Modelling nanostructure-enhanced upconversion for photovoltaic applications

PhD Thesis

Emil Haldrup Eriksen
Department of Physics and Astronomy
Aarhus University, Denmark
This thesis has been submitted to the Graduate School of Science and Technology (GSST) at Aarhus University, in order to fulfill the requirements for obtaining a PhD in Physics. The work has been carried out under the supervision of Professor Peter Balling and Associate Professor Søren Peder Madsen at the Department of Physics and Astronomy (IFA) at Aarhus University, Denmark.

This document was compiled and typeset in \LaTeX

September, 2018
Abstract

Electricity generation is currently transitioning from being based almost exclusively on fossil fuels towards more renewable sources. One of the key technologies in this transition is photovoltaics (PV). The majority of commercial PV systems are based on silicon cells with efficiencies around 20%. The main losses are due to the spectral mismatch between the solar radiation and the band gap of the absorber. Photons with an energy higher than the band gap are not converted efficiently, while photons with a lower energy are not converted at all. The latter type of losses can be addressed via upconversion (UC), i.e. the merging of multiple photons into one photon with a higher energy.

Due to the non-linear nature of UC processes, a much higher energy density than what is available at one sun is necessary to achieve efficient UC. Light focusing via conventional lenses is impractical and economically unfeasible. In the SunTune project, which this work is a part of, the possibility of utilizing nanostructures to achieve light concentration is explored. The focus of this work is on the prediction of the influence of particular plasmonic and/or photonic nanostructures on near-by UC systems through numerical modelling.

A framework has been developed, which couples the local changes in energy density and decay rates via a rate-equation model. The main focus has been on plasmonic single-particle systems, though a two-particle model was formulated to enable an assessment of particle-particle coupling effects. Furthermore, a branch for treating one-dimensional photonic crystals was implemented. As our in-house electron-beam lithography system was unable to fabricate the nanostructures with sufficient accuracy, a resolution-enhancing dose correction scheme was developed.

It was found that while the near-field of weakly-coupled plasmonic metal nanoparticles enable high local energy densities, a strongly correlated increase in quenching limits applicability for enhancing UC efficiency. A periodic design less prone to quenching was developed, yielding a UC luminescence enhancement of nearly two orders of magnitude in simulation. With the first experimental prototype achieving more than 20 times enhancement, the design seems promising. One-dimensional photonic structures yielding up to 300 times UC luminescence enhancement were identified. For this seemingly superior device, the main challenge lies in achieving sufficient accuracy in fabrication.
Dansk resumé

Elektricitetsproduktionen bevæger sig lige nu fra at være baseret primært på fossile brændstoffer til at inkludere en voksende andel af vedvarende energikilder. En af nøgleteknologierne i denne omstilling er solceller. Størstedelen af de kommercielle solceller er baseret på silicium og har en effektivitet på omkring 20%. De primære tab skyldes, at solens spektrum ikke passer med båndgabet af det absorberende materiale. Fotoner med en højere energi end båndgabet konverteres ikke effektivt, mens fotoner med en lavere energi slet ikke absorberes. Det sidstnævnte tab kan mindskes ved hjælp af opkonvertering, dvs. sammensmeltning af flere fotoner til én foton med en højere energi.


Beregninger viser, at svagt kobledes plasmoniske metalnanopartikler muliggør høj energiæthed i nærfeltet, men en korreleret forøgelse af quenching begrænser potentiotalet for anvendelse indenfor opkonvertering. Et periodisk design, der lader mindre af quenching, blev udviklet. Den beregnede forbedring i opkonverteringsluminescens er næsten to størrelsesordner, og da den første prototype allerede viser en forbedring på mere end en faktor 20, virker designet lovende. En-dimensionelle fotoniske krystaller, der giver op til 300 gange forbedring i opkonverteringsluminescens blev identificeret. For denne tilsyneladende overlegne struktur er udfordringen fremstilling med den nødvendige præcision.
Acknowledgements

First and foremost, I must express my sincere gratitude to my main and co supervisors, Peter Balling and Søren Peder Madsen, for their guidance. For always taking time for discussions, and for pushing me in the right direction when I was heading off-road. Knowing that your doors were always open has been invaluable.

I would like to thank the semiconductor group for providing a productive, yet cheerful, working environment. In particular Brian Julsgaard for insightful scientific dialogues, Joakim Vester-Petersen for numerical discussions, Harish Lakhotiya for the fabrication of upconverting materials, and Adnan Nazir for making the physical realization of plasmonic nanostructures possible. A special thanks goes to my office roommates Søren Hanghøj Møller, Jeppe Christiansen and Søren Roesgaard. Not only for carrying out optical measurements that support this work, but also for their continuos engagement in scientific (and non-scientific) discussions, coffee drinking and cake eating.

I also need to express my gratitude to the Novel Solar Cell Concepts group at Fraunhofer ISE in Freiburg for making an amazing research stay abroad possible. Especially Jan Christoph Goldschmidt for his guidance and inspiration, and Clarissa Hofmann for being not only a brilliant collaborator, but a true friend.

Finally, I am grateful for the love, encouragement, and support I have received from Astrid Jacobsen Harpøth, my better half, who has been at my side throughout this journey.

The Innovation Fund Denmark is acknowledged for the funding that made SunTune, and thus this PhD project, possible.

Aarhus, September 2018
Emil Haldrup Eriksen
List of publications

Publication I

Publication II

Publication III

Publication IV

Publication V
Co-author publications

Publication VI

Publication VII

Publication VIII

Publication IX

Publication X
Publication XI

Publication XII

Publication XIII

Conference papers

Publication XIV

Publication XV
**ACRONYMS**

**BC**  boundary condition.

**DBR**  distributed Bragg reflector.

**EBL**  electron-beam lithography.

**EOT**  extraordinary optical transmission.

**ESA**  excited-state absorption.

**ET**  energy transfer.

**FEM**  finite-element method.

**FRET**  Förster resonant energy transfer.

**GSA**  ground-state absorption.

**GUI**  graphical user interface.

**LDOS**  local density of optical states.

**LSPR**  localized surface plasmon resonance.

**MC**  monte carlo.

**MPR**  multi-phonon relaxation.

**NP**  nanoparticle.
PBG  photonic band gap.
P EC  perfect electric conductor.
P MC  perfect magnetic conductor.
PML  perfectly matched layer.
PV  photovoltaics.
QY  quantum yield.
REM  rate-equation model.
SBC  scattering boundary condition.
SEM  scanning-electron microscopy.
SP  surface plasmon.
SPE  spontaneous emission.
TEM  transmission-electron microscopy.
TMM  transfer-matrix method.
UC  upconversion.
UCL  upconversion luminescence.
UCNC upconverting nanocrystal.
UCQY upconversion quantum yield.
Introduction

Since the invention of the light bulb by Thomas Edison in 1879, humans have been consuming an ever increasing amount of electricity. In the past, the electricity was produced almost exclusively by burning fossil fuels, but during the last decades various renewable electricity sources have been commercialized. As of 2015, roughly 25% of the global electricity consumption is covered by renewables, see figure 1.1, and the renewable penetration is expected to continue rising in future[1, 2]. The transition towards an electricity system based on renewables is driven by the increasing awareness of the unpleasant environmental side effects of burning fossil fuels, as well as the realization that the fossil resources will be depleted within a foreseeable future. Nearly all countries have put forward policies supporting the development and deployment of renewable technologies[1, 2].

The current key players are hydro power, biomass, wind power and photovoltaics (PV), see figure 1.1, of which the last two hold the greatest potential for large scale expansion. With almost 540 GW of globally installed capacity, wind power is in the lead, but the growth is slowing down. On the contrary, PV has experienced an exponential growth since the mid-1990s, and the this trend is expected to continue in the
Chapter 1. Introduction

Figure 1.1: Estimated share of global electricity production. The first two charts show end-2017 numbers, while the last is based on data from 2016[2, 3].

years to come. The cost of solar modules keeps decreasing, roughly following a learning curve with a 20% reduction in price for each doubling of the cumulative shipped capacity[2, 3]. Meanwhile, the costs of other system elements are not reduced at the same pace. As the module price accounts for a decreasing fraction of the system cost, less than 20% as of 2017, the alternative strategy of increasing the module efficiency (rather than reducing the cost) becomes increasingly favorable.

The main losses in current single-junction solar cells are caused by the spectral mismatch between the solar spectrum and the band gap of the absorber. If the photon energy is larger than the band gap, excess energy is dissipated as heat (thermalization losses). If the photon energy is smaller than the band gap, the photon is not absorbed at all (transmission losses). Crystalline silicon, which accounts for more than 90% of the market (see figure 1.1), has a band gap of 1.1 eV (1100 nm). The resulting efficiency limit, commonly denoted as the Shockley-Queisser (SQ) limit[4], is roughly 29% under one sun. With 25% efficiency achieved in 1999[5], not much headroom was left for further improvements. Indeed, the progress has been only marginal since, yielding a current world record efficiency of 26.3%[6]. Slightly higher efficiencies can be achieved using gallium arsenide, see figure 1.2, but the commercial application is hindered by a significantly higher material cost as compared to silicon.

The stagnation of the efficiency of conventional silicon-based cells has facilitated research into concepts that enables efficiencies beyond the
Chapter 1. Introduction

Figure 1.2: Evolution of solar cell efficiencies over time. This plot is courtesy of the National Renewable Energy Laboratory, Golden, CO.

SQ limit. The most successful approach is multi-junction cells, where a combination of junctions with different band gaps work together to enable efficient conversion of a larger fraction of the solar spectrum. The current world record is held by a 4-junction cell, yielding an efficiency of 46% at 500 suns, see figure 1.2. However, the cost of multi-junction cells remains too high for commercial applications, except maybe for concentrated PV.

The other main path to overcome the SQ limit is via spectral conversion. That is, rather than tailoring the cell to match the solar spectrum (multi-junction cells), the solar spectrum is tuned to match the range of efficient current generation of the cell. Two main approaches exist; quantum cutting, where one high energy photon is split into multiple photons with a lower energy, and upconversion (UC), where multiple photons are merged into a single photon with a higher energy. While UC can limit transmission losses, quantum cutting can potentially reduce thermalization losses. Since quantum cutting involves photons, which are energetic enough to cross the band gap in the first place, the quantum cutter must be placed between the cell and the excitation source (to
avoid absorption in the cell), making integration without losses difficult. In addition, a quantum efficiency > 100% (including any losses related to integration) is needed to achieve a net improvement. UC, on the other hand, involves only photons which cannot be absorbed by cell in the first place. The upconverter can therefore be placed behind the cell, and any quantum efficiency > 0% will result in a net improvement.

The SunTune project, which this work is a part of, focus on UC. In particular on UC processes suitable for improving the efficiency of solar cells based on crystalline silicon. For such cells, an ideal upconverter raises the theoretical efficiency limit by almost 10 percent points\cite{7}. Erbium-based compounds have been identified as spectrally suitable, enabling UC from around 1500 nm to 980 nm (the exact numbers depend on the host material). Adding a downshifter, the spectral range that can be upconverted can be extended beyond the absorption range of erbium up to the silicon band gap. Both processes are shown in figure 1.3. The insert illustrates a simple scheme for integrating the upconverter and the downshifter with the solar cell. The realistically achievable gain using such a configuration is expected to be in the range of few percent points\cite{8}.

The main hindrance for commercial exploitation of UC in solar cells is the low efficiency of current UC materials. Due to the non-linear nature of the process, a much higher energy density than what is available at standard operation conditions is required to achieve efficient UC. While light concentration at macroscopic scale is possible, e.g. via Fresnel lenses as in concentrated PV systems\cite{9}, it is not suitable for standard solar cells due to practical as well as economical considerations. Plasmonic metal nanoparticles (NPs) enable light concentration below the diffraction limit, and thus extremely high energy densities. Such nanoscale light concentrators, possibly integrated within the upconverter, pose a more viable option for increasing the UC efficiency. One of the main goals of the SunTune project is to investigate this possibility, and ultimately to determine how such NPs should be designed.
1.1 State of the art

The concept of upconversion dates back to the 1950s[10], but it was not until the 1990s that applicability within PV was demonstrated[11]. Since the 1980s, it has been known that fluorescence enhancement via plasmonic structures is possible [12]. The application to upconversion processes was demonstrated a few centuries later[13, 14], and since then a large body of studies on plasmonically enhanced UC has materialized[15, 16]. The main focus has been on excitation at 980 nm, with only few studies considering excitation at 1500 nm. To the knowledge of this author, the highest UC enhancement (for 1500 nm excitation) was reported by Verhagen et al. [17]. Using erbium ions embedded in a sapphire substrate to probe the field enhancement near a structured
gold film (due to the excitation of surface plasmons), a 450-fold enhancement of the 980 nm emission was achieved. However, the enhancement is measured only across thin slice (a few tens of nm). In PV applications, the important quantity is the external upconversion quantum yield (UCQY), and the enhancement must thus be extended across a volume large enough to absorb most of the incident radiation. Verhagen et al. \cite{17} does not provide any UCQY values. It thus remains unclear whether UC enhancements in this range can be realized in PV applicable devices. Theoretical work on plasmonically enhanced UC has been carried out by Fischer et al. \cite{18,19}, who has evaluated potential benefits of placing the upconverter close to spherical gold nanoparticles. It was found that even though the UCQY was enhanced in some regions, the decrease observed in others regions result in a net decrease in the UCQY upon integration across the simulation volume.

1.2 Main objectives of this work

As part of the SunTune project, a key objective of this work is the development of a modelling setup capable of predicting how the presence of plasmonic structures influences the UC process in erbium-based compounds. Enabling rapid and cost-effective assessment of potential designs, as compared to physical prototype fabrication, accurate modelling capabilities are of crucial importance to the SunTune project.

The dynamics of the UC process can be described via a rate-equation model (REM). From a given set of initial conditions (and model parameters), it enables the prediction of the main figures of merit, i.e. the UCQY as well as the upconversion luminescence (UCL). A REM thus enters as a key component in the modelling setup. Based on previously determined parameters for $\beta$-NaYF$_4$:Er$^{3+}$ by Fischer et al. \cite{20}, a REM is implemented as part of this work. As a simpler alternative, and to address a broader range of materials system, approximate closed-form estimates based on the solution of a simplified REM are also considered. In the vicinity of plasmonic structures, UC systems are subject to a modification of the excitation conditions as well as the local photonic environment\cite{19}. Assuming that the structures are not too small\cite{21}, both effects can be treated via Maxwells equations. A broad variety of meth-
ods have been developed for solving these equations, each superior for particular types of problems. In SunTune, the main family of structures considered is thin-film stacks with the plasmonic structure situated on top. Therefore, a modelling framework tailored for this type of system is developed. As the calculations can be computationally complex, a key property of the setup is seamless integration with cluster computation facilities. Besides developing the modelling tools, an integral part of this work is the application of the tools to various systems. Both to support other parts of the SunTune project, e.g. to determine geometrical parameters to achieve desired plasmonic properties, but also to carry out independent simulation-based studies.

1.3 Outline

Chapter 2 provides an introduction to the concept of UC, focusing on erbium-based compounds, along with an outline of the basic structure of the REM setup. Furthermore, the main concepts for increasing the UC efficiency, i.e. via plasmonic or photonic structures, are presented. Chapter 3 contains a short review of relevant concepts from classical electrodynamics. The main modelling approaches, the transfer-matrix method (TMM) and the finite-element method (FEM), are described in chapter 4. Since a large part of my time has been spent on writing code, a brief overview of some of the resulting software has been included in chapter 5. The following five chapters cover my main publications. In each chapter, the paper context is outlined and my contributions are specified before the paper is included, either in published or pre-print form depending on the state of completion. Chapter 11 contains a summary of each of my co-author publications along with an outline of my contributions to each paper. The main results are summarized in chapter 12.

In accordance with GSST rules, parts of this thesis was also used in the progress report for the qualifying examination.
This chapter contains a short introduction to the concept of upconversion as well as the rate-equation-modelling framework used in this work. Furthermore, the two strategies considered in this work for increasing the upconversion efficiency, i.e. via plasmonic structures or photonic crystals, are outlined.
2.1 Introduction

The upconversion (UC) concept was developed in the late 1950s, where Bloembergen\cite{10} proposed a scheme for infrared-to-visible UC to enable detection of infrared photons via excited-state absorption (ESA), see figure 2.1a. A few years later, the process was realized in experiment using Pr$^{3+}$ ions embedded in a LaCl$_3$ lattice\cite{22, 23}. In 1966, Auzel suggested that UC efficiencies could be increased by exploitation of energy transfer (ET) processes, see figure 2.1b, between different types of ions\cite{24}. Utilizing this concept, in 1972 Menyuk\cite{25} achieved an UC efficiency of a few percent (at 980 nm excitation) using a combination of Er$^{3+}$ and Yb$^{3+}$ ions embedded in a NaYF$_4$ lattice.

![Figure 2.1: Schematic illustration of two different UC mechanisms. In (a), an electron is excited from the ground state $|g\rangle$ to the final state $|f\rangle$ via ground-state absorption (GSA) followed by ESA. As the electron decays from $|f\rangle$ directly to $|g\rangle$, an upconverted photon is emitted. In (b), two electrons (in separate ions) are promoted to the intermediate state, $|i\rangle$, through GSA. Subsequently, the donor (D) decays back to $|g\rangle$ by transferring its energy non-radiatively to the acceptor (A), thereby promoting it to $|f\rangle$. Finally, an upconverted photon is emitted as the electron in $|f\rangle$ decays directly to $|g\rangle$. Radiative (non-radiative) transitions are indicated by solid (dashed) lines.](image)

To this day, fluorides remain among the most popular host materials due to their low phonon energies (to limit non-radiative decays) and their chemical stability. The preferred dopants are (combinations of) lanthanides due to the ladder-like energy spacing of their 4f electronic states. For UC applications in silicon-based solar cells, the Er$^{3+}$ ion is currently considered to be the most well-suited. It enables UC from
around 1500 nm to 980 nm, thus bridging the silicon band gap. The most efficient materials are monocrystalline BaY$_2$F$_8$:Er$^{3+}$, along with microcrystalline Gd$_2$O$_2$S:Er$^{3+}$ and β-NaYF$_4$:Er$^{3+}$, all with reported internal quantum efficiencies exceeding 10%[26]. The efficiency of nanomaterials tends to be significantly lower due to increased surface quenching[27], though this effect can be counteracted to some degree via surface passivation[28].

2.2 Modelling upconversion in erbium-based compounds

Figure 2.2 shows the first seven energy levels of the 4f electronic states in Er$^{3+}$. The numbers are based on β-NaYF$_4$:Er$^{3+}$, but due to the shielding from the crystal field of the host lattice provided by the outer electrons (5s, 5p), the 4f electronic states are relatively insensitive to the choice of host material. Excitation at 1523 nm enables GSA from level 1 to level 2 and ESA from level 2 to level 4. Alternatively level 4 can be reached via GSA in two separate ions followed by ET. When an electron reaches level 4, a (fast) multi-phonon relaxation (MPR) brings it to level 3. The main UC emission occurs as the electron decays from level 3 back to level 1 through spontaneous emission (SPE). The dynamics of the upconversion process can be modeled via a rate-equation model (REM). A REM is a set of coupled differential equations that describes the population of (relevant) energy levels. To model the UC process at hand, at least the first four levels must be included. Assuming that the system is excited monochromatically (at frequency $\nu$) with an intensity $I$, a simple model can be expressed as
2.2 Modelling upconversion in erbium-based compounds

\[ \dot{\rho}_1 = -\sigma_{12} \frac{I}{h\nu} \rho_1 + \left[ \gamma_{21} + \sigma_{21} \frac{I}{h\nu} \right] \rho_2 + \gamma_{31} \rho_3 + \gamma_{41} \rho_4 \]  
\[ + W_{22,14} \rho_2^2 - W_{14,22} \rho_1 \rho_4 \]  
\[ \dot{\rho}_2 = + \sigma_{12} \frac{I}{h\nu} \rho_1 - \left[ \gamma_{21} + (\sigma_{21} + \sigma_{24}) \frac{I}{h\nu} \right] \rho_2 + \gamma_{32} \rho_3 \]  
\[ + \left[ \gamma_{42} + \sigma_{42} \frac{I}{h\nu} \right] \rho_4 - 2W_{22,14} \rho_2^2 + 2W_{14,22} \rho_1 \rho_4 \]  
\[ \dot{\rho}_3 = - \left[ \gamma_{31} + \gamma_{32} \right] \rho_3 + \gamma_{43} \rho_4 \]  
\[ \dot{\rho}_4 = + \sigma_{24} \frac{I}{h\nu} \rho_2 - \left[ \gamma_{41} + \gamma_{42} + \gamma_{43} + \sigma_{42} \frac{I}{h\nu} \right] \rho_4 \]  
\[ + W_{22,14} \rho_2^2 - W_{14,22} \rho_1 \rho_4, \]  

where \( \rho_i \) is the population density of level \( i \) (and the dot denotes differentiation with respect to time), \( \sigma_{ij} \) the absorption cross section for the transition from level \( i \) to level \( j \), \( \frac{I}{h\nu} \) the photon flux, \( \gamma_{ij} \) the rate of decay from level \( i \) to level \( j \) and \( W_{ij,kl} \) rates of ET from levels \( i \) and \( j \) to levels \( k \) and \( l \). In general, the decay rates \( \gamma_{ij} \) can contain both radiative and non-radiative components,

\[ \gamma_{ij} = \gamma^R_{ij} + \gamma^{nr}_{ij}. \]  

The system parameters \( (\sigma_{ij}, \gamma_{ij}, W_{ij,kl}) \) must be derived from experiments, as a purely theoretical evaluation is not possible.

The main figure of merit, the upconversion luminescence (UCL), can be calculated from the steady-state solution of the REM. As a simpler alternative, an approximate analytical expression can be derived[29]. First, it is assumed that the concentration of \( \text{Er}^{3+} \) ions is high (\( \geq 25\% \)), so that ET will be the dominant UC mechanism[30], i.e. the \( \sigma_{24} \) term can be neglected. As the MPR decay \( \gamma_{43} \) is fast, the \( \gamma_{42}, \gamma_{41} \) and \( \sigma_{42} \) terms can be dropped. From studies of a more complex REM considering the first seven levels of \( \beta\text{-NaYF}_4:\text{Er}^{3+} \)[20] it was found that \( \rho_2 \ll \rho_1 \approx 1 \) unless the excitation intensity is very high (\( I \gg 100 \text{ kW/m}^2 \)), i.e. the \( \sigma_{21} \) term can be neglected and \( \rho_1 \) replaced by unity. Furthermore, \( \gamma_{32} \ll \gamma_{31} \), so
Chapter 2. Upconversion

Figure 2.2: The first seven energy levels of the Er\(^{3+}\) ion. In (a), the main emission lines are indicated. Processes retained in the simplified REM, equations (2.6) to (2.9) are shown in (b). The upper levels are greyed out as their dynamics are not considered. Radiative (non-radiative) transitions are indicated by solid (dashed) lines. Transition pairs that are part of the same ET process are indicated by the same color.

The retained processes are sketched in figure 2.2b. The steady-state solution of the simplified system can be determined analytically. Defining the saturation intensity as

\[
I_{\text{sat}} = h\nu \frac{\gamma_{21} (W_{14,22} + \gamma_{43})}{8\sigma_{12} W_{22,14} \gamma_{43}}, \quad (2.10)
\]
the steady-state population density of level three can be expressed as

$$\rho_3 = \frac{\sigma_{12} I_{\text{sat}}}{h \nu \gamma_{31}} \left( 1 - \sqrt{1 + \frac{I}{I_{\text{sat}}} + \frac{1}{2} \frac{I}{I_{\text{sat}}}} \right). \quad (2.11)$$

As the objective is ultimately to determine UCL enhancements, it is sufficient to determine the scaling of UCL with the system parameters, i.e. constant terms can be dropped. To simplify the estimate, changes in ET rates are neglected (see section 2.3.3) and $\gamma_{43}$ and $\sigma_{12}$ are assumed constant. From equation (2.11) it is clear that $\rho_3$ scales with $\gamma_{31}^{-1}$, but the scaling of with $\gamma_{21}$ (and $I$) depends on the ratio $I/I_{\text{sat}}$. In experiments, the UCL is observed to scale approximately as $I^{1.5}[31]$. In can be shown that in this case, the scaling of $\rho_3$ with $\gamma_{21}$ will be as $\gamma_{21}^{-1}$ [32]. The resulting scaling of the UCL, which is proportional to $\rho_3$ times $\gamma_{31}^{-1}$, is

$$\text{UCL} \propto \frac{\gamma_{31}^{-1}}{\gamma_{31} \gamma_{21}} I^{1.5} = QY_{31} \tau_{21} I^{1.5}. \quad (2.12)$$

The first term is the quantum yield (QY) of the main UC emission. In the absence of non-radiative decay channels for the transition from level 3 to level 1, the QY is unity. The second term accounts for the main loss mechanism, the decay from level 2 to level 1. A decrease in the lifetime $\tau$ results in increased losses and thus a lower UCL. Finally, the $I^m$ term, with $m = 1.5$, expresses the non-linear dependence on excitation power.

### 2.3 Increasing the efficiency of upconversion

The simple UCL estimate, equation (2.12), indicates the main paths to improve the efficiency of the UC process. One can either increase the radiative decay rate of the main UC emission (in a real system, non-radiative decay channels will typically be present), decrease the rate of the main loss emission and/or increase the local energy density and thus the photon flux. Manipulation of these properties can be achieved via plasmonic and/or photonic structures.

#### 2.3.1 Plasmonic structures

Surface plasmons (SPs) are coherent oscillations of the free conduction electrons in metals, typically observed at metal-dielectric interfaces. As
they are spatially confined near the metal surface, metallic nanoparticles (NPs) enable localization of the optical energy on a length scale $L$ defined by the size of the NP system, which can be smaller than the excitation wavelength $\lambda$. This might seem contradictory at first, as $\lambda$ is typically considered as the defining length scale for the confinement of electromagnetic energy. However, when $L \ll \lambda$, most of the energy is in fact electromechanical (rather than electromagnetic) as thus not bound by $\lambda$[33]. On resonance, the resulting sub-wavelength focusing of the optical field enables extremely high energy densities.

To illustrate a few key properties of localized surface plasmon resonances (LSPRs), the case of a spherical metal NP, for which the optical response can be described via simple, analytical expressions, is considered. A schematic illustration is shown in figure 2.3. In the quasi-static limit ($L \ll \lambda$), the absorption and scattering cross sections are[34]

$$\sigma_{\text{abs}} = 4\pi k L^3 \text{Im} \left[ \frac{\epsilon - \epsilon_m}{\epsilon + 2\epsilon_m} \right]$$

$$\sigma_{\text{sct}} = \frac{8}{3\pi k^4 L^6} \left( \frac{\epsilon - \epsilon_m}{\epsilon + 2\epsilon_m} \right)^2,$$

where $k$ denotes the wave vector in the medium, $L$ the radius of the sphere, while $\epsilon$ and $\epsilon_m$ denote the dielectric permittivity of the NP and the surrounding medium. Obviously, if $\epsilon = \epsilon_m$, i.e. the two materials are the same, $\sigma_{\text{abs}} = \sigma_{\text{sct}} = 0$. When the condition

$$\epsilon + 2\epsilon_m = 0$$

is fulfilled, $\sigma_{\text{abs}}$ and $\sigma_{\text{sct}}$ diverges, i.e. an ideal LSPR is achieved. Since the surrounding (dielectric) medium has $\text{Re}(\epsilon_m) > 0$ and $\text{Im}(\epsilon_m) \geq 0$, equation (2.15) implies that $\text{Re}(\epsilon) < 0$ and $\text{Im}(\epsilon) \leq 0$. The first condition is fulfilled for metals in the visible region. However, the simultaneous fulfillment of the second condition (zero losses) is prohibited by the Kramers-Kronig relations[34]. Hence, $\sigma_{\text{abs}}$ and $\sigma_{\text{sct}}$ remain finite, with the most prominent LSPRs expected for low loss materials ($\text{Im}[\epsilon]$ close to zero). Indeed, strong LSPRs are observed in the visible region for metal NPs made of low loss materials such as silver and gold.
If the dielectric function of the metal is described via a simple Drude model (losses are neglected for the moment),

\[
\frac{\epsilon (\omega)}{\epsilon_0} = 1 - \frac{\omega_p^2}{\omega^2},
\]

(2.16)

where \(\omega_p\) is the bulk plasmon frequency and \(\omega\) denotes the excitation frequency, equation (2.15) can be rewritten as

\[
1 - \frac{\omega_p^2}{\omega^2} + 2\frac{\epsilon_m}{\epsilon_0} = 0 \iff \omega = \frac{\omega_p}{\sqrt{1 + 2\frac{\epsilon_m}{\epsilon_0}}}. 
\]

(2.17)

From equation (2.17) it is clear that the LSPR is redshifted with respect to the bulk plasmon frequency, with the redshift increasing for higher \(\epsilon_m\). While real metals are not lossless, the same tendency is in fact observed in real systems. Similarly, upon lifting the quasi-static approximation, it can be shown that the LSPR redshifts with increasing size of the NP system[34].

### 2.3.2 Photonic crystals

Photonic crystals are materials with a spatially periodic variation in the dielectric function. The periodicity can be in one-, two- or three dimensions as illustrated in figure 2.4. A key property of photonic crystals is
their ability to form photonic band gaps (PBGs). That is, for some frequency window(s) light propagation is prohibited along certain directions. As particles (in this case photons) only interact with a periodic environment if the length scale of the periodicity is comparable to its wavelength, a periodicity in the range of 100 nm to 1 µm is needed to achieve photonic crystals operating at visible frequencies. Except for gemstone opals, photonic crystal do not form naturally. Modern nanofabrication techniques have, however, enabled artificial realization.

Figure 2.4: Conceptual illustration of (a) one-, (b) two-, and (c) three-dimensional photonic crystals. The different colors indicate different values of the dielectric function.

Photonic crystals are of great practical importance for the construction of various optoelectronic devices. If a photonic crystal exhibits a PBG across some desired frequency range, light propagation will be confined to any defective regions. This basic concept allows the formation of photonic wave guides (extended defects), capable of transporting the light at around tight corners with low losses, optical cavities with high quality factors (localized defects) as well as more complex devices such as diplexers[35].

A side effect of the PBG is perfect reflectance. This property can be leveraged for the construction of high quality reflectors. A commonly encountered example is the distributed Bragg reflector (DBR), which consists of alternating quarter-wave layers of two different dielectric materials (a one-dimensional photonic crystal). In this case, the low reflection can also be understood intuitively as destructive interference between the reflected waves at each interface. When such a stack is excited just below the PBG, light can be coupled in efficiently. The result of the
interfering waves is instead the formation of an oscillatory energy density pattern inside the crystal, yielding high energy densities at points of constructive interference.

2.3.3 Evaluation of upconversion enhancement

To assess the UC enhancement due to a plasmonic/photonic structure, all quantities will be stated relative to the values in a reference system. Here, the simple case of homogeneous UC material is considered and the reference values are marked by a 0-index. In the REM formulation of section 2.2, the excitation of the system was specified through the intensity $I$, which is convenient in an experimental context. However, for the systems considered in the work, the electric field can vary in three dimensions. Therefore, a more convenient measure is the energy density $u$, which scales with the electric field amplitude to the second power. It is related to the intensity as $u = \frac{n}{c_0} I$. Since the refractive index $n$ of the upconverter remains unchanged, the modified energy density can be included in the REM via the scaling,

$$I \rightarrow I(r) = \frac{|E(r)|^2}{|E_0|^2} I,$$

where $|E_0|$ can be evaluated analytically. In the general case, $E(r)$ can be calculated using the finite-element method (FEM), but for stacks of planar layers, e.g. finite one-dimensional photonic crystals, the transfer-matrix method (TMM) is much more efficient (see sections 4.1 and 4.2). Note that since $I$ becomes spatially dependent, rather than solving the REM once, it must now be solved at each spatial point.

The presence of plasmonic/photonic structures does not only change the excitation conditions, but also the dielectric environment as expressed by the dyadic Green function $\mathbf{G}(\mathbf{r}, \mathbf{r}_0; \omega)$. It describes the electric field $E$ generated at $\mathbf{r}$ due to a radiating dipole $\mathbf{p}$ located at $\mathbf{r}_0$[36],

$$\mathbf{E}(\mathbf{r}) = \frac{\omega^2 \mu}{4\pi} \mathbf{G}(\mathbf{r}, \mathbf{r}_0; \omega) \mathbf{p}.$$  

(2.19)

The modified dielectric environment will affect the UC dynamics through a change in the local density of optical states (LDOS), which scales as

$$\text{LDOS}(\mathbf{r}; \omega_{ij}) \propto \mathbf{n} \cdot \text{Im} \left[ \mathbf{G}(\mathbf{r}, \mathbf{r}; \omega_{ij}) \right] \cdot \mathbf{n},$$

(2.20)
where $\mathbf{n}$ is a normal vector indicating the polarization of the emission and $\omega_{ij}$ the transition frequency. Furthermore, the ET rates scale as

$$
\text{ET}(\mathbf{r}_A, \mathbf{r}_D; \omega_{ij}) \propto \left| \mathbf{n}_A \cdot \mathbf{G}(\mathbf{r}_A, \mathbf{r}_B; \omega_{ij}) \cdot \mathbf{n}_D \right|^2,
$$

(2.21)

where $\mathbf{r}_A (\mathbf{r}_D)$ indicates the position of the acceptor (donor) and $\mathbf{n}_A (\mathbf{n}_B)$ the polarization[36].

As per Fermi’s golden rule[37], the SPE probability scales with the LDOS, so the modified LDOS can be included in the REM via the scaling

$$
\gamma_{ij} \rightarrow \gamma(\mathbf{r})_{ij} = \frac{\text{LDOS}(\mathbf{r}; \omega_{ij})}{\text{LDOS}_0(\omega_{ij})} \cdot \gamma_{ij}.
$$

(2.22)

If metal is present (plasmonic structures), a fraction of the emitted energy will be dissipated as heat. This loss can be represented as a non-radiative decay channel, i.e. $\gamma_{ii}^{nr} \neq 0$. Both $\gamma_{ij}$ and the splitting into radiative and non-radiative components can be evaluated via FEM calculations (see section 4.2). However, for periodic lossless structures, e.g. photonic crystals, eigenmode calculations provide a more efficient alternative[38].

Changes in ET rates have not been considered in this work, but they could be incorporated in the REM via a scaling of the $W_{ij,kl}$ parameters. For plasmonic structures, which are the main focus of this work, large changes in the ET rates are typically limited to the immediate vicinity of metal surfaces, where the quenching is so high that the UCL is essentially zero[39, 40]. In this case, changes in ET rates can be neglected.
This chapter provides a short review of the classical electrodynamic concepts that provide the foundation for the optical modelling methods, which are discussed in chapter 4. Furthermore, various notation and conventions applied throughout this work, are introduced here. If the reader is familiar with classical electromagnetic theory, this chapter can be omitted.
3.1 Maxwell’s Equations

Maxwell’s equations were established by James Clerk Maxwell in 1873. Together with the Lorentz force law, they serve as the foundation of classical electrodynamics. In differential, microscopic form they read

\[ \nabla \cdot \mathbf{E}(\mathbf{r}, t) = \frac{\rho(\mathbf{r}, t)}{\epsilon_0} \]  
\[ \nabla \cdot \mathbf{B}(\mathbf{r}, t) = 0 \]  
\[ \nabla \times \mathbf{E}(\mathbf{r}, t) = -\frac{\partial \mathbf{B}(\mathbf{r}, t)}{\partial t} \]  
\[ \nabla \times \mathbf{B}(\mathbf{r}, t) = \mu_0 \left( \mathbf{J}(\mathbf{r}, t) + \epsilon_0 \frac{\partial \mathbf{E}(\mathbf{r}, t)}{\partial t} \right), \]

where \( \mathbf{E} \) and \( \mathbf{B} \) are the electric and magnetic fields, \( \rho \) and \( \mathbf{J} \) the charge and current densities, while \( \epsilon_0 \) and \( \mu_0 \) denote the vacuum permittivity and permeability. To describe the macroscopic response of a material, it is helpful define the auxiliary fields

\[ \mathbf{D}(\mathbf{r}, t) = \epsilon_0 \mathbf{E}(\mathbf{r}, t) + \mathbf{P}(\mathbf{r}, t) \]  
\[ \mathbf{H}(\mathbf{r}, t) = \frac{1}{\mu_0} \mathbf{B}(\mathbf{r}, t) - \mathbf{M}(\mathbf{r}, t), \]

where \( \mathbf{P} \) is the polarization field and \( \mathbf{M} \) the magnetization field. The bound charge/current densities are related to the auxiliary fields as

\[ \rho_{\text{bnd}} = -\nabla \cdot \mathbf{P}(\mathbf{r}, t) \]  
\[ \mathbf{J}_{\text{bnd}} = \nabla \times \mathbf{M}(\mathbf{r}, t) + \frac{\partial \mathbf{P}(\mathbf{r}, t)}{\partial t} = \mathbf{J}_{\text{mag}}(\mathbf{r}, t) + \mathbf{J}_{\text{pol}}(\mathbf{r}, t). \]

The total charge/current density is the sum of the bound part and any free contributions (see [41] for additional discussion). With these definitions, Maxwell’s equations can be cast into their macroscopic form

\[ \nabla \cdot \mathbf{D}(\mathbf{r}, t) = \rho_{\text{free}}(\mathbf{r}, t) \]  
\[ \nabla \cdot \mathbf{B}(\mathbf{r}, t) = 0 \]  
\[ \nabla \times \mathbf{E}(\mathbf{r}, t) = -\frac{\partial \mathbf{B}(\mathbf{r}, t)}{\partial t} \]  
\[ \nabla \times \mathbf{H}(\mathbf{r}, t) = \mathbf{J}_{\text{free}}(\mathbf{r}, t) + \frac{\partial \mathbf{D}(\mathbf{r}, t)}{\partial t}. \]
Before the macroscopic Maxwell equations can be applied, a description of the material response is required. The equations specifying the response are called constitutive relations. If the response is linear, the constitutive relations are

\[
D = \varepsilon E, \quad H = \frac{1}{\mu} B,
\]

(3.13)

where \(\varepsilon\) and \(\mu\) are the permittivity and permeability of the material. In optics it is common to work with the refractive index \(n\) rather than \(\varepsilon\) and \(\mu\). Their relation is

\[
\varepsilon \mu = (\varepsilon_0 \mu_0) \varepsilon_r \mu_r = \left(\frac{n}{c_0}\right)^2,
\]

(3.14)

where \(\varepsilon_r\) and \(\mu_r\) denote the relative permittivity and permeability, while \(c_0\) is the speed of light in vacuum. In reality, the constitutive relations are not linear (except in vacuum). However, for the displacement field, the approximation remains valid in all but the most extreme cases (e.g. pulsed lasers). For the magnetizing field on the other hand, the linear approximation can break down for even common materials like iron.

### 3.1.1 Time harmonic fields

A time-harmonic, monochromatic field can be expressed in the form\(^1\)

\[
V(r, t) = \text{Re}\{V(r)e^{-i\omega t}\}.
\]

(3.15)

Note that while the spatial amplitude \(V(r)\) is complex, \(V(r, t)\) is real. Inserting the time harmonic ansatz, for the \(E, D, B\) and \(H\) fields into equations (3.9) to (3.12), their time harmonic form is obtained

\[
\nabla \cdot D(r) = \rho_{\text{free}}(r)
\]

(3.16)

\[
\nabla \cdot B(r) = 0
\]

(3.17)

\[
\nabla \times E(r) = i\omega B(r)
\]

(3.18)

\[
\nabla \times H(r) = J_{\text{free}}(r) - i\omega D(r).
\]

(3.19)

\(^1\)Engineers tend to use the opposite sign convention in the exponential function.
3.2 The wave equation

Maxwell’s equations in the form (3.16) to (3.19) are a set of coupled, first-order differential equations. They can be decoupled by taking the curl of equation (3.18) and inserting equation (3.19),

\[ \nabla \times \left( \frac{1}{\mu} \nabla \times \mathbf{E} \right) = i\omega \left( \nabla \times \mathbf{H} \right) = i\omega \left( \mathbf{J}_{\text{free}} - i\omega \mathbf{D} \right) \]  

(3.20)

Assuming that no external currents are present so that \( \mathbf{J}_{\text{free}} \) is solely due to a material response following Ohms law, equation (3.20) can be rewritten as

\[ \nabla \times \left( \frac{1}{\mu} \nabla \times \mathbf{E} \right) = i\omega \left( \sigma - i\omega \epsilon \right) \mathbf{E} = \left( \epsilon + i\frac{\sigma}{\omega} \right) \omega^2 \mathbf{E}. \]  

(3.21)

Identifying the content of the last bracket as a complex permittivity,

\[ \epsilon \leftrightarrow \epsilon + i\frac{\sigma}{\omega}, \]  

(3.22)

so as to take into account the combined effect of bound charges and conduction current [42], equation (3.21) is reduced to

\[ \nabla \times \left( \frac{1}{\mu} \nabla \times \mathbf{E} \right) - \epsilon \omega^2 \mathbf{E} = 0. \]  

(3.23)

This is the wave equation for the electric field. It can be cast into a slightly simpler form by applying the vector identity

\[ \nabla \times (\nabla \times \mathbf{V}) = \nabla (\nabla \cdot \mathbf{V}) - \nabla^2 \mathbf{V}. \]  

(3.24)

As no external charges are present, the divergence of \( \mathbf{D} \) vanishes. If \( \epsilon \) is independent of position \( \nabla \mathbf{E} \) is zero, and equation (3.23) is reduced to

\[ \nabla^2 \mathbf{E} - \epsilon k_0^2 \mathbf{E} = 0 \]  

(3.25)

with \( k_0 = \omega/c_0 \). The wave equation will be the governing differential equation for the formulation of all boundary value problems in this work. Solutions for these problems are obtained using the finite-element method (FEM), which is presented in section 4.2. It should be noted that for the problems treated in this work, \( \epsilon \) is generally not independent of position. However, it is piecewise constant (it changes only at material interfaces), which is good enough, since in the FEM formulation boundary conditions are enforced explicitly at the boundaries.
3.3 Boundary conditions

Boundary conditions for the interfaces between different media can be derived directly from Maxwell’s equations. The standard approach is to convert them into integral form by applying the divergence theorem and Stokes’s theorem, after which they are applied to an infinitesimal Gaussian pillbox at the boundary surface between the two media. A derivation can be found in Jackson [43]. For later reference, the result reads

\[(D_2 - D_1) \cdot \hat{n} = \sigma\]  
\[(B_2 - B_1) \cdot \hat{n} = 0\]  
\[\hat{n} \times (E_2 - E_1) = 0\]  
\[\hat{n} \times (H_2 - H_1) = K\]

where \(\hat{n}\) is the surface normal vector, \(\sigma\) and \(K\) denote the surface charge and current density, and the indices 1, 2 indicate the two different materials.

3.3.1 Assumptions

In this work, all media will be assumed isotropic, non-magnetic and the displacement field will be assumed linear. Hence the relevant constitutive relations are

\[D = \varepsilon E, \quad H = \frac{1}{\mu_0} B,\]

with \(\varepsilon\) being a (complex) scalar\(^2\). If the permittivity is complex, the refractive index will be complex as well. Since the non-magnetic assumption implies \(\mu_r = 1\), the relation between permittivity and refractive index, equation (3.14), simply becomes

\[n = \sqrt{\varepsilon_r} = n' + i n''.\]

The imaginary part, \(n''\), is called the extinction coefficient. It is assumed that no free charges/currents are present except for those accounted for by the complex permittivity. Also, all fields are assumed time harmonic, and complex notation is used.

\(^2\)For anisotropic materials \(\varepsilon\) (and possibly \(\mu\)) is a second rank tensor.
3.4 Polarization conventions

In most problems considered in this work, the excitation is expressed in terms of a plane wave,

$$E(r) = E_0 e^{i k \cdot r}, \quad (3.32)$$

incident upon a sample. The sample plane is chosen as the xy-plane so that a wave propagating along the z-axis will be normally incident. The wave vector is chosen as

$$k_f = n (\hat{z} \cos \theta + \hat{x} \sin \theta) k_0, \quad (3.33)$$

$$k_b = n (-\hat{z} \cos \theta + \hat{x} \sin \theta) k_0, \quad (3.34)$$

where $\theta$ is the polar angle with respect to the sample surface normal $\hat{n}$ and the indices $f/b$ indicate forward/backward propagation. With these definitions, the plane of incidence, spanned by $\hat{n}$ and $k$, will be the $xz$-plane. An arbitrary polarization state can be decomposed into components with the $E$-field perpendicular (s-polarization) and parallel (p-polarization) to the plane of incidence as illustrated in figure 3.1. In the case of s-polarization, the $E$-field will be in the $y$-direction. The $H$-field

![Figure 3.1: Plane wave incident on an interface between two different media. Directions of the $E$ (blue) and $H$ (red) fields are indicated for (a) s- and (b) p-polarization. The amplitudes of the incident $E_0$, reflected $E_r$ and transmitted $E_t$ waves are related through the complex Fresnel coefficients (see section 3.5).](image-url)
3.4 Polarization conventions

can be derived from the \( \mathbf{E} \)-field using equation (3.18),

\[
\mathbf{H} = \frac{1}{i \omega \mu_0} \left( \nabla \times \mathbf{E} \right) = \frac{E}{\omega \mu_0} \left( -\hat{x}k_z + \hat{z}k_x \right) = \frac{En}{c_0\mu_0} \left( \mp \hat{x} \cos \theta + \hat{z} \sin \theta \right).
\]

(3.35)

For p-polarization, the \( \mathbf{H} \)-field is in the \( y \)-direction. Equation (3.19) allows the evaluation of the \( \mathbf{E} \)-field from the \( \mathbf{H} \)-field,

\[
\mathbf{E} = \frac{-1}{i \omega \varepsilon} \left( \nabla \times \mathbf{H} \right) = \frac{-H}{\omega \varepsilon_0 n^2} \left( -\hat{x}k_z + \hat{z}k_x \right) = \frac{-Hc_0\mu_0}{n} \left( \mp \hat{x} \cos \theta + \hat{z} \sin \theta \right).
\]

(3.36)

Writing out the explicit expressions for the \( \mathbf{E} \) and \( \mathbf{H} \)-fields, one must decide whether the sign of the backward traveling wave should be based on the relative phase of (1) the \( \mathbf{E} \)-field or (2) the \( \mathbf{H} \)-field. For s-polarization there is agreement in the literature to choose convention (1),

\[
\mathbf{E} = +E\hat{y}, \quad \mathbf{H} = \frac{nE}{c_0\mu_0} \left( \mp \hat{x} \cos \theta + \hat{z} \sin \theta \right) \quad (s-pol).
\]

(3.37)

For p-polarization on the other hand, the textbooks are split almost evenly [44]. Since in this work the \( \mathbf{E} \)-field is of main interest, convention (1) will be adapted also for p-polarization

\[
\mathbf{H} = \pm H\hat{y}, \quad \mathbf{E} = \frac{Hc_0\mu_0}{n} \left( \hat{x} \cos \theta \mp \hat{z} \sin \theta \right) \quad (p-pol).
\]

(3.38)

When \( n \) is constant (which is true everywhere except at the boundaries), the p-polarized fields can be expressed in terms of the amplitude of the \( \mathbf{E} \)-field as

\[
\mathbf{H} = \pm \frac{En}{c_0\mu_0} \hat{y}, \quad \mathbf{E} = E \left( \hat{x} \cos \theta \mp \hat{z} \sin \theta \right) \quad (p-pol).
\]

(3.39)

Since equation (3.39) is not valid at the boundaries, equation (3.38) must be applied when boundary conditions are evaluated.

3.4.1 Energy flux

The instantaneous power flow is described by the poynting vector \( \mathbf{S} \),

\[
\mathbf{S} = \mathbf{E} \times \mathbf{H}.
\]

(3.40)
In most cases it is not the instantaneous value, but rather a time average, which is of interest. For time harmonic fields,

\[
S(t) = \text{Re} \left[ E e^{-i\omega t} \right] \times \text{Re} \left[ H e^{-i\omega t} \right] = \frac{1}{2} \left( \text{Re} \left[ \mathbf{E} \times \mathbf{H}^* \right] + \text{Re} \left[ \mathbf{E} \times \mathbf{H} e^{-2i\omega t} \right] \right).
\] (3.41)

Calculating the average power flow (per cycle), the fast-oscillating term drops outs,

\[
\langle S \rangle = \frac{1}{T} \int_0^T S(t) dt = \frac{1}{2} \text{Re} \left[ \mathbf{E} \times \mathbf{H}^* \right].
\] (3.42)

Combining equation (3.42) and equations (3.37) and (3.38), explicit expressions for s- and p-polarization can be obtained. Textbooks often list the relation

\[
\langle S \rangle = \frac{n}{2c_0\mu_0} |E_0|^2 \mathbf{n},
\] (3.43)

where \( \mathbf{n} \) is a unit vector indicating the direction of the energy flow. This expression is true for a forward propagating wave in a non-absorbing medium. In the more general case of a superposition of a forward and a backward travelling wave,

\[
\mathbf{E}(\mathbf{r}) = E_f e^{i\mathbf{k}_f \cdot \mathbf{r}} + E_b e^{i\mathbf{k}_b \cdot \mathbf{r}},
\] (3.44)

propagating in absorbing media, the expressions are

\[
\langle S \rangle_s = \frac{1}{2c_0\mu_0} \text{Re} \left[ \mathbf{k} |E_f + E_b|^2 (n \sin \theta)^* + \mathbf{z} (E_f + E_b)^* (n \cos \theta)^* \right]
\] (3.45)

\[
\langle S \rangle_p = \frac{1}{2c_0\mu_0} \text{Re} \left[ \mathbf{k} |E_f - E_b|^2 n^* \sin \theta + \mathbf{z} (E_f + E_b)^* (E_f - E_b)^* n^* \cos \theta \right].
\] (3.46)

### 3.5 The Fresnel equations

The Fresnel equations describe the reflection and transmission of electromagnetic waves at a planar interface as shown in figure 3.1. Assuming that no charge is present at the interface, equations (3.26) and (3.27) imply that the normal component of the \( \mathbf{D} \) and \( \mathbf{B} \) fields are continuos
across an interface. Applying equation (3.30), the conditions can be expressed in \( \mathbf{E} \) and \( \mathbf{H} \) directly,
\[
\epsilon_1 E_1^\perp - \epsilon_2 E_2^\perp = 0, \tag{3.47}
\]
\[
H_1^\perp - H_2^\perp = 0. \tag{3.48}
\]

Similarly, equations (3.28) and (3.29) and the assumption of no surface current imply that the tangential components of the \( \mathbf{E} \) and \( \mathbf{H} \) fields are continuous,
\[
E_1^\parallel - E_2^\parallel = 0, \tag{3.49}
\]
\[
H_1^\parallel - H_2^\parallel = 0. \tag{3.50}
\]
The Fresnel coefficients, \( r \) and \( t \), are obtained by applying the boundary conditions equations (3.47) to (3.50) one by one. The result is
\[
t_s = \frac{2n_i \cos \theta_i}{n_i \cos \theta_i + n_j \cos \theta_j}, \quad r_s = \frac{n_i \cos \theta_i - n_j \cos \theta_j}{n_i \cos \theta_i + n_j \cos \theta_j}, \tag{3.51}
\]
\[
t_p = \frac{2n_i \cos \theta_i}{n_j \cos \theta_i + n_i \cos \theta_j}, \quad r_p = \frac{n_i \cos \theta_j - n_j \cos \theta_i}{n_j \cos \theta_i + n_i \cos \theta_j}. \tag{3.52}
\]
along with Snell’s law
\[
n_i \sin \theta_i = n_j \sin \theta_j. \tag{3.53}
\]

### 3.5.1 Energy reflectance/transmittance

The energy reflectance/transmittance, \( R \) and \( T \), is the ratio between the \( z \)-components of incoming and the reflected/transmitted intensity (or flux). Applying equation (3.45) they are found to be
\[
R = |r|^2, \quad T = |t|^2 \gamma \tag{3.54}
\]
where the polarization dependent factor \( \gamma \) is
\[
\gamma_s = \frac{\text{Re} \left[ (n_2 \cos \theta_2)^* \right]}{\text{Re} \left[ (n_1 \cos \theta_1)^* \right]}, \quad \gamma_p = \frac{\text{Re} \left[ n_2^* \cos \theta_2 \right]}{\text{Re} \left[ n_1^* \cos \theta_1 \right]}. \tag{3.55}
\]
This chapter provides an overview of the main optical modelling techniques employed in this work. Since a transfer-matrix method (TMM) implementation was forged as part of this work, the method is described in detail. On the contrary, a commercial finite-element method (FEM) code was used, so only the basic principles of FEM are covered.
4.1 The transfer-matrix method

The standard TMM treats a monochromatic, coherent, plane wave incident upon stack of planar, homogeneous layers. At any point in the stack, the electric field can be represented as a superposition of a forward and a backward traveling wave,

\[ E(r) = E_f e^{ik_f r} + E_b e^{ik_b r}. \] (4.1)

For notational convenience, the amplitude of the forward/backward propagating wave will be denoted \( v \) and \( w \) respectively. Expressing the amplitudes as a vector, the traversal of the optical system can be expressed as a matrix equation

\[ \begin{pmatrix} v_0 \\ w_0 \end{pmatrix} = S \begin{pmatrix} v \\ w \end{pmatrix}, \] (4.2)

where \( S \) is the system scattering matrix. As a simple example, consider a wave traveling through a layer indexed by \( i \). The forward traveling wave is shifted in phase by

\[ \beta_i = \left( \frac{2\pi d_in_i}{\lambda_0} \right) \cos \theta_i \] (4.3)

where the (complex) angle \( \theta_i \) is calculated from Snell’s law, equation (3.53). Likewise the backward traveling wave is shifted by \(-\beta_i\),

\[ v = v_0 e^{i\beta_i} \] (4.4)
\[ w = w_0 e^{-i\beta_i}. \] (4.5)

Expressed in the form of equation (4.2), the layer propagation matrix is

\[ L_i = L_i(\beta_i) = \begin{pmatrix} \exp(-\beta_i) & 0 \\ 0 & \exp(\beta_i) \end{pmatrix}. \] (4.6)

Next, consider an interface between layers \( i \) and \( j \). By matching the amplitudes at each side of the interface, the relations

\[ v = v_0 t_{i,j} + w r_{j,i} \] (4.7)
\[ w_0 = wt_{j,i} + v_0 r_{i,j} \] (4.8)
are obtained with \( r_{i,j} \) and \( t_{i,j} \) being the (complex) Fresnel coefficients when going from layer \( i \) to layer \( j \). Rearranging terms,

\[
v_0 = \frac{1}{t_{i,j}} \left( v - w r_{j,i} \right) \tag{4.9}
\]

\[
w_0 = \frac{1}{t_{i,j}} \left( w(t_{i,j} t_{j,i} - r_{i,j} r_{j,i}) + v r_{i,j} \right) . \tag{4.10}
\]

Expressed in the form of equation (4.2), the interface matrix is

\[
\mathbb{I}_{i,j} = \frac{1}{t_{i,j}} \begin{pmatrix} 1 & -r_{j,i} \\ r_{i,j} & t_{i,j} t_{j,i} - r_{i,j} r_{j,i} \end{pmatrix} . \tag{4.11}
\]

The scattering matrix for an arbitrary layer stack is constructed by matrix multiplication of the element matrices. For a stack of \( m \) layers

\[
S = \mathbb{I}_{0,1} \mathbb{I}_{1,2} \ldots \mathbb{I}_{m-1,m} . \tag{4.12}
\]

### 4.1.1 Reflection/transmission coefficients

The (complex) reflection/transmission coefficients of the complete stack can be derived directly from the system scattering matrix. First, consider light incident from the left. In this case, the wave traveling backward on the right side of the structure \( (w) \) must be zero. Dividing by the incident field \( v_0 \), the matrix equation reads

\[
\begin{pmatrix} 1 \\ r \end{pmatrix} = S \begin{pmatrix} t \\ 0 \end{pmatrix}, \tag{4.13}
\]

where \( r/t \) denote the reflectance/transmittance coefficients defined as

\[
r = \frac{w_0}{v_0} = \frac{S_{21}}{S_{11}}, \quad t = \frac{v}{v_0} = \frac{1}{S_{11}} . \tag{4.14}
\]

Next, consider light incident from the right. In this case the wave traveling forward on the left side of the structure \( (v_0) \) must be zero. Dividing by the incident field \( w \), the matrix equation reads

\[
\begin{pmatrix} 0 \\ r' \end{pmatrix} = S \begin{pmatrix} r' \\ 1 \end{pmatrix}, \tag{4.15}
\]
where \( r'/t' \) denote the back-reflectance/transmittance coefficients defined as

\[
r' = \frac{v}{w} = \frac{S_{12}}{S_{11}}, \quad t' = \frac{w_0}{w} = \frac{\det S}{S_{11}}.
\] (4.16)

Combining equations (4.14) and (4.16), the scattering matrix can be expressed in terms of the reflection/transmission coefficients,

\[
S = \frac{1}{t} \begin{pmatrix} 1 & -r' \\ r & tt' - rr' \end{pmatrix}.
\] (4.17)

Note the similarity with equation (4.11). It implies that the scattering matrix can be interpreted as an effective interface matrix for the complete structure.

### 4.1.2 Energy flux

The power flow due to a superposition of a forward and a backward travelling was calculated in section 3.4.1. Since the structure is assumed infinite in the sample plane, the power flow (or energy flux) of interest is only that in the \( z \)-direction. From equation (3.45), the expressions are found to be

\[
\langle S \rangle_s \mathbf{\hat{z}} \propto \Re \left[ \left( E_f + E_b \right) \left( E_f - E_b \right)^* \left( n \cos \theta \right)^* \right],
\] (4.18)

\[
\langle S \rangle_p \mathbf{\hat{z}} \propto \Re \left[ \left( E_f + E_b \right) \left( E_f - E_b \right)^* \tilde{n}^* \cos \theta \right].
\] (4.19)

### 4.1.3 Incoherence

In the standard TMM, the light is assumed to be perfectly coherent. While this can be a reasonable assumption for some excitation/structure combinations, the assumption is not generally true. If the layer thickness is large compared to the coherence length of the light, the assumption breaks down. Even if this is not the case, imperfect interfaces and/or variations in layer thicknesses can cause loss of coherence. A number of different approaches have been proposed to extend TMM to deal with
complete[45] and partial[46–49] incoherence. All of the methods presented in the references are implemented in the \textit{tmmpy} package[50] developed as part of this work (see section 5.2). However, details are omitted here as all of the results presented in this work are based on standard TMM calculations.

4.2 The finite-element method

The finite-element method (FEM) is a numerical technique for obtaining approximate solutions to boundary value problems. It has been around since the 1940s, and today the method is applied in a variety of fields including structural analysis, fluid dynamics and electromagnetics. Many great books are available on the subject, both legible [51] and comprehensive [52]. Here, a brief introduction to the basic principles of FEM is given and the differential equations and boundary conditions used in this work are presented.

Modern FEM implementations are based on either the Ritz variational method or the Galerkin method (see appendix A). Both are classical methods for solving boundary value problems. The basic idea is to expand the unknown solution in some basis after which the boundary value problem is converted into a linear system of equations for the (unknown) expansion coefficients. A very important step in the process is the choice of basis functions. For most one dimensional problems it is possible to find appropriate basis functions, but for problems in two and three dimensions such functions might not be readily available. The core idea of FEM is to overcome this obstacle by splitting the domain into smaller parts called \textit{elements}. Instead of using complicated basis functions defined on the complete domain \(\Omega\), simple basis functions (typically low order polynomials) defined on each element \(\Omega_n \in \Omega\) are used. Applying the Galerkin or the Ritz method for each element and summing over all elements, a linear system of equations is obtained. The process of constructing this (linear) system of equations, including the enforcement of boundary conditions, is referred to as \textit{assembly}. The FEM can thus be summarized in three main steps

- \textbf{Discretization} of the domain into elements
4.2 The finite-element method

- **Assembly** of the linear system of equations

- **Solution** of the system

Since an actual FEM implementation has not been carried out as part of this work, details on the assembly process is beyond the scope of this text. Instead, the reader is referred to [52]. The discretization and solution steps on the other hand are performance critical and must be hand tuned. A few comments on each of these steps are provided in the following paragraphs.

### 4.2.1 Discretization

The discretization (or meshing) step is of uttermost importance. It affects the required memory, the computation time and the accuracy of the solution. In general terms, a higher mesh resolution improves the accuracy of the solution at the cost of increased memory consumption and computation time. However, the *mesh quality* is just as important. A higher mesh quality implies obtaining a higher accuracy without increasing the number of mesh elements. The generation of high quality meshes is a science of its own, but typical tricks include (1) enforcing physical boundaries in the mesh construction and (2) increasing the mesh density in regions where gradients are expected to be high.

For the simulation of electric fields, the so-called Nédélec elements are preferred. They enforce tangential continuity across element interfaces, i.e. equation (3.28) is enforced automatically. In addition, as the mesh should always be fine enough to resolve the field oscillation, at least 6 elements per wavelength should be present. In this work the COMSOL mesh generator was used to generate tetrahedral meshes of second order Nédélec elements. Since the wavelength depends linearly on the refractive index $n$, the mesh resolution was scaled accordingly as illustrated in figure 4.1a. Additionally, as illustrated in figure 4.1b, the mesh resolution was increased along the boundary of metal nanoparticles, as a high field gradient is expected in this region.
Figure 4.1: Examples of custom mesh generation. At the interface between air and silicon (a) the mesh resolution changes due to the higher value of $n$ for silicon (blue). Close to a gold sphere (b) a high field gradient is expected, hence the mesh resolution is increased at the sphere surface (blue).

**Solution**

Solving the linear system of equations is the final and typically the most time-consuming step. The FEM system matrix is sparse and often symmetric. The solver should always be chosen to exploit such properties. The process of solving large, sparse, linear systems of equations is a standard linear algebra operation for which efficient solvers are freely available [53]. Two classes exist, direct and iterative solvers. Direct solvers solve the problem in one step, often by LU decomposition. The advantage of direct solvers are their stability, while the trade-off is a high memory consumption. For very large problems, the memory consumption might exceed the available amount of RAM, and one is forced to employ an iterative solver instead. Iterative solvers start from some solution guess and refines it iteratively. The advantage of iterative solvers is reduced memory usage, but it comes at the cost of stability. Not all iterative solvers arrive at the same solution, and often they do not even converge.
4.3 Scattered-field formulation

The wave equation, equation (3.25), was derived in section 3.2. It is the core differential equation for all FEM calculations carried out in this work. Since it implies solving for the total field, it is referred to as the full-field formulation.

A slightly different approach designed specifically for scattering problems is the scattered-field formulation. To illustrate the basic idea, consider a structure with some scattering element(s). In the absence of the scattering element(s), the structure is characterized by some permittivity function $\epsilon_B$. The background field $E_B$ is then defined as the solution to the wave equation

$$\nabla \times \left( \frac{1}{\mu_0} \nabla \times E_B \right) - \epsilon_B \omega^2 E_B = 0.$$ (4.20)

Applying the superposition principle, the full field $E$ can be written as

$$E = E_B + E_S$$ (4.21)

where $E_S$ is the scattered field due to the presence of the scattering element(s). Inserting equation (4.21) into equation (3.25) and solving for the scattered field, the scattered-field formulation is recovered

$$\nabla \times \left( \frac{1}{\mu_0} \nabla \times E_S \right) - \epsilon \omega^2 E_S = -\nabla \times \left( \frac{1}{\mu_0} \nabla \times E_B \right) + \epsilon \omega^2 E_B.$$ (4.22)

In the scattered-field formulation, the field which is solved for is not the total field $E$, but the scattered field $E_S$. Contrary to the full-field formulation, the source term (the right hand side) is not zero. To clarify the meaning of the source term, equation (4.20) is substituted into equation (4.22),

$$\nabla \times \left( \frac{1}{\mu_0} \nabla \times E_S \right) - \epsilon \omega^2 E_S = (\epsilon - \epsilon_B) \omega^2 E_B.$$ (4.23)

From equation (4.23), the background field can be interpreted as a source of excitation in regions where $E_B \neq 0$ and $\epsilon \neq \epsilon_B$. 


4.3.1 Boundary conditions

To complete the formulation of a boundary value problem, boundary conditions (BCs) must be supplied in addition to the differential equation. The basic principles of the BCs used in this work are outlined in the following paragraphs.

**Perfect conductor**

For a perfect electric conductor (PEC) the internal electric field is zero, while for a perfect magnetic conductor (PMC) the internal magnetic field is zero. In this case, equations (3.26) to (3.29) are reduced to

\[
\begin{align*}
D \cdot \hat{n} &= \sigma \quad \text{(PEC)} \\
E \times \hat{n} &= 0 \quad \text{(PEC)} \\
B \cdot \hat{n} &= 0 \quad \text{(PMC)} \\
H \times \hat{n} &= K \quad \text{(PMC)},
\end{align*}
\]

where \(\sigma/K\) is the induced charge/current on the surface. The PEC BC can be used to approximate metallic objects, the advantage being to avoid meshing the internals of the object. Additionally, the PEC/PMC BCs can be used to represent symmetry planes for the electric/magnetic field.

**Scattering boundary condition**

In scattering problems, it is often of interest to simulate objects embedded in infinite space. Since infinite space cannot be simulated explicitly, open boundaries are needed. One approach to obtain such behavior is the scattering boundary condition (SBC). As an example, consider a plane wave incident upon a planar boundary in the \(xy\)-plane. Since the normal derivative

\[
\frac{\partial}{\partial z} E(r) = ik_z E(r)
\]

is known at all points on the boundary, an exact BC can be applied. However, the unknown scattered field is rarely a plane wave propagating in a particular direction. Assuming normal incidence, an approximate condition is obtained

\[
\frac{\partial}{\partial z} E(r) \approx i n k_0 E(r).
\]
This is the first order SBC for planar boundaries. It has zero reflection for normally incident plane waves, but the reflection is non-zero for oblique incidence. At grazing incidence, all light is reflected. Higher order SBCs and SBCs for other types of surfaces can be constructed by requiring the field to (approximately) satisfy the Sommerfeld radiation condition \[54\]. Details on such derivations are found e.g. in \[52\]. In the models considered in this work, only the first order SBC for planar surfaces has been used. While a second order SBC is available in COMSOL, the first order implementation was found to provide better stability.

**Perfectly matched layer**

Like the SBC, the perfectly matched layer (PML) is a construction designed to emulate an open boundary. However, the PML is not an actual BC, but rather a fictitious absorber with a perfectly matched interface. Mathematically, the PML can be interpreted as a coordinate stretching in complex space \[55\], physically as an anisotropic absorber \[56\]. Upon entering the PML, the wave is attenuated in the direction normal to the PML surface implying that reflections are avoided (provided that the PML is thick enough). In continuum space, the PML is thus truly reflectionless, but unfortunately this property does not carry over to discretized space \[52\]. Since numerical methods (like FEM) work in discretized space, PMLs will not be perfectly reflectionless. However, the combination of a PML and a SBC will result in low reflections across a wide range of angles. The trade-off compared to just using a SBC is that the PML must be meshed, i.e. the domain size is increased. In the simulations in this work a PML thickness of \(\lambda/2\) was found to be sufficient.

### 4.3.2 Probes

The immediate result of a FEM calculation is the field distribution across the entire geometry. This result is typically too large to save directly, so instead the quantities of interest are "measured" using probes. Examples of probes are reflection/transmission coefficients

\[
R = 1 - \frac{1}{P_0} \int_{\Omega_{\text{inc}}} (\langle S \rangle \cdot \hat{n}) \, dA, \quad T = \frac{1}{P_0} \int_{\Omega_{\text{out}}} (\langle S \rangle \cdot \hat{n}) \, dA
\]  

\[(4.30)\]
where $\Omega_{\text{inc}}/\Omega_{\text{out}}$ denote incoming/outgoing surfaces to which $\hat{n}$ is a normal vector, $\langle S \rangle$ is the time-averaged poynting vector and $P_0$ the incident power. For plasmonic particles, the scattering, absorption and extinction cross sections, $\sigma_{\text{sct}}, \sigma_{\text{abs}}$ and $\sigma_{\text{ext}}$, are typically probed,

\[
\sigma_{\text{sct}} = \frac{1}{I_0} \int_{\delta V} \langle S_{\text{sct}} \rangle \cdot \hat{n} \, dA, \\
\sigma_{\text{abs}} = \frac{1}{I_0} \int_{\delta V} \langle S \rangle \cdot \hat{n} \, dA, \\
\sigma_{\text{ext}} = \sigma_{\text{sct}} + \sigma_{\text{abs}}.
\]

Here $\delta V$ denote the surface of the particle, while $\langle S \rangle$ and $\langle S_{\text{sct}} \rangle$ denote the time-averaged poynting vector calculated for the full and the scattered field, respectively.

### 4.4 Decay rates

The evaluation of electric-field distributions for arbitrary excitation sources in inhomogeneous environments is possible via FEM. Using equation (2.19), the Green dyadic function can thus be reconstructed from the solutions of three separate FEM problems with a dipole source (with $\mathbf{p}$ along to the $x$-, $y$- and $z$-axis respectively) located at $\mathbf{r}_0$. Subsequently, the change in decay rates can be evaluated as per equations (2.20) and (2.22).

It can be shown\cite{36} that the change in decay rate of a two-level quantum system at $\mathbf{r}_0$ is proportional to the change in power $P$ emitted by a dipole source located at $\mathbf{r}_0$. The power can be evaluated via the integral

\[
P = \int_{\Omega} \langle S \rangle \cdot \mathbf{n} \, d\Omega,
\]

where $\langle S \rangle$ denotes the time-averaged poynting vector, $\mathbf{n}$ the unit normal vector to the surface element $d\Omega$, and $\Omega$ the integration surface. Choosing a small surface encompassing just the dipole, $\Omega_2$ in figure 4.2, the total power emitted can be found, and thus the total decay rate, $\gamma$. The radiative part of the decay rate, $\gamma^r$, corresponds to the power that escapes to infinity - hence a large integration surface, encompassing both the nanoparticle (NP) and the dipole emitter, $\Omega_1$ in figure 4.2, should be considered. From $\gamma$ and $\gamma^r$, the non-radiative part of the decay rate, $\gamma^{nr}$,
4.4 Decay rates

can be calculated. Alternatively, it can be evaluated explicitly by considering an integration surface encompassing only the NP, $\Omega_3$ in figure 4.2. In this case, the integral will yield the power dissipated in the NP.

Figure 4.2: Illustration of integration surfaces used in decay rate calculations. The surface $\Omega_1$ (red) encloses both the dipole (black dot) and the nanoparticle. The surface $\Omega_2$ (very small, barely visible) encompass only the dipole. The surface $\Omega_3$ (blue) encompass only the nanoparticle; in this example it is chosen simply as the surface of the nanoparticle.

Compared to point-wise evaluation via equations (2.19), (2.20) and (2.22), evaluation via the integral expression, equation (4.34), is numerically more stable. Therefore, all decay rate evaluations in this work are carried out via equation (4.34). For periodic, lossless media, eigenmode calculations provide a much more efficient way to calculate the change in decay rate(s). As I have only been using this method briefly, and only via an external tool, MIT Photonic bands[38], it will not be accounted for here. Some details are available in Publication II, while a more extended discussion is provided by [35].
Since a large amount of my time has been spent on the development of various codes and software packages, a chapter has been dedicated to document some of this work. Should anyone consider reusing code written as part of this work, this chapter would be a good place to start.
5.1 Code structure and development environment

The code is divided into projects organized in separate folders. Each folder is setup as a repository using the Git version control system and linked to a GitLab server hosted by Aarhus University. To readers unfamiliar with the Git, the following lines outline the fundamental concept (a comprehensive guide takes up a book of its own[57]). The most basic operations in Git are *commit*, *push* and *pull*. Changes to the code are saved together with an (informative) message via the *commit* operation, while the *push* operation propagates the committed changes from the local repository to the server. Similarly, the *pull* operation fetches changes from the server and merges them into the local repository. The complete code history is held indefinitely both locally and remotely, making it possible to compare (or revert) the current code to any previous state, with the navigation eased by the commit messages.

Some projects consist of scripts, e.g. for manipulating or plotting data, while other projects contain reusable functionality. The latter type of project will be denoted as a *package*. Besides the source code itself, a package contains a suite of unit tests, a continuous integration script, a readme file and potentially a number of resources as sketched in figure 5.1. The purpose of the unit tests is to assess the validity of the functionality provided by the package. If the package provided the functionality of a calculator, a simple example of a unit test could be to check that result returned by the *add* method is 4 when invoked with the arguments (2, 2). A strong suite of unit tests helps detecting introduced errors. This is particularly important for larger projects, where a code change might impact functionality which seems unrelated at first glance. The continuous integration script, denoted as `.gitlab-ci.yml` in figure 5.1, contains logic which is executed when code is pushed to the remote server. Typically, the logic will include an invocation of the unit test suite. If any of the unit tests fails, an email is dispatched indicating which tests failed and what commit caused the failure. The automated testing helps detecting and correcting errors at an early stage. To ensure that the unit tests are executed on a well-defined reference system, Docker containers are used[58]. Docker containers are somewhat similar to a virtual machines, but much more light weight. Using customized Docker im-
Figure 5.1: Example of package structure. As most of the packages deve-
dveloped as part in this work are written in Python, .py extensions are
used for the source files in this example.

ages, the invocation of special software such as COMSOL Multiphysics
from within the unit tests is possible.

5.2 Packages

An overview on the main software elements used to obtain results pre-
sented in this work is sketched in figure 5.2. All elements except those
marked as (external) have been developed as part of this work. Unless
marked otherwise in (), the implementation language is Python. The
functionality of each element is outlined in the following sections.

5.2.1 tmmpy

The tmmpy package[50] provides an object-oriented implementation of
the transfer-matrix method (TMM), as outlined in section 4.1. A simple
example of its usage is shown in figure 5.3. The main object is the Stack,
which is created in line 2. In line 3, a Material object is created with
refractive index 1.5 and extinction coefficient 0. In addition to the op-
tion of creating material objects manually, bindings to the refractivein-
dex.info database were implemented. In line 4 a Layer object is created

```
packagename/
    |-- res/
    |    |-- resource.png
    |    |-- resource.csv
    |-- src/
    |    |-- sourcecode.py
    |-- test/
    |    |-- unittests.py
    |-- readme.txt
    |-- .gitlab-ci.yml
```
Figure 5.2: Sketch of main software elements used to obtained results presented in this work. The arrows indicate the direction of data flow.

from the material (of thickness 1000 nm) and added to the stack. Using this syntax, any one-dimensional structure can be created. In line 5, the TMM problem is solved based on the provided excitation conditions ($\lambda = 1500$ nm, s-polarization and normal (0 degrees) incidence). The desired optical response parameters, such as the Fresnel transmission and reflect coefficients, are assigned to the result parameter. Both s- and p-

Figure 5.3: A simple example demonstrating the syntax of tmmpy.

from tmmpy import Stack, Material, Layer, solve_tmm
stack = Stack()
glass = Material(1.5, 0)
stack.append(Layer(glass, 1000))
result = solve_tmm(stack, 1500, 's', 0)

polarization, oblique angle of incidence and complex refractive indices are supported by tmmpy. Interfaces are provided to perform spatially resolved evaluation across the stack, e.g. for evaluation of energy flux and absorption profiles. Furthermore, various extensions of TMM dealing with complete\cite{45} and partial\cite{49, 59} incoherence are included.
**tmmpy_mc** is an extension of *tmmpy* that enables more efficient monte carlo (MC) simulations. In each MC step, the layer thickness is varied. This variation could in principle be carried using the *tmmpy* package, but the layer modification would trigger an (unnecessary) reevaluation of the interface matrices. Omitting this reevaluation enables a vast increase in performance.

**tmmpy_bridge** is an extension of *tmmpy* which generates the necessary configuration files from a *tmmpy* input (excitation conditions and a stack object) to construct the corresponding COMSOL model of the system. The configuration files holds both information of the excitation source, the geometry and the involved materials, as well as the evaluated field across the stack, which are used to setup the background field in COMSOL (see section 4.3).

### 5.2.2 bragg_structure

The *bragg_structure* package was developed during my stay in Freiburg. It contains a Python implementation of the rate-equation model (REM) originally published by Fischer et al. [20] along with Python bindings to the MIT Photonic Bands (MPB) package [38] (the code is prior to the release of an official Python interface) and *tmmpy_mc*. The code is tailored to assess the impact on the upconversion (UC) process, when the UC material is embedded in a finite, one-dimensional photonic crystal.

To achieve a reasonable performance, the REM code was optimized using the *cprofile* tool, and parallel parameter scanning was implemented (Linux only). Furthermore, the separate calculations steps (local energy density, local density of optical states (LDOS), and the solution of the REM itself) were decoupled. The decoupling makes it possible to avoid reevaluation of sub-results that remain unchanged by a parameter change. Changing the excitation intensity, for example, both the LDOS and the energy density data are reused (the energy density changes only by a constant scaling factor). A change in the angle of incidence, on the other hand, triggers a reevaluation of the energy density, while a change in the physical structure (layer thicknesses, material refractive indices) triggers a reevaluation of both the LDOS and the energy density.
5.2 Packages

All results in Publications II and XV were calculated using this code. Additionally, functionality to deal with oblique incidence (for both s- and p-polarization), broadband excitation sources, and fractional LDOS (FLDOS)\cite{60} was implemented. These extensions will be used for the calculations in the next paper(s) by Clarissa Hofmann. One of the main topics will be the angular distribution of the UC emission, which will be measured in experiment, and compared to FLDOS-based calculations.

5.2.3 comsol_interface

All finite-element method (FEM) calculations in this work were carried out using the COMSOL Multiphysics software\cite{61}. COMSOL Multiphysics is developed in Java with a clear separation between the graphical user interface (GUI) and the underlying calculation core. Compiling against the relevant .jar files, the API of the calculation core can be accessed without invoking the GUI. Utilizing this API, the comsol_interface package creates COMSOL models natively via Java code. Setting up the model via code, functions can be defined to specify commonly reused functionality. Examples of functionality implemented in the comsol_interface package include the setup of probes, solvers, particle geometries, symmetry operations and boundary conditions. The validity of this functionality is tested continuously via unit tests. Furthermore, the code setup makes it possible to inject information from external files.

At the center of the comsol_interface package is the ModelConfig (abstract) class. It defines a skeleton of functionality and options relevant across all problems considered in this work. Particular families of problems are treated in subclasses, which specify functionality and operations relevant to each family. Via a range of utility classes, the model configuration is translated into a sequence of calls to the COMSOL Multiphysics API, which are compiled to a .class file. Subsequently, the problem is solved by running this file with options injected via command line arguments. As an example, we look at the PolygonSctModel which considers a nanoparticle (NP) resting on a layered stack excited by an incident plane wave. The syntax for running the model (via bash) is shown in figure 5.4. A range of arguments are passed, specifying both general properties of the system and the excitation source, as well as the
Table 5.1: Selected command line arguments of the PolygonSctModel class. The arguments are grouped into general properties (-sym, -bc and -lmb), properties specific to a scattering calculation (-pol, -ang and -tpm), and properties specific to problems involving a nanoparticle (-geom, -ph and -mat).

<table>
<thead>
<tr>
<th>Argument</th>
<th>Values</th>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td>-sym</td>
<td>{x, y, xy}</td>
<td>mirror symmetries of the system</td>
</tr>
<tr>
<td>-bc</td>
<td>{sbc, pml, periodic}</td>
<td>boundary conditions</td>
</tr>
<tr>
<td>-lmb</td>
<td>(R_{&gt;0})</td>
<td>wavelength of the excitation source</td>
</tr>
<tr>
<td>-pol</td>
<td>{s, p}</td>
<td>polarization of the incident field</td>
</tr>
<tr>
<td>-ang</td>
<td>([0, \frac{\pi}{2}])</td>
<td>angle of incidence</td>
</tr>
<tr>
<td>-tpm</td>
<td>file name</td>
<td>\textit{tmmpy_bridge} configuration file</td>
</tr>
<tr>
<td>-geom</td>
<td>file name</td>
<td>particle geometry (coordinates)</td>
</tr>
<tr>
<td>-ph</td>
<td>(R_{&gt;0})</td>
<td>height of nanoparticle</td>
</tr>
<tr>
<td>-mat</td>
<td>material name</td>
<td>nanoparticle material</td>
</tr>
</tbody>
</table>

NP geometry (the extrusion of a two-dimensional polygon).

Figure 5.4: Simplified example of invocation of a precompiled model via bash. The arguments are explained in table 5.1.

```bash
$ comsolbatch PolygonSctModel.class -sym xy -pol s -ang 0 -lmb 1523 -bc sbc -geom disk_dia400.json -ph 50 -mat au_rakic_bb -tpm air_glass
```

5.2.4 jobgen

The jobgen module provides a Python interface for running COMSOL models via the \textit{comsol\_interface} package. A simplified example of a script for doing a scan in wavelength for two polarizations is outlined in figure 5.5. The command line arguments are passed as lists (line 3), which makes it possible to perform (multi-dimensional) parameter scans simply by specifying multiple values in the list(s). When the \textit{run\_config} method is called (line 4), a working directory with a copy of all resources
Figure 5.5: Simplified example of a script for running a (two-dimensional) parameter scan.

```python
import jobgen
model_name, cfg_name = "PolygonSctModel", "example"

cfg = {
    "-sym": ["xy"],
    "-bc": ["sbc"],
    "-lmb": ["{\{format(1) for l in range(500, 2000)}"]},
    "-pol": ["s", "p"],
    "-ang": ["0"],
    "-tpm": ["air_glass"],
    "-ph": ["50"],
    "-geom": ["disk_dia400.json"],
    "-mat": ["au_rakic-bb"]
}

jobgen.run_config(cfg, cfg_name, model_name)
```

(.class files, tmmpy models, particle geometries, etc.) is created. The command line arguments for each parameter combination are written to a separate .json file. Subsequently, the COMSOL model is invoked automatically via a predefined bash script. By using different bash scripts (controlled via a simple switch in the Python interface), the calculations can be run either directly (for local execution and debugging), or submitted to the SLURM job scheduling system (for remote execution on clusters)[62]. In the latter case, all jobs are submitted simultaneously, thus enabling parallel execution.

5.2.5 sampling

The decay rate of a dipole emitter can vary rapidly near metallic surfaces, but is otherwise relatively smooth in space. The sampling module contains the adaptive sampling logic used in Publications IV and V that exploits this insight to enable a faster evaluation of a spatially resolved decay rate. The basic idea is to do an initial evaluation of the decay rate on a coarse grid (of spatial points), and then use the acquired information to determine where the next points should be assigned (see Publication IV for algorithmic details). A simplified example of a script for evaluating the decay rate adaptively across the yz-plane is outlined in figure 5.6. Lines 2-3 are similar to the example script for running a parameter scan.
Figure 5.6: Simplified example of a script for doing an adaptive decay rate evaluation.

```python
from sampling import AdaptiveDelaunay, get_regular_grid, prepare_surfaces, dispatch_surfaces

model_name, cfg_name = "PolygonRateModel", "example"

cfg = ...

smp = AdaptiveDelaunay(get_regular_grid([0, 750, 0, 100]))

srf = {"pmax": 1000, "count": 10, "tag": "yz_plane", "sampling": smp, "transform": lambda points:
    [(0, point[0], point[1] + 1000) for point in points]}

srfs = prepare_surfaces(cfg, cfg_name, model_name, [srf])

dispatch_surfaces(srfs)
```

In line 4, the adaptive sampling object is created with the initial coarse grid passed as a parameter. In the example, the grid is a rectangle from 0/0 to 750/100 in the x/y-direction. Next, in line 5, all relevant information for doing the evaluation across the desired surface is assigned to the `srf` variable. This includes termination criteria, e.g. the maximum number of points to evaluate, `pmax`, the number of points to add in each refinement step, `count`, a `tag` to identify the surface, a `transform` function and possibly a number of symmetry transforms (not shown in the example). The transform function maps the two-dimensional sampling space to real, three-dimensional space. In the example, the mapping is simply \((x, y) \rightarrow (0, x, y + 1000)\) as we are interested in the yz-plane just below the substrate surface, which is located at \(z = 1000\) (the z-axis points downwards). Multiple surfaces can be created, which is useful e.g. for doing three-dimensional calculations in two-dimensional slices. When the `dispatch_surfaces` method is called in line 7, separate calculations for each surface are launched (in the above example, only a single surface is defined) using the `jobgen` method. If the calculations run on a cluster, all jobs are submitted immediately to the SLURM job scheduling system to enable maximum parallel execution. Subsequently, the main process keeps running, monitoring the progress of each calculation by scanning relevant files every 10th second. When all calculations for a surface are completed, the main process carries out the relevant calculations to determine the assignment of the next points, and the resulting calculations
are submitted automatically to the SLURM job scheduling system. The main process exits, when the termination criteria has been reached for all surfaces.

5.2.6 rempy

The *rempy* package is a Python implementation of the REM discussed in section 2.2. A small example script is shown in figure 5.7. The physical parameters are loaded in line 3, in this case the ones used by Hofmann *et al.* [63]. Prior to running the REM, the (relative) energy density and decay rates should be known (e.g. from FEM calculations) across some spatial grid. To avoid obscuring the example code with data loading logic, zero- and one-valued dummy arrays are used instead (line 4-5). Depending on the dimensionality of the physical system, the arrays can be one-, two-, or three-dimensional. In the example, only the decay rate of the 1558 nm transition is considered. Additional transitions are included simply by adding them to the map. The call to the *run_rem* method (line 7) initiates the actual REM calculation(s). If the spatial grid contains more than one point (in the example $3 \times 3 \times 3 = 9$ points), the spatial points will be evaluated in parallel (on Linux only). The output, including the steady-state solution, is written both to the working directory (denoted "example" in the example) and assigned to the *result* variable.

Figure 5.7: Simple example of a script for launching a REM calculation.

```python
import rem_core
import numpy as np
params = rem_core.hofmann2016()
u_rel = np.ones((3, 3, 3))
gamma_r, gamma_nr = np.ones((3, 3, 3)), np.zeros((3, 3, 3))
g_map = {1558: {"gamma_r": gamma_r, "gamma_nr": gamma_nr}}
result = rem_core.run_rem("example", params, u_rel, g_map)
```
5.2.7 pecpy

The pecpy package is a Python implementation of the proximity-effect correction scheme presented in Publication III. A small example demonstrating how to run the dose optimization routine is shown in figure 5.8. The $\alpha$ and $\eta$ parameters (which must be fitted in advance) and the $\beta$ steps are initialized in line 2. The dose stepper object, which determines how the dose is modified in each step, is created in line 3 using selected $\alpha_0$ and $\eta_0$ parameters (see Publication III for details). Next, the solver is initialized, line 4, with termination criteria in terms of tolerances as well as a limit on the number of iterations in each $\beta$-step. The path to the image specifying the intended design is set in line 5. Finally, in line 6, the optimize_dose call launches the optimization. On completion, a dictionary holding the target design, the optimized dose pattern as well as the predicted error is assigned to the result variable.

Figure 5.8: Dose optimization script demonstrating the syntax of pecpy.

```
from pecpy import DoseStepper, LBFGSB, optimize_dose
alpha, eta, betas = 20, 0.1, [2**i for i in range(11)]
stp = DoseStepper(alpha0=10, eta0s=[0.1])
sol = LBFGSB(ftol=1e-8, gtol=1e6, maxiter=100)
img = "path_to_image"
result = optimize_dose(img, alpha, eta, betas, stp, sol)
```

gdsutils is a utility module for mapping between gds files (used by the electron-beam lithography (EBL) system) and greyscale images (used internally by pecpy). In the gds format, a polygon (including potential holes) is represented by a coordinate list, which describes "how to draw the polygon without lifting the pen from the paper". While the mapping from gds to image is achieved simply by drawing the geometrical elements, the inverse mapping involves edge tracing, filtering and careful assembly of the coordinate list(s).
In the first SunTune publication[31], split peaks in the extinction spectrum for certain samples were observed in experiment, which were not predicted by the applied single-particle model. This publication is the result of my efforts trying to come up with a new model that encompass these observations.
6.1 Context

In real life, plasmonic devices are based on ensembles of metallic nanoparticles (NPs) rather than a single NP. Depending on the inter-particle spacing, the arrangement of the NPs, and the surrounding dielectric environment, the ensemble response can differ significantly from the single-particle response due to particle-particle coupling effects. The first SunTune publication[31] considered ensembles of well-separated NPs (disks), randomly distributed across a surface. Under these conditions, particle-particle coupling effects are typically weak[64]. Nonetheless, a particle-density dependent peak-splitting in the extinction cross section spectrum was observed, indicating the presence of non-negligible particle-particle interactions.

While periodic arrays of NPs can be treated efficiently by considering a single unit cell subject to appropriate boundary conditions, ensembles of randomly distributed NPs are more challenging. Applying the finite-element method (FEM), see section 4.2, the available memory on a modern PC limits the domain size to a few micron, which is typically sufficient for a unit cell, but far too small for a complete NP ensemble. In an initial brute-force attempt to explain the observed peak-splitting, a cluster (with a few hundred gigabytes of RAM) was utilized to extend the domain size by a few additional micron. A number of simulations were then carried out considering randomly selected sub-squares of the $100 \times 100 \ \mu m^2$ pattern. As the extinction cross section spectra resulting from the incoherent averaging of a few tens of such simulations were unable to reproduce the observed peak splitting, this approach was abandoned. Instead, attention was directed towards a search for a simpler model system, which incorporates enough of the important physics to predict the observed experimental results. Such a system was identified in terms of two identical particles with a separation equal to the mean particle spacing in the ensemble.

6.2 Contribution

My contributions to this work include writing essentially all of the text, carrying out all calculations and creating all figures.
Particle-particle interactions in large, sparse arrays of randomly distributed plasmonic metal nanoparticles: a two-particle model

EMIL H. ERIKSEN,¹,* BRIAN JULSGAARD,¹,² SØREN P. MADSEN,³ HARISH LAKHOTIYA,¹ ADNAN NAZIR,¹,² AND PETER BALLING¹,²

¹Department of Physics and Astronomy, Aarhus University, Ny Munkegade 120, 8000 Aarhus C, Denmark
²Interdisciplinary Nanoscience Center, Aarhus University, Gustav Wieds Vej 14, 8000 Aarhus C, Denmark
³Department of Engineering, Aarhus University, Inge Lehmanns Gade 10, 8000 Aarhus C, Denmark
*emher@au.dk

Abstract: A two-particle model is proposed which enables the assessment of particle-particle interactions in large, sparse arrays of randomly distributed plasmonic metal nanoparticles of arbitrary geometry in inhomogeneous environments. The two-particle model predicts experimentally observed peak splittings in the extinction cross section spectrum for randomly distributed gold nanocones on a TiO2:Er3+ thin film with average center-to-center spacings of 3-5 diameters. The main physical mechanism responsible is found to be interference between the incident field and the far-field component of the single-particle scattered field which is guided along the film.

© 2017 Optical Society of America

OCIS codes: (250.5403) Plasmonics; (290.5850) Scattering, particles; (240.0310) Thin films.

References and links

1. Introduction

Localized surface plasmons (LSPs) are collective excitations of the free electrons in metal nanoparticles (NPs). At resonance, a resulting near-field enhancement enables light concentration well below the diffraction limit [1]. The LSP resonance frequency depends not only on the geometry of the NPs [2], but also on the dielectric environment [3] and the NP material [4]. Manipulation of these parameters allows tuning of the plasmon frequency from ultraviolet to mid-infrared [5, 6]. The unique properties of LSPs make plasmonic devices, based on arrays of metal NPs, relevant in a variety of applications. Examples include light concentrators [1, 6], plasmonic wave guides [7], nano-scale polarization control [8], imaging below the diffraction limit [9], negative-refractive-index materials [10] and nanosensors [11].

A key element in the plasmonic-device design process is optical modeling. If both the NP dimensions and separations are larger than 10 nm, quantum effects can be neglected [12], and the problem is reduced to solving Maxwell’s equations. While the simple case of a spherical NP can be solved analytically using Mie theory [13], the general case of an arbitrary NP geometry can only be solved numerically. For sheets of periodically arranged NPs, an ensemble can be modeled by considering a unit cell domain to which boundary conditions that enforce the appropriate translational symmetry are applied. Non-periodic structures are intrinsically more diverse and can generally only be modeled by considering the full multi-scattering problem. A broad variety of non-periodic sheet-like structures have been studied including random planar composites [14], large disordered clusters [15] and deterministic aperiodic nanostructures [16] with ensemble properties ranging from chaotic to tuneable. For spheres, the multi-scattering problem can be solved analytically using generalized Mie theory [17]. Non-spherical NPs can be treated using the T-matrix method [18], but convergence remains poor for non-spheroids [19]. NPs of arbitrary geometry in inhomogeneous environments can be simulated using the boundary element method (BEM) [20], the finite-difference time-domain method (FDTD) [21], or the frequency-domain finite-element method (FEM) [22]. All of these methods discretize Maxwell’s equations on a mesh, BEM across the involved surfaces and FEM/FDTD throughout the simulation volume. As the volume mesh must be fine enough to resolve the electric-field oscillation, the available
memory on a modern computer limits the domain size for FEM/FDTD to the micrometer scale for visible wavelengths. In the case of BEM, the resulting system of equations is smaller, but typically dense, which causes the memory consumption to increase rapidly with increasing number and/or complexity of the NPs. The memory-imposed restrictions are usually not a problem for a unit cell representation, but it severely limits the applicability of these methods to NP ensembles. While small NP arrays have been simulated using FDTD [23], large arrays cannot be treated with currently available computational resources.

At sufficiently large inter-particle spacings, particle-particle interactions become negligible. In this quasi-single-particle case, the ensemble response can be predicted from single-particle simulations. A such one-particle model was employed by [4] and [6] which consider similarly sized nanodisk ensembles with center-to-center spacings of several diameters. While in the former case a very good fit with experimental data was obtained, important features in the extinction spectrum remained unaccounted for in the latter. In this paper a two-particle model is proposed which relaxes the assumption of no particle-particle interactions in the simplest possible way. While being sufficiently simple in concept to provide physical insight and in computational complexity to run on a modern personal computer, it captures the main deviations in the experiment [6] versus the one-particle model.

2. Methods

The field distribution in the vicinity of metal NPs can be obtained by solving Maxwell’s equations. For non-magnetic, isotropic, linear media in the absence of external charges and currents, the wave equation for a time-harmonic field is

\[ \nabla \times (\nabla \times E) - \tilde{\epsilon} k_0^2 E = 0, \]

where \( k_0 \) is the wave vector in vacuum and \( \tilde{\epsilon} \) is the complex relative permittivity. Applying the superposition principle, the total field \( E \) can be written as the sum of a background term \( E_B \) and a scattered term \( E_S \),

\[ E = E_B + E_S, \]

where \( E_B \) is chosen to satisfy Eq. (1) with \( \tilde{\epsilon} = \tilde{\epsilon}_B \). The scattered-field formulation is obtained by inserting Eq. (2) into Eq. (1) and solving for \( E_S \),

\[ \nabla \times (\nabla \times E_S) - \tilde{\epsilon} k_0^2 E_S = (\tilde{\epsilon} - \tilde{\epsilon}_B) k_0^2 E_B. \]

In this work, the structure of interest is a thin-film stack (background structure) onto which the NPs (scattering elements) are placed. Without the scattering elements present, the electric field across the stack has an analytical form, which can be evaluated using the transfer matrix method [24]. Subsequently Eq. (3) can be solved numerically for the scattered field. The scattered-field calculations were carried out using COMSOL Multiphysics [25]. The extinction cross section was calculated from the field distribution as

\[ \sigma_{\text{ext}} = \sigma_{\text{sc}} + \sigma_{\text{abs}} = \frac{1}{I_0} \int_{\Omega} (\langle S_S \rangle - \langle S \rangle) \, dA, \]

where \( I_0 \) is the incident intensity, \( \sigma_{\text{sc}} \) and \( \sigma_{\text{abs}} \) are the scattering and absorption cross sections, \( \Omega \) is the NP surface while \( \langle S_S \rangle \) and \( \langle S \rangle \) denote the time-averaged poynting vectors of the scattered and the full field, respectively.

In accordance with the samples considered in [6], the example structure is chosen as truncated gold nanocones of 50 nm height placed on a 100 nm TiO\textsubscript{2}:Er\textsuperscript{3+} thin film on a SiO\textsubscript{2} substrate. The geometries for the one- and two-particle models are shown in Fig. 1. To emulate that the NPs are embedded in infinite space, all boundaries are truncated by perfectly matched layers (PMLs).
Fig. 1. Schematic illustration of the $xz$-plane at $y = 0$ for the (a) one- and (b) two-particle models. Origo is placed at the air/film interface at (a) the NP center and (b) midway between the NPs. The bottom NP diameter $D$ and the two-particle center-to-center distance $d$ are indicated. The NPs are modeled as truncated cones with top diameter $0.9D$.

The background field is normal incident along the positive $z$-axis, i.e. $E_B$ lies in the $xy$-plane. Relative to the line connecting the NPs in the two-particle geometry, polarization states with the electric field parallel ($\parallel$) and orthogonal ($\perp$) can be defined. With the NPs located along the $x$-axis the $\parallel$ and $\perp$ polarization states correspond to $x$- and $y$-polarization, respectively.

3. Simulation results

The NP diameter was selected as $D = 270 \text{ nm}$, one of the diameters for which a peak splitting was observed experimentally by [6]. Calculated field distributions for the two-particle geometry for $d = 2D$ are shown in Fig. 2(a) and 2(b). While the near-field component is most intense at the NP edges along the axis of polarization, the far-field component is most pronounced orthogonal to it. For dimer-like two-NP systems with small interparticle spacings ($d < 1.5D$) the near field dominates, and the maximum coupling is obtained for $\parallel$ polarization [26]. However, for $d \geq 2D$ as considered in this work, the far-field prevails and the strongest coupling is observed for $\perp$ polarization. To emphasize the difference between the two models, all calculations in this section are shown for $\perp$ polarization rather than a polarization average.

Figure 2(c) shows calculated extinction spectra for selected values of $d$ along with a single-particle calculation. Alongside, the phase difference between the background field and the scattered field emitted by a single particle, $\Delta \phi$, is plotted in Fig. 2(d) and 2(e). Intuitively, the strongest coupling is expected when the second particle is located at points of in-phase addition [27], i.e., along the $\Delta \phi = 0$ lines. When $d$ is small, a blueshift in the extinction cross section (Fig. 2(c)) is observed, but as $d$ increases the blueshift decreases. At $d = 580$ nm the two-particle peak aligns with the single-particle peak resulting in a peak sharpening. From Fig. 2(d) we see that the position $d = 580$ nm (white shading) agrees well with the first $\Delta \phi = 0$ line. Subsequently, the peak is redshifted and around $d = 900$ nm a new, blueshifted peak is formed while the amplitude of the redshifted peak decreases. At $d = 1100$ nm, the amplitudes of the two peaks become similar in magnitude and they appear as a split peak. Figure 2(e) illustrates that while the $\Delta \phi = 0$ condition is not satisfied at the single-particle resonance, it is at higher/lower $\lambda$-values (intersections of $\Delta \phi = 0$ lines with the black, shaded column) which explains the observed peak splitting. Upon increasing $d$ further, the peak movements repeat continuously. Spectra for the first two/three distances for which a peak sharpening/splitting is observed are included in Fig. 2(c). At large values of $d$, the two-particle model curves converge towards the single-particle curve apart from small wiggles.

4. Comparison to experimental results

To assess the applicability of the two-particle model for random NP arrays, the simulation results are compared with previously published experimental data [6]. The data are grouped in three series denoted S4k, S6k and S8k according to their NP densities of $N = 4000$, 6000, and 8000 NPs per $100 \times 100 \mu m^2$. The NPs where fabricated by electron-beam lithography with the
Fig. 2. (a,b) Simulated field distributions at $z = 0$ for (a) $\parallel$ and (b) $\perp$ polarization, $D = 270$ nm, $d = 2D$ and $\lambda = 1290$ nm (single particle resonance). To allow visualization of the far-field component, the color scale has been truncated at 2.5 even though the local field enhancement in (b) exceeds 20 within 1 nm of the NP edges. (c) Extinction cross section $\sigma_{\text{ext}}$ in units of the geometrical cross section $\sigma_{\text{geo}}$ as a function of $\lambda$ for one- (dashed lines) and two-particle (solid lines) models for different $d$. The curves for which a peak sharpening/splitting is observed are shown in blue/red. (d,e) Simulated phase distributions. Shadings indicate particle positions for which peak sharpening (white) and splitting (black) is observed. Shown are $\Delta \phi$ in (d) the $xy$-plane at the NP center ($z = -25$ nm) for $\lambda = 1290$ nm and (e) along the line $(y, z) = (0, -25 \text{ nm})$ for different values of $\lambda$. The two plots (d,e) coincide along the dashed, white line.

pseudo-random distributions constructed using a random number generator. The resulting mean distance to the four nearest neighbors, a measure analog to the nearest-neighbor distance in a square lattice, appears normally distributed with a mean value very close to the characteristic particle spacing $1/\sqrt{N}$ and a standard deviation around 10% of the mean value. For each sample, the extinction cross section was measured as a function of $\lambda$, as exemplified in Fig. 3(a). In all three cases, the main resonance peak is formed by two subpeaks. Together with the experimental data, simulated spectra calculated from the one- and two-particle models are shown. In the two-particle model $d$ was chosen as $1/\sqrt{N}$. As the orientation of the electric field varies randomly for the nearest-neighbor pairs across the sample, polarization averaged calculations were performed.

While the one-particle model reproduces the overall peak location fairly well, the absence of additional features clearly demonstrates it being too simplistic. The two-particle model, on the other hand, is able to predict the main features in the spectrum including the positions of the subpeaks. However, a disagreement in scale remains. A possible explanation could be additional light trapping in the TiO$_2$:Er$^{3+}$ thin film due to presence of metal NPs on the surface. The peak location(s) were estimated by fitting a (double) Gaussian function to the spectra for both experiment and calculation. Applying this procedure for all samples, Fig. 3(b) was obtained. For each data series, the peak location(s) are shown as a function of the inverse of the NP diameter. The one-particle model predicts a single curve independent of particle density, which approaches a straight line for large diameters [4]. Overall it fits the data reasonably well, but the peak splitting is not captured. On contrary, the two-particle model includes the necessary symmetry breaking to reproduce the peak splitting, thus providing a much better fit with the experimental data.

When NPs are deposited directly on a substrate, particle-particle couplings are strongly suppressed due to the lack of phase matching between waves propagating along each side of the interface. As reported by [28–30], the suppression can be lifted by applying an index-matching liquid as superstrate. Two-particle model simulations (not shown) successfully reproduce these
Fig. 3. Comparison of one- (dashed line) and two-particle models (solid lines) with experimental data (symbols, connected by a thin line in (a)). (a) Extinction spectra for a selected diameter in each series, 345 nm for S4k (left), 315 nm for S6k (center) and 270 nm for S8k (right). (b) Fitted peak position(s) as a function of inverse diameter $D$ for the S4k, S6k and S8k series. The selected diameters in (a) are marked by grey, vertical shadings. The inserts show SEM images of a sample from each series.

observations. The stronger couplings observed in this work compared to [4], where similar NP distributions are considered, are due the presence of the TiO$_2$ thin film. As the scattered field is guided along the film, a significant particle-particle coupling is observed even in the absence of an index-matched superstrate. The reproduction of the guiding effect, in terms of a simulated non-negligible particle-particle coupling, demonstrates the applicability of the two-particle model to NP ensembles in inhomogeneous environments that can otherwise be challenging to simulate.

5. Summary and discussion

Particle-particle interactions have been investigated in a two-particle system for center-to-center spacings larger than two diameters. It was found that the coupling is dominated by the far field, and that it depends strongly on the dielectric environment. The applicability of a two-particle model for investigating particle-particle couplings in large, sparse arrays of randomly distributed nanoparticles in inhomogeneous dielectric environments was assessed by comparison to experimental data. While the agreement in absolute extinction cross section is not perfect, the two-particle model succeeds in predicting the main spectral features including the observed peak splitting. At the same time, the computational complexity remains low enough that calculations can be carried out on a personal computer.

By tuning the inter-particle spacing $d$, the particle-particle interactions can be exploited to selectively sharpen (or broaden) the plasmonic resonance of the ensemble. The effect is analogous to diffraction in periodic arrays, but weaker and less sensitive to incidence angle. Randomly distributed nanoparticles can be manufactured inexpensively on industrial scale by self-assembly processes, while periodic arrays typically require expensive, low-throughput processing. The combination of cost-effective assembly and low angular sensitivity make plasmonic devices based on random arrangements of nanoparticles favorable for many real life applications, e.g. solar cells. The two-particle model provides a valuable tool for tailoring the particle-particle interactions in such devices.

Funding

Innovation Fund Denmark under the SunTune project.
This publication is the fruit of my work during my stay abroad at the Fraunhofer Institute for Solar Energy Systems (ISE) in Freiburg, where I was working with Clarissa Hofmann in the Novel Solar Cell Concepts group under the supervision of Jan Christoph Goldschmidt.
7.1 Context

Similar to the SunTune project, the Novel Solar Cell Concepts group works with upconversion (UC) in erbium-based compounds. However, rather than looking at plasmonic structures for improving the UC efficiency, their focus is on photonic crystals (see section 2.3.2). During my stay, I worked with Clarissa on her project – improving upconversion efficiency using a particular type of one-dimensional photonic crystal, the so-called Bragg structure. The Bragg structure is essentially a quarter-wave stack, which has been modified slightly to enable better incoupling of broad band excitation sources[63]. It consists of alternating layers of the UC material and a (transparent) high refractive index spacer material. While Clarissa does both theoretical and experimental work, my efforts were focused on the development of a new simulation framework. By coupling my previously developed transfer-matrix method (TMM) code, see section 4.1, with the MIT photonic bands software[38] and the rate equation modelling logic by Fischer[20], a coherent and efficient setup was implemented. Utilizing the new setup, I studied the effect on UC of various system parameters, including the refractive index of the spacer material, excitation conditions, and the number of layers in the stack. Furthermore, I considered the influence of realistic production tolerances via monte carlo (MC) simulations.

7.2 Contribution

My contributions to this work include writing most of the text, carrying out all calculations and creating all figures.
Enhanced upconversion in one-dimensional photonic crystals: a simulation-based assessment within realistic material and fabrication constraints

CLARISSA L. M. HOFMANN,1,2,6,7 EMIL H. ERIKSEN,3,6,8 STEFAN FISCHER,4 BRYCE S. RICHARDS,2,5 PETER BALLING,3 AND JAN CHRISTOPH GOLDSCHMIDT1

1 Fraunhofer Institute for Solar Energy Systems ISE, Heidenhofstrasse 2, 79110 Freiburg, Germany
2 Institute of Microstructure Technology (IMT), Karlsruhe Institute of Technology, Hermann-von-Helmholtz-Platz 1, 76344 Eggenstein-Leopoldshafen, Germany
3 Department of Physics and Astronomy, Aarhus University, Ny Munkegade 120, DK-8000 Aarhus C, Denmark
4 Department of Chemistry, University of California - Berkeley, Berkeley, CA 94720, USA
5 Light Technology Institute (LTI), Karlsruhe Institute of Technology (KIT), Engesserstrasse 13, 76131 Karlsruhe, Germany
6 These authors contributed equally to this work
7 clarissa.hofmann@ise.fraunhofer.de
8 emher@au.dk

Abstract: This paper presents a simulation-based assessment of the potential for improving the upconversion efficiency of β-NaYF4:Er3+ by embedding the upconverter in a one-dimensional photonic crystal. The considered family of structures consists of alternating quarter-wave layers of the upconverter material and a spacer material with a higher refractive index. The two photonic effects of the structures, a modified local energy density and a modified local density of optical states, are considered within a rate-equation-modeling framework, which describes the internal dynamics of the upconversion process. Optimal designs are identified, while taking into account production tolerances via Monte Carlo simulations. To determine the maximum upconversion efficiency across all realistically attainable structures, the refractive index of the spacer material is varied within the range of existing materials. Assuming a production tolerance of σ = 1 nm, the optimized structures enable more than 300-fold upconversion photoluminescence enhancements under one sun and upconversion quantum yields exceeding 15% under 30 suns concentration.

© 2018 Optical Society of America under the terms of the OSA Open Access Publishing Agreement

OCIS codes: (190.7220) Upconversion; (230.5298) Photonic crystals; (230.1480) Bragg reflectors; (160.5690) Rare-earth-doped materials; (240.0310) Thin films; (260.3800) Luminescence; (310.6628) Subwavelength structures, nanostructures; (310.6845) Thin film devices and applications.

References and links


30. E. Verhagen, L. Kuipers, and A. Polman, “Field enhancement in metallic subwavelength aperture arrays probed by...

1. Introduction

Upconversion (UC) is the process of converting two (or more) photons into one photon of higher energy. It is relevant in a broad range of applications, from anti counterfeiting [1, 2], plastic recycling [3], bioimaging [4, 5] and theranostics [5–7] to photovoltaics [8–11]. In photovoltaics, UC enables the utilization of sub-bandgap photons, which for silicon-based solar cells raises the theoretical efficiency limit from ≈ 30% [12] to ≈ 40% [8]. Spectrally suitable materials capable of harvesting near-infrared light have been identified among the trivalent rare-earth materials, in particular the Er$^{3+}$ ion [9, 13–16]. In this study, we consider erbium-doped hexagonal sodium yttrium fluoride ($\beta$-NaYF$_4$:Er$^{3+}$), which ranges among the best performing materials for UC applications in the context of silicon-based photovoltaics [17–22]. However, even for this material system, the UC efficiency remains too low for commercial exploitation [11].

UC is a multi-photon process and thus of non-linear nature [23]. For UC to take place, it requires that at least two photons are absorbed in a small volume within a short period of time. Therefore, in the unsaturated regime, to which non-concentrated solar illumination belongs [24], the UC efficiency can be increased by light concentration. Besides focusing by conventional means, such as lenses, strong local field enhancements can be achieved through plasmonic structures [25–30]. The strong field leads to an increased excitation density in the upconverter and consequently to a higher UC efficiency. A major drawback of plasmonic structures, however, is parasitic absorption, which reduces the energy density of the incident field available for upconversion. Furthermore, additional non-radiative loss channels are introduced, which can dissipate energy from the excited upconverter ions. In the immediate vicinity of plasmonic structures, the non-radiative losses strongly reduce the UC emission, which is particularly problematic as it is typically here the maximum field enhancement occurs [25, 26]. Another way to obtain strong local field enhancements is through dielectric photonic structures [31–41]. Unlike plasmonic structures, photonic structures do not inherently suffer from parasitic absorption, but they do modify the local density of optical states (LDOS). As the LDOS affects the probability of spontaneous emission, a modified LDOS can potentially increase UC emission and/or suppress loss channels, all of which can be tailored through structural engineering [31, 32].

The literature on the application of photonic crystals to improve UC luminescence is dominated
by inverse opal [37, 40–42] and opal structures [36, 38]. For those structures, an experimentally measured UC luminescence enhancements of up to 43 times has been reported [37]. Moreover, a very high UC luminescence enhancement of $10^4$ has been reported for a waveguide structure [43]. Recently, photonic-plasmonic hybrid structures have been experimentally demonstrated, yielding UC luminescence enhancement factors up to three orders of magnitude [44, 45]. Additionally, an experimental study of one-dimensional photonic crystals was carried out by [34], where the crystal was formed from Er$^{3+}$-doped porous silicon with the periodic variation of the refractive index obtained by varying the silicon porosity. In that case, a maximum UC luminescence enhancement of a factor of 26.6 was reported.

In most studies, the design of the photonic structure is not optimized, but simply chosen such that the excitation lies at the photonic band edge [36–38]. In a few reports the actual local field is simulated, based on which an optimal design is realized [33, 34, 43, 44]. However, the effect of the LDOS and the dynamics of the UC process are rarely considered. To address this issue, a rate-equation-modeling framework has been developed as part of our previous work. The model considers the combined effects of the modified LDOS and the modified local energy density [31, 32, 46]. In a theoretical study of the photonic effects on UC, it has been demonstrated that it is indeed important to include the effect of the modified LDOS [32]. Furthermore, the so-called Bragg structure was identified as a promising design for increasing UC efficiency [32, 47]. In contrast to structures that exploit surface effects, the amount of upconverting material in a Bragg structure can be increased by adding more layers in the design. In applications, this is a very important property, in particular for the weakly-absorbing rare-earth materials.

In this work, we present a simulation-based assessment of the potential of the Bragg structure for increasing the UC efficiency of 1523 nm excitation light via the use of $\beta$-NaYF$_4$:Er$^{3+}$ nanophosphors. The Bragg structure consists of alternating quarter-wave layers of two materials with refractive indices $n_{\text{low}}$ and $n_{\text{high}}$, with respect to a design wavelength $\lambda_D$, where $d$ is the layer thickness and $n$ the refractive index, as shown in Fig. 1(a). The outer-most layers have a reduced optical thickness of only $\lambda_D/(8n)$, which enables a more efficient in-coupling of broad-band excitation. In principle, the upconverting material could be embedded in either layer. However, to avoid scattering, which would cause decoherence and thus suppress the desired photonic effects, the upconverting material must be index-matched with respect to the host layer. Since the refractive index of $\beta$-NaYF$_4$:Er$^{3+}$ ($\approx 1.5$ [48]) is relatively low, choosing it as the high index layer would allow only for a very low refractive index contrast. Therefore, we choose the low index layers as the active ones, containing the upconverter material, thus fixing $n_{\text{low}} = 1.5$. The outer-most layers of reduced thickness are assumed passive. The reference structure consists of a single, homogeneous layer containing the same amount of active material as the corresponding Bragg structure.

![Fig. 1. Structural sketches of the Bragg structure (a) and the reference structure (b). The Bragg structure consists of alternating quarter-wave layers, with respect to a design wavelength $\lambda_D$, of an active and a spacer material with refractive indices $n_{\text{low}}$ and $n_{\text{high}}$. The outermost layers have a reduced optical thickness of $\lambda_D/(8n)$ and are assumed passive. The reference structure consists of a single, homogeneous layer containing the same amount of active material as the corresponding Bragg structure.](image-url)
we consider the range $1.5 < n_{\text{high}} \leq 4.0$. Examples of high refractive index materials that remain transparent in the region of the main UC emission include TiO$_2$ and a-Si:H. As the reference, we choose a homogeneous structure with refractive index $n_{\text{low}}$, containing the same amount of active material as the corresponding Bragg structure, see Fig. 1(b).

While ideal photonic crystals are infinite and perfectly periodic, the experimental realizations are finite and suffer from production imperfections. As demonstrated previously [32], the beneficial peak in the energy density enhancement is very narrow in $\lambda_D$, or equivalently the layer thickness, implying a high sensitivity to fabrication tolerances. For production of Bragg-like structures a variety of thin film fabrication methods are available. A fast and efficient production method is spin-coating from solution. With this process, layer uniformities of 1-2% of the total layer thickness can be achieved [49]. Higher layer uniformities can be reached in chemical vapour deposition processes. Using metalorganic vapour phase epitaxy (MOVPE), a layer thickness control of 0.1-0.5% has been reported [50]. Another high-precision production method is atomic layer deposition (ALD), where a thickness accuracy of 0.5% is possible for TiO$_2$ thin films at deposition temperatures above 250 °C [51]. To enable a realistic assessment of the performance of fabricated structures in this work, we take into account finite production tolerances via Monte Carlo simulations. Based on achieved thickness control of 0.5% reported in the literature [50,51], a realistic target tolerance of 0.5% is assumed in the main analysis of this work, which corresponds to an absolute value of $\approx 1$ nm. We choose a value at the limit of what is currently possible, imagining that such accuracy will be enabled in routine production by future advances in nanofabrication techniques. One option could be, for example, to produce the whole layer stack with ALD processes.

We first present the methods for evaluating the local energy density, including the Monte Carlo method, in Section 2.1, after which the LDOS calculations are introduced in Section 2.2. The photonic effects are coupled via a rate equation model (REM), presented in Section 2.3, which yields the upconversion quantum yield (UCQY) and the upconversion photoluminescence (UCPL). The potential for increasing the local energy density, assuming different production accuracies, is investigated in Section 3.1, while the modified LDOS is discussed in Section 3.2. Finally, in Section 3.3, we present the calculated UCQY and UCPL enhancements, scanning through all design parameters. Additionally, different irradiance regimes, targeting different application scenarios, are considered.

2. Methods

2.1. Local energy density

For stacks of planar, homogeneous layers an analytical solution for the electrical field distribution, $E(x)$, exists. It can be found efficiently using the transfer matrix method [52] from which the local energy density of the electric field, $u(x)$, can be calculated as

$$u(x) = \frac{1}{2} \varepsilon(x) |E(x)|^2. \tag{1}$$

We define the relative local energy density as

$$u_{\text{rel}}(\tilde{x}) = \frac{u_{\text{seg}}(\tilde{x})}{u_{\text{ref}}(\tilde{x})}, \tag{2}$$

where for the Bragg structure the $\tilde{x}$ coordinate runs inside the active layers only, as indicated in Fig. 1. For visualization purposes we define also the average relative energy density

$$\bar{u}_{\text{rel}} = \frac{\int u_{\text{seg}}(\tilde{x})d\tilde{x}}{\int u_{\text{ref}}(\tilde{x})d\tilde{x}}. \tag{3}$$
To account for final production tolerances, Monte Carlo simulations are carried out. A number of calculations are performed where for each layer the thickness $d$ is modified as

$$d \rightarrow d + \delta d,$$

(4)

where $\delta d$ is drawn from a Gaussian distribution with a standard deviation $\sigma$ representing the production accuracy. Finally, the energy density is determined from the (incoherent) average across all calculations.

2.2. Local density of optical states

For infinite periodic structures, i.e. ideal photonic crystals, the local density of optical states (LDOS) can be derived from eigenmode calculations \[53–55\]. In this work, we use the MIT Photonic bands \[53\] software package. While the ideal crystal assumption is not accurate for Bragg structures with only a small number of layers \[56,57\], the structures of main interest in this work have 10 active layers or more. Additionally, the eigenmode approach is relevant in a future perspective, as it permits calculation of an angularly resolved LDOS \[58\], which is important for our future work on the modified directionality of upconversion (UC) emission in a Bragg structure.

Due to the scale invariance of the problem, we consider dimensionless quantities

$$k' = \frac{k a}{2\pi}, \quad \omega' = \frac{\omega a n}{2\pi c_0},$$

(5)

where $k$ is the wave vector, $a$ the size of the Wigner-Seitz unit cell and $c_0$ the speed of light in vacuum. For a given eigenvector $k'_j$, an eigenmode calculation yields the mode frequency $\omega'_{b,k'_j}$ and the electric field profile $E_{b,k'_j}(x)$ with $b$ being the band index. Applying the histogramming method and exploiting the in-plane symmetry of the Bragg structure \[58\], the quasi-three-dimensional LDOS can be calculated as

$$\text{LDOS}(x, \omega') = \sum_b \sum_{k',\omega'} |E_{b,k'}(x)|^2 \cdot 2\pi k'_y$$

(6)

with $K_{b,\omega'} = \{k'_j | \omega' \leq \omega_{b,k'_j} \leq \omega' + \Delta \omega'\}$.

where $k'_j$ is sampled on an equidistant grid with spacing $\Delta k'$ in the first quadrant of the $xy$-plane in reciprocal space. While the sampling in the $x$-direction is bound by the edge of the First Brillouin Zone at $k'_x = 0.5$, $k'_y$ is unbound due to the lack of translational symmetry. In this work, all modes across the first seven bands in the range $0 < \omega' < 1.4$ were calculated. To obtain all contributing modes for the relevant emission frequencies and material combinations, the range $0 \leq k'_y \leq 5.4$ was considered. An optional output in MIT Photonic bands is the electric field energy density, from which the squared amplitude of the electric field can be computed,

$$|E_{b,k'(x)}|^2 = \frac{2\mu_{b,k}(x)}{\epsilon(x)},$$

(7)

Subsequently, the LDOS can be calculated according to Eq. (6). Across all calculations presented in this work, discretization steps of $\Delta k' = 10^{-3}$ and $\Delta \omega' = 10^{-3}$ were used. The limited resolution in $k'$-space caused a significant amount of binning noise. To reduce the noise, the LDOS was smoothed along the frequency axis using a Gaussian filter with $\sigma = 5$.

As the reference is homogeneous, the LDOS is independent of the position. Hence, it is equal to the DOS up to a multiplicative constant which depends on the discretization of the Wigner-Seitz unit cell. The DOS for a homogeneous medium has an analytical form \[59\],

$$\text{DOS3D}(\omega') = \frac{4\pi n^3}{\Delta k'^2} \omega'^2.$$

(8)
Applying the same binning procedure as in the numerical calculation for the Bragg structure, the binned reference LDOS can be calculated as

$$LDOS_{\text{ref}}(\omega') = \int_{\omega'}^{\omega'+\Delta\omega'} \text{DOS3D}(\omega'')d\omega'' = \frac{4\pi n^3 \Delta\omega'}{\Delta k'^2} \left( \omega'^2 + \omega' \Delta\omega + \frac{\Delta\omega'^2}{3} \right)$$

with $n = n_{\text{low}}$ [58]. We define the relative LDOS in the Wigner-Seitz unit cell as

$$LDOS_{\text{rel}}(x, \omega') = \frac{LDOS_{\text{brg}}(x, \omega')}{LDOS_{\text{ref}}(\omega')}.$$  

For visualization purposes we define also the average relative LDOS across the active layers of the Bragg structure,

$$\overline{LDOS}_{\text{rel}}(\omega') = \frac{\int LDOS_{\text{brg}}(x, \omega') dx}{\int LDOS_{\text{ref}}(\omega') dx}.$$  

The relative LDOS for a particular transition $i \rightarrow f$ and design, characterized by $\lambda_{fi}$ and $\lambda_D$, respectively, is mapped to the dimensionless transition frequency $\omega_{fi}'$ as [32]

$$\omega_{fi}' = \frac{n_{\text{low}} + n_{\text{high}} \lambda_D}{4n_{\text{low}}n_{\text{high}} \lambda_{fi}}.$$  

### 2.3. Rate equation model

We describe the dynamics of the UC process using a modified version of a rate equation model (REM) originally developed for homogeneous media [46]. All experimental parameters are measured for an excitation wavelength of 1523 nm, for which the main UC emission lies at 984 nm. The first seven energy levels of $\beta$-NaYF$_4$:Er$^{3+}$, shown in Fig. 2, are considered. The energy levels of the $^2H_{11/2}$ and the $^2S_{3/2}$ states are treated as one effective energy level, due to their close proximity. The occupation of each level is described as an element of the occupation density vector.

![Energy Levels Diagram](image-url)

**Fig. 2.** Schematic of the first seven energy levels in $\beta$-NaYF$_4$:Er$^{3+}$ included in the rate equation model, along with most important transitions for the UC process. The highest two energy levels are treated as one due to their close proximity. For the considered excitation of 1523 nm wavelength, the main UC emission lies at 984 nm.
The linear processes included in the REM are ground state absorption (GSA), excited state absorption (ESA), stimulated emission (STE), spontaneous emission (SPE) and multi-phonon relaxation (MPR). Additionally, the non-linear Förster energy transfer processes, energy transfer upconversion (ETU) and cross relaxation (CR), are considered. The most important processes are indicated in Fig. 2. The rate of change of the occupation density vector yields

$$\dot{\mathbf{n}} = [M_{\text{GSA}} + M_{\text{ESA}} + M_{\text{STE}} + M_{\text{SPE}} + M_{\text{MPR}}] \mathbf{n} + \mathbf{v}_{\text{ETU}}(\mathbf{n}) + \mathbf{v}_{\text{CR}}(\mathbf{n}),$$  (13)

where $M$ denotes transition matrices and $\mathbf{v}$ vector functions. Additional details on the REM for the case of a homogeneous medium are available in the original work [46].

The modifications of the REM due to the changes in the photonic environment imposed by the photonic structure can thus be taken into account by scaling the corresponding transition matrices by the relative local energy density

$$M_{\text{GSA}} \rightarrow M_{\text{GSA}} \cdot \text{LDOS}_{\text{rel}}(x, \omega_{f_i}), \quad M_{\text{ESA}} \rightarrow M_{\text{ESA}} \cdot \text{LDOS}_{\text{rel}}(x, \omega_{f_i}), \quad M_{\text{STE}} \rightarrow M_{\text{STE}} \cdot \text{LDOS}_{\text{rel}}(x, \omega_{f_i}).$$  (15)

The spontaneous emission probability $P_{fi}$ is governed by Fermi’s golden rule [60]

$$P_{fi} = \frac{2\pi}{\hbar} |\langle f | H_{\text{int}} | i \rangle|^2 \text{LDOS}(x, \omega_{f_i}) \propto \text{LDOS}(x, \omega_{f_i}),$$  (16)

where $H_{\text{int}}$ is the interaction Hamiltonian between initial, $|i\rangle$, and final, $|f\rangle$, electronic states. To incorporate the effect of the modified LDOS into the REM, the Einstein coefficients for spontaneous emission are scaled as

$$A_{fi} \rightarrow A_{fi} \cdot \text{LDOS}_{\text{rel}}(x, \omega_{f_i}).$$  (17)

It remains a subject of discussion in the literature whether Förster energy transfer processes, which are crucial to the UC process in lanthanide-doped materials, are also influenced by changes in the local photonic environment [61–63]. In this work we follow the arguments of [61,64] and neglect any such effects.

The output of the REM is a steady-state occupation density vector $\mathbf{N}$ from which the main figures of merit, the upconversion photoluminescence (UCPL) and the internal upconversion quantum yield (UCQY), can be calculated. The photoluminescence (PL) for each transition is

$$\text{PL}_{fi} = \int A_{fi}(x) N_{i}(x) dx$$  (18)

where $N_{i}$ is the steady-state occupation density of level $i$. The energy level numbering scheme is shown in Fig. 2. In the UCPL, we consider only the main UC emission of the $3 \rightarrow 1$ transition,

$$\text{UCPL} = \text{PL}_{31}.$$  (19)

This approximation enables a simpler analysis going forward, and the associated error is small, as the $3 \rightarrow 1$ transition accounts for more than 95% of the emitted, upconverted photons. To obtain the UCQY, the UCPL is divided by the number of absorbed photons,

$$\text{UCQY} = \frac{\text{UCPL}}{N_{1} M_{\text{GSA,12}} + N_{2} M_{\text{ESA,24}} + N_{4} M_{\text{ESA,46}}}.$$  (20)
To enable clear visualization of the effect of the Bragg structure, we define also the relative UCPL

$$\text{UCPL}_{\text{rel}} = \frac{\text{UCPL}_{\text{brg}}}{\text{UCPL}_{\text{ref}}}.$$  \hfill (21)

### 3. Results

#### 3.1. Local energy density

The average relative energy density, $\bar{u}_{\text{rel}}$, for an exemplary Bragg structure with $n_{\text{high}} = 2.3$ and a number of active UC layers ($#_{\text{al}}$) of 25 is plotted in Fig. 3(a) for normal incidence and monochromatic excitation at $\lambda_{\text{exc}} = 1523$ nm in dependence on the design wavelength $\lambda_D$. An exemplary value of $n_{\text{high}} = 2.3$ was chosen in accordance with a recent experimental study where the high index layer was made of TiO$_2$ with $n = 2.3$ at 1523 nm [47]. The graph shows that for a certain design wavelength, $\lambda_D^{\text{max}}$, a very large enhancement of $\bar{u}_{\text{rel}}$ occurs. As is known from the literature, a high field enhancement occurs at the photonic band edge [34, 65]. In Fig. 3(b), the reflectance for $\lambda_D = \lambda_D^{\text{max}}$ shows the position of the photonic band gap (PBG) relative to the excitation wavelength $\lambda_{\text{exc}}$. As expected, $\lambda_{\text{exc}}$ lies close to the photonic band edge. In addition to the Bragg structure, the reflectance is also shown for a pure quarter-wave stack. Comparing the two curves, it is clear that the reduced thickness of the outer layers of the Bragg structure significantly affects the performance of the UCPL.
structure causes a suppression of the side lobes of the reflectance peak. While this feature is not necessary for a perfect simulated structure under monochromatic excitation, it enables a more efficient in-coupling in experiments, in particular for broad-band excitation sources. The spatial dependence of \( u \) within the structure is illustrated in Fig. 3(c) for \( \lambda_D = \lambda_D^{\text{max}} \), along with the refractive index profile. A strong increase in \( u \) is observed inside the active layers of the structure. The enhancement can be explained in the context of photonic crystals as slow light piling up \[66\] or in a more classical context as the formation of a standing wave due to interference between the forward and backward propagating waves. It should also be noted that practically all energy density is located in the active layers.

Figure 3(d) shows \( \bar{u}_{\text{rel}} \) for the range of considered structures with \( 1.5 < n_{\text{high}} \leq 4.0 \) and \( 1 \leq \#_{\text{al}} \leq 50 \), evaluated at \( \lambda_D = \lambda_D^{\text{max}} \) for each structure. As the number of active layers and/or the refractive index contrast increases, the photonic effects become stronger and \( \bar{u}_{\text{rel}} \) increases. The associated increase in sharpness of the peak on the \( \lambda_D \) axis with respect to the layer thickness causes a correspondingly increasing sensitivity to structural imperfections. In fact, to obtain the maximum value of \( \bar{u}_{\text{rel}} \approx 400 \) shown in Fig. 3(d), the production accuracy must be subatomic. In this case, the peak of \( \bar{u}_{\text{rel}} \) on the \( \lambda_D \) axis reaches a sharpness in the sub-Ångström-range. Such accuracy can never be realized, and the extremely high \( \bar{u}_{\text{rel}} \) value is thus non-physical.

To take into account production tolerances, we apply a Monte Carlo method as described in section 2.1. The simulation results for four different \( \sigma \)-values are shown in Fig. 4. As can be seen,

![Fig. 4](image_url)

Fig. 4. Average relative energy density \( \bar{u}_{\text{rel}} \) across the active layers of a Bragg structure as a function of \( \#_{\text{al}} \) and \( n_{\text{high}} \). \( \bar{u}_{\text{rel}} \) is shown for four different production accuracies, simulated using a Monte Carlo method (see section 2.1). For each pixel, 50,000 separate calculations were carried out. The black contours indicate 99% (solid line) and 95% (dashed line) of the maximum. The panels show \( \sigma \)-values of 0.1 nm (a), 0.5 nm (b), 1.0 nm (c), and 5.0 nm (d). The non-zero production tolerances limit the realistically achievable value of \( \bar{u}_{\text{rel}} \) severely. Even with Ångström precision, as displayed in Fig. 4(a), corresponding to the thickness of a single
atomic layer, the maximum value of $u_{rel}$ drops to around 125. With 0.5 nm and 1 nm precision, Figs. 4(b) and 4(c), the maximum decreases further to values around 35 and 20, respectively. Figure 5 shows the associated spread in $u_{rel}$ for an exemplary $n_{high}=2.3$. As more layers are added, the increasing sharpness of $u_{rel}$ for the ideal structure causes a higher spread, which in turn lowers and broadens the mean-value maximum. The position of the mean-value maximum shifts towards structures with fewer layers, as exemplified in Fig. 5(b), and/or smaller refractive index contrast (see Fig. 4). As the peak in $u_{rel}$ for the ideal version of these structures is broader, they are less sensitive to structural imperfections compared to structures with more layers and/or a higher refractive index contrast. For $\sigma = 5$ nm, Fig. 4(d), the maximum moves further down, especially to lower $\#_{al}$, and the maximum value drops to around 6. Hence, to obtain high enhancement factors, high-precision manufacturing is of uttermost importance. Going forward, a production accuracy of $\sigma = 1$ nm is assumed, which is realistically attainable with current high-precision manufacturing methods.

### 3.2. Local density of optical states

The band structure and the relative local density of optical states (LDOS$_{rel}$) within the Wigner-Seitz unit cell is shown in Fig. 6 for two example structures with $n_{high}=2.3$ (a) and $n_{high}=3.0$ (b). Comparing the two structures, a compression of the band structure along the frequency axis when going from $n_{high}=2.3$ to $n_{high}=3.0$ is observed due to the increasing effective refractive index. In addition, the features become more pronounced and the size of the band gaps increases due to the increase in refractive index contrast. The LDOS influences the probability of all spontaneous emission processes (see section 2.3). In this work, the upconversion (UC) emission is almost exclusively caused by the spontaneous emission from energy level $3 \rightarrow 1$ (SPE31), while the radiative losses are dominated by the $2 \rightarrow 1$ emission (SPE21), see Fig. 2. To allow for a simple assessment of the LDOS effects of different structures, the ratio between the LDOS$_{rel}$ for SPE31 and SPE21 is plotted in Fig. 7. Since, ideally, the UC emission should be enhanced and the loss emission suppressed, this ratio should be as high as possible. In Fig. 7(a), the LDOS$_{rel}$ is plotted for an example structure with $n_{high}=2.3$. The $x$-axis of Fig. 7 is $\lambda_D$, representing the scaling of the unit cell in position space. The orange and blue shaded regions indicate the position of the first photonic bandgap (1PBG) for SPE31 and SPE21, respectively. LDOS$_{rel}$ is strongly reduced when the respective emission falls into the 1PBG. Therefore, the maximum ratio is observed when SPE21 is located in the region of the 1PBG, which is very close to $\lambda_D^{max}$, the design wavelength that maximizes the relative energy density $u_{rel}$. It is important to note...
Fig. 6. Illustration of the LDOS for two different high index materials, \( n_{\text{high}} = 2.3 \) (a) and \( n_{\text{high}} = 3.0 \) (b), along with the associated photonic band structure for \( k'_y = 0 \). The regions of \( n_{\text{high}} \) and \( n_{\text{low}} \) (the active region) within the Wigner-Seitz unit cell are indicated on the top. The band gaps are marked by blue shadings. To avoid washing out features in the left panel, the scale is truncated at 2.0 even though the maximum value in the right panel is 2.5.

Fig. 7. (a) LDOS rel for SPE31 (main UC emission) and SPE21 (loss emission), as well as their ratio for an example structure with \( n_{\text{high}} = 2.3 \) as a function of the design wavelength \( \lambda_D \). The shaded regions indicate where the respective transition falls within the first photonic bandgap (1PBG). (b) Ratio of LDOS rel for SPE31 and SPE21 as function of \( n_{\text{high}} \) and \( \lambda_D \). The solid lines indicate where the respective transition falls within the 1PBG.

that the emission wavelength of 1558 nm is significantly stokes-shifted relative to the excitation wavelength \( \lambda_{\text{exc}} = 1523 \) nm. Therefore, \( \lambda_{\text{exc}} \) can be efficiently coupled into the structure at the edge of the 1PBG, while the emission wavelength is suppressed in the 1PBG.

In Fig. 7(b), \( n_{\text{high}} \) is varied on the y-axis. The grey dashed lines indicate the example structure shown in Fig. 7(a). The orange and blue lines indicate the edges of the 1PBG for SPE31 and
SPE21, respectively. For all \( n_{\text{high}} \), the behavior is similar to Fig. 7(a). The ratio is small when the UC emission SPE31 falls into the 1PBG and large when the loss emission SPE21 lies in the region of maximum suppression in the 1PBG. Additionally, the ratio increases with increasing \( n_{\text{high}} \). This can be understood from Fig. 6. As \( n_{\text{high}} \) increases, the features of the LDOS become more pronounced. Thereby, the contrast in the LDOS increases between the region of the 1PBG and the band edges surrounding the 1PBG. We conclude that the most favorable design is obtained by placing SPE21 in the 1PBG while utilizing materials with the largest possible refractive index contrast.

3.3. Upconversion photoluminescence and quantum yield

Having analyzed the photonic effects of the local energy density enhancement and the local density of optical states (LDOS) separately, we now turn to their effect on UC as modeled using the rate equation model described in section 2.3. A detailed analysis of the photonic effects within the rate-equation-modeling framework, leading up to this work, can be found in [32]. In this section, we want to give an overview of the effects a Bragg structure can have on the upconversion quantum yield (UCQY) and upconversion photoluminescence (UCPL). To limit the analysis to physically realizable structures, a production accuracy of \( \sigma = 1 \text{ nm} \) is assumed as discussed in section 3.1.

In Fig. 8, we illustrate the dependence of the UCQY on the incident irradiance. To limit the parameter space, we fix \( \lambda_D \) at \( \lambda_{D}^{\text{max}} \), the design wavelength where the enhancement of the energy density is at its maximum. For low irradiances, this is a reasonable approximation, as the energy density enhancement is the most important effect in this regime. Additionally, as discussed in sections 3.1 and 3.2, \( \lambda_{D}^{\text{max}} \) is typically close to the \( \lambda_D \) for which the benefit of the LDOS, i.e. the ratio between \( \text{LDOS}_{\text{rel}} \) for SPE31 and SPE21, is at its maximum. Since the peak of this ratio in \( \lambda_D \) space is much broader than that of \( \bar{u}_{\text{rel}} \), \( \lambda_{D}^{\text{max}} \) will typically yield a UCQY value close to the maximum possible value.

![Fig. 8. UCQY as a function of incident irradiance I for exemplary families of Bragg structures with (a) \( n_{\text{high}} \) fixed at 2.3 while \#al is varied and (b) \#al fixed at 10 while \( n_{\text{high}} \) is varied. In both cases \( \lambda_D = \lambda_{D}^{\text{max}} \). For the Bragg structures, higher UCQY values at much lower irradiances are achievable. The maximum for each design is marked with a large dot.](image-url)

In Fig. 8(a), we fix \( n_{\text{high}} \) at 2.3 and vary only the number of active layers, \#al. This allows studying the effect of \( \bar{u}_{\text{rel}} \), as the calculated LDOS varies only slightly with \#al (due to the change in \( \lambda_{D}^{\text{max}} \)). To reach the maximum UCQY for the reference, an irradiance of 11600 W/m² is needed. This optimal irradiance is characteristic for the regarded material system. At higher

In Fig. 8(b), we fixed \( n_{\text{high}} \) at 2.3 and varied only the number of active layers, \#al. This allows studying the effect of \( \bar{u}_{\text{rel}} \), as the calculated LDOS varies only slightly with \#al (due to the change in \( \lambda_{D}^{\text{max}} \)). To reach the maximum UCQY for the reference, an irradiance of 11600 W/m² is needed. This optimal irradiance is characteristic for the regarded material system. At higher
irradiances, the population of higher energy levels becomes more dominant, such that the UCQY decreases. For the Bragg structures, the UCQY curve is compressed along the irradiance axis, i.e. a lower incident irradiance is needed to achieve the optimal irradiance at the position of the upconverter. From Fig. 4(c) we know that for \( n_{\text{high}} = 2.3 \), \( n_{\text{rel}} \) increases with increasing \#_n. Therefore, the compression of the UCQY curve on the irradiance axis is stronger for higher \#_n. Additionally, the maxima for the Bragg structures are slightly higher than for the reference due to the modified LDOS. For additional discussion on this point, we refer to [32]. In Fig. 8(b), \#_n is fixed at 10 while \( n_{\text{high}} \) is varied. From Fig. 4(c) we know that at \#_n = 10, \( n_{\text{rel}} \) increases with increasing \( n_{\text{high}} \). That is, as the refractive index contrast increases, fewer layers are needed to reach the same energy density enhancement. Similar to the case of increasing \#_n, a compression along the irradiance axis occurs. Additionally, the Bragg structure maximum point goes to higher UCQY values with increasing \( n_{\text{high}} \) due to the increasing strength of the LDOS modification as illustrated in Figs. 6 and 7.

To enable an assessment of structural designs in two dimensions, we reduce the dimensionality of the parameter space by fixing the irradiance at specific application scenarios. For non-concentrated sunlight, the irradiance available from the air-mass 1.5 global spectrum within the absorption range of Er\(^{3+}\) (from 1450 nm to 1600 nm) is approximately 30 W/m\(^2\) [19]. A higher incident irradiance case of 1000 W/m\(^2\) could be reached by combined spectral and geometrical concentration [67]. To also investigate the photonic upconverter system in the high irradiance regime, we regard the case of 10000 W/m\(^2\) (1 W/cm\(^2\)).

For each scenario, the UCQY is shown as a function of \( n_{\text{high}} \) and \#_n in the left panels of Fig. 9. In the first panel, \( I = 30 \text{ W/m}^2 \), the incident irradiance is so low, that the UCQY is determined almost exclusively by the energy density enhancement. This explains the structural resemblance of Fig. 9(a) to Fig. 4(c). The highest UCQY reached for this irradiance is 6.7%. At \( I = 1000 \text{ W/m}^2 \), Fig. 9(c), UCQY values up to 15.4% become possible. With \( n_{\text{high}} = 2.3 \), 95% of this maximum value can be reached with 20 layers, while for \( n_{\text{high}} = 3.0 \) only 10 layers are needed. Because of the saturation of the UCQY that is clearly visible in Fig. 8 for the reference, the UCQY does not feature a strong dependence on the irradiance. Hence, the observed maximum of the UCQY is rather broad. Going to even higher irradiances, Fig. 9(e), saturation occurs and the UCQY starts decreasing (see also Fig. 8). Hence, in concentrated-solar applications, the benefit of the Bragg structure decreases regarding the UCQY. However, it should be noted that higher UCQY values than shown are possible by tuning the design wavelength. In the case of a very high incident irradiance, the assumption of \( A_D = \lambda_{\text{max}}^u \) being an almost-ideal choice for maximizing the UCQY, is no longer valid.

For some applications, the UCPL is of more interest than the UCQY. Therefore, we also investigate the relative UCPL (UCPL\(_{\text{rel}}\)) as shown in the right panels of Fig. 9, again as a function of \( n_{\text{high}} \) and \#_n. While the absorption enhancement is directly given by the energy density enhancement, the UCPL depends non-linearly on the local energy density. This non-linearity saturates to linearity at a characteristic irradiance threshold. Therefore, at low irradiances far from saturation, the UCPL can be increased by orders of magnitude compared to the performance of the reference. For the very low irradiance of \( I = 30 \text{ W/m}^2 \), Fig. 9(b), the Bragg structure enables a 330-fold UCPL enhancement. At \( I = 1000 \text{ W/m}^2 \), Fig. 9(d), the maximum UCPL enhancement has decreased to \( \approx 40\)-fold. At this irradiance the relative effect of the Bragg structure is lower because the UCQY of the reference is already much higher. Proceeding to the case of \( I = 10000 \text{ W/m}^2 \) (1 W/cm\(^2\)) plotted in Fig. 9(f), the enhancement factor drops further to a maximum of \( \approx 12 \). At this high irradiance, the UCQY is in fact lower for some Bragg structures compared to the reference, which causes the UCPL enhancement to drop below the energy density enhancement. In consequence, the ideal design for the UCPL enhancement is a different one than for the UCQY enhancement.
Fig. 9. UCQY (left) and relative UCPL (right) as a function of \( n_{\text{high}} \) and \( \#_{\text{al}} \) for \( \lambda_D = \lambda_{\text{umax}} \).

The black contour lines indicate 99% (solid line) and 95% (dashed line) of the maximum in each plot. The rows show different irradiance scenarios of \( I = 30 \text{ W/m}^2 \), 1000 \text{ W/m}^2, and 10000 \text{ W/m}^2 (1 \text{ W/cm}^2).

3.4. Discussion of results

Experimental work on erbium-doped distributed Bragg reflectors has previously been published by Johnson et al. [34]. Their structures were made of Er\(^{3+}\)-doped porous silicon, where the periodic variation of the refractive index was achieved by varying the silicon porosity. At an excitation wavelength of 1550 nm and a laser power of 200 mW, an UCPL enhancement of 26.6 is reported for the 550 nm UC emission. The main UC emission of 980 nm was enhanced by a factor of around 5. The investigated structure by Johnson et al. corresponds to our simulated Bragg structure of \( \#_{\text{al}} = 30 \), featuring refractive indices of \( n_{\text{low}} \approx 1.5 \) and \( n_{\text{high}} \approx 2.2 \). Johnson et al.
Johnson et al. report difficulties in controlling the layer thickness accuracy in the structures, hence we draw a comparison to our lowest simulated accuracy. Johnson et al. do not report the exact irradiance, but that the 200 mW laser beam is focused on the sample. Therefore a high irradiance can be assumed and we draw a comparison to our highest simulated irradiance. At a production accuracy of $\sigma = 5$ nm and an irradiance of 10000 W/m$^2$, we calculated an UCPL enhancement of 4.8. This is very well in line with the measured enhancement of a factor of 5 by Johnson et al., especially, when taking into account the complexity of both the simulation model and the experiment.

Works on other photonic structures are not directly comparable. Nevertheless, it is possible to draw a comparison in terms of advantages and disadvantages of the different structures. In a waveguide structure a very high enhancement of $10^6$ has been reported by Lin et al. [43]. The downside of this very effective device, at least in application in photovoltaics, is that the enhancement occurs only within a very narrow range in excitation wavelength and incident angle (approximately 1 degree). Opal photonic crystal structures often exploit surface effects, which allow only for a thin layer of upconverter material to be deposited on top of the photonic structure. This limits the amount of upconverter material affected by the enhancement and therefore the total UC signal. Enhancement factors of up to 30 have been reported by Niu et al. [38] and Yin et al. [36]. Later, Yin et al. also included gold nanorods in opal photonic crystal structures to additionally exploit plasmonic effects, raising the UC enhancement factor to three orders of magnitude [44]. In inverse opal photonic crystal structures, the voids of the structure can be filled with UC material, allowing for more upconverter material to be included in the device [37,40]. Here, the performance of the structures deviates by an order of magnitude with Zhang et al. measuring a maximum UC enhancement factor of 4.6 [40], while Xu et al. report a factor of 43 [37]. Recently Shao et al. demonstrated the additional usage of plasmonic effects, embedding gold nanoparticles in inverse opal photonic structures. With a measured UC enhancement factor of 10 [45], no clear benefit compared to pure photonic structures was observed. In comparison to all these reported structures, the Bragg structure is a promising device. The amount of upconverter material affected by the enhancement can be varied, and indeed can be quite large. Simultaneously, high UCPL enhancement factors can be reached.

Other factors that will affect the applicability of any photonic structure, especially in photovoltaics, is its sensitivity to spectral and/or angular changes of the excitation, as well as the angular characteristics of the emission. Investigating these dependencies is beyond the scope of this paper, but the topics will be addressed in our future work.

4. Summary

We have presented a simulation-based analysis of the photonic effects of a Bragg structure on the upconversion quantum yield (UCQY) and the upconversion photoluminescence (UCPL) of the embedded upconverter, $\beta$-NaYF$_4$:Er$^{3+}$. The change in local energy density and the local density of optical states were considered within a rate-equation-modeling framework. To include all realistically attainable structures in the analysis, the refractive index of the spacer material, $n_{\text{high}}$, was varied within the refractive index range of naturally occurring materials, $1.5 < n_{\text{high}} \leq 4.0$, and up to 50 active layers were considered. Furthermore, manufacturing imperfections were incorporated in the analysis via Monte Carlo simulations.

Neglecting production imperfections, the energy density across the active layers of the Bragg structure can be enhanced indefinitely, simply by adding more layers. For 50 layers, a more than 400-fold average enhancement was found for $n_{\text{high}} = 4.0$. However, when realistic production accuracies are taken into account, the maximum shifts to fewer layers and takes on a finite value. At $\sigma = 1$ nm and $\sigma = 5$ nm, for example, the enhancement drops to 20- and 6-fold, and the number of active layers needed are only 30 and 15, respectively. These observations underline the crucial importance for high-precision manufacturing methods in order to realize efficient Bragg structures experimentally. Furthermore, we showed that for all applications it is highly
profitable to use the maximum possible refractive index for the spacer layer. The higher refractive index contrast significantly increases the beneficial photonic effects. Additionally, it decreases the number of layers needed to reach a given energy density enhancement, which decreases the complexity of the fabrication process.

For an optimized Bragg structure, a 330-fold UCPL enhancement was predicted at an illumination of one sun for a production accuracy of $\sigma = 1$ nm. However, due to a non-ideal effective energy density, the UCQY remains below 7%. At an incident irradiance equivalent to 30 suns, possible e.g. using a combination of down-shifting and geometrical concentration [67], the UCQY of the optimized Bragg structure is raised to a near-maximum value of 15.4%, while the reference remains at only 8.0%.

In conclusion, we find that the Bragg structure is a very promising candidate for increasing the efficiency of upconversion processes in a broad range of applications. In particular within the context of non-concentrated photovoltaics, the combination of a high UCPL enhancement and the possibility of including a large amount of upconverting material makes the Bragg structure exceedingly favorable.

**Funding**

Baden-Württemberg Ministry of Science, Research and Arts; Baden-Württemberg Ministry of Finance and Economy; Sustainability Center Freiburg under the NaLuWiLeS project; Innovation Fund Denmark under the SunTune project; Heinrich-Böll Stiftung; Helmholtz Association under the STN research program.
As the first topologically optimized designs materialized, it became clear that our electron-beam lithography (EBL) system was unable to reproduce the designs with sufficient accuracy. This publication presents my work on resolution-enhancing software to address this problem.
8.1 Context

As the plasmonic properties of a metal nanoparticle (NP) depends strongly on the NP geometry, it is crucial that the NPs can be fabricated with high accuracy. In SunTune, the NPs have been fabricated by EBL, see figure 8.1. The resolution of EBL systems is limited by electron scattering processes. Various approaches have been proposed to limit these effects, ranging from physical modifications of the EBL system[65, 66] and correction exposure(s)[67] to dose-pattern corrections[68–70]. For the EBL system used in this work, pattern-correction software was commercially available, but it was costly, and the manufacturer would not provide any guarantee that it would actually work for nanoscale structures. Therefore, a project was initiated to develop a proximity-effect correction tool, specifically targeted at the correction of nanostructures.

![Figure 8.1: Schematic illustration of the EBL process. The substrate (a) is spin-coated with an electron-sensitive resist (b) and selected areas are exposed by an electron beam (c). In a subsequent development step the exposed areas are removed (d) and a thin-film of gold is deposited on top (e). Finally, the remaining resist is removed along with the gold attached to it, leaving only the desired pattern on top of the substrate (f).](image)

8.2 Contribution

My contributions to this work include writing essentially all of the text, carrying out all calculations and creating all figures.
Research paper

Dose regularization via filtering and projection: An open-source code for optimization-based proximity-effect-correction for nanoscale lithography

Emil H. Eriksen⁎, Adnan Nazirb, Peter Ballingb, Joakim Vester-Petersenc, Rasmus E. Christiansend, Ole Sigmundd, Søren P. Madsenc

⁎ Corresponding author.
E-mail address: emher@au.dk (E.H. Eriksen).

Keywords: Proximity effect correction Optimization

ARTICLE INFO

Abstract

A new method for dose regularization in optimization-based proximity-effect-correction is proposed. In contrast to the commonly adopted approach of adding penalty terms to the objective function, a modified scheme is demonstrated where dose regularization is achieved via filtering and projection techniques.

The resulting dose patterns are simple and two-toned, and can thus readily be applied in production. Furthermore, existing extensions developed in the context of topology optimization that build on top of the filtering framework, such as robust optimization and strict length scale control, can be adopted directly. The validity of the scheme is assessed in experiments, where the resolvable feature size of the considered 30 kV electron-beam lithography system is decreased from around 100 nm to a few tens of nm. A Python implementation of the scheme is made freely available.

1. Introduction

In electron-beam lithography (EBL), an electron beam is scanned over the surface of a thin layer of electron-sensitive resist. The exposure alters the physico-chemical structure of the resist, which enables selective removal of the exposed/unexposed areas in a subsequent development step (positive/negative resist). The resolution of EBL systems is limited by forward scattering of electrons as they travel through the resist along with back scattering of electrons from the substrate. While the range of the backscattered electrons is relatively long (~10 μm), the forward scattering is more local with a range of some tens of nanometers [1]. These effects, often denoted jointly as proximity effects, cause unintended exposure of regions adjacent to the intended dose pattern.

Since the 1970s, a number of proximity-effect-correction (PEC) schemes have been proposed. Dose-modification schemes, such as the self-consistent method [2,3], the pattern-area density map [1], and various transform- [4] and optimization-based [5-7] approaches, modify the dose pattern based on an underlying model of the EBL process. Shape-modification schemes employ a more empirical approach where the pattern shapes are modified according to experimentally-determined design tables [8,9]. Other powerful techniques include background equalization methods, such as GHOST [10], and various hybrid approaches [11-13]. While the early literature considers micrometer-sized structures, the focus of more recent research is shifting towards the nanoscale. At the nanoscale, the main contribution of the proximity effect is from forward scattering, which requires highly localized PEC. Recent studies have been addressing the sub-100-nm regime utilizing different concepts such as geometry assisted PEC design rules [14] and minimal variation of critical-development time [15].

In this work, we consider structures with a size below 1 μm and feature sizes smaller than 100 nm. We employ an image-based representation of the dose pattern and formulate the PEC scheme as an optimization problem. The main drawback of optimization-based approaches is that they tend to yield complex dose patterns, which are difficult to realize in practice. This issue is commonly addressed by regularization via the addition of penalty terms to the objective function [6,7]. In this work, an alternative dose regularization strategy, based on the filtering and projection techniques developed in the context of topology optimization [16,17], is demonstrated. The resulting dose patterns are simple and two-toned, and can thus readily be adopted in production. To demonstrate the practical applicability of the resulting PEC scheme, test geometries with sub-50-nm features are...
fabricated experimentally. In the hope of stimulating the development of open-source PEC tools, a Python implementation of the scheme is made freely available.

2. Modeling

The blueprint design $B$ and the dose pattern $D$ are both two-dimensional quantities. The resulting physical structure $P$ can be approximated as a two-dimensional pattern, which is extruded perpendicular to the resist surface to form a three-dimensional object. We will use the term “physical structure” and the symbol $P$ to refer to this two-dimensional pattern. To enable a discretized formulation, applicable to numerical optimization, the resist surface is divided into $n \times m$ square elements, or pixels. The elements are arranged in a regular two-dimensional grid, like the pixels of an image. Each of the quantities $(B, D, P) \in Q$ can then be expressed in discrete form as $n \times m$ element matrices $x^0$, where each entry $x^0_{ij}$ holds the value of $Q$ evaluated at the center of pixel $(i,j)$. The blueprint design and the physical structure are both strictly binary, with $0(1)$ indicating the absence(presence) of material. Similarly, the dose can be considered binary with $0(1)$ indicating that the e-beam is off(on). However, some EBL systems support multiple dose steps, in which case the dose can take the discrete values $(0, 1/\ell, ..., 1)$, where $\ell$ is the number of dose steps.

2.1. Dose optimization

The relation between the dose pattern $x^0$ and the resulting physical structure $x^D$ can formally be expressed as

$$x^D = \mathbf{T}(x^0),$$

where the operator $\mathbf{T}$ represents the forward model, i.e. a known physical model of the EBL process. The forward model used in this paper is presented in Section 2.3. Defining the difference between the desired blueprint design $x^0$ and physical structure as the error, the PEC scheme can be formulated as a least-squares optimization problem,

$$\begin{align*}
\text{minimize} & \quad f(x^D) = \frac{1}{nm} \sum_{ij} |x^D_{ij} - \mathbf{T}(x^0_{ij})|^2, \\
\text{s.t.} & \quad x^0_{ij} \in \{0, 1/\ell, ..., 1\}.
\end{align*}$$

(2)

Due to the discrete nature of $x^0_{ij}$ along with a possibly non-differentiable $\mathbf{T}$ operator, the problem in Eq. (2) is inherently non-smooth. However, the problem can be cast into a smooth form by relaxing the constraints on $x^D_{ij}$ to the continuous interval $[0,1]$ while replacing any non-differentiable operations in $\mathbf{T}$ with smooth alternatives. A key advantage of this approach is that it enables the use of highly efficient gradient-based optimization techniques. As the problem size can be large, e.g. $1 \text{ mio.}$ degrees of freedom for a $1 \times 1 \mu\text{m}^2$ domain resolved by $1 \times 1 \mu\text{m}^2$ pixels, the efficiency of the optimization is essential. The main drawback is the added problem of mapping the values of $x^0$ from the continuous interval $[0,1]$ to the discrete values $(0, 1/\ell, ..., 1)$. The continuation scheme facilitating this process is outlined in Section 2.4.

2.2. Dose regularization

Since the solution of the equation $x^D = \mathbf{T}(x^0)$ is not unique, i.e. the same physical structure can be obtained from different dose patterns, Eq. (2) is ill-posed. Solving it directly can result in noisy dose patterns with features on the length scale of the pixel size, which can be difficult to realize in practice. This problem, as well as the mapping of $x^0$ to the allowed, discrete values, can be addressed through regularization. A commonly adopted approach is via the addition of penalty terms to the objective function [6,7]. However, in this work a different strategy is proposed. It is inspired by the regularization filtering approaches employed in topology optimization [16,17].

We restate the optimization problem, Eq. (2), in terms of a new virtual field $x^V$,

$$\begin{align*}
\text{minimize} & \quad f(x^V) = \frac{1}{nm} \sum_{ij} \mathbf{I}(x^0_{ij})^2, \\
\text{s.t.} & \quad x^0_{ij} \in \{0, 1/\ell, ..., 1\}.
\end{align*}$$

(3)

The virtual field has no physical meaning, but merely serves as a proxy for expressing the dose. The dose is calculated from $x^V$ by convolution with a kernel $K$ and projection via element-wise composition with a Heaviside step function,

$$x^D = H_0(x^V*K),$$

(4)

where $H_0$ denotes the threshold value. The projection ensures that a binary dose pattern is obtained in the end. If the EBL system supports multiple dose levels, an $l$-step dose pattern can be achieved simply by replacing the Heaviside step function in Eq. (4) by its $l$-step equivalent. In this work, we will limit our attention to binary dose pattern, i.e. $\ell = 2$. The kernel $K$ is taken to be a Gaussian of width $\sigma_0$.

The combination of $\sigma_0$ and $H_0$ provides a handle to control the complexity of the dose pattern independently of the pixel size, making it possible to wash out discretization-dependent noise. Additionally, the reduced mesh dependence smooths the landscape of the objective function, increasing both convergence speed and stability of the optimization. However, if $\sigma_0$ is chosen too large, small features will be rounded. For $\sigma_0 = 0/1$, a strict enforcement of length scale is possible for the solid/void phase [17]. Simultaneous control of the length scale of both phases can be achieved by imposing geometrical constraints [18] or by further projections [19].

2.3. The EBL process

To facilitate a clear presentation of the examples, a simple model of the EBL process is employed. However, it must be emphasized that the proposed framework can straightforwardly be applied to problems with more complex forward models. We assume that the resist is thin, so that the dependence of the exposure $E$ on the distance to the substrate can be neglected [20]. It can then be modeled in two-dimensional space as the convolution of the dose aimed at the resist surface with a point spread function PSF that describes the electron-scattering effects,

$$x^D = x^V \otimes \text{PSF}.$$  

(5)

When both forward and backward scattering are considered, a double-Gaussian is typically employed [10], but as our main focus is forward scattering, the PSF is, without loss of generality, chosen as a single Gaussian in the remainder of this work. As customary in the PEC literature, the variance $\sigma^2$ of the Gaussian distribution is expressed via a beam-broadening parameter $\alpha = \sqrt{2\sigma}$. Following exposure, the resist is developed, at which point areas with exposure below a threshold value $\eta$ are removed (a positive resist is assumed, but the PEC scheme could be formulated for a negative resist as well). This step is modeled by composition with another Heaviside step function thus yielding the predicted physical material distribution,

$$x^P = H_0(x^V \otimes \text{PSF}),$$

(7)

where $H_0$ is applied element-wise.

2.4. Implementation

The PEC scheme was implemented in Python. The convolution operations were carried out as multiplications in Fourier space with the FFT operations performed by the FFTW library [21]. The resulting complexity of the forward model is $O(nm \log(nm))$ due to the involved FFT operations. The optimizations were carried out using the L-BFGS-B
The gradients were evaluated via the expression
\[
\frac{\partial f(x^0)}{\partial x^0} = \frac{1}{nm} \left[ \frac{\partial H_\xi(\xi)}{\partial \xi} \right]_{\xi=x^0} \left[ 2(x^0 - x^0) \frac{\partial H_\eta(\eta)}{\partial \eta} \right]_{\eta=0} \ast \text{PSF} \ast K,
\]
with \( \frac{\partial H_\xi(\xi)}{\partial \xi} \) applied element-wise, derived using the chain rule. In this form, the complexity of evaluating the Jacobian is \( O(nm \log(nm)) \). To enable differentiation of the (non-differentiable) Heaviside step function, it is replaced by a smooth version \[23\],
\[
H_\xi^\beta(\xi) = \frac{\tanh(\beta \eta) + \tanh(\beta(\xi - \eta))}{\tanh(\beta \eta) + \tanh(\beta(1 - \eta))}.
\]

The steepness of the projection is controlled by \( \beta \). Low values of \( \beta \) indicate a soft projection while \( \lim_{\beta \to 0} H_\xi^\beta(\xi) = H_\xi(\xi) \). To limit the non-differentiability and the restriction put on design freedom by the projection we start the optimization at a low value of \( \beta \), where the problem is smooth, and the dose pattern takes values in the range \( [0;1] \). By increasing \( \beta \) continuously, using the result of the previous iteration as the initial guess for the next, the dose pattern is gradually pushed towards the discrete values \( \{0,1\} \). The resulting PEC algorithm is sketched in Algorithm 1.

Algorithm 1. Outline of the PEC algorithm. The minimization problem in line 5 is obtained by combining Eqs. (3), (4), and (7). To achieve a perfectly binary design, the final projection is carried out using the standard (non-smooth) Heaviside step function.

The full code is available on pypi under the package name pecpy.

3. Numerical examples

As the first example geometry (G1), we consider the blueprint design of Fig. 1a. It bears resemblance to the classical example of Haslam and McDonald [4], but the overall size has been reduced to yield a nanoscale structure. It is represented on a 768 × 768 nm² domain, discretized into \( 1 \times 1 \) nm² pixels. This domain is used in all examples. The resulting PEC runtime was on the order of minutes on an Intel i7-4770 CPU.

The physical structure obtained without PEC is shown in Fig. 1e. The proximity effect is clearly visible. The large rectangles merge and their corners are rounded, while the cross is reduced to an almost rhombic shape. Applying PEC without regularization, the blueprint design is reproduced almost exactly, see Fig. 1f. However, the dose pattern, Fig. 1b, contains blurry features and grey-scale gradients, which are difficult to realize in practice. Fig. 1c, g shows the effect of the proposed regularization scheme, with \( \eta_0 = 0.1 \) and \( \sigma_0 = 10\sqrt{2} \) nm. These parameter values were found to provide a good compromise between noise reduction and feature blurring and have been used throughout this work. Compared to the case without regularization, Fig. 1b, f, the error increases slightly, but it remains practically negligible. The dose pattern, on the other hand, changes drastically. The blurry features and the greyscale gradients vanish, leaving a two-toned, production-friendly dose pattern. Note that even though a particular length scale is not strictly enforced, the dose pattern appears simple and free from tiny features. These observations are qualitatively in agreement with Christiansen et al. [24], where a double-filtering approach was found to yield better length scale control in practice for related topology optimization problems.

For comparison, a PEC scheme with regularization via the addition of penalty terms in the objective function was implemented. We follow the approach of Poonawala and Milanfar [6] and apply a combination of a gradient term (to decrease the dose pattern complexity) and a quadratic term (to achieve a binary dose pattern). To prevent stagnation, their weight factors were turned up gradually with \( \beta \). The gradient (quadratic) term improves (reduces) the smoothness of the objective function. Their individual weights must be chosen carefully to keep the optimization stable while avoiding excessive rounding of the dose pattern, which occurs if the weight of the gradient term becomes too large. Furthermore, as the magnitude of the gradient norm is geometry dependent, a separate calibration of the weight parameters must be carried out for each geometry to achieve optimal performance. Employing optimized weight parameters, the dose pattern in Fig. 1d, h was obtained. Overall, the pattern complexity as well as the error is similar.
to Fig. 1c, g. These observations imply that the performance of the proposed regularization scheme is on par with established approaches, but with less parameter tuning needed (the same $\eta$- and $\sigma$-values were found be to efficient across all considered geometries). However, in contrast to the established approaches, the proposed scheme offers a suite of powerful extensions, developed in the context of topology optimization, building on top of the filtering framework. Besides strict length scale enforcement [18,19] and multi-level dose projections, as mentioned in the previous sections, other relevant techniques include robust optimization [25], dimensional reduction [26], and unification of the device design and PEC process [7].

4. Test geometries

In addition to G1, three other test geometries were selected to assess the performance of the proposed PEC scheme. The second geometry (G2) is a punctured disk with circular voids of varying size, Fig. 2a, making it possible to estimate the minimum resolvable hole size. The third and fourth geometries (G3, G4) go beyond circular and rectangular shapes and include features in the sub-50-nm range, see Fig. 2b, c. G3 was derived from an in-house topologically optimized plasmonic antenna design, while G4 was inspired by Ginzburg et al. [27]. G3 and G4 enable an assessment of the applicability of the PEC scheme to more exotic geometries.

For G1, the optimized dose pattern, Fig. 1c, is characterized by two distinct features. First, to avoid the large rectangles merging, their ends are shrunk close to the narrow gap. Similarly, dose is removed near the center of the small cross to avoid legs merging. Second, to increase the sharpness of the corners, spots of dose are added near the corners. Inspecting previous iterations of the optimization (not shown), it was found that pointy features similar to the corner-assist features proposed by Ocola et al. [14] develop from the corners. In the final stages of the optimization, these features separate from the parent geometry and turn into the observed spots.

The optimized dose pattern for G2, Fig. 2d, turns out to be rather intuitive. To avoid making the enclosing disk too large, the over-all size is reduced. Meanwhile, the size of the holes is increased to prevent overexposure. For G3 and G4, the optimized dose patterns, Fig. 2e, f, are much less intuitive. While similar strategic elements are apparent (size-reduction, fracturing), the resulting dose patterns are rather complex. They are unlike anything one would expect from manual refinement or rule-based corrections. In both cases, the blueprint design is reproduced within ±2 pixels of its perimeter.

5. Experiment

The test geometries were fabricated in gold on a silicon substrate. The substrate was spin coated with positive resist (950PMMA) and post-baked at 180 °C for 5 min, yielding a 100 nm thin film. The exposure was carried out using a FEI Magellan 400 SEM system (30 kV accelerating voltage, 13 pA current) equipped with a Raith pattern generator (100 × 100 μm² writing field, 4 nm step size). Post-exposure, the samples were developed for 30 s in a solution (1:3) of methyl isobutyl ketone (MIBK) and isopropyl alcohol (IPA), whereas pure IPA was used as a stopper for 30 s. Subsequently, 2 nm of titanium (adhesion layer) and 20 nm of gold were deposited through the mask using a physical vapor deposition system equipped with an e-beam-gun source (deposition rate 0.3 Å/s). Finally, the samples were soaked in acetone to remove unwanted material.

5.1. Parameter fitting

To fit the model parameters, a two-dimensional array of structures, corrected with $\eta$- and $\sigma$-values varying along each axis, was fabricated. Inspecting the resulting structures in SEM, the most accurate representation of the blueprint design was identified, and the best parameter choice was determined from its row/column number, see Fig. 3. In general, the model parameters can depend on all aspects of the EBL setup (substrate material, resist type and thickness, acceleration voltage and current of the electron beam, the development process, etc.). However, they do not depend on the blueprint design, so once the parameters have been determined for a particular EBL setup, any...
structure can be corrected. The independence on the blueprint design was confirmed in experiment with the best parameter choice ($\eta = 0.15 \pm 0.03$ and $a = 20 \pm 2 \text{ nm}$) being equal across all geometries within measurement uncertainties.

### 5.2. Experimental results

Fig. 4 shows the fabricated test structures with and without PEC. For G1, the proximity effect is clearly visible with the large rectangles merging, see Fig. 4a. Additionally, the corners are rounded, and the cross is reduced to a curvy diamond-like shape. Employing PEC, Fig. 4e,

![Fig. 3. Schematic illustration of a fabricated array of structures corrected using different values of $\eta$ and $a$. To guide the eye, a semi-transparent outline of the intended blueprint design, Fig. 2a, is shown in green. Clearly, the best match is the center panel, i.e. the best parameter choice is $(\eta, a) = (0.12, 20 \text{ nm})$. The scale bar in each panel is 100 nm. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)](image-url)

![Fig. 4. SEM images of the fabricated test structures without (a–d) and with PEC (e–h). In the correction, fitted parameter values of $\eta = 0.15 \pm 0.03$ and $a = 20 \pm 2 \text{ nm}$ were used. To guide the eye, a semi-transparent outline of the intended blueprint design is shown in green. The scale bar in each panel is 100 nm. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)](image-url)
the sharpness of the corners is increased significantly, and the large rectangles detach. The edges of the cross remain slightly rounded, but the over-all cross shape is well resolved. In the case of G2, the uncorrected dose results in a disk which is slightly too large, and with only two holes resolved, see Fig. 4b. With PEC, Fig. 4f, the diameter of the enclosing disk is corrected, and three additional holes are resolved. We suspect that the failure to resolve the last hole is due to problems in the lift-off process.

Going to the more complex geometries, G3 and G4, features smaller than ≈100 nm tend to merge without PEC. While the over-all contours of G4 remain visible, Fig. 4d, G3 is reduced to an almost blob-like shape, see Fig. 4c. Upon turning on PEC, Fig. 4g, h, all essential features of both G3 and G4 are reproduced well. The small arms of G3 now appear separated from the main body, and even the small bump at the back has become visible. The contours of G4 appear crisp, with only slight deviations near the center.

6. Conclusion and outlook

We have demonstrated how a new combined filtering and projection technique can be applied for dose regularization in the context of optimization-based PEC. The resulting PEC scheme yields binary, production-friendly dose patterns. The performance of the scheme was validated in experiments considering four different test geometries. A consistent performance was observed across all structures, decreasing the minimum resolvable feature size from about 100 nm to a few tens of nm.

By adopting the filtering technique for dose regularization in PEC, other methods building on top of the filtering framework can be applied directly. Particularly interesting examples are multi-level dose projections, robust optimization [25], strict length scale enforcement [18,19], dimensional reduction [26], and unification of the device design and PEC process [7].

Acknowledgments

This research was supported by the Innovation Fund Denmark under the SunTune project.

References

To assess the role of the nanoparticle (NP) geometry on upconversion (UC), a three-dimensional simulation framework was developed taking into account not only the change in the local energy density, but also the modification of decay rates. In this paper, the framework is utilized to study a range of different NP shapes.
9.1 Context

One of the main objectives of SunTune is to gain a better understanding on how the NP shape affects on the UC process. Ultimately, this knowledge should translate into the identification of an ideal plasmonic structure via topological optimization; this is the subject of one of the other PhD projects in SunTune. While the key advantage of plasmonic structures is the strong local energy density enhancement, the decay rates of the optical transitions are also modified. The changes are particularly pronounced close to the metal surface, where a large non-radiative channel (heat dissipation in the NP) dominates. The interplay between these different effects and their combined impact on the UC process, evaluated within a rate-equation-modelling framework as outline in section 2.2, is the subject of Publication IV.

9.2 Contribution

My contributions to this work include writing all of the text, carrying out all calculations and creating all figures.
Enhanced upconversion via plasmonic near-field effects: role of the particle shape

Emil H. Eriksen,1 Søren P. Madsen,2 Brian Julsgaard,2 Clarissa L. M. Hofmann,3 Jan Christoph Goldschmidt,3 and Peter Balling1,4

1) Department of Physics and Astronomy, Aarhus University, Ny Munkegade 120, DK-8000 Aarhus C, Denmark
2) Department of Engineering, Aarhus University, Inge Lehmanns Gade 10, DK-8000 Aarhus C, Denmark
3) Fraunhofer Institute for Solar Energy Systems ISE, Heidenhofstrasse 2, 79110 Freiburg, Germany
4) Interdisciplinary Nanoscience Center, Aarhus University, Gustav Wieds Vej 14, DK-8000 Aarhus C, Denmark

The large energy-density enhancements, associated with the near-field of plasmonic metal nanoparticles, can potentially be utilized to increase the efficiency of non-linear processes such as upconversion. A drawback of employing metallic structures for upconversion applications is luminescence quenching, i.e. the transfer of energy from the upconverter material to the metal, where it is dissipated as heat. In this study, a rate-equation model is applied to study the interplay between near-field enhancement and luminescence quenching for a range of different geometries. It is found that while shapes that incorporate pointy features and/or narrow gaps support stronger near-field enhancements, they also suffer more severely from luminescence quenching. Due to the strong correlation between the two effects, the predicted enhancement in upconversion luminescence is similar across all considered geometries ranging from 1 to 3. Our results indicate that the near-field of plasmonic metal nanoparticles might not be suitable for increasing upconversion efficiency.
I. INTRODUCTION

Upconversion (UC) is the process in which two (or more) photons are merged to form one photon with a higher energy. The concept has found applications in various fields ranging from anti counterfeiting and plastic recycling to bioimaging and theranostics. Another potential application is photovoltaics, where UC enables efficiencies beyond the Shockley Queisser limit via reduction of transmission losses. UC materials based on the Er$^{3+}$ ion are particularly interesting as they support UC from around 1500 nm to 980 nm (the exact wavelengths depend on the host material) thus bridging the band gap of silicon, which is currently the dominating solar-cell material. The main drawbacks of erbium-based compounds are a low absorption cross section and/or a too low internal upconversion quantum yield (UCQY) at excitation power densities relevant for solar applications.

Plasmonic structures enable light confinement below the diffraction limit and hence very large local energy densities. As the absorption scales with the energy density, plasmonic structures present a possible path for overcoming the low absorption in erbium. Due to the non-linear nature of the UC process, the increase in energy density can potentially also enhance the UCQY. In addition, the resulting inhomogeneous dielectric environment affects the dynamics of the UC process. Plasmonic structures can also be designed to exploit this effect to improve the UC efficiency, e.g. by enhancing the spontaneous emission rate for the UC transition(s).

The literature contains a large body of experimental studies on plasmonically-enhanced UC. Different classes of structures have been considered, ranging from island films across aperture arrays to metal nanoparticles (MNPs) of various shapes. The most commonly encountered shapes are spheres, shells, disks and rods, either arranged in periodic patterns, randomly or attached to UC nanoparticles. The reported enhancement factors range from below 1 (net loss) to several hundreds. Some studies optimize the structure so as to support a plasmon resonance at the UC emission wavelength, some at the excitation wavelength, while others perform no optimization at all. The different design objectives along with frequently missing data on excitation intensity makes it hard to compare enhancements factors across studies. Obviously, as the UC emission increases non-linearly with energy density, higher enhancement factors are obtained at low excitation power. With each study considering only one or two shapes, it is thus difficult to draw
FIG. 1. Schematic of the considered model system in 2D (a) and 3D (b). The gold nanoparticle, a 2D polygon extruded to a height of 50 nm, rests on a $\beta$-NaYF$_4$:Er$^{3+}$ substrate.

definite conclusions on the effect of the MNP shape on UC.

In this study, the performance of various MNP shapes are assessed via a simulation-based approach. To enable a fair comparison, all shapes are optimized with respect to the same objective and subject to the same excitation conditions. With current technology, arbitrary three-dimensional MNPs are difficult to realize. However, the subspace spanned by extrusions of two-dimensional polygons has been made accessible by the development of high-precision top-down fabrication techniques such as electron-beam lithography (EBL). To limit the design space to manufacturable structures, we choose as our model system such MNPs situated on an upconverting substrate, see Fig. 1. Since our main focus is the role of the MNP shape, and not particle coupling effects, a single-particle model is employed. In experiment, the condition of negligible particle-particle coupling is valid if the MNPs are distributed randomly across the substrate and the inter-particle spacing is not too small\textsuperscript{17,18}.

The upconverting material is taken to be $\beta$-NaYF$_4$:Er$^{3+}$ (refractive index of $n = 1.5$, main absorption at 1523 nm, main UC emission at 984 nm\textsuperscript{19}), which ranges among the most efficient UC materials\textsuperscript{8}.

II. METHODS

The presence of a plasmonic MNP modifies both the local energy density of the incident field and the local dielectric environment as expressed by the dyadic Green function $\vec{G}(\mathbf{r}, \mathbf{r'})$. The energy density, which is proportional to the square of the electric field amplitude, affects the rate of stimulated processes $W_{fi}$,

$$W_{fi}(\mathbf{r}) \propto |\mathbf{E}(\mathbf{r}, \omega_{fi})|^2,$$

(1)
where f/i denotes the final/initial states and \( \omega_{fi} \) the transition frequency. The dielectric environment influences the local density of optical states (LDOS) as well as the rate of Förster resonant energy transfer (FRET)\(^2\). The modified LDOS causes a change in the spontaneous decay rate \( \gamma_{fi} \) as per Fermi’s golden rule,

\[
\gamma_{fi}(r) \propto \text{LDOS}_{fi}(r).
\] (2)

In non-periodic systems, significant change in the Förster energy transfer rate is limited to within a few nm of the metal surface\(^{21,22}\). For the structures considered in this work, we find that the luminescence quenching in these regions is so severe that the upconversion luminescence (UCL) is essentially zero. Hence, we can safely neglect changes in the FRET rate.

A. Local energy density

As the MNPs considered in this work are large enough that quantum effects can be neglected\(^{23}\), the local energy density can be obtained by solving Maxwell’s equations. We apply the finite-element method (FEM) in the scattered-field formulation with open boundaries, and the background field is evaluated via the transfer-matrix method (TMM)\(^{17}\). All calculations are carried out with monochromatic excitation (\( \lambda = 1523 \) nm) at normal incidence, and the results are polarization averaged. The FEM calculations are performed in COMSOL Multiphysics\(^{24}\), while the TMM problem is solved using an in-house developed code\(^{25}\). Details are available in our previous work\(^{17,26}\).

B. Decay rates

Since the dyadic Green function is not known for arbitrary inhomogeneous systems, we evaluate the decay rate numerically. Approximating each transition as a two-level quantum system, the decay rate can be obtained via the identity\(^{20}\)

\[
\gamma(r) = \gamma^0 \cdot \frac{P(r)}{P^0}.
\] (3)

Here \( P(r) \) and \( P^0 \) are the power emitted by a dipole in the arbitrary inhomogeneous system and in the bulk UC material, respectively, while \( \gamma^0 \) denotes the decay rate in the bulk UC
material. For numerical evaluation, it is convenient to express $P(r)$ as the integral of the time-averaged poynting vector $\langle\mathbf{S}\rangle$ over the surface $\Omega$ of a sphere enclosing the dipole,

$$P(r) = \int_{\Omega} \langle\mathbf{S}\rangle \cdot \mathbf{n} \, d\Omega,$$

(4)

where $\mathbf{n}$ denotes the unit vector normal to the surface element $d\Omega$ and $r$ the location of the dipole. When the MNP is present, it is important to distinguish between two families of spheres; those which encompass it and those which do not. In the former case, the integral will include only the power $P^r$, which is radiated to infinity. The difference, $P - P^r$, is absorbed by the MNP and dissipated as heat. The decay rate can thus be split into a radiative and a non-radiative part,

$$\gamma(r) = \gamma^r(r) + \gamma^{nr}(r),$$

(5)

where

$$\gamma^r(r) = \gamma^0 \cdot \frac{P^r(r)}{P_0}, \quad \gamma^{nr}(r) = \gamma^0 \cdot \frac{P(r) - P^r(r)}{P_0}.$$  

(6)

The quantum yield (QY) of the emitter is given by the ratio $\gamma^r/\gamma$.

To evaluate the integral in Eq. 4, we first solve a full-field FEM problem with a dipole excitation source. The spatial dependence of $\gamma$ can be mapped out by sweeping the dipole across the domain, carrying out three FEM calculations (one for each possible polarization) at each spatial point. Unless the MNP is rotationally symmetric, this step is computationally very intensive. Employing a regular grid with 5 nm resolution, to resolve the rapid changes in $\gamma$ near the MNP edges, a domain of e.g. 1000 nm x 1000 nm x 100 nm would require a total of 2.4 million FEM calculations. Since $\gamma$ varies only slowly away from the MNP surface, this number can be reduced significantly via an appropriate adaptive sampling strategy. In this work, a simple scheme, which assigns points based on the local variation in the QY and the current point density, was used. An example of a resulting sampling is shown in Fig. 2. Algorithmic details are available in appendix A.

C. Upconversion luminescence

The combined effect of the modified energy density and decay rates on the UCL can be assessed via a rate-equation model (REM), which considers the dynamics of the UC process. A REM taking into account the first seven energy levels of $\beta$-NaYF$_4$:Er$^{3+}$ has been developed
FIG. 2. Illustration of the adaptive sampling scheme applied in the $xy$-plane 5 nm below a dimer (305 nm in diameter) for 984 nm emission. Panel (a) shows the sampled points (grey dots at line intersections) and the resulting triangulation (blue). Only the first quadrant is evaluated, the remaining domain (semi transparent) is reconstructed via mirror operations. The particle geometry, not known by the sampling algorithm, is indicated by black shadings. Panel (b) shows the resulting (interpolated) quantum yield.

by Fischer et al.\textsuperscript{27}. It was initially intended for homogeneous media, but later extended to photonic\textsuperscript{19,28,29} and plasmonic materials\textsuperscript{30,31}. A less complex model, considering only the first four energy levels, was later published by Christiansen et al.\textsuperscript{32}. The advantages of the latter model are less fitting parameters and more physical insight.

From the steady-state solution of the REM, the UCL of each transition from an infinitesimal volume $dV$ can be calculated as $A_f \cdot N_i(r) \cdot dV$, where $A_f$ denotes the Einstein coefficient of emission and $N_i$ the (steady-state) population density of the initial level. When the MNP is present, the $A_f$ coefficient must be scaled proportional to the resulting radiative decay rate,

$$A_f = A_f \frac{\gamma_f(r)}{\bar{\gamma}_f}.$$  \hspace{1cm} (7)

For the system considered in this work, the UCL can be estimated as the luminescence from the $^4I_{11/2} \rightarrow ^4I_{15/2}$ (984 nm) transition. Due to the complexity of the full REM by Fischer et al.\textsuperscript{27}, the steady-state solution can only be obtained numerically. Meanwhile, the lower complexity the of REM by Christiansen et al.\textsuperscript{32} makes it possible to derive an approximate analytical expression for the UCL enhancement. For intermediate excitation intensity, it
reads
\[
\frac{UCL_{\text{est}}(r)}{UCL^0_{\text{est}}} = \frac{\gamma_{984}(r)}{\gamma_{984}(r)} \frac{\gamma_{1558}(r)}{\gamma_{1558}(r)} \frac{|E(r)|^3}{|E^0|^3} = QY_{984}(r) \frac{\tau_{1558}(r)}{\tau_{1558}(r)} \frac{|E(r)|^3}{|E^0|^3}.
\] (8)

Besides the QY of the main UC emission and the energy density, the estimate depends on the lifetime (\(\tau = 1/\gamma\)) of the main loss transition \(4I_{13/2} \rightarrow 4I_{15/2}\) (1558 nm). To simplify the notation, variables have been indexed by their emission wavelength, rather than the initial/final states. This convention will be adopted throughout this work.

III. RESULTS

The main desireable effect of a plasmonic MNP is the local increase in energy density, which is strongest when the MNP is excited close to resonance. The resonance frequency depends on the shape and size of the MNP, as well as the MNP material and the surrounding dielectric environment. In this study, the latter two parameters are kept fixed (see Fig. 1). For a particular shape, this leaves size as the only free parameter. To enable a fair comparison between different shapes, the individual sizes are chosen so as to maximize a common objective. A possible choice of objective is the extinction cross section. However, it turns out that the peak in field enhancement is red-shifted significantly relative to the peak in extinction, which makes explicit integration of the energy density (to some power \(m\)) across the UC volume a more appropriate choice. Since the discretized domain of the FEM model is finite, a finite volume must be selected for the integration. For the structures considered in this work, the near-field enhancement was found to be limited to within a distance of \(\lambda/(n \cdot 10) \approx 100\) nm of the MNP. Hence for the integration of the UCL we choose a volume extending 100 nm from the edges of the MNP and 100 nm into the film, see Fig. 3b.

A. Disk

The simplest and most well-studied MNP geometry is a sphere. We therefore begin our analysis with its two-dimensional analog, a disk. Since a disk is rotationally symmetric and excitation and emission are assumed unpolarized, a cut through the plane spanned by the \(z\)-axis and a vector perpendicular to it (the radial direction \(r\)) holds all information. We first consider the two main desireable effects of the MNP, the change in energy density and in the radiative decay rate of the main UC emission, see Fig 3a,c. A significant increase
FIG. 3. Visualization of relative energy density (a), radiative rate of the main UC emission (c), QY of the main UC emission (d), lifetime of the main loss transition (e), the simplified UCL estimate (f), and the UCL calculated via the full REM (g) below a plasmonic disk ($d = 420$ nm). The dashed white lines indicate the borders of the integration region, the average (non-logarithmic) enhancement across this region is shown in the right corner. The structure is sketched in (b) with the region shown in the other panels indicated by a black shading.

in energy density occurs at the edge of the MNP. However, it drops off quickly reaching a value close to unity around 100 nm from the edge. A similar trend is observed for $\gamma_{984}$, though the enhancement is lower. Unfortunately, the advantageous effects are accompanied by severe luminescence quenching close to the MNP. Within 20 nm of the particle, the QY drops significantly, and the lifetime of the first excited state decreases, see Fig 3d-e.

The rotational symmetry of the disk decreases the computational complexity of the decay rate calculations (the most time-consuming step in the UCL calculation) dramatically. The disk is thus an ideal candidate for doing a comparison of the UCL predicted by the full
REM with the simplified estimate. To evaluate the effect of the MNP via the full REM, \( \gamma \) must be calculated for all (15) transitions. However, not all transitions affect the UCL significantly. To assess the importance of each transition, a number of calculations were carried out scanning \( \gamma/\gamma^0 \) from 1 to 10 (values of \( \gamma/\gamma^0 > 10 \) occur only very close to the MNP where luminescence quenching dominates) while monitoring the UCL. Neglecting transitions yielding a change in the UCL below 1\%, six important transitions remain - the main loss transition (1558 nm), the main UC emission (984 nm), and additionally the \( ^4I_{9/2} \rightarrow ^4I_{15/2} \) (814 nm), \( ^4F_{9/2} \rightarrow ^4I_{15/2} \) (660 nm), \( ^4I_{11/2} \rightarrow ^4I_{13/2} \) (2672 nm) and \( ^4I_{9/2} \rightarrow ^4I_{13/2} \) (1705 nm) transitions. Employing the full REM for an incident intensity of 30 W/m\(^2\) (the energy available at one sun within the absorption band of erbium\(^{35}\)) while including the change in decay rate for these transitions, the UCL map of Fig. 3g was obtained. Despite luminescence quenching effects, the UCL is enhanced significantly near the particle edge. However, away from the edge the UCL is suppressed, in particular below the MNP. Integrated across the selected integration volume (indicated by white, dashed lines), the MNP does increase the UCL, but the enhancement is only by a factor of \( \approx 2.5 \). For comparison, the simplified UCL estimate calculated via Eq. 8 is shown in Fig. 3f. The overall agreement with the more elaborate model is good. This observation confirms that the simplified model captures the important physics for the considered system. Furthermore, as the simplified estimate does not depend on any parameters specific to \( \beta\)-NaYF\(_4\):Er\(^{3+}\), conclusions drawn based on the simplified UCL estimate can be expected to hold more generally for erbium-based upconverters as compared to the more elaborate UCL measure. For the remainder of this paper, the UCL will be calculated via the simplified estimate.

B. Single-element shapes

In addition to the disk, a suite of different shapes has been considered. Since infinitesimal gaps and perfectly sharp corners cannot be fabricated, all shapes have been rounded to a minimum radius of curvature of 10 nm and all gaps are set to a minimum spacing of 10 nm. Furthermore, the minimum feature width is set to 50 nm. These numbers represent typical fabrication tolerances for modern EBL systems.

It is generally accepted that pointy structures support larger enhancements of the electric field as compared to rounded structures\(^{20,36}\). To investigate the influence on the UCL, we
FIG. 4. Visualization of the $z$-directional average (from the surface and 100 nm into the UC material) of $|E|^3/|E_0|^3$ (a,e,i), QY of the main UC emission (b,f,j), lifetime of the main loss transition (c,g,h), and the simplified UCL estimate (d,h,l) across the $xy$-plane for a disk (a-d), a rod (e-h) and a triangle (i-l). The dashed white lines indicate the the borders of the integration region, the average enhancement across this region is shown in the right corner. The scale bar is 200 nm.

compare each of the terms in Eq. 8 (and the resulting UCL enhancement) for a disk, a rod and a triangle in Fig. 4. As expected, the sharper corners of the latter two structures enable much higher field enhancements as compared to the disk. Hot spots are located at points of small radius of curvature, i.e. at the ends of the rod and at the corners of the triangle. The largest average enhancement is achieved by the rod, partly facilitated by its smaller volume to surface ratio. The triangle, on the other hand, supports the largest point-wise field enhancement. In both cases, the large field enhancements are accompanied by severe luminescence quenching; the two effects show a strong spatial correlation. The luminescence quenching counteracts the field enhancement, so that the resulting average UCL enhancements for the triangle and the rod become comparable to that of the disk.

Another path to achieving high field enhancements is via gap plasmons. In Fig. 5, we consider a few commonly encountered gap geometries, a dimer, a split rod and a bowtie. As
FIG. 5. Visualization of the z-directional average (from the surface and 100 nm into the UC material) of $|E|^3/|E_0|^3$ (a,e,i), QY of the main UC emission (b,f,j), lifetime of the main loss transition (c,g,h), and the simplified UCL estimate (d,h,l) across the $xy$-plane for a dimer (a-d), a split rod (e-h) and a bowtie (i-l). In all cases, the gap is 10 nm. The dashed white lines indicate the borders of the integration region, the average enhancement across this region is shown in the right corner. The scale bar is 200 nm.

compared to their single-element siblings, the gap structures yield larger field enhancements; both on average and point-wise. The largest average enhancement, exceeding two orders of magnitude, is achieved for the split rod. Meanwhile, the largest point-wise enhancement is observed for the bowtie. The observed tendencies for the gap geometries are thus similar to their single-element siblings. Inside the gaps, the UC enhancement is large, but it decreases elsewhere. As a result, the average UC enhancement factor is in fact lower for the gap geometries in all cases.

Comparing all of the considered shapes, a wide range of field enhancement factors is observed. However, the difference in the average UC enhancement is small. The structures that achieve high field enhancements suffer from higher luminescence quenching and vice versa. These observations are in qualitative agreement with previous studies considering...
a gold sphere\textsuperscript{31}, where regions of high field enhancement were identified as regions of high luminescence quenching too. However, our observations suggest that this correlation is not unique to spherical MNPs, but rather a more general property of plasmonic MNPs. This indicates that possible plasmonic efficiency-enhancing elements in future upconversion devices should be based on effects beyond the near-field of isolated structures, e.g. via periodic and/or wave-guiding structures. It is important to note that plasmonic near-field effects can still boost the efficiency of higher-order processes, such as surface-enhanced Raman scattering, which scales with the electric field to the power of four\textsuperscript{20}. The highly localized field enhancements might also be useful for other purposes, e.g. in localized sensing applications. Plasmonic MNPs can also cause scattering of the light, which can increase the effective path length in the surrounding absorber, thereby improving the efficiency. Experimentally observed UCL enhancements, larger than predicted by near-field calculations, have previously been attributed to this effect\textsuperscript{14}.

IV. SUMMARY AND CONCLUSION

The potential for increasing the upconversion efficiency of $\beta$-NaYF$_4$:Er$^{3+}$ via plasmonic near-field effects has been assessed for a range of metal nanoparticles of different shapes. To enable a fair comparison between the different shapes, the overall size of each particle was optimized with the respect to the same objective. While shapes with pointy features and/or gaps were able to support larger field enhancements, they were also subject to stronger luminescence quenching. Taking both effects into account via a rate-equation modeling approach, the resulting upconversion enhancement was found to be similar, numerically between 1 and 3, across all geometries. These observations indicate that the near-field of weakly-coupled plasmonic metal nanoparticles might not be suitable for increasing the efficiency of upconversion.

V. ACKNOWLEDGMENTS

This research was supported by the Innovation Fund Denmark under the SunTune project. Stefan Fischer is acknowledged for providing insightful comments on the manuscript.
REFERENCES


28Barbara Herter, Sebastian Wolf, Stefan Fischer, Johannes Gutmann, Benedikt Bläsi, and Jan Christoph Goldschmidt. Increased upconversion quantum yield in photonic structures due to local field enhancement and modification of the local density of states—s a simulation-based analysis. *Optics Express*, 21(105):A883–A900, 2013.


Appendix A: Adaptive sampling

To simplify the adaptive sampling routine, the three-dimensional problem is split into two-dimensional sub-problems. As the MNP polygon is located in the $xy$-plane, it is natural to do the slicing along the $z$-axis. The resulting two-dimensional slices are evaluated via a two-dimensional adaptive sampling scheme as outlined in Alg. 1.

**Algorithm 1** Basic principle of the adaptive sampling algorithm. The $a$ parameter controls the respective weighting of the triangle area (density of sampling points) versus the local variation in the QY of the emitter. The $b$ parameter controls the degree of parallelization. Setting $b > 1$, multiple points are evaluated in parallel. The cost is a slightly less ideal sampling. The overall sampling density is determined by the convergence tolerance $tol$.

1: function ADAPTIVE2D($a, b, tol$)
2: vertices ← vertices of the initial coarse grid
3: do
4:   $\gamma_r, \gamma_{nr} \leftarrow$ calculate $\gamma_r, \gamma_{nr}$ at each vertex in the grid using FEM
5:  triangles ← Delaunay triangulation of vertices
6:  for $i = 0$ to length of triangles do
7:      area, QYs = area of triangles[$i$], $\gamma_r/(\gamma_r + \gamma_{nr})$ at each vertex of triangles[$i$]
8:      weights[$i$] ← area/$a$ + standard deviation of QYs
9:  new_vertices ← center of the longest edge of the $b$ triangles with the highest weight
10: vertices ← vertices + new_vertices
11: while Max(weights) > tol
12: return vertices, $\gamma_r, \gamma_{nr}$
The phenomena of extraordinary optical transmission (EOT) through sub-wavelength apertures in thin metallic films is associated with surface plasmons at the metal-dielectric interface. In this paper, the possibility of utilizing an EOT-inspired design to achieve plasmonically enhanced upconversion (UC) with minimal quenching is investigated.
10.1 Context

My interest for exotic material systems was awoken by a talk on epsilon-near-zero materials that I attended at a conference in Boulder, Colorado in fall of 2017. However, as the proposed structure was nearly impossible to fabricate, I started looking for alternative options. It was during this search that I stumbled across the concept of EOT. Starting from a standard EOT structure, i.e. a periodic array of sub-wavelength apertures in a thin metallic film, I came up with a number of designs and carried out numerical assessments of their performances (in terms of up-conversion luminescence (UCL) enhancement potential). Utilizing the knowledge gathered through this process, a production-friendly variant was formed based on discussions with colleagues working within the field of nanoscale fabrication.

Even though manufacturability was considered a key property in the design process, problems with fabrication were encountered. As a result, the first prototype structure was not realized until July 2018. Besides not adhering exactly to specifications (with respect to the optimized design parameters), the initial prototype suffered from the formation of small pillars inside the holes during the etching step (see the paper draft). The root cause is probably the micromasking effect[71], i.e. the aluminum mask is unintentionally sputtered and redeposited inside the holes. A possible solution could be to use a different material for the mask, e.g. chromium or aluminum nitride. However, due to time constrains, it was not possible to create more samples. Therefore, the experimental section in the paper draft is based exclusively on the initial prototype sample.

10.2 Contribution

My contributions to this work include writing all of the text, carrying out all calculations and creating all figures.
Leveraging extraordinary optical transmission to enhance upconversion efficiency

Emil H. Eriksen,1 Adnan Nazir,2 Jeppe D. Christiansen,1 Søren H. Møller,1 Søren P. Madsen,3 Brian Julsgaard,1 and Peter Balling1,2

1) Department of Physics and Astronomy, Aarhus University, Ny Munkegade 120, DK-8000 Aarhus C, Denmark
2) Interdisciplinary Nanoscience Center, Aarhus University, Gustav Wieds Vej 14, DK-8000 Aarhus C, Denmark
3) Department of Engineering, Aarhus University, Inge Lehmanns Gade 10, DK-8000 Aarhus C, Denmark

Plasmonically-enhanced upconversion is achieved via a new type of structure, inspired by the phenomena of extraordinary transmission as observed in sub-wavelength aperture arrays in metallic films. By strategically embedding the upconverter material, TiO₂:Er³⁺, between a such film and the inverse pattern (an array of gold disks), upconversion enhancements up to nearly two orders of magnitude are achieved in simulations. An experimental realization of a suboptimal device yields an enhancement factor of more than 20, thus supporting the validity of the proposed concept.
I. INTRODUCTION

In the geometrical limit, the fraction of energy transmitted through an aperture array in a metallic film is equal to the relative area covered by the holes, i.e. the transmission efficiency is one. The transmission efficiency through sub-wavelength apertures was predicted by Bethe\textsuperscript{1} to decrease drastically with the hole radius $r$, scaling as $(r/\lambda)^4$. However, a few decades ago, transmission efficiencies exceeding unity were observed in experiment\textsuperscript{2}. The phenomena is now known to be caused by coupling to surface plasmons at the metal-dielectric interface, and is referred to as extraordinary optical transmission (EOT)\textsuperscript{2–4}.

Similar to other plasmonic structures, the sub-wavelength aperture array enables light confinement below the diffraction limit. The resulting extremely high local energy density allows for increasing the efficiency of non-linear processes. Commonly encountered examples include surface-enhanced Raman spectroscopy (SERS) and upconversion (UC). The latter has found application across various fields, ranging from anti counterfeiting\textsuperscript{5,6} across bioimaging\textsuperscript{7,8} to photovoltaics\textsuperscript{9–11}. In this work, we consider the TiO$_2$:Er$^{3+}$ material system excited at 1523 nm, resulting in a main UC emission at 984 nm. Bridging the band gap of silicon, this process is an excellent candidate for UC applications in the context of silicon-based photovoltaics. However, due to the low absorption cross section of erbium, the UC efficiency is currently too low for commercial exploitation\textsuperscript{11}. Nanoscale light concentration via plasmonic structures could pose a path to overcome this issue.

Besides changing the excitation conditions, the presence of metallic structures also modifies the dielectric environment of the UC system. As a result, the spontaneous emission rates change due to a modified local density of optical states (LDOS). Furthermore, non-radiative decay channels arise via coupling to the metal. Energy, which is transferred to the metal, is subsequently dissipated as heat (ohmic loss) and thus lost to the UC process. For plasmonic nanoparticles, this luminescence quenching effect tends to correlate with the energy density enhancement\textsuperscript{12,13}. In this work, we explore the possibility of leveraging EOT-inspired designs to achieve regions of high energy density away from the metal surface, so as achieve a higher UC efficiency via reduced quenching.
II. NEAR-FIELD ENHANCEMENT

The most common form of an EOT system is a thin metallic film perforated periodically by circular holes. For appropriate choices of hole diameter \( d \) and period \( p \), EOT can be achieved\(^2\). Under these conditions, strong near-field enhancements are present near the film, and the UC material should thus be placed here. A drawback of such a configuration is that the transmitted light cannot be utilized by the UC process. A possible solution is to place a reflector behind the perforated metallic film, so the transmitted light can be reflected back through the holes, potentially increasing the near-field enhancement. If the reflector is chosen as a continuous film, as shown in Fig. 1a, a significant amount of the upconverted light will be trapped inside the structure. A better option is to choose the reflector as the inverse pattern of the perforated metallic film, i.e., an array of disks, see Fig. 1b. An alternative design that builds on the same principles, but which can be fabricated more easily (see Sec. V), is shown in Fig. 1c. To facilitate physical realization of the device, we focus our attention on the latter design using a SiO\(_2\) substrate with TiO\(_2\):Er\(^{3+}\) as the UC material and gold for the metallic elements (the perforated film and the reflector disks).

![Fig. 1. Schematic illustration of EOT-inspired designs for possible UC enhancement structures. The designs are periodic, each panel shows a single unit cell.](image)

The family of structures, Fig. 1c, can be parametrized via the thicknesses of the layers of gold, \( h_{\text{Au}} \), and TiO\(_2\):Er\(^{3+}\), \( h_{\text{UC}} \), in addition to \( p \) and \( d \). For intermediate excitation power, the upconversion luminescence (UCL) scales approximately with the electric field amplitude \( |E| \) to the power of three\(^1\). Hence, we consider as an initial figure of merit (to be further developed later in the manuscript) the enhancement factor of the integral of \( |E|^3 \) across the volume \( V \) of the UC material (details on the evaluation of electric field distributions are
available in Appendix A). As the reference system, we choose a TiO$_2$:Er$^{3+}$ film of the same thickness $h_{UC}$ resting directly on the substrate. Quantities evaluated in the reference system are marked by a 0-index. To ease notation, we define the operator

$$\hat{I}(Q) = \frac{\int V Q^3 dV}{\int V^3 Q_0^3 dV},$$ (1)

so that the initial figure of merit can be written as $\hat{I}(|E|^3)$. Choosing exemplary values of $h_{Au} = 150$ nm and $h_{UC} = 210$ nm, the effect of varying $p$ and $d$ can be visualized in a two-dimensional map, see Fig. 2a. The maximum enhancement factor of around 350 is achieved for $(p,d) = (1150$ nm, 750 nm). To enable a visual assessment of the resulting field distribution, cross-sectional plots of $|E|$ in units in the incident field amplitude are shown in Fig. 2b-d. While the highest values (around 10) are located near the metal surface, Fig. 2d, large values (up to $\approx 8$) are also observed inside the UC material far (more than 100 nm) from the metal surface. These observations indicate that the EOT-inspired design might be less prone to quenching as compared to nanoparticles$^{12,13}$.

It should be noted that the electric field amplitude across the reference structure (not shown) is below unity (it varies between 0.4 and 0.7) due to the high refractive index of the UC material and reflections from the film surface. The increase in the electric field amplitude due to the EOT-inspired design is thus larger than values indicated in Fig. 2c,d.

### III. UPCONVERSION LUMINESCENCE

The dynamics of the UC process can be modeled via a rate-equation model (REM). Considering a such model, Christiansen et al.$^{14}$ found that the UCL under intermediate excitation power scales approximately as

$$\text{UCL} \propto |E|^3 \left(\frac{\gamma_984}{\gamma_984} \right) \left(\frac{1}{\gamma_1558}\right) = |E|^3 QY_{984}\tau_{1558},$$ (2)

where $\gamma_X$ denotes the decay rate of transition $X$, $\gamma^r_X$ the radiative part of the latter, $QY_{984}$ the quantum yield of the main UC emission at 984 nm and $\tau_{1558}$ the lifetime of the main loss transition at 1558 nm (details on the evaluation of these quantities are available in Appendix A). Using this expression, we define our final figure of merit $\hat{I} (\text{UCL}) = \hat{I} (|E|^3 QY_{984}\tau_{1558})$. To gain a better understanding of the role of each term, the panels of Fig. 3 show the spatial variation of each of them across the $xy$-plane.
FIG. 2. Panel (a) shows the dependence of $\hat{I} (|E|^3)$ (see text) on the diameter $d$ and period $p$ for the EOT-inspired design, Fig. 1c. The maximum is marked by a white, shaded circle. Panels (c,d) show plots of $|E|$ evaluated across the planes indicated in (b) for the optimal parameter choice, $(p,d) = (1150 \text{ nm}, 750 \text{ nm})$, with the incident field polarized along the $y$-axis. The white shadings indicate regions of UC material. All values are evaluated for normally incident monochromatic excitation at $\lambda = 1523 \text{ nm}$.

The first term, $|E|^3$, is shown in Fig. 3a. The largest values, around 500, are found at the center. Additionally, regions with values exceeding 100 are present near the edge of the unit cell close to the $x$ and $y$ axes. The resulting integral enhancement factor is $\approx 350$ (in accordance with Fig. 2). The decay rate of the main UC transition (not shown) show some fluctuations across the domain. However, as the variation in the radiative and the non-radiative part are strongly correlated, the QY remains nearly constant across the hole as well as the outer region, see Fig. 3b. Meanwhile, the direction of emission (not shown) differs between the two regions. Light emitted from within the hole is predominantly emitted towards the back (transmitted), while light emitted from the outer part is mainly emitted towards the front (reflected). The lifetime of the main loss transition is reduced, as illustrated in Fig 3c. The largest reduction is at the center where the field enhancement is largest. The
FIG. 3. Visualization of the different terms in the UCL estimate, Eq. 2. The structure is that of Fig. 1c with $h_{\text{Au}} = 150$ nm, $h_{\text{UC}} = 210$ nm, $p = 1150$ nm and $d = 750$ nm. To enable a two-dimensional representation, the plots show the $z$-directional average across the region containing the UC material; outside the white circle for $z \in [0:210\text{nm}]$ and inside for $z \in [360\text{nm}:570\text{nm}]$ (see Fig. 2b,c). The panels show (a) the electric field amplitude to the third power, (b) the quantum yield of the main UC emission at 984 nm, (c) the lifetime of the main loss transition at 1558 nm, and (d) the UCL estimate.

UCL, Fig 3d, is therefore less focused at the center as compared to the $|E|^3$ measure. The resulting integral enhancement is just below two orders of magnitude. Hence even though quenching does lower the UCL (by a factor of nearly 4), a significant UCL enhancement remains. The EOT-inspired design thus seems like a promising alternative to plasmonic nanoparticles$^{12,13}$.
IV. SENSITIVITY ANALYSIS

The absorption band of the Er$^{3+}$ ion has a FWHM of $\approx 50$ nm (depending on the host material and the doping concentration)\(^1\). It would thus be desirable if the enhancement in the electric field is not too narrow with respect to the excitation wavelength, $\lambda$. Furthermore, from a manufacturing perspective, it is important that the enhancement in the electric field (and ultimately the UCL) is not too sensitive to changes in the design parameters. Since the computation of the full UCL measure is computationally complex, we focus the sensitivity analysis on $\hat{I} (|E|^3)$. Figure 4 shows the sensitivity with respect to all system parameters (except $p$ and $d$, which are shown in Fig. 2).

![Graphs showing sensitivity analysis](image)

**FIG. 4.** Dependence of $\hat{I} (|E|^3)$ on (a) the excitation wavelength, (b) the height of the gold layer, (c) the height of the upconverting layer, and (d) the hole depth (see text). The orange circles represent calculated values, the blue line is a guide to the eye.

Fig. 4a shows a FWHM in $\lambda$-space of roughly 40 nm for $\hat{I} (|E|^3)$. While a broader peak would be preferable, a reasonable enhancement will be present across most of the erbium absorption band. The sensitivity with respect to $h_{Au}$, Fig. 4b, is rather low. As long as the gold is thick enough, larger than $\approx 100$ nm, $\hat{I} (|E|^3)$ is only influenced weakly. On the other hand, it is extremely important that $h_{UC}$ is realized with high accuracy, see Fig. 4c.
A deviation of only $\pm 5$ nm from the target thickness decreases the value of $\hat{I}(|E|^3)$ by almost 50%. In our fabrication approach, the inner cylinder is created by etching a hole in the substrate (see Sec. V). Ideally, the hole depth should be $h_{UC} + h_{Au} = 360$ nm. The sensitivity to hole depth is shown in Fig. 4d. The FWHM is around 30 nm, indicating that up to $\approx 10$ nm offset from the specification can be tolerated. The performance is also somewhat sensitive to changes in $p$ and $d$, but since these parameters are realized through a mask created by electron-beam lithography (EBL), sufficient accuracy is readily achieved (see Sec. V). From a fabrication point of view, the most critical parameter is thus $h_{UC}$.

V. EXPERIMENT

The main reason for choosing the design of Fig. 1c rather than Fig. 1b is that the former design can be achieved using a single EBL step, and alignment issues are thus avoided. The fabrication strategy used in this work is shown in Fig. 5. A cylindrical pattern of resist was defined via EBL, after which a thin film of aluminum was deposited by physical vapour deposition (PVD). Soaking the sample in stripper, the resist cylinders were lift-off, leaving only the aluminum mask (Fig. 5a). Next, holes were etched using reactive-ion etching (RIE), and residual aluminum was removed by immersing the sample in aluminum etchant (Fig 5b). Subsequently, the gold layer was deposited via thermal evaporation (Fig 5c). Finally, the UC material was sputtered on top using an RF-magnetron sputtering system (Fig 5d).

![FIG. 5. Schematic of the fabrication process. A mask of aluminum is deposited on the substrate (a) to enable selective etching of the substrate (b). Subsequently, gold is deposited thereby forming the perforated gold film and the back reflector in one step (c). Finally, the UC material is sputtered on top (d). Each panel shows a single unit cell of the periodic pattern.](image-url)
Fig. 6 shows SEM images of the fabricated sample at each step in the process. The aluminum mask, Fig. 6a, is deposited as intended. After the etching, Fig. 6b, small pillars have formed inside the holes. While some holes suffer from severe contamination by the pillars, others appear clean. Overall, the distribution of pillars seems random. After gold deposition, Fig. 6c, the pillars remain visible, and even more so after the sputtering of the UC material, Fig. 6d. Meanwhile, the hole diameter has decrease from around 740 nm after etching to just 610 nm after the UC material has been sputtered.

FIG. 6. SEM images taken at each fabrication step, after (a) aluminum deposition, (b) etching, (c) gold deposition, and (d) sputtering of UC material. To visualize the pillars, which form during the etching step (b), more clearly, the zoom is increased in panels (b-d). The hole diameter is $\approx 740$ nm in (b) and $\approx 610$ nm in (d). The scalebar is 1500 nm in panel (a) and 500 nm in panels (b-d).

To assess the optical properties of the fabricated structure, the transmittance $T$ was measured as a function of wavelength, see Fig. 7a. For comparison, a simulated spectrum based on parameters extracted from the SEM images, $(p,d) = (1100 \text{ nm}, 675 \text{ nm})$, is shown. A peak in (simulated) transmittance is observed just below 1500 nm. It is accompanied by an even larger peak around $\lambda = 1900$ nm. From the field distributions (not shown), it is clear that the latter peak is not suited for UC enhancement. Via coupling to surface plasmons, the light travels across the back reflector, similar to how it travels through the hole array. Hence there is no focusing of the energy density on top of the back reflector where the UC material is located. The experimental data are in qualitative agreement with the simulations. Both the large peak around $\lambda = 1900$ nm and the peak related to the UC enhancement are present, although the latter is redshifted and the magnitude much lower. Likely explanations for the discrepancies are the contamination by the pillars (see Fig. 6), and the resulting surface roughness, along with the curved height profile of the holes as illustrated in Fig. 7b.
The UCL was measured at 1500 nm excitation using an integrating sphere and normalized with respect to a reference sample produced in the same UC material deposition run. In case of the reference, the UC material is deposited on a clean, flat substrate, which might yield a better material quality as compared to the EOT-inspired structure. This would lead to lower enhancement factors. The UCL enhancement is plotted in Fig. 7c as a function of incidence angle. At normal incidence, the UCL is enhanced by a factor of 13. This is a fairly good result considering the deviations of the experimentally realized geometry from the model (contamination by pillars, rounded height profile, \((p, d)\) offset relative to optimum, etc.). At oblique angle of incidence, the enhancement increases, reaching a factor of 21 at 25 degrees. A possible explanation could be that structure is excited off-resonance (as per the transmission plot, the region of UC enhancement is around 1600 nm, while the excitation is at 1500 nm). As \(\theta\) increases, the resulting decrease in the component of the wave vector normal to the sample surface might result in a better matching of the resonance condition. From these results, it does not seem unlikely that a performance close to the theoretical \(\approx\) two orders of magnitude can be reached, if a sample adhering to the specifications was produced.
A new type of design for achieving plasmonically-enhanced upconversion has been demonstrated. Inspired by the extraordinary transmission observed from sub-wavelength aperture arrays, a simple hole-reflector design was developed. Simulations of the upconversion luminescence (UCL), including quenching effects, yield enhancement factors up to nearly two orders of magnitude. An experimental realization of a sub-optimal design reached a UCL enhancement of more than 20. If the ideal design was fabricated, it is not unlikely that an enhancement near the predicted two orders of magnitude can be reached. The proposed structure thus seems to be a better candidate for realizing plasmonically-enhanced upconversion than conventional nanoparticles\textsuperscript{13,16}. Even higher UCL enhancement factors (exceeding 300) have been predicted for one-dimensional photonic crystals, but only if they can be fabricated with extreme accuracy\textsuperscript{17}. Hence, even though the fabrication of the proposed device is not trivial, we believe that the lower demands on accuracy, as compared to the photonic structure, makes it a better candidate for upconversion applications.

REFERENCES


Appendix A: Methods

The electric-field distributions were calculated using the finite-element method (FEM) in the scattered-field formulation via COMSOL Multiphysics\textsuperscript{18}. The calculations were carried out only for the first quadrant (due to $xy$-symmetry) of the unit cell shown in Fig. 8a. Details are available in our previous work\textsuperscript{13,16}. To evaluate the modifications of the decay rates of the loss and UC transitions, we assume that each transition can be approximated as a two-level quantum system. In this case, the decay rate $\gamma$ is related to the emitted power $P$ by a dipole emitter as\textsuperscript{19},

$$\frac{\gamma(r)}{\gamma_0} = \frac{P(r)}{P_0},$$  \hspace{1cm} (A1)

where the 0-indices indicate values for a reference system. After solving a FEM problem with a dipole source, $P$ is evaluated as the integral of the time-averaged poynting vector across a small sphere enclosing the dipole. The power radiated towards infinity, $P_r$, is calculated in the same way, but with the integration surface taken as the large spheroid shown in Fig. 8b. The QY is given by the ratio $P_r/P$, while $\tau = \gamma^{-1}$. The spatial dependence of $\gamma$ is mapped out by sweeping the dipole across the domain using an adaptive sampling strategy\textsuperscript{13}.

![Fig. 8](image-url)

FIG. 8. Renderings of the considered geometries for (a) the electric-field calculations and (b) the decay rate calculations. In the latter case, the (infinite) lattice is truncated to 9 holes, and the decay rate evaluations are carried out at the center hole. One fourth of the integration spheroid is hidden to enable a more clear visualization of the hole array. The colors indicate the materials, TiO$_2$:Er$^{3+}$ (green), Au (gold), and SiO$_2$ (light blue). All elements have been made semi-transparent to enable a better visualization of the geometry.
Co-author publications

This chapter provides a brief summary of each of the co-authored manuscripts that form part of this work. Furthermore, my contributions to each publication is outlined.
**Publication VI**

This paper is the first SunTune publication. It presents the main concept of SunTune, plasmonically-enhanced upconversion at $\approx 1500$ nm excitation via Er$^{3+}$ ions. The considered system is an upconverting thin-film of TiO$_2$:Er$^{3+}$ resting on a quartz substrate. On top of the film, randomly distributed nanodisks with different diameters and mean inter-particle spacings are realized via electron-beam lithography (EBL). The experimental results are compared with finite-element method (FEM) simulations using a single-particle model. A reasonable agreement is found although an unexplained splitting on the extinction peak is observed in some regimes. The expected non-linear dependence of the upconversion luminescence (UCL) is confirmed with the UCL scaling with the electric field amplitude to the power of $\approx 3$. UCL enhancements up to 7-fold are observed.

Besides participation in the discussion and the interpretation of the results, my contributions to this work include FEM calculations, e.g. the ones used to create figure 2. Additionally, I developed a piece of software that enables random positioning of structures on our EBL system (only periodic arrangements were supported by the standard software). This software was used for creating all samples considered in this paper.

**Publication VII**

This paper presents a method for mapping out the near-field of metallic nanoparticles (NPs) via femtosecond laser ablation. The technique thus enables direct experimental validation of the FEM near-field calculations. The basic idea is to irradiate a NP below a predetermined ablation threshold, e.g. at 50% of the threshold fluence. If ablation occurs in region(s) near the nanoparticle, the near-field enhancement must have been at least two, and the ablated region(s) can be interpreted as the contour line(s) of two-fold fluence enhancement. If the laser pulse has a Gaussian shape, the decrease in fluence with distance from the center has a well-known form, making it possible to extract a complete near-field enhancement contour map from a single shot. The resulting
near-field maps of the considered star-shaped NPs are compared to FEM calculations both on and off resonance. The measurements are found to agree semi-quantitatively with the FEM calculations.

My main contribution to this work has been FEM calculations, yielding the results shown figure 4 and figure 5d-g. To enable accurate calculations, the FEM particle geometry was created from an outline a physical NP realization obtained via edge-tracing of a scanning-electron microscopy (SEM) image of the sample. As edge tilting is not supported by COMSOL, I wrote a custom mesh generator to create the tilted mesh shown in figure 4. Furthermore, I provided panel (a) of figure 1 and the mesh illustrations of figure 4.

Publication VIII

This paper considers nanotexturing of thin-film solar cells. Appropriate texturing scatters the light, thereby enhancing the effective path length in the absorber. As a result, the absorption (and thus efficiency) is improved. Light-management concepts, such as surface texturing, are particularly important for thin-film solar cells, where it is a key challenge to achieve sufficient absorption. In this paper, a thin gold film is deposited on pre-stretched polystyrene sheets, which after subsequent heat-treatment shrinks, thereby forming nanowrinkles. To assess the performance of the texture, silicon-based thin-film solar cells are fabricated on top. The increased absorption due to the wrinkles raises the efficiency from 6.8% (for the flat cell) to 9.5%. The experimental results are compared to transfer-matrix method (TMM) (for the flat reference) and FEM calculations.

My main contribution to this work has been discussion of the results in addition to the FEM calculations. To construct an appropriate geometrical model, the layer shapes were constructed via edge-tracing of cross-sectional transmission-electron microscopy (TEM) images. I also wrote the modelling section and created figure 2d-k.
Publication IX

This paper provides an overview of SunTune activities. It incorporates elements from all members, so the topics are diverse. They include fabrication of upconversion (UC) materials, EBL processing, UC in organic solar cells, optical modelling approaches, rate-equation models and topology optimization as well as experimental techniques for measuring UCL and plasmonic near-fields.

I contributed the section on electric-field calculations.

Publication X

Due to surface quenching, the efficiency of upconverting nanocrystals (UCNCs) is significantly lower as compared to bulk crystalline materials[27]. In this paper, the effect of surface passivation by enclosing the UCNCs in an inert shell (through chemical synthesis) is investigated, both in colloidal solution and in monolayer form. In the monolayer configuration, the core-shell structures yield UCL enhancements up to two orders of magnitude as compared to naked UCNCs. In an attempt to enhance the UCL further, plasmonic nanodisks engineered to yield a resonance at 1500 nm were embedded below/placed on top of the monolayer. However, only moderate enhancements, ranging from 1.5 (thick shell) to 5 (naked UCNC), were observed.

My main contributions to this work has been through discussions and FEM calculations. The latter has guided the fabrication of the nanodisks towards achieving a plasmonic resonance at 1500 nm. Data from some of these calculations are shown in figure 3b.
**Publication XI**

The optical phase change in Ge$_2$Sb$_2$Te$_2$ (GST) observed upon irradiation with laser pulses of sufficient fluence is commonly associated with the transition from crystalline to amorphous phase. In this study, the threshold for phase change and ablation (occurring at even higher irradiances) in GST thin-films is investigated using femtosecond laser pulses. Experimental threshold values are derived from SEM and TEM images and compared to two-temperature model and TMM simulations. The results reveal that the optical phase change occurs prior to amorphization, with no additional change in the optical properties upon amorphization. The observations thus suggest that amorphization is not necessary to achieve a permanent, optical phase-change.

My contributions to this work has been via discussion of modelling approaches and by supplying the code used for the TMM calculations.

**Publication XII**

A rate-equation model (REM) for upconversion of 1500 nm light to 980 nm using erbium is presented. It considers the first four energy levels along with all transitions relevant for the UC process, including Förster resonant energy transfer (FRET). Adopting a number of simplifying assumptions, an approximate analytical solution for the UCL is derived. To assess its validity, the predicted power dependence is compared to experiment, and a good agreement is observed over several orders of magnitude of excitation power.

My main contribution to this work has been via discussions of the REM, e.g. on which terms to include and which to neglect. Furthermore, the model design has been aided by comparison with results obtained via my implementation of the REM by Fischer et al. [20].
Publication XIII

In this study, a number of optimized Bragg stacks (the same type of structure considered in Publication II) are realized in experiment. The UC layers are formed by spin-coating of $\beta$-NaYF$_4$:Er$^{3+}$ nanocrystals in PMMA solution, while the TiO$_2$ spacer layers are deposited by atomic layer deposition. The resulting structures are characterized both in terms of single-layer properties as well as the overall UCL enhancement at oblique incidence. The measured UCL enhancements are compared to calculated values derived from solutions of a REM that takes into account non-zero production tolerances via monte carlo (MC) simulations.

My main contributions to this work has been through discussions and via the development of the modelling tools utilized to calculate/simulate the UCL enhancement values.
Summary and outlook

Nearly 20% of the energy in the solar spectrum is not absorbed by modern silicon-based solar cells. These so-called transmission losses can be addressed via upconversion (UC), i.e. by merging two (or more) low-energy photons into one photon with a higher energy. As UC is inherently a non-linear process, a much larger energy density than available under one sun is needed to achieve a high efficiency. In the SunTune project, which this work is a part of, the possibility of achieving light concentration via nanostructures, as an alternative to expensive and impractical conventional optical elements, is investigated. The main focus of this work has been on the development of modelling tools to enable prediction of the influence of various nanostructures on the UC process.

A two-particle model has been developed to enable the assessment of particle-particle coupling effects in sparse arrays of plasmonic metal nanoparticles (NPs) distributed randomly across a surface. It was found that the main coupling can be attributed to interference between the incident field and the far-field component of the single-particle scattered field. Furthermore, a significant dependence of the coupling strength on the dielectric environment was identified. The model was able to pre-
dict previously unexplained features in measured extinction spectra for randomly distributed plasmonic gold nanodisks situated on a TiO$_2$:Er$^{3+}$ thin film on a quartz substrate.

As an alternative to plasmonic metal NPs, a selected family of one-dimensional photonic crystals was investigated. The main benefits of the considered structures, as compared to plasmonic metal NPs, are the absence quenching (no metal is present) and the ability to incorporate large amounts of UC material (by adding more layers). The drawbacks are high sensitivity to the excitation wavelength and incidence angle, as well the thickness of the layers in the structure. Assuming a fabrication accuracy equal to that of modern atomic layer deposition systems, simulations predict up to 300 times upconversion luminescence (UCL) enhancement under one sun.

Physical realization of the plasmonic structures was achieved through electron-beam lithography (EBL). However, as the resolution of our in-house EBL system was not sufficient, a resolution-enhancing dose correction scheme was developed. For the considered 30 kV EBL system, the minimum resolvable features size was improved from around 100 nm to a few tens of nm. As the properties of plasmonic metal NPs depend strongly on the NP shape, better accuracy in fabrication may enable increased device performance.

To assess the influence of the NP shape on performance in terms of UCL enhancement, a range of different shapes has been studied numerically. Via a rate-equation model, the combined effect of a locally modified energy density and locally modified decay rates was considered. It was found that shapes with pointy features support large near-field enhancements in regions of high surface curvature. However, due to an associated increase in quenching, the resulting UCL enhancement was not affect much by the NP shape (enhancements values range between 1 and 3 across the considered geometries). These observations imply that future plasmonic UCL enhancement devices should exploit effect beyond the single-particle near field, e.g. via periodic structures and/or coupling to waveguides.

In an attempt to come up with a plasmonic structure less prone to quenching, a design drawing inspiration from the sub-wavelength apertures supporting extraordinary optical transmission (EOT) was devel-
oped. It yields a UCL enhancement of almost two order of magnitude in simulation. While the structure is somewhat sensitive to variations in some geometrical parameters, it is much less sensitive than the previously mentioned photonic structure and thus less complex to manufacture. As the initial prototype shows a UCL enhancement of more than 20 times in experiment, the EOT-inspired design seems like a promising candidate for UCL enhancement applications.

With both the photonic crystals and the EOT-inspired designs showing great performance in simulation, future work could be in either direction. However, the EOT-inspired designs are particularly interesting as they can be fabricated more easily. An obvious extension of this work would be to manufacture and characterize a device adhering to the optimized parameters determined in this work. Furthermore, it would be interesting to look into additional variations of the design, e.g. varying the size of the backreflector, changing the contour of the holes beyond simple circles and/or increasing the size of the UC layer. If the large UCL enhancement could be sustained across a thickness large enough to absorb most of the incident light, the EOT-inspired design could enter as a key component in achieving a UC efficiency high enough for photovoltaic applications.
The Galerkin method

Consider a differential equation defined on a domain $\Omega$,

$$\mathcal{L}\phi = f$$  (A.1)

where $\mathcal{L}$ is a differential operator, $f$ a source term and $\phi$ the unknown solution to be solved for. Given some approximate solution $\tilde{\phi}$,

$$\tilde{\phi} = \sum_j c_j v_j$$  (A.2)

where $v_j$ are basis functions on $\Omega$ and $c_j$ are constant coefficients, the residual $r$ is

$$r = L\tilde{\phi} - f.$$  (A.3)

Improving the accuracy of the approximate solution $\tilde{\phi}$ is equivalent to reducing $r$ at all points on $\Omega$. An effective scheme to minimize $r$ is the method of weighted residuals which enforces the condition(s)

$$R_i = \int_\Omega w_i r d\Omega = 0.$$  (A.4)

Here $R_i$ is the weighted residual integrand and $w_i$ are chosen weight functions. Choosing the weight functions as $w_i = v_i$, equation (A.4)
Appendix A. The Galerkin method

takes the form

\[ R_i = \int_\Omega v_i \left( \mathcal{L} \sum_j c_j v_j - f \right) d\Omega = 0. \]  \hspace{1cm} (A.5)

Equation (A.5) can be rewritten as a matrix equation

\[ Sc = b \]  \hspace{1cm} (A.6)

where the elements of \( S \) and \( b \) are

\[ S_{ij} = \int_\Omega v_i \mathcal{L} v_j d\Omega, \quad b_j = \int_\Omega v_j f d\Omega. \]  \hspace{1cm} (A.7)

The (numerical) solution of the differential equation has thus been cast into the problem of solving a system of linear equations.


