Title
Topological Inspirations in Photonic Devices

Author
Gong, Zilun

Publication Date
2020

Peer reviewed|Thesis/dissertation

Topological Inspirations in Photonic Devices

By

Zilun Gong

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

Engineering - Materials Science and Engineering

in the

Graduate Division

of the

University of California, Berkeley

Committee in Charge:

Professor Jie Yao, Chair

Professor Lane Martin

Professor Boubacar Kanté

Spring 2020

Abstract

Topological Inspirations in Photonic Devices

by

Zilun Gong

Doctor of Philosophy in Materials Science and Engineering

University of California, Berkeley

Professor Jie Yao, Chair

Material topology is an exotic degree of freedom in the condensed matter physics. It was initially proposed to explain the electron transportation in ferroelectric materials in 1954, and eventually made its way to the 2016 Nobel Prize in Physics. In recent decades, researchers are dedicated to transplanting the concepts in the condensed matter to the optics and photonics realm. Nowadays, the field of topological photonics is thriving, which allows us to mimic the behaviors of the electrons using photons and discover phenomena exclusive to the Bosonic systems.

In this dissertation, I present my work to show how we can get inspirations from the electronic system, then design the photonic device with new functionalities. Chapter 1 is dedicated to the cornerstone of material topology: geometric phase. The geometric phase determines the topological invariant in a crystal, but we can also use this concept to engineer the phase front in the diffractive optics devices. I show the first-ever metasurface for ultra-violet wavelengths using the geometric phase. In Chapter 2, I discuss the one-dimensional topological insulator and how to realize it in an optical waveguide array. Besides, I reveal the relation between the one-dimensional model and the topological edge states in the two-dimensional nanoribbons. It deepens our understanding of topological behaviors and the Bloch theorem. Finally, in Chapter 3, I investigate the bound-state-in-continuum, which is a topological singularity in the photonic crystals. Then I show the photonic integrated circuits that utilize this concept and result in a versatile optical filter. All the devices proposed are made of silicon, which is a promising material choice in terms of fabrication and scalability.

In general, I introduce the essential concepts in topological photonics and explain the physical pictures to my best knowledge, in the hope of inspiring readers to explore this field and design novel photonic devices.

Table of Contents

Introduction

In 2016, the Nobel prize of physics was awarded to three scientists for their contributions to the topological theory in solids[1]. The discovery of the topological insulator has been one of the biggest topics in the last decade[2]. The root of the topological phenomena, however, was discussed decades ago when people were trying to understand the abnormal electron velocity in the ferromagnetic materials[3]. The geometric phase, later referred to as the Berry phase, was proposed as an extra phase beside the propagation phase. People at that time did not realize that this concept would completely change the landscape of physics, materials science, even photonics in the twenty-first century[4].

The geometric phase is the extra phase that electrons pick up during an adiabatic process in the *k*-space. Originally, it was considered to be gauge-dependent and thus not observable. However, a closed-loop evolution gives rise to the gauge-independent Berry phase, which directly links to the "monopoles" of Berry flux in the reciprocal space[5]. This discovery defines a new degree of freedom in the energy band that is characterized by the topological invariant, i.e., the number of those "monopoles." The topological insulators with non-zero topological invariants support conducting channels on the edges of the material, even though the bulk material is an insulator[6,7]. Numerous researches have been dedicated to understanding different classes of topology based on symmetry arguments[8], observing various quantum Hall effects based on the theoretical predictions[9–14], and proposing devices with novel functionality such as ferroelectric conducting channels[15]. It is fair to say that the discussions of the material topology are one of the greatest topics in the last decade.

The photonic system, as a bosonic counterpart to the condensed matter, joined the cause as well[16–18]. In experiments, it is usually more convenient to construct an artificial lattice for photons since their wavelengths range from hundreds of nanometers to several microns. In a more general definition of electromagnetic waves, the microwave signals are also photons with centimeter-order wavelengths. Indeed, some early demonstrations of the topological photonics used metallic lattices for microwaves[19,20]. Later, more devices working at visible and near-infrared frequencies were proposed based on optical lattices[21–24]. In the wake of the photonic integrated circuit technology, it is reasonable to believe that topological devices will finally enter the realm of practical electrical engineering[25].

Throughout my Ph.D. researches, I have had the chance to investigate many topological behaviors in the photonic systems. In this dissertation, I will break down different aspects of my researches and how the topology arguments in the condensed matter could bring about new opportunities in photonic devices.

Chapter 1 will focus on the fundaments of topological phenomena, which is the geometric phase. I will discuss how to engineer the geometric phase of photons via optical antennas. The major result is a metasurface that works for ultraviolet frequencies. Based on the phase modulation, I will show applications, including beam steering and hologram. In this chapter, a circuit model will be proposed to explain the material choice and provide clues to enhance the efficiency of optical antennas.

Chapter 2 is about the central concepts of topology, which is the topological phase and edge states. Specifically, I will show a particular type of optical lattices, i.e., waveguide arrays, that can

demonstrate the one-dimensional topological insulator. Then a formalism will be introduced that maps the one-dimensional lattice to two-dimensional nanoribbon, with a realization of optical waveguide arrays. Essentially, it is an approach to connect physical systems with different dimensions. It provides a powerful tool to study some exotic behaviors in the solids with a controlled environment using photonic devices.

Chapter 3 is an attempt to generalize the topological phase in photonics and design a practical device. The typical topological phase originates from the singularity of the Berry flux. However, people have shown that the polarization singularities in the photonic crystal slabs, i.e., the "bound states in the continuum," also carry the topological charges similar to the topological insulators[26]. This chapter contains discussions of such bound states in the photonic crystal slabs and their realizations in the silicon photonics. Eventually, I want to show the feasibility of applying the concept of topology in an integrated photonics device.

I choose to touch upon a wide range of topics without getting too deep and buried with technical details. Overall, this dissertation is dedicated to providing clear physical pictures to readers who are hoping to carry out researches on topological photonics. Throughout three chapters, I will introduce many essential concepts and models, such as the Berry phase, the optical antenna model, the topological phase, the Su-Schrieffer-Heeger model, the optical waveguide array, photonic crystals, group theory, bound states in the continuum, and the coupled-mode theory. Hopefully, I shall give clear and accurate explanations to the best of my knowledge, provide thorough references, so that it could become a good starting point for the readers' future reading and study.

Also, it is worth mentioning that all the devices involved are based on silicon. Because of the constant development and refinement in silicon processing technologies, it is the best choice to design photonic devices using silicon or silicon-on-insulator wafers. This material is not only compatible with the "complementary metal oxide semiconductor" technology for mass productions but also has a transparency window in the near-infrared band, which is ideal for communication applications. Generally, silicon is the "future-proof" material for photonics.

At last, I sincerely hope that my knowledge, experiences, and results can inspire the readers to perform their own researches. In that case, all my efforts would be quite worthwhile.

References

1. Gibney, E. & Castelvecchi, D. Nobel for 2D exotic matter Physics award goes to theorists who used topology to explain strange phenomena. *Nature* **538**, 18 (2016).
2. Moore, J. E. The birth of topological insulators. *Nature* **464**, 194–198 (2010).
3. Karplus, R. & Luttinger, J. M. Hall effect in ferromagnetics. *Phys. Rev.* **95**, 1154–1160 (1954).
4. Berry, M. V. The adiabatic phase and Pancharatnam's phase for polarized light. *J. Mod. Opt.* **34**, 1401–1407 (1987).
5. Berry, M. V. Quantal phase factors accompanying adiabatic changes. *Proc. R. Soc. London. A. Math. Phys. Sci.* **392**, 45–57 (1984).
6. Asbóth, J. K., Oroszlány, L. & Pályi, A. *A Short Course on Topological Insulators*. vol. 919 (Springer International Publishing, 2016).

7. Haldane, F. D. M. Model for a quantum hall effect without landau levels: Condensed-matter realization of the 'parity anomaly'. *Phys. Rev. Lett.* **61**, 2015–2018 (1988).

8. Chiu, C. K., Teo, J. C. Y., Schnyder, A. P. & Ryu, S. Classification of topological quantum matter with symmetries. *Rev. Mod. Phys.* **88**, (2016).

9. Li, J., Chu, R. L., Jain, J. K. & Shen, S. Q. Topological anderson insulator. *Phys. Rev. Lett.* **102**, 1–4 (2009).

10. Burkov, A. A. & Balents, L. Weyl semimetal in a topological insulator multilayer. *Phys. Rev. Lett.* **107**, 1–4 (2011).

11. Fu, L. & Kane, C. L. Topological insulators with inversion symmetry. *Phys. Rev. B - Condens. Matter Mater. Phys.* **76**, 1–17 (2007).

12. Qi, X. L. & Zhang, S. C. The quantum spin Hall effect and topological insulators. *Phys. Today* **63**, 33–38 (2010).

13. Hsieh, D. *et al.* A topological Dirac insulator in a quantum spin Hall phase. *Nature* **452**, 970–974 (2008).

14. Chang, C. Z. *et al.* Experimental observation of the quantum anomalous Hall effect in a magnetic topological Insulator. *Science (80-.).* **340**, 167–170 (2013).

15. Son, Y. W., Cohen, M. L. & Louie, S. G. Half-metallic graphene nanoribbons. *Nature* **444**, 347–349 (2006).

16. Lu, L., Joannopoulos, J. D. & Soljačić, M. Topological photonics. *Nat. Photonics* **8**, 821–829 (2014).

17. Khanikaev, A. B. & Shvets, G. Two-dimensional topological photonics. *Nature Photonics* vol. 11 (2017).

18. Ozawa, T. *et al.* Topological photonics. *Rev. Mod. Phys.* **91**, 15006 (2019).

19. Poo, Y., Wu, R. X., Lin, Z., Yang, Y. & Chan, C. T. Experimental realization of self-guiding unidirectional electromagnetic edge states. *Phys. Rev. Lett.* **106**, 1–4 (2011).

20. Lu, L. *et al.* Experimental observation of Weyl points. *Science (80-.).* **349**, 622–624 (2015).

21. Wu, L. H. & Hu, X. Scheme for achieving a topological photonic crystal by using dielectric material. *Phys. Rev. Lett.* **114**, 1–5 (2015).

22. Rechtsman, M. C. *et al.* Photonic Floquet topological insulators. *Nature* **496**, 196–200 (2013).

23. Plotnik, Y. *et al.* Observation of unconventional edge states in 'photonic graphene'. *Nat. Mater.* **13**, 57–62 (2014).

24. Malkova, N., Hromada, I., Wang, X., Bryant, G. & Chen, Z. Observation of Optical Shockley-Like Surface States in Photonic Superlattices. *Opt. Lett.* **34**, 1633 (2009).

25. Mittal, S., Ganeshan, S., Fan, J., Vaezi, A. & Hafezi, M. Measurement of topological invariants in a 2D photonic system. *Nat. Photonics* **10**, 180–183 (2016).

26. Zhen, B., Hsu, C. W., Lu, L., Stone, A. D. & Soljačić, M. Topological nature of optical bound states in the continuum. *Phys. Rev. Lett.* **113**, 1–5 (2014).

Acknowledgments

It is quite a journey when I look back at this point. It is a true privilege to join one of the most renowned universities in the world, spend five years live and study on the same campus that has witnessed so many great figures.

I would like first to thank my advisor Professor Jie Yao, who guided me through all the challenges and shared all my excitement. My projects would not be possible without my collaborators: Dr. André Anders from Lawrence Berkeley National Lab (now in Leibniz Institute of Surface Engineering), Professor Chris Bartone from Ohio University, and Professor Stefan Preble from Rochester Institute of Technology. The academic life would not be complete without my colleagues and friends in the Yao Group. I consider myself very lucky to know these amazing people that we can learn and progress together.

Finally, I cannot thank my parents enough for their unconditional support, not only for the last five years but for my entire life. They are and will always be my inspiration.

Chapter 1: Control the Geometric Phase of Light

§1.1 Introduction of the geometric phase in crystals

In recent years, the topological properties of materials have attracted much attention to the research community[1–10]. The pioneers of this field also won the 2016 Nobel Prize in Physics[11–17]. The field of topology was substantially developed in mathematics in the 20th century. The first successful implementation of topology in condensed matter physics is related to the explanation of anomalous velocity of Bloch electrons. From the Bloch theorem, the semiclassical motion of electron follows the velocity of Bloch wavepacket[18]:

$$\boldsymbol{v}(\boldsymbol{k}) = \dot{\boldsymbol{r}} = \frac{1}{\hbar}\nabla_{\boldsymbol{k}} E_n(\boldsymbol{k}) \tag{1-1}$$

where $\boldsymbol{k}$ is the wavevector, E_n is the nth electron band. However, Karplus and Luttinger pointed out that in ferromagnetic materials, there is an additional term, i.e., anomalous velocity, to the equation[19,20]. The complete equation of motion is:

$$\boldsymbol{v}(\boldsymbol{k}) = \dot{\boldsymbol{r}} = \frac{1}{\hbar}\nabla_{\boldsymbol{k}} E_n(\boldsymbol{k}) + \dot{\boldsymbol{k}} \times (\nabla_{\boldsymbol{k}} \times \boldsymbol{A}_{\boldsymbol{k}}) \tag{1-2}$$

where $\boldsymbol{A}_{\boldsymbol{k}}$ is defined by

$$\boldsymbol{A}_{\boldsymbol{k}} = \frac{1}{i}\int_{\Omega} d^3\boldsymbol{r} \cdot u^*_{\boldsymbol{k}}(\mathbf{r})\nabla_{\boldsymbol{k}} u_{\boldsymbol{k}}(\boldsymbol{r}) \tag{1-3}$$

where $u_{\boldsymbol{k}}$ is the Bloch function, and the integral is taken over the unit cell Ω.

The anomalous velocity explains the extraordinary Hall coefficient of ferromagnetic materials. Phenomenologically, $\boldsymbol{A}_{\boldsymbol{k}}$ is an effective vector potential in the k-space similar to that of an external magnetic field in real space: $\boldsymbol{B} = \nabla \times \boldsymbol{A}$. Besides the extraordinary Hall coefficient, this result has more profound indications.

In 1995 and 1996, Chang and Niu re-derived the anomalous velocity of Bloch electrons[21]. They pointed out that $\boldsymbol{A}_{\boldsymbol{k}}$ is an important quantity called "Berry connection." It is related to the geometric phase, i.e., Pancharatnam–Berry phase, or "Berry phase," that the electron picks up in the n^{th} electron band during an adiabatic process[22,23]:

$$\gamma_n = \int_{\boldsymbol{l}} d\boldsymbol{l} \cdot \nabla_{\boldsymbol{k}} \times \boldsymbol{A}_{\boldsymbol{k}} \tag{1-4}$$

where $\boldsymbol{l}$ is an adiabatic path in the parameter space (specifically the k-space). It is worth noting that local $\boldsymbol{A}_{\boldsymbol{k}}$ is a gauge-dependent quantity; therefore, it is not observable. However, if $\boldsymbol{l}$ is a closed-loop $\boldsymbol{c}$ in the parameter space, the Berry phase becomes observable and well-defined. The real-space counterpart of the Berry phase gives rise to the famous Aharonov–Bohm effect[24], where the geometric phase is observable from interference when electrons travel around a magnetic flux.

The most extraordinary feature of the geometric phase in the condensed matter is quantization. Specifically, Berry phase γ_n is an integer multiple of 2π, $\gamma_n = m \cdot 2\pi$, and m is the topological invariant of the n^{th} electron band. In other words, the geometric phase one electron picks up in the

closed-loop adiabatic process is only determined by the topology of the band. This discovery opened up a whole field of topological insulators, topological classifications, and corresponding bandstructure engineering[25].

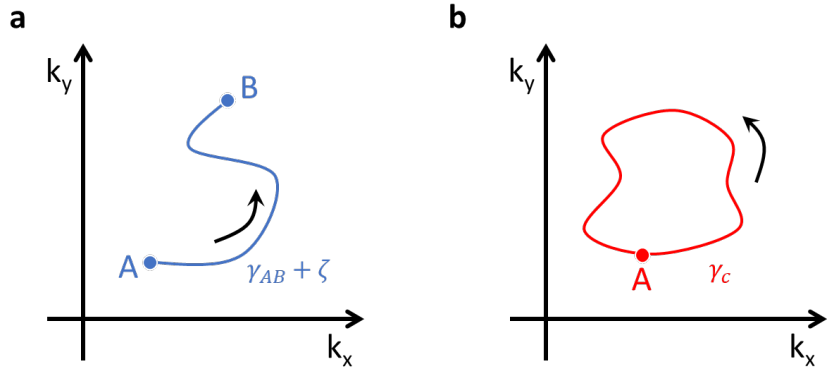

Figure 1-1 | a. The geometric phase γ_{AB} is a path-dependent phase when system evolves from A to B in the *k*-space. ζ is the gauge-dependent phase. **b.** The geometric phase is observable (gauge independent) if the path forms a closed-loop.

Following the thriving discussions of topological properties in the condensed matter physics, people also turned to the photonic system for their counterparts, and the natural platform is photonic crystals[26]. Photonic crystals are artificial periodic structures that mimic the periodic potentials in condensed matter. Because of the periodic scattering of the light, the photons behave similarly as electrons in crystals, while the resonance mode is the counterpart of the Bloch state. As a result, photonic crystals have their own "bandstructure" as well, which can be engineered thanks to the long wavelengths of photons compared to electrons.

If we engineer the bandstructure of the photonic crystals, potentially, we can create the non-trivial topological states at the boundary of the photonic crystals, similar to topological insulators. There have been many successful demonstrations of the edge states in photonic crystals. Readers can refer to the topological photonics review and the references therein[27–29]. We need to point out that there are some fundamental differences between photonic crystals and solid-state crystals, mainly because electrons are fermions while photons are bosons. It affects some important symmetry arguments that are utterly important for the topological phenomena. For example, the spin quantum Hall effect in electronic systems is protected by the time-reversal symmetry. Breaking of the time-reversal symmetry will lead to the breaking of spin channels and the emergence of charge channels at the boundary. On the contrary, photonic crystals must break, or effectively break, the time-reversal symmetry to demonstrate the spin quantum Hall effect, where circular polarizations are pseudo-spins of photons.

§1.2 Metasurface and phase control using geometric phase

Besides the Berry phase hidden in the bandstructure of photonic crystals, another application of the geometric phase in photonic systems is metasurface[30,31]. In general, metasurface refers to a category of devices that encode spatially varying optical response in a thin layer. The target results include amplitude, polarization, and phase manipulation, while the phase control is the most common scenario to date. Optical engineers have been engineering the phase front since the first

optical lens was made. One of the primary missions of the metasurface is to thin down the bulky lens and other optical components, so that beam steering, imaging, or even hologram can be achieved compactly. Readers can refer to the review papers and references therein. We would like to limit the following discussions to the phase control.

The general idea of phase control via metasurface is realized by discrete "meta-atoms" that are usually subwavelength in size. Each meta-atom introduces a local phase shift as a function of incident polarization by design; therefore, the metasurface creates an artificial wavefront. Note that the meta-atoms are not supposed to crosstalk or "see" each other in a metasurface; otherwise, the device would be working as a photonic crystal or photonic crystal slab.

Since each meta-atom works independently, we can model one meta-atom using a Jones matrix of a conventional, linearly birefringent wave plate[32]:

$$J = R(-\theta)\begin{pmatrix} e^{i\varphi_x} & 0 \\ 0 & e^{i\varphi_y} \end{pmatrix} R(\theta) \tag{1-5}$$

where R is the rotation matrix, θ is the angel between local coordinates of the meta-atom and the laboratory coordinates. The meta-atom imposes phase shift of φ_x and φ_y on light linearly polarized along its "fast" and "slow" optical axis, respectively. This Jones matrix assumes the basis of orthogonal polarizations, which can be two linear polarizations, two circular polarizations, or even two elliptical polarization if they are orthogonal.

From the Jones matrix, there are two approaches to meta-atom construction: control rotation angle and control local phase shift, which bring about geometric phase shift and propagation phase shift, respectively. Table.1-1 summarizes the differences between these two approaches. We should point out that these two approaches are not exclusive to each other. One can combine those two methods to achieve versatile functionalities[32].

Propagation Phase Control	**Geometric Phase Control**
Change φ_x and φ_y	Change θ
Independent control over two linear polarization incidence	Equal and opposite control over two circular polarization incidence
Based on resonance inside meta-atom	Based on geometric (Berry) phase
Oval-shape cross-section, waveguide-like along the propagation direction	Bar-shape cross-section, negligible depth along the propagation direction

Table 1-1 | The comparison between propagation phase control and geometric phase control in metasurfaces.

The propagation phase control is equivalent to introducing local delay lines at each meta-atom position. One can engineer the geometry of each meta-atom to construct the desired phase profile for a particular linear polarization incidence. The geometric phase control, on the other hand, takes advantage of the Berry phase upon polarization change[33,34]. In contrast to the Berry phase in the

condensed matter, where the electron evolves in the *k*-space, here the parameter space is the polarization space of light, which can be visualized using the Poincaré sphere.

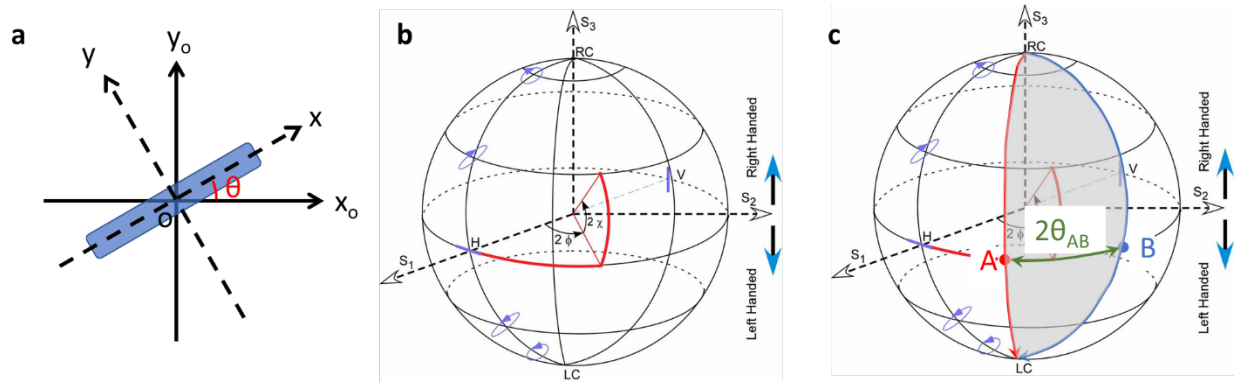

Figure 1-2 | a. The angle between local coordinates of the meta-atom and the laboratory coordinates is θ. **b.** The Poincaré sphere allows us to describe any polarization using the spherical coordinates ϕ and χ. **c.** Two meta-atoms with θ_A and θ_B result in the polarization going through point A and B on the equator of the Poincaré sphere, respectively. The relative geometric phase equals half of the solid angle of the shaded area, the polar angle of which is $2|\theta_A - \theta_B| = 2\theta_{AB}$.

The Poincaré sphere is a parameter space that represents all possible polarizations of light, where the completely polarized states lie on the surface of the sphere. The north pole and south pole represent right-hand circular polarization (RCP) and left-hand circular polarization (LCP), respectively. These two polarizations span the parameter space in a similar way to the spin representation in quantum mechanics. RCP and LCP are two bases of the Poincaré sphere so that we can use two base vectors $\widehat{e_1}$ and $\widehat{e_2}$ to represent them, respectively. Then polarization status can be written using the spherical coordinates[35]:

$$\hat{e} = \sin(2\chi) \cdot \widehat{e_1} + e^{i(2\phi)} \cos(2\chi) \cdot \widehat{e_2} \tag{1-6}$$

where 2χ and 2ϕ are azimuthal angle and polar angle, respectively. The factors of two correspond to the fact that any polarization ellipse is indistinguishable from one rotated by 180°. If the meta-atoms we use are bar structures that only respond to the polarization aligned with itself, then the polar angle ϕ is exactly the rotation angle θ of the meta-atom relative to the laboratory coordinates.

The function of the meta-atoms on the geometric-phase metasurface is to change the polarization of incident light from one pole to the other via a specific point on the equator. The rotation angle θ of that meta-atom determines the path of polarization change, that polar angle ψ must match θ. If we can design the spatial distribution of meta-atoms, we can control the path of polarization change at each point. According to the theory of Pancharatnam–Berry phase, two different paths (polar angle difference being $2\theta_{AB}$) will introduce a phase difference, which is precisely half of the solid angle of the enclosed surface[33]:

$$\gamma_{AB} = \frac{1}{2}\Omega_{AB} = \frac{1}{2}\iint \frac{ds}{r^2} = \frac{1}{2} \cdot \frac{2\theta_{AB} \cdot 4\pi r^2}{2\pi \cdot r^2} = 2\theta_{AB} \tag{1-7}$$

In conclusion, the phase difference between two meta-atoms is simply twice the rotation-angle difference. It brings much convenience to device designs.

From a mathematic perspective, the phase shift imposed by a meta-atom should have π difference between φ_x and φ_y. Without loss of generality, we can assume the overall Jones matrix is:

$$J(\theta) = \begin{pmatrix} cos\theta & sin\theta \\ -sin\theta & cos\theta \end{pmatrix} \begin{pmatrix} e^0 & 0 \\ 0 & e^{i\pi} \end{pmatrix} \begin{pmatrix} cos\theta & -sin\theta \\ sin\theta & cos\theta \end{pmatrix} = \begin{pmatrix} cos2\theta & -sin2\theta \\ -sin2\theta & -cos2\theta \end{pmatrix} \tag{1-8}$$

The LCP and RCP polarization under cartesian coordinates are:

$$\begin{cases} |L\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ +i \end{pmatrix} \\ |R\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix} \end{cases} \tag{1-9}$$

Therefore, imposing the Jones matrix to the circular-polarized light yields:

$$J(\theta) \cdot \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ \pm i \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} cos2\theta \mp i \cdot sin2\theta \\ \mp i \cdot cos2\theta - sin2\theta \end{pmatrix} = e^{\mp i \cdot 2\theta} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ \mp i \end{pmatrix} \tag{1-10}$$

Or:

$$\begin{cases} J(\theta)|L\rangle = e^{-i \cdot 2\theta}|R\rangle \\ J(\theta)|R\rangle = e^{+i \cdot 2\theta}|L\rangle \end{cases} \tag{1-11}$$

In other words, the meta-atom can change the polarization with a 2θ additional phase. LCP and RCP incidence will have opposite phase delay/advance. This result is consistent with that derived from the Poincaré sphere.

§1.3 The plasmonic response of optical antennas

Because the meta-atoms scatter light, in other words, they "radiate" energy, we shall refer to them as optical antennas. Firstly, we need to understand how optical antennas work under external drive.

Optical antennas are antennas for light. The term "antenna" was directly borrowed from the familiar radio frequency (RF) antennas. We shall first examine a typical RF dipole antenna[36] in Figure 1-3.

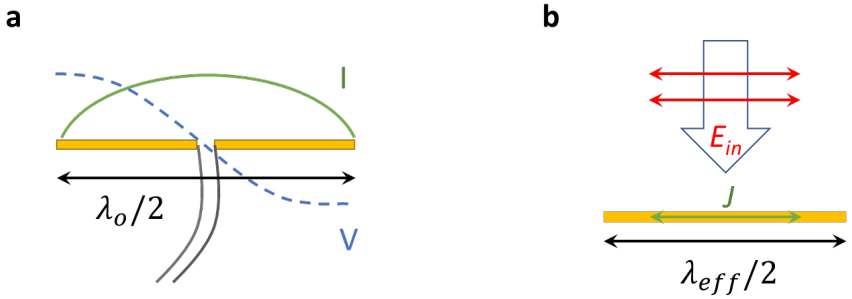

Figure 1-3 | a. Half-wave dipole antenna model for RF. **b.** Optical antenna model with effective wavelength scaling.

The dipole antenna is made of two metallic arms, with a total length of half the wavelength of the free-space electromagnetic (EM) wave. Note that here we only consider air as the environment, as the dielectric loading will increase the effective length of the antenna[37]. Therefore, the resonance condition can be satisfied with the current maxima at the center, where the feed connects to a coaxial cable. However, this description has a significant yet less discussed assumption, that the wavelength of electron resonance on the antenna equals the wavelength of EM wave in the free space. Unfortunately, this assumption is only correct conditionally.

We must understand this behavior from the plasma perspective. When the incident radiation reaches the antenna, it drives the free electrons in the metal. Typically, the free electron gas has its characteristic plasma frequency as a function of electron density n:

$$\omega_p = \sqrt{\frac{ne^2}{m^*\varepsilon_o}} \tag{1-12}$$

where e is the electron charge, m^* is the electron effective mass, and ε_o is the dielectric constant of vacuum[38]. For a typical metal, the plasma frequency is at the order of 10^{16} Hz or 10^4 THz, far beyond the RF frequency that usually ranges from 0.1 GHz to 10 GHz.

The plasma frequency describes how well the electron "keeps up" with the incident EM wave. When the incident frequency is very low compared to plasma frequency, such as the RF antenna case, the electrons in the metal can completely keep up with the incident wave. It is worth noting that under low-frequency condition (DC-limit), the metal behaves as a "good conductor" because induced current $\boldsymbol{J}$ and incident electric field $\boldsymbol{E_{in}}$ has π/2 phase difference, so the dissipation energy power is $\frac{1}{2}Re(\boldsymbol{J}^* \cdot \boldsymbol{E_{in}}) = 0$.

However, when the driving frequency increases, the electrons will have trouble "keeping up" with the driving field because of their inertia, and the phase difference will decrease. Therefore, the dielectric loss will emerge (here we still refer to this dissipation as "dielectric loss" even though the object is metallic because we intend to include the behavior into the complex permittivity). If we refer to the dissipation power equation above, the two vectors are no longer perpendicular to each other. Eventually, if the driving frequency reaches the plasma frequency of the free electron gas, the phase difference becomes 0, all incident energy will be lost to heat dissipation.

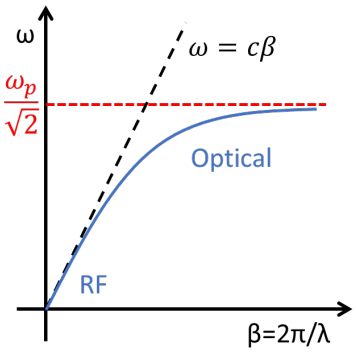

Figure 1-4 | Dispersion of surface plasma polariton (blue curve) assuming environment is vacuum. Black dashed line is the dispersion in free space.

To be more specific, the electron oscillation on the antenna is called surface plasma polariton, or SPP[39]. It describes the collective oscillation, i.e., quasi-particle, on the surface of metal when coupled with EM wave. SPP is different from bulk plasma, where the electrons oscillate inside the metal body, mainly because of the dispersion relation. From the schematic dispersion plot shown in Figure 1-4, at the long-wavelength limit, the SPP mode is very close to the free space dispersion line. Therefore, it explains why the wavelength of SPP on the RF antenna equals the wavelength of the EM wave in the free space.

However, as the frequency rises, the SPP dispersion deviates from the free space and approaching the asymptote, which the surface plasma frequency ω_{sp}:

$$\omega_{sp} = \frac{\omega_p}{\sqrt{1+\varepsilon_{env}}} \tag{1-13}$$

where ε_{env} is the relative permittivity of the environment. If we assume the environment is air, then the relation is simply $\omega_{sp} = \frac{\omega_p}{\sqrt{2}}$.

It shows that when the frequency approaches ω_{sp}, the propagation constant can be huge, and the SPP wavelength is significantly smaller than λ_o. Thus, if we design a "half-wave antenna" in the optical frequency with metals, the antenna length should be smaller than $\lambda_o/2$. In the hope of carrying out the same methodology as in RF antenna design, the optical antenna community came up with an empirical approach called "effective wavelength scaling"[40]. It considers the inertia of electron at high frequency and scales the λ_o to an effective wavelength λ_{eff}:

$$\lambda_{eff} = c_1 + c_2\left(\frac{\lambda_p}{\lambda_o}\right) \tag{1-14}$$

where λ_p is the plasma wavelength, c_1 and c_2 are fitting parameters.

This method was well-received, especially when designing metallic optical antennas for near-infrared and visible frequencies. However, it fails to deliver a physical picture in terms of how the change of permittivity affects the wavelength scaling. Also, dielectric materials were not included, which makes this method less universal. The RF antenna theory is beyond just half-wave resonance. In our works, we intend to build a more generic and specific model that is based on the RF antenna theory but works for both metal and dielectric across the spectrum. Hopefully, it provides a new perspective for optical antenna modeling.

Before we discuss the antenna model, it is beneficial to clarify the conventions of complex (relative) permittivity and complex conductivity. They both can be written in forms of the real part and the imaginary part:

$$\varepsilon = \varepsilon_r + i\varepsilon_i \; ; \; \sigma = \sigma_r + i\sigma_i \tag{1-15}$$

In general, the goal is to describe the electrons' response under EM excitations. As discussed earlier in this section, the response can be separated into two parts: one is called "displacement current" that has $\pi/2$ phase difference with the incident field, and it does not introduce loss; the other is called "conduction current" that is in-phase with the incident field, and it causes conduction and energy dissipation.

There are three conventions to describe this behavior, and they can be readily converted to each other using Ampere's law:

$$\nabla \times \boldsymbol{H} = \overbrace{-i\omega\varepsilon_0\varepsilon_r\boldsymbol{E}}^{\text{Displacement}} + \overbrace{\sigma_r\boldsymbol{E}}^{\text{Conduction}} = -i\omega\varepsilon_0\varepsilon\boldsymbol{E} = -i\omega\varepsilon_0\boldsymbol{E} + \sigma\boldsymbol{E} \tag{1-16}$$

	ε_r	ε_i	σ_r	σ_i
Convention 1	Displacement	0	Conduction	0
Convention 2	Displacement	$\varepsilon_i = \sigma_r/(\omega \cdot \varepsilon_o)$	0	0
Convention 3	0	0	Conduction	$\sigma_i = \omega\varepsilon_o(1-\varepsilon_r)$

Table 1-2 | The three conventions describing displacement and conduction currents.

First, use ε_r to describe displacement current, use σ_r to describe conduction current. It is the most intuitive method. Second, use ε_r to describe displacement current, use ε_i to describe conduction current. It is convenient for dielectrics where the loss is small. Third, use σ_i to describe displacement current, use σ_r to describe conduction current. It is popular among studies of metals under low frequency, where the conduction dominates. Table 1-2 summarizes the three conventions and their conversions. We will use the second convention throughout the following discussions.

§1.4 The circuit model of optical antennas

Achieving a high-performance metasurface requires investigating the scattering efficiency of an optical antenna[41], and we need a mathematical model to do so. The optical antenna model we investigate has the shape of a nanorod. Assuming the cross-section is small compared to length, we can apply the thin-wire antenna model borrowed from RF antennas[36].

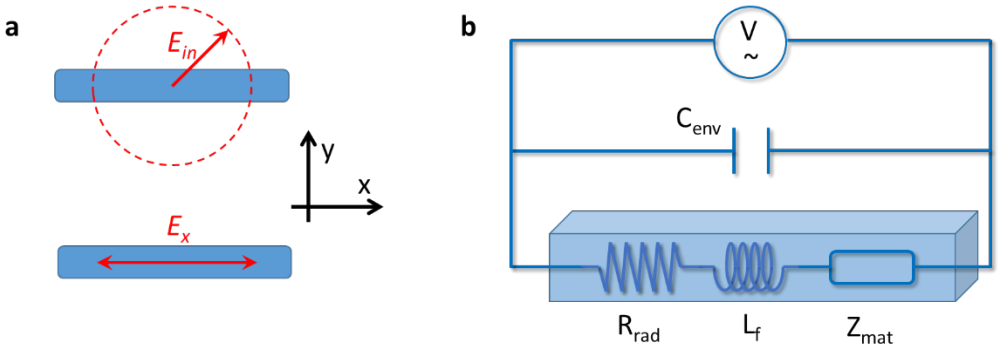

Figure 1-5 | a. The circular-polarization can be modeled as linear-polarization along the antenna under thin-wire model. **b.** The circuit model of the optical antenna in the shape of a nanorod.

Since the cross-section is small (thin-wire), we shall only consider the responses to the electrical field that is aligned with the bar. As discussed in previous sections, the incident light is circular-polarized. Because the circular-polarized light can be decomposed to two linear-polarized light, it is equivalent to consider the excitation with linear-polarized incidence $\boldsymbol{E_x}$ (Figure.1-5(a)). The scattered $\boldsymbol{E_x}$ will have a π phase delay and become the cross-polarization of the output (with the un-modulated $\boldsymbol{E_y}$ component). Therefore, the total efficiency of the optical antenna is the scattering efficiency of the $\boldsymbol{E_x}$ component. In this section, we will only consider linear excitation.

Parameter	Definition
C_{env}	Capacitance from the environment
R_{rad}	Radiation resistance
L_f	Self-inductance
Z_{mat}	Material impedance
Z_o	Vacuum impedance, 377 Ω
P_{rad}	Radiation (Scattering) power
l, w, h	Length, width, and height of nanorod
a	Lattice constant in metasurface

Table 1-3 | The definitions of parameters used in the circuit model.

To simplify the derivation, we define all the parameters in Table.1-3. The circuit model[42] for the optical antenna is shown in Figure.1-5. In general, we inherit the radiation resistance and self-inductance as in the RF antenna circuit model. The critical modification is the material impedance since now the antenna material is no longer a good conductor nor a lossless dielectric. We have to include a more general term Z_{mat} to describe the material response.

The incident field is equivalent to an AC voltage source applied to the nanorod:

$$\mathrm{V} = l \cdot E_x \tag{1-17}$$

R_{rad} and L_f are only determined by the antenna geometry[43]:

$$R_{rad} = \frac{Z_o}{2\pi sin^2(\frac{kl}{2})}\left\{\gamma_e + \ln(kl) - Ci(kl) + \frac{1}{2}\sin(kl)\,[Si(2kl) - 2Si(kl)] + \frac{1}{2}\cos(kl)[Ci(2kl) - 2Ci(kl) + \gamma_e + ln\left(\frac{kl}{2}\right)]\right\}$$

$$L_f = \frac{Z_o}{2\pi\omega sin^2(\frac{kl}{2})}\left\{Si(kl) + \frac{1}{2}\cos(kl)\,[2Si(kl) - Si(2kl)] + \frac{1}{2}\sin(kl)[Ci(2kl) - 2Ci(kl) + Ci\left(\frac{2kr_{eff}^2}{l}\right)]\right\} \tag{1-18}$$

where *Si(x)* and *Ci(x)* are sine and cosine integral functions, respectively, γ_e is the Euler's constant ($\gamma_e = 0.57721566\ldots$), r_{eff} is the effective radius of the nanorod that can be fitted, k is the wavevector in the vacuum. We can also combine these two parts as radiation impedance:

$$Z_{rad} = R_{rad} - i\omega L_f \tag{1-19}$$

The material impedance (i.e., complex resistance) is calculated from complex conductivity:

$$\nabla \times \boldsymbol{H} = -i\omega\varepsilon_0\varepsilon\boldsymbol{E} = -i\omega\varepsilon_0\boldsymbol{E} + \sigma\boldsymbol{E} \tag{1-20}$$

Therefore,

$$Z_{mat} = \frac{1}{\sigma}\cdot\frac{l}{A} = \frac{i}{\omega(\varepsilon-1)}\cdot\frac{l}{w\cdot h} \tag{1-21}$$

where A is the nanorod cross-section, $Re(Z_{mat})$ is the familiar Ohmic resistance from ε_i, while $Im(Z_{mat})$ can be positive or negative depends on ε_r. If $\varepsilon_r>1$, the material is capacitive, which is typical for dielectrics; if $\varepsilon_r<1$, the material is inductive, which is typical for metals. Therefore, this circuit model accommodates for both metal and dielectrics, which is better than effective wavelength scaling.

The power scattered from the nanorod is equal to the power "dissipated" at R_{rad}, which is:

$$P_{rad} = (\frac{V}{Z_{rad}+Z_{mat}})^2\cdot R_{rad} = (\frac{E_x\cdot l}{Z_{rad}+Z_{mat}})^2\cdot R_{rad} \tag{1-22}$$

In the real device, the substrate is silica with a refractive index of 1.5. According to Fermi's Golden Rule[18], the scattering probability is proportional to the refractive index of the environment. Correcting that factor gives us the forward scattering power to the air:

$$P_{rad}^{forward} = (\frac{E_x\cdot l}{Z_{rad}+Z_{mat}})^2\cdot R_{rad}\cdot\frac{1}{2.5} \tag{1-23}$$

The incident power on each unit cell can be calculated from the time-averaged Poynting vector:

$$P_{in}^{cell} = \frac{1}{Z_o}E_x^2\cdot a^2 \tag{1-24}$$

The total forward scattering efficiency:

$$\eta_{forward} = \frac{P_{rad}^{forward}}{P_{in}^{cell}} = \frac{1}{2.5}\times\left(\frac{l}{a}\right)^2\cdot\frac{R_{rad}\cdot Z_o}{(Z_{rad}+Z_{mat})^2} \tag{1-25}$$

$\eta_{forward}$ is also the expected conversion efficiency of the metasurface.

We can also express the efficiency of a single antenna in terms of scattering cross-section:

$$\sigma_s = \frac{P_{rad}^{forward}}{P_{rod}} = \frac{1}{2.5}\times\frac{l}{w}\cdot\frac{R_{rad}\cdot Z_o}{(Z_{rad}+Z_{mat})^2} \tag{1-26}$$

where is P_{rod} is the incident power on the footprint of the antenna:

$$P_{rod} = \frac{1}{Z_o}E_x^2\cdot l\cdot w \tag{1-27}$$

§1.4 Choice of material and fabrication

Previously, the metasurface based on the geometric phase had many successful implementations targeting a broad spectrum range from THz to visible[44–49]. The primary materials of choice are metals with good conductivity, such as gold (Au), silver (Ag), and aluminum (Al). From the

plasmonic perspective, high electron density leads to high plasma frequency, which needs to be significantly higher than the incident light to result in functional efficiency. However, this formula does not work for ultraviolet (UV) light (λ<400 nm) because of material limitations. Gold has an interband transition that makes it very absorbing for light whose wavelength is short than 500 nm. Silver has deteriorated performance for wavelengths small than 310 nm as well. Al potentially has a better performance for UV, but it suffers from oxidation-induced degradation and fabrication challenges because of large grain size[50]. Therefore, the demonstration of metasurface for UV based on the geometric phase was lacking.

Based on the circuit model argument, however, the high efficiency does not require a metallic material. If a dielectric material with ε_r>1 but has a low overall impedance Z_{mat}, then the radiation power will still be high. Without loss of generality, we use the circuit model to calculate the scattering cross-section of a single antenna as a function of material permittivity ε. Note that the incident wavelength is 380 nm, the length, width, and height of the nanorod are 130 nm, 30 nm, and 30 nm, respectively. These parameters are chosen to represent the actual device used in the experiment.

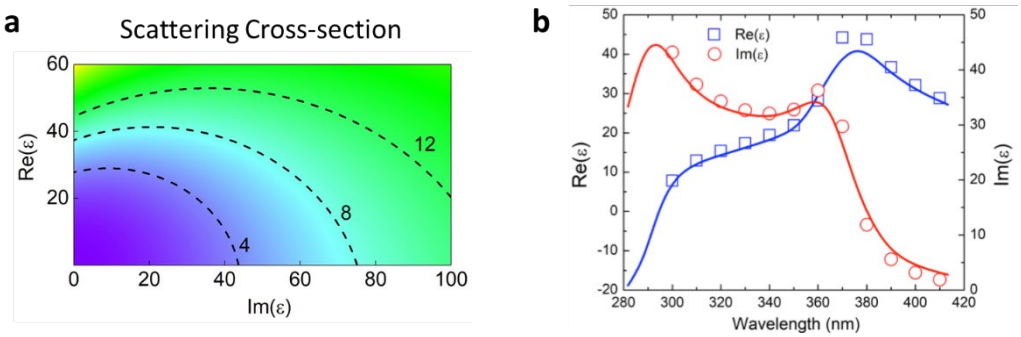

Figure 1-6 | a. The scattering cross-section plot as the function of material permittivity. **b.** The permittivity dispersion of single-crystal Si in the wavelength range of interest. The scatters are measured results, while the curves are Lorentz model fits.

We can observe a trend that a dielectric material with large ε_i can potentially have high efficiency. It is a bit counter-intuitive because the ε_i usually has a positive correlation with loss in dielectrics. We should understand this behavior using the overall impedance. The actual parameter responsible for dielectric loss is Ohmic resistance. From Equation 1-21, the rigorous expression for Ohmic resistance is:

$$R_\Omega = Re(Z_{mat}) = \frac{\varepsilon_i}{(\varepsilon_r - 1)^2 + \varepsilon_i^2} \cdot \frac{l}{\omega \cdot wh} \tag{1-28}$$

The positive correlation between ε_i and R_Ω is valid when $\varepsilon_r >> \varepsilon_i$. However, if ε_r is small, the correlation can be negative, which functions in our favor.

Fortunately, for the wavelength range of interest, single-crystal silicon (Si) happens to be the right candidate. Figure 1-6 shows the permittivity of single-crystal Si in UV, which has clear Lorentz peaks because of the interband transition. Also, the processing of Si is very convenient in modern cleanrooms, which makes the device more feasible.

However, there is still one technical challenge remains: how to obtain a single crystal Si film on a UV-transparent substrate, e.g., fused silica? We developed a transfer process like that used in the 2D-material community to address this issue[51]. The schematic figure in Figure 1-7 shows the key steps in the transfer process. We started from a transmission electron microscope grid (SiMPore Inc), which was back-etched on a silicon-on-insulator (SOI) wafer to form a suspending single-crystal Si film (~35 nm thickness). We punctured the corners of the Si window with tungsten probe tips, then applied a poly(dimethylsiloxane) (PDMS) stamp under controlled temperature. The Van der Waals force between PDMS and Si is large enough that we can mechanically break the film at the window boundaries. Then the film is transferred to a quartz substrate. The size of the transferred Si film is approximately 90 um by 90 um. Figure 1-7 also shows an optical image of the sample after this transfer process. Subsequently, standard electron beam lithography (EBL) and reactive ion etching (RIE) were used to pattern the nano-antennas with the optimized parameters.

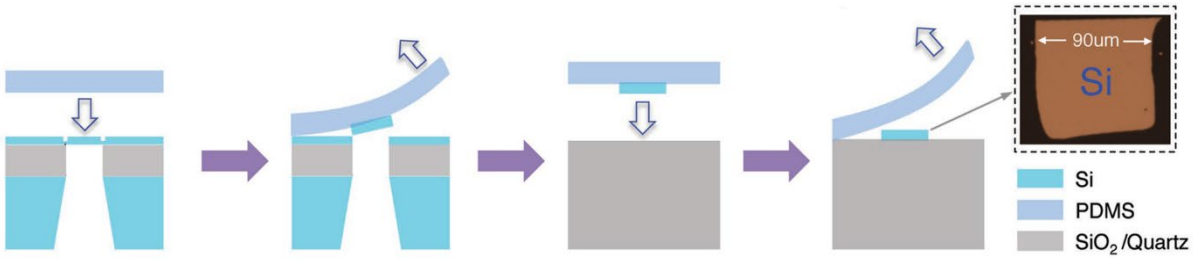

Figure 1-7 | The transfer process of single-crystal Si membrane.

§1.5 Si metasurface for the ultraviolet light

We demonstrated two metasurface devices for broad-band UV applications[51]. The first is a beam steering device; the second is a hologram phase plate. In the demonstration of UV beam steering, each unit cell of the metasurface is composed of eight optimized Si antennas with different orientation angles lying on a quartz substrate, as illustrated in Figure 1-8(a). The distance between the centers of adjacent antennas is 160 nm. The length, width, and height of the fabricated Si antenna are 140±2 nm, 30±2 nm, and 35±2 nm, respectively.

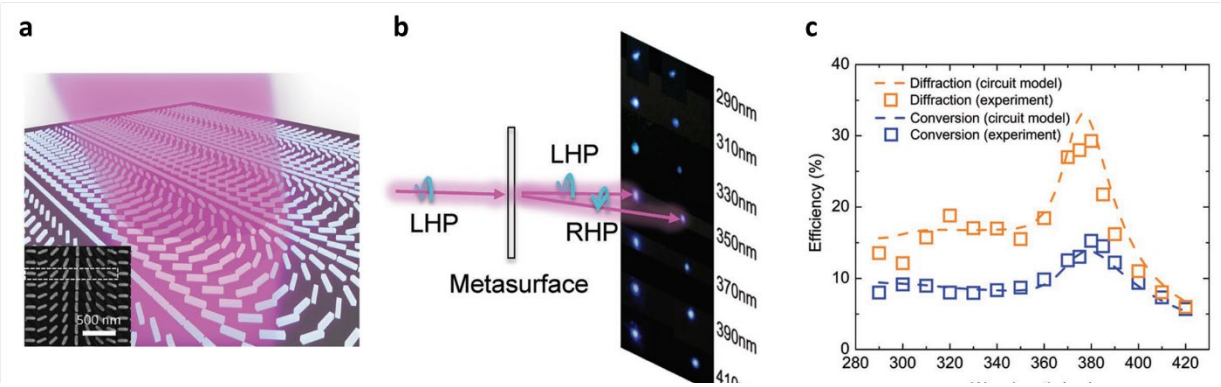

Figure 1-8 | a. The schematic of beam-steering metasurface. Insert shows the SEM image of the actual sample. The white dashed box shows one unit cell with eight antennas. **b.** Beam steering image for different wavelengths. **c.** efficiency plot with circuit model fitting.

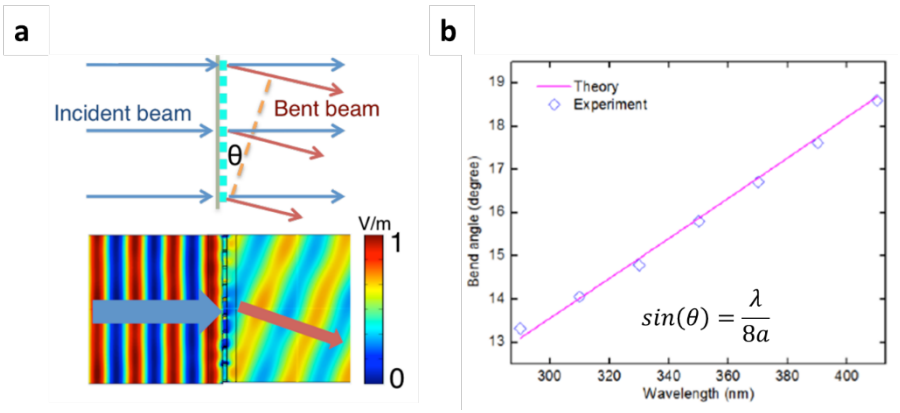

Figure 1-9 | a. Beam steering schematic and simulation. **b.** The calculation of beam steering angle. λ is the wavelength, *a* is the lattice constant, which is 160 nm.

Each unit cell creates 2π phase gradience along one direction, leading to a bending angle of approximately 13.1 degrees at 290 nm wavelength and 17.3 degrees at 380 nm. The measured bending angle for different wavelength follows the theory well (Figure 1-9).

The conversion efficiency, which is defined as the ratio of the power of the bent beam to the total incident power, coincides well with the circuit model calculations (Figure 1-8(c)) Over the wavelength range from 280 nm to 410 nm, the conversion efficiency maintains above 7% and the highest efficiency reaches around 15% at 380 nm. Following discussions in the literature, we also plot the "diffraction efficiency," which is the ratio between the power of the bent beam over that of the unmodulated beam. It is a useful parameter when the loss on the device itself is not a big concern. Overall, it is the first successful demonstration of the beam steering of UV light based on the geometric phase. It proves the feasibility of using dielectric material with large permittivity, real and imaginary, as the scattering medium.

Also, the manipulation of (relative) phase for each antenna allows the implementation of computer-generated holography (CGH). We demonstrated holographic lithography using UV metasurfaces. The device operated at the high-efficiency 380 nm wavelength, which is also close to the mercury i-line, allowing the utilization of various widely used photoresists (e.g., S1800 series, SPR220 series, and AZ 7800 series). The target image is a "CAL" pattern (UC Berkeley logo), and the pixel resolutions of both the metasurface and the target image were designed to be 300 × 300. An improved Gerchberg-Saxton algorithm was used to generate the discrete phase distribution (32 phase levels covering 2π)[52]. Figure 1-10(b) shows a high-resolution SEM image of the fabricated metasurface for holography. The optical setup is depicted in Figure 1-10(a), and a photoresist (S1805) was placed at the imaging plane of the sample. To effectively diminish laser speckle noises, we placed a 4*f* system in the optical path with a rotating diffuser in the Fourier plane. It reduces temporal coherence while maintaining spatial coherence. An imaging system was set up to image the hologram by a camera (Figure 1-10(e)). After development, the exposed pattern on S1805 showed high fidelity of the lithography process (Figure 1-10(f)), which further proves the feasibility of UV hologram using Si metasurface.

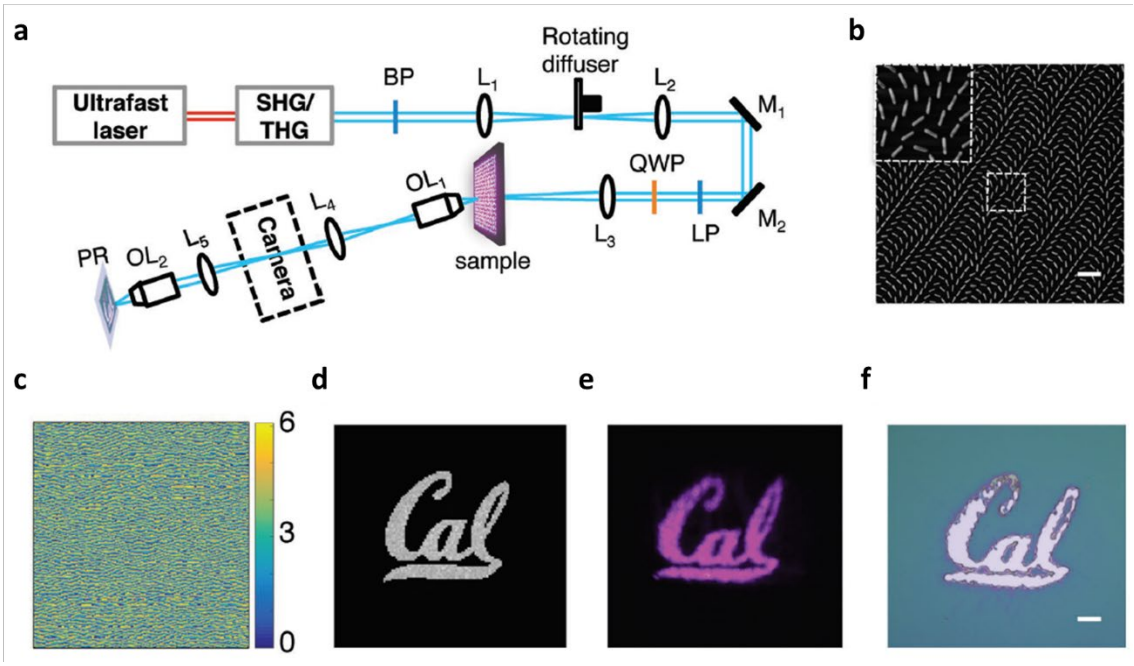

Figure 1-10 | a. Optical setup for holographic lithography. BP: band pass filter; $L_1 \sim L_5$: lenses; M_1 & M_2: mirrors; LP: linear polarizer; QWP: quarter wave plate; OL_1 & OL_2: objective lenses; PR: photoresist (S1805). **b.** SEM image of the metasurface sample. Scale bar is 500 nm. **c.** Phase distribution calculated with the improved Gerchberg-Saxton algorithm. Color bar is the phase. **d.** Calculated target pattern. **e.** Hologram imaged by camera. **f.** Pattern transferred to photoresist after development. Scale bar is 10 um.

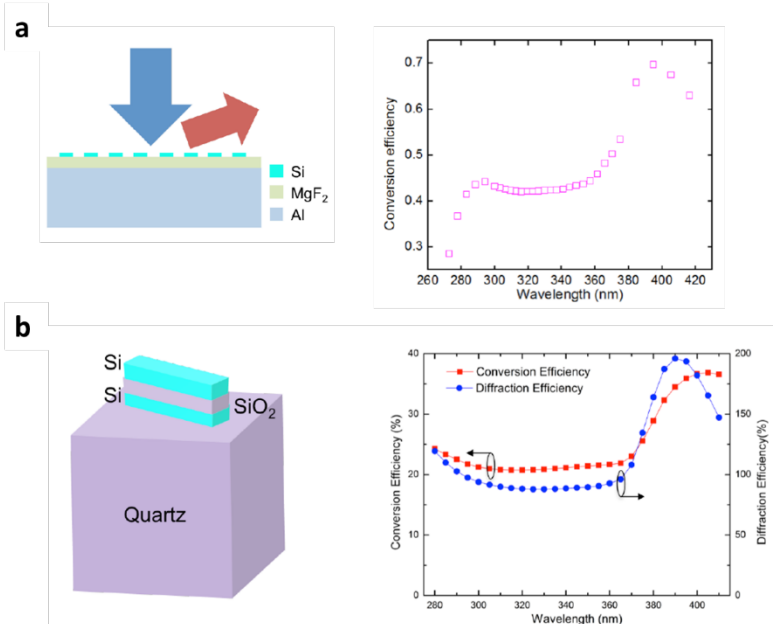

Figure 1-11 | a. Reflective metasurface design and efficiency plot. **b.** Double-bar metasurface design and efficiency plot.

It is worth mentioning that this device still has room for improvement in terms of efficiency. If process permits, one can design the Si metasurface to work at reflection mode[46]. Essentially, a resonance cavity with the carefully designed size is formed, and light can resonate and interact with the antenna multiple times. This design is achievable either via a reflective substrate (e.g., Al) where light resonates between the substrate and antenna layer or via a double-bar structure where light resonates between two antenna layers (Figure 1-11). We used numerical simulations to confirm high efficiency from both designs, which leaves an opportunity for future improvements.

§1.6 Fluorescence enhancement with Si optical antennas

In the previous sections, we elaborate on the applications of Si optical antennas used in the UV spectrum. Specifically, we take advantage of the geometric phase upon polarization conversion. In the final section of this chapter, we want to include another important application of optical antennas: near-field enhancement. This application is not related to the geometric phase, so this section serves as a supplementary part of this chapter and completes the discussion of Si optical antennas.

One of the original motivations of optical antenna researches was to enhance light-matter interaction[41,53]. It is pivotal in applications such as surface-enhanced Raman spectroscopy (SERS) and surface-enhanced fluorescence. On the one hand, Raman scattering is the inelastic scattering between lattice phonon and probing photon, which is a weak interaction; on the other hand, many biomaterials have low intrinsic quantum efficiencies. Thus, it is usually challenging to observe such weak light-matter interactions, especially if the sample has a low damage threshold[54].

When appropriately designed, optical antennas can enhance the Raman or fluorescence signals by one to three orders[55,56]. From the field distribution perspective, optical antennas can focus the probe beam and creates a high field intensity at the location of interest. Besides, optical antennas provide a high (local) optical density of states, so that we can "pump out" more emitting photons. In the following discussion, we want to focus on the fluorescence enhancement with UV excitation.

To understand the full picture of fluorescence enhancement, we need to borrow the concept of impedance matching in RF antennas[36]. In the transmission line theory, all the electrical components can be modeled using the complex impedance, so that energy reflection/ transmission, as well as phase shift, can be conveniently calculated. The typical coaxial cable has a characteristic impedance of 50 Ω or 75 Ω, while the characteristic impedance of vacuum is $Z_o = \frac{1}{\varepsilon_o c} = 377\ \Omega$. This impedance mismatch would cause a significant energy reflection at the interface and low-efficiency transmission. The device that mitigates this problem is the antenna. Therefore, the antenna should be considered as an impedance-matching device between two media. So what is precisely the antenna's impedance? It depends on how one "looks" at it. If one looks into the (on-resonance) antenna from the transmission line side, the antenna will appear to be a resistor with only the radiation resistance, which is close to the impedance of a transmission line. If one looks into the (on-resonance) antenna from the vacuum, then the antenna has a resistance of 377 Ω. It is like transformers where the voltage depends on the side one "looks into."

In the applications of fluorescence enhancement, optical antennas help match the impedance between the molecule and vacuum[42,57,58]. The emission of a typical fluorescence molecule is in the visible spectrum, but the size of the molecule is usually several nanometers. It can be modeled as a small oscillating dipole with the size of a few nanometers. The difference between the molecule

size and wavelength causes a considerable impedance mismatch. The molecule will have a minimal radiation resistance and a large imaginary impedance so that its fluorescence efficiency can be small. However, if an optical antenna is introduced, it increases the field intensity in the excitation process, while physically "enlarge" the size of the dipole in the emission process. Even with the cost of loss in the antenna itself, the overall fluorescence efficiency can be significantly enhanced.

Unfortunately, many biomolecules need to be excited via mid- to deep-UV light. For example, tryptophan, as one of the building bricks of proteins, has excitation and emission peaks around 280 nm and 370 nm, respectively (in polar solutions such as water)[59]. The tryptophan molecules in debris disintegrated from the protein can be the indicator of target protein in fluorescence measurements. However, the intrinsic quantum yield of tryptophan is not high enough compared to dedicated fluorescence dyes, so that using optical antennas to enhance its fluorescence without labeling have been highly desirable.

Following the same argument in the previous sections, noble metals fail to work efficiently in the UV, while the Si can be a good candidate material. It is worth noting that there have been some demonstrations of using Al particles for enhancement[60–63], but this solution remains less practical because of fabrication precision and material degradations. We will use the tryptophan as the target molecule and optimize the antenna geometry to enhance its overall fluorescence efficiency.

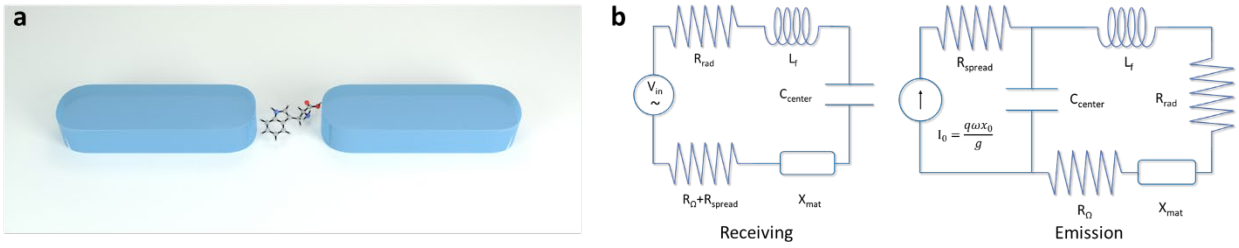

Figure 1-12 | a. Rendering of Si optical antenna with a tryptophan molecule in the gap. **b.** The circuit model for receiving and emission processes. C_{center}: the capacitance in the gap; R_{spread}: the spread resistance near the gap; x_o: size of the molecule; g: gap size; q: electron charge; ω: emission frequency. The rest of the definitions can be found in Table 1-3.

Considering the feasibility of EBL fabrications, the width, height, and gap of the Si antenna are set to be 30 nm, 40 nm, and 15 nm, respectively. We established two circuit models for the receiving and emission processes, respectively[42]. Note that we rewrite the $Re(Z_{mat})$ and $Im(Z_{mat})$ as $R\Omega$ and X_{mat}, respectively, for the convenience of discussion. In the receiving circuit, the source is an AC voltage source, assuming the incident light is linear polarized, which is the same as the metasurface:

$$V_{in} = l \times E_{in} \tag{1-29}$$

In the emission circuit, the molecule is modeled as a current source[42]. This oscillating dipole induces current in the antenna body, hence emits photons.

We shall follow the established formalism to calculate the fluorescence enhancement for receiving and emission processes[64]. For the receiving mode, enhancement is given by the electric power density:

$$f_{rec} = \frac{|E_{center}|^2}{|E_{in}|^2} \tag{1-30}$$

where E_{center} is the electric field in the gap, and E_{in} is the incidence electric field.

For the emission process, enhancement comes from the increase of quantum yield:

$$f_{emi} = \frac{{P_{rad}}/{P_o}}{{P_{rad}}/{P_o} + {P_{loss}}/{P_o} + {(1-\eta_o)}/{\eta_o}} \times \frac{1}{\eta_o} \tag{1-31}$$

where P_{rad} is the radiation power, P_{loss} is the power lost in R_{spread} and $R\Omega$, P_o is the power radiated from the emitter without an antenna, η_o is the intrinsic quantum yield of the emitter, which is 0.2 for tryptophan. The emitter in the gap is considered as a short dipole whose moment is aligned with the antenna. Additionally, the radiation efficiency of the emission process can also be defined conveniently:

$$\eta_{rad} = \frac{P_{rad}}{P_{rad} + P_{loss}} \tag{1-32}$$

Note that the fabricated Si rods have round tips, so does the model we used in the calculation. It reduces center capacitance and increases the corresponding center impedance compared to flat-ends case. This structure will increase the voltage drop across the C_{center} in the receiving process, and the shunt current that goes through it in the emission mode is minimized. Overall, round tips in the dipole antennas are beneficial for fluorescence enhancement.

Finally, the total fluorescence enhancement is:

$$f_{tot} = f_{rec} \times f_{emi} \tag{1-23}$$

We used CST Studio Suite® to simulate the fluorescence enhancement process. The E_{center}, P_{rad}, P_{loss}, and P_o can all be extracted from the simulations directly. We summarize the results given by the circuit model and CST simulations in Figure 1-13.

The simulation results and the circuit model coincide very well. It shows that we can have accurate analytical predictions about antenna performance. The circuit model goes beyond the plasmonic model that is usually found in optical antenna literature. It applies to both dielectric and metallic materials and provides more insight into the physical origins behind antenna performance.

When appropriately designed, the Si antenna can give ~90 times of total fluorescence enhancement to tryptophan. We also used Al antennas as a material benchmark and plotted in the same figure. Surprisingly, the Si outperforms Al even though it is less practical to fabrication Al nano-antennas in the first place. The peak emission efficiency of the Al antenna can be higher than the Si antenna, but the receiving enhancement dominates and crowns the Si.

We can calculate the exact impedance of Si and Al at the receiving and emission wavelengths. The real and imaginary parts of normalized Z_{mat} are listed in Table 1-4.

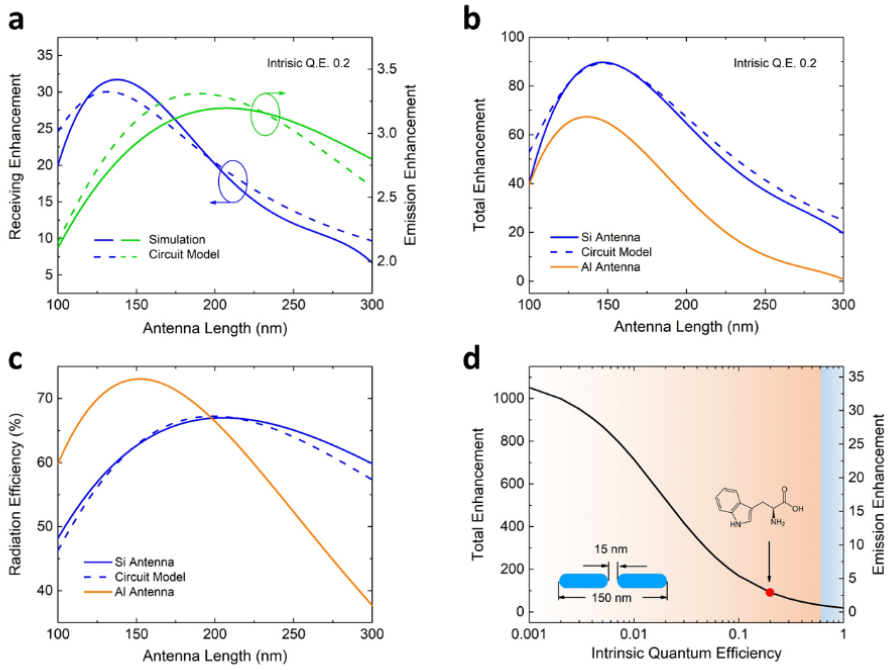

Figure 1-13 | a. Optimizing the antenna length to achieve the maximum enhancement. **b.** Total enhancement factor of Si and Al antennas. **c.** Radiation efficiency of Si and Al antennas. **d.** Total enhancement factor of the Si antenna as a function of intrinsic quantum efficiency. Tryptophan (η_o=0.2) is marked by the red dot.

$\times 10^2$ (Ω·nm)	$\lambda = 280\ nm$		$\lambda = 370\ nm$	
	$R_\Omega \cdot A/l$	$X_{mat} \cdot A/l$	$R_\Omega \cdot A/l$	$X_{mat} \cdot A/l$
Si	3.74	2.22	2.49	-3.70
Al	2.01	18.01	2.08	13.51

Table 1-4 | The normalized material impedances of Si and Al, where A is the cross-section of the nanorod, l is the length of the nanorod.

From the table above, we can conclude, although the Al is less lossy than Si, its large reactance limits the induced current in the antenna arms and eventually compromises the enhancement. Another factor worth mentioning is that the Si is capacitive in the emission mode (X_{mat}<0). According to the effective wavelength scaling, the extra capacitance will increase the physical size of the antenna for a particular resonance wavelength[40]. Longer antenna length not only relaxes the fabrication requirement but also minimizes the Wheeler's limit, which is inversely proportional to the antenna volume[65]. Wheeler's limit gives the minimum achievable quality factor for the antenna.

Lower the limit will result in larger antenna bandwidth. In Figure 1-13(c), the Si antenna has a much broader radiation efficiency peak than the Al antenna. It is of great importance when the excitation and emission wavelengths are far apart, that larger bandwidth will allow one antenna length to serve both processes better.

Since the fluorescence enhancement is a function of the intrinsic quantum yield (Equation 1-31), it is necessary to investigate their relationship to get a complete picture of antenna performance. We fix the Si antenna length at 150 nm, which is the optimal length from previous calculations, then plot the total enhancement as a function of intrinsic quantum yield. In Figure 1-13(d), the orange shade represents emission enhancement, while the blue shade shows emission quenching. The dielectric loss in antenna arms could counteract the enhancement for the high-efficiency emitters (>60% in our case). When the intrinsic quantum yield drops, the enhancement factor increases rapidly. Potentially, our Si antenna can reach up to three orders of total fluorescence enhancement with very low efficient emitters. This result has practical implications[59]: first, the quantum yield of biomolecules, tryptophan included, is a function of solution pH, so that the change of enhancement provides extra information regarding the chemical environment; second, in the applications involving very inefficient emitters, such as DNA bases[66], the Si optical antenna can be a potent tool.

In conclusion, we investigate the application of Si optical antennas in this section. Our circuit model successfully predicts the performance of the antenna and provides quantitative explanations from the perspective of material properties. Overall, Si optical antennas can have great potential in fluorescence enhancement, where the UV light is required to excite the fluorophore. The fabrication convenience and biology compatibility would make it a more attractive approach.

Chapter 2: Topological States in Optical Waveguide Arrays

§2.1 Topological states in Su-Schrieffer-Heeger model

In Chapter 1, we have discussed engineering the wavefront of light using a metasurface, which can generate arbitrary geometric phase differences from the orientations of individual nano-antennas. However, in the crystals, the quantized geometric phase is usually the subject of investigations. As described in Figure.1-1, in an adiabatic process, the geometric phase acquired by the electron is gauge dependent, unless the process is a closed-loop in the parameter space.

Specifically, in a 1D crystal, the geometric phase, i.e., the Zak phase[67,68] should be an integer multiple of π if the electron evolves through a closed-loop:

$$\varphi_{\mathrm{Zak}} = m \cdot \pi \tag{2-1}$$

where m is called the Chern number of the system, which is a topological invariant. We shall limit the following discussion to 1D and visualize the Chern number.

The model we use here is the Su-Schrieffer-Heeger (SSH) model[69,70]. It is the simplest model with non-trivial topological states based on polyacetylene. Firstly, we shall consider a 1D crystal with carbon (C) atoms using the tight-binding model.

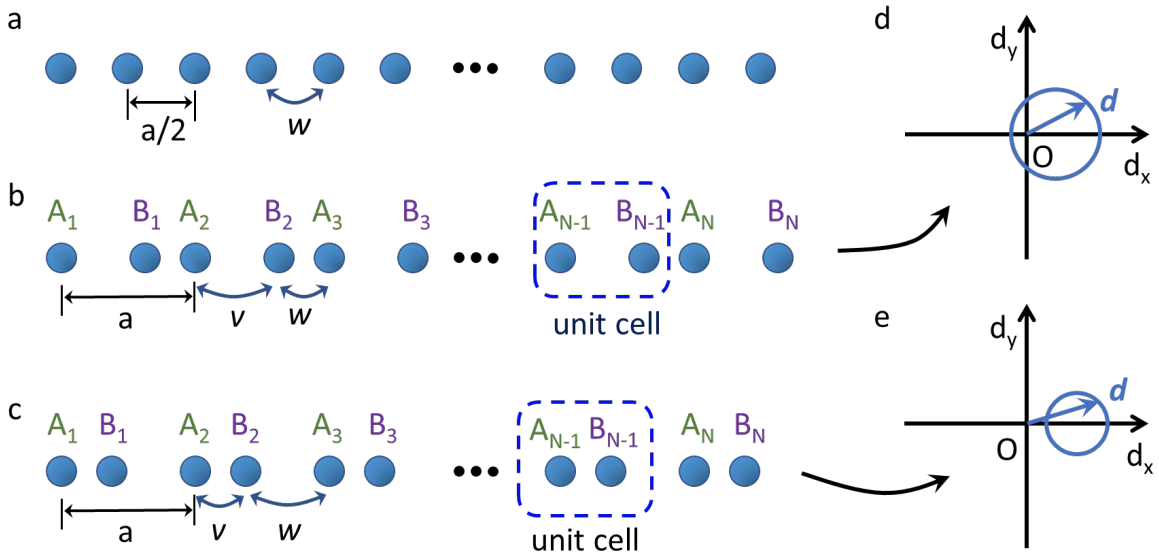

Figure 2-1 | **a.** Equally spaced 1D atomic chain. **b&c.** Dimerized 1D atomic chain, corresponding to non-trivial and trivial topological states, respectively. **d&e.** The parameter spaces show the trace of vector ***d*** as *k* sweeps the Brillouin zone. In the non-trivial state (d), the trace encloses the origin, while in the trivial state (e) it does not.

If the atoms are equally spaced with a real hopping coefficient of *w* between nearest neighbors Figure 2-1(a), the dispersion relation is:

$$E(k) = -2w \cdot \cos\left(\frac{ka}{2}\right) \tag{2-2}$$

However, because of Peierls instability, the system tends to go through a dimerization process (Peierls transition), where the atomic chain becomes Figure 2-1(b) or Figure 2-1(c). The new dispersion relation is:

$$E(k) = \pm\sqrt{v^2 + w^2 + 2vw \cdot \cos(ka)} \tag{2-3}$$

It opens an energy gap of $2|v - w|$ so that the occupied states below the Fermi level will have lower total energy.

From the bulk crystal perspective or using the periodic boundary condition, Figure 2-1(b) and Figure 2-1(c) are the same. However, the topological states of them are different. We can set the on-site energy to be 0 and write down the Hamiltonian after dimerization:

$$H(k) = \begin{pmatrix} 0 & v + we^{-ika} \\ v + we^{ika} & 0 \end{pmatrix} \tag{2-4}$$

Like any two-band model, we can also write down this Hamiltonian using Pauli matrices:

$$H(k) = \boldsymbol{d} \cdot \boldsymbol{\sigma} = d_x \cdot \sigma_x + d_y \cdot \sigma_y + d_z \cdot \sigma_z \tag{2-5}$$

$$\begin{cases} d_x = v + w \cdot \cos(ka) \\ d_y = w \cdot \sin(ka) \\ d_z = 0 \end{cases} \tag{2-6}$$

As the wavenumber *k* runs through the entire Brillouin zone (BZ), the endpoint of the vector ***d****(k)* traces out a closed circle of radius *w* on the d_x, d_y plane, centered at (*v*, 0). If $v<w$ (Figure 2-1(b)), the path encloses the origin (Figure 2-1(d)); if $v>w$ (Figure 2-1(c)), then the path excludes the origin (Figure 2-1(e)). The bulk winding number counts the number of times the loop winds around the origin of the d_x, d_y plane. The winding number is also the topological invariant (Chern number) of the system. Therefore, now we have two topological states:

$v < w$	$v > w$
Chern number: 1	Chern number: 0
Topological non-trivial	Topological trivial

Table 2-1 | The comparison between two topological states.

Note that in this derivation, we have assumed the crystal to have N unit cells with *v* as the hopping coefficient at the edges. When the atomic chain is terminated by a weak bond, we can expect it to be the non-trivial topological phase. From the argument of bulk-boundary correspondence, the non-zero bulk winding number indicates the existence of the edge states.

Indeed, the edge atom can host a zero-state in the non-trivial case, that the exact energy is at the order of $e^{-N \cdot log(\frac{w}{v})}$, and a characteristic decay length of $1/log(\frac{w}{v})$. Edge modes decay exponentially from the boundaries to the bulk: the left edge mode only exists on A sites, while the right edge mode only exists on B sites. The exact derivation of the edge mode can be found in the first chapter of the textbook[70] or some helpful introductory papers[71]. In this chapter, we intend to focus on the realization of the topological states in the photonic systems.

§2.2 Topological states in the optical waveguide arrays

Although polyacetylene is an excellent example in the SSH model, the experimental realization is very challenging in crystals. It comes down to the fact that we cannot manipulate the atomic bond conveniently. Fortunately, the photonic systems provide a viable approach to mimic the condensed matter.

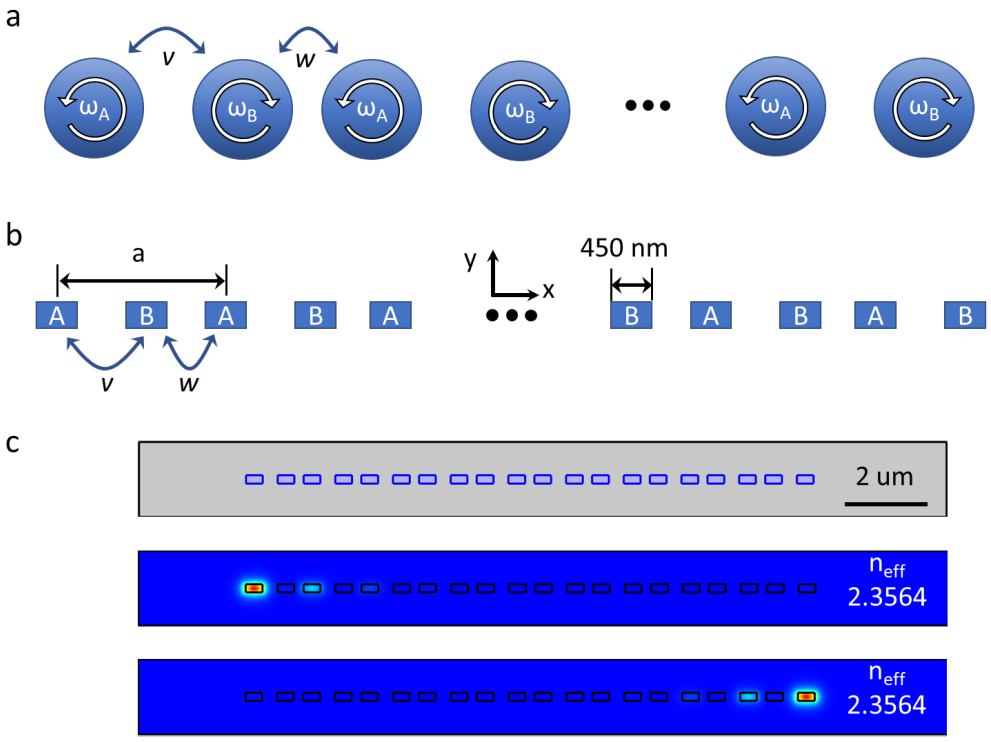

Figure 2-2 | a. Using a chain of coupled optical resonators to mimic a 1D crystal. **b.** Using an optical waveguide array to mimic a 1D crystal. **c.** An example of an optical waveguide array with non-trivial edge states. The top panel is the device set-up where the grey color is silica and purple color is silicon. The two edge states both decay exponentially where the red color is strong electric field while the blue color means no field. They share the same effective index which is also close to the "zero-energy" of the system.

In our investigation of the SSH model, we can generalize it to a chain of coupled resonators. The energy is equivalent to the resonance frequency, while the coupling coefficient or bond strength can be converted to decay (coupling) rate, lifetime, or quality factor.

$$\begin{cases} E(energy) \Leftrightarrow \omega(frequency) \\ v(or\ w)(coupling) \Leftrightarrow \frac{1}{\tau}(lifetime) \Leftrightarrow \frac{1}{Q}(quality\ factor) \end{cases}$$

Naturally, we can set up a series of optical resonators to simulate a chain of atoms (Figure 2-2(a)). Each resonator has its resonance frequency ω_i, which is the "on-site" energy. The nearest neighbors can crosstalk so that photons can hop between them, which is similar to chemical bonds. We can design and fabricate the resonators and their arrangement so that we would have full

control of the on-site energies and coupling coefficient, therefore the Hamiltonian. Coupled resonator arrays have been proven to be an effective way to simulate crystals[72–75]. Here we would like to introduce a special kind of "resonator," which is the optical waveguide (Figure 2-2(b)).

In contrary to a ring resonator or a cavity working at whispering-gallery mode, the photons in waveguides do not resonant at a particular frequency; instead, each frequency is associated with an effective mode index. The mode in the optical waveguide can be described by the paraxial approximation (linear medium)[76]:

$$-i\frac{\partial \boldsymbol{E}(x,y,z)}{\partial z} = \{\frac{1}{2k_o}\nabla_T^2 + k_o(n(x,y) - n_o)\}\boldsymbol{E}(x,y,z) \tag{2-7}$$

where z is the propagation direction, k_o is the wavevector in the free space, n_o is the refractive index of the environment, $n(x,y)$ is the refractive index profile of the cross-section, ∇^2 is the Laplace operator: $\nabla_T^2 = \frac{\partial^2}{\partial^2 x} + \frac{\partial^2}{\partial^2 y}$. We can rewrite the electric vector using the propagation constant β_z:

$$\boldsymbol{E}(x,y,z) = \boldsymbol{E}(x,y)e^{i\beta_z z} = \boldsymbol{E}(\boldsymbol{r})e^{i\beta_z z} \tag{2-8}$$

Then the wave equation above immediately becomes the familiar Schrödinger's equation with the following substitutions:

$$\begin{cases} z \to t \\ \beta_z \to \omega \\ n(\boldsymbol{r}) \to V(\boldsymbol{r}) \end{cases} \tag{2-9}$$

In other words, the propagation distance is the effective time, the propagation constant is effective frequency or energy, and the refractive index is the effective potential. Therefore, we can use a waveguide to simulate an atom. To simplify the discussion, we can even use the effective index of the mode as the effective energy for monochrome light, since $\beta_z = n_{eff}k_o$.

It is convenient to set up an optical waveguide array as the 1D atomic chain used in the SSH model. Without loss of generality, the dynamic equation is:

$$\begin{cases} i\frac{\partial \boldsymbol{E}_{i,A}}{\partial z} + \beta_A \boldsymbol{E}_{i,A} + we^{-ika}\boldsymbol{E}_{i-1,B} + v\boldsymbol{E}_{i,B} = 0 \\ i\frac{\partial \boldsymbol{E}_{i,B}}{\partial z} + \beta_B \boldsymbol{E}_{i,B} + we^{ika}\boldsymbol{E}_{i+1,A} + v\boldsymbol{E}_{i,A} = 0 \end{cases} \tag{2-10}$$

Where $\boldsymbol{E}_i$ is the electric field in the i^{th} unit cell, $\beta_{A(B)}$ is the intrinsic propagation constant, i.e., the on-site energy. In the standard SSH model, A and B sites are made of identical waveguides, so that $\beta_A = \beta_B$. This equation will yield the bulk dispersion exactly as Equation 2-3. from the SSH model, we can expect that edge states would emerge if the system is topologically non-trivial.

In Figure 2-2(c), we show an example of the topological edge states. The waveguides are made from the 220-nm silicon-on-insulator (SOI) wafer with a 450-nm width. The cladding is silica, and the working wavelength is 1550 nm. We can arrange the waveguide array so that the larger gap corresponds to the smaller coupling coefficient, and vice versa. We can immediately observe the edges states at the left and right boundaries, respectively. More importantly, the n_{eff} of the edge states are very close to the "zero energy," which is the intrinsic n_{eff} of an isolated 450-nm waveguide (2.3563). This result further proves that the edge states are topological.

Note that the waveguide we use here is the "single-mode" waveguide for the wavelength of interest. Only fundamental modes, i.e., TE_0 and TM_0, can propagate. Without loss of generality, all the modes discussed here are TE modes, where E_x dominates over E_y. To be more specific, the TE and TE modes in a strip or ridge waveguide should be called "TE-like" and "TM-like" because the polarization cannot be purely transverse. Readers can refer to a textbook[77] on this topic.

§2.3 Mimicking topological states in the graphene nanoribbons

Indeed, it is interesting to see the realization of the SSH model in an optical waveguide array, but the photonic system can be a much more powerful tool to study topological phenomena. Fundamentally, it provides a way to engineer the coupling coefficient at will, which brings about a lot of versatility. In this section, we want to demonstrate this point by introducing a mapping between the parameter space of the waveguide array to the *k*-space of a graphene nanoribbon. We will show that it is possible to mimic any momentum and the topological states in the nanoribbon using a waveguide array, while the Brillouin zone (BZ) of the nanoribbon can be reproduced in by a set of waveguide arrays.

Since the discovery and production of single-layer graphene and other 2D materials, new opportunities have emerged from the material family of nanoribbons[78,79]. For instance, the finite termination of graphene along one direction results in quasi-1D nanoribbon structures. They are categorized based on the types of edges, namely zigzag or armchair, with a supercell across the width. In experiments, bottom-up synthesis can produce such nanoribbons with smooth edges[80]. Electronically, a graphene nanoribbon usually exhibits a subset of the 2D graphene band structure[81].

However, the zigzag graphene nanoribbon (ZGNR) possesses a unique feature that does not exist in 2D graphene: the edge states (flat bands) near the BZ boundary. This phenomenon was predicted theoretically[82,83] and later observed by the scanning tunneling microscope (STM)[84]. The topological origin of such edge states was also discussed, that the Zak phase and Chern number become non-trivial when the momentum of the electron is large enough[85,86]. Despite the direct observation of edge states with sophisticated instruments, the study of such edge states remains challenging. Firstly, it requires skillful synthesis and expensive STM; secondly, one can only retrieve the energy information of the edge edges, but not the dispersion relation without angular-resolved spectroscopy[87]; thirdly, some theoretically interesting structure, e.g., beard edge nanoribbons[86], is too unstable to be produced.

To address these challenges, researchers have found an approach to mimic solid-state matter with photonic systems. Explicitly, under paraxial approximation, the optical waveguide array can simulate a chain of atoms[88]. To study the intriguing edge states in graphene nanoribbons, Blanco-Redondo et al. employed femtosecond laser writing techniques to construct a 2D waveguide array, with each waveguide representing one carbon atom in the nanoribbon model[89]. However, the fabrication of such optical waveguide arrays can be a challenge by itself.

In this section, we propose a formalism that allows us to map the *k*-space of a nanoribbon system to the parameter space of a 1D diatomic chain. It also gives a perspective of connecting physical systems of different dimensions[90]. Such 1D systems can be realized by 1D optical waveguide arrays, which are compatible with fabrications on typical silicon-on-insulator (SOI) wafers. This approach provides an excellent tight-binding-model realization and allows us to reproduce the

band structures of nanoribbon models. We believe this formalism will also deepen our understandings of the topological nature in the nanoribbons and give predictions of novel phenomena.

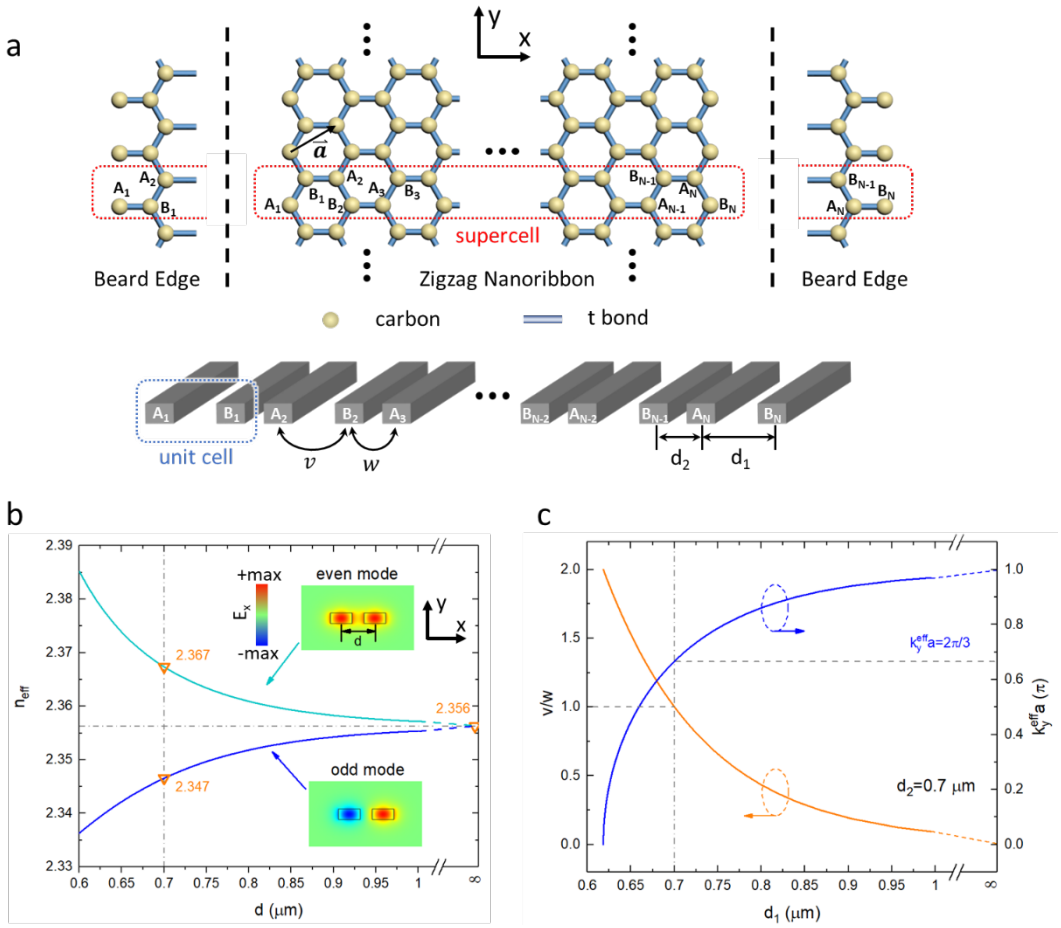

Figure 2-3 | a. Schematic drawing of the graphene nanoribbon and 1D optical waveguide array (diatomic chain model). $\vec{a}$ is the lattice constant in the nanoribbon. In the 1D model, v and w are intracell and intercell hopping constants, respectively; d_1 and d_2 are intracell and intercell distances, respectively. **b.** Effective mode indices (eigenvalues) of directional couplers. The inset shows cross-section of the waveguides. **c.** The mapping relation between order parameter v/w and normalized wavevector as a function of d_1.

We shall first write down the 2N × 2N Hamiltonian of a ZGNR with 2N atoms in the supercell using the tight-binding model (Figure 2-3(a)):

$$H_{ZGNR}(k_y) = \begin{pmatrix} 0 & tf & & & & & & \\ tf & 0 & t & & & & & \\ & t & 0 & tf & & & & \\ & & tf & 0 & & & & \\ & & & & \ddots & & & \\ & & & & & 0 & t & \\ & & & & & t & 0 & tf \\ & & & & & & tf & 0 \end{pmatrix} \tag{2-11}$$

where t is the hopping coefficient of nearest neighbors, f is the Bloch term: $f = 2\cos\left(\frac{k_y a}{2}\right)$, k_y is the wavevector along the translation direction of the nanoribbon, a is the magnitude of the lattice vector. Note that we assume a proper gauge to make the hopping coefficient a real number, and we set on-site energy to be zero.

Next, we shall write down the 2N x 2N Hamiltonian of a 1D diatomic chain with N unit cells (2N atoms) in real space[70] (Figure 2-3(a)):

$$H_{SSH} = \begin{pmatrix} 0 & v & & & & & & \\ v & 0 & w & & & & & \\ & w & 0 & v & & & & \\ & & v & 0 & & & & \\ & & & & \ddots & & & \\ & & & & & 0 & w & \\ & & & & & w & 0 & v \\ & & & & & & v & 0 \end{pmatrix} \tag{2-12}$$

where v and w are intracell and intercell hopping coefficients, respectively. H_{ZGNR} and H_{SSH} are equivalent if:

$$\begin{cases} w = t \\ v = 2t \cdot cos\left(\frac{k_y a}{2}\right) \end{cases} \tag{2-13}$$

Here we can draw an important conclusion: if we take an SSH model, fix the intercell hopping coefficient w, and change intracell hopping coefficient v in the range of $0 \le v \le 2w$, we can effectively sweep the entire BZ of the ZGNR and reconstruct its band structure. It is essentially a mapping relation between a 1D diatomic chain and ZGNR model that each k_y corresponds to an instance of the diatomic chain:

$$a \cdot k_y = 2 \cdot cos^{-1}\left(\frac{v}{2w}\right) \tag{2-14}$$

It is effectively mapping the k-space in the nanoribbon to the parameter space (v/w) in the 1D diatomic chain.

We can understand this mapping relation using the Bloch theorem. Along y-direction in the nanoribbon, the translation symmetry preserves. Because of the periodic scattering potential, the hopping coefficient of every other bond is modulated by the Bloch wave, with the first and last bonds subject to modulation. Therefore, changing k_y is equivalent to physically changing the hopping coefficient of those modulated bonds in a 1D diatomic chain along the x-direction.

With this mapping relation, the topological origin of the edge states in ZGNR can be described by the standard SSH model. In the collection of all the SSH instances that constitute the BZ of ZGNR, the ones where $v < w$ are topologically non-trivial with edge states. From Equation 2-14, this condition converts to $|a \cdot k_y| > \frac{2\pi}{3}$ in ZGNR, which coincides with the theoretical calculations[86,91].

This mapping relation provides a powerful tool to reconstruct and study the topological states in the nanoribbons if one can artificially control the hopping coefficient between atoms. Indeed, it is not likely to arbitrary alter the chemical bonds in the solids, but optical waveguide arrays are a feasible platform.

The mode propagation in a single-mode optical waveguide can be described by the paraxial approximation where the effective mode index (n_{eff}) is the "energy level" of the lattice[88]. We numerically retrieve the coupling coefficient $p(d)$ as a function of the distance (d) between two waveguides (center to center) from the eigenvalues of the directional coupler Hamiltonian:

$$H_{DC}(d) = \begin{pmatrix} n_{eff} & p(d) \\ p^*(d) & n_{eff} \end{pmatrix} \tag{2-15}$$

where n_{eff} is the effective index of the fundamental TE mode in one waveguide.

Without loss of generality, we use silicon waveguides as "atoms," which are 450 nm in width, 220 nm in height, and surrounded by silica. The working wavelength is 1550 nm. This setup represents a typical single-mode silicon waveguide fabricated on an SOI wafer, which is also compatible with complementary metal oxide semiconductor (CMOS) technology[92]. Here we only consider the fundamental TE mode ($n_{eff} = 2.3563$) and nearest-neighbor coupling. The simulations are carried out in Comsol Multiphysics®.

The eigenvalues of directional couplers are plotted in (Figure 2-3(b)): the odd and even modes split apart when d decreases, and they converge to the single waveguide n_{eff} when $d \to \infty$. The order parameter of the topological phase is v/w rather than the absolute value of v and w. So we choose $w = p_{d=0.7\mu m}$, then we normalize $v = p(d)$ to w (Figure 2-3(c)). In other words, we fix the intercell distance $d_2 = 0.7\ \mu m$ while changing intracell distance d_1. Finally, Equation 2-14 can give us the mapping relation between $k_y^{eff} \cdot a$ and d_1, as plotted in Figure 2-3(c). Note that we need to extrapolate at the region close to the BZ boundary because $d_1 \to \infty$. The topological phase transition occurs at $d_1 = d_2 = 0.7\ \mu m$, where $k_y^{eff} \cdot a = {}^{2\pi}/_{3}$. The right-hand side of the boundary represents non-trivial topological states, while the left-hand side represents trivial states.

Now we can construct the band structure of ZGNR. We sweep $k_y^{eff} \cdot a$ from 0 to π (including extrapolation) by changing d_1 to the corresponding values. The band structure with 15 unit cells is plotted in Figure 2-4(a), with edge states marked in red when $|k_y^{eff} \cdot a| > {}^{2\pi}/_{3}$. This band structure coincides very well with the tight-binding calculation of ZGNR[86,91,93]. There are two features of the band structure worth further discussion: first, the degeneracy at BZ boundary; second, the slight dispersion of "zero-modes" near the BZ boundary.

At BZ boundary, where $k_y^{eff} \cdot a = \pi$, $d_2 = 0.7\ \mu m$, $d_1 \to \infty$, the optical waveguide array constitutes of N-1 pairs of waveguides (waveguides are separated by 0.7 um in each pair), and two stand-alone waveguides on both ends. Therefore, the system must degenerate into three eigenvalues that are marked with orange triangles in Figure 2-3(b). This configuration represents

the fully dimerized SSH model (ν=0), whose edge states are "exact" zero-modes rather than "near" zero-modes. It is true even if we include a non-negligible overlap integral in Schrodinger's equation. This unique feature of the ZGNR model is usually overlooked, but it is easy to understand using our mapping method. The degeneracy at the BZ boundary indicates that the optical waveguide array system is a cleaner tight-binding model realization compared to graphene nanoribbon. Both density function theory (DFT) calculation and STM measurement confirmed that the degeneracy of at the BZ boundary in ZGNR is lifted due to Coulomb repulsion, spin interaction, and chemical potentials at the edge[81,91,94].

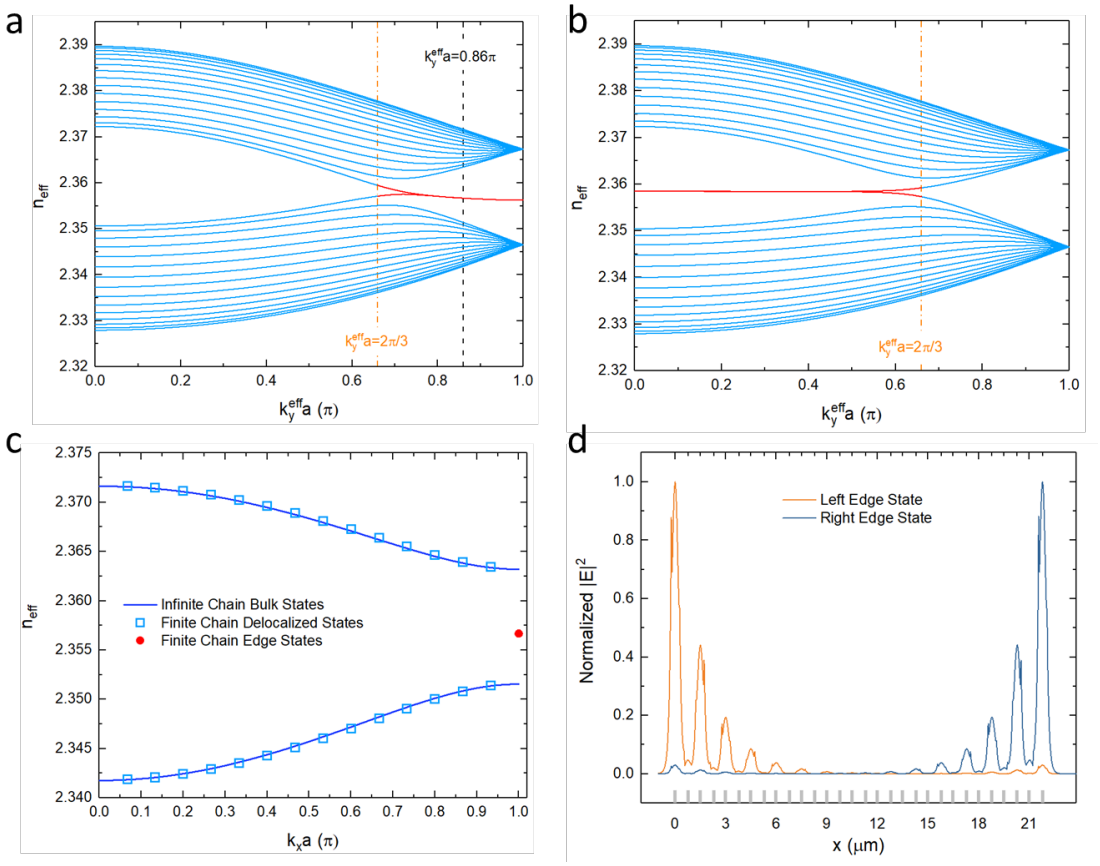

Figure 2-4 | a&b. The bandstructure of the ZGNR and BGNR models, respectively, constructed by optical waveguide arrays. **c.** Dispersion relation of a non-trivial waveguide array that corresponds to the black dashed line in panel a. **d.** Wavefunction of edge states marked by the red dot in panel c. The grey bars on horizontal axis are the positions of waveguides.

The slight dispersion of the zero-mode here is very similar to that in calculations of graphene[91]. Discussions attributed this phenomenon to second-nearest-neighbor hopping[95]. Nevertheless, we can consider this dispersion as a more general consequence of breaking particle-hole symmetry; that is, the upper and lower bands are not symmetrical with respect to "zero energy" (Figure 2-4(a)). Adding second-nearest-neighbor hopping is one way to break particle-hole symmetry, but in a well-confined waveguide, this effect is weak. We believe the primary reason to break particle-

hole symmetry here is the substantial non-zero overlap integral between adjacent waveguides. This feature will be further discussed in the next section.

To further prove that each k_y^{eff} corresponds to an instance of the SSH model, we take a vertical slice in Figure 2-4(a) at $k_y^{eff} \cdot a = 0.82\pi$, then replot all the eigenvalues in Figure 2-4(c). Now the *x*-axis is k_x, which is the wavevector of the 1D diatomic chain. The delocalized modes fall entirely on the bulk bands, which are calculated using one unit cell and periodic Bloch boundary conditions. The (degenerate) edge states are marked with the red dot that appears in the bandgap. We can also plot the wavefunction of the two edge states in Figure 2-4(d). The patterns are typical edge states, with exponential decay from the ends and finite electric field on every other site. Because the system is not large enough, we can also observe some coupling between the left and right edge states. It is a distinctive feature of the SSH model[70].

An essential advantage of the photonic system is allowing us to mimic electronic states that are challenging to be observed in solids. In the graphene nanoribbon system, researchers have considered a model with beard edges, i.e., dangling bonds that terminate the nanoribbon[86] (Figure 2-3(a)). This structure is unstable, so there is no report of its fabrication or measurement. The Hamiltonian of the beard graphene nanoribbon (BGNR) is similar to H_{ZGNR} that we only need to swap *t* and *tf* in Equation 11 because the bonds at both ends are not modulated by Bloch wave. To map the BGNR model to a 1D optical waveguide array, we need to set $v = p_{d=0.7\mu m}$ while changing $w = p(d)$. In other words, now we fix $d_1 = 0.7\ \mu m$ and sweep d_2. Naturally, the condition for non-trivial topological states ($\left|k_y^{eff} \cdot a\right| < {}^{2\pi}/_{3}$) is the complement of that in ZGNR to comply with the same requirement of non-trivial states ($v < w$ or equivalently $d_1 > d_2$). Again, the bandstructrue we calculated (Figure 2-4(b)) coincides very well with the tight-binding calculations in the literature[86,93]. The mapping method provides a clear and intuitive physical picture to understand nanoribbons with different edge types.

While the ZGNR model promises for exact zero modes at the BZ boundary, it is not the case in the BGNR model. When a finite system is not fully dimerized, the zero-mode in the SSH model is not precisely zero energy[70], but instead has an eigenvalue of $E \sim \pm\, e^{-N \cdot \log\left|\frac{w}{v}\right|}$. Besides, a finite overlap integral of adjacent orbitals can shift the energy band, which results in a shift of zero-modes as well. In the BGNR model, there is no fully dimerized point, so the exact zero-mode does not exist. In Figure 2-4(b), the flat band does not cross the n_{eff} of the single waveguide, which is 2.3563.

Besides studying the hypothetical edge type, we can also construct nanoribbon models with artificial bonds using 1D optical arrays. Specifically, the ZGNR only has one carbon-carbon bond, and we use *t* as the only hopping coefficient (Figure 2-3(a) & Equation 2-11). Then the Bloch theorem determines the restriction that $0 \le v \le 2w$ (Equation 2-13). However, the 1D optical array can simulate a ZGNR-like model with two chemical bonds (Figure 2-5(a)). If we generalize Equation 2-13 to:

$$\begin{cases} w = t_2 \\ v = 2t_1 \cdot cos\left(\dfrac{k_y^{eff} a}{2}\right) \end{cases} \tag{2-16}$$

The topological phase transition condition (v=w) gives:

$$\cos\left(\frac{k_y^{eff}a}{2}\right)=\frac{t_2}{2t_1} \tag{2-17}$$

We can plot the phase transition boundary as a function of dimerization (t_1/t_2) in the supercell of the nanoribbon (Figure 2-5(b)). We also need to map the nanoribbon BZ to different ranges in the parameter space of the 1D system. The BZ center starts at different d_1 values, while the BZ boundary always corresponds to $d_1 \rightarrow \infty$ (Figure 2-5(c)). Therefore, a weaker t_1 bond (compared to t_2) will result in a larger range of k_y where the topological states exist, and vice versa. Specifically, if $t_1 < 0.5t_2$, topological states emerge across the entire BZ. These results are consistent with the theory work based on the black phosphorus nanoribbon model[96].

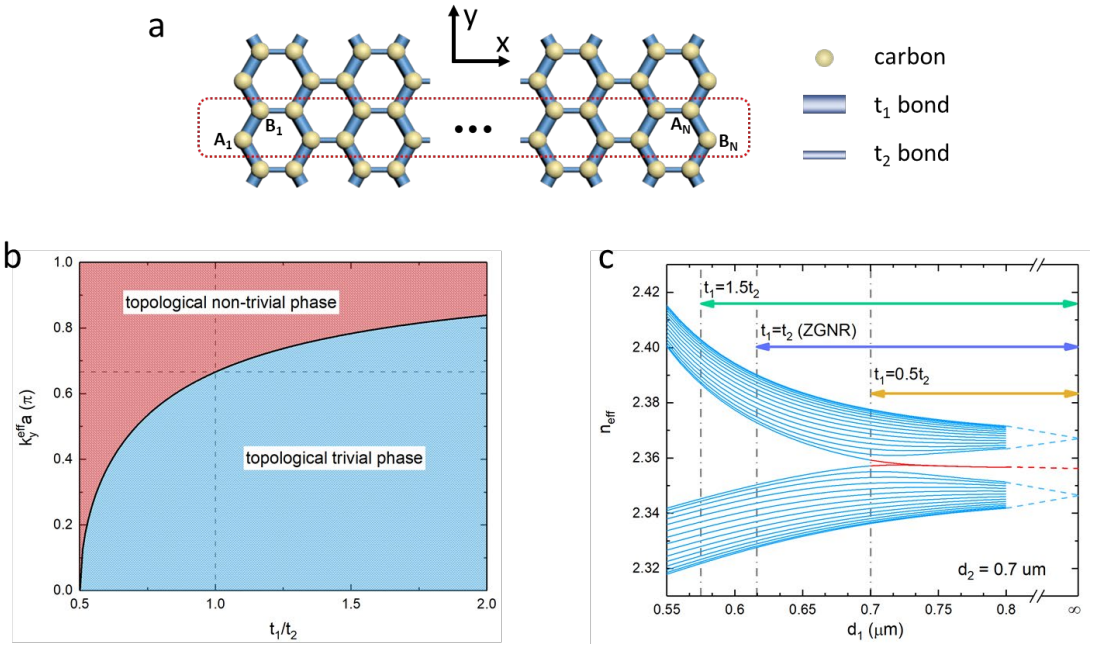

Figure 2-5 | **a.** ZGNR model with a dimerized supercell. **b.** Topological phase diagram as a function of supercell dimerization. The dashed lines denote the normal ZGNR model. **c.** The double-head arrows show the range of d_1 needed to map the entire BZ under different supercell dimerization. Blue curves are the eigenstates of waveguide arrays, red curves are the edge states. Dashed lines are extrapolations.

§2.4 Staggered waveguide arrays and control of topological phase transitions

The discussions above are all based on the ZGNR model, where on-site energy is set to zero since carbon is the only element. If we introduce a staggered waveguide array that consists of two kinds of waveguides, we can simulate a material with two elements, such as hexagonal boron nitride (*h*BN). *h*BN nanoribbon model has drawn a lot of attention recently[79,97], but the difficulty of obtaining single-layer *h*BN limits the experimental explorations. 2D *h*BN is supposed to be an

insulator, but zigzag *h*BN nanoribbon (ZBNNR) can theoretically show conducting edges similar to the ZGNR[98,99]. To simulate the band structure of ZBNNR, we use 420-nm-width waveguides to replace all the B sites in Figure 2-3(a); then, we repeat the process as in the ZGNR case. The resulting band structure is shown in Figure 2-6(b), with edge states marked in red.

A significant feature of this band structure is that two edge states have different phase transition points: the upper edge states emerger when $\left|k_y^{eff} \cdot a\right| > 0.4\pi$, but the lower edge states only appear when $\left|k_y^{eff} \cdot a\right| > 0.88\pi$. In other words, if $0.4\pi < k_y^{eff} \cdot a < 0.88\pi$, this waveguide array will only have a one-side edge state. The calculation result of the ZBNNR already shows the "asymmetrical" edge states[100]. It is a consequence of breaking both chiral symmetry and particle-hole symmetry. We can use this example to discuss the non-zero overlap integral in optical waveguide arrays.

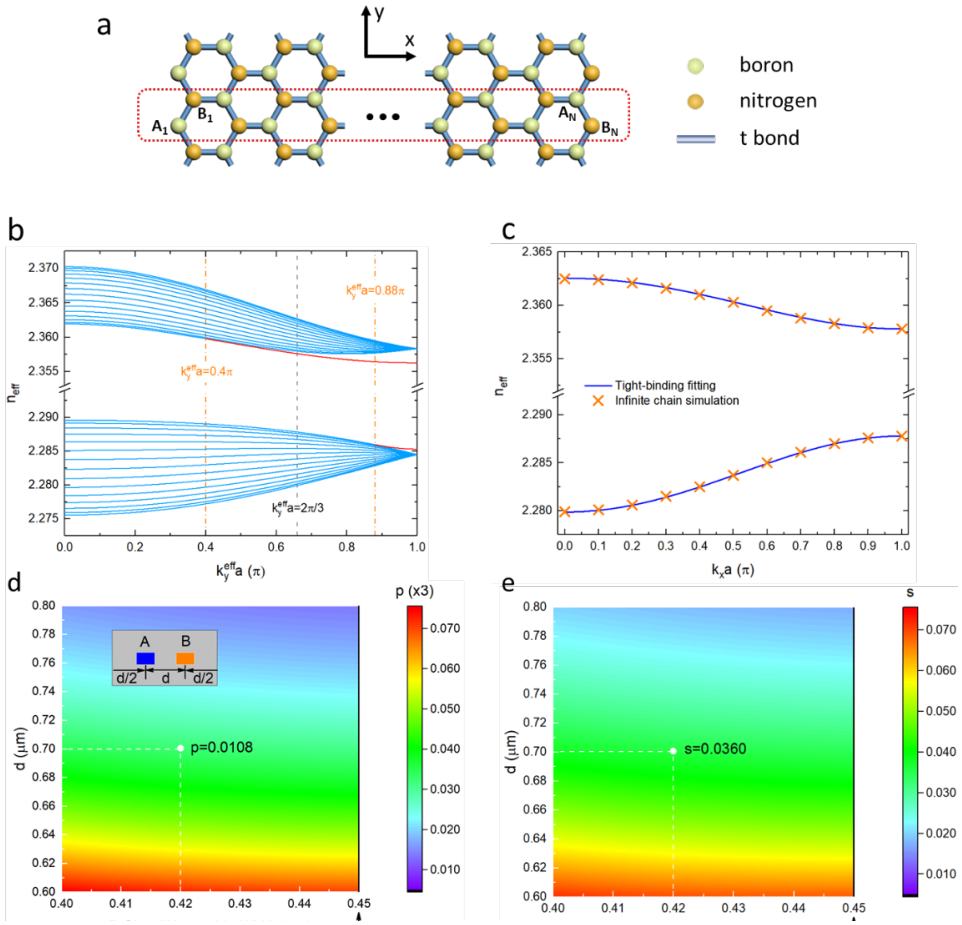

Figure 2-6 | a. ZBNNR model with staggered on-site energy. Red rectangle indicates the supercell. Hydrogen terminations are omitted here. **b.** The bandstructure of ZBNNR model constructed by the optical waveguide arrays. **c.** Tight-binding fitting of the waveguide array that corresponds to the black dashed line in the bandstructure. **d & e.** Coupling coefficient *p* (×3) and overlap integral *s*, respectively, as functions of waveguide distance d and B-sites waveguide width. The A-sites waveguide width is 450 nm. Inset shows schematic of simulation unit cell, where left and right boundaries are subject to Bloch boundary conditions. White dots show the parameters used in panel c.

The standard SSH model possessed the chiral symmetry:

$$\sigma_z H_{SSH} \sigma_z = -H_{SSH} \tag{2-18}$$

where σ_z is the third Pauli matrix, which is the chiral operator here. If we only break chiral symmetry by adding the staggered on-site energy, it becomes the Rice-Mele model[70]. Note that the Rice-Mele model could bring about exotic phenomena such as Thouless pumping in the waveguide array[101,102] and non-integer Chern number of the topological phase[103], here we only want to focus on the topological phase transition. The resulting band structure will still be symmetrical[100], and topological phase transition still happens at $\left|k_y^{eff} \cdot a\right| = \frac{2\pi}{3}$. But the edge states will have non-zero eigenvalues.

However, in the optical waveguide array system, the particle-hole symmetry is also broken by a finite overlap integral *s*. Similar to the coupling coefficient *p* between nearest neighbors, *s* is a function of distance *d* and on-site energy of waveguides. We apply periodic boundary condition to a unit-cell (inset of Figure 2-6(d)), which is described by the complete Schrodinger's equation:

$$H_{RM} \begin{bmatrix} \varphi_A^{k_x} \\ \varphi_B^{k_x} \end{bmatrix} = E_k S_k \begin{bmatrix} \varphi_A^{k_x} \\ \varphi_B^{k_x} \end{bmatrix} \tag{2-19}$$

$$\begin{bmatrix} \Delta & 2p \cdot \cos(k_x d) \\ 2p \cdot \cos(k_x d) & -\Delta \end{bmatrix} \begin{bmatrix} \varphi_A^{k_x} \\ \varphi_B^{k_x} \end{bmatrix} = E_k \begin{bmatrix} 1 & 2s \cdot \cos(k_x d) \\ 2s \cdot \cos(k_x d) & 1 \end{bmatrix} \begin{bmatrix} \varphi_A^{k_x} \\ \varphi_B^{k_x} \end{bmatrix} \tag{2-20}$$

where Δ is the on-site energy offset. Figure 2-6(d&e) show the fitting result of *p* and *s* under different conditions. The fitted parameters also give rise to an analytical band structure that coincides well with simulations (Figure 2-6(c)). Contradicting the norm in the electronic system, where *s* is usually significantly smaller than *p*, the overlap integral is not negligible here. Therefore, the band structure in Figure 2-6(b) is distorted (breaking the particle-hole symmetry). It is a unique feature in optical waveguide arrays because the modes in adjacent waveguides need to overlap so that the integral *s* can be substantial. It also indicates that overlap integral *s* could be a new degree of freedom in controlling topological phases in waveguide arrays.

Moreover, we shall consider breaking particle-hole symmetry while preserving chiral symmetry. This case is the photonic realization of the ZGNR model (Figure 2-4(a)), where the fitted *p* and *s* are at the right-hand-side boundaries in Figure 2-6(d&e) (noted by black arrows). Similar to the staggered waveguide array, there is non-negligible overlap integral s, which distorts the bands and causes the dispersion of the flat band in the ZGNR model. However, the chiral symmetry forces the system to have a pair of edge states if there are any[86]. Intuitively, the two edges are not distinguishable under chiral symmetry, so that edge states must emerge/disappear simultaneously.

We now can consider in a more general way that how the energy offset, the coupling coefficient, and the overlap integral together affect the topological phase transition point. The question we are trying to answer here is: for a bulk Hamiltonian and overlap matrix as the following, what is the topological phase transition condition for upper and lower band, respectively.

$$\begin{cases} H = \begin{bmatrix} \Delta & v \cdot e^{ik_x d_1} + w \cdot e^{-ik_x d_2} \\ v \cdot e^{-ik_x d_1} + w \cdot e^{ik_x d_2} & -\Delta \end{bmatrix} \\ S = \begin{bmatrix} 1 & s_1 \cdot e^{ik_x d_1} + s_2 \cdot e^{-ik_x d_2} \\ s_1 \cdot e^{-ik_x d_1} + s_2 \cdot e^{ik_x d_2} & 1 \end{bmatrix} \end{cases} \tag{2-21}$$

In the simulations, we fix the width of A-site waveguides to be 450 nm and decrease the width of the B-site waveguides to tune the energy offset Δ. Note that a library of the width-n_{eff} relation should be acquired beforehand. Then we fix d_2=700 nm while sweep d_1, observing topological phase transition points for the left edge (high energy side) and the right edge (low energy side), respectively. From the result shown in Figure 2-6(d&e), we can translate the distance to the coupling coefficient and overlap integral. The result is summarized in Figure 2-7.

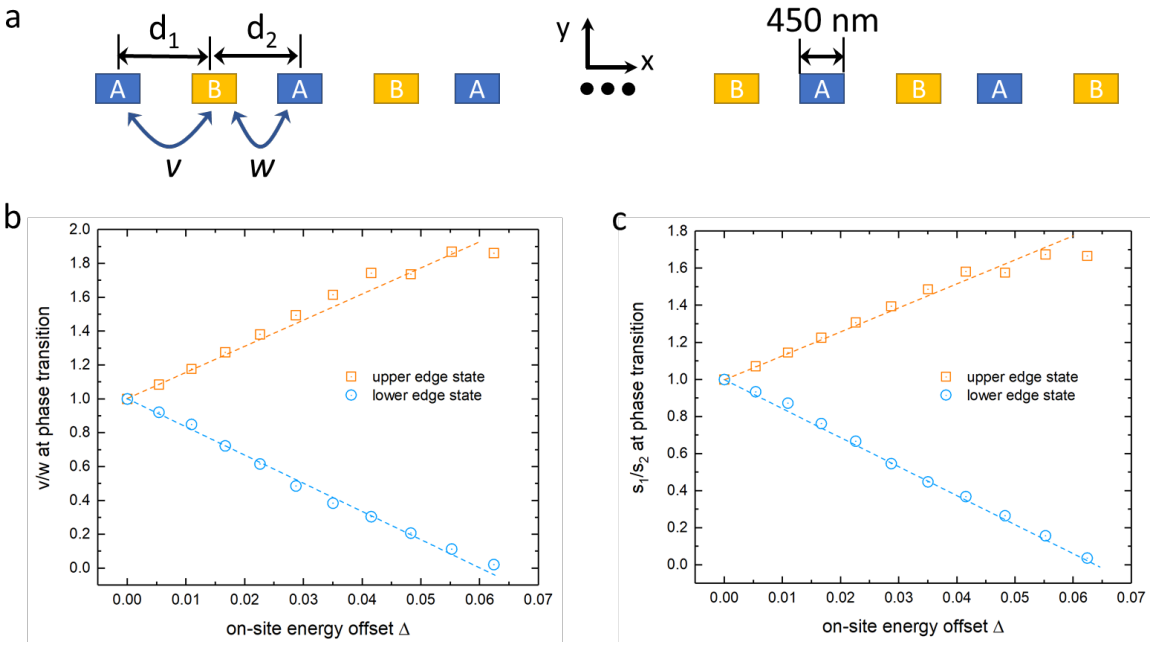

Figure 2-7 | a. Schematic of the waveguide array, where d_2 is fixed at 700 nm, A-site waveguides have a fixed width of 450 nm. **b&c.** The order parameter v/w and overlap integral ratio at the topological phase transition as a function of energy offset Δ. The fitting indicate a linear relationship.

When Δ=0, the topological phase transition occurs at v=w, which is anticipated from the SSH model. As the energy offset increases, the transition points on the parameter space v/w split between the high energy and low energy sides. The fitting indicates a linear relationship between the transition point and Δ. The plot of the energy integral s at the transition point follows a similar trend. A more rigorous derivation is necessary to understand this phenomenon further. It should be considered as a generalized SSH model, and we believe it could deepen our understanding of the topological phases.

Finally, we want to remark on the more general on-site energy setting, which is complex on-site energies. By introducing *PT*-symmetry breaking in the optical waveguide array, it is possible to create exact zero-mode and eliminate coupling between two edge states[104]. We can also predict

that if a hypothetical ZGNR lattice has complex on-site energy that breaks *PT*-symmetry, the edge states should also be exact zero-modes without crosstalk. Although intriguing, the discussion of *PT*-symmetry is beyond the scope of this chapter.

§2.5 Topological channels without crosstalk

In this section, we would like to discuss a potential application of the staggered waveguide arrays. In the realization of the SSH model, the so-called energy states on both ends are degenerate, because they are both "zero-energy" state. There is an essential claim in the topological photonics community that we can use topologically protected states for communications. However, if the edge states can crosstalk when the decay length is comparable to the system size, the communication channel could be compromised. Conversely, an extensive array would defeat the purpose of photonic integration. Here we proposed to use the staggered waveguide array so that the degeneracy of the edge states is lifted.

We begin by investigating a waveguide array with one kind of waveguide. It is similar to the system we used in section 2-3, but with a smaller size. Figure 2-8(a) is the schematic of the device on the 220-nm SOI wafer; the waveguide width is 450 nm, the cladding is silica, the working wavelength is 1550 nm.

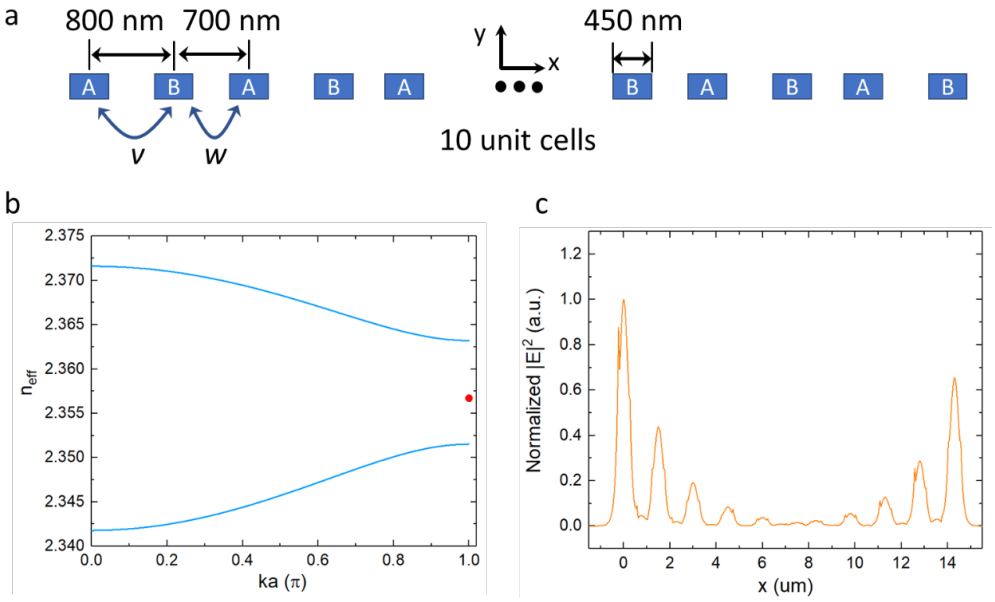

Figure 2-8 | a. Schematic of a compact waveguide array without energy offset. **b.** The bulk bandstructure of the array (blue curve) and the position of the edge state (red dot). **c.** The wavefunction plot of the left edge state. There is substantial energy couples to the opposite edge.

The edge state wavefunction in Figure 2-8(c)decays exponentially from the left edge, and mainly exists on the odd sites. However, there is a substantial mode profile at the right edge because the

states on both edges are degenerate. In other words, this device allows the light lanched into one topological channel to bounce between two boundaries, which is not ideal for communications.

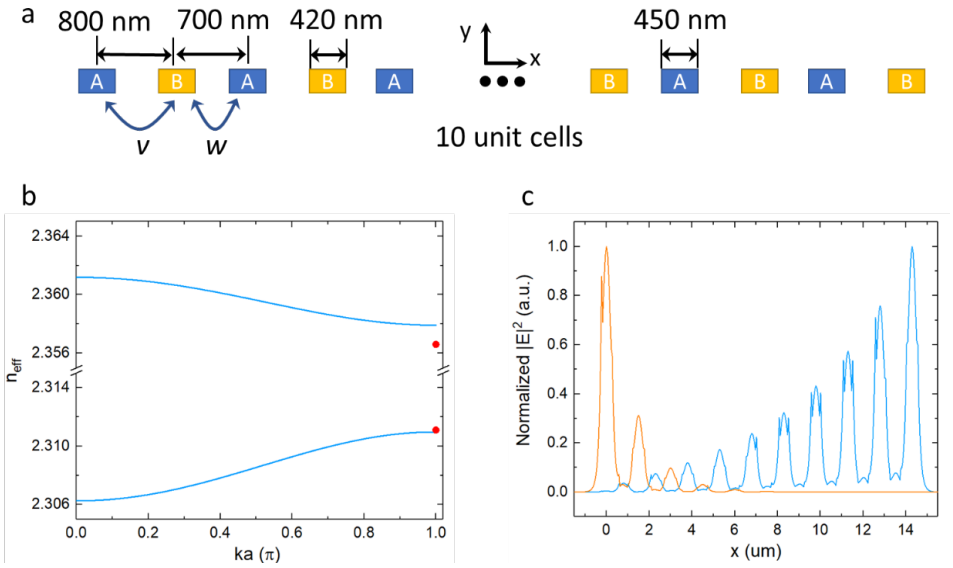

Figure 2-9 | a. Schematic of a compact waveguide array with energy offset. **b.** The bulk bandstructure of the array (blue curve) and the position of the edge states (red dot). Note that the two edge states decouple. **c.** The wavefunction plot of the left (orange) and right (blue) edge states. There is no crosstalk between the two edges.

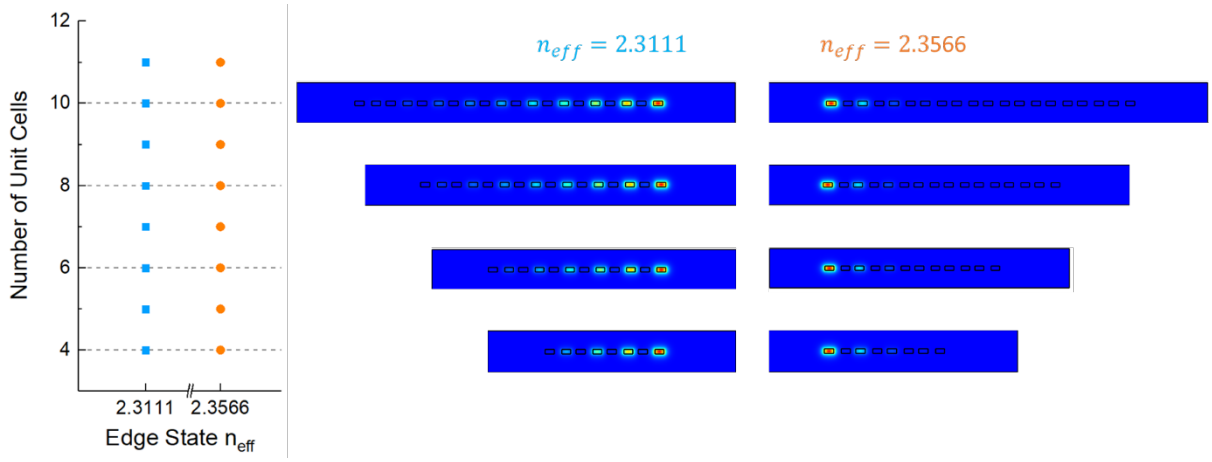

Figure 2-10 | The eigenvalues (n_{eff}) of the two topological states remain constant while decreasing the system size (number of unit cells). The field plot also confirms that the communication channels remain robust without crosstalk.

Figure 2-9(a)shows the schematic drawing of the optical array with the staggered on-site energies. It is the same as the previous device except for we replacing all the even sites with 430-nm-wide waveguides. It supports edge states at the non-trivial topological phase (Figure 2-8(b)), but the energies of those two edge states are ±Δ, respectively. The chiral symmetry of the SSH model is

broken by staggered on-site energy, and the degeneracy of the edge states is lifted. As a result, those two edges states do not cross-talk (Figure 2-9(c)). It allows us to design a compact optical lattice and still preserve the topologically protected states.

To demonstrate the feasibility of the topological states with a small footprint, we can reduce the size of the lattice (number of unit cells) and examine the edge states. Figure 2-10 shows that the mode indices of two edge states remain constant even we only have four unit cells. Their corresponding wavefunctions still demonstrate robust edges states when we decrease the system size. This result is promising for realizing topological states in a compact device.

Chapter 3: Bound-states-in-continuum on a Chip

§3.1 Review of the photonic crystal and group theory

In Chapter 1, we have discussed the continuous control of the geometric phase; in Chapter 2, we also investigated the realization of topological states in the photonic lattice as a counterpart of condensed matter. The non-trivial topological state is always associated with a non-zero topological invariant, which is the Chern number in the previous example. It counts the number of "charges" that causes the Berry flux, which is like the magnetic monopole in the reciprocal space[27,29]. If one energy band encloses one such "monopole" with a "charge" of +1, it means the band's Chern number is +1, and so forth. In other words, the typical topological state is associated with the singularity of the Berry flux. Based on this conclusion, we can generalize the concept and look for other singularities. For example, the polarization singularities in the photonic crystal (PhC) slabs can also carry topological charges[27,105]. Those states are called "bound-states-in-continuum" or BICs.

In general, BICs are confined resonances amid a continuous spectrum of radiating channels. This phenomenon was first proposed in the quantum mechanics but recently proven to be a much generic behavior of all physical waves[106]. Notably, the PhC slabs have become a promising platform for BIC studies[107]. Initially, the symmetry mismatch between some modes at high-symmetry points, e.g., Γ points, of the PhC slabs, and the free space polarizations will introduce robust BIC states protected by the symmetry[26]. Recently, a fine-tuned BIC off the Γ point was proposed as a result of the perfect cancellation of out-going waves[108].

Besides giving birth to rich physics, BIC states and near BIC states are inherently high quality factor (high-Q) resonance modes, which bring about potential applications such as lasing[109,110]. Also, the nature of the polarization singularity at BIC states promises a lasing source with polarization vortex, which gives rise to a new degree of freedom in optical communications[111,112].

In this chapter, we will discuss the BICs in the photonic crystal and the realization in the silicon photonics. To better understand this phenomenon, we need to review the group theory and the symmetry arguments in PhCs.

Photonic crystals are artificial structures that can scatter light periodically[107]. It uses the mode of resonance to mimic the Bloch wave in the conventional crystals. The periodic refractive index distribution provides the periodic potential for photons. Therefore, each photon frequency (energy) is associated with one or a few k-vectors (momentum), which results in the dispersion of the PhC. The 1D PhC is widely applied as the "Bragg reflector" or "dielectric mirror." In other words, we use the bandgap of the 1D PhC to reflect light. Similar to semiconductors, we can introduce defects in the PhC and create defect modes. In the case of Bragg reflector, the defect layer can bring about a sharp transmission peak, and the device is called "Bragg filter"[113]. This idea can be extended to 2D or 3D, where the refractive index has 2D or 3D periodicity. The defect in the PhC can also give rise to optical cavities, whose working frequency is in the bandgap of the matrix PhC[114].

We shall narrow our scope to 2D PhCs, where the BIC concept is most relevant. Figure 3-1(a) shows the cross-section view of the square lattice PhC. Typically the PhC consists of two materials with low and high indices, respectively. In the BIC applications, the two materials are both transparent at the frequency of interest, while in some microwave demonstrations, metal or water

can be used as high index materials[115–117]. In a 2D PhC, the index distribution is on the *xy*-plane, while the *z*-direction is assumed to be infinitely long. There are two ways to construct a 2D PhC, using the high index material as the matrix (background) and low index material as the pattern or vice versa. In a practical device at optical frequencies, it is often more convenient to start from a bulk matrix and etch into it rather than assembling high index wires. Therefore, without loss of generality, we shall focus on the cases where the high index material serves as a matrix. In terms of material choices, silicon (Si) or silicon nitride (SiN) are often used as high index media, while silica (SiO_2) or air is the typical low index media.

Before proceeding, we must define the polarization of the modes inside PhCs[26]. Based on the schematic, we shall call the *xy*-plane as "in-plane" while *z*-direction as "out-of-plane direction." The electric field and magnetic field for TE and TM modes in the PhC are summarized in Table 3-1. Note that rigorously these modes should be called TE-like and TM-like modes, as in the waveguides, because they are not purely transverse. For convenience, we shall just address them as TE and TM modes in the following context.

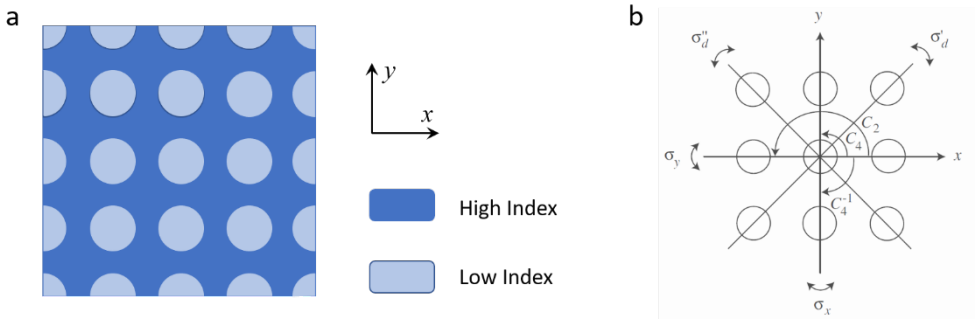

Figure 3-1 | a. Cross-section of a 2D photonic crytal with a square lattice. The matrix is made of high index material while the holes (cylindars) are low index. **b.** The symmetry operations in the square lattice. They all belong to the C_{4v} group.

	TE (like)	TM (like)
E-field	In-plane (E_x, E_y)	Out-of-plane (E_z)
H-field	Out-of-plane (H_z)	In-plane (H_x, H_y)
Other Names	*H*-polarization	*E*-polarization

Table 3-1 | The polarizations in the 2D photnic crystals.

Next, we need to understand the role of symmetry in defining the modes in the PhC. Following the schematic of the square lattice, we shall plot all the symmetry operations of the lattice in Figure 3-1(b). Together with the identity operation (basically no operation), all the symmetry operations constitute the symmetry group of the square lattice, which is C_{4v}.

$$C_{4v} = \{E, C_4, C_4^{-1}, C_2, \sigma_x, \sigma_y, \sigma_d', \sigma_d''\} \tag{3-1}$$

A group is a set of elements that satisfy four fundamental properties: closure, associativity, the identity property, and the inverse property. The closure property means that the product of any two

elements in the group is also an element in the group. The associativity means that the multiplication of elements is associative, just like matrix operations. The identity property means that there is an identity element in the group, multiply which with any element in the group results in the element itself. The inverse property means that any element in the group has its inverse element, which is also in the group. The product of one element and its inverse element is the identity element.

Apparently, all the symmetry operations in the C_{4v} group satisfy the properties. We can conveniently construct a multiplication table, describing the product of any two elements in the group based on the closure property (the first column and the first row are the first and second operations, respectively).

C_{4v}	E	C_4	C_4^{-1}	C_2	σ_x	σ_y	σ_d'	σ_d''
E	E	C_4	C_4^{-1}	C_2	σ_x	σ_y	σ_d'	σ_d''
C_4	C_4	C_2	E	C_4^{-1}	σ_d''	σ_d'	σ_x	σ_y
C_4^{-1}	C_4^{-1}	E	C_2	C_4	σ_d'	σ_d''	σ_y	σ_x
C_2	C_2	C_4^{-1}	C_4	E	σ_y	σ_x	σ_d''	σ_d'
σ_x	σ_x	σ_d'	σ_d''	σ_y	E	C_2	C_4	C_4^{-1}
σ_y	σ_y	σ_d''	σ_d'	σ_x	C_2	E	C_4^{-1}	C_4
σ_d'	σ_d'	σ_y	σ_x	σ_d''	C_4^{-1}	C_4	E	C_2
σ_d''	σ_d''	σ_x	σ_y	σ_d'	C_4	C_4^{-1}	C_2	E

Table 3-2 | The multiplication table of the C_{4v} group.

Based on the multiplication table, it is possible to find a set of eight matrices that satisfy the entire multiplication table. This set is called a representation of the group. There are infinitely many representations that a group has, but there are finite irreducible ones that are entirely made of irreducible matrices. We can summarize the irreducible representations and the trace of each matrix. The group theory tells us that the trace of the matrix is precisely the character of this operation under a particular representation. This table is called the character table of a group, where the first row lists all the elements in the group, and the first column lists all the irreducible representations.

C_{4v}	E	$2C_4$	C_2	$2\sigma_v$	$2\sigma_d$	Functions
A_1	1	1	1	1	1	z, x^2+y^2, z^2
A_2	1	1	1	-1	-1	R_z
B_1	1	-1	1	1	-1	x^2-y^2
B_2	1	-1	1	-1	1	xy
E	2	0	-2	0	0	$(x,y),(R_x,R_y),(xz,yz)$

Table 3-3 | The character table of the C_{4v} group.

Note that we have combined some operations for concise notations: two 90-degree rotations are combined, as well as two vertical (v) mirror reflections and two diagonal (d) mirror reflections. From the table, we shall see that there are five irreducible representations for the C_{4v} group. Since the matrix representation of an identity operation $\boldsymbol{E}$ is always an identity matrix, the shaded column (the trace of the identity matrix) gives the information of matrix dimensions of each representation. Interestingly, all the representation except for E is one-dimensional, while E is two-dimensional. Note that we shall not confuse the identity operation $\boldsymbol{E}$ with the representation E. In the last column, we list some functions that "transform according to each representation." It means that the function follows the spatial symmetry of that representation.

Without loss of generality, we can use the TM mode as an example. The wave equation in the PhC is expressed as:

$$\mathcal{L}_{TM}E_z(\boldsymbol{r}) = -\frac{1}{\varepsilon(\boldsymbol{r})}\nabla_{\boldsymbol{r}}^2 E_z(\boldsymbol{r}) = \frac{\omega^2}{c^2}E_z(\boldsymbol{r}) \tag{3-2}$$

where ε is the permittivity, $\boldsymbol{r}$ is the in-plane position vector. It can be proved that all the symmetry operators commute with the $\mathcal{L}_{TM}$. This result has very profound indications: $\mathcal{L}_{TM}$ and symmetry operations share the eigenvalues and eigenvectors. In other words, we can assign each resonance mode with a unique irreducible representation, and the mode must follow the spatial symmetry of that representation, or transform according to that representation.

Eventually, we want to investigate the symmetry of modes on the PhC dispersion curves, so instead of the group for the real space crystal, we must know the group in the reciprocal space. Figure 3-2(a) shows a schematic reciprocal space of the 2D square lattice. The Γ point always retains all the symmetry operations as the real crystal, so its symmetry group is C_{4v}. The high symmetry points X and M are C_{2v} and C_{4v}, respectively. The paths between the high symmetry groups are all C_{1h} with lower symmetry. The irreducible representations are C_{2v}={A_1, A_2, B_1, B_2}; C_{1h}={A, B}.

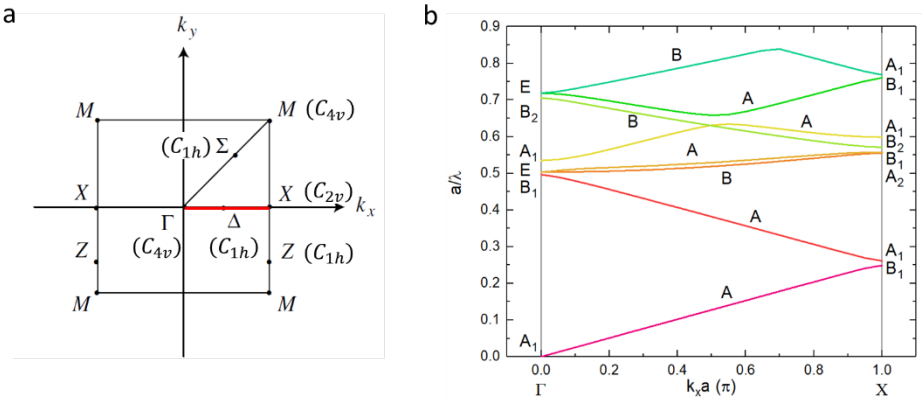

Figure 3-2 | a. The first Brillouin zone of the square lattice. Notations include the high symmetry points, paths, and their corresponding group. The red line is the path that used in the band structure calculation. **b.** The TM band structure of the 2D square lattice along the Δ path. All the modes are assigned by their irreducible representations.

Next, we can assign the modes on the dispersion curves with one of the irreducible representations of that k-point. In Figure 3-2(b), we provide the example of the normalized TM bandstructure of the square lattice. The radius is 0.2; the lattice constant is 1. The low refractive index is 1.46 (SiO_2);

the high refractive index is 2.02 (Si_3N_4). Simulations were performed with the finite-difference time-domain (FDTD) method, using an open-source software package *Meep*[118]. The code used in this calculation is in Appendix A.

Any mode on the dispersion should follow the spatial symmetry of its unique irreducible representation, and it is also the way of assigning these representations. For example, the B_1 mode should be symmetrical under C_2 and σ_v operations, but anti-symmetrical under C_4 and σ_d operations. More importantly, E representation is the only two-dimensional representation, so E modes must be doubly degenerate. Indeed, we need to assign two doubly degenerate modes as E modes. The degeneracy lifts as the mode departs from the Γ point because E mode is a degenerate representation in the C_{1h} group and will reduce to A and B modes.

The character table is a powerful tool to understand the symmetry in the PhCs. We need to point out that the symmetry argument is not exclusive to the PhC. On the contrary, this methodology is carried over from the condensed matter. The electronic states in the solid also need to transform according to the representations of their symmetry group, because the Hamiltonian commutes with symmetry operations. Nevertheless, this argument is so essential in PhCs, since an optical simulation can clearly show the symmetry of the mode, which enables convenient assignments of the representations. In the condensed matter, the symmetry argument cannot take into account effects such as Columb interactions, so the simulations might not reflect the characteristics predicted by symmetries. In the next section, we will discuss how the character table (Table 3-3) has already predicted the existence of BICs and their nature of polarization singularities.

§3.2 Bound-states-in-continuum in the photonic crystal slabs

In the last section, the majority of discussions are based on 2D PhCs, where the light always propagates inside the PhC. In other words, the modes are all "bound states," and there is no coupling with the "continuum," i.e., air, or any homogeneous medium. However, if we restrain the z-direction of the 2D PhC into a finite thickness h, the 2D PhC becomes the so-called 2D PhC slab, and there is a chance that light inside the PhC slab can leak into the environment, i.e., the continuum.

We will keep using a square lattice as an example (Figure 3-3(a)). In the lateral view, any mode that propagates in the slab falls into one of the two categories: one (red arrows) is called "bound state," where the light goes through total internal reflection, and there is no leakage to the continuum; the other (green arrows) is called "leaky mode," where light can be refracted into the continuum.

The boundary between the bound state and the leaky mode is determined by the critical angle in the total internal reflection. Here we would like to give a more straightforward approach. Assume the environment refractive index is n_{env}. The in-plane wavevector of the mode in the slab is $k_{//}$ with frequency ω. The bound state must have an evanescent out-of-plane wavevector in the environment, which means the $k_{//}$ must be larger than the total wavevector in the environment. So the condition of the bound state is:

$$k_{//} > \frac{\omega \cdot n_{env}}{c} \tag{3-3}$$

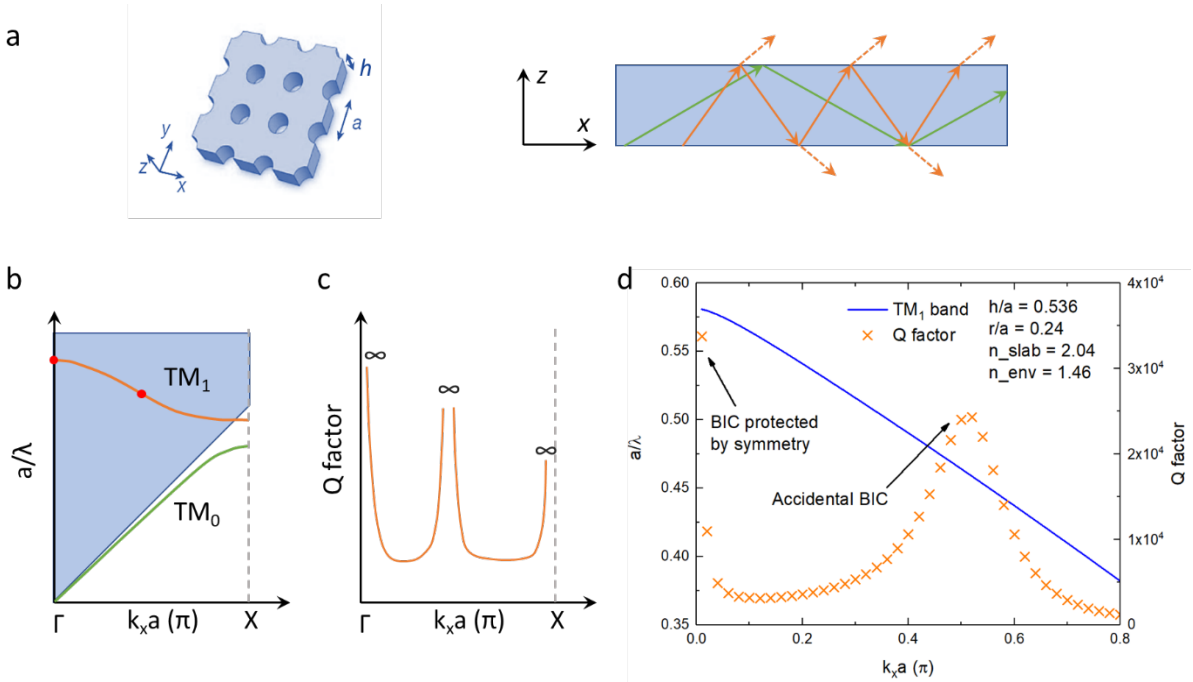

Figure 3-3 | a. The 2D PhC slab is created by limiting the *z*-direction to a finite thickness *h*. The lateral view shows the bound state propagation (green) and leak mode propagation (orange). **b.** The schematic TM band structure of the slab along the Δ path. The shaded area is the light cone and the red dots are the BIC modes. **c.** The schematic Q factor of the modes on the TM_1 band. Three ∞ points/region from left to right correspond to symmetry-protected BIC, accidental BIC, and modes outside the light cone. **d.** Lumerical FDTD simulation of the TM_1 band disperstion and the Q factor. Two BIC modes show exceptionally large Q.

In other words, the boundary between leaky modes and the bound states in the bandstructure has a slope of c/n_{env}. In Figure 3-3(b), we use the shaded area to show this "light cone." The modes inside the light cone are leaky modes (finite decay rate, finite lifetime), while the ones outside the light cone are bound states (zero decay rate, infinite lifetime), assuming lossless materials. However, there are exceptions to the leaky modes, which are the BICs in the PhC slabs. These modes do not decay into the continuum despite that they are inside the light cone.

The first kind of BIC is protected by symmetry or symmetry mismatch. It is common for modes at high-symmetry points, such as the Γ point. In the last section, we have shown that each mode at the Γ point must transform according to its representation. To leak (couple) from the PhC slab to the continuum, the symmetry of the mode must match the symmetry of light in the continuum. In other words, the mode in the slab and the refracted wave in the continuum must transform according to the same representation. In the case of the Γ point, the polarization vector of light in the continuum follows the function of (x, y), which means that it transforms according to the E representation in the square lattice[26] (See Table 3-3). Therefore, all the modes at the Γ point inside the light cone cannot decay except for the one with E representation, i.e., the doubly degenerate modes.

We plot the schematic TM_0 and TM_1 bands of the square lattice in Figure 3-3(b), and the TM_1 band has a mode of B_1 representation at the Γ point. It results in a BIC protected by the symmetry (the left red dot in Figure 3-3(b)). The symmetry mismatch has further implications: the

polarization of such BIC is undefined. People have successfully demonstrated that the BIC point is the crossing of the two polarization nodal lines, where E_x=0 and E_y=0, respectively[105]. So the BIC point is the singularity of the polarization. We can define the topological charge of such singularity:

$$q = \frac{1}{2\pi}\oint d\boldsymbol{k}\cdot\nabla_{\boldsymbol{k}}\varphi(\boldsymbol{k})\ , q\in\mathbb{Z} \tag{3-4}$$

where φ is the polarization angle. If we substitute the gradient of φ with the Berry connection, then q is exactly the Chern number in topological insulators. This equation also means that in the reciprocal space, there is a polarization vortex around the BIC. It predicts the vortex-polarization lasing from the BIC modes[111,112].

Another type of BIC in the PhC slab is the accidental BIC, marked by the red dot in the middle of the TM_1 band in Figure 3-3(b). There are two kinds of resonances that coexist in the PhC slab. The first is the Fabry–Pérot resonance from the homogeneous slab (with averaged refractive index), the second is the PhC modes. The former one can be considered as the background scattering, where the amplitude varies slowly with energy. However, the resonance of the PhC modes changes both magnitude and phase quickly. Those two resonance couples together to create so-called Fano resonance.

It can be proved using the coupled-mode-theory (CMT)[107] that if the decay rates above and below the slab are the same (by matching the refractive indices), it is possible that the outgoing waves entirely cancel out due to destructive interference. In this case, the radiative decay rate becomes zero, and the resonance mode becomes a bound state.

The Fano feature equation can be expressed as[108]:

$$f(\omega) = \frac{Q_r^{-1}}{2i\left(1-\frac{\omega}{\omega_o}\right)+Q_r^{-1}+Q_{nr}^{-1}}(r_{slab}-t_{slab}) \tag{3-5}$$

where Q_r is the quality factor of the resonance, Q_{nr} is the quality factor due to intrinsic loss, ω is the probe beam frequency, ω_o is the resonance frequency, $r_{slab,}$ and t_{slab} are the reflectivity and transmissivity of a homogeneous slab, respectively. At the BIC condition, $Q_r \rightarrow \infty$, $f(\omega_o)\rightarrow 0$, the Fano feature disappears, the tilted incident light that matches the in-plane k-vector of BIC mode can only "see" a homogeneous slab, but not the mode of the PhC. Note that zero decay rate also means zero coupling rate, as required by time-reversal symmetry. This feature of the BIC is the principle of the device proposed in the next section.

If we follow the TM_1 band from Γ point to X point, the schematic Q factor will show three peaks (to infinity), as shown in Figure 3-3(c). The first peak is due to the symmetry-protected BIC, and the second peak is due to the accidental BIC. Then the band extends outside the light cone, and all the modes have infinite lifetimes.

We should point out that the topological nature of symmetry-protected BICs is also true for accidental BICs[105]. BICs are robust topological charges and polarization singularities in the reciprocal space for both cases. By engineering the device, we can shift the position of accidental BICs, in which process the topological charges can be created or annihilated following the symmetry of the lattice, while the total charge remains conserved. Those features create

connections between the energy band topology and polarization of BICs, which is quite a fantastic physical picture.

Finally, we can simulate the BIC mode using the FDTD method. The simulation is carried out in Lumerical FDTD, although the Meep package should have the capability to achieve similar results, as shown in papers cited in this section. The codes used to control the simulation is in Appendix B. Firstly, we need to use the 2D bandstructure result (Figure 3-2(b)) to define the frequency range for the TM_1 band. Then, the FDTD solver scans the k-vector and find the eigenfrequency of each k. As we are concerned about the modes inside the light cone, the eigenfrequency should be a complex number except for BIC states:

$$\omega_{TM_1} = \omega_r + i\omega_i \tag{3-6}$$

where ω_r is the resonance frequency, and ω_i implies the decay rate. The Q factor is:

$$Q = \frac{\omega_r}{2\omega_i} \tag{3-7}$$

It is worth mentioning that this method can be tricky around high-Q value modes. The FDTD solver excites mode using multiple dipoles, and monitor the lifetime of the mode, which can be translated to ω_i. However, if the decay rate is very low, as in the BIC or near BIC states, the mode cannot "die out" at the end of the simulations time, which will create artifacts in the result. The solution is to track the slope of the electric field as a function of time, rather than waiting for the mode to dissipate completely. It requires careful apodization when processing the raw data,

The specifications of the slab (square lattice), the TM_1 band, and quality factor are all plotted in Figure 3-3(d). The dispersion agrees well with the 2D simulation, and we can see the two peaks of Q factors inside the light cone suggesting the two BIC states. Note that the Q factor peak for the accidental BIC is broader compared to the symmetry-protected BIC. People have used this feature to engineer the device with fabrication tolerances because near-BIC states still have very large Q values[109,110]. Also, one can merge multiple BIC states by engineering the PhC slab so that the resulting Q factor peak can be further broadened[119].

§3.3 Bound-states-in-continuum in silicon photonics

To date, most of the photonic BIC demonstrations and its applications have been realized in the PhC slab systems. Although PhC slabs have their engineering flexibilities, such as material, lattice symmetry, and band structure engineering, the success of BIC realizations in PhC slabs requires a large area of low defect periodic structures[108,120]. It poses a substantial challenge to the application of BIC states in the photonic integrated circuits (PICs). Additionally, the dynamic tuning of PhCs remains impractical, which further shadows the feasibility of BIC applications.

Besides the enhanced light-matter interaction in lasing applications, however, the high-Q nature of BICs also poses an opportunity of versatile filters, which is compatible with PIC and complementary metal oxide semiconductor (CMOS) technology[92]. By definition, high-Q modes have long lifetimes and low coupling rates (coefficients) to/from the radiation continuum. At the BIC or near-BIC point, the photon can hardly "see" the resonator system but instead passes through[108], which would become the fundament of a transparency window.

In this section, we demonstrated such filter devices on the active PIC platform, where the existence of the transmission peak, together with the peak position, peak width (Q-factor), can all be independently and dynamically controlled. We believe it is a leap forward towards the practical applications of BIC states, and it provides new methodologies to PIC device designs.

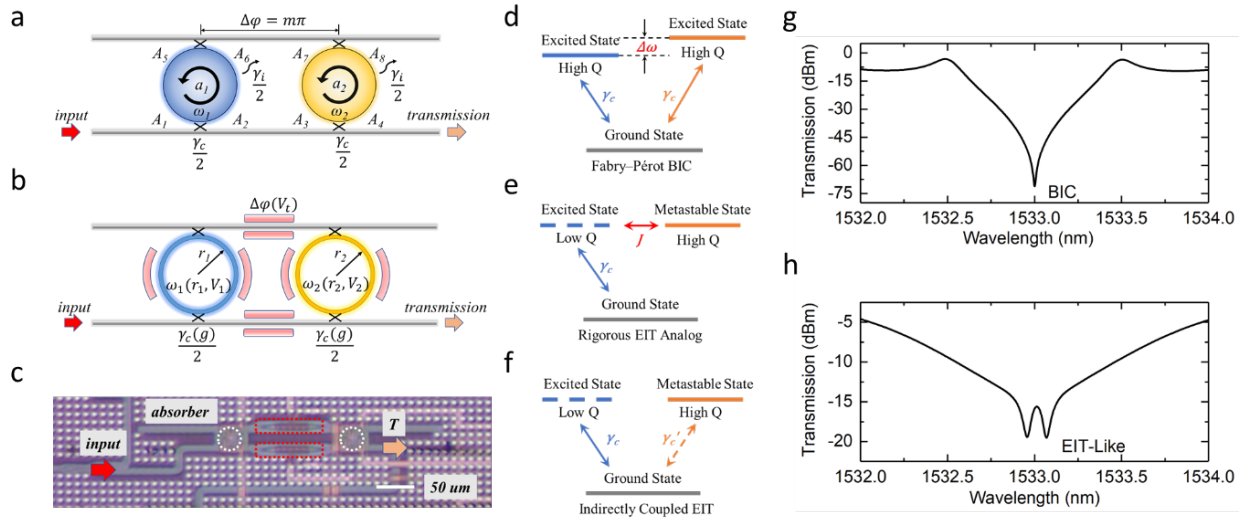

Figure 3-4 | a. The schematic model of the dual-resonator system that demonstrates Fabry–Pérot BIC. A_1~A_8 are the electric field amplitudes in the bus near the coupling ports, a_1 and a_2 are the amplitudes of the resonance modes at steady state. ω_1, ω_2 are the intrinsic resonance frequencies of the two resonators. The intrinsic decay rate is $\gamma_i/2$ for both resonators; the coupling rate is $\gamma_c/2$ at all the gaps. **b.** Schematic of a silicon photonics realization. Two ring resonators have slightly different radii (r_1 & r_2) coupling through two bus waveguides. The gaps (g) between rings and buses are of the same size to enforce same coupling rate $\gamma_c(g)/2$. Thermo-optic heaters (pink bars) tune ω_1, ω_2 and $\Delta\varphi$ with independent control voltages (V_1, V_2 & V_t). **c.** Optical image from the fabricated dual-resonator device. White circles denote the ring resonators, red rectangles are the thermo-optic phase shifters. **d - f.** Energy level diagrams of Fabry–Pérot BIC, rigorous EIT analog, and indirectly coupled EIT, respectively. $\Delta\omega$ is the frequency offset between two resonators, J is the direct coupling rate between two resonators, $\gamma_c(\gamma_c')$ is the coupling rate between ground state and excited states. **g & h.** Simulation of BIC & EIT-like (Fano resonance) transmission with zero energy offset ($\Delta\omega$=0) between two resonators.

We utilize the Fabry–Pérot (FP) BIC states in this section[106]. Figure 3-4(a) shows a schematic structure of its realization: the two cascaded resonators have their intrinsic resonance frequencies ω_1 and ω_2, respectively. The two resonators do not crosstalk directly, instead they only couple through the radiation continuum (the bus waveguides). If we assume the coupling rates at all gaps are $\gamma_c/2$, and the intrinsic loss rate $\gamma_i/2$ is negligible, the Hamiltonian of the resonator system is:

$$H = \begin{pmatrix} \omega_o + \Delta & 0 \\ 0 & \omega_o - \Delta \end{pmatrix} - i\frac{\gamma_c}{2}\begin{pmatrix} 1 & e^{i\Delta\varphi} \\ e^{i\Delta\varphi} & 1 \end{pmatrix} \tag{3-8}$$

where $\omega_o=(\omega_1+\omega_2)/2$ is the center frequency, $\Delta=(\omega_1-\omega_2)/2$ is the resonance detuning, $\Delta\varphi$ is the phase shift between the two resonators. Then we can further simplify the model by assuming $\Delta\varphi$ is an integer multiple of π, i.e., $\Delta\varphi=m\pi$, so that the round-trip satisfies the FP resonance condition (we shall call it "system on-resonance"). The eigenvalue of the resonator system is, therefore:

$$\omega_{\pm} = \omega_o - i\frac{\gamma_c}{2} \pm \sqrt{\Delta^2 - \frac{{\gamma_c}^2}{4}} \tag{3-9}$$

If there is no resonance detuning, i.e., $\Delta=0$, ω_o is the real eigenvalue of the system. It means the life-time of this mode is infinitely long, and this is the FP BIC state[106].

However, if we allow a small frequency detuning compared to the coupling rate ($\gamma_c/2 \gg \Delta$), we can still write the eigenvalue as

$$\omega_{\pm} \approx \omega_o - i\frac{\gamma_c}{2} \pm i\frac{\gamma_c}{2} \tag{3-10}$$

Then ω_o is a near-BIC state of the resonator system, which promises for high transmission.

Now we can derive the transmission characteristics for the resonator system using the temporal-coupled-mode theory[121–123]. The probe beam's frequency is ω, and the detune frequencies between the probe beam and two resonators are $\Delta_{1(2)}=\omega-\omega_{1(2)}$, respectively. The intrinsic decay rate is $\gamma_i/2$ for both resonators; the coupling rate is $\gamma_c/2$ at all the gaps. Following the electric field amplitude denoted in the schematic drawing (Figure 3-4(a)), we can describe the amplitude relations at steady state using a system of equations[121]:

$$\begin{cases} \left(i\Delta_1 - \left(\frac{\gamma_i}{2}+\gamma_c\right)\right)a_1 - \sqrt{\gamma_c}A_1 - \sqrt{\gamma_c}A_6 = 0 \\ \left(i\Delta_2 - \left(\frac{\gamma_i}{2}+\gamma_c\right)\right)a_2 - \sqrt{\gamma_c}A_3 = 0 \\ A_3 = A_2 = A_1 + \sqrt{\gamma_c}a_1 \\ A_6 = A_7 = \sqrt{\gamma_c}a_2 \\ A_4 = A_3 + \sqrt{\gamma_c}a_2 \end{cases} \tag{3-11}$$

The transmission is:

$$T = \left|\frac{A_4}{A_1}\right|^2 = t^2 \tag{3-12}$$

$$t = \frac{-\Delta_1\Delta_2 - i\frac{\gamma_i}{2}(\Delta_1+\Delta_2) + \frac{\gamma_i^2}{4}}{-\Delta_1\Delta_2 - i(\gamma_c+\frac{\gamma_i}{2})(\Delta_1+\Delta_2) + \gamma_c\gamma_i + \frac{\gamma_i^2}{4}} \tag{3-13}$$

If we assume $\gamma_c \gg \Delta \gg \gamma_i \rightarrow 0$, at $\omega=\omega_o$ & $\Delta_1=\Delta_2=\Delta$:

$$t_{BIC} = \frac{\Delta^2 + \frac{\gamma_i^2}{4}}{\Delta^2 + \gamma_c\gamma_i + \frac{\gamma_i^2}{4}} \approx \frac{\Delta^2}{\Delta^2 + \gamma_c\gamma_i} \tag{3-14}$$

Therefore, the transmission at the BIC point is protected, with $T \rightarrow 1$. On the contrary, if the probe frequency coincides with one of the intrinsic resonance frequencies of the two resonators, we have:

$$t_{\omega_1(\omega_2)} = \frac{\pm i\gamma_i\Delta + \frac{\gamma_i^2}{4}}{\pm i\gamma_i\Delta \pm 2i\gamma_c\Delta + \gamma_c\gamma_i + \frac{\gamma_i^2}{4}} \approx \frac{1}{1 + 2\frac{\gamma_c}{\gamma_i} \mp i\frac{\gamma_c}{\Delta}} \rightarrow 0 \qquad (3\text{-}15)$$

To summarize, in a cascaded resonator system, if we fine-tune the phase delay between two resonators to make the system on-resonance, the center frequency becomes a near-BIC point. Coupling into this mode from the continuum is difficult, which gives rise to high transmission. Additionally, the transmission at ω_1 and ω_2 is low. Therefore, we can obtain an artificial transparency window that can be fully controlled by ω_1 and ω_2: ω_o determines the peak position, while Δ controls the peak width and Q-factor. Besides, we can also turn the peak off by breaking the resonance condition or diminishing resonance detuning Δ.

Note that these transmission characteristics are often referred to in the literature as the optical analog to the electromagnetically induced transparency (EIT) in atomic physics[124–130]. However, we believe this system is not a rigorous EIT analog. Figure 3-4(d) is the energy level diagram of the FP BIC. The two energy levels (resonators) have similar Q-factors that couple indirectly through the radiation continuum. The transmission window requires the energy offset $\Delta\omega$. Figure 3-4(g) is the simulation result of the BIC state with two identical resonators ($\Delta\omega$ =0); both have a resonance wavelength of 1533 nm. The transmission spectrum shows total reflection at the BIC point.

Figure 3-4(e) is the energy level diagram of the rigorous optical analog of EIT. A low-Q resonator couples to the radiation continuum, while a high-Q resonator only couples to the low-Q resonator with a coupling rate *J*. *J* is the analog of the Rabi frequency in the original EIT[124]. This system can be realized using two directly-coupled resonators and a single bus. EIT is a particular Fano resonance where two resonators have disparate Q-factors but the same resonance frequency. In other words, their energy levels should have zero offsets[131,132]. It is possible to construct an EIT analog with indirectly-coupled resonators. However, the system still needs to follow the Fano resonance condition. We can simulate this effect by tweaking the setup following the energy level diagram in Figure 3-4(f). We shall lower the Q-factor of the first resonator and reduce the coupling rate of the second resonator while fixing their resonance wavelengths at 1533 nm. Here the Fano resonance results in the transparency window (Figure 3-4(h)). The shape of the peak is affected by the Q-factor difference and symmetry of the feedback loop[132].

In summary, the transparency window in a rigorous EIT analog is a result of Fano resonance (with two overlapping resonances), which is not valid for the near-BIC state. Phenomenologically, we can refer to both systems as "EIT-like" in terms of transmission characteristics.

Besides, we would like to address the seemingly contradicting phenomena that the exact-BIC state leads to total reflection while the near-BIC state has high transmission. BIC and near-BIC states are high-Q states where coupling from the radiation continuum is difficult (energy stays in the continuum), which does not guarantee the direction of the output energy. From Equation 3-14, the direction of output is a competing result between Δ and γ_i: for exact BIC state Δ=0, so it shows high reflection; for near-BIC state $\Delta \gg \gamma_i$, so it shows high transmission instead.

§3.4 Device realization in photonic integrated circuits

From the discussions in the last section, we shall conclude the essential requirements for a versatile BIC filter: first, the intrinsic loss needs to be small, otherwise the height of the transmission peak decreases; second, the coupling rate between the resonator and the bus should be high enough to ensure $\gamma_c \gg \Delta$; third, we need to fine-tune the intrinsic resonance frequencies and the phase delay between resonators.

We designed the proposed device using the silicon photonics platform with 220-nm silicon-on-insulator (SOI) wafers (Figure 3-4(b)). Two ring resonators have slightly different radii (r_1=14.98 um; r_2=15 um), and they are controlled by thermo-optic heaters. (controlled by voltage V_1 and V_2). Thermo-optic phase shifters connect two ring resonators, and they are controlled by voltage V_t. The bus waveguides and resonator waveguides all have a base width of 300 nm. All the coupling gaps (g) between bus waveguides and ring resonators are 300 nm to enforce identical coupling rates ($\gamma_c/2$). AIM Photonics fabricated the device through the multi-project wafer (MPW) run[133]. The simulations were carried out in Lumerical MODE and INTERCONNECT[134]. Figure 3-4(c)is the optical image of the fabricated device.

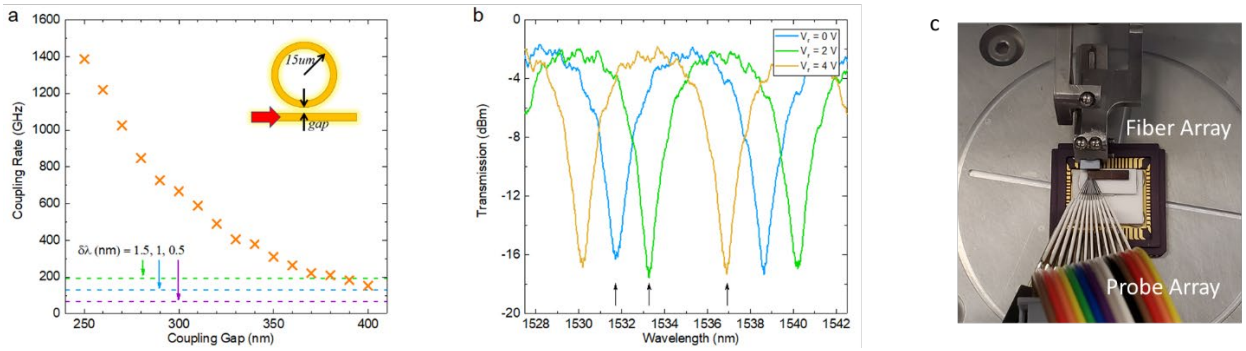

Figure 3-5 | a. Simulation of the coupling rate as a function of gap size using a 15-um-radius ring and a single bus. Waveguide width is 300 nm. Resoannce wavelength is 1533 nm. Dased lines are frequency detuning converted from wavelength detuning around 1533 nm. **b.** Measured transmission spectra of a single 15-um-radius ring resonator with double buses. The black arrow shows the same transmission dip under different ring heater voltages. **c.** Photo of the measurement setup.

We simulated the coupling rate between a 15-um ring and a single bus waveguide in Figure 3-5(a) as a function of gap size (g) with the resonance wavelength being 1533 nm. The coupling rate increases exponentially as the gap size decreases. At g=300 nm, we can satisfy the condition that $\gamma_c \gg \Delta$. In Figure 3-5(b), we show the ring heater's capability by tuning a single ring device (r=15 um) to cover most of the free spectrum range (FSR). It lays a solid foundation for the following demonstrations. Figure 3-5(c) is a photo of the measurement setup in the Rochester Institute of Technology. The fiber array and probe array are in the opposite direction of the silicon chip.

The transmission characteristics of the device shown in Figure 3-5(c) are summarized in Figure 3-6. The upper curve in Figure 3-6(a) is the example of a BIC filter, which is also the reference curve in other panels. Two rings have their intrinsic resonance wavelengths of 1532.4 nm and 1533.4 nm, respectively. The phase delay between the two rings is tuned to make the system on-resonance. The transmission at intrinsic resonance wavelengths is low, while the center wavelength is the

protected near-BIC state with high transmission. This result coincides with the model very well. If we increase the phase delay between the two rings by increasing the voltage of the phase shifter, the FP resonance condition fails. The transparency window red-shifts and the peak becomes asymmetric.

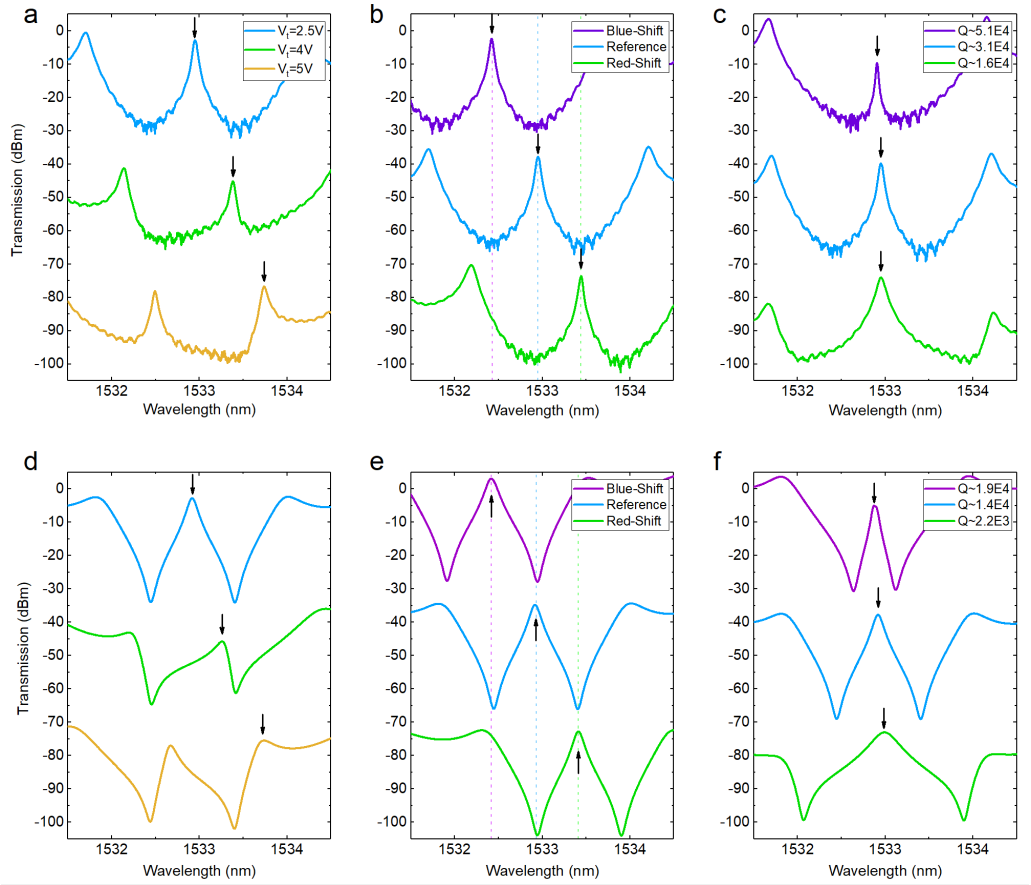

Figure 3-6 | Transmission spectra of BIC filters. **a.** Tuning the phase delay between two rings. The first curve is the on-resonace case with a symmetrical transparency window. **b.** Tuning the peak position with the fixed width. **c.** Tuning the peak width (Q-factor) with the fixed position. Spectra in **b** and **c** are all on-resonance. Blue curves in **a**, **b**, and **c** are the same. Black arrows are used to guide eyes to the peaks of interest in each panel. All the curves are shifted vertically for clear illustration. **d - f.** Simulation reproduction of **a**, **b**, and **c**, respectively.

We can independently control the peak position and peak width by tuning two ring resonators. In Figure 3-6(b), we show that we can maintain the width of the transparency window while changing the position by red-shifting or blue-shifting two rings together. We can also change the width and Q-factor of the transmission peak while fixing the position (Figure 3-6(c)). Bringing two ring resonators closer in terms of resonance wavelengths results in a narrower transparency window (higher Q-factor), and vice versa. The height of the transmission peak is a function of both detuning and loss (Equation 3-14). Thus a high-Q transparency window may have lower overall

transmission. Designing the device with low-loss waveguides could mitigate this issue. Note that every scenario in Figure 3-6(b) and Figure 3-6(c) is on-resonance, so the phase shifters between two rings are tuned when we change the peak position.

For a filter design involving a single resonator, the resonance wavelength and Q-factor are usually dependent on each other, because the change of loss often accompanies the tuning process. In the BIC filter, however, the transparency window is not the mode of an individual resonator, but the mode of the near-BIC state. Therefore, the position and width of the transparency window are independent of each other. This control flexibility is very desirable in the PICs[135].

If one intends to use the BIC filter as an optical switch, there are three approaches. We shall regard the first curve in Figure 3-6(a) as the "On" state for 1533 nm wavelength; then we can turn the transmission off by 1. breaking the FP resonance (Figure 3-6(a), 22 dB on-off ratio); 2. shifting the peak position (Figure 3-6(b), 25 dB on-off ratio); 3. constructing exact-BIC state (Figure 3-4 (g) & Figure 3-6(f), 75 dB on-off ratio).

Figure 3-6(d), (e), (f) are the simulation results reproducing the experimental data. The simulations show apparent transmission dips indicating intrinsic resonance wavelengths of the rings. Note that the curvature near the peak is slightly different in the simulations compared to the experiments, which also result in larger Q-factors for the windows (Figure 3-6(f)). It is because of the additional loss in the fabricated device that causes the transmission to fall off quickly around near-BIC points.

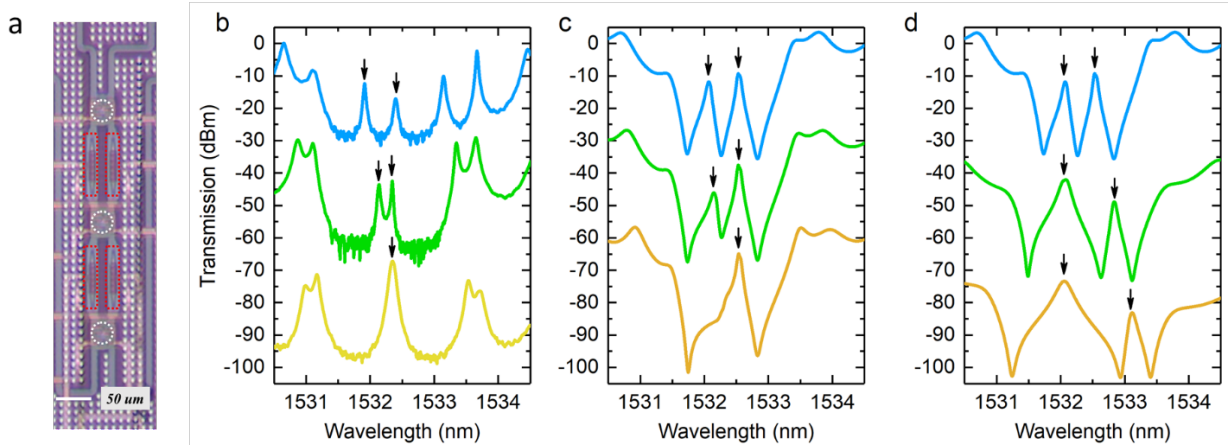

Figure 3-7 | a. The optical image of the triple-resonator device. White circles denote the ring resonators, red rectangles are the thermo-optic phase shifters. The upper and middle ring are the first loop, the middle and the lower ring are the second loop. **b.** Tuning the phase delay in the first loop. The first curve is the on-resonace case for both loops with two transmission peaks. **c.** Simulation reproduction of **a**. **d.** Tuning the right peak's position with the fixed width while tuning the left peak's width (Q-factor) with the fixed position. Spectra in **d** are all on-resonance in both loops. Blue curves in **c** and **d** are the same. Black arrows are used to guide eyes to the peaks of interest in each panel. All the curves are shifted vertically for clear illustration.

Although the discussions above are based on dual-resonator systems, we can cascade more detuned resonators sharing the same buses so that the transmission spectrum will show a series of tunable peaks[128]. As a demonstration of concept, we designed the triple-resonator device with an additional ring (r_3=15.02 um) shown in Figure 3-7(a). The two loops between two adjacent rings give rise to two transparency windows when the system is on-resonance (first curve in Figure 3-7(b)).

In Figure 3-7(b), we show the effect when we fix the resonance wavelengths of three rings and the phase delay in the second loop while increasing the phase delay in the first loop. The second peak is fixed in place, but the first peak red-shifts until it merges into the right peak. It promises the ability to have two transparency windows or kill either one of them. Figure 3-7(c) is the simulation result of the same process, confirming the effect of tuning phase delay in one of the loops.

Because of the thermal cross-talk in the fabricated device, we used simulation to demonstrate the potentials of individual peak tuning in Figure 3-7(d). Starting from two peaks with similar widths, we can control the resonances of three rings so that the center frequency in the first loop is constant while the detuning in the second loop remains fixed. The result is that the left peak has a fixed position but varying width, while the right peak has a fixed width at different positions. Note that every scenario here is on-resonance for both loops. This phenomenon follows the same principle of the dual-resonator device and further proves the versatility of the BIC filters.

In conclusion, we have demonstrated the near-BIC state using the PIC platform. The cascaded resonators bring about the near-BIC state at the average resonance wavelength, which has high transmission and is used to construct a transparency window. We argue that this phenomenon fits the BIC formalism rather than the rigorous optical analog of EIT. The devices we demonstrated are versatile filters whose peak position and width can be independently controlled. They can also be used as optical switches. We believe the results could deepen our understandings of the general BIC concept and help lower the threshold of its applications.

Conclusions & Outlooks

In conclusion, we have reviewed and discussed three aspects as to how the topological behaviors in solid-state physics can inspire the novel photonic devices. Fundamentally, those researches intend to build the connection between a Fermion system and a Boson system.

In Chapter 1, the phase manipulation of light is based on the fact that the polarization of photons can directly translate to electron spins. Therefore, the solid angle on the Poincaré sphere describes the geometric phase of a photon. It allows us to bypass the propagation phase manipulation that requires a low loss scattering medium, which is a big obstacle in the ultraviolet frequency band. Instead, we managed to use processing-friendly silicon to fabricate a metasurface that works for ultraviolet light.

This work implies that high-index dielectric materials could be candidates when metals fail to function. Our circuit model gives analytical results as to why the large imaginary permittivity of silicon does not necessarily lead to large dissipations. In the last section of Chapter 1, we proposed the silicon optical antenna for fluorescence enhancement of biomolecules, which requires ultraviolet light as the excitation source. Silicon antennas could be a promising future direction in optical antenna researches.

In Chapter 2, we reviewed the one-dimensional topological insulator model, i.e., the Su-Schrieffer-Heeger model, and how the topological phase transits. Then we investigated the photonic counterpart using a one-dimensional waveguide array, which allowed us to engineer the bonding strength and on-site energy. It leads to a clear mapping relation between the parameter space of the waveguide array and the k-space of a two-dimensional nanoribbon, which unveils the topological origins in the nanoribbons. Based on the formalism, we successfully reproduced and studied the topological edge states of the graphene nanoribbon and hexagonal boron nitride nanoribbon models, which coincide very well with tight-binding calculations.

More importantly, this method allows us to make predictions about the electronic counterparts based on the phenomena in the photonic systems. For example, the introduction of PT-symmetry could bring about special zero modes in the hypothetical nanoribbon lattice. Also, the engineering of the on-site energy indicates a possible path to control the topological phase transition point, which is exciting both theoretically and experimentally.

In Chapter 3, we discussed the group theory in the photonic crystals and the rise of the bound-state-in-continuum in the two-dimensional photonic crystal slabs. The group theory predicts that those symmetry-protected bound states are polarization singularities. This feature carried on to the accidental bound-state-in-continuum and proved its universal topological nature. Later in the chapter, we proposed the silicon photonic device with cascaded resonators to realize such particular states on a chip. Along with the powerful technology of the photonic integrated circuit, the result is a versatile filter that can be electrically tuned in every aspect. The idea of utilizing the bound state can pave the way to novel devices in photonic integration.

The idea of bound-state-in-continuum originates in quantum mechanics and has been well-received in the crystals (both electronic and photonic). The principal value of our work is to generalize this concept into a more practical platform, such as silicon photonics. We are looking forward to more filters, modulators, optical switches based on this idea in the future.

Overall, this dissertation is dedicated to discussing some essential physical pictures in the topological phenomena, and how they inspired interesting photonic devices. Through three chapters, we can see the connections between electronic and photonic systems, between low dimension and high dimension systems, and between various topological invariants. One key goal of physics researches is to study the universality of physical laws, and topology might be the lastest and greatest approach to do that.

References

1. Moore, J. E. The birth of topological insulators. *Nature* **464**, 194–198 (2010).
2. Ando, Y. Topological insulator materials. *J. Phys. Soc. Japan* **82**, 1–32 (2013).
3. Fu, L. & Kane, C. L. Topological insulators with inversion symmetry. *Phys. Rev. B - Condens. Matter Mater. Phys.* **76**, 1–17 (2007).
4. Qi, X. L. & Zhang, S. C. The quantum spin Hall effect and topological insulators. *Phys. Today* **63**, 33–38 (2010).
5. Chang, C. Z. *et al.* Experimental observation of the quantum anomalous Hall effect in a magnetic topological Insulator. *Science (80-.).* **340**, 167–170 (2013).
6. Li, J., Chu, R. L., Jain, J. K. & Shen, S. Q. Topological anderson insulator. *Phys. Rev. Lett.* **102**, 1–4 (2009).
7. Yu, R. *et al.* Quantized Anomalous Hall Effect in Magnetic Topological Insulators. *Science (80-.).* **329**, 61–65 (2010).
8. Lindner, N. H., Refael, G. & Galitski, V. Floquet topological insulator in semiconductor quantum wells. *Nat. Phys.* **7**, 490–495 (2011).
9. Hsieh, D. *et al.* A topological Dirac insulator in a quantum spin Hall phase. *Nature* **452**, 970–974 (2008).
10. Burkov, A. A. & Balents, L. Weyl semimetal in a topological insulator multilayer. *Phys. Rev. Lett.* **107**, 1–4 (2011).
11. Gibney, E. & Castelvecchi, D. Nobel for 2D exotic matter Physics award goes to theorists who used topology to explain strange phenomena. *Nature* **538**, 18 (2016).
12. Haldane, F. D. M. Model for a quantum hall effect without landau levels: Condensed-matter realization of the 'parity anomaly'. *Phys. Rev. Lett.* **61**, 2015–2018 (1988).
13. 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 (1983).
14. Haldane, F. D. M. Continuum dynamics of the 1-D Heisenberg antiferromagnet: Identification with the O(3) nonlinear sigma model. *Phys. Lett. A* **93**, 464–468 (1983).
15. kosterlitz, J. M. & Thouless, D. J. Ordering, metastability and phase transitions in two-dimensional systems. *J. Phys. C Solid State Phys.* **6**, 1181–1203 (1973).
16. Kosterlitz, J. M. & Thouless, D. J. Long range order and metastability in two dimensional solids and superfluids. (Application of dislocation theory). *J. Phys. C Solid State Phys.* **5**, (1972).
17. den Nijs, M. Quantized Hall conductance in a two dimensional periodic potential. *Phys. A Stat. Mech. its Appl.* **124**, 199–210 (1984).
18. Kittel, C. & Fong, C.-Y. *Quantum theory of solids*. *Quantum Theory of Solids* (Wiley, 1963). doi:10.1201/9781315273198.

19. Karplus, R. & Luttinger, J. M. Hall effect in ferromagnetics. *Phys. Rev.* **95**, 1154–1160 (1954).

20. Chong, Y. D. Berry's phase and the anomalous velocity of Bloch wavepackets. *Phys. Rev. B - Condens. Matter Mater. Phys.* **81**, (2010).

21. Chang, M. C. & Niu, Q. Berry phase, hyperorbits, and the Hofstadter spectrum. *Phys. Rev. Lett.* **75**, 1348–1351 (1995).

22. Simon, B. Holonomy, the quantum adiabatic theorem, and Berry's phase. *Phys. Rev. Lett.* **51**, 2167–2170 (1983).

23. Berry, M. V. Quantal phase factors accompanying adiabatic changes. *Proc. R. Soc. London. A. Math. Phys. Sci.* **392**, 45–57 (1984).

24. Aharonov, Y. & Bohm, D. Significance of Electromagnetic Potentials in the Quantum Theory. *Phys. Rev.* **115**, 485–491 (1959).

25. McGinley, M. & Cooper, N. R. Classification of topological insulators and superconductors out of equilibrium. *Phys. Rev. B* **99**, (2019).

26. Sakoda, K. Optical properties of photonic crystals. *Springer Sci. Bus. Media* **80**, (2004).

27. Lu, L., Joannopoulos, J. D. & Soljačić, M. Topological photonics. *Nat. Photonics* **8**, 821–829 (2014).

28. Ozawa, T. *et al.* Topological photonics. *Rev. Mod. Phys.* **91**, 15006 (2019).

29. Khanikaev, A. B. & Shvets, G. Two-dimensional topological photonics. *Nature Photonics* vol. 11 (2017).

30. Chen, H. T., Taylor, A. J. & Yu, N. A review of metasurfaces: Physics and applications. *Reports Prog. Phys.* **79**, 0–59 (2016).

31. Yu, N. & Capasso, F. Flat optics with designer metasurfaces. *Nat. Mater.* **13**, 139–150 (2014).

32. Balthasar Mueller, J. P., Rubin, N. A., Devlin, R. C., Groever, B. & Capasso, F. Metasurface Polarization Optics: Independent Phase Control of Arbitrary Orthogonal States of Polarization. *Phys. Rev. Lett.* **118**, 1–5 (2017).

33. Berry, M. V. The adiabatic phase and Pancharatnam's phase for polarized light. *J. Mod. Opt.* **34**, 1401–1407 (1987).

34. Chen, K. *et al.* Geometric phase coded metasurface: From polarization dependent directive electromagnetic wave scattering to diffusion-like scattering. *Sci. Rep.* **6**, 1–10 (2016).

35. van Dijk, T., Schouten, H. F. & Visser, T. D. Geometric interpretation of the Pancharatnam connection and non-cyclic polarization changes. *J. Opt. Soc. Am. A* **27**, 1972 (2010).

36. Balanis, C. A. *Antenna theory: analysis and design*. (John wiley & sons, 2016).

37. Gong, Z., Bartone, C., Yang, F. & Yao, J. High-efficiency water-loaded microwave antenna in ultra-high-frequency band. *Appl. Phys. Lett.* **112**, (2018).

38. Jackson, J. D. & Fox, R. F. Classical Electrodynamics, 3rd ed . . *Am. J. Phys.* **67**, 841–842 (1999).

39. Maier, S. A. *Plasmonics: fundamentals and applications*. (Springer Science & Business Media, 2007).

40. Novotny, L. Effective wavelength scaling for optical antennas. *Phys. Rev. Lett.* **98**, 1–4 (2007).

41. Novotny, L. & Van Hulst, N. Antennas for light. *Nat. Photonics* **5**, 83–90 (2011).

42. Eggleston, M. S., Messer, K., Zhang, L., Yablonovitch, E. & Wu, M. C. Optical antenna enhanced spontaneous emission. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 1704–1709 (2015).

43. Silver, S. *Microwave antenna theory and design*. (Peter Peregrinus Ltd., 1984).

44. Ni, X., Emani, N. K., Kildishev, A. V., Boltasseva, A. & Shalaev, V. M. Broadband light bending with plasmonic nanoantennas. *Science (80-.).* **335**, 427 (2012).

45. Yu, N. *et al.* Light Propagation with Phase Reflection and Refraction. *Science (80-.).* **334**, 333–337 (2011).

46. Zheng, G. *et al.* Metasurface holograms reaching 80% efficiency. *Nat. Nanotechnol.* **10**, 308–312 (2015).

47. Lin, J., Genevet, P., Kats, M. A., Antoniou, N. & Capasso, F. Nanostructured holograms for broadband manipulation of vector beams. *Nano Lett.* **13**, 4269–4274 (2013).

48. Yin, X., Ye, Z., Rho, J., Wang, Y. & Zhang, X. Photonic spin hall effect at metasurfaces. *Science (80-.).* **339**, 1405–1407 (2013).

49. Wu, P. C. *et al.* Versatile Polarization Generation with an Aluminum Plasmonic Metasurface. *Nano Lett.* **17**, 445–452 (2017).

50. Knight, M. W. *et al.* Aluminum for plasmonics. *ACS Nano* **8**, 834–840 (2014).

51. Deng, Y. *et al.* All-Silicon Broadband Ultraviolet Metasurfaces. *Adv. Mater.* **30**, 1–6 (2018).

52. Zhang, J., Pégard, N., Zhong, J., Adesnik, H. & Waller, L. 3D computer-generated holography by non-convex optimization. *Optica* **4**, 1306–1313 (2017).

53. Bharadwaj, P., Deutsch, B. & Novotny, L. Optical Antennas. *Adv. Opt. Photonics 1* 438–483 (2009) doi:10.1364/aop.1.000438.

54. Kerwin, B. A. & Remmele, R. L. Protect from Light: Photodegradation and Protein Biologics. *J. Pharm. Sci.* **96**, 1468–1479 (2007).

55. Fort, E. & Grésillon, S. Surface enhanced fluorescence. *J. Phys. D. Appl. Phys.* **41**, (2008).

56. Stiles, P. L., Dieringer, J. A., Shah, N. C. & Van Duyne, R. P. Surface-Enhanced Raman Spectroscopy. *Annu. Rev. Anal. Chem* **1**, 601–26 (2008).

57. Xu, Y., Tucker, E., Boreman, G., Raschke, M. B. & Lail, B. A. Optical Nanoantenna Input Impedance. *ACS Photonics* **3**, 881–885 (2016).

58. De Arquer, F. P. G., Volski, V., Verellen, N., Vandenbosch, G. A. E. & Moshchalkov, V. V. Engineering the input impedance of optical nano dipole antennas: Materials, geometry and excitation effect. *IEEE Trans. Antennas Propag.* **59**, 3144–3153 (2011).

59. Barenboĭm, G. M., Domanskiĭ, A. N. & Turoverov, K. K. *Luminescence of biopolymers and cells*. (Springer, 2013).

60. Chowdhury, M. H., Ray, K., Gray, S. K., Pond, J. & Lakowicz, J. R. Aluminum nanoparticles as substrates for metal-enhanced fluorescence in the ultraviolet for the label-free detection of biomolecules. *Anal. Chem.* **81**, 1397–1403 (2009).

61. Jha, S. K., Ahmed, Z., Agio, M., Ekinci, Y. & Löffler, J. F. Deep-UV surface-enhanced resonance Raman scattering of adenine on aluminum nanoparticle arrays. *J. Am. Chem. Soc.* **134**, 1966–1969 (2012).

62. Jha, S. K., Mojarad, N., Agio, M., Löffler, J. F. & Ekinci, Y. Enhancement of the intrinsic fluorescence of adenine using aluminum nanoparticle arrays. *Opt. Express* **23**, 24719 (2015).

63. Dorh, N. *et al.* Nanoantenna arrays combining enhancement and beam control for fluorescence-based sensing applications. *Appl. Opt.* **56**, 8252 (2017).

64. Kinkhabwala, A. *et al.* Large single-molecule fluorescence enhancements produced by a bowtie nanoantenna. *Nat. Photonics* **3**, 654–657 (2009).

65. Miller, O. D. *et al.* Fundamental limits to optical response in absorptive systems. *Opt. Express* **24**, 3329 (2016).

66. EASTMAN, J. W. & ROSA, E. J. the Fluorescence of Adenine. the Effects of Solvent and Temperature on the Quantum Yield. *Photochem. Photobiol.* **7**, 189–201 (1968).

67. Zak, J. Berrys phase for energy bands in solids. *Phys. Rev. Lett.* **62**, 2747–2750 (1989).

68. Longhi, S. Zak phase of photons in optical waveguide lattices. *Opt. Lett.* **38**, 3716 (2013).

69. Su, W. P., Schrieffer, J. R. & Heeger, A. J. Solitons in polyacetylene. *Phys. Rev. Lett.* **42**, 1698–1701 (1979).

70. Asbóth, J. K., Oroszlány, L. & Pályi, A. *A Short Course on Topological Insulators*. vol. 919 (Springer International Publishing, 2016).

71. Batra, N. & Sheet, G. Understanding Basic Concepts of Topological Insulators Through Su-Schrieffer-Heeger (SSH) Model. *arXiv Prepr.* 1–8 (2019).

72. Notomi, M., Kuramochi, E. & Tanabe, T. Large-scale arrays of ultrahigh-Q coupled nanocavities. *Nat. Photonics* **2**, 741–747 (2008).

73. Hafezi, M., Demler, E. A., Lukin, M. D. & Taylor, J. M. Robust optical delay lines with topological protection. *Nat. Phys.* **7**, 907–912 (2011).

74. Malzard, S., Poli, C. & Schomerus, H. Topologically Protected Defect States in Open Photonic Systems with Non-Hermitian Charge-Conjugation and Parity-Time Symmetry. *Phys. Rev. Lett.* **115**, 1–6 (2015).

75. Altug, H. & Vučković, J. Two-dimensional coupled photonic crystal resonator arrays. *Appl. Phys. Lett.* **84**, 161–163 (2004).

76. Hudock, J. Optical wave propagation in discrete waveguide arrays. (2005).

77. Pollock, C. R. & Lipson, M. *Integrated photonics*. (Springer Science+Business Media, LLC, 2003).

78. Dutta, S. & Pati, S. K. Novel properties of graphene nanoribbons: A review. *J. Mater. Chem.* **20**, 8207–8223 (2010).

79. Zeng, H. *et al.* 'White graphenes': Boron nitride nanoribbons via boron nitride nanotube unwrapping. *Nano Lett.* **10**, 5049–5055 (2010).

80. Ruffieux, P. *et al.* On-surface synthesis of graphene nanoribbons with zigzag edge topology. *Nature* **531**, 489–492 (2016).

81. Wakabayashi, K., Sasaki, K. I., Nakanishi, T. & Enoki, T. Electronic states of graphene nanoribbons and analytical solutions. *Sci. Technol. Adv. Mater.* **11**, (2010).

82. Nakada, K., Fujita, M., Dresselhaus, G. & Dresselhaus, M. S. Edge state in graphene ribbons: Nanometer size effect and edge shape dependence. *Phys. Rev. B - Condens. Matter Mater. Phys.* **54**, 954–961 (1996).

83. Fujita, M., Wakabayashi, K., Nakada, K. & Kusakabe, K. Peculiar Localized State at Zigzag Graphite Edge. *J. Phys. Soc. Japan* **65**, 1920–1923 (1996).

84. Niimi, Y. *et al.* Scanning tunneling microscopy and spectroscopy of the electronic local density of states of graphite surfaces near monoatomic step edges. *Phys. Rev. B - Condens. Matter Mater. Phys.* **73**, 1–8 (2006).

85. Delplace, P., Ullmo, D. & Montambaux, G. Zak phase and the existence of edge states in graphene. *Phys. Rev. B - Condens. Matter Mater. Phys.* **84**, 1–13 (2011).

86. Ryu, S. & Hatsugai, Y. Topological Origin of Zero-Energy Edge States in Particle-Hole Symmetric Systems. *Phys. Rev. Lett.* **89**, 4–7 (2002).

87. Sugawara, K., Sato, T., Souma, S., Takahashi, T. & Suematsu, H. Fermi surface and edge-localized states in graphite studied by high-resolution angle-resolved photoemission spectroscopy. *Phys. Rev. B - Condens. Matter Mater. Phys.* **73**, 3–6 (2006).

88. Blanco-Redondo, A. *et al.* Topological Optical Waveguiding in Silicon and the Transition between Topological and Trivial Defect States. *Phys. Rev. Lett.* **116**, 1–5 (2016).

89. Plotnik, Y. *et al.* Observation of unconventional edge states in 'photonic graphene'. *Nat. Mater.* **13**, 57–62 (2014).

90. Maczewsky, L. J. *et al.* Synthesizing multi-dimensional excitation dynamics and localization transition in one-dimensional lattices. *Nat. Photonics* (2019) doi:10.1038/s41566-019-0562-8.

91. Son, Y. W., Cohen, M. L. & Louie, S. G. Half-metallic graphene nanoribbons. *Nature* **444**, 347–349 (2006).

92. Bogaerts, W. *et al.* Nanophotonic waveguides in silicon-on-insulator fabricated with CMOS technology. *J. Light. Technol.* **23**, 401–412 (2005).

93. Kohmoto, M. & Hasegawa, Y. Zero modes and edge states of the honeycomb lattice. *Phys. Rev. B - Condens. Matter Mater. Phys.* **76**, 1–6 (2007).

94. Wassmann, T., Seitsonen, A. P., Saitta, A. M., Lazzeri, M. & Mauri, F. Structure, stability, edge states, and aromaticity of graphene ribbons. *Phys. Rev. Lett.* **101**, 1–4 (2008).

95. Sasaki, K., Murakami, S. & Saito, R. Stabilization mechanism of edge states in graphene. *Appl. Phys. Lett.* **88**, 10–13 (2006).

96. Ezawa, M. Topological origin of quasi-flat edge band in phosphorene. *New J. Phys.* **16**, 0–

13 (2014).

97. Zhang, K., Feng, Y., Wang, F., Yang, Z. & Wang, J. Two dimensional hexagonal boron nitride (2D-hBN): Synthesis, properties and applications. *J. Mater. Chem. C* **5**, 11992–12022 (2017).

98. Lopez-Bezanilla, A., Huang, J., Terrones, H. & Sumpter, B. G. Boron nitride nanoribbons become metallic. *Nano Lett.* **11**, 3267–3273 (2011).

99. Zheng, F. *et al.* Half metallicity along the edge of zigzag boron nitride nanoribbons. *Phys. Rev. B - Condens. Matter Mater. Phys.* **78**, 1–5 (2008).

100. Zheng, F., Sasaki, K. I., Saito, R., Duan, W. & Gu, B. L. Edge states of zigzag boron nitride nanoribbons. *J. Phys. Soc. Japan* **78**, 1–6 (2009).

101. Nakajima, S. *et al.* Topological Thouless pumping of ultracold fermions. *Nat. Phys.* **12**, 296–300 (2016).

102. Ke, Y. *et al.* Topological phase transitions and Thouless pumping of light in photonic waveguide arrays. *Laser Photonics Rev.* **10**, 995–1001 (2016).

103. Atala, M. *et al.* Direct measurement of the Zak phase in topological Bloch bands. *Nat. Phys.* **9**, 795–800 (2013).

104. Song, W. *et al.* Breakup and Recovery of Topological Zero Modes in Finite Non-Hermitian Optical Lattices. *Phys. Rev. Lett.* **123**, 165701 (2019).

105. Zhen, B., Hsu, C. W., Lu, L., Stone, A. D. & Soljačić, M. Topological nature of optical bound states in the continuum. *Phys. Rev. Lett.* **113**, 1–5 (2014).

106. Hsu, C. W., Zhen, B., Stone, A. D., Joannopoulos, J. D. & Soljac, M. Bound states in the continuum. *Nat. Rev. Mater.* **1**, 16048 (2016).

107. Joannopoulos, J. D., Johnson, S. G., Winn, J. N. & Meade, R. D. *Molding the Flow of Light*. (Princeton University Press, 2008).

108. Hsu, C. W. *et al.* Observation of trapped light within the radiation continuum. *Nature* **499**, 188–191 (2013).

109. Kodigala, A. *et al.* Lasing action from photonic bound states in continuum. *Nature* **541**, 196–199 (2017).

110. Ha, S. T. *et al.* Directional lasing in resonant semiconductor nanoantenna arrays. *Nat. Nanotechnol.* **13**, 1042–1047 (2018).

111. Zhang, Y. *et al.* Observation of Polarization Vortices in Momentum Space. *Phys. Rev. Lett.* **120**, 186103 (2018).

112. Doeleman, H. M., Monticone, F., Den Hollander, W., Alù, A. & Koenderink, A. F. Experimental observation of a polarization vortex at an optical bound state in the continuum. *Nat. Photonics* **12**, 397–401 (2018).

113. Wang, X. *et al.* Tunable Bragg filters with a phase transition material defect layer. *Opt. Express* **24**, 20365 (2016).

114. Akahane, Y., Asano, T., Song, B. S. & Noda, S. High-Q photonic nanocavity in a two-

dimensional photonic crystal. *Nature* **425**, 944–947 (2003).

115. Poo, Y., Wu, R. X., Lin, Z., Yang, Y. & Chan, C. T. Experimental realization of self-guiding unidirectional electromagnetic edge states. *Phys. Rev. Lett.* **106**, 1–4 (2011).

116. Cheng, X. *et al.* Robust reconfigurable electromagnetic pathways within a photonic topological insulator. *Nat. Mater.* **15**, 542–548 (2016).

117. Rybin, M. V. *et al.* Phase diagram for the transition from photonic crystals to dielectric metamaterials. *Nat. Commun.* **6**, 6–11 (2015).

118. Oskooi, A. F. *et al.* Meep: A flexible free-software package for electromagnetic simulations by the FDTD method. *Comput. Phys. Commun.* **181**, 687–702 (2010).

119. Jin, J. *et al.* Topologically enabled ultrahigh-Q guided resonances robust to out-of-plane scattering. *Nature* **574**, 501–504 (2019).

120. Lee, J. *et al.* Observation and differentiation of unique high-Q optical resonances near zero wave vector in macroscopic photonic crystal slabs. *Phys. Rev. Lett.* **109**, 1–5 (2012).

121. Li, Q., Wang, T., Su, Y., Yan, M. & Qiu, M. Coupled mode theory analysis of mode-splitting in coupled cavity system. *Opt. Express* **18**, 8367 (2010).

122. Wang, Z., Fan, S. & Suh, W. Temporal coupled-mode theory and the presence of non-orthogonal modes in lossless multimode cavities. *IEEE J. Quantum Electron.* **40**, 1511–1518 (2004).

123. Fan, S. & Suh, W. Temporal coupled-mode theory for the Fano resonance in optical resonators. *J. Opt. Soc. Am. A* **20**, 569–572 (2003).

124. Liu, Y. C., Li, B. B. & Xiao, Y. F. Electromagnetically induced transparency in optical microcavities. *Nanophotonics* **6**, 789–811 (2017).

125. Manipatruni, S., Dong, P., Xu, Q. & Lipson, M. Tunable superluminal propagation on a silicon chip. *Opt. Lett.* **33**, 2928–2930 (2008).

126. Pan, J. *et al.* Tuning the coherent interaction in an on-chip photonic-crystal waveguide-resonator system. *Appl. Phys. Lett.* **97**, 1–4 (2010).

127. Zhou, L., Ye, T. & Chen, J. Coherent interference induced transparency in self-coupled optical waveguide-based resonators. *Opt. Lett.* **36**, 13 (2011).

128. Xiao, Y. F., Zou, X. B., Jiang, W., Chen, Y. L. & Guo, G. C. Analog to multiple electromagnetically induced transparency in all-optical drop-filter systems. *Phys. Rev. A - At. Mol. Opt. Phys.* **75**, 1–4 (2007).

129. Xu, Q. *et al.* Experimental realization of an on-chip all-optical analogue to electromagnetically induced transparency. *Phys. Rev. Lett.* **96**, 123901 (2006).

130. Mancinelli, M., Bettotti, P., Fedeli, J. M. & Pavesi, L. Reconfigurable optical routers based on Coupled Resonator Induced Transparency resonances. *Opt. Express* **20**, 23856 (2012).

131. Peng, B., Özdemir, Ş. K., Chen, W., Nori, F. & Yang, L. What is and what is not electromagnetically induced transparency in whispering-gallery microcavities. *Nat. Commun.* **5**, 1–9 (2014).

132. Wang, C., Jiang, X., Zhao, G., Zhang, M. & Wei, C. Electromagnetically induced transparency at a chiral exceptional point. *Nat. Phys.* 1–7 (2020) doi:10.1038/s41567-019-0746-7.

133. Timurdogan, E. *et al.* AIM process design kit (AIMPDKv2.0): Silicon photonics passive and active component libraries on a 300mm wafer. *2018 Opt. Fiber Commun. Conf. Expo. OFC 2018 - Proc.* 1–3 (2018).

134. Lumerical Inc. https://www.lumerical.com/products/.

135. Chen, L., Sherwood-Droz, N. & Lipson, M. Compact bandwidth tunable microring resonators. *Opt. InfoBase Conf. Pap.* **32**, 3361–3363 (2008).

Appendix A

```
#Meep code under Linux environment for calculation of 2D photonic crystal dispersion
# Define data output path
export PATH=/home/user/miniconda3/bin:$PATH
source activate mp
python

# Estimated calculation time using Core i5-7260U: 1 min.
import meep as mp
from meep import mpb
import math

# number of energy bands need to be calculated
num_bands=10
# The path of k in the first Brillouin zone. Here we choose the Δ path from Γ to X
k_points = [mp.Vector3(), mp.Vector3(0.5)]
# Number of k points need to be calculated
k_points = mp.interpolate(19, k_points)
# Define the low index cylinder
geometry = [mp.Cylinder(radius=0.2, material=mp.Medium(epsilon=2.1316)) ]
# Define the unit cell
geometry_lattice = mp.Lattice(size=mp.Vector3(1, 1))
# Define geometry resolution
resolution = 32
# Define matrix (background) material
default_material=mp.Medium(epsilon=4.08)
# Define ModeSolver using all the parameters defined above
ms = mpb.ModeSolver(num_bands=num_bands, k_points=k_points, geometry=geometry,
geometry_lattice=geometry_lattice, resolution=resolution, default_material=default_material)

ms.run_tm()      # Run simulation
```

Appendix B

```
# Lumerical FDTD code for TM1 band and quality factor calculation.
# This code is derived from the Lumerical Knowledge Base:
# https://apps.lumerical.com/diffractive_optics_pc_bandstructure.html
# This script only calculate one band of interest, must know the band structure first

########################################################
# preparation steps
########################################################
# simplify input variable names by removing spaces
number_resonances = %number resonances%;
make_plots = %make plots%;
n_monitors = %n monitors%;
resonance_trials = 4; # find three peaks fist then go to the lowest peak
min_filter_width = 1; # min width of filter in units of resonance FWHM
max_filter_width = 8; # min width of filter in units of resonance FWHM
filter_width_test_points = 20;
zero_pad = 2^17; # fft zero padding
# Note fft zero pad should be a power of 2,
# and larger gives more resolution in the
#frequency domain.

for(N=0; (N+1) <= n_monitors; 0)
{
N=N+1;
} # up to the sum of # of monitors

########################################################
# get the monitor data for the first monitor
########################################################
t = getdata("m3","t");
```

```
field0_t_Ex = pinch(getdata("m3","Ex"));
field0_t_Ey = pinch(getdata("m3","Ey"));
field0_t_Ez = pinch(getdata("m3","Ez"));

####################################################
# do fft to frequency domain for all monitors
####################################################
signal = 0;
w = fftw(t,1,zero_pad);
field_w = matrix(length(w),6*N);
for(i=1:N)
{
mname = "m" + num2str(i);
for(j=1:6)
{
if(almostequal(j,1))
{
component = "Ex";
}
if(almostequal(j,2))
{
component = "Ey";
}
if(almostequal(j,3))
{
component = "Ez";
}
if(almostequal(j,4))
{
component = "Hx";
}
if(almostequal(j,5))
```

```
{
component = "Hy";
}
if(almostequal(j,6))
{
component = "Hz";
}
if(j > 3.5)
{
extra_factor = sqrt(mu0/eps0);
}
else
{
extra_factor = 1;
}
if(havedata(mname,component))
{
signal = pinch(getdata(mname,component));
# signal = signal*exp( - 0.5*(t-max(t)*apod_center)^2/(apod_width*max(t))^2);
field_w(1:length(w),6*(i-1)+j) = 2*extra_factor*( (1:length(w)) <=(length(w)/2+0.1)) *
fft(signal,1,zero_pad);
}
}
}

#####################################################
# find resonant peaks, including all monitors, within the frequency of interest
#####################################################
f_spectrum = sum(abs(field_w)^2,2);
f1 = find(w, fmin_norm*3e8/period*2*pi);
f2 = find(w, fmax_norm*3e8/period*2*pi);
p = findpeaks(f_spectrum(f1:f2),number_resonances);
p = p + f1 - 1;
```

```
f0 = w(p)/2/pi;

# plot(w(f1:f2),f_spectrum(f1:f2));
# plot(w,f_spectrum);
?"lowest frequency =" + "position:" + num2str(f1) + "omega:" + num2str(w(f1));
?"highest freqeuncy =" + "position" + num2str(f2) + "omega:" + num2str(w(f2));
?"peak position = " + num2str(p) + " frequency: " + num2str(f0) + "Hz";

#####################################################
# find quality factors
#####################################################
# reserve memory for results
peak_spectra = matrix(length(w),number_resonances);
peak_filters2 = matrix(length(w),number_resonances);

# calculate slope of decay using 20-80% of time signal
tp1 = round(0.2*length(t));
tp2 = round(0.8*length(t));
t2 = t(tp1:tp2);
log_field_all = matrix(tp2-tp1+1,number_resonances);

Q = matrix(number_resonances);
delta_Q = matrix(number_resonances)+1e300;

# loop over each peak
for(i=1:number_resonances)
{
# find FWHM of peak
peak_val = f_spectrum(p(i));
continue_search = 1;
for(p1=p(i)-1; (p1>1) & continue_search ; 1)
{
```

```
if(f_spectrum(p1)<=peak_val/2)
{
continue_search = 0;
}
else
{
p1 = p1-1;
}
}

continue_search = 1;
for(p2=p(i)+1; (p2<length(w))& continue_search; 1)
{
if(f_spectrum(p2)<=peak_val/2)
{
continue_search = 0;
}
else
{
p2 = p2+1;
}
}

if(p1 < 1)
{
p1 = 1;
}
if(p2 > length(w))
{
p2 = length(w);
}
```

```
FWHM = w(p2)-w(p1);
for(filter_width=linspace(min_filter_width,max_filter_width,filter_width_test_points))
{
# calculate the filter for the peak
peak_filter = exp( -0.5*(w-w(p(i)))^2/(filter_width*FWHM)^2 );
# inverse fft to get data in time domain
field2_t = 0;
for(mcount=1:6*N)
{
field2_t = field2_t + abs(invfft(pinch(field_w,2,mcount)*peak_filter))^2;
}
field2_t = field2_t(tp1:tp2);
log_field = log10(abs(field2_t));

# calculate slope and Q from the slope of the decay
# estimate error from the slope
slope = (log_field(2:length(t2))-log_field(1:length(t2)-1))/(t(2:length(t2))-t(1:length(t2)-1));
slope_mean = sum(slope)/length(slope);
slope_delta = sqrt( sum((slope-slope_mean)^2)/length(slope) );
Q_test = -w(p(i))*log10(exp(1))/(slope_mean);
delta_Q_test = abs(slope_delta/slope_mean*Q_test);
if(delta_Q_test < delta_Q(i))
{
Q(i) = Q_test;
delta_Q(i) = delta_Q_test;

# collect data for final plot
peak_spectra(1:length(w),i) = f_spectrum * peak_filter^2;
peak_filters2(1:length(w),i) = peak_filter^2;
log_field_all(1:length(t2),i) = log_field;

}
```

```
}
# output summary of peak results to script window
?"Resonance " + num2str(i) + ":";
?" frequency = " + num2str(w(p(i))/(2*pi)*1e-12) + "THz, or "+num2str(2*pi*c/w(p(i))*1e9)+"
nm";
?" Q = " + num2str(Q(i)) +" +/- " + num2str(delta_Q(i));
}

Q_matrix=Q;



######################################################
# create dataset and plot the results
######################################################
spectrum = matrixdataset("spectrum");
spectrum.addparameter("f",w/2/pi);
spectrum.addattribute("spec",f_spectrum);

band_count=matrix(number_resonances);
band_count=[1:number_resonances];

bandstructure = matrixdataset("bandstructure");
bandstructure.addparameter("band_count", band_count);
bandstructure.addattribute("freq",f0); #resonance position in absolute frequency
bandstructure.addattribute("Q", Q_matrix);
bandstructure.addattribute("dQ", delta_Q);

if (make_plots)
{
# # plot signal and envelope for first monitor
field_t_Ex = invfft(pinch(field_w,2,1));
field_t_Ex = field_t_Ex(1:length(t));
field_t_Ey = invfft(pinch(field_w,2,2));
```

```
field_t_Ey = field_t_Ey(1:length(t));
field_t_Ez = invfft(pinch(field_w,2,3));
field_t_Ez = field_t_Ez(1:length(t));
plot(t*1e15,field0_t_Ex,abs(field_t_Ex),field0_t_Ey,abs(field_t_Ey),field0_t_Ez,abs(field_t_Ez),
"time (fs)","field envelope"); legend("field (Ex)","envelope (Ex)","field (Ey)","envelope
(Ey)","field (Ez)","envelope (Ez)");


# plot the slopes of the decaying fields
plot(t2*1e15,log_field_all,"time (fs)","log10(|field(t)|)","Decay for each resonance");


# plot spectra
p1 = find(w,0.8*min(w(p)));
p2 = find(w,1.2*max(w(p)));
f = w/(2*pi);
plot(f(p1:p2)*1e-12,peak_spectra(p1:p2,1:number_resonances)/max(f_spectrum),"frequency
(THz)","Arbitrary units","Spectrum of resonances");
plot(f(p1:p2)*1e12,f_spectrum(p1:p2)/max(f_spectrum),peak_filters2(p1:p2,1:number_resonanc
es),"frequency (THz)","Arbitrary units","Spectrum and filters");
}
```