Diffractive coupling for optical neural networks

Sheler Maktoobi

THÈSE DE DOCTORAT DE L'ÉTABLISSEMENT UNIVERSITÉ BOURGOGNE FRANCHE-COMTÉ

PRÉPARÉE À L'UNIVERSITÉ DE FRANCHE-COMTÉ

École doctorale n°37

Sciences Pour l'Ingénieur et Microtechniques

Doctorat d'Optique

par

SHELER MAKTOOBI

Couplage Diffractif Pour Réseaux de Neurones Optiques

Diffractive Coupling for Optical Neural Networks

Thèse présentée et soutenue à Besançon, le 21 juillet 2020

Composition du Jury :

BARBAY SYLVAIN	Directeur de recherche au CNRS, Université Paris-Saclay, Palaiseau	Rapporteur
ALONSO MIGUEL	Professeur, Université Rochester, Rochester (USA)/Professeur, Université Aix-Marseille, Marseille (France)	Rapporteur
RONTANI DAMIEN	Maître de Conférences, Ecole Supérieure d'Electricité, Metz	Examinateur
BRUNNER DANIEL	Chargé de recherche au CNRS, Institut FEMTO-ST, Besançon	Directeur
JACQUOT MAXIME	Professeur, Université Bourgogne Franche-Comté, Besançon	Codirecteur
FROEHLY LUC	Chargé de recherche au CNRS, Institut FEMTO-ST, Besançon	Codirecteur

N° | X | X | X |

école doctorale **sciences pour l'ingénieur et microtechniques**

Titre : Couplage Diffractif Pour Réseaux de Neurones Optiques

Résumé :

Les réseaux photoniques à haute performance peuvent être considérés comme des supports pour les futurs systèmes de calcul. Contrairement à l'électronique, les systèmes photoniques offrent des avantages intéressants, par exemple la possibilité de réaliser des réseaux complètement parallèles. Récemment, les réseaux de neurones ont attiré l'attention de la communauté photonique. L'une des difficultés les plus importantes, en matière de réseaux photoniques parallèles à grande échelle, est la réalisation des connexions. La diffraction est exploitée ici comme méthode pour traiter les connexions entre les nœuds (couplage) dans les réseaux de neurones optiques. Dans cette thèse, nous étudions l'extensibilité d'un couplage diffractif en détails de la façon suivante :

Tout d'abord, nous commençons par une introduction générale à propos de l'intelligence artificielle, de l'apprentissage machine, des réseaux de neurones artificiels et des réseaux de neurones photoniques. Lors de la conception d'un réseau neuronal fonctionnel, les règles de l'apprentissage machine sont des éléments essentiels pour optimiser une configuration et ainsi obtenir une faible erreur du système, donc les règles de l'apprentissage sont introduites (chapitre 1). Nous étudions les concepts fondamentaux du couplage diffractif dans notre réservoir spatio-temporel. Dans ce cas, la théorie de la diffraction est expliquée. Nous utilisons un schéma analytique pour fournir les limites en termes de taille des réseaux diffractifs, qui font partie intégrante de notre réseau neuronal photonique (chapitre 2). Les concepts du couplage diffractif sont étudiés expérimentalement dans deux expériences différentes afin de confirmer les limites obtenues analytiquement, et pour déterminer le nombre maximum de nœuds pouvant être couplés dans le réseau photonique (Chapitre 3). Les simulations numériques d'une telle expérience sont basées sur deux schémas différents pour calculer numériquement la taille maximale du réseau, qui approche une surface de 100 mm^2 (chapitre 4). Enfin, l'ensemble du réseau neuronal photonique est démontré. Nous concevons un réservoir spatialement étendu sur 900 nœuds. En conséquence, notre système généralise la prédiction pour la séquence chaotique de Mackey-Glass (chapitre 5).

Mots-clés : Réseaux optiques, Diffraction, Couplage, Réseaux neuronaux photoniques.

Title: Diffractive Coupling for Optical Neural Networks

Abstract:

Photonic networks with high performance can be considered as substrates for future computing systems. In comparison with electronics, photonic systems have substantial privileges, for instance the possibility of a fully parallel implementation of networks. Recently, neural networks have moved into the center of attention of the photonic community. One of the most important requirements for parallel large-scale photonic networks is to realize the connectivities. Diffraction is considered as a method to process the connections between the nodes (coupling) in optical neural networks. In the current thesis, we evaluate the scalability of a diffractive coupling in more details as follow:

First, we begin with a general introductions for artificial intelligence, machine learning, artificial neural network and photonic neural networks. To establish a working neural network, learning rules are an essential part to optimize a configuration for obtaining a low error from the system, hence learning rules are introduced (Chapter 1). We investigate the fundamental concepts of diffractive coupling in our spatio-temporal reservoir. In that case, theory of diffraction is explained. We use an analytical scheme to provide the limits for the size of diffractive networks which is a part of our photonic neural network (Chapter 2). The concepts of diffractive coupling are investigated experimentally by two different experiments to confirm the analytical limits and to obtain maximum number of nodes which can be coupled in the photonic network (Chapter 3). Numerical simulations for such an experimental setup is modeled in two different schemes to obtain the maximum size of network numerically, which approaches a surface of 100 mm^2 (Chapter 4). Finally, the complete photonic neural network is demonstrated. We design a spatially extended reservoir for 900 nodes. Consequently, our system generalizes the prediction for the chaotic Mackey–Glass sequence (Chapter 5).

Keywords: Optical networks, Diffraction, Coupling, Photonic neural networks.

UBFC UNIVERSITÉ BOURGOGNE FRANCHE-COMTÉ

Université Bourgogne Franche-Comté
32, avenue de l'Observatoire
25000 Besançon, France

ACKNOWLEDGMENTS

I would like to thank the following people who have helped me throughout the last three years:
My supervisors Dr. Daniel Brunner, Dr. Luc Froehly and Dr. Maxime Jacquot for their enthusiasm in this project, consistent supports and encouragement.
The Institute of Femto-St, Bourgogne Franche-Comté University, for all staff members, research group members in optics department, director of institute Laurent Larger, head of doctoral school Thérèse Leblois, and also Hervé Maillotte who are always energetic.
Aurea technology with its fabulous director Dr. Johann Cussey and friendly engineers made such an inspiring experience for me in a short period of collaboration during my thesis.
My colleagues Louis Andreoli, Xavier Porte, Nadezhda Semenova and Vladimir Semenova, Bogdan Penkovsky, Bicky Marquez, Julian Bueno, and my friends Asma Chargui, Laetitia Raguin, Solveig Perret, Vincent Pecheur, Rémi Meyer, Cyril Millon, Florian Behague, Coraline Lapre, Thomas Daugey, Clément Eustache, Amar Nath Ghosh, Nicolas Laforge and Kirthanaa Indumathi for all your collaborative effort and your generosity over my researches.
And my best friend Benoit Morel who helped me sincerely along this way.

My kind sister and my wonderful brothers,
my extraordinary parents,
you set me off on the road to this phd thesis a long time ago,
special thanks.

Contents

Introduction 1

0.1 Introduction . . . 1
0.2 Structure of the thesis . . . 2
0.3 Publications . . . 4
0.4 Conference presentations . . . 4
0.5 poster presentations . . . 4

1 Optical neural networks 5

1.1 Neural networks . . . 5
1.1.1 Introduction . . . 5
1.1.1.1 Perceptron . . . 6
1.1.2 Artificial neural networks . . . 6
1.1.2.1 Feedforward neural networks . . . 7
1.1.2.2 Recurrent neural network . . . 7
1.1.3 Neural networks training . . . 8
1.1.3.1 Supervised learning . . . 9
1.1.3.2 Feedforward neural network training . . . 9
1.1.3.3 Recurrent neural network training . . . 9
1.1.3.4 Unsupervised learning . . . 10
1.1.4 Reservoir computing . . . 11
1.2 Photonic neural networks . . . 11
1.2.1 Photonic delay reservoirs . . . 12
1.2.2 Photonic spatio-temporal reservoirs . . . 13
1.3 Conclusion . . . 14

2 Diffractive coupling 15

2.1 Theory of diffraction . . . 15
2.1.1 Maxwell's equations . . . 15
2.1.2 Propagation equation . . . 16
2.1.3 Scalar propagation equation . . . 17

2.1.4 Helmholtz equation . . . 17
2.1.4.1 Propagation of plane-waves . . . 17
2.1.4.2 Velocity of light in a medium . . . 19
2.1.5 Huygens principle . . . 20
2.1.6 Rayleigh-Sommerfeld integral . . . 21
2.1.7 Fresnel diffraction integral . . . 22
2.1.8 Fraunhofer diffraction integral . . . 23
2.2 Plane-wave spectrum . . . 23
2.2.1 Steps of plane-wave spectrum for propagation . . . 24
2.2.2 Example: diffraction by gratings . . . 28
2.2.2.1 The optical grating in ray-optics . . . 28
2.3 Principle of diffractive coupling . . . 31
2.4 Conclusion . . . 35

3 Experimental characterization of diffractive coupling 37
3.1 Translate point source along the (x, y)-plane . . . 37
3.1.1 Experimental setup . . . 37
3.1.1.1 Image magnification . . . 38
3.1.1.2 Impact of DOE . . . 39
3.1.1.3 Vignetting . . . 41
3.1.1.4 Impact of the distance between two microscope objectives 42
3.1.1.5 Reduce distance between MOs . . . 44
3.1.2 Results with emulating optical emitters . . . 44
3.1.2.1 Impact of beam vignetting . . . 45
3.2 Options of solving the restriction of beam vignetting . . . 47
3.3 Emulate emitter positions through DOE-tilting . . . 49
3.3.1 Experimental setup . . . 49
3.3.1.1 Coupling mismatch $|\Delta|$ . . . 50
3.3.2 Limitation of the setup . . . 52
3.4 conclusion . . . 52

4 Numerical model 55
4.1 Beam propagation . . . 56
4.1.1 Absorbing boundary . . . 56
4.2 Microscope objectives . . . 58
4.2.1 Debye integral . . . 59

4.2.1.1 Numerical implementation . . . 61
4.3 Re-scaling of image plane . . . 62
4.3.1 Coordinates under paraxial approximations . . . 62
4.3.2 Non-paraxial coordinates . . . 63
4.4 The DOE . . . 64
4.5 Numerical implementation . . . 66
4.5.1 Using low magnification microscope objectives with a low numerical aperture . . . 67
4.5.2 Results . . . 67
4.5.2.1 Distance between diffractive orders . . . 67
4.5.2.2 Distance mismatch . . . 68
4.5.3 Using low magnification microscope objectives with a large numerical aperture . . . 69
4.5.4 Results . . . 70
4.5.4.1 Distance between two diffractive orders . . . 70
4.5.4.2 Distance mismatch . . . 71
4.5.4.3 Width of diffractive orders . . . 71
4.5.4.4 Distance mismatch with and without re-scaling the image planes . . . 72
4.6 Conclusion . . . 74

5 Photonic neural network 75
5.1 Introduction to recurrent neural networks . . . 75
5.2 Experimental setup and diffractive network . . . 75
5.2.1 Readout . . . 78
5.2.2 Photonic learning . . . 79
5.3 Conclusion . . . 80

6 Conclusion and Outlook 81
6.1 Conclusion . . . 81
6.2 Outlook . . . 83

Bibliography 85

List of figures 93

INTRODUCTION

0.1/ INTRODUCTION

Computing was interesting for human before the technological revolution. One of the first computers was the Antikythera-mechanism [1]. That mechanism calculates moon phases and planetary position within the lunisolar calendar. In the Antikythera hardware, there is an algorithm about lunar and planetary motion which was realized by Hipparchos of Rhodes. This mechanical computer was a great challenge since it consists of 37 bronze gears and this multistage system required high manufacturing accuracy for it's reliable operation. This system needed significant skills which were obtained via human progress in agriculture, religious, the planets and lunar cycles in that historical period. This shows progress in computing since two millennia ago. In 1941 a re-programmable digital electronic-computer $Z3$ was realized for the practical re-programmability benefits. As an important challenge the $Z3$ was applied in the mechanical properties of airplane wings. The transition from mechanics to electronic-computer set the basis for the current, Turing complete, integrated and powerful computing processors. The developments of technology and concepts therefore have created toady's powerful computers. In that case, one considers another transition from electronics to photonics or merger of both. Nowadays, people are familiar with typical electronic computers, while the optical ones are still relatively unknown, even by representing over five decades ago [2]. For instance an optical experiment was created by Farhat et al [3] which was an optical content addressable memory. That system was able to provide the content of stored memory even if the original question is noise-corrupted or not accurate.

Different type of experiments such as electrical, electro-optical and all optical computing demonstrate benefits and drawbacks of each hardware. For instance, Lohmann in 1990 discussed with the fundamental requirements to compute and their connection to physical properties of their hardware substrates [4]. Then, he identified some important ingredients such as communication or logic for binary logic operations. Electrons are able to implement nonlinear transformations very well, but the strong interaction between electrons decrease their possibilities for information transmission. Moreover, modulation bandwidth of electronic transmission is limited by the information carries' electrical charge. But transmission of information using photons is possible without strong interaction, which for instance is important in the implementation of artificial neural networks (ANNs) with large scale connectivities [5]. In addition, modulation bandwidth is not limited by the transmission length. Pockels and Kerr effects can be used as the optical nonlinearities [6]. In another example, the bandwidth of an electro-optical modulation can arrive at Terra Hertz [7].

Optical logic for computing has received substantial attention because of the success of computing based on Boolean logic operations. Replacement of electrons with photons

has three important reasons. First, by using photons, one can avoid the opto-electronic conversions since almost all communication is today done optically. Second, the potentially high bandwidth of optical processes allows to achieve pico second switching times. While today's transistors has the same switching time, in the global system the bandwidth limitation due to electronics interconnects strongly limits the overall speed. Third, parallel and spatially distribution of optical signal transmission is another important factor in optical computers, which are discussed intensely today [8]. For instance, the article by Caulfield et al. [9] or Miller et al. [10] highlight the optical transistors as principal units for real world applications and for all optical computing logic.

Another approach for optical computing is based on spatial transformations. For instance, an optical lens is able to transform optical waves fully in parallel and in a large space bandwidth. Compared with serial optical binary logic, spatial transformations in parallel can be considered as a significant advantage. Another example is realizing the subtraction of two images by using Michelson interferometry [11]. After creating spatial light modulators for the first time [12], the technology of this optical component brings a wide range of applications for such optical systems.

Photonic neural networks are an attractive field because they have caused the potential for next-generation computing. Today neural networks have a revolution in computing. One of the most important points is that information can propagate between the network's nodes based on parallel communication [8]. In this work, we study a photonic system which is a part of a neural network. In the reduced interpretation, the neurons and their connectivities are the important parts of these networks, hence we intend to obtain the maximum size of this neural network in which the connectivities between the nodes are formed based on diffraction.

0.2/ STRUCTURE OF THE THESIS

In this thesis we investigate fundamental and practical limits to the size of diffractive coupling for a spatio-temporal neural network as follow:
In Chapter 1, we describe neural networks as inspired by human brain and the concept of a biological neuron. We present a brief history of development of artificial neural networks which have been in the center of researches attentions. Based on the structure of the connections, two classes of neural networks (Feedforward and Recurrent) are explained. To optimize the configuration in neural network, we require learning rules. Hence, we will explain two classes of leaning (supervised and unsupervised) which is followed by a discussion about training two different neural network typologies (Feedforward and Recurrent). Since the two discussed training strategies are slow and difficult in their algorithms, reservoir computing is proposed. We will introduce two type of reservoirs, photonic delay reservoirs and photonic spatio-temporal reservoir. An experimental example is given for photonic delay reservoir but a photonic spatio-temporal reservoir will be studied in the fifth Chapter of this thesis as our neural network.

In Chapter 2, connectivity for a large number of photonic nodes is investigated. To find the limitation for diffractive networks' size, diffractive coupling will be studied. For the

goal, theory of diffraction is needed. In that case, we begin from Maxwell equations to obtain the diffraction theory. During this Chapter, we represent the important concepts and equations for plane waves, the Huygens principle, the Fresnel integral and the plane-wave spectrum. Beam propagation between optical elements is computed by using the plane-wave spectrum. The Fourier transform and inverse Fourier transform will be introduced during this propagation part. Finally, diffraction of light, transmission and reflection gratings are described. We continue by defining the principle of diffractive coupling which creates connectivities between the neurons in a photonic neural network. We utilize a single diffractive optical element (DOE) to evaluate coupling and analyze the coupling by using simple analytical calculations. The result of the analytical simulation shows the maximum size of neural network in which neurons can be successfully coupled. As a result, maximum number of neurons are obtained.

In Chapter 3, two simple experiments are introduced to confirm the diffractive coupling experimentally. First, all optical elements in this experiment are introduced in detail. Then, in the first experiment, a single mode optical fiber emulates optical emitters of a network. In this method, there are some limitations originating from the components (i.e. microscope objectives causing beam vignetting). To overcome the restrictions of the first experiment, we have iterated and tested a large number of optical components, but the effect is continued to be present. Finally, we therefore introduce the second scheme, where we tilt the DOE, which allows to go beyond the mentioned limitations. We compare our experimental data with analytical simulation to confirm the obtained results from this Chapter.

In Chapter 4, we numerically model the system's full beam propagation, diffraction and collimation/imaging to observe the effects found in the experiment. To implement all relevant effects which could limit our approach, paraxial approximation is avoided. The applied DOE which is based on phase-shifts was characterized using phase-retrieval measurements of the element. We use the Debye-approximation for simulating the imaging microscope objectives. In this Chapter two setups are presented. In first setup, low magnification microscope objective with low numerical aperture were used. The result of this scheme will be compared with the first experiment. In this scheme, we see the vignetting effects from microscope objectives. To solve this problem, we developed the second setup which has low magnification microscope objective with a high numerical aperture. We make a comparison between second experiment, second numerical scheme and analytical simulation. The good agreement between analytical, numerical and experimental data points shows a successful coupling.

In Chapter 5, our optical neural network will be introduced. First, all optical components are described. We obtain the result of network coupling in a matrix for 900 nodes in this experiment. Then, we implement readout connection by using a digital micro-mirror device. Finally, learning prediction evaluates for Mackey-Glass equation.

0.3/ PUBLICATIONS

- S. Maktoobi, L. Froehly, L. Andreoli, J. Porte, M. Jacquot, L. Larger, and D. Brunner ; "Diffractive Coupling For Photonic Networks: How Big Can We Go", IEEE Journal of Selected Topics in Quantum Electronics, Volume 26, Issue (1), July 2019.
- D. Brunner, J. Bueno, J. Porte, S. Maktoobi and L. Andreoli ; "Photonic Reservoir Computing", De Gruyter, pp. 83-115, July 2019.
- J. Bueno, S. Maktoobi, L. Froehly, I. Fischer, M. Jacquot, L. Larger, and D. Brunner ; "Reinforcement Learning in a large scale photonic Recurrent Neural Network", Optica (OSA), vol.5, pp 756-760, June 2018.

0.4/ CONFERENCE PRESENTATIONS

- D. Brunner, S. Maktoobi, L. Andreoli, J. Porte, M. Jacquot, S. Reitzenstein, L. Larger ; "Limits and applications of diffractive coupling", OSA Mathematics in Imaging, Jun 2019, Munich, Germany.
- D. Brunner, L. Andreoli, S. Maktoobi, J. Bueno, M. Jacquot, I. Fischer, L. Larger ; "Reinforcement learning in a large scale photonic network", SPIE Photonics West, Feb 2019, San Francisco, United States.
- J. Bueno, S. Maktoobi, L. Froehly, I. Fischer, M. Jacquot, L. Larger, D. Brunner ; "Large scale spatio-temporal networks of nonlinear oscillators for neuromorphic computing", International Symposium on Nonlinear Theory and its Application, 2018, Tarragona, Spain.
- S. Maktoobi, L. Andreoli, L. Larger, M. Jacquot, D. Brunner ; "Diffractive coupling for large scale photonic Reservoir Computer", OSA Advanced Photonics Conference, 2-5th July 2018, Zurich, Switzerland.
- S. Maktoobi, L. Froehly, M. Jacquot, D. Brunner ; "Reinforcement learning in a large scale photonic network", Conference on Lasers and Electro-Optics, 2018, San Jose, United States.
- S. Maktoobi, L. Froehly, M. Jacquot, D. Brunner ; "Diffractive Coupling in Optical Network", Photonics Day, 1st December 2017, Dijon, France.

0.5/ POSTER PRESENTATIONS

- S. Maktoobi, L. Andreoli, L. Froehly, M. Jacquot, D. Brunner ; "Scalable Optical Neural Network via Diffractive Coupling", Cognitive Computing - merging concepts with hardware, 2018, Hannover, Germany.
- S. Maktoobi, L. Froehly, M. Jacquot, D. Brunner ; "Diffractive Coupling for Optical Neural Network", 6th International symposium in Optics and its Applications, 17-20th February 2018, Trento, Italy. **(Best Poster Presentation Award)**.
- S. Maktoobi, L. Andréoli, S. Reitzenstein, L. Froehly, M. Jacquot, D. Brunner ; " Neuromorphic Computing using QD-Networks NeuroQNet", Workshop BioComp Summer School, 25th June –1st July 2017, Roscoff, France.

1

OPTICAL NEURAL NETWORKS

1.1/ NEURAL NETWORKS

1.1.1/ INTRODUCTION

Artificial Intelligence (AI), as a branch of computer science, has been developed based on a very rough approximate of the most fundamental properties of the human brain [13, 14]. In that case, human's intelligence are defined in a specific way and machines are able to mimic this way. Artificial intelligence can be used in different tasks for instance in health care [15] or finance [16]. There are two classifications for AI, the first class is weak AI, which carries on the simple tasks while, the second class, strong AI, tends to solve more complex tasks.
Machine learning is a subset of AI. In machine learning, the machines are not programmed but learn from the past data. To obtain an accurate prediction, it is essential to know how an logarithm learns. Machine learning is classified into the two categories of supervised learning and unsupervised learning which will be explained more in details in the following sections.
It was in 1943 that Mc Culloch (neuro-physiologist) and Pitts (logician) proposed the first notions of a model-neuron [17, 18]. This concept was then networked with an input and output layer by Rosenblatt in 1959 to simulate retinal functionality in order to recognize shapes [19]. This was the origin of the perceptron. In 1969 Minsky and Papert provided a rigorous analysis of the perceptron [20]. Werbos developed the back-propagation learning algorithm in 1971 [21]. In a back-propagation learning algorithm weights are updated based on error gradient calculated according to the chain rule of differentiation.
In a very simplistic way, a biological neuron is a cell that is characterized by [22, 23]:

- The synapses, connection points with other neurons, the nerve or the muscle fibres.
- The neural inputs.
- The axons, or outputs from the neuron to other neurons or muscle fibres.
- The core that activates the outputs according to the input stimuli.

Figure 1.1 schematically illustrates such an analog neuron [24]. For this figure, we can describe a model neuron mathematically according to

$$y = f\left(\sum_{j=1}^{k} w_j u_j\right), \tag{1.1}$$

where y is the neuron's output, u_j are the inputs from other neurons in the network, w_j are

weights of the corresponding input connections, f is the nonlinear activation function and k is the number of network's input. Neural networks are composed of neurons as their elementary computational units.

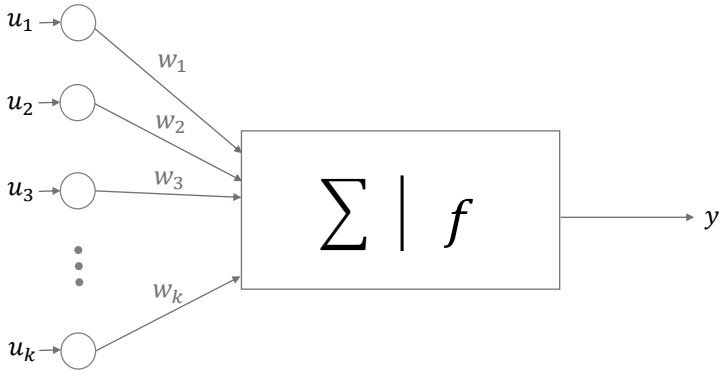

FIGURE 1.1 – Graph of an analog neuron with a non-linear function $y = f(u_j, w_j)$. u_j are called the inputs and w_j are weights.

1.1.1.1/ PERCEPTRON

Perceptrons are binary classifiers that take a linear combination of inputs and produce a binary output. Perceptrons made the functionality for the neural network models from 1980s until 1990s. The real value input vector $u \in \mathbb{R}^k$ is mapped to a binary output value $y(u)$

$$y(u) = \begin{cases} 1, & \text{if } w \cdot u + w_0 > 0 \\ -1, & \text{otherwise,} \end{cases} \tag{1.2}$$

with

$$w \cdot u = \sum_{j=1}^{k} w_j u_j, \tag{1.3}$$

where w_j,..., w_0 are the parameters of the separating hyperplane or decision boundary. A single layer perceptron is a simple feedforward neural network [25]. However, a single layer perceptron in Eq. 1.2 can solve only linearly separable patterns. For instance, they cannot solve the Boolean *XOR* problem. To solve this problem, an additional layer of artificial neurons has to be introduced. Still, perceptrons have been used in several applications such as image or speech recognition [26].

1.1.2/ ARTIFICIAL NEURAL NETWORKS

An ANN system consists of a number of artificial neurons and a large number of interconnections among them. During the past decade, ANNs have been one of the leading information processing concepts. Their neuro-inspired architectures make them different from Neumann computing concepts. In 1980s the potential of optics in the ANNs was first discussed. Today, photonic ANNs are at the centre of researchers' attention. According to the structure of the connections, two different classes of neural network are identified: feedforward neural network and recurrent neural network. In next sections we will describe these two classes.

1.1.2.1/ FEEDFORWARD NEURAL NETWORKS

The first class of neural network to be discussed here is feedforward networks. This network consists of at least three layers as illustrated in Fig. 1.2. Neurons are connected from one layer to next one and the information only travels forward in the layers. The first and last layers are the input layer and output layer, respectively. The layer that are not linked to an input or an output is hidden layer. However, a feedforward networks can consist of several hidden layers in other examples. Connection between the neurons has a connection strength which is known as weight. This weight defines how much the output of a neuron effects another one. For example, a single-layer perceptron model in Eq. 1.2 has only one layer, with a feedforward signal moving from a layer to an individual node. Multi-layer perceptron models are also feedforward [25, 27]. A feedforward network with k inputs, h hidden neurons and m output neurons is shown in Fig. 1.2. This network computes m nonlinear function of its k input variables [27]. In the first stage, the hidden layer **H** is a vector calculated as

$$\mathbf{H} = f_1\left(w_{\mathrm{i}}\mathbf{u}\right), \tag{1.4}$$

where **u** is the input vector of k input variables, and w_i is the input weights matrix. The result of matrix-vector product $(w_i\mathbf{u})$ is a vector, which is transformed by the perceptron's transfer function f_1. In the second stage, the output of the network **y** is calculated as a perceptron which receives the values from the hidden layer

$$\mathbf{y} = f_2\left(w_{\mathrm{r}}\mathbf{H}\right), \tag{1.5}$$

where w_r is the readout matrix and f_2 is another transfer function. As a result, a feedforward neural network in multi-layer perceptron is a nonlinear function of the inputs and the parameters.

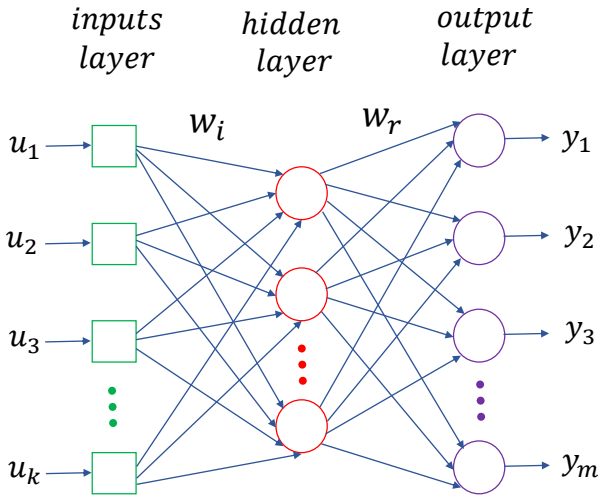

FIGURE 1.2 – Scheme of a feedforward neural network with k inputs as input layer, m output neurons as output layer and h hidden neurons as hidden layer.

1.1.2.2/ RECURRENT NEURAL NETWORK

A second class of neural networks to be discussed here are recurrent neural networks (RNN) [28, 29]. While a feedforward neural network has a one-directional information

processing flow, in RNN there is at least one loop inside its connection topology. In such a loop the connection returns back to its origin. A RNN presents a network in which the connections allow cycles. The hidden layer in the middle of RNNs is known as the recurrent layer. A schematic of RNN is shown in Fig. 1.3. Recurrent neural networks can

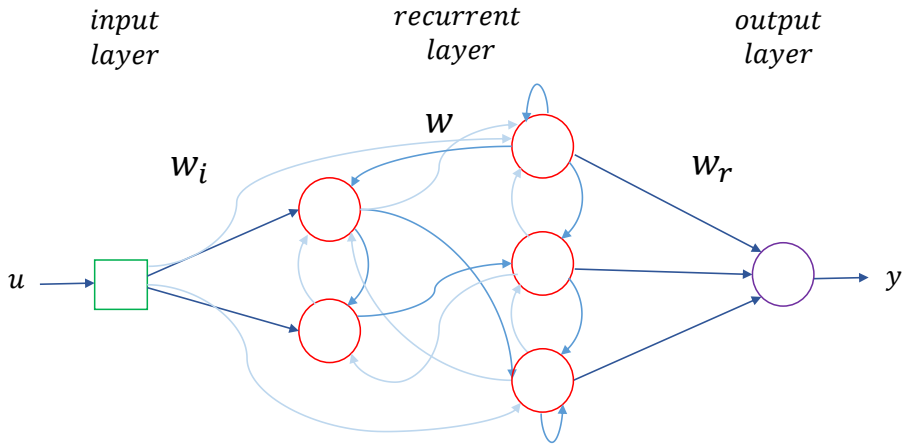

FIGURE 1.3 – Graph of a recurrent neural network (RNN). The recurrent layer in the middle with an internal connections keeps the information. For this reason, the output is depending on the previous inputs.

be defined as

$$\mathbf{x}(n) = f\left(w_{\mathrm{i}}\mathbf{u}(n) + w\mathbf{x}(n-1)\right), \tag{1.6}$$

here **u** is the input vector, w_i is the input weights matrix, w is the map of the previous state of the recurrent layer, and $\mathbf{x}(n)$ is the internal state of a network at discrete time n. The result of computation $\mathbf{y}(n)$ is obtained as

$$\mathbf{y}(n) = w_{\mathrm{r}}\mathbf{x}(n), \tag{1.7}$$

here w_r is considered as a readout map. Equation 1.6 illustrated that the internal state of RNN $\mathbf{x}(n)$ depends on the current and previous inputs. Hence in RNNs the input information does not propagate along a single direction towards the output. Feedforward neural networks are general function approximators, but RNNs can be regarded as algorithms. Recurrent neural networks are also known as universal approximators of dynamical systems [30]. Another considerable difference can be memory. Recurrent neural networks have memory due to the internal connection loops, while feedforward neural networks lack it. The Long Short-Term Memory (LSTM) networks are perhaps the most successful RNN [31]. There are two significant tasks in LSTM first, remember the all past knowledge second, forget the irrelevant data. To do the mentioned tasks, different activation function layers which are known gates are used.

1.1.3/ NEURAL NETWORKS TRAINING

Artificial neural networks have the great advances in control and processing the tasks such as decision-making and pattern recognition. In a successful ANN, there are important parameters such as number of layers, number of nodes in the layers and optimal configurations of connections weights. We require an optimized configuration in order to minimize the errors made by an ANN in a special task. This configuration is known as

learning, during which the ANN is trained. In this section, we will describe learning which is classified into two groups. The goal of these training methods is to obtain a functionality of $y(u)$ for network weights such that the computational error

$$E^{error} = \|y(\mathbf{u}) - y^{\text{target}}(\mathbf{u})\| \tag{1.8}$$

is minimized. Here, $\mathbf{u}$ is a set of input and output has a set $y^{\text{target}}(\mathbf{u})$. In general, the learning process can be classified [32]

- Learning with a teacher which mentions to supervised learning.
- Learning without a teacher which mentions to unsupervised learning.

1.1.3.1/ SUPERVISED LEARNING

When we know input and output variables and we need an algorithm to learn the mapping function from the input to the output, we implement supervised learning. In that case, a mapping function is approximated such that for new input data the ANN can correctly predict the desired output [32]. For instance in handwriting recognition, the input is defined as a database of known faces and the mapping function has to identify these faces. In another example for spam detection, the input is defined as the characters of a message and the mapping function has to identify that the email is spam or not by returning true or false. In last examples the goal is to obtain a mapping function which corresponds to identifying the specific relationships between the input data and the correct response. Since in supervised learning we have correct example outputs we can compute the gradients of the output error. Therefore error back propagation is a common training concept in supervised learning.

1.1.3.2/ FEEDFORWARD NEURAL NETWORK TRAINING

As we explained, to process information for a spacial task, the ANNs need to find a configuration of weights. Backpropagation can be considered as one of the popular supervised learning methods in feedforward neural network training [33]. Before training, all the system's connection matrices are typically initialized randomly. The backpropagation is executed iteratively [34]. In each iteration, a forward pass is calculated, which provides the current computation error. Then, during the backward pass, weights are updated from the output to the input layer. Based on the forward error and the chain rule the error gradient are computed from the final towards the first layer. This procedure causes that the error is dramatically decreased.
There are several modification of backpropagation algorithm, yet there remain strong challenges. For instance, reaching the global error minimum is not guaranteed. The training speed is also a limitation of this method, since the training is slowly converging. To train a neural network, lasting multiple iterations may need substantial computation resources.

1.1.3.3/ RECURRENT NEURAL NETWORK TRAINING

Training a RNN is similar to train a feedforward neural network with a large number of hidden layers. Figure 1.4 illustrates a RNN. By using the map w, the input $u(n)$ (in black

arrow) and the recurrent feedback $x(n)$ (in red arrow) are transferred. As a result the input $u(n)$, the recurrent feedback $x(n)$ and the map w establish a recurrent layer, which produces the internal network state $x(n+1)$. The result of this network state $x(n+1)$ is considered as the input for the following layer. Eventually, the readout map w_r returns the network's answer $y(n+1)$.

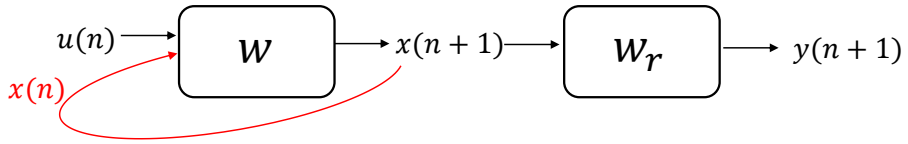

FIGURE 1.4 – Scheme of a simple RNN. w and w_r are the map matrix and the readout map, respectively.

Feedforward neural network can be constructed when the RNN is unfolded in time as shown in Fig. 1.5. In this figure, we replicate the recurrent layer k times to display it in k steps. The results with k layers get k inputs $u(n), u(n+1), ..., u(n+k-1)$. As a result, the internal state x depends on all the last inputs. This also demonstrated how RNNs store the information from the previous inputs.

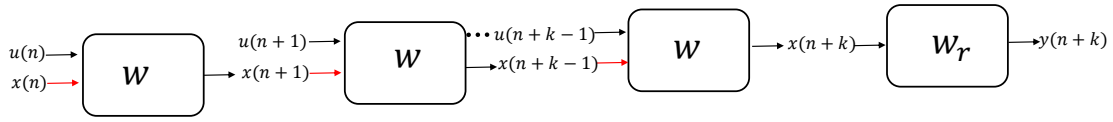

FIGURE 1.5 – Scheme of a RNN where w and w_r are the map matrix and the readout map, respectively. This network is unfolded in time for k time steps to construct a feedforward neural network.

In order to train this network, we can apply the method which is known back-propagation through time (BPTT). During the back-propagation, instead of propagating gradient through the different layers, they are propagated through different time steps hence, gradient can be increased or decreased at each time step. Unfortunately, this can be results in vanishing or exploding gradients, respectively. Hence, due to potentially vanishing or exploding gradients, there are difficulties to learn long-range temporal dependencies [35]. There are alternative solutions to this challenge such as teacher forcing [36]. As a results, the computational expenses in RNNs' training is typically larger than in feedforward neural networks with the same number of layers.

1.1.3.4/ UNSUPERVISED LEARNING

In unsupervised learning, there is only input data and no output target value. The goal is to modify statistical structure or distribution of the injection information based on autonomous optimization to learn more about the input information. The most common unsupervised learning problem is clustering where it is attempted to separate data into groups such as category, class and cluster. Another typical unsupervised is association rules in which relationships between variables are analyzed.

1.1.4/ RESERVOIR COMPUTING

Around the year 2000, two approaches were proposed to solve the problems of training complexity and slow convergence in RNN training [37]. These concepts could increase the performances of RNNs by only training the output weights. There were mainly two approaches considered:

- Liquid state machines (LSMs) are spiking and more biologically inspired networks which were introduced by Maass [38].
- Echo state networks (ESNs) consist of discrete-time analogue sigmoidal or hyperbolic tangent neurons. This concept was introduced by Jaeger [35].

By two solutions above, the number of trained weights is reduced and training corresponds to a single step matrix inversion. As a result, RNNs learning was simpler and the expenses were reduced. Both concepts, the LSM and ESN were later unified under the umbrella term Reservoir Computing (RC). The principal scheme of RC as illustrated in Fig. 1.6 consists of [39, 40]:

- The input layer: randomly maps the inputs to the internal nodes of the reservoir.
- A recurrent system (reservoir): non linearly transforms the data according to its internal connectivity.
- Output layer: linearly recombines the internal states of the reservoir.

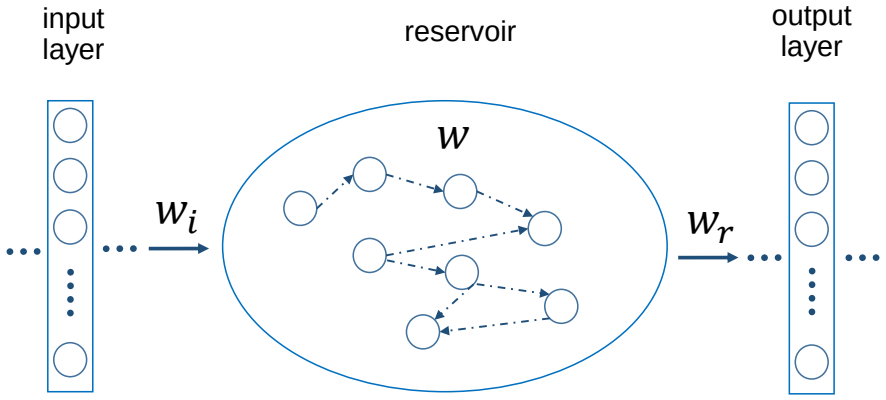

FIGURE 1.6 – The architecture of a reservoir computer consist of: an input layer, the reservoir, and an output layer. Data are masked with a randomly generated mask w_i, then the second layer with random connections called reservoir, and the third layer which obtained by multiplication with readout matrix w_r.

Reservoir computers can non-linearly map the injected information surprisingly in a large dimensional space, which shows their high computational power. Reservoirs are usually established of hundreds of nodes or neurons to make an area with higher dimensionality, which can increase the chance of good performance [41]. Optical systems can provide benefits in ANNs. In next sections, we will represent the benefits of photonic neural networks, followed by the discussion of two common typologies for photonic RC.

1.2/ PHOTONIC NEURAL NETWORKS

Hardware implementations of ANNs in optical systems have along history. Optical systems have advantages such as high processing speed due to the high speed in optical

devices. Artificial neural networks require a large connectivities in a small area to distribute the state of one node to several other fully in parallel [8], this can be possible by photonic systems. Interestingly, different approaches are successfully investigated in order to connect the photonic nodes and control the weights [42, 43]. Optical devices are also capable of establishing the nonlinear nodes and so the new implementations of this operation [44]. For instance thin films optical amplifiers and optical bistability are introduced because of nonlinear performance [45]. Different optical devices are represented in neuromorphic computing, including semiconductor Laser [46], saturable absorber laser [47], semiconductor optical amplifiers [48], photo-detectors [49], grating [50] and modulators [51]. All the systems above have demonstrated abilities of ANNs to process the information, these modern implementations and concepts constantly being published for better and larger implementations. This shows the high interest, potential and future of the optical computing with ANNs. There are two common typologies for photonic RC: first, system using the time delay reservoir (TDR) concept to create a reservoir by using a single nonlinear element coupled to a linear delay line and temporal multiplexing [52], second, spatio-temporal reservoirs where the each network neuron is implemented with a separate photonic components, and the network is multiplexed in space [8]. We will describe both typologies in next sections.

1.2.1/ PHOTONIC DELAY RESERVOIRS

In previous section, we discussed the importance and interest for optical ANNs implementations. One of the approaches of RC is known as a delay-based RC [52]. This concept is of particular interest because it can be implemented via a single nonlinear node instead and a delay line. The approach makes use of the high dimensionality of delay systems to create the state space required for RCs. By using time multiplexing, the neurons in reservoir layer are emulated, typically during one delay interval.
An example scheme of a hardware optoelectronic RC is shown in Fig. 1.7 [53]. The red parts in scheme 1.7 show the optical parts, the green elements the electrical parts. A laser as a continuous source light is located behind a Mach-Zehnder modulator (MZ) which is used to provide a nonlinear transformation $(\sin^2)$, and the delay loop in this setup is a fiber spool. At the output of the fibre spool, a photodiode converts the light intensity to a voltage. The output voltage from the photodiode is combined with an input voltage which encodes the input information (generated by a function generator). The obtained result is amplified, then used to drive the intensity modulator. A bias offset φ is adjusted by using a DC voltage to change the operating point of the MZ modulator. After the photodiode, a digitizer (oscilloscope) records the response of the system.
The schematic illustration in Fig. 1.7 demonstrates the various stages of processing of the reservoir nodes and inputs. The internal state of the reservoir is $x(t)$. After a temporal retardation by delay τ, data will be $x(t-\tau)$. The input information $u(t)$ is injected with an input gain β. The input information is combined with data from previous node state $x(t-\tau)$. Then the established signal is amplified by the value α. Function sin^2 has been chosen as nonlinear function to establish new state reservoir. This reservoir state can be written

$$x(t) = \sin^2\left(\alpha x(t-\tau) + \beta u(t) + \varphi\right). \tag{1.9}$$

All the network states are gathered in readout section, then they are multiplied by respective weights W, finally they are added together to obtain the successful output. This hardware reservoir was applied in nonlinear channel equalization [54] and speech recognition [55–57]. The excellent obtained results from these experiments illustrated its

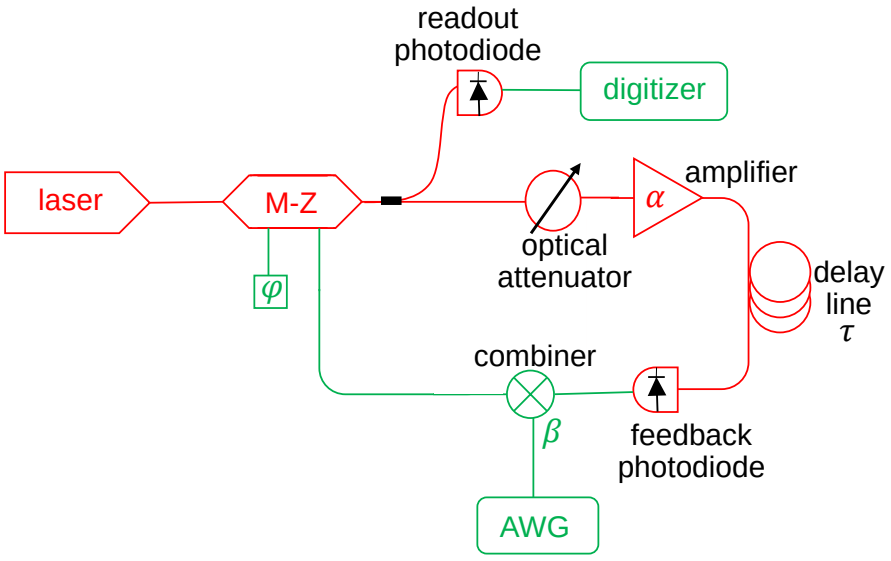

FIGURE 1.7 – Scheme of an optoelectronic delay RC. There are two fundamental parts in this scheme. Optical and electrical parts are in red and green respectively. An optical fibre is used to create the delay loop. The input information is produced by computer and injected to the system by AWG (arbitrary waveform generator). The digitizer (oscilloscope) records the response of the system.

relevance on tasks of practical importance. Moreover, the information processing was enough fast for real time information processing. By re-optimizing the output layer (new readout weight W_r) and by changing operating points (the feedback gain α, the input gain β, and possibly the bias φ), these reservoirs can be used for different tasks.

1.2.2/ PHOTONIC SPATIO-TEMPORAL RESERVOIRS

Here, I will introduce an optical RC where the reservoir is extended in space. Spatially extended of optical neural networks are a well established concept [52, 58, 59] and interest idea grew fast due to, first, the possibilities of realizing photonic neural networks with numerous nonlinear nodes fully in parallel [42, 58], second, the high connectivity that could be achieved with optical systems because, optical fields can cross without interacting [60, 61].
One benefit of a spatially extended reservoir is that the nodes respond simultaneously to injected information, and such the network therefore avoids the speed limitation which occurred because of time multiplexing in time delay system [62]. A spatially multiplexed photonic reservoir has to consist of a network of photonic nonlinear elements for which we are able to extract the information for the final output. This network should consist of a large number of nonlinear nodes to solve difficult tasks, and additionally all nodes need to be able to potentially exhibit nonlinear dynamics.
In Chapter 5, a spatially extended large optical reservoir in an experiment will be introduced and developed. To develop this reservoir, we investigate experimentally, numerically and analytically the fundamental limits to the size of this photonic network from Chapter 2 to Chapter 4. In chapter 5, I will present a reservoir with almost 10^3 nonlinear nodes, interconnected with neighboring nodes by using a single passive optical element and we

will apply the learning techniques to optimize the output weights. As a result, this system will be successfully trained to predict the chaotic Mackey-Glass (MG) timeseries and will report on most recent results [8].

1.3/ CONCLUSION

Neural networks have a biologically neuro-inspired origin. Artificial neural networks therefore try to mimic the most basic structures of brain. Neuron was modeled in ANNs. One of the basic ANNs that was explained in this Chapter is feedforward neural networks with three layers: input layer, hidden layer and output layer, where connections are created from one layer to the following layer. Such feedforward neural networks are memory-less. Another ANNs in this Chapter was RNN consisting of three layers: input, recurrent, and output layer. In RNNs, there are connectivity among neurons in the hidden layer, also referred to as self connections or cyclic connections. In RNNs, information store in its' dynamical state which creates an internal memory. The internal state of RNN depends on current and previous inputs.
To use a physical system as a neural network, it is necessary to apply learning rules. These learning rules optimize the configuration to solve a particular task. We presented supervised learning vice-versa unsupervised learning and I described feedforward neural network training and RNN training to minimize an ANN's computing error.
While RNNs are powerful, their training algorithm is neither simple nor quick. For this reason RC was developed in the communities of ANN and neuro-science to solve the mentioned problems. Reservoir Computers have three layers: input layer, reservoir layer and readout layer. Random connections and that training only modifies the readout. In this Chapter, two photonic reservoirs were introduced. Photonic delay reservoirs use a delay system to create the state space for RC. This hardware was successfully used in nonlinear channel equalization and speech recognition. In addition, this setup was fast for real time information processing. Furthermore, I described the optical spatial reservoir. In this reservoir we explained the benefits and challenges in compared with delay reservoir. We will develop and realize a photonic spatio-temporal reservoir in Chapter 5 of this thesis. But first, I will explain the fundamental limits to the size of this photonic network analytically, experimentally and numerically in Chapter 2 to 4. Then, we will describe completely in details the setup of photonic spatio-temporal reservoir in Chapter 5.

2

DIFFRACTIVE COUPLING

In the previous chapter, we explained the objective for developing an optical hardware neural network. The main concept to create the connectivity in an optical ANN to be demonstrated in this Thesis is based on diffractive coupling, we therefore the principle ideas and concepts of scalar diffraction theory. At the end of this chapter, we will describe the diffractive coupling technique used in our optical neural network.

2.1/ THEORY OF DIFFRACTION

The motivation of this section is to start from Maxwell equations and derive the main equations and consequences of the diffraction theory. The goal is to explain the assumptions of the different theoretical models for light such that we are able to derive an adequate description of the diffractive coupling concept. During the explanation of theoretical models, we will introduce the following different important notions: plane waves, spherical waves, the Huygens principle, the Fresnel integral and the plane-wave spectrum [63, 64].

2.1.1/ MAXWELL'S EQUATIONS

Maxwell's equations are the starting point of our reasoning. This set of coupled partial differential equations describes the spatial and temporal evaluations of electromagnetic fields [65, 66]

$$\nabla \cdot \mathbf{D} = \rho \tag{2.1}$$

$$\nabla \cdot \mathbf{B} = 0 \tag{2.2}$$

$$\nabla \times \mathbf{H} - \frac{\partial \mathbf{D}}{\partial t} = \mathbf{J} \tag{2.3}$$

$$\nabla \times \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} = 0 \tag{2.4}$$

where **E** is the electric field vector, **H** is the magnetic field vector, **B** is magnetic induction and **D** electric displacement, ρ is charge density and **J** is electric current density. Equations 2.1 and 2.2 characterize the divergence properties of fields, whereas Eq. 2.3 and Eq. 2.4 treat the coupling between magnetic field and electric fields.
Vector quantities are the basic field vectors such as electric field: $\mathbf{E} = (E_x, E_y, E_z)$ and magnetic Field: $\mathbf{H} = (H_x, H_y, H_z)$. Whereas **D** and **B** are describing the influence of matter on

the electromagnetic fields such as electric displacement: $\mathbf{D} = (D_x, D_y, D_z)$ and magnetic induction: $\mathbf{B} = (B_x, B_y, B_z)$. Maxwell equations therefore connect the five basic quantities **E**, **H**, **B**, **D** and **J** and the temporal derivatives of **E** and **B**. To allow a unique determination of the field vectors from a given distribution of currents and charges, these equations must be supplemented by relationship which describe the behavior of matter under the influence of the field, which are also known as constitutive relations. If the material is isotropic, linear and homogeneous, they take the relatively simple form

$$\mathbf{J} = \sigma\mathbf{E} \tag{2.5}$$

$$\mathbf{D} = \epsilon\mathbf{E} \tag{2.6}$$

$$\mathbf{B} = \mu\mathbf{H} \tag{2.7}$$

where σ is called the conductivity, ε is the dielectric permittivity of the medium and μ is the magnetic permeability of the medium. Those quantities characterize the response of a medium to an electric and magnetic excitation. In many situations they are more complex, for example when the medium is inhomogenous, non-isotropic and non-linear. Indeed, if the medium is inhomogenous medium, these quantities are dependant of the spatial coordinates. If the medium is non-isotropic, these quantities become tensors and if the medium is non-linear, these quantities become functions of the fields' amplitudes.
For the following, we will consider a non-magnetic medium. In this kind of medium, the magnetic permeability of the medium is the magnetic permeability of the vacuum μ_0

$$\mu = \mu_0. \tag{2.8}$$

Moreover, we will consider neutral medium ($\rho = 0$) without free current density ($\mathbf{J} = \mathbf{0}$).

2.1.2/ PROPAGATION EQUATION

In this section, we will derive propagation equations from Maxwell's equations. In Eq. 2.3, the temporal variation of the electric displacement vector is linked to the spatial variation of the magnetic field vector [66]. In Eq. 2.4, the temporal variation of the magnetic induction vector is linked to the spatial variation of the electric field vector, and we want to obtain differential equations which each of the basic fields must separately satisfy. For that purpose, we substitute for **B** from the constitutive Eq. 2.7 into the Eq. 2.4, and we divide both side by μ_0 and apply the Curl operator

$$\nabla \times (\frac{1}{\mu_0}\nabla \times \mathbf{E}) + \nabla \times \frac{\partial \mathbf{H}}{\partial t} = 0. \tag{2.9}$$

In parallel, we differentiate Eq. 2.3 with respect to time and use the constitutive Eq. 2.6 for **D**. We replace $\left(\nabla \times \frac{\partial \mathbf{H}}{\partial t}\right)$ in Eq. 2.9 by $(\frac{\epsilon\partial^2\mathbf{E}}{\partial t^2})$ and obtain the following equation

$$\nabla \times (\frac{1}{\mu_0}\nabla \times \mathbf{E}) + \frac{\epsilon\partial^2\mathbf{E}}{\partial t^2} = 0. \tag{2.10}$$

Using the vector identity $\nabla \times (\nabla \times \mathbf{U}) = \nabla(\nabla \cdot \mathbf{U})–\nabla^2\mathbf{U}$ we obtain the following equation

$$\nabla^2\mathbf{E} - \epsilon\mu_0\frac{\partial^2\mathbf{E}}{\partial t^2} = 0. \tag{2.11}$$

Equation 2.11 is called *the propagation equation.* It links the spatial and temporal variations of the electric field in an homogeneous medium. In a similar manner, the propagation equation for the magnetic field can be obtained

$$\nabla^2 \mathbf{H} - \epsilon\mu_0 \frac{\partial^2 \mathbf{H}}{\partial t^2} = 0. \tag{2.12}$$

2.1.3/ SCALAR PROPAGATION EQUATION

In our work, we will consider only linear polarisation and can therefore describe the electric field by a scalar function E(**r**,t)

$$\mathbf{E}(\mathbf{r}, t) = E(\mathbf{r}, t)\mathbf{u}_0, \tag{2.13}$$

where $\mathbf{u}_0$ is a unit vector which indicates the polarization axis. Thus, we can write the scalar propagation equation for the function E(**r**,t)

$$\nabla^2 E - \epsilon\mu_0 \frac{\partial^2 E}{\partial t^2} = 0. \tag{2.14}$$

2.1.4/ HELMHOLTZ EQUATION

The Helmholtz equation represents a time-independent form of the propagation equation. It can be derived Helmholtz equation by applying the technique of separation of variables in order to reduce the complexity of the analysis. We consider an electric field in which we assume that the function $E(\mathbf{r}, t)$ is in fact separable according to

$$E(\mathbf{r}, t) = A(\mathbf{r})C(t), \tag{2.15}$$

where $A(\mathbf{r})$ is an amplitude term which depends only on spatial position **r**, and $C(t)$ is an amplitude term which depends only on time. In order to obtain a time-independent form, we will consider a monochromatic wave

$$C(t) = e^{j\omega t}, \tag{2.16}$$

and the scalar propagation Eq. 2.14 becomes

$$\nabla^2 A + \omega^2 \epsilon\mu_0 A = 0, \tag{2.17}$$

which is the *scalar Helmholtz equation*. This equation describes the spatial evolution of a monochromatic field in a medium with a dielectric permittivity ϵ.

2.1.4.1/ PROPAGATION OF PLANE-WAVES

The simplest solution to the Helmholtz equation is a plane wave. Plane waves carry a given amount of energy per area per time that depends on the amplitude of the electric and magnetic fields [67,68]. Plane waves are produced by a monochromatic point source located at infinity. Figure 2.1 consists of the blue planes as positive amplitude planes whereas red ones as negative amplitude planes. In Fig. 2.1 the field is oscillating in x and

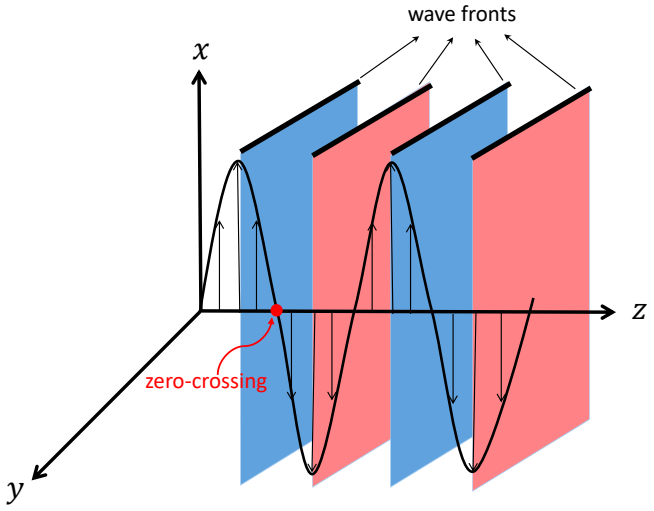

FIGURE 2.1 – Structure of a plane wave. Blue planes are positive amplitude planes. Red planes are negative amplitude planes. Black sine wave shows the direction of plane wave propagation in z. In zero-crossing point the sign of a mathematical function changes (from positive to negative).

y direction, and spatial variation only happens along propagation axis z. Such types of waves are called transversal waves.

Plane waves are characterised by an amplitude and a phase. The amplitude is the peak magnitude of the oscillation $(\pm M)$ as shown in Fig. 2.2(a). The phase of the wave corresponds to the position along the propagation direction z. If we think of a wave as having local maxima and minima with a zero-crossing between them, the relative phase of the wave is defined from a reference (for example the first zero-crossing) and a considered point. The notion of phase is shown in Fig. 2.2(a) in time for a fixed spatial position. The difference between two maxima in time defines the period of the wave T, and frequency f is the inverse of the period and characterised the number of period per second, whereas wave's angular frequency is defined by $\omega = \frac{2\pi}{T}$. The notion of phase is shown in Fig. 2.2(b) in space at a given time. Mathematically, a plane wave is written as a constant amplitude term M modulated by a complex exponential

$$E(z,t) = Me^{i(\omega t - kz)}. \tag{2.18}$$

To be a solution of the Helmholtz Eq. 2.17, the wave number k must be equal to

$$k^2 = \omega^2 \epsilon \mu_0. \tag{2.19}$$

We define the dielectric permittivity as follows

$$\epsilon = \epsilon_r \epsilon_0, \tag{2.20}$$

where ϵ_r is a relative permittivity specific to the medium and ϵ_0 is the vacuum permittivity. Since the velocity of light in vacuum is defined as $c = 1/\sqrt{\epsilon_0 \mu_0}$, the wave number k becomes

$$k = \frac{\sqrt{\epsilon_r}\omega}{c}. \tag{2.21}$$

The refractive index n of a medium is

$$n = \sqrt{\epsilon_r}, \tag{2.22}$$

and thus, wavenumber k becomes

$$k = \frac{n\omega}{c}. \tag{2.23}$$

The wavelength λ inside the medium is directly linked to the wavenumber by

$$\lambda = \frac{2\pi}{k}, \tag{2.24}$$

where k is the wave's wavenumber and distance between the peaks is given by wavelength λ. Equation 2.23 is known as the dispersion relation of a plane wave in a medium of refractive index n. It allows to link the optical oscillation with angular frequency ω to the wave number in the medium k.

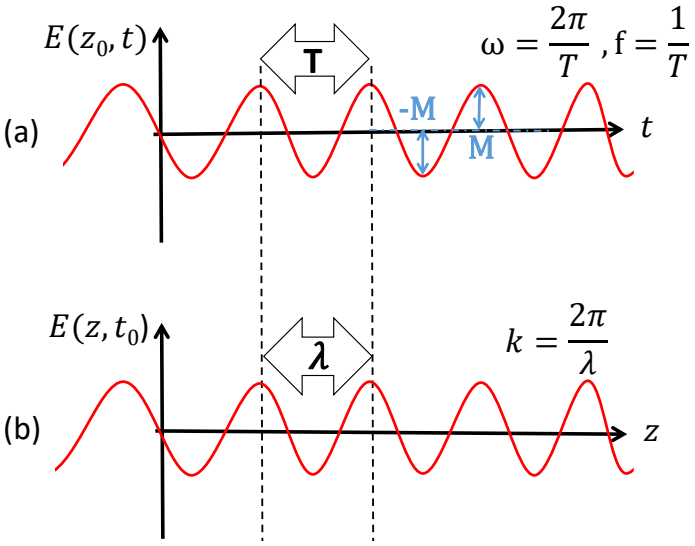

FIGURE 2.2 – (a) Amplitude and phase of a plane wave for a given spatial positions. The time between two peaks is T, so angular frequency is $\frac{2\pi}{T}$. (b) Amplitude and phase of a plane wave at a given time. Distance between two peaks is λ, so the wave number is $\frac{2\pi}{k}$.

The superposition principle states that in linear systems the net response caused by two or more stimuli is the sum of the responses to each stimulus individually. Therefore, if A_1 and A_2 are a solution of the Helmholtz equation, then $A_1 + A_2$ will have a solution. As a single plane wave is a solution for the Helmholtz equation, then the same therefore is true for the sum of plane waves. Furthermore, any complex beam can be decomposed into a sum of plane waves. Consequently, if we are able to decompose a complex optical beam into a sum of plane waves, we can find of the complex beam propagation according to Helmholtz equation. Monochromatic plane waves are the most simple solution, however, more complex solutions do exist.

2.1.4.2/ VELOCITY OF LIGHT IN A MEDIUM

The phase of the plane wave in Eq. 2.18 by setting $\varphi = \omega t - kz$ can be written

$$\frac{\omega}{k} = \frac{\frac{\Delta\varphi}{\Delta t}}{\frac{\Delta\varphi}{\Delta z}} = \frac{\Delta z}{\Delta t} = v_p, \tag{2.25}$$

here v_{p} is the phase velocity of a plane wave in a medium. From Eq. 2.23, we therefore obtain

$$v_p = \frac{\omega}{k} = \frac{c}{n}. \tag{2.26}$$

For a known n one therefore can obtain the phase's velocity of propagating in the medium. The dielectric constant ϵ is usually greater than unity. Consequently, a wave propagates slower in a medium than in vacuum.

2.1.5/ HUYGENS PRINCIPLE

There are two traditional ways to find the main results of scalar diffraction theory for the Rayleigh-Sommerfeld diffraction integrals:

1. The first way is based on the Green theorem to deduce the Kirchhoff diffraction integral from the Helmholtz equation.
2. The second way is based on the Huygens principle. In this thesis, we implement the Huygens principle to obtain the Rayleigh-Sommerfeld diffraction integrals, the Fresnel diffraction formula and the Fraunhofer diffraction formula.

To understand the Huygens principle, we start from the notion of a wave front [69,70]. The wave front is an imaginary line joining points on a wave that are in phase. Three black lines in Fig. 2.3 illustrate the fronts created by the shown waves. The distance between each wave front is λ. The Huygens principle states that every point on a wavefront is itself

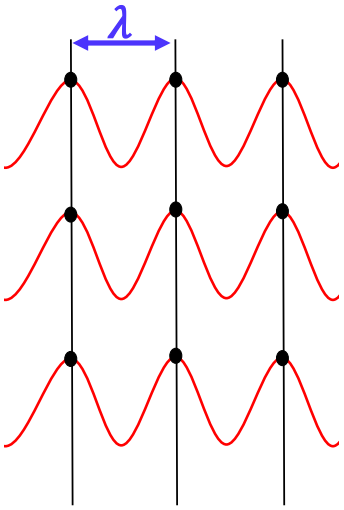

FIGURE 2.3 – Wave front description. Distance between the peaks is wavelength λ.

the source of a spherical wavelet. There is a point where all of these spherical wavelets touch and form a line, which is the next wave front or the future position of the wave.
When the plane waves go through a slit (see Fig. 2.4), the points on the plane wave inside the slit therefore emit spherical wavelets. We will describe this phenomenon by starting from a simple plane wave

$$E(r) = e^{ikr}, \tag{2.27}$$

where r is the distance between the point source and the observation point. Physically, we observe that the amplitude of the wave decreases with distance, and therefore we

include another factor

$$E(r) = \frac{1}{r} e^{ikr}. \tag{2.28}$$

We can describe this effect of wave propagation by including a factor $\cos(\theta)$ which was derived from Green's function and the first Rayleigh Sommerfeld solution

$$E(r) = \frac{\cos(\theta)}{r} e^{ikr}, \tag{2.29}$$

where θ is the angle between the Z axis and and r. This can be rewritten more convenient using $\cos(\theta) = \frac{z}{r}$, hence

$$E(r) = \frac{z}{r^2} e^{ikr}. \tag{2.30}$$

We will apply Eq. 2.30 in the Rayleigh-Sommerfeld integral in next section.

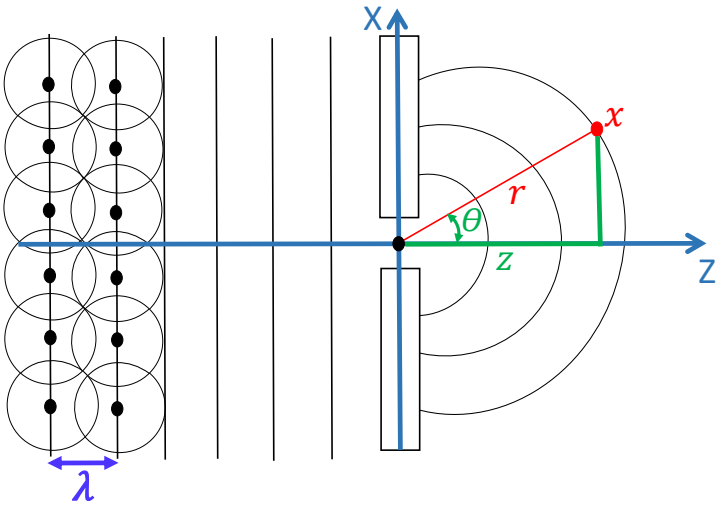

FIGURE 2.4 – Huygens principle. Every point on a wave front acts as an independent source to generate wavelets for the next wave front. Propagation of light is in Z direction and distance between two wave fronts is λ. Where waves approaches a barrier with a small slit, the points of plane waves emit spherical wavelets.

2.1.6/ RAYLEIGH-SOMMERFELD INTEGRAL

While the boundary conditions on the field strength and the normal derivative are required to be imposed in Kirchhoff theory, Sommerfeld eliminated these necessities in the Rayleigh Sommerfeld theory. For instance in Fig. 2.5, the Rayleigh-Sommerfeld integral adds the contributions from all point sources in one plane to find the field in another plane that is separated by distance z. Using $x_1^{'}$ and $y_1^{'}$ to denote two dimensional position vector in source plane, and x and y denote the two dimensional position vector in the observation plane [71,72], the distance between position $(x_1^{'}, y_1^{'})$ and position (x, y) is given by r which we can calculate. To find the total field at the observation point (x_1, y_1), we simply integrate over all the point sources in the plane source $z^{'} = 0$ is therefore

$$E(x_1, y_1, z) = \frac{1}{i\lambda} \int\int E(x_1^{'}, y_1^{'}, 0) \frac{z}{r^2} e^{ikr} dx_1^{'} dy_1^{'}, \tag{2.31}$$

with

$$r = \sqrt{(x_1 - x_1^{'})^2 + (y_1 - y_1^{'})^2 + z^2}, \tag{2.32}$$

where the term $(\frac{z}{r^2}e^{ikr})$ was obtained from Eq. 2.30 in the last section.

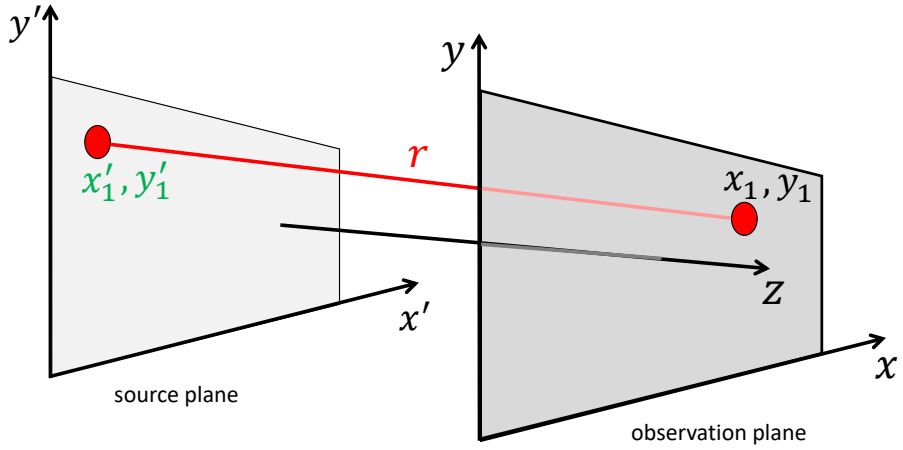

FIGURE 2.5 – Source plane and observation plane used in the Rayleigh-Sommerfeld integral, where $x_1^{'}$ and $y_1^{'}$ are position vector in source plane, x and y are the position vector in the observation plane. The distance between position $(x_1^{'},y_1^{'})$ and position (x, y) is given by r.

2.1.7/ FRESNEL DIFFRACTION INTEGRAL

The Rayleigh-Sommerfeld integral is our starting point to deduce the well known Fresnel and Fraunhofer formulas [66, 73, 74]. Our aim is to simplify the Rayleigh-Sommerfeld integral Eq. 2.31 under the assumption that the propagation distance z is larger than the range of the transverse coordinates $(x_1^{'}, y_1^{'})$. In particular, with this assumption we can simplify the expression for r. As we see r occurs twice in the integral. One in the denominator and once in the complex exponential. In the denominator we can employ the following approximation

$$\frac{1}{r^2} \approx \frac{1}{z^2}, \tag{2.33}$$

and can therefore rewrite r from Eq. 2.32 as follow

$$r = z\sqrt{1 + \frac{(x_1 - x_1^{'})^2 + (y_1 - y_1^{'})^2}{z^2}}. \tag{2.34}$$

By using a first order Taylor expansion ($\sqrt{1+a} = 1 + \frac{a}{2} - \frac{a^2}{8} + ...$), we can approximate Eq. 2.34 as

$$r \approx z\left(1 + \frac{(x_1 - x_1^{'})^2 + (y_1 - y_1^{'})^2}{2z^2}\right), \tag{2.35}$$

which we can expand r in four terms

$$r \approx z + \frac{x_1^2 + y_1^2}{2z} + \frac{x_1^{'2} + y_1^{'}}{2z} + \frac{-x_1 x_1^{'} - y_1 y_1^{'2}}{z}. \tag{2.36}$$

By inserting the simplified expression for r of (Eq. 2.36) into the Rayleigh-Sommerfeld integral (Eq. 2.31), we obtain

$$E(x_1, y_1, z) = \frac{e^{ikz} e^{ik\frac{x_1^2+y_1^2}{2z}}}{\lambda z i} \int\int E(x_1^{'}, y_1^{'}, 0) e^{ik\frac{x_1^{'2}+y_1^{'2}}{2z}} e^{ik\frac{-x_1 x_1^{'}-y_1 y_1^{'}}{z}} dx_1^{'} dy_1^{'}. \tag{2.37}$$

Since $k = \frac{2\pi}{\lambda}$, we can write

$$E(x_1, y_1, z) = \frac{e^{ikz} e^{ik\frac{x_1^2+y_1^2}{2z}}}{\lambda z i} \int\int E(x_1^{'}, y_1^{'}, 0) e^{ik\frac{x_1^{'2}+y_1^{'2}}{2z}} e^{-2\pi i\left(x_1^{'}\frac{x_1}{\lambda z}+y_1^{'}\frac{y_1}{\lambda z}\right)} dx_1^{'} dy_1^{'}. \tag{2.38}$$

Inside of integral in Eq. 2.38, there are two terms which depend on the integration variables. The first term is a quadratic phase factor, the second one the complex exponentials which has similarities with the definition of Fourier transform

$$F\{f(t)\} \equiv \int f(t) e^{2\pi i \omega t} dt. \tag{2.39}$$

By using the Fourier transform definition, we can therefore rewrite Eq. 2.38

$$E(x_1, y_1, z) = \frac{e^{ikz} e^{ik\frac{x_1^2+y_1^2}{2z}}}{\lambda z i} F\left\{E(x_1^{'}, y_1^{'}, 0) e^{ik\frac{x_1^{'2}+y_1^{'2}}{2z}}\right\}\left(\frac{x_1}{\lambda z}, \frac{y_1}{\lambda z}\right), \tag{2.40}$$

and this equation is known as the *Fresnel diffraction integral*.

2.1.8/ FRAUNHOFER DIFFRACTION INTEGRAL

If the propagation distance z is much larger than the transverse distances $(x_1^{'}, y_1^{'})$, then the quadratic phase factor inside the Fourier transform's argument in Eq. 2.40 can be ignored. Thus, the propagated field $E(x_1, y_1, z)$ is simply the Fourier transform of initial field multiplied by a simple factor

$$E(x_1, y_1, z) = \frac{e^{ikz} e^{ik\frac{x_1^2+y_1^2}{2z}}}{\lambda z i} F\left\{E(x_1^{'}, y_1^{'}, 0)\right\}\left(\frac{x_1}{\lambda z}, \frac{y_1}{\lambda z}\right). \tag{2.41}$$

Equation 2.41 is known as the *Fraunhofer diffraction integral* [75, 76]. which is the common approximation for the optical far field.

2.2/ PLANE-WAVE SPECTRUM

In this section, we will describe the plane-wave spectrum. This approach will be used in the numerical simulation of the following Chapters. We will derive this method from the Helmholtz equation Eq. 2.17. At the end of this section, we will create the link between the plane-wave spectrum and the other approaches (typically Fresnel integral) [77, 78].
First, we consider the propagation of an electromagnetic wave $E(x, y, z)$ in z direction between $z = 0$ and z. We can write the two dimensional Fourier transform as follows

$$E(u, v, z) \equiv \int\int E(x, y, z) \cdot e^{-2\pi i(ux+vy)} dx dy. \tag{2.42}$$

In the same frame of mind, we can define the inverse Fourier transform $E(u,v,z)$

$$E(x,y,z) \equiv \int\int E(u,v,z)\cdot e^{2\pi i(ux+vy)}dudv, \tag{2.43}$$

where u and v are respectively the spatial frequencies in the x and y direction. Starting from Helmholtz equation Eq. 2.17 to describe the propagation of the field $E(x,y,z)$ in z direction

$$\left(\nabla^2 + k^2\right)\cdot E(x,y,z) = 0, \tag{2.44}$$

we can express the field as function of its Fourier transform (Eq. 2.43)

$$\nabla^2\left(\int\int E(u,v,z)\cdot e^{2\pi i(ux+vy)}dudv\right) + k^2\left(\int\int E(u,v,z)\cdot e^{2\pi i(ux+vy)}dudv\right) = 0. \tag{2.45}$$

By performing the transverse spatial derivatives, Eq. 2.45 becomes

$$\frac{\partial^2 E}{\partial z^2}(u,v,z) + k^2(1-\lambda^2u^2-\lambda^2v^2)\cdot E(u,v,z) = 0, \tag{2.46}$$

for which the mathematical solution simply is

$$E(u,v,z) = E(u,v,0)e^{ikz\sqrt{1-\lambda^2u^2-\lambda^2v^2}}. \tag{2.47}$$

Equation 2.47 shows in particular that the propagation of a complex electromagnetic wave is simple in Fourier space. To propagate a field in Fourier space, we just have to multiply the initial field $E(u,v,0)$ with $e^{ikz\sqrt{1-\lambda^2u^2-\lambda^2v^2}}$. The plane-wave spectrum method is based on Eq. 2.47. The propagation of a optical field based on this method can be summarized in 3 steps:

1. Calculate the Fourier transform $E(u,v,0)$ of the initial field $E(x,y,0)$. The idea of the Fourier transform is a plane wave decomposition: it explains notably the name of the method *plane-wave spectrum*.
2. Multiply the field $E(u,v,0)$ in the Fourier space with factor $e^{ikz\sqrt{1-\lambda^2u^2-\lambda^2v^2}}$ (see Eq. 2.47).
3. We can go back in the direct space (x,y) with the inverse Fourier transform to have access to the propagate field $E(x,y,z)$.

Finally, the plane-wave spectrum is a more general approach than Fresnel diffraction integral of Eq. 2.40 and Fraunhofer diffraction integral of Eq. 2.41. Indeed, these integrals can be derived from the plane-wave spectrum. Because of its simplicity and accuracy, the plane-wave spectrum is in particular well suited for our problem.

2.2.1/ STEPS OF PLANE-WAVE SPECTRUM FOR PROPAGATION

In last section, we introduced the 3 summarized steps to propagate the field by the plane-wave spectrum. In this section, we investigate these steps in details to find the function and the algorithms for our numerical simulation:

1- The first step starts with the initial field distributed along x and y as $E(x,y,z=0)$ and to calculate the Fourier transform of this field $E(u,v,z=0)$. Fourier methods are also widely used for numerical computations [79]. For this reason we will describe Fourier transform

further in next part.
The Fourier transform is a mathematical operation which transforms data from a spatial or time domain into a frequency domain. For any function $f(x)$, the Fourier transform can be defined $F(u)$ [80, 81]. The Fourier transform in continuous space or time is known as continuous Fourier transform. In this thesis, we measure x in space so u corresponds to inverse space or spatial frequency. The Fourier transform is defined by

$$F(u) \equiv \int_{-\infty}^{+\infty} f(x) e^{-2\pi i u x} dx, \tag{2.48}$$

which is a harmonic function which can be considered as sum of the plane waves with different amplitudes, frequencies and phases. In Eq. 2.48, $f(x)$ is relative amplitude and the term $e^{-2\pi i u x}$ is the plane waves. In two dimensions the continuous Fourier transformation is given by

$$F(u,v) \equiv \int_{-\infty}^{+\infty} f(x,y) e^{-2\pi i (ux+vy)} dx dy, \tag{2.49}$$

where the variables x and y represent space, u and v are their corresponding spatial frequencies.
Next, we introduce a discrete version of the Fourier transform. In that case, we deal with signals which are discretely sampled, usually at constant intervals. For such data, only a finite number of sinusoids is needed and the Discrete Fourier Transform (DFT) is the corresponding operation [82, 83]. For all Fourier transform theorems, there are related theorems for the DFT. To use the DFT we sample the data points at a sufficiently high rate. The DFT of N uniformly sampled data points x_{i} where $i = 0, 1, ..., N-1$ is determined by

$$X_k = \sum_{i=0}^{N-1} x_j e^{-2\pi i jk/N}, \tag{2.50}$$

if the signal is sampled uniformly, then the frequency corresponding to one-half that rate is called the Nyquist frequency as [84–88]

$$\nu_{N/2} = \frac{1}{(2\Delta x)}. \tag{2.51}$$

The sampling rate of $(\frac{1}{\Delta x})$ is known as Nyquist rate, where Δx is the largest allowed spacing between samples to still reconstruct correctly a signal. If discrete sampling is below the Nyquist frequency (insufficient sampling) aliasing will appear, which results in incorrect frequencies and/or amplitudes [89]. The Fast Fourier transform (FFT) algorithm is using to translate the DFT in our codes [90, 91]. In our numerical simulation, we use the algorithm of (FFT) to transfer our initial signal in spatial domain into frequency domain. Fourier transform has a significant property (the inverse transform) which can be determined by

$$f(x) \equiv \int_{-\infty}^{+\infty} F(v) e^{2\pi i v x} dv, \tag{2.52}$$

and in two dimension

$$f(x,y) \equiv \int_{-\infty}^{+\infty} F(u,v) e^{2\pi i (ux+vy)} du dv. \tag{2.53}$$

In two dimensions the discrete Fourier transform is

$$x_j = \sum_{k=0}^{N-1} X_k e^{2\pi i jk/N}. \tag{2.54}$$

To compute DFT from the Eq. 2.54, another algorithm is introduced. By using the algorithm (F^{-1}), we are able to achieve the result in space domain numerically.

2- The second step applies the $e^{ikz\sqrt{1-\lambda^2u^2-\lambda^2v^2}}$ phase shift for each spatial frequency component at a position z. This term was obtained by the non-paraxial plane-wave spectrum. In this section, we explain that how we derive the phase shift numerically. We therefore calculate this term also in paraxial approximation, allowing to make a comparison between both results.

Paraxial approximation

The paraxial approximation corresponds to the linear approximation of geometric optics valid under certain conditions. One of these conditions is that the incidence angles of rays with regards to the optical axis of an element are small. For obtaining analytical solutions it is often necessary to apply these approximations in order to simplify the equations [92–94]. However, in numerical implementations we can use the non-paraxial form. In this part we start calculating the propagation of a plane wave in paraxial situation [95]. The phase shift of each plane wave is simply $k_r\vec{r}$, with vector r in one dimension x illustrated in Fig. 2.6. Integrating the propagated plane wave in z direction can then be written as

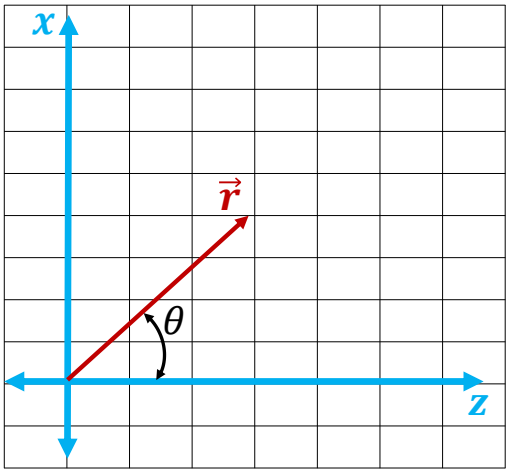

FIGURE 2.6 – Vector $\vec{r}$ in the coordinate system x and z. The plane wave propagates along z direction and θ is the angle between vector $\vec{r}$ and z component.

$$E(x,z) = \int E(v)e^{ik_r\vec{r}}dr, \tag{2.55}$$

which we can transform into the relevant dimensions by rewriting $\vec{r}$ according to

$$E(x,z) = \int E(v)e^{ik(x\sin\theta+z\cos\theta)}d\theta. \tag{2.56}$$

This is commonly substituted by the Taylor approximations according to $\sin\theta = \theta$ or $\cos\theta = 1 - \frac{\theta^2}{2}$ [96]

$$E(x,z) = \int E(v)e^{ik(x\theta+z(1-\frac{\theta^2}{2}))}d\theta. \tag{2.57}$$

By comparing Eq. 2.52 and Eq. 2.57, we find the relation between θ and v as $(\theta = \lambda v)$, which leads to

$$E(x,z) = e^{\frac{i2\pi z}{\lambda}} \int E(v)e^{i2\pi xv}e^{-i\pi z\lambda v^2}dv. \tag{2.58}$$

Equation 2.58 can be rearranged to

$$E(x,z) = e^{\frac{i2\pi z}{\lambda}} F^{-1}\left\{E(v)e^{-i\pi z\lambda v^2}\right\}, \tag{2.59}$$

which in two dimensions corresponds to the paraxial approximation according to

$$E(x,y,z) = e^{\frac{i2\pi z}{\lambda}} F^{-1}\left\{E(u,v)e^{i\pi z\lambda(-u^2-v^2)}\right\}. \tag{2.60}$$

Non-paraxial propagation

In numerical calculations the increased complexity for fully applying the $\sin\theta$ or $\cos\theta$ terms in Eq. 2.56 is small. Also, later target will show a high sensitivity to the precise position in space x and y. Hence, we will introduce the plane wave in non-paraxial propagation along z [97]. In Eq. 2.56, we replace $(\cos\theta)$ with ($\sqrt{1-\sin^2\theta}$)

$$E(x,z) = \int E(v)e^{ik(x\sin\theta + z\sqrt{1-sin^2\theta})}d\theta. \tag{2.61}$$

By comparing Eq. 2.52 and Eq. 2.61 the relation between angle θ and spatial frequency v is $(\sin\theta = \lambda v)$, which leads to

$$E(x,z) = \int E(v)e^{i2\pi xv}e^{ikz\sqrt{1-\lambda^2v^2}}dv, \tag{2.62}$$

and a numerical implementation according to

$$E(x,z) = F^{-1}\left\{E(v)e^{ikz\sqrt{1-\lambda^2v^2}}\right\}. \tag{2.63}$$

In the (x,y)-plane, beam propagation along the z-direction is therefore computed according to

$$E(x,y,z) = F^{-1}\left\{E(u,v)e^{ikz\sqrt{1-\lambda^2u^2-\lambda^2v^2}}\right\}. \tag{2.64}$$

By comparing to Eq. 2.60 and 2.64, we can see the obvious difference between propagated plane wave in paraxial approximation and non-paraxial. In the diffractive coupling context to be introduced later, we need an accuracy of around 1 μm in the coupling position. For a focal distance of 18 mm, this corresponds to an angle accuracy of $\theta = tan^{-1}(\frac{1\times10^{-6}}{18\times10^{-3}}) = 0.003$ radians, showing that the error between the angle in paraxial approximations is larger than $\theta = 0.003$ radians. Therefore, paraxial effects need to be removed in the whole numerical simulation. As a result in second step, by multiplying the obtained phase shift from Eq. 2.64 (non-paraxial propagation) and initial beam in Fourier space, we propagate the beam $E(u,v,0)$ in the Fourier space and along z direction until $E(u,v,z)$.

3- The third step is a solution to transfer the propagated signal from Fourier domain into space. For this reason we need to apply inverse Fourier from second step $E(u,v,z)$ to obtain $E(x,y,z)$. By explained three steps above, in our numerical simulations in next Chapters, we propagate our original beam in the defined distance. Therefore in our setup in next Chapters wherever the propagation of beam is required, the three steps will be applied.

2.2.2/ EXAMPLE: DIFFRACTION BY GRATINGS

In this section, we intend to derive the grating formula. This example will be of relevance, since our diffractive coupling is based on optical gratings. We will describe an intuitive approach of the diffraction grating and finally we will demonstrate the grating equation with the plane-wave spectrum method.

2.2.2.1/ THE OPTICAL GRATING IN RAY-OPTICS

A grating is an optical device composed of a periodic structures (transmission grating or reflection grating), all spaced at the same distance, called the ”pitch” of the grating. In general, any object with a periodic pattern on its surface can act as a grating. A grating for visible light typically has a pitch in the order of the optical wavelengths [98, 99].
A grating diffracts or splits light into several beams in different directions. Directions of these beams depend on the wavelength of the light and the size of the slits and their spacing (pitch). Those diffracted waves interfere with each other to give a diffraction pattern on the screen [100–102]. The principle of the theoretical study of the grating is the same as that of the *Thomas Young's slits*, except that instead of having only two slits side by side, we have a greater number, see Fig. 2.7.

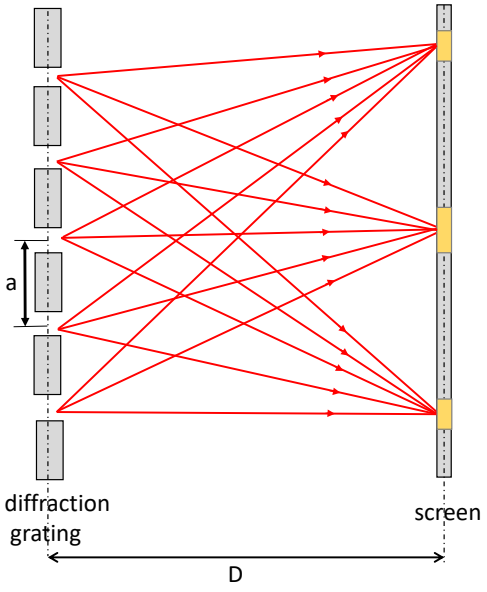

FIGURE 2.7 – Diffraction of light on an optical grating. In the left side the diffraction grating is located with period a. The right side is a screen to show the diffraction pattern. Distance between diffraction grating and the screen is D.

In this first approach, we try to know precisely in which directions the diffracted waves will interfere constructively. To do this, we study more precisely what happens for two neighbouring slits as showed on Fig. 2.8. We assume that the distance D between the grating and the screen is very large compared to distance a. The difference of distance between the two diffracted waves in Fig. 2.8 is

$$\delta = a \sin\theta, \tag{2.65}$$

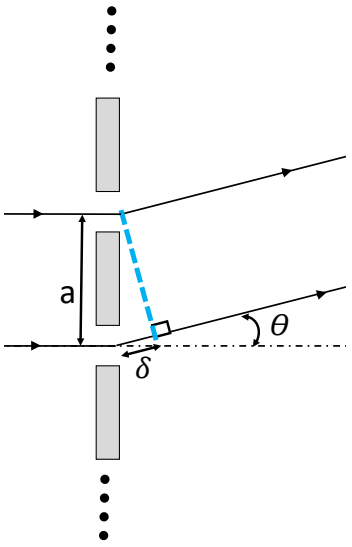

FIGURE 2.8 – Zoom on 2 slits in the optical grating. Distance between two slits is a. $|\delta|$ is difference of distance between the two diffracted waves. Angle θ is the angle between diffracted beam and optical axis.

where a is the grating period. The interference will be constructive on the screen for

$$\delta = m\lambda, \tag{2.66}$$

where m is an integer. From Eq. 2.65 and Eq. 2.66, the constructive interference is therefore characterised by

$$a \sin\theta = m\lambda. \tag{2.67}$$

In case of a transmission grating where the incident angle θ_0 is non zero, see Fig. 2.9, we can calculate as

$$\delta = mn - st, \tag{2.68}$$

by replacing mn and st we obtain

$$\delta = a\sin(\theta_0) - a\sin(|\ \theta_m\ |), \tag{2.69}$$

if $\theta_0 > 0$ then

$$\delta = a\sin(\theta_m) - a\sin(\theta_0), \tag{2.70}$$

by applying Eq. 2.66 in Eq. 2.70 we write

$$\sin(\theta_m) - \sin(\theta_0) = m\frac{\lambda}{a}. \tag{2.71}$$

For a reflection grating, as shown on Fig. 2.10, the grating equation is

$$\delta = mn - st, \tag{2.72}$$

by replacing mn and st we have

$$\delta = a\sin(\theta_m) - a\sin(|\ \theta_0\ |), \tag{2.73}$$

if $\theta_0 < 0$ hence

$$\delta = a\sin(\theta_m) + a\sin(\theta_0), \tag{2.74}$$

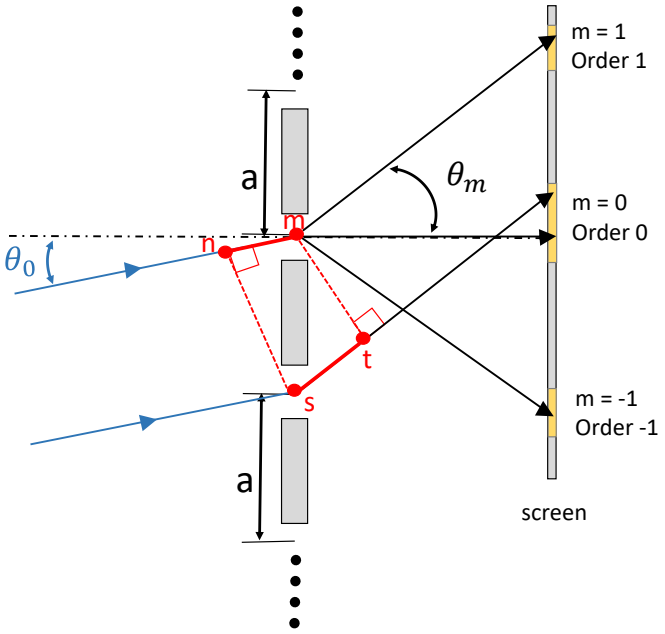

FIGURE 2.9 – Diffraction of light on two slits of the transmission grating. θ_0 is the incident beam which is non zero. θ_m is the angle between diffracted beam and optical axis. On the screen the diffractive orders $-1, 0, 1$ are established.

by applying Eq. 2.66 in Eq. 2.74 we write

$$\sin(\theta_m) + \sin(\theta_0) = m\frac{\lambda}{a}. \tag{2.75}$$

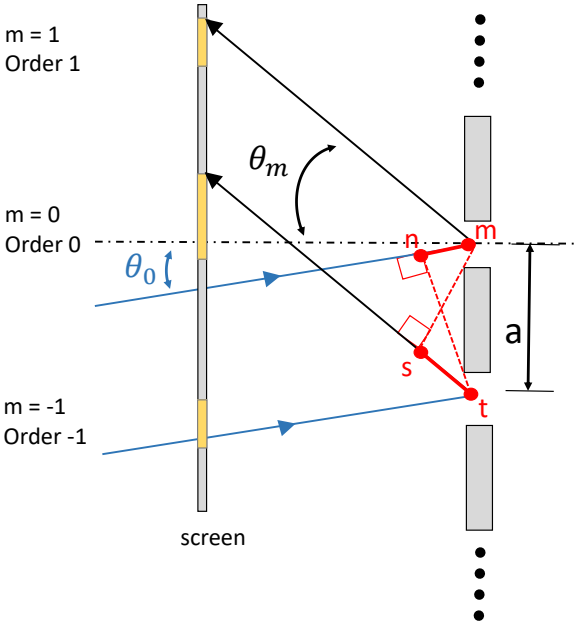

FIGURE 2.10 – Diffraction of light on two slits of reflection grating. θ_0 is the incident beam which is non zero. θ_m is the angle between diffracted beam and optical axis. On the screen the diffractive orders $-1, 0, 1$ are established.

2.3/ PRINCIPLE OF DIFFRACTIVE COUPLING

In the previous section, it was demonstrated the main results of diffraction theory. Using the plane-wave spectrum approach could deal with propagation in a uniform medium and the diffraction of a wave by a periodically structured object. The previous sections will be utilized to describe the principle of the diffractive coupling in this section, and ultimately arrive at how to model the connections between neurons in our optical hardware network. We implement a DOE, which will be considered as a transmission grating [103–105]. The idea of diffractive coupling is relatively simple: a matrix of optical emitters is sent through a grating (DOE), creating several diffractive orders [106–108]. Under certain conditions, we can observe an overlapping between different orders of diffraction of different emitters. This overlapping will be the basis of the coupling which defines the connectivity between neurons in our optical hardware network. Coupling among the neurons means that the states of the neurons affect the states of other neurons which separate information throughout the network. To model analytically the diffractive coupling, we will split this section in three parts. First, We present the scheme of diffractive coupling, the elements and their spatial positions. Second, We analyze the system by deriving of the rules and equations from the previous sections. We will explain the coupling procedure in a simple example in two steps. Third, We present the results of a simple analytical simulation, for explained configuration in the first part. This shows the principle and limits of diffractive coupling.

The concept of our diffractive coupling is schematically illustrated in Fig. 2.11. This figure is restricted to the y and z dimensions. The properties along x are identical to y and can therefore be omitted for illustration-simplicity. In this figure, there is an array of discrete optical emitters spaced in the (x, y)-plane with the period p^{array}. So, we define a $2D$ array of optical emitters such as Gaussian beam. The width of each emitter is $w = 3.5\ \mu$m at $\frac{1}{e^2}$, which is a common modal width for optical single mode emitters. The positions of emitters inside the array is given by $r_{\text{i,j}} = p^{\text{array}}$, where $i, j \in \{-N, N\}$ assign physical node positions to integer indices.
Emitters are located at z_{a} a focal distance away from the first lens or microscope objective (MO_1). We use infinity-corrected MOs and emulate the DOE as a periodic phase modulation with period of p^{DOE}. The DOE is located at z_{DOE}. After the DOE, all diffractive orders are established. Each diffractive order has a spatial position along y.
At a distance $(f_1 + D)$ of the emitters, the second microscope objective (MO_2) is situated, which images all diffractive orders onto the focal plane located at a distance of focal length f_2. The diffractive orders are located at the distance $(f_1 + D + f_2)$ of the emitters. Based on the spatial position of each diffractive order along y, we define the values $0, +1, -1$ as illustrated in Fig. 2.11. For instance the blue circle at z_{m} in the center got the value 0. The upper diffractive order (blue circle) has the value $+1$ and the lowest one has value -1.

As shown in the Fig. 2.11, the analysis of diffractive coupling is the distribution of the original emitter's field across various diffraction orders in plane $z = z_{\text{a}}$. If microscope objectives are ideal and without aberration we have a trigonometric equation for imaging

$$\tan(\Theta_{\text{i}}^{\text{im}}) = \frac{i \cdot p^{\text{array}}}{f_1}, \tag{2.76}$$

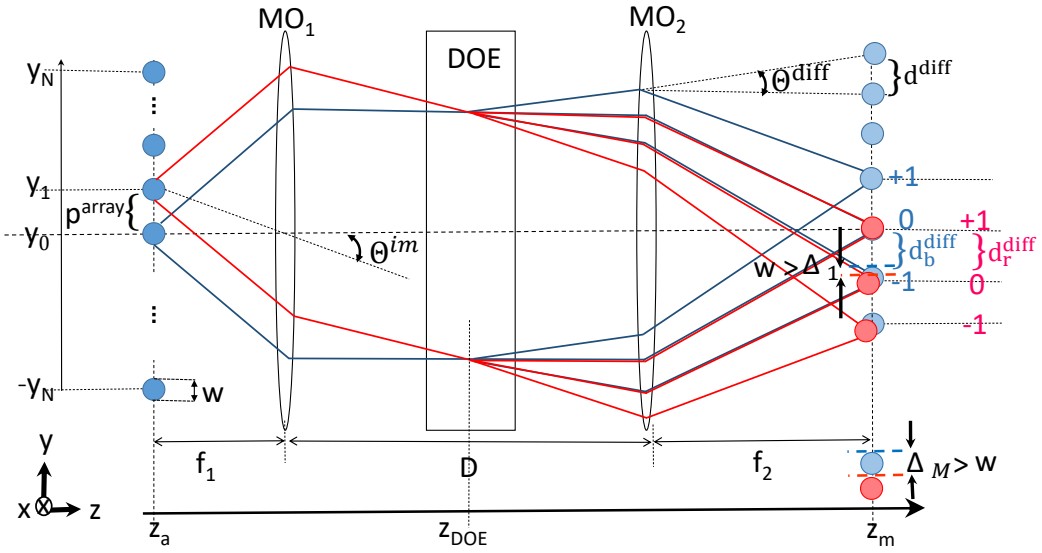

FIGURE 2.11 – Principle of diffractive coupling. Two emitters are imaged via f_1 and f_2, where f_1 is focal length of MO_1 and f_2 is focal length of MO_2. The DOE is situated between MOs and creates three discrete diffractive orders $-1, 0$ and $+1$. The width of emitters are called w which is the Gaussian beam width at $1/e^2$. $\Theta_{\mathrm{i}}^{\mathrm{im}}$ is the angle between the principal rays of neighboring emitters and Θ^{diff} is the angle between diffractive orders. The distance between two emitters is p^{array} and between two diffractive orders is d^{diff}. $d_{\mathrm{b}}^{\mathrm{diff}}$ is the distance between orders $m = 0$ and $m = -1$ for emitter y_0 in blue color. $d_{\mathrm{r}}^{\mathrm{diff}}$ is the distance between orders $m = 0$ and $m = 1$ for emitter y_1 in red color.

where Θ^{im} is the imaging angle. Based on grating Eq. 2.71, we can write

$$\sin(\Theta_{\mathrm{i,m}}^{\mathrm{diff}}) - \sin(\Theta_{\mathrm{i}}^{\mathrm{im}}) = m\frac{\lambda}{p^{\mathrm{DOE}}}, \tag{2.77}$$

where Θ^{diff} is the diffraction angle after the grating. We can define a trigonometric equation for diffraction

$$\tan(\Theta^{\mathrm{diff}}) = \frac{d^{\mathrm{diff}}}{f_2}, \tag{2.78}$$

where d^{diff} is distance between two diffractive orders. To understand better the procedure of diffractive coupling, we will provide here an example. We consider an emitter being in center of y with indice $i = 0$ in Fig. 2.11. In this figure, the emitter y_0 is collimated in front of MO_1. After, collimated beam goes through the DOE. Next, light is diffracted from the DOE in directions corresponding to $0, +1$ and -1 (blue circles in Fig. 2.11). The second microscope objective images the created diffractive orders in plane z_{m}. In Eq. 2.76, in location y_0 we know $p^{\mathrm{array}} = 0$ therefore, $\Theta^{\mathrm{im}} = 0$ and for diffractive order in position $m = -1$ we rewrite Eq. 2.77 as

$$\sin(\Theta_{0,-1}^{\mathrm{diff}}) = -\frac{\lambda}{p^{\mathrm{DOE}}}. \tag{2.79}$$

Equation 2.78 for diffractive order $m = -1$ for the blue node is

$$\tan(\Theta_{0,-1}^{\mathrm{diff}}) = \frac{d_{\mathrm{b}}^{\mathrm{diff}}}{f_2}, \tag{2.80}$$

where $d_{\mathrm{b}}^{\mathrm{diff}}$ is the distance between orders $m = 0$ and $m = -1$ for emitter y_0 in blue color and ($d_{\mathrm{b}}^{\mathrm{diff}} = 12.2\ \mu\mathrm{m}$) which has the same value as our experiment and $f_2 = 45$ mm, we can compute

$$\Theta_{0,-1}^{\mathrm{diff}} = \tan^{-1}\left(\frac{12.2 \cdot 10^{-6}}{45 \cdot 10^{-3}}\right) = 0.27\ radiant. \tag{2.81}$$

Now by replacing $\Theta_{0,-1}^{\mathrm{diff}}$ in Eq. 2.79 and $\lambda = 660$ nm, we can compute the DOE's period as

$$p^{\mathrm{DOE}} = \frac{660 \cdot 10^{-9}}{\sin(0.27)} = 2.4\ \mu m. \tag{2.82}$$

The values (f_1, f_2, λ) from Fig. 2.11 were selected identical to our experimental setup, which will be explained in the next Chapter. Now, we use the same process from the example above for the emitter with indice $i = 1$ located at $y_1 = p^{\mathrm{array}}$ in Fig. 2.11. For this emitter the established diffractive orders in z_{m} are $0, +1$ and -1 (in red color). In location $y_1 = p^{\mathrm{array}}$ we know $\Theta^{\mathrm{im}} \neq 0$ so we can rewrite Eq. 2.76 as

$$\tan(\Theta_1^{\mathrm{im}}) = \frac{p^{\mathrm{array}}}{f_1}, \tag{2.83}$$

where $p^{\mathrm{array}} = d_b^{\mathrm{diff}}$, Eq. 2.83 is

$$(\Theta_1^{\mathrm{im}}) = \tan^{-1}\left(\frac{12.2 \cdot 10^{-6}}{18 \cdot 10^{-3}}\right) = \tan^{-1}\left(0.67 \cdot 10^{-3}\right). \tag{2.84}$$

In Fig. 2.11 and after the DOE, Eq. 2.76 for emitter located in y_1 and $m = 0$ will be

$$\sin(\Theta_{1,0}^{\mathrm{diff}}) - \sin(\Theta_1^{\mathrm{im}}) = 0. \tag{2.85}$$

Equation 2.78 for red diffractive order in position $m = 0$ will be

$$\tan(\Theta_{1,0}^{\mathrm{diff}}) = \frac{d_{\mathrm{r}}^{\mathrm{diff}}}{f_2}, \tag{2.86}$$

from equation 2.85, we obtained $\sin(\Theta_{1,0}^{\mathrm{diff}}) = \sin(\Theta_1^{\mathrm{im}})$. As a result, we adjust the angle between the diffractive orders to the angle from collimating the emission of neighboring nodes. We also calculated $\sin(\Theta_1^{\mathrm{im}})$ in Eq. 2.84. So, the value of $d_{\mathrm{r}}^{\mathrm{diff}}$ is

$$d_{\mathrm{r}}^{\mathrm{diff}} = \tan(0.67 \times 10^{-3}) \times 45 \cdot 10^{-3} = 30.5\ \mu m. \tag{2.87}$$

The magnification effect of both MOs can be computed as

$$M = \frac{f_2}{f_1} = \frac{45 \cdot 10^{-3}}{18 \cdot 10^{-3}} = 2.5. \tag{2.88}$$

By dividing our obtained result by magnification M, we obtain $d_{\mathrm{r}}^{\mathrm{diff}} = \frac{30.5 \cdot 10^{-6}}{2.5} = 12.2\ \mu m$. In imaged plane $\mathrm{z_m}$ there will be overlapping for instance between -1 order (blue circle) from $\mathrm{y_0}$ with 0 order (red circle) from $\mathrm{y_1}$. In this example we define a coupling criteria between (diffractive order $m = -1$ from emitter $i = 0$) and (diffractive order $m = 0$ from emitter $i = 1$). This criteria for coupling between diffractive orders can be defined as

$$|\Delta_1| = |d_{\mathrm{r}}^{\mathrm{diff}} - d_{\mathrm{b}}^{\mathrm{diff}}|, \tag{2.89}$$

where $|\Delta_1|$ is the mismatch coupling between diffractive order $m = -1$ in blue circle and diffractive order $m = 0$ in red circle, showed in Fig. 2.11. This value illustrates the coupling mismatch between two mentioned diffractive orders. As we know in this example $|\Delta_1| = 0$ and the two diffractive orders (blue circle in $m = -1$ and red circle in $m = 0$) have 100 percent overlap (coupling). As a result, coupling was optimized for the central emitter. The goal of the last example in two steps was description of coupling with a simple method. But the analytical computation of Fig. 2.11 is not limited to this simple example. We require to calculate coupling mismatch between diffractive orders for all emitter positions $\{-y_N, ..., y_N\}$.
For emitters placed away from the optimal position (central emitter), the overlapping between neighboring emitters will gradually be reduced. That is due to the different trigonometric relationships for imaging and diffraction as we can see in Eqs. 2.76 and 2.79. In these equations we avoided the paraxial approximations so $tan\theta \neq \theta$, $sin\theta \neq \theta$ and $|tan\theta - sin\theta|$ is a nonlinear function. Therefore, that gradual reduction of overlapping between neighboring emitters will be nonlinear.
We defined $w = 3.5$ μm as the width of each emitters. As we explained the overlapping between neighboring emitters will be reduced. We intend to find the limitation of the overlapping between diffractive orders. Two diffractive orders can have the overlapping where mismatch coupling $|\Delta| < 3.5$ μm. If the size of $|\Delta|$ is larger than the width of an emitter, coupling between the neurons is terminated. Fig 2.11 shows two diffractive orders far from center without any overlapping since ($|\Delta_M| > w$) in this location.
In next part we will display the result of a simple analytical simulation and the limitation of diffractive coupling from parameter $|\Delta|$. We will simulate the parameter $|\Delta|$ for all emitter positions in z_a.

We show the limit of our diffractive coupling technique used in our optical neural network. In this simulation we used experimental parameters for the wavelength, the first focal length and the second focal length as we calculated during the previous paragraphs. We consider here an array of 1681 emitters with the constant periodic $p^{\text{array}} = 12.2$ μm. To measure analytically mismatch $|\Delta|$, we use Eq. 2.89.
In Fig. 2.12, we plot the mismatch $|\Delta|$ as function of emitter position y. We can see that the mismatch increases when the emitters position diverges far from the optical axis. When the mismatch becomes greater than the emitters size w, we cannot generate the coupling which means that there are no connections between neurons (see Fig. 2.12).
The result of this simulation determine our limitation of coupling mismatch. In our experimental set up, the size of emitters $w = 3.5$ μm. Thus, to have overlapping, the mismatch must be less than 3.5 μm. This criteria will be satisfied for all the emitters located within between -8.2 mm $< y < 8.2$ mm. The mismatch coupling between -8.2 mm $< y < 8.2$ mm is less than the width of emitter so, inside this area there is overlapping between the neurons. But outside of -8.2 mm $< y < +8.2$ mm, neurons do not have overlapping since $|\Delta| > w$. As a result, the coupling in our hardware neural network is valid for emitters located inside a range of 16.4 mm $(\pm 8.2\text{mm})$. In this space where distance between diffractive order is 12.2 μm, we can calculate the maximum number of neurons by $\left(\frac{16.4 \cdot 10^{-3}}{12.2 \cdot 10^{-6}} = 1340\right)$. As a result, in an area of $(16.4 \text{ mm})^2$ we can obtain $1,795,600$ neurons.

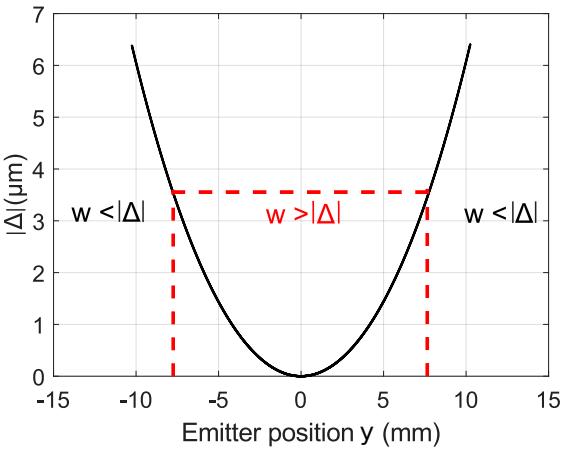

FIGURE 2.12 – Mismatch coupling between diffractive orders versus emitter position in z_a. Red dash lines identify the limitation of coupling in our hardware neural network. Inside the red dash line, there is coupling for emitters in the position of -8.2 mm $< y < +8.2$ mm. While outside of red dash line, there is no overlapping between the diffractive orders.

2.4/ CONCLUSION

In this Chapter, we derived the main results of the diffraction theory that we will use to develop our optical hardware neural network. We started from Maxwell equations in matter, and derived the propagation equation and the Helmholtz equation. We reminded the Huygens principle, and showed how to derive the Rayleigh-Sommerfeld integral, the Fresnel diffraction formula and the Fraunhofer formula. Then, we detailed the plane-wave spectrum technique which will be the basic tool in the following chapters. We highlighted the principle of the optical grating since the diffractive coupling used in our optical hardware neural network is based on such devices.
We finished this Chapter by describing the diffractive coupling. We used a simple scheme to explain the concept of diffractive coupling. Then, we explored the principles and limits of diffractive coupling by applying the analytical solution. We investigated how many neurons can be maximally coupled, which is limited by the coupling mismatch parameter $|\Delta|$. We obtained that if $|\Delta|$ is increased, the strength of our coupling is reduced. Finally, there is a point where coupling was terminated. In our setup where $|\Delta| > w$, (w is size of an emitter), coupling was terminated. Results of the simulation demonstrated that the coupling was terminated for $| y |< 8.2$ mm. We conclude that the coupling is valid for all emitters located in 16.4 mm and for more than 1340 photonic neurons. Such this area $(16.4 \text{ mm})^2$ allows the creation of neural network larger than $1,795,600$ photonic neurons.

3

EXPERIMENTAL CHARACTERIZATION OF DIFFRACTIVE COUPLING

In this Chapter, we experimentally investigate the concept of diffractive coupling which was presented analytically in Chapter 2. We intend to confirm the maximum number of photonic nodes which can be coupled successfully. We will in particular highlight the experimental limitations and introduce and evaluate ways to avoid them. We will present two different experimental setups: once we emulate photonic neurons by translating a single mode optical emitter across the objective plane, once we emulate the effect of photonic neurons at different positions by tilting the grating. Both approaches isolate different features of the concept. The setup describes in this Chapter is an essential part of our hardware reservoir computing system in Chapter 5.

3.1/ TRANSLATE POINT SOURCE ALONG THE (x, y)-PLANE

In this section, we evaluate diffractive coupling of an emitters array via a DOE with a periodic structure. In order to emulate the different emitters of the photonic neuron array, a single mode source in translated across the (x, y)-plane will be translated. The goal is to determine the maximum size of an emitter array which allows diffractive coupling and to quantify the physical mechanisms behind potential limitations.

3.1.1/ EXPERIMENTAL SETUP

The experimental setup is illustrated in Fig. 3.1 [50]. In this setup, we use a single mode fibre to emulate different elements inside an array of emitters. The idea is to translate the source (optical mode of the fiber) in the (x, y)-plane to characterize the coupling and to identify the impact of different emitter positions. The single mode fiber is located in the focal plane of the first microscope objective (MO_1), and hence a photonic neuron's emission is collimated behind MO_1. The beam propagates through the DOE, which has a structure with period p^{DOE}. All diffracted orders are imaged on to the camera in the (x', y') plane by a second microscope objective (MO_2). This system is symmetric between x and y, the experiment is limited to investigate effects of translation along y; translation along x is assumed to have an identical effect. In this setup, we use a laser diode (Thorlabs LP660 SF20, $\lambda = 662.1$ nm) as our source [109, 110]. A single mode fiber guides the emission of the laser with mode field diameter at its output of (4 ± 0.5) μm. To move the optical fibre,

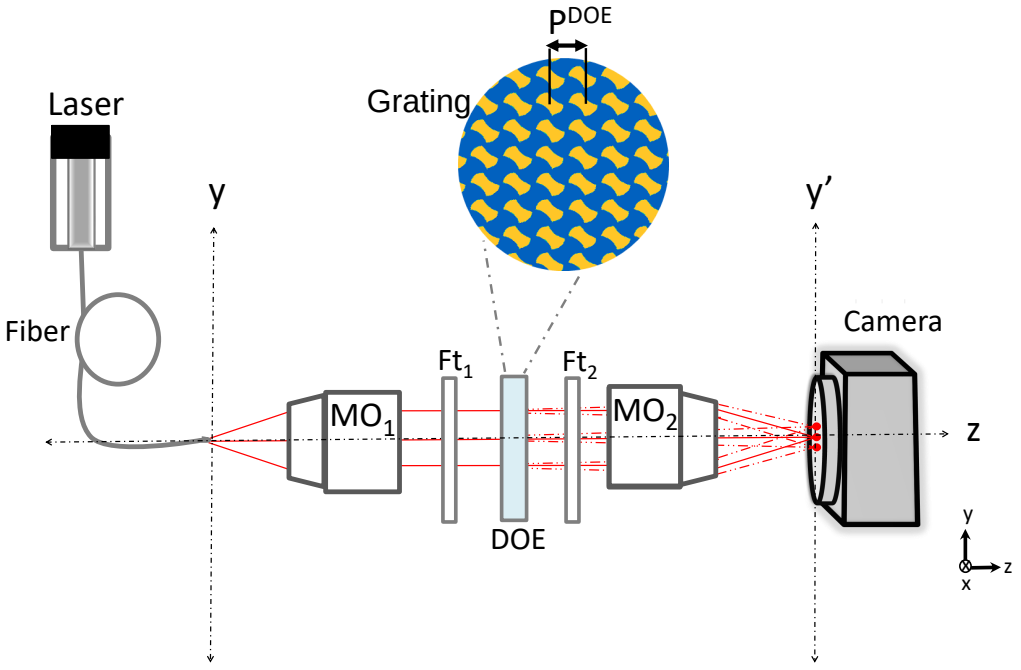

FIGURE 3.1 – Experimental scheme to test limits of diffractive coupling.

it is necessary to have an $(x-y)$-stage to determine precisely (micrometer precision) the displacement in each direction. The fibre is therefore mounted on a differential $(x-y)$-stage (Thorlabs ST1XY-D/M). The MO_1 (Olympus RMS10X) has a focal length $f_1 = 18$ mm, a numerical aperture $\mathrm{NA}_1 = 0.25$ and working distance WD $= 18.5$ mm. The DOE used to distribute the incident beam along an array of (3×3) of diffractive orders has the following reference HOLOOR MS-443-650-Y-X. The MO_2 (Olympus RMS4X) has a focal length $f_2 = 45$ mm, a numerical aperture $\mathrm{NA}_2 = 0.1$ and a working distance WD $= 10.6$ mm. We use two neutral density filters (Ft_1, Ft_2) to avoid saturation of the camera (Thorlabs NE 40A, NE 30A), both with a clear aperture diameter of 25 mm. Finally, a CMOS camera (IDS U3-3482LE-M) is used to record the image inside the focal plane of MO_2. This camera has 2560×1920 pixels with a pixel spacing of 2.2 μm.

3.1.1.1/ IMAGE MAGNIFICATION

At the initial stage we determine the magnification of our experimental setup [111, 112]. For that, we remove temporarily the DOE as shown in Fig. 3.2. We set the single mode fiber on the optical axis of the setup and in the focal plane of MO_1. The image is recorded by the camera and shown in Fig. 3.3. The recorded beam at the central region of the camera corresponds to $(0,0)$ in the plane $(x'-y')$. According to Fig. 3.3, the recorded beam has a width of 11 μm at $\frac{1}{e^2}$. Furthermore, MO_1 has a magnification of $\mathrm{M}_1 = 10$ whereas MO_2 has a magnification $\mathrm{M}_2 = 4$. Thus, the magnification effect M of the setup is

$$M = \frac{f_2}{f_1} = \frac{M_1}{M_2} = 2.5. \tag{3.1}$$

Since, the mode field diameter provided by the manufacturer is (4 ± 0.5) μm, by considering the magnification effect, we must expect an image beam with a width of (10 ± 1.25) μm

on the camera. The experimental width measured previously 11 μm is well within the range provided by the manufacturer. To conclude, the magnification characterized by the experiment is in agreement with the theoretical expectation.

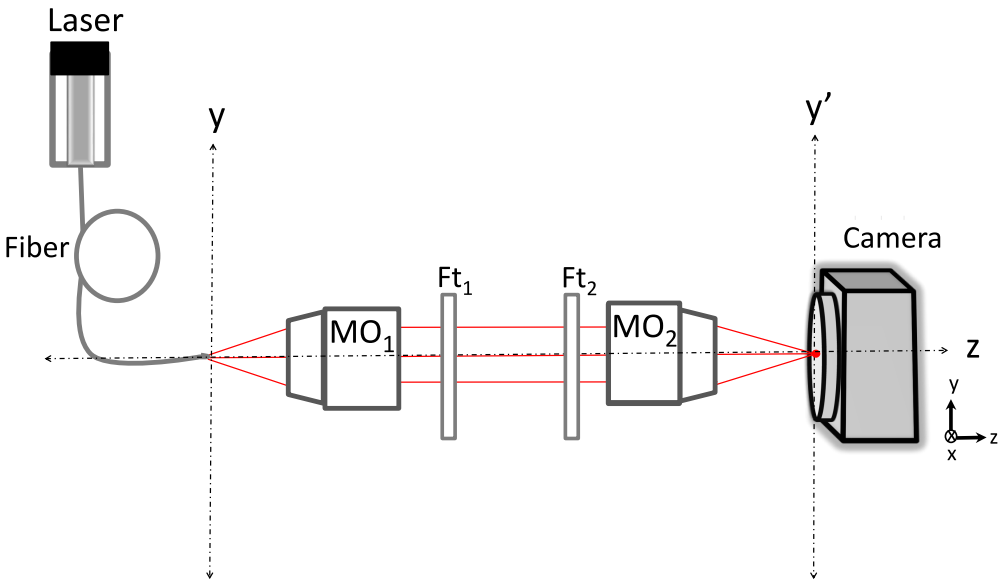

FIGURE 3.2 – Experimental setup without the DOE to determine the magnification.

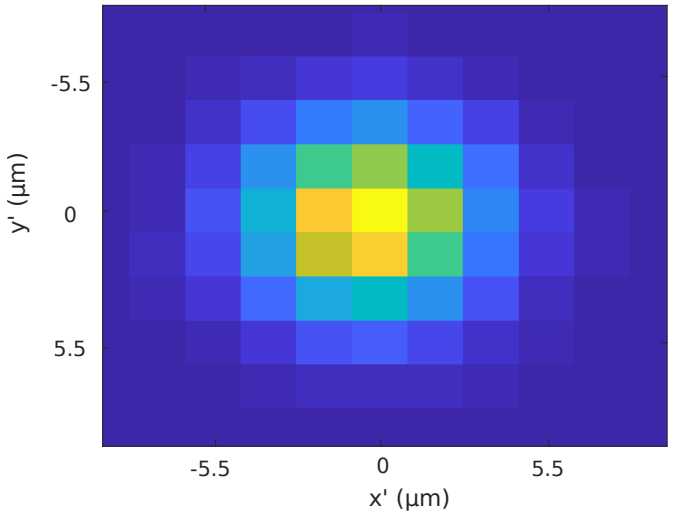

FIGURE 3.3 – Recorded beam in center of camera without the DOE effect.

3.1.1.2/ IMPACT OF DOE

To characterize the effect of the DOE, we include it in the middle of collimated area, i.e. located centered between the two microscope objectives, see in Fig. 3.4. The emitter is once again at position $(0, 0)$ of the $(x - y)$-stage, and the diffractive orders are imaged on camera as displayed in Fig. 3.5. The central position of camera corresponds to pixel

$(1187, 967)$, which therefore sets the reference position $(0, 0)$ in the $(x' - y')$-plane. We observe 9 diffractive orders on the camera and from Fig. 3.5, we can calculate the distance between two diffractive orders. The physical distance between two diffractive orders is around 30.8 μm as illustrated in Fig. 3.5. Furthermore, compensating for magnification M this corresponds to $|d^{\text{diff}}| \sim \frac{30.8\ \mu m}{2.5} \sim 12.3$ μm. All these calculations are of low resolution as they are based on simply identifying the pixel in each diffractive order with the highest intensity.

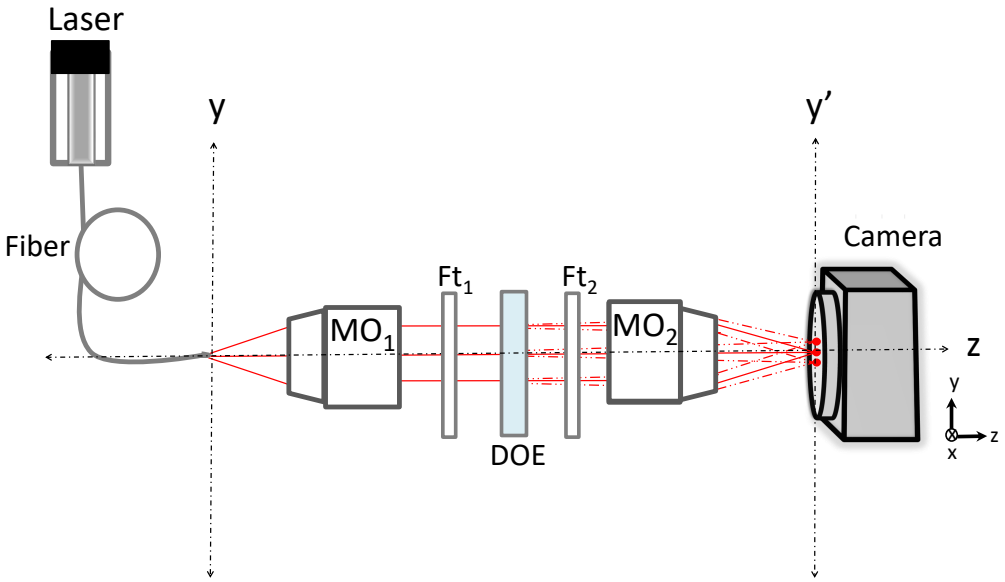

FIGURE 3.4 – Experimental setup to analyze diffractive orders on the camera.

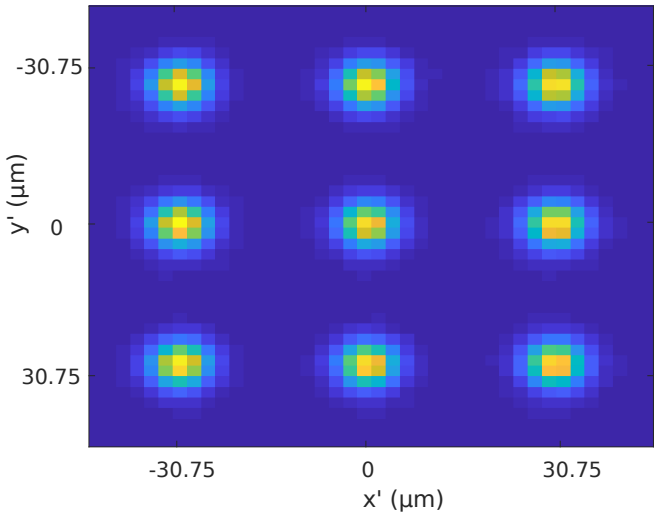

FIGURE 3.5 – Experimental diffractive orders imaged on the camera.

3.1.1.3/ VIGNETTING

We will now focus on the properties of the imaging generated by off-axis emitter position. We will also translate the emitter source along the y direction and image the generated beam on the camera (see Fig. 3.6). First we use an emitter being placed in $y_1 = 0.88$ mm. The collimation distance between the two MOs is $L = 100$ mm. As we can see in Fig. 3.6, L is an important parameter because if this distance is too long, there will be a strong vignetting on MO_2. Indeed, the beam will pass through the outer edges of MO_2 and such vignetting will impact the quality of the diffractive orders as shown in Fig. 3.7 for $y_1 = 0.88$ mm. For an emitter position far from the optical axis $y_2 = 1.1$ mm, beam vignetting is so strong that the diffractive orders can not any longer be clearly identified (see Fig. 3.8).

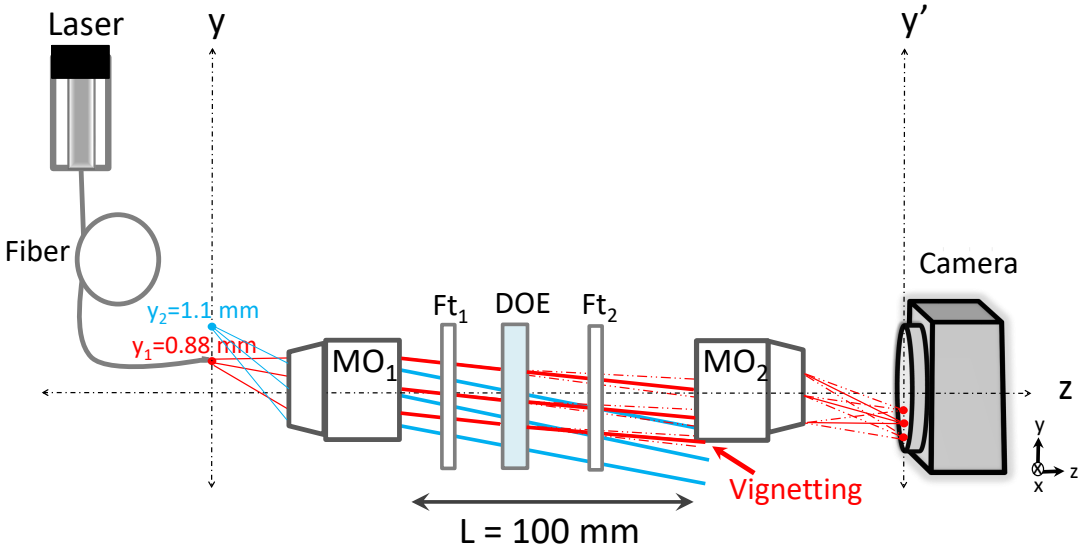

FIGURE 3.6 – Experimental setup using off-axis emitters.

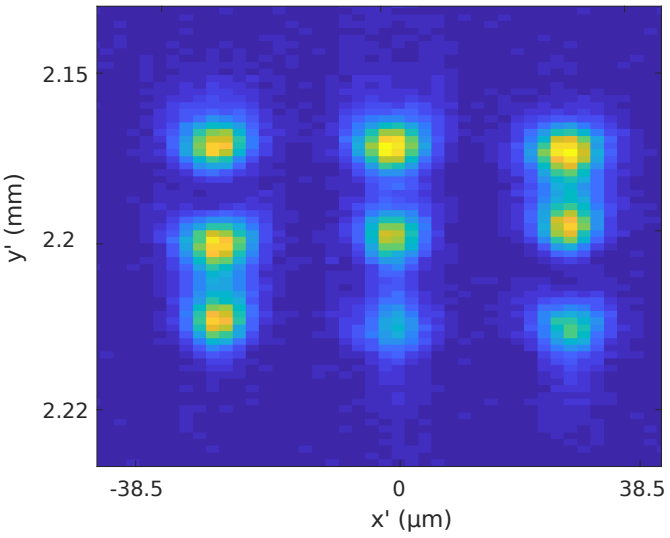

FIGURE 3.7 – Diffractive orders for $y_1 = 0.88$ mm.

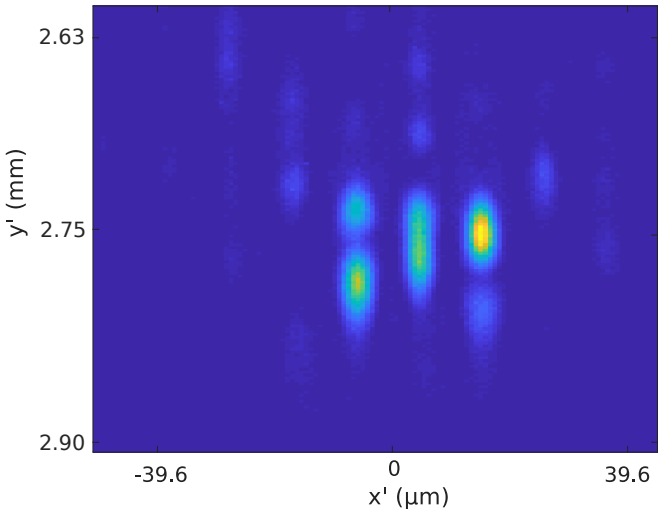

FIGURE 3.8 – Diffractive orders for $y_2 = 1.1$ mm.

3.1.1.4/ IMPACT OF THE DISTANCE BETWEEN TWO MICROSCOPE OBJECTIVES

To understand the problem of vignetting, we study the relevant optical rays as shown in Fig. 3.9. In this figure, θ is the acceptance half angle of the objective, R_1 is radius of MO_1 and R_2 is radius of MO_2. We investigate here at which distance L the beam vignetting occurs for a given emitter position y. The numerical aperture NA_1 of the first microscope objective is given by

$$NA_1 = n_\mathrm{i} \sin\theta, \tag{3.2}$$

where n_i is the index of refraction of the immersion medium, which in our experiment is air, hence $n_\mathrm{i} = 1$. We know

$$R_1 = f_1 \tan\theta, \tag{3.3}$$

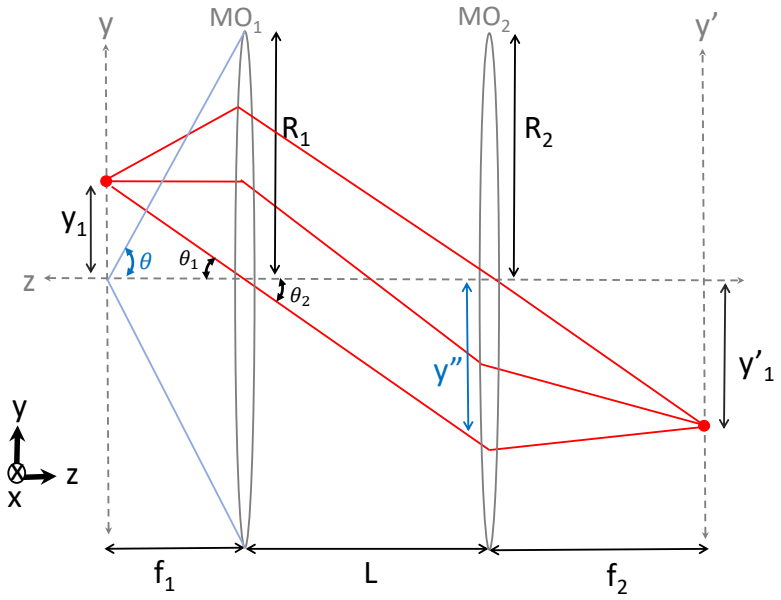

FIGURE 3.9 – Scheme to adjust the maximum distance between MOs (L) for emitters far from center and based on the diameter of MO_2 (R_2).

and inserting Eq. 3.2 into Eq. 3.3, this result in

$$R_1 = f_1 \tan\left(\sin^{-1}(NA_1)\right). \tag{3.4}$$

We obtain $R_1 = 18$ mm $\times \tan\left(\sin^{-1}(0.25)\right) = 4.6$ mm and $R_2 = 45$ mm $\times \tan\left(\sin^{-1}(0.1)\right) = 4.5$ mm as the clear aperture radius of both microscope objectives, which are the maximum extend of an optical input wave the MOs can accept without vignetting. If by translating the emitter the radius of its collimated beam remains below R_1 and R_2, we avoid the vignetting effect. On the other side we know

$$\tan\theta_1 = \frac{y_1}{f_1}, \tag{3.5}$$

and

$$\tan\theta_2 = \frac{y''}{L}. \tag{3.6}$$

Based on the Snell's law, relationship between the angles of incidence and refraction for MO_1 in Fig. 3.9 is $\sin\theta_1 = \sin\theta_2$ and therefore $\theta_1 = \theta_2$. Then, from Eq. 3.5 and Eq. 3.6 we can write

$$y_1 L = y'' f_1, \tag{3.7}$$

we will consider fixed y_1 and f_1. If the distance L increases, the collimated beam's waist radius y'' at the clear aperture input of MO_2 will increase. Because of limited size of R_2, beam vignetting can potentially occur. The maximum value of L is $y''_{\max} = R_2$ and by using Eq. 3.7 we obtain

$$L_{\max} = \frac{f_1 R_2}{y_1}, \tag{3.8}$$

thus, for $y_1 = 0.88$ $L_{\max} = 94$ mm and for $y_2 = 0.98$ $L_{\max} = 84$ mm. These numbers clearly demonstrate why for $L = 100$ mm, the beam of an emitter located at y_1 only exhibits slight beam vignetting, while for y_2 the effect is very pronounced.

3.1.1.5/ REDUCE DISTANCE BETWEEN MOS

In our experiment, we seek the fundamental limitation of diffractive coupling due to the increasing relevance of non-paraxial and to achieve this objective, we must move the point source around 8 mm as we obtained in Chapter 2. In order to achieve this purpose, we must significantly reduce L. In that case, to reduce L we remove filters Ft_1 and Ft_2 of Fig. 3.1 and instead use fiber based attenuators (Thorlabs Attenuator FA25T) with 25 dB and (Thorlabs Attenuator FA10T) with 10 dB. In Fig. 3.10 the experimental setup is described. Distance L is reduced to 55 mm and the DOE again located in the middle between the two MOs. Nine diffractive orders are imaged on the camera (see Fig. 3.11). For $L_{\max} = 55$ mm, Eq. 3.8 gives $y_1 = \frac{R_2 f_1}{L_{\max}} = 1.4$ mm, which corresponds to the maximum displacement allowing to avoid strong vignetting effect.
However, we also can be limited by the size of the camera's detection array. Due to the imaging system's to magnification $(M = 2.5)$, the extension of the emitter array which can still be imaged onto the camera in y is $\frac{2560\times2.2\ \mu m}{2.5} = \pm 1.2$ mm and in x is $\frac{1920\times2.2\ \mu m}{2.5} = \pm 0.85$ mm. The imaged beam in Fig. 3.11 showed the 3×3 diffractive orders without vignetting effect for emitter in the position $y_1 = +1.16$ mm. This displacement corresponds approximately to our limit to record the diffractive orders. This limit due to the size of camera therefore prohibits the full characterization of the limitations expected of diffractive coupling. In order to overcome this practical limitation, we can mount the camera to an $(x - y - z)$ stage. Whenever imaged beam comes to the edge of the camera by moving emitter in y direction, we can shift camera further in the y direction by the stage.

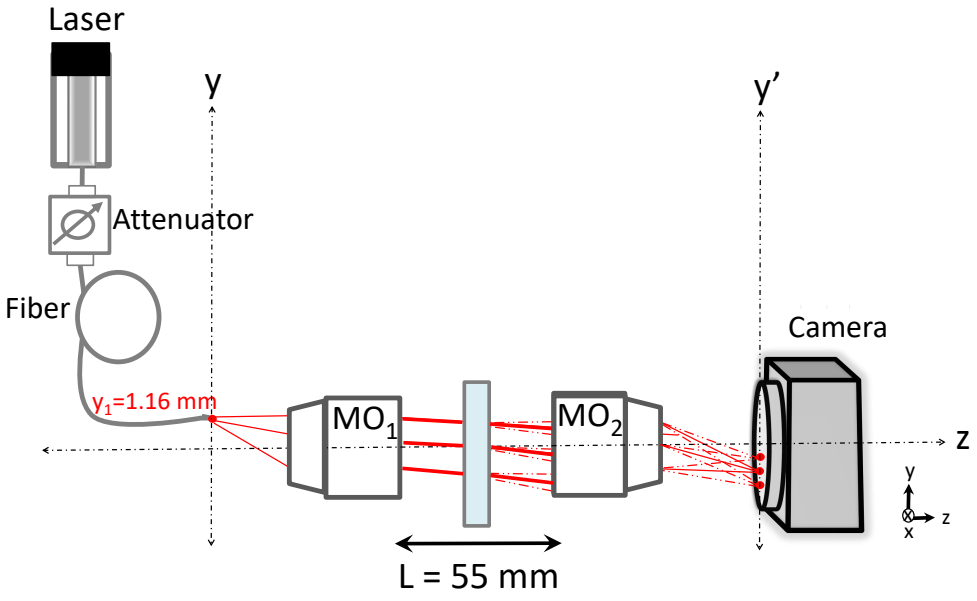

FIGURE 3.10 – Setup after removing filters and adding the attenuator.

3.1.2/ RESULTS WITH EMULATING OPTICAL EMITTERS

In last sections, we studied the limitations (e.i. vignetting and the size of camera) of our setup by translating of the optical fiber's position. In this section, we obtain the detailed result from moving a single mode fibre along the (x, y)-plane. We measure in detail the

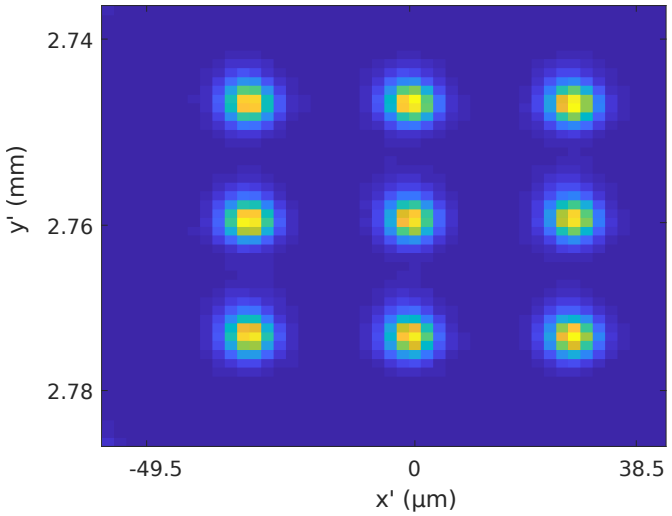

FIGURE 3.11 – Diffractive orders on the camera for emitter in y_1 = 1.16 mm.

distance between diffractive orders, the diffractive coupling mismatch and the width of diffractive orders.

3.1.2.1/ IMPACT OF BEAM VIGNETTING

In this section, we consider the following case: we emulate an array of 10 discrete sources with different positions in y direction. The displacements in y direction are in the range $(-1.14, +1.01)$ mm. First, we register on the camera the position of each source point by removing the DOE. Then, we add the DOE and record the 9 diffractive orders on the camera for each of the 10 emitters. We process the obtained images by fitting each diffractive order with a two dimensional Gaussian

$$I(x,y) = I_0 \; exp\left(\frac{x - x_0}{w_1}\right)^2 exp\left(\frac{y - y_0}{w_2}\right)^2, \tag{3.9}$$

where I_0 is the maximum intensity, x_0 and y_0 are the center position inside the x and y plane, respectively. w_1 and w_2 are radius in x and y direction, respectively. Figure 3.12 shows the distance between 9 diffractive orders and nomenclature in which the orders are labeled. Six distances between neighbouring diffractive orders can be measured in y direction (red arrows) with distance $|d^{\text{diff}}|$. To obtain a strong optical coupling in our device, the distances $|d^{\text{diff}}|$ must remain constant. In next steps, we explain 3 different measurements from our experimental data points. In order to measure accurately the distance between two orders, we fit the intensity profile and the results are shown in Fig. 3.13 for different emitter positions. The black dash line in the Fig. 3.13 is the constant distance ($|d^{\text{diff}}| \sim 12.2$ μm) for which in a periodic array coupling is optimal for emitter position between $y_0 = 0$ mm and $y_{10} = 1.14$ mm. In Fig. 3.13 for emitters $y < 1$ mm, experimental distances between diffractive orders $|d^{\text{diff}}|$ follow the black dash line with a slight difference. This difference is due to:

- First, the detection noise of the camera which will be obtained as a correct value in next Chapter.

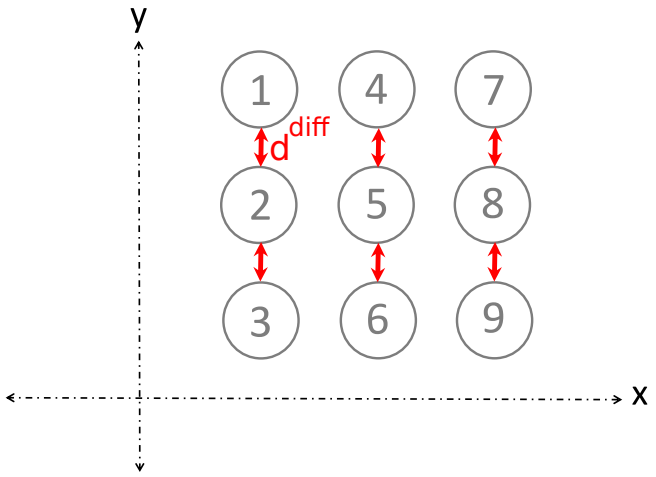

FIGURE 3.12 – Distance between diffractive orders $|d^{\text{diff}}|$ in the y plane (red arrows) for 9 diffractive orders.

- Second, the limited precision of the fitting routine. This limitation is the consequence of the numerical algorithm to fit the experimental data with Eq. 3.9.

However, for emitters $y > 1$ mm, this difference starts growing significantly. This behaviour occurred due to vignetting effect in our experiment for $y > 1$ mm, introduced in the section 4.1.1.3. As shown previously, vignetting destroyed the quality of diffracted orders on camera.

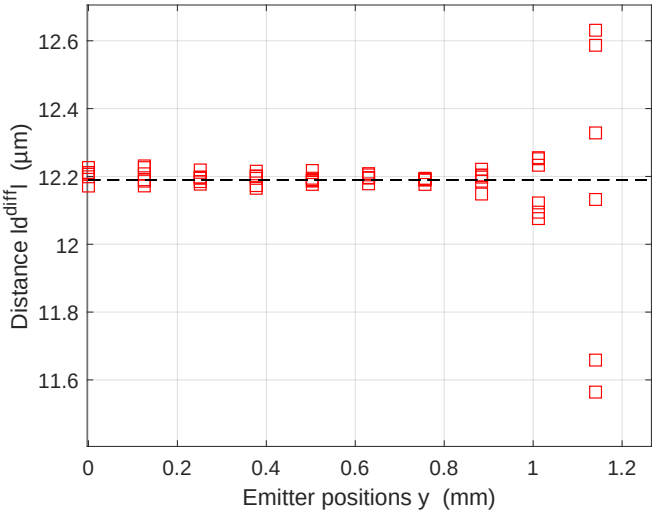

FIGURE 3.13 – Distances between neighbouring diffractive orders $|d^{\text{diff}}|$ as a function of the emitter position. Black dash line: the constant distance for which in a periodic array's coupling is optimal.

In another measurement in Fig. 3.14 mismatch $|\Delta|$ versus emitter positions y from experimental data are plotted in a double-logarithmic scale with red stares while, the analytical limit is plotted as the black dash line. The analytic curve in Fig. 3.14 shows a clear polynomial increase of mismatch $|\Delta|$. However, we observe a number of additional features in the experimental data. For emitter displacements $y < 0.5$ mm, mismatch $|\Delta|$ is below 30 nm but still larger than the mismatch predicted analytically, which is due to noise of the

camera and the limited precision of the fit algorithm. In contrast, for $y > 0.5$ mm in Fig. 3.14 we see a strong divergence of the experimentally obtained $|\Delta|$ from the theoretical prediction (due to effect of noise and fit algorithm). This divergence arrives at $|\Delta| = 0.3\ \mu$m for $y_{10} = 1.14$ mm because of the vignetting effect in our setup. We therefore need to solve vignetting problem caused by the small radius ($\mathrm{R}_1 = 4.6$ mm and $\mathrm{R}_2 = 4.5$ mm) or low NA ($\mathrm{NA}_1 = 0.25$ and $\mathrm{NA}_2 = 0.1$) of the MOs.

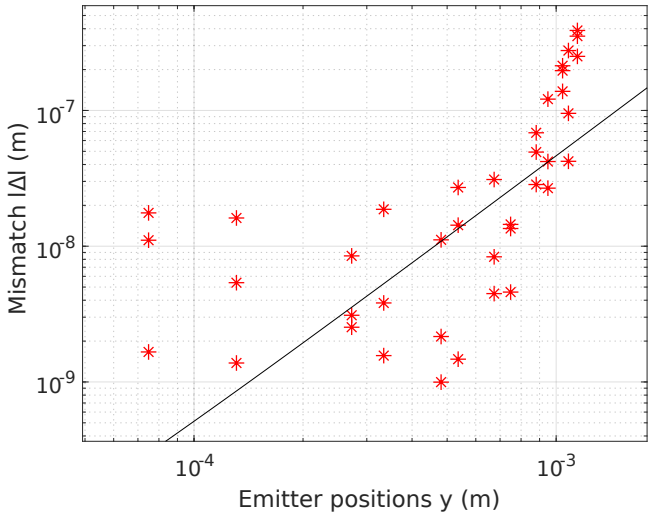

FIGURE 3.14 – Mismatch $|\Delta|$ versus emitter positions y from experimental data in a double-logarithmic scale (red stars). Theoretical data are plotted as a black line.

The quality of optical coupling depends on several parameters in the system. One of this significant parameters is the width of the focal spot [113]. In this last measurement, the width of all established diffractive orders is investigated. For this reason, we present the obtained data from fitted function Eq. 3.9 for width parameter at $\frac{1}{e^2}$ and for all diffractive orders on the camera. w_1 and w_2 from Eq. 3.9 are the width of every diffractive order in x and y plane, respectively. Figure 3.15 shows the w_2 for all diffractive orders in red circles. The predicted analytical width is shown in 3.15 as the black dashed line. From our setup we see for emitters $y < 1$ mm that the change of width is smooth and achieves $0.8\ \mu$m difference from black dash line (effect of noise and fit algorithm). While, for emitters $y > 1$ mm the width size is dramatically increasing. Indeed, for emitter's displacement from $y_0 = 1$ mm to $y_{10} = 1.14$ mm, width in $w_2 = 3.4\ \mu$m arrives at $w_2 = 5.2\ \mu$m which is linked to the beginning of vignetting. Such a vignetting problem reduces the coupling efficiency in our setup hence, we will suggest a solution to remove this impact in next section.

3.2/ OPTIONS OF SOLVING THE RESTRICTION OF BEAM VIGNETTING

As we explained in last sections, due to a limited diameter of microscope objectives, vignetting effect can occur. To address this problem, we selected several kinds of MOs with different NAs. For instance, we exchange MO_2 with another MO (Olympus RMS20x with $\mathrm{NA}_2 = 0.5$). But since the minimum distance of mounting the optical elements from MO_1 to

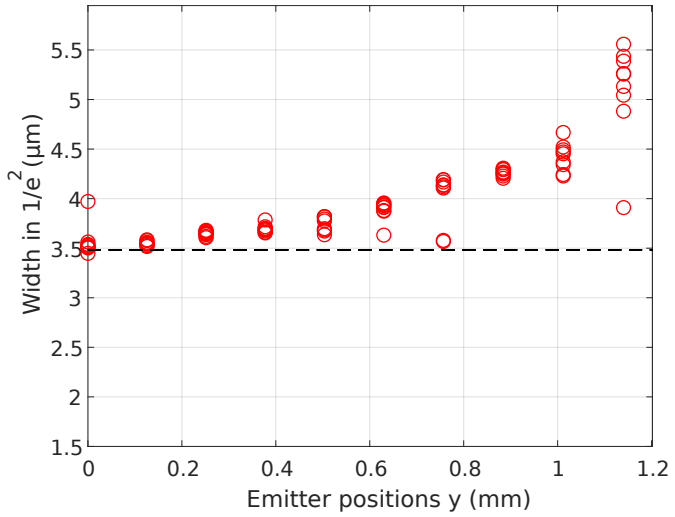

FIGURE 3.15 – Width for 90 diffractive orders in experiment (red circles) relative to the diffraction limit (black dashed line) versus emitter positions y. Vignetting leads to a strong increase of the width for $y > 1$ mm.

MO_2 is at least 55 mm, vignetting impact will arrive quickly for non-centered beams. We tested achromatic doublets such as the Thorlabs AC254-030-B-ML instead of MO_1 and Thorlabs AC254-060-B-ML or Thorlabs AC254-080-B-ML instead of MO_2. The problem of achromatic doublet was that they exhibit strong spherical aberrations, which tend to the deformation of a flat object. Figure 3.16 illustrated the aberration effect on our beam for achromatic doublet Thorlabs AC254-150-B-ML in the replacement of MO_2 for placing the single mode fiber at $y = 1.57$ mm.

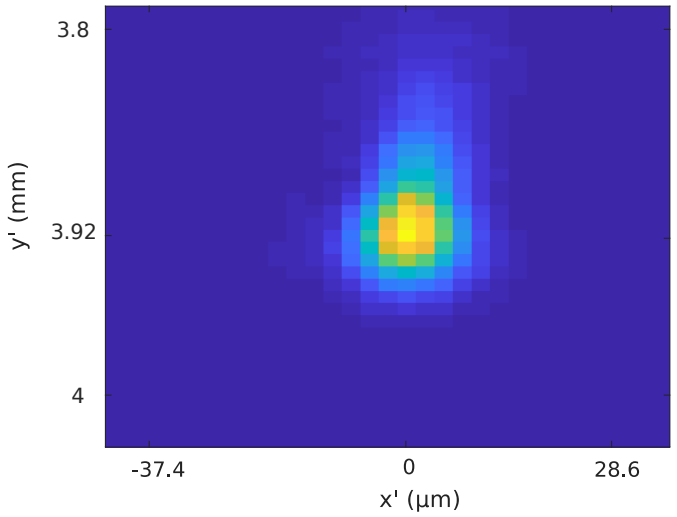

FIGURE 3.16 – Effect of aberrations on a beam after replacing MO_2 with an achromatic doublet.

In another test, we replace MO_2 with a tube lens for infinity corrected microscope objective (Nikon, 58-520) which typically provide homogeneous, flat images with negligible

aberration. The focal length of this tube lens is $f = 200$ mm. The entrance pupil diameter in tube lens is 32 mm, which is enough large for our purpose. The advantage of these tubes lenses is that their larger clear aperture's diameters allows more space between objective and tube lens. In this tube lens as long as the distance between the tube lens and objective is less than 232.7 mm, there will be no vignetting. Figure 3.17 shows the diffractive orders from an emitter on the camera. Aberration from the tube lens changed the shape of diffractive orders. Since we need the accurate results in our system, this solution can not be accepted. For all mentioned reasons, we seek another method to isolate the emulation of diffactive effects from imaging aberrations.

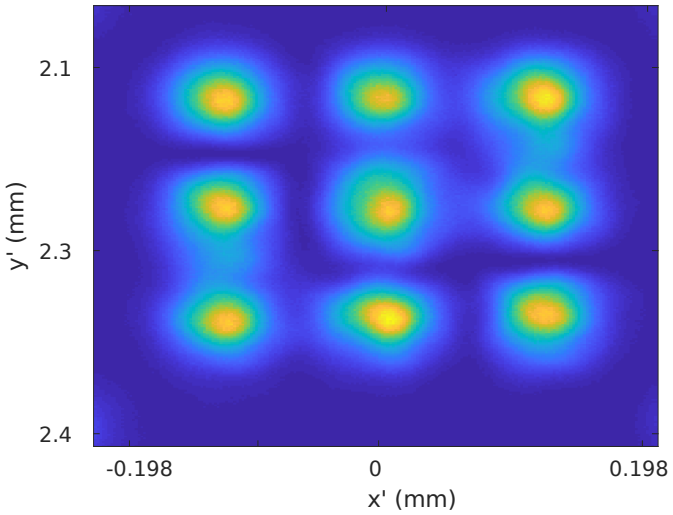

FIGURE 3.17 – Diffractive orders obtained by using tube lens. The effect of aberration changes the shape of diffractive orders.

3.3/ EMULATE EMITTER POSITIONS THROUGH DOE-TILTING

3.3.1/ EXPERIMENTAL SETUP

In this section, instead of translating the fibre, we tilt the DOE, as demonstrated in Fig. 3.18. This approach continuously keeps the collimated beam at the center position of the setup and camera. Shifting in $x - y$ results in an angle of the collimated beam, and its effect on diffractive coupling, we can emulate by keeping the collimated beam the same by tilting the grating. We use the following equation to explain the position of emitter on $(x - y)$-stage by tilting the DOE

$$y = f_1 \cdot \tan(\theta^{\mathrm{i}}). \tag{3.10}$$

In Eq. 3.10 not that the beam stays center, but that tilting θ^{i} emulates positions y, and after that for such an approach the beam remains centered. In this experiment, the position of optical fibre is always at the center of the $(x - y)$-stage. In this setup we modified MO_1, which now has $f_1 = 20$ mm (MO_1: Nikon N10X-PF) while all other optical elements are identical to the first experiment shown in Fig. 3.10.

For each new recording in this setup, we tilt the DOE from $\theta^{\mathrm{i}} = 0$ to $\theta^{\mathrm{i}} = 22$ degrees, which is equivalent to emulating emitter positions ranging from $y_0 = 0$ mm to $y_{lim} = 8$ mm. The

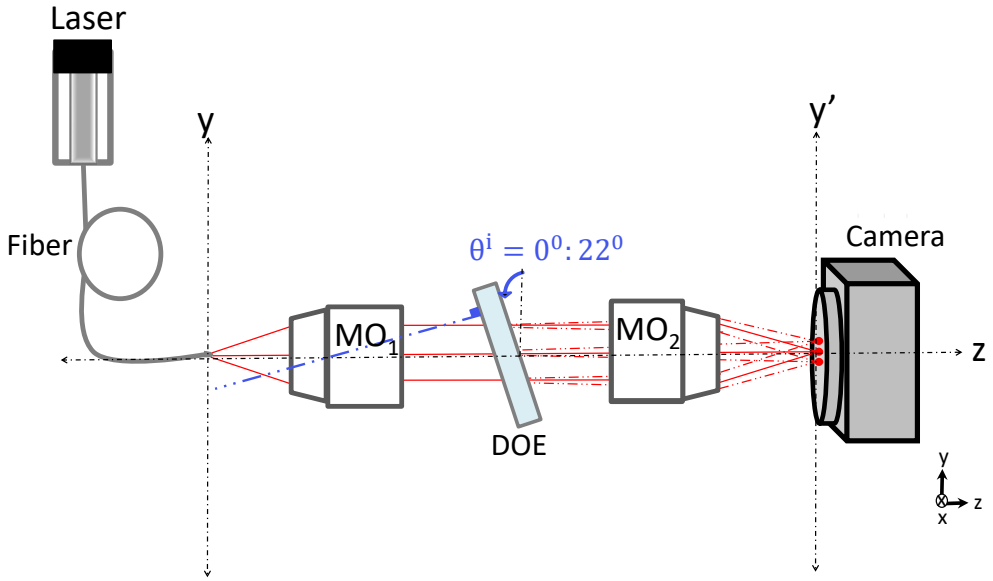

FIGURE 3.18 – Experimental setup to test limits of diffractive coupling by tilting the DOE angle from 0 degree to 22 degrees.

optical coupling limit obtained from the analytical model was found to be $y_{lim} = 8$ mm. For this reason, after 22 degree, we stop recording the data on the camera. In next section we will present experimentally 3 measurements (distance between diffractive orders, mismatch $|\Delta|$ and width of diffractive orders) for Fig. 3.18 [50].

3.3.1.1/ COUPLING MISMATCH $|\Delta|$

To show the result of tilting the DOE, we start with distance between diffractive orders as discussed in Fig. 3.12. Different emitter positions in the range of 0 mm $< y <$ 8 mm are emulated via tilting the DOE. If we got 20 source points in y direction, number of distances in red arrows $|d^{\text{diff}}|$ from Fig. 3.12 for this setup will be $(20 \times 6 = 120)$. The physical properties of the diffractive orders were characterized based on the the same numerical fit with a Gaussian as in the previous sections. Distances between diffractive orders $|d^{\text{diff}}|$ for different emitter positions y are shown in Fig. 3.19. The black dashed line is the constant distance for which in a periodic array coupling is optimal. In Fig. 3.19 for positions from $y_0 = 0$ mm to $y_{lim} = 8$ mm, distance is increasing from $|d_0^{\text{diff}}| \sim 12.21$ μm until $|d_{end}^{\text{diff}}| \sim 14$ μm.
The difference between red rectangles and the black dashed line in $y_{lim} = 8$ mm arrives 1.8 μm which is the maximum disagreement between experimental and the alignment condition. This difference shows that the coupling effect reduces smoothly by moving emitters far from center. As a result, we have coupling in $y_{lim} = 8$ mm which is less strong than the coupling in $y_0 = 0$ mm.

In another measurement, Fig. 3.20 illustrates mismatch $|\Delta|$ for different emulated emitter positions in y in a double-logarithmic scale. In this figure, red stars show experimental data and analytical data is the black line. We can divide this plot into two sections with characteristic behaviour. The first part is related to positions located in $y < 0.5$ mm. As before the mismatch $|\Delta|$ remains below 30 nm, which as previously discussed is below the

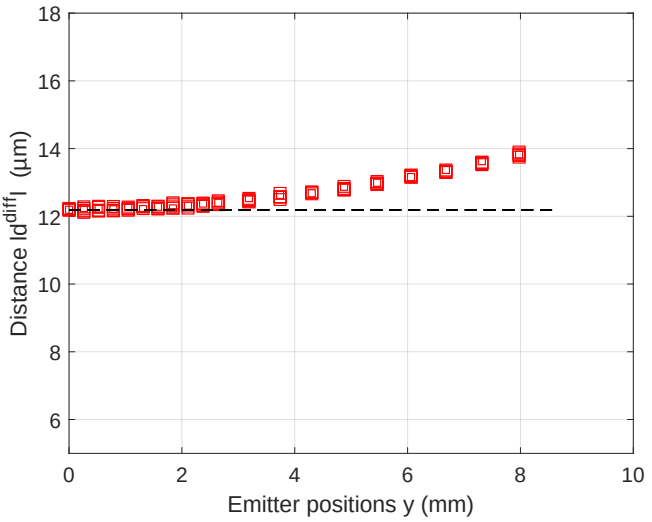

FIGURE 3.19 – Red rectangles: 120 distances between $|d^{\text{diff}}|$ diffractive orders versus emitter positions y from $y_0 = 0$ mm to $y_{lim} = 8$ mm. Black dashed line: distance between diffractive orders in analytical solution.

data acquisition and analysis accuracy of our system.
The second part corresponds to the range of 0.5 mm $< y <$ 8 mm. In this range, the red stars overlap perfectly with the black line. This excellent result demonstrates the high potential of diffractive coupling if vignetting is avoided. To conclude, the agreement of the distance mismatch for tilt angles ranging from 0 to 22 degree with analytical data confirm the validity of diffractive coupling for arrays with emitters inside an area of 8 mm.

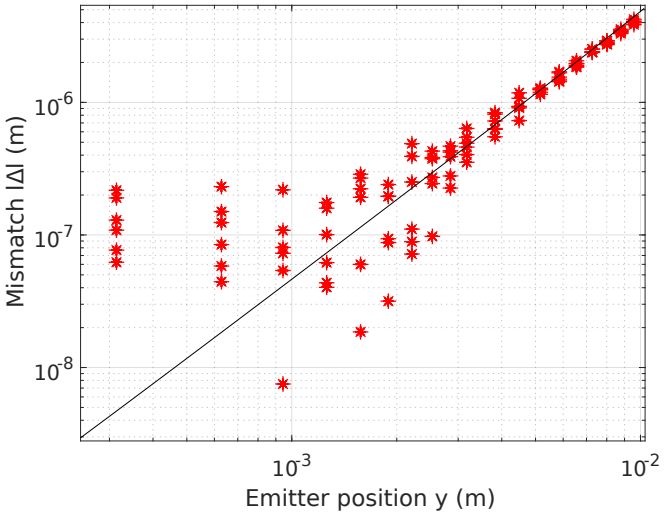

FIGURE 3.20 – The diffractive coupling mismatch $|\Delta|$ obtained by the experiment (red stars) and the analytical limit from Eq. (1.10) (black line) excellently agree.

As we explained before, the diameter of the focused beam is one of the significant factor to establish high-quality optical coupling. For this reason in another measurement based on Eq. 3.9, w_2 for all diffractive orders is obtained. The width of the 180 diffractive orders at $\frac{1}{e^2}$ for emitter positions from $y_0 = 0$ mm to $y_{lim} = 8$ mm are displayed in Fig. 3.21 and in red

circles ; the diffraction limit is shown as the black dashed line to which the experimental data excellently agrees. Small mismatches between the red circles and black dashed line are again due to the noise of the system and the limited precision of the fit functions. For the distance of 8 mm and where distance between two diffractive orders is 12.2 μm, we can maximally couple ($\frac{8 \cdot 10^{-3}}{12.2 \cdot 10^{-6}} = 655$) emitters or photonic nodes.

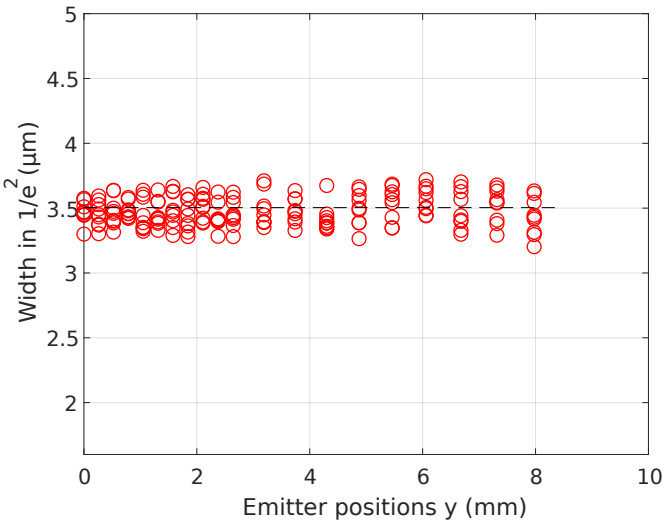

FIGURE 3.21 – Distribution of width for 180 diffractive orders in 1/e^2 from $y_0 = 0$ mm to $y_{lim} = 8$ mm in red circles. The diffraction limit is in black dashed line.

3.3.2/ LIMITATION OF THE SETUP

As we saw in the last section, by keeping emitter constant in the center, we achieved the successful coupling from $\theta^{\text{i}} = 0$ degree to $\theta^{\text{i}} = 22$ degree without vignetting effect. Another great advantage was the roughly constant size of the width for all diffractive orders. But, since emitter did not move, physical effects of system such as aberration could not be recognized on the camera. Therefore, this system is not sensitive to off-axis beam induced imperfections of imaging.

3.4/ CONCLUSION

In this Chapter, based on two experiments we investigated the practical limits to the size of photonic networks coupled by diffraction. In the first experiment, a single-mode fiber's output was translated by a $(x-y)$-stage in order to imitate an array of single mode emitters. The plotted results illustrated that this setup is sensitive to aberration. We identified the following effects limiting the size of an emitter array which can be potentially coupled by diffraction:

- For off-axis emitters, beam-vignetting was observed in mismatch distance $|\Delta|$ because of a too low NA of the MOs.
- Strong vignetting was also found in focus spot's width.
- Replacing MOs with achromatic doublet or tube lens deteriorate imaged beam because of optical aberrations.

To characterize maximum size of optical coupling in a setup without beam vignetting, we created a second characterization methodology. We used the same optical elements for new setup except MO_1. The single-mode fiber's output was located on the optical axis for measurements, and different emitter positions in an array were emulated by tilting the DOE from $\theta^{\mathrm{i}} = 0$ to $\theta^{\mathrm{i}} = 22$ degrees, which corresponds to emitter located between $y_0 = 0$ mm to $y_{lim} = 8$ mm. By applying this technique, the vignetting effect was avoided. However, this system is not sensitive to aberration effects since the location of all emitters is constant. According to experimental and analytical data, diffractive coupling can be obtained for emitters located in a space of 8 mm. Mismatch $|\Delta|$ was less than 4 μm which is a typical size of an optical nod. This achievement determined successful coupling by second experiment and the high quality of system to create the diffractive orders. In another measurement, we showed the quality of system, we plotted the width of all diffractive orders at $\frac{1}{e^2}$. Crucially, the focal spot's width is constant for all recorded diffractive orders inside the 8 mm and close to the diffraction limit. The emitter positions y at all figures in Chapter 3 were selected as an example to confirm the size of photonic networks coupled. After this confirmation, hence inside an area of $(16 \text{ mm})^2$ we can coupled $1,719,000$ photonic nodes.

4

NUMERICAL MODEL

In the previous Chapter we investigated experimentally the limitations of diffractive coupling. During our measurements, we observed numerous effects introducing limitations. For instance a strong divergence of mismatch distance $|\Delta|$ and an increase of the focal spot waist radius induced beam vignetting. These effects cannot be accurately described with the simple analytical treatment introduced in Chapter 2. We require a treatment considering effects of coherent light propagation through the system for any emitter position. For this reason, in this Chapter we will model the experimental setup and numerically reproduce the beam propagation, collimation, focusing and diffraction. During our simulation and calculations different techniques are applied. These techniques are indicated in Fig. 4.1 which include propagation, the Debye integral for emulating high NA MOs, absorbing boundaries, masking at the MOs' clear aperture and rescaling the coordinate systems in the object and image plane [50].

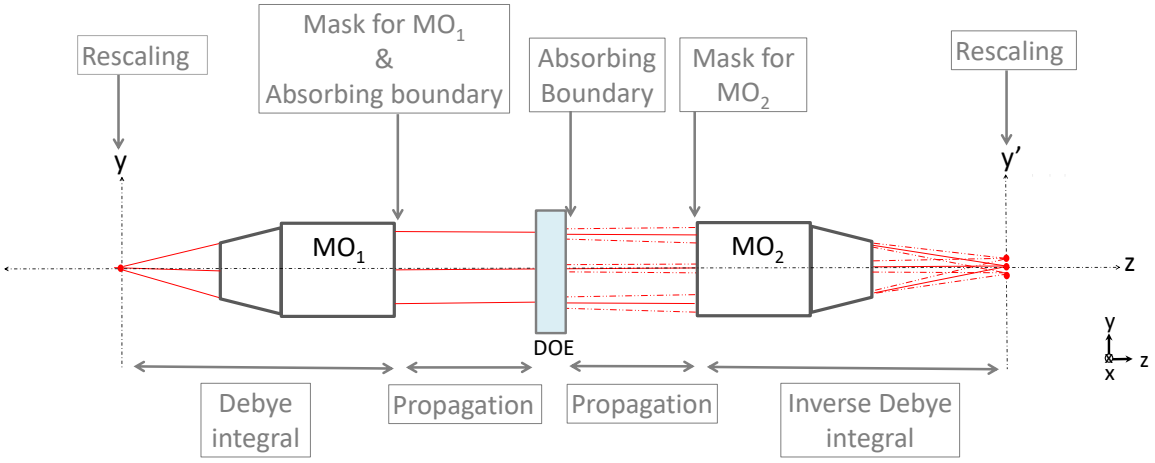

FIGURE 4.1 – The scheme of numerical simulation. The mentioned explanations are the techniques which are used during our simulation.

Within this Chapter, we will first explain in detail the techniques indicated in Fig 4.1. Second, we adjust the numerical model that we use parameters from the experiment. Third, we suggest a solution to overcome the limitations induced by the vignetting effects.

4.1/ BEAM PROPAGATION

In order to propagate an optical wave through an optical setup, we introduced several methods in the second Chapter, among others the Helmholtz equation (Eq. 2.17) which represents a time-independent form of the propagation equation. Then we derived a simple solution for this equation, which uses plane waves to propagate the complex optical fields. In the numerical simulation we utilize the plane waves spectrum which was studied in details in Chapter 2.

4.1.1/ ABSORBING BOUNDARY

After applying the three steps of wave propagation in our numerical simulation, as explained in Chapter 2, we see that the FFT algorithm induces a periodicity in our calculation window's boundary condition. Due to this periodicity, a beam which is going out on one side enters into the window by the other side, as illustrated in Fig. 4.2, which shows the propagation of a Gaussian beam with an angle [114–116]. This propagation is along z axis and an angle of the Gaussian beam of $\theta = 53$ degree. As a result from Fig. 4.2, the beam re-entered on the other side of the propagation. To avoid this numerical artifact, we implement the absorbing boundary. In this method, we define a Hyper Gaussian in Eq. 4.1, inside which the signal is maintained, while outside of the propagating beam will be strongly attenuated and absorbed. In Eq. 4.2 the absorbing boundary is defined for two approaches first the hyper Gaussian and second the rectangular absorbing boundary

$$Abs(x,y) = exp\left(\frac{x-x_0}{w_1}\right)^m exp\left(\frac{y-y_0}{w_2}\right)^m, \tag{4.1}$$

when

$$Abs(x,y) = \begin{cases} \text{Hyper Gaussian,} & \text{if } 2 < m < 10 \\ \text{Rectangular,} & \text{if } 10 < m \end{cases} \tag{4.2}$$

Absorbing boundary approaches are modeled in Fig. 4.3 for -5 mm $< x < +5$ mm. Figure 4.3(a) illustrates an hyper Gaussian absorbing boundary where $m = 6$ in Eq. 4.1 while, Fig. 4.3(b) is a rectangular absorbing boundary for $m = 60$. In this thesis, we use hyper Gaussian function since the absorption of the boundary occurs smoothly than the simple rectangular function. Therefore, the final equation of wave propagation including the absorbing boundary will be

$$E(x,y,z) = Abs(x,y)\ F^{-1}\left\{E(u,v)e^{ikz\sqrt{1-\lambda^2u^2-\lambda^2v^2}}\right\}. \tag{4.3}$$

The result of wave propagation after implying the Eq. 4.3 in the simulation is displayed in Fig. 4.4 for hyper Gaussian absorbing boundary and rectangular absorbing boundary. This method removes the problem of periodicity in our numerical simulation. In Fig. 4.4(a)

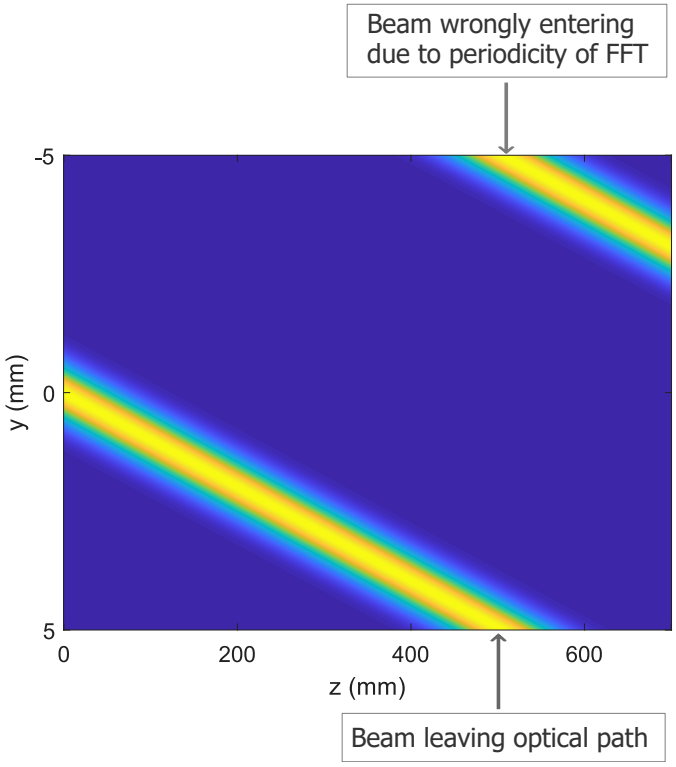

FIGURE 4.2 – Observation of the periodic boundaries introduced by the FFT during beam propagation. A Gaussian beam in angle $\theta = 53$ degree is propagated along the z direction. It leaves the simulated window from one side (bottom) and re-enters from the another side (top of window). This is a numerical and un-physical artefact.

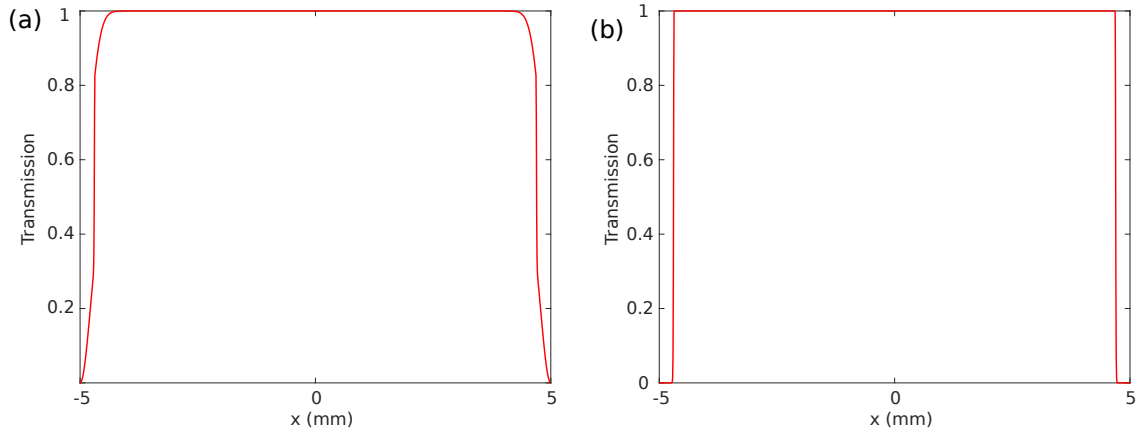

FIGURE 4.3 – Proposed models to remove the boundary effect during beam propagation. (a) The hyper Gaussian absorbing boundary. (b) The rectangular absorbing boundary.

the beam propagates along z and the edge of window is absorbed smoothly by a hyper Gaussian absorbing boundary and Fig. 4.4(b) demonstrates a zoom to see better the form of hyper Gaussian absorbing boundary at the edge. While, in Fig. 4.4(c) the beam propagates along z and is absorbed at the edge of the window in form of a rectangle and a zoom into the edge of window for this state is shown in Fig. 4.4(d). Therefore, we apply this technique (hyper Gaussian absorbing boundary) during propagation of the field in the

simulation to solve the periodic boundaries introduced by the FFT.

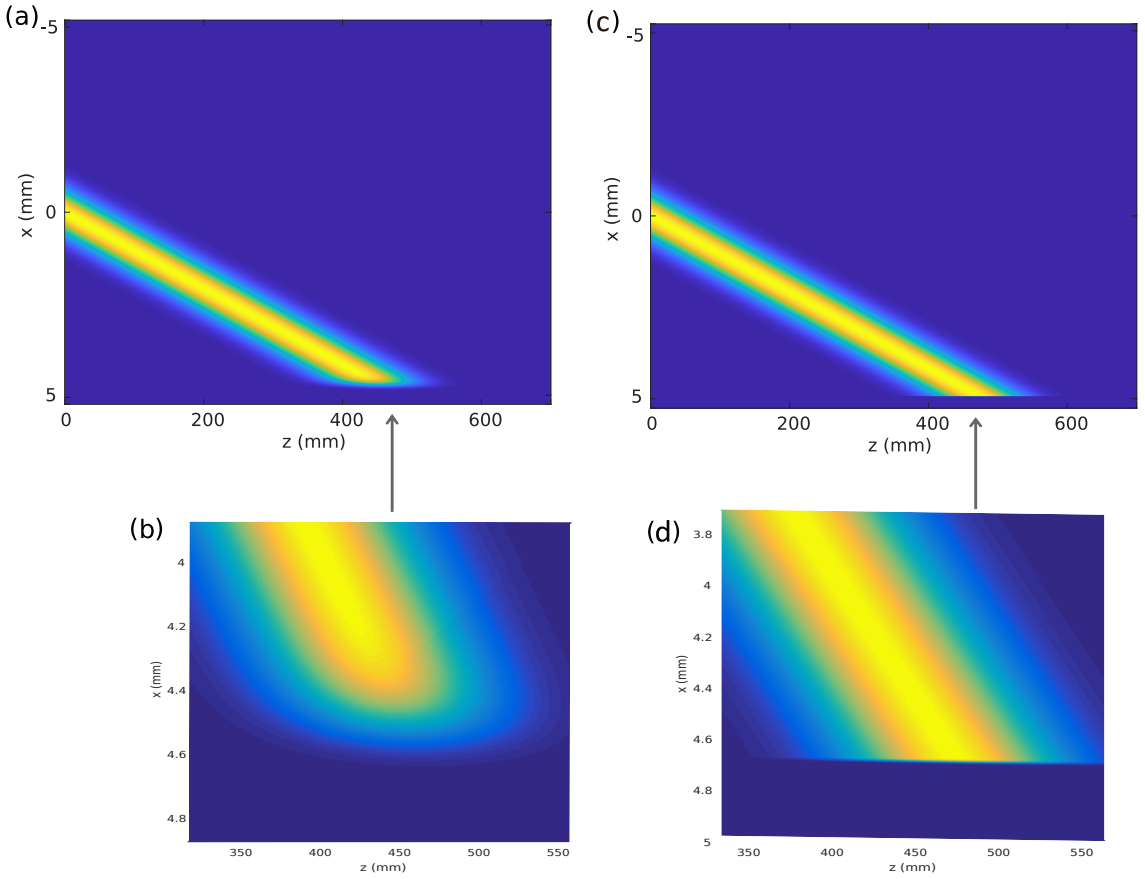

FIGURE 4.4 – Absorbing boundary solution for a Gaussian beam in angle $\theta = 53$ degree is propagated along z direction in Fig. 4.2. (a) The hyper Gaussian to absorb the beam smoothly at the edge of window. (b) The inset showing a zoom into the edge of window in hyper Gaussian absorbing boundary. (c) The rectangular absorbing boundary to absorb the beam rectangulary at the edge of window. (d) The inset showing a zoom into the edge of window in rectangular absorbing boundary.

4.2/ MICROSCOPE OBJECTIVES

Typical microscopes contain many lenses [117–119]. There are two parameters commonly used to characterize a microscope objective are their magnification such as 4X, 10X or 20X and numerical aperture (NA). For an entrance pupil radius of microscope objective R and its focal length f where n is index of refraction of the immersion medium we can write

$$\mathrm{NA} = n \sin\left(\tan^{-1}\left(\frac{R}{f}\right)\right), \tag{4.4}$$

by replacing $\left(\tan^{-1}\left(\frac{R}{f}\right)\right)$ with θ in Eq. 4.4 we obtain

$$\mathrm{NA} = n \sin\theta. \tag{4.5}$$

In Eq. 4.4 if the pupil radius is larger and f is constant, NA will be higher. We can describe a thin lens in the (x, y)-plane in Eq. 4.6 as

$$t(r) = a_0 \exp\left(-i\frac{\pi}{\lambda f}(r^2)\right), \tag{4.6}$$

where $t(r)$ is phase transmission function, f is the focal length of the lens, and $r^2 = x^2 + y^2$. In Eq. 4.6, $a_0 = \exp(-ik_0 r_0)$ is a constant phase factor and the term $\exp\left(-i\frac{\pi}{\lambda f}(r^2)\right)$ changes with the square of the distance r from the optical axis. For point sources arranged with such an radially changing phase, all fields will constructively interfere at distance f. According to the Huygens principle, a lens therefore has to implement a radial phase retardation in order to focus. There is a direct method which calculates the electromagnetic field at focal point of a lens. However, such a direct solution of the propagating plane-wave spectrum exhibits a challenge because in $(x - y)$-plane, we need a potentially extremely large number of samples to simulate a lens (a curve shape). To obtain the number of samples for a lens, we set an example: an optical beam with $\lambda = 660$ nm propagates through a lens with $f = 20$ mm creating N phase-slips

$$N \cdot \lambda = f, \tag{4.7}$$

here $N = 3 \times 10^4$ for this lens. Considering at least 5 samples per 2π, we therefore require more than $N = 1.5 \times 10^5$ data points. Hence, such a direct method is well suited to calculate the focus field for a lens with low NA [120]. In order to find the limitation of coupling in our setup, we are very sensitive to aberrations and therefore need a dense sampling of the curved wavefront. In that case since the memory of computer is limited, we are not able to realize a sufficient number of samples. Moreover, the duration of calculation for the highest possible number of samples would be very long. For this reason, we attempt to find another solution to propagate plane wave in high NA microscope.
In experiments, the aperture stop of the MO defines the clear aperture of MOs. In numerical simulation, we limit the transmission of a MO to this area, inside the clear aperture's diameter a beam is allowed to pass while outside of this area beam can not pass. We determined a mask function f_{m} in MOs for position $r^2 = x^2 + y^2$ according to

$$f_{\mathrm{m}} = \begin{cases} 1, & \text{if } r < R \\ 0, & \text{otherwise} \end{cases} \tag{4.8}$$

Function f_m is multiplied onto the field at the MOs aperture stop.

4.2.1/ DEBYE INTEGRAL

Instead of numerically implementing the direct propagation through the MOs we will employ the Debye integral. The Debye integral is a method to calculate the electromagnetic field at the focal plane as well as the field of the collimated beam. In this method, one projects the incident beam into cylindrical coordinates then spherical coordinates to obtain the image's focal point. By applying the Debye integral, we can simulate the microscope objective with high NA and without an excessive number of samples [120–127]. Figure 4.5 demonstrates a basic scheme explaining individual steps of the Debye integral method [120]. Behind the aperture stop A with radius R, the monochromatic light is parallel

to the optical axis. The optical field propagates toward the principal plane $\mathbb{P}_1$. In principal plane $\mathbb{P}_1$

$$x = r\cos(\phi), \tag{4.9}$$

and

$$y = r\sin(\phi), \tag{4.10}$$

where r is the off-axis coordinate of incident beam $\left(r = \sqrt{x^2 + y^2}\right)$ and ϕ is the azimuth angle around the z-axis

$$\phi = \arctan\left(\frac{y}{x}\right). \tag{4.11}$$

In Fig 4.5, the focal length f can be defined as

$$f = \overline{\mathrm{F_1V_1}}, \tag{4.12}$$

where F_1 is focal point and the principal plane $\mathbb{P}_1$ has vertex point V_1. Principal plane $\mathbb{P}_2$ is curved according to the thin lens formula, hence after passing trough the plane $\mathbb{P}_2$, the field focuses to focal point F_2. The principal planes $\mathbb{P}_2$ has the vertex point V_2. [120]. Calculating imaging by a microscope objective based on the Debye integral requires:

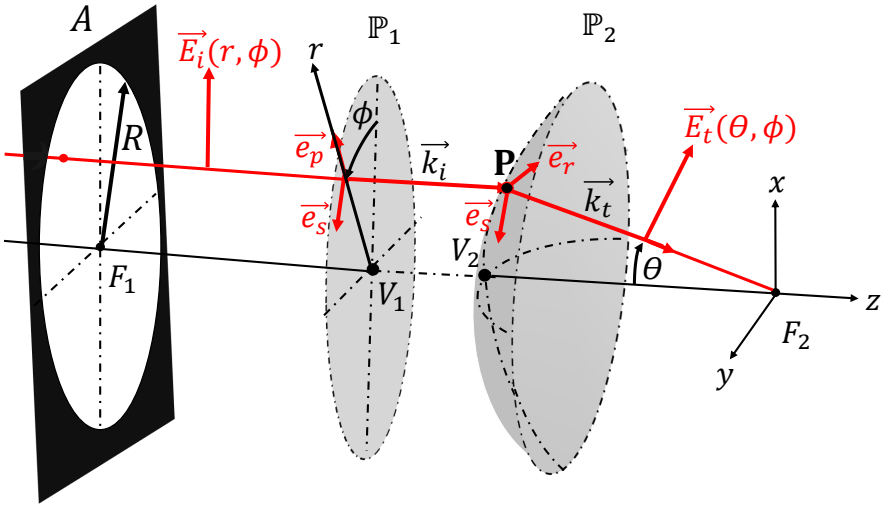

FIGURE 4.5 – Principal of the Debye integral. Optical field components are projected onto the microscope objective principal plane $\mathbb{P}_1$ then $\mathbb{P}_2$. The focal length f is calculated as $f = \overline{\mathrm{F_1V_1}}$. The point **P** lies on the plane $\mathbb{P}_2$ and shows the focusing of a ray at the focal point F_2, reproduced from [120].

First, projecting the incident field $\vec{E}_{\mathrm{i}}(x,y)$ onto cylindrical coordinates. These coordinates create $\vec{E}_{\mathrm{i}}(r,\phi)$ at plane $\mathbb{P}_1$, $\vec{E}_{\mathrm{i}}(r,\phi)$ is decomposed into the optical field's radial component (p-polarized) and tangential component (s-polarized). The unit vectors for p-polarization is

$$\vec{e}_p = \begin{bmatrix} \cos\phi \\ \sin\phi \\ 0 \end{bmatrix}, \tag{4.13}$$

and in s-polarization is

$$\vec{e}_s = \begin{bmatrix} -\sin\phi \\ \cos\phi \\ 0 \end{bmatrix}. \tag{4.14}$$

Second, the transformed field is projected onto the microscope objective's principle plane $\mathbb{P}_2$ in spherical coordinates. In the plane $\mathbb{P}_2$, $\vec{e}_{\mathrm{p}}$ is deflected by angle θ and creates

$$\vec{e}_r = \begin{bmatrix} \cos\phi\cos\theta \\ \sin\phi\cos\theta \\ \sin\theta \end{bmatrix}, \tag{4.15}$$

where n_t is refractive index behind the $\mathbb{P}_2$. In the point of **P** we use

$$\sin\theta = \frac{r}{R}\frac{NA}{n_t}. \tag{4.16}$$

As a result, amplitude, phase and polarization of the transmitted field in plane $\mathbb{P}_2$ are given by

$$\vec{E}_t(\theta,\phi) = t_p\left(\vec{E}_i\cdot\vec{e}_p\right)\vec{e}_r + t_s\left(\vec{E}_i\cdot\vec{e}_s\right)\vec{e}_s, \tag{4.17}$$

where t_{p} is the transmission coefficients for p-polarization and t_{s} is the transmission coefficients for s-polarization. In our calculations we use $t_{\mathrm{p}} = 1$ and $t_{\mathrm{s}} = 1$.

Third, after these transformations, the electric field $\vec{E}$ at the point (x, y, f) near the focus is calculated by integrating the propagation of plane waves along z axis

$$\vec{E}(x,y,z) = \frac{-if}{\lambda_0}\iint_{\Omega}\vec{E}_t(\theta,\phi)e^{i(k_z z - k_x x - k_y y)}d\Omega. \tag{4.18}$$

The phase term $e^{\mathrm{ik_z f}}$ is a phase accumulation and $e^{-\mathrm{i(k_x x + k_y y)}}$ determines the phase difference of the beam from the on-axis point $(0, 0, f)$ to the off-axis point (x, y, f). By introducing the two limits of $0 < \phi < 2\pi$ and $0 < \theta < \Theta$ in Eq. 4.18 we rewrite

$$\vec{E}(x,y,z) = \frac{-if}{\lambda_0}\int_0^{\Theta}\sin(\theta)\int_0^{2\pi}\vec{E}_t(\theta,\phi)e^{i(k_z z - k_x x - k_y y)}d\theta d\phi, \tag{4.19}$$

where in focal point F_2, we observed $(\sin\Theta = NA)$. We can rewrite Eq. 4.19 in the form of a Fourier transform

$$\vec{E}(x,y,z) = \frac{-if}{\lambda_0 k_t^2}\iint_{r<R}\left(\vec{E}_t(\theta,\phi)e^{ik_z z}/\cos\theta\right)e^{-i(k_x x + k_y y)}dk_x dk_y, \tag{4.20}$$

which leads to

$$\vec{E}(x,y,z) = \frac{-if}{\lambda_0 k_t^2}FFT\left\{\frac{\vec{E}_t(\theta,\phi)e^{i(k_z z)}}{\cos\theta}\right\}, \tag{4.21}$$

where in spherical coordinates θ, ϕ and in focal point F_2, we have $k_{\mathrm{t}} = \mathrm{k}_0$.

4.2.1.1/ NUMERICAL IMPLEMENTATION

Based on the final calculation from the Debye diffraction integral as an inverse Fourier transform, we can now develop a numerical procedure to compute the field of an imaged or focused wavefront by a microscope objective with high NA. For numerical simulation of the Fourier transform, our algorithms operate in discrete which requires the data points

as the samples. We defined S as the sampling points over the aperture radius. By using $k_x = m\Delta k$ and $k_y = n\Delta k$ where $\Delta k = \frac{k_0 NA}{S}$, we obtain sampling positions in plane $\mathbb{P}_2$ at

$$\theta_{mn} = \arcsin\left(\frac{\Delta k}{k_t}\sqrt{x^2 + y^2}\right), \tag{4.22}$$

and

$$\phi_{mn} = \arctan\left(\frac{n}{m}\right), \tag{4.23}$$

for $\mid m \mid, \mid n \mid \leq$ S. As a result in the numerical implementation, we express the Debye integral for a microscope objective by

$$\vec{E}(x_k, y_k, f) = \frac{-iR^2}{\lambda_0 f S^2} FFT\left\{\frac{\vec{E}_t(\theta_{mn}, \phi_{mn}) e^{ik_{zmn}z}}{\cos\theta}\right\}. \tag{4.24}$$

Equation 4.24 computes the optical field in the image plane after a microscope objective. In order to obtain the collimated beam, we employ the inverse Fourier transform of incident beam and hence, complete the propagation between the microscope objectives $(x, y, z = 55\ mm)$ and its image plane $(x, y, z = f)$

$$\vec{E}(x_k, y_k, f) = \frac{-iR^2}{\lambda_0 f S^2} F^{-1}\left\{\frac{\vec{E}_t(\theta_{mn}, \phi_{mn}) e^{ik_{zmn}z}}{\cos\theta}\right\}, \tag{4.25}$$

the obtained $\vec{E}(x_k, y_k, f)$ from Eq. 4.25 is replaced $\vec{E}_i$ in Eq. 4.17

$$\vec{E}_t(\theta, \phi) = t_p\left(\vec{E}(x_k, y_k, f) \cdot \vec{e}_p\right)\vec{e}_r + t_s\left(\vec{E}(x_k, y_k, f) \cdot \vec{e}_s\right)\vec{e}_s, \tag{4.26}$$

by replacing $\vec{E}_t(\theta_{mn}, \phi_{mn})$ inside of Eq. 4.17, we can achieve the collimated beam.

4.3/ RE-SCALING OF IMAGE PLANE

During the second step of plane-wave spectrum propagation (2.2.1), we explained the non-paraxial effects in beam propagation. Because of high sensitivity to optical inaccuracies in our setup, it was required to avoid any approximation. In this section, we show how the usage of the Debye integral concept unfortunately induces an approximation between frequencies of the plane-wave spectrum and position space. Finally, we re-scale numerically the image plane in order to compensate for this effect.
For this goal, we illustrate the relationship between spatial positions (x, y) and spatial frequencies (u, v) in paraxial and non-paraxial scale. Figure 4.6 is an example where the microscope objective is located in spatial frequency space with coordinate v. The collimated beam in the frequency plane will be focused on the x plane in focal length f.

4.3.1/ COORDINATES UNDER PARAXIAL APPROXIMATIONS

In paraxial equations, one typically approximates $\sin\theta \simeq \tan\theta \simeq \theta$. From Fig 4.6, we obtain

$$\tan\theta = \frac{x_0}{f}, \tag{4.27}$$

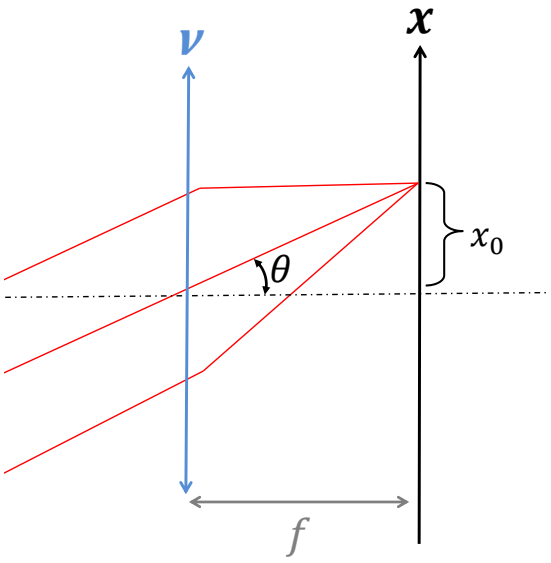

FIGURE 4.6 – An imaging system in our optical setup. Image plane coordinate is x and frequency plane coordinate is v. Beam is imaged in focal length f.

using the link between spatial frequency and angle θ based on the paraxial approximation, introduced in the propagation section Eq. 2.52 and Eq. 2.57, leads to

$$\theta = \lambda v, \tag{4.28}$$

the left parts of Eq. 4.27 and Eq. 4.28 are equal leading to

$$x_0 = \lambda v f. \tag{4.29}$$

Equation 4.29 provides the relationship between the image coordinate plane and spatial frequencies.

4.3.2/ NON-PARAXIAL COORDINATES

For the general context not relying on approximations one has to assume $\sin\theta \neq \tan\theta$. Combined with Eq. 4.27 one obtains

$$\frac{\sin\theta}{\cos\theta} = \frac{x_0}{f}, \tag{4.30}$$

Combined with the angle versus spatial frequency relationship given by Eq. 2.52 and Eq. 2.61

$$\sin\theta = \lambda v, \tag{4.31}$$

we obtain

$$x_0 = \frac{f\sin\theta}{\cos\theta} = \frac{\lambda v f}{\cos\theta}, \tag{4.32}$$

this can further be simplified by using $\cos\theta = \sqrt{1-\sin^2\theta} = \sqrt{1-\lambda^2 v^2}$

$$x_0 = \frac{\lambda v f}{\sqrt{1-\lambda^2 v^2}}. \tag{4.33}$$

The numerical coordinate system is connected via Eq. 4.33. By comparing Eq. 4.29 and Eq. 4.33, we find that the non-paraxial coordinate plane can be obtained by a rescaling through dividing paraxial spatial frequency coordinates with the term $\sqrt{1-\lambda^2 v^2}$. This term is non-linear and hence its effects in our system will be most dominant at positions far away from the center.
Figure 4.7 is a part of an array of Gaussian beams (0 mm $< y <$ 4.5 mm) for different emitter positions imaged according to the setup in Fig. 4.1 located in -4.5 mm $< y <$ 4.5 mm. Gaussian emitters are compared in paraxial and non-paraxial coordinate systems in red and blue colors. Red emitters illustrate in paraxial coordinate planes and blue Gaussian emitters in non-paraxial coordinate planes. For emitters where $y < 2$ mm, difference between beams in red and blue beams is small, but for $y > 2$ mm is significant. For instance at $y_1 = 0.16$ mm the difference between those two planes is 0.3 μm, while at $y_{end} = 4.5$ mm is 0.151 mm, which certainly can not be neglected. The non-paraxial re-scaling is therefore applied to all our numerical simulation.

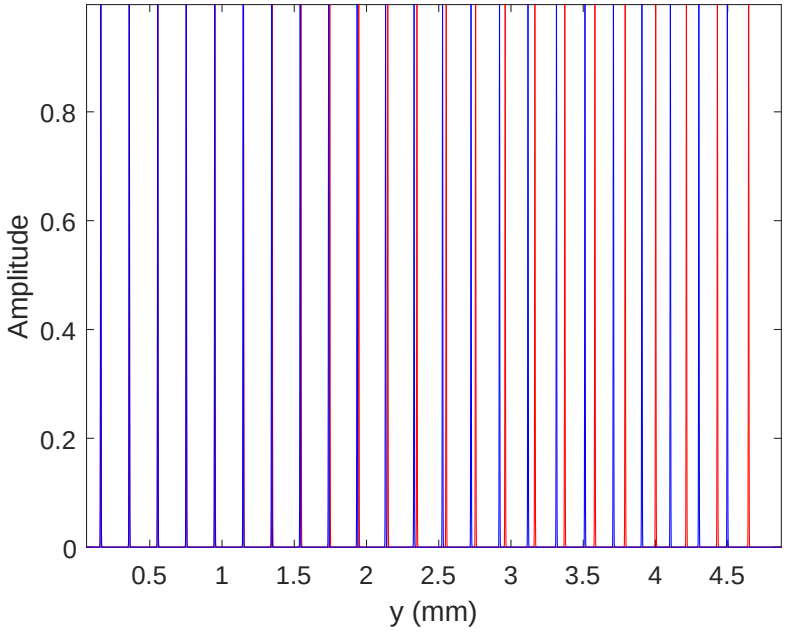

FIGURE 4.7 – Results of an array of Gaussian emitters in paraxial and non-paraxial coordinate systems. Blue color are non-paraxial and red color are paraxial simulations. Difference between both coordinate systems are considerable at $y > 2$ mm.

4.4/ THE DOE

The diffractive coupling employed in this work is based on a DOE, which is a phase only object. Our aim is to model the experiment as close as possible to reality. In order to do so, the phase of the DOE that we use in our simulation should corresponds to the phase of the DOE element. The procedure to measure this phase uses an adaptation of the Gerschberg Saxton (GS) phase retrieval algorithm [128, 129]. The principle of this method relies on the fact that there is only one phase possible for two known beam intensity distributions separated by a distance z. This is essentially due to the Fourier

transform property which is here involved as a tool to propagate from one plane to the other. It is an iterative algorithm that converges even faster when the intensity variations between the two planes are significant.
Figure 4.8 illustrates the experimental setup. The coherent light source is a HeNe laser. We expand the beam of the HeNe in order to illuminate an area, which was a round Gaussian beam with 7 mm diameter on the DOE what involves a magnification via the beam expander by $\times 17$. After the DOE an imaging system composed of a lens, a microscope objective and a CCD detector produces a reduced image $(\frac{1}{20})$ of a plane after the DOE. The imaged plane can be changed since the imaging system is mounted on a translation stage with a position precision better than the micrometer (Newport M ILS 100LM). Different images planes at different z positions after the beam is passed through the DOE are recorded.

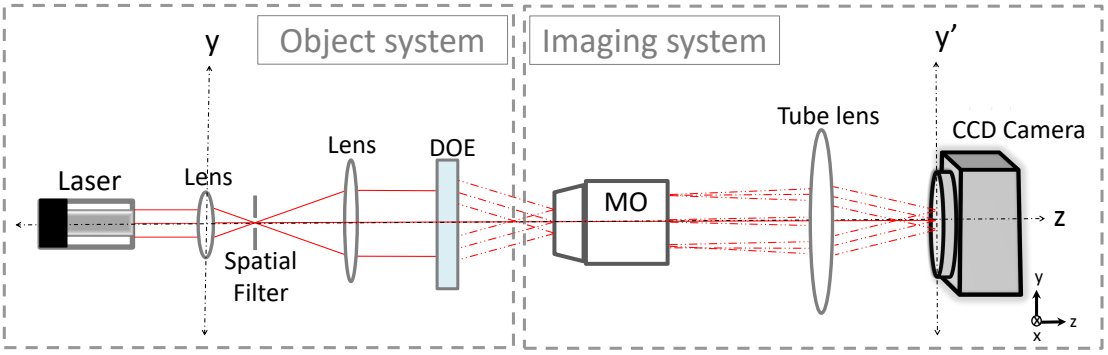

FIGURE 4.8 – Experimental setup to obtain numerically the structure of the DOE by using a phase retrieval technique.

The phase retrieval itself consists in an iterative procedure. Whereas the GS is applied between conjugated Fourier planes here it is applied between the direct and frequency space at a finite distance z. The procedure can be describe as follow. The intensity in different plane z_n is measured in our experiment and called I_n, yet phase ϕ_n is unknown. We start with I_0 which is the measured intensity in plane z_0. Then a non-paraxial propagation of the beam to the next z_1 plane is computed via the planar wave spectrum model [77]. This model consists in applying the Fourier transform to the beam in the direct space which is then multiplied by a phase term $\exp\left(i2\pi z\sqrt{\frac{1}{\lambda^2}-v_x^2-v_y^2}\right)$ where v_x and v_y are the spatial frequencies of the planar wave spectrum. Then the result of intensity profile is back Fourier transform leading to the new field $\sqrt{I_1'}exp(i\phi_1)$ in x, y coordinates at the z_1 plane. Then I_1' is replaced by I_1 which is known from measurement whereas the computed phase term is kept unchanged. Then, $\sqrt{I_1}exp(i\phi_1)$ is retro-propagated until plane z_0 using the same procedure. For several planes $(z_2, ..., z_n)$ the operation is repeated until convergence. By using this operation, the phase in plane z_0 will be found.
We require phase retrieval to obtain the phase distribution from the DOE. In that case, we apply the described procedure. Figure 4.9 shows the volume of recorded propagation in five planes $(z_0, ..., z_4)$. The measured structure presents the diffraction angle, and our final numerical scheme from phase retrieval technique is shown in Fig. 4.10.

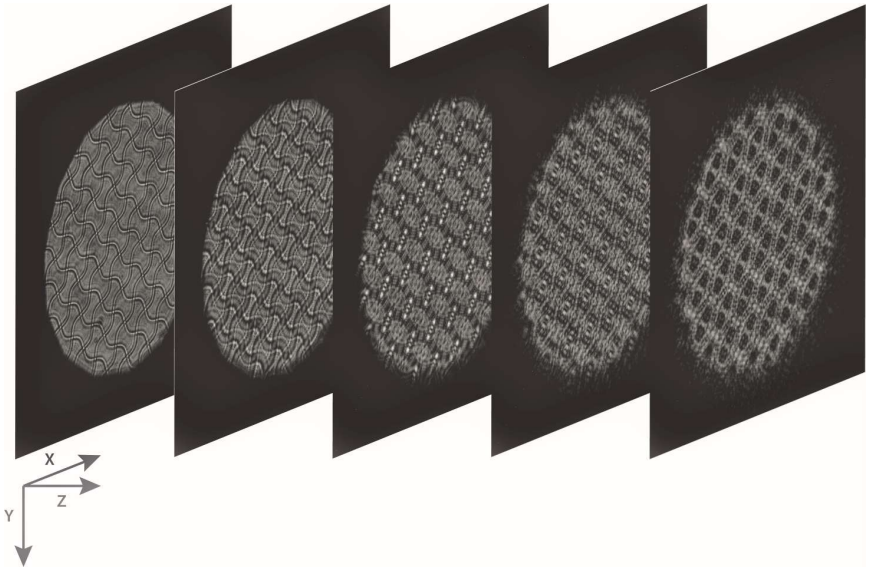

FIGURE 4.9 – Volume of recorded propagation during five plane $(z_0, ..., z_4)$ by implementing the phase retrieval technique.

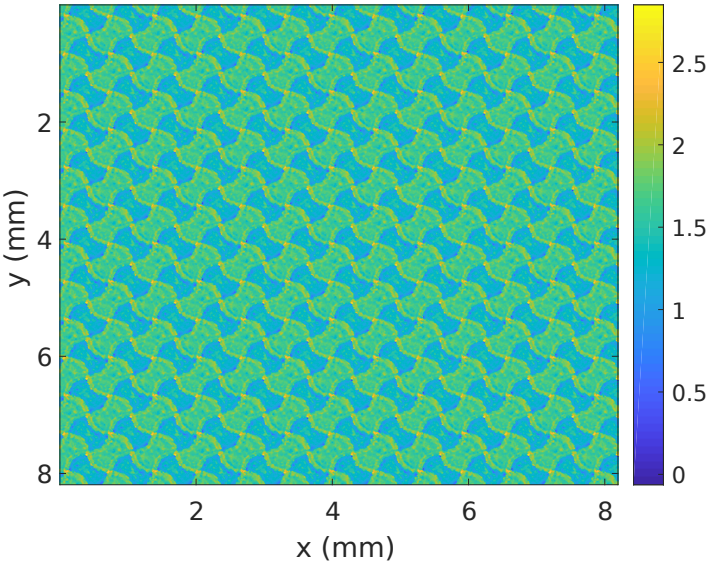

FIGURE 4.10 – Periodic pattern in our DOE which is obtained based on a phase retrieval technique. The color bar axis is for intensity in arbitrary unit.

4.5/ NUMERICAL IMPLEMENTATION

According to all explained techniques in this Chapter, we start setting up our numerical simulation. At first, we make a simulation for the first experiment from Fig. 3.10. Since this simulation incorporates the major limitations of the optical components, we should observe the physical effects such as vignetting. We then modify the setup in order to overcome these limitations using commercially available components. The obtained data points from numerical simulations will be compared with analytic.

4.5.1/ USING LOW MAGNIFICATION MICROSCOPE OBJECTIVES WITH A LOW NUMERICAL APERTURE

As a first system we investigate the setup schematically illustrated in Fig. 4.11 in our numerical simulation. All elements and their parameters are selected from the first experimental setup Fig. 3.10. As a special characteristic in this setup, microscope objectives have low magnifications and low NAs. All components of this figure are described in the numerical model.

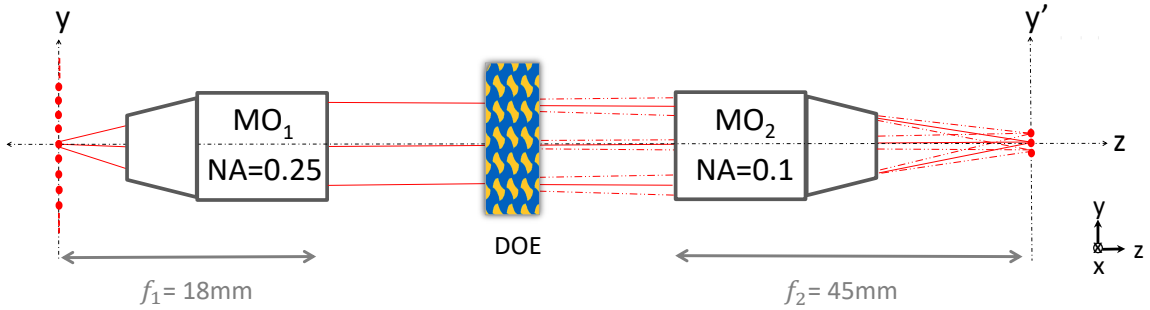

FIGURE 4.11 – Numerical setup with an array of emitters in plane y. All components and their parameters are the same as in the experiment of Fig. 3.10. MO_1 has a NA $= 0.25$ and $f_1 = 18$ mm, MO_2 has a NA $= 0.1$ and a $f_1 = 45$ mm. The periodic DOE is located between the two MOs.

In Fig. 4.11, we create an array of Gaussian emitters in an area of $(-1.2 \text{ mm} < y < +1.2 \text{ mm})$ with a constant position in the x plane. As in the experiment, the width of each Gaussian emitter is $w = 3.5$ μm at $\frac{1}{e^2}$ and wavelength is $\lambda = 660$ nm. Since we have an image plane, we require to re-scale our image plane (x, y) according to in Eq. 4.33. Therefor, the (x, y) coordinate plane of our array is divided by $\sqrt{1 - \lambda^2 v^2}$. In this case the parameter v can be obtained from the information of MO_1. We define MO_1 with the focal length $f_1 = 18$ mm and NA $= 0.25$. From Eq. 4.4 we can obtained the pupil radius of MO_1 as $R_1 = 4.64$ mm. Consequently, the parameter v_1 can be introduced as $(v_1 = \frac{1}{2R^{max}} = 107 \text{ mm}^{-1})$. In MO_2 the focal length is $f_2 = 45$ mm and NA $= 0.1$. Radius pupil of this objective is calculated in Eq. 4.4 as $R_2 = 4.52$ mm. Hence, parameter v_1 will be $(v_1 = \frac{1}{2R^{max}} = 110 \text{ mm}^{-1})$.

4.5.2/ RESULTS

After fitting the created data in numerical simulation, the results can be showed. These results were obtained from two low magnification MOs with low NAs. As the experimental setup, we are seeking the data for the distance between diffractive orders, mismatch distance $|\Delta|$ and width of each diffractive order.

4.5.2.1/ DISTANCE BETWEEN DIFFRACTIVE ORDERS

In Fig. 4.12, for emitters $|y| < 0.8$ mm, we see the numerical data which follow the horizontal line with a distance of $|d_0^{\text{diff}}| = 12.2$ μm, which hence is the condition to have coupling in a periodic array of emitters. For emitters in $|y| > 0.8$ mm however, there is a divergence

away from the coupling condition. This divergence is dramatically increasing with the distance from the center of the array. For instance, at $|y_0| = 0.85$ mm, distance between the diffractive orders is $(12 \pm 0.2\ \mu\text{m})$. Already at $|y_{end}| = 1.21$ mm this distance has increased to $(12 \pm 1.2\ \mu\text{m})$.

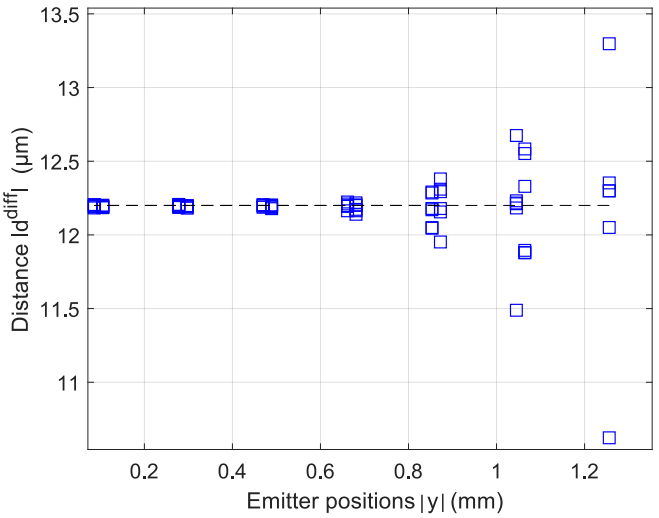

FIGURE 4.12 – Distances between diffractive orders $|d^{\text{diff}}|$ versus the emitter positions $|y|$. Blue rectangles are numerical data, the black dashed line is the distance for coupling a periodic array.

4.5.2.2/ DISTANCE MISMATCH

Another significant factor which was investigated in second Chapter was mismatch distance $|\Delta|$. To find the coupling limitation numerically, $|\Delta|$ is one of the main parameters in the entire simulation. In Fig. 4.13, mismatch $|\Delta|$ is shown for the full array of emitters in double-logarithmic scale. The numerical data points are the blue stars, the experimental data are red stars (from the third Chapter), while the analytical result from the second Chapter is the black line. Numerical data in Fig. 4.13 show three behaviours:
First behaviour, for the range of $|y| < 0.4$ mm the mismatch $|\Delta|$ is constant. This behaviour occurs because of the fit algorithm limited precision. Second behaviour, for distance 0.4 mm $< y < 0.9$ mm mismatch has a good agreement with analytical line. Third behaviour, for the range of $|y| > 0.9$ mm the mismatch has a strong divergence from the analytical solution. In numerical mismatch where $|y_0| = 1.2$ mm, $|\Delta|$ arrives $2\ \mu$m. This is the previously discussed vignetting effect of the MOs with a low NA.
We can compare the figures of $|\Delta|$ for the low NA in numerical simulation and the first experiment. The experimental results are demonstrated at the Fig. 4.13 in red stars. Now we can have a comparison for the area of $|y| < 0.4$ mm numerically and experimentally. In numerical simulation the limited precision of the fitting routine in 7 nm occurred. But in experimental data, additional uncertainty is added due to the detection noise. Both effects combined resulted in a position uncertainty of 17 nm.
It is important to point out that such high levels of position accuracy can only be achieved since we have accurate knowledge of the emitters' mode profile. This position accuracy is far beyond the diffraction limit, and our setup would certainly not be able to resolve emitters spaced by such a small distance. However, the accurate mapping of our imaging

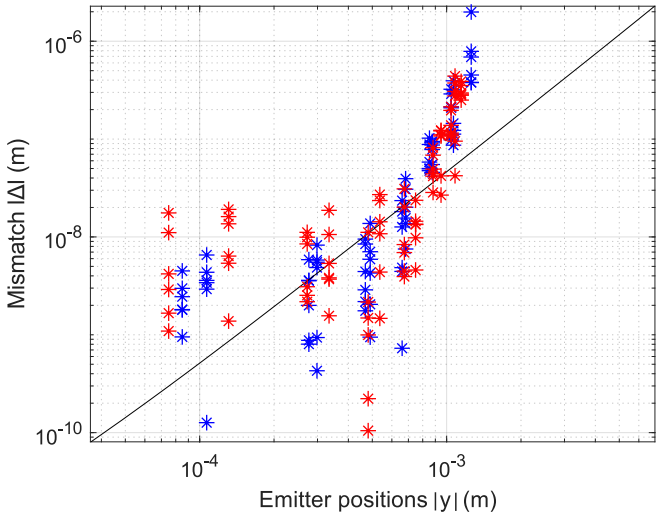

FIGURE 4.13 – Scheme of mismatch distance $|\Delta|$ versus emitter positions $|y|$ in a double-logarithmic scale. Blue stars are $|\Delta|$ in numerical simulation for array of emitter positions $|y|$. Black dashed line is analytical mismatch for array $|y|$. Red stars are $|\Delta|$ in experiment for array of emitter positions y.The good agreement between the red and the blue data.

setup, the low detection noise level and the precise knowledge of the mode profile allow an accurate position determination with this precision.
In Fig. 4.13 the numerical simulations excellently reproduce the divergences induced by optical beam vignetting as found in the first experiment from the previous Chapter. The properties of their imaging are obtained from the same fitting routine as in the previous section, the results of width for all diffractive orders are demonstrated in Fig. 4.14. In addition, the analytical indicating the diffraction limit for this imaging setup is shown as the black dashed line. For emitters $|y| < 0.87$ mm we have a good agreements between the analytical and numerical simulations and the diffraction limit. For $|y| > 0.87$ mm, there is a high divergence for numerical and experimental data points. At $|y_0| = 1.26$ mm the width of the individual imaged diffractive orders has reached 6.1 μm. From Fig. 4.14 we see that the numerical simulations and experimental data really agree perfectly well. This validates the model and we can fully understand the underlying mechanisms.

4.5.3/ USING LOW MAGNIFICATION MICROSCOPE OBJECTIVES WITH A LARGE NUMERICAL APERTURE

In order to solve the physical effects due to vignetting in our scheme, we decided to modify the setup. In the new scheme, we change the MOs' NAs as illustrated in Fig. 4.15 [50]. All other parameters are the same as Fig. 4.11
For MO_1, we increase the NA to $NA = 0.45$, while f_1 remains 18 mm as the Fig. 4.11. In this case, by considering the new NA and Eq. 4.4 we can calculate $R_1 = 9.1$ mm. The new R_1 is almost two times larger than R_1 from Fig. 4.11.
For MO_2, we change the NA to $NA = 0.2$ but f_2 is same as Fig. 4.11. Again by using Eq. 4.4 and the new parameter for NA we can write $R_2 = 9.2$ mm.

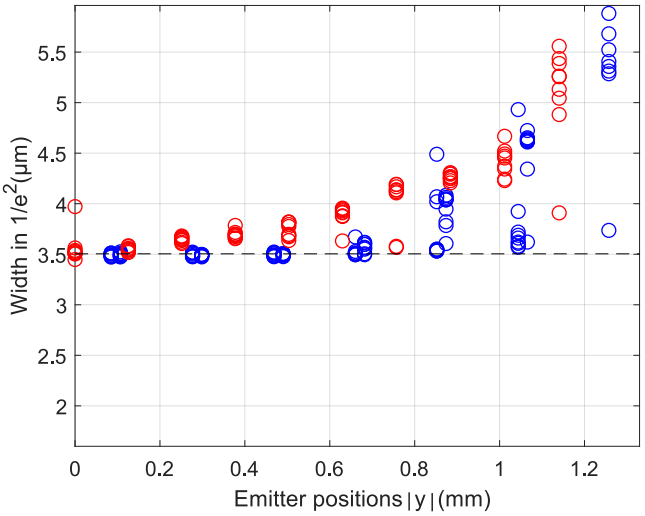

FIGURE 4.14 – Width of all diffractive orders in plane y' versus emitter positions $|y|$. Blue circles are width of diffractive orders in numerical simulation for array of emitter positions $|y|$. Red circles are width of diffractive orders in experiment for array of emitter positions y. Black dashed line is the width in analytical simulation for array $|y|$.

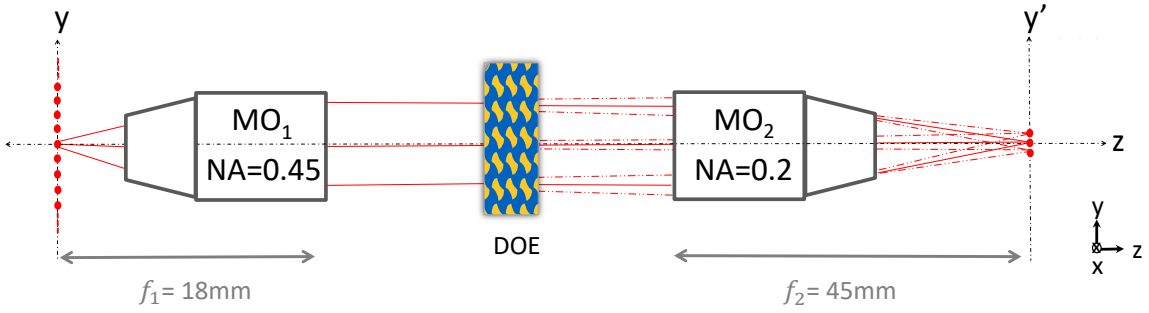

FIGURE 4.15 – Numerical setup with an array of emitters in plane y. This setup has the low magnifications in MOs and high NAs. MO_1 has high $NA = 0.45$ in focal length $f_1 = 18$ mm. MO_2 has a $NA = 0.2$ in focal length $f_2 = 45$ mm. The periodic DOE is located between both MOs.

4.5.4/ RESULTS

In next sections, we will present precisely the measurements of distance between diffractive orders, mismatch $|\Delta|$ and width of diffractive orders for the new scheme.

4.5.4.1/ DISTANCE BETWEEN TWO DIFFRACTIVE ORDERS

In the Fig. 4.15 and in y plane we create 49 emitters as our array. These emitters are distributed in a distance of 10 mm. Figure 4.16 shows the distances between diffractive orders $|d^{\text{diff}}|$ for the absolute emitter array of $(-5 \text{ mm} < y < +5 \text{ mm})$ as blue rectangles. The coupling condition for a periodic array is given by the black dashed line $|d_0^{\text{diff}}| = 12.2\ \mu$m. For all emitters the distances between diffractive orders are distributed from $|d_0^{\text{diff}}| = 12.21\ \mu$m until $|d_{end}^{\text{diff}}| = 14\ \mu$m. The maximum difference between the blue rectangles and the dashed

line is $1.8\ \mu$m. In the range of $(-1\text{ mm} < y < +1\text{ mm})$ we have almost 100 percent coupling between the orders. Due to the diffraction at $|y_0| = 5$ mm we have only 50 percent coupling where the width of diffractive orders are $3.5 \pm 0.5\ \mu$m. But the distance between diffractive orders $|d^{\text{diff}}|$ remain close to the required condition such that coupling is still maintained, though it could be substantially reduced to the outer parts of the array.

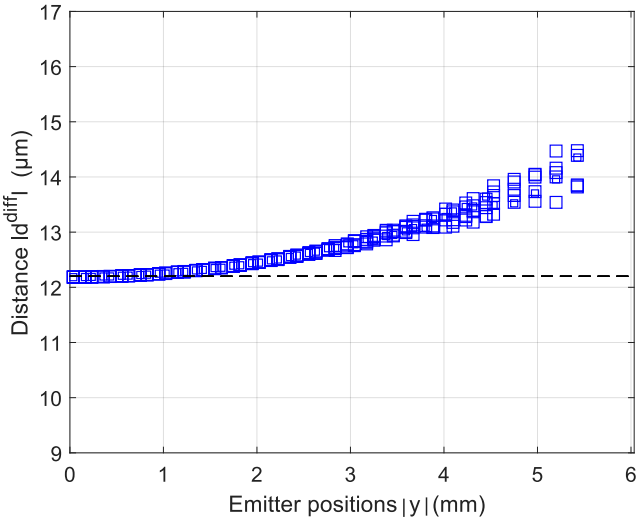

FIGURE 4.16 – Distance between diffractive orders $|d^{\text{diff}}|$ versus emitter positions $|y|$. Blue rectangles: distribution of the distance in numerical simulation. The black dashed line: the constant distance of the alignment condition in a periodic array.

4.5.4.2/ DISTANCE MISMATCH

Figure 4.17 illustrates the mismatch distance $|\Delta|$ for all diffractive orders in absolute y axis $(-5\text{ mm} < y < +5\text{ mm})$. Mismatches $|\Delta|$ are plotted in a double-logarithmic scale for numerical data points in blue stars. The analytical results are in a black line.

Such as our experiment we can divide this results in some parts based on data behaviors. For $|y| < 0.4$ mm the mismatch distance is constant in 7 nm. However, for $|y| > 0.4$ mm there is a good agreement between analytical and numerical data. By moving to the new NAs, we could eliminate the vignetting problem. In addition, the agreement between data from Fig. 4.17 confirmed the coupling in our setup.

4.5.4.3/ WIDTH OF DIFFRACTIVE ORDERS

The width of each Gaussian beam in our array is $w = 3.5\ \mu$m. We investigate the width for each diffractive order after propagation in plane y'. The result of width are shown in Fig.4.18. In this figure the numerical data are in blue circle and the width of the diffraction limit is given by the black dashed line. In this setup for $|y| < 2$ mm the width of the diffracted order images remain close to the diffraction limit. For $|y| > 2$ mm this agreement is reduced slowly and the difference between numerical and analytical data points arrives at $2\ \mu$m in $|y_0| = 5$ mm. This difference is of the order of half the diameter of our emitters mode profile, and hence still a significant part of coupling should be maintained. Finally, the

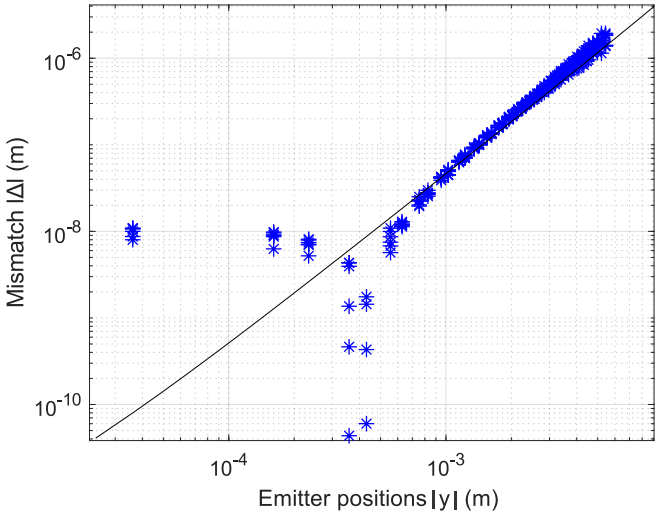

FIGURE 4.17 – Mismatch $|\Delta|$ versus emitter positions $|y|$ in a double-logarithmic scale. Blue stars are distribution of numerical mismatch after fit. Black line is analytical mismatch.

beam diameters remain close to the required condition that coupling is still maintained, though it could be substantially reduced to the outer parts of the array.

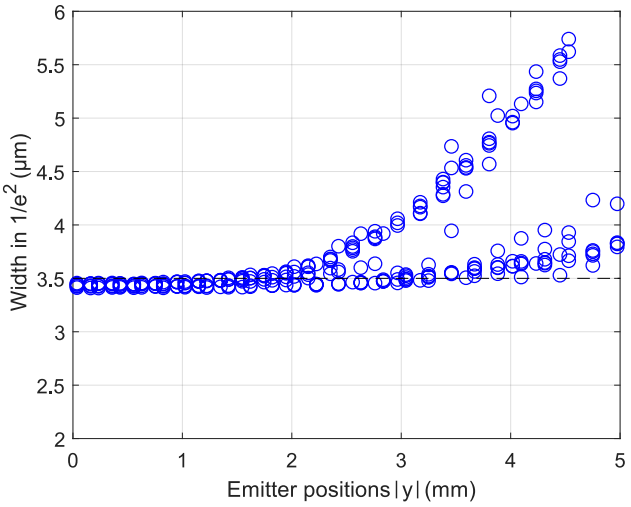

FIGURE 4.18 – Width for all diffractive orders in $\frac{1}{e^2}$ versus emitter positions $|y|$. Blue stars are distribution of numerical width after fit in array of emitter positions $|y|$. Black dashed line is diffraction limit.

4.5.4.4/ DISTANCE MISMATCH WITH AND WITHOUT RE-SCALING THE IMAGE PLANES

To show the importance of re-scaling of the image plane (dividing the correction factor to (x, y) plane in paraxial state), we plotted $|\Delta|$ in a double-logarithmic scale for numerical data points with re-scaling the planes in Fig. 4.19 in blue stars, $|\Delta|$ without re-scaling are as brown stars in Fig. 4.19. The brown stars show a periodic emitter position $|y|$, which occurred because we applied the paraxial approximations $(\tan\theta = \sin\theta = \theta)$ in our

calculation to obtain the image planes. While, in blue stars the computed image planes are based on non-paraxial re-scaling $(\tan\theta \neq \sin\theta \neq \theta)$.

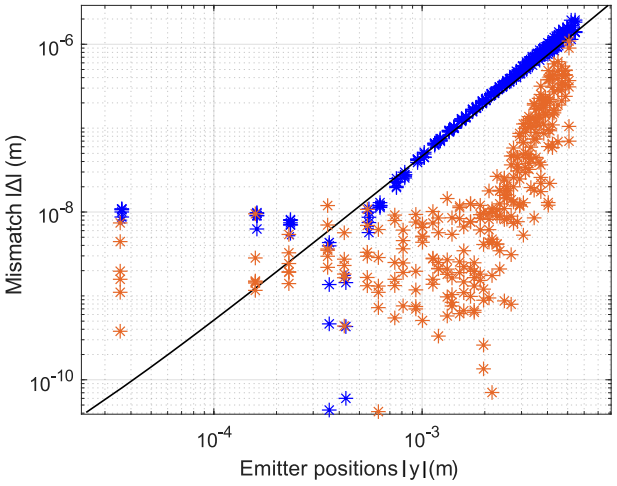

FIGURE 4.19 – Mismatch $|\Delta|$ versus emitter positions $|y|$ in a double-logarithmic scale. Blue stars are the distribution of numerical mismatch with non-paraxial re-scaling of the image planes. Brown stars are distribution of numerical mismatch with paraxial calculation in the image planes. Black line is analytical mismatch.

4.6/ CONCLUSION

In this Chapter we investigated coupling in numerical simulations for two setups. We reproduced the experimental setup in detail in our numerical simulations. We used several techniques during our simulations:

- In order to include the relevant effects which might limit our approach, it is crucial that the paraxial approximation is avoided during our simulation.
- To simulate numerically the propagation of beam, we introduced discrete Fourier transform (DFT). To compute the DFT we had to consider the number of samples and Nyquist frequency. During the propagation of optical field we included an absorbing boundary to remove the effect of a periodic boundary which appeared due to the periodic boundaries introduced by the FFT.
- Simulation of a lens or MO requires sampling the wave's phase profile with a large number of sampling data points, which renders the use of a standard desktop PC unpractical. We introduced the Debye integral as a solution to compute the field at the end of focal point based on a smaller number of sampling points. Calculating imaging by a MO based on the Debye's integral requires projecting the incident field on to cylindrical coordinates. This creates a decomposition into the optical field's radial (p-polarized) and tangential (s-polarized) components. The transformed field is then projected onto the MO's principle plane in spherical coordinates. After these transformations, the electric field E at a point (x, y, f) near the focus is obtained by integrating the propagated plane waves.
- We modulated the optical beam based on DOE's phase profile. This phase profile was obtained from the optical component by using a phase retrieval method.
- Each MO's clear aperture is limited by applying a circular mask with a transparent center and an opaque outer area. The radius of this circle is defined by the radius of the MO.
- Since the Debye integral method utilized the paraxial approximation, we required to re-scale the images planes. The principle was a coordinate re-scaling which accounts for the relation between spatial positions and spatial frequencies.

After explaining the techniques, we obtained the distance between diffractive orders, mismatch distance $|\Delta|$ and width of each diffractive order then compared the experimental results to the numerical model. In numerical simulations we observed the vignetting effects comparable to the experimental data of the previous Chapter. That vignetting events occurred because of low NA in MOs. We then numerically investigated another system where we used the same configuration but we increase the NAs for both MOs using $NA_1 = 0.45$ and $NA_2 = 0.2$. This augmented the MOs' diameters, the problem of vignetting was solved.

We explored the fundamental limit of coupling based on diffraction and investigated how many photonic neurons can maximally be coupled. For the high NA version of the concept the distance between diffractive orders remains inside $12.2\ \mu\text{m} < |d^{\text{diff}}| < 14\ \mu\text{m}$ for all emitters in distance 10 mm. In two dimensions and for the resulting area of 100 mm^2, the optical emitters can therefore be coupled for $671,300$ emitters.

5

PHOTONIC NEURAL NETWORK

In previous Chapters, we evaluated experimentally and numerically diffractive coupling in a simple setup. In the first Chapter, we introduced a spatio-temporal reservoir computer [37, 130] and in this Chapter we describe the experimental setup to build such a system [8]. First, we present a simplified scheme of the experiment which illustrates the relevant components. In this study we use a reinforcement learning based process consisting of multiple iteration steps and we used the MG prediction task to determine its performance. Finally, our photonic RNN successfully generalized the target system's properties.

5.1/ INTRODUCTION TO RECURRENT NEURAL NETWORKS

Figure 5.1 shows a single layer of a recurrent neural network. Circles are considered as neurons, and the connections which are illustrated as arrows. In our neural network, information enters the system via a single input node, as illustrated in Fig. 5.1, which is coupled to a recurrently connected network of nonlinear nodes, which are realized with a spatial light modulator (SLM). Recurrent and complex network connections are implemented via the DOE-based diffractive coupling discussed during the previous Chapters. The resulting connectivity is matrix W^{DOE}, which we realize using a digital micro-mirror device (DMD). The network's state is summed according to weight matrix W^{DMD}. Then, the computational result is provided at the single output node. We will explain W^{DMD} and W^{DOE} in more details in the next section.

5.2/ EXPERIMENTAL SETUP AND DIFFRACTIVE NETWORK

In Fig. 5.2, we schematically illustrate our hardware neural network, which consists of the following optical elements [8]:

- BS: A beam splitter is designed to split the incident beam into two beams with two different directions at a Reflection/Transmission $(R = 0.5/T = 0.5)$ ratio.
- PBS: A polarizing beam splitter divides light into two polarized light transmitting p-polarized light and reflecting s-polarized light.
- SLM: A spatial light modulator modulates intensity or phase of light. Spatial light modulators are two dimensional array of pixels distributed across a liquid crystal cells with the possibility of addressing them individually. In our case, the SLM is operated with an auxiliary computer. The SLM reflects light by a reflective surface

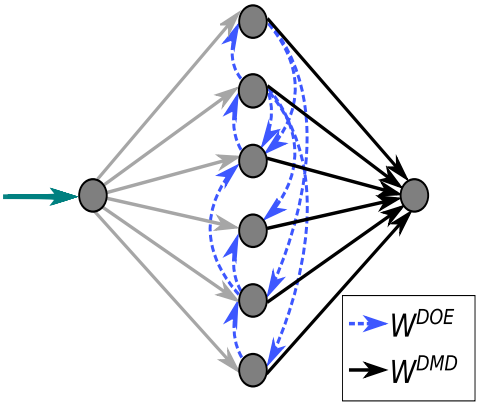

FIGURE 5.1 – Schematic illustration of a recurrent neural network with the input (green arrow), which is coupled to the nonlinear nodes. The recurrent connections of nodes are illustrated by blue arrows. The system's output as created by the weighting the network states with W^{DMD}.

in the back to reflect the incoming light. In our reservoir, each cell of the SLM is considered as a node. In this experiment, the lateral size of the SLMs' cells which are squares, is $12.5\ \mu$m [131, 132].

- DMD: A digital micro-mirror device is an array of aluminum micro-mirrors to reflect the light. Each micro-mirror has two electrodes underneath, and their controllable charge can tilt the mirror between the two extreme locations. For the mirrors, there are two tilting angles where the output values, which are Boolean are defined based on these angles. [133].

An illumination laser (Thorlabs LP660-SF20, $\lambda = 661.2$ nm, $I_{bias} = 89.69$ mA, T=23^{o}C) is adjusted to *s*-polarization. A consecutive $50/50$ BS reflects the illumination beam toward the PBS, which reflects the *s*-polarization into MO_1 (MO_1, Nikon CFI Plan Achro10×). When the illumination laser focuses on the back focal plane of MO_1, then the SLM pixels (*i=1,...,N*) are illuminated by a plane wave of amplitude E_i^0. A $\frac{\lambda}{2}$ plate between PBS and MO_1 is used to control the polarization of the entrance beam in order to optimize for intensity modulation by the SLM. Finally, the optical field in *p*-polarization is transmitted through the PBS, and at integer time n, the field behind the PBS is obtained by

$$E_i(n) = E_i^0 \cos\left(\frac{2\pi}{k_{\mathrm{SLM}}}\left(x_i^{\mathrm{SLM}}(n) + \theta_i^0\right)\right), \tag{5.1}$$

here $x_i^{SLM}(n)$ is the SLM pixel's gray scale with defined value $x_i^{SLM}(n) \in \{0, 1, ..., 255\}$. $k_{SLM} = 244.6 \pm 1.6$ is the conversion between polarization angle in radiance and the pixel gray scale. $\theta_i^0 = 11.1 \pm 1.1$ is considered as a gray scale offset, which is constant. The values k_{SLM} and θ_i^0 are experimentally measured for the 900 pixels of the network.

The transmitted beam passes the DOE (HOLOOR MS-443-650-Y-X) and each pixel of the SLM is diffracted into 3 by 3 diffractive orders. All diffractive orders are transmitted through MO_2 and a mirror reflects the full signal back for a second pass through the DOE. Every individual pass creates 3 by 3 diffractive orders and the double pass creates the last diffraction pattern which is a convolution of the diffraction pattern with itself. On average these results are observed in a 5×5 diffraction pattern for each node. This pattern presented a coupling matrix W^{DOE} in Fig. 5.3 [8]. Self-coupling in the network is possible where the imaging setup is in a 4f architecture. For this reason, a $\frac{\lambda}{4}$ wave-plate is included, hence when the beam passes two times through $\frac{\lambda}{4}$, its polarization changes

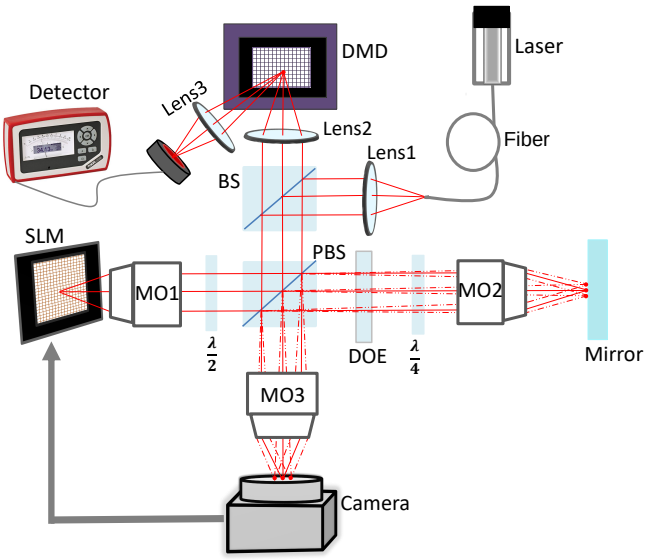

FIGURE 5.2 – Scheme of a recurrent neural network. The optical nodes are establishes by pixels of a SLM. Nonlinear nodes are created by a PBS which is a polarization filtering. Coupling between the nodes is fulfilled by a DOE. Before detection at the system's output, there is a digital micro-mirror device, which creates a spatially modulated image of the SLMs state.

from *p*-polarized to *s*-polarized. After going through the PBS, the now *s*-polarized beam is reflected towards the MO_3 (Nikon CFI Plan Fluor $4\times$) and imaged on the camera (CAM, Thorlabs DCC1545M). We require the camera's image in order to drive SLM. To create a recurrent network, a closed loop is needed which is found between the SLM and the camera. Camera state is given by

$$x_i^C(n) = \alpha \mid E_i \mid^2, \tag{5.2}$$

with

$$\alpha = \frac{GS}{I_{sat}} ND, \tag{5.3}$$

where $x_i^C(n) \in \{0, 1, ..., GS\}$ and GS= 255 is the 8-bit camera gray scale and I_{sat} is the camera's saturation intensity. To obtain the maximum dynamical range of the camera while avoiding overexposure, transmission through a neutral density filter (ND) is selected. In this experiment the magnification is 2.5 and size of each SLM's pixel is 12.5 μm, while this value is 5.2 μm for camera's pixel. The optical field on the camera is given by

$$E_i^C = \sum_j^N W_{i,j}^{DOE} E_j, \tag{5.4}$$

by using the SLM and the camera, we can obtain W^{DOE} (network coupling matrix) experimentally. Figure 5.3 illustrates W^{DOE} for a photonic network of 900 nodes. This matrix represents the 5×5 structure because of double pass trough DOE. Black expresses high coupling strength, white absent of coupling. To show more details, we added the inset zoom of this matrix for a smaller region in Fig. 5.3. From this inset, we can observe that the connectivity strengths diversify. The DOE's area that is passed by the collimated beam

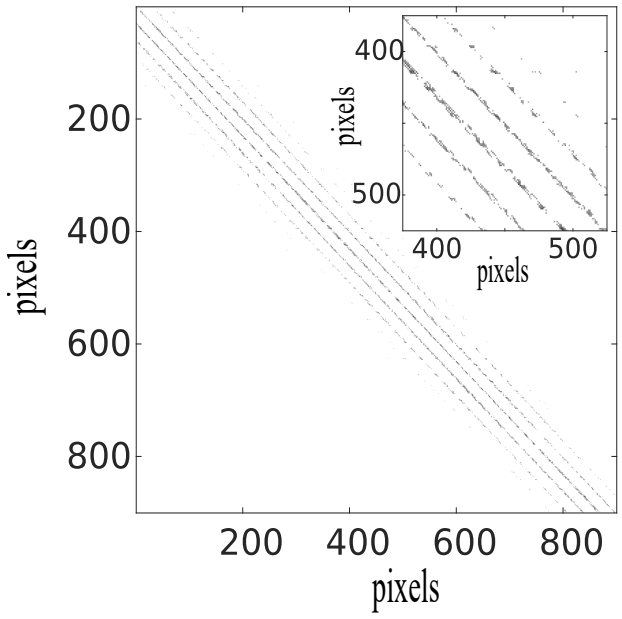

FIGURE 5.3 – Coupling matrix in our RNN for 900 nodes where black presents high coupling strength and white presents absent of coupling. A zoom for a smaller region shows better this coupling strength.

shifts slightly from pixel to pixel, resulting in a modification for the beam's phase modulation and as a consequence the intensity distribution between diffractive orders varies [8]. There are also external information, which must be coupled to the system by a matrix of injection W^{inj}, whose entires are randomly and uniformly drawn from $[0, 1]$. Each state $x(n+1)$ of photonic network is computed according to

$$x_i(n+1) = \alpha \mid E_i^0 \mid^2 \cos^2\left[\beta \cdot \alpha \left| \sum_j^N W_{i,j}^{DOE} E_j(n) \right|^2 + \gamma W_i^{inj} u(n+1) + \theta_i\right], \tag{5.5}$$

where $u(n+1)$ is the input information being injected to the network, γ is the injection gain and β is the scalar feedback gain. To control the instruments and to up date the network, MATLAB was using.

5.2.1/ READOUT

In order to achieve the desired computation in our system we require to adjust this system. Generally, it is possible by several learning rules to modify the connection weights [134, 135]. Our RNN nodes are spatially distributed. By using an array of micro-mirrors (DLi4120 XGA, pitch $13.68\ \mu$m), the optical signal can be flipped in two states either $+12°$ or $-12°$. In our experiments, where the micro-mirrors are in the state $-12°$, the output signal is measured by a detector (DET, Thorlabs PM100A, S150C). As a result, the DMD can be considered as an optical spatial filter to implement the output weights. For the system where readout weights were Boolean, the signal at the detector is described by

$$y^{out}(n+1) \propto \left| \sum_i^N W_{i,k}^{DMD} \left(E_i^0 - E_i(n+1)\right) \right|^2, \tag{5.6}$$

where $W^{\mathrm{DMD}}_{i=1,..,N,k}$ is a square matrix and k is the current learning iteration or learning epoch. Unlike delay reservoir, here the weights are not temporal modulations and Eq. 5.6 is fulfilled for all parameters in parallel.

5.2.2/ PHOTONIC LEARNING

In this section, we evaluate the learning procedure and prediction of the MG series [136, 137]. First, we introduce the MG equation. The MG equation is nonlinear time delay differential equation

$$\frac{dx}{dt} = B\frac{x(t-r)}{1+x(t-r)^n} - Ax, \tag{5.7}$$

for instance $A = 1$, $B = 2$, $r = 1$ and $n = 9.65$. Based on the values of parameters, this equation displays a range of periodic and chaotic dynamics. In Eq. 5.6, $W^{DMD}_{i=1,..,N,k}$ got $(k = 1, 2, ..., K)$ learning iterations, in that case $y^{out}(n+1)$ approximates the desired response for $u(n+2)$ at $k = K$.
We used two hundred points of the chaotic Mackey–Glass sequence as $u(n+1)$. At each iteration we modify RNN's output weights. In this modification if the error is reduced, a reward signal is taken, otherwise it will be rejected. We therefore teach our photonic RNN [138–140]. By this method, our photonic RNN can generalize the target system's properties. The good prediction performance can be seen in Fig. 5.4. On the left y axis of this figure in blue line, data shows the recorded output power. Data on the right y axis in red dots is according to the normalized prediction target signal. A difference between blue line and red dots is hardly visible. The prediction error ϵ in yellow dashed line is also very small [8]. As a result, our system was successfully trained for chaotic signal prediction. We therefore realized learning fully implemented photonic hardware.

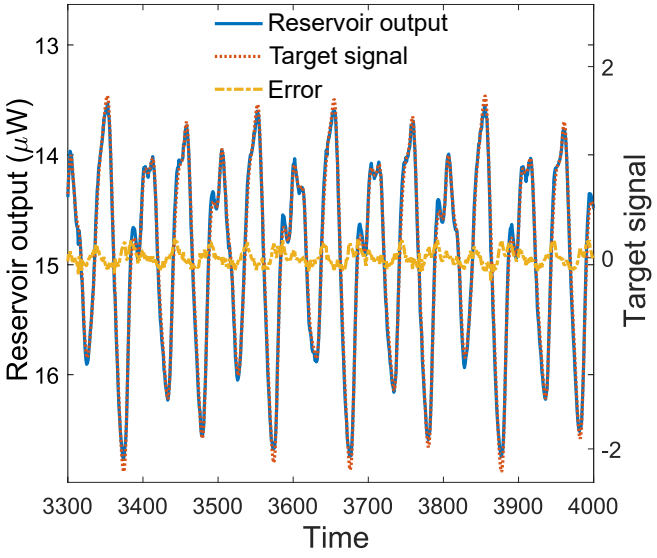

FIGURE 5.4 – Reservoir output in μm in blue line and prediction target in red dots. Red dots and blue line can hardly be diffracted. Prediction error is in yellow line.

5.3/ CONCLUSION

In this Chapter, a concept of photonic recurrent neural network was presented. Internal connections between neurons were established by diffractive coupling, which constructed the connectivity matrix of W^{DOE}. We successfully obtained W^{DOE} for a network of 900 nodes in our experiment. If we apply this system as a neural network, which requires learning rules to optimize the configuration of W^{DMD}. The single output is the sum of network's state. Besides the network internal connections, the system was coupled to information injection. The signal reflected by the PBS was imaged onto the surface of a DMD, which allowed the detection of a weighted network state. Readout connection was via a digital micro-mirror device, which could be flipped between $\pm 12^{\circ}$. Weights W^{DMD} were adjusted according to reinforcement learning. For measuring the system's performance, we introduced the MG system and we evaluated one step ahead prediction for this equation. This system was successfully trained for chaotic signal prediction and our recorded output got a good agreement with signal target prediction and the prediction error was small.

6

Conclusion and Outlook

6.1/ Conclusion

In this thesis we have analytical, experimental and numerical investigation to obtain the maximum size of an optical network coupled via diffraction. Analytically, we evaluated the fundamental limits of coupling in our network by using a single DOE. Experimentally, two different setups were developed to go beyond these limitations. Numerically, different techniques in our simulation implemented to achieve the desirable performance. In a second try, we modified the numerical model to explore the maximum number of nodes which can be coupled. Finally, the confirmation of coupling in the system allowed us to study the whole spatial RNN in order to predict the chaotic MG sequence. This thesis was organized as follows:

In Chapter 1, we explained ANNs as a substrate of machine learning, which tries to crudely mimic the structure of human brain which was a complex network of neurons. We started with a highly simplistic model of a neuron, the perceptron. One of the most basic ANN is the feed-forward neural network. In this model, neurons are arranged in layers and connections between the neurons are always in the forward direction. Another important class of ANNs are RNNs, which are circular networks including unidirectional coupling of neurons. In RNNs information in the network is encoded in its dynamical state, creating an internal memory. We then studied neural networks training to find an optimal configuration of connection weights which minimize the error of a computation to be carried out. We presented feed forward and recurrent neural network training, and reservoir computing as a new method reducing number of weights that are modified was discussed. Photonic neural networks have become popular reservoir computing systems because of high processing speed and a large connectivitys in a limited space achievable with photonic devices. We presented two photonic reservoir computing concepts: the photonic delay reservoir and the photonic spatio-temporal reservoir. In photonic delay reservoirs a single non-linear node under the influence of delayed feedback implements a neural network in a temporally multiplexed system. Hardware reservoir computing based on delay reservoir was demonstrated which was applied to channel equalization and speech recognition. As a second topology we introduced the creation of optical reservoir computing where the reservoir is extended in space. A primary motivation for optical spatio-temporal reservoir technology is to establish connectivity fully in parallel for a large number of photonic nodes. Hardware spatio-temporal reservoir will be discussed more in Chapter 5.

In Chapter 2 we were able to derive the fundamental limit for the size of photonic networks coupled via diffraction based on a simple analytical description. We started from the Maxwell equations to derive the propagation equation and the scalar Helmholtz equation and presented plane waves as solutions of the Helmholtz equation. The Kirchhoff diffraction integral and the Huygens principle are two ways to obtain the diffraction theory. In this thesis, the Huygens principle allows to derive the Rayleigh-Sommerfeld diffraction integrals, the Fresnel diffraction formula and the Fraunhofer diffraction formula. We continued to study the plane-wave spectrum as another method to propagate the complex optical fields. The plane-wave spectrum will be implemented in Chapter 4 in our numerical simulations of diffractive coupling. Finally, as an example light diffraction, the transmission and reflection grating were presented and the equations were obtained to utilize in our analytical description to approximate the effect of a DOE by a transmission grating.
We analytically explored the principles and limits of diffractive coupling by considering coupling arrays of discrete photonic emitters as the network's nodes. Imaging a periodic array through a periodic diffraction grating creates spatially multiplexed images. We obtained the maximum number of neurons which can be coupled. By using parameters mismatch coupling $|\Delta|$ and width of diffractive orders w, we were able to show where the coupling was terminated. Finally, for diffractive networks' size of $(16.4\text{ mm})^2$, we confirmed at least $1,795,600$ photonic nodes.

In Chapter 3, we investigated practical limits to the size of photonic networks coupled via diffraction. We characterized these based on two independent experiments which investigate fundamentally different limitations. In the first setup, a single mode optical fiber emulated optical nodes, and coupling for different spatial positions within a network was characterized by translation of the optical fiber's position. The beams diffracted by the DOE were imaged on the camera. The results from this setup were fit and characterized through the distance between diffractive orders, the mismatch distance and width of diffractive orders. We found that beam vignetting by the MOs induced strong aberrations. For this reason, we develop the second experiment in order to eliminate this practical limitations and to experimentally confirm the analytical limit. In a second setup instead of translating the fiber we tilt the DOE which is linked to position of emitters where each emitter position is equal to multiplication of the focal length of MO into tangent of the tilted angle of DOE. We obtained three parameters distance between diffractive orders, mismatch distance and width of diffractive orders for tilted angles in a range of 0 degree to 22 degree. Tilting the DOE therefore emulates different node positions without the effect of beam vignetting and we obtained an excellent agreement between the analytical model and the experimental data. Results confirm the validity of diffractive coupling far beyond what was in the first experiment and for an area of $(16\text{ mm})^2$.

In Chapter 4, we numerically modeled beam propagation, diffraction and collimation/imaging from the first experiment. A correct description of the underlying optical processes requires considering effects of coherent light propagation through the imaging system. We avoided paraxial approximations such as $\sin\theta \simeq \tan\theta \simeq \theta$ and the plane-wave spectrum was the method for describing propagation of an electromagnetic field during the simulation. We employed the Debye integral method to simulate the propagation through high NA microscope objectives. A rescaling of image plane was applied to all

numerical simulations which accounted for the relation between spatial position and spatial frequencies in an imaging system beyond the paraxial approximation limit. To include the effect of the DOE, we modulated the optical beam between both microscope objectives according to the DOE's phase profile. The phase profile was obtained directly from the component based on a phase retrieval technique. We accurately reproduced the first experimental setup, including the directly measured distances from the MO_1 to the DOE and from the DOE to the MO_2, plus the physical properties of both microscope objectives as obtained from their data-sheets. In addition, we used the same fitting and data-extraction routing as for the experiment. The results of distance between diffractive orders, mismatch distance and width of diffractive orders were demonstrated and compared with the first experiment, and the numerical method excellently reproduced the vignetting effect as encountered in the experiment. We therefore modified the numerical simulation and replaced MO_1 with a high NA, low magnification microscope objective and simulations show that in this configuration beam vignetting would strongly be reduced. Moreover, the numerical simulation confirmed the validity of diffractive coupling under conditions comparable to a realistic experimental setting. Finally, our experimental, numerical and analytical investigation provided the first systematic analysis of size-limits for optical networks coupled via diffraction.

In Chapter 5, we discussed the development of our hardware spatial recurrent neural network. The goal was to develop and optically implement learning tools, information processing (by using reservoir computing) and the programmable readout layer. This RNN includes hundreds of nonlinear nodes by using a spatial light modulator where each pixel of the SLM established an individual node and in combination with a PBS realizes a nonlinear mapping for each node. We utilized coupling according to the previously introduced diffractive concept based on a DOE. By measuring the state of each node and feeding it back to SLM, the dynamics were obtained. We recorded the state of network with a CMOS camera and the input layer was realized digitally. We utilized the learning rules based on reinforcement learning. These learning rules controlled the state of a DMD upon which the network state was imaged and whose mirrors acted as output weights. We trained our system to predict the chaotic MG sequence and the resulting prediction error was very low. We also demonstrated a network with 900 nonlinear nodes.

6.2/ OUTLOOK

In this thesis, we were able to obtain the interesting methods and results in our optical system. There are several methods which can be investigated in future studies being listed in below:
In numerical simulations instead of phase retrieval method for DOE phase modulation, that structure can be obtained by mathematical equations. Changing the DOE pattern also can be highly interesting in order to compare with the diffractive coupling in this thesis.
Regarding our studies, we can recalculate our measurements in ray optics. In that case, the result of numerical simulation can easily be achieved by the commercial software such as optic studio (Zmax). The significant points in this application is that one can change easily all parameters of the lenses such as diameter, thickness, material, etc. As a result, we can design and optimize system with desirable parameters to reduce

exceedingly the aberration factor and to establish a larger number of nodes by enlarging the area of diffraction coupling. Finally, the discussed lenses or MOs can be fabricated by some companies to utilize in our experimental setups.
In Chapter 5, the spatially extended reservoir was investigated for prediction the chaotic MG sequence. This research can be extended by changing the learning rules for different applications such as image or voice recognition. Regarding the use of SLM in our reservoir computing, if it is replaced with an optical element like an optically addressed spatial light modulator (OASLM) to modulate light, we can explore an all optical reservoir computing.

BIBLIOGRAPHY

[1] T. Freeth and al, “Decoding the ancient greek astronomical calculator known as the antikythera mechanism,” *Nature*, 2006.

[2] P. Ambs, “Optical computing: A 60-year adventure,” *Advances in Optical Technologies*, 2010.

[3] N. Farhat, D. Psaltis, A. Prata, and E. Paek, “Optical implementation of the hopfield model,” *Applied optics*, 1985.

[4] A. Lohmann, “Principles of optical computing,” *Springer*, 1990.

[5] I. Markov, “Limits on fundamental limits to computation,” *Nature*, 2014.

[6] B. Eggleton, B. Luther-Davies, and K. Richardson, “Chalcogenide photonics,” *Nature photonics*, 2011.

[7] K. Liu, C. Ye, S. Khan, and V. Sorger, “Review and perspective on ultrafast wavelength-size electro-optic modulators,” *Laser Photonics Reviews*, 2015.

[8] J. Bueno, S. Maktoobi, L. Froehly, I. Fischer, M. Jacquot, L. Larger, and D. Brunner, “Reinforcement learning in a large scale photonic recurrent neural network,” *Optica*, vol. 5, p. 756–760, 2018.

[9] H. Caulfield, R. Soref, and C. Vikram, “Universal reconfigurable optical logic with silicon-on-insulator resonant structures,” *Photonics and Nanostructures (Elsevier)*, 2007.

[10] D. Miller, S. Smith, and C. Seaton, “Optical bistability in semiconductors,” *IEEE Journal of Quantum Electronics*, vol. 17, 1981.

[11] A. E. Chiou, P. Yeh, and M. Khoshnevisan, “Coherent image subtraction using phase conjugate interferometry,” *Nonlinear Optics and Applications*, 1986.

[12] J. Borel, J. Deutsch, G. Labrunie, and J. Robert, “Liquid crystal diffraction grating,” *US Patent*, 1974.

[13] P. C. James, L. D. William, and S. Sudeshna, “Introduction to focus issue: Intrinsic and designed computation: Information processing in dynamical systems-beyond the digital hegemony,” *Chaos*, 2010.

[14] L. Indiveri, L. Giacomo, and C. Shih, “Memory and information processing in neuromorphic systems,” *Proceedings of the IEEE*, vol. 103, p. 1379–1397, 2015.

[15] K.-H. Yu, A. L. Beam, and I. S. Kohane, “Artificial intelligence in healthcare,” *Nature Biomedical Engineering*, vol. 2, p. 719–731, 2018.

[16] A. Bahrammirzaee, “A comparative survey of artificial intelligence applications in finance: Artificial neural networks, expert system and hybrid intelligent systems,” *Neural Computing and Applications*, vol. 19, pp. 1165–1195, 2010.

[17] p. Antonik and al., “Online training of an opto-electronic reservoir computing,” *International conference on neural information processing*, vol. 9490, p. 233, 2015.

[18] W. Mcculloch and W. Pitts, “A logical calculus of the ideas immanent in nerous activity (reprinted from 1943),” *Bulletin of Mathematical Biology*, vol. 5, pp. 115–133, 1943.

[19] F. Rosenblatt, “The perceptron: A probabilistic model for information storage and organization in the brain,” *Psychological Review*, vol. 65, p. 386–408, 1958.

[20] M. Minsky and S. Papert, “Perceptrons an introduction to computational geometry,” *MIT press*, vol. 165, p. 386–408, 1969.

[21] M. Nielsen, “How the backpropagation algorithm works,” *Neural Networks and Deep Learning*, pp. 1–18, 2014.

[22] W. Gerstner, W.Kistler, R. Naud, and L. Paninski, “Neuronal dynamics,” *Cambridge Univ. Press*, pp. 14–17, 2015.

[23] D. Hassabis, D. Kumaran, and C. Summerfield, “Neuroscience-inspired artificial intelligence,” *Neuron*, 2017.

[24] L. Chenchen, Y. Q. andZ. Chi, J. Hao, W. Qing, and L. Hai, “A memristor-based neuromorphic engine with a current sensing scheme for artificial neural network applications,” *22nd Asia and South Pacific Design Automation Conference (ASP-DAC)*, pp. 647–652, 2017.

[25] G. Cybenko, “Approximation by superpositions of a sigmoidal function,” *Math. Control Signals Systems*, p. 303–314, 1989.

[26] C. Fernando and S. Sojakka, “Pattern recognition in a bucket,” *In Proceedings of the 7th European Conference on Artificial Life*, p. 588–597, 2003.

[27] K. Hornik, “Approximation capabilities of multilayer feedforward networks,” *Neural Networks*, vol. 4, p. 251–257, 1991.

[28] R. Douglas and K. Martin, “Recurrent neuronal circuits in the neurocortex,” *Current biology*, vol. 27, p. 496–500, 2004.

[29] K. Gregor, I. Danihelka, and A. Graves, “A recurrent neural network for image generation,” *32nd International Conference on Machine Learning, ICML 2015*, 2015.

[30] K. Funahashi and Y. Nakamura, “Approximation of dynamical systems by continuous time recurrent neural networks,” *Neural Networks.*, vol. 6, pp. 801–806, 1993.

[31] H. Jaeger, “The “echo state” approach to analysing and training recurrent neural networks,” *J. Comput. Phys.*, 2001.

[32] A. M. Andrew, *Reinforcement Learning: An Introduction*. Kybernetes, 1998.

[33] F. Pineda, “Generalization of back-propagation to recurrent neural networks,” *German National Research Center for Information Technology GMD Technical Report*, vol. 14, p. 148, 1987.

[34] M. Nielsen, *How the backpropagation algorithm works*. Neural Networks and Deep Learning, 2014.

[35] H. Jaeger, “A tutorial on training recurrent neural networks, covering bppt, rtrl, ekf and the ”echo state network” approach,” *German National Research Center for Information Technology*, vol. GMD Report 159, p. 1–46, 2002.

[36] K. Vandoorne, W. Dierckx, B. Schrauwen, D. Verstraeten, R. Baets, P. Bienstman, and J. Campenhout, “A tutorial on training recurrent neural networks, covering bppt, rtrl, ekf and the echo state network approach,” *German National Research Center for Information Technology*, pp. 1–46, 2002.

[37] K. Vandoorne, W. Dierckx, B. Schrauwen, D. Verstraeten, R. Baets, P. Bienstman, and J. V. Campenhout, "Toward optical signal processing using photonic reservoir computing," *Opt.Express*, vol. 16, p. 11182–11192, 2008.

[38] W. Maass, T. Natschläger, and H. Markram, "Real-time computing without stable states: a new framework for neural computation based on perturbations," *Neural Comput*, vol. 14, p. 2531–2560, 2002.

[39] Y. Paquot and al., "Optoelectronic reservoir computing," *Scientific reports*, vol. 2, 2012.

[40] F. Duport, B.Schneider, A. Smerieri, M. Haelterman, and S. Massar, "All-optical reservoir computing," *Optics express*, vol. 20, p. 22783–22795, 2012.

[41] T. Cover, "Geometrical and statistical properties of systems of linear inequalities with applications in pattern recognition," *IEEE Transactions on Electronic Computers*, pp. 326–334, 1965.

[42] D. Brunner and I. Fischer, "Reconfigurable semiconductor laser networks based on diffractive coupling," *Opt. Lett.*, vol. 40, p. 3854–3857, 2015.

[43] D. Brunner, B. Penkovsky, R. Levchenko, E. Schöll, L. Larger, and Y. Maistrenko, "Photonic neural networks in delay systems," *Journal of applied physics*, vol. 124, 2018.

[44] D. Brunner, M. Jacquot, I. Fischer, and L. Lager, "Photonic networks for neuromorphic computing," *OSA Technical Digest*, 2017.

[45] D. Psaltis and N. Farhat, "Optical information processing based on an associative-memory model of neural nets with thresholding and feedback," *Opt.Letters*, vol. 10, pp. 98–100, 1985.

[46] D. Brunner, M. Soriano, and I. Fischer, "High-speed optical vector and matrix operations using a semiconductor laser," *IEEE Photonics Technology Letters*, vol. 25, p. 1680, 2013.

[47] M. Nahmias, B. J. Shastri, A. N. Tait, and P. Prucnal, "A leaky integrate-and-fire laser neuron for ultrafast cognitive computing," *IEEE (JSTQE)*, vol. 19, p. 35, 2013.

[48] K. Vandoorne, W. Dierckx, B. Schrauwen, D. Verstraeten, R. Baets, P. Bienstman, and J. Campenhout, "Toward optical signal processing using photonic reservoir computing," *Opt. Express*, vol. 16, pp. 11182–11192, 2008.

[49] A. Akram, A. Piotr, H. Marc, and M. Serge, "Towards autonomous photonic reservoir computer based on frequency parallelism of neurons," *Optica*, vol. 2, pp. 438–446, 2015.

[50] S. Maktoobi, L. Froehly, L. Anreoli, X. Porte, M. Jacquot, L. Larger, and D. Brunner, "Diffractive coupling for photonic networks: how big can we go ?," *IEEE Journal of Selected Topics in Quantum Electronics)*, vol. 26, 2019.

[51] L.Larger, A.Baylón-Fuentes, R.Martinenghi, V.Udaltsov, Y.Chembo, and M.Jacquot, "High-speed photonic reservoir computing using a time-delay-based architecture: million words per second classification," *Physical Review*, vol. 7, 2017.

[52] J. Bueno, D. Brunner, M. Soriano, and I. Fischer, "Conditions for reservoir computing performance using semiconductor lasers with delayed optical feedback," *Opt. Express*, vol. 25, p. 2401–2412, 2017.

[53] Y. Paquot, F. Duport, A. Smerieri, J. Dambre, B. Schrauwen, M. Haelterman, and S. Massar, "Optoelectronic reservoir computing," *nature*, vol. 2, p. 78–80, 2012.

[54] H. Jaeger and H. Haas, “Harnessing nonlinearity: predicting chaotic systems and saving energy in wireless communication,” *Science*, vol. 304, p. 78–80, 2004.

[55] F. Triefenbach, A. Jalalvand, B. Schrauwen, and J. Martens, “recognition with large hierarchical reservoirs,” *Advances in Neural information processing systems*, vol. 23, pp. 1–9, 2010.

[56] H. Jaeger, M. Lukosevicius, D. Popovici, and U. Siewert, “Optimization and applications of echo state networks with leaky-integrator neurons,” *Neural Netw*, vol. 20, p. 335–352, 2007.

[57] L. Larger, M. Soriano, D. Brunner, L. Appeltant, J. Gutierrez, L. Pesquera, C. Mirasso, and I. Fischer, “Photonic information processing beyond turing: an optoelectronic implementation of reservoir computing,” *Opt. Express*, vol. 20, pp. 3241–3249, 2012.

[58] D. Brunner, M. Soriano, C. Mirasso, and I. Fischer, “Parallel photonic information processing,” *Nature Communications*, vol. 4, p. 1364–1367, 2013.

[59] C. Denz, *Optical neural networks*. Springer, 2013.

[60] K. Wagner and D. Psaltis, “Multilayer optical learning networks,” *Applied Optics*, 1987.

[61] Y. Abu-Mostafa and D. Psaltis, “Optical neural computers,” *Applied Optics*, 1987.

[62] M. Naruse, Y. Terashima, A. Uchida, and S.-J. Kim, “Ultrafast photonic reinforcement learning based on laser chaos,” *Science*, vol. 7, p. 8772, 2017.

[63] M. Born and E. Wolf, “Principles of optics: electromagnetic theory of propagation, interference and diffraction of light,” *Elsevier, 2013*, 2013.

[64] J. D. Jackson, *Classical electrodynamics (3rd ed.)*. Wiley, 1998.

[65] J. Maxwell, “Understanding and validity in qualitative research,” *Harvard Educational Review*, 1992.

[66] M. Born and E. Wolf, *Principles of Optics*. Cambridge University Press, 1999.

[67] M. Salam, *Uniform Plane Waves*. Springer, 2014.

[68] P. Clemmow, *The plane wave spectrum representation of electromagnetic fields*. The Plane Wave Spectrum Representation of Electromagnetic Fields, 1996.

[69] C. Huyghens, *Traité de la Lumière*. Journal of the Röntgen Society, 1920.

[70] G. O. Reynolds, J. B. DeVelis, G. B. Parrent, and B. J. Thompson, *Huygens’ Principle: The new physical optics notebook: Tutorials in Fourier optics*. SPIE, 1989.

[71] A. Dubra and J. Ferrari, “Diffracted field by an arbitrary aperture,” *American Journal of Physics - AMER J PHYS*, vol. 67, pp. 87–92, 1999.

[72] F. Shen and A. Wang, “Fast-fourier-transform based numerical integration method for the rayleigh–sommerfeld diffraction formula,” *Optical Society of America*, vol. 45, pp. 1102–1110, 2006.

[73] Y. G. Soskind, *Fresnel diffraction: Field guide to diffractive optics*. SPIE, 2011.

[74] F. W. Sears, *Optics*. Addison-Wesley, 1948.

[75] A. Lipson, S. Lipson, and H. Lipson, *Optical Physics*. Cambridge University Press, 2011.

[76] H. Eugene, *Optics*. Addison Wesley, 2002.

[77] J. Goodman, *Introduction to Fourier Optics 2ed*. 1968.

[78] P. Clemmow, *The plane wave spectrum representation of electromagnetic Fields*. Wiley-IEEE Press, 1997.

[79] C. V. Loan, *Computational Frameworks for the Fast Fourier Transform*. 1992.

[80] M. Heideman, D. Johnson, and C. Burrus, "Gauss and the history of the fast fourier transform," *IEEE ASSP Magazine*, vol. 4, p. 14–21, 1984.

[81] W. Cochran, J. Cooley, and L. David, "What is the fast fourier transform ?," *IEEE Transactions on Audio and Electroacoustics*, 1967.

[82] F. Harris, "On the use of windows for harmonic analysis with the discrete fourier transform," *Proceedings of the IEEE*, 1978.

[83] B. Usevitch, *The discrete fourier transform*. Mobile Communications Handbook, Third Edition, 2017.

[84] H. Nyquist, "Certain factors affecting telegraph speed," *Bell System Technical Journal*, vol. 3, pp. 324–346, 1924.

[85] M. Mintchev, *Sampling theorem and aliasing in biomedical signal processing*. Wiley, 2006.

[86] H. Nyquist, "Certain topics in telegraph transmission theory," *A.I.E.E. Transactions*, vol. 47, pp. 617–644, 1928.

[87] M. Mintchev, *Sampling theorem and aliasing in biomedical signal processing*. Wiley, 2006.

[88] A. Oppenheim, R. Schafer, and J. Buck, *Discrete-Time Signal Processing 3rd*. Discrete-time Signal Processing, 1989.

[89] E. Brigham and R. Morrow, *The fast Fourier transform*. IEEE spectrum, 1967.

[90] J. W. Cooley, P. Lewis, and P. Welch, "The fast fourier transform and its applications," *IEEE Transactions on Education*, 1969.

[91] P. Duhamel and M. Vetterli, "Fast fourier transforms: A tutorial review and a state of the art," *Signal Processing*, 1990.

[92] M. Lax, W. Louisell, W. McKnight, and B. William, "From maxwell to paraxial wave optics," *Physical Review A*, 1975.

[93] M. Bertolotti, "Waves and fields in optoelectronics," *Optica Acta: International Journal of Optics*, 1985.

[94] W. Cochran, J. Cooley, and L. David, "Optica acta: International journal of optics," *IEEE Transactions on Audio and Electroacoustics*, 1967.

[95] L. Rosenfeld, "Principles of optics. electromagnetic theory of propagation, interference and diffraction of light," *Nuclear Physics*, 1960.

[96] C. Huyghens, *Taylor expansions and applications. In: Mathematical Analysis*. Springer, 2008.

[97] P. Chamorro-Posada, G. McDonald, and G. New, "Non-paraxial beam propagation methods," *Optics Communications*, 2001.

[98] E. B. Treacy, "Optical pulse compression with diffraction gratings," *IEEE J. Quantum Electron*, vol. 5, 1969.

[99] N. Bonod and J. Neauport, "Diffraction gratings: from principles to applications in high-intensity lasers," *Advances in Optics and Photonics*, 2016.

[100] P. Christopher, *Diffraction grating handbook*. Journal of the Optical Society of America, 2005.

[101] R. Longhurst, *Geometrical and physical optics*. Longmans, 1967.

[102] F. Jenkins and H. White, *Fundamentals of optics*. McGraw Hill, 1957.

[103] V. Nascov, D. Apostol, and V. Damian, "On a diffractive optical element," *Proceedings of the International Semiconductor Conference, CAS*, 2004.

[104] H. Kim and B. Lee, "Iterative fourier transform algorithm with adaptive regularization parameter distribution for the optimal design of diffractive optical elements," *Japanese Journal of Applied Physics*, vol. 43, pp. 702–705, 2004.

[105] J. Turunen and F. Wyrowski, *Diffractive optics for industrial and commercial applications*. Wiley, 1997.

[106] B. Kress and P. Meyrueis, *Digital diffractive optics: An introduction to planar diffractive optics and related technology*. Wiley, 2000.

[107] J. Bengtsson, P. Modh, J. Backlund, H. Lindberg, and A. Larsson, "Progress in diffractive integrated optics, in diffractive optics and micro-optics," *Optical Society of America*, vol. 75, pp. 17–19, 2002.

[108] H. Kim, B. Yang, and B. Lee, "Iterative fourier transform algorithm with regularization for the optimal design of diffractive optical elements," *Optical Society of America A*, vol. 21, pp. 2353–2365, 2004.

[109] M. Csele, *Fundamentals of Light Sources and Lasers*. Wiley, 2004.

[110] W. T. Silfvast, *Laser Fundamentals*. Cambridge University Press, 1996.

[111] F. Harris, "On the use of windows for harmonic analysis with the discrete fourier transform," *Proceedings of the IEEE*, 1978.

[112] F. R. Sidney, *Applied Photographic Optics: Lenses and Optical Systems for Photography, Film, Video, Electronic and Digital Imaging*. Focal, 2002.

[113] S. Orazio, *Principles of Lasers*. Springer, 2010.

[114] B. Engquist and A. Majda, "Absorbing boundary conditions for the numerical simulation of waves," *MATHEMATICS OF COMPUTATION*, vol. 31, pp. 629–651, 1977.

[115] A. Majda and S. Osher, "Reflection of singularities at the boundary," *Communications on Pure and Applied Mathematics*, vol. 28, p. 479–499, 1975.

[116] H. Kreiss, "Initial boundary value problems for hyperbolic systems," *Communications on Pure and Applied Mathematics*, vol. 23, p. 277–298, 1970.

[117] S. C. Kim and Chung, "Light fantastic: The art and design of stage lighting," *Building and Environment*, vol. 46, pp. 210–222, 2011.

[118] M. Abramowitz, K. Spring, H. Keller, and M. Davidson, "Basic principles of microscope objectives," *BioTechniques*, vol. 33, pp. 772–781, 2002.

[119] J. LoBiondo, M. Abramowitz, and M. Friedman, "Microscope objectives," *Current Protocols in Cytometry*, 2011.

[120] M. Leutenegger, R. Rao, R. A. Leitgeb, and T. Lasser, "Fast focus field calculations," *Optics Express*, vol. 14, p. 11277, 2006.

[121] P. Torok and P. Varga, ""electromagnetic diffraction of light focused through a stratified medium," *Appl. Opt*, vol. 36, p. 2305–2312, 1997.

[122] J. Stamnes, *Waves in Focal Regions: propagation, diffraction and focusing of light, sound and water waves*. Hilger, Bristol UK, 1986.

[123] G. Mikula, A. Kolodziejczyk, M. Makowski, C. Prokopowicz, and M. Sypek, "Diffractive elements for imaging with extended depth of focus," *Opt. Eng.*, vol. 44, 2005.

[124] N. Huse, A. Schonle, and S. Hell, "Z-polarized confocal microscopy," *J. Biomed. Opt.*, vol. 6, p. 480–484, 2001.

[125] M. Mansuripur, "Certain computational aspects of vector diffraction problem," *J. Opt. Soc. Am*, vol. 6, p. 786–805, 1989.

[126] C. J. R. Sheppard, "Validity of the debye approximation," *Optics letters*, vol. 25, p. 1660–1662, 2000.

[127] E. Wolf and Y. Li, "Conditions for the validity of the debye integral representation of focused fields," *Optics Communications*, vol. 39, p. 205–210, 1981.

[128] R. Gerchberg and W. Saxton, "A practical algorithm for the determination of the phase from image and diffraction plane pictures," *Optik (Stuttgart)*, vol. 35, pp. 237–250, 1972.

[129] W. H. Press, W. T. Vetterling, S. Teukolsky, and B. Flannery, "Numerical recipes in fortran: The art of scientific computing.numerical recipes in c: The art of scientific computing," *Mathematics of Computation*, 1974.

[130] L. Appeltant, M. C. Soriano, G. V. der Sande, J. Danckaert, S. Massar, J. Dambre, B. Schrauwen, C. R. Mirasso, and I. Fischer, "Information processing using a single dynamical node as complex system," *Nat. Commun*, vol. 2, p. 468, 2011.

[131] N. Mukohzaka, N. Yoshida, H. Toyoda, Y. Kobayashi, and T. Hara, "Diffraction efficiency analysis of a parallel-aligned nematic-liquid-crystal spatial light modulator," *The Optical Society*, vol. 33, p. 2804, 1994.

[132] S. D. Coomber, C. D. Cameron, and J. R. Hughes, "Optically addressed spatial light modulators for replaying computer-generated holograms," *SPIE*, 2001.

[133] L. Hornbeck, "Multi-level digital micromirror device," *US Patent*, 1996.

[134] P. Antonik, M. Haelterman, and S. Massar, "Online training for highperformance analogue readout layers in photonic reservoir computers," *Cognit. Comput*, vol. 9, p. 297–306, 2017.

[135] H. Jaeger and H. Haas, ""harnessing nonlinearity: predicting chaotic systems and saving energy in wireless communication," *Science*, vol. 304, p. 78–80, 2004.

[136] J. D. Farmer, "Chaotic attractors of an infinite-dimensional dynamical system," *Physica D: Nonlinear Phenomena*, 1982.

[137] J. D. Farmer and J. Sidorowich, "Predicting chaotic time series," *Phys. Rev. Lett*, vol. 59, pp. 845–848, 1987.

[138] J. Bueno, D. Brunner, M. Soriano, and I. Fischer, "Conditions for reservoir computing performance using semiconductor lasers with delayed optical feedback," *Opt. Express*, vol. 25, p. 2401–2412, 2017.

[139] M. C. Ortín, S. Soriano, L. Pesquera, D. Brunner, D. San-Martín, I. Fischer, C. R. Mirasso, and J. M. Gutiérrez, "A unified framework for reservoir computing and extreme learning machines based on a single time delayed neuron," *Sci. Rep*, vol. 5, p. 14945, 2015.

[140] M. Naruse, Y. Terashima, A. Uchida, and S. Kim, "Ultrafast photonic reinforcement learning based on laser chaos," *Sci. Rep*, vol. 7, p. 8772, 2017.

LIST OF FIGURES

1.1 Graph of an analog neuron with a non-linear function $y = f(u_j, w_j)$. u_j are called the inputs and w_j are weights. 6

1.2 Scheme of a feedforward neural network with k inputs as input layer, m output neurons as output layer and h hidden neurons as hidden layer. . . . 7

1.3 Graph of a recurrent neural network (RNN). The recurrent layer in the middle with an internal connections keeps the information. For this reason, the output is depending on the previous inputs. 8

1.4 Scheme of a simple RNN. w and w_r are the map matrix and the readout map, respectively. 10

1.5 Scheme of a RNN where w and w_r are the map matrix and the readout map, respectively. This network is unfolded in time for k time steps to construct a feedforward neural network. 10

1.6 The architecture of a reservoir computer consist of: an input layer, the reservoir, and an output layer. Data are masked with a randomly generated mask w_i, then the second layer with random connections called reservoir, and the third layer which obtained by multiplication with readout matrix w_r. . 11

1.7 Scheme of an optoelectronic delay RC. There are two fundamental parts in this scheme. Optical and electrical parts are in red and green respectively. An optical fibre is used to create the delay loop. The input information is produced by computer and injected to the system by AWG (arbitrary waveform generator). The digitizer (oscilloscope) records the response of the system. 13

2.1 Structure of a plane wave. Blue planes are positive amplitude planes. Red planes are negative amplitude planes. Black sine wave shows the direction of plane wave propagation in z. In zero-crossing point the sign of a mathematical function changes (from positive to negative). 18

2.2 (a) Amplitude and phase of a plane wave for a given spatial positions. The time between two peaks is T, so angular frequency is $\frac{2\pi}{T}$. (b) Amplitude and phase of a plane wave at a given time. Distance between two peaks is λ, so the wave number is $\frac{2\pi}{k}$. 19

2.3 Wave front description. Distance between the peaks is wavelength λ. 20

2.4 Huygens principle. Every point on a wave front acts as an independent source to generate wavelets for the next wave front. Propagation of light is in Z direction and distance between two wave fronts is λ. Where waves approaches a barrier with a small slit, the points of plane waves emit spherical wavelets. 21

2.5 Source plane and observation plane used in the Rayleigh-Sommerfeld integral, where $x_1^{'}$ and $y_1^{'}$ are position vector in source plane, x and y are the position vector in the observation plane. The distance between position $(x_1^{'},y_1^{'})$ and position (x,y) is given by r. 22

2.6 Vector $\vec{r}$ in the coordinate system x and z. The plane wave propagates along z direction and θ is the angle between vector $\vec{r}$ and z component. . . 26

2.7 Diffraction of light on an optical grating. In the left side the diffraction grating is located with period a. The right side is a screen to show the diffraction pattern. Distance between diffraction grating and the screen is D. 28

2.8 Zoom on 2 slits in the optical grating. Distance between two slits is a. $|\delta|$ is difference of distance between the two diffracted waves. Angle θ is the angle between diffracted beam and optical axis. 29

2.9 Diffraction of light on two slits of the transmission grating. θ_0 is the incident beam which is non zero. θ_m is the angle between diffracted beam and optical axis. On the screen the diffractive orders $-1, 0, 1$ are established. . . 30

2.10 Diffraction of light on two slits of reflection grating. θ_0 is the incident beam which is non zero. θ_m is the angle between diffracted beam and optical axis. On the screen the diffractive orders $-1, 0, 1$ are established. 30

2.11 Principle of diffractive coupling. Two emitters are imaged via f_1 and f_2, where f_1 is focal length of MO_1 and f_2 is focal length of MO_2. The DOE is situated between MOs and creates three discrete diffractive orders $-1, 0$ and $+1$. The width of emitters are called w which is the Gaussian beam width at $1/e^2$. $\Theta_{\mathrm{i}}^{\mathrm{im}}$ is the angle between the principal rays of neighboring emitters and Θ^{diff} is the angle between diffractive orders. The distance between two emitters is p^{array} and between two diffractive orders is d^{diff}. $d_{\mathrm{b}}^{\mathrm{diff}}$ is the distance between orders $m = 0$ and $m = -1$ for emitter y_0 in blue color. $d_{\mathrm{r}}^{\mathrm{diff}}$ is the distance between orders $m = 0$ and $m = 1$ for emitter y_1 in red color. 32

2.12 Mismatch coupling between diffractive orders versus emitter position in z_{a}. Red dash lines identify the limitation of coupling in our hardware neural network. Inside the red dash line, there is coupling for emitters in the position of -8.2 mm $< y <$ $+8.2$ mm. While outside of red dash line, there is no overlapping between the diffractive orders. 35

3.1 Experimental scheme to test limits of diffractive coupling. 38

3.2 Experimental setup without the DOE to determine the magnification. 39

3.3 Recorded beam in center of camera without the DOE effect. 39

3.4 Experimental setup to analyze diffractive orders on the camera. 40

3.5 Experimental diffractive orders imaged on the camera. 40

3.6 Experimental setup using off-axis emitters. 41

3.7 Diffractive orders for $y_1 = 0.88$ mm. 42

3.8 Diffractive orders for $y_2 = 1.1$ mm. 42

3.9 Scheme to adjust the maximum distance between MOs (L) for emitters far from center and based on the diameter of MO_2 (R_2). 43

3.10 Setup after removing filters and adding the attenuator. 44

3.11 Diffractive orders on the camera for emitter in y_1 = 1.16 mm. 45

3.12 Distance between diffractive orders $|d^{\text{diff}}|$ in the y plane (red arrows) for 9 diffractive orders. 46

3.13 Distances between neighbouring diffractive orders $|d^{\text{diff}}|$ as a function of the emitter position. Black dash line: the constant distance for which in a periodic array's coupling is optimal. 46

3.14 Mismatch $|\Delta|$ versus emitter positions y from experimental data in a double-logarithmic scale (red stars). Theoretical data are plotted as a black line. . . 47

3.15 Width for 90 diffractive orders in experiment (red circles) relative to the diffraction limit (black dashed line) versus emitter positions y. Vignetting leads to a strong increase of the width for $y > 1$ mm. 48

3.16 Effect of aberrations on a beam after replacing MO_2 with an achromatic doublet. 48

3.17 Diffractive orders obtained by using tube lens. The effect of aberration changes the shape of diffractive orders. 49

3.18 Experimental setup to test limits of diffractive coupling by tilting the DOE angle from 0 degree to 22 degrees. 50

3.19 Red rectangles: 120 distances between $|d^{\text{diff}}|$ diffractive orders versus emitter positions y from $y_0 = 0$ mm to $y_{lim} = 8$ mm. Black dashed line: distance between diffractive orders in analytical solution. 51

3.20 The diffractive coupling mismatch $|\Delta|$ obtained by the experiment (red stars) and the analytical limit from Eq. (1.10) (black line) excellently agree. 51

3.21 Distribution of width for 180 diffractive orders in $1/e^2$ from $y_0 = 0$ mm to $y_{lim} = 8$ mm in red circles. The diffraction limit is in black dashed line. 52

4.1 The scheme of numerical simulation. The mentioned explanations are the techniques which are used during our simulation. 55

4.2 Observation of the periodic boundaries introduced by the FFT during beam propagation. A Gaussian beam in angle $\theta = 53$ degree is propagated along the z direction. It leaves the simulated window from one side (bottom) and re-enters from the another side (top of window). This is a numerical and un-physical artefact. 57

4.3 Proposed models to remove the boundary effect during beam propagation. (a) The hyper Gaussian absorbing boundary. (b) The rectangular absorbing boundary. 57

4.4 Absorbing boundary solution for a Gaussian beam in angle $\theta = 53$ degree is propagated along z direction in Fig. 4.2. (a) The hyper Gaussian to absorb the beam smoothly at the edge of window. (b) The inset showing a zoom into the edge of window in hyper Gaussian absorbing boundary. (c) The rectangular absorbing boundary to absorb the beam rectangulary at the edge of window. (d) The inset showing a zoom into the edge of window in rectangular absorbing boundary. 58

4.5 Principal of the Debye integral. Optical field components are projected onto the microscope objective principal plane $\mathbb{P}_1$ then $\mathbb{P}_2$. The focal length f is calculated as $f = \overline{\mathrm{F_1V_1}}$. The point **P** lies on the plane $\mathbb{P}_2$ and shows the focusing of a ray at the focal point F_2, reproduced from [120]. 60

4.6 An imaging system in our optical setup. Image plane coordinate is x and frequency plane coordinate is v. Beam is imaged in focal length f. 63

4.7 Results of an array of Gaussian emitters in paraxial and non-paraxial coordinate systems. Blue color are non-paraxial and red color are paraxial simulations. Difference between both coordinate systems are considerable at $y > 2$ mm. 64

4.8 Experimental setup to obtain numerically the structure of the DOE by using a phase retrieval technique. 65

4.9 Volume of recorded propagation during five plane $(z_0, ..., z_4)$ by implementing the phase retrieval technique. 66

4.10 Periodic pattern in our DOE which is obtained based on a phase retrieval technique. The color bar axis is for intensity in arbitrary unit. 66

4.11 Numerical setup with an array of emitters in plane y. All components and their parameters are the same as in the experiment of Fig. 3.10. MO_1 has a NA $= 0.25$ and $f_1 = 18$ mm, MO_2 has a NA $= 0.1$ and a $f_1 = 45$ mm. The periodic DOE is located between the two MOs. 67

4.12 Distances between diffractive orders $|d^{\mathrm{diff}}|$ versus the emitter positions $|y|$. Blue rectangles are numerical data, the black dashed line is the distance for coupling a periodic array. 68

4.13 Scheme of mismatch distance $|\Delta|$ versus emitter positions $|y|$ in a double-logarithmic scale. Blue stars are $|\Delta|$ in numerical simulation for array of emitter positions $|y|$. Black dashed line is analytical mismatch for array $|y|$. Red stars are $|\Delta|$ in experiment for array of emitter positions y.The good agreement between the red and the blue data. 69

4.14 Width of all diffractive orders in plane y' versus emitter positions $|y|$. Blue circles are width of diffractive orders in numerical simulation for array of emitter positions $|y|$. Red circles are width of diffractive orders in experiment for array of emitter positions y. Black dashed line is the width in analytical simulation for array $|y|$. 70

4.15 Numerical setup with an array of emitters in plane y. This setup has the low magnifications in MOs and high NAs. MO_1 has high $NA = 0.45$ in focal length $f_1 = 18$ mm. MO_2 has a $NA = 0.2$ in focal length $f_2 = 45$ mm. The periodic DOE is located between both MOs. 70

4.16 Distance between diffractive orders $|d^{\mathrm{diff}}|$ versus emitter positions $|y|$. Blue rectangles: distribution of the distance in numerical simulation. The black dashed line: the constant distance of the alignment condition in a periodic array. 71

4.17 Mismatch $|\Delta|$ versus emitter positions $|y|$ in a double-logarithmic scale. Blue stars are distribution of numerical mismatch after fit. Black line is analytical mismatch. 72

4.18 Width for all diffractive orders in $\frac{1}{e^2}$ versus emitter positions $|y|$. Blue stars are distribution of numerical width after fit in array of emitter positions $|y|$. Black dashed line is diffraction limit. 72

4.19 Mismatch $|\Delta|$ versus emitter positions $|y|$ in a double-logarithmic scale. Blue stars are the distribution of numerical mismatch with non-paraxial rescaling of the image planes. Brown stars are distribution of numerical mismatch with paraxial calculation in the image planes. Black line is analytical mismatch. 73

5.1 Schematic illustration of a recurrent neural network with the input (green arrow), which is coupled to the nonlinear nodes. The recurrent connections of nodes are illustrated by blue arrows. The system's output as created by the weighting the network states with W^{DMD}. 76

5.2 Scheme of a recurrent neural network. The optical nodes are establishes by pixels of a SLM. Nonlinear nodes are created by a PBS which is a polarization filtering. Coupling between the nodes is fulfilled by a DOE. Before detection at the system's output, there is a digital micro-mirror device, which creates a spatially modulated image of the SLMs state. 77

5.3 Coupling matrix in our RNN for 900 nodes where black presents high coupling strength and white presents absent of coupling. A zoom for a smaller region shows better this coupling strength. 78

5.4 Reservoir output in μm in blue line and prediction target in red dots. Red dots and blue line can hardly be diffracted. Prediction error is in yellow line. 79

Document réalisé avec LaTeX et :
le style LaTeX pour Thèse de Doctorat créé par S. Galland — http://www.multiagent.fr/ThesisStyle
la collection de paquets `tex-upmethodology` — http://www.arakhne.org/tex-upmethodology/