Supplementary Information

Supplementary Note 1: Dimensionality crisis.

In any data driven approach to an inverse problem, the number N_p of free parameters to be considered, has to be taken small enough to have a treatable design process. This is particularly true in ML techniques, where the "curse of dimensionality" is a well known problem.

ML requires training from a finite number of data samples, N_s, where each feature f_i, with $i = 1, ..., N_p$, have a range R_i of possible values. For the regressor to be effective, the distance between neighboring points for each feature have to be less than some value d_i, which depends on the scale of feature variations. This means that one needs $n_i = R_i/d_i$ points for each feature f_i and $N_s = \prod_{i=1}^{N_p} R_i/d_i$ data for the whole training. As a consequence, as the space feature dimensionality increases, the number of training points required for a good regressor grows exponentially.

So an interesting question is which features among ϵ_A, ϵ_B, β, χ and ξ should be used and how many points are needed for each feature.

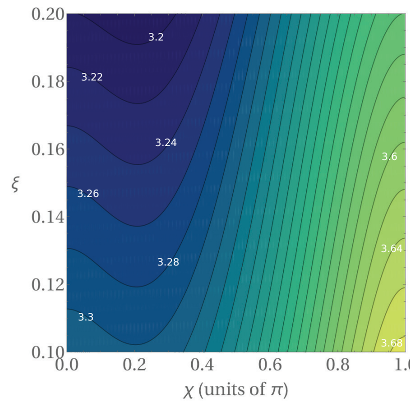
From the point of view of practical applications the choice of the dielectric function values can be made on the basis of the desired properties for the un-modulated structure; in particular position and width of its bandgaps, mainly determined by the refractive index contrast of high and low index layers: $\Delta\epsilon = \epsilon_A - \epsilon_B$.

Regarding the β parameter one has to consider that in order to have nontrivial Chern numbers and then topological edge states the structure must lack an inversion center[1]. The minimum number of layers A in the unit cell to achieve this is three, so the choice $\beta = 1/3$ realizes the simplest topological structure.

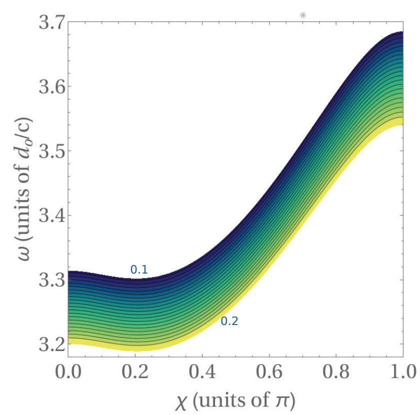
One is then left with the free parameters χ and ξ as part of the feature's set. The iso-frequency lines shown in Supplementary Figure 1 allows to estimate the number of points for each feature χ and ξ to have the optimal training set.

The iso-frequency lines are the intersections of constant frequency ω-planes to a dispersion surface: specifically the m_{31}^{+} mode. They quantify all allowed χ and ξ values and their corresponding frequencies showing a stronger anisotropy with respect to the χ feature than the ξ one. This justify our choice of a training set more sparse in the ξ direction. See Fig.3a) of the main text.

In the same way the iso-ξ and iso-χ contours give the allowed couples (ω,χ), Supplementary Figure 2, and (ω,ξ), Supplementary Figure 3, for the design of the topological structure. They visually reveal the presence of feature ranges where the inverse problem solution does not exist. Moreover, the lack of uniqueness of solutions to the inverse problem is rendered by the corresponding forward problem maps; see Supplementary Figure 1.



Supplementary Figure 1. Isofrequency lines. The contour lines are labeled with ω values in units of d_o/c.



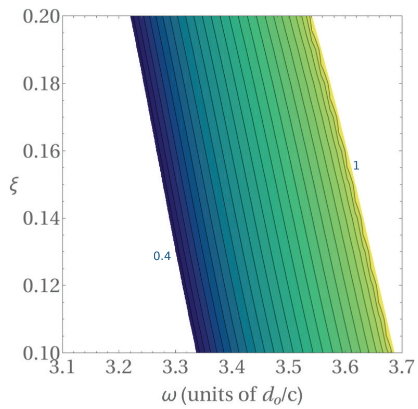
Supplementary Figure 2. Iso-ξ lines. The ξ values 0.1 and 0.2 of the first and the last contour lines are shown.

Supplementary Note 2: Consistency cycle.

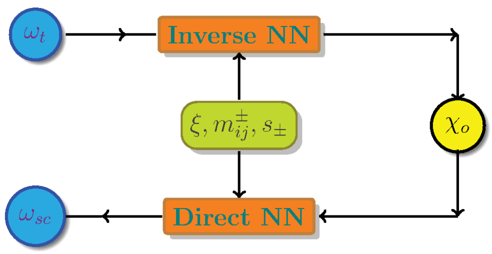
As reported in the main manuscript, we cannot take the solutions provided by the inverse NN as valid without verifying their consistency, because the NN may generate solutions that are not physical.

Let us consider as target a smooth, non monotic function $\omega_t(\chi)$. Without loss of generality, $\omega_t(\chi)$ may have two regions of monoticity: it decreases from $\omega_t = \omega'_{max}$ to $\omega_t = \omega_{min}$ for $\chi \in (0, \chi_1)$, and increases from $\omega_t = \omega_{min}$ to $\omega_t = \omega_{max}$ for $\chi \in (\chi_1, 1)$, with $\omega_{min} < \omega'_{max} < \omega_{max}$, $0 < \chi_1 < 1$. When we use an inverse NN to compute $\chi_o(\omega_t)$, ML is engineered in a way such that it generalizes the solutions and associates to every $\omega_t \in [\omega_{min}, \omega_{max}]$ a value $\chi_o \in [0, 1]$ for each branch, producing new, unphysical values for $\omega_t \in (\omega'_{max}, \omega_{max}]$. These values must be eliminated. To avoid the solution generalization, we test the self-consistency by the procedure sketched in Supplementary Figure 4.

After the training stage, implemented through numerically computed data (ω, χ, ξ), we use an inverse NN with input (ω_t, ξ) to attain an output χ_o. In order to have the unfolding and determine in advance the slope of $\omega_t(\chi)$ in



Supplementary Figure 3. Iso-χ lines. The χ values 0.4 and 1 of the first and the last contour lines are shown.



Supplementary Figure 4. Scheme of the self-consistent procedure for the machine learning inverse problem solution.

its different branches, we add the mode index $m_{ij}^{\pm}$ and the trend $s_{\pm}$ (defined in the main manuscript) as inputs. To clarify the idea, with reference to Supplementary Figure 5, let us fix the normalized thickness ξ and the mode index $m_{ij}^{\pm}$: we have two correspondences, i.e., on one hand

$$(\omega_t, s_-) \longrightarrow (\chi_o, s_-) \tag{S1}$$

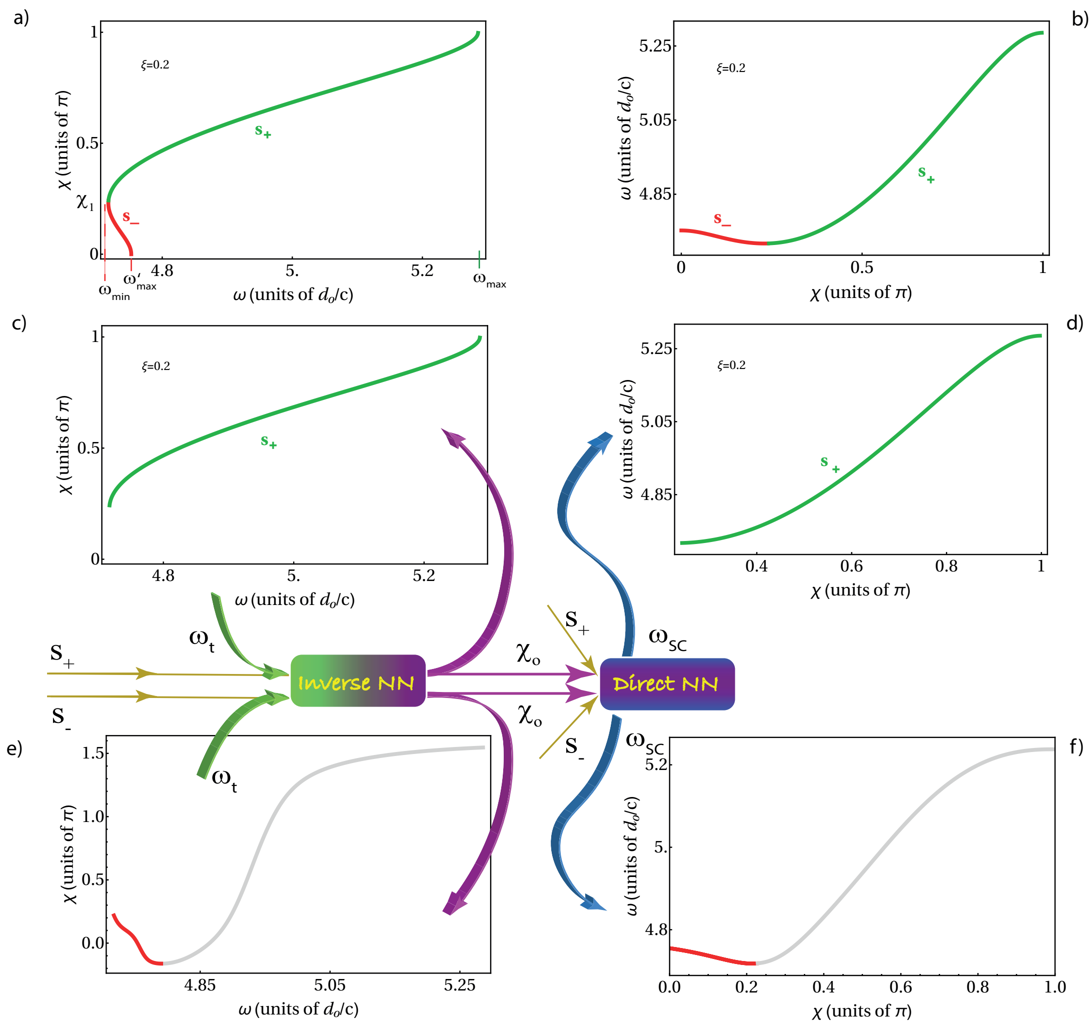
for the red branch, on the other hand

$$(\omega_t, s_+) \longrightarrow (\chi_o, s_+) \tag{S2}$$

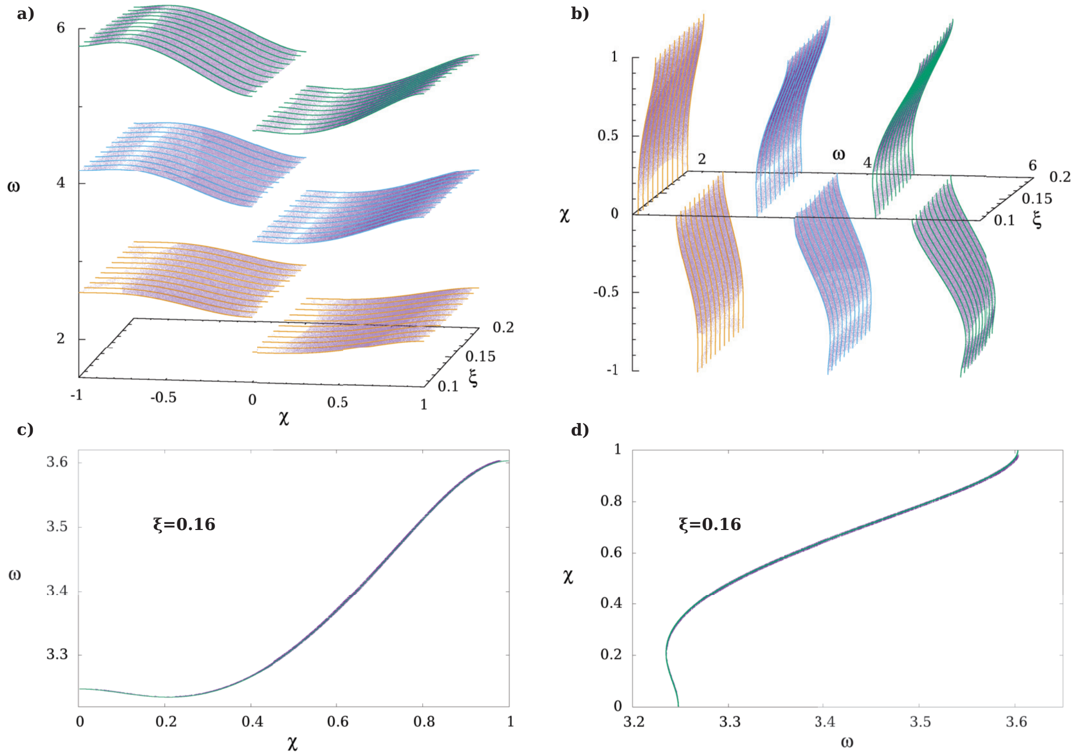
for the green branch, as illustrated in Supplementary Figures 5 a),b). While Eq. (S2) associates $\omega_t \in [\omega_{min}, \omega_{max}]$ to $\chi_o \in [\chi_1, 1]$, as in Supplementary Figure 5c, Eq. (S1) should associate just $\omega_t \in [\omega_{min}, \omega'_{max}]$ to $\chi_o \in [0, \chi_1]$, but in Supplementary Figure 5e is clairly shown that ML produces extra values, represented by the gray branch. Hence, how may we be sure that χ_o corresponds to a physical solution, i.e., χ_o does not belong to a forbidden region? We need to test the validity of our result by a direct NN that takes the previous values $(\chi_o, \xi, m_{ij}^{\pm}, s_{\pm})$ as input and gives back a self-consistent output ω_{sc}, reported in Supplementary Figures 5 d),f).
In order to complete our cycle, we need to compare the resulting value ω_{sc} with the input ω_t by the choice of a tolerance δ, which affects the model accuracy. Finally, if $|\omega_{sc} - \omega_t| < \delta$, then χ_o is accepted as solution of the inverse model, otherwise it is discarded.

Supplementary Note 3: Inverse problem results. Supplementary Figure 6 shows further data concerning the solution of the inverse problem in the main text. In particular we show the uncertainty in the determination of the physically relevant solutions determined by the parameter δ.

Supplementary reference.
1 Poshakinskiy A. V., Poddubny A. N., Pilozzi L., Ivchenko E. L., *Radiative topological states in resonant photonic crystals*, Phy. Rev. Lett. 112, 107403 (2014)



Supplementary Figure 5. **a)** Multivalued function $\chi_o(\omega_t)$, i.e., the final result of our ML consistency cycle. **b)** Non monotic function $\omega_t(\chi)$. It owns two monoticity regions, corresponding to a decreasing red branch and an increasing green branch. **c)** Result of the inverse NN with input $(\omega_t, \xi, m_{ij}^{\pm}, s_+)$: ML computes $\chi_o(\omega_t)$ for $\omega_t \in [\omega_{min}, \omega_{max}]$ and $\frac{\partial \omega_t}{\partial \chi} > 0$. Since, with this trend, the direct function $\omega_t(\chi)$ spans the whole interval $[\omega_{min}, \omega_{max}]$, we have that $\chi_o(\omega_t)$ coincides with $\omega_t^{-1}(\chi_o)$. **d)** Result of the direct NN with input $(\chi_o, \xi, m_{ij}^{\pm}, s_+)$: ML computes $\omega_{sc}(\chi_o)$ for $\chi_o \in [0,1]$ and $\frac{\partial \omega_{sc}}{\partial \chi_o} > 0$. **e)** Result of the inverse NN with input $(\omega_t, \xi, m_{ij}^{\pm}, s_-)$: ML computes $\chi_o(\omega_t)$ for every $\omega_t \in [\omega_{min}, \omega_{max}]$. Since the direct function $\omega_t(\chi)$ does not span the whole interval $[\omega_{min}, \omega_{max}]$ with trend s_- (it spans only $[\omega_{min}, \omega'_{max}]$), we have that $\chi_o(\omega_t)$ coincides with the red branch $\omega_t^{-1}(\chi_o)$ only for $\chi_o \in [0, \chi_1]$. Every point (ω_t, χ_o) of the gray branch is unphysical. Values of χ_o outside $[0,1]$ are discarded because corresponding to a forbidden range, as well as the points laying on the gray curve violating the condition $\frac{\partial \omega_t}{\partial \chi} < 0$. **f)** Result of the direct NN with input $(\chi_o, \xi, m_{ij}^{\pm}, s_-)$: ML computes $\omega_{sc}(\chi_o)$ for $\chi_o \in [0,1]$ and $\frac{\partial \omega_{sc}}{\partial \chi_o} < 0$. Also here, the gray branch represents unphysical values: $\omega_t(\chi)$ should span only $[\omega_{min}, \omega'_{max}]$ for the current trend, but since the input includes the interval $[\chi_1, 1]$ as well, the correspondence $[\chi_1, 1] \longrightarrow [\omega_{min}, \omega_{max}]$ is not physical.



Supplementary Figure 6. **a,b)** as in main text; **c)** (ω, χ) section of the inverse problem solution for $\xi = 0.16$; multiple lines correspond to the very small uncertainty in the solution resulting from the δ parameter that is fixed in the supervision step; **d)** (χ, ω) section of the inverse problem solution for $\xi = 0.16$; multiple lines correspond to the very small uncertainty in the solution resulting from the δ parameter that is fixed in the supervision step.