UNIVERSITY OF SÃO PAULO

SÃO CARLOS SCHOOL OF ENGINEERING

DEPARTMENT OF ELECTRICAL AND COMPUTATION ENGINEERING

POST-GRADUATION PROGRAM OF THE ELECTRICAL DEPARTMENT

Metasurfaces for control of light propagation and diffractive optics applications

Augusto Martins

São Carlos – SP

2021

UNIVERSITY OF SÃO PAULO

SÃO CARLOS SCHOOL OF ENGINEERING

DEPARTMENT OF ELECTRICAL AND COMPUTATION ENGINEERING

POST-GRADUATION PROGRAM OF THE ELECTRICAL DEPARTMENT

Metasurfaces for control of light propagation and diffractive optics applications

Author: Augusto Martins

Advisor: Prof. Dr. Ben-Hur Viana Borges

Co-advisor: Prof. Dr. Emiliano Rezende Martins

Doctoral thesis submitted in fulfilment of the requirements for the academic degree of Doctor in Engineering and Doctor in Science.

São Carlos – SP

2021

Trata-se da versão corrigida da tese. A versão original se encontra disponível na EESC/USP que aloja o Programa de Pós-graduação de Engenharia elétrica

Ficha catalográfica elaborada pela Biblioteca Prof. Dr. Sérgio Rodrigues Fontes da EESC/USP com os dados inseridos pelo(a) autor(a).

Martins, Augusto
M379m Metasurfaces for control of light propagation and diffractive optics applications / Augusto Martins; orientador Ben-Hur Viana Borges; coorientador Emiliano Rezende Martins. São Carlos, 2021.

Tese (Doutorado) - Programa de Pós-Graduação em Engenharia Elétrica e Área de Concentração em Telecomunicações -- Escola de Engenharia de São Carlos da Universidade de São Paulo, 2021.

1. Fotônica. 2. Metassuperfície. 3. Holografia digital. 4. Metalente. 5. Células solares. I. Título.

Eduardo Graziosi Silva - CRB - 8/8907

FOLHA DE JULGAMENTO

Candidato: Engenheiro **AUGUSTO MARTINS.**

Título da tese: "Metassuperfícies para o controle da propagação da luz e aplicações em óptica difrativa".

Data da defesa: 08/03/2021.

Comissão Julgadora	**Resultado**
Prof. Dr. **Emiliano R Martins** **(Orientador)** (Escola de Engenharia de São Carlos – EESC/USP)	**APROVADO**
Prof. **Andrea Di Falco** (University of St Andrews)	**APROVADO**
Prof. Dr. **Thiago Pedro Mayer Alegre** (Universidade Estadual de Campinas/UNICAMP)	**APROVADO**
Prof. **Carlos Ourivio Escobar** (Universidade Estadual de Campinas/UNICAMP)	**APROVADO**
Dr. **Alexandre Manoel Pereira Alves da Silva** (Universidade Estadual de Campinas/UNICAMP)	**APROVADO**

Coordenador do Programa de Pós-Graduação em Engenharia Elétrica:
Prof. Titular **Ivan Nunes da Silva**

Presidente da Comissão de Pós-Graduação:
Prof. Titular **Murilo Araujo Romero**

Acknowledgements

Firstly, I would like to express my gratitude to my parents José and Rita de Cassia and to my siblings Thalles Martins and Fernanda Martins for all their support and love.

To my girlfriend Giuliane for all the love, patience and support. I am very grateful for her long distance support and love during my exchange at York.

To my advisor Prof. Ben-Hur Viana Borges for his dedication throughout my academic endeavour since when I was an undergraduate.

To my co-advisor Prof. Emiliano Rezende Martins for his dedication.

To Prof. Thomas F. Krauss for supervising me during one year in his group at the University of York.

I am very grateful for all their guidance, support and confidence, which were fundamental for the execution of this thesis.

To Professor Juntao Li for the essential contribution in this thesis fabricating several metasurfaces for me.

To Professor Luiz Gonçalves Neto for his support and advices.

To Achiles F. da Mota, Donato C., Fernando L. T., Haowen Li, João P. do C., and Yin W. for all the support and help in my publications.

To Prof. Luís Alberto Mijam Barêa and Prof. Newton Cesario Frateschi for all the support.

To my friend Pedro Paulo Justino da Silva Arantes for all the discussions and support since when we were undergraduates.

To my friend Norbert Vanek for all the insightful and crazy discussions about everything and for all his support.

To my friend Kezheng Li for his friendship, support and help when I was in York and also for teaching me nanofabrication techniques.

To my friends at USP: Achiles F. da Mota for our long term and fruitful collaboration; Mateuzin R., Rodrigo R., Rodrigo Gounella, Nicão, Gustavo C., Guilherme A., Talita C. and Vinicius P. for all the help, discussions and leisure times.

To my friends and colleagues at York: Manuel, George, Kallum, Donato, Pepe, Isabel, Elena, Chris, Yue Wang, Alex, Josh, Giampaolo and Ben.

To the São Paulo Research Foundation (FAPESP) (Grants 2015/21455-1 and 2018/25372-1) for the financial support.

Abstract

MARTINS, A. Metasurfaces for control of light propagation and diffractive optics applications. Thesis (Ph.D.), São Carlos School of Engineering, University of São Paulo, São Carlos, 2021.

This PhD thesis describes the design, modelling, fabrication, and characterization of metasurfaces capable of controlling the propagation of light beams with low insertion losses. Metasurfaces are planar subwavelength structures that allow local control of phase, amplitude and/or polarization of light. These structures have proven to be extremely versatile, finding applications in imaging, holography, polarization optics and sensing, to mention only a few. One key aspect in the design of a metasurface is the material choice of its constituents, as it plays a significant role in defining the physical mechanism underlining its operation. In this sense, we can divide metasurfaces into two groups: plasmonic and dielectric. Plasmonic metasurfaces, which use metallic structures, were the first metasurfaces demonstrated in the literature. Nevertheless, the efficiencies of these metasurfaces are severely impacted by Ohmic losses and are theoretically limited in 25% when operating in transmission mode. For example, it is shown in this thesis that the transmission efficiencies of plasmonic metasurfaces based on aluminium are of the order of ~ 13%, which is typically too low for holography, for instance.

Recently, all-dielectric metasurfaces based on high refractive index materials have been proposed as an alternative to circumvent the low transmission problem of plasmonic metasurfaces. In this thesis, it is shown how the transmission efficiencies of metasurfaces are dramatically improved by dielectric materials. The dielectric of choice in this thesis is crystalline silicon (c-Si), which has a combination of advantageous properties, such as: high refractive index, ease of patterning, and low absorption in the visible (as compared to amorphous silicon). Two metasurface designs are then proposed for holography applications. The first design uses cylindrical nanoposts to impose a phase modulation in the transmitted light. The hologram shows high fidelity and high efficiency, with measured transmission and diffraction efficiencies of ~65% and ~40%, respectively. Although originally designed to achieve full phase control in the range [0-2π] at 532 nm, these holograms have also performed well at 444.9 nm and 635 nm. The high tolerance to both fabrication and wavelength variations demonstrate that holograms based on c-Si metasurfaces are quite attractive for diffractive optics applications, and particularly for full-colour holograms.

The second design uses elliptical cross-section nanoposts that are form-birefringent, that is, they provide independent control of phase for two orthogonal polarizations in the visible spectrum. Relying on these properties, a holographic stereogram was encoded in the metasurface. Briefly, a stereoscopic image (stereogram) is composed of a pair of orthogonally polarized images

taken from the same scene but recorded in slightly shifted positions to replicate the natural parallax of the human eye. For the stereoscopic effect (depth perception) to occur, each of these two images has to be directed to each of the user's eyes separately with the help of cross-polarized glasses. The stereoscopic effect is obtained by combining two holograms on the same metasurface (one for each polarization). The hologram was encoded with four phase levels. Two additional non-stereoscopic holograms using two uncorrelated images were also fabricated to help assessing polarization cross-talk. The reconstruction plane consists of a fine-sanded aluminium surface to preserve the polarization of the scattered light. The stereoscopic view is obtained with a pair of cross-polarized filters (or glasses) placed in front of the observers' eyes. The theoretical bandwidth is 110 nm with a signal to noise ratio (SNR) >15 dB. The measured transmission and diffraction efficiencies are about 70% and 15%, respectively, at 532 nm. Such high efficiency is due to a combination of low absorption and high index of c-Si at visible: the index is sufficiently high to enable sufficiently small posts to alleviate the material losses. We also investigated the metasurfaces at 444.9 nm and 635 nm to experimentally assess their bandwidth performance. The quality of the stereoscopic effect is surprisingly high at 444.9 nm (but not so much at 635 nm) with transmission and diffraction efficiencies around 70% and 18%, respectively. The proposed structure was able to successfully capture the depth effect on the reconstructed images, with potential applications in diverse areas such as visual arts, entertainment, and security. The latter, in particular, will certainly benefit from the increased degree-of-freedom conveyed by stereoscopic information.

Leveraging on the experience obtained with the research on holograms, we focused on the problem of monochromatic aberrations on metalenses. Metalenses are nanostructured surfaces that mimic the functionality of optical elements. Many exciting demonstrations had already been made, for example, focusing into diffraction-limited spots or achromatic operation over a wide wavelength range. The key functionality that was yet missing, however, and that is most important for applications such as smartphones or virtual reality, is the ability to perform the imaging function with a single element over a wide field of view. Thus, relaxing the constraint on diffraction-limited resolution, we demonstrated the ability of single-layer metalenses to perform wide field of view (WFOV) imaging while maintaining high resolution suitable for most applications. We also discussed the WFOV physical properties and, in particular, we showed that such a WFOV metalens mimics a spherical lens in the limit of infinite radius of curvature and infinite refractive index.

Finally, we explored the expertise acquired with the design of nanostructures to address an important problem in the renewable energy community: how to improve the performance of solar cells using nanostructures. In particular, we analysed the impact of these structures on the performance of a new class of solar cells: the tandem solar cell employing perovskites and silicon. Such tandem solar cells require careful photon management for optimum performance, which can

be achieved with intermediate photonic structures. We first identified that a photonic intermediate structure in a perovskite/c-Si tandem solar cell should act as an optical impedance matching layer at the perovskite-silicon interface. This conclusion did not tally with the perception in the tandem community at the time, which tented to ascribe the role of a tailored reflector to intermediate structures. Relying on the new insights gained, we analysed two simple designs and compared their performances with intermediate reflectors based on Distributed Bragg Reflectors (DBR). Our conclusion was that the intermediate structures acting only as an optical impedance matching layer show similar performance as the DBR reflectors but are much simpler. We completed the analysis by simulating a realistic device configuration and showed that optical impedance matching alone can increase the short circuit current of the silicon solar cell by 18.5% (corresponding to a boost of 2.8 mA/cm^2), thus resulting in an expected tandem efficiency in excess of 30%.

Keywords: photonic crystals, photonic nanostructures, dielectric metasurfac, metalens, wide field of view metalens, computer holography, form birefringence, stereoscopy, perovskite, tandem solar cells, crystalline silicon

Resumo

MARTINS, A. Metasuperfícies para o controle da propagação da luz e aplicações em ótica difrativa. Tese (Doutorado), Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2021.

Esta tese de doutorado descreve o projeto, a modelagem, a fabricação e a caracterização de metasuperfícies para o controle da propagação de feixes de luz com baixas perdas. Metassuperfícies são estruturas planas compostas de estruturas menores que o comprimento de onda operante que permitem o controle local da fase, amplitude e/ou polarização da luz. Tais estruturas se provaram extremamente versáteis com aplicações demonstradas em imageamento, holografia, polarização da luz e sensoriamento, por exemplo. Uma característica fundamental no projeto de uma metassuperfície é a escolha material de seus elementos, pois ele dita o mecanismo físico no qual ela se baseia. Dessa forma, podemos agrupar as metassuperfícies em duas categorias: as plasmônicas e as dielétricas. As metassuperfícies plasmônicas, que são compostas de estruturas metálicas, foram as primeiras metassuperfícies demonstradas na literatura. Porém, suas eficiências são afetadas por perdas ôhmicas e teoreticamente limitadas em 25% quando operando em transmissão. Por exemplo, nesta tese é mostrado que a eficiência de transmissão de metassuperfícies plasmônicas feitas em alumínio é da ordem de 13%, o que é muito baixo para muitas aplicações como holografia.

Recentemente, metassuperfícies baseadas em materiais dielétricos de alto índice de refração foram propostas como uma alternativa para solucionar o problema da baixa transmissão das metassuperfícies plasmônicas. Nesta tese, nós demonstramos que as metassuperfícies dielétricas apresentam, de fato, uma melhora significativa na eficiência de transmissão quando comparadas com as plasmônicas. Para tanto, utilizamos como material dielétrico o silício cristalino (c-Si), que possui uma combinação de propriedades favoráveis, tais como: alto índice de refração, facilidade de corrugação e baixas perdas no visível quando comparadas com outros tipos de silício como o amorfo e o policristalino. Assim, foram propostas e projetadas duas metassuperfícies para aplicações em holografia. A primeira é baseada em nanopostes cilíndricos capazes de modular a fase de feixes não polarizados transmitidos através da metassuperfície. Os hologramas apresentam alta fidelidade e alta eficiência, com eficiências de transmissão e difração aproximadamente de 65% e 40%, respectivamente, medidas experimentalmente. Apesar de terem sido projetadas para operar em 532 nm, os hologramas também apresentaram bons resultados em comprimentos de onda de 444.9 nm e 635 nm. Portanto, as altas tolerâncias a variações na fabricação e comprimento de onda evidenciam que hologramas baseados em metassuperfícies de silício cristalino são ótimos candidatos para aplicações em ótica difrativa e, particularmente, para hologramas coloridos.

O segundo projeto utiliza nanopostes com seção transversal elíptica que apresentam birrefringência de forma no visível. Ou seja, tais nanopostes modulam diferentemente a fase da luz transmitida de acordo com o estado de polarização da luz incidente. Dessa forma, um estereograma holográfico foi gerado com tal metassuperfície. Resumidamente, uma imagem esterocópica (estereograma) é composto de duas imagens tomadas de uma mesma sena mas fotografadas em posições diferentes para replicar a paralaxe natural da visão humana. Para o efeito estereoscópico (percepção de profundidade) ocorrer, cada uma dessas imagens deve ser vista independentemente por cada um dos olhos do observador. Para tanto, decidimos realizar dois hologramas com quatro níveis de fase cada, um para cada imagem do estereograma e numa mesma metassuperfície birrefringente em cada um dos dois estados de polarização ortogonal. Assim, o efeito estereoscópico pode ser visto na reconstrução birrefringente com o uso de óculos cujas lentes apresentam polarizadores ortogonais. Além disso, projetamos os hologramas de duas imagens diferentes para facilitar a análise de efeitos de *cross-talk* na polarização. O plano de reconstrução, para os hologramas estereoscópicos, consiste de uma superfície de alumínio lixada levemente para preservar o estado de polarização da luz espalhada. A largura de banda estimada teoreticamente é de 110 nm com uma relação de sinal ruído maior que 15 dB. As medidas de eficiências de transmissão e difração são da ordem de 70% e 15%, respectivamente, no comprimento de onda de 532 nm. Tais valores são consequências das baixas perdas e alto índice de refração do silício cristalino no visível. Ou seja, o índice de refração é alto o suficiente para minimizar as perdas materiais. As metassuperfícies foram investigadas experimentalmente quando iluminadas com lasers em 444.9 nm e 635 nm para avaliar experimentalmente sua largura de banda. A qualidade do efeito estereoscópico é surpreendentemente alta em 444.9 nm com eficiências de transmissão e difração de 70% e 18%. Já em 635 nm, as reconstruções não foram tão boas. Dessa forma, verificamos que a estrutura proposta foi capaz de demonstrar o efeito estereoscópico nas reconstruções com potencial para aplicações em diversas áreas como artes visuais, entretenimento e segurança. A última, em particular, certamente irá beneficiar do grau de liberdade adicional fornecido pela informação birrefringente.

Com base na experiência obtida na pesquisa de metassuperfícies holográficas, decidimos focar no problema de aberrações monocromáticas em metalentes. Metalentes são metassuperfícies que reproduzem as funcionalidades de elementos óticos. Muitas demonstrações surpreendentes já foram demonstradas tais como foco em pontos no limite de difração e operação acromática em uma larga banda de comprimentos de onda. Uma característica importante que ainda não havia sido propriamente solucionada e que é fundamental em aplicações como *smartphones* e realidade virtual é a capacidade de formar imagens com alto campo de visão e apenas uma metalente. Para tanto, abdicando a operação no limite de difração, nós demonstramos a habilidade de apenas metalente obter imagens com alto campo de visão (WFOV) com resoluções altas o suficiente para grande parte das aplicações em imageamento. Também são

discutidas as propriedades físicas de tais metalentes e, em particular, é mostrado que elas simulam uma lente esférica no limite de raio de curvatura e índice de refração interno infinitos.

Por fim, a experiência no projeto de nanoestruturas para o controle da luz foi utilizado para resolver um problema importante no contexto de energia renovável: como aprimorar o desempenho de células solares com nanoestruturas. Em particular, foram analisados os impactos de tais estruturas na performance de uma nova classe de células solares tandem que utilizam peroviskitas e silício. Esse tipo de célula solar tandem requer um cuidados controle fotônico para obter o melhor desempenho e isso pode ser realizado com estruturas fotônicas intermediárias, ou seja, postas entre as camadas de silício e peroviskita. Primeiramente, nós identificamos que, para essa classe de célula solar tandem, a estrutura fotônica intermediária deve casar a impedância entre a peroviskita e o silício. Tal conclusão não concorda com a percepção da comunidade científica, que era a de que deveria utilizado refletor otimizado como estrutura intermediária. Com base na conclusão que obtivemos, nós decidimos avaliar duas estruturas simples, que agem como casadoras de impedância, e comparar seu desempenho com refletores intermediárioas baseados em refletores Bragg distribuídos. Assim, concluímos que, de fato, estruturas intermediárias baseadas em casadores de impedância ótica mostram desempenhos muito semelhantes aos dos refletores intermediários mas com a vantagem de serem muito mais simples.

Palavras-chave: cristais fotônicos, metassuperfícies dielétricas, metalentes, metalentes com alto campo de visão, holografia computacional, birrefringência de forma, estereoscopia, peroviskita, células solares tandem, silício cristalino.

Index

I. Introduction

I.1. Physics of Light - Historical Introduction

The curiosity about the properties of light has driven several scientific and technological innovations throughout the history of mankind. Since the early days of classical antiquity, Greek philosophers such as Empedocles and Euclid were already familiar with some of the properties of light and applications, such as burning glasses, reflection, refraction and rectilinear propagation. At that time, they tried to explain those phenomena based on an emission (corpuscular) theory. The development of optical instrumentation experienced a boom during the Renaissance, thanks to the relentless efforts of science giants like Galileo Galilei (1564-1642), [1]. Improved telescopes and microscopes turned the far places of the universe and the uncanny world at small scales observable to the eyes of humanity, marking turning points in the fields of astronomy and biology.

Thus, under the rigorous treatment of the scientific method, several discoveries contributed to build the knowledge of the properties of light from the seventeenth to nineteenth centuries. For instance, interference fringes (Robert Boyle (1627-1691) and Robert Hooke (1635-1703)), diffraction (Francesco Maria Grimaldi (1618-1663)), polarization (attributed to Christian Huygens (1629-1692)), finite propagation velocity (Olaf Römer (1644-1710)) are some discoveries that contributed to the development of applied and fundamental optics [1, 2]. This period also marked a great conflict in the understanding of what is light. Indeed, the corpuscular theory of light, which had Newton as its main supporter and it was fiercely defended at that time, could not explain the interference and diffraction phenomena.

Huygens tried to explain these phenomena by treating light as a wave rather than a corpuscular entity. He proposed his famous principle in which every point in the wavefront can be regarded as a secondary spherical source and the sum of all these sources build up the wavefront at a later instant. Nevertheless, the theory was not accepted for almost a century, mostly due to the unique influence held by Newton at that time. Only a century after Huygens' principle had been enunciated that the wave theory regained significant attention. Thomas Young (1773-1829) in the beginning of the nineteenth century introduced the principle of interference, predicting the signature of destructive interference only allowed by a wave theory. Furthermore, the theoretical work of Augustin Jean Fresnel (1788-1827) had a decisive impact on establishing the wave theory of light, corroborated by several experimental observations (such as the Poisson's spot) [1, 3].

In the nineteenth and twentieth centuries, the understanding of what is light suffered a turning point fuelled by two breakthrough discoveries. The first of them stemmed from the fields of electricity and magnetism and is epitomized by James Clerk Maxwell (1831-1879) and his seminal 1864 paper. Maxwell not only rigorously unified all previous experimental observations in a set of equations that now bear his name, but also used them to predict the existence of electromagnetic waves. Surprisingly, his equations predicted that these waves should travel at a speed very close to the speed of light. Maxwell did not consider this fact as a mere coincidence and proposed that light is also made of coupled and oscillating electric and magnetic fields. Although not readily accepted at the time, it did not take a long time for Maxwell's predictions to be experimentally proved. In 1888, Heinrich Hertz (1857-1894) proceeded with his famous experiment proving the existence of electromagnetic waves. In other words, Maxwell's equations led to the extraordinary prediction that light is physically equivalent to radio waves, infrared radiation, ultraviolet radiation, x-rays and gamma radiation. All the different phenomena related to these different types of radiation could be in principle equally treated with Maxwell's theory. The main difference between them is their frequency of operation, which has great influence on their interaction with matter. Figure I-1 shows a simplified chart of the electromagnetic spectrum with the different kinds of radiations and their frequencies and wavelengths. Visible light is just a small patch in the rich electromagnetic spectrum.

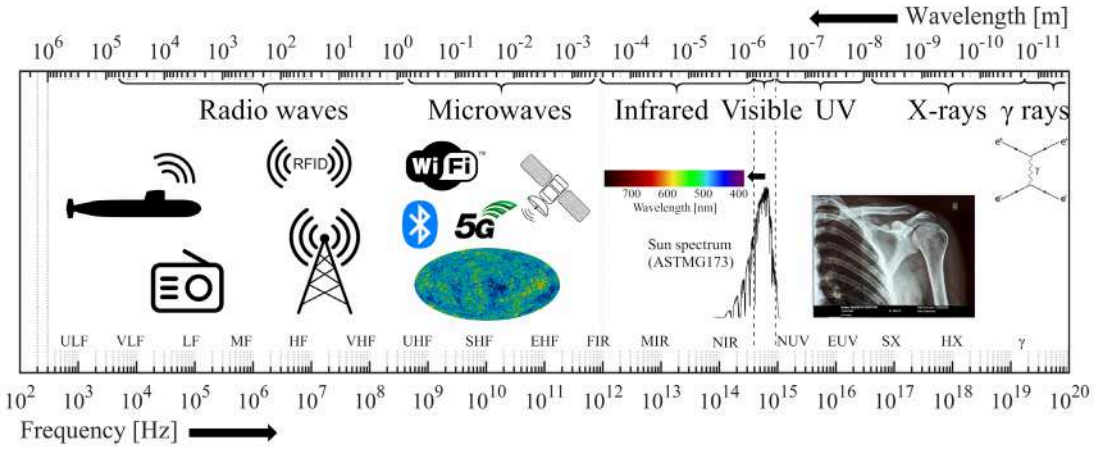


Figure I-1 – The electromagnetic spectrum. Based on [4]. Acronyms: ULF – Ultra low frequency; VLF – Very low frequency; LF – Low frequency; MF – medium frequency; HF – high frequency; VHF very high frequency; UHF – ultra high frequency; SHF – super high frequency; EHF – extremely high frequency; FIR – far infrared; MIR – mid infrared; NID – near infrared; NUV – near ultraviolet; EUV – extreme ultraviolet; SX – soft X-rays; HX – hard X-rays; γ - gamma rays.

Although Maxwell seemed to have demonstrated unequivocally that light is indeed a wave phenomenon, the second breakthrough challenged this interpretation and revived the old particle versus wave debate. Indeed, the quantum theory, which emerged in 1900 almost by accident with Max Planck's (1858-1947) research on thermal emission, shook the classical physics in its apotheosis. Newton's and Maxwell's descriptions of nature seemed to rule almost all observable physical phenomena. Nevertheless, several experiments such as blackbody radiation and

observation of discrete lines in the spectra of light crossing gaseous chemical elements could not be explained by those theories. In a creative fashion, Planck was able to theoretically estimate the blackbody radiation under the assumption that light can only assume energy values that are integers multiples of the fundamental quantum hf, where f is the operating frequency and h the Planck's constant [5]. Inspired by Planck's take on blackbody radiation, Albert Einstein (1879-1955) revived in 1905 the corpuscular theory of light. Einstein assumed that light is quantized in particles (now called photons) whose energy is given by Planck's fundamental quantum hf. That is, the corpuscular theory of light was back in the game, and with robust experimental validation, since Einstein successfully described the photoelectric effect based on this assumption. Therefore, light has characteristics of both particles and waves. From there, quantum mechanics was born, and the twentieth century witnessed its quick developments. Not long after, the field of quantum optics emerged through the work of Paul Dirac (1902-1984).

I.2. Light manipulation at the wavelength scale

Photonics is the field that studies the manipulation of light and its interaction with matter. It deals with the generation, transmission, modulation, amplification and detection of light [6]. That is, the ultimate goal of photonics is to manipulate light for scientific and technological applications. For instance, telescopes and microscopes unveil the universe at different scales; Fourier optics improves imaging systems; waveguides and optical fibres enable the transmission of light over long distances. The field was greatly impacted by the development of electronics as light and electrons interact with each other. Lasers, diodes, solar cells and electro-optical modulators are examples of applications spanning both photonics and electronics. Nevertheless, these devices rely most on the manipulation of electrical properties of the materials and/or on their natural optical properties.

What about the optical properties? How to control them?

The past few decades saw a huge effort to provide a solution for the problem of controlling the optical properties of optoelectronic and photonic devices. A clever solution was to bring crystals (periodic arrangement of atoms or molecules) from the electronic to the photonic domain. As it is well known, the electronic properties of a crystal depend on the lattice size and shape of the crystals, their constituents and their arrangements in the unit cell [7]. Evidently, the crystalline arrangement of the atoms also has an impact on the optical properties of the materials. But as the wavelength of the light is much larger than the periodicity of the crystals (the former is in the order of hundreds of nanometers whereas the latter of a few angstroms), the optical properties are impacted differently from their electronic counterparts. In fact, a crystal is effectively a homogenous medium for light and thus can be described by macroscopic parameters such as the refractive index. Consequently, light interaction with these materials are appropriately described by the macroscopic form of Maxwell's equations [8]. In this form, all the fields and charges are assumed to be averaged over a scale much smaller than the wavelength of the light but much larger than the separation of the charges. Therefore, the electromagnetic fields are described by macroscopic (averaged) field quantities whereas the material properties are described in terms of effective polarization ($\boldsymbol{P}$) and magnetization ($\boldsymbol{M}$) vectors, i.e., by a refractive index [8]. Electrons, on the contrary, have De Broglie wavelengths in the order of the lattice size of the crystals and require a more sophisticated description than averaging. The theory that explains how electrons behave in crystals is well stablished in quantum mechanics and is based on Bloch's (also called Floquet's) theorem. In short, this theorem states that the eigenfunctions of a periodic system are given by a plane wave modulated by a periodic function with the same periodicity of the system [7]. The properties of electrons are dictated by the periodic potential in the crystal.

Analogously, a similar feature could be obtained with electromagnetic waves (note that it does not need necessarily to be light) by patterning a bulk material in lattices of the order of the operating wavelength. In this case, it is the spatial variation of the macroscopic permittivity (or the refractive index) that rules the wave interaction with matter. The resulting periodic structures are called "photonic crystals", as represented in Figure I-2. One of the beauties of this approach is that several observations made for electrons in crystals also apply for photons in photonic crystals [9]. The appearance of bandgaps, for example, are also possible in the band diagram of a photonic crystal, which prevents the propagation of the wave in certain directions, as shown in Figure I-3. There is a myriad of applications explored with photonic crystals such as: waveguide couplers, filters, photonic crystal fibers and waveguides, cavities, solar cell efficiency enhancement and sensing, to mention but a few.

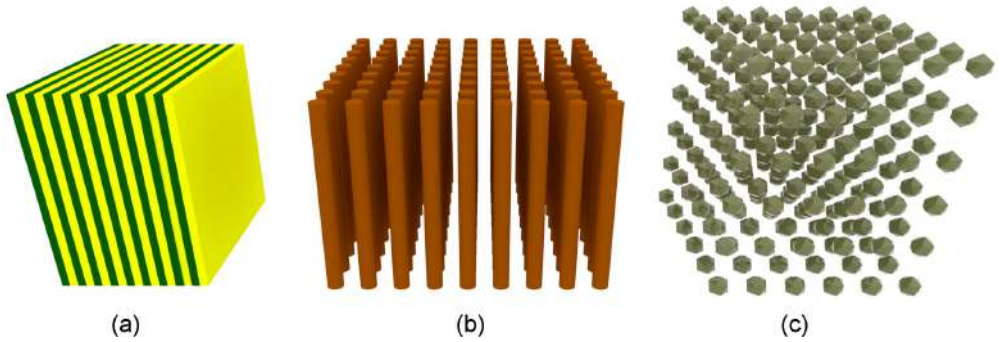


Figure I-2 – Examples of (a) one-, (b) two- and (c) three-dimensional photonic crystals.

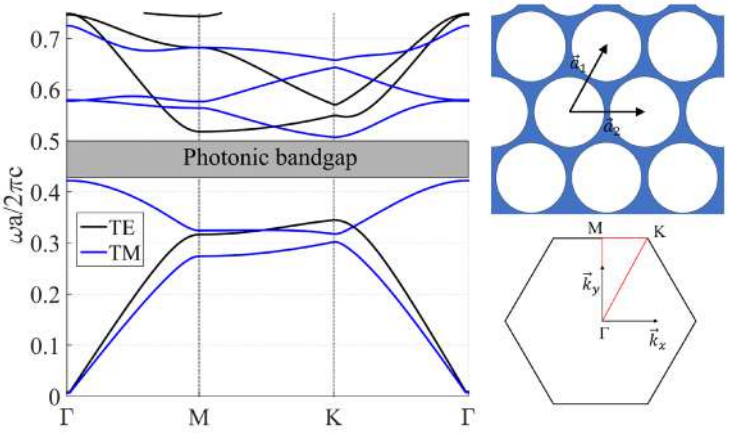


Figure I-3 – In-plane ($k_z = 0$) band diagram (left) of a 2D photonic crystal with a triangular lattice (upper right). Note the presence of a photonic bandgap in the middle of the band diagram. The primitive vectors are $\vec{a}_1 = a\left(\frac{\sqrt{3}}{2}, \frac{1}{2}\right)$ and $\vec{a}_2 = a(1, 0)$. The photonic crystal in this example is made up of holes in a medium with a permittivity of 13. The bottom right figure shows the irreducible Brillouin zone of the crystal with the symmetry points (Γ, M and K) at the corners. This example was extracted from [9]. The band diagram was calculated using a RCWA code written by the author. The TE (even) and TM (odd) modes have only (E_x,E_y,H_z) and (H_x,H_y,E_z) components, respectively. This is the definition according to the mirror symmetry of the modes with respect to the *x*-*y* plane.

Another approach to manipulate light that has resulted in several applications is the coupling between photons and electrons in a metal/dielectric interface. In the visible and infrared ranges, metals support coherent oscillations of electrons driven by electromagnetic waves, the so-called surface plasmon polaritons (SPPs), as illustrated in Figure I-4 (a). SPPs are in-plane propagating modes confined in the metal/dielectric interface [10]. The coupling between electrons and photons finds several interesting applications and can lead to novel physical phenomena, so much so that a whole field was named after it: the field of Plasmonics. Biological and chemical sensing and spectroscopy [11-14] are some examples of the vast range of applications that resulted from SPPs. Due to its guided nature, the dispersion line of a SPP mode lies below the cone of light, as represented in Figure I-4 (c) and cannot couple directly with free space radiation. That is, they need an additional mechanism to provide enough momentum to the incoming light and reach the required phase matching. For instance, this can be done by tunnelling of light in a prism or with a grating [10].

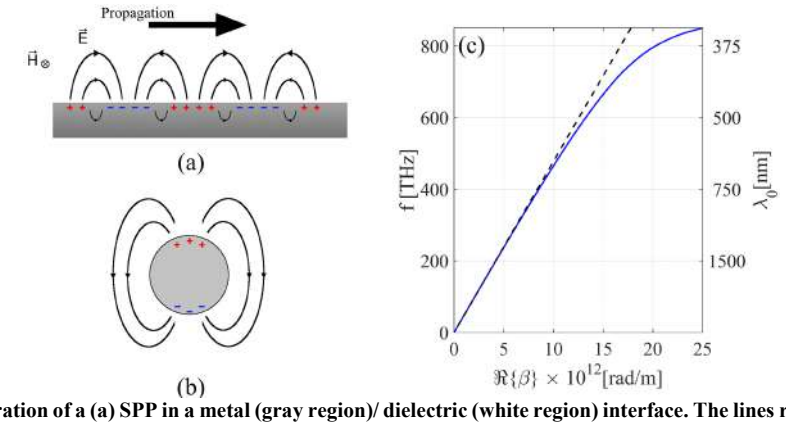


Figure I-4 – Illustration of a (a) SPP in a metal (gray region)/ dielectric (white region) interface. The lines represent the electric field and the signals the charge density inside the metal. (b) Represents the charge density and electric field lines of a LSPR of a metallic sphere. (c) Shows the dispersion line (solid blue) of a SPP in a Silver/Air interface before the silver frequency of plasma and interband transitions [10]. Experimental data by Johnson et al. was used for the permittivity distribution of Silver [15]. The dashed line marks the dispersion of plane waves in air. Propagation in air is allowed only above this line.

This issue can be circumvented by "trapping" the surface plasmon polariton in a cavity-like structure. For example, this can be done with metallic nanospheres [16], nanorods [12, 17] and nanoantennas [18-21]. In these configurations there is no need for an external phase matching mechanism and the plasmon can be coupled directly with free space radiation. Differently from common cavities, where the electromagnetic field is confined inside them, a SPP "cavity" supports its waves on the metal/dielectric interface; in other words, the plasmon oscillates locally around (electromagnetic waves) and on the surface (charges) of the metallic structure. This mode receives the name of Localized Surface Plasmon (LSP) and it typically shows resonant behaviour; its resonance being called a Localized Surface Plasmon Resonance (LSPR) [10]. A LSPR of a metallic sphere is represented in Figure I-4 (b). Typically, these structures manifest a large enhancement of both the extinction cross section and the near field of the nanostructure on resonance. The near-field enhancement is usually extremely high, making the LSPR very sensitive to environmental changes. Thus, LSPRs are promising candidates for sensing

applications. Furthermore, the high near-field enhancement also finds applications based on non-linear phenomena such as Surface-enhanced Raman scattering [22, 23].

Another class of structures that has received a great deal of attention in the past few decades are the metamaterials. Metamaterials are artificially engineered bulk materials whose electromagnetic properties can be manipulated in unprecedented ways [24-26]. Note that, differently from photonic crystals, metamaterials are formed by structures lying in subwavelength regime and whose properties are homogenized. A band diagram representing the regimes where these structures are typically found is shown in Figure I-5 (a). Differently from a photonic crystal, a metamaterial has only one mode for each state of polarization and it is possible to define effective wave parameters such as refractive index, permittivity and permeability parameters. A generic metamaterial is illustrated in Figure I-5 (b). These structures present flexibility on the tunability of their electrical properties making them appealing for several scientific and technological applications such as sensors [27], antennas [28], perfect lenses [29] and even invisibility cloaks [30-32]. Negative refraction, for instance, is a property that is not found in nature but that is possible in metamaterials [31] and photonic crystals near the bandgap [33]. Unfortunately, high insertion losses and fabrication challenges associated with bulk metamaterials hinder their practical applications.

Metasurfaces, the two-dimensional (2D) equivalent of volumetric metamaterials, are planar subwavelength structures that allow local control of phase, amplitude and/or polarization of light [34-39]. Differently from metamaterials, whose functionalities rely on the propagation of electromagnetic waves through a bulk media, metasurfaces operate as a discontinuity capable of imposing abrupt changes on wavefronts [34, 37]. Therefore, metasurfaces can overcome the propagation loss and fabrication issues associated with bulk metamaterials. A representation of a metasurface is shown in Figure I-5 (c).

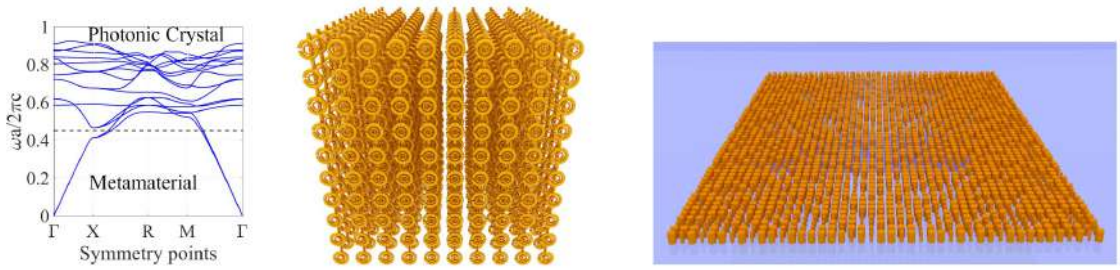


Figure I-5 – (a) Bloch modes diagram calculated of *n*=3 oblate spheroids (x and z semi axes: *a/2*; y semi axis: *a/4*) in simple square lattice (period *a*) embedded in air. For details on the symmetry points, check [7]. The energy diagram was calculated in Comsol. Representation of a metamaterial and a metasurface (b) and (c), respectively. Note that differently from a photonic crystal, the unit cell size of a metamaterial is several times smaller than the wavelength.

The first metasurfaces proposed in the literature were based on metallic nanoresonators whose functionality relied on the excitation of LSPRs [36-38, 40-45]. Unfortunately, LSPRs suffer from high losses at visible wavelengths that severely limit the metasurface efficiency when operating in transmission-mode [17, 36, 43, 45, 46]. For instance, the maximum reported

efficiency for a transmission based metasurface hologram based on LSPR is about ~10% at near-infrared wavelengths [43]. Furthermore, LSPR-based metasurfaces do not afford full-phase shift (0-2π) between incident and transmitted fields through resonance frequency tuning only [36], which is an major limitation in many applications. In contrast, when operating in reflection mode, LSPR based metallic metasurfaces can reach efficiencies as high as 80% [44].

In view of all these problems, all-dielectric metasurfaces based on high refractive index materials have been proposed as an alternative to circumvent the low transmission problem of metallic metasurfaces. All-dielectric metasurfaces can be divided in two classes: high- and low-contrast structures, with contrast defined as the difference between the refractive index of the nano inclusion and that of the surrounding medium. High-contrast metasurfaces structured as low-aspect ratio resonators may exhibit only electric and magnetic dipolar Mie resonances at optical wavelengths [47], and these resonances can be tuned and brought into spectral overlap [48], thus allowing manipulation of the amplitude and phase of the light [49-51]. However, if both the contrast and aspect-ratio are high, the structure may exhibit several multipole resonances at optical wavelengths [47]. In this case, it is preferable to treat each structure as a truncated waveguide (exhibiting Fabry-Pérot resonances [52]) where the waveguide effective index can be tailored by adjusting the fill-factor of the structures. This last feature offers an attractive degree-of-freedom in the design of diffraction gratings as it allows them to be made with either high or low index contrast. For instance, a [0-2π] phase control has been achieved with tall cylinders (nanoposts, see Figure I-6 (a)) and high transmission [52-55]. In fact, the high diffraction efficiency provided by dielectric metasurfaces has been explored in many other classical applications, such as lenses [56-58], holograms [50-52, 55, 59-63], wave plates [64], anomalous refraction generation [53, 65], and vortex beam generation [65, 66]. Figures I-6 (b) and (c) illustrate examples of lenses and holograms, respectively, made with metasurfaces based on high contrast gratings.

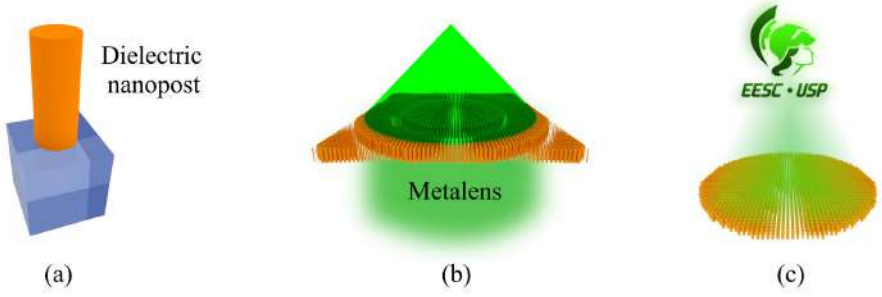


Figure I-6 – (a) Representation of a dielectric cylinder used to phase shift light locally. (b) Representation of a metalens made of dielectric nanoposts. Note that in this case the nanoposts are arranged to mimic the phase distribution of a lens. (c) Representation of a holographic metasurface. This image is not to scale.

The choice of materials plays an important role in the design of the strucures. High index materials such as titanium dioxide (TiO_2) [58, 62, 67, 68], silicon nitride (Si_3N_4) [54], and silicon [50-53, 59-61, 63-65, 69] are usually the preferred choices. Silicon is particularly interesting for

metasurface applications not only for its compatibility with CMOS processes, but also for its high refractive index that allows high-contrast gratings to be fabricated with low aspect ratio. For instance, metasurface holograms with diffraction efficiencies in excess of 90% were obtained in the infrared with poly-silicon (p-Si) metasurfaces based on Mie resonances [63] and with amorphous silicon (a-Si) based on high-contrast grating [52]. However, the high absorption of these two types of silicon severely limits their application in the visible range. For example, a full color hologram has been demonstrated with diffraction efficiencies of only 3.6%(blue), 5.2%(green) and 18% (red) [61] using amorphous silicon (a-Si). Furthermore, using poly-silicon (p-Si) a hologram operating at 532 nm has been demonstrated with a diffraction efficiency of only 6% [55].

In this scenario, crystalline silicon (c-Si) can be advantageous as it has lower absorption in the visible range. This feature motivated the recent demonstration of a polarization independent metasurface based on c-Si, which achieved a high transmission efficiency of 71% at 532 *nm* with an aspect ratio of only 3.4 [53]. In contrast, a TiO_2-based metasurface hologram requires aspect ratios larger than 10 [62] to achieve a similar efficiency.

I.3. From classical to digital holography

Holography is a technique that generates an image by modulating the amplitude and/or phase of a light beam. This concept was initially developed in the context of optical microscopy by Dennis Gabor [70] in 1948 to improve the imaging quality. In 1971, Gabor was awarded a Nobel prize for the invention of holography. In his work, Gabor used the interference pattern of two coherent sources at the recording process: the scattered field by an object and a reference field, as shown in Figure I-7. It can be shown that the amplitude of the resulting interference pattern is modulated by both the phase and amplitude of the field scattered by the object [3, 70]. The interference pattern is then recorded at a medium sensitive to the field intensity. Therefore, when a coherent source, with the same wavelength and at the same incidence angle as the reference wave used to record the hologram, impinges at the recorded media, the scattered field will reconstruct the original object field pattern creating an image of it. Furthermore, spurious scattering will also be generated: a zero order that is the portion of wave passing the recorded media directly and a twin image that will create a virtual image of the original object. In practice, this procedure needs an optical setup with high precision to make sure that all elements are aligned.

Gabor's work did not receive too much attention during the first decade after its publication due to practical reasons that even hindered the experimental realization of the technique at that time. Nevertheless, the theory received improvements after Gabor's initial work by other researchers [70]. In 1960s, several scientists mainly at the Soviet Union and University of Michigan's Radar Laboratory applied the emerging technology of lasers to holography, providing a huge development in the method. For instance, Leith and Upatnieks improved the recording technique to remove the twin images superposition that is present in the Gabor Hologram [71].

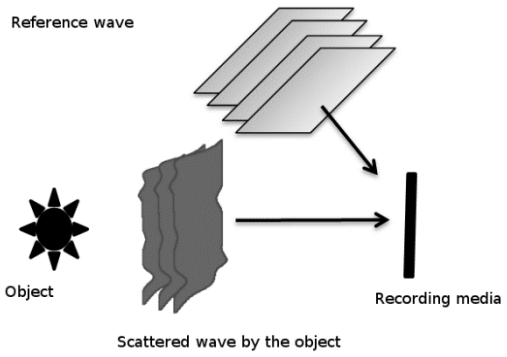


Figure I-7 – Pictorial representation of the recording of a classical hologram. The scattered and reference wavefronts are coherent and interfere at the recording media that is sensible to the field intensity [3].

With the advent and ongoing improvement of digital computers, several algorithms were developed for generating holograms digitally. The resulting computer-generated holograms (CGHs) are matrices storing the field phase and amplitude as its classical counterpart. The main difference between them is that it is much easier to generate a CGH than a classical hologram because the former does not need any optical setup for its generation, but only a computer, whereas the latter requires a precise setup. The diffraction of the field can be calculated, under some approximations – as described in appendix *D.1 Diffraction in Fourier Optics - Mathematical Background* – by Fourier Transforming the phase/amplitude spatial distribution imparted by the recording media. Reversely, the field distribution in the hologram plane can be obtained, under the same approximations, by inverse Fourier Transforming the original field distribution. These operations can be easily performed numerically with the fast Fourier Transform (FFT) algorithm, that is a highly optimized algorithm.

Different from classical holograms where a real scattered field from a physical object is needed to be recorded, in a CGH any digital image or three-dimensional scenario can be, in principle, recorded. Therefore, one can make a hologram from any field distribution in theory which may widen the applicability of holography. Another important difference in comparison with classical holography is that a CGH is stored numerically in a computer memory rather than in a physical medium. Therefore, in order to reconstruct a CGH, it is necessary to encode it in an array of diffractive optical elements (DOEs). There are several well-developed techniques for fabricating the CGHs such as pen plotters, laser printers, photolithography, electron-beam lithography, focused ion beam milling and nanoimprinting, to mention a few.

Unfortunately, it is challenging to achieve full locally control the phase and amplitude of the light with DOEs. Therefore, the phase and/or the amplitude of the hologram must be quantized in a suitable way that does not spoil the information being encoded. Usually, the quality of the encoded CGH is quantified in terms of its reconstruction signal-to-noise ratio (SNR) and diffraction efficiency (DE). Several algorithms were developed to maximize those values such as Detour Phase Hologram, Direct Binary Search and the Iterative Fourier Transform Algorithm (IFTA).

In this work I adopted the Iterative Fourier Transform Algorithm (IFTA) [72-75], because it outperforms both the Detour Phase Hologram and the Direct Binary Search in terms of calculation time and efficiency [75]. A detailed explanation of the IFTA is presented in section *D.3 Iterative Fourier Transform Algorithm (IFTA)* of the appendix, which was extracted and adopted from Neto [75]. The main goal of this thesis is to explore CGHs encoded on metasurfaces and study their performances, main properties and limitations. Both plasmonic and dielectric metasurfaces were considered but, due their far better performances on transmission mode, the thesis focuses on the latter.

I.4. Text organization

This thesis is organized as follows. Chapter II presents the physics and design concepts behind metasurfaces in general, and of c-Si based metasurfaces in particular. In section II.1 an overview of metasurfaces is presented where the main physical mechanisms employed in their design are presented. Section II.2 presents an example of a metasurface design. In sections III.1 and III.2 the design, fabrication and experimental characterization of c-Si based holographic metasurfaces are shown. Section III.3 presents some concluding remarks and future works. Chapter IV presents the design, fabrication and characterization of metalenses with wide field of view.

Chapter V shows how nanostructures can impact the performance of a promising new class of solar cells, the tandem solar cells employing perovskites and silicon. The chapter closes with some concluding remarks and future works.

It should be noted that a plasmonic holographic metasurface design based on aluminium nanoantennas is described in Appendix A. Appendix B contains additional information regarding the metasurface design. Appendix C contains additional information regarding the metalens design. The Appendix D presents the theoretical and numerical foundations of digital holography.

I.5. Published works

The published works that I have been involved in during my PhD are listed below. All were published in international periodicals but the last one, which I presented in an Optical Society of America (OSA) congress. Publications 2, 3, 4, 11 and 12 are results from this thesis. The remaining are results from collaborations in other projects related to mine.

1. A. F. Mota, A. Martins, J. Weiner, F. L. Teixeira, and B.-H. V. Borges, "Constitutive parameter retrieval for uniaxial metamaterials with spatial dispersion," Physical Review B, vol. 94, p. 115410, 2016.
2. A. Martins, B. H. V. Borges, J. Li, T. F. Krauss, and E. R. Martins, "Photonic Intermediate Structures for Perovskite/c-Silicon Four Terminal Tandem Solar Cells," IEEE Journal of Photovoltaics, vol. 7, pp. 1190-1196, 2017.
3. A. F. d. Mota, A. Martins, H. Ottevaere, W. Meulebroeck, E. R. Martins, J. Weiner, F. L. Teixeira, and B.-H. V. Borges, "Semianalytical Model for Design and Analysis of Grating-Assisted Radiation Emission of Quantum Emitters in Hyperbolic Metamaterials," ACS Photonics, vol. 5, pp. 1951-1959, 2018.

4. A. Martins, J. Li, A. F. da Mota, Y. Wang, L. G. Neto, J. P. do Carmo, F. L. Teixeira, E. R. Martins, and B. V. Borges, "Highly efficient holograms based on c-Si metasurfaces in the visible range," Opt Express, vol. 26, pp. 9573-9583, Apr 16 2018.
5. A. Martins, J. Li, A. F. da Mota, V. M. Pepino, Y. Wang, L. G. Neto, F. L. Teixeira, E. R. Martins, and B.-H. V. Borges, "Broadband c-Si metasurfaces with polarization control at visible wavelengths: applications to 3D stereoscopic holography," Optics Express, vol. 26, pp. 30740-30752, 2018/11/12 2018.
6. V. M. Pepino, A. F. d. Mota, A. Martins, and B. V. Borges, "3-D-Printed Dielectric Metasurfaces for Antenna Gain Improvement in the Ka-Band," IEEE Antennas and Wireless Propagation Letters, vol. 17, pp. 2133-2136, 2018.
7. A. F. da Mota, A. Martins, V. Pepino, H. Ottevaere, W. Meulebroeck, F. L. Teixeira, and B.-H. V. Borges, "Semianalytical modeling of arbitrarily distributed quantum emitters embedded in nanopatterned hyperbolic metamaterials," JOSA B, vol. 36, pp. 1273-1287, 2019.
8. A. Abass, A. Martins, S. Nanz, B.-H. Borges, E. Martins, and C. Rockstuhl, "Perturbing beyond the shallow amplitude regime: Green's function scattering formalism with Bloch modes," JOSA B, vol. 36, pp. F89-F98, 2019.
9. A. Safdar, Y. Wang, C. Reardon, J. Li, G. S. De Arruda, A. Martins, E. R. Martins, and T. F. Krauss, "Interplay between optical and electrical properties of nanostructured surfaces in crystalline silicon solar cells," IEEE Photonics Journal, 2019.
10. G. S. Arruda, J. Li, A. Martins, K. Li, T. F. Krauss, and E. R. Martins, "Reducing the surface area of black silicon by optically equivalent structures," IEEE Journal of Photovoltaics, vol. 10, pp. 41-45, 2019.
11. H. Liang, A. Martins, B.-H. V. Borges, J. Zhou, E. R. Martins, J. Li, and T. F. Krauss, "High performance metalenses: numerical aperture, aberrations, chromaticity, and trade-offs," Optica, vol. 6, pp. 1461-1470, 2019.
12. A. Martins, K. Li, J. Li, H. Liang, D. Conteduca, B.-H. V. Borges, T. F. Krauss, and E. R. Martins, "On Metalenses with Arbitrarily Wide Field of View," Acs Photonics, vol. 7, pp. 2073-2079, 2020/08/19 2020.
13. A. F. da Mota, A. Martins, J. Weiner, P. Courteille, E. R. Martins, and B.-H. V. Borges, "Design and analysis of nanopatterned graphene-based structures for trapping applications," Physical Review B, vol. 102, p. 085415, 2020.
14. K. Li, S. Haque, A. Martins, E. Fortunato, R. Martins, M. J. Mendes, and C. S. Schuster, "Light trapping in solar cells: simple design rules to maximize absorption," Optica, vol. 7, pp. 1377-1384, 2020.

15. A. Martins, J. Li, A. F. da Mota, Y. Wang, L. G. Neto, J. P. do Carmo, F. L. Teixeira, E. R. Martins, and B.-H. V. Borges, "Crystalline Silicon (c-Si) Metasurface Holograms in the Visible Range," in Imaging and Applied Optics 2018 (3D, AO, AIO, COSI, DH, IS, LACSEA, LS&C, MATH, pcAOP), Orlando, Florida, 2018, p. DTh2E.5.

II. Light control with metasurfaces

II.1. Introduction

Metasurfaces are planar subwavelength structures that allow local control of phase, amplitude and/or polarization of light [34-38]. Its functionality relies on the Huygens principle, which states that every point of the wavefront can be regarded as a secondary spherical source and the sum of all these sources build the wavefront at a later instant. In this sense, each element of a metasurface can be regarded as an engineered secondary source generator. Therefore, with a judicious arrangement of these structures, a phase, amplitude and/or polarization profile can be imprinted on the transmitted or reflected light [39]. Figure II-1 illustrates the basics of a metasurface. The yellow cylinders represent the metasurface building blocks that are excited by the red wavefront from below. After transmission, the wavefronts are all distorted because the metasurface imprinted a phase and/or amplitude distribution on the transmitted light. Note that a metasurface can also work on reflection. With this principle a metasurface can be used to encode holograms, metalenses, waveplates and beam shapers, to mention only a few.

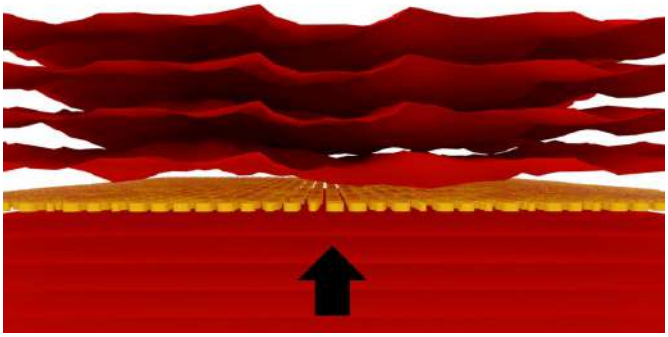


Figure II-1 – Pictorial representation of the operation of a metasurface composed by yellow cyllinders. The structure is excited from below in this example and the transmitted light acquires phase and amplitude signatures imparted by the metasurface.

The design of a metasurface starts by first determining how and which properties of light (phase, amplitude and/or polarization) will be manipulated after transmission and/or reflection in a desired wavelength. The next step is to choose what physical mechanism will rule the property manipulation and which material is more suitable for that in the chosen wavelength. To this intent, the most common approaches are resonant tuning, Pancharatnam-Berry (PB) phase and hybrid Resonant Tuning with PB phase [36].

The first method is based on tunning a local resonance of the inclusion by changing its size, as illustrated in Figure II-2, where the radius of an arbitrary shaped resonator is changed. Note that the scattering of a nanostructure is strongly enhanced when it is in the resonance, which is also accompanied by an enhancement of its absorption if the material is lossy [76]. Thus, with

small geometrical adjustments, the inclusion may go from a non-resonant to a resonant condition that can radically change the phase and/or amplitude of the scattered light, as represented in Figure II-2. Thus, local control of light may be achieved by shaping the nanostructure in a controlled way. In this approach, there is no polarization conversion between the incident and scattered light and the structure can even present form-birefringence [52, 77]. This last feature can be achieved by resorting to asymmetrical structures that present non-degenerate modes excited with light polarized along two orthogonal axes.

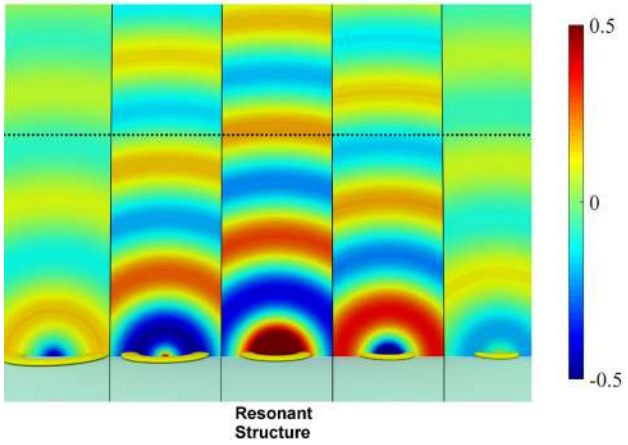


Figure II-2 – Representation of the resonant tuning mechanism by changing the size of the structure. The field distributions are only representations and not the actual scattering distributions by these structures. Each field distribution is the scattering of different structures plane wave excitation coming either from below or above. The resonant structure presents the higher scattered field. Furthermore, note that the phase shift of the radiated light by different structures is different (use the dotted line as reference).

The PB phase, also called geometric phase, is a way to obtain a local phase shift on transmitted or reflected light by rotating the scatterer [78, 79], as represented in Figure II-3. When a circularly polarized light is incident on a nanostructure, it will transmit light not only with the same polarization of the incident light but also with the other handedness (cross-polarized term). If the structure is rotated about its centre by an angle θ, the transmitted cross-polarized term will acquire a phase shift of 2θ [78, 79]. Thus, full phase-control in the range [0-2π] can be achieved by simply rotating the structure in the range [0-π]. On the contrary, if the reflected light is analysed, it is the co-polarized reflected term that acquires a phase shift of 2θ. Figure II-3 illustrates the cross-polarized scattered field by a rod-like structure excited with circular polarization. In this example, each structure is [counter] clockwise rotated 22.5 degrees with respect to its [left] right neighbour, thus, the phase difference between them is [-] +45 degrees. Although the phase control does not depend on any resonance that the structure may present, the scattering cross section does. Therefore, the design of structures aims to maximize the scattered cross(co)-polarization if transmissive (reflective) metasurfaces are of interest and resonant behaviour is usually required. In other words, these structures should behave as half-wave plates (in reflection, a half-wave plate keeps the circular polarization the same).

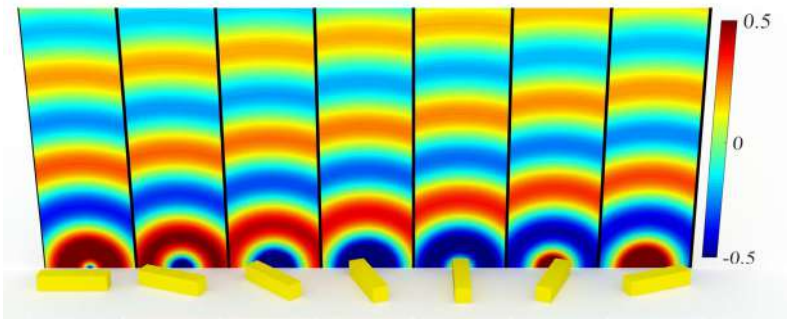


Figure II-3 – Representation of the scattered field by structures operating under the Pancharatnam-Berry phase (these are not the actual field scattered by the represented structures). It is assumed that the exciting field is polarized circularly (can income either from above or below) and the shown distributions are circularly polarized in the other handedness. Note that, in this example, each structure is [counter] clockwise rotated 22.5 degrees with respect to its [left] right neighbor, thus, the phase difference between them is [-] +45 degrees. Based on [80].

Finally, with the hybrid method (that explores both resonant tuning and PB phase) it is possible to design structures that present the phase shift flexibility allowed by the geometric phase and also amplitude and polarization control [68, 81]. This method can operate not only with circular polarization but also with any orthogonal pair of polarization states [68].

All these mechanisms have been successfully demonstrated with all-dielectric structures [36]. Furthermore, when shaped with low aspect-ratio, high-index dielectric structures may present localized electric and magnetic dipolar Mie resonances (see Figure II-4**)** with similar resonant wavelengths and comparable scattering cross-sections [76]. When these resonances overlap spectrally and have similar strengths, several interesting physical phenomena arise. For instance, they support the Kerker effect (named after Milton Kerker for theoretically predicting it [82]) in which the overall backward scattering is suppressed due to the interference between the partial scattering of the electric and magnetic dipole modes [48, 83]. Additionally, it was discovered that this condition also allows the full phase control [0-2π] of the scattered light acting as Huygen's sources [49]. Metasurfaces that rely on this mechanism are named Huygens' metasurfaces.

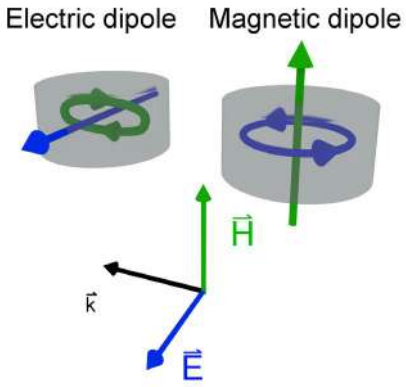


Figure II-4 -Representation of the electric and magnetic dipole Mie resonances excited by a plane wave that is being represented by the $\vec{k}$ (wavevector), $\vec{E}$ (electric field) and $\vec{H}$ (magnetic field) triad. Note that the electric (magnetic) field circulates inside the structure at the magnetic (electric) resonance. This figure was adapted from [83].

Increasing the aspect ratio of high contrast dielectric structures enhances the scattering contributions of high order multipole modes [47], thus making the multipole description of their scattering properties cumbersome and non-intuitive. These structures, however have been studied using High Contrast Gratings (HCGs), which are periodic by nature, long before the development of the metasurface concept [67]. Several applications have been successfully demonstrated with HCGs in the past such as high reflecting mirrors for vertical-cavity surface-emitting lasers (VCSELs)[84], high-Q resonator [85] and lenses [86], to mention only a few. Although periodic structures are fully described by the Bloch's theorem and are non-local by nature, the high contrast index of HCG gives rise to local resonances with high mode confinement inside the high index region [52, 56, 86]. These modes can be treated as Fabry-Perot resonances of each array element, which is acting as a truncated waveguide [52]. Thus, due to the high modal localization, the optical response of these structures can be obtained by treating them as HCG, as will be shown in the next section. More details about the modes confinement in HCG structures are shown in *Appendix B – Field confinement in dielectric* nanoposts.

II.2. Design of a metasurface

The design procedure is applied to the case of beam steering metasurfaces, which can bend light by an arbitrary angle θ with respect to the normal incidence (this could also be done in reflection). This example lies in the core of metasurfaces as it was the catalyst for the development of the concept [80]. How should a metasurface operate on light to create this effect? This was answered with the aid of Fermat's principle that states that light follows the optical path that is at an extremum under variational changes between two points (let's call these points A and B) [6]

$$\delta \int_A^B n(\vec{x})dx = 0 \qquad \text{(II-1)}$$

where $n(\vec{x})$ is the refractive index along the path. Equation (II-1) is valid for light propagating without discontinuities. What if there is an open surface *S* between A and B that imposes a phase discontinuity $\Phi(\vec{x}), \vec{x} \in S$ in the path, such as a metasurface? In this case, Equation (II-1) should be rewritten as [80]

$$\delta \left[\Phi(\vec{x}) + \int_A^B n(\vec{x})dx \right] = 0 \qquad \text{(II-2)}$$

It will be further assumed that *S* is a plane, the light travels normally from point A to the discontinuity and that the two media separated by S are homogeneous with refractive indexes n_a and n_b where points A and B stand, respectively. Working under these assumptions, one arrives at [80]

$$\frac{d\Phi}{dx} = n_b k_0 \sin\theta \qquad \text{(II-3)}$$

where $k_0 = \frac{2\pi}{\lambda_0}$ is the freespace wavenumber and λ_0 the operating wavelength. Equation (II-3) is the normal incidence case of the generalized law of refraction [80]. The importance of Equation (II-3) should not be underestimated. Equation (II-3) can be used to analyse the behaviour of a generic metasurface where the steering angle is not constant. Note that the right part is essentially the local in-plane momentum acquired by the transmitted beam. In other words, Equation (II-3) tells what is the local spatial frequency [3] of the transmitted beam. For instance, a hyperbolic metalens with focal length *f* needs $\theta(x) = -\operatorname{asin}\frac{x}{\sqrt{x^2+f^2}}$ [87]**.** Inserting this equation in equation (II-3), the Hyperbolic phase profile is readily obtained [87]

$$\phi(x) = -n_b k_0 \sqrt{x^2 + f^2} \qquad \text{(II-4)}$$

The same rule applies to other devices that operate under similar conditions such as axicons and holograms. Turning our attention back to the simple case of constant θ, it follows from equation (II-3)

$$\Phi(x) = x n_b k_0 \sin\theta \quad \text{(II-5)}$$

Note that the integration constant was neglected because a constant phase value does not alter the anomalous angle of refraction given by equation (II-3). Therefore, a metasurface needs to impose this linear phase distribution on the reflected or transmitted light to simply steer the incoming beam, as represented in Figure II-3.

Suppose that a given design has pixels with unit cell size *P* and can provide a full range phase shift with *N* different structures. That is, there are *N* different phase levels available whose discretization is $\Delta\Phi = \frac{2\pi}{N}$. Therefore,

$$\frac{d\Phi}{dx} \cong \frac{\Delta\Phi}{P} = \frac{2\pi}{NP} \quad \text{(II-6)}$$

Substituting (II-3) in (II-6), the following restriction is imposed

$$n_b k_0 \sin\theta = \frac{2\pi}{NP} \quad \text{(II-7)}$$

Interestingly, equation (II-7) is exactly the grating equation for the first diffraction order for a grating with periodicity *NP*, under normal incidence and towards a medium with refractive index n_b. Furthermore, equation (II-7) imposes a trade-off between the desired steering angle, the number of phase samplings and the metasurface unit cell size.

Equations (II-6) and (II-7) give the necessary information about the phase distribution that the proposed metasurface should impart. From equation (II-6), we can conclude that the metasurface should impart a linear phase distribution on the metasurface. That is, two adjacent elements should have absolute phase difference of $2\pi/NP$. Furthermore, from equation (II-7) we can obtain the required *NP* value for this design.

In this work, the phase shift imparted on the transmitted light is the parameter of interest. Furthermore, the mechanism chosen to drive this control is the resonant tuning of dielectric nanoposts. Therefore, phase and transmission maps as functions of the geometrical parameters of the structures need to be obtained. The phase map is obtained by calculating the transmission of a normal plane wave incident on an infinite array of nanoposts with the rigorous coupled-wave analysis (RCWA) [88]. This semi-analytical method relies on Bloch's theorem, which, in short, states that the eigenfunctions of a periodic system, in this case a photonic one, are given by a plane wave modulated by a periodic function with the same periodicity of the system [9]. Thus, RCWA already assumes that the system is periodic. Furthermore, the silicon high index guarantees that the field is highly confined inside the structure, thus making the coupling between adjacent structures small [52]. Therefore, the phase and transmission changes as function of the geometrical parameters of each structure in the array are similar to the changes for the same parametrization on an isolated structure. This is necessary because the holographic metasurface assumes that the response of each pixel is local in the sense that it is not coupled to adjacent pixels.

The advantage of using an infinite array is because Bloch's theorem greatly facilitates the numerical modelling. However, the numerical simulation of the electromagnetic response of an isolated structure under the excitation of a plane wave requires the computational domain to be extended beyond its unit cell, so evanescent fields excited by the structure do not touch the boundaries of the domain, and therefore tends to have higher computational cost. Normally, these simulations are carried out with finite elements method (FEM) or finite-difference time-domain method (FDTD) and require a perfect match layer (PML) surrounding the computational domain to absorb the scattered light and simulate an infinite region.

Continuing our example, the structure represented in Figure II-5 (a) will be used to achieve the desired phase control. It consists of a dielectric ridge made of silicon on a silicon substrate. Note that the actual calculations are carried with the array depicted in Figure II-5 (b) in RCWA, as explained above. The operating wavelength was chosen to be 900 *nm* and the silicon index is n_{Si}=3.74. The incidence is normal from above with TM polarization (electric field along $\hat{x}$, see Figure II-5 (a)).

The phase and efficiency of the transmitted zero-order by the grating into the substrate was evaluated as function of the grating period (*p*) and fill factor (*ff).* The fill factor is defined as the ratio between the ridge width to unit cell size (*ff=w/p).* The unit cell size is swept in the range [50 nm,225 nm] and the fill factor from 0 to 1.

Three height values were chosen: 200 nm, 300 nm and 400 nm. Figure II-5 contains the zero-order transmission and relative phase for the chosen height values as functions of grating period and fill factor. The first column (Figures II-5 (c), (e) and (g)) has the amplitude of the transmitted zero order by the grating whereas the second (Figures II-5 (d), (f) and (h)) has the corresponding relative phase as functions of the grating period and fill factor. The first (Figures II-5 (c) and (d)), second (Figures II-5 (e) and (f)) and third (Figures IIII.2-5 (g) and (h)) lines show the results for grating height (*h*) equals 200 nm, 300 nm and 400 nm, respectively. The relative phase maps were normalized by a factor of 2π. Note that we seek to choose a unit cell size for which it is possible to phase shift light in the full range [0-2π] by changing the fill factor of the structures.

The structure that is 200 nm tall cannot make this control as Figure II-5 (d) shows. The maximum and minimum values of normalized phase shift are of approximately 0.3 and -0.25, respectively, giving a difference slightly bigger than π rad, a half than what is required. For a grating height of 300 nm the relative phase map almost covers the necessary phase range with a maximum phase difference of about 1.72π rad, as shown in Figure II-5 (f). Finally, for a grating 400 nm tall, full phase difference coverage is possible with high transmission, as shown in Figure II-5 (g) and (h). Thus, this is the chosen design to build our metasurface.

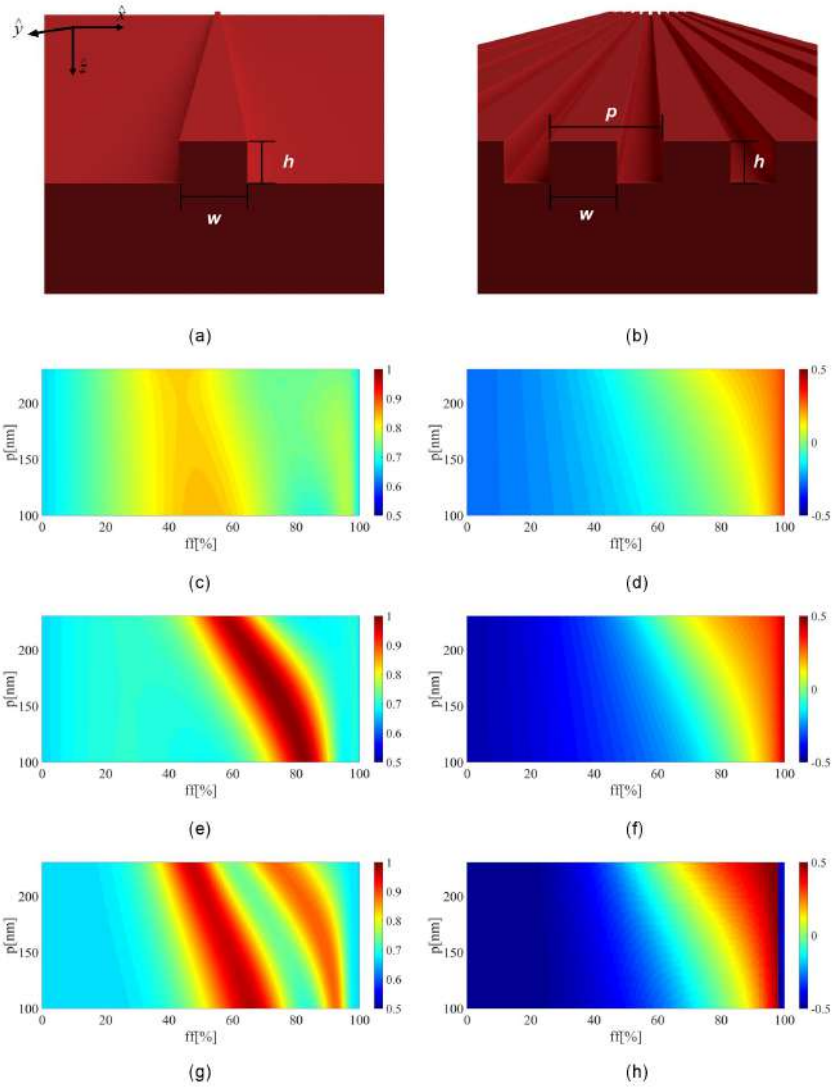


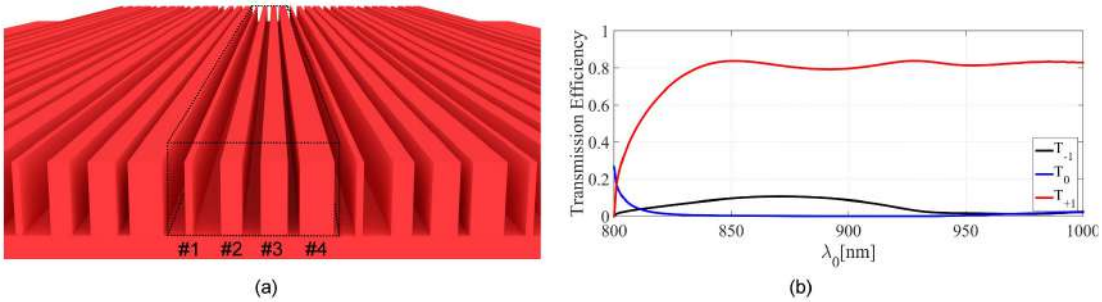
Figure II-5 - Results from the first silicon high contrast grating design. The chosen structure is depicted in (a) and the simulated structure in (b). The fill factor (ff) is defined as the ratio *w*/*p*. The first column (figures c, e and g) shows the amplitude of the transmitted zero order by the grating whereas the second (figures d, f and h) shows the corresponding relative phase as function of the grating period and fill factor. The relative phase maps in radian were normalized by a factor of 2π. The first, second and third rows have the results for grating height (*h*) equals 200 nm, 300 nm and 400 nm, respectively. The operating wavelength is 900nm at normal incidence for TM polarization (electric field polarized along x). The superstrate is air and the substrate silicon. It was used n_{Si}=3.62 for the refractive index of silicon.

Now suppose that we want to bend the transmitted light by an angle of 17.85°, using four elements. Thus, according to equation (II-7), the required unit cell size is 200 nm. Table II-1 lists the obtained structures in terms of its widths and their relative phase and transmission, obtained from the maps shown in Figures II-5 (g) and (h).

Table II-1 - Relative phase and transmission of high contrast gratings made of silicon (400 nm of height) with different values of fill factor and unit cell size of 200 nm. The wavelength is 900 nm with light coming from air to the substrate, made of silicon.

Element	*w* [nm]	Relative phase [rad/2π]	Transmission
#1	34	0.25	0.67
#2	100	0.5	0.96
#3	128	0.75	0.74
#4	160	1	0.88

Assembling the elements in Table II-1 in a supercell, results in a metasurface with positive phase gradient (note that the relative phase increases with the increasing in the element number). Figure II-6 (a) shows the assembled metasurface, the dotted line highlights one supercell. The diffraction efficiency of the metasurface was calculated with RCWA again. Figure II-6 (b) shows the resulting transmission diffraction efficiencies spectra of the -1,0 and +1 orders by the metasurface. Light is normally incident from air with TM polarization and there is nothing below the metasurface besides the substrate of silicon. As desired, most of the light is diffracted to the +1 order of the supercell, which, as we saw from equation (II-7) is equivalent to the target angle for anomalous refraction. Therefore, the proposed design is capable of phase shifting the light locally, even though the simulations were carried out using an infinite array of structures. Furthermore, this design could be applied to impart any phase profile along one direction and with the required phase levels. Nevertheless, it should be noted that a trade-off between number of phase levels and fabrication challenge must be considered. The more phase levels are required, the smaller is the minimum fill factor (ridge size) difference between the structures. Thus, a limitation is imposed by the resolution of the fabrication method used to create the metasurface.



(a) (b)

Figure II-6 - (a) Pictorial representation of the silicon metasurface supercell made of high contrast gratings. The fill factor of each numbered element is shown in Table II-1 (b) Metasurface transmission diffraction efficiency of the +1, 0 and -1 orders as function of the wavelength. Light is incident from air and below the metasurface is the silicon substrate.

III. Holographic metasurfaces

III.1. Highly efficient holograms based on c-Si metasurfaces in the visible range

Adapted with permissions from

Augusto Martins, et.al , "Highly efficient holograms based on c-Si metasurfaces in the visible range," Opt. Express 26, 9573-9583 (2018) [77] © The Optical Society

and

A. Martins, et. al, "Crystalline Silicon (c-Si) Metasurface Holograms in the Visible Range," in Imaging and Applied Optics 2018 (3D, AO, AIO, COSI, DH, IS, LACSEA, LS&C, MATH, pcAOP), OSA Technical Digest (Optical Society of America, 2018), paper DTh2E.5: [89] © The Optical Society.

A copy of the permission can be found in appendix E.1.

III.1.1. Introduction

This chapter describes the design and implementation of a dielectric metasurface for holographic applications. In 2017, we started a collaboration with Prof. Dr. Juntao Li of the State Key Laboratory of Optoelectronic Materials & Technologies from Sun-Yat Sen university at China. At the beginning of 2017, the group of Prof. Dr. Li demonstrated a c-Si based metasurface with a low aspect ratio of 2.7, reaching efficiencies of 71% in transmission at 532 nm [53]. This result is of relevant scientific impact because c-Si stands out as arguably possessing the best combination of properties for dielectric metasurfaces: low losses with high refractive index at the visible range. It should be noted that, the higher is the refractive index, the smaller is the effective wavelength inside a given structure, thus resulting in a higher field confinement and smaller structures (see section *B.1 Bloch modes* on the Appendix for more details). For instance, TiO_2 based structures, which presents a smaller refractive index than c-Si, require aspect ratios larger than 10 [62], which are difficult to fabricate. Furthermore, most of the published works on silicon based metasurfaces relied either on its amorphous (a-Si) [63] or on its polycrystalline (p-Si) [52] varieties, which is disadvantageous since both present high losses in the visible spectrum [61].

Therefore, it was decided to exploit c-Si based holographic metasurfaces in collaboration with Prof. Dr. Juntao Li's group. In the collaboration, I designed the metasurfaces and sent the design to Prof. Li's group for fabrication. Both the numerical and experimental characterizations were performed by the author. Two designs have been proposed: cylindrical and elliptical ones.

Both have been successfully proved to be a suitable platform for highly efficient holograms operating in the visible range.

III.1.2. Metasurface design

The metasurfaces were designed using cylindrical c-Si nanopost as the individual resonators, as shown in Figure III-1 (a). Also shown in this figure is a representation of the two substrates used: quartz and sapphire. In the first design, the cylinders were patterned on a 220 nm thick layer of c-Si and are over a 1 μm thick silica layer that is bonded to the quartz substrate by an also 1 μm thick NOA61 adhesive. The additional silica and adhesive layers are residual from the fabrication process, which is based on the layer-transfer technique (see section *III.1.6.a.1 Fabrication methods – Quartz* or [53] for details). The second design, with sapphire substrate is represented in Figure III-1 (d), was fabricated on a commercially available 230 nm thick c-Si (100), epitaxially grown on a sapphire substrate from UniversityWafer (see section *III.1.7.a.1 Fabrication method – Sapphire* details).

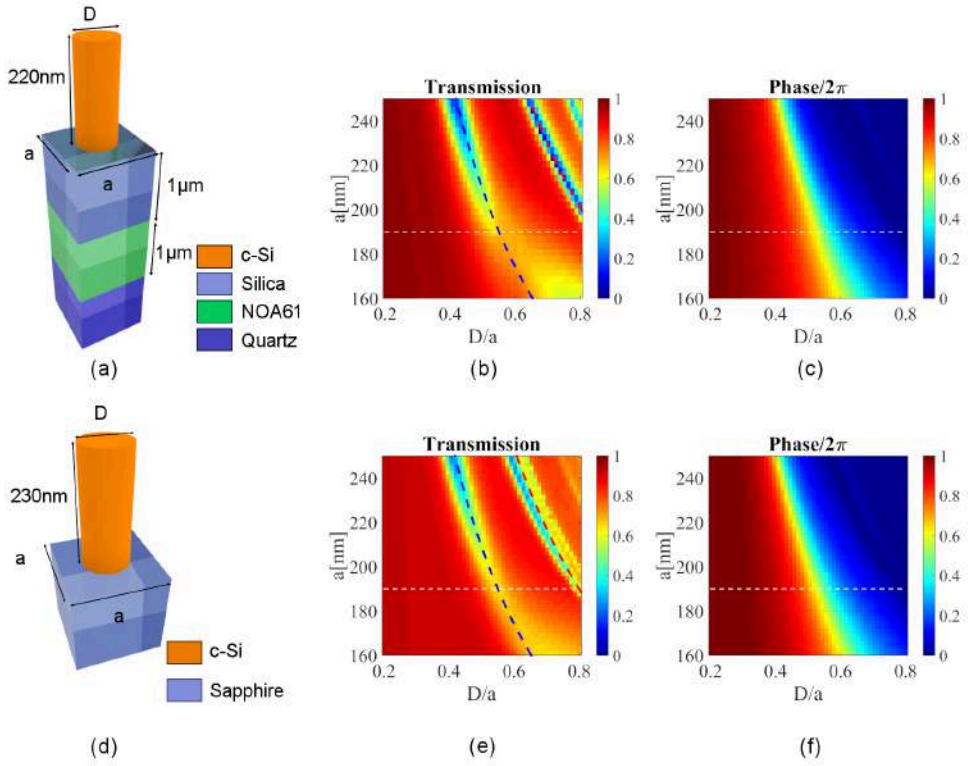


Figure III-1 – (a) and (d) show, respectively, the unit cell of the cylindrical c-Si structures over quartz and sapphire substrates. (b) and (c) [(e) and (f)] show the zero order transmission and relative phase divided by 2π, respectively, of the design over quartz [sapphire] substrate as functions of the unit cell size and the nanopost diameter to unit cell size ratio, *D/a*. The dashed blue [red] lines in (b) and (e) mark the regions where the diameter is equal to 105 nm [105 nm] and 158 nm [160 nm], respectively. The dashed white lines mark the regions with $a = 190$ nm, used for both designs. The diffraction efficiencies and relative phase shift along the dashed lines are shown in Figure III-2. The operating wavelength is 532 nm.

The operating wavelength was chosen to be 532 nm for all cases. In order to assess the metasurface sensitivity to wavelength variations, after the design was optimized for the wavelength of 532 nm, additional simulations were performed at 444.9 nm and 633 nm. The complex refractive indexes of c-Si at 444.9 nm, 532 nm, and 635 nm are, respectively, n = 3.875 + j0.0158, n = 4.141 + j0.032 and n = 4.733 + j0.099 [90], with the index of the remaining materials assumed constant (quartz n = 1.45, silica n = 1.45, adhesive n = 1.56 and sapphire n = 1.77).

The resulting zero order transmission efficiency and phase shift maps as a function of the cylinder *D* diameter to unit cell size *a* ratio, *D/a*, and of the nanopost diameter, *D*, are shown in Figures III-1 (b) and (c) for the structures on the quartz substrate. Figures III-1 (e) and (f) show the same parameters, but for the structures on sapphire substrate. From the phase maps of both structures (Figures III-1 (c) and (f)), by varying the diameter of the nanoposts, full-phase control [0-2π] is attained for all *a* in the range analysed. Nevertheless, the transmission maps (Figures III-III.1-1 (b) and (e)) present a valley that gets more accentuated for periods higher than ~200 nm. Therefore, it was decided to use *a* = 190 nm (marked with a dashed white line in Figures III-1 (b), (c), (d) and (f)). Apart from the high transmission, another important aspect taken into account for the choice of *a* = 190 nm was the resulting low aspect ratios.

The valleys on both designs are due to the Fabry-Perot resonance in the post for those particular structures. Interestingly, the first resonance occurs for fixed diameters of 105 nm (blue dashed lines on Figures III-1 (b) and (e)) for periods larger than 170 nm, approximately. The second resonance, which is marked with red dashed lines on Figures III-1 (b) and (e), occur for diameters of 150 *nm* (nanopost over quartz) and 160 nm (nanopost over sapphire), respectively. In other words, this resonance depends mainly on the nanopost geometry and material and not much on the array. This conclusion is corroborated by the field profile of the Bloch modes supported by this structure. Since the posts have high index, the field tends to be more concentrated inside the posts, thus minimizing the interaction between adjacent structures. As the period is reduced, the overlap of the evanescent field outside adjacent nanoposts increases, which also increases the coupling among posts and affects the resonances. A detailed analysis of this design in terms of Bloch modes is found in Appendix B, section *B.1 Bloch modes* and the field distributions of the resonant nanoposts for unit cell sizes of 190 nm and 220 nm are shown in Appendix B, section *B.2 Nanoposts resonances*. At the resonance, light bounces back and forth inside the structures, which enhances the phase accumulation and the field inside the nanopost. The increased field inside the structures also increases their absorption losses, which can be seen by the minima in Figures III-1 (b) and (e).

Finally, in Figures III-2 (b) and (e), it is shown the transmitted zero order phase and efficiency of the designs with posts over quartz and sapphire substrates, respectively, and *a* = 190 nm as function of *D/a* at 532 nm. Additionally, Figure III-2 shows the absorption and reflection

efficiencies as function of the diameter. It is known that, the more phase levels are used, the better are the CGH reconstruction signal to noise ratio and diffraction efficiency (see appendix *D.3.1 Example: Fourier hologram calculation*). Note, however, how that the phase curve is highly slanted for diameter values in the range [80 nm, 120 nm] (see Figures III-2 (b) and (e)). Therefore, the more phase levels are used, the smaller is the diameter differences between different structures in this range. This imposes a trade-off between CGH quality and fabrication feasibility. We found that a good trade-off between quality and fabrication constrains can be reached with eight phase levels (see appendix *D.3.1 Example: Fourier hologram calculation*). The corresponding selected phases and transmission values are marked with dots in Figure III-2.

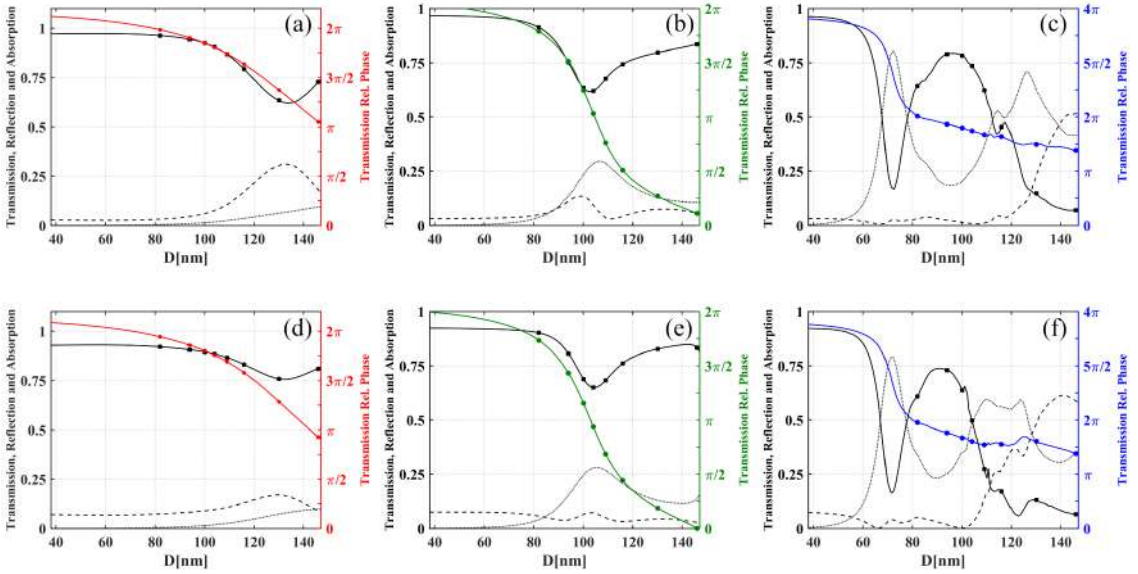


Figure III-2 – Zero order transmission efficiency (solid black) and phase (solid coloured), reflection efficiency (dashed) and absorption (dotted) of the metasurface designs over a quartz (upper row) and sapphire (bottom row) substrates as function of the diameter of the nanopost with unit cell size a = 190 nm. The results are shown for operating wavelengths of 633 nm, 532 nm and 444.9 nm at the first, second and third columns.

To assess the bandwidth of the holograms, Figure III-2 also shows the same results operating at 633 nm (Figures III-2 (a) and (d)) and at 444.9 nm (Figures III-2 (c) and (f)). Clearly, at these wavelengths, the phase control is reduced. Nevertheless, the transmissions are still quite high, reaching 0.79 at 449.9 nm and 0.98 at 633 nm for the chosen structures. Additionally, note that all curves present only one Fabry-Pérot resonance that is shifted toward smaller diameters as the wavelength is reduced. As shown in section *B.2 Nanoposts resonances* of the appendix, this resonance is characterized by an enhancement of the field inside the structure that leads to an increasement in the absorption because silicon is lossy. Additionally, this resonance is marked by a strong phase shift at the transmitted light. These features can be seen by the valleys of the transmission functions (maximum of absorptions) at 130 nm (Figures III-2 (a) and (d)), 100 nm (Figures III-2 (b) and (e)) and at 70 nm (Figures III-2 (c) and (f)). Note how the absorption peak increases at the resonance as the wavelength reduces since the imaginary part of c-Si increases as

well. The transmission minimum for diameters bigger than 100 nm and operating at 449 nm is only due to the increased fill factor of c-Si, whose absorption is much larger at this wavelength.

III.1.3. Hologram designs

The holograms were calculated with the iterative Fourier Transform algorithm (IFTA), described in appendix *D.3 Iterative Fourier Transform Algorithm (IFTA)*. Two CGHs with eight phase-levels, number of pixels 512×512 and 1024×1024 pixels, were calculated using target images shown in Figures IIIIII.1-3 (a) and (b), respectively. These images have resolutions of, respectively, 123×159 and 257×213 pixels. The image window (that is, the region of the reconstruction plane where the image is placed and is defined at Figure D.2 (b) in appendix D.3) of the CGH relative to Figure III-3 (a) was centralized with the reconstruction plane, whereas the image window relative to Figure III-3 (b) was vertically displaced by 190 pixels to separate the image from the spurious zero order that appears in the reconstruction of the metasurface hologram.

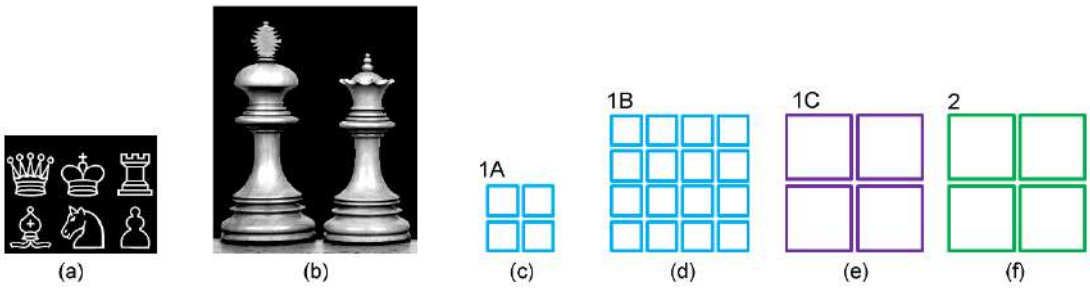


Figure III-3 – (a) and (b) show the target images of the first and second holograms, respectively. (c)-(f) show the metasurfaces assembly composed with arrays of the calculated CGHs. Each square corresponds to a CGH. In (c) and (d) they represent the first hologram, with 512×512 pixels, recorded on a c-Si metasurface and with each pixel represented by a single nanopost cell with dimension of 190 nm × 190 nm. The squares at (e) also represent the first CGH but encoded in a c-Si metasurface where each pixel is represented by an array of 2×2 nanopost cells with a total dimension of 380 nm × 380 nm. Finally, in (f) each square represents the second CGH, with 1024×1024 pixels, recorded on a c-Si metasurface where each pixel is a single nanopost cell with dimensions of 190 nm × 190 nm. Therefore, metasurfaces 1B, 1C and 2 all have dimensions of 389.12 µm × 389.12 µm whereas 1A has dimensions of 194.56 µm × 194.56 µm.

It should be noted that the first metasurface that was fabricated encoded the CGH related to Figure III-3 (a) and, during its experimental and numerical reconstructions, we noted a wide-angle aberration, as explained in appendix *D.4 Rigorous Rayleigh Sommerfeld (RS) Diffraction*. Thus, in the other CGH, a wide-angle correction was performed prior to its calculation to avoid this issue. This correction depends on several parameters related to the CGH design such as pixel size, operating wavelength, propagation distance and image window position in the reconstruction plane [91].

The resulting reconstruction field distributions amplitudes obtained from the IFTA algorithm using the Fresnel diffraction approach of equation (D.13) are shown in Figure III-4. Note how the reconstruction of the second hologram, shown in Figure III-4 (b), looks distorted in comparison with its target image, shown in Figure III-3 (b). This is due to the wide-angle correction that was performed on it. Furthermore, because the Fresnel Diffraction does not suffer from wide angle aberration, its reconstruction does not compensate for it.

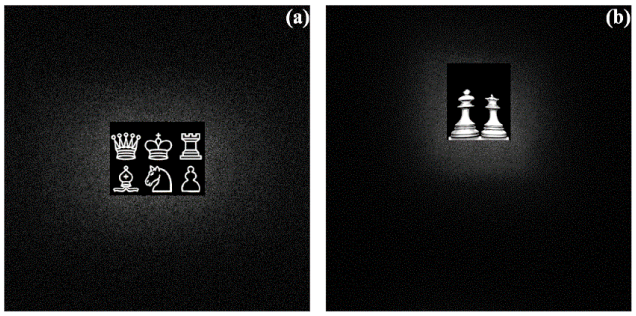


Figure III-4 – (a) and (b) field distribution amplitude at the reconstruction plane of the first and second CGS, respectively. Figures are not to scale.

The resulting figures-of-merit of these CGHs are listed in Table III-1. The image efficiencies, that is, the amount of power that is cast inside the image window, are over 70% for both CGHs. Note that such high efficiencies are accompanied of elevated SNRs, whose base ten logarithms are over 34.

Table III-1 – Obtained figures of merit, defined in Table D-1 in the appendix, of the CGHs whose target images are shown in Figures III-3 (a) and (b).

CGH Target Image	Figure III-3 (a)	Figure III-3 (b)
Image Efficiency (%)	73.48	77.3
$\log_{10}$(SNR)	34.04	35.55
Mean Squared error	3.9e-4	2.78e-4

Each hologram was encoded in a c-Si metasurface by correlating its phase distribution with a given nanopost diameter according to Figure III-2 (b). This procedure was made using four different approaches to analyse their performances in terms of efficiency and reconstruction quality and are shown in Figures III-3 (c)-(f). The first CGH, whose target image is shown in Figure III-3 (a), was encoded in metasurfaces 1A,1B and 1C. Metasurfaces 1A (Figure III-3 (c)) and 1B (Figure III-3 (d)) are made of arrays with, respectively, 2×2 and 4×4 CGHs and each pixel of the CGHs is represented by a single cell of size 190 nm × 190 nm with one nanopost. 1A was fabricated only over the quartz substrate and the remaining over both substrates. The metasurface 1C (Figure III-3 (e)), in turn, is composed by an array of 2×2 CGHs and each pixel of the CGHs is composed of an array of 2×2 nanoposts having a size of 380 nm×380 nm. Finally, metasurface 2 (Figure III-3 (f)), encodes an array of 2×2 CGHs whose target image is shown in Figure III-3

(b), where each pixel of the CGHs is represented by a single nanopost cell sizing 190 nm×190 nm. Therefore, as the first and second CGH have 512×512 and 1024×1024 pixels each, metasurfaces 1B,1C and 2 have dimensions of 389.12 μm × 389.12 μm whereas 1A has dimensions of 194.56 μm × 194.56 μm.

III.1.4. Numerical Characterization

The numerical characterization of the holograms was made using the Rayleigh-Sommerfeld (RS) integration [3], with the reconstruction plane positioned 20 cm away from the hologram. The calculation details are explained in appendix *D.4 Rigorous Rayleigh Sommerfeld (RS) Diffraction*. Note that this method is computationally very slow as the involved field matrices have millions of pixels. Therefore, an optimization procedure based on some symmetries between the reconstruction and hologram planes was developed. The procedure is thoroughly described in appendix *D.4.1 RS Integration.* A performance comparison is provided in section *D.4.1.f. Performance comparison*. The reconstructions were performed for wavelengths of 444.9 nm, of 532 nm and of 633 nm. Two approaches were used to perform the numerical reconstructions. In one, called "c-Si approximation", the phase and amplitude values of each CGH pixel follow the values of the metasurface design according to Figure III-2. In the other approach, called "Ideal", it is assumed that the transmission of each pixel is unitary, and the phase shift of each pixel follows the exact value from the CGH. The reconstructions were made only for designs using a quartz substrate, as their transmission and phase maps are very similar to those of the designs using a sapphire substrate. The transmission efficiency is defined as the ratio between the transmitted power by the hologram to the transmitted power with the metasurface removed [53], and the diffraction efficiency is defined as the ratio between the power in the image window to the power transmitted with the metasurface removed. The definition of the image efficiency is the same as the definition of the IFTA: the ratio between the power at the image window to the total power transmitted with the metasurface. The power calculations are described in Appendix *D.4.2 Power calculation of the diffracted fields*. The resulting numerical reconstructed fields will be shown side by side with the experimental ones in the next section.

The calculated diffraction and transmission efficiencies of the analysed metasurfaces are listed in Table III-2. Note that the Ideal metasurfaces have transmission efficiencies of 100% for all wavelengths, but this is not listed in Table III-2. Therefore, the diffraction efficiency of the Ideal metasurface has the same result of the image efficiency used as a figure of merit to quantify the CGH performance in the IFTA. For the first CGH, used in metasurfaces 1A, 1B and 1C, the expected image efficiency from the IFTA was 73.48%, which is almost equal to the result obtained by RS integration operating at 444.9 nm and 532 nm, according to Table III-2. At the wavelength of 633 nm, the diffraction efficiencies were 68.3% and 66.4% with metasurfaces 1A

and 1B, respectively, which are still relatively high values. On the contrary, the c-Si metasurface has lower efficiencies because it considers the transmission and phase values of the design into the CGH. The main effect of the transmission modulation in the pixels is the reduction of the overall transmission, which had marginal variations between different metasurfaces, remaining near 53%, 72% and 85% for the wavelengths of 444.9 nm, 532 nm and 633 nm, respectively. These results manifest directly the fact that c-Si losses are larger for smaller wavelengths [92]. But that does not mean that the operation of the metasurface was enhanced at 633 nm. In fact, this behaviour does not appear in the diffraction efficiency, which reaches 16.2%, 52.2% and 39.8% when operating at 444.9 nm, 532 nm and 633 nm respectively. This result is a consequence of the phase shift imparted by the nanoposts, optimized to operate at 532 nm. At the other wavelengths, the resulting phase control is worsened, as can be seen in Figure III-2, with the worst control at 444.9 nm. The poorer phase modulation deteriorates the CGH phase distribution at the metasurface, which results in more energy being casted as noise, zero order and the symmetric Hermitian copy of the image, which are not considered in the diffraction and image efficiencies. Furthermore, the reduced diffraction efficiency of the metasurface operating at 532 nm is only affected by the fact that the pixels are faded due to the transmission curve of the c-Si nanoposts. In fact, the image efficiency, that is, the ratio between energy in the image window to total energy transmitted by the device, of the c-Si metasurfaces at this wavelength is almost equal to the image efficiency of the ideal one, reaching 72.5% for metasurface 1C. In contrast, the image efficiencies of the metasurfaces at 633 nm [444.9 nm] range from 25% to 46.7% [9.76% to 16.2%], evidencing once again the manifestation of the poorer phase modulation of the metasurfaces at these wavelengths.

Table III-2 – Numerical transmission, diffraction and image efficiencies of the designed metasurfaces operating at 444.9 nm, 532 nm and 633 nm. The ideal metasurfaces have transmission efficiencies of 100%.

	Ideal			c-Si Approximation								
	Difraction | Image			Transmission			Diffraction			Image		
λ_0 (nm)	444.9	532	633	444.9	532	633	444.9	532	633	444.9	532	633
1A	71.8	70.2	68.3	53.1	72.0	85.1	14.5	50.5	32.4	27.3	70.1	38.0
1B	71.1	70.2	66.4	53.1	72.0	85.1	13.6	50.5	31.6	25.6	70.1	37.1
1C	72.6	72.9	71.9	53.1	72.0	85.1	16.2	52.2	39.8	30.5	72.5	46.7
2	77.2	70.9	68.0	53.0	72.0	85.0	9.76	50.8	21.3	18.4	70.5	25.0

Finally, in order to quantify the hologram fidelity, we calculated the signal-to-noise ratio (SNR) of the proposed metasurfaces. This quantity is defined as [75]

$$SNR = 10\,log_{10}\left(\frac{\sum_{\Omega}|M|^2}{\sum_{\Omega}(|I|-|M|)^2}\right)dB = 10\,log_{10}\left(\frac{1}{\sum_{\Omega}(|I|-|M|)^2}\right)\,dB \qquad \textbf{(III-1)}$$

where Ω is the image window, and *I* and *M* are the normalized field distributions of the ideal and c-Si metasurfaces, respectively. The *sinc* function that appears due to the zero-order

diffraction in the origin is neglected in this calculation. It should be noted that both *I* and *M* are normalized in such a way that their power in Ω are the same (this is why the sum in the numerator equals 1). Thus, the mean squared error gives a measure of how much the metasurface reconstruction deviates from the ideal case without taking the intensity of the reconstructed image into account.

The results obtained with the SNR calculations are listed in Table III-3. The SNR is larger than 34 dB for all cases at 532 nm, resulting in high reconstruction fidelity. This means that the amplitude modulation has little impact on the image reconstruction. In contrast, the SNR reach ~11 dB and ~10 dB at 444.9 nm and 635 nm, respectively.

Table III-3 – Signal-to-noise ratio (SNR) of the four metasurfaces investigated in this work at different wavelengths.

	SNR (dB)		
Samples	444.9 nm	532 nm	633 nm
1A	10.67	34.75	7.72
1B	10.45	34.76	7.79
1C	10.63	34.79	8.14
2	11.77	37.05	10.62

III.1.5. Experimental Characterization

The holograms are characterized with two different optical setups, one for reconstructing their images and other for measuring their transmitted and diffracted powers. The first setup, shown in Figure III-5 (a), consists of a solid-state laser, an iris (to block unwanted scattered light by the optical interfaces), and two lenses (L_1 and L_2 with focus f_1=7.5 cm and f_2=2.5 cm, respectively). The lenses are arranged as a Keplerian telescope to reduce the beam waist size to a collimated spot diameter of ~400 μm. All metasurfaces have an area of 389.12 μm × 389.12 μm. The reconstruction plane is located 20 cm away from the hologram in all cases. All reconstructions are captured with a camera (Nikon Coolpix p100) positioned in front of the reconstruction plane. The power measurement procedure is carried out as illustrated in Figure III-5 (b), with the same optical source of (a), a lens L_3 (f_3=25 cm), and an iris. The sample is positioned near the lens focus so that a beam waist of ~200 μm is obtained (smaller than the hologram size) [44]. The total transmitted power is measured with the power meter detection head (Thorlabs S120C) positioned in front of the metasurface at position P_1 in Figure III-5 (b) (this guarantees that the power is focused onto the detector surface). The zero-order transmitted power is measured by moving the detection head to the zero-order spot position represented by position P2 in Figure III-5 (b). Figures III-5 (c) and (d) show photos of the used reconstruction and power measurement setups.

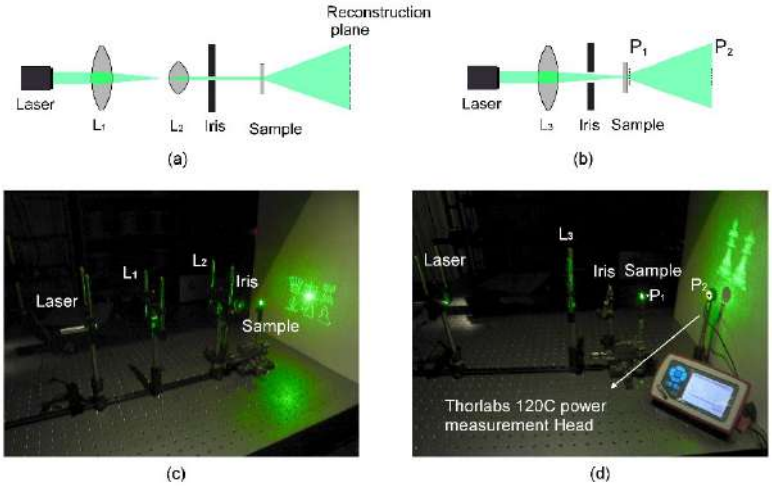


Figure III-5 - (a) and (b) measurement setup used for the holograms' image reconstruction and power measurements, respectively. Note that in (a) the lenses L1 and L2 are arranged as a Keplerian telescope. (c) and (d) photos of the used reconstruction and power measurement setups.

The experimental reconstruction of the metasurfaces and the measured efficiencies will be shown next. The results relative to each design (quartz and sapphire substrates) are shown separately.

III.1.6. Quartz Substrate

a) Fabricated Metasurfaces – Quartz

a.1. *Fabrication methods – Quartz*

The fabrication steps for the metasurfaces on the quartz substrate are represented in Figure III-6 and are based on the transfer layer technique [53]. In this process, the c-Si layer from a SOITEC SOI wafer (Figure III-6 (a)) is transferred to a quartz substrate. Initially, a 1 μm silica is deposited over a wafer with a c-Si layer (220 nm height) over the insulating silica layer (1 μm) for protection (Figure III-6 (b)). A UV light curable NOA61 adhesive is then spin coated on the silica layer (Figure III-6 (c)), followed by bonding to a quartz substrate (Figure III-6 (d)). The structure is subsequently illuminated with UV light to cure the adhesive for two hours and then it is baked for two days to increase the adherence (Figure III-6 (d)). The silica layer is exposed by milling the silicon handle followed by DRIE (deep reactive ion etching) (Figure III-6 (e)). Finally, the c-Si layer is exposed by removing the silica isolating layer using wet etching (Figure III-6 (f)).

The structure at this point is now ready to be patterned. The structure is patterned on a HSQ electron beam resist using an electron beam generator (Vistec EBPG 5000+ at 100 KeV - Figure III-6 (g)-(h)). After developing the resist, the pattern is transferred from the resist to the silicon layer using inductively coupled plasma etching (ICP) (PlasmaPro System 100ICP180) (Figure

III-6 (i)). The remaining HSQ is removed with Hydrofluoric acid. The scanning electron micrographs (SEM) of the metasurfaces are shown in the next section.

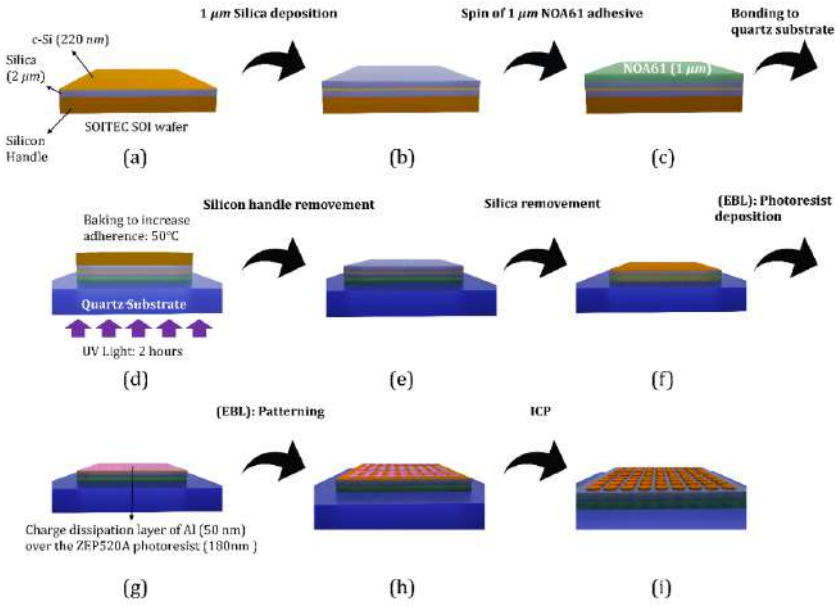


Figure III-6 - Fabrication steps of the metasurfaces fabricated over quartz substrate. (a) SOITEC SOI wafer. (b)-(i) shows the steps involved on the fabrication. This image was based on [53]. (EBL) Electron beam lithography. (ICP) Inductively coupled plasma.

a.2. *Metasurface 1A*

The metasurface 1A was fabricated twice on the same sample. Figure III-7 shows pictures of the sample with the metasurfaces and their SEM micrographs under different scales. The SEMs indicate that the metasurfaces present good quality in their fabrications (see in Figures III-7 (b) and (c)).

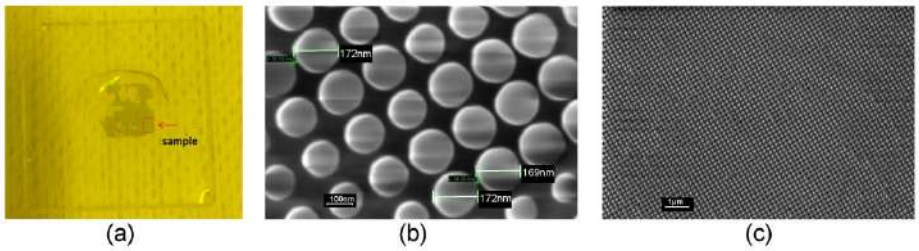


Figure III-7 – (a) Picture of the sample where the two 1A metasurfaces were fabricated. The red region arrow points their locations on the sample. (b) and (c) show SEM micrographs of one of the structures in different scales.

a.3. *Metasurfaces 1B, 1C and 2*

The metasurfaces 1B, 1C and 2 were fabricated on the same substrate but on four different samples, that will be called Q1, Q2, Q3 and Q4. Figures III-8 – III-15 show pictures and SEM micrographs of the fabricated samples. The quality of the metasurfaces, as shown by the micrographs, impact the performance of the metasurfaces as our efficiencies and reconstructions measurements show in the next sections. Note that, unfortunately, there are some coalescence in some regions of the metasurfaces while in others the nanoposts fell on the substrate. Furthermore,

we noted that there is some variation between the target diameters and the fabricated ones of about 10 nm, which also impacts the performance of the structure. Nevertheless, there are regions where the posts present good quality despite the diameter variation. Thus, even with these imperfections, we obtained satisfactory results in terms of both diffraction efficiency (reaching 29.5% in 532 nm) and image quality.

a.3.1. *Sample Q1*

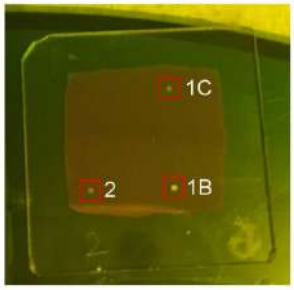


Figure III-8 – Picture of sample Q1 with metasurfaces 1B, 1C and 2, circumscribed by red squares.

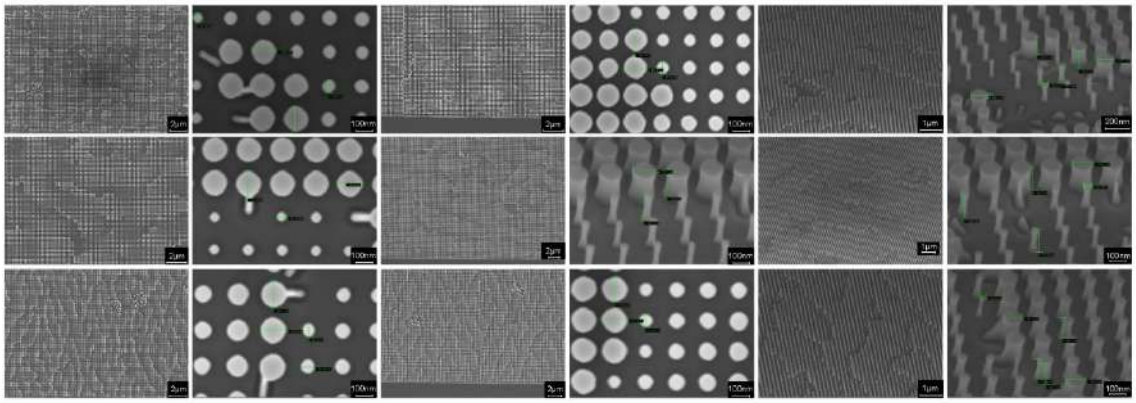


Figure III-9 – SEM micrograph of metasurfaces 1B (first row), 1C (second row) and 2 (third row) at different scales and positions of Q1 sample.

a.3.2. *Sample Q2*

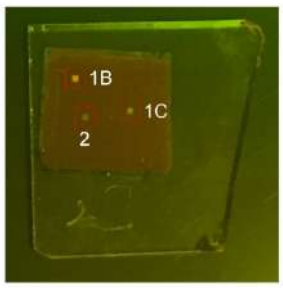


Figure III-10 – Picture of sample Q2 with metasurfaces 1B, 1C and 2, circumscribed by red squares.

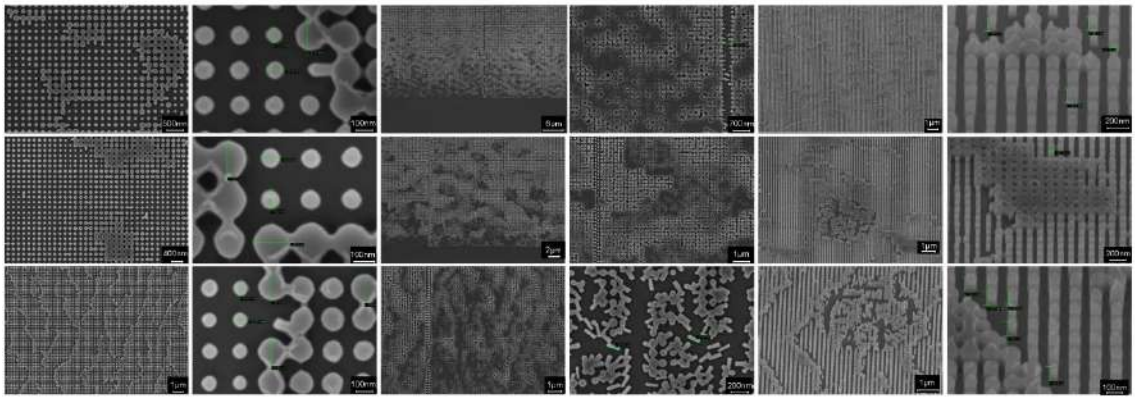


Figure III-11 – SEM micrograph of metasurfaces 1B (first row), 1C (second row) and 2 (third row) at different scales and positions of Q2 sample.

a.3.3. *Sample Q3*

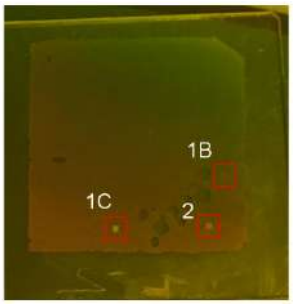


Figure III-12 – Picture of sample Q3 with metasurfaces 1B, 1C and 2, circumscribed by red squares.

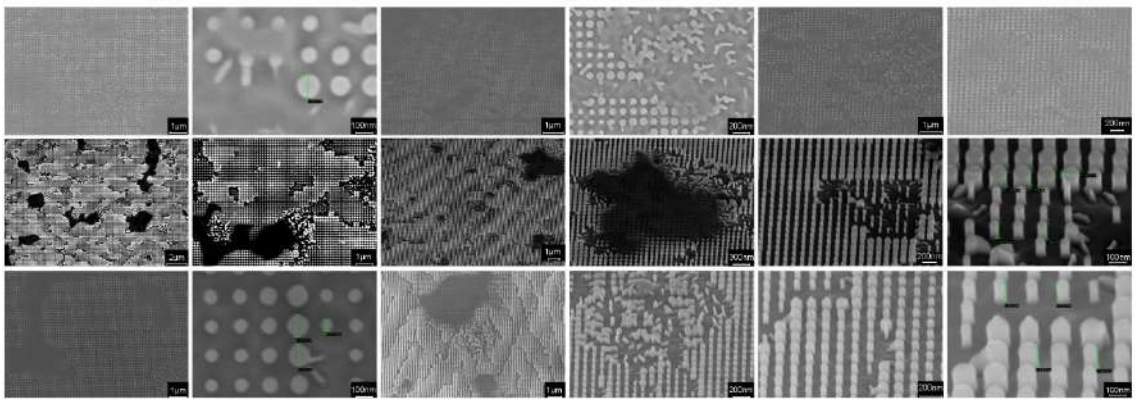


Figure III-13 – SEM micrograph of metasurfaces 1B (first row), 1C (second row) and 2 (third row) at different scales and positions of Q3 sample.

a.3.4. *Sample Q4*

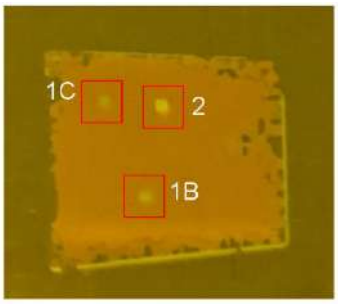


Figure III-14 – Picture of sample Q4 with metasurfaces 1B, 1C and 2, circumscribed by red squares.

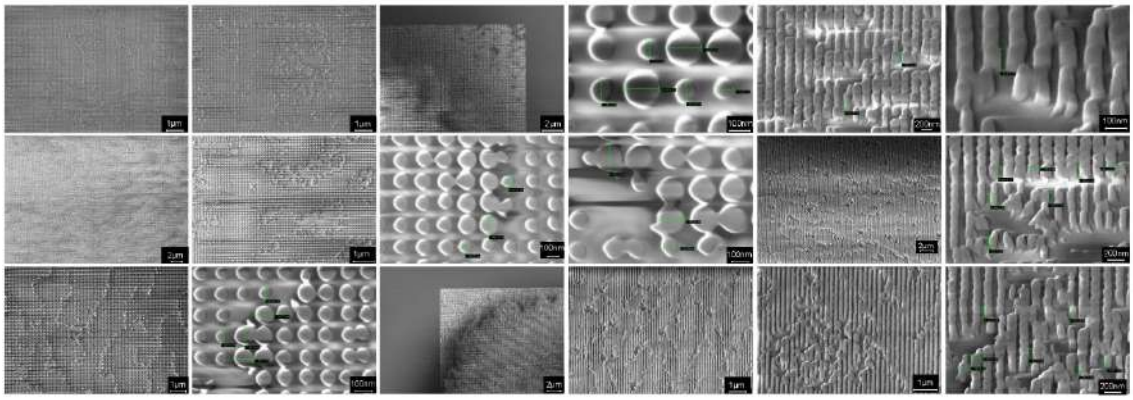


Figure III-15 – SEM micrograph of metasurfaces 1B (first row), 1C (second row) and 2 (third row) at different scales and positions of Q4 sample.

b) Efficiencies

The measured transmission and diffraction efficiencies of metasurface 1A and samples Q1, Q2, Q3 and Q4, with metasurfaces 1B, 1C and 2 are listed in Tables III-4, III-5, and III-6 for the wavelengths of 532 nm, 444.9 nm and 633 nm respectively. Three main factors have an impact on the diffraction efficiency: 1) The amplitude and phase maps do not afford the ideal modulation required by the hologram. 2) The local character of the resonances: the amplitude and phase maps refer to infinite periodic arrays, but only a few periods are used in the metasurface to create the local resonance (often only a single period is used). Here, materials with high refractive index, such as c-Si, offer an obvious advantage, as they promote localization of the resonance. 3) Fabrication imperfections that induce noise in the phase and amplitude modulation.

As theoretically predicted (see Figure III-2 and Table III-2), the transmission efficiency increases with wavelength for all cases. This behaviour is also observed for the measured transmission and diffraction efficiencies. It is worth noting that the measured diffraction efficiencies are smaller than the numerical ones at the target wavelength of 532 nm, and the opposite is true for the transmission efficiencies. This is mostly because the fabricated metasurfaces have nanoposts whose diameters are on average 10 nm smaller than originally designed. Since the nanopost phase-diameter response is not linear (see Figure III-2), its relative phase is not maintained when the diameter is varied. The metasurface then loses more power to the zero order and noise, therefore reducing its diffraction efficiency. Note that, by modulating the pixel intensity via the c-Si nanopost array transmission, the diffraction efficiency is reduced by almost 20% for all metasurfaces at 532 nm (compare the c-Si and Ideal metasurfaces approaches in Table III-2). Nonetheless, the diffraction efficiency is still high at this wavelength.

Although the phase control is worsened at 444.9 nm and 633 nm (see Figure III-2), the obtained transmission efficiencies still reach 51.8% and ~63.9%, respectively, which is quite remarkable for the visible range and shows the suitability of c-Si for metasurfaces operating in the visible. Note that the diffraction efficiencies at 444.9 nm and 633 nm are about half of those

at 532 nm. This means that most of the transmitted energy does not contribute to the image reconstruction at these wavelengths, since it is essentially lost to the zero-order beam at the origin due to the poor phase modulation.

Table III-4 – Measured transmission and diffraction efficiencies of the fabricated metasurfaces over quartz substrate operating at 532 nm.

Sample	Metasurface	Transmission	Diffraction
First Sample (metasurface 1A)	Upper	51.8%	25.0%
	Bottom	42.9%	30.4%
Q1	1B	27.6%	9.8%
	1C	33.3%	14.3%
	2	25.7%	9.5%
Q2	1B	27.9%	16.9%
	1C	29.9%	20.6%
	2	28.4%	20.3%
Q3	1B	No reconstruction	
	1C	28.6%	19.7%
	2	27.1%	19.4%
Q4	1B	41.9%	29.5%
	C	38.1%	28.6%
	2	37.6%	26.2%

Table III-5 – Measured transmission and diffraction efficiencies of the fabricated metasurfaces over quartz substrate and operating at 444.9 nm.

Sample	Metasurface	Transmission	Diffraction
First Sample (metasurface 1A)	Upper	24,9%	11,2%
	Bottom	17,3%	7,5%
Q1	1B	17,8%	8,3%
	1C	18,7%	8,7%
	2	17,4%	8,7%
Q2	1B	10,8%	7,6%
	1C	13,9%	10,3%
	2	11,7%	7,4%
Q3	1B	No reconstruction	
	1C	14,7%	10,5%
	2	12,9%	8,9%
Q4	1B	51,8%	16,4%
	1C	50,0%	18,6%
	2	49,5%	16,8%

Table III-6 – Measured transmission and diffraction efficiencies of the fabricated metasurfaces over quartz substrate and operating at 633 nm.

Sample	Metasurface	Transmission	Diffraction
First Sample (metasurface 1A)	Upper	50,0%	26,3%
	Bottom	52,6%	26,3%
Q1	1B	48,6%	8,6%
	1C	45,3%	8,3%
	2	42,9%	8,6%
Q2	1B	37,1%	13,1%
	1C	42,9%	23,7%
	2	34,3%	19,7%
Q3	1B	No reconstruction	
	1C	47,2%	8,3%
	2	41,7%	8,3%
Q4	1B	58,3%	25,0%
	1C	52,8%	19,4%
	2	63,9%	30,6%

c) Reconstructions - Quartz

Only the reconstructions of metasurfaces 1A upper and Q4 1B, 1C and 2 are shown because the other ones are very similar to those but noisier and with stronger zero orders. In other words, only the best reconstructions are shown. Figures III-16, III-17 and III-18 show the numerical ideal and experimental reconstructions of those metasurfaces operating at 532 nm, 444.9 nm and 633 nm, respectively (these figures are not to scale). Samples 1A and 1B use only one nanopost per pixel and the wide-angle aberration is present operating in all wavelengths [Figures III-16, III-17 and III-18 (a)-(d)], these figures are not to scale. Note that the wide-angle aberration increases with wavelength because the angular spread also does. The reconstructions from sample 1C, in turn, [Figures III-16, III-17 and III-18 (e)-(f)] do not appear distorted even without wide-angle correction at all wavelengths. This is due to its larger pixel size (2×2 array of nanoposts per pixel), which results in a smaller angular spread [3].

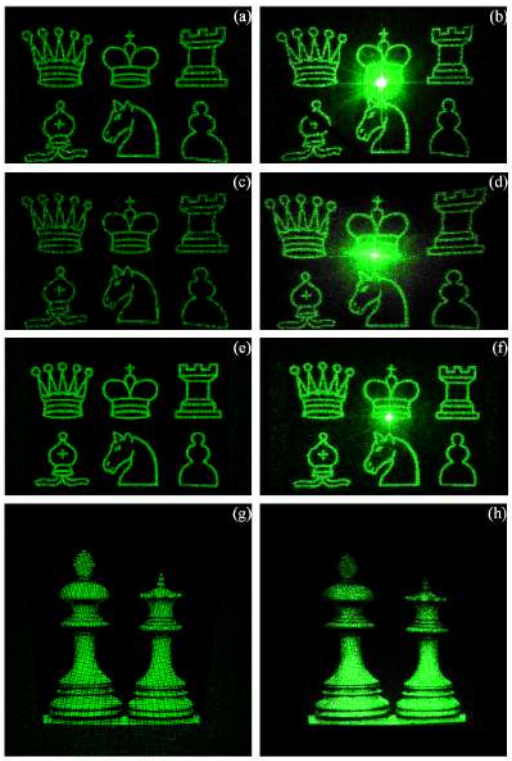


Figure III-16 – 1A [(a) and (b)], 1B [(c) and (d)], 1C [(e) and (f)] and 2 [(g) and (h)] numerical (first column) and experimental (second column) reconstructions operating at 532 nm.

Metasurface 2 uses only one nanopost per pixel and wide-angle distortion correction [Figures III-16, III-17 and III-18 (g)-(h)]. Observe that the correction performed in this sample is carried out only for the operation wavelength (532 nm), which makes the reconstructed images from it to appear distorted for the other wavelengths, namely 444.9 nm [Figures III-17 (g)-(h)] and 633 nm [Figures III-18 (g)-(h)]. Note how the image is warped at 444.9 nm (compare with Figure III-3 (b)). In this case, the reconstruction size is smaller than at 532 nm. Therefore, the wide-angle distortion is weaker and does not compensate for the correction performed previously on the target image. On the other hand, at 633 nm, the reconstruction of metasurface 2 [Figures III-18 (g)-(h)] is wider than at 532 nm and the wide-angle correction is not enough to compensate for the wide-angle distortion. Therefore, the reconstruction has a stretched aspect (compare with Figure III-3 (b)).

The holograms at the wavelengths of 444.9 nm and 633 nm still display relatively high brightness and quality, as shown in Figures III-17 (g)-(h) and III-18 (g)-(h), respectively, which is an interesting observation given how far these wavelengths are from the target wavelength. These results show that holograms based on c-Si metasurfaces can be effectively applied for full-color holograms.

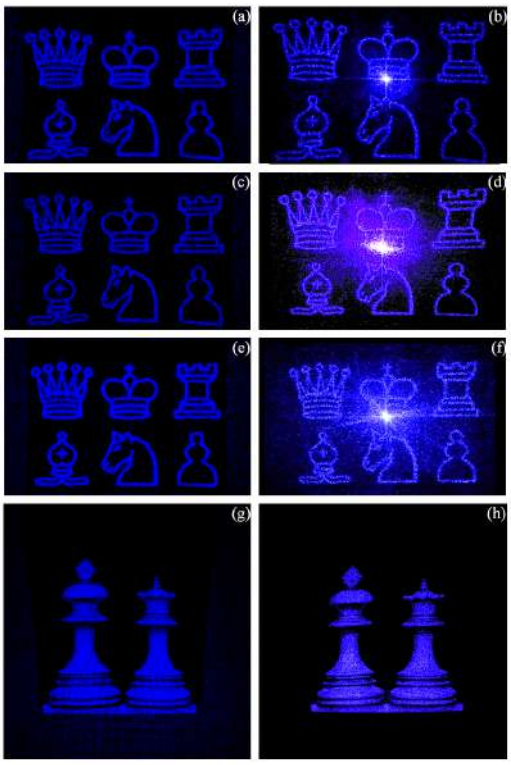


Figure III-17 – 1A [(a) and (b)], 1B [(c) and (d)], 1C [(e) and (f)] and 2 [(g) and (h)] numerical (first column) and experimental (second column) reconstructions operating at 444.9 nm.

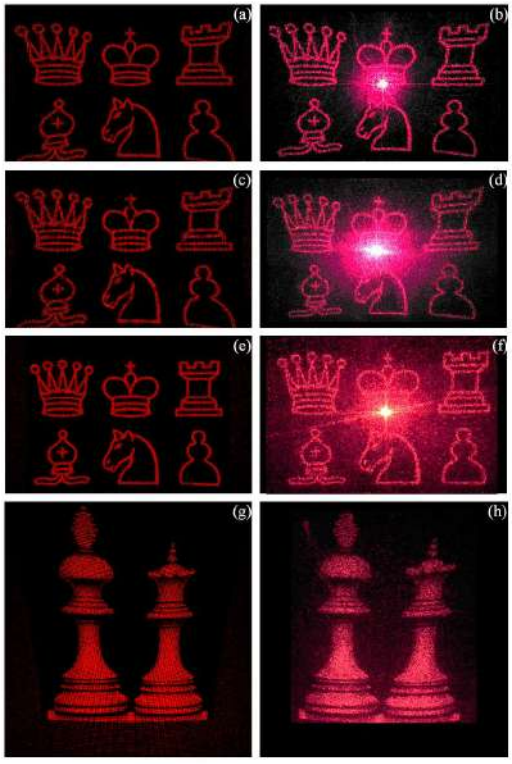


Figure III-18 – 1A [(a) and (b)], 1B [(c) and (d)], 1C [(e) and (f)] and 2 [(g) and (h)] numerical (first column) and experimental (second column) reconstructions operating at 632 nm.

The difference between 1A and 1B is not clear in Figures III-16, III-17 and III-18, but can be visualized by taking a closer look at one portion of the reconstruction. Figures III-19 (a)-(c)

show numerical reconstructions zooming in the queen for 1A, 1B and 1C, respectively, at 532 nm, these figures are to scale. 1A and 1B reconstructions have the same size, but the former is composed of bigger dots than the latter. This behaviour was previously seem [44] and is a direct manifestation of the fact that the hologram were calculated by the discrete Fourier transform (DFT) as is explained at Appendix *D.2.3 Diffraction of the sampled hologram*. In short, DFT assumes that the hologram is periodic but only a few periods are in fact fabricated. Therefore, a windowing effect happens that mathematically is equivalent to multiplying the periodic hologram by a *top rect* function. As the reconstruction of the hologram is given by its Fourier transform (neglecting the wide-angle aberration), the actual reconstruction of the windowed CGH will be given by the convolution between the DFT of the periodic CGH (that is, its reconstruction) convoluted with a sinc function, that is, the Fourier Transform of the top *rect* function. Due to the convolution, at each pixel of the discrete image there is a *sinc* function modulated by the pixel intensity. Thus, the more repetitions the holograms have, the larger is the top *rect* function and the smaller are the *sinc* functions waists on the reconstructions. Finally, the reconstruction of 1C, shown in Figure III-19 (c), has half the size in each direction than 1A and 1B, because it uses larger pixel dimensions, which in turn eliminates wide angle aberration. Furthermore, 1C is composed of a 2×2 CGH array as 1A so the dots composing both reconstructions have the same relative dimension in comparison with the queen size. In absolute terms, the dots at the reconstruction of 1C have half the size of the ones at the reconstruction of 1A along each direction, because the reconstruction of the former is also smaller than the reconstruction of the latter.

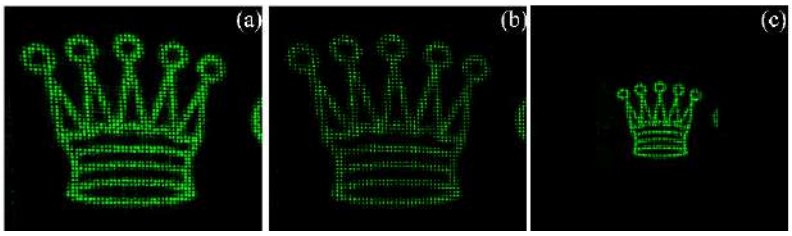


Figure III-19 - Numerical reconstructions of (a) 1A (b) 1B and (c) 1C metasurfaces zoomed in the queen. Operating at 532 nm. These figures are to scale.

III.1.7. Sapphire Substrate

a) Fabricated metasurfaces - Sapphire

a.1. *Fabrication method – Sapphire*

The metasurfaces were fabricated on a commercially available 230 nm thick c-Si (100) epitaxially grown on a sapphire substrate (from UniversityWafer, Inc.), shown in Figure III-20 (a). The structure was patterned on a HSQ electron beam resist using electron beam generator Vistec EBPG 5000 +at 100 KeV (Figures III-20 (b)-(c)). After developing the resist (Figure III-20 (c)), the pattern was transferred from the resist into the silicon layer using inductively coupled plasma etching (PlasmaPro System 100ICP180) (Figure III-20 (d)). The remaining HSQ is removed with Hydrofluoric acid. The scanning electron micrographs (SEM) of the metasurfaces are shown in the next section.

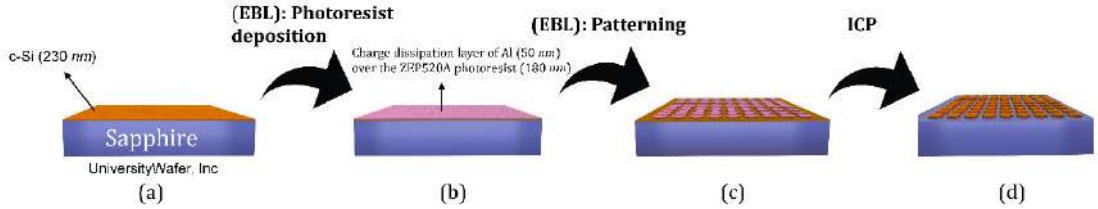


Figure III-20 - Fabrication steps of the metasurfaces fabricated over quartz substrate. (a) UniversityWafer silicon on sapphire waver. (b)-(c) shows the steps involved on the fabrication. (EBL) Electron beam lithography. (ICP) Inductively coupled plasma.

a.2. *Metasurfaces 1B, 1C and 2*

The metasurfaces fabricated on the sapphire substrate were made in the same substrate using the same substrate in three different samples that will be called S1, S2 and S3. Figures III-21 – III-24 show pictures and SEM micrographs of the fabricated samples.

From the micrograph over a larger area in Figure III-21 it is already possible to see that some of the metasurfaces present fabrication defects. Note that, ideally, they should be almost transparent as metasurfaces 1B of samples S1 and S2 and all metasurfaces 2. On the other hand, 1B from sample S1 and all 1C metasurfaces present dark regions suggesting that there are imperfections in the structures. Nevertheless, taking a closer look in the good regions of the samples, shown in the micrographs in Figures III-22 – III-24**,** it is clear that there are neither coalescence nor fallen nanoposts as we saw for the quartz design. Furthermore, as we saw for the quartz design, the diameter of the fabricated also differed from the designed ones. In summary, the pictures of these micrographs show that the fabricated metasurfaces over sapphire substrate present better quality than the previous one. Consequently, the measured diffraction efficiencies reach 40% in 532 nm (10% higher than the design on quartz) and the image reconstructions better quality.

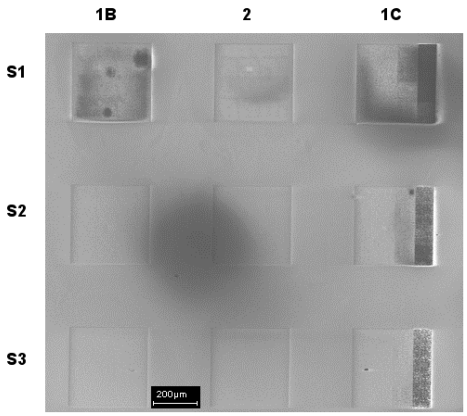


Figure III-21 – SEM micrograph of the sample with the metasurfaces fabricated over sapphire substrate. The first, second and third rows show the metasufaces S1, S2 and S3, respectively. The first, second and third columns show metasurfaces 1B, 2 and 1C, respectively.

a.2.1. *Sample S1*

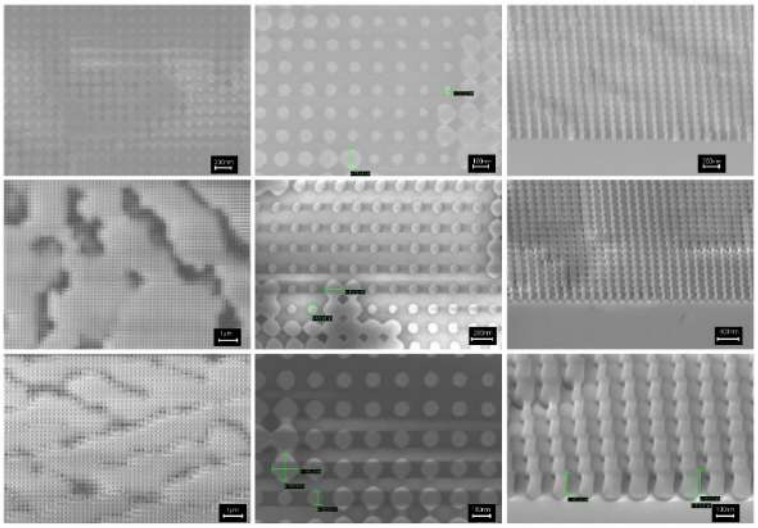


Figure III-22 – SEM micrographs of metasurfaces 1B (first row), 1C (second row) and 2 (third row) of sample S1.

a.2.2. *Sample S2*

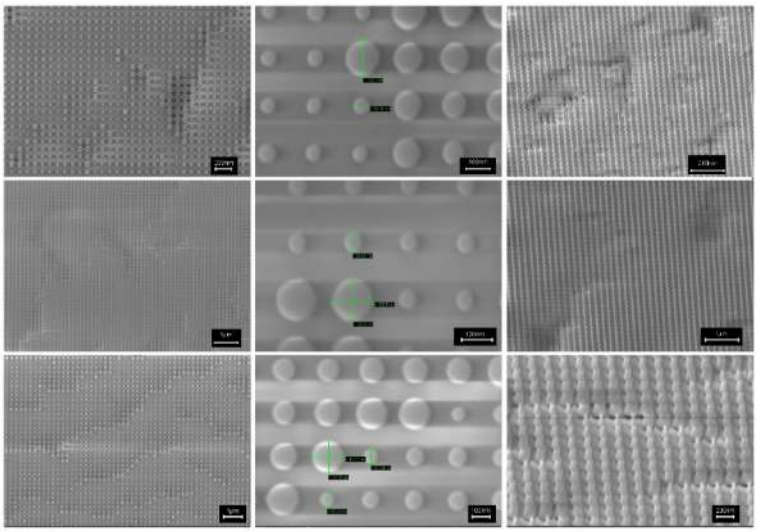


Figure III-23 – SEM micrographs of metasurfaces 1B (first row), 1C (second row) and 2 (third row) of sample S2.

a.2.3. ***Sample S3***

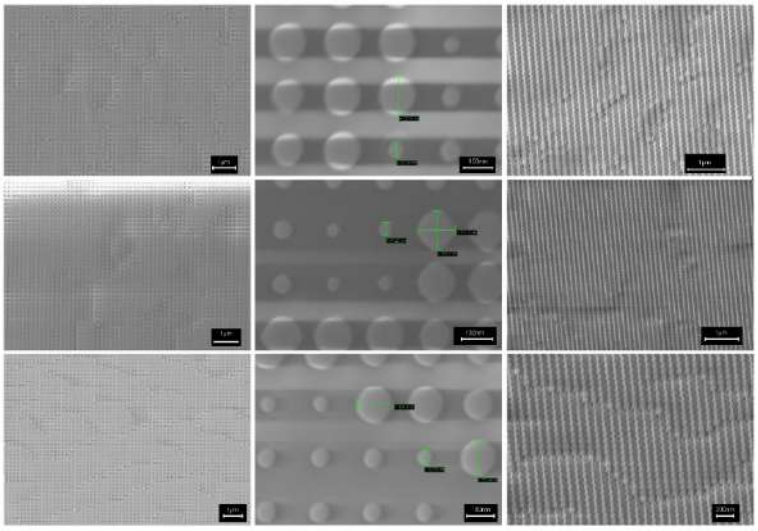


Figure III-24 – SEM micrographs of metasurfaces 1B (first row), 1C (second row) and 2 (third row) of sample S3.

b) Efficiencies

The measured transmission and diffraction efficiencies of Samples S1, S2 and S3 are listed in Tables III-7, III-8 and III-9 for the wavelengths of 532 nm, of 633 nm, and of 444.9 nm, respectively. The same fabrication issues discussed for the metasurfaces on quartz substrate apply here (see quartz section *III.1.6.b) Efficiencies*). But the efficiencies of metasurfaces on sapphire are higher than the designs on quartz. For the target wavelength of 532 nm, the higher transmission (diffraction) measurement is 68.57% (41.14%) which is 17% (10%) larger than the best results on quartz. Furthermore, the highest transmission (diffraction) efficiency reduced only 4% (10%) when compared to the numerical results with the c-Si approximation listed in Table III-2. These better results are consequence of the better quality in the samples, as can be seen in the SEMs micrographs of Figures III-23 and IIIIII.1-24 (compare with the SEMs of the design on the quartz substrate, shown in Figures III-9, III-11, III-13 and III-15). The same discussion also applies to the results of the other wavelengths, which also surpass the corresponding ones obtained for the design on the quartz substrate.

Table III-7 – Measured transmission and diffraction efficiencies of the metasurfaces over sapphire substrate operating at 532 nm.

Sample	Hologram	Transmission efficiency	Diffraction efficiency
S1	1B	68,57%	34,57%
	1C	65,71%	31,43%
	2	62,86%	29,71%
S2	1B	64,86%	39,66%
	1C	60,00%	30,94%
	2	62,86%	33,14%
S3	1B	61,14%	41,14%
	1C	45,43%	30,23%
	2	58,29%	38,20%

Table III-8 – Measured transmission and diffraction efficiencies of the metasurfaces over sapphire substrate operating at 633 *nm*.

Sample	Hologram	Transmission efficiency	Diffraction efficiency
S1	1B	76,36%	27,27%
	1C	83,78%	27,97%
	2	74,55%	18,36%
S2	1B	72,73%	29,27%
	1C	68,78%	19,59%
	2	77,03%	25,68%
S3	1B	65,45%	38,18%
	1C	56,76%	30,54%
	2	63,38%	33,11%

Table III-9 – Measured transmission and diffraction efficiencies of the metasurfaces over sapphire substrate operating at 444 *nm*.

Sample	Hologram	Transmission efficiency	Diffraction efficiency
S1	1B	43,33%	15,00%
	1C	66,67%	11,67%
	2	40,00%	15,00%
S2	1B	50,00%	23,33%
	1C	41,67%	13,33%
	2	46,67%	18,33%
S3	1B	43,33%	16,67%
	1C	35,00%	16,67%
	2	35,00%	10,00%

c) Reconstructions - Sapphire

Only the reconstructions of S2 are shown because the other ones, although very similar, are noisier and with stronger zero orders. In other words, only the best reconstructions are shown. Figures III-25, III-26 and III-27 show the numerical ideal and experimental reconstructions of those metasurfaces operating at 532 nm, 444.9 nm, and 633 nm, respectively (these figures are not to scale). These reconstructions are better in terms of noise and zero order power than the ones made for design on quartz (compare them with Figures III-16, III-17 and III-18**)**. These results tally with the efficiency measurements. Finally, the issues related to wide-angle aberration and differences between each hologram discussed for the quartz substrate (see section *III.1.6.c) Reconstructions - Quartz* of chapter *II* for details) also apply to the sapphire substrate.

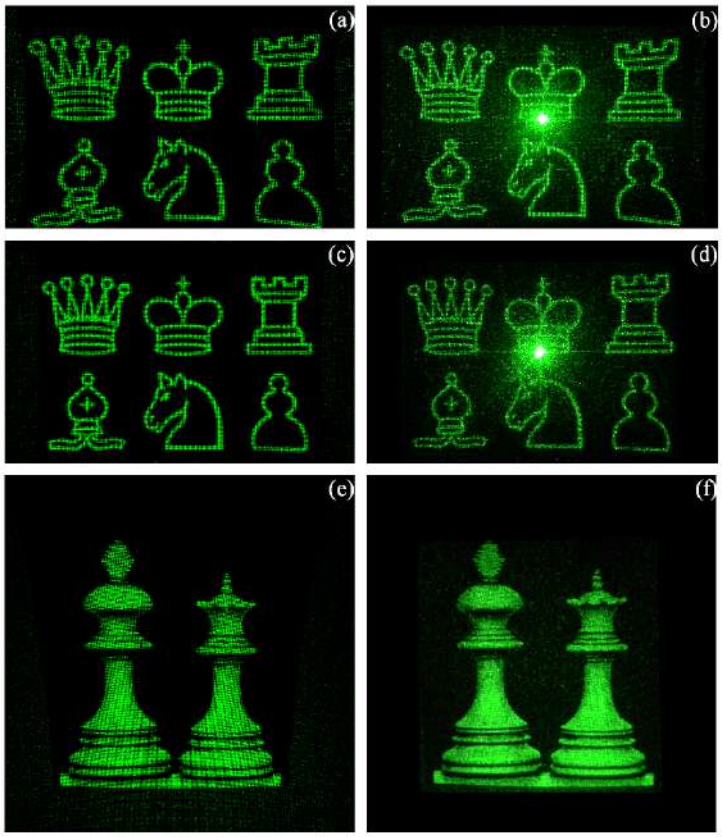


Figure III-25 – 1B [(a) and (b)], 1C [(c) and (d)] and 2 [(e) and (f)] numerical (first column) and experimental (second column) reconstructions operating at 532 nm.

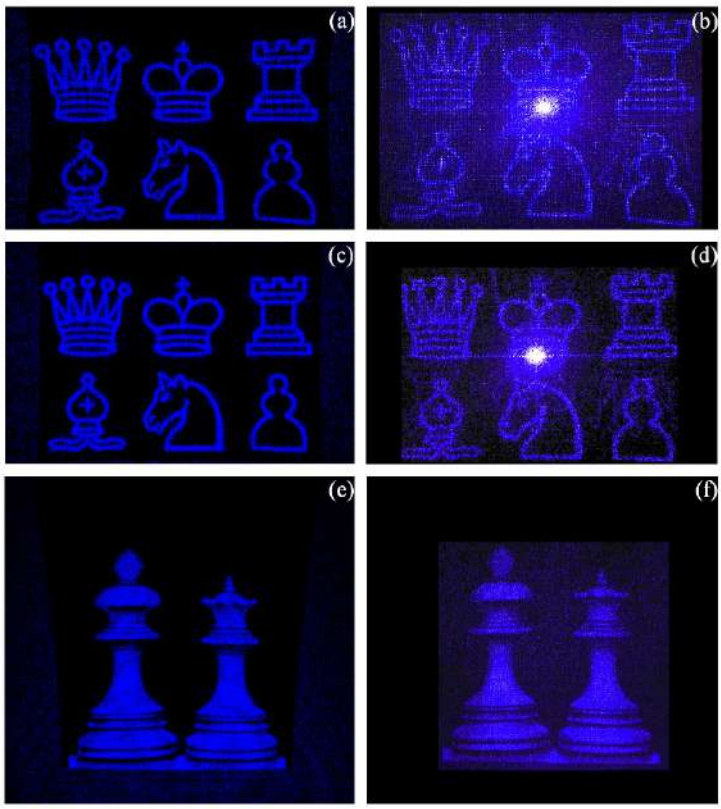


Figure III-26 – 1B [(a) and (b)], 1C [(c) and (d)] and 2 [(e) and (f)] numerical (first column) and experimental (second column) reconstructions operating at 444.9 nm.

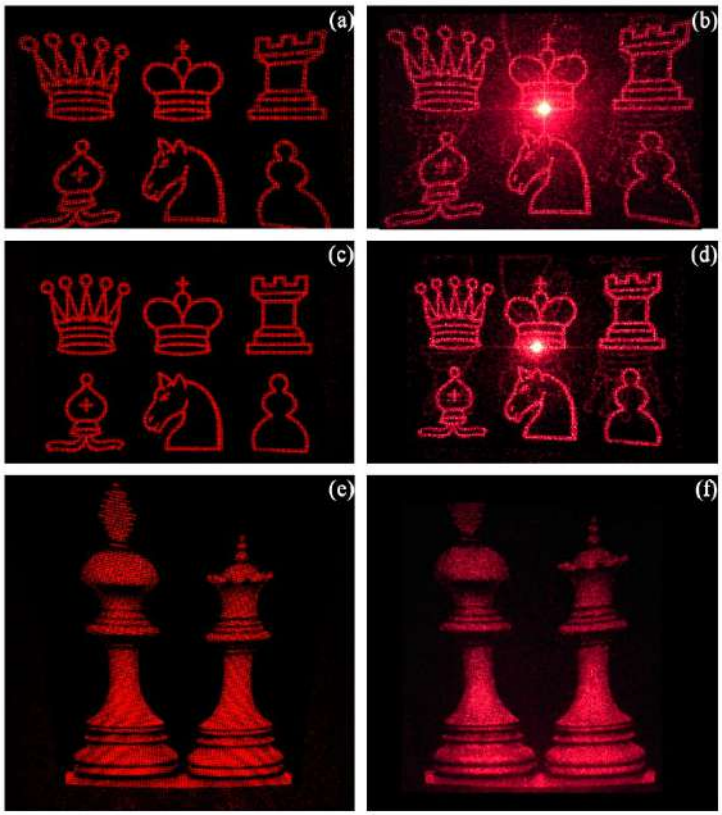


Figure III-27 – 1B [(a) and (b)], 1C [(c) and (d)] and 2 [(e) and (f)] numerical (first column) and experimental (second column) reconstructions operating at 633 nm.

III.2. Broadband c-Si metasurfaces with polarization control at visible wavelengths: Applications to 3D stereoscopic holography

III.2.1. Introduction

In the past few years, visual arts and entertainment related industries, as well as many research groups around the world, have invested a great deal of effort and resources directed at improving users' experience by means of state-of-the-art visualization techniques, such as telemedicine, holographic projections, three-dimensional (3D) cinema screens, and 3D mobile displays, to mention a few. These technologies have benefited from rapid advances in nanofabrication techniques that have had a direct impact on the pixel resolution and on the miniaturization of optical systems. One remarkable example of optical miniaturization is the concept of planar photonic components, or flat optics, enabled through metasurfaces, i.e. a dense arrangement of subwavelength resonators designed to modify the optical response of the interface [94]. Metasurfaces, the two-dimensional (2D) equivalent of (volumetric) metamaterials, are periodic or semi-periodic materials that allow light properties (phase, amplitude and/or polarization) to be tailored by geometrical adjustments on the size and/or orientation of the subwavelength unit-cell geometry.

Interestingly enough, a decades-old 3D visualization technique known as stereoscopic viewing has only recently been theoretically explored with metasurfaces [95]. Briefly, a stereoscopic image (stereogram) is composed of a pair of orthogonally polarized images taken from the same scene but recorded in slightly shifted positions to replicate the natural parallax of the human eye. For the stereoscopic effect to occur, each of these two images should be directed to one of our eyes separately with the help of cross-polarized glasses. Most of the proposed metasurfaces to this date operate either with single [42, 49, 69] or multiple discrete wavelengths [45, 77, 96, 97] in the visible and infrared ranges. Nevertheless, many applications such as metalenses, holography and polarization optics could benefit from or even require broadband operation. This issue has already been addressed both in reflection [98] and transmission mode [62, 64, 99-101] metasurfaces. However, high efficiency at the blue region of the visible spectrum was achieved only with TiO_2 based structures [62].

In the present work, we proposed and fabricated the first broadband holographic stereogram made with a birefringent crystalline silicon (c-Si) metasurface at visible wavelengths. The metasurface design is based on c-Si nanoposts with elliptical cross-section and different effective indices, and excited by linearly polarized light aligned to each of its semi-axes[52]. The holograms (one for each polarization) are combined on the same metasurface and encoded with four phase levels each, calculated via modified Gerchberg-Saxton (G-S) phase-retrieval algorithm to maximize the signal-to-noise-ratio (SNR) [73, 75]. Furthermore, a coordinate transformation correction is performed on the target images to avoid wide angle aberration of the reconstructions [91]. Two additional metasurface holograms, whose reconstructions yield two completely distinct images, are also fabricated aimed at addressing possible cross-talk between the two polarizations. We show that the relative phase shift of the nanostructures has a high tolerance over a broad bandwidth and a high diffraction efficiency even at the blue region of the spectrum. Diffraction efficiencies >20% within a 110 nm bandwidth are numerically obtained with a SNR>15 dB. We also show that the operating band of the structure can be blueshifted by simply reducing the nanoposts diameter with minimum impact on the relative phase. Even in this case the calculated diffraction efficiency remains >20% within an 80 nm bandwidth with a SNR>10 dB.

The experimental stereographic reconstruction is obtained on a fine-sanded aluminium surface. This choice not only preserves the polarization of the scattered light but also minimizes the mirror effect of the aluminium surface, which increases the 3D perception. The measurements are carried out at three different wavelengths, namely 444.9 nm, 532 nm (the design wavelength) and 635 *nm*, to assess the metasurface bandwidth performance. The measured transmission and diffraction efficiency maxima at 532 nm are 70% and 15% for both polarizations, respectively. The reconstructed images at 444.9 nm are as good as at 532 nm in terms of noise, polarization cross-talk, and depth perception. The measured transmission and diffraction efficiencies in this case are 69.5% and 18.5%, respectively. This good performance at lower wavelengths can be attributed to the reduced semi-axis size of the fabricated nanoposts that cause the metasurface operating point to blueshift. The reconstructed images at 635 nm, in contrast, are noisier with stronger polarization cross-talk when compared to those at 532 nm, making it difficult to observe the stereoscopic effect. Consequently, the diffraction efficiencies are small (<10%), despite the high transmission (>80%) due to the c-Si low absorption at this wavelength.

III.2.2. Metasurface design

The unit cell geometry of the proposed birefringent metasurface structure is shown in Figure III-28 (a). It consists of c-Si elliptical nanoposts 230 nm high on top of a sapphire substrate, with the design wavelength of 532 nm. This structure can be regarded as a truncated waveguide with elliptical cross section exhibiting polarized modes with different effective wavelengths along each

semi-axis of the ellipsis [52]. In other words, the transmitted light experiences different phase changes when polarized along each semi-axis. Therefore, independent phase control can be achieved for each polarization by tuning the semi-axes sizes. The design is carried out by calculating the transmission phase and efficiency of an infinite array of identical nanoposts under normal plane-wave incidence from air. We use a square unit cell 190 nm in size and sweep both semi-axes (D_x and D_y) to determine the optimum phase levels and transmission efficiencies of the array using the rigorous coupled-wave analysis (RCWA) [88]. The refractive indices of c-Si and sapphire adopted in the simulations are, respectively, $n_{c\text{-}si}$ =4.3+j0.072[90] and n_S=1.77[102].

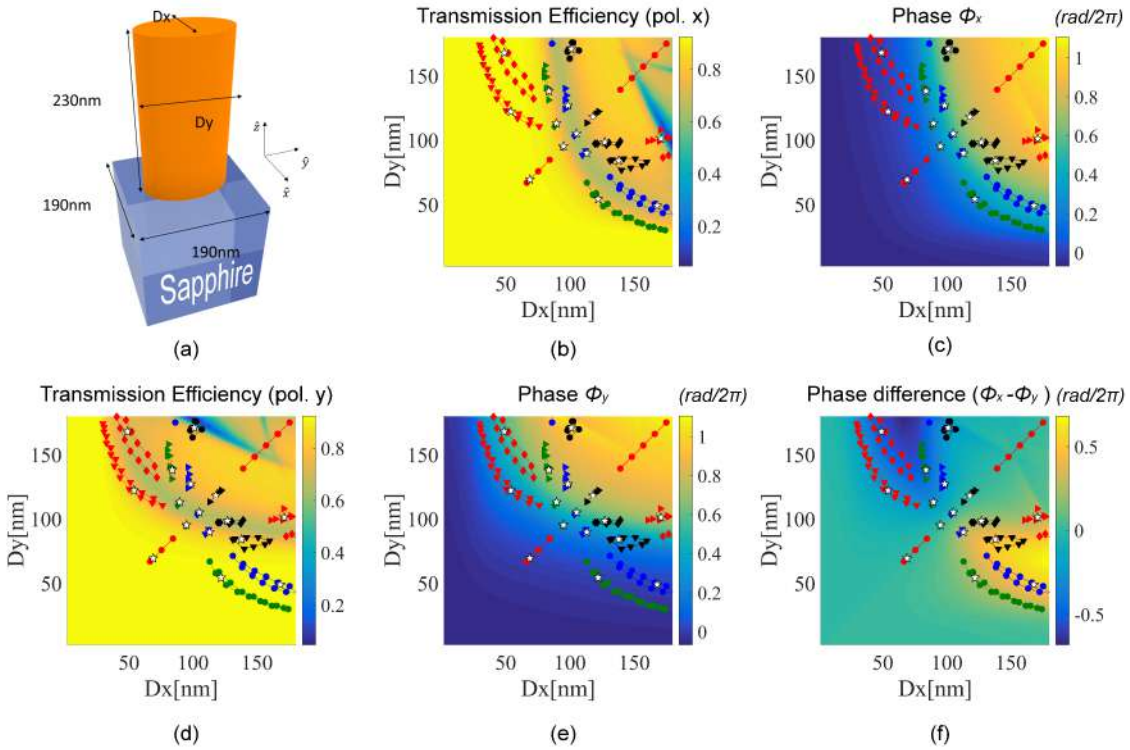


Figure III-28 -(a) Schematic of the unit cell for high birefringent contrast grating (not to scale). Transmission efficiency (b) and relative phase maps (c) of the c-Si nanopost array as function of their semi-axes for light polarized along *x*. (d) and (e) represent analogous results for light polarized along *y*. The transmitted phase difference (Φ_x- Φ_y) is shown in (f). The operating wavelength is 532 nm in all cases. The colored symbols represent the map boundaries for achieving four phase levels. Their correspondences are listed in Table III-10. The stars mark the chosen structures inside each region.

The transmission efficiency and phase maps (in units of rad/2π) as function of D_x and D_y are shown, respectively, in Figures III-28 (b) and (c) for polarization along *x*, and in Figures III-28 (d) and (e) for polarization along *y*. Full phase control [0-2π] is obtained for each polarization with high transmission efficiency. Moreover, these maps present as major features both birefringence and high transmission efficiency. The next step is to check whether independent full phase control [0-2π] for both polarizations can be achieved with these maps. In other words, if *n* phase levels are needed for each polarization, then a total of n^2 different elements are required to satisfy all possible phase combinations ($\Phi_{x,i}$, $\Phi_{y,j}$), $i, j \in [1,n]$. Therefore, the phase difference map (Φ_x- Φ_y) needs to cover the range [-π, π], as shown in Figure III-28 (f).

The final step is to check which structures satisfy these phase level requirements. Assuming a 10% tolerance on the phase value, independent phase control is achieved with $n = 4$ phase levels, resulting in sixteen different structures. The symbols in Figures III-28 (b)-(f) (listed in Table III-10) delimit the boundaries where a given structure geometry satisfies the required phase level. The chosen structures are marked with white stars in Figures III-28 (b)-(f). The pertinent parameters of the 16 structures that satisfy the stablished design criteria, namely, the length of the semi axes, transmission phase and efficiency values, are listed in Table III-11 for each polarization.

Table III-10 - Symbol definition relative to Figures III-28 (b) – (f). Each color (symbol) corresponds to a relative phase level obtained for electric field polarized along *x or y* direction.

Target Phase [rad]	Color (x)	Symbol (y)
-0.5π or 1.5π	black	Down triangles
0 or 2π	red	Left triangles
0.5 π	green	Diamonds
-π or π	blue	circles

Table III-11 - Target and obtained phase and transmission values of each pixel with four phase level control. The operating wavelength is 532 nm. The chosen structures are marked with white stars in Figures III-28 (b)-(f).

Target x phase ϕ_x [1/2π]	Target y phase ϕ_y [1/2π]	D_x [nm]	D_y [nm]	Obtained x phase ϕ_x [1/2π]	Obtained y phase ϕ_y [1/2π]	$\lvert T_x \rvert$	$\lvert T_y \rvert$
0 or 1	0 or 1	69	69	0.01	0.01	0.92	0.92
0 or 1	0.25	54.1	122	0.01	0.25	0.92	0.79
0 or 1	0.5	49	168	0.01	0.50	0.92	0.7
0 or 1	0.75	171	101	0.99	0.73	0.63	0.75
0.25	0 or 1	122	54.1	0.25	0.01	0.79	0.92
0.25	0.25	95	95	0.26	0.26	0.78	0.78
0.25	0.5	89.6	113	0.26	0.51	0.79	0.65
0.25	0.75	83.8	138	0.26	0.74	0.81	0.77
0.5	0 or 1	168	49	0.50	0.01	0.7	0.95
0.5	0.25	113	89.6	0.51	0.26	0.65	0.79
0.5	0.5	105	105	0.48	0.50	0.65	0.65
0.5	0.75	98.3	127	0.48	0.75	0.66	0.77
0.75	0 or 1	101	171	0.73	0.99	0.75	0.63
0.75	0.25	138	83.8	0.74	0.26	0.77	0.81
0.75	0.5	127	98.3	0.75	0.48	0.77	0.66
0.75	0.75	118	118	0.76	0.76	0.76	0.76

Next, the metasurface bandwidth was assessed in terms of both relative phase and transmission efficiency. As shown in the previous section, c-Si metasurfaces do have broadband behaviour in the visible spectrum. However, here we address this issue by averaging the spectra of the metasurfaces that provide the same phase level ϕ_x at 532 nm according to Table III-11 (a similar result can be obtained for ϕ_y). The resulting averaged spectra are shown in Figure III-29 (a) for the following phase levels: 0 (red), 0.25 (blue), 0.5 (black) and 0.75 (green) rad/2π. The dots in Figure III-29 (a) indicate the optimum phase levels at 532 nm. The error bars indicate the

phase and transmission standard deviations of the average values. Note that the phase curves relative separation does not change significantly over a broad spectral region around 532 nm. Note further that the corresponding averaged transmission spectra, shown in Figure III-29 (b), are not only high but also broadband, therefore confirming the broadband nature of c-Si metasurfaces.

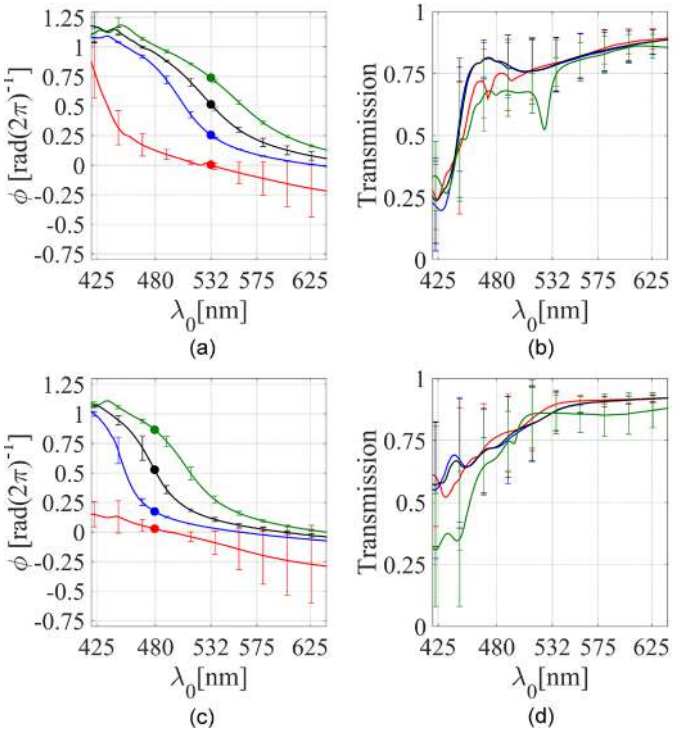


Figure III-29 - Averaged phase (a) [(c)] and transmission (b) [(d)] of the unperturbed [perturbed, nanopost semi-axes reduced by 20 nm] structures designed to give relative phase levels of 0 (red), 0.25 (blue), 0.5 (black) and 0.75 (green) rad/2π. The average is taken over the spectra of different metasurfaces that provide the same phase level in each polarization.

This analysis is now further extended assuming a 20 nm reduction of the semi-axes, which is compatible with our nanofabrication resolution. The phase and transmission spectra are shown in Figure III-29 (c) and (d), respectively. Note that the metasurface optimum operating point (dots in Figure III-29 (c)) blueshifts by approximately 52 nm (to ~480 nm). Moreover, this occurs without significantly changing the relative phase separation over a broad band (see Figure III-29 (c)) and with high transmission (>60%, as shown in Figure III-29 (d)). Finally, Figure III-29 (a)-(d) indicate that variations in transmission (wide error bars) at a given wavelength are less detrimental to the metasurface SNR and diffraction efficiency than variations in phase.

III.2.3. Hologram design

The performance of the birefringent metasurfaces is assessed by way of two different hologram designs (Designs 1 and 2). Each design encodes two target images as illustrated in Figure III-30. In Design 1, we use two different superposed target images (images A and B) with the purpose of assessing the cross-talk between each polarization and comparing their efficiencies. In Design 2 we use two similar images (C and D) which are slightly separated in space with

respect to each other. This is done to mimic the human eye parallax and thus guarantee stereoscopic reproduction. This effect is created with the open source 3D computer-graphic software Blender [103] (to see the depth effect in the Design 2 anaglyph hold a blue filter in front of the **right** eye and a red filter in front of the **left** eye). However, in the proposed image reconstruction setup these images (both at the same wavelength) must be orthogonally polarized to each other and separated at the observer's eyes with the help of a pair of cross-polarized filters. This arrangement successfully reproduces the depth perception characteristic of 3D projections.

The computer generated holograms (CGHs) of each target image A-D, all with 1024×1024 pixels and pixel dimension of 190 nm×190 nm are calculated via the iterative Fourier Transform algorithm (IFTA), described in appendix *D.3 Iterative Fourier Transform Algorithm (IFTA)*. The total size of each CGH is 194.56 μm×194.56 μm. The small pixel dimension compared to the operating wavelength (532 nm) causes the reconstructed images to suffer from wide angle aberration. Thus, a correction is required for each target image to avoid this effect prior to hologram calculation [91]. Finally, the target images are vertically displaced to avoid cross-talk with the unwanted zeroth-order beam. The resulting CGHs of A and B (similarly for C and D) are then encoded in a c-Si birefringent metasurface for **x** and **y** polarized light, respectively, according to Table III-11. Each encoded CGH is then assembled as 2×2 arrays D1 and D2 as the final metasurfaces with total size of 389.12 μm×389.12 μm.

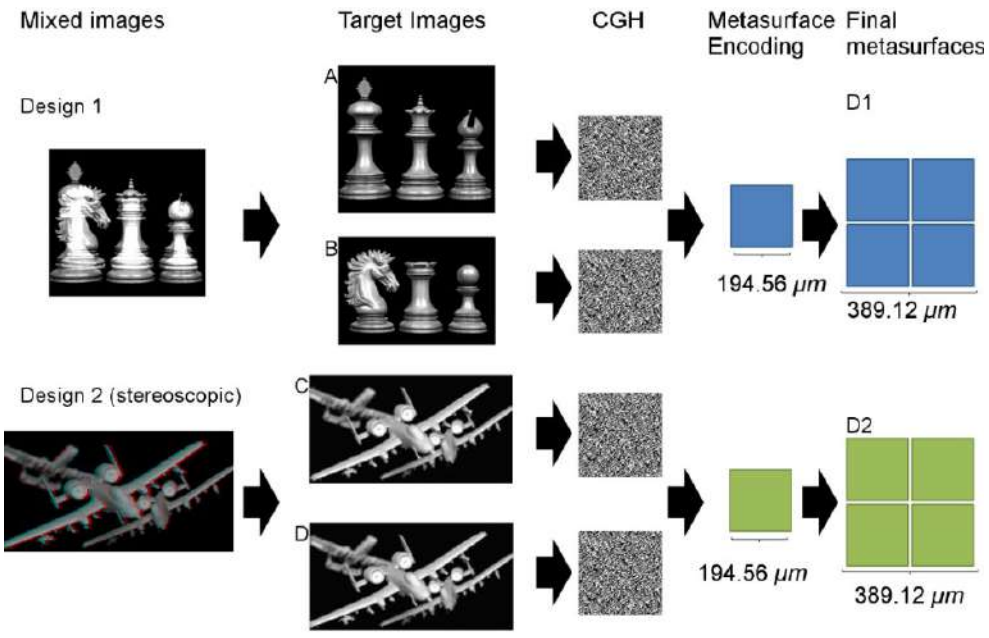


Figure III-30 - Concept of the two birefringent c-Si metasurfaces. Design 1 consists of two different target images (A encoded for x- and B for y-polarization) with the purpose of accessing the efficiency and polarization cross talk of both images. Design 2 consists of two similar parallax-separated images (C encoded for x- and D for y-polarization) for 3D stereoscopic reconstruction.

III.2.4. Numerical Characterization

A broadband reconstruction analysis of metasurface D1 was carried using the RS integration (explained in appendix *D.4 Rigorous Rayleigh Sommerfeld (RS) Diffraction*) by simulating its SNR, diffraction and transmission efficiencies spectra. As usual, the transmission efficiency is defined as the ratio between the transmitted power through the hologram to the transmitted power through the substrate, while the diffraction efficiency is defined as the ratio between the power at the image window to the power transmitted through the substrate. The calculated efficiencies and SNR are shown in Figures III-31 (a) and (b), respectively. Note in both plots that the operating point of the metasurface is blueshifted to 480 nm when the semi-axes are reduced by 20 nm. Furthermore, the diffraction efficiencies (dotted and continuous lines in Figure III-31 (a) from CGHs A and B, respectively) are >20% within a 110 nm [80 nm] region for the unperturbed [perturbed, semi-axes reduced by 20 nm] metasurfaces, as indicated by the blue [black] curves. In this region, the transmission efficiency remains >70% [>50%] and the SNR, shown in Figure III-31 (b), >15dB[>10dB], resulting in high quality reconstructions. These results confirm once again the broadband nature of c-Si metasurfaces as suggested by the phase and transmission spectra in Figures III-29 (a) and (b), respectively. More importantly, it demonstrates that by simply rescaling the nanoposts it is possible to tune the operating bandwidth of the proposed design.

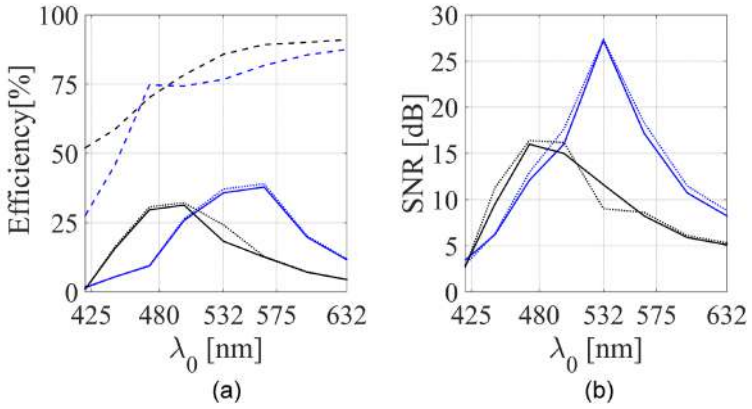


Figure III-31 - (a) Transmission (dashed lines) and diffraction (continuous and dotted for CGH A and B, respectively) efficiencies spectra for the design without perturbation (blue) and reduced by 20 nm (black). The transmission efficiencies spectra of A and B are overlapped. (b) SNR of the unperturbed (blue lines) and reduced (black lines) designs. The continuous and dotted lines show the result of CGHs A and B, respectively.

III.2.5. Experimental Characterization

a) Fabricated Metasurfaces

a.1. *Fabrication methods*

The fabrication procedure is the same used to fabricate the previous metasurfaces over sapphire substrate. See section *III.1.7.a.1 Fabrication method – Sapphire* for the detailed steps involved.

a.2. *Metasurfaces D1 and D2*

The metasurfaces D1 and D2 were fabricated in the same substrate but with different doses of exposure. Figures III-32 – III-34 show pictures and SEM micrographs of the fabricated samples.

The SEM micrographs show that all the fabricated metasurfaces present good quality. Unfortunately, we noted that the fabricated posts have smaller posts than the designed ones. We estimated a reduction of 20 nm in average. Therefore, this caused the operation of the structure to blueshift, as our experimental and numerical characterizations show. Nevertheless, the reconstructions and efficiency measurements were good operating at both 532 nm and 444.9 nm, which showed us that this design present a broadband behaviour.

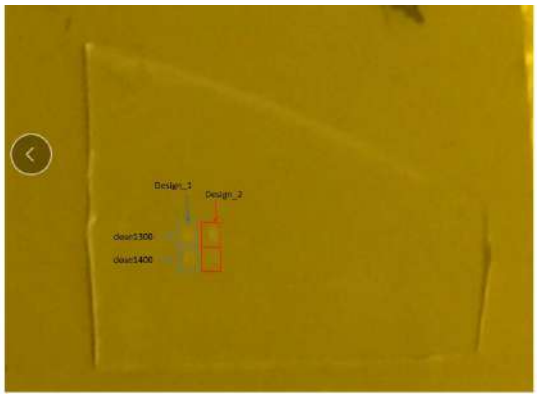


Figure III-32 – Picture of the sample with metasurfaces D1 (cyan squares) and D2 (red squares) fabricated with different doses (1300 and 1400) of ebeam exposure.

a.2.1. ***Dose1300***

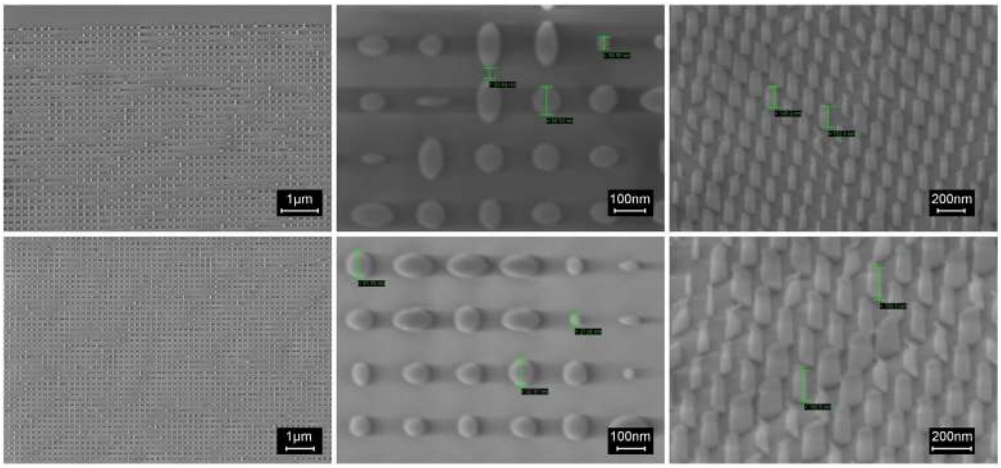


Figure III-33 - SEM micrograph of metasurfaces D1 (first row) and D2 (second row) at different scales and positions of Dose1300 sample.

a.2.2. ***Dose1400***

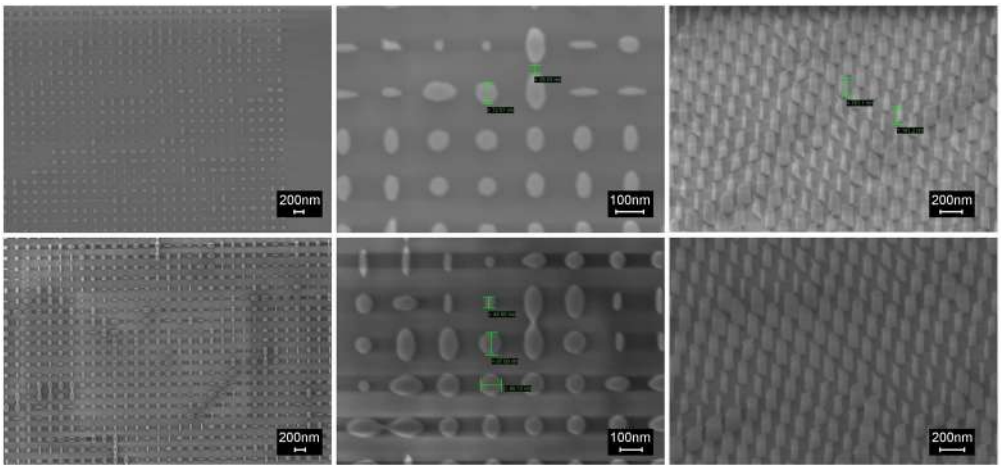


Figure III-34 - SEM micrograph of metasurfaces D1 (first row) and D2 (second row) at different scales and positions of Dose1400 sample.

b) Efficiencies

The experimental characterization is first made at the design wavelength of 532 nm. The measured transmission and diffraction efficiencies are listed in Table III-12 for both metasurfaces and polarizations. The transmission efficiencies of all metasurfaces are very high because of c-Si low absorption at the operating wavelength when compared to amorphous or polycrystalline silicon [77]. However, the diffraction efficiency is modest for all samples because the fabricated metasurfaces suffer from semi-axis size deviations of their elliptical nanoposts. As a result, the phase shift imparted by each nanopost on the transmitted light deviates from the desired values listed in Table III-11, as shown in Figure III-28. In this sense, more light is cast as noise in addition to being lost to the zeroth-order beam. Both these effects contribute to reduce the diffraction efficiency and to increase the polarization cross-talk.

Table III-12 - Measured and simulated efficiencies at the holograms reconstructions plane and different wavelengths.

		Wavelength	444.9 nm		532 nm		635 nm	
Sample	Metasurface	Hologram (polarization)	Transmission Efficiency	Diffraction Efficiency	Transmission Efficiency	Diffraction Efficiency	Transmission Efficiency	Diffraction Efficiency
Dose1300	D1	A (x)	69.36%	18.50%	67.95%	10.26%	85.88%	9.41%
		B (y)	69.52%	14.29%	73.08%	12.82%	77.53%	5.62%
	D2	C (x)	69.65%	10.17%	60.26%	8.97%	80.00%	9.41%
		D (y)	69.33%	14.00%	57.69%	14.10%	77.75%	4.83%
Dose1400	D1	A (x)	72.83%	11.56%	71.79%	2.56%	85.6%	3.53%
		B (y)	79.05%	10.48%	76.92%	5.13%	82.02%	6.74%
	D2	C (x)	63.58%	14.34%	65.38%	12.82%	82.53%	7.41%
		D (y)	70.52%	13.33%	69.23%	16.67%	82.22%	5.62%

Next, the metasurfaces are investigated at 444.9 nm and 635 nm to assess their bandwidth performance. Table III-12 lists the resulting diffraction and transmission efficiencies for both metasurfaces and polarizations. Surprisingly, the efficiencies at 444.9 nm are as high as those at 532 nm, even though c-Si is more absorptive at this wavelength. This result can be attributed to variations in the semi-axes of the nanoposts during fabrication, that caused the metasurface operating point to blueshift, as shown in Figures III-31 (a) and (b).

c) Reconstructions

Only the reconstructions of dose1300 sample are shown, as they resulted in the better images. The experimental reconstructions of metasurface D1 are shown in Figures III-35 (a) and (d) for x- and y-polarized light, respectively. Both present good reconstruction quality despite the noise caused by variations on the size of the nanoposts. Part of this noise comes from the hologram encoded on the other polarization. Note that the reconstructions are polarized along one direction, x or y, meaning that these spurious reconstructions are essentially due to polarization cross-talk. To provide a phenomenological explanation for the mechanism behind this effect, we carry out numerical reconstructions, by RS integration, of the holograms. Initially, it is assumed that each CGH pixel has a homogeneous field distribution whose phase and amplitude accounts for the nanopost response according to Table III-11. Since CGHs are close to ideal, the reconstructions for x and y polarizations shown in Figures III-35 (b) and (e), respectively, are very good and without the polarization cross-talk image. Next, we perturb the semi-axes of the nanoposts by 20 nm (within the error from the SEM measurements), leading to new phase and amplitude values for each pixel read from the maps in Figure III-28. As expected, this perturbation diverts more power to both the zeroth-order beam and the Hermitian copy of the reconstruction for x and y polarizations, respectively, which increases the noise and reduces the diffraction efficiency as a result. Note also that the shadow of the orthogonal polarization reconstruction is visible around the desired image in this perturbed numerical reconstruction. Therefore, these numerical results,

mimicking the fabrication error and its impact on both the reconstruction noise and diffraction efficiency, fully corroborate the experimental observations.

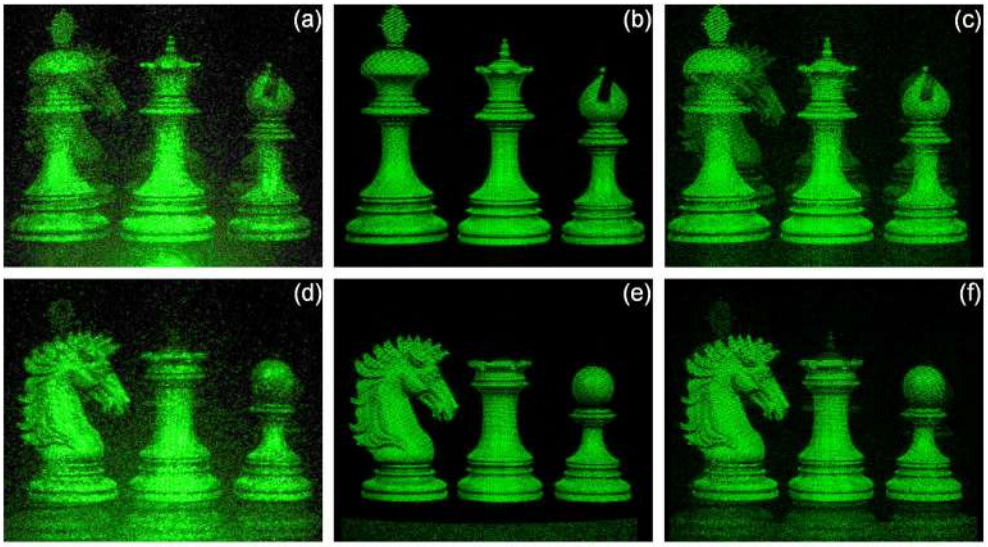


Figure III-35 - Experimental reconstructions of metasurfaces D1 for (a) x- and (d) y- polarized light. (b) and (e) [(c) and (f)] show the numerical reconstructions of the unperturbed [perturbed] metasurface. The reconstructions are taken 20 cm away from the metasurface. These images are not stereoscopic and are used to assess polarization cross-talk. The operating wavelength is 532 nm.

The experimental stereoscopic reconstructions of metasurface D2 are shown in Figures III-36 (a) and (b) for x- and y-polarized light, respectively. The stereoscopic reconstruction requires the hologram to be illuminated with x+y polarized light. The image is then reconstructed on a fine-sanded aluminium plate surface, used as reconstruction plane, to preserve the polarization of the scattered light. The stereoscopic view, however, requires the observer to wear a pair of cross-polarized filters in front of his/her eyes as used in our experiment (or a pair of cross-polarized glasses). Note that the cross-polarized spurious reconstructions are also present, but they are subtler to observe due to the similarity between images.

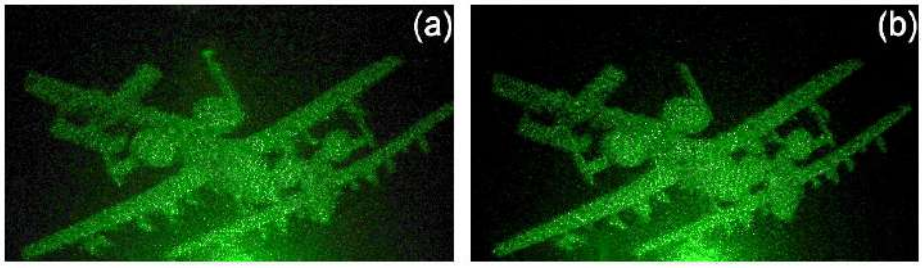


Figure III-36 - Experimental reconstructions of each image from the holographic stereogram for (a) x- and (b) y-polarized light. The reconstructions are taken 20 *cm* away from the metasurface. The operating wavelength is 532 nm.

The image reconstructions at 444.9 nm are shown in Figures III-37 (a), (b), (e) and (f), where (a) and (e) are not stereoscopic and used only to assess polarization cross-talk while (b) and (f) are obtained from the holographic stereogram. These images exhibit not only low noise, but also virtually inexistent cross-talk, indicating that phase control is more efficient at this wavelength resulting in improved depth perception. In contrast, the stereoscopic effect is not observed at 635

nm despite the metasurface high transmission efficiencies (>80%) due to c-Si low absorption at this wavelength. Moreover, the low diffraction efficiencies (<10%) are due to the poorer phase control at this wavelength, leading to higher noise and stronger polarization cross-talk compared to lower wavelengths, as seen in Figures III-37 (c), (d), (g) and (h). Similarly, (c) and (g) are not stereoscopic and used only to assess polarization cross-talk while (d) and (h) are obtained from the holographic stereogram. It is important to point out that the reconstructions at 444.9 nm and 632 nm are distorted because the wide-angle correction is wavelength dependent and is carried only for 532 nm.

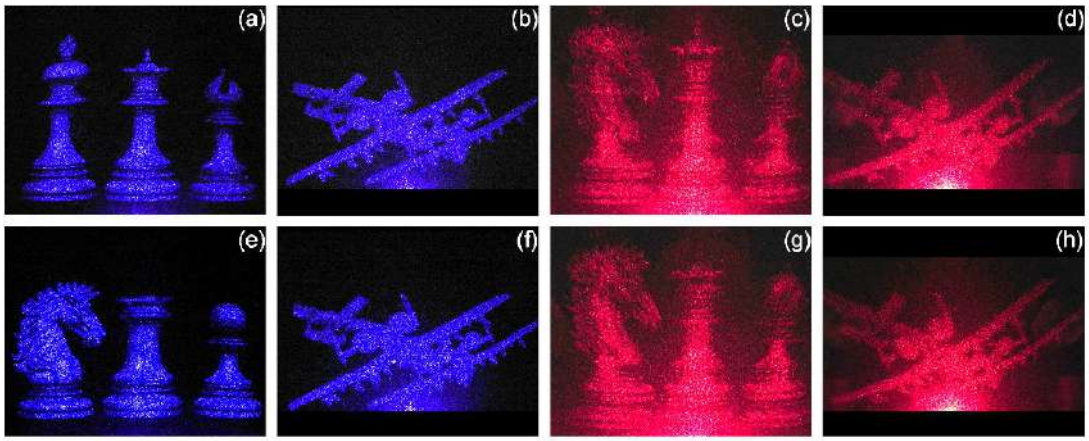


Figure III-37 - Experimental reconstructions for x- (first row) and y- (second row) polarized light at 444.9 nm (first two columns) and 635 nm (last two columns). The reconstructions are taken 20 cm away from the metasurface. (a), (e), (c) and (g) are not stereoscopic and are used to assess polarization cross-talk. (b), (f), (d) and (h) are obtained from the holographic stereogram.

III.3. Conclusions and future works

The design proposed in section *III Highly efficient holograms based on c-Si metasurfaces in the visible range* is the first c-Si metasurface hologram designed to operate at 532 *nm* in transmission-mode with maximum transmission efficiency of ~65% and diffraction efficiency of ~40%. We have also investigated the performance of the metasurfaces, both numerically and experimentally at different wavelengths to assess their operation bandwidth. Surprisingly, the reconstructed images at these wavelengths show good quality despite their reduced theoretical transmission efficiency (~19% at 444.9 nm, and ~51% at 635 nm). We used the numerically reconstructed images, obtained with the RS integration, as benchmark to assess the performance of the fabricated metasurfaces in terms of transmission and diffraction efficiencies. The numerical calculations showed that the transmission(diffraction) efficiency can reach up to 53.1%(16.2%), 72%(52.2%) and 85% (39.8%) at 444.9 nm, 532 nm and 635 nm, respectively. There is plenty of room for improvements in future designs by optimizing of the metastructures at each wavelength. More importantly, these results show that c-Si-based hologram metasurfaces are tolerant to fabrication and wavelength variations, making them quite attractive not only for diffractive optics applications in general, but particularly for full-color holograms.

After concluding the first metasurface design with c-Si, we decided to explore another one more sophisticated that features independent phase control for a pair of orthogonal linear polarization of light. Thus, we have presented the first broadband stereoscopic hologram fabricated with a birefringent c-Si/Saphire metasurface designed to operate at 532 nm. The birefringent metasurface was designed with c-Si nanoposts with elliptical cross-section and different effective indices, excited with linearly polarized light aligned to each of its semi-axis. The stereoscopic holograms (one for each polarization) were combined on the same metasurface and encoded with four phase levels each, calculated via the modified Gerchberg-Saxton (G-S) phase-retrieval algorithm to maximize the signal-to-noise-ratio (SNR). A coordinate transformation correction was carried out on the target images to avoid wide-angle aberration. The reconstruction plane used for the stereogram consisted of a fine-sanded aluminum surface to preserve the polarization of the scattered light and to minimize the mirror effect of the aluminum surface. We also fabricated two additional metasurface holograms using two completely uncorrelated images to address polarization cross-talk issues. The broadband nature of the metasurfaces was assessed by means of phase and transmission spectra calculation. We noticed that the wavelength band can be easily tuned simply by reducing the diameters of the nanoposts without significantly affecting their phase response. We also performed numerical reconstructions to assess the metasurface wavelength response in the visible range. The results showed diffraction efficiencies >20% within a 110 nm bandwidth with SNR>15 dB. Moreover, when the semi-axes

of the nanoposts were reduced by 20 nm (to account for fabrication errors), the diffraction efficiency remained >20% within a 80 nm bandwidth with SNR>10 dB. The measured transmission efficiencies were very high (about 70%) because of the c-Si low absorption at the operating wavelength when compared to amorphous or polycrystalline silicon. The diffraction efficiency, in contrast, was modest because the semi-axes size of the fabricated nanoposts were smaller than originally designed. This caused more light to be diverted to noise and to the zeroth-order, in addition to increasing the polarization cross-talk. We have also investigated the metasurfaces' performance at 444.9 nm and 635 nm to assess their bandwidth. The stereoscopic effect was surprisingly good at 444.9 nm with transmission and diffraction efficiencies as high as 70% and 18%, respectively, with good depth perception. The same was not true at 635 nm due to low diffraction efficiencies (<10%, due to poor phase control), despite high transmission efficiencies (>80%, due to c-Si low absorption at this wavelength). This occurred because the semi-axes size variation blueshifted the operating wavelength. Nonetheless, the proposed structure was able to successfully capture depth effect on the reconstructed images, with potential applications in diverse areas such as visual arts, entertainment, and security. The latter in particular will certainly benefit from the increased degree-of-freedom conveyed by stereoscopic information.

Most of the proposed metasurfaces in the literature are static, which means that it is unchangeable once a functionality is encoded on it. This constrain could be removed if the wave-matter interaction of the structure could be tuned. Therefore, reconfigurable metasurfaces could be encompass several exciting applications such as dynamical holography, switchable optical devices such as metalenses with controllable focus, beam steering to mention a few. There are several ways to perform this control and it has been already reported with liquid crystals [104], phase change materials [105, 106], graphene [107], carrier density control [108], mechanical control [109] to mention a few. Nevertheless, an ultrathin dynamical holographic display based on metasurface concept still have not been achieved yet, with a proof-of-concept demonstrated only at microwave regime [110]. Therefore, as a future work, we seek to contribute to this area and study approaches for the development of this concept at optical regime.

Furthermore, the versatility allowed by a metasurface and its ultrathin thickness make it a platform suitable for biosensing applications. For instance, by using structures with resonances possessing high quality factors, a refractive index and temperatures metasurface sensor with sensitivities of of 242.44nm/RIU and 50.47 pm/°C was reported [111]. Another work reported a plasmonic metasurface capable of sensing nonspecific BSA protein binding [112]. Therefore, this platform has been proven suitable for biosensing and another future work aims to develop a metasurface platform to operate as compact and highly sensitive biosensors.

IV. Design, fabrication and characterization of aberration-free metalenses

IV.1. Introduction

The current advances in optical device integration have allowed unprecedented miniaturization of optical systems in the most diverse applications, such as high numerical aperture (NA) microlens arrays [115], ultracompact multi-lens objectives [116], micro-optical elements [117], to mention a few. However, these devices still rely on conventional refractive or planar lenses, which poses a limit to further miniaturization. Moreover, conventional lenses require laborious fabrication steps to mould their surfaces properly, partially alleviated with the use of spherically shaped lenses despite their worsening of third order (Seidel) aberration. These aberrations usually require several lenses in tandem to eliminate them [1], posing an additional limit to miniaturization. Conventional planar lenses, such as diffractive and Fresnel plate, demand even more complex fabrication methods that make them inadequate for optical integration [118]. Additionally, diffractive lenses suffer from the generation of spurious diffractive orders that lead to unwanted focuses and ghost imaging.

Metalenses are an emerging technology that uses nanostructures to modulate the amplitude and phase of an optical beam, thus achieving optical functionalities that can replace or improve bulk optical systems [36, 37, 42, 114, 118, 119]. A canonical example for this capability is the development of the hyperbolic metalens, which is free from spherical aberrations [120, 121], a property that is difficult to obtain in bulk optics, especially with single elements. The metalens paradigm, in contrast, can be readily designed to impose a hyperbolic phase profile, thus leading to diffraction limited resolution [58, 118, 121-125]. More interestingly, metalenses can also be designed to realise functionalities that are impossible to achieve in bulk optics. We exploit this new degree of freedom for metalens design by addressing the problem of field of view (FOV), which has not received much attention so far. Indeed, in many applications, such as imaging

systems, FOV is the crucial parameter, but metalens designs that have been put forward to date are limited to procedures that reduce off-axis aberration via numerical optimisation [126], often combined with a doublet system [127, 128]. While ingenious, these approaches are limited to moderate Numerical Apertures (NA) and do not explore the physical mechanisms limiting the FOV. An alternative strategy is to relax the requirement of diffraction limited resolution in order to achieve a wide FOV (WFOV) in a singlet system [129-131].

In this project, we demonstrate WFOV imaging with a single layer metalens, reaching a $\mathrm{FOV} > 170^{\circ}$ and an NA of 0.8. The design trades off FOV against the diffraction limited resolution equivalent to an NA of 0.27, which is however sufficient for most imaging applications, including smartphones. I discuss the physical properties of the WFOV metalens and, using Fourier analysis, we show that the WFOV metalens achieves an arbitrarily large FOV by mimicking a bulk spherical lens with infinite refractive index and infinite radius of curvature. The WFOV is therefore unique in that its bulk counterpart is impossible to obtain. Finally, we compare the WFOV metalens to the diffraction limited (DL) hyperbolic metalens, showing that they are complementary in the sense that the FOV of the former and the resolution of the latter improve monotonically with the NA.

These findings were published in the journal Optica [114] and ACS photonics [113].

IV.2. Metalenses field of view (FOV) analysis

The analysis starts by assessing how the oblique incidence affects the point spread function (PSF) of metalenses with three different phase profiles: Hyperbolic, numerically optimized and spherical profile. The first, shown in equation (IV-1), where n is the background refractive index, k_0 is the freespace wavenumber and f is the focal length of the lens, was successfully adopted by the metasurface community because it does not suffer from spherical aberrations and can reach diffraction limited focusing at normal incidence [58, 120]. For a high NA, however, this advantage is offset by the introduction of coma and other off-axis aberrations, which severely limit the FOV. To overcome this problem, we optimize the phase profile to minimize the off-axis aberration by allowing some spherical aberration to exist. Following the approach used in the literature, this is done by including polynomial functions of the radial coordinate as shown in equation (IV-2). The optimization process is described in the appendix. Finally, we investigate a metalens with the phase profile equivalent to a spherical lens with radius of curvature ρ, as shown in equation (IV-3) [3]. Notice that the radius of curvature and the focal length are constrained by the relation $\Delta n = \frac{\rho}{2f}$, where Δn is the difference between the equivalent lenses inner refractive index and the surrounding refractive index.

$$\phi_{hyp}(r) = -k_0 n\left(\sqrt{f^2 + r^2} - f\right) \quad \text{(IV-1)}$$

$$\phi_{optim}(r) = -k_0 n\left(\sqrt{f^2 + r^2} - f\right) + \sum_{i=1}^{5} a_i \left(\frac{r}{R}\right)^{2i} \quad \text{(IV-2)}$$

$$\phi_{sph}(r) = -k_0 n \frac{\rho^2}{f}\left(1 - \sqrt{1 - \frac{r^2}{\rho^2}}\right) \quad \text{(IV-3)}$$

Figure IV-1 shows longitudinal cuts of the focused fields by the three lenses. The focal length is 15 μm and NA = 0.8, where $NA = \frac{R}{\sqrt{R^2+f^2}}$, R is the radius of the metalens and $\Delta n = 0.5$ corresponding to a glass lens for the spherical profile. These simulations were done using the angular spectrum formalism [3]. More details are given in section *C.4 Angular spectrum formalism* of the appendix.

For normal incidence, the optical field is tightly focused at the focal distance for all cases [Figures IV-1 (a), (d) and (g)]. The plots show that the hyperbolic has the tightest aberration-free focal spot. Nevertheless, when the angle of incidence is increased to 30°, the focal spot of the hyperbolic lens is strongly distorted due to off-axis aberrations, as seen in Figure IV-1 (b). This aberration is also clearly apparent from the PSFs cuts along the x direction shown in Figure IV-1 (c) for different angles of incidence. Notice how more energy is cast into the lateral lobes of the PSF as the angle of incidence is increased and how the peak decreases accordingly.

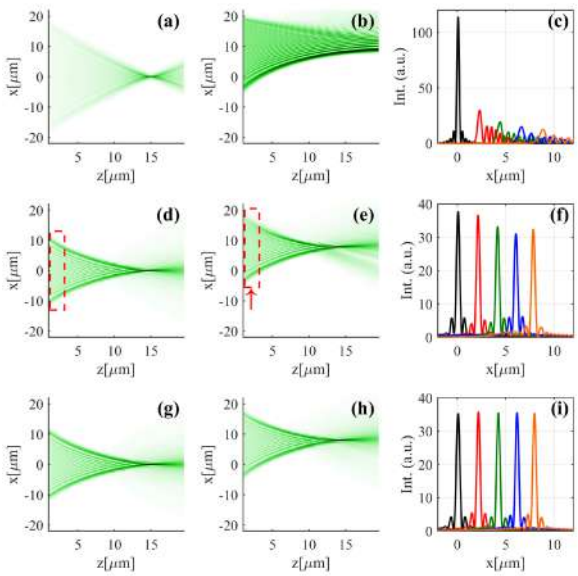


Figure IV-1 – Diffraction reconstructions of three metalenses (f=15 µm) with the following phase profiles: (a-c) Hyperbolic, (d-f) hyperbolic superimposed with optimized polynomial and (g-i) spherical. The first and second columns show the longitudinal field distributions at normal and at oblique (30°) incidence, respectively. The third column shows the resulting point spread function (PSF) at 0° (black), 7.5° (red), 15° (green), 22.5° (blue) and 30° (orange). The operating wavelength is 532 nm. The dashed boxes in (d) and (e) highlight that the corresponding phase profiles impose an effective aperture onto the lens. These simulations were done using the angular spectrum formalism [3]. More details are given in section *C.4 Angular spectrum formalism* of the appendix.

The problem of coma is clearly important and a substantial research effort has already been poured on it [120, 126-128, 132, 133]. Most attempts to reduce coma, however, have relied on numerical optimisation [127, 128, 133]. Typically, a polynomial function is superimposed onto the hyperbolic phase profile and its coefficients are optimised, as exemplified by the second term of equation (IV-3). An example of a field profile obtained with such a procedure is shown for perpendicular and angular incidence in Figures IV-1 (d) – (f), clearly showing that the coma is much reduced, although careful observation of the PSF reveals that the focal spots are slightly broadened compared to Figure IV-1 (c). The reason for this broadening is apparent from Figure IV-1 (d); even though the incident plane wave covers the entire metalens, the rays contributing to the focus emerge only from a limited area, corresponding to approximately half the radius of the lens. This means that the optimised phase profile induces an effective aperture onto the metalens and reduces the NA. Interestingly, this effective aperture (highlighted by the dashed squares) is laterally displaced when the angle of incidence is increased, as indicated in Figure IV-1 (e).

In fact, the appearance of an effective aperture and its displacement as function of incidence angle is the signature of a spherical phase profile (equation (IV-3)). Accordingly, I added a spherical profile to the comparison; Figures IV-1 (g) and (h) show the corresponding field profiles and effective aperture. The “moving” aperture appears again and the focal spot is unchanged irrespective of the angle of incidence, so the lens is free of coma. This observation is even more

apparent from the PSFs cuts shown in Figure IV-1 (i), which are nearly identical for all angles of incidence between 0 and 30 degrees.

This comparison points to the conclusion that the numerical optimisation procedure essentially converts the hyperbolic into a spherical phase profile, which points to a fundamental limitation of this approach. It also highlights the difficulty of correcting both spherical and off-axis aberrations simultaneously in a single metasurface. For example, the hyperbolic phase profile corrects for spherical aberrations yet introduces coma, while the spherical profile does the opposite. Multi-layered metasurfaces, in contrast, can correct for both aberrations at the same time [127, 128, 133] at the cost of fabrication complexity. The trade-offs that are most favourable will depend on the specific application.

Studying the Fourier Transform (FT) of the field distribution helps to understand the problem. Figures IV-2 (a) and (e) show the two-dimensional FT of the field distribution right after the metalens with spherical and hyperbolic phase profiles, respectively, as a function of *k*-vector. The light line is shown as a dashed white circle: any *k*-vector component outside the circle is evanescent and therefore does not contribute to focusing. Figures IV-2 (c) and (g) (red lines) show a corresponding line plot through the origin with the light line now indicated as a dashed black line. The hyperbolic phase profile has its Fourier components tightly confined inside the light line, and so gives the best performance for perpendicular incidence. The fact that the amplitude of its *k*-vector components increases towards the light line explains its high performance at normal incidence, because a large amplitude of high k-vector components ensures a tight focal spot. The spherical phase profile, on the other hand, has components outside the light line, i.e. the lens imposes *k*-components above $\sqrt{k_x^2 + k_y^2} > k_0$ onto the incoming beam. These components are evanescent and do not contribute to the focal spot. Furthermore, in real space, they are generated at the edge of the metalens, and are the cause of the effective aperture; so, at first sight, they appear wasted. When the input angle changes, however, it is precisely the presence of these high *k*-components that avoids coma, because the high-*k* components become available at higher input angles and then contribute to image formation. This effect is apparent when comparing *k*-space for perpendicular incidence (red lines in Figures IV-2 (c) and (g)) with angular incidence (blue lines in Figures IV-2 (c) and (g)). The effect of angular incidence is to increase the *k*-vector in that particular direction, thereby moving the spectrum sideways (see also Figures IV-2 (b) and (f)). The hyperbolic lens now loses k-vectors, thereby losing focusing power; in particular, the *k*-vector distribution becomes very asymmetric, which is a clear signature of coma and other off-axis aberrations. For the spherical lens, however, the *k*-vector components that were previously evanescent now come into play and move into the light line. Overall, its *k*-vector distribution remains largely similar to the un-shifted distribution at perpendicular incidence, so the focusing

performance will also be similar. Therefore, the imposition of the effective aperture by high *k*-vector components and the reduction of coma are directly related and are not accidental.

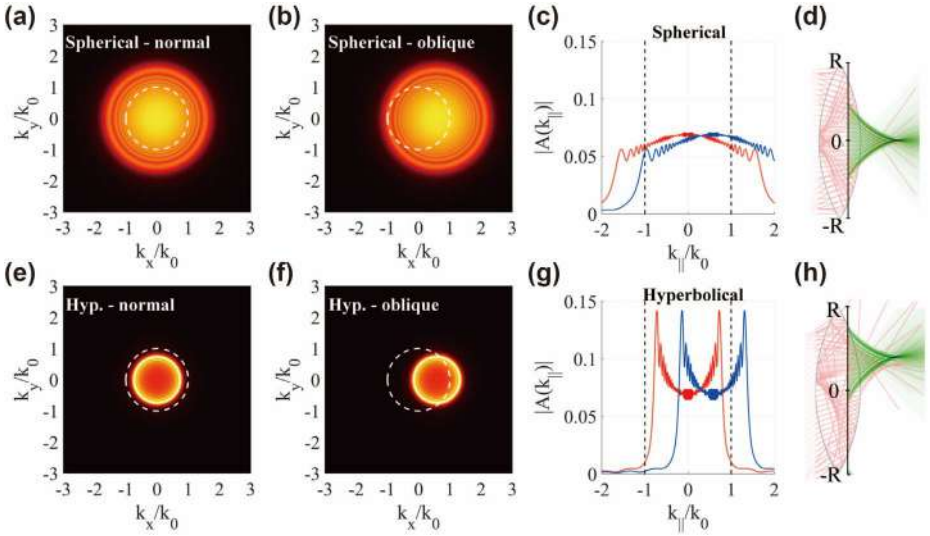


Figure IV-2 - Fourier transform amplitude of the field at the exit of the spherical metalens at normal (a) and oblique (b) incidence as a function of normalised k-vector. (c) Corresponding line plot through the origin showing normal incidence (red) and oblique incidence (blue) (e-g) show the same for the hyperbolic metalens. The light line is shown by the white dashed circle in (a), (b), (e), and (f) and as dashed black lines in (c) and (g). Ray tracing of a bulk spherical lens is shown in (d) for perpendicular incidence and (h) for oblique incidence, superimposed onto the field distributions of an equivalent metalens with spherical phase profile.

The origin of the effective NA can also be appreciated from a ray optics point of view, as we illustrate with a conventional ("bulk") lens with spherical profile. Figure IV-2 (d) shows the ray tracing picture of such a lens for normal incidence and we superimpose the field profile of an equivalent metasurface with spherical phase profile (as in equation (IV-3)). Notice that the effective aperture of the metalens coincides with the area of the rays forming the focal spot of the bulk lens. Figure IV-2 (h) shows the same for incidence at 30 degrees and we note that the analogy largely holds. So, a conventional spherical lens also imposes an effective aperture. This effective aperture does not act as a physical aperture, however, since light entering the high gradient region may pass through undiffracted, producing haze and reducing image contrast. Therefore, it may be advantageous to use a physical aperture, as e.g. provided by a multilayer lens [127, 128].

It is worth emphasizing at this point that it is not possible to obtain a single phase profile that corrects perfectly for coma and for spherical aberration at the same time [120, 126]. Any phase profile will be a trade-off between NA, spherical aberration and coma, so the challenge is to find the best compromise.

An alternative compromise between NA, spherical aberration, and coma is to immerse a spherical profile metalens into a high refractive index medium. The image quality of this arrangement would be limited by spherical aberrations, yet high NA would be achievable without coma. For example, we found that immersion in an oil of refractive index n = 1.56 raises the maximum effective NA to 1.56×0.707 = 1.1, without image quality deterioration compared to the

non-immersed case. The possibility of oil immersion as offered by high refractive index materials, therefore, is important for achieving coma-free and high NA metalenses.

IV.3. Spherical-phase profile analysis

As shown in the previous section, the diffraction limited hyperbolic profile suffers from strong off axis aberrations and trying to fix it with optimization boils down to the spherical profile. But a phase profile equivalent to what radius of curvature would give result in the best metalens? It is known that the focal length and radius of curvature are constrained by the refractive index contrast of the lens, as described by the equation $\Delta n = \frac{\rho}{2f}$. Nevertheless, a metalens can impose any phase profile and is not constrained by this relation whereas bulk lenses are subject to its inner refractive index.

In order to understand how the physics of a spherical phase profile based metalens relates to bulk optics, consider a conventional spherical lens with light incident at oblique incidence at some angle $\theta = 30°$, as shown in Figure IV-3 (a) for a conventional lens with R = 500 μm and f = 500 μm (ray tracing obtained using Comsol). The focal distance f of the spherical lens is given by the ratio of the radius of curvature R and the refractive index contrast Δn between the lens and the surrounding medium ($f = \rho/2\Delta n$). It is obvious from the equation that the same focal length can be achieved by keeping the ratio of radius and index contrast constant. Conventional optical systems are constrained to index contrasts of typically $\Delta n < 1$ by the availability of suitable materials, so this ratio is not explored in practise. Metalenses do not have this constraint, however, and they can be designed to mimic any radius and any refractive index. Figure IV-3 explores this opportunity. For example, Figure IV-3 (b) shows the example of a lens with twice the radius and twice the index contrast (R = 1000 μm, Δn = 1) of the conventional lens. Note that the ray convergence is already improved compared to the lens of Figure IV-3 (a). The convergence then improves further as the radius of curvature is increased while keeping the focal length constant, as illustrated in Figure IV-3 (c), for which R = 1500 μm, Δn = 1.5.

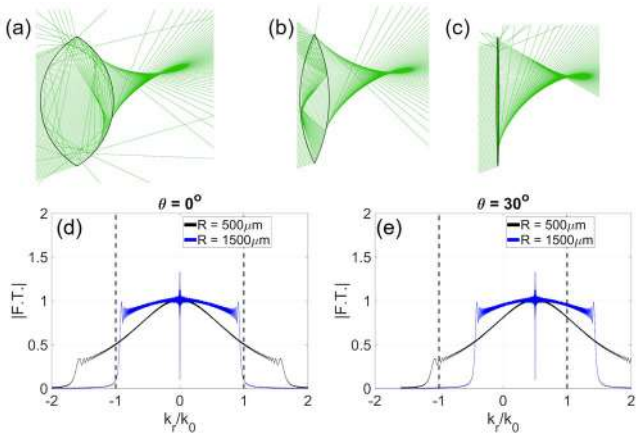


Figure IV-3 - Comparison between the WFOV metalens and spherical bulk lens. Ray traces for a spherical bulk lens (NA=0.65) for oblique incidence (θ=30°) and a radius of curvature of (a) 500 µm, (b) 1000 µm, and (c) 1500 µm. (d) FT amplitude of the field generated by the spherical lenses in (a) (black line) and (c) (blue line); the FT in (d) is for perpendicular incidence and (e) for oblique incidence (θ=30°).

This improvement of convergence with increasing radius and refractive index can be readily understood via the Fourier Transform (FT) of the field distribution immediately after the lens for perpendicular incidence (Figure IV-3 (d)), i.e. by considering *k*-space. The FT of the large radius lens (Figure IV-3 (d)- blue line) is flatter than that of the conventional lens (Figure IV-3 (d) - black line), which increases the symmetry of the Fourier components (see sections C.2 and C.3 on the Appendix for more details). Symmetry in k-space is important because it ensures that the rays contributing to image formation from opposite sides of the lens have equal strength, thus forming a well-defined spot. For oblique incidence (Figure IV-3 (e)), this symmetry is compromised, unless the FT is flat, which is the case for large radius spherical lenses (see also Figures IV-5 (a) – (b). Note that the asymmetry in k-space is particularly pronounced for the hyperbolic lens, which explains its very limited field of view [114].

These insights clarify the optical performance of the WFOV lens. Indeed, the parabolic profile is obtained in the limit of both ρ and Δn going to infinity while keeping $f = \rho/2\Delta n$ constant, as shown in Equation (IV-4).

$$\phi(r) = \lim_{\substack{(\rho,\Delta n)\to(+\infty,+\infty) \\ f\ const.}} -k_0 n \frac{\rho^2}{f}\left(1-\sqrt{1-\frac{r^2}{\rho^2}}\right) = -\frac{\pi r^2}{\lambda_0 f} n \qquad \text{(IV-4)}$$

The spectral flattening of the spherical lens can be readily seen by the Fourier Transform of the spherical phase profile as shown in appendix *C.3 Spectrum flattening of the spherical phase profile.*

Since a metalens with parabolic profile corresponds to the limit of a spherical lens with infinite radius and infinite refractive index, we conclude that the WFOV metalens does not have a bulk counterpart. We notice that a bulk lens with parabolic profile is not equivalent to the WFOV lens, since their optical characteristics coincide only for paraxial rays.

Notice that the limit of the spherical lens when the radius and refractive index go to infinity, keeping the focal length fixed, is given by a quadratic profile. One might then think that the WFOV lens behaves just like a bulk lens with a quadratic profile. However, I emphasize that this is not the case, however, because their behaviour is only equivalent for paraxial rays, as illustrated in Figure IV-4, which shows a ray tracing comparison between the WFOV metalens (left) and quadratic bulk lens (right).

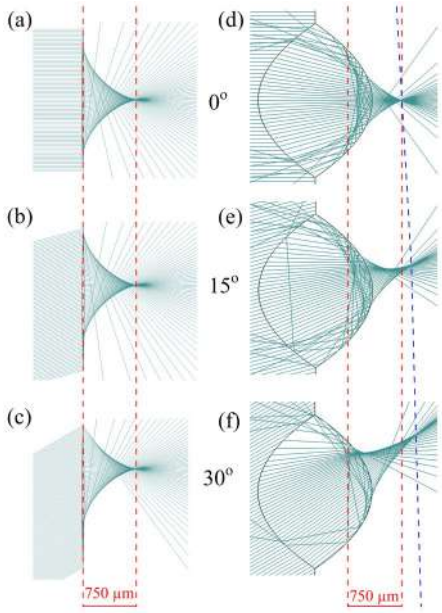


Figure IV-4 - Comparison between the WFOV metalens (first column) and an equivalent bulk quadratic lens (second column) at different angles of incidence . The red dashed lines mark the output aperture of the lenses and focal positions at normal incidence, which coincides at oblique incidence for the quadratic metalens. The blue dashes line marks the bulk lens focal position for different angles of incidence. All lenses have a focal length of 750 µm with NA = 0.8.

Notice that the WFOV metalens, for being flat, focuses light at the same distance (750 µm), irrespective of the angle of incidence. The focal point of the bulk quadratic lens, in contrast, is dependent of the angle of incidence (which is a manifestation of Petzval curvature aberration). This effect is highlighted by comparing the blue dashed line in Figures IV-4 (d)-(f), that mark the focal point for different angles of incidence, with the red dashed line that marks the focal distance at normal incidence. Notice how the focal length of the bulk lens increases at oblique incidence, tracing the Petzval field curvature.

IV.4. Optical performance of a WFOV metalens

The optical characteristic of the WFOV lens can also be understood intuitively by conducting a geometrical analysis, where the effect of angular incidence is to add a linear term to the phase profile. As shown by Pu et.al. [129], the parabolic profile converts the linear phase into a change of the origin of the coordinate system, so that the effect of angular incidence is only a translation of the focus spot away from the centre.

The FT of the WFOV lens is further explored in Figure IV-5 (a) for NAs of 0.4 (black line), 0.65 (blue line) and 0.9 (red line). Note that, the larger the NA, the wider the FT. In fact, for an NA ≈ 0.9 the FT covers the range of normalized *k* vectors from -2 to 2, i.e., it covers a region twice as large as the light line, which makes the FT invariant for oblique incidence (Figure IV-5 (b)), thus resulting in a FOV = 180°. We note a very interesting difference of lens performance on NA; for most lens designs, including the hyperbolic design, it is the spatial resolution which increases with increasing NA. For the WFOV lens, however, it is the FOV that increases with increasing NA. This interesting behaviour is explored in Figures IV-5 (c)-(d). As expected, the FWHM of the DL (hyperbolic) lens (Figure IV-5 (c) – black line) decreases monotonically as the NA increases, while the FWHM for the spherical (blue line) and WFOV (red line) lenses remain constant beyond an NA of 0.2 (spherical) and ≈ 0.3 (WFOV). The FTs of the WFOV and DL lenses are very similar up to an NA ≈ 0.3 (see section *C.2 Fourier Transforms of quadratic (WFOV) and hyperbolic (DL) field profiles.* on the appendix for the Fourier Transform profiles), and consequently their FWHM are identical in the low NA regime; for higher NAs, however, their FTs differ, which is a manifestation of spherical aberrations and limits the FWHM of the WFOV lens.

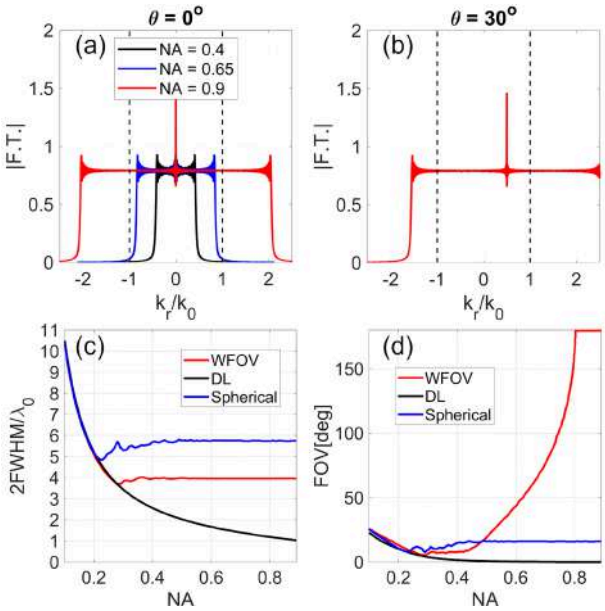


Figure IV-5 – (a) Fourier Transforms of the WFOV lens for an NA of 0.4 (black), 0.65 (blue) and 0.9 (red), for perpendicular incidence. (b) same as red line in (a), but for oblique incidence. (c) Spatial resolution (2FWHM/λ_0) and (d) FOV as a function of NA for the DL (black), spherical (blue) and WFOV (red) metalenses. The operating wavelength is 532 nm and focal length is 500 µm for all cases. All simulations were performed using the angular spectrum formalism. More details are given in section *C.4 Angular spectrum formalism* of the appendix.

In terms of field of view, however, the roles are reversed. The FOV of the DL lens now rapidly decreases as the NA increases (Figures IV-5 (d) – black line – see section *C.5 FOV characterization* on the appendix for detailed information on how the FOV is quantified), while the FOV of the WFOV lens (red line) increases monotonically with NA, until it reaches 180° for

NA ≈ 0.8. Note that the FOV of the spherical lens (blue line) saturates at a value of ≈ 20°. This limitation of the spherical lens highlights the unique advantage of the WFOV design: even though the WFOV lens is fundamentally based on the spherical design, it achieves much higher performance by going to the $R,n \rightarrow +\infty$ limit. Therefore, while the spatial resolution of the DL lens increases with NA (at the expense of FOV), it is the FOV that increases with NA for the WFOV lens. We note, however, that the spatial resolution of the WFOV lens can be readily improved by using oil immersion, while the limited FOV of the DL lens can only be improved with a suitable phase corrector, e.g. in a doublet configuration [127, 128].

IV.5. Characterization

IV.5.1. Materials and methods

a) Metalens design

The metalenses are realised using an array of 230 nm tall c-Si posts on a sapphire substrate, similar to the design I have used for the metasurface holograms, as show in section *III.1.2 Metasurface design*. More details are provided in the Appendix, Section *C.1 Phase and transmission maps,* including the angle-dependent phase maps. Micrographs and a photograph are shown in Figures IV-6 (a)-(c), highlighting the excellent fabrication quality despite the very small period of 190 nm (see item *b) Fabrication* of this section, for details on fabrication and the imaging system, and for more micrographs). The imaging system (Figure IV-6 (d)) integrates the metalens with a complementary metal oxide semiconductor (CMOS) sensor (Sony IMX219) in a 3D printed box. I fabricated two metalenses: a WFOV using a quadratic profile [129-131], and a diffraction limited (DL) metalens using a hyperbolic profile as a reference. The phase profiles of the metalenses are defined in equations (IV-1) and (IV-4).

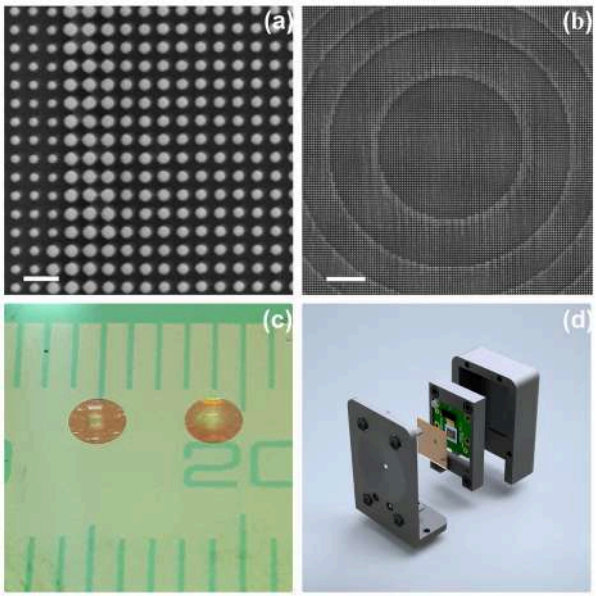


Figure IV-6 - (a) and (b) show SEM micrographs of the array of c-Si nanoposts forming the metalens. The scale bars are 400 nm (a) and 3 µm (b). (c) photograph of the fabricated metalenses. The units of the ruler are in mm indicating the 2 mm diameter. (d) Setup of the 3D-printed system encapsulating the metalens with a CCD camera.

b) Fabrication

The metalenses were fabricated by me on commercially available 230 nm thick c-Si (100) wafers epitaxially grown on a sapphire substrate (from The Roditi International Corporation Limited.). The sample was cleaned using acetone, isopropyl alcohol (IPA) and oxygen plasma. It was subsequently spin-coated with a 300 nm positive electron beam resist layer (AR-P 6200.13,

AllResist GmbH) followed by a 60 nm charge dissipation layer (AR-PC 5090, AllResist GmbH). The structure was then patterned using an e-beam system (Voyager, Raith GmbH) at a dose of 117 μC/cm^2 followed by a resist development in xylene for 2 minutes. The pattern was transferred to silicon using reactive ion etching. SEM micrographs of finished metalenses are shown in Figure IV-7.

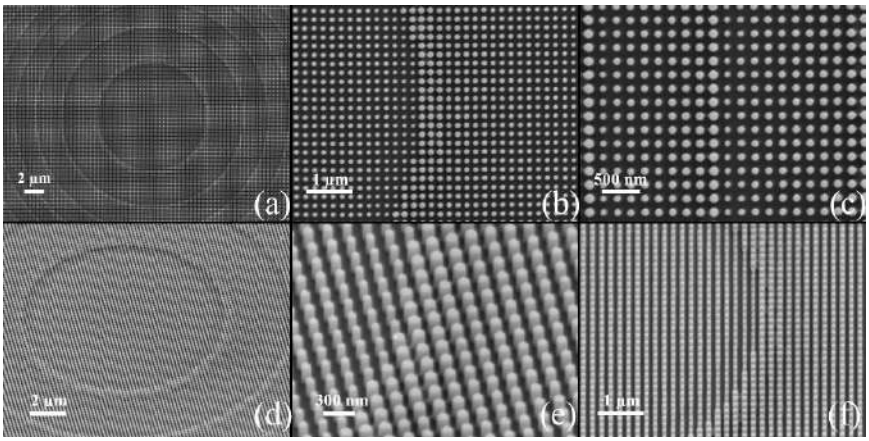


Figure IV-7 - SEM micrographs of the fabricated metalenses.

After the metalens was fabricated, I proceeded to the mask fabrication using lift off. The procedure is illustrated in Figure IV-8. I spun LOR as sacrificial layer for 60 seconds at 5000 rpm and finally baked at 180°C for 5 minutes. Then I spun negative resist SU8 2000.5 at 500 rpm for 10 seconds followed by 60 seconds at 5000 rpm and finally baked the sample at 65 °C, 95 °C and 65 °C for 1 minute, 2 minutes and 1 minute, respectively. The aperture was then patterned onto the resist using an e-beam system (Voyager, Raith GmbH) at a dose of 6.75 μC/cm^2. This dose overexposes the SU8 but since the features were in mm (apertures) scale they were not significantly affected and it was safer to ensure that the whole exposed region remained for the following steps. It should be noticed that some markers were left on the sample during the metalens fabrication to align the mask with it. Then a post-exposure bake (2 minutes at 95 °C plus two 1-minute long ramps at 65 °C) and development for 2 minutes in EC solvent were made. After this step, only the exposed areas remained on the SU8 with a layer of LOR covering the whole sample under it. The uncovered LOR regions were then developed on MF-319 for ~ 16s and rinsed in DI water. Then a layer of 100 nm of Aluminium was evaporated over the sample. Finally, the remaining LOR+SU8 layer was lift off on MICROPOSIT® REMOVER 1165 at 60 °C. Figure IV-9 shows a picture with some metalenses samples with the aluminium aperture. The defects around the bottom metalens are due the exposure the resist suffered by simply imaging the sample on the e-beam when I was searching for the markers.

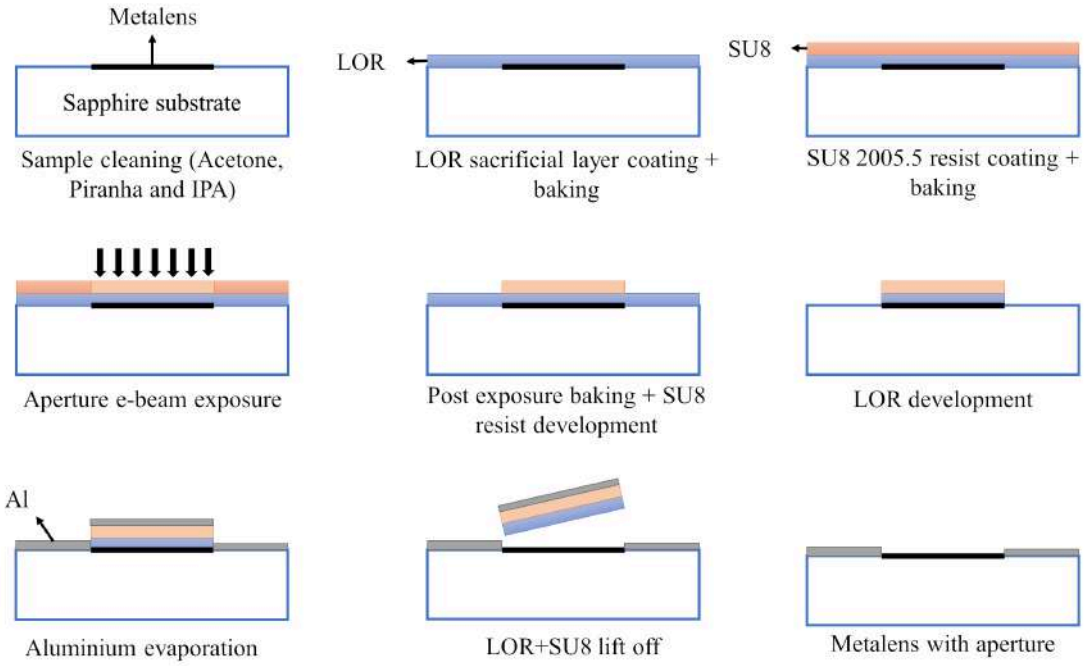


Figure IV-8 – Lift-off fabrication steps.

Figure IV-9 - Metalenses with aluminium apertures fabricated with lift-off technique.

c) Camera (imaging system)

The metalens was integrated with a CMOS sensor (Sony IMX219 8-megapixel sensor) to form a metalens camera. The camera set-up is shown in Figures IV-10 (a)-(c). Figure IV-10 (a) shows a schematic of the metalens position with respect to the CMOS sensor, where the red arrow indicates that the CMOS position is adjustable to focus the image. Figures IV-10 (b)-(c) show the integrated system without and with the lid on, respectively. The imaging system uses a monochromatic light source, as indicated in Figure IV-10 (d). In order to better show the FOV, parts of the image were magnified with the system shown in Figure IV-10 (e).

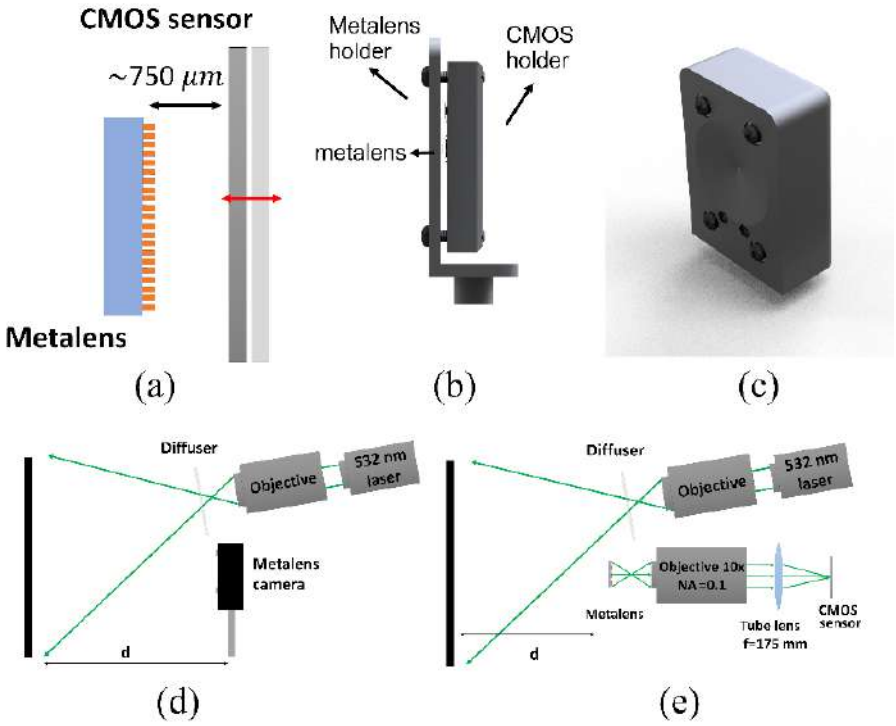


Figure IV-10 - (a) shows a representation of the 3D printed metalens camera focusing. The metalens is fixed in the front part (metalens holder in (b), which shows a side view of the inner part of the camera) and the focus is adjusted by moving the CMOS holder position through the front screws, as pictorially represented by the red arrow in (a). (c) Shows a render image of the metalens camera mounted with the lid. The monochromatic imaging setups are drawn in (d). (e) System for 10x magnified imaging, using an objective and a tube lens. The objective and diffuser in front of the laser are used to diverge the beam and eliminate speckle noise.

d) Point Spread Function (PSF) characterization setup

The PSFs were measured using a rotation stage, as shown in Figure IV-11. The metalens was illuminated by a laser at a wavelength of 532 nm, and the corresponding PSFs were then imaged onto the CMOS sensor.

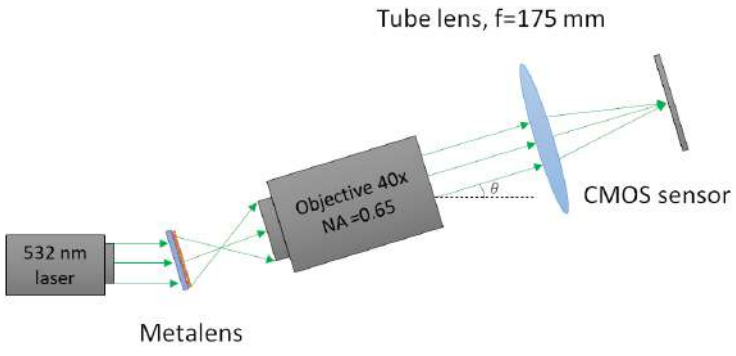


Figure IV-11 - Point spread function characterization setup. The metalens, objective, tube lens and camera were mounted in a rotation stage.

IV.5.2. Point spread function (PSF) characterization

The extremely large FOV of the WFOV metalens is demonstrated by analysing the Point Spread Function (PSF) shown in Figures IV-12 (a) and (b) for different angles of incidence (see section *IV.5.1.d) Point Spread Function (PSF) characterization setup* for details on the experimental setup). The normalized PSF of the WFOV metalens is virtually unchanged for an incoming angle as high as 89° (Figure IV-12 (a)), while the normalized PSF of the DL lens is already badly distorted for an angle of only 2° (Figure IV-12 (b)). The focusing efficiency of the WFOV metalens is virtually unchanged up to an angle of incidence of 20° and reduces for higher angles due to a reduction in the transmission efficiency. A more detailed analysis of the efficiency is shown in section *IV.5.3 Focusing efficiency* . The angular dependence of the PSFs full width at half maximum (FWHM) in the x-direction in Figures IV-12 (a)-(b) is shown in Figure IV-12 (c), comparing experimental and theoretical (see Appendix, Section *C.4 Angular spectrum formalism*, for details on the PSF simulation method). For perpendicular incidence (0°), the DL lens has a diffraction limited FWHM, which explains its high performance at the very centre of the image, while the FWHM of the WFOV lens is twice as large (notice in Figure 12 (c) that the measured FWHM of the DL lens is larger than the theoretical, which I attribute mainly to the objective used in the imaging of the PSF). The FWHM of the DL lens, however, rapidly increases for small angles of incidence, while the FWHM of the WFOV lens is constant up to angles as high as ±89°.

Notice that the WFOV single layer metalens achieves a FOV > 170° and a FWHM ~ $2\lambda_0$, which is comparable to bulk optics WFOV lenses. For example, the wide and ultra-wide lenses of the iPhone 11 Pro Max have a FOV = 60°, FWHM ~ $1.87\lambda_0$ and a FOV = 120°, FWHM ~ $2.26\lambda_0$, respectively [134, 135]. As another example, the Nikkon AS-Fisheye NIKKOR 8-15mm f/3.5-4.5E ED system has a FOV = 60° and FWHM ~ $3.54\lambda_0$.

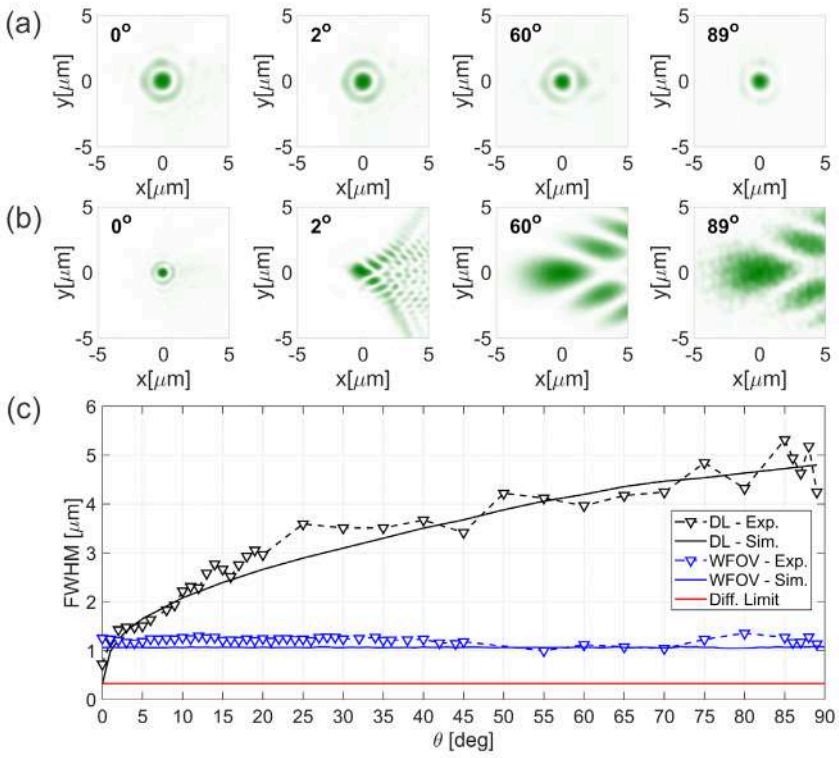


Figure IV-12 - PSFs for different angles of incidence for the WFOV (a) and the DL (b) metalenses. The x axis in (a) and (b) is centred at the maximum of each PSF. (c) Dependence of the FWHM on the angle of incidence for both lenses. Focal length is 750 µm, diameter is 2 mm, operating wavelength is 532 nm and NA=0.8.

IV.5.3. Focusing efficiency

a) Experimental

The angle dependent focusing and transmission efficiencies of the quadratic and hyperbolic metalenses are shown in Figure IV-13 (a) (quadratic) and Figure IV-13 (b) (hyperbolic) for both *s* (black lines) and *p* (blue lines) polarizations. The theoretical efficiencies, determined by numerical simulations, are shown in Figure IV-14 of the next item of this subsection. The focusing efficiencies were determined by integration of the energy in the focal spot using a circular aperture with a radius of 7.5 µm. The dotted lines show the transmission efficiencies. To better assess the metalens performance, the experimental focusing efficiency with respect to both incident (solid lines) and transmitted power (dashed lines) are shown in Figure IV-13. Comparing these two efficiencies is helpful to determine whether a reduction in the focusing efficiency comes from a loss of the metalens ability to focus light, or from a loss of the transmitted power itself. Therefore, the reduction of the focusing efficiency with respect to the incoming power (Figure IV-13 (a), solid lines) is not due to a loss of ability to focus light, but it is rather due to a loss of transmitted power itself (Figure IV-13 (a), dotted lines), which is mostly due to the cosine dependent projected area. Nevertheless, the efficiency with respect to incoming power (Figure IV-13 (a), solid lines) is virtually unchanged up to an angle of 20º.

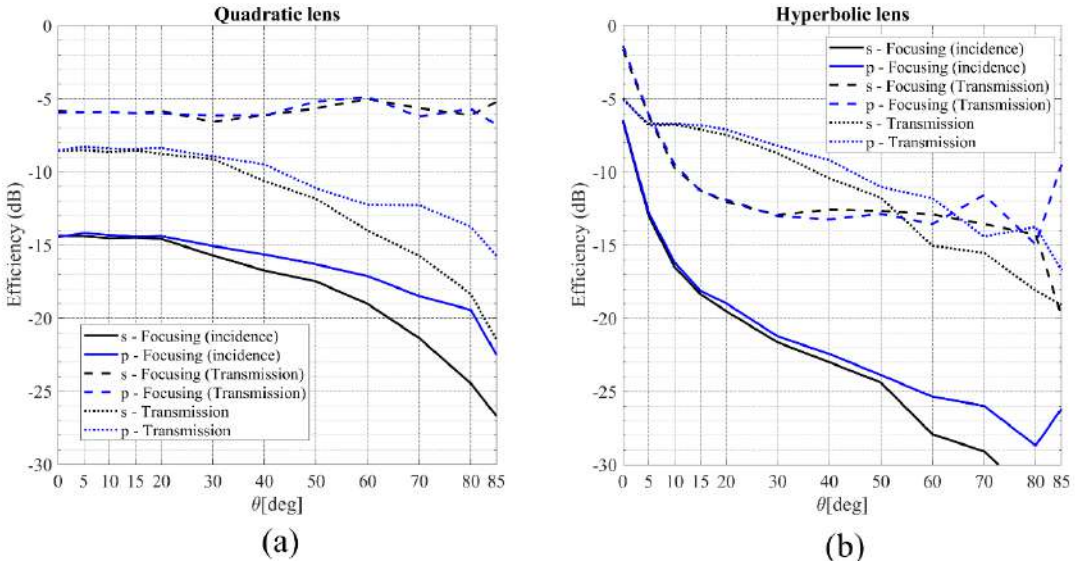


Figure IV-13 - Measured transmission (dotted lines) and focusing efficiencies (dashed and solid lines) as a function of the angle of incidence for both s and p polarisations. The focusing efficiency is determined with respect to both incoming power (solid lines) and transmission power (dashed lines). (a) efficiencies of the quadratic metalens. At perpendicular incidence, the transmission efficiency is 14%, the focusing efficiency with respect to incoming power is 3.5% and the focusing efficiency with respect to transmitted power is 25%. (b) efficiencies of the hyperbolic metalens. At perpendicular incidence, the transmission efficiency is 31%, the focusing efficiency with respect to incoming power is 23%, and the focusing efficiency with respect to transmitted power is 74%.

As expected, the efficiencies of the quadratic lens are more tolerant at oblique incidence than the efficiencies of the hyperbolic lenses. Notice, however, that the efficiencies of the quadratic lens (3.5% for focusing with respect to incoming power and 14% for transmission) are much lower than the efficiencies of the hyperbolic lens (23% for focusing with respect to incoming power and 31% for transmission). This is an intrinsic consequence of a combination of spherical aberrations and the effective numerical aperture of the quadratic lens [129], (for a more detailed analysis, see the next item). This feature highlights the need of using low loss materials, such as c-Si, in the design of WFOV metalenses. Notice that the simulated focusing efficiency (shown in Figure IV-14, solid lines) is -13 dB, (i.e., 5%), as compared to the -14.5 dB (i.e., 3.5%) of the experimental efficiency (Figure IV-13 (a), solid lines). We therefore achieved an experimental focusing efficiency of 3.5/5*100% = 70% compared to the simulated efficiency.

b) Theoretical

The measured efficiencies are shown in Figure IV-13 and their corresponding simulated values are shown in Figure IV-14.

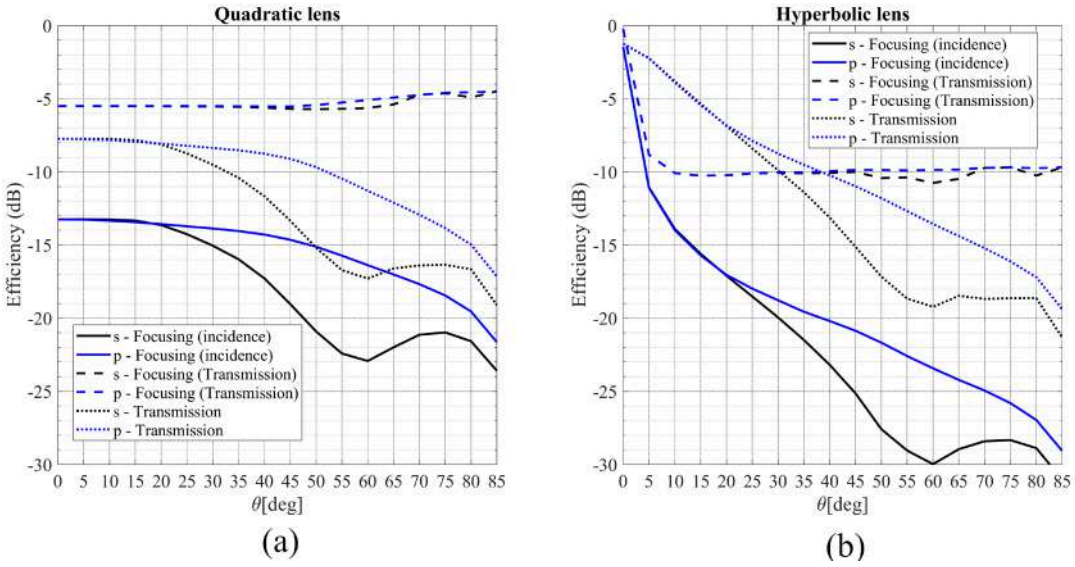


Figure IV-14 - Simulated transmission (dotted lines) and focusing efficiencies as function of the angle of incidence for both s and p polarisations. The focusing efficiency is determined with respect to both incoming power (continuous lines) and transmission power (dashed lines). (a) efficiencies of the quadratic metalens. At perpendicular incidence, the transmission efficiency is 34%, the focusing efficiency with respect to incoming power is 5%, and the focusing efficiency with respect to transmitted power is 14%. (b) efficiencies of the hyperbolic metalens. At perpendicular incidence, the transmission efficiency is 75%, the focusing efficiency with respect to incoming power is 71%, and the focusing efficiency with respect to transmitted power is 95%.

Now the physical origin of the low efficiencies of quadratic metalenses is investigated. I consider a metalens with NA = 0.8 (Figure IV-15 (a)) and a metalens with NA = 0.71 (Figure IV-15 (b)). In Figure IV-15 (a) (NA = 0.8), it is highlighted the effective numerical aperture with the green colour. The metalens of Figure IV-15 (b) (NA = 0.71) is essentially "cut out" from the effective numerical aperture region of Figure IV-15 (a), so that the entire metalens area of Figure IV-15 (b) lies inside the effective numerical aperture.

Light that impinges on the lens outside the numerical aperture region is mostly reflected, while the light that impinges on the lens inside the numerical aperture is mostly transmitted, as illustrated in Figures IV-15 (c) and (d) for the metalenses with NA = 0.8 and NA = 0.71, respectively; in a hyperbolic lens, however, the entire area contributes to the focusing, as shown in Figures IV-15 (e) and (f).

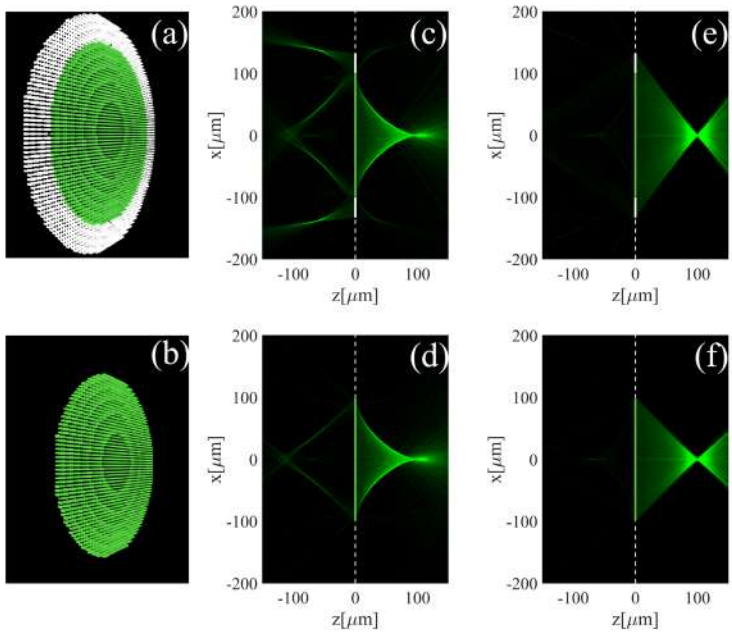


Figure IV-15 – (a) and (b) show a representation of a metalens with NA = 0.8 and NA=0.71, respectively. The radii of the metalenses are 133 μm (NA = 0.8) and 100 μm (NA = 0.71). The green and white parts mark the effective aperture and evanescent regions, respectively, of the quadratic phase profile. (c) and (d) show the transmitted and reflected field intensities (incident field not shown), respectively, of a quadratic metalens with NA = 0.8 and NA =0.71, respectively. The focal length is 100 μm and the operating wavelength is 532 nm. (e) and (f) show the same for a hyperbolic metalens with the same parameters. The field distributions were obtained through FDTD and angular spectrum simulations. These plots are saturated to highlight the intensity distribution.

In order to better understand the origin of the lower efficiency, I considered four distinct situations: 1) an ideal quadratic lens, 2) an ideal hyperbolic lens, 3) a c-Si quadratic lens, and 4) a c-Si hyperbolic lens. The ideal metalens is a surface that imparts the desired phase discontinuity, while the c-Si metasurface is a surface that imparts the phase and transmission obtained from the transmission/phase maps. Table IV-1 shows the transmission and focusing efficiencies of these four cases.

We first consider the metalens with NA = 0.71, where the entire lens lies inside the effective numerical aperture. Notice that the ideal transmissions (quoted under parenthesis) are very similar: 96.5% and 97.0% for the ideal quadratic and hyperbolic metalenses, respectively. The focusing efficiencies, however, are very different: 10.6% and 89.4% for the ideal quadratic and hyperbolic metalenses, respectively. Since in this case there is no effect of effective numerical aperture, and that the transmission efficiencies are very similar, we conclude that the lower focusing efficiency of the quadratic lens is due to spherical aberrations.

Table IV-1 – Simulated ideal and c-Si based metalenses focusing (transmission) efficiencies with quadratic and hyperbolic phase profiles with NA = 0.71 and NA = 0.8. The ideal metalens is a surface that imparts the desired phase discontinuity. The focal length is 750 *μm* for all lenses. The operating wavelength is 532 nm.

Aperture	NA = 0.71		NA = 0.8	
Type	Ideal	c-Si	Ideal	c-Si
Quadratic	10.6% (96.5%)	8.4% (76.7%)	6% (55%)	5% (34%)
Hyperbolic	89.4% (97.0%)	70.9% (77.2%)	89.9% (97%)	71% (75%)

I then compared the ideal efficiencies with the c-Si metalens efficiencies. The results are summarised in Table IV-1. Notice that the focusing efficiency of the quadratic c-Si metalens is 8.4%, which is close to the ideal 10.6%. These results corroborate the claim that the c-Si is a suitable material for WFOV metalens.

Next, I performed the same calculations, but now considering a metalens with NA = 0.8. The results are shown in Table IV-1. Comparing the transmission efficiencies of the ideal quadratic lenses, we notice that the transmission is 96.5% and 55% for the NA = 0.71 and NA = 0.8, respectively. This drop in transmission comes from the numerical aperture effect, a conclusion that can be supported by noticing that the ratio between the metalenses areas is $0.75^2 = 0.56$, which is almost the same ratio of the transmission efficiencies (55/96.5 = 0.57). This lower transmission efficiency ensues in a lower focusing efficiency of the ideal quadratic lens, which is 6% for NA = 0.8, as compared to 10.6% for NA = 0.71.

Finally, I noticed that the focusing efficiency of the c-Si metalens with NA = 0.8 is 5%, which is close to the ideal focusing efficiency of 6%, once more corroborating the advantages of using c-Si.

Thus, the lower focusing efficiency of the quadratic metalens is due to a combination of spherical aberration and effective numerical aperture.

IV.5.4. Imaging

The performance of the WFOV lens is shown in Figure IV-16 (a), in comparison to a DL design (Figure IV-16 (b)) using a USAF 1964 chart 6 cm from the lens (see Appendix, section *C.6 USAF 1951 chart,* for a description of the chart). Figure IV-17 shows the same but with another chart at ~5.7 mm from the lens (higher magnification). The metalenses have the same numerical aperture of NA = 0.8 which is here defined as $\mathrm{D}/\sqrt{\mathrm{D}^2 + 4\mathrm{f}^2}$, where *D* and *f* are the lens aperture diameter and focal length, respectively. For all lenses, I used a focal length of *f = 750* μm and a diameter of *D* = 2 mm. It is clear that the image obtained with the WFOV metalens is virtually free of aberrations, except for the barrel distortion or "fish-eye" effect, which is typical of optical systems with a wide field of view [127]. The barrel distortion arises due to a mismatch between the actual displacement of the focal spot (which depends on the *sine* of the angle of incidence), and the paraxial displacement (which depends on the tangent of the angle). The measured displacements are shown in Figure IV-16 (d). Because this displacement is predictable, it can be corrected in post-processing [136], so it does not constitute a fundamental limitation of the lens.

The image of the DL metalens, in contrast, is strongly blurred by off-axis aberrations. The only area where the DL lens achieves superior performance is at the very centre of the image. The imaging quality of the lenses is further evidenced on Figures IV-18 (a)-(b) that shows a

comparison between the WFOV and DL, respectively, reinforcing the previous conclusions. The comparison is highlighted by the rulers, that are zoomed in and shown Figures IV-18 (b)-(d) for the WFOV and DL metalens, respectively. Notice that they are well defined when placed in all regions and imaged with WFOV whereas only the on axis (viewing angle of 0°) is not blurred when imaged by the DL metalens.

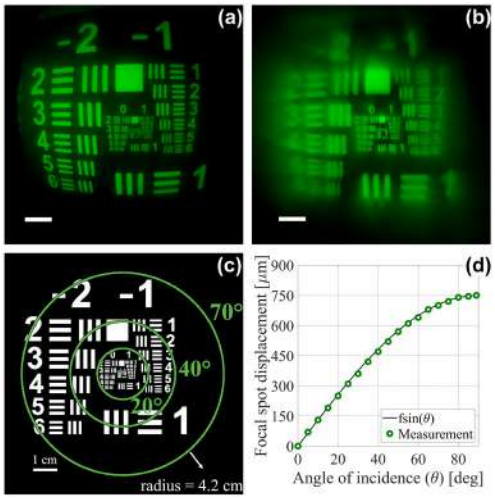


Figure IV-16 - (a) Image of a USAF 1964 chart obtained with the WFOV (quadratic) metalens using the setup of Figure IV-10 (a); the scale bar is 100 μm long. (b) Image of a USAF 1964 chart obtained with the DL (hyperbolic) metalens using the setup of Figure IV-10 (a); the scale bar is 100 μm long and the field of view is 70°. (c) Picture of the USAF 1964 chart with the field of view zones highlighted by dashed circles. See Appendix, section *C.6 USAF 1951 chart,* for a description of the chart. The field of view zones marked by the solid circles (from innermost to outermost) are 20°, 40° and 70°. The radius of the outermost circle is 4.2 cm. (d) Focal spot displacement as function of the incident angle. The continuous line shows the product of the sine function with the focal length. The distance between the lens and the object is 6 cm.

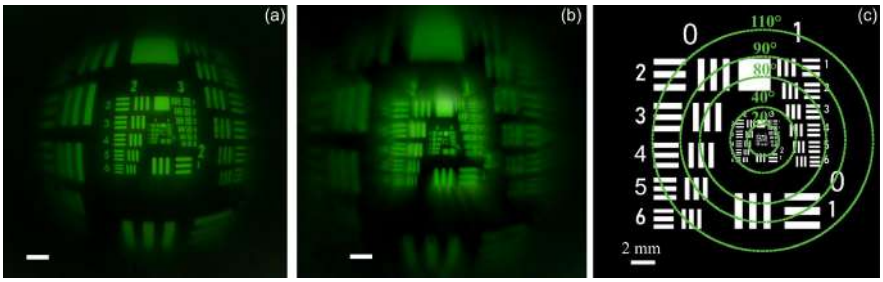


Figure IV-17 - (a) Image of a USAF 1964 chart obtained with the WFOV (quadratic) metalens; the scale bar is 100 μm long. (b) Image of a USAF 1964 chart obtained with the DL (hyperbolic) metalens; the scale bar is 100 μm long and the field of view is 140°. The distance between the lens and the object is 5.7 mm. (c) Picture of the USAF 1964 chart with the field of view zones highlighted by green circles. The field of view zones marked by the circles (from innermost to outermost) are 20°, 40°, 80°, 90°, 110°. The radius of the outermost dashed circle is 8.14 mm. See Appendix, section *C.6 USAF 1951 chart,* for a description of the chart.

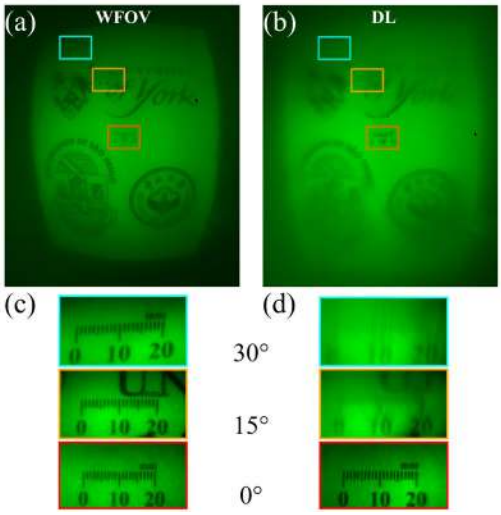


Figure IV-18 - Imaging comparison between WFOV and DL metalenses. (a) and (b) show the obtained images with the WFOV and DL metalenses, respectively, using the system of Figure IV-10 (d). (c) and (d) show a zoom in of the scales placed at viewing angles of 0, 15 and 30 degrees. These zoomed images were taken with the system of Figure IV-10 (e). The rulers unit is in mm.

Finally, I tested the WFOV metalens imaging under incoherent illumination (ambient light). To minimize the chromatic aberration, I placed a green filter in front of the camera. Figure IV-19 shows a photo of myself in the optics lab of our group at the Department of Electrical and Computer engineering at the University of Sao Paulo taken with the metalens with the filtered incoherent illumination. Despite the presence of the barrel distortion, the image presents a good quality.

Figure IV-19 – Photo taken with WFOV metalens camera under incoherent illumination (ambient light). To minimize the chromatic aberration, a green filter was used.

IV.6. Conclusions and future works

In conclusion, I have demonstrated WFOV imaging with a single layer metalens by relaxing the constraint on diffraction limited resolution; in particular, I showed that WFOV metalenses mimic a spherical lens in the limit of infinite radius of curvature and infinite refractive index. This metalens has an arbitrarily large FOV and can be seen as complementary to the established DL (hyperbolic) lens design, in the sense that the FOV of the WFOV lens increases with the NA, while it is the resolution that increases with NA in DL lenses. Notice that the WFOV metalens achieves a FWHM $\approx 2\lambda_0$, which is sufficient for all but the most demanding imaging applications, including the very best smartphone cameras, which have a maximum NA of ~ 0.3 (notice that FWHM $\approx 2\lambda$ is obtained in the DL lens with NA = 0.27). The WFOV metalens opens an unprecedented degree of freedom to control the FOV, with achromatic operation presenting the next challenge. Overall, I believe that the WFOV metalens will make an important contribution to the development of novel wide-field and high-resolution imaging applications, including smartphone imaging, virtual reality viewers and miniaturised high-end scientific imaging systems.

V. Photon Management in Tandem Si/Perovskite Solar Cells

V.1. Introduction

Control over light propagation is of paramount importance for solar cells. Traditionally, highly efficient silicon solar cells use micron-sized pyramids to improve absorption in the active layer [138], but recently there has been intense research on employing nanostructures to achieve cheaper and/or more efficient solar cells. In this chapter, I show how nanostructures can impact the performance of a promising new class of solar cells: the tandem solar cell employing perovskites and silicon [139].

For solar cells to become competitive with traditional energy sources, their cost per Watt of energy needs to be reduced. This cost depends mainly on the installation and manufacturing process as well as on the power conversion efficiency. As the average price of silicon has already dropped sharply in the last decade [140], the manufacturing costs are already very low and the installation costs are difficult to reduce, much of the research effort is now focusing on the efficiency problem. Since the efficiency of single junction silicon cells is already approaching its theoretical limit of 30% [141, 142], it is essential to seek low cost alternatives to boost the efficiency of silicon solar cells beyond their single junction limit. As a result, there has been a surge in interest in tandem solar cells using silicon as the low band gap absorber and perovskite as the high band gap absorber [139, 143-147], with the goal of exploiting the higher open circuit voltage of perovskites. This interest is justified by the combination of the mature silicon technology with the huge potential of perovskites to deliver low cost and highly efficient solar cells [139, 148, 149]. Theoretical analyses suggest that, in principle, tandem silicon perovskite cells can have efficiencies in excess of 30% [143, 144, 150], while there have been experimental reports of tandem silicon perovskite cells reaching 23.4%[145], 18% [147], 25% [151] and even 28% [152], the latter using an optical splitting system. In order to be technologically viable, tandem cells need to be realized as stacked structures, in which case the performance is highly dependent on photon management. For example, in two-terminal tandem cells, photon management is used to balance the absorption in the different layers in order to match the currents

delivered by each cell. In the alternative four-terminal configuration, photon management is used to maximize the absorption in the top cell, which has the highest band gap and therefore delivers the highest open circuit voltage (V_{oc}). Photon management can be achieved by placing intermediate photonic reflectors between the two cells, with the reflectance designed to match the top cell absorption band. Examples of intermediate reflectors include homogeneous layers [153, 154], photonic crystals [153-156], metallic nanoparticles [157], stacked layers [153, 154] and a combination of those with randomly texturized surfaces [158]. It is interesting to note that most of the attention has been paid to the reflectance at shorter wavelengths, motivated by the desire to optimise the top cell performance [153-158], with only a few papers addressing the problem of optical impedance matching into the lower cell [156]. Here, we take the optical impedance matching concept further and demonstrate that it is, in fact, the dominant effect in determining tandem cell efficiency.

The majority of intermediate reflectors proposed so far has been targeted at micromorph solar cells, with their potential for perovskite/silicon tandem cells yet to be explored [139]. As an example of this potential, Lal *et al*. have concluded that a combination of an intermediate reflector with a Lambertian scatterer can overcome the 30% limit in a perovskite/silicon tandem cell [150]. However, it is still not clear what the ideal properties of an intermediate structure should be and whether these properties can be met by realistic structures. Here, we address this problem by first identifying the ideal requirement for intermediate photonic structures in perovskite/silicon tandem solar cells in Chapter V.2. Importantly, and counter-intuitively, we show that the intermediate structure reflectance into the high band-gap material (perovskite) absorption window does not have a significant impact on the overall tandem performance. This conclusion differs from the requirement for micromorph solar cells [152, 159]. Instead of boosting the reflectance into the perovskite window, we show that the intermediate structure should mainly act as an optical impedance matching layer for the spectral region where the perovskite is transparent, i.e., the intermediate structure should maximize the optical transmission between the top cell and bottom cell. After carefully identifying this requirement, we design simple and robust photonic structures that provide broad-band optical impedance matching between top and bottom cells. The designs are thoroughly described in Chapter V.3. Finally, in Chapter V.4 the performance of different optical impedance matching structures is compared to that of intermediate reflectors, analyzing the impact of both classes of photonic structures on the overall performance of the tandem solar cell. Chapter V.5 closes this section with some concluding remarks. The conclusion is that simple and efficient structures are obtained when only an optical impedance matching layer is used. As an example, we show that by varying the reflectivity and cut-off wavelength of the intermediate reflector, the highest increase in short circuit current, assuming there is no light trapping, is 18.5% compared to a structure without intermediate reflector, and it is achieved for $R \approx 0$, i.e. for an optical impedance matching layer.

V.2. Ideal Intermediate Photonic Structure

We begin by identifying the ideal requirements for the intermediate photonic structure, which is placed between the perovskite and silicon absorbing layers of a 4-terminal tandem solar cell, as shown in Figure V-1 (a). The choice of a 4-terminal configuration has two important advantages: 1) it does not require current matching and 2) it affords the optical separation between the top and bottom cells by means of an optical buffer. This scheme is particularly convenient because it allows the cells to be fabricated separately and be subsequently bonded together. The top cell is comprised of an Anti-Reflection (AR) coating, a 400 nm thick perovskite layer and a 100 nm thick Indium Tin Oxide (ITO) electrode. The top and bottom cells are optically separated by a 1 µm silicon dioxide (SiO_2) buffer, followed by the photonic intermediate structure. Finally, the cell is terminated by a 400 µm thick crystalline silicon (c-Si) absorber covered with a perfect mirror. The refractive index of the AR coating is set to 1.45 and the dispersion of all materials is shown in Figure V-2. As the choice of the optimum transparent front contact is still under active research [160-164], we opted to first perform the calculations without any particular choice of front contact and transport layers, so that the results can be kept as general as possible. However, the performance characterization for a complete device using ITO as the front contact and multiple perovskite thicknesses is provided in section *V.4.2 Complete device characterization*.

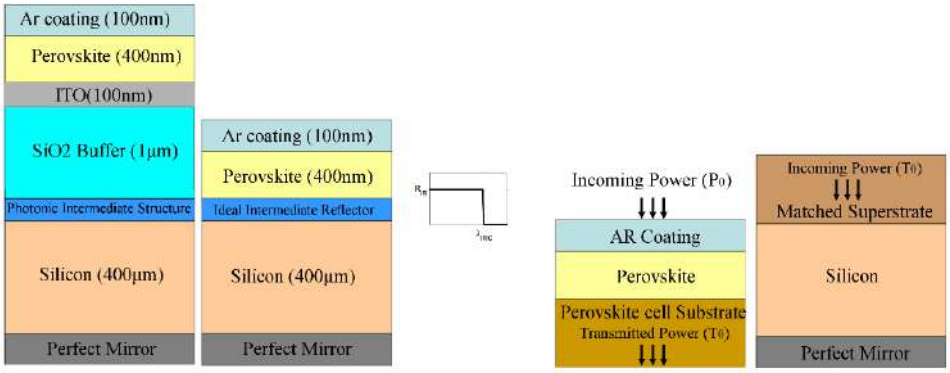


Figure V-1 – (a) Intermediate photonic structure on a perovskite/c-Si tandem solar cell. The thickness of each layer is stated in parenthesis in front of the material. The refractive index of the AR coating is set to 1.45. First, the optimum properties of an ideal intermediate reflector are identified. The ideal reflector is depicted in (b). In (c), the absorption in the perovskite layer is calculated by choosing a substrate such that the Fresnel reflection coefficient between the perovskite layer and the substrate gives the desired ideal reflectance - according to the top inset of Figure V-3 (a). In the second step, (d), the power transmitted into the substrate of (c) is transposed to a matched superstrate (with the same real part of refractive index as silicon), so as to avoid reflection from the silicon layer. The structure of (d) is then used to calculate the absorption in the silicon layer. The light reflected from (d) is considered as loss and does not reach (c) again.

The ideal properties of the intermediate structure can be identified by considering an ideal reflector placed between the perovskite and silicon layers, as shown in Figure V-1 (b). The ideal reflector spectrum is shown in top the inset of Figure V-3 (a): it has a fixed reflectance $\boldsymbol{R_{ir}}$ up to a cut-off wavelength $\boldsymbol{\lambda_{irc}}$, after which the reflectance drops sharply to zero. The impact of the

ideal reflector on the solar cell efficiency can thus be assessed by varying the reflectance $\boldsymbol{R_{ir}}$ and the cut-off wavelength $\boldsymbol{\lambda_{irc}}$.

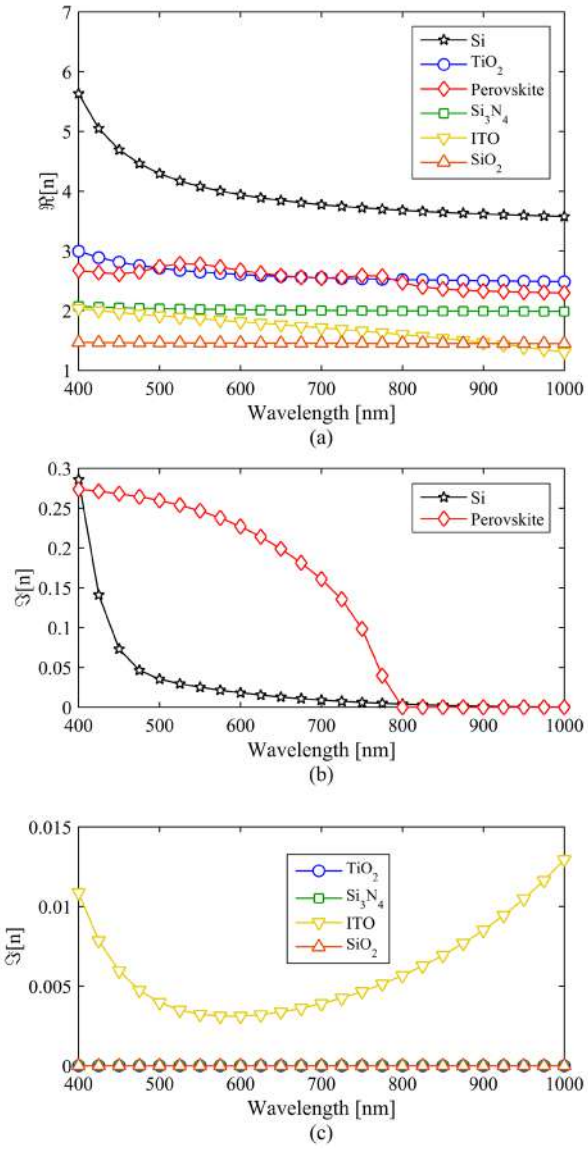


Figure V-2 - (a) Real and (b), (c) imaginary parts of the refractive indices used in the calculations.

The absorption spectra of the perovskite and silicon layers in the presence of the ideal reflector are then calculated in two steps. Firstly, the perovskite absorption and transmission are calculated by replacing the entire bottom cell, including the optical buffer, by a substrate whose refractive index is calculated from the Fresnel equations to provide the chosen reflectance $\boldsymbol{R_{ir}}$, as shown in Figure V-1 (c). Secondly, the power transmitted into the substrate of Figure V-1 (c) is assumed to be the incident power on the matched superstrate of Figure V-1 (d). Notice that the power incoming from the superstrate is the total incident power minus the power lost to both reflection and absorption in the perovskite layer. The configuration of Figure V-1 (d) is then used to calculate the absorption in the silicon layer. The calculations assume perpendicular incidence

and AM1.5G solar spectrum. All optical calculations were performed using the Rigorous Coupled Wave Analysis (RCWA) [88].

The calculated absorptions were then used to determine the performance of the tandem cell. Following the procedure described in [143], the bottom silicon solar cell parameters are chosen based on the c-Si PERL solar cell [165]; they are: charge carrier collection probability of 0.978, fill factor of 82.8% and open circuit voltage V_{oc}^{Si} given by the following diode equation.

$$V_{oc}^{Si} = \frac{kT}{q} \ln\left(\frac{J_{sc}^{Si}}{J_0} + 1\right) \qquad \text{(V-1)}$$

where k is the Boltzmann constant, T is the temperature in K, q is the fundamental charge, J_{sc}^{Si} is the short circuit current density in units of mA/cm^2 and $J_0 = 4.9 \times 10^{-11}$ mA/cm^2.

The top perovskite solar cell short-circuit current, open circuit voltage and efficiency are also calculated according to [143]. Following Ref. [143], the parameters used in the calculations are: diffusion length L_d = 100 nm, luminescence efficiency of 0.55, fill factor of 80% and band-gap of 1.55 eV. Even though the optimum band-gap is around 1.8 eV., we opted to choose a band-gap that is more similar to current perovskite materials.

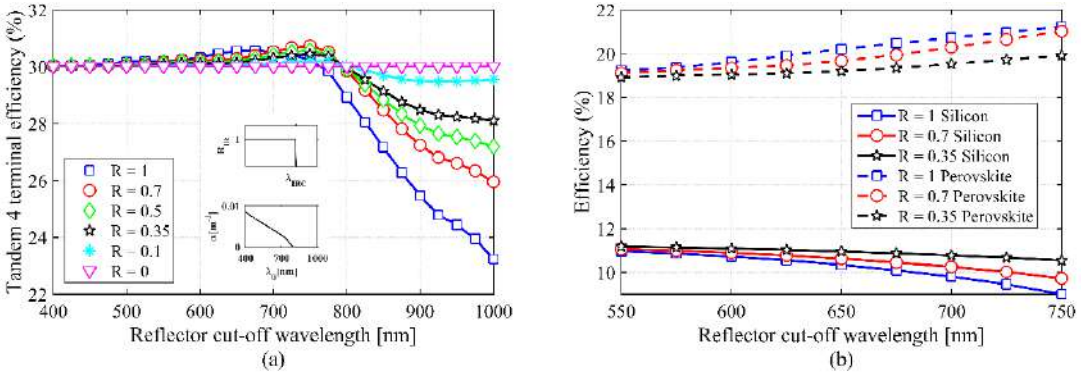


Figure V-3 - (a) Dependence of the tandem solar cell efficiency on the ideal reflector cut-off wavelength. The efficiency is not strongly dependent on the reflectance for cut-off wavelengths inside the perovskite absorption window, but drops sharply for high reflectance when the cut-off wavelength reaches the perovskite transparency window. The top inset shows the reflectance of the ideal reflector and the bottom inset shows the absorption coefficient of perovskite. (b) Dependency of the silicon (continuous line) and perovskite (dashed lines) layers' efficiency on the ideal reflector cut-off wavelength.

Figure V-3 (a) shows the tandem solar cell efficiency as a function of the cut-off wavelength $\boldsymbol{\lambda_{irc}}$ for different values of $\boldsymbol{R_{ir}}$. According to the ideal reflector spectrum (top inset of Figure V-3 (a)), the reflectance between the perovskite and silicon layers is fixed to $\boldsymbol{R_{ir}}$ up to the cut-off wavelength. For wavelengths above $\boldsymbol{\lambda_{irc}}$, the ideal reflector acts a perfect optical impedance matching layer between perovskite and silicon. As perovskite absorbs up to 800 nm, it is reasonable to assume that the ideal reflector optimum parameters should be $\boldsymbol{R_{ir}} = 1$ and $\boldsymbol{\lambda_{irc}} \approx 800$ nm, since this is the condition that maximizes absorption in the high band-gap (perovskite) layer. The results of Figure V-3 (a), however, contradict this assumption: it is clear that the total efficiency is not maximized for $\boldsymbol{R_{ir}} = 1$ and, more importantly, it depends only weakly on $\boldsymbol{R_{ir}}$ for

λ_{irc} up to ~750 nm. In the wavelength region near cut-off (550 nm < λ_{irc} < 800 nm), one would expect a strong dependence of the tandem efficiency on R_{ir} because of the strong overlap of high solar photon density with the perovskite absorption window. According to Figure V-3 (a), however, this dependence is not pronounced and, moreover, it peaks at R_{ir} = 0.7 and λ_{irc} = 750 nm instead of the expected R_{ir} = 1, with only a weak dependence on R_{ir}. Therefore, these results lead to the conclusion that high reflectance in the perovskite absorption window is not a strong requirement.

We propose that this counter-intuitive behaviour is a consequence of the balance between reflection and thermalization losses, which is best explained by means of an example: assuming a perfect AR coating, for R_{ir} = 0, the absorption at the wavelength of 700 nm is about 80% in the perovskite layer and 20% in the silicon layer (at this wavelength, the 400 μm thick silicon layer absorbs virtually all photons that reach it). For R_{ir} = 1, however, and at the same wavelength of 700 nm, the absorption in the perovskite layer is about 90%, while in the silicon layer it is 0% (since no photon reaches it). This means that R_{ir} = 1 results in higher absorption in the perovskite layer and, consequently, lower thermalization losses. Yet, this improvement is more than outweighed by the loss of efficiency in the silicon. Consequently, there is a trade-off between reflection and thermalization losses that cannot be overlooked in the intermediate structure design; essentially, while any value of R_{ir} > 0 increases the absorption in the perovskite layer, it also increases the reflection back into free space, so photons are being lost rather than being absorbed in the silicon. It is this trade-off that accounts for the low dependency of the tandem efficiency on R_{ir} in the spectral region 550 nm < λ_{irc} < 800 nm. This compensation is clearly seen when the efficiencies for the perovskite and silicon solar cells are plotted independently (Figure V-3 (b)): as R_{ir} increases, the efficiency of the perovskite top cell also increases by approximately the same amount as the efficiency of the silicon bottom cell decreases. Consequently, the tandem efficiency is not strongly affected by a change of R_{ir}.

Even though the tandem efficiency is fairly constant for λ_{irc} up to 750 nm, the reflection loss argument starts to become much more relevant in the long cut-off wavelength region (λ_{irc} > 750 nm), where the perovskite layer becomes transparent. In this region, any value of R_{ir} different from zero results in reflection losses, so it is of paramount importance to identify a minimum requirement for R_{ir}. This requirement can be readily identified through the results of Figure V-3 (a): it is apparent that the tandem efficiency is almost independent of wavelength for R_{ir} = 0.1, thus indicating that the reflection losses should not be higher than 10% in the spectral region where perovskite is transparent.

We have thus gained two key insights into the analysis of the ideal intermediate reflector: 1) The device performance is almost independent on the intermediate structure reflectance in the perovskite absorption window; 2) the intermediate structure should act predominantly as an optical impedance matching layer in the spectral region where perovskite is transparent, with the

requirement that the reflection arising from impedance mismatch be kept below 10%. In the next section, these insights are used to guide the design of the intermediate photonic structures.

V.3. Intermediate Photonic Structure design

As discussed in section V.2, we have now identified that the main role of the photonic intermediate structure in perovskite/c_Si tandem cells should be to provide optical impedance matching between the top and bottom cells. In order to better understand the role of intermediate photonic structures, we now introduce four different designs, shown in Figures V-4 (a)-(d). The structures of Figures V-4 (a)-(b) are designed to act as optical impedance matching layers between the top and bottom cells, whereas the structures of Figures V-4 (c)-(d) are designed as intermediate reflectors. Figure V-4 (a) consists of only a layer of silicon nitride (Si_3N_4) sandwiched between the SiO_2 superstrate and the silicon substrate. The structure of Figures V-4 (b) shows a patterned double layer photonic structure, where corrugated titanium dioxide (TiO_2) layer is combined with a homogeneous layer of the same material and thickness. The intermediate structures shown in Figures V-4 (c) and (d) are Distributed Bragg Reflectors (DBR), comprised of 2 and 4 layers, respectively. The design procedures of each structure will be shown separately in the following sections.

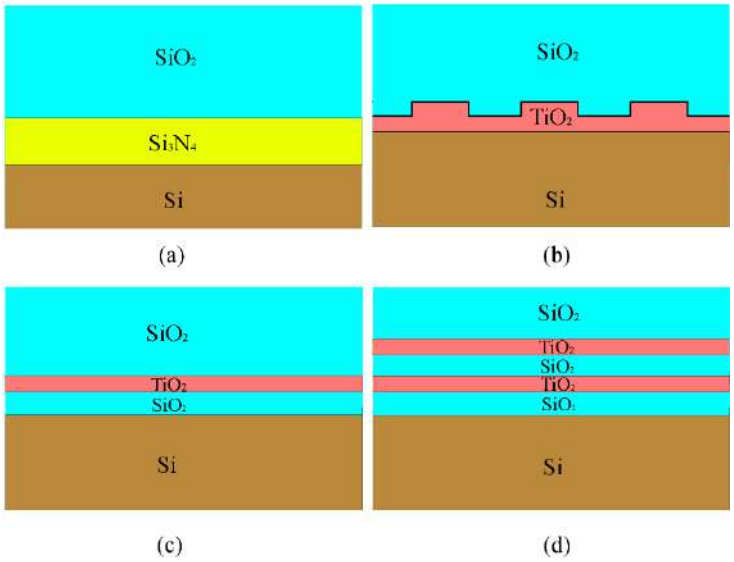


Figure V-4 - (a)-(d) Four different intermediate structure designs. (a), (b) Intermediate structures designed as impedance matching layer: (a) A single 80 nm thick silicon nitride impedance matching layer. (b) double layer structure consisting of a corrugated TiO_2 layer, with a period of 300 nm, followed by a homogenous layer of the same material. There are no propagating diffraction orders and the grating acts only as an effective index layer. (c)-(d) Intermediate structures designed as intermediate reflectors. (c) a 2-layer DBR with parameters fine-tuned to optimize the tandem solar cell performance. (d) a 4-layer DBR with parameters fine-tuned to optimize the tandem solar cell performance.

V.3.1. Homogeneous Si_3N_4 layer

This intermediate structure is shown in Figure V-4 (a) and its design consists only in adjusting the Si_3N_4 thickness, so it has a Fabry-Perot resonance at a desired wavelength λ_{max}. This structure should act as an impedance matching layer to maximize light transmission into silicon. Therefore, light reflected at the SiO_2/Si_3N_4 interface must interfere destructively with light reflected at the Si_3N_4/Si interface, which means that their phase difference must be $m\pi$ rad, where m is an integer number. The phase acquired upon reflection at SiO_2/Si_3N_4 interface is π rad, as the index of reflection of silica is smaller than Si_3N_4. The light that is reflected from Si_3N_4/Si interface, and acquires a π phase shift upon reflection, must travel through the Si_3N_4 layer twice before getting back to the SiO_2 buffer. So, the total phase accumulated by light that gets reflected by the Si_3N_4/Si is

$$\phi = \pi + 2H_{Si_3N_4} n_{Si_3N_4} k_{max} \quad \text{(V-2)}$$

where $2H_{Si_3N_4} n_{Si_3N_4} k_{max}$ is the phase accumulated during propagation through the Si_3N_4 layer twice. $H_{Si_3N_4}$ and $n_{Si_3N_4}$ are height and refractive index of Si_3N_4 layer and $k_{max} = \frac{2\pi}{\lambda_{max}}$ is the wavenumber. Therefore, the phase difference between the portions reflected at the two Si_3N_4 interfaces, and which must equal $m\pi$ rad, is

$$\phi - \pi = 2H_{Si_3N_4} n_{Si_3N_4} k_{max} = m\pi \quad \text{(V-3)}$$

Isolating $H_{Si_3N_4}$ in equation (V-3), it follows that

$$H_{Si_3N_4} = \frac{m\lambda_{max}}{4n_{Si_3N_4}} \quad \text{(V-4)}$$

Equation (V-4) provides a first guess for the thickness of the Si_3N_4 layer acting as an optical impedance matcher. Figure V-5 shows $H_{Si_3N_4}$ as function of the resonance wavelength, λ_{max}, for m=1 (this value of m was chosen because it gives the smaller possible height for the layer).

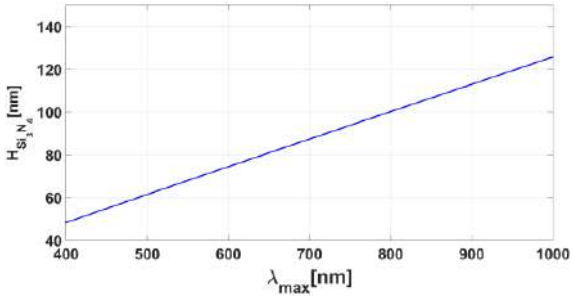


Figure V-5 -Graphic of equation (V-4) for m=1.

The solar cell efficiency of the silicon layer was calculated as function of the Si_3N_4 height to find which value maximizes the efficiency. To this intent, we calculated the absorption in a 400

μm thick silicon layer under the intermediate structure and the SiO_2 buffer and terminated by an ideal mirror on bottom. The calculations were performed using Si_3N_4 layer height in the range [0 nm, 150 nm], that was based on Figure V-5. The range starts at 0 nm to access the efficiency without the intermediate structure. Figure V-6 shows the silicon solar cell efficiency as function of $H_{Si_3N_4}$. Note that it presents a maximum near 80 nm, which we chose to implement in the tandem solar cell. Furthermore, this structure increases the efficiency of the silicon bottom solar cell by approximately 7% in comparison with the case without it.

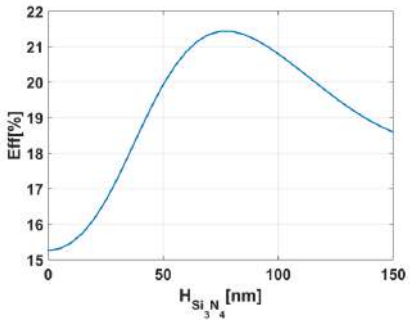


Figure V-6 – Silica solar cell efficiency as function of the intermediate Si3N4 homogeneous layer thickness. In the absorption calculation, the system consists of a silica layer 400 μm height under the intermediate structure and the silicon buffer and terminated by an ideal mirror on bottom. Light comes from the silicon buffer. The calculations assume perpendicular incidence and AM1.5G solar spectrum.

V.3.2. Corrugated TiO_2 layer

The second approach used for impedance matching between silicon and silica layer consists of a TiO_2 grating, height h_g, on top of a homogeneous layer of the same material with height given by h_l, as represented in Figure V-7 (a). The grating period is p, the ridge width is d and, therefore, the fill factor is $ff = \frac{d}{p}100\%$. The transmission of the structure was calculated using RCWA assuming normal incidence from the silica dioxide buffer.

The optimization process starts by analyzing the response of the structure as function of the fill factor of the grating. To this intent, the other parameters were set as

- p = 200 nm
- h_l = 50 nm
- h_g = 50 nm

Figure V-7 (b) shows the calculated transmission spectrum as function of the fill factor. Note that there is a region in which the transmission is over 0.9 in a 300 nm wavelength bandwidth that redshifts when the fill factor is increased. Since the wavelength is much larger than its period, the grating does not present diffractive behaviour for $\lambda_0 > 800$ nm. Consequently, the wave perceives the grating as an effective medium whose index is modulated by the fill factor. In this sense, the response of the system is almost period independent, as can be seen in Figure V-7 (c), where the transmission spectrum is shown as function of the period. The grating and homogeneous layers heights were fixed at 50 *nm* and the fill factor at 60%, because it provides

maximum transmission for $\lambda_0 = 800\ nm$ according to Figure V-7 (b)**.** Diffraction in TiO_2 only starts in the region above the white line and in the buffer only above the black dots in Figure V-7 (c). Nevertheless, it does not impact the high transmission band that starts at $\lambda_0 = 700\ nm$ significantly. The major perturbation is the onset of diffraction (Rayleigh anomaly) in the buffer, that leaves a resonant signature in the spectrum.

Thus, with the presented results it is already possible to choose a design for the final structure and the following parameters will be used:

- ff = 60%, h_l = 50 nm, h_g = 50 nm
- p = 300 nm, because with this period the grating does not present diffractive phenomena in the region with maximum transmission ($\lambda_0 > 800$ nm) and also presents large grooves that do not pose any fabrication challenge.

Finally, it will be analysed the sensibility of the structures with respect to variations in the grating and homogeneous layers heights. For this purpose, Figure V-7 (d) shows the transmission spectra a function of the homogeneous layer height while maintaining the grating height equals 50 *nm*. Figure V-7 (e) shows the same plot but with now with the homogenous layer fixed at 50 *nm* and with the grating height varying. Note that for both cases the transmission maximum redshifts when the height is increased but affecting neither the bandwidth nor the peak.

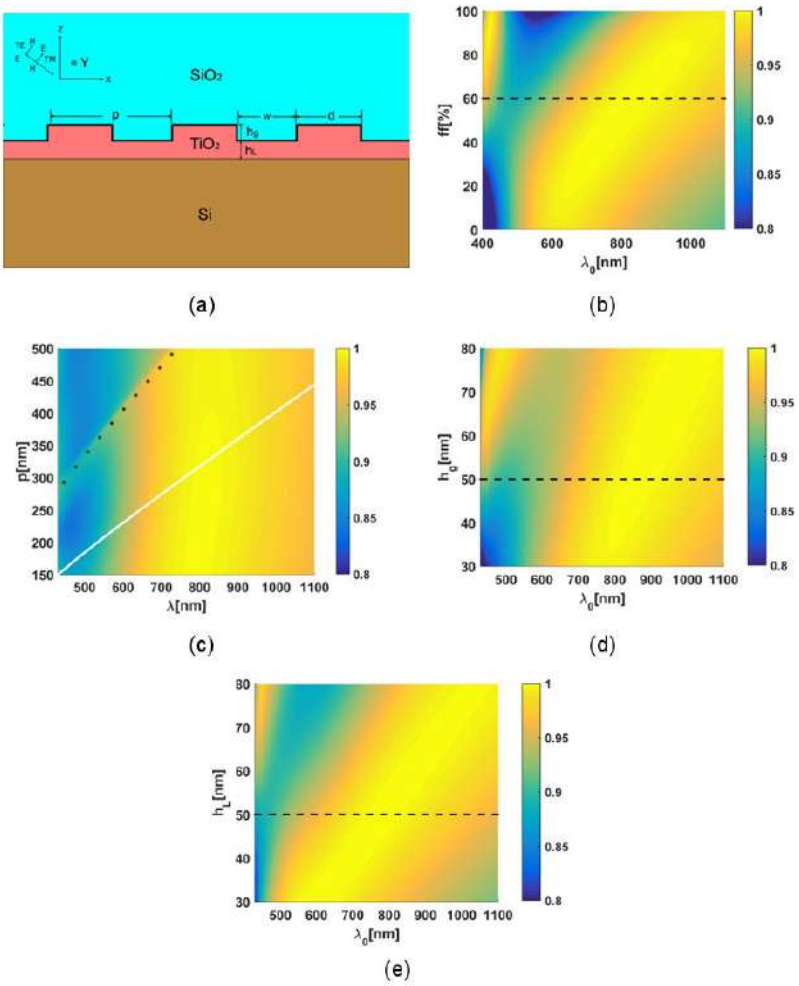


Figure V-7 – (a) Corrugated TiO_2 layer system. Vectors E and H represent the electric and magnetic fields, respectively, for TE and TM polarizations. Unpolarized light is assumed to impinge normally on the structure. (b) Transmission spectrum of the structure as function of fill-factor (ff) with p=200 nm and $h_g = h_l$ = 50 nm. (c) Transmission spectrum of the structure as function of the unit cell size (p), with ff=60%, and $h_g = h_l$ = 50 nm. (d) and (e) Show the transmission spectrum as function of h_g and h_l, respectively. The fill factor was maintained constant for both cases at 60% and h_l [h_g] was fixed in 50 nm in (d) [(e)]. The dashed lines in (b), (d), and (e) mark the spectrum of the final design. The white line in (c) show the border for diffraction on TiO_2 whereas the dotted line the Rayleigh anomaly on the buffer.

V.3.3. Intermediate Reflectors: Bragg reflectors with 2 and 4 layers

To complete the analysis, we also designed two structures to act as intermediate reflectors, as is usually done in the literature. To this intent, distributed Bragg Reflectors (DBR) with two and four layers were chosen. These structures are composed of multiple homogeneous dielectric layers, A and B, disposed alternately. Figures V-8 (a) and (b) illustrate two DBR filters with two and four layers, respectively, where it is assumed that the index of A (n_A) is higher than that of B (n_B) and of SiO_2. Furthermore, the thicknesses of layers with the same material have the same height (h_A and h_B for layers A and B, respectively).

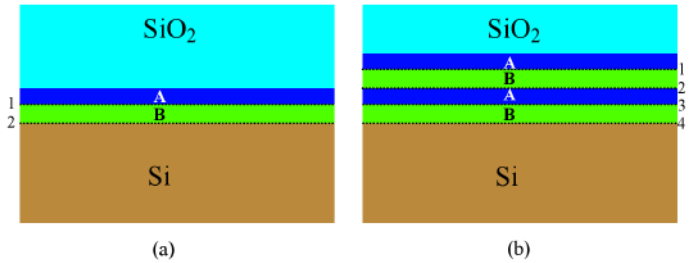


Figure V-8 – DBR filters designed with two (a) and four (layers). The thicknesses of the layers A and B are h_A and h_B, respectively. The numbers are used to enumerate the interfaces between the layers.

The DBR design follows the same line of reasoning made for the homogeneous Si_3N_4 layer design (section *V.3.1 Homogeneous Si3N4 layer*); i.e., it is necessary to guarantee that the reflected wave at each layer interface suffers constructive interference at the SiO_2 buffer. The wave reflected immediately on the (SiO_2/A) acquires phase of π rad. Thus, the waves reflected on the remaining interfaces must return to the buffer with accumulated phase given by $\pi+2\pi m$, $m \in \mathbb{Z}$. The waves reflected at odd interfaces (see the interface numbering at Figure V-8) do not acquire π *rad* as they happen in a medium with higher refractive index. Thus, the phase acquired by the wave returning to the SiO_2 buffer after being reflected by an odd interface are due only to propagation and is given by

$$\phi_l = 2(l-1)h_B k_{max} n_B + 2lh_A k_{max} n_A \,, l\ odd \qquad (V\text{-}5)$$

where l is the interface number and $k_{max} = \frac{2\pi}{\lambda_{max}}$ is the wavenumber for maximum reflection operating at a wavelength λ_{max}.

The waves reflected at even interfaces, in turn, acquire additional π rad because they happen in the medium with smaller refractive index. Thus, the total phase acquired by a wave returning to the SiO_2 buffer after being reflected by an even interface is given by

$$\phi_l = 2lk_{max} n_B + 2lh_A k_{max} n_A + \pi, l\ even \qquad (V\text{-}6)$$

Imposing the constructive interference condition at equations (V-5) and (V-6), and, after some algebraic manipulations, it follows that, respectively,

$$l(n_B h_B + h_A n_A) - n_B h_B = \frac{\lambda_{max}}{2}\left(\frac{1}{2} + m\right) \,, l\ odd, m \in \mathbb{Z} \qquad (V\text{-}7)$$

$$l(n_B h_B + h_A n_A) = \frac{\lambda_{max}}{2} m \,, l\ even, m \in \mathbb{Z} \qquad (V\text{-}8)$$

The operating wavelength λ_{max} is chosen by design and equations (V-7) and (V-8) must hold for all possible l, with m as a free parameter. Thus, the layers heights and refractive indexes must be judiciously chosen to satisfy equations (V-7) and (V-8). One possible way of doing it is by imposing the following restriction

$$n_B h_B = n_A h_A = \frac{\lambda_{max}}{4} \quad \text{(V-9)}$$

Thus, by following equation (V-9), three different DBR filters are proposed with TiO_2 used on layers A for all cases with target at $\lambda_{max} = 550$ nm. For layer B it was analysed the filter performance with alumina (Al_2O_3), SiO_2 and Si_3N_4. The resulting thicknesses of each design are listed on Table V-1.

Table V-1 – Thicknesses of the designed DBR filters following equation (V-9) with target wavelength at $\lambda_{max} = 550$ nm.

	Material	Thickness (nm)
A	TiO2	51
B	Si3N4	68
	Alumina	82
	SiO2	94

The reflectance of the proposed DBRs were calculated using the RCWA [88] for the cases with two and four layers to assess their performances and the results are shown in Figures V-9 (a) and (b), respectively. Both cases with four and two layers present high reflectance at the region where perovskite is absorptive ($\lambda_0 < 800$ nm) and low reflectance for higher wavelengths, where perovskite is transparent, as expected for an intermediate reflector. Nevertheless, the four-layered filters have better performance as its reflectance is higher with a more well-defined bandwidth. Furthermore, silica is the material presenting higher reflectance for both two- and four- layered systems. Thus, the DBRs of SiO_2/TiO_2 will be chosen for application on the tandem solar cell.

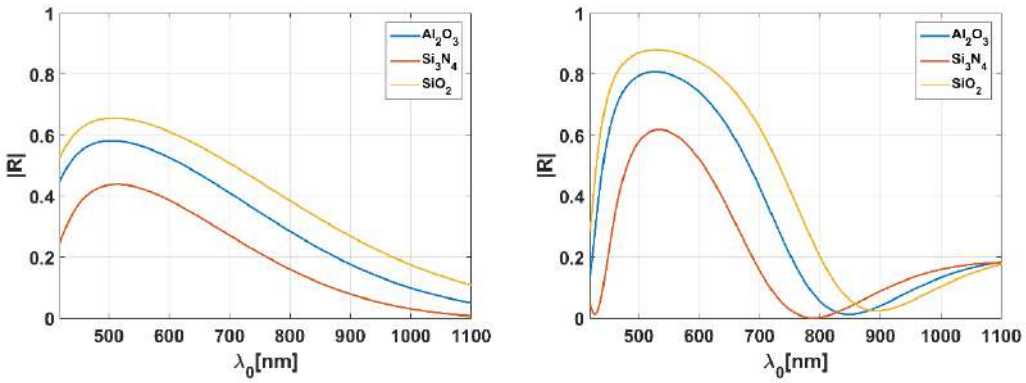


Figure V-9 – Reflectances of DBR filters with (a) two and (b) four layers. The layers thicknesses are listed on Table V-1. The high index material is TiO2 for all DBRs.

V.4. Application of intermediate structures on Si/Perovskite tandem solar cells

V.4.1. Simplified structure characterization

In this section the performance of the tandem solar cell in the presence of the designed intermediate structures will be evaluated. The final structures are shown in Figures V-10 (a)-(d). Figure V-10 (a) consists of only a single 80 nm thick layer of silicon nitride (Si_3N_4) sandwiched between the SiO_2 superstrate and the silicon substrate. The structure of Figure V-10 (b) shows a patterned double layer photonic structure, where a 50 nm thick corrugated titanium dioxide (TiO_2) layer is combined with a homogeneous layer of the same material and thickness. The period of the TiO_2 grating is 300 nm and the fill factor is 60%. Notice that the small grating period ensures that the grating acts only as an effective medium and that there are no propagating diffraction orders. As shown in Figure V-10 (e), the transmittance of the double layer structure is higher than the transmittance of the single layer structure, but the difference is not large: both structures provide broad-band high transmittance (> 90%), spanning the spectral range between 500 and 1100 nm.

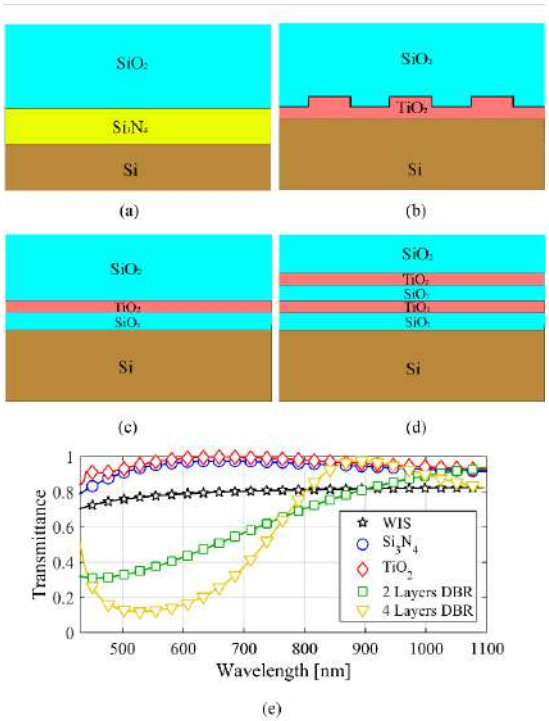


Figure V-10 - (a)-(d) Four different intermediate structure designs. (a), (b) Intermediate structures designed as impedance matching layer: (a) A single 80 nm thick silicon nitride impedance matching layer. (b) double layer structure consisting of a corrugated TiO_2 layer, with a period of 300 nm, followed by a homogenous layer of the same material. There are no propagating diffraction orders and the grating acts only as an effective index layer. (c)-(d) Intermediate structures designed as intermediate reflectors. (c) a 2-layer DBR with parameters fine-tuned to optimize the tandem solar cell performance. d) a 4-layer DBR with parameters fine-tuned to optimize the tandem solar cell performance. (e) Transmittance of all four intermediate photonic structures. The transmittance without any intermediate structure (WIS) is also shown for comparison. All transmittances are between the silicon dioxide superstrate and the silicon substrate.

The intermediate structures shown in Figures V-10 (c) and (d) are Distributed Bragg Reflectors (DBR), comprised of 2 and 4 layers, respectively. As described in section *V.3.3 Intermediate Reflectors: Bragg reflectors with 2 and 4 layers,* the DBRs are first optimized to provide high reflectance into the perovskite absorption spectral window and, at the same time, low reflectance in the perovskite transparency window. The optimised DBR final thicknesses are 50nm for the TiO_2 layers and 95nm for the SiO_2 layers.

Figures V-11 (a)-(d) show the tandem solar cells with the corresponding intermediate photonic structures, whereas Figures V-11 (e)-(f) show the absorption in the perovskite and silicon layers, respectively. As expected, the perovskite absorption is higher in the structures with intermediate reflectors, as shown in green squares and yellow triangles in Figures V-11 (e). The perovskite absorption spectra for the intermediate reflectors also show more pronounced Fabry-Perot oscillations when compared to the optical impedance matching layers. We use the integrated absorption as a figure of merit to assess the performance of the structures. The integrated absorption takes into account the photon density in the solar spectrum and is defined as the total amount of absorbed solar photons divided by the total amount of incoming solar photons.

The integrated absorptions in the perovskite layer for the solar cells of Figures V-11 (a)-(d) are, respectively: 46.2% (optical impedance matching), 46.2% (optical impedance matching), 49.3% (DBR), 50.2% (DBR); while the integrated absorption in the perovskite layer without intermediate structure (WIS) is 47.4%. The absorption in the silicon layer shown in Figure V-11 (f), on the other hand, is higher for the impedance matched structures of Figures V-11 (a)-(b).

The integrated absorptions in the silicon layer for the structures of Figures V-11 (a)-(d) are, respectively: 41.9% (optical impedance matching), 43.5% (optical impedance matching), 34.6% (DBR), 34.8% (DBR); while the integrated absorption in the silicon layer without intermediate structure (WIS) is 36.85%. The calculated efficiencies and short circuit currents are shown in Table V-2 for two different values of the perovskite charge carrier diffusion length: $L_d = 100$ and $L_d = 400\ nm$.

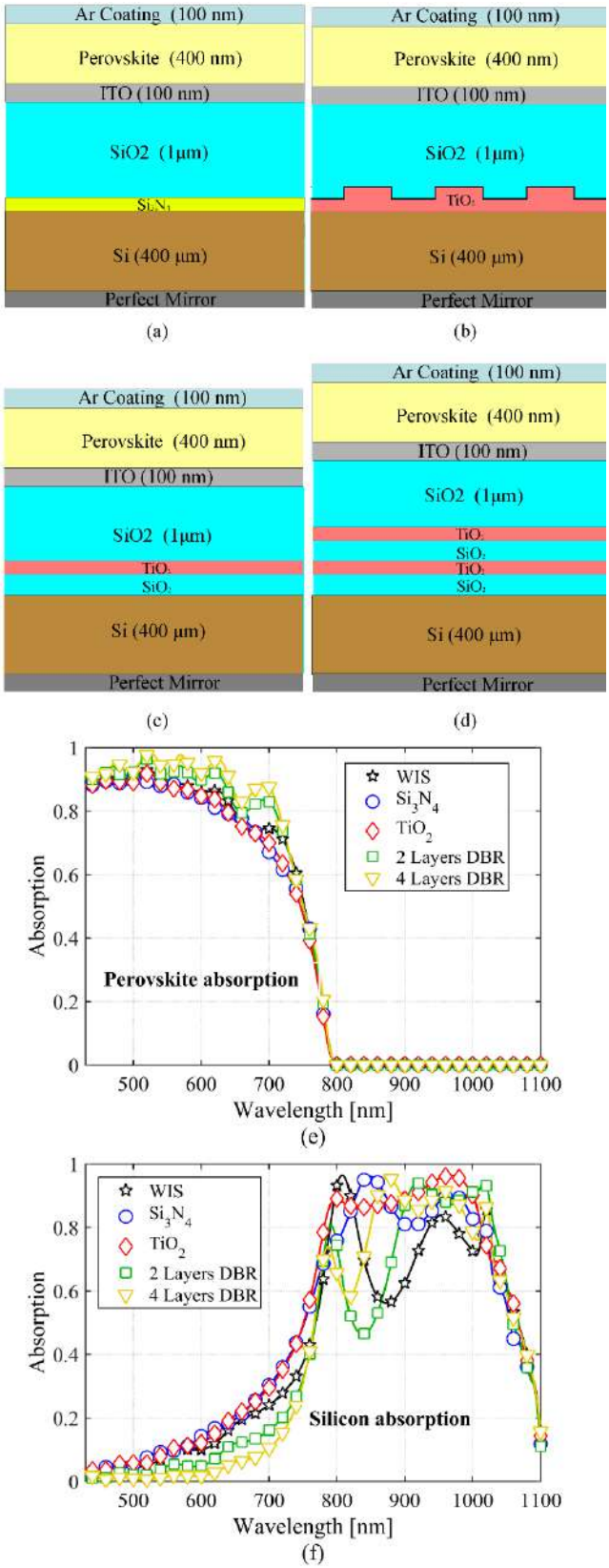


Figure V-11 - Illustration of the complete solar cell with the intermediate photonic structure. (a)-(b) Photonic intermediate structures designed as impedance matching layers. (c)-(d) Photonic intermediate structures designed as intermediate reflectors. (e) Absorption in the perovskite layer. (f) Absorption in the silicon layer.

Table V-2- Solar cell parameters for L_D=100 nm and for L_D=400 nm.

	L_d =100 nm					L_d =400 nm				
	Impedance matching layers		Intermediate Reflectors		Without intermediate structure	Impedance matching layers		Intermediate Reflectors		Without intermediate structure
	Si_3N_4	TiO_2	2 DBR	4 DBR		Si_3N_4	TiO_2	2 DBR	4 DBR	
Silicon J_{sc} (mA/cm^2)	17.81	18.50	14.74	14.81	15.68	17.81	18.50	14.74	14.81	15.68
Perovskite J_{sc} (mA/cm^2)	18.79	18.79	20.05	20.57	19.29	19.98	19.98	21.32	21.87	20.50
Silicon Efficiency (%)	10.08	10.49	8.29	8.32	8.83	10.08	10.49	8.29	8.32	8.83
Perovskite Efficiency (%)	19.01	19.03	20.27	20.79	19.52	20.23	20.26	21.57	22.13	20.78
Tandem Efficiency (%)	29.09	29.52	28.56	29.11	28.35	30.31	30.74	29.86	30.45	29.61

The 4-layer DBR shows high transmittance for wavelengths larger than 800 nm (Figure V-10 (e)), which results in high absorption in the silicon layer for these wavelengths (Figure V-11 (f)). The 4-layer DBR, therefore, is acting as both an intermediate reflector for the perovskite absorption window, and also as an optical impedance matching layer for wavelengths larger than 800 nm (i.e., the perovskite transparency window).

This double role of the DBR structure might justify the assumption that the 4-layer DBR is the optimum structure. As shown in Table V-2, however, the tandem solar cell efficiency is not highest for the 4-layer DBR. Indeed, the best overall efficiency is achieved for the optical impedance matching structure comprised of the TiO_2 grating (Figure V-11 (b)). It is important to notice that, even when compared to the simplest optical impedance matching layer (Figure V-11 (a)), the tandem efficiency of the 4-layer DBR is only marginally higher. These counter-intuitive results can be understood by comparing Figure V-10 (e) with Figure V-11 (f): even though the transmittance for the 4-layer DBR rises quickly between 700 and 800 nm, this transition is not sharp enough to mitigate reflection losses in the silicon layer. Consequently, the absorption in the silicon layer (Figure V-11 (f)) in the short-wavelength region (< 850 nm) is higher for the structures that provide only optical impedance matching. It is this difference that accounts for the similar performances of the tandem solar cells, listed in Table V-2. These results, therefore, corroborate the previous conclusion that it is preferable that intermediate structures be designed to provide only optical impedance matching between the top and bottom cells. Indeed, the DBR

intermediate structures are more complex and yet show similar performance when compared to the simplest optical impedance matching layer.

Our calculations indicate that both optical impedance matching layers result in efficiencies exceeding 30% for L_d = 400 nm. Naturally, we expect that the tandem device performance can be further increased by applying light trapping concepts [146, 150] and that the design of such light trapping structures will benefit from the optical impedance matching principles we have outlined here. When compared to the solar cell without intermediate structure, the silicon solar cell short circuit current is increased by 18.5% for the TiO_2 grating. It is important to notice that these high efficiencies were achieved without any light trapping scheme, which is expected to boost the efficiencies even further [143, 150].

Resonant structures depend on specific phase accumulation requirements that tend to limit their angular performance. One additional advantage of employing optical impedance matching layers is that their non-resonant behavior results in a very angular tolerant performance. This feature can be seen in Figures V-12 (a)-(b), which show, respectively, the absorption in the silicon and perovskite layers versus incidence angle (θ) and wavelength for the system with the TiO_2 grating acting as optical impedance matching layer. The dependence of the solar cell efficiencies on the incidence angle is shown in Figure V-12 (c). Impressively, the efficiencies are virtually constant up to an angle of 60 degrees, implying an acceptance cone of at least 120 degrees.

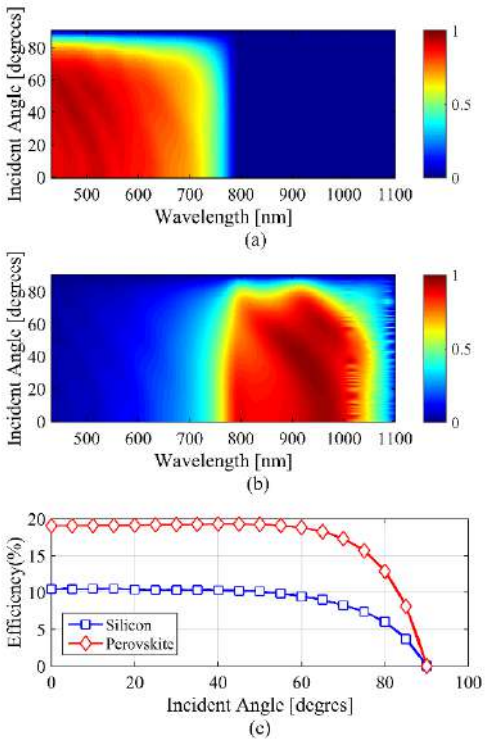


Figure V-12 - Angular dependence of the solar cell performance. (a) and (b) show the absorption spectra of the perovskite and c-Si, respectively, as a function of incident angle and wavelength. (c) Efficiency of the silicon and perovskite solar cells as a function of the incident angle.

V.4.2. Complete device characterization

Additional simulations were performed to characterize the solar cell performance by taking into account a front contact and also transport layers. Therefore, a front ITO contact and also charge transport layers were included into the model, taking their real and imaginary indexes into account (refractive indexes shown in Figure V-2). We find that while the performance is somewhat reduced, the physical principles and conclusions from the previous section remain unchanged.

In order to support these arguments, the complete structure is shown in Figure V-13 and the absorptions for two intermediate structures in both simplified and complete devices in Figure V-14. The simplified device refers to the device used in the previous section (Figures V-10 (b) and (d)). The label TiO_2 refers to the optical impedance matching layer comprised of a TiO_2 grating and the 4 DBR label refers to the intermediate reflector. As expected, the absorption in both perovskite and silicon layers are lower in the complete structure (Figure V-14), mostly due to parasitic absorption in the front ITO contact. However, the effects of parasitic absorption are similar in both impedance matching (TiO2) and intermediate reflectors (4 DBR). For example, the absorption in the silicon layer around 950 nm (Figure V-14 (b)) exhibits a gap in the complete device for both optical impedance matching (TiO_2) and intermediate reflector (4 DBR), when compared to the simplified cell (i.e., without front ITO and charge transport layers). Such a gap is not due to an increase in the material parasitic absorption, but due to thin-film interference in the top solar cell and its effects on the devices with both classes of photonic intermediate structures are very similar.

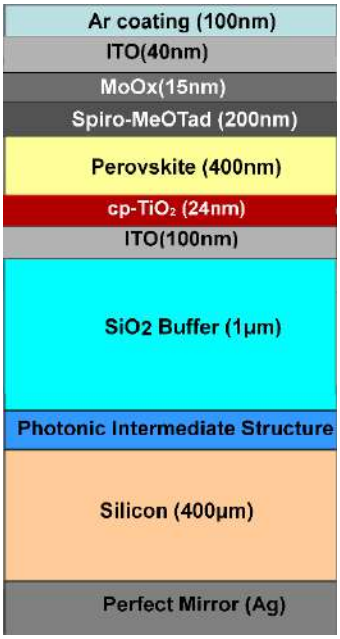


Figure V-13 – Complete device structure

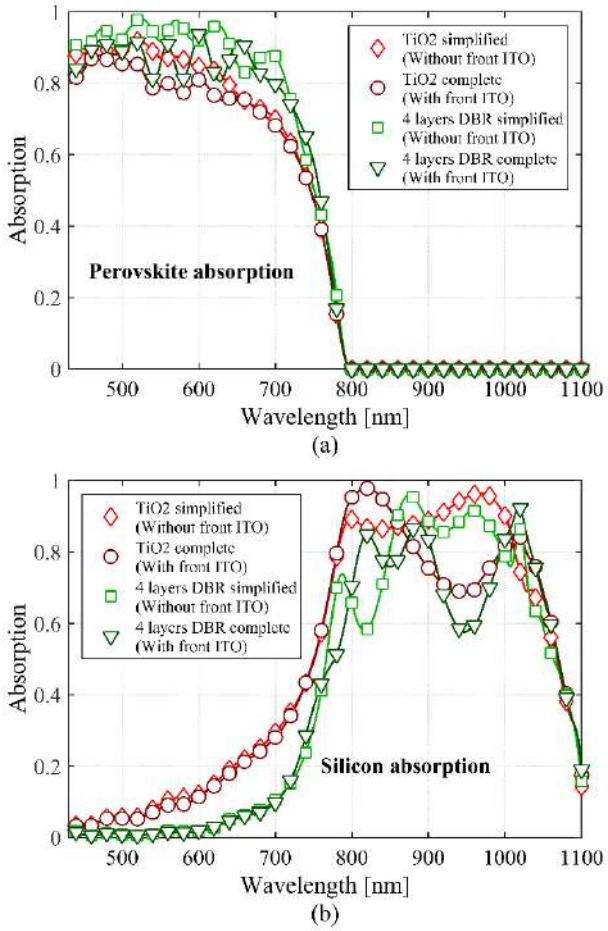


Figure V-14 - Absorptions in the perovskite (a) and silicon (b) layers for two classes of photonic intermediate structures: a TiO2 grating acting as an impedance matching layer and a 4 layers DBR acting as an intermediate reflector. The simplified structures do not include the front ITO and transport layers. The complete structure includes the front ITO and transport layers. A comparison between the absorptions in the simplified and complete devices shows that the absorption penalty is very similar for both classes of photonic intermediate structures.

A comparison between the parameters of the solar cells can be found in Table V-3. In all cases, the parasitic absorption reduced the efficiencies of the cells. For example, if we compare the performance of the device using the TiO2 grating (optical impedance matching layer) with the device using 4 DBR (intermediate reflector) for L_d = 400 nm, we find that changing from the simplified to the complete device lowers the silicon short circuit current from 18.5 mA/cm^2 to 17.96 mA/cm^2 for the device with TiO_2 grating and from 14.81 mA/cm^2 to 14.39 mA/cm^2 for the device with 4 DBR. These reductions correspond to a factor of 0.971 for the TiO_2 device and 0.972 for the 4 DBR device: in other words, the silicon short circuit current penalty is virtually the same in both devices. This is also true for the perovskite short circuit current penalty: the current in the TiO2 device is reduced by a factor of 0.89 while the current in the 4 DBR is reduced by a factor of 0.9. Therefore, the efficiencies of both devices are reduced by very similar amounts: the efficiency of the TiO2 devices goes from 30.74% to 28.12% (a reduction of 2.62%) and the efficiency of the 4 DBR goes from 30.45% to 27.91% (a reduction of 2.54%). In conclusion, the message of this work is unaltered: the optical impedance matching layers perform equally or even

slightly better than the intermediate reflectors. It is important to emphasize that, since the main conclusion of this work does not depend on the particular choice of device, the simplified structure can convey our message more straightforwardly.

Table V-3- Solar cell parameters for L_D=100 nm and for L_D=400 nm. Comparison between the performances of the simplified and complete devices. The simplified structure does not include the front ITO and transport layers whereas the complete structure does.

	L_d=100 nm Simplified | *Complete*				
	Impedance Matching layers		Intermediate Reflectors		Without intermediate structure
	Si_3N_4	TiO_2	2 Layer DBR	4 Layer DBR	
Silicon J_{sc} (mA/cm^2)	17.81|*17.57*	18.50|*17.96*	14.74|*14.01*	14.81|*14.39*	15.68|*15.72*
Perovskite J_{sc} (mA/cm^2)	18.79|*16.67*	18.79|*16.70*	20.05|*17.95*	20.57|*18.46*	19.29|*17.18*
Silicon Efficiency (%)	10.08|*9.94*	10.49|*10.17*	8.29|*7.86*	8.32|*8.08*	8.83|*8.86*
Perovskite Efficiency (%)	19.01|*16.82*	19.03|*16.87*	20.27|*18.11*	20.79|*18.63*	19.52|*17.32*
Tandem Efficiency (%)	29.09|*26.76*	29.52|*27.04*	28.56|*25.97*	29.11|*26.71*	28.35|*26.18*
	L_d=400 nm Simplified | *Complete*				
	Impedance Matching layers		Intermediate Reflectors		Without intermediate structure
	Si_3N_4	TiO_2	2 Layer DBR	4 Layer DBR	
Silicon J_{sc} (mA/cm^2)	17.81|*17.57*	18.50|*17.96*	14.74|*14.01*	14.81|*14.39*	15.68|*15.72*
Perovskite J_{sc} (mA/cm^2)	19.98|*17.72*	19.98|*17.75*	21.32|*19.08*	21.87|*19.62*	20.50|*18.26*
Silicon Efficiency (%)	10.08|*9.94*	10.49|*10.17*	8.29|*7.86*	8.32|*8.08*	8.83|*8.86*
Perovskite Efficiency (%)	20.23|*17.90*	20.26|*17.95*	21.57|*19.28*	22.13|*19.83*	20.78|*18.44*
Tandem Efficiency (%)	30.31|*27.84*	30.74|*28.12*	29.86|*27.14*	30.45|*27.91*	29.61|*27.30*

V.5. Conclusions and future work

It was identified that a photonic intermediate structure in a perovskite/c-Si tandem solar cell should act as an optical impedance matching layer at the perovskite-silicon interface. The reason for this somewhat unexpected behavior is that by increasing the reflectivity, the reflection loss back into free space tends to outweigh the improvement in absorption in the top layer. This insight affords the relaxation of the photonic structure reflectance in the perovskite absorption window, which leads to very simple and robust designs for the intermediate structure. Accordingly, two simple designs were analyzed, and their performances compared with DBR based intermediate reflectors. The conclusion is that intermediate structures acting only as optical impedance matching layers are much simpler than the DBR structures yet showing similar performances. This insight was implemented by simulating a realistic device configuration and show that optical impedance matching alone can increase the short circuit current of the silicon solar cell by 18.5% (corresponding to a boost of 2.8 mA/cm^2), thus resulting in an expected tandem efficiency in excess of 30%.

The obvious next step for a future work is to combine these insights with light trapping structures, acting to scatter light on either (or both) perovskite and silicon substrate.

List of Figures

Main text

Appendices

List of Tables

Main text

Appendices

List of acronyms and abbreviations

a-Si – Amorphous Silicon
CGH - Computer Generated Hologram
c-Si – Crystalline Silicon
DBR – Distributed Bragg Reflectors
DE – Diffraction Efficiency
DFT – Discrete Fourier Transform
DL – Diffraction limited
DOE - Diffractive Optical Element
DRIE – Deep reactive ion etching
EBL – Electron beam lithography
FFT – Fast Fourier Transform
FOV – Field of view
FWHM – Full width at half maximum
HCG – High Contrast Grating
ICP – Inductively coupled plasma
IE – Image Efficiency
IFFT – Inverse Fast Fourier Transform
IFTA – Iterative Fourier Transform Algorithm
LSPR - localized surface plasmon resonances
NA – Numerical aperture
PB – Pancharatnam Berry
PSF – Point spread function
p-Si – Polycrystalline Silicon
RCWA – Rigorous Coupled Wave Analysis
RS – Rayleigh Sommerfeld
SNR – Signal to Noise Ratio
SOI – Silicon on Insulator
SPP – Surface Plasmon Polariton
UV – Ultra Violet
VCSEL – Vertical-Cavity Surface-Emitting Lasers
WFOV – Wide field of view

Appendices

Appendix A – Optical Plasmonic Metasurface Hologram

A.1. Introduction

In this chapter, it is described the design and measurement of a plasmonic holographic metasurface operating in the optical regime. The metasurface is composed of an array of complementary aluminum nanorods on a silica substrate. Nanorods antennas are structures that may present two localized surface plasmon resonances (LSPRs) at optical frequencies: one that arises from a longitudinal plasmon mode and the other from a transverse plasmon mode [17]. When excited at the resonance frequency, the scattering cross-section of the structure is enhanced, which may be useful for tailoring the properties of light. In this project, this characteristic is used to enhance the cross-polarization efficiency of the transmitted light, that is, the amount of energy that is scattered with polarization orthogonal to the incident field polarization. The reason for that is because two complementary nanorod antennas that are 90° rotated with respect to each other scatters cross-polarized light with a phase difference of π *rads* [80]. Therefore, they can be used as a binary phase wavefront modulator that works with cross polarization. This concept is then used to design a dual color holographic plasmonic metasurface.

This section starts with the design of the nanoantennas. In A.2.1, it is analyzed the theoretical limit of the cross-polarization efficiency and the binary phase control of the complementary nanorod structure. Then, in A.2.2, numerical simulations are used to optimize the efficiency of the structure operating at 444 nm or 635 nm. In A.2.3, a dual color metasurface arrangement is proposed. Micrographs of the fabricated metasurfaces are shown in A.3. Finally, A.4 shows the optical reconstructions and it is discussed the results. This section closes at A.5 where the main conclusions of this work are exposed.

A.2. Metasurface Design

A.2.1. Theoretical Analysis of the Plasmonic Metasurface Design

The binary phase modulation is performed with complementary nanorod antennas, as shown in Figure A.1 (a). The nanorod shown in Figure A.1 (b) could also be used to achieve binary phase control but the complementary structure was used because it reduces the co-polarized light transmission, which increases the signal to noise ratio of the hologram [43]. A pair of those nanoantennas with angular orientations β_1 and β_2 differing by $\frac{\pi}{2}$, as shown in Figure A.1 (c), scatter the cross polarized light with relative phase difference of π [80].

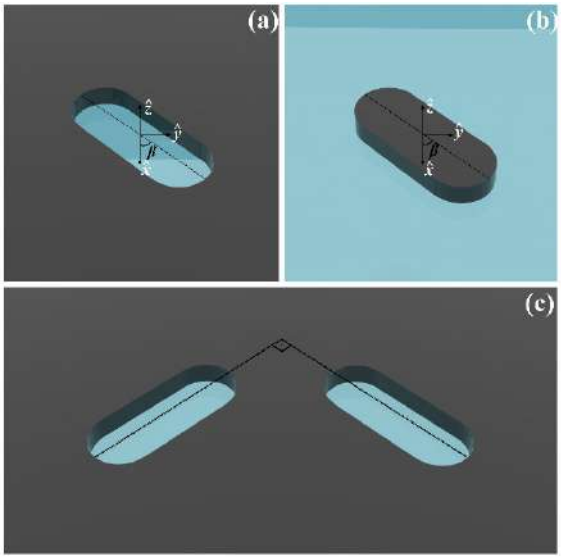


Figure A.1 - Representation of (a) a complementary and (b) a regular nanorod antennas oriented at an angle β with respect the *x* axis. (c) A pair of complementary nanorods with angular orientations differing by $\frac{\pi}{2}$ rad. Gray parts represent the metal portion of the structure and cyan the substrate.

The binary phase control of these structures can be understood by considering the Jones matrix formulation. A nanorod may be viewed as a birefringent layer with transmittance functions T_L and T_T for light polarized along the major and minor length directions of the structure, respectively. The subscripts *L* and *T* stand for the excited plasmon modes longitudinal and transversal, respectively, as they rule the antennas scattering along these directions. For convenience, it will be supposed that the major length of the antenna is parallel to the x axis while the minor to the *y* axis. If the nanoantenna is rotated by an angle β, the resulting transmitted field by the antenna will be given by [166]

$$\begin{pmatrix} A_x \\ A_y \end{pmatrix} = \begin{pmatrix} \cos(\beta) & -\sin(\beta) \\ \sin(\beta) & \cos(\beta) \end{pmatrix} \begin{pmatrix} T_L & 0 \\ 0 & T_T \end{pmatrix} \begin{pmatrix} \cos(\beta) & \sin(\beta) \\ -\sin(\beta) & \cos(\beta) \end{pmatrix} \begin{pmatrix} I_x \\ I_y \end{pmatrix} \quad \textbf{(A.1)}$$

where, A_x and A_y are the transmitted field amplitudes polarized along $\hat{x}$ and $\hat{y}$, respectively; Ix and Iy are the incident field amplitudes polarized along $\hat{x}$ and $\hat{y}$, respectively. Developing (A.1), and considering the incident field polarized along $\hat{x}$, that is, $I_y = 0$, follows that

$$\begin{pmatrix} T_{xx} \\ T_{yx} \end{pmatrix} = \begin{pmatrix} \cos^2(\beta)\, T_L + \sin^2(\beta)\, T_T \\ \frac{1}{2}\sin(2\beta)\,(T_L - T_T) \end{pmatrix} \quad \textbf{(A.2)}$$

Where, $T_{xx} \equiv \frac{A_x}{I_x}$ and $T_{yx} \equiv \frac{A_y}{I_x}$. From equation (A.2) it is possible to infer that the cross polarized transmission, T_{yx}, is maximum, in terms of β, when $\sin(2\beta)$ is at maximum, that is for $\beta_1 = \frac{\pi}{4}$ and $\beta_2 = \frac{3\pi}{4}$ (note that the $\beta_2 - \beta_1 = \frac{\pi}{2}$) . In other words, the nanoantenna must be oriented at an angle of $\pm\frac{\pi}{4}$ relative to the incident polarization. Furthermore, if the antenna is rotated by $\frac{\pi}{2}$ with respect to an initial rotation of β, the cross polarized transmission, from (A.2), reduces to:

$$T_{yx} = \frac{1}{2}\sin\left(2\left(\beta + \frac{\pi}{2}\right)\right)(T_L - T_T) = -\frac{1}{2}\sin(2\beta)(T_L - T_T) \quad \textbf{(A.3)}$$

Therefore, rotating the antenna by π/2 causes a delay of π *rad* in the cross-polarized field.

The cross-polarization energy conversion, $\mathcal{E}$, is defined as the ratio between the transmitted cross-polarized energy flux, $\mathcal{T}_{yx}$, and incident energy flux, $\mathcal{I}_x$, that is

$$\mathcal{E} = \frac{\mathcal{T}_{yx}}{\mathcal{I}_x} \quad \textbf{(A.4)}$$

Assuming that the transmitted and incident waves are plane waves and the antennas are arranged in a subwavelength array, then the power fluxes can be calculated by the plane wave Poynting Vector amplitude [8], that is

$$\mathcal{T}_{yx} = \frac{|A_y|^2}{2n_s\eta_0} = \frac{1}{8n_s\eta_0}\sin^2(2\beta)\,|T_L - T_T|^2|I_x|^2 \quad \textbf{(A.5)}$$

and

$$\mathcal{I}_x = \frac{|I_x|^2}{2\eta_0} \tag{A.6}$$

where, n_s is the substrate refractive index and η_0 is the vacuum impedance. Using (A.5) and (A.6) in (A.4) follows that

$$\mathcal{E} = \frac{1}{4n_s}\sin^2(2\beta)\,|T_L - T_T|^2 \tag{A.7}$$

Therefore, the cross-polarization energy conversion depends on the difference between the transmission of light that is scattered by the longitudinal and transverse plasmon modes of the antenna. The transmission parameters, T_L and T_T, and the cross-polarization energy conversion will be numerically calculated by using the commercial software HFSS on next section. Nevertheless, it is possible to infer the maximum conversion of energy from equation (A.7). For the complementary nanorod, the longitudinal and transverse LSPRs appears as peaks in the dispersion of T_L and T_T [167]. As the transverse resonance is at smaller wavelengths than the longitudinal one, we can assume that $|T_L| > |T_T|$ at the longitudinal LSPR condition. Therefore, at the longitudinal LSPR we can approximate (A.7) by

$$\mathcal{E} \cong \frac{1}{4n_s}\sin^2(2\beta)\,|T_L|^2 \tag{A.8}$$

Furthermore, the power flux of the radiated field by the longitudinal mode is given by

$$\mathcal{I}_L = \frac{|T_L|^2}{2n_s\eta_0} \tag{A.9}$$

From conservation of energy it follows that

$$\mathcal{I}_L \leq \mathcal{I}_x \tag{A.10}$$

Thus, substituting (A.6) and (A.9) in (A.10) and, after some algebraic manipulations, it is concluded that

$$|T_L|^2 \leq n_s \tag{A.11}$$

Using this fact on (A.8), it follows that

$$\mathcal{E} \leq \frac{1}{4}\sin^2(2\beta) \tag{A.12}$$

If the antenna is oriented so that polarization conversion is at maximum, that is with $\beta = \frac{\pi}{4}$ or $\beta = \frac{3\pi}{4}$, from (A.12) follows that the energy conversion of the proposed design is limited by

$$\mathcal{E} \leq 0.25 \tag{A.13}$$

This limitation is in good agreement with the numerical results that will be shown in the next section.

A.2.2. Numerical simulations

The design goal is to find complementary nanorod geometries that present high polarization $\mathcal{E}$ operating at 444 nm and 635 nm. As discussed previously, the nanorods are made of aluminium on a silica substrate. The calculations were performed assuming a constant refractive index for the silica equals to n_s=1.45 and the dispersion of aluminium was taken from Palik [92]. The thickness of the aluminium was fixed at $H = 50$ nm and the nanorod orientation was fixed at $\beta = \frac{\pi}{4}$. As discussed in the previous section, the conversion efficiency will be mostly dependent on the scattering of the longitudinal LSPR, according to equation (A.7). Therefore, before starting the numerical simulations, an estimation of the required complementary nanorod lengths and widths so the LSPR is at 444 nm for one set and at 635 nm for the other will be performed. Note that the LSPR of the nanorod and its complementary counterpart are almost in tune [167]. Therefore, the estimation of the LSPR of the nanorods will be performed assuming that it behaves as the first Fabry-Perot resonance of a truncated aluminium ridge waveguide on a silica substrate with length L and terminated by ideal mirrors. It follows that the resonance condition is

$$\Re[k_g]L = \pi \tag{A.14}$$

where, $k_g = \frac{2\pi}{\lambda_0} n_{eff}$ is the propagation constant of the fundamental mode of the ridge with λ_0 being the operating wavelength and $n_{eff} = n'_{eff} - jn''_{eff}$ is the effective complex index of refraction of the mode. From (A.14), it follows that

$$L = \frac{\lambda_0}{2n'_{eff}} \tag{A.15}$$

The estimation of (A.15) does not take into account the round curvature of the termination of the nanorod and the retardation effects on the extremities, since the field, at optical frequencies, extends beyond the cavity [168]. This result will be only used as a first guess in the full numerical simulation. The effective index of the ridge was calculated using the commercial software COMSOL and the results for each wavelength are listed in Table A-1. The simulations were done using the mode analysis module in COMSOL, which calculates the propagation constant of a waveguide. As a first approximation, the effective index was calculated assuming an infinite rod with constant cross section. Thus, in COMSOL, only the waveguide cross section is necessary for the propagation constant (effective index) calculation. A representation of the simulation region used is shown in Figure A.2.

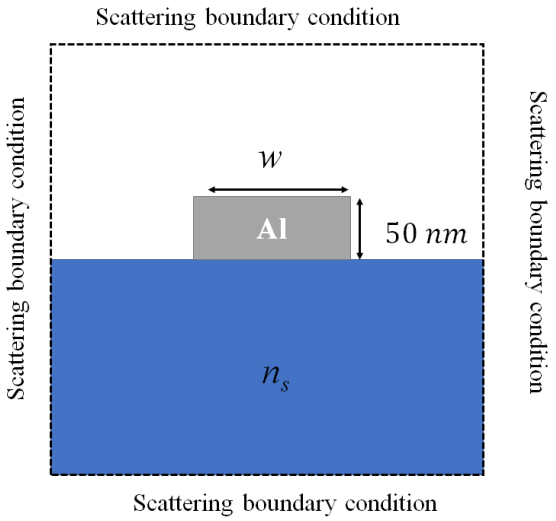


Figure A.2 –Geometry used for the calculation of the nanorod effective index in the mode analysis module of Comsol. Note that, since the mode is assumed to be confined around the Al region, the simulation region must be large enough so the evanescent fields do not touch the outer boundaries.

The calculation was performed with widths *w* ranging from 80 nm to 150 *nm*. Since, as shown in Table A-1, the effective index does not change significantly as a function of *w*, using (A.15), one finds that the resonance occurs for $L_B \cong 149$ nm , at 444 nm, and $L_R \cong 218$ nm, at 635 nm. To minimize the overlapping between the transverse and longitudinal plasmon modes, the width should be smaller than the length. Thus, the width will be fixed at 100 nm, for the design operating at 444 nm, and 150 nm, for the design operating at 635 nm.

Table A-1 - Effective index (n_{eff}) of the aluminum waveguide ridge (50 *nm* tall) over a silica substrate for different widths at two wavelengths (444 *nm* and 635 nm).

w (nm)	n_{eff} *at* $\lambda_0 = 444$ nm	n_{eff} *at* $\lambda_0 = 635$ nm
80	1.512-0.038j	1.457-0.036j
90	1.509-0.035j	1.453-0.032j
100	1.507-0.034j	1.453-0.030j
110	1.506-0.032j	1.453-0.029j
120	1.505-0.031j	1.452-0.027j
130	1.504-0.030j	1.453-0.026j
140	1.504-0.0291j	1.453-0.025j
150	1.504-0.028j	1.450-0.025j

With the initial design dimensions determined, we can start the rigorous simulation of the complementary nanoantenna to determine the geometry that optimizes the cross-polarization energy conversion for the required wavelengths. The simulation was performed on the commercial software HFSS using Bloch boundaries on the lateral walls to simulate an infinite array of complementary nanorods and the system was illuminated by a normal incident plane wave. The unit cell size was taken as a square with dimensions of $\sqrt{2}L \times \sqrt{2}L$ to minimize the coupling between the antennas of the array. The simulated unit cell is shown at Figure A.3 (a) and the resulting cross-polarization energy conversion maps are shown at Figure A.3 (b) and (c) as function of the length, L, and operating wavelength, λ_0, for widths of 100 nm and 150 nm, respectively. Note that the resonances of each case appear near the proposed wavelengths but deviate from the obtained guesses L_R and L_B, as it was expected. In other words, according to the white dotted lines in Figure A.3 (b) [(c)], the resonance is at 444 nm [635 nm] when L= 165 nm [L = 255 nm], which is only 15 nm [32 nm] bigger than the L_B[L_R] predicted by the Fabry-Perot model. Finally, note that the length for maximum energy conversion depends linearly on the wavelength for both widths, as predicted by the Fabry-Perot cavity model in equation (A.15). Finally, Figure A.3 (d) shows the cross-polarization energy conversion dispersion for the optimum structures resonating at 444 nm (blue line) and 635 nm (red line). Note that the maximum conversion for both structures is almost 0.12, which is half the limit imposed in (A.13). This difference arises due to Ohmic losses in the aluminum and the interferences between the longitudinal and transverse plasmon modes scattering that are not considered in (A.13).

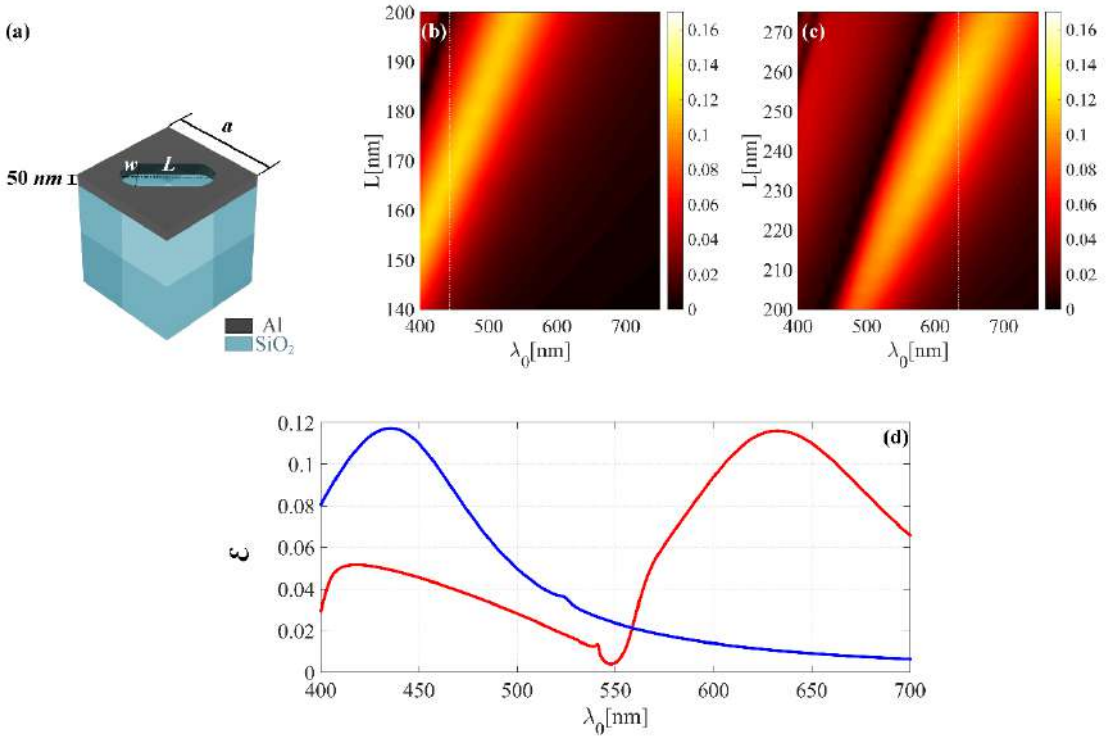


Figure A.3 - (a) Representation of the unit cell containing the complementary nanorod. The incidence is normal from the cover, above the metallic layer. The period was taken as $\sqrt{2}\,L$. The resulting cross-polarization energy conversion is shown at (b) and (c) for widths (w) equal 100 nm and 150 nm, respectively, as functions of the length, L, and the operating wavelength, λ_0. The dotted lines in (b) and (c) mark the regions with wavelengths of 444 nm and 635 nm, respectively. (d) cuts of the cross-polarization energy conversion dispersions of (b) and (c) in blue (L= 165 nm) and red (L = 255 nm), respectively.

A.2.3. Dual Color Plasmonic Holographic Metasurface

The designed plasmonic metasurface structures can work as a static (in the sense that once its functionality is encoded it is unchangeable) wavefront modulator with binary phase control. For instance, it could be used as a Fresnel Zone Plate, a hologram or a beam shape converter, to mention but a few. We employed the metasurface to encode two CGH operating at each of those wavelengths. The CGHs were calculated as Fourier Holograms, see *Appendix D – Digital* Holography for details on the calculation, that are reconstructed in the far-field. The target images are shown in Figure A.4 (a) for the hologram that will be reconstructed with a red laser (wavelength of 635 nm) and Figure A.4 (b) for the blue laser (wavelength of 444 nm). Note that these images are the complementary of each other. Thus, if the reconstructions of each hologram is aligned it would result in a dual color image, as represented in Figure A.4 (c).

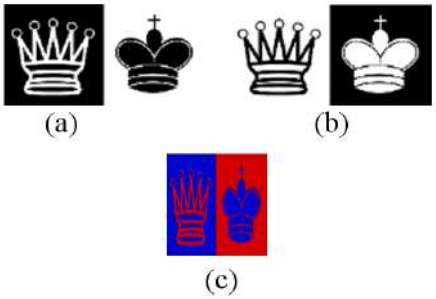


Figure A.4 - Target images for the red (a) and blue (b) holograms. (c) Shows a representation of the expected reconstruction of the merged metasurface when both holograms are excited.

To relax the image alignment requirements for the reconstructions, the metasurfaces containing the CGHs for each color were merged in only one structure following the pattern of a Chess Table, as shown in Figure A.5. Note that this arrangement makes the hologram stretched along the horizontal dimension. This will impact the reconstruction by making it thinner along the horizontal dimension according to the similarity principle of the Fourier Transform, that rules the diffraction of the Fourier Hologram according to (D.6), as represented in Figure A.4 (c). Furthermore, according to the spectra of each nanoantenna design, shown in Figure A.3 (d), the red (blue) design has a cross-polarization energy conversion of about 6% (1%) when operating at 444 nm (635 nm*)*. Therefore, there will be a noticeable cross-talk between the two wavelengths mainly due to the red design. The red CGH will be reconstructed also by light operating at 444 nm with an efficiency of approximately a half of the reconstruction efficiency of the red CGH with light at 635 nm. The blue CGH design, on the other hand, will be reconstructed at 635 nm with an efficiency of only 1%, that is ten times smaller than its efficiency at 444 nm.

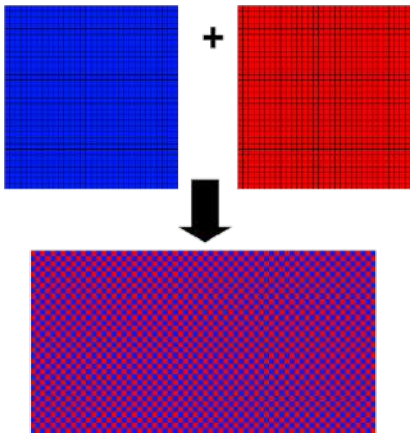


Figure A.5 - Representation of the merging between the CGHs designed to operate at 444 nm (blue square matrix on top) and 635 nm (red square matrix on top) resulting in the final Chess Table arrangement (below). Note that the arrangement is made by threading the pixels of each original CGH.

Unfortunately, the Chess Table arrangement has the drawback that the pixels for each color should have the same size but the unit cell design of the nanoantenna for each color have different sizes. Thus, two solutions were employed to solve this challenge. The first consists in adapting the blue nanoantenna design unit cell size to match the red one, which is larger. That is, it will be

used 386 nm for the pixel size of both structures, as shown in Figure A.6 (a). The second procedure consists in imposing the CGHs pixel size as 1930 nm × 1930 nm and fill this area with an array of 5×5 of nanoantennas, as depicted in Figure A.6 (b).

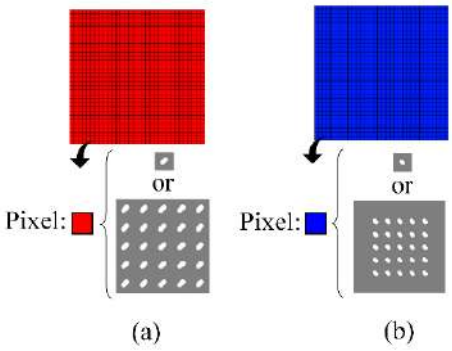


Figure A.6 - Arrangements used to represent each pixel of the CGHs for the red (a) and (b) holograms. The unit cell size of the isolated nanoantennas is equal to the red design, that is 386 *nm* × 386 *nm* while for the 5×5 arrays is 1930 nm × 1930 nm.

Finally, the metasurface mask was assembled using the commercial software TannerTools, which allows the generation of masks in GDSII format, required for posterior fabrication. Initially, the pixels represented in Figure A.6 were drawn on the software for each arrangement (one nanoantenna and 5×5 array of nanoantennas per pixel). Note that the binary phase CGHs for each color are separated and independent numerical matrices containing only 1s and -1s that correspond to the phases of the resulting distribution field from the IFTA. The numbers 1 and -1 represent phases 0 and π *rad*, respectively. Then, these matrices are merged according to the schemes shown in Figure A.6. This procedure is further represented in Figure A.7 (a) where the blue (red) numbers correspond to the CGH calculated to operate with blue (red) light. The final merged matrix is then inserted into a script on TannerTools that translates each the 1s and -1s to the corresponding representation by the nanoantennas. Therefore, if in a given position it is required a blue (red) +1 pixel, the previously drawn reference representation of the blue (red) pixel is inserted at that position. Instead, if a -1 pixel is required, the reference is rotated about its center by $\frac{\pi}{2}$ rad. The resulting masks of the scheme in Figure A.7 (a) are shown in Figures A**.7** (b) and (c) for metasurfaces that use one antenna per pixel and the array with 5×5 nanoantennas, respectively.

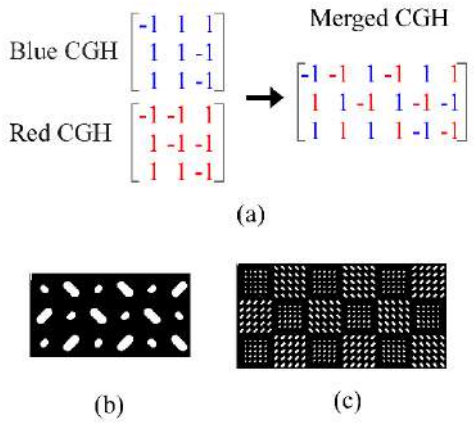


Figure A.7 — (a) Matrix representation of the merging of the blue and red CGHs. (b) and (c) show the resulting masks of the merged CGH shown in (a) with one nanoantenna per pixel and an array of 5×5 nanoantennas per pixel, respectively.

A.3. Fabrication

The fabrication of the metasurface was made in collaboration with Prof. Dr. Luis Alberto Barêa (UFSCar/UNICAMP), from the group of Prof. Dr. Newton Frateschi (IFGW/UNICAMP). As I did not participate directly in the fabrication process, its details were omitted.

Scanning electron microscopy (SEM) micrographs of parts of the fabricated metasurfaces can be seen in Figures A.8 (a) and (b) for the metasurfaces with one nanoantenna per pixel and an array of 5×5 nanoantennas per pixel, respectively.

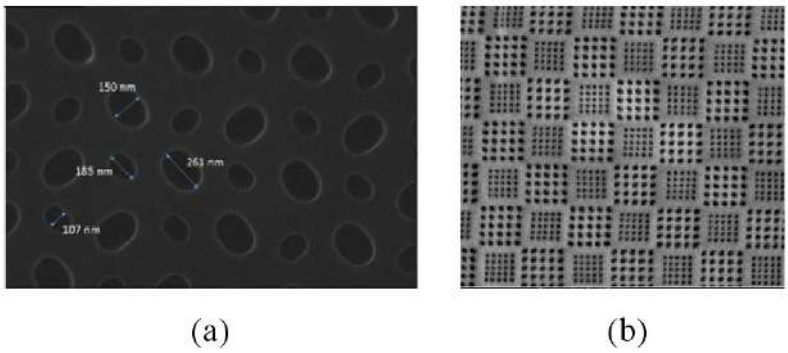


Figure A.8 – SEM micrographs of the (a) one nanoantenna per pixel metasurface and (b) 5×5 nanoantenna array per pixel.

A.4. Optical reconstruction

Unfortunately, only the design with the 5×5 nanoantenna array was available for optical characterization. The setup used for the measurements is shown in Figure A.9.

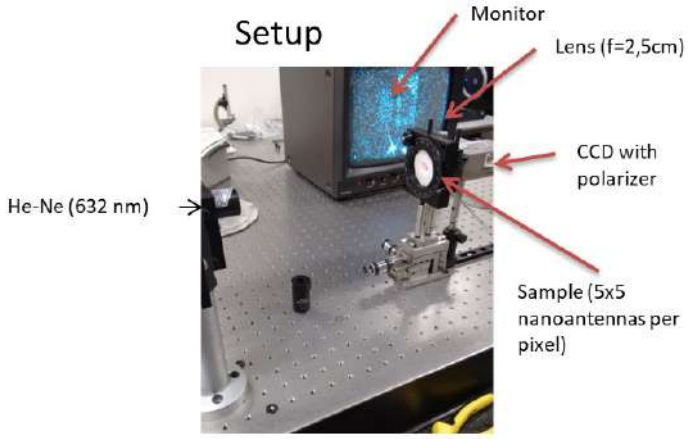


Figure A.9 – Optical setup used for the reconstruction of the CGH. The He-Ne laser is polarized horizontally. The lens is used to reconstruct the hologram on the focus of the CCD camera [3]. In front of the CCD, there is a polarizer that only transmits the cross-polarized light.

The He-Ne laser (632 nm) is already vertically polarized and a polarizer is placed in front of the CCD to allow only the cross-polarized light to impinge on it. Furthermore, a convergent lens (f=2.5 cm) was used so the reconstruction is formed on its focus, where the CCD is placed, instead of forming in the far-field [3]. The reconstructed image is shown in Figure A.10 (a) along with the numerical reconstruction in Figure A.10 (b), obtained via rigorous RS-Integration (see section *D.4.1 - RS Integration* for details), operating at 632 nm. As expected, both images are thinner (compare Figure A.10 (b) with Figure A.4 (a)) along the horizontal dimension because the merged hologram stretches along this dimension. The experimental reconstruction is extremely noisy, and one can hardly identify the target image. I hypothesize that the poor image quality is due to some pixels not transmitting light with the required phase by the CGH, thus degrading the reconstruction. This hypothesis was numerically tested by reconstructing the CGH, via RS integration, with a random phase added to its pixel. The resulting reconstruction is shown in Figure A.10 (c) and, as expected, it is severely noisier than the numerical reconstruction of the original CGH, in Figure A.10 (b), indicating that the presence of dead pixels on the fabricated metasurface may have contributed for the degeneration of its reconstruction. Unfortunately, the reconstruction was not present when the sample was excited with a wavelength of 444 nm.

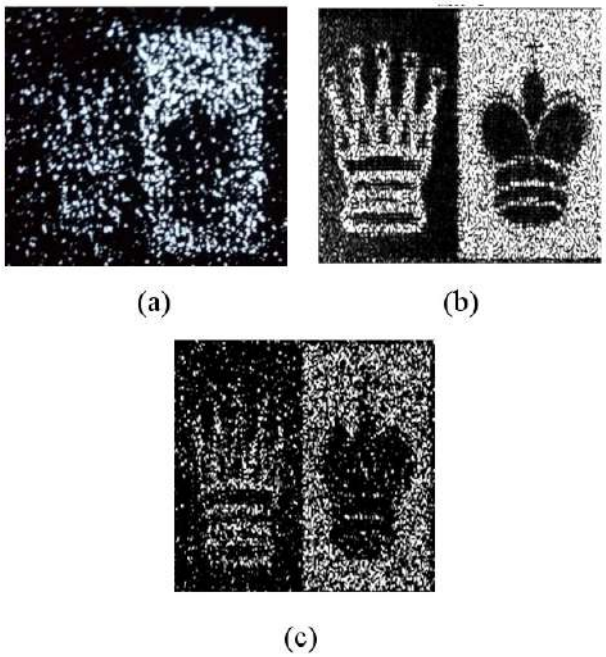


Figure A.10 - (a) Experimental and (b) numerical reconstructions of the metasurface operating at 632 nm. (c) Shows the reconstruction of the CGH, operating also at 632 nm, when a random phase noise is applied on it showing that the poor quality of the experimental reconstruction in (a) is due to phase variations on the metasurface.

According to Prof. Dr. Luis Barêa, in charge of the fabrications, the machine used to record the masks had temperature issues because the air conditioner of the room where it was installed was broken at that time. Therefore, this issue may have affected the geometries of the nanoantennas which, in turn, changed they electromagnetic response in an unpredictable way that harmed its functionality.

A.5. Conclusion

In this work it was projected, fabricated and experimentally measured the first holographic metasurfaces of this project. The used structures to compose the array were simple complementary nanorod antennas made of aluminum on a silica substrate. When excited at the LSPR, these structures may present a cross polarization conversion that can be used to achieve a binary phase control. Unfortunately, they present a theoretical limit conversion efficiency of 25%, neglecting losses. Thus, two designs were proposed that resonate at 444 nm and 635 nm each with a maximum cross-polarization efficiency of 12% both. These designs were used to assemble two dual color holographic metasurface designs that follows a chess table distribution of the pixels. The CGH were previously calculated with the IFTA, described at section *D.3 Iterative Fourier Transform Algorithm (IFTA)*. In one of the metasurface designs, the CGH pixel was represented by a single nanoantenna whereas in the other by a 5×5 nanoantenna array.

The designed metasurfaces were fabricated in collaboration with Prof. Dr. Luis Alberto Barêa (UFSCar/UNICAMP), and the group of Prof. Dr. Newton Frateschi (IFGW/UNICAMP). Unfortunately, we only received the 5×5 nanoantenna array design that suffered from temperature issues that deteriorated its optical performance, making the reconstruction at 632 nm extremely noisy and the one 444 nm completely washed out. Therefore, these issues and the small theoretical limit of the conversion efficiency (25%) motivated us to seek other concepts and materials for the holographic metasurfaces. Nevertheless, the presented work was important for the development of this thesis because it was the most directly attempt to its original proposal. It was a milestone in terms of learning experimental and design tools related to metasurfaces, electromagnetism and diffractive optics.

Appendix B – Field confinement in dielectric nanoposts

The starting point of the metasurface design is the calculation of the scattering parameters of its building blocks. Here, I was interested in manipulating the phase shift imparted by these structures onto the transmitted light. Therefore, I calculated the phase shift of the transmitted light and its efficiency by an array of identical structures. These simulations rely on Bloch's theorem and greatly reduces the computational requirements as discussed in section *II.2 Design of a metasurface.* However, this procedure assumes that the array and isolated nanopost responses match. That is, the response of the nanopost must be local in the sense that each nanopost is weakly coupled to its neighbours. To check if that is indeed the case, we need to analyse the modes supported by the array structure and check how confined they are in the nanoposts.

This section explores in depth the modes and resonances of the cylindrical nanoposts used in the holographic metasurfaces and metalenses design. First, the dispersion and field distributions of the Bloch modes supported by the structures are calculated. Additionally, the Bloch modes parameters of the same geometries are calculated assuming TiO_2 and a medium with refractive index of 1.5. It will be shown that the higher is the cylinder refractive index the higher is the field confinement of the fundamental Bloch modes in the cylinder. Then, a brief analysis is made on the field distribution of the Fabry-Perot modes found in the nanoposts that are the building blocks of the metasurfaces. Similar analysis and conclusions can be drawn with the elliptical nanoposts that are used in the birefringent designs with the difference that TM and TE modes are no longer degenerated. Thus, the analysis of the elliptical nanoposts is omitted.

B.1. Bloch modes

The Bloch modes are calculated using a RCWA code written by me and based on the works of Whittaker et al. [88] and Popov et al. [169]. The eigenvalue system is solved to find the out-of-plane (z direction) propagation constant of a periodic array of circles in a square lattice at a given frequency. The unit cell size with the relevant geometrical parameters is shown in Figure A.11**.** The embedding medium is always air and different materials are used within the cylinder. All calculations are done for out-of-plane propagation at the Γ point ($k_x, k_y = 0$).

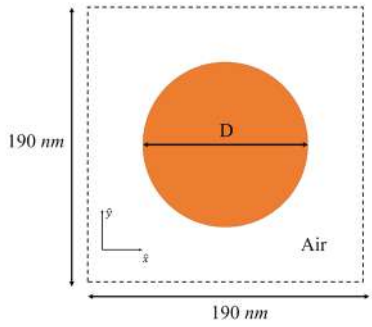


Figure A.11 - Unit cell representation.

Figures B.2, B.3 and B.4 show the out-of-plane effective indices of the first 20 Bloch modes with cylinders made of constant medium with refractive index 1.5, TiO_2 and c-Si, respectively. The blue diamonds mark the dispersion of the two fundamental modes (since they are degenerated, only one curve is visible) and the black dots the dispersions of the remaining modes. The TiO_2 and c-Si wavelength-dependent refractive indices were taken from [170] and [90], respectively. Since we are interested in the existence of propagating (non-evanescent) Bloch modes, only the real parts of the refractive indices were accounted for. In all figures, the upper and bottom rows show the real and imaginary parts of the effective indices. Structures with three diameters were analysed: 100 nm, 130 nm and 156 nm, which are shown in (a)-(b), (c)-(d) and (e)-(f), respectively, in Figures B.2, B.3 and B.4.

All SiO_2 based arrays allow only two propagating modes, which are the degenerated fundamental modes, due to the small refractive index contrast supported by these structures, as shown in Figure A.12. Therefore, these arrays can be treated in terms of an effective medium with refractive index given by the effective index of the fundamental modes. The TiO_2 structures present a similar behaviour except for the largest diameter at low wavelength. That is, it supports high order propagating modes for wavelengths smaller than 450 nm when the diameter is equal to 156 nm, as shown in Figure B.3. At this condition, the effective medium theory no longer holds, and the structure behaves as a photonic crystal. Finally, the c-Si based array of cylinders supports propagating higher order Bloch modes in a given region of the spectrum, as shown in Figure A.14. The onset of the propagating higher order modes is at 500 nm, 600 nm and 630 nm with arrays of cylinders having diameters of 100 nm, 130 nm and 156 nm, respectively. These are the wavelengths upon which the imaginary part of the following two higher order modes kicks off and become evanescent as shown by Figures B.4 (b), (d) and (f).

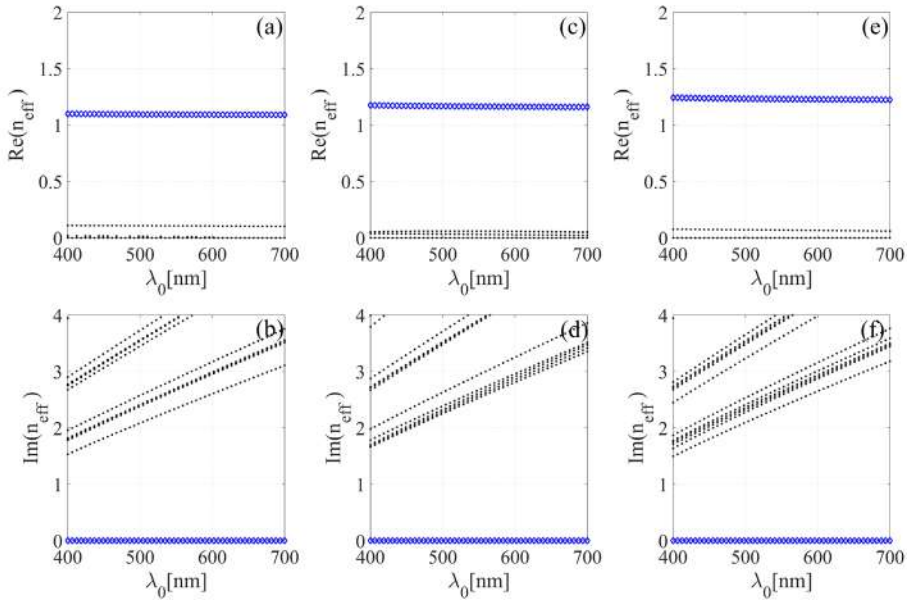


Figure A.12 – Out-of-plane effective indices of the first 20 Bloch modes with cylinders made of constant medium with a refractive index of 1.5. The blue diamonds mark the dispersion of the two fundamental modes (since they are degenerated only one curve is visible) and the black dots the dispersions of the remaining modes. In all figures, the upper and bottom rows show the real and imaginary parts of the effective indices. Structures with three diameters were analysed: 100 nm, 130 nm and 156 nm, which are shown in (a)-(b), (c)-(d) and (e)-(f), respectively.

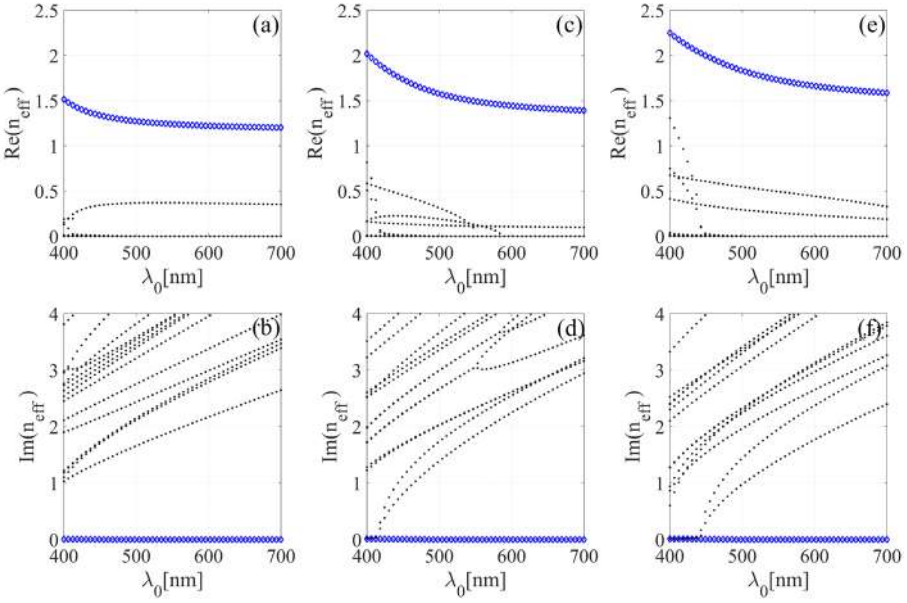


Figure A.13 – Out-of-plane effective indices of the first 20 Bloch modes with TiO_2 cylinders. The blue diamonds mark the dispersion of the two fundamental modes (since they are degenerated only one curve is visible) and the black dots the dispersions of the remaining modes. The TiO_2 refractive index dispersion was taken from [170]. In all figures, the upper and bottom rows show the real and imaginary parts of the effective indices. Structures with three diameters were analysed: 100 nm, 130 nm and 156 nm, which are shown in (a)-(b), (c)-(d) and (e)-(f), respectively.

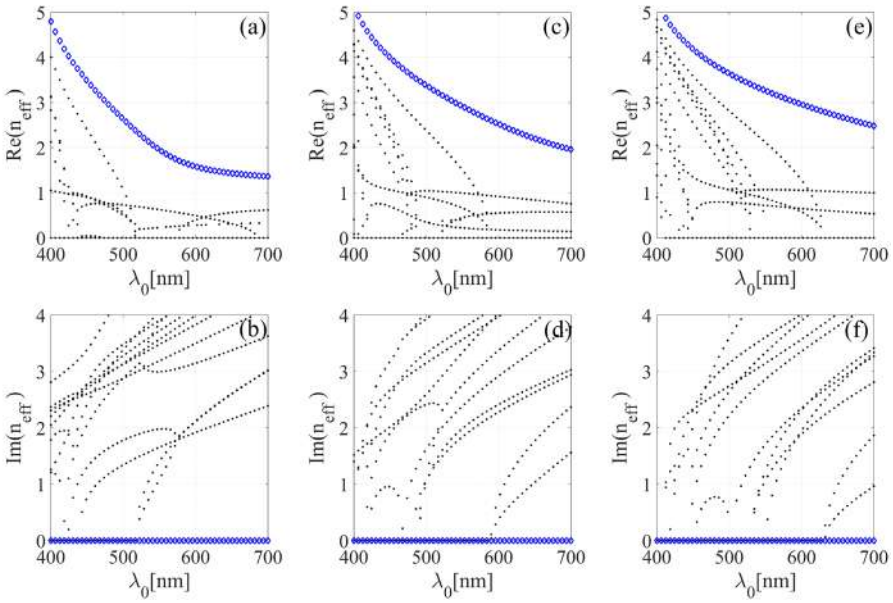


Figure A.14 – Out-of-plane effective indices of the first 20 Bloch modes with c-Si cylinders. The blue diamonds mark the dispersion of the two fundamental modes (since they are degenerated only one curve is visible) and the black dots the dispersions of the remaining modes. The c-Si refractive index dispersion was taken from [90]. In all figures, the upper and bottom rows show the real and imaginary parts of the effective indices. Since we are interested in the existence of propagating (non-evanescent) Bloch modes, only the real parts of the refractive indices were accounted for. Structures with three diameters were analysed: 100 nm, 130 nm and 156 nm, which are shown in (a)-(b), (c)-(d) and (e)-(f), respectively.

Thus, it seems that SiO_2 and TiO_2 are better than c-Si because they can support only one mode and can be treated in terms of an effective index medium. However, as a matter of fact, this is not a desired feature because the effective medium effects only hold when the low index cylinders are arranged in the array to create the required fundamental Bloch mode. This can be seen by the electric field amplitude of the fundamental x-oriented mode (stronger component along x direction), as shown in Figure A.15. The results for the structures made of SiO_2, TiO_2 and c-Si are shown in Figures B.5 (a), (b), and (c), respectively. Note that the field is mainly around the cylinders in the first two cases for all diameters analysed but has more field inside the cylinder for the TiO_2 cases. Thus, this indicates that the posts are strongly coupled to each other and the response of an isolated structure strongly differs from the array. The structures made of c-Si present most of the fields inside the cylinder for all cases but the one with the smaller diameter (100 nm) as shown in Figure B.5 (c). Therefore, most of light propagates in the nanopost without significant interaction between neighbour structures, which give these structures a local character.

However, we saw in Figure B.4 that the c-Si based structures can support higher order modes even on the desired region of the spectrum. In particular, at 532 nm, there are two additional modes for the structures with diameters of 130 nm and 156 nm that could give a photonic like behaviour for the structure. However, these modes present asymmetric distributions (see Figure A.16) and can couple with neither of the fundamental modes, which are symmetric, nor with the normally incident plane wave. Therefore, these are dark modes that will not affect the optical response. Therefore, the response of the array of truncated waveguides is ruled by the fundamental

Bloch mode, which is highly concentrated inside the structure. Thus, it is expected that the response of each c-Si nanopost in terms of phase accumulation and energy transmission better resembles that of the array than the other cases.

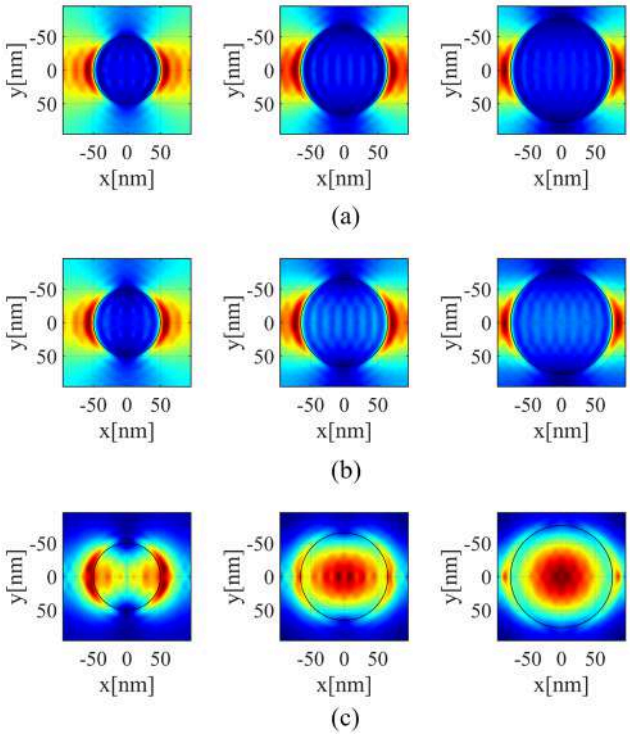


Figure A.15 – Fundamental x-polarized (stronger component along x) Bloch mode field amplitude distribution of a SiO_2 (a), TiO_2 (b) and c-Si (c) array. From left to the right, the cylinders diameters are 100 nm, 130 nm and 156 nm. The unit cell size is 190 nm and the operating wavelength is 532 nm.

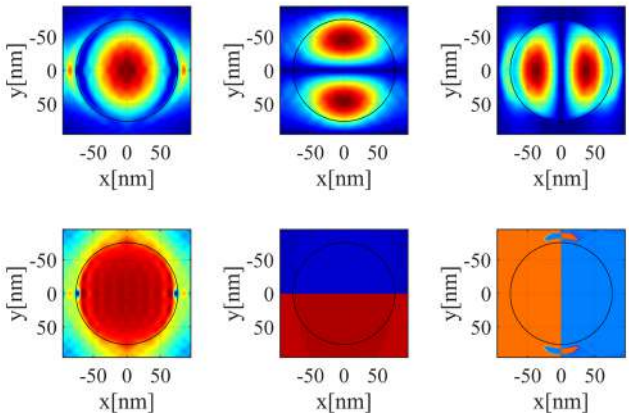


Figure A.16 – x-component field amplitude (top) and phase (bottom) distributions of the three Bloch modes with lowest effective index on a c-Si based array. From left to the right, the cylinders diameters are 100 nm, 130 nm and 156 nm. The unit cell size is 190 nm and the operating wavelength is 532 nm.

B.2. Nanoposts resonances

In the previous section it was shown that the field distribution of c-Si based 2D photonic is highly confined inside the high index region. Thus, the response of each individual post is similar to the response of the array in terms of the transmitted light phase delay and power. When this array is finite along the out-of-plane direction it behaves as a truncated waveguide (leaky cavity) that can support modes along its axis. As it was shown in Figures III-1 and III-2 of section *III.1.2 Metasurface design* in the main text, these truncated waveguides demonstrated a resonant behaviour for certain diameters, which were characterized by reduced transmissions in both designs. Since the physics is the same for both designs over Sapphire and Quartz, only the former is analysed here for simplicity. The simulations were carried out in Comsol using periodic boundary condition (Floquet) in the XY and YZ boundaries. The structures are excited with a x-polarized plane wave at normal incidence and operating at 532 nm.

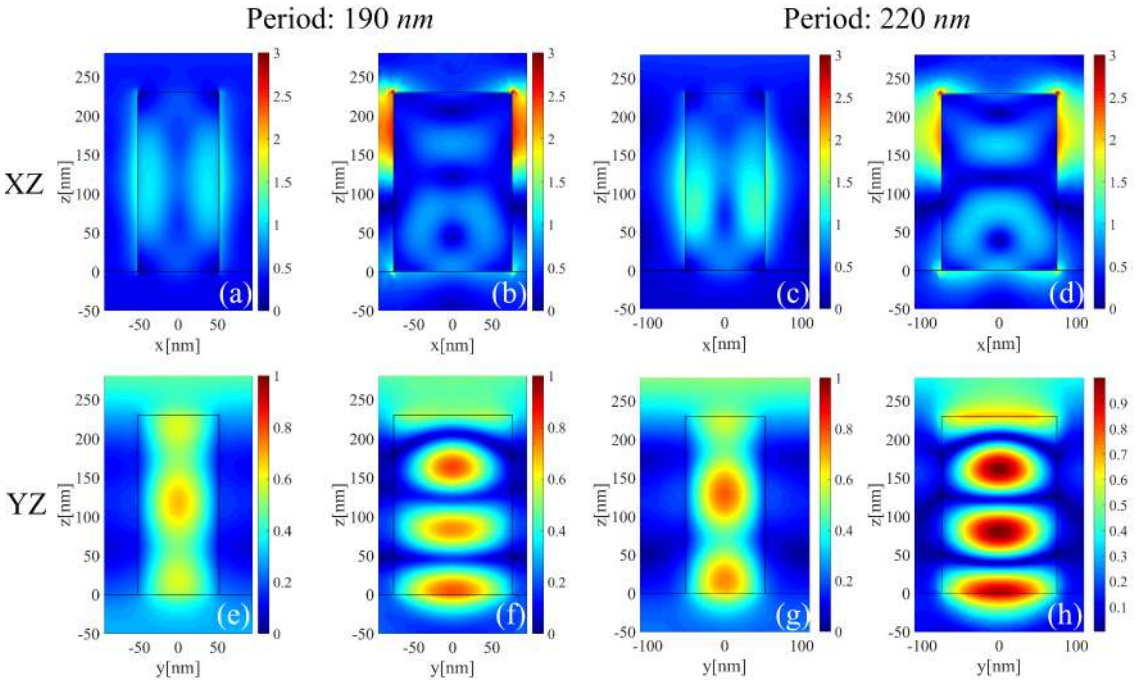


Figure A.17 - Electric field amplitude distributions of the structures with diameters 190 nm ((a), (b), (e) and (f)) and 220 nm ((c), (d), (g) and (h)). The top row ((a) – (d)) shows a cut along the XZ plane and the bottom row first row ((e) – (h)) along the YZ plane. Finally, (a), (c), (e) and (g) shows the results for posts with diameters of 105 nm and (b), (d), (f) and (h) shows the results for posts with diameters of 150 nm.

Figure A.17 shows the field distributions of the structures with diameters 190 nm (Figure A.17 (a), (b), (e) and (f)) and 220 nm (Figure A.17 (c), (d), (g) and (h)). The top row (Figure A.17 (a) – (d)) shows a cut along the XZ plane and the bottom row first row (Figure A.17 (e) – (h)) along the YZ plane. Finally, Figures A0 (a), (c), (e) and (g) shows the results for posts with diameters of 105 nm and Figure A.17 (b), (d), (f) and (h) shows the results for posts with diameters of 150 nm. Note that this resonance are marked by the presence of intense region of maximum

that can be better seen along the YZ plane. Finally, note that all fields distributions are very intense inside the post, which is the reason for the enhanced absorption under the resonance as shown in Figure 2 of section *III.1.2 Metasurface design* in the main text.

.

Appendix C – Metalenses additional informations

C.1. Phase and transmission maps

The metalenses are comprised of eight different meta-atoms to encode eight phase levels. Table C-1 shows the corresponding diameters with the ideal phase levels at normal incidence.

Table C-1 - Metalens nanoposts diameters and ideal relative phases

Id	#1	#2	#3	#4	#5	#6	#7	#8
Diameter [nm]	71	88	96	101	106	111	120	137
Ideal relative phase [rad]	0	$\pi/4$	$\pi/2$	$3\pi/4$	π	$5\pi/4$	$3\pi/2$	$7\pi/4$

Figures C.1 and C.2 [C.3 and C.4] show the relative phase and transmittance maps, respectively, of the eight meta-atoms (labeled according to Table S3), as function of the normalized parallel wavevector components k_x and k_y at s and p polarizations. The operating wavelength is 532 nm. It is clear that, except for nanoposts #6 and #8, both the phase and transmissions are largely independent of the angle of incidence, thus guaranteeing a WFOV. Nanoposts #6 and #8 showed a more pronounced dependency on the angle of incidence, but this dependency did not perceptibly affect the overall metalens performance.

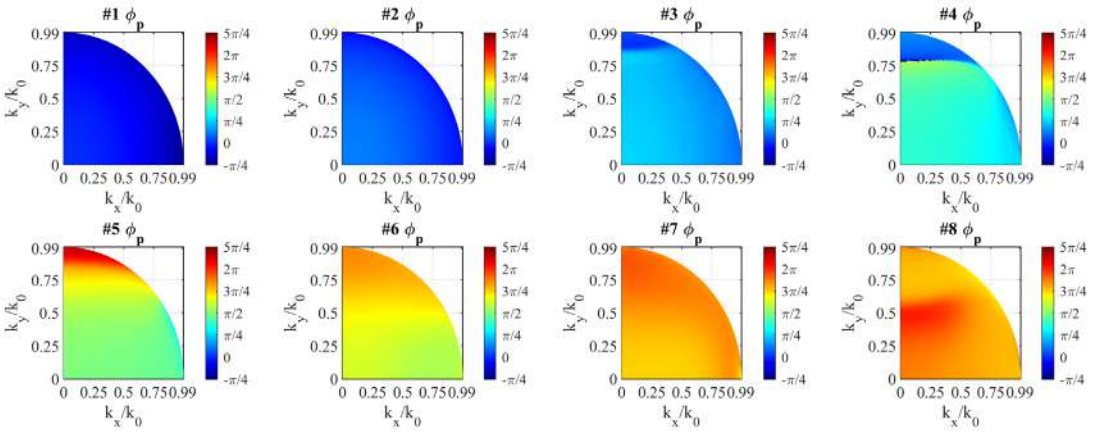


Figure C.1 - Phase maps, in units of rad, of the eight structures used in the c-Si metalens at p polarisation. The operating wavelength is 532 nm.

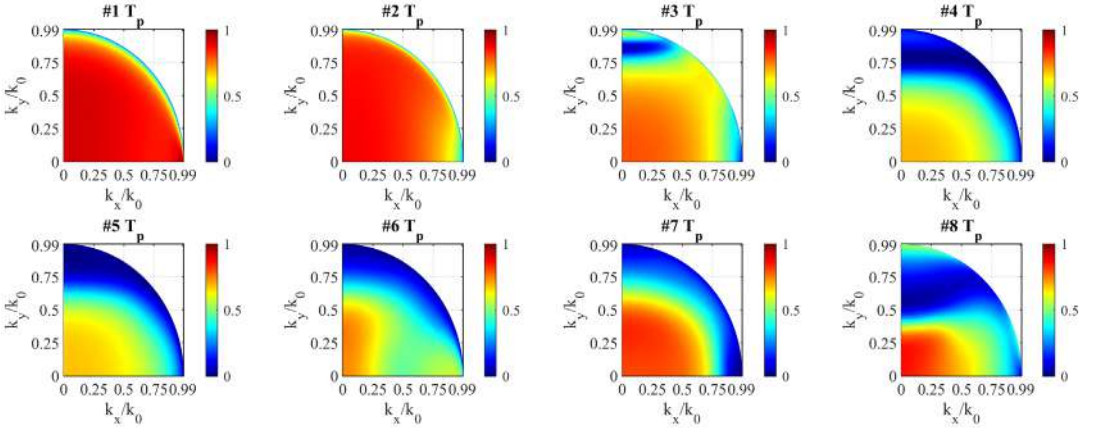


Figure C.2 - Transmittance maps of the eight structures used in the c-Si metalens at p polarization. The operating wavelength is 532 nm.

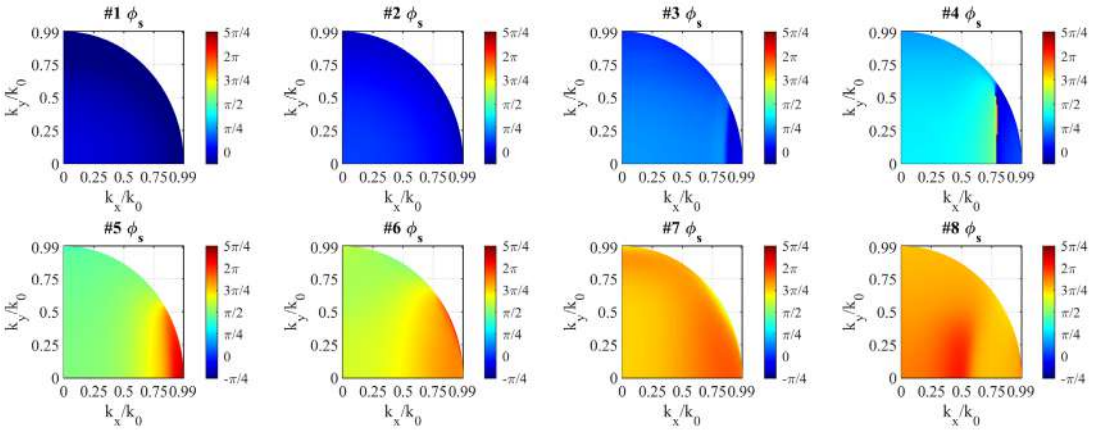


Figure C.3 - Phase maps, in units of rad, of the eight structures used in the c-Si metalens at s polarisation. The operating wavelength is 532 nm.

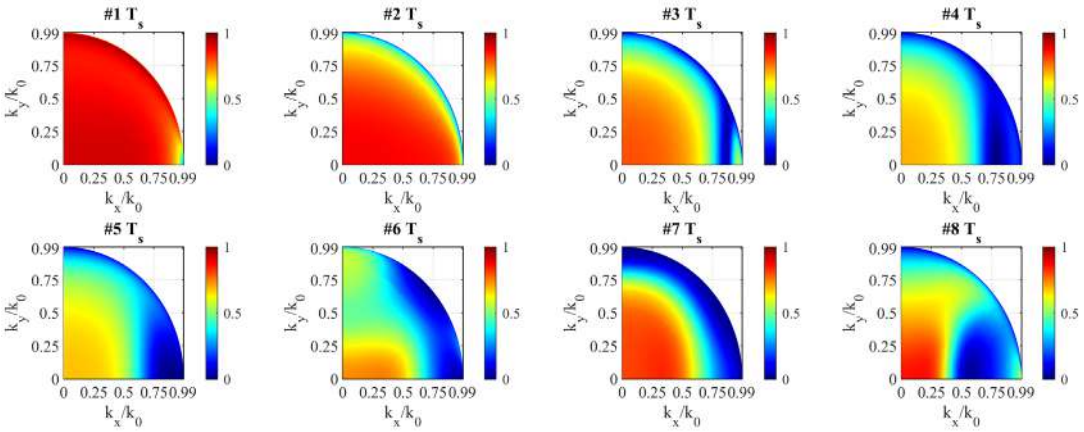


Figure C.4 - Transmittance maps of the eight structures used in the c-Si metalens at s polarisation. The operating wavelength is 532 nm.

C.2. Fourier Transforms of quadratic (WFOV) and hyperbolic (DL) field profiles.

The Fourier Transform (FT) amplitude of the field profile generated by the WFOV and DL metalenses are shown in Figures C.5 (a) and (b), respectively, as function of the radial component of the wavevector and the NAs of the metalenses. The dashed white lines correspond to the numerical aperture equation NA = ± sin $\theta = \pm k_r/k_0$. Note that the hyperbolic spectrum is totally confined inside the cone defined by those lenses (Figure C.5 (b)) whereas the quadratic is only confined up to NA ≈ 0.3 (Figures C.5 (a)). As the NA increases, the hyperbolic FT has a higher intensity around the cone edges, which explains its diffraction limits capacity. The FT of the quadratic profile, however, spreads the energy at high frequency components outside the cone, which explains its high field of view [114].

In Figure C.5 (c), it is shown a cut of Figures C.5 (a)-(b) along NA = 0.3. Up to NA ≈ 0.3, the spectra are very similar, as clearly seen in Figure C.5 (c). Figure C.5 (d) shows a cut along NA=0.8 of Figures C.5 (a) and (b).

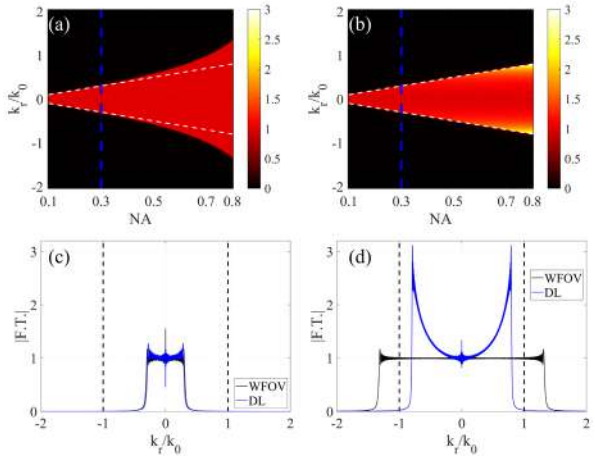


Figure C.5 - (a) Quadratic and (b) hyperbolic field profiles Fourier transform comparison as function of the metalenses 'NAs and radial wavevector component. The white dashed lines mark the equations $NA=\pm kr/k0$. (c) [(d)] shows a cut of (a) and (b) in black and blue lines, respectively, at an NA of 0.3 [0.8]. The operating wavelength is 532 nm and the focal length 750 μm.

C.3. Spectrum flattening of the spherical phase profile

The magnitude of the Fourier Transform of the spherical phase profile is given by:

$$\left|\mathcal{F}\left(e^{i\emptyset_{sph}}\right)\right| = 2\pi\frac{k_0 R}{f}\frac{1}{k_\perp^2 + \frac{k_0^2 R^2}{f^2}}\sqrt{R^2 + \frac{1}{k_\perp^2 + \frac{k_0^2 R^2}{f^2}}} \qquad \text{(C.1)}$$

where R is the radius of curvature, f is the focal length and $k_\perp^2 = k_x^2 + k_y^2$. The dependency of the FT magnitude on $k_\perp$ vanishes in the limit of R going to infinity:

$$\lim_{R\to\infty}\left|\mathcal{F}\left(e^{i\emptyset_{sph}}\right)\right| = 2\pi\frac{f}{k_0} \qquad \text{(C.2)}$$

which accounts for the spectral flattening. Note that equations (C.1)and (C.2) assume an infinite large metalens. In reality, they would be convoluted with a *jinc* function due to the limited numerical aperture of the metalens.

C.4. Angular spectrum formalism

All simulations done in chapter IV of the main text involved a rigorous wave-optics treatment which does not involve the paraxial approximation. It was assumed a background media with refractive index *n* and an operating free-space wavelength λ_0. The beam is propagated using the angular spectrum formalism [3]. That is, given a field in a plane $\Pi = [z = z_1, \forall(x,y,z) \in \mathbb{R}^3]$, the resulting diffracted field in a plane $\Gamma = [z = z_2, \forall(x,y,z) \in \mathbb{R}^3]$ which is parallel to Γ, is given by

$$U(x,y;z_2) = F^{-1}\left\{F\left[U(x,y;z_1)\right](k_x,k_y)\,e^{ik_z(z_2-z_1)}\right\}(x,y) \qquad \textbf{(C.3)}$$

where, $U(x,y;z)$ is the field distribution, F and F^{-1} are the bi dimensional Fourier and inverse Fourier transforms, respectively, given by

$$F\left[g(x,y)\right](k_x,k_y) = G(k_x,k_y) = \iint_{\mathbb{R}^2} g(x,y)\,e^{-j(x,y)\cdot(k_x,k_y)}dxdy \qquad \textbf{(C.4)}$$

$$F^{-1}\left[G(k_x,k_y)\right](x,y) = g(x,y) = \frac{1}{(2\pi)^2}\iint_{\mathbb{R}^2} G(k_x,k_y)\,e^{j(x,y)\cdot(k_x,k_y)}dk_x dk_y \qquad \textbf{(C.5)}$$

Finally, k_z is the z component of the wave-vector $\vec{k} = (k_x, k_y, k_z)$ ($\left|\vec{k}\right| \equiv k = 2\pi/\lambda_0 n$)

$$k_z = \sqrt{k^2 - k_x^2 - k_y^2} \qquad \textbf{(C.6)}$$

The advantage of this approach is that it can be calculated using Fast Fourier Transform (FFT) algorithm. Thus, this method is faster and more exact than other approaches, such as the Rayleigh-Sommerfeld (RS) integral calculation. Next section explores the differences between the RS integral and Angular spectrum formalisms.

C.4.1. Equivalence to the Rayleigh-Sommerfeld equation

Although there is no question about the validity of equation (C.3), I had not seen an equivalency among it and the Rayleigh-Sommerfeld (RS) integral (shown in equation (D.50) [3], which is also a rigorous diffraction equation under the wave-optics approximation. As the RS integral is a more stablished equation for diffraction calculations, it is convenient to draw the link with the angular spectrum and check their similarities.

For convenience, the RS equation is rewritten bellow

$$U_{RS}(x,y;z_2) = -\frac{jk}{2\pi}\iint_{\mathbb{R}^2} dx'dy'\, U(x',y';z_1)\frac{z e^{jkr(x'-x,y'-y,z_2-z1)}}{r^2(x'-x,y'-y,z_2-z_1)} \quad \text{(C.7)}$$

where,

$$r(x,y;z) = \sqrt{x^2+y^2+z^2} \quad \text{(C.8)}$$

Thus, I begin by applying the convolution theorem in the angular spectrum equation (C.3) , which results in

$$U(x,y;z_2) = U(x,y;z_1) \otimes \mathcal{F}^{-1}\left(e^{ik_z(z_2-z_1)}\right) \quad \text{(C.9)}$$

But, according to Weyl's identity [171],

$$\frac{e^{jk\sqrt{x^2+y^2+z^2}}}{\sqrt{x^2+y^2+z^2}} = j2\pi\mathcal{F}^{-1}\left[\frac{1}{k_z}e^{ik_z z}\right](x,y) \quad \text{(C.10)}$$

Deriving both sides of (C.10) by z, it follows that

$$\mathcal{F}^{-1}\left[e^{ik_z z}\right](x,y) = -\frac{1}{2\pi}\frac{z e^{jk\sqrt{x^2+y^2+z^2}}}{x^2+y^2+z^2}\left(jk - \frac{1}{\sqrt{x^2+y^2+z^2}}\right) \quad \text{(C.11)}$$

Substituting (C.11) in (C.9) and defining $r \equiv \sqrt{x^2+y^2+z^2}$

$$U(x,y;z_2) = U(x,y;z_1) \otimes \left[-\frac{1}{2\pi}\frac{z e^{jkr}}{r^2}\left(jk-\frac{1}{r}\right)\right] =$$

$$= -\frac{1}{2\pi}\iint_{\mathbb{R}^2} dx'dy'\, U(x',y';z_1)\left\{\frac{z e^{jkr(x'-x,y'-y,z_2-z_1)}}{r^2(x'-x,y'-y,z_2-z_1)}\left[jk - \frac{1}{r(x'-x,y'-y,z_2-z_1)}\right]\right\} \quad \text{(C.12)}$$

Comparing the RS integral (Eq. (C.7)) with (C.12), one concludes that

$$U(x,y;z_2) = U_{RS}(x,y;z_2) + \frac{1}{2\pi}\iint_{\mathbb{R}^2} dx'dy'\, U(x',y';z_1)\frac{z e^{jkr(x'-x,y'-y,z_2-z_1)}}{r^3(x'-x,y'-y,z_2-z_1)} \quad \text{(C.13)}$$

That is, the angular spectrum equation contains both the RS integral and another integral, whose integrand is proportional to $\frac{z}{r^3} \sim \frac{1}{distance^2}$ and do not radiate. Therefore, the angular spectrum integral is more accurate than the RS integral, especially in the near field domain. This is a

reasonable result since the angular spectrum equation is derived from Helmholtz equation with only the scalar wave approximation whereas the RS integral carries also the radiation approximation [3].

C.5. FOV characterization

In order to quantify the field of view, we define a function that expresses the distortion suffered by the point spread function at oblique incidence. The obvious approach is to calculate the energy of the difference between the PSF amplitudes at normal and oblique incidences. That is,

$$\xi(\theta) = \int_{\Gamma} dS \left| \frac{|P(x,y,0)|}{\sqrt{\rho(0)}} - \frac{|P(x-\Delta x, y-\Delta y, \theta)|}{\sqrt{\rho(\theta)}} \right|^2 \qquad \text{(C.14)}$$

where, θ is the angle of incidence, $P(x,y,\theta)$, is the PSF field amplitude at that angle and $\rho(\theta)$ its focusing efficiency. Note that the oblique incidence PSF was displaced by (Δx,Δy) to align the centre of its peak with the P(x,y,0) (see Figure C.6). We observed that the point spread function is well defined for. The FOV is thus defined as the angular cone for which $\xi(\theta) < 0.1$.

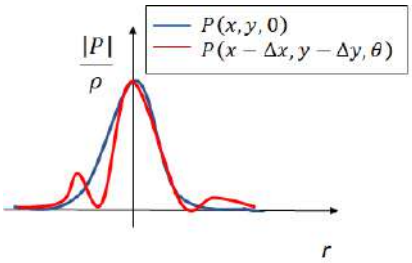


Figure C.6 - Representation of the cross section of the PSF at normal (blue) and oblique incidence (red). Notice that the later was displaced to align with the former.

C.6. USAF 1951 chart

The USAF 1951 resolution chart is a test target created by the U. S. Air force to comply with the MIL-STD-150A standard [172]. It consists of several pairs of elements orthogonally arranged. Each element has three rectangles with a length to width ratio of 5 to 1. The elements are arranged in groups of six as shown in Figure C.7. Each element within a group is numbered from 1 to 6. Note, in Figure C.7 that the groups are arranged in pairs with the even to the left (red box) and the odd to the right (blue box).

The width (w) of the elements change according to equation (C.15)

$$w = \frac{2^{-\left(G+\frac{E-1}{6}\right)}}{2}\ mm \qquad \textbf{(C.15)}$$

Where E is the element number and G the group number. Table C-2 lists the width of the elements from the group -1 to 7.

Table C-2 – Element width (w) in the USAF chart according to (C.15). The units are in µm.

Element (E)	Group Number (G)								
	−1	0	1	2	3	4	5	6	7
1	1000.00	500.00	250.00	125.00	62.50	31.25	15.63	7.81	3.91
2	890.90	445.45	222.72	111.36	55.68	27.84	13.92	6.96	3.48
3	793.70	396.85	198.43	99.21	49.61	24.80	12.40	6.20	3.10
4	707.11	353.55	176.78	88.39	44.19	22.10	11.05	5.52	2.76
5	629.96	314.98	157.49	78.75	39.37	19.69	9.84	4.92	2.46
6	561.23	280.62	140.31	70.15	35.08	17.54	8.77	4.38	2.19

Note that pairs of groups from (C.15)

group exponentially to the power of 2 and can, therefore, be arranged concentrically, as shown in Figure C.8. For example, all the groups of would fit inside the pair shown in Figure C.7.

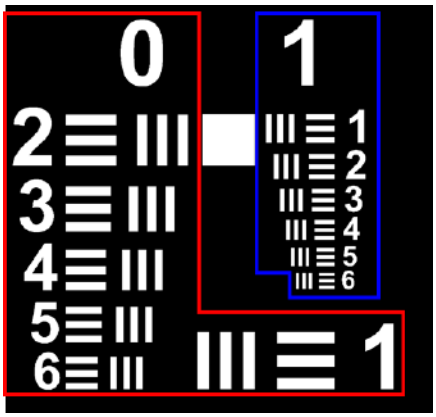


Figure C.7 – USAF 1951 chart with the groups 0 (red box) and 1 (blue box). This figure is not to scale. Table C-2 lists the elements widths.

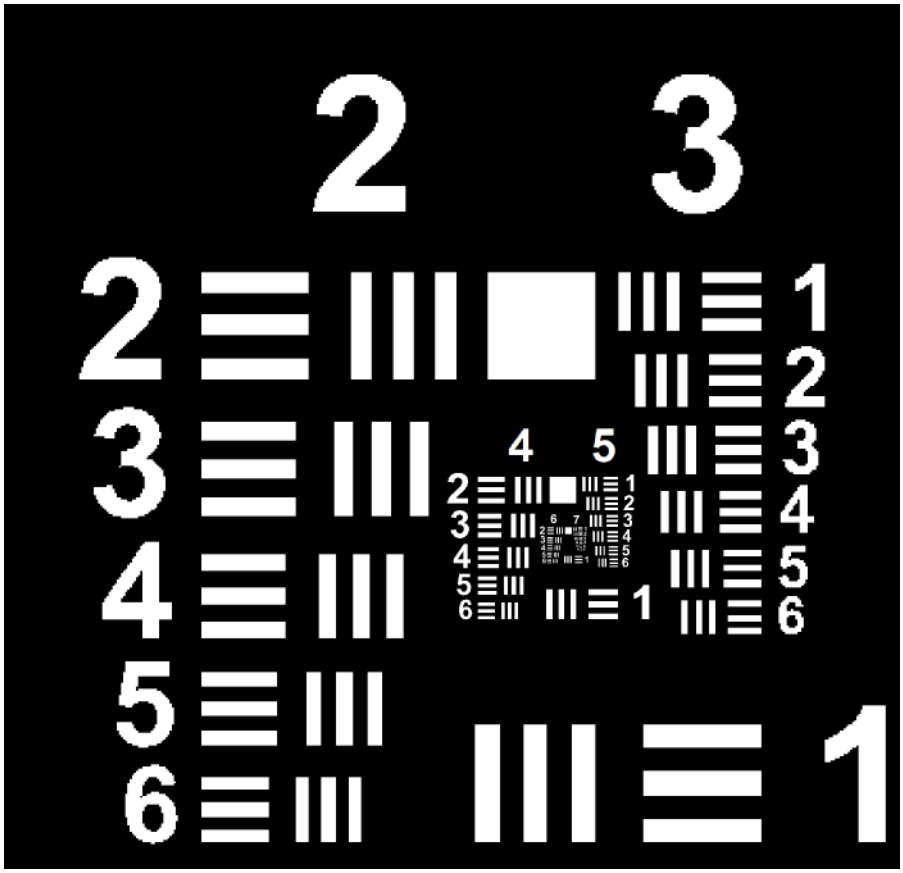


Figure C.8 – USAF 1951 chart with groups 2 to 7. This figure is not to scale. Table C-2 lists the elements widths.

Appendix D – Digital Holography

This section explains thoroughly how the *Iterative Fourier Transform Algorithm* (IFTA) algorithm is implemented. The method was initially proposed by Gerchberg and Saxton [73] and it was followed by several improvements [74, 75]. The presented implementation was taken and adapted from [75], where most of the improvements are described. Additionally, it was implemented three algorithms for the diffraction calculation of the CGHs by the RS (Rayleigh Sommerfeld) integration, which better reproduces the real diffraction. These algorithms exploit some symmetries in the hologram and reconstruction planes that can reduce the calculation time by a factor of 5 (note that some reconstructions may take days using the RS integration). Furthermore, the RS integration manifested a wide angle aberration in the numerical reconstruction of the optical metasurface holograms that is not expected by the IFTA algorithm because the IFTA is based on either Fresnel or Fourier diffractions [75]. In short, the designed CGHs have subwavelength pixels, which causes their diffraction to spread over wide angles thus violating the paraxial approximation that is assumed in the Fresnel and Fourier diffractions [3]. As the IFTA diffraction is based on Fourier transforming the fields in each plane, the wide-angle distortion does not happen. Therefore, a correction, based on an approach proposed in [91], was implemented to compensate for the wide angle aberration.

Appendix C is organized as follows: in section D.1 the mathematical foundations behind the scalar diffraction are defined. In section D.2 all diffraction equations defined in section 1 are discretized and some useful relations are obtained. Section D.3 describes the IFTA algorithm and has some examples. Finally, in section D.4, the RS integration method is described and it is shown how to reduce its calculation time.

D.1. Diffraction in Fourier Optics - Mathematical Background

In this section, the analytical equations of diffraction will be defined and it is not intended to provide a rigorous demonstration of any equation, which can be found in [3]. The objective of this chapter is to define the equations that relate a given field distribution (hologram) to its diffracted field (reconstruction or object field). The hologram and object fields are defined at planes Σ and Π, respectively, as shown in Figure D.1, and represented by $G(x,y)$ and $U(\xi,v)$, respectively.

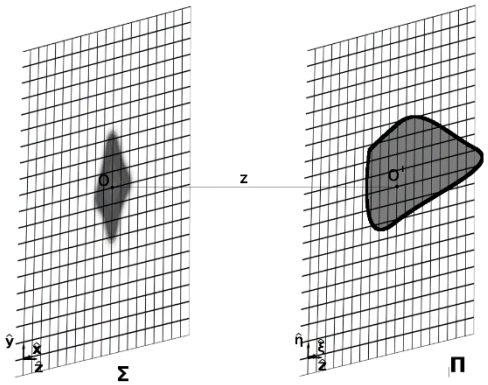


Figure D.1 - Analysed diffraction geometry with the hologram (Σ) and reconstruction (Π) planes. The origin of these planes are the points O and O', respectively, whose distance is $\overline{OO'} = z$,

Assuming that the scalar theory of diffraction holds, the diffraction of *U* from Π to Σ is rigorously given by [3]

$$G(x,y) = -\frac{z}{j\lambda_0}\int\int_{\Pi} U(\xi,\eta)\frac{e^{-jk_0 r(\xi,\eta;x,y)}}{r^2(\xi,\eta;x,y)}\,d\xi d\eta \qquad \textbf{(D.1)}$$

where λ_0 is the working wavelength, $k_0 = \frac{2\pi}{\lambda_0}$ the freespace wavenumber and *r* is the distance between the points *(x,y)*, at Σ, and (ξ,η), at Π, defined as

$$r(\xi,\eta;x,y) = \sqrt{(x-\xi)^2 + (y-\eta)^2 + z^2} \qquad \textbf{(D.2)}$$

By simply changing the propagation direction, we can also express *U* in terms of *G* as

$$U(\xi,\eta) = \frac{z}{j\lambda_0}\int\int_{\Pi} G(x,y)\frac{e^{jk_0 r(\xi,\eta;x,y)}}{r^2(\xi,\eta;x,y)}\,dxdy \qquad \textbf{(D.3)}$$

If the paraxial approximation holds ($z \gg x,y,\xi,\eta$), the following approximation can be made for the denominators of equations (D.1) and (D.3)

$$\frac{1}{r^2(\xi,\eta;x,y)} = \frac{1}{z^2}\frac{1}{\left[\frac{(x-\xi)^2}{z^2}+\frac{(y-\eta)^2}{z^2}+1\right]} \cong \frac{1}{z^2} \qquad \textbf{(D.4)}$$

And the following for the complex exponentials at equations (D.1) and (D.3)

$$r(\xi,\eta;x,y) = z\sqrt{\frac{(x-\xi)^2}{z^2}+\frac{(y-\eta)^2}{z^2}+1} \cong z\left[1+\frac{1}{2}\frac{(x-\xi)^2}{z^2}+\frac{1}{2}\frac{(y-\eta)^2}{z^2}\right] \qquad \textbf{(D.5)}$$

Note that equations (D.4) and (D.5) are the binomial expansions of $r(\xi,\eta;x,y)$ with only one and two terms, respectively. The exponential requires additional terms because it is multiplied by *k*, that normally is a large number, and small variations in the exponent can change the value of the exponential significantly [3].

Substituting the approximations (D.4)and (D.5) on (D.1) and, after a few algebraic manipulations, one obtains the Fresnel diffraction equation given by [3]

$$G(x,y) = -\frac{e^{-jk_0z}}{j\lambda_0 z}e^{-j\frac{k_0}{2z}(x^2+y^2)}\mathcal{F}^{-1}\left\{U(\xi,v)e^{-\frac{jk_0}{2z}(\xi^2+v^2)}\right\}\left(\frac{x}{\lambda_0 z},\frac{y}{\lambda_0 z}\right) \qquad \textbf{(D.6)}$$

where $\mathcal{F}$ is the two-dimensional Fourier Transform operator defined by

$$\mathcal{F}\{f\}(u,v) \equiv G(f_x,f_y) = \int_{\mathbb{R}^2} dxdyf(x,y)e^{-2\pi j(f_x,f_y)\cdot(x,y)} \qquad \textbf{(D.7)}$$

For a given function $f\colon \mathbb{R}^2 \to \mathbb{R}$. Similarly, we define the inverse Fourier Transform of $\mathrm{G}(\mathrm{f_x},\mathrm{f_y})$, represented by $\mathcal{F}^{-1}\{\mathrm{G}\}$, as

$$\mathcal{F}^{-1}\{G\}(x,y) \equiv f(x,y) = \int_{\mathbb{R}^2} df_x df_y G(f_x,f_y)e^{2\pi j(f_x,f_y)\cdot(x,y)} \qquad \textbf{(D.8)}$$

We can also write a simplified expression for *U* in terms of *G* by either applying the approximations (D.5) and (D.6) at (D.3) or by simply changing the direction of propagation at (D.6). Either way, one finds [3]

$$U(\xi,\eta) = \frac{e^{jk_0z}}{j\lambda_0 z}e^{\frac{jk_0}{2z}(\xi^2+v^2)}\mathcal{F}\left\{G(x,y)e^{j\frac{k_0}{2z}(x^2+y^2)}\right\}\left(\frac{\xi}{\lambda_0 z},\frac{\eta}{\lambda_0 z}\right) \qquad \textbf{(D.9)}$$

Finally, these equations can be further simplified to the Fraunhofer diffraction under either one of the following conditions [3]:

1. $z \gg \frac{k_0(\xi^2+v^2)_{max}}{2}$

2. The object and the hologram plane are on the focuses of a convergent lens with focal length f.

The Fraunhofer approximation is obtained by neglecting the quadratic phase terms (given by $e^{\pm \frac{jk_0}{2z}(u^2+v^2)}$) at (D.6) or (D.9). Therefore, besides the paraxial approximation, the diffracted field will be in the far field zone. For both cases, (D.6) reduces to

$$G(x,y) = -\frac{e^{-jk_0 z}}{j\lambda_0 z}\mathcal{F}^{-1}\{U(\xi,v)\}\left(\frac{x}{\lambda_0 z},\frac{y}{\lambda_0 z}\right) \tag{D.10}$$

that is, the diffraction is obtained by simply Fourier Transforming its field distribution, scaling it, and multiplying it by phase and amplitude factors. Furthermore, we can also apply the Fraunhoffer approximation to (D.9) obtaining

$$U(\xi,v) = \frac{e^{jk_0 z}}{j\lambda_0 z}\mathcal{F}\{G(x,y)\}\left(\frac{\xi}{\lambda_0 z},\frac{\eta}{\lambda_0 z}\right) \tag{D.11}$$

These relations also hold for case 2 with a lens with $z = f$. Therefore, equations(D.6) and (D.9) will be used to relate the field distributions between the Hologram and Reconstruction planes for **Fresnel Holograms**. Equations (D.10) and (D.11), in turn, do the same but for **Fourier Holograms**. Note that all these equations perform a Fourier Transformation, which can be calculated by the Fast Fourier Transform (FFT) algorithm in a computer. Therefore, the IFTA algorithm will be very fast because it benefits from the FFTs fast calculations. For simplicity, equations (D.6) and (D.10), that relates G in terms of $\mathcal{F}^{-1}\{U\}$, will be redefined as

$$G(x,y) = A_{-}[f\ (x,y)]^{-1}\mathcal{F}^{-1}\{U(\xi,v)[f(\xi,v)]^{-1}\}\left(\frac{x}{\lambda_0 z},\frac{y}{\lambda_0 z}\right) \tag{D.12}$$

Whereas, equations (D.9) and (D.11), that relates U in terms of $\mathcal{F}\{G\}$ as

$$U(\xi,v) = A_{+}f(\xi,\eta)\mathcal{F}\{G\ (x,y)\ f(x,y)\}\left(\frac{\xi}{\lambda_0 z},\frac{\eta}{\lambda_0 z}\right) \tag{D.13}$$

where $A_{\pm} = \pm\frac{e^{\pm jk_0 z}}{jk_0 z}$ and the function $f: \mathbb{R}^2 \to \mathbb{C}$ is defined as

$$f(a,b) \equiv f = \begin{cases} e^{\frac{jk_0}{2z}(a^2+b^2)} & for\ Fresnel\ Holograms \\ 1 & for\ Fourier\ Holograms \end{cases} \tag{D.14}$$

D.2. Computational development of the diffraction relations

In a computer we are only able to store a finite amount of data, but equations (D.12) and (D.13) relate complex functions of the kind $f:\mathbb{R}^2 \rightarrow \mathbb{C}$. Therefore, to represent digitally the field distributions, we need not only discretize equations (D.12) and (D.13) but also limit the calculation to a finite region at the hologram (Σ) and reconstruction (Π) planes. Thus, the hologram region will be limited to a rectangle centered on the origin of Σ discretized by *M* and *N* points along directions *x* and *y*, respectively, each separated by P_x and P_y along those directions. The reconstruction plane will also be limited in a rectangle centered at the origin of Π and discretized by *M* and *N* points with separations that depends on P_ξ, P_η, operating wavelength and the separation between the planes (*z*) as will be shown. Therefore, the field distributions of the hologram (*G*) and of the object (*U*) will be numerically stored in matrices G_{MxN} and U_{MxN}, respectively.

This section starts by sampling the field distributions at each plane and calculating the diffracted field of each one. The effect of the sampling of the hologram on the diffraction will be discussed. The next part deals with the relationship between the discretized fields at both planes, i.e., it will be shown that they relate to each other by the discrete Fourier transform, as expected.

D.2.1. Sampling the fields

In this step we will first sample the field distributions at each plane in a finite matrix of square pixels. Therefore, the field at each pixel of the sampled distribution is assumed to have constant amplitude and phase given by

$$g[m,n] = \begin{cases} G(mP_x, nP_y)\left[f(mP_x, nP_y)\right]^{-1}, & m \in \left[-\frac{M}{2}, \frac{M}{2}-1\right] and\ n \in \left[-\frac{N}{2}, \frac{N}{2}-1\right] \\ 0, & otherwise \end{cases} \quad \textbf{(D.15)}$$

Thus, the sampled field distribution is given by $G_s(x,y)$:

$$G_s(x,y) \equiv G_s = \sum_{m=-\frac{M}{2}}^{\frac{M}{2}-1} \sum_{n=-\frac{N}{2}}^{\frac{N}{2}-1} rect\left(\frac{x}{P_x} - m, \frac{y}{P_y} - n\right) g[m,n]$$

$$= rect\left(\frac{x}{MP_x} + \frac{1}{2M}, \frac{y}{NP_y} + \frac{1}{2N}\right) \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} rect\left(\frac{x}{P_x} - m, \frac{y}{P_y} - n\right) g[m,n] \quad \textbf{(D.16)}$$

where,

$$rect(x,y) = \begin{cases} 1, & |x| < \frac{1}{2} \text{ e } |y| < \frac{1}{2} \\ 0, & \text{c. c} \end{cases} \quad \text{(D.17)}$$

We can rewrite (D.16) as

$$G_s(x,y) = rect\left(\frac{x}{MP_x}+\frac{1}{2M},\frac{y}{NP_y}+\frac{1}{2N}\right)\left[\sum_{m=-\infty}^{\infty}\sum_{n=-\infty}^{\infty}\delta(x-mP_x,y-nP_y)g[m,n]\right] \otimes rect\left(\frac{x}{P_x},\frac{y}{P_y}\right)$$

$$= rect\left(\frac{x}{MP_x}+\frac{1}{2M},\frac{y}{NP_y}+\frac{1}{2N}\right) G_d(x,y) \otimes rect\left(\frac{x}{P_x},\frac{y}{P_y}\right) \quad \text{(D.18)}$$

$G_d(x,y)$ is the discrete field distribution defined as

$$G_d(x,y) \equiv G_d = \sum_{m=-\infty}^{\infty}\sum_{n=-\infty}^{\infty}\delta(x-mP_x,y-nP_y)g[m,n] \quad \text{(D.19)}$$

Using (D.18) in (D.13), we obtain (for here on it is assumed that $\mathcal{F}$ is calculated at $\left(\frac{\xi}{\lambda_0 z},\frac{\eta}{\lambda_0 z}\right)$)

$$U_{G_s}(\xi,v) = A_+ f\ \mathcal{F}\left\{rect\left(\frac{x}{MP_x}+\frac{1}{2M},\frac{y}{NP_y}+\frac{1}{2N}\right)\right\} \otimes \left\{\left[\mathcal{F}\{G_d\}\mathcal{F}\left\{rect\left(\frac{x}{P_x},\frac{y}{P_y}\right)\right\}\right]\right\} \quad \text{(D.20)}$$

In (D.20) it was used the convolution theorem of the Fourier transform. Note that U_{G_s} is the field diffracted by a sampled field distribution. Evaluating the tabulated inverse Fourier transforms [3]

$$U_{G_s}(\xi,\eta) = A_+ f\ MNP_xP_y e^{j\left(\frac{\xi}{\lambda_0 z},\frac{\eta}{\lambda_0 z}\right)\cdot\left(\frac{P_x}{2},\frac{P_y}{2}\right)} sinc\left(\frac{\xi MP_x}{\lambda_0 z},\frac{\eta NP_y}{\lambda_0 z}\right) \otimes \left[\mathcal{F}\{G_d\}P_xP_y sinc\left(\frac{\xi P_x}{\lambda_0 z},\frac{\eta P_y}{\lambda_0 z}\right)\right] \quad \text{(D.21)}$$

Equation (D.21) contains already useful information regarding the diffraction by the sampled field. First, $\mathcal{F}\{G_d\}$ is just the discrete Fourier transform of G_d and we can relate it with $\mathcal{F}\{G\}$ by substituting (D. 15) in (D.19) and calculating the Fourier Transform

$$\begin{aligned} \mathcal{F}\{G_d\} &= \mathcal{F}\left\{\sum_{m=-\infty}^{\infty}\sum_{n=-\infty}^{\infty}\delta(x-mP_x,y-nP_y)G\left(mP_x,nP_y\right)f\left(mP_x,nP_y\right)\right\} \\ &= \mathcal{F}\left\{\left[\sum_{m=-\infty}^{\infty}\sum_{n=-\infty}^{\infty}\delta(x-mP_x,y-nP_y)\right] \otimes \left[G(x,y)f\ (x,y)\right]\right\} \\ &= \frac{(\lambda_0 z)^2}{P_xP_y}\sum_{m=-\infty}^{\infty}\sum_{n=-\infty}^{\infty}\delta\left(\xi-\frac{m\lambda_0 z}{P_x},\eta-\frac{n\lambda_0 z}{P_y}\right) \otimes \mathcal{F}\{Gf\} \\ &= \frac{(\lambda_0 z)^2}{P_xP_y}\sum_{m=-\infty}^{\infty}\sum_{n=-\infty}^{\infty}\mathcal{F}\{Gf\}\left(\frac{\xi}{\lambda_0 z}-\frac{m}{P_x},\frac{\eta}{\lambda_0 z}-\frac{n}{P_y}\right) \end{aligned} \quad \text{(D.22)}$$

Therefore, the Fourier transform of the discretized field distribution is obtained by an infinite superposition of displaced copies of the Fourier Transform $\mathcal{F}\{Gf\}$ by $\frac{\lambda_0 z}{P_x}$ and $\frac{\lambda_0 z}{P_y}$ in directions ξ and η, respectively. This is a commonly known result from digital signal processing. Therefore, we can take as our reduced reconstruction plane one of these repetitions with pixel sizes $P_\xi = \frac{\lambda_0 z}{MP_x}$ and $P_\eta = \frac{\lambda_0 z}{NP_y}$ in directions ξ and η, respectively. Furthermore, in (D.21), the term $\mathcal{F}\{G_d\}$ is multiplied by a sinc function that came from the finite size of the pixels (its effect on the calculated hologram will be discussed later). Finally, all these terms are convolved with another *sinc* that is skewed, that is due to the finite size of the hologram, and its effect will also be analyzed later.

Analogously, we can sample U with a matrix of square pixels sizing $P_\xi \times P_\eta$ along the directions ξ and η, respectively, on the Π plane. Again, the field at each pixel of the sampled distribution is assumed to have constant amplitude and phase given by

$$U[i,j] \equiv u[m,n] = U\left(mP_\xi, nP_\eta\right)\left[f\left(mP_\xi, nP_\eta\right)\right]^{-1} \quad \textbf{(D.23)}$$

Proceeding with the same steps made from (D.16) to (D.21) but using the diffraction equation (D.12) instead we arrive at

$$G_{U_s}(x,y) = A\_f\ MNP_\xi P_\eta e^{-j\left(\frac{x}{\lambda_0 z},\frac{y}{\lambda_0 z}\right)\cdot\left(\frac{P_\xi}{2},\frac{P_\eta}{2}\right)} sinc\left(\frac{xMP_\xi}{\lambda_0 z},\frac{yNP_\eta}{\lambda_0 z}\right) \otimes \left[\mathcal{F}^{-1}\{U_d\}P_\xi P_\eta sinc\left(\frac{xP_\xi}{z\lambda_0},\frac{yP_\eta}{z\lambda_0}\right)\right] \quad \textbf{(D.24)}$$

where $\mathcal{F}^{-1}\{\ \} \equiv \mathcal{F}^{-1}\left(\frac{x}{\lambda_0 z},\frac{y}{\lambda_0 z}\right)$ and U_d is defined as

$$U_d(x,y) \equiv U_d = \sum_{m=-\infty}^{\infty}\sum_{n=-\infty}^{\infty} \delta\left(x - mP_\xi, y - nP_\eta\right)u[m,n] \quad \textbf{(D.25)}$$

All the same observations of (D.21) applies to (D.24). Therefore,

$$\mathcal{F}^{-1}\{U_d\} = \frac{(\lambda_0 z)^2}{P_\xi P_\eta}\sum_{m=-\infty}^{\infty}\sum_{n=-\infty}^{\infty} \mathcal{F}^{-1}\{Uf^{-1}\}\left(\frac{x}{\lambda_0 z} - \frac{m}{P_\xi}, \frac{y}{\lambda_0 z} - \frac{n}{P_\eta}\right) \quad \textbf{(D.26)}$$

Until now, we sampled the field distributions at the hologram and reconstruction planes and found the diffraction equations by those fields. The resulting diffracted field still needs to be discretized to be used for calculations in a computer, which will be done in the next section.

D.2.2. Discretizing the diffraction equations

We need to discretize equations (D.21) and (D.24) to relate them with each other and calculate them with the aid of the FFT. Let's define the discrete version of $G_{U_s}(x,y)$ as

$$G_{U_s}^{d}(x,y) = G_{U_s}(x,y)\sum_{k=-\frac{M}{2}}^{\frac{M}{2}-1}\sum_{l=-\frac{N}{2}}^{\frac{N}{2}-1}\delta\left(x-kP_x, y-lP_y\right)$$

$$= \sum_{k=-\frac{M}{2}}^{\frac{M}{2}-1}\sum_{l=-\frac{N}{2}}^{\frac{N}{2}-1}\delta\left(x-kP_x, y-lP_y\right)G_{U_s}(kP_x, lP_y) \quad \textbf{(D.27)}$$

and the discrete version of $U_{G_s}(\xi,\eta)$ as

$$U_{G_s}^{d}(\xi,\eta) = U_{G_s}(\xi,\eta)\sum_{k=-\frac{M}{2}}^{\frac{M}{2}-1}\sum_{l=-\frac{N}{2}}^{\frac{N}{2}-1}\delta\left(\xi-kP_\xi, \eta-lP_\eta\right)$$

$$= \sum_{k=-\frac{M}{2}}^{\frac{M}{2}-1}\sum_{l=-\frac{N}{2}}^{\frac{N}{2}-1}\delta\left(\xi-kP_\xi, \eta-lP_\eta\right)U_{G_s}(kP_\xi, lP_\eta) \quad \textbf{(D.28)}$$

Substituting (D.28) in (D.12), we obtain

$$G_d'(x,y) = A_-[f(x,y)]^{-1}\mathcal{F}^{-1}\left\{\sum_{k=-\frac{M}{2}}^{\frac{M}{2}-1}\sum_{l=-\frac{N}{2}}^{\frac{N}{2}-1}\delta(\xi-kP_\xi, \eta-lP_\nu)\ U_{G_s}(kP_\xi, lP_\eta)\right\}\left(\frac{x}{\lambda_0 z}, \frac{y}{\lambda_0 z}\right)$$

$$= A_-[f(x,y)]^{-1}\sum_{k=-\frac{M}{2}}^{\frac{M}{2}-1}\sum_{l=-\frac{N}{2}}^{\frac{N}{2}-1}U_{G_s}(kP_\xi, lP_\eta)\, e^{j2\pi\left(\frac{x}{\lambda_0 z}, \frac{y}{\lambda_0 z}\right)\cdot(kP_\xi, lP_\eta)} \quad \textbf{(D.29)}$$

Substituting (D.27) in (D.13) we obtain

$$U'_d(\xi,\nu) = A_+ f(\xi,\eta)\mathcal{F}\left\{\sum_{k=-\frac{M}{2}}^{\frac{M}{2}-1}\sum_{l=-\frac{N}{2}}^{\frac{N}{2}-1}\delta\left(x-kP_x, y-lP_y\right)G_{U_s}(kP_x,lP_y)\right\}\left(\frac{\xi}{\lambda_0 z},\frac{\eta}{\lambda_0 z}\right)$$

$$= A_+ f(\xi,\eta)\sum_{k=-\frac{M}{2}}^{\frac{M}{2}-1}\sum_{l=-\frac{N}{2}}^{\frac{N}{2}-1}G_{U_s}(kP_x,lP_y)\,e^{-j2\pi\left(\frac{\xi}{\lambda_0 z},\frac{\eta}{\lambda_0 z}\right)\cdot(kP_x,lP_y)} \quad \textbf{(D.30)}$$

Note that G'_d and U'_d are the diffraction of the discrete field distributions U and G, respectively. Finally, we discretize $\mathrm{G'_d(x,y)}$ and $U'_d(\xi,\nu)$ by defining the discrete variables $G'_d[m,n]$ and $U'_d[m,n]$ as, respectively

$$G'_d[m,n] \equiv G'_d(mP_x,nP_y) \quad \textbf{(D.31)}$$

and

$$U'_d[m,n] \equiv U'_d(mP_\xi,nP_\eta) \quad \textbf{(D.32)}$$

Substituting (D.29) in (D.31) and (D.30) in (D.32) we obtain the following relations, respectively

$$G'_d[m,n] = A_-\left[f(mP_x,nP_y)\right]^{-1}\sum_{k=-\frac{M}{2}}^{\frac{M}{2}-1}\sum_{l=-\frac{N}{2}}^{\frac{N}{2}-1}U_{G_s}(kP_\xi,lP_\eta)\,e^{j2\pi\left(\frac{mP_x}{\lambda_0 z},\frac{nP_y}{\lambda_0 z}\right)\cdot(kP_\xi,lP_\eta)} \quad \textbf{(D.33)}$$

and

$$U'_d[m,n] = A_+(mI_x,nP_y)\sum_{k=-\frac{M}{2}}^{\frac{M}{2}-1}\sum_{l=-\frac{N}{2}}^{\frac{N}{2}-1}G_{U_s}(kP_x,lP_y)\,e^{-j2\pi\left(\frac{I_x m}{\lambda_0 z},\frac{I_y n}{\lambda_0 z}\right)\cdot(kP_x,lP_y)} \quad \textbf{(D.34)}$$

But $U_{G_s}\left(\mathrm{kP_\xi, lP_\eta}\right)$ is the field distribution at point $\left(\mathrm{kP_\xi, lP_\eta}\right)$ in the reconstruction plane obtained from the sampled version of the hologram distribution (G_s). If the Nyquist theorem holds on (D.22), we could assume that it reproduces the field distribution at (D.21):

$$U_{G_s}(kI_x,lI_y) = U(kI_x,lI_y) \quad \textbf{(D.35)}$$

Analogously, $G_{U_s}(kP_x,lP_y)$ is the field distribution at point (kP_x,lP_y) on the hologram plane obtained from the sampled version of the hologram distribution (U_s). If the Nyquist theorem holds on (D.26), we could assume that $G_{U_s}(x,y)$ reproduces the field distribution at (D.24):

$$G_{U_s}(kP_x, lP_y) = G(kP_x, lP_y) \tag{D.36}$$

Therefore, as the Nyquist theorem is respected for both field distributions by hypothesis, it is satisfied at (D.29) and (D.30) as well. Therefore,

$$U'_d[m,n] = U(mI_x, nI_y) \tag{D.37}$$

and

$$G'_d[m,n] = G(mP_x, nP_y) \tag{D.38}$$

Thus, equations (D.33) and (D.34) can be rewritten as a Discrete Fourier Transform (DFT) pair:

$$G'_d[m,n] = A_-\left[f(mP_x, nP_y)\right]^{-1} \sum_{k=-\frac{M}{2}}^{\frac{M}{2}-1} \sum_{l=-\frac{N}{2}}^{\frac{N}{2}-1} U'_d[k,l]\, e^{j2\pi \frac{(m,n)}{MN}\cdot(k,l)} \tag{D.39}$$

and

$$U'_d[m,n] = A_+ f(mP_\xi, nP_\eta) \sum_{k=-\frac{M}{2}}^{\frac{M}{2}-1} \sum_{l=-\frac{N}{2}}^{\frac{N}{2}-1} G'_d[k,l]\, e^{-j2\pi \frac{(m,n)}{MN}\cdot(k,l)} \tag{D.40}$$

Note that the relations $P_\xi = \frac{\lambda_0 z}{MP_x}$ and $P_\eta = \frac{\lambda_0 z}{NP_y}$ were used to obtain the DFTs. Therefore, we can calculate the diffraction of the fields via equations (D.39) and (D.40) in a computer by using the Fast Fourier Transform (FFT) algorithm to compute the DFTs. These equations will be used to implement the iterative Fourier transform algorithm (IFTA) to calculate the hologram (matrix G_{MxN}) of a given discretized field distribution (matrix U_{MxN}) .

D.2.3. Diffraction of the sampled hologram

For now, let's suppose that we are given a discretized field distribution (U_{MxN} could be a digital image for instance) and obtain the discrete hologram matrix by (D.39). If we implement this hologram on a wavefront modulator and illuminate it with a coherent light source, how the reconstructed field relates to U_{MxN}? To answer this question, it will be assumed that the wavefront modulator has a transfer function given by (D.16). Therefore, its diffraction is given by (D.21). Substituting (D.39) on (D.19) and the result on (D.21) we have

$$U_R(\xi,\eta) = A_+ f\ MNP_xP_y e^{j\left(\frac{\xi}{\lambda_0 z},\frac{\eta}{\lambda_0 z}\right)\cdot\left(\frac{P_x}{2},\frac{P_y}{2}\right)} sinc\left(\frac{\xi MP_x}{\lambda_0 z},\frac{\eta NP_y}{\lambda_0 z}\right)$$

$$\otimes\left[\mathcal{F}\left\{\sum_{k=-\frac{M}{2}}^{\frac{M}{2}-1}\sum_{l=-\frac{N}{2}}^{\frac{N}{2}-1}\delta\left(x-kP_x, y-lP_y\ \right)G_d'[k,l]\right\}P_xP_y sinc\left(\frac{\xi P_x}{\lambda_0 z},\frac{\eta P_y}{\lambda_0 z}\right)\right] \quad \textbf{(D.41)}$$

But

$$\mathcal{F}\left\{\sum_{i=-\frac{M}{2}}^{\frac{M}{2}-1}\sum_{j=-\frac{N}{2}}^{\frac{N}{2}-1}\delta\left(x-iP_x, y-jP_y\ \right)G_d'[i,j]\right\} = \sum_{i=-\frac{M}{2}}^{\frac{M}{2}-1}\sum_{j=-\frac{N}{2}}^{\frac{N}{2}-1} e^{-j\left(\frac{\xi}{\lambda_0 z},\frac{\eta}{\lambda_0 z}\right)\cdot\left(iP_x jP_y\right)}G_d'[i,j]$$

$$\overset{(D.30)}{=} U_d(\xi,\nu)[f(\xi,\eta)]^{-1}A_+^{-1} \quad \textbf{(D.42)}$$

Substituting (D.42) in (D.41)

$$U_R(\xi,\eta) = f\ MN\left(P_xP_y\right)^2 e^{j\left(\frac{\xi}{\lambda_0 z},\frac{\eta}{\lambda_0 z}\right)\cdot\left(\frac{P_x}{2},\frac{P_y}{2}\right)} sinc\left(\frac{\xi MP_x}{\lambda_0 z},\frac{\eta NP_y}{\lambda_0 z}\right)\otimes\left[U_d(\xi,\nu)[f(\xi,\eta)]^{-1} sinc\left(\frac{\xi P_x}{\lambda_0 z},\frac{\eta P_y}{\lambda_0 z}\right)\right] \quad \textbf{(D.43)}$$

We can analyze the effect of those *sincs* on the reconstruction by evaluating them with the discrete distribution $\mathrm{U_d}(\xi,\nu)$. The inner *sinc* (that is multiplied to U_d directly), is due to the finite size of the pixel and has the effect of modulate the amplitude of the reconstruction. As we saw in the previous section, the sampling of the hologram plane causes the reconstruction to be a superposition of the diffraction separated by $\frac{\lambda_0 z}{P_x}$ and $\frac{\lambda_0 z}{P_y}$ at direction ξ and η, respectively. Therefore, to satisfy the Nyquist theorem, $\mathrm{U_d}(\xi,\nu)$ is limited in the rectangle given by

$$\begin{cases} -\dfrac{\lambda_0 z}{2P_x} < \xi < \dfrac{\lambda_0 z}{2P_x} \\ -\dfrac{\lambda_0 z}{2P_y} < \eta < \dfrac{\lambda_0 z}{2P_y} \end{cases} \quad \textbf{(D.44)}$$

This region is inside the first lobe of the internal *sinc*, which decreases as $|\xi|$ and/or $|\eta|$ increases. Therefore, its minimum value occurs at a vertex of rectangle the (D.44). Therefore, at $(\xi,\eta) = \left(-\frac{\lambda_0 z}{2P_x},-\frac{\lambda_0 z}{2P_y}\right)$ it values $\mathrm{sinc}\left(\frac{1}{2},\frac{1}{2}\right) = \frac{4}{\pi^2} \cong 0.4$. This issue can be circumvented by pre-modulating $\mathrm{U_d}(\xi,\nu)$ with $\left(\mathrm{sinc}\left(\frac{\xi \mathrm{MP_x}}{\lambda_0 z},\frac{\eta \mathrm{NP_y}}{\lambda_0 z}\right)\right)^{-1}$.

The effect of the external *sinc* (that is convoluting U_d) to the reconstruction has another interesting effect on the reconstruction. Supposing that the inner *sinc* was corrected and substituting (D.25) in (D.43), we have

$$U_R(\xi,\eta) = f\,MN\left(P_xP_y\right)^2 e^{j\left(\frac{\xi}{\lambda_0 z},\frac{\eta}{\lambda_0 z}\right)\cdot\left(\frac{P_x}{2},\frac{P_y}{2}\right)} sinc\left(\frac{\xi M}{l_x},\frac{\eta N}{l_y}\right) \otimes \left[\sum_{k=-\frac{M}{2}}^{\frac{M}{2}-1}\sum_{l=-\frac{N}{2}}^{\frac{N}{2}-1} \delta\left(\xi - kP_\xi, \eta - lP_\eta\ \right) U_d[k,l]\right]$$

$$= f\,MN\left(P_xP_y\right)^2 e^{j\left(\frac{\xi}{\lambda_0 z},\frac{\eta}{\lambda_0 z}\right)\cdot\left(\frac{P_x}{2},\frac{P_y}{2}\right)} \sum_{k=-\frac{M}{2}}^{\frac{M}{2}-1}\sum_{l=-\frac{N}{2}}^{\frac{N}{2}-1} sinc\left(M\frac{\xi - kP_\xi}{P_\xi}, N\frac{\eta - lP_\eta}{P_\eta}\right) U_d[k,l] \qquad \textbf{(D.45)}$$

Therefore, the reconstruction of the sampled hologram is the superposition of *sinc* functions on the sample grid of the reconstruction plane modulated by the discretized field matrix U_{MxN}. For fixed P_ξ and P_η, increasing the size of the hologram, i.e., *M* and *N*, the *sinc* function gets skewed and the reconstruction seems as a sum of tiny dots. These results were numerically and experimentally verified on the optical photonic metasurface project (see *Chapter III.1.6.c) Reconstructions - Quartz).*

D.3. Iterative Fourier Transform Algorithm (IFTA)

Equations (D.39) and (D.40) are already the simplified diffraction equations that can be used to calculate a hologram and its reconstruction under either Fourier or Fraunhofer approximations. However, the resulting hologram is a complex-valued matrix whose phase and amplitude are only constrained by the floating-point representation of the complex numbers on the computer. From a computational point of view this pose no problem at all, but it is a challenge to be reproduced by a wavefront modulator. It is a difficult task to implement discrete optical elements (DOE) that have a full phase control in the range [0-2π] with also amplitude modulation capability. In practice, most DOEs rely on the phase control only, which is already not an easy task to do as well, with only a finite number of phase levels. Therefore, the hologram must be encoded in such a way that it has constant amplitude and only a finite number of phase levels, so it can be well reproduced by the wavefront modulator. This necessity gave rise to the field of computer-generated holograms (CGH) algorithms as briefly described in section *I.3 From classical to digital holography* of the main text. Here it will be described the Iterative Fourier Transform Algorithm (IFTA) that is based on [75].

The task of the algorithm is to obtain a phase distribution matrix G_{kl}, that is the hologram, with constant and unitary amplitude whose diffraction at the reconstruction plane, g_{mn}, have an intensity distribution that resembles a known field pattern given by U_{mn}. The discrete version of the hologram, Σ, and reconstruction, Π, planes are represented in Figure D.2. Inside the reconstruction plane, in Figure D.2 (b), the target field distribution is placed inside a region Ω, called image window. This is necessary to increase the degree of freedom of the method and increase the final reconstruction quality [75].

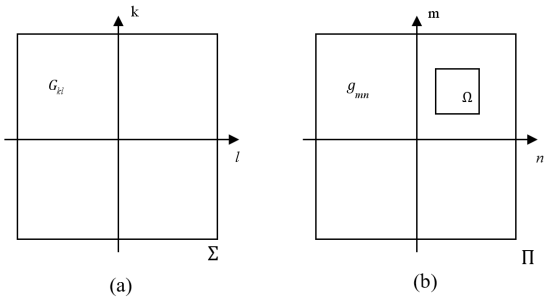


Figure D.2 – Discrete representations of the hologram (a) and reconstruction (b) planes. Each one is a matrix with MxN elements representing the field distributions. In (b), Ω represents the image window where the target field pattern U_{mn} is inserted.

Based on those definitions, the following figures of merit, listed on Table D-1, are defined to assess the quality of the calculated reconstructions

Table D-1 - List of figures of merit to account for the quality of the reconstructions

Mean squared error (MSE)	$e = \sum_{\Omega} [\lvert U\rvert - \lvert g'\rvert]^2$	**(D.46)**
Image efficiency (IE)	$\rho = \frac{\sum_{\Omega} \lvert g'\rvert^2}{\sum_{\Pi} \lvert g'\rvert^2}$	**(D.47)**
SNR	$SNR = \frac{\sum_{\Omega} \lvert g'\rvert^2}{e} = \frac{\sum_{\Omega} \lvert g'\rvert^2}{\sum_{\Omega} [\lvert U\rvert - \lvert g'\rvert]^2}$	**(D.48)**

The mean squared error gives a quantification of the inserted noise energy into the reconstructed field due the phase quantization. The image efficiency gives the resulting energy that is being transmitted to the image window. Finally, the signal to noise ratio (SNR), is self-explanatory, and will be used as our main figure of merit.

The algorithm starts by inserting the target image in the image window with all unoccupied positions on the reconstruction plane initialized as zero. Then, a random phase noise is added to this field distribution followed by the calculation of the inverse diffraction of the resulting matrix, that is g_{mn}. As explained in section *D.2.2 Discretizing the diffraction equations* of the Appendix, we will use the use the Fast-Fourier transform (FFT) pair to calculate the diffraction between the fields at each plane. Therefore, the hologram matrix, G_{kl}, is obtained by Fast Fourier Transforming g_{mn}. Then, the phase of G_{kl} is quantized while the amplitude is normalized to a constant value and the result is stored in G'_{kl}. Subsequently, G'_{kl} is Inverse Fast Fourier transformed to the reconstruction plane [73, 75]. The resulting matrix, g'_{mn} (containing the reconstructed field with quantized phases and normalized amplitudes) is then reinforced at the image window with the original image multiplied by a scale factor, therefore increasing the SNR, given by [75]

$$\alpha = \frac{\sum_{\Omega} \lvert g'\rvert^2}{\sum_{\Omega} \lvert U g'\rvert} \qquad \textbf{(D.49)}$$

Thus, the result is stored on matrix g_{mn} that will be inverse Fast Fourier Transformed to obtain G_{kl} that will be again quantized, and the previous steps repeated iteratively as shown in Figure D.3 [75]. These processes are carried out iteratively until the phase quantization is complete. The phase quantization at the plane of the hologram is not carried out in one step [173]. Instead, it makes use of a stepwise operator that restricts the allowed phase values on the hologram in each iteration. In the first iteration, all possible phase levels are allowed and the number of levels is iteratively reduced until a discrete number of values is obtained after *k* iterations [173]. This process increases the degree-of-freedom of the algorithm resulting in better efficiency and SNR [75, 173]. At each step, the three figures of merit given by equations (D.46) - (D.48) are

evaluated. Note that, during the algorithm evaluation, the phase of g_{mn} can have any value in Ω while the noise field outside this region can have any value, which is exploited as another degree of freedom for the algorithm [75]. This is made by forcing $g'_{mn} = 0, \forall\ (m,n) \notin \Omega$ for the first c iterations and then letting it vary freely. Therefore, the hologram calculation is performed several times in a sweep with increasing values of c in the range $[1, \psi], with\ 1 \leq \psi \leq k$. For each CGH, all the figures of merit are evaluated and stored in the memory. Then, the last value of SNR is compared with the biggest one, stored in a dummy variable that is initialized with zero. If it is bigger, the dummy variable is updated to the last value, and the calculated hologram is stored in a dummy matrix (used to store the best hologram in terms of SNR, note that it will starts with the hologram for *c=1* because the evaluated SNR is always greater than 0). Then, if $c < k$, c is increased, and the hologram is recalculated, otherwise the loop halts. On the other hand, if the last SNR is 10% smaller than the best case, the algorithm stops, and the convergence is achieved. This criterion will be justified on section *D.3.1 Example: Fourier hologram calculation*. Finally, if it is smaller, but by less than 10%, the loop is performed again if $c < k$. The algorithm is represented in a flow diagram in Figure D.3.

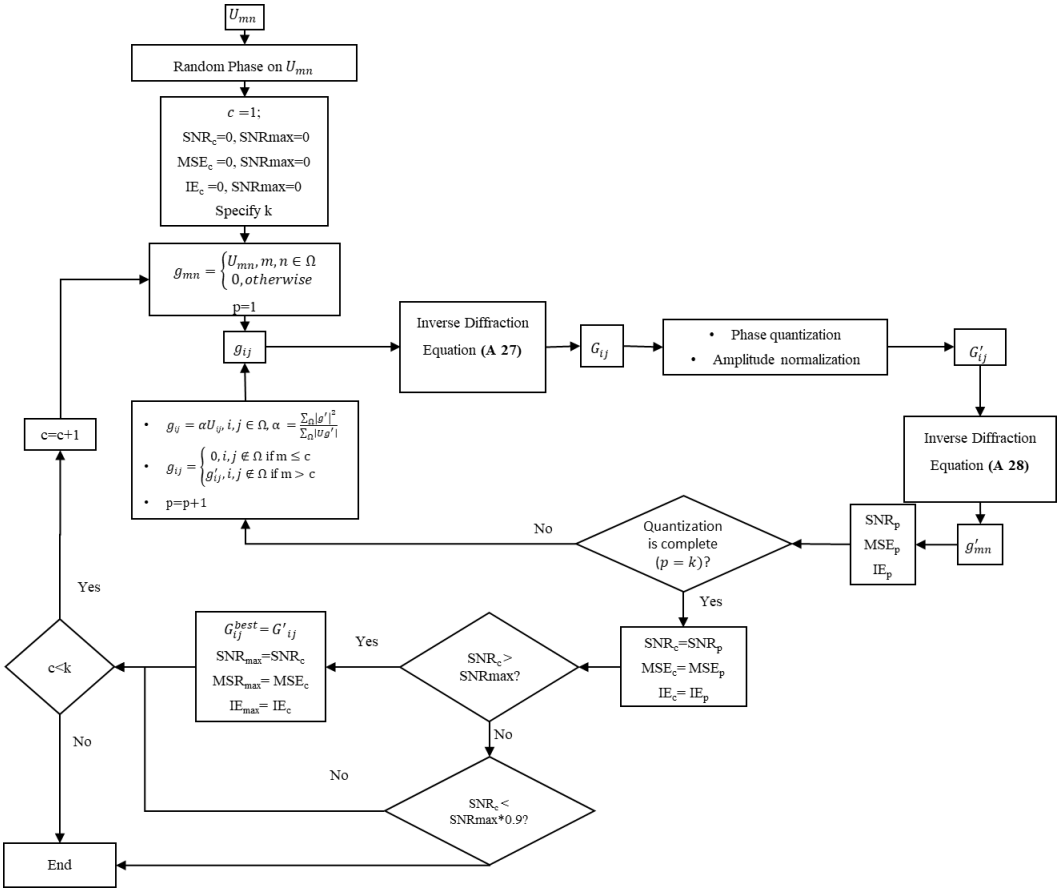


Figure D.3 - Flow diagram of the implemented IFTA. Adapted from [75].

D.3.1. Example: Fourier hologram calculation

In this section, several CGHs will be calculated using the IFTA algorithm. All holograms, except the ones with two phase levels, will be obtained from the same field distribution U_{mn}, which corresponds to the logo of the EESC (*Escola de Engenharia de São Carlos),* shown in Figure D.4 (a). For two phase levels, the hologram is real, therefore its reconstruction must be Hermitian symmetric because its diffraction is given by a FFT. Thus, the image window was displaced from the reconstruction plane origin to avoid superposition between the image and its Hermitian twin. By setting different discrete phase levels and number of quantization steps the resulting holograms will be analyzed in terms of the SNR and IE. Therefore, it will be justified the convergence criterion, in terms of SNR, proposed in the last section. This will be done by analyzing the behavior of the resulting SNR $\forall\, c \in [1, k]$, that is, by disabling the convergence criteria. The resolutions of the reconstruction and hologram plane are both equal 256x256 and of the image window is 111x105. For Fourier holograms, there is no need to specify the operating wavelength, distance between the planes neither the pixel sizes because the quadratic phase function *f*, is equal to one. Therefore, the field at one plane is obtained by Fast Fourier Transforming the other. Nonetheless, these parameters can be specified at the fabrication step, provided that the Fraunhofer approximation is valid, and will affect both its size and the size of its reconstruction.

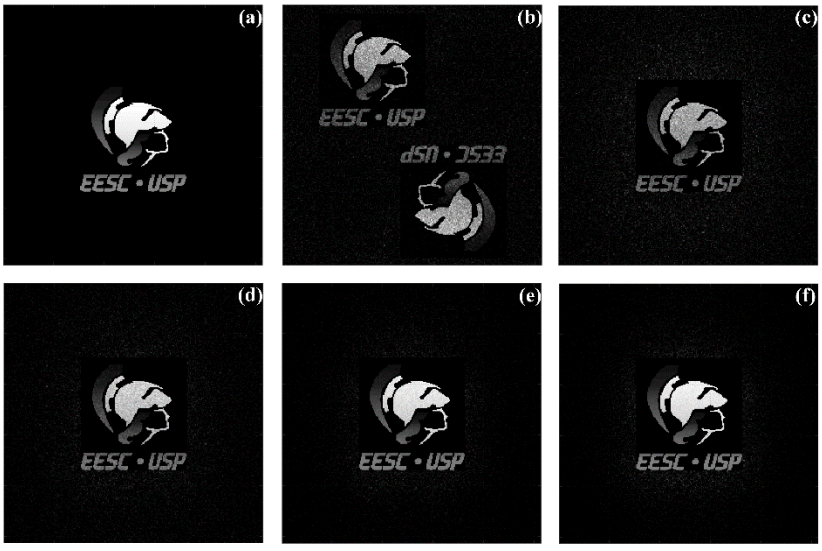


Figure D.4 – (a) Target field distribution at the reconstruction plane. (b)-(f) Show the reconstructed image from the algorithm obtained with 2,3,4,8 and 16 phase levels for $k = 100$. Note that for 2 phase levels the reconstruction presents Hermitian Symmetry because the Hologram is real. Therefore, the image window was displaced to avoid superposition of the target image with its Hermitian twin.

The resulting SNR [IE] is shown as function of c/k at Figures D.5 (a),(c),(e),(g) and (i) [(b),(d),(f),(h) and (j)] for 2, 3, 4, 8 and 16 phase levels, respectively. For each case, the quantization steps (k) of 5, 20, 50 and 100 are shown in black, blue, red, and green, respectively. The diamonds at each line mark the maximum SNR for each case, that is the target of the algorithm. Note that, for this target image, using 2, 3, 4 and 8 phase levels, increasing c/k makes the SNR reaches a maximum then start to reduce. For the other cases, the SNR is already at maximum for $c = 1$. The presence of only one maximum for SNR as function of c/k, happened even for other target images with different hologram dimensions. Therefore, if the SNR is, in average, decreasing as function of c/k, we can infer that the stored maximum is in fact the best hologram in terms of SNR. As the SNR function is not monotonic, see Figure D.5 (a) for instance, it is not possible to rely on its derivative to tell if the maximum was reached. Nevertheless, based on the calculation of several other holograms, when the SNR decreases by more than 10% it indicates that the maximum has already been reached and the function is at the decrescent portion. Thus, the algorithm can be halted if the actual SNR is more than 10% smaller than the highest one, that is stored in the memory.

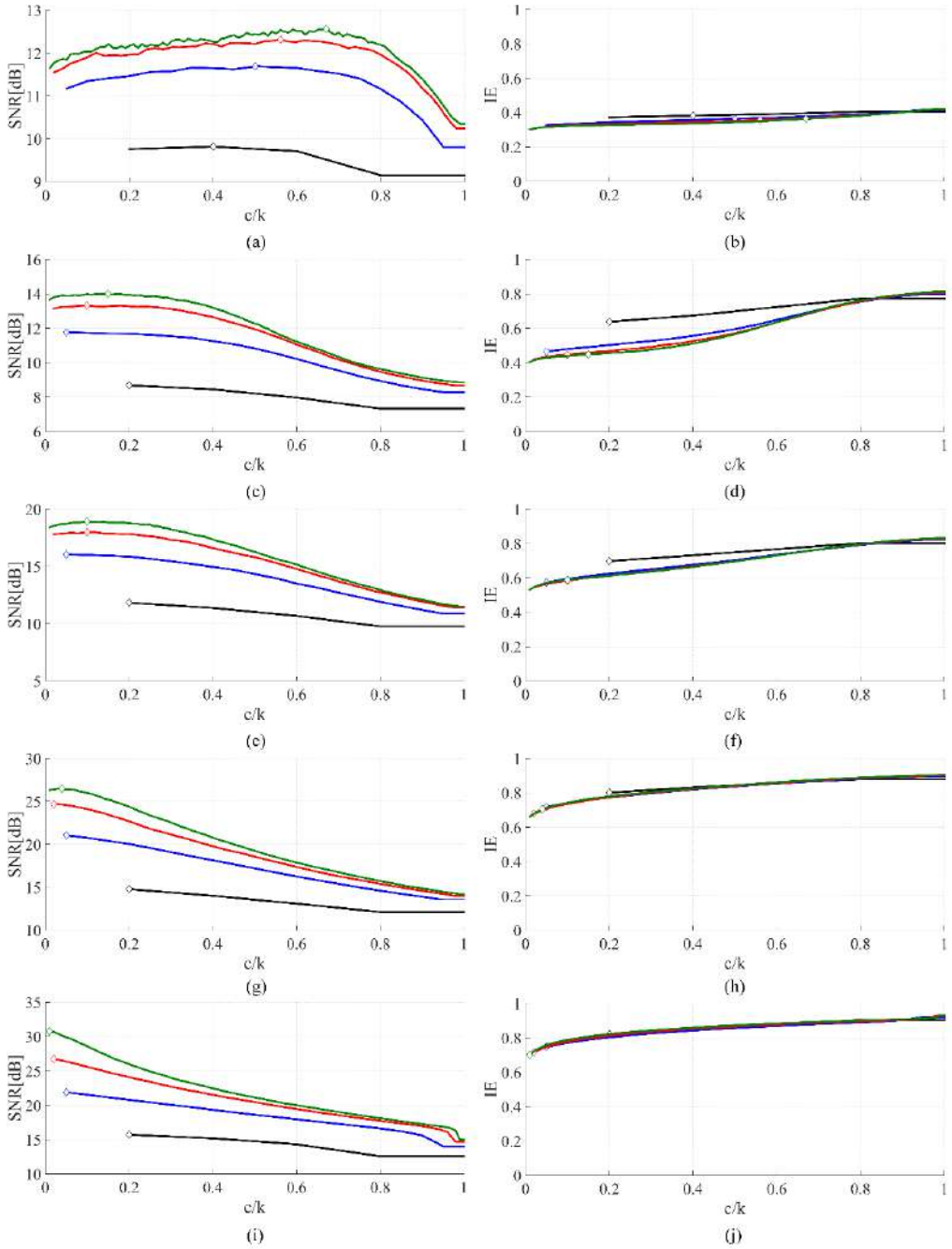


Figure D.5 - (a), (c), (e), (g) and (i) [(b),(d),(f),(h) and (j)] show the SNR [IE] as function of the ratio c/k for 2, 3, 4, 8 and 16 phase levels, respectively. For each case, the quantization steps (k) of 5, 20, 50 and 100 are shown in black, blue, red, and green, respectively. The diamonds at each line mark the maximum SNR for each case, that is the target of the algorithm. Note that for two phase levels, the Hermitian is treated as noise and does not enter in the efficiency calculation.

D.4. Rigorous Rayleigh Sommerfeld (RS) Diffraction

In the previous section, the diffraction was performed assuming the paraxial approximation validity and applied to IFTA. The advantage of this procedure is that diffraction becomes a Fourier Transform which can be performed almost instantaneously with FFT depending on the hologram size. Nevertheless, the designed metasurfaces, that will be used as wavefront modulator for the holograms, have subwavelength pixel sizes. Therefore, the diffracted image will be wide-angle, thus violating the paraxial approximation. This observation is evident from (D.44), which shows the dimensions of the reconstruction plane. If we assume that $P_x = P_y = P < \lambda_0$, the opening angle, α, of the reconstruction will satisfy

$$\alpha = \operatorname{atan}\left(\frac{\max(\xi)}{z}\right) = \operatorname{atan}\left(\frac{\lambda_0}{2P}\right) \geq \operatorname{atan}\left(\frac{1}{2}\right) \cong 26^o$$

where it was used the fact the *atan* function is crescent. Therefore, the reconstruction obtained from the IFTA algorithm will not represent the physical diffraction of the metasurface containing that hologram. This is because the procedure of Fourier transforming the near-field (hologram) to obtain the far-field (reconstruction) assumes the latter sufficiently far away from the former so that its projection plane can be calculated as a spherical surface whose origin is at the hologram position [2]. In the Fourier diffraction it is further assumed that the projection plane is planar, which is valid only in the paraxial approximation. Therefore, if the hologram is calculated using the Fourier transform and its period is small enough for the paraxial approximation to be used, the reconstruction will be formed at the surface of a sphere and not at a planar screen as we would expect. This issue can be overcome by mapping the target image onto a spherical surface in the hologram design via a coordinate transformation [91] and causes the reconstruction on a spherical surface to be distorted but it correctly reconstructs the image on a planar screen.

For instance, Figures D.6 (a)-(c) show the target field distribution, the obtained Fresnel diffraction and the RS diffraction at the reconstruction plane, respectively, of an 8-phase level CGH obtained by the IFTA with $c = 100$. The RS integration was performed assuming an operating wavelength of 532 nm, hologram pixel size of 190 nm and the reconstruction plane at the Fraunhofer region. For these conditions, the opening angle of the reconstruction plane is of $54^o \times 54^o$, which is far from the paraxial approximation. Therefore, the RS reconstruction at Figure D.6 (c) gets distorted at distant regions from the origin by comparing it with the target field distribution at Figure D.6 (a). On contrary, this feature does not happen for Fresnel

Diffraction, shown at Figure D.6 (b), whose reconstruction resembles the target image, at Figure D.6 (a), without distortion.

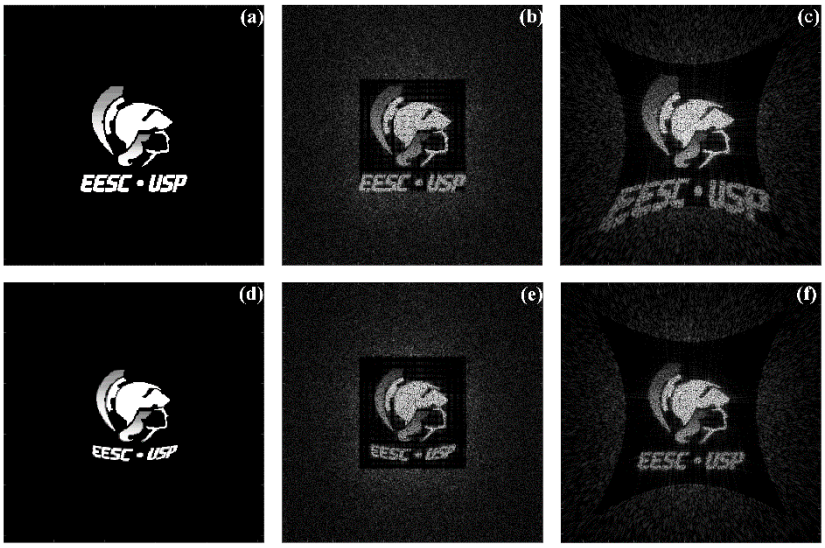


Figure D.6 – (a) and (d) show the original and the wide-angle corrected, respectively, target reconstruction planes. (b) and (e) [(c) and (f)] show the Fresnel [RS] Diffractions of (a) and (d), respectively. The operating wavelength is 532 nm.

To correct this issue, the wide-angle correction was applied at the reconstruction plane and the resulting field distribution, shown at Figure D.6 (d), was used to calculate a new CGH. The resulting Fresnel and RS diffractions of the fixed hologram are shown at Figures D.6 (e) and (f), respectively. Note that now the roles have inverted: the RS reconstruction image is less distorted while the Fresnel one is more distorted, when compared with Figure D.6 (a). Because the RS integration compensates the wide-angle correction while the Fresnel diffraction does not and resembles the wide-angle corrected image of Figure D.6 (d).

D.4.1. RS Integration

The RS needs to be performed for each point at the reconstruction plane separately, which can take from several hours to few days depending on the hologram size and the required resolution of the reconstruction. In the search of minimizing the computational effort it was discovered that the slowest part of the calculation is the creation of the matrix containing the term $e^{jk_0 r(\xi,\eta;x,y)}$ for each point at the reconstruction plane. Nevertheless, the distance function between two points at each plane $r(\xi,\eta;x,y) \equiv r$, defined at (D.2), may possess mirror symmetries with respect to some axes as will be explained in the next section. This observation allows to use the same exponential matrix for two or eight different points depending on the symmetry group of the function, reducing the necessary calculation time up to two and five times, respectively.

This section starts by analyzing the symmetries properties of the RS integration. Then, the implementation of RS integration without resorting to any symmetry of r is used followed by the

implementation of the integration with the symmetries. Then, the performance of each case will be compared and an equation for estimating the integration time of each one will be proposed. The section closes by describing a way of calculating the energy of the diffracted field.

a. Symmetry properties of the RS integration

We can calculate the reconstruction at any point in Σ, by simply specifying its coordinates. Therefore, lets define the observation region as a square in Σ centered at $(\xi - \Delta\xi, \eta - \Delta\eta)$, as shown in Figure D.7 (a)**.** It will be shown that when $\Delta\xi$ ($\Delta\eta$) is zero the function *r* possesses mirror symmetry with respect the axis $\xi = 0$ ($\eta = 0$). Furthermore, if both are zero it will have mirror symmetry with respect the axis $\xi = 0, \eta = 0, \eta = \xi \; and \; \eta = -\xi$. The RS diffraction equation (D.3) will be rewritten in terms of a coordinate system centered at the observation region, with coordinates $\xi' = \xi - \Delta\xi$ and $\eta' = \eta - \Delta\eta$:

$$U(\xi', \eta') = \frac{z}{j\lambda_0} \int \int_{\Pi} G(x, y) \frac{e^{jk_0 r(\xi', \eta'; x, y)}}{r^2(\xi', \eta'; x, y)} \; dxdy \qquad \textbf{(D.50)}$$

Therefore, it will be stated three symmetry properties of the RS integral:

- **S1**: If $\Delta\xi = 0$ and $\Delta\eta \neq 0$, then we can calculate $U(\xi', \eta')$ and $U(-\xi', \eta')$ by using the same function $\frac{e^{jk_0 r(\xi', \eta'; x, y)}}{r^2(\xi', \eta'; x, y)}$ in the RS integrand. This property arises because $r(\xi, \eta'; x, y) = r(-\xi, \eta'; -x, y)$, as depicted at Figure D.7 (b).

- **S2**: If $\Delta\xi \neq 0$ and $\Delta\eta = 0$, then we can calculate $U(\xi', \eta')$ and $U(\xi', -\eta')$ by using the same function $\frac{e^{jk_0 r(\xi', \eta'; x, y)}}{r^2(\xi', \eta'; x, y)}$ in the integrand. This property arises because $r(\xi, \eta'; x, y) = r(\xi, -\eta'; x, -y)$, as depicted at Figure D.7 (c).

- **S3**: If $\Delta\xi = 0$ and $\Delta\eta = 0$ then, we can calculate $U(\sigma\xi', \sigma\eta'), U(\sigma\eta', -\sigma\xi'), U(\sigma\eta', \sigma\xi') \; and \; U(\sigma\eta', -\sigma\xi'), \; \sigma = \pm 1$ by using the same function $\frac{e^{jk_0 r(\xi', \eta'; x, y)}}{r^2(\xi', \eta'; x, y)}$ in the integrand. This property arises because $r(\sigma\xi \, , \sigma\eta \, ; \sigma x, \sigma y) = r(-\sigma\xi \, , \sigma\eta \, ; -\sigma x, \sigma y) = r(\sigma\eta \, , \sigma\xi \, ; \sigma y, \sigma x) = r(-\sigma\eta \, , \sigma\xi \, ; -\sigma y, \sigma x)$, as depicted at Figure D.7 (d).

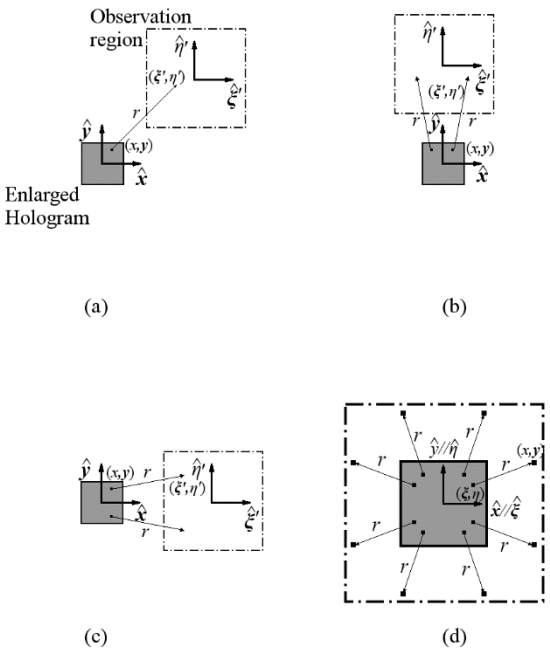


Figure D.7 – Representation of the RS Diffraction regions of integration (enlarged hologram) and of observation window(dotted-dashed square) for (a) $\Delta\xi \neq 0$ *and* $\Delta\eta \neq 0$, (b) $\Delta\xi = 0$ *and* $\Delta\eta \neq 0$, (c) $\Delta\xi \neq 0$ *and* $\Delta\eta = 0$ and (d) $\Delta\xi = \Delta\eta = 0$. $\Delta\xi$ *and* $\Delta\eta$ are the displacements of the center of the observation window with respect to the center of the reconstruction, or equivalently hologram, plane. In (b), the points *(x,y)* and (ξ', η') were mirrored with respect the axes $\xi' = 0$ and $x = 0$, respectively, but the coordinates were omitted to clean the figure. In (c), the points are also mirrored but with respect the axes $\eta' = 0$ and $y = 0$. Note that the distance between the original pair of points and the mirrored ones is the same. In (d), the hologram and the reconstruction window were enlarged to avoid overcrowding of information. Furthermore, the point *(x,y)* was mirrored with respect the axes $x = 0, y = 0, x = -y, x = y, x = y$ followed by $x = 0$ *or* $y = 0$ *or* $x = -y$ while the point (ξ', η') with respect the axes $\xi = 0, \eta = 0, \xi = -\eta, \xi = \eta, \xi = \eta$ followed by $\xi = 0$ *or* $\eta = 0$ *or* $\xi = -\eta$. Note that the distance, *r*, between the original pair of points and the mirrored ones is the same.

Proofs:

- S1:

Assumption: $\Delta\xi = 0,\ \xi = \xi'$

The diffraction at $(-\xi, \eta')$ is given by, using (D.50),

$$U(-\xi, \eta') = -\frac{z}{j\lambda_0}\int\int_{\Pi} G(u,y)\frac{e^{jk_0 r(-\xi,\eta';-x,y)}}{r^2(-\xi,\eta';-x,y)}\ dudy \qquad \textbf{(D.51)}$$

where the change of variables $x = -u$ was made. From (D.2) , we have that $r(-\xi, \eta'; -x, y) = r(\xi, \eta'; x, y)$. Substituting this result on

(D.51)

$$U(-\xi',\eta') = -\frac{z}{j\lambda_0}\int\int_{\Pi} G(u,y)\frac{e^{jk_0 r(\xi,\eta';u,y)}}{r^2(\xi,\eta';u,y)}\ dudy \qquad \textbf{(D.52)}$$

Therefore, if $\Delta\xi = 0$, we can use the same function $\frac{e^{jk_0 r(\xi,\eta;u,y)}}{r^2(\xi,\eta';u,y)}$ to calculate the field at points (ξ,η') and $(-\xi,\eta')$.

- S2:

Assumption: $\Delta\eta = 0,\ \eta = \eta'$

The diffraction at $(\xi',-\eta)$ is given by, using (D.50),

$$U(\xi',-\eta) = -\frac{z}{j\lambda_0}\int\int_{\Pi} G(x,v)\frac{e^{jk_0 r(\xi',-\eta;x,-v)}}{r^2(\xi',-\eta;x,-v)}\ dxdv \qquad \textbf{(D.53)}$$

Where the change of variables $y = -v$ was made. From (D.2), we have that $r(\xi',-\eta;x,-y) = r(\xi,\eta;x,y)$. Therefore, substituting this result on (D.53)

$$U(\xi',-\eta) = -\frac{z}{j\lambda_0}\int\int_{\Pi} G(x,v)\frac{e^{jk_0 r(\xi',\eta;x,v)}}{r^2(\xi',\eta;x,v)}\ dxdv \qquad \textbf{(D.54)}$$

Therefore, if $\Delta\xi = 0$, we can use the same function $\frac{e^{jk_0 r(\xi',\eta;x,v)}}{r^2(\xi',\eta;x,v)}$ to calculate the field at points (ξ',η) and $(\xi',-\eta)$.

- S3:

Assumption: $\Delta\xi = 0,\ \xi = \xi'$ and $\Delta\eta = 0, \eta = \eta'$

The properties S1 and S2 hold if S3 is satisfied. Therefore, (D.54) and (D.56) already hold. If S1 followed by S2 are applied at (D.50), follows that

$$U(-\xi,-\eta) = \frac{z}{j\lambda_0}\int\int_{\Pi} G(u,b)\frac{e^{jk_0 r(\xi,\eta;u,v)}}{r^2(\xi,\eta;u,v)}\ dudv \qquad \textbf{(D.55)}$$

Next, we need to check the mirror symmetry with respect $\xi = \eta$ followed by the applications of S1 and S2 at it. Thus, the diffraction at (η, ξ) is given by, using (D.50),

$$U(\eta,\xi) = \frac{z}{j\lambda_0}\int\int_{\Pi} G(x,y)\frac{e^{jk_0 r(\eta,\xi;x,y)}}{r^2(\eta,\xi;x,y)}\ dxdy \qquad \textbf{(D.56)}$$

Now, we mirror the (x,y) plane with respect the $x = y$ axis. This can be done by the following change of coordinates $(x', y') = (y, x)$, whose Jacobian determinant can be calculated easily and is given by $|J(x', y')| = -1$. Furthermore, from (D.2) is easy to check that $r(\eta, \xi; y', x') = \mathrm{r}(\xi, \eta; x', y')$. Applying the coordinate change at (D.56)

$$U(\eta,\xi) = \frac{z}{j\lambda_0}\int\int_{\Pi} G(y',x')\frac{e^{jk_0 r(\xi,\eta;x',y')}}{r^2(\xi,\eta;x',y')}\ |J(x',y')|dx'dy'$$

$$= -\frac{z}{j\lambda_0}\int\int_{\Pi} G(y',x')\frac{e^{jk_0 r(\xi,\eta;x',y')}}{r^2(\xi,\eta;x',y')}\ dx'dy' \qquad \textbf{(D.57)}$$

If we apply S1 in (D.57), follows that

$$U(-\eta,\xi) = \frac{z}{j\lambda_0}\int\int_{\Pi} G(v',x')\frac{e^{jk_0 r(\xi,\eta;x',v')}}{r^2(\xi,\eta;x',v')}\ dx'dv' \qquad \textbf{(D.58)}$$

If we apply S2 in (D.57), follows that

$$U(\eta,-\xi) = \frac{z}{j\lambda_0}\int\int_{\Pi} G(y',u')\frac{e^{jk_0 r(\xi,\eta;u',y')}}{r^2(\xi,\eta;u',y')}\ du'dy' \qquad \textbf{(D.59)}$$

Finally, we apply both S1 and S2 in (D.57), obtaining

$$U(-\eta,-\xi) = -\frac{z}{j\lambda_0}\int\int_{\Pi} G(v',u')\frac{e^{jk_0 r(\xi,\eta;u',v')}}{r^2(\xi,\eta;u',v')}\ du'dv' \qquad \textbf{(D.60)}$$

Therefore, if $\Delta\xi = \Delta\eta = 0$, from equations (D.50), (D.52), (D.54), (D.55), (D.57), (D.58), (D.59) and (D.60) we can calculate $U(\sigma\xi', \sigma\eta'), U(\sigma\eta', -\sigma\xi'), U(\sigma\eta', \sigma\xi')\ and\ U(\sigma\eta', -\sigma\xi'),\ \sigma = \pm 1$ by using the same function $\frac{e^{jk_0 r(\xi',\eta';x,y)}}{r^2(\xi',\eta';x,y)}$ in the integrand.

b. Normal discrete RS Integration

The calculation will be performed for the sampled hologram assuming a constant field distribution across each pixel. Therefore, equation (D.3) can be approximated by

$$U_{kl} = \frac{z}{j\lambda_0} P^2 \sum_{m=1}^{M} \sum_{n=1}^{M} G_{mn} \frac{e^{jk_0 r(\xi_k, \eta_l; x_m, y_n)}}{r^2(\xi_k, \eta_l; x_m, y_n)} \quad \text{(D.61)}$$

where the CGH is stored in matrix G_{mn}, with MxM pixels with dimensions PxP each; the axes of the hologram plane are stored in vectors x_m and y_n; U_{kl} is the reconstructed field at point (ξ_k, η_l). The operating wavelength is λ_0 and the distance between the planes is z. To perform (D.61) in MATLAB, for each U_{kl}, it is faster to first create the following auxiliary matrix

$$A_{mn}(\xi_k, \eta_l) \equiv \frac{e^{jk_0 r(\xi_k, \eta_l; x_m, y_n)}}{r^2(\xi_k, \eta_l; x_m, y_n)} \quad \text{(D.62)}$$

and then perform the element-wise multiplication with G_{mn} and sum all elements of the resulting matrix. Therefore, if the reconstruction window has $N \times N$ points, the matrix A_{mn} must be created N^2 times.

c. RS discrete Integration: S1 symmetry

Applying

(D.51) at (D.61), follows

$$U_{N-k,l} = -\frac{z}{j\lambda_0} P^2 \sum_{m=1}^{M} \sum_{n=1}^{M} G_{M-m,n} A_{mn}(\xi_k, \eta_l) \quad \text{(D.63)}$$

Note that, from (D.61) and (D.63), $U_{k,l}$ $U_{N-k,l}$ are calculated from the same matrix $A_{mn}(\xi_k, \eta_l)$. Furthermore, the summations in (D.63) are performed by element-wise multiplication between $G_{M-m,n}$, that is the G_{mn} flipped up-down, and A_{mn} and summing the resulting matrix. Therefore, if the reconstruction window has $N \times N$ points, the matrix A_{mn} must be created $\frac{N^2}{2}$ times.

d. RS discrete Integration: S2 symmetry

Applying (D.52) at (D.61), follows

$$U_{k,N-l} = -\frac{z}{j\lambda_0} P^2 \sum_{m=1}^{M} \sum_{n=1}^{M} G_{m,M-n} A_{mn}(\xi_k, \eta_l) \quad \text{(D.64)}$$

Note that, from (D.64) and (D.63), $U_{k,l}$ $U_{k,N-l}$ are calculated from the same matrix $A_{mn}(\xi_k, \eta_l)$. Furthermore, the summations in (D.63) are performed by element-wise multiplication between $G_{m,M-n}$, that is the G_{mn} flipped left-right, and A_{mn} and summing the resulting matrix. Therefore, if the reconstruction window has $N \times N$ points, the matrix A_{mn} must be created $\frac{N^2}{2}$ times.

e. RS discrete Integration: S2 symmetry

Applying equations (D.50), (D.52), (D.54), (D.55), (D.57), (D.58), (D.59) and (D.60) at (D.61), follows, respectively that

$$U_{N-k,l} = -\frac{z}{j\lambda_0} P^2 \sum_{m=1}^{M} \sum_{n=1}^{M} G_{M-m,n} A_{mn}(\xi_k, \eta_l) \quad \text{(D.65)}$$

$$U_{k,N-l} = -\frac{z}{j\lambda_0} P^2 \sum_{m=1}^{M} \sum_{n=1}^{M} G_{m,M-n} A_{mn}(\xi_k, \eta_l) \quad \text{(D.66)}$$

$$U_{N-k,N-l} = \frac{z}{j\lambda_0} P^2 \sum_{m=1}^{M} \sum_{n=1}^{M} G_{M-m,M-n} A_{mn}(\xi_k, \eta_l) \quad \text{(D.67)}$$

$$U_{lk} = -\frac{z}{j\lambda_0} P^2 \sum_{m=1}^{M} \sum_{n=1}^{M} G_{nm} A_{mn}(\xi_k, \eta_l) \quad \text{(D.68)}$$

$$U_{l-N,k} = \frac{z}{j\lambda_0} P^2 \sum_{m=1}^{M} \sum_{n=1}^{M} G_{n-M,m} A_{mn}(\xi_k, \eta_l) \quad \text{(D.69)}$$

$$U_{l-N,k} = \frac{z}{j\lambda_0} P^2 \sum_{m=1}^{M} \sum_{n=1}^{M} G_{n,m-N} A_{mn}(\xi_k, \eta_l) \quad \text{(D.70)}$$

$$U_{l-N,k-N} = -\frac{z}{j\lambda_0} P^2 \sum_{m=1}^{M} \sum_{n=1}^{M} G_{n-N,m-N} A_{mn}(\xi_k, \eta_l) \quad \text{(D.71)}$$

In equations (D.61) and (D.65)-(D.71) the same matrix $A_{mn}(\xi_k, \eta_l)$ is used to obtain the diffraction at different points. These operations can also be performed by element-wise

multiplication of $A_{mn}(\xi_k, \eta_l)$ with a transformed G_{mn} followed by the summation of all elements of the resulting matrix. In (D.65) the multiplication is with $G_{M-m,n}$, that is the flipped left-right G_{mn}; in (D.66) the multiplication is with $G_{m,n-M}$, that is the flipped up-down G_{mn}; in (D.67) the multiplication is with $G_{M-m,n-N}$, that is the flipped left-right and up-down G_{mn}; in (D.68) the multiplication is with $G_{n,m}$, that is the transpose of G_{mn}; finally, in (D.69), (D.70),(D.71) the multiplication is with, respectively, the flipped left-right, flipped up-down and flipped left-right and up-down of the transpose of G_{mn}. These matrix transformations are fast and therefore do not appreciably increases the computational time.

f. Performance comparison

In this section the implementations of the normal, S1 and S3 RS integrations will be compared. Note that S1 and S2 should have the same performance under the same circumstances so only the former will be shown. The hologram that will be reconstructed is an 8 phase levels CGH with wide angle correction obtained by the IFTA with $c = 100$ as described at section D.4. The RS integration is performed assuming an operating wavelength of 532 nm, hologram pixel size of 190 nm and the reconstruction plane at the Fraunhofer region. The hologram has dimensions of 1024×1024 pixels and the observation window have the same size of the reconstruction region and is centered at the reconstruction plane, that is, $\Delta\xi = \Delta\eta = 0$. Table D-2 lists the time taken to perform the RS integration under the normal, S1 and S3 methods for observation windows with different resolutions. Note that, for all methods, the computational time is proportional to N^2. Furthermore, S1 is twice as faster as the Normal method and the S3 is almost five times faster than the normal method. Although not included, the performance of the S2 method should be the same of S1 because both algorithms use the same operations. Therefore, depending on the location of projected image, it is better to place the observation window in such a way that preferable S3 or S1 or S2 is used to reduce the computational time.

Table D-2 – Time expended to complete the RS integration of a hologram with 1024×1024 pixels in an observation window with different resolutions. Three methods were used: Normal, S1 and S3.

Observation window resolution (N×N)	Computational time (s)		
	Normal	S1	S3
16×16	24	13.18	5.1
32×32	99.51	51.5	18.9
64×64	382	203.5	73.3
128×128	1548.8	819.2	315
256×256	6092	3200	1185

D.4.2. Power calculation of the diffracted fields

The scalar approximation will be used to calculate the transmitted power over certain region S in the reconstruction plane. In this case, the Poynting vector, $\vec{P}$, at a point (ξ, η) in the reconstruction plane, will be radial and given by [3]

$$\vec{P}(\xi,\eta) = \frac{|U(\xi,\eta)|^2}{2}\hat{s} \qquad \text{(D.72)}$$

Where $U(\xi,\eta)$ is the reconstructed field distribution and $\hat{s}$ is the direction versor, $|\hat{s}| = 1$ of the Poynting vector. It is assumed that the reconstructed field is in the far-field zone and the holograms size is much smaller than the reconstruction plane. Therefore, the Poynting vector is assumed to be radial with center on the middle of the hologram, see Figure D.8.

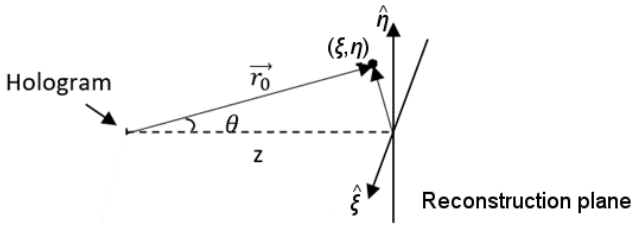


Figure D.8 – Representation of the reconstruction plane.

Follows that,

$$\hat{s} = \frac{\vec{r_0}}{|\vec{r_0}|} \qquad \text{(D.73)}$$

where $\vec{r_0} \equiv \xi\hat{\xi} + \eta\hat{\eta} + z\hat{z}$ is a radial vector that links the center of the hologram to a point (ξ,η) in the reconstruction plane. Therefore, the power flux crossing a region S inside the reconstruction plane is given by

$$\mathcal{E} = \int\int_S \vec{dS} \cdot \vec{P}(\xi,\eta) = \frac{1}{2}\int\int_S dS|U(\xi,\eta)|^2 \hat{s}\cdot\hat{z} \qquad \text{(D.74)}$$

Substituting (D.73) and (D.2) in (D.74), follows that

$$\mathcal{E} = \frac{1}{2}\int\int_S dS|U(\xi,\eta)|^2 \frac{z}{r(\xi,\eta;0,0)} \qquad \text{(D.75)}$$

The field distribution $U(\xi,\eta)$ is calculated numerically by the RS integration and stored in a $n_x \times n_y$ matrix U_{ij}. Thus, equation (D.75) is approximated by the following sum

$$\mathcal{E} = \frac{1}{2} z d\xi d\eta \sum_{i=1}^{n_x} \sum_{j=1}^{n_y} \left|U_{ij}\right|^2 \frac{1}{r\left(\xi_i, \eta_j; 0,0\right)} \tag{D.76}$$

where $d\xi$ and $d\eta$ are the sampling of the field on the reconstruction plane along the ξ and η directions.

Appendix E – Rights & Permissions

E.1. Holographic metasurfaces

Section III.1 of chapter III the main text was adapted with permissions from

Augusto Martins, et.al , "Highly efficient holograms based on c-Si metasurfaces in the visible range," Opt. Express 26, 9573-9583 (2018) [77] © The Optical Society

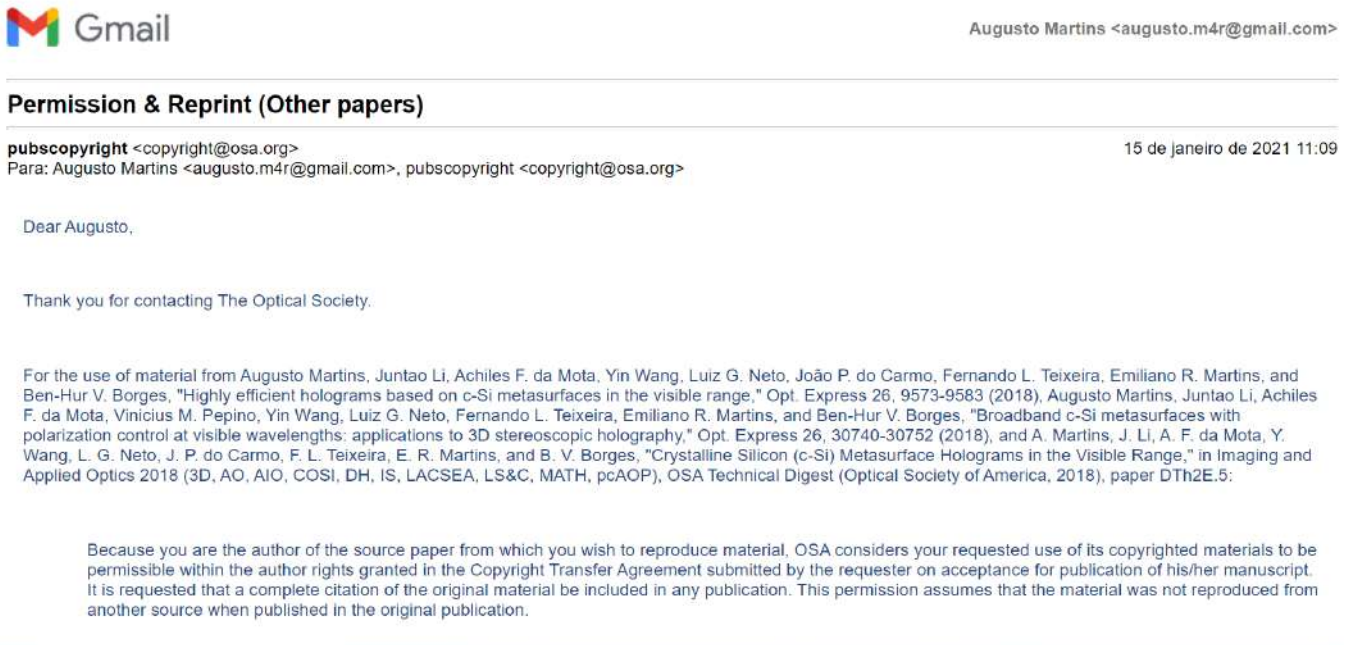
and

A. Martins, et. al, "Crystalline Silicon (c-Si) Metasurface Holograms in the Visible Range," in Imaging and Applied Optics 2018 (3D, AO, AIO, COSI, DH, IS, LACSEA, LS&C, MATH, pcAOP), OSA Technical Digest (Optical Society of America, 2018), paper DTh2E.5: [89] © The Optical Society.

Section III.2 of chapter III of the main text was adapted with permissions from

Augusto Martins, et. al, "Broadband c-Si metasurfaces with polarization control at visible wavelengths: applications to 3D stereoscopic holography," Opt. Express 26, 30740-30752 (2018) [93] © The Optical Society.

A copy of the permission is shown below.



Gmail Augusto Martins <augusto.m4r@gmail.com>

Permission & Reprint (Other papers)

pubscopyright <copyright@osa.org> 15 de janeiro de 2021 11:09
Para: Augusto Martins <augusto.m4r@gmail.com>, pubscopyright <copyright@osa.org>

Dear Augusto,

Thank you for contacting The Optical Society.

For the use of material from Augusto Martins, Juntao Li, Achiles F. da Mota, Yin Wang, Luiz G. Neto, João P. do Carmo, Fernando L. Teixeira, Emiliano R. Martins, and Ben-Hur V. Borges, "Highly efficient holograms based on c-Si metasurfaces in the visible range," Opt. Express 26, 9573-9583 (2018), Augusto Martins, Juntao Li, Achiles F. da Mota, Vinicius M. Pepino, Yin Wang, Luiz G. Neto, Fernando L. Teixeira, Emiliano R. Martins, and Ben-Hur V. Borges, "Broadband c-Si metasurfaces with polarization control at visible wavelengths: applications to 3D stereoscopic holography," Opt. Express 26, 30740-30752 (2018), and A. Martins, J. Li, A. F. da Mota, Y. Wang, L. G. Neto, J. P. do Carmo, F. L. Teixeira, E. R. Martins, and B. V. Borges, "Crystalline Silicon (c-Si) Metasurface Holograms in the Visible Range," in Imaging and Applied Optics 2018 (3D, AO, AIO, COSI, DH, IS, LACSEA, LS&C, MATH, pcAOP), OSA Technical Digest (Optical Society of America, 2018), paper DTh2E.5:

Because you are the author of the source paper from which you wish to reproduce material, OSA considers your requested use of its copyrighted materials to be permissible within the author rights granted in the Copyright Transfer Agreement submitted by the requester on acceptance for publication of his/her manuscript. It is requested that a complete citation of the original material be included in any publication. This permission assumes that the material was not reproduced from another source when published in the original publication.

E.2. On metalenses with arbitrarily wide field of view

Chapter IV of the main text was adapted with permission from

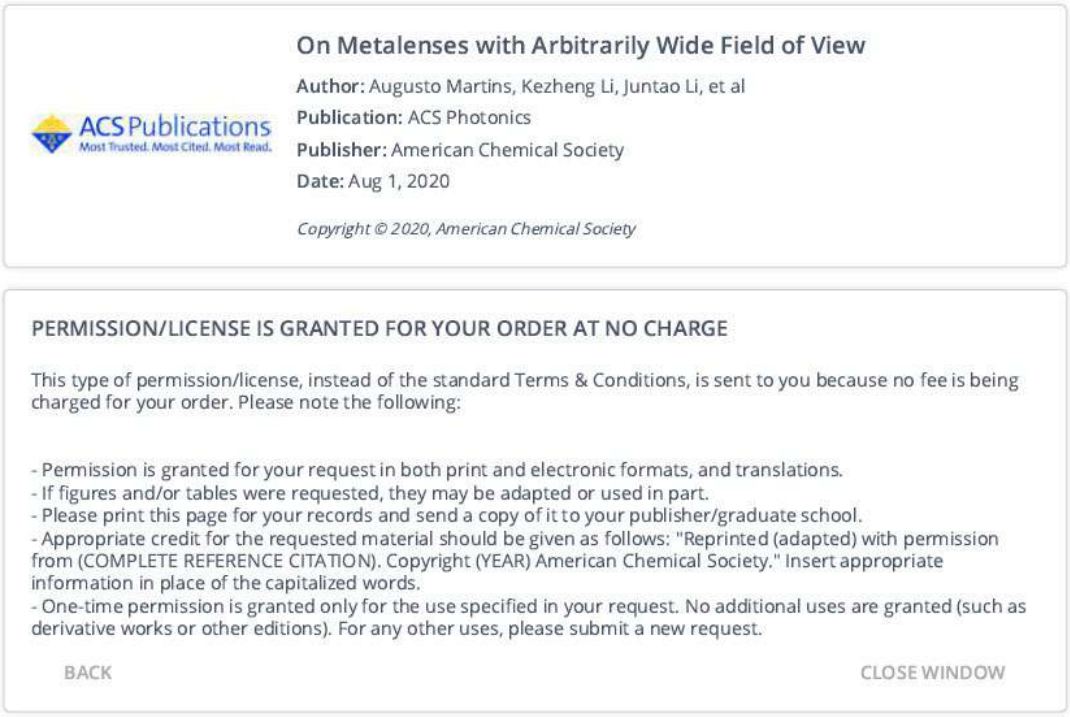
Martins, Augusto, et al. "On metalenses with arbitrarily wide field of view." ACS Photonics 7.8 (2020): 2073-2079. Copyright (2021) American Chemical Society.

and

Haowen Liang, et al."High performance metalenses: numerical aperture, aberrations, chromaticity, and trade-offs," Optica 6, 1461-1470 (2019) [114] © The Optical Society.

A copy of the permissions are shown below.

- **ACS:**



On Metalenses with Arbitrarily Wide Field of View

Author: Augusto Martins, Kezheng Li, Juntao Li, et al
Publication: ACS Photonics
Publisher: American Chemical Society
Date: Aug 1, 2020

Copyright © 2020, American Chemical Society

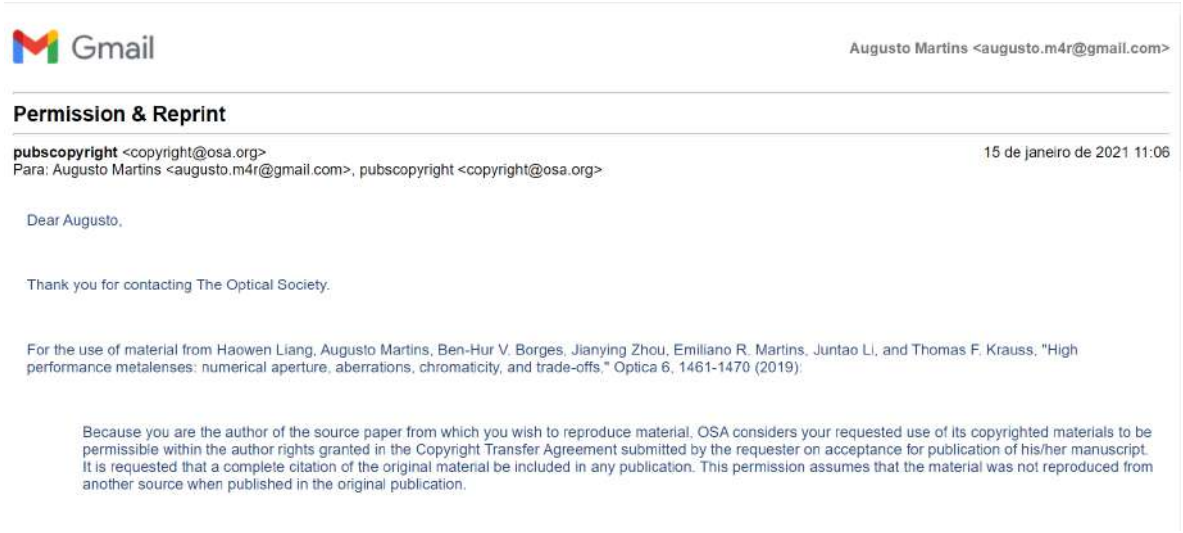
PERMISSION/LICENSE IS GRANTED FOR YOUR ORDER AT NO CHARGE

This type of permission/license, instead of the standard Terms & Conditions, is sent to you because no fee is being charged for your order. Please note the following:

- Permission is granted for your request in both print and electronic formats, and translations.
- If figures and/or tables were requested, they may be adapted or used in part.
- Please print this page for your records and send a copy of it to your publisher/graduate school.
- Appropriate credit for the requested material should be given as follows: "Reprinted (adapted) with permission from (COMPLETE REFERENCE CITATION). Copyright (YEAR) American Chemical Society." Insert appropriate information in place of the capitalized words.
- One-time permission is granted only for the use specified in your request. No additional uses are granted (such as derivative works or other editions). For any other uses, please submit a new request.

BACK CLOSE WINDOW

- **OSA**



Gmail Augusto Martins <augusto.m4r@gmail.com>

Permission & Reprint

pubscopyright <copyright@osa.org> 15 de janeiro de 2021 11:06
Para: Augusto Martins <augusto.m4r@gmail.com>, pubscopyright <copyright@osa.org>

Dear Augusto,

Thank you for contacting The Optical Society.

For the use of material from Haowen Liang, Augusto Martins, Ben-Hur V. Borges, Jianying Zhou, Emiliano R. Martins, Juntao Li, and Thomas F. Krauss, "High performance metalenses: numerical aperture, aberrations, chromaticity, and trade-offs," Optica 6, 1461-1470 (2019):

Because you are the author of the source paper from which you wish to reproduce material, OSA considers your requested use of its copyrighted materials to be permissible within the author rights granted in the Copyright Transfer Agreement submitted by the requester on acceptance for publication of his/her manuscript. It is requested that a complete citation of the original material be included in any publication. This permission assumes that the material was not reproduced from another source when published in the original publication.

E.3. Photon Management in Tandem Si/Perovskite Solar Cells

Chapter V of the main text was adapted with permission from *Martins, Augusto, et al. "Photonic intermediate structures for Perovskite/c-Silicon four terminal tandem solar cells." IEEE Journal of Photovoltaics 7.5 (2017): 1190-1196.*.

A copy of the permission is shown below.

IEEE
Requesting permission to reuse content from an IEEE publication

Photonic Intermediate Structures for Perovskite/c-Silicon Four Terminal Tandem Solar Cells

Author: Augusto Martins
Publication: IEEE Journal of Photovoltaics
Publisher: IEEE
Date: Sept. 2017

Copyright © 2017, IEEE

Thesis / Dissertation Reuse

The IEEE does not require individuals working on a thesis to obtain a formal reuse license, however, you may print out this statement to be used as a permission grant:

Requirements to be followed when using any portion (e.g., figure, graph, table, or textual material) of an IEEE copyrighted paper in a thesis:

1) In the case of textual material (e.g., using short quotes or referring to the work within these papers) users must give full credit to the original source (author, paper, publication) followed by the IEEE copyright line © 2011 IEEE.
2) In the case of illustrations or tabular material, we require that the copyright line © [Year of original publication] IEEE appear prominently with each reprinted figure and/or table.
3) If a substantial portion of the original paper is to be used, and if you are not the senior author, also obtain the senior author's approval.

Requirements to be followed when using an entire IEEE copyrighted paper in a thesis:

1) The following IEEE copyright/ credit notice should be placed prominently in the references: © [year of original publication] IEEE. Reprinted, with permission, from [author names, paper title, IEEE publication title, and month/year of publication]
2) Only the accepted version of an IEEE copyrighted paper can be used when posting the paper or your thesis on-line.
3) In placing the thesis on the author's university website, please display the following message in a prominent place on the website: In reference to IEEE copyrighted material which is used with permission in this thesis, the IEEE does not endorse any of [university/educational entity's name goes here]'s products or services. Internal or personal use of this material is permitted. If interested in reprinting/republishing IEEE copyrighted material for advertising or promotional purposes or for creating new collective works for resale or redistribution, please go to http://www.ieee.org/publications_standards/publications/rights/rights_link.html to learn how to obtain a License from RightsLink.

If applicable, University Microfilms and/or ProQuest Library, or the Archives of Canada may supply single copies of the dissertation.

BACK CLOSE WINDOW

Bibliography

[1] M. Born, E. Wolf, and A. B. Bhatia, *Principles of optics : electromagnetic theory of propagation, interference and diffraction of light*, 7th (expanded) ed. Cambridge: Cambridge University Press, 2002.

[2] L. Novotny and B. Hecht, *Principles of nano-optics*. Cambridge: Cambridge University Press, 2006.

[3] J. W. Goodman, *Introduction to Fourier Optics* vol. 2: Roberts & Company Publishers, 2005.

[4] G. Keiser, "Optical fiber communications," *Wiley encyclopedia of telecommunications,* 2003.

[5] R. Loudon, *The quantum theory of light*, 3rd ed. Oxford: Clarendon Press ;, 2000.

[6] B. E. Saleh, M. C. Teich, and B. E. Saleh, *Fundamentals of photonics* vol. 22: Wiley New York, 1991.

[7] C. Kittel and P. McEuen, *Introduction to solid state physics*, 8th ed. New York ; Chichester: Wiley, 2005.

[8] J. D. Jackson, *Classical Electrodynamics*, 1999.

[9] J. D. Joannopoulos, *Photonic crystals : molding the flow of light*, 2nd ed. Princeton: Princeton University Press, 2008.

[10] S. A. Maier, *Plasmonics : fundamentals and applications*. Bath ; New York: Springer, 2006.

[11] C. G. Biris and N. C. Panoiu, "Excitation of dark plasmonic cavity modes via nonlinearly induced dipoles: applications to near-infrared plasmonic sensing," *Nanotechnology,* vol. 22, p. 235502, 2011.

[12] C. Sönnichsen and A. P. Alivisatos, "Gold Nanorods as Novel Nonbleaching Plasmon-Based Orientation Sensors for Polarized Single-Particle Microscopy," *Nano Letters,* vol. 5, pp. 301-304, 2005/02/01 2005.

[13] J. N. Anker, W. P. Hall, O. Lyandres, N. C. Shah, J. Zhao, and R. P. Van Duyne, "Biosensing with plasmonic nanosensors," *Nature Materials,* vol. 7, p. 442, 2008.

[14] K. A. Willets and R. P. Van Duyne, "Localized surface plasmon resonance spectroscopy and sensing," *Annu. Rev. Phys. Chem.,* vol. 58, pp. 267-297, 2007.

[15] P. B. Johnson and R.-W. Christy, "Optical constants of the noble metals," *Physical Review B,* vol. 6, p. 4370, 1972.

[16] J. Prikulis, F. Svedberg, M. Käll, J. Enger, K. Ramser, M. Goksör, and D. Hanstorp, "Optical spectroscopy of single trapped metal nanoparticles in solution," *Nano Letters,* vol. 4, pp. 115-118, 2004.

[17] M. W. Knight, L. Liu, Y. Wang, L. Brown, S. Mukherjee, N. S. King, H. O. Everitt, P. Nordlander, and a. N. J. Halas, "Aluminum Plasmonic Nanoantennas," *Nano Letters,* vol. 12, pp. 6000–6004, October 2012.

[18] J. Ye and P. Van Dorpe, "Plasmonic behaviors of gold dimers perturbed by a single nanoparticle in the gap," *Nanoscale,* vol. 4, pp. 7205-7211, 2012.

[19] V. Giannini, A. I. Fernández-Domínguez, S. C. Heck, and S. A. Maier, "Plasmonic Nanoantennas: Fundamentals and Their Use in Controlling the Radiative Properties of Nanoemitters," *Chem. Re.,* vol. 111, pp. 3888-3912, March 2011.

[20] H. Fischer and O. J. F. Martin, "Engineering the optical response of plasmonic nanoantennas," *Optics Express,* vol. 16, pp. 9144-9154, 2008/06/09 2008.

[21] G. W. Bryant, J. G. d. Abajo, and J. AIzpurua, "Mapping the Plasmon Resonances of Metallic Nanoantennas," *Nano Letters,* vol. 8, pp. 631-636, January 2008.

[22] J. B. Jackson and N. J. Halas, "Surface-enhanced Raman scattering on tunable plasmonic nanoparticle substrates," *Proceedings of the National Academy of Sciences,* vol. 101, pp. 17930-17935, 2004.

[23] D. Wang, W. Zhu, M. D. Best, J. P. Camden, and K. B. Crozier, "Directional Raman scattering from single molecules in the feed gaps of optical antennas," *Nano Lett,* vol. 13, pp. 2194-8, May 8 2013.
[24] E. Shamonina and L. Solymar, "Metamaterials: How the subject started," *Metamaterials,* vol. 1, pp. 12-18, 2007.
[25] N. Engheta and R. W. Ziolkowski, *Metamaterials: physics and engineering explorations*: John Wiley & Sons, 2006.
[26] V. G. Veselago, "The Electrodynamics of Substances with Simultaneously Negative Values of ε and μ," *Sov. Phys. Usp.,* vol. 10, p. 509, January-February 1968.
[27] J. J. Yang, M. Huang, and J. Sun, "Double negative metamaterial sensor based on microring resonator," *IEEE Sensors Journal,* vol. 11, pp. 2254-2259, 2011.
[28] R. W. Ziolkowski and A. D. Kipple, "Application of double negative materials to increase the power radiated by electrically small antennas," *IEEE Transactions on Antennas and Propagation,* vol. 51, pp. 2626-2640, 2003.
[29] J. B. Pendry, "Negative Refraction Makes a Perfect Lens," *Physical Review Letters,* vol. 85, pp. 3966-3969, Outubro 2000.
[30] U. Leonhardt, "Optical Conformal Mapping," *Science,* vol. 312, pp. 1777-1780, June 2006.
[31] J. B. Pendry, D. Schurig, and D. R. Smith, "Controlling Electromagnetic Fields," *Science,* vol. 312, pp. 1780-1782, June 2006.
[32] T. Ergin, N. Stenger, P. Brenner, J. B. Pendry, and M. Wegener, "Three-dimensional invisibility cloak at optical wavelengths," *Science,* vol. 328, pp. 337-339, 2010.
[33] E. Cubukcu, K. Aydin, E. Ozbay, S. Foteinopoulou, and C. M. Soukoulis, "Negative refraction by photonic crystals," *Nature,* vol. 423, pp. 604-605, 2003.
[34] C. L. Holloway, E. F. Kuester, J. A. Gordon, J. O'Hara, J. Booth, and D. R. Smith, "An Overview of the Theory and Applications of Metasurfaces: The Two-Dimensional Equivalents of Metamaterials," *IEEE Antennas and Propagation Magazine,* vol. 54, pp. 10–35, April 2012.
[35] P. Genevet and F. Capasso, "Holographic Optical metasurfaces: a review of current progress," *Reports on progress in Physics,* vol. 78, p. 024401, January 2015.
[36] P. Genevet, F. Capasso, F. Aieta, M. Khorasaninejad, and R. Devlin, "Recent advances in planar optics: from plasmonic to dielectric metasurfaces," *Optica,* vol. 4, pp. 139-152, 2017.
[37] N. Yu and F. Capasso, "Flat optics with designer metasurfaces," *Nature Materials,* vol. 13, pp. 139-150, January 2014.
[38] Y. Zhao, X.-X. Liu, and A. Alù, "Recent advances on optical metasurfaces," *Journal of Optics,* vol. 16, p. 123001, 2014.
[39] W. T. Chen, A. Y. Zhu, and F. Capasso, "Flat optics with dispersion-engineered metasurfaces," *Nature Reviews Materials,* vol. 5, pp. 604-620, 2020.
[40] F. Aieta, P. Genevet, N. Yu, M. A. Kats, Z. Gaburro, and F. Capasso, "Out-of-Plane Reflection and Refraction of Light by Anisotropic Optical Antenna Metasurfaces with Phase Discontinuities," *Nano Letters,* vol. 12, pp. 1702-1706, February 2012.
[41] F. Aieta, P. Genevet, M. A. Kats, N. Yu , R. Blanchard, Z. Gaburro, and F. Capasso, "Aberration-Free Ultrathin Flat Lenses and Axicons at Telecom Wavelengths Based on Plasmonic Metasurfaces," *Nano Letters,* vol. 12, pp. 4932-4936, August 2012.
[42] A. V. Kildishev, A. Boltasseva, and V. M. Shalaev, "Planar Photonics with Metasurfaces," *Science,* vol. 339, p. 1232009, March 2013.
[43] X. Ni, A. V. Kildishev, and V. M. Shalaev, "Metasurface holograms for visible light," *Nature Communications,* vol. 4, p. 2807, 2013.
[44] G. Zheng, H. Mühlenbernd, M. Kenney, G. Li, T. Zentgraf, and S. Zhang, "Metasurface holograms reaching 80% efficiency," *Nature Nanotechnology,* vol. 10, pp. 308-312, 2015.
[45] W. Wan, J. Gao, and X. Yang, "Full-Color Plasmonic Metasurface Holograms," *ACS Nano,* vol. 10, pp. 10671-10680, Dec 27 2016.

[46] Y.-W. Huang, W. T. Chen, W.-Y. Tsai, P. C. Wu, C.-M. Wang, G. Sun, and D. P. Tsai, "Aluminum plasmonic multicolor meta-hologram," *Nano Letters,* vol. 15, pp. 3122-3127, 2015.

[47] A. B. Evlyukhin, C. Reinhardt, and B. N. Chichkov, "Multipole light scattering by nonspherical nanoparticles in the discrete dipole approximation," *Physical Review B,* vol. 84, p. 235429, 2011.

[48] I. Staude, A. E. Miroshnichenko, M. Decker, N. T. Fofang, S. Liu, E. Gonzales, J. Dominguez, T. S. Luk, D. N. Neshev, I. Brener, and Y. Kivshar, "Tailoring Directional Scattering through Magnetic and Electric Resonances in Subwavelength Silicon Nanodisks," *ACS Nano,* vol. 7, pp. 7824–7832, 2013/09/24 2013.

[49] M. Decker, I. Staude, M. Falkner, J. Dominguez, D. N. Neshev, I. Brener, T. Pertsch, and Y. S. Kivshar, "High-efficiency dielectric Huygens' surfaces," *Advanced Optical Materials,* vol. 3, pp. 813-820, 2015.

[50] W. Zhao, H. Jiang, B. Liu, J. Song, Y. Jiang, C. Tang, and J. Li, "Dielectric Huygens' Metasurface for High-Efficiency Hologram Operating in Transmission Mode," *Scientific Reports,* vol. 6, p. 30613, 2016.

[51] K. E. Chong, L. Wang, I. Staude, A. R. James, J. Dominguez, S. Liu, G. S. Subramania, M. Decker, D. N. Neshev, and I. Brener, "Efficient polarization-insensitive complex wavefront control using Huygens' metasurfaces based on dielectric resonant meta-atoms," *ACS Photonics,* vol. 3, pp. 514-519, 2016.

[52] A. Arbabi, Y. Horie, M. Bagheri, and A. Faraon, "Dielectric metasurfaces for complete control of phase and polarization with subwavelength spatial resolution and high transmission," *Nature Nanotechnology,* vol. 10, pp. 937-943, 2015.

[53] Z. Zhou, J. Li, R. Su, B. Yao, H. Fang, K. Li, L. Zhou, J. Liu, D. Stellinga, C. P. Reardon, T. F. Krauss, and X. Wang, "Efficient Silicon Metasurfaces for Visible Light," *ACS Photonics,* vol. 4, pp. 544–551, 2017.

[54] A. Zhan, S. Colburn, R. Trivedi, T. K. Fryett, C. M. Dodson, and A. Majumdar, "Low-contrast dielectric metasurface optics," *ACS Photonics,* vol. 3, pp. 209–214, 2016.

[55] G. Yoon, D. Lee, K. T. Nam, and J. Rho, "Pragmatic Metasurface Hologram at Visible Wavelength: The Balance between Diffraction Efficiency and Fabrication Compatibility," *ACS Photonics,* 2017/12/20 2017.

[56] A. Arbabi, R. M. Briggs, Y. Horie, M. Bagheri, and A. Faraon, "Efficient dielectric metasurface collimating lenses for mid-infrared quantum cascade lasers," *Optics Express,* vol. 23, pp. 33310-33317, 2015.

[57] M. Khorasaninejad, F. Aieta, P. Kanhaiya, M. A. Kats, P. Genevet, D. Rousso, and F. Capasso, "Achromatic metasurface lens at telecommunication wavelengths," *Nano Letters,* vol. 15, pp. 5358-5362, 2015.

[58] 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, pp. 1190-1194, 2016.

[59] K. Huang, Z. Dong, S. Mei, L. Zhang, Y. Liu, H. Liu, H. Zhu, J. Teng, B. Luk'yanchuk, and J. K. Yang, "Silicon multi-meta-holograms for the broadband visible light," *Laser & Photonics Reviews,* vol. 10, pp. 500-509, 2016.

[60] Q.-T. Li, F. Dong, B. Wang, F. Gan, J. Chen, Z. Song, L. Xu, W. Chu, Y.-F. Xiao, and Q. Gong, "Polarization-independent and high-efficiency dielectric metasurfaces for visible light," *Optics Express,* vol. 24, pp. 16309-16319, 2016.

[61] B. Wang, F. Dong, Q.-T. Li, D. Yang, C. Sun, J. Chen, Z. Song, L. Xu, W. Chu, and Y.-F. Xiao, "Visible-frequency dielectric metasurfaces for multiwavelength achromatic and highly dispersive holograms," *Nano Letters,* vol. 16, pp. 5235-5240, 2016.

[62] 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 of the United States of America,* vol. 113, pp. 10473-10478, September 20, 2016 2016.

[63] L. Wang, S. Kruk, H. Tang, T. Li, I. Kravchenko, D. N. Neshev, and Y. S. Kivshar, "Grayscale transparent metasurface holograms," *Optica,* vol. 3, pp. 1504–1505, 2016/12/20 2016.
[64] S. Kruk, B. Hopkins, I. I. Kravchenko, A. Miroshnichenko, D. N. Neshev, and Y. S. Kivshar, "Invited Article: Broadband highly efficient dielectric metadevices for polarization control," *APL Photonics,* vol. 1, p. 030801, 2016.
[65] M. I. Shalaev, J. Sun, A. Tsukernik, A. Pandey, K. Nikolskiy, and N. M. Litchinitser, "High-efficiency all-dielectric metasurfaces for ultracompact beam manipulation in transmission mode," *Nano Letters,* vol. 15, pp. 6261–6266, 2015.
[66] K. E. Chong, I. Staude, A. James, J. Dominguez, S. Liu, S. Campione, G. S. Subramania, T. S. Luk, M. Decker, and D. N. Neshev, "Polarization-independent silicon metadevices for efficient optical wavefront control," *Nano Letters,* vol. 15, pp. 5369-5374, 2015.
[67] P. Lalanne, S. Astilean, P. Chavel, E. Cambril, and H. Launois, "Blazed binary subwavelength gratings with efficiencies larger than those of conventional échelette gratings," *Optics Letters,* vol. 23, pp. 1081–1083, 1998/07/15 1998.
[68] J. B. Mueller, N. A. Rubin, R. C. Devlin, B. Groever, and F. Capasso, "Metasurface Polarization Optics: Independent Phase Control of Arbitrary Orthogonal States of Polarization," *Physical Review Letters,* vol. 118, p. 113901, 2017.
[69] Y. F. Yu, A. Y. Zhu, R. Paniagua-Domínguez, Y. H. Fu, B. Luk'yanchuk, and A. I. Kuznetsov, "High-transmission dielectric metasurface with 2π phase control at visible wavelengths," *Laser & Photonics Reviews,* vol. 9, pp. 412–418, 2015.
[70] D. Gabor, "A new microscopic principle," *Nature,* vol. 161, pp. 777-778, 1948.
[71] E. N. Leith and J. Upatnieks, "Wavefront reconstruction and communication theory," *J. Opt. Soc. Am.,* vol. 52, pp. 1123-1130, 1962.
[72] P. M. Hirch, J. A. Jordan, and L. B. Lesem, "Method of making an object-dependent diffuser," 3619022, 1971.
[73] R. W. Gerchberg and W. O. Saxton, "A practical algorithm for the determination of phase from image and diffraction plane pictures," *Optik,* vol. 35, pp. 237-246, 1972.
[74] F. Wirowski and O. Bryngdahl, "Iterative Fourier-Transform algorithm applied to computer holography," *Journal of Optical Society of America A,* vol. 5, pp. 1058-1066, 1988.
[75] L. G. Neto, "Optical real-time holograms using liquid crystal television and computer interative design," Doctoral Doctoral Thesis, Faculté des Sciences et Génie, Université Laval, Canada, 1995.
[76] C. F. Bohren, *Absorption and scattering of light by small particles*, 2nd ed. Weinheim Chichester: Wiley-VCH ; John Wiley [distributor], 2010.
[77] A. Martins, J. Li, A. F. da Mota, Y. Wang, L. G. Neto, J. P. do Carmo, F. L. Teixeira, E. R. Martins, and B. V. Borges, "Highly efficient holograms based on c-Si metasurfaces in the visible range," *Opt Express,* vol. 26, pp. 9573-9583, Apr 16 2018.
[78] S. Pancharatnam, "Generalized theory of interference, and its applications.," *Proc. Indian Acad. Sci. Sect. A* vol. 44, pp. 247-262, 1956.
[79] M. V. Berry, "Quantal phase factors accompanying adiabatic changes.," *Proc. R. Soc. London A* vol. 392, pp. 45-57, 1984.
[80] N. Yu, P. Genevet, M. A. Kats, F. Aieta, and J.-P. Tetienne, "Light Propagation with Phase Discontinuities: Generalized Laws of Reflection and Refraction," *Science,* vol. 21, pp. 333-337, September 2011.
[81] Q. Wang, X. Zhang, Y. Xu, J. Gu, Y. Li, Z. Tian, R. Singh, S. Zhang, J. Han, and W. Zhang, "Broadband metasurface holograms: toward complete phase and amplitude engineering," *Scientific Reports,* vol. 6, 2016.
[82] M. Kerker, D.-S. Wang, and C. Giles, "Electromagnetic scattering by magnetic spheres," *JOSA,* vol. 73, pp. 765-767, 1983.
[83] S. Kruk and Y. Kivshar, "Functional Meta-Optics and Nanophotonics Governed by Mie Resonances," *ACS Photonics,* vol. 4, pp. 2638-2649, 2017/11/15 2017.

[84] M. C. Y. Huang, Y. Zhou, and C. J. Chang-Hasnain, "A surface-emitting laser incorporating a high-index-contrast subwavelength grating," *Nature Photonics,* vol. 1, p. 119, 2007.
[85] J. Wang, I. Glesk, and L. R. Chen, "Subwavelength grating filtering devices," *Optics Express,* vol. 22, pp. 15335-15345, 2014/06/30 2014.
[86] D. Fattal, J. Li, Z. Peng, M. Fiorentino, and R. G. Beausoleil, "Flat dielectric grating reflectors with focusing abilities," *Nat Photon,* vol. 4, pp. 466-470, 2010.
[87] X. Chen, L. Huang, H. Mühlenbernd, G. Li, B. Bai, Q. Tan, G. Jin, C.-W. Qiu, S. Zhang, and T. Zentgraf, "Dual-polarity plasmonic metalens for visible light," *Nature communications,* vol. 3, p. 1198, 2012.
[88] D. M. Whittaker and I. S. Culshaw, "Scattering-matrix treatment of patterned multilayer photonic structures," *Physical Review B,* vol. 60, pp. 2610–2618, 07/15/ 1999.
[89] A. Martins, J. Li, A. F. da Mota, Y. Wang, L. G. Neto, J. P. do Carmo, F. L. Teixeira, E. R. Martins, and B.-H. V. Borges, "Crystalline Silicon (c-Si) Metasurface Holograms in the Visible Range," in *Imaging and Applied Optics 2018 (3D, AO, AIO, COSI, DH, IS, LACSEA, LS&C, MATH, pcAOP)*, Orlando, Florida, 2018, p. DTh2E.5.
[90] D. E. Aspnes and A. Studna, "Dielectric functions and optical parameters of si, ge, gap, gaas, gasb, inp, inas, and insb from 1.5 to 6.0 ev," *Physical review B,* vol. 27, p. 985, 1983.
[91] H. Pang, S. Yin, Q. Deng, Q. Qiu, and C. Du, "A novel method for the design of diffractive optical elements based on the Rayleigh–Sommerfeld integral," *Optics and Lasers in Engineering,* vol. 70, pp. 38-44, 2015.
[92] E. D. Palik, *Handbook of optical constants of solids*. Orlando: Academic Press, 1985.
[93] A. Martins, J. Li, A. F. da Mota, V. M. Pepino, Y. Wang, L. G. Neto, F. L. Teixeira, E. R. Martins, and B.-H. V. Borges, "Broadband c-Si metasurfaces with polarization control at visible wavelengths: applications to 3D stereoscopic holography," *Optics Express,* vol. 26, pp. 30740-30752, 2018/11/12 2018.
[94] F. Aieta, M. A. Kats, P. Genevet, and F. Capasso, "Multiwavelength achromatic metasurfaces by dispersive phase compensation," *Science,* vol. 347, pp. 1342-1345, 2015.
[95] J. Deng, Z. Li, G. Zheng, J. Tao, Q. Dai, L. Deng, P. a. He, Q. Deng, and Q. Mao, "Depth perception based 3D holograms enabled with polarization-independent metasurfaces," *Optics Express,* vol. 26, pp. 11843-11849, 2018.
[96] W. Zhao, B. Liu, H. Jiang, J. Song, Y. Pei, and Y. Jiang, "Full-color hologram using spatial multiplexing of dielectric metasurface," *Optics letters,* vol. 41, pp. 147-150, 2016.
[97] 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, and D. P. Tsai, "GaN Metalens for Pixel-Level Full-Color Routing at Visible Light," *Nano Letters,* vol. 17, pp. 6345-6352, 2017/10/11 2017.
[98] Y. Yang, W. Wang, P. Moitra, I. I. Kravchenko, D. P. Briggs, and J. Valentine, "Dielectric meta-reflectarray for broadband linear polarization conversion and optical vortex generation," *Nano letters,* vol. 14, pp. 1394-1399, 2014.
[99] Z. Li, T. Zhang, Y. Wang, W. Kong, J. Zhang, Y. Huang, C. Wang, X. Li, M. Pu, and X. Luo, "Achromatic Broadband Super-Resolution Imaging by Super-Oscillatory Metasurface," *Laser & Photonics Reviews,* p. 1800064.
[100] F. Qin, Z. Liu, Z. Zhang, Q. Zhang, and J. Xiao, "Broadband full-color multichannel hologram with geometric metasurface," *Optics Express,* vol. 26, pp. 11577-11586, 2018.
[101] D. Wen, F. Yue, G. Li, G. Zheng, K. Chan, S. Chen, M. Chen, K. F. Li, P. W. H. Wong, and K. W. Cheah, "Helicity multiplexed broadband metasurface holograms," *Nature Communications,* vol. 6, p. 8241, 2015.
[102] I. H. Malitson, "Refraction and Dispersion of Synthetic Sapphire," *Journal of the Optical Society of America,* vol. 52, pp. 1377-1379, 1962/12/01 1962.
[103] "Blender, www.blender.com," ed.
[104] A. Komar, R. Paniagua-Domínguez, A. Miroshnichenko, Y. F. Yu, Y. S. Kivshar, A. I. Kuznetsov, and D. Neshev, "Dynamic Beam Switching by Liquid Crystal Tunable Dielectric Metasurfaces," *ACS Photonics,* 2018/02/08 2018.

[105] K. A. Matthew J. Dicken1, Imogen M. Pryce1, Luke A. Sweatlock, Elizabeth M. Boyd, Sameer Walavalkar, James Ma, and Harry A. Atwater, "Frequency tunable near-infrared metamaterials based on VO2 phase transition " *Optics Express,* vol. 17, pp. 18330- 18339, September 2009.
[106] Q. Wang, E. T. F. Rogers, B. Gholipour, C.-M. Wang, G. Yuan, J. Teng, and N. I. Zheludev, "Optically reconfigurable metasurfaces and photonic devices based on phase change materials," *Nature Photonics,* vol. 10, p. 60, 2015.
[107] L. Ju, B. Geng, J. Horng, C. Girit, M. Martin, Z. Hao, H. A. Bechtel, X. Liang, A. Zettl, Y. R. Shen, and F. Wang, "Graphene plasmonics for tunable terahertz metamaterials," *Nature Nanotechnology,* vol. 6, p. 630, 2011.
[108] Y.-W. Huang, H. W. H. Lee, R. Sokhoyan, R. A. Pala, K. Thyagarajan, S. Han, D. P. Tsai, and H. A. Atwater, "Gate-Tunable Conducting Oxide Metasurfaces," *Nano Letters,* vol. 16, pp. 5319-5325, 2016/09/14 2016.
[109] A. She, S. Zhang, S. Shian, D. R. Clarke, and F. Capasso, "Adaptive metalenses with simultaneous electrical control of focal length, astigmatism, and shift," *Science Advances,* vol. 4, 2018.
[110] L. Li, T. Jun Cui, W. Ji, S. Liu, J. Ding, X. Wan, Y. Bo Li, M. Jiang, C.-W. Qiu, and S. Zhang, "Electromagnetic reprogrammable coding-metasurface holograms," *Nature Communications,* vol. 8, p. 197, 2017/08/04 2017.
[111] J. Hu, T. Lang, and G.-h. Shi, "Simultaneous measurement of refractive index and temperature based on all-dielectric metasurface," *Optics Express,* vol. 25, pp. 15241-15251, 2017/06/26 2017.
[112] X. Xu, B. Peng, D. Li, J. Zhang, L. M. Wong, Q. Zhang, S. Wang, and Q. Xiong, "Flexible Visible–Infrared Metamaterials and Their Applications in Highly Sensitive Chemical and Biological Sensing," *Nano Letters,* vol. 11, pp. 3232-3238, 2011/08/10 2011.
[113] A. Martins, K. Li, J. Li, H. Liang, D. Conteduca, B.-H. V. Borges, T. F. Krauss, and E. R. Martins, "On Metalenses with Arbitrarily Wide Field of View," *Acs Photonics,* vol. 7, pp. 2073-2079, 2020/08/19 2020.
[114] H. Liang, A. Martins, B.-H. V. Borges, J. Zhou, E. R. Martins, J. Li, and T. F. Krauss, "High performance metalenses: numerical aperture, aberrations, chromaticity, and trade-offs," *Optica,* vol. 6, pp. 1461-1470, 2019.
[115] J. Y. Kim, N. B. Brauer, V. Fakhfouri, D. L. Boiko, E. Charbon, G. Grutzner, and J. Brugger, "Hybrid polymer microlens arrays with high numerical apertures fabricated using simple ink-jet printing technique," *Optical Materials Express,* vol. 1, pp. 259-269, 2011/06/01 2011.
[116] T. Gissibl, S. Thiele, A. Herkommer, and H. Giessen, "Two-photon direct laser writing of ultracompact multi-lens objectives," *Nature Photonics,* vol. 10, p. 554, 2016.
[117] T. Gissibl, S. Thiele, A. Herkommer, and H. Giessen, "Sub-micrometre accurate free-form optics by three-dimensional printing on single-mode fibres," *Nature communications,* vol. 7, p. 11763, 2016.
[118] M. Khorasaninejad and F. Capasso, "Metalenses: Versatile multifunctional photonic components," *Science,* vol. 358, p. eaam8100, 2017.
[119] N. A. Rubin, G. D'Aversa, P. Chevalier, Z. Shi, W. T. Chen, and F. Capasso, "Matrix Fourier optics enables a compact full-Stokes polarization camera," *Science,* vol. 365, p. eaax1839, 2019.
[120] F. Aieta, P. Genevet, M. Kats, and F. Capasso, "Aberrations of flat lenses and aplanatic metasurfaces," *Optics Express,* vol. 21, pp. 31530-31539, 2013/12/16 2013.
[121] R. Paniagua-Domínguez, Y. F. Yu, E. Khaidarov, S. Choi, V. Leong, R. M. Bakker, X. Liang, Y. H. Fu, V. Valuckas, L. A. Krivitsky, and A. I. Kuznetsov, "A Metalens with a Near-Unity Numerical Aperture," *Nano Letters,* vol. 18, pp. 2124-2132, 2018/03/14 2018.
[122] P. Lalanne and P. Chavel, "Metalenses at visible wavelengths: past, present, perspectives," *Laser & Photonics Reviews,* vol. 11, p. 1600295, 2017.

[123] H. Liang, Q. Lin, X. Xie, Q. Sun, Y. Wang, L. Zhou, L. Liu, X. Yu, J. Zhou, T. F. Krauss, and J. Li, "Ultrahigh Numerical Aperture Metalens at Visible Wavelengths," *Nano Letters,* vol. 18, pp. 4460-4466, 2018/07/11 2018.
[124] A. Arbabi, Y. Horie, A. J. Ball, M. Bagheri, and A. Faraon, "Subwavelength-thick lenses with high numerical apertures and large efficiency based on high-contrast transmitarrays," *Nature Communications,* vol. 6, p. 7069, 2015.
[125] W. T. Chen, A. Y. Zhu, M. Khorasaninejad, Z. Shi, V. Sanjeev, and F. Capasso, "Immersion meta-lenses at visible wavelengths for nanoscale imaging," *Nano letters,* vol. 17, pp. 3188-3194, 2017.
[126] A. Kalvach and Z. Szabó, "Aberration-free flat lens design for a wide range of incident angles," *JOSA B,* vol. 33, pp. A66-A71, 2016.
[127] 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, p. 13682, 2016.
[128] B. Groever, W. T. Chen, and F. Capasso, "Meta-Lens Doublet in the Visible Region," *Nano Letters,* vol. 17, pp. 4902-4907, 2017/08/09 2017.
[129] M. Pu, X. Li, Y. Guo, X. Ma, and X. Luo, "Nanoapertures with ordered rotations: symmetry transformation and wide-angle flat lensing," *Optics Express,* vol. 25, pp. 31471-31477, 2017/12/11 2017.
[130] Y. Guo, X. Ma, M. Pu, X. Li, Z. Zhao, and X. Luo, "Ultrathin Metalenses: High-Efficiency and Wide-Angle Beam Steering Based on Catenary Optical Fields in Ultrathin Metalens (Advanced Optical Materials 19/2018)," *Advanced Optical Materials,* vol. 6, p. 1870073, 2018.
[131] J. Engelberg, C. Zhou, N. Mazurski, J. Bar-David, A. Kristensen, and U. Levy, "Near-IR wide-field-of-view Huygens metalens for outdoor imaging applications," *Nanophotonics,* vol. 9, pp. 361-370, 2020.
[132] B. Groever, C. Roques-Carmes, S. J. Byrnes, and F. Capasso, "Substrate aberration and correction for meta-lens imaging: an analytical approach," *Applied Optics,* vol. 57, pp. 2973-2980, 2018/04/20 2018.
[133] Z. Lin, B. Groever, F. Capasso, A. W. Rodriguez, and M. Lončar, "Topology-Optimized Multilayered Metaoptics," *Physical Review Applied,* vol. 9, p. 044030, 2018.
[134] (Jan 20, 2020). *GSM Arena - iPhone 11 Pro - Technical Specifications*. Available: https://www.gsmarena.com/apple_iphone_11_pro_max-9846.php
[135] (Jan 20, 2020). *Apple iPhone 11 Pro - Technical Specifications*. Available: https://www.apple.com/iphone-11-pro/specs/
[136] M. Lee, H. Kim, and J. Paik, "Correction of barrel distortion in fisheye lens images using image-based estimation of distortion parameters," *IEEE ACCESS,* vol. 7, pp. 45723-45733, 2019.
[137] A. Martins, B. H. V. Borges, J. Li, T. F. Krauss, and E. R. Martins, "Photonic Intermediate Structures for Perovskite/c-Silicon Four Terminal Tandem Solar Cells," *IEEE Journal of Photovoltaics,* vol. 7, pp. 1190-1196, 2017.
[138] J. Zhao, A. Wang, and M. A. Green, "24·5% Efficiency silicon PERT cells on MCZ substrates and 24·7% efficiency PERL cells on FZ substrates," *Progress in Photovoltaics: Research and Applications,* vol. 7, pp. 471-474, 1999.
[139] O. Graydon, "The race for tandems," *Nature Photonics,* vol. 10, pp. 754-755, Dec 2016.
[140] A. Louwen, W. Van Sark, R. Schropp, and A. Faaij, "A cost roadmap for silicon heterojunction solar cells," *Solar Energy Materials and Solar Cells,* vol. 147, pp. 295-314, 2016.
[141] P. K. Nayak and D. Cahen, "Updated assessment of possibilities and limits for solar cells," *Advanced Materials,* vol. 26, pp. 1622-1628, 2014.
[142] M. A. Green, K. Emery, Y. Hishikawa, W. Warta, and E. D. Dunlop, "Solar cell efficiency tables (version 47)," *Progress in Photovoltaics: Research and Applications,* vol. 24, 2016.

[143] T. P. White, N. N. Lal, and K. R. Catchpole, "Tandem solar cells based on high-efficiency c-Si bottom cells: top cell requirements for> 30% efficiency," *IEEE Journal of Photovoltaics,* vol. 4, pp. 208-214, 2014.
[144] M. Filipič, P. Löper, B. Niesen, S. De Wolf, J. Krč, C. Ballif, and M. Topič, "CH3NH3PbI3 perovskite/silicon tandem solar cells: characterization based optical simulations," *Optics express,* vol. 23, pp. A263-A278, 2015.
[145] R. Sheng, A. W. Ho-Baillie, S. Huang, M. Keevers, X. Hao, L. Jiang, Y.-B. Cheng, and M. A. Green, "Four-Terminal Tandem Solar Cells Using CH3NH3PbBr3 by Spectrum Splitting," *The journal of physical chemistry letters,* vol. 6, pp. 3931-3934, 2015.
[146] D.-L. Wang, H.-J. Cui, G.-J. Hou, Z.-G. Zhu, Q.-B. Yan, and G. Su, "Highly efficient light management for perovskite solar cells," *Scientific reports,* vol. 6, 2016.
[147] S. Albrecht, M. Saliba, J. P. C. Baena, F. Lang, L. Kegelmann, M. Mews, L. Steier, A. Abate, J. Rappich, and L. Korte, "Monolithic perovskite/silicon-heterojunction tandem solar cells processed at low temperature," *Energy & Environmental Science,* vol. 9, pp. 81-88, 2016.
[148] G. Hodes, "Perovskite-based solar cells," *Science,* vol. 342, pp. 317-318, 2013.
[149] Y. Li, L. Meng, Y. M. Yang, G. Xu, Z. Hong, Q. Chen, J. You, G. Li, Y. Yang, and Y. Li, "High-efficiency robust perovskite solar cells on ultrathin flexible substrates," *Nature communications,* vol. 7, 2016.
[150] N. N. Lal, T. P. White, and K. R. Catchpole, "Optics and light trapping for tandem solar cells on silicon," *IEEE Journal of Photovoltaics,* vol. 4, pp. 1380-1386, 2014.
[151] J. Werner, L. Barraud, A. Walter, M. Bräuninger, F. Sahli, D. Sacchetto, N. Tétreault, B. Paviet-Salomon, S.-J. Moon, and C. Allebé, "Efficient Near-Infrared-Transparent Perovskite Solar Cells Enabling Direct Comparison of 4-Terminal and Monolithic Perovskite/Silicon Tandem Cells," 2016.
[152] H. Uzu, M. Ichikawa, M. Hino, K. Nakano, T. Meguro, J. L. Hernández, H.-S. Kim, N.-G. Park, and K. Yamamoto, "High efficiency solar cells combining a perovskite and a silicon heterojunction solar cells via an optical splitting system," *Applied Physics Letters,* vol. 106, p. 013506, 2015.
[153] A. Bielawny, C. Rockstuhl, F. Lederer, and R. B. Wehrspohn, "Intermediate reflectors for enhanced top cell performance in photovoltaic thin-film tandem cells," *Optics express,* vol. 17, pp. 8439-8446, 2009.
[154] P. G. O'Brien, A. Chutinan, K. Leong, N. P. Kherani, G. A. Ozin, and S. Zukotynski, "Photonic crystal intermediate reflectors for micromorph solar cells: a comparative study," *Optics express,* vol. 18, pp. 4478-4490, 2010.
[155] A. Bielawny, J. Üpping, P. T. Miclea, R. B. Wehrspohn, C. Rockstuhl, F. Lederer, M. Peters, L. Steidl, R. Zentel, and S. M. Lee, "3D photonic crystal intermediate reflector for micromorph thin-film tandem solar cell," *physica status solidi (a),* vol. 205, pp. 2796-2810, 2008.
[156] S. Fahr, C. Rockstuhl, and F. Lederer, "Sandwiching intermediate reflectors in tandem solar cells for improved photon management," *Applied Physics Letters,* vol. 101, p. 133904, 2012.
[157] S. Fahr, C. Rockstuhl, and F. Lederer, "Metallic nanoparticles as intermediate reflectors in tandem solar cells," *Applied Physics Letters,* vol. 95, p. 121105, 2009.
[158] S. Fahr, C. Rockstuhl, and F. Lederer, "The interplay of intermediate reflectors and randomly textured surfaces in tandem solar cells," *Applied Physics Letters,* vol. 97, p. 173510, 2010.
[159] T. Todorov, O. Gunawan, and S. Guha, "A road towards 25% efficiency and beyond: perovskite tandem solar cells," *Molecular Systems Design & Engineering,* 2016.
[160] P. Löper, S.-J. Moon, S. M. De Nicolas, B. Niesen, M. Ledinsky, S. Nicolay, J. Bailat, J.-H. Yum, S. De Wolf, and C. Ballif, "Organic–inorganic halide perovskite/crystalline silicon four-terminal tandem solar cells," *Physical Chemistry Chemical Physics,* vol. 17, pp. 1619-1629, 2015.

[161] G. E. Eperon, V. M. Burlakov, A. Goriely, and H. J. Snaith, "Neutral Color Semitransparent Microstructured Perovskite Solar Cells," *ACS Nano,* vol. 8, pp. 591-598, 2014/01/28 2014.
[162] C. Roldán-Carmona, O. Malinkiewicz, A. Soriano, G. M. Espallargas, A. Garcia, P. Reinecke, T. Kroyer, M. I. Dar, M. K. Nazeeruddin, and H. J. Bolink, "Flexible high efficiency perovskite solar cells," *Energy & Environmental Science,* vol. 7, pp. 994-997, 2014.
[163] Z. M. Beiley, M. G. Christoforo, P. Gratia, A. R. Bowring, P. Eberspacher, G. Y. Margulis, C. Cabanetos, P. M. Beaujuge, A. Salleo, and M. D. McGehee, "Semi-Transparent Polymer Solar Cells with Excellent Sub-Bandgap Transmission for Third Generation Photovoltaics," *Advanced Materials,* vol. 25, pp. 7020-7026, 2013.
[164] G. Y. Margulis, M. G. Christoforo, D. Lam, Z. M. Beiley, A. R. Bowring, C. D. Bailie, A. Salleo, and M. D. McGehee, "Spray Deposition of Silver Nanowire Electrodes for Semitransparent Solid-State Dye-Sensitized Solar Cells," *Advanced Energy Materials,* vol. 3, pp. 1657-1663, 2013.
[165] J. Zhao, A. Wang, M. A. Green, and F. Ferrazza, "19.8% efficient “honeycomb” textured multicrystalline and 24.4% monocrystalline silicon solar cells," *Applied Physics Letters,* vol. 73, pp. 1991-1993, 1998.
[166] A. Yariv and P. Yeh, *Optical waves in crystals* vol. 5: Wiley New York, 1984.
[167] T. Zentgraf, T. Meyrath, A. Seidel, S. Kaiser, H. Giessen, C. Rockstuhl, and F. Lederer, "Babinet’s principle for optical frequency metamaterials and nanoantennas," *Physical Review B,* vol. 76, p. 033407, 2007.
[168] E. Feigenbaum and M. Orenstein, "Ultrasmall volume plasmons, yet with complete retardation effects," *Physical review letters,* vol. 101, p. 163902, 2008.
[169] E. Popov and M. Nevière, "Maxwell equations in Fourier space: fast-converging formulation for diffraction by arbitrary shaped, periodic, anisotropic media," *Journal of the Optical Society of America A,* vol. 18, pp. 2886-2894, 2001/11/01 2001.
[170] J. R. DeVore, "Refractive Indices of Rutile and Sphalerite," *Journal of the Optical Society of America,* vol. 41, pp. 416-419, 1951/06/01 1951.
[171] W. C. Chew, *Waves and fields in inhomogeneous media*: IEEE press, 1995.
[172] Wikipedia. (2021/01/07). *1951 USAF resolution test chart*. Available: https://en.wikipedia.org/wiki/1951_USAF_resolution_test_chart
[173] F. Wyrowski, "Diffractive optical elements: iterative calculation of quantized, blazed phase structures," *Journal of the Optical Society of America A,* vol. 7, pp. 961–969, 1990/06/01 1990.