



Deep Learning the Quantum Phase Transitions in Random Two-Dimensional Electron Systems

Tomoki Ohtsuki^{1*} and Tomi Ohtsuki^{2†}

¹NTT DATA Mathematical Systems Inc., Shinjuku, Tokyo 160-0016, Japan

²Physics Division, Sophia University, Chiyoda, Tokyo 102-8554, Japan

(Received October 3, 2016; accepted October 28, 2016; published online November 18, 2016)

Random electron systems show rich phases such as Anderson insulator, diffusive metal, quantum Hall and quantum anomalous Hall insulators, Weyl semimetal, as well as strong/weak topological insulators. Eigenfunctions of each matter phase have specific features, but owing to the random nature of systems, determining the matter phase from eigenfunctions is difficult. Here, we propose the deep learning algorithm to capture the features of eigenfunctions. Localization–delocalization transition, as well as disordered Chern insulator–Anderson insulator transition, is discussed.

Introduction—More than half a century has passed since the discovery of Anderson localization,¹⁾ and the random electron systems continue to attract theoretical as well as experimental interest. Symmetry classification of topological insulators^{2–5)} based on the universality classes of random noninteracting electron systems^{6,7)} gives rise to a fundamental question: can we distinguish the random topological insulator from Anderson insulators? Note that topological numbers are usually defined in the randomness free systems via the integration of the Berry curvature of Bloch function over the Brillouin zone, although topological numbers in random systems have recently been proposed.^{8,9)}

Determining the phase diagram and the critical exponents requires large-scale numerical simulation combined with detailed finite size scaling analyses.^{10–13)} This is because, owing to large fluctuations of wavefunction amplitudes, it is almost impossible to judge whether the eigenfunction obtained by diagonalizing small systems is localized or delocalized, or whether the eigenfunction is a chiral/helical edge state of a topological insulator. In fact, it often happens that eigenfunctions in the localized phase seem less localized than those in the delocalized phase [see Figs. 1(b) and 1(c) for example].

Recently, there has been great progress on image recognition algorithms¹⁴⁾ based on deep machine learning.^{15,16)} Machine learning has recently been applied to several problems of condensed matter physics such as Ising and spin ice models^{17,18)} and strongly correlated systems.^{19–24)}

In this letter, we test the image recognition algorithm to determine whether the eigenfunctions for relatively small systems are localized/delocalized, and topological/nontopological. As examples, we test two types of two-dimensional (2D) quantum phase transitions: Anderson-type localization–delocalization transition in symplectic systems, and disordered Chern insulator to Anderson insulator transition in unitary systems.

Distinguishing localized states from delocalized ones—We start with a 2D symplectic system, which is realized in the presence of spin–orbit scattering. We use the SU(2) Hamiltonian²⁵⁾ that describes the 2D electron on a square lattice with nearest-neighbor hopping,

$$H = \sum_{i,\sigma} \epsilon_i c_{i,\sigma}^\dagger c_{i,\sigma} - \sum_{\langle i,j \rangle, \sigma, \sigma'} R(i,j)_{\sigma, \sigma'} c_{i,\sigma}^\dagger c_{j, \sigma'}, \quad (1)$$

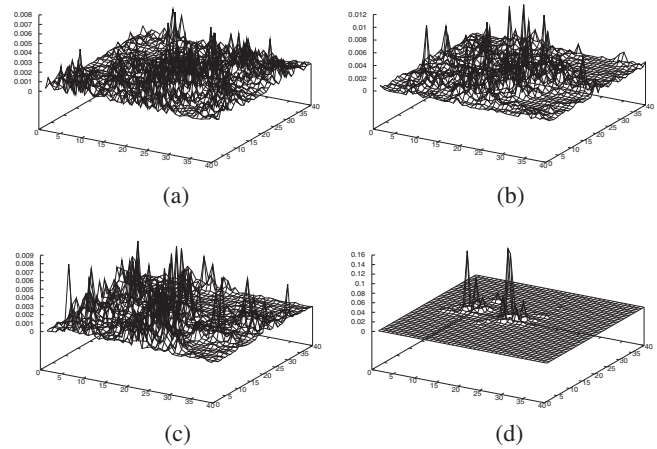


Fig. 1. Examples of eigenfunction modulus squared $|\psi(x, y)|^2$ for (a) $W = 0.124 \approx W_c^{\text{SU2}}/50$, (b) $W = 5.58 \approx 0.9W_c^{\text{SU2}}$, (c) $W = 6.82 \approx 1.1W_c^{\text{SU2}}$, and (d) $W = 12.4 \approx 2W_c^{\text{SU2}}$. Peak positions are shifted to the center of the systems.

where $c_{i,\sigma}^\dagger$ ($c_{i,\sigma}$) denotes the creation (annihilation) operator of an electron at site $i = (x, y)$ with spin σ , and ϵ_i denotes the random potential at site i . We assume a box distribution with each ϵ_i uniformly and independently distributed on the interval $[-W/2, W/2]$. The modulus of the transfer energy is taken to be the energy unit. $R(i, j)$ is an SU(2) matrix,

$$R(i, j) = \begin{pmatrix} e^{i\alpha_{i,j}} \cos \beta_{i,j} & e^{i\gamma_{i,j}} \sin \beta_{i,j} \\ -e^{-i\gamma_{i,j}} \sin \beta_{i,j} & e^{-i\alpha_{i,j}} \cos \beta_{i,j} \end{pmatrix}, \quad (2)$$

with α and γ uniformly distributed in the range $[0, 2\pi)$. The probability density $P(\beta)$ is

$$P(\beta) = \begin{cases} \sin(2\beta) & 0 \leq \beta \leq \pi/2 \\ 0 & \text{otherwise} \end{cases}. \quad (3)$$

Examples of the eigenfunctions in delocalized [Figs. 1(a) and 1(b)] and localized phases [Figs. 1(c) and 1(d)] are shown in Fig. 1.

For $E = 0$ (band center), from the finite size scaling analyses of the quasi-1D localization length,^{25,26)} it is known that the states are delocalized when $W < W_c^{\text{SU2}}$ (≈ 6.20), while they are localized when $W > W_c^{\text{SU2}}$. We impose periodic boundary conditions in x - and y -directions, and diagonalize systems of 40×40 . From the resulting 3200 eigenstates with Kramers degeneracy, we pick up the 1600th eigenstate (i.e., a state close to the band center). For



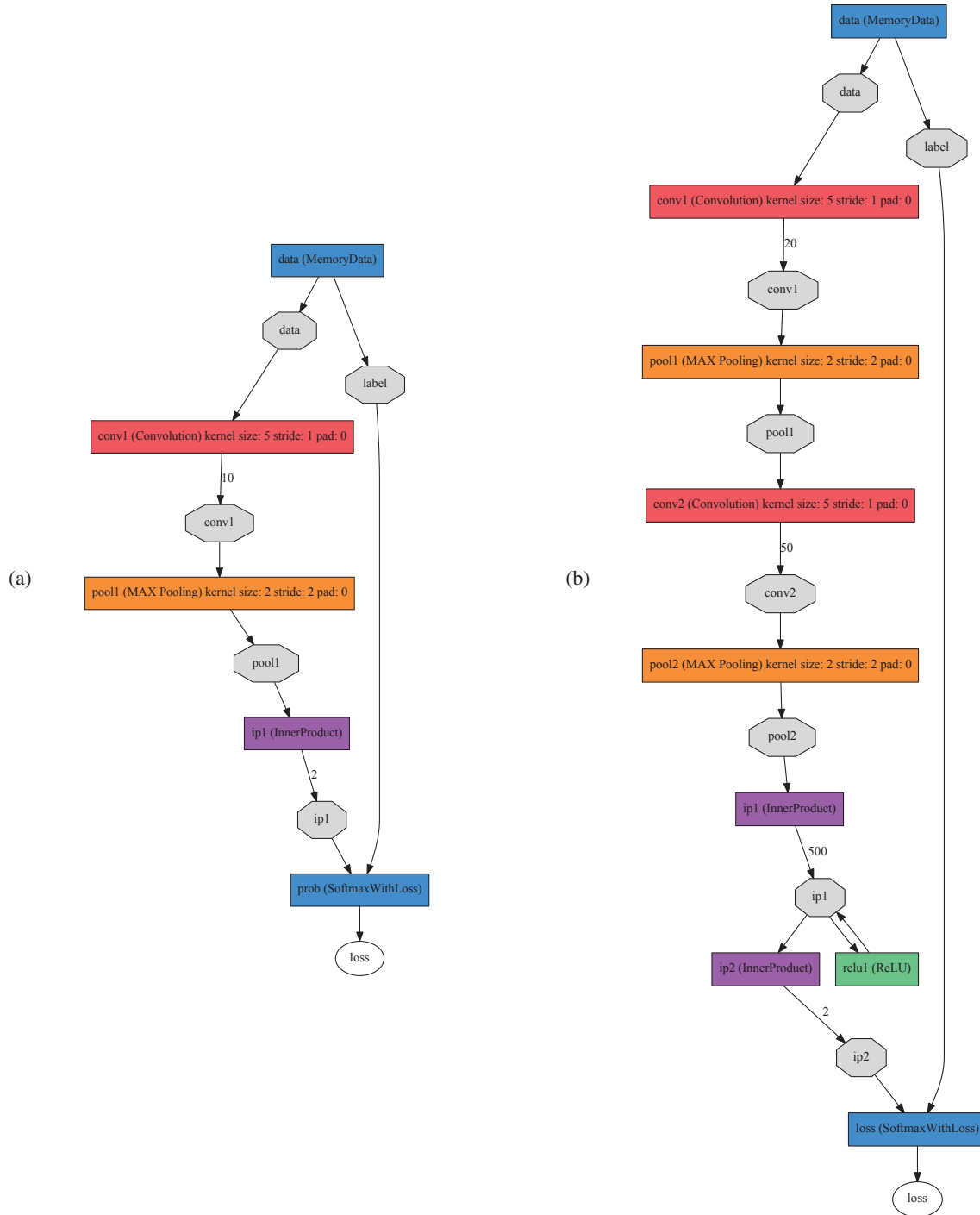


Fig. 2. (Color) Network architectures used in this work. (a) A simple two-weight-layer CNN, which consists of convolution and max pooling, followed by dense linear transformation. (b) LeNet-like architecture with ReLU activation.

simplicity, the maximum modulus of the eigenfunction is shifted to the center of the system. Changing W and the seed of the random number stream (Intel MKL MT2023), we prepare 2000 samples of states, i.e., 1000 for $W < W_c^{\text{SU}2}$ and 1000 for $W > W_c^{\text{SU}2}$. We then teach the machine whether the states belong to the localized (delocalized) phase.

For our network architecture, we consider two types of simple convolutional neural network (CNN), which output two real numbers, i.e., probabilities for each phase, given 40×40 input eigenfunction. The first one is a very simple network with two weight layers, which first convolves the input with a 5×5 filter with stride 1 to 10 channels, then

applies max pooling with a kernel size of 2×2 and stride 2, and finally performs fully connected linear transformation to output the learned probabilities. The loss function can then be defined by the cross entropy of probabilities and the localized/delocalized labels. The second, rather deep one with four weight layers is a variant of LeNet²⁷⁾ included in Caffe²⁸⁾ (with the input size changed to 40×40), which utilizes rectified linear unit (ReLU) as its activation function. See Fig. 2 for illustration and detailed parameters. The network weight parameters (to be trained) are sampled from gaussian distribution, the scale of which is determined by the number of input and output dimensions,²⁹⁾ except for the first

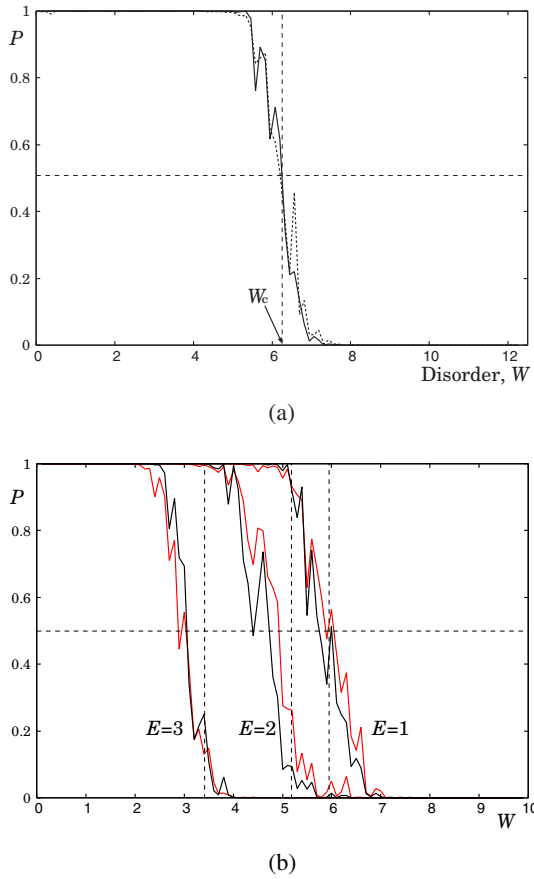


Fig. 3. (Color online) Probability of eigenfunction to be judged delocalized as a function of disorder W . Averages over five samples are taken. (a) Band Center $E = 0$. Critical disorder $W_c^{\text{SU}2} \approx 6.20$, as well as 50% probability, is indicated by the dashed lines. The dotted line is for two-weight-layer network [Fig. 2(a)], while the solid one is for four-weight-layer network [Fig. 2(b)]. (b) For $E = 1.0, 2.0$, and 3.0 . The red line is for two-weight-layer network, while the black line is for four-weight-layer network. The values of $W_c^{\text{SU}2}$ for $E = 1.0, 2.0$, and 3.0 estimated via the finite size scaling of the localization length²⁶⁾ are 5.953, 5.165, and 3.394, respectively, which are indicated by the vertical dashed lines.

convolution layer connected to the raw input: since we are dealing with eigenfunctions, whose typical values at each lattice site are much smaller than those of gray-scale images, we have manually chosen the weight initialization scale to be 100, which worked better in practice for the two networks. As the stochastic gradient descent solver, we have used the RMSProp solver³⁰⁾ with the parameters in the Caffe MNIST example (which is contained as `examples/mnist/lenet_solver_rmsprop.prototxt` in the Caffe source). Before the training, we always partition the training data into 90 and 10%, and use the latter as the validation set during the training. The solver performs enough iterations so that the validation error becomes stationary. We have used a workstation: Intel Xeon E5-1620 v4, single CPU with 4 cores with GPU Quadro K420 and GPGPU TESLA K40 running on Linux CentOS 6.8.

We then test 5 sets of ensemble, each consisting of 100 eigenstates, and let the machine judge whether the states are localized or not. The resulting probability for eigenfunction to be delocalized, P , is shown in Fig. 3(a).

We then apply the results of the learning around $E = 0$ to judge whether the states around $E = 1.0, 2.0$, and 3.0 are delocalized. Results are shown in Fig. 3(b), in which we

observe that, with increasing E , that is, as we move from band center to band edge, the electron begins to be localized with a smaller strength of the disorder W , qualitatively consistent with the finite size scaling analysis.²⁶⁾ There seems to be, however, a systematic deviation of the 50% criterion of localization–delocalization transition and the actual critical point with increasing E . This may be due to the appearance of bound states near the band edge, which is absent in the machine learning around $E = 0$. We have further applied the results of SU(2) model machine learning for the Ando model,³¹⁾ and verified that once the machine learns the eigenfunction features in certain systems, it can be applied to other systems belonging to the same class of quantum phase transition (see Supplemental Material for detail³²⁾).

Distinguishing topological edge states from non-topological ones— We next study the topological Chern insulator to nontopological Anderson insulator transition.^{33–35)} We use a spinless two-orbital tight-binding model on a square lattice, which consists of s -orbital and $p \equiv p_x + ip_y$ orbital,³⁶⁾

$$H = \sum_{\mathbf{x}} ((\epsilon_s + v_s(\mathbf{x}))c_{\mathbf{x},s}^\dagger c_{\mathbf{x},s} + (\epsilon_p + v_p(\mathbf{x}))c_{\mathbf{x},p}^\dagger c_{\mathbf{x},p}) + \sum_{\mathbf{x}} \left(- \sum_{\mu=x,y} (t_s c_{\mathbf{x}+\mathbf{e}_\mu,s}^\dagger c_{\mathbf{x},s} - t_p c_{\mathbf{x}+\mathbf{e}_\mu,p}^\dagger c_{\mathbf{x},p}) + t_{sp}(c_{\mathbf{x}+\mathbf{e}_x,p}^\dagger - c_{\mathbf{x}-\mathbf{e}_x,p}^\dagger)c_{\mathbf{x},s} - it_{sp}(c_{\mathbf{x}+\mathbf{e}_y,p}^\dagger - c_{\mathbf{x}-\mathbf{e}_y,p}^\dagger)c_{\mathbf{x},s} + \text{h.c.} \right),$$

where ϵ_s , $v_s(\mathbf{x})$, ϵ_p , and $v_p(\mathbf{x})$ denote atomic energy and disorder potential for the s - and p -orbitals, respectively. Both $v_s(\mathbf{x})$ and $v_p(\mathbf{x})$ are uniformly distributed within $[-W/2, W/2]$ with identical and independent probability distribution. t_s , t_p , and t_{sp} are transfer integrals between neighboring s -orbitals, p -orbitals, and that between s - and p -orbitals, respectively.

In the absence of disorder, the system is a Chern insulator when the band inversion condition is satisfied: $0 < |\epsilon_s - \epsilon_p| < 4(t_s + t_p)$. We set $\epsilon_s - \epsilon_p = -2(t_s + t_p)$, $\epsilon_s = -\epsilon_p < 0$, and $t_s = t_p > 0$ so that this condition is satisfied, and set $t_{sp} = 4t_s/3$. The energy unit is set to $4t_s$. A bulk band gap appears in $|E| < E_g = 0.5$ where chiral edge states exist.

For $E = 0$, the system remains as a Chern insulator for $W < W_c^{\text{CI}} \approx 3.2$,³⁴⁾ while it is an Anderson insulator for $W > W_c^{\text{CI}}$. [Unfortunately, the estimate of W_c^{CI} is less precise than the SU(2) model.] We impose fixed boundary conditions in the x - and y -directions, so that the edge states appear if the system is a topological insulator.

We diagonalize square systems of 40×40 sites, and from the resulting 3200 eigenstates, we pick up the 1600th eigenstate. Examples of the eigenfunctions in topological Chern [Figs. 4(a) and 4(b)] and nontopological Anderson insulators [Figs. 4(c) and 4(d)] are shown in Fig. 4. As shown in Fig. 4, it is difficult to judge whether the state is an edge state or not when W is close to W_c^{CI} : see, for example, $W = 0.9W_c^{\text{CI}}$ [Fig. 4(b), Chern insulator phase] and $W = 1.1W_c^{\text{CI}}$ [Fig. 4(c), Anderson insulator phase]. In fact, learning 1000 samples for each phase gives 93% validation accuracy for four-weight-layer network compared with 98% or more as in the SU(2) model. The difficulty may be due to the fixed boundary condition where shifting the locus of the

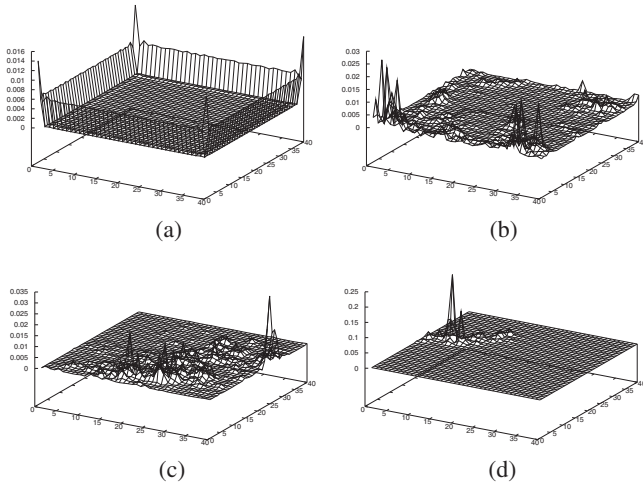


Fig. 4. Eigenfunction modulus squared $|\psi(x, y)|^2$ for (a) $W = 0.064 \approx W_c^{\text{CI}}/50$, (b) $W = 2.88 \approx 0.9W_c^{\text{CI}}$, (c) $W = 3.52 \approx 1.1W_c^{\text{CI}}$, and (d) $W = 6.4 \approx 2W_c^{\text{CI}}$.

maximum of the eigenfunction amplitude is not allowed. Another reason for difficulty is that the bulk of the systems are localized in both topological and nontopological regions. To overcome these difficulties, we increased the number of samples: 27000 samples belonging to the topological phase, and 27000 to the nontopological phase. We have also increased the number of hidden units to be 32 for the first convolution layer (“conv1” in Fig. 2), 128 for the second (“conv2”), and 512 for the hidden dense connection layer (“ip1”).

In Fig. 5(a), we plot the probability of the eigenfunction to be judged topological. A new ensemble of eigenfunctions with different random number sequences has been prepared to test this method. As in the case of delocalization–localization transition, the probability fluctuates near the critical point and vanishes in the nontopological region. The validation accuracy is 90% for the case of two layers of network (dotted line), and 97% for four layers of network (solid line), which demonstrates clearly that a deeper network exhibits better performance.

We next apply the result of the deep learning around $E = 0$ to judge the states in the bulk band gap region at zero disorder, $|E| < E_g = 0.5$. We diagonalize a system for $W = 1 < W_c^{\text{CI}}$ and $W = 6 > W_c^{\text{CI}}$, take all the eigenstates with $|E| < E_g$, and let the machine judge them. Figure 5(b) shows that topological edge states other than $E = 0$ are also well distinguished from nontopological ones based on the learning around $E = 0$.

Concluding remarks—In this paper, we focused on 2D random electron systems. We have demonstrated the validity of deep learning for distinguishing various random electron states in quantum phase transitions. For strong enough and weak enough randomness, the precision of judgement is 0.99999..., while in the critical regime, the judgement becomes less accurate. This region is related to the critical region where the characteristic length scale ξ is comparable to or longer than the system size L . That is, the probability P for the eigenfunction to be judged delocalized/topological obeys the scaling law, $P(W, L) = f[(W - W_c)L^{1/\nu}]$, although determining the exponent ν is beyond the scope of this letter.

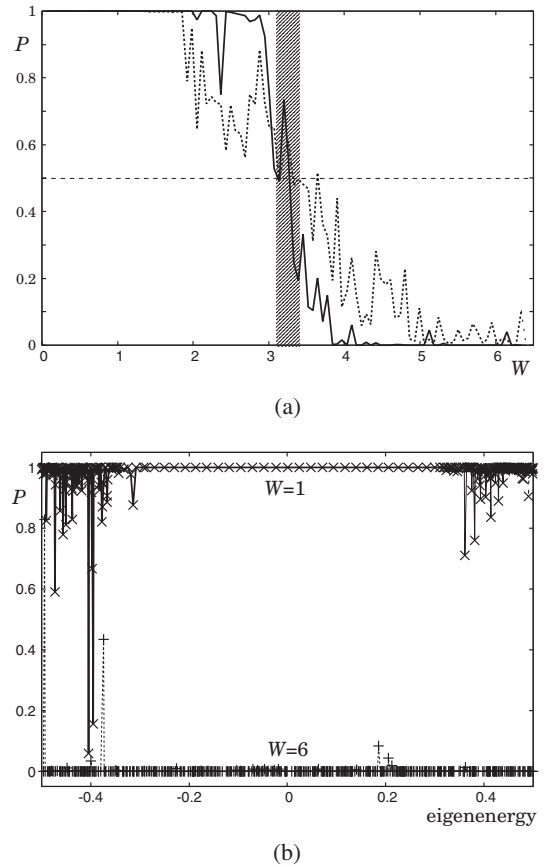


Fig. 5. (a) Probability of eigenfunction around $E = 0$ to be judged topological edge states as a function of disorder W . Averages over five samples are taken. 50% probability is indicated as the horizontal dashed line. Since the critical disorder is less accurate, $W_c^{\text{CI}} = 3.25 \pm 0.1$ is shown as a shaded region. The dotted line is for a two-weight-layer network, while the solid one is for a four-weight-layer one. (b) Same quantity but as a function of eigenenergy E inside the bulk band gap region $|E| < E_g = 0.5$. Results for $W = 1 < W_c^{\text{CI}}$ (x, solid line) and $W = 6 > W_c^{\text{CI}}$ (+, dotted line) are shown.

Since all we need to calculate are eigenfunctions with relatively small systems, the method will work for systems where the transfer matrix method is not applicable (localization problems on random^{37–40} and fractal lattices,⁴¹ for example).

We have used the known values of critical disorder to teach the machine. After learning the feature of eigenfunctions near the band center, the machine could capture localized/delocalized and topological/nontopological features away from the band center. We have also verified that the results of the SU(2) model learning can be applied to the Ando model.³²⁾

In the cases of Anderson transition near the band edge in the SU(2) model [Fig. 3(b)] and that at the band center in the Ando model, the machine tends to predict the transition for a slightly smaller disorder than the estimate of finite size scaling analyses.^{31,42)} We have extracted the features in the middle layers to explain this tendency,³²⁾ but could not clarify how the machine judges phases. The details of judgement should be clarified in the future.

We have focused on the amplitude of eigenfunction in 2D. In higher dimensions, the same algorithm will be applicable via dimensional reduction: integration of $|\psi|^2$ over certain directions, reducing the image to two dimensions. The dimensional reduction will also work for disordered 3D

strong and weak topological insulators.⁴³⁾ Other interesting quantities for machine learning are phase and spin texture of eigenfunctions in random electron systems. Classical waves (photon, phonon) in random media^{44–46)} as well as disordered magnon⁴⁷⁾ are also worth machine learning.

Acknowledgments The authors would like to thank Keith Slevin, Koji Kobayashi, and Ken-Ichiro Imura for useful discussions. This work was partly supported by JSPS KAKENHI Grant Number JP15H03700.

*ootsuki.t@msi.co.jp

†ohtsuki@sophia.ac.jp

- 1) P. W. Anderson, *Phys. Rev.* **109**, 1492 (1958).
- 2) A. P. Schnyder, S. Ryu, A. Furusaki, and A. W. W. Ludwig, *Phys. Rev. B* **78**, 195125 (2008).
- 3) A. Kitaev, *AIP Conf. Proc.* **1134**, 22 (2009).
- 4) M. Z. Hasan and C. L. Kane, *Rev. Mod. Phys.* **82**, 3045 (2010).
- 5) X.-L. Qi and S.-C. Zhang, *Rev. Mod. Phys.* **83**, 1057 (2011).
- 6) M. R. Zirnbauer, *J. Math. Phys. (N.Y.)* **37**, 4986 (1996).
- 7) A. Altland and M. R. Zirnbauer, *Phys. Rev. B* **55**, 1142 (1997).
- 8) B. Sbierski and P. W. Brouwer, *Phys. Rev. B* **89**, 155311 (2014).
- 9) H. Katsura and T. Koma, *J. Math. Phys. (N.Y.)* **57**, 021903 (2016); H. Katsura and T. Koma, [arXiv:1611.01928](#).
- 10) A. Rodríguez, L. J. Vasquez, K. Slevin, and R. A. Romer, *Phys. Rev. Lett.* **105**, 046403 (2010).
- 11) A. Rodríguez, L. J. Vasquez, K. Slevin, and R. A. Romer, *Phys. Rev. B* **84**, 134209 (2011).
- 12) K. Slevin and T. Ohtsuki, *New J. Phys.* **16**, 015012 (2014).
- 13) L. Ujfalusi and I. Varga, *Phys. Rev. B* **91**, 184206 (2015).
- 14) T. Obuchi, H. Koma, and M. Yasuda, *J. Phys. Soc. Jpn.* **85**, 114803 (2016).
- 15) Y. LeCun, Y. Bengio, and G. Hinton, *Nature* **521**, 436 (2015).
- 16) D. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. van den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot, S. Dieleman, D. Grewe, J. Nham, N. Kalchbrenner, I. Sutskever, T. Lillicrap, M. Leach, K. Kavukcuoglu, T. Graepel, and D. Hassabis, *Nature* **529**, 484 (2016).
- 17) J. Carrasquilla and R. G. Melko, [arXiv:1605.01735](#).
- 18) A. Tanaka and A. Tomiya, [arXiv:1609.09087](#).
- 19) G. Carleo and M. Troyer, [arXiv:1606.02318](#).
- 20) P. Broecker, J. Carrasquilla, R. G. Melko, and S. Trebst, [arXiv:1608.07848](#).
- 21) K. Ch'ng, J. Carrasquilla, R. G. Mello, and E. Khatami, [arXiv:1609.02552](#).
- 22) L. Li, T. E. Baker, S. R. White, and K. Burke, [arXiv:1609.03705](#).
- 23) E. P. van Nieuwenburg, Y.-H. Liu, and S. D. Huber, [arXiv:1610.02048](#).
- 24) L. Huang and L. Wang, [arXiv:1610.02746](#).
- 25) Y. Asada, K. Slevin, and T. Ohtsuki, *Phys. Rev. Lett.* **89**, 256601 (2002).
- 26) Y. Asada, K. Slevin, and T. Ohtsuki, *Phys. Rev. B* **70**, 035115 (2004).
- 27) Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, *Proc. IEEE* **86**, 2278 (1998).
- 28) Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, S. Guadarrama, and T. Darrell, [arXiv:1408.5093](#).
- 29) X. Glorot and Y. Bengio, *Aistats* **9**, 249 (2010).
- 30) T. Tieleman and G. Hinton, COURSE: Neural Networks for Machine Learning (2012) Lect. 6.5.
- 31) T. Ando, *Phys. Rev. B* **40**, 5325 (1989).
- 32) Results for Ando model and intermediate feature maps are provided online as [Supplemental Material](#).
- 33) J. P. Dahlhaus, J. M. Edge, J. Tworzydło, and C. W. J. Beenakker, *Phys. Rev. B* **84**, 115133 (2011).
- 34) S. Liu, T. Ohtsuki, and R. Shindou, *Phys. Rev. Lett.* **116**, 066401 (2016).
- 35) C.-Z. Chang, W. Zhao, J. Li, J. K. Jain, C. Liu, J. S. Moodera, and M. H. W. Chan, *Phys. Rev. Lett.* **117**, 126802 (2016).
- 36) X.-L. Qi, T. L. Hughes, and S.-C. Zhang, *Phys. Rev. B* **78**, 195424 (2008).
- 37) Y. Avishai and J. M. Luck, *Phys. Rev. B* **45**, 1074 (1992).
- 38) R. Berkovits and Y. Avishai, *Phys. Rev. B* **53**, R16125 (1996).
- 39) A. Kaneko and T. Ohtsuki, *J. Phys. Soc. Jpn.* **68**, 1488 (1999).
- 40) L. Ujfalusi and I. Varga, *Phys. Rev. B* **90**, 174203 (2014).
- 41) Y. Asada, K. Slevin, and T. Ohtsuki, *Phys. Rev. B* **73**, 041102 (2006).
- 42) U. Fastenrath, G. Adams, R. Bundschuh, T. Hermes, B. Raab, I. Schlosser, T. Wehner, and T. Wichmann, *Physica A* **172**, 302 (1991).
- 43) K. Kobayashi, T. Ohtsuki, and K.-I. Imura, *Phys. Rev. Lett.* **110**, 236803 (2013).
- 44) P. Sheng, *Introduction to Wave Scattering, Localization, and Mesoscopic Phenomena* (Academic Press, San Diego, CA, 1995).
- 45) C. M. Aegerter, M. Störzer, and G. Maret, *Europhys. Lett.* **75**, 562 (2006).
- 46) S. Faez, A. Strybulevych, J. H. Page, A. Lagendijk, and B. A. van Tiggelen, *Phys. Rev. Lett.* **103**, 155703 (2009).
- 47) B. Xu, T. Ohtsuki, and R. Shindou, [arXiv:1606.02839](#).