Real-space mapping of topological invariants using artificial neural networks

D. Carvalho,[1] N. A. García-Martínez,[2] J. L. Lado,[2,3] and J. Fernández-Rossier[2,*]

[1]*Physics department, Universidade de Aveiro, 3810-193 Aveiro, Portugal*
[2]*Quantalab, International Iberian Nanotechnology Laboratory, Avenida Mestre José Veiga, 4715-330 Braga, Portugal*
[3]*Institute for Theoretical Physics, ETH Zurich, 8093 Zurich, Switzerland*

Topological invariants allow one to characterize Hamiltonians, predicting the existence of topologically protected in-gap modes. Those invariants can be computed by tracing the evolution of the occupied wave functions under twisted boundary conditions. However, those procedures do not allow one to calculate a topological invariant by evaluating the system locally, and thus require information about the wave functions in the whole system. Here we show that artificial neural networks can be trained to identify the topological order by evaluating a local projection of the density matrix. We demonstrate this for two different models, a one-dimensional topological superconductor and a two-dimensional quantum anomalous Hall state, both with spatially modulated parameters. Our neural network correctly identifies the different topological domains in real space, predicting the location of in-gap states. By combining a neural network with a calculation of the electronic states that uses the kernel polynomial method, we show that the local evaluation of the invariant can be carried out by evaluating a local quantity, in particular for systems without translational symmetry consisting of tens of thousands of atoms. Our results show that supervised learning is an efficient methodology to characterize the local topology of a system.

I. INTRODUCTION

The study of topological electronic phases is one of the central topics in modern condensed-matter physics. Depending on the symmetry class different topological states exist, with the paradigmatic examples of time-reversal topological insulators [1], topological superconductors [2], topological crystal insulators [3], topological Kondo insulators [4], and topological Mott insulators [5] among others. The most fundamental quantity to characterize these states is the so-called topological invariant, the value of which determines the topological class of the system. In particular, interfaces between systems with different topological invariants show topologically protected excitations, resilient towards perturbations respecting the symmetry class of the system. Computationally, the calculation of the topological invariant usually requires the explicit knowledge of the wave functions of the entire system [6–8]. In particular, topological invariants can be calculated as the winding number of the occupied wave functions under twisted boundary conditions [6–8]. In that way, these methods generically require computing the full wave functions, that becomes a cumbersome task for systems without translational symmetry consisting of thousands of atoms.

In several situations of experimental relevance, translational symmetry is broken and systems are able to show different phases in real space due to the spatial modulation of the effective parameters. This situation might lead to protected modes between different regions of the system, dramatically changing the low-energy properties of the whole material.

This is the natural scenario in van der Waals heterostructures, where Moire patterns [9–11] could coexist with any topological state [12,13]. A more controlled situation is the proposals for topological superconductivity involving nanowires, where the topological state is controlled *locally* by electric gates [14,15]. Even though real-space formulations for the topological invariant do exist [16–20], their computation requires an integration over the whole space. Thus, there is not a simple methodology to obtain a topological invariant in inhomogeneous systems by evaluating solely their local properties.

Application of machine learning (ML) methods in condensed-matter physics is a growing area. A significant advantage of these techniques is that they are capable of finding the important degrees of freedom of a dataset without needing a profound insight of the treated problem. The identification of phase transitions [21–26] and the study of the ground state and correlations in different quantum many-body problems [27–32] are just some of the problems that machine learning has helped tackle in the past few years. Techniques have even been used in combination with *ab initio* calculations, allowing a broader and more accurate understanding of materials [25,33,34]. Within the language of machine learning, the calculation of topological invariants is understood as a simple classification algorithm [35,36], that could be efficiently tackled with the so-called artificial neural networks (ANNs) [37–42].

In this paper we show that ANNs are capable of characterizing the local topology of a system using as input a restricted amount of real-space information. In particular, we show that a trained ANN identifies correctly the local topological character in spatially varying Hamiltonians that create topologically different regions in space. Importantly, we show that this technique, used in conjunction with the

---

*On leave from Departamento de Física Aplicada, Universidad de Alicante, Alicante, Spain.

kernel polynomial method (KPM), allows one to compute local topological invariants with an algorithm the computational cost of which scales just linearly with the size of the system.

The rest of the paper is organized as follows. In Sec. II we review the basics of artificial neural networks (Sec. II A) and summarize the use of the kernel polynomial method to efficiently compute density matrices (Sec. II B). In Sec. III we apply the combined ANN-KPM technique both to a model Hamiltonian for a one-dimensional (1D) topological superconductor (Sec. III A) and to a two-dimensional (2D) anomalous Hall insulator (Sec. III B). Finally, in Sec. IV we present our conclusions.

II. METHOD

A. Artificial neural networks

Machine learning is a broad field that includes many different approaches, goals, and methods [43]. The defining property of ML algorithms is that they allow computers to perform specific tasks without being explicitly programed for each one of them [44]. Within the vast variety of ML algorithms, we will focus on supervised learning algorithms, which require a training dataset to fit the parameters in the model. One of the most common models of supervised learning is the ANN, which has been proven very useful to model patterns and correlations of complex problems that cannot be modeled analytically such as image or sound recognition [37–39], and even natural language processing [40,41]. In our case, we aim to use an artificial neural network to characterize locally the topological state of a 1D [Fig. 1(a)] or 2D [Fig. 1(b)] system. The objective of the procedure is to have a neural network that, given local information about the system, returns the topological invariant as sketched in Fig. 1(c). The local information that will be provided is a local block of the density matrix of the system, as we will discuss later.

ANNs are loosely based on parts of the brain, consisting of neurons, modeled as perceptrons [45], and synapses as shown in Fig. 1(a). The neurons in an ANN do not attempt to model the actual structure or behavior of the biological cells [46]. Instead, they mimic one of their main features, the activation function. This activation function, σ, sketched in Fig. 1(b) provides the output of each neuron based on the received inputs and an external parameter (bias). For computational convenience, σ should be any smooth and differentiable function defined over $\mathbb{R}$ but with its range restricted to a closed interval, namely, $\sigma \in [-1,1]$, as depicted in Fig. 1(e). Usually, these functions are either the tanh or the sigmoid function, but others might be used without loss of generality or functionality since these models are only weakly sensitive to these details [47].

The inputs X entering each neuron are weighted by the synapses W and shifted by the bias b. The synapses' weights are parameters to be tuned and they can be arranged as rectangular matrices, W^α, so the output L of the layer α can be obtained simply as $L_\alpha = \sigma(X_\alpha \cdot W_\alpha + b_\alpha)$, where X_α is the input of the layer α (note that for the hidden layers $X_\alpha = L_{\alpha-1}$). As a formative example the outputs of every layer of the toy model sketched in Fig. 1(d) can be calculated as follows:

$$L_0 = \sigma(X_0) \quad \text{or just} \quad L_0 = X_0,$$
$$L_1 = \sigma(L_0 \cdot W_1 + b_1), \quad L_2 = \sigma(L_1 \cdot W_2 + b_2) = \hat{y} \quad (1)$$
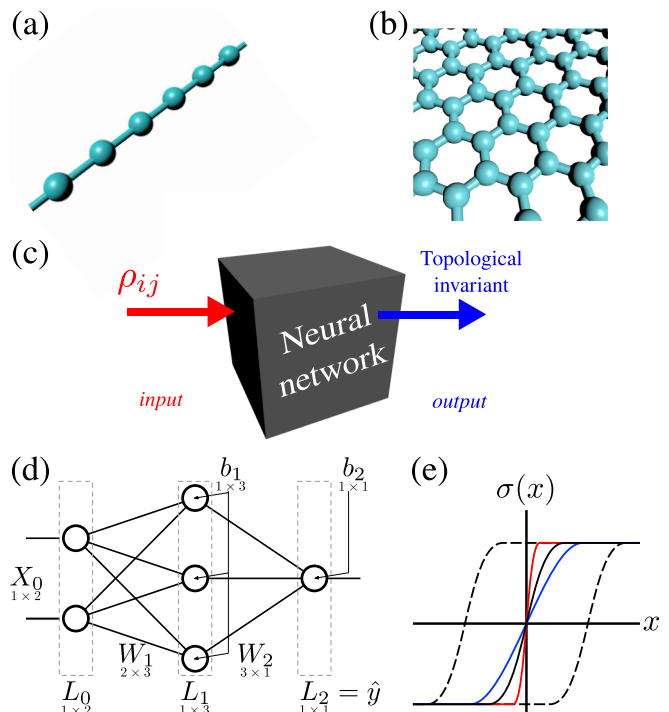

FIG. 1. (a, b) The two different geometries of the model Hamiltonians considered below, a one-dimensional topological superconductor (a) and a two-dimensional quantum anomalous Hall insulator (b). (c) A schematic sketch of our procedure: a trained neural network will take as input a local density matrix, and it will return the topological invariant of the system. (d) An artificial neural network as described in the text. (e) The standard behavior of an activation function of a neuron, $\sigma(x)$, for different weights (colors) and bias (dashed lines).

where X_0 is the input fed to the ANN. The matrices W_α and the arrays b_α are the parameters to be fitted during the training process in order to modify the activation functions of each of the neurons in the ways showed in Fig. 1(e). Note that the number of parameters in ANN models grows very quickly with the size (number of neurons per layer) and depth (number of layers) of the network.

An artificial neural network, like every supervised learning algorithm, consists of three phases. First, the architecture of the model (i.e., the number of layers and neurons per layer) is decided depending on the complexity of the problem addressed. Second, the model is trained. In this process, several input-output pairs are provided to the model the parameters of which are fitted to mimic the correlations present in the user-provided data. Finally, when the training is completed, the model can be used to evaluate new (unseen) input data.

Supervised learning algorithms require a training dataset to optimize the parameters of the models. The training is performed by minimizing a cost function, $\mathcal{E}$, usually proportional to the squared difference between the expected output, y, and the actual output of the network, $\hat{y}$:

$$\mathcal{E} = \tfrac{1}{2}(y - \hat{y})^2. \quad (2)$$

Notice that y is a constant defined by the (user-provided) training dataset while $\hat{y}$ depends on all the parameters of the network (weights and bias). The minimization of $\mathcal{E}$ is, then,

performed by iteratively modifying the values of all the weights and bias in the network until the desired output is obtained. This is a computationally complex and expensive process since the number of parameters can range from a few tens to millions. In fact, it was not until 1986 that an efficient method was developed for such a purpose [48]. We use the gradient descent with the back-propagation algorithm to train the ANN, which is the most common approach nowadays. We used the open source library PyBrain [42], to create, train, and evaluate the ANN.

B. Correlation functions with the kernel polynomial method

In this section we review how real-space correlation functions can be efficiently calculated using the KPM [49]. We will focus the discussion on the case of a normal electronic system, since the case of a superconductor can be treated in an analogous way. The main task that we have to perform is to obtain the density matrix, evaluated in a restricted area of real space, of a certain (very large) Hamiltonian. In terms of the eigenfunctions $|\Psi_k\rangle$ of the Hamiltonian H, the elements of the density matrix can be written as

$$\rho_{ij} = \int_{-\infty}^{E_F} \langle i|\Psi_k\rangle\langle\Psi_k|j\rangle\delta(E_k - \omega)d\omega \tag{3}$$

where $|i\rangle$ and $|j\rangle$ are the elements of the basis for the Hamiltonian H and E_F is the Fermi energy. The diagonal elements of the matrix, ρ_{ii}, are the integrated local density of states. In the gapped state, the off-diagonal elements are expected to decay exponentially with distance. So, when the Fermi energy E_F lies in the gap, the density matrix is properly described by restricting the calculation to a set of neighboring sites. Generically, calculating the previous matrix requires diagonalizing the full Hamiltonian to obtain the occupied wave functions, a task that scales with $O(N^3)$, with N the system size of the system. The kernel polynomial method allows the computation of ρ_{ij}, for a restricted set of neighboring sites, with a computational cost that scales only as $O(N)$.

The KPM allows one to compute the quantity

$$g_{ij}(\omega) = \sum_k \langle i|\Psi_k\rangle\langle\Psi_k|j\rangle\delta(E_k - \omega) \tag{4}$$

which can easily be integrated to obtain the density matrix (3). The central idea is that g_{ij} can be expressed in matrix form as

$$g_{ij}(\omega) = \langle i|\delta(H - \omega)|j\rangle. \tag{5}$$

The KPM consists of expanding Eq. (5) in terms of Chebyshev polynomials $T_n(\omega)$. To do so, the Hamiltonian is first rescaled so that all the eigenenergies lie in the interval $\mathcal{E}_k \in (-1, 1)$. The rescaled Hamiltonian is denoted as $\mathcal{H}$. The corresponding spectral function is calculated as

$$g_{ij}(\omega) = \frac{1}{\pi\sqrt{1 - \omega^2}}\left(\bar{\mu}_n + 2\sum_{n=1}^{N} \tilde{\mu}_n T_n(\omega)\right). \tag{6}$$

The coefficients $\tilde{\mu}_n$ determine the expansion of a certain element g_{ij}, and are calculated as

$$\tilde{\mu}_n = g_n^N \mu_n \tag{7}$$

where μ_n are the coefficients calculated from the Hamiltonian $\mathcal{H}$ and g_n^N denotes the Jackson kernel that improves the convergence of the series [49]:

$$g_n^N = \frac{(N - n - 1)\cos\frac{\pi n}{N+1} + \sin\frac{\pi n}{N+1}\cot\frac{\pi}{N+1}}{N + 1}. \tag{8}$$

Given two sites i and j, we define two vectors located in those sites v_i and v_j. The coefficients μ_n would be calculated as a conventional functional expansion:

$$\mu_n = \langle v_i| \int_{-1}^{1} \delta(\mathcal{H} - \omega)T_n(\mathcal{H})d\omega|v_j\rangle, \tag{9}$$

which in the diagonal basis reads

$$\mu_n = \int_{-1}^{1} \langle v_i|\Psi_k\rangle\delta(\mathcal{E}_k - \omega)\langle\Psi_k|v_j\rangle T_n(\omega)d\omega. \tag{10}$$

Performing the integration over ω we get

$$\mu_n = \langle v_i|\Psi_k\rangle T_n(\mathcal{E}_k)\langle\Psi_k|v_j\rangle = \langle v_i|T_n(\mathcal{H})|v_j\rangle. \tag{11}$$

Therefore, the coefficients μ_n can be calculated as the overlap of two vectors:

$$\mu_n = \langle v_j|\alpha_n\rangle \tag{12}$$

where α_n is calculated with the recursion relations associated to the Chebyshev polynomials

$$\begin{aligned}|\alpha_0\rangle &= |v_i\rangle, \\ |\alpha_1\rangle &= \mathcal{H}|\alpha_0\rangle, \\ |\alpha_{n+1}\rangle &= 2\mathcal{H}|\alpha_n\rangle - |\alpha_{n-1}\rangle. \end{aligned} \tag{13}$$

This procedure thus involves matrix vector products to calculate the coefficients. For a sparse matrix, as is the case of a tight-binding Hamiltonian, the number of nonzero elements scales linearly with the system size, so the computational cost of calculating the density matrix for a fixed number of sites also scales linearly. This method allows one to compute g_{ij} at every energy simultaneously, so that ρ_{ij} can be calculated by integration up to the Fermi energy.

For small systems, the density matrix can be calculated also by exact diagonalization of the full Hamiltonian. In principle, that procedure allows one to calculate the correlation function of relatively large one-dimensional systems. However, for a two-dimensional system, the dimension of the matrix will be too large in general. It is in that situation when the kernel polynomial method is especially suitable.

III. TOPOLOGICAL INVARIANTS WITH SUPERVISED LEARNING

We now describe the procedure to characterize the local topological character of a system using ANNs. We choose as input the elements of the density matrix that involve one site and its nearest neighbors (ANNs). This procedure allows one to naturally treat systems without translation symmetry and with disorder, as the calculation of the density matrices is not more computationally expensive in those situations using the previous procedure. The process to calculate the density matrix was discussed in Sec. II B but in order to use the density matrix as input for our ANN some processing is required. The density matrix is, in general, a complex Hermitian matrix, so
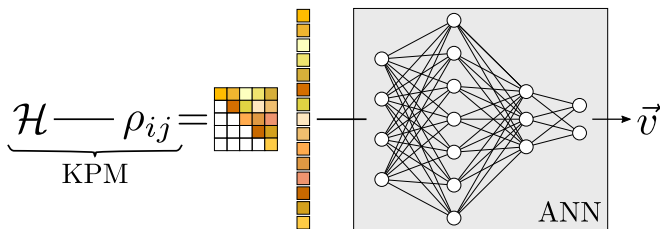
FIG. 2. The process to evaluate the topological character of a local region of space. The density matrix corresponding to a certain area in space is calculated using the KPM; after removing the redundant elements the matrix is rearranged in a 1D array that is used as input for a neural network that will provide the corresponding topological invariant as output.

we will remove the redundant elements (namely, the lower triangle) and arrange the remaining elements in a 1D array, concatenating the real and imaginary parts. Furthermore, we included the eigenvalues of the density matrix as part of the input. Strictly, the inclusion of the eigenvalues is a redundant operation that could be avoided by increasing the size and/or depth of the ANN, yet we found that it helped the optimization of the model with a negligible computational overhead. The output of the ANN will be the topological invariant of the system. The calculation of the corresponding output is done by constructing a translationally invariant Hamiltonian in which the corresponding topological invariant is well defined and can be calculated in a standard way. Finally, since we are using the ANN as a classifier, it is convenient to encode the possible outputs as linearly independent vectors, $\vec{v}$, rather than use a single scalar. The use of vectors allows the discrimination between wrong answers and false positives. This whole architecture is sketched in Fig. 2.

In order to train the ANN we generate a large number of realizations of a family of Hamiltonians, exploring their parameter space. For a given choice of parameters, we compute the topological invariant of the corresponding pristine case and its local density matrix. This procedure allows us to generate a set of inputs and outputs, which are used to train the ANN. Once the ANN is trained, the model is ready to be evaluated with new data that the network has never tried to test the accuracy of the network. The last step is to create a new Hamiltonian with spatially dependent parameters, and evaluate the ANN with the local density matrix corresponding to a neighborhood of every lattice site. In this way, we have a procedure to locally evaluate the topological invariant of a system lacking translational symmetry.

A. One-dimensional topological superconductor

In the following, we will consider a lattice model Hamiltonian for a one-dimensional electron gas that is able to host both trivial and topological superconducting states. The corresponding topological invariant is a Z_2 number that can be calculated as a Berry phase [50]. Such an effective one-dimensional system, in particular the superconducting topological phase, is realized in semiconducting nanowires deposited on top of a *s*-wave superconductor [51–57]. The model describes electrons in a 1D chain, in the presence of Zeeman field, Rashba
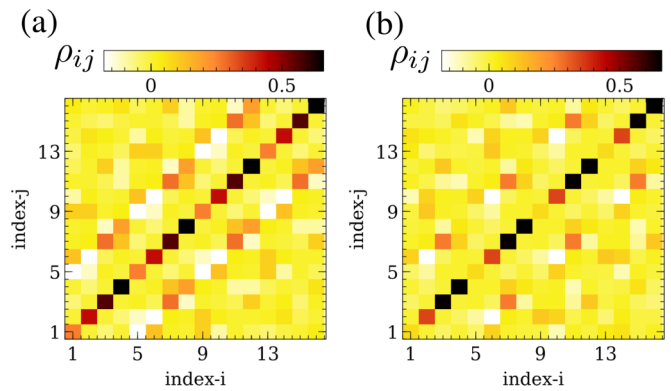

FIG. 3. Image representation of the density matrix for two particular different states for the superconducting 1D system, trivial (a) and topological (b). In terms of the matrices shown, the task of the neural network can be understood as an image recognition algorithm capable of distinguishing a trivial input from a topological input, for different parameters of the Hamiltonian chosen randomly. The different indices in the *x* and *y* axis run over spin and electron-hole sectors in the closest sites.

spin-orbit coupling, superconducting proximity effect, and a sublattice imbalance term. Thus, the model has six different parameters: a spin-conserving hopping t, chemical potential μ, Rashba spin orbit t_R, external Zeeman field B_z, on-site pairing term Δ, and a trivial mass m. Moreover, we also include the possibility of having finite Anderson disorder W_i, so that the full Hamiltonian reads

$$
\begin{aligned}
\mathcal{H} = {} & -t \sum_{\langle ij \rangle_\alpha} c_{i\alpha}^\dagger c_{j\alpha} + it_R \sum_{\langle ij \rangle_{\alpha\beta}} \hat{e}_z \cdot (\vec{\sigma}_{\alpha\beta} \times \vec{d}_{ij}) c_{i\alpha}^\dagger c_{j\beta} \\
& + B_z \sum_{i\alpha} c_{i\alpha}^\dagger \sigma_z c_{i\alpha} + \Delta \sum_i [c_{i\uparrow} c_{i\downarrow} + c_{i\downarrow}^\dagger c_{i\uparrow}^\dagger] \\
& + \mu \sum_{i,\alpha} c_{i\alpha}^\dagger c_{i\alpha} + m \sum_i \tau_i c_i^\dagger c_i + \sum_{i,\alpha} W_i c_{i\alpha}^\dagger c_{i\alpha}. \quad (14)
\end{aligned}
$$

The previous Hamiltonian can have topological and trivial phases. In a nutshell, a topological phase may arise when the Zeeman term B_z is such that the chemical potential μ crosses only one of the spin channels, so that a small pairing Δ and Rashba field t_R give rise to a spinless *p*-wave superconductor [55]. In the absence of both Zeeman and Rashba couplings, the induced superconducting gap is trivial.

The Hamiltonian (14) is solved in the Nambu representation by defining a spinor wave function as $\Psi^\dagger = (c_\uparrow^\dagger, \ c_\downarrow^\dagger, \ c_\downarrow, \ -c_\uparrow)$ which gives rise to a Bogoliubov–de Gennes Hamiltonian $\mathcal{H} = \frac{1}{2} \Psi^\dagger H \Psi$. The matrix H is used to calculate the correlation functions $\langle c_{i,s} c_{j,s'} \rangle$ and $\langle c_{i,s}^\dagger c_{j,s'} \rangle$, as introduced in Sec. II B, by integrating the different $g_{ij}(\omega)$ from $\omega = -\infty$ up to $\omega = 0$. In Fig. 3 we show an example of two different input data from the training dataset, for a topological [Fig. 3(a)] and a trivial [Fig. 3(b)] state computed for an open chain with $N = 400$ sites using the KPM. It is evident that simple inspection is not enough to distinguish between the two of them.

In order to generate the training dataset we considered different Hamiltonians for a bipartite chain with 400 sites by

varying the different values for the off-plane Zeeman field B_z, Rashba λ_R, chemical potential μ, superconducting pairing Δ, and sublattice imbalance m. In order to prove the robustness of our procedure, we also switch on the Anderson on-site disorder [$W \in (0.0, 0.4t)$], with a magnitude comparable to the other energy scales. For the training dataset we generated 1000 different Hamiltonians with parameters randomly chosen in the following ranges: $t_R \in [-0.3t, 0]$, $B_z \in [0.2t, 0.8t]$, $\mu \in [t, 2t]$, $\Delta \in [0.1t, 0.3t]$, $m \in [-0.2t, 0.2t]$, yielding a five-dimensional phase space. Using the generated Hamiltonians we calculate the density matrix of the central atom in the nanowire, ρ_{ij}, and its three closest neighbors. Since the Hamiltonian in Eq. (14) only involves two Pauli matrices for a linear chain, the Hamiltonian in real space can be chosen to be purely real, so that its density matrix will be also real. For each example the Z_2 topological invariant is calculated for the pristine system ($W_i = 0$) defined by that particular set of parameters, which is used as expected output. Since this topological invariant only has two possibilities, we encode the Z_2 invariant as a two-dimensional vector v, so that the topological case corresponds to $v = (1, 0)$ and the trivial case corresponds to $v = (0, 1)$. With this methodology a single element of the training dataset has a 152-dimensional input and a two-dimensional output. We took two hidden layers with 101 and 21 neurons. After training, a validation set with 200 new samples is generated to test the accuracy of the ANN yielding an accuracy of $\sim 97\%$. In order to gain some insight on the ANN capabilities, we run a simple test by freezing all the parameters in the Hamiltonian (14) but the chemical potential and comparing the actual Z_2 with the output provided by the ANN. In Fig. 4(a) we see that even for unseen data the ANN is able to provide the correct topological invariant.

Once the network is trained, it is ready to be used in the case of an inhomogeneous system. We now generate a one-dimensional system following Eq. (14) with spatially varying couplings. In particular, we modulate the chemical potential along the chain as shown in Fig. 4(b). Such modulation is feasible by means of local gates in the experimental realizations involving semiconducting nanowires [15]. With such modulation, we observe the emergence of zero-energy modes in the local density of states [Fig. 4(c)], which are expected to be a signature of a boundary between a trivial and topological phase. The evaluation of the topological invariant on every atomic position of the chain can be carried out by feeding the local density matrix to the trained neural network. Our network shows that the different regions of the space have different topological invariants as shown in Fig. 4(d). It is observed that the points of space where the topological invariant changes in Fig. 4(d) correspond to the location of the zero-energy Majorana modes, as seen in Fig. 4(c), validating the performance of our neural network.

The success of the neural network in describing the topological order of the different phases implies that, locally, the density matrix carries enough information to distinguish between the two cases. In particular, the elements of ρ_{ij} involving $\langle c_{i,s} c_{j,s'} \rangle$ encode information about the induced superconducting order parameter, both in the s- and p-wave channels, which physically is expected to determine the topological phase. If, in comparison, only the diagonal part of the density matrix was used as input for the neural network,
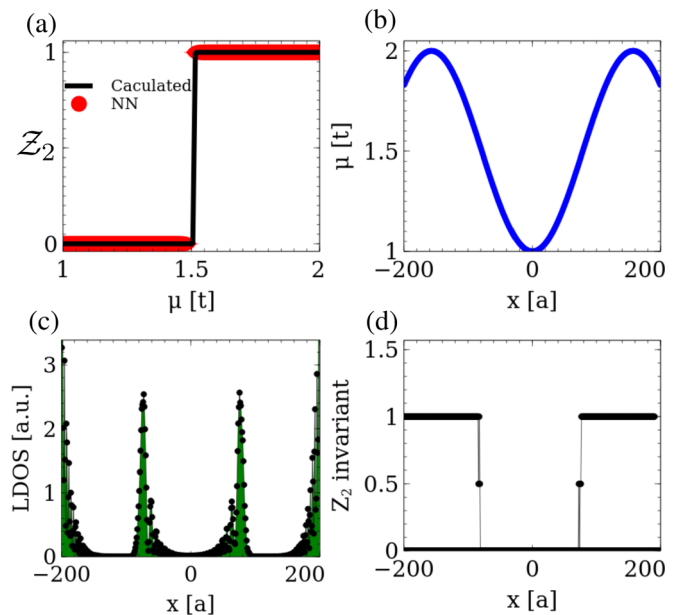

FIG. 4. (a) Comparison of the topological invariant computed exactly with the one predicted by the trained neural network in a pristine system, showing that the ANN perfectly captures the phase transitions in the homogeneous system. Afterwards, we create an inhomogeneous system with modulated chemical potential as shown in panel (b). Such modulation creates trivial and topological zones, with Majorana modes pinned at the transition points (c). The neural network is then evaluated in every point of the space, yielding a site-dependent topological invariant shown in panel (d). The topological transitions shown in panel (d) mark the existence of zero Majorana modes obtained in panel (c). The parameters used are $\lambda_R = -0.3t$, $B_z = 0.5t$, $m = 0$, and $\Delta = 0.1t$.

it would not be possible to distinguish between trivial and topological states. This is easily understood taking into account that the diagonal part of ρ_{ij} accounts for the total occupation numbers and two topologically inequivalent band structures can have arbitrarily similar density of states.

B. Two-dimensional Chern insulator

In this section we will use an analogous methodology to study a topological two-dimensional state. In particular, we consider a model Hamiltonian for electrons moving in a honeycomb lattice with Rashba spin-orbit coupling t_R and off-plane exchange B_z, that is known to result in a two-dimensional quantum anomalous Hall state [58]:

$$H = -t \sum_{\langle ij \rangle_\alpha} c_{i\alpha}^\dagger c_{j\alpha} + i t_R \sum_{\langle ij \rangle_{\alpha\beta}} \hat{e}_z \cdot (\vec{\sigma}_{\alpha\beta} \times \vec{d}_{ij}) c_{i\alpha}^\dagger c_{j\beta}$$
$$+ B_z \sum_{i\alpha} c_{i\alpha}^\dagger \sigma_z c_{i\alpha} + m \sum_{i,\alpha} \tau_i c_{i\alpha}^\dagger c_{i\alpha} + \sum_{i,\alpha} W_i c_{i\alpha}^\dagger c_{i\alpha}$$

(15)

where t_R is the Rashba coupling, $\vec{\sigma}$ are the spin Pauli matrices, B_z is the external Zeeman field, and $\tau_i = \pm 1$ is the sublattice operator. The first term is the usual tight-binding hopping term, the second one describes the Rashba interaction [58,59], and the third term is the so-called exchange or Zeeman term which

couples to the spin degree of freedom. The fourth term is a trivial mass term that assigns an opposite on-site energy for the atoms in each of the sublattices, that we introduce in order to have a trivial insulator phase in the model. Finally, the last term is an Anderson disorder term that we introduce to prove the robustness of the procedure. For $m = 0$, and $B_z \neq 0$ and $t_R \neq 0$, the model has a topological gap with Chern number $\mathcal{C} = \pm 2$. For $m \neq 0$ and $B_z = 0$, the model has a trivial ($\mathcal{C} = 0$) gap.

Each of these Hamiltonian terms can effectively describe different experimental situations. The sublattice imbalance could arise for a graphene monolayer deposited on boron nitride in a commensurate fashion [11,60]. The Rashba and exchange fields naturally arise for a graphene monolayer deposited over a ferromagnetic insulator, such as YIG [61,62], EuO [63], or CrI$_3$ [64,65]. Furthermore, the noncommensuration of graphene with the substrate creates Moire patterns, resulting in an effective spatial modulation of the different contributions [9–11].

It is worth mentioning two important differences with respect to the model presented in Sec. III A. On one hand, now the Hamiltonian involves the three Pauli matrices, so in general it will be complex. This implies that the calculated density matrices will also be complex, so that the neural network will receive as input both the real and imaginary components. On the other hand, since we are dealing now with a two-dimensional system, a finite island will have L^2 sites, with L the typical size of the island. In particular, the calculation of the density matrix with the wave functions of an island with side $L \approx 300$ would require the diagonalization of a matrix of dimension $L^2 \approx 90\,000$, the computational complexity of which is L^6. It is in this situation where the KPM will be especially useful, as it allows us to calculate the density matrix with a computational complexity of the number of sites, L^2.

We now move to apply our methodology to the system defined by Eq. (15). First, to train the neural network, we generate different spatially uniform Hamiltonians by choosing randomly each of the coupling parameters. The Zeeman and Rashba were randomly generated in the interval $t_R \in [-0.4t, 0.4t]$ and the mass between $m \in [0, 0.4t]$. Again, random Anderson-like disorder comparable to the other interactions is introduced all across the system $W_i \in [0.0, 0.4t]$. The training dataset consisted of 564 samples. For every set of parameters, we built the Hamiltonian as in Eq. (15) and calculated the local density matrix for the central atom and its three first neighbors which are used as input of the network, in this case a 128-dimensional array. Again we chose having two hidden layers with 101 and 21 neurons. It is worth considering again the challenging task of distinguishing between different inputs as those shown in Fig. 5, which highlights that the classification of topological and trivial phases based only in local properties is far from being a trivial task.

The output for each input was obtained by calculating the Chern number of the ground state of the system integrating the Berry curvature in the Brillouin zone of a translational invariant ($W_i = 0$) Hamiltonian with the same parameters. Once the network was trained, we tested its accuracy on a validation dataset with 586 samples randomly generated, showing an accuracy of ~92%. The comparison of the result predicted

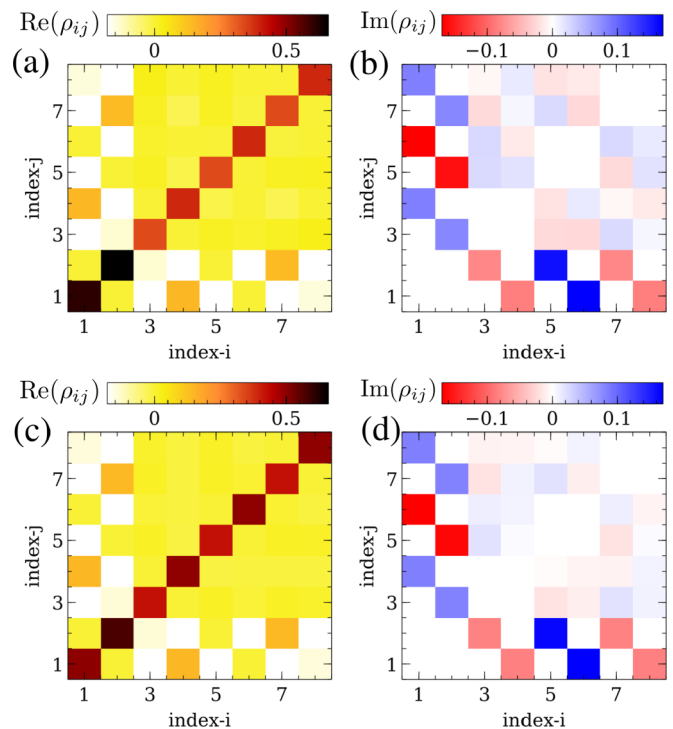FIG. 5. Real and imaginary parts of the density matrix for a trivial $\mathcal{C} = 0$ (a, b) and a topological $\mathcal{C} = 2$ (c, d) two-dimensional system. In this case, the neural network will implement an image recognition algorithm, where the inputs are the two images representing the real and imaginary parts.

by the network and the one calculated exactly in a system with translational invariance is shown in Fig. 6(a) for the different topological phases.

After the training, we generated a graphene nanoisland with 7400 atoms. In this island, we choose a spatially modulated exchange field of the form $B_z(x, y) = 0.1t[\cos(0.15x) + \cos(0.15y) + 2]$, a modulated mass term of the form $m(x, y) = 0.1t[\sin(0.15x) + \sin(0.15y) + 2]$ [shown in Fig. 6(b)], and a constant Rashba coupling $\lambda_R = 0.2$. The previous modulations are expected to create neighboring trivial and topological areas depending on which is the dominant contribution, mass or exchange and Rashba couplings. With such a Hamiltonian, we calculated the local density matrix using the kernel polynomial method, that was used as input of the neural network. The result of the evaluation of the neural network across the sample is shown in Fig. 6(c). It can be seen that different regions with different Chern number appear according to the spatial modulation of the Hamiltonian parameters. The significance of the different regions becomes clear once the in-gap density of states is calculated in Fig. 6(d). This shows both in-gap modes precisely at the boundary between different regions, as expected form the bulk-boundary correspondence, as well as edge states all around the sample.

This result highlights that the artificial neural network faithfully distinguishes between the different phases based solely on local information, providing a useful method to calculate the topological invariant in systems without translational symmetry.
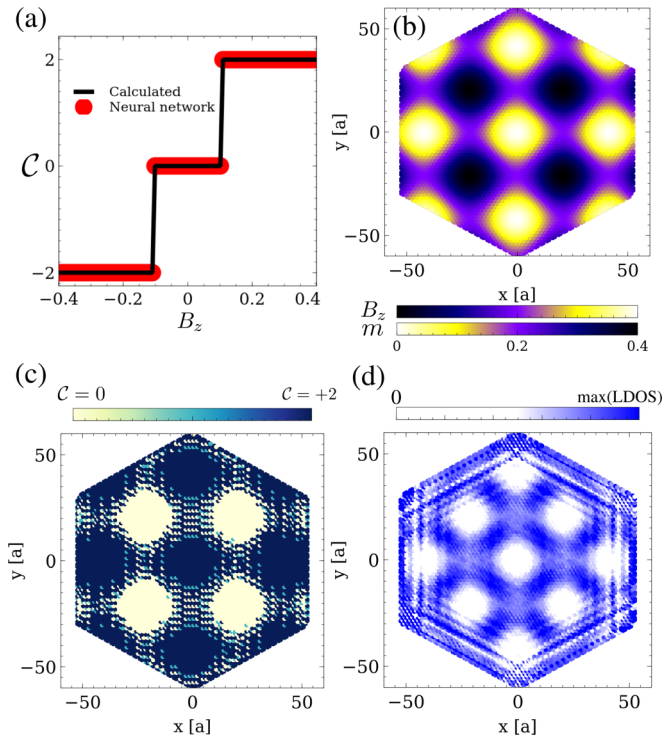
FIG. 6. (a) Comparison between the exact Chern number (black) and the prediction of the trained neural network (red) using as input the local density matrix. Once the accuracy of the network has been checked, we created a big graphene island with modulated mass and exchange term as shown in panel (b). The neural network is used to evaluate the topological invariant in each atom, yielding the result shown in panel (c). The boundary between different topological phases is expected to give rise to in-gap states, which is confirmed by calculating the in-gap spectral function as shown in panel (d).

IV. CONCLUSIONS

We have shown that an artificial neural network is capable of predicting the topological nature of different model Hamiltonians using as an input a local sector of the density matrix, i.e., evaluating solely *local properties*. Our procedure consisted of training an artificial neural network using as input the subspace of the density matrix corresponding to a local area of the sample, and as output the topological invariant that an analogous (pristine and translational invariant) Hamiltonian with the same effective parameters would have.

We applied this procedure to two well-known models, a 1D topological superconductor and a 2D topological insulator. In both cases we considered finite systems with a space dependent Hamiltonian that create regions with both topological and trivial character. By evaluating the network with local quantities for each Hamiltonian we showed that the different topological domains are accurately identified by the network, even when the inhomogeneous systems have Anderson-like disorder, proving that this methodology can be applied for disordered systems.

It is worth remarking that the training procedure is carried out for a specific model, and tested in that same model for different parameters, including local modulations in space. An open question is whether this methodology can be extended to cases with the same topological classes but different geometries. Finally, it is interesting to note that an analogous methodology could be applied to interacting systems, so that similar procedures could be exploited to identify quantum spin liquid states in two-dimensional spin systems.

[1] M. Z. Hasan and C. L. Kane, Rev. Mod. Phys. **82**, 3045 (2010).

[2] X.-L. Qi and S.-C. Zhang, Rev. Mod. Phys. **83**, 1057 (2011).

[3] L. Fu, Phys. Rev. Lett. **106**, 106802 (2011).

[4] M. Dzero, K. Sun, V. Galitski, and P. Coleman, Phys. Rev. Lett. **104**, 106408 (2010).

[5] S. Raghu, X.-L. Qi, C. Honerkamp, and S.-C. Zhang, Phys. Rev. Lett. **100**, 156401 (2008).

[6] A. A. Soluyanov and D. Vanderbilt, Phys. Rev. B **83**, 235401 (2011).

[7] D. Gresch, G. Autès, O. V. Yazyev, M. Troyer, D. Vanderbilt, B. A. Bernevig, and A. A. Soluyanov, Phys. Rev. B **95**, 075146 (2017).

[8] Q. Wu, S. Zhang, H.-F. Song, M. Troyer, and A. A. Soluyanov, Comput. Phys. Commun. **224**, 405 (2018).

[9] P. San-Jose, A. Gutiérrez-Rubio, M. Sturla, and F. Guinea, Phys. Rev. B **90**, 075428 (2014).

[10] J. Jung, E. Laksono, A. M. DaSilva, A. H. MacDonald, M. Mucha-Kruczyński, and S. Adam, Phys. Rev. B **96**, 085442 (2017).

[11] E. Wang, X. Lu, S. Ding, W. Yao, M. Yan, G. Wan, K. Deng, S. Wang, G. Chen, L. Ma *et al.*, Nat. Phys. **12**, 1111 (2016).

[12] J. D. Sanchez-Yamagishi, J. Y. Luo, A. F. Young, B. M. Hunt, K. Watanabe, T. Taniguchi, R. C. Ashoori, and P. Jarillo-Herrero, Nat. Nanotechnol. **12**, 118 (2017).

[13] A. F. Young, J. D. Sanchez-Yamagishi, B. Hunt, S. H. Choi, K. Watanabe, T. Taniguchi, R. C. Ashoori, and P. Jarillo-Herrero, Nature (London) **505**, 528 (2014).

[14] J. Alicea, Y. Oreg, G. Refael, F. von Oppen, and M. P. A. Fisher, Nat. Phys. **7**, 412 (2011).

[15] Ö. Gül, H. Zhang, J. D. S. Bommer, M. W. A. de Moor, D. Car, S. R. Plissard, E. P. A. M. Bakkers, A. Geresdi, K. Watanabe, T. Taniguchi *et al.*, Nat. Commun. **13**, 192 (2018).

[16] R. Bianco and R. Resta, Phys. Rev. B **84**, 241106 (2011).

[17] A. Marrazzo and R. Resta, Phys. Rev. B **95**, 121114 (2017).

[18] T. A. Loring, Ann. Phys. **356**, 383 (2015).

[19] N. P. Mitchell, L. M. Nash, D. Hexner, A. M. Turner, and W. T. M. Irvine, Nat. Phys. **1** (2018).

[20] I. C. Fulga, D. I. Pikulin, and T. A. Loring, Phys. Rev. Lett. **116**, 257002 (2016).

[21] E. P. L. van Nieuwenburg, Y.-H. Liu, and S. D. Huber, Nat. Phys. **13**, 435 (2017).

[22] K. Ch'ng, J. Carrasquilla, R. G. Melko, and E. Khatami, Phys. Rev. X **7**, 031038 (2017).

[23] T. Ohtsuki and T. Ohtsuki, J. Phys. Soc. Jpn. **85**, 123706 (2016).

[24] W. Hu, R. R. P. Singh, and R. T. Scalettar, Phys. Rev. E **95**, 062122 (2017).

[25] A. P. Bartók, S. De, C. Poelking, N. Bernstein, J. R. Kermode, G. Csányi, and M. Ceriotti, Sci. Adv. **3**, e1701816 (2017).

[26] M. Koch-Janusz and Z. Ringel, arXiv:1704.06279 (2017).

[27] G. Carleo and M. Troyer, Science **355**, 602 (2017).

[28] D.-L. Deng, X. Li, and S. D. Sarma, Phys. Rev. B **96**, 195145 (2017).

[29] D.-L. Deng, X. Li, and S. Das Sarma, Phys. Rev. X **7**, 021021 (2017).

[30] Y. Zhang and E.-A. Kim, Phys. Rev. Lett. **118**, 216401 (2017).

[31] Y. Nagai, H. Shen, Y. Qi, J. Liu, and L. Fu, Phys. Rev. B **96**, 161102 (2017).

[32] K. Ch'ng, N. Vazquez, and E. Khatami, Phys. Rev. E **97**, 013306 (2018).

[33] M. Rupp, A. Tkatchenko, K.-R. Müller, and O. A. von Lilienfeld, Phys. Rev. Lett. **108**, 058301 (2012).

[34] T. Gao, H. Li, W. Li, L. Li, C. Fang, H. Li, L. Hu, Y. Lu, and Z.-M. Su, J. Cheminf. **8**, 24 (2016).

[35] P. Zhang, H. Shen, and H. Zhai, Phys. Rev. Lett. **120**, 066401 (2018).

[36] N. Yoshioka, Y. Akagi, and H. Katsura, arXiv:1709.05790 (2017).

[37] A. Krizhevsky, I. Sutskever, and G. E. Hinton, in *Advances in Neural Information Processing Systems* (Curran Associates, Inc., USA 2012), Vol. 25, pp. 1097–1105.

[38] G. Dede and M. H. Sazlı, Digital Signal Processing **20**, 763 (2010).

[39] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, Proc. IEEE **86**, 2278 (1998).

[40] Y. Goldberg, arXiv:1510.00726v1 (2015).

[41] Y. Bengio, R. Ducharme, P. Vincent, and C. Janvin, J. Mach. Learn. Res. **3**, 1137 (2003).

[42] T. Schaul, J. Bayer, D. Wierstra, Y. Sun, M. Felder, F. Sehnke, T. Rückstieß, and J. Schmidhuber, J. Mach. Learn. Res. **11**, 743 (2010).

[43] R. Solomonoff, IRE Conv. Rec. **2**, 56 (1957).

[44] A. L. Samuel, IBM J. Res. Dev. **3**, 210 (1959).

[45] F. Rosenblatt, Psychol. Rev. **65**, 386 (1958).

[46] A. L. Hodgkin and A. F. Huxley, J. Physiol. **117**, 500 (1952).

[47] J. J. Hopfield, Proc. Natl. Acad. Sci. U.S.A. **79**, 2554 (1982).

[48] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, Nature (London) **323**, 533 (1986).

[49] A. Weiße, G. Wellein, A. Alvermann, and H. Fehske, Rev. Mod. Phys. **78**, 275 (2006).

[50] J. C. Budich and E. Ardonne, Phys. Rev. B **88**, 075419 (2013).

[51] R. M. Lutchyn, J. D. Sau, and S. Das Sarma, Phys. Rev. Lett. **105**, 077001 (2010).

[52] Y. Oreg, G. Refael, and F. von Oppen, Phys. Rev. Lett. **105**, 177002 (2010).

[53] V. Mourik, K. Zuo, S. M. Frolov, S. Plissard, E. Bakkers, and L. P. Kouwenhoven, Science **336**, 1003 (2012).

[54] T. D. Stanescu, R. M. Lutchyn, and S. Das Sarma, Phys. Rev. B **84**, 144522 (2011).

[55] R. M. Lutchyn, T. D. Stanescu, and S. Das Sarma, Phys. Rev. Lett. **106**, 127001 (2011).

[56] R. M. Lutchyn, E. P. A. M. Bakkers, L. P. Kouwenhoven, P. Krogstrup, C. M. Marcus, and Y. Oreg, arXiv:1707.04899 (2017).

[57] R. Aguado, Riv. Nuovo Cimento **40**, 523 (2017).

[58] Z. Qiao, S. A. Yang, W. Feng, W.-K. Tse, J. Ding, Y. Yao, J. Wang, and Q. Niu, Phys. Rev. B **82**, 161414 (2010).

[59] H. Min, J. E. Hill, N. A. Sinitsyn, B. R. Sahu, L. Kleinman, and A. H. MacDonald, Phys. Rev. B **74**, 165310 (2006).

[60] J. L. Lado, J. W. González, and J. Fernández-Rossier, Phys. Rev. B **88**, 035448 (2013).

[61] C. Tang, B. Cheng, M. Aldosary, Z. Wang, Z. Jiang, K. Watanabe, T. Taniguchi, M. Bockrath, and J. Shi, arXiv:1710.04179 (2017).

[62] Z. Wang, C. Tang, R. Sachs, Y. Barlas, and J. Shi, Phys. Rev. Lett. **114**, 016603 (2015).

[63] S. Su, Y. Barlas, J. Li, J. Shi, and R. K. Lake, Phys. Rev. B **95**, 075418 (2017).

[64] B. Huang, G. Clark, E. Navarro-Moratalla, D. R. Klein, R. Cheng, K. L. Seyler, D. Zhong, E. Schmidgall, M. A. McGuire, D. H. Cobden *et al.*, Nature (London) **546**, 270 (2017).

[65] J. Zhang, B. Zhao, T. Zhou, Y. Xue, C. Ma, and Z. Yang, Phys. Rev. B **97**, 085401 (2018).