Supplementary Material for
Deep learning empowered synthetic dimension dynamics: morphing of light into topological modes

Shiqi Xia[1+], Sihong Lei[1+], Daohong Song[1,2], Luigi Di Lauro[3], Imtiaz Alamgir[3], Liqin Tang[1,2],

Jingjun Xu[1], Roberto Morandotti[3], Hrvoje Buljan[1,4*], Zhigang Chen[1,2*]

1 The MOE Key Laboratory of Weak-Light Nonlinear Photonics, TEDA Institute of Applied Physics and School of Physics, Nankai University, Tianjin 300457, China
2 Collaborative Innovation Center of Extreme Optics, Shanxi University, Taiyuan, Shanxi 030006, China
3 INRS-EMT, 1650 Blvd. Lionel-Boulet, Varennes, Quebec J3X 1S2, Canada
4 Department of Physics, Faculty of Science, University of Zagreb, Bijenička c. 32, Zagreb 10000, Croatia
[+]These authors contributed equally to this work
* hbuljan.phy@pmf.hr, zgchen@nankai.edu.cn

Note 1: ANN architecture for the inversed diagonalization problem: $E \rightarrow H$

The neural network is constructed with connections of nodes, which are elements that correspond to the neurons of the brain. There are three different types of layers in a given ANN, i.e., the input layer, the hidden layers, and the output layer [34]. Each layer has nodes and connection weights (Fig. S1a). The input data are multiplied by the weight before they reach the nodes of the next layer, which is called the weighted sum and is calculated as follows:

$$h_j^s = \sum_{k=1}^{M} w_{jk}\xi_k + b_j$$

where h_j^s is the j-th parameter is s-th hidden layer, and $s = 1$ denotes the first hidden layer. ξ_k is the input data, w_{jk} is the connection weight between the k-th data in the previous layer ξ_k and the j-th data in current layer h_j^s, b_j is the bias acting as an additional factor associated with the storage of information. M is the total number of nodes in the hidden layer. Finally, the j-th node takes the weighted sum h_j^s into the activation function in the node and yields its output:

$$\mathcal{H}_j^s = f\left(h_j^s\right)$$

Many activation functions can be chosen in each layer that determine the behavior of nodes. The data generated via the procedure stated above are the input of the next layer. Based on the preassigned number of hidden layers and activation functions, properly chosen weights can construct ANNs for

the solution of the $E \rightarrow H$ problem. More precisely, here we address the inverse problem of matrix diagonalization: for a given the spectrum E (energies or propagation constants here denoted with β), what the matrix H (Hamiltonian) should be to possess such a spectrum, subject to constraints; the only constraint used here is for the Hamiltonian to have nearest neighbor coupling, or in other words, we demand the matrix to have only 3-diagonal terms.

The process to adjust weights on every node is the training of ANNs. The backpropagation algorithm provides a systematic method to determine the error and further adjust the weights of the hidden nodes. Back-propagation means that the data move backward from the output layer to the input layer. We first prepare a set of input–output data as the training data. The evaluation of errors in output data then feedbacks the adjustment of weights along propagation [34].
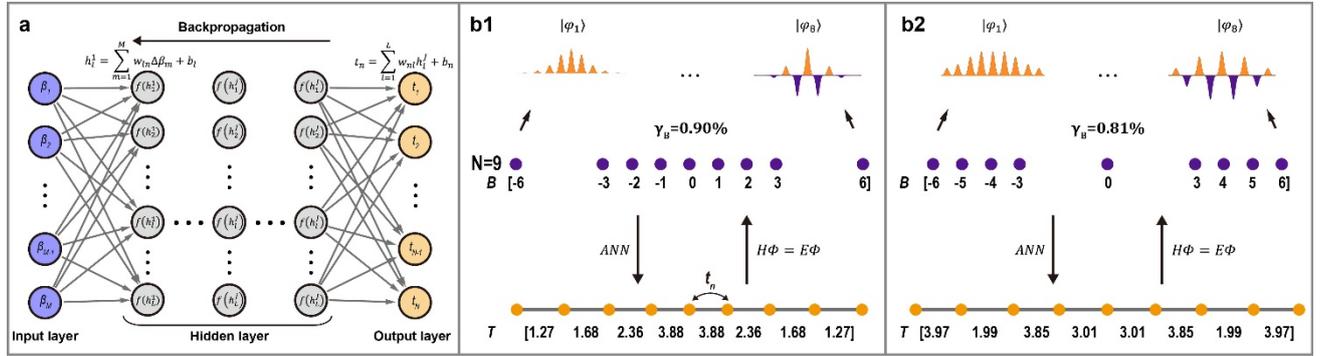

Fig. S1 An ANN architecture for reversed calculation of $E \rightarrow H$. (a) The ANN architecture for coupling array calculations composed of an input layer (B array), a number of hidden layers and an output layer (coupling array). (b) Demonstration of mode array calculation via ANNs. (b1) Top panel is the sketch of eigenvalue distribution B and corresponding eigenmodes $|\varphi_i\rangle$. The bottom panel shows couplings between sites calculated by ANNs. T is the coupling array with total lattice sites $N = 9$. (b2) has the same layout as (b1) but with different preassigned eigenvalue differences. γ_B is the figure of merit to test the effectiveness of our ANN method as defined in the main text.

To solve a specific inverse problem of $E \rightarrow H$, we write down the following procedure to establish and evaluate ANNs:

1) We first generate the input–output data via a tight-binding model and diagonalization of the corresponding Hamiltonian, $\Phi^\dagger H \Phi = E$. To simplify and obtain similar eigenmode distributions, eigenvalue (B) arrays and coupling (T) arrays are set as input and output layers, where T satisfies the inversion symmetry. We generate 800 sets of input–output data for training.

2) The built-in program in MATLAB is used to establish and train ANNs. We have 4 hidden layers, and the nodes on layers are set to be [16 16 16 10]. Since the lattices have $N = 8$ sites in the main text, the numbers of input (B) and output data (T) are 8 and 7, respectively.

3) After training, we send a preassigned spectrum $B = [\beta_1, ..., \beta_8]$ to the ANNs and obtain the couplings array $T = [t_1, ..., t_7]$ as the output. The tight-binding Hamiltonian can be formulated based on the output data. γ_B defined in the main text can evaluate the effectiveness of the $E \rightarrow H$ deep learning process (noted in Fig. S1 and Fig. S2).

Moreover, systems with different total lattice sites N can also be considered. Here, we briefly show the output data and γ_B when $N = 9$ (Fig. S1b) and $N = 10$ (Fig. S2).
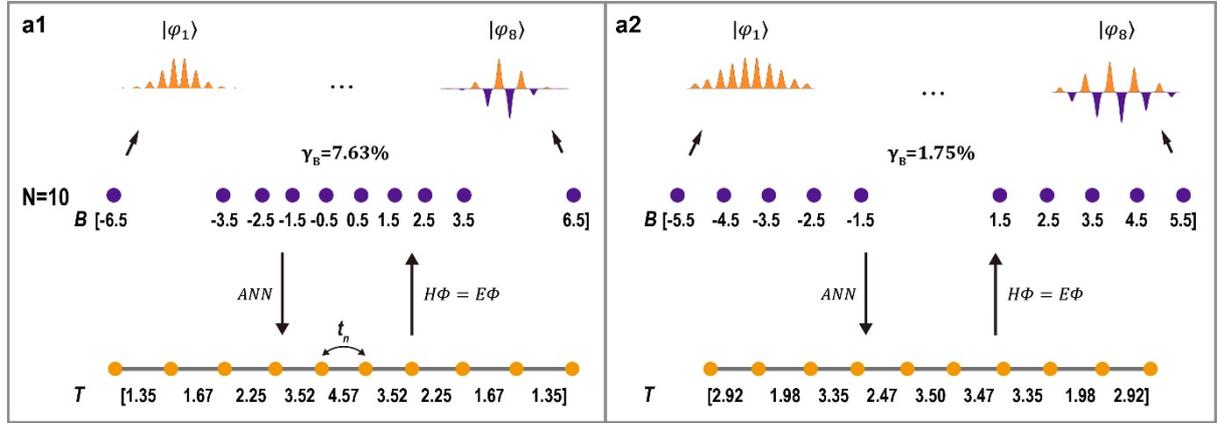

Fig. S2. Demonstration of difference in two preassigned mode arrays calculated via ANNs for $N = 10$. Figure S2 has the same layout as Fig. S1b.

Note 2: Construction and analysis of synthetic mode dimensions

Here we follow the derivation describing wiggling waveguides outlined in Ref [20], which is done in the vector potential gauge to describe the wiggling of the waveguides (in the main text we use the scalar potential gauge). The light traveling through the photonic lattice can be generally described by the paraxial Schrödinger equation:

$$i\frac{\partial}{\partial z}\psi(x,y,z) = -\frac{1}{2k_0}\nabla^2\psi(x,y,z) - \frac{k_0\Delta n(x,y,z)}{n_0}\psi(x,y,z) \qquad (S1)$$

where $\psi(x,y,z)$ is the complex amplitude of the linearly polarized electric field of the probe beam and z is the propagation direction, $k_0 = 2\pi n_0/\lambda$ is the wavenumber in the medium, and λ and n_0 are the wavelength of the probe beam in vacuum and the unperturbed refractive index of the crystal, respectively. Here $\Delta n(x,y,z)$ denotes the z-modulated waveguide arrays caused by refractive index

change in the crystal. The z-direction wiggling of waveguide arrays is embodied in

$$V(x,y,z) = -\frac{k_0\Delta n(x,y,z)}{n_0} = -\frac{k_0}{n_0}\sum_{n=1}^{N}\Delta n_{site}(x - D_n - Rsin(\Omega z + \theta), y, z),$$

where $\Delta n_{site}(x)$ is the profile of a single waveguide, $D_n = \sum_{i=1}^{n-1} d_i$; R, Ω and θ are the amplitude, wiggling frequency, and the initial phase of the oscillation of waveguides, respectively. Then, Eq. S1 can be written as

$$i\frac{\partial}{\partial z}\psi(x,y,z) = H\psi(x,y,z), \qquad H = -\frac{1}{2k_0}\nabla^2 + V(x,y,z). \tag{S2}$$

With the following substitution, $x' = x - Rsin(\Omega z + \theta), y' = y, z' = z$, we can obtain

$$i\frac{\partial}{\partial z}\chi = -\frac{1}{2k_0}(\nabla + i\mathbf{A})^2\chi + \frac{k_0\Delta n}{n_0}\chi \tag{S3}$$

where $\chi = \psi'\exp\left(-i\int \frac{1}{2}k_0^2R^2\Omega^2\cos^2(\Omega z + \theta)\,dz\right)$, $\mathbf{A} = -k_0R\Omega\cos(\Omega z + \theta)\hat{x}$. Under the tight-binding approximation and a Peierls substitution, Eq. S3 leads to the discrete Hamiltonian

$$H_W^A(z) = \sum_{n=1}^{N-1} t_n \exp(i\phi_n)\, c_{n+1}^{\dagger}c_n + H.c., \tag{S4}$$

where $\phi_n = \int_0^{d_n} \mathbf{A}\,dx = -d_nk_0R\Omega\cos(\Omega z + \theta)$, c_n is the annihilation operator on the nth site, and t_n is the coupling coefficient illustrated in main text Fig. 2. By using a unitary transformation $c_n = a_n\exp(-iD_nk_0R\Omega\cos(\Omega z + \theta))$, the wiggling Hamiltonian in the vector potential gauge $H_W^A(z)$ (Eq. S4) can be separated into two parts $H_0 + H_1(z)$:

$$H_0 = \sum_{n=1}^{N-1} t_n a_{n+1}^{\dagger}a_n + H.c., \qquad H_1(z) = \sum_{n=1}^{N} D_nk_0\Omega^2 Rsin(\Omega z + \theta)\, a_n^{\dagger}a_n + H.c. \tag{S5}$$

which is exactly the wiggling Hamiltonian in the scalar potential gauge used in the main text (Eq. 3). In the main text, we have for convenience centered the zero point of the scalar potential term $H_1(z)$ in the central point of the lattice $D_N/2$. It is evident that the dynamics is equivalent in the vector potential gauge and in the scalar potential gauge, as expected.

Now we address the lattice with equal spacing of eigenvalues in synthetic space. If t_n in Eq. S5 or Eq. 4 of the main text satisfies the coupling distribution of the J_x matrix, with $t_{n+j-1} = p_n = \frac{p}{2}\sqrt{j-n}\sqrt{j+n+1}$, p is a constant value, $j = (N+1)/2$, $n = -j, -j+1, \ldots, j-1, j$, then, we can have an analytic expression of $\Phi_A^{\dagger}(H_0 + H_1(z))\Phi_A$:

$$\mathcal{H} = \sum_{m=1}^{N-1} \left((p_m/p) d_m R\Omega^2 \cos(\Omega z + \theta) \, b_{m+1}^\dagger b_m + H.c. \right) + \sum_{m=1}^{N} pm b_m^\dagger b_m \qquad (S6)$$

where $b_m^\dagger (b_m)$ is the creation (annihilation) operator on the m-th mode in synthetic space [20].

Moreover, the diagonal term $\sum_{m=1}^{N} pn b_m^\dagger b_m$ (Fig. S3a1) can be removed by $u_m = b_m \exp(i\Omega mz), \Omega = p$ substitution. With the fast-rotating wave approximation,

$$\mathcal{H}' = \sum_{m=1}^{N-1} \left((p_m/2p) d_m R\Omega^2 e^{i\theta} u_{m+1}^\dagger u_m + H.c. \right) \qquad (S7)$$

Therefore, the J_x lattice can be constructed in synthetic space where modes only have couplings between nearest neighbors (Fig. S3a2). For mode arrays with unequal spacing, even though, we do not find an analytical expression equivalent to Eq. S6, the diagonal term can still be partially removed by $u_m = b_m \exp(i\Omega mz), \Omega = p$ substitution (Fig. S3b). Such array can realize the mode confinement due to the coupling blockade shown in Fig. S3b3.
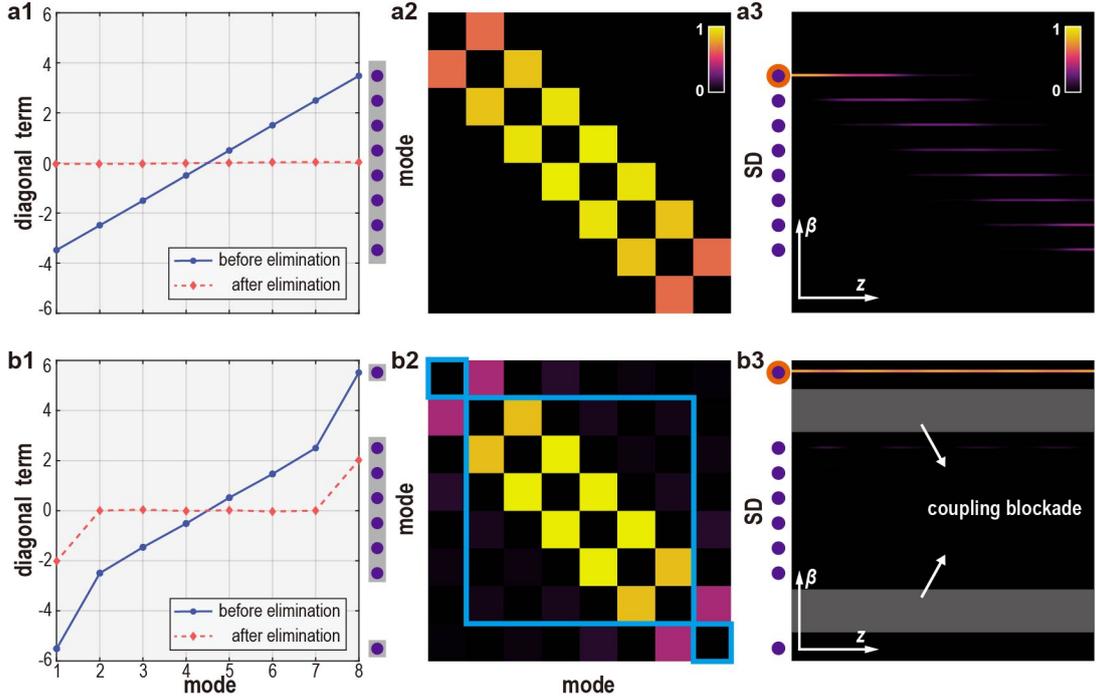

Fig. S3 Illustration of mode dynamics in different mode arrays. (a1-a3) Illustration of mode arrays with equal spacing of eigenvalues β_m. (a1) Distribution of $\mathrm{diag}(\Phi_A^\dagger H_0 \Phi_A)$ before (solid blue lines) and after (dashed red lines) elimination of the potential difference between eigenmodes by using a proper rotating wave frame. (a2) The off-diagonal values of $\Phi_A^\dagger H_1(z) \Phi_A$ indicate the couplings between eigenmodes in SD; we plot the normalized absolute values of the complex matrix elements.

(a3) Mode evolution in SD. The orange circle indicates the initially excited mode. (b1-b3) have the same layout as (a1-a3), except that they are for the mode arrays with outlying edges, showing that the excited mode is well confined in SD. The modes in the squared region indicated in (b2) are mutually coupled, and the shaded zones in (b3) show the coupling blockade between the edge and bulk modes in SD.

Note 3: Parameter tuning and phase extraction used in experiment

The coupling c in the tight-binding mode and waveguide spacing d in the simulation and experiment obey an exponential relation: $d = -d_0 \ln(c/c_0) + C_0$, where d and c are the distance and coupling between two waveguides, respectively. The parameters d_0, c_0, C_0 are constants employed in the simulation and experiment to establish the correspondence between couplings and waveguide distance. Specifically, in the present experiment, the values of d_0 and C_0 are set to $2.5 * 10^{-6}m$ and $12.4 * 10^{-6}m$. Since the minimum spacing between waveguides that we can write in the experiment is approximately $14.9 \mu m$, we set $c_0 = 5.48$ for the mode arrays with equal spacing and $c_0 = 11.42$ for mode arrays with outlying edges to have the minimum spacing between waveguides in the experiment. Therefore, the waveguide spacing is $\Delta d = [15.90, 15.25, 14.99, 14.90, 14.99, 15.25, 15.90] \mu m$ in mode arrays with equal spacing and $\Delta d =[17.99, 17.18, 15.77, 14.90, 15.77, 17.18, 17.99] \mu m$ in mode arrays with outlying edges. Both have the minimum spacing between waveguides $\min(\Delta d) = 14.9 \mu m$, which guarantees the maximum coupling strength along propagation for the mode arrays in the experiment. The refractive index change of the waveguide is around $\Delta n = 3.6 * 10^{-4}$.

On the other hand, to analyze long-distance light evolution in synthetic dimension (SD), we need to extract both the intensity and phase distributions of the output beams in the experiment, and then use them as the input to the next section of the waveguide arrays. Therefore, we apply the Fourier transform of the interferogram of output beams [47]. Assume we have the interferogram of the beams shown in Fig. S4a, which can be described by $I(x,y) = I_0(x,y) + I_1(x,y) + I_1^*(x,y)$, where $I_1(x,y) = \mathcal{I}(x,y) \exp[i\phi(x,y)]$ and $\mathcal{I}(x,y)$ and $\phi(x,y)$ are the amplitude and phase distributiuons, respectively. The Fourier transform of $I(x,y)$ yields three parts $\mathcal{F}(I(x,y)) = G_0(k) + G_+(k) + G_-(k)$ denoted in Fig. S4a2, where $G_+(k)$ contains the phase of the output beams. The inverse

Fourier transform of $g_+(x) = \mathcal{F}^{-1}\big(G_+(k)\big)$ uncovers the phase distribution of $I(x,y)$. Moreover,

$$g_+(x) = g_{+0}(x) \exp\left(i\left(\phi_p(x) + \phi_r(x)\right) \right) \tag{S9}$$

where $g_{+0}(x)$ is the intensity component, $\phi_p(x)$ is the phase carried by the probe beam and $\phi_r(x)$ is the phase of the reference beam. Notably, since only the phase of the probe beam is necessary, the interferogram between a broad beam (quasi-plane wave) and the reference beam
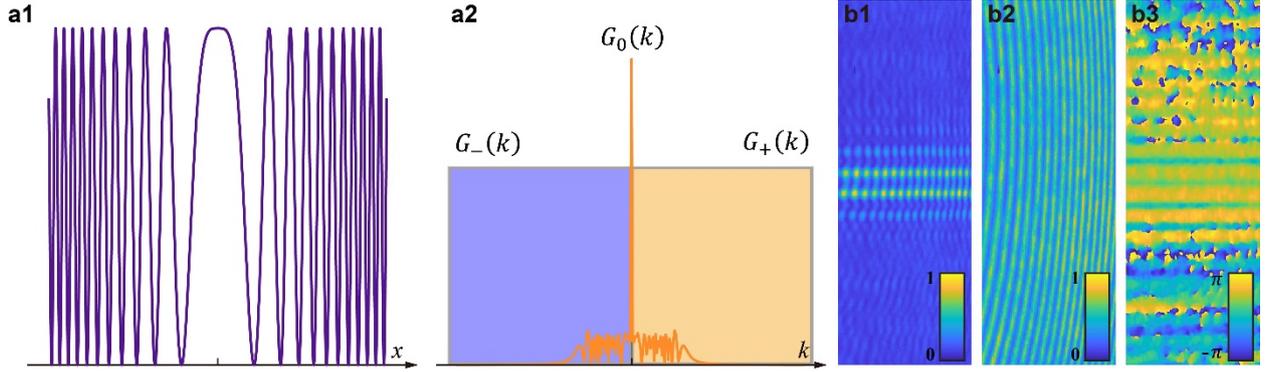

Fig. S4 Demonstration of phase extraction via the Fourier transform method. (a) An example to illustrate the Fourier transform method. (a1) A sample intensity distribution $I(x,y)$ as obtained from interferogram. (a2) The Fourier transform of $I(x,y)$. (b) The interferogram of an output beam obtained in the experiment. (b1, b2) The interferograms obtained with a reference beam interfering with the output beam (b1) and a quasi-plane-wave beam (b2). (b3) The extracted phase distribution of the probe beam.

Note 4: Illustration of the cw-laser writing and cascade probing method

As we stated in the main text, the amplitude R and wiggling frequency Ω of writing beams (532nm) can be tuned by the spatial light modulator (SLM), when the light is ordinarily polarized, as shown in Fig. 3a. However, limited by the length of the crystal $L = 20mm$, we need to truncate the writing beams into several parts to write the waveguide arrays by section. The length of each part is $20mm$ which fits the long side of the crystal (Fig. 3a). For section 1 denoted in Fig. 3a, we generate the first part of the writing beams via the SLM. The waveguide arrays can be written into the crystal subsequently due to the 'memory effect' of the nonlinear photorefractive crystal. The position of the writing beams is precisely controlled by SLM.

Once the lattice is written in the crystal, the probe beam is generated by the SLM. The polarization of the probe beam with extraordinary polarization is adjusted by a halfwave plate mounted before the nonlinear crystal. Both the intensity and phase distribution of the probe beam are captured at the output

in the experiment. In section 2, the second part of the writing beams is applied to write the subsequent section of the waveguide array. The intensity and phase distributions of the output beam captured in section 1 are properly duplicated by SLM and serve as the input beam of section 2 (Fig. 3a). Therefore, although light in each step propagates only 20mm, the output of section 2 has effectively 40 mm propagation distance, thanks to the cascade probing method described above. The propagation distance of light can be further elongated to 80 mm within the allowed error range caused by the extraction-duplication process via SLM.

Figure S4 shows the results obtained with the cascade probing method in the experiment. The intensity and phase distributions of the output beam at $z = 20mm$ match the input beam of section 2 ($z = 40mm$) duplicated by SLM. The consistency between the output beam and the following input beam guarantees the effectiveness of the cascade probing method.
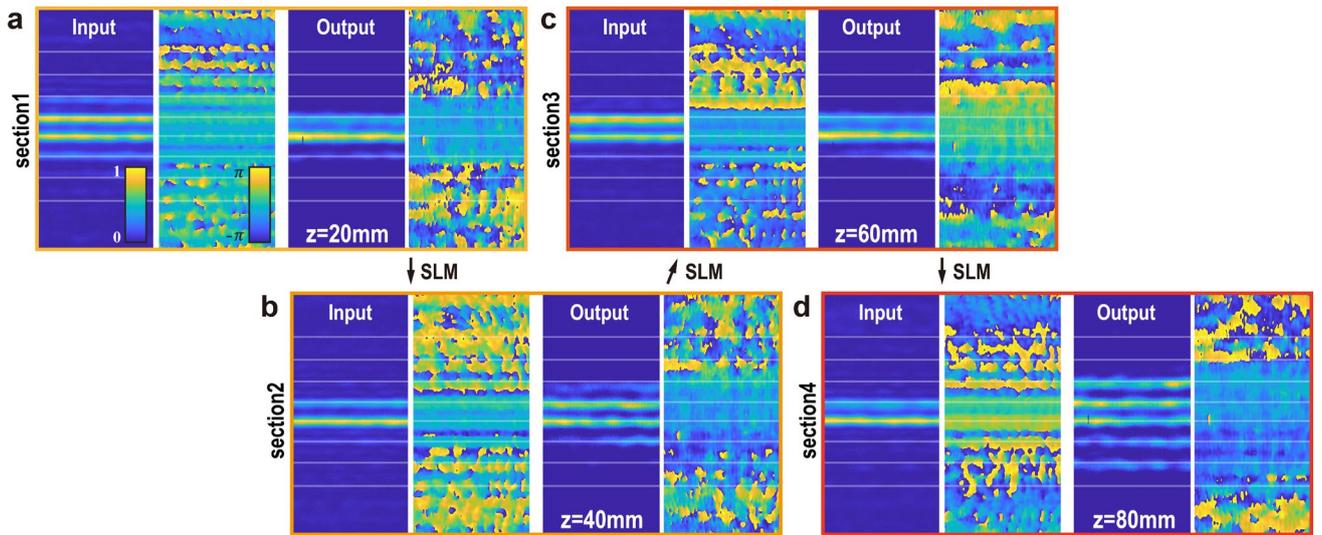

Fig. S5 Experimental results of the cascade probing method, showing input–output relation of the probe beam for the mode array with outlying edges. The arrows show the process of the extraction-duplication for recycling the probe beam. (a) to (d) are the intensity and phase distribution of the input and output beams in four sections at $z = 20, 40, 60, 80mm$. Each section of the waveguide arrays is 20 mm long, corresponding to the crystal length.

Note 5: Simulations of single mode excitation using the continuous model

The simulations of single mode excitation are governed by the paraxial Schrödinger equation (Eq. S1). We first construct the lattice by establishing the relation between D and T, where the waveguide spacing is $\Delta d = [15.90, 15.25, 14.99, 14.90, 14.99, 15.25, 15.90]\ \mu m$ in mode arrays with equal

spacing (Fig. S6a) and Δd =[17.99, 17.18, 15.77, 14.90, 15.77, 17.18, 17.99] μm in mode arrays with outlying edges (Fig. S6b), as stated in Note 3. Then, one mode is programmed as the probe beam. By numerically solving Eq. S1 with the beam propagation method, the sideview of light propagation in real space is obtained numerically and shown in the top row of Fig. S6. By calculating $\eta_i(z)$, the mode evolution in the continuous model is plotted in SD (bottom row of Fig. S6). The simulation parameters are taken from experiments. The agreement between Fig. 3 and Fig. S6 confirms the effectiveness of the tight-binding model for the discussion of mode evolution in SD.
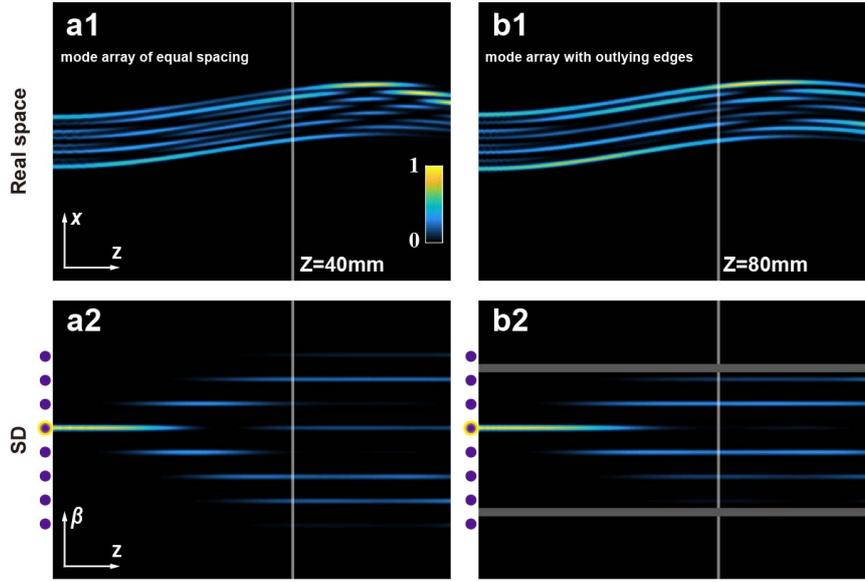

Fig. S6 Simulations of single mode excitation using the continuous model. (a) Light evolving in real space (top row) and SD (bottom row). Yellow circles marked in the left column indicate the excited modes. (b) have the same layouts as (a) but for differently selected mode excitations in different mode arrays: (a) are for mode arrays with equal spacing. (b) are for mode arrays with outlying edges corresponding to Fig. 3. The shaded zones indicate the coupling blockade. The propagation distance at the vertical line in (a) and (b) are $Z = 40mm$ and $Z = 80mm$.

Section 6: Modification of trivial and nontrivial SSH structures by ANNs.

The Su-Schrieffer-Heeger (SSH) model features fixed intra-cell (t_{intra}) and inter-cell (t_{inter}) couplings, resulting in two SSH structures with distinct topological properties (Fig. S7a1 and b1). In momentum space, the eigenvalues of bulk modes in the standard SSH model are not equally spaced, presenting a challenge for mode switching in SD. To address this, ANNs are employed to adjust the coupling coefficients in real space, yielding an equal-spacing eigenvalue distribution with

approximately linear dispersion shown in Fig. S7a3 and Fig. S7b3 for the topological trivial and nontrivial SSH structures, respectively. Importantly, such modification does not alter the topological properties of the modes, as the gap remains unclosed and the topological invariant remains unchanged (Figs. S6a2 and b2). Thus, the defect modes depicted in Fig. S7b3 retain their topological protection.
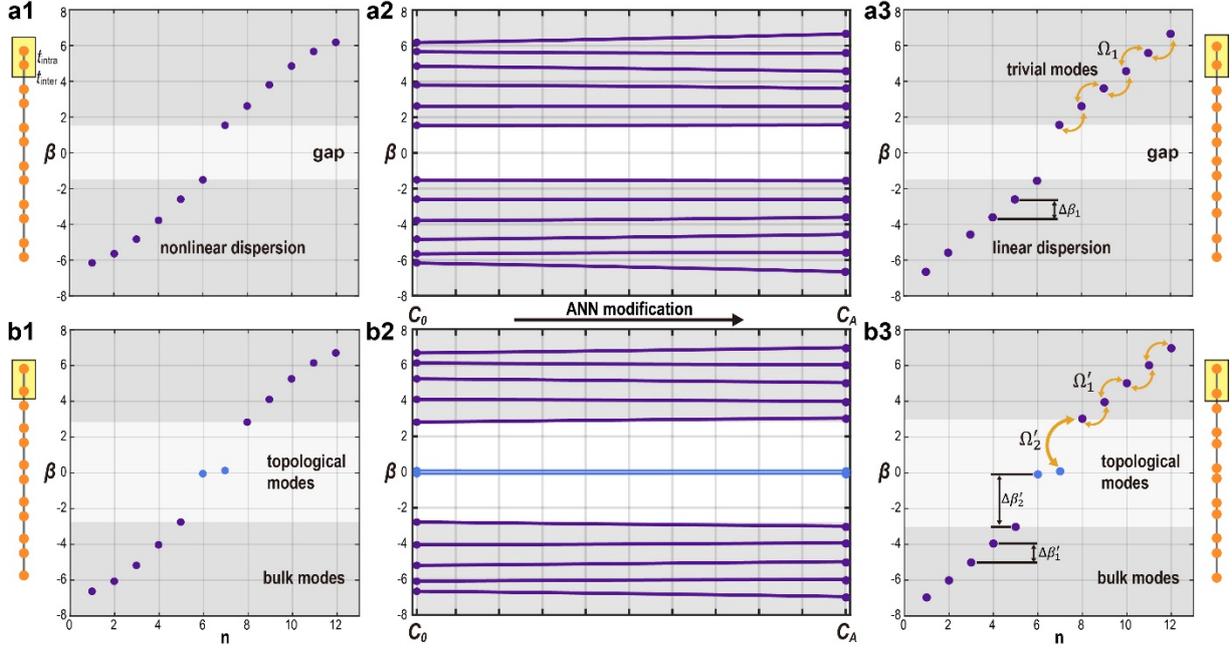

Fig. S7 Illustration of the SSH structures modified by ANNs. (a1) shows the standard trivial SSH structure with fixed coupling coefficients $t_{\text{intra}} = 3.6969$ and $t_{\text{inter}} = 2.6505$ (far left column), and its corresponding eigenvalue with nonlinear dispersion. The yellow box illustrates the unit-cell. (a2) shows the continuous modification of the trivial finite SSH lattice without gap closing. (a3) shows the trivial SSH structure modified by ANNs, resulting in an equal-spaced eigenvalue distribution, for which the coupling coefficients $T = [3.2542, 2.1986, 3.7560, 2.8882, 4.0805, 3.0787, 4.0805, 2.8882, 3.7560, 2.1986, 3.2542]$ form the lattice from top to bottom (far right column). (b1-b3) have the same layout as (a1-a3) but for the nontrivial case. (b1) displays the intra-cell and inter-cell couplings, which are $t_{\text{intra}} = 2.4434$ and $t_{\text{inter}} = 4.4211$, respectively. The modified nontrivial SSH structure with coupling arrays $T = [2.4189, 3.8624, 2.4353, 4.7121, 2.4760, 4.9567, 2.4760, 4.7121, 2.4353, 3.8624, 2.4189]$ is shown in (b3).