = 0)} \end{aligned} \tag{2.28}$$

when light is shone in the top input. They are both 0 for an ideal MZI.

2.5.1 Intensity output

It's interesting to study how the output optical powers behave in function of the parameters of the problem. If the input light is represented by a vector $v = \begin{pmatrix} a \\ b \end{pmatrix}$ the top intensity output is going to be

$$P_1 = |aT_{11} + bT_{12}|^2 = |ae^{i\phi}(Ae^{i\theta} - B) + ib(Ce^{i\theta} + D)|^2 \tag{2.29}$$

which by Euler's identity and trigonometric formulas can be rewritten as

$$\begin{aligned} P_1 = |&(aA\cos(\theta + \phi) - aB\cos(\phi) - bC\sin(\theta)) + \\ &i(aA\sin(\theta + \phi) - aB\sin(\phi) + bC\cos(\theta) + bD)|^2 \\ = &(aA\cos(\theta + \phi) - aB\cos(\phi) - bC\sin(\theta))^2 + \\ &(aA\sin(\theta + \phi) - aB\sin(\phi) + bC\cos(\theta) + bD)^2 \end{aligned} \tag{2.30}$$

by developing the squares, trigonometric formulas and noticing that $AB = CD$

$$\begin{aligned} P_1 = a^2(A^2+B^2) + b^2(C^2+D^2) + 2AB\cos(\theta)(b^2-a^2) + \\ 2ab(\sin(\phi)(AC-BD) + BC\sin(\theta-\phi) + AD\sin(\theta+\phi)) \end{aligned} \tag{2.31}$$

which by simple substitution of the values of A, B, C, D becomes

$$\begin{aligned} P_1 = a^2 + (b^2-a^2)\left[\frac{1}{2} - 2\delta k_1\delta k_2 + 2\sqrt{(\frac{1}{4}-\delta^2 k_1)(\frac{1}{4}-\delta^2 k_2)}\cos(\theta)\right] \\ + 2ab\left[\sin(\phi)(-2\delta k_2)\sqrt{\frac{1}{4}-\delta^2 k_1} + \right. \\ \left. \sqrt{\frac{1}{4}-\delta^2 k_2}\left((\frac{1}{2}+\delta k_1)\sin(\theta-\phi) + (\frac{1}{2}-\delta k_1)sin(\theta+\phi)\right)\right] \end{aligned} \tag{2.32}$$

The bottom port intensity can be derived as $P_2 = a^2 + b^2 - P_1$. Lets consider some particular cases of this formula.

$\mathbf{b = 0}$

How does the MZI behave when just one input is illuminated?
For P_1 and P_2 we get

$$\begin{aligned} P_1 = a^2\left(\frac{1}{2} + 2\delta k_1\delta k_2 - 2\sqrt{(\frac{1}{4}-\delta k_1^2)(\frac{1}{4}-\delta k_2^2)}\cos(\theta)\right) \\ P_2 = a^2\left(\frac{1}{2} - 2\delta k_1\delta k_2 + 2\sqrt{(\frac{1}{4}-\delta k_1^2)(\frac{1}{4}-\delta k_2^2)}\cos(\theta)\right) \end{aligned} \tag{2.33}$$

In Fig 2.7 we can observe the plot of the power in function of θ for $k_1 = 0.6, k_2 = 0.8$. The output power still has a functional form of the shape of Eq 2.25.

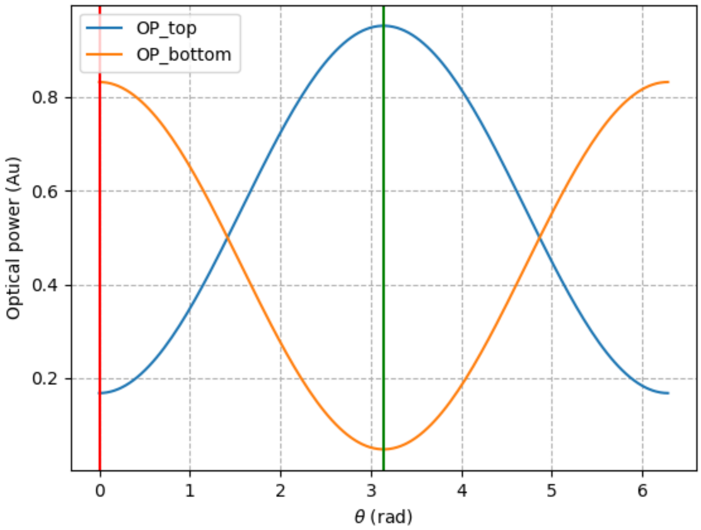


Figure 2.7: Power output in function of θ for $k_1 = 0.6, k_2 = 0.8, a = 1, b = 0$.
When compared to Fig. 2.6, we can see that the curves appear scaled and shifted one towards the top and one towards the bottom. Independently on the applied phase, the bar and cross state cannot be produced

$\mathbf{a = b}$

This configuration greatly simplifies the functional form of the output powers, while maintaining dependencies on all of the parameters of the problem.

$$P_1 = A + B\cos(\theta) + C\sin(\theta) \tag{2.34}$$

where

$$\begin{aligned} A &= a^2(1 - 4\delta k_2\sqrt{\frac{1}{4} - \delta k_1^2}\sin(\phi)) \\ B &= -4a^2\delta k_1\sqrt{\frac{1}{4} - \delta k_2^2}\sin(\phi) \\ C &= 2a^2\sqrt{\frac{1}{4} - \delta k_2^2}\cos(\phi) \end{aligned} \tag{2.35}$$

$\mathbf{a = b, \boldsymbol{\theta} = \boldsymbol{\pi}/2}$

This leads to

$$\begin{aligned} P_1 &= a^2\left(1 - 4\delta k_2\sqrt{\frac{1}{4} - \delta^2 k_1}\sin(\phi) + 2\cos(\phi)\sqrt{\frac{1}{4} - \delta^2 k_2}\right) \\ &= a^2\left(1 - 2\sqrt{\frac{1}{4} - \delta^2 k_1 \delta^2 k_2}\ cos(\phi - \chi)\right) \end{aligned} \tag{2.36}$$

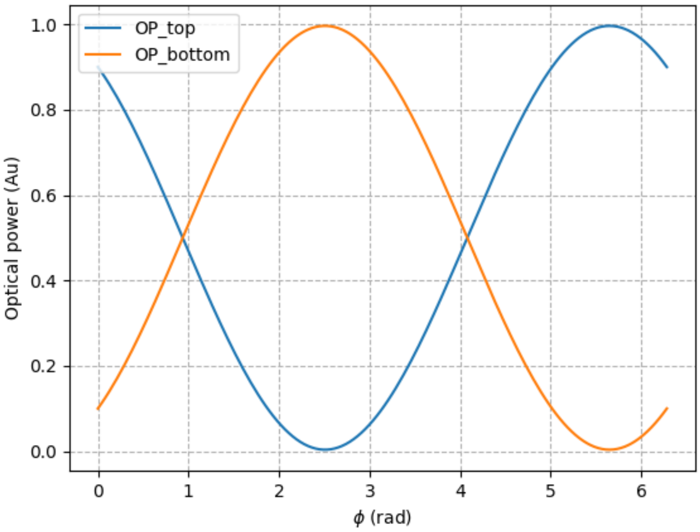


Figure 2.8: Optical power outputs in function of ϕ for $k_1 = 0.6, k_2 = 0.8, a = b, \theta = \frac{\pi}{2}$. When compared to Fig. 2.6, we can see that the curves appear shifted towards the left.

where

$$\begin{aligned} \cos(\chi) &= \frac{\sqrt{\frac{1}{4} - \delta^2 k_2}}{\sqrt{\frac{1}{4} - \delta^2 k_1 \delta^2 k_2}} \\ \sin(\chi) &= \frac{-2\delta k_2 \sqrt{\frac{1}{4} - \delta^2 k_1}}{\sqrt{\frac{1}{4} - \delta^2 k_1 \delta^2 k_2}} \end{aligned} \tag{2.37}$$

The optical power outputs don't follow Eq. 2.25 since the unbalancements act as a shift in the phase, a plot of it for $k_1 = 0.6$ and $k_2 = 0.8$ can be seen in fig 2.8

2.5.2 Programming an imperfect MZI

Let us consider a simple problem, programming a specific matrix on an MZI. For an ideal device we can easily find the necessary phases ϕ and θ, but what about an unbalanced MZI?
If we try to apply the same phases, as one can see from the previous section, the resulting transformation will differ from the desired one. So, is there a way to correct the error introduced by non-idealities? Are there some transformations that's not possible to implement anymore, even in principle?
The answer, as shown in [4], is yes to both: it's possible, if the unbalacements are known, to find new phases which will restore the original desired transformation, and the set of implementable matrices will be 'smaller' then for a perfect MZI.
An ideal MZI implements a transformation $T(\phi, \theta)$, represented by

$$T\left(\phi, \theta\right) = e^{i\frac{\theta}{2}} \begin{pmatrix} e^{i\phi} \sin\frac{\theta}{2} & \cos\frac{\theta}{2} \\ e^{i\phi} \cos\frac{\theta}{2} & -\sin\frac{\theta}{2} \end{pmatrix} \tag{2.38}$$

In presence of unbalancements, using the angular notation for the unbalancement, applying the same phases will yield a different transformation

$$T'(\alpha, \beta, \phi, \theta) = \begin{pmatrix} \cos(\beta) & i\sin(\beta) \\ i\sin(\beta) & \cos(\beta) \end{pmatrix} T(\phi, \theta) \begin{pmatrix} \cos(\alpha) & ie^{-i\phi}\sin(\alpha) \\ ie^{i\phi}\sin(\alpha) & \cos(\alpha) \end{pmatrix} \tag{2.39}$$

Where α, β are the splitting coefficients of the DCs in the MZI. To retrieve the original transformation it's necessary to find a new pair ϕ', θ' such that $T'(\alpha, \beta, \phi', \theta') = T(\phi, \theta)$, which translates to

$T'(\alpha,\beta,\phi',\theta')_{i,j} = T(\phi,\theta)_{i,j} \quad i = 1,2 \quad j = 1,2.$
Remembering that the matrices are complex valued we have to equate both the amplitude and the phase of every element. The equivalence between the amplitudes leads to

$$\theta' = 2\arcsin\sqrt{\frac{\sin^2(\theta/2) - \sin^2(\alpha+\beta)}{\cos^2(\alpha-\beta)} - \sin^2(\alpha+\beta)} \tag{2.40}$$

which has a solution only if $\sin^2(\theta/2) > \sin^2(\alpha+\beta)$ and if $\sin^2(\theta/2) < \cos^2(\alpha-\beta)$, which restricts the range of θ to

$$2|\alpha+\beta| < \theta < \pi - 2|\alpha-\beta| \tag{2.41}$$

If θ falls into this range we can substitute the value of θ' in the matrix to get

$$T'(\alpha,\beta,\phi,\theta') = ie^{i\theta'/2}\begin{pmatrix} e^{i\xi_b} & 0 \\ 0 & e^{i\xi_d} \end{pmatrix}\begin{pmatrix} e^{i\phi+\xi_a-\xi_b}\sin\frac{\theta}{2} & \cos\frac{\theta}{2} \\ e^{i\phi+\xi_a-\xi_b}\cos\frac{\theta}{2} & -\sin\frac{\theta}{2} \end{pmatrix} \tag{2.42}$$

where $\xi_a, \xi_b, \xi_c, \xi_d$ are the phases of the four elements of $T'(\alpha,\beta,\phi,\theta')$ (minus a factor $ie^{i\theta'/2}$, ξ_c has been rewritten if function of the other three in eq. 2.42)
To match the phases one has then to apply two additional phase shifts $\psi_1 = -\xi_b + (\theta-\theta')/2$ and $\psi_2 = -\xi_d + (\theta-\theta')/2$ to the top and bottom output modes(we point out that this is unnecessary if only intensity of the outputs will be measured, one could argue otherwise that it's not really possible to restore the starting matrix without adding new components to the device), and to choose $\phi' = \phi + \xi_b - \xi_a$. Comparing the expression for the matrices gives the values of ϕ', ψ_1, ψ_2 to be:

$$\begin{aligned} \phi' =& \phi + \arctan\left(\frac{\sin(\alpha-\beta)}{\cos(\alpha+\beta)}\tan(\theta'/2)\right) \\ & - \arctan\left(\frac{\sin(\alpha+\beta)}{\cos(\alpha-\beta)}\cot(\theta'/2)\right) \end{aligned} \tag{2.43}$$

$$\psi_1 = -\arctan\left(\frac{\sin(\alpha-\beta)}{\cos(\alpha+\beta)}\tan(\theta'/2)\right) + (\theta-\theta')/2 \tag{2.44}$$

$$\psi_2 = \arctan\left(\frac{\sin(\alpha+\beta)}{\cos(\alpha-\beta)}\cot(\theta'/2)\right) + (\theta-\theta')/2 \tag{2.45}$$

If the value of θ falls off the range of Eq. 2.41, its not possible to achieve $T(\theta,\phi)$ with the considered MZI, its best approximation will be achieved by setting θ to 0 or π, whichever is the closest. This bound connects the implementable matrices on an imperfect MZI to its unbalancements.
From Eq. 2.41 we can also derive the necessary condition to be able to implement identity(I) or the swap operator(X), in particular for the former it's necessary that $\delta k_1 = \delta k_2$ while for the latter that $\delta k_1 = -\delta k_2$. Only an ideal MZI can implement both of them.
A commonly used figure of merit between matrices is the fidelity [5, 8], defined as

$$F(A,B) = \left|\frac{Tr\left[AB^\dagger\right]}{\sqrt{Tr\left[AA^\dagger\right]Tr\left[BB^\dagger\right]}}\right|^2 \tag{2.46}$$

It is possible to calculate the maximum fidelity achievable between the bar(cross) state and I(X) in function of the unbalancements, which finds

$$\begin{aligned} F(I)_{max} &= \frac{1}{2} + 2\delta k_1\delta k_2 + 2\sqrt{(\frac{1}{4}-\delta k_1^2)(\frac{1}{4}-\delta k_2^2)} \\ F(X)_{max} &= \frac{1}{2} - 2\delta k_1\delta k_2 + 2\sqrt{(\frac{1}{4}-\delta k_1^2)(\frac{1}{4}-\delta k_2^2)} \end{aligned} \tag{2.47}$$

a plot of the level curves of $F(I)_{max}$ can be seen in Fig. 2.9

An analogous argument can be made for an MZI with the external phase shifter on the output, the result is similar but requires the two additional phase shifts to be added in front of the device.

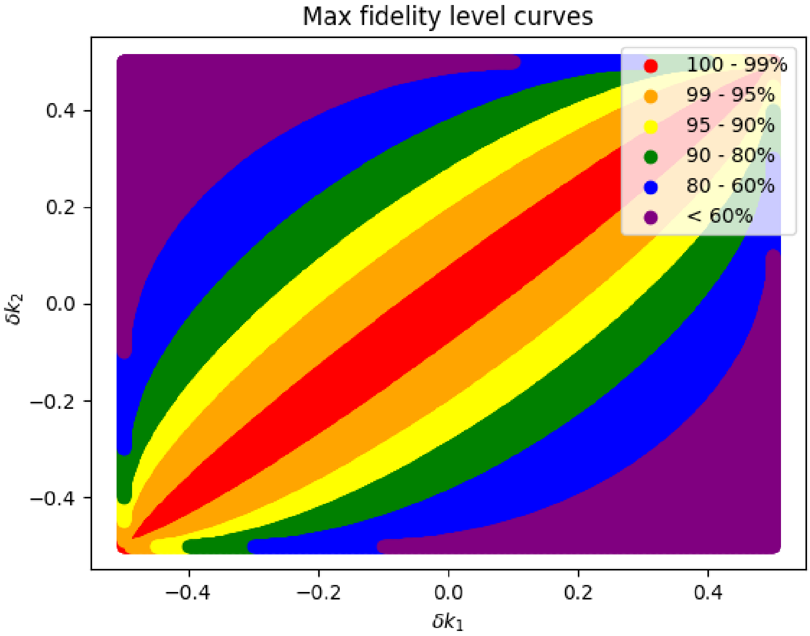


Figure 2.9: A plot of the curve levels of the maximum fidelity with the identity in function $\delta k1$ and $\delta k2$, we clearly see that $F = 1$ along the main diagonal

Chapter 3

Mathematical description of a universal multiport interferometer

Any unitary matrix U can be decomposed as the product of two levels unitary matrices $U_{m,n}$: unitary matrices acting non trivially only on two or fewer vector components, where m, n refers to those components [18].
The principle behind this decomposition is that it's possible to find a two level unitary $U_{m,n}$ such that

$$U' = U_{m,n}U \tag{3.1}$$

has one element equal to zero, an operation called nullification.
By repeating this process a series of $U_{m,n}$ can be found such that

$$I = \prod_{m,n} U_{m,n}U \tag{3.2}$$

allowing to rewrite U as

$$U = \prod_{m,n} U^{-1}_{m,n} \tag{3.3}$$

The basic idea of a UMI is to recreate this decomposition using matrices implementable by an MZI, which can implement only a subspace of the two level unitaries.
It can be shown [22] that the matrix $T_{m,n}(\phi, \theta)$ implemented by a MZI can be used to nullify an element of a unitary matrix U from the right ($U' = UT_{m,n}(\phi, \theta)$) for a suitable choice of ϕ and θ, while it's inverse $T^{-1}_{m,n}(\phi, \theta)$ can do the same from the left($U' = T^{-1}_{m,n}(\phi, \theta)\,U$, from now on we will leave implicit the dependence of this matrices from the phases).

This allows to rewrite U as

$$U = \left(\prod_{m,n} T^{-1}_{m,n}\right) D \left(\prod_{m,n} T_{m,n}\right) \tag{3.4}$$

where D is a unitary diagonal matrix and the products are over a suitable choice of m, n. It can also be proven [22] that $T^{-1}_{m,n}D = D'T_{m,n}$, where D' is a new diagonal unitary matrix, which allows to shift it to the left of the equation, so that

$$U = D' \prod_{m,n} T_{m,n} \tag{3.5}$$

giving a way of decomposing U as a product of $T_{m,n}$ and of a diagonal unitary matrix.
At the end of this procedure, we are left with a series of phases that can be used to physically implement U on a succession of MZIs, followed at the end by a final array of phase shifters.
It should be noticed that for most applications the phase of the output is irrelevant since usually only the intensity of the light is measured and the last row of phase shifters can be left out.
It's also important to observe that to nullify an element U_{ij} we can both use $T_{m,n}$ with $m = i$ or

$n = i$, or $T^{-1}_{m,n}$ with $m = j$ or $n = j$.
The freedom in the choice of the values of m, n and in the order of nullification allows for different geometries for the mesh of MZIs.

3.1 Clement's architecture

The most widespread design for a UMI is the Clement's architecture [5], it is based on a rectangular geometry and employs $\frac{N(N-1)}{2}$ MZIs, the minimal number needed to create a UMI, where N is the number of input modes. Also, improving on it's predecessor [22], all paths leading to the same MZI have the same length, minimizing distortions of the interferometric pattern due to losses [5].
The order of nullification to achieve this geometry is obtained by nullifying elements of the desired matrix diagonally starting from the bottom left corner and switching between T and T^{-1} at the end of each diagonal. The choice of m, n is made to create a correspondence between the diagonals in the matrix and the diagonals in the rectangular mesh. A schematic representation of how this works for a 5×5 matrix can be seen in Fig 3.1.

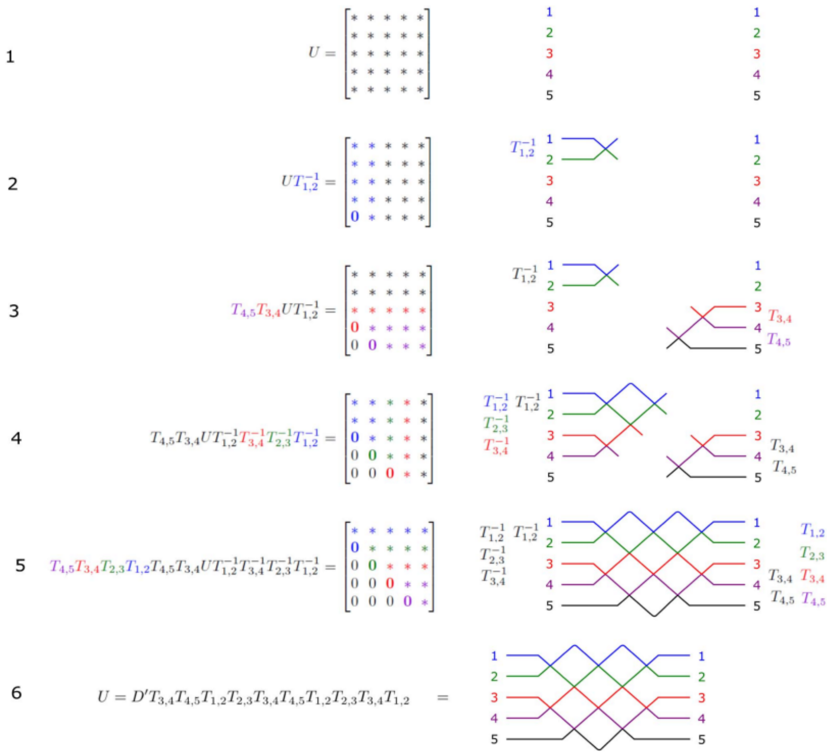


Figure 3.1: A schematic representation of how the clement decomposition works, copied from [5]. On the left the order of nullification, while on the right the correspondence with the rectangular mesh. Each crossing represents an MZI.

3.1.1 Two designs for the MZIs

As said above, the external phase shifter might be positioned both on an input or an output mode. The two configurations can't implement the same set of transformations. The former, by definition, implements the transformations we called $T_{m,n}$, while the latter implements the ones we referred to as $T^{-1}_{m,n}$ (minus a global phase factor which doesn't matter for the nullification of an element of U). If we use the second configuration we will have to use a different decomposition, from Eq. 3.4 we can push the diagonal towards the right, and rewrite U as

$$U = \left(\prod_{m,n} T_{m,n}^{-1} \right) D' \tag{3.6}$$

This allows to keep the same geometry for the mesh when switching from the two configurations, requiring only to move the final array of phase shifters from the end to the start of the circuit.

3.1.2 Geometry and nomenclature

Following the Clement's decomposition, the MZIs end up arranged in a rectangular pattern with N rows and N columns. If N is even the number of MZI in each column alternates between $N/2$ and $N/2-1$, while it stays constant at $\frac{N-1}{2}$ if N is odd. A schematic of the mesh can be seen in fig. 3.2.

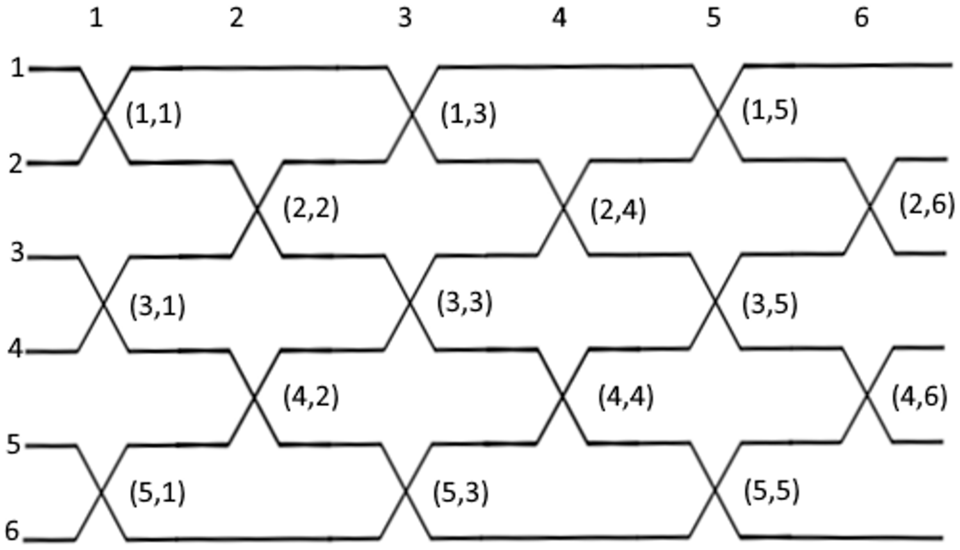


Figure 3.2: A schematic of a 6 × 6 mesh. Each crossing represents an MZI. Columns ,rows and nomenclature are indicated.

To identify the rows, waveguides are numbered from 1 to N (in a physical device this can be done in two possible ways and will decide what is considered the 'top' and 'bottom' of the device. To stay coherent with the theoretical discussion, where the TOPS have been considered in the 'top' arm, waveguide 1 should be the one with TOPS on top).

If the number of modes is even, column 1, at the left of the chip in 3.2, will have an even number of MZI, while if the number of modes is odd the system is symmetric, except for the position of the external phase shifter. This allows to identify each MZI by a pair of numbers, where the first one represents the number of the waveguide of their top arm and the second one their column.

It is immediate to observe that not every pair of numbers has a corresponding MZI, for example (1,2) and (2,1) don't have a corresponding MZI in any mesh size (It is easy to see since they don't appear in the 3×3 mesh, and a mesh of size N contains every smaller one).

3.1.3 Two directions of use

The device is normally intended to be used with column 1 as input but, given its ability to implement any unitary matrix, can also be used the other way around.

To program a matrix U when column N is used as input, it's sufficient to set the device to act as U^T if it were used the standard way. The reason for this can clearly be seen by taking into consideration eq. 3.5 and that the transfer matrices of phase shifter and DC are invariant under transposition, which implies

$$U^T = D\left(\prod_{m',n'} T^{-1}_{m',n'}\right) \tag{3.7}$$

where the order of the product has been reversed. Since the right side of the equation is the transfer matrix of the device used backwards, transposing is equivalent to changing the input side.

3.1.4 Clement's decomposition with unbalanced MZIs

The Clement's algorithm finds a collection of phases that would implement the target matrix on an ideal device but, as we showed previously, an imperfect device can behave in a significantly different way than an ideal one.
If the values of the splitting coefficients of each DC in the device are known there is a simple procedure, introduced in [4], that can be followed to correctly program the device.
Lets take into consideration eq. 3.5, remembering from section 2.5.2 that

$$T_{m,n}\left(\phi,\theta\right) = D_{m,n}\left(\psi_1,\psi_2\right) T'_{m,n}\left(\phi',\theta'\right) \tag{3.8}$$

where $T'_{m,n}$ is the transfer matrix of an imperfect MZI and $D_{m,n}$ are diagonal unitary matrices collecting two additional phases ψ_1, ψ_2, we can decompose any unitary matrix as

$$U = D\left(\prod_{m,n} D_{m,n}\left(\psi_1,\psi_2\right) T'_{m,n}\left(\phi',\theta'\right)\right) \tag{3.9}$$

where $\theta', \phi', \psi_1, \psi_2$ are given respectively by equations 2.40, 2.43, 2.44 , 2.45.
It can also be shown [4] that

$$T'_{m,n}\left(\phi',\theta'\right) D_{m,n}\left(\psi_1,\psi_2\right) = D_{m,n}\left(\psi_1,\psi_2\right) T_{m,n}\left(\phi'+\psi_1-\psi_2,\theta'\right) \tag{3.10}$$

which allows, starting from the end of eq. 3.9, to translate every $D_{m,n}$ to the beginning of the equation by appropriately changing the values of the external phases, leading in the end to

$$U = D\left(\prod_{m,n} T'_{m,n}\left(\phi'',\theta'\right)\right) \tag{3.11}$$

which results in a new set of phases that will implement the desired matrix on an imperfect device. We point out that it might not be possible to perform this procedure perfectly if one or more phases exit the range of eq. 2.41.

3.2 Thermal crosstalk

As we discussed in the previous chapter, fabrication imperfections will change the expected behavior of the chip's components, a fact that has to be considered when working with a real device. But in addition to fabrication imperfection there are other unexpected behaviors of the device which emerge only when it is considered as a whole.
In particular when a TOPS is being used, it might interact with nearby TOPS. This can happen mainly in two ways: the electric signal of the TOPS might 'leak' in the connection of another one, and/or the TOPS might heat surrounding waveguides.
This phenomenon, known as crosstalk, changes how the device behaves. It is necessary to find a way to mathematically model it an account for its effect. This section will cover how to describe crosstalk and how to program the chip when it's present.
While we expect thermal crosstalk to be dominant in the chip, the results of this section apply to any source of crosstalk between the TOPS.

This discussion is based on [8].

3.2.1 Reduced description

The state of the an ideal Clement's interferometer of size N can be described by the $N(N-1)$ applied phases which, given the nature of the device under consideration, can be related to the power dissipated by their respective TOPS as

$$\chi_i = k_i P_i \tag{3.12}$$

where χ_i is used to refer to any phase in the circuit, without distinction between internal or external, k_i is the self-heating coefficient and P_i is the power dissipated by its associated TOPS.
To model crosstalk (under the assumption that it behaves linearly with dissipated power) we can introduce crosstalk coefficients k_{ji} defined by the relation

$$\chi_j = k_{ji} P_i \tag{3.13}$$

If we collect the phases and the powers in a vector (the order in which this is done is irrelevant) we can rewrite this relation as

$$\boldsymbol{\chi} = K\mathbf{P} \tag{3.14}$$

where K is a $N(N-1) \times N(N-1)$ matrix, which we'll call crosstalk matrix. But, as discussed in the previous chapter, TOPS will also heat waveguides without a heater on them, introducing phases that are not accounted for in the vector $\boldsymbol{\chi}$.
We may ask ourselves if this simple model is enough to also describe the thermal component of crosstalk and surprisingly, under the right assumptions, the answer is yes.

3.2.2 Extended description

Up until now we described the device with two phases for MZI but, in reality, each segment of waveguide contained between two DC can introduce a phase shift that should be considered, effectively doubling the number of phases that have to be taken into account and potentially adding an offset to the phase power relation, which can be rewritten as:

$$\boldsymbol{\chi}_{ext} = K_{ext}\mathbf{P} + \boldsymbol{\chi}_{ext0} \tag{3.15}$$

where $\boldsymbol{\chi}_{ext}$ and $\boldsymbol{\chi}_{ext0}$ are vectors double the length of $\mathbf{P}$, which also implies that K_{ext}, which we'll call the extended crosstalk matrix, is a rectangular matrix of size $2N(N-1) \times N(N-1)$.
Now, to model thermal crosstalk we can, in a similar fashion as done before, introduce non diagonal terms to K_{ext}, which will now account for crosstalk on waveguides with and without a heater.
Compared to the restricted description the number of phases to consider has doubled, and the crosstalk matrix in not square anymore, making it non invertible.

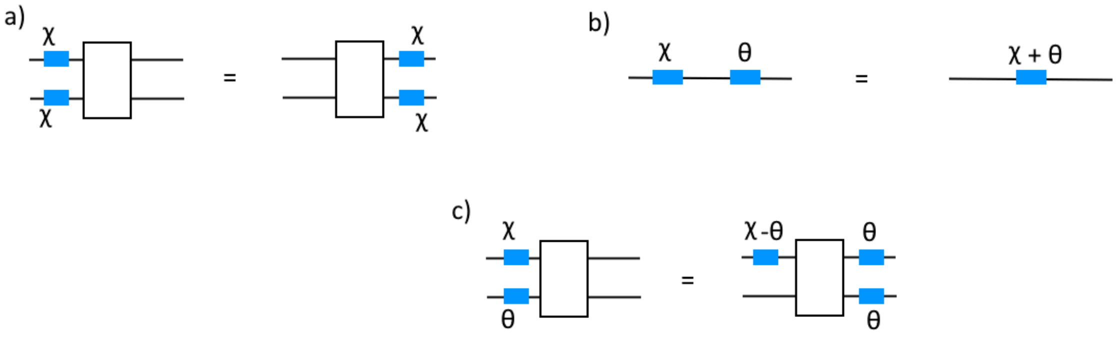


Figure 3.3: a) Schematic representation of the shift identity b) Schematic representation of the merge identity c) Combination of shift and merge to translate forward undesired phases

Under this new extended description it is not clear anymore how to program the device. In general, even if all the parameters of the problem were known, having doubled the number of phases to consider puts us outside the mathematical framework of the Clement's decomposition.
Ideally we would like to find a way to reduce again the phases needed to describe the device to the original ones, to return to a relation of the form of Eq. 3.14. As it is shown in [7] this is possible if our application is phase insensitive at the output and input of the circuit, meaning that the results only depend on the intensity at the input and output.
The idea is to find a set of phases in the reduced description that equivalently describe the extended one by mathematically translating the undesired phases onto the original ones. This is achieved thanks to two identities:

- A shift identity: an identical phase shift between two arms can be translated across a DC and the mathematical description on the device doesn't change.
- A merge identity: Two phase shifts on the same segment can be considered as one equal to their sum.

Let's consider a MZI with an undesired phase on the bottom input arm, it can be translated forward by subtracting an identical one from the top input arm, now we have two identical phases on the outputs of the MZI, a schematic drawing of this procedure can be seen in fig. 3.3 .

If the MZI is at the end of the mesh, since we are working under the hypothesis of an experiment insensitive to input/output phases, the translated phases become irrelevant. Inside the mesh a phase on the bottom output arm of an MZI is on the same arm as the external phase shifter of a following MZI and can be summed to its phase, while the phase on the top output of an MZI will be again on the bottom input arm of another MZI. In this way the phases can be pushed trough the circuit until the edge is reached. An analog reasoning can be done if the unwanted phase is on the bottom path between the two DCs of the MZI.
This shows that a phase offset on a segment of waveguide without phase shifter is mathematically identical to a phase offset on multiple segments with phase shifters, in particular it will sum to the phase of some MZI in the circuit and subtract from others.
This allows to describe the real device characterized by the phase vector $\boldsymbol{\chi}_{ext}$ as an ideal device with a vector of phases $\boldsymbol{\chi}$ by applying this procedure to all the undesired ones. Given eq. 3.15, this can be identically done to the matrix K_{ext} by summing and subtracting the corresponding rows and to the vector $\boldsymbol{\chi}_{ext0}$ leading to a phase power relation of the form:

$$\boldsymbol{\chi} = K\mathbf{P} + \boldsymbol{\chi}_0 \tag{3.16}$$

To program the device, if K is invertible, it is sufficient to apply the powers

$$\mathbf{P} = K^{-1}(\boldsymbol{\chi} - \boldsymbol{\chi}_0) \tag{3.17}$$

If K is not invertible, in general it will not be posible to program an arbitrary matrix on the device. It has been observed that stronger crosstalk leads to a higher possibility of this happening. The size of the device also plays a role, with bigger devices requiring smaller crosstalk to be fully programmable [8].
It has to be noted that if the position of the external phase shifter is at the output of the MZI the undesired phases have to be translated towards the front, instead of towards the end.

3.2.3 Non-linearity of the phase power relation

For simplicity the discussion above was done considering a linear power-phase relation. But this assumption is not necessary since the two used identities don't depend on it. Phase insensitivity at the input/output is a necessary and sufficient condition to describe the device in the reduced description. If the TOPS don't behave linearly in power the phases will just follow a more complex relation of the form

$$\boldsymbol{\chi} = f\left(\mathbf{P}\right) + \boldsymbol{\chi}_0 \tag{3.18}$$

We point out that equation 3.16 is recovered by performing a first order Taylor expansion of 3.18. If we consider the power dissipated by a TOPS to follow

$$P \propto \left(\sqrt{1 + \left(\frac{V}{V_0}\right)^2} - 1\right) \tag{3.19}$$

as previously shown, than we can Taylor expand 3.18 using the voltage as independent variable to find

$$\boldsymbol{\chi} = \sum_i K_i \mathbf{V}^{2i} + \boldsymbol{\chi}_0 \tag{3.20}$$

an expansion that presents only even terms.

Chapter 4

Calibration algorithms and characterization

Given the high number of components in the device, and the intrinsic difficulty in measuring their properties due to the complex geometry of the system, a multitude of approaches have been developed to use a Clement's interferometer as a black box, preferring to ignore the complex internal structure of the device.
Between them we find quite different solutions: from machine learning based programming algorithms [6, 29] to self-configuration[10], and on chip training for photonic neural networks [29]. In their variety, they all share a key characteristic: they need to perform long measurements on the device each time it has to programmed.
Though, applications like photonic classical and quantum computing need to efficiently program the device to any desired transformation. The necessity to perform time consuming measurements on the device each time its implemented matrix has to be changed would be detrimental, especially if the chip is embedded in a larger system, where it may not possible to access it directly.
Instead, a full characterization of the device would allow to calculate a priori the necessary voltages to be applied for each transformation. Therefore, for these applications, developing a long and complex calibration algorithm, which can be run only one time to measure all the properties of the device, is justified.
We decided to adopt this approach in our work, mainly for two reasons: as said, it is advantageous in quantum computing, but it also gives more insights on the inner working mechanism of the chip and on the quality of fabrication, since it retrieves all the specifics of the device's components. While all these information may be unnecessary to operate the chip, they are useful to understand its limits.
The calibration algorithm that can be considered the state of the art is the one used by Quandela for their 12×12 quantum processor [7, 15], which mixes interferometric measurements and machine learning to create a simulation of the device.

This chapter will cover how to characterize a Clement's interferometer: starting from the calibration of a single MZI, it will show how to circumvent the problems arising from their complex arrangement in the device, and discuss the effects of fabrication imperfection on the calibration and how to mitigate them.

4.1 Calibration of an MZI

The response curves of the phase shifters of an MZI can be empirically determined by measuring the optical power in one of its output modes [16].

The core of the proposed method, which we'll refer to as "cosinusidal proportional calibration " in accordance with [3], is to find a configuration in which the output optical power in function of the voltage has the following functional form

$$P(V) = A + B\cos(\alpha(V)) \tag{4.1}$$

where $\alpha(V)$ is the phase voltage relation.
Since A and B can be extracted from the data as $A = P_{max} + P_{min}$ and $B = P_{max} - P_{min}$ ($B = P_{min} - P_{max}$ if B is negative), the value $\alpha(V)$ can be found for each applied V by inverting eq. 4.1 as

$$\begin{aligned} \alpha(V) &= \arccos\left(\frac{P(V)-A}{B}\right) \ \alpha \in [0,\pi] \\ \alpha(V) &= 2\pi - \arccos\left(\frac{P(V)-A}{B}\right) \ \alpha \in [\pi, 2\pi] \end{aligned} \tag{4.2}$$

where the two cases can be distinguished, after an assumption on the sign of B is done, by the sign of the derivative of P (We point out the implicit assumptions on the behavior of $\alpha(V)$, which should be continuous, monotonic and differentiable).

As mentioned in Ch. 2, shining light in only one input channel of an MZI finds this functional form for the internal phase, while injecting light in both inputs and setting $\theta = \frac{\pi}{2}$ finds it for the external one.
It should be noted that this method doesn't require knowledge of the input optical power, and the output optical power can be scaled by any constant factor; an important characteristic for the calibration of MZIs inside the mesh, where there is no direct access to their inputs and outputs.

4.1.1 Calibration of an unbalanced MZI

As shown by Eq. 2.36, unbalacements act as a shift on the external phase in the configuration used to calibrate it, which, if not accounted for, will introduce an error on the offset found by the previously explained calibration algorithm. It would be convenient to find a calibration procedure for the external phase shifter that doesn't require previous knowledge of the splitting coefficients. We present two methods that can find the correct offset, at the cost of an increased number of data that has to be acquired.
The first one can be divided in three steps:

1 Perform a calibration identical to the one for the ideal case($a = b, \theta = \pi/2$), which will return a response function with a constant offset from the true one. This calibration can be used to find the necessary voltage to increase the external phase shift by a certain value.

2 Perform a measurement of the output optical power in function of θ with $a = b$ and $\phi = \phi_0$ (V = 0 V) and find the coefficients A, B, C of equation 2.34.

3 Repeat the measurement of step 2 but, by using the calibration found in step 1, setting $\phi = \phi_0 + \frac{\pi}{2}$ to get three new coefficients A', B', C'.

The correct offset and the splitting coefficients can now be found as

$$\begin{aligned} \phi_0 &= \arctan\left(\frac{-C'}{C}\right) \\ \delta k_1 &= \frac{B}{2C'} = -\frac{B'}{2C} \\ \delta k_2 &= \frac{\alpha}{2}\sqrt{\frac{1}{1-4\delta k_1^2+\alpha^2}} \end{aligned} \tag{4.3}$$

where $\alpha = \frac{A-A'}{C+C'}$.
The function arctan returns values in the range $[-\frac{\pi}{2}, \frac{\pi}{2}]$, but the signs of C' and C allow to correctly place the value of ϕ_0 in the range $[0, 2\pi]$.
This approach has one drawback: the amplitude of the oscillations of the output power of step two and three depends on the starting ϕ_0. To mitigate this effect different measurements can be performed shifting ϕ_0 by a known amount (using the calibration from step 1), until a decent oscillation is found, but this will inevitably increase the number of measurement to be performed.
The coefficients A, B, C of equation 2.34 can be found from the experimental data using the 'co-sinusidal proportional calibration', by observing that the sum of a sine and cosine can always be

rewritten as a scaled and translated cosine.

The second approach requires to:

1 Perform a calibration identical to the one for the ideal case ($a = b, \theta = \frac{\pi}{2}$) to find the calibration with an offset $\phi_1 = \phi_0 + \chi$ and coefficients A, B (refer to eq. 4.1 and eq. 2.36).

2 Perform another cosinusidal proportional calibration with $a = b, \theta = \frac{3\pi}{2}$, which will change the sign of the offset caused by the unbalancement, giving the same calibration but with an offset of $\phi_2 = \phi_0 - \chi$ and coefficients A', B' (which should be equal to A, B).

The correct offset and the splitting coefficients can now be found as

$$\begin{aligned} \phi_0 &= \frac{\phi_1 + \phi_2}{2} \\ \chi &= \frac{\phi_1 - \phi_2}{2} \\ \delta k_2 &= -sign(\sin(\chi))\sqrt{\frac{1}{4} - (B\cos(\chi))^2} \\ |\delta k_1| &= \sqrt{\frac{1}{4} - \left(\frac{B\sin(\chi)}{2\delta k_2}\right)^2} \end{aligned} \tag{4.4}$$

While this method may look easier to implement, it requires additional assumptions on the entity of the unbalancement. Let us imagine we find $\phi_1 = 0.56°$ and $\phi_2 = 359°$: this experimental result might be explained both by large unbalancement, for which we would conclude $\phi_0 \simeq 180°$ or by small unbalancement and a value of ϕ_0 close to $0°(360°)$. In the physical device we expect the latter to be the case, still making it a viable option. Also, the sign of δk_1 cannot be found this way, since the system is invariant under its change.

These two methods allow to find the correct calibration for the external phase shifter of an imperfect MZI, and also to calculate the splitting coefficients of its DC, which, with the calibration of the internal phase shifter, completely characterizes the device. For its few advantages over the other, we will adopt the first one in this work.

4.2 Calibration of the mesh

The inputs and outputs of an MZI inside the mesh are generally not directly accessible, limiting the possibility to calibrate the device with the methods explained above. While photodetectors might be added in the design just for the purpose of calibrating the device, they would significantly increase its complexity both in term of photonic and electrical structures.

To circumvent this limitation, in [3] they propose an algorithm to calibrate an ideal Clement's interferometer which requires only to measure the optical power outputs of the device. The idea is to control previously calibrated MZIs in the mesh to create a direct path for light from the inputs and outputs of the chip to the target MZI, as to apply the 'cosinusidal proportianal calibration'.

The algorithm is divided in two steps: the first one finds the calibration of all internal phase shifters, while the second one finds the calibration for the external ones. The idea behind the calibration of the internal phase shifters is to use previously calibrated MZI to 'isolate' the target one. With 'isolated' MZI we refer to an MZI in the mesh where light enters only one input, and of which one output can be observed at the output of the device, at most scaled by a constant factor. In other terms an isolated MZI is one for which the 'cosinusidal proportional calibration' can be applied directly for its internal phase shifter.

At the beginning of the calibration, when there is no information about the circuit, only the MZIs on the main diagonal fall under this definition. In fact, if light is injected in the first input of the device, it will enter each MZI on the main diagonal only in its top input, and, at the last

output of the device, their output power can be observed, at most scaled by a constant factor(see Fig. 4.1a).
Therefore, it is possible to calibrate the internal phase shifters across the first diagonal by applying the 'co-sinusidal proportonal calibration'. To isolate the remaining MZIs, in [3] they introduce a 'node isolation alhorithm'. It can be divided in two phases: one for the diagonals above the main one ('top' diagonals), and one for the diagonals below the main one ('bottom' diagonals).
For the 'top' diagonals:

1. Set all MZIs in the main diagonal to the cross state.
2. Inject light in the first input.
3. Observe the second last output
4. Set the last MZI in the main diagonal to the bar state, this isolates one MZI in the following diagonal.
5. Calibrate the isolated MZI and set it to the cross state.
6. Repeat step 3 and 4 shifting the MZI set to the bar state up by one until all the MZIs in the following diagonal are calibrated(see Fig 4.1b).
7. Set all the MZIs in the just calibrated diagonal to the bar state.
8. Set all the MZIs in the main diagonal to the cross state.
9. Repeat steps 3,4,5,6 until all the top diagonals are calibrated. At each cycle shift the observed output by two towards the top. Also the first MZI which will be set to the bar state in the main diagonal is shifted towards the top by two (see Fig 4.1c).

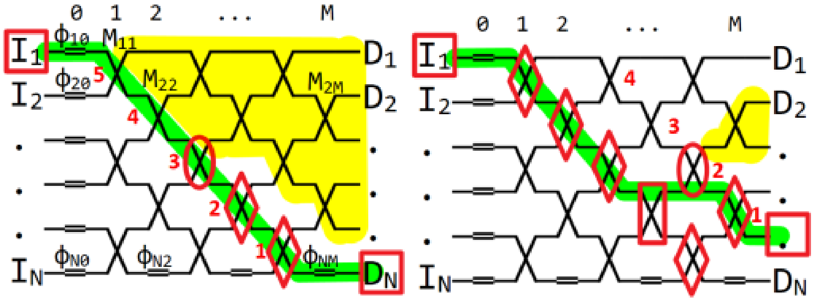
The calibration of the 'bottom' diagonals is performed in a different fashion:

1. Set all the calibrated MZIs to the bar state, observe the last output.
2. Inject light in input 3.
3. Calibrate the now isolated bottom diagonal. Once finished set its last MZI to the bar state (see Fig 4.1d).
4. Repeat step 2 and 3 until the end of the circuit is reached. At each cycle shift the input by two towards the bottom.

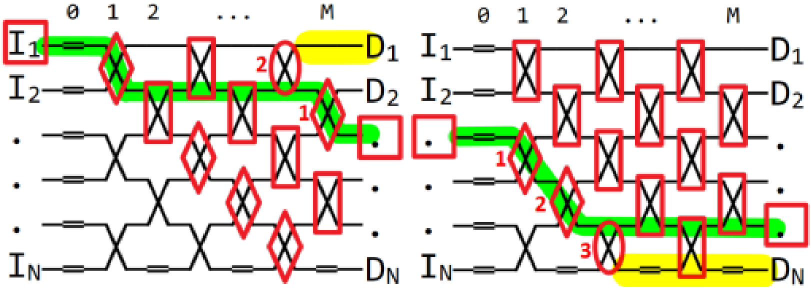
A schematic representation of the calibration algorithm for the internal phase shifters can be seen in fig. 4.1.
The calibration of the external phase shifters is more straightforward. Since the internal ones have already been calibrated we can:

1. Set each MZI to the bar state, so light will travel straight across the device
2. Calibrate the external phase shifter of each MZI in the first column, using the "cosinusidal proportional calibration"(We recall that to calibrate the external phase shifter light has to be injected in the two input channels, while the internal phase is set to $\frac{pi}{2}$; in presence of imperfections the more complex methods presented in Ch. 4.1.1 should be used).
3. Set each MZI in the first column to the identity ($\phi = \pi, \theta = \pi$).
4. Repeat step 2 and 3 for each subsequent column.

In presence of imperfections in the DCs, as it was shown before, it's not possible to perfectly implement the identity or the swap operator on an MZI. Consequently, some light will 'leak' from the desired path and enter the wrong inputs during the calibration, distorting the interferometric patterns.



(a) 3rd node, main diagonal (b) 2nd node, 1st top diagonal



(c) 2nd node, 2nd top diagonal (d) 3rd node, 1st bottom diagonal

Figure 4.1: Schematic representation of the node isolation algorithm, figure taken from [3].
In green is the direct path of light we are considering, in yellow light disregarded during the calibration. MZIs in a rectangle are in the bar state, ones in a rombus are in the cross state, while the circled MZI is the one being calibrated .

4.2.1 A different node isolation algorithm

Other ways can be found to isolate the MZIs in the mesh, which might be more robust to imperfections.
Instead of following the procedure explained above, for the 'top' diagonals one may (see Fig. 4.2)

1. Set MZI (1,1) to the bar state, inject light in the first input.
2. Observe the second last output.
3. Calibrate the now isolated 'top' diagonal and set its first MZI to the bar state.
4. Repeat step 2 and 3 for all the other 'top' diagonals. At each cycle shift the observed output towards the top by two.

This node isolation algorithm is analogous to the one used for the 'bottom' diagonals, where the role of the input and output has been reversed. A similar method is also used by Quandela [7] to calibrate their 12 × 12 quantum processor [15].

When compared to the original node isolation algorithm, it presents one major drawback: light travels through non calibrated MZIs before reaching the to be calibrated one (in the original one it happened only for the calibration of the main and bottom diagonal). An initial optimization step can be performed to increase the fraction of light that reaches the target input. This may be done, for example, by applying random phases to the internal phase shifters on the to be calibrated diagonal, or by performing a coarse sweep of their phase, until enough light reaches the observed output.

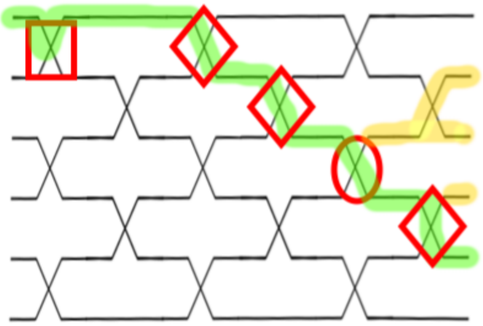


Figure 4.2: A schematic of the second node isolation algorithm.
The color notation is the same of Fig. 4.1

4.2.2 Parasitic light in the two isolation algorithms

Two factors limit the ability to implement a specific matrix on an MZI: the precision with which it's possible to control the phase shifts and deviations from the ideal splitting coefficients of the DCs.
As discussed above, the unbalancements in an MZI restrict the set of implementable matrices. The swap operator and the identity are especially sensitive to imperfections, as they require specific conditions on the unbalancements to be implementable (we recall that the swap operator requires $\delta k_1 = -\delta k_2$, while the identity $\delta k_1 = \delta k_2$. Both are perfectly implementable at the same time only in the ideal case).
Due to due fabrication processes DC close together on the device should have similar value of $\delta k \neq 0$. Therefore we expect the MZIs to reach a higher fidelity with the identity rather than with the swap operator, when programmed to their best approximations. Consequently we are incentivezed to look for a calibration method which relies more on the bar state rather than on the cross state.

We recall that the purpose of these algorithms is to shine light in only one input of each MZI in the mesh, and to observe one of its outputs, which may be scaled by a constant multiplicative factor.
Failure in implementing exactly the bar and cross state will degrade the calibration, due to light deviating from the correct path. Therefore, the fraction of input entering a wrong input in a to be calibrated MZI is a good figure of merit to evaluate the 'node isolation algorithms'.

With these considerations in mind we can move on to analyzing the robustness of the two different node isolation techniques to imperfections. Let us start with the first presented algorithm and, for the sake of simplicity, consider the first top diagonal. What is the average intensity of light that will enter the wrong ports on average?
Let us call

- p the average percentage of light that exits the bottom output when a calibrated MZI is set in the cross state, and light enters the top input.
- q the average percentage of light that exits the top output when a calibrated MZI is set in the bar state, and light enters the top input.
- N the number of MZI above the one being analyzed ($N = row - 1$).
- I the intensity of injected light.

Let us also assume that, for symmetry reasons, if an MZI is not controlled, on average half of the light will exit the top output and half the bottom one, which is equivalent to assuming a random offsets on its internal phase. We will see that in the device under consideration these phases tend to be small, biasing the MZIs towards the cross state when uncontrolled.
From the first MZI of the main diagonal $(1-p)I$ of the light will exit the wrong output, and will

subsequently cross N uncontrolled MZI before reaching the wrong input of the one being calibrated, contributing in total to $2^{-N}(1-p)$ of the initial input light I to enter the wrong port. In the second MZI of the main diagonal it will enter light of intensity pI, which will leak $p(1-p)2^{1-N}$ of the initial light in the wrong input. We can repeat the same reasoning for each MZI on the main diagonal coming before the one set to the identity. Summing all their contributions we find that, in total, parasitic light will have average an intensity of

$$I_{wrong} = I\frac{(1-p)}{2^N}\sum_{n=0}^{N}(2p)^n = \frac{(1-p)}{2^N}\frac{1-(2p)^{N+1}}{1-2p}I \tag{4.5}$$

The right input will be reached by light of intensity

$$I_{right} = p^{N+1}qI \tag{4.6}$$

the fraction light in the wrong input will be

$$R = \frac{I_{wrong}}{I_{wrong}+I_{right}} = \frac{1}{1+\frac{I_{right}}{I_{wrong}}} = \frac{1}{1+\frac{p^{N+1}q}{\frac{(1-p)}{2^N}\frac{1-(2p)^{N+1}}{1-2p}}} \tag{4.7}$$

We can study the denominator in function of N

$$\begin{aligned}\lim_{N\to\infty}\frac{p^{N+1}q}{\frac{(1-p)}{2^N}\frac{1-(2p)^{N+1}}{1-2p}} &= \frac{1-2p}{2-2p}q\lim_{N\to\infty}\frac{(2p)^{N+1}}{1-(2p)^{N+1}}\\ &= -\frac{1-2p}{2-2p}q \quad if \quad p>1/2\end{aligned} \tag{4.8}$$

$$\frac{d(\frac{1-2p}{2-2p}q\frac{(2p)^{N+1}}{1-(2p)^{N+1}})}{dN} = \frac{N(2p)^N}{(1-(2p)^{N+1})^2} < 0 \quad if \quad p>1/2 \tag{4.9}$$

The percentage of light in the wrong mode increases with N until a limit value.
Let us consider some optimistic values, $p = 0.99$ and $q = 0.99$. In this scenario R varies from $R(1) = 0.01$ to $R(\infty) = 0.02$. In every MZI in the first diagonal more than 1% of the incoming intensity would enter the wrong input.

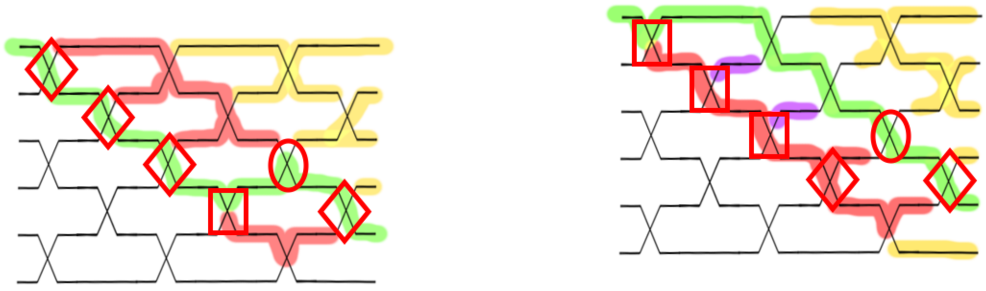
In the second node isolation technique, since the parasitic light is going through already calibrated MZI, we can use them to redirect it to the right path. To achieve this, we set all the already calibrated MZIs in the main diagonal above the one being calibrated to the bar state, and the one leading the wrong input to the cross state. If we now calculate the intensity of light getting in the wrong input we find

$$I_{wrong} = (1-q)^N(1-p)I \tag{4.10}$$

Which, for similar parameters, will be smaller than the value found for the previous algorithm. On the other hand, since light is passing through non calibrated MZIs before reaching the to be calibrated one, it is not possible to know a priori which fraction of the light will enter it from the correct input. As mention in Ch 4.2.1, we can increase the observed output power via an optimization step, e.g. by applying random voltages to these non calibrated MZIs, until sufficient light is observe, or via a coarse sweep of their phase. This should allow to direct at least half of the light in the correct direction at each MZI(as pointed out before the device studied in this work has MZI biased towards the cross state, making this fraction usually greater than half even without any optimization). Under this assumption R becomes

$$R \simeq \frac{1}{1+\frac{0.5^N}{(1-q)^N(1-p)}} \tag{4.11}$$

which decreases with N and quickly approaches 0 if $q > 1/2$.



(a) Parasitic light in the first node isolation algorithm

(b) Parasitic light in the second node isolation algorithm

Figure 4.3: In red the parasitic light which disturbs the calibration in the node isolation algorithms, in purple parasitic light redirected to the correct path

The second method shows a stronger dependence on q and allows an exponential decrease in the parasitic light with the length of the diagonal, making it more appealing to calibrate imperfect meshes. A schematic drawing of the parasitic light in the two algorithm can be seen in Fig. 4.3 .

For the sake of simplicity, we didn't consider the error introduced by parasitic light below the MZI being calibrated, which will create a distortion of the output signal.
We also observe that for the second method to be exact it's sufficient that $q = 1$, while the first algorithm requires $p = 1, q = 1(q = 1$ is less important in the calibration of θ, since it only removes parasitic light in the MZI following the one being calibrated in the diagonal, but it's still relevant for calibrating ϕ).

4.2.3 Impact of thermal crosstalk

Crosstalk also impacts the accuracy of the calibration, and how this happens depends strongly on its nature. Let us make some assumptions on its behavior in the device: we consider thermal diffusion to be the only source of crosstalk, also we assume that it happens only 'vertically', as to say a TOPS only heats waveguide segments on the same column as itself. Both these caveats are reasonable in the context of the device studied: if we have again a look at the design in fig. 2.1b we see that vertical spacing between waveguides is significantly smaller then the length of their segments between the DCs. Also, we did not observe other sources of crosstalk on the chip other than heat diffusion.
Furthermore, we observe that an unwanted internal phase on the bottom path between two DC of an MZI, when translated at the end of the circuit following the procedure described in the previous chapter, subtracts to its internal phase and then only influences external phase shifters, while external unwanted phases only sum and subtract to other external phases.
Therefore, during the calibration, thermal crosstalk will:

- Reduce the found self-heating coefficients(see Ch. 3.2): specifically the calibration algorithm will find the self heating-coefficients of the reduced crosstalk matrix
- Slightly impact parasitic light reduction in the calibration of the internal phase shifters: MZIs on the same column of the one being calibrated should not be used to not affect the calibration.
- Affect the calibration of the external phase shifters: setting the MZIs to the bar state to perform the calibration of the external phase shifters will introduce additional offsets on the external phases, also setting multiple MZI on the same column to the bar state without taking into account crosstalk will make them deviate from the true bar state.

A basic strategy to mitigate the effect of crosstalk is to not utilize any unnecessary MZI, in particular during the second step of the calibration. An example might be to not use the calibrated external TOPS to set the identity (in step 3 of their calibration only set $\theta = \pi$), and instead

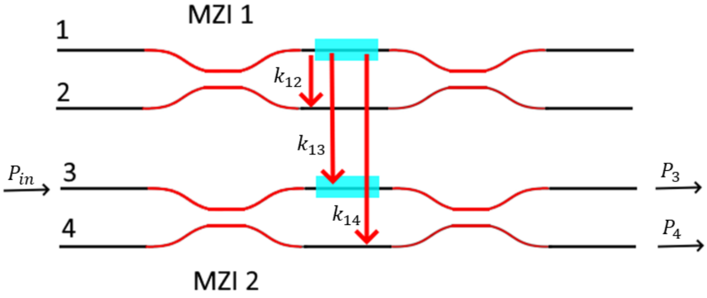


Figure 4.4: Schematic representation of the characterization procedure for thermal crosstalk between two vertically stacked MZIs. Red arrows represent heat diffusion, with their respective crosstalk coefficients, black arrows indicate input and output of light.

mathematically account for the found phase shifts.
Another measure that can be taken to reduce the effect of crosstalk on the calibration of the external phase shifter is to split light directly inside the device, instead of using two inputs (remember that the calibration of the external phases requires to use two inputs of the device, as seen in Ch 4.1). This is done by setting the preceding MZI in the same row as the one being calibrated to act as a beamspliter($\theta = \frac{\pi}{4}$), which allows to program less MZIs in the circuit to the bar state. Additionally it removes eventual errors introduced by different coupling coefficients at the input of the chip.
Characterizing the extended crosstalk matrix (see Ch. 3.2.2) on the go can also be used to effectively correct the calibration, but strategies to do so can be lengthy and complex.

4.3 Crosstalk matrix characterization

Exactly characterizing the extended crosstalk matrix with interferometric measurements poses a new, non trivial challenge. We first cover a general method to measure thermal crosstalk between two adjacent MZIs, then we discuss how it can be applied for MZIs of the mesh.

Suppose two MZIs vertically stacked on top of each other as shown in fig. 4.4, of which the calibration is known. To study the effect of crosstalk we inject light in the top arm of MZI 2 and observe how its power outputs change while swiping the voltage of the internal TOPS of MZI 1. The optical output powers of MZI 2 will follow

$$\begin{aligned} P_3 &= A + B\cos(\theta_{02} + f(P_{j1})) \\ P_4 &= A' - B\cos(\theta_{02} + f(P_{j1})) \end{aligned} \tag{4.12}$$

where P_3, P_4 are respectively the top and bottom optical power outputs of MZI 2, P_{J1} is the power dissipated by Joule effect in MZI 1, θ_{02} is the internal phase offset of MZI 2 and f is a function describing the thermal crostalk .
The 'cosinusidal proportional calibration' can now be used to retrieve $f(P_{J1})$ (see Ch. 4.1), since A ,A' and B are known from the calibration. To increase the precision of this method the internal phase of MZI 2 can be set to $\frac{\pi}{2}$ to maximize the change in P_3 and P_4 in function of P_{j1}.

Under the assumption of linearity of the thermal crosstalk, $f(P_{J1})$ will take the form

$$f(P_{J1}) = (k_{13} - k_{14})P_{J1} \tag{4.13}$$

allowing to find $k_{13} - k_{14}$, where k_{1j} are the thermal crosstalk coefficients between the TOPS on waveguide 1 and waveguide j. It has to be noted that for phase insensitive experiments only their difference has a physical importance. Generally this method doesn't allow to find k_{12}, since it only has an impact on the output phase.

In an ideal device this method can be applied straightforwardly be applied to the mesh after the calibration of the internal phase shifters is found since light may be now redirected in any desired way, but it will not be able to characterize all the coefficients: the first neighbor thermal crosstalk coefficients of the TOPS on the first two rows cannot be recovered this way (these are the coefficients corresponding to k_{12} in Fig. 4.4). Again the main source of error in an imperfect mesh will be parasitic light.

Similar strategies may be found for the external phase shifters.

4.3.1 Crosstalk matrix characterization via stochastic gradient descent

The calibration methods we have covered so far, based on the analytical analysis of the interferometric patterns of the device, show clear limits in their accuracy in presence of imperfection. While more and more complex measurements may be devised to take into account every non-ideality, a switch in the paradigm behind how we perform the calibration may lead to better results with simpler measurements.

The method which can be considered the state of the art to characterize a Clement's interferometer, presented in [8], uses machine learning to determine the crosstalk matrix, the splitting coefficients and the coupling coefficients of the device. The idea is to create a simulation of the chip, and to optimize its parameters to make it behave as similarly as possible to the real one.

A first rough characterization of the phase offsets and self-heating coefficients is obtained via the interferometric calibration and used as the starting point for the simulated chip, then the machine learning step is performed to complete the characterization of the device.

As for every machine learning algorithm, we have to define:

- A cost function, which evaluates the performance of the simulated device.
- A data set on which to train it.
- An optimization algorithm.

Since the goal is to simulate the behavior of the physical device, each data point of the data set is obtained from the chip by shining light in a random input, applying random voltages and measuring the power outputs. It will therefore be the collection $[I, V, P]$ where

- I is the input where light is shone
- V is the set of applied voltages
- P is the output power distribution

The cost function chosen is

$$f(p) = \frac{\sum_i (P_{iexp}(V_i, I_i) - P_{isim}(V_i, I_i))^2}{N} \tag{4.14}$$

where P_{iexp} are the (normalized) experimental power distributions and P_{isim} are the simulated ones, the sum is over the data set and N is its size. This cost function evaluates how similar, on average, the simulated output power distributions are to the real ones, given the input and the applied voltages.

Each parameter of the model, apart from the phase offsets, is optimized via the Adam [12] variation of stochastic gradient descent (SGD) to minimize the cost function.
After the machine learning step converges, we are left with a improved description of the device. This new knowledge can then be used to repeat the interferometric measurements, to find a better estimate of the values of the phase offsets. It is then possible to continuously cycle between interferometric measurement and SGD optimization to improve the results.

The machine learning step needs to optimize $8(N(N-1))^2$ coefficients of the crosstalk matrix, $(N(N-1))$ splitting coefficients and N coupling coefficients (the ones at the input are not relevant when just one input is used). An appropriate data set should have a least as many data points as the number of parameters of the problem.

4.4 Interferometric measurement of the fiber-chip insertion losses

To conclude the discussion on the characterization of the device we present a simple method to find the insertion losses (IL), describing the chip-fiber interface.

Let us consider an MZI where light enters the top arm, and where the two outputs have two insertion losses, k_1 and k_2. The two optical power outputs P_1 and P_2 in function of θ will follow

$$\begin{aligned} P_1 &= k_1(A + B\cos(\theta)) \\ P_2 &= k_2(A' - B\cos(\theta)) \end{aligned} \tag{4.15}$$

where A, A' and B are arbitrary constants (see eq. 2.33). We can write P_1 in function of P_2 as

$$P_1 = k_1(A + A') - \frac{k_1}{k_2}P_2 \tag{4.16}$$

we can now find the ratio $\frac{k_1}{k_2}$ as the slope of this line.
This is analogous to finding two coefficients a, b such that $P_1 + bP_2 = a$. It is well known that for two set of data point the best approximation (the values that minimize the square distance between $y = a - bP_2$ and P_1) for a, b are:

$$\begin{aligned} b &= -\frac{Covar(P_1, P_2)}{Var(P_2)} = \frac{k_1}{k_2} \\ a &=< P_1 > +b < P_2 >= k_1(A + A') \end{aligned} \tag{4.17}$$

It is not necessary for light to enter only one input, since for conservation of energy it is generally true that the optical power outputs of an MZI sum to a constant.

The just presented method can straightforwardly be applied for the last column of MZIs to calculate half of the ratios. To fully characterize the ILs this method can be applied to the second-last column of MZIs, therefore creating a system of three MZIs (or two if next to the edge of the device), of which a schematic drawing can be seen in fig. 4.5.
We will call P_{Ij} the input powers, P_{mj} the output powers of the first MZI, P_{Oj} the output powers of the device and k_j the coupling coefficients, where $j = 1, 2, 3, 4$ refers to the waveguide number.
We can find the ratio $\frac{k_1}{k_3}$ swiping the internal phase of the first MZI, and reconstructing

$$\begin{aligned} P_1 &= k_1(P_{i1} + P_{m2}) = P_{O1} + \frac{k_1}{k_2}P_{O2} \\ P_2 &= k_2(P_{i4} + P_{m3}) = P_{O3} + \frac{k_3}{k_4}P_{O2} \end{aligned} \tag{4.18}$$

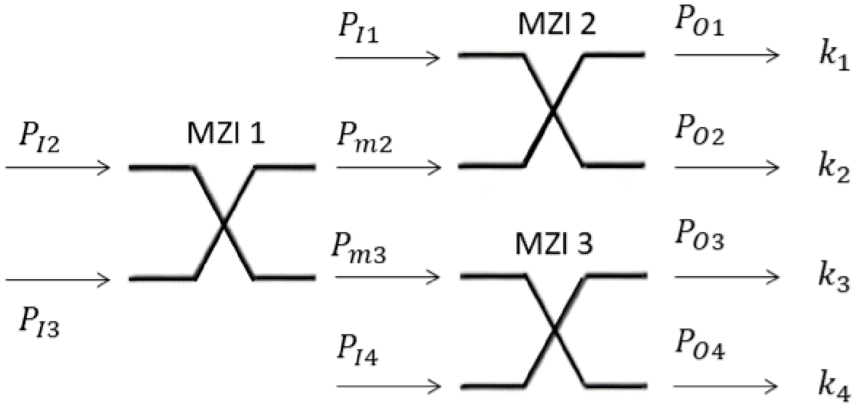


Figure 4.5: Schematic representation of the characterization procedure for the coupling coefficients of the outputs of the device. Black arrows representing the input/output light with its respective power. Initially the internal phase of MZI 2 and MZI 3 are swept, allowing to find $\frac{k_1}{k_2}$ and $\frac{k_3}{k_4}$. Then the internal phase MZI 1 is swept to find $\frac{k_1}{k_3}$.

since $P_{I1}, P_{I2}, P_{I3}, P_{I4}$ are constant, $P_1 + P_2$ will also be constant, allowing to use the method explained above to find $\frac{k_1}{k_3}$. Thermal crosstalk though will make this last assumption false in general, due to P_{I1}, P_{I4} being affected by the internal TOPS of MZI 1. To circumvent this problem, the characterization of the coupling coefficients at the output of the device can be performed after its calibration, since the algorithms presented above are not affected by them. Therefore, it is possible to use the calibration to input all light in the first MZI, as to make $P_{I1} = P_{I4} = 0$.

This method assumes path and θ independent losses inside the circuit, an assumption already made for the calibration of the device.

4.5 Characterization of the transfer matrix

One question we avoided so far is how to evaluate the quality of a calibration. An appropriate figure of merit is fundamental to validate and compare the found results, and should summarize how 'good' is the calibration in a single number. But what does it mean for a calibration to be good?
We should remember the purpose that motivated the full characterization of the circuit: the ability to program it without performing any additional optimization. Therefore, an appropriate quantity to evaluate a calibration should describe how closely, on average, the programmed matrix will be to the desired one.
This property is well summarized by the average fidelity (see Ch. 2.5.2) over all unitaries

$$\mathscr{F} = < F(U, U_{exp}) > \tag{4.19}$$

where U is a unitary matrix and U_{exp} is its implementation on the device. An estimate of this value can be obtained calculating the average over a large enough sample of random matrices.

A way to measure the implemented matrix is then necessary to estimate $\mathscr{F}$. We followed the procedure explained in [21], which we will now briefly cover.

Let T be the transfer matrix of the device, if light of intensity I is shone in input i, the j-th power output will be

$$P_{ij} = I|T_{ij}|^2 \tag{4.20}$$

since T is unitary it holds that

$$\sum_j |T_{ij}|^2 = 1 \quad \forall i \tag{4.21}$$

the absolute value of T_{ij} can therefore be reconstructed as

$$|T_{ij}| = \sqrt{\frac{P_{ij}}{\sum_j P_{ij}}} \tag{4.22}$$

This straightforward procedure allows to reconstruct $|T|$, where $|\cdot|$ is the element-wise absolute value.

The characterization of T can then be finished by finding the phases of its elements, which requires additional components and more complex measurements [21] (see appendix). To avoid the necessity of performing this step we tweak the previously introduced figure of merit, changing it to

$$\mathscr{F}_{abs} =< F_{abs}(U, U_{exp}) > \tag{4.23}$$

where $F_{abs}(U, U_{exp}) = F(|U|, |U_{exp}|)$, is the absolute fidelity.
It can be shown that

$$1 \geq F_{abs}(U, U') \geq F(U, U') \tag{4.24}$$

which also implies $|\mathscr{F}| \geq \mathscr{F}$. Using the absolute fidelity 'overestimates' the similarity between matrices, since it disregards the information about the phases.

The absolute fidelity is commonly used [5, 7] instead of the fidelity to evaluate the similarity between matrices, as a mean to avoid performing the full characterization of the transfer matrix of the physical device.

Chapter 5

Simulation of the device

There are ample of reasons to build a simulation of the device: it can be used to test the characterization and calibration algorithms, it allows to study the effect of fabrication imperfection, it helps attribute a confidence interval to the experimental results, and can also be used, as shown in section 4.3.1, for a more complex and precise characterization method.

The simulation was built in python, where the device is represented as the product of the transfer matrices of its individual components, under the extended description (see section 3.2.2) to allow a more accurate description of thermal crosstalk. The simulated device supports a phase voltage relation of the form

$$\boldsymbol{\chi} = K\mathbf{V}^{\mathbf{n}} + \boldsymbol{\chi}_0 \tag{5.1}$$

where the exponentiation between the vector V and the vector n is intended element-wise ($(\mathbf{V}^{\mathbf{n}})_i = V_i^{n_i}$). Therefore, it is defined by:

- $2N(N-1)$ phases χ_0
- $N(N-1)$ splitting coefficients k_i
- $N(N-1)$ voltages V_i
- $N(N-1)$ exponents n_i
- A $2N(N-1) \times N(N-1)$ crosstalk matrix

where N is the device size. Different coupling coefficients were not considered in the simulation, but can be easily be added.

5.1 Comparison between the two node isolation algorithms

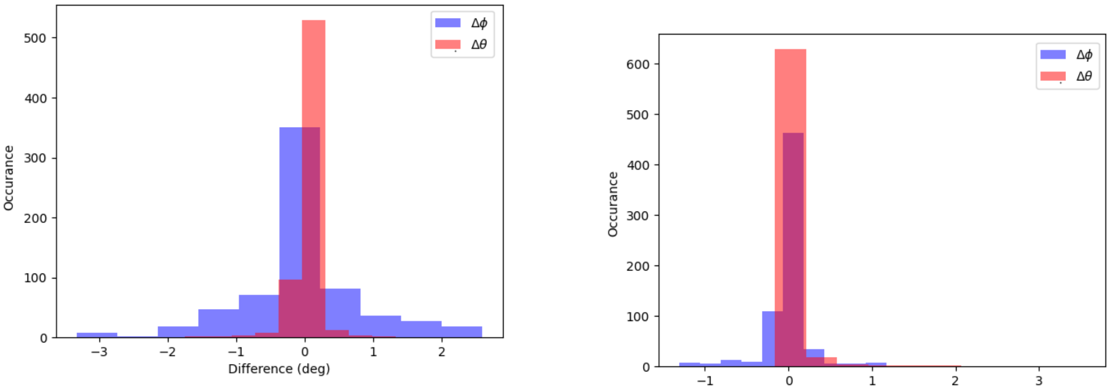
Our first objective was to compare the two 'node isolation algorithms', to see which one performed better on an imperfect device. For this purpose, a simple linear phase response function of the form

$$\boldsymbol{\chi} = \mathbf{V} + \boldsymbol{\chi}_0 \tag{5.2}$$

was chosen, so every phase shifter could be characterized by only an offset, which can be found from the 'cosinusidal proportional calibration' as $\chi_i(V = 0)$.

To test that the algorithms were working correctly, we ran them on a simulated perfect 12×12 mesh. The offsets where randomly extracted from uniform distributions: from 0 to π for the internal phases and from 0 to 2π for the external ones. Histograms of the difference between the offsets found by the algorithm and the true offsets in ten runs can be seen in fig 5.1, which shows great accuracy for both algorithms (for each run a new device was simulated, the accuracy shown by the algorithms is the expected one from the 'cosinusidal proportional calibration' given the number of points used).

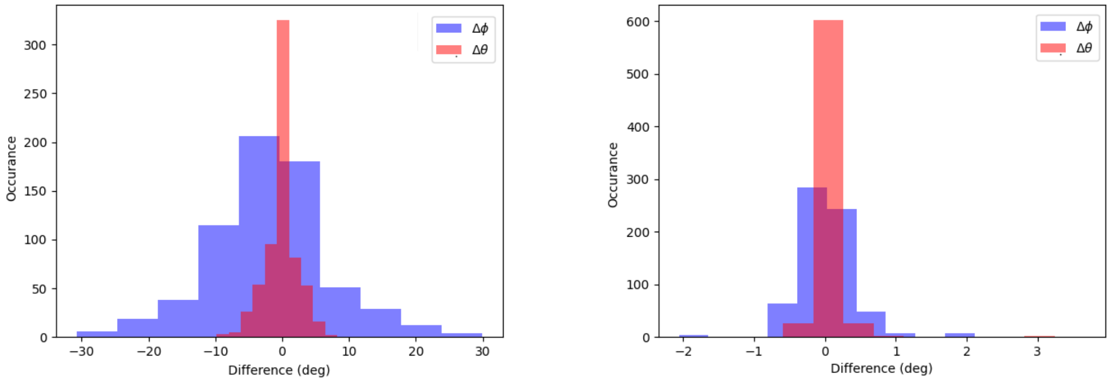
Then, the two 'node isolation algorithms' were benchmarked on a simulated imperfect mesh, where the unbalancements were extracted from a normal distribution centered around 0.02 with $\sigma = 0.005$ (these values have been chosen since they had been previously measured in [6] on similarly built MZIs), the results obtained from ten runs can be seen in fig. 5.2. As expected, the second node isolation algorithm is more robust to imperfections of the device.



(a) Results using the first introduced node isolation algorithm.

(b) Results using the second node isolation algorithm.

Figure 5.1: Errors on the offsets found by the calibration of 10 12 × 12 simulated devices with perfect beam splitters. 100 points were used for the co-sinusidal proportional calibration.



(a) Results using the first introduced node isolation algorithm, the calibration of the external phase did not account for the unbalancements.

(b) Results using the second node isolation algorithm, the calibration of the external phase accounted for the unbalancements.

Figure 5.2: Errors on the offsets found by the calibration of 10 12 × 12 simulated devices with unbalanced beam splitters. 100 points were used for the 'cosinusidal proportional calibration'. We clearly observe a higher spread in the errors for the first node isolation algorithm, for both the internal and external phases. The large errors on the external phases are due to a poor calibration of the internal ones, and for not accounting for the different splitting ratios in their calibration.

5.2 Study of the performance of the algorithm

It can be interesting to study how the performance of the device and of the calibration are affected by different values of the unbalancements, to assign a bound on the tolerable entity of non-idealities. In section 4.5 we already introduced a figure of merit to quantify the 'goodness' of a calibration, the average absolute fidelity $\mathscr{F}$, but let us introduce a few more:

parameter	distribution	mean	std var
k	uniform(0.5,1.5)	1	$\frac{\sqrt{3}}{6}$
n	normal	2	0.01
ϕ_0	uniform(0,2π)	π	$\frac{\sqrt{3}\pi}{3}$
θ_0	uniform(0,π)	$\frac{\pi}{2}$	$\frac{\sqrt{3}\pi}{6}$

Table 5.1

1. F_I: the absolute fidelity between the identity matrix I and the implementation of I.
2. $< \Delta\theta >$ and $< \Delta\phi >$: the average distance from the true values of θ_0 and ϕ_0 and the measured ones.
3. $\mathscr{F}_{max}$: $\mathscr{F}$ calculated using the perfect calibration.
4. $\Delta\mathscr{F} = \mathscr{F}_{max} - \mathscr{F}$.

were the the first one is a simplified version of $\mathscr{F}$, which mostly takes into account the quality of the calibration of the internal phase shifters, but also highlights the fidelity achievable with a fundamental transformation: the identity. The other newly introduced figures of merit are only measurable in a simulation, but can be quite helpful in understanding the performance of the device. The second one is a simple indicator of how much the calibration is wrong. It could be calculated for all the parameters of the simulation, though we notice that the found offsets are particularly sensitive to imperfections, and are the main source of error when programming the device. To understand the purpose of the last two we should differentiate between the performance of the device and the performance of the calibration. In fact, an imperfect chip cannot, even in principle (see section 3.1.4), implement any arbitrary transformation. Therefore, the possibility to accurately program a target matrix will be limited by two factors: the accuracy of the calibration and the quality of the device. These two are both accounted for by $\mathscr{F}$, while they are individually evaluated by $\mathscr{F}_{max}$ and $\Delta\mathscr{F}$. In fact, the former only depends on the parameters of the device, while the latter estimates only the quality of the calibration by comparing the obtained performance with the maximum achievable one. $\mathscr{F}$ can be rewritten as

$$\mathscr{F} = \mathscr{F}_{max} - \Delta\mathscr{F} = 1 - (1 - \mathscr{F}_{max}) - \Delta\mathscr{F} \tag{5.3}$$

clearly isolating the error introduced by the calibration ($\Delta\mathscr{F}$) and by the device ($1 - \mathscr{F}_{max}$).

5.2.1 Changes to the simulation and algorithms

To better simulate a physical device, we introduced different self-heating coefficients and exponents in the phase voltage relation of the phase shifter, modifying the phase voltage relation to:

$$\boldsymbol{\chi} = K\mathbf{V}^{\mathbf{n}} + \boldsymbol{\chi}_0 \tag{5.4}$$

where K is a diagonal matrix. The values for the parameters used in the simulation were extracted from random distributions described in table 5.1.
Three parameters are then necessary to describe a TOPS. To retrieve them from the data we perform two fits: a first one on the calibration obtained from the 'cosinusidal calibration algorithm', and a second one on the measured optical power with a curve shaped like

$$P = A + B\cos(kV^n + \theta_0) \tag{5.5}$$

using the values found for k, n and θ_0 from the previous fit as starting points to get a better result, and where A and B are arbitrary parameters of the fit (see eq. 2.33). In a similar fashion, a fit is used to find the values of the coefficients C and C' used to calculate ϕ_0 and the splitting coefficients of the DC (see Ch. 4.1.1).
Estimating the parameters through a fit should also reduce the effect of noise on the calibration.

5.2.2 Results

We extracted the values of the splitting coefficients of the simulated circuit from a normal distribution, and studied the previously mentioned figures of merit in function of the average $< \delta k >$ and of the standard deviation $\sigma(\delta k)$ of the distribution. To reduce random fluctuation, for each point three simulations were ran, and the results were averaged.

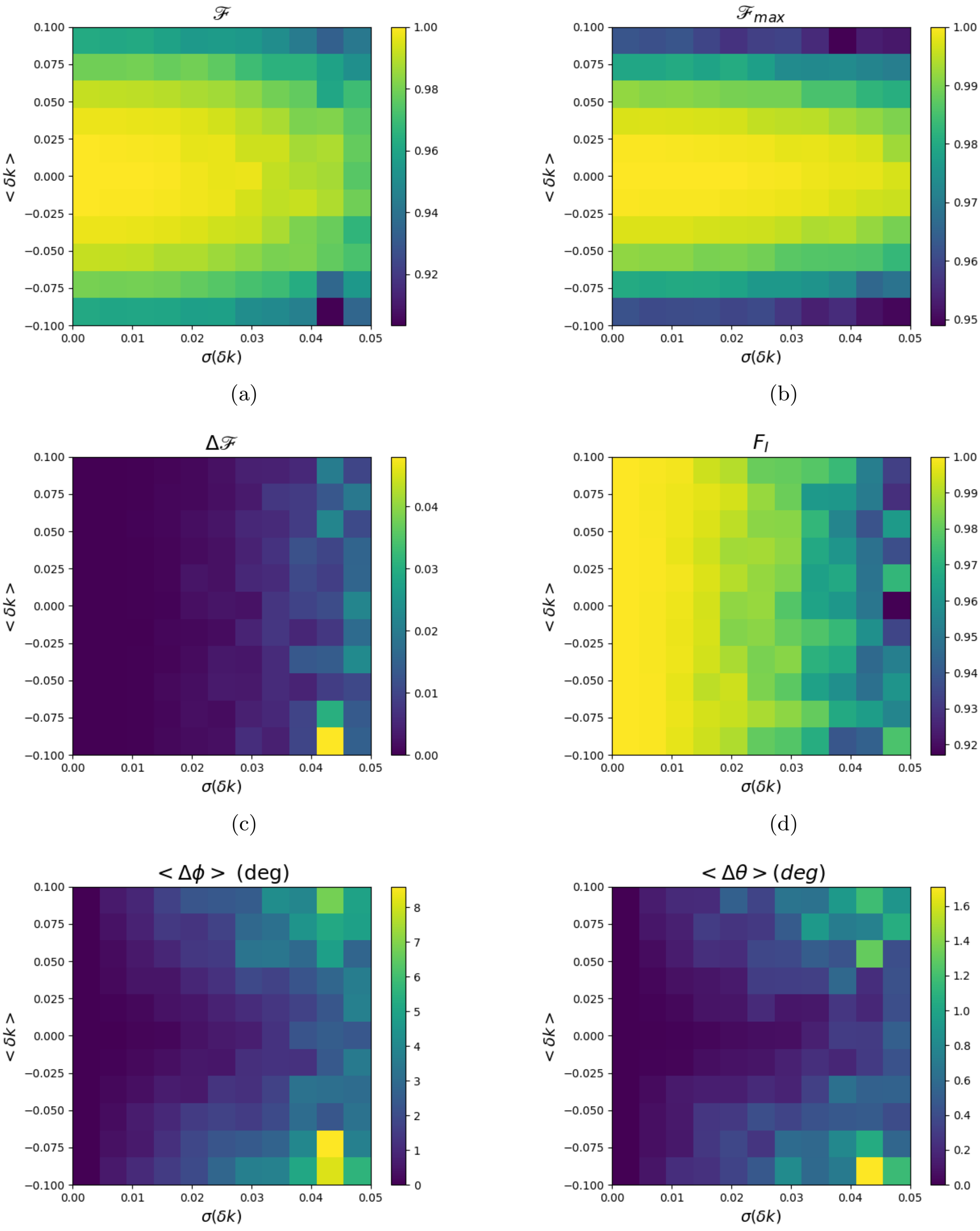
Fig 5.3 shows heat maps representations of the found figures of merit in function of $< \delta k >$ and $\sigma(\delta k)$. Comparing 5.3a and 5.3b, we see that the performance of the device quickly degrades with the average of the unbalancements, mainly due to the impossibility of perfectly implementing the desired transformation. In fact, from figs. 5.3c, 5.3e and 5.3f we can observe that the quality of the calibration degrades mainly with the σ of the distribution. The precision at which it is possible to implement the identity, shown in fig. 5.3d, also degrades mainly with σ.

So, why is this the case?

Let us have a look again at how $\mathscr{F}$ is calculated. It is the average absolute fidelity reached with a sample (we used a sample of size 100) of random matrices uniformly extracted from the unitary matrices. But, as shown in [19], uniformly extracting a random unitary biases the internal phases on the device towards small values. Also, non-ideal MZIs require $\delta k_1 \simeq -\delta k_2$ to exactly implement transformations with internal phases close to zero (see section 2.5.2). So, under this prior, the favorable device is one where $< \delta k > \simeq 0$, while $\sigma(\delta k)$ has more freedom to vary. In contrary, the accuracy of the calibration depends on the fidelity of the bar state, which is maximized by $\delta k_1 = \delta k_2$. This is also reflected on F_I, which degrades mainly with $\sigma(\delta k_1)$, due to the decreasing fidelity achievable by the MZIs set to the bar state with the identity.

In other terms, the spread of the splitting coefficients affects mostly the quality of the calibration, while their average impacts the average quality of the device more (under the assumption that a random sample of unitary matrices uniformly extracted is a good representative of our use case)

Comparing 5.3e and 5.3f also shows that the calibration of the external phases is significantly more affected by unbalancements than the one for the internal phases.



(a) (b)

(c) (d)

(e) (f)

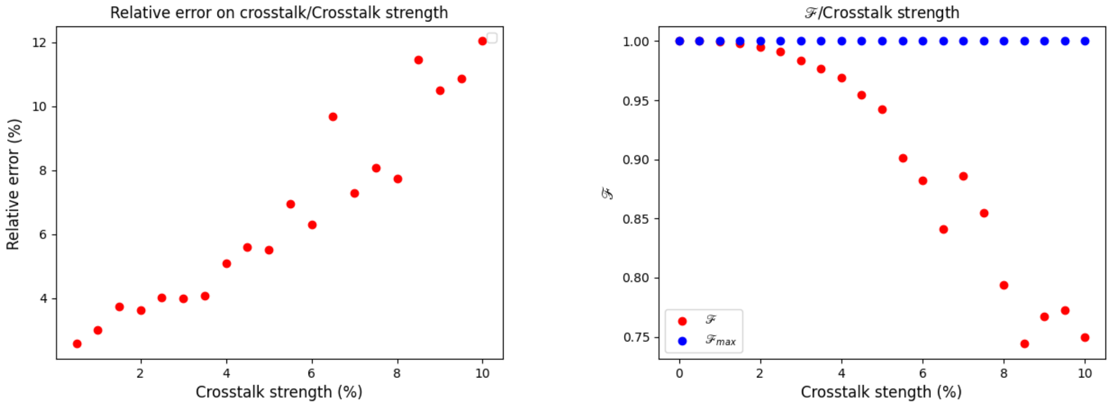
Figure 5.3: Various figures of merit obtained from the simulation of the device in function of $< \delta k >$ and $\sigma(\delta k)$: a) $\mathscr{F}$ b) $\mathscr{F}_{max}$ c) $\Delta\mathscr{F}$ d) F_I e) $< \Delta\phi >$ f) $< \Delta\theta >$. Each point is the average of three different simulations, to reduce the effect of randomness.

5.3 Simulation and study of thermal crosstalk

Thermal crosstalk was introduced under some simple assumptions:

- It only affected first neighbor waveguides
- It was considered symmetric: given a TOPS, its thermal crosstalk coefficient with the waveguide above is equal to the crosstalk coefficient with the waveguide below.

The calibration algorithm was expanded to characterize the first neighbor crosstalk coefficients of the TOPS (except the ones of MZIs in the first two rows, which, as previously mentioned in Ch. 4.3, are not trivial to find. Also, only two data points were used to find the crosstalk coefficients, since simulating crosstalk significantly increases the computation time. Increasing the number of used data points to characterize crosstalk would also yield an increased accuracy). Thermal crosstalk coefficients were extracted from the same distribution of self-heating coefficients, and then scaled by a multiplicative factor, which we will refer to as crosstalk strength. The error on the characterization of crosstalk and the obtained values for $\mathscr{F}$ in function of the crosstalk strength can be seen in fig. 5.4. We observe that a poor characterization of the crosstalk matrix leads quickly to low performances of the device.



(a) (b)

Figure 5.4: Results of the study of the effect of crosstalk. To the left: average relative error on the found crosstalk matrix in function of the crostalk strength. To the right: $\mathscr{F}_{max}$ and $\mathscr{F}$ in function of the crosstalk strenght. $\mathscr{F}_{max}$ is constantly one. In fact, if the crosstalk matrix were to be known exactly, the device could be operated perfectly. To reduce the effect of randomness, each point is the average of three simulations. These results are pessimistic, since an approximated method has been used to retrieve the crosstalk coefficients. Still, they show how a poor characterization of the crosstalk matrix leads to bad performances of the device.

5.3.1 Simulation of the SGD crosstalk matrix characterization

We ran a simulation of the stochastic gradient descent (SGD) crosstalk matrix characterization algorithm covered in section 4.3.1. The parameters chosen were

- Size of the device = 8
- Crosstalk strength = 5%
- Relative error on the measurements = 1%
- $< \delta k >= 0$
- $\sigma(\delta k) = 0.005$
- Batch size = 4

- Train set size = 224
- Test set size = 56
- Optimization steps = 500

also each data point contained the measured transfer matrix, instead of only one column of it (see section 4.3.1). The cost function was changed to:

$$C = 1 - < F_{abs}(U_m, U_s) > \tag{5.6}$$

where U_m is the measured transfer matrix and U_s is the simulated one, and the average is calculated over the train set (the "measured" values are of course simulated).
The small size of the data sets is justified by the reduced number of parameters of the problem, since we assumed first neighbor crosstalk. We see a significant improvement on the estimate of the crosstalk matrix and of the unbalancemets found by the interferometric calibration. A plot of the performance calculated over the test set in function of the optimization step can be seen in fig. 5.5. We also report an improvement of $\mathscr{F}$ from

$$\mathscr{F}_{init} = 0.94 \tag{5.7}$$

to

$$\mathscr{F}_{final} = 0.998 \tag{5.8}$$

with a $\mathscr{F}_{max}$ of the device of

$$\mathscr{F}_{max} = 0.999 \tag{5.9}$$

The SDG optimization of the parameters proved to be effective in improving the quality of the calibration.

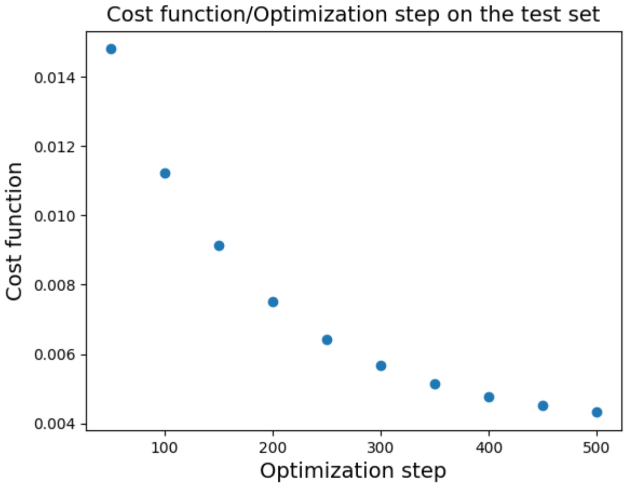


Figure 5.5: Results of the SGD optimization: plot of the cost function evaluated on the test set in function of the optimization step. We see a clear improvement in performance.

Chapter 6

Experimental characterization

Due to the high number of optical and electrical connections in the device, and given the complex nature of the calibration algorithm, only two chips were packaged for testing: one in house (by Olivier Castany) and one by Tyndall. Before testing on the chips, preliminary measurements were performed on a single MZI to test the "co-sinusidal calibration algorithm" and to observe the phase voltage relation and $I-V$ curve of the TOPS, the we recall are formed by a heater placed on top of the a waveguide.

Fit of the experimental data will often be used in this chapter. They were performed to minimize the square distance between the data and the model function. The reduced coefficient of determination R^2_{red} has been used a a figure of merit to evaluate the 'goodness' of a fit.

6.1 Preliminary characterization: a single MZI

We tested the behavior of the heaters and the calibration algorithm with data previously obtained (by Nicholas Dunoyer) from a MZI formed by a MMI and a DC, of which the design can be seen in fig. 6.1.
This device behaves as a normal MZI, where light enters only the top input. Two thermal phase shifters are used to control the internal phase difference, one on each path. We tested both, following the convention that the one on the top path will introduce a positive phase shift, while the one on the bottom a negative one.

While light was injected in the chip the voltage applied to the heaters was swept from $0V$ to $40V$ and the applied voltage, current and the optical power of the outputs were measured.

To check the validity of Eq. 2.17 we performed a fit of the $I-V$ curves of the heaters, shown in fig. 6.2. We extracted the resistance at room temperature, R_0, and we calculated the resistivity $\rho = \frac{R_0}{L}$ at room temperature of the heaters to be:

$\rho_{top} = 13.46 \pm 0.03\Omega/\mu m$

$\rho_{bottom} = 13.25 \pm 0.03\Omega/\mu m$

Then, we calibrated the TOPS with the "cosinusidal proportional algorithm" using as independent variables "x" the measured voltage, the measured current and the dissipated power obtained as the product of the two, and fitted them with

$$\theta = \theta_0 + kx^n \tag{6.1}$$

2.19 where θ_0 is the offset on the internal phase, k is the self-heating coefficient of the TOPS and n is parameter describing the shape of its phase-voltage relation. The results of the fits were subsequently used as initial values to fit the measured optical power OP with

$$OP = A + B\cos(kx^n + \theta_0) \tag{6.2}$$

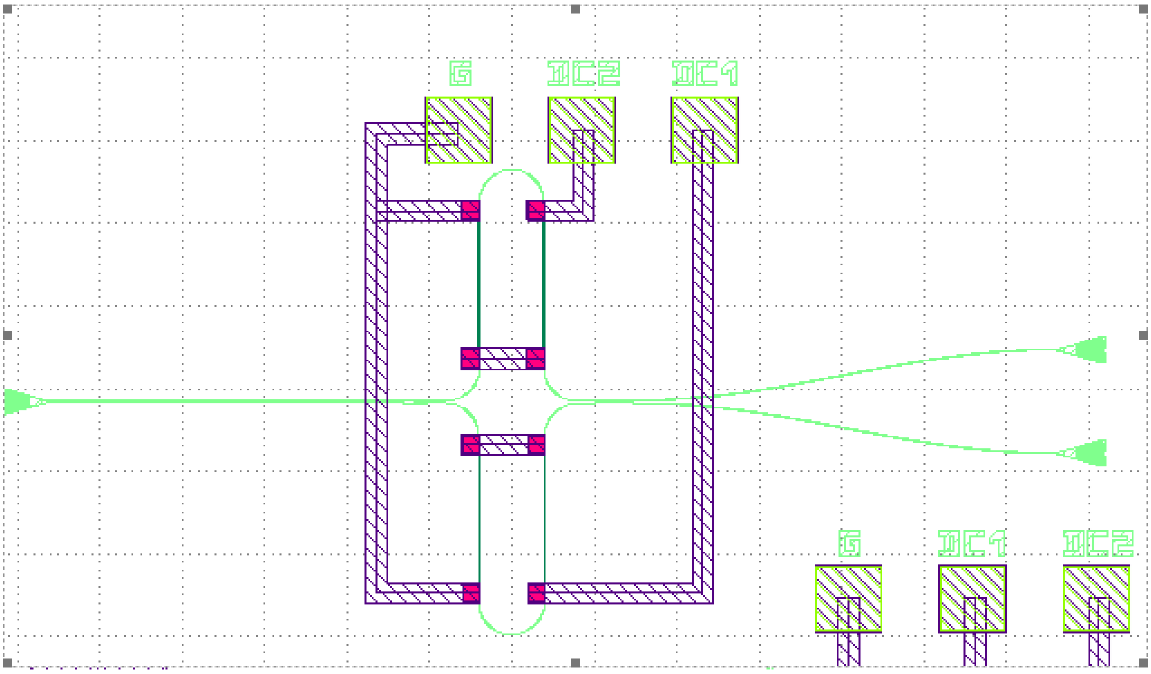


Figure 6.1: Design of the considered MZI. It can be seen that two TOPS are presents, one for each arm of the MZI.

where A, B are parameters of the fit. Fig. 6.3 shows the results of the fit of the optical power in function of the dissipated power and the voltage of the top phase shifter.
Using the voltage or the current as the independent variable "x" finds n to be $\simeq 1.6$ and 2.7 respectively, while using the power finds $n \simeq 1$ as expected.

The values found for the parameters of the fits (using the dissipated power as independent variable) are reported in table 6.1. The two TOPS have similar values for n and k, and the sum of the two offsets $\theta_{top} + \theta_{bottom}$ is $6.19 \sim 2\pi$, as expected since the starting offset is the same for both of them.
From k and n we can calculate the power needed to induce a π phase shift, a commonly used figure of merit for the TOPSs, as

$$P_\pi = \left(\frac{\pi}{k}\right)^{\frac{1}{n}} \tag{6.3}$$

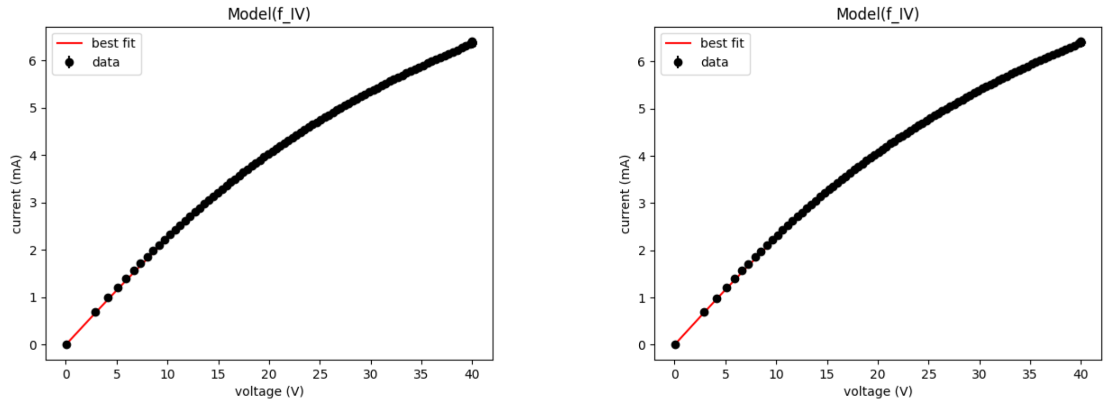
We find

$$P_{\pi top} = 143 \pm 8mW$$

$$P_{\pi bottom} = 142 \pm 14mW$$

parameter	value	err
n_{top}	1.04	0.01
k_{top}	0.018 rad/mW	0.001 rad/mW
θ_{0top}	4.28 rad	0.02 rad
n_{bottom}	1.02	0.02
k_{bottom}	0.02 rad/mW	0.002 rad/mW
$\theta_{0bottom}$	1.91 rad	0.03 rad

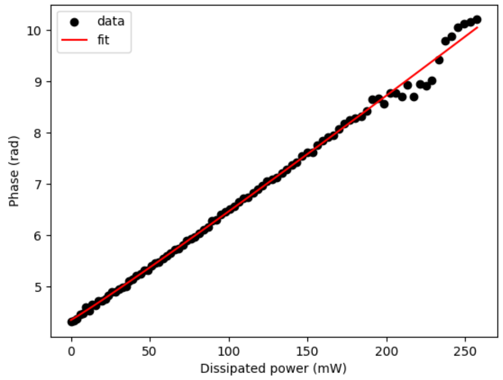
Table 6.1: Values for the calibration curve extracted from the fit of the optical power in function of the dissipated power. The fit for the top phase shifter can be seen in fig. 6.3



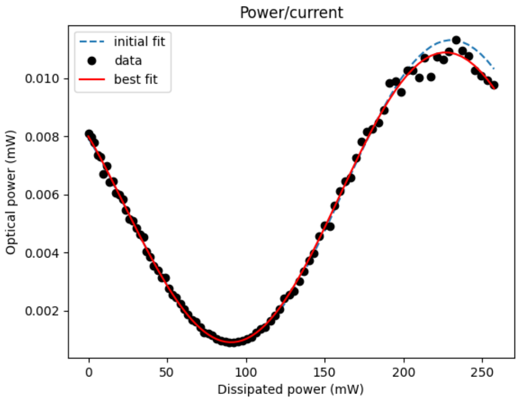
(a) Fit with Eq 2.17 of the $I - V$ curve for the top thermal phase shifter, which finds $R_0 = 4.31 \pm 0.01 k\Omega$ and $V_0 = 24.3 \pm 0.1V$

(b) Fit with Eq 2.17 of the $I - V$ curve for the bottom thermal phase shifter, which finds $R_0 = 4.24 \pm 0.01 k\Omega$ and $V_0 = 23.8 \pm 0.1V$

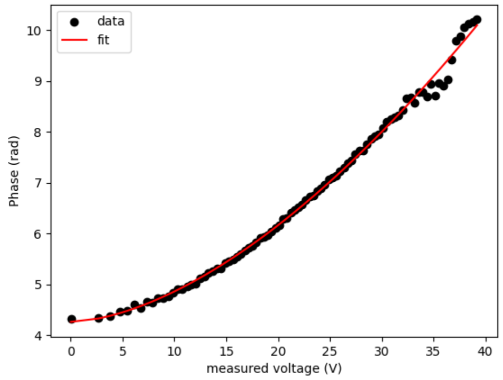
Figure 6.2



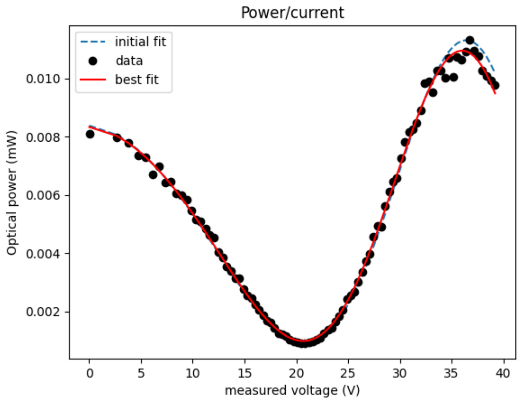
(a) Fit of the phase-voltage relation in function of the dissipated power(mW) with Eq. 2.19 for the top phase shifter.



(b) Fit of the top output optical power in function of the dissipated power(mW) for the top phase shifter. The initial fit is done using the parameters found by the fit of phase found by the 'cosinusidal proportional calibration'.



(c) Fit of the phase-voltage relation in function of the voltage(V) applied the top phase shifter with Eq. 2.19.



(d) Fit of the top output optical power in function of the voltage(V) applied the top phase shifter. The initial fit is done using the parameters found by the fit of phase found by the 'cosinusidal proportional calibration'.

Figure 6.3

6.2 Experimental setup

The experimental setup, of which a schematic representation can be seen in fig. 6.4, was designed with the calibration algorithms in mind, which needs access to all optical inputs and outputs of the device and to all its electrical connections.

A laser was connected to a polarization controller, followed by three 1x4 optical switches arranged to form a 1x8 switch, which were then connected to the 8 optical inputs of the device. The 8 optical outputs were connected to a power meter that can measure them simultaneously. Voltages were applied to the circuit via a source measurement unit (SMU) with 120 channels able to provide up to $40V$ and $300mA$ per channel and powered by a separate power supply.

The instruments used are here listed

- Power meter: Santech OPM150 12channels extended range
- Polarization controller: General Photonics PSY-201
- Laser: OSICS ECL tunable laser module
- Switches: OSICS SWT Option:PM13
- SMU: XPOW-120AX-CCvCV-U
- Power supply: RS PRO Digital, Bench Power Supply, 150W, 1 Output, 0 → 30V 0 → 5A (Ref.: 175-7368)
- Optical fibers: Thorlabs P3-1550PM-FC-1, P3-SMF28E-FC-1
- Mainframe: Yenista Optics OSICS

The laser and the switches were mounted inside the OSICS mainframe.

While most of the experiments were performed with this setup, some required to modify it. These changes will be pointed out when describing the related measurements.

6.2.1 Study of errors

In the experimental apparatus described above, three measurement instruments are present: the polarization controller, which also measures the actual polarization, the SMU, which also measures the effective applied currents and voltages, and the optical power meter. The accuracy of the polarization controller will have little importance in the experiments we will cover, so we can focus on analyzing the remaining two.

The optical power meter has a declared resolution of $0.001dB$, which we also observe experimentally.

The SMU has a declared resolution of 16 bits, which translates to an accuracy on the current of

$$\Delta I = \frac{300mA}{2^{16}} \simeq 4.5\mu A$$

an on the voltage of

$$\Delta V = \frac{40V}{2^{16}} \simeq 0.6mV$$

Though, we observe a resolution of $\simeq 1mV$, corresponding to 15 bits, and of $\simeq 30\mu A$, corresponding to 13 bits. Furthermore, we observe that in the range between $0.25V$ and $1.25V$ the instrument is unreliable, with each channel behaving differently, as shown in fig. 6.5 . Above $1.25V$ the instrument behave as expected. During this work, there was no need to account for this effect, since we usually operated at higher voltages.

To test which factor dominates the error in our setup, between noise and instrumental accuracy, we set a fixed state on the device and continuously measured it. This was done by injected light

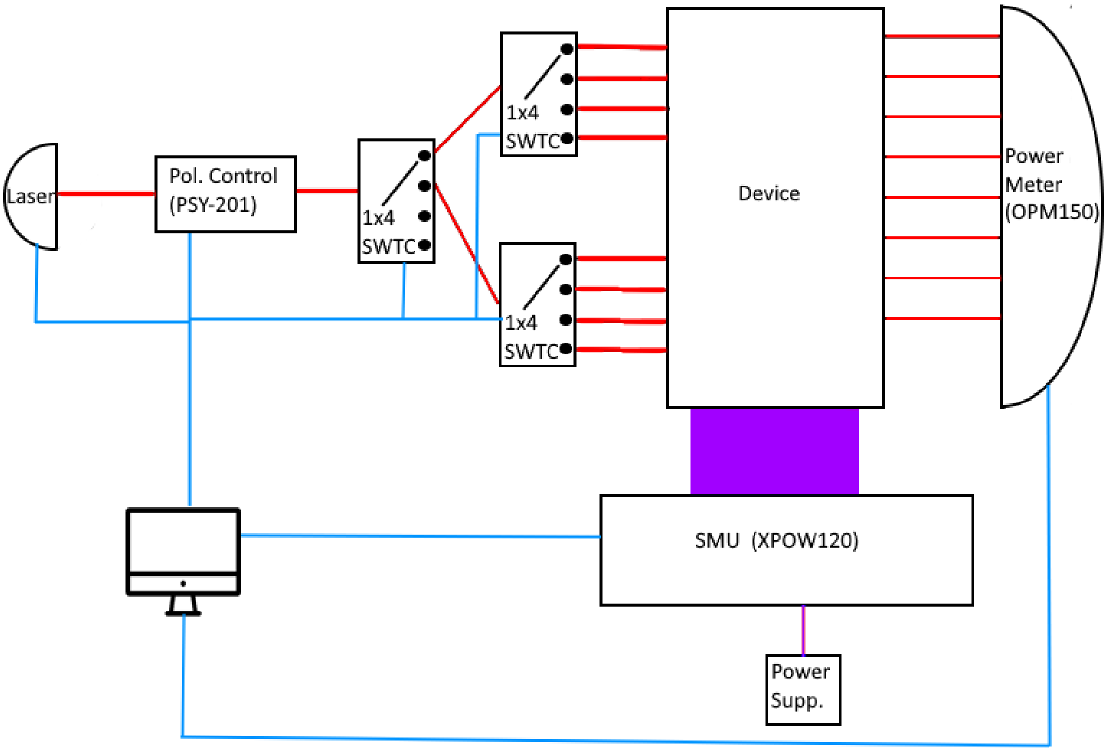
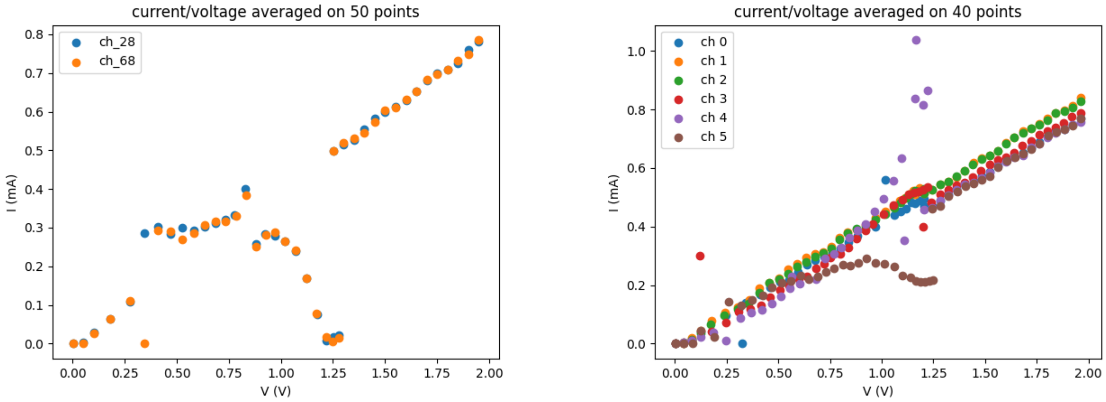


Figure 6.4: Schematic of the experimental apparatus. Red lines are optical connections, in purple electrical connections and in light blue data connections.



(a) Measured current voltage relation of two heaters in the range $0 - 2V$, each point is the average of 50 data points.

(b) Measured current voltage relation of a $2.4k\Omega$ resistor with six different channels in the range $0 - 2V$, each point is the average of 40 data points.

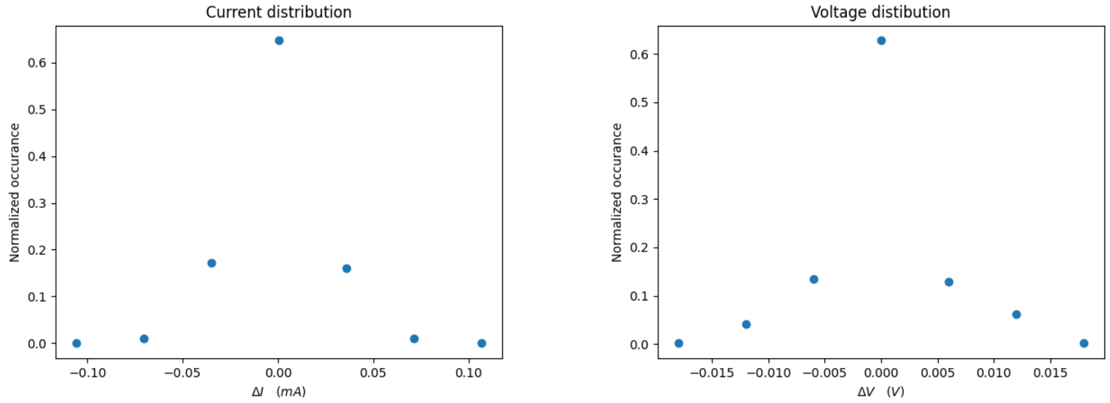
Figure 6.5: Testing the SMU: a) on-chip b) on a simple resistance. We observe that the measured currents and voltages under $1.25V$ don't seem to follow their expected linear behavior. Furthermore, there doesn't seem to be a common trend between different SMU channels. Data points are labeled with the number of the channel used on the SMU.

inside the chip, setting a voltage on a TOPS (which has an influence on the output optical powers) and continuously measuring the current, voltage and optical power outputs for a fixed amount of time.

We took 100 measurement of $\sim 140s$ (500 data points). The TOPS used was the internal one of an MZI on the last column of the circuit. Fig. 6.6a and 6.6b show the distributions of the oscillations of the current and voltage, we observe that around $60\% \sim 70\%$ of the measured values fall in the central bin. The instrumental accuracy is the dominant source of error.

The measurement of optical power was dominated by noise, which also showed a strong dependence on the measured value, with lower powers exhibiting higher relative noise. For each measurement (500 data points) we calculated the average measured power and its standard variation,. The plot of the standard variation in function of the average power is shown in fig. 6.7a, while the plot of the distribution of the power oscillations can be seen in fig. 6.7b .

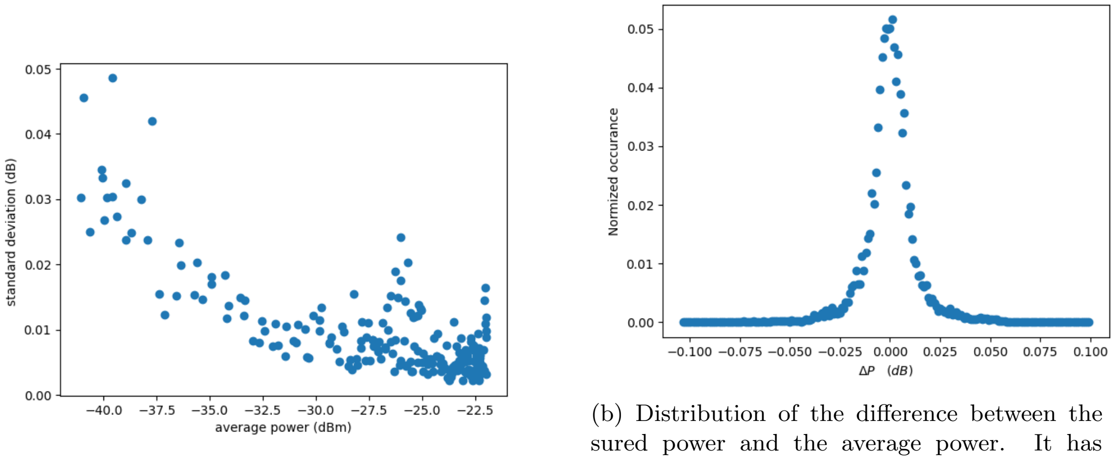
Usually time integration provides a clear way to deal with noise, at the cost of an increase of measurement time. Unfortunately the calibration algorithms already requires a large amount of data. Consequently, it would be convenient to model the behavior of the noise, as to assign a confidence interval to the measured power without having to perform lengthy measurement for each data point. A first idea might to use the distribution shown in fig. 6.7b, but it wouldn't take into account the clear dependence of noise on the power. We used a linear fit of the data in fig. 6.7a to approximately model the confidence interval in function of the power. This will be useful to attribute weights to the fit of the optical powers that will follow in this chapter.



(a) Distribution of the oscillation of the measured current

(b) Distribution of the oscillation of the measured voltage

Figure 6.6: Distribution of the oscillations (difference between the measured voltage and its average over 500 data point) in the measurement of current and voltage. We see that most values fall in the central bin, showing a dominance of the instrumental accuracy over noise.



(a) Plot of the measured standard deviation of the optical power in function of its measured average value.

(b) Distribution of the difference between the measured power and the average power. It has $\sigma = 0.012dB$, this disregards the relation between the error and the measured power.

Figure 6.7

6.3 Tyndall packaged chip

Two chips were package for testing: one by Tyndall, and one by CEA. The Tyndall packaged device, shown in fig. 6.9, has an array of 12 SMF 28 fibers on both side to optically interface with the chip, while its electrical connections are arranged as two 2×16 grids with pace $0.5cm$.

On the chip, numbering of the waveguides and of the electrical connections follows the one shown in fig. 6.8. Each side of electrical connection has its own ground, connected to the two leftmost pins. The top row of connections controls the internal TOPS, while the bottom one is connected to the external ones. We distinguish between the two sides of optical connections with a letter: A, to the left in fig. 6.8, and B, to the right. Therefore each optical connection can be identified by a letter and a number. On both side A and B, two alignment loops are present used during the packaging. They can be also used to estimate the insertion losses of the device.
MZIs row number goes from 1, to the bottom, to 8, to the top, while their column number goes from 1, at side A, to 8, at side B. Therefore, MZIs can be connected to their electrical pins with the formulas:

$$\begin{aligned} pin_\theta &= (\lceil col/2 \rceil - 1)(N-1) + ((row+1)\%2)(N/2-1) + \lfloor row/2 \rfloor + 3 \\ pin_\phi &= 28 - [(\lceil col/2 \rceil - 1)(N-1) + (((N-row)+1)\%2)(N/2-1) \\ &\quad + \lfloor (N-row)/2) \rfloor] \end{aligned} \tag{6.4}$$

where $\%2$ represents division modulo 2, $\lceil x \rceil$ is the closest bigger integer to x, $\lfloor x \rfloor$ its closest smaller integer and N is the size of the mesh (8 for this chip, the formula is valid for N even).

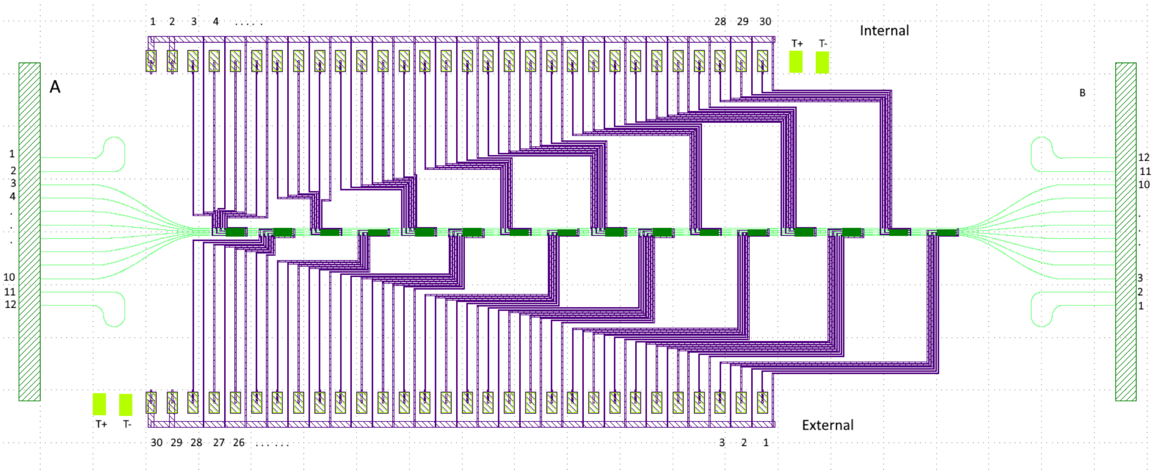


Figure 6.8: Design of the chip with naming convection of the Tyndall packaged chip. On the top: row of electrical connections of the internal TOPS. On the bottom: row of electrical connections of the external TOPS. We point out that internal and external phase shifters have a separate ground, connected to the first two pins to the left on each row of connections. Two additional unconnected pins $T+$ and $T-$ are present on each row of connections. To the left: array of 12 fibers, 4 of which form 2 alignment loops, identified by the letter A. To the right: array of 12 fibers, 4 of which form 2 alignment loops, identified by the letter B.

Custom cables, shown in Fig. 6.10, have been built to connect the device to the SMU: given the packaging and the SMU they had to have a 2×16 connector on one side, and two 2×20 connectors on the other. Green cables have been used for the ground, while yellow ones for the signals.

In fig. 6.9 we can see the electrical connections on the device: on each connector two additional unconnected pins, named T_+ and T_-, can be found, which, together with the ground pins, make them asymmetric. Therefore, they have to be connected with the proper custom cable, which are not interchangeable, and are identifiable by the position of the empty pins and the ground.
On the SMU side, the connector should be placed to have the ground (in green, but also indicated by

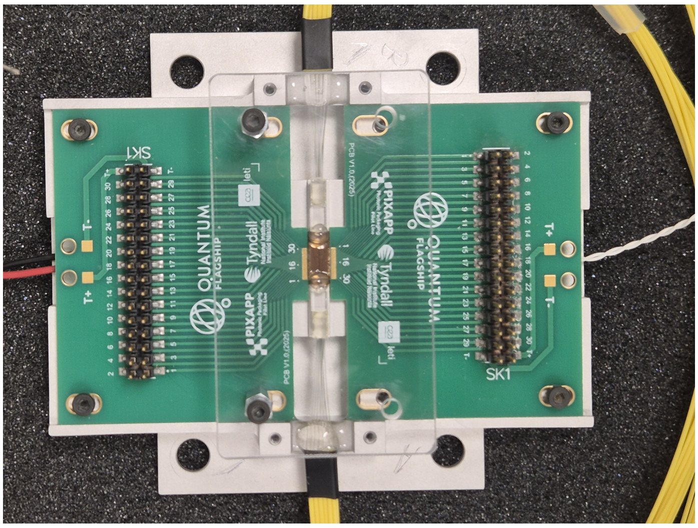
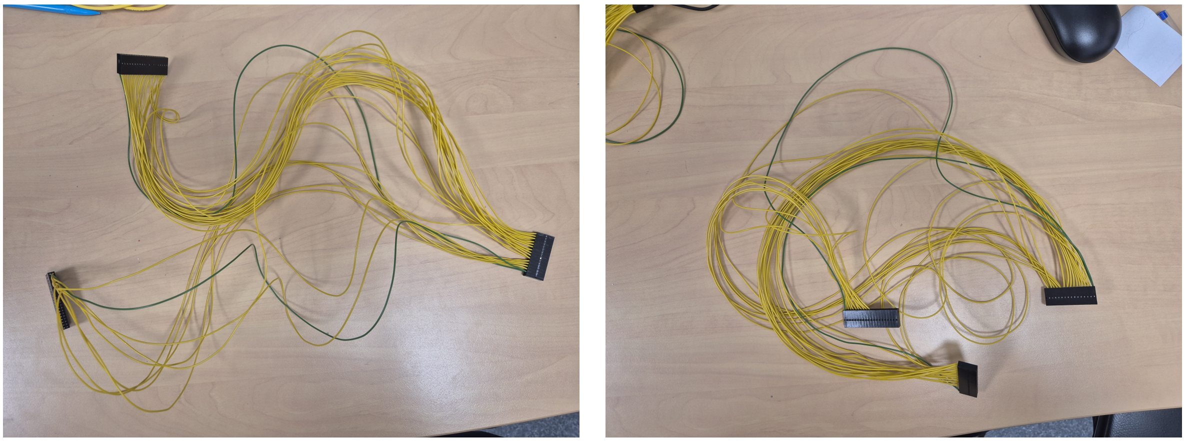


Figure 6.9: A photo of the Tyndall packaged chip



(a) Internal phase shifters custom cable

(b) External phase shifters custom cable

Figure 6.10

the mark on the connector) towards the lower numbers. Following this rule preserves the number of the pins (pin n on the chip will be connected to pin n on the SMU, with the exception of the ground).

Continuity on the custom cables was checked with a tester before use. We also checked continuity between the pins on the chip, which showed a resistance of $\simeq 2.4k\Omega$ between them and the ground. Pin 16 of the external phase shifter side showed no connection.

6.3.1 Electrical characterization

We measured the $I - V$ curve for each heater on the chip, of which a plot can be seen in fig. 6.11. One clearly sees the unconnected pin 16, having zero current in pic 6.11b.

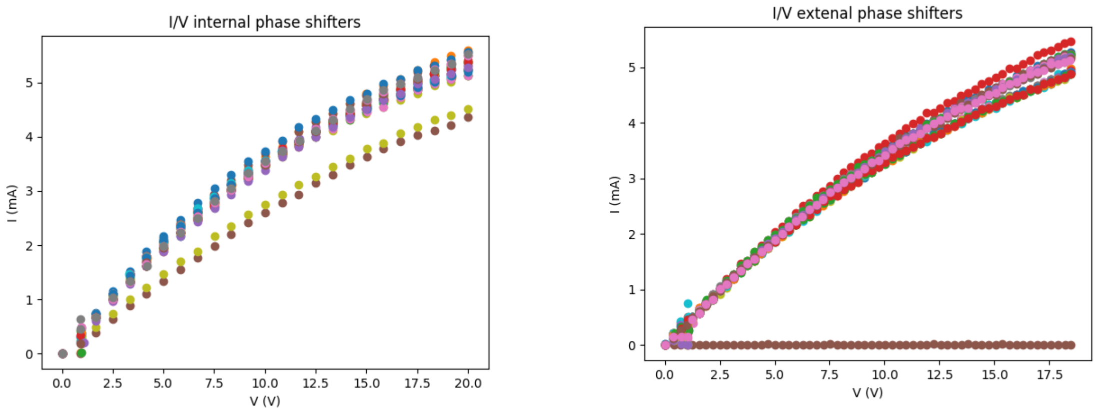
Fitting them with eq. 2.17 gives values for R_0 and V_0 shown in the histograms of fig. 6.12. We find

$$< R_0 >= 2.49k\Omega \qquad \sigma(R_0) = 0.22k\Omega$$

$$< V_0 >= 11.12V \qquad \sigma(V_0) = 1.41V$$

which, considering their nominal length of $180\mu m$, leads to

$$< \rho >= 13.83\Omega/\mu m \qquad \sigma(\rho) = 1.22\Omega/\mu m$$



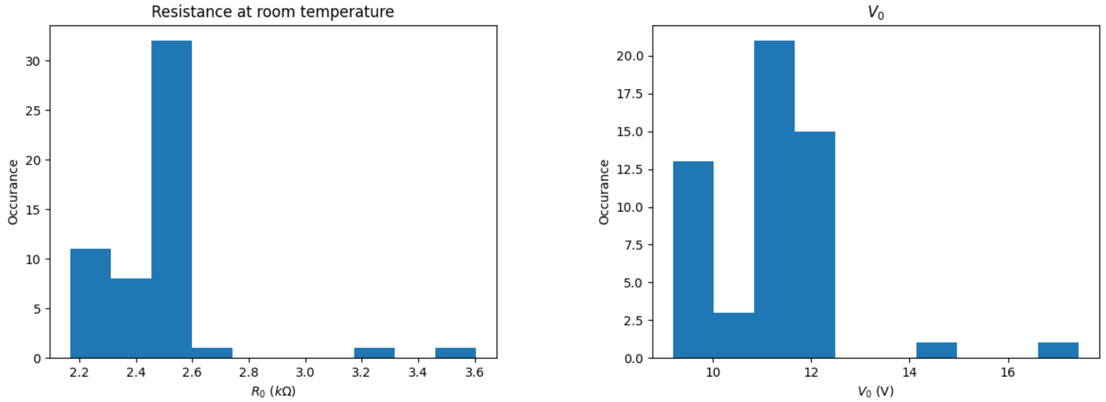
(a) $I - V$ curves of the internal phase shifters

(b) $I - V$ curves of the external phase shifters

Figure 6.11

compatible with the values found for the TOPS studied in section 6.1.

Two of the internal heaters show a significantly higher resistance than the others, probably due to poor connections on the custom cable.



(a) Histogram of the R_0 values found from the fit for the Tyndall packaged chip.

(b) Histogram of the V_0 values found from the fit for the Tyndall packaged chip.

Figure 6.12

6.3.2 Insertion losses

To calculate the insertion losses (IL) on the four alignment loops (identified by the letter of their side and the numbers of their fibers), we injected light in them and measured the output power in function of the polarization.

The experimental setup was changed from the one described above to remove the switches. The path of light during this measurement is schematized below.

Laser → polarization controller → chip (into the alignment loops) → OPM150

With the polarization controller we swept the Stokes parameters $S1$ and $S2$, while $S3$ was fixed to 0 and registered the output optical power (sweep of the linear polarization). As reference we repeated the measurement substituting the chip with a 1m SMF 28 fiber.

The results are reported in the table 6.2, while the plot of the output power of loop $A1 \rightarrow A2$ in function of the swipe parameter ξ ($S1 = \cos(\xi), S2 = \sin(\xi)$) can be seen in fig. 6.13. ILs appear to be similar for loops on the same side, while they differ significantly from A to B. We assume constant insertion losses over the optical connections on one side. The better alignment of side B was also reported by Tyndall. The average IL of each fiber-chip connection are

Side A: 4.25 dB per interface

Side B: 2.3 dB per interface

Aligment loop	Insertion losses (dB)	Pol. Dep. Losses (dB)
$A1 \rightarrow A2$	8.5	0.92
$A11 \rightarrow A12$	8.5	0.8
$B1 \rightarrow B2$	4.5	1.2
$B11 \rightarrow B12$	4.8	0.4

Table 6.2: Insertion losses and polarization dependent losses measured for the 4 alignment loops. ILs are obtained subtracting the maximum of the output optical power in function of the polarization to the reference power. Polarization dependent losses are calculated as the difference between the maximum and minimum output optical power. Average insertion losses of a single chip-fiber interface can be found dividing the values by two.

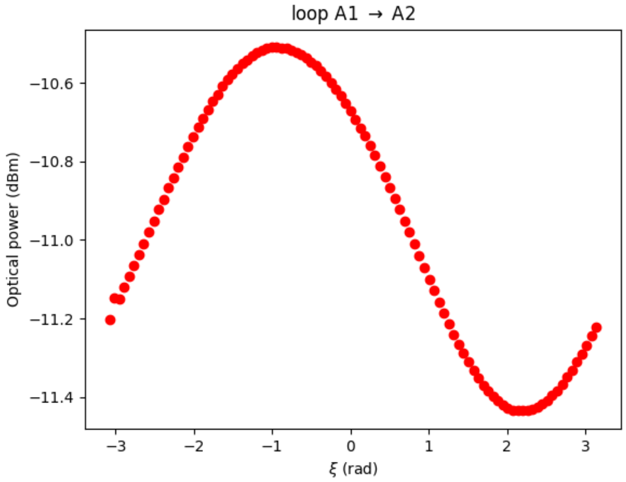


Figure 6.13: Output power of loop $A1 \rightarrow A2$ in function of ξ, where $S1 = \cos(\xi)$ ans $S2 = \sin(\xi)$.

6.3.3 Phase voltage relation

We were interested in comparing different fitting functions for the phase-voltage relation of the TOPS, to find which one empirically described their behavior better. In particular we tried 7 different functions:

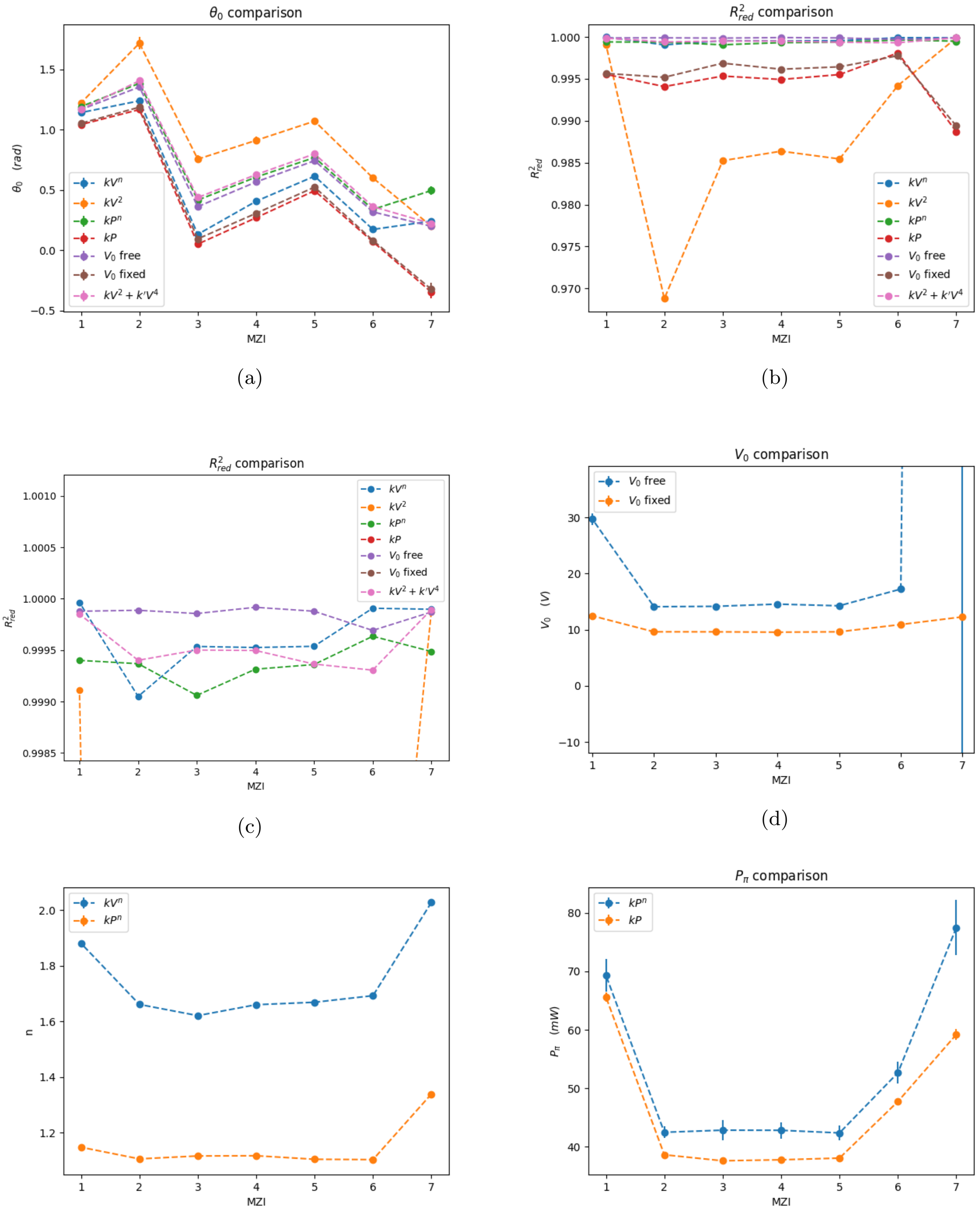
1. $\theta = \theta_0 + kV^2$
2. $\theta = \theta_0 + kV^n$
3. $\theta = \theta_0 + kP_j$
4. $\theta = \theta_0 + kP_j^n$

5. $\theta = \theta_0 + k\left(\sqrt{1+\left(\frac{V}{V_0}\right)^2} - 1\right) \quad V_0 \quad \text{fixed}$
6. $\theta = \theta_0 + k\left(\sqrt{1+\left(\frac{V}{V_0}\right)^2} - 1\right) \quad V_0 \quad \text{free}$
7. $\theta = \theta_0 + kV^2 + k'V^4$

Where the difference between functions 5 and 6 is that in 5 the value of V_0 is fixed to the value previously found by a fit of the $I - V$ curve, while in 6 V_0 is left free to change.

To compare them we calibrated the main diagonal of the mesh, as explained in Ch. 4, and used the different routines to fit the data. Comparison between the found values of θ_0 , n , R^2_{red}, P_π and V_0 from the fits can be seen in fig. 6.14. From fig. 6.14b we observe that methods 2, 4, 6 and 7 fit better the experimental data than the other three. In particular the better accordance of 4 over 3, and 6 over 5 suggest a non linear relation between the power dissipated due to Joule effect by the heater and the temperature of the waveguide.

Between the four better performing methods, two come with their additional problems: 4 uses the dissipated power as the independent variable and 6 shows a high correlation between its parameters, which can lead to absurd results, as seen in fig 6.14d. Therefore, we are left with two viable candidate functions to model the TOPS: 2 and 7. Between them, if we have a look at fig. 6.14a, 7 seems to find values for θ_0 compatible to the ones found by 4 and 6, while 2 seems to find values in between the ones found by 4, 6, and 7 and the ones found by 3 and 5. Also 7 allows for a more complex crosstalk description, since it is the truncation of Eq. 3.20 to the second order. On the other hand this added complexity might be unnecessary if the simpler relation 2 can describe the system equally well.
That said, we preferred to adopt 2 as the standard empirical function to describe the TOPS, for its simpler form. Its main drawback is the lack of a direct connection to the physical properties of the system. Though, even if we were to assign a physical interpretation to the parameters of relation 7, our lack of knowledge about some key properties of the device (the temperature coefficient of the resistance of the heaters, the heat diffusion coefficient) would prevent us from extracting meaningful information from them. In other words, apart from the possibility to have a more detailed description of thermal crosstalk, we don't find any benefits into adopting a more complex relation. Still, as it will become clearer later, it is already challenging to measure the first order crosstalk matrix of Eq. 3.20. Characterizing the second order one is too complex and would provide limited benefits.



(a) (b)

(c) (d)

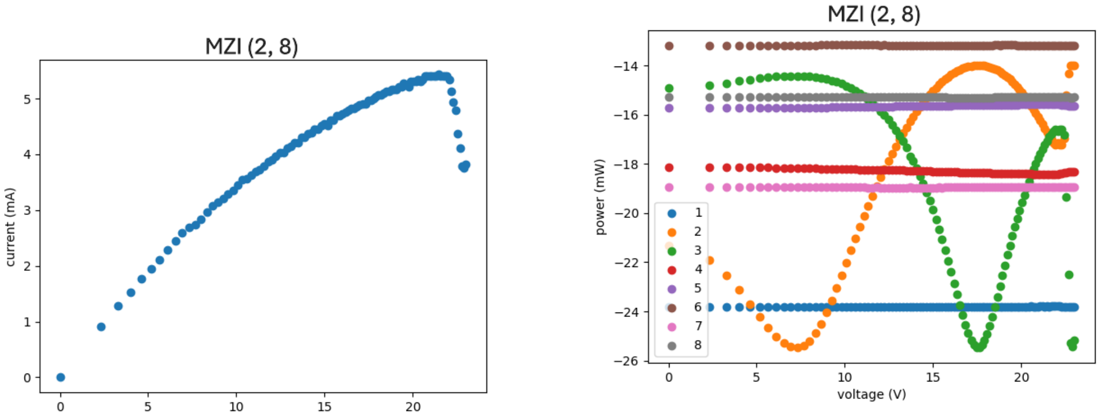
(e) (f)

Figure 6.14: Comparison between the different fit functions. The x axis refers to the considered MZI, which, being on the main diagonal, can be identified by a single number. a) θ_0 found for the MZIs with the different methods. b) R^2_{red} found for the different fits for the output power of the MZIs. c) Zoom of b) d) Values of V_0 found with two methods. The value for MZI (7,7) (out of the graph) is the absurd $V_0 = 1073 \pm 13129V$. Nonetheless, we observe a good agreement between the experimental data and the fit, this is a consequence of the high correlation found by the fit between the parameters V_0 and k. e) n found for the MZIs with the different methods. f) P_π found for the MZIs with the different methods.

6.3.4 Heaters burning point and observation of thermal crosstalk

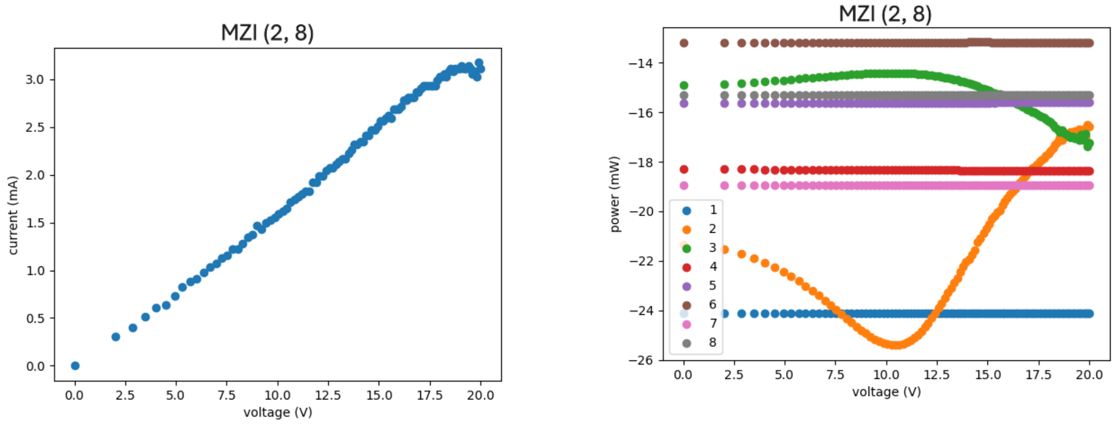
From the previous measurement, we observed that limiting the applied voltage to $20V$ did not allow to reach a complete 2π phase swipe for some MZIs. Therefore, we tried to raise the maximum voltage to $22V$, which proved to be too high, causing some heaters to burn. We observed the $I-V$ curve and optical power outputs of the chip while heaters started burning, of which a plot is shown in figs. 6.15 and 6.16. The TOPSs seem to burn at around $20.5-21V$ and, after burning, their resistance increases, as shown in fig. 6.15c, additionally limiting the achievable phase offset. We decided to limit the maximum voltage applied to the heaters to the conservative $16.5V$ to prevent damage to other TOPS.

The effect of thermal crosstalk can also be clearly seen in figs. 6.16b . In fact, we observe that controlling the internal phase shifter of MZI (4,8), which should only affect the optical power of outputs 4 and 5, also influences 2 and 3.



(a) $I-V$ curve of the internal TOPS of MZI (2,8) when it burned

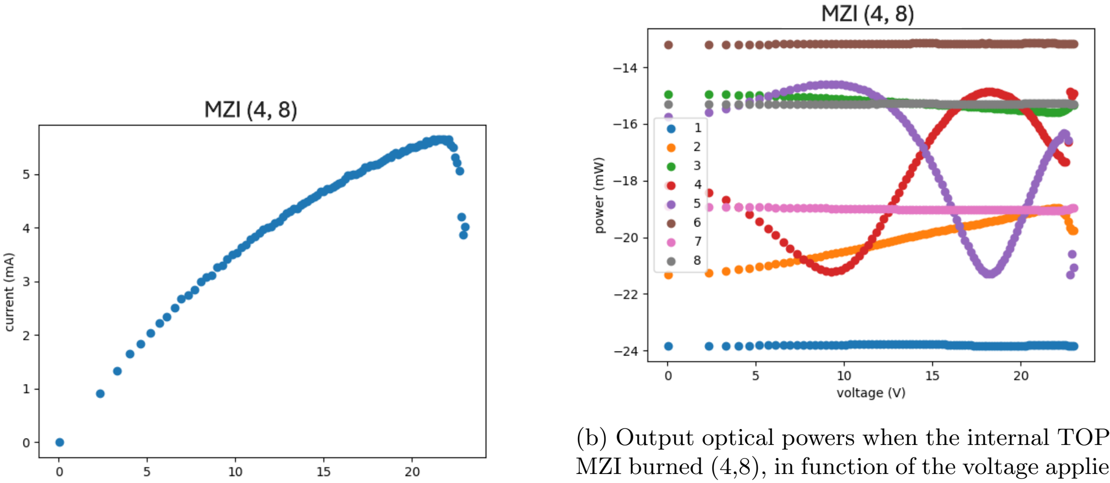
(b) Output optical powers when the internal TOPS of MZI (2,8) burned, in function of the voltage applied to it.



(c) $I-V$ curve of the internal TOPS of MZI (2,8) after it burned.

(d) Output optical powers after the internal TOPS of MZI (2,8) burned, in function of the voltage applied to it.

Figure 6.15: On the left: I-V curves of the TOPS. On the right: related optical power outputs, the curves are named after their optical output.



(a) $I - V$ curve of the internal TOPS of MZI (4,8) when it burned

(b) Output optical powers when the internal TOPS of MZI burned (4,8), in function of the voltage applied to it. We also clearly see first neighbor thermal crosstalk, since the optical powers of output 2 and 3 also are impacted by heating MZI (4,8).

Figure 6.16: On the left: I-V curves of the TOPS. On the right: related optical power outputs, the curves are named after their optical output.

6.4 In-house packaged chip

Another chip was packaged in house by Olivier Castany from CEA Leti. It differs from the Tyndall packaged one by the fact that, to connect side A, polarization maintaining PM780-HP fibers have been used, which allowed to know the polarization of the injected light. They were chosen to reduce insertion losses, by better matching the mode diameter of the fiber to the mode diameter on the chip. On the other hand, since they are not meant for 1550 nm light, small bends result in high radiative losses. Therefore, extra precaution had to be taken to ensure that they were sufficiently straight during measurement.

Fig 6.17 shows the design of the chip with the nomenclature of the optical and electrical connections found on the in-house packaged chip. MZI numbering follows the same described for the Tyndall packaged chip.

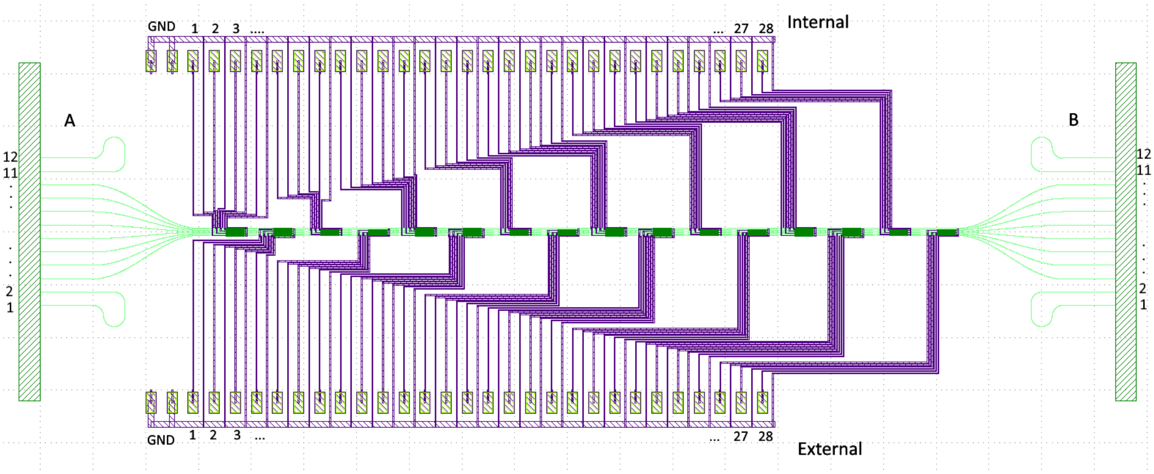


Figure 6.17: Design of the chip with the numbering present on the in-house packaged chip. On the top: row of electrical connections of the internal TOPS. On the bottom: row of electrical connections of the external TOPS. We point out that internal and external phase shifters have a separate ground. To the left: array of 12 fibers, 4 of which form 2 alignment loops, identified by the letter A. To the right: array of 12 fibers, 4 of which form 2 alignment loops, identified by the letter B. Input/output fibers are numbered from the bottom to the top. Electrical connections are numbered from left to right, skipping the first two pins, which connect to the ground.

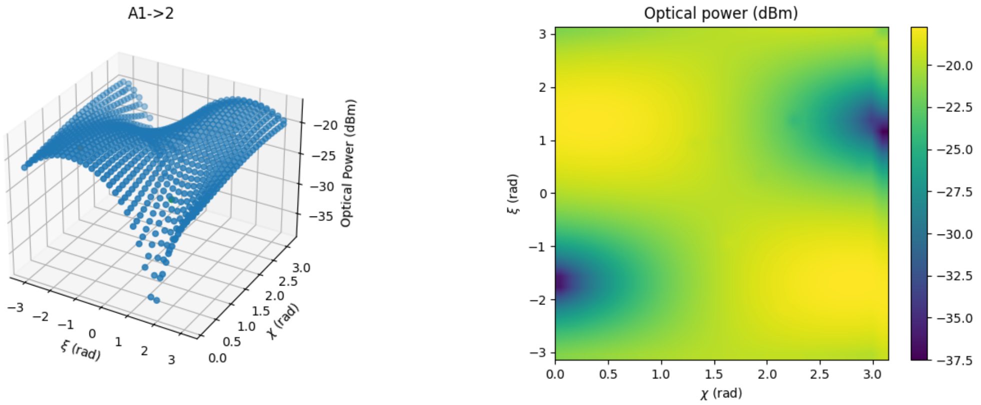
An MZI (row, col) is related to its electrical pins via the formulas:

$$\begin{aligned} pin_\theta &= (\lceil col/2 \rceil - 1)(N-1) + ((row+1)\%2)(N/2-1) + \lfloor row/2 \rfloor + 1 \\ pin_\phi &= (\lceil col/2 \rceil - 1)(N-1) + (((N-row)+1)\%2)(N/2-1) \\ &+ \lfloor (N-row)/2) \rfloor + 1 \end{aligned} \tag{6.5}$$

Electrical connection were made to preserve the number of the pins between the circuit and the SMU (following the covention that the mark on the connector should be attached towards the lower numbers).

6.4.1 Insertion losses

We measured the insertion losses on the alignment loops of the chip, following a similar approach as the one used for the Tyndall chip, with the difference that polarization was swept across the whole Poincarré sphere(a sweep of the linear polarization would have been sufficient). To measure the reference value to calculte the ILs, for side A, a $2m$ PM-780 fiber was used, but, due to the high dependence of losses on its position, an accurate measurement was not possible. The results are reported in table 6.3 .



(a) (b)

Figure 6.18: a) Measured optical power in function of the polarization angles ξ and χ, where $S1 = \cos(\xi)\cos(\chi)$; $S2 = \sin(\xi)\cos(\chi)$; $S3 = \sin(\chi)$. b)Its heat map representation.

Aligment loop	Insertion losses (dBm)	Pol. Dep. Losses (dBm)
$A1 \rightarrow A2$	–*	25.3
$A11 \rightarrow A12$	–*	24.9
$B1 \rightarrow B2$	10.6	2.3
$B11 \rightarrow B12$	10.4	5.2

Table 6.3: Insertion and polarization dependent losses measured for the 4 alignment loops. To find the average losses per single fiber chip interface the values have to be divided by two. High uncertainty on the reference value used to calculate the insertion losses of side A did not allow to find their value.

Side A exhibits high polarization dependent losses, introduced by the interface between PM1550 and PM780 fibers (this is confirmed by the fact that similar losses are observed also when the chip is replaced by a PM780 fiber to measure the IL reference value). Fig. 6.18 shows the output optical power of alignment loop $A1 \rightarrow A2$ in function of the polarization angles. Again, on side B, we observe similar values for the insertion losses for the two alignment loops. We assume them to be constant through the inputs and outputs of the device, which will be important when multiple optical power outputs of the chip will be considered simultaneously. Since only one optical input has been used at a time, it is less important to characterize the ILs on side A. For side B we report IL = 5.5 dB per interface.

We observe the maximum of transmission at $S2 \simeq -1$, while its minimum at $S2 \simeq 1$. We interpret this polarization states to be TE and TM respectively.

6.4.2 Electrical characterization

As done for the Tyndall packaged chip, the $I - V$ curve of each TOPS has been measured and fitted to find V_0 and R_0. A histogram of the results is shown in fig. 6.19. We find

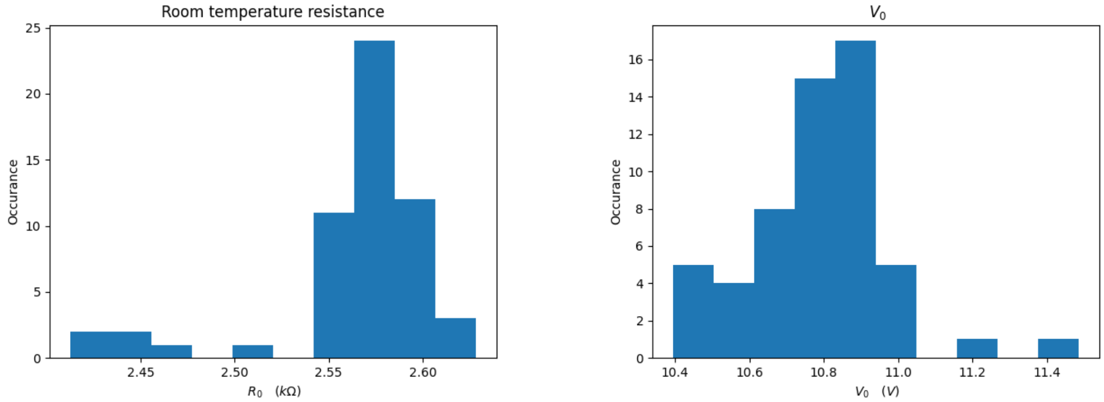
$$< R_0 >= 2.56k\Omega \qquad \sigma(R_0) = 0.04k\Omega$$

$$< V_0 >= 10.78V \qquad \sigma(V_0) = 0.18V$$

which, considering their nominal length of $180\mu m$, leads to

$$< \rho >= 14.22\Omega/\mu m \qquad \sigma(\rho) = 0.2\Omega/\mu m$$

compatible with the values found for the Tyndall chip.



(a) Histogram of R_0 values found for the CEA Leti packaged chip.

(b) Histogram of V_0 values found for the CEA Leti packaged chip.

Figure 6.19: Histograms of the parameters found from the fit of the $I-V$ curves of the heaters with equation 2.17.

6.4.3 Splitting coefficients characterization

We characterized the splitting coefficients of the DC of MZIs (2,8) and (6,8), using the method described in [25] and a fit, under the assumption of path independent losses inside the chip.

The two MZIs were chosen since they are relatively easy to isolate, the procedure followed was:

1. Inject light in input 1 (8).
2. By swiping the voltage applied to the internal TOPS of MZI (1,1) ((7,1)), minimize the optical power of output 8 (1).
3. By swiping the voltage applied to the internal TOPS of MZI $(i, i+2)$ for $1 \leq i \leq 5$ ($(8-i, 8-i+2)$ for $1 \leq i \leq 5$) maximize the sum of the optical powers of outputs 6 and 7 (2 and 3).
4. By swiping the voltage applied to the internal TOPS of MZI (i, i) for $1 \leq i \leq 5$ ($(8-i, 8-i)$ for $1 \leq i \leq 5$) minimize the the optical powers of output 8 (1).
5. Swipe the voltage of MZI (6,8) ((2,8)) and measure the optical power of outputs 6 and 7 (2 and 3).
6. Correct isolation of the MZI can be checked by swiping the preceding external phase shifter, which does not impact the power outputs whenever lights enters the MZI in only one input.

This procedure redirects most of the light into one input of the MZI, while minimizing light entering from the other one, allowing to observe both output power of an isolated MZI in function of the voltage applied to its internal TOPS .

The splitting coefficients, adjusting the results of [25] to symmetric MZIs, can be found as

$$\begin{aligned}\delta k_1 &= (-1)^l \sqrt{\frac{\frac{1}{4} + P^2 - B^2 + (-1)^m \sqrt{(B^2 - P^2 - \frac{1}{4})^2 - 4P^2}}{2}} \\ \delta k_2 &= \frac{P}{\delta k_1}\end{aligned} \tag{6.6}$$

where

$$\begin{aligned}
B &= \sqrt{(\frac{1}{4}-\delta k_1^2)(\frac{1}{4}-\delta k_2^2)} = \frac{\frac{R_{14}-1}{R_{14}+1}}{1+\frac{(R_{13}+1)(R_{14}-1)}{(R_{13}-1)(R_{14}+1)}} \\
P &= \delta k_1 \delta k_2 = \frac{1}{4}\frac{\frac{(R_{13}+1)(R_{14}-1)}{(R_{13}-1)(R_{14}+1)}-1}{\frac{(R_{13}+1)(R_{14}-1)}{(R_{13}-1)(R_{14}+1)}+1} \\
R_{13} &= \frac{P_{1max}}{P_{1min}} \\
R_{14} &= \frac{P_{2max}}{P_{2min}}
\end{aligned} \tag{6.7}$$

and where l, m can be 0 or 1, giving four possible solutions, a consequence of the fact that the system is invariant under the exchange of δk_1 and δk_2, or under the change of both of their signs (we chose the pair $l = 1, m = 1$). Bootstrapping using normal distributions with σ extracted from fig. 6.7a was used to assign a confidence interval to the results.

To validate the result, the splitting coefficients were also extracted via a fit of the optical powers P_1 and P_2 (in function of the dissipated power) with equation:

$$\begin{aligned}
P_1 &= A\left(\frac{1}{2}+2\delta k_1\delta k_2 - 2\sqrt{(\frac{1}{4}-\delta k_2^2)(\frac{1}{4}-\delta k_1^2)}\cos(kP_j^n+\theta_0)\right) \\
P_2 &= A\left(\frac{1}{2}-2\delta k_1\delta k_2 + 2\sqrt{(\frac{1}{4}-\delta k_2^2)(\frac{1}{4}-\delta k_1^2)}\cos(kP_j^n+\theta_0)\right)
\end{aligned} \tag{6.8}$$

P_2 was scaled beforehand to account for the slightly different insertion losses. The plots of the fits and of the distributions found via bootstrapping are shown in fig. 6.20. The found values are reported in table 6.4.

Splitting coefficient	Bootstrapping	Fit
(2, 8) $\delta k1$	-0.0223 ± 0.0003	-0.0206 ± 0.0006
(2, 8) $\delta k2$	0.0437 ± 0.0003	0.0399 ± 0.0006
(6, 8) $\delta k1$	0.0328 ± 0.0003	0.0311 ± 0.0006
(6, 8) $\delta k2$	0.0405 ± 0.0003	0.0391 ± 0.0006

Table 6.4: Values found for the spitting coefficients of MZI (2, 8) and (6, 8) by the two different methods. Except for δk_2 of MZI (2,8), the values measured by the two different methods fall between 1.5 and 2 σ. The two values found for δk_2 of MZI (2,8) differ significantly. We might have underestimated the confidence intervals, or additional problems may have arised. Nonetheless, their average should give a good estimate of the true value.

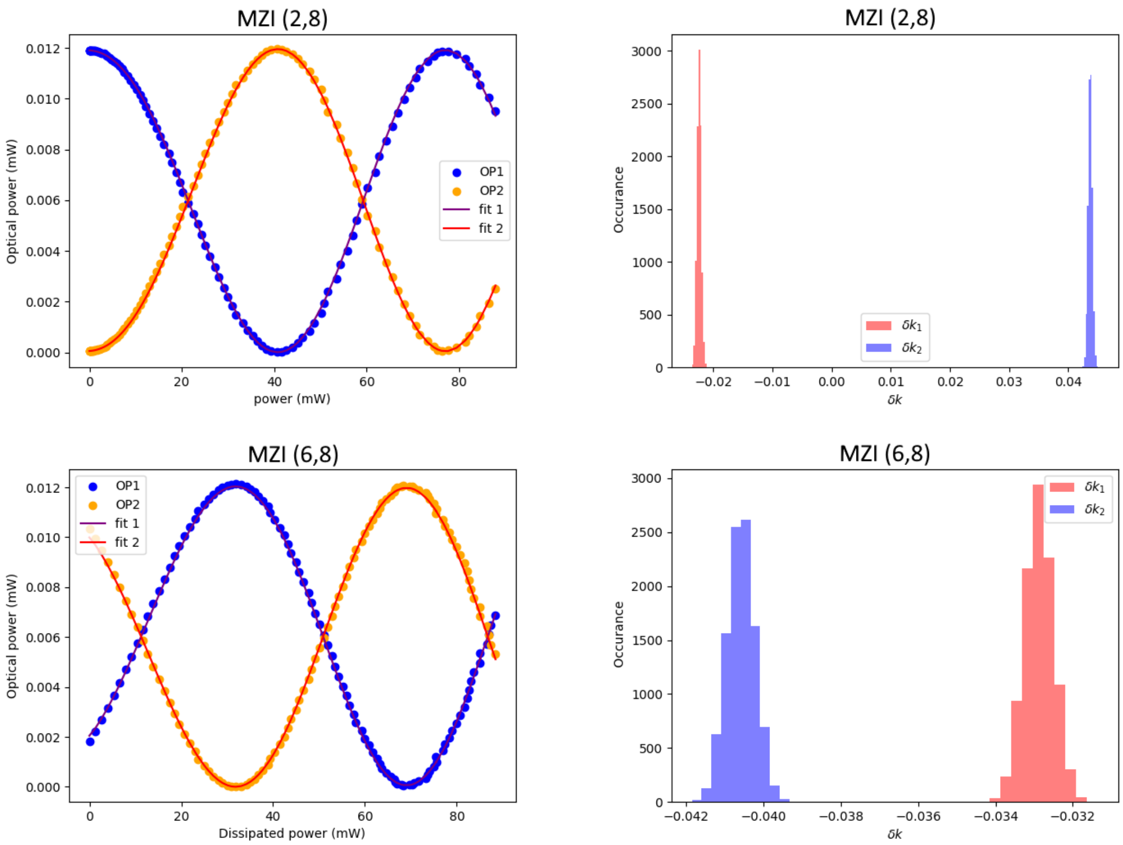


Figure 6.20: To the left: the output optical powers and their fit of the respective MZIs. To the right: histogram of $\delta k_1, \delta k_2$ found via bootstrapping on method of [25].

6.4.4 Calibration of the internal phase shifters

We performed the calibration algorithm for the internal phase shifters, assuming a phase voltage relation of the form

$$\theta = \theta_0 + kV^n \tag{6.9}$$

Histograms of the found values for the offsets θ_0, self heating coefficients k, exponents n and a heat map of the reduced crosstalk matrix can be seen in fig. 6.22. Crosstalk was characterized only for first neighbors. We observe small offsets ($< \pi$) and around 1.5% of crosstalk between first neighbors.

The characterization of the crosstalk matrix is significantly impacted by the unbalancements in the device. fig 6.21 shows the plot of the data used to calibrate crosstalk between two pairs of neighboring MZIs. In the first calibration, the measured optical power behaved as expected. The second one shows unexpected oscillations in the measured optical power due to parasitic light, which degrades the results. Still, the number of affected measurements was limited.

A direct consequence of performing a partial calibration is a limit in the matrices that is possible to implement. In particular without knowledge of the workings of the external phases, only the permutations can be programmed on the device. Consequently, the calibration was tested by programming the identity(I) and the swap operator (X, with the swap we refer to the operator sending input i to output $9-i$), both disregarding and taking into account crosstalk. In the former case we found the absolute fidelity

$$F_{abs}(I, I_{exp}) = 0.92 \qquad F_{abs}(X, X_{exp}) = 0.86$$

while in the latter

$$F_{abs}(I, I_{exp}) = 0.94 \qquad F_{abs}(X, X_{exp}) = 0.87$$

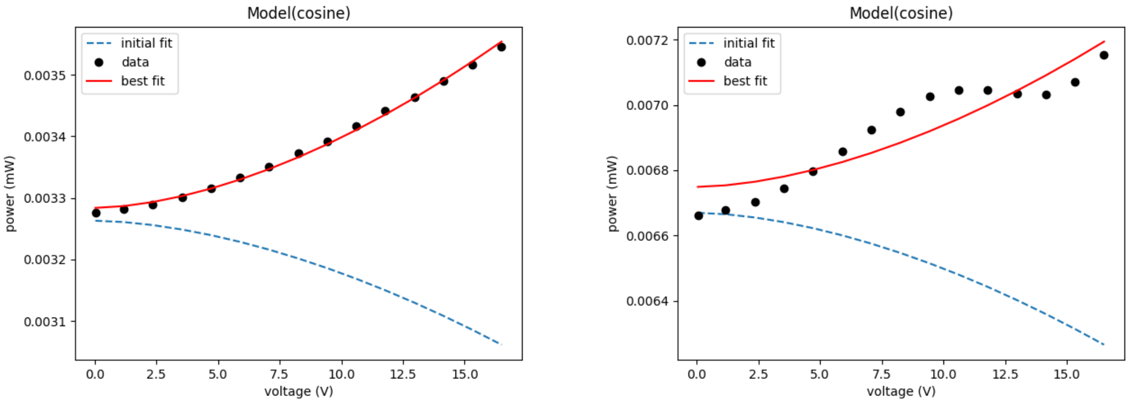
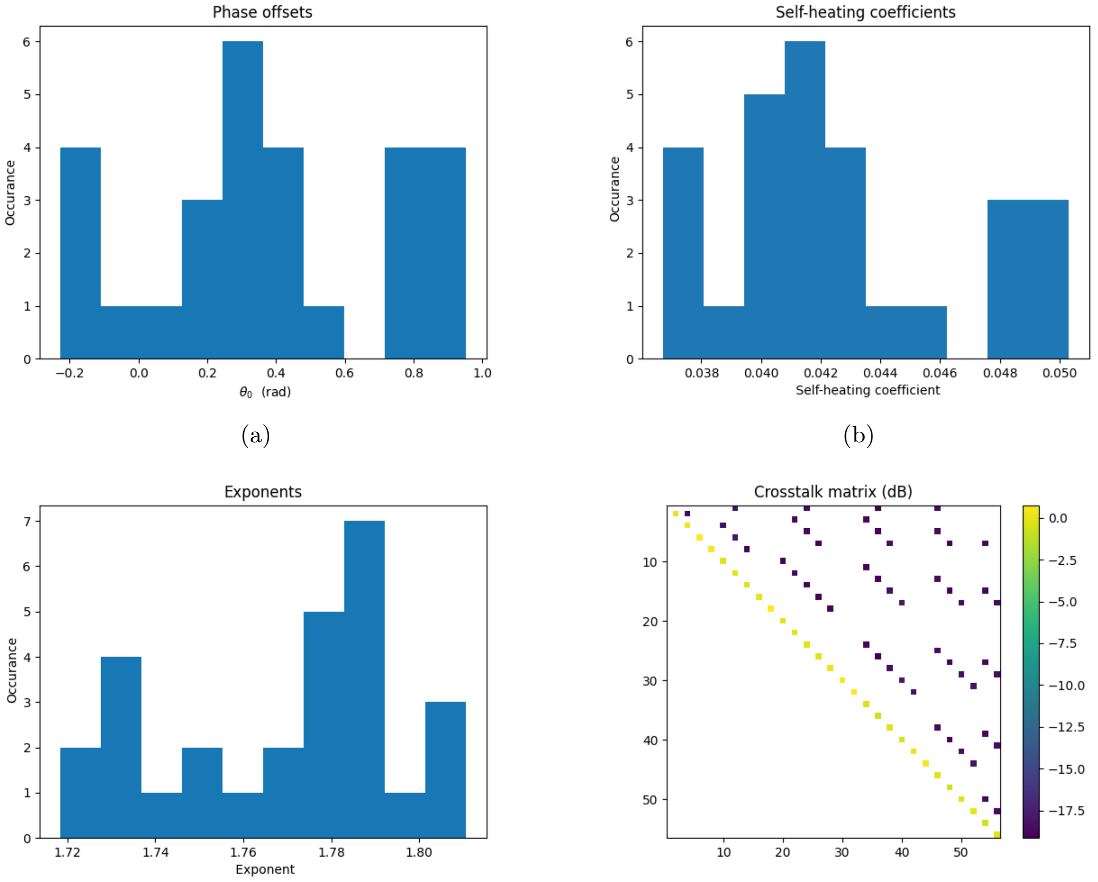


Figure 6.21: Illustration of the effect of parasitic light in the characterization of thermal crosstalk. Data used to characterize crosstalk between two pairs of MZIs: on the left an example of a correct characterization. On the right we observe a significant deviation of the data from the expected behavior, due to parasitic light. The initial fit was performed assuming $\sim 1.5\%$ of crosstalk, it appears mirrored with respect to the data due to a wrong assumption on the sign of one of its coefficients. This error does not impact the final result.

where the subscript $_{exp}$ refers to the measured transfer matrix of the device, when programmed to that transformation. Even under such simple assumption on the crosstalk's nature, and also considering the errors on its characterization, we observe a slight improvement in fidelity when thermal crosstalk is taken into account. Heat maps representation of the measured experimental transfer matrices can be seen in fig. 6.23 .

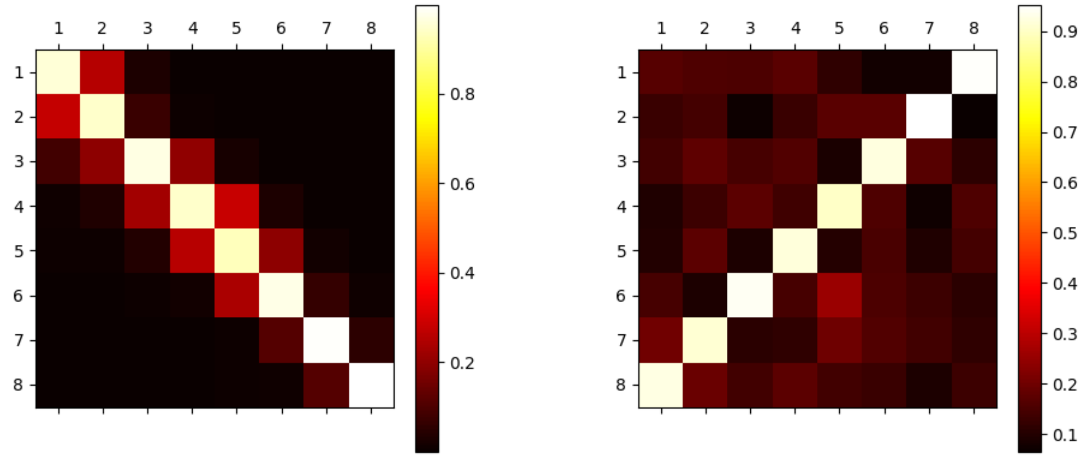
The calibration proved to be stable in time, still programming the device with a comparable fidelity even after weeks of operation. Similar fidelities were also obtained using the different phase voltage relations previously discussed. Since they did not show any significant improvements in the results, we decided to continue using the simpler relation of Eq. 6.9.



(a) (b)

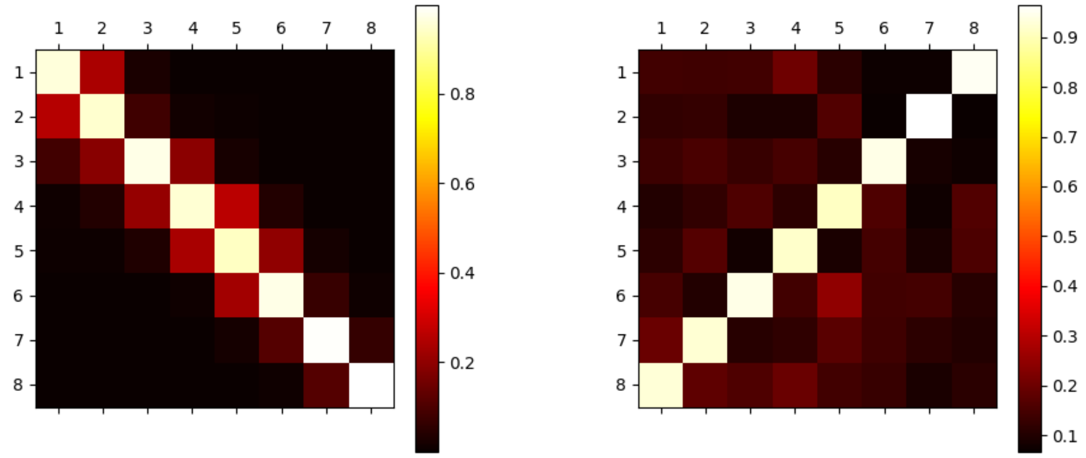
(c) (d)

Figure 6.22: Values found by the calibration: a) Histogram of the phase offsets θ_0. b) Histogram of the self heating coefficients k. c) Histogram of the exponents$n$. d) Heat map representation of the reduced crosstalk matrix (in dB), normalized over the average of the self heating coefficients.



(a) Identity disregarding crosstalk

(b) Swap disregarding crosstalk



(c) Identity taking into account crosstalk

(d) Swap taking into account crosstalk

Figure 6.23: Heat maps representation of the transfer matrices of the device when programmed to different transformations: to the left identity, to the right swap.

6.4.5 Gradient descent optimization

We tested the possibility to use gradient descent optimization on the chip as a mean to improve the outcome of the programming via the calibration obtained as described before. A simple goal was chosen as benchmark: sending light from input n to output n, which, in a perfect device, has the trivial solution of programming the MZIs in rows n and $n-1$ to the bar state.
We applied the proper voltages to the internal TOPS using the found calibration, then they were optimized to maximize light in output n.

The plot of the applied voltages and performance (intensity of light in output n, normalized) in function of the optimization step for rows $n = 1, 2, 3, 4$ can be seen in fig. 6.24. We observe that the same MZIs, when optimized for different n, converge to different values for the applied voltage. This behavior suggest that the final found states are not the ones where all the MZIs are in the bar state. Instead, the found solutions are the better ones 'globally', where the MZIs correct each others errors.

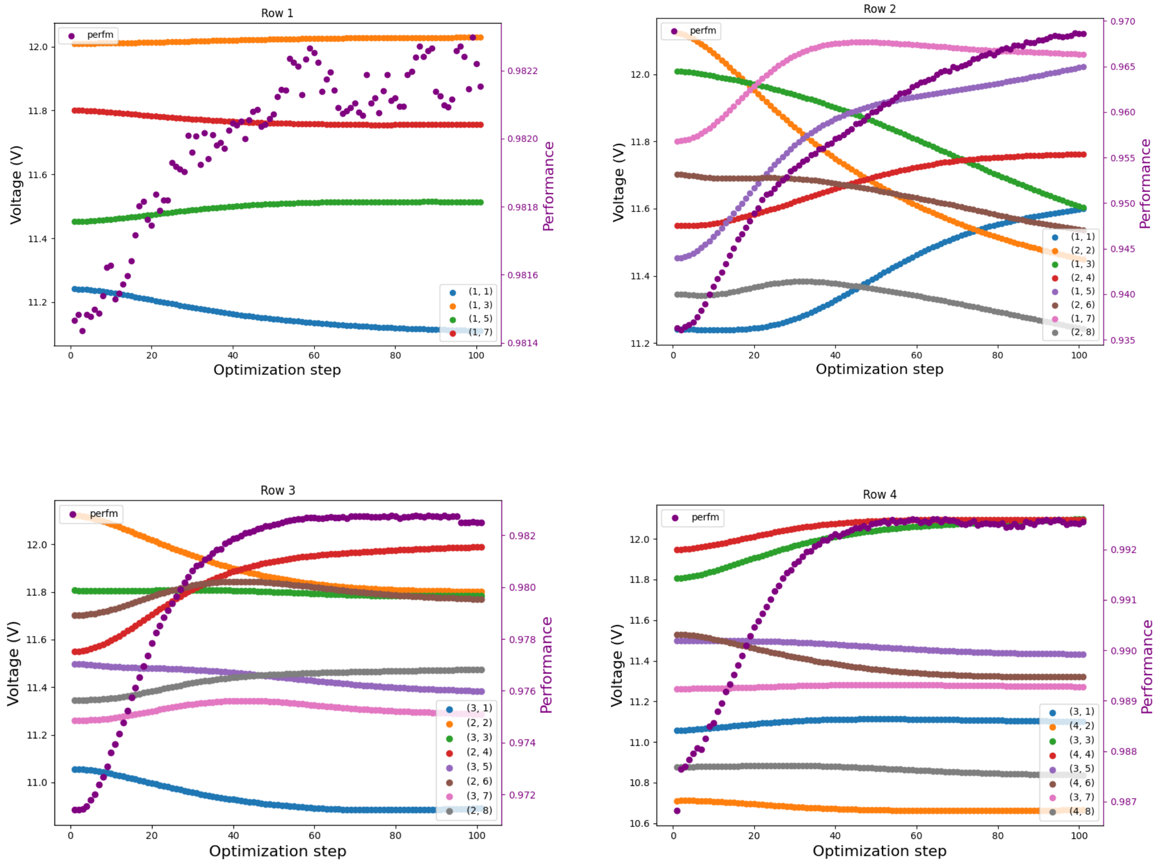


Figure 6.24: Applied voltages to the internal TOPS (labeled after their MZI) and performance of four rows of MZIs in function of the optimization step. We can see that the same MZI can converge to a different voltage for different rows, e.g. comparing the result for $n = 1$ and $n = 2$.

6.4.6 Programming an arbitrary unitary

As discussed before, the purpose of performing a full characterization of the device was the possibility to program any arbitrary matrix, without the need to use a subsequent optimization algorithm. Unfortunately, we were not able to calibrate the external phase shifters, but we were still curious to try to implement an arbitrary transformations.
To circumvent our lack of knowledge of the full chip, we assumed all DC to be perfect and the external phase shifters to behave as the internal ones of the same MZI. The unitary matrix was randomly

chosen, under the constraint that it should be implementable with voltages under $16V$, as to not be limited by the physical constraints of the device. This was achieved by reverse engineering the Clement's decomposition: 56 phases were extracted randomly in a range normally implementable by the TOPS on the device, than the target matrix U_T was constructed from them (as shown [19], uniformly extracting the phases of the Clement's decomposition will not yield an matrix uniformly extracted from the unitary matrices. This does not concern us particularly, since we are not trying to build a statistic and we are also limited by the maximum voltage applicable to the device) .

With the calibration obtained we programmed U_T on the device. The measured implemented transfer matrix U_{Texp} showed an absolute fidelity with the target of

$$F_{abs}\left(U_T, U_{Texp}\right) = 0.86.$$

Certainly its value would be higher if the external TOPS were precisely calibrated.

Subsequently, we tested the genetic optimization algorithm, explained in [6], on the device to improve F_{abs}. We chose to let it run for 90 generations with a population of 20 individuals. We observed a steady increase of $0.1 (\sim 11\%)$ in the absolute fidelity. The constraint of time forced us to choose a small population size and to let it run for liùmited number generations, though we observe that the performance did not reach a plateau. If continued, this optimization should have further improved the result.

A plot of the heat maps of the target matrix, the initially programmed transfer matrix, the final transfer matrix found by the genetic optimization and the fidelity in function of the optimization cycle can bee seen in fig. 6.26.

We then implemented 500 random unitary matrices, without performing any optimization, to find an approximate measurement of $\mathscr{F}_{abs}$. The matrices were extracted as explained above. As already mentioned this leads to non uniformly distributed random unitary matrices, which may bias the found estimate of $\mathscr{F}_{abs}$. We found

$$\mathscr{F}_{abs} = 0.8 \qquad \sigma(F_{abs}) = 0.05$$

an histogram of the results can be seen in fig. 6.25

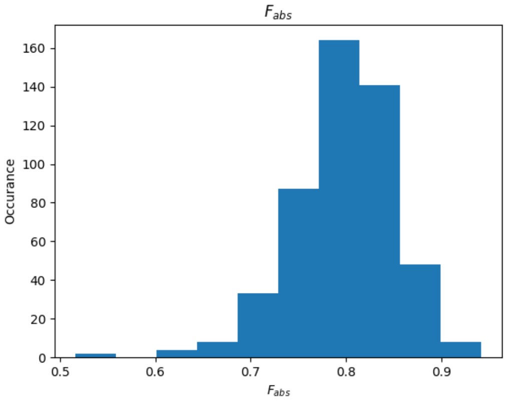
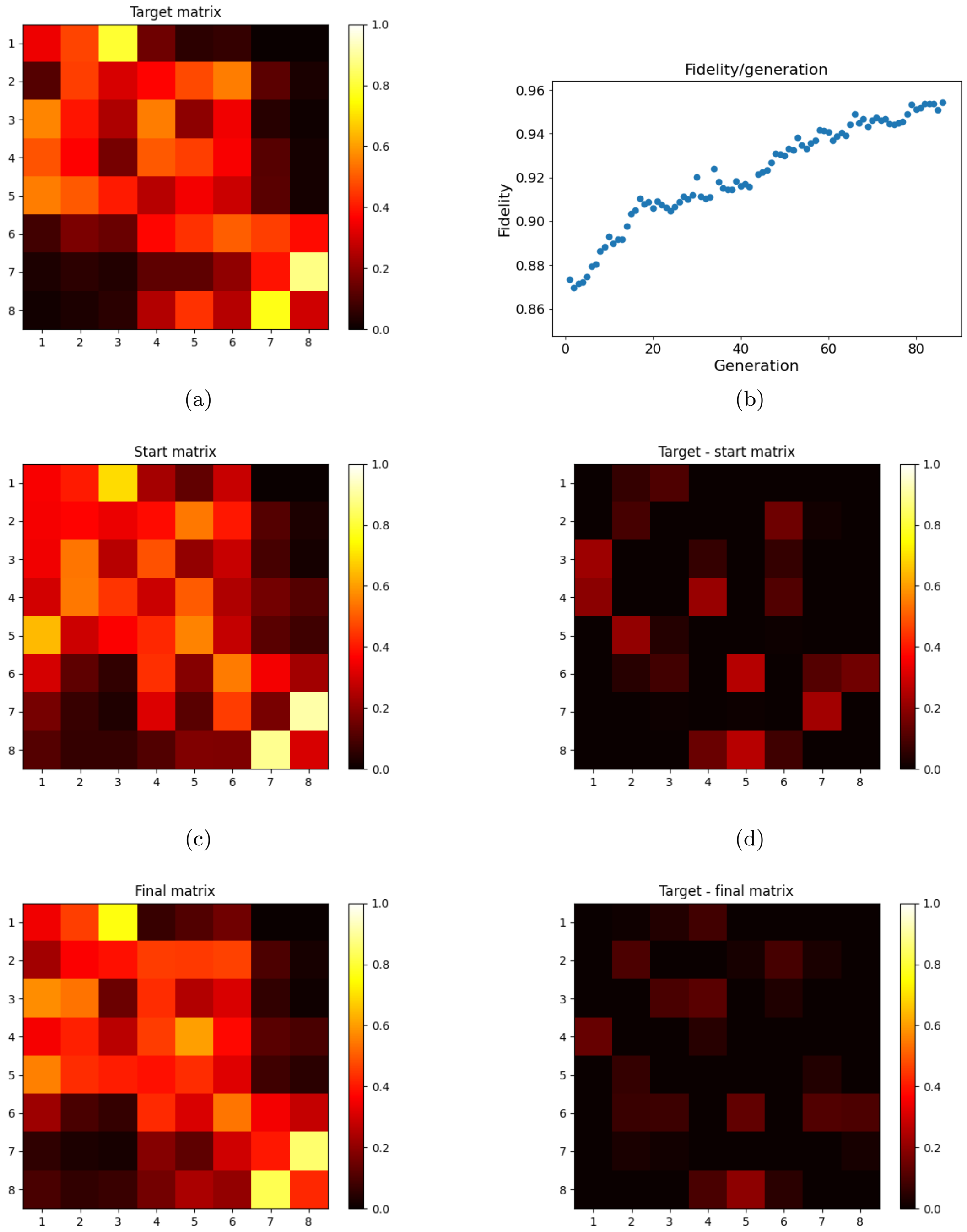


Figure 6.25: Histogram of the found F_{abs} for 500 randomly selected unitary matrices.



(a)

(b)

(c)

(d)

(e)

(f)

Figure 6.26: a) Heat map representation of the target matrix b) plot of the absolute fidelity between the best performing individual of the generation and the target matrix in function of the generation c) Starting transfer matrix found via the calibration and also used as the starting point for the genetic algorithm d) Difference between the target matrix and the starting one e) Transfer matrix of the best performing individual of the last generation f) Difference between the target matrix and the final one

6.5 Summary of the results

We performed a first basic characterization of the device which allowed us to study both the properties of the single components of the device (TOPS, DC) and the their collective behavior.

The TOPS showed significant variance in their parameters across the chip, and also exhibited a slightly non linear relation between the dissipated power and the introduced phase. In general, not all heaters are able to perform a full 2π range before burning; while this may not be necessary for the internal phases, it still limited our ability to reach the cross state in some MZIs. Also, for a full programming, the external phase need to access the full range.

Even though too little DC were characterized to build a significant statistic, we observed a significant variance in the values of the few measured splitting coefficients. Such differences, arising from the fabrication processes, significantly impacted our ability to calibrate the external phase shifters.

The chip was programmed to some simple transformations using the found calibration, with good results. We observed the presence of thermal crosstalk, which, although small, had an observable impact on the performances of the device. A first attempt at implementing an arbitrary matrix was performed, but the quality of the result was limited by our partial knowledge of the chip. Also, two different optimization algorithms, gradient descent and the genetic algorithm, have been tested to improve the performance of the programmed states. They were able to significantly improve the obtained results.

Chapter 7

Conclusions

After an in-depth mathematical description of the device was presented, different calibration approaches were introduced and compared, and their strength and weaknesses were highlighted. Then, through a simulation, we studied the performance of the device and of the calibration in regard to the entity of the non-idealities of the chip. One calibration algorithm demonstrated an intrinsic robustness to unbalancements. In contrary, thermal crosstalk proved to be a significant challenge to characterize with a simple analysis of the interferometric patterns of the device. Also, we observed that a poor characterization of the crosstalk matrix severely impacts the achievable fidelity with a target matrix. Though, machine learning based calibration methods allow to significantly improve the accuracy of the crosstalk matrix characterization, allowing excellent control over the device.

Two chips were packaged for testing, which allowed to measure key characteristics of their components, e.g. the splitting coefficients of some directional couplers. We also performed a partial calibration of one of the devices, and were able to program multiple arbitrary matrices with an average absolute fidelity of $F_{abs} = 0.8$. Then, we tested the possibility to improve the quality of the programming on chip via gradient descent and genetic algorithms. In particular, the genetic algorithm showed promising results in improving the fidelity between the target and the implemented matrix while being relatively simple to implement. In fact, it was able to boost the absolute fidelity achieved with a randomly chosen unitary matrix from $F_{abs} = 0.86$ to $F_{abs} = 0.96$, in only 90 generations. Instead, on chip gradient descent showed promising results in finding accurate solutions to simpler problems involving few MZIs.

Still, work is far from being finished. Further optimization of the simulation of the chip is necessary to reduce computing times. Faster simulations would allow a more in-depth study of the performance of the device under more general assumptions on the behavior of crosstalk, of which the complexity in this work was mainly limited by the computational resources. In particular, GPUs may be used to significantly speed up the matrix multiplications at the root of the simulation. The faster computation time would also enable machine learning based calibration methods for the calibration of the physical chip.

On the experimental side, the calibration of the external TOPS still has to be performed. If done, it would open the possibility of more refined measurements on the chip and more precise operation. In fact, the calibration of the external TOPS would enable the full characterization of the directional couplers, as well as the implementation of the machine learning based calibration on the physical chip. Also, it would increase the reached fidelity when programming the device to a target matrix. Only the insertion losses over the alignment loops of the device were measured, while the insertion losses across the other fiber to chip interfaces still have to be measured.

As a final remark, we observe that, as the size of the chips increases, the fabrication precision will have to follow to maintain suitable performances. In fact, as observed from the simulations, the average performances of the device are intrinsically related to the fabrication precision of its directional couplers.

Bibliography

[1] Sufi Ahmed, Reza Baghdadi, Mikhail Bernadskiy, Nate Bowman, Ryan Braid, Jim Carr, Chen Chen, Pietro Ciccarella, Matthew Cole, John Cooke, Kishor Desai, Carlos Dorta, Jonathan Elmhurst, Bryce Gardiner, Elliot Greenwald, Shashank Gupta, Parry Husbands, Brian Jones, Anthony Kopa, and Nicholas Harris. Universal photonic artificial intelligence acceleration. *Nature*, 640:368–374, 04 2025.

[2] Koen Alexander, Andrea Bahgat, Avishai Benyamini, Dylan Black, Damien Bonneau, Stanley Burgos, Ben Burridge, Geoff Campbell, Gabriel Catalano, Alex Ceballos, Chia-Ming Chang, CJ Chung, Fariba Danesh, Tom Dauer, Michael Davis, Eric Dudley, Ping Er-Xuan, Josep Fargas, Alessandro Farsi, Colleen Fenrich, Jonathan Frazer, Masaya Fukami, Yogeeswaran Ganesan, Gary Gibson, Mercedes Gimeno-Segovia, Sebastian Goeldi, Patrick Goley, Ryan Haislmaier, Sami Halimi, Paul Hansen, Sam Hardy, Jason Horng, Matthew House, Hong Hu, Mehdi Jadidi, Henrik Johansson, Thomas Jones, Vimal Kamineni, Nicholas Kelez, Ravi Koustuban, George Kovall, Peter Krogen, Nikhil Kumar, Yong Liang, Nicholas LiCausi, Dan Llewellyn, Kimberly Lokovic, Michael Lovelady, Vitor Manfrinato, Ann Melnichuk, Mario Souza, Gabriel Mendoza, Brad Moores, Shaunak Mukherjee, Joseph Munns, Francois-Xavier Musalem, Faraz Najafi, Jeremy L. O'Brien, J. Elliott Ortmann, Sunil Pai, Bryan Park, Hsuan-Tung Peng, Nicholas Penthorn, Brennan Peterson, Matt Poush, Geoff J. Pryde, Tarun Ramprasad, Gareth Ray, Angelita Rodriguez, Brian Roxworthy, Terry Rudolph, Dylan J. Saunders, Pete Shadbolt, Deesha Shah, Hyungki Shin, Jake Smith, Ben Sohn, Young-Ik Sohn, Gyeongho Son, Chris Sparrow, Matteo Staffaroni, Camille Stavrakas, Vijay Sukumaran, Davide Tamborini, Mark G. Thompson, Khanh Tran, Mark Triplet, Maryann Tung, Alexey Vert, Mihai D. Vidrighin, Ilya Vorobeichik, Peter Weigel, Mathhew Wingert, Jamie Wooding, and Xinran Zhou. A manufacturable platform for photonic quantum computing, 2024.

[3] Christofer Alexiev, Jason C. C. Mak, Wesley D. Sacher, and JoyceE K. S. Poon. Calibrating rectangular interferometer meshes with external photodetectors. *OSA Continuum*, 4(11), 2021.

[4] Saumil Bandyopadhyay, Ryan Hamerly, and Dirk R. Englund. Hardware error correction for programmable photonics. *CoRR*, abs/2103.04993, 2021.

[5] William R. Clements, Peter C. Humphreys, Benjamin J. Metcalf, W. Steven Kolthammer, and Ian A. Walmsley. Optimal design for universal multiport interferometers. *Optica*, 3(12), 2016.

[6] Giuliano Coppola. Calibration of an integrated universal multiport interferometer. Master's thesis, Università di Pavia, 2024.

[7] Andreas Fyrillas, Olivier Faure, Nicolas Maring, Jean Senellart, and Nadia Belabas. Scalable machine learning-assisted clear-box characterization for optimally controlled photonic circuits. *Optica*, 11(3):427–436, Mar 2024.

[8] Andreas Fyrillas, Nicolas Heurtel, Simone Piacentini, Nicolas Maring, Jean Senellart, and Nadia Belabas. Resource-efficient crosstalk mitigation for the high-fidelity operation of photonic integrated circuits with induced phase shifters, 2025.

[9] T. Goh, S. Suzuki, and A. Sugita. Estimation of waveguide phase error in silica-based waveguides. *Journal of Lightwave Technology*, 15(11):2107–2113, 1997.

[10] Ryan Hamerly, Saumil Bandyopadhyay, and Dirk Englund. Accurate self-configuration of rectangular multiport interferometers. *Physical Review Applied*, 18(2), August 2022.

[11] Maxime Jacques, Alireza Samani, Eslam El-Fiky, David Patel, Zhenping Xing, and David V. Plant. Optimization of thermo-optic phase-shifter design and mitigation of thermal crosstalk on the soi platform. *Opt. Express*, 27(8):10456–10471, Apr 2019.

[12] Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization, 2017.

[13] E. Knill, R. Laflamme, and G. Milburn. Efficient linear optics quantum computation, 2000.

[14] Pieter Kok, W. J. Munro, Kae Nemoto, T. C. Ralph, Jonathan P. Dowling, and G. J. Milburn. Linear optical quantum computing. *Reviews of Modern Physics*, 79(1):135–174, January 2007.

[15] Nicolas Maring, Andreas Fyrillas, Mathias Pont, Edouard Ivanov, Petr Stepanov, Nico Margaria, William Hease, Anton Pishchagin, Thi Huong Au, Sébastien Boissier, Eric Bertasi, Aurélien Baert, Mario Valdivia, Marie Billard, Ozan Acar, Alexandre Brieussel, Rawad Mezher, Stephen C. Wein, Alexia Salavrakos, Patrick Sinnott, Dario A. Fioretto, Pierre-Emmanuel Emeriau, Nadia Belabas, Shane Mansfield, Pascale Senellart, Jean Senellart, and Niccolo Somaschi. A general-purpose single-photon-based quantum computing platform, 2023.

[16] David A. B. Miller. Supplementary document for analyzing and generating multimode optical fields using self-configuring networks - 4600959.pdf. 7 2020.

[17] Kaveh (Hassan) Rahbardar Mojaver, Bokun Zhao, Edward Leung, S. Mohammad Reza Safaee, and Odile Liboiron-Ladouceur. Addressing the programming challenges of practical interferometric mesh based optical processors. *Opt. Express*, 31(15):23851–23866, Jul 2023.

[18] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2000.

[19] Sunil Pai, Ben Bartlett, Olav Solgaard, and David A. B. Miller. Matrix optimization on universal unitary photonic devices. *Phys. Rev. Appl.*, 11:064044, Jun 2019.

[20] H. Rad, T. Ainsworth, R. Alexander, B. Altieri, M. Askarani, Reenu Baby, Leonardo Banchi, Ben Baragiola, J. Bourassa, R. Chadwick, I. Charania, H. Chen, M. Collins, P. Contu, N. D'Arcy, Guillaume Dauphinais, Robbe De Prins, D. Deschenes, I. Luch, and Y. Zhang. Scaling and networking a modular photonic quantum computer. *Nature*, 638:912–919, 01 2025.

[21] Saleh Rahimi-Keshari, Matthew A. Broome, Robert Fickler, Alessandro Fedrizzi, Timothy C. Ralph, and Andrew G. White. Direct characterization of linear-optical networks. *Optics Express*, 21(11):13450, May 2013.

[22] Michael Reck and Anton Zeilinger. Experimental realization of any discrete unitary operator. *physical review letters*, 73(1), 1994.

[23] Farhad Shokraneh, Simon Geoffroy-gagnon, and Odile Liboiron-Ladouceur. The diamond mesh, a phase-error- and loss-tolerant field-programmable mzi-based optical processor for optical neural networks. *Opt. Express*, 28(16):23495–23508, Aug 2020.

[24] Sergei Slussarenko and Geoff Pryde. Photonic quantum information processing: A concise review. *Applied Physics Reviews*, 6:041303, 12 2019.

[25] Minh A. Tran, Tin Komljenovic, Jared C. Hulme, Michael L. Davenport, and John E. Bowers. A robust method for characterization of optical waveguides and couplers. *IEEE PHOTONICS TECHNOLOGY LETTERS*, 28(14), 2016.

[26] Jianwei Wang, Fabio Sciarrino, Anthony Laing, and Mark G. Thompson. Integrated photonic quantum technologies. *Nature Photonics*, 14(5):273–284, October 2019.

[27] Quentin Wilmart, Houssein el Dirani, Nicola Tyler, Daivid Fowler, Stéphane Malhouitre, Stephanie García, Marco Casale, Sébastien Kerdiles, Karim Hassan, Christelle Monat, Xavier Letartre, Ayman Kamel, Kresten Yvind, L.K. Oxenlowe, Wilfried Rabaud, Corrado Sciancalepore, Bertrand Szelag, and Ségolène Olivier. A versatile silicon-silicon nitride photonics platform for enhanced functionalities and applications. *Applied Sciences*, 9:255, 01 2019.

[28] Haoran Zhang, Yuhang Song, Shifan Chen, Yunping Bai, Xingyuan Xu, Chaoran Huang, Jian Wang, Hongwei Chen, David Moss, and Kun Xu. Integrated platforms and techniques for photonic neural networks. *npj Nanophotonics*, 2, 10 2025.

[29] Hui Zhang, Jayne Thompson, Mile Gu, Xu Jiang, H. Cai, Patricia Liu, Yuzhi Shi, Yi Zhang, Muhammad Karim, Patrick Lo, Xianshu Luo, Bin Dong, Leong Kwek, and Ai-qun Liu. Efficient on-chip training of optical neural networks using genetic algorithm. *ACS Photonics*, 8, 04 2021.