Quantum topology identification with deep neural networks and quantum walks

Yurui Ming,¹ Chin-Teng Lin,¹ Stephen D. Bartlett,² and Wei-Wei Zhang² *

¹Centre for Artificial Intelligence, School of Software, University of Technology Sydney, Sydney, Australia
²Centre for Engineered Quantum Systems, School of Physics, The University of Sydney, Sydney, Australia

(Dated: December 3, 2018)

Abstract

Continuous-time quantum walks (CTQW) in topological quantum systems. The coherent dynamics of particles, with motion dependent on an internal degree of freedom such as spin, are described as quantum walks. Along with providing an advanced tool for building quantum algorithms, quantum walks also provide a platform to simulate and analyse complex physical systems. Discrete-time quantum walks have been shown to be closely related with the topology of the driving system. The experimental observations of particle localisation at the boundary between materials possessing different topological ordering and its robustness to the defects have been used to prove the existence of topologically protected edge modes. The moments of the probability distribution for the walker’s position after many steps is an experimental signature of a topological quantum phase transition in one-dimensional quantum walks. Continuous-time quantum walks with a two-dimensional spin-orbit lattice Hamiltonian can also reveal topological phase transitions, a fact supported by recent experiments. In such CTQW, the resulting density profile of an initially localized particle is expected to contain a wealth of information to identify the topological order of the underlying quantum system, provided one can extract this information efficiently.

Introduction

Topologically ordered materials may serve as a platform for new quantum technologies such as fault-tolerant quantum computers. To fulfill this promise, efficient and general methods are needed to discover and classify new topological phases of matter. We demonstrate that deep neural networks augmented with external memory can use the density profiles formed in quantum walks to efficiently identify properties of a topological phase as well as phase transitions. On a trial topological ordered model, our method’s accuracy of topological phase identification reaches 97%, and is shown to be robust to noise on the data. Our approach is generally applicable and may be used for topology identification with a variety of quantum materials.

Topological phases we consider are described by a parameterised Hamiltonian on a two-dimensional lattice model with spin-orbit coupling. Our method demonstrates high identification accuracy, greater than 97%, and is robust to noise on the input data. Our results provide a powerful tool for the efficient discovery and analysis of novel topological quantum systems, and therefore the design of robust quantum technologies.

Results

Continuous-time quantum walks in topological quantum systems. The coherent dynamics of particles, with motion dependent on an internal degree of freedom such as spin, are described as quantum walks. Along with providing an advanced tool for building quantum algorithms, quantum walks also provide a platform to simulate and analyse complex physical systems. Discrete-time quantum walks have been shown to be closely related with the topology of the driving system. The experimental observations of particle localisation at the boundary between materials possessing different topological ordering and its robustness to the defects have been used to prove the existence of topologically protected edge modes. The moments of the probability distribution for the walker’s position after many steps is an experimental signature of a topological quantum phase transition in one-dimensional quantum walks. Continuous-time quantum walks (CTQW) with a two-dimensional spin-orbit lattice Hamiltonian can also reveal topological phase transitions, a fact supported by recent experiments. In such CTQW, the resulting density profile of an initially localized particle is expected to contain a wealth of information to identify the topological order of the underlying quantum system, provided one can extract this information efficiently.

The topological phases we consider are described by a parameterised Hamiltonian on a two-dimensional lattice (599 × 599 in our simulation). Following Ref. [17], we use a continuous-time quantum walk (CTQW) for a initially localised single particle under this Hamiltonian, where the behaviour of the distribution of the quantum state after long times provides a signature of the topological phase. We will investigate the use of both the particle’s spatial distribution $p(x, y)$ as well as its momentum space distribution $p(k_x, k_y)$, marginal-
Deep learning has achieved breakthroughs across many presentations are computed in terms of less abstract ones. The adaptability in this regard. Each concept is defined in a nested hierarchy of concepts, providing great capability and set of machine learning which represents the data as a underlying characteristics of a physical system even without prior human knowledge. Inspired by the hierarchical biostructures in vision systems, deep neural networks (DNN) can automatically extract the most suitable representations from input data and make accurate predictions. Generally speaking, during the end-to-end learning process, the representations of data will emerge rather than being discovered or manually crafted.

We will apply DNN to the problem of topological identification by providing the network with the density profiles from a CTQW as input. As described above, the density profiles contain a wealth of information about the topological phase of the system, but identifying which features are important is challenging, especially for higher order phases. A DNN with external memory has the capacity to solve complex, structural tasks that are inaccessible to stand-alone neural networks, and has shown the ability to answer synthetic questions designed to emulate reasoning and inference problems. The architecture of our deep neural network is shown in Fig. 2, which consists of multiple computation blocks (CB) and fully connected layers (computation network), as well as an external memory coupled to the last convolutional layer (memory network). The computation network is of a supervised-learning paradigm and the memory network is of an unsupervised-learning paradigm. Both are jointly trained during the process.

Our experiment consists of four steps: data preparation, neural network training, validation, and testing. The data preparation stage is based on numerically simulations of CTQW with different Hamiltonian parameters, and is described in the Methods. The data corresponding to different topological phases is randomised and split into three sets with the ratio 0.8 : 0.1 : 0.1 for training, validation and testing respectively. The prepared data is reused three times to evaluate the network. As the performance indicator for the corresponding prepared data, the accuracy in our results is the average over three independent randomisation sets.

We illustrate the outcome of our experiments using the principal component analysis (PCA) of memory, a t-distributed stochastic neighbour embedding (t-SNE) of the computation network output, and the statistical accuracy of the test. Both the PCA and the t-SNE are visualisation results, and the accuracy is a statistical evaluation. The t-SNE shows the topological classification of input data corresponding to different Chern numbers. The PCA demonstrates how the input data is clustered according to its correlation by self-organisation, which reveals the different topological phases of the input data. The accuracy shows the identification accuracy calculated from the test data.

The PCA and t-SNE based on the data—the density profiles in momentum and position space—are shown in Fig. 3, where the DNN identification forms separated clusters associated with the five topological phases of our model Hamiltonian system. For the momentum space data, the identification clearly reveals five clusters corre-
sponding to each of the distinct topological phases of the
Hamiltonian. For the position space data, only four clusters
are identified; the topological phases corresponding to
Chern numbers $C = \pm 2$ are not distinguished based
on this data.

The statistical accuracy of our test, i.e., the ratio be-
tween the number of testing samples classified into cor-
correct topological phases and the total number of testing
samples, is shown in Table I. We see that, when based
on momentum space distributions, we obtain a very high
accuracy for data covering both the whole phase diagram
region and a restriction to the region around the phase
transition (97% and 91.7% respectively). Position space
distributions lead to identification with relatively lower
accuracy for the same regions (70.3% and 71.4% respec-
tively), especially for the case $C = 2$. The unbalanced
accuracies for $C = -2$ and $C = 2$ cases are likely due to
the relatively smaller region for $|C| = 2$ in our phase di-
agram of Fig. 1. The relatively low accuracy for data in
position space indicated the data in position space is less
directly connected with the Chern number of the Hamil-
tonian as defined in Eq. 3. As the density profiles in
position and momentum spaces are directly related to
each other via the Fourier Transformation, and so funda-
mentally carry the same information, we conjecture that
to obtain a higher accuracy with position density pro-
fils as input data a much larger size of DNN is required
to compensate the Fourier Transformation of the infor-
data augmentation will also help as discussed below.

Noisy data as input. Quantum walks on engineered
topological quantum materials have been realised in dif-
f erent physical platforms including photonics systems [14]
[19] [22] [39] and cold atoms [13], amongst others. For our
method to be useful on experimental data, it must be
robust to noise. Here, we test the performance of our
method with noisy input data for our trained DNN. We
add Gaussian noise to our simulated data, at a level com-
parable with current experimental techniques in optical
systems [22] [39] and cold atoms systems [10] [18] [45]; de-
tails are discussed in the Methods.

In these tests, the accuracy statistics for topological
phase identification shows limited degradation as indi-
cated in Table I. We can see that the accuracy only drops
0.020 on average with momentum density profiles, which
could potentially be offset by increasing the size of the
network. One unexpected outcome is that, with noise on
the density profiles in position space, the accuracy is
0.008 and 0.030 higher that the accuracy obtained with
the data without noise. The performance increase with
noise data in position space is potentially due to the ef-
fect of a data augmentation for training deep neural net-
work [12] [44]. The Gaussian noise introduced in our data
increases the margin of the decision boundary, given an
accuracy of around 70%.

Conclusion and outlook.

We have demonstrated a universal automatic method
for the identification of distinct topological phases of
quantum materials. Our simulated experimental results
show that the combination of the particle’s density profile
from a CTQW and DNN augmented with external mem-
ory is a reliable and efficient method to identify topo-
logical phases and phase transitions in our trial system, even
for the high order $C = \pm 2$ and noisy data. Our approach
is generally applicable and may be used for the identi-
fication of topological phases with a variety of quantum
materials.

Methods

Here we present the trial topological Hamiltonian sys-
tem, and describe the generation of a particle’s density
profile as used as the input data for our DNN. We also
provide the details of the architecture of our DNN.

The topological system in our simulated experiments.
The two-dimensional spin-orbit lattice Hamiltonian we
consider here is [17] [18] [39] [40]

$$
H = \sum_{x,y} \left[ \frac{1}{2} \left( c_{x,y}^\dagger \sigma_3 c_{x,y} + \text{h.c.} \right) + t_{1x} \sigma_1 - i t_2 \sigma_3 \right] c_{x+1,y} + t_{1y} \sigma_2 + i t_2 \sigma_3 c_{x,y} \\
+ \sum_{k_x,k_y} \tilde{h} \cdot \sigma \left| k_x, k_y \right> \left< k_x, k_y \right|
$$

(1)

using $\{m, t_{1x}, t_{1y}, t_2, t_3\}$ as the coupling parameters, $i \in \{1, 2, 3\}$, $\sigma = \{\sigma_1, \sigma_2, \sigma_3\}$ as the Pauli operators and $\tilde{h} = (h_1, h_2, h_3)$. The last line of Eq. 1 is obtained by using

translation invariance and the Fourier Transformation

$$
|k_x\rangle = \frac{1}{\sqrt{2\pi}} \sum_x e^{-i k_x x} |x\rangle, \quad |k_y\rangle = \frac{1}{\sqrt{2\pi}} \sum_y e^{-i k_y y} |y\rangle,
$$

the $2 \times 2$ block diagnosis Hamiltonian in momentum space

$$
\tilde{h} \cdot \sigma = 2t_{1x} \cos k_x \sigma_1 + 2t_{1y} \cos k_y \sigma_2 \\
+ \{m + 2t_3 \left( \sin k_x + 2 \sin k_y \right) + 2t_2 \cos \left( k_x + k_y \right) \} \sigma_3.
$$

(2)

This Hamiltonian supports the topological phases with
$C \in \{0, \pm 1, \pm 2\}$ by varying the coupling parameters $m, t_3$ while fixing $t_{1x} = t_{1y} = 1, t_2 = 5$, with the definition of
Chern number as

$$
C = \frac{1}{4\pi} \int_{BZ} d^2k \ \tilde{h} \cdot \left( \partial_{k_x} \tilde{h} \times \partial_{k_y} \tilde{h} \right)
$$

(3)

with $\tilde{h} = \tilde{h}/|\tilde{h}|$. The different topological phases labelled by Chern number $C$, as a function of Hamiltonian
parameters, is shown in Fig. 1.

The formation of particle’s density profile in both mo-
momentum and position spaces. In CTQW evolutions, a
particle with spin up, initially localised in the centre of
a two-dimensional lattice in position space, spreads out
and gradually occupies a larger area of the lattice. Equiv-
ally, the particle is initially uniformly distributed in
momentum space and during the evolution the particle’s
components at every momenta oscillates between spin
up and spin down components. The particle’s probability distributions in both position and momentum spaces form a certain pattern which is closely related with the Hamiltonian.

At evolution time $t$, the state of the particle initially localised at the centre of two-dimensional lattice is (setting $\hbar = 1$)

$$\psi(t) = \sum_k (\alpha_{k\uparrow} |\uparrow\rangle + \alpha_{k\downarrow} |\downarrow\rangle) |k\rangle$$

$$= \sum_k \left( \frac{\hbar s\sin(E_k t)}{E_k} - \cos(E_k t) \right) |k\rangle$$

(4)

where $E_k = \sqrt{h_1^2 + h_2^2 + h_3^2} \neq 0$ is the eigenenergy of system’s Hamiltonian. When $E_k = 0$ we have $\alpha_{k\uparrow} = 1$ and $\alpha_{k\downarrow} = 0$, which is the case at Dirac point while the system is under topological phase transition. The particle’s state represented in position space is the Fourier transform of the corresponding spin components.

From the expression of Eq. (4) for particle’s state at time $t$, the amplitude and the relative phase of both spin up and spin down components are closely related with the energy $E_k$ and sensitive to the band gap of the system which is min$\{2E_k\}$ as discussed in Ref. [17] [18]. The topological phase of the system characterised with Chern number is revealed by the band structure of the system. Therefore, the particle’s density profile is a competitive candidate for the topological detection, even for higher order phases.

Here, we generate two sets of density profiles. One is in momentum space and the other is in position space. For the training of the neural network, we decompose the complex values of both spin up and spin down components into two real values and map the amplitude and relative phase matrices into image representation. With this process, the input data set consists of the set of spatial or momentum distributions for the particle’s final states.

Dataset generation for our deep neural network identifying the topology of quantum matters. Our system supports topological phases with $C = \{0, \pm 1, \pm 2\}$ as described above. The diagram showing the distribution of Chern number $C$ with respect to $m, t_3$ and fixed $t_{1x} = t_{1y} = 1, t_2 = 5$ is shown in Fig. 1, where the shaded area represents the parameter area for the dataset labeled as “whole” and the dotted area represents the parameter area for the dataset labeled as “transition” in our tables. The dataset for $C = 2$ is generated with the same $m, t_3, t_{1x}, t_2$ as $C = -2$, but with $t_{1y} = -1$. The sizes of our dataset generated for the whole phase diagram are $\{1449, 1478, 1486, 1488, 1449\}$ and for the phase transition area of the diagram are $\{2542, 1262, 1474, 1564, 2542\}$ corresponding to $C = \{-2, -1, 0, 1, 2\}$ respectively. The evolution time we chose to be a time which enables the particle’s density profile occupy 80% of the lattice area.

The method to add the noise to our density profiles are different for the data collected in different measurement spaces, i.e. momentum or position. The experimental momentum data measurement can be implemented in cold atom systems as in Ref. [10] [11], where the noise in the data are the shot-noise and Gaussian white noise. The standard deviation of Gaussian noise is set to be 0.02 in our simulated data, which is a reasonable estimation for current technology based on the error bar ranges in Ref. [15]. The experimental position data measurement can be implemented in cold atom systems as in Ref. [13] and photonics systems as in Ref. [22] [24] by encoding the position of a walker in either time-bins or spatial modes. The noise in position data includes shot-noise and device noise resulting in the uncertainty in both relative phase and amplitude of the state, which is realised by the convolution between the perfect state and the point-spread function (PSF) of the system. In our noisy data, the PSF we used is a Gaussian with 0 as its mean and 2 as the standard deviation which is also within current experimental techniques level [17] [18].

The configurations of our deep neural network for topological phase identification of quantum systems. We use a deep neural network coupled with an external memory for identification of topological phases from the distributions from CTQWs. We take advantage of the most of up-to-date techniques for our computation network design. For the memory network, the simplification of memory operations is achieved by using a self-organising map (SOM), which is endowed with effective memory addressing and allocation mechanisms. A hybrid learning approach is devised to optimise the network for obtaining promising results.

The detailed architecture and the configuration of our network is illustrated in Fig. 2 and Table I. There are 6 computation blocks (2 with size $8 \times 8$, 2 with size $16 \times 16$ and 2 with size $32 \times 32$), 2 fully connected layers and an external memory. During the training process, the learning rates (LR) for computation network and memory network are 0.0001 and 0.4 respectively. The batch size is set as 64 and the network is trained 1000 iterations. The learning rate decay factor in our computation network is 0.9 for every 100 iterations. The time constant for SOM is the number of iterations divided by the natural logarithm of initial radius (128 in our experiment). The labels for memory clusters are probed by tracking the corresponding coordinates of a few typical data from different topological phases.

Our experiments run on a GPU cluster with three nodes. Each node is with two Intel CPUs of model E5-2680 and 128GB physical memory. For computing acceleration, each CPU manages a separate PCIe slot in which an NVIDIA Quadro P5000 GPU card with 16GB on-board memory installed.
Acknowledgments

This work is supported by the Australian Research Council via the Centre of Excellence in Engineered Quantum Systems project number CE170100009 and Discovery Project numbers DP170103073, DP180100670 and DP180100656. The authors acknowledge discussions about noisy experimental data with Wei Sun, Chao Chen, Yu He, and Steven Flammia, and comments from Robin Harper and John Manion. The authors acknowledge the University of Sydney and University of Technology Sydney for providing HPC resources that have contributed to the research results reported in this paper.

Author Contributions

YM and CL designed and performed experiments. WZ and SB provided theoretical support. WZ prepared the training data. All authors contribute to write the paper.
FIG. 2: Network architecture: the computation network is constructed from six computation blocks of a supervised learning paradigm, the memory network is of an unsupervised learning paradigm. “SOM” represents for self-organising map, “conv” represents for convolution. The input of our DNN is the density profiles and the outputs are PCA, t-SNE and statistical accuracy.

FIG. 3: PCA of memory based on momentum (a) and position (c) density profiles, which illustrate the self-organised clusters formed in memory during our training process and shows the clustering of the input data. The size of PCA is same as the size of the memory in our DNN (256 × 256). The RGB colour is obtained by projecting the 32 dimensional vector of each memory pixel into a 3 channel colour representation. In our experiments, the scree plots indicate that the first three components explain around 80% of variance, and an “elbow”, the cutting-off point, appears at the third principal component. This justifies our choice of first three principal components in our experiments. (b) and (d) are classification visualisations of momentum and position samples via t-SNE, which is a projection from the 5 dimensional DNN output vector into the location indices of a 2 dimensional space.
TABLE I: The statistical accuracy for the topological identification using our DNN. The accuracy is obtained by averaging over three randomised data sets, where every data set is trained three times.

| Phase Diagram Area | Density Profile Data | Measurement Domain | C                | Overall |
|-------------------|---------------------|--------------------|-----------------|---------|
|                   |                     |                    | -2   | -1  | 0   | 1   | 2   |
| Whole             | Momentum            |                     | 0.979 | 0.954 | 0.965 | 0.952 | 1.000 | 0.970 |
|                   | Position            |                     | 0.726 | 0.961 | 0.692 | 0.931 | 0.195 | 0.703 |
| Transition        | Momentum            |                     | 0.994 | 0.869 | 0.965 | 0.647 | 1.000 | 0.917 |
|                   | Position            |                     | 0.94  | 0.672 | 0.748 | 0.587 | 0.566 | 0.714 |

The statistical accuracy with noisy input data

| Phase Diagram Area | Density Profile Data | Measurement Domain | C                | Overall |
|-------------------|---------------------|--------------------|-----------------|---------|
|                   |                     |                    | -2   | -1  | 0   | 1   | 2   |
| Whole             | Momentum            |                     | 0.952 | 0.960 | 0.955 | 0.919 | 1.000 | 0.957 |
|                   | Position            |                     | 0.290 | 0.955 | 0.832 | 0.990 | 0.554 | 0.711 |
| Transition        | Momentum            |                     | 0.959 | 0.880 | 0.882 | 0.625 | 1.000 | 0.891 |
|                   | Position            |                     | 0.924 | 0.843 | 0.805 | 0.648 | 0.541 | 0.744 |

TABLE II: DNN architecture configuration with LR as learning rate.

| Computation Network with LR = 0.0001 |
|-------------------------------------|
| Block | Layer | Filter Size | Activation | Padding | Repetition |
|-------|-------|-------------|------------|---------|------------|
| 1st   | AvgPool | – (2,2) | – | Valid | 2 |
|       | Conv2D | 8 (5,5) | – | Valid |
|       | BatchNorm | – | – | – |
|       | SeparableConv2D | 8 (5,5) | ELU | Same |
| 2nd   | AvgPool | – (2,2) | – | Valid | 2 |
|       | Conv2D | 16 (5,5) | – | Valid |
|       | BatchNorm | – | – | – |
|       | SeparableConv2D | 16 (5,5) | ELU | Same |
| 3rd   | AvgPool | – (2,2) | – | Valid | 2 |
|       | Conv2D | 32 (5,5) | – | Valid |
|       | BatchNorm | – | – | – |
|       | SeparableConv2D | 32 (5,5) | ELU | Same |
| 4th   | Linear | – 256 | ReLU | – | 1 |
| 5th   | Linear | – 5 | Softmax | – | 1 |

Memory Network with LR = 0.4

| Height | Width | Element Size | Decay Factor of LR | Initial Radius |
|--------|-------|--------------|--------------------|---------------|
| 256    | 256   | 32           | 0.9               | 128           |