Machine learning electron correlation in a disordered medium

Jianhua Ma,[1,*] Puhan Zhang,[2] Yaohua Tan,[1] Avik W. Ghosh,[1,2,†] and Gia-Wei Chern[2,‡]
[1]*Charles L. Brown Department of Electrical and Computer Engineering, University of Virginia, Charlottesville, Virginia 22904, USA*
[2]*Department of Physics, University of Virginia, Charlottesville, Virginia 22904, USA*

(Received 16 October 2018; revised manuscript received 1 February 2019; published 11 February 2019)

Learning from data has led to a paradigm shift in computational materials science. In particular, it has been shown that neural networks can learn the potential energy surface and interatomic forces through examples, thus bypassing the computationally expensive density functional theory calculations. Combining many-body techniques with a deep-learning approach, we demonstrate that a fully connected neural network is able to learn the complex collective behavior of electrons in strongly correlated systems. Specifically, we consider the Anderson-Hubbard (AH) model, which is a canonical system for studying the interplay between electron correlation and strong localization. The ground states of the AH model on a square lattice are obtained using the real-space Gutzwiller method. The obtained solutions are used to train a multitask multilayer neural network, which subsequently can accurately predict quantities such as the local probability of double occupation and the quasiparticle weight, given the disorder potential in the neighborhood as the input.

DOI: 10.1103/PhysRevB.99.085118

I. INTRODUCTION

Machine learning (ML) [1–3] is one of today's most rapidly growing interdisciplinary fields. The deep-learning neural network (NN) provides a powerful universal method for finding patterns and regularities in high-dimensional data [4,5]. It has found successful applications in a wide variety of fields. In condensed-matter physics and materials science, notable applications include using ML to guide materials design [6–8] and for identification and classification of crystalline structures [9–13]. Recently, ML techniques have also been taken up by researchers in the area of strongly correlated systems. The majority of such activities focus on using ML to identify phases and phase transitions in many-body systems ranging from classical statistical models [14–17] and quantum fermionic Hamiltonians [18,19] to topological phases [20] and many-body localization [21]. In these studies, a deep-learning NN, trained with data from classical or quantum Monte Carlo simulations, is shown to be able to correctly distinguish phases and predict phase diagrams. ML trained NNs can also represent thermodynamic phases in equilibrium (Boltzmann machines) [22], or ground-state wave functions of quantum many-body systems [23,24].

In this paper, we demonstrate another application of ML in correlated electron systems, namely, using NN as an efficient emulator for many-body problem solvers. Specifically, our goal is to investigate whether deep-learning NN can be trained to predict electron correlation, such as the probability of double occupation, in a disordered medium. Our approach here is similar in spirit to those adopted in quantum chemistry and materials science communities, where the ML trained NN is used to bypass the time-consuming density functional theory (DFT) calculations [25–31]. Such activities have led to the fast prediction of molecular atomization energies [32,33] and efficient parametrization of interatomic force fields [34–38], to name a few. We note in passing that similar ideas of bypassing expensive numerical calculations with the ML model have also been explored in correlated electron systems, such as using ML to replace the impurity solver for dynamical mean-field theory (DMFT) [39], or to speed up the total energy calculation in Monte Carlo simulations [40–42].

We consider the disordered Hubbard model in two dimensions,

$$\mathcal{H} = -t \sum_{ij,\sigma} \hat{c}^{\dagger}_{i,\sigma} \hat{c}_{j,\sigma} + \sum_{i,\sigma} \epsilon_i \hat{n}_{i,\sigma} + U \sum_{i} \hat{n}_{i,\uparrow} \hat{n}_{i,\downarrow}, \qquad (1)$$

where $\hat{c}^{\dagger}_{i,\sigma}$ is the electron creation operator with spin $\sigma = \uparrow, \downarrow$ at site i, and $\hat{n}_{i,\sigma} \equiv \hat{c}^{\dagger}_{i,\sigma} \hat{c}_{i,\sigma}$ is the corresponding number operator. The first term describes the nearest-neighbor hopping of electrons. The second term denotes the random local potential. The last term is the on-site Hubbard repulsion. As in the standard Anderson model, here the site energy ϵ_i is a random number drawn uniformly from the interval $[-W/2, +W/2]$. We work at half filling on an $L \times L$ square lattice with periodic boundary conditions. The Hamiltonian Eq. (1), also known as the Anderson-Hubbard (AH) model, is considered a paradigmatic model for studying the interplay between strong electron correlation and disorder.

The AH model has been intensively studied by several numerical methods, including Hartree-Fock calculations [43,44], quantum Monte Carlo simulations [45–48], and extended DMFT [49–52]. In particular, an intrinsic metal-insulator transition *without* magnetic order can be quantitatively calculated within the DMFT framework [53]. For application to disordered systems, DMFT can be readily combined with the typical medium theory (TMT) in which a

*jm9yq@virginia.edu
†ag7rq@virginia.edu
‡gc6u@virginia.edu

geometrically averaged local density of states is used to construct the electron bath [54]. The nonmagnetic phase diagram of the AH model obtained from the TMT-DMFT method includes three distinct phases: a correlated metallic phase, a Mott insulating phase, and an Anderson insulating phase [51,52]. Importantly, the two insulating phases of the AH model have very different characters. The Mott insulator results from the strong correlation effect which prohibits electrons from hopping to the neighboring sites. On the other hand, strong disorder weakens the constructive interference that allows an electron wave packet to propagate coherently in a periodic potential, leading to the Anderson insulator. TMT-DMFT calculation shows that these two insulating phases are continuously connected [51,52].

Real-space approaches such as variational Monte Carlo (VMC) simulations [47,48], statistical DMFT [49,55], and the Gutzwiller methods [56,57] can better cope with the crucial spatial fluctuations in low-dimensional systems. Applying VMC to the two-dimensional (2D) AH model finds a continuous transition that separates the Mott insulator from the Anderson insulator in the nonmagnetic phase diagram [47,48]. It is worth noting that there is no sharp distinction between correlated metal and the Anderson insulator in 2D. Interestingly, detailed large-scale simulations of the 2D AH model within the Brinkman-Rice formalism, where the efficient Gutzwiller method can be applied, showed that strong spatial inhomogeneity gives rise to an electronic Griffiths phase that precedes the metal-insulator transition [56].

II. REAL-SPACE GUTZWILLER METHOD

Here, we employ the Gutzwiller method to solve the AH model on a square lattice. In its original formulation, a variational wave function $|\Psi_G\rangle = \mathcal{P}_G|\Psi_0\rangle$ is constructed by applying a real-space projector $\mathcal{P}_G = \prod_i \mathcal{P}_i$ on the Slater determinant $|\Psi_0\rangle$ obtained from the noninteracting electron Hamiltonian [58]. Optimization of $|\Psi_G\rangle$ can be efficiently carried out with the so-called Gutzwiller approximation (GA) [58], which becomes exact in the infinite-dimension limit. Moreover, GA corresponds to the zero-temperature saddle-point solution of the slave-boson (SB) method [59]. Indeed, by factoring out the occupation probability P_i^0 of the uncorrelated state, the local projector can be expressed as $\mathcal{P}_i \equiv \sum_{\alpha,\beta} \Phi_{i,\alpha\beta}/(P_{i,\beta}^0)^{-1/2}|\alpha\rangle\langle\beta|$, where α, β are the local many-electron states, and the elements of the variational matrix Φ_i correspond to the SB coherent-state amplitude [60,61]. For the single-band Hubbard model, Φ_i is a diagonal matrix of dimension 4, i.e., $\Phi_i = \mathrm{diag}(e_i, p_{i,\uparrow}, p_{i,\downarrow}, d_i)$, and the square of these diagonal elements corresponds to the probability of empty, single (with spin $\sigma = \uparrow, \downarrow$), and double-occupied states, respectively. In the following, we consider the nonmagnetic solutions of the AH model and assume $p_{i,\uparrow} = p_{i,\downarrow} = p_i$.

The GA solution for the AH model in Eq. (1) is obtained by minimizing the following energy functional,

$$\mathcal{E}(\rho_{ij}, \Phi_i) = -2t \sum_{\langle ij\rangle} \mathcal{R}_i \mathcal{R}_j \rho_{ij} + 2\sum_i \epsilon_i \rho_{ii}, \\ + U\sum_i d_i^2 + 2\sum_i \mu_i\left(\rho_{ii} - p_i^2 - d_i^2\right). \quad (2)$$

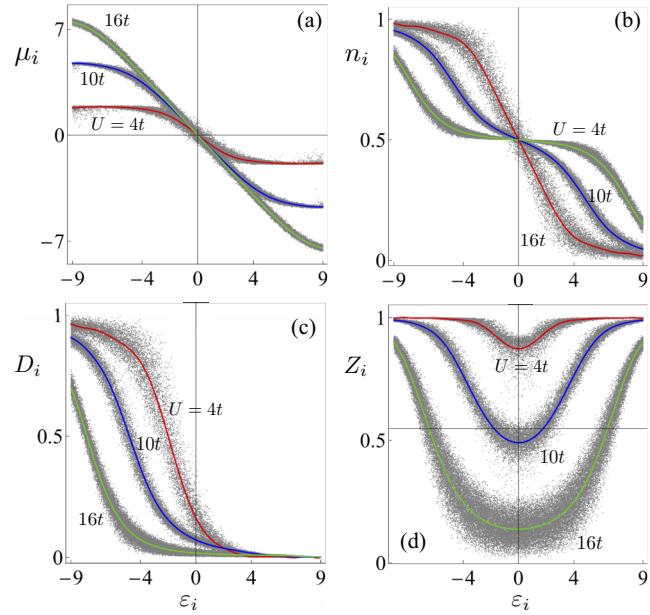

FIG. 1. Summary of the GA solution for the AH model on a 30×30 square lattice. The panels show the scatter diagram of (a) local energy correction μ_i, (b) site electron density n_i, (c) probability of double occupation $D_i = d_i^2$, and (d) local quasiparticle weight $Z_i = \mathcal{R}_i^2$ vs the random site energy ε_i. The data points were obtained from calculations with random strength $W/t = 6, 10, 14, 18$ and three different $U = 4t$, $10t$, and $16t$. The smooth curves showing the underlying overall trend for a given U were obtained using polynomial regression with up to 14th-order polynomials. The red, blue, and green curves correspond to $U = 4t$, $10t$, and $16t$, respectively.

Here, the prefactor 2 accounts for the spin degeneracy, $\rho_{ij} = \langle\Psi_0|c_j^\dagger c_i|\Psi_0\rangle$ is the single-particle density matrix, $\mathcal{R}_i = (e_i p_i + p_i d_i)/\sqrt{n_i(1-n_i)}$ is the Gutzwiller renormalization factor [58], $n_i = n_{i,\uparrow} = n_{i,\downarrow}$ is the local electron density, and μ_i is the Lagrangian multiplier that enforces the Gutzwiller constraint $n_i = p_i^2 + d_i^2 = \rho_{ii}$ [60,61]. The optimization of the density matrix, or equivalently of the wave function $|\Psi_0\rangle$, amounts to solving the following renormalized tight-binding Hamiltonian,

$$\hat{\mathcal{H}}^* = -t\sum_{\langle ij\rangle} \mathcal{R}_i\mathcal{R}_j \hat{c}_i^\dagger \hat{c}_j + \sum_i (\varepsilon_i + \mu_i)\hat{n}_i. \quad (3)$$

The minimization with respect to SB amplitudes $\partial\mathcal{E}/\partial\Phi_i = 0$, subject to constraint $e_i^2 + 2p_i^2 + d_i^2 = 1$, can be recast into an eigenvalue problem for each site. These two steps, optimization of Ψ_0 and Φ_i, have to be iterated until convergence is reached.

Using the above GA solver on a $L = 30$ square lattice, large data sets were generated with various disorder strengths $W/t = 6, 10, 14, 18$ and Hubbard parameters $U/t = 2, 4, \ldots, 16$. The scatter plots in Fig. 1 show the various local quantities versus the random site energy ε_i obtained from the GA solution with three different values of Hubbard repulsion. The local quantities are the site Lagrangian multiplier μ_i, the local electron density n_i, the double-occupation probability $D_i = d_i^2$, and the local quasiparticle weight $Z_i = \mathcal{R}_i^2$. Interestingly, for a given U, the data points cluster around a smooth

curve, indicating an underlying continuous trend. More quantitatively, we used polynomial regression to determine the overall dependence of the local quantities on the site energy ε_i; see the solid curves in Fig. 1.

Extensive studies on the statistics of electron correlation in the 2D AH model have been carried out using SB or real-space DMFT methods [55–57]. One interesting phenomenon is the screening of the impurity potential due to electron correlations, especially close to the metal-insulator transition. Our result shown in Fig. 1(a) clearly demonstrates this trend. Indeed, from Eq. (3), one can define a renormalized site potential as $\tilde{\varepsilon}_i = \varepsilon_i + \mu_i$. The anticorrelation between μ_i and ε_i thus results in a reduced effective site potential. Moreover, the local density n_i exhibits a more homogeneous distribution in the vicinity of the Fermi energy with increasing U; see Fig. 1(b).

The overall behavior of local quasiparticle weight versus ε is consistent with the result obtained from TMT-DMFT using the SB method as the impurity solver [52]. As shown in Fig. 1(d), electrons at large $|\varepsilon_i|$ get less renormalization, i.e., retain a larger Z_i, compared with those close to the Fermi energy ($\varepsilon \sim 0$). Moreover, the difference between large and small Z_i increases as one approaches the Mott transition boundary. This behavior also indicates a strong spatial inhomogeneity. While electrons in some regions become localized magnetic moments characterized by a vanishing Z_i, electrons in other regions undergo an Anderson localization transition and maintain a large value of Z_i.

III. NEURAL NETWORK MODEL

In order to capture the spatial site-to-site fluctuations of the electron correlation, we next employ deep-learning techniques to predict the local electronic properties of the AH model. More specifically, our goal is to predict the local quantities μ, n, D, and Z at a randomly picked site, say, site 0, with the site potentials ε_j in its neighborhood within a cutoff radius r_c as the input; see Fig. 2(a). This, of course, is based on the assumption of locality which implies that correlation functions decay strongly with the distance. In general, the single-particle density matrix exhibits an exponential and a power decay for insulators and metals, respectively. The localization of electron wave functions due to disorder also enhances the decay of correlation functions, especially in 2D. To quantify this locality approximation, we have repeated our ML training with various r_c, and have verified that the predictions of the NNs are not sensitive to the cutoff radius. The results presented below were obtained by including up to 14th nearest neighbors with a total of 89 sites within the cutoff.

A proper representation of the site energies ε_j is crucial in order to provide a description of the neighborhood that is invariant under fundamental transformations of the lattice symmetry. To this end, we first decompose all ε_j into irreducible representations (irrep) of point group D_4, which is the site symmetry group of a square lattice. The neighboring sites can be classified into three different invariant subsets, as shown in Fig. 2(b). Decomposition of these subsets into the corresponding irreps is straightforward. Taking the square as an example, there are three irreps: $x_{A_1} = \varepsilon_a + \varepsilon_b + \varepsilon_c + \varepsilon_d$, $x_{B_1} = \varepsilon_a - \varepsilon_b + \varepsilon_c - \varepsilon_d$, and $\mathbf{x}_E = (\varepsilon_a - \varepsilon_c, \varepsilon_b - \varepsilon_d)$. The amplitudes of each irrep and their relative phases are then used as the input for the NN. For example, consider all doublet irreps: $\mathbf{x}_m$ with $m = 1, 2, \ldots, M$, where M is the total number. The amplitudes $|\mathbf{x}_m|$, and relative angle $\cos\theta_{mn} = \mathbf{x}_m \cdot \mathbf{x}_n / |\mathbf{x}_m|\,|\mathbf{x}_n|$, are invariant under symmetry operations. We note that this descriptor of the site environment is similar to the atom-centered symmetry functions used in ML potentials for quantum molecular dynamics simulations [34,37].

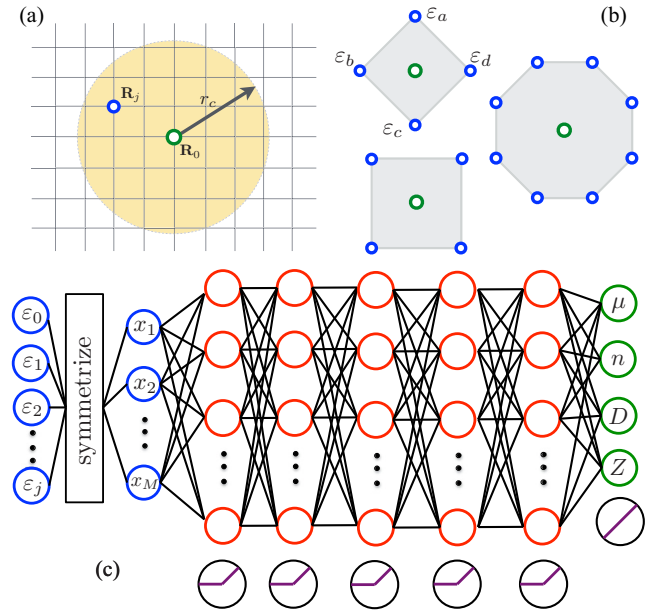

FIG. 2. (a) Schematic showing the target site at $\mathbf{R}_0$ on the square lattice. Random potentials ε_j of neighbors up to some cutoff radius r_c are used as the input to the neural network (NN). (b) Basic invariant subgroups of neighbors include two types of squares and a octagons. (c) Architecture of the fully connected NN for the disordered correlated systems. For input, we use all the random-distributed on-site energies in the certain circle. The ReLU activation function is used in the five hidden-feature extraction layers with $512 \times 256 \times 256 \times 128 \times 64$ nodes. The linear activation function is used to predict local quantities including μ, n, D, and Z.

We design a fully connected neural network (NN) with five hidden layers consisting of $n = 512 \times 256 \times 256 \times 128 \times 64$ rectified linear unit (ReLU) neurons [62]. The input layer is the symmetrized neighborhood ε as discussed above. The NN performs a sequence of transformations on the input that is illustrated in Fig. 2(c). In the mth layer, the nth neuron processes the activation $\mathbf{a}^{(m-1)}$ from the $(m-1)$th layer through independent weights and biases $\mathbf{w}^{(m-1)}\mathbf{a}^{(m-1)} + \mathbf{b}^{(m-1)}$. After the ReLU functions, the outcome is fed forward to be processed by the output neuron with the linear activation function. Importantly, here we adopt the multitask ML technique [63] that forces the NN to learn multiple local electron properties simultaneously. The additional constraints coming from the multitask setup helps the search for the true ML model because of the smaller set of models that can fit all properties simultaneously.

We use the mean absolute error (MAE) as the cost function with the L2 regularization [64] to avoid overfitting and a

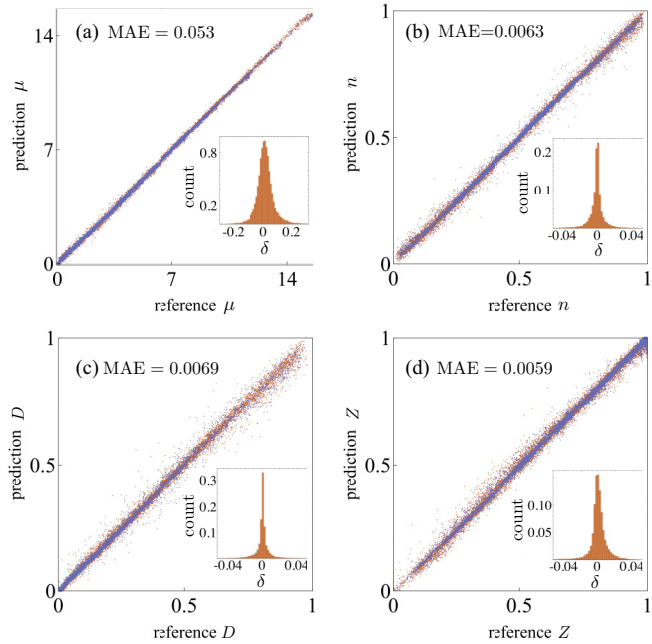

FIG. 3. Comparison of the ML predictions with references obtained from the GA solvers, for (a) the local potential renormalization μ_i, (b) site electron density n_i, (c) probability of double occupation D_i, and (d) local quasiparticle weight Z_i. The blue and orange data points denote predictions for training and test data sets, respectively. The insets show the normalized count of the error δ defined as the difference between prediction and reference values.

minimum batch size of 100. We use randomly mixed 900 000 data samples as the training set and perform a fivefold cross-validation during the training. The Glorot uniform initializer [65] and Adam optimizer [66] with a learning rate of 0.000 01 is applied for the training process. Once the training process is successful, the trained neural network can rapidly predict the 237 600 test data samples. Figure 3 compares the ML prediction with the GA solutions for all accumulated configurations, i.e., those used in the training phase and the remaining configurations used for validation. For all four local quantities, the NN gives rather good predictions as attested by the small MAE, which is of the order of less than 1% of the mean values for all quantities.

IV. DISCUSSION AND OUTLOOK

To summarize, we have introduced a ML model for predicting the local electron correlation of the Anderson-Hubbard Hamiltonian based on training a deep multitask NN in configuration space. In order to describe the spatial inhomogeneity of the electronic structure, we use the real-space Gutzwiller method to numerically solve the AH model on a square lattice. Using the disorder potential in the neighborhood as the input, our ML trained NN is able to predict local electron properties such as double occupancy and quasiparticle weight. Interesting phenomena such as the correlation-induced screening of the disorder potential and the local Mott transition can be accurately predicted by our ML model. Our work provides a proof-of-principle study showing that deep NNs can serve as an efficient many-body problem solver for the large-scale computation of strongly correlated systems. For example, instead of the Gutzwiller solutions, one can train the NNs with data sets obtained from the real-space DMFT or the VMC methods for the AH model. Although more computational effort is required to generate the training data, more accurate predictions can be achieved with the resultant NN model.

As discussed above, a primary motivation for ML trained NN is to bypass the expensive DFT calculation that is required in simulations such as *ab initio* molecular dynamics. Similarly, our proposed ML model as an efficient GA solver also has a direct application for the molecular dynamics simulations of the so-called Holstein-Hubbard model [67–69], in which the site potential $\varepsilon_i = -gX_i$ is related to the amplitude of the local phonon mode X_i, where g is the electron-phonon coupling constant. In such simulations [68], forces acting on the local elastic modes are proportional to the local electron density $F_i = g\,n_i$, which can be efficiently computed using the trained NN. Another related application is to the recently proposed Gutzwiller molecular dynamics (GMD) [70]. The atomic forces in this method are computed from the optimized Gutzwiller many-electron wave function at every time step. Contrary to DFT-based molecular dynamics, GMD simulations allow one to investigate the effects of electron correlation on atomic structural dynamics [70]. Our work shows that ML techniques can be applied to develop a NN that efficiently emulates a GA solver. Preliminary results [71] indeed show that ML is a promising approach for such applications.

ACKNOWLEDGMENTS

We thank Kipton Barros for useful discussions on ML methods. G.W.C. thanks Vladmir Dobrosavljević for insightful discussions on slave-boson and Gutzwiller methods for disordered Hubbard models. P.Z. and G.W.C. are partially supported by the Center for Materials Theory as a part of the Computational Materials Science (CMS) program, funded by the U.S. Department of Energy, Office of Science, Basic Energy Sciences, Materials Sciences and Engineering Division. J.M., Y.T., and A.W.G. are thankful for support from NSF-DMREF 1235230. The authors also acknowledge the Advanced Research Computing Services at the University of Virginia for providing technical support that has contributed to the results in this paper.

[1] T. M. Mitchell, *Machine Learning* (McGraw-Hill, New York, 1997).
[2] B. Schölkopf and A. J. Smola, *Learning with Kernels* (MIT Press, Cambridge, MA, 2002).
[3] M. Nielsen, *Neural Networks and Deep Learning* (Determination Press, 2015).
[4] R. Duda, P. Hart, and D. Stork, *Pattern Classification*, 2nd ed. (Wiley, New York, 2001).

[5] C. Bishop, *Pattern Recognition and Machine Learning* (Springer, Berlin, 2006).

[6] B. G. Sumpter and D. W. Noid, On the design, analysis, and characterization of materials using computational neural networks, Annu. Rev. Mater. Sci. **26**, 223 (1996).

[7] S. V. Kalinin, B. G. Sumpter, and R. K. Archibald, Big-deep-smart data in imaging for guiding materials design, Nat. Mater. **14**, 973 (2015).

[8] P. V. Balachandran, B. Kowalski, A. Sehirlioglu, and T. Lookman, Experimental search for high-temperature ferroelectric perovskites guided by two-step machine learning, Nat. Commun. **9**, 1668 (2018).

[9] W. F. Reinhart, A. W. Long, M. P. Howard, A. L. Ferguson, and A. Z. Panagiotopoulos, Machine learning for autonomous crystal structure identification, Soft Matter **13**, 4733 (2017).

[10] C. Dietz, T. Kretz, and M. H. Thoma, Machine-learning approach for local classification of crystalline structures in multiphase systems, Phys. Rev. E **96**, 011301(R) (2017).

[11] N. Lubbers, T. Lookman, and K. Barros, Inferring low-dimensional microstructure representations using convolutional neural networks, Phys. Rev. E **96**, 052111 (2017).

[12] E. D. Cubuk, S. S. Schoenholz, J. M. Rieser, B. D. Malone, J. Rottler, D. J. Durian, E. Kaxiras, and A. J. Liu, Identifying Structural Flow Defects in Disordered Solids Using Machine-Learning Methods, Phys. Rev. Lett. **114**, 108001 (2015).

[13] S. S. Schoenholz, E. D. Cubuk, D. M. Sussman, E. Kaxiras, and A. J. Liu, A structural approach to relaxation in glassy liquids, Nat. Phys. **12**, 469 (2016).

[14] L. Wang, Discovering phase transitions with unsupervised learning, Phys. Rev. B **94**, 195105 (2016).

[15] J. Carrasquilla and R. G. Melko, Machine learning phases of matter, Nat. Phys. **13**, 431 (2017).

[16] S. J. Wetzel, Unsupervised learning of phase transitions: From principal component analysis to variational autoencoders, Phys. Rev. E **96**, 022140 (2017).

[17] W. Hu, R. R. P. Singh, and R. T. Scalettar, Discovering phases, phase transitions, and crossovers through unsupervised machine learning: A critical examination, Phys. Rev. E **95**, 062122 (2017).

[18] K. Ch'ng, J. Carrasquilla, R. G. Melko, and E. Khatami, Machine Learning Phases of Strongly Correlated Fermions, Phys. Rev. X **7**, 031038 (2017).

[19] P. Broecker, J. Carrasquilla, R. G. Melko, and S. Trebst, Machine learning quantum phases of matter beyond the fermion sign problem, Sci. Rep. **7**, 8823 (2017).

[20] Y. Zhang and E.-A. Kim, Quantum Loop Topography for Machine Learning, Phys. Rev. Lett. **118**, 216401 (2017).

[21] F. Schindler, N. Regnault, and T. Neupert, Probing many-body localization with neural networks, Phys. Rev. B **95**, 245134 (2017).

[22] G. Torlai and R. G. Melko, Learning thermodynamics with Boltzmann machines, Phys. Rev. B **94**, 165134 (2016).

[23] G. Carleo and M. Troyer, Solving the quantum many-body problem with artificial neural networks, Science **355**, 602 (2017).

[24] D.-L. Deng, X. Li, and S. Das Sarma, Quantum Entanglement in Neural Network States, Phys. Rev. X **7**, 021021 (2017).

[25] F. Brockherde, L. Li, K. Burke, and K.-R. Müller, Bypassing the Kohn-Sham equations with machine learning, Nature Commun. **8**, 872 (2017).

[26] J. C. Snyder, M. Rupp, K. Hansen, K.-R. Müller, and K. Burke, Finding Density Functionals with Machine Learning, Phys. Rev. Lett. **108**, 253002 (2012).

[27] J. C. Snyder, M. Rupp, K. Hansen, L. Blooston, K.-R. Müller, and K. Burke, Orbital-free bond breaking via machine learning, J. Chem. Phys. **139**, 224104 (2013).

[28] L. Li, J. C. Snyder, I. M. Pelaschier, J. Huang, U.-N. Niranjan, P. Duncan, M. Rupp, K.-R. Müller, and K. Burke, Understanding machine-learned density functionals, Int. J. Quantum Chem. **116**, 819 (2016).

[29] K. Yao and J. Parkhill, Kinetic energy of hydrocarbons as a function of electron density and convolutional neural networks, J. Chem. Theory Comput. **12**, 1139 (2016).

[30] L. Li, T. E. Baker, S. R. White, and K. Burke, Pure density functional for strong correlation and the thermodynamic limit from machine learning, Phys. Rev. B **94**, 245129 (2016).

[31] K. T. Schütt, F. Arbabzadah, S. Chmiela, K. R. Müller, and A. Tkatchenko, Quantum-chemical insights from deep tensor neural networks, Nat. Commun. **8**, 13890 (2017).

[32] M. Rupp, A. Tkatchenko, K.-R. Müller, and O. A. von Lilienfeld, Fast and Accurate Modeling of Molecular Atomization Energies with Machine Learning, Phys. Rev. Lett. **108**, 058301 (2012).

[33] K. Hansen, G. Montavon, F. Biegler, S. Fazli, M. Rupp, M. Scheffler, O. A. von Lilienfeld, A. Tkatchenko, and K. B. Müller, Assessment and validation of machine learning methods for predicting molecular atomization energies, J. Chem. Theory Comput. **9**, 3404 (2013).

[34] J. Behler and M. Parrinello, Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces, Phys. Rev. Lett. **98**, 146401 (2007).

[35] A. P. Bartók, M. C. Payne, R. Kondor, and G. Csányi, Gaussian Approximation Potentials: The Accuracy of Quantum Mechanics, without the Electrons, Phys. Rev. Lett. **104**, 136403 (2010).

[36] Z. Li, J. R. Kermode, and A. De Vita, Molecular Dynamics with On-the-Fly Machine Learning of Quantum-Mechanical Forces, Phys. Rev. Lett. **114**, 096405 (2015).

[37] V. Botu and R. Ramprasad, Learning scheme to predict atomic forces and accelerate materials simulations, Phys. Rev. B **92**, 094306 (2015).

[38] S. Chmiela, A. Tkatchenko, H. E. Sauceda, I. Poltavsky, K. T. Schütt, and K.-R. Müller, Machine learning of accurate energy-conserving molecular force fields, Sci. Adv. **3**, e1603015 (2017).

[39] L.-F. Arsenault, A. Lopez-Bezanilla, O. A. von Lilienfeld, and A. J. Millis, Machine learning for many-body physics: The case of the Anderson impurity model, Phys. Rev. B **90**, 155136 (2014).

[40] J. Liu, Y. Qi, Z. Y. Meng, and L. Fu, Self-learning Monte Carlo method, Phys. Rev. B **95**, 041101(R) (2017).

[41] X. Y. Xu, Y. Qi, J. Liu, L. Fu, and Z. Y. Meng, Self-learning quantum Monte Carlo method in interacting fermion systems, Phys. Rev. B **96**, 041119(R) (2017).

[42] L. Huang and L. Wang, Accelerated Monte Carlo simulations with restricted Boltzmann machines, Phys. Rev. B **95**, 035105 (2017).

[43] D. Heidarian and N. Trivedi, Inhomogeneous Metallic Phase in a Disordered Mott Insulator in Two Dimensions, Phys. Rev. Lett. **93**, 126401 (2004).

[44] H. Shinaoka and M. Imada, Single-particle excitations under coexisting electron correlation and disorder: A numerical study of the Anderson-Hubbard model, J. Phys. Soc. Jpn. **78**, 094708 (2009).

[45] M. Ulmke, V. Janis, and D. Vollhardt, Anderson-Hubbard model in infinite dimensions, Phys. Rev. B **51**, 10411 (1995).

[46] M. Ulmke and R. T. Scalettar, Magnetic correlations in the two-dimensional Anderson-Hubbard model, Phys. Rev. B **55**, 4149 (1997).

[47] M. E. Pezzoli, F. Becca, M. Fabrizio, and G. Santoro, Local moments and magnetic order in the two-dimensional Anderson-Mott transition, Phys. Rev. B **79**, 033111 (2009).

[48] M. E. Pezzoli and F. Becca, Ground-state properties of the disordered Hubbard model in two dimensions, Phys. Rev. B **81**, 075106 (2010).

[49] V. Dobrosavljević and G. Kotliar, Mean Field Theory of the Mott-Anderson Transition, Phys. Rev. Lett. **78**, 3943 (1997).

[50] V. Dobrosavljević, D. Tanasković, and A. A. Pastor, Glassy Behavior of Electrons Near Metal-Insulator Transitions, Phys. Rev. Lett. **90**, 016402 (2003).

[51] K. Byczuk, W. Hofstetter, and D. Vollhardt, Mott-Hubbard Transition versus Anderson Localization in Correlated Electron Systems with Disorder, Phys. Rev. Lett. **94**, 056404 (2005).

[52] M. C. O. Aguiar, V. Dobrosavljević, E. Abrahams, and G. Kotliar, Critical Behavior at the Mott-Anderson Transition: A Typical-Medium Theory Perspective, Phys. Rev. Lett. **102**, 156402 (2009).

[53] A. Georges, G. Kotliar, W. Krauth, and M. J. Rozenberg, Dynamical mean-field theory of strongly correlated fermion systems and the limit of infinite dimensions, Rev. Mod. Phys. **68**, 13 (1996).

[54] V. Dobrosavljević, A. A. Pastor, and B. K. Nikolić, Typical medium theory of Anderson localization: A local order parameter approach to strong-disorder effects, Europhys. Lett. **62**, 76 (2003).

[55] D. Tanasković, V. Dobrosavljević, E. Abrahams, and G. Kotliar, Disorder Screening in Strongly Correlated Systems, Phys. Rev. Lett. **91**, 066603 (2003).

[56] E. C. Andrade, E. Miranda, and V. Dobrosavljević, Electronic Griffiths Phase of the $d = 2$ Mott Transition, Phys. Rev. Lett. **102**, 206403 (2009).

[57] E. C. Andrade, E. Miranda, and V. Dobrosavljević, Quantum Ripples in Strongly Correlated Metals, Phys. Rev. Lett. **104**, 236401 (2010).

[58] M. C. Gutzwiller, Effect of Correlation on the Ferromagnetism of Transition Metals, Phys. Rev. Lett. **10**, 159 (1963).

[59] G. Kotliar and A. E. Ruckenstein, New Functional Integral Approach to Strongly Correlated Fermi Systems: The Gutzwiller Approximation as a Saddle Point, Phys. Rev. Lett. **57**, 1362 (1986).

[60] N. Lanatá, H. U. R. Strand, X. Dai, and B. Hellsing, Efficient implementation of the Gutzwiller variational method, Phys. Rev. B **85**, 035133 (2012).

[61] N. Lanatá, Y.-X. Yao, C.-Z. Wang, K.-M. Ho, and G. Kotliar, Phase Diagram and Electronic Structure of Praseodymium and Plutonium, Phys. Rev. X **5**, 011008 (2015).

[62] V. Nair and G. E. Hinton, in *Proceedings of the 27th International Conference on International Conference on Machine Learning, ICML'10* (Omnipress, Madison, WI, 2010), pp. 807–814.

[63] R. Caruana, Multitask learning, Mach. Learn. **28**, 41 (1997).

[64] A. Y. Ng, in *Proceedings of the Twenty-first International Conference on Machine Learning, ICML'04* (ACM, New York, 2004), p. 78.

[65] X. Glorot and Y. Bengio, Understanding the difficulty of training deep feedforward neural networks, Proc. Mach. Learn. Res. **9**, 249 (2010).

[66] D. P. Kingma and J. Ba, Adam: A method for stochastic optimization, arXiv:1412.6980.

[67] T. Holstein, Studies of polaron motion: Part I. The molecular-crystal model, Ann. Phys. **8**, 325 (1959).

[68] S. Pradhan and G. V. Pai, Holstein-Hubbard model at half filling: A static auxiliary field study, Phys. Rev. B **92**, 165124 (2015).

[69] D. Di Sante, S. Fratini, V. Dobrosavljević, and S. Ciuchi, Disorder-Driven Metal-Insulator Transitions in Deformable Lattices, Phys. Rev. Lett. **118**, 036602 (2017).

[70] G.-W. Chern, K. Barros, C. D. Batista, J. D. Kress, and G. Kotliar, Mott Transition in a Metallic Liquid: Gutzwiller Molecular Dynamics Simulations, Phys. Rev. Lett. **118**, 226401 (2017).

[71] H. Suwa, J. S. Smith, N. Lubbers, C. D. Batista, G.-W. Chern, and K. Barros, Machine learning for molecular dynamics with strongly correlated electrons, arXiv:1811.01914 (2018).