**Letters**

# Phase Diagrams of Three-Dimensional Anderson and Quantum Percolation Models Using Deep Three-Dimensional Convolutional Neural Network

Tomohiro Mano and Tomi Ohtsuki*

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

The three-dimensional Anderson model is a well-studied model of disordered electron systems that shows the delocalization–localization transition. As in our previous papers on two- and three-dimensional (2D, 3D) quantum phase transitions [J. Phys. Soc. Jpn. **85**, 123706 (2016), **86**, 044708 (2017)], we used an image recognition algorithm based on a multilayered convolutional neural network. However, in contrast to previous papers in which 2D image recognition was used, we applied 3D image recognition to analyze entire 3D wave functions. We show that a full phase diagram of the disorder-energy plane is obtained once the 3D convolutional neural network has been trained at the band center. We further demonstrate that the full phase diagram for 3D quantum bond and site percolations can be drawn by training the 3D Anderson model at the band center.

*Introduction*— Applying machine learning methods to solve problems in condensed matter physics has proven to be successful. Ising and spin ice models,[1,2] low dimensional topological systems,[3,4] strongly correlated systems,[5–14] as well as random two- and three-dimensional (2D, 3D) topological and non-topological systems,[15–17] have been studied using machine learning.

In previous papers,[15,16] we presented studies of 2D and 3D random electron systems via a multilayer convolutional neural network (CNN) approach called deep learning, and the features of quantum phase transitions such as delocalization–localization transitions (the Anderson metal–insulator transition) and topological–nontopological insulator transitions were shown to be captured. We diagonalized the Hamiltonian, obtained the eigenfunctions $\Psi_\nu(r)$, and trained the neural network by feeding the electron density $|\Psi_\nu(r)|^2$ for a specific quantum phase. We used 2D image recognition, and for 3D systems, integration over one direction was performed to reduce the 3D electron density to 2D. The advantage of reducing the electron density from 3D to 2D is that we can then use standard 2D image recognition. The drawback is that we lose the information of electron density distribution along one direction.

In this letter, we apply 3D image recognition via deep 3D CNN to analyze a random 3D electron system. The Anderson model for delocalization–localization transition[18] and the 3D quantum bond and site percolation models were studied. We show that the full phase diagram of the disorder-energy ($W$–$E$) plane is available once the 3D CNN has been trained at the band center, i.e., along $E = 0$ [see arrows in Fig. 2(b)]. We further demonstrate that the two-dimensional phase diagrams (see Fig. 3) for the 3D quantum percolation model, both bond and site types, can be drawn by training the 3D Anderson model at the band center, without learning the features of quantum percolations.

*Models and method*— We consider the following 3D Anderson Hamiltonian,[18]

$$H = \sum_x v_x c_x^\dagger c_x + \sum_{\langle x, x' \rangle} c_x^\dagger c_{x'}, \tag{1}$$

where $c_x^\dagger$ ($c_x$) denotes the creation (annihilation) operator of an electron at a site $x = (x, y, z)$ that forms a simple cubic lattice, and $v_x$ denotes the random potential at $x$. $\langle \cdots \rangle$

indicates the nearest-neighbor hopping. A box distribution is assumed with each $v_x$ uniformly and independently distributed at the interval $[-W/2, W/2]$. At energy $E = 0$, i.e., at the center of the band, the wave functions are delocalized when $W < W_c$ and the system is a diffusive metal. For $W > W_c$, the wave functions are exponentially localized and the system is an Anderson insulator. Here, the critical disorder $W_c$ at $E = 0$ is estimated to be $16.54 \pm 0.01$[19] by the finite size scaling analysis of the Lyapunov exponent calculated from the transfer matrix method.[20–23]

Next, we consider the 3D quantum bond and site percolation models described by the following Hamiltonian,[24–27] where electrons are allowed to move on connected sites of a classical percolation,[28–30]

$$H = \sum_x v_x c_x^\dagger c_x + \sum_{\langle x, x' \rangle} t_{xx'} c_x^\dagger c_{x'}. \tag{2}$$

The hopping parameter $t_{xx'}$ is defined as

$$t_{xx'} = \begin{cases} 1 & \text{(for connected bond)} \\ 0 & \text{(for disconnected bond)} \end{cases}, \tag{3}$$

where the energy unit is the transfer energy between connected sites. In the case of bond percolation, bonds are randomly connected with a probability $p_B$. In the case of site percolation, each site is filled with a probability $p_S$, and bonds are connected only if the sites on both sides of the bond are filled. For each realization of bond and site percolations, we identify the maximally connected cluster with a depth-first search, and analyze only the states in this cluster.

According to the studies on quantum percolation,[27,31–33] many strongly localized states, so-called molecular states, appear at energies, e.g., $E = 0, \pm 1, \pm\sqrt{2}$, when we set $v_x = 0$. These states are peculiar to the quantum percolation model, and are degenerate, resulting in strong peaks in the density of states. Due to this degeneracy, any linear combination is possible, which results in difficulty in judging the delocalized/localized phases. We therefore assume a very weak site random potential of the order of $10^{-3}$, namely, $v_x \in [-10^{-3}/2, 10^{-3}/2]$, and thereby lift the degeneracy.

Random numbers were generated by the Mersenne Twister algorithm,[34] MT2203. We imposed the periodic boundary condition, and the maximum modulus of the eigenfunction

was shifted to the center of the system to improve the accuracy of the 3D image recognition. For all calculations, we considered a $40 \times 40 \times 40$ lattice, and diagonalized the Hamiltonian. For training data and the data for Fig. 2(a), we used the sparse matrix diagonalization Intel MKL/FEAST,[35] since we restricted the energy range to around $E = 0$. For drawing the phase diagrams, we used the standard Linear Algebra package LAPACK[36] for diagonalization, and obtained all the eigenvalues and eigenvectors.

To calculate the probabilities of delocalized and localized phases, $P_{\text{deloc}}$ and $P_{\text{loc}}$ ($= 1 - P_{\text{deloc}}$) for a given eigenfunction, we adopted a 3D CNN using Keras[37] as frontend and TensorFlow as backend. NVIDIA GTX GEFORCE 1080 and 1080 Ti graphics cards were used for GPU calculation. The configuration of our CNN and detailed parameters are shown schematically in Fig. 1. The input shape was $40 \times 40 \times 40$, corresponding to the system size of our numerical simulation. No bias parameters are included in the weight parameters. Through the hidden layers that consist of a convolution layer, max pooling layer, and densely-connected layer, our neural network outputs two real numbers, i.e., the probability of a delocalized (localized) phase $P_{\text{deloc}}$ ($P_{\text{loc}}$) in the case of the Anderson and quantum percolation models. The network utilizes the AdaDelta solver[38] as the stochastic gradient decent solver and rectified linear unit (ReLU) as its activation function except for the last layer, which is activated by the softmax function.[39] The loss function is defined by the cross entropy (which is *categorical_crossentropy* in Keras), which was minimized during the training. To avoid overfitting, dropout processes, which randomly drop half of the inputs, were implemented in some layers (see Fig. 1). Before training, we randomly chose 10% of the training data as a validation data set.

We then trained our neural network based on data from the 3D Anderson model, and determined the network weight parameters, which captured the features of the delocalized/localized phase. For training data, we restricted the energy region to the vicinity of the band center, $E = 0$, in the Anderson model. The main purpose of this work was to obtain global delocalization–localization phase diagrams of 3D random electron systems, and we assumed that the critical disorder at $E = 0$, $W_c = 16.54$, was known in advance. As the training data of the delocalized phase, we prepared 4000 eigenstates in the range, $W \in [14.0, 16.0]$. As the training data for the localized phase, we prepared 4000 eigenstates in the range, $W \in [17.0, 19.0]$. These training regions are indicated by arrows in Fig. 2(b).

To confirm that the network sufficiently captures the features of the eigenstates, we prepared test data from the Anderson model for various $W$ and $E$, and determined the phase diagram in $W$–$E$ parameter space and then compared it with those obtained by other researchers.[40–42] Once the results were confirmed, we applied this trained CNN to bond and site quantum percolation models, the phase diagrams of which are less well-known. 5 independent samples were prepared for the test data from the Anderson and quantum percolation models, and the output probabilities are averaged over the 5 samples.

*Result for Anderson model*— In Fig. 2(a), we show the probabilities $P_{\text{deloc}}$ and $P_{\text{loc}}$ at $E = 0$ as functions of $W$. The test data consisted of $100 \times 5$ eigenfunctions closest to
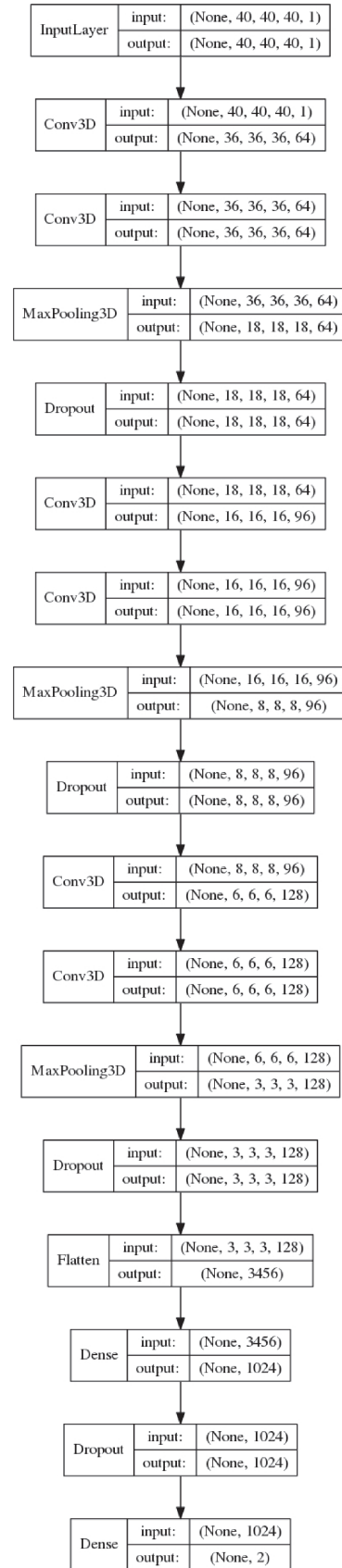


**Fig. 1.** Schematic of the deep 3D CNN structure, where the neural network parameters are indicated. 6 convolutions have been performed, with the 1st and 2nd convolution $5 \times 5 \times 5$ with stride 1, and the other convolutions, $3 \times 3 \times 3$ with stride 1. For the 2nd, 4th, and 6th convolutions, zero padding is applied so that the input and the output image sizes are the same. The activation function ReLU was performed after each convolution layer as well as the densely-connected layer, except for the last layer where softmax was used.
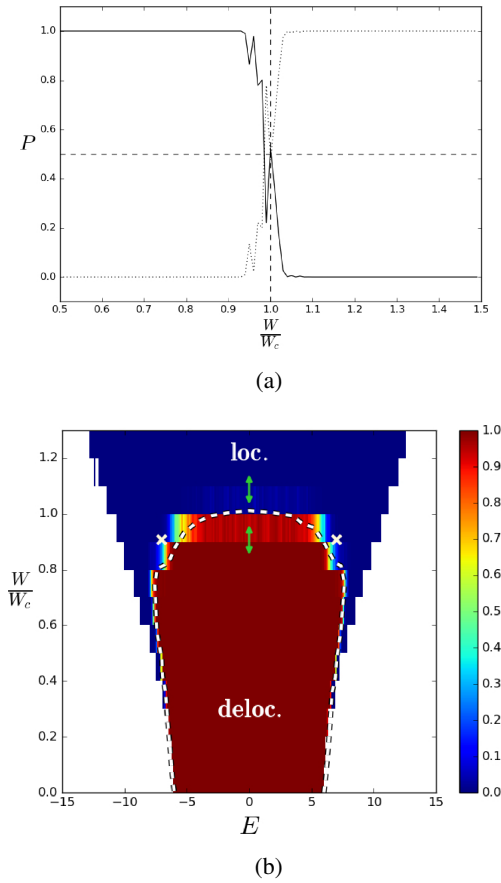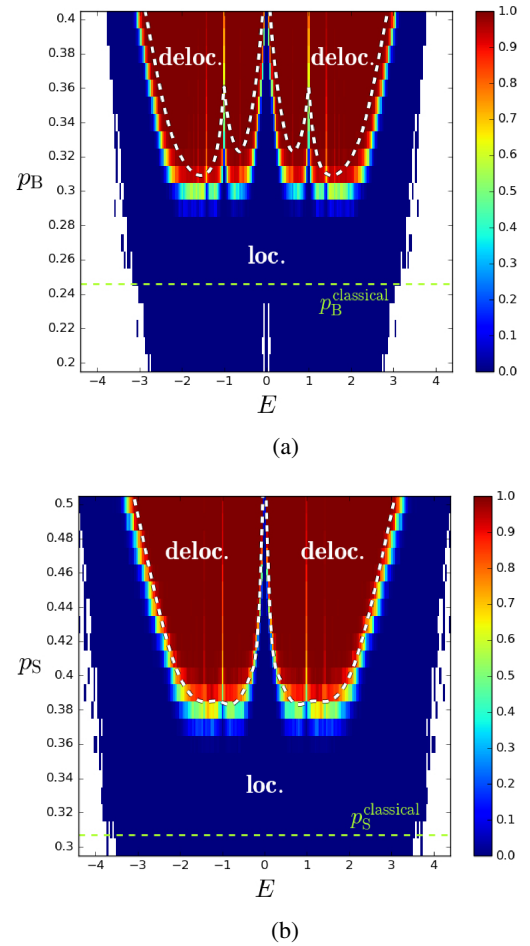
**Fig. 2.** (Color online) The probability that the states are delocalized/localized for the Anderson model. $W_c = 16.54$ is the critical disorder for $E = 0$. The training regions, which are 1D line segments, are indicated by the arrows. (a) $P_{\text{deloc}}$ (solid line) and $P_{\text{loc}}$ (dotted line) when the energy $E$ is close to the band center. 50% probability as well as $W = W_c$ is indicated as dashed lines as a guide to the eye. (b) Color map of the intensity $P_{\text{deloc}}$ for all the energy spectrum. In both figures, an average over 5 samples was performed. Dashed line is from Ref. 42, while the crosses (×) are from Ref. 41.





**Fig. 3.** (Color online) Phase diagram of the (a) bond and (b) site quantum percolations. The intensity $P_{\text{deloc}}$ is shown. Classical percolation thresholds for the bond percolation, $p_B^{\text{classical}}$, and for the site percolation, $p_S^{\text{classical}}$, are indicated by green dashed lines. In both figures, the average over 5 samples has been performed. The white dashed line in (a) is a spline interpolation from the data in Ref. 48, while that in (b) is the spline interpolation from the data in Ref. 27.

$E = 0$: 100 values of $W \in [\frac{1}{2} W_c, \frac{3}{2} W_c]$, and 5 independent samples. The average over 5 samples was taken as in Fig. 1 of Ref. 16. From this figure, the Anderson metal–insulator transition from delocalized (metal) phase to localized (insulator) phase has been observed around $W_c$, from which we can confirm that the CNN correctly detected the Anderson transition.

We then prepared eigenfunctions all over the energy spectra with varying $W$, and let the CNN determine the phase. The result is shown in Fig. 2(b) where the intensity $P_{\text{deloc}}$ is plotted as a color map, which is in good agreement with the previous results from the transfer matrix.[40,41] The flat phase boundary around $W_c$, $|E| \lesssim 6$, together with the rapidly changing critical disorder outside the band edge $|E| = 6$ in the absence of randomness, are reproduced.

For relatively small disorder, the CNN judges that eigenstates near the band edge are localized. With increasing $W$ but fixed $E$ around $E = 7$, the localized states near the band edge are delocalized and then localized again, a re-entrant phenomenon known for the Anderson model with the box distributed site energies.[40–42]

Note that the states near the band edge with small disorders look rather like bound states trapped by the potential fluctuation, and smoothly change in space. It is interesting that the CNN trained only around $E = 0$, where the random quantum interference was strong and the wave function amplitudes were rapidly fluctuating, correctly judged the smooth bound states as localized.

*Application to quantum percolation*— Figures 3(a) and 3(b) show the phase diagrams of the bond and site quantum percolation models, respectively. Eigen energies of the maximal clusters are plotted on the horizontal axis, while the probability $p_B$ or $p_S$ is plotted on the vertical axis. All over the energy spectra, the quantum percolation transition occurs well above the classical percolation threshold, $p_B^{\text{classical}} \approx 0.2488$ and $p_S^{\text{classical}} \approx 0.3116$.[43,44] The energy dependence of the quantum site percolation threshold is consistent with the finite size scaling analysis of multi-fractality[45–47] by Ujfalusi and Varga,[27] where a strong energy dependence is observed near $E = 0$ and near the band edge.[48–50]

In the quantum percolation models, even on the maximal cluster, there appear energetically degenerate molecular states, the amplitudes of which are localized on finite sites, while being exactly zero on other sites.[31,32] As stated while

explaining the model, we lifted the degeneracy by introducing a very weak site randomness. The molecular states can coexist with other states, but their numbers are much smaller than other non-molecular states except at specific energies $E = 0, \pm 1, \pm\sqrt{2}, \dots$ where the shapes of molecular states are relatively simple. Because we average the probability $P_{\text{deloc}}$ over a finite energy region, the phase is judged to be delocalized (metal) if the other non-molecular states are delocalized.

*Summary and concluding remarks*— In this letter, we applied 3D image recognition to draw the global phase diagrams of 3D random electron systems. We have shown that once the 3D CNN is trained in a small region of the Anderson transition, the global phase diagram can be obtained by using this training. Furthermore, the CNN trained for the Anderson model at $E = 0$ allows us to draw the phase diagrams of the quantum bond and site percolation models. Thus, the present work has opened a way to use 3D deep CNN for condensed matter physics. Application to strongly correlated 3D electron systems with disorder[51–53] is one of the interesting problems that could benefit from this approach. Another interesting application is to distinguish rare event states in Weyl/Dirac semimetals[54,55] from other types of states, which is especially important in discussing the Weyl/Dirac semimetal to metal transition.[56–59]

The present method can be extended easily to analyze other quantum percolation problems, such as systems with magnetic fields or spin–orbit interactions, i.e., systems belonging to different symmetry classes.[60–63] In these systems, the quantum percolation threshold is expected to be lowered due to the breaking of the time reversal or spin–rotational symmetries.[26] How the energy-dependent quantum percolation threshold is changed by the breaking of the symmetries is an interesting problem that remains to be studied in the future.

Let us conclude this letter by comparing the present method with the transfer matrix method,[19–23] and stating that the present method and the transfer matrix method are complementary. Due to the zero connections between sites that appear randomly, the transfer matrix analysis is not applicable to the quantum percolation problem. This situation is similar to topological insulators with a specific set of parameters.[16] The method presented here is, in this sense, more general than the transfer matrix method. In addition, we can draw the phase diagram in the $W$–$E$ plane easily using the present method, while the transfer matrix method requires a one-dimensional parameter change, hence drawing the global $W$–$E$ phase diagram is more numerically demanding. On the other hand, the transfer matrix method gives a more precise estimate of the critical point than the present method. The estimate of the critical exponent is also possible in the finite size scaling analysis of the Lyapunov exponent obtained by the transfer matrix method, which is so far impossible using the present method. A possible way to estimate the critical exponent by the present method is to change the system size $L$ and observe how the quantities, such as probabilities of delocalized/localized states and/or convergence time vary with $L$. The independent estimate of critical disorder $W_c$, with the aid of unsupervised training, is also an important topic for the future.

*ohtsuki@sophia.ac.jp

1) J. Carrasquilla and R. G. Melko, Nat. Phys. **13**, 431 (2017).
2) A. Tanaka and A. Tomiya, J. Phys. Soc. Jpn. **86**, 063001 (2017).
3) Y. Zhang and E.-A. Kim, Phys. Rev. Lett. **118**, 216401 (2017).
4) P. Zhang, H. Shen, and H. Zhai, arXiv:1708.09401.
5) G. Carleo and M. Troyer, Science **355**, 602 (2017).
6) P. Broecker, J. Carrasquilla, R. G. Melko, and S. Trebst, Sci. Rep. **7**, 8823 (2017).
7) K. Ch'ng, J. Carrasquilla, R. G. Melko, and E. Khatami, Phys. Rev. X **7**, 031038 (2017).
8) L. Li, T. E. Baker, S. R. White, and K. Burke, Phys. Rev. B **94**, 245129 (2016).
9) E. P. van Nieuwenburg, Y.-H. Liu, and S. D. Huber, Nat. Phys. **13**, 435 (2017).
10) L. Huang and L. Wang, Phys. Rev. B **95**, 035105 (2017).
11) F. Schindler, N. Regnault, and T. Neupert, Phys. Rev. B **95**, 245134 (2017).
12) H. Saito, J. Phys. Soc. Jpn. **86**, 093001 (2017).
13) H. Saito and M. Kato, arXiv:1709.05468.
14) H. Fujita, Y. Nakagawa, S. Sugiura, and M. Oshikawa, arXiv:1705.05372.
15) T. Ohtsuki and T. Ohtsuki, J. Phys. Soc. Jpn. **85**, 123706 (2016).
16) T. Ohtsuki and T. Ohtsuki, J. Phys. Soc. Jpn. **86**, 044708 (2017).
17) N. Yoshioka, Y. Akagi, and H. Katsura, arXiv:1709.05790.
18) P. W. Anderson, Phys. Rev. **109**, 1492 (1958).
19) K. Slevin and T. Ohtsuki, New J. Phys. **16**, 015012 (2014).
20) A. MacKinnon and B. Kramer, Phys. Rev. Lett. **47**, 1546 (1981).
21) A. MacKinnon and B. Kramer, Z. Phys. B **53**, 1 (1983).
22) J.-L. Pichard and G. Sarma, J. Phys. C **14**, L127 (1981).
23) B. Kramer and A. MacKinnon, Rep. Prog. Phys. **56**, 1469 (1993).
24) Y. Avishai and J. M. Luck, Phys. Rev. B **45**, 1074 (1992).
25) R. Berkovits and Y. Avishai, Phys. Rev. B **53**, R16125 (1996).
26) A. Kaneko and T. Ohtsuki, J. Phys. Soc. Jpn. **68**, 1488 (1999).
27) L. Ujfalusi and I. Varga, Phys. Rev. B **90**, 174203 (2014).
28) A. Aharony and D. Stauffer, *Introduction to Percolation Theory* (Taylor & Francis, London, 1994) 2nd ed.
29) *Quantum and Semi-classical Percolation and Breakdown in Disordered Solids*, ed. B. K. Chakrabarti, K. K. Bardhan, and A. K. Sen (Springer, Berlin, 2009).
30) A. A. Saberi, Phys. Rep. **578**, 1 (2015).
31) S. Kirkpatrick and T. P. Eggarter, Phys. Rev. B **6**, 3598 (1972).
32) J. T. Chayes, L. Chayes, J. R. Franz, J. P. Sethna, and S. A. Trugman, J. Phys. A **19**, L1173 (1986).
33) G. Schubert and H. Fehske, *Quantum Percolation in Disordered Structures* (Springer, Berlin, 2009) Chap. 5, p. 135.
34) M. Matsumoto and T. Nishimura, ACM Trans. Modeling Comput. Simulation **8**, 3 (1998).
35) E. Polizzi, Phys. Rev. B **79**, 115112 (2009).
36) LAPACK — Linear Algebra PACKage. [http://www.netlib.org/lapack].
37) F. Chollet, Keras. [https://github.com/fchollet/keras] (2015).
38) M. D. Zeiler, arXiv:1212.5701.
39) I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning* (MIT Press, Cambridge, MA, 2016).
40) B. Bulka, M. Schreiber, and B. Kramer, Z. Phys. B **66**, 21 (1987).
41) S. L. A. de Queiroz, Phys. Rev. B **63**, 214202 (2001).
42) G. Schubert, A. Weiße, G. Wellein, and H. Fehske, in *HQS@HPC: Comparative Numerical Study of Anderson Localisation in Disordered Electron Systems*, ed. A. Bode and F. Durst (Springer, Berlin/Heidelberg, 2005) p. 237.
43) A. Sur, J. L. Lebowitz, J. Marro, M. H. Kalos, and S. Kirkpatrick, J. Stat. Phys. **15**, 345 (1976).
44) J. Wang, Z. Zhou, W. Zhang, T. M. Garoni, and Y. Deng, Phys. Rev. E **87**, 052107 (2013).
45) A. Rodriguez, L. J. Vasquez, K. Slevin, and R. A. Romer, Phys. Rev. Lett. **105**, 046403 (2010).
46) A. Rodriguez, L. J. Vasquez, K. Slevin, and R. A. Romer, Phys. Rev. B

**84**, 134209 (2011).

47) L. Ujfalusi and I. Varga, Phys. Rev. B **91**, 184206 (2015).

48) C. M. Soukoulis, Q. Li, and G. S. Grest, Phys. Rev. B **45**, 7724 (1992).

49) A. Kusy, A. Stadler, G. Haldas, and R. Sikora, Physica A **241**, 403 (1997).

50) I. Travěnec, Int. J. Mod. Phys. B **22**, 5217 (2008).

51) H. Shinaoka and M. Imada, J. Phys. Soc. Jpn. **78**, 094708 (2009).

52) H. Shinaoka and M. Imada, J. Phys. Soc. Jpn. **79**, 113703 (2010).

53) Y. Harashima and K. Slevin, Phys. Rev. B **89**, 205108 (2014).

54) R. Nandkishore, D. A. Huse, and S. L. Sondhi, Phys. Rev. B **89**, 245110 (2014).

55) J. H. Pixley, D. A. Huse, and S. Das Sarma, Phys. Rev. X **6**, 021042 (2016).

56) K. Kobayashi, T. Ohtsuki, K.-I. Imura, and I. F. Herbut, Phys. Rev. Lett. **112**, 016402 (2014).

57) Y. Ominato and M. Koshino, Phys. Rev. B **89**, 054202 (2014).

58) S. V. Syzranov, V. Gurarie, and L. Radzihovsky, Phys. Rev. B **91**, 035133 (2015).

59) S. Liu, T. Ohtsuki, and R. Shindou, Phys. Rev. Lett. **116**, 066401 (2016).

60) F. J. Dyson, J. Math. Phys. **3**, 140 (1962).

61) F. J. Dyson, J. Math. Phys. **3**, 1199 (1962).

62) A. Altland and M. R. Zirnbauer, Phys. Rev. B **55**, 1142 (1997).

63) F. Evers and A. D. Mirlin, Rev. Mod. Phys. **80**, 1355 (2008).