MASTER THESIS

# Inverse Design of an Achromatic Metalens

**Author**: Daniele De Gregorio
**Supervisor**: Univ.-Prof. Dr. Stefan Maier
**Co-Supervisor**: Dr. Haoran Ren

Chair in Hybrid Nanosystems
Faculty of Physics
Ludwig Maximilian University of Munich

30 June 2021

MASTERARBEIT

# Inverses Design einer achromatischen Metalinse

**Autor**: Daniele De Gregorio
**Betreuer**: Univ.-Prof. Dr. Stefan Maier
**Ko-Betreuer**: Dr. Haoran Ren

Lehrstuhl für Hybride Nanosysteme
Fakultät für Physik
Ludwig-Maximilian-Universität München

30. Juni 2021

# Acknowledgements

First and foremost, I would like to thank Professor Stefan Maier who offered me the opportunity to work alongside him and his team in the beautiful facilities of the Nanoinsitute Munich next to the English garden. I am especially grateful to you, Stefan, that you allowed me to start my studies at the institute as a working student before beginning with the master's project.

Then, I would like to acknowledge my colleagues and friends from the cleanroom: Andreas, Benni, Julius, Juan, Lin, Lucca, Ludwig, Thomas and Namvar. Working next to you in a lab with all this advanced and expensive machinery was a privilege and an experience I will not forget soon. Special thanks go to Lucca, Ludwig, Namvar and Thomas, who introduced me excellently to the machines and never showed a sign of fatigue for my endless questions.

I would also like to thank the members of the chiral team: Andres, Emi, Fedja, Johannes, Leo, Stefan and Seunghoon. The discussions we had about chiral metamaterials and chiral mirrors were very stimulating and insightful. Getting to know how experienced scientists tackle problems was inspiring indeed. Regarding the topic of chirality, I would like to acknowledge Ilka, from the Photonics and Optoelectronics Group, for a wonderful collaboration on chiral metamaterials, which was the first research project I had at the institute.

Concerning the project of this thesis, I owe great debts of gratitude to Dr. Juan Durillo from the Leibniz Supercomputing Centre (LRZ) in Munich, who supported me with the training of the neural networks on the LRZ servers. In connection with this, I gratefully acknowledge the computational and data resources provided by the LRZ.

In addition, I would like to thank my supervisor, Dr. Haoran Ren, whose expertise was instrumental in coming up with the idea of incorporating machine learning into the design process of metalenses. Thank you, Haoran, for your wonderful guidance throughout this master's project. I wish you all the best in your new position at the Macquarie University in Australia and I sincerely hope to see you again one day.

I also want to thank my friends Jonas, Emma and Ian, who generously spent their free time proofreading this thesis. In the end, I express my deep gratitude towards my family for supporting me emotionally during the writing phase of this work.

# Abstract

In recent years rapid progress has been made towards expanding the operational bandwidth of metalenses from discrete wavelengths to broadband ranges. Current solutions to designing broadband achromatic metalenses involve searching through extensive libraries of meta-atom geometries and their corresponding wavelength-dependent phases and group delays to find the designs that closest match the desired optical response. The purpose of this work is to examine whether the appropriate meta-atoms can be more efficiently selected than conventional brute-force techniques with the assistance of machine learning. Moreover, it is investigated if machine learning algorithms could use the design libraries to learn the complex nanostructure-performance relationship and potentially enable the on-demand design of metalenses with arbitrary numerical apertures and operational bandwidths.

Among the machine learning algorithms studied, decision trees proved most suitable for solving the inverse design problem. Compared to neural networks, they were easier to implement, faster to train and could more accurately map the employed nanopillars to their optical responses. Nevertheless, decision trees were not able to predict and extrapolate to metalenses outside of their training sets. Hence, to construct four achromatic metalenses with various numerical apertures and different bandwidths, a custom-made design library had to be prepared for each metalens. Once the decision trees were trained, they produced achromatic metalenses that closely fit the predefined target phase profiles. Remarkably, the trees were able to accelerate the design process by predicting the appropriate metalenses up to 200 times faster than conventional brute-force optimization algorithms.

# Contents

# Chapter 1

# Introduction

## 1.1 Conventional refractive lenses

From unlocking the vast cosmos with Galileo Galilei observing the moons of Jupiter to the microscale with van Leeuwenhoek inspecting and experimenting with microbes for the first time – lenses have played a critical part in the progress of scientific understanding.

Conventional optical lenses form the basis for technologies such as cameras, microscopes, telescopes and optical lithography. As single refractive lenses do not form perfect images, an important challenge in designing optical lenses is the correction of aberrations. Aberrations cause the light to spread out over some region of space rather than directing it towards a focal point and can be both monochromatic and chromatic [1]. Chromatic aberrations originate from the wavelength-dependent refractive index of materials and can be compensated by combining multiple lenses with complementary dispersion characteristics. However, this causes conventional lenses to become bulkier and more expensive, presenting a significant limitation as technology advances and demands more and more compact optical components.

## 1.2 Metalenses and achromatic metalenses

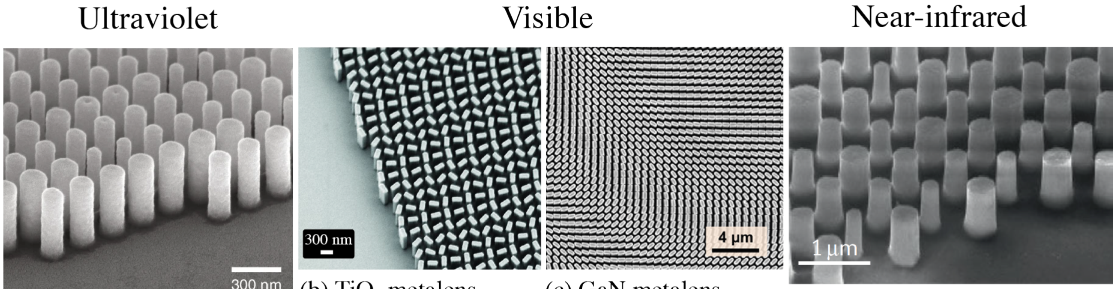
Metasurfaces, two-dimensional arrays of subwavelength antennas, could be used to deal with this challenge. By engineering the physical shape of the antennas, metasurfaces enable the control over basic properties of light such as the phase, amplitude and polarization. Several components have been developed based on metasurfaces such as holograms [2], quarter-wave plates [3], carpet cloaks [4], vortex plates [5], thin absorbers [6], or sensors [7]. One device that particularly has received wide interest is the metasurface lens or metalens. Compared to the traditional bulky lens, metalenses are ultrathin, thinner than a sheet of paper, and can be mass-produced using microfabrication techniques similar to those used in the semiconductor industry [8]. Metalenses are usually designed based on nanostructures that have a high index of refraction and are lossless in the wavelength regime of interest [9]. High-index materials are selected because they provide better light confinement which improves the overall efficiency and facilitates the fabricability of the metalens.

Following these two material properties, several chromatic metalenses across various regions of the electromagnetic spectrum have been demonstrated [9]. In recent work, Zhang et al. [10] used $HfO_2$, a UV-transparent high-refractive-index dielectric material, to implement metalenses operating in the UV at 325 nm and 363 nm with focusing efficiencies[1] of 55% and 56%, respectively (Fig.1.1a). For metalenses in the visible spectrum, typical materials include $TiO_2$ [11, 12], GaN [13, 14] and single-crystal silicon [15, 16, 17]. $TiO_2$ metalenses that achieved diffraction-limited focusing at wavelengths of 405, 532, and 660 nm with corresponding efficiencies of 86%, 73%, and 66% have been demonstrated (Fig.1.1b) [18]. In [14], the authors reported three GaN-based metalenses with operation efficiencies of 87%, 91.6%, and 50.6% for blue (430 nm), green (532 nm) and red (633 nm) light, respectively (Fig.1.1c). In another attempt [19], a single-crystal silicon metalens was designed for 532 nm with an efficiency of 67% (Fig.1.1d). Notably, Liang et al. preferred single-crystal silicon over amorphous silicon due to its lower optical loss at visible wavelengths. For near-infrared wavelengths, however, amorphous silicon becomes the material of choice. A Huygens nanoantenna-based metalens was presented at a wavelength of 825 nm that reached an efficiency of 20% (Fig. 1.1e) [20]. Moreover, a polarization-insensitive metalens composed of silicon nano-posts with a focusing efficiency of up to 82% at 1550 nm was introduced [21]. Metalenses have also been extended to the mid-infrared region with Si [22, 23, 24] and PbTe [25, 26] as primary materials.

The aforementioned metalenses operate at discrete wavelengths only and consequently are not yet suitable for full-colour imaging systems such as photography or microscopy. In recent years there has been growing interest to expand the operational bandwiths of

---

[1] The focusing efficiency was defined as the ratio of the optical power of the focused spot to the total power illuminating the metalens.

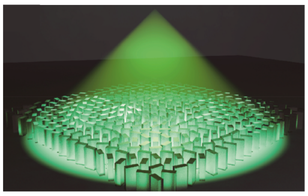


(a) $HfO_2$ metalens

(b) $TiO_2$ metalens

(c) GaN metalens

(e) α-Si metalens



(d) Single-crystal Si metalens

Figure 1.1: **Chromatic metalenses across different parts of the electromagnetic spectrum.** (a) A $HfO_2$ metalens operating in the ultraviolet [10]. (b)-(d) Metalenses working at visible wavelengths are typically made out of (b) TiO2 [18], (c) GaN [14] and (d) single-crystal silicon [19]. (e) A Si-based metalens adapted for the NIR [20].

metalenses from discrete to broadband ranges. The first systematic studies on metalenses with a continuous wavelength range were conducted in 2015 by Wang et al. [27], who demonstrated a cylindrical lens with achromatic focusing in the entire visible band. However, their lens had a low numerical aperture (NA) of 0.013 and was constrained to the photoresists used in grayscale lithography as constituent materials. Subsequently, in 2016 Khorasaninejad et al. [28] reported an achromatic metalens consisting of titanium dioxide nanopillars on a dielectric spacer layer atop a metallic mirror (Fig. 1.2a). Their metalens worked in reflection mode and achieved achromatic focusing over a bandwidth of 60 nm in the visible with NA = 0.2. Despite successfully suppressing the chromatic aberration over 60 nm, much work remained to be done to extend the working bandwidth of the metalens for practical purposes. In [29], the authors experimentally realized a broadband achromatic metalens over a continuous wavelength region from 1200 to 1680 nm by employing metallic integrated-resonant unit elements (IRUs) to control the dispersion (Fig. 1.2b).

However, these metalenses work in a reflection scheme. Various approaches have been put forward to design transmissive metalenses, which are much more desirable for practical applications. In a follow-up study of [29], Wang et al. introduced a design comprising of GaN-based IRUs (Fig. 1.3a) to construct an achromatic metalens that almost cov-

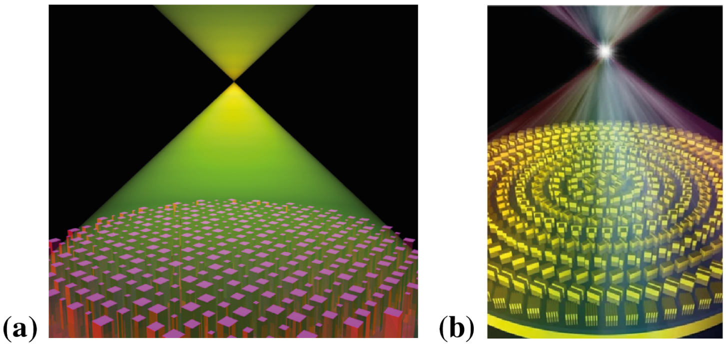


Figure 1.2: **Broadband polarization-sensitive achromatic metalenses operating in reflection mode.** (a) Schematic of a nanopillar-based achromatic metalens that focuses a collimated light beam into a spot [28]. (b) Schematic of an IRUs-based achromatic metalens operating in the NIR [29].

ered the entire visible region ($400 - 660\,\text{nm}$) in transmission mode [30]. The fabricated metalens reached an average efficiency of about 40% with NA = 0.106. To satisfy the phase compensation required for achromatic focusing, the authors utilized waveguide-like resonant modes in the GaN nanostructures. In a different approach, Chen et al. corrected chromatic aberrations by simultaneously controlling the phase, group delay and group delay dispersion (see section 2.2 for a more detailed explanation) of light by the judicious design of nanofins on a surface (Fig. 1.3b) [31]. This is distinct from the strategy mentioned previously in [29], where the authors employed plasmonic resonances and neglected the group delay dispersion. The proposed metalens (NA = 0.2) operated between $470 - 670\,\text{nm}$ with an efficiency of 20% at 500 nm.

The transmissive metalenses introduced so far are limited to incident light of one particular polarization. Initial studies on broadband polarization-insensitive achromatic metalenses were carried out in [32]. The authors prepared a library of isotropic meta-atoms with complex cross-sectional geometries to satisfy the phase requirements at all design wavelengths - a crucial step for realizing achromatic metalenses across wide bandwidths. Notably, by introducing meta-unit geometries with four-fold rather than just rotational symmetry [33, 34, 28], Shrestha et al. could enhance the degrees of freedom for fine-tuning the dispersion without sacrificing the polarization-independent performance. The experimentally demonstrated metalenses reached polarization-independent focusing efficiencies up to 50% with a near-constant focal length over $1200 - 1600\,\text{nm}$ (Fig. 1.4a). While employing isotropic nanostructures circumvents the polarization sensitivity in metalenses, an essential degree of freedom for engineering the dispersion is lost. In a

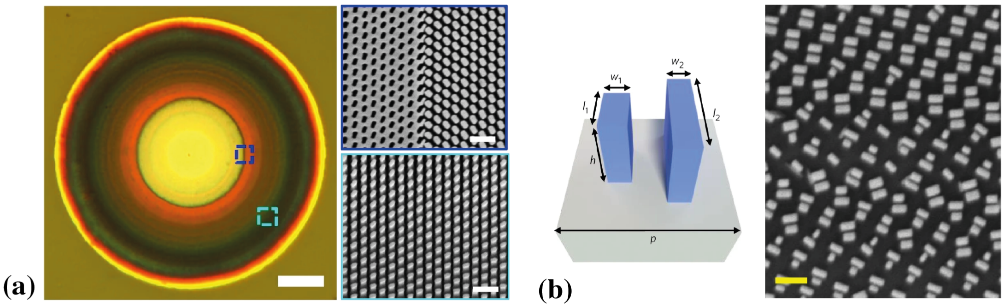


Figure 1.3: **Broadband polarization-sensitive achromatic metalenses operating in transmission mode.** (a) Optical image of the transmissive achromatic metalens. The SEM images show the GaN-based IRUs at different positions of the metalens (indicated by the dashed square in the optical image) [30]. (b) Schematic of the $TiO_2$-based metalens building block and SEM image of the fabricated metalens at a specific region [31].

subsequent study, Chen et al. [35] showed that simultaneous polarization-independent focusing was indeed possible using anisotropic nanostructures. They overcame the unwanted polarization sensitivity by limiting the rotation angle of their anisotropic nanofins to 0 and 90 degrees (Fig. 1.4b). The reported polarization-insensitive metalens operated at wavelengths ranging from $460 - 700\,\text{nm}$ with an efficiency of 35%.

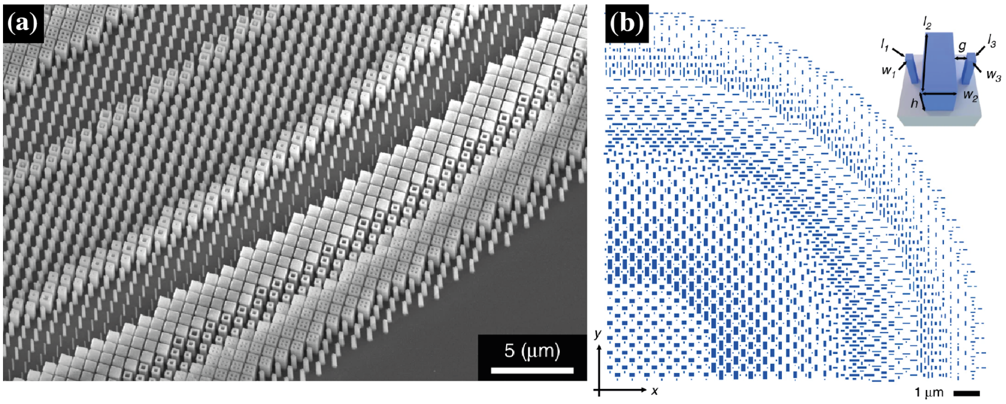


Figure 1.4: **Broadband polarization-insensitive achromatic metalenses.** (a) SEM image of the fabricated metalens using meta-unit geometries with four-fold symmetry [32]. (b) Layout of one quadrant of the metalens using anisotropic nanofins as building blocks [35].

The most recent developments at the time of writing have led towards achromatic metalenses with higher focusing efficiencies and broader working bandwidths. In [36], an achromatic polarization-insensitive metalens with average efficiencies of over 70% in the continuous band from 640 – 1200 nm was reported (Fig. 1.5a) - thereby showing a significant improvement over the aforementioned state-of-the-art metalenses. Achromatic performance crucially depends on the achieved group delays by the metalens nanostructures. Previous efforts have focused on two-dimensional meta-atoms with exotic shapes and high aspect ratios to sufficiently control the dispersion. By contrast, in [8], the authors relaxed the height constraint adopted by most metalenses and gained an additional degree of freedom by accessing 3D geometries using multiphoton instead of grayscale lithography and etching processes (Fig. 1.5b). The proposed metalenses exhibited polarization-insensitive achromatic focusing over the broad spectral region from 1000 to 1800 nm with average efficiencies greater than 60% and maximum efficiencies of up to 80%. Finally, to further illustrate the benefits of embracing true 3D geometries for planar optics, Balli et al. designed and fabricated achromatic metalenses operating from the visible to the short-wave infrared, 450 – 1700 nm, with average focusing efficiencies of 60% [37]. By replacing the nanopillars from their former study [8] with a novel nanohole structure, they could reduce fabrication difficulties and thereby extend the achromatic correction towards the visible spectrum.

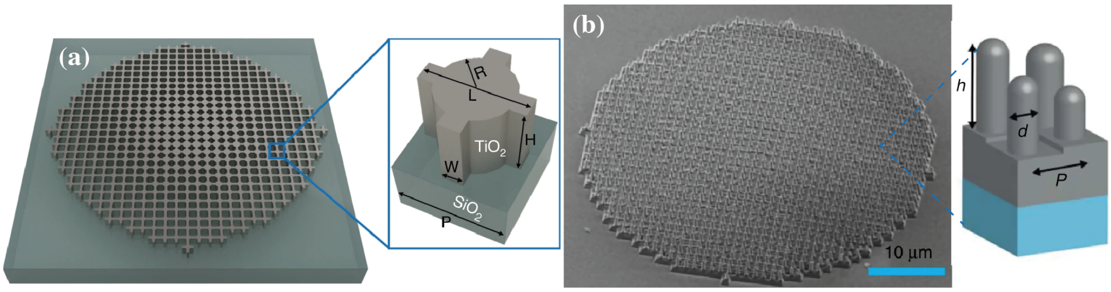


Figure 1.5: **Current state-of-the-art achromatic metalenses.** (a) Schematic of the fishnet-type achromatic metalens and its corresponding unit cell [36]. (b) SEM image of the proposed hybrid achromatic metalens and a schematic showing the meta-atom [8].

## 1.3 Machine learning and its applications in photonics

Similar to how advancements in nanofabrication techniques have enabled the realization of metalenses, affordable data storage and cheaper and more powerful computers have boosted growth in artificial intelligence (AI) and its subset machine learning. Games provide a good measure to grasp the rapid pace of progress in AI as they offer a well-defined problem with a certain level of difficulty. In 1994 ”Chinook” became the first computer program to win a world championship in checkers [38]. Three years later, IBM developed ”Deep Blue” to beat the former world chess champion, Garry Kasparov, in chess [39]. The trend of AI to tackle games and problems with increasing degrees of freedom continued, reaching its pinnacle year in 2016 when Google Deepmind introduced ”AlphaGo”, the first computer program to defeat a professional human Go player in Go, a board game that was known to be the most challenging classical game for AI because of its complexity [40]. Moving away from the environment of games, AI is gaining increasing relevance in addressing scientific questions too. Just recently, in 2020, the AlphaFold project at Google Deepmind has solved a 50-year-old challenge in biology by demonstrating accurate predictions of how proteins fold based on their genetic sequence alone [41]. This is a technological breakthrough and shows the potential AI can have, to accelerate scientific progress and deepen our understanding of the universe.

Very recently, AI has also entered the field of nanophotonics. There, forward modeling networks have been trained to predict the optical properties of nanophotonic structures, by effectively approximating Maxwell’s equations instead of solving them as conventional numerical methods do. As a natural extension to the forward modeling of nanophotonics, machine learning has also been introduced to solve complex inverse design problems. In contrast to forward modeling networks, inverse design networks receive as input the desired optical property and output the corresponding nanophotonic device. Lately, machine learning-assisted inverse design has gained increasing attention as a powerful method to design unintuitive photonic devices with unprecedented performances.

### 1.3.1 Machine learning for forward nanophotonic modelling

The conventional approach for describing the light propagation in nanophotonic systems relies on solving Maxwell’s equations either analytically [42] or numerically [43] [44]. Analytical solutions to electromagnetic waves passing through stratified media can be obtained by the transfer matrix method [45]. Semi-analytical methods such as the rigorous coupled-wave analysis (RCWA) solve Maxwell’s equations in Fourier space and are well suited for approximating the light scattering from periodically structured surfaces [46]. However, for photonic systems of higher complexity (e.g. three-dimensional structures), more sophisticated approaches such as the finite-element (FET) and finite-difference time-domain (FDTD) method have to be utilized to get an accurate representation of the physical reality of the system [44]. Both of these methods are based on subdividing the considered objects into fine meshes and then solving Maxwell's equations at each spatial location [44]. Although these approaches can handle a variety of problems in nanophotonics, they demand more and more computational resources as increasingly complex structures are investigated [43].

In recent years, data-driven forward modeling networks have attracted widespread interest as they have been shown to predict the optical properties of nanophotonic devices much faster than traditional full-wave simulations [47]. The training of these networks initially demands a training set, which is provided by electromagnetic simulations. Although preparing the training set is computationally expensive, it has to be emphasized that this is only a one-time cost [48]. Once the networks have successfully learned to approximate Maxwell's equations, they can be applied, for example, to accelerate conventional inverse design techniques in nanophotonics [49].

In one of the initial studies on data-driven forward modeling, Peurifoy et al. [47] presented an artificial neural network (ANN) that could accurately predict the scattering cross-section spectra of multilayer nanoparticles. The nanoparticles under study were composed of silica cores with eight alternating $TiO_2$ and silica shells, where each shell had a thickness between $30 - 70\,\mathrm{nm}$. The training set for the proposed ANN consisted of 50,000 nanoparticles and their associated scattering responses, which were analytically obtained via the transfer matrix method. The authors used a fully connected network that received as input the thickness of each nanoparticle shell and returned as output the scattering cross-section at different wavelengths of the scattering spectrum (Fig. 1.6a) [47]. Once the training was complete, the ANN's performance was validated on a multilayer nanoparticle it had not encountered previously. Strikingly, the approximated spectra agreed well with the target response from the simulation, even though the ANN was trained on only 50,000 nanoparticles. Furthermore, by comparing the predicted response of the target nanoparticle with the closest examples from the training set, the authors showed that the trained model did not simply interpolate between the data points but instead learned some of the underlying physics of the system (Fig. 1.6b).

To learn the nonlinear mapping between the input and the output data, machine learning algorithms rely on a vast amount of labelled data points that reasonably capture the complexity of the problem. Preparing a sufficiently big training set, however, is time-consuming and requires computationally expensive numerical simulations. As a possible solution to reduce the burden of data collection, Qu et al. [50] applied transfer learning to improve the predictive accuracy of their ANNs trained on a small training set.

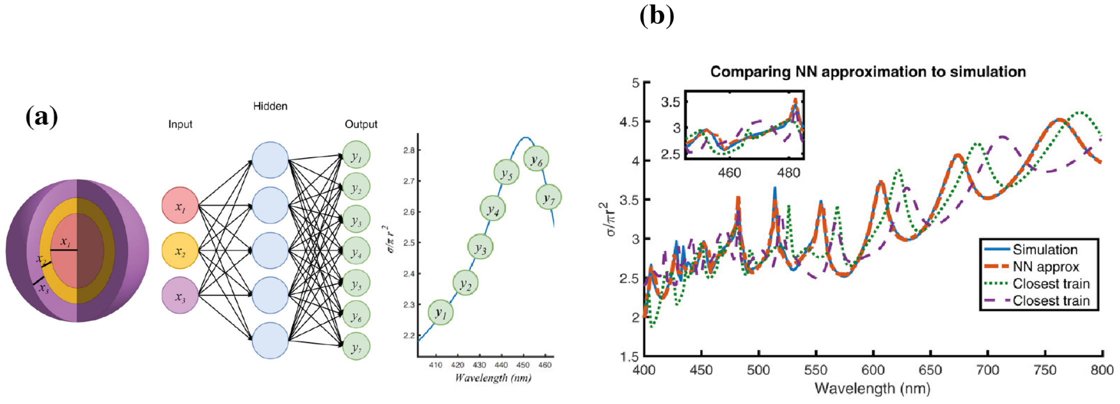


Figure 1.6: **Neural network-assisted forward modeling of multilayer nanoparticles.** (a) Fully connected neural network receives as input the thickness of each nanoparticle shell and predicts as output the scattering cross section at different wavelengths between 400 and 800 nm. (b) Spectrum approximation of the neural network. The prediction of the ANN is compared to the real spectrum and the closest spectra from the training set [47].

Transfer learning is based on utilizing the knowledge obtained from solving one source task and then applying it to a different but related target task (Fig. 1.7a) [50]. It proves particularly useful for scenarios in which the source data is large and inexpensive, and the target data is small and expensive. The authors observed that they could reduce the predictive error rate by 50.5% when migrating the knowledge obtained from a 10-layer multilayer film to predict the transmission from an 8-layer film (Fig. 1.7b). Remarkably, transfer learning could also be applied between physically very different scenarios such as the scattering from core-shell nanoparticles and transmission from multilayer films. The error rate of predicting the transmissivity from the 8-layer films was significantly reduced by 19.7% by reusing the physical rules extracted from the 8-layer nanoparticle scattering problem. The success of this approach has demonstrated how neural networks could be combined with transfer learning when the training set is small or difficult to acquire.

The previously introduced papers approximated the optical properties of simple shapes with limited geometric degrees of freedom. Inampudi et al. extended the suitability for machine learning-enabled forward modeling by predicting the response of devices with arbitrary unit shapes [51]. More specifically, they trained a neural network to predict the diffraction efficiency (DE) of 13 diffraction orders from an arbitrarily shaped metagrating atop a reflective surface. The unit cell of the periodic structure was given by a sixteen-

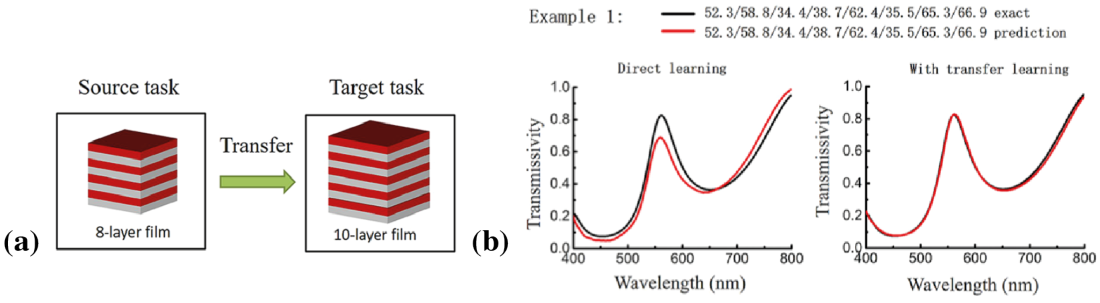


Figure 1.7: **Migrating knowledge between physically similar systems.** (a) Knowledge is transferred from 8-layer to 10-layer films. (b) The predicted transmission spectrum with and without transfer learning. The exact spectra are shown in black [50].

sided polygon, whose shape was determined by the radius coordinates of the sixteen vertices (Fig. 1.8a). The proposed ANN architecture consisted of three layers in total and was trained on a data set containing the diffraction efficiencies of 90,000 different unit cells. The preparation of the training set was carried out using an RCWA solver and took around 50 hours to complete. After the training was finished, the network could near-instantaneously identify the trends of the DEs for previously unseen geometries (Fig. 1.8b).

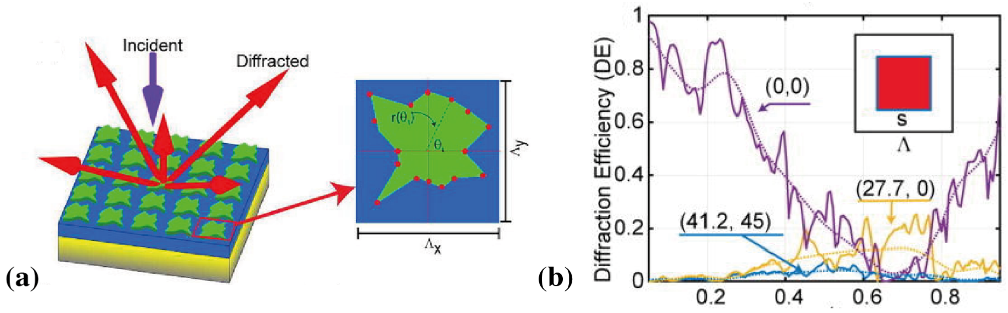


Figure 1.8: **Forward modeling of devices with arbitrary shapes.** (a) Schematic of the metagrating based on a 16-sided polygon unit cell. (b) Comparison of the predicted (dotted line) and simulated DEs (solid line) of three diffraction orders along the $(\theta, \phi)$ direction for square-shaped unit cells [51].

In a similar attempt to extend the accurate forward-modelling to more complicated shapes, Sajedian et al. proposed a model composed of both a convolutional and a recurrent neural network to find the absorption spectra of plasmonic structures based on 2D images [52]. The ResNet-type convolutional neural network [53] analyzed the input images and passed critical spatial features to the recurrent neural network, which was taught to find the correlation between the extracted features and the absorption spectra. The combined networks were trained on a data set consisting of 100,000 random structures connected to their associated responses via the commercial FDTD simulation package Lumerical. Once the model was trained, it could almost perfectly reproduce the 10,000 absorption spectra from previously unseen structures within a fraction of a second (Fig. 1.9).

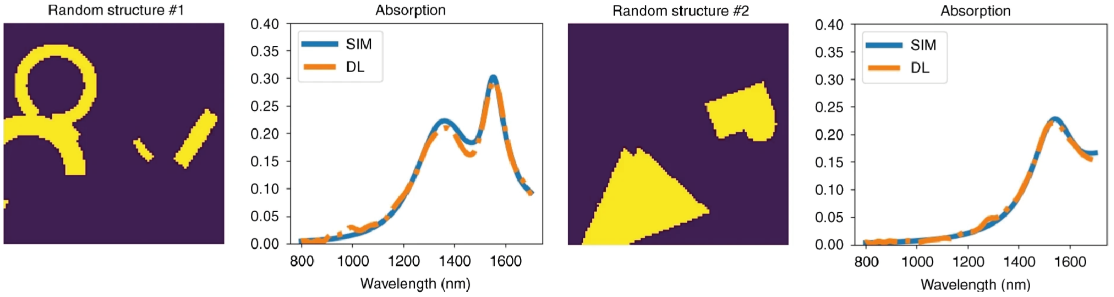


Figure 1.9: **Results of applying the trained forward-modeling network on previously unseen structures.** The absorption spectra obtained from the simulation (solid blue line) are compared to the predictions of the deep learning model (dotted orange line). The trained model reproduces the absorption spectra of almost all 10,000 structures perfectly [53].

### 1.3.2 Machine learning nanophotonic inverse design

Most of the early breakthroughs in nanophotonics have been achieved with designs based on scientific intuition, experience obtained from prior practice and knowledge acquired from simplified analytical models. For example, it is commonly known that polarization sensitivity can be introduced in symmetrical nanoparticles by elongating one of the sides or that ring-like structures support strong magnetic resonances when the incident magnetic field is perpendicular to the ring. Apart from structures obtained from physical intuition, several photonic designs have been derived from analytical models. For instance, the optical properties of spherically shaped plasmonic or dielectric nanoparticles can be accurately described by Mie theory [54]. Following these physics-inspired guidelines, early nanophotonics research has achieved breakthrough performances with devices such as the split ring resonator [55], gammadions [56], bowtie- [57] and V-shaped antennas [58] to name just a few.

For simple photonic structures with limited degrees of freedom, the designs with the desired performance can often be found by analytical solutions or by conducting various parameter sweeps using numerical methods for iteratively monitoring the optical response. However, as advanced nanofabrication techniques allow for increasingly complex structures, the parameter space grows and with it the challenge of finding the optimal design through trial and error increases. For this purpose, various algorithms have been introduced to assist in the design process. Common approaches include evolutionary methods such as genetic algorithms [59] and particle swarm optimization [60], and gradient-based methods such as topology optimization [61]. Compared to conventional physics-based approaches, these techniques open up a much broader parameter space which facilitates the search for non-intuitive photonic structures with unprecedented performances. However, as the optimization process often involves hundreds or thousands of electromagnetic simulations to approach the targeted response, design procedures based on optimization algorithms are inevitably slow and computationally expensive. Moreover, in the case that the optimization goal has to be adjusted during the design process (e.g. because of added fabrication constraints), the whole optimization has to be carried out once again [44]. Machine learning-assisted inverse design seeks to overcome the drawbacks of conventional optimization methods. Unlike optimization-based approaches, data-driven methods can extract useful information from large datasets. Once an inverse design network has learned the complex relationship between the input and the output data, it can predict new structures almost instantaneously.

The comparatively higher optimization speed in data-driven methods over numerical inverse design algorithms was highlighted by Peurifoy and his group. After they successfully trained their ANN to approximate the light scattering from multilayer nanoparticles (as discussed in subsection 1.3.1), the authors attempted to solve the inverse design problem by running the ANN backwards [47]. More specifically, they fixed the weights of the ANN and set the output to the desired spectrum. Subsequently, they used backpropagation to iterate through the input geometries that most closely reproduced the result. Fig. 1.10a shows that the nanoparticle predicted by the ANN approached the desired spectrum much closer than the interior-point method, i.e. the most effective optimization algorithm they could find for this type of problem. Comparing the optimization runtime of both approaches revealed that the ANN completed the inverse design two orders of magnitude faster than the algorithm. Notably, the runtime in the ANN increased linearly with the design complexity (i.e. the number of particle shells), whereas for the numerical method, it increased polynomially (Fig. 1.10b). Lastly, the authors further demonstrated the suitability of their ANN as an optimization tool by realizing nanoparticles that satisfied distinct design requirements, such as high scattering for single and broadband wavelengths [47].

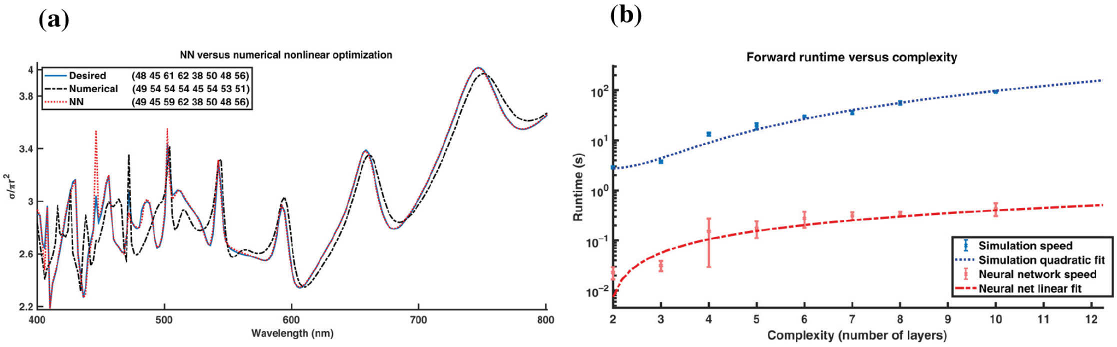


Figure 1.10: **Data-driven inverse design of multilayer nanoparticles.** (a) Inverse design of an eight-shell nanoparticle. The solutions of the ANN and the numerical optimization are compared to the desired spectrum in blue. The legend gives the thickness of each particle shell. (b) Runtime comparison between the ANN and the interior-point method for different design complexities [47].

Artificial neural networks have also been applied to study the transmission of multilayer thin films [48]. More specifically, Liu et al. attempted to optimize the thickness of each of the 16 alternating layers of $SiO_2$ and $Si_3N_4$ in their thin film (Fig. 1.11a) to achieve specific target transmission spectra. As datasets grow to address increasingly complicated systems, ANNs often fail to converge due to the existence of nonunique training instances (i.e. vastly different devices can produce similar optical responses). To overcome this fundamental issue, the authors proposed a tandem architecture consisting of an inverse design network connected to the input of a pretrained forward modeling network (Fig 1.11b). The forward modeling network was trained on 500,000 labelled data pairs to map the designs to the transmission spectra, and its weights were later fixed in the tandem configuration. The tandem network solved the data inconsistency problem as the inverse network was not required to reproduce the real designs from the training set, but instead was forced to converge to devices that minimized the error between the predicted and target response. The example test result in Figure 1.11c demonstrates that the spectrum of the designed multilayer film showed good agreements with the desired response.

In a very similar approach to overcome the non-uniqueness problem, Malkiel and colleagues introduced a deep neural network that combined a geometry-predicting-network (GPN) and a spectrum-predicting-network (SPN) to function both as a fast approximator and an inverse design tool for ”H”-shaped gold nanoparticles [62].

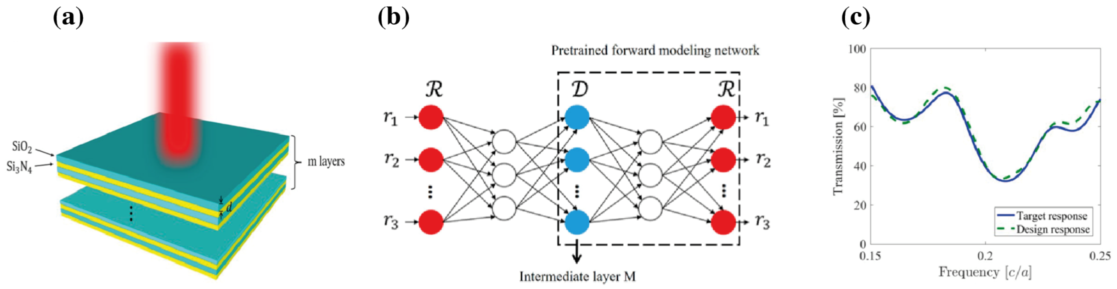


Figure 1.11: **Inverse design of multilayer thin films.** (a) The multilayer film consists of alternating layers of $SiO_2$ and $Si_3N_4$. (b) The tandem network architecture comprises a forward modeling network and an inverse design network. (c) Comparison of the target and test response from the tandem network [48].

In the meantime, Ma and coworkers have reported a model comprising two bidirectional neural networks, denoted as primary and auxiliary network, to achieve the on-demand design of three-dimensional chiral metamaterials with strong chiroptical responses [63]. The chiral metamaterial under consideration consisted of stacked gold split-ring resonators (SRRs) twisted at an angle with a gold reflector at the bottom (Fig. 1.12a). The meta-atom was fully-determined by the following 5 parameters: the top and bottom SSR size $l_1$ and $l_2$, the top and bottom spacer thickness $t_1$ and $t_2$ and the twisting angle $\alpha$. During the training, the primary network learned the intricate relationship between the optical spectra and the metamaterial design parameters, whereas the auxiliary network associated the design with the circular dichroism (CD) response. By connecting the individual networks through an ensembled learning strategy, both the forward prediction and the inverse retrieval of the designs improved significantly. As shown in Fig. 1.12b., the combined network was able to retrieve the design parameters that best fit the required CD spectrum. Interestingly, the model suggested a near-symmetric structure, which intuitively should not yield strong chiroptical behaviour.

The machine learning architectures mentioned above fall into the category of discriminative models. Such models are trained on labelled data that captures the complexity of the problem to infer the relationship between the designs and the associated optical properties. However, as systems with thousands or more degrees of freedom are considered, it becomes infeasible to prepare enough data for training. What is more, once a model is trained, it only provides a single design for a given spectrum, whereby in reality many

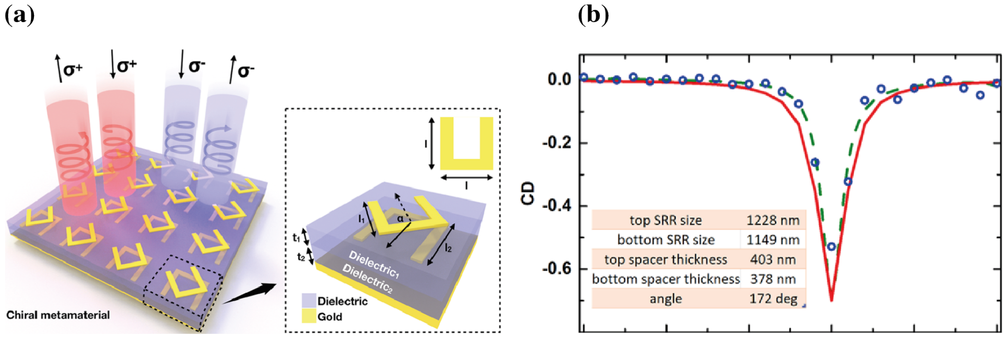


Figure 1.12: **Inverse design of 3D chiral metamaterials.** (a) Schematic of the designed chiral unit cell. The inset shows the five design parameters of the single meta-atom. (b) Desired (red solid line), predicted (blue circles) and simulated (green dashed line) CD spectra. The inset lists the parameters of the designed chiral meta-atom [63].

different solutions with similar responses would exist [64]. Recently, generative models such as generative adversarial networks (GANs) and variational autoencoders (VAEs) have been utilized to reduce the dimensionality of the design space and to enable the one-to-many mapping of photonic structures [65].

Initial studies for incorporating GANs into the inverse design of metasurfaces were conducted by Liu et al. in 2018 [64]. They proposed an architecture consisting of three convolutional networks: a generator, a critic and a pretrained simulator. The overall objective of the generator was to produce realistic structures whose transmittance spectra resembled the user-defined input spectra. The critic served as feedback to the generator to produce designs similar to the input geometries. To carry out fast simulations of the generated patterns, the authors trained a simulator on a dataset containing 6,500 full-wave FEM simulations of metasurfaces with diverse shapes. After the training, the generator illustrated its competence to perform inverse design by producing a nanostructure that fitted the randomly generated Gaussian-like test spectrum (Fig. 1.13).

In [66], a variational autoencoder (VAE), another type of generative model, was utilized to solve the one-to-many mapping issue commonly present in several supervised learning schemes. The proposed encoder-decoder network compressed the input structural designs with their optical responses into a space of reduced dimensionality termed latent space. From the latent space, the latent variables were stochastically sampled. By feeding the latent variables and required optical responses into the decoder network, the authors achieved inverse retrieval of meta-atom geometries with desired properties.

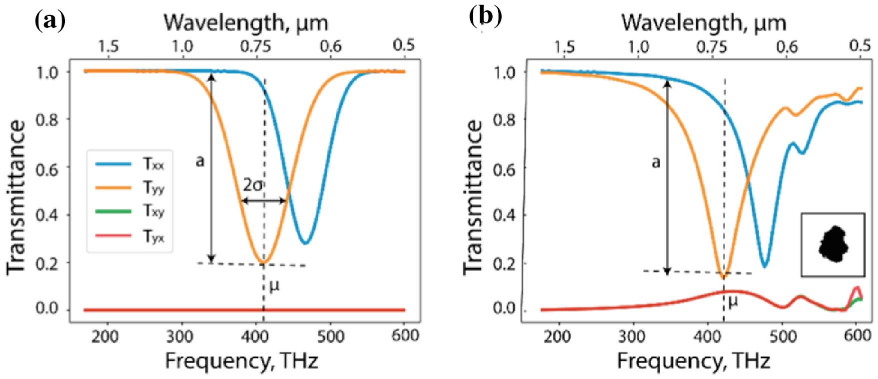


Figure 1.13: **GAN-based inverse design of metasurfaces.** (a) Desired transmission coefficients as the input to the trained generator. (b) Transmittance spectrum of the generated unit cell structure [64].

Remarkably, due to the flexibility of the sampling process, the group demonstrated structures with diverse geometries but similar optical properties (Fig. 3.3) and thus overcame the one-to-many mapping problem in nanophotonic inverse design.

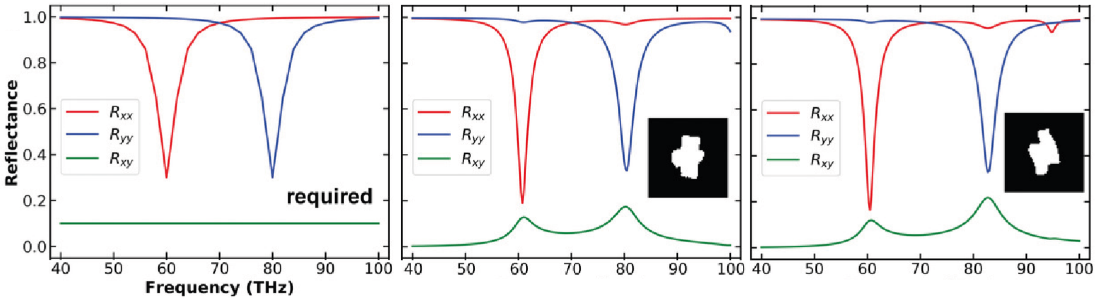


Figure 1.14: **Overcoming the one-to-many mapping problem through VAE-mediated inverse design.** Trained VAE enables the design of multiple geometries for a given reflection spectrum [66].

Lastly, Jiaqi Jiang and Jonathan Fan have incorporated adjoint variable calculations into a GAN for designing metagratings operating across a wide range of wavelengths and deflection angles [67]. The hybridization of the data-based and optimization-based approach produced not only devices with higher deflection efficiencies than conventional adjoint-based topology optimizations, but it also did so for 10% of the computational cost.

# Chapter 2

# Conventional forward modelling of an achromatic metalens

This chapter begins by examining how the fundamental goal of a lens, directing incident light into a target point, is realized in conventional refractive lenses and metalenses. Subsequently, a distinction between metalenses based on plasmonic and dielectric materials is made. The advantages and disadvantages of each type of metalens are discussed, and their underlying working mechanisms are explained in more detail based on examples from the literature. The chapter concludes with a comprehensive step-by-step design process commonly applied in recent state-of-the-art metalenses.

## 2.1 Design methods of an achromatic metalens

The fundamental objective of a lens, be it a conventional-, diffractive- or metalens, is to focus incident light beams at a target point [1]. Conventional lenses rely on refraction to bend the light towards the focal point. Thereby, the phase accumulates via the light propagating through the lens material. To obtain the required gradual phase change, conventional refractive lenses rely on polished surfaces with specific shapes.

Metalenses, however, are flat and, as a consequence, have to modify the phase differently. Contrary to traditional lenses, in metalenses, the phase delays are induced by the interaction between the metalens nanostructures and the incident light. In the last few years, various phase modulating methods have been applied for the design of broadband achromatic metalenses across the electromagnetic spectrum. Commonly used phase inducing mechanisms include resonant effects such as plasmonic resonances and non-resonant effects such as the propagation or the geometric phase. Typically, achromatic metalenses are designed based on a combination of several mechanisms. Furthermore, the choice of material is a critical factor in the design process. The ideal material has both a high refractive index and is lossless in the spectral region of interest [9]. Depending on the material composition, metalenses can be divided into either plasmonic or dielectric metalenses. In the following, the advantages and disadvantages of plasmonic and dielectric metalenses are listed. Examples of plasmonic and dielectric metalenses from the literature are given, and their underlying working mechanisms are discussed.

### 2.1.1 Plasmonic and dielectric metalenses

The early investigations on metalenses concentrated on plasmonic nanostructures, designed to operate near a scattering resonance to introduce the required phase delay [68]. Plasmonic metalenses can be made ultrathin (a fraction of a wavelength) and therefore are easier to fabricate than dielectric metalenses that consist of high-aspect-ratio nanostructures [1]. However, since only the phase delay is corrected without considering higher-order terms such as the group delay and group delay dispersion (see section 2.2 for a more detailed explanation), most plasmonic metalenses focus only single or a discrete set of wavelengths [69, 70, 71]. Moreover, the focusing efficiencies of the lenses are severely limited to 25% [72] due to the high intrinsic losses experienced by plasmonic materials. Consequently, in recent years dielectrics have become the material of choice for state-of-the-art metalenses, allowing both larger bandwidths and focusing efficiencies up to 90% [73, 74, 75].

#### Plasmonic metalenses

Plasmonic metalenses with broadband achromatic focusing capabilities have been reported in [76, 29, 77]. In [76], the authors demonstrated a plasmonic metalens that compensated for chromatic aberration by leveraging the structure dispersion resulting from surface plasmon polaritons (SPP) in a metal-insulator-metal (MIM) waveguide (Fig. 2.1a). The proposed metalens proved both in theory and numerical simulations achromatic performance at a wavelength ranging from 1000 and 2000 nm.

When TM polarized light impinges onto a MIM slit of subwavelength width, only the

SPP mode exists [76]. The phase retardation can be modulated by varying the SPP propagation constant inside the silver nanoslit-waveguides. By solving Maxwell's equation under appropriate boundary conditions, an expression for the SPP propagation constant $\beta$ can be derived [78]:

$$\tanh\left(\frac{\sqrt{\beta^2 - k_0^2\epsilon_d}\, w}{2}\right) = -\frac{\epsilon_d\sqrt{\beta^2 - k_0^2\epsilon_m}}{\epsilon_m\sqrt{\beta^2 - k_0^2\epsilon_d}}, \tag{2.1}$$

where $k_0$ is the optical wave vector in free space, $w$ is the width of the slit, and $\epsilon_d$ and $\epsilon_m$ are the dielectric and metal related permittivities, respectively. The permittivity of the metal $\epsilon_m(\omega)$ is given by the Drude model [78]:

$$\epsilon_m(\omega) = \epsilon_\infty - \frac{\omega_p^2}{\omega^2 + i\omega\gamma}, \tag{2.2}$$

where $\epsilon_\infty$ is the permittivity at infinite angular frequency, $\omega$ is the frequency of the incident radiation, $\omega_p$ is the plasma frequency of the free electron gas and $\gamma$ is the collision frequency. By considering only frequencies $\omega \ll \omega_p$ and $\omega \gg \gamma$, the permittivity $\epsilon_m(\omega)$ can be approximated to [76]:

$$\epsilon_m(w) \approx -\frac{\omega_p^2}{\omega^2} \ll -1. \tag{2.3}$$

Inserting equation 2.3 into 2.1, simplifies the propagation constant $\beta$ as follows [76]:

$$\beta = \frac{2\pi}{\gamma}\sqrt{\epsilon_d\left(\frac{2c}{\omega_p w} + 1\right)} = \text{constant}. \tag{2.4}$$

The phase retardation $\Delta\varphi$ experienced by electromagnetic waves traversing the subwavelength metallic slits, is defined as [79]:

$$\Delta\varphi = 2m\pi + \mathrm{Re}(\beta h) + \theta, \tag{2.5}$$

where $h$ represents the length of the nano-slit waveguides and $\theta = \arg\left[1 - \left(\frac{1-\beta/k_0}{1+\beta/k_0}\right)^2 e^{i2\beta h}\right]$ describes the multiple SPP reflections that occur between the entrance and exit interfaces [79]. However, according to the numerical simulations conducted by the authors, the phase shift $\Delta\varphi$ is mainly affected by the $\mathrm{Re}(\beta h)$ factor. Equation 2.5 thus shows that the phase retardations can be controlled by altering the slit widths $w$ and slit lengths $h$ of the MIM waveguides.

Based on these findings, Li et al. numerically simulated a nanoslit-based metalens consisting of 51 slits of constant length and varying widths. The period of the MIM unit cell was set to 200 nm and the phase distribution was designed for a focal length of $f = 5\,\mu\text{m}$. The achromaticity of the plasmonic lens becomes apparent from Fig. 2.1b, in which a near-constant focal length is shown for different wavelengths of light. Even though the MIM waveguides maintain a nearly constant effective index of refraction over the entire bandwidth, with a diameter of approximately 10 um the lens is small and challenging to fabricate. In addition, the focusing efficiency of the metalens was not reported. However, plasmonic metalenses usually suffer from high optical losses and reflections at the surface, which greatly inhibit the efficiency of the lens [9].

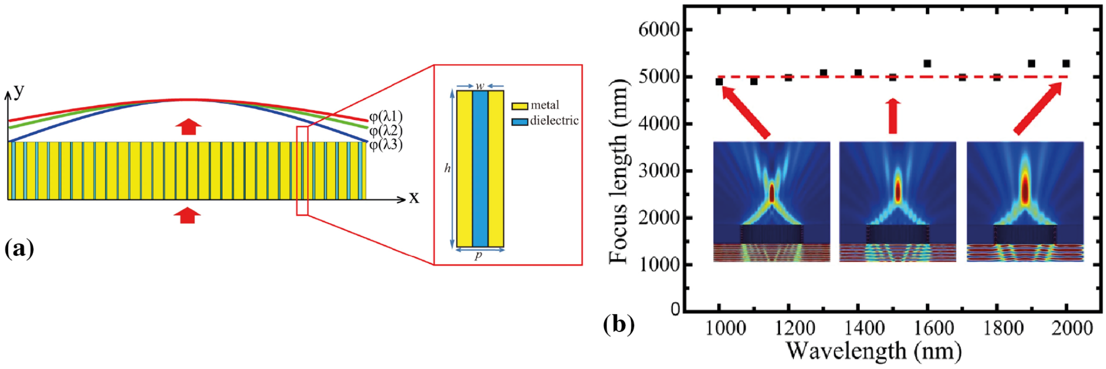


Figure 2.1: **A nanoslit-based achromatic metalens.** (a) Side view of the plasmonic metalens. 51 MIM waveguides, as shown in the inset, were selected to achieve broadband achromaticity. (b) Focal length of the designed metalens at different wavelengths. The insets show the electric field intensity $|E|^2$ distributions at $\lambda = 1000, 1500$ and $2000\,\text{nm}$ [76].

In a different approach, Wang et al. [29] realized broadband achromaticity by combining the geometric phase with the smooth and linear phase dispersion from metallic integrated-resonant unit elements (IRUs). As target phase for their metalens, the authors selected the hyperbolic phase profile given by:

$$\varphi(R,\lambda) = -\left[2\pi\left(\sqrt{R^2+f^2}-f\right)\right]\frac{1}{\lambda} \tag{2.6}$$

where $R$ is the radial distance to the center of the lens, $\lambda$ is the working wavelength in free space and $f$ is the designed focal length. They proceeded by splitting the required phase retardation in equation 2.6 into a wavelength-dependent and wavelength-independent term:

$$\varphi_{\mathrm{Lens}}(R,\lambda) = \varphi(R,\lambda_{\max}) + \Delta\varphi(R,\lambda) \tag{2.7}$$

with

$$\Delta\varphi(R,\lambda) = -\left[2\pi\left(\sqrt{R^2+f^2}-f\right)\right]\left(\frac{1}{\lambda}-\frac{1}{\lambda_{\max}}\right) \tag{2.8}$$

and with $\lambda_{\min}$ and $\lambda_{\max}$ being the boundaries of the wavelength band of interest. The wavelength-independent phase profile, $\varphi(R,\lambda_{\max})$, is realized by modulating the geometric or Pancharatnam–Berry phase [80, 81], which originates from the rotation of the anisotropic IRUs when illuminated by circularly polarized light. In contrast, the latter term in equation 2.7, $\Delta\varphi(R,\lambda)$, has a linear relation with $1/\lambda$, and can be obtained by tuning the integrated-resonant phase response of each unit element. Since both modulation methods work independently from each other, they can be applied simultaneously [29] to correct the chromatic aberration. As building blocks for the metalens, the authors used coupled metallic nano-rods that support various plasmonic resonances such as the fundamental dipolar mode and several higher-order modes. However, when the plasmonic resonances are excited, an abrupt phase shift occurs, which does not satisfy the linear phase difference $\Delta\varphi(R,\lambda)$ from equation 2.8. To overcome this issue, the authors concentrated on the near-linear phase distribution between the resonant modes for the chromatic correction [29].

Following this strategy, they experimentally realized a metalens, shown in Fig. 2.2a, with a diameter of 55.55 µm, a designed focal length of 100 µm and broadband achromaticity in the NIR between $1200 - 1680$ nm. Fig 2.2b illustrates the close agreement between the measured and simulated light intensity profiles along the cross-section plane and highlights that the focal spot remains at a near-constant 100 µm for all measured wavelengths [29]. Nevertheless, the proposed plasmonic metalens achieved a maximum focusing efficiency of only 12% and operated in reflection mode. Since transmissive optical devices are more suitable for practical applications, in the following, two dielectric-based metalenses operating in transmission mode are presented.

### Dielectric metalenses

In their seminal paper of 2018 [32], Shrestha et al. constructed polarization-independent achromatic metalenses with the assistance of a design libary consisting of numerous meta-units with complex cross-sectional geometries. Rather than incorporating plasmonic resonances, the authors utilized diverse waveguide modes in their isotropic nanos-

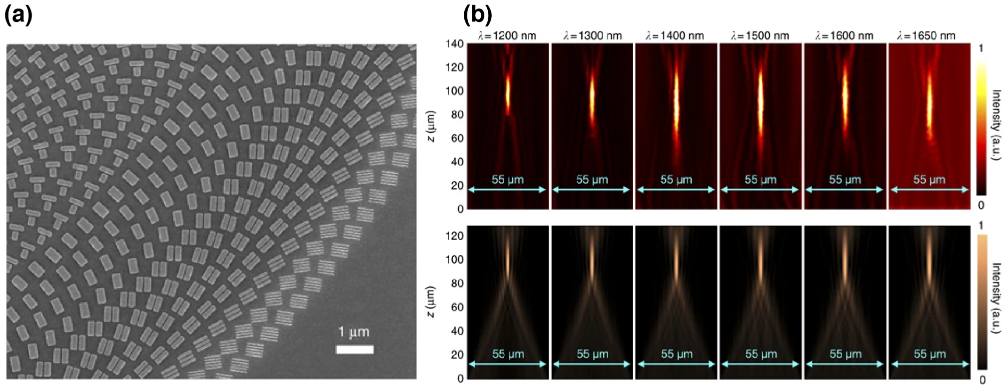


Figure 2.2: **A broadband plasmonic metalens operating in reflection mode.** (a) SEM image of the fabricated metalens. (b) Experimental (top row) and numerical (bottom row) intensity profiles of the plasmonic metalens for different incident wavelengths [29].

tructures for phase- and dispersion control. In the waveguide model, the propagation phase, $\phi(\omega, H)$, is written as follows:

$$\phi(\omega, H) = \frac{\omega}{c} n_{\text{eff}}\,(\omega) H, \tag{2.9}$$

where $\omega$ is the angular frequency of light, $n_{\text{eff}}$ is the effective index of refraction, $H$ is the height of the meta-unit and $c$ is the speed of light. The dispersion or group delay is given by the derivative of the propagation phase with respect to the angular frequency:

$$\frac{d\phi(\omega, H)}{d\omega} = \frac{1}{c}\left(n_{\text{eff}}(\omega) + \omega\frac{dn_{\text{eff}}(\omega)}{d\omega}\right) H. \tag{2.10}$$

Thus, the maximal range of group delays $\Delta\Phi'$ covered by a meta-unit library is given by [32]:

$$\Delta\Phi' = \left(\left.\frac{d\phi}{d\omega}\right|_{\omega_{\text{max}}} - \left.\frac{d\phi}{d\omega}\right|_{\omega_{\text{min}}}\right) = \frac{1}{c}\left(n_{\text{eff}}\left(\omega_{\text{max}}\right)\omega_{\text{max}} - n_{\text{eff}}\left(\omega_{\text{min}}\right)\omega_{\text{min}}\right) H. \tag{2.11}$$

As the meta-unit height $H$ and the speed of light $c$ are constant, the group delay $\frac{d\phi}{d\omega}$ is mainly influenced by the effective refractive index $n_{\text{eff}}(\omega)$, which can be altered by the transverse dimensions of the nanostructures. Since the group delay required at the center of the metalens monotonically increases with the lens radius, an upper bound for the radius $R_{\max}$ can be derived [32]:

$$R_{\max} \leq \frac{\Delta\Phi' c}{\Delta\omega\left(\frac{1}{\mathrm{NA}} - \sqrt{\frac{1}{\mathrm{NA}^2} - 1}\right)}, \tag{2.12}$$

which for a numerical aperture of $\mathrm{NA} \ll 1$ reduces to

$$R_{max}\mathrm{NA}\Delta\omega \leq 2c\Delta\Phi'. \tag{2.13}$$

Equation 2.13 illustrates that the basic properties of the metalens such as the radius, NA and spectral bandwidth correlate with each other and are limited by the range of group delays provided by the meta-unit library.

To meet the challenge of extending the range of dispersive responses as much as possible, the authors developed three different meta-unit libraries (referred to as Generation 1A, 1B and 2). Generation 1A (Fig. 2.3a) comprises singular and annular pillars and concentric rings made out of amorphous silicon with a height of 800 nm. Strikingly, by increasing the height of the nanostructures to 1400 nm (Fig. 2.3b), the dispersion range doubled in the Generation 1B library (Fig 2.3e). Lastly, to provide even more degrees of freedom for engineering the dispersion, the authors introduced the Generation 2 library, which has the same height as Generation 1B but uses geometries with fourfold symmetry rather than just rotational symmetry (Fig. 2.3c).

To evaluate the impact of the meta-unit libraries on the metalens performance, they subsequently fabricated metalenses based on all three generations of libraries [32]. The authors attributed the limited performance of the Generation 1A metalens to the insufficient group delay range covered by the 800 nm tall nanostructures (Fig. 2.3d). The metalens based on the Generation 1B library yielded achromatic performance across $1200 - 1650$ nm and had a diameter of 200 µm ($\mathrm{NA} \approx 0.13$) (Fig. 2.3g). Lastly, by using the Generation 2 library, a metalens with an increased NA of 0.88 but a reduced bandwidth of $\lambda = 1200 - 1400$ nm was constructed (Fig. 2.3h).

A change in the material platform from amorphous silicon to $TiO_2$ has enabled Chen et al. for the first time to accomplish broadband achromaticity over the entire visible spectrum [31]. They were able to engineer the target phase profile $\varphi(r, \omega)$ by separately tuning the group delay $\frac{\partial\varphi}{\partial\omega}$ and group delay dispersion $\frac{\partial^2\varphi}{\partial\omega^2}$ of each nanostructure, independently from its phase $\varphi(r, \omega_d)$ [31]. More specifically, the required phase profile was realized based on the wavelength-independent geometric phase [80, 81], whereas the

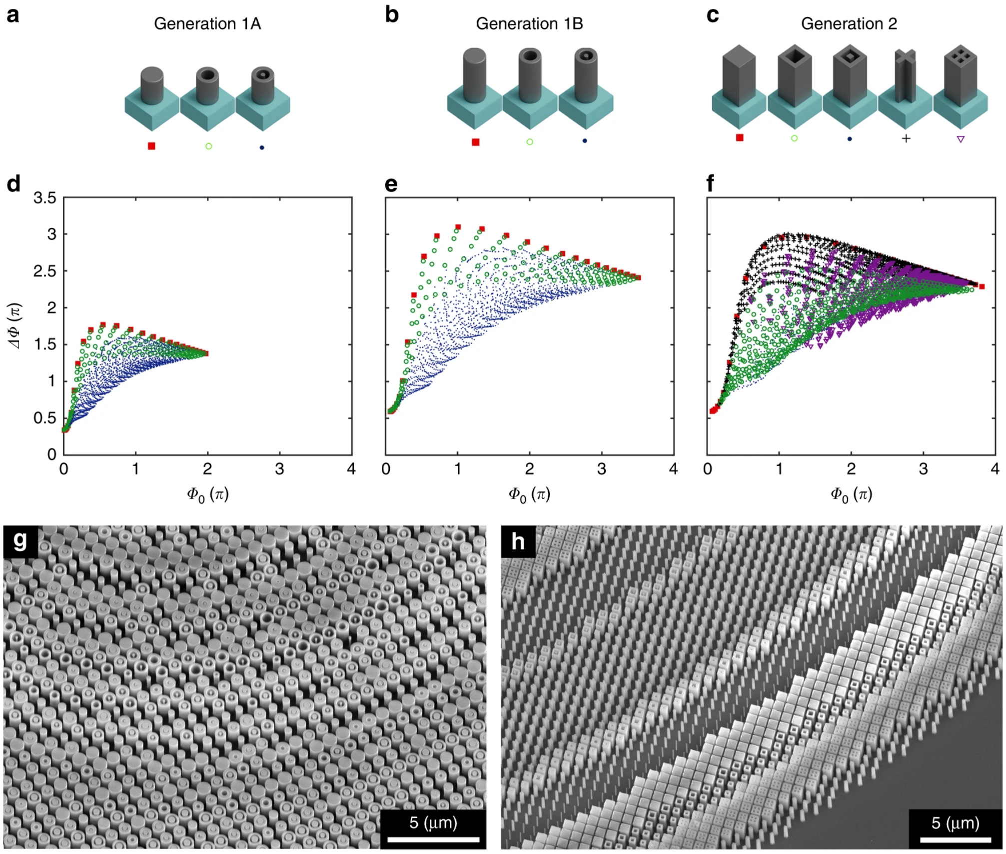


Figure 2.3: **A dielectric-based polarization-insensitive achromatic metalens.** (a)-(c) Schematics showing the different meta-atom archetypes used for building the Generation 1A, 1B and 2 meta-unit libraries. (d)-(f) Calculated phase $\Phi_0$, of the largest wavelength $\lambda = 1.6\,\mu\text{m}$ and dispersion $\Delta\Phi = \frac{d\phi}{d\omega}\Delta\omega$, of each meta-atom in the Generation 1A, 1B and 2 libraries. (g), (h) SEM images of the fabricated metalenses using the Generation 1A (g) and Generation 2 (h) meta-atoms, respectively. [32].

metalens dispersion was corrected via the different waveguide modes forming within the nanostructures. To more easily fulfil the multiple conditions of phase, group delay and group delay dispersion at each lens coordinate, they increased the degrees of freedom in their meta-atom design by utilizing coupled phase-shift elements, i.e. two anisotropic nanofins in close proximity (Fig 2.4a). The optical response of a single $TiO_2$ nanofin agrees well with the one of a truncated waveguide. A small change to equation 2.9 has to be made, however, since linearly polarized light experiences different effective indices of refraction, $n_l$ and $n_s$, depending on whether it passes through the short or long axes of the nanofin:

$$\phi_{l,s}(\omega, h) = \frac{\omega}{c} n_{l,s}(\omega) h. \tag{2.14}$$

The complex transmission coefficients along the long and short axes, $t_l$ and $t_s$, can be written as [9]:

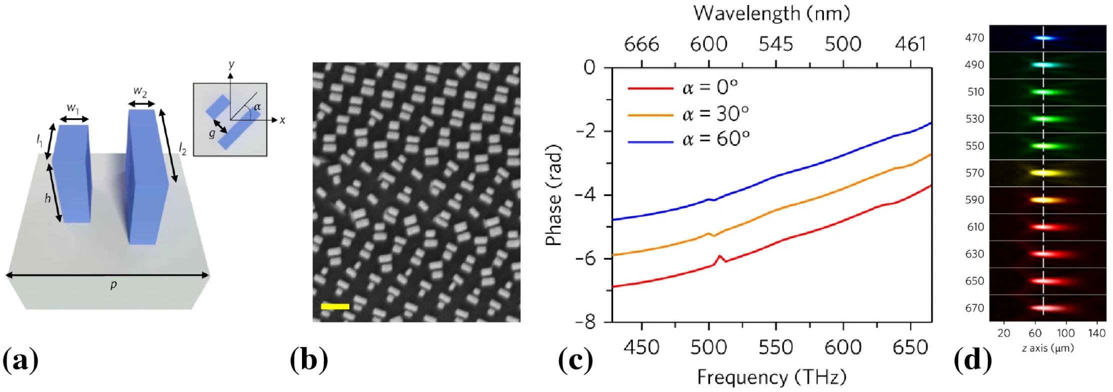
$$t_{l,s} = |t_{l,s}| \exp\left[i\frac{2\pi}{\lambda}(n_{l,s} - 1)h\right], \tag{2.15}$$

where $h$ is the height of the nanofin and $\lambda$ is the incident wavelength in vacuum. Under circularly polarized light illumination $(1 \pm i)^T$ (the positive sign denotes left-handed polarized light), the transmitted electric field can be described by the Jones vector [9, 82]:

$$\frac{t_l + t_s}{2}\begin{bmatrix} 1 \\ \pm i \end{bmatrix} + \frac{t_l - t_s}{2} e^{\pm i2\alpha}\begin{bmatrix} 1 \\ \mp i \end{bmatrix}. \tag{2.16}$$

Part of the transmitted light flips its polarization state and receives an additional phase shift of $\arg\left(\frac{t_l - t_s}{2}\right) \pm 2\alpha$, where $2\alpha$ is the wavelength-independent geometric or Pancharatnam–Berry phase [80, 81], equal to twice the rotation angle of the nanofins. This part of the transmitted light is known as the cross-polarization. The remaining part, the co-polarization, maintains its polarization state and acquires a phase shift of $\arg\left(\frac{t_l + t_s}{2}\right)$, where arg refers to the argument of a complex number [9]. Since the geometric phase $2\alpha$ is independent of the wavelength, it allows the decoupling of the target phase from the required group delay and group delay dispersion. Moreover, it enabled the authors to impart a relative phase delay from 0 to $2\pi$ by purely rotating the nanofins (Fig 2.4c) [9, 31].

For the design of the achromatic metalens, the authors proceeded similarly to Shrestha et al. [32] by selecting from a library the nanofin dimensions that closest match the phase and dispersion requirements. The fabricated metalens demonstrated diffraction-limited achromatic focusing from 470 to 670 nm with an efficiency of up to 20% and NA of 0.2. Because the acquired geometric phase depends on the helicity of the incident light, geometric phase-based metalenses are generally polarization-insensitive (equation 2.16) [9]. In a later work, however, the group overcame the polarization sensitivity by restricting the rotation angle of the anisotropic nanofins to 0 and 90 degrees, thus eliminating the conjugate values resulting from the geometric phase [35]. Lastly, the proposed metalens implemented nanofins that achieved a group delay range of about 5 fs. An extended group delay range - a crucial necessity for realizing lenses with larger diameters, bandwidths and NAs (see equation 2.13) - was obtained in recent state-of-the-art metalenses with more complex nanostructures of varying heights [8, 36, 37].



**(a)** **(b)** **(c)** **(d)**

Figure 2.4: **Dielectric-based achromatic metalens operating in the visible region.** (a) Schematic of the $TiO_2$ nanofin. Each unit cell consist of one or more nanofins, which can vary in the transverse dimension (length and width) and rotation angle $\alpha$, but stay constant in height ($h = 600\,\text{nm}$) and pitch ($p = 400\,\text{nm}$). (b) SEM image of a region on the fabricated metalens. (c) Phase as a function of the frequency for a single nanofin ($l = 250\,\text{nm}$ and $w = 80\,\text{nm}$) and varying rotation angles $\alpha$. (d) Measured intensity distributions of the achromatic metalens in linear scale in the $x - z$ plane [31].

## 2.2 Principle of achromatic focusing

The design process of a broadband achromatic metalens typically consists of the following three steps: First, the target phase profile $\varphi(r, \omega)$ at a design frequency $\omega_d$ needs to be computed ($r$ represents the radial distance from the metalens center). Afterwards, a library containing multiple nanostructure geometries together with their optical responses is developed via electromagnetic solvers. Finally, the nanostructures that fulfil the phase and dispersion requirements of the target phase the closest are selected out of the library for each metalens coordinate $r$ [9].

### Calculation of the target phase profile

Depending on the number of aberrations that need to be corrected simultaneously (e.g. coma, spherical aberration, distortion, astigmatism etc.) [83], advanced ray-tracing software is necessary to derive an expression for the target phase profile. However, to focus a broadband beam of light impinging at a right angle onto a metalens into a diffraction-limited spot, the following analytical expression of the hyperbolic phase profile is sufficient [69]:

$$\varphi(r,\omega) = -\frac{\omega}{c}\left(\sqrt{r^2+F^2}-F\right), \tag{2.17}$$

where $r$, $\omega$ and $F = f \cdot \omega^n$ ($n = 0$ for an achromatic lens) are the lens radial coordinate, angular frequency, and focal length, respectively. Since the target phase is frequency-dependent, the metalens has to impart a specific phase profile for each wavelength of the incident light. Equation (2.17) can be Taylor expanded around a design frequency $\omega_d$ which is usually chosen at the center of the bandwidth of interest [9]:

$$\varphi(r,\omega) = \varphi(r,\omega_d) + \left.\frac{\partial\varphi(r,\omega)}{\partial\omega}\right|_{\omega=\omega_d}(\omega-\omega_d) + \frac{1}{2}\left.\frac{\partial^2\varphi(r,\omega)}{\partial\omega^2}\right|_{\omega=\omega_d}(\omega-\omega_d)^2 + \ldots, \tag{2.18}$$

where $\varphi(r,\omega_d)$ represents the required phase, and $\frac{\partial\varphi(r,\omega)}{\partial\omega}$ and $\frac{\partial^2\varphi(r,\omega)}{\partial\omega^2}$ denote the group delay and group delay dispersion, respectively. These terms can be intuitively understood by viewing the incoming polychromatic light as individual wavepackets (Fig 2.5). The required phase directs the wavepackets towards the focal point, whereas the group delay compensates for the difference in optical path lengths and thus ensures that wavepackets from different radial coordinates arrive at the focal point simultaneously to constructively interfere [9]. The group delay must be independent of the wavelength to realize the correction of the chromatic aberration. In that case, as suggested by Chen et al. [31], the group delay dispersion and other higher-order terms become negligibly small. An expression for the group delay can be found by taking the derivative of equation 2.17 with respect to the angular frequency $\omega$ of the light:

$$\frac{\partial\varphi(r,\omega)}{\partial\omega} = -\frac{1}{c}\left(\sqrt{r^2+F^2}-F\right). \tag{2.19}$$

The challenge in the design of a broadband achromatic metalens, is to find nanostructures that satisfy both the phase $\varphi(r,\omega)$ and group delay requirements $\frac{\partial\varphi(r,\omega)}{\partial\omega}$ at every metalens coordinate $r$.

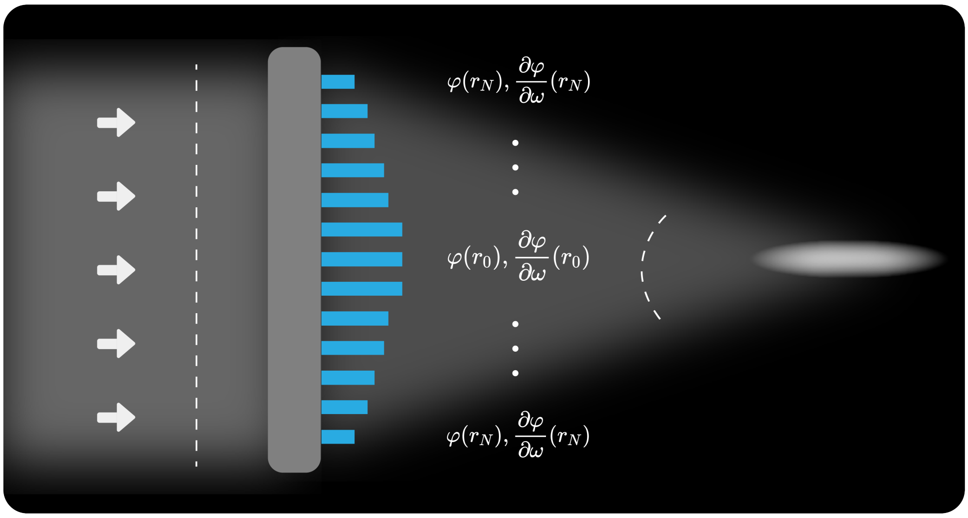


Figure 2.5: **Principle of achromatic focusing**. To realize achromatic focusing, the building blocks of a metalens must both satisfy the required phase $\varphi(r,\omega)$ and group delay $\partial_\omega\varphi(r,\omega)$ at every radial coordinate $r$. While the phase $\varphi(r,\omega)$ directs the incoming wavepackets towards the focal point, the group delay $\partial_\omega\varphi(r,\omega)$ ensures simultaneous arrival.

### Development of a design library

Once the target phase has been determined, the next step is to develop a library of nanostructures with their corresponding phases and group delays. The linkage between a specific meta-atom geometry and its optical response is generally established through rigorous coupled-wave analysis, finite element methods, or finite-difference time-domain methods [1]. With regard to the choice of the nanostructure, initial metalenses were made out of deeply subwavelength meta-atoms that modulated the phase based on resonant effects [68]. These metalenses suffered from severe chromatic aberrations and operated at single frequencies only [69, 70, 71], since the derivate terms in equation 2.18 were not considered. In contrast, commonly used nanostructures today have thicknesses on the order of the wavelength and function similar to truncated waveguides [68]. To expand the nanostructure library and therefore facilitate the implementation of the target phase profile, the trend in state-of-the-art metalenses is going towards more complicated meta-atom geometries with high aspect ratios. Examples of nanostructures with increased geometric degrees of freedom include coupled nanofins [35], isotropic nanopillars with fourfold symmetry [32], cross-circle waveguides [36] and nanopillars of variable heights [8].

### Brutal force-based meta-atom selection

In the last step of the metalens design, an optimization algorithm selects a nanostructure from the constructed library, whose phase and dispersion fit the target phase profile $\varphi(r_i, \omega)$ as closely as possible. This process is then repeated for every metalens coordinate $r$ [9]. As metalenses grow larger and reach milimeter- to centimeter-scales [84, 85], the number of the constituent subwavelength meta-atoms increases significantly and with it the difficulty in selecting the right structures with conventional brute-force methods. This thesis therefore examines whether the appropriate nanostructures can be more accurately and efficiently identified from the design library by leveraging the power of machine learning.

# Chapter 3

# Principle of machine learning inverse design of an achromatic metalens

This chapter begins with an overview of the commonly used machine learning techniques for tackling the inverse design problem in nanophotonics. Here, a distinction between discriminative and generative models will be made and the weaknesses and strengths of each learning strategy will be illustrated with examples from the literature. Based on this knowledge, various machine learning architectures will be investigated to realize the inverse design of achromatic metalenses. Particular attention will be given to the decision tree model. It could model the relationship between the metalens building blocks and the corresponding optical response more adequately than conventionally used neural network architectures. In the last section of this chapter, the working principle of the decision tree algorithm will be explained in more detail.

## 3.1 Inverse design models used in nanophotonics

In recent years, a wide variety of machine learning models have been investigated to tackle the inverse design of complex nanophotonic structures. In general, two types of machine learning models are utilized: discriminative and generative models. Discriminative models predict the probability of a target variable $y$, given an input $x$, i.e., $P(y|x)$ [65]. Since $y$ can be a class label or a continuous variable, discriminative models are commonly used for both regression and classification tasks. To establish the desired mapping between the input $x$ and the output data $y$, discriminate models rely on a labelled training set. In contrast, generative models learn the joint probability distribution of $x$ and $y$, $P(x, y)$, and usually belong to unsupervised learning schemes where no labels are required in the training process [65]. One of the advantages of generative over discriminative models is that they can produce data similar to the existing training set by sampling from the joint probability distribution $P(x, y)$ [65]. In the context of nanophotonics, generative models can also assist in reducing the dimensionality of the design space, which can accelerate global optimization searches [86]. Examples of frequently used generative models include generative adversarial networks (GANs) and variational autoencoders (VAEs). Machine learning algorithms such as logistic regression, support vector machines, decision trees, random forests and neural networks, on the other hand, are typical discriminative models. In the following, the strength and weaknesses of discriminative and generative models will be further illustrated with concrete examples from the literature.

### 3.1.1 Discriminative Models

In one of the initial investigations on data-driven inverse design, Peurifoy and colleagues demonstrated the successful implementation of discriminative machine learning models for the design of layered photonic structures [47]. More precisely, they reported a fully connected neural network (FCN) architecture that could accurately predict the light scattering of multilayer nanoparticles consisting of silica cores with eight alternating $TiO_2$ and silica shells (Fig. 3.1a). Additionally, they proposed a solution to the inverse design of multilayer nanoparticles by running their trained forward modeling network backwards. Before going into more details on the employed model architecture, the working principle of an FCN is briefly discussed.

In general, a fully connected neural network consists of one input layer, one or several hidden layers, and one output layer [49]. Each layer consists of multiple neurons that connect to the neurons of the adjacent layer. The single neurons perform certain transformations to the input data through their associated weights and biases and pass the output along to the next layer of the network. As an example, to find the output of the $j$th neuron in the $m$th layer, written as $a_j^m$, the weighted sum of the outputs from the previous layer, $z_j^m$, needs to be computed and modified by a nonlinear activation function, denoted as $\sigma(\cdot)$ [49]:

$$a_j^m = \sigma\left(\sum_k w_{jk}^m \, a_k^{m-1} + b_j^m\right) = \sigma\left(z_j^m\right), \tag{3.1}$$

where $w_{jk}^m$ refers to the weight from the $k$th neuron in the $(m-1)$th layer to the $j$th neuron in the $m$th layer, and $b_j^m$ to the bias of the $j$th neuron in the $m$th layer. The incorporation of a nonlinear activation function $\sigma(\cdot)$ was biologically inspired by the rate of action potential firing inside a brain cell [87] and is critical for enabling neural networks to solve nontrivial problems. Thanks to the introduced nonlinearity, neural networks comprising only two hidden layers are theoretically proven to fit any continuous function, according to the universal approximation theorem [88]. Commonly used activation functions include the sigmoid, $\sigma(x) = \frac{1}{1+e^{-x}}$, the hyberbolic tangent, $\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$, and the rectified linear unit (ReLU), $\max\{0, x\} = x \, \mathbb{1}_{x>0}$. The process of passing the input data from one layer to the next is continued until the output layer is reached, and the predictions from the output neurons are retrieved (forward pass) [49]. To measure the network's prediction error, a cost function $C(\cdot)$ is defined, which computes the difference between the desired and the actual output of the network [49]. Subsequently, the backpropagation learning algorithm traverses the network from the output to the input layer to measure the error contribution from each neuron [89]. The connection weights are finally adjusted to reduce the error from the previous forward pass. Typically, the training of neural networks involves numerous weight updates till a satisfactory performance is achieved.

Coming back to the work of Peurifoy et al., a fully connected neural network consisting of one input and one output layer and additional four hidden layers was adopted. The FCN received as input the normalized (zero mean and unit variance) thickness of each nanoparticle shell and predicted as output the scattering cross-section for 200 points between 400 and 800 nm. For nanoparticles with two alternating layers, they selected 100 neurons per hidden layer. To model the optical response of particles with ten shells, the authors increased the network size to 300 neurons per hidden layer. The training set of the FCN consisted of 50,000 nanoparticles and their associated scattering responses, previously obtained from the transfer matrix method. As a cost function, they chose the mean square error, and it took around 1,000 to 2,000 weight updates until the model's performance stopped improving. Once the model was trained to approximate the light scattering from multilayer nanoparticles, the authors attempted to solve the inverse design problem by fixing the weights of the network and setting the output to the desired response. The input was set as trainable and was modified by backpropagation until the geometry that reproduced the desired result was suggested. Results indicated that for nanoparticles with five to ten shells, the FCN could solve the inverse design more accurately and more quickly than conventional optimization algorithms [47].

Another successful implementation of FCNs was reported by Liu et al. [48], who proposed a tandem architecture consisting of a pretrained forward modeling network connected to an inverse design network to study the transmission of multilayer thin films. Importantly, by combining both FCNs, they could successfully overcome the nonuniqueness problem that negatively affects the training of neural networks and typically arises when more complicated systems with larger datasets are considered. To build the tandem network, the authors began by training a forward modeling network that maps a design (i.e. the thickness of each layer of the thin film) to its associated transmission spectrum [48]. Later, in the tandem configuration, the weights of the pretrained forward modeling networks were fixed, and the weights of the inverse design network were trained to minimize the cost function defined as $E = \frac{1}{2}\sum_i (r_i - \sigma_i)^2$. Here, $r_i \in \mathcal{R} = [r_1, r_2, \ldots, r_{200}]$ is the target response taken as the input to the tandem network, and $\sigma_i \in [\sigma_1, \sigma_2, \ldots \sigma_{200}]$ is the approximated spectrum of the forward modeling part in the output of the combined network. The inverse network, which produces the designs $\mathcal{D}$ in the intermediate layer, is composed of four layers with 200-500-200-20 neurons per layer (Fig. 3.1b). Crucially, by not forcing the inverse network to reproduce the real designs from the training set, but instead asking it to minimize the cost between the target and predicted response from the forward modeling network, the authors could overcome the nonuniqueness problem.

### 3.1.2 Generative Models

While discriminative models have shown remarkable accuracy in approximating Maxwell's equations, and have demonstrated methods to overcome the data inconsistency in inverse design problems, it gets increasingly difficult to prepare an adequate training set once the complexity of the system rises. Moreover, since discriminative models are deterministic, they produce only one device for a given optical response and thereby contrast physical intuition that diverse designs can have similar properties.

To circumvent these drawbacks, Liu et al. incorporated GANs, an unsupervised learning method, into the inverse design of metasurfaces [64]. The complete architecture consisted of three convolutional neural networks: a generator and a critic, together constituting

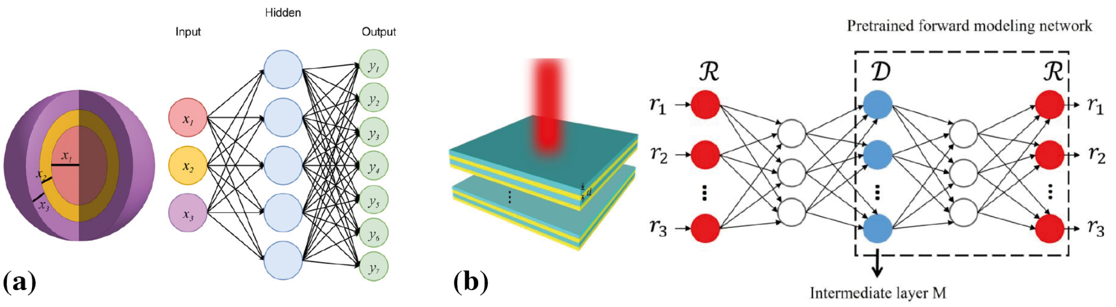


Figure 3.1: **Nanophotonic inverse design based on discriminative models.** (a) FCN-assisted forward modeling and inverse design of multilayer nanoparticles. Nanoparticles comprise silica cores with alternating layers of $TiO_2$ and silica shells. The forward modeling network predicts the scattering cross-section at different wavelengths for given shell thicknesses as the input. For illustration purposes a simplified FCN architecture is shown [47]. (b) Proposed tandem architecture, composed of a forward modeling network and an inverse design network, for investigating the transmission and design of multilayer thin films. For illustration purposes a simplified tandem network is shown [48].

the GAN, and a pretrained simulator (Fig. 3.2). The generator received as input the spectra $T$ and some noise $z$ and produced an arbitrary pattern of the unit cell structure, represented as a binary image of 64x64 pixels. The associated transmittance spectra $\hat{T}$ of the generated metasurface patterns could be retrieved from the simulator, a forward modeling network that was pretrained on 6,500 full wave FEM simulations. The critic of the proposed GAN finally compared the generated patterns to the user-defined geometric data and computed the distance between the two distributions. The overall objective was to train a generator to produce fabrication-friendly structures with transmission spectra similar to the user-defined input spectra. To achieve this goal, the weights of the generator were adjusted to minimize both the loss defined by the simulator (i.e. the Euclidean distance between the spectra of the generated pattern $T'$ and the input spectra $T$) and the critic (i.e. the distance of the distributions between the geometric data and the generated pattern). The role of the critic was essential, as it restricted the generator to construct patterns that resembled the user-defined input geometries. After the training was complete, the authors demonstrated the performance of the generator by designing a unit cell structure that successfully replicated a randomly generated Gaussian-like response [64].

In [66], a variational autoencoder (VAE), another class of generative models, was used to tackle the inverse design of metamaterials and solve the one-to-many mapping problem

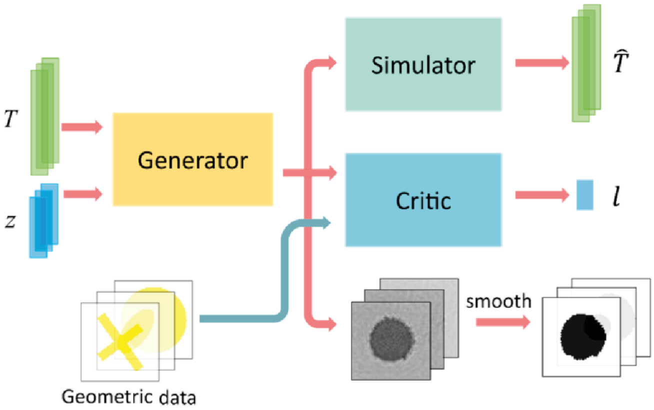


Figure 3.2: **GAN mediated inverse design of metasurfaces.** The architecture of the proposed network consists of a generator, a critic and a pretrained simulator. The generator accepts the spectra $T$ and noise $z$ as input and aims to produce images of realistic designs. The critic compares the generated with the actual patterns and computes the distance $l$ between the distributions of the two sets of geometries. The pretrained simulator approximates the optical response $\hat{T}$ of a given structure [64].

prevalent in deterministic learning schemes.

To facilitate the understanding of the proposed model architecture, the working principle of VAEs is briefly discussed. VAEs fundamentally consist of an encoder-decoder network. The encoder compresses the input data $x$ into a predefined prior distribution $p(z|x)$ (e.g. a standard Gaussian distribution) over the latent space (i.e. the space representing the compressed data). Afterwards, a point $z$ (also denoted as a latent variable) is stochastically sampled from the latent distribution $p(z|x)$. The decoder then attempts to reconstruct the initial input data, $x \sim d(z)$, from the sampled point $z$ [90]. VAEs are often referred to as probabilistically generative models due to the fact that the encoders model the input as a distribution $p(z|x)$ over the latent space instead of a single point $z$ [90].

To realize the VAE, Ma and coworkers proposed an architecture composed of three submodules: the recognition model, the prediction model and the generation model [66]. As shown in Fig. 3.3b, said submodules were implemented by four neural networks, namely a feature extraction network, a prediction network, a recognition network and a reconstruction network. The recognition network, which served as the aforementioned encoder, compressed the input data into a standard Gaussian distribution over a low-dimensional latent space. Here, as input data, a metal-dielectric-metal sandwich structure (as shown in Fig. 3.3a), represented as a 2D binary image of 64x64 pixels, and its associated reflection coefficients $R_{xx}, R_{yy}$ and $R_{xy}$ were considered. The prediction model was used for quickly approximating the reflection spectra of the metamaterials. The generation model, operating as the decoder of the VAE, finally produced the new designs when fed with the latent variables (previously sampled from the latent space) and the desired optical response. Crucially, due to the stochasticity of the sampling process, the decoder allowed the design of multiple geometries for fixed optical requirements and thereby provided a solution to the one-to-many mapping issue [66]. Compared to the GAN approach mentioned earlier, the proposed VAE did not rely on a pretrained simulator for the inverse design. What is more, the latent space representation of the structural designs and the corresponding optical responses could potentially be explored to more intuitively understand the complex design-performance relationship in nanophotonics [66].

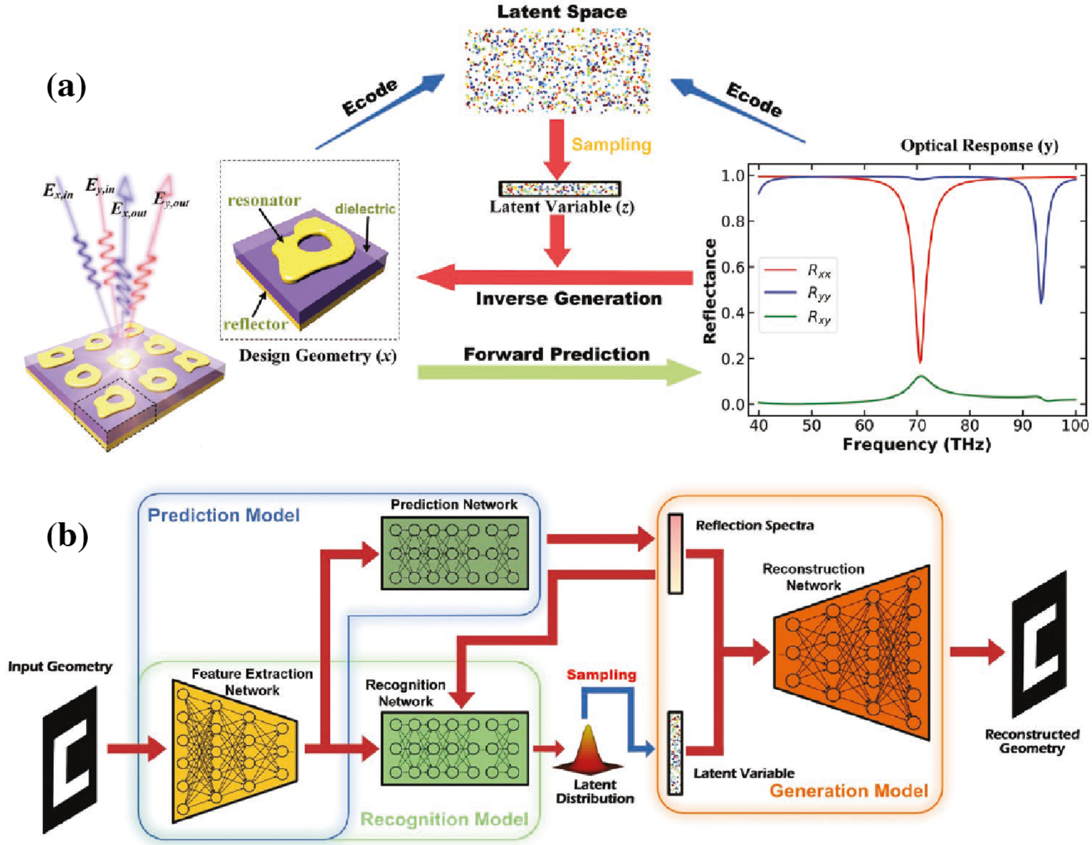


Figure 3.3: **VAE-based forward modeling and inverse design of metal-dielectric-metal nanostructures.** (a) Both structural designs and optical responses are encoded into a low-dimensional latent space with a predefined Gaussian distribution. Latent variables are sampled from the latent space and are utilized for inverse generation. (b) Three submodules, a recognition model, a prediction model and a generation model consitute the complete architecture of the proposed VAE and are formed by four different neural networks. The recognition model compresses the structural designs and the corresponding optical responses into a latent space. The prediction model approximates the optical response of the investigated metamaterials and the generation model produces a design based on a sampled latent variable and a required optical response [66].

## 3.2 Comparison between neural networks and decision tree methods for the inverse design of an achromatic metalens

Although generative models have shown great potential in producing unintuitively shaped metasurfaces with promising optical performances, this thesis is primarily concentrated on utilizing discriminative models to achieve the inverse design of achromatic metalenses. To successfully carry out the inverse design, discriminative models need to establish the linkage between the metalens building blocks and their associated optical responses. In the following chapter, this linkage will be obtained by training a machine learning model on a design library consisting of numerous meta-atom geometries and their corresponding phase $\varphi$ and group delay responses $\frac{\partial\varphi}{\partial\omega}$. The meta-atoms used to build up this library are symmetric and asymmetric (square- and rectangular-shaped) nanopillars of varying heights. As shown in Fig. 3.4, the employed nanopillars are fully characterized by the incident wavelength $\lambda$, the pitch P, the height $H$, the length $L$ and the width $W$. By exploiting three-dimensional instead of the commonly used planar geometries, one gains both an added degree of freedom for fine-tuning the dispersion and, more crucially, an increased capability for modulating the group delays. As the nanopillars can be treated as truncated waveguides, their group delays increase linearly with the pillar heights, whereas the slope of the group delays is controlled by the transverse dimensions (length and width) of the nanopillars (see equation 2.1.1). Once a machine learning model has been successfully trained on the design library, it can near-instantaneously predict the nanostructures fitting a predefined target phase $\varphi$ and group delay profile $\frac{\partial\varphi}{\partial\omega}$.

During the course of this work, different machine learning algorithms have been tested on their ability to map the meta-atoms to their corresponding optical properties. Initial investigations have been conducted based on neural networks since they have a proven track record of success in solving inverse design tasks in nanophotonics. Fig. 3.5 summarizes the three main architectures that were implemented. The first architecture (Fig. 3.5a) consisted of a fully connected neural network with three input neurons for the optical response (i.e. the phase, the group delay and the transmitted amplitude) and five output neurons for the design parameters of the nanopillar (i.e. $H, L, W, P$ and $\lambda$). Between the input and the output layer, there were additional four hidden layers with 1024 neurons each. The second architecture, shown in Fig. 3.5b, was inspired by the work of Peurifoy et al. [47] (see subsection 3.1.1) and consisted of the pretrained forward modeling network shaded in blue. Since the input layer did not contain any trainable

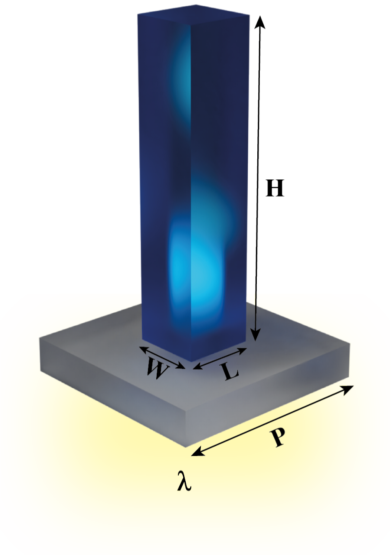


Figure 3.4: **Schematic of the 3D meta-atom.** The nanopillar is characterized by its height $H$, length $L$, width $W$ and pitch $P$. To build the design library, multiple nanopillar geometries were simulated at different incident wavelengths, $\lambda$, of light.

weights, an additional dense layer (coloured in red) was added to the network from which the designs could be retrieved. For the last architecture (Fig. 3.5c), a tandem network, as it was introduced by Liu et al. [48] (see subsection 3.1.1), was employed. In this architecture, a pretrained forward modeling network was combined with an inverse design network. Here, the predicted designs could be retrieved from the intermediate layer shown in red.

Despite their prior success in inverse design problems, neural networks were not able to predict the nanopillars based on their optical responses with sufficient accuracy. Instead, it was discovered that decision trees were much more successful in modelling the design-performance relationship. Besides an increased accuracy in predicting the nanopillars, decision trees trained much faster[1] than neural networks. An additional advantage of the decision tree is that it allows for easy interpretability of its predictions, as will become apparent in the next section [89]. Such models are often referred to as white-box models. In contrast, neural networks are generally considered black-box models because it is usually hard to explain why they made their predictions [89]. Finally, when deciding

---

[1] Without considering any hyperparameters, the decision tree completed the training in approximately one second on a MacBook Pro 2015 with an 2,7 GHz Dual-Core Intel Core i5 processor. By contrast, the neural networks described in Fig 3.5 took several hours to carry out 1,000 iterations of training on the latest NVIDIA Tesla P100 graphics card provided by the Leibniz Supercomputing Centre.

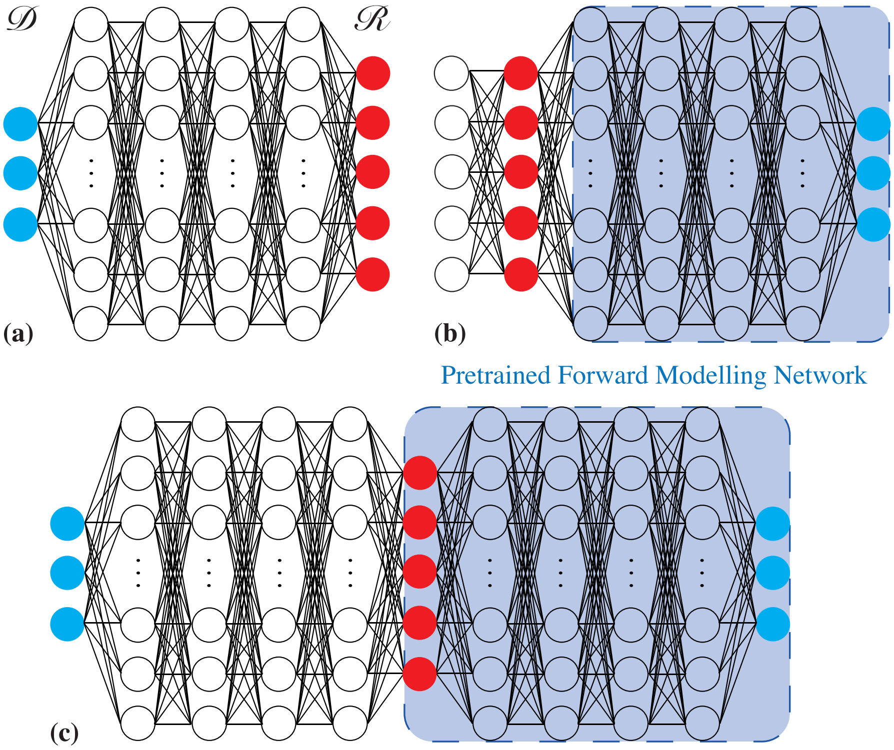


Figure 3.5: **Failed attempts of solving the inverse design with neural networks.** (a) Neural network architecture comprising one input and one ouput layer and four hidden layers, with 1024 neurons per layer, in between. Here, $\mathcal{R}$ denotes the desired optical response and $\mathcal{D}$ refers to the predicted nanopillar design. (b) Inspired by the work of Peurifoy et al. [47], an architecture containing a pretrained forward modeling network was utilized. (c) Tandem architecture, in which the output of an inverse design network is connected to the input of a pretrained forward modeling network [48].

which machine learning algorithm to apply for a given task, some valuable insight can be extracted from Kaggle, a website hosting data science competitions with more than 6 million users worldwide. By evaluating the winning solutions, Kaggle's CEO concluded that while neural networks appear to be best suited for unstructured data (i.e. audios, images and videos), tree-based algorithms shine for structured data (i.e. tabular data as used in this work) [91, 92].

## 3.3 Principle of a decision tree-based algorithm

Before applying the decision tree for the metalens design, it is crucial to understand its underlying working principle first. The decision trees are implemented with the machine learning library scikit-learn [93], which applies the Classification and Regression Tree (CART) algorithm to train (or "grow") its trees [89]. At the beginning of the training process, the CART algorithm starts by splitting the training set (e.g. the design library) into two subsets using a single input feature $k$ (e.g. the phase or the group delay) and a threshold $t_k$ (e.g. group delay $< 13\,\text{fs}$) [89]. The algorithm essentially subdivides the data by asking a true or false question. But it does not ask any arbitrary true or false question, but the ones that produce the subsets with the smallest mean squared error (MSE). More specifically, the CART algorithm chooses the pair of $(k, t_k)$ that minimize the cost function given by

$$J(k,t_k) = \frac{m_{\text{left}}}{m}\,\text{MSE}_{\text{left}} + \frac{m_{\text{right}}}{m}\,\text{MSE}_{\text{right}}, \quad \text{where} \begin{cases} \text{MSE}_{\text{node}} & = \sum\limits_{i \in \text{node}} (\hat{y}_{\text{node}} - y^{(i)})^2 \\ \hat{y}_{\text{node}} & = \frac{1}{m_{\text{node}}} \sum\limits_{i \in \text{node}} y^{(i)} \end{cases} . \tag{3.2}$$

Here, $m_{\text{left/right}}$ is the number of instances in the left/right subset and $\text{MSE}_{\text{left/right}}$ is the mean squared error of the left and right node, respectively [89]. Afterwards, it continues splitting the subsets in a similar manner. This iterating process stops once the tree has reached a certain maximum depth (defined by the `max_depth` hyperparameter) or when it cannot find any splits to further reduce the MSE. The nodes that cannot be divided into smaller subsets are said to be pure and are referred to as leaf nodes [89].

After the tree has been trained and has developed its full structure, it can be used to make predictions on previously unseen data points. For instance, suppose one wants to ask the tree for the nanopillar dimensions that correspond to a phase of 0.5 and a group delay of 1 fs (Fig. 3.6). Starting from the top node (i.e. root node), the tree is traversed until a bottom node is reached, where a certain nanopillar is predicted. The prediction made here is simply the average of the nanopillars of the training set associated with that particular bottom node. Importantly, since newly predicted pillars are just averages of existing pillars from the training set, decision trees can only predict the nanostructures that lie within the parameter space of their training sets. This means that, even if decision trees are capable of learning the structure-performance relationship, they are fundamentally constrained to predict metalenses that fall within the space of their corresponding design library.

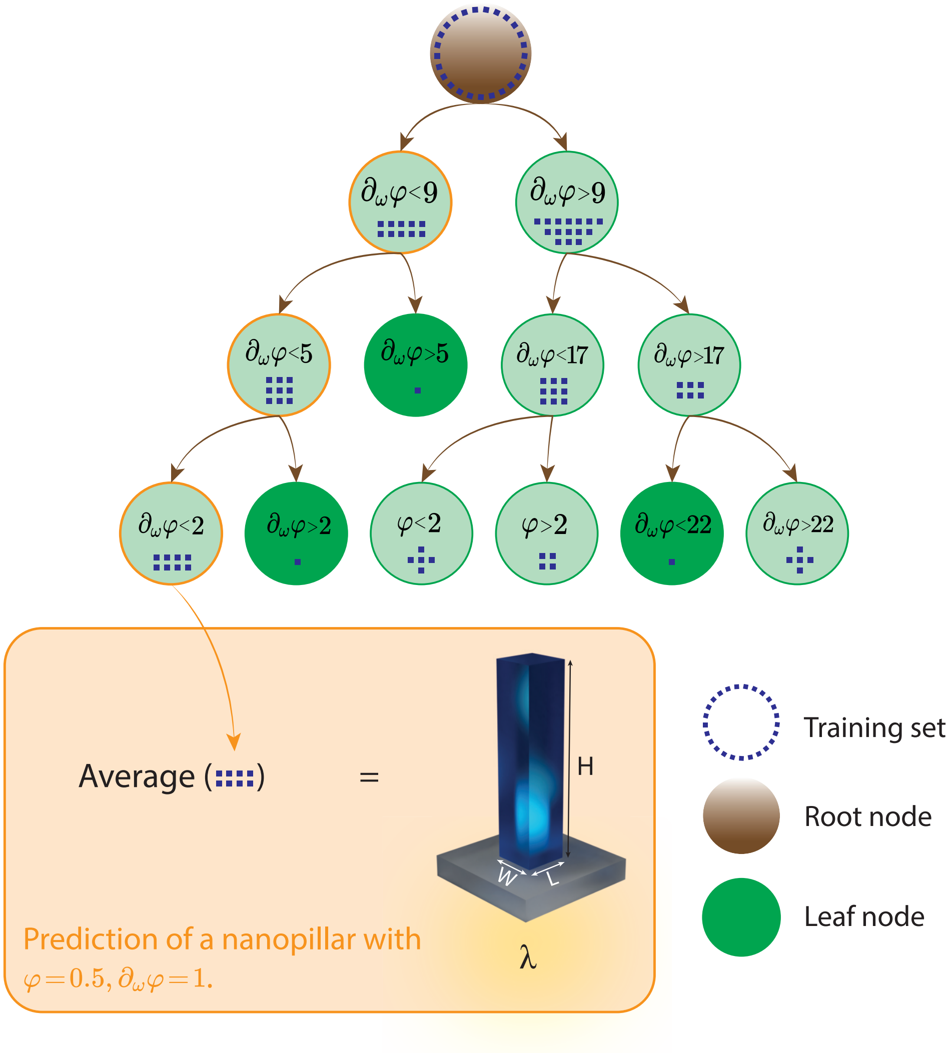


Figure 3.6: **Visualization of how decision trees train and make predictions.** For illustration purposes, the training process is demonstrated on a dataset consisting of 25 nanopillars. The actual design libraries will contain several thousands of nanopillar geometries. During training, the CART algorithm splits the training set into smaller subsets by asking true or false questions with respect to the input features. It proceeds iteratively till a given tree depth is reached or till it cannot divide the data any further. Once the tree structure has been established, predictions of new pillars can be made by traversing the tree from the root to the bottom nodes.

# Chapter 4

# Inverse design of achromatic metalenses

In this chapter, the decision tree-based inverse design of four achromatic metalenses will be presented. Throughout this chapter, the core workflow of most supervised machine learning problems will become visible. First, a design library will be prepared for each of the four metalenses. After that the training and testing of the decision trees follows on the previously established datasets. In the end, the trained models will be asked to predict four achromatic metalenses with given optical properties.

## 4.1 Preparation of the training sets

To demonstrate the applicability of the decision trees for the inverse design, four metalenses designed for achromatic, polarization-insensitive and diffraction-limited focusing across the entire biological transparency window ($700 - 1100\,\mathrm{nm}$) [94] and near-infrared telecommunication band ($1250 - 1650\,\mathrm{nm}$) were constructed. For the sake of better readability, the four metalenses are identified as symmetric ($700 - 1100\,\mathrm{nm}$), symmetric ($1250 - 1650\,\mathrm{nm}$), asymmetric ($700 - 1100\,\mathrm{nm}$) and asymmetric ($1250 - 1650\,\mathrm{nm}$). Each name provides information about the type of nanopillar used and the wavelength region addressed.

As the metalenses employed nanopillars occupying different parts of the parameter space, a tailor-made design library was prepared for each metalens. The datasets were generated based on an in-house rigorous coupled-wave analysis (RCWA) model and contain the phases, group delays and transmitted amplitudes in the co- and cross-polarization responses of the nanopillars. The parameter sweeps that were conducted for the metalenses in the lower (i.e. $700-1100\,\text{nm}$) and higher wavelength range (i.e. $1250-1650\,\text{nm}$) considered nanopillars made of an IP-L 780 polymer (Nanoscribe GmbH, Germany) with a constant period of $P = 1\,\mu\text{m}$ and $2\,\mu\text{m}$, a length of $L = 0.35-0.8\,\mu\text{m}$ and $0.5-1.6\,\mu\text{m}$, and a height of $H = 1-9\,\mu\text{m}$ and $1-15\,\mu\text{m}$, respectively. For the asymmetric metalenses, nanopillars with in-plane aspect ratios ranging from $R = 0.3-0.8$ were selected. The above-mentioned geometries were mapped to their associated responses for various incident wavelengths of light. By customizing the RCWA algorithm to run in parallel utilizing the parallel computing toolbox of MATLAB, the simulation of the nanopillars could be accelerated significantly. The resulting design libraries for the symmetric and asymmetric metalenses consisted of 1,125,000 and 4,320,000 data points and took approximately 16 and 66 hours to complete on a machine with a 3 GHz Intel Core i7-9700 processor. Fig. 4.1 shows the design library of the symmetric ($700-1100\,\text{nm}$) metalens, containing the phase and group delay responses of a total of 22,500 nanopillars[1]. Here, the phases of the co-polarization responses were simulated for a fixed incident wavelength of $\lambda = 1100\,\text{nm}$.

After applying some fabrications constraints, the design libraries of the symmetric ($700-1100\,\text{nm}$), symmetric ($1250-1650\,\text{nm}$), asymmetric ($700-1100\,\text{nm}$) and asymmetric ($1250-1650\,\text{nm}$) metalens were further restricted to nanopillars that exhibited a transmission efficiency of greater than 90%, 90%, 50% and 30%. Even if this added constraint reduced the size of the dataset considerably to approximately 350,000, 170,000, 300,000 and 90,000 points, it ensured that the decision trees, later, were trained on nanopillars with high transmission efficiencies only.

## 4.2 Training the decision trees

Before the training of the four decision trees started, the design libraries were randomly split into two subsets of 80% and 20% for training and testing purposes, respectively. The training was very quick, taking less than a second for each tree[2]. To be precise, for the symmetric ($700-1100\,\text{nm}$), symmetric ($1250-1650\,\text{nm}$), asymmetric ($700-1100\,\text{nm}$) and asymmetric ($1250-1650\,\text{nm}$) metalens, it completed, on average, in about

---

[1] Please refer to the Appendix A for the libraries of the remaining three metalenses.

[2] The same computer was used in the training as already in the preparation of the design libraries.

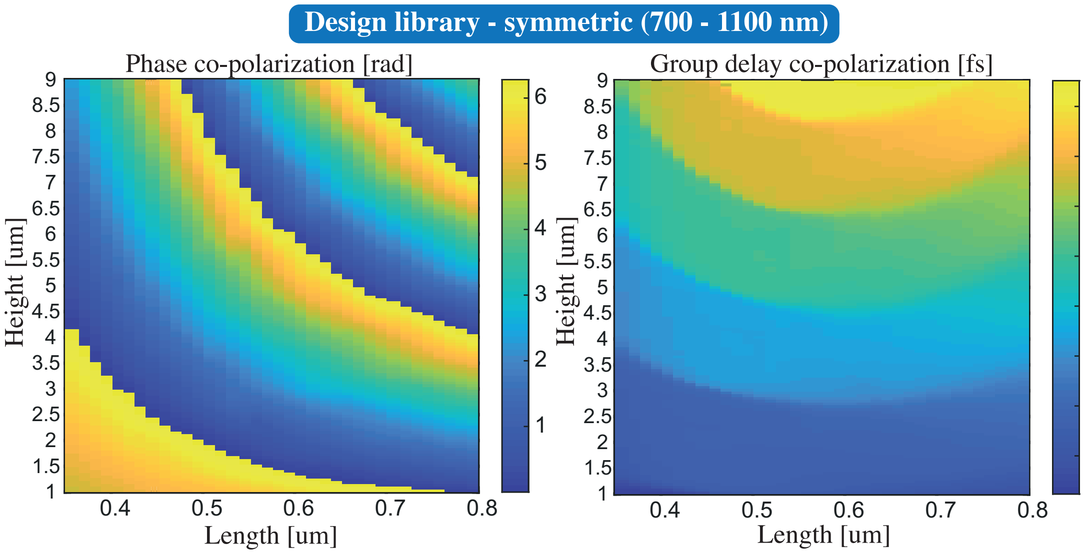


Figure 4.1: **Design library of the symmetric $(700 - 1100\,\mathrm{nm})$ metalens.** Phase and group delay in the co-polarization response of various nanopillar geometries. The phases were obtained for an incident wavelength of $\lambda = 1100\,\mathrm{nm}$. As the 3D nanopillars can be approximated as truncated waveguides, the group delays increase with the height of the pillars (see equation 2.1.1).

736 ms, 361 ms, 964 ms and 211 ms, respectively. Since no hyperparameters were applied, the decision trees could grow without any restrictions until all their bottom nodes turned into leaf nodes and could not be split further. The maximum tree depths that were reached during training can be seen in Fig. 4.2. After training, it became noticeable that out of the two input features (i.e. the phase and the group delay), the tree structures were mainly determined by the group delay, which reached a feature importance of above 99% for all four models. That means that 99% of the true or false questions asked to grow the trees related to the group delay.

## 4.3 Model evaluation

After the training, the decision trees were evaluated based on how well they could predict the previously unseen nanopillars from the test sets. As for performance metrics, both the coefficient of determination ($R^2$) and the root mean squared error (RMSE) were calculated. The RMSE is given by

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n}} \tag{4.1}$$

and the $R^2$ is defined as

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \hat{y})^2}. \tag{4.2}$$

Here, $y$ represents the labels, $\hat{y}$ the predictions and $\overline{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$ the mean of the labels. In the best case, when the predictions of a trained model exactly match the labels, $R^2$ reaches a value of 1. However, a constant model that always predicts $\overline{y}$ will have an $R^2$ of 0.

Remarkably, when evaluated on the test sets of the symmetric ($700 - 1100\,\text{nm}$), symmetric ($1250 - 1650\,\text{nm}$), asymmetric ($700 - 1100\,\text{nm}$) and asymmetric ($1250 - 1650\,\text{nm}$) metalens, the decision trees attained near-perfect $R^2$ scores of 0.996, 0.984, 0.767 and 0.804 at tree depths of 54, 51, 65 and 54, respectively (Fig. 4.2). To further demonstrate the predictive accuracy of the trees, the RMSEs of the target variables (i.e. the wavelength, the height, the length and the width of the nanopillar) were calculated at the aforementioned depths. A summary of the obtained RMSE values can be found in the insets of Fig. 4.2.

## 4.4 Inverse design of four different achromatic metalenses

In the final step, the trained decision trees were asked to predict the nanopillars that matched the phase and group delay requirements for an achromatic metalens[3].

The prediction of the appropriate nanopillars did not require much computational power and occurred near-instantaneously, taking approximately 4 ms for each metalens to complete. Subsequently, the predicted nanopillars were processed, and a CAD image for each metalens was generated (Fig. 4.3). The designs, in which the grayscale value provides information about the pillar height, can later be used by a faculty-owned commercial photolithography system (Nanoscribe Photonic Profession GT) to fabricate the metalenses via two-photon polymerization. If fabrication difficulties arise due to the large fill factor in the predicted metalenses, additional constraints could be imposed onto the

---

[3] The target phase profile is given by equation 2.17. The required group delay can be derived from equation 2.19.

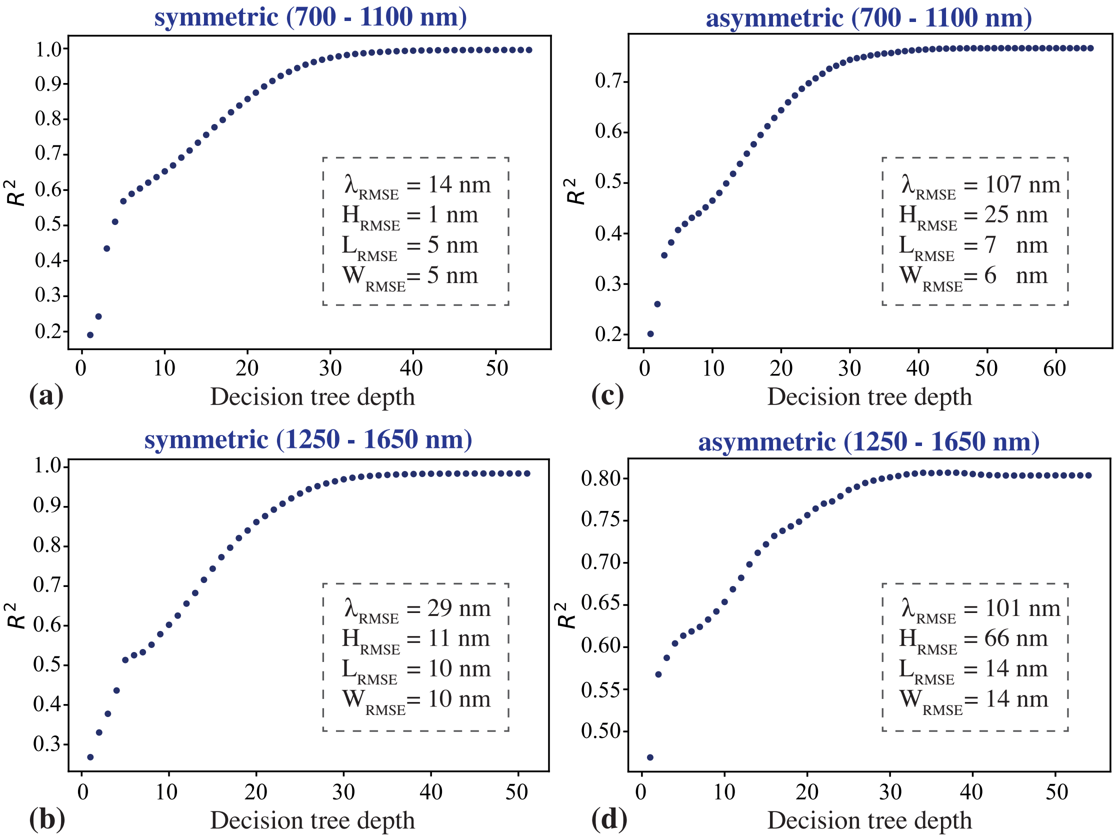


Figure 4.2: **Evaluating the trained decision trees on the unseen data of the test sets.** (a)-(d) The $R^2$ scores were calculated for each decision tree at varying tree depths. Maximum $R^2$ scores of (a) 0.996, (b) 0.984, (c) 0.767 and (d) 0.804 were achieved at tree depths of (a) 54, (b) 51, (c) 65 and (d) 54, respectively. The insets show the RMSEs of the target variables at the aforementioned tree depths.

training sets by allowing nanopillars with a certain cross-sectional area only.

To compare the achieved with the desired optical response, simulations of the designed achromatic metalenses were carried out with the RCWA model. The simulations of the predicted meta-atoms show that the required and the obtained group delays agree very well for all four metalenses - so do the phases, which, however, reveal some noticeable deviations in the center of the lenses (Fig. 4.4). The diameters of the metalenses were limited by the group delays accessible in the meta-unit libraries (see equation 2.13). Since the datasets of the symmetric ($700 - 1100\,\mathrm{nm}$), symmetric ($1250 - 1650\,\mathrm{nm}$), asymmetric ($700 - 1100\,\mathrm{nm}$) and asymmetric ($1250 - 1650\,\mathrm{nm}$) metalens provided group delays with a range of $17\,\mathrm{fs}, 28\,\mathrm{fs}, 13\,\mathrm{fs}$ and $26\,\mathrm{fs}$, the maximal lens diameters were restricted to $94\,\mu\mathrm{m}, 120\,\mu\mathrm{m}, 62\,\mu\mathrm{m}$ and $88\,\mu\mathrm{m}$ resulting in numerical apertures (NA) of 0.10, 0.13, 0.12 and 0.17, respectively. Strikingly, as the training sets were constrained to efficient nanopillars only, the simulated metalenses exhibited transmission efficiencies greater than 90%, 90%, 50% and 30%, respectively.

Finally, to find out whether the proposed decision trees select the desired nanostructures more efficiently from the library than conventional brute-force methods, the runtime for designing a metalens was compared between the two approaches. Both the decision tree and the in-house implemented optimization algorithm were asked to produce the asymmetric ($700 - 1100\,\mathrm{nm}$) and asymmetric ($1250 - 1650\,\mathrm{nm}$) metalens based on a dataset containing a total of 108,000 nanopillars. Although both methods used the same machine, the decision tree completed the design of the asymmetric ($700 - 1100\,\mathrm{nm}$) and asymmetric ($1250 - 1650\,\mathrm{nm}$) metalens, on average, in $319\,\mathrm{ms}$ and $297\,\mathrm{ms}$, whereas the optimization algorithm took around $64\,\mathrm{s}$ and $20\,\mathrm{s}$. This result clearly shows the potential decision trees have in accelerating the design process presenting an up to 200 times faster inverse retrieval over conventional brute-force techniques.

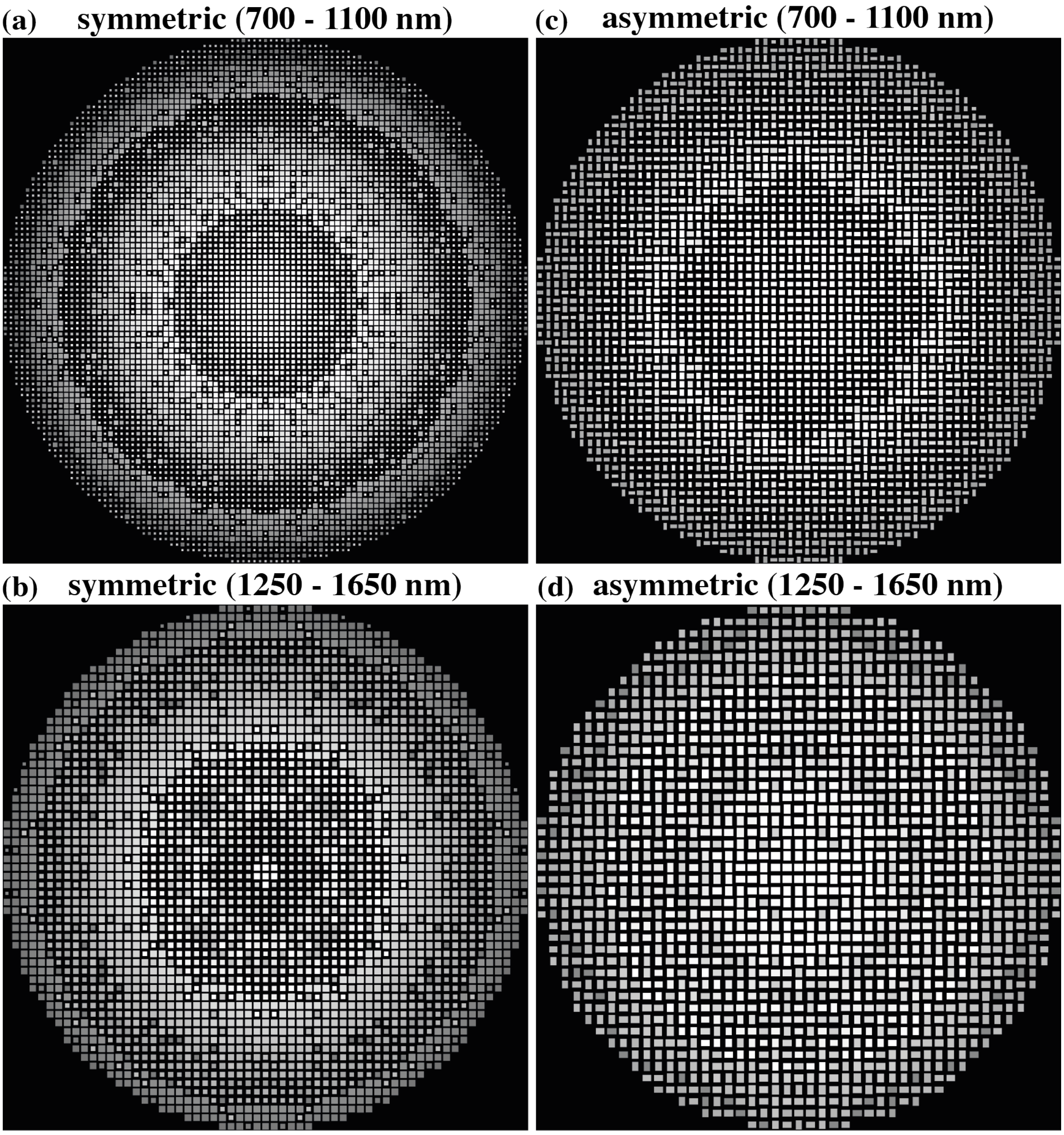


Figure 4.3: **CAD images of the machine learning predicted achromatic metalenses.** The metalenses were designed for achromatic and polarization-insensitive performance across (a), (c) 700 – 1100 nm and (b), (d) 1250 – 1650 nm, and were based on (a), (b) symmetric and (c), (d) asymmetric nanopillars. The lenses have diameters of (a) 94 µm, (b) 120 µm, (c) 62 µm and (d) 88 µm and numerical apertures of (a) 0.10, (b) 0.13, (c) 0.12 and (d) 0.17. According to the simulation, the metalenses reach transmission efficiencies greater than (a) 90%, (b) 90%, (c) 50% and (d) 30%. In the CAD image, the grayscale values relate to the height of the nanopillars, where the brighter tones refer to the taller and the darker tones to the smaller pillars.

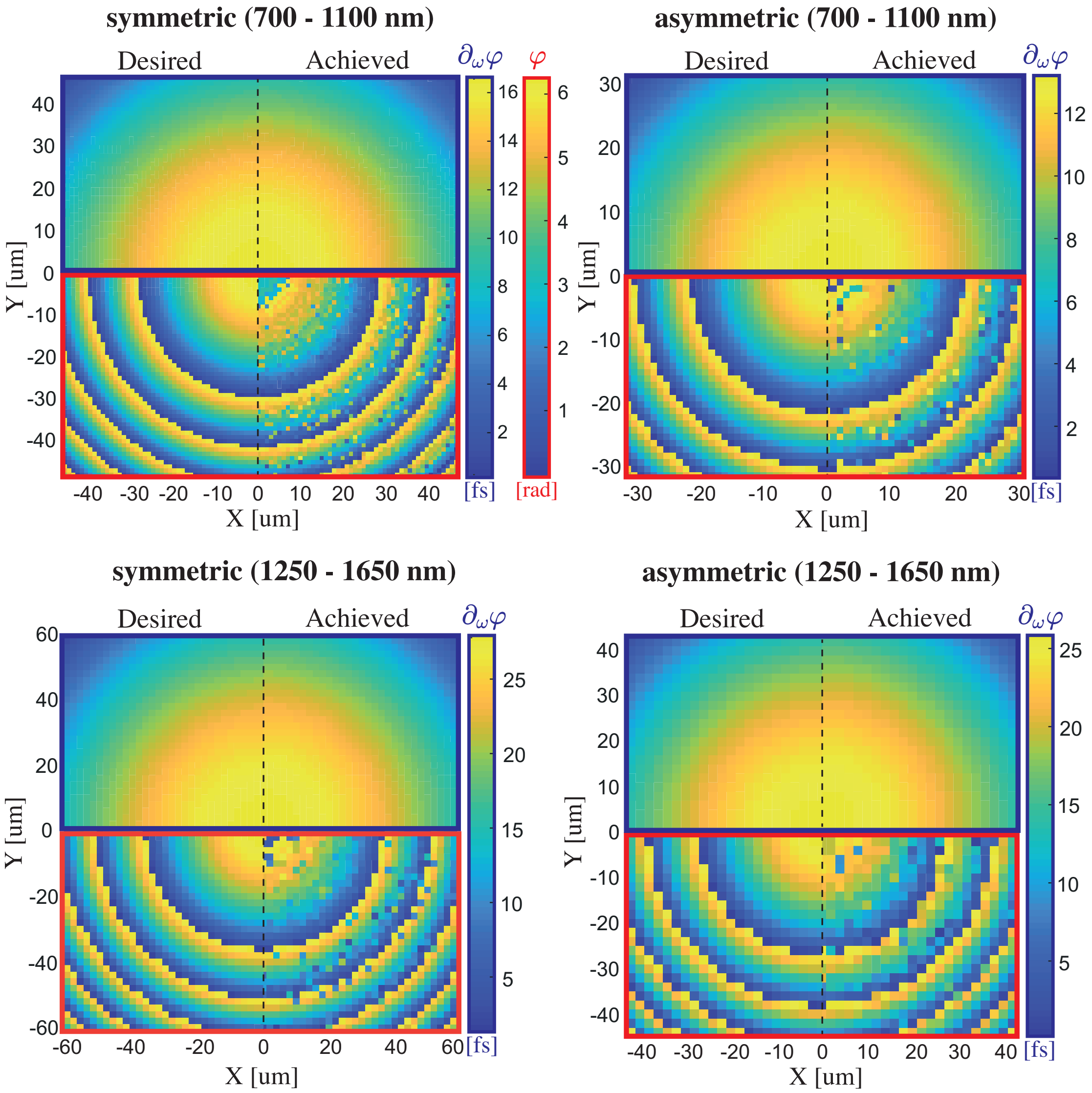


Figure 4.4: **Comparison of the achieved and desired optical responses.** The predicted metalenses were simulated based on the RCWA model, and their optical responses were subsequently compared to the target phase and group delay profiles. The red (blue) boxes denote the desired and achieved phase (group delay) profiles. The black dashed line in the center of each metalens separates the target from the achieved responses. The colour bar of the phase occurs only once since it is the same for all four metalenses.

# Chapter 5

# Conclusion and Outlook

The overall goal of this thesis was to incorporate data-driven machine learning methods into the inverse design of achromatic metalenses. More specifically, it was examined whether machine learning models could more efficiently identify the metalens nanostructures from the design library than conventional brute-force based optimization algorithms. Furthermore, it was studied whether the trained models could be used for the on-demand design of achromatic metalenses with arbitrary NAs and bandwidths after they learned the critical nanostructure-performance relationship from the design library.

During the course of this work, it was found that decision trees were most suitable for learning the nontrivial mapping between the employed nanopillars and their corresponding phases and group delays. Compared to neural networks, decision trees were easier to implement, faster to train and, more crucially, showed greater accuracies in predicting the nanopillar dimensions. However, as the predicted nanopillars are just averages of already existing pillar geometries in the training set, decision trees can only design metalenses that fall within the parameter space of their respective design libraries. Hence, to realize the design of four achromatic metalenses with various numerical apertures and distinct operating bandwidths, a custom-made design library had to be assembled for each metalens. This apparent weakness of the decision tree can also be seen as its strength since it facilitated the application of fabrication constraints and enabled precise control on the efficiency of the designed metalenses. Results showed that the transmission efficiency thresholds imposed onto the nanostructure libraries proved effective as they directed the decision trees to predict highly efficient metalenses only.

While decision trees could successfully learn the structure-performance relationship, they treated the metalens design as a regression problem and thus produced only one design for a given optical response. This one-to-one mapping of metalenses contrasts physical intuition and could potentially be overcome by employing generative models such as generative adversarial networks and variational autoencoders. Generative models are worth considering for future work also because they have shown to predict human-unintuitive nanophotonic structures with promising optical performances while not having to rely on a labelled dataset. Loosening the dependency on a computationally expensive training set is an important endeavour, especially as increasingly complex meta-atoms are considered in the construction of achromatic metalenses.

To conclude, the simulations of all four predicted achromatic metalenses agreed well with the desired optical responses derived from the lens equation. Remarkably, the trained decision trees could not only produce the appropriate designs, but they did so up to 200 times faster than conventional brute-force methods. This drastic speed advantage over conventional optimization algorithms becomes particularly noticeable when metalenses reach millimeter- to centimeter scales, thereby causing the number of the constituent subwavelength meta-atoms to rise significantly.

# Bibliography

[1] S.-W. Moon, Y. Kim, G. Yoon, and J. Rho, “Recent Progress on Ultrathin Metalenses for Flat Optics,” *Iscience*, p. 101877, 2020.

[2] G. Zheng, H. Mühlenbernd, M. Kenney, G. Li, T. Zentgraf, and S. Zhang, “Metasurface holograms reaching 80% efficiency,” *Nature nanotechnology*, vol. 10, no. 4, pp. 308–312, 2015.

[3] N. Yu, F. Aieta, P. Genevet, M. A. Kats, Z. Gaburro, and F. Capasso, “A broadband, background-free quarter-wave plate based on plasmonic metasurfaces,” *Nano letters*, vol. 12, no. 12, pp. 6328–6333, 2012.

[4] L. Hsu, T. Lepetit, and B. Kanté, “Extremely thin dielectric metasurface for carpet cloaking,” *arXiv preprint arXiv:1503.08486*, 2015.

[5] R. C. Devlin, A. Ambrosio, N. A. Rubin, J. B. Mueller, and F. Capasso, “Arbitrary spin-to–orbital angular momentum conversion of light,” *Science*, vol. 358, no. 6365, pp. 896–901, 2017.

[6] Q. Ding, S. F. Barna, K. Jacobs, A. Choubal, G. Mensing, Z. Zhang, K. Yamada, N. Kincaid, G. Zhu, R. Tirawat, *et al.*, “Feasibility analysis of nanostructured planar focusing collectors for concentrating solar power applications,” *ACS Applied Energy Materials*, vol. 1, no. 12, pp. 6927–6935, 2018.

[7] A. M. Shaltout, K. G. Lagoudakis, J. van de Groep, S. J. Kim, J. Vučković, V. M. Shalaev, and M. L. Brongersma, “Spatiotemporal light control with frequency-gradient metasurfaces,” *Science*, vol. 365, no. 6451, pp. 374–377, 2019.

[8] F. Balli, M. Sultan, S. K. Lami, and J. T. Hastings, “A hybrid achromatic metalens,” *Nature communications*, vol. 11, no. 1, pp. 1–8, 2020.

[9] W. T. Chen, A. Y. Zhu, and F. Capasso, “Flat optics with dispersion-engineered metasurfaces,” *Nature Reviews Materials*, vol. 5, no. 8, pp. 604–620, 2020.

[10] C. Zhang, S. Divitt, Q. Fan, W. Zhu, A. Agrawal, Y. Lu, T. Xu, and H. J. Lezec, “Low-loss metasurface optics down to the deep ultraviolet region,” *Light: Science & Applications*, vol. 9, no. 1, pp. 1–10, 2020.

[11] R. C. Devlin, M. Khorasaninejad, W. T. Chen, J. Oh, and F. Capasso, “Broadband high-efficiency dielectric metasurfaces for the visible spectrum,” *Proceedings of the National Academy of Sciences*, vol. 113, no. 38, pp. 10473–10478, 2016.

[12] Y. Liang, H. Liu, F. Wang, H. Meng, J. Guo, J. Li, and Z. Wei, “High-efficiency, near-diffraction limited, dielectric metasurface lenses based on crystalline titanium dioxide at visible wavelengths,” *Nanomaterials*, vol. 8, no. 5, p. 288, 2018.

[13] G. Brière, P. Ni, S. Héron, S. Chenot, S. Vézian, V. Brändli, B. Damilano, J.-Y. Duboz, M. Iwanaga, and P. Genevet, “An etching-free approach toward large-scale light-emitting metasurfaces,” *Advanced Optical Materials*, vol. 7, no. 14, p. 1801271, 2019.

[14] B. H. Chen, P. C. Wu, V.-C. Su, Y.-C. Lai, C. H. Chu, I. C. Lee, J.-W. Chen, Y. H. Chen, Y.-C. Lan, C.-H. Kuan, *et al.*, “Gan metalens for pixel-level full-color routing at visible light,” *Nano letters*, vol. 17, no. 10, pp. 6345–6352, 2017.

[15] D. Sell, J. Yang, S. Doshay, K. Zhang, and J. A. Fan, “Visible light metasurfaces based on single-crystal silicon,” *Acs Photonics*, vol. 3, no. 10, pp. 1919–1925, 2016.

[16] Z. Zhou, J. Li, R. Su, B. Yao, H. Fang, K. Li, L. Zhou, J. Liu, D. Stellinga, C. P. Reardon, *et al.*, “Efficient silicon metasurfaces for visible light,” *Acs Photonics*, vol. 4, no. 3, pp. 544–551, 2017.

[17] B. Wang, F. Dong, Q.-T. Li, D. Yang, C. Sun, J. Chen, Z. Song, L. Xu, W. Chu, Y.-F. Xiao, *et al.*, “Visible-frequency dielectric metasurfaces for multiwavelength achromatic and highly dispersive holograms,” *Nano letters*, vol. 16, no. 8, pp. 5235–5240, 2016.

[18] M. Khorasaninejad, W. T. Chen, R. C. Devlin, J. Oh, A. Y. Zhu, and F. Capasso, “Metalenses at visible wavelengths: Diffraction-limited focusing and subwavelength resolution imaging,” *Science*, vol. 352, no. 6290, pp. 1190–1194, 2016.

[19] H. Liang, Q. Lin, X. Xie, Q. Sun, Y. Wang, L. Zhou, L. Liu, X. Yu, J. Zhou, T. F. Krauss, *et al.*, "Ultrahigh numerical aperture metalens at visible wavelengths," *Nano letters*, vol. 18, no. 7, pp. 4460–4466, 2018.

[20] J. Engelberg, C. Zhou, N. Mazurski, J. Bar-David, U. Levy, and A. Kristensen, "Near-ir wide field-of-view huygens metalens for outdoor imaging applications," in *2019 Conference on Lasers and Electro-Optics (CLEO)*, pp. 1–2, IEEE, 2019.

[21] A. Arbabi, E. Arbabi, S. M. Kamali, Y. Horie, S. Han, and A. Faraon, "Miniature optical planar camera based on a wide-angle metasurface doublet corrected for monochromatic aberrations," *Nature communications*, vol. 7, no. 1, pp. 1–9, 2016.

[22] Q. Fan, M. Liu, C. Yang, L. Yu, F. Yan, and T. Xu, "A high numerical aperture, polarization-insensitive metalens for long-wavelength infrared imaging," *Applied Physics Letters*, vol. 113, no. 20, p. 201104, 2018.

[23] H. Zuo, D.-Y. Choi, X. Gai, P. Ma, L. Xu, D. N. Neshev, B. Zhang, and B. Luther-Davies, "High-efficiency all-dielectric metalenses for mid-infrared imaging," *Advanced Optical Materials*, vol. 5, no. 23, p. 1700585, 2017.

[24] C. Yan, X. Li, M. Pu, X. Ma, F. Zhang, P. Gao, K. Liu, and X. Luo, "Midinfrared real-time polarization imaging with all-dielectric metasurfaces," *Applied Physics Letters*, vol. 114, no. 16, p. 161904, 2019.

[25] L. Zhang, J. Ding, H. Zheng, S. An, H. Lin, B. Zheng, Q. Du, G. Yin, J. Michon, Y. Zhang, *et al.*, "Ultra-thin high-efficiency mid-infrared transmissive huygens meta-optics," *Nature communications*, vol. 9, no. 1, pp. 1–9, 2018.

[26] M. Y. Shalaginov, S. An, F. Yang, P. Su, D. Lyzwa, A. Agarwal, H. Zhang, J. Hu, and T. Gu, "A single-layer panoramic metalens with¿ 170 {\deg} diffraction-limited field of view," *arXiv preprint arXiv:1908.03626*, 2019.

[27] P. Wang, N. Mohammad, and R. Menon, "Chromatic-aberration-corrected diffractive lenses for ultra-broadband focusing," *Scientific reports*, vol. 6, no. 1, pp. 1–7, 2016.

[28] M. Khorasaninejad, Z. Shi, A. Y. Zhu, W.-T. Chen, V. Sanjeev, A. Zaidi, and F. Capasso, "Achromatic metalens over 60 nm bandwidth in the visible and metalens with reverse chromatic dispersion," *Nano letters*, vol. 17, no. 3, pp. 1819–1824, 2017.

[29] S. Wang, P. C. Wu, V.-C. Su, Y.-C. Lai, C. H. Chu, J.-W. Chen, S.-H. Lu, J. Chen, B. Xu, C.-H. Kuan, *et al.*, “Broadband achromatic optical metasurface devices,” *Nature communications*, vol. 8, no. 1, pp. 1–9, 2017.

[30] S. Wang, P. C. Wu, V.-C. Su, Y.-C. Lai, M.-K. Chen, H. Y. Kuo, B. H. Chen, Y. H. Chen, T.-T. Huang, J.-H. Wang, *et al.*, “A broadband achromatic metalens in the visible,” *Nature nanotechnology*, vol. 13, no. 3, pp. 227–232, 2018.

[31] W. T. Chen, A. Y. Zhu, V. Sanjeev, M. Khorasaninejad, Z. Shi, E. Lee, and F. Capasso, “A broadband achromatic metalens for focusing and imaging in the visible,” *Nature nanotechnology*, vol. 13, no. 3, pp. 220–226, 2018.

[32] S. Shrestha, A. C. Overvig, M. Lu, A. Stein, and N. Yu, “Broadband achromatic dielectric metalenses,” *Light: Science & Applications*, vol. 7, no. 1, pp. 1–11, 2018.

[33] E. Arbabi, A. Arbabi, S. M. Kamali, Y. Horie, and A. Faraon, “Multiwavelength polarization-insensitive lenses based on dielectric metasurfaces with metamolecules,” *Optica*, vol. 3, no. 6, pp. 628–633, 2016.

[34] E. Arbabi, A. Arbabi, S. M. Kamali, Y. Horie, and A. Faraon, “Controlling the sign of chromatic dispersion in diffractive optics with dielectric metasurfaces,” *Optica*, vol. 4, no. 6, pp. 625–632, 2017.

[35] W. T. Chen, A. Y. Zhu, J. Sisler, Z. Bharwani, and F. Capasso, “A broadband achromatic polarization-insensitive metalens consisting of anisotropic nanostructures,” *Nature communications*, vol. 10, no. 1, pp. 1–7, 2019.

[36] A. Ndao, L. Hsu, J. Ha, J.-H. Park, C. Chang-Hasnain, and B. Kanté, “Octave bandwidth photonic fishnet-achromatic-metalens,” *Nature communications*, vol. 11, no. 1, pp. 1–6, 2020.

[37] F. Balli, M. A. Sultan, A. Ozdemir, and J. T. Hastings, “An ultrabroadband 3d achromatic metalens,” *Nanophotonics*, vol. 10, no. 4, pp. 1259–1264, 2021.

[38] M. A.C., “How checkers was solved. The Atlantic..” `https://www.theatlantic.com/technology/archive/2017/07/marion-tinsley-checkers/534111/`, July 2017.

[39] M. Campbell, A. Hoane, and F. hsiung Hsu, “Deep blue,” *Artificial Intelligence*, vol. 134, no. 1, pp. 57–83, 2002.

[40] D. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. Van Den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot, *et al.*, "Mastering the game of go with deep neural networks and tree search," *nature*, vol. 529, no. 7587, pp. 484–489, 2016.

[41] A. W. Senior, R. Evans, J. Jumper, J. Kirkpatrick, L. Sifre, T. Green, C. Qin, A. Žídek, A. W. Nelson, A. Bridgland, *et al.*, "Improved protein structure prediction using potentials from deep learning," *Nature*, vol. 577, no. 7792, pp. 706–710, 2020.

[42] J. Petschulat, C. Menzel, A. Chipouline, C. Rockstuhl, A. Tünnermann, F. Lederer, and T. Pertsch, "Multipole approach to metamaterials," *Physical Review A*, vol. 78, no. 4, p. 043811, 2008.

[43] B. Gallinet, J. Butet, and O. J. Martin, "Numerical methods for nanophotonics: standard problems and future challenges," *Laser & Photonics Reviews*, vol. 9, no. 6, pp. 577–603, 2015.

[44] S. So, T. Badloe, J. Noh, J. Bravo-Abad, and J. Rho, "Deep learning enabled inverse design in nanophotonics," *Nanophotonics*, vol. 9, no. 5, pp. 1041–1057, 2020.

[45] M. C. Troparevsky, A. S. Sabau, A. R. Lupini, and Z. Zhang, "Transfer-matrix formalism for the calculation of optical response in multilayer systems: from coherent to incoherent interference," *Optics express*, vol. 18, no. 24, pp. 24715–24721, 2010.

[46] J. Gou, H. Cansizoglu, C. Bartolo-Perez, S. Ghandiparsi, A. S. Mayet, H. Rabiee-Golgir, Y. Gao, J. Wang, T. Yamada, E. P. Devine, A. F. Elrefaie, S.-Y. Wang, and M. S. Islam, "Rigorous coupled-wave analysis of absorption enhancement in vertically illuminated silicon photodiodes with photon-trapping hole arrays," *Nanophotonics*, vol. 8, no. 10, pp. 1747–1756, 2019.

[47] J. Peurifoy, Y. Shen, L. Jing, Y. Yang, F. Cano-Renteria, B. G. DeLacy, J. D. Joannopoulos, M. Tegmark, and M. Soljačić, "Nanophotonic particle simulation and inverse design using artificial neural networks," *Science advances*, vol. 4, no. 6, p. eaar4206, 2018.

[48] D. Liu, Y. Tan, E. Khoram, and Z. Yu, "Training deep neural networks for the inverse design of nanophotonic structures," *Acs Photonics*, vol. 5, no. 4, pp. 1365–1369, 2018.

[49] W. Ma, Z. Liu, Z. A. Kudyshev, A. Boltasseva, W. Cai, and Y. Liu, "Deep learning for the design of photonic structures," *Nature Photonics*, pp. 1–14, 2020.

[50] Y. Qu, L. Jing, Y. Shen, M. Qiu, and M. Soljacic, “Migrating knowledge between physical scenarios based on artificial neural networks,” *ACS Photonics*, vol. 6, no. 5, pp. 1168–1174, 2019.

[51] S. Inampudi and H. Mosallaei, “Neural network based design of metagratings,” *Applied Physics Letters*, vol. 112, no. 24, p. 241102, 2018.

[52] I. Sajedian, J. Kim, and J. Rho, “Finding the optical properties of plasmonic structures by image processing using a combination of convolutional neural networks and recurrent neural networks,” *Microsystems & nanoengineering*, vol. 5, no. 1, pp. 1–8, 2019.

[53] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778, 2016.

[54] C. F. Bohren and D. R. Huffman, *Absorption and scattering of light by small particles.* John Wiley & Sons, 2008.

[55] F. Martın, J. Bonache, F. a. Falcone, M. Sorolla, and R. Marqués, “Split ring resonator-based left-handed coplanar waveguide,” *Applied Physics Letters*, vol. 83, no. 22, pp. 4652–4654, 2003.

[56] W. Zhang, W. Zhu, E. Chia, Z. Shen, H. Cai, Y. Gu, W. Ser, and A. Liu, “A pseudo-planar metasurface for a polarization rotator,” *Optics express*, vol. 22, no. 9, pp. 10446–10454, 2014.

[57] P. Melchior, D. Bayer, C. Schneider, A. Fischer, M. Rohmer, W. Pfeiffer, and M. Aeschlimann, “Optical near-field interference in the excitation of a bowtie nanoantenna,” *Physical Review B*, vol. 83, no. 23, p. 235407, 2011.

[58] R. Blanchard, G. Aoust, P. Genevet, N. Yu, M. A. Kats, Z. Gaburro, and F. Capasso, “Modeling nanoscale v-shaped antennas for the design of optical phased arrays,” *Physical Review B*, vol. 85, no. 15, p. 155457, 2012.

[59] D. E. Goldberg and J. H. Holland, “Genetic algorithms and machine learning,” 1988.

[60] J. Robinson and Y. Rahmat-Samii, “Particle swarm optimization in electromagnetics,” *IEEE transactions on antennas and propagation*, vol. 52, no. 2, pp. 397–407, 2004.

[61] M. P. Bendsoe and O. Sigmund, *Topology optimization: theory, methods, and applications.* Springer Science & Business Media, 2013.

[62] I. Malkiel, M. Mrejen, A. Nagler, U. Arieli, L. Wolf, and H. Suchowski, "Plasmonic nanostructure design and characterization via deep learning," *Light: Science & Applications*, vol. 7, no. 1, pp. 1–8, 2018.

[63] W. Ma, F. Cheng, and Y. Liu, "Deep-learning-enabled on-demand design of chiral metamaterials," *ACS nano*, vol. 12, no. 6, pp. 6326–6334, 2018.

[64] Z. Liu, D. Zhu, S. P. Rodrigues, K.-T. Lee, and W. Cai, "Generative model for the inverse design of metasurfaces," *Nano letters*, vol. 18, no. 10, pp. 6570–6576, 2018.

[65] Z. Liu, D. Zhu, L. Raju, and W. Cai, "Tackling photonic inverse design with machine learning," *Advanced Science*, vol. 8, no. 5, p. 2002923, 2021.

[66] W. Ma, F. Cheng, Y. Xu, Q. Wen, and Y. Liu, "Probabilistic representation and inverse design of metamaterials based on a deep generative model with semi-supervised learning strategy," *Advanced Materials*, vol. 31, no. 35, p. 1901111, 2019.

[67] J. Jiang and J. A. Fan, "Global optimization of dielectric metasurfaces using a physics-driven neural network," *Nano letters*, vol. 19, no. 8, pp. 5366–5372, 2019.

[68] F. Presutti and F. Monticone, "Focusing on bandwidth: achromatic metalens limits," *Optica*, vol. 7, no. 6, pp. 624–631, 2020.

[69] M. Khorasaninejad and F. Capasso, "Metalenses: Versatile multifunctional photonic components," *Science*, vol. 358, no. 6367, 2017.

[70] M. L. Tseng, H.-H. Hsiao, C. H. Chu, M. K. Chen, G. Sun, A.-Q. Liu, and D. P. Tsai, "Metalenses: advances and applications," *Advanced Optical Materials*, vol. 6, no. 18, p. 1800554, 2018.

[71] P. Lalanne and P. Chavel, "Metalenses at visible wavelengths: past, present, perspectives," *Laser & Photonics Reviews*, vol. 11, no. 3, p. 1600295, 2017.

[72] J. Chen, F. Zhang, Q. Li, J. Wu, and L. Wu, "A high-efficiency dual-wavelength achromatic metalens based on pancharatnam-berry phase manipulation," *Optics express*, vol. 26, no. 26, pp. 34919–34927, 2018.

[73] P. Genevet, F. Capasso, F. Aieta, M. Khorasaninejad, and R. Devlin, "Recent advances in planar optics: from plasmonic to dielectric metasurfaces," *Optica*, vol. 4, no. 1, pp. 139–152, 2017.

[74] P. R. West, J. L. Stewart, A. V. Kildishev, V. M. Shalaev, V. V. Shkunov, F. Strohkendl, Y. A. Zakharenkov, R. K. Dodds, and R. Byren, "All-dielectric subwavelength metasurface focusing lens," *Optics express*, vol. 22, no. 21, pp. 26212–26221, 2014.

[75] M. Khorasaninejad, A. Y. Zhu, C. Roques-Carmes, W. T. Chen, J. Oh, I. Mishra, R. C. Devlin, and F. Capasso, "Polarization-insensitive metalenses at visible wavelengths," *Nano letters*, vol. 16, no. 11, pp. 7229–7234, 2016.

[76] Y. Li, X. Li, M. Pu, Z. Zhao, X. Ma, Y. Wang, and X. Luo, "Achromatic flat optical components via compensation between structure and material dispersions," *Scientific Reports*, vol. 6, no. 1, pp. 1–7, 2016.

[77] F. Zhang, M. Zhang, J. Cai, Y. Ou, and H. Yu, "Metasurfaces for broadband dispersion engineering through custom-tailored multi-resonances," *Applied Physics Express*, vol. 11, no. 8, p. 082004, 2018.

[78] S. A. Maier, *Plasmonics: fundamentals and applications.* Springer Science & Business Media, 2007.

[79] X. Luo, "Principles of electromagnetic waves in metasurfaces," *Science China Physics, Mechanics & Astronomy*, vol. 58, no. 9, pp. 1–18, 2015.

[80] M. V. Berry, "The adiabatic phase and pancharatnam's phase for polarized light," *Journal of Modern Optics*, vol. 34, no. 11, pp. 1401–1407, 1987.

[81] S. Pancharatnam, "Generalized theory of interference and its applications," in *Proceedings of the Indian Academy of Sciences-Section A*, vol. 44, pp. 398–417, Springer, 1956.

[82] L. Nikolova and P. S. Ramanujam, *Polarization holography.* Cambridge University Press, 2009.

[83] F. Aieta, P. Genevet, M. Kats, and F. Capasso, "Aberrations of flat lenses and aplanatic metasurfaces," *Optics express*, vol. 21, no. 25, pp. 31530–31539, 2013.

[84] A. She, S. Zhang, S. Shian, D. R. Clarke, and F. Capasso, "Large area metalenses: design, characterization, and mass manufacturing," *Optics express*, vol. 26, no. 2, pp. 1573–1585, 2018.

[85] M. Kenney, J. Grant, D. Hao, K. Docherty, G. Mills, G. Jeffrey, D. Macleod, D. Henry, P. MacKay, M. Sorel, *et al.*, "Large area metasurface lenses in the nir

region," in *Modeling Aspects in Optical Metrology VII*, vol. 11057, p. 110570C, International Society for Optics and Photonics, 2019.

[86] Z. A. Kudyshev, A. V. Kildishev, V. M. Shalaev, and A. Boltasseva, "Machine-learning-assisted metasurface design for high-efficiency thermal emitter optimization," *Applied Physics Reviews*, vol. 7, no. 2, p. 021407, 2020.

[87] A. L. Hodgkin and A. F. Huxley, "A quantitative description of membrane current and its application to conduction and excitation in nerve," *The Journal of physiology*, vol. 117, no. 4, pp. 500–544, 1952.

[88] G. Cybenko, "Approximation by superpositions of a sigmoidal function," *Mathematics of control, signals and systems*, vol. 2, no. 4, pp. 303–314, 1989.

[89] A. Géron, "Hands-on machine learning with scikit-learn, keras, and tensorflow: Concepts, tools, and techniques to build intelligent systems," 2017.

[90] R. J., "Understanding variational autoencoders (VAEs). towards data science.." `https://towardsdatascience.com/understanding-variational-autoencoders-vaes-f70510919f73`, Sept. 2019.

[91] O'Reilly. (2017, May 25), "What Kaggle has learned from almost a million data scientists – Anthony Goldbloom (Kaggle) [Video]. YouTube.." `https://youtu.be/jmHbS8z57yI`.

[92] Biewald, L. [Weights & Biases]. (2020, September 9), "How to win Kaggle competitions with Anthony Goldbloom [Video]. YouTube.." `https://youtu.be/0ZJQ2Vsgwf0`.

[93] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, *et al.*, "Scikit-learn: Machine learning in python," *the Journal of machine Learning research*, vol. 12, pp. 2825–2830, 2011.

[94] S. Chinnathambi and N. Shirahata, "Recent advances on fluorescent biomarkers of near-infrared quantum dots for in vitro and in vivo imaging," *Science and technology of advanced materials*, vol. 20, no. 1, pp. 337–355, 2019.

# List of Figures

# Appendix A

# Inverse design of achromatic metalenses

## A.1 Preparation of the training sets

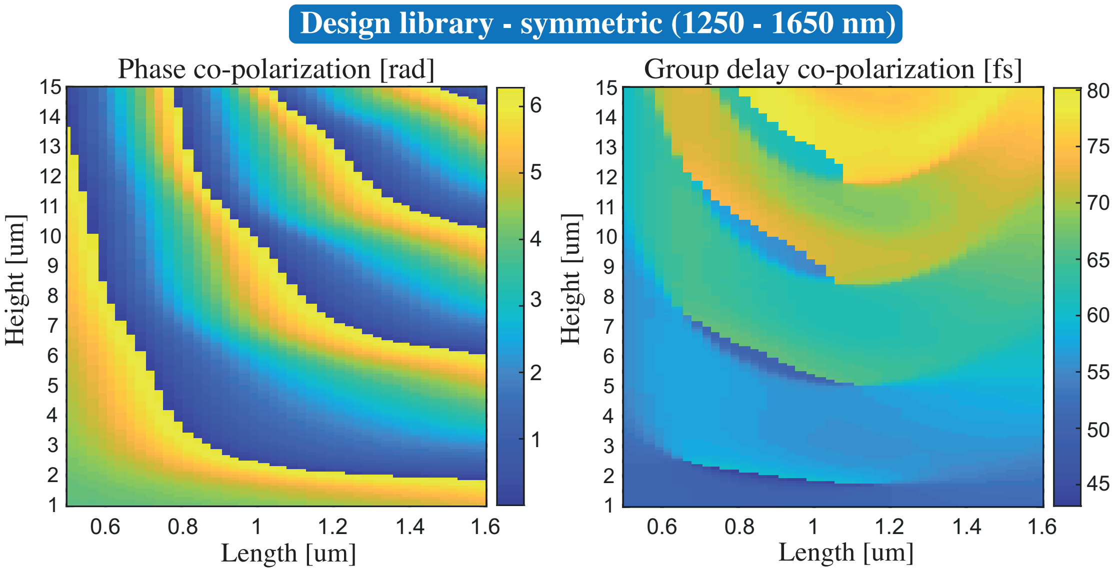


Figure A.1: **Design library of the symmetric ($1250 - 1650\,\mathrm{nm}$) metalens.** Phases and group delays in the co-polarization response of various nanopillar geometries. The phases were obtained for an incident wavelength of $\lambda = 1650\,\mathrm{nm}$.

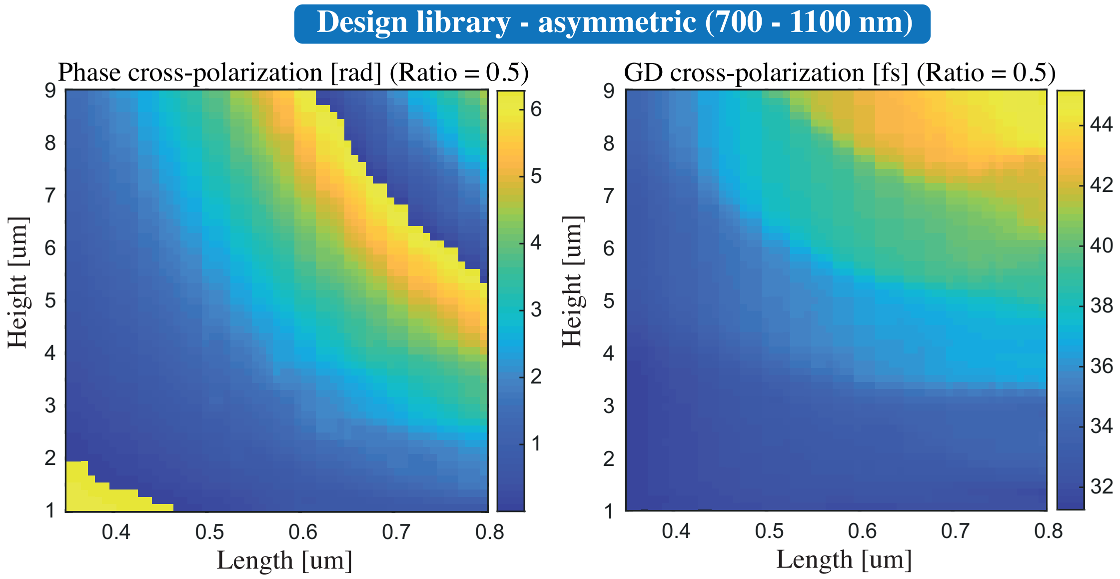


Figure A.2: **Design library of the asymmetric (700 − 1100 nm) metalens.** Phases and group delays in the cross-polarization response of various nanopillar geometries. The phases were obtained for an incident wavelength of $\lambda = 1100\,\mathrm{nm}$. The in-plane aspect ratio was fixed to $R = 0.5$.

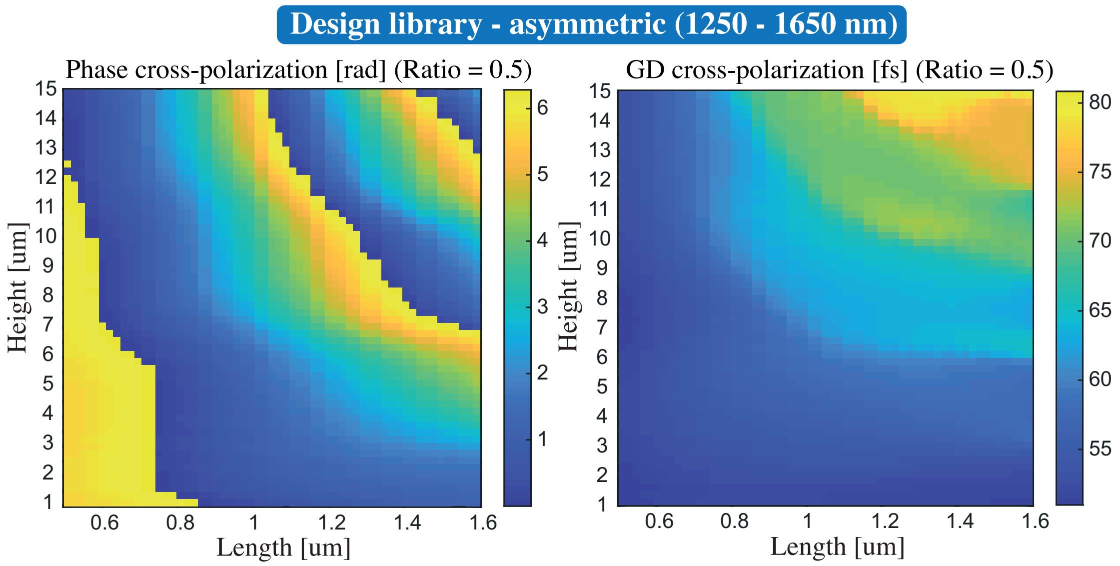


Figure A.3: **Design library of the asymmetric (1250 − 1650 nm) metalens.** Phases and group delays in the cross-polarization response of various nanopillar geometries. The phases were obtained for an incident wavelength of $\lambda = 1650\,\mathrm{nm}$. The in-plane aspect ratio was fixed to $R = 0.5$

# Eidesstattliche Erklärung

Hiermit erkläre ich, die vorliegende Arbeit selbständig verfasst zu haben und keine anderen als die in der Arbeit angegebenen Quellen und Hilfsmittel benutzt zu haben.

München, den 30.6.2021

Unterschrift (Daniele De Gregorio)