



POLITECNICO DI MILANO
DEPARTMENT OF PHYSICS
DOCTORAL PROGRAMME IN PHYSICS

SCALABLE CHARACTERIZATION TECHNIQUES FOR INTEGRATED PHOTONIC DEVICES IN A FEMTOSECOND LASER WRITTEN PLATFORM

Doctoral Dissertation of:
Niki Di Giano

Supervisor:
Prof. Francesco Ceccarelli

Tutor:
Prof. Andrea Crespi

The Chair of the Doctoral Program:
Prof. Roberto Osellame

Year 2025/2026 – XXXVIII Cycle

Abstract

LINEAR optical networks constitute a core technology for both classical and quantum photonics, enabling applications that range from optical signal processing and analog optical computing to quantum interferometry, communication, and computation. Universal Photonic Processors (UPPs) are integrated optical devices that provide a scalable hardware platform for such applications by implementing arbitrary unitary transformations through meshes of reconfigurable 2-by-2 unit cells. «Calibration» in this sense means obtaining an accurate model representing the relationship between the control variables of the reconfigurable cells and the implemented unitary transformation.

Femtosecond Laser Waveguide Writing (FLWW) in glass is a fabrication platform that offers several advantages such as rapid prototyping, three-dimensional optical structure writing, and substrate machining at the micro-scale. This platform, together with photolithography-based methods to fabricate resistive heaters on top of the glass substrate, enables the creation of highly performant UPPs. However, the behavior of FLWW UPPs is affected by fabrication imperfections and, critically, by thermal cross-talk between the densely packed resistive heaters, which complicates modeling and calibration. All these factors contribute to making high fidelities more difficult to achieve.

This work presents a comprehensive experimental study of calibration and operation strategies for FLWW UPPs, with a focus on scalability, automation, and high-fidelity performance. Fully automated calibration pipelines are developed, reducing the total calibration time from multiple

days to a few hours, while enabling completely automated and scalable optimization algorithms. On a 6-mode processor, the proposed approach achieves an average amplitude fidelity of 0.9979 ± 0.0009 . Leveraging the speed of the automated calibration, a single-matrix optimization method based on current fine-tuning is demonstrated, increasing the mean amplitude fidelity for selected target matrices up to 0.9996 on average, highlighting fine-tuning as a viable strategy to push past the limits of the calibration in FLWW UPPs for a variety of applications.

The robustness of the calibrated devices is further investigated with respect to polarization effects. Despite waveguide birefringence that is intrinsic to the fabrication platform, the experiments show that the polarization state of the input light alters the average fidelity of the UPP only by a negligible amount even when confronting opposite polarization states, without the need to recalibrate the device.

Scalability is addressed through the extension of the calibration setup to larger meshes and through the introduction of machine-learning-based «digital twin» models trained on large matrix datasets. The proposed setups are capable of both gathering the data and applying the algorithms autonomously. Using these methods and setups, a 6-mode processor achieves an average amplitude fidelity of 0.9978 ± 0.0008 , an 8-mode processor achieves an average amplitude fidelity of 0.9948 ± 0.002 , while a 12-mode device achieves 0.9907 ± 0.0063 , with single matrix acquisition times below 10 seconds.

From the point of view of mitigating side-effects in UPPs, thermal cross-talk is investigated in detail through vacuum-enhanced device operation. By operating the UPP at medium vacuum conditions (3×10^{-2} mbar), heat transfer is reduced, leading to a marked increase in thermal isolation between adjacent heaters. Quantitative cross-talk metrics show an improvement of an order of magnitude when transitioning from air to vacuum. While vacuum conditions modify the thermal dynamics of the system, introducing longer thermalization times for individual heaters, simulations show that the effective global reconfiguration time (Time-To-Matrix, TTM) remains comparable. These results highlight vacuum operation as a viable pathway to suppressing thermal cross-talk in large-scale FLWW UPPs without fundamentally compromising reconfiguration speed.

Beyond amplitude-only characterization, an interferometric phase retrieval technique based on a dedicated interface chip is introduced, enabling access to physically meaningful phase information and full fidelity evaluation. This approach allows the measurement of complex matrix elements, constituting one of the first demonstrations of full (amplitude and phase)

matrix and fidelity measurements in a FLWW UPP with a Clements-type architecture. Using the proposed calibration methods and measurement technique for the UPP, fidelities between 0.9935 and 0.9993 are measured for selected matrices, fully accounting for the complex nature of the unitaries.

Finally, a complementary investigation of a novel methodology for the characterization of multimode waveguides is presented. By acquiring large datasets of output intensity patterns under varying excitation conditions and applying Principal Component Analysis (PCA), the number of modes supported by the structure is inferred from the dominant principal components. Experimental results on two different fibers (1550BHP and 780HP) at five different wavelengths ranging from 405 nm to 925 nm are carried out, demonstrating qualitative discrimination of multimodal behavior which is consistent with theoretical expectations. This approach establishes a low-cost, data-driven framework for multimode waveguide characterization and provides a foundation for future extensions toward quantitative mode analysis.

Overall, this work demonstrates a scalable and experimentally validated toolbox for the calibration and high-fidelity operation of FLWW integrated devices, providing a solid foundation for their deployment in advanced classical and quantum photonic applications.

Contents

1	Introduction and Theoretical Background	1
1.1	Introduction	1
1.2	Motivation	1
1.3	Integrated Photonics Theory	3
1.3.1	Integrated Waveguide	3
1.3.2	Integrated Coupled Waveguide	5
1.3.3	Directional Coupler	9
1.4	Programmable Photonics	11
1.4.1	Integrated Mach-Zehnder Interferometers	11
1.4.2	Universal Photonic Processors	13
1.4.3	Performance Evaluation of UPPs	15
2	Materials and Methods	17
2.1	Introduction	17
2.2	UPP Fabrication	17
2.2.1	Femtosecond Laser Waveguide Writing	17
2.2.2	Post-FLWW Processing of Integrated Photonic Devices	18
2.2.3	General Structure of UPP Devices	19
2.3	UPP Calibration	21
2.3.1	Temperature and Resistance Stability	23
2.4	Operation of UPPs	24
2.4.1	Matrix Decomposition	24
2.4.2	Matrix Amplitude Measurement	27
2.5	Experimental Instruments Used in the Work	28

2.5.1	Electronics and Current Control	28
2.5.2	Optical and Motion Control	30
2.5.3	Temperature Control	31
3	Universal Photonic Processors	33
3.1	Introduction	33
3.2	UPP Calibration	33
3.2.1	Motivation and State of the Art	33
3.2.2	Calibration Setups	38
3.2.3	Single-Matrix Optimization	41
3.2.4	Polarization Independence	44
3.2.5	Full Machine Learning Calibration	47
3.2.6	6-mode Automated Calibration	51
3.2.7	8-mode Automated Calibration	53
3.2.8	12-mode Automated Calibration	54
3.3	Vacuum-Enhanced Calibration of UPPs	58
3.3.1	Motivation	58
3.3.2	Method and Setup	59
3.4	Measurement of Complex Matrix Elements	67
3.4.1	Motivation	67
3.4.2	UPPs and Real Bordered Matrices	68
3.4.3	Method and Setup	69
3.4.4	Experimental Validation	75
4	Multimode Waveguide Characterization	81
4.1	Introduction	81
4.2	Motivation	81
4.3	State of the Art	82
4.3.1	Cylindrical Waveguides Theory	82
4.3.2	Principal Component Analysis	84
4.4	Multiple Mode Excitations	86
4.5	Method and Setup	87
4.6	Data Processing and Results	90
4.6.1	Reconstruction of Rotating Phase Profiles	93
5	Conclusions and Future Developments	99
	Bibliography	117

CHAPTER 1

Introduction and Theoretical Background

1.1 Introduction

In this chapter, I give a brief introduction to build up some crucial aspects of the theory of integrated optics that underpin the rest of the work. In Section 1.2 I describe some of the main goals of integrated optics to contextualize the work; in Sections 1.3 and 1.4 I provide the theory for integrated optics with a focus on programmable integrated optical components. In Section 1.4.2 I introduce some additional theory for specific integrated optical devices that will be important throughout the work.

1.2 Motivation

Integrated optics is a branch of science and engineering that focuses on the design and fabrication of miniaturized optical components, with the goal of integrating these components into a single compact device. In this regard, it is driven by similar interests compared to microelectronics for electrical components. More in particular, *integrated quantum photonics* is the field of integrated optics centered around applications involving the quantum nature of light.

The field has seen substantial growth since seminal work at Bell Labs from 1969 [1]; it offers clear benefits in well-established fields of optics such as signal processing, spectroscopy, and microscopy, but also in more recent fields such as optical computing. In fact, integrated photonics focuses on devices with dimensions down to the micro- or even nano-scale. At these scales the optical, mechanical and thermal properties of integrated photonic devices - as well as their instabilities - are easier to control compared to their bulk optics counterparts.

Several fabrication platforms are available for integrated photonics. *Lithographic* technologies leverage the well-established fabrication methods of microelectronic components, such as CMOS, to construct integrated photonic devices layer by layer; on the other hand, *direct writing* technologies exploit the local modification of a material to inscribe components right into the substrate.

Understandably, integrated photonics comes with its own set of challenges from the engineering point of view. Miniaturization of optical systems must be done with very precise tools, so that the repeatability and scalability play a crucial role in the engineering of compact integrated photonic devices. On top of that, the goal of integrating several components with different roles in the same device inevitably leads to the necessity to connect together objects created through several different technologies or fabrication platforms, which can be a huge challenge.

For this reason, it becomes increasingly important to properly characterize the behavior of each component individually before they are combined together; for more complex integrated devices, in fact, it becomes increasingly difficult to obtain relevant information about each component after assembly. Among the aims of this work is to also provide methods to untangle the effects of these components, so that even after the device is completely assembled it is still possible to retrieve all the important parameters. In a similar vein, once a proper characterization has been obtained, it becomes necessary to quantify the quality of this characterization.

The main aim of this work is to introduce novel methods to characterize the behavior of both integrated photonic devices fabricated via direct writing, as well as to show how these methods are applied in practice for both UPPs (in Chapter 3) and more general multimode structures (in Chapter 4). Before providing further details about the work, I will introduce some of the most important theoretical concepts underpinning the methods.

1.3 Integrated Photonics Theory

1.3.1 Integrated Waveguide

The simplest building block of integrated photonics is the *integrated optical waveguide* [2]. By the phenomenon of total internal reflection, the optical waveguide's working principle is to confine the electromagnetic field within the internal area - the *core*. The field remains mostly confined within the core and only evanescent tails leak out to the outside area - the *cladding*. An example of integrated optical waveguide is represented in Figure 1.1. In this device, the core is characterized by a refractive index $n_2 > n_1$. Maxwell's equations read:

$$\begin{cases} \nabla \cdot \vec{D} = \rho, & \nabla \times \vec{E} + \frac{\partial \vec{B}}{\partial t} = \vec{0}, \\ \nabla \cdot \vec{B} = 0, & \nabla \times \vec{H} - \frac{\partial \vec{D}}{\partial t} = \vec{J}. \end{cases} \quad (1.1)$$

Here $\vec{H}$ and $\vec{D}$ are respectively the magnetic field and electric displacement field, which in linear (non-magnetic) media relate to the magnetic induction field and electric field via the dielectric permittivity of the material ε and the magnetic permeability of vacuum μ_0 as

$$\vec{H} = \frac{\vec{B}}{\mu_0}, \quad \vec{D} = \varepsilon \vec{E}. \quad (1.2)$$

From these equations we can obtain

$$\nabla^2 \vec{E} - \varepsilon \mu_0 \frac{\partial^2 \vec{E}}{\partial t^2} = -\nabla \left(\vec{E} \cdot \frac{\nabla \varepsilon}{\varepsilon} \right), \quad (1.3)$$

both in the cladding and the core.

The type of waveguides that will be relevant for the rest of this work satisfy the *weak guiding condition* $\Delta n \ll n_1$: in the devices that will be analyzed in this work, $\Delta n = n_2 - n_1$ is in the order of 10^{-3} - 10^{-4} . This implies that $\nabla \varepsilon \ll \varepsilon$ and therefore we are left with

$$\nabla^2 \vec{E} - \varepsilon \mu_0 \frac{\partial^2 \vec{E}}{\partial t^2} = 0 \quad (1.4)$$

which is the usual wave equation.

Note also that the electric permittivity is directly related to the refractive index of the material via $\varepsilon = n^2$ by definition, and therefore we have two

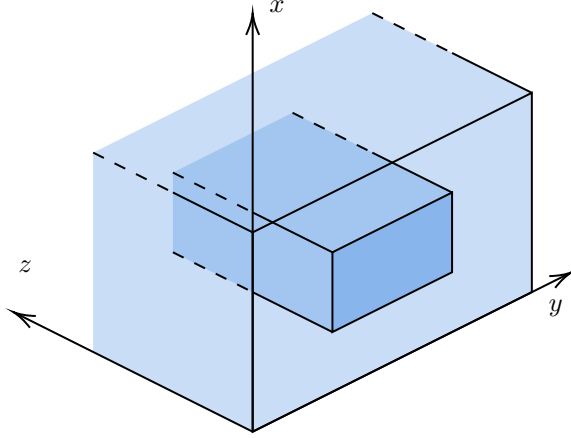


Figure 1.1: A pictorial representation of an integrated waveguide. The darker area is the core (refractive index n_2), surrounded by the cladding (refractive index n_1). To achieve the guiding action, $n_2 > n_1$.

equations

$$\begin{aligned} \nabla^2 \vec{E} - \varepsilon_1 \mu_0 \frac{\partial^2 \vec{E}}{\partial t^2} &= -\nabla \left(\vec{E} \cdot \frac{\nabla \varepsilon_1}{\varepsilon_1} \right) \quad \text{in the cladding,} \\ \nabla^2 \vec{E} - \varepsilon_2 \mu_0 \frac{\partial^2 \vec{E}}{\partial t^2} &= -\nabla \left(\vec{E} \cdot \frac{\nabla \varepsilon_2}{\varepsilon_2} \right) \quad \text{in the core,} \end{aligned} \quad (1.5)$$

that need to be satisfied by the field at the same time. Additional requirements imposed by Equation 1.1 specify the behavior of the field at the boundary, but a full derivation of the solutions is outside of the scope of this work.

A similar derivation can be obtained for the magnetic field, and the configurations of fields $(\vec{E}, \vec{H})$ that satisfy all equations at the same time are the so-called «modes» of the structure.

In general, these modes are indexed by one or more integer numbers, with an additional label P that distinguishes between two possible polarization states of the mode. In a rectangular structure such as the one shown in Figure 1.1, the modes are characterized by the polarization state P and a pair of integer numbers m, n . To each solution is associated a propagation constant β_{mn}^P , and their analytical form is

$$\vec{E}_{mn}^P(x, y, z, t) = \vec{e}_P A_{mn}^P(x, y) e^{i(\omega t - \beta_{mn}^P z)} \quad (1.6)$$

assuming the wave is propagating along the positive z axis. The vector $\vec{e}_P$ determines the polarization state of the wave, while the field ampli-

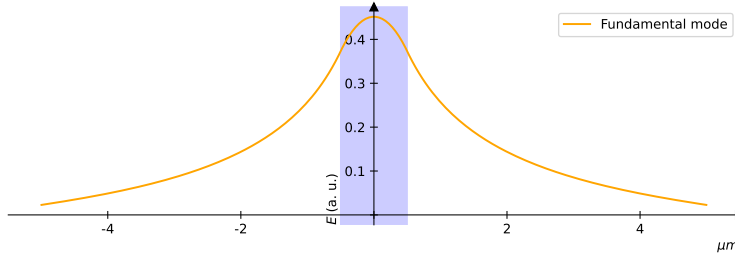


Figure 1.2: The plot (arbitrary units) of the electric field profile of the fundamental mode for the analyzed structure, along the x axis at fixed $y = 0$. The core region is shaded in blue. This graph was obtained by numerically solving Equation 1.4 for a core of width $1\text{ }\mu\text{m}$, $\Delta n = 10^{-2}$ and $\lambda = 633\text{ nm}$.

tude profile $A_{mn}^P(x, y)$ determines the shape of the field along the plane orthogonal to the propagation axis. Figure 1.2 shows the fundamental mode ($m = n = 0$, any P) for the simple structure we have considered so far.

The number of solutions for this structure depends on ω . For $\omega = 0$ there are no propagating wave solutions of the form given in Equation 1.6, but as ω increases the number of solutions increases as well. There is a range of values for ω such that the only available solution is the fundamental mode ($m = n = 0$) for each polarization label P : in this range, the waveguide is called «singlemode». Otherwise, if solutions with larger m and n exist, the waveguide is called «multimode».

1.3.2 Integrated Coupled Waveguide

Figure 1.2 shows the profile of the fundamental mode of a waveguide. The electric field leaks outside of the boundaries of the core through exponentially decaying tails. Therefore, if we were to put a second waveguide next to the original one as in Figure 1.3, the field would leak into the core of the second waveguide, leading to coupling between the two structures. A more rigorous analysis would require us to go back to Equations 1.1 and obtain solutions for the overall structure; with this approach, it's possible to show that for two coupled singlemode waveguides there are always at least two fundamental solutions to the differential equations: we call these solutions $\vec{S}$ and $\vec{A}$ (Figure 1.4). It can be shown that if the waveguides are sufficiently far apart from one another [2] they can be well approximated by symmetric and antisymmetric linear combinations of the fundamental modes for each

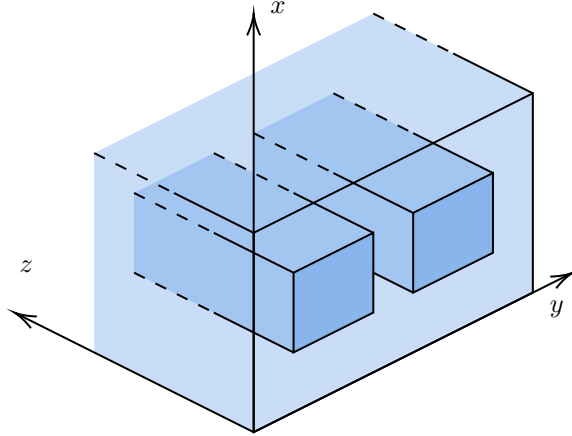


Figure 1.3: A pictorial representation of integrated coupled waveguides. The two cores have the same refractive index.

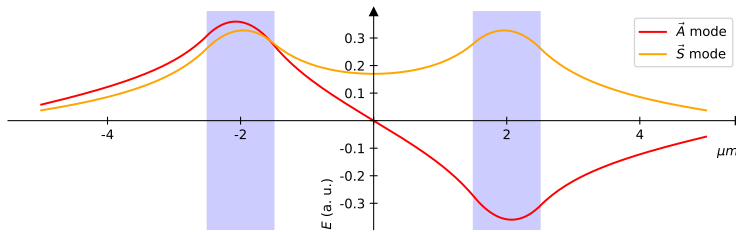


Figure 1.4: The electric field profiles of the $\vec{S}$ and $\vec{A}$ modes for the coupled waveguides. The parameters used to obtain these profiles are the same as Figure 1.2, with both waveguides having a width of $1 \mu\text{m}$. The modes resemble a symmetric and antisymmetric sum of the fundamental mode of each waveguide.

structure taken as an individual waveguide. In short:

$$\vec{S} = \vec{\epsilon}_P \frac{\vec{A}_{00,1}^P + \vec{A}_{00,2}^P}{\sqrt{2}} e^{i(\omega t - \beta_S^P z)}, \quad (1.7)$$

$$\vec{A} = \vec{\epsilon}_P \frac{\vec{A}_{00,1}^P - \vec{A}_{00,2}^P}{\sqrt{2}} e^{i(\omega t - \beta_A^P z)}. \quad (1.8)$$

If both waveguides are singlemode, then these fields are a complete basis for all the fields supported by the structure.

Let us consider a field that, at fixed time t and position z , is identical to the fundamental mode of the first waveguide. By definition of $\vec{S}$ and $\vec{A}$ we can write the field as

$$\vec{E}_1(z, t) = \frac{\vec{S}(z, t) + \vec{A}(z, t)}{\sqrt{2}}. \quad (1.9)$$

The intensity of a field $\vec{E}$ that oscillates in time is defined as

$$I = \frac{1}{2} cn \left\langle \left\| \vec{E} \right\|^2 \right\rangle \quad (1.10)$$

where the angle brackets denote average over time (over the period of the oscillation). This quantity has units of W m^{-2} , and it represents the surface density of the power carried by the field across a section of the structure. As Figure 1.5a shows, the field $E_1(z, t)$ has most of its intensity concentrated in the first waveguide. If we consider how the field propagates for a generic length Δz , we obtain

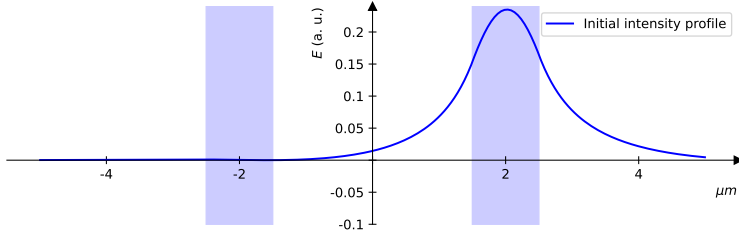
$$\vec{E}_1(z + \Delta z, t) = \frac{\vec{S}(z, t) e^{-i\beta_S^P \Delta z} + \vec{A}(z, t) e^{-i\beta_A^P \Delta z}}{\sqrt{2}} \quad (1.11)$$

$$= \frac{e^{-i\beta_S^P \Delta z}}{\sqrt{2}} \left(\vec{S}(z, t) + \vec{A}(z, t) e^{-i\Delta\beta \Delta z} \right) \quad (1.12)$$

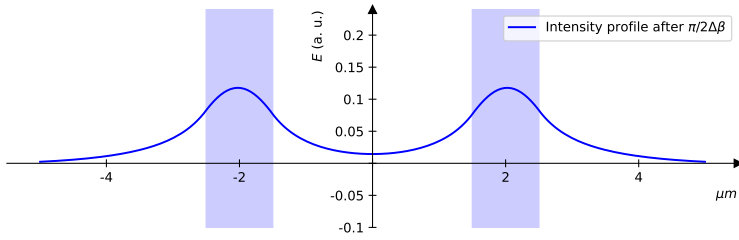
where $\Delta\beta = \beta_A - \beta_S$. At $\Delta z = \pi/\Delta\beta$, the field becomes

$$\vec{E}_1(x, y, z + \pi/\Delta\beta, t) = e^{-i\frac{\beta_S^P \pi}{\Delta\beta}} \frac{\vec{S}(z, t) - \vec{A}(z, t)}{\sqrt{2}}. \quad (1.13)$$

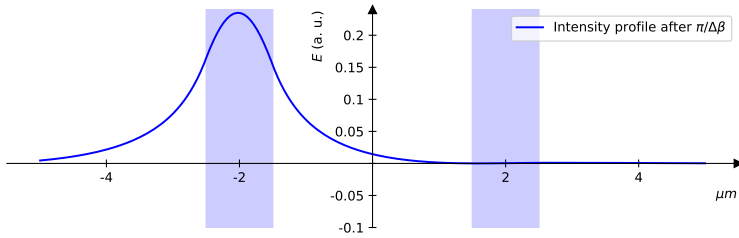
The intensity profile of this field is shown in Figure 1.5. The intensity in this case is entirely concentrated within the second waveguide, and the mode resembles the fundamental mode of the second waveguide taken alone.



(a)



(b)



(c)

Figure 1.5: The intensity profile of the $\vec{E}_1(z + \pi/2\Delta\beta, t)$ field, at fixed time. After a propagation of $\pi/2\Delta\beta$, the intensity is split evenly between the waveguides.

In other words, there is a propagation length $\pi/\Delta\beta$ such that the field seems to «skip» from one waveguide to the other as it propagates in space. This cycle repeats indefinitely if the waveguides extend to infinity. If we were to look at the propagation after a half of this length, $\Delta z = \pi/2\Delta\beta$, we would obtain

$$\vec{E}_1(z + \pi/2\Delta\beta, t) = e^{-i\frac{\beta_S^P \pi}{2\Delta\beta}} \frac{\vec{S}(z, t) + i\vec{A}(z, t)}{\sqrt{2}} \quad (1.14)$$

and the intensity profile of this field is shown in Figure ?? . As can be seen from the Figure, the field intensity in this case splits evenly between the two waveguides; this implies that the optical power carried by the field is split evenly as well.

More generally any power split in between can be achieved with a pair of coupled waveguides, which makes them fundamentally important as building blocks of integrated optics.

1.3.3 Directional Coupler

The effect described in the previous section can be exploited to obtain an integrated *directional coupler*: a device capable of splitting the electric field between input and output. While in the following work most of the directional couplers involved in the integrated photonic devices are designed to split the field evenly - a configuration that will be called 3 dB *coupler* from now on - a directional coupler can be created to provide any power split in principle. The standard design of a directional coupler is schematically represented in Figure 1.6 and it involves three major sections: an input stage where the two integrated waveguides are far apart enough to have negligible coupling between them; an interaction stage where the waveguides are allowed to exchange energy for a certain distance; an output stage where the waveguides drift apart once again. The interaction distance determines the splitting of the power at the output stage of the device. Between these sections the waveguides must curve in order to come together or apart; this generally influences the fields by introducing losses and changing the polarization state of the light.

The two modes at the input of the device are each described by Equation 1.6, and for the purposes of analysing the effect of the device, the state of the light may be described as a 2-dimensional complex vector $[w_0, w_1]$ representing their amplitudes and phases. Then since the directional coupler mixes the modes linearly, its effect can be described simply as a matrix

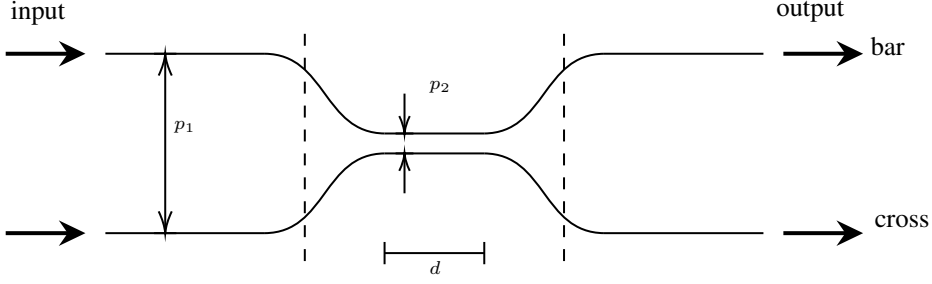


Figure 1.6: Schematic of a directional coupler. The black lines represent integrated waveguides. p_1 and p_2 are the waveguide distances in the non-interacting and interacting region respectively. The geometric interaction distance is d .

multiplication:

$$\begin{bmatrix} w'_0 \\ w'_1 \end{bmatrix} = \begin{bmatrix} d_{00} & d_{01} \\ d_{10} & d_{11} \end{bmatrix} \begin{bmatrix} w_0 \\ w_1 \end{bmatrix} \quad (1.15)$$

where, if we neglect losses, the matrix D must preserve the total power, $w_0^2 + w_1^2 = w_0'^2 + w_1'^2$; therefore, the matrix must be unitary. Every 2-by-2 unitary matrix, up to external phase terms, can be described by a single parameter η as

$$T = \begin{bmatrix} \cos(\eta) & i \sin(\eta) \\ i \sin(\eta) & \cos(\eta) \end{bmatrix}. \quad (1.16)$$

If we were to inject light only at one input port - say, the first, so that the input state of light would be $[A_0, 0]$ - then the resulting vector would be

$$\begin{bmatrix} w'_0 \\ w'_1 \end{bmatrix} = \begin{bmatrix} A_0 \cos(\eta) \\ i A_0 \sin(\eta) \end{bmatrix} \quad (1.17)$$

and to obtain the optical power at the first or second output it is sufficient to take the square of the corresponding element in the light state vector:

$$P_0 = A_0^2 \cos^2(\eta), \quad P_1 = A_0^2 \sin^2(\eta). \quad (1.18)$$

In general the η parameter is related to the optical and geometrical parameters of the device. For a directional coupler as pictured in Figure 1.6, it can be shown that

$$\eta = \Delta\beta L/2, \quad (1.19)$$

where $\Delta\beta$ is the differential propagation constant and L is the interaction length of the coupler. Setting $\eta = \pi/4$ gives the transfer function of a 3 dB coupler:

$$T_{3\text{dB}} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix}. \quad (1.20)$$

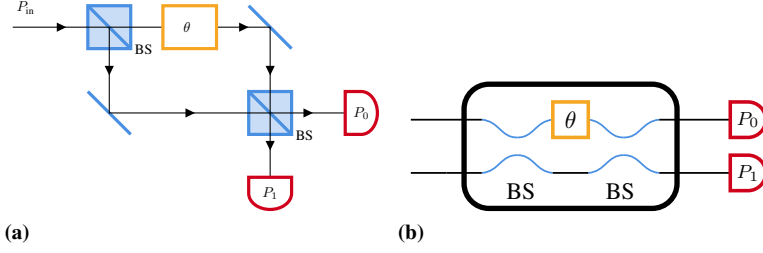


Figure 1.7: Schematic representations of Mach-Zehnder interferometers (a) in bulk optics and (b) in integrated optics.

With this setting, the output powers for a 3 dB coupler turn out to be equal to $A_0^2/2$, which is exactly half of the input power.

1.4 Programmable Photonics

1.4.1 Integrated Mach-Zehnder Interferometers

While a directional coupler is useful in its own right, being able to tune or «program» the η parameter would allow one to choose an arbitrary power split, so that multiple functionalities can be fulfilled by a single device. As mentioned, the η parameter is connected to the optical and geometrical properties of the device. For several fabrication platforms, changing either of those properties in a controlled manner can be impractical. On the other hand, in bulk optics one can find *beam splitters*, who similarly to 3 dB couplers are able to split the incoming light evenly from input to output, and allow interferometry between light beams if they impinge from both inputs at once. In bulk optics, a Mach-Zehnder Interferometer (MZI) is a system comprising two beam splitters and a delay line (often implemented through moving mirror stages, see Figure 1.7a); it can be shown that the power split at the two outputs of such systems is given by

$$P_0 = A_0^2 \sin^2 \left(\frac{\Delta\theta}{2} \right), \quad P_1 = A_0^2 \cos^2 \left(\frac{\Delta\theta}{2} \right). \quad (1.21)$$

where $\Delta\theta$ is the relative phase accumulated by one beam with respect to the other. Since 3 dB couplers function the same as beam splitters for integrated photonics, it is possible to create an integrated MZI only using integrated photonics building blocks (Figure 1.7). This new device consists of a 3 dB coupler, followed by a stage where one waveguide imparts an additional phase shift $\Delta\theta$ to the light, followed by a second 3 dB coupler. The final

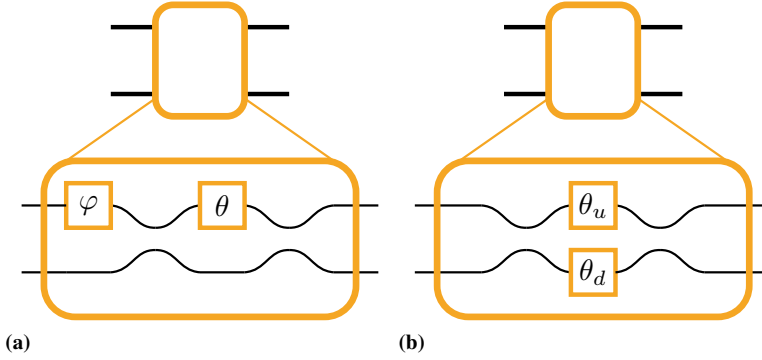


Figure 1.8: Schematic representations of programmable MZIs in the (a) external-internal and (b) push-pull configurations.

transfer function of this device is therefore

$$T_{\text{MZI}} = \sqrt{\frac{1}{2}} \begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix} \cdot \begin{bmatrix} e^{i\Delta\theta} & 0 \\ 0 & 1 \end{bmatrix} \cdot \sqrt{\frac{1}{2}} \begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix} \quad (1.22)$$

which may be simplified down to

$$T_{\text{MZI}} = ie^{i\frac{\Delta\theta}{2}} \begin{bmatrix} \sin\left(\frac{\Delta\theta}{2}\right) & \cos\left(\frac{\Delta\theta}{2}\right) \\ \cos\left(\frac{\Delta\theta}{2}\right) & -\sin\left(\frac{\Delta\theta}{2}\right) \end{bmatrix}. \quad (1.23)$$

which gives precisely the same power split as in Equation 1.21.

We may actually go one step further, adding another phase shifting element to the MZI either in an external-internal configuration (Figure 1.8a) or push-pull configuration (Figure 1.8b). Including the additional effect of this phase shift in the first case gives

$$T_{2,c} = ie^{i\frac{\theta}{2}} \begin{bmatrix} e^{i\varphi} \sin\left(\frac{\theta}{2}\right) & \cos\left(\frac{\theta}{2}\right) \\ e^{i\varphi} \cos\left(\frac{\theta}{2}\right) & -\sin\left(\frac{\theta}{2}\right) \end{bmatrix} \quad (1.24)$$

where θ is the internal phase shift and φ the external phase shift; the second case gives

$$T_{2,b} = ie^{i\frac{\theta_u + \theta_d}{2}} \begin{bmatrix} \sin\left(\frac{\theta_u - \theta_d}{2}\right) & \cos\left(\frac{\theta_u - \theta_d}{2}\right) \\ \cos\left(\frac{\theta_u - \theta_d}{2}\right) & -\sin\left(\frac{\theta_u - \theta_d}{2}\right) \end{bmatrix} \quad (1.25)$$

where θ_u is the upper arm phase shift and θ_d is the lower arm phase shift.

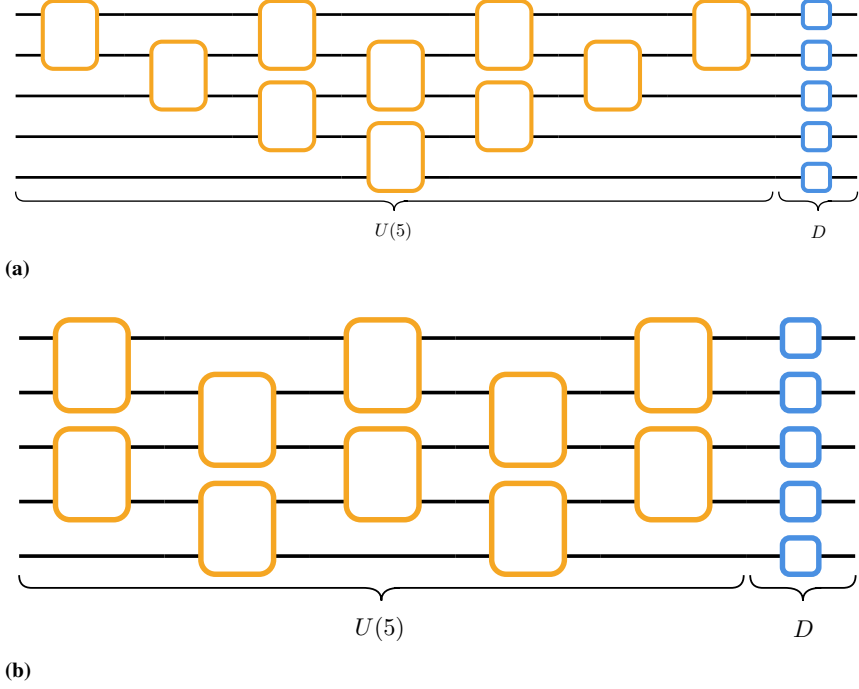


Figure 1.9: Representations of the Reck (a) and Clements (b) structures. The unit cell is represented in Figure 1.8a. The blue squares represent the extra phase terms, and may be physically controlled by implementing singular phase shifters.

1.4.2 Universal Photonic Processors

While the ability to reconfigure the power split and relative phase shift between just two modes finds several applications in photonics and optics [3–5], this work is entirely focused on devices designed for those applications that involve the mixing of more than two modes via unitary matrix multiplication. Such cases would require a device capable of implementing a unitary matrix $U \in U(N)$, where N is some fixed number of modes and $U(N)$ is the space of complex unitary matrices; notice that $D_{2,C}$ in Equation 1.24 and $D_{2,B}$ in Equation 1.24 are both elements of $U(2)$.

It was shown initially by Reck *et al.* [6] that such a device can be realized using a network of programmable MZIs. The architecture proposed by Reck includes MZIs in an external-internal configuration as the unit cell (Figure 1.8a), each implementing a transformation on a pair of modes as in Equation 1.24, arranged in a triangular layout (Figure 1.9a). Reck showed that this structure can implement any unitary matrix up to a set of N global phases, that is, if U_t is the target unitary matrix, then the system is capable

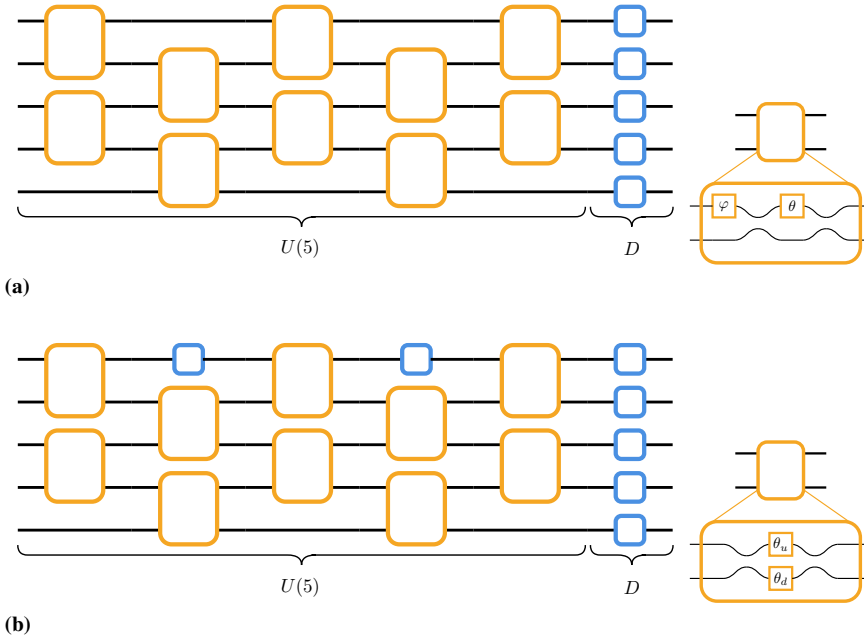


Figure 1.10: Representations of the Clements (a) and Bell (a) architectures, and their respective unit cells. The blue squares represent some extra phase terms, and may be physically controlled by implementing singular phase shifters.

of implementing DU_t where D is some diagonal matrix consisting of only phase elements. We may call this property «universality», and the photonic devices that possess this property are called *Universal Photonic Processors* (UPPs).

Later on, Clements *et al.* improved on the Reck design by introducing a square layout (Figure 1.9b), which they proved to be universal [7]. This architecture has the advantage that each optical path travels through at most N interferometers, which means that not only are the maximum losses (from bending and propagation) reduced compared to the Reck structure but also the lengths of the paths the light can take are more uniform, as is the number of MZIs placed along these paths. Bell *et al.* introduced a similar architecture (Figure 1.10b) that made use of the push-pull MZI design as unit cell, together with only a few additional external phase shifters, and proved that the structure is equivalent to the one proposed by Clements [8] but can be made even smaller, since the external phase shifters are removed.

1.4.3 Performance Evaluation of UPPs

In order to quantify the quality of the implementation and the measurement setup, the similarity between the target unitary U_t and the measured U_m needs to be evaluated. Measurements setups are a central discussion of this work, and further details about how to measure unitary matrices from UPPs are given in Chapter 2.

Once the matrix U_m has been obtained, a figure of merit that can represent the similarity between two unitary matrices is the *fidelity*

$$\mathcal{F}(U_t, U_m) = \frac{1}{N} \text{Tr} \left\{ \left| U_t^\dagger U_m \right| \right\} \quad (1.26)$$

and where a complete characterization of the complex phases of the unitary matrix isn't possible, so that only the amplitudes can be measured, it is still possible to obtain a figure of merit by taking the *amplitude fidelity*

$$\mathcal{F}_{\text{ampl}}(U_t, U_m) = \frac{1}{N} \text{Tr} \left\{ |U_t^\dagger| |U_m| \right\}. \quad (1.27)$$

The two quantities range between 0 and 1, where 1 is achieved if and only if the two matrices are identical.

Next, in order to be able to compare the performance between similar devices, it is necessary to pick a standardized set of matrices where either of the two metrics can be evaluated. In this context, two main types of matrices are typically explored.

Switching matrices are a subset of the space of unitary matrices where every element of the set can be obtained via a permutation of the columns (or rows) of the identity matrix. These matrices are typically taken at random in order to evaluate the performance of a UPP over tasks such as light routing and port switching [9]. These matrices may be picked at random over the $N!$ possible permutations to provide an average value of the performance metric over these tasks.

On the other hand, there is a unique uniform function called the *Haar measure* that allows one to pick unitary matrices uniformly at random [10]. Taking the average value of the performance metric over a large collection of matrices picked in this way therefore provides a standard figure of merit for the performance of the implementation for any task. In Chapter 3 this figure of metric will be the main metric to ultimately compare this work with other similar results found in the literature in the context of UPPs.

CHAPTER 2

Materials and Methods

2.1 Introduction

In this chapter, I introduce the main methods with which UPPs and other integrated devices studied throughout the work are fabricated, controlled, and calibrated. In Section 2.2 I introduce Femtosecond Laser Waveguide Writing and Femtosecond Laser Micromachining, which are the main fabrication platforms for the UPPs studied in this work. In Section 2.3 I introduce the main theoretical background and assumptions behind the process of calibration as well as a generic overview of known calibration methods. In Section 2.4 I introduce the primary method to obtain the measurement of a unitary matrix starting from the target matrix. Finally, in Section 2.5 I list and describe the devices used to perform the experiments and measurements described in the main body of the work.

2.2 UPP Fabrication

2.2.1 Femtosecond Laser Waveguide Writing

By «Femtosecond Laser Waveguide Writing» (FLWW) we mean the set of techniques for the fabrication of waveguides in transparent media through

the usage of femtosecond laser pulses. Similarly, by «Femtosecond Laser Micromachining» (FLM) we mean the set of techniques for the machining of materials at the micro-scale through the usage of femtosecond laser pulses.

The principle at the base of both platforms is the same: a high intensity (10 TW cm^{-2}) femtosecond pulse induces multiphoton absorption in the material, which creates a high-energy electron plasma. The energy is transferred to the material in a few picoseconds, permanently modifying its properties.

Two regimes can be identified in this method. A low-repetition-rate regime ($\lesssim 1 \text{ MHz}$) where transparent media see a local modification of their refractive index in the order of $10^{-2} - 10^{-3}$, which is enough to fabricate waveguide structures of the kind described in Chapter 1.3. At a high-repetition-rate regime instead, the nonlinear absorption induces an avalanche breakdown of the material, enabling the technology to perform machining at the micro-scale.

The main advantages of FLWW and FLM include the speed of the fabrication processes to produce optical device prototypes, which is much higher compared to planar platforms such as Silicon on Insulator or Silica on Silicon; on top of that, both techniques can leverage the third dimension to produce more complex structures such as stacked waveguide layers [11], trenches and waveguide bridges [12].

2.2.2 Post-FLWW Processing of Integrated Photonic Devices

After FLWW and FLM, the irradiated material has accumulated a lot of internal stress, which increases the birefringence of the glass. This effect can be quenched through annealing to obtain a birefringence as low as 10^{-6} [13], so that the final glass structure is ready for the next steps of the fabrication process.

After the waveguide writing and micromachining via femtosecond laser pulses, the next step is to fabricate the active elements that induce phase shifting on the guided light. While several solutions can be found in the literature, such as plasma dispersion in silicon based fabrication platforms [14], by far the method to apply phase shifting in UPP that is most compatible with FLWW is thermal phase shifting.

To achieve thermal phase shifting the main solution consists in creating a set of resistive elements that can heat the glass via Joule effect. These resistive elements are fabricated by sputtering gold [15] or using photolithographic methods to pattern chromium-copper layers [16] on the glass. The

metal film may also be patterned into individual resistive elements using FLM, if not directly through the photolithographic process.

Finally, in particular for chromium-copper based designs, a passivation layer is added on top of the device in order to prevent oxidation [16].

2.2.3 General Structure of UPP Devices

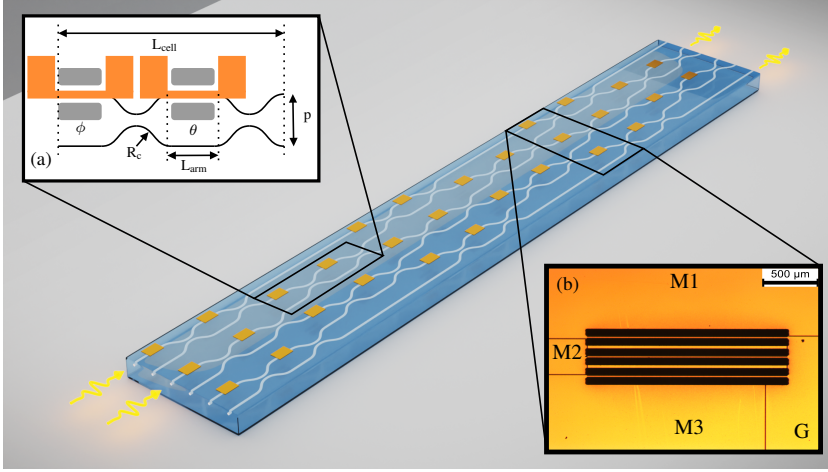


Figure 2.1: A render of the typical structure of a UPP. In this picture, the structure is the one of a 6-mode Clements layout. Inset (a) shows a scheme of the unit cell with the active elements; inset (b) shows a microscope picture comprising 3 heating elements in a column (M1, M2, M3 and their common ground G); the grey rectangular areas are trench structures dug into the glass, leaving thin slivers where the waveguides have been written.

An example structure of a typical UPP device fabricated through FLWW and FLM can be seen in Figure 2.1. The waveguide structure of the UPP is written in a borosilicate glass substrate (typically Corning EagleXG and Eagle2000 or Schott AF32), and a set of insulation structures in the form of deep trenches or bridges is machined around the waveguides (see Figure 2.2). As will be further discussed in the following chapter, these structures are important to limit the unwanted side effects of thermo-optical phase shifting.

After the writing and machining processes are completed, the glass substrate undergoes annealing. This process quenches the stress accumulated in the substrate from the FLWW and FLM processes, relaxing the structure and reducing birefringence in the glass. After this stage, gold or chromium-copper resistive heaters are fabricated on top of the glass structure.

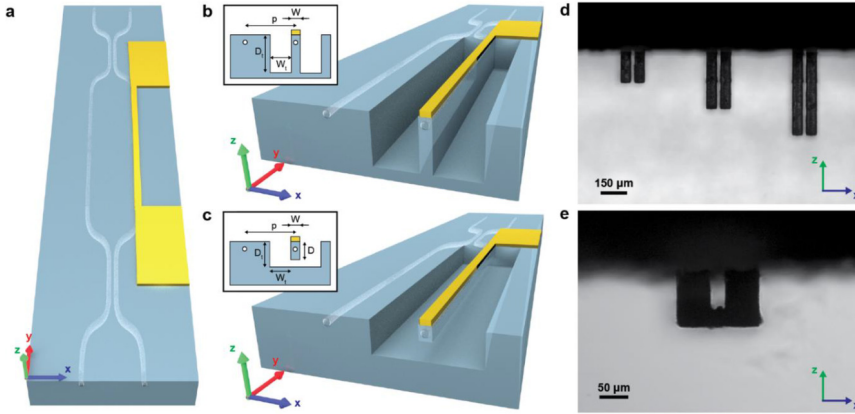


Figure 2.2: Renders of the typical structure of an MZI written with FLWW and FLM. Figure (a) shows the structure of a Clements unit cell from the top; Figures (b) and (d) show a render and a microscope image of a section of trench structure fabricated with FLWW. Figures (c) and (e) show the same images for a bridge structure instead. Figures adapted from Ref. [17].

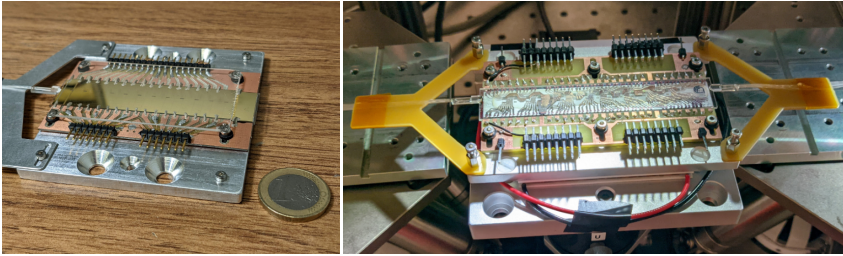


Figure 2.3: Two assembled UPP devices. On the left, a 6-mode UPP with gold heating elements and its size compared to a 1 Euro coin; on the right, a fully assembled 8-mode UPP with chromium-copper heating elements after the pig-tailing process.

At this stage, the UPP is already functional. However, in order to be able to actually control it and use it, additional interfaces are necessary to connect the device to power supplies, light sources, and temperature controllers. To this end, custom-made PCBs are fabricated and connected to the device via either electrically conductive glue or wire-bonding; thermally conductive glue is used to bond the device to a metal base that provides a large surface upon which to attach the PCBs, Peltier cells as thermo-electric actuators, and additional supports. The metal base also acts as a heat sink for the glass structure.

Finally, the UPP can be equipped with fiber arrays at input and output to easily inject and retrieve light in the waveguide network. This step is

performed by gluing («pig-tailing») the fiber arrays with transparent UV glue; the design of the waveguide network includes a fan-out section to take into account the common spacing of commercially available fiber arrays, typically 127 nm.

Pictures of two fully assembled devices can be seen in Figure 2.3.

2.3 UPP Calibration

A high-fidelity operation of a photonic processor involves an accurate modeling of the relationship between the physical control variables of the device and the optical phases implemented on the chip. In the work that follows, the control variables in question are primarily the currents flowing through the thermal phase shifters. Other state variables, such as temperature and pressure, are treated as additional parameters but their effect isn't modeled directly. In this context, *calibrating* means modeling and measuring the function $\phi(I)$ at a given temperature and pressure.

In the devices considered in this work, the main physical phenomenon of the optical phase shift is the thermo-optic effect, which can be modeled with a linear relation between the increase in temperature of the material and the refractive index change:

$$\Delta n = \beta(T_0)(T - T_0) = \beta(T_0)\Delta T \quad (2.1)$$

where T_0 is the temperature setpoint of the material and $\beta(T_0)$ is the thermo-optical coefficient of the material (which generally slightly depends on the temperature setpoint). The local temperature change in the material is proportional to the dissipated power P via Fourier's law

$$\Delta T = kP \quad (2.2)$$

where k is a coefficient with the dimensions of a thermal resistance. The dissipated power in turn is related to the electrical current flowing through the resistive heating element I via Joule's law

$$P = RI^2 \quad (2.3)$$

with R being the electrical resistance of the heating element. If L is the length of a thermal phase shifter and λ is the wavelength of the guided electromagnetic wave, the associated induced phase shift is

$$\Delta\varphi = \frac{2\pi L}{\lambda} \Delta n \quad (2.4)$$

which yields a quadratic relationship between the phase shift induced by a thermal phase shifter and its current, in the absence of other phase shifters:

$$\Delta\varphi = \frac{2\pi L\beta k R}{\lambda} I^2. \quad (2.5)$$

Because heat conduction in our UPPs is due to heat traveling through both the glass and air (due to the existence of insulation structures, that are typically filled with air), the thermal conductivity depends on the pressure setpoint P_0 of the environment, $k = k(P_0)$. We may therefore define the *tuning coefficient* $\alpha(T_0, P_0, \lambda)$ to encapsulate this quadratic relation as

$$\Delta\varphi = \alpha(T_0, P_0, \lambda) I^2. \quad (2.6)$$

In a network, however, the phase shift across a thermal phase shifter doesn't depend only on the power dissipated on itself. Some heat may be transported from a phase shifter to the region under the influence of another, for example by air conduction, inducing a change in optical phase. We call this effect «thermal cross-talk» and it is one of the main issues to take into account for the calibration of UPPs as we will thoroughly explore in the next chapter. We may model this effect as a linear combination of all other thermal shifters in the network with appropriate weights:

$$\Delta\varphi_i = \sum_{j=0}^{S-1} \alpha_{ij}(T_0, P_0, \lambda) I_j^2 \quad (2.7)$$

with S equal to the number of phase shifters in the network. There are also S such relations, one per shifter, and therefore this expression can be succinctly written as a matrix multiplication:

$$\Delta\varphi = \alpha \mathbf{I}^2 \quad (2.8)$$

where the vector $\mathbf{I}^2$ is the element-wise square of the vector of the currents, and α is called *cross-talk matrix*. The elements of this matrix depend on the temperature setpoint T_0 and the pressure P_0 of the environment as will be clear in the next chapter.

The complete description of the phases on the device must include also the phases that are naturally induced by propagation of the light in the waveguide, which is affected by fabrication tolerances. These phases are called *static phases* and they modify Equation 2.8 as

$$\varphi = \varphi_0 + \alpha \mathbf{I}^2. \quad (2.9)$$

Note that the derivation above hinges among other things on the proportionality between the dissipated power P and the squared current I^2 , i.e. Ohm's law. If this law isn't verified, higher order terms may appear in Equation 2.9:

$$\varphi = \varphi_0 + \alpha I^2 + \beta I^4 + \dots \quad (2.10)$$

2.3.1 Temperature and Resistance Stability

Ensuring the stability of thermal and electrical properties of the UPP is fundamental since the devices need to be operated for hours or even days during the calibration process or even during the actual implementation of the unitary matrix.

For this reason, before any calibration can be performed it is important to ensure that the parameters of the device do not change during the process. As described previously, the cross-talk parameters of the device may vary depending on the temperature setpoint of the whole device. For this reason, we employ temperature control systems (see Section 2.5.3) to keep a steady temperature; for reference, the standard temperature chosen in this work is 27 °C.

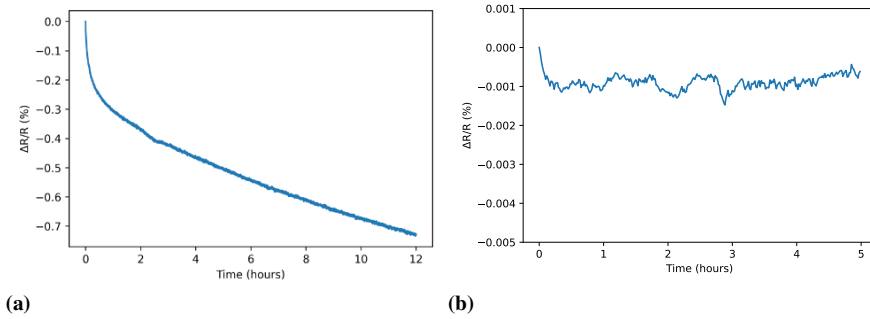


Figure 2.4: Measurement of resistance stability, showing the resistance drift as a relative percentage over time. In Figure 2.4a the experiment was performed over 12 hours, actuating a single resistive element fabricated with sputtered gold. In Figure 2.4b, the experiment was performed over 5 hours on a resistive element fabricated with chromium-copper deposition.

After the metal deposition processes, the resistive elements have not yet reached complete stability from the electrical point of view. In fact, as can be seen from Figure 2.4, running current through them for many hours causes the resistance to settle on a lower value as the dissipated heat induces a permanent resistance drift. To ensure that the drift is minimal during the actual operation of the device, then, we set the currents in the device to

dissipate the expected maximum power (typically around 60-70 mW depending on the wavelength) for up to 15 hours depending on the material. In particular, as shown in Figure 2.4, for gold this process requires a much longer time than chromium.

2.4 Operation of UPPs

2.4.1 Matrix Decomposition

Applications that involve the use of UPPs require the implementation of either a specific unitary matrix U or, if the structure of the UPP is known, specific phase shifts φ . Either way, this process requires solving Equation 2.9 for I if the φ are known, but if the application only specifies the matrix U , then some procedure is required to decompose the matrix U into the phase shifts φ .

In a Clements architecture, U is a function of the vector of internal phases θ and the vector of external phases ϕ :

$$U = U(\theta, \phi). \quad (2.11)$$

Clements *et al.* [7] have proven that a unitary matrix U can always be decomposed into a product

$$U = D \left(\prod_{m,n \in S} T_{mn} \right). \quad (2.12)$$

where D is a diagonal unitary matrix, S is a set of pairs of indices between 0 and $N - 1$, while T_{mn} is an elementary matrix of the form

$$T_{mn}(\theta, \phi) = \begin{bmatrix} 1 & 0 & & & & 0 & 0 \\ 0 & 1 & & & & 0 & 0 \\ & & \ddots & & & & \\ & & & ie^{i\phi} \sin \theta & i \cos \theta & & \\ & & & ie^{i\phi} \cos \theta & -i \sin \theta & & \\ & & & & & \ddots & \\ 0 & 0 & & & & & 1 & 0 \\ 0 & 0 & & & & & 0 & 1 \end{bmatrix}. \quad (2.13)$$

for a pair θ, ϕ of phase shifts that belong to the same MZI; T_{mn} are N -by- N unitary matrices that mix the m and n modes of the structure (in the

example above, $n = m + 1$). The proof is constructive: each T_{mn} is found explicitly by nulling off-diagonal entries sequentially in the unitary matrix U . Pseudo-code for the decomposition algorithm is provided in Algorithm 1.

Algorithm 1 Clements' Decomposition Algorithm

```

for  $i$  from 1 to  $N - 1$  do
  if  $i$  is odd then
    for  $j$  from 0 to  $i - 1$  do
      find  $T_{i-j, i-j+1}$  such that the  $(N - j, i - j)$  entry of  $UT_{i-j, i-j+1}^{-1}$  is zero
       $U \leftarrow UT_{i-j, i-j+1}^{-1}$ 
    end for
  else
    for  $j$  from 1 to  $i$  do
      find  $T_{N+j-i-1, N+j-i}$  such that the  $(N + j - i, j)$  entry of  $T_{N+j-i-1, N+j-i}U$  is zero
       $U \leftarrow T_{N+j-i-1, N+j-i}U$ 
    end for
  end if
end for

```

Moreover, it can be shown that for each MZI in the Clements architecture there is a corresponding elementary matrix T_{mn} in the decomposition. Since the Clements and Bell architectures are equivalent, the same algorithm with some adjustments can be applied to the Bell architecture [8].

In short, the Bell architecture features a unit cell (refer to Figure 1.8b) that is similar but not identical to the Clements counterpart (Figure 1.8a). Besides the obvious remapping

$$\theta = \theta_u - \theta_d \quad (2.14)$$

there is no way to remap φ to a combination of θ_u and θ_d within the unit cell alone. Notably, if a phase shift $+\Delta\theta$ is added to both θ_u and θ_d , the equivalent internal phase $\theta = \theta_u - \theta_d$ stays the same but the overall effect is the addition of an extra phase term common to both optical paths. This is different compared to the external-internal unit cell structure, where the external phase φ is a differential phase shift between the optical paths. In fact, the remapping can only be achieved when these unit cells are organized into a network.

Figure 2.5 shows the principle of the remapping for a generic UPP structure. Suppose the Clements decomposition algorithm for a matrix U requires a specific external phase φ at a physical location in the Clements structure. By actuating θ_u and θ_d together we can obtain a common mode

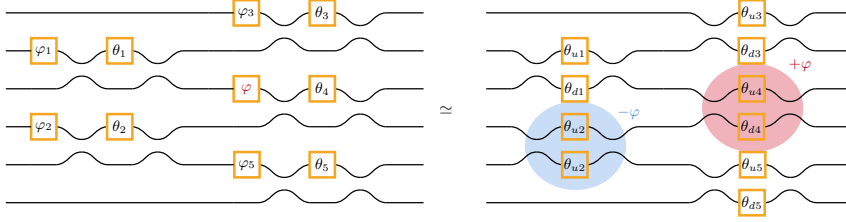


Figure 2.5: The basic idea that allows the Bell and Clements structures to be remapped onto one another. If a certain phase φ (in red) needs to be achieved on a physical location of the Clements structure, this can be created by adding and subtracting a common mode phase φ on two different MZIs of the Bell structure.

phase equal to φ , but the common mode phase applies to both optical paths whereas the Clements structure would only apply it to the upper path. In order to solve this, a common mode phase of $-\varphi$ is added to a previous MZI in the network. The end result is equivalent to placing φ only on the top path, as required.

This remapping of the phases allows any set of required phases φ for the Clements structure to be remapped into a set of equivalent required phases for the Bell structure. In this way, Algorithm 1 can be used for both structures.

What remains to do, for both structures, is to express the φ in terms of our control variables $\mathbf{I}$. A naive approach would be to invert Equation 2.9; assuming the matrix can be inverted in the first place, this isn't enough to ensure that the resulting vector

$$\mathbf{I}^2 = \alpha^{-1}(\varphi - \varphi_0) \quad (2.15)$$

has only positive entries. Since $\mathbf{I}^2$ represents the square of the currents to supply to the thermal phase shifters, this is of course a problem.

Fortunately for the purposes of solving this problem we may recall that the vector φ represents phase shifts, and therefore 2π -periodic: we can add a vector of integer multiples of 2π to φ and the resulting implemented transformation U will be the same. In other words, instead of trying to solve for $\mathbf{I}^2$ in Equation 2.9 naively, we try to solve the following problem:

$$\begin{aligned} &\text{solve} \quad \alpha \mathbf{I}^2 + 2\pi \mathbf{n} = \varphi - \varphi_0 \\ &\text{subject to} \quad I_i \geq 0 \quad \forall i, \end{aligned} \quad (2.16)$$

where $\mathbf{n}$ is an array of integers. This makes the inversion problem a Mixed Integer Linear Programming (MILP) problem, which is an NP-hard problem [18]. While sophisticated methods can be used to find solutions with

minimal power, i.e. solutions that minimize the sum of elements of I^2 , brute-force methods can be used to quickly find acceptable solutions to the problem despite not guaranteeing optimality.

The brute-force solution method I used in the work consists in the following: first, the «naive» I^2 is computed (Equation 2.15). For any squared current that isn't positive, we add 2π to the corresponding phase in φ (or, equivalently, we decrease the corresponding entry of the vector of integers n by 1). Running the naive computation with this new φ drives the currents towards becoming more positive, although it does not guarantee finding the solution with the smallest currents. The algorithm is iterative, only stopping when the whole vector I^2 is composed of positive numbers. This method is summarized in Algorithm 2.

Algorithm 2 Brute-Force Algorithm for the Currents

```

 $I^2 \leftarrow \alpha^{-1}(\varphi - \varphi_0)$ 
while not all elements of  $I^2$  are positive do
  for each element  $q_k$  of  $I^2$  do
    if  $q_k < 0$  then
       $\varphi_k \leftarrow \varphi_k + 2\pi$ 
    end if
  end for
   $I^2 \leftarrow \alpha^{-1}(\varphi - \varphi_0)$ 
end while
 $I \leftarrow \sqrt{I^2}$ 

```

2.4.2 Matrix Amplitude Measurement

To measure the amplitudes of every element of an implemented matrix, it is sufficient to have a setup that is capable of retrieving the optical power from any output port while choosing to inject light into the UPP from any input port. By measuring the optical power from every pair of inputs-outputs, we obtain a raw measurement of the squares of the amplitudes of the matrix.

The measured values, however, do not directly correspond to the squared amplitudes of the unitary matrix implemented by the UPP on two accounts. On the one hand, this measurement contains units (of power) and must therefore be normalized. On the other hand, losses at both the input and output side of the setup influence the measurement and skew the results away from the real matrix U . In particular, if M is the measured matrix and A is the matrix of amplitudes of the implemented unitary matrix, then

$$M = L_1 A^2 L_0 \quad (2.17)$$

where L_0 and L_1 are diagonal matrices that encode the effects of losses and noise on the measurement.

By unitarity of U we must have $U^\dagger U = I$. This implies that the amplitudes matrix A must have the following property:

$$\sum_j A_{ij}^2 = 1 \quad \text{and} \quad \sum_i A_{ij}^2 = 1, \quad (2.18)$$

that is, the sum of the squared amplitudes over any row and over any column must be 1. If a matrix does not have this property, it cannot represent the amplitudes of a unitary matrix. Due to losses in the measurement setup and noise represented by L_0 and L_1 , this property will in fact never be verified in practice; despite this, it's always possible to refine the raw measurement of the squared amplitudes into a matrix using Sinkhorn's Algorithm (Algorithm 3):

Algorithm 3 Sinkhorn's Algorithm

```

while  $\sum M_{ij} - N > \varepsilon$  do
  for  $i$  from 0 to  $N$  do
    Define  $R$  as the diagonal matrix where the  $i$ -th diagonal element is the sum of the
    elements on the  $i$ -th row,  $\sum_j M_{ij}$ .
     $M \leftarrow MR^{-1}$ 
  end for
  for  $j$  from 0 to  $N$  do
    Define  $C$  as the diagonal matrix where the  $j$ -th diagonal element is the sum of the
    elements on the  $j$ -th column,  $\sum_i M_{ij}$ .
     $M \leftarrow C^{-1}M$ 
  end for
end while

```

This algorithm allows one to project the (N -by- N matrix) M into an amplitude squared matrix A^2 to a level of precision ε . Applying the square root element-wise is sufficient to obtain the actual amplitude matrix. This algorithm may also be used to obtain an estimate of the losses L_0 and L_1 , although generally these values can be measured directly as described in Section 2.5.2.

2.5 Experimental Instruments Used in the Work

2.5.1 Electronics and Current Control

In the context of UPPs that employ thermal phase shifting as a way of actuating the programmable MZIs in the network, precise control of many different electrical components at the same time is crucial.

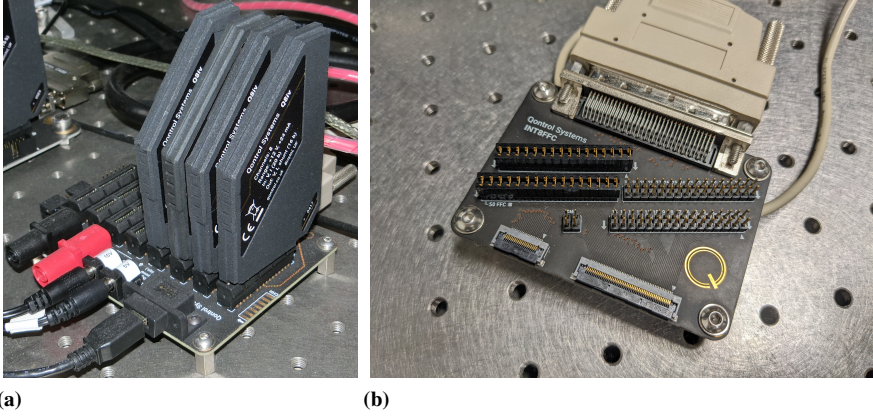


Figure 2.6: The devices used for current control. Figure 2.6a shows the Q8iv driver modules, which act as multichannel current sources; for the pictured device, up to 64 resistive elements channels can be controlled at the same time. Figure 2.6b in an interposer board, which exposes the channels of the drivers with a standard interface to more easily connect the current control to the UPP device.

Qontrol Q8iv current driver modules have been used throughout the work in order to control the current applied to the resistive heating elements. These drivers in the context of reconfigurable photonic integrated devices are extremely helpful to streamline the operation and calibration steps, as they provide the means to apply currents to many different channels independently at once.

On each channel, the Qontrol Q8iv can provide a current up to 24 mA and a voltage up to 12 V at 16-bit precision ($\pm 180 \mu\text{V}$, $\pm 370 \text{ nA}$). These devices can be controlled via Python code; the interface provides basic functionality to apply or measure current and voltage of the channels.

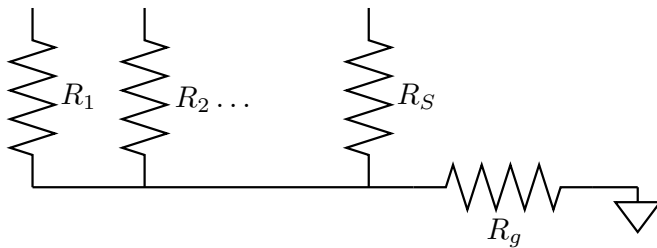


Figure 2.7: The electrical scheme of a UPP. There are numerous resistors, with resistance values R_i , all connected to the same ground node. In reality, between the common node of the resistors and the ground a small resistor should be considered, with a ground resistance R_g .

As explained by Ref. [17], controlling the resistive heating elements by imposing currents and not voltages is essential to a correct operation of the device. As shown in Figure 2.7, a set of resistors with a common ground will deviate from the ideal case by the inclusion of a parasitic ground resistance. If the resistors are actuated via setting voltages, the total current induced on the ground resistance will affect the voltage V_i across the i -th resistor as a voltage divider:

$$V_i = V_{set} - R_g \sum_j \frac{V_j}{R_j} \quad (2.19)$$

where V_{set} is the ideal set voltage on the resistor, R_g is the ground resistance, R_j the resistance of the j -th resistor and V_j the voltage across it. This electrical cross-talk effect can be removed entirely if the electrical controllers set the currents I_i instead.

2.5.2 Optical and Motion Control

For the experiments detailed in this work, we employed several different laser sources since each UPP was designed for some specific wavelength, depending on the application. In particular, wavelengths of 785 nm, 808 nm, 925 nm and 1550 nm have been used in this work. Some setups include a free-space laser that is coupled via a lens into an optical fiber, while other setups include in-fiber lasers. The details, when relevant, will be discussed within the appropriate section.

Fiber switches were employed to inject light into or retrieve the light coming out of each UPP, thanks to their simple interface and reliable performance. Throughout the work we have used multiport mechanical fiber switches (LFiber LF-OSW-925 with 16 780HP fibers, and LFiber LF-OSW-1550 with 20 SM1550 fibers) in a one-to-many configuration to redirect input light from the laser sources to any port of the device, and in a many-to-one configuration to collect the light coming from one output port at a time.

Throughout the work, three main methods of retrieving the optical power have been used:

- For single-mode optical power measurements I used Thorlabs PM16-120 and PM16-122 power meters, at the appropriate working wavelengths.
- For multi-mode optical power measurements I used a Santec OPM150 24-channel power meter, which can retrieve the optical power of up to 24 ports at the same time.

- For camera imaging, at a wavelength of 1550 nm I used a Xenics Bobcat 640 GigE camera. For imaging at wavelengths between 600 and 925 nm I used a IDS uEye SE USB Camera, model UI-2220SE-M-GL.

All these devices were connected to the central computer system in order to retrieve the optical power or the image digitally.

To collect the light at the output of the device, two main methods have been utilized:

- For devices without an optical fiber array glued to the output, a wide numerical aperture lens (NA=0.68) was used to collect the output light. A pair of automated translation stages (PI Compact Linear Stages M-112 and M-111) is used to control the position of the lens in the xy plane; this stage has a travel range of 100 mm and a resolution lower than 0.05 μm . A PI Mercury C-863 DC motor controller is used to send the commands from the PC to the stage, one for each axis.
- For devices with an optical fiber array, where multi-mode optical power measurements were not performed, a matching LFiber optical fiber switch was used in a many-to-one configuration to select one output of the UPP.

It is important to remember that every optical element between the source and the photonic circuit, as well as between the photonic circuit and a power meter, will introduce «insertion losses». For this reason, the total losses (in dB) will be

$$L_{\text{tot}} = L_{\text{UPP}} + L_{\text{ins}} = 10 \log_{10} \frac{P}{P_s} \quad (2.20)$$

where P_s is the power of the optical source and P the measured optical power with the full setup. In order to measure the insertion and UPP losses, a simple method is to take the optical setup as-is and remove the UPP, connecting each output of the first fiber switch to each input of the second fiber switch. We can then measure the new total optical power P' in this configuration. Then:

$$L_{\text{ins}} = 10 \log_{10} \frac{P'}{P_s} \quad \text{and} \quad L_{\text{UPP}} = L_{\text{tot}} - L_{\text{ins}}. \quad (2.21)$$

2.5.3 Temperature Control

To control the temperature of the devices across the different experiments, I employed a Thorlabs MTD1020T TEC controller. This device is capable of

providing up to 2 A, monitoring the temperature via a 10 k Ω thermistor. The current is sent to a Peltier cell (in this work, all cells were 62 mm $\times$ 62 mm, but in general the size and number of cells required to control the temperature of a UPP may vary). The TEC controller regulates the temperature with a software-based PID feedback loop; more in particular, the equation that governs the output current $A(t)$ is

$$A(t) = P[T(t) - T_0] + I \int_{t-W}^t [T(\tilde{t}) - T_0] d\tilde{t} + D \frac{dT}{dt} \quad (2.22)$$

where $T(t)$ is the temperature signal coming from the thermistor, T_0 is the temperature setpoint, W is the time window for integration, and P, I, D are three real numbers that must be tuned for this feedback looped system to be able to control the temperature. The parameters are tuned through a Thorlabs proprietary software that controls the board via USB communication.

CHAPTER 3

Universal Photonic Processors

3.1 Introduction

In this section we explore the main body of the research on the topic of Universal Photonic Processors (UPPs). In Section 3.2.1 I will summarize the state of the art in calibration methods and fabrication platforms of UPPs, to give context in regards to the past and present state of the research on the subject. In Section 3.2 I explore the main improvements that I introduced in the context of calibration, from the point of view of complete measurement setups and methods, as well as a deep exploration of the effects of thermal-cross talk and how they can be attenuated, and listing the results that have been experimentally obtained on actual UPP devices. In Section 3.4 I introduce a novel method to measure the actual implementation of a unitary matrix with the inclusion of the angles of the complex elements.

3.2 UPP Calibration

3.2.1 Motivation and State of the Art

The fundamental problem that underlies the correct implementation of quantum algorithms, communication protocols and signal processing in

UPPs is the problem of calibration. As explained in Section 2.3, by «calibration» we mean the act of modeling the functional relationship between the control variables of the UPP (for instance, the electrical currents flowing through the thermo-optical phase shifters) and the implemented unitary transformation.

The performance of the calibration can be quantified by introducing a metric on the space of unitary matrices. Where the full complex matrix is available, the most commonly used metric in the literature is the *fidelity* (Equation 1.26); in contexts where only the amplitudes of the elements for the implemented matrix are available however, such as when only optical power measurements over each individual mode can be performed, the most commonly used metric is the *amplitude fidelity* (Equation 1.27).

State of the Art of Calibration Methods

Several complementary directions have emerged in the literature to provide effective methods of calibrating a UPP. Since the need to calibrate the devices emerged at the same time as the UPPs themselves were being fabricated and used, the earliest forms of calibration were tailor-made for the specific design of the network. The earliest such example can be found directly in Reck's original article that introduced the triangular structure for UPPs [6], providing a simple method to obtain the parameters of the structure by individually isolating and measuring the transfer function of each MZI in the network.

This method was one of the earliest examples of «white-box» methods: complete knowledge of the physical structure of the UPP is assumed, and calibrations are typically carried out by retrieving each element of the cross-talk matrix and static phases (see Equation 2.9) via progressive algorithms.

In more complex meshes such as the Clements layout, where the effect of each individual MZI cannot be easily isolated from the others, the earliest forms of calibration methods featured detectors at every node of the network [19] to individually measure the transfer function of each MZI. This required the addition of several components to the UPP design, which is generally prone to introducing additional issues down the line. Other methods can perform this type of measurement non-invasively, but they rely on specific properties of the fabrication material such as measuring the change in impedance in silicon waveguides due to surface absorption [20].

For this reason, the research focused on methods that can retrieve the desired parameters only using external detectors, i.e. power meters only at the output of the UPP. This had already been demonstrated in the more limited context of optical switching applications [21], but external-detector

white-box methods were eventually extended to the complete measurement of all calibration parameters [22].

White-box methods of this kind employ a sequence of algorithmic steps to isolate and measure the effect of each MZI in the network, extracting the parameters one at a time by fitting the curve between the measured optical power at the output and the control variables (for example, the currents to supply to the thermo-optical phase shifters).

External-detector methods like these still operate on a white-box level, leveraging the full knowledge of the geometry and design of the device to retrieve the elements of the cross-talk matrix. This is a massive advantage since the model is considering a set of known equations that govern the actual integrated circuit. On the other hand, the precision of these methods is directly related to the level of detail of the model; while the earliest methods were fabrication-platform-agnostic [22], the unique challenges presented by each fabrication platform have to be considered to adapt the method. For example, for UPPs that employ thermal phase shifting it is important to consider the effects of thermal cross-talk [17].

Moreover, every step typically relies on the data retrieved at previous steps, which causes possible errors in the fitting or measurement to propagate downstream.

At the same time, research was also pushing the boundaries of UPP operation in a different direction: since the main advantage of UPPs is the ability to arbitrarily reconfigure the implemented transformation of light states, one can leverage this fact by making the UPP «self-configure» into the target transformation [23, 24]. This can be achieved by using a closed feedback loop that tunes the control variables until the target is achieved with the desired degree of accuracy. These methods are particularly useful for example in applications where a small number of target configurations must be achieved by the UPP, or if the calibration parameters drift over time and a full re-calibration of the device is impractical.

These methods, in principle, do not require complete knowledge of the physical model of the UPP, thus being applicable to a variety of structures. They are limited by the fact that they do not provide a full understanding of the UPP, and they don't retrieve the parameters of the calibration (i.e. cross-talk matrix and static phases), limiting their usefulness in several applications. Nevertheless, these methods employ a completely novel idea in that the reconfigurability of the UPPs can be used to apply optimization algorithms for their calibration. From this line of thinking was born a different class of methods which may be called «black-box».

These methods retrieve the cross-talk matrices and static phases by em-

playing an optimization algorithm based on a random dataset (I, U) of currents and matrices. The algorithm tunes the values of the internal model (for example Equation 2.9 and the fact that U depends on φ) to minimize a suitable loss function.

They offer the obvious advantage of retrieving the model parameters without needing to develop ad-hoc procedures, instead relying on general optimization algorithms to obtain an acceptable model. The main disadvantage of these models is that as the number of control variables increases, so does the amount of data needed to optimize the model and so does the computational complexity of every step of the optimization; for large N , they may require substantial computational power, data and time to produce a performant model. Moreover, there is no guarantee that the produced model is actually simulating the actual physics of the device; in order to obtain optimized models that more closely follow sensible physical laws, additional constraints on the parameters can be added and explicit cross-talk data can be measured from the devices directly. These methods may be considered *gray-box* methods, and they offer an alternative that sits in the middle between the two calibration methods.

An example of performant and widely applicable method is the one introduced by Ref. [25], which employs a digital simulator of the UPP to apply a machine-learning-based optimization algorithm to obtain the calibration parameters α, φ_0 . This method was applied to a physical 12-mode circuit fabricated in silicon nitride, obtaining an average amplitude fidelity (evaluated over a set of 1000 Haar random matrices) of 0.9977. The method has the advantage of being easily adaptable to different sizes and materials, provided the simulator is tuned for the specific challenges presented by each fabrication platform.

In the following sections I will be detailing how I applied these methods to FLWW UPPs, how I handled the challenge of keeping the advantages of these methods as the devices became larger and more complex, and the experimental results that I obtained from this research work.

State of the Art of Fabrication Methods and Performance

A complete discussion in the context of UPPs would be missing an important aspect if no information was reported about the actual performance of real UPPs found in the literature. As the calibration method of choice and the solutions to emergent problems are typically linked with the fabrication technology and the design of the UPP, a proper description of the environment must be given to contextualize the results provided in this work.

Silica-on-silicon (SOS) is a platform where waveguides are fabricated

using CMOS-compatible photolithographic methods, leveraging a mature technology to create integrated optical devices. Doped silica is deposited on a silicon wafer, which is then etched and buried in a silica cladding layer. The main advantage of this technology is in the very low propagation losses (0.05 dB cm^{-1}) and the compatibility with CMOS processes for the integration into larger and more complex systems. This was one of the first platforms to produce a fully featured UPP [26], a 6-mode device with a Reck structure. Ref. [26] reports the performance over different processes involving multiphoton states. Over 100 Haar random matrices the reported fidelity is $\mu + \sigma = 0.999 \pm 0.001$, although this figure of merit is evaluated only by looking at the photon distribution in the output when injecting single photons into one input. Still, among the reported performance figures of merit this one is the closest to an average amplitude fidelity and can be used as a benchmark to compare the performance of other 6-mode UPPs.

Silicon Nitride (SiN) is a similarly CMOS-compatible platform that employs Si_3N_4 and makes use of both its thermo-optical coefficient for shifting and its nonlinear properties to create various linear and nonlinear circuits. Several different integrated optical chips have been developed with this technology; most notably in the context of universal photonics, 12-mode [27] and 20-mode [28] UPPs have been created, showcasing how the platform is capable of integrating hundreds of optical elements on a small footprint thanks to the very low bending radii ($100 \mu\text{m}$) compared to other platforms (15 mm for FLWW). For the 12-mode, Ref. [27] reports an average amplitude fidelity over 1000 Haar random unitaries of 0.987 ± 0.006 . For the 20-mode UPP the reported average amplitude fidelity is 0.974 ± 0.005 for the same dataset size.

Silicon-on-insulator (SOI) is by far the most developed fabrication platform for integrated optics. The technology uses a silicon guiding core buried in a silica cladding; the high index contrast provides very small bending radii ($1 \mu\text{m}$ [29]) which allows for extreme miniaturization of optical devices and very high density of integration. The major downside of SOI is the difficulty in integrating it with other fabrication platforms, since the high index contrast means low compatibility with low contrast technologies such as glass optical fibers. So far, several different large-scale integrated optical devices have been designed and used for specific quantum algorithms and protocols [30,31] which does highlight the potential for this technology to produce UPPs with large number of modes.

For what concerns *direct writing techniques* such as FLWW, major advancements in the calibration and operation of UPPs modes were achieved by me with the research described in this work, some of which was con-

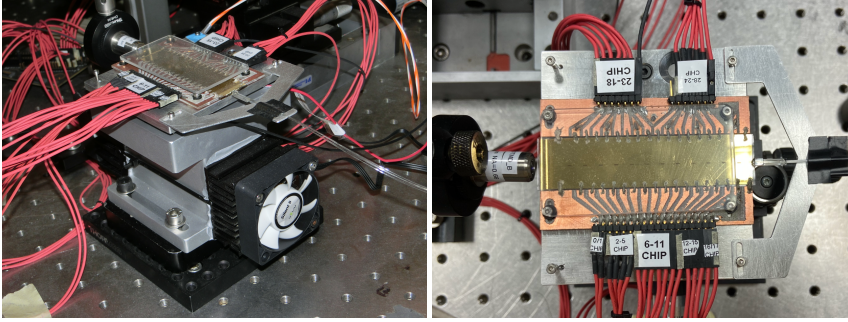


Figure 3.1: *The first setup to measure the inputs and outputs of a UPP. This setup is fully automated, with a fiber switch at the input (not pictured) and an automated stage at the output (not pictured) controlling the lens that collects the output light.*

ducted in collaboration with Ephos inc.; in parallel with the research, the state of the art for FLWW in terms of complexity and performance was pushed forward by Ephos Inc. with a 24-mode UPP [32]. This device was written in an alumino-borosilicate glass slab using FLWW, and features 576 chromium-copper heating elements for thermo-optical phase shifting. The UPP’s performance was tested on a set of 2000 Haar random matrices, obtaining an average amplitude fidelity of 0.997.

In the rest of this work, I will be exploring the methods and results of the calibrations that I have performed during the three years of my Ph. D., and how they compare to the environment described so far.

3.2.2 Calibration Setups

Fast Automated Calibration

While calibration methods that range from white-box to black-box methods are known and used in the literature, it is important to be able to calibrate UPPs with setups that are stable, fast and automated, in such a way that the performance of the UPP across a wide range of applications can be measured easily.

At the beginning of our research, the state of the art UPP produced using our research group’s FLWW setups was a 6-mode UPP (Figure 3.1). This device features a network of 15 reconfigurable MZIs, in a Clements configuration, tuned for operation at 785 nm wavelength, where phase shifting is controlled by 30 resistive heating elements. The elements consist of a thin and narrow strip of gold, deposited on the waveguides during the fabrication process via metal sputtering and patterned using femtosecond laser pulse ablation as described in Chapter 2.

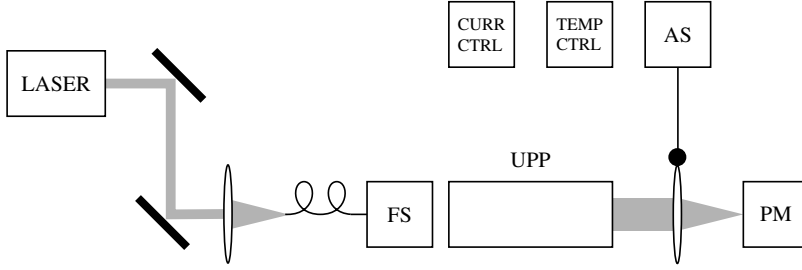


Figure 3.2: The setup used for the first automated measurements. The light from the laser source is focused into a fiber, while the fiber switch (FS) selects which input of the UPP the light is injected to. Then, the light is collected at the output by a lens mounted on an automated stage (AS) and the optical power is read by a power meter (PM). The choice of using an automated stage for the output was made in accordance with the specifications of the UPP, since it did not feature an output fiber array.

Initially, this device was slotted into a setup that included a free space laser coupled into a fiber, which was then coupled into the UPP via a manual translation stage in a butt coupling configuration; the light at the output was collected by a wide numerical aperture ($NA=0.68$) lens directly focusing the image of the chip's output ports into a power meter. The current flowing through the resistive elements on the device was controlled using a series of Qontrol Q8iv microcontrollers.

This device could be calibrated and operated manually by a human using the described setup. The calibration method chosen for this UPP was a white-box method similar to Ref. [22], developed during my Master's thesis and applied to this device.

The microcontrollers could be operated via Python code, but a lot of time was required to change the input and output ports by hand and the feedback required to ensure that the calibration process did not halt. For this reason, I introduced a series of changes to the calibration setup that built towards a full automation of the setup. In particular, the manual translation stages for both the fiber and the lens have been replaced by automated translation stages (a pair of PI Mikromove M-155.11) and an optical fiber switch at the input (LFiber LF-OSW-925) operated via Python code. This enhancement of the setup allowed me to fully automate the collection of power at each input and each output of the UPP, which is the requirement for a full calibration of the UPP.

As explained in Section 2.3, the calibration depends also on the temperature setpoint of the chip. For this reason, I added a Peltier cell below the base of the UPP, with a PID controller to monitor and regulate the temper-

ature of the UPP. This allowed me to stabilize the temperature setpoint, so that the calibration parameters of the UPP are reproducible across different environments even outside of the laboratory. A scheme of the full setup is shown in Figure 3.2.

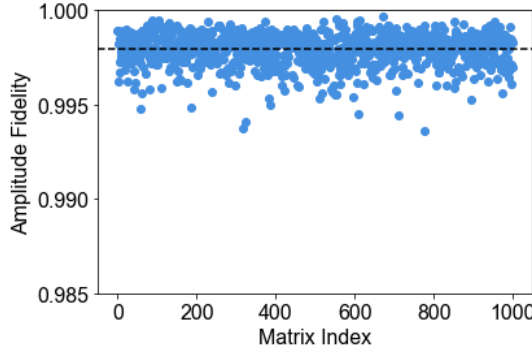


Figure 3.3: A graph of the amplitude fidelity evaluated over 1000 Haar random unitary matrices. The average value of fidelity over this set is 0.9979, highlighted by the dashed black line.

Using this setup, coupled with better and more capable software that could coordinate all of the different devices simultaneously, the time to calibrate this 6-mode UPP was reduced from several days with human input down to under 6 hours without any human input. I tested the performance of this calibration over 1000 Haar random matrices, and the result is reported in Figure 3.3; the average amplitude fidelity over this set is $\mu + \sigma = 0.9979 \pm 0.0009$, which is comparable to the 6-mode in silica-on-silicon by Ref. [26].

The temperature of the UPP is an important factor in ensuring the calibration provides reproducible results. To evaluate the impact of the temperature on the calibration, a set of 100 unitary matrices was chosen; implementing these matrices at 20 °C resulted in an average amplitude fidelity of 0.9975 ± 0.0016 . On the other hand, implementing the same matrices at 25 °C, without changing calibration parameters, resulted in an average amplitude fidelity of 0.9937 ± 0.0036 .

While an amplitude fidelity equal to one would denote identical matrices, in practice this value is affected by errors and noise that are intrinsic to the measurement process, such as detector noise, electrical noise, and more. For this reason, there exists an upper bound, an «experimental limit of fidelity», above which two matrices are considered to be essentially the same as any difference between them could be ascribed primarily to experimental

noise. This experimental limit, for this setup, was estimated by measuring the same implemented matrix 100 times, denoted by U_n , and comparing the average amplitude fidelity across all possible pairs (U_i, U_j) with $i \neq j$; this value is found to be 0.9997 ± 0.0002 for the described experimental setup.

The difference between the average amplitude fidelity over the Haar random set and the experimental limit suggests that more work can be done to optimize the calibration; for this reason, I turned my attention to researching methods to further optimize and refine the calibration and operation of the UPPs.

3.2.3 Single-Matrix Optimization

Motivation

Some applications of UPPs do not require the full power of their reconfigurability. For example, tasks such as input switching, path-encoded quantum key distribution, or multiphoton Hong-Ou-Mandel interferometry may only require a limited set of phase shift configurations since they need to employ only a finite, discrete set of unitary matrices. In these applications it is crucial to achieve the highest possible fidelity of implementation for the chosen set, rather than obtaining a more complex model that tries to maximize the performance over the whole space of unitaries.

Moreover, achieving a more refined implementation of single matrices would confirm the hypothesis that the small gap between the average fidelity obtained with the calibration and the experimental limit measurable within the setup is not due to physical limitations of the device in question, so that this gap may be filled by a more refined process that goes beyond the calibration alone.

In this context, the speedup and ease of use of the automated setups was so significant that it opened up the possibility of employing new optimization methods to further optimize the performance of the UPP in this context that require large amounts of data. In particular, tailor-made optimization methods could be applied to refine the implementation of a single matrix. To this end, I introduced an optimization algorithm based on the Nelder-Mead method [33] to take a single implemented matrix U_0 and, through iterative steps, refine the implementation of the matrix to push the amplitude fidelity of the matrix at the i -th step, U_i , as close to 1 as possible.

Method and Setup

We begin from a fully calibrated UPP, but in principle this method could be applied even if the calibration is not fully completed as long as the initial

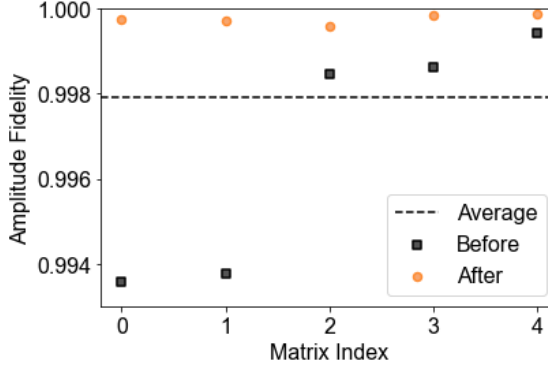


Figure 3.4: The amplitude fidelity before and after applying the optimization process as described. The black dashed line corresponds to the value of the average (amplitude) fidelity of the original Haar random set.

guess is reasonably close. We introduce a loss function $L(\phi)$ defined as follows:

$$L(\phi) = 1 - \mathcal{F}_{\text{ampl}}(U_t, U(\phi)) \quad (3.1)$$

where ϕ is the set of phase shifts implemented by the UPP, and $U(\phi)$ is the matrix that the UPP implements with those phase shifts. We can then employ any optimization method to minimize this loss function. Conceptually, the optimization algorithm minimizes the function L by tuning the parameters ϕ , which in turn are a function of the physical currents in the UPP I . So in practice we don't minimize over the phase shifts themselves, but over the currents:

$$L(I) = 1 - \mathcal{F}_{\text{ampl}}(U_t, U(\phi(I))). \quad (3.2)$$

The caveat is that in order to evaluate $L(I)$ at the i -th step of the optimization process the fidelity $\mathcal{F}_{\text{ampl}}(U_t, U_i)$ must be evaluated, and therefore the implemented matrix U_i must be measured. This is a process whose execution time can scale quadratically with the number of modes, and if the measurement process is slow, the repeated application of the iterative optimization might take a long time. However, the increase in speed of the automated setup makes this process reasonably fast: with the described automated stages, a single 6-by-6 matrix could be measured in 20 to 30 seconds. This rate can be increased considerably by employing faster port switching devices and parallelizing the output measurement, as will be further detailed in Sections 3.2.7 and 3.2.8.

After considering various options for the optimization algorithm and evaluating which one was the best suited for the task, I chose the Nelder-

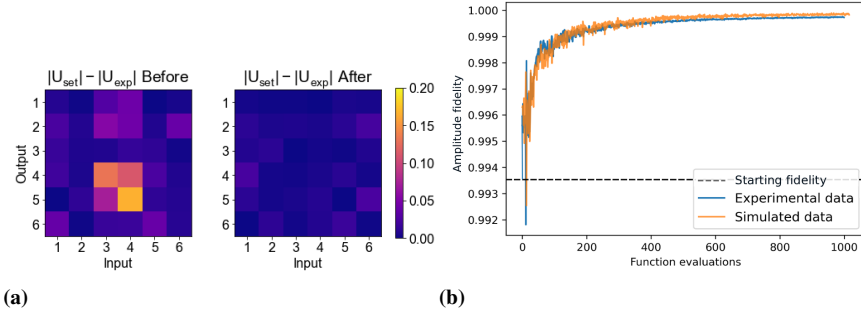


Figure 3.5: Effects of the single-matrix optimization algorithm on the worst implemented matrix out of 1000. Figure 3.5a shows a comparison of the difference between the amplitudes of the target matrix before and after the optimization; the algorithm has completely removed a 2 by 2 submatrix that carried the biggest error. In this picture, the considered matrix is the worst of the whole dataset, whose fidelity improved from 0.9934 to 0.9997. Figure 3.5b shows the improvement of the fidelity over the iterations of the algorithm; a simulation of the optimization was also performed on a digital twin of the UPP (see Section 3.2.5).

Mead algorithm and applied this method to 5 unitary matrices. The 5 matrices were picked based on the amplitude fidelity of the implementation after the full calibration of the UPP: I picked the two worst implemented matrices, one matrix with a fidelity on the average, and two among the best implemented matrices. As can be seen from the graph in Figure 3.4, the process increased the amplitude fidelity from 0.9976 to 0.9996 on average, which is also within the experimental limit of this setup.

This a successful application of the machine learning optimization method to provide a higher fidelity implementation of these matrices, and it provides the first successful instance of a self-configuring calibration method being applied to a FLWW UPP.

Moreover, it demonstrates that FLWW UPPs can in principle reach extremely good performances, up to the experimental limit of distinguishability in this setup. This result clearly highlights that, since the UPP is capable of implementing selected matrices extremely well, the major factor in providing high-fidelity implementation across a wide range of matrices is still the quality of the calibration.

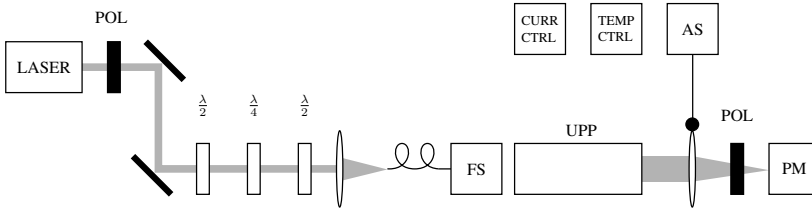


Figure 3.6: The setup used for the measurements with polarization state selection. The light from the laser source passes through a polarizer and a series of waveplates before being focused into a fiber; the fiber switch (FS) selects which input of the UPP the light is injected to. Then, the light is collected at the output by a lens mounted on an automated stage (AS) and the optical power is read by a power meter (PM).

3.2.4 Polarization Independence

Motivation

While the automated calibration can achieve extremely good results with minimal human intervention, and the single-matrix optimization can further refine the fidelity, the question remains about whether or not these processes are robust against changes in the physical properties of the light, polarization in particular. In all previous experiments, in fact, the polarization state of the light was not controlled.

Waveguides fabricated via FLWW feature birefringence, enough to be able to create integrated devices that can precisely control and transform the polarization [34]. It is therefore important to test whether the polarization state of the light has an impact over the behavior of the UPPs, testing the reproducibility of the calibration across a wide range of polarization states.

Method and Setup

I introduced a simple variation of the described setup that could control the polarization state of light in the device, in order to measure its performance across several polarization states. The setup consists of setting up a polarizer and a set of three polarization rotating plates (one half-wave plate, one quarter-wave plate, and another half-wave plate) in free-space before the laser light is coupled into the fiber. The output of the setup is mostly unchanged, with the addition of a polarizer before the power meter. A schematic representation of the setup is shown in Figure 3.6.

With this setup, it is possible to select a specific polarization state (for example, the linear vertical V state) by setting the polarizer at the output of the setup to the orthogonal position H , and minimizing the optical power measured by the power meter by rotating the waveplates. Indeed,

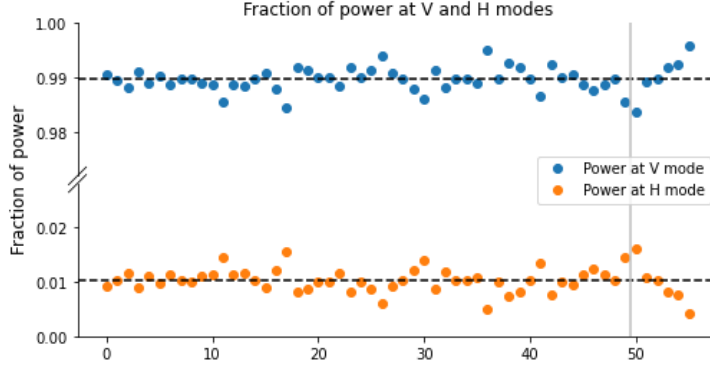


Figure 3.7: Scatterplot of the fraction of optical power between the V and H states measured by selecting the polarization state at the output to be H . The grey vertical line separates the Haar random set from the switching set.

any polarization state of light can be achieved by rotating the plates, and this procedure ensures that the light at the output of the device is in the V polarization state. This allows us to set a specific polarization state for the light within the device, so that it's possible to swap between orthogonal polarization states by rotating a $\lambda/2$ waveplate by $\pi/4$.

Experimental Results

Firstly, I measured the stability of the polarization state when the UPP is actuated, while the polarizer is still set to the orthogonal direction H and the state V of the light is selected. I implemented a set M of 56 matrices, 50 Haar random matrices and 6 randomly selected switching matrices, as shown in Figure 3.7. The measurement reports that the fraction of power at the output is stable around 1%, which is also the value that could be measured when the UPP is completely passive and no current is flowing through the resistors; in other words, turning on the resistors and changing the path taken by the light in the UPP does not appreciably change the overall birefringence of the UPP. This is an important result for the stability of the UPP's behavior, especially in the context of the method that will be described in Section 3.4.

After verifying the stability, I moved on to evaluating the performance of the calibration across several different polarization states.

1. To provide a baseline value, I removed the polarizer and waveplates, letting the polarization state of the light be the same as the one used for the calibration - which we may call A for «arbitrary». I then im-

plemented the same set M of matrices, measuring their amplitudes.

2. After putting the setup back together, I selected the polarization state V and measured the whole set M .
3. After rotating the half-wave plate by $\pi/4$, I was able to select the polarization state H . I measured the whole set M again.

With three different polarization states A, V, H I could then start to evaluate the fidelity of the implementation depending on the polarization, and how different the matrices really are between each polarization state.

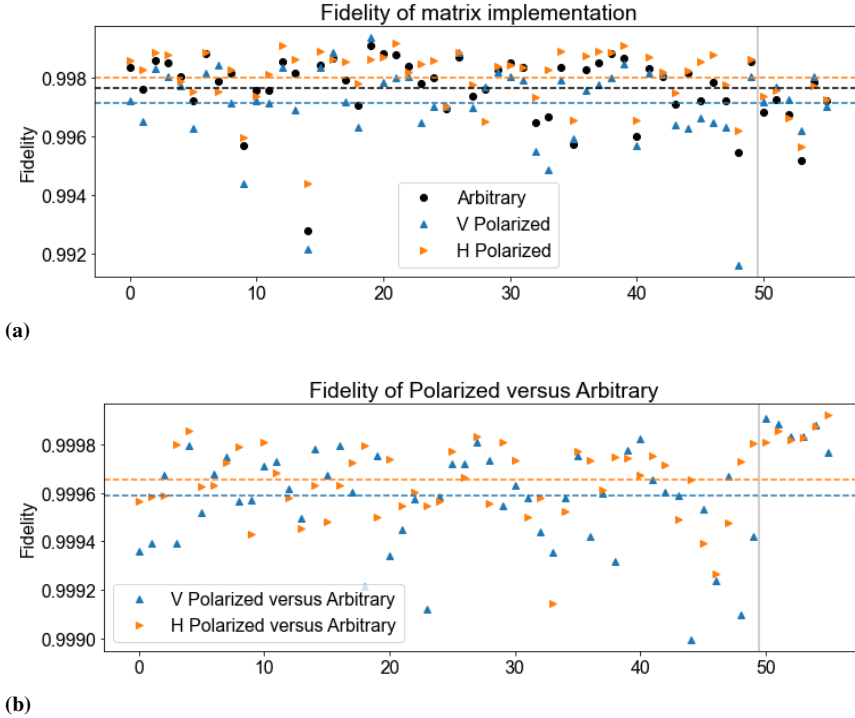


Figure 3.8: Scatterplots of the amplitude fidelity for the first experiment (panel (a)) and second experiment (panel (b)). The grey vertical line separates the Haar random set from the switching set.

The graphs of the main comparisons can be found in Figure 3.8. First, for each matrix in M I evaluated the amplitude fidelity in the three polarization states against the target matrix U ; the average amplitude fidelity was 0.9976, 0.9980, 0.9971 for the states A, V, H respectively. Interestingly, the H state performed consistently better than the V or even the A state.

Secondly, I evaluated the amplitude fidelity between the A -state-measured matrix against the V and H states separately. This gives a rough

estimate of the similarity of the «selected» states against the arbitrary state, obtaining 0.99960 and 0.99966 for the V and H states respectively. The fact that these fidelities are higher than in the previous case suggests that the polarization is not the main effect responsible for the measured value of amplitude fidelity against the target matrix, especially since the difference in performance between two opposite states can mostly be found at the fourth decimal place. Comparing this with the experimental limit (0.9997 ± 0.0002), we conclude that the two matrices are almost indistinguishable from the point of view of the setup.

In other words, it is possible to deduce from these experiments that the polarization can in fact play a role in the performance of the calibration, but for the UPP under scrutiny this effect is small enough as to be negligible.

3.2.5 Full Machine Learning Calibration

Digital Twins for Simulations and Machine Learning

To further speed up the data analysis of the calibration and to simulate the UPPs in a controlled environment, I created a full digital simulator for UPPs. This simulator is capable of generating the basic 2-mode unit cell of a Clements or a Bell layout, including fabrication tolerances on the reflectivity of the beam splitters. To model the error on the beam splitters, it is important to correctly model the transfer function of MZIs featuring directional couplers that deviate from the 3 dB coupler behavior. As reported in Equation 1.16, the more general equation can be described with an angle η that determines the power split of the directional coupler.

The full transfer matrix of an integrated MZI, for the external-internal design, is therefore the matrix product

$$T_{2,C} = \begin{bmatrix} \cos(\eta_2) & i \sin(\eta_2) \\ i \sin(\eta_2) & \cos(\eta_2) \end{bmatrix} \begin{bmatrix} e^{i\theta} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \cos(\eta_1) & i \sin(\eta_1) \\ i \sin(\eta_1) & \cos(\eta_1) \end{bmatrix} \begin{bmatrix} e^{i\varphi} & 0 \\ 0 & 1 \end{bmatrix} \quad (3.3)$$

where η_1 is the angle relative to the first directional coupler and η_2 the one relative to the second. To shorten this expression we may write $\cos(\eta_i)$ and $\sin(\eta_i)$ as r_i and t_i respectively. The expression then evaluates to

$$T_{2,C} = \begin{bmatrix} e^{i\theta} e^{i\varphi} r_2 r_1 - e^{i\varphi} t_2 t_1 & i e^{i\theta} t_2 r_1 + i r_2 t_1 \\ i e^{i\theta} e^{i\varphi} r_2 t_1 + i e^{i\varphi} t_2 r_1 & r_2 r_1 - e^{i\theta} t_2 t_1 \end{bmatrix}. \quad (3.4)$$

For the push-pull design, the same calculation can be performed:

$$T_{2,B} = \begin{bmatrix} r_2 & i t_2 \\ i t_2 & r_2 \end{bmatrix} \begin{bmatrix} e^{i\theta_u} & 0 \\ 0 & e^{i\theta_d} \end{bmatrix} \begin{bmatrix} r_1 & i t_1 \\ i t_1 & r_1 \end{bmatrix} \quad (3.5)$$

which evaluates to

$$T_{2,B} = \begin{bmatrix} e^{i\theta_u} r_2 r_1 - e^{i\theta_d} t_2 t_1 & i e^{i\theta_u} t_2 r_1 + i e^{i\theta_d} r_2 t_1 \\ i e^{i\theta_u} r_2 t_1 + i e^{i\theta_d} t_2 r_1 & e^{i\theta_d} r_2 r_1 - e^{i\theta_u} t_2 t_1 \end{bmatrix}. \quad (3.6)$$

Note that r_i and t_i are respectively the absolute values of the reflectivity and transmissivity of the coupler, i.e. the square root of the fraction of power at bar and cross respectively. For any coupler we can define the error, or deviation from 3 dB behavior, as

$$\varepsilon_i = r_i^2 - \frac{1}{2} \quad (3.7)$$

and we can compile them into a vector ε , which enters the digital twin model of the UPP as an additional set of parameters.

The software can then piece together these unit cells in order to recreate the network of MZIs for any number of modes taking into account thermal cross-talk effects, as described in Section 2.3. By including the Calibration Equation into the model (Equation 2.9) it can generate, parametrically, the full map from currents I to unitary matrix U .

This software-based simulator is therefore capable of generating unitary matrices $U(I, \alpha, \varphi_0, \varepsilon)$ starting from a complete physical model of the chip, the currents I supplied to each thermal phase shifter, and a set of tunable parameters $\alpha, \varphi_0, \varepsilon$ that encode cross-talk effects, static phases and the fabrication tolerances respectively. This simulator may be equipped with the cross-talk parameters originating directly from a physical UPP, obtained via calibration; in this case, we call it a *digital twin* of the physical object.

The digital twin offers the possibility of testing machine learning optimization algorithms in a controlled environment. For example, it can be used to test the behavior of the single-matrix optimization algorithm to obtain an estimate of the improvement that can be achieved on a specific unitary matrix without having to run the optimization algorithm in the laboratory (see Figure 3.5b).

Machine Learning Calibration of UPPs

An automated setup can be left to collect data for a long time, and a properly automated setup can have a very high rate of data collection. It becomes then possible to use large amounts of data to not only improve the performance over a discrete set of matrices, but calibrate the entire device in one go.

The main method I introduced to the machine learning calibration of UPPs consists in employing the method of digital twins to fully calibrate a physical UPP. The method consists of the following steps:

1. Collect a large amount of matrices from the physical UPP. In our devices that employ thermal phase shifting, these matrices can be randomly picked by selecting values for the current at random so that the square of the current is uniformly distributed. This is equivalent to picking a phase at random for each phase shifter, although thermal cross-talk skews the distribution away from uniformity. At the end of this step, a data set $\{(U, I^2)\}$ of matrix-current square pairs is obtained.
2. Create a digital twin of the UPP, equipping it with an initial guess for the cross-talk matrix α . This initial guess can originate from a previous calibration of the device, or simply by assuming no cross-talk between adjacent MZIs is present and using the values of the resistances to evaluate the cross-talk coefficient (see Equation 2.5).
3. Apply an optimization algorithm to maximize the fidelity of the digital twin across the dataset. A fraction of the dataset may be kept as test set, to evaluate the performance of the calibration on unseen data. The speed of this step is not limited by the speed of the measurement setup, but only by the speed of the hardware on which the optimization algorithm and the digital twin are running.
4. At the end of the previous step, we have a fully calibrated digital twin. The parameters of the digital twin can then be applied to the physical twin, which completes the calibration process.

At the end of this process, the UPP is fully calibrated and the parameters $(\alpha, \varphi_0, \varepsilon)$ have been retrieved so that the device is ready to be operated. Examples of cross-talk matrix α for different structures are shown in Figure 3.9, while an example of coupler error vector ε is shown in Figure 3.10.

This method was applied to UPPs with varying numbers of modes: 6, 8, and even 12 modes. A comparison of the performance over a representative of each design is shown in Figure 3.30. The fidelities for all designs are very high, especially when compared to similarly-sized UPPs in the literature; in the following sections I will go in further detail over how each UPP performs, taking a representative of each size and comparing the performance with state of the art devices of the same number of modes.

These results are a testament to the usefulness and flexibility of a platform like FLWW, or more in general direct writing in glass substrates, for

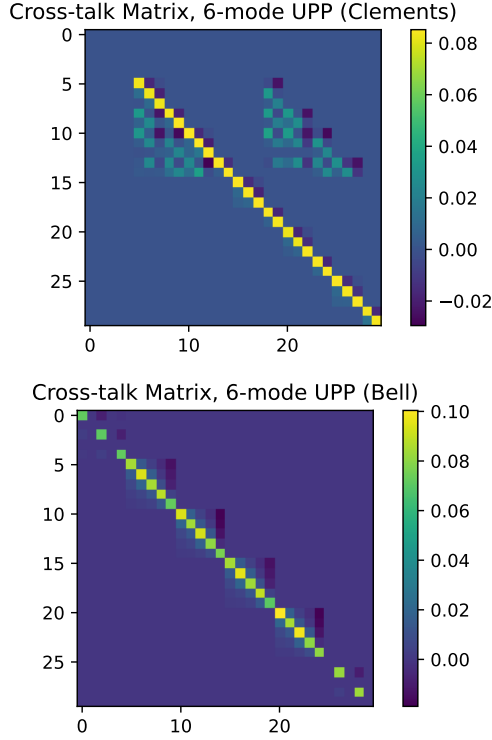


Figure 3.9: Two different cross-talk matrices α , for a Clements-type structure (left) and a Bell-type structure (right); the off-diagonal elements represent cross-talk terms while the diagonal elements are self-heating terms. Both of these were obtained from the calibration of 6-mode UPPs. Empty rows and columns in these matrices correspond to heating elements that are not actually used during operation.

the fabrication of integrated optical processors, but a few observations are in order. From the Figure we can see that the amplitude fidelity trends downwards as the number of modes increases; this is understandable on several fronts.

The ability for a UPP to implement any unitary, even in principle, depends on the extinction ratio of the MZIs on the central portion of the layout when they are set to their bar configuration (that is, when the relative phase shift between two arms of the interferometer is $\Delta\phi = \pi$). If the extinction ratio is nonzero, the UPP cannot implement all unitary matrices as some portions of the space of unitaries become inaccessible - which implies that unit fidelity cannot be achieved for all matrices, not even in principle; this is typically not particularly relevant for our UPPs, but as the number of modes increases this requirement becomes more and more stringent [35]

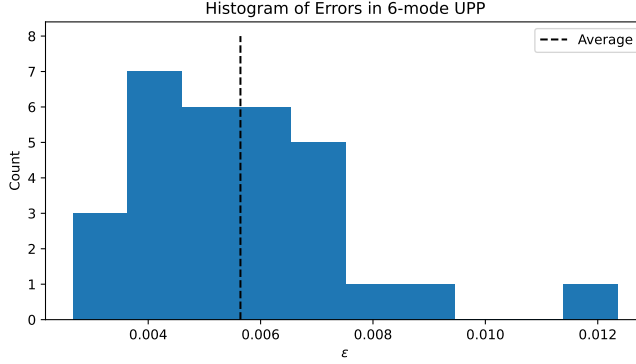


Figure 3.10: Histogram of deviations from 3 dB behavior for each coupler in a 6-mode. The average, marked by the dashed line, is 0.0056.

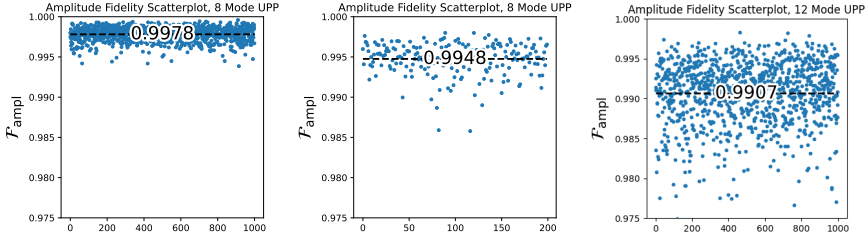


Figure 3.11: A comparison between the performance after automated calibration of 6, 8 and 12 mode UPPs.

and larger devices are increasingly difficult to fabricate while maintaining accuracy and precision over the whole area.

For this reason, it becomes difficult to maintain the same level of fidelity, even the theoretically achievable one, as the number of modes increases, beyond what the calibration can compensate for.

3.2.6 6-mode Automated Calibration

The automated calibration has been performed across several different 6-mode devices throughout my work, each of which was designed and built for different applications. As a representative of this set I will describe the results obtained on the 6-mode device pictured in Figure 3.12.

This a UPP device featuring a Clements layout with internal-external unit cells. The waveguides were fabricated by optimizing the guiding behavior at the 808 nm range; the losses are around 3 dB on average over all modes before pig-tailing and packaging, but they rose to between 5-10 dB after packaging. Indeed, the packaging was specifically developed for the

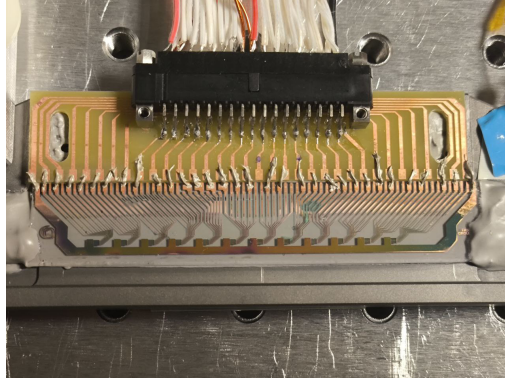


Figure 3.12: A picture of the 6-mode UPP seen from the top.

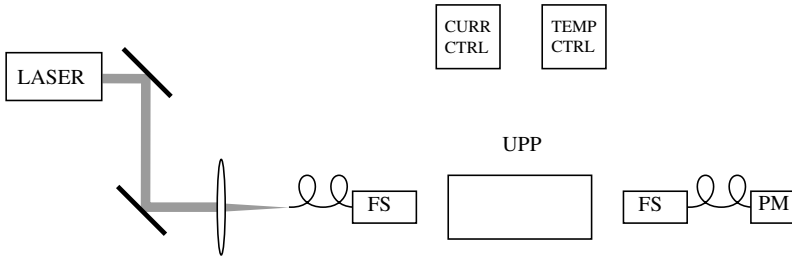


Figure 3.13: The automated setup that I introduced to calibrate UPPs with fiber arrays at both input and output.

target application, namely space deployment, and therefore prioritized mechanical robustness and reliability over optical loss minimization, as ultra-low losses were not a strict requirement for the intended use case. The 72 phase shifters on the device have been fabricated with chromium-copper metal layer deposition; the pattern was designed to be able to interface the phase shifters with a custom PCB for the electrical connections. As is the case for almost all the UPPs described in this work, redundant phase shifters were fabricated that ensure a backup in case some shifters were to stop working; during operation, in fact, only 25 phase shifters are strictly necessary (Figure 3.14). During assembly, the UPP was equipped with input and output fiber arrays, and the whole chip was glued to a titanium heatsink specifically designed for the UPP.

To accommodate the features of this and similarly-built UPPs, the automated setup was slightly altered to include a second optical fiber switch, to select the output of the device. A scheme of the new setup is shown in Figure 3.13. The calibration was performed by taking a training dataset of 1000 Haar random matrices and letting the digital twin based algorithm optimize

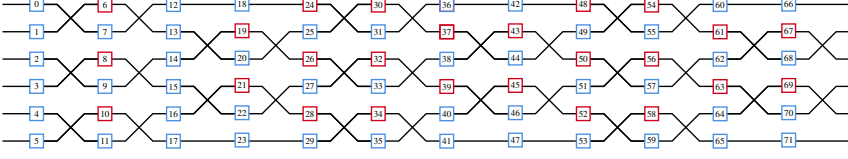


Figure 3.14: The waveguide structure of the 6-mode UPP; the location of each phase shifter is highlighted and numbered. Blue shifters correspond to those that were fabricated, but are not connected to the PCB; red shifters are those that are connected to the PCB and actually used during calibration and operation. Every crossing corresponds to a single 3 dB coupler.

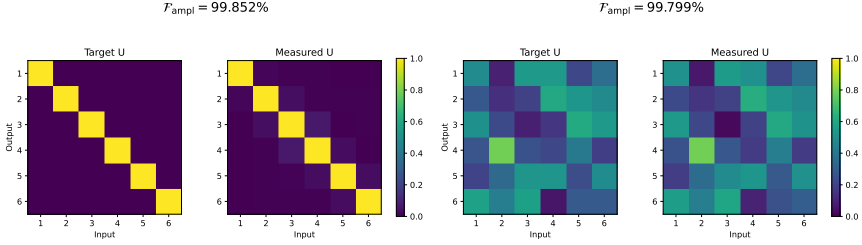


Figure 3.15: Example of performance of the 6-mode UPP over the identity matrix (left) and a randomly selected Haar random matrix from the test set (right).

the parameters before transferring them to the UPP, as described previously. This resulted in a value of amplitude fidelity of around 0.9978 ± 0.0008 , comparable to the 6-mode UPP described in the previous section.

3.2.7 8-mode Automated Calibration

Following the same principles with which the 6-mode UPPs were calibrated, I used the exact same setup with an 8-mode device as well. This device is fabricated using the same methods as the showcased 6-mode from the previous section, only differing in complexity as the increased number of modes also implies an increased number of phase shifters (64) and an increased amount of unitary matrices to include in the training dataset (2000). The waveguides are optimized for operation at a wavelength of 925 nm.

The average amplitude fidelity measured over 200 Haar random unitaries after the calibration was 0.9948 ± 0.002 . We may be able to compare the performance of this UPP with a device with a similar size in other fabrication platforms, such as the 8-mode UPP by Ref. [9]. For this device there is no strict comparison, as a figure of merit that is exactly equivalent to the amplitude fidelity averaged over a large number of randomly picked unitary matrices. However, a value of 0.962 was given for the performance of this

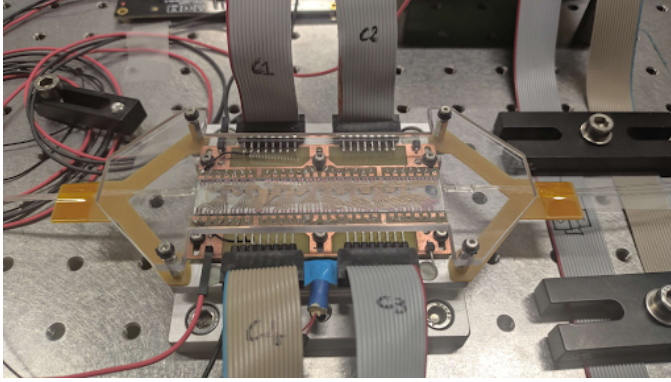


Figure 3.16: *The 8-mode UPP.*

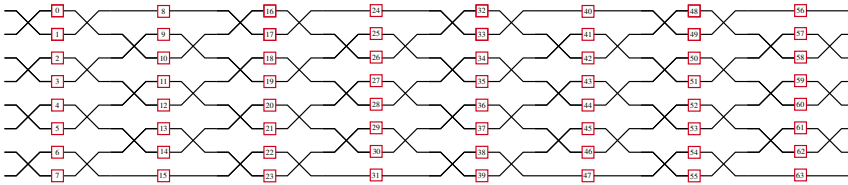


Figure 3.17: *The waveguide structure of the 8-mode UPP; the location of each phase shifter is highlighted and numbered. Every crossing corresponds to a single 3 dB coupler.*

UPP over a specific set of matrices (so-called «six dimensional X^i gates»)). Comparing this value with the obtained fidelity of the 8-mode showcased in this work highlights the quality of the calibration and implementation, as well as the flexibility of the method as it can maintain this advantage as the number of modes increases.

It is worth pointing out that nothing has changed in the setup compared to Figure 3.13, but the results are still excellent compared to the state of the art, which is a demonstration of the wide applicability and flexibility of this method.

3.2.8 12-mode Automated Calibration

Moving up in complexity, by far the largest and most complex UPP that has been calibrated using the automated setup is a 12-mode UPP (pictured in Figure 3.19). The optical layout of the device, like all the other devices described in this work, was directly written using FLWW inside of a glass slab; the losses were measured to be 3.5 dB on average for all the modes. This UPP was designed for applications involving light at a wavelength of

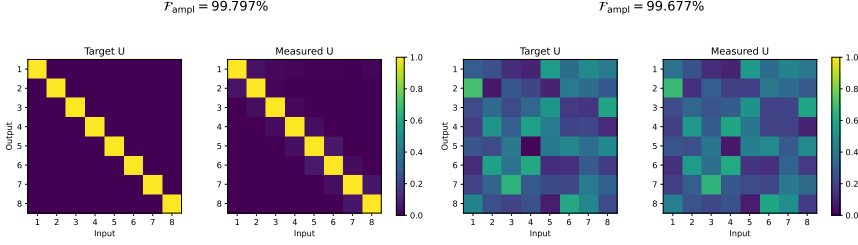


Figure 3.18: Example of performance of the 8-mode UPP over the identity matrix (left) and a randomly selected Haar random matrix from the test set (right).

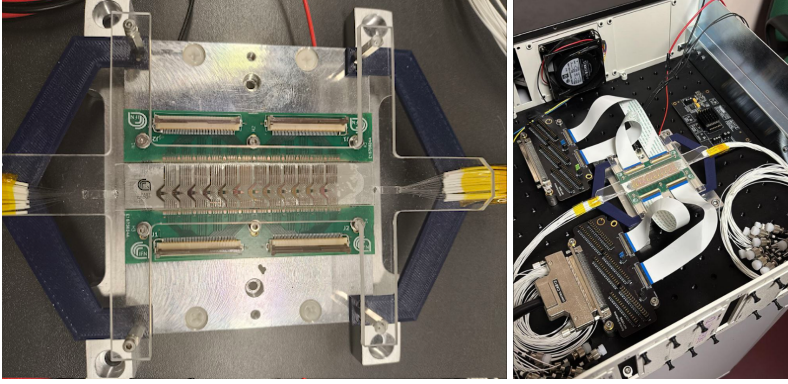


Figure 3.19: The 12-mode UPP used for this example. On the left, a picture of the UPP seen from the top. On the right, the UPP slotted within a calibration and operation setup, compact enough to fit inside a drawer on a 70 cm by 94 cm server rack.

925 nm.

This UPP features 144 resistive heating elements, fabricated in chromium-copper, with a distribution of resistances centered around the average value of 315Ω ; the design of the unit cell is the push-pull configuration (Figure 1.8b). A pair of custom PCBs have been glued to either side of the device, providing the connection for all the resistances. Each PCB is connected to an interposer board, which in turn is connected to a separate Qontrol BP12 control board where 11 Q8iv microcontrollers are slotted in each.

A single matrix measurement for this device needs to fill 144 elements of a 12-by-12 matrix. A setup like the one shown in Figure 3.13 can be operated quickly, but because of the number of times that the fiber switches have to swap port - at least 144 times for either the input or the output - even operating the fiber switches at their fastest possible speed of 300 ms according to the datasheet, the minimum measurement time amounts to

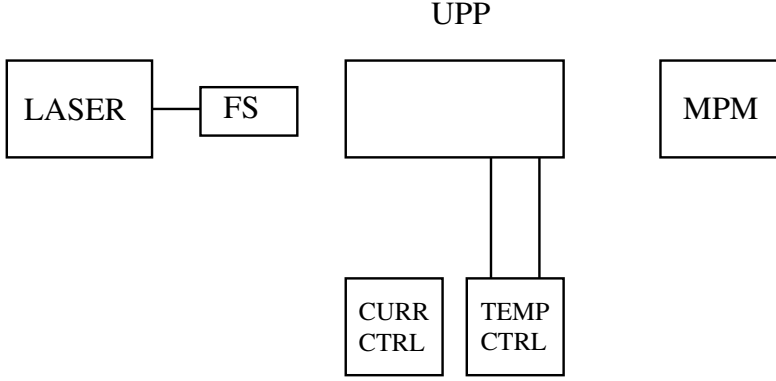


Figure 3.20: The parallelized automated setup used for the calibration of the 12-mode UPP.

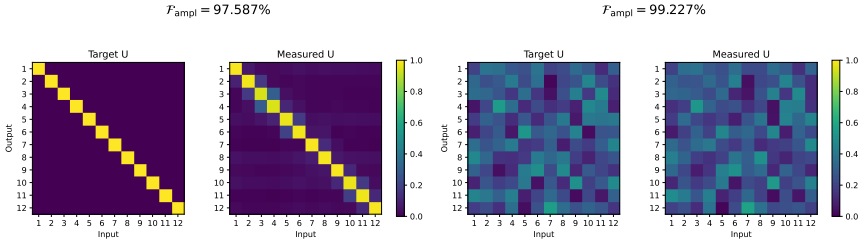


Figure 3.21: Example of performance of the 12-mode UPP over the identity matrix (left) and a randomly selected Haar random matrix from the test set (right).

43.2 s. In the laboratory, during the actual experiment, the time needed to measure a full matrix in that setup was higher than one minute.

For this reason, I expanded the setup by adding a multichannel optical power meter (a Santec OPM150); this power meter features 24 optical channels that can measure the power at the same time. This pushed the time needed in the laboratory to measure a full matrix down to under 10 s, including the time needed to actuate all of the electrical controls as well. In general, the introduction of a multichannel optical power meter reduces the time complexity of the single matrix measurement from $O(N^2)$ to $O(N)$, at the cost of a more expensive device in the experimental setup.

With this setup I was able to gather a dataset of 10^4 unitary matrices to be used to train the digital twin of the UPP. After training, I tested the performance of the calibration, obtaining a value of amplitude fidelity over 1000 Haar random matrices of 0.9907 ± 0.0063 (see Figure 3.30). This value is vastly superior to the closest device found in the literature of the same size, a 12-mode UPP fabricated with silicon nitride [36], which features an aver-

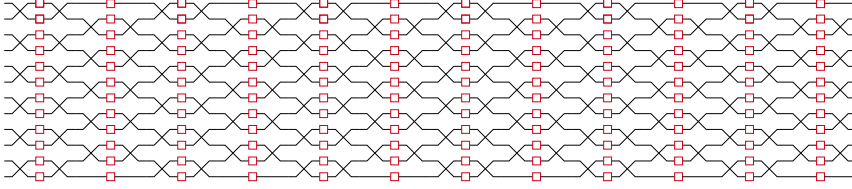


Figure 3.22: The waveguide structure of the 12-mode UPP; the location of each phase shifter is highlighted. Every crossing corresponds to a single 3 dB coupler.

age amplitude fidelity over 1000 Haar random matrices of 0.904 ± 0.0024 .

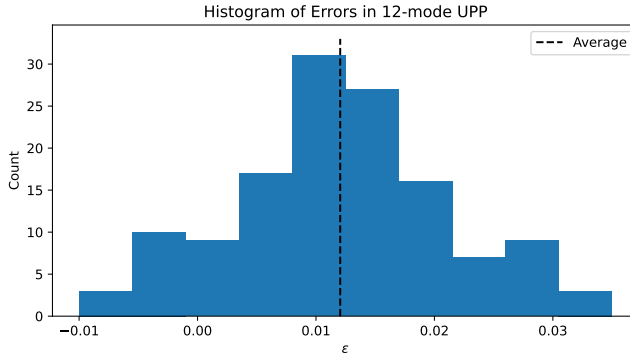


Figure 3.23: Histogram of deviations from 3 dB behavior for each coupler in the network. The average, marked by the dashed line, is 0.012.

It is important to note here that a similar 12-mode device was also calibrated using the state of the art method reported by Ref. [25], which is an analogue «digital twin» gray-box method, to obtain a fidelity of 0.9977. The discrepancy between the fidelity obtained with this work compared to the state of the art is likely to be found in the way the matrices are implemented after the calibration: in fact, where my method of implementing the matrices follows Clements' decomposition algorithm - suitably adapted to the structure of the 12-mode, to include corrections due to deviation from ideality of the reflectivities - the reported methodology applies a global optimization method to retrieve the phases that reportedly performs better when the reflectivities deviate from the ideal value. The reflectivities can actually be retrieved using the digital twin method, since the optimization algorithm also includes the coupler errors ε as a parameter: the histogram of errors for the MZIs in this device is shown in Figure 3.23. In this device, the reflectivities are skewed consistently towards higher values than 0.5.

Despite the application of a different method for the retrieval of the

phases (and thus, the solutions to Equation 2.9), these results still among the best for devices of equal size. These results are another proof of the power and flexibility of the automated calibration method, especially as the size of the UPP increases.

3.3 Vacuum-Enhanced Calibration of UPPs

3.3.1 Motivation

One of the main sources of nonidealities in our UPPs is thermal phase shifting. In fact, despite the mitigation strategies used in FLWW UPPs to reduce the amount of heat transfer between adjacent waveguides such as deep trenches and bridge structures [12], thermal cross-talk remains significant.

The calibration process, including the machine-learning methods described in the previous section, is built to handle and compensate the thermal cross-talk. Despite this, thermal cross-talk still causes issues that the calibration alone cannot solve.

Firstly, when actuating a single MZI, thermal cross-talk induces a phase shift on both arms of the interferometer. The output distribution is governed not by the absolute phase shift, but rather the relative phase shift between the two arms of the MZI $\Delta\phi$. Therefore, if P is the power dissipated on only one arm and γ the fraction of power that is transferred to the other arm, we have

$$\Delta\phi \propto \frac{2\pi}{\lambda} P(1 - \gamma). \quad (3.8)$$

This means that significantly more power must be dissipated in order to achieve a given $\Delta\phi$ for γ close to 1. This implies large power dissipation when actuating the whole UPP at once. State of the art UPP devices have a power dissipation close to 385 mW to achieve a phase shift of 2π [37] at a wavelength of 1550 nm; in our best performing UPPs, this value is between 60 mW and 70 mW. As the number of modes increases, the total power dissipation on our processor is bound to increase as well; the number of phase shifters in the network scales quadratically with the number of modes, and therefore so does the dissipated power all else being equal. For this reason, there is interest in researching methods to reduce - not just compensate, as the calibration does - the effects of thermal cross-talk in UPP devices.

Secondly, compensating for thermal cross-talk implies measuring every single term of the cross-talk matrix α according to the model described in Section 2.3. Both ends of the spectrum of calibration processes, from white-box to black-box, benefit from a reduction in the amount of parameters

that must be measured. If thermal cross-talk is reduced to zero, it can be shown that the cross-talk matrix α becomes a sparse matrix. For a large number of modes N , this can significantly speed up the calibration process by taking the number of parameters involved from $O(N^4)$ to $O(N^2)$.

Clearly, reduced thermal transfer in the UPP comes with the downside of slowing the heat dissipation process. It is to be expected that, if cross-talk is reduced, then so is the speed at which the device can be operated. So the question of speeding up the calibration must be evaluated against the slowing down of the operation (and the calibration itself, since it depends on the speed of the reconfiguration as well).

3.3.2 Method and Setup

Cross-talk and Power Dissipation Enhancement

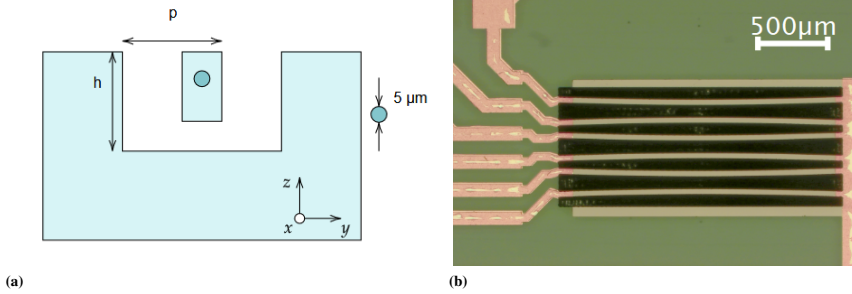


Figure 3.24: The typical structure of the micromachined glass in our UPPs. Figure (a) shows a section of the bridge-and-trench structure in the plane orthogonal to light propagation; the darker dot is the waveguide core. Figure (b) shows a microscope picture from the top of a column of bridges in the UPP. The pink areas are the copper-chromium depositions, while the light gray areas are the portions where only chromium is present; this concentrates the resistance of the metal deposition on the bridge only, since chromium is more resistive than copper.

Since our state of the art UPPs feature very streamlined and minimal structures, we can safely assume that the main vector of heat transfer between adjacent waveguides must be conductive heat transfer through air.

For this reason, I introduced a vacuum chamber into the calibration setup in order to explore the effects of vacuum on the thermal cross-talk. A scheme of the setup is shown in Figure 3.25. The vacuum chamber is a steel box with several feedthrough outlets to transfer light and electrical signals in and out of the vacuum environment. The pressure in the vacuum chamber is controlled by an oil pump and monitored via a digital pressure

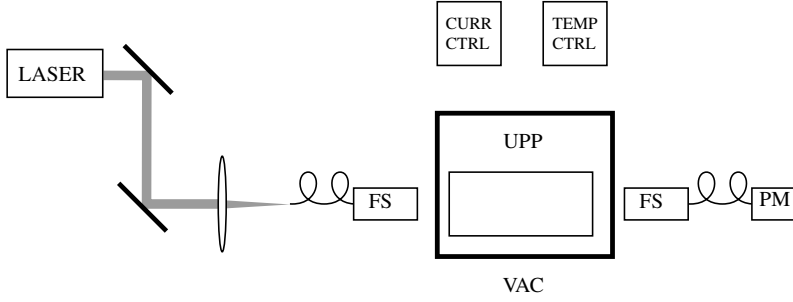


Figure 3.25: *The setup used for the vacuum experiments.*

sensor; the chamber can reach pressures in the order of 3×10^{-2} mbar while the pump is running. This value for the vacuum condition is used throughout the rest of the work.

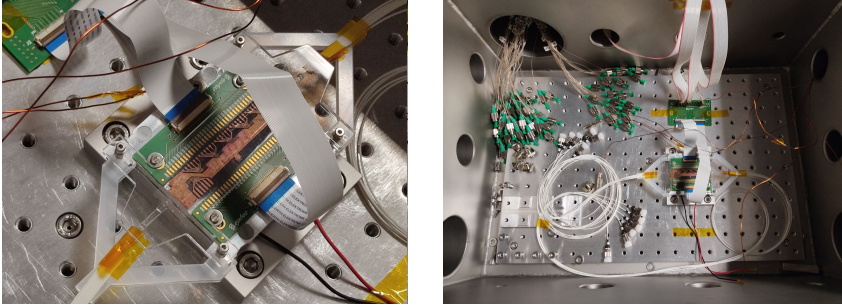


Figure 3.26: *A closeup of the UPP used for the vacuum experiments (left) and the inside of the vacuum chamber during assembly (right).*

The UPP used for this experiment is a 6-mode device featuring a Clements layout with push-pull unit cells. The device is optimized for guiding light at a wavelength of 1550 nm; I measured losses of around 2.6 dB on average. It comprises 36 resistive elements fabricated with chromium-copper deposition, however during the calibration process only 30 elements were controlled. Pictures of the UPP and the vacuum chamber setup are shown in Figure 3.26.

Initially, I measured the value of $P_{2\pi}$, the power required to achieve a 2π phase shift. At standard pressure, this value is 68 mW averaged over all phase shifters; in vacuum, this value is 14 mW due to the suppression of thermal cross-talk.

I then performed the automated calibration of this device at both pressure conditions. To obtain a figure of merit for the cross-talk that could be compared between air and vacuum, I evaluated $\alpha_{ii}/\alpha_{ii+1}$, the ratio between

the phase shift induced by a shifter on itself, and by the same shifter on a first neighbor. In air this value is 13.2% while in vacuum this value is 1.8%; a comparison between the cross-talk matrices in air and vacuum is shown in Figure 3.27.

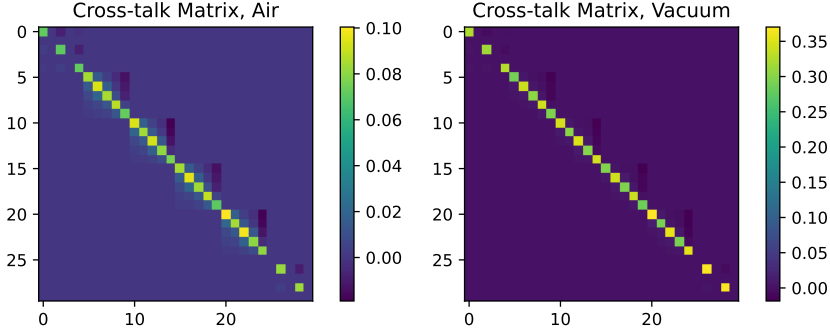


Figure 3.27: The crosstalk matrix of the 6-mode UPP in air (left) versus vacuum (right). The diagonal elements are larger in the case of vacuum, and the ratio between diagonal and off-diagonal elements is smaller.

As predicted, therefore, the introduction of vacuum in our UPPs does in fact reduce the power dissipation and the thermal cross-talk. This is a first, encouraging result that points us in the right direction when it comes to reducing this effect in FLWW UPPs.

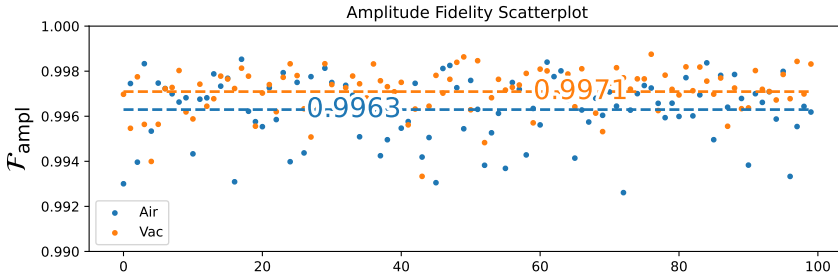


Figure 3.28: A comparison between the fidelity in air and the fidelity in vacuum, using the same optimization algorithm and the same model for both.

As shown in Figure 3.28, the value of the average amplitude fidelity in both environments was essentially unchanged, with a slight improvement in vacuum condition. In air, the amplitude fidelity was 0.9961 ± 0.0014 , while in vacuum it was 0.9971 ± 0.0010 . The observation that the two values are very similar suggests that the process of calibration is able to compensate for cross-talk in both conditions equally as efficiently.

One might imagine that the advantage of reduced cross-talk is traded off with having to operate the UPP in a vacuum environment. On the one hand, this exploration of the behavior in vacuum of the UPP is useful for applications where vacuum is already present, such as operation in space. On the other hand, it is possible to design vacuum sealed packaging for the UPP devices exploiting the same photolithographic processes that are used to fabricate the heating elements: an alloy of chromium, gold and nickel can be deposited on top of the passivated device in a vacuum environment, followed by a glass cap to apply a vacuum tight seal [38]. In this way, all the advantages of vacuum can be exploited in FLWW UPPs without having to accommodate for bulky vacuum chamber setups.

Time To Calibration Analysis

As mentioned briefly, with a reduction in cross-talk comes a heavy reduction of the number of parameters that a cross-talk matrix would need to take into account. This has the added effect of removing complexity from the calibration equation (Equation 2.9) which greatly simplifies and speeds up all calibration methods across the board.

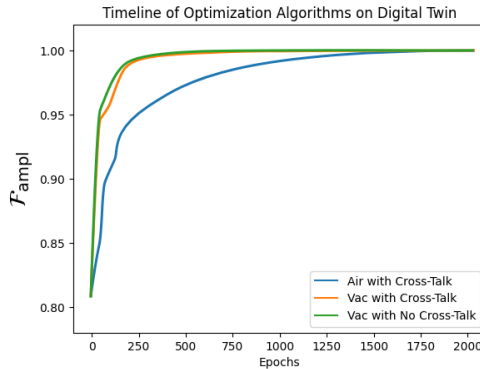


Figure 3.29: Comparison between the amplitude fidelity versus epoch of the optimization models for the digital twin. While all fidelities eventually converge to unit value, the model in vacuum requires less epochs compared to air. If the model is made simpler, such as the case of a model that does not include cross talk, the model converges even faster.

In fact, while calibrating the 6-mode UPP in both pressure environments, I observed that the calibration in vacuum converged to a stable value of fidelity in a smaller number of iterations compared to the same calibration performed in air. This can be explained by the fact that the algorithm was able to focus on a smaller number of parameters in order to maximize

the performance over the whole dataset, since most of the free parameters are very small and the cross-talk matrix is similar to a sparse matrix. A comparison between the number of iterations, or «epochs», of the machine learning algorithm in the two conditions is shown in Figure 3.29. As is evident from this graph, the machine learning algorithm requires less time to converge in the case of vacuum compared to air; we can say that it has a lower «time to calibration» (TTC).

A faster calibration is of course an additional advantage to the introduction of vacuum in FLWW UPPs, as the computational power and time needed to calibrate larger UPPs grows very quickly.

I decided to check whether this advantage could be pushed even further by simply adapting the machine learning algorithm to model the UPP as an ideal device with no cross-talk whatsoever. This forces the algorithm to optimize only a sparse set of elements in the matrix, leading to another shortening of the TTC. A comparison between the epochs in the case of a full model versus the simplified, no-cross-talk model is shown also in Figure 3.29.

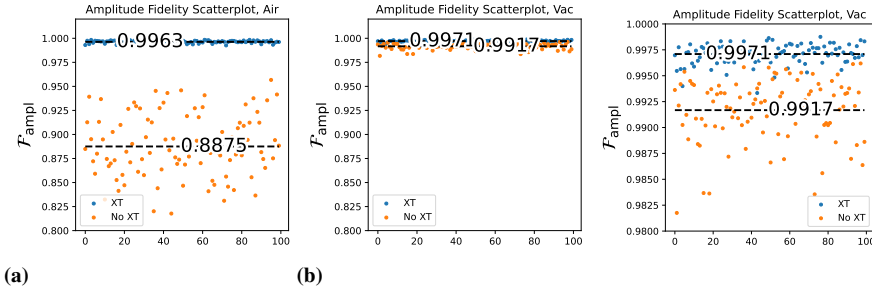


Figure 3.30: A comparison between the performance of the calibration using the two models, one where the cross-talk is fitted by the model versus one where it isn't. Figure 3.30a shows the change in performance of the calibration in air, while Figure 3.30b shows the change in performance in vacuum; a closeup on the overlap region for Figure 3.30b is provided on the right for better visibility.

Clearly, this begs the question of whether removing the cross-talk from the model entirely is legitimate, and whether this action hurts the performance of the calibration. To verify this fact, I tested both models in air and in vacuum. The results are reported in Figure 3.30: in air, moving from the larger model to the smaller model brings the amplitude fidelity down from 0.9961 to 0.8875 which is unacceptably low; in vacuum however, the fidelity goes from 0.9971 to 0.9917 which is still a very good result considering the fact that cross-talk is non-negligible in this UPP even in the

vacuum environment.

This suggests that if the cross-talk can be removed when moving to the vacuum, then it is possible to obtain a calibration that is fast, accurate and less computationally expensive, and further lower the TTC.

Time To Matrix Analysis

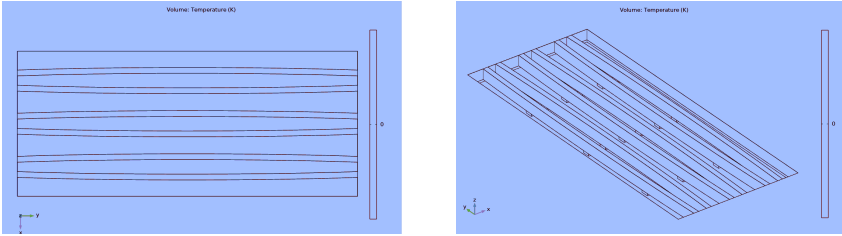


Figure 3.31: *The Comsol simulation of the column of the UPP, while every heater is turned off, from the top (left) and in isometric view (right). There are a series of pillars that support the bridge structure. The scale is measured in terms of temperature difference with respect to the equilibrium.*

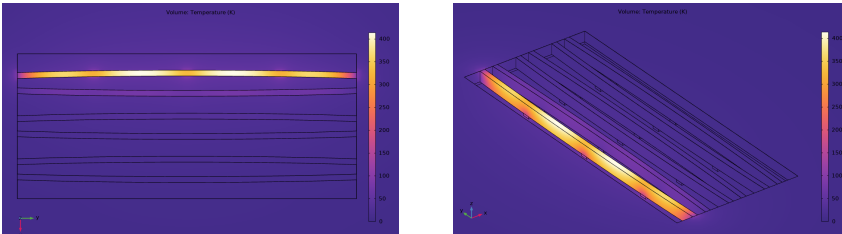


Figure 3.32: *The Comsol simulation of the column of the UPP in air, while the top heater is dissipating heat, from the top (left) and in isometric view (right). Dips in temperature can be seen where the pillars are transferring heat from the bridge to the bottom of the structure. The scale is measured in terms of temperature difference with respect to the equilibrium.*

The question remains whether the advantage of speeding up the calibration process is a trade-off considering that the device should be significantly slower in vacuum. To test the time response and explore the effects of vacuum more in detail, I simulated a single column of the UPP using the Comsol MultiPhysics simulation suite. Pictures of the simulated column are shown in Figures 3.31 and 3.32.

In our UPPs, resistive elements are spaced apart 40 to 80 μm vertically and 5 to 6 mm horizontally. As a result, the heating elements are tightly packed in the vertical direction (see for example Figure ??) but essentially

thermally isolated from all others horizontally. It is therefore possible to simulate the behavior of the whole device by just focusing on modeling and simulating a single column, with the assumption that all columns would then act independently.

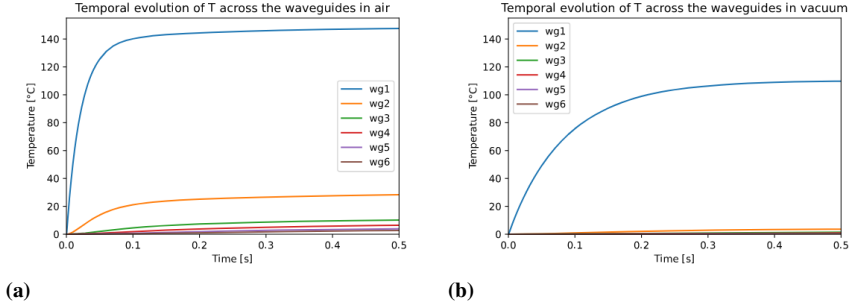


Figure 3.33: The simulated temperature profiles over time (a) in air and (b) in vacuum, for each of the six waveguides under consideration.

I performed two simulations. First, I set the simulation up to dissipate power on a single heating element and simulated the time evolution of the whole column, keeping track of the average temperature change over each waveguide. The typical behavior of the temperature, as shown in Figure 3.33, is of course reminiscent of the step response of a linear system, with a slight delay due to the fact that heat needs to travel either in the air or in the glass before it can reach the waveguide. Based on this observation, I evaluated a time constant τ for each temperature profile by evaluating the time interval the curve takes to go from 10% to 90% of its asymptotic value.

The simulation shows that the time constant for a single shifter is 67 ms in air, and 192 ms in vacuum. This is of course to be expected, since vacuum significantly slows down the conduction of heat in the UPP as the energy must travel through the glass structure instead of air. Figure 3.34 the time constants for all six waveguides in the simulated column.

It is important to remark here that while very large τ are obtained here for this simulation, they do not necessarily represent the time needed for the whole column to reach thermal equilibrium during regular operations. Moreover, since the cross-talk between the first and last waveguide is extremely small (of the order 10^{-5} even in air, essentially zero for vacuum), it's difficult to extract a meaningful value for the time constant, hence why it appears that the last waveguide thermalizes faster than the second-to-last in Figure 3.34b.

For this reason, I set the simulation to dissipate power over all six heat-

ing elements at the same time, tracking the change over time for all waveguides as before. This experiment runs closer to the actual behavior of the UPP when it implements an actual matrix. The values of the time constants are reported in Figure 3.34. We may notice that while for the single-element operation the slowdown from air to vacuum is almost a factor 4, for the whole-column the slowdown from air to vacuum is essentially not present.

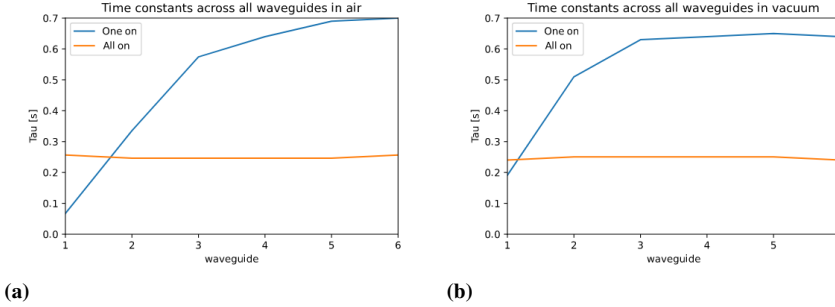


Figure 3.34: The values of the time constants for each waveguide while actuating only one resistive element versus actuating all phase shifters, (a) in air and (b) in vacuum.

This is a counterintuitive result, which however can be explained with a simple model. We can assume that in the second experiment, when all heaters are actuated at once, the temperature profile of a single waveguide is given by a superposition of the temperature profiles due to each of the other heating elements:

$$T_0(t) = \sum_j T_{\text{single},j}(t), \quad (3.9)$$

where $T_{\text{single},j}(t)$ is the profile corresponding to the j -th waveguide when only the first element is actuated. To provide a qualitative model for the temporal dynamics of the temperature, we may assume that the temperature profiles follow a simple saturating exponential behavior:

$$T_{\text{single},j}(t) = A_j \left(1 - e^{-\frac{t}{\tau_j}} \right). \quad (3.10)$$

We may also assume that the sum of the profiles is an exponential as well; this is not strictly true, but it is a good enough approximation that simplifies calculations in what follows. We have:

$$T_0(t) = A_{\text{tot}} \left(1 - e^{-\frac{t}{\tau_{\text{tot}}}} \right). \quad (3.11)$$

From this equation we may notice that we can also extract an estimate for the value of the time constant. By evaluating the derivative at the origin on both sides of Equation 3.9 we obtain:

$$\frac{1}{\tau_{\text{tot}}} = \frac{1}{A_{\text{tot}}} \sum_j A_j \frac{1}{\tau_{\text{tot}}} \quad (3.12)$$

and therefore

$$\tau_{\text{tot}} = A_{\text{tot}} \left(\sum_j A_j \frac{1}{\tau_{\text{tot}}} \right)^{-1} \quad (3.13)$$

the ratio between A_0 and A_1 is much smaller in air (4.93) compared to vacuum (56.2). Therefore, the linear combination in Equation 3.13 for air contains many small time constants that are summed together with comparable weights; for vacuum, the linear combination contains many large constants but the weights are small except for A_0 . These two effects compensate, in such a way that we obtain a similar value for vacuum and air of τ_{tot} .

In other words, the existence of non-negligible cross-talk makes it so that the whole column needs to thermalize before equilibrium is reached, which can be a slow process. On the other hand, if cross-talk is negligible or severely reduced, each shifter can thermalize independently. This trade-off between speed and independence significantly dampens the slowing effect of vacuum compared to air when the whole UPP needs to be actuated at once. This time is an estimate of the amount of time that a UPP needs to implement a full matrix, which we might call «time to matrix» (TTM).

This is very beneficial for the vacuum calibration of UPPs, since the reduced effect of cross-talk does not significantly alter the TTM, while maintaining the advantage of allowing faster calibration.

3.4 Measurement of Complex Matrix Elements

3.4.1 Motivation

As mentioned in the previous sections and chapters, two metrics of performance are typically given for UPPs in the literature: the fidelity (Equation 1.26) and the amplitude fidelity (Equation 1.27). The main reason for this split is that works that focus on the theoretical aspects of the UPP have the whole unitary matrix at their disposal, including the angles of the complex elements, while works that focus on the physical aspects of the UPP are

usually only able to measure the amplitudes of these elements. The angles of the complex element of the matrix, in fact, are tied to the relative phase between different modes of the UPP and are generally much harder to measure in a practical setting.

Several setups to measure the relative phase between multiple coherent modes exist in the literature [39–41]. Most of these methods involve the exploration of the angles of the unitary matrix by, in mathematical terms, analyzing the behavior of the implemented matrix by studying separately many of its 2-by-2 submatrices. Alternatively, other methods employ non-classical states of light to explore essentially the same behavior [42, 43].

These methods often involve expensive measurement setups or specify precise requirements of the state of light at the input, and none of them can easily and cheaply be adapted to the fast calibration setups described in the preceding sections. The aim of this part of my research work was to find a new method to obtain the measurement of the fidelity that is simple to use in practice and compatible with the experimental setups already described for FLWW UPPs.

3.4.2 UPPs and Real Bordered Matrices

It is important to give some theoretical background with regards to the measurement of phases in UPPs. In general, the setup in which an UPP is used will introduce additional transformations on the state of light both before and after the device. These transformations come in the form of losses and phase shifts; in other words, while the UPP may be implementing a unitary transformation U , the whole setup from source to detector will implement a transformation M described by

$$M = GUG' \quad (3.14)$$

where G, G' can in principle be any pair of matrices. If the setup does not allow any mode to mix before or after the UPP, then G and G' are diagonal; if losses are present, then G and G' have elements with moduli less than 1. As explained previously, Sinkhorn's Algorithm (Algorithm 3) normalizes away the effect of losses. However, the diagonal elements of G and G' might still contain a phase term, which represents the additional phase shift coming from other optical elements in the setup (fiber arrays before and after the UPP, light traveling in free space, waveplates, fiber switches...). These terms are essentially arbitrary and do not contain any meaningful information from the point of view of actual UPP applications.

It is necessary to find a standard way in which to remove these terms and obtain a matrix that represents the implemented matrix U in a physi-

cally meaningful way. As it turns out, one can choose the «real bordered representation» of M to solve this problem.

Any unitary matrix $M \in U(N)$ can be reduced to its real bordered representation R by multiplying it before and after by two diagonal matrices P and P'

$$R = PMP' \quad (3.15)$$

in such a way that the first row and the first column of R are all real numbers.

It can be shown that the real bordered representation of a matrix is unique, preserves unitarity, and that taking the real bordered representation is an equivalence relation on the space $U(N)$. Moreover, any experiment with photon number states that are transformed by some unitary matrix M yields the same results as its real bordered representation R [41].

In other words, taking the real bordered representation of a matrix divides the space of unitary matrices into equivalence classes where any experiment within that class will yield the same result. For this reason, angle measurements don't retrieve the full matrix M , but only the real bordered representation R .

The process to obtain R from M consists of these steps: let $M \in U(N)$ be a unitary matrix. This matrix's elements can be written out more explicitly as

$$M_{ij} = |M_{ij}|e^{ia_{ij}}. \quad (3.16)$$

Then let R be the matrix whose elements are written as

$$R_{ij} = |M_{ij}|e^{ib_{ij}} \quad (3.17)$$

where

$$b_{ij} = a_{ij} - a_{1j} - a_{i1} + a_{11}. \quad (3.18)$$

Notice that if i or j is equal to 1, then b_{ij} is zero and R_{ij} is real. Therefore, R is the real bordered representation of M .

3.4.3 Method and Setup

In order to measure the relative angle between at least two modes by measuring the optical power distribution, it is necessary to apply a method that employs interference. Our UPPs feature fiber arrays at both input and output, and therefore free-space interferometry is difficult to achieve. The ideal experimental setup for this task should be able to keep the UPP as-is without applying any modifications compared to the previously described setups, and on top of that, we would like to perform minimal edits to the

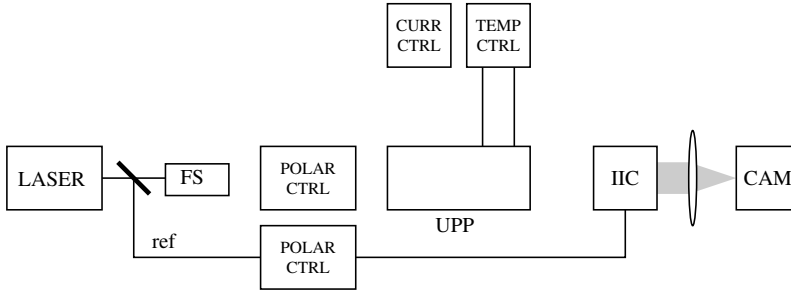


Figure 3.35: The setup used to measure the angles of the complex elements of the implemented matrix. A portion of the laser light is divided into a reference branch and a UPP branch; a set of polarization controls is added on both branches of the setup. At the output of the UPP, the IIC is added together with a cylindrical lens and a camera.

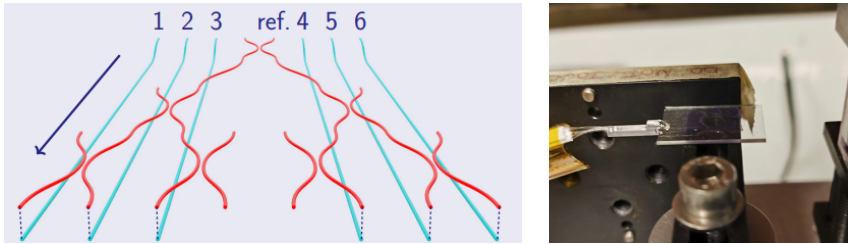


Figure 3.36: A schematic representation of the waveguide design of the IIC (left) and a picture of the IIC in the setup (right).

automated calibration setup in such a way that the calibration methods described in the previous sections can still be performed.

For this reason, I introduced a new setup that simply extends the previously established ones by adding an additional photonic chip, a cylindrical lens, a camera and a set of paddles to control the polarization state at the input of the UPP. The design of the extension chip (which will be called Interferometric Interface Chip from here on out, IIC), the reason for polarization control and the general methodology of the experiment will be described in the remainder of the section.

The IIC is a passive device fabricated through FLWW, the same method used to fabricate the MZI network in the UPP. It consists of two layers (Figure 3.36): one layer (the «beam splitter layer») comprises a series of 3 dB couplers, two input ports and several output ports; the other layer is a set of straight waveguides. Each output port of the straight waveguide layer is paired with a corresponding one in the other layer, forming a series of Young interferometers.

Light from the laser source is split off into two branches: a fraction goes

to the fiber switch and thus to the UPP; another fraction goes to the beam splitter layer of the IIC. Each output of the UPP is connected to a straight waveguide of the IIC.

For this experiment, we employed an IIC with 8 input ports (6 inputs for the top layer and 2 for the beam splitter layer as detailed in Figure 3.36, although only 1 input is connectorized and actually used for the beam splitter layer), which can measure the simultaneous relative angle between up to 6 modes. The layers are spaced $200\text{ }\mu\text{m}$ vertically while each mode is spaced $127\text{ }\mu\text{m}$ apart from the other. At the input of the device, the input of the beam splitter layer is made to line up with the input of the straight waveguides so that the distribution of the modes at the input can match common fiber array layouts; in order to easily couple the light from the UPP to the IIC, the input of the IIC has been pig-tailed with a standard 8-mode fiber array.

The general idea of the experiment is to retrieve the far-field interferometry profile of the modes, pairing up each output from the straight waveguide layer with each output from the beam splitter layer. In fact, both the output of the straight waveguide layer and the beam splitter layer can be considered a series of point-like sources. The output of the IIC impinges on a cylindrical lens, which focuses the light onto a camera; the camera then displays the far-field interferometric profile of the field.

We inject light from the laser source into the beam splitter layer input of the IIC, and we inject the output of the UPP into the straight waveguide layer. An example of the power distribution captured by the camera is shown in Figure 3.39; here we can clearly see the interferometric fringes coming from the pairing of the two layers.

According to the theory of Young's double-slit experiment, when two point-like coherent sources of electromagnetic waves impinge on a screen at a distance L much larger than their separation d (Figure 3.37), we can calculate the resulting interference profile by applying the definition of intensity

$$I = cn \langle \|E_{\text{tot}}\|^2 \rangle \quad (3.19)$$

where $n \simeq 1$ is the refractive index of air, c is the speed of light, and the brackets denote time average over the typical oscillation time of the field. The total field is given by the superposition of the field coming from the

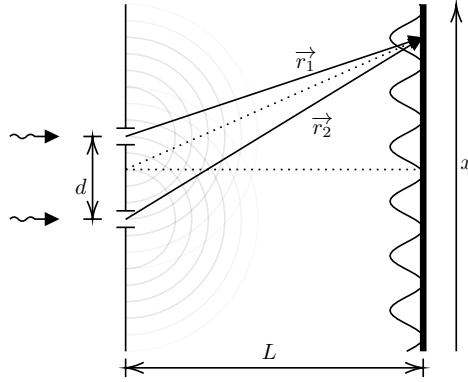


Figure 3.37: A scheme representing the typical analysis of Young's double-slit experiment for waves. The two incoming coherent light beams (on the left) impinge on a pair of point-like slits separated by a distance d , impinging on a screen at distance $L \gg d$ (on the right). The x axis marks the position on the screen, with the origin $x = 0$ corresponding to the middle point between the slits.

two point-like sources, in the far-field limit $L \gg d$:

$$E_1 = \vec{J}_1 e^{i\omega t - ik \cdot \vec{r}_1 + i\Delta\xi} = \begin{bmatrix} J_{1x} \\ J_{1y} \end{bmatrix} e^{i\omega t - ik \cdot \vec{r}_1 + i\Delta\xi}$$

$$E_2 = \vec{J}_2 e^{i\omega t - ik \cdot \vec{r}_2} = \begin{bmatrix} J_{2x} \\ J_{2y} \end{bmatrix} e^{i\omega t - ik \cdot \vec{r}_2}$$

where $\Delta\xi$ is the relative phase between the coherent point-like sources, J_i the amplitude and polarization state of the i -th field written as a Jones vector. In this representation of the electric field, the intensity is simply equal to

$$I = \frac{1}{2} cn \left\| \vec{J} \right\|^2 = \frac{1}{2} cn (|J_x|^2 + |J_y|^2). \quad (3.20)$$

The far-field condition gives

$$\|\vec{r}_2 - \vec{r}_1\| \simeq \frac{d}{L} x = \Delta L \quad (3.21)$$

where x is the distance from the center of the screen and $\vec{r}_1, \vec{r}_2$ the vectors that connect each slit to the point on the screen at x (Figure 3.37). There-

fore:

$$\begin{aligned} E_{\text{tot}} &= \begin{bmatrix} J_{1x}e^{-ik \cdot r_1 + i\Delta\xi} + J_{2x}e^{-ik \cdot r_2} \\ J_{1y}e^{-ik \cdot r_1 + i\Delta\xi} + J_{2y}e^{-ik \cdot r_2} \end{bmatrix} e^{i\omega t} \\ &= \begin{bmatrix} J_{1x}e^{ik\Delta L + i\Delta\xi} + J_{2x} \\ J_{1y}e^{ik\Delta L + i\Delta\xi} + J_{2y} \end{bmatrix} e^{i\omega t - ik \cdot r_2} \end{aligned}$$

and hence by Equation 3.20:

$$\begin{aligned} I &= \frac{1}{2}c \left[|J_{1x}e^{ik\Delta L + i\Delta\xi} + J_{2x}|^2 + |J_{1y}e^{ik\Delta L + i\Delta\xi} + J_{2y}|^2 \right] \\ &= \frac{1}{2}c \left[|J_{1x}|^2 + |J_{2x}|^2 + |J_{1y}|^2 + |J_{2y}|^2 \right] + \\ &\quad + c \left[\text{Re}[(J_{1x}J_{2x} + J_{1y}J_{2y})e^{ik\Delta L + i\Delta\xi}] \right] \end{aligned}$$

In other words, if we write the term $J_{1x}J_{2x} + J_{1y}J_{2y}$ as a generic complex number $z = \rho e^{i\Delta\zeta}$ and group up similar terms:

$$I = I_0 + I_m \cos\left(\frac{kd}{L}x + \Delta\xi + \Delta\zeta\right). \quad (3.22)$$

Here, I_0 encapsulates all those terms that only depend on the intensities of each source while I_m is the amplitude of the intensity modulation due to interference.

From this analysis we notice a few things. On the one hand, the interference profile is a sinusoidal function of x , and there are several methods to retrieve the value of the parameters for such a function (Fourier transform, curve fitting, etc.). On the other hand, not only does the profile feature a phase term $\Delta\xi$ which is precisely the phase difference between the two sources, it also features an «extra angle» $\Delta\zeta$ term that depends on the polarization state of the two sources.

The phase difference $\Delta\xi$ is affected by the length of the fibers between UPP and IIC. However, the effect of the fiber length can be modeled as additional phase shifts that are outside the UPP, and therefore they do not affect the measurement of the real bordered matrix.

On the other hand, the extra angle term $\Delta\zeta$ depends on the polarization state at each output of the IIC. The main issue with this is that the polarization of the light that reaches the IIC depends on the polarization that enters the UPP, and this may depend on the input chosen to inject the light into. In other words, $\Delta\zeta$ cannot be modeled as extra phase terms and passing to the real bordered matrix does not compensate for them, because $\Delta\zeta$ would depend on both the chosen input and the chosen output.

$\Delta\zeta$ can be compensated in two ways:

1. by ensuring that the polarization state of light at the input of the UPP is the same for all inputs, or
2. by ensuring that the polarization state of light at the input of the IIC is the same for all inputs.

The two methods are equivalent. In both cases, the $\Delta\zeta$ stop being dependent on either the input (first case) or the output of the UPP (second case) and thus can be described as extra phase terms that are compensated away by passing to the real bordered representation.

In the experiment, I used a set of paddles to control the polarization state of light in the fibers; in order to add as little fiber as possible between the UPP and IIC, I decided to place the polarization control at the input of the UPP instead. In fact, in that region I only ever inject light into one fiber at a time and thus any phase instabilities coming from the long fiber windings would result in a variation of the global phase, which is easier to compensate for with data analysis.

Notably, as shown in previous experiments, the device's calibration is resistant to changes in the input polarization so a full recalibration of the UPP is not necessary.

It is worth pointing out that the stability of the polarization state is a crucial factor if the fiber arrays are not Polarization-Maintaining (PM). In fact, the setup features several meters of fiber overall, meaning that mechanical and thermal stability are especially important factors to keep the stability of the polarization during the experiment under control.

If this polarization state is controlled and stabilized, it is possible to retrieve at once all the $\Delta\xi$ for each output of the UPP at the same time. To prepare the polarization state, it is sufficient to have polarization controllers at the input of the UPP. The methodology is as follows:

1. Remove all light from the straight waveguide layer of the IIC.
2. Inject the split-off reference light from the coherent source into the beam splitter layer. Using the polarization controllers, select a polarization state for the beam splitter layer of the IIC (with the same process detailed in Section 3.2.4).
3. Remove all light from the beam splitter layer of the IIC.
4. Make the UPP implement the identity transformation ($U = I$), and inject light into the first input port of the UPP. Using the polarization controllers, select the same polarization state.

5. Make the UPP implement a switching matrix such that the second input is redirected towards the first output, and inject light into the second input port of the UPP. Using the polarization controllers, select the same polarization state.
6. Repeat the same process as above for all input ports.

This process ensures that the polarization of the light injected into any input port of the UPP is always the same. Because of this, the extra angles $\Delta\zeta$ now only depend on the output, and they can be treated as additional phase terms that will be removed by passing to the real bordered representation.

3.4.4 Experimental Validation

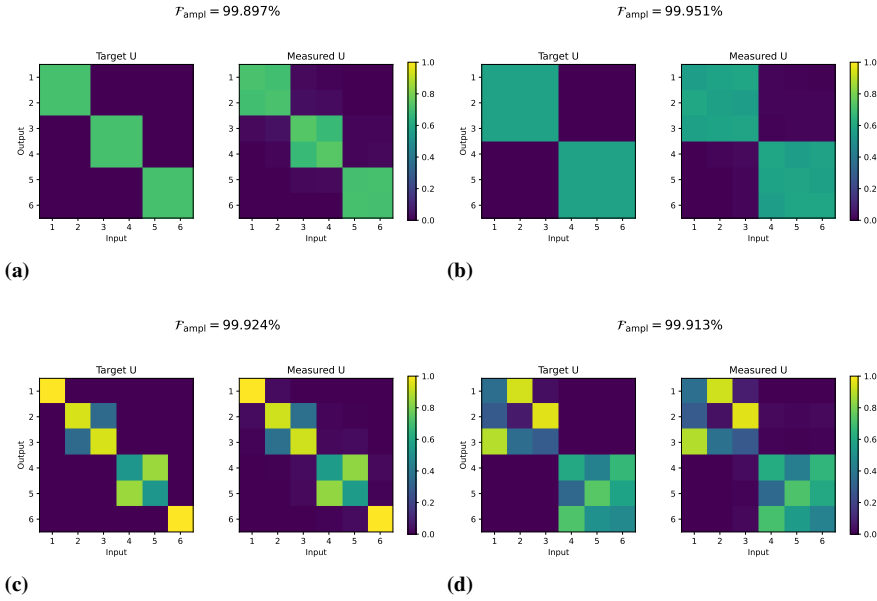


Figure 3.38: The measured amplitudes and (amplitude) fidelities of each of the four selected matrices. Matrix (a) and (b) feature uniform submatrices, while (c) and (d) feature random submatrices. In this work, only the top left submatrices have been measured.

In order to validate this methodology with an experiment, it is necessary to compare any measurement against a set of unitary matrices whose angles are easy to verify independently of the UPP.

From the theory of unitary matrices it is possible to prove an important

property of matrices in $U(2)$ and $U(3)$: their angles are completely determined by their moduli, up to conjugation and real bordered representation.

More in particular, given a generic 2-by-2 unitary matrix M , we can write it explicitly in terms of amplitudes A_{ij} and angles ϕ_{ij} :

$$M = \begin{bmatrix} A_{11}e^{i\phi_{11}} & A_{12}e^{i\phi_{12}} \\ A_{21}e^{i\phi_{21}} & A_{22}e^{i\phi_{22}} \end{bmatrix}. \quad (3.23)$$

Then, the real bordered representation of this matrix turns out to be

$$M_{\text{RB}} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22}e^{i(\phi_{22}+\phi_{11}-\phi_{12}-\phi_{21})} \end{bmatrix}. \quad (3.24)$$

We may call the angle of the second element of the second row the «characteristic angle» of M , $\phi_c = \phi_{22} + \phi_{11} - \phi_{12} - \phi_{21}$. This characteristic angle is completely determined by the amplitudes A_{ij} of the matrix, since by unitarity

$$M_{\text{RB}}^\dagger M_{\text{RB}} = I \implies A_{11}A_{12} + A_{21}A_{22}e^{i\phi_c} = 0 \quad (3.25)$$

Which specifies the angle completely (in particular, it must be an odd multiple of π). Similarly for a 3-by-3 matrix there are four characteristic angles:

$$M_{\text{RB}} = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22}e^{i\phi_{c,1}} & A_{23}e^{i\phi_{c,2}} \\ A_{31} & A_{32}e^{i\phi_{c,3}} & A_{33}e^{i\phi_{c,4}} \end{bmatrix} \quad (3.26)$$

and by unitarity we obtain similar equations that fully determine them from amplitudes alone.

Since we can already measure the moduli of the matrices with high precision in our setup, we can use this property to obtain a set of expected angles against which to verify the experiment. I selected four matrices out of $U(6)$, shown in Figure 3.38. These matrices feature full 2-by-2 and 3-by-3 submatrices, both uniform and random, whose angles can be retrieved from the optical power distribution measurement according to the theory.

I carried out and verified the experimental methodology using a 6-mode UPP, the same one as in Section 3.3. In order to control the polarization of multiple modes at the same time, a set of four single-fiber polarization control systems (Thorlabs FPC031) were used. Only four were available at the time of this experiment, so I placed one on the reference branch and three on the first three inputs of the UPP. This ensured that I could control the polarization of up to three modes plus the reference at the same time.

3.4. Measurement of Complex Matrix Elements

As detailed in the previous paragraphs, experimental validation of the method can be done with 2-by-2 and 3-by-3 submatrices. In this experiment, only the first three inputs and outputs of the UPP have been measured, effectively restricting the range of view to just the top left 3-by-3 submatrix.

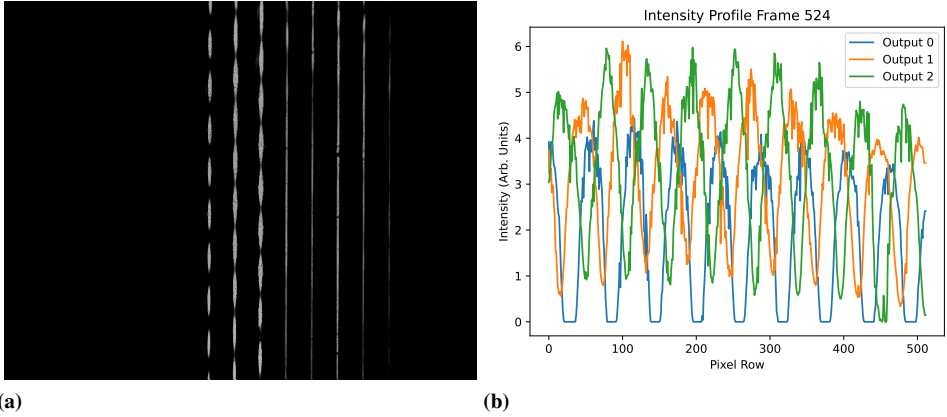


Figure 3.39: The intensity profiles as measured with the camera. Figure (a) shows the unedited picture from the camera, where the UPP is implementing a uniform 3-by-3 matrix. Figure (b) shows the raw profiles obtained from this image.

An example of an intensity profile obtained with the experimental setup is shown in Figure 3.39. I retrieve the intensity profile by integrating the values from each pixel on a narrow strip of pixels containing the whole interference figure for one output mode.

In order to obtain a good guess for the phase and the frequency with which to fit the sinusoidal function of the profile, I first pass each profile through a Discrete Fourier Transform (DFT) algorithm that gives an estimate for the frequency components and phase of the profiles. Then, I fit a sinusoidal curve using the frequency as a starting guess to retrieve the phase.

As can be seen from Figure 3.40, the DFT shows that tiny variations in the picked frequency induce large variations in the phase. This is one of the main sources of error in the method, as tiny deviations in fitting the sinusoidal function tend to induce large variations in the retrieved phase.

Several optical elements with significant optical length were placed on the UPP arm of the setup: the input and output fiber switches with their optical fibers, the UPP itself with its fiber arrays, the fiber array of the IIC. This resulted in a significant optical length imbalance between the split-off

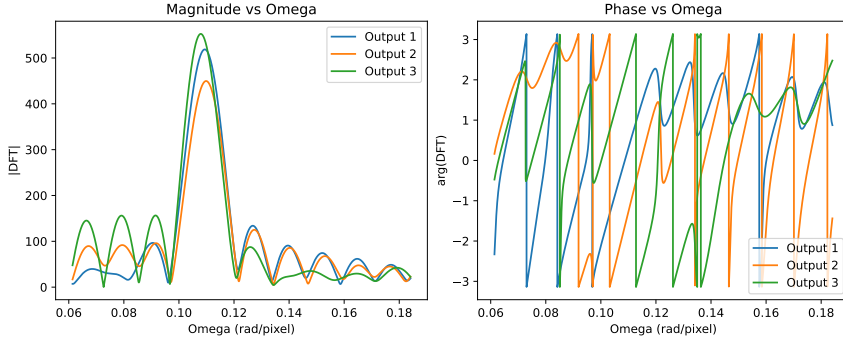


Figure 3.40: *The Discrete Fourier Transform applied to the intensity profiles. From these graphs we can see that the phase of the signal depends strongly on the frequency in the region of interest.*

coherent reference and the rest of the setup. This caused the interference fringes to be visibly unstable when displayed on the camera. For this reason I added a 3m fiber bracket (SM1550) to match the optical length of both arms. This step visibly improved the stability of the interference fringes on the camera, enabling me to begin the experiment.

In the experiment, I started by measuring the stability of the retrieved angles across the typical time of the measurement. I found this value to be 0.12 rad.

Next, I applied the method described in the previous section to retrieve the angles of each matrix in the set. In Tables 3.1 and 3.2 I report the values of the angles retrieved for the four matrices in the chosen set, adjusted for the real bordered representation, compared with the expected values from the theory.

Matrix	Ch. Angle
2-by-2 Uniform (Theory)	3.1419
2-by-2 Uniform (Measured)	3.0739
2-by-2 Random (Theory)	3.1419
2-by-2 Random (Measured)	3.2101

Table 3.1: *The measured values of the angles for the two 2-by-2 matrices implemented by the UPP.*

We can notice that most of the obtained angles deviate from the theory within the stability range. This constitutes a successful experimental validation of the method, which can then be applied in principle to any number of modes.

However, one of the angles for random 3-by-3 matrix does not match

3.4. Measurement of Complex Matrix Elements

Matrix	Ch. Angle 1	Ch. Angle 2	Ch. Angle 3	Ch. Angle 4
3-by-3 Uniform (Theory)	4.1888	2.0944	2.0944	4.1888
3-by-3 Uniform (Measured)	4.1488	2.1213	2.1326	4.1768
3-by-3 Random (Theory)	3.1984	3.0311	3.1379	6.167
3-by-3 Random (Measured)	0.3158	3.2086	3.2766	6.1224

Table 3.2: *The measured values of the angles for the two 3-by-3 matrices implemented by the UPP.*

with the theory. This mismatch can actually be explained by looking at the intensity profiles measured by the camera, or even by looking at the 3-by-3 random matrix itself (Figure 3.38d, top left submatrix) : that element has very low intensity compared to the others and compared to the noise, therefore it becomes difficult to retrieve its value when fitting the intensity profiles.

This exposes a criticality in the method: the intensity of the light at the output of the UPP needs to be high enough for the imaging system to be able to retrieve a good intensity profile. Fortunately, this is not a critical issue from the point of view of the fidelity: in fact, if the intensity of a specific element is low, then its weight in the fidelity calculation is also small, and the error in measuring the fidelity is attenuated away. This phenomenon is clearly illustrated by the calculated values of fidelity from these measurements: even though the experiment does not retrieve the expected angle, the fidelity is still very high overall.

Matrix	Measured $\mathcal{F}_{\text{ampl}}$	Measured $\mathcal{F}$
2-by-2 Uniform	0.9990	0.9987
2-by-2 Random	0.9992	0.9945
3-by-3 Uniform	0.9995	0.9993
3-by-3 Random	0.9991	0.9935

Table 3.3: *The measured amplitude fidelity $\mathcal{F}_{\text{ampl}}$ versus the value of fidelity $\mathcal{F}$ obtained with this method.*

After obtaining the values of the angles for all matrices, I calculated the fidelity figure of merit (Equation 1.26). The values for each matrix are reported in Table 3.3. This constitutes one of the first experimental measurements of the fidelity for a physical UPP with a Clements layout, the first of its kind for a UPP fabricated via FLWW.

CHAPTER 4

Multimode Waveguide Characterization

4.1 Introduction

In this chapter, I introduce additional research work that I have conducted in parallel with the main body of work described in the previous chapter. Here I present an experimental study of a novel characterization method for optical devices, to characterize their behavior in terms of the number of guided modes that the structure can support.

4.2 Motivation

Over the years, an increased demand for higher data throughput in optical communication systems has pushed the research in the direction of encoding information in several independent properties of light, be it polarization, intensity, or wavelength [44]. In particular, Mode Division Multiplexing (MDM) offers a scalable solution by allowing the same optical channel to transmit a larger amount of data using a single physical medium.

As mentioned in previous chapters, Femtosecond Laser Waveguide Writing (FLWW) is an established fabrication platform for guiding structures with unprecedented levels of flexibility; owing to this versatility,

FLWW is capable of writing even multimode waveguide structures [45]. A critical question in the fabrication of these guiding structures is whether there are simple methods to characterize their behavior. In the case of experimental fabrication of waveguides, checking the presence of multiple modes may be fairly easy. To do so, a laser can be coupled into the waveguide under test, and the evolution of the output intensity distribution can be qualitatively observed by varying the coupling conditions at the waveguide input. If the intensity profile varies, the waveguide can be considered multimode; otherwise, it is considered singlemode. The main limitation of this technique lies in its binary nature. In the case of a multimode waveguide, it is not simple to know how many modes are supported, even approximately. This lack of information is even more problematic when there is no theoretical model that gives the number of modes supported.

4.3 State of the Art

In the literature, several solutions are found to precisely quantify the number of modes supported by some optical medium. For example, a swept-wavelength interferometer can be used [46], employing a set of beam splitters, fiber splitters and waveplates to characterize precisely the modal structure of a multimode optical system. Similarly, transfer-matrix inversion [47] relies on light state manipulation components such as spatial light modulators to shape the light in order to explore the full modal structure.

While these methods can provide very detailed answers to the question of how many - or even exactly which - spatial modes the structure is capable of supporting, they have the major downside of adding complexity and expensive components to the experimental setup.

The end goal is to determine whether information about the modal structure of a waveguide can be extracted from statistical correlations alone, without requiring complex interferometric reconstruction techniques or phase retrieval, and avoiding the need for complex and overly expensive laboratory equipment.

4.3.1 Cylindrical Waveguides Theory

Before analyzing the experimental results, it is important to provide theoretical context on which the methodology described in this work relies.

In general, for a guiding structure, we can represent optical spatial modes as time-periodic solutions to Maxwell's equations of the following form:

$$\vec{E}(\vec{r}, t) = \vec{E}_\omega(x, y, z)e^{-i\omega t}. \quad (4.1)$$

A guided mode propagating along the z -axis is written

$$\vec{E}_\omega(x, y, z) = \vec{E}_0(x, y)e^{i\beta z}, \quad (4.2)$$

where β is the propagation constant, as described in Section 1.3.

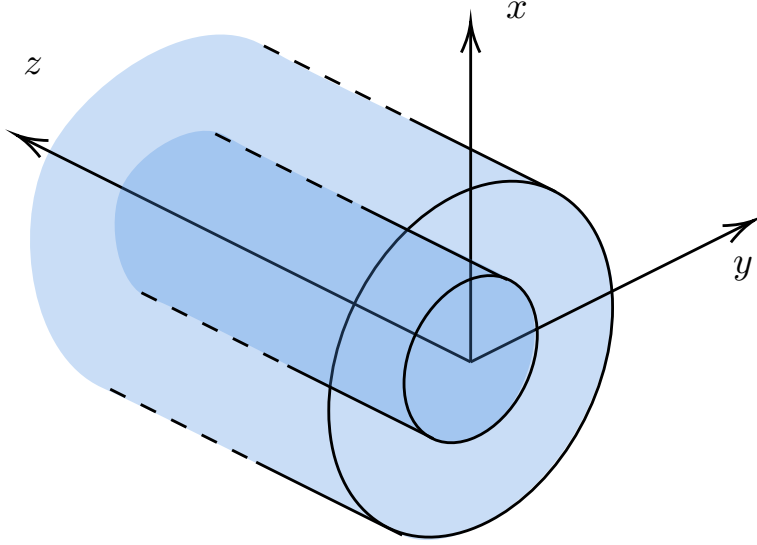


Figure 4.1: Pictorial representation of a cylindrical waveguide. The darker region (the core) has higher refractive index than the outside (the cladding).

Cylindrical waveguides are light guiding structures displaying complete translational symmetry along z and radial symmetry in the xy plane (Figure 4.1). The optical modes sustained by the guiding structure can therefore be described in terms of cylindrical coordinates. These waveguides are characterized by a refractive index $n(r, \varphi)$ that does not depend on the azimuthal angle φ ; like many other waveguides, they feature a central region of higher refractive index (the core) surrounded by a region of lower refractive index (the cladding).

The number of modes that the structure can sustain, that is, the number of solutions to the wave equation (Equation 1.4) depends on the frequency ω of the wave and therefore on the wavelength λ . It can be shown that the modes of a cylindrical waveguide can be labeled via a pair of integer numbers, l and m , respectively called the «azimuthal order» and «radial order». The azimuthal order is an integer between 0 and $l_{\max}$, with $l_{\max}$ also depending on λ .

We typically express the number of modes sustained by the waveguide with a normalized frequency parameter V that depends on the wavelength

and the geometrical properties of the waveguide. In particular,

$$V = \frac{2\pi\rho}{\lambda} \text{NA} \quad (4.3)$$

where ρ is the radius of the waveguide's core and NA is the numerical aperture. Figure 4.2 shows the dependence of the number of modes sustained by a cylindrical waveguide with fixed NA as the wavelength increases.

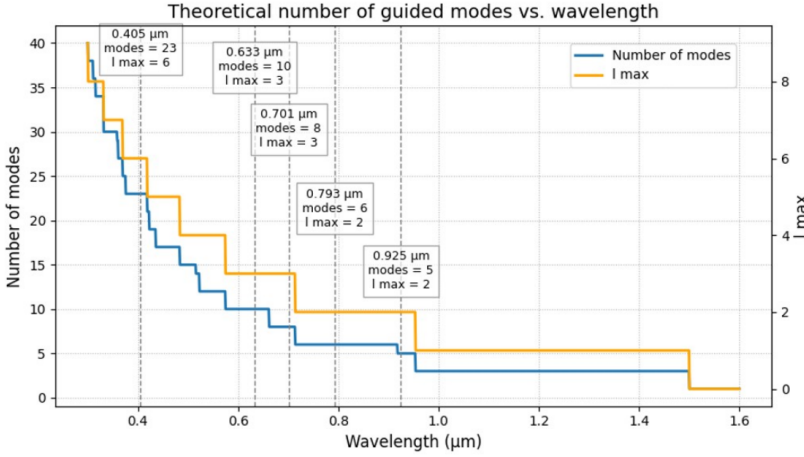


Figure 4.2: Theoretical number of guided modes and their cutoff dependence on wavelength for a step-index fiber with numerical aperture $\text{NA} = 0.13$, tuned for single-mode operation at 1550 nm. This graph was obtained by finding the solutions to the wave equations via numerical methods.

4.3.2 Principal Component Analysis

In data analysis, Principal Component Analysis (PCA) is a linear dimensionality reduction technique that represents data by its set of statistical correlations.

By «data» in this context we mean a set of n -dimensional vectors $\{\vec{p}_i\}$, i.e. a subset of $\mathbb{R}^n$. Given a basis $\{\vec{e}_j\}$, every data point $\vec{p}_j$ can be written as a superposition

$$\vec{p}_i = \sum_n c_{ij} \vec{e}_j. \quad (4.4)$$

PCA is a technique to apply a basis change to the dataset in such a way that each data point $\vec{p}_i$ can be written as a linear combination of an orthonormal basis, where each basis element induces the largest possible variance in the coefficients c_{ij} . More concretely, it provides a new orthonormal basis

$\{w_j\}$ called the «principal components», with which to express the data-points $\{\vec{p}_i\}$ in such a way that the new linear combination coefficients c'_{ij} obtained by projecting each data point onto the principal components have maximal variance.

Given the data points $\{\vec{p}_i\}$, we may start by assuming that each element of the datapoints has null average:

$$\frac{1}{S} \sum_{j=1}^S (\vec{p}_i)_j = 0 \quad \forall i \quad (4.5)$$

where S is the size of the dataset. If this is not verified, we can subtract the average and obtain a new dataset with this property. Notice that the data is a set of vectors, so they can be arranged into a matrix P where

$$P_{ij} = (\vec{p}_i)_j. \quad (4.6)$$

and the null-average requirement is equivalent to asking that the sum of all columns of P is the zero vector.

The first principal component $\vec{w}_1$ is a vector with the following property: the coefficients c_{i1} have the largest variance when the data points $\vec{p}_i$ are projected onto it. In mathematical terms, this variance is calculated as

$$\sigma_1^2 = \frac{1}{S} \sum_{i=1}^S c_{i1}^2 \quad (4.7)$$

where

$$c_{i1} = \vec{p}_i \cdot \vec{w}_1. \quad (4.8)$$

Therefore, $\vec{w}_1$ is the (normalized) vector such that σ_1^2 is maximised; in mathematical terms, this can be written as

$$w_1 = \arg \max_{\|\vec{v}\|=1} \frac{1}{S} \sum_{j=1}^S (\vec{p}_j \cdot \vec{v})^2. \quad (4.9)$$

This can be rewritten using the P matrix:

$$w_1 = \arg \max_{\|\vec{v}\|=1} \frac{1}{S} \|P\vec{v}\|^2 \quad (4.10)$$

and since $\frac{1}{S} \|P\vec{v}\|^2$ is maximized if and only if $\|P\vec{v}\|^2$ is, we can remove the factor $\frac{1}{S}$. Finally, since

$$\|P\vec{v}\|^2 = \vec{v}^T P^T P \vec{v} \quad (4.11)$$

we have

$$w_1 = \arg \max_{\|\vec{v}\|=1} \vec{v}^T P^T P \vec{v}. \quad (4.12)$$

It can be shown that for a positive (semi-)definite matrix, which $P^T P$ always is if P is real, the above equation is equivalent to finding the largest eigenvalue λ_1 of $P^T P$, and $\vec{w}_1$ would be the corresponding eigenvector. In fact, *all* the $\{\vec{w}_j\}$ can be found in this way, and are ordered starting from the largest eigenvalue to the smallest.

This procedure is formally equivalent to taking the singular value decomposition of P :

$$P = U \Sigma W^T \quad (4.13)$$

where Σ is a diagonal matrix of P 's «singular values», that is, the square roots of the eigenvalues of $P^T P$. These elements represent the «weight» each principal component has in making up the data, or how much the data depends on the variation of each principal component. The principal components would therefore be the eigenvectors of $P^T P$.

One of the main uses of PCA is to perform dimensionality reduction, i.e. reducing the data to a smaller subset of basis elements that explain most of the variance. In fact, since the singular values describe how much of the data's variance is «caused» by the corresponding principal component, it implies that most of the information contained in the data is contained only in the components with largest singular values. We can then cut off all the low-singular-value components from the data and reduce the dimensionality while maintaining most of the information. Methodologies by which to decide the boundary point for dimensionality reduction are widely discussed in the literature [48]; these methods include criteria based on cumulative explained variance, information-theoretic approaches, and statistical tests. The choice of method to obtain the boundary point is discussed in Section 4.6.

4.4 Multiple Mode Excitations

Let us consider a situation where several modes of a guiding structure are excited simultaneously, each with complex amplitude A_j . Then, the total field is described by a superposition

$$E_{\text{tot}}(x, y) = \sum_j A_j E_j(x, y) \quad (4.14)$$

where $E_j(x, y)$ are real components of the field and j runs over all pairs of indices supported by the structure; in cylindrical waveguides, this means j

runs over all supported l, m . The intensity of this field is proportional to the square modulus

$$I(x, y) \propto |E_{\text{tot}}(x, y)|^2 = \left(\sum_i A_i E_i(x, y) \right) \left(\sum_j A_j E_j(x, y) \right). \quad (4.15)$$

Since the fields are real, i.e. $E_i(x, y) = E_i^*(x, y)$, expanding the product yields a sum of $N(N + 1)/2$ distinct spatial contributions. These are the intensities of each mode taken individually plus all pairwise interference terms:

$$I(x, y) = \sum_i^N \sum_{j \geq i}^N \gamma_{ij} w_{ij}(x, y) \quad (4.16)$$

where

$$w_{ij} = \frac{1}{2} c E_i(x, y) E_j(x, y) \quad \text{and} \quad \gamma_{ij} = (\overline{A_i} A_j + \overline{A_j} A_i). \quad (4.17)$$

At the output of a waveguide, therefore, the interference pattern of the field is described by Equation 4.16 which is composed of up to $N(N + 1)/2$ terms.

4.5 Method and Setup

The key idea of the method lies in the fact that the interference patterns at the output of a structure must be a superposition of the form described in Equation 4.16. Many different interference patterns, obtained by exciting the modes of the structure in different ways, would still be the result of the same kind of superposition but with different weights. In a real experiment, a measured interference pattern will be generated by a basis of $N(N + 1)/2$ many terms plus noise; these $N(N + 1)/2$ will explain most of the variance of the data, and therefore PCA is in principle capable of picking them up as the highest-principal-value (most «relevant») principal components.

By gathering a large amount of different interference pattern images as datapoints and applying PCA, the relevant principal components should generate a basis with approximately the same dimension as the number of elements of the sum in Equation 4.16, the rest being noise. This, in turn, gives a direct measurement of the number N of modes sustained by the structure.

The most important requirement of this method is to ensure that the dataset sufficiently samples the modal excitation space, otherwise the number of relevant principal components might not correlate with the number of underlying physical contributions.

In order to test all of these ideas experimentally, I put together a setup that leverages many of the same devices of the automated calibration setups described in Chapter 3.

The experimental apparatus uses five lasers (405 nm, 633 nm, 701 nm, 793 nm, 925 nm), a coupling objective mounted on a motorized translation stage, and an imaging objective projecting the output onto a CCD camera. For this experiment, I employed two single mode fibers as sample: 780HP, and 1550BHP.

A scheme of the setup is shown in Figure 4.3. After a section where the laser light is in free space, the light is focused into a single mode fiber via an objective lens (lens with $NA = 0.40$ and focal length $f = 6.24$ mm) mounted on the translation stage. On the opposite end of the sample a second objective lens ($NA = 0.68$, $f = 3.1$ mm) focuses the light onto a UI-224xSE-M CCD camera in order to perform the imaging of the output intensity distribution.

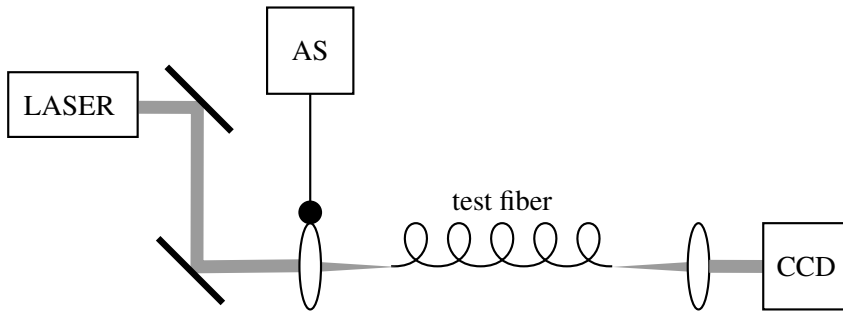


Figure 4.3: Schematic of the experimental setup including the scanning objective, the characterized fiber, the imaging objective, and the CCD camera.

The main idea of the method is to use the automated translation stages to perform a scan over many different positions in the xy plane. If the light is focused in different positions at the input of the structure, it will induce different modal excitations which in turn provide different interference patterns once the light is collected by the imaging objective and retrieved by the camera.

The input objective performs a scan within a circle of radius $5\text{ }\mu\text{m}$, each xy coordinate sampled from a grid with step size $0.25\text{ }\mu\text{m}$, producing approximately 1250 images per scan. A schematic representation of the positions taken by the input objective over the course of one measurement is shown in Figure 4.4. The specific instruments used in the measurement setup (PI Compact Linear Stage M-112 and M-111) feature $3\text{ }\mu\text{m}$ of backlash, which needs to be compensated in order to obtain a reproducible po-

sition across many experiments; the backlash compensation, also shown in Figure 4.4, consists of overshooting the position by $5\text{ }\mu\text{m}$ and returning to the intended target afterwards.

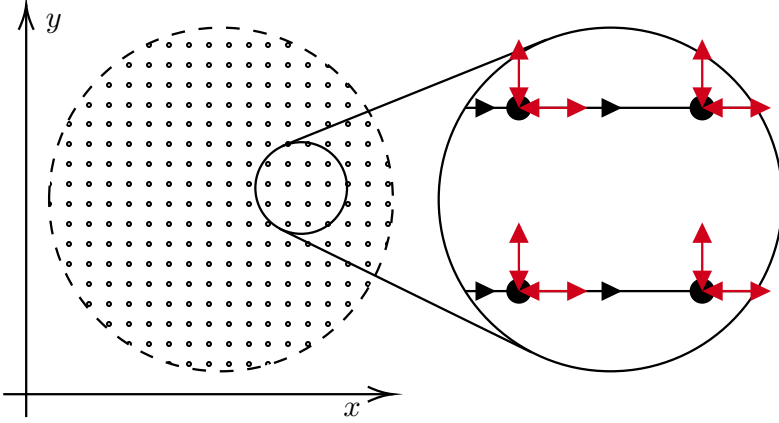


Figure 4.4: Circular scanning pattern used to vary modal excitation. Each dot in the larger circle denotes an acquisition point. The automated stage stops on every point on the same row from left to right (along increasing x) before moving up to the next row (along increasing y); the red arrows indicate the additional movements that the stage performs in order to compensate backlash (not to scale).

A few examples of images obtained by the CCD camera during this experiment are shown in Figure 4.5. Each acquired intensity pattern is cropped to a square pixel region centered on the brightest pixels. To reduce the dimensionality of the obtained image, I performed binning on the images to obtain a smaller representative of the image that is 32×32 pixels in size.

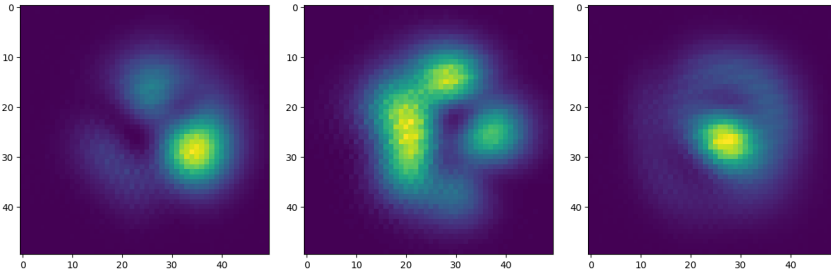


Figure 4.5: Examples of images captured by the CCD camera, after binning. These images were obtained with laser light at 793 nm focused inside of the 1550BHP fiber.

Each image is flattened into a 1024-element vector, yielding a dataset

matrix of size $S = 1250$ and dimensionality $n = 1024$. After subtracting the average from each pixel, PCA is applied to the dataset to recover the eigenvalues and eigenvectors of the $P^T P$ matrix.

4.6 Data Processing and Results

In this experiment, two fibers were tested: Thorlabs 1550BHP (core radius $4.5\text{ }\mu\text{m}$, $\text{NA} = 0.13$), and Thorlabs 780HP (core radius $2.2\text{ }\mu\text{m}$, $\text{NA} = 0.13$). Each fiber was tested with all laser sources, obtaining data for their behavior with 405 nm, 633 nm, 701 nm, 793 nm and 925 nm.

The retrieved eigenvalue curves have several common features:

- A steep initial drop (which corresponds to the region with dominant modal terms),
- A quasi-linear decay region (which corresponds to the region with noise terms),
- A rapid transition to noise-dominated components.

Example pictures of the retrieved eigenvectors for both the modal region and the noise region are shown in Figure 4.6, while a graph of the eigenvalues obtained from the experiment is shown in Figure 4.7.

To determine the number of physically meaningful principal components, it was necessary to choose which thresholding methods to apply in order to distinguish between the «modal» region and the «noise» region of singular values. I decided to apply two separate methods, comparing their ability to identify the correct boundary for the two regions.

First, I applied simple eigenvalue thresholding. In short, this method consists of choosing a threshold Y value for the eigenvalues σ_i^2 corresponding to each principal component, and the boundary is set to the smallest i such that $\sigma_i^2 < Y$. In other words, any singular value below $\sqrt{Y}$ is discarded.

Secondly, I applied Logarithmic Eigenvalue (LEV) analysis. This method hinges on a property widely shared by many graphs obtained from PCA when plotted in logarithmic scale: they feature a region with a sharp decrease of the eigenvalues, followed by an inflection point (called the «knee») and a transition into a linear region (such as in Figure 4.7). LEV is a good technique to use in order to distinguish between these two regions and therefore separate the relevant principal components from the irrelevant ones, as the linear region is also to be identified with the noise region.

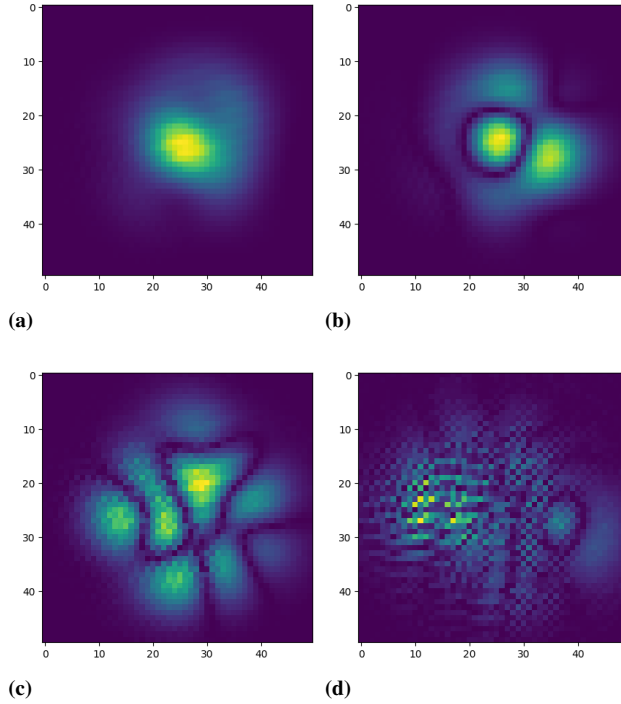


Figure 4.6: Examples of principal components retrieved using PCA. Panel (a) is the first principal component; panels (b), (c), and (d) are respectively the third, tenth and twentieth principal components. In particular, panel (d) represents a principal component classified as noise term. These images were obtained with laser light at 793 nm focused inside of the 1550BHP fiber.

To implement it, I first took a sliding window of the data in the logarithmic scale, and I evaluated a linear fit of the data within each window. Then, for each window I evaluated the coefficient of determination R^2 between the data and the linear fit; this method allowed me to identify the knee of each graph, and the boundary point of the modal region is set to where R^2 exceeded a predefined threshold. The parameters for this method were also arbitrarily chosen: the window was set to 100 eigenvalues and a threshold of 0.98 was chosen for R^2 .

A summary of all the data and a comparison between them and the expected number of contributions is shown in Figure 4.8. This Figure shows, for both fibers, a comparison between the theoretical number of intensity contributions $N(N + 1)/2$ and the retrieved boundary points with eigenvalue thresholding and LEV across all the observed wavelengths. The theoretical number is marked with a gray triangle, while eigenvalue threshold-

Chapter 4. Multimode Waveguide Characterization

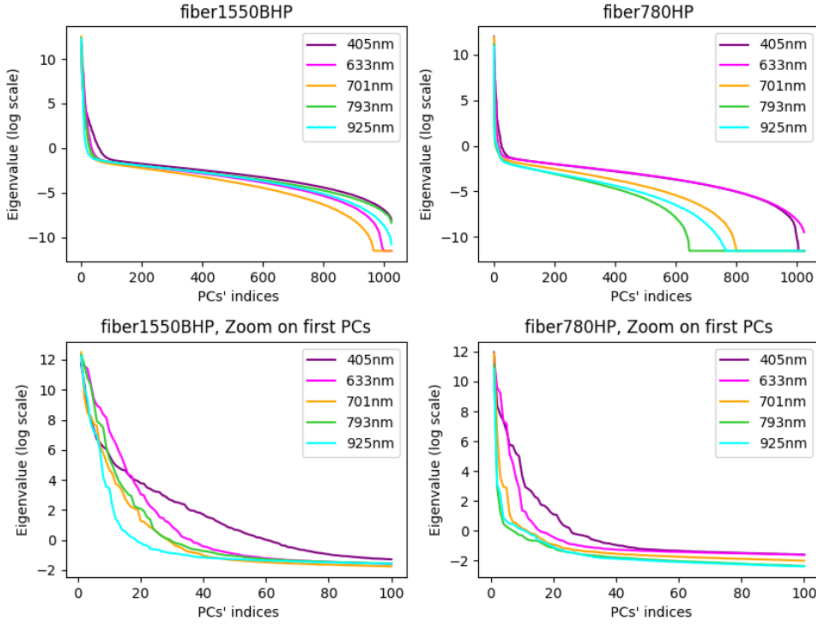


Figure 4.7: A graph of the eigenvalues retrieved using PCA, for both fibers, in logarithmic scale. Below each graph, the same curves restricted to just the first 100 eigenvalues.

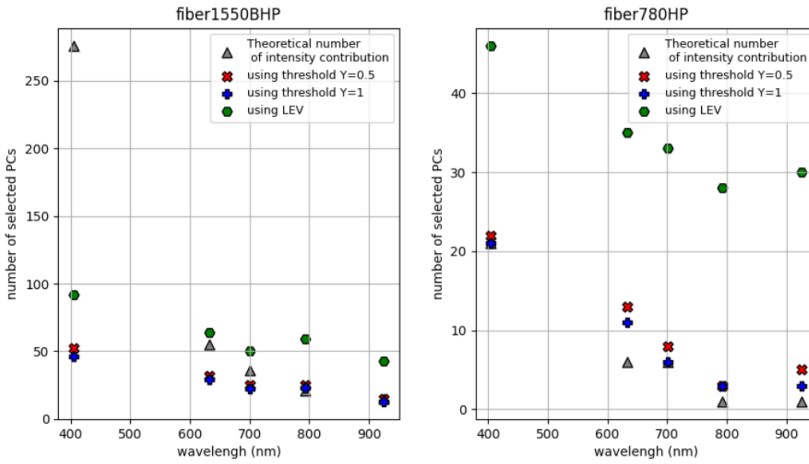


Figure 4.8: Comparison between the obtained boundary points with the two methods, focused on a comparison between the two methods and the expected values.

ing is marked with either a red or a blue cross; the boundary point identified with LEV with the described parameters is marked with a green hexagon.

The Figure shows that while simple thresholding is unable to retrieve

the expected number $N(N + 1)/2$ exactly, the obtained values follow the expected trend. It is important to remember that this method includes an arbitrary choice of Y . Notably, the number obtained with LEV analysis is always higher than the theoretical number. A potential explanation for this result is that LEV analysis is implemented with parameters that prioritize minimizing information loss over reducing the number of principal components. The only exception to this is the case of the 1550BHP fiber at 405 nm, where the number is much lower. This can be explained by the fact that the number of modes sustained by the fiber at this wavelength is very high (23) and the scanning method might be unable to sufficiently excite the entire modal space.

While neither method is able to retrieve the exact number of contributions to the sum expected from the theory, they both appear to follow the same trend. This is encouraging, and it provides evidence that this method is able to capture large-grain differences between the modal spaces; this opens the door to further investigation on this topic as will be detailed in the next section.

4.6.1 Reconstruction of Rotating Phase Profiles

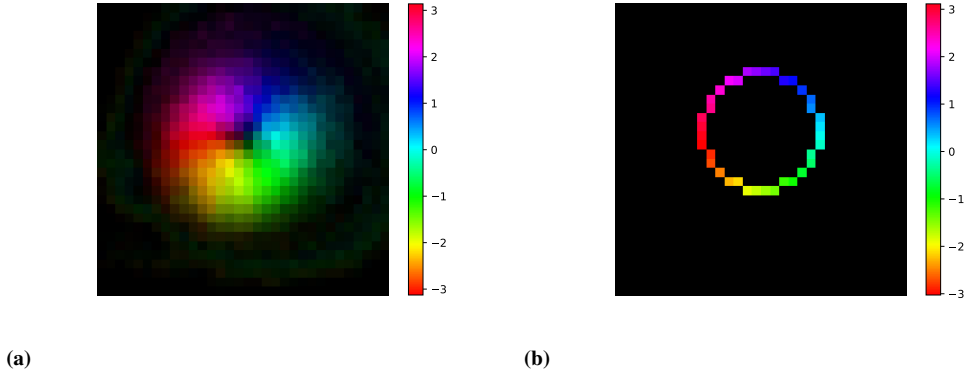


Figure 4.9: Rotating phase field with $l = 1$. Figure (a) shows the typical profile of a rotating phase field supported by the waveguide, reconstructed from the principal components obtained on the 1550BHP waveguide at 925 nm. Figure (b) shows the target for the reconstruction. The colors represent the phase of the field (in the range $[-\pi, \pi)$); black pixels represent small amplitudes and colorful pixels large amplitudes.

Because the reduced PCA basis should span the same subspace generated by the modal interference patterns, it should be possible to recon-

struct synthetic phase patterns from the principal components themselves. This, in turn, allows for further qualitative classification between different waveguides.

One interesting phase pattern to attempt to reconstruct, that can provide information about the properties of the guided modes of the structure, is the so-called «rotating phase» pattern. These patterns are characterized by an intensity profile in the shape of an annulus, while the phase depends on the radial angle measured from the center of the annulus. More concretely, rotating phase fields are of the form

$$I(r, \varphi) = I_0(r)e^{il\varphi} \quad (4.18)$$

with l an integer. An example of such phase profile is shown in Figure 4.9a. Using the PCA method it is possible to probe the maximum l ($l_{\max}$) that can be achieved via linear combination of the principal components before the reconstruction fails. This evaluation would allow us to further distinguish between different guiding structures.

It is important to note that while these profiles are artificial, they still contain physically meaningful information about the underlying modes. In fact, a rotating phase profile with $l = 1$ obviously cannot be created as a linear combination of just one electric field, meaning that a singlemode waveguide should not be able to reconstruct an $l = 1$ profile. Likewise, the higher l and the larger the variety of modes that need to be included in order to reconstruct it.

For this reason, we expect $l_{\max}$ to increase with the number of modes in the waveguide N .

The quality of the reconstruction needs to be evaluated by establishing a figure of merit, so that we can quantify what it means to be able to reconstruct these profiles. For this reason I introduced a new type of fidelity, in a similar fashion to the fidelity functions as defined for UPPs. This «reconstruction fidelity» is computed as:

$$F = \frac{\text{Tr}\left\{\left|m_t^\dagger m_r\right|\right\}}{\sqrt{(m_t^\dagger m_t)(m_r^\dagger m_r)}} \quad (4.19)$$

where m_t is the target matrix and m_r the PCA-based reconstruction. Once again, the reconstruction fidelity is 1 when the two matrices are exactly the same. The targets in question are simplified profiles of rotating phase fields: each target matrix was generated by taking the center of mass of the first principal component and evaluating its width and height to obtain a pair of

dimensions σ_x, σ_y . The width and height are taken as simply the standard deviation of the distribution in the x and y direction respectively.

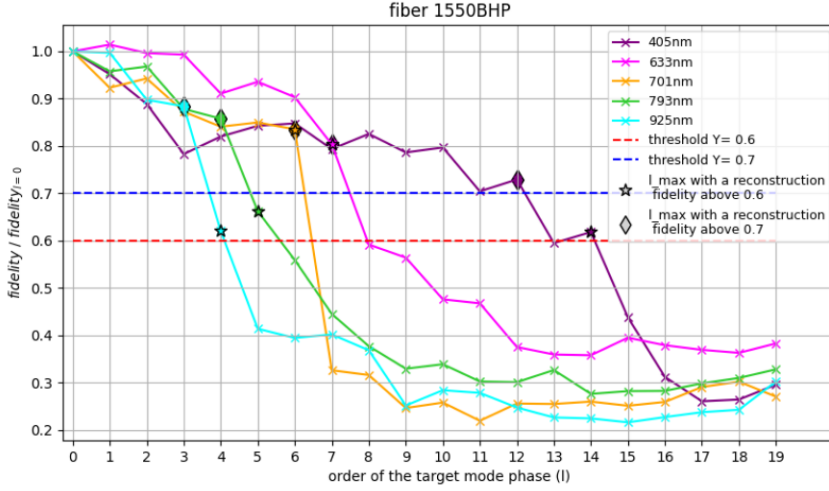
Then, the target image is constructed by initializing every pixel to zero, and only setting to 1 those pixels that lie on an ellipse (with axes σ_x, σ_y), centered on the center of mass of the first principal component. Then, for each element I evaluate the angle θ as the angle between the x axis and the line segment connecting the center with the x, y position of the pixel. Then, I multiply the pixel by $e^{il\theta}$, where l is the order that I would like to test.

Finally, I take the relevant principal components obtained via PCA as a basis, only up to the boundary defined using LEV analysis as detailed in the previous section. With this basis, I then apply simple gradient descent to maximise the reconstruction fidelity over all complex linear combinations of the basis. I applied this method for each fiber and each wavelength, evaluating the reconstruction fidelity of rotating phase profiles with increasing l .

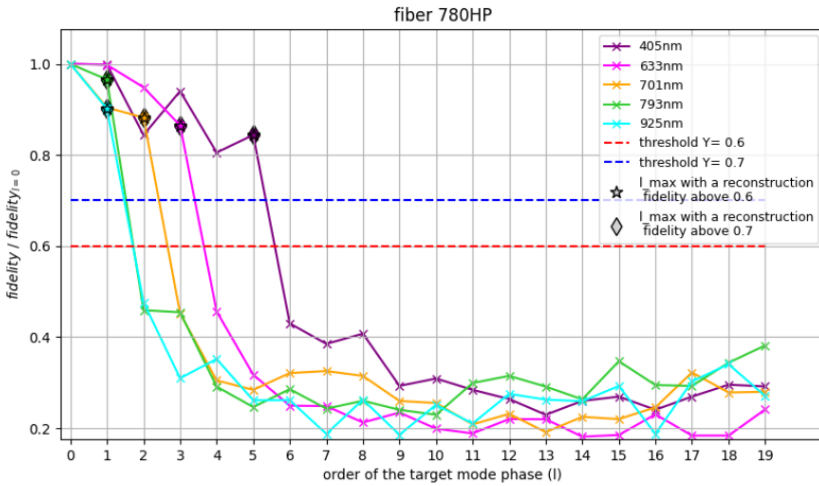
Figure 4.10 shows the fidelity obtained in this way. It is worth noting that in the Figure all fidelities are normalized with respect to the $l = 0$ case; this is because the absolute value of the fidelity (or how close it is to 1) does not matter, since a perfect reconstruction of the target matrix is irrelevant. All that matters is where this fidelity shows a sharp drop relative to the previous values of l , which would allow us to identify $l_{\max}$.

As can be seen in the Figure, all the curves have a similar trend. In fact, for all the tested wavelengths and for both waveguides the curves feature a plateau of very similar values, followed by a sharp decrease that leads to a second plateau characterized by lower fidelity values. This behavior is consistent with the theoretical predictions: in principle, this sharp decrease corresponds with the maximum azimuthal angle l such that the principal component basis can no longer consistently generate the rotating phase profile. With lower wavelength, this maximum l increases, which matches the expected trend since a cylindrical waveguide can sustain higher-order modes at higher normalized frequencies (and thus, lower wavelengths). This is strong evidence that the eigenvectors found with PCA in fact contain physical information about the modal structure of the waveguide, and that this method of probing rotating phase patterns is able to distinguish between waveguides.

For a more quantitative analysis, a threshold was chosen to distinguish between successful and failed reconstructions as shown in Figure 4.10. We can observe that these thresholds give values for $l_{\max}$ that follow the expected trend, but there are a few caveats that suggest more work needs to be done to retrieve more refined information from this experiment.



(a)



(b)

Figure 4.10: The reconstruction fidelity evaluated using the principal components selected via LEV analysis, as a function of l , and normalized with respect to the $l = 0$ case for clarity. All the graphs show a plateau followed by a sharp drop in fidelity, which matches the expected maximum azimuthal angle supported by the guide.

For example, further study could help to explain the case of the wavelengths 633 nm and 701 nm in the 780HP fiber, where the theoretical contribution is expected to be similar, yet the maximum reconstruction order differs experimentally. In the current state, it seems difficult to determine

whether this discrepancy originates from the characterization method or from the modeling used to obtain the theoretical results.

Another observation, still concerning the 780HP fiber, is that 793 nm and 925 nm both fall in the spectral region where the fiber is singlemode. As such, the maximum value $l_{\max}$ should be zero, but the experimentally retrieved one is $l_{\max} = 1$ in both cases. However, this behavior can be explained by looking at the number of modes that LEV analysis retrieves for these wavelengths: as is the case for all wavelengths of the 780HP fiber, LEV cuts off at a much higher boundary point than expected (Figure 4.8), 28 for 793 nm and 30 for 925 nm instead of just 1. This leads to an overfitting of the target matrix which is the main cause of this high fidelity value.

All of these observations reveal certain limitations of the method, which currently make precise quantitative use impractical. However, they also provide several avenues for improvement that could eventually lead to a usable and more efficient version of the method. Moreover, the interpretation of the general behavior observed (whether through the PCA eigenvalues alone or through the results of the phase profile reconstruction tests) reasonably suggests that these principal components contain meaningful physical information. It is therefore reasonable to pursue the development of a method capable of extracting this information more effectively.

CHAPTER 5

Conclusions and Future Developments

In this work, a comprehensive framework for the characterization of FLWW multimode integrated optical devices was demonstrated and experimentally validated.

In Chapter 3 a fully automated calibration pipeline for UPPs was proposed, enabling the measurement of all the elements of the cross-talk matrix of the UPP with minimal human intervention and reducing calibration times from multiple days down to a few hours. The same setup was shown to be compatible with both white-box and gray-box calibration techniques; the experiments demonstrate that the method is robust and repeatable across devices of increasing size, achieving high performance operation over 6, 8, and 12-mode UPPs. This powerful methodology is easily scalable to devices of even larger size, such as the 24-mode FLWW UPP described in Ref. [32], provided the number of modes is supported by the fiber switches and - if included - multi-channel power meters.

The speed and reliability of the automated calibration have further enabled fidelity optimization methods based on the fine-tuning of the currents flowing through the thermo-optical phase shifters. While not intended as a replacement for full calibration, the demonstrated approach shows that high-fidelity operation can be reached by FLWW UPPS in applications

where a finite number of matrices is required.

Thermal cross-talk as a bottleneck for scalability is investigated through an analysis of the effects of vacuum on our FLWW UPPs. Experimental results show that operating under reduced pressure substantially increases thermal isolation between heaters, leading to a suppression of thermal cross-talk by an order of magnitude and a reduction in dissipated power by a factor 4. Importantly, while vacuum operation modifies the thermal dynamics of the system, the overall effective time required to reach a stable target matrix remains comparable, highlighting low-pressure environments as a viable route for improving calibration quality and operational stability in large-scale processors.

Additionally, the framework expands beyond just optical power measurements as a novel technique for the retrieval of the complex elements of the implemented unitaries is introduced. This, in turn, enables the evaluation of the fidelity taking into account the complex nature of the implemented unitaries. The use of a dedicated IIC enables simultaneous multi-port phase measurements and represents a crucial step toward complete measurement of complex-valued unitary matrices in FLWW UPPs. While this method has not yet been applied to 4-by-4 matrices or larger, the included analysis of the experimental validation of this technique offers solid proof that the method can be extended to the measurement of full matrices of any size.

Looking forward, several directions naturally emerge from the work presented in this thesis. On the calibration side, the integration of the phase retrieval into the automated pipeline would enable full complex-matrix data-driven algorithms at scale. In fact, including the phase information into the optimization method and applying the algorithms on this more comprehensive dataset would lead to a more informed calibration capable of modeling the full behavior of the UPP.

From the point of view of the actual implementation of unitary matrices, refining the decomposition pipeline could improve fidelities further. More sophisticated phase shift retrieval methods, for example those discussed in Ref. [25], could help in reproducing the high-fidelity performance seen in smaller UPPs when the number of modes is increased and the devices are more sensitive to deviations in the reflectivity of the couplers.

On the side of scalability, the number of modes of the UPP is also a driver of the physical dimensions of the glass chip, which incentivizes the miniaturization and compactification of the components and thus an increase in thermal cross-talk. Vacuum operation (and vacuum packaging) as a means of severely reducing the cross-talk would allow smaller physical

footprints for UPPs while preserving their high performance. Conversely, reducing the thermal cross-talk without changing the geometric parameters of the UPP is useful in its own right to alleviate the heavy computational costs of the machine-learning algorithms, allowing the models to optimize over a smaller number of parameters.

More broadly, the demonstrated framework provides a concrete foundation for deploying FLWW UPPs in increasingly demanding quantum photonic applications. In this context, this research work on the calibration of UPPs was instrumental in the realization and experimental validation of advanced quantum algorithms and protocols.

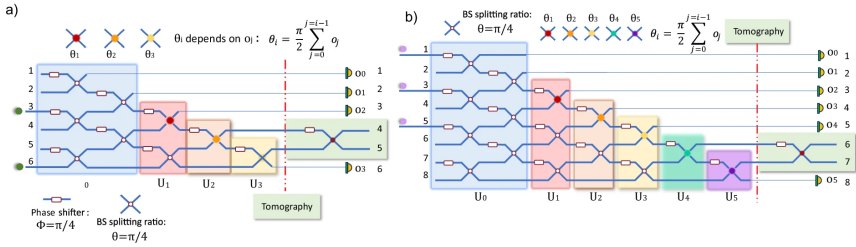


Figure 5.1: The experimental use of two calibrated UPPs to implement adaptive boson sampling protocols. A 6-mode UPP is programmed to implement a sequence of conditional unitaries (a), where subsequent transformations are dynamically reconfigured by detection events, enabling heralded qubit state preparation and tomography with high kernel fidelity. This approach is then scaled to an 8-mode UPP (b), implementing an adaptive architecture with up to 15 unitaries selected by the detection events, showcasing the ability of the UPPs to support feed-forward quantum photonic protocols. Figure adapted from Ref. [49].

Among the most recent developments on this front, an Adaptive Boson Sampling (ABS) algorithm for Quantum Machine Learning (QML) was demonstrated [49], employing both a 6-mode and an 8-mode UPP of the exact kind described in this work to implement a QML-based binary classification task using multi-photon states (Figure 5.1). More concretely, the ability of the UPPs to be reconfigured on the fly was instrumental in implementing ABS for the estimation of the pairwise similarities between encoded quantum states of the photons - i.e. the so-called «quantum kernel matrix» - which enabled the UPP to successfully complete classification tasks in a fully quantum pipeline.

The UPPs described in this work were also used in bleeding edge research in quantum information. In this context it is worth mentioning the work reported by Ref. [50] for the implementation of threshold Quantum State Tomography (tQST) on the 8-mode UPP, and the work by Ref. [51]

where both the 8-mode and 12-mode devices are deployed to create a Photonic Quantum Convolutional Neural Network, experimentally validating the capabilities of this structure on the task of binary classification of 4-by-4 images.

Additional work that is yet to be published as of time of writing also makes extensive use of these calibrated devices to push the boundaries of research in fundamental physics, quantum information and algorithms [52–54].

In all these cases, the calibration, characterization, and control framework developed in this chapter was a key component to enable these advances in integrated photonics research, providing a rigorous and experimentally validated pipeline to go from a fabricated and packaged interferometric mesh to a completely characterized and operable device, with simple specifications in terms of setup complexity and software implementation.

In this way, protocols that require unitary transformations of the state of light across increasingly many modes - be it classical or quantum - can be implemented, benchmarked, and reproduced across different devices with high fidelity.

In Chapter 4, I presented a PCA-based technique for classifying multi-mode waveguides through spatial scanning and imaging of the output interference patterns. Under the assumption that the dataset explores the space of modal excitations sufficiently, the method can be used to discern between different numbers of supported modes by reducing the dimensionality of the dataset to the number of contributions to the interference patterns. It's important to remark that the proposed pipeline is, in principle, completely general and structure-agnostic. Some knowledge of the guiding structure is leveraged only in the choice of the reconstruction tests, which nevertheless must be physically informed in order to be meaningful.

While neither of the implemented thresholding methods (eigenvalue thresholding versus LEV) returned the exact theoretical number of expected contributions, both were found to reproduce the correct qualitative trends across wavelength and fiber type, providing experimental evidence that the PCA basis does in fact capture large-grain differences in the underlying modal space. This analysis was further refined by analyzing the space of achievable patterns starting from the selected principal components, using synthetic phase and intensity profiles to probe the ability of the basis to reproduce specific patterns that contain information about the modal structure. The experimentally observed behavior matches theoretical expectations and therefore constitutes strong evidence that the principal components extracted from the interference datasets contain physically

meaningful information about the guiding structure as a whole.

At the same time, the experiments also highlight that the current implementation is not yet quantitatively robust. In particular, the chosen thresholding methods overestimate the number of relevant components and - on the opposite side - insufficient modal excitation may lead to an underestimation of the number and is difficult to detect without prior knowledge of the expected theoretical values. These issues propagate to the reconstruction tests, where an overestimated basis may lead to overfitting and yielding unphysical results even in nominally singlemode conditions.

More refined techniques may therefore be necessary to achieve a more precise classification of multimode structures. For example, a more varied method of excitation of the modes might be required such as varying not just the position of the focused light but also the angle at which it enters the guiding structure.

The thresholding selection method contributes heavily to the method's validity and precision. While LEV provides good results across a wide range of wavelengths, as shown by the reconstruction tests, the present implementation relies on arbitrary window sizes and parameter choices. A natural next step is to employ more principled thresholding criteria, for example by using established approaches such as cumulative explained variance, information-theoretic criteria, or statistical tests discussed in the PCA literature [48], or even incorporate a new model tuned to the specific setup in order to further refine the eigenvalue analysis.

The rotating phase reconstruction tests do provide meaningful insights into the structure of the analysed waveguides. Additional tests with different shapes can also contribute to the consolidation of the results into a more cohesive framework, even though more diverse tests may require prior knowledge of the underlying modal structure.

Finally, the method could be tested on a broader variety of guiding structures, not just cylindrical waveguides, including FLWW multimode waveguides and other integrated multimode optical components. For these structures the theoretical predictions may be less reliable and a fast and low-cost characterization - especially given its general applicability with minimal tuning of the setup - would be particularly valuable.

List of Figures

1.1	A pictorial representation of an integrated waveguide. The darker area is the core (refractive index n_2), surrounded by the cladding (refractive index n_1). To achieve the guiding action, $n_2 > n_1$	4
1.2	The plot (arbitrary units) of the electric field profile of the fundamental mode for the analyzed structure, along the x axis at fixed $y = 0$. The core region is shaded in blue. This graph was obtained by numerically solving Equation 1.4 for a core of width $1\text{ }\mu\text{m}$, $\Delta n = 10^{-2}$ and $\lambda = 633\text{ nm}$	5
1.3	A pictorial representation of integrated coupled waveguides. The two cores have the same refractive index.	6
1.4	The electric field profiles of the $\vec{S}$ and $\vec{A}$ modes for the coupled waveguides. The parameters used to obtain these profiles are the same as Figure 1.2, with both waveguides having a width of $1\text{ }\mu\text{m}$. The modes resemble a symmetric and anti-symmetric sum of the fundamental mode of each waveguide.	6
1.5	The intensity profile of the $\vec{E}_1(z + \pi/2\Delta\beta, t)$ field, at fixed time. After a propagation of $\pi/2\Delta\beta$, the intensity is split evenly between the waveguides.	8
1.6	Schematic of a directional coupler. The black lines represent integrated waveguides. p_1 and p_2 are the waveguide distances in the non-interacting and interacting region respectively. The geometric interaction distance is d	10

List of Figures

1.7	Schematic representations of Mach-Zehnder interferometers (a) in bulk optics and (b) in integrated optics.	11
1.8	Schematic representations of programmable MZIs in the (a) external-internal and (b) push-pull configurations.	12
1.9	Representations of the Reck (a) and Clements (b) structures. The unit cell is represented in Figure 1.8a. The blue squares represent the extra phase terms, and may be physically controlled by implementing singular phase shifters.	13
1.10	Representations of the Clements (a) and Bell (a) architectures, and their respective unit cells. The blue squares represent some extra phase terms, and may be physically controlled by implementing singular phase shifters.	14
2.1	A render of the typical structure of a UPP. In this picture, the structure is the one of a 6-mode Clements layout. Inset (a) shows a scheme of the unit cell with the active elements; inset (b) shows a microscope picture comprising 3 heating elements in a column (M1, M2, M3 and their common ground G); the grey rectangular areas are trench structures dug into the glass, leaving thin slivers where the waveguides have been written.	19
2.2	Renders of the typical structure of an MZI written with FLWW and FLM. Figure (a) shows the structure of a Clements unit cell from the top; Figures (b) and (d) show a render and a microscope image of a section of trench structure fabricated with FLWW. Figures (c) and (e) show the same images for a bridge structure instead. Figures adapted from Ref. [17].	20
2.3	Two assembled UPP devices. On the left, a 6-mode UPP with gold heating elements and its size compared to a 1 Euro coin; on the right, a fully assembled 8-mode UPP with chromium-copper heating elements after the pig-tailing process.	20
2.4	Measurement of resistance stability, showing the resistance drift as a relative percentage over time. In Figure 2.4a the experiment was performed over 12 hours, actuating a single resistive element fabricated with sputtered gold. In Figure 2.4b, the experiment was performed over 5 hours on a resistive element fabricated with chromium-copper deposition. .	23

2.5	The basic idea that allows the Bell and Clements structures to be remapped onto one another. If a certain phase φ (in red) needs to be achieved on a physical location of the Clements structure, this can be created by adding and subtracting a common mode phase φ on two different MZIs of the Bell structure.	26
2.6	The devices used for current control. Figure 2.6a shows the Q8iv driver modules, which act as multichannel current sources; for the pictured device, up to 64 resistive elements channels can be controlled at the same time. Figure 2.6b in an interposer board, which exposes the channels of the drivers with a standard interface to more easily connect the current control to the UPP device.	29
2.7	The electrical scheme of a UPP. There are numerous resistors, with resistance values R_i , all connected to the same ground node. In reality, between the common node of the resistors and the ground a small resistor should be considered, with a ground resistance R_g	29
3.1	The first setup to measure the inputs and outputs of a UPP. This setup is fully automated, with a fiber switch at the input (not pictured) and an automated stage at the output (not pictured) controlling the lens that collects the output light.	38
3.2	The setup used for the first automated measurements. The light from the laser source is focused into a fiber, while the fiber switch (FS) selects which input of the UPP the light is injected to. Then, the light is collected at the output by a lens mounted on an automated stage (AS) and the optical power is read by a power meter (PM). The choice of using an automated stage for the output was made in accordance with the specifications of the UPP, since it did not feature an output fiber array.	39
3.3	A graph of the amplitude fidelity evaluated over 1000 Haar random unitary matrices. The average value of fidelity over this set is 0.9979, highlighted by the dashed black line.	40
3.4	The amplitude fidelity before and after applying the optimization process as described. The black dashed line corresponds to the value of the average (amplitude) fidelity of the original Haar random set.	42

3.5	Effects of the single-matrix optimization algorithm on the worst implemented matrix out of 1000. Figure 3.5a shows a comparison of the difference between the amplitudes of the target matrix before and after the optimization; the algorithm has completely removed a 2 by 2 submatrix that carried the biggest error. In this picture, the considered matrix is the worst of the whole dataset, whose fidelity improved from 0.9934 to 0.9997. Figure 3.5b shows the improvement of the fidelity over the iterations of the algorithm; a simulation of the optimization was also performed on a digital twin of the UPP (see Section 3.2.5).	43
3.6	The setup used for the measurements with polarization state selection. The light from the laser source passes through a polarizer and a series of waveplates before being focused into a fiber; the fiber switch (FS) selects which input of the UPP the light is injected to. Then, the light is collected at the output by a lens mounted on an automated stage (AS) and the optical power is read by a power meter (PM).	44
3.7	Scatterplot of the fraction of optical power between the V and H states measured by selecting the polarization state at the output to be H . The grey vertical line separates the Haar random set from the switching set.	45
3.8	Scatterplots of the amplitude fidelity for the first experiment (panel (a)) and second experiment (panel (b)). The grey vertical line separates the Haar random set from the switching set.	46
3.9	Two different cross-talk matrices α , for a Clements-type structure (left) and a Bell-type structure (right); the off-diagonal elements represent cross-talk terms while the diagonal elements are self-heating terms. Both of these were obtained from the calibration of 6-mode UPPs. Empty rows and columns in these matrices correspond to heating elements that are not actually used during operation.	50
3.10	Histogram of deviations from 3 dB behavior for each coupler in a 6-mode. The average, marked by the dashed line, is 0.0056.	51
3.11	A comparison between the performance after automated calibration of 6, 8 and 12 mode UPPs.	51
3.12	A picture of the 6-mode UPP seen from the top.	52

3.13 The automated setup that I introduced to calibrate UPPs with fiber arrays at both input and output.	52
3.14 The waveguide structure of the 6-mode UPP; the location of each phase shifter is highlighted and numbered. Blue shifters correspond to those that were fabricated, but are not connected to the PCB; red shifters are those that are connected to the PCB and actually used during calibration and operation. Every crossing corresponds to a single 3 dB coupler.	53
3.15 Example of performance of the 6-mode UPP over the identity matrix (left) and a randomly selected Haar random matrix from the test set (right).	53
3.16 The 8-mode UPP.	54
3.17 The waveguide structure of the 8-mode UPP; the location of each phase shifter is highlighted and numbered. Every crossing corresponds to a single 3 dB coupler.	54
3.18 Example of performance of the 8-mode UPP over the identity matrix (left) and a randomly selected Haar random matrix from the test set (right).	55
3.19 The 12-mode UPP used for this example. On the left, a picture of the UPP seen from the top. On the right, the UPP slotted within a calibration and operation setup, compact enough to fit inside a drawer on a 70 cm by 94 cm server rack. . . .	55
3.20 The parallelized automated setup used for the calibration of the 12-mode UPP.	56
3.21 Example of performance of the 12-mode UPP over the identity matrix (left) and a randomly selected Haar random matrix from the test set (right).	56
3.22 The waveguide structure of the 12-mode UPP; the location of each phase shifter is highlighted. Every crossing corresponds to a single 3 dB coupler.	57
3.23 Histogram of deviations from 3 dB behavior for each coupler in the network. The average, marked by the dashed line, is 0.012.	57

3.24	The typical structure of the micromachined glass in our UPPs. Figure (a) shows a section of the bridge-and-trench structure in the plane orthogonal to light propagation; the darker dot is the waveguide core. Figure (b) shows a microscope picture from the top of a column of bridges in the UPP. The pink areas are the copper-chromium depositions, while the light gray areas are the portions where only chromium is present; this concentrates the resistance of the metal deposition on the bridge only, since chromium is more resistive than copper.	59
3.25	The setup used for the vacuum experiments.	60
3.26	A closeup of the UPP used for the vacuum experiments (left) and the inside of the vacuum chamber during assembly (right).	60
3.27	The crosstalk matrix of the 6-mode UPP in air (left) versus vacuum (right). The diagonal elements are larger in the case of vacuum, and the ratio between diagonal and off-diagonal elements is smaller.	61
3.28	A comparison between the fidelity in air and the fidelity in vacuum, using the same optimization algorithm and the same model for both.	61
3.29	Comparison between the amplitude fidelity versus epoch of the optimization models for the digital twin. While all fidelities eventually converge to unit value, the model in vacuum requires less epochs compared to air. If the model is made simpler, such as the case of a model that does not include cross talk, the model converges even faster.	62
3.30	A comparison between the performance of the calibration using the two models, one where the cross-talk is fitted by the model versus one where it isn't. Figure 3.30a shows the change in performance of the calibration in air, while Figure 3.30b shows the change in performance in vacuum; a closeup on the overlap region for Figure 3.30b is provided on the right for better visibility.	63
3.31	The Comsol simulation of the column of the UPP, while every heater is turned off, from the top (left) and in isometric view (right). There are a series of pillars that support the bridge structure. The scale is measured in terms of temperature difference with respect to the equilibrium.	64

3.32	The Comsol simulation of the column of the UPP in air, while the top heater is dissipating heat, from the top (left) and in isometric view (right). Dips in temperature can be seen where the pillars are transferring heat from the bridge to the bottom of the structure. The scale is measured in terms of temperature difference with respect to the equilibrium.	64
3.33	The simulated temperature profiles over time (a) in air and (b) in vacuum, for each of the six waveguides under consideration.	65
3.34	The values of the time constants for each waveguide while actuating only one resistive element versus actuating all phase shifters, (a) in air and (b) in vacuum.	66
3.35	The setup used to measure the angles of the complex elements of the implemented matrix. A portion of the laser light is divided into a reference branch and a UPP branch; a set of polarization controls is added on both branches of the setup. At the output of the UPP, the IIC is added together with a cylindrical lens and a camera.	70
3.36	A schematic representation of the waveguide design of the IIC (left) and a picture of the IIC in the setup (right).	70
3.37	A scheme representing the typical analysis of Young's double-slit experiment for waves. The two incoming coherent light beams (on the left) impinge on a pair of point-like slits separated by a distance d , impinging on a screen at distance $L \gg d$ (on the right). The x axis marks the position on the screen, with the origin $x = 0$ corresponding to the middle point between the slits.	72
3.38	The measured amplitudes and (amplitude) fidelities of each of the four selected matrices. Matrix (a) and (b) feature uniform submatrices, while (c) and (d) feature random submatrices. In this work, only the top left submatrices have been measured.	75
3.39	The intensity profiles as measured with the camera. Figure (a) shows the unedited picture from the camera, where the UPP is implementing a uniform 3-by-3 matrix. Figure (b) shows the raw profiles obtained from this image.	77
3.40	The Discrete Fourier Transform applied to the intensity profiles. From these graphs we can see that the phase of the signal depends strongly on the frequency in the region of interest.	78

List of Figures

4.1	Pictorial representation of a cylindrical waveguide. The darker region (the core) has higher refractive index than the outside (the cladding).	83
4.2	Theoretical number of guided modes and their cutoff dependence on wavelength for a step-index fiber with numerical aperture $NA = 0.13$, tuned for single-mode operation at 1550 nm. This graph was obtained by finding the solutions to the wave equations via numerical methods.	84
4.3	Schematic of the experimental setup including the scanning objective, the characterized fiber, the imaging objective, and the CCD camera.	88
4.4	Circular scanning pattern used to vary modal excitation. Each dot in the larger circle denotes an acquisition point. The automated stage stops on every point on the same row from left to right (along increasing x) before moving up to the next row (along increasing y); the red arrows indicate the additional movements that the stage performs in order to compensate backlash (not to scale).	89
4.5	Examples of images captured by the CCD camera, after binning. These images were obtained with laser light at 793 nm focused inside of the 1550BHP fiber.	89
4.6	Examples of principal components retrieved using PCA. Panel (a) is the first principal component; panels (b), (c), and (d) are respectively the third, tenth and twentieth principal components. In particular, panel (d) represents a principal component classified as noise term. These images were obtained with laser light at 793 nm focused inside of the 1550BHP fiber.	91
4.7	A graph of the eigenvalues retrieved using PCA, for both fibers, in logarithmic scale. Below each graph, the same curves restricted to just the first 100 eigenvalues.	92
4.8	Comparison between the obtained boundary points with the two methods, focused on a comparison between the two methods and the expected values.	92

4.9	Rotating phase field with $l = 1$. Figure (a) shows the typical profile of a rotating phase field supported by the waveguide, reconstructed from the principal components obtained on the 1550BHP waveguide at 925 nm. Figure (b) shows the target for the reconstruction. The colors represent the phase of the field (in the range $[-\pi, \pi)$); black pixels represent small amplitudes and colorful pixels large amplitudes.	93
4.10	The reconstruction fidelity evaluated using the principal components selected via LEV analysis, as a function of l , and normalized with respect to the $l = 0$ case for clarity. All the graphs show a plateau followed by a sharp drop in fidelity, which matches the expected maximum azimuthal angle supported by the guide.	96
5.1	The experimental use of two calibrated UPPs to implement adaptive boson sampling protocols. A 6-mode UPP is programmed to implement a sequence of conditional unitaries (a), where subsequent transformations are dynamically re-configured by detection events, enabling heralded qubit state preparation and tomography with high kernel fidelity. This approach is then scaled to an 8-mode UPP (b), implementing an adaptive architecture with up to 15 unitaries selected by the detection events, showcasing the ability of the UPPs to support feed-forward quantum photonic protocols. Figure adapted from Ref. [49].	101

List of Tables

3.1	The measured values of the angles for the two 2-by-2 matrices implemented by the UPP.	78
3.2	The measured values of the angles for the two 3-by-3 matrices implemented by the UPP.	79
3.3	The measured amplitude fidelity $\mathcal{F}_{\text{ampl}}$ versus the value of fidelity $\mathcal{F}$ obtained with this method.	79

Bibliography

- [1] E. Miller *et al.*, “Integrated optics: An introduction,” *The Bell System technical journal*, vol. 48, pp. 2059–2069, 1969.
- [2] R. G. Hunsperger, *Integrated Optics, Theory and Technology*. Springer, 6 ed., 2009.
- [3] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O’Brien, “A variational eigenvalue solver on a photonic quantum processor,” *Nature Communications*, vol. 5, p. 4213, 7 2014.
- [4] J. A. Dolphin, T. K. Paraíso, H. Du, R. I. Woodward, D. G. Marangon, and A. J. Shields, “A hybrid integrated quantum key distribution transceiver chip,” *npj Quantum Information*, vol. 9, 9 2023.
- [5] H. Mahmudlu, R. Johanning, A. Van Rees, A. K. Kashi, J. P. Epping, R. Haldar, K.-J. Boller, and M. Kues, “Fully on-chip photonic turnkey quantum source for entangled qubit/qudit state generation,” *Nature Photonics*, vol. 17, pp. 518–524, 4 2023.
- [6] M. Reck *et al.*, “Experimental realization of any discrete unitary operator,” *Physical Review Letters*, vol. 73, pp. 58–61, 1994.
- [7] W. R. Clements *et al.*, “Optimal design for universal multiport interferometers,” *Optica*, vol. 3, pp. 1460–1465, 2016.
- [8] B. A. Bell and I. A. Walmsley, “Further compactifying linear optical unitaries,” *APL Photonics*, vol. 6, no. 7, 2021.
- [9] C. Taballione *et al.*, “8x8 reconfigurable quantum photonic processor based on silicon nitride waveguides,” *Opt. Express*, vol. 27, pp. 26842–26857, 9 2019.
- [10] P. R. Halmos, *Measure theory*. Springer, 12 2013.
- [11] F. Hoch, S. Piacentini, T. Giordani, Z.-N. Tian, M. Iuliano, C. Esposito, A. Camillini, G. Carvacho, F. Ceccarelli, N. Spagnolo, A. Crespi, F. Sciarrino, and R. Osellame, “Boson sampling in a reconfigurable continuously-coupled 3d photonic circuit,” 2021.
- [12] F. Ceccarelli, S. Atzeni, C. Pentangelo, F. Pellegatta, A. Crespi, and R. Osellame, “Low power reconfigurability and reduced crosstalk in integrated photonic circuits fabricated by femtosecond laser micromachining,” *Laser & Photonics Reviews*, vol. 14, no. 10, p. 2000024, 2020.

- [13] G. Corrielli, S. Atzeni, S. Piacentini, I. Pitsios, A. Crespi, and R. Osellame, “Symmetric polarization-insensitive directional couplers fabricated by femtosecond laser writing,” *Optics Express*, vol. 26, pp. 15101–15109, June 2018.
- [14] G. T. Reed, G. Mashanovich, F. Y. Gardes, and D. J. Thomson, “Silicon optical modulators,” *Nature Photonics*, vol. 4, pp. 518–526, Aug. 2010.
- [15] F. Flamini, L. Magrini, A. S. Rab, N. Spagnolo, V. D’Ambrosio, P. Mataloni, F. Sciarrino, T. Zandrini, A. Crespi, R. Ramponi, and R. Osellame, “Thermally reconfigurable quantum photonic circuits at telecom wavelength by femtosecond laser micromachining,” *Light: Science & Applications*, vol. 4, no. 11, pp. e354–e354, 2015.
- [16] R. Albiero, C. Pentangelo, M. Gardina, S. Atzeni, F. Ceccarelli, and R. Osellame, “Toward higher integration density in femtosecond-laser-written programmable photonic circuits,” *Micromachines*, vol. 13, no. 7, p. 1145, 2022.
- [17] F. Ceccarelli, S. Atzeni, A. Prencipe, R. Farinaro, and R. Osellame, “Thermal phase shifters for femtosecond laser written photonic integrated circuits,” *Journal of Lightwave Technology*, vol. 37, no. 17, pp. 4275–4281, 2019.
- [18] M. Conforti, G. Cornuéjols, and G. Zambelli, *Integer Programming*. 1 2014.
- [19] F. Shokraneh, S. Geoffroy-Gagnon, and O. Liboiron-Ladouceur, “Towards phase-error- and loss-tolerant programmable MZI-based optical processors for optical neural networks,” in *2020 IEEE Photonics Conference (IPC)*, pp. 1–2, 2020.
- [20] B.-X. Ke, K.-H. Lo, D.-W. Tsai, and M.-C. M. Lee, “Non-invasively monitoring tunable 2-by-2 Mach-Zehnder interferometers in programmable Si photonic integrated circuits,” 10 2024.
- [21] M. S. Hai, M. M. P. Fard, D. An, F. Gambini, S. Faralli, G. B. Preve, G. W. Roberts, and O. Liboiron-Ladouceur, “Automated characterization of SiP MZI-based switches,” in *2015 IEEE Optical Interconnects Conference (OI)*, pp. 94–95, 2015.
- [22] C. Alexiev, J. C. C. Mak, W. D. Sacher, and J. K. S. Poon, “Calibrating rectangular interferometer meshes with external photodetectors,” *OSA Continuum*, vol. 4, no. 11, pp. 2892–2904, 2021.
- [23] D. Pérez-López, A. López, P. DasMahapatra, and J. Capmany, “Multipurpose self-configuration of programmable photonic circuits,” *Nature Communications*, vol. 11, no. 1, p. 6359, 2020.
- [24] R. Hamerly, S. Bandyopadhyay, and D. Englund, “Accurate Self-Configuration of rectangular multiport interferometers,” *Physical Review Applied*, vol. 18, 8 2022.
- [25] A. Fyrrillas, O. Faure, N. Maring, J. Senellart, and N. Belabas, “Scalable machine learning-assisted clear-box characterization for optimally controlled photonic circuits,” *Optica*, vol. 11, p. 427, 2 2024.
- [26] J. Carolan, C. Harrold, C. Sparrow, E. Martín-López, N. J. Russell, J. W. Silverstone, P. J. Shadbolt, N. Matsuda, M. Oguma, M. Itoh, *et al.*, “Universal linear optics,” *Science*, vol. 349, no. 6249, pp. 711–716, 2015.
- [27] D. G. Michiel, H. Snijders, P. Venderbosch, B. Kassenberg, N. Kannan, D. H. Smith, C. Taballione, J. P. Epping, H. Van Den Vlekkert, and J. J. Renema, “High fidelity 12-Mode Quantum Photonic processor operating at INGAAS Quantum Dot Wavelength,” 4 2022.
- [28] C. Taballione, M. C. Anguita, M. De Goede, P. Venderbosch, B. Kassenberg, H. Snijders, N. Kannan, W. L. Vleeshouwers, D. Smith, J. P. Epping, R. Van Der Meer, P. W. H. Pinkse, H. Van Den Vlekkert, and J. J. Renema, “20-Mode Universal Quantum Photonic Processor,” *Quantum*, vol. 7, p. 1071, 8 2023.

- [29] N. Hoppe, W. S. Zaoui, L. Rathgeber, Y. Wang, R. H. Klenk, W. Vogel, M. Kaschel, S. L. Portalupi, J. Burghartz, and M. Berroth, "Ultra-efficient silicon-on-insulator grating couplers with backside metal mirrors," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 26, no. 2, pp. 1–6, 2020.
- [30] D. Llewellyn, Y. Ding, I. I. Faruque, S. Paesani, D. Bacco, R. Santagati, Y.-J. Qian, Y. Li, Y.-F. Xiao, M. Huber, M. Malik, G. F. Sinclair, X. Zhou, K. Rottwitt, J. L. O'Brien, J. G. Rarity, Q. Gong, L. K. Oxenlowe, J. Wang, M. G. Thompson, D. Llewellyn, Y. Ding, I. I. Faruque, S. Paesani, D. Bacco, R. Santagati, Y.-J. Qian, Y. Li, Y.-F. Xiao, M. Huber, M. Malik, G. F. Sinclair, X. Zhou, K. Rottwitt, J. L. O'Brien, J. G. Rarity, Q. Gong, L. K. Oxenlowe, J. Wang, and M. G. Thompson, "Chip-to-chip quantum teleportation and multi-photon entanglement in silicon," *Nature Physics*, vol. 16, pp. 148–153, 12 2019.
- [31] R. Santagati, J. Wang, A. A. Gentile, S. Paesani, N. Wiebe, J. R. McClean, S. Morley-Short, P. J. Shadbolt, D. Bonneau, J. W. Silverstone, D. P. Tew, X. Zhou, J. L. O'Brien, and M. G. Thompson, "Witnessing eigenstates for quantum simulation of Hamiltonian spectra," *Science Advances*, vol. 4, p. eaap9646, 1 2018.
- [32] A. Barzaghi, M. Bénéfice, F. Ceccarelli, G. Corrielli, V. Galli, M. Gardina, V. Grimaldi, J. Kaczorowski, F. Malaspina, R. Osellame, C. Pentangelo, A. Rocchetto, and A. Rudi, "A low-loss, 24-mode laser-written universal photonic processor in a glass-based platform," *arXiv (Cornell University)*, 5 2025.
- [33] J. A. Nelder and R. Mead, "A simplex method for function minimization," *The Computer Journal*, vol. 7, pp. 308–313, 1 1965.
- [34] G. Corrielli, A. Crespi, R. Geremia, R. Ramponi, L. Sansoni, A. Santinelli, P. Mataloni, F. Sciarrino, and R. Osellame, "Rotated waveplates in integrated waveguide optics," *Nature Communications*, vol. 5, p. 4249, June 2014.
- [35] R. Burgwal, W. R. Clements, D. H. Smith, J. C. Gates, W. S. Kolthammer, J. J. Renema, and I. A. Walmsley, "Using an imperfect photonic network to implement random unitaries," *Optics Express*, vol. 25, no. 23, pp. 28236–28245, 2017.
- [36] C. Taballione, R. Van Der Meer, H. J. Snijders, P. Hooijschuur, J. P. Epping, M. De Goede, B. Kassenberg, P. Venderbosch, C. Toebe, H. Van Den Vlekkert, P. W. H. Pinkse, and J. J. Renema, "A universal fully reconfigurable 12-mode quantum photonic processor," *Materials for Quantum Technology*, vol. 1, p. 035002, 7 2021.
- [37] C. Taballione, M. C. Anguita, M. de Goede, P. Venderbosch, B. Kassenberg, H. Snijders, N. Kannan, W. L. Vleeshouwers, D. Smith, J. P. Epping, *et al.*, "20-mode universal quantum photonic processor," *Quantum*, vol. 7, p. 1071, 2023.
- [38] J. Yufeng and Z. Jiaxun, "Mems vacuum packaging technology and applications," in *2005 6th International Conference on Electronic Packaging Technology*, pp. 1–5, 2005.
- [39] S. Rahimi-Keshari, M. A. Broome, R. Fickler, A. Fedrizzi, T. C. Ralph, and A. G. White, "Direct characterization of linear-optical networks," *Opt. Express*, vol. 21, pp. 13450–13458, 6 2013.
- [40] R. Heilmann, M. Gräfe, S. Nolte, and A. Szameit, "A novel integrated quantum circuit for high-order W-state generation and its highly precise characterization," *Science Bulletin*, vol. 60, no. 1, pp. 96–100, 2015.
- [41] F. Hoch, T. Giordani, N. Spagnolo, A. Crespi, R. Osellame, and F. Sciarrino, "Characterization of multimode linear optical networks," *Advanced Photonics Nexus*, vol. 2, no. 1, p. 016007, 2023.
- [42] A. Laing and J. L. O'Brien, "Super-stable tomography of any linear optical device," 2012.

- [43] I. Dhand, A. Khalid, H. Lu, and B. C. Sanders, “Accurate and precise characterization of linear optical interferometers,” *Journal of Optics*, vol. 18, p. 035204, feb 2016.
- [44] I. Cristiani, C. Lacava, G. Rademacher, B. J. Puttnam, R. S. Luís, C. Antonelli, A. Mecozzi, M. Shtaif, D. Cozzolino, D. Bacco, L. K. Oxenløwe, J. Wang, Y. Jung, D. J. Richardson, S. Ramachandran, M. Guasoni, K. Krupa, D. Kharenko, A. Tonello, S. Wabnitz, D. B. Phillips, D. Faccio, T. G. Euser, S. Xie, P. S. J. Russell, D. Dai, Y. Yu, P. Petropoulos, F. Gardes, and F. Parmigiani, “Roadmap on multimode photonics,” *Journal of Optics*, vol. 24, p. 083001, jul 2022.
- [45] W. Watanabe, Y. Note, and K. Itoh, “Fabrication of multimode interference waveguides in glass by use of a femtosecond laser,” *Optics Letters*, vol. 30, p. 2888, 11 2005.
- [46] N. K. Fontaine, R. Ryf, M. A. Mestre, B. Guan, X. Palou, S. Randel, Y. Sun, L. Grüner-Nielsen, R. V. Jensen, and R. Lingle, “Characterization of space-division multiplexing systems using a swept-wavelength interferometer,” in *Optical Fiber Communication Conference/National Fiber Optic Engineers Conference 2013*, p. OW1K.2, Optica Publishing Group, 2013.
- [47] J. Carpenter, B. J. Eggleton, and J. Schröder, “110x110 optical mode transfer matrix inversion,” *Optics Express*, vol. 22, pp. 96–101, 2014.
- [48] I. T. Jolliffe, *Principal Component Analysis*. New York, NY, USA: Springer, 2nd ed., 2002.
- [49] F. Hoch, E. Caruccio, G. Rodari, T. Francalanci, A. Suprano, T. Giordani, G. Carvacho, N. Spagnolo, S. Koudia, M. Proietti, C. Liorni, F. Cerocchi, R. Albiero, N. Di Giano, M. Gardina, F. Ceccarelli, G. Corrielli, U. Chabaud, R. Osellame, M. Dispenza, and F. Sciarrino, “Quantum machine learning with Adaptive Boson Sampling via post-selection,” *Nature Communications*, vol. 16, p. 902, 1 2025.
- [50] E. Caruccio, D. Maragnano, G. Rodari, D. Picus, G. Garberoglio, D. Binosi, R. Albiero, N. Di Giano, F. Ceccarelli, G. Corrielli, N. Spagnolo, R. Osellame, M. Dapor, M. Liscidini, and F. Sciarrino, “Experimental verification of threshold quantum state tomography on a fully-reconfigurable photonic integrated circuit,” *npj Quantum Information*, vol. 11, 11 2025.
- [51] L. Monbroussou, B. Polacchi, G. Carvacho, T. Giordani, M. Bossi, A. Rajan, N. Di Giano, R. Osellame, and E. Kashefi, “Photonic quantum convolutional neural networks with adaptive state injection,” *Advanced Photonics*, vol. 7, 11 2025.
- [52] G. Rodari, C. Fernandes, E. Caruccio, A. Suprano, F. Hoch, T. Giordani, G. Carvacho, R. Albiero, N. Di Giano, G. Corrielli, F. Ceccarelli, R. Osellame, D. J. Brod, L. Novo, N. Spagnolo, E. F. Galvão, and F. Sciarrino, “Experimental observation of counter-intuitive features of photonic bunching,” 10 2024.
- [53] G. Rodari, T. Francalanci, E. Caruccio, F. Hoch, G. Carvacho, T. Giordani, N. Spagnolo, R. Albiero, N. Di Giano, F. Ceccarelli, G. Corrielli, A. Crespi, R. Osellame, U. Chabaud, and F. Sciarrino, “Observation of Lie algebraic invariants in Quantum Linear Optics,” 5 2025.
- [54] F. Hoch, A. Camillini, G. Rodari, E. Caruccio, G. Carvacho, T. Giordani, R. Albiero, N. Di Giano, G. Corrielli, F. Ceccarelli, R. Osellame, M. Robbio, L. Novo, N. Spagnolo, E. F. Galvão, and F. Sciarrino, “Optimal distillation of photonic indistinguishability,” 9 2025.