A perspective on machine learning and data science for strongly correlated electron problems

Steven Johnston [a,b,*], Ehsan Khatami [c], Richard Scalettar [d]

[a] Department of Physics and Astronomy, The University of Tennessee, Knoxville, TN 37996, USA
[b] Institute for Advanced Materials and Manufacturing, University of Tennessee, Knoxville, TN 37996, USA
[c] Department of Physics and Astronomy, San Jose State University, San Jose, CA 95192, USA
[d] Department of Physics and Astronomy, University of California, Davis, CA 95616, USA

ARTICLE INFO

ABSTRACT

Numerical approaches to the correlated electron problem have achieved considerable success, yet are still constrained by several bottlenecks, including high order polynomial or exponential scaling in system size, long autocorrelation times, challenges in recognizing novel phases, and the Fermion sign problem. Methods in machine learning (ML), artificial intelligence, and data science promise to help address these limitations and open up a new frontier in strongly correlated quantum system simulations. In this paper, we review some of the progress in this area. We begin by examining these approaches in the context of classical models, where their underpinnings and application can be easily illustrated and benchmarked. We then discuss cases where ML methods have enabled scientific discovery. Finally, we will examine their applications in accelerating model solutions in state-of-the-art quantum many-body methods like quantum Monte Carlo and discuss potential future research directions.

1. Introduction

The development of nonperturbative numerical methods and steady growth in available computational power achieved over the last half-century have provided us with the tools necessary to obtain reliable solutions to simplified correlated electron models. Through the combined use of methods like quantum Monte Carlo (QMC), density matrix renormalization group (DMRG), dynamical mean-field theory (DMFT) and its cluster extensions, and others, we now have a great deal of insight into the physics of the single-band Hubbard [1–18], Holstein [19–25], Su–Schrieffer–Heeger [26–33], and periodic Anderson models [34–39]. Substantial progress is also being made towards understanding models where several different, competing, interactions are present [40–49], as well as those that extend beyond the "standard" interactions [49–55]. For example, while not everything is settled, it is now clear that the single-band Hubbard model in the intermediate coupling regime,[1] is a minimal model for the low-temperature electronic properties of cuprates, and many aspects of their phase diagram [3–5,7–9,11,12,15–18,48,56] (with the important possible exception of superconductivity [9,13,14]). This knowledge, however, was only won after decades of method development and research studies by groups from around the world.

While impressive progress has been made toward understanding multi-orbital extensions of these models, achieving the same level of success for other strongly correlated materials will require new approaches. For example, the suitable minimal effective models remain unknown in many cases. And even when they are known, they often cannot be treated using existing numerical methods owing to issues like the Fermion "sign problem" [57–64].

Many researchers hope that methods in data science, machine learning (ML), and artificial intelligence (AI) can be exploited to reach the next stage of quantum simulations. Research in this area has developed along with several particularly opportune areas, including (but not limited to) using ML to (1) implement general schemes to perform "global moves" – analogs of loop and Wolff–Swendsen–Wang methods [65–67] – in classical and QMC simu-

* Corresponding author at: Department of Physics and Astronomy, The University of Tennessee, Knoxville, TN 37996, USA.
E-mail address: sjohn145@utk.edu (S. Johnston).

[1] $U/W \approx 1$ where U is the strength of the Hubbard interaction and W is the electronic bandwidth, see Section 2.

lations that update many degrees of freedom simultaneously with high acceptance rates; (2) recognize phase transitions in models, especially in cases when the relevant order parameter(s) are unknown; (3) determine optimal constrained wave functions in variational approaches to deal with the sign problem with the minimal introduction of bias; and (4) extract suitable low-energy effective models from experimental data or high-energy models.

In response to this rapidly emerging field, Oak Ridge National Laboratory held a workshop on *Artificial Intelligence in Multi-Fidelity, Multi-Scale and Multi-Physics Simulations of Materials* in August of 2021 as part of the Joint Nanoscience and Neutron Scattering User Meeting. This special issue on *Digital Twins in Materials and Chemical Sciences* is the result of that workshop. At that meeting, we gave consecutive talks on applications of ML for the many-body problem, drawing from our research areas. In that spirit, this article synthesizes our talks in a mini-review, focusing on where these methods intersect with our work and interests in magnetic, charge, and pairing order in classical and quantum many-body systems. The fields of correlated electrons and ML applications are broad and rapidly evolving. We have not attempted to be exhaustive in our references or discussion, especially given the rapid pace at which this field is developing. For the interested reader, we note that several comprehensive reviews and books of ML applications in physics have recently been published [68–74]. We apologize in advance for any work that we have left omitted.

The organization of this paper is as follows. Section 2 will provide a brief overview of the models discussed throughout this paper. Since the majority of our research involving ML methods focuses on Markov Chain Monte Carlo methods, Section 3 provides a brief description, including the path integral mapping of the quantum partition function in d spatial dimensions to an equivalent classical problem in $d + 1$ dimensions. Section 4 then provides an overview of the ML methods discussed throughout this work while Section 5 discusses several proof-of-concept studies applying these methods to problems in statistical and correlated electron physics. We then highlight a few new scientific discoveries that have been enabled by ML methods in Section 6. Section 7 focuses on self-learning Monte Carlo methods, emphasizing recent applications to the Holstein model as an example. Finally, Section 8 provides our perspective on future work in this area.

2. Models

We will discuss several ML and data science applications to understanding correlated electron systems, often tested on canonical model Hamiltonians. This section provides an overview of the relevant models to acquaint the reader and establish our notation, beginning with classical systems. Due to the nature of this review, our discussion will be brief. The reader familiar with this material can safely skip to the next section.

2.1. The Ising model

The Ising Hamiltonian

$$H = -\sum_{ij} J_{ij} s_i s_j - B \sum_i s_i, \tag{1}$$

describes a collection of localized classical degrees of freedom, which are often understood to represent the magnetic moments (or spins) of individual atoms on a lattice of sites $i = 1, 2, \ldots, N$. These spins are constrained to point in one of two directions such that $s_i = \pm 1$. Each spin individually couples to an external magnetic field B so that the orientation $s_i = +1$ is energetically favored. Pairs of spins interact via a magnetic coupling J_{ij} between different lattice sites. For $J_{ij} > 0 \,(< 0)$, the interactions are ferromagnetic (antiferromagnetic) in nature.

The Ising model is perhaps the simplest model exhibiting spontaneous symmetry breaking, which occurs at a phase transition. For example, in the *absence* of a field $B = 0$, the total energy is invariant under a global change of variables $s_i \leftrightarrow -s_i$. Nevertheless in $d \geq 2$, the system falls into one of these two equivalent states as T is lowered, and for the ferromagnetic case, the magnetization $m = 1/N \sum_i \langle s_i \rangle$ becomes non-zero. This behavior occurs only in the thermodynamic limit, $N \to \infty$, so one of the challenges of Monte Carlo and in the use of ML is in the finite-size scaling needed to extrapolate finite N simulations.

The Ising model has known analytical solutions in certain cases [75]. For example, on a square lattice for nearest-neighbor ferromagnetic coupling and in zero field, it undergoes a second-order transition between paramagnetic ($m = 0$) and ferromagnetic ($m \neq 0$) phases at $T_c = 2.269J$. The correlation length ξ, which measures the decay in the magnetic correlations $\langle s_i s_{i+r} \rangle \sim e^{-r/\xi} \langle s_i s_{i+r} \rangle \sim e^{-r/\xi}$, diverges at this transition as $\xi \sim 1/(T - T_c)^\nu$, as does the magnetic susceptibility $\chi = [\langle m^2 \rangle - \langle m \rangle^2]/T \sim 1/(T - T_c)^\gamma$. The exponents $\nu = 1$ and $\gamma = 7/4$ are known analytically, as is the exponent $\beta = 1/8$ describing the onset of nonzero magnetization $m \sim (T_c - T)^\beta m \sim (T_c - T)^\beta$ below T_c. There is no known analytic solution for the Ising model in $d > 2$.

The Ising model combines simplicity with deep conceptual physics. It also provides a textbook application for Markov chain Monte Carlo methods, through which researchers can readily produce training and validation data in large quantities. The fact that the model has an exact solution in $d \leq 2$ while being unsolved in general means that researchers can use it as a high-precision benchmark for new numerical approaches while also offering the opportunity to sharpen our knowledge of unknown critical points and exponents. These aspects have positioned the Ising model as an essential test for many new ML-based approaches [76–82].

Because the (ferromagnetic) Ising model on other 2D lattices falls into the same universality class (i.e. identical critical exponents), it is probably sufficient to consider the square lattice in testing ML approaches to this problem. It is worth noting, however, that realizations of the Ising model on different lattices exist, which offer rich opportunities to compare traditional simulation methods with ML-driven methodologies. These include Kagome and triangular lattices [83–85], which exhibit frustration in the antiferromagnetic (AF) case of $J_{ij} < 0$, and spin-glass behavior when the J_{ij} are chosen with random ferromagnetic and antiferromagnetic signs [86–88].

2.2. The Blume–Capel model

The Blume–Capel model [89,90] is a generalization of the Ising model, where the magnetic moments can align parallel, antiparallel, or orthogonal to an external magnetic field. Its Hamiltonian is

$$H = -\Delta \sum_i \left(1 - s_i^2\right) - \sum_{i,j} J_{ij} s_i s_j - B \sum_i s_i. \tag{2}$$

where $s_i = 0, \pm 1$ and Δ is the zero-field splitting, which measures the energy difference between the singlet $s_i = 0$ and doublet $s_i = \pm 1$ moments.

The parameter Δ controls the density of $s_i = 0$ sites. When $\Delta \to -\infty$ such sites are energetically highly unfavorable. The Blume–Capel model approaches the Ising model in this limit, exhibiting the same second-order magnetic phase transition. As Δ increases, more $s_i = 0$ sites are introduced and T_c decreases. Ultimately, at $\Delta = J/2$ there is no long-range order at any $T > 0$, i.e. the critical temperature $T_c = 0$. A fascinating feature of this model is that along the phase boundary in the $T - \Delta$ plane, the nature of the transition changes from second to first order at a "tricritical point." Thus the Blume–Capel model offers several new vistas for ML: first, in providing a simple model to explore methods

that can locate a one-dimensional locus of critical temperatures, as opposed to an isolated critical point, and second, as means to examine whether these methods can distinguish different types of phase transitions.

2.3. The Hubbard model

The Hubbard Hamiltonian [91] is a simple model for describing the effect of electron-electron interactions on itinerant lattice electrons. In the case of a single band,

$$\hat{H} = \hat{H}_0 + \hat{H}_U, \tag{3}$$

where

$$\hat{H}_0 = \sum_{i,j,\sigma} t_{i,j} c_{i,\sigma}^\dagger c_{j,\sigma} - \mu \sum_{i,\sigma} \hat{n}_{i,\sigma} \tag{4}$$

describes electrons propagating through a one-orbital lattice and

$$\hat{H}_U = U \sum_i \hat{n}_{i,\uparrow} \hat{n}_{i,\downarrow} \tag{5}$$

is a local Coulomb interaction between electrons on the same site. Here, $c_{i,\sigma}^\dagger$ ($c_{i,\sigma}$) creates (destroys) an electron with spin σ on lattice site i, $\hat{n}_{i,\sigma}$ is the Fermion number operator at site i, μ is the chemical potential, $t_{i,j}$ is hopping integral between sites i and j, and U is the strength of the Hubbard interaction, which can either be repulsive ($U > 0$) or attractive ($U < 0$). The range of values of the energy levels of H_0 is referred to as the bandwidth W.

Despite its simplicity, analytical solutions to the Hubbard model have only been obtained in one dimension [92]. In higher dimensions, the most reliable results have been obtained with nonperturbative numerical methods, particularly in the intermediate coupling regime $U/W \approx 1$. Nevertheless, significant progress has been made toward understanding the physics of the Hubbard model in this regime. For example, the doped two-dimensional Hubbard model contains robust antiferromagnetic Mott correlations [1,18,93], spin- and charge-stripes [8,11–13,16,18,94,95], strange metal behavior [64,96], a pseudogap [18,93,97], and unconventional superconductivity [9,97–99].

The Hubbard model can also be easily generalized to include longer-range interactions,

$$\hat{H}_V = \sum_{i \neq j, \sigma, \sigma'} V_{i,j} \hat{n}_{i,\sigma} \hat{n}_{j,\sigma'}, \tag{6}$$

or additional orbital degrees of freedom by modifying the underlying tight-binding model. From a computational perspective, the inclusion of further orbitals is 'trivial' in the determinant QMC formalism on which we focus - the orbital index plays an identical role as a site label. However, longer-range interactions have a dramatic effect. They necessitate a significant restructuring of the Hubbard–Stratonovich transformation, described below, and make the fermion sign problem much worse.

2.4. The Holstein model

The Holstein Hamiltonian is a simplified model for describing electrons coupled to the lattice [100]. Like the Hubbard model, it approximates the noninteracting electronic degrees of freedom using a single orbital tight-binding model. The noninteracting lattice degrees of freedom are described using simple harmonic oscillators at each site, while the electron-lattice interaction is introduced through a linear coupling between the on-site electron density and the lattice displacement. The corresponding Hamiltonian is

$$\hat{H} = \hat{H}_0 + \hat{H}_{\text{lat}} + \hat{H}_{e-\text{lat}}, \tag{7}$$

where H_0 describes the electronic degrees of freedom as in Eq. (4),

$$\hat{H}_{\text{lat}} = \sum_i \left[\frac{\hat{P}_i^2}{2M} + \frac{1}{2} M \omega_0^2 \hat{X}_i^2 \right] \tag{8}$$

describes the noninteracting lattice degrees of freedom and

$$\hat{H}_{e-\text{lat}} = \sum_{i,\sigma} \alpha \, \hat{n}_{i,\sigma} \hat{X}_i \tag{9}$$

describes the electron-phonon interaction. Here, $\hat{P}_i$ and $\hat{X}_i$ are the momentum and position operators for the oscillator at site i, ω_0 is the energy of the oscillator ($\hbar = 1$), and α is the electron-phonon coupling strength.

Two dimensionless ratios are frequently quoted when simulating the Holstein model. The first is the so-called adiabatic ratio ω_0/E_F, which measures the relative energies of the lattice and electron degrees of freedom. The second is the dimensionless e-ph coupling constant $\lambda = 2\alpha^2 / (MW\omega_0^2)$, where W is again the noninteracting bandwidth.

Like the Hubbard model, the Holstein model exhibits a rich collection of phases. The model has a metal to a charge-density-wave (CDW) insulator transition near half-filling [24,101], conventional superconductivity away from half-filling [102], and polaron and bipolaron formation [23,24,103]. In the ML context, this model provides an excellent platform for testing new algorithms where electrons are coupled to continuous fields (the lattice displacements) and where there can be significant differences in the time scales associated with the electron and lattice dynamics. It is also worth noting that a fast and scalable hybrid quantum Monte Carlo algorithm has recently been made available for studying Holstein models and other electron-phonon coupled systems [104]. This approach can treat these models on large system sizes, thus providing additional opportunities for testing proposed ML-accelerated algorithms.

2.5. The periodic Anderson model

The periodic Anderson model (PAM) is a variant of the Hubbard Hamiltonian containing two orbitals per site. In this case, one orbital ($c_{i,\sigma}$) is 'metallic' and has no on-site interaction, while the other orbital ($f_{i,\sigma}$) is 'localized' and has a large on-site Hubbard repulsion $U_f \neq 0$. Its Hamiltonian is

$$\hat{H} = -t \sum_{\langle i,j \rangle} c_{i,\sigma}^\dagger c_{j,\sigma} - \mu \sum_{i,\sigma} \left[n_{i,\sigma}^c + n_{i,\sigma}^f \right]$$
$$+ V \sum_{i,\sigma} \left[c_{i,\sigma}^\dagger f_{i,\sigma} + h.c. \right] + U_f \sum_i n_{i,\uparrow}^f n_{i,\downarrow}^f, \tag{10}$$

where V is the hybridization between the localized and metallic orbitals and $n_{i,\sigma}^c = c_{i,\sigma}^\dagger c_{i,\sigma}$ and $n_{i,\sigma}^f = f_{i,\sigma}^\dagger f_{i,\sigma}$ are the number operators for each type of orbital.

In its ground state and at half-filling, the PAM undergoes a quantum phase transition as a function of V. Here, the system transitions from a state with antiferromagnetic order on the f-orbitals at small V to a singlet phase at large V. In the latter, the conduction and localized f electrons form local spin-0 singlets on a site, which disrupts the long-range antiferromagnetic order. Previous determinant QMC studies have placed the quantum critical point (QCP) for this transition at $V_c \sim tV_c \sim t$ for $U_f \sim 4tU_f \sim 4t$ [35].

3. Monte Carlo methods

3.1. Overview

Markov chain Monte Carlo (MCMC) methods are a powerful class of algorithms for simulating physical systems and have found widespread use in condensed matter physics [105,106]. These techniques perform a random walk through some configuration space while statistically sampling the relevant observables in a way that guarantees the correct probability distribution is generated.
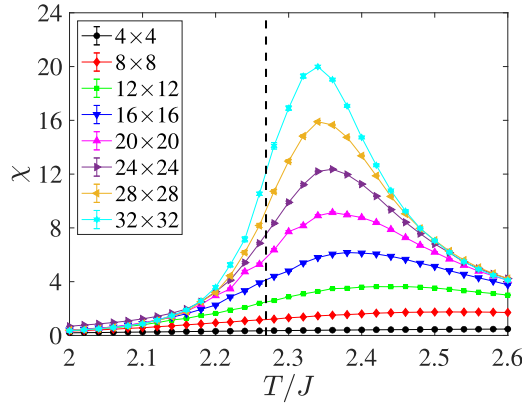
Fig. 1. The magnetic susceptibility $\chi = d\langle M \rangle / dB = \beta(\langle M^2 \rangle - \langle M \rangle^2)$ of the 2D Ising model using a primitive single spin flip Metropolis Monte Carlo algorithm. χ is maximal in the vicinity of the known $T_c = 2.269$, and the peak shifts towards T_c as the lattice size N grows.

Fig. 2. Analysis of the crossing of the Binder ratio $B = 1 - \langle M^4 \rangle / 3 \langle M^2 \rangle^2$ gives a more accurate estimate of T_c. The data here could be vastly improved by using a cluster method to avoid critical slowing down; however, even with this straightforward implementation, T_c is given to accuracy of one percent or so, with simulation run times on the order of minutes on a desktop computer.

3.2. Classical Monte Carlo

In a classical Monte Carlo simulation, one aims to evaluate the thermodynamic expectation value of an observable O with respect to a set of microstates $|m\rangle$ that follow a Boltzmann probability distribution,

$$\langle O \rangle = \frac{1}{\mathcal{Z}} \sum_m e^{-\beta E_m} \langle m | O | m \rangle. \tag{11}$$

Here, E_m denotes the energy of the microstate, O_m denotes the value of the observable for the microstate, and $\mathcal{Z} = \sum_m \exp(-\beta E_m)$ is the partition function.

The sum in Eq. (11) must be taken over all accessible microstates, which is intractable for most systems of interest. Instead, one uses MCMC methods to evaluate the sum stochastically. In a classical simulation of the Ising model, for example, a standard procedure is to start with a random lattice of up or down spins and then select individual spins to flip. These flips may be done either by visiting each site in a random order or by going through the lattice in some specified pattern. The change in energy ΔE that is induced by flipping the spin is then evaluated, and the proposed move is accepted with probability $p = \min(1, e^{-\beta \Delta E})$, where $\beta = 1/T$ is the inverse temperature in units where $k_B = 1$. This prescription is the 'Metropolis-Hastings' algorithm [107], and it ensures that the statistical distribution of the spins follows a Boltzmann distribution. Alternatively, one can use a 'heat-bath' probability $p = e^{-\beta \Delta E} / (1 + e^{-\beta \Delta E})$.

A crucial feature of classical Monte Carlo is that updating the entire lattice scales *linearly* with system size N, making simulations of large lattices practical.[2] This linear scaling follows from the locality of the energy, which implies that evaluating ΔE is independent of N. In contrast, QMC generally scales as N^3 for interacting fermions. As we shall discuss later, this scaling results from the fact that the action determining the probability is *non-local*, involving the determinant of a matrix of dimension N.

Fig. 1 shows a typical result for a classical simulation obtained using the traditional classical Monte Carlo methodology. In this case, the magnetic susceptibility χ of the 2D Ising model is plotted as a function of temperature. Scale-invariant measurements such

as the Binder ratio provides improved ways to locate the critical point on finite lattices [108], as indicated in Fig. 2. The key take-away is that a determination of the critical temperature to one percent accuracy is easily attained using these simplest classical models, even with unsophisticated single spin flip moves, and typical desktop computers. These models, therefore, form interesting testbeds for ML methods and provided the first evidence of their potential. However, ML, while adding interesting insight, are unlikely to constitute breakthrough applications in this context, given the efficacy of existing tools.

3.3. Quantum Monte Carlo: the DQMC method

This section presents a brief overview of the standard determinant quantum Monte Carlo (DQMC) algorithm while highlighting the critical aspects for understanding the ML applications discussed later. For more complete discussions of the DQMC algorithm, we refer the reader to Refs. [1,47,104,105,109–111].

The goal of DQMC is to evaluate the expectation values of thermodynamic observables of quantum many-body Hamiltonians such as the Hubbard, periodic Anderson, and Holstein models. That is, it allows the computation, within the grand canonical ensemble, of

$$O \equiv \langle \hat{O} \rangle = \frac{1}{\mathcal{Z}} \text{Tr} \, \hat{O} \, e^{-\beta \hat{H}},$$

where $\mathcal{Z} = \text{Tr} \, e^{-\beta \hat{H}}$ is the grand partition function. Observables of interest include the Hamiltonian itself, giving the energy and specific heat, as well as pairing, charge, and spin correlation function and their susceptibilities, which signal transitions into low-temperature ordered phases. We will first focus on evaluating $\mathcal{Z}$. Once the recipe is established, it is straightforward to generalize it to evaluate O.

It is convenient first to partition the Hamiltonian into two parts, $\hat{H} = \hat{H}_0 + \hat{H}_{\text{int}}$, where $\hat{H}_0$ contains the noninteracting terms (i.e. those which are quadratic in the fermion creation and destruction operators) and $\hat{H}_{\text{int}}$ contains the interaction terms. (Note that if lattice terms are present in the model - e.g., as in the Holstein model – then the Hamiltonian is further partitioned as $\hat{H} = \hat{H}_0 + \hat{H}_{\text{lat}} + \hat{H}_{\text{int}}$.) Next, we divide the imaginary time interval into L discrete imaginary time steps such that $\tau = l \Delta \tau$, with $l = 0, 1, \ldots, L - 1$ and $\Delta \tau = \beta / L$. We can then approximate $\mathcal{Z}$ using the Trotter formula [112–114],

$$\mathcal{Z} = \text{Tr} \, e^{-\beta \hat{H}} = \text{Tr} \, (e^{-\Delta \tau \hat{H}_0} e^{-\Delta \tau \hat{H}_{\text{int}}})^L + O(\Delta \tau^2) \tag{12}$$

Here, $O(\Delta \tau^2)$ is a controllable Suzuki-Trotter error introduced by the neglected commutation of the Hamiltonian terms.

---

[2] There can, however, be 'hidden' factors of N. Most commonly, near a critical point, the autocorrelation time τ of the simulation diverges as L^z, where L is the linear system size and z is the dynamical critical exponent of the algorithm being used. z often takes the value $z \sim 2$, leading to very long τ. However, in many cases, special larger-scale "global" or "block" moves have been developed to address this [65–67], leading to $z \sim 0$.

We first consider the case where the model only has local Hubbard interactions. In this case, one needs to recast the quartic terms in the Hamiltonian into quadratic ones by introducing an auxiliary field $s_{i,l}$ at every spatial i ($= 1, 2, \ldots, N$) and imaginary time l point. While many auxiliary fields are possible, in DQMC, one usually adopts the discrete Hubbard–Stratonovich (HS) transformation for an on-site Hubbard interaction [115]. For $U > 0$, $s_{i,l}$ couples to the z-component of spin,

$$e^{-U\Delta\tau\left(n_{i,\uparrow}-\frac{1}{2}\right)\left(n_{i,\downarrow}-\frac{1}{2}\right)} = \frac{1}{2}e^{-\frac{\Delta\tau U}{4}}\sum_{s_{i,l}}e^{\lambda s_{i,l}(n_{i,\uparrow}-n_{i,\downarrow})}$$

and for $U = -|U| < 0$, $s_{i,l}$ couples to the charge,

$$e^{|U|\Delta\tau\left(n_{i,\uparrow}-\frac{1}{2}\right)\left(n_{i,\downarrow}-\frac{1}{2}\right)} = \frac{1}{2}e^{-\frac{\Delta\tau|U|}{4}}\sum_{s_{i,l}}e^{\lambda s_{i,l}(n_{i,\uparrow}+n_{i,\downarrow}-1)}.$$

Regardless of the sign of U, $s_{i,l} = \pm 1$ and λ is a constant satisfying $\cosh(\lambda) = \exp(\Delta\tau|U|/2)$. From this point forward, we will focus on the repulsive case with $U > 0$. The generalization to negative values of U is straightforward.

Once the HS transformation is applied, the partition function only involves quadratic terms in the fermion operators, which comes at the expense of having to sum over the auxiliary fields. For a fixed configuration $\{s_{i,l}\}$, the trace over the fermion degrees of freedom can be evaluated analytically [116] to yield a product of determinants (hence the name). The traces over up and down fermions yield separate determinants as long as $\hat{H}_0$ contains no terms that hybridize the two spin species. The partition function thus reduces to

$$\mathcal{Z} = \sum_{\{s_{i,l}\}}\det(M_\uparrow)\det(M_\downarrow) \equiv \sum_{\{s_{i,l}\}}W(\{s_{i,l}\}). \tag{13}$$

Here, the matrices M_σ are defined as

$$M_\sigma = \mathbb{I} + B_\sigma(L-1)_\sigma B_\sigma(L-2)\ldots B_\sigma(0),$$

where $B_\sigma(l) = e^{\mp\Delta\tau\lambda\nu(l)}e^{-\Delta\tau K}$, K is the matrix representation of H_0, $\mathbb{I}$ is the $N \times N$ identity matrix, and $\nu(l)$ is a diagonal matrix whose elements are given by $[\nu(l)]_{i,j} = \delta_{i,j}s_{i,l}$. The $\mp$ in the exponential refers to spin $\sigma = \uparrow, \downarrow$.

All that remains is to evaluate the summation over all HS configurations appearing in Eq. (13), which is accomplished using standard MCMC methods. A central quantity in this sampling procedure is the equal time Green's function $G^\sigma_{ij}(\tau = l\Delta\tau) = \langle c_{j,\sigma}(\tau)c^\dagger_{i,\sigma}(\tau)\rangle$. For a given auxiliary field configuration, it can be expressed as

$$G^\sigma_{i,j}(l) = [\mathbb{I} + B_\sigma(l)_\sigma B(l-1)\ldots B_\sigma(0) \times B_\sigma(L-1)\ldots B_\sigma(l+1)]^{-1}_{i,j}. \tag{14}$$

This quantity can be interpreted as describing the propagation of a free electron through the potential established by the given HS field configuration. Note that Eq. (14) implies that the Green's function on successive time slices satisfies the equation

$$G^\sigma(l+1) = B_\sigma(l+1)G^\sigma(l)B^{-1}_\sigma(l+1). \tag{15}$$

It also establishes that the weight of the Monte Carlo configuration can be equated to the product of the determinant of the inverse Green's functions

$$W(\{s_{i,l}\}) = \det(G^{-1}_\uparrow)\det(G^{-1}_\downarrow).$$

These relationships form the basis for the MCMC sampling procedure, where one performs a random walk through the configurations $\{s_{i,l}\}$ following the probability distribution $W(\{s_{i,l}\})$. First, the time index is fixed to a particular slice l, and the corresponding equal time Green's functions $G_\sigma(l\Delta\tau)$ are computed. Next, one visits every spatial site i in the lattice, proposing changes in the local auxiliary fields $s_{i,l} \to s'_{i,l}$. (For the case

of the HS fields, this is a local spin flip move $s_{i,l} \to -s_{i,l}$.) These moves can be accepted with the Metropolis probability $p = \min[1, W(\{s'_{i,l}\})/W(\{s_{i,l}\})]$, or with the heat-bath prescription, $p = W(\{s'_{i,l}\})/\left(W(\{s'_{i,l}\}) + W(\{s_{i,l}\})\right)$, both of which obey detailed balance. Similarities with the Monte Carlo procedure for a classical model (like Ising) are evident. Two key differences are (i) that simulation of the original quantum model in d dimensions requires sampling a field $s_{i,l}$ with an 'imaginary time' index l in addition to the spatial site label i; and (ii) the weight involves a non-local quantity - the fermion determinants.

In principle, updating a single $s_{i,l}$ requires the $O(N^3)$ operations required in evaluating a determinant. However, the computation of p involves only the ratio of the determinants, for which there is a simple expression in terms of the equal time Green's functions. Once updates have been proposed at every site, the algorithm advances to the next time slice using Eq. (15), and the process is repeated. Because of the locality of the change in the matrix, the determinant ratios can be evaluated in $O(N^2)$ operations, so that updating all NL components $s_{i,l}$ of the HS field takes $O(N^3L)$ steps. This is the fundamental system-size scaling of the DQMC algorithm.

The Holstein model does not have $e - e$ interaction terms and thus does not necessitate the introduction of the Ising-like HS fields. Instead, one must deal with the e-ph terms, which couple the local (quadratic) fermion density operator to the lattice displacement. We again begin with Eq. (12), but this time we insert a complete set of phonon position and momentum states at each spacetime point (i, l). At this stage, the trace over the phonon momenta and quadratic electron degrees of freedom can be performed analytically, reducing the partition function to a familiar form

$$\mathcal{Z} = \int dX e^{S_{\text{lat}}}\det(M_\uparrow)\det(M_\downarrow) = \int dX\, W(\{X_{i,l}\}). \tag{16}$$

In this case, the matrices $B_\sigma(l) = e^{\Delta\tau gX(l)}e^{-\Delta\tau K}$, where $X(l)$ denotes a diagonal matrix whose diagonal elements are $[X(l)]_{i,j} = X_{i,l}\delta_{i,j}$, $\{X_{i,l}\}$ denotes a given configuration of the lattice sites, and dX is short hand for an $N \times L$ multi-dimensional integral over all lattice displacements $\{X_{i,l}\}$. The structure is very similar to that of the Hubbard Hamiltonian. Note, the additional term $\exp(S_{\text{lat}})$ where

$$S_{\text{lat}} = \Delta\tau\left[\frac{\omega_0^2}{2}\sum_{i,l}X_{i,l} + \sum_{i,l}\left(\frac{X_{i,l+1} - X_{i,l}}{2\Delta\tau}\right)^2\right]$$

in the configuration weight resulting from noninteracting lattice terms $\hat{H}_{\text{lat}}$. Apart from these changes, the remainder of the DQMC algorithm is unchanged.

The evaluation of S_{lat} is very rapid compared to dealing with the fermion determinant, so it contributes very little to the computational workload. Indeed, its simplicity also makes adaptation of the Holstein model to other forms of $\hat{H}_{\text{lat}}$, e.g. including anharmonic terms in the phonon potential energy, almost trivial. S_{lat} does play a profound role in controlling the imaginary time fluctuations of the phonon field in comparison to the HS field, for which an analog of S_{lat} is absent. As a consequence, the fermion matrices are much better conditioned, opening the door for efficient Langevin updates of electron-phonon Hamiltonians [25,31,104], which are much more challenging in the Hubbard model.

It is important to stress that we have skipped over many technical details in this brief overview, which must be addressed when implementing the DQMC algorithm. For example, the $B_\sigma(l)$ matrices are stiff, especially as the product $U\beta$ becomes large. To evaluate the long products of these matrices, one must use QR factorizations and other numerical tricks to stabilize the numer-

ics. Additional details on these technical details can be found in Refs. [1,47,104,105,109–111,117].

We will briefly mention one issue since it is the primary limitation to the application of DQMC: the signs of the configuration weights $\det(M_\uparrow)\det(M_\downarrow)$ are not always positive-definite and, therefore, cannot be directly interpreted as a probability. When this occurs, the auxiliary fields are sampled according to a new probability distribution given by the absolute value of the original distribution $|W(\{s_{i,l}\})|$. The two distributions are related by $W = \text{sign}W\,|W|$. This change requires us to re-weight any observable as

$$\langle O \rangle_W = \frac{\langle O\,\text{sign}W \rangle_{|W|}}{\langle \text{sign}W \rangle_{|W|}}, \qquad (17)$$

where the subscript $|W|$ in the expectation value emphasizes the configurations are now generated with probability $\propto |W|$. When the system size increases or the temperature decreases, the sign of W is positive and negative with nearly equal probability, causing $\langle \text{sign}W \rangle \to 0$. The numerator must also vanish exponentially since $\langle O \rangle$ is well defined. The process of taking the ratio of two very small quantities, each with finite error bars ultimately yields values that are meaningless. This behavior is a reflection of the Fermion sign problem [57,59]. (The sign problem is not restricted to fermionic models, but also occurs for frustrated quantum spins [118].)

It is worth noting that the determinants $\det(M_\uparrow)=\det(M_\downarrow)$ for the Holstein model because the phonon field couples in the same way to the spin up and spin down fermions. This is also true of the attractive Hubbard model at any filling, and of the half-filled Hubbard with nearest-neighbor hopping only. Other 'sign-problem free' models and the symmetries and choices of bases from which they originate have been discovered [58,61–63,119–132]. In such cases, $W(\{X_{i,l}\})$ is always greater than zero and DQMC can access the physics down to arbitrarily low temperatures. In these cases, one can obtain essentially exact solutions of the corresponding quantum many-body problem.

3.4. Challenges and limitations

Despite their power and versatility, MCMC methods are limited in several notable ways. First, they frequently employ finite-size clusters when studying correlated electron models like those discussed in Section 2. Because of this, a finite-size analysis is required to extrapolate results to the thermodynamic limit. The thermodynamic limit can be approached more directly by embedding the cluster in a self-consistent dynamical mean field that approximates correlations beyond the cluster [133,134]. However, such embedding schemes can still exhibit considerable finite-size effects [9]. Because of this, addressing the thermodynamic limit requires large system sizes, which can be prohibitively expensive for many classes of MC algorithms.

Another fundamental limitation of any MC method is its decorrelation time. A MC algorithm must draw measurements from statistically independent configurations to obtain unbiased estimators for an observable and its statistical error. Because of this, the autocorrelation time τ - defined as the number of updates needed to generate such configurations - is a crucial measure of a MC simulation's efficiency. Many MC applications suffer from prohibitively long autocorrelation times (e.g., *e*-ph or frustrated spin models, continuum limit lattice gauge theory simulations, and confined quantum liquids).

Autocorrelation times can also depend strongly on the parameter regime of a particular model and the sampling method. The Holstein model in the adiabatic limit ($\omega_0/t \ll 1$) is quite challenging for traditional Metropolis-Hastings sampling but more amenable to hybrid Monte Carlo methods [104]. Autocorrelation times also tend to grow in the vicinity of a phase transition, where correlation lengths extend beyond the size of the cluster and single-site updates are no longer capable of efficiently moving MC configurations out of meta-stable minima. This latter problem is known as "critical slowing down." In some cases, "global" MC moves can be performed to mitigate the autocorrelation time. In such schemes, an extended region of configuration space is updated simultaneously to generate independent configurations quickly. However, the form of these updates is only known in some special cases. Examples include the Wolff update [135] for the classical Ising model as well as $\omega_n = 0$ shifts of auxiliary fields in QMC simulations of the Hubbard [101] and Holstein models [47], or "swap" updates where lattice configurations on neighboring sites are interchanged [104]. Importantly, these global moves can fail to reduce autocorrelation time if they only achieve low acceptance rates. It is usually not obvious how global MC moves should be proposed for a given model.

Finally, QMC simulations must also contend with the aforementioned Fermion sign problem.

These factors have prevented the widespread deployment of QMC algorithms for many effective models relevant to current materials of interest. However, it is hoped that ML methods can provide new routes forward. For example, the problem of constructing generalized global updates can be addressed using so-called self-learning Monte Carlo methods [76], as discussed in Section 7.

4. Machine learning methods

4.1. Artificial neural networks

Artificial neural networks (ANNs) are data structures capable of encoding highly-nonlinear functions of their input features. Originally motivated by models for the brain, ANNs usually consist of several interconnected layers of perceptrons. Like a neuron, a perceptron is an element of decision-making, providing an output (y) based on the weighted average of a set of input values ($\mathbf{x}$)

$$y = f(\mathbf{x} \cdot \mathbf{W} + b), \qquad (18)$$

where $\mathbf{W}$ are weights associated with the input values, b is a bias variable, and f is called the activation function, usually a nonlinear function such as tanh, sigmoid, or the rectified linear unit (ReLU). The idea is that more complex decisions can be made by having a large number of these perceptrons in deeper (in terms of the number of layers) and more complicated ANNs.

ANNs are usually designed for specific tasks or to make particular decisions, e.g., categorizing a large number of inputs. Training ANNs to make correct decisions takes place through observation. In supervised learning, a supervisor provides many inputs and their "labels" (correct categories) to the ANN. During this process, through backpropagation, the network gradually adjusts its many fitting parameters (weights $\mathbf{W}$'s and biases b's) to match its output with the desired output (labels). For more details, see Refs. [136,137].

4.2. Convolutional neural networks

Convolutional neural networks (CNNs) are a group of ANNs that use one or more convolutional layers in their architecture. In a convolutional layer, a kernel (also known as a filter), usually with the same dimensionality as the input data, sweeps across each input and convolves with portions of it. After going through an activation function, the results of these convulsions are combined in a "feature map", which is passed to the next layer of the neural network. Pixel values of the kernel are considered weights that can be adjusted during the training process. But importantly, the same weights are used for every convolution for a given kernel. A CNN

can have more than one convolutional layer, with subsequent layers acting on feature maps of previous layers, or more than one kernel in each convolutional layer, where each kernel has a unique set of weights and creates its feature map.

The idea behind convolutional layers is that one can work with the input data in their original shape and take advantage of spatial correlations or patterns that may exist in them. Kernels in the first convolutional layers usually pick up the most basic features in the data that can be used for categorization, and subsequent layers use those features to create more complicated patterns. Having convolutional layers generally improves the training accuracy of neural networks, especially if the input data contain important spatial features (e.g., translational invariance in physical systems). For this reason, CNNs are widely used in image recognition. However, in Section 6, we will see examples of how the information encoded in convolutional kernels can be used to infer nonlocal correlations from snapshots of fermions on a lattice.

4.3. Principal component analysis

Principal component analysis (PCA) is perhaps conceptually the simplest unsupervised ML approach. Within the context of classical statistical mechanics, it begins with a set of M 'snapshots' of the lattice, e.g. a collection of N-component vectors $\mathbf{S}_\gamma$, with $\gamma = 1, 2, 3, \ldots, M$, each representing a given configuration generated during a Monte Carlo simulation. (For example, the components $\mathbf{S}_\gamma$ could represent the $N = L \times L$ spin orientations at a given instance of a simulation of the 2D Ising model.) These M vectors are assigned to rows of a matrix $\mathcal{X}$, with dimension $M \times N$. The M vectors are typically chosen from simulations at n_T different temperatures $\{T_i\}$ that transit T_c. At each T_i, R configurations are chosen, so that $M = n_T R$. To implement the PCA, the eigenvalues λ_α and eigenvectors $\mathbf{w}_\alpha$ of the $N \times N$ covariance matrix $\mathcal{X}^T \mathcal{X}$ are determined

$$\mathcal{X}^T \mathcal{X} \mathbf{w}_\alpha = \lambda_\alpha \mathbf{w}_\alpha. \tag{19}$$

The overlaps of each configuration $\mathbf{S}_\gamma$ with a given eigenvector $\mathbf{w}_\alpha$ are then computed to define weights, or principal components, $p_{\gamma\alpha} = \mathbf{S}_\gamma \cdot \mathbf{w}_\alpha$.

As we shall see below, the topology of scatter plots of $p_{\gamma\alpha}$ for the first few largest eigenvalues λ_α changes decisively through T_c. The method works best when these eigenvalues are much larger than the remaining ones, a condition that holds in many interesting cases. In such a scenario, the original N-dimensional data contained in $\mathcal{X}$, the M vectors $\mathbf{S}_\gamma$ of length N, have been projected to a (much) smaller dimensional space of $p_{\gamma\alpha}$ with, for example, $\gamma = 1, 2$. PCA is an unsupervised ML method. No labeling of $\mathbf{S}_\gamma$ as being below or above T_c is required, and one only utilizes the raw spin configurations. The critical point emerges spontaneously as a change in the nature of the principal components in passing through the phase transition.

4.4. Autoencoders

Like PCA, the autoencoder (AE) method is an unsupervised method for performing dimensional reduction. However, it is generally more powerful because it employs nonlinear transformations. In its simplest form, an AE consists of two ANNs, the "encoder" and the "decoder", connected through a hidden layer in the middle, as shown in Fig. 3. An AE analysis begins, like the PCA, with a collection $\mathbf{S}_\gamma$ of snapshots of the system. These snapshots are used as the 'inputs' to the encoder from which the values $\mathbf{U}$ of a subsequent ('hidden') layer are computed. Denoting the components of a given $\mathbf{S}_\gamma$ as S_i (suppressing the label γ), and the associated hidden layer values by U_j, one computes for each γ
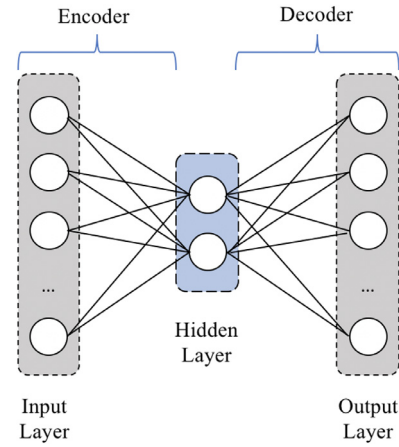
Fig. 3. The topology of an autoencoder. High dimensional input data are compressed through a hidden layer with (far) fewer degrees of freedom, followed by a demand that the high dimensional output reproduces the input when fed back into the decoder layers.

$$U_j = f\left(\sum_i W_{ji}^{(1)} S_i + b_j^{(1)}\right). \tag{20}$$

as in Eq. (18). Similarly, the values V_k on an 'output' layer out of the decoder are obtained from those on the hidden layer using

$$V_k = f\left(\sum_j W_{kj}^{(2)} U_j + b_k^{(2)}\right). \tag{21}$$

The weights $W^{(1)}$ between the input and hidden layers, and $W^{(2)}$ between hidden and output layers, as well as the biases $b^{(1)}$ and $b^{(2)}$ are adjusted, e.g. through backpropagation so that the output V matches the input S. (Hence the name 'autoencoder.')

In the AE approach, separate networks are trained for distinct parameter choices (e.g. temperature). The values characterizing the small number of hidden layer nodes, e.g. their activations, can be analyzed as a function of the control parameter. Typically these values change distinctively through a phase transition.

4.5. T-distributed stochastic neighbor embedding

Like PCA, t-distributed stochastic neighbor embedding (tSNE) is an unsupervised method used to reduce the dimensionality of data and represent them using a few projected values. But unlike PCA, tSNE does this nonlinearly by minimizing the difference between pairwise conditional probability distributions representing the similarity of points in the high- and low-dimensional spaces. The distributions are based on Student's t-distribution functions centered at each point and have widths adjusted to keep the number of effective neighbors of each point fixed throughout the configuration space. The user sets the latter as the "perplexity" number. A typical tSNE analysis starts with an initial PCA to reduce the dimensionality of data from the original value to around 50 before applying the tSNE algorithm, as it can be costly to work directly with an input dimension of hundreds or thousands. More details about the tSNE method can be found in Refs. [138,139].

4.6. Random trees embedding

Random trees embedding is another unsupervised learning method that uses the notion of a tree to extract features from data. A tree refers to a graph with nodes repeatedly branching out in

one direction. The parent node (the root of the tree) has all the data, while the data is divided into subsets at subsequent nodes representing tree branches. Each node corresponds to a feature in the data, so branches farther from the parent node containing smaller subsets of data correspond to finer features.

In random trees embedding, data are projected onto an ensemble of random trees whose total number and maximum branching depth are parameters that can be tuned. Each tree makes an independent observation regarding the features, and the ensemble of trees "votes" for dominant features, judged by the amount of overlap between nodes at a certain depth from different trees. For more details see Refs. [140,141].

5. Proofs of concept and benchmarks

5.1. Classical models

In May and June of 2016, a series of groundbreaking papers [142–145] came out that demonstrated the power and potential of machine learning techniques in encoding information about the statistics of classical and quantum many-body systems and how they may be used for physics discovery. These works showed for the first time that one could think of the degrees of freedom in many-body systems – e.g., individual spin orientations in the Ising model or auxiliary field configurations in QMC simulations – as "features" in ML algorithms. This realization paved the way for utilizing ML methods developed and refined for industry applications to learn new physics.

Carrasquilla and Melko [142] employed fully-connected and convolutional neural networks to study phase transitions in models for magnetic systems. By coloring their 2D spin configurations obtained from a Monte Carlo simulation as hot or cold (referring to whether they were obtained at a temperature above the critical temperature of the model or not), they were able to train the networks to classify never-before-seen configurations and pinpoint T_c with a high degree of accuracy. They further showed that predicted values of T_c approached the analytical value in the thermodynamic limit as their system sizes increased and explored applications of training with models exhibiting topological orders. They also demonstrated that a fully-connected neural network simply learns to compute the magnetization of the Ising model and uses it as a metric for classification. This observation helped explain its success in transferring the knowledge learned on a square lattice geometry to a triangular lattice geometry.

Lei Wang's application of the PCA technique to the Ising model [143] led to a similar conclusion: physical properties, such as magnetization or magnetic susceptibility, emerge in the first two principal components. What was remarkable about Wang's findings was that these properties could be inferred without providing knowledge about the problem's physics to the machine. That is, the spin configurations were not labeled in the PCA study. Later these ideas were applied to study frustrated classical magnetic models, such as in Ref. [147].

To illustrate these approaches more concretely, we now highlight a specific example, the 2D ferromagnetic Ising model [146,148,149]. Fig. 4 shows results obtained for the phase transition using a PCA [146]. Fig. 4(a) plots the eigenvalues of $\mathcal{X}^T \mathcal{X}$, which fall off rapidly, a condition for the data compression inherent in a PCA to be effective. A scatter plot of pairs $(p_{1\gamma}, p_{2\gamma})$ of projections of the spin configurations $\mathbf{S}_\gamma$ on the PCA eigenvectors $\mathbf{w}_1, \mathbf{w}_2$ with the largest and next-largest eigenvalues [Fig. 4(b)] shows an evolution from a single clump centered at the origin for $T > T_c$ to a bimodal distribution for $T < T_c$. The average of $|p_{1\gamma}|$, shown in Fig. 4(c), behaves like an order parameter (the magnetization), evolving from zero at high T to a non-zero value at low T. The transition becomes increasingly sharp with L
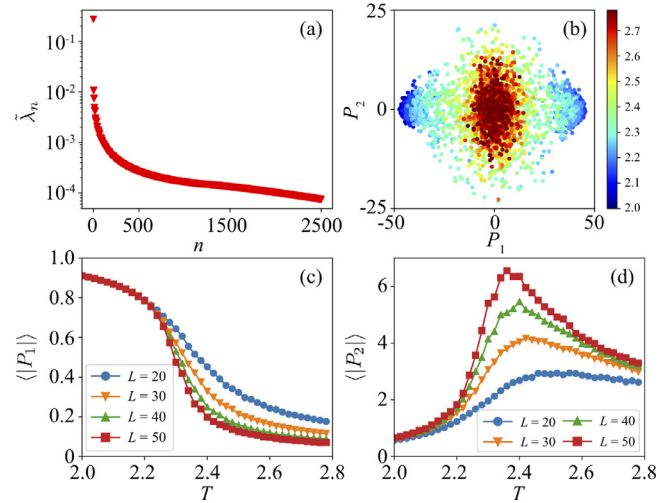
Fig. 4. PCA results for the square lattice Ising model, which has an analytically known $T_c/J = 2.269$. (a) Distribution of the lowest 2500 eigenvalues of the matrix $\mathcal{X}^T \mathcal{X}$. (b) Scatter plot of pairs $(p_{1\gamma}, p_{2\gamma})$ of overlaps of Ising configurations $\mathbf{S}_\gamma$ with the first two eigenvectors $\mathbf{w}_1, \mathbf{w}_2$ of $\mathcal{X}^T \mathcal{X}$. Results are shown for 100 configurations at each temperature. (c) Average overlap $p_{1\gamma}$ and (d) $p_{2\gamma}$ as a function of temperature for four linear lattice sizes L. Adapted from Ref. [146] with minor modifications.
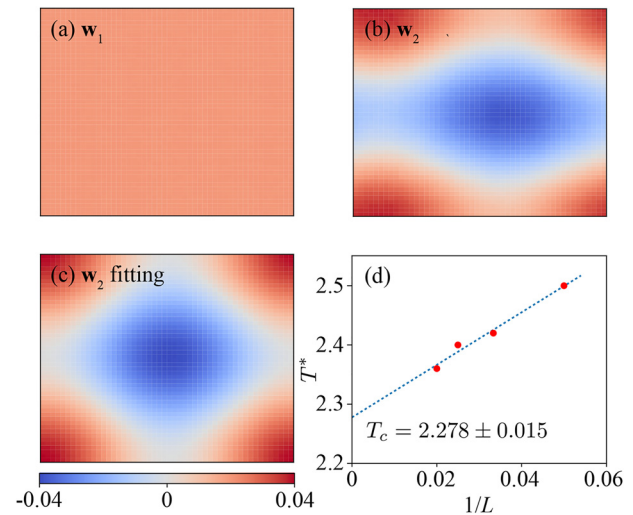

Fig. 5. (a) The eigenvector $\mathbf{w}_1$ of largest eigenvalue in a PCA is nearly uniform ($\mathbf{k} = (0, 0)$), so that its dot products $p_{1\gamma}$ with spin configurations $\mathbf{S}_\gamma$ have the physical interpretation of the magnetization. (b) The eigenvector $\mathbf{w}_2$ of next-largest eigenvalue in a PCA exhibits domain walls between spin-up and spin-down regions. It can be well fit by a vector consisting of a sum of two plane waves with $k_1 = (0, 2\pi/L)$ and $k_2 = (2\pi/L, 0)$, i.e. the k values closest to the origin, as seen in panel (c). (d) shows the evolution of the peak positions of the average of $p_{2\gamma}$ (Fig. 4(d)) with inverse linear lattice size. Adapted from Ref. [146] with minor modifications.

(the total number of sites $N = L \times L$) and occurs near the known $T_c = 2.269$ [75]. The average of $|p_{2\gamma}|$ [see Fig. 4(d)] behaves like the susceptibility, peaking near T_c (compare with Fig. 1).

Fig. 5 (a) shows the first eigenvector $\mathbf{w}_1$. It is nearly uniform, so that $p_{1\gamma} = \mathbf{w}_1 \cdot S_\gamma$ is essentially the magnetization, as already implied by Fig. 4(c). One of the most promising possibilities of ML approaches for statistical mechanics is the possibility that examining the learning mechanism might lend insight into the nature of phase transitions, especially in cases where the order parameter is unknown. Fig. 5(b) shows, similarly, the second eigenvector, which can be compared with a linear combination of vectors with domain walls in the vertical and horizontal directions, i.e. $\mathbf{k}_1 = (0, 2\pi/L)$ and $\mathbf{k}_2 = (2\pi/L, 0)$, as shown in Fig. 5(c). To-
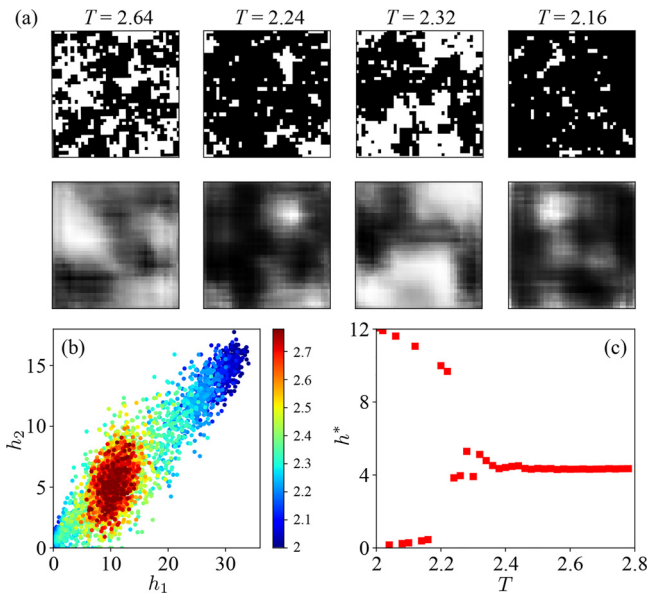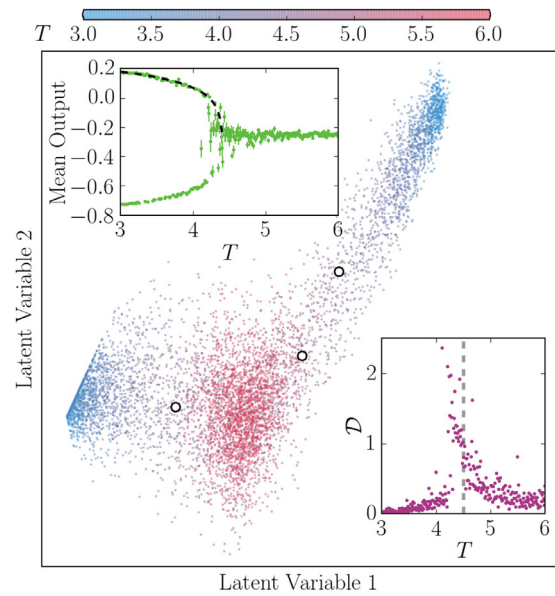
Fig. 6. Autoencoder results for the 2D square lattice Ising model. (a) After compression of the spin configuration of an $N = 40 \times 40$ lattice through 200 hidden layer neurons, the basic features of the input can still be replicated in the output. (b) Scatter plots of the activations (h_1, h_2) of a hidden layer with two neurons. The distribution bifurcates at T_c. (c) The activation h^* of a single hidden layer neuron follows a single trajectory for $T > T_c$ but then separates into two branches for $T < T_c$. Adapted from Ref. [146] with minor modifications.


Fig. 7. Projection of 3D Ising data for an $N = 8 \times 8 \times 8$ lattice on the space of two latent variables (hidden neurons) of an AE. CNN's are used for the encoder/decoder of the AE. Different colors in the scatter plot correspond to different temperatures. Top inset: The output of a different AE, in which the hidden layer consists of one single neuron, as a function of temperature. The dashed line is a fit to $A(B - T)^\beta + C$ with $A = 0.38$, $B = 4.55$, $\beta = 0.34$, and C kept fixed at -0.25, which is the average output overall T. Bottom inset: Temperature dependence of a measure for the spread of data in the main panel. The vertical dashed lines mark the location of T_c. Taken from Ref. [146].

gether with the $\mathbf{k} = (0, 0)$ structure of the first eigenvector shown in Fig. 5(a), one concludes PCA is constructing the low energy (small $\mathbf{k}$) Fourier components of the spin configurations. Finally, Fig. 5(d) shows an extrapolation of the peaks in Fig. 4(d) in inverse linear lattice size. The temperature intercept in the thermodynamic limit is within error bars (of about 1%) of the exact $T_c = 2.269$. This result illustrates that ML can be combined with finite size scaling to reach the thermodynamic limit (as with older methods), but also that quite accurate results can be obtained without too much effort (i.e., using the simple PCA).

As mentioned, the AE method can be viewed as a nonlinear generalization of PCA. Therefore, it should be no surprise that the AE method is also an effective means to study the Ising phase transition, as shown in Fig. 6. Panel (a) illustrates the data compression of the AE by showing the actual input features (top row), in this case, $N = 40 \times 40$ spin values for configurations at four different temperatures, together with their replication at the output stage (bottom row). Here, the AE uses two hundred hidden neurons, almost an order of magnitude reduction over the number of sites in the model. Figs. 6(b) and (c) demonstrate that the AE retains the ability to detect the phase transition even when the data is highly compressed. For example, in Fig. 6(b), the hidden layer has only *two* neurons, yet a scatter plot in the plane of their activation (h_1, h_2) bifurcates in a manner similar to a PCA of Fig. 4. The temperature at which this bifurcation occurs yields an estimate of T_c; however, even the activation h^* of a single hidden layer neuron, panel Fig. 6(c), shows a clear signal of the phase transition as T is reduced.

Similar results have also been obtained for the 3D version of the antiferromagnetic Ising model [150] (see Fig. 7). The top inset in Fig. 7 shows that the single latent variable of a convolutional autoencoder can act as the order parameter; however, other indicators can also be defined, based on the distribution of more than one latent variable, that point to the transition temperature and even correlate with physical properties. For example, in the bottom inset of Fig. 7, the spread of the data in the space of two

latent variables peaks around the critical temperature and closely follows the susceptibility curve. The mean output can be fit to give a critical temperature of $T_c = 4.55$ and critical exponent of $\beta = 0.34$. These values should be compared with values $T_c = 4.5115$ and $\beta = 0.326$ obtained by Monte Carlo simulation [151].

Agrawal et al. have also used autoencoders to examine the related problem of detecting and identifying which symmetry is broken spontaneously across a phase transition [82]. To this end, they introduced an architecture called the group-equivariant autoencoder (GE-autoencoder). In this application, one first deduces a set of symmetries that will remain intact in all phases at all temperatures using group theory. This information is then used to constrain the hyperparameters of the GE-autoencoder such that the encoder learns an order parameter invariant to these symmetries. Benchmarking their method for the ferromagnetic and antiferromagnetic Ising model in 2D, they could construct GE-autoencoders whose size remained independent of the system size. By including additional symmetry regularization terms in the loss function, they found that the GE-autoencoder learns an order parameter that satisfied the remaining symmetries of the system. The authors extracted information about the associated spontaneous symmetry breaking by examining the group representation by which the learned order parameter transforms. The GE-autoencoder was also able to produce estimates for T_c in the thermodynamic limit with greater accuracy, robustness, and time efficiency than a symmetry-agnostic autoencoder discussed above.

The studies of the Ising model (Figs. 4, 5, and 7) focused on using configurations at different *temperatures* to determine T_c. In the Blume–Capel model, however, the phase boundary can be crossed by varying the zero-field splitting Δ. PCA is also effective in such 'parameter-tuned' transitions, as shown in Figs. 8 and 9. The Blume–Capel model has a tricritical point at $(T/J, \Delta/J) = (0.609, 1.965)$ along its phase boundary. Fig. 8 cuts across at a second-order transition $(T/J = 1.0)$, while Fig. 9 cuts across at a
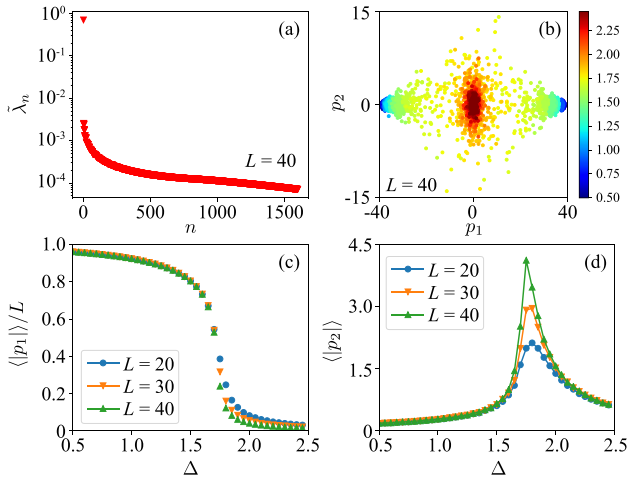
Fig. 8. Same as Fig. 4 except for the Blume–Capel model and sweeping Δ at fixed $J = T = 1$. As in the Ising model, there is a crisp separation of eigenvalue scales (panel a) and a bifurcation of the scatter plot at Δ_c (panel b). The average overlaps once again serve as proxies for the magnetization ($\langle|p_1|\rangle$, panel c) and susceptibility ($\langle|p_2|\rangle$, panel d). The transition is second order for this $T = 1$ sweep. Adapted from Ref. [146] with minor modifications.
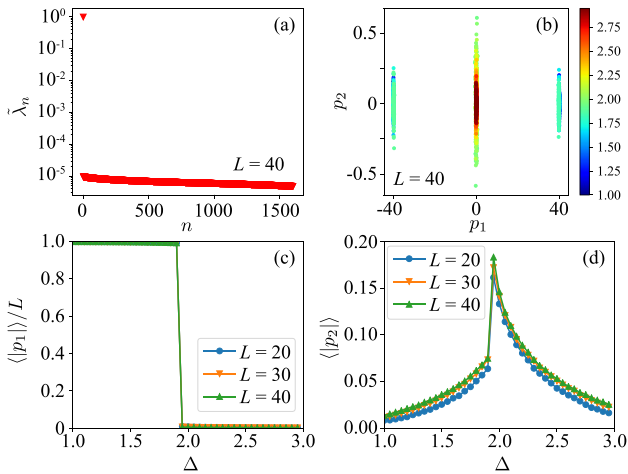
Fig. 9. Similar to Fig. 8, except the transition is first order for this $T = 0.4$ sweep. Adapted from Ref. [146] with minor modifications.

first-order transition ($T/J = 0.4$). Not only are the critical points easily identified, but their orders are readily apparent.

For pedagogical reasons, we have focused here on the use of ML for the well-known and characterized Ising and Blume–Capel models. It is worth noting, however, that these methods have also been used to elucidate the behavior of other, more challenging, classical models. These include the biquadratic spin exchange spin-1 Ising model of He^3-He^4 mixtures [146,152], frustrated magnetism [147], the Kosterlitz–Thouless transition of the 2D XY Hamiltonian [146,153,154], and topological order in Ising gauge theories [155]. Together, the studies discussed here illustrate the rapid pace at which powerful variants of ML methods are being developed, in analogy with the (much longer) history of the evolution of Monte Carlo methods for studying phase transitions.

5.2. Quantum models

The use of restricted Boltzmann machines (RBM) as an ansatz for representing ground state wavefunctions of quantum many-body systems have been another exciting and fruitful approach. In
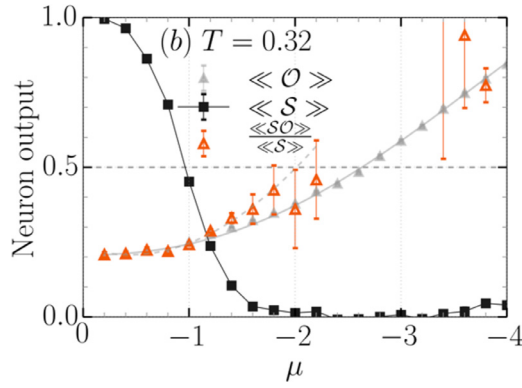
a novel study, Carleo and Troyer [144] used reinforcement learning to train their RBMs by minimizing the ground state energy of the transverse-field Ising and quantum Heisenberg models. In doing so, they showed that they could surpass the performance of then state-of-the-art conventional variational techniques by systematically increasing the density of the hidden layer.

These works laid the groundwork for and inspired many other studies that followed shortly after [156]. For example, ideas of Ref. [142] were extended and applied to quantum many-body systems like the Fermi Hubbard model on the honeycomb and cubic lattices [157,158] to learn quantum or thermal phase transitions. Broecker et al. [157] showed that the knowledge of the physics of the problem at the extremes, in that case, deep in the semi-metal or AFM phases of the honeycomb lattice Hubbard model, can lead to an accurate estimate of the transition temperature by the artificial neural networks. They also found that input data engineering to guide the neural networks toward physical properties of interest can significantly affect the training.

Ch'ng et al. [158] used raw auxiliary field configurations of the cubic lattice Hubbard model in 3 + 1 dimensions and by treating the time slices along the quantum dimension as "color channels" in their CNNs, allowing the method to learn the finite-temperature Néel transition at half-filling. They then demonstrated the power of transfer learning by training a CNN using a mix of configurations from two different U values in the weak- and strong-coupling regimes, which was then used to estimate the nontrivial shape of the AFM phase boundary in the temperature-interaction strength phase diagram of the model (see Fig. 10).

Both of these works touched on the fermion sign problem in Fermi–Hubbard models away from symmetry points and their implications for machine learning. Ref. [157] demonstrated that, at least for some models and properties of interest, the sign could essentially be ignored in the training and classification. In Ref. [158], Ch'ng et al. avoided training their machines in the sign-problematic parameter region away from half-filling. Instead, in using those CNNs trained at half-filling to track the magnetic transition away from half-filling, they treated their network output as another physical observable, arguing that the sign of the auxiliary field configurations should be incorporated into their averages. As shown in Fig. 11, ignoring the sign can lead to small but significant differences in the results in this case.

Early on, it was shown that topological states of matter could also be studied using artificial neural networks. RMBs were first


Fig. 10. Prediction of the AF transition temperature in the 3D Hubbard model by a CNN. The CNN is trained using DQMC's auxiliary field data at different temperatures for fixed values of U. The prediction is made after three pieces of training: One with data from $U = 5$, one from $U = 16$, and one with mixed data from $U = 5$ and 16. The latter yields the nontrivial shape of the phase boundary when providing the CNN with data from other U values it has not seen before. Estimates for T_N in the thermodynamic limit are from past DQMC and numerical linked-cluster expansion calculations. Taken from Ref. [158] .

Fig. 11. Average neuron output $\mathcal{O}$ calculated with and without taking the fermion sign into account at $T = 0.32$ ($\langle\langle\mathcal{SO}\rangle\rangle/\langle\langle\mathcal{S}\rangle\rangle$ and $\langle\langle\mathcal{O}\rangle\rangle$, respectively) when using a CNN trained at half-filling to detect the Néel transition away from half-filling. 0.5 crossing indicates the location of the transition. The black line shows the average fermion sign. Taken from Ref. [158].

used to represent topological states in one, two, and three dimensions [159], and it was found that the number of hidden parameters needed scales only linearly with the system size. It was further shown that RBMs could find the topological ground states of generic nonintegrable Hamiltonians through reinforcement learning and identify their topological phase transitions. RBMs were also used as a decoder of topological codes [160]. The challenge of capturing nonlocal properties of topological phases with neural networks led Zhang and Kim to introduce quantum loop topography, a procedure to construct a multidimensional image of the wavefunction based on two-point correlation functions that form loops, which was then used as input to neural networks to distinguish Chern insulators from trivial insulators [161].

Algorithms for the unsupervised learning of phases and phase transitions (with no specific knowledge of the nature of phases or the whereabouts of the transition) that were based on supervised machine learning methods attracted much interest. In the "confusion" method [162], neural networks are trained with data that have been deliberately mislabeled. Monitoring how the training accuracy varies as different locations for the phase transition are proposed allows one to identify the correct labels for the data and pinpoint the transition. Another method also used the training accuracy as a function of the tuning parameter, but for training performed on data from consecutive tuning parameters [163]. In this approach, a peak in the accuracy would indicate a sudden change in the character at the location of the true phase transition.

Traditional unsupervised learning methods, such as PCA, autoencoders, tSNE, and random trees embedding, showed a remarkable ability to reveal phase changes in the presence of quantum fluctuations. While an early application of the PCA to QMC data for the Heisenberg model led to no discernible features in the reduced dimensional space [162], a thorough analysis of the raw auxiliary field DQMC data for the 2D and 3D Hubbard model demonstrated outstanding potential for nonlinear methods to shed light on the phases and phase transitions of quantum lattice models. It also showed that indicators that correlate with conventional properties could be defined using the data projected onto the reduced dimensional space [150].

Figs. 12 and 13 highlight some of those early findings. Fig. 12(a) shows the results of tNSE applied to raw auxiliary field data for the half-filled 2D Hubbard model with $U = 4t$ on a 10×10 square lattice. The color indicates temperature and makes it clear that the data projected to the 2D space evolve from a large symmetric cluster at relatively high temperatures to two smaller ones on either side of the hot cluster at lower temperatures. They correspond to
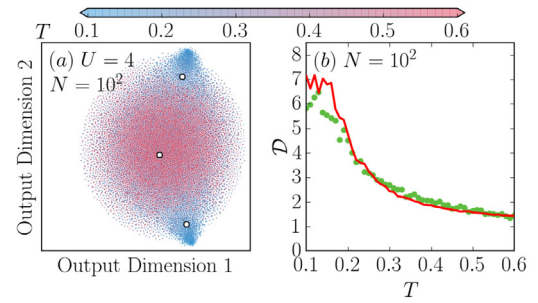
Fig. 12. Projection of the auxiliary field data for the half-filled 2D Hubbard model to a two-dimensional latent space using the tSNE algorithm. The system size is $N = 10^2$ and $U = 4t$. Eight hundred configurations per temperature in a uniform grid of T/t between 0.1 and 0.60 are used. (b) The indicator (green symbols – see text) and the AFM structure factor (red line) as a function of temperature. Image reproduced from Ref. [150]. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
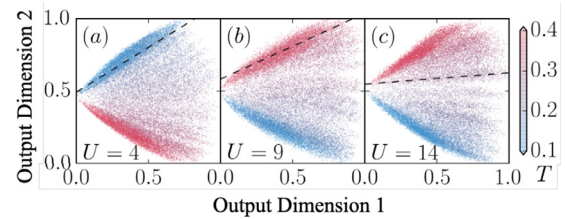

Fig. 13. Projection of the four latent variables of a fully-connected AE to two dimensions using the random trees embedding algorithm. The original data used to train the AE are auxiliary fields for the 3D Hubbard model with (a) $U = 4$, (b) $U = 9$, and (c) $U = 14$ at half-filling. The dashed lines are line fits to data at the estimated Néel temperature for each U. Image reproduced from Ref. [150] .

the two possible sublattice orientations of the Néel ordered phase. When applied to the data, the k-means algorithm identifies three clusters at each temperature. Fig. 12(b) shows how the average distance of a cluster's center from the mean location of data closely follows the magnetic structure factor as a function of temperature. Note that no knowledge about the physics of the problem has been provided to the machine.

Fig. 13 shows temperature gradients in the 3D Hubbard data from an AE that have been further analyzed by the random trees embedding method. Here, the AE reduces the original auxiliary field data to four latent variables, and then the random trees embedding projects the latent variables to a 2D space. In this case, the output separates data points from different temperatures.

ML methods have been applied to many other quantum models. For example, PCA has also been widely used to locate phase transitions in various quantum Hamiltonians, with some notable successes and failures, which we now discuss. We begin with the finite temperature CDW phase transition in the Holstein model and then examine magnetic quantum phase transitions tuned by changing model parameters in several different contexts, including: (1) the inter-orbital hybridization in the periodic Anderson model, (2) the on-site interaction in the Hubbard model on a honeycomb lattice, and (3) the density in the Hubbard model on a Lieb lattice. 'Topological data analysis' is a related ML method recently applied to (1) and (2) [164]. In all these cases, the input features are provided to the PCA are the Hubbard–Stratonovich field variables $\{s_{i,l}\}$ obtained from DQMC simulations of these models. With these success stories established, we then discuss the challenges encountered in studying the Kosterlitz–Thouless transition to a superconducting phase in the doped two-dimensional attractive Hubbard model [165]. We note that for quantum models, DQMC simulations work in a path integral representation of the partition function so that they sample a $d + 1$ dimensional space-imaginary time lattice,
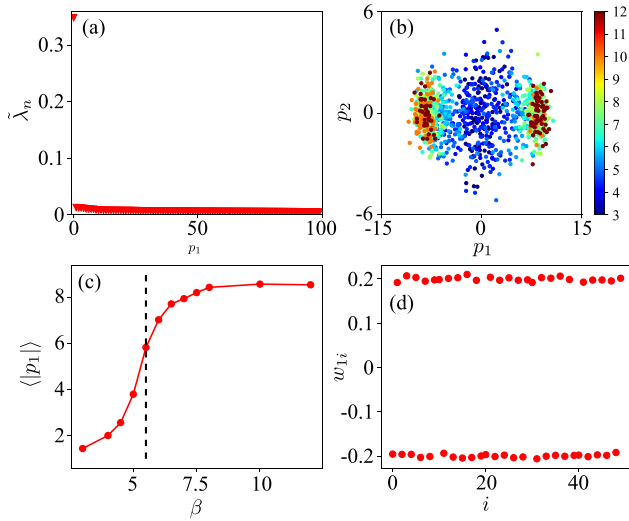
Fig. 14. PCA for the Holstein Hamiltonian on a 10×10 lattice at half-filling, with $\omega_0 = 1$ and $\lambda = 1/2$. The principal eigenvalues of panel (a) exhibit a rapid fall-off, suggesting data compression should be effective. The vertical dashed line gives the critical temperature for the CDQ phase transition obtained from a traditional finite-size scaling analysis of the order parameter. The ML calculation shows a sharp rise in the first principal component (panel c) and the development of a two-peak structure in a (p_1, p_2) scatter plot (panel b) at the same T_c. The overlaps of the spin configurations with the first principal eigenvector having the staggered structure of CDW order (panel d). Adapted from Ref. [166] with minor modifications.

where d is the spatial dimension. Unless otherwise indicated, the configurational vectors used in the PCA discussion that follow contain the entire lattice. In principal, one might also study the performance of the PCA for different discretizations $\beta = L\Delta\tau$ of imaginary time; however, the results do not appear to be very sensitive to the size of the Suzuki–Trotter errors, provided $\Delta\tau$ is reasonably small.

The Holstein Hamiltonian poses special difficulties to QMC simulations owing to its long autocorrelation times. ML methods play an especially useful role in the acceleration of the simulations, as discussed in Section 7. In this section, we will confine ourselves to PCA's use in analyzing configurations generated by the conventional DQMC method.

Fig. 14 shows PCA analysis of the CDW transition in the Holstein model in a close analogy to that presented in Fig. 4. As with the previous examples, the sharp drop-off in the eigenvalues shown in Fig. 14(a) (even more dramatic than that in the Ising case) suggests that the PCA method will be able to compress the configuration space efficiently. The projection onto the plane of the two principal components is shown in panel (b). Here, the bifurcation of the distribution is observed as T is lowered, reflecting the transition to the CDW phase. In this case, the average of the first principal components as a function of (inverse) temperature, shown in panel (c), serves as an order parameter. At the same time, the spatial structure of the components of the first eigenvector w_{1i} at low T oscillates in sign, reflecting the two sublattice structure of the CDW phase whose ordering wave-vector is $\mathbf{Q}_{CDW} = (\pi, \pi)$.

Fig. 15 examines PCA's ability to discern the quantum phase transition of the PAM (Section 2.5) [166]. In this case, DQMC simulations at low temperature ($T = t/24$) are used to generate data for varying hybridization values V. Applying the PCA to this data set at fixed T produces results very similar to the Holstein model; the principal eigenvalues fall off rapidly [Fig. 15(a)], demonstrating that PCA can indeed achieve a large degree of data compression. Similarly, the scatter plot of (p_1, p_2) bifurcates as a function of the model parameter V. At the same time, the first principal component p_1 [Fig. 15(c)] behaves like an order parameter for the anti-
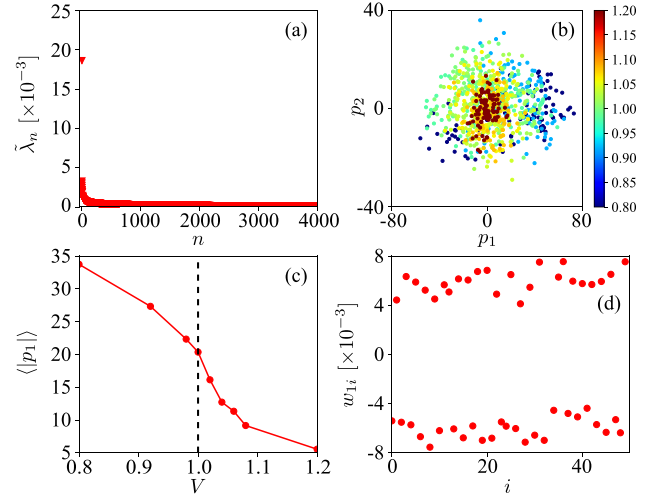
Fig. 15. PCA for the periodic Anderson Model (PAM) on a 12×12 lattice with $U_f = 4t$ and $\beta t = 24$. The inter-orbital hybridization V is the tuning parameter. The PCA can isolate the QPT, which occurs between AF order at small V and the singlet phase at large V. Adapted from Ref. [166] with minor modifications.

ferromagnetic phase in that it goes to zero as V increases across the known $V_c \sim 1$ [35]. At low V, the first eigenvector exhibits a clear oscillatory pattern, reflecting strong AFM correlations.

The Hubbard model on a square lattice has AFM order at half-filling for all values of the on-site repulsion U [1,167]. For weak coupling, this ordering is a consequence of the 'perfect-nesting' of the Fermi surface (FS), where a large number of points $\mathbf{k}$ and $\mathbf{k} + (\pi, \pi)$ lie on the FS leading to an enhanced instability to AFM order. The logarithmic van-Hove singularity of the density of states $N(\omega = 0)$ contributes to this process. No such nesting occurs on a honeycomb lattice, and $N(\omega = 0)$ vanishes linearly with ω at half-filling (the so-called 'Dirac spectrum') [168]. This electronic structure leads to a finite U_c for the AF order. The physics of the semi-metal to AF transition has engendered a great deal of investigation with numerical methods [46,168–171], including early studies of a possible intervening spin-liquid phase [172] that does not appear to occur [173]. Ref. [166] revisited this issue and studied the semi-metal/AF transition of Dirac fermions on a honeycomb lattice using PCA, with clear indications of a transition to an AFM state at a U_c in agreement with the most accurate value $U_c = 3.87t$ found in the literature [173].

The results above demonstrate that a PCA can resolve temperature- and model-parameter-driven (quantum) phase transitions. However, many quantum materials can be doped, and the fermionic carrier density often functions as another tuning parameter. With this in mind, we now examine simulations of the Hubbard model on a Lieb lattice, the geometry of the CuO_2 planes in cuprate superconductors, as a function of carrier concentration. The Lieb lattice has a square array of d orbital 'Cu' sites that are bridged by intervening p orbital 'O' sites, as shown in the inset of Fig. 16(a). Each unit cell thus has three orbitals and the filling corresponding to the AF parent compounds of the cuprates is $\rho = 1/3$ hole per unit cell. The model is most commonly studied with $U_d > U_p$, and on-site energy for the oxygen sites is higher (for holes) than for the copper sites, as is the case for cuprates [56,174,175].

Fig. 16 shows the density-tuned transition through AF order in the Lieb lattice. A somewhat different perspective here is obtained by showing the distribution of principal components both in the SDW phase $\rho = 1/3$ (panel c), as well as below $\rho < 1/3$ (panel b) and panel $\rho > 1/3$ (panel d). The tightness of the cluster at $\rho = 1/3$, in contrast to the bracketing densities, indicates the
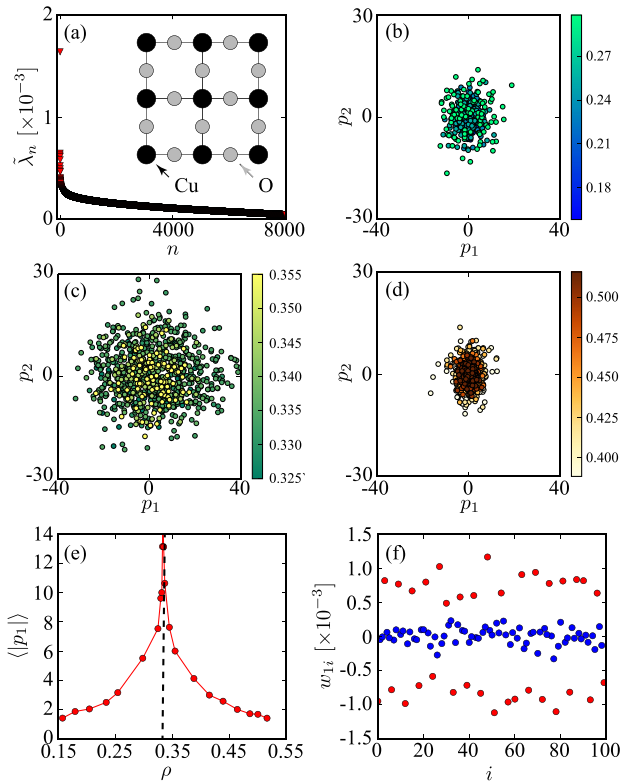
Fig. 16. PCA for the antiferromagnetic transition on a Lieb lattice. The geometry consisted of 10×10 unit cells, each with three orbitals, as shown in the inset of panel (a). The inverse temperature $\beta t = 20$. The on-site $U_d = 4t$ for the 'copper sites' and $U_p = 0$ for the 'oxygen sites'. The charge-transfer energy $\epsilon_{pd} = 2t$. Here the first eigenvector, shown in panel f, contains precise information about the nature of the ordered phase. The copper atoms (red circles) hold the AF order, whereas the oxygen sites (blue circles) do not participate. The convention for the density is such that $\rho = 1/3$ corresponds to one hole per (three sites) unit cell. Adapted from Ref. [166] with minor modifications. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

nature of the transition is disorder-order-disorder. This is a distinction from the previous cases where the critical point separates disordered from ordered regions. Another difference is the relative closeness of the first principal component to the succeeding ones - it is only larger by a factor of two.

Application of PCA to Hubbard models has a sign problem in cases where the problem is not particle-hole symmetric [57,59], as is the case for the Lieb lattice [174]. Thus, the results of Fig. 16 carry the additional implication that ML methods have some potential to address phase transitions in a quantum model, evading the sign problem. This aspect is an especially intriguing possibility since the ML analysis does not involve the measurement of the noisy ratio of two quantities that are both exponentially small. Other approaches investigating novel observables that avoid the sign problem have recently been proposed [64,176].

Our final example of using ML to examine fermionic quantum phase transitions concerns the transition into a superconducting phase in the attractive Hubbard model when doped away from half-filling. This transition has long been a challenging problem in the field because the transition is in the Kosterlitz–Thouless universality class; the nature of the ordered phase, where the correlation functions decay as power laws, is more delicate than with a "true" long-range order. Larger lattices are required to treat such phases accurately, and it is fair to say they are much less understood than the phase transitions discussed above. Indeed, quantita-
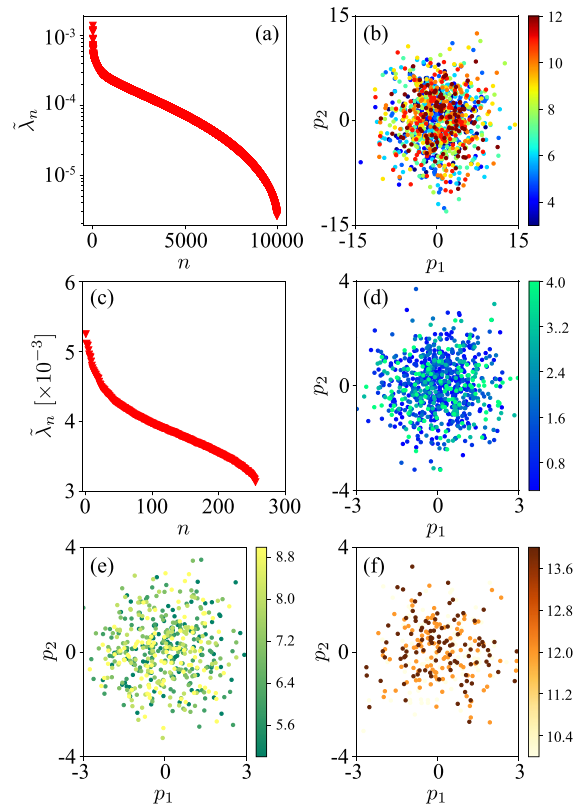
Fig. 17. PCA for the attractive Hubbard model on a square lattice with $U = -4t$ and $\rho = 0.80$. Panels a,b are for a 12×12 lattice and panels c–f for a 16×16 lattice. In the latter case, the distribution has been separated into three plots. The critical value of the inverse temperature is $\beta_c \sim 7.5$, but the analysis provides no signal there. Adapted from Ref. [166] with minor modifications.

tive values of superconducting T_c in the attractive Hubbard model have varied by 20–30% in various QMC studies [165,177–182].

In Fig. 17, we observe that the same difficulty is encountered in a PCA. For one, the eigenvalue distribution [Fig. 17(a)] does not exhibit a gap but instead decays slowly. Similarly, the principal component distribution [Fig. 17(b)] shows no clear signal as the inverse temperature β is tuned through $\beta_c \sim 7.5$, where the superconductivity is believed to onset. In an attempt to discern the transition, analyses using input feature vectors containing the entire space-time lattice (panels a & b) or just a single time slice (panels c–f) have been attempted with no success.

Figs. 18 and 19 present further efforts to discern the superconducting transition in the attractive Hubbard model. In the former, rather than providing the HS field configurations as the vectors $\mathbf{S}_\gamma$, the authors instead use the equal time Greens function $G_{ij}^\sigma = \langle c_{i\sigma} c_{j\sigma}^\dagger \rangle$. In the latter, they used the equal time pair correlation function, $\mathcal{P}_{ij} = \langle \Delta_i \Delta_j^\dagger \rangle$ with $\Delta_i = c_{i\uparrow} c_{i\uparrow}$.[3] Only the second approach seems capable of capturing the transition. For example, the average of the first principal component in Fig. 19(c) has a maximum slope near the known value of β_c.

6. AI-assisted phase discovery in strongly correlated systems

The establishment of ML methods as viable tools for categorizing quantum data led scientists to utilize them for phase detection and physics discovery. Here, we highlight two studies in

---

[3] Both quantities require measuring physical observables, and would be impacted by the fermion sign if this approach is applied to models with a sign problem.
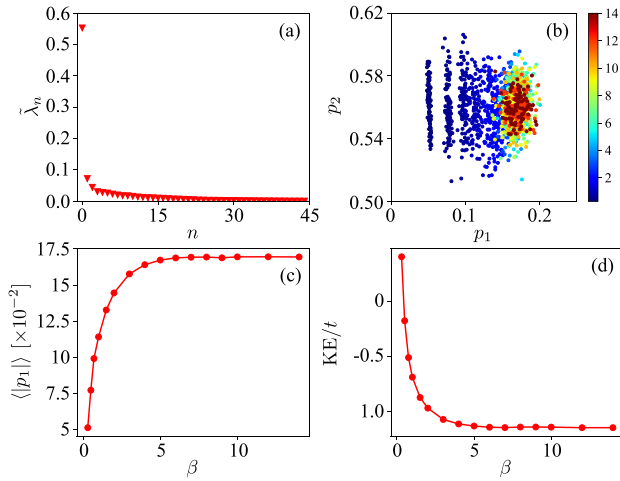
Fig. 18. PCA for the attractive Hubbard model using vectors comprised of the equal time Greens function. Although there is a gap in the eigenspectrum and an evolution of the scatter plot of the first two principal components, there is no evidence of the transition. As might be expected, the average of the first principal component tracks the kinetic energy (panel d) rather than the pairing order parameter. Adapted from Ref. [166] with minor modifications.
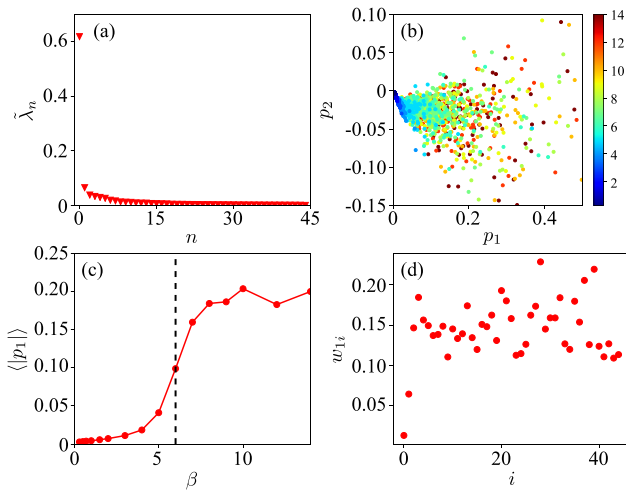
Fig. 19. PCA for the attractive Hubbard model using vectors comprised of the equal time pair correlation function. Here, finally, there is some indication of the superconducting transition. Adapted from Ref. [166] with minor modifications.

which experimental data were analyzed by AI to yield insight into the physics of strongly correlated systems. In the first study, Zhang et al. [183] used an ensemble of fully-connected neural networks to identify the dominant pattern in an archive of real scanning tunneling microscopy (STM) images of lightly-doped cuprates. These networks were trained on synthetic STM images created using theoretical models to represent four different categories of electronic ordering patterns. The authors showed that the ANNs could discover a lattice-commensurate, four-unit-cell periodic, translational-symmetry-breaking phase in the noisy experimental data. Moreover, they established the unidirectionality of the ordering pattern and how its dominance depends on the electron energy. Fig. 20 shows a sample STM image analyzed by the ensemble of ANNs (top left panel). The linear Fourier transform (top right panel) reflects the level of noise and complexity that exists in the image and does not point to any particularly dominant **q** vector. On the other hand, the output of the ANNs (lower panel) demonstrates
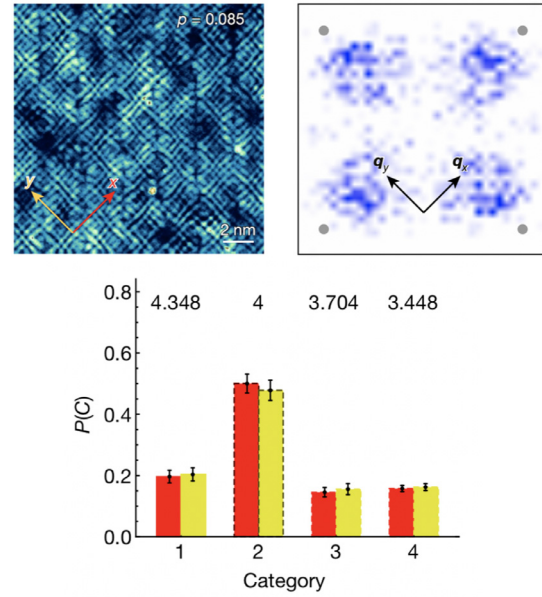
Fig. 20. Top Left: Measured 440 pixel $\times$ 440 pixel STM image of $Bi_2Sr_2CaCu_2O_8$ at 8.5% hole doping. Top Right: The d-symmetry Fourier transform of the image on the left. Bottom: Average output category of 81 ANNs. The numbers on top show the wavelength of each electronic ordering category in units of the lattice spacing. Red and yellow indicate the outputs for two different orientations of the input image, 90° rotated relative to each other. Image reproduced from Ref. [183]. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

that the second category with a wavelength four times the lattice spacing is clearly dominant.

Linear Fourier transforms have been traditionally used for decades to analyze such images. The key to the success of ANNs in this study was the existence of non-linearities, allowing them to look beyond what a Fourier transform can provide.

The second study [184] aimed for the CNNs to have an unbiased take (not guided by any theory) on the ordering patterns and possible correlations of strongly correlated fermions in optical lattices. Following an early application of CNNs to help decide which one of two theories better describes patterns in snapshots taken in the pseudogap regime of the Hubbard model using quantum gas microscopy [185], Khatami et al. [184] designed a simple CNN architecture and showed that patterns formed in filters of a CNN, trained to distinguish snapshots taken at low temperatures from those taken at high temperatures, can reflect the correlations favored by the systems as the temperature is lowered. Fig. 21 shows an example of the CNN used in their study. Having one/few filter(s) in the usually only convolutional layer directly connected to the input (physical snapshots of fermions) allows the scheme to work. The idea is that since there are no correlations at high temperatures when the system is entirely unordered, the filter will likely pick up patterns formed in the snapshots of the system at low temperatures to carry out the categorization accurately. By studying the trained filters, one can then infer relevant electronic correlations.

Two types of experimental snapshots were available at the time, mainly around the strange metal phase of the Hubbard model, and were used in this study: (1) Those of a single species of fermions and (2) those of the two species together, minus the double occupations, which would show up as empty sites for technical reasons. (2) are often called "singles" snapshots and can also be thought of as snapshots of local moments. In addition, theory snapshots were generated by periodically pausing the DQMC simulations of the 2D Hubbard model and calculating average site oc-
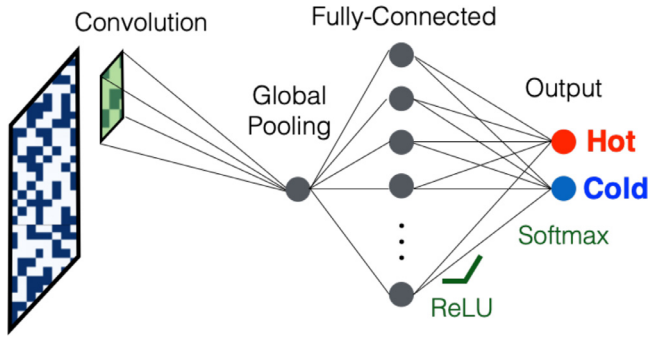
Fig. 21. The main CNN used in Ref. [184]. It has one convolutional layer with one filter, followed by other pooling and fully-connected layers. It is trained to distinguish input snapshots of fermions at two extreme temperatures. After the training, patterns developed in the filter can correspond to physical correlations in the system at low temperatures. Image reproduced from Ref. [184].
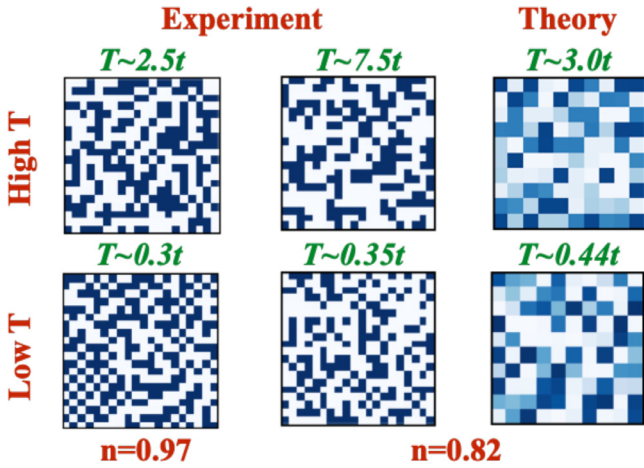

Fig. 22. Sample snapshots used in the training of the CNN in Fig. 21. Shown are four from the experiments for densities $n = 0.97$ and $n = 0.82$ at the two extreme temperatures available, and two from DQMC for $n = 0.82$. Image partly reproduced from Ref. [184].

cupations. The latter leads to non-binary snapshots (non-projective measurements), which nevertheless proved useful for theory comparisons. Fig. 22 shows a sample of these snapshots at different temperatures.

Fig. 23 shows a sample of 5-pixel by 5-pixel filters from four different CNNs that have been separately trained to distinguish single-species snapshots at half-filling at two different temperatures of $T = 0.35t$ and $2.5t$. The long-range checkerboard patterns in these filters hint at developing antiferromagnetic correlations at temperatures below $T = t$. Typical patterns drawn from other training sets using snapshots of local moments are shown in Fig. 24 for densities $\langle n \rangle = 0.58$ and 0.97. They show the anti-correlation of nearest-neighbor fermions for a density close to quarter-filling as expected and a pattern that can be interpreted as significant nearest-neighbor doublon-hole fluctuations, which are known to be large near half-filling [186]. In local moment snapshots, empty sites could represent holes or double occupancies.

Similar results obtained at a density of 0.82, where a strange metallic behavior was directly observed through the dynamical properties of the same system [187]. They point to short-range antiferromagnetic correlations when training with single-species snapshots, and more or less random patterns when training with snapshots of local moments. Training with theory snapshots leads to similar results. To shed some light on density correlations (or more accurately, local moment correlations) that might be specific to the strange metal phase, the authors designed a CNN with six
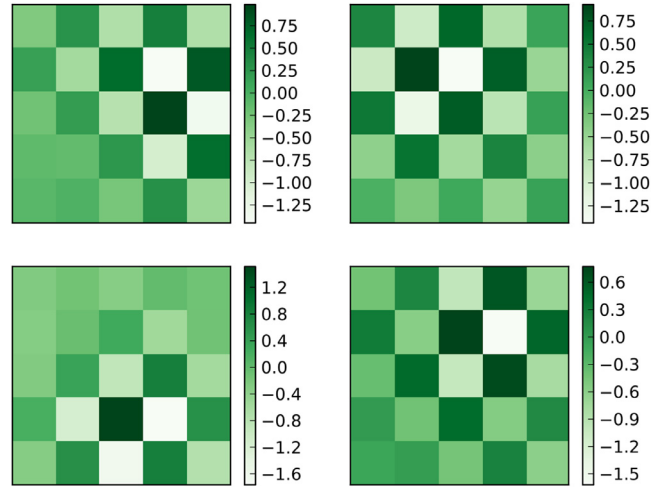
Fig. 23. Sample filters after separate training using experimental single-species snapshots near half filling. The observed checkerboard patterns point to long-range antiferromagnetic correlations in the system at low temperatures.
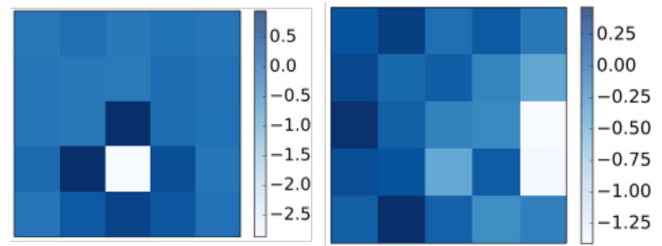
Fig. 24. Example of two filters after training using experimental local moment snapshots at densities $n = 0.58$ (left) and 0.97 (right). Signatures of nearest-neighbor anti-correlation and doublon-hole fluctuations are seen in the corresponding filters, respectively. Image reproduced from Ref. [184].

larger 7-pixel by 7-pixel filters and were able to show that the CNN trained at the density of 0.82 can as a whole act as an order parameter for the strange metal phase with a signal that decreases as one moves away from this density. They eliminated the density of particles as an obvious indicator by subtracting the network signal for the real snapshots of local moments from that obtained for the same snapshots but with randomized pixels (fake snapshots). This work demonstrated that by studying the inner workings of neural networks, one could gain unbiased insight into the physics of the problem.

Until now, this perspective has focused squarely on solving low-energy effective Hamiltonians for strongly correlated materials. However, it is important to keep in mind that the right effective models are not known or are still being debated for many classes of models. In this context, an important task is to extract effective models from inelastic scattering data by solving the *inverse* scattering problem, where the high-dimensional inelastic neutron or X-ray scattering data for a dynamical structure factor $S(\mathbf{q}, \omega)$ is used to determine the parameters of an effective model. ML methods can help with this task by formulating it as a supervised learning problem [72,188]. Essentially, one formulates a model with parameters $\boldsymbol{\theta} = [\theta_1, \theta_2, \ldots, \theta_n]$ that can be used to generate a model spectra $S_{\text{model}}(\mathbf{q}, \omega, \boldsymbol{\theta})$. One then adjusts the parameters to minimize

$$\chi^2 = ||S_{\text{model}}(\mathbf{q}, \omega, \boldsymbol{\theta}) - S(\mathbf{q}, \omega)||^2, \tag{22}$$

where $|| \cdot ||$ defines a norm [72]. Eq. (22) is minimized using AI and ML-based optimization methods. In this context, autoencoders have been particularly useful for reducing the complexity of the optimization problem [72,188,189]. For this approach to be success-

ful, however, one must rapidly generate dynamic structure factors for a given candidate model. Therefore, a crucial component here is the existence of fast solvers for the *direct* scattering problem. In many cases, this component is the bottleneck; however, methods for treating a large class of quantum magnets have recently been developed that generalize semiclassical Landau–Lifshitz dynamics to SU(N) [190–192]. These kinds of solvers have enabled the successful extraction of model parameters for $Dy_2Ti_2O_7$ [189] and α-$RuCl_3$ [193] from inelastic neutron scattering data. Achieving a similar level of success for resonant inelastic X-ray scattering experiments, with its more complex and challenging cross-section, will require additional work.

Another popular approach for obtaining model parameters relies on down-folding ab initio electronic structure calculations onto low-energy target subspaces. These calculations typically project the bands near the Fermi level onto a minimal set of maximally localized Wannier functions to obtain tight-binding parameters [194]. A set of constrained random phase approximation calculations can then be conducted to compute the corresponding interaction parameters [195,196]. ML and AI methods can also play a role here [197]. For example, techniques from these fields have been used to construct functionals [198] and accelerate density functional theory simulations [199]. They have also been applied towards extracting tight-binding [200] and interaction [201] parameters, as well as impurity and defect parameters [202]. These methods thus provide an alternative route towards deriving the appropriate effective models for many-body simulations.

7. Accelerated Monte Carlo simulations

The term "self-learning" Monte Carlo (SLMC) algorithms refers to a powerful class of ML–accelerated MCMC methods that have been developed and refined in recent years. These methods were first introduced by Liu et al. [76,203] in the context of classical MC simulations of the Ising model. Since then, the method has been expanded to a much broader class of correlated electron models, different flavors of classical and quantum MC algorithms, and ML frameworks [185,204–213].

7.1. Overview of self-learning Monte Carlo method

The objective of an SLMC algorithm is to learn an accurate proxy function for the transition probabilities between different MC configurations. In other words, the algorithm learns an effective energy E_{eff} such that

$$e^{-\beta \Delta E_{eff}(\{X\},\{X'\})} \approx P(\{X\} \rightarrow \{X'\}) = \frac{W(\{X'\})}{W(\{X\})},$$

where $W(\{X\})$ is the true MC weight of the system for a given configuration $\{X\}$ (see Section 3).

If the proxy energy $E_{eff}(\{X\})$ can be evaluated more efficiently than $W(\{X\})$, then it can be used to quickly evolve the Markov chain between largely *uncorrelated* configurations. Most implementations to date perform this task through a series of local updates in the configuration space, following the standard Metropolis–Hastings scheme. Specifically, one sweeps through all sites in configuration space proposing local updates to the MC configuration that are accepted or rejected with a probability estimated by the learned effective model

$$p_l = \min\left[1, e^{-\beta \Delta E_{eff}(\{X\},\{X'\})}\right].$$

After many sweeps through configuration space, this procedure should produce a new MC configuration $\{X'\}$ that is very far removed from the starting one. This procedure can thus be viewed as a means for constructing non-trivial *global* updates of the MC configurations. However, to maintain detailed balance [76], a final

(cumulative) acceptance step is required where the entire move $\{X\} \rightarrow \{X'\}$ is accepted with probability

$$p_c = \min\left[1, \frac{W(\{X\})}{W(\{X'\})} e^{-\beta \Delta E_{eff}(\{X\},\{X'\})}\right]. \tag{23}$$

While this final step requires evaluating the proper MC weights, it occurs infrequently enough that a considerable algorithm speedup can often be obtained, assuming p_c is not too small. A large p_c will be achieved when $e^{-\beta \Delta E_{eff}(\{X\},\{X'\})}$ is close to $W(\{X'\})/W(\{X\})$.

To illustrate this procedure, consider an MC simulation of a spin-Fermion model, where itinerant fermions are coupled to a classical spin $\mathbf{S}_i$ [203]

$$H = -\sum_{i,j,\sigma} t_{i,j} c_{i,\sigma}^\dagger c_{j,\sigma} - J \sum_{i,\sigma,\sigma'} \mathbf{S}_i \cdot c_{i,\sigma}^\dagger \boldsymbol{\tau}_{\sigma,\sigma'} c_{i,\sigma'}, \tag{24}$$

where $\boldsymbol{\tau}$ is a vector of Pauli matrices. For a given configuration of the classical spins $\{\mathbf{S}_i\}$, Eq. (24) can be diagonalized exactly to obtain the energy $E(\{\mathbf{S}_i\})$, which is an order $O(N^3)$ operation. One then samples the classical spin configurations by proposing updates of $\mathbf{S}_i \rightarrow \mathbf{S}_i'$ at each site such that the total computational cost of a full Monte Carlo sweep is $O(N^4)$.

In an SLMC implementation, one assumes that the itinerant electrons mediate an effective RKKY-like spin-spin interaction of the form

$$H_{eff} = J_0 + \sum_{\langle i,j \rangle_n} J_n \mathbf{S}_i \cdot \mathbf{S}_j + \ldots, \tag{25}$$

where J_n is the effective coupling between pairs of nth-neighbor spins. To determine the coefficients, one fits Eq. (25) to a large set of training data obtained using standard MCMC simulations of the original Hamiltonian. Once trained, the proxy energy provided by Eq. (25) can be evaluated in $O(N)$ operations as opposed to the original $O(N^3)$ operations that are necessary to evaluate the DQMC determinant ratios in $W(\{X'\})/W(\{X\})$. The effective Hamiltonian can also be easily extended to include more complicated three- and four-body interactions [76]. One can also enforce specific symmetries into the form of H_{eff} if these are known [209].

The efficiency of the SLMC approach depends crucially on the accuracy of the underlying effective model. For example, if the effective model is not sufficiently flexible to capture the training data patterns, it can be challenging to train an accurate surrogate H_{eff}. Moreover, a specified effective model's ability to accurately capture the MC weights can depend very strongly on the parameters of the full model or the simulation temperature. One might expect this limitation to some extent, as different low-energy effective models describe different ordered phases. When the effective model cannot describe the true model's physics, the SLMC algorithm will attempt to construct updates that the original MCMC algorithm would typically reject. Because of this, the final cumulative update given Eq. (23) will begin to have a high rejection rate.

In some cases, one can improve the quality of the surrogate model by including longer-range interactions, many-body interactions, or additional symmetries in the underlying model. However, no well-defined procedure exists for systematically deriving the correct effective models. To overcome this shortcoming, several research groups have attempted to implement deep learning frameworks where an artificial neural network tries to learn the form of the effective model [207,210].

7.2. Self-learning Monte Carlo for the Holstein model

To overcome the need to have a priori information about the underlying effective model in SLMC applications, some of the current authors implemented a set of artificial neural networks for performing SLMC simulations of lattice QMC simulations [210]. As
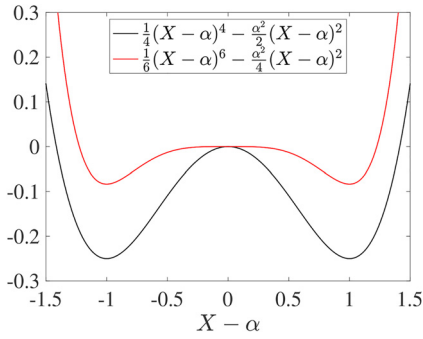
Fig. 25. The symmetric functions used to enforce the double-well potential present in the Holstein model near half-filling for the case $\alpha = -g/\omega_0^2 = 1$. Adapted from Ref. [209].

Fig. 26. A finite-size scaling analysis of the $\mathbf{q} = (\pi, \pi)$ CDW structure factor $S(\mathbf{q})N_x^{2\gamma/\nu-2}$ vs T/t. The CDW transition is in the 2D Ising universality class with critical exponents $\gamma = \frac{1}{8}$ and $\nu = 1$. The inset shows the collapse of the data for a critical temperature $T_c = 0.244$. Reproduced from Ref. [210].

a proof-of-principle, we applied this approach to simulations of the CDW transition in the two-dimensional half-filled Holstein model. We will, therefore, first give a brief overview of DQMC simulations for this challenging model.

The integral over fields in Eq. (16) is accomplished using MCMC sampling; however, two different types of update procedures are required. The first is the standard *local* updates, where the displacements are changed individually at each (i, l) site. This update can be performed on a single location with a computational cost that scales like $O(N^2)$ using an efficient Sherman–Morrison updating scheme [1]. An entire sweep of single-site updates costs $O(N^3L)$ operations to perform. The second class of updates is so-called *block* or *global* updates, where the displacements at all time slices are shifted by the same amount in a single update [47]. These updates are nonlocal in imaginary time and require one to explicitly compute the weights in Eq. (16). Since this task required evaluating a matrix determinant, the nominal cost of performing a block update at a single site is $O(N^3)$ while an entire sweep costs $O(N^4)$. Despite their higher computational costs, block updates are needed to achieve reasonably short autocorrelation times in simulations of the Holstein model. (The same is true for simulations of the Hubbard model whenever $U\beta$ is large [101]).

Reference [209] was the first to attempt to conduct SLMC simulations of the Holstein model. In this approach, the authors defined an effective model in the spirit of Eq. (25) but with terms designed to reflect the known $\mathbb{Z}_2$ ($X \to -X$) symmetry of the model. Specifically, they adopted

$$-\beta H_{\text{eff}} = J_k \sum_{i,l} (X_{i,l+1} - X_{i,l})^2 \tag{26}$$
$$+ J_p \sum_{i,l} \left(\frac{1}{4}(X-\alpha)^4 - \frac{\alpha^2}{2}(X-\alpha)^2 \right)$$
$$+ J_p' \sum_{i,l} \left(\frac{1}{6}(X-\alpha)^6 - \frac{\alpha^2}{4}(X-\alpha)^4 \right)$$
$$+ J_{nn} \sum_{\langle i,j \rangle, l} (X_{i,l} - \alpha)(X_{j,l} - \alpha)$$
$$+ J_{nn} \sum_{i, \langle l,l' \rangle} (X_{i,l} - \alpha)(X_{i,l'} - \alpha),$$

where $\alpha = -\frac{g}{\omega_0^2}$. The first term reflects the usual kinetic energy of the phonon fields, while the fourth and fifth terms reflect nearest-neighbor interactions between the lattice fields in both space and imaginary time. The second and third terms, plotted in Fig. 25, allow for the presence of a double-well potential in the Holstein model, which can become deep in the strong coupling limit.

Chen et al. [209] used this effective model to study the CDW transition in the half-filled Holstein model for $\omega_0 = t/2$ on large
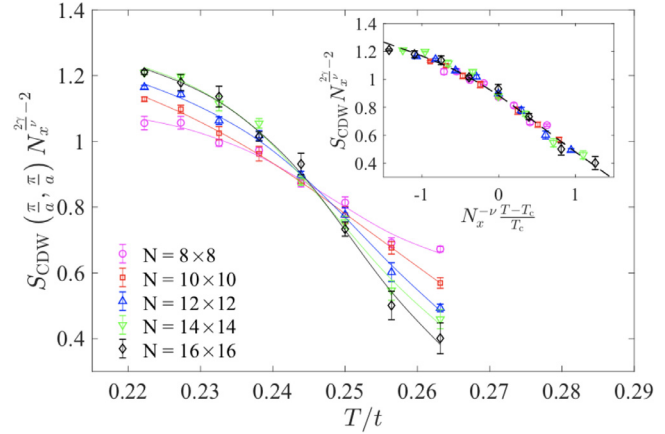
16×16 lattices, large enough to perform reliable scaling analysis to estimate T_{CDW}. The authors also reported a significant reduction in the autocorrelation time of the CDW structure factor using the SLMC method compared to the conventional DQMC algorithm. We note, however, that the authors defined a full MC sweep in the DQMC case as proposing single-site updates at all spacetime points together with four block updates at (presumably) randomly chosen sites. In the SLMC case, however, they defined an MC sweep as a sweep of local updates followed by a Wolff-cluster update. The sampling procedures differ, so it is unclear what the ultimate mechanisms of the autocorrelation reduction were.

Li et al. [210] later employed artificial neural networks to perform SLMC simulations of the Holstein model, where the form of the effective Hamiltonian was learned directly from the training data. In this case, the authors trained two neural networks: a fully connected network to perform local single-site updates and a convolutional neural network to act as a proxy for global moves. Like the effective model of Chen et al. [209], both networks used the proposed change in displacement $\Delta X_{i,l}$ (or ΔX for global updates) together with *local* nearest neighbor information about the surrounding fields as input features. The networks could thus be trained using data generated using cheaper small clusters before being deployed on larger systems. For example, Li et al. trained their networks on 6×6 clusters before deploying them on system sizes up to 16×16.

Li et al. [210], benchmarked their method for the challenging half-filled model with $\omega_0 = t/2$. They found that the neural networks were able to reproduce the results of the conventional DQMC algorithm accurately and obtained a thermodynamic value of T_{CDW} similar to Chen et al. [209] (see Fig. 26). Li et al. also examined the autocorrelation time for their approach. Because they were able to perform both global and local updates within the SLMC framework, the authors could make more meaningful comparisons with DQMC by defining an MC sweep in the same way for each case (a full sweep of single-site updates, followed by a block update performed at every site). In doing so, they found that the SLMC approach produced *identical* autocorrelation times compared to the conventional algorithm. This result makes intuitive sense. Suppose the SLMC algorithm is accurately reproducing the accept/reject steps of the DQMC algorithm. In that case, it should produce the same autocorrelations in the Monte Carlo configurations, provided the SLMC and DQMC algorithms perform the same updates. This notion was reinforced by the observed behavior in the cumulative acceptance rates. For the neural net-

works, the cumulative acceptance rate remained near 100%, independent of the temperature of the simulation. On the other hand, Chen et al. [209] obtained cumulative acceptance rates that fell off rapidly as the temperature was lowered.

Despite these successes near half-filling, it is essential to note that the neural-network-based method has not yet been generalized to the metallic state that appears with doping. This failure occurs even though the networks often accurately reproduce the validation data sets. We speculate that this issue might be related to longer-range effective interactions in space or imaginary time that are not captured by the local input features used in the neural networks. For example, it is possible that no local models like Eq. (26) or Eq. (25) exists for this problem. One possible solution would be to develop neural networks that use the entire set of auxiliary fields as input features. While costly, modern GPU resources would make this feasible. Additional research is needed to push this technology further.

8. Outlook and conclusions

From this short perspective, we hope it is clear that over the last several years, methods in artificial intelligence, machine learning, and data science have leaped from novel tools for studying *classical* phase transitions to providing powerful new means to solve *quantum* many-body problems. Moreover, these methods can provide a pathway forward to solving new and challenging forefront questions in strongly correlated materials. However, we also believe it is important to take a step back and remind ourselves that there remains work to be done in this area.

The first issue concerns the 'value added'. For example, it is clear that PCA can capture phase transitions of many interesting quantum models of interacting fermions. However, it is important to consider whether it (or other ML and AI methods) can do so *more effectively* than data analysis tools that have been refined over the last several decades centered on sophisticated finite-size scaling (FSS) of appropriate correlation functions. That is, in what ways do ML and AI 'revolutionize' the field of computational quantum many-body physics? In the examples noted above, the characterizations of the critical points with ML and AI are *less precise*. Part of this is, of course, due to the time frames of the approaches; ML and AI have not yet matured in these investigations. There is roughly an order of magnitude difference in the time frames over which "traditional" and ML approaches have been under development. It seems likely that, given the opportunity, ML will approach the accuracy of FSS-based methods, but it is too soon to know if they will surpass them.

More promising, ML and AI-based methods can provide additional information, such as the ordering patterns in experimental images of fermions or identifying the presence of broken symmetries in MCMC simulations without having to know which observable to measure in advance. This problem is one area where these methods can provide new insights; similarly, one can envision hybrid approaches where a ML method quick surveys the model phase space to identify promising areas that are then expanded on using conventional methods.

We believe that a particularly important set of opportunities involves deploying ML and AI to attack the bottlenecks of quantum simulations. These include 1) speeding up the generation of independent samples in regimes of long autocorrelation times, 2) identifying complex intertwined orders, and 3) (possibly) helping to mitigate the sign problem. ML applications to address the first are perhaps the most mature with the rapid development of the SLMC method. Nevertheless, SLMC methods have yet to be widely adopted by QMC practitioners. This is likely because the approach has only been validated on a small subset of problems and model parameters and often fails to generalize to other regimes. Addi-

tional research will be needed to develop robust design and deployment methods. For example, determining ways to quantify the level of trust in the model proxy's ability to sample the entire configuration space would be crucial. For example, it is entirely possible that the effective model used in an SLMC simulation can introduce more local minima in the MCMC configuration space that could introduce ergodicity issues. Another promising line of research involves the extension of matrix product states and tensor networks using CNNs or restricted Boltzmann machines [214–217].

Finally, using ML and AI-methods to solve the inverse scattering problem, where model parameters are extracted from inelastic neutron and X-ray scattering data, is a growing and promising avenue of research. This application, however, urgently needs fast methods for solving the direct scattering problem. Of course, such solvers can be AI-accelerated, but other avenues should also be explored [104,191,192].

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability

Data will be made available on request.

References

[1] S.R. White, D.J. Scalapino, R.L. Sugar, E.Y. Loh, J.E. Gubernatis, R.T. Scalettar, Numerical study of the two-dimensional Hubbard model, Phys. Rev. B 40 (1989) 506–516, doi:10.1103/PhysRevB.40.506.

[2] H. Park, K. Haule, G. Kotliar, Cluster dynamical mean field theory of the Mott transition, Phys. Rev. Lett. 101 (2008) 186403, doi:10.1103/PhysRevLett.101.186403.

[3] M. Jarrell, T. Maier, M.H. Hettler, A.N. Tahvildarzadeh, Phase diagram of the Hubbard model: beyond the dynamical mean field, Europhys. Lett. (EPL) 56 (4) (2001) 563–569, doi:10.1209/epl/i2001-00557-x.

[4] K. Ido, T. Ohgoe, M. Imada, Competition among various charge-inhomogeneous states and d-wave superconducting state in Hubbard models on square lattices, Phys. Rev. B 97 (2018) 045138, doi:10.1103/PhysRevB.97.045138.

[5] S.R. White, D.J. Scalapino, Density matrix renormalization group study of the striped phase in the 2D $t - J$ model, Phys. Rev. Lett. 80 (1998) 1272–1275, doi:10.1103/PhysRevLett.80.1272.

[6] A. Macridin, M. Jarrell, Bond excitations in the pseudogap phase of the Hubbard model, Phys. Rev. B 78 (2008) 241101, doi:10.1103/PhysRevB.78.241101.

[7] S.R. White, D.J. Scalapino, Energetics of domain walls in the 2D $t - J$ model, Phys. Rev. Lett. 81 (1998) 3227–3230, doi:10.1103/PhysRevLett.81.3227.

[8] E.W. Huang, C.B. Mendl, H.-C. Jiang, B. Moritz, T.P. Devereaux, Stripe order from the perspective of the Hubbard model, npj Quantum Mater. 3 (1) (2018) 22.

[9] T.A. Maier, M. Jarrell, T.C. Schulthess, P.R.C. Kent, J.B. White, Systematic study of d-wave superconductivity in the 2D repulsive Hubbard model, Phys. Rev. Lett. 95 (2005) 237001.

[10] S. Li, A. Nocera, U. Kumar, S. Johnston, Particle-hole asymmetry in the dynamical spin and charge responses of corner-shared 1D cuprates, Commun. Phys. 4 (1) (2021) 217, doi:10.1038/s42005-021-00718-w.

[11] P. Mai, S. Karakuzu, G. Balduzzi, S. Johnston, T.A. Maier, Intertwined spin, charge, and pair correlations in the two-dimensional Hubbard model in the thermodynamic limit, Proc. Natl. Acad. Sci. 119 (7) (2022), doi:10.1073/pnas.2112806119. e2112806119

[12] E.W. Huang, T. Liu, W.O. Wang, J.-C. Jiang, P. Mai, T.A. Maier, S. Johnston, B. Moritz, T.P. Devereaux, Fluctuating intertwined stripes in the strange metal regime of the Hubbard model, arXiv:2202.08845(2022).

[13] M. Qin, C.-M. Chung, H. Shi, E. Vitali, C. Hubig, U. Schollwöck, S.R. White, S. Zhang, Absence of superconductivity in the pure two-dimensional Hubbard model, Phys. Rev. X 10 (2020) 031016, doi:10.1103/PhysRevX.10.031016.

[14] S. Jiang, D.J. Scalapino, S.R. White, Ground-state phase diagram of the $t - t' - J$ model, Proc. Natl. Acad. Sci. 118 (44) (2021), doi:10.1073/pnas.2109978118. e2109978118

[15] B.-X. Zheng, C.-M. Chung, P. Corboz, G. Ehlers, M.-P. Qin, R.M. Noack, H. Shi, S.R. White, S. Zhang, G.K.-L. Chan, Stripe order in the underdoped region of the two-dimensional Hubbard model, Science 358 (6367) (2017) 1155–1160, doi:10.1126/science.aam7127.

[16] S.R. White, D.J. Scalapino, Stripes on a 6-leg Hubbard ladder, Phys. Rev. Lett. 91 (2003) 136403.

[17] B. Xiao, Y.-Y. He, A. Georges, S. Zhang, Temperature dependence of spin and charge orders in the doped two-dimensional Hubbard model, arXiv:2202. 11741(2022).

[18] A. Wietek, Y.-Y. He, S.R. White, A. Georges, E.M. Stoudenmire, Stripes, antiferromagnetism, and the pseudogap in the doped Hubbard model at finite temperature, Phys. Rev. X 11 (2021) 031007, doi:10.1103/PhysRevX.11.031007.

[19] R.T. Scalettar, N.E. Bickers, D.J. Scalapino, Competition of pairing and Peierls–charge-density-wave correlations in a two-dimensional electron-phonon model, Phys. Rev. B 40 (1989) 197–200, doi:10.1103/PhysRevB.40.197.

[20] M. Capone, P. Carta, S. Ciuchi, Dynamical mean field theory of polarons and bipolarons in the half-filled Holstein model, Phys. Rev. B 74 (2006) 045106, doi:10.1103/PhysRevB.74.045106.

[21] J.P. Hague, Superconducting states of the quasi-2D Holstein model: effects of vertex and non-local corrections, J. Phys. 17 (37) (2005) 5663–5676, doi:10.1088/0953-8984/17/37/005.

[22] P.M. Dee, J. Coulter, K.G. Kleiner, S. Johnston, Relative importance of nonlinear electron-phonon coupling and vertex corrections in the Holstein model, Commun. Phys. 3 (1) (2020) 145, doi:10.1038/s42005-020-00413-2.

[23] B. Nosarzewski, E.W. Huang, P.M. Dee, I. Esterlis, B. Moritz, S.A. Kivelson, S. Johnston, T.P. Devereaux, Superconductivity, charge density waves, and bipolarons in the Holstein model, Phys. Rev. B 103 (2021) 235156, doi:10.1103/PhysRevB.103.235156.

[24] O. Bradley, G.G. Batrouni, R.T. Scalettar, Superconductivity and charge density wave order in the two-dimensional Holstein model, Phys. Rev. B 103 (2021) 235104, doi:10.1103/PhysRevB.103.235104.

[25] B. Cohen-Stead, K. Barros, Z.Y. Meng, C. Chen, R.T. Scalettar, G.G. Batrouni, Langevin simulations of the half-filled cubic Holstein model, Phys. Rev. B 102 (2020) 161108, doi:10.1103/PhysRevB.102.161108.

[26] D.J.J. Marchand, G. De Filippis, V. Cataudella, M. Berciu, N. Nagaosa, N.V. Prokof'ev, A.S. Mishchenko, P.C.E. Stamp, Sharp transition for single polarons in the one-dimensional Su–Schrieffer–Heeger model, Phys. Rev. Lett. 105 (2010) 266605, doi:10.1103/PhysRevLett.105.266605.

[27] M. Weber, F.F. Assaad, M. Hohenadler, Excitation spectra and correlation functions of quantum Su–Schrieffer–Heeger models, Phys. Rev. B 91 (2015) 245147, doi:10.1103/PhysRevB.91.245147.

[28] J. Sous, M. Chakraborty, R.V. Krems, M. Berciu, Light bipolarons stabilized by Peierls electron-phonon coupling, Phys. Rev. Lett. 121 (2018) 247001, doi:10.1103/PhysRevLett.121.247001.

[29] S. Li, S. Johnston, Quantum Monte Carlo study of lattice polarons in the two-dimensional three-orbital Su–Schrieffer–Heeger model, npj Quantum Mater. 5 (1) (2020) 40, doi:10.1038/s41535-020-0243-3.

[30] X. Cai, Z.-X. Li, H. Yao, Robustness of Antiferromagnetism in the Su–Schrieffer–Heeger–Hubbard model, arXiv:2112.14744(2021).

[31] A. Götz, S. Beyl, M. Hohenadler, F.F. Assaad, Valence-bond solid to antiferromagnet transition in the two-dimensional Su–Schrieffer–Heeger model by Langevin dynamics, Phys. Rev. B 105 (2022) 085151, doi:10.1103/PhysRevB.105.085151.

[32] B. Xing, W.-T. Chiu, D. Poletti, R.T. Scalettar, G. Batrouni, Quantum Monte Carlo simulations of the 2D Su–Schrieffer–Heeger model, Phys. Rev. Lett. 126 (2021) 017601, doi:10.1103/PhysRevLett.126.017601.

[33] C. Zhang, J. Sous, D.R. Reichman, M. Berciu, A.J. Millis, N.V. Prokof'ev, B.V. Svistunov, Bipolaronic high-temperature superconductivity, arXiv:2203. 07380(2022).

[34] M. Jarrell, H. Akhlaghpour, T. Pruschke, Periodic Anderson model in infinite dimensions, Phys. Rev. Lett. 70 (1993) 1670–1673, doi:10.1103/PhysRevLett.70. 1670.

[35] M. Vekić, J.W. Cannon, D.J. Scalapino, R.T. Scalettar, R.L. Sugar, Competition between antiferromagnetic order and spin-liquid behavior in the two-dimensional periodic Anderson model at half filling, Phys. Rev. Lett. 74 (12) (1995) 2367.

[36] K. Held, C. Huscroft, R.T. Scalettar, A.K. McMahan, Similarities between the Hubbard and periodic Anderson models at finite temperatures, Phys. Rev. Lett. 85 (2) (2000) 373.

[37] P. Sun, G. Kotliar, Extended dynamical mean field theory study of the periodic Anderson model, Phys. Rev. Lett. 91 (3) (2003) 037209.

[38] D.J. Luitz, F.F. Assaad, Weak coupling continuous time quantum Monte Carlo study of the single impurity and periodic Anderson models with s-wave superconducting baths, Phys. Rev. B 81 (2) (2010) 024509.

[39] W. Wu, A.-M.-S. Tremblay, D-wave superconductivity in the frustrated two-dimensional periodic Anderson model, Phys. Rev. X 5 (1) (2015) 011019.

[40] P. Werner, A.J. Millis, Efficient dynamical mean field simulation of the Holstein–Hubbard model, Phys. Rev. Lett. 99 (2007) 146404, doi:10.1103/PhysRevLett.99.146404.

[41] H. Fehske, G. Hager, E. Jeckelmann, Metallicity in the half-filled Holstein–Hubbardmodel, Europhys. Lett. 84 (2008) 57001, doi:10.1209/0295-5075/84/57001.

[42] A. Nocera, M. Soltanieh-ha, C.A. Perroni, V. Cataudella, A.E. Feiguin, Interplay of charge, spin, and lattice degrees of freedom in the spectral properties of the one-dimensional Hubbard–Holstein model, Phys. Rev. B 90 (2014) 195134, doi:10.1103/PhysRevB.90.195134.

[43] J. Greitemann, S. Hesselmann, S. Wessel, F.F. Assaad, M. Hohenadler, Finite-size effects in Luther–Emery phases of Holstein and Hubbard models, Phys. Rev. B 92 (2015) 245132, doi:10.1103/PhysRevB.92.245132.

[44] M. Hohenadler, Interplay of site and bond electron-phonon coupling in one dimension, Phys. Rev. Lett. 117 (2016) 206404, doi:10.1103/PhysRevLett.117. 206404.

[45] Y. Wang, I. Esterlis, T. Shi, J.I. Cirac, E. Demler, Zero-temperature phases of the two-dimensional Hubbard–Holstein model: a non-Gaussian exact diagonalization study, Phys. Rev. Res. 2 (2020) 043258, doi:10.1103/PhysRevResearch. 2.043258.

[46] N.C. Costa, K. Seki, S. Sorella, Magnetism and charge order in the honeycomb lattice, Phys. Rev. Lett. 126 (2021) 107205, doi:10.1103/PhysRevLett.126. 107205.

[47] S. Johnston, E.A. Nowadnick, Y.F. Kung, B. Moritz, R.T. Scalettar, T.P. Devereaux, Determinant quantum Monte Carlo study of the two-dimensional single-band Hubbard–Holstein model, Phys. Rev. B 87 (2013) 235133, doi:10.1103/PhysRevB.87.235133.

[48] S. Karakuzu, A.T. Ly, P. Mai, J. Neuhaus, T.A. Maier, S. Johnston, Stripe correlations in the two-dimensional Hubbard–Holstein model, arXiv:2205. 15464(2022).

[49] S. Li, S. Johnston, Suppressed superexchange interactions in the cuprates by bond-stretching oxygen phonons, arXiv:2205.12678(2022).

[50] J.E. Hirsch, Quantum Monte Carlo and exact diagonalization study of a dynamic Hubbard model, Phys. Rev. B 65 (2002) 214510, doi:10.1103/PhysRevB. 65.214510.

[51] S. Li, E.A. Nowadnick, S. Johnston, Quasiparticle properties of the nonlinear Holstein model at finite doping and temperature, Phys. Rev. B 92 (2015) 064301, doi:10.1103/PhysRevB.92.064301.

[52] T. Ayral, S. Biermann, P. Werner, Screening and nonlocal correlations in the extended Hubbard model from self-consistent combined GW and dynamical mean field theory, Phys. Rev. B 87 (2013) 125149, doi:10.1103/PhysRevB.87. 125149.

[53] J. Paki, H. Terletska, S. Iskakov, E. Gull, Charge order and antiferromagnetism in the extended Hubbard model, Phys. Rev. B 99 (2019) 245146, doi:10.1103/PhysRevB.99.245146.

[54] P.M. Dee, J. Coulter, K.G. Kleiner, S. Johnston, Relative importance of nonlinear electron-phonon coupling and vertex corrections in the Holstein model, Commun. Phys. 3 (1) (2020) 1–7, doi:10.1038/s42005-020-00413-2.

[55] Z. Chen, Y. Wang, S.N. Rebec, T. Jia, M. Hashimoto, D. Lu, B. Moritz, R.G. Moore, T.P. Devereaux, Z.-X. Shen, Anomalously strong near-neighbor attraction in doped 1D cuprate chains, Science 373 (6560) (2021) 1235–1239, doi:10.1126/science.abf5174.

[56] E.W. Huang, C.B. Mendl, S. Liu, S. Johnston, H.-C. Jiang, B. Moritz, T.P. Devereaux, Numerical evidence of fluctuating stripes in the normal state of high-T_c cuprate superconductors, Science 358 (6367) (2017) 1161–1164, doi:10. 1126/science.aak9546.

[57] E.Y. Loh, J.E. Gubernatis, R.T. Scalettar, S.R. White, D.J. Scalapino, R.L. Sugar, Sign problem in the numerical simulation of many-electron systems, Phys. Rev. B 41 (1990) 9301–9307, doi:10.1103/PhysRevB.41.9301.

[58] C. Wu, S.-C. Zhang, Sufficient condition for absence of the sign problem in the fermionic quantum Monte Carlo algorithm, Phys. Rev. B 71 (2005) 155115, doi:10.1103/PhysRevB.71.155115.

[59] M. Troyer, U.-J. Wiese, Computational complexity and fundamental limitations to fermionic quantum Monte Carlo simulations, Phys. Rev. Lett. 94 (2005) 170201, doi:10.1103/PhysRevLett.94.170201.

[60] S. Chandrasekharan, Fermion bag approach to lattice field theories, Phys. Rev. D 82 (2010) 025007, doi:10.1103/PhysRevD.82.025007.

[61] Z.-X. Li, Y.-F. Jiang, H. Yao, Majorana-time-reversal symmetries: a fundamental principle for sign-problem-free quantum Monte Carlo simulations, Phys. Rev. Lett. 117 (2016) 267002, doi:10.1103/PhysRevLett.117.267002.

[62] M. Iazzi, A.A. Soluyanov, M. Troyer, Topological origin of the fermion sign problem, Phys. Rev. B 93 (2016) 115102, doi:10.1103/PhysRevB.93.115102.

[63] D. Hangleiter, I. Roth, D. Nagaj, J. Eisert, Easing the Monte Carlo sign problem, Sci. Adv. 6 (33) (2020), doi:10.1126/sciadv.abb8341.

[64] R. Mondaini, S. Tarat, R.T. Scalettar, Quantum critical points and the sign problem, Science 375 (6579) (2022) 418–424, doi:10.1126/science.abg9299.

[65] R.H. Swendsen, J.-S. Wang, Nonuniversal critical dynamics in Monte Carlo simulations, Phys. Rev. Lett. 58 (1987) 86–88, doi:10.1103/PhysRevLett.58.86.

[66] U. Wolff, Lattice field theory as a percolation process, Phys. Rev. Lett. 60 (1988) 1461–1463, doi:10.1103/PhysRevLett.60.1461.

[67] R.G. Edwards, A.D. Sokal, Generalization of the Fortuin–Kasteleyn–Swendsen–Wang representation and Monte Carlo algorithm, Phys. Rev. D 38 (1988) 2009–2012, doi:10.1103/PhysRevD.38.2009.

[68] G. Carleo, I. Cirac, K. Cranmer, L. Daudet, M. Schuld, N. Tishby, L. Vogt-Maranto, L. Zdeborová, Machine learning and the physical sciences, Rev. Mod. Phys. 91 (2019) 045002, doi:10.1103/RevModPhys.91.045002.

[69] J. Carrasquilla, Machine learning for quantum matter, Adv. Phys. 5 (1) (2020) 1797528, doi:10.1080/23746149.2020.1797528.

[70] M. Feickert, B. Nachman, A living review of machine learning for particle physics, arXiv:2102.02770s(2021).

[71] G.E. Karniadakis, I.G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang, L. Yang, Physics-informed machine learning, Nat. Rev. Phys. 3 (6) (2021) 422–440, doi:10.1038/s42254-021-00314-5.

[72] Z. Chen, N. Andrejevic, N.C. Drucker, T. Nguyen, R.P. Xian, T. Smidt, Y. Wang, R. Ernstorfer, D.A. Tennant, M. Chan, M. Li, Machine learning on neutron and X-ray scattering and spectroscopies, Chem. Phys. Rev. 2 (3) (2021) 031301, doi:10.1063/5.0049111.

[73] D.A. Roberts, S. Yaida, B. Hanin, The Principles of Deep Learning Theory, Cambridge University Press, 2022.

[74] A. Dawid, J. Arnold, B. Requena, A. Gresch, M. Płodzień, K. Donatella, K.A. Nicoli, P. Stornati, R. Koch, M. Büttner, R. Okuła, G. Muñoz-Gil, R.A. Vargas-Hernández, A. Cervera-Lierta, J. Carrasquilla, V. Dunjko, M. Gabrié, P. Huembeli, E. van Nieuwenburg, F. Vicentini, L. Wang, S.J. Wetzel, G. Carleo, E. Greplová, R. Krems, F. Marquardt, M. Tomza, M. Lewenstein, A. Dauphin, Modern applications of machine learning in quantum sciences, arXiv:2204.04198(2022).

[75] L. Onsager, Crystal statistics. I. A two-dimensional model with an order-disorder transition, Phys. Rev. 65 (3–4) (1944) 117–149, doi:10.1103/PhysRev.65.117.

[76] J. Liu, Y. Qi, Z.Y. Meng, L. Fu, Self-learning Monte Carlo method, Phys. Rev. B 95 (2017) 041101, doi:10.1103/PhysRevB.95.041101.

[77] S.J. Wetzel, M. Scherzer, Machine learning of explicit order parameters: from the Ising model to SU(2) lattice gauge theory, Phys. Rev. B 96 (2017) 184410, doi:10.1103/PhysRevB.96.184410.

[78] D. Kim, D.-H. Kim, Smallest neural network to learn the Ising criticality, Phys. Rev. E 98 (2018) 022138, doi:10.1103/PhysRevE.98.022138.

[79] A. Morningstar, R.G. Melko, Deep learning the Ising model near criticality, J. Mach. Learn. Res. 18 (163) (2018) 1–17.

[80] C. Alexandrou, A. Athenodorou, C. Chrysostomou, S. Paul, The critical temperature of the 2D-Ising model through deep learning autoencoders, Eur. Phys. J. B 93 (12) (2020) 1140, doi:10.1140/epjb/e2020-100506-5.

[81] D. Yevick, Variational autoencoder analysis of Ising model statistical distributions and phase transitions, arXiv:2104.06368(2021).

[82] D. Agrawal, A. Del Maestro, S. Johnston, J. Ostrowski, A group-equivariant autoencoder for identifying spontaneously broken symmetries in the Ising model, arXiv:2202.06319(2022).

[83] J. Stephenson, Ising-model spin correlations on the triangular lattice, J. Math. Phys. 5 (8) (1964) 1009–1024.

[84] D.P. Landau, Critical and multicritical behavior in a triangular-lattice-gas Ising model: repulsive nearest-neighbor and attractive next-nearest-neighbor coupling, Phys. Rev. B 27 (9) (1983) 5604.

[85] R. Moessner, S.L. Sondhi, Ising models of quantum frustration, Phys. Rev. B 63 (22) (2001) 224401.

[86] D. Sherrington, S. Kirkpatrick, Solvable model of a spin-glass, Phys. Rev. Lett. 35 (26) (1975) 1792.

[87] R.R.P. Singh, S. Chakravarty, Critical behavior of an Ising spin-glass, Phys. Rev. Lett. 57 (2) (1986) 245.

[88] W.L. McMillan, Scaling theory of Ising spin glasses, J. Phys. C 17 (18) (1984) 3179.

[89] M. Blume, Theory of the first-order magnetic phase change in UO_2, Phys. Rev. 141 (1966) 517–524, doi:10.1103/PhysRev.141.517.

[90] H.W. Capel, On the possibility of first-order phase transitions in Ising systems of triplet ions with zero-field splitting, Physica 32 (5) (1966) 966–988, doi:10.1016/0031-8914(66)90027-9.

[91] J. Hubbard, B.H. Flowers, Electron correlations in narrow energy bands, Proc. R. Soc. Lond. Ser. A 276 (1365) (1963) 238–257, doi:10.1098/rspa.1963.0204.

[92] E.H. Lieb, F.Y. Wu, Absence of Mott transition in an exact solution of the short-range, one-band model in one dimension, Phys. Rev. Lett. 21 (1968) 192, doi:10.1103/PhysRevLett.21.192.2.

[93] G. Sordi, P. Sémon, K. Haule, A.M.S. Tremblay, Strong coupling superconductivity, pseudogap, and Mott transition, Phys. Rev. Lett. 108 (2012) 216401.

[94] R. Peters, N. Kawakami, Spin density waves in the Hubbard model: a DMFT approach, Phys. Rev. B 89 (2014) 155134, doi:10.1103/PhysRevB.89.155134.

[95] B.-X. Zheng, C.-M. Chung, P. Corboz, G. Ehlers, M.-P. Qin, R.M. Noack, H. Shi, S.R. White, S. Zhang, G.K.-L. Chan, Stripe order in the underdoped region of the two-dimensional Hubbard model, Science 358 (6367) (2017) 1155–1160.

[96] E.W. Huang, R. Sheppard, B. Moritz, T.P. Devereaux, Strange metallicity in the doped Hubbard model, Science 366 (6468) (2019) 987–990, doi:10.1126/science.aau7063.

[97] E. Gull, O. Parcollet, A.J. Millis, Superconductivity and the pseudogap in the two-dimensional Hubbard model, Phys. Rev. Lett. 110 (2013) 216405, doi:10.1103/PhysRevLett.110.216405.

[98] C.-M. Chung, M. Qin, S. Zhang, U. Schollwöck, S.R. White, Plaquette versus ordinary d-wave pairing in the t'-Hubbard model on a width-4 cylinder, Phys. Rev. B 102 (2020) 041106.

[99] P. Mai, G. Balduzzi, S. Johnston, T. Maier, Orbital structure of the effective pairing interaction in the high-temperature superconducting cuprates, npj Quantum Mater. 6 (2021) 26.

[100] T. Holstein, Studies of polaron motion: part I. The molecular-crystal model, Ann. Phys. 8 (3) (1959) 325–342.

[101] R.T. Scalettar, R.M. Noack, R.R.P. Singh, Ergodicity at large couplings with the determinant Monte Carlo algorithm, Phys. Rev. B 44 (1991) 10502–10507, doi:10.1103/PhysRevB.44.10502.

[102] P.M. Dee, K. Nakatsukasa, Y. Wang, S. Johnston, Temperature-filling phase diagram of the two-dimensional Holstein model in the thermodynamic limit by self-consistent Migdal approximation, Phys. Rev. B 99 (2019) 024514, doi:10.1103/PhysRevB.99.024514.

[103] I. Esterlis, B. Nosarzewski, E.W. Huang, B. Moritz, T.P. Devereaux, D.J. Scalapino, S.A. Kivelson, Breakdown of the Migdal–Eliashberg theory: adeterminant quantum Monte Carlo study, Phys. Rev. B 97 (2018) 140501, doi:10.1103/PhysRevB.97.140501.

[104] B. Cohen-Stead, O. Bradley, C. Miles, G. Batrouni, R. Scalettar, K. Barros, Fast and scalable quantum Monte Carlo simulations of electron-phonon models, Phys. Rev. E 105 (2022) 065302, doi:10.1103/PhysRevE.105.065302.

[105] J. Gubernatis, N. Kawashima, P. Werner, Quantum Monte Carlo methods: Algorithms for Lattice models, Cambridge University Press, 2016, doi:10.1017/CBO9780511902581.

[106] F. Becca, S. Sorella, Quantum Monte Carlo approaches for correlated systems, Cambridge University Press, 2017, doi:10.1017/9781316417041.

[107] W.K. Hastings, et al., Monte Carlo sampling methods using Markov chains and their applications, Biometrika 57 (1) (1970) 97–109, doi:10.1093/biomet/57.1.97.

[108] K. Binder, D. Heermann, L. Roelofs, A.J. Mallinckrodt, S. McKay, Monte Carlo simulation in statistical physics, Comput. Phys. 7 (2) (1993) 156–157.

[109] F.F. Assaad, Quantum Monte Carlo methods on lattices: the determinantal approach, Quantum Simul. Complex Many-Body Syst. 10 (2002) 99–147.

[110] R.R.d. Santos, Introduction to quantum Monte Carlo simulations for fermionic systems, Braz. J. Phys. 33 (2003) 36.

[111] F.F. Assaad, H.G. Evertz, World-line and determinantal quantum Monte Carlo methods for spins, phonons and electrons, in: Computational Many-Particle Physics, Springer, 2008, pp. 277–356.

[112] H.F. Trotter, On the product of semi-groups of operators, Proc. Am. Math. Soc. 10 (4) (1959) 545–551.

[113] M. Suzuki, Relationship between d-dimensional quantal spin systems and $(d+1)$-dimensional Ising systems: equivalence, critical exponents and systematic approximants of the partition function and spin correlations, Prog. Theor. Phys. 56 (5) (1976) 1454–1469.

[114] R.M. Fye, New results on Trotter-like approximations, Phys. Rev. B 33 (9) (1986) 6271.

[115] J.E. Hirsch, Discrete Hubbard–Stratonovich transformation for fermion lattice models, Phys. Rev. B 28 (1983) 4059–4061, doi:10.1103/PhysRevB.28.4059.

[116] R. Blankenbecler, D.J. Scalapino, R.L. Sugar, Monte Carlo calculations of coupled boson-fermion systems. I, Phys. Rev. D 24 (1981) 2278–2286, doi:10.1103/PhysRevD.24.2278.

[117] A. Tomas, C.-C. Chang, R. Scalettar, Z. Bai, Advancing large scale many-body QMC simulations on GPU accelerated multicore systems, in: 2012 IEEE 26th International Parallel and Distributed Processing Symposium, 2012, pp. 308–319, doi:10.1109/IPDPS.2012.37.

[118] P. Henelius, A.W. Sandvik, Sign problem in Monte Carlo simulations of frustrated quantum spin systems, Phys. Rev. B 62 (2000) 1102–1113, doi:10.1103/PhysRevB.62.1102.

[119] S. Chandrasekharan, U.-J. Wiese, Meron-cluster solution of fermion sign problems, Phys. Rev. Lett. 83 (1999) 3116–3119, doi:10.1103/PhysRevLett.83.3116.

[120] C. Wu, J.-p. Hu, S.-c. Zhang, Exact SO(5) symmetry in the spin-3/2 fermionic system, Phys. Rev. Lett. 91 (2003) 186402, doi:10.1103/PhysRevLett.91.186402.

[121] E. Berg, M.A. Metlitski, S. Sachdev, Sign-problem–free quantum Monte Carlo of the onset of antiferromagnetism in metals, Science 338 (6114) (2012) 1606–1609, doi:10.1126/science.1227769.

[122] S. Chandrasekharan, Solutions to sign problems in lattice Yukawa models, Phys. Rev. D 86 (2012) 021701, doi:10.1103/PhysRevD.86.021701.

[123] Z. Cai, H.-H. Hung, L. Wang, C. Wu, Quantum magnetic properties of the $SU(2N)$ Hubbard model in the square lattice: a quantum Monte Carlo study, Phys. Rev. B 88 (2013) 125108, doi:10.1103/PhysRevB.88.125108.

[124] E.F. Huffman, S. Chandrasekharan, Solution to sign problems in half-filled spin-polarized electronic systems, Phys. Rev. B 89 (2014) 111101, doi:10.1103/PhysRevB.89.111101.

[125] L. Wang, Y.-H. Liu, M. Iazzi, M. Troyer, G. Harcos, Split orthogonal group: a guiding principle for sign-problem-free fermionic simulations, Phys. Rev. Lett. 115 (2015) 250601, doi:10.1103/PhysRevLett.115.250601.

[126] R.K. Kaul, Marshall-positive SU(N) quantum spin systems and classical loop models: a practical strategy to design sign-problem-free spin Hamiltonians, Phys. Rev. B 91 (2015) 054413, doi:10.1103/PhysRevB.91.054413.

[127] Z.-X. Li, Y.-F. Jiang, H. Yao, Solving the fermion sign problem in quantum Monte Carlo simulations by Majorana representation, Phys. Rev. B 91 (2015) 241117, doi:10.1103/PhysRevB.91.241117.

[128] Z.C. Wei, C. Wu, Y. Li, S. Zhang, T. Xiang, Majorana positivity and the fermion sign problem of quantum Monte Carlo simulations, Phys. Rev. Lett. 116 (2016) 250601, doi:10.1103/PhysRevLett.116.250601.

[129] Z.-X. Li, H. Yao, Sign-problem-free fermionic quantum Monte Carlo: developments and applications, Annu. Rev. Condens. Matter Phys. 10 (1) (2019) 337–356, doi:10.1146/annurev-conmatphys-033117-054307.

[130] A.J. Kim, P. Werner, R. Valentí, Alleviating the sign problem in quantum Monte Carlo simulations of spin-orbit-coupled multiorbital Hubbard models, Phys. Rev. B 101 (2020) 045108, doi:10.1103/PhysRevB.101.045108.

[131] R. Levy, B.K. Clark, Mitigating the sign problem through basis rotations, Phys. Rev. Lett. 126 (2021) 216401, doi:10.1103/PhysRevLett.126.216401.

[132] X. Zhang, G. Pan, X.Y. Xu, Z.Y. Meng, Sign problem finds its bounds, arXiv preprint arXiv:2112.06139(2021).

[133] T. Maier, M. Jarrell, T. Pruschke, M.H. Hettler, Quantum cluster theories, Rev. Mod. Phys. 77 (2005) 1027–1080, doi:10.1103/RevModPhys.77.1027.

[134] A. Georges, G. Kotliar, W. Krauth, M.J. Rozenberg, Dynamical mean-field theory of strongly correlated fermion systems and the limit of infinite dimensions, Rev. Mod. Phys. 68 (1996) 13–125, doi:10.1103/RevModPhys.68.13.

[135] U. Wolff, Collective Monte Carlo updating for spin systems, Phys. Rev. Lett. 62 (1989) 361–364, doi:10.1103/PhysRevLett.62.361.

[136] M. Neilsen, Neural Networks and Deep Learning, http://neuralnetworksanddeeplearning.com/.

[137] 3Blue1Brown Youtube Channel, But what is a neural network? | Chapter 1, Deep learning, https://www.youtube.com/watch?v=aircAruvnKk.

[138] L. van der Maaten, G. Hinton, Visualizing data using t-SNE, J. Mach. Learn. Res. 9 (86) (2008) 2579–2605.

[139] A tutorial on how to use t-SNE effectively can be found at https://distill.pub/2016/misread-tsne.

[140] F. Moosmann, B. Triggs, F. Jurie, Fast discriminative visual codebooks using randomized clustering forests, in: B. Schölkopf, J. Platt, T. Hoffman (Eds.), Advances in Neural Information Processing Systems, vol. 19, MIT Press, 2006.

[141] For a code, examples, and more references see https://scikit-learn.org/stable/modules/ensemble.html#random-trees-embedding.

[142] J. Carrasquilla, R.G. Melko, Machine learning phases of matter, Nat. Phys. 13 (5) (2017) 431–434.

[143] L. Wang, Discovering phase transitions with unsupervised learning, Phys. Rev. B 94 (2016) 195105, doi:10.1103/PhysRevB.94.195105.

[144] G. Carleo, M. Troyer, Solving the quantum many-body problem with artificial neural networks, Science 355 (6325) (2017) 602–606, doi:10.1126/science.aag2302.

[145] G. Torlai, R.G. Melko, Learning thermodynamics with Boltzmann machines, Phys. Rev. B 94 (2016) 165134, doi:10.1103/PhysRevB.94.165134.

[146] W. Hu, R.R.P. Singh, R.T. Scalettar, Discovering phases, phase transitions, and crossovers through unsupervised machine learning: a critical examination, Phys. Rev. E 95 (2017) 062122, doi:10.1103/PhysRevE.95.062122.

[147] C. Wang, H. Zhai, Machine learning of frustrated classical spin models. I. Principal component analysis, Phys. Rev. B 96 (2017) 144432, doi:10.1103/PhysRevB.96.144432.

[148] S.J. Wetzel, Unsupervised learning of phase transitions: from principal component analysis to variational autoencoders, Phys. Rev. E 96 (2) (2017) 022140.

[149] A. Tanaka, A. Tomiya, Detection of phase transition via convolutional neural networks, J. Phys. Soc. Jpn. 86 (6) (2017) 063001.

[150] K. Ch'ng, N. Vazquez, E. Khatami, Unsupervised machine learning account of magnetic transitions in the Hubbard model, Phys. Rev. E 97 (2018) 013306, doi:10.1103/PhysRevE.97.013306.

[151] K. Binder, E. Luijten, Monte carlo tests of renormalization-group predictions for critical phenomena in Ising models, Phys. Rep. 344 (4–6) (2001) 179–253.

[152] M. Blume, V.J. Emery, R.B. Griffiths, Ising model for the λ transition and phase separation in He^3-He^4 mixtures, Phys. Rev. A 4 (1971) 1071–1077, doi:10.1103/PhysRevA.4.1071.

[153] C. Wang, H. Zhai, Machine learning of frustrated classical spin models (II): kernel principal component analysis, Front. Phys. 13 (5) (2018) 1–7.

[154] M.J.S. Beach, A. Golubeva, R.G. Melko, Machine learning vortices at the Kosterlitz–Thouless transition, Phys. Rev. B 97 (2018) 045207, doi:10.1103/PhysRevB.97.045207.

[155] J.F. Rodriguez-Nieva, M.S. Scheurer, Identifying topological order through unsupervised machine learning, Nat. Phys. 15 (8) (2019) 790–795.

[156] For an incomplete early list, see "Applying Machine Learning to Physics" in Dr. Anna Golubeva's website at https://github.com/AnnaGolubeva/physicsml.github.io/blob/master/develop/content/pages/papers.md.

[157] P. Broecker, J. Carrasquilla, R.G. Melko, S. Trebst, Machine learning quantum phases of matter beyond the fermion sign problem, Sci. Rep. 7 (1) (2017) 8823, doi:10.1038/s41598-017-09098-0.

[158] K. Ch'ng, J. Carrasquilla, R.G. Melko, E. Khatami, Machine learning phases of strongly correlated fermions, Phys. Rev. X 7 (2017) 031038, doi:10.1103/PhysRevX.7.031038.

[159] D.-L. Deng, X. Li, S. Das Sarma, Machine learning topological states, Phys. Rev. B 96 (2017) 195145, doi:10.1103/PhysRevB.96.195145.

[160] G. Torlai, R.G. Melko, Neural decoder for topological codes, Phys. Rev. Lett. 119 (2017) 030501, doi:10.1103/PhysRevLett.119.030501.

[161] Y. Zhang, E.-A. Kim, Quantum loop topography for machine learning, Phys. Rev. Lett. 118 (2017) 216401, doi:10.1103/PhysRevLett.118.216401.

[162] E.P.L. van Nieuwenburg, Y.-H. Liu, S.D. Huber, Learning phase transitions by confusion, Nat. Phys. 13 (5) (2017) 435–439, doi:10.1038/nphys4037.

[163] P. Broecker, F.F. Assaad, S. Trebst, Quantum phase recognition via unsupervised machine learning, arXiv:1707.00663(2017).

[164] A. Tirelli, N.C. Costa, Learning quantum phase transitions through topological data analysis, Phys. Rev. B 104 (2021) 235146, doi:10.1103/PhysRevB.104.235146.

[165] R.A. Fontenele, N.C. Costa, R.R. dos Santos, T. Paiva, Two-dimensional attractive Hubbard model and the BCS-BEC crossover, Phys. Rev. B 105 (2022) 184502, doi:10.1103/PhysRevB.105.184502.

[166] N.C. Costa, W. Hu, Z.J. Bai, R.T. Scalettar, R.R.P. Singh, Principal component analysis for fermionic critical points, Phys. Rev. B 96 (2017) 195138, doi:10.1103/PhysRevB.96.195138.

[167] J.E. Hirsch, Two-dimensional Hubbard model: numerical simulation study, Phys. Rev. B 31 (1985) 4403–4419, doi:10.1103/PhysRevB.31.4403.

[168] S. Sorella, E. Tosatti, Semi-metal-insulator transition of the Hubbard model in the honeycomb lattice, Europhys. Lett. (EPL) 19 (8) (1992) 699–704, doi:10.1209/0295-5075/19/8/007.

[169] T. Paiva, R.T. Scalettar, W. Zheng, R.R.P. Singh, J. Oitmaa, Ground-state and finite-temperature signatures of quantum phase transitions in the half-filled Hubbard model on a honeycomb lattice, Phys. Rev. B 72 (2005) 085123, doi:10.1103/PhysRevB.72.085123.

[170] Y. Otsuka, S. Yunoki, S. Sorella, Universal quantum criticality in the metal-insulator transition of two-dimensional interacting Dirac electrons, Phys. Rev. X 6 (2016) 011029, doi:10.1103/PhysRevX.6.011029.

[171] M. Raczkowski, R. Peters, T.T. Phùng, N. Takemori, F.F. Assaad, A. Honecker, J. Vahedi, Hubbard model on the honeycomb lattice: from static and dynamical mean-field theories to lattice quantum Monte Carlo simulations, Phys. Rev. B 101 (12) (2020) 125103.

[172] Z. Meng, S. Wessel, A. Muramatsu, T.C. Lang, F.F. Assaad, Quantum spin liquid emerging in two-dimensional correlated Dirac fermions, Nature 464 (2010) 847–851, doi:10.1038/nature08942.

[173] S. Sorella, Y. Otsuka, S. Yunoki, Absence of a spin liquid phase in the Hubbard model on the honeycomb lattice, Sci. Rep. 2 (2012) 992, doi:10.1038/srep00992.

[174] G. Dopf, A. Muramatsu, W. Hanke, Three-band Hubbard model: a Monte Carlo study, Phys. Rev. B 41 (1990) 9264–9275, doi:10.1103/PhysRevB.41.9264.

[175] Y.F. Kung, C.-C. Chen, Y. Wang, E.W. Huang, E.A. Nowadnick, B. Moritz, R.T. Scalettar, S. Johnston, T.P. Devereaux, Characterizing the three-orbital Hubbard model with determinant quantum Monte Carlo, Phys. Rev. B 93 (2016) 155166, doi:10.1103/PhysRevB.93.155166.

[176] T.-C. Yi, R.T. Scalettar, R. Mondaini, Hamming distance and the onset of quantum criticality, arXiv:2111.12936(2021).

[177] A. Moreo, D.J. Scalapino, Two-dimensional negative-U Hubbard model, Phys. Rev. Lett. 66 (1991) 946–948, doi:10.1103/PhysRevLett.66.946.

[178] D.J. Scalapino, S.R. White, S. Zhang, Insulator, metal, or superconductor: the criteria, Phys. Rev. B 47 (13) (1993) 7995.

[179] J.M. Singer, M.H. Pedersen, T. Schneider, H. Beck, H.-G. Matuttis, From BCS–like superconductivity to condensation of local pairs: a numerical study of the attractive Hubbard model, Phys. Rev. B 54 (2) (1996) 1286.

[180] B. Kyung, S. Allen, A.-M. Tremblay, Pairing fluctuations and pseudogaps in the attractive Hubbard model, Phys. Rev. B 64 (7) (2001) 075116.

[181] T. Paiva, R.R. Dos Santos, R.T. Scalettar, P. Denteneer, Critical temperature for the two-dimensional attractive Hubbard model, Phys. Rev. B 69 (18) (2004) 184501.

[182] S. Karakuzu, K. Seki, S. Sorella, Study of the superconducting order parameter in the two-dimensional negative-U Hubbard model by grand-canonical twist-averaged boundary conditions, Phys. Rev. B 98 (7) (2018) 075156.

[183] Y. Zhang, A. Mesaros, K. Fujita, S.D. Edkins, M.H. Hamidian, K. Ch'ng, H. Eisaki, S. Uchida, J.C.S. Davis, E. Khatami, E.-A. Kim, Machine learning in electronic-quantum-matter imaging experiments, Nature 570 (7762) (2019) 484–490, doi:10.1038/s41586-019-1319-8.

[184] E. Khatami, E. Guardado-Sanchez, B.M. Spar, J.F. Carrasquilla, W.S. Bakr, R.T. Scalettar, Visualizing strange metallic correlations in the two-dimensional Fermi–Hubbard model with artificial intelligence, Phys. Rev. A 102 (2020) 033326, doi:10.1103/PhysRevA.102.033326.

[185] A. Bohrdt, C.S. Chiu, G. Ji, M. Xu, D. Greif, M. Greiner, E. Demler, F. Grusdt, M. Knap, Classifying snapshots of the doped Hubbard model with machine learning, Nat. Phys. 15 (9) (2019) 921–924, doi:10.1038/s41567-019-0565-x.

[186] L.W. Cheuk, M.A. Nichols, K.R. Lawrence, M. Okan, H. Zhang, E. Khatami, N. Trivedi, T. Paiva, M. Rigol, M.W. Zwierlein, Observation of spatial charge and spin correlations in the 2D Fermi–Hubbard model, Science 353 (6305) (2016) 1260–1264, doi:10.1126/science.aag3349.

[187] P.T. Brown, D. Mitra, E. Guardado-Sanchez, R. Nourafkan, A. Reymbaut, C.-D. Hébert, S. Bergeron, A.-M.S. Tremblay, J. Kokalj, D.A. Huse, P. Schauß, W.S. Bakr, Bad metallic transport in a cold atom Fermi–Hubbard system, Science 363 (6425) (2019) 379–382, doi:10.1126/science.aat4134.

[188] A.M. Samarakoon, D.A. Tennant, Machine learning for magnetic phase diagrams and inverse scattering problems, J. Phys. 34 (4) (2021) 044002, doi:10.1088/1361-648x/abe818.

[189] A.M. Samarakoon, K. Barros, Y.W. Li, M. Eisenbach, Q. Zhang, F. Ye, V. Sharma, Z.L. Dun, H. Zhou, S.A. Grigera, C.D. Batista, D.A. Tennant, Machine-learning-assisted insight into spin ice Dy$_2$Ti$_2$O$_7$, Nat. Commun. 11 (1) (2020) 892, doi:10.1038/s41467-020-14660-y.

[190] H. Zhang, C.D. Batista, Classical spin dynamics based on SU(N) coherent states, Phys. Rev. B 104 (2021) 104409, doi:10.1103/PhysRevB.104.104409.

[191] D. Dahlbom, H. Zhang, C. Miles, X. Bai, C.D. Batista, K. Barros, Geometric integration of classical spin dynamics via a mean-field Schrödinger equation, Phys. Rev. B 106 (2022) 054423, doi:10.1103/PhysRevB.106.054423.

[192] D. Dahlbom, C. Miles, H. Zhang, C.D. Batista, K. Barros, Langevin dynamics of generalized spins as SU(N) coherent states, arXiv:2209.01265(2022b).

[193] A.M. Samarakoon, P. Laurell, C. Balz, A. Banerjee, P. Lampen-Kelley, D. Mandrus, S.E. Nagler, S. Okamoto, D.A. Tennant, Extraction of interaction parameters for α-RuCl$_3$ from neutron data using machine learning, Phys. Rev. Res. 4 (2022) L022061, doi:10.1103/PhysRevResearch.4.L022061.

[194] N. Marzari, A.A. Mostofi, J.R. Yates, I. Souza, D. Vanderbilt, Maximally localized Wannier functions: theory and applications, Rev. Mod. Phys. 84 (2012) 1419–1475, doi:10.1103/RevModPhys.84.1419.

[195] F. Aryasetiawan, T. Miyake, R. Sakuma, The constrained RPA method for calculating the Hubbard U from first-principles, in: E. Pavarini, E. Koch, D. Voll-

hardt, A. Lichtenstein (Eds.), The LDA+DMFT Approach to Strongly Correlated Matter, Forschungszentrum Jülich, 2011.

[196] C. Eichstaedt, Y. Zhang, P. Laurell, S. Okamoto, A.G. Eguiluz, T. Berlijn, Deriving models for the Kitaev spin-liquid candidate material α-RuCl$_3$ from first principles, Phys. Rev. B 100 (2019) 075110, doi:10.1103/PhysRevB.100.075110.

[197] H.J. Kulik, T. Hammerschmidt, J. Schmidt, S. Botti, M.A.L. Marques, M. Boley, M. Scheffler, M. Todorović, P. Rinke, C. Oses, A. Smolyanyuk, S. Curtarolo, A. Tkatchenko, A.P. Bartók, S. Manzhos, M. Ihara, T. Carrington, J. Behler, O. Isayev, M. Veit, A. Grisafi, J. Nigam, M. Ceriotti, K.T. Schütt, J. Westermayr, M. Gastegger, R.J. Maurer, B. Kalita, K. Burke, R. Nagai, R. Akashi, O. Sugino, J. Hermann, F. Noé, S. Pilati, C. Draxl, M. Kuban, S. Rigamonti, M. Scheidgen, M. Esters, D. Hicks, C. Toher, P.V. Balachandran, I. Tamblyn, S. Whitelam, C. Bellinger, L.M. Ghiringhelli, Roadmap on machine learning in electronic structure, Electron. Struct. 4 (2) (2022) 023004, doi:10.1088/2516-1075/ac572f.

[198] F. Brockherde, L. Vogt, L. Li, M.E. Tuckerman, K. Burke, K.-R. Müller, Bypassing the Kohn–Sham equations with machine learning, Nat. Commun. 8 (1) (2017) 872, doi:10.1038/s41467-017-00839-3.

[199] A. Chandrasekaran, D. Kamal, R. Batra, C. Kim, L. Chen, R. Ramprasad, Solving the electronic structure problem with machine learning, npj Comput. Mater. 5 (1) (2019) 22, doi:10.1038/s41524-019-0162-7.

[200] Z. Wang, S. Ye, H. Wang, J. He, Q. Huang, S. Chang, Machine learning method for tight-binding hamiltonian parameterization from ab-initio band structure, npj Comput. Mater. 7 (1) (2021) 11, doi:10.1038/s41524-020-00490-5.

[201] M. Yu, S. Yang, C. Wu, N. Marom, Machine learning the Hubbard U parameter in DFT + U using Bayesian optimization, npj Comput. Mater. 6 (1) (2020) 180, doi:10.1038/s41524-020-00446-9.

[202] C. Schattauer, M. Todorović, K. Ghosh, P. Rinke, F. Libisch, Machine learning sparse tight-binding parameters for defects, npj Comput. Mater. 8 (1) (2022) 116, doi:10.1038/s41524-022-00791-x.

[203] J. Liu, H. Shen, Y. Qi, Z.Y. Meng, L. Fu, Self-learning Monte Carlo method and cumulative update in fermion systems, Phys. Rev. B 95 (2017) 241104, doi:10.1103/PhysRevB.95.241104.

[204] L. Huang, Y.-f. Yang, L. Wang, Recommender engine for continuous-time quantum Monte Carlo methods, Phys. Rev. E 95 (2017) 031301, doi:10.1103/PhysRevE.95.031301.

[205] Y. Nagai, H. Shen, Y. Qi, J. Liu, L. Fu, Self-learning Monte Carlo method: continuous-time algorithm, Phys. Rev. B 96 (2017) 161102, doi:10.1103/PhysRevB.96.161102.

[206] X.Y. Xu, Y. Qi, J. Liu, L. Fu, Z.Y. Meng, Self-learning quantum Monte Carlo method in interacting fermion systems, Phys. Rev. B 96 (2017) 041119, doi:10.1103/PhysRevB.96.041119.

[207] H. Shen, J. Liu, L. Fu, Self-learning Monte Carlo with deep neural networks, Phys. Rev. B 97 (2018) 205140, doi:10.1103/PhysRevB.97.205140.

[208] Z.H. Liu, X.Y. Xu, Y. Qi, K. Sun, Z.Y. Meng, Itinerant quantum critical point with frustration and a non-Fermi liquid, Phys. Rev. B 98 (2018) 045116, doi:10.1103/PhysRevB.98.045116.

[209] C. Chen, X.Y. Xu, J. Liu, G. Batrouni, R. Scalettar, Z.Y. Meng, Symmetry-enforced self-learning Monte Carlo method applied to the Holstein model, Phys. Rev. B 98 (2018) 041102, doi:10.1103/PhysRevB.98.041102.

[210] S. Li, P.M. Dee, E. Khatami, S. Johnston, Accelerating lattice quantum Monte Carlo simulations using artificial neural networks: application to the Holstein model, Phys. Rev. B 100 (2019) 020302, doi:10.1103/PhysRevB.100.020302.

[211] A. Tanaka, A. Tomiya, Towards reduction of autocorrelation in HMC by machine learning, arXiv:1712.03893(2017).

[212] H. Kohshiro, Y. Nagai, Effective Ruderman–Kittel–Kasuya–Yosida-like interaction in diluted double-exchange model: self-learning Monte Carlo approach, J. Phys. Soc. Jpn. 90 (3) (2021) 034711, doi:10.7566/JPSJ.90.034711.

[213] J.I. Monroe, V.K. Shen, Learning efficient, collective Monte Carlo moves with variational autoencoders, J. Chem. Theory Comput. 18 (6) (2022) 3622–3636, doi:10.1021/acs.jctc.2c00110.

[214] Y. Levine, O. Sharir, N. Cohen, A. Shashua, Quantum entanglement in deep learning architectures, Phys. Rev. Lett. 122 (2019) 065301, doi:10.1103/PhysRevLett.122.065301.

[215] O. Sharir, A. Shashua, G. Carleo, Neural tensor contractions and the expressive power of deep neural quantum states, 2021. 10.48550/ARXIV.2103.10293

[216] J. Chen, S. Cheng, H. Xie, L. Wang, T. Xiang, Equivalence of restricted Boltzmann machines and tensor network states, Phys. Rev. B 97 (2018) 085104, doi:10.1103/PhysRevB.97.085104.

[217] Y. Zheng, H. He, N. Regnault, B.A. Bernevig, Restricted Boltzmann machines and matrix product states of one-dimensional translationally invariant stabilizer codes, Phys. Rev. B 99 (2019) 155129, doi:10.1103/PhysRevB.99.155129.