Applied Physics Letters

Inverse design of photonic topological state via machine learning

Cite as: Appl. Phys. Lett. **114**, 181105 (2019); https://doi.org/10.1063/1.5094838
Submitted: 06 March 2019 . Accepted: 22 April 2019 . Published Online: 08 May 2019

Yang Long, Jie Ren, Yunhui Li, and Hong Chen

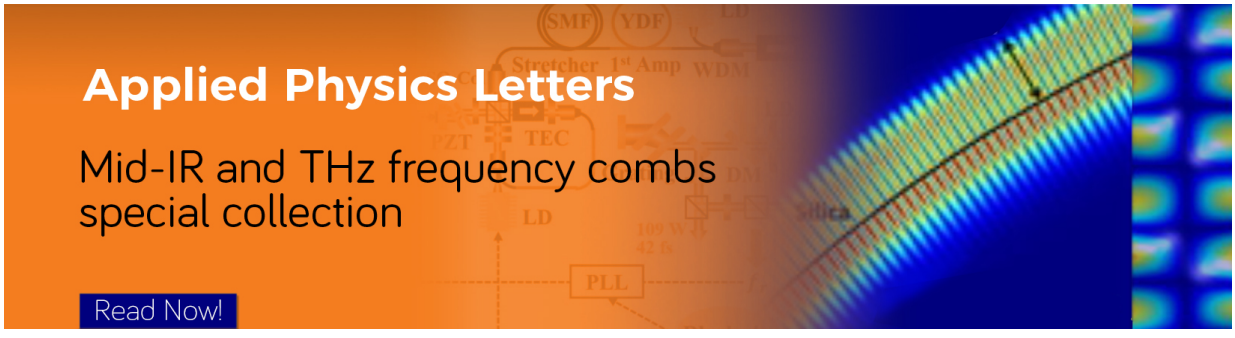
View Online Export Citation CrossMark

ARTICLES YOU MAY BE INTERESTED IN

Nonlinearity-based circulator
Applied Physics Letters **114**, 181102 (2019); https://doi.org/10.1063/1.5094736

Multiple non-diffracting beams by reflective surface based on admittance superposition
Applied Physics Letters **114**, 181104 (2019); https://doi.org/10.1063/1.5090318

All-angle Brewster effect observed on a terahertz metasurface
Applied Physics Letters **114**, 191902 (2019); https://doi.org/10.1063/1.5097742



© 2019 Author(s).

Applied Physics Letters ARTICLE scitation.org/journal/apl

Inverse design of photonic topological state via machine learning

Cite as: Appl. Phys. Lett. **114**, 181105 (2019); doi: 10.1063/1.5094838
Submitted: 6 March 2019 · Accepted: 22 April 2019 ·
Published Online: 8 May 2019

View Online Export Citation CrossMark

Yang Long,[1,a)] Jie Ren,[1,b)] Yunhui Li,[2] and Hong Chen[1,2]

AFFILIATIONS

[1]Center for Phononics and Thermal Energy Science, China-EU Joint Center for Nanophononics, Shanghai Key Laboratory of Special Artificial Microstructure Materials and Technology, School of Physics Sciences and Engineering, Tongji University, Shanghai 200092, China

[2]Key Laboratory of Advanced Micro-Structure Materials, MOE, School of Physics Science and Engineering, Tongji University, Shanghai 200092, China

[a)]**Electronic mail:** longyang_123@yeah.net
[b)]**Electronic mail:** xonics@tongji.edu.cn

ABSTRACT

The photonics topological state plays an important role in recent optical physics and has led to devices with robust properties. However, the design of optical structures with the target topological states is a challenge for current research. Here, we propose an approach to achieve this goal by exploiting machine learning technologies. In our work, we focus on Zak phases, which are the topological properties of one-dimensional photonics crystals. After learning the principle between the geometrical parameters and the Zak phases, the neural network can obtain the appropriate structures of photonics crystals by applying the objective Zak phase properties. Our work would give more insights into the application of machine learning on the inverse design of the complex material properties and could be extended to other fields, i.e., advanced phononics devices.

Published under license by AIP Publishing. https://doi.org/10.1063/1.5094838

Photonics topology physics unveils several unconventional optical properties in topological materials.[1–4] The topological properties in the photonics system relate to several optical phenomena: The one-dimensional (1D) topologically induced interface states,[5–9] the backscattering-immune wave transports,[10,11] and the pseudo-spin based wave splitter.[12] Recently, high-order topological physics has attracted a lot of attention due to its topological related quadrupole physics[13–16] with non-trivial bulk quadrupole moment, lower dimension topological states, and corner states being induced.[17] The development of topological optics would widen the optical application and enhance the robustness of device performance in the future. However, the photonics topological states are highly based on the complex optical response and geometrical structure of materials.[1–3,6,18] The design of the photonics device with the targeted topological state is still a challenge.

Recently, machine learning technologies have shown the power on the exploration of the expert game and the optimization for the complex strategy,[19–21] including the master of the Go game,[22,23] the detection of critical diseases,[24] and the automatic control of advanced robots.[25,26] Machine learning has been applied on the explorations of quantum states,[27–32] optimizations of metamaterial structures,[33,34] and inverse designs of optical nano-devices.[34–37] The 2016 Nobel prize has been awarded to scientists in the study of topological physics, and recently, the Turing award has been won by the pioneers in machine learning. It would be interesting to see what would happen after combining the Nobel prize and the Turing award. Here, we will achieve the design goal for the objective photonics topological states based on machine learning.

In this letter, we exploit the machine learning to design the photonics device with the target topological properties. In our work, we focus on the experimentally feasible one-dimensional (1D) dielectric photonics crystal (PC) system. Our core design approach includes designing a proper equivalent description about the photonics topological properties of the Zak phase,[5,38,39] training the network model for forward prediction, achieving an inverse design model based on the pre-trained tandem pipeline, and finally, one can obtain the optical structure by applying the target topological properties on the inverse network. These exploring approaches can be extended to other research fields, i.e., acoustic waves and elastic waves. Our work has shown that the machine learning technologies can be applied on the design of complex optical functions, such as topological properties, and can be extended to the design of phononics devices.

First, we will review the topological properties in 1D PC. The band structure of a dielectric binary PC shown in Fig. 1 can be described as follows:[40]

Published under license by AIP Publishing

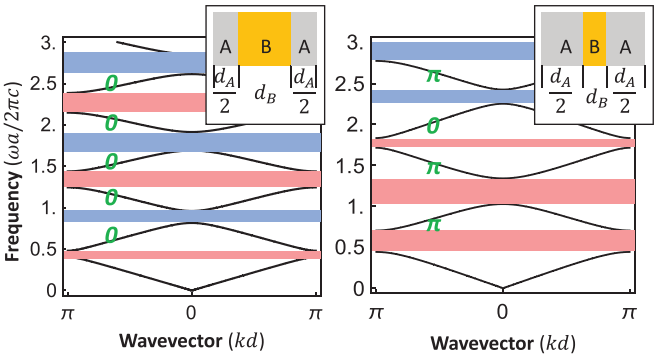


FIG. 1. Geometrical effects on the topological properties of the 1D photonics crystal with inversion symmetry. Two examples have been demonstrated to show that the relationship between the length of each layer in the unit cell and the topological properties of the band is complicated. Here, $n_A = 2.5$ and $n_B = 1$ for both cases. For the lengths, $d_A = 0.4a$ and $d_B = 0.1a$ for left, $d_A = 0.1a$ and $d_B = 0.6a$ for right, $d = d_A + d_B$ is the lattice constant, and a is an arbitrary length constant. The Zak phase of each individual band is labeled in green. The blue strip represents the gap with the negative reflection phase $\phi < 0$, while the red strip represents the gap with the positive reflection phase $\phi > 0$.

$$\cos(kd) = \cos(n_A k_0 d_A)\cos(n_B k_0 d_B) - \frac{1}{2}\left(\frac{n_A}{n_B} + \frac{n_B}{n_A}\right)\sin(n_A k_0 d_A)\sin(n_B k_0 d_B), \quad (1)$$

where $n_{A/B}$ are the refractive indexes, $d = d_A + d_B$ is the lattice constant, and $k_0 = \frac{\omega}{c}$ is the wave vector of light in vacuum. We have calculated the band dispersions of two structures with the same media but different geometrical parameters, shown in Fig. 1. It is clear that several features, i.e.,, the center frequency and size of the gap, are highly dependent on the length of each layer in the unit cell.

Importantly, the topological properties of 1D PC are also affected by the geometrical settings. The Zak phase of each band is exploited to express its topological properties of 1D PC[5–7]

$$\theta_n = \int_{-\pi/a}^{\pi/a} \left(i \int_{cell} \varepsilon(x) u_{n,k}^*(x) \partial_k u_{n,k}(x) \right) dk, \quad (2)$$

where $u_{n,k}(x)$ is the eigen-state of the n-th band, $\varepsilon(x) = n^2(x)$. In particular, the Zak phase of the lowest 0th band is determined by the sign of $(1 - n_A/n_B) : \exp(i\phi_0) = \text{sign}(1 - n_A/n_B)$.[5,7] We calculated the Zak phase for the cases shown in Fig. 1. From the results, we can see that the relationship between the Zak phase and the geometrical settings is complicated and cannot be described by some simple mathematical formulas. Thus, finding the proper geometrical parameters to realize the target topological properties in the frequency region of interest is hard and complex. In the following, we will show you the approach to design the photonics topological states by the machine learning technologies. Moreover, these physical considerations can also be applied on the acoustic and phononics research.

In our work, we will focus on two dielectric media: SiO_2 and Si, which are mostly usable optical materials in the on-chip integrated optics. We set $n_{SiO_2} = 1.5$ and $n_{Si} = 3.5$ for simplicity. We focus on the geometrical properties of 1D PC with inversion symmetry,[5] and we introduce the state vector $\boldsymbol{d}$, which is composed of the length of each layer in the unit cell, $\boldsymbol{d} = (d_1, d_2, \ldots, d_M)^T$, where $M = 4$ is the number of parameters such that the unit cell is symmetric with the configuration (d1/2, d2/2,…, dM,…, d2/2, d1/2), as shown in Fig. 2(a). To characterize the topological properties, we use the reflection phase features of gaps instead of using the Zak phase of bands because the topological properties of each band would be associated with the reflection phase of gaps[5,7]

$$\text{sign}(\phi_{n-1})/\text{sign}(\phi_n) = -e^{i\theta_n}, \quad \theta_n = 0, \pi, \quad (3)$$

where ϕ_n is the reflection phase of the n-th gap, $\phi_n \in (-\pi, \pi)$, $\text{sign}(\phi_n) = -1$ for $\phi_n \in (-\pi, 0)$; otherwise, $\text{sign}(\phi_n) = 1$. We divide the frequency region of interest $[\omega_{min}, \omega_{max}]$ into N parts ω_j, $j = 1 \ldots N$. Here, we set $N = 100$. We introduce the label vector $\boldsymbol{\beta}$, which is the $N \times 1$ vector. It is composed of the reflection phase properties for the each frequency point f_j: $\beta_j = \text{sign}(\phi(\omega_j))$ for the gap; otherwise, $\beta_j = 0$ for the band, shown in Fig. 2(b). Obviously, the

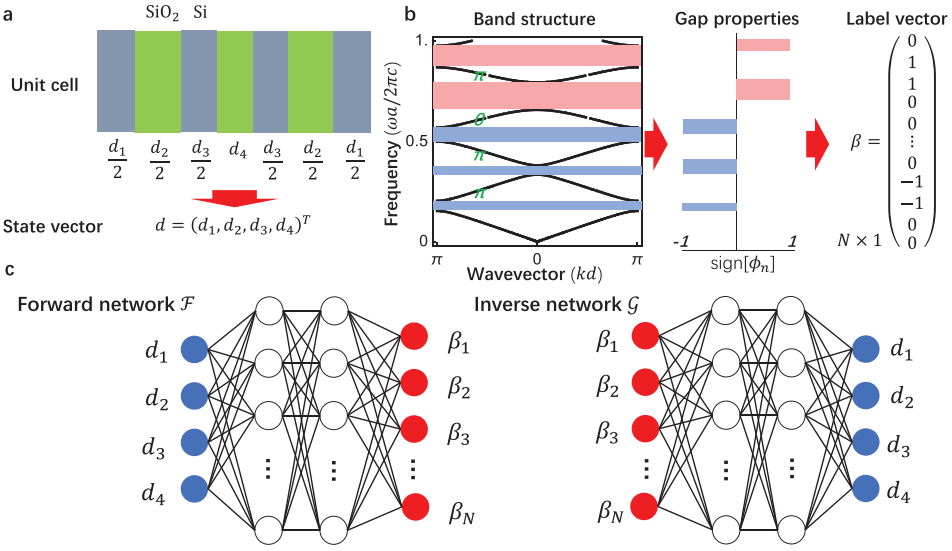


FIG. 2. The schematic representation of the state vector, the label vector, and the neural networks. (a) The state vector composed of the length of each layer in the unit cell. (b) The label vector obtained by the sign of the reflection phases of gaps. The signs of the reflection phases are represented by the colorful strip: Red for $\phi > 0$ and blue for $\phi < 0$. (c) Two neural networks: The forward network $\mathcal{F}$ that predicts the label vector $\boldsymbol{\beta}$ based on the input state vector $\boldsymbol{d}$; the inverse network $\mathcal{G}$ that predicts the state vector $\boldsymbol{d}$ based on the input label vector $\boldsymbol{\beta}$.

Published under license by AIP Publishing

definition of the label vector will not only contain the band structure properties but also include its topological properties.

With the state vector $\boldsymbol{d}$ and label vector $\boldsymbol{\beta}$, we will construct two network models to achieve the forward prediction and inverse design, shown in Fig. 2(c)

$$\boldsymbol{\beta} = \mathscr{F}(\boldsymbol{d}), \quad \boldsymbol{d} = \mathscr{G}(\boldsymbol{\beta}), \tag{4}$$

where $\mathscr{F}$ is the forward prediction network and $\mathscr{G}$ is the inverse design network. The $\mathscr{F}$ network has 5 layers with hidden units of 512 for each layer. The cost function has been set as $E_{\mathscr{F}} = \frac{1}{L}\sum_i |\boldsymbol{\beta}_i - \hat{\boldsymbol{\beta}}_i|^2$, where $\hat{\boldsymbol{\beta}}_i$ is the predicted results and L is the number of training samples in each batch. We train the forward network with the 5×10^6 sample dataset (90% for the train set and 10% for the test set). The optimizer is the Adam optimizer with a learning rate of 0.001 and a decay of 10^{-6}. It would converge quickly (nearly after 3000 epochs) for training $\mathscr{F}$ because one optical structure will definitely correspond to only one form of band dispersion: One-to-one. However, a similar approach for training the inverse network $\mathscr{G}$ will make the convergence slow or diverge at all because the one form of band dispersion will not correspond to only one case of the optical structure: One-to-many question. The presence of such an inconsistency is difficult to detect and cannot be eliminated by the filtering method.[35]

Here, we exploit the training pipeline introduced in Ref. 35 to overcome these issues. This training pipeline will construct a tandem network, which is composed of an inverse network followed by a pre-trained forward network, shown in Fig. 3. It means that without training $\mathscr{G}$ directly with the dataset $\{\boldsymbol{\beta}, \boldsymbol{d}\}$, we replace the training dataset with $\{\boldsymbol{\beta}, \mathscr{F}(\boldsymbol{d})\}$ for the combined network. This pipeline model resembles an autoencoder,[41,42] in which $\mathscr{G}$ can encode the label vector into the state vector, and then, this state vector will be decoded into the original label vector by $\mathscr{F}$. Analogy to the generative models in the natural language processing,[43] $\mathscr{F}$ acts as the effective discriminative model, which maps the words or phrases into the word vectors in semantic space, and $\mathscr{G}$ will behave as the effective generative model that generates the sentences with the target meanings as close as possible. But different from the conventional autocoders, the state vectors predicted by $\mathscr{G}$ have their concrete physical meanings: The length parameters in the unit cell. A similar autocoder scheme based on the dimensionality reduction is applied to deal with the non-uniqueness in the design of metasurfaces.[44]

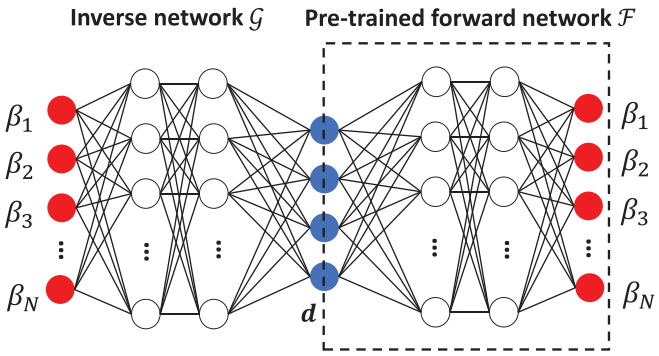


FIG. 3. The schematic representation of training the inverse network $\mathscr{G}$. A tandem network is composed of an inverse network $\mathscr{G}$ connected to a forward network $\mathscr{F}$. The forward network $\mathscr{F}$ is trained in advance. During training this pipeline, only the weights in the $\mathscr{G}$ are updated to reduce the cost while the weights in the pre-trained $\mathscr{F}$ are fixed.

The training process of the tandem network in Fig. 3 can be regarded as finding $\mathscr{G}$ to minimize the loss between the label vectors and the predicted ones

$$\min_{\mathscr{G}} \frac{1}{L}\sum_i |\mathscr{F}(\mathscr{G}(\boldsymbol{\beta}_i)) - \boldsymbol{\beta}_i|^2. \tag{5}$$

Only the units in $\mathscr{G}$ will be updated during the training. This training pipeline has shown its validity on dealing with the inconsistent issue in the training data.[35] The $\mathscr{G}$ network has 6 layers with hidden units of 512 for each layer. The training of the combined network is done by minimizing the cost function of the tandem network defined as $E_{\mathscr{G}} = \frac{1}{L}\sum_i |\boldsymbol{\beta}_i - \hat{\boldsymbol{\beta}}_i|^2$, where $\hat{\boldsymbol{\beta}}_i$ is the predicted results from the combined network. The Adam optimizer with a learning rate of 0.001 and a decay of 10^{-6} is exploited. The combined network will converge after nearly 5000 epochs.

Here, we show some examples of designing the 1D PC composed of SiO_2 and Si with the target topological properties in the frequency region of interest. Here, we set the frequency region as $\omega \in [0, 0.4]\frac{2\pi c}{a}$. The unit cell of 1D PC has 7 layers with inversion symmetry and thus with the state vector $\boldsymbol{d} = (d_1, d_2, d_3, d_4)^T$, as shown in Fig. 2(a). The label vector $\boldsymbol{\beta}$ will contain the $N = 100$ elements mentioned before. After training the forward network and then obtaining the inverse network by the approach shown in Fig. 3, two cases are demonstrated in Fig. 4. We can set the signs of the label vectors (reflection phases) of gaps on the both sides of the target band to control its Zak phase:[5,7] (i) $\theta_n = \pi$ for the same signs, $\text{sign}(\phi_{n-1})/\text{sign}(\phi_n) = 1$; and (ii) $\theta_n = 0$ for the opposite signs, $\text{sign}(\phi_{n-1})/\text{sign}(\phi_n) = -1$. In Figs. 4(a) and 4(d), the target label vectors are represented by the blue point, and the label vectors of the predicted optical structures by the inverse network are shown as the red point. From these results, we can see that the inverse network is able to predict the proper optical structures to satisfy the target label vectors. The first case in Fig. 4(a) realizes $\theta_1 = \theta_2 = 0$. The second case in Fig. 4(d) will have $\theta_1 = \theta_2 = \pi$. The Zak phase properties of these optical structures predicted by the inverse network are verified by the reflection spectrum and reflection phases using the transfer matrix methods, shown in Figs. 4(b) and 4(e), respectively. Based on these results, we can obtain the band dispersions and the topological properties of these optical structures in Figs. 4(c) and 4(f). The Zak phases of bands are labeled in green color.

To summarize, we demonstrate a machine learning based method to design the optical structure with the target topological properties. In our work, we focus on the Zak phase of each band, which is the topological invariant of the 1D photonic crystal. We introduce the state vectors that contain the geometrical information and the label vectors that reflect the reflection phase properties associated with the Zak phase. With these two descriptors, we train the forward network in advance and the inverse network by comprising a tandem pipeline to overcome the non-uniqueness issue. Several inverse design cases are shown in our paper. Our work would supply some ideas for designing optical topological materials and can be extended to higher dimensional photonic crystals. It could also be exploited in exploring the topological properties of other research fields, i.e., topological phononics.

Published under license by AIP Publishing

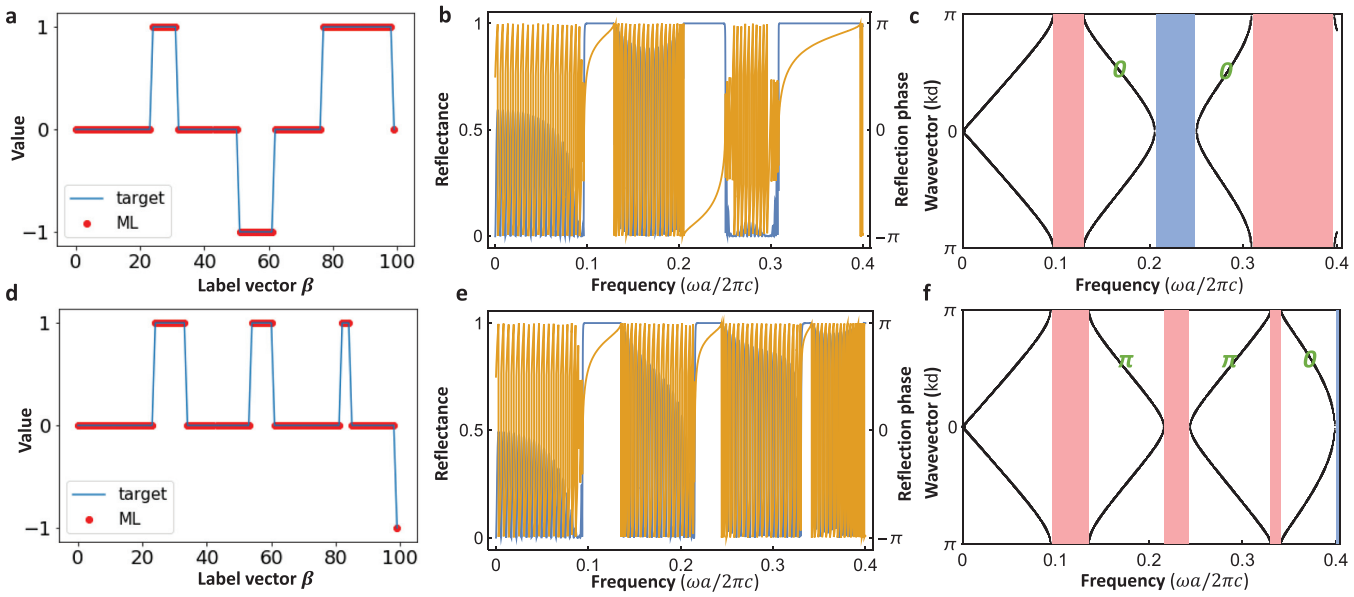


FIG. 4. Examples designed by the inverse network $\mathscr{G}$. Two cases have been demonstrated in (a) and (d). The target label vectors and the label vectors of optical structures obtained by $\mathscr{G}$ are represented in the blue and red colors, respectively. The length of each layer in the unit cell is (a) $\boldsymbol{d} = (0.375, 0.147, 0.510, 0.595)a$; (d) $\boldsymbol{d} = (0.509, 0.715, 0.148, 0.519)a$, where a is a length constant. The corresponding reflection spectrum (Blue line) and reflection phases (Orange line) are calculated numerically in (b) and (e). The band dispersion (black line) and the Zak phases (labeled in green color) of these optical structures are shown in (c) and (f). The signs of the reflection phases are represented by the colorful strip: Red for $\phi > 0$ and blue for $\phi < 0$.

We acknowledge the support from the National Natural Science Foundation of China (Nos. 11775159 and 61621001), the National Key Research Program of China (No. 2016YFA0301101), the Shanghai Science and Technology Committee (Nos. 18ZR1442800 and 18JC1410900), the Opening Project of Shanghai Key Laboratory of Special Artificial Microstructure Materials and Technology, the Fundamental Research Funds for the Central Universities, and the National Youth 1000 Talents Program in China.

REFERENCES

[1]L. Lu, J. D. Joannopoulos, and M. Soljačić, Nat. Photonics **8**, 821 (2014).
[2]L. Lu, J. D. Joannopoulos, and M. Soljačić, Nat. Phys. **12**, 626 (2016).
[3]B.-Y. Xie, H.-F. Wang, X.-Y. Zhu, M.-H. Lu, Z. Wang, and Y.-F. Chen, Opt. Express **26**, 24531 (2018).
[4]M. Z. Hasan and C. L. Kane, Rev. Mod. Phys. **82**, 3045 (2010).
[5]M. Xiao, Z. Zhang, and C. T. Chan, Phys. Rev. X **4**, 021017 (2014).
[6]W. Zhu, Y.-Q. Ding, J. Ren, Y. Sun, Y. Li, H. Jiang, and H. Chen, Phys. Rev. B **97**, 195307 (2018).
[7]W. S. Gao, M. Xiao, C. Chan, and W. Y. Tam, Opt. Lett. **40**, 5259 (2015).
[8]F. Cardano, A. D'Errico, A. Dauphin, M. Maffei, B. Piccirillo, C. de Lisio, G. De Filippis, V. Cataudella, E. Santamato, L. Marrucci *et al.*, Nat. Commun. **8**, 15516 (2017).
[9]M. Atala, M. Aidelsburger, J. T. Barreiro, D. Abanin, T. Kitagawa, E. Demler, and I. Bloch, Nat. Phys. **9**, 795 (2013).
[10]A. B. Khanikaev, S. H. Mousavi, W.-K. Tse, M. Kargarian, A. H. MacDonald, and G. Shvets, Nat. Matter. **12**, 233 (2013).
[11]M. C. Rechtsman, J. M. Zeuner, Y. Plotnik, Y. Lumer, D. Podolsky, F. Dreisow, S. Nolte, M. Segev, and A. Szameit, Nature **496**, 196 (2013).
[12]X. Cheng, C. Jouvaud, X. Ni, S. H. Mousavi, A. Z. Genack, and A. B. Khanikaev, Nat. Matter. **15**, 542 (2016).
[13]M. Serra-Garcia, V. Peri, R. Süsstrunk, O. R. Bilal, T. Larsen, L. G. Villanueva, and S. D. Huber, Nature **555**, 342 (2018).
[14]C. W. Peterson, W. A. Benalcazar, T. L. Hughes, and G. Bahl, Nature **555**, 346 (2018).
[15]S. Imhof, C. Berger, F. Bayer, J. Brehm, L. W. Molenkamp, T. Kiessling, F. Schindler, C. H. Lee, M. Greiter, T. Neupert *et al.*, Nat. Phys. **14**, 925 (2018).
[16]B.-Y. Xie, H.-F. Wang, H.-X. Wang, X.-Y. Zhu, J.-H. Jiang, M.-H. Lu, and Y.-F. Chen, Phys. Rev. B **98**, 205147 (2018).
[17]W. A. Benalcazar, B. A. Bernevig, and T. L. Hughes, Science **357**, 61 (2017).
[18]Z. Guo, H. Jiang, Y. Sun, Y. Li, and H. Chen, Opt. Lett. **43**, 5142 (2018).
[19]Y. LeCun, Y. Bengio, and G. Hinton, Nature **521**, 436 (2015).
[20]Z. Ghahramani, Nature **521**, 452 (2015).
[21]M. I. Jordan and T. M. Mitchell, Science **349**, 255 (2015).
[22]D. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. Van Den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot *et al.*, Nature **529**, 484 (2016).
[23]D. Silver, J. Schrittwieser, K. Simonyan, I. Antonoglou, A. Huang, A. Guez, T. Hubert, L. Baker, M. Lai, A. Bolton *et al.*, Nature **550**, 354 (2017).
[24]H. Liang, B. Y. Tsui, H. Ni, C. C. Valentim, S. L. Baxter, G. Liu, W. Cai, D. S. Kermany, X. Sun, J. Chen *et al.*, Nat. Med. **25**(3), 433–438 (2019).
[25]Y. Duan, X. Chen, R. Houthooft, J. Schulman, and P. Abbeel, in *Proceedings of the 33rd International Conference on Machine Learning (ICML), 2016,* pp. 1329–1338.
[26]K. Noda, H. Arie, Y. Suga, and T. Ogata, Rob. Autom. Syst. **62**, 721 (2014).
[27]G. Carleo and M. Troyer, Science **355**, 602 (2017).
[28]J. Carrasquilla and R. G. Melko, Nat. Phys. **13**, 431 (2017).
[29]E. P. Van Nieuwenburg, Y.-H. Liu, and S. D. Huber, Nat. Phys. **13**, 435 (2017).
[30]P. Zhang, H. Shen, and H. Zhai, Phys. Rev. Lett. **120**, 066401 (2018).
[31]J. Venderley, V. Khemani, and E.-A. Kim, Phys. Rev. Lett. **120**, 257204 (2018).
[32]Y. Zhang and E.-A. Kim, Phys. Rev. Lett. **118**, 216401 (2017).
[33]W. Ma, F. Cheng, and Y. Liu, ACS Nano **12**, 6326 (2018).
[34]Z. Liu, D. Zhu, S. P. Rodrigues, K.-T. Lee, and W. Cai, Nano Lett. **18**, 6570 (2018).
[35]D. Liu, Y. Tan, E. Khoram, and Z. Yu, ACS Photonics **5**, 1365 (2018).
[36]L. Pilozzi, F. A. Farrelly, G. Marcucci, and C. Conti, Commun. Phys. **1**, 57 (2018).

Published under license by AIP Publishing

[37]J. Peurifoy, Y. Shen, L. Jing, Y. Yang, F. Cano-Renteria, B. G. DeLacy, J. D. Joannopoulos, M. Tegmark, and M. Soljačić, Sci. Adv. **4**, eaar4206 (2018).
[38]A. Bansil, H. Lin, and T. Das, Rev. Mod. Phys. **88**, 021004 (2016).
[39]J. Zak, Phys. Rev. Lett **62**, 2747 (1989).
[40]A. Yariv and P. Yeh, *Optical Waves in Crystals* (Wiley New York, 1984), Vol. 5.
[41]G. E. Hinton and R. R. Salakhutdinov, Science **313**, 504 (2006).
[42]C.-Y. Liou, W.-C. Cheng, J.-W. Liou, and D.-R. Liou, Neurocomputing **139**, 84 (2014).
[43]W. Lu, H. T. Ng, W. S. Lee, and L. S. Zettlemoyer, in *Proceedings of the Conference on Empirical Methods in Natural Language Processing (Association for Computational Linguistics*, 2008), pp. 783–792.
[44]Y. Kiarashinejad, S. Abdollahramezani, and A. Adibi, preprint arXiv:1902.03865 (2019).

Published under license by AIP Publishing