Letter

Ground-state properties via machine learning quantum constraints

Pei-Lin Zheng,* Si-Jing Du,* and Yi Zhang†
International Center for Quantum Materials, Peking University, Beijing 100871, China
and School of Physics, Peking University, Beijing 100871, China

(Received 11 June 2021; accepted 12 September 2022; published 22 September 2022)

Ground-state properties are central to our understanding of quantum many-body systems. At first glance, it seems natural and essential to obtain the ground state before analyzing its properties; however, its exponentially large Hilbert space has made such studies costly, if not prohibitive, on sufficiently large system sizes. Here, we propose an alternative strategy based upon the expectation values of an ensemble of operators and the elusive yet vital quantum constraints between them where the search for ground-state properties simply equates to classical constrained minimization. These quantum constraints are generally obtainable via sampling and then machine learning on a large number of systematically consistent quantum many-body states. We showcase our perspective on one-dimensional fermion chains and spin chains for applicability, effectiveness, caveats, and unique advantages especially for strongly correlated systems, thermodynamic-limit systems, property designs, etc.

DOI: 10.1103/PhysRevResearch.4.L032043

Introduction. The collective behaviors of quantum many-body systems are central to various cutting-edge fields in condensed-matter physics and beyond. Despite the nominal simplicity of certain quantum Hamiltonians, e.g., the Hubbard model [1–3], the noncommuting quantum operators squander any advantageous basis, and the exponentially large Hilbert space renders the solutions and characterizations of ground states costly, limiting the system size and geometry in numerical techniques, e.g., exact diagonalization and density matrix renormalization group (DMRG) [4,5]. While quantum Monte Carlo methods introduce efficient samplings, they are limited to sign-problem-free cases [6–8]. Also, the ground-state solution usually starts from scratch upon slight model modifications, making the systematic study of a complex phase diagram, not uncommon in condensed-matter physics [9], even more expensive.

Rather than the abstract quantum many-body ground state, we are usually interested in its properties, such as the ground-state energy and spontaneous-symmetry-breaking order parameters- (linear combinations of) expectation values of target observables. Considering that the minimum energy criteria also concern expectation values, one would be prompt to establish a study based solely on expectation values and cut out the ground state. However, the quantum operators follow nontrivial commutation relations and, as a result, enforce nontrivial quantum constraints upon their expectation values—a role played by the ground state as the mediator. Expectation values violating these quantum constraints do not have an underlying quantum state and may not reflect the true nature of the quantum many-body system. Therefore, such quantum constraints are complicated yet essential for proper expectation-value-based considerations. An example of such quantum constraints is the conformal bootstrap for conformal field theories [10], beyond which, however, the bootstrap reduces to mere bounds and no longer offers a controlled analysis of ground-state properties [11,12].

On the other hand, recent developments in machine learning [13,14] have revolutionized data analysis such as image recognition, spam, fraud detection, and autonomous driving [15]. Artificial neural networks (ANNs) can grasp the key yet complex and hidden rules in big datasets and generalize accurately for future scenarios [13–15]. Recently, machine learning has witnessed many explorations at the quantum many-body physics frontier, including quantum state tomography [16,17], quantum phase recognition [18–25], neural network states [26,27], experiment interpretations [28–30], etc.

In this Letter, we propose studying the ground-state properties of quantum many-body systems within a classical expectation-value framework with quantum constraints over an ensemble of important operators. Then, the ground-state properties amount to constrained minimization. We can generally encode such quantum constraints as ANNs via supervised machine learning on example quantum states. Without loss of generality, we showcase the unique advantages of our strategy on one-dimensional (1D) fermion and spin-1/2 models: (1) Compared with the expensive procedure of solving quantum many-body states, evaluations of expectation values are efficient and easily parallelizable among multiple operators and

*P.-L. Zheng and S.-J. Du are responsible for one-dimensional fermion chains and spin chains, respectively, and contributed equally.
†frankzhangyi@gmail.com

quantum states. (2) Our main effort is to extract and apply the quantum constraints through a large classical dataset of expectation values where machine learning techniques excel compatibly and proficiently. (3) Given a sufficiently diverse and representative training set, the obtained quantum constraints work for all Hamiltonians with different parameters where we iterate the classical constrained minimization with respect to the expectation values of different Hamiltonians. (4) We embed systematic properties, such as system size, geometry, and dimensions into the sample quantum many-body states and are rarely limited by them. (5) The quantum constraints also exhibit the competition and symbiosis between the observables, offering recipes for engineering models for desired ground-state properties, even emergent phases.

Algorithm. Our approach consists of steps as follows:

(1) Start with a large and representative ensemble of quantum many-body states $\{|\Phi\rangle_\alpha\}$ systematically consistent with the potential ground state, namely, obeying the expected symmetries and the area law.

(2) For each $|\Phi\rangle_\alpha$, evaluate the expectation values of a set of operators $\{\hat{O}_j\}$ and contribute a physical data point $\langle\hat{\mathbf{O}}\rangle_\alpha = (\langle\hat{O}_1\rangle_\alpha, \langle\hat{O}_2\rangle_\alpha, \ldots)$ in the $\langle\hat{\mathbf{O}}\rangle$ space. Operators with lower orders and spatial extents receive priority due to larger relevance and compatibility with local Hamiltonians. For comparison, unphysical data is obtained by considering deviations from the physical data [31].

(3) Via supervised machine learning on the training set $\{\langle\hat{\mathbf{O}}\rangle_\alpha\}$, train ANNs $f(\langle\hat{\mathbf{O}}\rangle)$ to distinguish physical (unphysical) values of $\langle\hat{\mathbf{O}}\rangle$ that is allowed (disallowed) by the quantum constraints.

(4) For the Hamiltonian $\hat{H} = \sum_j a_j \hat{O}_j$, search the constrained minimum of the energy $E = \sum_j a_j \langle\hat{O}_j\rangle$ with the quantum constraints $f(\langle\hat{\mathbf{O}}\rangle)$. The coordinates $\langle\hat{\mathbf{O}}\rangle_0$ of the resulting minimum offer the expectation values that characterize the ground state.

The first three steps yield the quantum constraints $f(\langle\hat{\mathbf{O}}\rangle)$ that mark the physical manifold in the classical $\langle\hat{\mathbf{O}}\rangle$ space. We expect a relatively smooth and continuous manifold as the adiabatic theorem ensures that the quantum many-body ground states and the corresponding $\langle\hat{\mathbf{O}}\rangle$ evolve continuously in the absence of first-order phase transitions. Importantly, the area law and symmetries vastly reduce the pool of quantum many-body states from the original Hilbert space, and machine learning can summarize and generalize from a limited number of training samples [13,14], making it feasible to extract the quantum constraints through a polynomial amount of sample states. Since evaluating expectation values is simple and efficient, the key is to obtain a diverse training set representative of the candidate parts of the Hilbert space, e.g., by teaming up multiple quantum many-body ansatzes.

Only the final step that applies the quantum constraints to the model Hamiltonians is repeated throughout a parameter space. Although the classical constrained optimizations are not guaranteed to be fully straightforward, compared with the exponential expense of brute-force quantum algorithms, the overall cost can be much less, especially given the available optimization algorithms and physical intuitions. For example, the solution $\langle\hat{\mathbf{O}}\rangle_0$ for one set of model parameters helps to initialize searches for its neighbors as $\langle\hat{\mathbf{O}}\rangle_0$ changes continuously in the absence of transitions. For efficiency, we can start with models with exact solutions or controlled approximations [32] and move progressively into other parts of the parameter space, tracking $\langle\hat{\mathbf{O}}\rangle_0$ successively in the process [31].

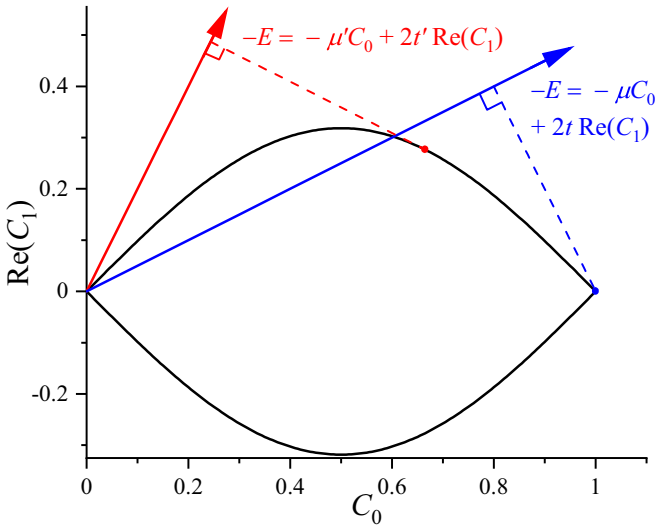

FIG. 1. The black contour shows the physical expectation values $[\mathrm{Re}(C_1), C_0]$ consistent with the quantum constraint in Eq. (2) for $\mathrm{Im}(C_1) = 0$. Those expectation values not on the contour are unphysical, i.e., no quantum many-body state can realize them. Given different Hamiltonians, e.g., $t = 1$, $\mu = -4$ (blue) and $t' = 1$, $\mu' = -1$ (red), their energy expectation values correspond to projections along different directions. The coordinates of the physical point with the lowest energy characterize the ground-state properties.

An important question is the choice of observables for which we suggest two criteria: (1) Is the observable likely to appear in target Hamiltonians? (2) Does the observable represent a physical quantity we are interested in? These favor local low-order operators, and the more, the better, although with added costs. Also, these criteria are soft: irrespective of chosen observables, quantum constraints address the physical realizability of their expectation values in a yes/no fashion; they do lose capacity without certain observables, e.g., tell corresponding degeneracy, but encounter no algorithmic breakdown.

A heuristic example. First, let us consider a 1D Fermi sea between $k_L = k_0 - k_F$ and $k_R = k_0 + k_F$ where we have a simple analytical expression for the quantum constraints. Its expectation values of two-point correlators are as follows:

$$C_0 = \langle c_x^\dagger c_x\rangle = k_F/\pi,$$
$$C_r = \langle c_{x+r}^\dagger c_x\rangle = \sin(k_F r) e^{ik_0 r}/\pi r, \quad r \neq 0, \tag{1}$$

irrespective of x due to the translation symmetry. Expectation values of higher-order operators depend fully on C_r's through Wick's theorem. In particular, the following quantum constraint holds between the most dominant real-valued C_0 and complex-valued C_1,

$$\pm\pi|C_1| = \sin(\pi C_0), \tag{2}$$

as illustrated in Fig. 1, which holds as long as there is one Fermi sea and no spontaneous translation symmetry breaking.

We note that the physical manifold (black contour) is smooth except for the two end points at $C_0 = 0, 1$, corresponding to Van Hove singularities.

Now, let us consider a tight-binding Hamiltonian with nearest-neighbor hopping $t \in \mathbb{R}$ [33] and Fermi energy μ,

$$\hat{H} = \sum_x -t(c_{x+1}^\dagger c_x + c_x^\dagger c_{x+1}) + \mu c_x^\dagger c_x,$$
$$E = [-2t\,\mathrm{Re}(C_1) + \mu C_0]N, \tag{3}$$

where N is the system size and we set $t = 1$ as our unit of energy hereinafter. As the constraint in Eq. (2) only concerns $|C_1|$, we set $\mathrm{Im}(C_1) = 0$ to allow maximal range for $\mathrm{Re}(C_1)$. A schematic for the solutions of optimal values of $[\mathrm{Re}(C_1), C_0]$ is in Fig. 1. More rigorously, we define $\pi C_0 = y \in (0, \pi)$, and $\mathrm{Re}(C_1)/C_0 = f(y) = \sin(y)/y \in (0, 1)$ is a single-valued function following the quantum constraint. To minimize $E \propto [\mu - 2tf(y)]y$, the subsequent solutions,

$$\begin{aligned} 2tf(y) - \mu + 2tf'(y)y = 0 \quad &\Rightarrow \quad 2t\,\cos(y) = \mu, \\ C_0 = y/\pi &= \arccos(\mu/2t)/\pi, \\ \mathrm{Re}(C_1) = yf(y)/\pi &= \sin(y)/\pi \\ &= \mathrm{sgn}(t)\sqrt{1 - \mu^2/4t^2}/\pi, \end{aligned} \tag{4}$$

which are consistent with the exact results obtained in the momentum space $H = \sum_k [-2t\,\cos(k) + \mu]c_k^\dagger c_k$ for generic values of t and μ.

For ground-state properties of Hamiltonians with upto nth-nearest-neighbor hopping, we need to employ quantum constraint $C_0 = f(C_1/C_0, C_2/C_0, \ldots)$ on the expectation values C_i, $i = 0, 1, 2, \ldots, n_{fs}$, which can be represented by ANNs and trained via supervised machine learning on quantum states with multiple Fermi seas [31]. For general quantum many-body systems, we may not formulate the quantum constraints as a function between the expectation values. It is more convenient to establish a "penalty" function $f(\langle\hat{\mathbf{O}}\rangle)$ that measures the extent of $\langle\hat{\mathbf{O}}\rangle$'s violations to the quantum constraints [31], which is also advantageous for allowing more freedom in choices of $\hat{\mathbf{O}}$. We will examine such formalism next.

Benchmark examples. We consider 1D fermion insulators with a bipartite unit cell, whose Bloch states take a general form $u(k) = [\cos(\theta_k/2), \sin(\theta_k/2)\exp(i\varphi_k)]^T$ with the first (second) component denoting the A (B) sublattice. The expectation values of two-point correlators are as follows:

$$\begin{aligned} C_0^{AA(BB)} &= 0.5 \pm g_0/2, \\ C_r^{AA(BB)} &= \pm g_r/2, \quad r \in \mathbb{Z}^+, \\ C_{r'}^{AB} &= \tilde{g}_{r'}/2, \quad r' \in \mathbb{Z} + 1/2, \end{aligned} \tag{5}$$

the rest obtainable via complex conjugation. $g_r = \int_0^{2\pi} \frac{dk}{2\pi} e^{ikr}\cos(\theta_k)$, $\tilde{g}_{r'} = \int_0^{2\pi} \frac{dk}{2\pi} e^{i(kr'+\varphi_k)}\sin(\theta_k)$, over which we establish the following quantum constraints,

$$\sum_r g_r g_{r+s}^* + \sum_{r'} \tilde{g}_{r'} \cdot \tilde{g}_{r'+s}^* = \delta_s. \tag{6}$$

g_r and $\tilde{g}_{r'}$, related to correlations in insulators, are fast decaying functions of r and r', allowing us to truncate at a finite distance $\Lambda = 20$ unless noted otherwise. We can, thus, define a positive-definite penalty function,

$$f(g_r, \tilde{g}_{r'}) = \sum_{s=0}^{\Lambda/2} \left[\sum_r g_r g_{r+s}^* + \sum_{r'} \tilde{g}_{r'} \cdot \tilde{g}_{r'+s}^* - \delta_s \right]^2, \tag{7}$$

which yields ~ 0 if and only if $\{g_r, \tilde{g}_{r'}\}$ are consistent with the quantum constraints. We note that the derivation of an expression as Eq. (7) is unavailable in generic quantum scenarios. Here for noninteracting fermions, it offers benchmarks to our strategy via machine learning quantum constraints in the following paragraphs.

Starting from random $u(k)$, we obtain 1.92×10^6 samples of $\{g_r, \tilde{g}_{r'}\}$ consistent with the quantum constraints and no penalty. We also include in the dataset 7.68×10^6 contrasting samples with small random deviations to $\{g_r, \tilde{g}_{r'}\}$ and corresponding penalties [31]. Besides, we utilize the gauge equivalence to reduce the degrees of freedom [31]. Then, we apply supervised machine learning [13,14] to train ANNs on the quantum constraints of $\{g_r, \tilde{g}_{r'}\}$ in the neighborhood of small or no violations [31]. In practice, we use the average output of multiple independent ANNs $f^*(g_r, \tilde{g}_{r'})$ as the approximate penalty and their max output as an acceptance threshold to avoid unphysical regions.

To test out these quantum constraints, we study the mean-field solutions of a 1D interacting fermion Hamiltonian at half-filling,

$$\hat{H} = \sum_x -t(c_{x+1}^\dagger c_x + c_x^\dagger c_{x+1}) + V c_{x+1}^\dagger c_{x+1} c_x^\dagger c_x. \tag{8}$$

The underlying assumptions of $f(g_r, \tilde{g}_{r'})$ and $f^*(g_r, \tilde{g}_{r'})$ are that the ground state takes a noninteracting fermion framework, hence, the Hartree-Fock approximation, and an emergent bipartite order parameter may spontaneously break the translation symmetry. Likewise, while our strategy straightforwardly applies to any quantum many-body ansatz, e.g., matrix product states [34–36], neural network states [26,27], quantum Monte Carlo methods, *ab initio* wave functions, even multiple ansatzes at the same time, the resulting quantum constraints will inherit the underlying assumptions and skip lower-energy scenarios beyond such assumptions, if any.

Under these circumstances, the energy expectation value is as follows:

$$\begin{aligned} E = \langle\hat{H}\rangle = \{&-t[\mathrm{Re}(\tilde{g}_{1/2}) + \mathrm{Re}(\tilde{g}_{-1/2})] \\ &+ 0.25V(2 - 2g_0^2 - |g_1|^2 - |g_{-1}|^2)\}N/2. \end{aligned} \tag{9}$$

We look for the constrained minimum $\{g_r, \tilde{g}_{r'}\}$ by minimizing either $L = \bar{E} + \eta f$ or $L = \bar{E} + \eta f^*$, where $\bar{E} = E/(N/2)$. η controls the weight of the quantum constraints, and the optimized results approach the physical limit asymptotically when $\eta \to \infty$. In practice, we should balance η between too large to allow an efficient search acceptance rate and too small to prevent the search from exiting regions represented by the samples. The extrapolation of η may offer a more systematic analysis, and an example is shown in the Supplemental Material [31].

Also, we use the expectation values of $\{g_r, \tilde{g}_{r'}\}$ at V to initialize the search at $V + \delta V$, and so on so forth. In practice, we start from $V = 2$ with an interval of $\delta V = -0.01$ [37]. The

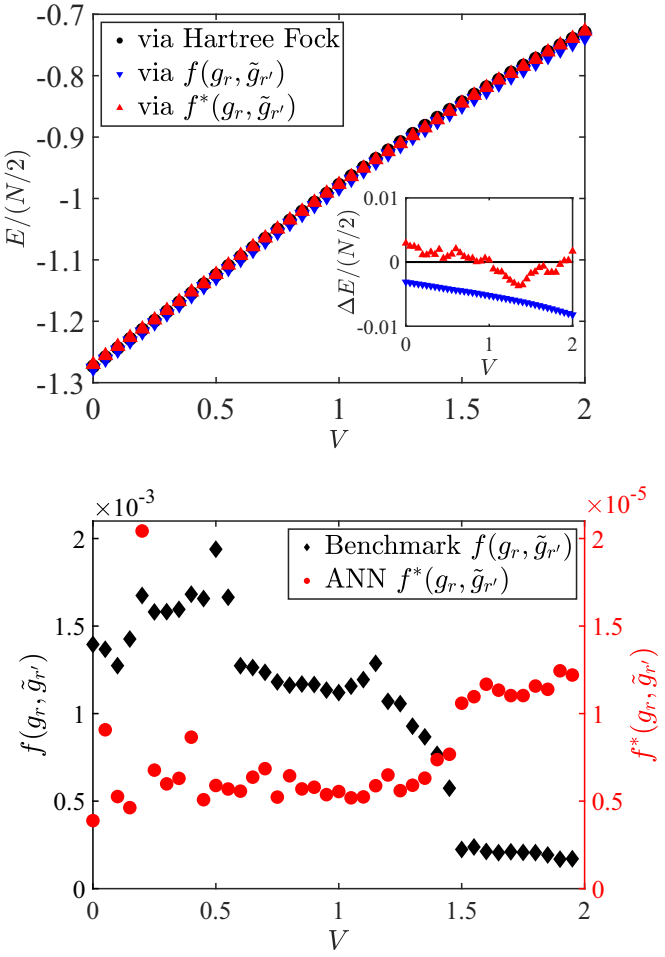

FIG. 2. Top: The ground-state energies and (the inset) their relative difference of the quantum many-body system in Eq. (8) from the Hartree-Fock approximation as well as the constrained minimization of $L = \bar{E} + \eta f$ and $L = \bar{E} + \eta f^*$, respectively, $\eta = 1000$. Note that the energies obtained with our strategy may end up slightly below the theoretical values since we have slacked the quantum constraints for improved efficiency. Bottom: The ANN outputs $f^*(g_r, \tilde{g}_{r'})$ for the constrained minimum as well as the benchmark $f(g_r, \tilde{g}_{r'})$ for the same $\{g_r, \tilde{g}_{r'}\}$ show consistency with the quantum constraints (very small penalty values) throughout the V range as we gradually lower from $V = 2$.

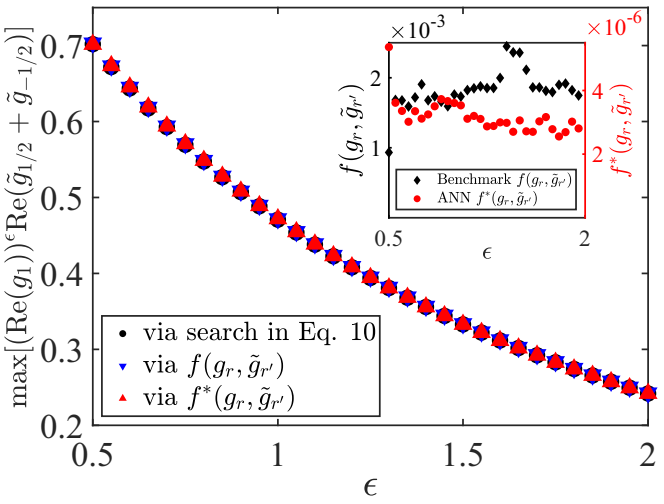

FIG. 3. The maximum of $[\mathrm{Re}(g_1)]^\epsilon \mathrm{Re}(\tilde{g}_{1/2} + \tilde{g}_{-1/2})$ for 1D fermion insulators on bipartite lattices shows good consistency between constrained optimization via the quantum constraints and searches among the variational Hamiltonian H_{var} in Eq. (10). Individual expectation values for realizing the maximum, e.g., g_1 and $\tilde{g}_{1/2}$, also check out for each ϵ. For the quantum-constraints approaches, we gradually increase ϵ from $\epsilon = 0.5$ while keeping track of $\{g_r, \tilde{g}_{r'}\}$. Inset: the ANN $f^*(g_r, \tilde{g}_{r'})$ and the benchmark $f(g_r, \tilde{g}_{r'})$ suggest the obtained $\{g_r, \tilde{g}_{r'}\}$ indeed obey the quantum constraints, $\eta = 1000$.

benchmark results are summarized in Fig. 2 and Ref. [31], and their consistency indicates that given sufficient dataset and training, machine learning can offer a trustworthy path toward quantum constraints. It is worth noting that such soft quantum constraints offer a distinctive and complementary perspective to conventional variational approaches: While the latter bounds the ground states from above, given a search space generally smaller than necessary, our method may approach the ground state from below, where near-physical regions join our consideration yielding a search space larger than permitted.

The quantum-constraint perspective also allows us to design quantum many-body systems like never before. Say we wish to apply certain criteria to expectation values: Sometimes it is as simple as the inclusion of the corresponding observables into the Hamiltonian, yet sometimes the criteria do not possess simple interpretations or require a nontrivial origin, such as spontaneous symmetry breaking. For instance, to maximize $[\mathrm{Re}(g_1)]^\epsilon \mathrm{Re}(\tilde{g}_{1/2} + \tilde{g}_{-1/2})$, $\epsilon \in [0.5, 2]$ for 1D fermion insulators on bipartite lattices, we commonly need to consider a variational Hamiltonian, such as

$$H_{\mathrm{var}} = \sum_x -t(c_{x+1}^\dagger c_x + \mathrm{H.c.}) + (-1)^x \Delta (c_{x+2}^\dagger c_x + \mathrm{H.c.}), \tag{10}$$

which balances operators favoring $\mathrm{Re}(\tilde{g}_{1/2} + \tilde{g}_{-1/2})$ and $\mathrm{Re}(g_1)$, respectively, and Δ is a variational parameter for optimization. More generally, larger variational spaces with additional operators are preferred for thorough searches, and the ground-state solutions may bring additional complications. On the other hand, with the quantum constraints, we can circumvent such difficulties and resort to a constrained maximization. We compare our results in Fig. 3. Furthermore, we can establish the underlying Hamiltonians and quantum states via the strategy in Ref. [38].

Strongly correlated scenarios. Generality is another essential merit of our strategy, which applies straightforwardly to strongly correlated systems and outshines the conventional methods hanging on the mind-boggling quantum many-body ground states themselves. Here, we illustrate the quantum constraints of 1D interacting spin-1/2 chains in the thermodynamic limit. Since the ground states of local Hamiltonians obey the area law, we can use tensor network states [4,5,39,40] especially infinite matrix product states [34–36] for infinite system sizes as our representation of quantum many-body state samples for machine learning quantum constraints. We also emphasize that it is straightforward to generalize, and it is beneficial to include other quantum may-body ansatzes, such

as projected states via variational Monte Carlo methods; see the Supplemental Material for examples and results [31].

First, we sample quantum many-body states with random, translation symmetric matrices of dimension $\chi = 8$ [41]. Next, we evaluate the expectation values $\langle\hat{\mathbf{O}}\rangle$ of a series of low-order spin operators S_r^{λ} and $S_r^{\lambda}S_{r+l}^{\lambda'}$, $\lambda, \lambda' = x, y, z$ upon a section of length $l_{\max} = 6$. Other than 6.75×10^4 of these quantum-constraints-abiding samples, we also include 2.025×10^5 contrasting samples with small deviations and corresponding penalties [31]. Then, we perform supervised machine learning on the dataset and train ANNs $f^*(\langle\hat{\mathbf{O}}\rangle)$ to recognize how well a target $\langle\hat{\mathbf{O}}\rangle$ aligns with the quantum constraints. We note that the trained ANNs, as well as the previous ANNs $f^*(g_r, \tilde{g}_{r'})$ and benchmark $f(g_r, \tilde{g}_{r'})$ for 1D fermion insulators, penalize expectation values' departure from and thereby enforcing quantum constraints as intended [31].

Subsequently, we use the quantum constraints for ground-state properties of quantum spin Hamiltonians. For instance, we apply our strategy with $f^*(\langle\hat{\mathbf{O}}\rangle)$ to the 1D transverse field ($h = 0$) and the nonintegrable longitudinal-transverse field Ising models ($h \neq 0$),

$$H = \sum_j -JS_j^zS_{j+1}^z - gS_j^x - hS_j^z, \tag{11}$$

where wet set $J = 1$ as the unit of energy. The results on the energy expectation value per site $E/N = \langle H\rangle/N = -J\langle S_0^zS_1^z\rangle - g\langle S_0^x\rangle - h\langle S_0^z\rangle$ and beyond in the $N \to \infty$ thermodynamic limit are summarized in Fig. 4 and the Supplemental Material [31]. Such quantum constraints are also directly applicable to the spin-1/2 XXZ chains [31]. To summarize, we obtain quantitative results on ground-state properties, such as energies and short-range correlators yet qualitative trends only on longer-range correlators, making pinpointing phase transitions relatively tricky. An additional or different set of observables addressing critical behaviors may be helpful.

Discussions. We propose to analyze ground-state properties via machine learning quantum constraints on expectation values and complement conventional ground-state-based approaches. Other than the aforementioned advantages, we have yet to establish a controlled quantitative analysis of algorithmic uncertainties especially for relatively soft degrees of freedom, e.g., the order parameter of a spontaneous symmetry-breaking phase. Qualitative tendencies are

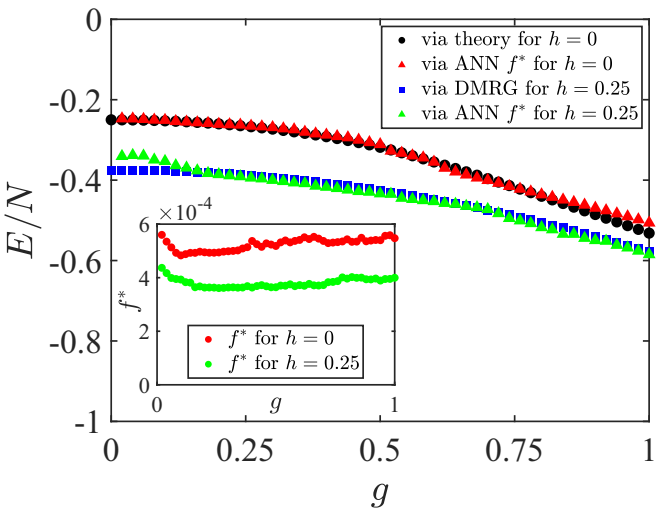

FIG. 4. The ground-state energies of the transverse field Ising model ($h = 0$) and longitudinal-transverse field Ising model ($h = 0.25$) in Eq. (11) obtained by theory solutions [42], DMRG, and the constrained minimization using ANN quantum constraints $f^*(\langle\hat{\mathbf{O}}\rangle)$ exhibit satisfactory consistency. For the latter, we start from $g = 0$ and gradually increase g. $\eta = 1000/3$ for $h = 0$ and $\eta = 400$ for $h = 2.5$, respectively. The inset: Small $f^*(\langle\hat{\mathbf{O}}\rangle)$ suggests that the obtained results well satisfies the quantum constraints.

observable, and extrapolation of η, the weight of quantum constraints, offers a partial solution [31]. Also, tensor network states in two dimensions and beyond may become costly, and other quantum many-body ansatzes may help complement the training data and simultaneously reduce biases originating from respective ansatz. Finally, while systematic presumptions, such as the area law and symmetries help narrow the questions and facilitate the calculations, such physics intuitions should sometimes be taken with a grain of salt.

Acknowledgements. We thank J. Shi, X. Yuan, Y. Qi, Z. Wang, and P.-C. Xie for insightful discussions. We are supported by the National Key R&D Program of China (Grant No. 2021YFA1401900) and the National Science Foundation of China (Grant No. 12174008). The computation was supported by the High-performance Computing Platform of Peking University. The source code is available in Ref. [60].

[1] R. T. Scalettar, E. Y. Loh, J. E. Gubernatis, A. Moreo, S. R. White, D. J. Scalapino, R. L. Sugar, and E. Dagotto, Phase Diagram of the Two-Dimensional Negative-U Hubbard Model, Phys. Rev. Lett. **62**, 1407 (1989).

[2] A. Moreo and D. J. Scalapino, Two-Dimensional Negative-U Hubbard Model, Phys. Rev. Lett. **66**, 946 (1991).

[3] J. P. F. LeBlanc, A. E. Antipov, F. Becca, I. W. Bulik, G. K.-L. Chan, C.-M. Chung, Y. Deng, M. Ferrero, T. M. Henderson, C. A. Jiménez-Hoyos, E. Kozik, X.-W. Liu, A. J. Millis, N. V. Prokof'ev, M. Qin, G. E. Scuseria, H. Shi, B. V. Svistunov, L. F. Tocchio, I. S. Tupitsyn *et al.*, Solutions of the Two-Dimensional Hubbard Model: Benchmarks and Results from a Wide Range of Numerical Algorithms, Phys. Rev. X **5**, 041041 (2015).

[4] S. R. White, Density Matrix Formulation for Quantum Renormalization Groups, Phys. Rev. Lett. **69**, 2863 (1992).

[5] U. Schollwöck, The density-matrix renormalization group, Rev. Mod. Phys. **77**, 259 (2005).

[6] E. Y. Loh, J. E. Gubernatis, R. T. Scalettar, S. R. White, D. J. Scalapino, and R. L. Sugar, Sign problem in the numerical simulation of many-electron systems, Phys. Rev. B **41**, 9301 (1990).

[7] W. M. C. Foulkes, L. Mitas, R. J. Needs, and G. Rajagopal, Quantum Monte Carlo simulations of solids, Rev. Mod. Phys. **73**, 33 (2001).

[8] M. Troyer and U.-J. Wiese, Computational Complexity and Fundamental Limitations to Fermionic Quantum Monte Carlo Simulations, Phys. Rev. Lett. **94**, 170201 (2005).

[9] E. Fradkin, S. A. Kivelson, and J. M. Tranquada, Colloquium, Rev. Mod. Phys. **87**, 457 (2015).

[10] D. Poland, S. Rychkov, and A. Vichi, The conformal bootstrap: Theory, numerical techniques, and applications, Rev. Mod. Phys. **91**, 015002 (2019).

[11] X. Han, Quantum many-body Bootstrap, arXiv:2006.06002.

[12] T. Deutsch, Exact solution of the many-body problem with a $\mathcal{O}(n^6)$ complexity, arXiv:2111.15281.

[13] M. A. Nielsen, *Neural Networks and Deep Learning* (Determination Press, 2015).

[14] Y. LeCun, Y. Bengio, and G. Hinton, Deep learning, Nature (London) **521**, 436 (2015).

[15] M. Jordan and T. Mitchell, Machine learning: Trends, perspectives, and prospects, Science **349**, 255 (2015).

[16] G. Torlai, G. Mazzola, J. Carrasquilla, M. Troyer, R. Melko, and G. Carleo, Neural-network quantum state tomography, Nat. Phys. **14**, 447 (2018).

[17] J. Carrasquilla, G. Torlai, R. G. Melko, and L. Aolita, Reconstructing quantum states with generative models, Nat. Mach. Intell. **1**, 155 (2019).

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

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

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

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

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

[23] Y. Zhang, R. G. Melko, and E.-A. Kim, Machine learning $\mathbb{Z}_2$ quantum spin liquids with quasiparticle statistics, Phys. Rev. B **96**, 245119 (2017).

[24] P. Zhang, H. Shen, and H. Zhai, Machine Learning Topological Invariants with Neural Networks, Phys. Rev. Lett. **120**, 066401 (2018).

[25] W. Lian, S.-T. Wang, S. Lu, Y. Huang, F. Wang, X. Yuan, W. Zhang, X. Ouyang, X. Wang, X. Huang, L. He, X. Chang, D.-L. Deng, and L. Duan, Machine Learning Topological Phases with a Solid-State Quantum Simulator, Phys. Rev. Lett. **122**, 210503 (2019).

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

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

[28] A. A. Melnikov, H. Poulsen Nautrup, M. Krenn, V. Dunjko, M. Tiersch, A. Zeilinger, and H. J. Briegel, Active learning machine learns to create new quantum experiments, Proc. Natl. Acad. Sci. USA **115**, 1221 (2018).

[29] Y. Zhang, A. Mesaros, K. Fujita, S. Edkins, M. Hamidian, K. Ch'ng, H. Eisaki, S. Uchida, J. S. Davis, E. Khatami *et al.*, Machine learning in electronic-quantum-matter imaging experiments, Nature (London) **570**, 484 (2019).

[30] C. Valagiannopoulos, Predicting the quantum texture from transmission probabilities, J. Appl. Phys. **127**, 174301 (2020).

[31] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevResearch.4.L032043 for further details on constrained optimization algorithms, generation of training samples, the settings of our ANNs, and additional examples and results on 1D spin and fermion chains, such as multiple Fermi seas, which include Refs. [43–56].

[32] Such scenarios also provide valuable additional training samples.

[33] The arguments can generalize to complex t straightforwardly noting that the minimum energy requires $\arg C_1 = -\arg t$.

[34] F. Verstraete, V. Murg, and J. Cirac, Matrix product states, projected entangled pair states, and variational renormalization group methods for quantum spin systems, Adv. Phys. **57**, 143 (2008).

[35] G. Vidal, Classical Simulation of Infinite-Size Quantum Lattice Systems in One Spatial Dimension, Phys. Rev. Lett. **98**, 070201 (2007).

[36] J. Jordan, R. Orús, G. Vidal, F. Verstraete, and J. I. Cirac, Classical Simulation of Infinite-Size Quantum Lattice Systems in Two Spatial Dimensions, Phys. Rev. Lett. **101**, 250602 (2008).

[37] Although the model at $V = 0$ also offers a good starting point, the truncation in r and r' for a metal state may cause issues.

[38] J.-B. Wang and Y. Zhang, Single-shot quantum measurements sketch quantum states, arXiv:2203.01348.

[39] M. Fannes, B. Nachtergaele, and R. F. Werner, Finitely correlated states on quantum spin chains, Commun. Math. Phys. **144**, 443 (1992).

[40] U. Schollwöck, The density-matrix renormalization group in the age of matrix product states, Annal. Phys. (NY) **326**, 96 (2011).

[41] We note that random matrix product states have emergent statistics [57–59] against sampling generality at large dimension χ, hence, our small χ choice.

[42] E. Lieb, T. Schultz, and D. Mattis, Two soluble models of an antiferromagnetic chain, Annal. Phys. (NY) **16**, 407 (1961).

[43] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Kopf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai *et al.*, Pytorch: An imperative style, high-performance deep learning library, in *Advances in Neural Information Processing Systems*, edited by H. Wallach, H. Larochelle, A. Beygelzimer, F. Alche-Buc, E. Fox, and R. Garnett (Curran, Vancouver, Canada, 2019), Vol. 32.

[44] J. Han and C. Moraga, The influence of the sigmoid function parameters on the speed of backpropagation learning, in *From Natural to Artificial Neural Computation*, edited by J. Mira and F. Sandoval (Springer, Berlin/Heidelberg, 1995), pp. 195–201.

[45] K. He, X. Zhang, S. Ren, and J. Sun, Deep residual learning for image recognition, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV* (IEEE, Piscataway, NJ, 2016).

[46] K. Fukushima, Visual feature extraction by a multilayered network of analog threshold elements, IEEE Trans. Syst. Sci. Cybern. **5**, 322 (1969).

[47] X. Glorot, A. Bordes, and Y. Bengio, Deep sparse rectifier neural networks, in *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*, Proceedings of Machine Learning Research, edited by G. Gordon, D. Dunson, and M. Dudík (PMLR, Fort Lauderdale, FL, 2011), Vol. 15, pp. 315–323.

[48] S. Ioffe and C. Szegedy, Batch normalization: Accelerating deep network training by reducing internal covariate shift, in *Proceedings of the 32nd International Conference on International Conference on Machine Learning, ICML'15* (JMLR.org, Lille, France, 2015), Vol. 37, pp. 448–456.

[49] J. Hauschild and F. Pollmann, Efficient numerical simulations with Tensor Networks: Tensor Network Python (TeNPy), SciPost Phys. Lect. Notes, 5 (2018).

[50] C. E. Brodley and M. A. Friedl, Identifying mislabeled training data, J. Artif. Intell. Res. **11**, 131 (1999).

[51] E. Beigman and B. B. Klebanov, Learning with annotation noise, in *ACL 2009, Proceedings of the 47th Annual Meeting of the Association for Computational Linguistics and the 4th International Joint Conference on Natural Language Processing of the AFNLP, Singapore*, edited by K. Su, J. Su, and J. Wiebe (The Association for Computer Linguistics, Singapore, 2009), pp. 280–287.

[52] D. Rolnick, A. Veit, S. J. Belongie, and N. Shavit, Deep learning is robust to massive label noise, arXiv:1705.10694.

[53] Y. Kim, J. Yim, J. Yun, and J. Kim, NLNL: negative learning for noisy labels, in *2019 IEEE/CVF International Conference on Computer Vision, ICCV 2019, Seoul, Korea (South), October 27 - November 2, 2019* (IEEE, Piscataway, NJ, 2019), pp. 101–110.

[54] M. Zhao, S. Zhong, X. Fu, B. Tang, and M. Pecht, Deep residual shrinkage networks for fault diagnosis, IEEE Trans. Indust. Inform. **16**, 4681 (2020).

[55] M. J. D. Powell, A direct search optimization method that models the objective and constraint functions by linear interpolation, in *Advances in Optimization and Numerical Analysis*, edited by S. Gomez and J.-P. Hennart (Springer, Dordrecht, 1994), pp. 51–67.

[56] M. J. D. Powell, Uobyqa: Unconstrained optimization by quadratic approximation, Math. Program. **92**, 555 (2002).

[57] S. Garnerone, T. R. de Oliveira, and P. Zanardi, Typicality in random matrix product states, Phys. Rev. A **81**, 032336 (2010).

[58] S. Garnerone, T. R. de Oliveira, S. Haas, and P. Zanardi, Statistical properties of random matrix product states, Phys. Rev. A **82**, 052312 (2010).

[59] J. Haferkamp, C. Bertoni, I. Roth, and J. Eisert, Emergent statistical mechanics from properties of disordered random matrix product states, PRX Quantum **2**, 040308 (2021).

[60] https://github.com/PeilinZHENG/MLQC