Chebyshev expansion of spectral functions using restricted Boltzmann machines

Douglas Hendry, Hongwei Chen, Phillip Weinberg, and Adrian E. Feiguin
Department of Physics, Northeastern University, Boston, Massachusetts 02115, USA

(Received 14 May 2021; revised 13 September 2021; accepted 3 November 2021; published 24 November 2021)

Calculating the spectral function of two-dimensional systems is arguably one of the most pressing challenges in modern computational condensed matter physics. While efficient techniques are available in lower dimensions, two-dimensional systems present insurmountable hurdles, ranging from the sign problem in quantum Monte Carlo methods to the entanglement area law in tensor-network-based methods. We hereby present a variational approach based on a Chebyshev expansion of the spectral function and a neural network representation for the wave functions. The Chebyshev moments are obtained by recursively applying the Hamiltonian and projecting on the space of variational states using a modified natural gradient descent method. We compare this approach with a modified approximation of the spectral function which uses a Krylov subspace constructed from the "Chebyshev wave functions." We present results for the one-dimensional and two-dimensional Heisenberg model on the square lattice and compare them with those obtained by other methods in the literature.

DOI: 10.1103/PhysRevB.104.205130

I. INTRODUCTION

The knowledge of the excitation spectrum of a quantum many-body system provides valuable information that can be directly compared with experiments. While much effort is dedicated to the calculation of ground-state wave functions, their energies do not have much experimental relevance since they cannot be measured. Experiments such as photoemission or inelastic neutron scattering spectroscopies, for instance, measure energy differences between excited and ground states, yielding a spectral function, or density of states. In addition, the integrated weight also yields the equal time correlations in the ground state, which can also indicate the existence of order or lack thereof, or the presence of quasiparticle excitations that could be used as fingerprints to identify quantum phases. Unfortunately, the numerical evaluation of dynamical correlation functions in strongly correlated systems is a very difficult undertaking, and the development of techniques that can be applied beyond one dimension is possibly one of the pressing challenges in computational condensed matter physics.

Since the advent of high-temperature superconductivity, progress has been marked by ingenuity to overcome the computational limitations of the time. The first techniques to emerge were based on exact diagonalization [1], which is limited to small clusters, and different variants of the quantum Monte Carlo method, which suffers from the sign problem and requires uncontrolled analytic continuations [2–9]. Soon after the invention of the density matrix renormalization group (DMRG) [10–14], several approaches to extract dynamical properties were proposed, which can be generically referred to as dynamical DMRG [15–20], or DDMRG. The time-dependent density matrix renormalization group and recent variations using Chebyshev expansions have been important developments, giving access to accurate spectra for very large one-dimensional (1D) systems [21–28] and quasi-2D cylinders [29,30] with moderate computational resources. An alternative approach consists of proposing variational forms for the excited states, which can be represented with matrix product states (MPSs) [31–34] or other variational wave functions that can be easily extended to higher dimensions and are free from the sign problem [35–39]. The variational Monte Carlo (VMC) approach relies on a variational ansatz for the ground state inspired by some physical insight (typically of the form of a Gutzwiller projected wave function) and can provide a few hundred discrete poles by diagonalizing a Hamiltonian matrix projected onto a subspace of single magnon excitations.

In this paper, we bring together some of these ideas translated in the context of recent advances in quantum machine learning. Attempts in this direction have already demonstrated very promising results: In Ref. [39], the concepts behind correction-vector dynamical DMRG were generalized to arbitrary variational wave functions and applied to the Heisenberg model using restricted Boltzmann machines (RBMs) [40–43]. RBMs are a type of artificial neural network widely used in machine learning that inspired Carleo and Troyer [40] to propose a variational wave function for a spin-$\frac{1}{2}$ system of N sites. The visible layer corresponds to the spin configurations $\vec{\sigma}^z = (\sigma_1^z, \sigma_2^z, \ldots, \sigma_N^z)$. Then the coefficients of the wave function are given as $\psi(\vec{\sigma}^z, \vec{a}, \vec{b}, W) = e^{\sum_{i=1}^N a_i \sigma_i^z} \prod_{j=1}^M 2\cosh(\theta_j)$, where $\theta_j = b_j + \sum_{i=1}^N W_{ij}\sigma_i^z$. The weights W_{ij} and the biases A and b in this expression are free parameters that are used to variationally minimize the energy $E = \langle\psi|H\psi\rangle/\langle\psi|\psi\rangle$. This procedure is carried out using the Monte Carlo method to sample over all possible spin configurations. The remarkable aspect of this type of wave function is that they can encode a great deal of information in a relatively small set of variational parameters, without making any assumptions about the physics of the problem [44].

The main goal of this paper is to variationally approximate the zero-temperature Green's function for a quantum many-

body system:

$$G_{ij}(z) = \langle\psi|\hat{A}_j^\dagger \frac{1}{z-\hat{H}}\hat{A}_i|\psi\rangle,$$

where A_i and A_j are some local operators of interest, E_0 is the ground-state energy, and $z = \omega + E_0 + i\eta$. The spectral function is thus obtained as $A_{ij}(\omega) = -\frac{1}{\pi}G_{ij}(z)$. The correction-vector DMRG method recasts this calculation as a complex system of equations for each frequency ω. In Ref. [39], the solution to this system of equations was encoded in the form of an RBM. The optimization is carried out by means of a natural gradient descent approach based on quantum geometry concepts that allows one to solve a large system of equations stochastically with RBMs, or any other variational wave function.

We here explore a different avenue by expanding the spectral function in terms of Chebyshev polynomials. In Ref. [26] it was shown that this approach, combined with DMRG, can yield very accurate results with a fraction of the effort compared with other methods. In fact, it comes with several advantages over other methods, such as that (i) it provides uniform resolution over the entire relevant frequency range with a relatively small number of Chebyshev moments and (ii) the information to calculate each moment is encoded in an RBM that can be stored for later use or systematically improved if needed.

In addition, we propose two alternative ideas to improve on the basic algorithm. The main observation that we make is that variational wave functions in general can have limited ability to accurately represent arbitrary many-body states, translating into systematic errors in the calculated quantities. This can be due to either the mathematical structure of the variational state or numerical errors in their optimization. Therefore we use a new set of basis states, constructed as linear combinations of RBM wave functions, that allows us to explore regions of the Hilbert space outside of the RBM manifold. This approach yields new, "corrected" results that considerably improve the accuracy of the expansion. Finally, the Hamiltonian, projected on this new basis, can be tridiagonalized to produce yet another estimate of the spectral function in terms of a continued fraction. The latter can be directly compared with other methods, such as dynamical DMRG.

The paper is organized as follows: In Sec. II we introduce the basic ideas behind the Chebyshev representation of the spectral function; in Sec. III we discuss the implementation using RBMs and the details of the variational optimization. Results can be improved by taking advantage of error correction measures described in Sec. IV. Section V covers an alternative approach using a representation of the spectral function in terms of a Krylov expansion of the basis. Results for both the one-dimensional and two-dimensional Heisenberg model on the square lattice are presented in Sec. VI. We finally close with a summary and discussion of possible strategies to improve the accuracy of the calculation.

II. CHEBYSHEV EXPANSION

In this section we briefly describe the expansion of the spectral function $A_{ij}(\omega)$ in terms of Chebyshev moments. We refer the reader to Ref. [26] for a more detailed discussion. An important aspect of Chebyshev polynomials is they are defined in the interval $[-1, 1]$. Hence we rescale the frequencies and the Hamiltonian such that the range of the expansion encompasses the region with the bulk of the spectral weight. If our problem has nonzero spectral weight in the interval $0 \leqslant \omega \leqslant W_A$, we define $\omega' = \omega/a - W'$ with $a = W_*/2W'$, and $H' = (H - E_0)/a - W'$. In these expressions, $W_* > W_A$ is an effective bandwidth larger than W_A, and W' is the rescaled bandwidth, which we chose, for safety reasons, to be slightly smaller than 1 ($W' = 1 - 0.0125$ in the following). In a nutshell, the spectral function can now be written as

$$A_{ij}(\omega) = \frac{2W'/W_*}{\pi\sqrt{1-\omega'^2}}\left[g_0\mu_0 + 2\sum_{n=1}^{N-1} g_n\mu_n T_n(\omega')\right], \quad (1)$$

where $T_n(x)$ is the nth Chebyshev polynomial of the first kind, the coefficients g_n are damping factors that affect the broadening, and μ_n are the corresponding Chebyshev moments:

$$\mu_n = \langle\psi|\hat{A}_j^\dagger T_n(\hat{H}')\hat{A}_i|\psi\rangle. \quad (2)$$

Luckily, one can exploit the definition of the Chebyshev polynomials and their recursion relation to recast the moments as

$$\mu_n = \langle\psi|\hat{A}_j^\dagger|t_n\rangle, \quad |t_n\rangle = T_n(\hat{H}')\hat{A}_i|\psi\rangle, \quad (3)$$

with

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x),$$
$$T_0(x) = 1, T_1(x) = x. \quad (4)$$

Hence, given an initial wave function $|t_0\rangle = \hat{A}_i|\psi\rangle$, we iteratively solve for a sequence of "Chebyshev wave functions" $|t_1\rangle, |t_2\rangle, \ldots$ such that $|t_n\rangle = T_n(\hat{H}')|t_0\rangle$. In the following section, we describe how to carry out this task using a variational representation of $|t_n\rangle$.

III. VARIATIONAL OPTIMIZATION

Let $|\phi\rangle$ be the target wave function for a holomorphic variational wave function $|\psi(\alpha)\rangle$ with variational parameters $\{\alpha_1, \alpha_2, \ldots, \alpha_M\} \in \mathbb{C}^M$ that need to be optimized to make $|\psi(\alpha)\rangle$ as equal to $|\phi\rangle$ as possible up to an overall constant. This is done by minimizing the Fubini-Study metric—essentially the angle—between the two wave functions, which is given by

$$\gamma(\psi, \phi) = \arccos\sqrt{\frac{\langle\psi|\phi\rangle\langle\phi|\psi\rangle}{\langle\psi|\psi\rangle\langle\phi|\phi\rangle}}.$$

Taking the gradient of γ^2 with respect to $|\psi\rangle$ gives

$$|\delta\psi\rangle = \delta\gamma \cdot \gamma(\psi, \phi)\cot[\gamma(\psi, \phi)]\left[\frac{\langle\psi|\psi\rangle}{\langle\psi|\phi\rangle}|\phi\rangle - |\psi\rangle\right],$$

which is a differential rotation by an angle $\delta\gamma \cdot \gamma$ towards $|\phi\rangle$ (with a change in the overall phase to match $|\psi\rangle$). As with rotations in real space, $\langle\psi|\delta\psi\rangle = 0$ and $\langle\delta\psi|\delta\psi\rangle = (\delta\gamma \cdot \gamma)^2\langle\psi|\psi\rangle$.

The procedure to update the variational coefficients α using natural gradient descent is described in detail in Ref. [39], and we hereby summarize it. The updates $\Delta\alpha$ correspond to projecting $|\delta\psi\rangle$ onto the manifold of possible variational wave functions of the proposed form (in our case, an RBM). To

calculate $\Delta\alpha$, one must solve the following linear system of equations:

$$\sum_{k'=1}^{M} S_{kk'}\Delta\alpha_{k'} = \frac{\langle\partial_k\psi|\delta\psi\rangle}{\langle\psi|\psi\rangle} - \frac{\langle\partial_k\psi|\psi\rangle}{\langle\psi|\psi\rangle}\frac{\langle\psi|\delta\psi\rangle}{\langle\psi|\psi\rangle},$$

where $|\partial_k\psi\rangle = \frac{\partial}{\partial\alpha_k}|\psi\rangle$ and S is given by

$$S_{kk'} = \frac{\langle\partial_k\psi|\partial_{k'}\psi\rangle}{\langle\psi|\psi\rangle} - \frac{\langle\partial_k\psi|\psi\rangle}{\langle\psi|\psi\rangle}\frac{\langle\psi|\partial_{k'}\psi\rangle}{\langle\psi|\psi\rangle}.$$

To estimate the the overlap between wave functions in the definition of S and in the linear system defined above, we use the VMC approach. The states $|s\rangle$ are sampled according $P(s) = |\psi_s|^2/\langle\psi|\psi\rangle$, where $\psi_s \equiv \langle s|\psi\rangle$. Using these sampled states, we define estimators such as the log-derivative $\mathcal{O}_k(s) = \frac{\partial\psi_s}{\partial\alpha_k}/\psi_s$ and the ratio $R(s) = \phi_s/\psi_s$ that are used to calculate the following relevant quantities:

$$\begin{aligned} S_{kk'} &= \langle\mathcal{O}_k(s)^*\mathcal{O}_{k'}(s)\rangle - \langle\mathcal{O}_k(s)\rangle^*\langle\mathcal{O}_{k'}(s)\rangle, \\ \gamma(\psi,\phi) &= \arccos\sqrt{\frac{|\langle R(s)\rangle|^2}{\langle|R(s)|^2\rangle}}, \\ \frac{\langle\partial_k\psi|\delta\psi\rangle}{\langle\psi|\psi\rangle} &= \delta\gamma\,\gamma(\psi,\phi)\cot[\gamma(\psi,\phi)] \\ &\quad\times\left[\frac{\langle\mathcal{O}_k(s)^*R(s)\rangle}{\langle R(s)\rangle} - \langle\mathcal{O}_k(s)^*\rangle\right]. \end{aligned} \tag{5}$$

Notice that when $|\psi\rangle \propto |\phi\rangle$, the variance of $R(s)$ goes to 0. As a consequence, the variance of the estimators above (except S) also goes to zero. Once the parameters have converged, the normalization constant β can be calculated as

$$\beta = \frac{\langle\psi|\phi\rangle}{\langle\psi|\psi\rangle} = \langle R(s)\rangle.$$

This optimization algorithm is applied to every step of the Chebyshev recursion: Each wave function in sequence is approximated by a non-normalized variational wave function $|\tilde{t}_n\rangle$ and corresponding normalization constant β_n. The wave functions are solved iteratively using the Chebyshev recurrence relation and are projected onto an RBM form such that the corresponding target wave functions are $|t_1\rangle = \beta_0 H'|\tilde{t}_0\rangle$ and $|t_n\rangle = 2\beta_{n-1}H'|\tilde{t}_{n-1}\rangle - \beta_{n-2}|\tilde{t}_{n-2}\rangle$ for $n \geqslant 2$, where $\beta_0 = 1$ and $|\tilde{t}_0\rangle = |t_0\rangle$. These wave functions can be easily stored, since the number of variational coefficients is $O(N^2)$. This allows the calculation to proceed in three steps: First, the wave functions $|t_n\rangle$ are obtained; second, the Chebyshev moments are calculated by sampling over the wave functions to obtain the overlaps, and, finally, the spectral function can be reconstructed.

IV. CHEBYSHEV MOMENT CORRECTION

In the absence of errors, only the overlap of the Chebyshev vectors with $|t_0\rangle$ is needed to obtain the Chebyshev moments. Due to numerical errors, limitations in the expressivity of the RBM wave function, and the fact that each successive Chebyshev wave function is obtained recursively, the variational Chebyshev wave functions will become a less accurate approximation, amplifying the error for the higher moments. However, we can obtain better estimates for the Chebyshev wave functions by generating a new orthogonal basis from the original Chebyshev wave functions. The orthogonalization requires calculating the overlap matrix between the RBM states, as we describe below. The new basis will be spanned by states that are linear combinations of RBMs, and are no longer RBMs. Once we have obtained the new basis, we recalculate the matrix elements of the Hamiltonian and use the new representation to obtain the corrected moments. Thus the resulting error no longer depends on the accuracy of the individual wave functions: Even if the individual RBM wave functions drift from their target Chebyshev wave function, the new basis will correct for these errors.

TABLE I. Workflow describing the procedure to calculate the corrected Chebyshev moments.

Chebyshev moment correction	
1	Obtain $\|t_n\rangle$ variationally using Chebyshev recursion
2	Calculate overlaps $\sigma_{nm} = \langle t_n\|t_m\rangle$ and matrix elements $H_{nm} = \langle t_n\|H\|t_m\rangle$
3	Define projector $P = \sum_{nm}(\sigma^{-1})_{nm}\|t_n\rangle\langle t_m\|$
4	Project $H \to H_P = P^\dagger HP$
5	Recalculate moments in the new basis using H_P, Eqs. (6)–(8)

The procedure to recalculate the Chebyshev moments is as follows (Table I): The projector onto the space spanned by the RBMs is given by $\hat{P} = \sum_{mn}(\sigma^{-1})_{mn}|t_m\rangle\langle t_n|$, $\sigma_{nm} = \langle t_n|t_m\rangle$, and $H_{nm} = \langle t_n|\hat{H}|t_m\rangle$. Let us project the Hamiltonian onto this basis as $\hat{H}_P = \hat{P}\hat{H}\hat{P}$. Then, the recalculated Chebyshev wave functions are obtained as

$$|t'_n\rangle = T_n(\hat{H}_P)|t'_0\rangle = \sum_m c_{mn}|t_m\rangle,$$

with $|t'_0\rangle = |t_0\rangle$ and thus $c_{0n} = \delta_{0n}$. The coefficients c_{mn} can be obtained from the Chebyshev recursion relations, which results in the following system of equations:

$$\sum_l \sigma_{ml}c_{l1} = \sum_l H_{ml}c_{l0}, \tag{6}$$

$$\sum_l \sigma_{ml}c_{l,n+1} = \sum_l 2H_{ml}c_{ln} - \sigma_{ml}c_{l,n-1}. \tag{7}$$

Finally, the moments are obtained as

$$\mu_n = \sum_{ml} c^*_{m0}\sigma_{ml}c_{ln}. \tag{8}$$

V. CONTINUED FRACTION

In the previous section we obtained a representation of the Hamiltonian in an orthogonal basis, $\hat{H}_P$. We can then generate a Lanczos recursion to cast the Hamiltonian matrix in tridiagonal form. The Lanczos wave functions $|v_0\rangle, \ldots, |v_{N-1}\rangle$ are obtained from the projected Hamiltonian [1] $\hat{H}_P$ with $|v_0\rangle = |t_0\rangle$:

$$\begin{aligned} |v_1\rangle &= \hat{H}_P|v_0\rangle - \frac{\langle v_0|\hat{H}_P|v_0\rangle}{\langle v_0|v_0\rangle}|v_0\rangle, \\ \|v_{n+1}\rangle &= \hat{H}_P|v_n\rangle - a_n|v_n\rangle - b_n^2|v_{n-1}\rangle, \end{aligned} \tag{9}$$

with

$$a_n = \frac{\langle v_n|\hat{H}_P|v_n\rangle}{\langle v_n|v_n\rangle}, \quad b_n^2 = \frac{\langle v_n|v_n\rangle}{\langle v_{n-1}|v_{n-1}\rangle}. \tag{10}$$

Finally, to obtain the spectral function from H_P, we use a continued-fraction expansion of the spectral function projected onto this subspace [1,15,19]:

$$A_{ij}(\omega) = -\frac{1}{\pi}\mathrm{Im}\left[\cfrac{\langle\psi|\hat{A}_j\hat{A}_i|\psi\rangle}{z - a_0 - \cfrac{b_1^2}{z - a_1 - \cfrac{b_2^2}{z - a_2 - \cdots}}}\right], \tag{11}$$

with $z = \omega + E_0 + i\eta$, where η is a small number (we take $\eta = 0.1$ in our calculations) that introduces an artificial broadening of the poles. This approach resembles the one proposed by Hallberg in Ref. [15] to calculate dynamical properties using DMRG, where the variational manifold is spanned by wave functions of the form of matrix product states instead of RBMs. This approach is simpler than the correction-vector method presented in Refs. [16,17,39] but, as we shall see below, may lose accuracy at high energies [19].

VI. RESULTS

For demonstration and benchmarking we focus on the spin-$\frac{1}{2}$ Heisenberg model in one and two dimensions on the square lattice:

$$\hat{H} = J\sum_{\langle ij\rangle}\vec{S}_i\cdot\vec{S}_j, \tag{12}$$

where $\vec{S} = (\hat{S}^x, \hat{S}^y, \hat{S}^z)$ are spin operators and the sum runs over pairs of nearest-neighbor sites on the lattice. We consider periodic boundary conditions and chose J as our unit of energy. We are interested in the longitudinal spin structure factor, defined as

$$S^z(\mathbf{k},\omega) = -\frac{1}{\pi N}\mathrm{Im}\sum_n e^{i\mathbf{k}\cdot\mathbf{r}_n}\langle\psi|\hat{S}_0^z\frac{1}{z-\hat{H}}\hat{S}_n^z|\psi\rangle,$$

where we have used translational invariance. In our calculations, unless otherwise stated, we consider $W' = 1 - \epsilon/4$ with $\epsilon = 0.05$; $W_* = 10W'$. In addition, we use the so-called "Jackson damping" in Eq. (1):

$$g_n = \frac{(N-n+1)\cos\frac{\pi n}{N+1} + \sin\frac{\pi n}{N+1}\cot\frac{\pi}{N+1}}{N+1},$$

where N is the number of moments used. We also implement spatial symmetries to improve the accuracy of our results, as described in Appendix A.

In Fig. 1 we show the evolution of the longitudinal structure factor of a 1D spin chain of length $L = 32$ for a choice of momentum $k = \pi/2$ as a function of the number of Chebyshev moments N. Unless otherwise stated, in all calculations we used L hidden variables. We clearly observe that the spectrum develops oscillations reproducing the well-resolved high-energy features in the finite system. These oscillations become more pronounced when increasing n over the range of energies corresponding to the spinon continuum. This shows that one can obtain very reasonable results over the entire

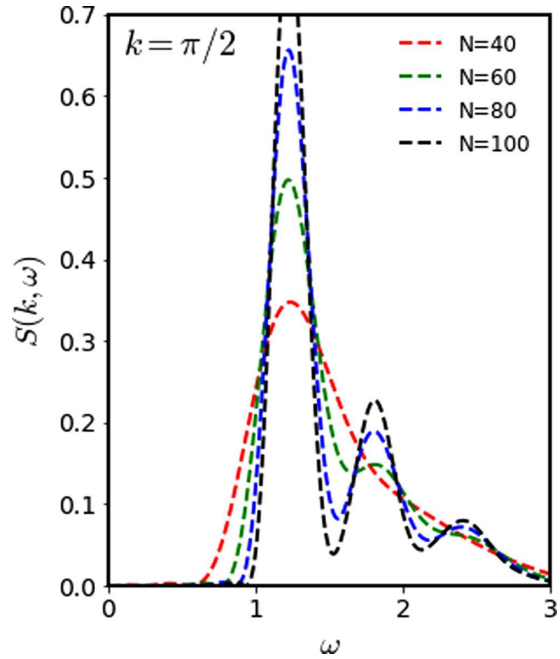

FIG. 1. Evolution of the spin dynamic structure factor of a Heisenberg chain with the number of Chebyshev moments using RBM wave functions for $L = 32$ at momentum $k = \pi/2$.

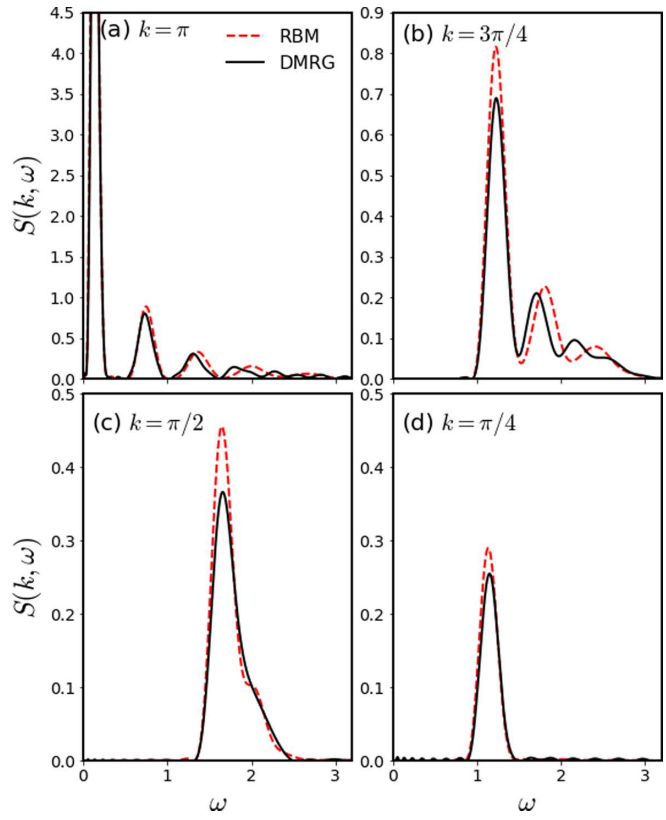

FIG. 2. (a)–(d) Spin dynamic structure factor for a Heisenberg chain obtained using $N = 100$ RBM wave functions, compared with DMRG results using a Chebyshev expansion for $L = 32$ and various values of momenta k.

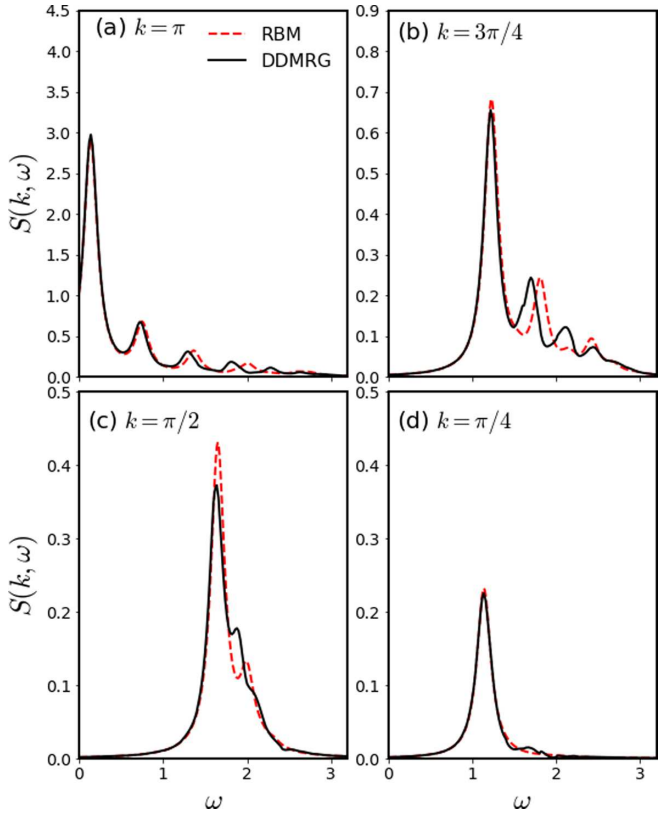

FIG. 3. (a)–(d) Spin dynamic structure factor for a Heisenberg chain obtained with a continued-fraction approach in the basis of $N = 100$ Chebyshev RBM wave functions, compared with dynamical (correction vector) DMRG, for $L = 32$ and various values of momenta k. An artificial broadening $\eta = 0.1$ is introduced.

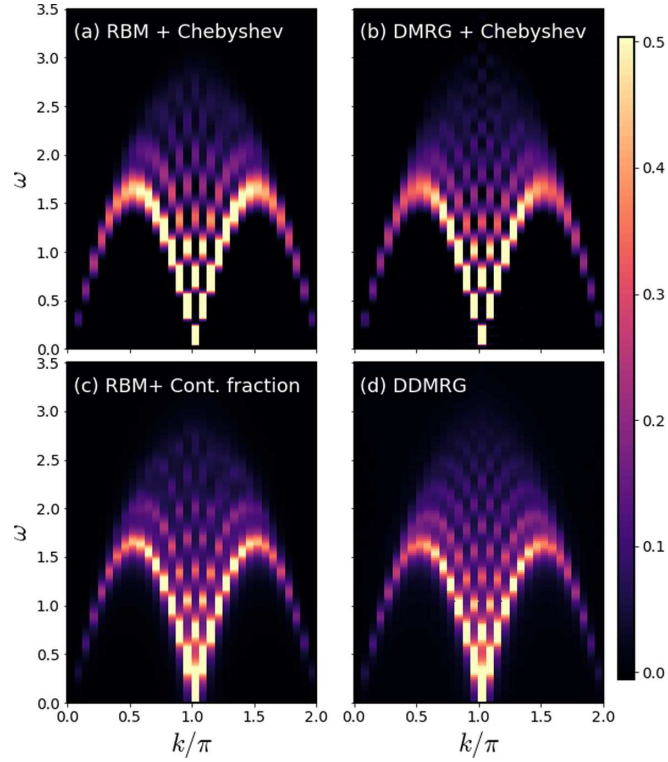

FIG. 4. Color density plots depicting the spin dynamic structure factor for a Heisenberg chain with $L = 32$ as a function of momentum and frequency, obtained by the four methods used in this paper: (a) Chebyshev moments with RBM wave functions, (b) Chebyshev moments using DMRG, (c) continued (Cont.) fraction using the Chebyshev RBM basis, and (d) dynamical DMRG. We use $N = 100$ Chebyshev moments.

frequency domain with moderate effort using a small number of wave functions, or Chebyshev iterations.

In Fig. 2 we compare results with $N = 100$ with those obtained using the same approach but with DMRG as a solver (i.e., using matrix product states as a variational ansatz). The agreement over the entire range of momenta and frequencies is remarkable. Notice that there is no straightforward way to quantify the error, except by comparing with such accurate calculations using other methods (DMRG in this case). We discuss the possible sources of the discrepancies in the discussion, Sec. VII.

Figure 3 shows the same spectra but using the continued-fraction approach, compared with dynamical DMRG. We note that while the positions of the peaks are the same, the profiles are different from those in Fig. 2 because the spectrum in this case is described by a superposition of Lorentzians. For all practical purposes, both techniques reproduce the same data. This is more dramatically illustrated in Fig. 4, where we compare all approaches as a function of momentum and frequency in color density maps. It is fair to say that they are practically indistinguishable.

While results in 1D serve as a benchmark and illustrate the accuracy of the approach, we aim at making this method also applicable in higher dimensions, where one can access richer physics, such as quantum spin liquids [45]. In this case, we focus on a 6×6 square lattice, where our variational results can be contrasted directly to exact diagonalization (see Appendix B). In Figs. 5 and 6 we display a comparison for several values of momenta on the two-dimensional Brillouin zone using both a Chebyshev expansion and a continued-fraction approach, respectively. We also include results with $2L$ hidden variables, showing very small differences indicating that the error is dominated by the Chebyshev truncation (the number of moments), as shown in Fig. 1. We find that while the first peak is very accurately described, the high-energy features with smaller spectral weight are partially lost, as also observed in the original dynamical DMRG using continued fractions [15].

VII. SUMMARY AND CONCLUSIONS

We have presented a highly efficient approach to calculate spectral functions of strongly correlated systems using a Chebyshev expansion of the Green's function and a variational representation of the many-body states in the form of restricted Boltzmann machines. Unlike a previous method introduced in the literature by the authors [39], these calculations are numerically inexpensive and yield very accurate results in 1D. We have also described remedial measures to improve the accuracy of the results, such as using spatial symmetries and correcting for the limited "expressivity" of the neural network. As mentioned earlier, the method provides

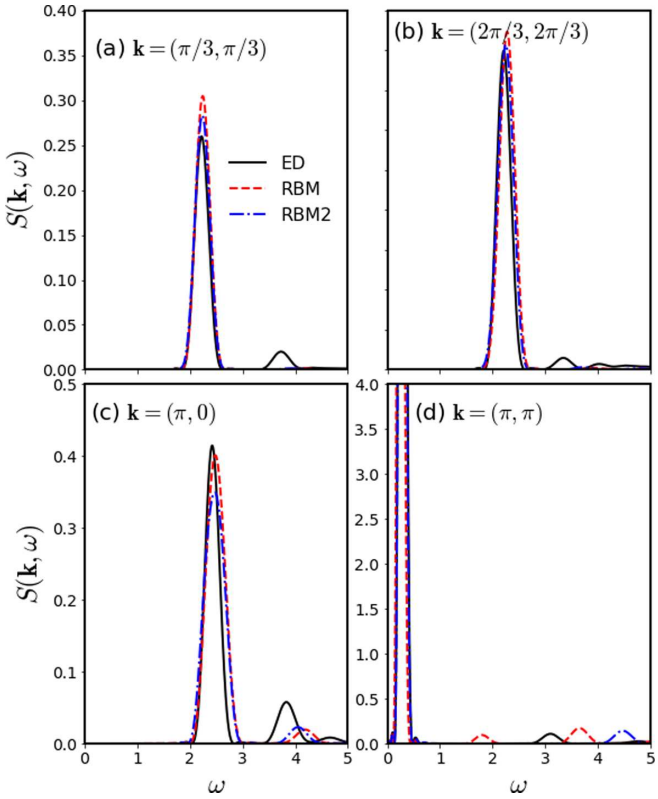

FIG. 5. (a)–(d) Spin dynamic structure factor for the Heisenberg model on a $L_x \times L_y = 6 \times 6$ square lattice using a Chebyshev expansion and exact diagonalization (ED). The curves with legends RBM and RBM2 correspond to a hidden layer with L and $2L$ variables, respectively.

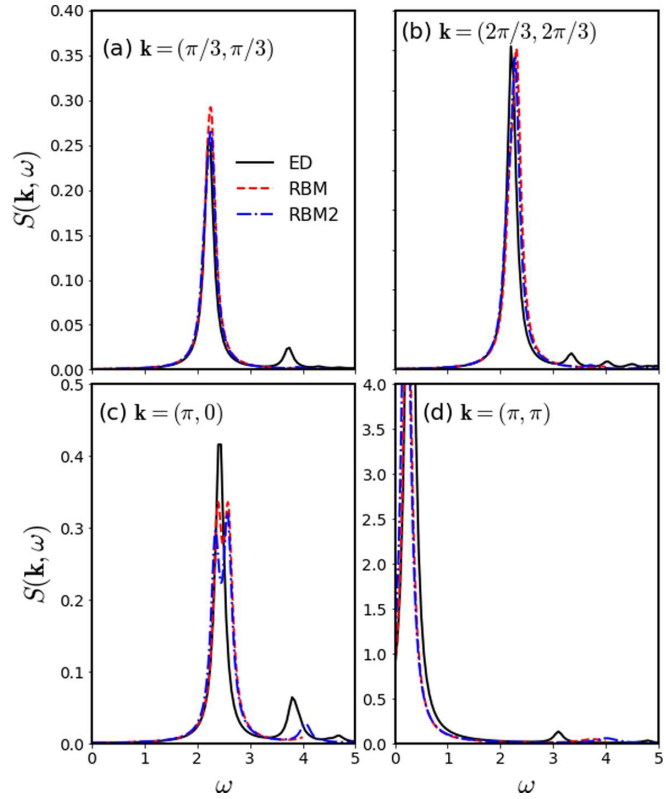

FIG. 6. (a)–(d) Spin dynamic structure factor for the Heisenberg model on a $L_x \times L_y = 6 \times 6$ square lattice using both continued fraction with RBM Chebyshev wave functions and exact diagonalization. The curves with legends RBM and RBM2 correspond to a hidden layer with L and $2L$ variables, respectively.

uniform resolution over the entire relevant frequency range with a relatively small number of Chebyshev moments, and the information to calculate each moment is encoded in an RBM that can be stored for postprocessing.

In order to understand the sources of error, we start by recalling that (i) our calculations rely on a particular form of the variational wave function, meaning that the results will depend on the representation power of an RBM determined primarily by the number of hidden variables [46]; (ii) the optimization of the wave function depends on our ability to accurately minimize our loss function with respect to the variational parameters (this loss function may have a complex landscape with local minima); and, in addition, (iii) we need to consider the errors introduced by the stochastic sampling of all the different quantities. Due to the recursive nature of the calculation, since every wave function depends on the previous ones, any errors are propagated down the sequence. We have shown that these errors can be mitigated by recalculating the moments in a Chebyshev basis that has support outside the space of single RBMs.

These ideas are general, as demonstrated by both prior studies using matrix product states and in this paper using RBMs. The main challenge moving forward is to scale these methods to larger systems, which will be primarily determined by our ability to train larger networks, or variational wave functions, with a growing number of free parameters.

It is also worth noting that one can take existing moments and approximate higher moments using extrapolation techniques [27,47]. This can help push calculations to larger system sizes by reducing the total number of moments required.

We finally point out that the overall technique, including the error correction approach we used to improve the Chebyshev wave functions, is generic and can be implemented with any variational form of wave function.

ACKNOWLEDGMENTS

A.E.F. and H.C. acknowledge the National Science Foundation for support under Grant No. DMR-1807814. D.H. is supported by a Northeastern Tier 1 grant.

APPENDIX A: IMPLEMENTING SYMMETRIES

The RBM wave functions are optimized to approximate $|t_{n00}\rangle = T_n(\hat{H})S^z_{00}|\psi_{gs}\rangle$ for the nth Chebyshev polynomial T_n and approximate ground state $|\psi_{gs}\rangle$. To obtain the Chebyshev wave functions for momentum $\mathbf{k} = (k_1, k_2)$, we make use of the translational symmetry of the problem by assuming that the ground state has zero momentum (translations do not result in a change in overall phase). Here, we use $\hat{T}_{m_1 m_2}$ to represent the translation operator on the square lattice with a shift of (m_1, m_2), while $T_n(\hat{H})$ denotes the Chebyshev polynomials

of the Hamiltonian. Using this notation, we can write

$$|\widetilde{t}_{nk_1k_2}\rangle = \frac{1}{L}T_n(\hat{H})\sum_{m_1,m_2=0}^{L-1} e^{-i\frac{2\pi}{L}(k_1m_1+k_2m_2)}\hat{T}_{m_1m_2}\hat{S}^z_{00}\hat{T}_{-m_1-m_2}|\psi_{gs}\rangle$$
$$= L\hat{P}^{(T)}_{k_1k_2}|t_{n00}\rangle.$$

In this equation, one can see that we are acting with the projection operator onto the momentum sector with momentum quantum numbers given by integers k_1, k_2 defined as

$$P^{(T)}_{k_1k_2} = \frac{1}{L^2}\sum_{m_1,m_2=0}^{L-1} e^{-i\frac{2\pi}{L}(k_1m_1+k_2m_2)}\hat{T}_{m_1m_2}.$$

Thus each momentum wave function can represented in terms of linear combinations of translations of a single RBM optimized for the local S^z_{00} operator. Additionally, since $\hat{P}^{(T)}_{k_1k_2}$ is a projector, the resulting state will have the correct momentum, and all the error from different momentum sectors will be projected out. In the same vein, rotational symmetry can be enforced by applying the following projector of rotations and translations.

$$\hat{P}^{(R)} = \tfrac{1}{4}(1 + \hat{T}_{10}\hat{R} + \hat{T}_{11}\hat{R}^2 + \hat{T}_{01}\hat{R}^3).$$

This projector is an average of each rotation followed by the translation which takes the S^z operator back to the origin or, equivalently, rotations about the origin. In the exact calculation for the Chebyshev wave functions, the application of $\hat{P}^{(R)}$ should leave the $|t_{n00}\rangle$ unchanged. However, the RBM representation will not necessarily have rotational symmetry; therefore the projector will correct any errors related to an absence of rotational symmetry. After including the rotation projector we find

$$|\widetilde{t}_{nk_1k_2}\rangle = L\hat{P}^{(T)}_{k_1k_2}\hat{P}^{(R)}|t_{n00}\rangle.$$

All that is needed to calculate the Chebyshev moments is the overlap:

$$\mu_{nk_1k_2} = \frac{\langle\widetilde{t}_{0k_1k_2}|\widetilde{t}_{nk_1k_2}\rangle}{\langle\psi_{gs}|\psi_{gs}\rangle}.$$

However, due to the recursive process of obtaining the Chebyshev wave functions, the error will propagate down the sequence of wave functions. As mentioned in the main text, we can correct some of this error by constructing new Chebyshev wave functions as linear combinations of the original RBM wave functions. This error correction requires us to calculate the following quantities:

$$G^{(k_1k_2)}_{nn'} = \frac{\langle\widetilde{t}_{nk_1k_2}|\widetilde{t}_{n'k_1k_2}\rangle}{\langle\psi_{gs}|\psi_{gs}\rangle},$$
$$H^{(k_1k_2)}_{nn'} = \frac{\langle\widetilde{t}_{nk_1k_2}|H|\widetilde{t}_{n'k_1k_2}\rangle}{\langle\psi_{gs}|\psi_{gs}\rangle}.$$

These matrices are calculated by sampling over spin configuration states $|s\rangle$ probability $P(s) = \frac{\langle\psi_{gs}|s\rangle\langle s|\psi_{gs}\rangle}{\langle\psi_{gs}|\psi_{gs}\rangle}$. Then, for each translation (m_1, m_2) and rotation m_r of s, we obtain

$$B_{nm_1m_2m_r}(s) = \frac{\langle s|\hat{T}_{m_1m_2}\hat{R}^{m_r}|t_{n00}\rangle}{\langle s|\psi_{gs}\rangle}.$$

From these values, the ratio $\widetilde{B}_{nk_1k_2}(s) = \langle s|\widetilde{t}_{nk_1k_2}\rangle/\langle s|\psi_{gs}\rangle$ can be obtained by Fourier-transforming B.

$$\widetilde{B}_{nk_1k_2}(s) = \frac{1}{L}\sum_{m_1,m_2=0}^{L-1} e^{-i\frac{2\pi}{L}(k_1m_1+k_2m_2)}\frac{1}{4}\big[B_{n,m_1,m_2,0}(s)$$
$$+ B_{n,m_1+1,m_2,1}(s) + B_{n,m_1+1,m_2+1,2}(s)$$
$$+ B_{n,m_1,m_2+1,3}(s)\big].$$

Finally, the average gives $G^{(k_1k_2)}_{nn'} = \langle\widetilde{B}_{nk_1k_2}(s)^*\widetilde{B}_{n'k_1k_2}(s)\rangle$. A similar procedure is followed to obtain $H^{(k_1k_2)}_{nn'}$.

APPENDIX B: EXACT CALCULATION FOR 2D HEISENBERG MODEL

In order to calculate the spectral function for a 6×6 lattice with periodic boundary conditions, we implement the translational and point-group symmetries to block-diagonalize the Hamiltonian using the QUSPIN exact diagonalization library [48,49]. Any local operator A_i explicitly breaks translational invariance, so instead we work with linear combinations of the local operators:

$$\hat{A}_{k_1,k_2} = \frac{1}{L}\sum_{n_1,n_2=0}^{L-1} e^{-i\frac{2\pi}{L}(k_1n_1+k_2n_2)}\hat{A}_{n_1,n_2}. \tag{B1}$$

The resulting operators have a well-defined quantum number for the momentum or point-group symmetries; therefore, when applying this operator to the ground state, the resulting state lives in the sector with the momentum given by the operator

$$|A(k_1,k_2)\rangle = \hat{A}_{k_1,k_2}|\psi_{gs}\rangle. \tag{B2}$$

With this state calculated we can calculate the Chebyshev moments as well as solve for the exact spectral function by using an iterative linear solver on the following linear equation:

$$(z - \hat{H})|A(k_1,k_2)\rangle = |B(k_1,k_2,z)\rangle, \tag{B3}$$

which can then be used to calculate the spectral function defined in momentum space:

$$G_{k_1,k_2}(z) = \frac{1}{\pi}\langle\psi_{gs}|\hat{A}^\dagger_{k_1,k_2}\frac{1}{z-\hat{H}}\hat{A}_{k_1,k_2}|\psi_{gs}\rangle$$
$$= \frac{1}{\pi}\langle A(k_1,k_2)|B(k_1,k_2,z)\rangle. \tag{B4}$$

[1] E. Dagotto, Rev. Mod. Phys. **66**, 763 (1994).
[2] H. B. Schüttler and D. J. Scalapino, Phys. Rev. B **34**, 4744 (1986).
[3] A. W. Sandvik, Phys. Rev. B **57**, 10287 (1998).
[4] R. N. Silver, D. S. Sivia, and J. E. Gubernatis, Phys. Rev. B **41**, 2380 (1990).

[5] J. E. Gubernatis, M. Jarrell, R. N. Silver, and D. S. Sivia, Phys. Rev. B **44**, 6011 (1991).
[6] O. F. Syljuåsen, Phys. Rev. B **78**, 174429 (2008).
[7] S. Fuchs, T. Pruschke, and M. Jarrell, Phys. Rev. E **81**, 056701 (2010).
[8] A. W. Sandvik, Phys. Rev. E **94**, 063308 (2016).
[9] H. Shao, Y. Q. Qin, S. Capponi, S. Chesi, Z. Y. Meng, and A. W. Sandvik, Phys. Rev. X **7**, 041072 (2017).
[10] S. R. White, Phys. Rev. Lett. **69**, 2863 (1992).
[11] S. R. White, Phys. Rev. B **48**, 10345 (1993).
[12] U. Schollwöck, Rev. Mod. Phys. **77**, 259 (2005).
[13] U. Schollwöck, Ann. Phys. (Amsterdam) **326**, 96 (2011).
[14] A. E. Feiguin, in *Strongly Correlated Systems: Numerical Methods*, edited by A. Avella and F. Mancini (Springer, Berlin, 2013), pp. 31–65.
[15] K. A. Hallberg, Phys. Rev. B **52**, R9827(R) (1995).
[16] T. D. Kühner and S. R. White, Phys. Rev. B **60**, 335 (1999).
[17] E. Jeckelmann, Phys. Rev. B **66**, 045114 (2002).
[18] P. E. Dargel, A. Honecker, R. Peters, R. M. Noack, and T. Pruschke, Phys. Rev. B **83**, 161104(R) (2011).
[19] P. E. Dargel, A. Wöllert, A. Honecker, I. P. McCulloch, U. Schollwöck, and T. Pruschke, Phys. Rev. B **85**, 205119 (2012).
[20] A. Nocera and G. Alvarez, Phys. Rev. E **94**, 053308 (2016).
[21] A. J. Daley, C. Kollath, U. Schollwöck, and G. Vidal, J. Stat. Mech.: Theor. Exp. (2004) P04005.
[22] S. R. White and A. E. Feiguin, Phys. Rev. Lett. **93**, 076401 (2004).
[23] A. E. Feiguin and S. R. White, Phys. Rev. B **72**, 020404(R) (2005).
[24] A. E. Feiguin, in *XV Training Course in the Physics of Strongly Correlated Systems*, edited by Adolfo Avella and Ferdinando Mancini, AIP Proceedings Vol. 1419 (AIP, Melville, NY, 2011), p. 5.
[25] A. E. Feiguin, in *Strongly Correlated Systems: Numerical Methods*, edited by A. Avella and F. Mancini (Springer, Berlin, 2013), pp. 131–152.
[26] A. Holzner, A. Weichselbaum, I. P. McCulloch, U. Schollwöck, and J. von Delft, Phys. Rev. B **83**, 195115 (2011).
[27] F. A. Wolf, J. A. Justiniano, I. P. McCulloch, and U. Schollwöck, Phys. Rev. B **91**, 115144 (2015).
[28] H. D. Xie, R. Z. Huang, X. J. Han, X. Yan, H. H. Zhao, Z. Y. Xie, H. J. Liao, and T. Xiang, Phys. Rev. B **97**, 075111 (2018).
[29] M. P. Zaletel, R. S. K. Mong, C. Karrasch, J. E. Moore, and F. Pollmann, Phys. Rev. B **91**, 165112 (2015).
[30] R. Verresen, F. Pollmann, and R. Moessner, Phys. Rev. B **98**, 155102 (2018).
[31] J. Haegeman, T. J. Osborne, and F. Verstraete, Phys. Rev. B **88**, 075133 (2013).
[32] L. Vanderstraeten, M. Mariën, F. Verstraete, and J. Haegeman, Phys. Rev. B **92**, 201111(R) (2015).
[33] L. Vanderstraeten, F. Verstraete, and J. Haegeman, Phys. Rev. B **92**, 125136 (2015).
[34] L. Vanderstraeten, J. Haegeman, and F. Verstraete, SciPost Phys. Lect. Notes 7 (2019), doi: 10.21468/SciPostPhysLectNotes.7.
[35] T. Li and F. Yang, Phys. Rev. B **81**, 214509 (2010).
[36] B. Dalla Piazza, M. Mourigal, N. B. Christensen, G. J. Nilsen, P. Tregenna-Piggott, T. G. Perring, M. Enderle, D. F. McMorrow, D. A. Ivanov, and H. M. Rønnow, Nat. Phys. **11**, 62 (2014).
[37] F. Ferrari, A. Parola, S. Sorella, and F. Becca, Phys. Rev. B **97**, 235103 (2018).
[38] F. Ferrari and F. Becca, Phys. Rev. B **98**, 100405(R) (2018).
[39] D. Hendry and A. E. Feiguin, Phys. Rev. B **100**, 245123 (2019).
[40] G. Carleo and M. Troyer, Science **355**, 602 (2017).
[41] H. Saito, J. Phys. Soc. Jpn. **86**, 093001 (2017).
[42] Z. Cai and J. Liu, Phys. Rev. B **97**, 035116 (2018).
[43] I. Glasser, N. Pancotti, M. August, I. D. Rodriguez, and J. I. Cirac, Phys. Rev. X **8**, 011006 (2018).
[44] J. Chen, S. Cheng, H. Xie, L. Wang, and T. Xiang, Phys. Rev. B **97**, 085104 (2018).
[45] A. Szabó and C. Castelnovo, Phys. Rev. Research **2**, 033075 (2020).
[46] N. Le Roux and Y. Bengio, Neural Comput. **20**, 1631 (2008).
[47] M. Ganahl, P. Thunström, F. Verstraete, K. Held, and H. G. Evertz, Phys. Rev. B **90**, 045144 (2014).
[48] P. Weinberg and M. Bukov, SciPost Phys. **2**, 003 (2017).
[49] P. Weinberg and M. Bukov, SciPost Phys. **7**, 020 (2019).