Computational Design of Periodic Photonic Materials: From Subwavelength Gratings to Topological Plasmonics

Norman Israel

Thesis submitted to the University of Ottawa in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Physics

Department of Physics
Faculty of Science
University of Ottawa

©Norman Israel, Ottawa, Canada, 2026

Examining Committee

The following people serve on the thesis Examination Committee:

External Examiner: Kenneth Chau
Associate Professor, School of Engineering, The University of British Columbia

Internal Examiners: Jacob Krich
Professor, Department of Physics, University of Ottawa

Thomas Brabec
Professor, Department of Physics, University of Ottawa

Seyda Ipek
Associate Professor, Department of Physics, Carleton University

Acknowledgements

I am eternally grateful to my spouse, Anathèle, whose love, patience, support, and mentally stimulating discussions have been instrumental in taking me to this stage of my PhD. I also wish to express my sincere gratitude to my friends and family, along with my spouse's family, in particular her mom, for their patience and unwavering support.

Sincere gratitude to my supervisor, Professor Ramunno, for her support and for providing me with the opportunity to pursue ideas that fascinate me. I would also like to express my sincere gratitude to my thesis advisory committee, Professor Krich and Professor Brabec, for their invaluable guidance and support in navigating the final stages of my PhD.

I would also like to thank Professor Ipek and Professor Chau for reviewing and providing valuable feedback on my thesis.

I would also like to thank Dr. Dan-Xia Xu and Dr. Yuri Grinberg from the National Research Council of Canada (NRC), as well as Dr. Flore Kunst and Julius Gohsrich at the Max Planck Institute for the Science of Light for their guidance and helpful discussions.

I am also grateful to Professor Alexandre Poulain and Dr Andrew Pelling for their invaluable guidance and support at the earlier stages of my entrepreneurial journey. In addition, I am grateful to Professor Trevor Charles of the University of Waterloo for providing further exposure to entrepreneurship and commercialization, along with valuable feedback and guidance for my entrepreneurial journey. They all reignited my interest in entrepreneurship, which I am finding to be highly valuable for the next stage of my career.

I am also thankful to the Canadian Black Scientist Network (CBSN) for their encouragement and support, and for letting me know that I am seen, heard, and valued.

The work in this thesis is possible due to support from the Nexus for Quantum Technologies Institute (NexQT), Natural Sciences and Engineering Research Council of Canada (NSERC), the uOttawa-Max Planck Centre for Extreme and Quantum Photonics, the Artificial Intelligence for Design Challenge program and Canada Research Chairs.

Declaration on AI Usage

Generative AI tools were used to improve Python coding scripts written by the author for analysis tasks such as extracting data from Comsol to make eigenmode plots and computing Zak phase. It was also used to improve scripts to extract data from Lumerical and to improve scripts that use this data to make plots. Generative AI was further used to make python scripts used to interact with Inkscape and Blender to help make diagrams. These scripts were later modified by the author to make the final diagrams. While AI was sometimes used to help find references, all references were checked and verified by the author. AI was also used to assist with software and code debugging.

Abstract

Periodic photonic materials are playing an important role in integrated photonic devices. Having strong theoretical and computational models is essential when designing these devices in order to have insight into their operation. Theoretical tools are often too idealized to describe realistic systems, so computational models are often used to address their shortcomings. However, computational tools are resource-intensive, particularly when we have to perform optimization in the large design spaces that are common in integrated photonics. This slows the design and development of integrated photonic devices. This thesis aims to address these issues by laying the foundation for two approaches that combine the best of both worlds: the low-resource requirements of theoretical models with the realistic and accurate modeling capabilities of traditional computational methods. First, we introduce a data-driven effective medium approximation for subwavelength grating structures relevant to integrated photonics. Deep neural network surrogate models, trained on full 3D electrodynamic simulations, are used to predict dispersive effective indices to reproduce the optical response of complex 3D structures using 2D models. Using six structures, we show that the method agrees with the full 3D simulations of the 3D structures across the telecommunication bandwidth, enabling computationally efficient exploration of the design space. Next, we developed a coupled-mode formalism for topological plasmonic nanoresonator chains, constructed by simulating a single resonator. Using one chain made of spherical nanoresonators, and another made of cuboid nanoresonators, we show how the formalism can be applied to study topology in plasmonic systems. Our approach has the potential to incorporate arbitrarily shaped nanoresonators, an arbitrary number of bands, and long-range couplings. Using this approach, alongside finite-element simulations of a full lattice, we unravel bulk and boundary properties for a plasmonic 1D Su-Schrieffer-Heeger (SSH) chain. The formalism has the potential to significantly reduce computational cost while providing physical insight into the emergence of topological phenomena in plasmonic platforms. This work opens up a pathway to explore and design periodic plasmonic materials, particularly those that host topological states. Together, these contributions provide relevant tools to guide the design of photonic devices in the plasmonics and silicon photonics domain.

Author Contributions

The following are the manuscripts, conference presentations, and the research commercialization activities that the author has contributed to and participated in for the PhD work presented in this thesis:

0.1 Manuscripts in Preparation

- **N. S. Israel**, D. Blankespoor, Q. Wang, D.-X. Xu, Y. Grinberg, and L. Ramunno, "Data-Driven Effective-Medium Approximation for Metamaterials in Device Designs," manuscript in preparation for *Optics Express*.

  - Presents an extension of the concept of effective-medium to finite gratings in Chapter 4. This work is more applied. It focuses on demonstrating how feasible our method is in filling the gap between full lattice simulations and conventional effective medium theories to model finite-sized silicon-on-silica gratings for applications in grating couplers.

- Qiang Wang, **Norman Israel**, Dan-Xia Xu, Yuri Grinberg, Martin Vachon, Shurui Wang and Lora Ramunno, "A Machine Learning-Based Framework for Pseudo-3D Optimization of Vertical Metamaterial Grating Couplers," manuscript in preparation for *Optics Express*.

  - Not presented in this thesis, however, it directly builds on work from Chapter 4 to design a grating of the same design presented in this thesis. It also provides some experimental validation of the work presented in Chapter 4.

- J. Gohsrich, **N. Israel**, F. Kunst, and L. Ramunno, "Topological Nanoresonator Plasmonics: A Quasinormal Mode Approach," manuscript in preparation for *Physical Review Research*.

  - Presents a coupled-mode theory for periodic topological plasmonic materials in Chapter 5. This work is exploratory. It focused on developing the foundations of

a theoretical formalism for topological plasmonics. The goal was to fill the gap between full lattice simulations and topological plasmonic theories, so as to improve our understanding of topological states in non-Hermitian systems.

- **N. Israel**, J. Gohsrich, L. Ramunno, and F. Kunst, "Topological Edge Modes from Modal Analysis of Nanoresonators," manuscript in preparation for *ACS Photonics*.

  - Presents full lattice simulations for the 1D SSH plasmonic chain in Chapter 6. This paper and the coupled-mode theory presented in Chapter 5 are complementary to each other.

0.2 Conference Presentations

- **N. Israel**, J. Gohsrich, F. Kunst, and L. Ramunno, "Non-Hermitian Topological Plasmonics with Metallic Nanoresonators," BE-STEMM Conference, 2026 [**Poster, won prize for "Outstanding Lightning Presentation"**].

- **N. Israel**, J. Gohsrich, F. Kunst, and L. Ramunno, "New Computational Design Methods for Photonic Metamaterials," Max Planck - uOttawa Centre for Extreme and Quantum Photonics (MPC) Meeting, Ottawa, Canada, 2026 [**Poster**].

- **N. Israel**, J. Gohsrich, F. Kunst, and L. Ramunno, "Non-Hermitian Topological Plasmonics with Metallic Nanoresonators," Max Planck - uOttawa Centre for Extreme and Quantum Photonics (MPC) Meeting, Erlangen, Germany, 2025 [**Poster and Talk**].

- Dan-Xia Xu [**Invited Speaker**], Yuri Grinberg, Siegfried Janz, Sergey Dedyulin, Jianhao Zhang, Martin Vachon, Shurui Wang, John Weber, Pavel Cheben, Jens Schmid, **Norman Israel**, Qiang Wang, Lora Ramunno, Dusan Gostimirovic, Andy Li, and Odile Liboiron-Ladouceur, "Advancing PIC Development Using Machine Learning: from Design to Fabrication to Optical Characterization," OFC, San Francisco, California United States, 2025 [**Invited Talk**].

- **N. S. Israel**, D. Blankespoor, D. Xu, Y. Grinberg, and L. Ramunno, "Inverse Design of Metamaterial Waveguide Gratings using Machine Learning," Black Excellence in Science, Technology, Engineering, Math, and Medicine (BE-STEMM) Conference, 2024 [**Talk**].

- **N. S. Israel**, D. Blankespoor, D. Xu, Y. Grinberg, and L. Ramunno, "Inverse Design of Metamaterial Waveguide Gratings using Machine Learning," Max Planck - uOttawa Centre for Extreme and Quantum Photonics (MPC) Meeting, Erlangen, Germany, 2024 [**Poster**].
- **N. Israel**, J. Gohsrich, F. Kunst, and L. Ramunno, "Non-Hermitian Topological Photonics and (nearly) Flatbands," Schawlow-Townes Meeting, Ottawa, Canada, 2023 [**Poster**].
- **N. Israel**, J. Gohsrich, F. Kunst, and L. Ramunno, "Non-Hermitian Topological Photonics and (nearly) Flatbands," Max Planck - uOttawa Centre for Extreme and Quantum Photonics (MPC) Meeting, Erlangen, Germany, 2023 [**Poster**].
- **N. S. Israel**, D. Blankespoor, D. Xu, Y. Grinberg, and L. Ramunno, "Inverse Design of Metamaterial Waveguide Gratings using Machine Learning," Photonics & AI Workshop, Barbados, 2023 [**Poster, won 2nd place for poster competition**].
- **N. S. Israel**, D. Blankespoor, D. Xu, Y. Grinberg, and L. Ramunno, "Machine Learning-Enabled Inverse Design of Subwavelength Waveguide Gratings using Effective Index Approximations," Photonics North, Montreal, Canada, 2023 [**Talk**].
- **N. S. Israel** and L. Ramunno, "Non-Hermitian Topological Photonics Beyond the Tight-Binding Regime," Photonics North, Montreal, Canada, 2023 [**Talk**].

0.3 Research Commercialization Activities

- Attended the Nobellum (an international startup accelerator) Innovator Summit & Expo (NISE) in Toronto to network and pitch my business idea from my PhD with investors and fund managers, 2026 [**Invited**].
- Received a Nobellum STEM Innovator Award at the NISE in Toronto for commercialization work related to my PhD, 2026.
- Participated in Nobellum pitch competition at BE-STEMM conference for my business idea related to my PhD work, 2026 [**Invited**].
- Entrepreneurship panelist for the BE-STEMM conference to discuss commercialization of academic research based on my PhD, 2026 [**Invited**].

- Participated in Lab2Market Discover Program to further look at commercializing my PhD work, 2026.
- Wrote a proposal on commercializing my PhD work, which led to my acceptance into the University of Ottawa's Graduate Entrepreneurship Experience Program (GEEP) to start exploring the commercialization of my PhD, 2023.

Contents

Examining Committee **ii**

Acknowledgements **iii**

Declaration on AI Usage **iv**

Abstract **v**

Author Contributions **vi**
0.1 Manuscripts in Preparation vi
0.2 Conference Presentations vii
0.3 Research Commercialization Activities viii

List of Figures **xiii**

List of Tables **xvi**

0 Prelude **xvii**

1 Introduction **1**

2 Background **10**
2.1 Machine Learning for Photonics 11
2.1.1 Perceptron 12
2.1.2 Artificial Neural Networks 14
2.1.3 Supervised and Unsupervised Learning 15
2.1.4 Training an ANN 16

2.2 Bloch Modes in Condensed Matter Systems . . . 18
2.3 Photonic Bloch Modes . . . 20
2.4 A Brief Discussion on SWGs . . . 23
2.4.1 Slab Model . . . 27
2.5 Topological Materials in Condensed Matter Systems . . . 30
2.5.1 The Tight-Binding Model . . . 32
2.5.2 Hermitian Topological Materials . . . 35
2.5.3 Su-Schrieffer-Heeger Model . . . 36
2.5.4 Zak Phase and Wilson Loop . . . 39
2.5.5 Non-Hermitian Topology . . . 42
2.6 Topological Photonics . . . 46
2.6.1 Photonic Tight-Binding Model . . . 46
2.6.2 Hermitian Photonic SSH Chain . . . 48
2.6.3 Topological Photonics in Open Systems . . . 49
2.7 Models for the dielectric permittivity of materials . . . 50
2.8 Metallic Nanoresonators and Quasinormal Modes (QNMs) . . . 51

3 Methods **58**
3.1 Numerical Methods for Classical Electromagnetism . . . 60
3.1.1 Finite Difference Time Domain (FDTD) . . . 60
3.1.2 Finite Element Method (FEM) . . . 68
3.2 Data-Driven Effective Medium Approximation . . . 77
3.2.1 TensorFlow . . . 78
3.2.2 Data-Driven Effective Medium on SWG Designs . . . 79
3.3 Theoretical and Computational Methods for Topological Plasmonics . . . 85
3.3.1 Numerical Implementation of Zak Phase using the Non-Hermitian Wilson Loop . . . 85
3.3.2 QNM Coupled-Mode Theory . . . 87

4 Data-Driven Effective Medium Approximation for Metamaterials in Device Designs **96**

5 Topological nanoresonator plasmonics: A quasinormal mode approach **129**

6 A Modal Analysis of Topological Plasmonic Nanoresonator Chains **154**

7 Metallic Nanoresonators on a Substrate **181**
7.1 Method 182
7.1.1 Some Preliminary Results 183

8 Ideas for Future Work **185**
8.1 Non-Hermitian Berry Phase from QNMs 186
8.2 Geometric Interpretation of Non-Hermitian Berry Phase 195
8.2.1 Non-Hermitian Wilson Loop 200

9 Conclusion & Outlook **203**

A Appendices **207**
A.1 Geometric Phases 207
A.2 Adiabatic Evolution of States 208
A.3 Connection with Topology 211
A.3.1 Hermitian Fibre Bundles 220
A.3.2 Hermitian Wilson Loop 226

References **232**

List of Figures

1.1 Schematic of periodic materials in the subwavelength (left), wavelength (middle), and geometric optics (right) regimes. 3

2.1 Illustration of a simple artificial neuron - a perceptron. From the left, we have input features, $x_1, ..., x_n$, which are data that go into the perceptron. The weights, $w_1, ..., w_n$, assign importance to input features. The bias, b, is a threshold for the activation function, $f(z)$, which determines if a signal makes it to the output, y. 13

2.2 Example of an artificial neural network made by combining multiple perceptrons. 15

2.3 Schematic of a simple subwavelength grating design. The arrow represents the direction of propagation. Waves are oriented parallel (left, TE modes) and perpendicular (right, TM modes) to the orientation. 24

2.4 SSH chain diagram of polyacetylene lattice. w and v are the intercell and intracell hoppings, respectively. They alternate through the chain. 37

2.5 SSH chain diagram. w is the coupling shown by the double lines, and v is for the single lines. 48

2.6 Illustration of the response of a metallic nanoparticle (sphere) when excited by an external field. Surface charges move in response to the excitation source (electric field, for example), which in turn causes the nanoparticle to produce an enhanced electric field in its vicinity at the resonance frequency. This is the quasinormal mode. Excitations can occur at different orders, for example, dipole (top-left), quadrupole (top-right), hexapole (bottom-left), and octopole (bottom-right). 52

3.1 Illustration of the Yee grid (staggering in space, with coordinates i, j and k) for the field components E_x, E_y and H_z evolving forward (staggered) in time. 64

3.2 Illustration of the Lumerical setup. 66
3.3 Illustration of a simple computational graph for the mathematical operations involved in computing $z = xy + b$. 79
3.4 Illustration of one DNN architecture for 2D. The points are DNNs, and the lines are the connectivity. 81
3.5 Illustration of one ANN architecture for 3D. The points are ANNs, and the lines are the connectivity. 83

7.1 A schematic of a 1D SSH chain of plasmonic cuboids on a glass substrate. The superstrate is air and the metallic particles are modeled by the Drude model. The unit cell depicted consists of two particles separated by a distance $\beta a/2$, where a is the lattice constant and β is a distance constant that controls the distance. 183
7.2 Preliminary open-boundary spectra for cuboids on a substrate for two configurations: non-trivial, with $\beta = 1.3$ (left) and trivial, with $\beta = 0.7$ (right). Two edge modes are clearly present in the gap for the non-trivial case. . . 184

A.1 Schematic of 2D sphere (left) and 2D torus (right) with loops. All loops on the sphere can be contracted to a point without cutting it or moving it off the surface. In the case of the torus, loops 1 and 2 cannot be contracted to a point without cutting it or having it fall off the surface. 212
A.2 Bandstructure (left) for different combinations of hopping parameters $(v, w) = (0.5, 1), (0.8, 1), (1, 1), (1, 0.5)$, represented by blue, orange, green and black curves, respectively. On the right is the corresponding representation in the (d_1, d_2)-plane. The singularity (band-crossing point) of the normalized d-vector, $\hat{d}$, is represented by the black point in the middle of the plane. . . 217
A.3 Schematic of parallel transport of a vector (blue) along a closed curve. It starts from the red arrow and travels clockwise. When it completes a full loop, its direction changes and it gains a phase. 221
A.4 Schematic of parallel transport of a vector along a closed curve. It starts from the blue arrow and travels clockwise. When it completes a full loop, its direction changes and it gains a phase (compare the red and blue arrows at A). The left image shows the tangent planes at three points in the loop, whereas the right image is without the tangent planes. 222
A.5 Schematic of fibre bundle with Brillouin zone/k-space (BZ) as the torus. The Hilbert space $\mathcal{H}_k$ is the fibres. 223

A.6 Schematic of triangulation of fibre bundle for Hamiltonian. 229

List of Tables

2.1 Analogy between photonic and condensed matter eigenproblems. The first column is the concept, while the second and third columns are the photonic and condensed matter setups, respectively. 22

3.1 Architecture of the hidden layers of the three DNN models used for the training of three distinct 2D and three distinct 3D forward models. Each row lists the number of neurons in the hidden layer at that position, where the input layer (not shown) would be on the left, and the output layer (not shown) would be on the right. 81

3.2 Input parameter information for DNNs. Λ is the period in the x-direction, which is the sum of L_1 to L_4, the lengths of the four segments in each unit cell. In the 3D case, Λ_z is the SWG pitch in the z-direction, and a_z is the SWG silicon width. In the 2D case, n_{swg} is the effective index of the segment in the unit cell that replaces the SWG, while λ is the wavelength of the incident light. 82

0

Prelude

It was the early 2000s, and I was still a high school student in Jamaica. I was sitting in a classroom with other students just like me from different parts of the country. We were all there to get support in preparing our application materials for foreign universities. Our mentor, a Jamaican physicist, suddenly walked into the classroom with a grin on his face. Dangling in his hands were some metal contraptions I had never seen before. I later learned that these were called vexiers: puzzles made of metal that require you to free a loop or a ring without cutting or bending the metal. One must use continuous deformations to solve them. He handed them around for us to solve. Later on, he told us that solving these puzzles exposed us to a mathematical field called topology. My fascination with the subject

was ignited then, and I later became eager to learn the subject and find applications for it in physics. Topology unravels the intrinsic global nature of spaces by studying which of their properties remain unchanged under continuous deformations.

I first found use for it when I was in high-energy physics, working on quantum gravity: the attempt to unravel the nature of gravity using quantum theory. I explored these questions using ideas at the interface of geometry and topology. During that period, I encountered the gauge-gravity duality, first elucidated by Juan Maldacena [1,2]. He showed that a quantum field theory that describes strongly interacting systems on the boundary of a universe with constant negative curvature can tell us about a gravitational theory in the interior (bulk) of this universe. While attending research seminars on topological insulators during my master's degree, I learned about an idea which bore some general conceptual similarity to this: bulk-boundary correspondence. Topological insulators resemble ordinary insulators upon inspection of the behaviour of electrons in their bulk. However, this is deceiving, as these materials conduct electric currents on their boundaries without backscattering, even when defects are present. The existence of these boundary conducting states can be inferred from the behaviour of electrons in the material's bulk. This is what is meant by bulk-boundary correspondence. It was introduced over the years

[1] Maldacena, J. M. The Large N limit of superconformal field theories and supergravity. *Adv. Theor. Math. Phys.* **2,** 231–252. arXiv: `hep-th/9711200` (1998)

[2] Ammon, M. & Erdmenger, J. in *Gauge/Gravity Duality: Foundations and Applications* 179–218 (Cambridge University Press, 2015)

by various people[3] [4,5,6,7] and thus, unlike gauge-gravity duality, cannot be attributed to a single person.

Importantly, the underlying physics between gauge-gravity duality and bulk-boundary correspondence is fundamentally different; bulk-boundary correspondence tells us we can infer the existence of boundary states from the physics of the bulk, while gauge-gravity duality relates two fundamentally different theories: a gravitational theory in the bulk and a quantum field theory living on the boundary. Despite such fundamental differences, my interest in bulk-boundary correspondence was influenced by my brief conceptual encounter with gauge-gravity duality due to the fact that they both highlight a non-trivial relationship between bulk and boundary. This higher-level feature of the two concepts is what I found most intellectually stimulating.

Even more fascinating was the fact that the boundary states of topological insulators were a manifestation of a different state of matter that cannot be understood within the Landau paradigm. The Landau paradigm has been a cornerstone of many physical phenomena I encountered often throughout my physics career. My first encounter came from learning about the water cycle, where the Landau paradigm is essential in understanding

[3]Thouless, D. J. *et al.* Quantized Hall Conductance in a Two-Dimensional Periodic Potential. *Phys. Rev. Lett.* **49,** 405–408 (6 1982).

[4]Halperin, B. I. Quantized Hall conductance, current-carrying edge states, and the existence of extended states in a two-dimensional disordered potential. *Phys. Rev. B* **25,** 2185–2190 (4 1982)

[5]Wen, X. G. Chiral Luttinger liquid and the edge excitations in the fractional quantum Hall states. *Phys. Rev. B* **41,** 12838–12844 (18 1990)

[6]Hatsugai, Y. Chern number and edge states in the integer quantum Hall effect. *Phys. Rev. Lett.* **71,** 3697–3700 (22 1993)

[7]Xiao-Gang Wen *et al.* Chiral operator product algebra and edge excitations of a fractional quantum Hall droplet. *Nuclear Physics B* **422,** 476–494. ISSN: 0550-3213 (1994)

the phase change of water at different parts of the cycle. I also encountered it in other situations, such as when learning about the origin of magnetism through the Ising model. It has also been relevant in high-energy physics to explain various phenomena, such as the formation of the majority of the mass of protons and neutrons. The Landau paradigm assumes that knowing the phase of a small region of a system is sufficient to know the phase of the entire system. Underlying it all are symmetries which are associated with different phases (orders). If two systems have the same symmetry, they are in the same phase. Changing phase leads to spontaneous symmetry-breaking. This all breaks down when dealing with topological matter. There, phases are not represented by symmetries, but by quantities that measure the global connectivity of the system. These quantities are computed using topology and are called topological invariants, as they are unchanged so long as the system is in the same topological state. Undergoing a phase change means changing the topology of the system, which is completely different from the spontaneous symmetry-breaking mechanism underlying the Landau paradigm.

I went on a bit of a journey to learn more about this subject, but then abandoned it for a while to continue pursuing my interests in gravitational physics. After returning to Jamaica, I eventually obtained a position to lecture physics at a university there. I successfully published a paper in modified gravity, after which I made the decision to return to my interest in topological materials. I wanted to see if it was possible to use machine learning to probe the bulk behaviour of electrons in materials and determine if they are topological. During this period, I learned how challenging it is to build and mass-produce technologies based on topological materials, resulting in me being skeptical of their

potential. I later met another Jamaican physicist who was doing work in optics. My various conversations with him sparked my interest in the field. I then started reading papers on the subject, where I learned about periodic photonic materials called metamaterials. Not too long after, I learned that these materials host topological states as well. Interestingly, the photonic topological states I encountered in the literature were of classical origin, whereas the condensed matter ones were of quantum origin. This was a surprise to me, as I had thought these states were uniquely quantum. Later, I realized the one thing that the classical and quantum systems had fundamentally in common was waves: the periodic photonic materials were concerned with classical electromagnetic waves, whereas the condensed matter systems were concerned with the wave-like nature of electrons in solids, which has a quantum origin.

These papers also touted potential applications which seemed more achievable in periodic photonic materials than in condensed matter systems. As a result, I began to see them as having more potential than electronic materials, as periodic photonic materials are more easily engineered compared with condensed matter systems. As a result, I became even more interested in the field and decided to pursue a PhD in photonics. As I read more papers on the topic, I became interested in non-Hermitian (open) topological systems. This led me into plasmonics, which studies how light interacts with electrons at the surface of metals. Plasmonic systems can easily incorporate non-Hermiticity, as they naturally have ohmic losses. Plasmonic platforms are more accessible than condensed matter platforms for probing non-Hermitian effects because of their higher design flexibility. This higher design flexibility also makes them attractive for designing and building devices that use

topological effects in their core functionality.

I became interested in entrepreneurship around this time, an interest which was initially ignited by being around various tech entrepreneurs in the Jamaican startup space. As a result, during my doctoral studies, I applied to the Graduate Entrepreneurship Experience Program (GEEP) at the University of Ottawa. After being accepted, I embarked on a journey to explore how I could commercialize technologies related to topological photonics and artificial materials in general. The goal of this thesis is to lay the foundation for my own research program aligned with this direction. It presents foundational work on theoretical and computational tools for one possible direction for balancing the computational complexity of full lattice simulations with the simplicity of theoretical modeling with the long-term goal of improving the efficiency of the design of subwavelength periodic photonic structures.

1

Introduction

In this thesis, our work is in the field of nanophotonics [1–3]. Nanophotonics is concerned with light-matter interactions at the nanoscale. The interactions can be between light and a single nanoparticle, or between light and a collection of nanoparticles that are either periodically arranged or nonperiodically arranged in a homogeneous or non-homogeneous medium. These nanoparticles and their background are usually part of a larger system we will call the universe. For the purpose of this thesis, we will divide the universe into two parts: (1) the nanophotonic system, and (2) the external universe. We define the nanophotonic system as the nanoparticle(s) together with the medium they are in. We define the external universe as the collection of everything else except the nanophotonic

system. In general, the nanophotonic system may or may not exchange energy with the external universe. If it does, we say the system is non-Hermitian (open); if not, we say it is Hermitian (closed).

The shape of the nanoparticles, their spatial arrangements, the material that makes up the medium, the material composition of the nanoparticles, as well as whether or not they are lossy, all determine the nature of the light-matter interaction. The complex light-matter interactions in nanophotonics can be modeled in two ways: (1) numerically and (2) theoretically. They are complementary to each other - numerical models can give us accurate predictions of the optical effects resulting from these light-matter interactions, while theoretical models help us gain physical insight into the interactions that lead to the predicted optical effects. The predicted optical effects often emerge from complex electromagnetic interactions. The complexity of these interactions increases when the number of particles and their complexity increases. This increases the demand on computational resources as well as the challenges for theoretical models to explain them.

Periodic subwavelength photonic materials are a subset of optical materials whose smallest feature size is smaller than the operating wavelength. They have attracted much interest in nanophotonics for technological applications, such as those reviewed in Halir et al. [4] and Cheben et al. [5], including waveguide couplers, sensors, and spectral filtering devices. In addition, they have applications connected to more fundamental physics, such as the imitation of spacetime geometries [6], black holes [7], and Hawking radiation [8] in the field of analogue gravity. Fig. 1.1 illustrates how the functionality of periodic photonic materials is fundamentally determined by the ratio of the period, a, to the operating wavelength, λ.

There are three main regimes: (1) $a >> \lambda$ (geometric optics), (2) $a \gtrsim \lambda$ (wave phenomena such as diffraction and Bragg scattering), and (3) $a < \lambda$ (subwavelength regime). In the geometric optics regime, the structure is larger than the operating wavelength. This is the realm of bulk devices such as mirrors and optical lenses. In the wave phenomena regime, the periodicity is smaller, down to the length scale of the operating wavelength. Here, we have photonic crystals, which are naturally occurring [9] as well as artificial [10] photonic periodic structures. They are responsible for the colour change of chameleons [11], peacock feathers [12], opals [13], as well as some butterfly wings [14]. As feature sizes shrink to below the wavelength, we are in the subwavelength regime where near-field light-matter interaction dominates, and diffraction is suppressed. The work of this thesis focuses on periodic photonic materials in the subwavelength regime. They can be engineered for a

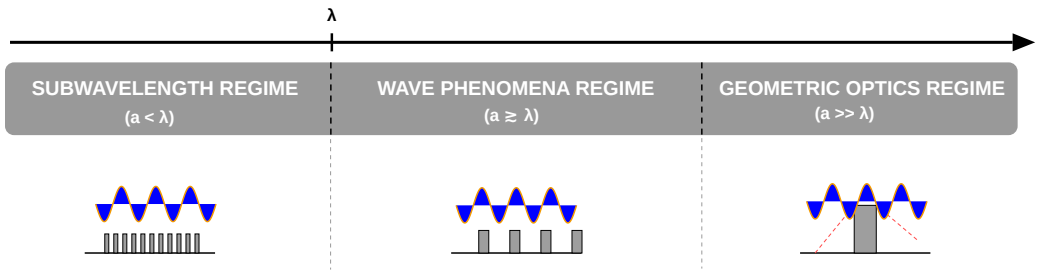


Figure 1.1: Schematic of periodic materials in the subwavelength (left), wavelength (middle), and geometric optics (right) regimes.

variety of purposes, in quantum computing [15–19], photonic neural networks [20], and photonic integrated circuits (PICs)/chips [21]. Engineering them for these applications often involves exploring their parameter spaces to find which designs are most suitable for

their application.

Though theoretical models can provide physical insights and rapid exploration of large parameter spaces, their assumptions are often too idealistic, limiting their ability to describe and design realistic devices. As a result, brute-force electromagnetic simulations are often performed. However, performing full electromagnetic simulations on a large parameter space requires significant compute, making the design process lengthy and slow. An important question to ask, then, is: how might we get the conceptual clarity and rapid exploration capabilities that theoretical models can provide, with the realistic structures that computational models provide, with neither the high compute cost of computational models nor the overly simplistic assumptions of theoretical models? The goal of this thesis is to establish a potential road map to achieve this balance.

In order to establish this road map, we explore two photonic periodic systems: subwavelength gratings (SWGs) embedded in vertical grating couplers, and topological plasmonic nanoresonator chains. These two artificially structured materials are of interest in this thesis due to their significant potential for technological applications and due to the significant differences in their inherent nature: the SWGs studied are made of silicon and silica, whereas the plasmonic system of interest consists of metallic nanoresonators. SWGs consist of a periodic arrangement of two or more alternating dielectric materials, with structural features of subwavelength scale [22–26]. Plasmonic nanoresonator topological lattices are near-field-coupled periodic structures whose size and spacings are subwavelength. The resonances of these metallic lattices are a result of the hybridization of resonances from their metallic constituents. Despite having physical differences, both SWGs and plasmonic

lattices share a common design challenge: their optical response depends nonlinearly on parameters of a high-dimensional design space. Analytical approximations provide insight but are often severely limited when designing devices. Simulations of the entire structure are accurate but computationally expensive. This thesis develops computational and semi-analytic frameworks to design photonic metamaterials, spanning both dielectric subwavelength gratings and plasmonic lattices that can host non-Hermitian topology. Using a data-driven effective medium approximation and simulations for SWGs, along with semi-analytic tools and eigenmode simulations of plasmonic systems, we demonstrate new ways of designing and probing periodic photonic materials.

One device where this might be suitable is in the design of PICs. PICs often have to communicate with other components of integrated photonic devices over short ranges [21, 27]. Oftentimes, this is done by coupling a standard single-mode fibre with waveguides of size on the order of hundreds of nanometers [28, 29]. However, single-mode fibres have mode-field diameters much larger than the size of the waveguides, resulting in significant challenges for efficient coupling. One way that has been developed to address this is the use of edge grating couplers. These couple with the fibre in-plane and have high coupling efficiency. When using edge grating couplers, the fibre is aligned with the chip facet in such a way that it couples horizontally with the waveguide. This makes wafer testing difficult, leads to poor fabrication tolerances, and also imposes significant restrictions on the placement of components [30]. As a result, edge grating [29] couplers are not the most suitable components for making compact integrated photonic devices. Another more suitable component is a vertical grating coupler [31, 32]. These couple out-of-plane with the fibre and

can be placed at arbitrary positions on a chip. This makes them easier to fabricate and makes wafer testing easier [30]. Another reason for their appeal is their high compatibility with widely available integrated photonic device components such as multi-core fibres and lasers. However, vertical grating couplers have lower coupling efficiency compared with edge couplers [31, 32]. Thus, improving the coupling efficiency of vertical grating couplers is an important design challenge, given their significant advantages. Designing them to optimize for coupling efficiency is an important step toward achieving this. Vertical grating couplers incorporating subwavelength gratings (SWGs) [4, 33] are attracting attention, as they have complex design features with the potential to have high coupling efficiency [34–36].

SWGs were first applied to control the effective index of waveguide cores in [37] and can be engineered to have a variety of optical properties tailored to the needs of the designer. They have been used in couplers [38, 39], polarization splitters [40, 41], as well as optical time-delay lines [42]. When optimizing SWGs, multiple design iterations are often explored, which is resource-consuming. Meanwhile, finite-size effects render traditional effective medium theories (EMTs) unreliable. A more efficient approach is needed to address these issues. Recently, data-driven approaches using artificial neural networks (ANNs) have been used for design in photonics [43–46]. In one work [47], ANNs were used to design subwavelength gratings for particle trapping, while in [48], they were used to design optical filters. They have also been used to enhance light-matter interactions [49–52], as well as to design grating couplers [53, 54]. In this thesis, we leverage ANNs, in particular, deep neural networks (DNNs), in the design of metamaterial gratings. We

build upon the work of Melati et al. [54] to reduce the design space of the metamaterial and build an effective medium approximation using DNN surrogate models.

Sometimes, these artificially engineered optical materials can host interesting optical phenomena. One such phenomenon is the existence of photonic topological states, which have been known to exist for over a decade [55–57]. They obey the bulk-boundary correspondence, which tells us that we can infer the existence of backscatterless boundary states from the system's bulk properties. These states have various technological applications, from robust optical delay lines [58], lasing [59, 60], high-density data transmission [61] to programmable chips [62, 63]. In general, topological systems can be divided into two types: Hermitian (they do not exchange energy with the rest of the universe) and non-Hermitian (they exchange energy with the rest of the universe). While Hermitian systems are interesting in their own right, there is a more diverse array of topological effects in non-Hermitian systems that never appear in Hermitian systems. One example is the non-Hermitian skin effect [64], where all the eigenstates of a system containing asymmetric hopping are spatially localized at the boundary. Another example is exceptional points [65], which are degeneracies where both the eigenvalues and eigenvectors coalesce. These have technological applications such as sensing [66–68], lasing [60, 69, 70], and wavefront shaping [71].

In photonics, non-Hermiticity naturally appears in plasmonics, where we focus on light-matter interactions at metal-dielectric interfaces. Since metals are lossy, they are excellent platforms to study non-Hermitian topological effects. Some work has already been done in this area. As with Hermitian systems, topological plasmonics theory considered nearest-

neighbour couplings first [72], with long-range and radiative effects included later [73]. Much of the theory has focused on implementing the dipole approximation, which is only relevant for resonators small enough compared to the wavelength of interest that they can be approximated as point-like. The dipole approximation neglects such as quadrupolar interactions and beyond. As a result, these approaches are significantly limited, especially if one were to use them for device design, as they are not able to handle larger particles sizes with arbitrary shapes.

This thesis focuses on topological phases in classical electromagnetism, specifically in arrays of near-field coupled metallic nanoresonators. We study these systems using a formalism that treats the resonances of the constituent plasmonic nanoresonators as basis states for many-particle systems [74]. Our approach was similar in a manner similar to the formalism of non-Hermitian quantum mechanics presented in Moiseyev [75], which is applicable, for example, to resonant phenomena, where particles are trapped temporarily in a potential. The resonances we focus on in this thesis are classical resonances known as quasinormal modes (QNMs). As the optical properties of these systems depend on the size, shape, spacing, and material compositions of their substructures, one therefore deals with large design spaces. A large design space has advantages, such as providing flexibility and diversity in technological applications, allowing applications to be tailored to specific needs. There are also disadvantages, however, as a large design space also means higher compute demands during optimization.

This thesis is organized as follows. In Chapter 2, we discuss the general subject of periodic photonic materials, introduce the concept of neural networks and how they have

been used in designing photonic materials, review the relevant background material on Hermitian and Non-Hermitian topological phases, and discuss photonic quasinormal modes (QNMs). In Chapter 3, we begin by reviewing the numerical tools that were essential for this thesis, mainly the finite-difference time-domain (FDTD) method and the finite element method (FEM). We also discuss the tools we used to train and develop our deep neural network surrogate models. We show how we use these surrogate models to develop our data-driven effective medium approximation. Finally, we end Chapter 3 with a more detailed discussion on quasinormal modes as well as the coupled-mode theory for quasinormal modes and how they apply to topological plasmonics. The main novel contributions of this thesis are presented in Chapters 4, 5 and 6. In Chapter 4, we present the work on our data-driven effective medium approximation for SWGs. In Chapter 5, we present the work on a QNM coupled-mode formalism for topological plasmonics. In Chapter 6, we present the QNM simulation work on topological plasmonics. Chapter 7 presents preliminary work on simulating topological plasmonic chains consisting of resonators on substrates. Chapter 8 outlines preliminary theoretical developments regarding my own interpretations of non-Hermitian Berry phases. Finally, in Chapter 9, we provide a conclusion and suggestions for future work.

2

Background

In this chapter, we discuss the background relevant to the numerical and theoretical tools that are developed in the thesis. In Section 2.1, we discuss machine learning, how it has been used in photonics, and how it will be used in this thesis. In Section 2.2, we discuss the concept of Bloch modes in condensed matter systems. The purpose of discussing Bloch modes in condensed matter systems is due to the fact that the concepts have been developed and established there. Next, in Section 2.3, we discuss how the concept of Bloch modes extends to photonic systems, where we then show the analogy between condensed matter and photonic periodic systems. In Section 2.4, we introduce SWGs. In Section 2.5 and 2.6, we discuss topological systems. As topological photonics concepts have been extensively

drawn from condensed matter physics, we first introduce the subject in condensed matter systems in Section 2.5, where we discuss the Hermitian and non-Hermitian cases. Next, we introduce topological photonics in Section 2.6, particularly for Hermitian systems, and compare the eigenvalue equations of the Hermitian condensed matter and photonic systems to help the reader understand how some relevant concepts transfer from condensed matter systems to periodic electromagnetic media. Then, in Section 2.7, we discuss the Drude model and its relevance to this thesis. Finally, in Section 2.8, we introduce quasinormal modes, which we use for our work on topological plasmonic systems.

2.1 Machine Learning for Photonics

Machine learning is playing a role in the design of periodic photonic materials and integrated photonic devices. Data-driven strategies to design periodic photonic materials are changing the design landscape from an intuitive, trial-and-error system to one that is more guided. As pointed out in [76], electromagnetic solvers have enormous design spaces, requiring significant computational resources and time. Solutions are often non-unique and involve probing high-dimensional nonlinear parameter spaces. These kinds of problems can be handled effectively with machine learning. Data-driven strategies are less time- and resource-consuming when deployed. They also have faster design cycles and can incorporate fabrication constraints well. They play an important role in accelerating the discovery and design of next-generation photonic devices. They can be quite powerful in performing tasks such as inverse design, where geometric parameters of a photonic material are ad-

justed until a desired optical performance is met. An important tool in machine learning is an artificial neural network (ANN). We discuss ANNs in more detail in Section 2.1.2.

There are many different kinds of ANNs: convolutional neural networks (CNNs), recurrent neural networks (RNNs), transformers, and generative adversarial networks (GANs) [77]. Convolutional neural networks (CNNs) are very good at extracting features from images and have been used, for example, to optimize the Q-factors of photonic crystals [78]. Recurrent neural networks (RNNs) are typically used to do feature extraction from sequential data and have been used, for example, to obtain the absorption spectra of plasmonic structures. We point the reader to some general books on machine learning [79–87]. There are also quite a few reviews on the application of machine learning to photonics [88–91].

2.1.1 Perceptron

Artificial neural networks are computational models that are inspired by how humans learn [83]. Humans have an advanced nervous system that consists of components known as neurons, which release electrical signals in response to stimuli such as touch. These electrical signals travel to the brain. The brain's neurons repeatedly interact with each other through these electrical signals. When learning a new task, we are repeatedly exposed to a stimulus. This repeated exposure causes the neurons to fire their electrical signals many times along multiple pathways. Over time, certain pathways carry stronger signals than others. They become the dominant pathways, which allow us to store and recall

information more efficiently. This motivated people to develop and explore artificial neural networks (ANNs).

A single unit of an ANN is known as a perceptron (or equivalently, an artificial neuron), as it is inspired by biological neurons. Fig. 2.1 shows a schematic of a perceptron. It accepts inputs $(x_1, x_2, x_3, ..., x_n)$, which are multiplied by corresponding weights $(w_1, w_2, w_3, .., w_n)$. The weights determine the importance of the features of the input. The weighted input values are summed (Σ), and a bias, b, is added as a threshold adjustment factor. The resulting value, $z = \sum_{j=1}^{N} x_j w_j + b$, is then passed to an activation function, $f(z)$, which returns the output value, y. The role of the activation function is to introduce nonlinearity, and it determines whether an input signal makes it to the output. One way to define the

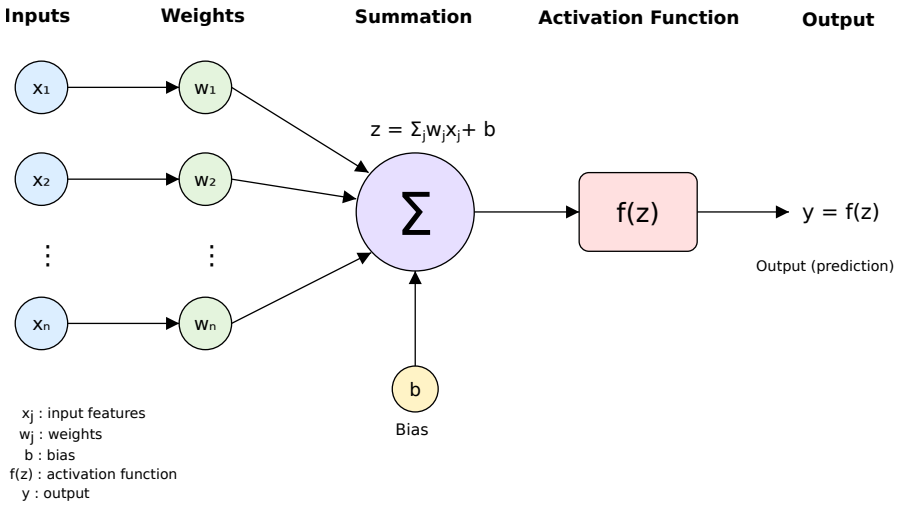


Figure 2.1: Illustration of a simple artificial neuron - a perceptron. From the left, we have input features, $x_1, ..., x_n$, which are data that go into the perceptron. The weights, $w_1, ..., w_n$, assign importance to input features. The bias, b, is a threshold for the activation function, $f(z)$, which determines if a signal makes it to the output, y.

output is as a Heaviside step function

$$f(z) = \begin{cases} 1 & \text{if } z > 0, \\ 0 & \text{otherwise.} \end{cases} \tag{2.1}$$

One can combine many of these to form an artificial neural network (ANN).

2.1.2 Artificial Neural Networks

Perceptrons [83] can be combined in many different ways to form different kinds of neural networks. Fig. 2.2 shows an example of an ANN. Each circle region corresponds to a perceptron, and the lines between them represent their connectivity. Notice that the outputs of one are the inputs of others, going from left to right in the diagram. The connectivity strengths between the perceptrons are represented by weights, which we discussed in Section 2.1.1. The larger the weight between two perceptrons, the stronger the connectivity between the two perceptrons. We can have multiple layers: an input layer, hidden layers, and an output layer. Each layer can have multiple neurons. The input layer accepts the input data; the middle layers adjust their weights while features are being learned, analogous to how biological neurons in the brain adjust their connectivity as we learn. The output layer generates the predictions of the ANN. For example, if the ANN was designed to classify objects, the output layer is what tells us the classification. When we have two or more hidden layers, we call this a deep neural network (DNN). It is this kind of neural network that we use in our work.

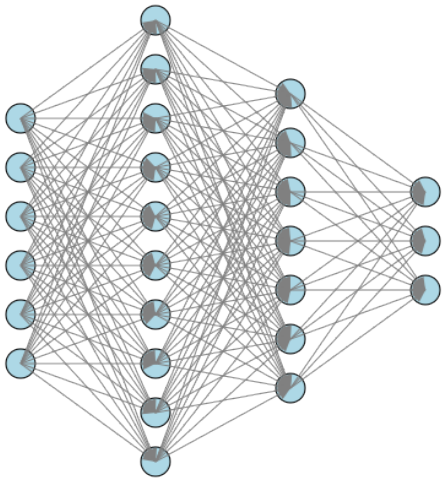


Figure 2.2: Example of an artificial neural network made by combining multiple perceptrons.

2.1.3 Supervised and Unsupervised Learning

In machine learning, we can assign labels to a set of input data that tell the machine what the correct output should be, and this can be used for training (that is, adjusting the weights). This is called supervised learning. For example, suppose we build a machine learning model that classifies images of cats and dogs. Training images with cats would be labeled as cats, and training images with dogs would be labeled as dogs. The machine eventually maps features of cats to cat and features of dogs to dog. Then, when one gives it images that are cats or dogs, it outputs the answer as cat or dog as its classification. Supervised learning is typically used when one wants to predict unknown outputs, Y, from a set of inputs, X, that is not contained within the training data. There are various tools

for this, such as linear regression, logistic regression, and ANNs [83], all of which can be trained to map X to Y. In unsupervised learning, the input training data is unlabeled, meaning that the output corresponding to the input is unknown. The algorithm focuses instead on learning features in a dataset and grouping them based on common features. Tools that perform this include clustering tools such as K-means clustering [79, 92], and dimensionality reduction tools such as principal component analysis (PCA) [79, 92]. In our work, we use supervised learning, and our main tool is the artificial neural network (ANN), which we discussed in Section. 2.1.2.

2.1.4 Training an ANN

Every ANN must be trained so that it learns features of a dataset that would give it predictive power. The training stage is very important; the better the ANN is trained, the better it is at providing accurate predictions. During this stage, one requires two datasets: a training dataset and a validation dataset. The training data is used to optimize the weights and biases and the validation data is what is used to see how well the model generalizes. If T is the training dataset and V is the validation dataset, it is better if $T \cap V = \emptyset$. Choosing a good training dataset is of high importance, as it significantly influences predictive power. If the training dataset is not a diverse representative sample of the kind of objects/features you wish to predict, then the ANN will perform poorly. A very large dataset can provide a diverse sample size, but increasing the size of the dataset does not necessarily improve the performance of the ANN. Rather, it is choosing

the minimum sample size that is a sufficient representation of the population of features that are important for predictive performance.

During the training of our DNN-based surrogate models in Chapter 4, we take training data, $S_i \subset T$, called mini-batches which are selected randomly. Doing it this way, is typically faster than using the entire dataset at once. At each training step, one [83]: (1) selects a mini-batch, S_i, (2) pushes it through each layer of the ANN and performs a prediction, (3) compares this prediction with the true values, by computing an error (loss, L) using for example, a mean-square-error function (4) perform backpropagation, where it works back through the network and computes how each weight contributed to the error by computing the gradients of the loss, then (5) it updates the weights appropriately, to reduce the error computed in step 3, using an update computed from gradient descent

$$w_{\text{new}} = w_{\text{old}} - \eta \frac{\partial L}{\partial w}, \tag{2.2}$$

where w_{new} and w_{old} is the new and old values of the weights, L is the loss computed in step 3, and η is the learning rate/step-size. It repeats these steps until the error computed in step 3 is minimized. There are smarter update rules than Eq. 2.2, such as adaptive moment estimation (Adam) and root mean square propagation (RMSProp) optimizers [83]. Standard gradient descent uses a single learning rate for an entire model, whereas Adam and RMSProp customize the learning rate. RMSProp adjusts the learning rate of the model based on the gradient history. It does this in such a way that parameters with larger gradients are assigned a smaller learning rate, while parameters with smaller

gradients are assigned a larger learning rate. Adam modifies the RMSProp approach by including gradient averaging and a bias correction [83].

2.2 Bloch Modes in Condensed Matter Systems

The concept of Bloch modes originated in condensed matter physics, describing electron behaviour in solids. Periodic solids in condensed matter physics are represented by a periodic lattice. A periodic lattice consists of a unit cell that repeats itself. In condensed matter physics, the electron behaviour in solids is governed by quantum mechanics. As a result, we start our discussion on solids with the Schrödinger equation

$$\hat{H}\left|\psi_n(\vec{r})\right\rangle = E_n\left|\psi_n(\vec{r})\right\rangle, \tag{2.3}$$

where $-\frac{\hbar^2}{2m}\nabla^2 + V_{\text{lat}}(\vec{r}) = \hat{H}$, m is the electron mass, $V_{\text{lat}}(\vec{r})$ is the lattice potential which is periodic [93, 94], $|\psi_n(\vec{r})\rangle$ is the eigenstate and E_n is the energy associated with the n^{th} eigenstate. We impose periodic boundary conditions and write the wavefunctions in terms of their Bloch modes

$$\left|\psi_{n\vec{k}}(\vec{r})\right\rangle = e^{i\vec{k}\cdot\vec{r}}\left|u_{n\vec{k}}(\vec{r})\right\rangle, \tag{2.4}$$

where $\vec{k}$ is the crystal momentum/wavevector in reciprocal space (also called k-space or the Brillouin zone) and the Bloch mode, $|u_{n\vec{k}}(\vec{r})\rangle$, satisfies

$$\left|u_{n\vec{k}}(\vec{r}+\vec{R})\right\rangle = \left|u_{n\vec{k}}(\vec{r})\right\rangle. \tag{2.5}$$

Substituting this into the Schrödinger equation gives [93, 94]

$$\hat{H}(\vec{k})\,|u_{n\vec{k}}(\vec{r})\rangle = E_n(\vec{k})\,|u_{n\vec{k}}(\vec{r})\rangle\,. \tag{2.6}$$

The eigenvalues, $E_n(\vec{k})$, tell us what energies a single active electron in this lattice is allowed to have. All other energies are forbidden. One typically plots the allowed eigenvalues as a function of the wavevector $\vec{k}$, giving the bandstructure which characterizes bulk properties. A particular curve of the bandstructure for a given lattice tells us that electrons are allowed to propagate with that energy through the lattice. When there is a gap in the bandstructure, it means that there is a range of energies for which there is no corresponding k, meaning electrons cannot exist at that energy. The occupation levels of the bands determine the electronic properties of the system, for example, whether it is a semiconductor, conductor, or insulator.

Sometimes an insulator might be hiding additional properties, such as boundary conducting states, where conducting electrons are localized along the boundary. Insulators with such a property are said to obey bulk-boundary correspondence, as their (bulk) bandstructure can in fact reveal the presence (or absence) of these boundary states. These insulators are called topological insulators, and they belong to a class of materials known as topological materials. Topological materials will be introduced in Section 2.5 for condensed matter systems.

2.3 Photonic Bloch Modes

Eli Yablonovitch [95] and Sajeev John [96] realized that Maxwell's equations in periodic media had behaviour similar to electrons in solids, indicating the presence of Bloch modes in periodic photonic systems.

Let us consider the source-free Maxwell equations, which are given by [97]:

$$\nabla \cdot \vec{D}(\vec{r}, t) = 0, \tag{2.7a}$$

$$\nabla \cdot \vec{B}(\vec{r}, t) = 0, \tag{2.7b}$$

$$\nabla \times \vec{H}(\vec{r}, t) = \frac{\partial \vec{D}(\vec{r}, t)}{\partial t}, \tag{2.7c}$$

$$\nabla \times \vec{E}(\vec{r}, t) = -\frac{\partial \vec{B}(\vec{r}, t)}{\partial t}, \tag{2.7d}$$

where $\vec{E}(\vec{r}, t)$ is the electric field, $\vec{D}(\vec{r}, t) = \epsilon \vec{E}(\vec{r}, t)$ is electric displacement vector in linear media, $\vec{B}(\vec{r}, t) = \mu_0^{-1} \vec{H}(\vec{r}, t)$ is the magnetic flux density for non-magnetic materials, and $\vec{H}(\vec{r}, t)$ is the magnetic field strength. ϵ and μ_0 are the electric permittivity and magnetic (free-space) permeability. We can write $\vec{E}(\vec{r}, t), \vec{H}(\vec{r}, t)$ as harmonic fields

$$\vec{E}(\vec{r}, t) = \vec{E}(\vec{r}) e^{i\omega t}, \tag{2.8a}$$

$$\vec{H}(\vec{r}, t) = \vec{H}(\vec{r}) e^{i\omega t}, \tag{2.8b}$$

where ω is the frequency associated with the field distributions $\vec{E}(\vec{r})$ and $\vec{H}(\vec{r})$. Then we

obtain

$$-i\mu_0^{-1}\nabla \times \vec{E}_n(\vec{r}) = \omega_n \vec{H}_n(\vec{r}), \tag{2.9a}$$

$$i\epsilon_0^{-1}\epsilon(\vec{r})^{-1}\nabla \times \vec{H}_n(\vec{r}) = \omega_n \vec{E}_n(\vec{r}), \tag{2.9b}$$

which can be written as

$$\hat{H}\left|\psi_n(\vec{r})\right\rangle = \omega_n \left|\psi_n(\vec{r})\right\rangle, \tag{2.10}$$

where

$$\left|\psi_n(\vec{r})\right\rangle = \begin{bmatrix} \vec{E}_n(\vec{r}) \\ \vec{H}_n(\vec{r}) \end{bmatrix}, \tag{2.11}$$

and

$$\hat{H} = \begin{bmatrix} 0 & i\epsilon_0^{-1}\epsilon(\vec{r})^{-1}\nabla\times \\ -i\mu_0^{-1}\nabla\times & 0 \end{bmatrix}. \tag{2.12}$$

This is the eigenvalue problem for dispersionless photonic systems. For periodic photonic systems, $\epsilon(\vec{r}+\vec{R}) = \epsilon(\vec{r})$. We can write the Bloch form in reciprocal space

$$\begin{bmatrix} 0 & i\epsilon_0^{-1}\epsilon(\vec{r})^{-1}(\nabla + i\vec{k})\times \\ -i\mu_0^{-1}(\nabla + i\vec{k})\times & 0 \end{bmatrix} \begin{bmatrix} \vec{u}_{n\vec{k}}(\vec{r}) \\ \vec{v}_{n\vec{k}}(\vec{r}) \end{bmatrix} = \omega_n \begin{bmatrix} \vec{u}_{n\vec{k}}(\vec{r}) \\ \vec{v}_{n\vec{k}}(\vec{r}) \end{bmatrix}. \tag{2.13}$$

Table 2.1 shows an analogy between the photonic and condensed matter eigenvalue

problems for periodic lattices. We see from the table, for example, that the components of the eigenstates for photonic systems are vectors, whereas for electronic systems, the components are scalars. The eigenvalues for the photonic system are angular frequencies, while for the condensed matter systems, they are energies. We also see that the Hamiltonians for the photonic setup are matrix-like and depend on the curl operator, whereas the condensed matter Hamiltonians are scalar-like and depend on the Laplacian operator. Notice that we can also express the photonic eigenproblem in Bloch form.

Concept	**Photonic System**	**Condensed Matter System**
State	$\|\psi_n(\vec{r})\rangle = \begin{bmatrix} \vec{E}_n(\vec{r}) \\ \vec{H}_n(\vec{r}) \end{bmatrix}$	$\|\psi_n(\vec{r})\rangle$
Eigenvalue Equation	$\hat{H}\|\psi_n(\vec{r})\rangle = \omega_n \|\psi_n(\vec{r})\rangle$	$\hat{H}\|\psi_n(\vec{r})\rangle = E_n \|\psi_n(\vec{r})\rangle$
Real Space Hamiltonian	$\hat{H} = \begin{bmatrix} 0 & i\epsilon_0^{-1}\epsilon(\vec{r})^{-1}\nabla\times \\ -i\mu_0^{-1}\nabla\times & 0 \end{bmatrix}$	$\hat{H} = -\frac{\hbar^2}{2m}\nabla^2 + V_{\text{lat}}(\vec{r})$
Eigenvalues	ω_n (frequency)	E_n (energy)
Potential	$\epsilon(\vec{r}), \mu(\vec{r})$	$V(\vec{r})$
Differential Operator	$\nabla\times$ (curl)	∇^2 (Laplacian)
Underlying Equation	Maxwell equations	Schrödinger equation
Periodic Medium	Photonic crystal/metamaterial	Atomic crystal
Band Structure	$\omega_n(\vec{k})$	$E_n(\vec{k})$
Band Gap Origin	Periodic dielectric contrast	Periodic lattice potential

Table 2.1: Analogy between photonic and condensed matter eigenproblems. The first column is the concept, while the second and third columns are the photonic and condensed matter setups, respectively.

2.4 A Brief Discussion on SWGs

One of the attractive features of subwavelength periodic structures is that they allow for the detailed engineering of the optical properties of materials. The relevant optical property engineered is the effective refractive index of the structure. Different methods have been developed to approximate the effective refractive index. One such class of tools is called effective medium theories (EMTs) [98, 99]. They have traditionally been used to approximate the effective refractive index of the subwavelength gratings (SWGs), which is then used to build a slab model.

In an SWG [22–26, 100], such as those depicted in Fig. 2.3, waves can propagate so that they are longitudinal (left) or perpendicular (right) to the propagation direction. In addition to the material makeup, the period (Λ), and thickness (a) of the ridges determine how the SWG interacts with light. More complex designs than Fig. 2.3 are also possible [31]. SWGs are important structures for integrated photonics. Understanding their optical behaviour is important for these applications.

Effective medium theories (EMTs) were developed to explain how macroscopic optical effects emerge from a collection of substructures that are spatially separated. One approach was invented by Garnett [101], leading to the Maxwell-Garnett formula, which was developed to understand how permittivity arises from microscopic polarizations. Another approach developed by Sergei Rytov used Maxwell's equations and Bloch waves to describe subwavelength periodic media as a homogeneous effective medium, with a

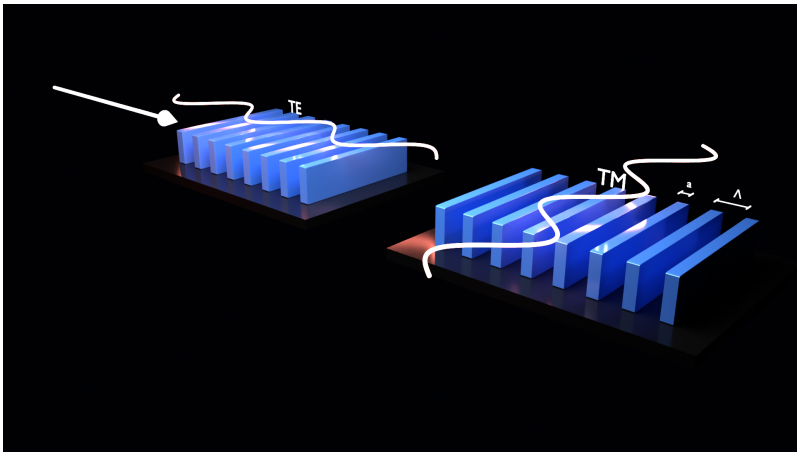


Figure 2.3: Schematic of a simple subwavelength grating design. The arrow represents the direction of propagation. Waves are oriented parallel (left, TE modes) and perpendicular (right, TM modes) to the orientation.

polarization-dependent effective refractive index [102]. Recall that for periodic structures, the position-dependent permittivity, $\epsilon(x)$, satisfies

$$\epsilon(x + \Lambda) = \epsilon(x), \tag{2.14}$$

where Λ is the period of the structure. The y-component of the electric field, $E_y(x)$, can be written as

$$E_y(x) = u_y(x)e^{ikx}, \tag{2.15}$$

where $u_y(x)$ is a Bloch mode with period Λ. In the long-wavelength limit [103],

$$k\Lambda \ll 1, \tag{2.16}$$

giving

$$e^{ikx} \approx 1. \tag{2.17}$$

To zeroth order, the periodic part of the Bloch function is constant, so we have

$$E_y(x) = u_y(x)e^{ikx} \approx E_y. \tag{2.18}$$

This now allows us to derive the TE mode expression:

$$\begin{aligned} \langle D_y \rangle &= \frac{1}{\Lambda} \int_0^{\Lambda} D_y(x)\, dx, \\ &= \frac{1}{\Lambda} \left[\int_{\mathrm{M1}} \epsilon_1 E_y(x)\, dx + \int_{\mathrm{M2}} \epsilon_2 E_y(x)\, dx \right], \end{aligned} \tag{2.19}$$

where M1 (of permittivity, ϵ_1) and M2 (of permittivity, ϵ_2) represent the two materials of the periodic structure. If we set $f = a/\Lambda$, we have

$$\langle D_y \rangle = \left[f\epsilon_1 + (1-f)\epsilon_2 \right] E_y, \tag{2.20}$$

which means that

$$\epsilon_{\text{eff}}^{\text{TE}} = f\epsilon_1 + (1-f)\epsilon_2. \tag{2.21}$$

In the case of TM modes (right image of Fig. 2.3),

$$\begin{aligned}\langle E_x \rangle &= \frac{1}{\Lambda}\int_0^{\Lambda} E_x(x)\,dx, \\ &= \frac{1}{\Lambda}\left[\int_{\text{M1}} \frac{D_x(x)}{\epsilon_1}\,dx + \int_{\text{M2}} \frac{D_x(x)}{\epsilon_2}\,dx\right].\end{aligned} \tag{2.22}$$

To zeroth order, we have (similarly to above)

$$D_x(x) \approx D_x. \tag{2.23}$$

This gives

$$\begin{aligned}\langle E_x \rangle &= \frac{1}{\Lambda}\left[\frac{D_x}{\epsilon_1}(f\Lambda) + \frac{D_x}{\epsilon_2}((1-f)\Lambda)\right], \\ &= \left[\frac{f}{\epsilon_1} + \frac{1-f}{\epsilon_2}\right] D_x,\end{aligned} \tag{2.24}$$

which means that

$$\epsilon_{\text{eff}^{\text{TM}}}^{-1} = f\epsilon_1^{-1} + (1-f)\epsilon_2^{-1}. \tag{2.25}$$

We will summarize the main equations above in a form familiar in many texts or papers.

We note that $n_1^2 = \epsilon_1$ and $n_2^2 = \epsilon_2$ and we note that above, we dropped the higher-order corrections. n_1 and n_2 are the refractive index of M1 and M2, respectively. In summary we have [102]

$$n_{\text{eff}}^{\text{TE}} = f n_1^2 + (1-f) n_2^2 + \mathcal{O}\left(\frac{\Lambda^2}{\lambda^2}\right), \tag{2.26a}$$

$$n_{\text{eff}}^{\text{TM}} = f n_1^{-2} + (1-f) n_2^{-2} + \mathcal{O}\left(\frac{\Lambda^2}{\lambda^2}\right), \tag{2.26b}$$

where λ is the wavelength. The fundamental approximation of the Rytov EMT (as well as other EMTs) is that the structures comprising the SWGs satisfy $\Lambda/\lambda \ll 1$. Another assumption is that the lowest Bloch mode dominates, while higher Bloch modes have negligible contributions. In addition, the structure must be locally periodic and not be strongly dependent on the wavevector.

Traditionally, an EMT such as the Rytov approximation, Eq. 2.26a-Eq. 2.26b, approximates the effective refractive index. The entire grating region of Fig. 2.3 is then replaced with a homogeneous structure of effective refractive index (a process called homogenization), $n_{\text{eff}}^{\text{TE/TM}}$. This is used to build the SWG slab model.

2.4.1 Slab Model

The Rytov approximation allows us to obtain the effective refractive index of an SWG of infinite extent. However, for integrated photonic applications, SWGs are often fabricated within a planar geometry; that is, they have finite height confined to a slab. The Rytov

approximation (or other bulk EMTs) can be used to obtain an estimate of the bulk index of the SWG within the slab. This is then used within the dielectric slab model for a planar waveguide (surrounded above and below by uniform materials with a smaller refractive index) to obtain the effective mode index. In this way, one can determine the effective index of a finite-height SWG.

Consider the TE mode of the first (left) structure of Fig. 2.3. If we use the Rytov approximation, Eq. 2.26a tells us the effective refractive index, $n_{\text{eff}}^{\text{TE}}$, to use in the slab model in place of the SWG. Maxwell's wave equation for the TE mode in this homogeneous slab core is

$$\frac{d^2 E_y}{dx^2} + \left[k_0^2 (n_{\text{eff}}^{\text{TE}})^2 - \beta^2\right] E_y = 0, \tag{2.27}$$

where β is the propagation constant and k_0 is the wavenumber. Under appropriate boundary conditions, Eq. 2.27 gives the dispersion relation for the even TE modes

$$\sqrt{k_0^2 (n_{\text{eff}}^{\text{TE}})^2 - \beta^2} \tan\left[\frac{\theta}{2}\sqrt{k_0^2 (n_{\text{eff}}^{\text{TE}})^2 - \beta^2}\right] = \sqrt{\beta^2 - k_0^2 n_{\text{clad}}^2}, \tag{2.28}$$

which is a transcendental equation that can be solved numerically for β. Here, θ is the thickness of the core. In turn, one then computes the effective refractive index of a mode, n_{mode}

$$n_{\text{mode}} = \frac{\beta}{k_0}. \tag{2.29}$$

This slab model approach to determining the effective refractive index of SWGs works for many designs. However, it can be insufficient for others, particularly those with finite

length as well as finite height and/or those that are highly radiative. In particular, this is the case for vertical grating couplers with embedded SWGs [31]. In the diffractive regime, higher-order Bloch modes cannot be ignored. These higher-order dependencies reduce the advantages of EMTs, and for every structure with designs similar to these or even with added complexity, a different number of terms has to be kept.

Due to the limitations of standard EMTs, people rely on simulations of the full 3D structure with brute-force methods, such as the finite-difference time-domain (FDTD) method and/or the finite element method (FEM), both of which are described in Chapter 3. Such simulations are time- and resource-consuming. In this thesis, we explore developing an approach that extends the concept of effective medium to structures that are finite in the propagation direction as well as finite in height, with the intent of balancing the physical insights and rapid exploration capabilities of effective medium theories with the numerical accuracy achieved by computational models, without inheriting the computational cost. We achieve this by using deep neural network surrogate models that are trained on FDTD data. This enables our group to perform and experimentally validate accurate inverse design of vertical grating couplers containing SWGs with 2D rather than 3D FDTD simulations[1]. Our approach is outlined in more detail in Chapter 3.

[1]Qiang Wang, Norman Israel, Dan-Xia Xu, Yuri Grinberg, Martin Vachon, Shurui Wang and Lora Ramunno, "A Machine Learning-Based Framework for Pseudo-3D Optimization of Vertical Metamaterial Grating Couplers," manuscript in preparation.

2.5 Topological Materials in Condensed Matter Systems

Although periodic photonic materials with conventional optical properties have interesting applications that suffer from the design challenges mentioned above, when these structures have exotic properties, those design challenges get magnified due to the large parameter space. The reason is that in order to observe some exotic properties, we have to find the right parameters in an already large parameter space. Designing realistic devices that leverage these exotic optical properties is best done using efficient methods to explore the parameter space. The exotic properties we are interested in here are photonic topological states, particularly those in open systems.

The theoretical formalism of topological photonics borrows extensively from the theoretical developments in condensed matter physics. As a result, we build upon our discussion in Section 2.2 to develop the relevant ideas in condensed matter systems. Later, we will develop the relevant equations for classical electromagnetism that are analogous to the condensed matter equations and compare the two directly so as to demonstrate which set of relevant basic concepts transfer from condensed matter systems to photonic systems.

Topological materials can be thought of as a phase of matter whose behaviour is described by the global properties of the eigenstate(s) of the system. Examples of topological materials include topological insulators, topological semimetals, and topological superconductors [104]. However, the phase transitions associated with topological materials do not

fit into the Landau paradigm [104–106], traditionally used to study phase changes of matter. In the Landau paradigm, we have a parameter containing information about the local properties of the system (local order parameter) [104–106]. The local order parameter is used to describe symmetry-breaking mechanisms that lead to phase transitions such as solid to liquid to gas (and vice versa). The Landau paradigm allowed for the classification of solids according to geometric symmetries. Topological phase transitions go beyond this paradigm by focusing on singularities of their eigenstates that are stable under smooth deformations of the eigenstates of the system. A topological phase transition occurs when we cross one of these singularities. The topological phase transition therefore happens, not because of some symmetry-breaking mechanism, but because of a discontinuous deformation of the system's eigenstate. This way of thinking signaled a new paradigm in our understanding of phase transitions.

One class of topological materials of much interest is topological insulators [104–106]. Topological insulators are materials whose electronic behaviour appears at first to be similar to that of an insulator - i.e., their (bulk) bandstructure has a bandgap resembling that of an ordinary insulator. However, they possess backscattering-free conducting states on the boundary. The origins of this effect seem to rest within the wave-like nature of the electrons in the lattice. An important phenomenon revealed by these systems is bulk-boundary correspondence, where the bulk properties of the system (bandstructure) reveal the existence (or lack of) boundary states in infinite systems.

The quantum Hall effect [107] became an important demonstration of the concept of bulk-boundary correspondence. Analogues have since been discovered in photonics [108,

109], electric circuits [110, 111] and mechanical systems [112, 113]. The potential applications include lasing [60, 69], sensing [68], and computing [114]. David Thouless and collaborators explained the quantum Hall effect using topology [115]. A paradigmatic example is a topological insulator consisting of a chain of spins known as the Su-Schrieffer-Heeger (SSH) chain [116, 117], which we will discuss more in Section 2.5.3. Further, Pankratov et al. predicted [118] that a system of mercury telluride sandwiched between cadmium telluride is a topological insulator. This was experimentally confirmed in 2007 [119] and was the first experimental verification of topological insulators. In 2005, the quantum spin Hall insulator was predicted [120, 121]. This is a two-dimensional topological insulator that was experimentally discovered [122] at low temperatures in quantum wells within mercury telluride. Since their work, much progress has been made to understand topological materials and topological phase transitions [104, 123].

2.5.1 The Tight-Binding Model

In Section 2.2, we introduced the concept of Bloch modes in condensed matter physics. We had then introduced the concept of bandstructure which are obtained from these Bloch modes. Electronic bandstructures, are of significant importance in condensed matter systems in order to understand the electronic properties of solids, such as their conduction properties. However, they can prove quite difficult to calculate directly, often requiring significant computing power, even with advanced algorithms. As a result, people sometimes turn to simple models to explore some of the basic physics before using more complex

tools. The tight-binding (TB) model [124] provides a simple and useful framework for modeling electrons in crystalline materials. It captures the essential mechanism of band formation by assuming that particles remain largely localized (tight) around lattice sites while weakly tunneling between neighbouring sites. Within this approximation, the total wavefunction is expressed as a linear combination of localized basis functions centered at each lattice site. A convenient choice of basis is given by the Wannier functions, which are orthonormal and spatially localized. Unlike atomic orbitals, Wannier functions centered at different lattice sites form an orthogonal set, making them particularly suitable for lattice Hamiltonians.

Restricting attention to one eigenmode per lattice site, we expand the field operators as [125]

$$|\psi(\vec{r})\rangle = \sum_{\alpha} a_{\alpha} |\phi_{\alpha}(\vec{r})\rangle , \tag{2.30}$$

where $|\phi_{\alpha}(\vec{r})\rangle$ denotes a Wannier orbital localized at lattice site α, and a_{α} is the corresponding creation operator. As the Wannier states form a complete orthonormal basis, we have the completeness relation

$$\hat{I} = \sum_{\alpha} |\phi_{\alpha}(\vec{r})\rangle \langle\phi_{\alpha}(\vec{r})| . \tag{2.31}$$

As a result, we can write the TB Hamiltonian as

$$\hat{H}_{\mathrm{TB}} = \hat{I}\hat{H}\hat{I}. \tag{2.32}$$

Substituting the completeness relation into Eq. 2.32, we have

$$\hat{H}_{\mathrm{TB}} = \sum_{\alpha,\beta} t_{\alpha\beta} \left|\phi_\alpha(\vec{r})\right\rangle \left\langle\phi_\beta(\vec{r})\right| , \tag{2.33}$$

where the tunneling (hopping) amplitudes that quantify the tunneling between localized states at sites α and β are given by

$$t_{\alpha\beta} = \left\langle\phi_\alpha(\vec{r})\right| \hat{H} \left|\phi_\beta(\vec{r})\right\rangle . \tag{2.34}$$

If we separate the diagonal terms out and consider onsite energies, $t_{\alpha\alpha} = \bar{E}_\alpha$, Eq. 2.33 becomes

$$\hat{H}_{\mathrm{TB}} = \bar{E}_\alpha + \sum_{\alpha\neq\beta} t_{\alpha\beta} \left|\phi_\alpha(\vec{r})\right\rangle \left\langle\phi_\beta(\vec{r})\right| . \tag{2.35}$$

In practice, the tight-binding model is often simplified by retaining only on-site energies and nearest-neighbor hopping terms. Despite this truncation, the model captures the essential features of bandstructure formation and provides a natural starting point for the study of topological phases in periodic systems. We will apply this in Section 2.5.3, where we review the SSH model.

2.5.2 Hermitian Topological Materials

Much of the work on understanding topological insulators has been done on Hermitian systems. Hermitian systems refer to those systems that do not exchange energy with the rest of the universe. Their Hamiltonians are self-adjoint, i.e., $\hat{H}^\dagger = \hat{H}$, resulting in real eigenvalues. Topological materials have been widely studied in these systems; however, the field is still quite active [104–107, 120, 121]. Reviews on this were already presented in the introduction to Section 2.5.

Recall that the bandstructure contains information about the behaviour of electrons in the material bulk. Also recall that this information can be captured by an effective Hamiltonian of the system. Hence, the effective Hamiltonian is a reasonable starting point to investigate topology in periodic lattices. Changing the parameters of a Hamiltonian can cause it to cross a region (a point, line) in the Brillouin zone where it becomes undefined. We will call these undefined regions defects. These defects cannot be removed using a smooth deformation of the Hamiltonian. Deforming the Hamiltonian after a certain point results in it crossing one of these defects, which amounts to two Bloch bands closing. Once the bands open again, the topological phase of the system changes. Hence, crossing the defect amounts to a topological transition quantified by a non-zero change in a quantity known as a topological invariant, which is computed from the eigenstates associated with the Bloch bands of interest.

Topological materials have a bulk-boundary correspondence, where the bulk properties (topological invariant) are related to boundary modes. In the case of topological insulators,

the boundary modes are defect-immune backscatterless states localized at the boundary of the lattice. These boundary states appear when the topological invariant has a non-zero value, and disappear when the topological invariant is zero.

Theoretical models make these ideas more explicit. For a Hermitian TB Hamiltonian, it is known that the hopping term is real, i.e.

$$t_{\alpha\beta} = t^*_{\beta\alpha}. \tag{2.36}$$

Substituting Eq. 2.30 into Eq. 2.35, we get

$$\hat{H} = \bar{E}_\alpha \delta_{ij} + \sum_{\alpha,\beta} \left(t_{\alpha\beta} a^\dagger_\alpha a_\beta + t_{\beta\alpha} a^\dagger_\beta a_\alpha \right). \tag{2.37}$$

Eq. 2.37 can be used to model the 1D SSH chain [126], where the bulk-boundary correspondence comes from the existence of chiral and inversion symmetry

2.5.3 Su-Schrieffer-Heeger Model

We illustrate how topology emerges in the Hermitian Su-Schrieffer-Heeger (SSH) model [126], noting that many of these ideas extend to non-Hermitian systems. An early example of the SSH chain is the polyacetylene chain [127]. Fig. 2.4 is a depiction of the polyacetylene chain. Its structure generally consists of carbon and hydrogen atoms, but only the carbon atoms are depicted in Fig. 2.4, as they are the ones that form the chain and are represented by black points. The carbon atoms consist of alternating double and single bonds that

represent intercell, w, and intracell, v, hoppings, respectively. The unit cells of the lattice

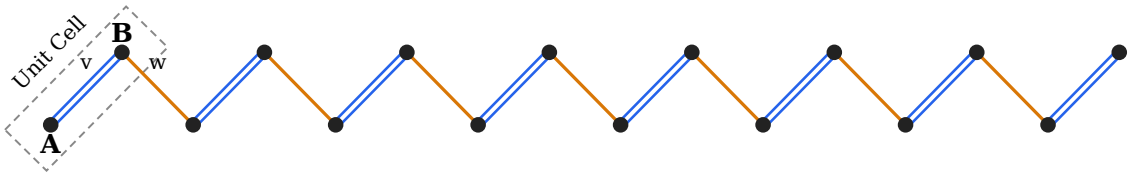


Figure 2.4: SSH chain diagram of polyacetylene lattice. w and v are the intercell and intracell hoppings, respectively. They alternate through the chain.

are indexed by i, each containing two orbitals (or sublattice sites), labeled A and B. We restrict ourselves to nearest-neighbour hopping. We assume the on-site energies are the same for A and B. This gives us the real-space SSH Hamiltonian

$$H = \bar{E}\delta_{jl} + \sum_{i=1}^{L} \left(v\, a^{\dagger}_{A,i} a_{B,i} + w\, a^{\dagger}_{B,i} a_{A,i+1} + \text{h.c.} \right), \tag{2.38}$$

where L is the number of unit cells and h.c. denotes the Hermitian conjugate.

Assuming periodic boundary conditions, we perform the Fourier transform to momentum space

$$a_{\alpha,i} = \frac{1}{\sqrt{L}} \sum_{k} e^{ikx_i} a_{\alpha,k}, \tag{2.39}$$

which yields

$$H = \sum_{k} \begin{bmatrix} a^{\dagger}_{A,k} & a^{\dagger}_{B,k} \end{bmatrix} h(k) \begin{bmatrix} a_{A,k} \\ a_{B,k} \end{bmatrix}, \tag{2.40}$$

with Bloch Hamiltonian

$$h(k) = \begin{bmatrix} \bar{E} & v + we^{ika} \\ v + we^{-ika} & \bar{E} \end{bmatrix}. \tag{2.41}$$

The energy eigenvalues are therefore

$$E_{\pm}(k) = \bar{E} \pm \sqrt{(v + w\cos ka)^2 + (w \sin ka)^2}. \tag{2.42}$$

A study of the energies given by Eq. 2.42 with respect to the parameters v and w can inform us of which phase our system is in. When $|v| > |w|$ or $|v| < |w|$, the bandgap is open. We can compute an integer quantity (the topological invariant) known as the Zak phase

$$\nu_n = \frac{1}{\pi} \oint \langle \psi_{nk}(\vec{r}) | \nabla_k | \psi_{nk}(\vec{r}) \rangle \, dk, \tag{2.43}$$

where $|\psi_{nk}(\vec{r})\rangle$ is the Bloch eigenstate of H. This integer-valued invariant can only have one of two possible values:

$$\nu = \begin{cases} 0 & |v| > |w|, \text{trivial phase, or} \\ 1 & |v| < |w|, \text{non-trivial phase.} \end{cases} \tag{2.44}$$

In the non-trivial phase, robust localized edge modes that exponentially decay into the chain appear. This is how bulk-boundary correspondence manifests itself in the SSH chain. The bulk information comes from the value of the Zak phase, which is computed using the

(bulk) bandstructure. The boundary information comes from the appearance/lack of edge modes: when the Zak phase is 1, localized edge modes appear, while when the Zak phase is 0, there are no localized edge modes. The transition from trivial to non-trivial or vice versa occurs at the transition point $v = w$, which is where the bandgap closes. All this works well provided that $H^\dagger = H$, i.e. the system is Hermitian. In the non-Hermitian case, however, this can get a bit more complicated, though some of the intuition transfers over.

2.5.4 Zak Phase and Wilson Loop

The Zak phase, Eq. 2.66, is presented in terms of derivatives. However, the eigenmodes from simulations may have small discontinuities that can present practical issues for derivatives. As a result, we compute the Zak phase using a mathematical tool known as a Wilson loop. The Hermitian Wilson loop is computed by comparing the eigenstates of the system at nearby points in the Brillouin zone. In this section, we present a simplified derivation of the Hermitian Wilson loop. Readers who would like a deeper understanding of the Hermitian Wilson loop are advised to look at Section A.3.2 of this thesis.

In a bandstructure, bands can be non-degenerate or degenerate, i.e., it does not overlap with another band. A non-degenerate band is isolated. A degenerate band is has at least two bands overlapping. We must consider the two situations when computing the Wilson loop. The non-degenerate case is the simpler of the two; as a result, we begin with this case.

Non-Degenerate Wilson Loop

We start by doing a Taylor expansion up to first order of the Bloch eigenstate for the n^{th} band

$$\langle\psi_{nk}(\vec{r})|\psi_{n(k+\delta k)}(\vec{r})\rangle \approx 1 + \langle\psi_{nk}(\vec{r})|\nabla_k|\psi_{nk}(\vec{r})\rangle\delta k, \tag{2.45}$$

where $|\psi_{nk}(\vec{r})\rangle$ is the eigenvector for band n at a particular point k, in the Brillouin zone, and we have defined $A_k = \langle\psi_{nk}(\vec{r})|\nabla_k|\psi_{nk}(\vec{r})\rangle$. We define the Wilson loop in terms of the overlap of two discrete adjacent states indexed 0 to 1 as $\bar{O}_{0,1}$, from 1 to 2 as $\bar{O}_{1,2}$ and so on. We have $\bar{O}_{0,1} = \langle\psi_{nk_0}(\vec{r})|\psi_{nk_1}(\vec{r})\rangle$, $\bar{O}_{1,2} = \langle\psi_{nk_1}(\vec{r})|\psi_{nk_2}(\vec{r})\rangle$, ..., $\bar{O}_{L-2,L-1} = \langle\psi_{nk_{L-2}}(\vec{r})|\psi_{nk_{L-1}}(\vec{r})\rangle$. The application of each overlap consecutively gives

$$\bar{O}_{L-2,L-1}\cdots\bar{O}_{2,3}\bar{O}_{1,2}\bar{O}_{0,1} = e^{-i\sum_{l=0}^{L-1} A_{k_{l-1,l}}\delta k}. \tag{2.46}$$

Since we are on a closed loop, the endpoints are set equal to each other, so the sum above is a sum with periodic boundary conditions. We have then that

$$\bar{O}_{L-2,L-1}\cdots\bar{O}_{2,3}\bar{O}_{1,2}\bar{O}_{0,1} \approx e^{-i\oint A_k\cdot dk} = W, \tag{2.47}$$

where W is the Wilson loop. Thus, we have

$$\oint A\cdot dk = \frac{i}{\pi}\ln[W]. \tag{2.48}$$

Eq. 2.48 is the Zak phase of non-degenerate bands. Next, we work with the degenerate case.

Degenerate Wilson Loop

In the degenerate case, instead of a single band, we consider the case where there are N degenerate bands. Following a similar approach as the non-degenerate case, we start with the Taylor expansion

$$\langle \psi_{mk}(\vec{r})|\psi_{n(k+\delta k)}(\vec{r})\rangle \approx 1 + \langle \psi_{mk}(\vec{r})|\nabla_k|\psi_{nk}(\vec{r})\rangle \delta k, \tag{2.49}$$

where this time $m, n = 1, ..., N$ and $\langle \psi_{mk}(\vec{r})|\nabla_k|\psi_{nk}(\vec{r})\rangle = (A_k)_{mn}$.

We have that $(\bar{O}_{0,1})_{mn} = \langle \psi_{mk_0}(\vec{r})|\psi_{nk_1}(\vec{r})\rangle$, $(\bar{O}_{1,2})_{mn} = \langle \psi_{mk_1}(\vec{r})|\psi_{nk_2}(\vec{r})\rangle$, ..., $(\bar{O}_{L-2,L-1})_{mn} = \langle \psi_{nk_{L-2}}(\vec{r})|\psi_{nk_{L-1}}(\vec{r})\rangle$, which gives

$$(\bar{O}_{L-2,L-1})_{mn} \cdots (\bar{O}_{2,1})_{mn}(\bar{O}_{0,1})_{mn} = e^{-i\sum_{l=0}^{L-1}(A_{k_{l-1,l}})_{mn}\delta k}. \tag{2.50}$$

As a result, we have

$$e^{-i\sum_{l=0}^{L-1}(A_{k_{l-1,l}})_{mn}\delta k} = e^{-i\oint (A_k)_{mn}\cdot dk}. \tag{2.51}$$

Since we are on a closed loop, the end points are set equal to each other, so the sum above

is a sum with periodic boundary conditions. We have then that

$$(\bar{O}_{L-2,L-1})_{mn} \cdots (\bar{O}_{2,3})_{mn}(\bar{O}_{1,2})_{mn}(\bar{O}_{0,1})_{mn} \approx e^{-i \oint (A_k)_{mn} \cdot dk} = \hat{W}. \tag{2.52}$$

where $\hat{W}$ is an $N \times N$ matrix. If we set $(A_k)_{mn} = \hat{A}$, then we have that the Wilson loop is

$$\hat{W} = e^{-i \oint \hat{A} \cdot dk}. \tag{2.53}$$

Thus, we have

$$\oint \hat{A} \cdot dk = \frac{i}{\pi} \ln[\hat{W}]. \tag{2.54}$$

Eq. 2.54 is the Zak phase of degenerate bands.

2.5.5 Non-Hermitian Topology

Non-Hermitian systems are those that exchange energy with the rest of the universe and therefore experience loss and gain. They have generated much interest due to the richness of effects that can be explored and their potential applications. Research in the subject is increasing with activity, as evident from various reviews on the subject [128–131].

Non-Hermitian topological systems demonstrate some unique phenomena, such as the non-Hermitian skin effect [132]. They also provide more accessible ways to create structured degeneracies such as knots [133, 134]. These features of non-Hermitian topological systems make them attractive for technological applications. These degeneracies can have

topologically non-trivial features which can, for example, form knotted or braided structures in the Brillouin zone [135–139].

One general feature of non-Hermitian topological systems is that generally, they have complex eigenvalues, though previous work [140] has shown that if a non-Hermitian Hamiltonian is PT-symmetric, then $\tilde{E}_n \in \mathbb{R}$. In the general case though, a non-Hermitian Hamiltonian has eigenvalues, $\tilde{E}_n \in \mathbb{C}$, where the imaginary part is the loss/gain rate associated with a given mode and the real part is interpreted as usual. This ultimately leads to a topological invariant that is a complex number. In this section, we will discuss non-Hermitian phenomena briefly. Much of this discussion will be kept general, but in the condensed matter context. We briefly discuss an important formalism used to handle non-Hermitian quantum mechanics [141]: biorthogonal quantum mechanics [75]. We restrict ourselves to finite-dimensional Hilbert spaces. This discussion is important to lay the conceptual groundwork for a parallel formalism we use to study non-Hermitian systems in the context of photonics.

In general, in order to understand topological states for open systems, we have to generalize the bulk-boundary correspondence to non-Hermitian Hamiltonians. Non-Hermitian Hamiltonians have left and right eigenvectors, and their eigenvalues are, in general, complex. Recall that for Hermitian Hamiltonians, the eigenvalues are real, and there is only one set of eigenvectors (because the left and right eigenvectors are equal). The Hermitian case satisfies the typical orthonormality condition

$$\langle \psi_m(\vec{r}) | \psi_n(\vec{r}) \rangle = \delta_{mn}. \tag{2.55}$$

The non-Hermitian eigenvectors do not satisfy the typical orthonormality conditions. In order to work with these, we use biorthogonal quantum mechanics [75]. Biorthogonal quantum mechanics has been applied to understand the non-Hermitian skin effect [141]. We will not spend much time on biorthogonal quantum mechanics here; however, we will introduce it briefly, as the approach used for this thesis is parallel to it. Our discussion will follow [75], where we talk about the application to resonances: in particular, metastable states that tend to decay over time.

The eigenfunctions for Hermitian operators vanish at infinity, making it straightforward to construct a well-behaved inner product. This is not the case for non-Hermitian operators. The eigenfunctions for non-Hermitian operators are associated with loss and lead to decaying resonant states. They have complex eigenvalues and a pair of eigenstates (left and right). In general, let H be the non-Hermitian Hamiltonian. We have

$$H\,|\tilde{\psi}_n^R(\vec{r})\rangle = \tilde{E}_n\,|\tilde{\psi}_n^R(\vec{r})\rangle\,, \tag{2.56a}$$

$$H^\dagger\,|\tilde{\psi}_n^L(\vec{r})\rangle = \tilde{E}_n^*\,|\tilde{\psi}_n^L(\vec{r})\rangle\,, \tag{2.56b}$$

where $\tilde{E}_n \in \mathbb{C}$ are the eigenvalues, while $|\tilde{\psi}_n^L\rangle \in \mathcal{H}^L$ and $|\tilde{\psi}_n^R\rangle \in \mathcal{H}^R$ are the left and right eigenvectors respectively, with $\mathcal{H}^L$ and $\mathcal{H}^R$ are the left and right Hilbert spaces respectively and n is the band index. We can write

$$\tilde{E}_n = E_n - i\Gamma_n/2, \tag{2.57}$$

where Γ_n is the decay rate of the mode and τ_n is the decay time, with

$$\tau_n = \frac{\hbar}{\Gamma_n}, \tag{2.58}$$

where $\hbar$ is the reduced Planck constant. The energy from the decaying modes are handled with the outgoing boundary conditions. Away from exceptional points, the biorthonormality condition is satisfied for the eigenmodes

$$\langle \tilde{\psi}_m^L(\vec{r}) | \tilde{\psi}_n^R(\vec{r}) \rangle = \delta_{mn}. \tag{2.59}$$

We see that the introduction of outgoing boundary conditions to the time-independent Schrödinger equation introduces complex eigenvalues (associated with decaying modes), and a pair of eigenmodes that, together, satisfy a biorthogonality condition. The result is a complexity of resonant phenomena, where the resonances are the eigenmodes of the time-independent Schrödinger equation. This makes the subject of resonant phenomena and thus non-Hermitian systems rich and exciting.

Many of the ideas presented in this section on condensed matter systems transfer to periodic photonic systems.

2.6 Topological Photonics

Since their appearance in electronic platforms, topological phenomena have also appeared in photonic [55–57, 142–147], acoustic [148–150], mechanical [151–153] and even electric circuit [154–156] systems. In this section, we only focus on topological states in periodic photonic media, as this is the focus of the topological work in this thesis.

Topological photonic platforms have gained significant interest due to being robust to certain fabrication defects. This is important because fabrication defects in integrated photonic device manufacturing diminish device performance, which in turn decreases production yields.

2.6.1 Photonic Tight-Binding Model

Just as with condensed matter systems, photonic periodic structures can be represented by a tight-binding model similar to Eq. 2.37. We begin by examining two lossless structures, A and B, interacting with each other through classical electromagnetic modes. We first define

$$|\phi_\alpha(\vec{r})\rangle = \begin{bmatrix} \vec{E}_\alpha(\vec{r}) \\ \vec{H}_\alpha(\vec{r}) \end{bmatrix}, \tag{2.60}$$

where $\alpha \in \{A, B\}$. If each structure's modes are indexed by n, we can express the modes of this combined system as a superposition of the modes of the two systems

$$|\psi(\vec{r})\rangle = \sum_{\alpha,n} a_{\alpha,n} \, |\phi_{\alpha,n}(\vec{r})\rangle \, . \tag{2.61}$$

To simplify things, we assume we are just concerned with one mode per system. Eq. 2.61 then simplifies to

$$|\psi(\vec{r})\rangle = \sum_{\alpha} a_{\alpha} \, |\phi(\vec{r})_{\alpha}\rangle \, , \tag{2.62}$$

which is the same as Eq. 2.30 in Section 2.5 for the electronic Wannier functions in condensed matter systems. Thus, $|\phi(\vec{r})_{\alpha}\rangle$ take the role of Wannier states.

The result is a TB Hamiltonian for a single mode resembling that of TB of condensed matter physics, given by

$$\hat{H} = \omega_n \delta_{jl} + \sum_{\alpha,\beta} \left(t_{\alpha\beta} a^*_{\alpha} a_{\beta} + t_{\beta\alpha} a^*_{\beta} a_{\alpha} \right) , \tag{2.63}$$

where ω_α are the on-site frequencies, a_α, a_β are the mode coefficients at site α, β. $t_{\alpha\beta}$ are the coupling of mode at site α with a mode at site β. If we are looking at the influence of an arbitrary mode, n, Eq. 2.63 becomes

$$\hat{H} = \omega_\alpha \delta_{jl} + \sum_{\alpha,\beta,n} \left(t_{\alpha\beta} a^*_{\alpha,n} a_{\beta,n} + t_{\beta\alpha} a^*_{\beta,n} a_{\alpha,n} \right) . \tag{2.64}$$

We can apply Eq. 2.64 to a Hermitian photonic SSH chain.

2.6.2 Hermitian Photonic SSH Chain

In Section 2.5.3, we discussed the polyacetylene chain with the alternating double and single bonds, which correspond to alternating electron hopping amplitudes. We now present the Hermitian photonic SSH chain, first observed experimentally in a semi-infinite photonic lattice [157]. In addition, programmable SSH lattices have been introduced [114]. Fig. 2.5 depicts an example of a photonic SSH lattice. We note that $t_{\alpha\beta} = v$, when $\alpha = A$ and $\beta = B$ from the same unit cell, while $t_{\alpha\beta} = w$, when $\alpha = A$ and $\beta = B$ for A and B in different (adjacent) cells. Here, they represent mode couplings instead of electron hopping (tunneling) amplitudes. We thus have the photonic SSH Hamiltonian

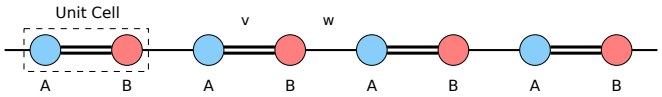


Figure 2.5: SSH chain diagram. w is the coupling shown by the double lines, and v is for the single lines.

$$H = \omega + \sum_{n} \sum_{i=1}^{L} \left(v\, a^{\dagger}_{A,i,n} a_{B,i,n} + w\, a^{\dagger}_{B,i,n} a_{A,i+1,n} + \text{h.c.} \right). \tag{2.65}$$

Its Zak phase is computed as in condensed matter systems

$$\nu_n = \frac{1}{\pi} \oint \langle \psi_{nk}(\vec{r}) | \nabla_k | \psi_{nk}(\vec{r}) \rangle \, dk, \tag{2.66}$$

where $|\psi_{nk}(\vec{r})\rangle$ are the electromagnetic Bloch modes. Similar to condensed matter systems, we can use the Wilson loop.

2.6.3 Topological Photonics in Open Systems

We have just discussed topological states in Hermitian photonic systems. In the last decade, however, there has been significant interest in non-Hermitian photonics. This is due to the possibility of designing exceptional points [158] and other features with potential technological applications such as on-chip reconfigurable optical devices [158] and lasing [159]. The subject is an active area of research, with work on bulk-edge correspondence in open photonic systems [160] and topological bound states in the continuum in non-Hermitian photonics [161] as some recent examples. Recall from Section 2.5.5 that a non-Hermitian system is described by a Hamiltonian that is not self-adjoint, i.e., $\hat{H} \neq \hat{H}^\dagger$. The eigenvalues are generally complex, i.e., $\tilde{E}_n \in \mathbb{C}$ and they have left and right eigenvectors, $\langle \psi_n^L(\vec{r}) |$ and $|\psi_n^R(\vec{r})\rangle$, respectively. These eigenvectors respect the biorthonormality condition. These properties also hold in non-Hermitian topological photonics, where the eigenvalues are the eigenfrequencies and the eigenvectors are the electric and magnetic field modes similar to Eq. 2.11, except now, we have two electric and magnetic field mode pairs: one in the left eigenspace and the other in the right eigenspace.

The Hamiltonian for the non-Hermitian SSH model in photonics is [158, 159]

$$H(k) = \omega_A \sum_n a_n^{A\dagger} a_n^A + \omega_B \sum_n a_n^{B\dagger} a_n^B + \sum_n \left[\kappa_1 \left(a_n^{B\dagger} a_n^A + a_n^{A\dagger} a_n^B \right) + \kappa_2 \left(a_{n-1}^{B\dagger} a_n^A + a_n^{A\dagger} a_{n-1}^B \right) \right], \tag{2.67}$$

where $\omega_{A/B} \in \mathbb{C}$ and $\kappa_{1/2}$ are the intracell and intercell couplings, respectively. In matrix form, we have

$$H(k) = \begin{bmatrix} \omega_A & \kappa_1 + \kappa_2 e^{-ik} \\ \kappa_1 + \kappa_2 e^{ik} & \omega_B \end{bmatrix}. \tag{2.68}$$

2.7 Models for the dielectric permittivity of materials

In solving numerical electromagnetic problems, materials can either be modeled with experimental data or analytic models of the dielectric permittivity (or, equivalently, the refractive index) as a function of frequency. In our work on SWGs, we used experimental refractive index data for silica and silicon to model the structures in FDTD using Ansys Lumerical. Lumerical interpolates the refractive index from tabulated data values for the materials. There are various experimental material data sets, such as Palik [162] and Johnson & Christy [163]. In our work, we used Palik data for both Silica and Silicon.

In plasmonics, the optical behaviour of the metallic resonators is mostly determined by the response of free electrons in the conduction band to an incoming electric field. As

pointed out by S.V. Gaponenko [1], we can model the free-electron response as a free-electron gas. The electron response can be modeled as an oscillator, with the electric field providing the driving force

$$\epsilon(\omega) = \epsilon_\infty - \frac{\omega_p^2}{\omega^2 - i\Gamma\omega}, \tag{2.69}$$

where ω is the operating frequency, ϵ_∞ is a term that accounts for the effect of bound electrons, and ω_p is the plasma frequency given by $\omega_p = \sqrt{\frac{Ne^2}{\epsilon_0 m}}$, where N is the electron number density, e is the electron charge, ϵ_0 is the free-space permittivity, and m is the effective mass of the electron in the lattice. We also have Γ, which is the damping term that captures electron-electron and electron-atom/molecular collisions. The inverse of the damping term gives the damping rate, $\tau = \frac{1}{\Gamma}$. The Drude model has various limitations; for example, it does not account for the response of bound electrons nor quantum effects. Nevertheless, it is a good starting point for a simple model of the optical behaviour of metals. As a result, in the topological plasmonics work, we use this model to build an auxiliary eigenvalue problem, where we chose the parameters of Pocock et al. [164].

2.8 Metallic Nanoresonators and Quasinormal Modes (QNMs)

The free electrons in metals are responsible for their heat and electrical conduction. In photonics, these electrons are also responsible for how these metals interact with light, and their collisions with other electrons and atoms determine how the energy from these light-

matter interactions leaks away from the system over time as thermal energy. Fig. 2.6 depicts possible optical responses to an incident electric field. As the field comes into contact with the particle, the electrons respond by redistributing themselves at the metal's surface such that there are positive (red circles) charges in some places and negative (blue circles) charges in others. The electric field then becomes enhanced in these regions. The effect is strongest at the resonance frequencies, which correspond to field distributions that are known as quasinormal modes, which, as the name suggests, are not strictly normal modes one would expect for Hermitian systems. The resonances depend on the size, shape, as well

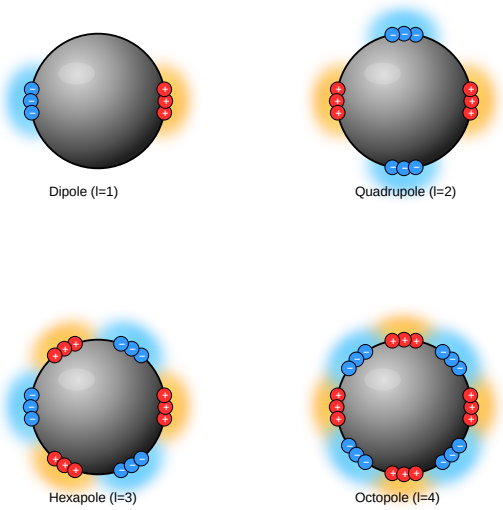


Figure 2.6: Illustration of the response of a metallic nanoparticle (sphere) when excited by an external field. Surface charges move in response to the excitation source (electric field, for example), which in turn causes the nanoparticle to produce an enhanced electric field in its vicinity at the resonance frequency. This is the quasinormal mode. Excitations can occur at different orders, for example, dipole (top-left), quadrupole (top-right), hexapole (bottom-left), and octopole (bottom-right).

as the material composition of the nanoparticles participating in the interaction. They can be thought of as the natural electromagnetic modes of the metallic nanoparticles and can be used to construct a basis set for a Hilbert space if we introduce absorbing boundary conditions. Their associated eigenfrequencies are complex numbers, whose imaginary part gives information on their decay rate, as they are leaky and hence non-Hermitian.

A periodic arrangement of these particles makes an excellent playground to study non-Hermitian topology. The overall interaction in such a periodic system is depicted in the bandstructure which one obtains from solving the eigenvalue problem.

Contrary to normal modes, in their typical form, quasinormal modes exponentially diverge at infinity. This means that one cannot naturally incorporate them into the Hilbert space formalism, as they are not square-integrable, and hence there is no way to define an inner product space. As a result, they are challenging to tame. Work needs to be done for this to be possible; they must be regularized, normalized, and complete. We will briefly discuss these in this section. This discussion is based on work presented in [74, 165].

We first start with regularization. In this process, the original space of QNMs is mapped to a space of finite, exponentially-decaying functions. In this space, they are now square-integrable, making it possible to define an inner product space. Recall that the QNMs take the form

$$|\tilde{\Psi}\rangle \propto e^{i\vec{k}\cdot\vec{r}}. \tag{2.70}$$

Fundamentally, the mapping is done using a coordinate transformation

$$\vec{r} \to \vec{r} + i\vec{\alpha}(\vec{r}). \tag{2.71}$$

The imaginary term from this leads to an exponentially decaying physical field, causing the QNMs to be square-integrable. Practically, this coordinate transformation is achieved via a perfectly matched layer (PML) which is commonly used in computational electrodynamics to implement absorbing boundary conditions. Regularization is a very important step to take in order to build up a Hilbert space formalism for QNMs.

Once this regularization process is done, one must normalize the QNMs. This leads to a definition of a biorthogonality condition, which, as discussed in Section 2.5.5, is the inner product norm for non-Hermitian systems. The standard normalization equations of QNMs actually come from the Lorentz reciprocity theorem. In order to demonstrate this, we consider two quasinormal modes of an open resonator, with fields $\vec{E}_m, \vec{H}_m$ and $\vec{E}_n, \vec{H}_n$ at eigenfrequencies $\tilde{\omega}_m$ and $\tilde{\omega}_n$ respectively. Their frequency-domain Maxwell equations are

$$\nabla \times \vec{E}_m = i\tilde{\omega}_m \, \mu(\tilde{\omega}_m)\vec{H}_m, \qquad \nabla \times \vec{H}_m = -i\tilde{\omega}_m \, \epsilon(\tilde{\omega}_m)\vec{E}_m, \tag{2.72a}$$

$$\nabla \times \vec{E}_n = i\tilde{\omega}_n \, \mu(\tilde{\omega}_n)\vec{H}_n, \qquad \nabla \times \vec{H}_n = -i\tilde{\omega}_n \, \epsilon(\tilde{\omega}_n)\vec{E}_n. \tag{2.72b}$$

Taking the scalar product of $\vec{E}_m$ with $\nabla \times \vec{H}_n$ and of $\vec{H}_n$ with $\nabla \times \vec{E}_m$, and subtracting

the corresponding expressions with m and n exchanged, one obtains

$$\nabla\cdot\Big(\vec{E}_m\times\vec{H}_n-\vec{E}_n\times\vec{H}_m\Big)=i\Big[\vec{H}_m\cdot(\tilde{\omega}_n\mu(\tilde{\omega}_n)-\tilde{\omega}_m\mu(\tilde{\omega}_m))\vec{H}_n-\vec{E}_m\cdot(\tilde{\omega}_n\epsilon(\tilde{\omega}_n)-\tilde{\omega}_m\epsilon(\tilde{\omega}_m))\vec{E}_n\Big]. \tag{2.73}$$

Integrating this identity over the whole space Ω and applying the divergence theorem yields

$$\begin{aligned}&\oint_{\partial\Omega}(\vec{E}_m\times\vec{H}_n-\vec{E}_n\times\vec{H}_m)\cdot d\vec{S}\\&\quad=i\iiint_\Omega\Big[\vec{H}_m\cdot(\tilde{\omega}_n\mu(\tilde{\omega}_n)-\tilde{\omega}_m\mu(\tilde{\omega}_m))\vec{H}_n-\vec{E}_m\cdot(\tilde{\omega}_n\epsilon(\tilde{\omega}_n)-\tilde{\omega}_m\epsilon(\tilde{\omega}_m))\vec{E}_n\Big]dV.\end{aligned} \tag{2.74}$$

For quasinormal modes, the surface term on the left-hand side vanishes: either because the fields satisfy outgoing-wave conditions at infinity, or because a perfectly matched layer (PML) maps the divergent fields to square-integrable ones. The result is therefore the orthogonality relation

$$\iiint_\Omega\Big[\vec{H}_m\cdot(\tilde{\omega}_n\mu(\tilde{\omega}_n)-\tilde{\omega}_m\mu(\tilde{\omega}_m))\vec{H}_n-\vec{E}_m\cdot(\tilde{\omega}_n\epsilon(\tilde{\omega}_n)-\tilde{\omega}_m\epsilon(\tilde{\omega}_m))\vec{E}_n\Big]dV=0,\qquad m\neq n. \tag{2.75}$$

To obtain the normalization condition, one sets $m=n$ but considers an infinitesimal difference of eigenfrequencies. Expanding

$$\tilde{\omega}_n\epsilon(\tilde{\omega}_n)-\tilde{\omega}_m\epsilon(\tilde{\omega}_m)\simeq(\tilde{\omega}_n-\tilde{\omega}_m)\,\frac{\partial}{\partial\omega}(\omega\epsilon(\omega))\bigg|_{\omega=\tilde{\omega}_n}, \tag{2.76}$$

and similarly for μ, dividing through by $(\tilde{\omega}_n - \tilde{\omega}_m)$ and taking the limit $\tilde{\omega}_m \to \tilde{\omega}_n$ leads to

$$\iiint_\Omega \left[\vec{E}_n \cdot \frac{\partial(\omega\epsilon)}{\partial\omega}\Big|_{\tilde{\omega}_n} \vec{E}_n - \vec{H}_n \cdot \frac{\partial(\omega\mu)}{\partial\omega}\Big|_{\tilde{\omega}_n} \vec{H}_n \right] dV \;=\; N_n. \tag{2.77}$$

This volume integral defines the normalization of the quasinormal modes $\vec{E}_n, \vec{H}_n$. N_n is a finite complex normalization factor.

Once we have a way to normalize the QNMs, we can now define the orthonormality of the QNMs. In order to do this, we define the QNM inner product between two modes $|\tilde{\psi}_m\rangle = [\vec{E}_m, \vec{H}_m]$ and $|\tilde{\psi}_n\rangle = [\vec{E}_n, \vec{H}_n]$ as

$$\langle\tilde{\psi}_m|\tilde{\psi}_n\rangle = \frac{1}{\sqrt{N_n N_m}} \iiint_\Omega \left[\vec{E}_m \cdot \frac{\partial(\omega\epsilon)}{\partial\omega}\Big|_{\tilde{\omega}_n} \vec{E}_n - \vec{H}_m \cdot \frac{\partial(\omega\mu)}{\partial\omega}\Big|_{\tilde{\omega}_n} \vec{H}_n \right] dV. \tag{2.78}$$

With this definition, the QNMs satisfy the orthonormality relation

$$\langle\tilde{\varphi}_m|\tilde{\psi}_n\rangle = \delta_{mn}. \tag{2.79}$$

This is, in fact, a biorthonormality relation, which we can put in terms of left $\langle\tilde{\varphi}_m|$ and right $|\tilde{\psi}_m\rangle$ eigenvectors, with the left and right eigenvectors being defined as

$$\langle\tilde{\varphi}_m| = \left[\vec{E}_m(\vec{r}) \frac{\partial(\omega\epsilon)}{\partial\omega}\Big|_{\tilde{\omega}_m} \quad -\vec{H}_m(\vec{r}) \frac{\partial(\omega\mu)}{\partial\omega}\Big|_{\tilde{\omega}_m} \right], \tag{2.80a}$$

$$|\tilde{\psi}_n\rangle = \begin{bmatrix} \vec{E}_n(\vec{r}) \\ \vec{H}_n(\vec{r}) \end{bmatrix}, \tag{2.80b}$$

respectively. Now that biorthonormality is established, we must demonstrate that the basis states are indeed sufficient to fully describe the physics of the system. This is known as completeness.

3

Methods

Advances in numerical methods and computing hardware over the many decades have played a significant role in establishing computational electromagnetics as a serious discipline. This, in turn, has helped to produce some of the technological advances we benefit from today. They are used to simulate components of integrated photonic devices, such as metamaterials, before they are made in a lab or foundry. The main advantage of doing this is to avoid significant trial and error in a lab, saving time, resources, and money, all of which are important in academia and industry. In Section 3.1, we discuss the two conventional numerical tools we use in this thesis: FDTD and FEM. FDTD was used for the metamaterial grating project in collaboration with the National Research Council, Canada,

presented in Chapter 4, while FEM was used for the topological plasmonics projects in collaboration with the Max Planck Institute for the Science of Light, Germany, presented in Chapters 5 and 6. Section 3.2 outlines how we use machine learning in Chapter 4. Section 3.3 gives more details about the theory and computations used for Chapters 5 and 6.

The overarching problem we worked on addressing is that, though full lattice simulations using either FDTD or FEM are very good at simulating metamaterials whose dimensions and features are used in devices, they are computationally costly and, by themselves, are limiting in providing insights into the physics of the metamaterials. The consequence is that people may work with theoretical models. These theoretical models have assumptions that are often too simple when we consider metamaterials being used for device design. These restrictions from both the computational modeling and theoretical models make it difficult to get relevant and rapid physical insights at low computational cost that could guide device design. These issues plague the design of finite subwavelength gratings, as conventional effective medium theories assume the system is infinite, and so FDTD simulations are often used to model them. These simulations are costly for 3D structures when many optimization procedures have to be done on multiple designs. This issue is compounded by the fact that the design space for SWGs and plasmonic metamaterials is quite large. A similar problem presents itself when designing plasmonic metamaterials to create realistic structures that leverage their topological properties for device operation. FEM simulations are computationally expensive, and proper theoretical models that account for resonator shape, mode contributions, etc., to study topological states in plasmonic systems

are lacking. This thesis aims to address this by developing: (1) an extension of the concept of effective medium to finite subwavelength gratings used in device design, and (2) a theoretical model for topological plasmonics that accounts for resonator shapes, mode contributions, long-range couplings, etc., supported by full lattice simulations. Together, the hope is that these tools help find a balance between computational complexity and theoretical models to guide the design and discovery of photonic metamaterials for device applications.

3.1 Numerical Methods for Classical Electromagnetism

In this section, we present the numerical setup and parameters for the FDTD and the FEM simulations used for the work of this thesis. We will discuss some of the technical details of both methods, starting with FDTD and then FEM. Readers should consult various sources to learn more about FDTD [166–172] and FEM [173–176] as well as the mathematics behind FEM and its adaptation to classical electromagnetism [177–182].

3.1.1 Finite Difference Time Domain (FDTD)

One of the major advances in computational electromagnetism was the development of the finite-difference time-domain (FDTD) method by Kane Yee [183], which exploded in popularity when increasingly powerful computational resources became more available.

It starts directly from Maxwell's equations, representing them as finite differences on a rectangular/square grid. The reader can consult various reviews and books on FDTD [166, 170, 172].

The FDTD method solves Maxwell's curl equations in the time domain

$$\nabla \times \vec{E} = -\mu \frac{\partial \vec{H}}{\partial t}, \tag{3.1a}$$

$$\nabla \times \vec{H} = \vec{J} + \epsilon \frac{\partial E}{\partial t}, \tag{3.1b}$$

using the finite difference method, which approximates derivatives as finite central differences

$$\frac{df(t)}{dt} \approx \frac{f(t_0 + \Delta t/2) - f(t_0 - \Delta t/2)}{\Delta t}, \tag{3.2}$$

up to second-order accuracy, such that the electric and magnetic fields are staggered in both space and time. The half represents a half-step in time. The Yee algorithm uses a leap-frog update scheme where the fields are updated on a staggered grid. Update equations compute future values of the field from past values. The electric field at the current time is used to find the magnetic field a half time-step ahead. These updated magnetic fields are then used in turn to compute the electric fields. These field computation steps are repeated until the electromagnetic fields have been computed over the desired time.

If we take each grid cell to have dimensions $\delta x, \delta y, \delta z$ respectively, with the labels i, j, k,

then we have the discretized form of Maxwell's equations, with the time step n as

$$\vec{H}^{\,n+\frac{1}{2}} = \vec{H}^{\,n-\frac{1}{2}} - \Delta t\,\mu^{-1}\left(\nabla_d \times \vec{E}^{\,n}\right), \tag{3.3a}$$

$$\vec{E}^{\,n+1} = \vec{E}^{\,n} + \Delta t\,\epsilon^{-1}\left(\nabla_d \times \vec{H}^{\,n+\frac{1}{2}}\right) - \Delta t\,\epsilon^{-1}\vec{J}^{\,n+\frac{1}{2}}, \tag{3.3b}$$

where $\nabla_d\times$ is the discrete curl operator, defined as $\nabla_d\times = \left(\partial_y^d - \partial_z^d\right)\hat{x} + \left(\partial_z^d - \partial_x^d\right)\hat{y} + \left(\partial_x^d - \partial_y^d\right)\hat{z}$. We also have that the field values at positions i, j and k in the arrays correspond to staggered locations in real physical space according to the Yee grid, that is

$$E_x \to \left(i+\frac{1}{2},\, j,\, k\right), \qquad H_x \to \left(i,\, j+\frac{1}{2},\, k+\frac{1}{2}\right), \tag{3.4a}$$

$$E_y \to \left(i,\, j+\frac{1}{2},\, k\right), \qquad H_y \to \left(i+\frac{1}{2},\, j,\, k+\frac{1}{2}\right), \tag{3.4b}$$

$$E_z \to \left(i,\, j,\, k+\frac{1}{2}\right), \qquad H_z \to \left(i+\frac{1}{2},\, j+\frac{1}{2},\, k\right). \tag{3.4c}$$

The FDTD scheme on the Yee grid with the leap-frog is illustrated in Fig. 3.1, while the Pseudocode is displayed in Algorithm 1. In order to have numerical stability, the condition

$$\Delta t \leq \frac{1}{c\sqrt{\frac{1}{\Delta x^2} + \frac{1}{\Delta y^2} + \frac{1}{\Delta z^2}}}, \tag{3.5}$$

must be satisfied. Here c is the speed of light in a vacuum, $\Delta x, \Delta y, \Delta z$ are the grid sizes, and Δt is the time-step. The Courant condition defines an upper bound on the time-step.

Algorithm 1 FDTD algorithm on the Yee grid. The Yee grid is staggered in both space and time, n.

1: Initialize the electric field $\vec{E}^{\,0}$ on the center of edges.
2: Initialize the magnetic field $\vec{H}^{\,-1/2}$ on the center of faces.
3: Initialize the current density $\vec{J}^{0}$.
4: **for** $n = 0 \to N - 1$ **do**
5: Update the magnetic field (Faraday's law)

$$\vec{H}^{\,n+\frac{1}{2}} = \vec{H}^{\,n-\frac{1}{2}} - \Delta t\, \mu^{-1} \left(\nabla_d \times \vec{E}^{\,n} \right)$$

.
6: Apply the magnetic field's boundary conditions.
7: Update the electric field (Ampère's law)

$$\vec{E}^{\,n+1} = \vec{E}^{\,n} + \Delta t\, \epsilon^{-1} \left(\nabla_d \times \vec{H}^{\,n+\frac{1}{2}} - \vec{J}^{n+\frac{1}{2}} \right)$$

.
8: Apply the electric field's boundary conditions.
9: Update sources and currents $\vec{J}^{n+\frac{1}{2}}$.
10: **end for**

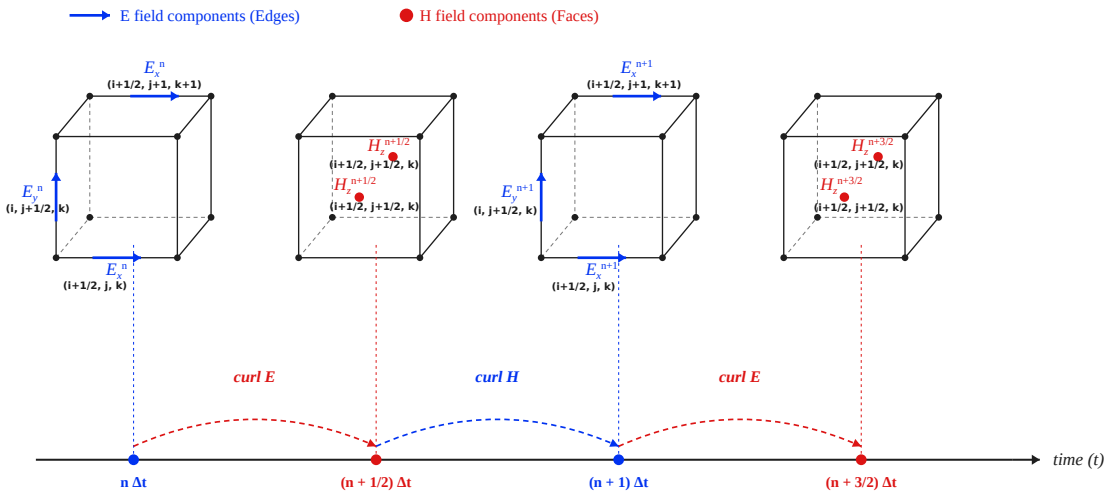


Figure 3.1: Illustration of the Yee grid (staggering in space, with coordinates i, j and k) for the field components E_x, E_y and H_z evolving forward (staggered) in time.

If the grid spacings are the same, i.e. $\Delta x = \Delta y = \Delta z = \Delta L$ then it simplifies to

$$\Delta t \leq \frac{\Delta L}{c\sqrt{3}}. \tag{3.6}$$

Boundary conditions must accompany these equations. In this thesis, we use, in particular, perfectly matched layers (PML), which are absorbing boundaries, and periodic boundary conditions (PBC), which must be enforced at every time step.

In order to extract useful optical data from the simulations, one requires data monitors to record field values on a subset of the grid. One can have point monitors, which record data at a point; line monitors, which record data along a line, plane monitors, which record

data in a plane; and volume monitors, which record data in a volume region. One can use frequency-domain power monitors that, as a simulation runs, perform discrete Fourier transforms at specified spatial locations and frequencies; the resulting data can then be used to compute the power flux through the monitor. This can then be used to compute the transmittance and reflectance (normalized power)

$$r(\lambda) = \frac{P_{\text{reflected}}(\lambda)}{P_{\text{incident}}(\lambda)}, \tag{3.7a}$$

$$t(\lambda) = \frac{P_{\text{transmitted}}(\lambda)}{P_{\text{incident}}(\lambda)}. \tag{3.7b}$$

This is the optical data that is relevant to our SWG work and is what we use to determine if two optical structures are equivalent. We discuss this further in Section 3.2.

In the FDTD work, we used Lumerical FDTD [184], where we simulate 3D structures with the grating feature. Then, we have the equivalent effective structure, where we replace the grating feature with a homogeneous structure with a dispersive effective index. This structure, because all features are now homogeneous, can be simulated in a 2D simulation space. Fig. 3.2 shows a conceptualization of how the structure is built in Lumerical. There are three major layers: (1) top layer, (2) middle layer, and (3) bottom layer. The top and bottom layers are silicon dioxide of thickness 1055 nm. The middle layer is 220 nm thick.

We divide the middle layer into 28 regions: 14 upper boxes and 14 lower boxes, so that partially etched structures can be included in the future and so that each box can be assigned a desired dimension and material appropriate to the physical structure of interest. In the case where we have a 3D simulation space, we have two sets of 28 boxes in the x-

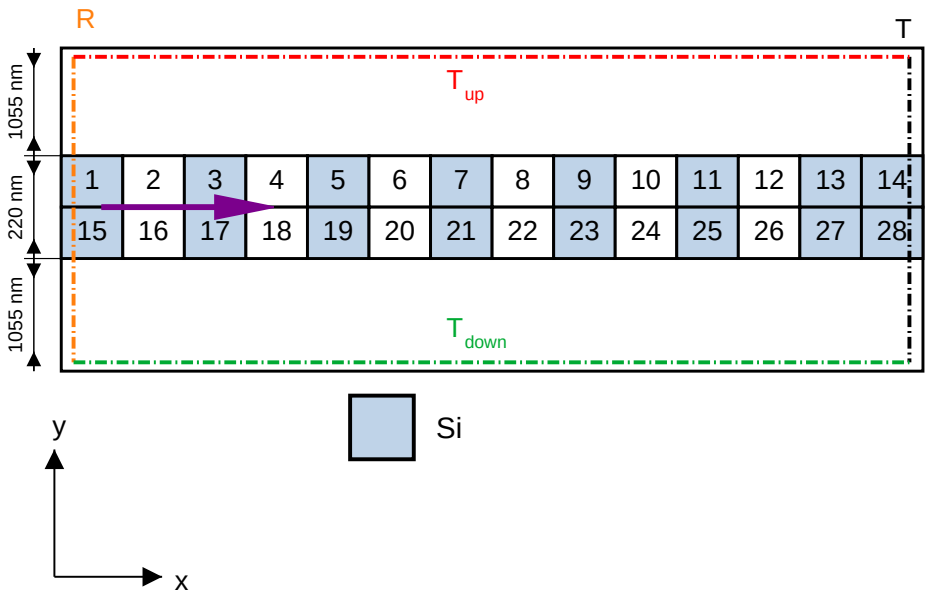


Figure 3.2: Illustration of the Lumerical setup.

direction (in order to capture the two parts of the SWG unit cell). In the 2D case, it is only 1 set of 28 boxes arranged in the x-direction. The input boxes (1 and 15) and the output boxes (14 and 28) are both made of silicon.

For the 3D simulations, a mode source (forward-pointing arrow in the middle layer) is placed in the input boxes and is injected in the forward x-direction. We choose the fundamental TM mode. The number of trial modes is set to 1000, and the number of field profile samples is set to 10. Our bandwidth is set to our range of interest, which is 1450 nm to 1650 nm. In the 3D structure, PML boundaries are used for the x and y directions. However, we used periodic boundaries in the z-direction, since we have a grating layer extending in that direction. The PML profile is set to the standard, with 8 layers, and is

a stretched-coordinate PML. The mesh type was set to uniform, with a conformal variant 0 mesh refinement. The best maximum mesh step in the x, y, and z directions was chosen to be 10 nm. This was determined from convergence tests done using a slab waveguide design.

In the 2D setup, the mesh type and mesh refinements, as well as the maximum mesh step, were the same as in the 3D setup, confirmed again with convergence tests. The boundaries were PML in the x and y directions, with stretched-coordinate PML with 8 layers and a standard profile. We used a 3000 fs simulation time. For the 2D simulations, a mode source of the fundamental TM mode was used, injected in the forward x direction. It is also positioned in the input boxes (1 and 15). The number of trial modes is also set to 1000, with the number of field profile samples set to 10. The bandwidth is as mentioned earlier.

We used four Frequency-Domain and Field Monitors to collect the transmission data horizontally (vertical broken lines in Fig. 3.2, R and T) and vertically (horizontal broken lines in Fig. 3.2, T_{up} and T_{down}). In the 2D case, we used linear Y-monitors to capture the fields scattered horizontally, and linear X-monitors to capture the fields scattered vertically. In the 3D structures, we used 2D X-normal and 2D Y-normal monitors to capture the fields scattered horizontally and vertically, respectively. We positioned the monitors as close to the boundaries as reasonably possible to ensure that the scattered fields are collected efficiently. Our simulation time is set to 1000 fs.

3.1.2 Finite Element Method (FEM)

Although the FDTD method is a powerful tool to solve Maxwell's equations in the time-domain, there are drawbacks. In order to compute the bandstructure using FDTD, one would simulate a unit cell using Bloch boundary conditions in the appropriate directions. One would then excite the system with a broadband dipole source. The resonances would be recorded using time monitors, and one would perform a fast Fourier transform (FFT) and look for peaks in the spectrum. However, FDTD only sees the frequencies that are excited by the source in the particular location in which it sits. Hence, it can miss bands in a bandstructure computation. Also, if two modes are too close, it might see them as the same. This gets worse as the losses of the system increase. Bandstructure computations are an eigenvalue problem, and they are important when working with topological photonics. As we are working with lossy systems, it is better to turn to a different tool other than FDTD to compute the bandstructure.

In this regard, we turned to the finite element method (FEM), which is a completely different approach [175, 185–187]. Instead of using a rectangular/square grid, it triangulates the structure. It converts Maxwell's equations into a variational form, approximates the electric and magnetic fields with basis functions on the triangulated space, and computes their coefficients. It can be used to solve eigenvalue problems, and it does not require a source. It gives the eigenfrequencies directly. It is the fundamental numerical tool behind COMSOL Multiphysics, which we used to tackle our non-Hermitian problem. The reader may consult various sources to learn more about FEM [173, 176]. Regarding the mathe-

matics behind FEM and its adaptation to classical electromagnetism, we point the reader to additional sources [177–182]. FEM is especially useful for solving eigenvalue problems, and it does not require a source as scattering-based methods do. It is the fundamental numerical tool behind COMSOL Multiphysics, which we used to tackle our topological plasmonics problem.

FEM involves decomposing a geometry into structures known as simplices, forming a simplicial complex. The d-simplex is the simplest d-dimensional object formed from $d+1$ points. For example, a 0-simplex is a point, a 1-simplex is a line, a 2-simplex is a triangle, a 3-simplex is a tetrahedron, and so on. If we have $d+1$ points, $v_0, v_1, ..., v_d$, the simplest geometric structure for this would have those points as its vertices, and they would satisfy the condition [182]

$$\sigma = \left\{ \sum_{i=0}^{d} \alpha_i v_i \ : \ \alpha_i \geq 0, \ \sum_{i=0}^{d} \alpha_i = 1 \right\}, \tag{3.8}$$

where α_i are called barycentric coordinates. A simplicial complex is a collection of simplices such that for every simplex within the structure, its faces are also included in the structure, and if two simplices are glued together, then the boundary along which they are glued is also part of the structure. More precisely, a simplicial complex $\mathcal{K}$ is a set of simplices, σ glued together such that if $\sigma \in \mathcal{K}$, then every face C of σ is such that $C \in \mathcal{K}$. Also, if $\sigma, \tau \in \mathcal{K}$, where σ, τ are two simplices, then $\sigma \cap \tau \in \mathcal{K}$.

In FEM, we discretize a geometry, Ω, with a simplicial complex, $\mathcal{K}$, such that

$$\Omega \approx \bigcup_{\sigma \in K} \sigma. \tag{3.9}$$

The simplicial complex is what we call the FEM mesh. Simplicial complexes are a natural way to discretize a geometry into simple building blocks of triangles or tetrahedra. In classical electromagnetism, this is particularly relevant as vector (basis) functions like the Nédélec basis functions used in FEM can be easily assigned to the edges of simplices in such a way as to maintain continuity. The electric fields are represented as line integrals along edges of simplices, magnetic flux on faces, and scalar quantities on the vertices. In the simplicial complex, each of our simplices has vertices associated with barycentric coordinates, α_i. One can construct first order Nédélec basis functions, $\vec{N}_{ij}$, for an oriented edge e_{ij}, between vertices i, j to give $\vec{N}_{ij} = \alpha_i \nabla \alpha_j - \alpha_j \nabla \alpha_i$. The electric field has its degrees of freedom represented as line integrals along the edges of the mesh, e_{ij}, as

$$e_{ij} = \int_{e_{ij}} \vec{E} \cdot d\vec{l}, \tag{3.10}$$

with the $\vec{N}_{ij}$ satisfying

$$\int_{e_{kl}} \vec{N}_{ij} \cdot d\vec{l} = \delta_{(ij),(kl)}, \tag{3.11}$$

which makes them dual to the edge degrees of freedom. As a result, we can approximate the electric field, $\vec{E}$, as

$$\vec{E} = \sum_e e_e \, \vec{N}_e. \tag{3.12}$$

The wave equation for the electric field is

$$\nabla \times \left(\mu_r^{-1} \nabla \times \vec{E} \right) - \frac{\omega^2}{c^2} \epsilon_r \vec{E} = 0, \tag{3.13}$$

which gives the weak-form PDE with a test function $\vec{v}$, giving

$$\int_\Omega \vec{v} \cdot (\nabla \times (\mu_r^{-1} \nabla \times \vec{E}))\, d\Omega - \frac{\omega^2}{c^2} \int_\Omega \epsilon_r\, \vec{v} \cdot \vec{E}\, d\Omega = 0. \tag{3.14}$$

After integration by parts, this becomes

$$\int_\Omega (\nabla \times \vec{v}) \cdot \mu_r^{-1} (\nabla \times \vec{E})\, d\Omega - \left(\frac{\omega}{c}\right)^2 \int_\Omega \epsilon_r\, \vec{v} \cdot \vec{E}\, d\Omega = 0. \tag{3.15}$$

When we expand the fields as Eq. 3.12 and choose test functions $\vec{v} = \vec{N}_m$ and sub into Eq. 3.15, we get

$$\sum_n e_n \sum_{\sigma \in \mathcal{K}} \left[\int_\sigma (\nabla \times \vec{N}_m) \cdot \mu_r^{-1} (\nabla \times \vec{N}_n)\, d\Omega - \left(\frac{\omega}{c}\right)^2 \int_\sigma \epsilon_r \vec{N}_m \cdot \vec{N}_n\, d\Omega \right] = 0. \tag{3.16}$$

and we set

$$K_{mn} = \sum_{\sigma \in \mathcal{K}} \int_\sigma (\nabla \times \vec{N}_m) \cdot \mu_r^{-1} (\nabla \times \vec{N}_n)\, d\Omega, \tag{3.17a}$$

$$M_{mn} = \sum_{\sigma \in \mathcal{K}} \int_\sigma \epsilon_r\, \vec{N}_m \cdot \vec{N}_n\, d\Omega, \tag{3.17b}$$

which gives the discrete Maxwell equations

$$\left(\mathbb{K} - \frac{\omega^2}{c^2} \mathbb{M} \right) e = 0, \tag{3.18}$$

where

$$e = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \end{bmatrix}. \tag{3.19}$$

The process is summarized in Algorithm 2.

Algorithm 2 Summary of FEM Maxwell eigenmode solver algorithm.

1: Generate simplicial mesh $\mathcal{K}$ of Ω.
2: Construct Nédélec edge basis functions $\{\vec{N}_e\}$.
3: Expand $\vec{E} = \sum_e e_e \vec{N}_e$.
4: **for all** $\sigma \in \mathcal{K}$ **do**
5: Compute element matrices $K^{(\sigma)}$ and $M^{(\sigma)}$.
6: Assemble into global matrices $\mathbb{K}$ and $\mathbb{M}$.
7: **end for**
8: Apply boundary conditions.
9: Solve

$$\mathbb{K}e = \left(\frac{\omega}{c}\right)^2 \mathbb{M}e$$

.
10: Reconstruct eigenmodes $\vec{E}(\vec{r})$ and postprocess.

There are two ways to compute resonance frequencies (eigenvalues) and resonant modes (eigenvectors): using the eigensolver of COMSOL, or doing a pole-search in COMSOL. The pole-search approach is slow and not very effective in finding all the eigenvalues. However, this approach works well with experimental material data. The eigensolver is computationally more efficient, as it gives all the eigenvalues at once, though it only works with analytic models. We chose the latter approach due to its efficiency, using the Drude model as our analytic model.

In our work on SSH chains of metallic nanoresonators, we performed frequency-domain simulations using the eigenmode solver of COMSOL Multiphysics [188]. We rely on COMSOL models using the weak-field approach from Wu et al. [189] to build our simulations. COMSOL Multiphysics solves the linear eigenvalue problem

$$\nabla \times \mu_0^{-1} \nabla \times \vec{E}_m - \omega_m^2 \epsilon(r, \omega_m) \vec{E}_m = 0. \tag{3.20}$$

This is fine for non-dispersive systems; however, in our case, where the system is dispersive, Eq. 3.20 becomes non-linear. As an example, we consider a Lorentz model for the dielectric permittivity and demonstrate how to rewrite Eq. 3.20 as a linear eigenvalue problem [190]. The differential equation for the induced polarization is

$$\frac{\partial^2 \tilde{P}(\vec{r}, t)}{\partial t^2} + \Gamma \frac{\partial \tilde{P}(\vec{r}, t)}{\partial t} + \omega_0^2 \tilde{P}(\vec{r}, t) = \bar{\omega}_p^2 \epsilon_\infty \tilde{E}(\vec{r}, t). \tag{3.21}$$

where $\bar{\omega}_p = \frac{1}{\sqrt{\epsilon_\infty}} \omega_p$, for plasma frequency ω_p, ω_0 is a material resonance frequency, and Γ is a damping factor. We then define

$$\tilde{J}(\vec{r}, t) = \frac{\partial \tilde{P}(\vec{r}, t)}{\partial t}. \tag{3.22}$$

This gives

$$\frac{\partial \tilde{J}(\vec{r}, t)}{\partial t} + \Gamma \tilde{J}(\vec{r}, t) + \omega_0^2 \tilde{P}(\vec{r}, t) = \bar{\omega}_p^2 \epsilon_\infty \tilde{E}(\vec{r}, t). \tag{3.23}$$

Note that for $\omega_0 = 0$, this is the Drude model. The Maxwell equations for these auxiliary

fields for a non-magnetic system can be written in the form

$$\hat{\mathcal{H}} \left|\tilde{\psi}(\vec{r},t)\right\rangle = i\frac{\partial}{\partial t}\left|\tilde{\psi}(\vec{r},t)\right\rangle, \tag{3.24}$$

where

$$\hat{\mathcal{H}} = \begin{bmatrix} 0 & -i\mu_0^{-1}\nabla\times & 0 & 0 \\ i\epsilon_\infty^{-1}\nabla\times & 0 & 0 & -i\epsilon_\infty^{-1} \\ 0 & 0 & 0 & i \\ 0 & i\bar{\omega}_p^2\epsilon_\infty & -i\omega_0^2 & -i\Gamma \end{bmatrix}, \tag{3.25a}$$

$$\left|\tilde{\psi}(\vec{r},t)\right\rangle = \begin{bmatrix} \tilde{H}(\vec{r},t) \\ \tilde{E}(\vec{r},t) \\ \tilde{P}(\vec{r},t) \\ \tilde{J}(\vec{r},t) \end{bmatrix}. \tag{3.25b}$$

In frequency space, this becomes

$$\hat{\mathcal{H}} \left|\tilde{\psi}_m(\vec{r})\right\rangle = \tilde{\omega}_m \left|\tilde{\psi}_m(\vec{r})\right\rangle, \tag{3.26}$$

where we have that $\tilde{\omega}_m \in \mathbb{C}$. The orthonormality is defined according to the inner product

$$\begin{aligned}\left\langle\tilde{\psi}_m\right| \hat{D} \left|\tilde{\psi}_n\right\rangle = \int_V \Bigg[&\epsilon_\infty \tilde{E}_m(\vec{r})\cdot\tilde{E}_n(\vec{r}) - \mu_0\tilde{H}_m(\vec{r})\cdot\tilde{H}_n(\vec{r}) + \frac{\omega_0^2}{\epsilon_\infty\bar{\omega}_p^2}\tilde{P}_m(\vec{r})\cdot\tilde{P}_n(\vec{r}) \\ &-\frac{1}{\epsilon_\infty\bar{\omega}_p^2}\tilde{J}_m(\vec{r})\cdot\tilde{J}_n(\vec{r})\Bigg] d^3r = \delta_{mn},\end{aligned} \tag{3.27}$$

where we have

$$\hat{D} = \begin{bmatrix} -\mu_0 & 0 & 0 & 0 \\ 0 & \epsilon_\infty & 0 & 0 \\ 0 & 0 & \frac{\omega_0^2}{\epsilon_\infty \bar{\omega}_p^2} & 0 \\ 0 & 0 & 0 & \frac{-1}{\epsilon_\infty \bar{\omega}_p^2} \end{bmatrix}. \tag{3.28}$$

Through direct multiplication, we can see that

$$\hat{\mathcal{H}}^\dagger \hat{D} = \hat{D}\hat{\mathcal{H}}^*. \tag{3.29}$$

It can be shown [74] that $\langle \tilde{\psi}_m(\vec{r})| \hat{D}$ is the adjoint eigenstate of $|\tilde{\psi}_n\rangle$. If we call $\langle \tilde{\varphi}_n| = \langle \tilde{\psi}_m| \hat{D}$, we have $\langle \tilde{\varphi}_m | \tilde{\psi}_n \rangle = \langle \tilde{\psi}_m| \hat{D} |\tilde{\psi}_n\rangle = \delta_{mn}$.

We use a QNM Eigenmode Solver, developed from [190], with implementation discussed in [189]. We convert Eq. 3.26 into a quadratic eigenvalue problem by noting that

$$\tilde{H}_m = -\frac{i}{\tilde{\omega}_m} \mu_0^{-1} \nabla \times \tilde{E}_m, \tag{3.30a}$$

$$\tilde{J}_m = -i\tilde{\omega}_m \tilde{P}_m. \tag{3.30b}$$

Substitute Eq. 3.30a-3.30b into Eq. 3.26, get

$$\nabla \times \mu_0^{-1} \nabla \times \tilde{E}_m - \omega_m^2 \epsilon_\infty \tilde{E}_m - \omega_m^2 \tilde{P}_m = 0, \tag{3.31a}$$

$$\epsilon_\infty \bar{\omega}_p^2 \tilde{E}_m - \omega_0^2 \tilde{P}_m + i\gamma\tilde{\omega}_m \tilde{P}_m + \tilde{\omega}_m^2 \tilde{P}_m = 0. \tag{3.31b}$$

If we set

$$\vec{x}_m = \begin{bmatrix} \tilde{E}_m \\ \tilde{P}_m \end{bmatrix}, \tag{3.32}$$

then we have

$$\hat{L}\vec{x}_m + \tilde{\omega}_m \hat{R}\vec{x}_m + \tilde{\omega}_m^2 \hat{M}\vec{x}_m = 0, \tag{3.33}$$

where

$$\hat{L} = \begin{bmatrix} \nabla \times \mu_0^{-1} \nabla \times & 0 \\ \epsilon_\infty \bar{\omega}_p^2 & -\omega_0^2 \end{bmatrix}, \tag{3.34a}$$

$$\hat{R} = \begin{bmatrix} 0 & 0 \\ 0 & i\gamma \end{bmatrix}, \tag{3.34b}$$

$$\hat{M} = \begin{bmatrix} -\epsilon_\infty & -1 \\ 0 & 1 \end{bmatrix}. \tag{3.34c}$$

In order to use COMSOL to compute the QNMs, we write Eq. 3.31a and Eq. 3.31b as

$$\int_V \left[\left(\nabla \times \vec{T}(\vec{r}) \right) \cdot \left(\mu_0^{-1} \nabla \times \tilde{E}_m \right) - \omega_m^2 \epsilon_\infty \vec{T}(\vec{r}) \cdot \tilde{E}_m - \omega_m^2 \vec{T}(\vec{r}) \cdot \tilde{P}_m \right] d^3r = 0, \tag{3.35a}$$

$$\int_V \left[\epsilon_\infty \bar{\omega}_p^2 \vec{T}(\vec{r}) \cdot \tilde{E}_m - \omega_0^2 \vec{T}(\vec{r}) \cdot \tilde{P}_m + i\gamma \tilde{\omega}_m \vec{T}(\vec{r}) \cdot \tilde{P}_m + \tilde{\omega}_m^2 \vec{T}(\vec{r}) \cdot \tilde{P}_m \right] d^3r = 0. \tag{3.35b}$$

Eq. 3.35a and Eq. 3.35b are known as weak-form PDEs, where $\vec{T}(\vec{r})$ is called a test function. We performed convergence tests for spherical and cuboid nanoparticles to determine the

best mesh parameters sufficient for our computational resources.

Using the Drude parameters discussed in Section 2.7, we look at two general situations: (1) eigenfrequencies when Bloch boundary conditions are implemented and (2) eigenfrequencies when we have open-boundary conditions. In both cases, we looked at two different SSH chains, nanoresonators made from: (1) spherical particles of radius 5 nm and (2) cuboid particles of dimensions 14 nm by 14 nm by 25 nm (lengths in x, y, and z directions, respectively). Both setups had a periodicity of 50 nm. We chose a background of 100 nm for depth and height. When we compute bandstructures, we enforce Bloch boundary conditions using Floquet periodicity in the x-direction, while PML boundaries are placed in the y and z directions. In order to improve the accuracy of the fields, we included a thin layer around each particle, setting the index of this region to be the same as the index of the background.

In order to establish consistency with bulk-boundary correspondence, we performed several sweeps of inter/intraparticle distances to search for edge modes in the open-boundary case and to compute the Zak phase. We also computed coupling strengths for the semi-analytic approach to construct our theoretical model.

3.2 Data-Driven Effective Medium Approximation

Different ANNs can have different numbers of neurons, layers, initializations, and shapes. Each of these is called an architecture. One can train and deploy multiple architectures

to work on the same kind of data, creating what are called ensemble networks. In the metamaterial grating work, we used an ensemble of neural networks [92, 191, 192], where we take an equally weighted average of the predictions of each member of the ensemble.

In our work, we use a tool developed by Google known as TensorFlow to build our models. Specifically, we used a Python wrapper known as Keras, which is a high-level version of TensorFlow.

3.2.1 TensorFlow

TensorFlow [193] is an open-source tool developed by Google to build, train, and deploy machine learning models. It is based on the idea of a tensor moving/flowing through a computational graph. Tensors can be thought of as generalizations of scalars, vectors, and matrices. A computational graph [194] is a structure with nodes and edges. The nodes represent operations such as addition, multiplication, etc or values such as parameters, inputs, etc., while the edges represent the flow of the data between these operations/values. Fig. 3.3 depicts a simple computational graph for the mathematical operations involved in computing $z = xy + b$. The nodes represent the variables x, y, b, z. The graph shows, starting from the left, that we take two variables x, y and multiply them. Then, once this is done, we add b to the result, which gives us z. Notice that the computational graph breaks down the operations step-by-step. This makes it easier to compute the gradients during backpropagation. This is especially important for large ANNs, which can have millions of parameters, making gradient descent approaches inefficient, especially for large-scale

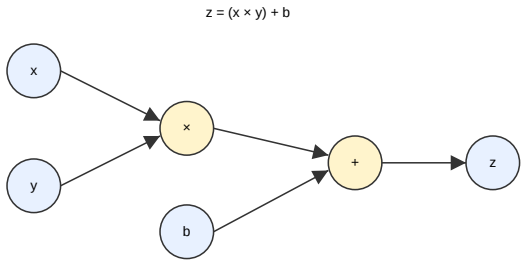


Figure 3.3: Illustration of a simple computational graph for the mathematical operations involved in computing $z = xy + b$.

models. This is one of the major strengths of TensorFlow and why we chose it at the time. TensorFlow can be complicated to use, making it difficult to build and deploy models, and it is more prone to user errors. Its coding structure is quite verbose and time-consuming. As a result, we turned to a high-level API for TensorFlow known as Keras [85], which was developed to make coding with TensorFlow much simpler.

3.2.2 Data-Driven Effective Medium on SWG Designs

In Chapter 4, we introduce an optimization procedure to obtain an effective refractive index of an SWG that is data-driven, using DNNs as surrogate models that were trained on FDTD simulation data. We use the DNNs as surrogate models for FDTD simulations because FDTD simulations would have been much too slow to use in the optimization procedure. Whereas a single FDTD simulation for a 3D geometry might take at least ten minutes, the DNN surrogate models for the 3D system will take less than a minute. The 3D and 2D surrogate ensemble models predict the forward and vertical transmissions,

along with the reflection, for a particular geometry and for different 2D effective indices at a given wavelength. The effective index corresponding to the 2D prediction that most closely matches the 3D prediction for the given wavelength is the effective index for the 2D structure with the best optical match to the 3D structure. We discuss this in more detail in Section 3.2.

Using Keras, we trained DNNs on 2D FDTD and 3D FDTD Lumerical simulations of the SWG structures. We created 3 DNNs for the 2D models and 3 DNNs for the 3D models. We then produced an ensemble network [92, 191, 192] for the 2D-trained structures and another for the 3D-trained structures. Once we had our models built with Keras, we built a data-driven effective medium theory that uses grid-search optimization from the predictions of our ensemble models for 2D and 3D structures to generate a dispersive effective index profile that takes into account scattering in the vertical and horizontal directions. One may adjust the parameters to create the models. These could be the architecture size or shape. One may also use different initialization seeds. These models may then be used together to create an ensemble of neural networks. One then combines the predictions of each network to make a single prediction. This can be done, for example, by doing a weighted average of the predictions of each model. Once we have a working model or ensemble, we can make predictions quite quickly - often in less than a second. In our case, we adjust the architecture and create three models.

The three models for 2D and 3D have the architectures M_1, M_2, and M_3 shown in Table 3.1, which presents only the hidden layers. The DNNs for each model are based on a kite-like architecture, which was found to be the simplest architecture suited for our

M_1 :	7	200	500	600	700	800	800	800	700	600	500	400	200	100	50	20
M_2 :	7	200	500	1200	1400	1600	1600	1600	1400	1200	500	400	200	100	50	20
M_3 :	7	200	500	1200	1400	1600	1600	1600	1400	1200	500	400	50	20		

Table 3.1: Architecture of the hidden layers of the three DNN models used for the training of three distinct 2D and three distinct 3D forward models. Each row lists the number of neurons in the hidden layer at that position, where the input layer (not shown) would be on the left, and the output layer (not shown) would be on the right.

problem. Fig. 3.4 is a visualization of M_1 for the 2D model. The input parameters

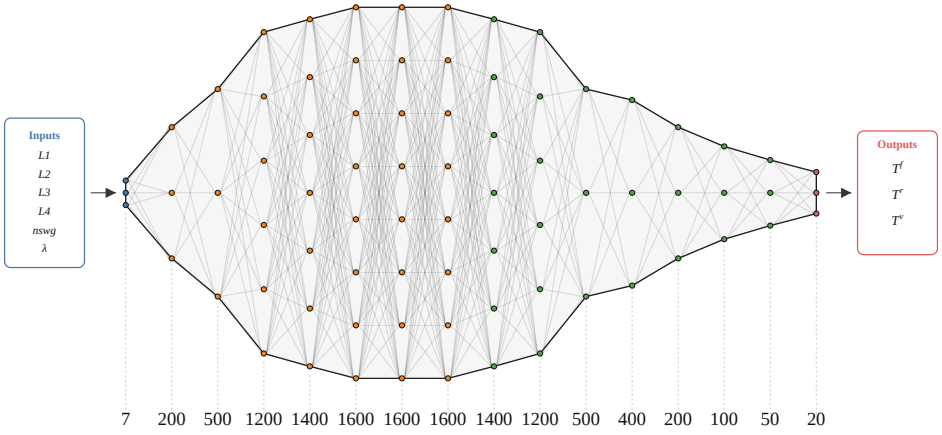


Figure 3.4: Illustration of one DNN architecture for 2D. The points are DNNs, and the lines are the connectivity.

for M_1, M_2 and M_3 are presented in Table 3.2. For the 2D models, we have 6 inputs: L_1, L_2, L_3 and L_4, which specify the geometry of the 2D structures; n_{nswg}, the effective index of the layer of length L_2; and the wavelength, λ. The outputs of the model are the forward transmission (T^f), reflection (T^r), and vertical transmission (T^v), respectively. In other words, let $\mathcal{I}^{2D} = \{L_1, L_2, L_3, L_4, n_{\text{swg}}, \lambda\}$ be the input of the 2D models, and

Parameter	**2D**	**3D**
Λ [nm]	400 to 1000	400 to 1000
L_1, L_2, L_3, L_4 [nm]	40 to Λ, $\sum L_i = \Lambda$	40 to Λ, $\sum L_i = \Lambda$
a_z [nm]	N/A	100 to 650
Λ_z [nm]	N/A	40 to $a_z - 40$
n	1.4 to 3.5	N/A
λ [nm]	1450 to 1650	1450 to 1650

Table 3.2: Input parameter information for DNNs. Λ is the period in the x-direction, which is the sum of L_1 to L_4, the lengths of the four segments in each unit cell. In the 3D case, Λ_z is the SWG pitch in the z-direction, and a_z is the SWG silicon width. In the 2D case, n_{swg} is the effective index of the segment in the unit cell that replaces the SWG, while λ is the wavelength of the incident light.

$\mathcal{O}_i^{2D} = \{T_i^f, T_i^r, T_i^v\}$ be the output of the said models. Then

$$M_i : \mathcal{I}^{2D} \to \mathcal{O}_i^{2D}, \tag{3.36}$$

where i labels the model number, which in our case is a different architecture. We have, more concretely

$$T_i^{(d)} = M_i^{2D}(L_1, L_2, L_3, L_4, n_{\text{swg}}, \lambda), \tag{3.37}$$

where $d \in \{f, r, v\}$.

Fig. 3.5 shows M_1 for the 3D case, where there are 7 inputs: L_1, L_2, L_3, L_4, pitch, and the length of the silica component in the grating section (along L_2),L_{SiO2}, which are the geometrical parameters of the 3D structures, and the wavelength, λ. In other words, let $\mathcal{I}^{3D} = \{L_1, L_2, L_3, L_4, \text{Pitch}, L_{\text{SiO2}}, \lambda\}$ be the input of the 3D models, and $\mathcal{O}_j^{3D} =$

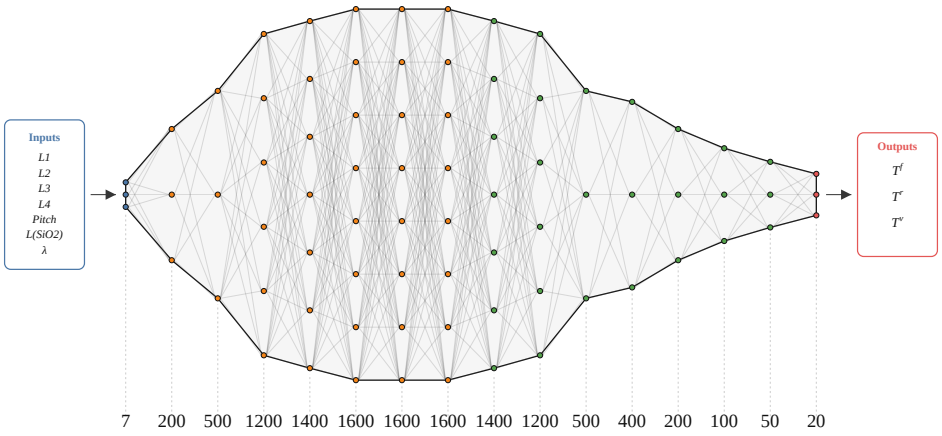


Figure 3.5: Illustration of one ANN architecture for 3D. The points are ANNs, and the lines are the connectivity.

$\{T_j^f, T_j^r, T_j^v\}$ be the output of the said models. Then, as before, we have

$$M_j : \mathcal{I}^{3D} \to \mathcal{O}_j^{3D}, \tag{3.38}$$

where j labels the model number. We have, more concretely

$$T_j^{(d)} = M_j^{3D}(L_1, L_2, L_3, L_4, \text{Pitch}, L_{\text{SiO2}}, \lambda), \tag{3.39}$$

where $d \in \{f, r, v\}$.

For the ensemble models, we take an equally weighted average of our solutions, i.e.

$$T_{2D}^{(d)} = \frac{1}{3}\sum_{i=1}^{3} M_i^{2D}, \tag{3.40a}$$

$$T_{3D}^{(d)} = \frac{1}{3}\sum_{j=1}^{3} M_j^{3D}. \tag{3.40b}$$

Now, let $\mathbb{O}$ be the collection of all the transmission information from $\mathcal{O}^{2D} = \{T_{2D}^f, T_{2D}^r, T_{2D}^v\}$ and $\mathcal{O}^{3D} = \{T_{3D}^f, T_{3D}^r, T_{3D}^v\}$ such that $\mathcal{O}_i^{2D}, \mathcal{O}_j^{3D} \in \mathbb{O}$, i.e. $\mathbb{O}$ is a set of sets. In our work, we are saying that at each wavelength and $\forall \mathcal{O}^{3D}$, there is at least one n_{swg}, so that $\mathcal{O}^{2D} = \mathcal{O}^{3D}$. We implemented an optimization process that demonstrates this by calculating an error function on a grid and selecting the optimal point on the grid (i.e., that minimizes the error). The optimization procedure allows us to identify candidate effective refractive indices as a function of T^f, T^v and T^r separately. However, not all candidates obtained are physically realistic. Selecting a physically realistic solution amounts to identifying where the candidate solutions of the effective indices for T^f, T^v and T^r overlap. Then, finally, we fit a polynomial to this overlap, giving us a single dispersive index, which, at least in some instances, is valid for all three directions.

3.3 Theoretical and Computational Methods for Topological Plasmonics

In this section, we discuss our approach to topological plasmonics in the context of nanoresonators. First, in Section 3.3.1, we discuss the numerical implementation of the Wilson loop to compute the Zak phase. This is relevant for the results presented in Chapter 6. Next, in Section 3.3.2, we present the QNM coupled-mode theory, which was important for the work presented in Chapter 5.

3.3.1 Numerical Implementation of Zak Phase using the Non-Hermitian Wilson Loop

Recall that Eq. 2.48 and Eq. 2.54 are used to compute the Hermitian Wilson loop for the non-degenerate and degenerate cases. The non-Hermitian Wilson loop is calculated in a similar manner, only now, left ($\langle\tilde{\varphi}(\vec{r})|$) and right ($|\tilde{\psi}(\vec{r})\rangle$) eigenvectors are used such that $(\bar{O}_{0,1})_{mn} = \langle\tilde{\varphi}_{mk_0}(\vec{r})|\tilde{\psi}_{nk_1}(\vec{r})\rangle$, $(\bar{O}_{1,2})_{mn} = \langle\tilde{\varphi}_{mk_1}(\vec{r})|\tilde{\psi}_{nk_2}(\vec{r})\rangle$, ..., $(\bar{O}_{L-2,L-1})_{mn} = \langle\tilde{\varphi}_{nk_{L-2}}(\vec{r})|\tilde{\psi}_{nk_{L-1}}(\vec{r})\rangle$. We note that for us, $\langle\tilde{\varphi}| = \langle\tilde{\psi}|\,\hat{D}$, where $\hat{D}$ is defined as in Eq. 3.28.

Using this non-Hermitian Wilson loop, we can compute the Zak phase, γ. We summarize how this is done in Algorithm 3. We compute the bandstructure in COMSOL over a certain number of k-points. For each k-point, we compute the overlaps $\bar{O}_{mn}$, then

accumulate the product. We get

$$\hat{W}^{\mathrm{NH}} = \prod_{l=1}^{L} (\bar{O}_{l,l+1})_{mn}, \tag{3.41}$$

so that the non-Hermitian Zak phase, ν^{NH}, is

$$\nu^{NH} = \frac{i}{\pi} \ln[\hat{W}^{\mathrm{NH}}], \tag{3.42}$$

Algorithm 3 Numerical Evaluation of the Zak Phase via Wilson Loop.

1: Choose your band.
2: Pick a closed path in k-space.
3: Discretize the path into L segments of equal length δk: you will have a finite number of k values.
4: **for** $l = 0$ to $L - 1$ **do** (i.e. loop through all the k values of the chosen band)
5: Compute the overlap

$$\bar{O}_{l,l+1} = \langle \tilde{\varphi}_{nk_l}(\vec{r}) \mid \tilde{\psi}_{nk_{L+1}}(\vec{r}) \rangle$$

.
6: Accumulate the product

$$\bar{\mathcal{O}} \leftarrow \bar{\mathcal{O}} \cdot \bar{O}_{l,l+1}$$

.
7: **end for**
8: Impose periodic boundary condition: $k_L \equiv k_0$, so that you close the loop.
9: Compute Zak phase.

$$\gamma \approx i \ln \mathcal{O}$$

.

It appears that a deep mathematical understanding of the non-Hermitian Wilson loop is still in development. In Section 8.2, we sketch a potential direction to put it on a more rigorous footing. Despite this, computing the non-Hermitian Zak phase has been done before; see Pocock et al. [73].

3.3.2 QNM Coupled-Mode Theory

In Section 2.8, we have established key features of the QNMs; we now move to the coupled-mode theory from Tao et al. [195]. This allows us to calculate the resonant modes and frequencies of only one particle and use them to reconstruct the resonant modes and frequencies of the full system, without simulating the entire system. Beyond this, it presents a recipe to estimate the coupling strengths between the metallic nanoparticles, as well as the contributions of individual resonances of a single particle to the hybridized modes of the entire system.

In this work, there are three key assumptions:

1. We can truncate the physical space using absorbing boundaries, perfectly matched layers (PML) and get physically meaningful results for the near-field.

2. In principle, there are an infinite number of modes, including physical and numerical (PML) modes, which are needed to form a complete basis. We assume that truncating to a finite number of modes, particularly a finite number of physical modes and neglecting the PML modes altogether, is sufficient to get physically meaningful results

pertinent to our problem.

3. We can truncate the coupling to a finite range and reconstruct the physically meaningful results relevant to us.

The modes of the full system can be written as a superposition of the modes of all individual particles, that is

$$|\tilde{\Psi}_l(\vec{r})\rangle = \sum_p \sum_m a_{p,m} |\tilde{\psi}_{p,m}(\vec{r})\rangle, \tag{3.43}$$

where $|\tilde{\Psi}_l(\vec{r})\rangle$ is obtained in COMSOL by solving the auxiliary eigenvalue problem of the full lattice. Computing the inner product $\langle\tilde{\psi}_{p,m}(\vec{r})|\hat{D}|\tilde{\psi}_l(\vec{r})\rangle$, we obtain the expansion coefficients $a_{p,m}$ quantifying the contribution of a particular mode, m, for a specific particle p, to the full system's l^{th} mode. This enables a modal decomposition of the particle, allowing us to determine, for example, if there are dipolar-like, quadrupolar-like, or octupolar-like components. It tells us which of these modes in the basis expansion are most dominant.

We now derive the coupled-mode expressions as in Tao et al. [195]. We will start with the Maxwell equations in the frequency domain with a current source, $\vec{J}(\vec{r},\omega)$, and free-space permeability

$$\nabla \times \vec{E}(\vec{r},\omega) = i\omega\mu_0 \vec{H}(\vec{r},\omega), \tag{3.44a}$$

$$\nabla \times \vec{H}(\vec{r},\omega) = -i\omega\epsilon(\vec{r},\omega)\vec{E}(\vec{r},\omega) + \vec{J}(\vec{r},\omega), \tag{3.44b}$$

where $\vec{E}(\vec{r},\omega)$ and $\vec{H}(\vec{r},\omega)$ are the electric and magnetic fields with frequency dependence,

ω. If we assume the electric and magnetic field modes

$$|\tilde{\Psi}(\vec{r},\omega)\rangle = \begin{bmatrix} \vec{E}(\vec{r},\omega) \\ \vec{H}(\vec{r},\omega) \end{bmatrix}, \tag{3.45}$$

form a complete basis set such that we can write any field $|\tilde{\Psi}(\vec{r},\omega)\rangle$ as

$$|\tilde{\Psi}(\vec{r},\omega)\rangle \approx \sum_m a_m(\omega)|\tilde{\psi}_m(\vec{r})\rangle, \tag{3.46}$$

from Sauvan et al. [165], we know that

$$a_m(\omega) = \frac{-i}{(\omega-\tilde{\omega}_m)N_m}\iiint_\Omega \vec{J}(\vec{r},\omega)\cdot\vec{E}_m(\vec{r}), \tag{3.47}$$

where N_m is defined as in Eq. 2.77, Ω represents the entire volume. The field $|\tilde{\Psi}(\vec{r},\omega)\rangle$ can be written as $|\tilde{\Psi}(\vec{r},\omega)\rangle = |\tilde{\Psi}^s(\vec{r},\omega)\rangle + |\tilde{\Psi}^b(\vec{r},\omega)\rangle$, where $|\tilde{\Psi}(\vec{r},\omega)^s\rangle$ is the scattered contribution and $|\tilde{\Psi}(\vec{r},\omega)^b\rangle$ is the background contribution.

We work with the electric fields alone, as they typically have larger magnitude. Noting that $\Delta\epsilon_p(\vec{r},\omega) = \epsilon_p(\vec{r},\omega) - \epsilon_b$, with $\epsilon_p(\vec{r},\omega)$ being the permittivity from the Drude model of particle, p. Meanwhile, ϵ_b is the permittivity of the medium surrounding the particles, treating the fields of all resonators except p as background fields and using Eq. 3.47, we

have

$$a_{p,m}(\omega) = \frac{-\omega \int_{V_p} \vec{E}_{p,m}(\vec{r}) \cdot \Delta\epsilon_p(\vec{r},\omega) \cdot \left[\sum_{q\neq p} \vec{E}^s_q(\vec{r},\omega) + \vec{E}^b(\vec{r},\omega)\right] d\vec{r}}{(\omega - \tilde{\omega}_{p,m})N_{p,m}}, \tag{3.48}$$

where V_p is the volume of resonator p. From this, we obtain

$$a_{p,m}(\omega) = \frac{-\omega \sum_{q\neq p}\sum_{r=1}^{M_q} a_{q,r}(\omega) \int_{V_p} \vec{E}_{p,m}(\vec{r}) \cdot \Delta\epsilon_p(\vec{r},\omega) \cdot \vec{E}_{q,r}(\vec{r}) d\vec{r}}{(\omega - \tilde{\omega}_{p,m})N_{p,m}} + \frac{-\omega \int_{V_p} \vec{E}_{p,m}(\vec{r}) \cdot \Delta\epsilon_p(\vec{r},\omega) \cdot \vec{E}^b(\vec{r},\omega) d\vec{r}}{(\omega - \tilde{\omega}_{p,m})N_{p,m}}, \tag{3.49}$$

with $N_{p,m}$ being the normalization constant defined as

$$N_{p,m} = \int_{\mathbb{R}^3} \left[\tilde{\vec{E}}_{p,m} \cdot \left.\frac{\partial(\omega\epsilon_p)}{\partial\omega}\right|_{\omega=\tilde{\omega}_{p,m}} \cdot \tilde{\vec{E}}_{p,m} - \mu_0 \tilde{\vec{H}}_{p,m} \cdot \tilde{\vec{H}}_{p,m}\right] \mathrm{d}^3 r. \tag{3.50}$$

Within equation Eq. 3.49, we can pull out the following factors for each term

$$\kappa_{(p,m)(q,r)}(\omega) = \frac{-\omega}{N_{p,m}} \int_{V_p} \vec{E}_{p,m}(\vec{r}) \cdot \Delta\epsilon_p(\vec{r},\omega) \cdot \vec{E}_{q,r}(\vec{r}) d\vec{r}, \tag{3.51a}$$

$$\bar{\beta}_{p,m}(\omega) = \frac{-\omega}{N_{p,m}} \int_{V_p} \vec{E}_{p,m}(\vec{r}) \cdot \Delta\epsilon_p(\vec{r},\omega) \cdot \vec{E}^b(\vec{r},\omega) d\vec{r}. \tag{3.51b}$$

which quantify the near field coupling between single particle modes Eq. 3.51a, and the decomposition of the background field with respect to the single particle modes Eq. 3.51b. If there is no externally applied field, then we can set $\vec{E}^b(\vec{r},\omega) = 0$, which gives the

inhomogeneous linear system of equations

$$0 = (\omega - \tilde{\omega}_{p,m})\, a_{p,m}(\omega) - \sum_{q \neq p} \sum_{n=1}^{M_q} \kappa_{(p,m),(q,n)}(\omega)\, a_{q,n}(\omega), \tag{3.52}$$

with ω being an arbitrary frequency, $\tilde{\omega}_{p,m}$ the on-site frequency and $a_{p,m}$ the coefficients. Collecting all coefficients, $a_{p,m}$ into $\vec{a}$ gives

$$[\omega \mathbb{I} - \mathbb{K}(\omega)]\, \vec{a} = 0, \tag{3.53}$$

where the diagonal entries of $\mathbb{K}$ are $\tilde{\omega}_{p,m}$ and the off-diagonal entries are the couplings $\kappa_{(p,m),(q,n)}(\omega)$. In the weak coupling limit, $\mathbb{K}(\omega) \approx \mathbb{K}$. We thus have

$$[\omega \mathbb{I} - \mathbb{K}]\, \vec{a} = 0. \tag{3.54}$$

We use Eq. 3.54 for our coupled mode theory, since Eq. 3.53 leads to transcendental equations, which are more complicated to solve.

We compute the near-field couplings, Eq. 3.51a, for the weak-coupling scenario, Eq. 3.54, using Algorithm 4, which uses the Gauss-Legendre approach to integration. We find the overlap integral

$$S_{mn} = \frac{1}{N_{A,m}} \int_{V_A} \vec{E}_m^{(A)}(\vec{r}) \cdot \vec{E}_n^{(B)}(\vec{r}) dV_A, \tag{3.55}$$

quantifies the influence of particle B at mode n on particle A at mode m. Notice the integral is not evaluated over the entire space, but only within the volume of particle, A.

We describe the full process in Pseudocode 4, where we apply the field of B at particle A's location (though in our work all particles are identical so that the same simulated modes can be used for both particles). We repeat this for different distances d between particles A and B, and different pairs of modes.

Algorithm 4 Computation of QNM Overlaps used to compute the coupling strength between resonators.

Get eigenfrequencies and eigenmodes $\{\omega_m, \vec{E}_m(\vec{r})\}$ from COMSOL.
Select modes of interest.
Choose subset $\{i_m\}_{m=1}^{i}$.
Extract fields.
for each mode m **do**
    Extract electric field $\vec{E}_m = (E_x, E_y, E_z)$.
    Extract normalization N_m.
end for
Compute overlap for each distance.
for each distance d **do**
    Shift coordinates: $\vec{r} \to \vec{r} + (d, 0, 0)$.
    Compute shifted fields $\vec{E}_n^{(B)}(\vec{r})$.
    for each mode pair (m, n) **do**,
        Compute dot product:

$$\vec{E}_m^{(A)} \cdot \vec{E}_n^{(B)} = E_{x,m}E_{x,n} + E_{y,m}E_{y,n} + E_{z,m}E_{z,n}$$

.
        Compute overlap:

$$S_{mn}(d) = \int \vec{E}_m^{(A)} \cdot \vec{E}_n^{(B)} \, dV_A$$

.
        Apply normalization:

$$S_{mn}(d) \leftarrow S_{mn}(d) \cdot N_m$$

.
    end for
end for
Post-processing.
Extract:

$$\mathrm{Re}[S], \quad \mathrm{Im}[S], \quad |S|$$

.

We then use Eq. 3.51a for the weak-coupling regime, which gives

$$\kappa_{(p,m)(q,n)} = \frac{-\tilde{\omega}_{p,m}}{N_{p,m}} S_{mn}. \tag{3.56}$$

If we explicitly look at the entries for $\mathbb{K}$ for a periodic SSH chain with two particles A and B in a unit cell, we have

$$\mathbb{K} = \begin{bmatrix} \tilde{\mathbf{\Omega}}^{A} & \mathbb{V}^{AB} + \mathbb{W}^{AB} e^{-ika} \\ \mathbb{V}^{BA} + \mathbb{W}^{BA} e^{ika} & \tilde{\mathbf{\Omega}}^{B} \end{bmatrix}, \tag{3.57}$$

where $\mathbf{\Omega}^{A}(\mathbf{\Omega}^{B})$ is a matrix containing the N most relevant resonant frequencies of A (B) on the diagonal, with zero everywhere else. We can write $\mathbf{\Omega}^{A}$ and $\mathbf{\Omega}^{B}$ as

$$\mathbf{\Omega}^{A,B} = \begin{bmatrix} \tilde{\omega}^{(1)}_{A,B} & 0 & \cdots & 0 \\ 0 & \tilde{\omega}^{(2)}_{A,B} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \tilde{\omega}^{(N)}_{A,B} \end{bmatrix}. \tag{3.58}$$

$\mathbb{V}^{AB}$ is a matrix that contains the nearest neighbor coupling coefficients κ (Eq. 3.56) between all N relevant modes of A with all N relevant modes of B such that A and B are not in the same unit cell (intercell). Whereas the matrix $\mathbb{W}^{AB}$ contains the nearest neighbor coupling coefficients κ between the N modes of A and B such that A and B are

in the same unit cell (intracell). They can be written as

$$\mathbb{V}^{AB} = \begin{bmatrix} V_{11}^{AB} & V_{12}^{AB} & \cdots & V_{1N}^{AB} \\ V_{21}^{AB} & V_{22}^{AB} & \cdots & V_{2N}^{AB} \\ \vdots & \vdots & \ddots & \vdots \\ V_{N1}^{AB} & V_{N2}^{AB} & \cdots & V_{NN}^{AB} \end{bmatrix}, \tag{3.59a}$$

$$\mathbb{W}^{AB} = \begin{bmatrix} W_{11}^{AB} & W_{12}^{AB} & \cdots & W_{1N}^{AB} \\ W_{21}^{AB} & W_{22}^{AB} & \cdots & W_{2N}^{AB} \\ \vdots & \vdots & \ddots & \vdots \\ W_{N1}^{AB} & W_{N2}^{AB} & \cdots & W_{NN}^{AB} \end{bmatrix}. \tag{3.59b}$$

This allows us to go beyond the traditional two-band model, where we can now incorporate N bands single particle modes (leading to N bands for infinite SSH chains.)

If we have a finite SSH chain of identical particles, then we just have

$$\mathbb{K} = \begin{bmatrix} \mathbf{\Omega}_1^A & \mathbb{V}_1^{AB} & 0 & 0 & \cdots & 0 \\ \mathbb{V}_1^{BA} & \mathbf{\Omega}_1^B & \mathbb{W}_1^{BA} & 0 & \cdots & 0 \\ 0 & \mathbb{W}_1^{AB} & \mathbf{\Omega}_2^A & \mathbb{V}_2^{AB} & \cdots & 0 \\ 0 & 0 & \mathbb{V}_2^{BA} & \mathbf{\Omega}_2^B & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \mathbb{V}_N^{AB} \\ 0 & 0 & 0 & \cdots & \mathbb{V}_N^{BA} & \mathbf{\Omega}_N^B \end{bmatrix}. \tag{3.60}$$

4

Data-Driven Effective Medium Approximation for Metamaterials in Device Designs

SWGs are attractive for technological applications as they allow for the engineering of complex electromagnetic interactions through periodic nanostructures. Conventional effective medium theories (EMT) have traditionally been used to model the optical properties of SWGs using an effective refractive index, with the assumption that SWGs are infinite. However, practical photonic devices are finite and strongly influenced by their electromag-

netic environment, making conventional EMTs insufficient. In this chapter, we present a data-driven effective medium approximation in more detail and present the results. Some of this was already discussed in chapter 3. This work was done in collaboration with the National Research Council (NRC), Canada. The main contribution of this work is that it extends the concept of effective medium to finite metamaterial gratings, in particular, subwavelength gratings (SWGs). Our main goal was to answer the question: can we extend the concept of effective medium to finite structures to reproduce the optical response of a three-dimensional metamaterial? Though the current work focused on a particular design relevant to SWGs used in grating couplers, we sought to explore whether it was possible to obtain more accurate effective descriptions at lower computational cost than repeated 3D simulations.

This project originated from work by our NRC collaborators[1]. The team had been developing perfectly vertical grating couplers for a range of applications, including optical communications and optical phased arrays. To expand the design space for controlling scattering strength, they investigated the incorporation of SWG metamaterials into the grating unit cells. The initial designs were developed using the widely adopted effective-medium approximation, which relates the effective refractive index to the segmentation pitch and duty cycle based on extended periodic structures. However, when several of these devices were experimentally characterized, many failed to perform as predicted by the two-dimensional simulations, exhibiting poor coupling efficiency and very high back-reflection. Interestingly, a small subset of devices performed exceptionally well. It was later

[1]Many thanks to Dr. Dan-Xia Xu for telling me the origin story presented here.

discovered that these structures had significantly different SWG duty cycles than intended, due to a coding error in the design process. Intrigued by this unexpected result, one of the collaborators, Dr. Dan-Xia Xu, began a deeper investigation to understand why these devices performed so well. Through comprehensive three-dimensional simulations, it was found that appropriate duty cycles could indeed yield excellent device performance, but the required values differed substantially from those predicted by the conventional effective-medium model. The revised designs were subsequently validated experimentally. This work ultimately led to the hypothesis that the effective-medium concept can be extended to highly truncated, fully integrated grating structures, provided that the surrounding environmental conditions are properly incorporated when establishing the relationship between the effective index, pitch, and duty cycle.

Rather than deriving an effective index solely through subwavelength averaging, in our approach, the effective response is defined through the equivalence of the transmission properties between three-dimensional and two-dimensional structures. Neural network surrogate models, trained on 2D and 3D FDTD data, accelerate the multidirectional transmission matching involving forward transmission, vertical transmission, and reflection. Using six structures representing varying pitch and varying length of the periodic regions, we showed that multidirectional overlap of the effective index indicates physically consistent effective-index solutions, which also show phase agreement across 2D and 3D optically equivalent structures despite phase not being explicitly included during training of the neural networks. The effective indices show sensitivity to surrounding geometry and are wavelength-dependent. Comparison with conventional slab EMT and three-dimensional

simulations demonstrates accurate performance even on structures where the slab model failed.

For this work, I designed the algorithms to select the geometries for constructing the Lumerical simulations that train the neural networks. I performed the two- and three-dimensional Lumerical simulations used to construct their corresponding neural networks. I also built, tested, and executed the final two-dimensional and three-dimensional neural networks along with their corresponding ensemble models. In addition, I performed the performance assessment of the neural networks. I also developed, tested, and executed an optimization procedure that matches the two- and three-dimensional transmission profiles that generated the effective indices. The method was developed, tested, and validated by me for the different structures. In addition, the data analysis and manuscript writing were done by me. A draft of the manuscript (in preparation) discussing our method and its results is presented in this chapter.

This work formed the foundation of two more papers being led by two other PhD students in the group. One is being led by Qiang Wang, which involved applying this approach to a design that has been built. The other paper, being led by Kieran Watts, involves using active learning to overcome the high cost of training data that the approach presented in this thesis faces.

Data-Driven Effective Medium Approach for Finite Subwavelength Gratings in Complex Environments

Norman S Israel[1,†], D. Blankespoor[1], Qiang Wang[1], D.-X. Xu[2], Y. Grinberg [3] and Lora Ramunno[1,*]

[1]*Department of Physics and Nexus for Quantum Technologies Institute, University of Ottawa, Canada*
[2]*Quantum and Nanotechnology Research Centre, National Research Council of Canada, Canada*
[3]*Digital Technologies Research Center, National Research Council of Canada, Canada*
[†]*nisra058@uottawa.ca*
[*]*lramunno@uottawa.ca*

Abstract: Subwavelength gratings (SWGs) provide a powerful platform for engineering optical properties in integrated photonics. Conventional effective-medium theories (EMTs) assign an effective refractive index to the subwavelength region based only on the internal geometry of the SWG itself. However, SWGs in photonic devices are often finite, radiative, and strongly influenced by their environment, making conventional EMTs increasingly insufficient. We introduce a data-driven effective medium approach for finite SWGs embedded in complex environments. Neural-network surrogate models are trained to predict their optical response, as well as the response of devices where the SWG region is replaced by a homogenous medium. The homogenous refractive index that produces the same optical response as the device with the SWG is taken to be the environment-aware effective refractive index, which we determine as a function of wavelength. We validate our approach with full 3D simulations, and show that it outperforms the slab model EMT.

© 2026 Optica Publishing Group under the terms of the Optica Publishing Group Publishing Agreement

1. Introduction

Integrated photonic devices have attracted significant interest due to their applications in quantum computing [1–3], light detection and ranging (LiDAR) [4], biosensors [5], and photonic neural networks [6], highlighting their importance for modern technologies. Subwavelength gratings (SWGs) expand the design space of integrated photonic devices by allowing the engineering of light-matter interaction using nanoscale geometric dielectric features [7–13]. True to their name, the features of SWGs are on subwavelength length scales, thus they behave as metamaterials with properties tunable via their geometry. SWGs have been employed in directional couplers [14, 15], polarization splitters [16, 17], as well as optical time-delay lines [18]. Their compatibility with industry-standard advanced manufacturing such as complementary metal-oxide-semiconductor (CMOS) technologies makes it possible to take SWG-based designs beyond the lab and into commercial applications [19–21].

Despite the growing application of SWGs, the design and modelling of devices containing SWGs, particularly those of finite size and within complex optical environments, remains challenging. Conventional effective medium theories (EMTs) have been used extensively to model the SWG region as a homogeneous medium with an effective refractive index that is primarily determined by the SWG geometry and material composition [22–25]. While highly successful in modelling many periodic metamaterials, they generally assume simplified propagation conditions and weak environmental dependence, assumptions that are not always applicable. This is because in many photonic devices, optical fields interact strongly with surrounding interfaces and radiative channels. This is typically true for finite and radiative SWGs,

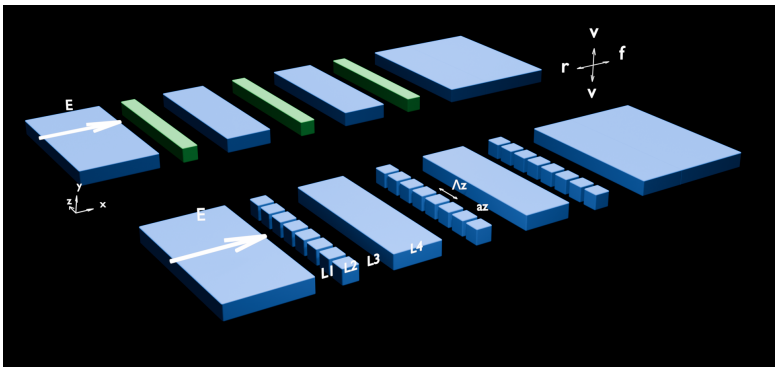


Fig. 1. Bottom: Schematic of three unit cells of a fully etched vertical grating coupler in SOI, where silicon (blue) is embedded within silica (not shown explicitly). Each unit cell contains four segments with lengths L_1 to L_4. The SWG segment (L_2) has pitch Λ_z and silicon width a_z. Top: Similar grating coupler design, but where the SWG segment is replaced by a homogeneous medium (green). Light is incident from the left through a silicon waveguide, as indicated by the arrow. Transmission is considered in the forward (f), vertical (v), and backward (r) directions, as shown.

where the assumptions of conventional EMTs can become invalid. One must then turn to full 3D computational simulations. While these generate accurate results, they often require large computational resources. This is particularly problematic and impractical for the device design process, for which many simulations are typically required.

In this paper, we show that effective medium descriptions remain useful even for finite SWGs in complex environments, enabling faster compute than full 3D simulations that explicitly incorporate each subwavelength feature. We introduce a data-driven effective medium approach (DD-EMA), a framework that constructs environment-aware effective-medium representations for finite, radiative SWG structures embedded in a complex optical environment. Rather than deriving the effective refractive index from an analytical homogenization theory, our DD-EMA defines the effective refractive index *operationally*. That is, for a device containing an SWG region, we seek an alternate scenario where the SWG region is replaced by a homogeneous medium. The effective refractive index of the SWG region is taken to be that for which the *optical response* of the device with the homogenous medium is the same as the device with the actual SWG. Crucially, the resulting effective refractive index is valid *only for that particular device*. The same SWG region embedded within a different device would potentially have a different effective refractive index; as our approach is environment aware, we are able to predict this different effective refractive index.

Data-driven methods employing deep neural networks (DNNs) in surrogate models that replace simulations has gained a lot of interest in recent years in the nanophotonics community [26–29], as they require significantly less computational resources when deployed [30–36]. To make our effective medium approach computationally efficient, our DD-EMA uses DNN-based surrogate models to predict the optical response of the devices of interest. One DNN surrogate model is trained for devices containing the full 3D SWG, and a second for devices containing a homogeneous medium in place of the SWG. This enables rapid and accurate evaluation of candidate effective-index solutions, opening the door to rapid and accurate inverse-design of complex photonic integrated devices containing SWGs.

We demonstrate the effectiveness of our DD-EMA in the design of vertical grating couplers containing SWGs. A major challenge in photonics device packaging is compactness, critical

for mobility and integration. Vertical grating couplers enable small-footprint coupling between integrated photonic circuit waveguides and off-chip components, such as other chips or optical fibers. Grating couplers can be placed at arbitrary positions on-chip, allowing for easy fabrication and facilitating wafer testing [37]. This is in contrast to edge couplers that impose significant restrictions in the placement of device components, and where the horizontal coupling leads to poor fabrication tolerances [37]. We consider here single-step-etch grating couplers in SOI containing SWG regions of finite height and finite longitudinal extent, with periodicity in the transverse direction, as sketched in Fig. 1. Given their finite size, and that grating couplers are designed to maximize out-of-plane radiation, we expect the optical response of the SWGs to depend significantly on their environment.

The paper is organized as follows. Section 2 details the training and validation of DNN-based surrogate models that predict the optical response of the two devices depicted in Fig. 1: one model for devices that contain SWG regions, and another for devices that contain a homogeneous medium in place of the SWG regions. This is done over telecommunications wavelengths and for wide range of the geometrical parameters indicated in Fig. 1. We use an ensemble of DNNs to improve prediction stability and accuracy. In Section 3, we use our surrogate models in an optimization procedure to determine the effective index as a function not only of wavelength and the internal SWG geometrical parameters (silicon post length a_z and pitch Λ_z), but also of the geometrical parameters of the device (lengths L_1 to L_4), that is,

$$n_{\text{swg}}(a_z, \Lambda_z, \lambda, L_1, L_2, L_3, L_4). \tag{1}$$

We determine separate effective refractive indices for forward transmission f, vertical transmission v, and reflection r, finding that multiple solutions can exist as phase was not included in the DNN training data. However, we identify the physical effective refractive index as the one where the solutions for all three directions overlap. In Section 4, our DD-EMA effective refractive index profiles are used in 2D device simulations, and are validated against full 3D device simulations. We find good agreement, indicating that reduced dimensional simulations are feasible with the DD-EMA approach. We also find that 2D device simulations based on the the conventional slab EMT model can be far less accurate, particularly for devices with geometrical parameters representing viable grating coupler devices. Finally, in Section 5, we provide conclusions.

2. DNN-based surrogate models for vertical grating couplers with SWGs

2.1. Data Acquisition

We consider the vertical grating coupler scheme depicted in Fig. 1 (bottom), a promising design paradigm for increased coupling efficiency (CE). It is a fully etched and encapsulated silicon-on-insulator (SOI) device, where the blue features in the diagram represent silicon, which are completely embedded in silica; the height (in the y direction) of all silicon features is 220 nm. Each unit cell contains four segments: (1) a silica segment of length L_1, (2) a SWG segment with length L_2, (3) a silica segment with length L_3, and (4) a silicon segment with length L_4. The sum of L_1 to L_4 gives the unit cell length, Λ. The SWG segment is composed of alternating silicon and silica segments along the z direction, with a silicon segment width of a_z, and a SWG pitch Λ_z. The top of Fig. 1 depicts the same grating coupler scheme, but where the SWG segment (L_2) is replaced by a homogeneous medium (green) with a uniform index of refraction n.

We simulated the two devices depicted in Fig. 1 for a large set of geometrical parameter values and (in the homogenous medium case) n values, chosen randomly within the ranges specified in Table 1. We enforced a minimum feature size of 40 nm, consistent with realistic fabrication constraints. The range for the unit cell length Λ was chosen according to the telecommunication band 1450 – 1650 nm. This placed an additional constraint on the choice of L_1 to L_4, which must sum to Λ; the effect was that L_1 to L_4 were not uniformly distributed, but rather had a distribution

Parameter	Device with SWG	Device with homogeneous medium
Λ [nm]	400 to 1000	400 to 1000
L_1, L_2, L_3, L_4 [nm]	40 to Λ, $\sum L_i = \Lambda$	40 to Λ, $\sum L_i = \Lambda$
Λ_z [nm]	100 to 650	N/A
a_z [nm]	40 to $\Lambda_z - 40$	N/A
n	N/A	1.4 to 3.5

Table 1. Range of values for the geometrical parameters indicated in Fig. 1 and for the refractive index n of the (green) homogeneous medium in the top device of Fig. 1.

skewed toward smaller values. The range for n was chosen such that it encompasses all values between the refractive index of silica and silicon within the wavelength range of interest.

We simulated three unit cells of each grating coupler instance, with an input and output silicon waveguide on either side of the three unit cells (as shown in Fig. 1). While in an actual grating coupler device there would be many more unit cells, the goal here is to develop a method to extract the environment-aware dispersive effective index that could be used in a full grating coupler simulation, not to do the full grating coupler simulation itself. For ease of computation, we also assumed an infinite extent in the transverse (z) direction. While realistic grating couplers have finite width, they are wide enough to consider this approximation to be negligible.

We used Ansys Lumerical to perform the simulations, which used a uniform mesh with conformal variant 0 mesh refinement, a mesh step size of 10 nm in all directions (determined to be sufficient through detailed convergence studies). Perfectly matched layer (PML) boundary conditions were applied in the x and y directions. For the device containing SWGs, 3D simulations were performed for a z span of Λ_z with periodic boundary conditions applied in z. For the device containing homogeneous medium regions, 2D FDTD simulations were performed. We injected TE-polarized light (E-field in the z-direction) in the x-direction from the input silicon waveguide using a mode source in the range of $1450 - 1650$ nm.

Transmission monitors were placed near the maximum and minimum x values to collect the forward transmitted (f) and reflected (r) intensity, as well as near the maximum and minimum y values to collect the transmitted intensity in the vertical (v) directions. Since a time-domain approach was used, only a single simulation was needed per structure to collect data for all desired wavelengths (100 points within $1450 - 1650$ nm).

Altogether, we simulated 23k 2D structures and 35k 3D structures, which provided 2.3 million and 3.5 million data points, respectively.

2.2. DNN Training

We train two families of DNNs to replace the 2D and 3D forward simulations via a supervised learning approach, where the DNNs take as input the parameters of Table 1 and the desired wavelength λ. The DNN output consists of three real-valued quantities: the forward (f) and vertical (v) transmissivity, and the reflectivity (r) at the input wavelength. Our training set for the 2D (homogenous medium) case consisted of simulation data from 20k unique structures at 100 wavelengths (2 million points), and for the 3D (SWG) case the training set consisted of 35k unique structures at 100 wavelengths (3.5 million points); the test data sets for both cases consisted of 3000 structures (300k points) distinct from the training sets.

In order to increase the accuracy of the DNN-based surrogate model predictions, we trained three distinct DNNs for 2D and for 3D that were used to implement ensemble averaging. The distinction between the three DNNs stems from them having variations in their architecture

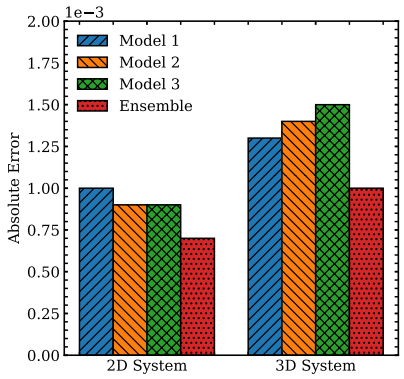


Fig. 2. Absolute error in ensemble model predictions (red) as well as the individual DNN predictions (blue, orange, green) for the 2D (left) and 3D (right) cases. The errors were determined by averaging the absolute error in all three outputs over the entire test data set, including all the wavelengths.

(number of hidden layers and/or number of neurons in each layer). Details of the architecture are presented in Table S1 of Section S2 in the Supplementary Section. The underlying architecture consists of a set of expanding layers from the input, which then contract into the output [38,39]. The gradual expansion of the first few layers in the DNN introduces a high-dimensional representation of the design space, allowing it to have flexibility in learning non-linear features [39]. Our architecture was not formally optimized, but instead was chosen based on trial and error.

2.3. DNN Performance

The ensemble model predictions for the 2D and 3D cases are obtained by averaging (with equal weights) the individual predictions from all three respective DNNs. The absolute errors of the ensemble model predictions are plotted in Fig. 2 for both 2D (left, red) and 3D (right, red) along with the absolute errors of the three individual 2D and 3D models (blue, orange and green). The absolute error was calculated by taking the absolute difference between the three output predictions (for the f, v, and r directions) and averaging over all test set data points (including all wavelengths). As the outputs take values between 0 and 1, this absolute error is comparable to (but not the same as) percent error. As expected, we see in both the 2D and 3D cases an improved performance for the ensemble model.

Fig. 3 further demonstrates ensemble model performance for 2D (left) and 3D (right) cases for the vertical transmission (v). The histograms in the top row show the distribution of the absolute error of the predicted forward (f) and vertical (v) transmission and reflection (r) over all wavelengths, with a logarithmic vertical axis for ease of viewing. The vertical lines indicate the average absolute error μ (red solid), as well as the 99th (orange dotted), 99.995th (green dashed), and 99.999th (blue dash-dotted) percentiles. One can see that most of the data is concentrated near the average, which represents a low error, and that even up to the 99.999th percentile, where there are very few counts, the absolute error remains far below one.

The bottom row contains scatter plots of the predicted vertical transmission versus the actual vertical transmission (from FDTD data) for all test data points over the wavelength range of interest. The colour of each scatter point marker has been set according to its corresponding wavelength, as indicated by the colour scale at the bottom, with very good agreement. Note that those plots highlight the outliers in particular, since the vast majority of data points (in total there are 300k data points) are plotted on top of each other along the diagonal line. Both the

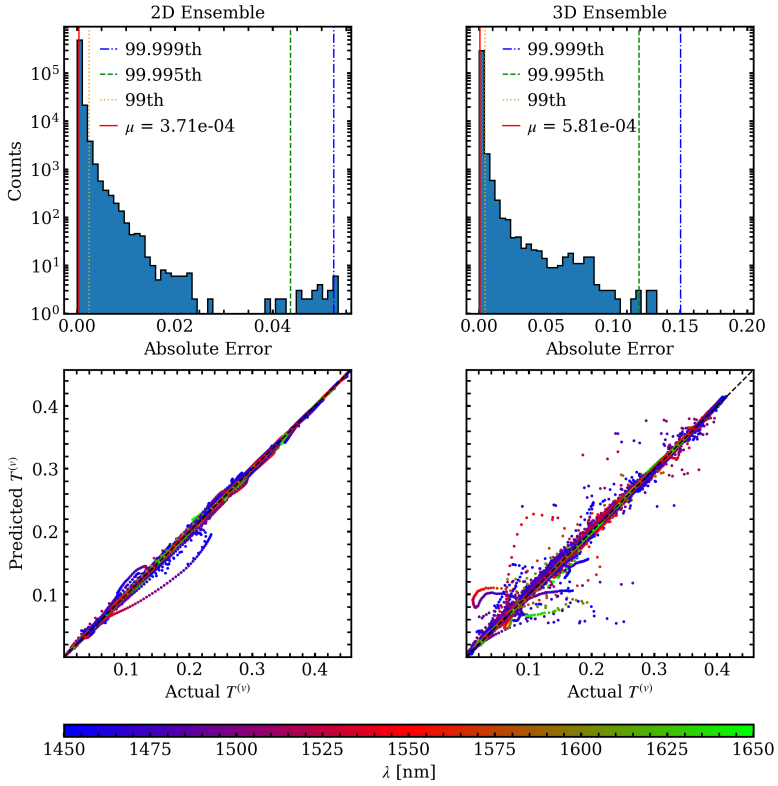


Fig. 3. Top row: Distribution of the absolute error of the ensemble model predictions for the 2D (left) and 3D (right) cases of vertical transmission. The vertical lines indicate the average absolute error, μ (red solid), as well as the 99th (orange dotted), 99.995th (green dashed), and 99.999th (blue dash-dotted) percentiles. Bottom row: Predicted vertical transmission versus the actual vertical transmission (from FDTD data) for all test data points over the wavelength range of interest. The colour of each data point indicates its wavelength, according to the colour scale shown.

histograms and predicted-actual scatter plots indicate that the 2D case outperforms the 3D, as we saw in Fig. 2.

The corresponding data for the predictions of forward transmission (t) and reflection (r) are presented in the Supplementary Section S2 in Figs. S1 and S2, respectively.

3. Effective Refractive Index Determination

In this section, we determine the (dispersive) effective refractive index n_{swg} of the SWG layer in the 3D grating coupler structure, where the effect of its environment and its own finite dimension is fully taken into account. The ultimate goal is to enable simulations where a uniform medium with index $n_{\text{swg}}(\lambda)$ replaces the 3D SWG segment. This would allow us to use 2D rather than 3D computational domains to simulate a larger grating coupler device, resulting in a significant compute cost reduction.

Structure	L_1 [nm]	L_2 [nm]	L_3 [nm]	L_4 [nm]	Λ_z [nm]	a_z [nm]	Duty Cycle
A	63	179	95	340	300	240	0.800
B	63	100	95	340	300	240	0.800
C	63	179	95	340	550	407	0.740

Table 2. Parameters for three 3D structures. A and B differ only in L_2, and A and C differ only in the SWG parameters Λ_z and a_z (and consequently Duty Cycle).

3.1. Procedure

The method to find the appropriate effective refractive index is outlined in Algorithm 1, which employs the 2D and 3D DNN-based surrogate ensemble models described in the previous section. We begin by choosing a 3D structure of interest - defined by a particular choice of geometrical parameters $L_1, L_2, L_3, L_4, \Lambda_z$ and a_z - for which we desire $n_{\text{swg}}(\lambda)$. We use the 3D ensemble model to predict the transmission values $T_{3D}^{(d)}$ for our chosen 3D structure, with $d \in \{f, v, r\}$ representing the forward, vertical, and backward directions, respectively. This is done for 100 wavelength values λ between 1450 and 1650 nm, and we refer to it as the 3D reference prediction.

We then use the 2D ensemble model to predict the transmission values $T_{2D}^{(d)}$ in each direction d for a 2D structure with L_1 to L_4 values identical to those of the 3D structure of interest, but where the SWG segment is replaced by a homogeneous medium with refractive index n; this is again done for each λ value. The goal is to find the value of n that minimizes the difference between the 3D reference prediction and the 2D prediction, for each λ.

We consider two approaches. In the first, we define the following error function

$$F(n, \lambda) = \sqrt{\sum_d \left(T_{3D}^{(d)}(\lambda) - T_{2D}^{(d)}(n, \lambda) \right)^2}. \tag{2}$$

As we seek to minimize this function with respect to only one parameter, n, the simplest approach is to create a set $\mathbb{M}$ of equally spaced n values, and run each through the 2D ensemble model to calculate its corresponding $T_{2D}^{(d)}(n, \lambda)$. For each λ, the optimal value of n is taken to be the one in $\mathbb{M}$ at which $F(n, \lambda)$ is smallest, that is, we take

$$n_{\text{swg}}(\lambda) = \underset{n \in \mathbb{M}}{\text{argmin}} \left(F(n, \lambda) \right). \tag{3}$$

The smaller the step between adjacent values of n, the more precise n_{swg} becomes. There are other optimization techniques one could use for this step (for example, grid search methods), but this was not done here as we desired to understand and visualize the features of the error function across the whole of the (n, λ) parameter space. As a single-parameter optimization problem, a simple parameter sweep is sufficient for this.

In the second approach, we relax the assumption that the effective index is necessarily the same for each direction. In this case, we define a separate error function for each direction $d \in \{f, v, r\}$, given by

$$F^{(d)}(n, \lambda) = \left| T_{3D}^{(d)}(\lambda) - T_{2D}^{(d)}(n, \lambda) \right|. \tag{4}$$

For each λ, we then seek three optimal values of n, one for each error function $F^{(d)}(n, \lambda)$. These are given by

$$n_{\text{swg}}^{(d)}(\lambda) = \underset{n \in \mathbb{M}}{\text{argmin}} \left(F^{(d)}(n, \lambda) \right), \tag{5}$$

We again use the set $\mathbb{M}$ of equally spaced n values.

3.2. Case studies and analysis

Now that we have defined our effective index extraction procedure, we turn to applying it to some concrete structures, as listed in Table 2. We present these structures to demonstrate how the optical behaviour changes with geometry, and how multiple effective index solution branches arise. A comparison of these structures probes the effect of increasing transverse period, while A and B isolate the optical sensitivity to the length. Structures A and C are grating coupler designs from earlier work [40] [41], in which the transverse SWG period was progressively increased to enlarge the minimum feature size while preserving the intended optical response. Structure A uses a 300 nm pitch with a duty cycle of 0.800, whereas Structure C uses a 550 nm pitch with a duty cycle of 0.740. These pitch/duty-cycle combinations were obtained from the earlier 3D optimization for approximately the same target homogeneous response. Structure B is a modified version of Structure A in which only the finite SWG length, L_2 (Fig. 1) is changed, while the transverse SWG geometry and the other longitudinal dimensions are kept fixed. Its role is to provide a controlled comparison for examining sensitivity to the finite longitudinal extent of the SWG segment.

We begin with the calculation of the error function F defined in Eq. 2 across the full wavelength range of interest (that is, equally distributed between 1450 and 1650 nm, as described in the previous subsection) and for n values in $\mathbb{M}$. Contour plots of $F(n, \lambda)$ for structures A (left), B (centre), and C (right) are shown in Fig. 4. We find that the contour plots for structures A and C are in fact very similar, suggesting that for two structures optimized for the same operating range, the effect of changing the pitch Λ_z is small. However, the contour plot is quite different for structure B, indicating that performance is much more sensitive to the SWG length L_2.

Though physically we would expect only one solution of n_{swg} for each wavelength, the contour plots in Fig. 4 also suggest the presence of multiple solutions. Two different n_{swg} values can reproduce one intensity channel locally, but only one lies within the overlap of the forward (f), vertical (v), and reflection (r) solutions and gives consistent agreement across the other channels. To explore this further, we next calculated the error functions for each transmission direction separately, that is, the $F^{(d)}(n, \lambda)$ for $d \in \{f, v, r\}$ as defined in Eq. 4. The $n \in \mathbb{M}$ that minimized $F^{(d)}$ for each direction and wavelength were extracted, and the resulting candidate $n_{\text{swg}}^{(d)}$ values are shown in the top row in Fig. 5 for structures A (left), B (middle) and C (right), for $d = f$ (forward transmission, blue circles), $d = v$ (vertical transmission, orange crosses), and $d = r$ (reflection, green three-pointed stars). Separate contour plots of the three $F^{(d)}(n, \lambda)$ for each of the structures A, B, and C are shown in Fig. S3 of the Supplementary Section S3.

Again, we find multiple solutions, but this time in each of the $n_{\text{swg}}^{(d)}$. Note that solutions appear to hop from one branch to another for different wavelengths. Since n is chosen according to a

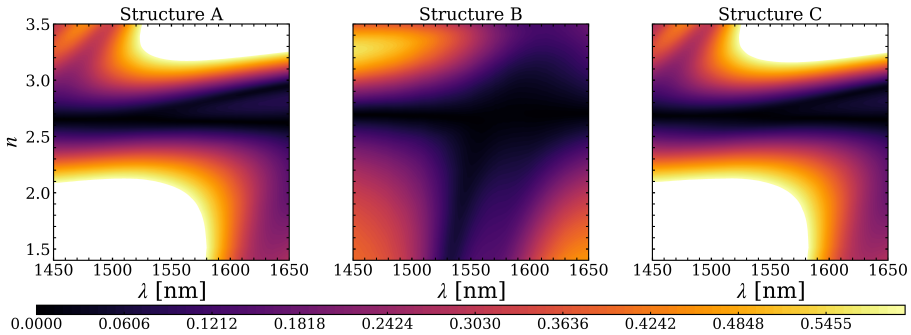


Fig. 4. Contour plot of $F(n, \lambda)$ for structures A (left), B (centre), and C (right).

fixed grid, and only one value of n is chosen for each wavelength, this seemingly random hopping between solutions is subject to numerical noise. To further explore the existence of multiple solutions, we examine more closely one of the curves in Fig. 5, in particular, the effective index solutions for $d = v$ for structure A. This is replotted in Fig. 6 (Left) for convenience. We choose a wavelength where there are two solution branches for further analysis, specifically $\lambda = 1509.1$ nm (vertical dashed line) where two possible effective index solutions (red stars) are $n = 2.663$ and $n = 2.843$.

The right plot of Fig. 6 shows the three error functions $F^{(d)}$ as a function of n for $\lambda = 1509.1$ nm, for all three directions $d = f, v, r$ (orange dashed, blue dotted, and bold green lines, respectively). The two red stars on the left correspond to the two dips in the forward curve on the right (orange dashed line). The curves for the error function in the two other directions both show a dip near $n = 2.663$, but not near $n = 2.843$.

The coincidence of the three dips at only one n value suggests that perhaps this n value is the true n_{swg} solution. To test this, we supplied both candidate n_{swg} values to the 2D ensemble model to obtain transmission predictions in all three directions. We then compared these spectra to the 3D ensemble model predictions for the actual structure; these are shown in Fig. 7. The 2D ensemble model predictions of transmission spectra in the f (left), v (centre) and r (right) directions are shown for $n_{\text{swg}} = 2.663$ (orange dashed line) and $n_{\text{swg}} = 2.843$ (green dash-dotted line), and these are compared to the 3D ensemble model predictions (blue solid line). For the forward direction, both values of n_{swg} lead to agreement with the 3D prediction at $\lambda = 1509.1$ nm. However, the $n_{\text{swg}} = 2.663$ (orange dashed line) shows much better overall agreement.

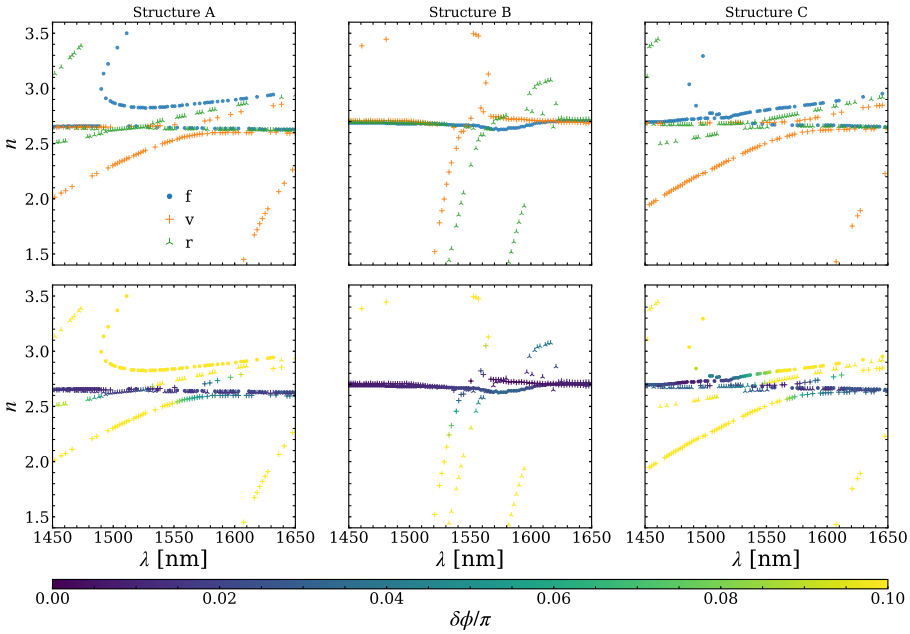


Fig. 5. Top row: Candidate solutions for $n_{\text{swg}}^{(d)}$ versus wavelength for structures A (left), B (middle), and C (right) of Table 2 for $d = f$ (forward transmission, blue circles), $d = v$ (vertical transmission, orange crosses), and $d = r$ (reflection, green three-pointed stars). Bottom row: Same data as top row, but where the marker colour indicates $\delta\phi^{(d)}$, which has been normalized by π.

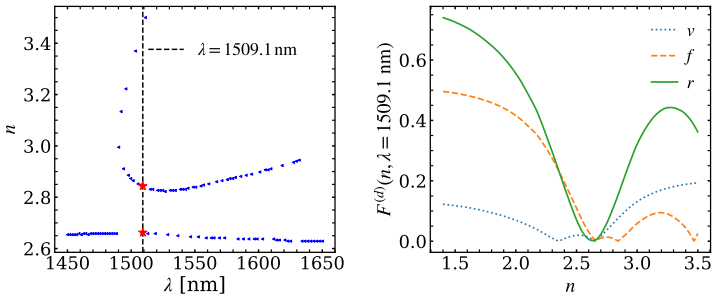


Fig. 6. Left: Candidate solutions for $n_{\text{swg}}^{(f)}$ for Structure A; the vertical line is at λ = 1509.1 nm, and two candidate $n_{\text{swg}}^{(f)}$ solutions are indicated by red stars. Right: $F^{(d)}$ as a function of n for λ = 1509.1 nm, for all three directions $d \in \{f, v, r\}$ (orange dashed, blue dotted, and bold green lines, respectively).

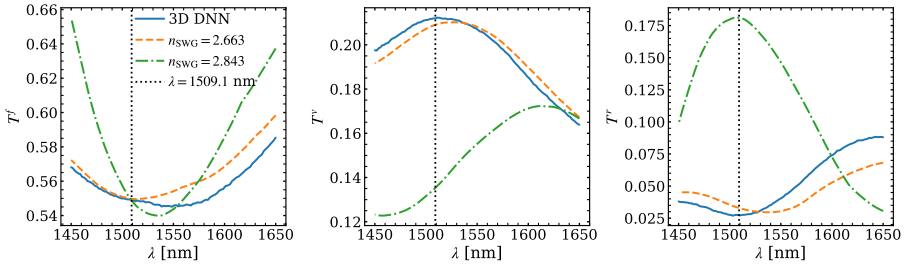


Fig. 7. Transmission curves in the forward (f, left), vertical (v, centre), and reflection (r, right) directions for Structure A at λ = 1509.1 nm as predicted by the 2D ensemble model with candidate n_{swg} values of 2.663 (orange dashed line), 2.843 (green dash-dotted line) as well as those predicted by the 3D ensemble model (blue solid line).

Interestingly, for the other two directions, only $n_{\text{swg}} = 2.663$ shows a reasonable overall agreement with the 3D prediction. This suggests that in cases where there are multiple solutions, one needs to identify the candidate $n_{\text{swg}}^{(d)}$ values that are nearly overlapping for the $d = f, v, r$ directions to obtain the true effective index solution.

As we illustrate below, the physical reason for the appearance of multiple solution branches is that our training data contained only transmission spectra and did not include the phase of the transmission coefficients. To show this, we compare the transmission phases calculated with 2D FDTD simulations with the predicted effective index values (including solutions of all the branches), and the transmission phases of the 3D FDTD simulation with the full structure. This phase comparison provides an independent physical check that the multi-directional overlap is consistent with a single physical solution. For each found n_{swg} value, a separate 2D FDTD simulation was run, and the complex transmission coefficients at the corresponding wavelength were collected at the output line. A single broadband 3D FDTD simulation was run for the full structure, and the complex transmission coefficients at the output plane was collected as functions of wavelength. The overall phase difference for each $d = f, v, r$ was then calculated via

$$\delta\phi^{(d)} = |\langle\phi^{(d)}_{\text{2D FDTD}}\rangle - \langle\phi^{(d)}_{\text{3D FDTD}}\rangle| \tag{6}$$

where the $\langle\,\rangle$ indicate averaging over the output line (in 2D) and output plane (in 3D) corresponding

to direction $d = f, v, r$.

The phase differences $\delta\phi^{(d)}$ between the 3D and 2D FDTD transmission phases are indicated in the plots in the bottom row of Fig. 5 for structures A, B, and C. These are the same scatter plots as in the top row, except that for each candidate $n_{\text{swg}}^{(d)}$ solution, the corresponding value of $\delta\phi^{(d)}$ is indicated by the colour of its marker. A small $\delta\phi^{(d)}$ indicates a high degree of agreement between 2D and 3D FDTD solutions, and thus those values of candidate $n_{\text{swg}}^{(d)}$ would be the true solutions. A large $\delta\phi^{(d)}$ indicates that the candidate $n_{\text{swg}}^{(d)}$ used for the 2D FDTD simulation is not the correct value for that wavelength.

The determination of $\delta\phi^{(d)}$ is not very convenient for identifying which of the multiple branches (such as those in Fig. 5) is the correct one. However, the figure suggests a much more feasible alternative. It shows, for all three structures, that candidate $n_{\text{swg}}^{(d)}$ solutions with small $\delta\phi^{(d)}$ solutions correspond to those solutions where there is significant overlap between $n_{\text{swg}}^{(d)}$ solutions for all three directions $d = f, v, r$. This, indeed, further supports our analysis for Figs. 6 and 7, which suggests that the $n_{\text{swg}}^{(d)}$ solutions that overlap are the correct ones.

In the Supplementary Section S4, we consider three additional structures, whose geometrical parameters are presented in Table S2, and do not represent optimal solutions. The candidate solutions for $n_{\text{swg}}^{(d)}$ also show multiple branches. Fig. S4 shows that for those structures, just like structures A, B, and C here, solutions with smaller $\delta\phi^{(d)}$ correspond to the regions of overlap of the $n_{\text{swg}}^{(d)}$ for $d = f, v, r$.

Overlap among the directional candidate solutions provides a practical phase-free criterion for selecting the effective index branch across wavelength. Although this is possible based on the phase information as well, in our work, we find this unnecessary, as the solution obtained by minimizing all transmission errors consistently coincides with using the phase information.

3.3. Determining the dispersive effective index function

That the true solutions (with small $\delta\phi^{(d)}$) correspond to overlapping $n_{\text{swg}}^{(d)}$ solutions for $d = f, v, r$ suggests a method for finding the true solution in situations where there are multiple solution branches. We need to develop a procedure for quantifying “overlap" that is robust and agrees with solutions identified by small $\delta\phi^{(d)}$. As well, we seek to fit our results to an analytical form for $n_{\text{swg}}(\lambda)$ that can be used in broadband simulations.

We start by determining a method for finding a functional form $n_{\text{swg}}(\lambda)$ that has the smallest $\delta\phi^{(d)}$ values. Based on the appearance of the $n_{\text{swg}}^{(d)}$ solutions in Fig. 5, we choose our fitting function to be a second-order polynomial. For each $n_{\text{swg}}^{(d)}$ data set, we choose only those values with $\delta\phi^{(d)}/\pi < \delta\phi$ for some $\delta\phi$. Then only these data points are used in the polynomial fit. This ensures that only low-error points are included. As it is a bit ambiguous how to choose the precise value of $\delta\phi$, we perform the fit for several $\delta\phi$ values. The corresponding fits are plotted in the top row of Fig. 8 for structures A (left), B (centre), and C (right). Also plotted is a second-order polynomial fit for the n_{swg} obtained by minimizing $F(n, \lambda)$, for comparison (black dots). One can see that the resulting polynomial fits, for all three structures, give results that are close (especially when considering the small range of n on the y-axes).

We repeat this procedure, but now selecting points based on how much they “overlap", that is, how close the $n_{\text{swg}}^{(d)}$ values are to each other for $d = f, v, r$. The algorithm for this is outlined in pseudocode in the Supplementary Section S5. Again, we define a δn that sets the maximum “distance" between $n_{\text{swg}}^{(d)}$ values, and select only those points that are within δn of each other. We then perform second-order polynomial fits for this subset of points. Again, we choose several values of δn and repeat this procedure. The results are plotted in the bottom row of Fig. 8. We see that this again produces results that are somewhat insensitive to the choice of δn, and whose curves are very similar to those that minimize $\delta\phi^{(d)}$ as shown in the top row.

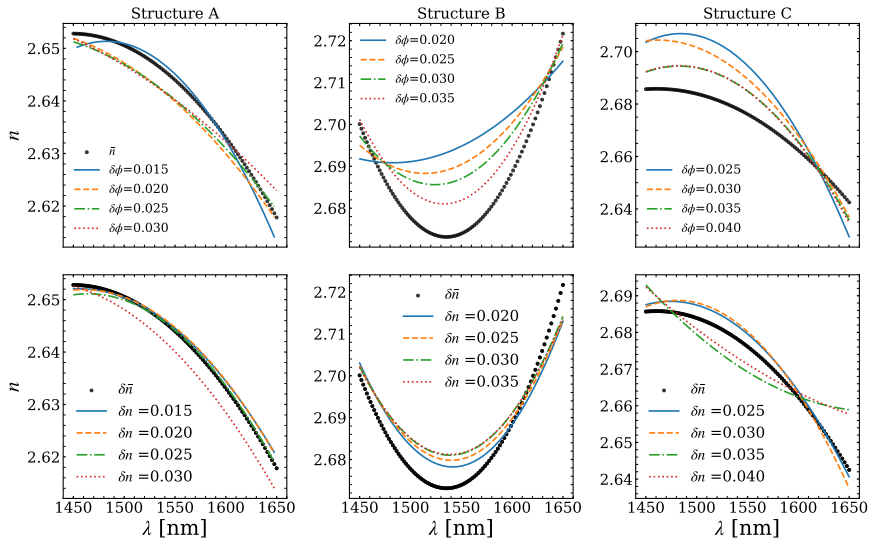


Fig. 8. Second-order polynomial fit of a structure based on phase difference (top row) and effective index separation (bottom row) for the three structures A (left row), B (middle row), and C (right row) in Table 2. $\delta\bar{n}$ is the fit from Eq. 2, whereas the others are the associated thresholds.

Fig. S6 in the Supplementary Section S4 shows the corresponding graphs for the second set of structures defined in Table S2, where similar agreement between the phase and overlap approaches is found, as is a similar insensitivity to the choice of the threshold value. With this, we arrive at a method for obtaining the n_{swg} that will be used for the remainder of this paper. It is defined as the average of all the polynomial fits computed for different values of δn, i.e., those plotted in the bottom row of Fig. 8 (and the bottom row of Fig. S6 in the Supplementary Section S4).

Overall, we see that the selected points form a smooth dispersive index as physically expected for an effective medium representation over this wavelength range. Comparing the results of the overlap and phase criteria, we notice a similar dispersive branch was obtained.

3.4. Comparison with Slab Model

The conventional slab model provides the natural reference for interpreting the extracted effective index. In Fig. 9, structures A (left), B (centre), and C (right) illustrate the effect of increasing transverse SWG period. Here, we fix the longitudinal geometry, while choosing the two pitch/duty-cycle combinations so that they have approximately the same target effective response. The effective index is obtained by averaging over all the δn curves in the bottom row of Fig. 8 (orange dashed lines). Also plotted for each structure is the prediction made by the slab model (bold blue lines). Structure A uses a 300 nm pitch with DC = 0.800, while C uses a 550 nm pitch with DC = 0.740; the much larger departure from the slab prediction for C is consistent with increasing deviation from the deeply subwavelength regime.

Structures A and B isolate a different effect, the finite longitudinal extent of the SWG segment: the transverse SWG geometry is identical, so the conventional slab model assigns the same local effective index, while only L_2 changes. The resulting change in the extracted index shows that, in

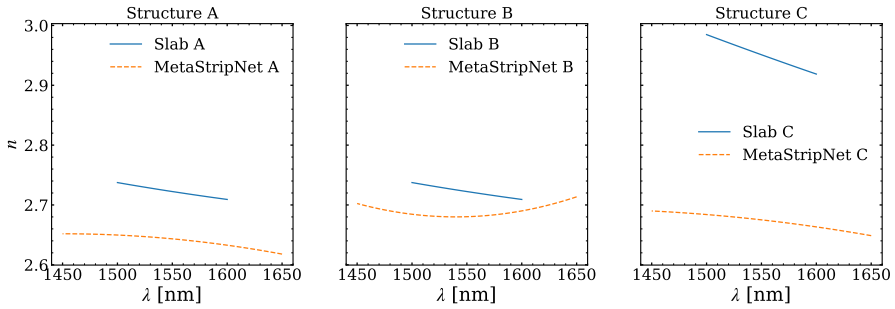


Fig. 9. Effective refractive index of slab model (blue bold line) compared with our prediction (orange dashed line) for structures A (left), B (centre), and C (right) of Table 2.

a finite structure, the effective representation can depend on the SWG segment length and is not determined by pitch and duty cycle alone. Together, these comparisons illustrate how departures from the local slab description can emerge through both transverse and finite-size effects, without implying a sharp boundary of validity. For the additional three structures in Table S2, Fig. S7 of Supplementary Section S4, we see behaviour further corroborating these observations.

Here, we have demonstrated a key result: the effective optical response of finite SWG structures cannot generally be described solely by local pitch and fill factor. Instead, the optical response depends strongly on the surrounding optical environment, which we observed by a systematic variation of L_2, which impacts the interaction with the surrounding environment. Variations in SWG pitch produce substantial changes in the dispersive effective index profiles, despite similarities between the local periodic structures. These observations demonstrate the limits of conventional slab models which assume plane wave propagation and weak a dependence on the surrounding environment. On the contrary, finite structures have sensitivity to the surrounding optical environment.

The comparisons between the conventional slab model and full 3D simulations above, show that PEMA more accurately reproduces multi-directional optical behaviour in finite structures. The extracted effective index should therefore be interpreted as an effective representation which is valid for certain structures.

4. Validation

In this section, we validate the extracted effective representation directly with FDTD. In order to demonstrate this, we show that the dispersive effective index, $n_{\mathrm{swg}}(\lambda)$, obtained from the matching procedure discussed earlier, allows the 2D structure to reproduce the optical behaviour of the full 3D SWG structure, and how this compares with the conventional slab representation.

We check this by inserting the fitted $n_{\mathrm{swg}}(\lambda)$ into the effective 2D structure and evaluating its optical behaviour using FDTD. We then compare the forward transmission, vertical transmission and reflection with the FDTD results for the 3D equivalent structure. We also compare these with the slab model, which provides a comparison baseline. In order to ensure that our results were robust, we also compared with the phase and overlap error predictions. The results are plotted in Fig. 10. The 3D Lumerical FDTD transmission spectra (bold green curve) were calculated using the full 3D structure, and thus serve as validation spectra. As described above, our $n_{\mathrm{swg}}(\lambda)$ is determined by averaging all the curves fit to data with different δn upper bounds (orange dashed line). We also include here another $n_{\mathrm{swg}}(\lambda)$ that is determined by averaging all the curves fit to

data with different $\delta\phi$ upper bounds (blue dashed line). Also shown are 2D FDTD results where the effective index of the SWG segment is the fit obtained with the $n_{\text{swg}}(\lambda)$ data obtained by minimizing the loss function $F(n, \lambda)$, corresponding to the black dots of Fig. 8 (black dotted line). Finally, we also present the 2D FDTD calculations with an effective SWG index as predicted by the slab model (red dotted line).

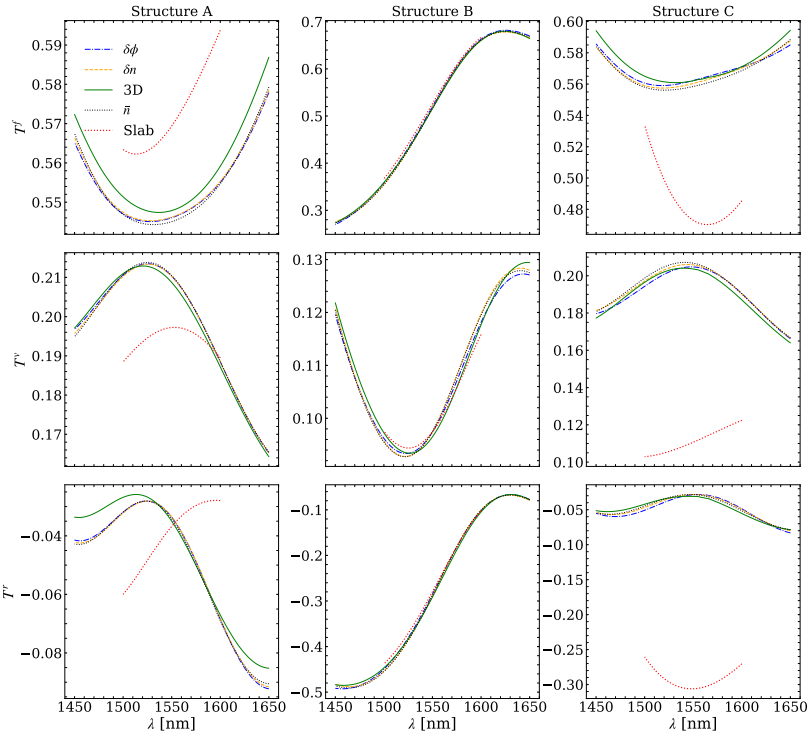


Fig. 10. Transmission plots for structures A, B, and C of Table 2 calculated using 2D and 3D FDTD simulations. $\delta\phi$ is the curve based on the mean of the phase, δn is the mean of the effective index separation, 3D is the 3D Lumerical result, $\bar{n}$ is from Eq. 2, and Slab is from the prediction from the slab model.

There is significant disagreement between the slab model and the 3D validation curve for structures A and C. This is consistent with the significant difference between the effective index predictions of the slab model and our prediction shown in Fig. 9. The transmission spectra calculated with each of our three different versions ($\delta\phi$, δn and $\bar{n}$) are significantly closer to the validation curves than the slab model spectra for these two structures. Structure C is particularly relevant because it represents the SWG with the larger pitch. Its good agreement with the full 3D calculation shows that the extracted effective medium representation remains useful even in a regime where the conventional slab description shows a larger discrepancy. This provides the natural link to using the approach for more fabricable grating coupler designs. In the case of structure B, we notice that the slab model and the 3D model agree well, as do all three of our approaches. Fig. S8 and Fig. S7 of Supplementary Section S4 plot these curves for the three structures defined in Table S2. Those plots further reinforce our conclusions.

5. Summary & Outlook

SWG metamaterials have become a mature design tool in integrated photonics, in large part because an effective-index description makes complex nanostructures accessible to conventional photonic design methods. At the same time, practical device development increasingly pushes toward larger feature sizes, broader accessible index ranges, and more strongly structured or radiative geometries. These trends make the limits of purely local homogenization progressively more relevant. As SWG structures move from idealized subwavelength media toward finite, fabricable device structures, there is increasing value in effective models that retain the efficiency and physical intuition of effective medium descriptions without requiring full 3D simulation throughout the design process.

The present work can be viewed as a further step in a hierarchy of effective descriptions. Conventional EMTs provide the starting point, while the slab model already incorporates the device thickness and vertical confinement. PEMA extends this idea further by allowing the effective representation to depend on the finite longitudinal geometry and environment of the SWG segment. Rather than abandoning the effective index picture when finite size effects appear, it constructs an effective representation from the optical response of the finite 3D structure. This preserves a physically meaningful design variable while incorporating information that is not captured by the local slab model. In this sense, the contribution of PEMA is broader than computational acceleration alone. It provides a way to retain the effective medium concepts in regimes where the relationship between geometry and effective response is no longer adequately described by the slab model. This is particularly relevant for metamaterial devices in which fabrication constraints encourage larger structural periods and where small local modeling errors can propagate into device performance.

A next step is to understand more systematically where increasingly local descriptions cease to be sufficient and how much environmental information an effective representation needs to retain. This includes studying other longitudinal environments, different metamaterial topologies, wider ranges of transverse period, and other classes of radiative or resonant photonic components. The companion grating coupler work provides one device design of this direction; more broadly, similar effective approaches may offer a practical bridge between full-wave modeling and scalable design of complex metamaterial photonic systems.

References

1. H. Aghaee Rad, T. Ainsworth, R. N. Alexander, B. Altieri, M. F. Askarani, R. Baby, L. Banchi, B. Q. Baragiola, J. E. Bourassa, R. S. Chadwick, I. Charania, H. Chen, M. J. Collins, P. Contu, N. D'Arcy, G. Dauphinais, R. De Prins, D. Deschenes, I. Di Luch *et al.*, "Scaling and networking a modular photonic quantum computer," Nature **638**, 912–919 (2025).
2. S. Yu, J. Sun, K.-C. Chen, Z.-H. Yang, Z. Li, E. Mer, Y. K. Alwehaibi, S. H. Winston, D. M. Lopena, Z.-C. Zhang, G. Yang, R. Tao, M. Zhou, G. J. Machado, Y. Dong, R. Bondesan, V. Vedral, M. S. Kim, I. A. Walmsley, and R. B. Patel, "Extensible universal photonic quantum computing with nonlinearity," Nat. Photonics **20**, 566–575 (2026).
3. H. Zhu, T. Chen, H. Ma, Z. Zhao, J. Guo, Q. Liang, Y. Tu, T. m. He, W. Luo, B. Wang, Q. Song, C. Jin, F. Liu, Y. Yang, Y. Wang, K. C. Leong, D. Wang, and J. Yang, "Recent progress towards large-scale integrated photonic quantum computation," npj Nanophotonics **3**, 1 (2026).
4. F. F. Settembrini, A. C. Gungor, A. Forrer, S. A. Fortune, A. Dell'Aquila, P. Padmanabhan, I. E. Opris, M. Sotto, N. Dordevic, Y. Perelman, and R. Nicolaescu, "A large-scale coherent 4d imaging sensor," Nature **651**, 364–370 (2026).
5. M. A. Butt, B. Imran Akca, and X. Mateos, "Integrated photonic biosensors: Enabling next-generation lab-on-a-chip platforms," Nanomaterials **15** (2025).
6. H. Zhang, Y. Song, S. Chen, Y. Bai, X. Xu, C. Huang, J. Wang, H. Chen, D. J. Moss, and K. Xu, "Integrated platforms and techniques for photonic neural networks," npj Nanophotonics **2**, 40 (2025).
7. P. Cheben, D. Xu, S. Janz, and A. Densmore, "Subwavelength waveguide grating for mode conversion and light coupling in integrated optics," Opt. express **14**, 4695–702 (2006).
8. J. M. Luque-González, A. Sánchez-Postigo, A. Hadij-ElHouati, A. Ortega-Moñux, J. G. Wangüemert-Pérez, J. H. Schmid, P. Cheben, Íñigo Molina-Fernández, and R. Halir, "A review of silicon subwavelength gratings: building break-through devices with anisotropic metamaterials," Nanophotonics **10**, 2765–2797 (2021).

9. J. M. Luque González, A. Sánchez-Postigo, A. Hadij-ElHouati, A. Ortega-Moñux, J. Wangüemert-Pérez, J. Schmid, P. Cheben, I. Molina-Fernández, and R. Halir, "A review of silicon subwavelength gratings: Building break-through devices with anisotropic metamaterials," Nanophotonics **10** (2021).
10. L. Sun, Y. Zhang, Y. He, H. Wang, and Y. Su, "Subwavelength structured silicon waveguides and photonic devices," Nanophotonics **9** (2020).
11. R. Halir, A. Ortega-Monux, D. Benedikovič, G. Mashanovich, J. Wanguemert-Perez, J. Schmid, I. Molina-Fernández, and P. Cheben, "Subwavelength-grating metamaterial structures for silicon photonic devices," Proc. IEEE **106**, 1–14 (2018).
12. P. Cheben, R. Halir, J. H. Schmid, H. A. Atwater, and D. R. Smith, "Subwavelength integrated photonics," Nature **560**, 565 – 572 (2018).
13. N. Kazanskiy, M. Butt, and S. Khonina, "Silicon photonic devices realized on refractive index engineered subwavelength grating waveguides-a review," Opt. Laser Technol. **138**, 106863 (2021).
14. R. Halir, A. Maese-Novo, A. Ortega-Moñux, I. Molina-Fernández, J. Wangüemert-Pérez, P. Cheben, D. Xu, J. Schmid, and S. Janz, "Colorless directional coupler with dispersion engineered sub-wavelength structure," Opt. Express **20**, 13470–7 (2012).
15. Y. Wang, Z. Lu, M. Ma, H. Yun, F. Zhang, N. A. F. Jaeger, and L. Chrostowski, "Compact broadband directional couplers using subwavelength gratings," IEEE Photonics J. **8**, 1–8 (2016).
16. Y. Xiong, J. G. Wangüemert-Pérez, D.-X. Xu, J. H. Schmid, P. Cheben, and W. N. Ye, "Polarization splitter and rotator with subwavelength grating for enhanced fabrication tolerance," Opt. Lett. **39**, 6931–6934 (2014).
17. Y. He, Y. Zhang, X. Wang, B. Liu, X. Jiang, C. Qiu, Y. Su, and R. Soref, "Silicon polarization splitter and rotator using a subwavelength grating based directional coupler," in *Optical Fiber Communication Conference,* (Optica Publishing Group, 2017), p. Th1G.6.
18. J. Wang, R. Ashrafi, R. Adams, I. Glesk, I. Gasulla, J. Capmany, and L. Chen, "Subwavelength grating enabled on-chip ultra-compact optical true time delay line," Sci. Reports **6**, 30235 (2016).
19. Y. Yang, Y. Jeon, Z. Dong, J. K. W. Yang, M. Haddadi Moghaddam, D.-S. Kim, D. K. Oh, J. Lee, M. Hentschel, H. Giessen, D. Kang, G. Kim, T. Tanaka, Y. Zhao, J. Burger, S. A. Maier, H. Ren, W. Jung, M. Choi, G. Bae, H. Chen, S. Jeon, J. Kim, E. Lee, H. Kang, Y. Park, D. Du Nguyen, I. Kim, P. Cencillo-Abad, D. Chanda, X. Jing, N. Liu, I. V. Martynenko, T. Liedl, Y. Kwak, J.-M. Nam, S.-M. Park, T. W. Odom, H.-E. Lee, R. M. Kim, K. T. Nam, H. Kwon, H.-H. Jeong, P. Fischer, J. Yoon, S.-H. Kim, S. Shim, D. Lee, L. A. Perez, X. Qi, A. Mihi, H. Keum, M. Shim, S. Kim, H. Jang, Y. S. Jung, C. Rossner, T. A. Konig, A. Fery, Z. Li, K. Aydin, C. A. Mirkin, J. Seong, N. Jeon, Z. Xu, T. Gu, J. Hu, H. Kwon, H. Jung, H. Alijani, I. Aharonovich, J. Kim, and J. Rho, "Nanofabrication for nanophotonics," ACS Nano **19**, 12491–12605 (2025).
20. L. Zhong, Y. Wang, X. Han, J. Hu, and J. Qiu, "Laser direct lithography of large-area three-dimensional integrated photonics: Technological challenges and advances," Chip **4**, 100157 (2025).
21. Y. Wan, W. He, J. Jaussi, L. Liao, D. Z. Pan, J. E. Bowers, and H. Rong, "Integrating silicon photonics with complementary metal–oxide–semiconductor technologies," Nat. Rev. Electr. Eng. **3**, 15–31 (2025).
22. A. Lakhtakia, *Selected Papers on Linear Optical Composite Materials* (1996).
23. D. Bruggeman, "Berechnung verschiedener physikalischer konstanten von heterogenen substanzen. i. dielektrizitatskonstanten und leitfahigkeiten der mischkorper aus isotropen substanzen," Annalen der Physik **416**, 636 – 664 (2006).
24. S. R. Jeeban Kumar Nayak, Partha Roy Chaudhuri and M. R. Sahoo, "A comprehensive review on effective medium theories to find effective dielectric constant of composites," J. Electromagn. Waves Appl. **37**, 282–322 (2023).
25. S. M. Rytov, "Electromagnetic properties of a finely stratified medium," (2014).
26. R. Selle, G. Vogt, T. Brixner, G. Gerber, R. Metzler, and W. Kinzel, "Modeling of light-matter interactions with neural networks," Phys. Rev. A **76**, 023810 (2007).
27. R. Selle, T. Brixner, T. Bayer, M. Wollenhaupt, and T. Baumert, "Modelling of ultrafast coherent strong-field dynamics in potassium with neural networks," J. Phys. B: At. Mol. Opt. Phys. **41**, 074019 (2008).
28. I. Malkiel, M. Mrejen, A. Nagler, U. Arieli, L. Wolf, and H. Suchowski, "Plasmonic nanostructure design and characterization via deep learning," Light. Sci. Appl. **7** (2018).
29. S. An, C. Fowler, B. Zheng, M. Y. Shalaginov, H. Tang, H. Li, L. Zhou, J. Ding, A. M. Agarwal, C. Rivero-Baleine, K. A. Richardson, T. Gu, J. Hu, and H. Zhang, "A deep learning approach for objective-driven all-dielectric metasurface design," ACS Photonics **6**, 3196–3207 (2019).
30. P. R. Wiecha, A. Arbouet, C. Girard, and O. L. Muskens, "Deep learning in nano-photonics: inverse design and beyond," Photon. Res. **9**, B182–B200 (2021).
31. R. Hegde, "Deep learning: A new tool for photonic nanostructure design," Nanoscale Adv. **2** (2020).
32. W. Ma, Z. Liu, Z. Kudyshev, A. Boltasseva, W. Cai, and Y. Liu, "Deep learning for the design of photonic structures," Nat. Photonics **15**, 1–14 (2020).
33. S. So, T. Badloe, J. Noh, J. Bravo-Abad, and J. Rho, "Deep learning enabled inverse design in nanophotonics," Nanophotonics **9**, 474 (2020).
34. T. Park, S. Mondal, and W. Cai, "Interfacing nanophotonics with deep neural networks: Ai for photonic design and photonic implementation of ai," Laser & Photonics Rev. **19**, 2401520 (2025).
35. H. Zhang, L. Kang, S. D. Campbell, J. T. Young, and D. H. Werner, "Data driven approaches in nanophotonics: a review of ai-enabled metadevices," Nanoscale **17**, 23788–23803 (2025).

36. J. Kim, J.-Y. Kim, J. Kim, Y. Hyeong, B. Neseli, J.-B. You, J. Shim, J. Shin, H.-H. Park, and H. Kurt, "Inverse design of nanophotonic devices enabled by optimization algorithms and deep learning: recent achievements and future prospects," Nanophotonics **14**, 121–151 (2025).
37. X. Zhou, Y. Xue, F. Ye, Z. Feng, J. He, X. Sun, K. M. Lau, and H. K. Tsang, "Fully etched low-back-reflection and high-efficiency silicon waveguide grating couplers with minimum feature size above 260 nm," J. lightwave technology **43**, 3854–3859 (2025).
38. W. Zhang, T. Du, and J. Wang, "Deep learning over multi-field categorical data," in *Advances in Information Retrieval,* N. Ferro, F. Crestani, M.-F. Moens, J. Mothe, F. Silvestri, G. M. Di Nunzio, C. Hauff, and G. Silvello, eds. (Springer International Publishing, Cham, 2016), pp. 45–57.
39. Q. Li and J. Chen, "An unsupervised network architecture search method for solving partial differential equations," J. Comput. Phys. **560**, 114954 (2026).
40. M. K. Dezfouli, Y. Grinberg, D. Melati, P. Cheben, J. H. Schmid, A. Sánchez-Postigo, A. O.-M. nux, G. Wangüemert-Pérez, R. Cheriton, S. Janz, and D.-X. Xu, "Perfectly vertical surface grating couplers using subwavelength engineering for increased feature sizes," Opt. Lett. **45**, 3701–3704 (2020).
41. J. Zhang, D. Melati, Y. Grinberg, M. Vachon, S. Wang, M. Al-Digeil, S. Janz, J. H. Schmid, P. Cheben, and D.-X. Xu, "Perfectly vertical silicon metamaterial grating couplers with large segmentation periods up to 650 nm," arXiv e-prints arXiv:2311.12057 (2023).

Supplementary: Data-Driven Effective Medium Approximation for Metamaterials in Device Designs

Norman S Israel[1,†], D. Blankespoor[1], Qiang Wang[1], D.-X. Xu[2], Y. Grinberg [3] and Lora Ramunno[1,*]

[1] *Department of Physics and Nexus for Quantum Technologies Institute, University of Ottawa, Canada*
[2] *Quantum and Nanotechnology Research Centre, National Research Council of Canada, Canada*
[3] *Digital Technologies Research Center, National Research Council of Canada, Canada*
[†] *nisra058@uottawa.ca*
[*] *lora.ramunno@uottawa.ca*

S1. Algorithm for finding the effective refractive index

Algorithm 1 Effective Index Determination via Error Function Minimization

Require: 3D DNN ensemble model, f_{3D}
Require: 2D DNN ensemble model, f_{2D}
Require: Input geometrical parameters, $\mathbb{P}_1 = L_1, L_2, L_3, L_4,\ \mathbb{P}_2 = \Lambda_z, a_z$
Require: Input wavelength set, $\mathbb{W} = \{1450, ..., 1650\}$ nm
Require: Input refractive index set, $\mathbb{M} = \{1.4, ..., 3.5\}$
Ensure: Effective refractive index

for each $\lambda \in \mathbb{W}$ **do**

$T_{3D}^{(d)}(\lambda)\ \forall\ d \in \{f, v, r\} \leftarrow f_{3D}(\mathbb{P}_1, \mathbb{P}_2, \lambda)$ ▷ 3D reference predictions

for each $n \in \mathbb{M}$ **do**

$T_{2D}^{(d)}(n, \lambda)\ \forall\ d \in \{f, v, r\} \leftarrow f_{2D}(\mathbb{P}_1, n, \lambda)$ ▷ 2D predictions for a given n

$F(n, \lambda) \leftarrow \sqrt{\sum_d \left(T_{3D}^{(d)}(\lambda) - T_{2D}^{(d)}(n, \lambda)\right)^2}$ ▷ Error functions calculation

$F^{(d)}(n, \lambda) \leftarrow \left|T_{3D}^{(d)}(\lambda) - T_{2D}^{(d)}(n, \lambda)\right|\ \ \forall\ d \in \{f, v, r\}$

end for

$n_{\text{swg}}(\lambda) \leftarrow \underset{n \in \mathbb{M}}{\text{argmin}}\, F(n, \lambda)$ ▷ Finding effective index values

$n_{\text{swg}}^{(d)}(\lambda) \leftarrow \underset{n \in \mathbb{M}}{\text{argmin}}\, F^{(d)}(n, \lambda)\ \ \forall\ d \in \{f, v, r\}$

end for

return $n_{\text{swg}}(\lambda),\ n_{\text{swg}}^{(d)}(\lambda)$

M_1 :	7	200	500	600	700	800	800	800	700	600	500	400	200	100	50	20
M_2 :	7	200	500	1200	1400	1600	1600	1600	1400	1200	500	400	200	100	50	20
M_3 :	7	200	500	1200	1400	1600	1600	1600	1400	1200	500	400	50	20		

Table S1. Architecture of the hidden layers of the three DNN models used for the training of three distinct 2D and three distinct 3D forward models. Each row lists the number of neurons in the hidden layer at that position, where the input layer (not shown) would be on the left, and the output layer (not shown) would be on the right.

S2. Model Performance

Here we present additional data pertaining to the performance of the DNN ensemble models discussed in Section 2.3.

In the case of the DNNs used, the number of neurons varies across the hidden layers, with the number increasing over several layers, remaining constant over several layers, and then decreasing. The layers with the larger number of neurons are the ones most suited to learn these non-linear features. This is important for the structures explored, as the electromagnetic interactions of these metamaterials have complex non-linear dependence on geometry and environmental features. The compressed layers at the end of the architecture reduce the number of features to only those that are relevant to our problem. Table S1 lists the number of neurons in the hidden layers for the three models (labeled as models M_1, M_2, and M_3) used for the 2D and 3D cases.

As in Fig. 3 of the main text, we present here absolute error histograms (top row) and prediction-truth scatter plots (bottom row) for the 2D (left) and 3D (right) DNN predictions. While we showed the results for the prediction of vertical transmission (v), here we show the additional results for the prediction of forward transmission (f) in Fig. S1 and reflection (r) in Fig. S2. The results are quite similar for all three cases.

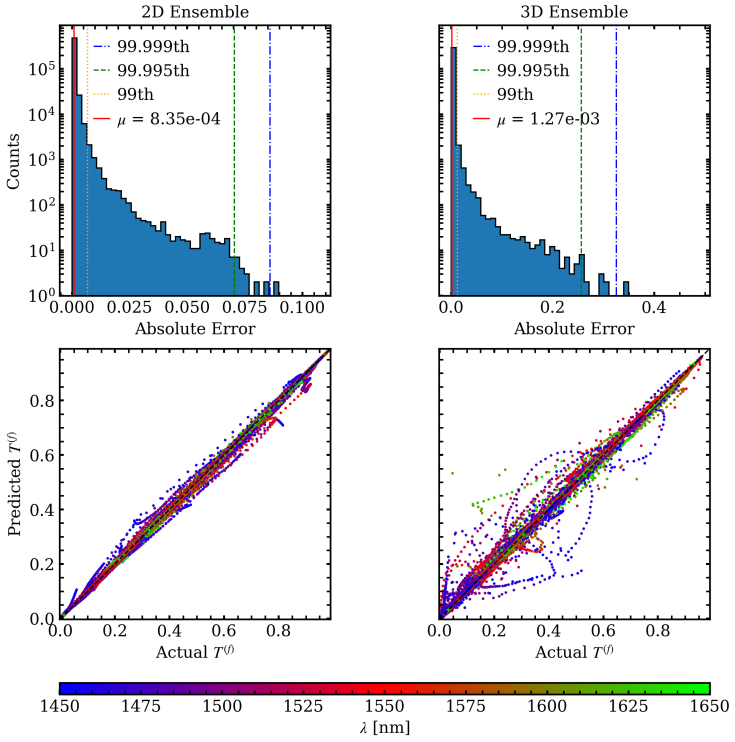


Fig. S1. Top row: Distribution of the absolute error of ensemble model predictions of forward transmission for 2D (left) and 3D (right). The vertical lines indicate the average (red solid), as well as the 99th (orange dotted), 99.995th (green dashed), and 99.999th (blue dash-dotted) percentiles. Bottom row: Predicted versus actual forward transmission over the wavelength range of interest. The colour of each data point indicates its wavelength, according to the colour scale.

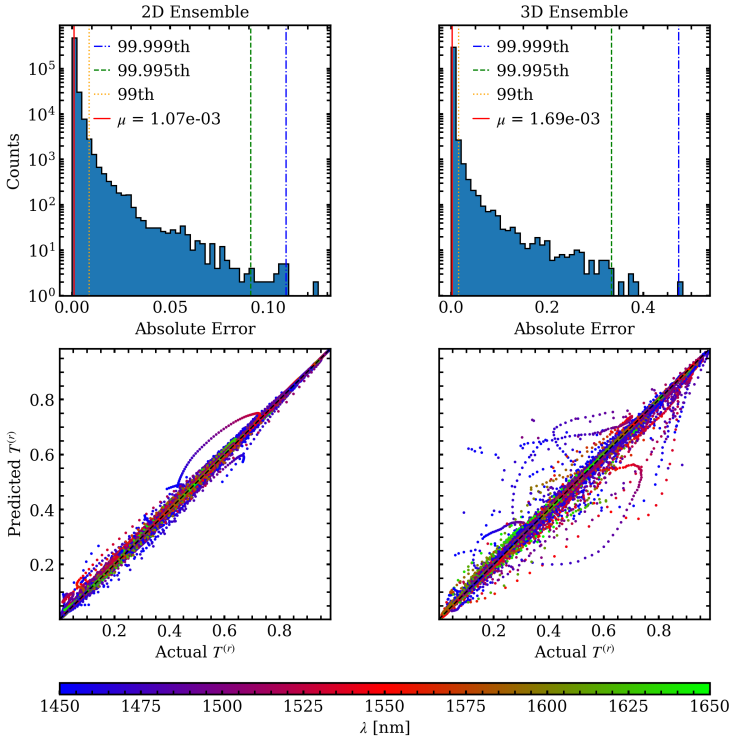


Fig. S2. Same as for Fig. S1, except that the data is for reflection predictions.

S3. Contour Maps of Error Functions for each direction for structures A, B, C

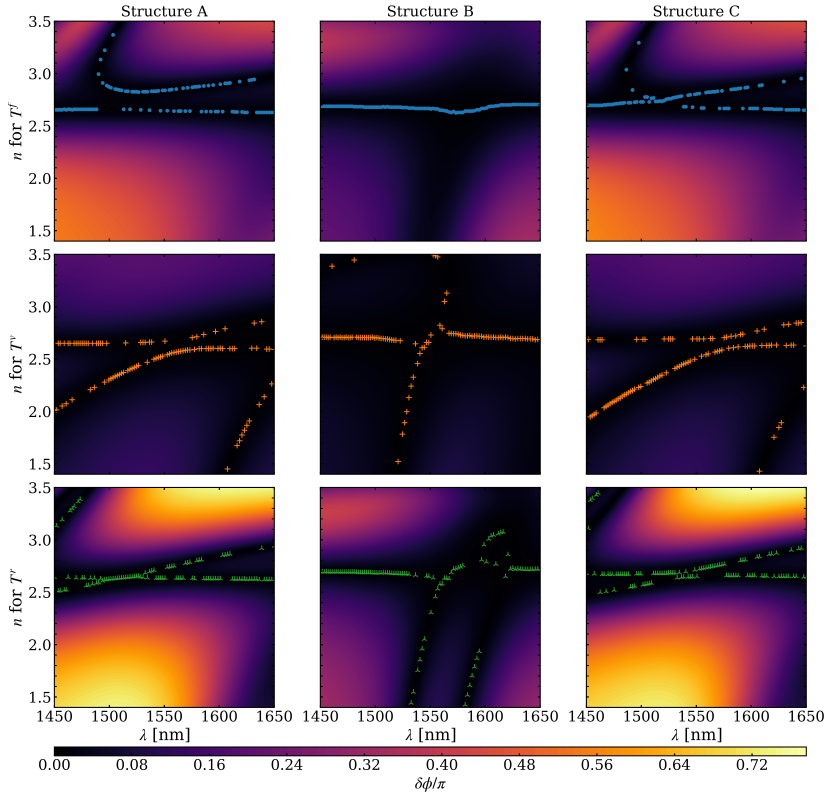


Fig. S3. Contour map $F^{(d)}(\lambda, n)$ for structures A (left), B (centre), and C (left) as a function of index n and wavelength λ where $d \in \{f, v, r\}$ represents the forward (top row), vertical (middle row) and reflection (bottom row), respectively; contours shown in linear scale. The markers represent the candidate $n_{\text{swg}}^{(d)}(\lambda)$ values for f (top row, blue circles), v (middle row, orange crosses), and r (bottom row, green three-pointed stars).

S4. Effective index for a second set of structures

To show that our results are robust, we repeat the procedure of Section 3 and 4 in the main text for another three structures whose parameters are listed Table S2. These were not chosen based on previously known optimal solutions (like those considered in the main text that minimize back-reflection). The parameters of Structure SA of Table S2 were chosen randomly, Structure SB is different from SA only in its value of SWG segment length L_2, and Structure SC is different from SA only in its value of SWG pitch Λ_z (and thus duty cycle).

Structure	L_1 [nm]	L_2 [nm]	L_3 [nm]	L_4 [nm]	Λ_z [nm]	a_z [nm]	Duty Cycle
SA	128	96	200	407	440	332	0.755
SB	128	150	200	407	440	332	0.755
SC	128	96	200	407	300	192	0.640

Table S2. Parameters for three arbitrary 3D structures. SA and SB differ only in L_2, and SA and SC differ only in the SWG parameters Λ_z and a_z (and consequently Duty Cycle).

Fig. S4 shows the candidate $n_{\text{swg}}^{(d)}$ values for structures SA (left), SB (centre) and SC (right), found from the $n \in \mathbb{M}$ that minimize $F^{(d)}(n, \lambda)$ for each λ, according to the procedure outlined

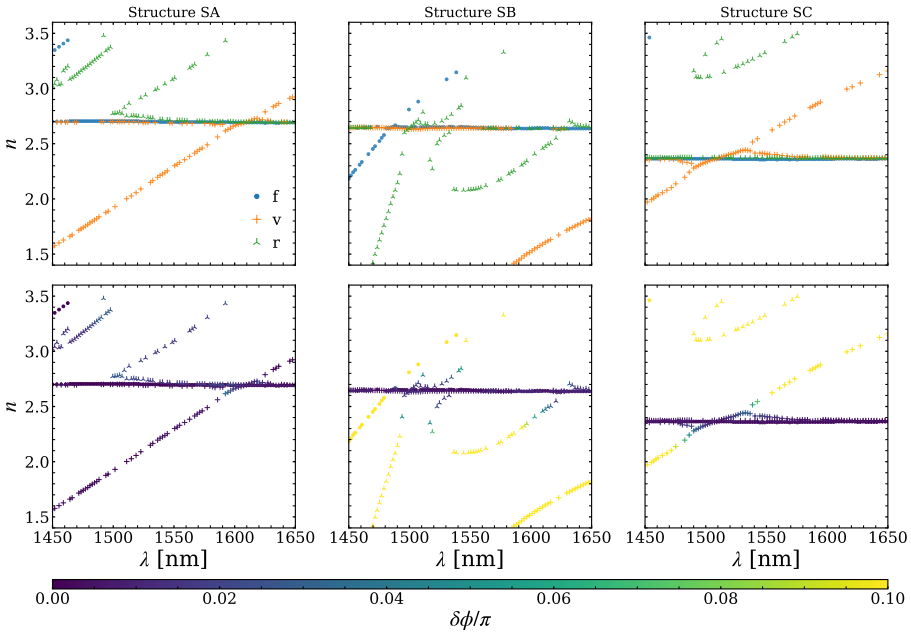


Fig. S4. Top row: Candidate $n_{\text{swg}}^{(d)}$ versus wavelength for structures SA (left), SB (middle), and SC (right) of Table S2 for $d = f$ (forward transmission, blue circles), $d = v$ (vertical transmission, orange crosses), and $d = r$ (reflection, green three pointed stars). Bottom row: Same data as top row, but where the marker colour indicates $\delta\phi^{(d)}$, which has been normalized by π.

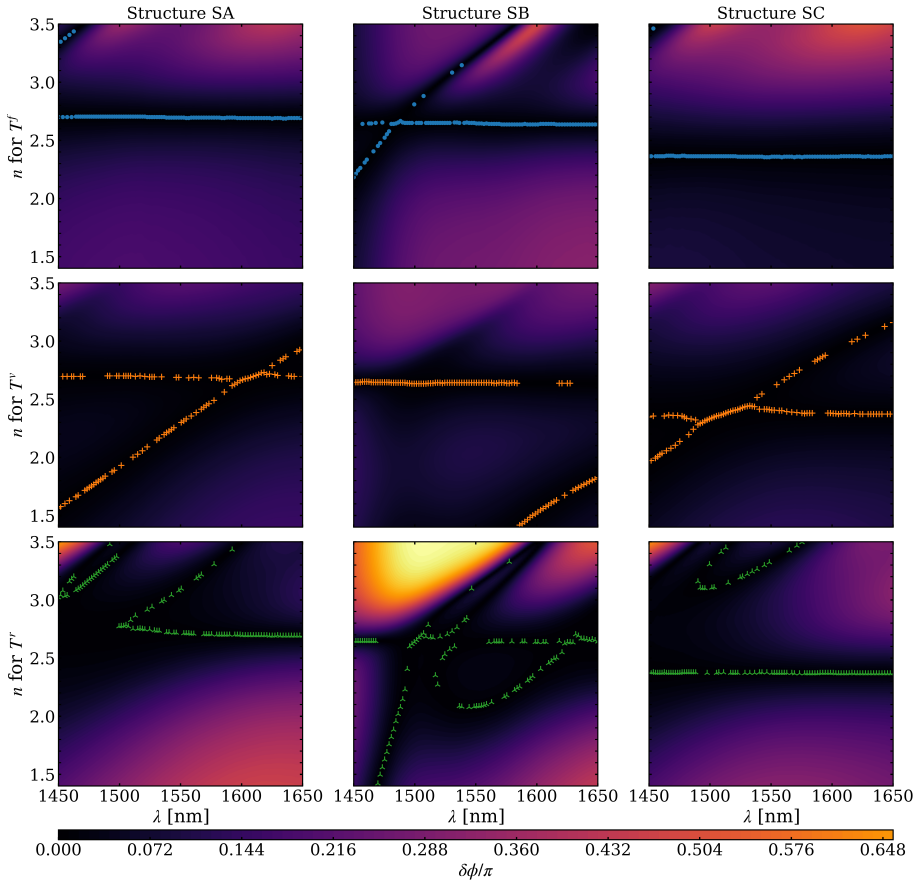


Fig. S5. Contour map of $F^{(d)}(n, \lambda)$ for structures SA (left), SB (middle), and SC (right) defined in Table S2.

in the main text. The markers of the top row are chosen according to the direction $d = f, v, r$ (blue circles, orange crosses, and green three-pointed stars, respectively). Contour maps for $F^{(d)}(n, \lambda)$ are shown in Fig. S5.

The colours of the markers in the bottom row of Fig. S4 are chosen according their values of $\delta\phi$, defined in Eq. 6 in the main text to be the difference in average transmission phase between 2D FDTD simulations with that particular candidate $n_{\text{swg}}^{(d)}$ solution, and 3D FDTD simulations of the full structure. Even for this set of structures, which are not good candidate grating coupler designs, we find that an overlap in candidate $n_{\text{swg}}^{(d)}$ values over $d = f, v, r$ corresponds to the lowest $\delta\phi$, indicating that these are the correct solutions.

Fig. S6 shows the fitting results for $\delta\phi$ (top) and δn (bottom) together with the fit for combining all directions (black points).

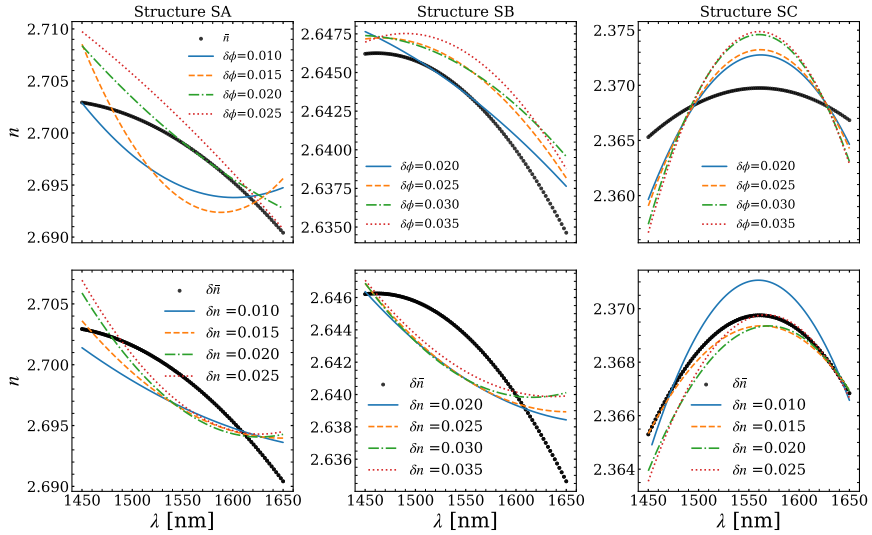


Fig. S6. Fit of thresholds for structures in Table S2 along with combining all directions.

Fig. S7 shows the slab model compared with our approach on structures SA, SB, and SC of Table S2.

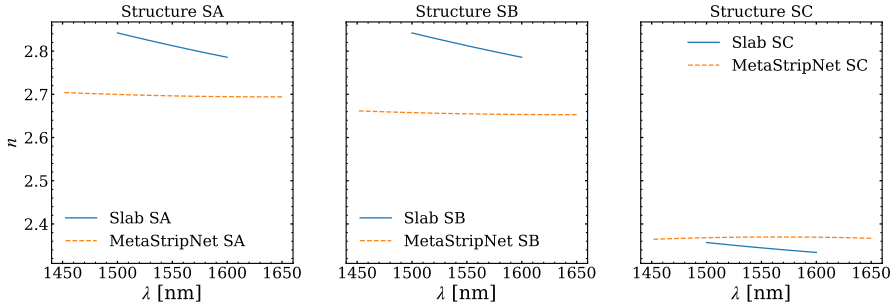


Fig. S7. Slab model (blue bold curve) compared with our Data-Driven Effective Medium Approximation (orange broken curve) for Table S2.

Fig. S8 shows the transmission plots for the structures of Table S2. As before, we have superior performance of our approach compared with the slab model.

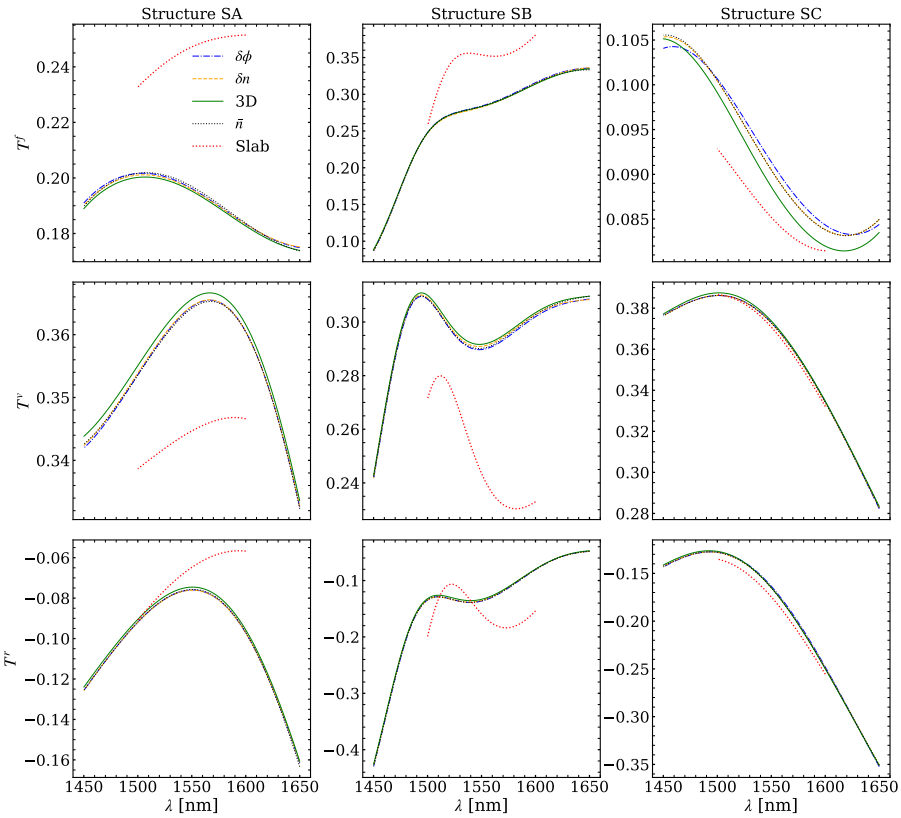


Fig. S8. Transmission plots for three more structures from Table S2.

S5. Algorithm for effective refractive index fitting procedure

Algorithm 2 Procedure for performing the fits for dispersive effective refractive index

Require: Directional $n_{\text{swg}}^{*(d)}(\lambda_i)$ from Pseudocode 1 ▷ $d \in \{f, v, r\}$ for forward transmission, vertical transmission and reflection respectively
Require: Wavelength grid $\{\lambda_i\}$
Require: Overlap tolerance ϵ
Ensure: Physically consistent dispersive index $\bar{n}_{\text{swg}}(\lambda)$

Step 1: Identify Directional Overlap
$\mathcal{S} \leftarrow \emptyset$
for each wavelength λ_i **do**
    $\delta n_{fr} \leftarrow |n_{\text{swg}}^{*(f)} - n_{\text{swg}}^{*(r)}|$
    $\delta n_{fv} \leftarrow |n_{\text{swg}}^{*(f)} - n_{\text{swg}}^{*(v)}|$
    $\delta n_{rv} \leftarrow |n_{\text{swg}}^{*(r)} - n_{\text{swg}}^{*(v)}|$
    $q \leftarrow 0$
    if $\delta n_{fr} < \epsilon$ **then**
        $q \leftarrow q + 1$
    end if
    if $\delta n_{fv} < \epsilon$ **then**
        $q \leftarrow q + 1$
    end if
    if $\delta n_{rv} < \epsilon$ **then**
        $q \leftarrow q + 1$
    end if
    if $q \geq 1$ **then** ▷ At least two curves overlap
        Add $(\lambda_i, n_{\text{swg}})$ to $\mathcal{S}$
    end if
end for
▷ $\mathcal{S}$ contains wavelengths where at least two directional solutions agree

Step 2: Polynomial Dispersion Fit
Fit a degree-p polynomial to points in $\mathcal{S}$:

$$\bar{n}_{\text{swg}}(\lambda) = \sum_{k=0}^{p} a_k \lambda^k$$

▷ $\bar{n}_{\text{swg}}(\lambda)$ is the dispersive index
return $\bar{n}_{\text{swg}}(\lambda)$

S6. Further Numerical Support: Running Average and Extra Minima

More precisely, the running average of the broken orange curve is smaller than the running average of the broken-dotted green curve for all directions. Let N be the number of wavelengths within a chosen window, and let n be a chosen index; the running average in this window is

$$\hat{F}^{(d)}(n_j^{(d)}, \lambda_i) = \frac{1}{N} \sum_{k=0}^{N-1} F^{(d)}(n_j^{(d)}, \lambda_{i+k}), \quad \forall d \in \{v, r, f\} \tag{S1a}$$

$$\hat{\bar{F}}(n_j, \lambda_i) = \frac{1}{N} \sum_{k=0}^{N-1} \bar{F}(n_j, \lambda_{i+k}) \tag{S1b}$$

Eq. S1a and Eq. S1b are the direction-dependent and combined running averages, respectively. We extract the indices after the running average is done, using the same approach presented in Eq. 5.

Practical implementation of our approach requires working without the phase, and thus numerical confirmation of the validity of the solutions of the overlap is preferred. There are numerical techniques to help with this. One of them is to perform a running average. A running average on the data of selected indices chosen from the global minima removes the outer bands of the dispersion, leaving the points along the overlap. This is demonstrated in Fig. S9, which is associated with the structures of Fig. S4.

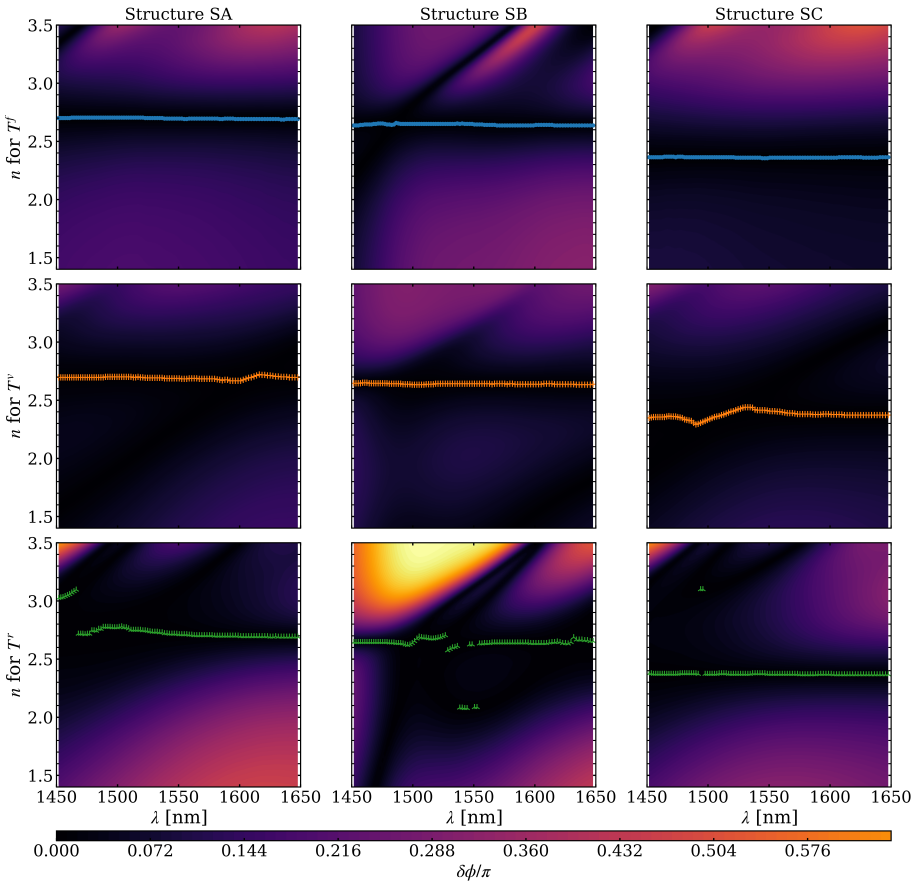


Fig. S9. Contour maps for selecting effective index without running average (top) and with 3-point running average (bottom) for Table S2.

5

Topological nanoresonator plasmonics: A quasinormal mode approach

In Chapter 4, we demonstrated on a few structures that it is realistically conceivable to extend the concept of effective medium to describe some SWGs. Though this is quite useful in its own right, it does not allow us to probe all kinds of optical properties in any periodic photonic material. For example, it would not allow us to probe exotic light-matter effects such as topological states on its own, especially in periodic plasmonic materials.

Topological states, in particular those that exchange energy with the rest of the universe (non-Hermitian/open systems), have interesting technological applications such as sensing, computing, and telecommunications. Periodic plasmonic materials are particularly attractive for studying open topological systems, as their ohmic losses render them naturally non-Hermitian. They allow for the engineering of the confinement of electromagnetic waves, which makes them attractive for technological applications.

Just like SWGs or any other periodic photonic material, periodic plasmonic materials have a large design space, which provides design flexibility, making them attractive for a wide variety of technological applications. This is in contrast with condensed matter systems that have a significantly smaller design space. As a result, in principle, we can design the topological properties of periodic plasmonic materials. However, they have proven to be quite a challenge to explore and understand, especially when it comes to building practical devices that leverage topological properties for operation. However, there are a limited number of theoretical approaches, each of which struggles to incorporate arbitrary shapes of resonators. In addition, they struggle to probe an arbitrary number of bands as well as the imaginary features of certain properties, often tied to long-range couplings, such as the imaginary part of the bandstructure.

Our main goal was to answer the question: in the case where we have a homogeneous background and the same type of particle (i.e., same shape, size, and material for the particles) in a unit cell, can we simulate only a single particle in that unit cell and build a theoretical model that can be used to guide physical intuition on the bulk-boundary behaviour of a periodic plasmonic material of interest? In the current work, we focused

on whether or not we could answer this question for spherical and cuboid-shaped nanoresonators in the paradigmatic one-dimensional SSH chain with air and glass backgrounds, respectively.

The foundations of this approach were discussed in Chapter 3, in particular, Section 3.3. The basis of this work was presented in [190] which we apply to study topology in periodic plasmonic materials. This approach uses the fundamental resonances, quasinormal modes, as basis states. We were in fact surprised that this tool had not yet been applied to study non-Hermitian topology in plasmonic systems, given the lack of formalism to include arbitrary nanoparticle shapes, long-range couplings, and other features which this formalism seems to handle well. Around the time of learning about this work, after realizing its potential in topological photonics, we saw a talk by Dr. Flore Kunst at one of the Max Planck-uOttawa meetings. There, she presented work going on in her group on electronic non-Hermitian topological systems. We reached out to her soon after, and she expressed interest in bringing the theoretical expertise from her group to study non-Hermitian topology in plasmonics. This established a collaboration between uOttawa and MPL, where uOttawa brought the computational nanoplasmonic expertise, and MPL brought the theoretical expertise. From the Kunst group, I worked with the PhD student, Julius Gohsrich. Such a synergy was important for this project, as even the theoretical formalism still required quite a bit of simulation expertise. This budding collaboration led to a trip to the MPL in September 2024. By then, I had worked out some initial calculations to show how work from [190] can be applied to topological plasmonics. I based my calculations on the work from [195], which had applied the work from [190] to a collection of metallic nanores-

onators. I applied this to a periodic chain of arbitrarily shaped metallic nanoparticles and compared it to both the plasmonic and condensed matter literature. This was then presented to the MPL group, who then extended and improved on my initial calculations.

We used the quasinormal modes of a single resonator along with the associated eigenfrequencies. The assumption is that the hybridized modes of the lattice can be written as a linear combination of the eigenmodes of a single resonator, where the scalar multiples of the eigenmodes are the excitation coefficients. We also terminate the linear combination as a finite sum, though the actual sum is infinite. The eigenmodes and eigenfrequencies together are used to compute the coupling between two arbitrary modes of interest. This information, along with the particle arrangements in the lattice, is used to construct a coupled-mode formalism for the system of interest. This formalism gives an eigenmode problem, whose eigenvalues are a set of eigenfrequencies that predict the bandstructure of the system. The eigenvectors are the associated excitation coefficients. Although I am listed as second author in this paper, I provided a significant contribution to the work that led to its development. I will explain this further below. I showed how we can write the equations from this formalism in a form where we can understand how this kind of system would evolve adiabatically. I used this argument to justify to myself that there must be the equivalent of a Berry phase in plasmonic systems. In parallel, I showed how the theoretical tools being developed could be solved numerically in order to get meaningful physical results. I used this to numerically compute the coupling strengths between the particles, which is necessary to compute everything from the bandstructure (bulk information) to the spectra of a finite chain, which we use to show the existence of edge modes

(boundary information). I then used this to demonstrate that we can simulate a single particle, whose eigenmodes and eigenfrequencies can be used, at low computational cost, to guide the exploration of topological states in plasmonic systems. Our work focused on structures embedded in a uniform background.

Once this was done, Julius Gohsrich expanded on my initial theoretical work to include long-range couplings and higher bands for both the periodic-boundary and open-boundary theoretical models. He first used Mie scattering to build the theory for spheres, which were embedded in air. These served to help establish a benchmark. These were later compared with COMSOL simulations. The next step was to do single-particle simulations for cuboid structures embedded in glass and, using the QNM coupled-mode formalism, predict the bands as well as the location of the edge modes in the spectra. The single-particle simulations and overlap integrals needed to build out the theoretical model were done by me. Julius used this to then build out the theoretical model for the cuboids. We studied the influence of particle separation and long-range couplings on the edge modes. In addition, we also studied edge modes in higher bands (beyond the lowest bands).

In summary, my contributions to this work involved the following: I conceived the main idea of applying coupled quasinormal mode (QNM) theory to topological plasmonics. I also constructed the initial nearest-neighbour coupled QNM model and computed the nearest-neighbour bandstructure for spherical and cuboid SSH chains, which was extended to long-range couplings by Julius. I then used this nearest-neighbour approximation to show that this QNM coupled-mode theory can predict edge modes and reproduce the general shape of the lowest bands. I also computed the mode couplings for arbitrary distances

from the single-particle simulations so that long-range couplings can be incorporated. I also helped to compare the spherical Mie scattering computations (done by Julius) with simulations. In addition, I co-analyzed the results.

Topological nanoresonator plasmonics: A quasinormal mode approach

Julius T. Gohsrich,[1,2,3,*] Norman S. Israel,[4,3] Lora Ramunno,[4,3] and Flore K. Kunst[1,2,3]

[1] *Max Planck Institute for the Science of Light, Staudtstraße 2, 91058 Erlangen, Germany*
[2] *Department of Physics, Friedrich-Alexander-Universität Erlangen-Nürnberg, Staudtstraße 7, 91058 Erlangen, Germany*
[3] *Max Planck – uOttawa Centre for Extreme and Quantum Photonics, University of Ottawa, Ottawa, ON K1N 6N5, Canada*
[4] *Department of Physics and Nexus for Quantum Technologies Institute, University of Ottawa, Ottawa, ON K1N 6N5, Canada*

(Dated: November 17, 2026)

Plasmonic nanoresonators allow for the subwavelength manipulation of light usable in several applications, such as sensing, microscopy and optical communication. Incorporating topological principles promises novel and robust functionalities. In this work, we set up a coupled quasinormal mode theory of topological arrays of arbitrarily shaped metallic nanoresonators, and analyze them using standard techniques from non-Hermitian physics. Within this framework, we introduce a plasmonic analog of the paradigmatic Su–Schrieffer–Heeger (SSH) model consisting of spherical metallic nanoresonators, and find full qualitative agreement with previous studies. In contrast to previous works, our framework is applicable beyond spherical nanoresonators, which we illustrate by discussing an SSH chain consisting of rectangular nanocuboids. In all presented cases, the quantized non-Hermitian Zak phase predicts the occurrence of exponentially localized edge states. We furthermore highlight the necessity of explicitly considering the long-range nature of the couplings between the resonators to properly capture important physical features. The presented methods allow for the optimization and design of topologically robust plasmonic nanostructures with potential applications in sensing, microscopy, and optical communication.

I. INTRODUCTION

Using topological principles to design structures to manipulate light has resulted in the vibrant field of topological photonics [1–4]. Inspired by condensed matter systems, such topological systems exhibit robust phenomena independent of geometric details, such as the presence of edge states, and may be protected from manufacturing imperfections or external perturbations. Photonic systems have a much larger design space and higher tunability compared to their condensed matter counterparts, allowing the implementation of topological models of greater variety. In this realm, plasmonics [5–8]—the study of the interaction between light and free electrons at metal-dielectric interfaces—promises novel functionalities due to the subwavelength manipulation of light, with potential applications in sensing, microscopy, and optical communication. Plasmonic systems inherently are dissipative due to intrinsic Ohmic losses, and thus fall in the category of non-Hermitian systems [9–11]. Having all these different ingredients raises the question on how to properly describe, analyze and characterize topological plasmonic systems.

One of the simplest topological systems is the paradigmatic Su–Schriefer–Heeger (SSH) model [12, 13], a one-dimensional chain having alternating couplings between the constituents of the chain, for example, by setting alternating distances as illustrated in Fig. 1. Hermitian SSH chains exhibit robust topological edge states, which are connected through a bulk–boundary correspondence

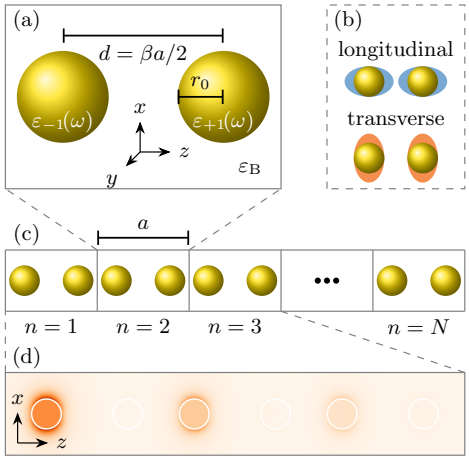


Figure 1. Plasmonic SSH chain. (a) The three-dimensional unit cell of the plasmonic SSH chain consists of two spherical metallic nanoresonators of radius r_0 and permittivity $\varepsilon_p(\omega)$ with $p = \pm 1$, which are embedded in a homogenous dispersionless background dielectric ε_B. The nanoresonators are separated by $d = \beta a/2$, where a is the length of the unit cell and β is the staggering factor; $\beta > 1$ corresponds to a topologically non-trivial chain featuring edge states. (b) All modes fall into two categories: longitudinal or transverse. Sketched is a false color map of the electric field magnitude $|\mathbf{E}(\mathbf{r})|$, and the color coding (blue and orange, respectively) is used throughout the manuscript. (c) To set up a plasmonic SSH chain, we repeat the unit cell along the z-direction, where n labels the N unit cells. Here, a topologically non-trivial configuration is illustrated. (d) The reconstructed QNM of the leftmost three unit cells ($n = 1, 2, 3$) of a transverse edge state, where $N = 8$, $\beta = 1.1$, $r_0 = 5\,\text{nm}$, and $a = 50\,\text{nm}$.

* julius.gohsrich@mpl.mpg.de

to a bulk winding number [13]. Non-Hermitian SSH chains can exhibit even richer phenomenology, such as the non-Hermitian skin effect and exceptional points—both truly non-Hermitian phenomena—resulting in a breakdown of the conventional bulk–boundary correspondence [9–11]. In the plasmonic realm, arrays of metallic nanoresonators [14] have been used to set up SSH chains. Such chains have been theoretically described and analyzed in several works [15–23], their connection to topological heat transport investigated [24, 25], and their topological edge states experimentally observed [26–29]. Other nanoresonator arrays implementing different topological geometries have also been discussed [30–38].

The aforementioned works on plasmonic SSH chains are usually modeled by coupled dipole equations [15–19, 23, 24]; experimental results are interpreted in this context and supplemented by full-wave simulations. However, these theoretical works are building on idealized geometries of the constituent nanoresonators, and coupled dipole equations require certain criteria to be met [39]. Using quasinormal modes (QNMs) [40–43], also referred to as resonant states [44–48], to model the properties of nanoresonators promises to overcome these challenges, as they are geometry-agnostic and allow for a straightforward multimodal treatment. Such QNMs are natural dissipative resonances and describe the response of a resonator to an external field. While often discussed in the context of single resonators, recent progress on understanding the interplay of multiple resonators has given rise to coupled QNM theories [49–51], see Table 1 in Ref. 51 for an overview.

In this work, we set up a coupled QNM theory [49] to analyze plasmonic analogs of the SHH model. We focus on chains of metallic nanospheres, as analytical expressions for the QNMs of the constituent metallic nanospheres are known [45–47], allowing an analytic treatment of the full SSH model and a direct comparison with previous theoretical studies [16, 17, 19]. As our approach is applicable to chains consisting of nanoresonators of arbitrary shape, whose associated QNMs may not be analytically known, we also analyze an SSH chain consisting of rectangular cuboids. We leverage computational methods to extract relevant quantities of the constituent cuboids [52], and perform a semi-analytical analysis of the full chain. In all cases, we find the anticipated edge states predicted by a quantized Zak phase, and we discuss the intricacies considering long-range effects.

II. QUASINORMAL MODE THEORY

In order to set up a coupled QNM theory for the plasmonic SSH chain, we start by reviewing and extending the relevant notions in the literature. In particular, we recount the theory of QNMs [40–43] for single metallic nanoresonators. Then, we discuss the relevant expressions for metallic nanospheres [45–47]. Thereafter, we introduce the coupled QNM theory from Ref. 49 and adapt it for our purposes, and finally, we introduce the relevant notions to handle periodic media.

To approach the topic of QNMs, let us start by stating the source-free Maxwell equations in the complex frequency domain

$$\begin{pmatrix} 0 & -\mathrm{i}\mu_0^{-1}\boldsymbol{\nabla}\times \\ \mathrm{i}[\varepsilon_0\varepsilon(\mathbf{r},\omega)]^{-1}\boldsymbol{\nabla}\times & 0 \end{pmatrix}\begin{pmatrix}\mathbf{H}(\mathbf{r})\\ \mathbf{E}(\mathbf{r})\end{pmatrix}=\omega\begin{pmatrix}\mathbf{H}(\mathbf{r})\\ \mathbf{E}(\mathbf{r})\end{pmatrix}. \quad (1)$$

Here, the time-independent fields $\mathbf{E}(\mathbf{r})$ and $\mathbf{H}(\mathbf{r})$ are the electric field and magnetic field strength at position $\mathbf{r}$, respectively, which are varying harmonically (e.g., $\mathbf{E}(\mathbf{r},t) = \mathrm{e}^{-\mathrm{i}\omega t}\mathbf{E}(\mathbf{r})$), and we analytically continue the frequency ω to be complex in general. Furthermore, ε_0 and μ_0 are the vacuum permittivity and permeability, respectively, and $\varepsilon(\mathbf{r},\omega)$ is the frequency-dependent relative permittivity. For our intends and purposes, this relative permittivity describes the spatial arrangement of P nanoresonators. Each resonator is labeled by $\boldsymbol{p}$, occupies a regions $V_{\boldsymbol{p}}$ in $\mathbb{R}^3$; has a surface $\partial V_{\boldsymbol{p}}$; and is modeled by a relative permittivity of any functional form $\varepsilon_{\boldsymbol{p}}(\omega)$ inside each resonator $\boldsymbol{p}$. These resonators are embedded in a uniform and dispersionless background with permittivity ε_B.

To define QNMs, one has to supplement Eq. (1) with the appropriate boundary conditions, namely: regular fields at the origin $\mathbf{r} = 0$; satisfied matching conditions at the metal-dielectric interfaces $\partial V_{\boldsymbol{p}}$; and outgoing fields for $|\mathbf{r}| \to \infty$ [40, 41, 44]. This makes Eq. (1) an eigenvalue equation, whose solutions are the QNMs. To proceed, we discuss the QNMs of a single resonator.

A. QNMs of a single resonator

We consider a single resonator (dropping $\boldsymbol{p}$ for now) which occupies a region V in $\mathbb{R}^3$ (including the origin), with surface ∂V describing the metal-dielectric interface, and relative permittivity

$$\varepsilon(\mathbf{r},\omega) = \begin{cases}\varepsilon(\omega), & \text{if } \mathbf{r}\in V,\\ \varepsilon_\mathrm{B}, & \text{else.}\end{cases} \quad (2)$$

Solving Eq. (1) using the discussed boundary conditions yields the resonance frequencies $\tilde{\omega}_{\boldsymbol{m}}$, the aforementioned natural dissipative resonances of the resonator, and the associated field modes $\tilde{\mathbf{E}}_{\boldsymbol{m}}(\mathbf{r})$ and $\tilde{\mathbf{H}}_{\boldsymbol{m}}(\mathbf{r})$. There are an infinite number of such modes indexed by $\boldsymbol{m}$, and we further decorate all quantities associated with a single resonator with a tilde ($\sim$) to distinguish them from all later occurring resonances and fields. Below, we will allow $\boldsymbol{m}$ to be a multi-index, but we set $\boldsymbol{m} = 1, 2, \ldots$ for now.

The advantage of using QNMs is that they allow one to express the solutions of a scattering problem as

$$\mathbf{E}(\mathbf{r}) = \sum_{\boldsymbol{m}=1}^{\infty} a_{\boldsymbol{m}}\,\tilde{\mathbf{E}}_{\boldsymbol{m}}(\mathbf{r}), \quad (3)$$

where $\mathbf{E}(\mathbf{r})$ is the scattered field, and the expansion coefficients $a_{\boldsymbol{m}}$ quantify how strong the field modes $\tilde{\mathbf{E}}_{\boldsymbol{m}}(\mathbf{r})$ are excited by the incident field. The validity of this expansion mathematically requires that the $\tilde{\mathbf{E}}_{\boldsymbol{m}}(\mathbf{r})$ are normalizable, complete and biorthogonal [42].

These requirements are met, for example, by solving the problem numerically [42]. Therefore, one discretizes space and introduces perfectly matched layers (PMLs), which are strongly absorbing, on all boundaries of the simulation domain. The solutions of such simulations fall into two categories: QNM modes and PML modes. The QNM modes are localized close to the resonator surface and thus excitable in a scattering experiment, whereas the PML modes are localized close to the PMLs and not excitable. However, the PML modes are necessary for the completeness of the basis, and thus relevant when determining physical observables [40].

This distinction allows one to split these two contributions in Eq. (3) as

$$\mathbf{E}(\mathbf{r}) = \mathbf{E}^{\mathrm{QNM}}(\mathbf{r}) + \mathbf{E}^{\mathrm{PML}}(\mathbf{r}). \tag{4}$$

In general, Eq. (3) is referred to as QNM expansion, even though one has to include the PML modes as highlighted by Eq. (4).

B. QNMs of a spherical nanoresonator

One analytically tractable problem is to find the QNMs of spherical nanoresonators. Here, we briefly revisit the relevant properties of the resonances and their associated field modes, discussed in Refs. 45–47; further details are deferred to App. A.

To solve Eq. (1) for spherical resonators of radius r_0, one makes an ansatz respecting the spherical symmetry and the boundary conditions at $\mathbf{r} = 0$ and $|\mathbf{r}| \to \infty$. This reduces the three-dimensional problem to one dimension, and requires matching conditions at $r = r_0$ only. Furthermore, this step naturally decouples the fields in two polarizations σ, namely, the transverse magnetic ($\sigma = \mathrm{TM}$) and the transverse electric ($\sigma = \mathrm{TE}$). For example, the matching conditions for TM modes (TE in App. A) are

$$\frac{j_{\ell+1}(nkr_0)}{n\, j_\ell(nkr_0)} - \frac{\ell+1}{n^2 kr_0} = \frac{h_{\ell+1}(n_{\mathrm{B}}kr_0)}{n_{\mathrm{B}}\, h_\ell(n_{\mathrm{B}}kr_0)} - \frac{\ell+1}{n_{\mathrm{B}}^2 kr_0}, \tag{5}$$

where $k = \tilde{\omega}/c_{\mathrm{vac}}$ is the resonant wavevector with c_{vac} the speed of light in vacuum; $n = \sqrt{\varepsilon(k)}$ and $n_{\mathrm{B}} = \sqrt{\varepsilon_{\mathrm{B}}}$ are the refractive indices inside and outside the spherical nanoresonator, respectively; and j_ℓ and h_ℓ are the spherical Bessel and spherical Hankel functions of the first kind, respectively.

Solving these transcendental equations numerically yields the resonance frequencies $\tilde{\omega}^\sigma_{\ell,\nu}$, from which the associated field modes $\tilde{\mathbf{E}}^\sigma_{\ell,\nu,m}(\mathbf{r})$ can be constructed as discussed in App. A. Beyond the polarization σ, these expressions are indexed by an orbital angular momentum number $\ell = 1, 2, \ldots$, an azimuthal angular momentum number $m = -\ell, \ldots, \ell$, and for each ℓ, there is an infinite number of resonances labeled by $\nu = 1, 2, \ldots$ [45, 46, 53, 54]. We also notice, that the resonances $\tilde{\omega}^\sigma_{\ell,\nu}$ are $(2\ell+1)$-fold degenerate with respect to m as a result of the spherical symmetry. Reducing the spherical symmetry to an axial one, e.g., considering an oblate spheroidal resonator, would partially lift this degeneracy (only $\pm m$ pairs remain degenerate), whereas breaking the axial symmetry would generically break all degeneracies [53, 54].

Now, we tie this back to the preceding discussion on single resonators by interpreting $\boldsymbol{m}$ as the multi-index $\boldsymbol{m} = (\sigma, \ell, \nu, m)$, so that $\tilde{\omega}_{\boldsymbol{m}} = \tilde{\omega}^\sigma_{\ell,\nu}$ and $\tilde{\mathbf{E}}_{\boldsymbol{m}}(\mathbf{r}) = \tilde{\mathbf{E}}^\sigma_{\ell,\nu,m}(\mathbf{r})$, where we have set σ as superscript for typographic reasons. This also allows us to write the QNM part in Eq. (4), in the case of a spherical nanoresonator, as

$$\mathbf{E}^{\mathrm{QNM}}(\mathbf{r}) = \sum_{\sigma=\mathrm{TE,TM}} \sum_{\ell=1}^{\infty} \sum_{\nu=1}^{\infty} \sum_{m=-\ell}^{\ell} a^\sigma_{\ell,\nu,m} \tilde{\mathbf{E}}^\sigma_{\ell,\nu,m}(\mathbf{r}). \tag{6}$$

In order to prevent confusion, we want to mention that the QNMs of spherical resonators are a cornerstone of the resonant state expansion (RSE) [44–48]. The RSE is an alternative approach to determine the QNMs of arbitrarily shaped nanoresonators, cf. Sec. II A, in a perturbative way. In this work, we will mainly use the language of the QNM expansion, i.e., the numerical approach discussed before, however, we will use the analytic expressions for the spherical resonators to illustrate our framework.

C. Coupled QNM theory

In general, the approach for a single resonator, cf. Sec. II A, can be used for (any arrangement of) several resonators. Therefore, one embeds several resonators $\boldsymbol{p}$ with relative permittivities $\varepsilon_{\boldsymbol{p}}(\omega)$ in a background medium with ε_{B} as outlined above. Again, imposing the appropriate boundary conditions, usually by employing PMLs, allows determining the QNMs of the full resonator arrangement by solving Eq. (1), which admits a QNM expansion as Eq. (3). However, this approach requires one to repeatedly determine the QNMs of the combined structure for any nanoresonator arrangement, resulting in a high computational demand. Furthermore, the determined collective QNMs might be hard to interpret as their physical nature is obscured.

The premise of coupled QNM theories [49–51] is to get around these limitations by determining the QNMs of the resonator arrangement in terms of their constituents QNMs as

$$\mathbf{E}(\mathbf{r}) = \sum_{\boldsymbol{p}} \sum_{\boldsymbol{m}} a_{\boldsymbol{p},\boldsymbol{m}}\, \tilde{\mathbf{E}}_{\boldsymbol{p},\boldsymbol{m}}(\mathbf{r} - \mathbf{r}_{\boldsymbol{p}}), \tag{7}$$

where the $\tilde{\mathbf{E}}_{\boldsymbol{p},\boldsymbol{m}}(\mathbf{r}-\mathbf{r}_{\boldsymbol{p}})$ are the field modes of the single nanoresonator, as discussed in Sec. II A, labeled by $\boldsymbol{p}$ and $\boldsymbol{m}$; localized at $\mathbf{r}_{\boldsymbol{p}}$; with resonance frequency $\tilde{\omega}_{\boldsymbol{p},\boldsymbol{m}}$. For now, we assume $\boldsymbol{p}=1,\ldots,P$, but we will later allow $\boldsymbol{p}$ to be a multi-index encoding all spatial degrees of freedom, in contrast to the multi-index $\boldsymbol{m}$, encoding all internal/modal degrees of freedom. Such an approach allows us to interpret the QNMs of the resonator arrangement as hybridized QNMs of the constituent resonators, and leads to a reduced computational demand.

In this work, we employ the coupled QNM theory from Ref. 49, which we slightly adapt for our purposes, and extend in Sec. II D. The central insight in Ref. 49 is that one can set up the coupled QNM theory using a multiple-scattering picture [41, 49]: All resonators excite each other, which is done by splitting the contribution of resonator $\boldsymbol{p}$ in Maxwell's equations, Eq. (1), from the contribution of all other resonators $\boldsymbol{p}'\neq\boldsymbol{p}$, which create source terms exciting $\boldsymbol{p}$. Using this insight and the properties of the QNMs, namely, their biorthonormality and completeness, yields the following coupled QNM equation [49]:

$$C\cdot\mathbf{a}=\omega\,\mathbf{a}, \tag{8}$$

where the coupling matrix C describes the coupling between the different QNMs, encoded in $\mathbf{a}$, the vector of coefficients $a_{\boldsymbol{p},\boldsymbol{m}}$ from Eq. (7). The coupling matrix C has elements

$$C^{\boldsymbol{p}',\boldsymbol{m}'}_{\boldsymbol{p},\boldsymbol{m}}=\tilde{\omega}_{\boldsymbol{p},\boldsymbol{m}}\delta^{\boldsymbol{p}'}_{\boldsymbol{p}}\,\delta^{\boldsymbol{m}'}_{\boldsymbol{m}}+(1-\delta^{\boldsymbol{p}'}_{\boldsymbol{p}})c^{\boldsymbol{p}',\boldsymbol{m}'}_{\boldsymbol{p},\boldsymbol{m}}, \tag{9}$$

where δ is the Kronecker delta. Writing out Eq. (8) with the coupling matrix elements of Eq. (9) reads

$$\sum_{\boldsymbol{p}'=1}^{P}\sum_{\boldsymbol{m}'=1}^{M_{\boldsymbol{p}'}}C^{\boldsymbol{p}',\boldsymbol{m}'}_{\boldsymbol{p},\boldsymbol{m}}\,a_{\boldsymbol{p}',\boldsymbol{m}'}=\omega\,a_{\boldsymbol{p},\boldsymbol{m}}.$$

Here, we have introduced a mode cutoff $M_{\boldsymbol{p}'}$ to restrict the expansion to the relevant modes of each resonator [40]; the exact limit is $M_{\boldsymbol{p}'}\to\infty$. The coupling coefficients $c^{\boldsymbol{p}',\boldsymbol{m}'}_{\boldsymbol{p},\boldsymbol{m}}$ in Eq. (9) are given by

$$c^{\boldsymbol{p}',\boldsymbol{m}'}_{\boldsymbol{p},\boldsymbol{m}}=-\tilde{\omega}_{\boldsymbol{p},\boldsymbol{m}}\Delta\varepsilon_{\boldsymbol{p}}(\tilde{\omega}_{\boldsymbol{p},\boldsymbol{m}})I^{\boldsymbol{p}',\boldsymbol{m}'}_{\boldsymbol{p},\boldsymbol{m}}, \tag{10}$$

where $\Delta\varepsilon_{\boldsymbol{p}}(\omega)=\varepsilon_{\boldsymbol{p}}(\omega)-\varepsilon_{\mathrm{B}}$. Here, we remind ourselves that $\varepsilon_{\boldsymbol{p}}(\omega)$ and ε_{B} are the relative permittivities inside and outside resonator $\boldsymbol{p}$, respectively. The overlap integrals $I^{\boldsymbol{p}',\boldsymbol{m}'}_{\boldsymbol{p},\boldsymbol{m}}$ are given by

$$I^{\boldsymbol{p}',\boldsymbol{m}'}_{\boldsymbol{p},\boldsymbol{m}}=\frac{\varepsilon_0}{\tilde{F}_{\boldsymbol{p},\boldsymbol{m}}}\iiint_{V_{\boldsymbol{p}}}\tilde{\mathbf{E}}_{\boldsymbol{p},\boldsymbol{m}}(\mathbf{r}-\mathbf{r}_{\boldsymbol{p}})\cdot\tilde{\mathbf{E}}_{\boldsymbol{p}',\boldsymbol{m}'}(\mathbf{r}-\mathbf{r}_{\boldsymbol{p}'})\,\mathrm{d}^3\mathbf{r}, \tag{11}$$

where $V_{\boldsymbol{p}}$ is the region resonator $\boldsymbol{p}$ occupies. The normalization coefficient $\tilde{F}_{\boldsymbol{p},\boldsymbol{m}}$ is given by

$$\tilde{F}_{\boldsymbol{p},\boldsymbol{m}}=\iiint\left\{\left[\frac{\partial(\omega\varepsilon_{\boldsymbol{p}}(\mathbf{r},\omega))}{\partial\omega}\right]_{\omega=\tilde{\omega}_{\boldsymbol{p},\boldsymbol{m}}}\varepsilon_0\,\tilde{\mathbf{E}}_{\boldsymbol{p},\boldsymbol{m}}(\mathbf{r})\cdot\tilde{\mathbf{E}}_{\boldsymbol{p},\boldsymbol{m}}(\mathbf{r})-\mu_0\,\tilde{\mathbf{H}}_{\boldsymbol{p},\boldsymbol{m}}(\mathbf{r})\cdot\tilde{\mathbf{H}}_{\boldsymbol{p},\boldsymbol{m}}(\mathbf{r})\right\}\mathrm{d}^3\mathbf{r},$$

where the volume integral incorporates the entire simulation domain, i.e., the nanoresonator region, the background region as well as the PML region [42]. This quantity is automatically determined when computing the QNMs in the commercial solver COMSOL Multiphysics.

Let us connect Eqs. (8-11) back to the aforementioned multiple scattering picture [41, 49]: The first term in Eq. (9) describes that resonator $\boldsymbol{p}$ creates a field distribution according to $\boldsymbol{m}$. All other resonators $\boldsymbol{p}'\neq\boldsymbol{p}$ also create a field distribution, and excite $\boldsymbol{p}$ due to their radiated fields according to the second term. The strength of this excitation of $\boldsymbol{p}$ is encoded in the coupling coefficients Eq. (10) via the overlaps Eq. (11): It is proportional to the scalar product of the field inside $\boldsymbol{p}$ and the radiated field of $\boldsymbol{p}'$, integrated over the region $V_{\boldsymbol{p}}$. Solving the eigenvalue equation, Eq. (8), determines solutions where all resonators are in a radiative equilibrium.

Setting up Eqs. (8-11) depends on two approximations: (i) We truncate the infinite series Eq. (7) by introducing the cutoffs M_p so that Eq. (8) is a proper eigenvalue equation, and (ii), we assume that C is slowly varying, i.e., $\varepsilon_p(\omega)$ is slowly varying, in the frequency region of interest. We note that the latter approximation leads to a different expression for the coupling coefficients $c^{\boldsymbol{p}',\boldsymbol{m}'}_{\boldsymbol{p},\boldsymbol{m}}$ in Eq. (10) compared to Ref. 49. There, the full expression is $c^{\boldsymbol{p}',\boldsymbol{m}'}_{\boldsymbol{p},\boldsymbol{m}}=-\omega\,\Delta\varepsilon_{\boldsymbol{p}}(\omega)\,I^{\boldsymbol{p}',\boldsymbol{m}'}_{\boldsymbol{p},\boldsymbol{m}}$. Thus, without our approximation, Eq. (8) is a nonlinear eigenvalue equation reading $C(\omega)\cdot\mathbf{a}=\omega\,\mathbf{a}$. This can also be understood by realizing that Eq. (1) seen as an eigenvalue equation is also nonlinear, as the left-hand side depends on ω via $\varepsilon(\mathbf{r},\omega)$. In Sec. III, we will assume a Drude model, where $\varepsilon_p(\omega)$ is slowly varying in the frequency region of interest. Thus, we use approximation (ii) for simplicity, as incorporating the nonlinearity would only lead to minor quantitative changes. When assuming such a Drude model, or an N-pole Lorentz-Drude model, one can alternatively map the full nonlinear eigenvalue equation back to a linear eigenvalue equation without approximations [51, 52, 55, 56].

Let us reiterate the utility of employing a coupled QNM theory: Having determined the relevant resonances $\tilde{\omega}_{\boldsymbol{p},\boldsymbol{m}}$ and their associated field modes $\tilde{\mathbf{E}}_{\boldsymbol{p},\boldsymbol{m}}(\mathbf{r})$ of all constituent resonators $\boldsymbol{p}$, the problem of finding the QNMs of the combined structure reduces to determining the overlap integrals, Eq. (11), and solving an eigenvalue equation, Eq. (8). This typically reduces the computational complexity or increase accuracy compared to full-wave simulations depending on the total number of QNMs $M=\sum_p M_p$ considered.

From here, one can set up the scattering theory [49] and determine the far-field scattering amplitudes [57]. However, in this work we are mainly concerned with the collective resonances ω.

D. Coupled QNM theory for periodic media

As we are ultimately interested in the properties of a plasmonic analog of the SSH model, we have to discuss

quasi one-dimensional arrangements of metallic nanoresonators. However, our approach is straightforwardly generalizable to two- or three-dimensional arrangements with standard prescriptions, which allows the analysis of a wide range of nanoparticle arrangements implementing lattice models [30–37]. We want to emphasize that the dimensionality of the lattice should not be confused with the dimensionality of its constituents—we are considering quasi one-dimensional lattices of three-dimensional resonators.

Let us start by considering a finite chain of metallic nanoresonators. Therefore, it is instructive to define a unit cell, which is repeated N times, without loss of generality, along the z-direction; the associated unit vector is $\hat{\mathbf{e}}_z$. We label each unit cell by n, and label the P resonators within the unit cell at positions $\mathbf{r}_p$ by p. Hence, each resonator in the chain is labeled by (n,p) and is located at

$$\mathbf{R}_{n,p} = na\hat{\mathbf{e}}_z + \mathbf{r}_p \tag{12}$$

with $n = 1,\dots,N$ and a the length of the unit cell. This is illustrated in Fig. 1(a,c) for $P = 2$, which is the relevant case for the plasmonic SSH model.

Having set the geometry of the chain, one can employ the equations discussed in the previous section by interpreting $\boldsymbol{p}$ as the spatial multi-index $\boldsymbol{p} = (n,p)$ and setting $\mathbf{r}_{\boldsymbol{p}} = \mathbf{R}_{n,p}$. Furthermore, we require that the QNM cutoffs for each resonator in each unit cell is the same, i.e., $M_{n,p} = M_p$. In the language of topological lattice models, this case is referred to as a finite chain under open boundary conditions (OBCs). We will use finite chain and system under OBCs interchangeably throughout this work.

In this context, the second relevant case is the system under periodic boundary conditions (PBCs), which we will also refer to as periodic system. Now we assume translational invariance in the z-direction, which allows us to go to a momentum-space description using standard techniques [13][58]. Therefore, we use

$$a_{n,p,\boldsymbol{m}} = \frac{1}{\sqrt{N}}\sum_q \mathrm{e}^{\mathrm{i}naq} a_{p,\boldsymbol{m}}(q), \quad \text{with } q = \frac{2\pi}{aN}j, \tag{13}$$

where $j = 0,\dots,N-1$. Inserting this into Eq. (8) and performing some algebra, this yields the momentum-space analogs of Eqs. (8) and (9),

$$C(q)\cdot\mathbf{a}(q) = \omega(q)\,\mathbf{a}(q), \tag{14}$$

and

$$C^{p',\boldsymbol{m}'}_{p,\boldsymbol{m}}(q) = \tilde{\omega}_{p,\boldsymbol{m}}\delta^{p'}_{p}\delta^{\boldsymbol{m}'}_{\boldsymbol{m}} + \sum_{n=-\infty}^{\infty}(1-\delta^n_0\delta^{p'}_p)c^{n,p',\boldsymbol{m}'}_{0,p,\boldsymbol{m}}\, e^{\mathrm{i}naq}, \tag{15}$$

respectively. Here, $C(q)$ is the coupling matrix in momentum space, its eigenvalues $\omega(q)$ are the frequency bands, and the eigenvectors $\mathbf{a}(q)$ encode the QNMs in momentum space. The form of the coupling coefficients and overlap integrals, Eqs. (10) and (11), stay unaltered except for setting $\boldsymbol{p} = (n,p)$ and $\mathbf{r}_{\boldsymbol{p}} = \mathbf{R}_{n,p}$.

This formulation allows us to establish a formal one-to-one correspondence between this coupled QNM theory and tight-binding models in solid state physics: The coupling matrix under PBCs $C(q)$ corresponds to a Bloch Hamiltonian $H(q)$, the frequency bands $\omega(q)$ to energy-bands $E(q)$, the fields $\tilde{\mathbf{E}}_{p,\boldsymbol{m}}(\mathbf{r})$ to Wannier functions, and the coupling coefficients $c^{\boldsymbol{p}',\boldsymbol{m}'}_{\boldsymbol{p},\boldsymbol{m}}$ correspond to hopping strengths in the tight-binding picture. A conceptional difference is that in the plasmonic case, the real-space fields $\tilde{\mathbf{E}}_{p,\boldsymbol{m}}(\mathbf{r})$ come directly from Eq. (1), whereas the Wannier functions are constructed from the momentum-space description. Importantly, the plasmonic system is non-Hermitian, i.e., $C(q)^\dagger \neq C(q)$ [9, 10], in the sense that it is lossy, as all resonance frequencies ω are complex with negative imaginary part. The necessary non-Hermitian treatment introduces some extra technicalities, which we discuss below.

III. THE PLASMONIC SSH CHAIN CONSISTING OF NANOSPHERES

Having discussed the relevant notions in the previous section, we now set up a coupled QNM theory of the simplest topological model—a plasmonic analog of the SSH model [12, 13], and analyze its properties.

We consider $P = 2$ spherical nanoresonators of radius $r_0 = 5\,\mathrm{nm}$ per unit cell as illustrated in Fig. 1(a). Within a unit cell, $p = \pm 1$ labels the two resonators, d is the distance between the centers of the resonators, and $a = 50\,\mathrm{nm}$ is the length of the unit cell. The finite chain consists of N unit cells, i.e., $2N$ nanospheres over all, as illustrated in Fig 1(c). All nanosphere centers are located along the z-axis as

$$\mathbf{R}_{n,p} = R_{n,p}\hat{\mathbf{e}}_z, \quad \text{with} \quad R_{n,p} = na + p\,d/2, \tag{16}$$

where the unit cells are labeled by $n = 1,2,\dots,N$, cf. Eq. (12).

For each nanoresonator, we chose the Drude model as simplest model to capture metallic behavior. Its permittivity is given by

$$\varepsilon(\omega) = \varepsilon_\infty - \frac{\omega_\mathrm{p}^2}{\omega(\omega+\mathrm{i}\gamma)}, \tag{17}$$

where ε_∞ is the high-frequency dielectric constant, ω_p is the plasma frequency, and γ is the damping rate. More specifically, we use the parameters for gold, i.e., $\varepsilon_\infty = 9.1$, $\omega_\mathrm{p} = 1.38\times10^{16}\,\mathrm{rad/s}$, and $\gamma = 1.08\times10^{14}\,\mathrm{rad/s}$ [17, 59], and we embed it in vacuum, i.e., we set $\varepsilon_\mathrm{B} = 1$. Such plasmonic SSH chains consisting of Drude nanospheres have been investigated previously in the context of coupled dipole or multipole equations [16, 17, 19, 20], allowing us to directly compare results.

Similar to the conventional SSH model, we only consider the lowest-frequency resonance of each individual nanosphere, and their associated QNMs. As discussed in Sec. II B, we solve Eq. (5) (and Eq. (A2)) for our parameters and find that the resonance of a single nanosphere with the lowest real part has the resonant frequency $\tilde{f}_{1,1}^{\mathrm{TM}} \approx (658.8 - 8.6\mathrm{i})\,\mathrm{THz}$ or resonant wavelength $\tilde{\lambda}_{1,1}^{\mathrm{TM}} \approx (455.0 + 5.9\mathrm{i})\,\mathrm{nm}$, with their associated field modes $\tilde{\mathbf{E}}_{1,1,m}^{\mathrm{TM}}(\mathbf{r})$ with $m = -1, 0, +1$. Let us remark that these QNMs are labeled by $\ell = 1$, which gives this low-frequency approximation a dipolar character—however, this approach should not be set equal to a dipole expansion [16, 17] (or multipole expansion for a cutoff with larger ℓ [20]), as all QNMs with $\ell = 1$ and arbitrary $\nu \geq 2$ have dipolar character. For brevity, we drop all fixed indices of the QNMs, i.e., $\tilde{\omega}_{\boldsymbol{m}} = \tilde{\omega}_{1,1}^{\mathrm{TM}} \equiv \tilde{\omega}$ and $\tilde{\mathbf{E}}_{\boldsymbol{m}}(\mathbf{r}) = \tilde{\mathbf{E}}_{1,1,m}^{\mathrm{TM}}(\mathbf{r}) \equiv \tilde{\mathbf{E}}_m(\mathbf{r})$ with $m = -1, 0, +1$.

The preceding considerations, i.e., setting the geometry, the material model, and the mode-cutoff in the previous paragraph (cf. discussion below Eq. (9)), fully set our problem.

A. Coupling coefficients

As a first step towards analyzing our plasmonic SSH chain, we are considering the coupling coefficients and overlap integrals in Eqs. (10) and (11), respectively. These quantities are used to construct the finite and periodic chains, and are thus instrumental to understand these systems.

We start by simplifying the overlap integrals from Eq. (11). Therefore, we substitute $\boldsymbol{m} = (\mathrm{TM}, 1, 1, m) \equiv m$, $\boldsymbol{p} = (n, p)$ as discussed in Sec. II D, and $\mathbf{r}_{\boldsymbol{p}} = \mathbf{R}_{n,p}$ from Eq. (16). Furthermore, as we consider a chain of spherical nanoresonators, which all have the same radius and same permittivity, we can drop the spatial indices $\boldsymbol{p}$ from the field modes, i.e., $\tilde{\mathbf{E}}_{\boldsymbol{p},\boldsymbol{m}}(\mathbf{r}-\mathbf{r}_{\boldsymbol{p}}) \equiv \tilde{\mathbf{E}}_{\boldsymbol{m}}(\mathbf{r}-\mathbf{r}_{\boldsymbol{p}})$, i.e., we have the same QNMs located at different positions. This yields

$$\begin{aligned} I_{n,p,m}^{n',p',m'} &= \frac{1}{2}\iiint_{V_{n,p}} \tilde{\mathbf{E}}_m(\mathbf{r}-\mathbf{R}_{n,p}) \cdot \tilde{\mathbf{E}}_{m'}(\mathbf{r}-\mathbf{R}_{n',p'})\,\mathrm{d}^3\mathbf{r}, \\ &= \frac{1}{2}\iiint_{\mathbb{B}} \tilde{\mathbf{E}}_m(\mathbf{r}) \cdot \tilde{\mathbf{E}}_{m'}(\mathbf{r}-R_{n'-n,p'-p}\hat{\mathbf{e}}_z)\,\mathrm{d}^3\mathbf{r}. \end{aligned} \tag{18}$$

For the second equality, we have spatially translated the argument of the modes and the integration domain along z, used $R_{n',p'} - R_{n,p} = R_{n'-n,p'-p}$, and introduced the shorthand notation $\mathbb{B}$ for the three-dimensional solid sphere of radius r_0 at the origin. Also note that we have set $\tilde{F}_m = 2\varepsilon_0$, yielding the factor 1/2 in front of the integral, to match the normalization in App. A [46].

Thus, the coupling coefficients, Eq. (10), are

$$\begin{aligned} c_{n,p,m}^{n',p',m'} &= -\tilde{\omega}\,\Delta\varepsilon(\tilde{\omega})\,I_{n,p,m}^{n',p',m'} = c_m^{m'}(R_{n'-n,p'-p}\hat{\mathbf{e}}_z) \\ &= c_m(R_{n'-n,p'-p})\,\delta_m^{m'}, \end{aligned} \tag{19}$$

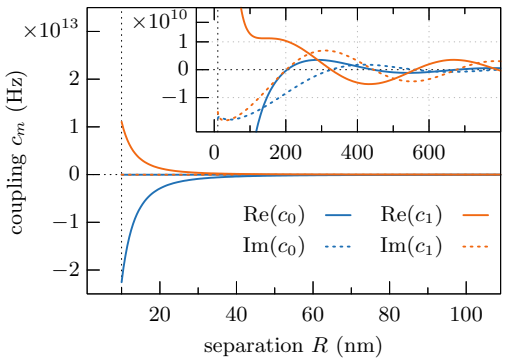


Figure 2. Coupling coefficients between two nanospheres. Real (solid) and imaginary (dashed) parts of the coupling coefficients c_0 (blue) and c_1 (orange) as function of the separation R. For $R \lesssim 100\,\mathrm{nm}$, the real parts of the couplings dominate. The vertical black dashed lines at $R = 10\,\mathrm{nm}$ corresponds to the limit where both nanospheres touch. The inset shows larger R with the y-axis rescaled to emphasize the imaginary parts of the couplings. It shows that the real and imaginary parts of the couplings are of similar magnitude for $R \gtrsim 200\,\mathrm{nm}$.

where we refer to the functions c_m as the couplings. The first line clarifies the structure of the coupling coefficients: $c_m^{m'}$ can be interpreted as a matrix encoding the internal degrees of freedom m and m', i.e., the modes, whereas the spatial degrees of freedom, (n, p) and (n', p'), are entirely encoded in the argument depending on $R_{n'-n,p'-p}$, the distance R between the two nanospheres. Due to the convention for the $\tilde{\mathbf{E}}_m(\mathbf{r})$ from App. A and our choice of aligning the chain along the z-direction, these matrix elements are diagonal ($\propto \delta_m^{m'}$) in the second line. Furthermore, this choice also implies that the couplings satisfy $c_{-1}(R) = c_{+1}(R)$. This is a result of the axial symmetry of our plasmonic SSH chain, as the two directions perpendicular to the chain (x- and y-directions in our case) must be equivalent. This distinction will give rise to the distinction between longitudinal modes ($m = 0$) and transverse modes ($m = \pm 1$).

Let us discuss the couplings. Their real and imaginary parts are shown in Fig. 2 as a function of the center-to-center separation R. For short separations ($R \lesssim 100\,\mathrm{nm}$), we see that the real parts of the couplings dominate and rapidly decay. The inset shows the couplings for larger separations, depicting a transition in the region $100\,\mathrm{nm} \lesssim R \lesssim 200\,\mathrm{nm}$ into an oscillatory decay of both the real and imaginary parts of the couplings for separations $R \gtrsim 200\,\mathrm{nm}$. Over all, this behavior of the couplings matches our expectation: For short separations, the real parts of the couplings dominate, whereas for longer separations, radiative couplings, i.e., the imaginary parts, come into play. The oscillatory decay matches the behavior of individual QNMs, which show similar behavior in their field amplitudes [42].

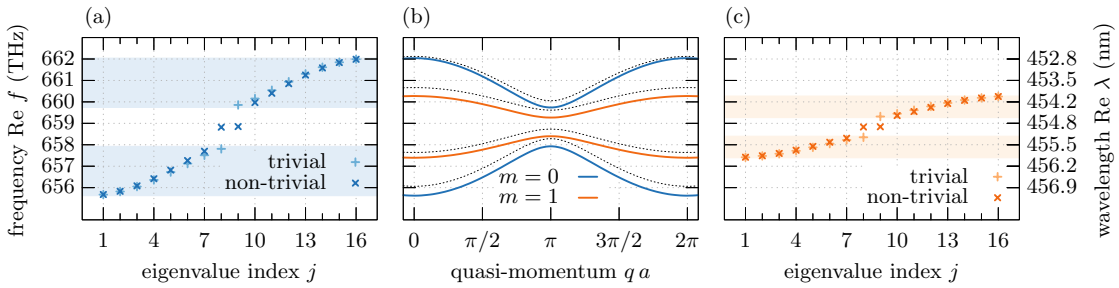


Figure 3. Spectra with nearest-neighbor coupling. (a) Spectra under OBCs of the longitudinal modes ($m = 0$) with 16 nanospheres. Light blue '+' markers indicate the spectrum for the topologically trivial configuration ($\beta = 0.9$), whereas dark blue '×' markers indicate the spectrum for the topologically non-trivial configuration ($\beta = 1.1$). The PBC band widths, cf. (b), are indicated by the colored regions. (b) Spectrum under PBCs of the longitudinal modes (blue lines, $m = 0$) and transverse modes (orange lines, $m = 1$) with $\beta = 0.9$, or equivalently, $\beta = 1.1$. The black dashed curves correspond to the results of Ref. 16. (c) Same as (a) for transverse modes ($m = 1$) with corresponding shades of orange.

B. Nearest-neighbor approximation

As the conventional SSH model considers nearest-neighbor couplings only, we first discuss this case before tackling the intricacies of long-range coupling in Sec. IV. To implement this, we truncate the infinite series in Eq. (15) to only include the couplings between nanospheres at $R_{n,-1}$ and at $R_{n,+1}$ via $c_m(d)$, and between nanospheres located at $R_{n,+1}$ and $R_{n+1,-1}$ via $c_m(a-d)$, cf. Eqs. (16) and (19). For convenience, and in order to closely match the notation of Ref. 13, we introduce

$$v_m = c_m(d), \quad w_m = c_m(a-d), \quad \text{and} \quad u_m = \tilde{\omega}, \tag{20}$$

where v_m is the intra-cell coupling, w_m the inter-cell coupling, and u_m is a constant onsite term. We also remind ourselves that $m = -1, 0, +1$, and the couplings $c_m(R)$ are identical for $m = \pm 1$.

C. Nearest-neighbor spectra under PBCs

Let us start by considering the periodic chain. Starting from Eq. (14), we find a 6×6 matrix $C(q)$. Due to the diagonal form of the coupling coefficients (Eq. (19)), and the constant diagonal (cf. first term in Eq. (15)), we can decouple $C(q)$ into three 2×2 matrices of the form

$$C_m(q) = \begin{pmatrix} u_m & v_m + w_m e^{-\mathrm{i}aq} \\ v_m + w_m e^{+\mathrm{i}aq} & u_m \end{pmatrix}. \tag{21}$$

For real v_m, w_m and u_m, Eq. (21) corresponds to the conventional, Hermitian, SSH model with a constant onsite term [13]. For our plasmonic SSH chain, however, the coupling coefficients are in general complex, and thus this model unsurprisingly corresponds to a non-Hermitian SSH model. Its spectrum is given by the two frequency bands (labeled by $\pm$)

$$\omega_{m,\pm}(q) = \tilde{\omega} \pm \sqrt{v_m^2 + w_m^2 + 2 v_m w_m \cos(aq)}, \tag{22}$$

where $m = 0$ corresponds to longitudinal modes, and $m = \pm 1$ to transverse modes, cf. Fig. 1(b). As the couplings for $m = \pm 1$ are identical, the corresponding spectra are degenerate. Thus, we do not further distinguish the two transverse cases, and simply refer to both transverse cases as $m = 1$.

In Fig. 3(b), the real parts of the spectra for $r_0 = 5\,\mathrm{nm}$, $a = 50\,\mathrm{nm}$ and $\beta = 2d/a = 1 \pm 0.1$ are shown. We show the frequencies f in Hertz and the corresponding wavelengths λ. The longitudinal modes ($m = 0$) and transverse modes ($m = 1$) are shown in blue and orange, respectively. Additionally, we show the results obtained from coupled dipole equations using a quasistatic approximation [16, 17, 19], which we refer to as quasistatic approximation, as black dashed curves. This approximation is applicable as long as $d, a - d \geq 3r_0$ [39].

We find great qualitative agreement between our coupled QNM approach and the quasistatic approximation: The main difference being a slight shift in the spectrum. This might at least be partially explained by the fact, that the coupled QNM theory is build on the full resonance of the metallic nanosphere known from Mie scattering, with $\mathrm{Re}\,\tilde{f} \approx 658.8\,\mathrm{THz}$, whereas the quasistatic approach approximates this by $\omega_\mathrm{p}/\sqrt{\varepsilon_\infty + 2\varepsilon_\mathrm{B}} \approx 659.2\,\mathrm{THz}$ [17]. The difference between these predictions for the resonance frequencies is of the order of the difference in the spectrum. The imaginary parts of the spectra are shown in App. B 1; we will discuss them in the context of long-range couplings in Sec. IV A.

D. Nearest-neighbor spectra under OBCs

Turning our attention to the finite chain, one can either repeat the previous considerations through Sec. II C, or know directly [13] that $C_m(q)$ under PBCs from Eq. (21) corresponds to

$$C_m = \begin{pmatrix} u_m & v_m & 0 & & & & \\ v_m & u_m & w_m & 0 & & & \\ 0 & w_m & u_m & v_m & 0 & & \\ & 0 & v_m & u_m & w_m & 0 & \\ & & \ddots & \ddots & \ddots & \ddots & \ddots \\ & & & 0 & w_m & u_m & v_m \\ & & & & 0 & v_m & u_m \end{pmatrix}_{2N\times 2N} \tag{23}$$

under OBCs for N unit cells.

Diagonalizing Eq. (23) yields the spectra of a finite chain. In Fig 3(a,c), we show the real parts of the OBC spectra for a chain for $N = 8$ (16 nanospheres in total) for longitudinal ($m = 0$, blue) and transverse ($m = 1$, orange) modes, respectively. On the x axes, we label each eigenvalue by an index j, and on the y axes, we show the real parts of the frequencies $\mathrm{Re}\, f$. Each of these 16 eigenvalues has a corresponding eigenvector, which we show and discuss in Sec. III F, which in turn corresponds to a collective mode of the full chain. Thus, j also labels the 16 modes. In both panels, $\beta = 0.9$ corresponds to the '+' markers, and $\beta = 1.1$ to the '×' markers. As we discuss in the next section, these choices for β correspond to the topologically trivial and non-trivial configurations, respectively. In the non-trivial configuration, we find two frequencies in the real band gap, which are absent in the trivial case. These in-gap frequencies are related to topological edge states, which we discuss in the following section. The remaining frequencies relate to bulk states, which are in the band widths of the PBC spectrum, indicated by the colored regions.

E. Topology, symmetry, and the bulk–boundary correspondence in the nearest-neighbor model

Let us discuss the notion of topology in nearest-neighbor SSH chains. In the Hermitian case [13], the existence of topologically protected edge states under OBCs is predicted by a bulk winding number, which depends on the eigenvectors from the system under PBCs, comprising a so-called bulk–boundary correspondence. The topological protection of such edge states stems from a chiral, or sublattice, symmetry of the system [60, 61]. These topological edge states come in pairs, where one state is exponentially localized at one end of the system, and the other one at the other end of the system.

Additionally, the Hermitian SSH chain has inversion, or parity, symmetry [61]. This on the one hand, allows for the simplification of expressions to determine the bulk winding number, while on the other hand, it allows for the definition of the so-called Zak phase [62], which we discuss in the context of longer-range coupling in Sec. IV C. Importantly, for the nearest-neighbor Hermitian SSH model, the bulk winding number and the Zak phase are equivalent topological quantities, which can be shown by direct computation.

As noted in Ref. 63, the discussed bulk–boundary correspondence as formulated for Hermitian systems generally breaks down in non-Hermitian systems [9, 10, 64], and the question on how to restore it was one of the driving forces behind the field of non-Hermitian physics in the last decade. The central insight was that due to the occurrence of the non-Hermitian skin effect and exceptional points [9–11]—both truly non-Hermitian phenomena—the conventional approach needs to be adapted [65]. By now, several methods to restore the bulk–boundary correspondence have been established [65–68].

Let us now come to the symmetries of our plasmonic SSH model. As conventional, we frame these symmetries in the language of the periodic model encoded in $C_m(q)$ (cf. Eq. (21)), however, these symmetries also apply to the finite system C_m (cf. Eq. (23)). In general, the constant diagonal in $C_m(q)$ corresponds to a shift in frequency and does not alter the eigenvectors of $C_m(q)$, which encode the system's topology. Thus, we set $\bar{C}(q) = C_m(q) - \tilde{\omega}\mathbb{1}_2$, where $\mathbb{1}_2$ is the identity matrix of size 2×2. This traceless matrix $\bar{C}(q)$, and by extension our plasmonic SSH chain, exhibits sublattice symmetry (SLS) and parity symmetry (P), as

$$\text{SLS}: \quad -\bar{C}(q) = \mathcal{S}\bar{C}(q)\mathcal{S}^{-1}, \qquad \mathcal{S} = \sigma_z, \tag{24a}$$

$$\text{P}: \quad \bar{C}(-q) = \mathcal{P}\bar{C}(q)\mathcal{P}^{-1}, \qquad \mathcal{P} = \sigma_x, \tag{24b}$$

where σ_x, σ_y, and σ_z, are the Pauli matrices. In Hermitian systems, sublattice symmetry and chiral symmetry are equivalent; parity symmetry and inversion symmetry as well. In non-Hermitian systems, both pairs of symmetries are inequivalent [69].

Due to the combination of sublattice and parity symmetry, the insights from the Hermitian case discussed in the beginning of this section carry over with only minor adjustments [70]. More specifically, having both these symmetries, it has been shown in Ref. 71 that $C_m(q)$ is topologically trivial if $|v_m| < |w_m|$, topologically non-trivial if $|v_m| > |w_m|$, and the topological phase transition occurs at $|v_m| = |w_m|$. For our plasmonic SSH chain, this directly translates to the geometry: As $|c_m(R)|$ is monotonously decaying with R for all relevant separations, cf. Fig. 2, $\beta < 1$ describes a topologically trivial chain, $\beta > 1$ a topologically non-trivial chain, and $\beta = 1$ is the topological phase transition. Therefore, we expect the occurrence of two in-gap frequencies, and associated with them, two topological edge states, for a finite chain in the topological configuration.

Before considering all eigenvectors, including these topological edge states, in the next section, we mention that the symmetries, Eq. (24), restrict the frequency-bands $\omega_{m,\pm}(q)$ of $C_m(q)$ to satisfy

$$\text{SLS}: \quad \omega_{m,\pm}(q) - \tilde{\omega} = -(\omega_{m,\mp}(q) - \tilde{\omega}), \tag{25a}$$

$$\text{P}: \quad \omega_{m,\pm}(\pi + q) = \omega_{m,\pm}(\pi - q), \tag{25b}$$

that is, the bands are symmetric with respect to $\omega = \tilde{\omega}$ and $q = \pi$, respectively. For the discussed nearest-neighbor couplings, this is apparent from their explicit form in Eq. (22) and in Fig. 3(b).

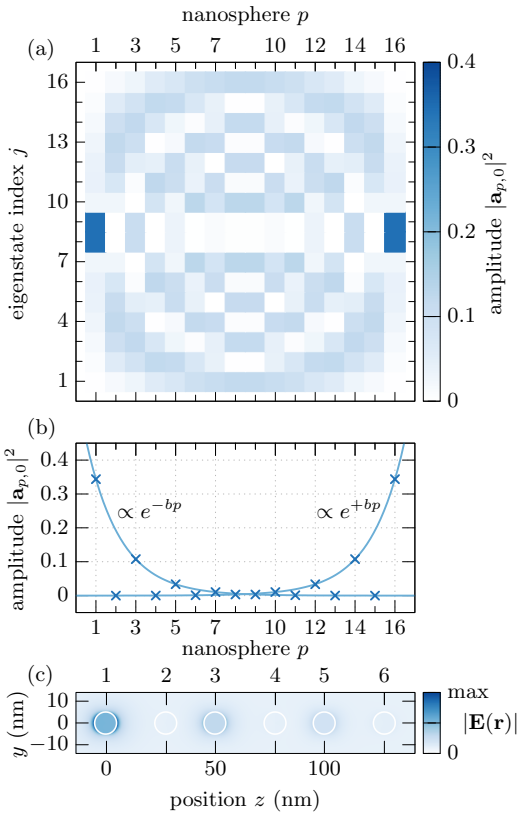


Figure 4. Longitudinal eigenvectors ($m = 0$) under OBCs in the topological configuration ($\beta = 1.1$), corresponding to Fig. 3(a). (a) Color map of all eigenvector amplitudes. The edge states ($j = 8, 9$) are clearly localized at the edges and distinguished from the delocalized bulk states ($j \neq 8, 9$). (b) Plot showing the amplitudes of one of the edge states ($j = 8$). The edge state follows the expected exponential behavior $\propto e^{-bp}$ ($\propto e^{+bp}$) for some real b on the odd (even) sublattice; the lines are included for better visibility. (c) Color map of the left end of the reconstructed QNM of the edge state in (b) with $x = 0$. The white circles indicate the metal-dielectric interfaces, which have the same electric field magnitude as the inside of the corresponding resonator. For the presented longitudinal edge mode, one can clearly see the alignment of the hot spots along the z-direction and increased field-amplitudes on the odd nanospheres.

F. Eigenvectors and topological edge states in the nearest-neighbor model

Now, with the notion of topology in the plasmonic SSH chain, we discuss the eigenvectors of the system under OBCs and show the topological edge states. The eigenvectors for $m = 0$ and $\beta = 1.1$, corresponding to Fig. 3(a) in the topological configuration, are shown in Fig. 4; the eigenvectors for $m = 1$ feature qualitative similar behavior (not shown). Figure 4(a) shows a color map of all eigenvector amplitudes of the system. The states corresponding to frequencies in the gap in Fig. 3(a) ($j = 8, 9$) are clearly localized towards the ends of the system. All other states ($j \neq 8, 9$), stemming from frequencies in the band widths, are delocalized. Thus, they do not feature the non-Hermitian skin effect as a result of the parity symmetry described by Eq. (24b) [11, 69]. Focusing on the topological edge states, we show their amplitudes in Fig. 4(b) in more detail. Due to the finite size of the chain, the two edge states are hybridized, and the amplitudes show the expected exponential decay away from the ends of the chain on the respective sublattices [13]. Finally, we show the amplitude of the electric field of the reconstructed QNM, $|\mathbf{E}(x = 0, y, z)|$, using Eq. (7). We depict the left end of the chain with one of the edge states in Fig. 4(c); Fig. 1(d) sketches the transverse ($m = 1$) case. This makes the distinction into longitudinal and transverse modes apparent: The hot spots outside the nanospheres are either aligned along the z-axis or are orthogonal to the z-axis, respectively. Furthermore, one can clearly see that for the depicted left end of the chain, only the QNMs of the nanospheres on odd sites are excited, which is a clear signature of the sublattice symmetry.

IV. THE PLASMONIC SSH CHAIN WITH LONG-RANGE COUPLING

Going beyond the nearest-neighbor SSH model discussed in the previous section, let us now introduce long-range couplings, which are necessary to more accurately capture the physics at play. The influence of long-range coupling has been, for example, discussed in the context of Hermitian SSH chains [72], in plasmonic chains using coupled dipole equations [17, 18], or in plasmonic chains in an open quantum systems description [22].

To set the notation, we introduce the long-range couplings generalizing the expressions in Eq. (20) as

$$v_m^{(n)} = c_m(na + d), \tag{26a}$$
$$w_m^{(n)} = c_m((n+1)a - d), \tag{26b}$$
$$u_m^{(n)} = c_m(na), \qquad \text{with} \qquad u_m^{(0)} = \tilde{\omega}, \tag{26c}$$

where $n \geq 0$; for $n = 0$, we have $v_m^{(0)} = v_m$, $w_m^{(0)} = w_m$ and $u_m^{(0)} = u_m$ consistent with Eq. (20). We remind ourselves that $m = -1, 0, +1$, and due to $c_{-1}(R) = c_{+1}(R)$ (cf. Sec. III A), it is enough to discuss the cases $m = 0$ (longitudinal) and $m = 1$ (transverse) as before.

A. Spectra with long-range coupling

Using Eq. (26), one can construct the coupling matrix under PBCs

$$C_m(q) = d_{m,0}(q)\,\mathbb{1}_2 + d_{m,x}(q)\,\sigma_x + d_{m,y}(q)\,\sigma_y \tag{27}$$

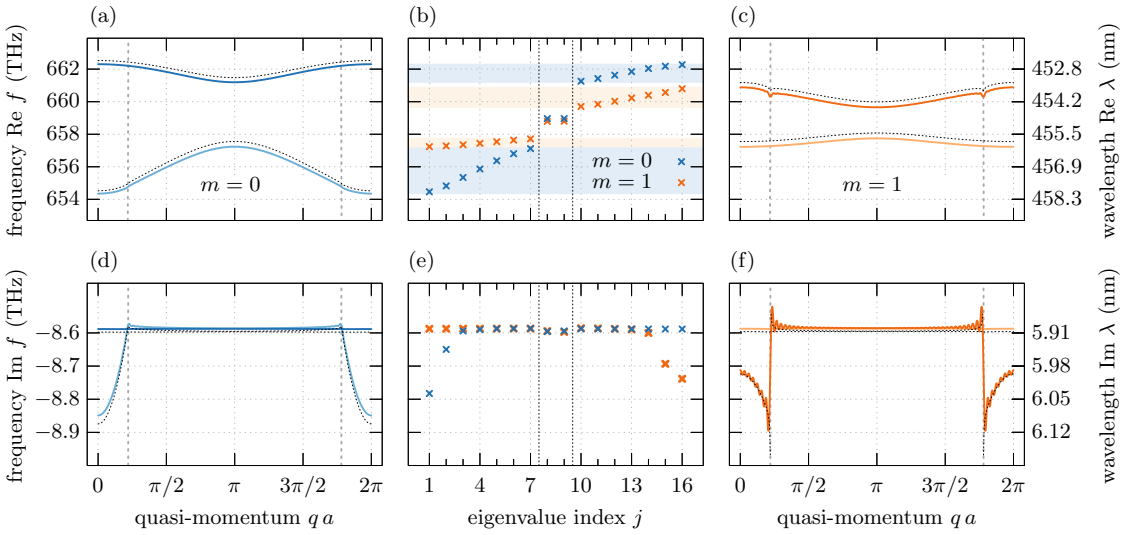


Figure 5. Spectra with long-range coupling. (a,c) Real parts of the spectrum under PBCs for (a) longitudinal (blue) and (c) transverse (orange) modes, respectively. The darker curves correspond to the upper bands and the lighter curves to the lower bands. The black dashed curves correspond to the results of Ref. 17, and the light lines are gray dashed. (b) In between the band spectra, the real part of the spectrum under OBCs for a chain of 16 nanospheres in the topological configuration, indicating the band widths under PBCs as colored regions. The vertical black dashed lines separate the bulk states corresponding to the lower band, the edge states, and the bulk states stemming from the upper band. (d-f) Imaginary parts of the spectra corresponding to the panels above. In (d,f), the vertical gray dashed lines correspond to the quasi-momenta where the real parts of the bands cross the light line in (a,c), respectively. For all spectra $\beta = 1.2$, corresponding to the topological configuration.

with frequency bands

$$\omega_{m,\pm}(q) = d_{m,0}(q) \pm \sqrt{d_{m,x}^2(q) + d_{m,y}^2(q)}. \tag{28}$$

In Eq. (27), we have introduced

$$d_{m,0}(q) = u_m^{(0)} + 2\sum_n u_m^{(n)} \cos(naq), \tag{29a}$$

$$d_{m,x}(q) = v_m^{(0)} + w_m^{(0)} \cos(aq) + \sum_n v_m^{(n)} \cos(naq) + w_m^{(n)} \cos((n+1)aq), \tag{29b}$$

$$d_{m,y}(q) = w_m^{(0)} \sin(aq) + \sum_n w_m^{(n)} \sin((n+1)aq) - v_m^{(n)} \sin(naq), \tag{29c}$$

where the terms before the sums correspond to the nearest-neighbor expression, whereas the sums encode the longer-range coupling. Importantly, the sums start at $n = 1$, and we introduce a cutoff $n_{\max}$ encoding the maximum coupling range, which restricts the sum over n in Eq. (15).

If one is interested in certain properties of the periodic system, it might be useful to consider the limit $n_{\max} \to \infty$ and rewrite the expressions in Eq. (29) as polylogarithms as in Refs. 73 and 17. However, within our framework, this creates two practical problems: (i) we do not have analytical expressions for the overlap integrals, and (ii) even if we had them, these overlap integrals are expected to be divergent as the associated QNMs are divergent, which would require a proper regularization technique. For now, we choose $n_{\max} = 63$ in order to qualitatively compare with the results in Ref. 17; we further comment on the choice of $n_{\max}$ in Sec. IV C.

Figure 5 shows the real and imaginary parts of the spectra with long-range coupling under PBCs and OBCs (with $N = 8$) for $\beta = 1.2$. In (a,c), the spectra for the longitudinal and transverse modes under PBCs are shown in blue and orange, respectively. Compared to the nearest-neighbor results, the symmetry between the upper and lower band, Eq. (25a), is broken, as sublattice symmetry, Eq. (24a), is broken due to the long-range coupling. Furthermore, we see that for longitudinal modes, the lower band ($\omega_{0,-}$) is more dispersive, i.e., has higher band width than the upper band ($\omega_{0,+}$), whereas this is the opposite for the transverse modes. Comparing to the black dashed curves, determined via the method described in Ref. 17, our bands are slightly shifted downwards. Again, this might be explained due to the approximation of resonance frequency of the individual nanosphere in Ref. 17, similar to the nearest-neighbor case. The (almost vertical) gray dashed lines correspond to the light lines, where $\omega = c_{\text{vac}} q/a$ and $\omega = c_{\text{vac}}(2\pi - q)/a$. In (c), we also see the dip at the light line as predicted in Ref. 17. In Fig. 5(b), we see the corresponding real parts of the spectra under OBCs with their respective band widths. Similar to the nearest-neighbor case, we sort all frequencies

by their real part. For $j = 1, \ldots, 7$, left of the left vertical black dashed line, and for $j = 10, \ldots, 16$, right of the right vertical black dashed line, the frequencies lie in the indicated band widths. We refer to these frequencies as bulk frequencies, and to their associated eigenvectors as bulk eigenvectors. Between the two vertical black dashed lines, $j = 8, 9$, we find two frequencies in the gap (for each $m = 0, 1$), which correspond to edge states as we will discuss below.

Coming to the imaginary parts of the spectra in Fig. 5(d-f), we start with the imaginary parts of the spectrum under PBCs of the longitudinal modes in (d). We also show imaginary parts determined by the method in Ref. 17 (but not shown therein) as black dashed curves, and we again find great qualitative agreement, where quantitatively, our approach delivers lower absolute values of the imaginary parts for similar reasons as discussed above. Compared to the real parts in (a), we find no gap between the bands in the imaginary part and again no symmetry between those bands. The imaginary part corresponding to the lower band in the real part $\omega_{0,-}$ (light blue curve), is almost constant between the two vertical gray dashed lines, whereas it is not constant outside this range. These vertical gray dashed lines correspond to the quasi-momenta where the real parts of the bands cross the light line in (a). The imaginary part of the upper band, $\omega_{0,+}$, appears to be flat throughout the plot with this y-axis scaling, but shows similar behavior when zooming in (cf. App. B 1). This overall behavior of the imaginary frequencies confirms our physical understanding of the light lines: Below the light line, corresponding to quasi-momenta between the two shown light lines in (a), the light is trapped inside the plasmonic structure and can only decay due to Ohmic losses. This inherent material loss is set by the damping rate γ of the Drude model, Eq. (17), resulting in an imaginary part of approximately $-\gamma/2 \approx -8.59\,\mathrm{THz}$ [74]. Above the light lines, however, light can additional radiate into free space, increasing losses and therefore increasing the amplitude of imaginary parts as apparent from the figure.

In Fig. 5(f), we show the imaginary part of the spectrum under PBCs of the longitudinal modes and see qualitatively similar behavior. One of the differences is that this time, $\omega_{1,-}$ is almost flat compared to $\omega_{1,+}$. While $\omega_{0,-}$ in (d) seems continuous, $\omega_{1,+}$ in (f) features a very sharp rise. Another noteworthy feature is that the spectrum features oscillations compared to the results in Ref. 17 as a result of the finite cutoff $n_{\max}$. In essence, we are expressing $\omega_{m,\pm}$ as a series of cosines (cf. Eqs. (28) and (29); the sine terms need to cancel due to parity symmetry, cf. Eqs. (24b) and (25b)). Hence, we are essentially using a truncated Fourier series, resulting in this oscillatory behavior. Especially close to the discontinuity at the light line, one can see the oscillations and the overshooting phenomenon, expected from a truncated Fourier series. Finally, (e) shows the corresponding imaginary parts of the spectra under OBCs. These imaginary parts are sorted according to their real parts in (b). As in (b), the frequencies in the bulk ($j = 1, \ldots, 7$ and $j = 10, \ldots, 16$) follow the behavior of the respective bands in (d,f). The frequencies between two vertical black dashed lines ($j = 8, 9$), corresponding to the in-gap frequencies in (b), have a slightly lower imaginary part then the bulk frequencies directly below and above real gap, e.g., $j = 7, 10$, respectively.

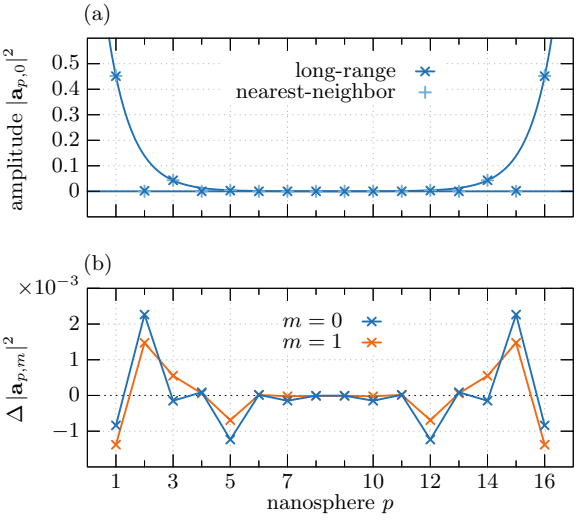


Figure 6. Influence of long-range coupling on the edge states. (a) Amplitudes of a longitudinal ($m = 0$) edge state including long-range coupling with $\beta = 1.2$ in dark blue, similar to Fig. 4(b). The light blue crosses mark the nearest-neighbor amplitudes, which are not distinguishable from the long-range amplitudes by eye. To make the influence of long-range coupling visible, we show the difference between the amplitudes with long-range and nearest-neighbor couplings in (b), for longitudinal ($m = 0$) and transverse ($m = 1$) edge states in blue and orange, respectively.

B. Eigenvectors and edge states with long-range couplings

Here we show that incorporating long-range coupling does not significantly alter the eigenvectors under OBCs. Figure 6 compares the amplitudes of the edge states as a function of nanosphere p; we discuss some bulk states in App. B 2. In (a), we show the amplitudes $|\mathbf{a}_{p,m=0}|^2$ of a longitudinal ($m = 0$) edge state including long-range coupling with $\beta = 1.2$ using dark blue '×' markers. The amplitudes with long-range coupling show the same qualitative behavior as the amplitudes with nearest-neighbor coupling only (light blue '+' markers), including an exponential decay away from the ends of the system on the respective sublattice, even though sublattice symmetry is broken; the transverse modes ($m = 1$) show the same qualitative behavior (not shown). As we are

deeper in topological phase with $\beta = 1.2$, compared to $\beta = 1.1$ in Fig. 4(b), the decay of the edge state is steeper. The transverse edge states feature qualitative similar behavior (not shown). In (b), we make the difference between long-range (superscript lr) and nearest-neighbor (superscript nn) coupling explicitly visible by showing $\Delta|\mathbf{a}_{p,m}|^2 = |\mathbf{a}^{\mathrm{lr}}_{p,m}|^2 - |\mathbf{a}^{\mathrm{nn}}_{p,m}|^2 \sim 10^{-3}$. The main changes in amplitude happen at the system ends.

C. Topology with long-range couplings

For nearest-neighbor coupling only, cf. Sec. III E, we discussed that the topological edge states stem from a bulk–boundary correspondence requiring sublattice symmetry, and parity symmetry simplifies the determination of the topological index. As already mentioned, longer-range couplings break sublattice symmetry, cf. Eq. (24a). However, the system is still parity symmetric with parity operator $\mathcal{P}$, cf. Eq. (24b), which allows us to define the non-Hermitian Zak phase

$$\phi_{m,\pm} = -\mathrm{i}\int_0^{2\pi/a} \langle \mathbf{a}^{\mathrm{L}}_{m,\pm}(q)|\partial_q|\mathbf{a}^{\mathrm{R}}_{m,\pm}(q)\rangle \,\mathrm{d}q, \tag{30}$$

$$= -\mathrm{i}\ln\Big[\langle \mathbf{a}^{\mathrm{L}}_{m,\pm}(\pi/a)|\mathcal{P}|\mathbf{a}^{\mathrm{R}}_{m,\pm}(\pi/a)\rangle \times \langle \mathbf{a}^{\mathrm{L}}_{m,\pm}(0)|\mathcal{P}|\mathbf{a}^{\mathrm{R}}_{m,\pm}(0)\rangle\Big], \tag{31}$$

which is quantized to 0 or π (mod 2π) for each band labeled by $\pm$ [62, 75]. In these expressions, $|\mathbf{a}^{\mathrm{R}}_{m,\pm}(q)\rangle$ and $\langle \mathbf{a}^{\mathrm{L}}_{m,\pm}(q)|$ are the right (superscript R) and left (superscript L) eigenvectors, respectively, satisfying

$$C_m(q)|\mathbf{a}^{\mathrm{R}}_{m,\pm}(q)\rangle = \omega_{m,\pm}(q)|\mathbf{a}^{\mathrm{R}}_{m,\pm}(q)\rangle, \tag{32a}$$

$$\langle \mathbf{a}^{\mathrm{L}}_{m,\pm}(q)|C_m(q) = \omega_{m,\pm}(q)\langle \mathbf{a}^{\mathrm{L}}_{m,\pm}(q)|, \tag{32b}$$

which are biorthonormalized, $\langle \mathbf{a}^{\mathrm{L}}_{m,\pm_1}(q)|\mathbf{a}^{\mathrm{R}}_{m,\pm_2}(q)\rangle = \delta^{\pm_1}_{\pm_2}$.

In the modern theory of polarization of electronic systems [76], a non-trivial Hermitian Zak phase has been related to a charge accumulation at opposite ends of a chain via a bulk–boundary correspondence [76, 77]. It also has been related to the emergence of in-gap states in the entanglement spectrum of electronic systems [78]. However, both of these bulk–boundary correspondences rely on the many-body properties of Hermitian electronic systems, and are thus not applicable to the plasmonic situation.

Due to the absence of sublattice symmetry one cannot expect the occurrence of topologically protected edge states in the system with long-range couplings. However, we observe that the non-Hermitian Zak phase predicts the occurrence of states exponentially localized at the edges of the system, cf. Fig. 6(a), as we discuss below. We refer to these states as edge states, and not topological edge states.

As a simple picture to understand how such edge states emerge, one can consider a fictitious sublattice-symmetric long-range SSH chain by setting all sublattice-symmetry-breaking terms to zero, i.e., setting all coefficients in Eq. (29a) to zero. This fictitious chain exhibits topological edge states predicted by the bulk winding number, which is again equivalent to the Zak phase, cf. Sec. III E. These topological edge states are robust against perturbations, and one can think of the sublattice-symmetry-breaking terms as perturbations. For the parameters we chose in this work, these perturbations are small in the sense that the edge states keep the relevant localization properties, cf. Fig. 6(a); choosing other parameters may destroy those properties [17–19, 72]. Thus, one can think of the observed edge states as the remnants of the topological edge states of the fictitious sublattice-symmetric chain.

This also allows us to discuss a subtlety overlooked in the literature. The determination of the Zak phase depends on which coupling matrix $C_m(q)$ in Eq. (32) one uses to determine the eigenvectors, as $C_m(q)$ implicitly depends on the choice of $n_{\max}$, cf. Eqs. (27) and (29). For two reasons, we fix the coupling range cutoff $n_{\max}$, without claiming a rigorous bulk–boundary correspondence. For a finite chain consisting of N unit cells, we require all-to-all coupling, which corresponds to setting $N = n_{\max} + 1$. Firstly, this directly connects to the fictitious sublattice-symmetric chain discussed above, allowing us to use its conventional non-Hermitian bulk–boundary correspondence, which requires a fixed $n_{\max}$ and considers the limit $N \to \infty$ [68]. Importantly, the Zak phase does not depend on the sublattice-symmetry-breaking terms [17], and thus is identical for the chain with and without sublattice symmetry for any cutoff $n_{\max}$. Secondly, introducing this cutoff makes our frequency bands $\omega_{m,\pm}(q)$ single-valued smooth functions (cf. Eq. (28)), giving rise to a well-defined Zak phase. This contrasts previous works, where the limit $N, n_{\max} \to \infty$ results in the discussed polariton-like splitting at the light lines [17, 73], rendering the Zak phase ill-defined.

Thus, when considering a finite chain with $N = 8$ unit cells under OBCs with all-to-all coupling as above, we determine Zak phases using $n_{\max} = 7$, so that $N = n_{\max}+1$; the corresponding band spectra are shown in App. B 1. Using Eq. (31) yields the anticipated result

$$\phi_{m,\pm} = \begin{cases} 0 \bmod 2\pi, & \text{if } \beta < 1, \\ \pi \bmod 2\pi, & \text{if } \beta > 1, \end{cases} \tag{33}$$

for all examples discussed in this manuscript. In Sec. IV A we have used the larger cutoff $n_{\max} = 63$ in order to compare our spectra with Ref. 17.

V. THE PLASMONIC SSH CHAIN CONSISTING OF NANOCUBOIDS

Finally, we illustrate that the coupled QNM theory employed in this work is agnostic to the shape of the constituent nanoresonators, and that we also do not need

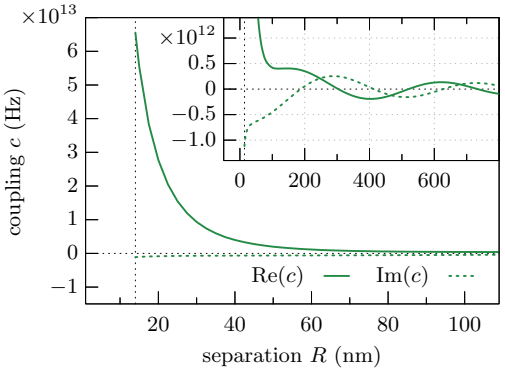


Figure 7. Coupling coefficients of metallic nanocuboids. Real (solid green) and imaginary (dashed green) parts of the coupling coefficients c of the rectangular cuboid as function of the center-to-center-separation R. The vertical black dashed line at $R = 14\,\mathrm{nm}$ corresponds to the limit where both nanocuboids touch.

analytical expressions for the QNMs of the constituents. Therefore, we discuss the low-frequency modes of a plasmonic SSH chain consisting of rectangular cuboids in glass ($\varepsilon_\mathrm{B} = 9/4$), using the same Drude model as before, cf. Eq. (17). In order to have a structure of a similar size as the nanospheres before (radius $r_0 = 5\,\mathrm{nm}$), we chose the side lengths $l_x = 25\,\mathrm{nm}$ and $l_y = l_z = 14\,\mathrm{nm}$ in the x-, y- and z-directions, respectively, with a rounding radius of 2 nm for all edges. For brevity, we refer to them as nanocuboids. Unlike for the case of spheres, no analytical solutions are known for such (rectangular) nanocuboids. Therefore, we have determined the resonance frequencies and associated field modes using COMSOL Multiphysics, where the asymmetry in the x-direction yields a non-degenerate resonance with lowest frequency. With this resonance frequency, and its associated field modes, one can determine the coupling coefficients solving the overlap integrals as discussed in Sec. III A. For details on the numerics, we refer the interested reader to Ref. 52.

Having determined the resonance frequencies and the relevant coupling coefficients allows a semi-analytic treatment of the chain of cuboids, paralleling the discussions in previous sections. This again shows the utility of using a coupled QNM theory: It breaks the determination of the QNMs of a plasmonic SSH chain, or in general any nanoresonator array, into small and interpretable steps. This also saves computational resources compared to a brute-force solution of the problem [52].

Let us discuss the results using the coupled QNM approach. The coupling coefficients are shown in Fig. 7 and show qualitatively similar behavior as the couplings of the spheres in Fig. 2. Due to the generally strong localization of the QNMs close to the nanoresonator surfaces, sketched in Fig. 1(b), the details of the surface geometry of the nanoresonators matters a lot for short separations. In our case, this results in larger coupling strengths for the nanocuboid compared to the nanospheres for short separations. Furthermore, one should keep in mind that the cuboid structure is embedded in glass, which essentially rescales the x-axis by the refractive index of glass, compared to the nanosphere in vacuum.

Figure 8 summarizes the results for the nanocuboid chain, which includes long-range couplings. In Fig. 8(a), we show the real part of the spectrum under PBCs for $\beta = 1.2$ with $n_\mathrm{max} = 11$ and unit cell size $a = 50\,\mathrm{nm}$. At the light lines, one can see slight dips as in the transverse case for the spheres, cf. Fig. 5(c). As argued in Sec. IV C, picking this cutoff n_max corresponds to an all-to-all coupling in a chain under OBCs with $N = n_\mathrm{max} + 1 = 12$ unit cells, i.e., 24 nanocuboids.

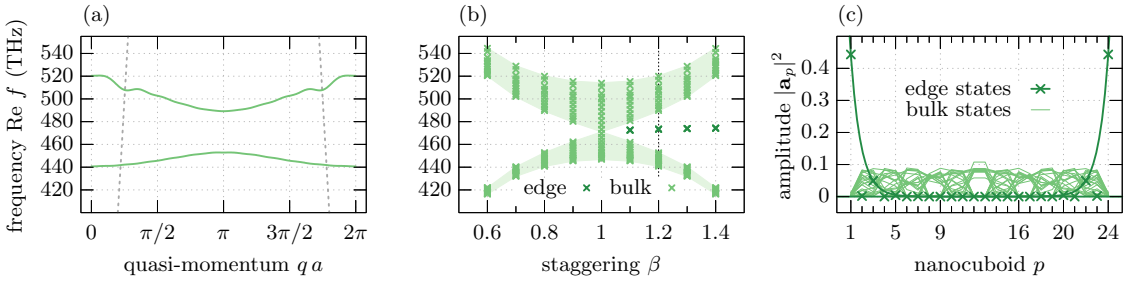


Figure 8. Results for the plasmonic SSH chain consisting of rectangular cuboids. (a) Spectrum under PBCs with $\beta = 1.2$ and long-range coupling. The light lines are again gray dashed. (b) Spectra under OBCs for different β tuning from topologically trivial ($\beta < 1$) to topologically non-trivial phase ($\beta > 1$) through the topological phase transition ($\beta = 1$) for 24 nanocuboids. Light green crosses in the spectra give rise to bulk states, whereas dark green crosses to edge states. The green regions correspond to the band widths under PBCs. The vertical dashed black line corresponds to $\beta = 1.2$, featured in (a,c). (c) Eigenvectors under OBCs for $\beta = 1.2$. Light green lines connect the amplitudes of bulk modes. Dark green crosses correspond to the amplitudes of the topological edge modes, the dark green lines are exponential fits similar to Fig. 4(b) showing the exponential localization on the corresponding sublattice.

Figure 8(b) shows the OBC spectra of such structures for different β. For each β on the x axis, we show the entire spectrum of the finite chain along y, tuning from the topologically trivial ($\beta < 1$) to topologically non-trivial phase ($\beta > 1$) via the topological phase transition at $\beta = 1$. One can clearly see the emergence of in-gap states in the topologically non-trivial phase.

Finally, Fig. 8(c) shows the eigenvectors under OBCs corresponding to $\beta = 1.2$. Like for the spheres, the in-gap frequencies correspond to edge states exponentially localized towards the system ends, whereas the bulk states are delocalized, cf. Fig. 4(a,b). Over all, the intuition we gained by analyzing the analytic example of the nanospheres transfers fully to the cuboid case, and exemplifies the power of our approach.

VI. CONCLUSION

In this work, we have analyzed plasmonic analogs of the SSH model consisting of either metallic nanospheres or rectangular nanocuboids using a coupled QNM approach. We modeled the metallic character of the resonators by assuming a Drude model, however, the approach can readily use N-pole Lorentz-Drude models [51, 55, 56] and can also be applied to dielectric nanoresonators. On top of that, the requirement of having an isotropic and non-magnetic material can be relaxed [48], and the QNM framework has recently been extended to time-varying problems [79, 80]. Furthermore, as illustrated by the semi-analytic treatment of the SSH chain of rectangular cuboids, other shapes for the constituent nanoresonators, such as core-shell nanoparticles [23], can be used. In general, the coupled QNM approach is also applicable to nanoparticle chains on a substrate [42, 81]. Also, the topological framework trivially extends to two- and three-dimensional nanoresonator arrays [30–37].

Beyond being a natural platform to investigate non-Hermitian topological phenomena [9–11, 35, 36], the coupled QNM theory promises to aid designing plasmonic metasurfaces for practical applications [43, 82–84]. For example, the exceptional-point-based metasurface in Ref. 82 for sensing is modeled using a coupled-mode picture [85], where the coupling coefficients are extracted from full-wave numerics. Here, our coupled QNM approach should not only be able to predict these coupling coefficients from first principles, but should also allow us to create a parametric model, which can be optimized for the desired functionality. Including more modes and being able to tune the geometry of the constituent nanoresonators [45, 54, 86] further increases the design space. Going beyond the proof of principle in our manuscript, the inclusion of more modes is essential to get high-precision results as generally understood [40, 41, 44] and further discussed in a complementary publication [52].

Over all, the presented methods allow for the optimization and design of topologically robust plasmonic nanostructures with potential applications in sensing, microscopy, and optical communication.

ACKNOWLEDGMENTS

J.T.G. and F.K.K. acknowledge funding from the Max Planck Society Lise Meitner Excellence Program 2.0. N.S.I. and L.R. acknowledge financial support from the National Research Council of Canada, and would like to thank the Digital Research Alliance of Canada (formerly Compute Canada) for providing computational resources and technical support.

AUTHOR CONTRIBUTIONS

J.T.G. carried out the theoretical work, supported by all co-authors through discussions. N.S.I. provided the numerical data for the rectangular nanocuboids, and seeded the idea to build the theoretical framework on a coupled QNM theory. L.R. and F.K.K. initiated and supervised the project. J.T.G. prepared the manuscript with input from all co-authors.

[1] F. D. M. Haldane and S. Raghu, Possible Realization of Directional Optical Waveguides in Photonic Crystals with Broken Time-Reversal Symmetry, Physical Review Letters **100**, 013904 (2008).
[2] S. Raghu and F. D. M. Haldane, Analogs of quantum-Hall-effect edge states in photonic crystals, Physical Review A **78**, 033834 (2008).
[3] L. Lu, J. D. Joannopoulos, and M. Soljačić, Topological photonics, Nature Photonics **8**, 821 (2014).
[4] T. Ozawa, H. M. Price, A. Amo, N. Goldman, M. Hafezi, L. Lu, M. C. Rechtsman, D. Schuster, J. Simon, O. Zilberberg, and I. Carusotto, Topological photonics, Reviews of Modern Physics **91**, 015006 (2019).
[5] S. A. Maier, *Plasmonics: Fundamentals and Applications* (Springer US, New York, NY, 2007).
[6] V. Giannini, A. I. Fernández-Domínguez, S. C. Heck, and S. A. Maier, Plasmonic Nanoantennas: Fundamentals and Their Use in Controlling the Radiative Properties of Nanoemitters, Chemical Reviews **111**, 3888 (2011).
[7] M. I. Stockman, K. Kneipp, S. I. Bozhevolnyi, S. Saha, A. Dutta, J. Ndukaife, N. Kinsey, H. Reddy, U. Guler, V. M. Shalaev, A. Boltasseva, B. Gholipour, H. N. S. Krishnamoorthy, K. F. MacDonald, C. Soci, N. I. Zheludev, V. Savinov, R. Singh, P. Groß, C. Lienau, M. Vadai, M. L. Solomon, D. R. Barton, M. Lawrence, J. A. Dionne, S. V. Boriskina, R. Esteban, J. Aizpurua, X. Zhang, S. Yang, D. Wang, W. Wang, T. W. Odom, N. Accanto,

P. M. de Roque, I. M. Hancu, L. Piatkowski, N. F. van Hulst, and M. F. Kling, Roadmap on plasmonics, Journal of Optics **20**, 043001 (2018).
[8] M. P. Mcoyi, K. T. Mpofu, M. Sekhwama, and P. Mthunzi-Kufa, Developments in Localized Surface Plasmon Resonance, Plasmonics **20**, 5481 (2025).
[9] Y. Ashida, Z. Gong, and M. Ueda, Non-Hermitian physics, Advances in Physics **69**, 249 (2020).
[10] E. J. Bergholtz, J. C. Budich, and F. K. Kunst, Exceptional topology of non-Hermitian systems, Reviews of Modern Physics **93**, 015005 (2021).
[11] J. T. Gohsrich, A. Banerjee, and F. K. Kunst, The non-Hermitian skin effect: A perspective, Europhysics Letters **150**, 60001 (2025).
[12] W. P. Su, J. R. Schrieffer, and A. J. Heeger, Solitons in Polyacetylene, Physical Review Letters **42**, 1698 (1979).
[13] J. K. Asbóth, L. Oroszlány, and A. Pályi, *A Short Course on Topological Insulators*, Lecture Notes in Physics, Vol. 919 (Springer International Publishing, Cham, 2016).
[14] M. S. Rider, Á. Buendía, D. R. Abujetas, P. A. Huidobro, J. A. Sánchez-Gil, and V. Giannini, Advances and Prospects in Topological Nanoparticle Photonics, ACS Photonics **9**, 1483 (2022).
[15] A. Poddubny, A. Miroshnichenko, A. Slobozhanyuk, and Y. Kivshar, Topological Majorana States in Zigzag Chains of Plasmonic Nanoparticles, ACS Photonics **1**, 101 (2014).
[16] C. W. Ling, M. Xiao, C. T. Chan, S. F. Yu, and K. H. Fung, Topological edge plasmon modes between diatomic chains of plasmonic nanoparticles, Optics Express **23**, 2021 (2015).
[17] S. R. Pocock, X. Xiao, P. A. Huidobro, and V. Giannini, Topological Plasmonic Chain with Retardation and Radiative Effects, ACS Photonics **5**, 2271 (2018).
[18] S. R. Pocock, P. A. Huidobro, and V. Giannini, Bulk-edge correspondence and long-range hopping in the topological plasmonic chain, Nanophotonics **8**, 1337 (2019).
[19] S. R. Pocock, *Topological Physics in One-Dimensional Chains of Metallic Nanoparticles*, Ph.D. thesis, Imperial College London (2020).
[20] R. P. H. Wu and H. C. Ong, Small mode volume topological photonic states in one-dimensional lattices with dipole-quadrupole interactions, Physical Review Research **4**, 023233 (2022).
[21] C. A. Downing and G. Weick, Topological collective plasmons in bipartite chains of metallic nanoparticles, Physical Review B **95**, 125426 (2017).
[22] C. A. Downing and G. Weick, Topological plasmons in dimerized chains of nanoparticles: Robustness against long-range quasistatic interactions and retardation effects, The European Physical Journal B **91**, 253 (2018).
[23] Á. Buendía and N. M. R. Peres, Coupled electric dipole model for a Su–Schrieffer–Heeger chain of optically resonant coreshell nanoparticles, Journal of Applied Physics **139**, 063110 (2026).
[24] A. Ott and S.-A. Biehs, Radiative heat flux through a topological Su-Schrieffer-Heeger chain of plasmonic nanoparticles, Physical Review B **102**, 115417 (2020).
[25] Z. Liu, P. Jin, M. Lei, C. Wang, F. Marchesoni, J.-H. Jiang, and J. Huang, Topological thermal transport, Nature Reviews Physics **6**, 554 (2024).
[26] I. S. Sinev, I. S. Mukhin, A. P. Slobozhanyuk, A. N. Poddubny, A. E. Miroshnichenko, A. K. Samusev, and Y. S. Kivshar, Mapping plasmonic topological states at the nanoscale, Nanoscale **7**, 11904 (2015).
[27] Q. Yan, E. Cao, Q. Sun, Y. Ao, X. Hu, X. Shi, Q. Gong, and H. Misawa, Near-Field Imaging and Time-Domain Dynamics of Photonic Topological Edge States in Plasmonic Nanochains, Nano Letters **21**, 9270 (2021).
[28] Y. Moritake, M. Ono, and M. Notomi, Far-field optical imaging of topological edge states in zigzag plasmonic chains, Nanophotonics **11**, 2183 (2022).
[29] B. Schurr, M. Hensen, L. Brenneis, P. Kessler, J. Qin, V. Lisinetskii, R. Thomale, T. Brixner, and B. Hecht, Plasmonic Su–Schrieffer–Heeger chains with strong coupling amplitudes, Science Advances **11**, eaea3844 (2025).
[30] G. Weick, C. Woollacott, W. L. Barnes, O. Hess, and E. Mariani, Dirac-like Plasmons in Honeycomb Lattices of Metallic Nanoparticles, Physical Review Letters **110**, 106801 (2013).
[31] M. Honari-Latifpour and L. Yousefi, Topological plasmonic edge states in a planar array of metallic nanoparticles, Nanophotonics **8**, 799 (2019).
[32] M. Kim and J. Rho, Topological edge and corner states in a two-dimensional photonic Su-Schrieffer-Heeger lattice, Nanophotonics **9**, 3227 (2020).
[33] M. Proctor, M. Blanco de Paz, D. Bercioux, A. García-Etxarri, and P. Arroyo Huidobro, Higher-order topology in plasmonic Kagome lattices, Applied Physics Letters **118**, 091105 (2021).
[34] L. Zhang, X.-M. Wang, X.-M. Qiu, Z. Wang, and J.-Y. Yan, Light controlled topological plasmonics in a graphene lattice arrayed by metal nanoparticles, Physical Review B **108**, 085402 (2023).
[35] J. Cuerda, J. M. Taskinen, N. Källman, L. Grabitz, and P. Törmä, Pseudospin-orbit coupling and non-Hermitian effects in the quantum geometric tensor of a plasmonic lattice, Physical Review B **109**, 165439 (2024).
[36] J. Cuerda, J. M. Taskinen, N. Källman, L. Grabitz, and P. Törmä, Observation of quantum metric and non-Hermitian Berry curvature in a plasmonic lattice, Physical Review Research **6**, L022020 (2024).
[37] Á. Buendía, J. L. Pura, V. Giannini, and J. A. S. Gil, Far-field radiation of bulk, edge and corner eigenmodes from a finite 2D Su-Schrieffer-Heeger plasmonic lattice (2025), arXiv:2510.08063.
[38] Q. Yan, B. Zhao, Q. Lyu, Y. Li, S. Chu, C. Lu, X. Hu, C. T. Chan, and Q. Gong, Near-field imaging of synthetic dimensional integrated plasmonic topological Harper nanochains, Nature Communications **16**, 2592 (2025).
[39] S. Y. Park and D. Stroud, Surface-plasmon dispersion relations in chains of metallic nanoparticles: An exact quasistatic calculation, Physical Review B **69**, 125418 (2004).
[40] P. Lalanne, W. Yan, K. Vynck, C. Sauvan, and J.-P. Hugonin, Light Interaction with Photonic and Plasmonic Resonances, Laser & Photonics Reviews **12**, 1700113 (2018).
[41] P. T. Kristensen, K. Herrmann, F. Intravaia, and K. Busch, Modeling electromagnetic resonators using quasinormal modes, Advances in Optics and Photonics **12**, 612 (2020).
[42] C. Sauvan, T. Wu, R. Zarouf, E. A. Muljarov, and P. Lalanne, Normalization, orthogonality, and completeness of quasinormal modes of open systems: The case of

electromagnetism, Optics Express **30**, 6846 (2022).
[43] T. Wu and P. Lalanne, Designing electromagnetic resonators with quasinormal modes, Frontiers in Physics **12**, 10.3389/fphy.2024.1461106 (2024).
[44] E. A. Muljarov, W. Langbein, and R. Zimmermann, Brillouin-Wigner perturbation theory in open electromagnetic systems, Europhysics Letters **92**, 50010 (2011).
[45] M. B. Doost, W. Langbein, and E. A. Muljarov, Resonant-state expansion applied to three-dimensional open optical systems, Physical Review A **90**, 013834 (2014).
[46] E. A. Muljarov and W. Langbein, Exact mode volume and Purcell factor of open optical systems, Physical Review B **94**, 235438 (2016).
[47] E. A. Muljarov and W. Langbein, Resonant-state expansion of dispersive open optical systems: Creating gold from sand, Physical Review B **93**, 075417 (2016).
[48] E. A. Muljarov and T. Weiss, Resonant-state expansion for open optical systems: Generalization to magnetic, chiral, and bi-anisotropic materials, Optics Letters **43**, 1978 (2018).
[49] C. Tao, J. Zhu, Y. Zhong, and H. Liu, Coupling theory of quasinormal modes for lossy and dispersive plasmonic nanoresonators, Physical Review B **102**, 045430 (2020).
[50] E. A. Muljarov, Rigorous Theory of Coupled Resonators, Physical Review Letters **136**, 023801 (2026).
[51] T. Wu and P. Lalanne, Dissipative coupling in photonic and plasmonic resonators, Advanced Photonics **7**, 056011 (2025).
[52] N. S. Israel, J. T. Gohsrich, F. K. Kunst, and L. Ramunno, A Modal Analysis of Topological Nanoresonator Chains (2026), (unpublished).
[53] A. Aiello, J. G. E. Harris, and F. Marquardt, Perturbation theory of optical resonances of deformed dielectric spheres, Physical Review A **100**, 023837 (2019).
[54] J. Gohsrich, T. Shah, and A. Aiello, Perturbation theory of nearly spherical dielectric optical resonators, Physical Review A **104**, 023516 (2021).
[55] A. Raman and S. Fan, Photonic Band Structure of Dispersive Metamaterials Formulated as a Hermitian Eigenvalue Problem, Physical Review Letters **104**, 087401 (2010).
[56] W. Yan, R. Faggiani, and P. Lalanne, Rigorous modal analysis of plasmonic nanoresonators, Physical Review B **97**, 205422 (2018).
[57] Q. Tao, Y. Su, C. Tao, Y. Zhong, and H. Liu, Efficient method for modeling large-scale arrays of optical nanoresonators based on the coupling theory of quasinormal mode, Optics Express **32**, 7171 (2024).
[58] The non-Hermitian nature of the matrix C from Eq. (8) necessitates a biorthogonal treatment, which is the straightforward generalization the Hermitian approach.
[59] A. Vial, A.-S. Grimault, D. Macías, D. Barchiesi, and M. L. de la Chapelle, Improved analytical fit of gold dispersion: Application to the modeling of extinction spectra with a finite-difference time-domain method, Physical Review B **71**, 085416 (2005).
[60] S. Ryu and Y. Hatsugai, Topological Origin of Zero-Energy Edge States in Particle-Hole Symmetric Systems, Physical Review Letters **89**, 077002 (2002).
[61] P. Delplace, D. Ullmo, and G. Montambaux, Zak phase and the existence of edge states in graphene, Physical Review B **84**, 195452 (2011).
[62] J. Zak, Berry's phase for energy bands in solids, Physical Review Letters **62**, 2747 (1989).
[63] T. E. Lee, Anomalous Edge State in a Non-Hermitian Lattice, Physical Review Letters **116**, 133903 (2016).
[64] Y. Xiong, Why does bulk boundary correspondence fail in some non-hermitian topological models, Journal of Physics Communications **2**, 035043 (2018).
[65] S. Yao and Z. Wang, Edge States and Topological Invariants of Non-Hermitian Systems, Physical Review Letters **121**, 086803 (2018).
[66] F. K. Kunst, E. Edvardsson, J. C. Budich, and E. J. Bergholtz, Biorthogonal Bulk-Boundary Correspondence in Non-Hermitian Systems, Physical Review Letters **121**, 026808 (2018).
[67] F. K. Kunst and V. Dwivedi, Non-Hermitian systems and topology: A transfer-matrix perspective, Physical Review B **99**, 245116 (2019).
[68] L. Koekenbier and H. Schulz-Baldes, Transfer matrix analysis of non-Hermitian Hamiltonians: Asymptotic spectra and topological eigenvalues, Journal of Spectral Theory **14**, 1563 (2024).
[69] K. Kawabata, K. Shiozaki, M. Ueda, and M. Sato, Symmetry and Topology in Non-Hermitian Physics, Physical Review X **9**, 041015 (2019).
[70] Due to parity symmetry, the generalized Brillouin zone (GBZ) defined in Ref. 65 reduces to the standard Brillouin zone [69]. Therefore, the only adjustment is the introduction of a biorthonormal basis when determining the bulk winding number, as we illustrate for the Zak phase in Sec. IV C.
[71] S. Lieu, Topological phases in the non-Hermitian Su-Schrieffer-Heeger model, Physical Review B **97**, 045106 (2018).
[72] B. Pérez-González, M. Bello, Á. Gómez-León, and G. Platero, Interplay between long-range hopping and disorder in topological systems, Physical Review B **99**, 035146 (2019).
[73] A. F. Koenderink and A. Polman, Complex response and polariton-like dispersion splitting in periodic metal nanoparticle chains, Physical Review B **74**, 033402 (2006).
[74] This value is also the prediction from the quasistatic approach in Ref. 16, but throughout all quasi-momenta and not only below the light lines. We also find a minor difference between this value and the actual value of the imaginary part as shown in App. B 1.
[75] S. Tsubota, H. Yang, Y. Akagi, and H. Katsura, Symmetry-protected quantization of complex Berry phases in non-Hermitian many-body systems, Physical Review B **105**, L201113 (2022).
[76] D. Vanderbilt, *Berry Phases in Electronic Structure Theory: Electric Polarization, Orbital Magnetization and Topological Insulators* (Cambridge University Press, Cambridge, 2018).
[77] J.-W. Rhim, J. Behrends, and J. H. Bardarson, Bulk-boundary correspondence from the intercellular Zak phase, Physical Review B **95**, 035421 (2017).
[78] T. L. Hughes, E. Prodan, and B. A. Bernevig, Inversion-symmetric topological insulators, Physical Review B **83**, 245132 (2011).
[79] A. C. Valero, S. Gladyshev, D. Globosits, S. Rotter, E. A. Muljarov, and T. Weiss, Revealing the Resonant Physics of Open Photonic Time Crystals, Laser & Photonics Re-

views , e71318 (2026).
[80] Y. Sun, S. Fan, and G. Hu, Quasinormal Mode Basis for Open Floquet Photonic Systems, Physical Review Letters **136**, 206903 (2026).
[81] F. Herz, A. Naeimi, and S.-A. Biehs, Substrate-driven topological engineering in plasmonic Su-Schrieffer-Heeger chains (2026), arXiv:2605.28463.
[82] J.-H. Park, A. Ndao, W. Cai, L. Hsu, A. Kodigala, T. Lepetit, Y.-H. Lo, and B. Kanté, Symmetry-breaking-induced plasmonic exceptional points and nanoscale sensing, Nature Physics **16**, 462 (2020).
[83] Q. Song, M. Odeh, J. Zúñiga-Pérez, B. Kanté, and P. Genevet, Plasmonic topological metasurface by encircling an exceptional point, Science **373**, 1133 (2021).
[84] R. Heilmann, J. Lehikoinen, S. Eyvazi, E. A. Mamonov, and P. Törmä, Chains of nanoparticles for flat-band emission and lasing (2026), arXiv:2602.21825.
[85] A. Kodigala, T. Lepetit, and B. Kanté, Exceptional points in three-dimensional plasmonic nanostructures, Physical Review B **94**, 201103 (2016).
[86] W. Yan, P. Lalanne, and M. Qiu, Shape Deformation of Nanoresonator: A Quasinormal-Mode Perturbation Theory, Physical Review Letters **125**, 013901 (2020).

Appendix A: Remaining expressions for the QNMs of spherical nanoresonators

Beyond the structure of the QNMs of spherical nanoresonators discussed in Sec. II B, we report the remaining expressions needed to reproduce all results in the main text. These expressions are derived in Refs. 45–47, and we include them here for completeness. For a more compact notation, we assume $\varepsilon_\mathrm{B} = 1$; the relative permittivity of the nanosphere of radius r_0 in spherical coordinates (r, θ, φ) then is

$$\varepsilon(r,k) = \begin{cases} \varepsilon(k), & r \le r_0, \\ 1, & r > r_0. \end{cases} \tag{A1}$$

In the main text, we have shown the secular equation for TM modes in Eq. (5); the secular equation for TE modes reads

$$\frac{1}{n(k)} \frac{j_{\ell+1}(n(k)kr_0)}{j_\ell(n(k)kr_0)} = \frac{h_{\ell+1}(kr_0)}{h_\ell(kr_0)}, \tag{A2}$$

where we remind ourselves that $\varepsilon(k) = (n(k))^2$.

Having determined the relevant resonant wavenumbers $k^\sigma_{\ell,\nu}$ from Eq. (A2) and/or Eq. (5), the corresponding fields in spherical coordinates are

$$\begin{aligned} \mathbf{E}^\mathrm{TE}_{\ell,m,\nu}(r,\theta,\varphi) &= A^\mathrm{TE}_\ell(k^\mathrm{TE}_{\ell,\nu}) \rho_\ell(r, k^\mathrm{TE}_{\ell,\nu}) \mathbf{v}^\mathrm{TE}_{\ell,m}(\theta,\varphi), \\ \mathbf{E}^\mathrm{TM}_{\ell,m,\nu}(r,\theta,\varphi) &= \frac{A^\mathrm{TM}_\ell(k^\mathrm{TM}_{\ell,\nu})}{\varepsilon(r,k^\mathrm{TM}_{\ell,\nu}) k^\mathrm{TM}_{\ell,\nu} r} \mathbf{v}^\mathrm{TM}_{\ell,m}(r, k^\mathrm{TM}_{\ell,\nu}, \theta,\varphi), \end{aligned} \tag{A3}$$

respectively. In these expressions, the normalizations A^σ_ℓ are given by

$$A^\mathrm{TE}_\ell(k) = \sqrt{\frac{2}{\ell(\ell+1) r_0^3 (\varepsilon(k)-1)}}, \tag{A4}$$

and

$$\begin{aligned} A^\mathrm{TM}_\ell(k) =& \frac{1}{n(k) A^\mathrm{TE}_\ell(k)} \\ &\times \left[\left(\frac{j_{\ell-1}(n(k)kr_0)}{j_\ell(n(k)kr_0)} - \frac{\ell}{n(k)kr_0} \right)^2 \right. \\ &\left. + \frac{\ell(\ell+1)}{(kr_0)^2} + \eta(k) C_\ell(k) \right]^{-1/2}, \end{aligned} \tag{A5}$$

respectively, where in the latter expression

$$\eta(k) = \frac{1}{\varepsilon(\omega)} \frac{\partial(\omega^2 \varepsilon(\omega))}{\partial(\omega^2)} - 1, \tag{A6}$$

and

$$\begin{aligned} C_\ell(k) =& \frac{1}{\varepsilon(k)-1} \left[\frac{2(\ell+1)}{(kr_0)^2} \right. \\ &\left. + \varepsilon(k) \left(\frac{j^2_{\ell+1}(n(k)kr_0)}{j^2_\ell(n(k)kr_0)} - \frac{j_{\ell+2}(n(k)kr_0)}{j_\ell(n(k)kr_0)} \right) \right]. \end{aligned} \tag{A7}$$

In case of the Drude model, Eq. (17), the derivative in Eq. (A6) evaluates to

$$\frac{\partial(\omega^2 \varepsilon(\omega))}{\partial(\omega^2)} = \varepsilon_\infty - \frac{\omega_\mathrm{p}^2}{2\omega(\omega + \mathrm{i}\gamma)^2/(\mathrm{i}\gamma)}. \tag{A8}$$

Going back to Eq. (A3), we furthermore have

$$\rho_\ell(r,k) = \begin{cases} j_\ell(n(k)kr)/j_\ell(n(k)kr_0), & r \le r_0, \\ h_\ell(kr)/h_\ell(kr_0), & r > r_0, \end{cases} \tag{A9}$$

and the two vectorial functions

$$\begin{aligned} \mathbf{v}^\mathrm{TE}_{\ell,m}(\theta,\varphi) =& \csc(\theta) \partial_\varphi Y_{\ell,m}(\theta,\phi) \hat{\mathbf{e}}_\theta \\ &- \partial_\theta Y_{\ell,m}(\theta,\phi) \hat{\mathbf{e}}_\varphi, \end{aligned} \tag{A10}$$

with $\csc(\theta) = 1/\sin(\theta)$, as well as

$$\begin{aligned} \mathbf{v}^\mathrm{TM}_{\ell,m}(r,k,\theta,\varphi) =& \left[\ell(\ell+1) \rho_\ell(r,k) Y_{\ell,m}(\theta,\varphi) \right] \hat{\mathbf{e}}_r \\ &+ \left[\partial_r (r \rho_\ell(r,k)) \partial_\theta Y_{\ell,m}(\theta,\varphi) \right] \hat{\mathbf{e}}_\theta \\ &+ \left[\partial_r (r \rho_\ell(r,k)) \csc(\theta) \partial_\varphi Y_{\ell,m}(\theta,\varphi) \right] \hat{\mathbf{e}}_\varphi. \end{aligned} \tag{A11}$$

In Eqs. (A10) and (A11), we have used the unit vectors in spherical coordinates

$$\hat{\mathbf{e}}_r = \begin{pmatrix} \sin\theta\cos\varphi \\ \sin\theta\sin\varphi \\ \cos\theta \end{pmatrix}, \quad \hat{\mathbf{e}}_\theta = \begin{pmatrix} \cos\theta\cos\varphi \\ \cos\theta\sin\varphi \\ -\sin\theta \end{pmatrix}, \quad \hat{\mathbf{e}}_\varphi = \begin{pmatrix} -\sin\varphi \\ \cos\varphi \\ 0 \end{pmatrix}, \tag{A12}$$

the real spherical harmonics

$$Y_{\ell,m}(\theta,\varphi) = \sqrt{\frac{2\ell+1}{2} \frac{(\ell-|m|)!}{(l+|m|)!}} P^{|m|}_\ell(\cos\theta) \chi_m(\varphi), \tag{A13}$$

where P^m_ℓ are the associated Legendre polynomials, and

$$\chi_m(\varphi) = \begin{cases} \pi^{-1/2} \sin(m\varphi), & m < 0, \\ (2\pi)^{-1/2}, & m = 0, \\ \pi^{-1/2} \sin(m\varphi), & m > 0. \end{cases} \tag{A14}$$

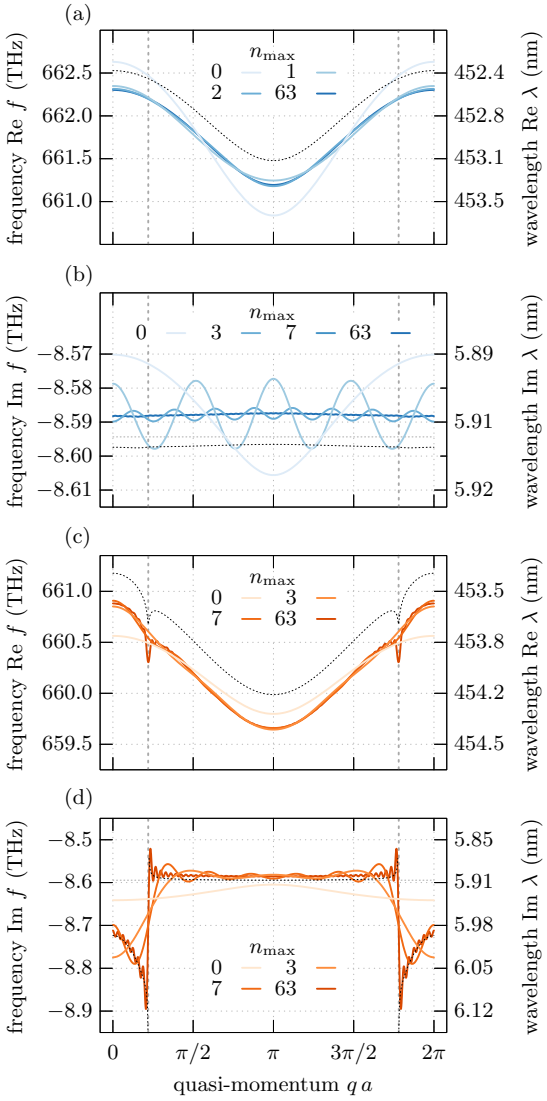


Figure A1. Influence of the cutoff $n_{\max}$ on the frequency bands $\omega_{m,+}(q)$. (a,b) The real and imaginary parts for longitudinal modes ($m = 0$), respectively, in different shades of blue. A larger cutoff corresponds to a darker shade of blue. (c,d) Same as (a,c) for transverse modes ($m = 1$) in oranges. The back dashed lines correspond to the predictions from Ref. 17; the (almost vertical) gray dashed lines in (a,c) to the light lines; the vertical gray dashed lines in (b,d) correspond to the quasi-momenta where the real parts of the bands cross the light line in (a,c), respectively; and the horizontal gray dashed lines in (b,d) to $-\gamma/2$ as predicted by Ref. 16, where γ is the damping rate from the Drude model.

Appendix B: Further analysis for long-range coupling

1. Influence of the coupling range cutoff on the frequency bands

Let us discuss some further details regarding the influence of the cutoff $n_{\max}$ on the frequency bands. Figure A1 shows real and imaginary parts of the upper bands $\omega_{m,+}(q)$. As a reminder, $n_{\max} = 0$ corresponds to nearest-neighbor coupling only; the real parts of the spectra are also shown in Fig. 3(b). Panels (a-d) with $n_{\max} = 63$ correspond to Fig. 5(a,d,c,f).

In (a,b), we show the real and imaginary parts of the spectrum of the longitudinal modes ($m = 0$), respectively, where a darker shade of blue corresponds to a larger cutoff $n_{\max}$. The same is shown in (c,d) for the transverse modes ($m = 1$) with different shades of orange. As in the main text, the black dashed curves correspond to the results obtained from Ref. 17, and the (almost vertical) gray dashed lines in the real parts of the frequency bands in (a,c) correspond to light lines. The horizontal gray dashed lines in the imaginary parts in (b,c) are $-\gamma/2$, relating to the intrinsic material losses from the Drude model.

As seen for the real parts of the longitudinal bands in (a), including next-to-nearest-neighbor couplings ($n_{\max} = 1$) already shows qualitative good agreement compared to the real parts with $n_{\max} = 63$. Long-range coupling has a minute effect on the real parts compared to the step from $n_{\max} = 0$ to $n_{\max} = 1$. For the imaginary parts in (b), the discussed oscillatory behavior becomes apparent. By increasing the cutoff the curves become more flat. However, even in the case of $n_{\max} = 63$, the band still shows oscillations, albeit with a tiny amplitude. This curve is not becoming entirely flat, as also seen in the case of the black dashed curve from Ref. 17. The horizontal gray dashed line encodes constant material loss, predicted by Ref. 16.

In (c,d), we show the real and imaginary parts of the transverse frequency bands, respectively. For the real parts of the bands in (c), we see good qualitative agreement away from the light lines starting from $n_{\max} = 3$. To actually capture the dip close to the light lines, one needs to include couplings beyond $n_{\max} = 7$, underlining that this phenomenon is truly originating from the long-range coupling. For the imaginary parts in (d), increasing $n_{\max}$ leads to a gradual change to capture the sharp transition related to the light lines. Having $n_{\max} = 63$, one can very clearly see the overshooting and subsequent drop close to this sharp transition. Unlike in the case of the longitudinal upper band, the imaginary parts are only close to $-\gamma/2$ between the light lines.

In Refs. 73 and 17, a polariton splitting at the light line is discussed, which resembles an avoided crossing. For the parameters we chose this is not apparent in the real parts of the bands, as the light lines are almost vertical, but it is clearly visible in the imaginary parts. As implied in

the main text, our approach cannot capture such a splitting, as our frequency bands $\omega_{m,+}(q)$ are single-valued, cf. Eq. (28). However, we find that such an avoided crossing is mimicked quite well by the overshooting discussed above, which is for example well visible in the region around $q = \pi/4$.

In summary, we see that it is enough to set $n_{\max} = 3$ to get qualitative good agreement in the real parts away from the light lines. However, to capture the features close to the light lines, as well as the imaginary parts, increasing $n_{\max}$ is necessary. As the real parts of the spectra are dominated by the real parts of the couplings, setting $n_{\max} = 3$ corresponds to including all couplings up to approximately 200 nm, which is the region where the real parts of the coupling coefficients are dominating as seen in Fig. 2. The oscillatory long-range tail of the couplings then is responsible for the features close to the light lines as well as the imaginary parts.

2. Influence on the bulk states

In the main text, we discuss the influence of long-range couplings on the edge states, cf. Fig. 6. In Fig. A2, we show some bulk states. Overall, the changes by including all-to-all coupling are minor, but larger compared to the edge states.

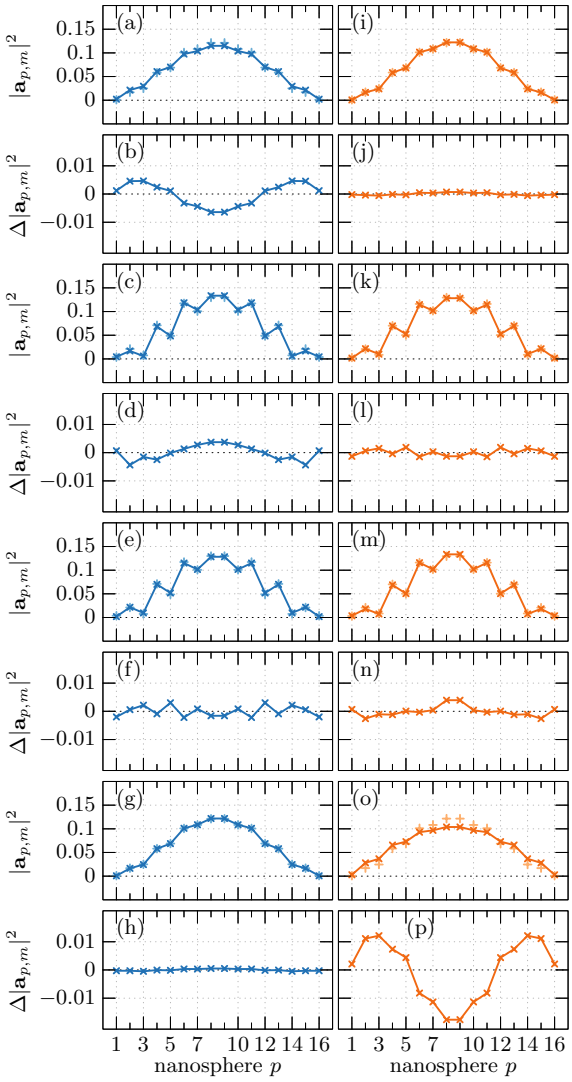


Figure A2. Influence of long-range (all-to-all) coupling on the bulk states. Similar representation as in Fig. 6, where the left column (a-h) corresponds to $m = 0$ in blue, and the right column (i-p) to $m = 1$ in orange. Panels (a,c,e,g,i,k,m,o) show the amplitudes $|\mathbf{a}_{p,m}|^2$ with long-range couplings (dark) and with nearest-neighbor coupling only (light), whereas panels (b,d,f,h,j,l,n,p) show the amplitude differences $\Delta|\mathbf{a}_{p,m}|^2$. The selected states correspond to the ones with the lowest real frequency (a,b,i,j), the frequency directly below the real gap (c,d,k,l), the frequency directly above the gap (e,f,m,n), and the highest frequency (g,h,o,p), corresponding to $j = 1, 7, 10, 16$ in Figs. 5(b,e), respectively.

6

A Modal Analysis of Topological Plasmonic Nanoresonator Chains

In Chapter 5, we discussed a coupled-mode formalism that utilized the resonances of a single simulated particle to predict the eigenspectra of the paradigmatic SSH chain of periodic plasmonic materials. We did this for spherical nanoresonators and cuboid nanoresonators. The formalism predicted edge modes in the eigenspectra. In this chapter, we discuss full lattice simulations and compare some of the results with those of the coupled-mode theory.

The coupled-mode formalism is considered complementary to the computational approach presented in this chapter. The foundation of the computational approach discussed here was presented in Chapter 3. The simulations were done for the SSH chains for spherical nanoresonators and cuboid nanoresonators. The main question we tried to answer was: Is bulk-boundary correspondence preserved for plasmonic SSH chains, and what basis resonance modes contribute to the edge modes of these systems?

The simulations were done with the COMSOL Multiphysics Radiofrequency (RF) module, which uses the finite element method (FEM), briefly discussed in Chapter 3. The COMSOL files were built using the setup from Comsol files from Lalanne's group [196]. The resonators were simulated such that COMSOL outputs complex eigenfrequencies and the appropriate eigenmodes. We focused on systems with a uniform background and with small resonators for simplicity. The bandstructure of these systems was computed, as well as the spectra for the finite chain. These were done for various particle spacings and hence couplings. The Zak phase was then computed from the bulk bands for the various spacings, while for the corresponding finite chain, we checked for the existence of edge modes to verify bulk-boundary correspondence. In addition, we compared the coupled-mode theory predictions with the full lattice simulations. We also used the complementary coupled-mode theory to estimate the contributions of the basis resonances of a single resonator to the hybridized modes of the full lattice. This tells us which basis modes are dominant in the edge modes. Overall, we observed band degeneracy for the spherical resonators, but clearly spaced bands for the cuboids. This comes from the spherical symmetry of the spherical resonators, which is lacking for the cuboids. The work also revealed that bulk-boundary

correspondence holds in the systems we investigated, as a non-trivial (non-zero) Zak phase was observed for the spacings (couplings) where edge modes were present. Meanwhile, no edge modes were observed where the Zak phase was trivial (zero). We also found that the coupled-mode theory performed better for the spherical resonators than for the cuboids. We conclude that this was due to the fact that too few modes were used to compute the bands for the cuboids for the coupled-mode theory. We used only one basis mode for the cuboid, whereas three were used for the spheres. This was further supported by the modal decomposition calculations, which showed that modes well beyond the first basis mode contribute to the edge modes of the cuboid. Further evsluating this is beyond the scope of the paper presented here. The modal decomposition of the spherical resonator chain revealed dipole mode contributions dominating the edge modes. We also checked for edge modes beyond the two lowest bands for the cuboid and found four, with the Zak phase again being non-trivial in this region.

I was also able to use the eigenmodes from the full-wave simulations to compute a topological invariant known as the Zak phase. These eigenmodes corresponded to the bands in the bandstructure. The topological invariant is hence bulk information. I then did simulations of an open chain and demonstrated the existence of edge modes (the boundary information). We saw that when the bulk information indicated a non-trivial topology (that is, when we obtained a non-zero Zak phase), we obtained edge modes. This hints at bulk-boundary correspondence. This is a non-trivial result, as it is not guaranteed that bulk-boundary correspondence exists in plasmonic systems.

For this work, I built and executed the computational models for full lattice simulations

in both periodic and open-boundary (finite) conditions. I constructed and executed the computational models for the single-particle simulations used in the coupled-mode theory. I developed and implemented the algorithms to extract the eigenmodes and compute the Zak phase. I performed all the computations used to demonstrate bulk-boundary correspondence. The idea of computing the basis resonance contributions to the hybridized mode of the full lattice system came from Julius Gohsrich, but the algorithms and computations developed to perform these computations were done by me. I also performed the analysis work and wrote this paper.

Both the coupled-mode formalism and the computational work are forming the foundations of future work in the group that will focus more on device applications.

A Modal Analysis of Topological Plasmonic Nanoresonator Chains

Norman S. Israel[1,2], Julius Gohsrich[2,3,4], Flore K. Kunst[2,3,4], and Lora Ramunno[1,2]

[1]Department of Physics and Nexus for Quantum Technologies Institute, University of Ottawa, Ottawa, ON K1N 6N5, Canada
[2]Max Planck - uOttawa Centre for Extreme and Quantum Photonics, University of Ottawa, Ottawa, ON K1N 6N5, Canada
[3]Max Planck Institute for the Science of Light, Staudtstraße 2, 91058 Erlangen, Germany
[4]Department of Physics, Friedrich-Alexander-Universität Erlangen-Nürnberg, Staudtstraße 7, 91058 Erlangen, Germany

*nisra058@uottawa.ca

Abstract

Topological plasmonic systems are commonly analyzed using the tight-binding and dipole approximations. Their eigenvalues are complex-valued due to loss being natural to plasmonics, making them natural hosts for non-Hermitian topological phenomena. However, the dipole approximation hides the influence of the nanoparticles' geometry on the topological nature of the system. We present Finite Element Method (FEM) simulation results for a framework that uses the fundamental resonances of spherical and cuboid nanoresonators to study topological plasmonic chains. The method combines full lattice eigenmode simulations with a tight-binding model that uses single-particle eigenmode simulations. Our approach allows us to determine the resonant mode contributions to the eigenspectra of a periodic arrangement of these nanoparticles. We studied Su-Schrieffer-Heeger (SSH) chains of Drude model spheres and cuboids. We compute the eigenspectra of these chains when periodic and open-boundary conditions are implemented. For spherical nanoparticles, we find 3 pairs of localized modes that are consistent with non-trivial topology, whereas for cuboid nanoparticles we find 1 pair of localized edge modes between the 2 lowest bands and 2 pairs between the next higher bands. We compute the contributions of the single-particle resonances to a finite-size lattice for both cases, where we find that dipole resonances dominate localized modes for the spheres, while higher-order resonances are relevant for cuboids. We also demonstrate the relevance of long-range contributions to both systems, using our tight-binding model. Our results demonstrate that the cuboid systems exhibit even stronger long-range couplings compared with the spherical systems. Surprisingly, we found that a non-zero Zak phase is consistent with the existence of localized states at the edge of the finite chains. Our results establish a shape- and environment-aware framework that can be used to study non-Hermitian topology in plasmonics, providing direct insight into the role of geometry and resonances in topological states. This approach has the potential to accelerate the discovery of non-Hermitian topological materials in plasmonic platforms that have a range of technological applications.

Topological materials have gained significant attention due to the various applications of their defect-immune backscatterless boundary states. The appearance of these states is dictated by the bulk-boundary correspondence, which uses the eigenfunctions obtained from an infinite bulk to compute how many states appear on the boundary of a finite version of the bulk system. They were first discovered in condensed matter systems[1–4] and show promising signs of applications.[5,6] Since their discovery, they have been experimentally realized in acoustic metamaterials,[7] electric circuits,[8,9] photonic metamaterials,[10–12] and mechanical systems.[13–16] Much of the focus has been on lossless (Hermitian) platforms. However, systems with either gain or loss (non-Hermitian) have attracted significant interest due to the exotic new topological phenomena that they host[17–19] such as the non-Hermitian skin effect[20–23] and exceptional points.[24–28] These phenomena have been harnessed for several applications such as sensing,[29] lasing,[30–32] wavefront shaping along with singular optics and other applications.[33] However, it is difficult to access non-Hermitian topology in condensed matter systems. This is where periodic metallic nanoresonator arrays offer an opportunity.

Metallic nanoresonators are subject to ohmic losses[34] and therefore are natural hosts for studying non-Hermitian phenomena. Contrary to electronic materials, they have a significantly larger design space. As a result, there is significant room to design and build nanoresonator lattices to probe non-Hermitian effects experimentally. When metallic nanoresonators interact with an electric field, the free electrons in the metals oscillate collectively. These collective oscillations, known as surface plasmons, have resonances at which the fields are confined and amplified in a manner dependent on the type of metal, as well as the geometry and the surrounding medium. Such resonances can be engineered to have desirable characteristics with applications ranging from sensing,[35–41] cell-engineering,[42] improving solar cell efficiency,[43] cancer phototherapy,[44,45] and surface-enhanced Raman spectroscopy,[46] to name a few. The resonant modes of these oscillations are called quasinormal modes (QNMs).

Indeed, the collective interactions between many metallic nanoresonators produce rich optical behaviour. When these resonators are periodically arranged, bandstructures can be obtained to study their topology, thus establishing a link between metallic nanoresonator arrays and condensed matter systems. In this context, theoretical[47–52] and experimental[53,54] studies of the one-dimensional Su-Schrieffer-Heeger (SSH) chain have been reported. Much of the theoretical work relied on the dipole approximation[47–51] or transfer matrix approach,[52] with the former being the most common. There have also been experimental studies of topological boundary modes in plasmonic SSH chains. The dipole approximation is restricted to metallic spheres in a homogeneous background within the regime $d \geq 3R$,[55] where R is the sphere's radius and d is the distance between two particles. Contrary to most condensed matter systems, long-range couplings play an important role in nanoresonator arrays.[48] The importance of long-range, retardation, and radiative contributions to plasmonic mode couplings and how they influence optical effects has been reported over the last two decades.[56–58] Despite this, a detailed shape-agnostic and unified approach is lacking. In addition, works presented so far[47–49,59] have not presented a unified formalism that allows for broad and efficient exploration of the behaviour of the multipolar resonances in topological plasmonic systems. As a result, explorations regarding the limitations of bulk-boundary correspondence in metallic nanoresonator arrays and non-Hermitian systems have been restrictive. The work we present in this paper sheds light on an approach that shows promise in reducing these restrictions, as it removes the limitations on shapes and environments that can be explored, while retaining the ability to probe the effect of long-range, multimode couplings as well as retardation and radiative effects beyond a 2-band system.

The goal of this paper is to develop a rigorous computational model based on a framework[60] which treats individual QNMs as part of a basis set that can be used to describe the physical properties of a collection of resonators. The computational framework is robust, complete, and can incorporate arbitrary shapes and materials, arbitrary coupling range, and can access an arbitrary number of bands in both homogeneous and substrate environments. The method we use has been discussed in detail in various manuscripts,[61,62] and has been applied to the optimization and design of nonlinear nanostructures,[63] but to the best of our knowledge, has never been applied to study topology in plasmonic arrays. In this work, we do not explicitly probe non-Hermitian phenomena such as exceptional points or the non-Hermitian skin effect in plasmonic systems; instead, we develop and present a framework for this purpose. We use the framework to study a reciprocal system containing Hermitian-like topology with some loss on top. Specifically, we study the Su-Schrieffer-Heeger (SSH) chain,[64] a paradigmatic example of a topological system from condensed matter physics, which hosts zero-energy states at the boundary in the topological regime. The bulk-boundary correspondence principle manifests itself in the fact that in this system, we use the bulk eigenstates to compute a quantity, the Zak phase, whose value determines the presence/absence of zero-energy states at the ends of the chain. The Zak phase is a topological invariant, as it remains constant as long as we do not cross a band closure. We consider SSH chains composed of spherical nanoparticles in air (depicted in Fig. 1) and cuboids in glass (depicted in Fig. 2), both described by a Drude model. Using finite-element method (FEM) simulations, we show the existence of topological edge modes in both. In related work,[65] we built a non-Hermitian tight-binding model that resembles those in condensed matter systems. We can explicitly compute short-range and long-range couplings and an arbitrary number of bands. Using the tight-binding models, we demonstrated that long-range couplings are relevant to both chains. The presence of edge modes even while long-range coupling is present, is surprising, as long-range coupling breaks chiral symmetry, which would result in no localized edge modes. We also show how we can use the tight-binding model to resolve the edge states to reveal the contribution of different resonances. This is important to understand which modes from the basis set contribute to the localized edge modes.

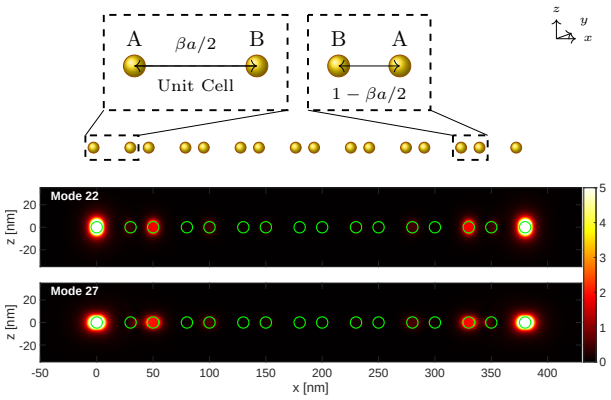


Figure 1: TOP: Schematic of an SSH chain of spherical nanoresonators in air for $\beta = 1.2$. BOTTOM: Mode plot of two edge modes in the xz-plane, scaled by 10^9 V/m.

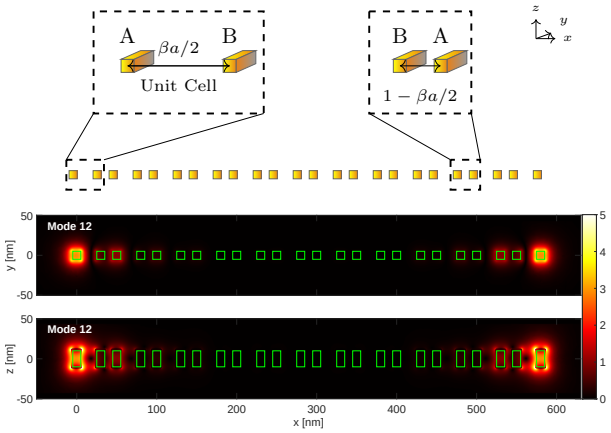


Figure 2: TOP: Schematic of an SSH chain of cuboid nanoresonators in glass for $\beta = 1.2$. BOTTOM: Mode plot of one edge mode in the xy-plane (first mode plot) and xz-plane (second mode plot), scaled by 10^9 V/m.

The remainder of our paper is structured as follows: In the Methods section, we discuss how we simulate and set up the models, including the parameters chosen, and discuss how we compute the coupling coefficients. In the Spheres and Cuboid sections, we discuss the results for the spheres and cuboids, respectively, for infinite and finite structures; then, in the Modal Decomposition section, we discuss the results demonstrating the single-particle mode influence on the hybridized modes of the full lattice. In the Bulk-Boundary Correspondence section, we show how our results are consistent with the bulk-boundary correspondence. Finally, in the Summary and Discussion section, we summarize our results and discuss future directions.

Methods

In this section, we discuss the numerical methods and the entire process, as well as how the single-particle simulations and the full-lattice simulations complement each other.

Numerical Methods

For our work, we use the COMSOL Multiphysics[66] eigenfrequency solver. We simulate two types of chains with homogeneous backgrounds consisting of: (1) spherical and (2) cuboid nanoresonators. We present the results of the simulation of an array of spherical and cuboid nanoparticles and compare them with the tight-binding model. We use a material model (Drude), nanoparticle geometry, and nanoparticle arrangements (i.e., if it is a chain, the dimensionality of the lattice, type of unit cells, etc.), and simulate an entire metallic nanoparticle array. One must also specify the boundary conditions: for a structure with an infinite number of particles, where we wish to compute the bandstructure, we use periodic boundary conditions (PBCs) along the x-direction of the chain. The other directions (y, z) have perfectly matched layers (PMLs), which are boundaries that absorb the fields as we move away from the resonator, making them decay as our distance increases. If we have a finite chain with N particles, we surround the simulation space with PMLs so that the fields are well-behaved in all directions. This is how we implement open-boundary conditions (OBCs). We then perform the simulation of an array. When we use PBC, we simulate a single unit cell, which for us is the two particles, A and B (depicted by the zoomed region of chains in Fig. 1 and Fig. 2). We then extract the eigenvectors (quasinormal Bloch modes) and their associated eigenvalues (the eigenfrequencies). Using these quasinormal modes, we compute the Zak phase.

Using open-boundary conditions, we simulate and extract the quasinormal modes and associated eigenfrequencies while we search for localized edge modes. We repeat this process for multiple values of the distance parameter, $\beta \in [0.8, 1.2]$, where β is defined in Figs. 1 and 2. Finally, we use the bulk modes to compute a quantity known as the Zak phase. It can either be π or 0. If it has a value of π, then our system is in a topological phase. At this value, we should observe edge modes in a finite chain representation of the system. If it has a value of 0, then the system is not in the topological phase. A finite chain representation would give no edge modes. We did this for the different values of β at which the band structure is evaluated. We also plot the eigenfrequencies of the finite chains as a function of β, and search for signs of the edge modes in the frequency gap.

Parameter Details

We start with spherical nanoparticles embedded in air, as they have exact solutions, which can be obtained using Mie theory. In topological systems, we look at the periodic boundary case to compute the bandstructure, and the open-boundary case to confirm the existence of edge modes. For our system, we work with Drude spheres of radius $R = 5$ nm and lattice constant $a = 50$ nm. For the cuboids, we use dimensions $14\text{ nm} \times 14\text{ nm} \times 25\text{ nm}$ $(L \times W \times H)$. The material of the nanoparticles is described by a Drude model, i.e.

$$\epsilon(\omega) = \epsilon_\infty - \frac{\bar{\omega}_p^2}{\omega^2 - i\omega\Gamma}. \tag{1}$$

We use the same parameters as in Pocock et al,[48] where $\epsilon_\infty = 9.1$, $\bar{\omega}_p = 1.38 \times 10^{16}$ rad/s and $\Gamma = 1.08 \times 10^{14}$ rad/s. Our background is air for the spheres, with $\epsilon_B = 1.00$, and glass for the cuboids, with $\epsilon_B = 2.25$. In COMSOL, we performed convergence tests for the spheres to determine our upper mesh and lower mesh values to be 1.0 nm and 0.5 nm, respectively, inside the resonators. For the cuboids, our upper and lower mesh values were 1.5 nm and 1.0 nm, respectively, inside the resonator. In both cases, the PML has a thickness of 15 nm. Results of the convergence tests are presented in Section , where we see convergence for the real and imaginary parts shown for the spheres in Fig. S1 and cubids in Fig. S2. These were done for a single resonator, and the values were chosen so that we are at the limits of our computational resources. Using these parameters, we simulated a periodic SSH chain as well as the finite chain.

Mathematical Problem

As our system is dispersive, we follow the approach taken in the modal analysis of nanoresonators (MAN) user manual[67] and solve the auxiliary eigenvalue problem

$$\hat{\mathcal{H}} \left|\tilde{\psi}_m(\vec{r})\right\rangle = \tilde{\omega}_m \left|\tilde{\psi}_m(\vec{r})\right\rangle, \tag{2}$$

where

$$\hat{\mathcal{H}} = \begin{bmatrix} 0 & -i\mu_0^{-1}\nabla\times & 0 & 0 \\ i\epsilon_\infty^{-1}\nabla\times & 0 & 0 & -i\epsilon_\infty^{-1} \\ 0 & 0 & 0 & i \\ 0 & i\omega_p^2\epsilon_\infty & -i\omega_0^2 & -i\Gamma \end{bmatrix}, \tag{3a}$$

$$|\tilde{\psi}_m(\vec{r})\rangle = \begin{bmatrix} \tilde{H}_m(\vec{r}) \\ \tilde{E}_m(\vec{r}) \\ \tilde{P}_m(\vec{r}) \\ \tilde{J}_m(\vec{r}) \end{bmatrix}. \tag{3b}$$

Here, $|\tilde{\psi}_m(\vec{r})\rangle$ in Eq. 3b are the quasinormal modes, with their associated eigenvalues being $\tilde{\omega}_m$ with the time-harmonic dependence $e^{-i\tilde{\omega}_m t}$. We also note that $\tilde{P}_m = \left(-\frac{\bar{\omega}_p^2}{\omega^2 - i\omega\Gamma}\right)\tilde{E}_m$ and $\tilde{J}_m = -i\omega P_m$. As $\hat{\mathcal{H}}$ in Eq. 3a is non-Hermitian, then $\tilde{\omega}_m \in \mathbb{C}$. In addition, non-Hermitian eigenvalue problems have two associated eigenstates, a right-eigenstate, given by Eq. 3b, and a left-eigenstate given by $\langle\tilde{\psi}_m^*(\vec{r})|\,\hat{D}$, where

$$\hat{D} = \begin{bmatrix} -\mu_0 & 0 & 0 & 0 \\ 0 & \epsilon_\infty & 0 & 0 \\ 0 & 0 & \frac{\omega_0^2}{\epsilon_\infty\bar{\omega}_p^2} & 0 \\ 0 & 0 & 0 & \frac{-1}{\epsilon_\infty\bar{\omega}_p^2} \end{bmatrix}. \tag{4}$$

In COMSOL, we use the Electromagnetic Waves, Frequency Domain Module along with the Weak Form PDE Module as presented in the MAN user manual.[67] In order to do this, we have to convert Eq. 2 into a quadratic eigenvalue problem. We do this by writing $\tilde{H}_m$ and $\tilde{J}_m$ in terms of $\tilde{E}_m$ and $\tilde{J}_m$ and substituting them into Eq. 2 to get the expressions

$$\nabla \times \mu_0^{-1}\nabla \times \tilde{E}_m - \tilde{\omega}_m^2\epsilon_\infty\tilde{E}_m - \tilde{\omega}_m^2\tilde{P}_m = 0, \tag{5a}$$

$$\epsilon_\infty\bar{\omega}_p^2\tilde{E}_m - \omega_0^2\tilde{P}_m + i\Gamma\tilde{\omega}_m\tilde{P}_m + \tilde{\omega}_m^2\tilde{P}_m = 0. \tag{5b}$$

If we set

$$\vec{x}_m = \begin{bmatrix} \tilde{E}_m \\ \tilde{P}_m \end{bmatrix}, \tag{5c}$$

we find

$$\hat{L}\vec{x}_m + \tilde{\omega}_m\hat{R}\vec{x}_m + \tilde{\omega}_m^2\hat{M}\vec{x}_m = 0, \tag{6}$$

where

$$\hat{L} = \begin{bmatrix} \nabla \times \mu_0^{-1}\nabla\times & 0 \\ \epsilon_\infty\bar{\omega}_p^2 & -\omega_0^2 \end{bmatrix}, \tag{7a}$$

$$\hat{R} = \begin{bmatrix} 0 & 0 \\ 0 & i\Gamma \end{bmatrix}, \tag{7b}$$

$$\hat{M} = \begin{bmatrix} -\epsilon_\infty & -1 \\ 0 & 1 \end{bmatrix}. \tag{7c}$$

Eq. 6 is the quadratic eigenvalue problem,[67] which we write as

$$\int_V \left[\left(\nabla \times \vec{T}(\vec{r})\right) \cdot \left(\mu_0^{-1}\nabla \times \tilde{E}_m\right) - \tilde{\omega}_m^2\epsilon_\infty\vec{T}(\vec{r}) \cdot \tilde{E}_m - \tilde{\omega}_m^2\vec{T}(\vec{r}) \cdot \tilde{P}_m \right] d^3r = 0, \tag{8a}$$

$$\int_V \left[\epsilon_\infty\bar{\omega}_p^2\vec{T}(\vec{r}) \cdot \tilde{E}_m - \omega_0^2\vec{T}(\vec{r}) \cdot \tilde{P}_m + i\Gamma\tilde{\omega}_m\vec{T}(\vec{r}) \cdot \tilde{P}_m + \tilde{\omega}_m^2\vec{T}(\vec{r}) \cdot \tilde{P}_m \right] d^3r = 0. \tag{8b}$$

Eq. 8a and Eq. 8b are known as weak-form partial differential equations (PDEs), where $\vec{T}(\vec{r})$ is called a test function. The volume integral is over the entire space. Comsol solves Eq. 8a and Eq. 8b with its Weak Form PDE module, obtaining the quasinormal modes and associated eigenfrequencies, which satisfy

$$\langle\tilde{\psi}_m^*(\vec{r})|\hat{D}|\tilde{\psi}_n(\vec{r})\rangle = \delta_{mn}. \tag{9}$$

Eq. 9 is known as the bio-orthonormal condition. It gives the results

$$\langle\tilde{\psi}_m^*(\vec{r})|\hat{D}|\tilde{\psi}_n(\vec{r})\rangle = \int_V \left[\epsilon_\infty \tilde{E}_m(\vec{r})\cdot\tilde{E}_n(\vec{r}) - \mu_0\tilde{H}_m(\vec{r})\cdot\tilde{H}_n(\vec{r}) + \frac{\omega_0^2}{\epsilon_\infty\bar{\omega}_p^2}\tilde{P}_m(\vec{r})\cdot\tilde{P}_n(\vec{r}) - \frac{1}{\epsilon_\infty\bar{\omega}_p^2}\tilde{J}_m(\vec{r})\cdot\tilde{J}_n(\vec{r})\right] d^3r = \delta_{mn}, \tag{10a}$$

Eq. 10a provides the inner product in terms of the fields and is derived from the Lorentz reciprocity theorem. We note that generally, for the systems we studied, $\tilde{H}_m(\vec{r}) \ll \tilde{E}_m(\vec{r})$, while the term with $\tilde{P}_m(\vec{r})$ is 0 because $\omega_0 = 0$ for the Drude model and the term with with $\tilde{J}_m(\vec{r})$ is very small since $\bar{\omega}_p$ is quite large. As a result, the dominant term is the $\tilde{E}_m(\vec{r})$ term.

In addition to the QNMs, we find so-called numerical modes (PML modes). We distinguish between the two kinds of modes by the fact that QNMs decay away from the resonator, whereas PML modes do not. More precisely, the energy density decays exponentially away from the resonator when we have a QNM, whereas for PML modes, the energy density increases when we are away from the resonator. All our work on computing the bandstructure and the finite eigenspectra for both the tight-binding results and the simulation of the entire array is done using only the QNM modes.

Computing the Zak Phase

Next, we use the QNMs corresponding to the bulk bands of the simulations of the full structure to compute the Zak phase of the system. The k-points are placed on a 1D grid, and the inner product at adjacent k-points is computed as

$$M_{mn}(k_i) = \langle\tilde{\varphi}_{mk_i}(\vec{r})|\tilde{\psi}_{nk_{i+1}}(\vec{r})\rangle\delta k. \tag{11}$$

This is used to compute the non-Hermitian Wilson loop

$$W = \prod_{i=1}^{N-1} M(k_i), \tag{12}$$

where N is the number of k-points, and the Zak phase γ_n for the band n is

$$\gamma_n = \Im\left[\ln\prod_{i=1}^{N-1} M(k_i)\right], \tag{13}$$

where $\gamma_n \in \mathbb{C}$ and $\langle\tilde{\varphi}_{mk_x}(\vec{r})| = \langle\tilde{\psi}^*_{nk_x}(\vec{r})|\hat{D}$ is the left eigenvector[60] and $|\tilde{\psi}_{nk_x}(\vec{r})\rangle$ is the right eigenvector. Details regarding a simplified derivation of the non-Hermitian Wilson loop, involving a Taylor expansion, are presented in Section . However, the concept more rigorously rests on mathematical spaces known as fibre bundles.

In Hermitian systems, $\gamma_n \in \mathbb{R}$ is due to the system undergoing unitary evolution, which preserves the magnitude of the inner product of the eigenstates. In the non-Hermitian case, the system undergoes non-unitary evolution, so while $\mathrm{Re}[\gamma_n]$ is the same quantity as in the Hermitian case, $\mathrm{Im}[\gamma_n]$ is a decay/damping term arising from the fact that the inner product of the eigenstates of a non-Hermitian system decreases.

Theoretical Methods

In the tight-binding model, only one particle is simulated. We input the material model (Drude) and nanoparticle geometry. We simulate the types of particles that make up a single unit cell. In our case, it is one eigenmode simulation for a structure of interest. This produces a set of modes and frequencies for that

particle. This has value in the design of plasmonic arrays, as eigenmode simulations of a complete structure of such systems consume a significant amount of compute. Single-particle simulations consume significantly less compute, saving time and resources, and hence speeding up the design process. The quasinormal modes resulting from the simulation of the entire system can be assumed to be the result of the hybridization of the modes of the single-particle simulations. This assumption is what is used to construct the tight-binding model. In another paper, we presented this tight-binding model incorporating long-range couplings and an arbitrary number of bands. This tight-binding model is shape- and environment-aware, as it involves the simulation of a single particle in the array. This way, we construct a semi-analytic model. After we perform the single-particle simulation, we extract the desired number of quasinormal modes and the associated eigenfrequencies. As a result, we truncate the number of modes and eigenfrequencies used to reconstruct the physics of the entire system. As a result, our tight-binding model not only relies on a truncated set of modes, but also a truncated range for the coupling. It also assumes frequency independence of the coupling. We then extract the eigenfrequencies and a set of modal coefficients which, when used in a modal expansion, give us the associated eigenmode.

We can then compare the tight-binding and full lattice results, where we use the full lattice results as the standard. The full lattice simulations do not truncate the modes or coupling range. We can then use the tight-binding model to estimate how much a particular mode of a single particle contributes to the modes of the simulation of the entire system; essentially performing a modal decomposition of the hybridized modes. This approach can be generalized to arbitrarily shaped nanoparticles and setups. For example, we can have different kinds of particles making up a single unit cell. We can also have two spheres of the same size but different materials, or the two particles could be made of the same material but be different shapes and sizes. We then use the QNMs and associated eigenfrequencies for this single particle to construct the tight-binding model in order to compute the spectral properties of the system, without simulating it in its entirety. Details are provided in the theory paper.

The Tight-Binding Model

The tight-binding model gives the equations for the periodic-boundary case as

$$\mathbb{H}_{\text{pbc}} = \begin{bmatrix} \tilde{\mathbf{\Omega}}^A & \mathbb{V}^{AB} + \mathbb{W}^{AB} e^{-ika} \\ \mathbb{V}^{BA} + \mathbb{W}^{BA} e^{ika} & \tilde{\mathbf{\Omega}}^B \end{bmatrix}, \tag{14}$$

where k is the wavenumber. Meanwhile, the open-boundary case is given by

$$\mathbb{H}_{\text{finite}} = \begin{bmatrix} \tilde{\mathbf{\Omega}}_1^A & \mathbb{V}_1^{AB} & 0 & 0 & \cdots & 0 \\ \mathbb{V}_1^{BA} & \tilde{\mathbf{\Omega}}_1^B & \mathbb{W}_1^{BA} & 0 & \cdots & 0 \\ 0 & \mathbb{W}_1^{AB} & \tilde{\mathbf{\Omega}}_2^A & \mathbb{V}_2^{AB} & \cdots & 0 \\ 0 & 0 & \mathbb{V}_2^{BA} & \tilde{\mathbf{\Omega}}_2^B & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \mathbb{V}_N^{AB} \\ 0 & 0 & 0 & \cdots & \mathbb{V}_N^{BA} & \tilde{\mathbf{\Omega}}_N^B \end{bmatrix}, \tag{15}$$

where $\tilde{\mathbf{\Omega}}^{A,B}$ are the on-site frequencies for particle A or B given by

$$\tilde{\mathbf{\Omega}}^{A,B} = \begin{bmatrix} \tilde{\omega}_{A,B}^{(1)} & 0 & \cdots & 0 \\ 0 & \tilde{\omega}_{A,B}^{(2)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \tilde{\omega}_{A,B}^{(n)} \end{bmatrix}, \tag{16}$$

where $\tilde{\omega}_{A,B}^{(n)}$ is the nth on-site eigenfrequency for particle A/B, $\mathbb{V}^{AB}$ and $\mathbb{W}^{AB}$ are the intercell and intracell couplings, respectively, over all relevant modes. They are defined as

$$\mathbb{V}^{AB} = \begin{bmatrix} V_{11}^{AB} & V_{12}^{AB} & \cdots & V_{1N}^{AB} \\ V_{21}^{AB} & V_{22}^{AB} & \cdots & V_{2N}^{AB} \\ \vdots & \vdots & \ddots & \vdots \\ V_{N1}^{AB} & V_{N2}^{AB} & \cdots & V_{NN}^{AB} \end{bmatrix}, \tag{17a}$$

$$\mathbb{W}^{AB} = \begin{bmatrix} W_{11}^{AB} & W_{12}^{AB} & \cdots & W_{1N}^{AB} \\ W_{21}^{AB} & W_{22}^{AB} & \cdots & W_{2N}^{AB} \\ \vdots & \vdots & \ddots & \vdots \\ W_{N1}^{AB} & W_{N2}^{AB} & \cdots & W_{NN}^{AB} \end{bmatrix}. \tag{17b}$$

V_{mn}^{AB} and W_{mn}^{AB} are computed using

$$W_{mn}^{pq'} = \frac{\bar{\omega}_p}{N_{p,m}} \int_{\tilde{\mathrm{V}},p} \tilde{E}_{m,p} \Delta\epsilon(\tilde{\omega}_p) \cdot \tilde{E}_{n,q'} \, d\vec{r}, \tag{18}$$

where $N_{p,m}$ is the normalization of the modes of particle p for mode m and Ω_{R} is the resonator volume. Eq. 18 is the coupling of the near-fields between two particles located some arbitrary distance apart. Eq. 18 is frequency-independent, as $\bar{\omega}_p$, the plasma frequency, is fixed for a single particle. Thus, $\kappa(\omega) \approx \kappa$. Without this approximation, the equations used for the tight-binding model would be more complex and are beyond the scope of this paper. Details of the algorithm to do this are presented in Section , but essentially the approach involves taking the single-particle modes and translating their position in space to the location where we imagine the second particle to be. We extract the field values of that shifted particle in the location of the resonator before it is shifted. We then evaluate Eq. 18 using Gauss-Legendre integration in the vicinity of the resonator of interest, p. In general, we get $\mathbb{V}^{AB}, \mathbb{W}^{AB} \in \mathbb{C}$, which already differs from typical Hermitian SSH models studied in condensed matter systems, where we would typically have $\mathbb{V}^{AB}, \mathbb{W}^{AB} \in \mathbb{R}$.

Computing Modal Signatures

Eq. 15, where we have open-boundary conditions (OBC), is where we have a finite number of particles. The eigenvectors in this case are the weights of the eigenmodes at each particle, and we assume our system can be represented using a modal expansion

$$\tilde{E}_l(\vec{r}) = \sum_p \sum_m a_{p,m} \tilde{E}_{p,m}(\vec{r}). \tag{19}$$

In Eq. 19, $|\tilde{\Psi}_l(\vec{r})\rangle$ are the hybridized modes and $|\tilde{\psi}_{p,m}(\vec{r})\rangle$ is the mth basis states of particle p. These only form a complete basis when we also include the PML modes. They are the QNMs of the single particle that we would simulate to build the theoretical model. The eigenvectors of Eq. 15 are the $a_{p,m}$, and the eigenfrequencies of this would ideally correspond to the eigenfrequencies of the full simulations done in Comsol. However, since we typically truncate Eq. 21 as well as coupling range in the theory to a finite number of modes, the accuracy of eigenfrequencies is affected. In the simulations of an entire structure, all modes and physical effects are considered. The simulations of the full system are thus considered the 'standard candle', and we compare our theoretical results to those simulations, using nearest-neighbour as well as long-range couplings. We also use Eq. 21 to approximate $a_{p,m}$ from the simulations in order to perform a modal decomposition, which allows us to determine which modes of a single particle contribute the most to eigenmodes generated from the simulation of the full system. Here, p is the site label and m is the basis (single-particle) mode, obtained from simulating a single mode. We use the fact that the modes are biorthonormal

$$\langle \tilde{\psi}_{q,n}^*(\vec{r}) | \hat{D} | \tilde{\psi}_{p,m}(\vec{r}) \rangle = \delta^{qp} \delta_{nm}, \tag{20}$$

where p, q are particle labels and m, n are mode labels. If we apply Eq. 20 to Eq. 19, we get the expression for the modal coefficients, $a_{q,n}$

$$a_{q,n} = \langle \tilde{\psi}^*_{q,n}(\vec{r}) | \hat{D} | \tilde{\Psi}_l(\vec{r}) \rangle \tag{21a}$$

$$\approx \int_V \epsilon_\infty \tilde{E}_{q,n}(\vec{r}) \cdot \tilde{E}_l(\vec{r})\, d\vec{r} \tag{21b}$$

We then plot the values computed by Eq. 21, and we use this to compute how far into the chain the decay proceeds. This is the decay length of the mode. We compute the decay length by performing an exponential fit

$$|a_m| = a_0 e^{-n/\xi}, \tag{22}$$

where ξ is the decay length and n is the site number. The decay length tells us how fast the mode decays along the chain.

Spheres

In order to benchmark our approach, we first studied a plasmonic array of spheres. Spherical plasmonic arrays can be solved analytically with Mie scattering, so spherical tight-binding results for analytic and semi-analytic approaches provide insight into performance. This comparison was presented in recent work [65] done by the authors of this manuscript. We chose spheres of radius $R = 5$ nm, so that we can also compare with the published results of Pockock et al. [48] In this section, we present the semi-analytic tight-binding predictions for the sphere. First, we present the results for the periodic boundary conditions; then we present the results for the open-boundary conditions.

We start by considering an infinite SSH chain consisting of a linear array of identical plasmonic particles with radius $R = 5$ nm. Periodic boundary conditions (PBCs) are applied, where the unit cell of length 50 nm, contains two particles separated from each other by length $\beta a/2$, with $\beta = 1$ consisting of the case where the particles are evenly spaced, while $\beta \neq 0$ consisting of the case where the particles are dimerized.

In Fig. 3, we present the bandstructure of the real eigenvalues of the array of spheres for $\beta = 1.2$. The imaginary eigenvalues are depicted in Fig. S5. The red points are numerical results obtained from Comsol, the green dashed curve is the quasinormal coupled-mode theory (the tight-binding model) with just nearest-neighbour coupling. The blue broken curve is the quasinormal coupled-mode theory with long-range couplings included (up to 8 unit cells). We find four bands, with the two middle bands doubly degenerate. As expected, a band gap is open for $\beta = 1.2$. The coupled-mode theory (green-solid and blue-dashed) reproduces the bands from the simulations of the full lattice well. We see that while the nearest-neighbour in the coupled-mode theory produced a good profile of the numerical results, long-range effects performed much better.

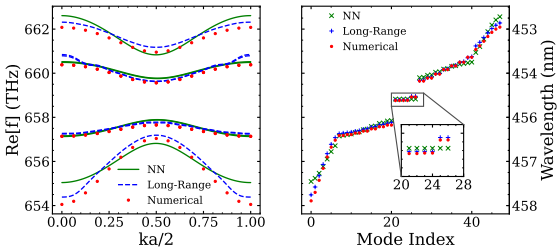


Figure 3: Real part of bandstructure (left) and open-boundary (right) for spherical nanoparticles when $\beta = 1.2$. Numerical simulations are red points. The green bold curve is the tight-binding nearest-neighbour prediction, while the blue broken curve is with long-range coupling. Green x's and blue crosses represent the nearest-neighbour and long-range coupling, respectively.

Next, we consider a finite chain of 16 nanospheres, by applying open boundary conditions. We take the particle radii, unit cell size, and separation lengths to be the same as the periodic boundary case, but

consider two different arrangements for the dimer: one with $\beta = 0.8$, where the separation of the particles within the unit cell is smaller than the separation of particles outside the unit cell, and another with $\beta = 1.2$, where the particles within the unit cell are farther apart from each other than from their neighbours from adjacent unit cells. For $\beta = 0.8$, each particle appears in a pair along the chain, whereas for $\beta = 1.2$, there are two "unpaired" particles at each end.

The right plot of Fig. 3 is the open-boundary (finite) chain of the same system. The red points, green, x's and blue crosses represent the numerical, nearest-neighbour, and long-range coupling for the coupled-mode theory (tight-binding model) when three modes were used from a single-particle simulation. The mode index is just a way to keep track of the mode. As expected, we find edge modes in the gap, as reflected in the open-boundary spectra (right panel in Fig. 3). There are six edge modes present, as we see in the inset. The mode profiles of two of them are depicted in the bottom image of Fig. 1. These edge modes have two flavours: longitudinal (mode 27) and transverse (mode 22). There are, in fact, four transverse edge modes and two longitudinal ones. This is consistent with the fact that the two degenerate bands in the PBC spectrum in Fig. 3 (left) are transverse modes, while the two outer bands are longitudinal modes. These modes are highly localized at the edge of the chain, and decay exponentially away from the edge. The edge modes decay from site A when coming from the left of the chain, and from site B when coming from the right. This is why the brightness decreases at alternating sites. Fig. 4 shows the spectra of the open chain when the system is

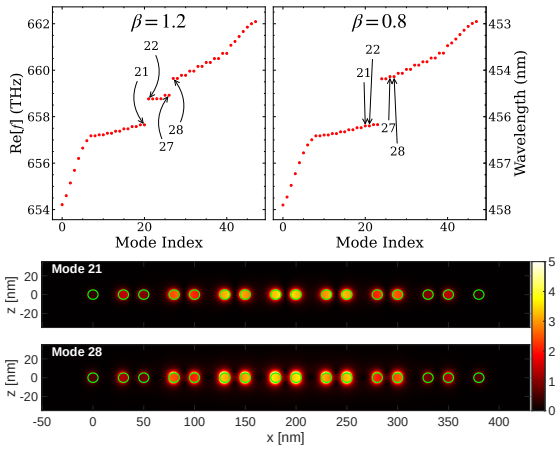


Figure 4: Comparison of simulation results for non-trivial ($\beta = 1.2$, left) and trivial ($\beta = 0.8$, right) of a spherical nanoparticle chain. Edge modes are present in the non-trivial setup; they move into the bulk in the trivial setup. Two bulk modes, 21 and 28, are depicted in the bottom plots, while two edge modes, 22 and 27, are depicted in the bottom plots of Fig 1. The mode plots are scaled by 10^9 V/m.

in a non-trivial state (top-left, shown on the right plot of Fig. 3) and a trivial state (top-right). The edge modes are present in the band gap, while in the trivial state, the edge modes move out of the gap and into the bulk. Two bulk modes of the chain are annotated in the figure, and their mode plots in the xz-plane are shown in the bottom image. We see that they are not localized like the edge modes are. We also note in the open-boundary spectra that there is a slight separation between the last two edge modes and the first four. This is what is shown in the zoomed region of the right plot of Fig. 3. Those four edge modes that are on the same line are consistent with the four degenerate transverse bands, while the other two are consistent with the two longitudinal bands. This offset is not explained with nearest-neighbour interactions (green x's), but is instead explained by long-range interactions (blue crosses). In general, we see good qualitative agreement with the coupled mode theory and the full numerical simulations.

Cuboid

Next, we move on to investigating nanostructures with a different shape, particularly one that would break spherical symmetry and also does not have an analytic solution. We chose nanocuboid arrays for simplicity. The cuboids were 14 nm × 14 nm × 25 nm, with curved corners consisting of 2 nm radius of curvature. In this section, we present the semi-analytic tight-binding results for the cuboid and compare them with the Comsol predictions. We first present the results for periodic boundary conditions (PBCs), then we present the results for the open-boundary conditions.

We start with the infinite nanocuboid SSH chain, with a unit cell of length 50 nm, consisting of two nanocuboids separated from each other by a length $\beta a/2$, where again, $\beta = 1$ corresponds to the case where the particles are evenly spaced, while $\beta \neq 0$ corresponds to the case where the particles are dimerized.

We start with PBCs. Fig. 5 shows the results for the real eigenvalues for the cuboid nanoparticles embedded in glass for the two lowest bands. The imaginary ones are depicted in Fig. S6. Again, the red points are numerical results, the green curve is the quasinormal coupled-mode theory (tight-binding model) with nearest-neighbour coupling, and the dashed blue curve is the quasinormal coupled-mode theory with long-range couplings included (up to 12 unit cells). The coupled-mode theory and full numerical simulations are qualitatively in good agreement, but they disagree qualitatively, especially in comparison to the spheres. We note that we only used one mode from a single particle, and expect to see better agreement upon including more modes. Nevertheless, we see improvements when we incorporate long-range couplings, highlighting the relevance of long-range behaviour. Fig. 6, top-right, depicts higher bands, which generally tend to be more

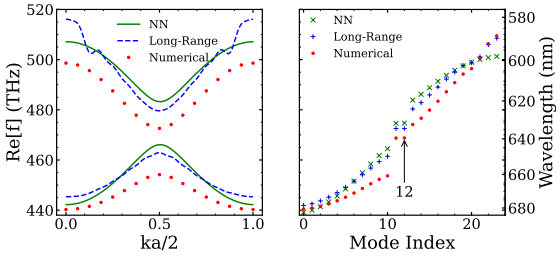


Figure 5: Lowest two bands for the real part of the bandstructure (left) and open-boundary (right) for cuboid nanoparticles when $\beta = 1.2$. Numerical simulations are red points. The green bold curve and the broken blue curve are the nearest-neighbour and long-range coupling. Green x's and blue crosses correspond to long-range coupling.

closely spaced. In order to see reasonably sized topological effects there, we have to open the band gap more by making $\beta = 1.3$. The result is bands that have a different shape than the lower bands. The character of the associated modes is also different from that of the lower bands, having a more asymmetric nature in the xy-plane, compared with the lower bands that have a more symmetric nature in the xy-plane. This is reminiscent of the asymmetric nature of the p-like orbitals and the symmetric nature of the s-like orbitals of molecular systems. We can therefore say that the lower bands have a p-like nature, while the upper bands presented have a p-like nature in the xy-plane. Meanwhile, in the yz-plane, they both have a p-like nature, whereas in the xz-plane, the nature in this molecular orbital system is ambiguous.

Next, we consider a finite chain of 24 nanocuboids, with open-boundary conditions. We take the nanocuboid sizes to be the same as the periodic boundary case. We consider two particle separations for the dimer: $\beta = 0.8$ and $\beta = 1.2$. As with the nanospheres, for $\beta = 0.8$, each particle appears in a pair along the chain, whereas for $\beta = 1.2$, there are two "unpaired" particles at each end.

Fig. 2, right, shows the edge modes in the non-trivial setup. The green x's and blue crosses in the right plot represent the nearest-neighbour and long-range coupling, respectively. There are two edge modes in the lower bands; one of which (mode 12) is depicted in the mode plot (bottom). We also note four more edge modes in Fig. 6, one (mode 15) of which is depicted in the mode plot. We note that the modal character of the edge mode presented in the bottom modes of Fig. 6 do not have the traditional exponential-like decay character. In addition, these edge modes appear when the Zak phase is π and vanish when the Zak phase is

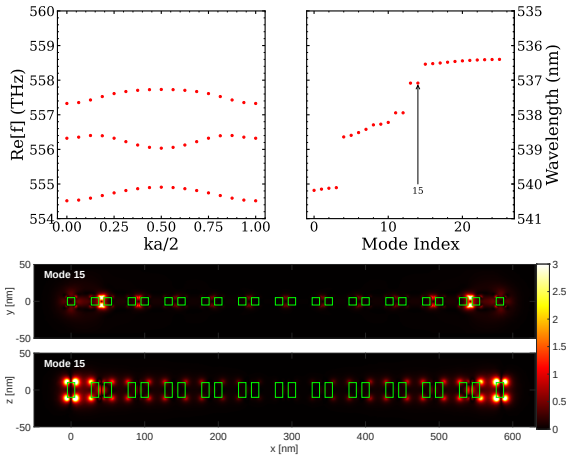


Figure 6: Numerical results for the real part of higher bands of a cuboid at $\beta = 1.3$ for periodic (left) and open boundary (right). The mode plots (bottom) are scaled by 10^9 V/m.

0.

Modal Decomposition

Each mode on a site is a hybridization of multiple basis modes. Finding which modes dominate the hybridized modes is useful to gain physical insight into the nature of the hybridized modes. In addition, it gives insight into how many modes we should keep when computing the bandstructure for the coupled-mode theory. This modal decomposition can have value in the design of devices where a particular resonance is relevant to the operation of the device. One such example is in sensing. The usual strategies to study nanoresonator array setups, such as the dipole approximation, make it difficult to perform a modal decomposition. In the work presented in this paper, we apply the coupled-mode theory to evaluate as many modal contributions as we wish (up to 15 modes in our case): all first-order, second-order, as well as third-order modes. The basis modes used are the resonances of a single particle in a unit cell. This is computed using eq. 21. We take the single-particle result, shifting the eigenmode to the same location of the particle of interest on the finite chain, where at each step, we use eq. 21.

Fig. 7 shows the results for the modal decomposition of the spherical nanoparticle array. The left plot is for one of the transverse edge modes of the chain, while the right plot is for one of the longitudinal edge modes. We clearly see first-order modes dominating, which have a dipole-like character. We also see noticeable contributions from second- and third-order modes, which have a quadrupolar and octupolar character. In the right plot, we note more contributions from octupoles. We note that the longitudinal modes are associated with the lowest and highest of the six bands, so some higher modes should indeed have influence. Modes of a quadrupolar and octupolar character have never been explicitly demonstrated, to the best of our knowledge, to contribute to plasmonic edge modes in plasmonic simulations. As we go further up the basis modes, we see they contribute less compared with the lower modes. This is why three modes were good enough to get high accuracy.

We look at the values associated with the colors of the colormap as a function of the site numbers and fit an exponential decay function to extract the decay length of the mode. We present this only for the transverse modes. This is depicted in the bottom plot of Fig. 7, where the blue boxes are first-order (sum of 'dipole'-like contributions), orange triangles are second-order (sum of 'quadrupole'-like contributions), and green diamonds are third-order (sum of 'octupole'-like contributions) modes. We note the longer decay length ξ, of the first-order mode, indicating its slower decay into the chain. A longer decay length seems directly related to the dominance of the mode. This is done for odd-numbered sites, but the results are the same for even-numbered sites. We also show the modal decomposition of the cuboid chain, depicted in

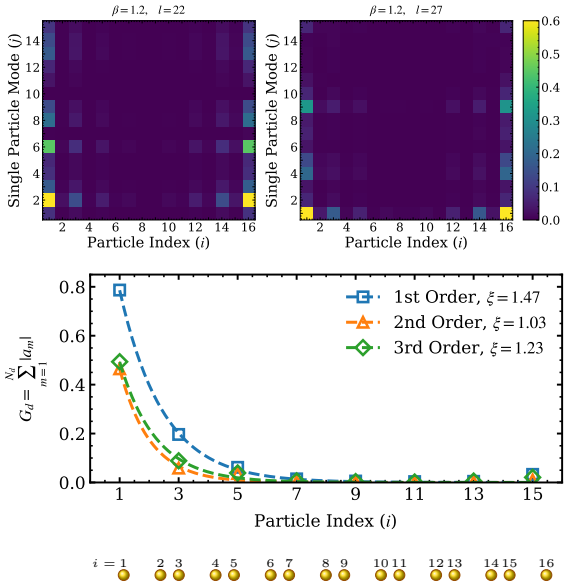


Figure 7: Modal decomposition (top) of a finite non-trivial chain of spheres (bottom) for two of six edge modes: transverse (left) and longitudinal (right) for $\beta = 1.2$. The particle index for the finite chain is i, while the basis (single-particle) modes are labeled j.

Fig. 8. We see that the first-order mode contribution is dominant at the edge. We in fact see that modes 12-15 seem to be influential enough that they should not be neglected. This is why our results did not agree well with simulations of the entire system; we only used one single-particle mode for cuboids.

The panel at the bottom of Fig. 8 shows the decay of the first-order (blue box), second-order (orange triangle), and third-order (green diamonds) contributions of one of the edge modes. We fit to the odd-numbered particles. The longer decay length and higher-than-normal weight on the second site for the dominant mode indicate the influence of long-range coupling. This has influence on bulk-boundary correspondence, which we will discuss next.

Bulk-Boundary Correspondence

We explore whether bulk-boundary correspondence is observed in both the spherical and cuboid nanoparticle chains.

Boundary

In order to do this, we perform a sweep over β and explicitly show the evolution of the eigenfrequencies as a function of β. We performed a sweep of β from 0.8 to 1.2 in steps of 0.05. Since β determines the separation of the particles, and since this in turn determines the coupling strength, sweeping β is equivalent to sweeping over the coupling strength. We do this for the lower bands of the cuboid and the six bands of the spherical nanoparticle chain. The results for the spheres and cuboids are shown in the left and right plots in Fig. 9, respectively. The red markers represent the edge modes, and the black markers represent the bulk modes. Due to the fact that the chain is finite in length and relatively short (16 particles for spheres and 24 for cuboids), the spectral separation between the edge modes and the bulk isn't always distinct, especially for β values closer to 1.0. Typically, in the nearest-neighbour tight-binding model for the SSH chain, the edge modes line up along a horizontal line for a β-sweep. In Fig. 9 (right), we see this is not the case. There is a tilt away from the horizontal. This is an effect due to long-range couplings and is explained in the theory

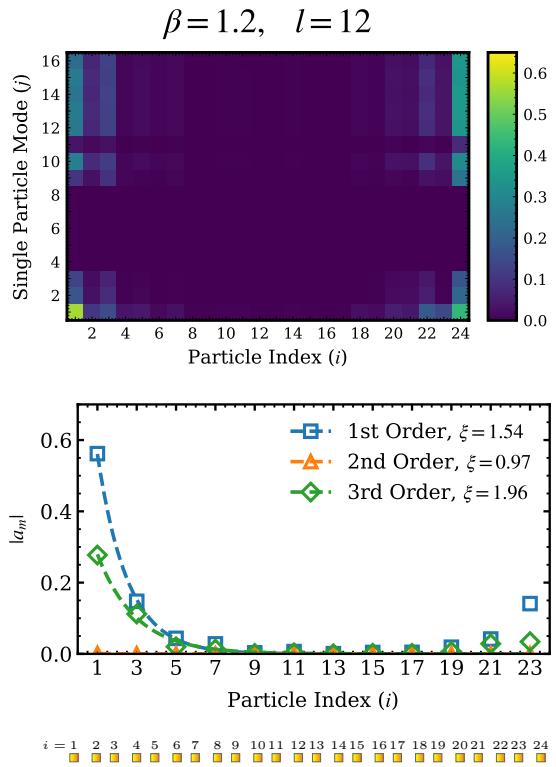


Figure 8: Modal decomposition (top) of finite non-trivial chain of cuboids (front-view at bottom) for one (1) of two edge modes for $\beta = 1.2$. The particle index for the finite chain is i, while the basis (single-particle) modes are labeled j.

paper. Next, we check to see if the bulk (bandstructure) has information regarding the presence (absence) of these edge states. We do this by computing the Zak phase from the bandstructure obtained from the simulations of the full system in Comsol.

Zak Phase

We use the QNMs of the associated bulk bands, computed by the Comsol simulations of the full structure, to compute the Zak phase. We see in Fig. 10 the computed Zak phase in different regions: trivial ($\beta < 1$) and non-trivial ($\beta > 1$) for the cuboids (right) and spheres (left). We see that the results for all bands are consistent with each other. In the non-trivial regime, the results are as expected ($\gamma = \pi$), and in the trivial regime, the results are also as expected $\gamma = 0$. There is a jump at $\beta = 1$, consistent with what is expected. All these behaviours are consistent with intuition from electronic topological materials. The results of the β-sweep and the Zak phase computations indicate the presence of a bulk-boundary correspondence. We note something important here. Chiral symmetry, when present, is responsible for the localized edge modes that we observe in the non-trivial state. Inversion symmetry is what gives the Zak phase π (0) for the non-trivial (trivial) state. We have shown that long-range coupling is relevant in our work. Long-range coupling breaks chiral symmetry in our work, while inversion symmetry remains present. This means that we should not see localized edge modes in the non-trivial state, even when the Zak phase gives π. Our results clearly show the presence of localized edge modes, when the Zak phase is π, which is consistent with a bulk-boundary correspondence. This is a surprising result, as breaking chiral symmetry should break bulk-boundary correspondence.

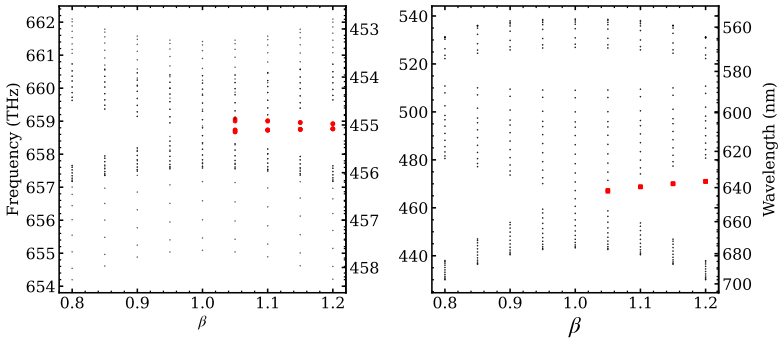


Figure 9: β sweep of a finite chain of spheres (left) and cuboid (right) from full structure simulations. The edge modes are the red markers. The edge (bulk) modes are shown in red (black).

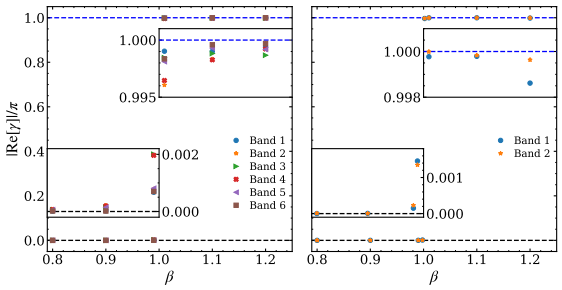


Figure 10: Bulk data (bandstructure) used to compute non-Hermitian Zak phase via Wilson loop. This plot is the absolute value of the real part of the Zak phase for a sphere (left) and a cuboid (right) as a function of the distance scaling parameter, β. When $\beta < 1$, it is 0 and if $\beta > 1$, it gives π. This is consistent with the open-boundary predictions of edge modes explained earlier.

Summary and Discussion

We have introduced a computational approach that is based on a shape-sensitive framework, which we used to model topological phases in inherently non-hermitian plasmonic systems. We simulated full plasmonic arrays and compared the results with a tight-binding description. We applied it to two SSH chains of Drude model spheres in air and cuboids in glass, demonstrating the presence of a bulk-boundary correspondence, even though it is expected that chiral symmetry is broken by long-range couplings.

For the spherical chain, our method predicts six topological edge states, originating from three pairs of transverse and longitudinal bands. The excellent agreement between the tight-binding reconstruction and full lattice simulations for the structure with the spherical nanoresonators highlights the strong potential of our approach and its scalability to do more complex geometries. In contrast, the cuboid chain supports only two edge states and exhibits good qualitative agreement between the tight-binding model and simulations of the entire array, which we directly attribute to our truncation of the coupling range and the number of modes used in the modal expansion.

We computed the Zak phase for both systems using a non-Hermitian Wilson loop and found values of π in the non-trivial regime and 0 in the trivial regime, fully consistent with nearest-neighbour SSH tight-binding models. Importantly, bulk-boundary behaviour remained even in the presence of long-range interactions. This is a surprising result as chiral symmetry is expected to be broken by long-range coupling, and only inversion symmetry is expected to remain. Typically, this would result in trivial and non-trivial topology revealed through Zak phase calculations, but no edge modes.

A key strength of our framework is its ability to perform a modal decomposition of both bulk and edge states. Beyond dominant dipolar-like contributions, we directly resolve higher-order modal fingerprints in

the edge modes of spherical chains, an effect not previously demonstrated in simulations of plasmonic SSH systems. This revealed not only which resonances contribute to the edge modes, but also the modal decay lengths.

Our results demonstrate that we can use our method to probe topology in inherently non-Hermitian plasmonic systems. It offers a platform that has the potential to accelerate the discovery of non-Hermitian topology in plasmonic systems, which we envision will one day have important technological applications from sensing to spectroscopy.

Acknowledgements

The authors acknowledge funding from the Max Planck Society Lise Meitner Excellence Program 2.0. We would also like to thank Julius Gohsrich and Flore Kunst for their theoretical contributions and stimulating discussions. We would also like to acknowledge the support of the Max Planck-uOttawa Centre for Extreme and Quantum Photonics.

References

[1] Klitzing, K. v.; Dorda, G.; Pepper, M. *Phys. Rev. Lett.* **1980**, *45*, 494–497.

[2] Hasan, M. Z.; Kane, C. L. *Rev. Mod. Phys.* **2010**, *82*, 3045–3067.

[3] Qi, X.-L.; Zhang, S.-C. *Rev. Mod. Phys.* **2011**, *83*, 1057–1110.

[4] Chiu, C.-K.; Teo, J. C. Y.; Schnyder, A. P.; Ryu, S. *Rev. Mod. Phys.* **2016**, *88*, 035005.

[5] Banerjee, N. et al. *Applied Physics Reviews* **2025**, *12*, 041328.

[6] Gilbert, M. J. *Communications physics* **2021**, *4*, 1–12.

[7] Alù, A.; Daraio, C.; Deymier, P.; Ruzzene, M. *Nature Communications* **2025**, *16*, 5680.

[8] Sahin, H.; Jalil, M. B. A.; Lee, C. H. *APL Electronic Devices* **2025**, *1*, 021503.

[9] Zhang, W.; Cao, W.; Qian, L.; Yuan, H.; Zhang, X. *Nature communications* **2025**, *16*, 198–11.

[10] Ozawa, T.; Price, H. M.; Amo, A.; Goldman, N.; Hafezi, M.; Lu, L.; Rechtsman, M. C.; Schuster, D.; Simon, J.; Zilberberg, O.; Carusotto, I. *Rev. Mod. Phys.* **2019**, *91*, 015006.

[11] Lan, Z.; Chen, M. L.; Gao, F.; Zhang, S.; Sha, W. E. *Reviews in Physics* **2022**, *9*, 100076.

[12] Jalali Mehrabad, M.; Mittal, S.; Hafezi, M. *Phys. Rev. A* **2023**, *108*, 040101.

[13] Rocklin, D. Z.; Zhou, S.; Sun, K.; Mao, X. *Nature Communications* **2015**, *8*.

[14] Kane, C. L.; Lubensky, T. C. *Nat. Phys.* **2014**, *10*, 39–45.

[15] ge Chen, B. G.; Upadhyaya, N.; Vitelli, V. *Proceedings of the National Academy of Sciences* **2014**, *111*, 13004–13009.

[16] Süsstrunk, R.; Huber, S. D. *Science* **2015**, *349*, 47–50.

[17] Bergholtz, E. J.; Budich, J. C.; Kunst, F. K. *Rev. Mod. Phys.* **2021**, *93*, 015005.

[18] Ashida, Y.; Gong, Z.; Ueda, M. *Advances in Physics* **2020**, *69*, 249–435.

[19] Okuma, N.; Sato, M. *Annual Review of Condensed Matter Physics* **2023**, *14*, 83–107.

[20] Yao, S.; Wang, Z. *Phys. Rev. Lett.* **2018**, *121*, 086803.

[21] Xiong, Y. *Journal of Physics Communications* **2018**, *2*, 035043.

[22] Zhang, X.; Zhang, T.; Lu, M.-H.; Chen, Y.-F. *Advances in Physics: X* **2022**, *7*, 2109431.

[23] Gohsrich, J. T.; Banerjee, A.; Kunst, F. K. *Europhysics Letters* **2025**, *150*, 60001.

[24] Meng, H.; Ang, Y. S.; Lee, C. H. *Applied Physics Letters* **2024**, *124*, 060502.

[25] Heiss, W. D. *Journal of Physics A: Mathematical and Theoretical* **2012**, *45*, 444016.

[26] Martinez Alvarez, V. M.; Barrios Vargas, J. E.; Foa Torres, L. E. F. *Phys. Rev. B* **2018**, *97*, 121401(R).

[27] Berry, M. V. *Czechoslovak Journal of Physics* **2004**, *54*, 1039–1047.

[28] Budich, J. C.; Carlstrom, J.; Kunst, F. K.; Bergholtz, E. J. *Phys. Rev. B* **2019**, *99*, 041406(R).

[29] Budich, J. C.; Bergholtz, E. J. *Phys. Rev. Lett.* **2020**, *125*, 180403.

[30] Bandres, M. A.; Wittek, S.; Harari, G.; Parto, M.; Ren, J.; Segev, M.; Christodoulides, D. N.; Khajavikhan, M. *Science* **2018**, *359*, eaar4005.

[31] Wu, B.; Ding, J.; Chen, F. *Nature Communications* **2026**, *17*, 112.

[32] Mao, W.; Zhu, B.; Zhang, Q.; Xu, W.; Jia, D.; Meng, Y.; Wang, C.; Li, F.; Bae, S. H.; Wang, Q. J.; Chong, Y.; Yang, L. *Nature Physics* **2026**, *22*, 1344–1349.

[33] Qin, H.; Lv, W.; Zhang, Z.; Yang, Z.; Li, J.; Li, M.; Li, B.; Zhou, J.; Fleury, R.; Genevet, P.; Song, Q.; Qiu, C.-W. *Nature Reviews Materials* **2026**, *11*, 453–468.

[34] Boriskina, S. V.; Cooper, T. A.; Zeng, L.; Ni, G.; Tong, J. K.; Tsurimaki, Y.; Huang, Y.; Meroueh, L.; Mahan, G.; Chen, G. *Adv. Opt. Photon.* **2017**, *9*, 775–827.

[35] Kant, K. et al. *Nanoscale Horiz.* **2024**, *9*, 2085–2166.

[36] Mayer, K. M.; Hafner, J. H. *Chemical Reviews* **2011**, *111*, 3828–3857.

[37] Zhao, J.; Zhang, X.; Yonzon, C. R.; Haes, A. J.; Van Duyne, R. P. *Nanomedicine* **2006**, *1*, 219–228.

[38] Petryayeva, E.; Krull, U. J. *Analytica Chimica Acta* **2011**, *706*, 8–24.

[39] Thioune, N.; Lidgi-Guigui, N.; Cottat, M.; Gabudean, A.-M.; Focsan, M.; Astilean, S.; Lamy de la Chapelle, M. *Gold Bulletin* **2013**, *46*, 275–281.

[40] Sarkawi, M.; Sharma, R.; Venkatesan, G.; Ahuja, T.; Panda, S.; Prakash, J.; Kuramitz, H.; devi, K. S. *Talanta Open* **2025**, *12*, 100554.

[41] Chirco, A.; Meacci, E.; Margheri, G. *ACS Bio Med Chem Au* **2025**, *5*, 922–946.

[42] Bahadori, A.; Oddershede, L. B.; Bendix, P. M. *Nano Research* **2017**, *10*, 2034–2045.

[43] Ali, A.; Kang, J. H.; Seo, J. H.; Walker, B. *Advanced Engineering Materials* **2020**, *22*, 1900976.

[44] Huang, X.; El-Sayed, M. A. *Alexandria Journal of Medicine* **2011**, *47*, 1–9.

[45] Wu, Y.; Ali, M.; Dong, B.; Han, T.; Chen, K.; Chen, J.; Tang, Y.; Fang, N.; Wang, F.; El-Sayed, M. *ACS Nano* **2018**, *12*.

[46] Langer, J. et al. *ACS Nano* **2020**, *14*, 28–117.

[47] Ling, C. W.; Xiao, M.; Chan, C. T.; Yu, S. F.; Fung, K. H. *Opt. Express* **2015**, *23*, 2021–2031.

[48] Pocock, S.; Xiao, X.; Huidobro, P.; Giannini, V. *ACS Photonics* **2018**, *5*.

[49] Honari-Latifpour, M.; Yousefi, L. *Nanophotonics* **2019**, *8*, 799–806.

[50] Buendia, A.; Sanchez-Gil, J. A.; Giannini, V. *ACS Photonics* **2023**, *10*, 464–474.

[51] Wu, H.; Wei, G.; Liu, Z.; Xiao, J.-J. *Phys. Rev. Appl.* **2025**, *23*, 024032.

[52] Miranda, D. A.; Bludov, Y. V.; Asger Mortensen, N.; Peres, N. M. R. *Journal of Optics* **2024**, *26*, 125001.

[53] Sinev, I. S.; Mukhin, I. S.; Slobozhanyuk, A. P.; Poddubny, A. N.; Miroshnichenko, A. E.; Samusev, A. K.; Kivshar, Y. S. *Nanoscale* **2015**, *7*, 11904–11908.

[54] Moritake, Y.; Ono, M.; Notomi, M. *Nanophotonics* **2022**, *11*, 2183–2189.

[55] Park, S. Y.; Stroud, D. *Phys. Rev. B* **2004**, *69*, 125418.

[56] Downing, C. A.; Mariani, E.; Weick, G. *Journal of Physics: Condensed Matter* **2017**, *30*, 025301.

[57] Koenderink, A. F.; Polman, A. *Phys. Rev. B* **2006**, *74*, 033402.

[58] Weber, W. H.; Ford, G. W. *Phys. Rev. B* **2004**, *70*, 125429.

[59] Downing, C. A.; Weick, G. *Phys. Rev. B* **2017**, *95*, 125426.

[60] Yan, W.; Faggiani, R.; Lalanne, P. *Phys. Rev. B* **2018**, *97*, 205422.

[61] Lalanne, P. et al. *J. Opt. Soc. Am. A* **2019**, *36*, 686–704.

[62] Sauvan, C.; Wu, T.; Zarouf, R.; Muljarov, E. A.; Lalanne, P. *Opt. Express* **2022**, *30*, 6846–6885.

[63] Gigli, C.; Wu, T.; Marino, G.; Borne, A.; Leo, G.; Lalanne, P. *ACS Photonics* **2020**, *7*, 1197–1205.

[64] Su, W. P.; Schrieffer, J. R.; Heeger, A. J. *Phys. Rev. Lett.* **1979**, *42*, 1698–1701.

[65] Gohsrich, J. T.; Israel, N. S.; Ramunno, L.; Kunst, F. K. Manuscript in preparation.

[66] COMSOL AB COMSOL Multiphysics. `https://comsol.com`, 2026; Stockholm, Sweden. Finite Element Analysis Simulation Software.

[67] Wu, T.; Arrivault, D.; Yan, W.; Lalanne, P. *Computer Physics Communications* **2023**, *284*, 108627.

[68] Tao, C.; Zhu, J.; Zhong, Y.; Liu, H. *Phys. Rev. B* **2020**, *102*, 045430.

[69] Bisharat, D.; Davis, R.; Zhou, Y.; Bandaru, P.; Sievenpiper, D. *IEEE Antennas and Propagation Magazine* **2021**, *63*, 112–124.

Supplementary: A Modal Analysis of Topological Plasmonic Nanoresonator Chains

Semi-Analytic Numerical Methods

While simulating the full structure can tell us whether or not our system is topologically trivial, some physical insights are more difficult to extract from first principles. For example, coupling strengths can only be obtained by fitting. We go beyond this, and compute the coupling from first-principles. In order to do this, we applied tools developed in[68] for our work, which, to the best of our knowledge, has not been done before. We start by simulating one particle in COMSOL. We extract a few eigenmodes and eigenfrequencies. We compute the coupling for each mode between each particle using the eigenmodes and eigenfrequencies of the single particle. We then use the coupling and the eigenfrequencies to build out an eigenvalue problem that has periodic boundary conditions, or open-boundary conditions. We then solve this problem to get the eigenfrequencies of the full system, which in the periodic case, gives us the bandstructure. The eigenvectors of this eigenvalue problem are modal expansion coefficients for reconstructing the full modal distribution of the system using the eigenmodes of a single particle. Visualizing the distribution of the modal coefficients alone however is sufficient to give us much of the information we need to tell if an edge mode is present or not. In this work, there are three key assumptions:

1. We can truncate the physical space using absorbing boundaries, perfectly matched layers (PML) and get physically meaningful results for the nearfield.
2. In principle, there are an infinite number of modes - including physical and numerical (PML) modes which are needed to form a complete basis. We assume that truncating to a finite number of modes, particularly, a finite number of physical modes and neglecting the PML modes altogether is sufficient to get physically meaningful results pertinent to our problem.
3. We can truncate the coupling to finite range and reconstruct the physically meaningful results relevent to us.

We use these assumptions to proceed in using the formalism to compute the coupling as a function of distance for the two different geometries, giving us insight of how the coupling is influenced by the geometry of the structure. In order to compute the coupling, we simulate a single particle only once. We then build an algorithm to shift the position of the particle and its modes once computed to the location that its neighbour is present at. We are only interested in the nearfield effects of our system. The algorithm we use is as follows

Algorithm S0.1 *Using a Gauss-Legendre approach to integration, we*

1. *Perform only one simulation of a single particle with an absorbing boundary,*
2. *extract pertinent eigenmodes and eigenfrequencies from one-particle simulation,*
3. *extract the Gauss points in the resonator at the location of the particle in the simulation,*
4. *extract the fields at those Gauss points,*
5. *shift those Gauss points to the position where the next particle is supposed to be,*
6. *extract the field at those points,*
7. *repeat steps 5 and 6 for as many locations as we need to compute the coupling for,*
8. *Use the results from steps 3 to 7 to compute the coupling with Eq. 18.*

Once we have the eigenmodes for the periodic system, we can compute the Zak phase. The Wilson loop approach was used. We sketch a quick derivation of it below.

Wilson Loop

As in,[69] we can do a Taylor series approximation of $|\psi_{nk}(\vec{r})\rangle$ as

$$|\tilde{\psi}_{n(k+\delta k)}(\vec{r})\rangle = |\tilde{\psi}_{nk}(\vec{r})\rangle + \nabla_k|\tilde{\psi}_{nk}(\vec{r})\rangle\delta k + \cdots \tag{S1}$$

If $\delta k << 1$, then we can keep up to the first-order term and get

$$|\tilde{\psi}_{n(k+\delta k)}(\vec{r})\rangle \approx |\tilde{\psi}_{nk}(\vec{r})\rangle + \nabla_k|\tilde{\psi}_{nk}(\vec{r})\rangle\delta k \tag{S2}$$

Now, we have

$$\begin{aligned}\langle\tilde{\varphi}_{nk}(\vec{r})|\tilde{\psi}_{n(k+\delta k)}(\vec{r})\rangle &\approx \langle\tilde{\varphi}_{nk}(\vec{r})|\tilde{\psi}_{nk}(\vec{r})\rangle \\ &+ \langle\tilde{\varphi}_{nk}(\vec{r})|\nabla_k|\tilde{\psi}_{nk}(\vec{r})\rangle\delta k \\ &= 1 + \langle\tilde{\varphi}_{nk}(\vec{r})|\nabla_k|\tilde{\psi}_{nk}(\vec{r})\rangle\delta k\end{aligned} \tag{S3}$$

This is the first order Taylor series approximation for e^{α}, where $\delta\alpha = \langle\tilde{\varphi}_{nk}(\vec{r})|\nabla_k|\tilde{\psi}_{nk}(\vec{r})\rangle\delta k$. Note that $\langle\tilde{\varphi}_{nk}(\vec{r})|\nabla_k|\tilde{\psi}_{nk}(\vec{r})\rangle = -iA_k$, so we have that $\delta\alpha = -iA_k\delta k$. This gives

$$\langle\tilde{\varphi}_{nk}(\vec{r})|\tilde{\psi}_{n(k+\delta k)}(\vec{r})\rangle \approx e^{-iA_k\delta k} \tag{S4}$$

We break the loop into very small segments based on the number of k-points we used to compute our bandstructure. We can rewrite as

$$\langle\tilde{\varphi}_{nk_i}(\vec{r})|\tilde{\psi}_{nk_{i+1}}(\vec{r})\rangle \approx e^{-iA_{k_i}\delta k} \tag{S5}$$

where we look at $k \in [k_i, k_{i+1}]$. Just as in the Hermitian case, we can do

$$M_n = \prod_{i=1}^{N-1}\langle\tilde{\varphi}_{nk_i}(\vec{r})|\tilde{\psi}_{nk_{i+1}}(\vec{r})\rangle\delta k \tag{S6a}$$

We get the Berry phase from this by

$$\phi_n = \sum_{i=1}^{N-1} -i\log\left[\langle\tilde{\varphi}_{nk_i}(\vec{r})|\tilde{\psi}_{nk_{i+1}}(\vec{r})\rangle\delta k\right] \tag{S6b}$$

Convergence Tests

In COMSOL, our minimum and maximum mesh settings were chosen to be around 1 nm within the resonator region, as we found this sufficient for this problem. For the spheres, our upper and lower mesh sizes were 1.0 nm and 0.5 nm, respectively, inside the resonators. The choice of the mesh for the spheres gives an absolute relative percent error of 0.02% and 0.01% for the real part. The results for this are shown in the top plot of Fig. S1. In the case of the imaginary part, we have 0.26% and 0.13% respectively for the absolute relative error. The results for this are shown in the bottom plot of Fig. S1.

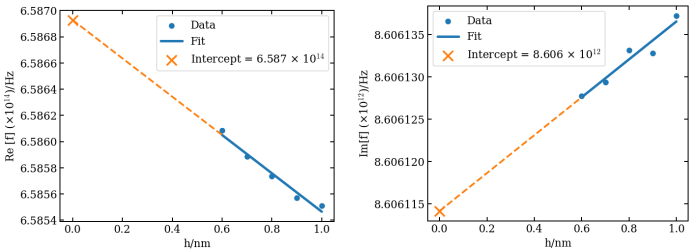


Figure S1: Convergence test results for the real part (left) of eigenfrequency for the sphere and the imaginary part (right) as a function of mesh size.

For the cuboids, our upper and lower meshes were 1.5 nm and 1.0 nm, respectively, inside the resonator. In the case of the real part, this gives an absolute relative error of 0.02% and 0.01%, respectively. The results for this are shown in the top plot of Fig. S2. In the case of the imaginary part, this gives 0.004% and 0.0027% respectively for the imaginary part. The results for this are shown in the bottom plot of Fig. S2.

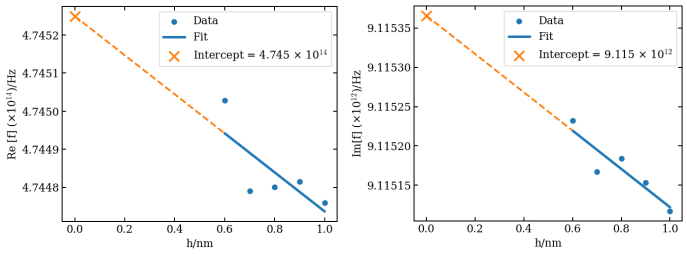


Figure S2: Convergence test results for the real part (left) of eigenfrequency for the cuboid and the imaginary part (right) as a function of mesh size.

Modes

We show examples of different modes for a single sphere. One can compare these with molecular orbitals and notice some similarities. Fig. S3 shows the first three modes, while Fig. S4 shows the next three modes. These are shown in the xz-plane for each component of $\tilde{E}$: $\tilde{E}_x, \tilde{E}_y$ and $\tilde{E}_z$.

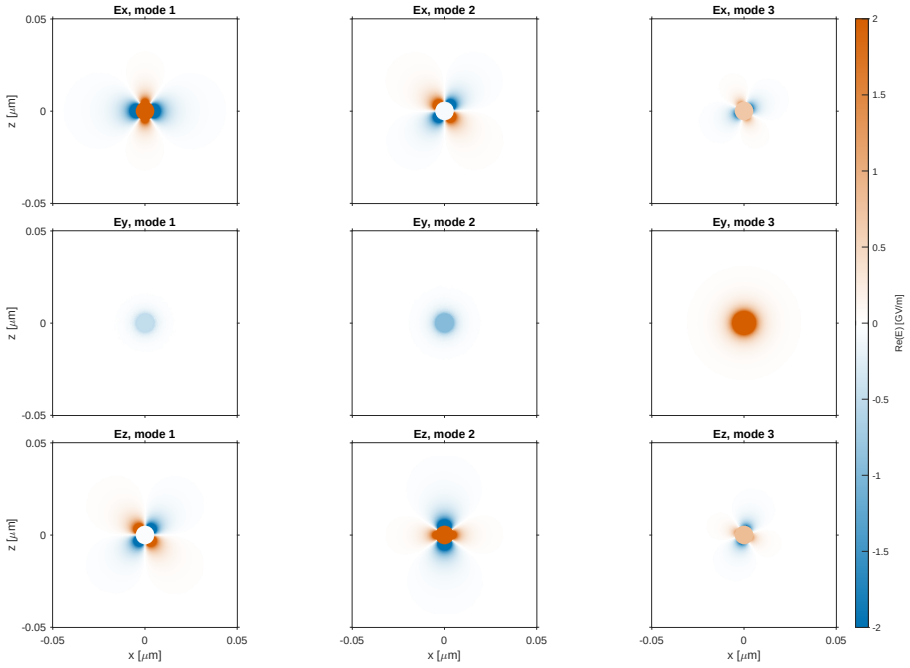


Figure S3: The first three modes in the xz-plane for a single sphere in the chain. E_x, E_y and E_z are shown from top to bottom, while modes one to three go from left to right.

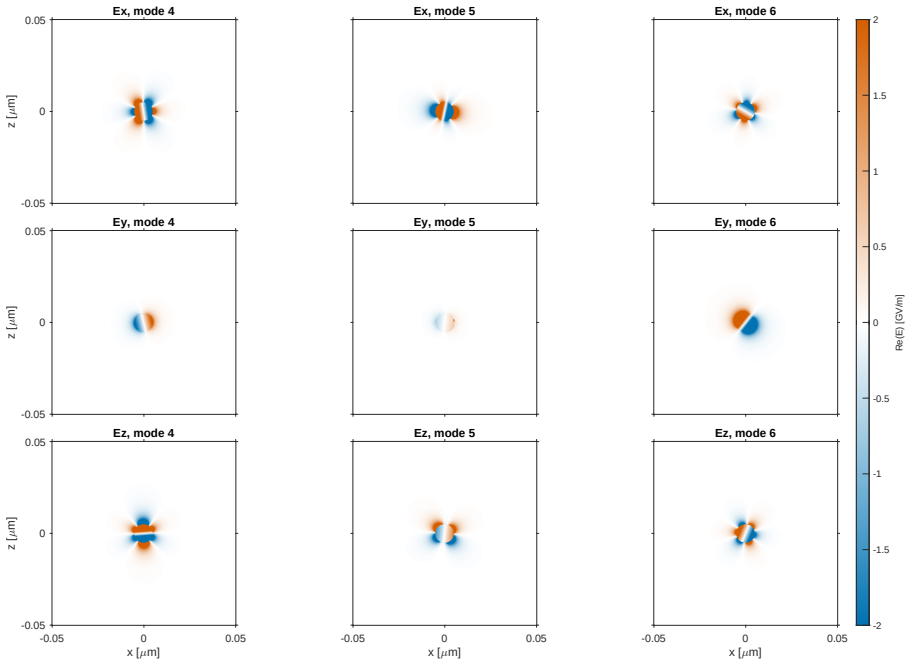


Figure S4: The next three modes in the xz-plane for a single sphere in the chain. E_x, E_y and E_z are shown from top to bottom, while modes one to three go from left to right.

Imaginary Part of Eigenfrequencies

We show the imaginary parts of the eigenfrequencies for the periodic boundary and finite chain case of the spherical particle chain are depicted in Fig. S5 and for the cuboid chain is in Fig. S6.

Imaginary Part of Zak Phase

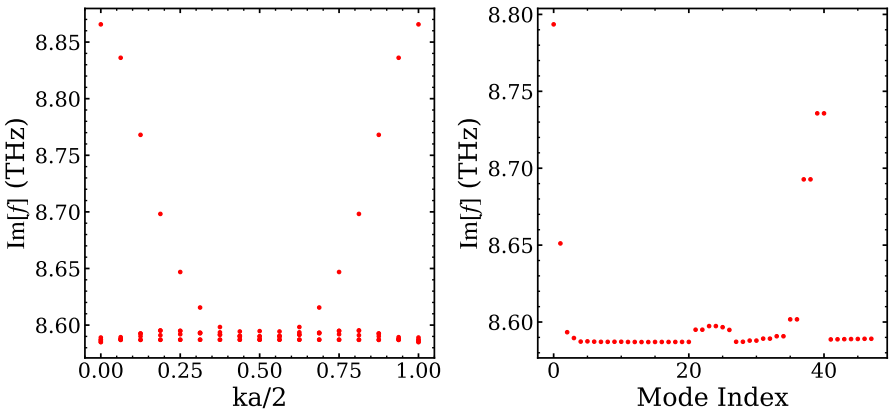


Figure S5: The first three modes in the xz-plane for a single sphere in the chain. E_x, E_y and E_z are shown from top to bottom, while modes one to three go from left to right.

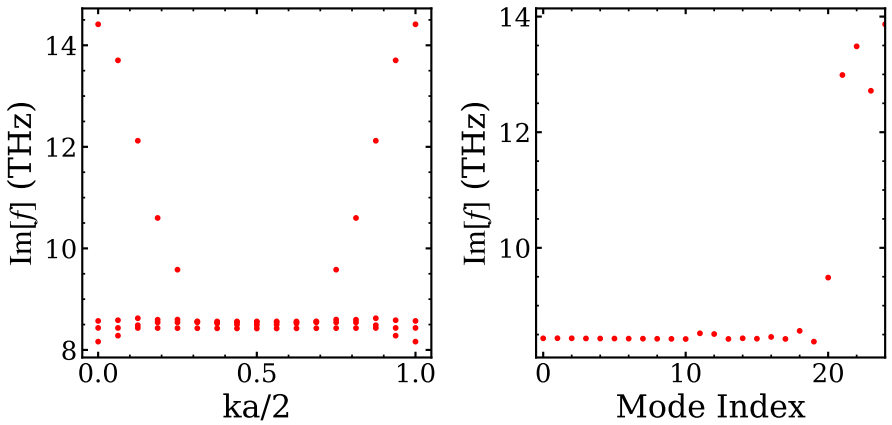


Figure S6: The next three modes in the xz-plane for a single sphere in the chain. E_x, E_y and E_z are shown from top to bottom, while modes one to three go from left to right.

7

Metallic Nanoresonators on a Substrate

One of the main motivations, beyond the basic physics, for developing the tools for topological plasmonics in this thesis was to be able to apply them to design devices whose fundamental operation exploits the exotic nature of topological states. Being able to explore topological plasmonics of nanoresonators on a substrate is one step in this direction. Exploring topology in plasmonic nanoresonators on a substrate presents various challenges. In the systems we have studied for this thesis, the background is homogeneous, so the interactions are relatively simple, and we only need to consider how this background modifies

the electromagnetic interactions of the particles on the chain. Traditional Hamiltonians from condensed matter physics, which inspired the topological work in plasmonics, are suited for this. When a substrate is present, we have an inhomogeneous background, as we now have the substrate as well as the region above the substrate (called a superstrate) modifying the interactions in their own way, making the electromagnetic interactions between the particles much more complex. Though this has the potential to introduce rich physics in the process, making the structures particularly fascinating for probing topological phenomena in plasmonics, and even more potential for interesting non-Hermitian topological phenomena, it also makes it much less tractable for traditional Hamiltonian approaches from condensed matter physics. As a result, few have tread this path [197, 198]. The approach we use for the topology work in this thesis, which is outlined in section 3.1.2, can handle resonators on a substrate [199]. As a result, we explore topological plasmonics of metallic resonators on a substrate using this method. Our system is an SSH chain.

7.1 Method

We used the Drude model parameters for eq. 2.69 in section 2.7. We have a rounded cuboid of dimensions 100 nm by 46.9 nm by 50 nm arranged periodically with a period of 400 nm. Each particle has a thin surrounding shell to improve the simulation's accuracy. We have air $\epsilon \approx 1.0$ as the superstrate and glass $\epsilon \approx 2.25$ as the substrate. A setup of what the system looks like in one configuration as a finite SSH chain is depicted in Fig. 7.1. Each particle in the unit cell, as in previous studies, is separated by a distance $\beta a/2$,

where β is a constant we use to control the distance. As with the other systems studied,

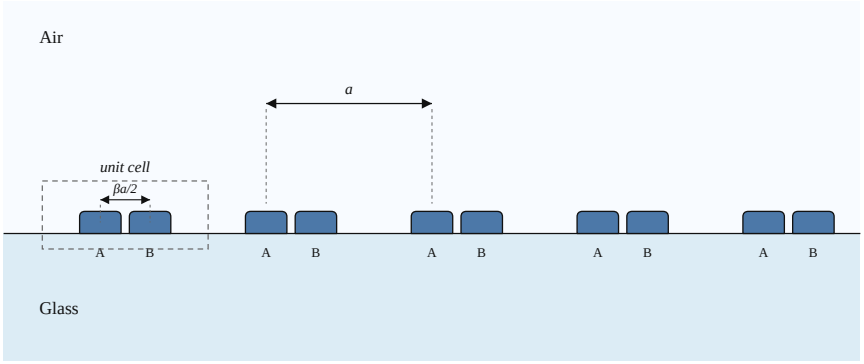


Figure 7.1: A schematic of a 1D SSH chain of plasmonic cuboids on a glass substrate. The superstrate is air and the metallic particles are modeled by the Drude model. The unit cell depicted consists of two particles separated by a distance $\beta a/2$, where a is the lattice constant and β is a distance constant that controls the distance.

we simulate with FEM using the COMSOL eigenmode solver. In order to perform these initial computations, we use extremely coarse mesh settings in COMSOL, but improvement on this will include improving the mesh. The results presented here are for a finite chain, and thus we surround our structure with PMLs, similar to how it was done in chapter 6. Here, we use a PML thickness of 55 nm. We present preliminary results for full lattice simulations.

7.1.1 Some Preliminary Results

We performed the initial study for $\beta = 1.3$ and $\beta = 0.7$, which correspond to the topologically non-trivial and trivial cases, respectively. The results for the open-boundary spectra are presented in Fig. 7.2 for the two configurations. In the non-trivial setup, we see two

edge modes present and clearly separated from the bulk, hinting that this system may very well preserve bulk-boundary correspondence. In order to demonstrate this, we must

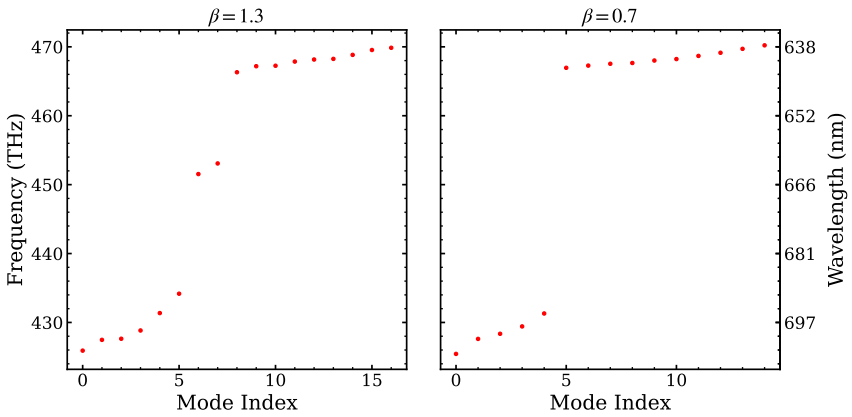


Figure 7.2: Preliminary open-boundary spectra for cuboids on a substrate for two configurations: non-trivial, with $\beta = 1.3$ (left) and trivial, with $\beta = 0.7$ (right). Two edge modes are clearly present in the gap for the non-trivial case.

compute the Zak phase, just as is done in chapter 6, from the bulk bands, which have not yet been computed for these specific configurations.

8

Ideas for Future Work

We present some work in relation to topological plasmonics that is currently in development. They build upon the main work of the thesis and will be further developed at a later date. Much of this work is theoretical, but there are some computational tasks as well. We first present work in which we sketch a derivation of non-Hermitian Berry phases through adiabatic evolution of plasmonic quasinormal modes. This is presented in Section 8.1. We then sketch a geometric interpretation of non-Hermitian Berry phases in the context of adiabatic evolution. This is presented in Section 8.2.

8.1 Non-Hermitian Berry Phase from QNMs

We follow [200] to derive the Berry phase in the context of quasinormal modes. We start by letting $\vec{k}(t)$ be a smooth parametrization of a closed path, C, in the Brillouin zone. We write the overall mode as a linear combination of the quasinormal modes

$$|\tilde{\Psi}(\vec{r},t)\rangle = \sum_n \beta_n(t)\,|\tilde{\psi}_n(\vec{r},\vec{k}(t))\rangle\,, \tag{8.1}$$

where $\beta_n(t)$ are the coefficients of the QNMs, $|\tilde{\psi}_n(\vec{r},\vec{k}(t))\rangle$. We note that in the literature, we normally write $|\tilde{\psi}_n(\vec{r},\vec{k}(t))\rangle = |\tilde{\psi}_{n\vec{k}}(\vec{r})\rangle$ so that we have

$$|\tilde{\Psi}(\vec{r},t)\rangle = \sum_n \beta_n(t)\,|\tilde{\psi}_{n\vec{k}}(\vec{r})\rangle\,. \tag{8.2}$$

We recall that

$$\hat{\mathcal{H}}\,|\tilde{\Psi}(\vec{r},t)\rangle = i\frac{d\,|\tilde{\Psi}(\vec{r},t)\rangle}{dt}. \tag{8.3}$$

This gives the result that

$$\langle\tilde{\varphi}_{m\vec{k}}(\vec{r})|\,\frac{d}{dt}\,|\tilde{\Psi}(\vec{r},t)\rangle = -i\tilde{\omega}_m(t)\beta_m(t). \tag{8.4}$$

Now, we take the time derivative and project the left eigenstate onto the result. That is, we do

$$\frac{d\left|\tilde{\Psi}(\vec{r},t)\right\rangle}{dt}=\sum_{n}\left(\left|\tilde{\psi}_{n\vec{k}}(\vec{r})\right\rangle\frac{d\beta_{n}(t)}{dt}+\beta_{n}(t)\frac{d\left|\tilde{\psi}_{n\vec{k}}(\vec{r})\right\rangle}{dt}\right), \tag{8.5}$$

then

$$-i\tilde{\omega}_{m}(t)\beta_{m}(t)=\frac{d\beta_{m}(t)}{dt}+\sum_{n\neq m}\beta_{n}(t)\left\langle\tilde{\varphi}_{m\vec{k}}(\vec{r})\right|\nabla_{\vec{k}}\left|\tilde{\psi}_{n\vec{k}}(\vec{r})\right\rangle\cdot\frac{d\vec{k}}{dt}. \tag{8.6}$$

In the weak coupling regime, where the coupling is constant (i.e. $\kappa(\omega)\approx\kappa$), the solution is

$$\beta_{n}(t)=\beta_{n}(0)e^{-i\int_{0}^{t}\tilde{\omega}_{n}(t')dt'+i\alpha_{n}(t)}, \tag{8.7}$$

where

$$\alpha_{n}=i\oint\left\langle\tilde{\varphi}_{m\vec{k}}(\vec{r})\right|\nabla_{\vec{k}}\left|\tilde{\psi}_{n\vec{k}}(\vec{r})\right\rangle\cdot d\vec{k}, \tag{8.8}$$

α_n is the geometric phase. Our parameter space consists of the k-vector and thus the state vectors used are the bloch states. Now, we must ask, under what conditions is such a system adiabatic? In order to explore this, we can start with the ansatz in k-space

$$\left|\tilde{\Psi}_{\vec{k}}\right\rangle=\sum_{n}\beta_{n\vec{k}}e^{-i\int_{k_{0}}^{k}\tilde{\omega}_{n}(k')dk'}\left|\tilde{\psi}_{n\vec{k}}\right\rangle. \tag{8.9}$$

Now, if we put this into Eq. 3.24 and take the time-derivative, and get

$$\begin{aligned}|\dot{\tilde{\Psi}}(\vec{r},t)\rangle &= i\sum_n \Big(\dot{\beta}_n(t)e^{-i\int_0^t \tilde{\omega}_n(t')dt'}\,|\tilde{\psi}_n(\vec{r},t)\rangle + \beta_n(t)e^{-i\int_0^t \tilde{\omega}_n(t')dt'}\,|\dot{\tilde{\psi}}_n(\vec{r},t)\rangle \\ &\quad + \tilde{\omega}_n(t)\beta_n(t)e^{-i\int_0^t \tilde{\omega}_n(t')dt'}\,|\tilde{\psi}_n(\vec{r},t)\rangle\Big), \\ &= \sum_n \tilde{\omega}_n(t)\beta_n(t)e^{-i\int_0^t \tilde{\omega}_n(t')dt'}\,|\tilde{\psi}_n(\vec{r},t)\rangle\,. \end{aligned} \tag{8.10}$$

This leads to

$$\begin{aligned}|\dot{\tilde{\Psi}}(\vec{r},t)\rangle &= i\sum_n \Big(\dot{\beta}_n(t)e^{-i\int_0^t \tilde{\omega}_n(t')dt'}\,|\tilde{\psi}_n(\vec{r},t)\rangle + \beta_n(t)e^{-i\int_0^t \tilde{\omega}_n(t')dt'}\,|\dot{\tilde{\psi}}_n(\vec{r},t)\rangle\Big), \\ &= 0. \end{aligned} \tag{8.11}$$

Taking the inner product with the left eigenvector, we get

$$\langle\tilde{\varphi}_m(\vec{r},t)|\dot{\tilde{\Psi}}(\vec{r},t)\rangle = i\sum_n \Big(\dot{\beta}_n(t)e^{-i\int_0^t \tilde{\omega}_n(t')dt'}\,\langle\tilde{\varphi}_m(\vec{r},t)|\tilde{\psi}_n(\vec{r},t)\rangle + \beta_n(t)e^{-i\int_0^t \tilde{\omega}_n(t')dt'}\,\langle\tilde{\varphi}_m(\vec{r},t)|\dot{\tilde{\psi}}_n(\vec{r},t)\rangle\Big), \tag{8.12}$$

when $n \neq m$, the first term on the left side vanish, only leaving terms where $n = m$. Thus we have

$$\begin{aligned}\langle\tilde{\varphi}_m(\vec{r},t)|\dot{\tilde{\Psi}}(\vec{r},t)\rangle = i\Big[&\dot{\beta}_m(t)e^{-i\int_0^t\tilde{\omega}_m(t')dt'} + \sum_{n\neq m}\beta_n(t)e^{-i\int_0^t\tilde{\omega}_n(t')dt'}\,\langle\varphi_m(\vec{r},t)|\dot{\tilde{\psi}}_n(\vec{r},t)\rangle \\ &+ \beta_n(t)e^{-i\int_0^t\tilde{\omega}_n(t')dt'}\,\langle\varphi_m(\vec{r},t)|\dot{\tilde{\psi}}_m(\vec{r},t)\rangle\Big], \\ = 0,&\end{aligned} \tag{8.13}$$

which gives

$$\begin{aligned}\dot{\beta}_m(t)e^{-i\int_0^t\tilde{\omega}_m(t')dt'} = &-\sum_{n\neq m}\beta_n(t)e^{-i\int_0^t\tilde{\omega}_n(t')dt'}\,\langle\varphi_m(\vec{r},t)|\dot{\tilde{\psi}}_n(\vec{r},t)\rangle \\ &-\beta_m(t)e^{-i\int_0^t\tilde{\omega}_m(t')dt'}\,\langle\varphi_m(\vec{r},t)|\dot{\tilde{\psi}}_m(\vec{r},t)\rangle\,,\end{aligned} \tag{8.14}$$

leading to

$$\dot{\beta}_m(t) = -\sum_{n\neq m}\beta_n(t)e^{-i\int_0^t(\tilde{\omega}_n(t')-\tilde{\omega}_m(t'))dt'}\,\langle\varphi_m(\vec{r},t)|\dot{\tilde{\psi}}_n(\vec{r},t)\rangle - \beta_m(t)\,\langle\varphi_m(\vec{r},t)|\dot{\tilde{\psi}}_m(\vec{r},t)\rangle\,. \tag{8.15}$$

The problematic term is the exponential term. In order to handle this in the Hermitian case (where $\tilde{\omega} \in \mathbb{R}$), we rely on the Riemann-Lebesgue Lemma [201–203]. We extend this application for when $\tilde{\omega} \in \mathbb{C}$. We note that $\tilde{\omega}_n - \tilde{\omega}_m = \omega_n - \omega_m - i\Gamma_n + i\Gamma_m$ and, since we

are working in k-space, we will make a change of the variables. This means that we have

$$- \sum_{n \neq m} \beta_{nk} e^{-i \int_{k_0}^{k} (\tilde{\omega}_n(k') - \tilde{\omega}_m(k'))dk'} \langle \varphi_{mk} | \dot{\tilde{\psi}}_{nk} \rangle \tag{8.16}$$

$$- \beta_{mk} \langle \varphi_{mk} | \dot{\tilde{\psi}}_{mk} \rangle$$

$$= - \sum_{n \neq m} \beta_{nk}, e^{-i \int_{k_0}^{k} (\omega_n(k') - \omega_m(k'))dk'} e^{- \int_{k_0}^{k} (\Gamma_n(k') - \Gamma_m(k'))dk'}$$

$$\times \langle \varphi_{mk} | \dot{\tilde{\psi}}_{nk} \rangle - \beta_{mk} \langle \varphi_{mk} | \dot{\tilde{\psi}}_{mk} \rangle . \tag{8.17}$$

Now, we are ready to state the conditions for the adiabatic condition. If $\Gamma_n(k') >> \Gamma_m(k')$, then adiabaticity holds under the following circumstances. In all instances, we take $\omega_n(k') >> \omega_m(k')$:

1. If $\langle \tilde{\varphi}_{mk} | \dot{\tilde{\psi}}_{mk} \rangle$ has polynomial growth.

2. If $\Delta\Gamma >> 0$ such that $\langle \tilde{\varphi}_{mk} | \dot{\tilde{\psi}}_{mk} \rangle << e^{- \int_{k_0}^{k} \Delta\Gamma dk}$.

It can also hold in more limited circumstances when $\Delta\Gamma < 0$.

Now, suppose our system is initially in an eigenstate $|\tilde{\Psi}(\vec{r}, 0)\rangle = |\tilde{\psi}_n(\vec{k}(0))\rangle$. If we do an adiabatic evolution, we have

$$|\tilde{\Psi}(\vec{r}, t)\rangle = e^{-i \int_0^t \tilde{\omega}_n(t')dt'} e^{i\alpha_n(t)} |\tilde{\psi}_n(\vec{k}(t))\rangle . \tag{8.18}$$

If we transport $|\tilde{\psi}_n(\vec{k}(0))\rangle$ along a closed curve, the basis state evolves with a phase factor,

that is $|\tilde{\psi}(\vec{k}(T)\rangle = e^{i\nu_n(T)}\,|\tilde{\psi}_n(\vec{k}(0))\rangle$. We have that

$$|\tilde{\Psi}(\vec{r},T)\rangle = e^{-i\int_0^T \tilde{\omega}_n(t')dt'} e^{i\alpha_n^B(T)}\,|\tilde{\psi}_n(\vec{k}(0))\rangle\,. \tag{8.19}$$

with $\alpha_n^B(T) = \alpha_n(T) + \nu_n(T)$. We want to show that α_n^B is independent of a scaling factor - i.e., independent of a choice of gauge. If we have the scaled basis

$$|\tilde{\tilde{\psi}}_n(\vec{k}(t))\rangle = \sigma_n(\vec{k}(t))\,|\tilde{\psi}_n(\vec{k}(t))\rangle\,, \tag{8.20a}$$

$$\langle\tilde{\tilde{\varphi}}_n(\vec{k}(t))| = \frac{1}{\sigma_n(\vec{k}(t))}\,\langle\tilde{\varphi}_n(\vec{k}(t))|\,, \tag{8.20b}$$

and we have a time-evolved state

$$|\tilde{\Psi}(\vec{r},T)\rangle = e^{-i\int_0^T \tilde{\omega}_n(t')dt'} e^{i\bar{\alpha}_n(T)}\,|\tilde{\tilde{\psi}}_n(\vec{k}(T))\rangle\,, \tag{8.21}$$

where

$$\bar{\alpha}_n(T) = i\int_0^T \langle\tilde{\tilde{\varphi}}_n(\vec{k}(t))|\,\nabla_{\vec{k}}\,|\tilde{\tilde{\psi}}_n(\vec{k}(t))\rangle\cdot\frac{d\vec{k}}{dt}dt. \tag{8.22}$$

Now

$$|\tilde{\psi}_n(\vec{k}(T))\rangle = \frac{1}{\sigma_n(\vec{k}(T))}\,|\tilde{\bar{\psi}}_n(\vec{k}(T))\rangle\,, \tag{8.23a}$$

$$\langle\tilde{\varphi}_n(\vec{k}(0))| = \langle\tilde{\bar{\varphi}}_n(\vec{k}(0))|\,\sigma_n(\vec{k}(0)), \tag{8.23b}$$

$$\implies \langle\tilde{\varphi}_n(\vec{k}(0))|\tilde{\psi}_n(\vec{k}(T))\rangle = \frac{\sigma_n(\vec{k}(0))}{\sigma_n(\vec{k}(T))}\,\langle\tilde{\bar{\varphi}}_n(\vec{k}(0))|\tilde{\bar{\psi}}_n(\vec{k}(T))\rangle\,. \tag{8.23c}$$

We note that

$$\langle\tilde{\varphi}_n(\vec{k}(0))|\tilde{\psi}_n(\vec{k}(T))\rangle = e^{i\alpha_n(T)}, \tag{8.24}$$

$$\langle\tilde{\bar{\varphi}}_n(\vec{k}(0))|\tilde{\bar{\psi}}_n(\vec{k}(T))\rangle = e^{i\bar{\alpha}_n(T)}, \tag{8.25}$$

$$\implies e^{i\alpha_n(T)} = \frac{\sigma_n(\vec{k}(0))}{\sigma_n(\vec{k}(T))}e^{i\bar{\nu}_n(T)}. \tag{8.26}$$

We have that

$$\begin{aligned}
\bar{\alpha}_n^B(T) &= \bar{\alpha}_n(T) + \bar{\nu}_n(T), \\
&= i\int_0^T \langle \tilde{\bar{\varphi}}_n(\vec{k}(t))| \nabla_{\vec{k}} |\tilde{\bar{\psi}}_n(\vec{k}(t))\rangle \cdot \frac{d\vec{k}}{dt} dt + \bar{\nu}_n(T), \\
&= i\int_0^T \frac{1}{\sigma_n(\vec{k}(t))} \langle \tilde{\varphi}_n(\vec{k}(t))| \nabla_{\vec{k}} \sigma_n(\vec{k}(t)) |\tilde{\psi}_n(\vec{k}(t))\rangle \cdot \frac{d\vec{k}}{dt} dt + \bar{\nu}_n(T), \\
&= i\int_0^T \langle \tilde{\varphi}_n(\vec{k}(t))| \nabla_{\vec{k}} |\tilde{\psi}_n(\vec{k}(t))\rangle \cdot \frac{d\vec{k}}{dt} dt + i\int_0^T \frac{\nabla_{\vec{k}}\sigma_n(\vec{k}(t))}{\sigma_n(\vec{k}(t))} \cdot \frac{d\vec{k}}{dt} dt + \nu_n(T) - i\ln\left(\frac{\sigma_n(\vec{k}(T))}{\sigma_n(\vec{k}(0))}\right), \\
&= i\int_0^T \langle \tilde{\varphi}_n(\vec{k}(t))| \nabla_{\vec{k}} |\tilde{\psi}_n(\vec{k}(t))\rangle \cdot \frac{d\vec{k}}{dt} dt + \nu_n(T), \\
&= \alpha_n(T) + \nu_n(T), \\
&= \alpha_n^B(T).
\end{aligned} \tag{8.27}$$

Thus, we see that $\bar{\alpha}_n^B(T) = \alpha_n^B(T)$. We will now develop the equations further for periodic systems. In our system, we have

$$\vec{k} = \begin{bmatrix} k_x \\ 0 \\ 0 \end{bmatrix}. \tag{8.28}$$

Then we have

$$\begin{aligned}\alpha_n &= i \oint \langle \tilde{\varphi}_{m\vec{k}}(\vec{r})| \nabla_{\vec{k}} |\tilde{\psi}_{n\vec{k}}(\vec{r})\rangle \cdot d\vec{k} \\ &= i \oint \langle \tilde{\varphi}_{mk_x}(\vec{r})| \frac{\partial}{\partial k_x} |\tilde{\psi}_{nk_x}(\vec{r})\rangle \, dk_x. \end{aligned} \tag{8.29}$$

Now, recall that our left eigenvector $\langle \tilde{\varphi}_{m\vec{k}}(\vec{r})|$ is defined in relation to the right eigenvector as $\langle \tilde{\varphi}_{m\vec{k}}(\vec{r})| = \langle \tilde{\psi}_{m\vec{k}}(\vec{r})| \, \hat{D}$ such that $\langle \tilde{\psi}_{m\vec{k}}(\vec{r})|\hat{D}|\tilde{\psi}_{n\vec{k}}(\vec{r})\rangle = \delta_{mn}$. From this, we have

$$\begin{aligned} &\frac{\partial}{\partial \vec{k}}(\langle \tilde{\psi}_{m\vec{k}}(\vec{r})|\hat{D}|\tilde{\psi}_{n\vec{k}}(\vec{r})\rangle) = 0, \\ \implies &\frac{\partial}{\partial \vec{k}}(\langle \tilde{\psi}_{m\vec{k}}(\vec{r})|\hat{D}|\tilde{\psi}_{n\vec{k}}(\vec{r})\rangle) = \frac{\partial}{\partial \vec{k}}(\langle \tilde{\psi}_{m\vec{k}}(\vec{r})|)\hat{D}|\tilde{\psi}_{n\vec{k}}(\vec{r})\rangle + \langle \tilde{\psi}_{m\vec{k}}(\vec{r})|\frac{\partial}{\partial \vec{k}}(\hat{D})|\tilde{\psi}_{n\vec{k}}(\vec{r})\rangle \\ &\qquad\qquad + \langle \tilde{\psi}_{m\vec{k}}(\vec{r})|\hat{D}\frac{\partial}{\partial \vec{k}}(|\tilde{\psi}_{n\vec{k}}(\vec{r})\rangle), \\ &\qquad\qquad = 0. \end{aligned} \tag{8.30}$$

which gives

$$\frac{\partial}{\partial \vec{k}}(\langle \tilde{\psi}_{m\vec{k}}(\vec{r})|)\hat{D}|\tilde{\psi}_{n\vec{k}}(\vec{r})\rangle + \langle \tilde{\psi}_{m\vec{k}}(\vec{r})|\hat{D}\frac{\partial}{\partial \vec{k}}(|\tilde{\psi}_{n\vec{k}}(\vec{r})\rangle) = -\langle \tilde{\psi}_{m\vec{k}}(\vec{r})|\frac{\partial}{\partial \vec{k}}(\hat{D})|\tilde{\psi}_{n\vec{k}}(\vec{r})\rangle, \tag{8.31}$$

and so

$$\Big[\langle \tilde{\psi}_{m\vec{k}}(\vec{r})|\hat{D}\frac{\partial}{\partial \vec{k}}(|\tilde{\psi}_{n\vec{k}}(\vec{r})\rangle)\Big]^* + \langle \tilde{\psi}_{m\vec{k}}(\vec{r})|\hat{D}\frac{\partial}{\partial \vec{k}}(|\tilde{\psi}_{n\vec{k}}(\vec{r})\rangle) = -\langle \tilde{\psi}_{m\vec{k}}(\vec{r})|\frac{\partial}{\partial \vec{k}}(\hat{D})|\tilde{\psi}_{n\vec{k}}(\vec{r})\rangle. \tag{8.32}$$

Noting that $\langle \tilde{\psi}_{m\vec{k}}(\vec{r})|\hat{D}\frac{\partial}{\partial \vec{k}}(|\tilde{\psi}_{n\vec{k}}(\vec{r})\rangle) = \vec{A}$, the vector potential, we see that generally, $\vec{A} \in \mathbb{C}$.

In our specific case, $\frac{\partial \hat{D}}{\partial \vec{k}} = 0$, which means $\vec{A} \in \mathbb{R}$. Hence $\alpha^{\mathrm{B}} \in \mathbb{R}$ and our gauge group is $U(1)$.

8.2 Geometric Interpretation of Non-Hermitian Berry Phase

In Section A.3.1 and Section A.3.2, we showed how the Hermitian Berry phase comes from Riemannian geometry, which preserves lengths during parallel transport. The non-Hermitian Berry phase cannot have its roots in Riemannian geometry, because loss and gain affect the length of the state vector as it is parallel transported along a curve. This requires a different geometric setting that relaxes length preservation during parallel transport. One such setting is Weyl geometry.

In this section, we discuss the non-Hermitian Berry phase in the context of Weyl geometry. Weyl geometry and its influence on modern physics are discussed in [204–206]. Our discussion on the topic is based on these sources.

We begin by reminding our readers that the discussion on Hermitian Berry phases presented in Section A.3.1 relied on a mathematical space known as a fibre bundle. We used a fibre bundle that is Riemannian-like, as inner products are preserved during parallel transport in such a bundle. This worked because for Hermitian systems, the traditional Hilbert space inner product is indeed preserved while the state undergoes unitary evolution. We demonstrated this in Section A.3.2. In non-Hermitian systems, this breaks down, and

we can no longer rely on a Riemannian-like geometry. There has been some work done to provide a geometric interpretation of non-Hermitian Berry phases [207]. We would also like to point the reader to a recent preprint [208]. Here, I present my own interpretation of the non-Hermitian Berry phase in the fibre bundle setup. While we still work in a fibre bundle, our transport operator, S in Eq. A.15, is no longer unitary. We now revisit this equation and ask: how is the geometry of the fibre bundle changed when S is non-unitary? The fundamental answer to this question tells us that we are now dealing with a geometry that is Weyl-like [204–206], where lengths are no longer preserved during parallel transport. Below, we provide a summary of the general idea of Weyl geometry based on [209], in order to set the foundation for our geometric interpretation of non-Hermitian adiabatic evolution in the context of Weyl geometry.

The gauge field, A_α, is a central idea in Weyl geometry [204–206]. We explore this by first reminding the reader that for a metric g on a Riemann manifold, we have

$$\nabla_\alpha g = 0, \tag{8.33}$$

which is the length preservation condition of vectors under parallel transport. g can be the inner product of vectors, and Eq. 8.33 is the metric condition with respect to the Riemann metric. When Weyl introduced the new gauge field, A_α, the result is the condition

$$\nabla_\alpha g_{\mu\nu} = -2A_\alpha g, \tag{8.34}$$

which meant that inner products were no longer preserved during parallel transport with respect to the Riemann metric. Weyl identified this gauge field as the electromagnetic potential. He showed that when we locally change scale, i.e. $g_{\mu\nu} \rightarrow e^{2\theta(x)} g_{\mu\nu}$, then $A_\alpha \rightarrow A_\alpha - \partial_\alpha \theta$. The electromagnetic field is just the curvature associated with the gauge potential. If we recall Section A.3.2, the gauge potential played a central role in the geometric interpretation of the Hermitian Berry phase. As discussed in [210], for non-Hermitian systems, this norm is not preserved under adiabatic evolution. Here, we reintroduce Weyl's original idea that lengths of vectors are not preserved during parallel transport. We sketch out a geometric interpretation of non-Hermitian adiabatic evolution in this setting. The goal of this is to semi-intuitively establish a geometric setting for how we compute and geometrically interpret the non-Hermitian Berry phase used in this thesis. After an extensive literature search, there appears to be no explicit interpretation of non-Hermitian adiabatic evolution in the language of Weyl geometry.

We start with the fact that S is non-unitary, then if we assume $|\psi\rangle, |\phi\rangle \in \mathcal{H}$, and $|\psi'\rangle = S|\psi\rangle$, while $|\phi'\rangle = S|\phi\rangle$. We have $\langle\psi'|\phi'\rangle = \langle\psi|S^{-1}S|\phi\rangle = \langle\psi|\phi\rangle$. Since S is non-unitary, then $S^{-1} \neq S^\dagger$. In the unitary case, the dual vectors $\langle\psi|$ evolved with $S^\dagger$. This meant the dual vector space and the vector space coincide. However, in this case, the vector space and its dual do not coincide. We have that S acts on the vectors in the vector space so that $t \rightarrow t + \Delta t$, while in the dual vector space, this is done by S^{-1}. Our system must therefore have two distinct vector spaces that are dual to each other. These vector spaces require independent bases for each vector space. As there are two independent vector spaces associated with the same system, we therefore have two independent sets

of basis vectors. These are the left and right eigenvectors mentioned in Section 2.5.5 as part of the biorthogonal framework for non-Hermitian systems. Hence, $|\psi\rangle, |\phi\rangle$ cannot belong to the same Hilbert space, $\mathcal{H}$, as initially assumed here. It must be that we have $\langle\psi| \in \mathcal{H}_L$ and $|\phi\rangle \in \mathcal{H}_R$, where $\mathcal{H}_L, \mathcal{H}_R$ are the left and right Hilbert spaces respectively. Thus, in the Hermitian setup, since the left and right eigenvectors coincide, we have one Hilbert space; in the non-Hermitian case, the left and right eigenvectors don't coincide, so we have two Hilbert spaces. This is because in non-Hermitian systems, we are dealing with non-unitary evolution. In the unitary case, $S \in U(N)$, but in the non-unitary case, we have $S \in GL(N, \mathbb{C})$, where $GL(N, \mathbb{C})$ is the general linear group over complex numbers.

We will now write the left and right eigenvectors as $\langle\psi_L|, |\psi_R\rangle$, respectively. We go back to Eq. A.20, and we have

$$\nabla_{\dot{\gamma}(t)} |\psi_R(\gamma(t))\rangle = \lim_{\Delta t \to 0} \frac{S^{-1}_{t+\Delta t \to t} |\psi_R(\gamma(t+\Delta t))\rangle - |\psi_R(\gamma(t))\rangle}{\Delta t}. \tag{8.35}$$

We resort to parameter space, giving

$$D_\mu |\psi_R\rangle = \lim_{\Delta x \to 0} \frac{S^{-1}_{x+\Delta x \to x} |\psi_R(x+\Delta x)\rangle - |\psi_R(x)\rangle}{\Delta x}, \tag{8.36}$$

and expand $S^{-1}_{x+\Delta x \to x}$ and get $S^{-1}_{x+\Delta x \to x} \approx \mathbb{I} + \bar{A}_\mu \Delta x + \mathcal{O}(\Delta x^2)$, thus $S_{x+\Delta x \to x} \approx \mathbb{I} - \bar{A}_\mu \Delta x + \mathcal{O}(\Delta x^2)$. If we assume $S^\dagger = S$ and say $|\psi_R(x+\Delta x)\rangle = S_{x \to x+\Delta x} |\psi_R(x)\rangle$, then $|\psi_R(x+\Delta x)\rangle = S^\dagger_{x \to x+\Delta x} |\psi_R(x)\rangle$. If we have $\langle\psi_R(x+\Delta x)|\psi_R(x+\Delta x)\rangle = \langle\psi_R(x)|S^\dagger S|\psi_R(x)\rangle = \langle\psi_R(x)|\psi_R(x)\rangle$. Here, we note carefully $S^\dagger S = (\mathbb{I} - \bar{A}^\dagger_\mu \Delta x)(\mathbb{I} - \bar{A}_\mu \Delta x) = \mathbb{I} - (\bar{A}^\dagger_\mu + \bar{A}_\mu)\Delta x +$

$\mathcal{O}(\Delta x^2)$. This means that $\langle\psi_R(x)|S^\dagger S|\psi_R(x)\rangle = \langle\psi_R(x)|\psi_R(x)\rangle - \langle\psi_R(x)|(\bar{A}_\mu^\dagger + \bar{A}_\mu)|\psi_R(x)\rangle\, dx^\mu$. This then tell us that $\partial_\mu \langle\psi_R(x)|\psi_R(x)\rangle = -\langle\psi_R(x)|(\bar{A}_\mu^\dagger + \bar{A}_\mu)|\psi_R(x)\rangle$. Now if we set $\bar{A}_\mu = A_\mu + Q_\mu$, where

$$A_\mu = \frac{\bar{A}_\mu - \bar{A}_\mu^\dagger}{2}, \tag{8.37a}$$

$$Q_\mu = \frac{\bar{A}_\mu + \bar{A}_\mu^\dagger}{2}. \tag{8.37b}$$

Thus, we have $\partial_\mu \langle\psi_R(x)|\psi_R(x)\rangle = -2\langle\psi_R(x)|Q_\mu|\psi_R(x)\rangle$. If it is a line bundle (i.e., where the fibres are 1D), then the vector space is 1D, and thus the operators are scalar, which gives the result that $-\langle\psi_R(x)|Q_\mu|\psi_R(x)\rangle = -2Q_\mu\langle\psi_R(x)|\psi_R(x)\rangle$. Thus

$$\partial_\mu \langle\psi_R(x)|\psi_R(x)\rangle = -2Q_\mu \langle\psi_R(x)|\psi_R(x)\rangle\,. \tag{8.38}$$

If we set $g = \langle\psi_R(x)|\psi_R(x)\rangle$, the metric, then we see that

$$\partial_\mu g = -2Q_\mu g, \tag{8.39}$$

which looks like the metric preservation condition for the Weyl metric as discussed earlier in Eq. 8.34. This is a consequence of non-unitarity, for it is the fact that $S^\dagger S \neq \mathbb{I}$ for non-Hermitian systems that leads us to this. It is also a consequence of how we define the metric: the inner product of the right eigenvector with itself. If we define a 'biorthogonal'

metric, $g_b = \langle\psi_L(x)|\psi_R(x)\rangle$, then we have

$$\partial_\mu g_b = 0, \tag{8.40}$$

where Eq. 8.40 is the condition of Eq. 8.33 satisfied for the biorthogonal metric, g_b. We have now built the foundations to compute the non-Hermitian Wilson loop.

8.2.1 Non-Hermitian Wilson Loop

We begin with a triangulated space as we did in Section A.3.2 and we have with $S_{ij} \in GL(N, \mathbb{C})$ that

$$\nabla_{ij} |\psi\rangle = \lim_{l_{ij}\to 0} \frac{S_{ij} |\psi_j^R\rangle - |\psi_i^R\rangle}{l_{ij}}. \tag{8.41}$$

We note that $S_{ij} |\psi_i^R\rangle = e^{i\beta_{ij}} |\psi_j^R\rangle$, where β_{ij} is the phase that is acquired when we move from $|\psi_i^R\rangle \in F_i$, and $|\psi_i^L\rangle \in F_i^*$. Also, we have that

$$S_{ij} = \langle\psi_i^L|\psi_j^R\rangle . \tag{8.42}$$

This phase is an infinitesimal amount. We will follow a similar line of reasoning as we did in Section A.3.1, with neighbouring vertices having coordinates $x_j^\mu = x_i^\mu + \Delta x^\mu$, with $\Delta x^\mu = l_{ij} a^\mu$. We Taylor expand as before and get $|\psi_j^R\rangle = |\psi_i^R\rangle + a^\mu l_{ij} \partial_\mu |\psi_i^R\rangle + \mathcal{O}(l_{ij}^2)$. Taking the biorthogonal inner product with $\langle\psi_i^L|$, we have

$$\langle\psi_i^L|\psi_j^R\rangle \approx 1 + a^\mu l_{ij} \langle\psi_i^L| \partial_\mu |\psi_i^R\rangle . \tag{8.43}$$

Unlike the Hermitian case, $\langle\psi_i^L|\,\partial_\mu\,|\psi_i^R\rangle \in \mathbb{C}$. Thus we write

$$\langle\psi_i^L|\,\partial_\mu\,|\psi_i^R\rangle = -i\bar{A}_\mu, \tag{8.44}$$

where $\bar{A}_\mu = A_\mu + Q_\mu$. Therefore,

$$\begin{aligned} S_{ij} &= \langle\psi_i^L|\psi_j^R\rangle \\ &\approx 1 - i\bar{A}_\mu(x_i)a^\mu l_{ij} \\ &\approx e^{-i\bar{A}_\mu(x_i)a^\mu l_{ij}}. \end{aligned} \tag{8.45}$$

We can then do

$$\begin{aligned} \nabla_{ij}\,|\psi_i^R\rangle &= \lim_{l_{ij}\to 0}\frac{S_{ij}\,|\psi_j^R\rangle - |\psi_i^R\rangle}{l_{ij}} \\ &\approx a^\mu\left(\partial_\mu + i\bar{A}_\mu(x_i)\right)|\psi_i^R\rangle\,. \end{aligned} \tag{8.46}$$

The Wilson loop is now the product of the biorthogonal overlaps along a closed path.

Setting $j = i + 1$, we have

$$\begin{aligned}
\lim_{l_{i,i+1}\to 0}\prod_{i=1}^{N}\langle\psi_i^L|\psi_{i+1}^R\rangle &= \lim_{l_{i,i+1}\to 0}\prod_{i=1}^{N} e^{-i\bar{A}_\mu(x_i)a^\mu l_{i,i+1}}, \\
&= e^{-i\lim_{l_{i,i+1}\to 0}\sum_{i=1}^{N}\bar{A}_\mu(x_i)a^\mu l_{i,i+1}}, \\
&= e^{-i\oint \bar{A}_\mu\, dk^\mu}, \\
&= e^{-i\oint (A_\mu+Q_\mu)dk^\mu}, \\
&= H(\gamma, D). \qquad (8.47)
\end{aligned}$$

Since Q_μ is purely imaginary, if we set $Q_\mu = -i\tilde{Q}_\mu$, then we have $H(\gamma, D) = e^{-\oint \tilde{Q}_\mu dk^\mu} e^{-i\oint A_\mu dk^\mu}$. The decaying exponential part of the holonomy ($H(\gamma, D)$) leads to a change in length when the state undergoes adiabatic evolution. The other term is just the part that leads to the phase gain during adiabatic evolution. We must note something crucial here: A_μ is the part of the potential that gives curvature to the bundle, while Q_μ is the part that gives dilation/contraction. Thus, unlike the Hermitian fibre bundle, the non-Hermitian bundle can be viewed as a manifold that not only has curvature, but also dilates/contracts. It is a measure of the distortion of the space.

9

Conclusion & Outlook

The long-term vision of this thesis is to establish a roadmap for the rapid design of photonic metamaterials with integrated photonics device and fundamental physics applications. The goal is to do this in a manner where we balance the conceptual clarity and rapid exploration capabilities that theoretical models provide with the ability of computational models to provide accurate insights into the operation of realistic structures, without the high computational cost of computational models and the overly simplistic assumptions of theoretical models. We chose subwavelength gratings and explored one-dimensional periodic plasmonic materials. These structures, with their significant technological potential and enough room to gain relevant physical insights, are complex enough to test our methods on, while simul-

taneously pushing them to their limits. To this end, we investigated topological plasmonic systems using the finite element method (FEM) and quasinormal mode theory, addressing key gaps in both design and physical understanding. In parallel, we developed a data-driven effective medium approximation for subwavelength gratings (SWGs) by training deep neural networks on finite-difference time-domain (FDTD) simulations, enabling the design and interpretation of silicon-based SWGs beyond the limits of traditional effective medium theories. Together, these contributions expand our understanding of topological plasmonic systems and complex metamaterials, while providing efficient tools to design and gain physical insight into their operation.

The first important contribution of this work is a data-driven effective medium approximation for subwavelength gratings. Critical for integrated photonic devices, SWGs can be used to design components such as grating couplers. However, their finite size coupled with the complex electromagnetic interactions they normally host makes traditional effective medium theories ineffective. Using surrogate neural networks on FDTD Lumerical simulations, we demonstrated, using six geometries of a single design, that the concept of effective medium can be extended to some of the finite SWGs used in integrated photonic devices. This allowed us to represent the three-dimensional structures with equivalent two-dimensional structures using effective index profiles. The qualitative performance of our work demonstrate its potential to help reduce computational cost, while preserving the accuracy of the optical responses of interest, i.e., vertical and forward transmission, as well as reflection.

Currently, the SWG work demonstrated our method's performance on a limited number

of structures of a single design. As a result, competitive performance over conventional methods have not been rigorously demonstrated and observations of computational cost savings has only been qualitative and not quantitative. The next steps involve demonstrating competitive performance rigorously, including extending the work beyond the number of geometries presented.

The second important contribution of this work was the application of quasinormal mode theory to periodic plasmonic materials. Previous work explored the optical response of single metallic nanoresonators, and in rare cases, periodic systems. This work extends the framework to topological systems. We pursued two approaches: (1) full-lattice simulations to identify edge modes and compute topological invariants, and (2) a coupled-mode theory based on single-particle simulations to reconstruct bulk and open bands and predict edge states. The second approach addresses the computational demand and limitations of full-lattice simulations. It enables controlled exploration of interactions, including long-range coupling, and provides insight into the contribution of basis resonances to hybridized states. Our results show that dipole interactions dominate in spherical systems, with higher-order multipoles playing secondary roles, while long-range interactions and higher-order basis resonances have significant influence on edge-mode frequencies in cuboid systems, indicating that higher resonances might play a significant role in the bandstructure, potentially explaining the discrepancy between the coupled-mode theory and the full-lattice simulations. Importantly, it is demonstrated that even in situations where the coupled-mode theory has significant disagreements with the full-lattice simulations, it still can guide users in the right direction on the physics of the system.

Though bulk-boundary correspondence was demonstrated in the topological plasmonics work, we did not demonstrate robustness in the boundary states. We have also not demonstrated explicitly how much improvement higher-order modes provide. In addition, we only studied two geometries in a homogeneous background. Also, although the systems were non-Hermitian, non-Hermitian topological effects such as exceptional points were not studied. The next goals involve addressing these issues, and steps have already been taken in addressing some of these, for example, work has already begun in investigating resonators on substrates.

Advancing integrated photonic technologies requires demonstrating clear advantages of each case over electronic platforms. Improving design capabilities are important steps to achieving this. The tools developed in this thesis lay the foundation to address this gap by outlining two approaches that feature a roadmap to efficiently design components of integrated photonic devices. Next steps will involve further benchmark testing of these methods, using them for designing new devices and perform inverse design.

A

Appendices

In this chapter, we present additional material to provide the reader with a deeper understanding of some relevant concepts. In Section A.1, we discuss geometric phases; Section A.2 discusses adiabatic evolution of states, and Section A.3 discusses how topology emerges from these concepts.

A.1 Geometric Phases

Geometric phases occur in many areas of physics. Although they are often discussed in the context of wave phenomena, fundamentally, they are rooted in phenomena that have

the concept of a state undergoing some kind of cyclic evolution, where the states can be compared at infinitesimal points of a parameter space. Other examples of geometric phases in nature include the Aharonov-Bohm effect [211], where an electric charge can experience an electromagnetic potential, even when the field strength is zero. This showed that electromagnetic potentials are not just a mathematical construct, but have direct physical consequences. There is also the Berry phase [212], where a slow (adiabatic) cyclic evolution of a state results in the acquisition of a phase upon completion of a cycle. There is also the Aharonov-Anandan phase [213, 214], which generalizes the Berry phase, loosening the adiabatic requirement. In addition, there is the Aharonov-Casher effect [215], where a phase shift occurs in neutral particles with dipole moments.

In physics, geometric phases lead to effects that are often explained with mathematical concepts at the intersection of differential geometry and topology. In our work, the more relevant geometric phase is the Zak phase, which describes adiabatic evolution of a state in a 1D Brillouin zone.

A.2 Adiabatic Evolution of States

When a system undergoes a cyclic adiabatic evolution, a phase can be acquired when it completes the full evolution in parameter space, where the parameter space can be the Brillouin zone. Due to the slow evolution, the system has time to adjust without transitioning to another state. When we say the evolution is slow, we mean with respect to its own internal dynamics, measured in terms of how fast the system takes to transition

from one state to the other, i.e., with respect to the timescale $\tau \approx \frac{\hbar}{\Delta E}$, where ΔE is the transition gap. The physical discussion of this can be found in many quantum texts. We give a brief explanation of the topic from this angle. There is, however, a much more mathematically rigorous formalism of the subject in mathematical structures known as fiber bundles. We present a high-level discussion of fiber bundles in section A.3.1.

We outline the traditional discussion of the Berry phase for Hermitian systems starting with a formal statement of the adiabatic theorem [212, 216].

Theorem 1 (Adiabatic Theorem)**.** *Let $H(t)$ be a time-dependent Hamiltonian defined on a finite-dimensional Hilbert space for $t \in [0, T]$, and assume:*

1. *$H(t)$ is continuously differentiable in t,*
2. *For each t, $H(t)$ has a nondegenerate instantaneous eigenvalue $E_n(t)$ with normalized eigenstate $|\psi_n(t)\rangle$, indexed by n*
3. *There exists a uniform spectral gap*

$$g(t) = \min_{m \neq n} |E_m(t) - E_n(t)| > 0$$

for all $t \in [0, T]$.

Suppose the system is prepared at $t = 0$ in the instantaneous eigenstate

$$|\Psi(0)\rangle = |\psi_n(0)\rangle.$$

Then, in the adiabatic limit $T \to \infty$, the solution of the time-dependent Schrödinger equation

$$i\hbar \frac{d}{dt}|\Psi(t)\rangle = H(t)|\Psi(t)\rangle$$

satisfies

$$|\Psi(t)\rangle = e^{i\gamma_n(t)} e^{-\frac{i}{\hbar}\int_0^t E_n(t')\,dt'} |\psi_n(t)\rangle + \mathcal{O}\left(\frac{1}{T}\right),$$

where

$$\gamma_n(t) = i\int_0^t \langle \psi_n(t')|\dot{\psi}_n(t')\rangle dt'$$

is the geometric phase.

We have the corollary

Corollary 1 (Cyclic Evolution)**.** *If, in addition, $H(T) = H(0)$, then the state returns to itself up to a total phase,*

$$|\Psi(T)\rangle = e^{i\alpha_n}|\psi_n(0)\rangle,$$

where

$$\alpha_n = -\frac{1}{\hbar}\int_0^T E_n(t)\,dt + \gamma_n(T).$$

In order to obtain more formal statements with proofs of the theorem, please see [217, 218]. Essentially, what it says is that if we start our quantum system in an initial state and evolve it slowly, it will stay close to that initial state (and acquire a phase factor). The phase factor it acquires contains pertinent information regarding the evolution of the system. It consists of a dynamical term and a geometric term. It is the geometric term

that leads to a plethora of interesting physical phenomena, in particular, those that have to do with topological phases.

A.3 Connection with Topology

Topology [219, 220] is concerned with properties of a space that remain unchanged after continuous deformations are made to the space, i.e., properties called topological invariants, that remain the same when the spaces are changed without cutting or gluing them.

We illustrate one way the topological equivalence of two spaces may be determined by using a torus and a sphere. A topologist would ask, for example, what is the difference between a torus and a sphere, and how can we answer this question in a systematic manner? One method to address this is to put loops on the surface and see if we can contract them to a point without cutting and/or gluing. Movement of the loop in the space is allowed as well as rotation or generally, any kind of continuous and smooth transformation. This is depicted in Fig. A.1, which shows loops on a 2D sphere and 2D torus. On the sphere, we can adjust any one of those loops so that it may be contracted to a point. Loops 2 and 3 can be contracted to a point on the sphere. Loop 1, however, has to be moved on the sphere and adjusted for it to be contracted. This can be done without cutting it or removing it from the surface. Hence, one can in fact continuously deform loop 1 into loop 2 or loop 3. In fact, if we take any arbitrary loop anywhere on the sphere, we can always continuously deform it into a contractible loop. On the torus, however, if we pick loop 2, for example, we cannot continuously deform it into a contractible loop. We have to cut it

and glue it together. Loop 1 cannot be contracted to a point on the torus either, without cutting and gluing. If we try to move this loop, the hole in the torus obstructs it. Only loop 3 is contractible on the torus. Thus, we can say that the sphere and the torus are topologically different.

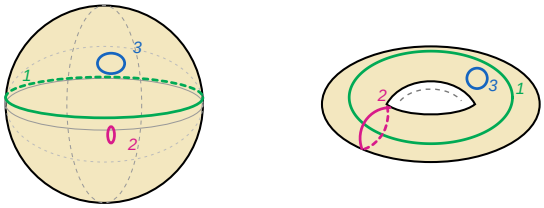


Figure A.1: Schematic of 2D sphere (left) and 2D torus (right) with loops. All loops on the sphere can be contracted to a point without cutting it or moving it off the surface. In the case of the torus, loops 1 and 2 cannot be contracted to a point without cutting it or having it fall off the surface.

One can see that asking questions about the contractibility of loops is important in topology. This question is also important in understanding topological materials. The idea of using the contractibility of loops to classify topological spaces falls under a branch of topology known as homotopy theory. Homotopy theory [219–221] is concerned with whether or not a loop in a space can be shrunk to a point to determine whether spaces are

equivalent. Homotopy uses group theory to classify spaces, which makes it important in the field of topology.

One of the most important groups in homotopy theory is the fundamental group, π. It is used to determine whether the set of loops at an arbitrary point in a space is equivalent. Two loops are equivalent if they can be continuously deformed into one another without cutting or gluing. In Fig. A.1, any loop drawn on the sphere can be continuously deformed into another; hence loops 1, 2 and 3 are all said to be homotopically equivalent to each other. As a result, there is only one type of loop. The fundamental group of the sphere is thus said to be trivial, $\pi(S^2) \cong 0$. If we look at a circle, S^1, on the other hand, there is a hole in the middle. A loop can wrap around it an integer number of times, so the fundamental group of the circle is the integers, i.e., $\pi(S^1) \cong \mathbb{Z}$.

On the torus, loops 1,2 and 3 are homotopically different from each other, because we have to either cut or glue to change one into the other. This is because the torus has two ways we can wrap a loop; one is to wrap a loop around the hole in the middle, like loop 1. The other is to wrap around the meridian, like loop 2. In both ways, one can wrap an integer number of times, so the fundamental group of the torus is $\pi(T^2) \cong \mathbb{Z} \times \mathbb{Z}$. We see then that the sphere and the torus are topologically different, for their loops belong to different homotopy groups.

Just as we can ask questions about the contractibility of loops on a sphere, torus, as well as other spaces, and use this to determine the topological relationship between two spaces, we can understand how two bulk Hamiltonians are topologically related by asking questions

about the contractibility of maps from the Brillouin zone into a space of Hamiltonians (or even the eigenspaces of these Hamiltonians). The Brillouin zone is a d-dimensional torus, T^d.

Let's illustrate how homotopy theory is useful for topological materials, using a concrete example, the 2-band SSH model, Eq. 2.41. We start by investigating the maps $f_t^{(j)}(\vec{x})$ to the space of all gapped Hamiltonians $\mathcal{G}$

$$f_t^{(j)}(\vec{x}) : T^d \to \mathcal{G} \tag{A.1}$$

and try to ascertain if they are contractible. Here, j is the label for the map, for example, $1, 2, 3, ...$, and t is a parameter(s) that we change to deform one curve into another. We express the SSH Hamiltonian, Eq. 2.41, in terms of the Pauli matrices. We do this by first noting that we can write the 2-band Hamiltonian as

$$H = \begin{bmatrix} d_3(k) + d_0(k) & d_1(k) - id_2(k) \\ d_1(k) + id_2(k) & -d_3(k) + d_0(k) \end{bmatrix} \tag{A.2}$$

and recall that the Pauli matrices are

$$\sigma_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \ \sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$
$$\sigma_2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \ \sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

Thus, we can write the general 2-band model as

$$H(k) = \vec{d}(k) \cdot \vec{\sigma} + d_0(k)\sigma_0 \tag{A.3a}$$

where

$$\vec{d}(k) = \begin{bmatrix} d_1(k) \\ d_2(k) \\ d_3(k) \end{bmatrix}, \quad \vec{\sigma} = \begin{bmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \end{bmatrix}$$

Now, if we apply chiral symmetry

$$\Gamma = \sigma_3 \tag{A.4}$$

then we have

$$\Gamma H \Gamma^{-1} = \sigma_3 H(k) \sigma_3^{-1}$$

and Eq. 2.41 now becomes

$$H(k) = d_1(k)\sigma_1 + d_2(k)\sigma_2, \tag{A.6}$$

where

$$d_1(k) = v + w\cos(ka), \tag{A.7}$$

$$d_2(k) = -w\sin(ka). \tag{A.8}$$

because of the periodic boundary condition. As a result, we see that the vector

$$\vec{d}(k) = (d_1(k), d_2(k), 0) \tag{A.9}$$

traces a circle in the (d_1, d_2)-plane as k traverses the Brillouin zone. If we normalize $\vec{d}(k)$, we have

$$\hat{d}(k) = \frac{\vec{d}(k)}{|\vec{d}(k)|}, \tag{A.10}$$

where we see that if $|\vec{d}(k)|= 0$, then $\hat{d}(k)$ is ill-defined. This happens when $\vec{d}(k) = (0, 0, 0)$. This is where the bandgap closes. Due to periodic boundary conditions and the fact that we are in the 1D Brillouin zone, the implication of the above is that

$$\hat{d} : S^1 \to \mathbb{R}^2\backslash\{0\} \implies H : S^1 \to \mathbb{R}^2\backslash\{0\} \tag{A.11a}$$

We depict the Hamiltonian for three different hopping parameters v and w in Fig. A.2. On the left is the Bloch sphere restricted to the equator, and on the right is the corresponding bandstructure. We see that in Fig. A.2

1. $|v|< |w|$ (orange and blue circle/bands): the curve winds once around the singularity (black point), and the bandgaps are open.

2. $|v|> |w|$ (black circle/bands): the curve does not encircle the singularity, but the bandgap is still open.

3. $|v|= |w|$ (green circle/bands): the curve and the singularity intersect, and the bandgap

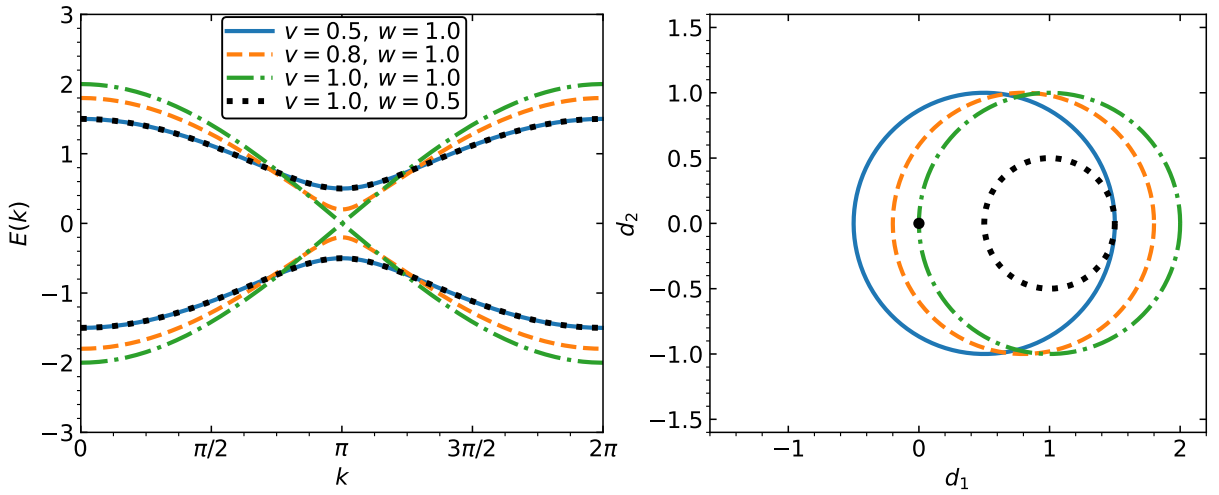


Figure A.2: Bandstructure (left) for different combinations of hopping parameters $(v, w) = (0.5, 1), (0.8, 1), (1, 1), (1, 0.5)$, represented by blue, orange, green and black curves, respectively. On the right is the corresponding representation in the (d_1, d_2)-plane. The singularity (band-crossing point) of the normalized d-vector, $\hat{d}$, is represented by the black point in the middle of the plane.

closes.

Traversing a blue band gives the blue circle, orange band gives orange circle, green band gives green circle, and black band gives black circle. As we can see, adjusting v, w determines which band we are on and thus which Hamiltonian we are dealing with. We navigate between Hamiltonians (and hence the loops) by changing the parameters v, w continuously. Now, suppose we focus on the blue bands. If we adjust v, w appropriately, we can end up on the orange band. This gets us from the blue loop to the orange loop. One did not cross a singularity (black point) in this process; hence the blue and orange loops are homotopic. As a result, the Hamiltonians that give the blue and orange bands are topologically the

same. If we continue to adjust v, w to get us to the black loop, then we have to cross the singularity. Looking at the bandstructure, this means we have to close the bandgap first (green band). This is why the green circle and the singularity intersect. As a result, moving from the blue bands to the black bands is a discontinuous transformation. This means the blue and black curves are not homotopic to each other. Thus, the Hamiltonian corresponding to the black bands and the Hamiltonian corresponding to the blue bands are not topologically the same. Hence, physically, one says that two Hamiltonians are topologically distinct if we can only get from one to the other by crossing a band closure. More precisely, comparing with Eq. A.1, if we set $\mathcal{G} = \mathbb{R}^2\backslash\{0\}$, $f_t^{(1)}(x) = \hat{d}^{(\text{blue})}_{(v,w)}(k)$, $f_t^{(2)}(x) = \hat{d}^{(\text{orange})}_{(v,w)}(k)$, $f_t^{(3)}(x) = \hat{d}^{(\text{green})}_{(v,w)}(k)$, $f_t^{(2)}(x) = \hat{d}^{(\text{black})}_{(v,w)}(k)$ and $T^d = S^1$. We have

$$\hat{d}^{(\text{blue})}_{(v,w)}(k) : S^1 \to \mathbb{R}^2\backslash\{0\} \implies H^{(\text{blue})}_{(v,w)}(k) : S^1 \to \mathbb{R}^2\backslash\{0\} \tag{A.12a}$$

$$\hat{d}^{(\text{orange})}_{(v,w)}(k) : S^1 \to \mathbb{R}^2\backslash\{0\} \implies H^{(\text{orange})}_{(v,w)}(k) : S^1 \to \mathbb{R}^2\backslash\{0\} \tag{A.12b}$$

$$\hat{d}^{(\text{green})}_{(v,w)}(k) : S^1 \to \mathbb{R}^2\backslash\{0\} \implies H^{(\text{green})}_{(v,w)}(k) : S^1 \to \mathbb{R}^2\backslash\{0\} \tag{A.12c}$$

$$\hat{d}^{(\text{black})}_{(v,w)}(k) : S^1 \to \mathbb{R}^2\backslash\{0\} \implies H^{(\text{black})}_{(v,w)}(k) : S^1 \to \mathbb{R}^2\backslash\{0\} \tag{A.12d}$$

Hence the loops in Fig. A.2 have the following relationship: $H^{(\text{blue})}_{(v,w)}(k) \ncong H^{(\text{black})}_{(v,w)}(k)$ and $H^{(\text{blue})}_{(v,w)}(k) \cong H^{(\text{orange})}_{(v,w)}(k)$, where $\cong$ and $\ncong$ means homotopic and non-homotopic respectively. We also have that $H^{(\text{blue})}_{(v,w)}(k) \ncong H^{(\text{green})}_{(v,w)}(k)$, $H^{(\text{orange})}_{(v,w)}(k) \ncong H^{(\text{black})}_{(v,w)}(k)$ and $H^{(\text{black})}_{(v,w)}(k) \ncong H^{(\text{green})}_{(v,w)}(k)$.

In order to classify the topological material, we find the fundamental group of the space,

$\mathcal{G}$, i.e., we find $\pi(\mathcal{G})$. In this case, we have

$$\pi(\mathcal{G}) \cong \pi(\mathbb{R}^2 \backslash \{0\})$$

but we note that $\mathbb{R}^2 \backslash \{0\} \sim S^1$ (where $\sim$ means similar to), hence

$$\pi(\mathbb{R}^2 \backslash \{0\}) = \pi(S^1) \sim \mathbb{Z}.$$

How is this related to the bulk information of the topological material? We use the band-structure (see $E(k)$ vs k in Fig. A.2), which contains information about the material's bulk properties. We use this to compute a topological invariant known as the winding number. It is an integer that tells us the number of times the loop (such as the blue, orange, green, or black loop, for example) encircles a singularity. It can wind clockwise or anti-clockwise. This determines the sign of the winding number. When we traverse the blue curve, the topological invariant is 1, because the blue circle goes once around the singularity in the clockwise direction. If we traverse the black curve, the winding number is 0, because the black loop doesn't encircle the singularity at all, so it doesn't wind around it. The fundamental group we computed above tells us the different possible windings we can get in general for an SSH system, and not just for our particular example. We showed above that the result $\pi(\mathbb{R}^2 \backslash \{0\}) \cong \mathbb{Z}$, which means that we can, in general, get a topological invariant of any integer value. Each value represents how many times one winds around a singularity and in what direction. Each winding number represents a distinct topological phase. In our case, we can only have a winding number, $\nu \in \{0, 1\} \subset \mathbb{Z}$. We take time to

note here that it is the (chiral) symmetry transformation that got us to this point. In fact, in general, it is the symmetry transformations that end up determining what homotopy groups a particular topological system belongs to, and hence, the nature of the topological invariant. In our above discussion, we use the simplest case of winding of a curve around a point. In more complex systems, one looks at the winding of an M-dimensional surface around an $(M-1)$-dimensional space.

A.3.1 Hermitian Fibre Bundles

In this section, we introduce the geometric interpretation of Berry phases. It turns out that the concept of geometric phases discussed in Section A.1 can be discussed in the context of curved spaces. This is relevant when computing them using the Wilson loop. Most curved spaces relevant for physics can be thought of as a combination of flat spaces (details of this are given later in this section). Flat spaces are understood in the framework of Euclidean geometry.

Euclidean geometry is concerned with spaces that are globally flat and have global coordinates, meaning that no matter where we are in the plane, the coordinates in one place are the same as those in another location in the space, i.e., the coordinates of the tangent spaces of two arbitrary points in a plane are the same. If we take the vector, $\vec{v}(t)$, along a curve, $\gamma(t)$, parametrized by a quantity, t, at two nearby points in the plane, this

can be done in a relatively straightforward way using the derivative

$$\frac{d\vec{v}(\gamma(t))}{dt} = \lim_{\Delta t \to 0} \frac{\vec{v}(\gamma(t+\Delta t)) - \vec{v}(\gamma(t))}{\Delta t}. \tag{A.14}$$

Fig. A.3 shows the effect of transporting a vector along a closed curve, $\gamma(t)$ (triangle with vertices A, B and C) in a plane so as to keep it parallel to itself. The vector doesn't change its magnitude or direction as the coordinates of the tangent space at every point are the same everywhere in the plane. This means that $\frac{d\vec{v}(\gamma(t))}{dt} = 0$.

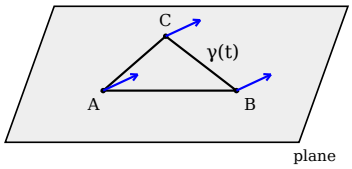


Figure A.3: Schematic of parallel transport of a vector (blue) along a closed curve. It starts from the red arrow and travels clockwise. When it completes a full loop, its direction changes and it gains a phase.

In a curved geometry [222, 223], the tangent space between two arbitrary points look different in general. Fig. A.4, left, shows an example of a sphere, with tangent spaces defined at A, B, and C. Notice the tangent spaces look different at the three points. As a consequence of this, we cannot define a global coordinate system for the space. We have to define local coordinates and develop tools to connect the coordinate systems between two

arbitrary tangent spaces. Riemannian geometry addresses this by treating curved spaces as

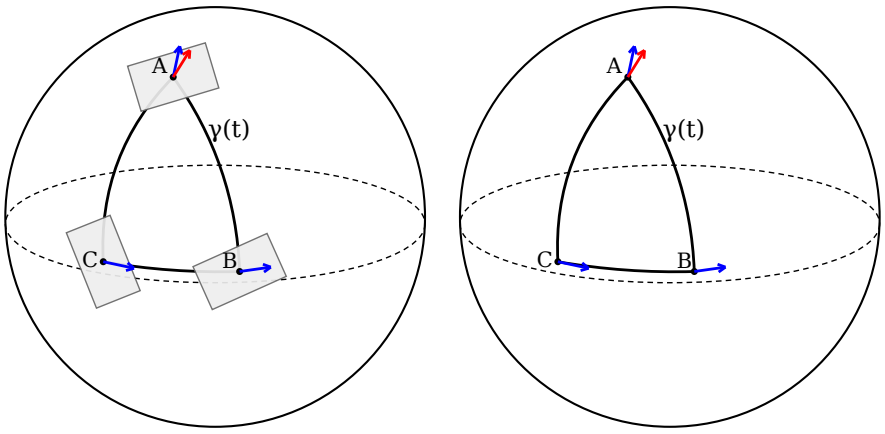


Figure A.4: Schematic of parallel transport of a vector along a closed curve. It starts from the blue arrow and travels clockwise. When it completes a full loop, its direction changes and it gains a phase (compare the red and blue arrows at A). The left image shows the tangent planes at three points in the loop, whereas the right image is without the tangent planes.

locally flat. More precisely, if we pick a point in the space and then choose all points very close to it, then the region containing all the chosen points would appear flat. This small region is called an ϵ-neighbourhood. In mathematical language, your ϵ-neighbourhood is locally flat and homeomorphic (similar) to a disk. In Riemannian geometry, spaces that are locally flat are called manifolds. The derivative defined in such a space would depend on how the two tangent spaces are connected/glued to each other. Therefore, one must add a correction to Eq. A.14 so that we can compare vectors at different points in the space.

This correction quantifies how those flat spaces must be glued/connected and is called a connection. This is what is used in place of a global coordinate system, as it precisely encodes the global properties of the space. It is used in differential geometry to compute geometric quantities such as the curvature of spaces. Suppose $\vec{v}(\gamma(t))$ is a vector at a point,

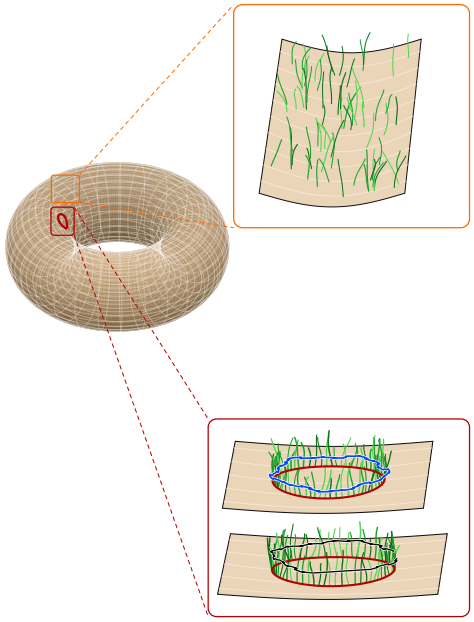


Figure A.5: Schematic of fibre bundle with Brillouin zone/k-space (BZ) as the torus. The Hilbert space $\mathcal{H}_k$ is the fibres.

p in a tangent space, and $\vec{v}(\gamma(t+\Delta t))$ is a vector in the tangent space of a point nearby q. In order to compare $\vec{v}(\gamma(t))$ and $\vec{v}(\gamma(t+\Delta t))$, we define an operator, $S^{-1}_{t+\Delta t\to t}$ to do this. The connection determines how this operator acts on the vector. We thus have the limit

definition of the derivative

$$\nabla_{\dot{\gamma}(t)}\vec{v} = \lim_{\Delta t \to 0} \frac{S^{-1}_{t+\Delta t \to t}\vec{v}(\gamma(t+\Delta t)) - \vec{v}(\gamma(t))}{\Delta t} \tag{A.15}$$

The derivative above is known as the covariant derivative. Just as in Eq. A.14, where the derivative is zero when the vector is constant, when we parallel transport a vector in the curved geometry along a closed curve, its covariant derivative is zero, i.e., $\nabla_{\dot{\gamma}(t)}\vec{v} = 0$.

In the geometric interpretation of gauge theories, the vector potential is identified as a connection for a specific kind of manifold known as a fibre bundle. It turns out that this manifold is precisely the kind that is relevant for a geometric interpretation of geometric phases as well. In the case of Hermitian systems specifically, we are referring to a fibre bundle with Riemannian-like geometry. The reason it is Riemannian-like can be explained by highlighting an important property of Riemannian manifolds. One of them is that Riemannian manifolds preserve the lengths of vectors during parallel transport. This means that the metric, $g(\vec{u}, \vec{v})$, which is the property used to measure lengths/distances on the manifold, is constant, i.e., $\partial_\mu g(\vec{u}, \vec{v}) = 0$. Now, since

$$\cos\theta = \frac{g(\vec{u}, \vec{v})}{\sqrt{g(\vec{u}, \vec{u})g(\vec{v}, \vec{v})}} \tag{A.16}$$

where θ is the angle between the vectors $\vec{u}$ and $\vec{v}$. Since $\partial_\mu g(\vec{u}, \vec{v}) = 0$, then it means that θ must be constant. Thus, if we parallel transport vectors along a curve on a Riemann manifold, the angle between them is also preserved.

Hermitian systems are lossless, and as a result, the norm of their eigenstates, $|\psi\rangle$, which defines the metric in Hilbert space, is constant, i.e., $\partial_\mu \langle\psi|\psi\rangle = 0$. Hence, the angle between two states is preserved under parallel transport as well. This motivates the case that the geometric setting for adiabatic transport in Hermitian systems must be Riemann-like. Thus, the fibre bundle we will work with for Hermitian systems must be a Riemann-like fibre bundle. In order to set the stage, we switch to a more position-coordinate case, where $\gamma(t) = x^\mu(t)$. If we do this, Eq. A.15 takes the form

$$\nabla_{\dot{\gamma}(t)}\vec{v} = \dot{x}_\mu D_\mu \vec{v} \tag{A.17}$$

where

$$D_\mu \vec{v} = \lim_{\Delta x \to 0} \frac{S^{-1}_{x+\Delta x \to x}\vec{v}(x+\Delta x) - \vec{v}(x)}{\Delta x} \tag{A.18}$$

This is the discrete form of the more familiar $D_\mu = \partial_\mu + A_\mu$ which describes parallel transport of vectors in curved spaces such as fibre bundles.

A fibre bundle [224, 225] consists of a base space (which could be any manifold) and spaces known as fibres. The fibres are attached to each point in the base. They can either be trivial or non-trivial. The geometry of the fibre bundle is determined by the nature of the inner product norm of vectors in the bundle as well as the preservation or lack thereof of the angles between vectors parallel transported in the bundle. Fig. A.5 shows an example of a fibre bundle relevant to topological materials. The torus is momentum space. The magnification of a region of the torus (orange box) shows protrusions from the torus. These are the fibres that are attached to each point of the torus. In our case, these

fibres are identical copies of Hilbert space. When we choose a Hamiltonian, $H^{(j)}(\vec{k})$, its eigenstates exist within the fibre at each point in k-space. So, when we find the eigenstates at each k, we are tracing out a path along a group of fibres. This is depicted in the red box of Fig. A.5. The path can either be closed (blue curve) or not closed (black curve) as it returns to the fibre it started from. When the path in the fibre is closed, it means the fibre bundle is trivial. When it is not closed, we have a non-trivial bundle. The way this path is traced is dependent on how the eigenspaces are related/glued to each other in the global structure of the bundle. This is measured by the connection. If the connection is trivial, the paths through the fibre will close. When it is non-trivial, the paths are open. The extent to which the paths fail to close is measured by a quantity known as the holonomy which is computed using the Wilson loop.

A.3.2 Hermitian Wilson Loop

When we do parallel transport in a fibre bundle, we are comparing fibres infinitesimally close to each other, instead of tangent spaces. In the case of topological materials, the fibres are identical copies of the Hilbert space. When we parallel transport a vector between the two fibres, we are moving an eigenstate from one fibre to the next, and comparing the eigenstates between the two. More precisely, suppose we have a curve $\mathcal{C}(t)$ in the base space, such as the red one on the torus in Fig. A.5. This traces a curve $\gamma(t)$ in the fibres, such as the blue or black curve in the red box of Fig. A.5. We find the eigenstate $|\psi(\gamma(t))\rangle$ in one fibre, directly above $\mathcal{C}(t)$. Then we find $|\psi(\gamma(t+\Delta t))\rangle$ in the adjacent fibre, directly

above $\mathcal{C}(t + \Delta t)$.

Just as when we compare vectors on a curve between tangent spaces, we compare states on a curve between fibres. We do this, by parallel transporting $|\psi(\gamma(t + \Delta t))\rangle$ into the same fibre as $|\psi(\gamma(t))\rangle$. We thus have the derivative

$$\nabla_{\dot{\gamma}(t)} |\psi(\gamma(t))\rangle = \lim_{\Delta t \to 0} \frac{S^{-1}_{t+\Delta t \to t} |\psi(\gamma(t + \Delta t))\rangle - |\psi(\gamma(t))\rangle}{\Delta t} \tag{A.19}$$

When we perform adiabatic transport along a closed curve, the eigenstate remains unchanged, i.e., $\nabla_{\dot{\gamma}(t)} |\psi(\gamma(t))\rangle = 0$. Thus, when one says that the eigenstate remains unchanged during a cyclic adiabatic evolution, geometrically, this means the covariant derivative of this state is zero along a closed curve in parameter space. The connection gives us the precise rule for how $S^{-1}_{t+\Delta t \to t}$ is to be used to compare the eigenstate from two adjacent fibres.

Recall that in Section A.3.1, we argued that for a Hermitian system, a Riemann-like bundle is the appropriate geometric setting. We show that the unitary operator is the relevant operator here, to be used in Eq. A.20, i.e. $S = \hat{U}$. More precisely, if we let $|\psi\rangle \in \mathcal{H}$ (where $\mathcal{H}$ is a Hilbert space) be an eigenstate for a Hermitian system at time t. Let S act on $|\psi\rangle$ such that $|\psi'\rangle = S |\psi\rangle$. Then S must be unitary. We can use the unitary property of S to show how the connection emerges, by first performing a triangulation of the base space of the fibre bundle. We will then show that the connection in this setting is in fact the Berry potential. From there, we will then obtain another geometric quantity known as a holonomy, which we compute by finding the Wilson loop. We will in fact show

that the holonomy is the Berry phase.

Due to the fact that manifolds are locally flat, even though they are globally curved, they admit a triangulation. This means that one can cut the curved space into flat (Euclidean) triangles that are infinitesimally small, and glue them back together to recover the space. This is the reason why, for example, complex shapes can be modeled in a finite element mesh using triangulations. Gluing these flat spaces back together to recover the correct space is non-trivial when the space is curved.

Recall earlier, we cannot naively compare states between two arbitrary fibres. As a consequence, the way these infinitesimal spaces are glued together affects how one compares the eigenstates between the fibres at the vertices of a triangle in the triangulation. As we are working with triangulated spaces, it will be helpful to resort to Eq. A.18, which is the covariant derivative in terms of position rather than time. In our case, x^μ is the parameter space, where the parameter could be, for example, the wavevector/momentum vector. Fig. A.6 shows a triangulated fibre bundle. Here, the base space is triangulated (left), and the vertices of the triangles are the locations of the fibres (right), labeled F_j. The triangles are oriented anti-clockwise, and their edges are oriented in such a manner as to respect a global orientation. U_{ij} is equivalent to S in Eq. A.20 and parallel transports a vector (state) between two adjacent fibres F_i and F_j.

With this in mind, we note that if we take the lengths of the triangle between vertices i, j to be l_{ij}, then the discrete covariant derivative of the state $|\psi\rangle$ on the triangulated

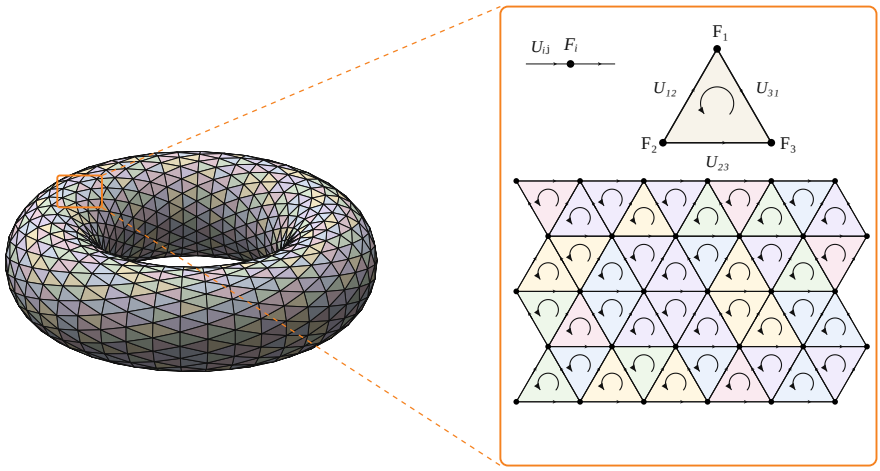


Figure A.6: Schematic of triangulation of fibre bundle for Hamiltonian.

bundle is

$$D_{ij}\left|\psi\right\rangle = \lim_{l_{ij}\to 0} \frac{U_{ji}\left|\psi_j\right\rangle - \left|\psi_i\right\rangle}{l_{ij}} \tag{A.20}$$

where $U_{ji} = U_{ij}^{-1}$. We define $U_{ij}\left|\psi_i\right\rangle = e^{i\alpha_{ij}}\left|\psi_j\right\rangle$, where α_{ij} is the phase that is acquired when we move from $\left|\psi_i\right\rangle \in F_i$ to $\left|\psi_j\right\rangle \in F_j$. This is an infinitesimal amount, since l_{ij} is infinitesimal. It is straightforward to show that

$$U_{ij} = \frac{\langle\psi_i|\psi_j\rangle}{|\langle\psi_i|\psi_j\rangle\,|} \tag{A.21}$$

Now, suppose neighbouring vertices have coordinates $x_j^\mu = x_i^\mu + \Delta x^\mu$, with $\Delta x^\mu = l_{ij}a^\mu$. Since the size of the triangles is infinitesimal, we can Taylor expand and get $\left|\psi_j\right\rangle = \left|\psi_i\right\rangle +$

$a^\mu l_{ij}\partial_\mu |\psi_i\rangle + \mathcal{O}(l_{ij}^2)$. Taking the inner product of the Taylor expansion with $\langle\psi_i|$, we have

$$\langle\psi_i|\psi_j\rangle \approx 1 + a^\mu l_{ij} \langle\psi_i| \partial_\mu |\psi_i\rangle \tag{A.22}$$

It is straightforward to show that $\langle\psi_i| \partial_\mu |\psi_i\rangle$ is purely imaginary, and hence can be written as $\langle\psi_i| \partial_\mu |\psi_i\rangle = -iA_\mu(x_i)$, where $A_\mu(x_i) \in \mathbb{R}$. $A_\mu(x_i)$ is in fact the connection, specifically, the Berry connection. Noting that $|\langle\psi_i|\psi_j\rangle| \approx 1$ at 1st order, we have

$$\langle\psi_i|\psi_j\rangle \approx e^{-iA_\mu(x_i)a^\mu l_{ij}} \tag{A.23}$$

Thus $U_{ij} \approx e^{-iA_\mu(x_i)a^\mu l_{ij}}$. Now, if we have

$$\begin{aligned} U_{ji} |\psi_j\rangle &\approx (1 + iA_\mu(x_i)a^\mu l_{ji})(|\psi_i\rangle + a^\nu l_{ij}\partial_\nu |\psi_i\rangle) + \mathcal{O}(l_{ij}) \\ &\approx |\psi_i\rangle + a^\mu l_{ij}(\partial_\mu + iA_\mu(x_i)) |\psi_i\rangle + \mathcal{O}(l_{ij}) \end{aligned}$$

This gives

$$\begin{aligned} \frac{U_{ji} |\psi_j\rangle - |\psi_i\rangle}{l_{ij}} &\approx a^\mu(\partial_\mu + iA_\mu(x_i)) |\psi_i\rangle + \mathcal{O}(l_{ij}) \\ \implies \lim_{l_{ij}\to 0} \frac{U_{ji} |\psi_j\rangle - |\psi_i\rangle}{l_{ij}} &\approx a^\mu(\partial_\mu + iA_\mu(x_i)) |\psi_i\rangle \end{aligned}$$

Recall from Eq. A.20 that the left side is just the covariant derivative. Thus, we have

$$\nabla_{ij} |\psi_i\rangle = a^\mu(\partial_\mu + iA_\mu(x_i)) |\psi_i\rangle \tag{A.25}$$

where we set $D_\mu = \partial_\mu + iA_\mu$, which is a more familiar expression of the covariant derivative in the continuum limit.

The Wilson loop is the product of $\langle\psi_i|\psi_j\rangle$ along a closed path, which means we set $j = i + 1$ and do

$$\begin{aligned}\lim_{l_{i,i+1}\to 0}\prod_{i=1}^{N}\langle\psi_i|\psi_j\rangle &= \lim_{l_{i,i+1}\to 0}\prod_{i=1}^{N} e^{-iA_\mu(x_i)a^\mu l_{i,i+1}} \\ &= e^{-i\lim_{l_{i,i+1}\to 0}\sum_{i=1}^{N} A_\mu(x_i)a^\mu l_{i,i+1}} \\ &= e^{-i\oint \vec{A}\cdot dk} \\ &= H(\gamma, D)\end{aligned}$$

This gives a geometric quantity known as the holonomy. We see that the term in the exponent of the holonomy is just the Berry phase. Due to the fact that we get a complex exponential, then we see that $H(\gamma, D) \in U(1)$. $U(1)$ is an abelian group, so what we derived is the abelian Berry phase. It is a solution to the covariant derivative equation, Eq. A.25, when the order of the paths doesn't matter. In the non-abelian case, where we have degenerate bands, the solution to this equation is the path-ordered exponential, i.e. $H(\gamma, D) = Pe^{-i\oint \vec{A}\cdot dk}$, where P just says that the order of the product matters the same way the order of a matrix product does. This generates a Wilson loop that is matrix-valued. The reason for this is that while the abelian case has 1D fibres, the non-abelian case has fibres of dimensions $n > 1$. This creates a matrix-valued connection, and since matrix products do not commute, the products of U_{ij} do not commute.

References

1. Gaponenko, S. V. *Introduction to Nanophotonics* (Cambridge University Press, 2010).
2. Xu, H. *Nanophotonics : Manipulating Light with Plasmons* First edition. eng. ISBN: 1-351-76757-7 (Taylor and Francis, London, 2017).
3. Benisty, H., Greffet, J.-J. & Lalanne, P. *Introduction to nanophotonics* eng. ISBN: 1-5231-5612-0 (Oxford University Press, Oxford, England, 2022).
4. Halir, R. *et al.* Subwavelength-Grating Metamaterial Structures for Silicon Photonic Devices. *Proceedings of the IEEE* **106,** 2144–2157 (2018).
5. Cheben, P. *et al.* Recent advances in metamaterial integrated photonics. *Adv. Opt. Photon.* **15,** 1033–1105 (2023).
6. Hou, T., Xiao, W. & Chen, H. Cosmology analogy for perfect hyperlens. *npj Nanophotonics* **2,** e12. arXiv: `2407.19506 [physics.optics]` (2025).
7. Chen, Y.-L., Cao, Q.-T. & Xiao, Y.-F. Trapping photons with optical black hole. *Light: Science & Applications* **12,** 73 (2023).
8. Procopio, L. M., Aguero-Santacruz, R., Bermudez, D. & Leonhardt, U. Backreaction of stimulated Hawking radiation in an optical analogue. *Nature* **655,** 336–341. arXiv: `2607.01118 [gr-qc]` (2026).
9. Cersonsky, R. K., Antonaglia, J., Dice, B. D. & Glotzer, S. C. The diversity of three-dimensional photonic crystals. *Nature Communications* **12,** 2543 (2021).
10. Peng, K. *et al.* Functional Photonic Crystal Materials for Biological Interfaces. eng. *ACS materials letters.* ISSN: 2639-4979 (2026).
11. Teyssier, J., Saenko, S. V., van der Marel, D. & Milinkovitch, M. C. Photonic crystals cause active colour change in chameleons. *Nature Communications* **6,** 6368 (2015).
12. Zi, J. *et al.* Coloration strategies in peacock feathers. *Proceedings of the National Academy of Sciences* **100,** 12576–12578. eprint: `https://www.pnas.org/doi/pdf/10.1073/pnas.2133313100` (2003).
13. Jones, J. B., Sanders, J. V. & Segnit, E. R. Structure of Opal. *Nature* **204,** 990–991 (1964).
14. Saranathan, V. *et al.* Structure, function, and self-assembly of single network gyroid ($I4_132$) photonic crystals in butterfly wing scales. *Proceedings of the National Academy of Sciences* **107,** 11676–11681. eprint: `https://pnas.org` (2010).

15. Matthews, J., Politi, A., Stefanov, A. & O'Brien, J. Manipulation of multiphoton entanglement in waveguide quantum circuits. *Nature Photonics - NAT PHOTONICS* **3,** 346–350 (May 2009).

16. Pernice, W. H. P. *et al.* High-speed and high-efficiency travelling wave single-photon detectors embedded in nanophotonic circuits. *Nature Commun.* **3,** 1325 (2012).

17. Politi, A., Cryan, M. J., Rarity, J. G., Yu, S. & OBrien, J. L. Silica-on-Silicon Waveguide Quantum Circuits. *Science* **320,** 646–649. eprint: `https://www.science.org/doi/pdf/10.1126/science.1155441` (2008).

18. Takesue, H. *et al.* Entanglement generation using silicon wire waveguide. *Applied Physics Letters* **91,** 201108. ISSN: 0003-6951. eprint: `https://pubs.aip.org/aip/apl/article-pdf/doi/10.1063/1.2814040/13385415/201108\_1\_online.pdf` (Nov. 2007).

19. Jin, H. *et al.* On-Chip Generation and Manipulation of Entangled Photons Based on Reconfigurable Lithium-Niobate Waveguide Circuits. *Phys. Rev. Lett.* **113,** 103601 (10 2014).

20. Yu, J. *et al.* Bioinspired mechano-photonic artificial synapse based on graphene/MoS¡sub¿2¡/sub¿ heterostructure. *Science Advances* **7,** eabd9117. eprint: `https://www.science.org/doi/pdf/10.1126/sciadv.abd9117` (2021).

21. Tsang, H. K. *et al.* Photonic integrated circuits and their applications. *National Science Review* **13,** nwag233. ISSN: 2095-5138. eprint: `https://academic.oup.com/nsr/article-pdf/13/19/nwag233/68106629/nwag233.pdf` (Oct. 2026).

22. Luque-González, J. M. *et al.* A review of silicon subwavelength gratings: building break-through devices with anisotropic metamaterials. *Nanophotonics* **10,** 2765–2797 (2021).

23. Sun, L., Zhang, Y., He, Y., Wang, H. & Su, Y. Subwavelength structured silicon waveguides and photonic devices. *Nanophotonics* **9** (May 2020).

24. Halir, R. *et al.* Subwavelength-Grating Metamaterial Structures for Silicon Photonic Devices. *Proceedings of the IEEE* **106,** 1–14 (Aug. 2018).

25. Cheben, P., Halir, R., Schmid, J. H., Atwater, H. A. & Smith, D. R. Subwavelength integrated photonics. *Nature* **560,** 565 –572 (2018).

26. Kazanskiy, N., Butt, M. & Khonina, S. Silicon photonic devices realized on refractive index engineered subwavelength grating waveguides-A review. *Optics Laser Technology* **138,** 106863. ISSN: 0030-3992 (2021).

27. Subbaraman, H. *et al.* Recent advances in silicon-based passive and active optical interconnects. *Opt. Express* **23,** 2487–2511 (2015).

28. Cheng, L., Mao, S., Li, Z., Han, Y. & Fu, H. Grating Couplers on Silicon Photonics: Design Principles, Emerging Trends and Practical Issues. *Micromachines* **11,** 666 (2020).

29. Mu, X., Wu, S., Cheng, L. & Fu, H. Edge Couplers in Silicon Photonic Integrated Circuits: A Review. *Applied Sciences* **10,** 1538 (2020).

30. Zhou, X. *et al.* Fully Etched Low-Back-Reflection and High-Efficiency Silicon Waveguide Grating Couplers With Minimum Feature Size Above 260 nm. eng. *Journal of lightwave technology* **43,** 3854–3859. ISSN: 0733-8724 (2025).

31. Kamandar Dezfouli, M. *et al.* Perfectly vertical surface grating couplers using subwavelength engineering for increased feature sizes. *Optics Letters* **45,** 3701–3704 (2020).

32. Taillaert, D., Bienstman, P. & Baets, R. Compact efficient broadband grating coupler for silicon-on-insulator waveguides. *Optics Letters* **29,** 2749–2751 (2004).

33. Das, H. M., Stewart, J. & Ye, W. N. Challenges and research opportunities in silicon-based subwavelength grating technologies. *Advanced Optical Technologies* **15,** 1916546 (2026).

34. Benedikovic, D. *et al.* Subwavelength index engineered surface grating coupler with sub-decibel efficiency for 220-nm silicon-on-insulator waveguides. *Optics Express* **23,** 22628–22635 (2015).

35. Benedikovic, D. *et al.* High-efficiency single etch step apodized surface grating coupler using subwavelength structure. *Laser & Photonics Reviews* **8,** L93–L97 (2014).

36. Dezfouli, M. K. *et al.* Perfectly vertical surface grating couplers using subwavelength engineering for increased feature sizes. *Optics Letters* **45,** 3701–3704 (2020).

37. Cheben, P., Xu, D., Janz, S. & Densmore, A. Subwavelength waveguide grating for mode conversion and light coupling in integrated optics. *Optics express* **14,** 4695–702 (June 2006).

38. Halir, R. *et al.* Colorless directional coupler with dispersion engineered sub-wavelength structure. *Optics Express* **20,** 13470–7 (May 2012).

39. Wang, Y. *et al.* Compact Broadband Directional Couplers Using Subwavelength Gratings. *IEEE Photonics Journal* **8,** 1–8 (2016).

40. Xiong, Y. *et al.* Polarization splitter and rotator with subwavelength grating for enhanced fabrication tolerance. *Opt. Lett.* **39,** 6931–6934 (2014).

41. He, Y. *et al.* Silicon Polarization Splitter and Rotator using a Subwavelength Grating based Directional Coupler. In *Optical Fiber Communication Conference* Th1G.6 (Optica Publishing Group, 2017).

42. Wang, J. *et al.* Subwavelength grating enabled on-chip ultra-compact optical true time delay line. *Scientific Reports* **6,** 30235 (July 2016).

43. Midtvedt, D. *et al.* Deep learning in light–matter interactions. *Nanophotonics* **11,** 3189–3214 (2022).

44. Wiecha, P., Arbouet, A., Christian, G. & Muskens, O. Deep learning in nano-photonics: inverse design and beyond. *Photonics Research* **9** (Jan. 2021).

45. Ma, W. *et al.* Deep learning for the design of photonic structures. *Nature Photonics* **15,** 77–90 (Feb. 2021).

46. Wiecha, P. R., Arbouet, A., Girard, C. & Muskens, O. L. Deep learning in nano-photonics: inverse design and beyond. *Photon. Res.* **9,** B182–B200 (2021).

47. Dong, T. *et al.* Artificial neural network assisted the design of subwavelength-grating waveguides for nanoparticles optical trapping. *Opt. Express* **32,** 9656–9670 (2024).

48. Guo, J.-H., Zhang, Y.-L., Zhang, S.-S., Cai, C.-L. & Liang, H.-F. A Design Method of Two-Dimensional Subwavelength Grating Filter Based on Deep Learning Series Feedback Neural Network. *Sensors* **22.** ISSN: 1424-8220 (2022).

49. Selle, R. *et al.* Modeling of light-matter interactions with neural networks. *Phys. Rev. A* **76,** 023810 (2 2007).

50. Selle, R, Brixner, T, Bayer, T, Wollenhaupt, M & Baumert, T. Modelling of ultrafast coherent strong-field dynamics in potassium with neural networks. *Journal of Physics B: Atomic, Molecular and Optical Physics* **41,** 074019 (2008).

51. Malkiel, I. *et al.* Plasmonic nanostructure design and characterization via Deep Learning. *Light: Science Applications* **7** (Dec. 2018).

52. An, S. *et al.* A Deep Learning Approach for Objective-Driven All-Dielectric Metasurface Design. *ACS Photonics* (2019).

53. Tu, X. *et al.* Analysis of Deep Neural Network Models for Inverse Design of Silicon Photonic Grating Coupler. *Journal of Lightwave Technology* **PP,** 1–1 (Feb. 2021).

54. Melati, D. *et al.* Mapping the global design space of nanophotonic components using machine learning pattern recognition. *Nature Communications* **10** (Oct. 2019).

55. Wang, Z., Chong, Y., Joannopoulos, J. D. & Soljačić, M. Observation of unidirectional backscattering-immune topological electromagnetic states. *Nature* **461,** 772–775 (2009).

56. Rechtsman, M. C. *et al.* Photonic Floquet topological insulators. *Nature* **496,** 196–200 (2013).

57. Hafezi, M., Mittal, S., Fan, J., Migdall, A. & Taylor, J. M. Imaging topological edge states in silicon photonics. *Nature Photonics* **7,** 1001–1005 (2013).

58. Hafezi, M., Demler, E. A., Lukin, M. D. & Taylor, J. M. Robust optical delay lines with topological protection. *Nature Physics* **7,** 907–912 (2011).

59. Bandres, M. A. *et al.* Topological insulator laser: Experiments. *Science* **359,** eaar4005 (2018).

60. Wu, B. *et al.* Reconfigurable all-solid-state topological lasing at arbitrary sites. *Nature Communications* **17,** 1426 (2026).

61. Cui, X. *et al.* Insulator-free topological photonic multi-lane highways. *Nature* **656,** 73–79 (2026).

62. Dai, T. *et al.* A programmable topological photonic chip. *Nature Materials* **23,** 928–936 (2024).

63. Jia, R. *et al.* On-Chip Active Non-Reciprocal Topological Photonics. *Advanced Materials* **37,** 2501711. eprint: `https://advanced.onlinelibrary.wiley.com/doi/pdf/10.1002/adma.202501711` (2025).

64. Helbig, T. *et al.* Generalized bulk–boundary correspondence in non-Hermitian topolectrical circuits. *Nature Physics* **16,** 747–750 (2020).

65. Lee, S.-B. *et al.* Observation of an Exceptional Point in a Chaotic Optical Microcavity. *Physical Review Letters* **103,** 134101 (2009).

66. Chen, W., Özdemir, Ş. K., Zhao, G., Wiersig, J. & Yang, L. Exceptional points enhance sensing in an optical microcavity. *Nature* **548,** 192–196 (2017).

67. Hodaei, H. *et al.* Enhanced sensitivity at higher-order exceptional points. *Nature* **548,** 187–191 (2017).

68. Budich, J. C. & Bergholtz, E. J. Non-Hermitian Topological Sensors. *Phys. Rev. Lett.* **125,** 180403 (18 2020).

69. Bandres, M. A. *et al.* Topological insulator laser: Experiments. *Science* **359,** eaar4005. eprint: `https://www.science.org/doi/pdf/10.1126/science.aar4005` (2018).

70. Mao, W. *et al.* Laser mode braiding on a chip. *Nature Physics* **22,** 1344–1349 (2026).

71. Qin, H. *et al.* Photonic exceptional points in engineered materials and their emerging applications. *Nature Reviews Materials* **11,** 453–468 (2026).

72. Ling, X. *et al.* Topological edge plasmon modes between diatomic chains of plasmonic nanoparticles. *Proceedings of the National Academy of Sciences* **112,** E1536–E1541 (2015).

73. Pocock, S. R., Huidobro, P. A. & Giannini, V. Topological Plasmonic Chain with Retardation and Radiative Effects. *ACS Photonics* **5,** 2271–2279 (2018).

74. Yan, W., Faggiani, R. & Lalanne, P. Rigorous modal analysis of plasmonic nanoresonators. *Phys. Rev. B* **97,** 205422 (20 2018).

75. Moiseyev, N. *Non-Hermitian Quantum Mechanics* (Cambridge University Press, 2011).

76. Wiecha, P. R., Arbouet, A., Girard, C. & Muskens, O. L. Deep learning in nano-photonics: inverse design and beyond. *Photon. Res.* **9,** B182–B200 (2021).

77. Zhang, A., Lipton, Z. C., Li, M. & Smola, A. J. *Dive into Deep Learning* (Cambridge University Press, 2023).

78. Asano, T. & Noda, S. Optimization of Photonic Crystal Nanocavities Based on Deep Learning. *Optics Express* **27,** 32708–32717 (2019).

79. Bishop, C. M. *Pattern Recognition and Machine Learning* (Springer, 2006).

80. Murphy, K. P. *Probabilistic Machine Learning: An Introduction* (MIT Press, 2022).

81. Murphy, K. P. *Probabilistic Machine Learning: Advanced Topics* (MIT Press, 2023).

82. Shalev-Shwartz, S. & Ben-David, S. *Understanding Machine Learning: From Theory to Algorithms* (Cambridge University Press, 2014).

83. Goodfellow, I., Bengio, Y. & Courville, A. *Deep Learning* (MIT Press, 2016).

84. Bishop, C. M. & Bishop, H. *Deep Learning: Foundations and Concepts* (Springer, 2024).

85. Chollet, F. *Deep Learning with Python* 2nd ed. (Manning, 2021).

86. Geron, A. *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow* 3rd ed. (O'Reilly, 2022).

87. Gulli, A., Kapoor, A. & Pal, S. *Deep Learning with TensorFlow and Keras* 3rd ed. (Packt, 2022).

88. Ma, W. *et al.* Deep Learning for the Design of Photonic Structures. *Nature Photonics* **15,** 77–90 (2021).

89. Molesky, S. *et al.* Inverse design in nanophotonics. *Nature Photonics* **12,** 659–670 (2018).

90. Liu, D., Tan, Y., Khoram, E. & Yu, Z. Training Deep Neural Networks for the Inverse Design of Nanophotonic Structures. *ACS Photonics* **5,** 1365–1369 (2018).

91. Kudyshev, Z. A., Kildishev, A. V., Shalaev, V. M. & Boltasseva, A. in *Frontiers in Optics and Photonics* (eds Capasso, F. & Couwenberg, D.) 381–394 (De Gruyter, Berlin, Boston, 2021).

92. Hastie, T., Tibshirani, R. & Friedman, J. *The Elements of Statistical Learning* 2nd ed. (Springer, 2009).

93. Ashcroft, N. W. & Mermin, N. D. *Solid State Physics* (Holt, Rinehart and Winston, 1976).

94. Kittel, C. *Introduction to Solid State Physics* 8th ed. (John Wiley & Sons, 2005).

95. Yablonovitch, E. Inhibited Spontaneous Emission in Solid-State Physics and Electronics. *Phys. Rev. Lett.* **58,** 2059–2062 (20 1987).

96. John, S. Strong localization of photons in certain disordered dielectric superlattices. *Phys. Rev. Lett.* **58,** 2486–2489 (23 1987).

97. Joannopoulos, J. D., Johnson, S. G., Winn, J. N. & Meade, R. D. *Photonic Crystals: Molding the Flow of Light - Second Edition* REV - Revised, 2. ISBN: 9780691124568 (Princeton University Press, 2008).

98. Cheben, P., Halir, R., Schmid, J. H., Atwater, H. A. & Smith, D. R. Subwavelength integrated photonics. *Nature* **560,** 565 –572 (2018).

99. Fu, X. & Cui, T. Recent progress on metamaterials: From effective medium model to real-time information processing system. *Progress in Quantum Electronics* **67** (May 2019).

100. Luque González, J. M. *et al.* A review of silicon subwavelength gratings: Building break-through devices with anisotropic metamaterials. *Nanophotonics* **10** (Aug. 2021).

101. Garnett, J. C. M. Colours in Metal Glasses and in Metallic Films. *Philosophical Transactions of the Royal Society of London A* **203,** 385–420 (1904).

102. Cheben, P., Halir, R., Schmid, J., Atwater, H. & Smith, D. Subwavelength integrated photonics. *Nature* **560** (Aug. 2018).

103. Rytov, S. M. Electromagnetic Properties of a Finely Stratified Medium. *Soviet Physics JETP* **2,** 466–475 (1956).

104. Wen, X.-G. Colloquium: Zoo of quantum-topological phases of matter. *Rev. Mod. Phys.* **89,** 041004 (4 2017).

105. Royal Swedish Academy of Sciences. *Strange phenomena in matter's flatlands* Popular Science Background for the Nobel Prize in Physics. Accessed: 2026-06-30. 2016.

106. Haldane, F. D. M. Nobel Lecture: Topological quantum matter. *Rev. Mod. Phys.* **89,** 040502 (4 2017).

107. Klitzing, K. v., Dorda, G. & Pepper, M. New Method for High-Accuracy Determination of the Fine-Structure Constant Based on Quantized Hall Resistance. *Phys. Rev. Lett.* **45,** 494–497 (6 1980).

108. Raghu, S. & Haldane, F. D. M. Analogs of quantum-Hall-effect edge states in photonic crystals. *Phys. Rev. A* **78,** 033834 (3 2008).

109. Ozawa, T. *et al.* Topological photonics. *Rev. Mod. Phys.* **91,** 015006 (1 2019).

110. Dong, J., Juričić, V. & Roy, B. Topolectric circuits: Theory and construction. *Phys. Rev. Res.* **3,** 023056 (2021).

111. Lee, C. H. *et al.* Topolectrical Circuits. *Communications Physics* **1,** 39. arXiv: 1705.01077 [cond-mat.mes-hall] (Dec. 2018).

112. Kane, C. L. & Lubensky, T. C. Topological boundary modes in isostatic lattices. *Nature Physics* **10,** 39–45. arXiv: 1308.0554 [cond-mat.mes-hall] (Jan. 2014).

113. Süsstrunk, R. & Huber, S. D. Observation of phononic helical edge states in a mechanical topological insulator. *Science* **349,** 47–50. arXiv: 1503.06808 [cond-mat.mes-hall] (July 2015).

114. On, M. B. *et al.* Programmable integrated photonics for topological Hamiltonians. *Nature Communications* **15,** 4939 (2024).

115. Thouless, D. J., Kohmoto, M., Nightingale, M. P. & den Nijs, M. Quantized Hall Conductance in a Two-Dimensional Periodic Potential. *Phys. Rev. Lett.* **49,** 405–408 (6 1982).

116. Haldane, F. Continuum dynamics of the 1-D Heisenberg antiferromagnet: Identification with the O(3) nonlinear sigma model. *Physics Letters A* **93,** 464 –468. ISSN: 0375-9601 (1983).

117. Haldane, F. D. M. Nonlinear Field Theory of Large-Spin Heisenberg Antiferromagnets: Semiclassically Quantized Solitons of the One-Dimensional Easy-Axis Néel State. *Phys. Rev. Lett.* **50,** 1153–1156 (15 1983).

118. Pankratov, O. A., Pakhomov, S. V. & Volkov, B. A. Supersymmetry in heterojunctions: Band-inverting contact on the basis of Pb $_{1x}$Sn $_{x}$Te and Hg $_{1x}$Cd $_{x}$Te. *Solid State Communications* **61,** 93–96 (Jan. 1987).

119. König, M. *et al.* Quantum Spin Hall Insulator State in HgTe Quantum Wells. *Science* **318,** 766. arXiv: 0710.0582 (Nov. 2007).

120. Kane, C. L. & Mele, E. J. Z2 Topological Order and the Quantum Spin Hall Effect. *Phys. Rev. Lett.* **95,** 146802 (14 2005).

121. Bernevig, B. A., Hughes, T. L. & Zhang, S.-C. Quantum Spin Hall Effect and Topological Phase Transition in HgTe Quantum Wells. *Science* **314,** 1757. eprint: cond-mat/0611399 (Dec. 2006).

122. König, M. *et al.* Quantum Spin Hall Insulator State in HgTe Quantum Wells. *Science* **318,** 766–770. ISSN: 0036-8075. eprint: http://science.sciencemag.org/content/318/5851/766.full.pdf (2007).

123. Kotetes, P. *Topological Insulators* ISBN: 978-1-68174-517-6 (Morgan Claypool Publishers, 2019).

124. Asbóth, J. K., Oroszlány, L. & Pályi, A. *A Short Course on Topological Insulators: Band-structure topology and edge states in one and two dimensions* ISBN: 978-3-319-25605-4 (Springer, Berlin; Heidelberg, 2016).

125. Grabovsky, D. *The Limits of the Hubbard Model* `https://web.physics.ucsb.edu/~davidgrabovsky/files-research/Hubbard.pdf`.

126. Grushin, A. *Introduction to topological Phases in Condensed Matter* `https://grushingroup.cnrs.fr/wp-content/uploads/2021/05/intro_to_topo-2.pdf`.

127. Su, W. P., Schrieffer, J. R. & Heeger, A. J. Solitons in Polyacetylene. *Phys. Rev. Lett.* **42,** 1698–1701 (25 1979).

128. Yan, Q. *et al.* Advances and applications on non-Hermitian topological photonics. *Nanophotonics* **12,** 2247–2271 (2023).

129. Nasari, H., Pyrialakos, G. G., Christodoulides, D. N. & Khajavikhan, M. Non-Hermitian topological photonics. *Opt. Mater. Express* **13,** 870–885 (2023).

130. Wang, H. *et al.* Topological physics of non-Hermitian optics and photonics: a review. *Journal of Optics* **23,** 123001 (2021).

131. Okuma, N. & Sato, M. Non-Hermitian Topological Phenomena: A Review. *Annual Review of Condensed Matter Physics* **14,** 83–107. ISSN: 1947-5462 (2023).

132. Bergholtz, E. J., Budich, J. C. & Kunst, F. K. Exceptional topology of non-Hermitian systems. *Rev. Mod. Phys.* **93,** 015005 (1 2021).

133. Bi, R., Yan, Z., Lu, L. & Wang, Z. Nodal-knot semimetals. *Phys. Rev. B* **96,** 201305(R) (20 2017).

134. Carlström, J., Stålhammar, M., Budich, J. C. & Bergholtz, E. J. Knotted non-Hermitian metals. *Phys. Rev. B* **99,** 161115(R) (16 2019).

135. Bi, R., Yan, Z., Lu, L. & Wang, Z. Nodal-knot semimetals. *Phys. Rev. B* **96,** 201305(R) (20 2017).

136. Patil, Y. S. S. *et al.* Measuring the knot of non-Hermitian degeneracies and non-commuting braids. *Nature* **607,** 271–275 (2022).

137. Yang, Z., Chiu, C.-K., Fang, C. & Hu, J. Jones Polynomial and Knot Transitions in Hermitian and non-Hermitian Topological Semimetals. *Physical Review Letters* **124,** 186402 (2020).

138. Stålhammar, M., Rødland, L., Arone, G., Budich, J. C. & Bergholtz, E. J. Hyperbolic nodal band structures and knot invariants. *SciPost Phys.* **7,** 019 (2019).

139. Chang, P.-Y. Nodal-line semimetals and their variance. *Materials Today Quantum* **8,** 100057. ISSN: 2950-2578 (2025).

140. Bender, C. M. & Boettcher, S. Real Spectra in Non-Hermitian Hamiltonians Having PT Symmetry. *Phys. Rev. Lett.* **80,** 5243–5246 (24 1998).

141. Kunst, F. K., Edvardsson, E., Budich, J. C. & Bergholtz, E. J. Biorthogonal Bulk-Boundary Correspondence in Non-Hermitian Systems. *Phys. Rev. Lett.* **121,** 026808 (2 2018).

142. Ozawa, T. *et al.* Topological photonics. *Reviews of Modern Physics* **91,** 015006 (2019).

143. Lu, L., Joannopoulos, J. D. & Soljačić, M. Topological photonics. *Nature Photonics* **8,** 821–829 (2014).

144. Hashemi, A., Zakeri, M. J., Jung, P. S. & Blanco-Redondo, A. Topological quantum photonics. *APL Photonics* **10,** 010903 (2025).

145. Mehrabad, M. J., Mittal, S. & Hafezi, M. Topological photonics: fundamental concepts, recent developments, and future directions. *Physical Review A* **108,** 040101 (2023).

146. Smirnova, D. A. *et al.* Topological photonics. *JETP Letters* **114,** 981–990 (2021).

147. Xiao, M. *et al.* Recent advances in topological photonics. *Nature Communications* **11,** — (2020).

148. He, C., Sun, X.-C., Liu, X.-P., Lu, M.-H. & Chen, Y.-F. Acoustic topological insulator and robust one-way sound transport. *Nature Physics* **12,** 1124–1129 (2016).

149. Yang, Z. *et al.* Topological Acoustics. *Physical Review Letters* **114,** 114301 (2015).

150. Ni, X. *et al.* Topologically protected one-way edge mode in networks of acoustic resonators. *Nature Communications* **6,** 8539 (2015).

151. Süsstrunk, R. & Huber, S. D. Observation of phononic helical edge states in a mechanical topological insulator. *Science* **349,** 47–50 (2015).

152. Nash, L. M. *et al.* Topological mechanics of gyroscopic metamaterials. *Proceedings of the National Academy of Sciences* **112,** 14495–14500 (2015).

153. Paulose, J., ge Chen, B. G. & Vitelli, V. Topological modes bound to dislocations in mechanical metamaterials. *Nature Physics* **11,** 153–156 (2015).

154. Ningyuan, J., Owens, C., Sommer, A., Schuster, D. & Simon, J. Time- and Site-Resolved Dynamics in a Topological Circuit. *Physical Review X* **5,** 021031 (2015).

155. Imhof, S. *et al.* Topolectrical-circuit realization of topological corner modes. *Nature Physics* **14,** 925–929 (2018).

156. Sahin, H., Jalil, M. B. A. & Lee, C. H. Topolectrical circuits: Recent experimental advances and developments. *AIP Advances / APL Photonics (or similar AIP journal)* (2025).

157. Malkova, N., Hromada, I., Wang, X., Bryant, G. & Chen, Z. Observation of optical Shockley-like surface states in photonic superlattices. *Opt. Lett.* **34,** 1633–1635 (2009).

158. Parto, M., Liu, Y. G. N., Bahari, B., Khajavikhan, M. & Christodoulides, D. N. Non-Hermitian and topological photonics: optics at an exceptional point. *Nanophotonics* **10,** 403–423. eprint: `https://onlinelibrary.wiley.com/doi/pdf/10.1515/nanoph-2020-0434` (2021).

159. Parto, M. *et al.* Edge-Mode Lasing in 1D Topological Active Arrays. *Phys. Rev. Lett.* **120,** 113901 (11 2018).

160. Ochiai, T. Bulk-edge correspondence in open photonic systems. *Waves in Random and Complex Media* **34,** 951–968. eprint: `https://doi.org/10.1080/17455030.2021.1923858` (2024).

161. Luo, Y. & Sun, X. Topological bound states in the continuum in a non-Hermitian photonic system. *Nanophotonics* **14,** 43–50. eprint: `https://onlinelibrary.wiley.com/doi/pdf/10.1515/nanoph-2024-0419` (2025).

162. *Handbook of Optical Constants of Solids* (ed Palik, E. D.) ISBN: 9780125444200 (Academic Press, Orlando, FL, 1985).

163. Johnson, P. B. & Christy, R. W. Optical Constants of the Noble Metals. *Physical Review B* **6,** 4370–4379 (1972).

164. Pocock, S. R., Xiao, X., Huidobro, P. A. & Giannini, V. Topological Plasmonic Chain with Retardation and Radiative Effects. *ACS Photonics* **5,** 2271–2279. eprint: `https://doi.org/10.1021/acsphotonics.8b00117` (2018).

165. Sauvan, C., Wu, T., Zarouf, R., Muljarov, E. A. & Lalanne, P. Normalization, orthogonality, and completeness of quasinormal modes of open systems: the case of electromagnetism. *Optics Express* **30,** 6846–6885 (2022).

166. Sullivan, D. M. A Brief History of the Finite-Difference Time-Domain Method. *IEEE Antennas and Propagation Magazine* **55,** 290–296 (2013).

167. Teixeira, F. L., Sarris, C. D. & Simpson, J. J. Finite-Difference Time-Domain Methods. *Nature Reviews Methods Primers* **3,** 76 (2023).

168. Sullivan, D. M. A Brief History of the Finite-Difference Time-Domain Method. *IEEE Antennas and Propagation Magazine* **55,** 290–296 (2013).

169. Kogon, A. J. & Sarris, C. D. FDTD Modeling of Periodic Structures: A Review. *IEEE Journal on Multiscale and Multiphysics Computational Techniques* **5,** 229–245 (2020).

170. Taflove, A., Oskooi, A. & Johnson, S. G. *Advances in FDTD Computational Electrodynamics: Photonics and Nanotechnology* 2nd ed. (Artech House, 2024).

171. Kunz, K. S. & Luebbers, R. J. *The Finite Difference Time Domain Method for Electromagnetics* (CRC Press, Boca Raton, FL, 1993).

172. Sullivan, D. M. *Electromagnetic Simulation Using the FDTD Method* 2nd ed. (Wiley-IEEE Press, 2013).

173. Silvester, P. P. & Ferrari, R. L. *Finite Elements for Electrical Engineers* 3rd ed. ISBN: 9780521449533 (Cambridge University Press, Cambridge, UK, 1996).

174. Jin, J.-M. *The Finite Element Method in Electromagnetics* 3rd ed. (Wiley-IEEE Press, 2015).

175. Liu, W. K., Li, S. & Park, H. S. Eighty Years of the Finite Element Method: Birth, Evolution, and Future. *Archives of Computational Methods in Engineering* **29,** 2991–3039 (2022).

176. Bontinck, Z., Corno, J., De Gersem, H., *et al.* Recent Advances of Isogeometric Analysis in Computational Electromagnetics. *COMPEL* **37,** 545–572 (2018).

177. Jin, J.-M. *The Finite Element Method in Electromagnetics* 3rd ed. ISBN: 9781118571370 (John Wiley & Sons, 2014).

178. Monk, P. *Finite Element Methods for Maxwell's Equations* ISBN: 9780198508887 (Oxford University Press, 2003).

179. Brenner, S. C. & Scott, R. *The Mathematical Theory of Finite Element Methods* 3rd ed. (Springer, 2008).

180. Webb, J. P. Edge Elements and What They Can Do for You. *IEEE Transactions on Magnetics* **29,** 1460–1465 (1993).

181. Bossavit, A. *Computational Electromagnetism: Variational Formulations, Complementarity, Edge Elements* ISBN: 9780121187103 (Academic Press, 1998).

182. Munkres, J. R. *Elements of Algebraic Topology* (Addison-Wesley, Menlo Park, CA, 1984).

183. Yee, K. Numerical solution of initial boundary value problems involving maxwell's equations in isotropic media. *IEEE Transactions on Antennas and Propagation* **14,** 302–307 (1966).

184. Ansys, Inc. *Ansys Lumerical FDTD* `https://ansys.com`. Finite-Difference Time-Domain Electromagnetic Simulator. 2024.

185. Winslow, A. M. & Silvester, P. P. Magnetic Field Calculation Using Finite Elements. *Proceedings of the Institution of Electrical Engineers* **116,** 1964–1970 (1969).

186. Silvester, P. P. & Chari, M. V. K. Finite Element Solution of Saturable Magnetic Field Problems. *IEEE Transactions on Power Apparatus and Systems* **89,** 1642–1651 (1970).

187. Ferrari, R. L. The Finite-Element Method, Part 2: P. P. Silvester, an Innovator in Electromagnetic Numerical Modeling. *IEEE Antennas and Propagation Magazine* **49,** 242–246 (2007).

188. COMSOL AB. *COMSOL Multiphysics* `https://comsol.com`. Stockholm, Sweden. Finite Element Analysis Simulation Software. 2026.

189. Wu, T., Arrivault, D., Yan, W. & Lalanne, P. Modal analysis of electromagnetic resonators: User guide for the MAN program. *Computer Physics Communications* **284,** 108627. ISSN: 0010-4655 (2023).

190. Yan, W., Faggiani, R. & Lalanne, P. Rigorous modal analysis of plasmonic nanoresonators. *Phys. Rev. B* **97,** 205422 (20 2018).

191. Zhou, Z.-H. *Ensemble Methods: Foundations and Algorithms* ISBN: 9781439830031 (CRC Press, 2012).

192. Prince, S. J. D. *Understanding Deep Learning* (MIT Press, 2023).

193. Abadi, M. *et al.* TensorFlow: A System for Large-Scale Machine Learning. In *12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16)* 265–283 (2016).

194. Baydin, A. G., Pearlmutter, B. A., Radul, A. A. & Siskind, J. M. Automatic Differentiation in Machine Learning: A Survey. *Journal of Machine Learning Research* **18,** 1–43 (2018).

195. Tao, C., Zhu, J., Zhong, Y. & Liu, H. Coupling theory of quasinormal modes for lossy and dispersive plasmonic nanoresonators. *Physical Review B* **102,** 045430 (2020).

196. Wu, T., Arrivault, D. & Lalanne, P. Modal analysis of electromagnetic resonators: MAN software. *Computer Physics Communications* **284,** 108627 (2023).

197. Rappoport, T. G., Bludov, Y. V., Koppens, F. H. L. & Peres, N. M. R. Topological Graphene Plasmons in a Plasmonic Realization of the Su–Schrieffer–Heeger Model. *ACS Photonics* **8,** 1817–1823. ISSN: 2330-4022. eprint: `https://pubs.acs.org/apchd5/article-pdf/8/6/1817/4609219/ph1c00417.pdf` (May 2021).

198. Herz, F., Naeimi, A. & Biehs, S.-A. Substrate-driven topological engineering in plasmonic Su-Schrieffer-Heeger chains. *arXiv e-prints,* arXiv:2605.28463. arXiv: `2605.28463 [cond-mat.mes-hall]` (May 2026).

199. Sauvan, C., Wu, T., Zarouf, R., Muljarov, E. A. & Lalanne, P. Normalization, orthogonality, and completeness of quasinormal modes of open systems: the case of electromagnetism [Invited]. *Opt. Express* **30,** 6846–6885 (2022).

200. Keck, F, Korsch, H. & Mossmann, S. Unfolding a diabolic point: A generalized crossing scenario. *Journal of Physics A: Mathematical and General* **36,** 2125 (Feb. 2003).

201. Rudin, W. *Real and Complex Analysis* 3rd ed. (McGraw–Hill, New York, 1987).

202. Folland, G. B. *Fourier Analysis and Its Applications* (Wadsworth & Brooks/Cole, Pacific Grove, CA, 1992).

203. Teufel, S. *Adiabatic Perturbation Theory in Quantum Dynamics* (Springer, Berlin, 2003).

204. Scholz, E. in *Beyond Einstein: Perspectives on Geometry, Gravitation, and Cosmology in the Twentieth Century* (eds Rowe, D. E., Sauer, T. & Walter, S. A.) 261–360 (Birkhäuser / Springer, New York, 2018).

205. Blagojević, M. *Gravitation and Gauge Symmetries* (Institute of Physics, 2002).

206. Cheng, T.-P. *Relativity, Gravitation and Cosmology: A Basic Introduction* 2nd ed. ISBN: 9780199573646 (Oxford University Press, 2010).

207. Cui, X.-D. & Zheng, Y. Geometric phases in non-Hermitian quantum mechanics. *Phys. Rev. A* **86,** 064104 (6 2012).

208. Arkhipov, I. I. Covariant formulation of the Berry connection in non-Hermitian systems. *arXiv e-prints,* arXiv:2601.19777. arXiv: `2601.19777 [quant-ph]` (Jan. 2026).

209. Goenner, H. F. M. On the History of Unified Field Theories. *Living Reviews in Relativity* **7,** 1–151 (2004).

210. Ozawa, T. & Schomerus, H. Geometric contribution to adiabatic amplification in non-Hermitian systems. *Phys. Rev. Res.* **7,** 013173 (1 2025).

211. Aharonov, Y. & Bohm, D. Significance of Electromagnetic Potentials in the Quantum Theory. *Phys. Rev.* **115,** 485–491 (3 1959).

212. Berry, M. V. Quantal phase factors accompanying adiabatic changes. *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences* **392,** 45–57. ISSN: 0080-4630. eprint: `https://royalsocietypublishing.org/rspa/article-pdf/392/1802/45/65511/rspa.1984.0023.pdf` (Mar. 1984).

213. Aharonov, Y. & Anandan, J. Phase change during a cyclic quantum evolution. *Phys. Rev. Lett.* **58,** 1593–1596 (16 1987).

214. Stanescu, T. D. *Introduction to topological quantum matter and quantum computation* eng. ISBN: 9781315181509 (CRC Press, Taylor and Francis Group, Boca Raton, FL, 2017 - 2017).

215. Aharonov, Y. & Casher, A. Topological Quantum Effects for Neutral Particles. *Phys. Rev. Lett.* **53,** 319–321 (4 1984).

216. Sakurai, J. J. & Napolitano, J. *Modern Quantum Mechanics* 3rd. ISBN: 978-1-108-47322-4 (Cambridge University Press, 2020).

217. Kato, T. On the Adiabatic Theorem of Quantum Mechanics. *Journal of the Physical Society of Japan* **5,** 435–439. eprint: `https://doi.org/10.1143/JPSJ.5.435` (1950).

218. Childs, A. *Lecture notes in Quantum Algorithms: The Quantum Adiabatic Theorem* 2008.

219. Kalajdzievski, S. *An illustrated introduction to topology and homotopy* 1st ed. eng. ISBN: 1-00-307323-9 (Chapman Hall/CRC, Boca Raton, 2017).

220. Armstrong, M. *Basic Topology* 1st ed. 1983. eng. ISBN: 1-4757-1793-8 (Springer Nature, 1983).

221. Kalajdzievski, S. *An illustrated introduction to topology and homotopy* eng. ISBN: 9781439848159 (CRC Press, Taylor and Francis Group, Boca Raton, 2015).

222. Nakahara, M. *Geometry, Topology and Physics* 2nd ed. (Taylor & Francis, 2003).

223. Pal, P. *A Physicists Introduction to Algebraic Structures: Vector Spaces, Groups, Topological spaces and more* ISBN: 9781108492201 (Cambridge University Press, 2019).

224. Baez, J. C. *Gauge fields, knots, and gravity* eng. ISBN: 9810217293 (World Scientific, Singapore ; 1994).

225. Husemoller, D. *Fibre Bundles* 3rd ed. (Springer, 1994).