POLITECNICO DI TORINO
Repository ISTITUZIONALE

Variational benchmarks for quantum many-body problems

Original
Variational benchmarks for quantum many-body problems / Wu, D., Rossi, R., Vicentini, F., Astrakhantsev, N., Becca, F., Cao, X., Carrasquilla, J., Ferrari, F., Georges, A., Hibat-Allah, M., Imada, M., Läuchli, A.M., Mazzola, G., Mezzacapo, A., Millis, A., Robledo Moreno, J., Neupert, T., Nomura, Y., Nys, J., Parcollet, O., et al.. - In: SCIENCE. - ISSN 0036-8075. - STAMPA. - 386:6719(2024), pp. 296-301. [10.1126/science.adg9774]

Availability:
This version is available at: 11583/2994200 since: 2024-11-11T13:53:10Z

Publisher:
American Association for the Advancement of Science

Published
DOI:10.1126/science.adg9774

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright
AAAS preprint/submitted version e/o post-print Author Accepted Manuscript

preprint/submitted version e/o post-print Author Accepted Manuscript

(Article begins on next page)

26 November 2026

Variational Benchmarks for Quantum Many-Body Problems

Dian Wu,[1,2] Riccardo Rossi,[1,3] Filippo Vicentini,[2,4,5] Nikita Astrakhantsev,[6] Federico Becca,[7] Xiaodong Cao,[8] Juan Carrasquilla,[9,10] Francesco Ferrari,[11] Antoine Georges,[5,8,4,12] Mohamed Hibat-Allah,[9,13,14,15] Masatoshi Imada,[16,17,18,19] Andreas M. Läuchli,[20,1] Guglielmo Mazzola,[21] Antonio Mezzacapo,[22] Andrew Millis,[23,8] Javier Robledo Moreno,[8,24] Titus Neupert,[6] Yusuke Nomura,[25,26] Jannes Nys,[1,2] Olivier Parcollet,[8,27] Rico Pohle,[17,19] Imelda Romero,[1,2] Michael Schmid,[17] J. Maxwell Silvester,[28] Sandro Sorella,[29] Luca F. Tocchio,[30] Lei Wang,[31,32] Steven R. White,[28] Alexander Wietek,[33] Qi Yang,[31,34] Yiqi Yang,[35] Shiwei Zhang,[8] and Giuseppe Carleo[1,2,*]

[1] *Institute of Physics, École Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland*
[2] *Center for Quantum Science and Engineering, École Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland*
[3] *Sorbonne Université, CNRS, Laboratoire de Physique Théorique de la Matière Condensée (LPTMC), F-75005 Paris, France*
[4] *CPHT, CNRS, École Polytechnique, Institut Polytechnique de Paris, F-91128 Palaiseau, France*
[5] *Collège de France, Université PSL, F-75005 Paris, France*
[6] *Department of Physics, University of Zurich, CH-8057 Zurich, Switzerland*
[7] *Dipartimento di Fisica, Università di Trieste, I-34151 Trieste, Italy*
[8] *Center for Computational Quantum Physics, Flatiron Institute, New York, New York 10010, USA*
[9] *Vector Institute, MaRS Centre, Toronto, Ontario M5G 1M1, Canada*
[10] *Institute for Theoretical Physics, ETH Zürich, CH-8093 Zurich, Switzerland*
[11] *Institut für Theoretische Physik, Goethe-Universität, 60438 Frankfurt am Main, Germany*
[12] *Department of Quantum Matter Physics, Université de Genève, CH-1211 Geneva, Switzerland*
[13] *Perimeter Institute for Theoretical Physics, Waterloo, Ontario N2L 2Y5, Canada*
[14] *Department of Physics and Astronomy, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada*
[15] *Department of Applied Mathematics, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada*
[16] *Toyota Physical and Chemical Research Institute, Nagakute, Aichi 480-1192, Japan*
[17] *Waseda Research Institute for Science and Engineering, Waseda University, Shinjuku-ku, Tokyo 169-8555, Japan*
[18] *Physics Division, Sophia University, Chiyoda-ku, Tokyo 102-8554, Japan*
[19] *Department of Applied Physics, University of Tokyo, Bunkyo-ku, Tokyo 113-8656, Japan*
[20] *Laboratory for Theoretical and Computational Physics, Paul Scherrer Institute, CH-5232 Villigen, Switzerland*
[21] *Institute for Computational Science, University of Zurich, CH-8057 Zurich, Switzerland*
[22] *IBM Quantum, T. J. Watson Research Center, Yorktown Heights, New York 10598, USA*
[23] *Department of Physics, Columbia University, New York, New York 10027, USA*
[24] *Center for Quantum Phenomena, Department of Physics, New York University, New York, New York 10003, USA*
[25] *Department of Applied Physics and Physico-Informatics, Keio University, Kohoku-ku, Yokohama 223-8522, Japan*
[26] *Institute for Materials Research, Tohoku University, Aoba-ku, Sendai 980-8577, Japan*
[27] *Université Paris-Saclay, CNRS, CEA, Institut de physique théorique, F-91191 Gif-sur-Yvette, France*
[28] *Department of Physics and Astronomy, University of California, Irvine, California 92697, USA*
[29] *SISSA, International School for Advanced Studies, I-34136 Trieste, Italy*
[30] *Institute for Condensed Matter Physics and Complex Systems, Department of Applied Science and Technology (DISAT), Politecnico di Torino, I-10129 Torino, Italy*
[31] *Institute of Physics, Chinese Academy of Sciences, Beijing 100190, China*
[32] *Songshan Lake Materials Laboratory, Dongguan, Guangdong 523808, China*
[33] *Max Planck Institute for the Physics of Complex Systems, 01187 Dresden, Germany*
[34] *University of Chinese Academy of Sciences, Beijing 100049, China*
[35] *Department of Physics, College of William and Mary, Williamsburg, Virginia 23187, USA*

(Dated: Sept. 12, 2024)

The continued development of computational approaches to many-body ground-state problems in physics and chemistry calls for a consistent way to assess its overall progress. In this work, we introduce a metric of variational accuracy, the V-score, obtained from the variational energy and its variance. We provide an extensive curated dataset of variational calculations of many-body quantum systems, identifying cases where state-of-the-art numerical approaches show limited accuracy, and future algorithms or computational platforms, such as quantum computing, could provide improved accuracy. The V-score can be used as a metric to assess the progress of quantum variational methods toward a quantum advantage for ground-state problems, especially in regimes where classical verifiability is impossible.

I. INTRODUCTION

A key aspect of the quantum many-body problem, for systems ranging from the subatomic to molecules and materials, is determining the ground-state properties and energy. Knowing the ground state, one can predict which systems are stable and whether these systems exhibit useful and exotic phases, such as superconductivity or spin liquid states. However, because of the exponential complexity of the quantum wave function, finding the ground state of a many-body system can be very challenging, which limits exact numerical studies to a small number of particles. Efficiently solving the general ground-state problem is largely believed to be intractable. However, this does not necessarily apply to any particular system or class of systems, which may admit powerful approximations for ground states. Decades of research have focused on devising computational methods to find approximate solutions for specific cases of interest.

These computational methods have widely varying degrees of accuracy, and typically each method is much more successful on some systems than on others. Some of the most widely used methods include quantum Monte Carlo (QMC) [1–3], tensor networks (TNs) [4, 5], and dynamical mean field theory (DMFT) and its extensions [6, 7]. It is known that the applicability of QMC methods is negatively affected by the frustration of the quantum system and particle statistics [8]; similarly, high entanglement and large correlation lengths limit the applicability of TN [9] and DMFT [7], respectively. Variational approaches based on physically motivated ansatzes [10, 11] or neural networks [12] are not explicitly affected by the aforementioned issues. However, it is more difficult to assess their applicability and accuracy for a given quantum many-body system.

Quantum computers provide an alternative platform to attack quantum many-body problems [13]. Notably, the dynamics of quantum many-body systems can be efficiently simulated by a digital quantum computer when the initial states are easy to prepare [14]. Besides dynamics, substantial attention has been devoted to preparing ground states that are difficult to study with classical algorithms. Quantum algorithms for this task include phase estimation [15], variational approaches [16–19], adiabatic passage [20], imaginary time evolution [21], and subspace and Lanczos methods [22, 23].

A fundamental challenge in assessing newly established computational methods based on classical or quantum computing is defining a consistent accuracy metric. Especially for ground-state problems, such a metric is necessary to clearly identify target Hamiltonians of broad interest, which cannot be solved with sufficient accuracy by existing methods. Also, this metric is crucial to quantify the improvements of computational approaches with time. In the context of assessing quantum computing-based methods, this issue pertains to the broader problem of determining in what cases quantum computers have an advantage over classical ones [24, 25].

Determining a consistent metric for physically and chemically relevant ground-state problems is one of the goals of this work. To this end, we provide a large, curated collection of variational and numerically exact results on strongly correlated lattice models obtained by both state-of-the-art and baseline methods. The data that we provide include multiple approaches, such as exact diagonalization (ED), QMC [1] in the auxiliary field algorithm [26–30], matrix product states (MPSs) [4], variational wave functions formulated on a lattice [31], and neural network-based methods [12]. In addition to providing the data, we introduce an indicator of the variational accuracy of these results, named the V-score, that is suitable for directly comparing classical and quantum computing-based variational approaches. The V-score, obtained as a combination of the mean energy and its variance of a given variational state, allows us to identify what Hamiltonians and regimes are hard to approximate with classical variational methods without prior knowledge of the exact solution. Furthermore, we argue that the V-score can be used as a controlled benchmark to quantify the continued progress of quantum algorithms and quantum hardware to simulate those challenging target Hamiltonians.

II. BENCHMARKING VARIATIONAL ALGORITHMS

We focus our study on benchmarking classical and quantum variational algorithms in approximating ground states of quantum many-body systems. On the classical side, these algorithms involve explicitly maintained variational representations of wave functions, such as TN or variational Monte Carlo (VMC)-based approaches. On the quantum side, the variational methods of major interest involve parameterized quantum circuits (PQCs) or other state preparation techniques based on local unitary transformations. In all cases, we assume that the methods to be benchmarked allow unbiased estimates of expectation values for Hamiltonians with few-body interactions (k-local operators, in the language of quantum information). Such expectation values can be obtained with a controllable statistical error, as in the case of classical Monte Carlo-based techniques, or as a result of statistical noise caused by measurements on quantum hardware.

A. Choice of problems

There is large freedom in the choice of many-body quantum problems that can be used to benchmark computational techniques. In this work, we have decided

* Corresponding author. Email: giuseppe.carleo@epfl.ch

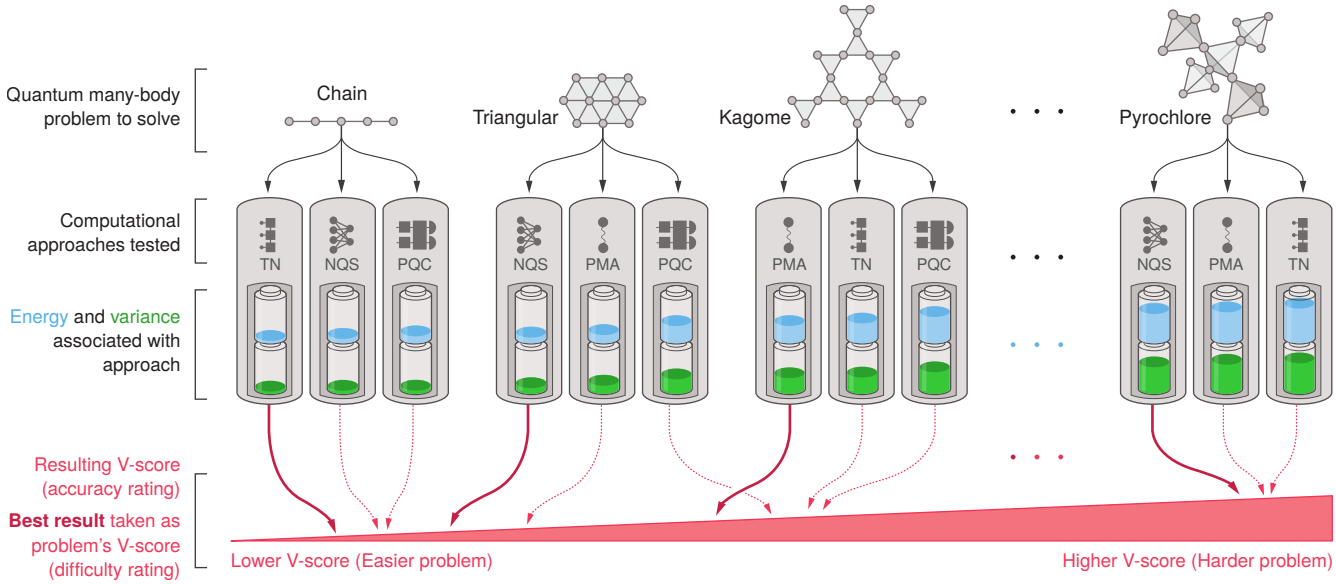


Figure 1. **Sketch of the V-score as a metric of simulation hardness.** In this work, we present an extensive dataset of computational results for quantum many-body ground-state problems (for this sketch, we have selected a spin-$\frac{1}{2}$ system on a chain, a kagome, and a pyrochlore lattice). For each Hamiltonian in the dataset, we compute the mean energy and its variance with different variational techniques, including physically motivated ansatzes (PMAs), neural quantum states (NQSs), tensor networks (TNs), and parameterized quantum circuits (PQCs). The energy and the variance are combined into the V-score, a metric of variational accuracy that we introduce in the main text. A low V-score is associated with high accuracy. The best result per Hamiltonian is then taken as the V-score for that Hamiltonian, and used to rank the Hamiltonians in terms of simulation accuracy, highlighting which quantum many-body models are hard to simulate with current methods.

to focus on lattice Hamiltonians. These are minimal models of strong correlations and typically capture the essence of many physical systems. Lattice models first rose to prominence within classical statistical mechanics with the definition of the Ising model [32]. Within solid-state physics, they stem from tight binding approaches to describe the electronic band structure [33]. More recently, within the second quantization formalism, they are routinely used in different areas of physics to understand the low-energy behavior of unconventional quantum phases and transitions among them [34, 35]. In this regard, the transverse-field Ising model (TFIM) provides the simplest example of a zero-temperature phase transition purely driven by quantum fluctuations between a paramagnet and a ferromagnet as seen, for example, in the Ising ferromagnet $LiHoF_4$ [34, 36]. Other prominent examples are the various quantum impurity models, in which a localized interacting degree of freedom is embedded into a non-interacting bulk, such as the Anderson impurity model [37]. Quantum impurity models are central to quantum embedding methods, such as DMFT [7], and have applications to nanoelectronic devices [38]. Their lattice generalizations, such as the Kondo lattice model, describe heavy fermion systems with 4f or 5f atoms, such as Ce or U [39].

Similarly, the Hubbard model [40–42] has been widely used to capture the essence of strong correlation in solids and has been shown to be relevant to the study of high-temperature superconductivity in cuprate compounds—e.g., $La_{2-x}Sr_xCuO_4$ [43]—and the Mott metal-insulator transitions in a variety of compounds [44]. A descendant of the Hubbard model, the Heisenberg model describes a wide range of magnetic phases, for example, with ferromagnetic or antiferromagnetic orders [45]. In addition, when defined on geometrically frustrated lattices, possibly with anisotropic super-exchange couplings, the Heisenberg model gives rise to a wealth of phenomena, including spin liquid phases with topological order and exotic critical points [46, 47]. In this respect, the rare earth compound $YbMgGaO_4$ [48] and the mineral herbertsmithite $ZnCu_3(OH)_6Cl_2$ [49] have offered examples for unconventional quantum phases on triangular and kagome lattices.

B. V-score

To quantify the accuracy of two or more variational methods applied on the same ground-state approximation task, a key indicator is the expectation value of the energy $E = \langle \hat{H} \rangle$, an unbiased metric to assess the relative accuracy of variational methods; here, $\hat{H}$ is the relevant Hamiltonian. Given, for example, two independent methods preparing approximate ground states with variational energies E_a and E_b, the one providing the lower energy can be considered more accurate. From a prac-

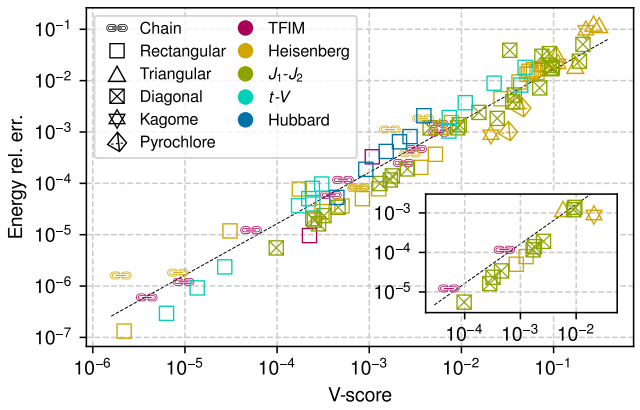


Figure 2. **Validation of V-score against exact results.** We compare V-scores versus energy relative errors on various strongly correlated models for which exact results (ED or QMC) are available. The black dashed line is a least-squares fit of $\log(\text{energy rel. err.}) = \log(\text{V-score}) + C$, where $C = -1.80 \pm 0.08$. The inset focuses on PQC results run on classical hardware (no shot noise included). The symbol shapes and colors correspond to the lattice geometry and model type, respectively (see legend).

tical point of view, however, it is preferable to have an absolute metric capable of predicting the accuracy of a method without comparing it with other methods. This would, for instance, allow comparing the performance of a given method on different tasks. Nonetheless, it is unlikely we could find such a metric that is provably applicable in all cases, because its existence would also allow the solution of NP-hard problems [51]. We are therefore forced to settle for an empirically applicable metric. Moreover, the metric should be easy to estimate with variational methods.

Apart from the mean energy, for most variational methods, we also have a controllable estimate of the energy variance $\operatorname{Var} E = \langle \hat{H}^2 \rangle - \langle \hat{H} \rangle^2$. It has the important property that it exactly vanishes if computed on the exact ground state. Therefore, $\operatorname{Var} E$ can be used to infer some information about the distance of the variational energy E from the exact, and a priori unknown, ground-state energy E_0. After early empirical observations [52], it has been shown that $\operatorname{Var} E$ scales linearly with the deviation $E - E_0$ [53–55], so it can be used as a measure of the accuracy of the variational state.

We can use E and $\operatorname{Var} E$ to create a dimensionless, intensive combination:

$$\text{V-score} := \frac{N \operatorname{Var} E}{(E - E_\infty)^2}, \tag{1}$$

where N is the number of degrees of freedom, which is the number of spins for spin models, and the number of particles for fermionic models. The constant E_∞ serves as a zero point of the energy, compensating for any global energy shift in the Hamiltonian definition. The V-score is dimensionless in energy units and system size for the variational states that we consider and is also invariant under energy shifts by construction. The procedure to benchmark the V-score is sketched in Fig. 1, and a detailed discussion of the definition of the V-score is presented in the supplementary materials [56].

To further justify the definition of the V-score, in Fig. 2, we present a comparison of this quantity against the energy relative error $(E - E_0)/(E_\infty - E_0)$ for a wide range of Hamiltonians and variational methods, where the ground-state energy E_0 is obtained by ED or numerically exact QMC. Despite the great diversity of Hamiltonians and variational methods considered, the V-score is a notably consistent and reliable estimator for the order of magnitude of the energy relative error, as shown by the linear fit in Fig. 2. In the inset of Fig. 2, we show that the same linear fit also well describes classically simulated PQCs, optimized with the variational quantum eigensolver (VQE) algorithm. These results validate the V-score as an absolute performance metric for both classical and quantum variational algorithms, at least for the Hamiltonians and the techniques we consider in the paper.

III. IDENTIFYING HARD PROBLEMS

We can now discuss which Hamiltonians are hard for the state-of-the-art variational methods presented in our collection. Given the intrinsic exponential complexity of the problem, it is no longer possible to obtain ED results on larger system sizes. Thus, using the V-score as a guide in this task is crucial. In Fig. 3 we show the V-score of all Hamiltonians and methods in our dataset. We first select the best method for each Hamiltonian by choosing the one with the lowest variational energy. We then use this method's V-score as an absolute hardness metric in the ground-state approximation task, which we refer to as the V-score of this Hamiltonian. The V-score of the best-performing method is marked in bold in Fig. 3. For additional clarity, in Fig. 4, we classify all those results by Hamiltonian types and lattice geometries.

It is well known that one-dimensional (1D) chain geometries are easy to solve with the density matrix renormalization group (DMRG). The small values of their V-scores in Fig. 3 and Fig. 4 clearly label them as accessible, particularly for spin models. Unfrustrated spin models typically also have small V-scores, ranging from 10^{-6} to 10^{-4} for the TFIM and the Heisenberg model on square lattices with open boundary conditions (OBCs). Moreover, these models can be efficiently simulated with unbiased stochastic techniques, such as QMC, and thus are easy to study on classical computers. On the other hand, the V-scores show that frustrated geometries, such as triangular, kagome, pyrochlore, and the J_1-J_2 square lattice, as well as fermionic models, such as the Hubbard model, are the most demanding for variational algorithms. With the help of Fig. 2, we can infer the order of magnitude of the energy relative error from the V-score.

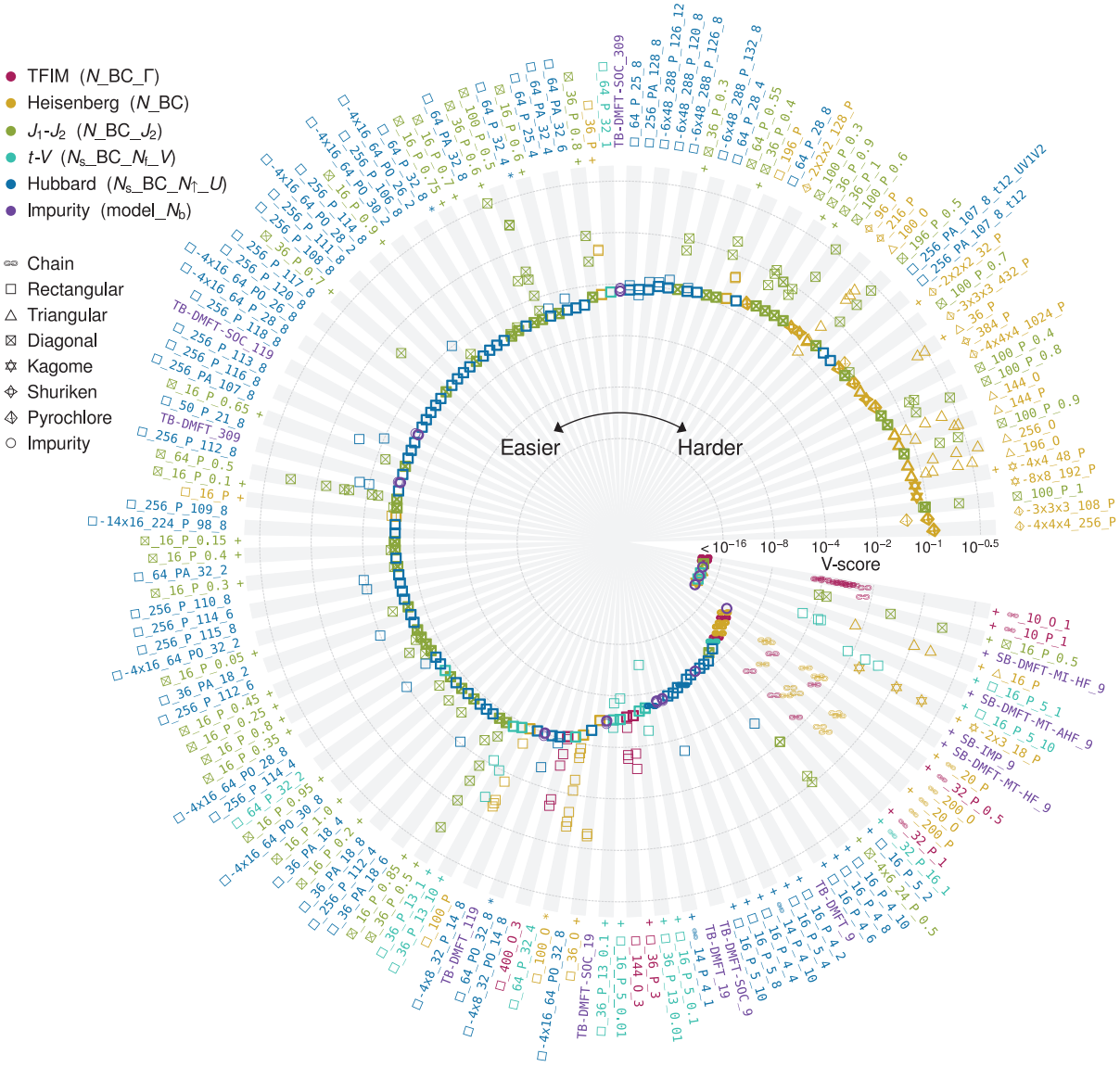


Figure 3. **V-score rankings of the complete dataset of variational benchmark results.** Each radial slice shows the results for a given Hamiltonian. The bold marker in the slice shows the method with the lowest variational energy (rather than the lowest V-score) for that Hamiltonian, and its V-score determines the clockwise ranking of that Hamiltonian. The radial axis uses the doubly log scale to show the results across a wide range of V-score magnitude. The results with V-scores less than 10^{-16} are indistinguishable from the exact solutions because of limited numerical precision. The label outside each slice is the name of the Hamiltonian in the dataset, which describes the lattice geometry, the lattice size N_s, the boundary conditions (BCs), and the Hamiltonian parameters such as the interaction strength (Γ, J_2, V, U) and the number of particles (N_f, $N_\uparrow$, N_b). If a rectangular or diagonal lattice is non-square, its two edge lengths are shown such as "□-4x8". Otherwise it is assumed to be square, and it can be tilted, allowing for numbers of spins such as 50. In a kagome or pyrochlore lattice, the number of copies of the unit cell in each spatial dimension is shown, such as "✡-4x4". The BC can be O (open), P (periodic), or A (anti-periodic) on each dimension. See legend for the meaning of the various parameters associated with each Hamiltonian type. Additionally, we include results on the generalized Hubbard model with second nearest neighbor hopping t_2, nearest neighbor repulse V_1, and second nearest neighbor repulse V_2, as indicated by the suffixes "_t12" and "_t12_UV1V2". Formal definitions of the Hamiltonian parameters and specific types of impurity models are provided in the supplementary materials [50]. The "+" marker near the label indicates that an ED solution is available for that Hamiltonian, and the "*" marker indicates a numerically exact QMC solution.

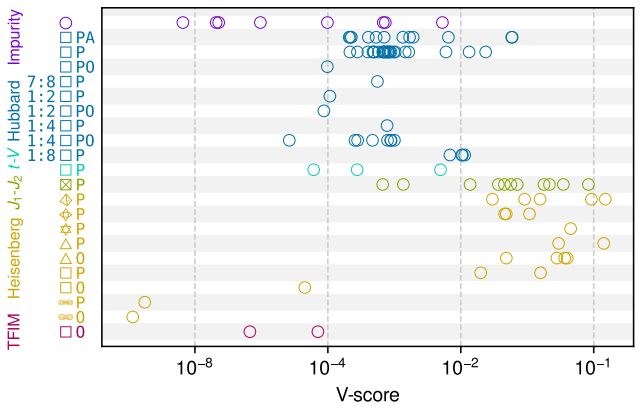


Figure 4. **V-scores of Hamiltonians,** classified by Hamiltonian types and lattice geometries. A Hamiltonian's V-score is defined as the V-score of the available variational method with the lowest energy on this Hamiltonian. Only the Hamiltonians without ED results are shown. The ratios to the left of the lattice icons are aspect ratios of rectangular lattices. The letters to the right are boundary conditions, which can be O (open), P (periodic), or A (anti-periodic) on each dimension.

When this correspondence is applied to those hard problems, it predicts that we cannot expect an accuracy on the energy better than one or two digits.

A. A perspective on quantum advantage

Many recent theoretical and experimental efforts have been dedicated to showing the computational advantage of quantum computers over classical computers. Informed by theoretical computer science arguments, random circuit sampling has been proposed as a specific task to show such advantage [57, 58]. Still, it is unclear whether current noisy experimental quantum computing platforms can provide a solid and scalable advantage over classical ones [59–61]. In addition to tasks of purely theoretical interest, there is also growing interest in finding practical quantum advantage [62], where quantum devices show a speedup for problems of scientific or technological relevance. The benchmarks introduced in this work belong to the family of approaches that can be useful to assess a quantum advantage that is practically relevant to physics.

In the context of variational ground-state algorithms, the V-score can readily be used as a good quality metric to assess quantum advantage. Furthermore, the V-score can be used as an absolute indicator of hardness for Hamiltonians. In this respect, Hamiltonians with large classical V-scores are identified as hard problems that are not yet satisfactorily solved by classical computers and can be targeted by quantum computations. Finally, in the absence of classical verifiability of the quantum solutions, the V-score can benchmark the progress of variational quantum computing-based approaches in solving ground-state problems relevant to physics.

In Fig. 3, we show the V-scores of the classical variational methods that we have analyzed. From these results, we can infer that there is little room for quantum advantage in 1D geometries, where DMRG already achieves V-scores less than 10^{-8}. In higher dimensions, unfrustrated spin models, such as the TFIM or the Heisenberg model on the square lattice, are similarly well approximated by existing classical methods, with V-scores less than 10^{-4}. On the contrary, specific regimes of higher-dimensional frustrated spin models constitute a clear challenge for existing classical methods. For example, the pyrochlore or kagome Heisenberg models typically present V-scores greater than 10^{-2}—substantially higher than their unfrustrated counterparts. A similar scenario emerges for the Hubbard model in two dimensions with V-scores on the order of 10^{-2}. In the specific regimes of interaction strengths, geometries, and frustration identified by large V-scores, these models represent natural targets for variational quantum algorithms. Impurity models with multiple bands also represent an ideal terrain for practical quantum advantage because the ability of classical algorithms to simulate them rapidly degrades upon increasing the number of bands, as Fig. 3 shows for three-band models, and because of their importance for materials science.

We can also provide an early assessment of the V-scores obtained by the type of variational states that can be efficiently prepared on quantum computers. In this respect, it is encouraging to remark that PQCs perform well compared with classical variational methods, as shown in the inset of Fig. 2, at least for the small system sizes that we consider, where PQCs can be classically simulated and ideally optimized. Applications on quantum hardware are much more challenging because of stochastic fluctuations and noise. However, the baseline of ideally optimized PQCs is promising for applications.

IV. DISCUSSION AND OUTLOOK

In this work, we have introduced the V-score, an empirical metric to quantify the absolute accuracy of variational solutions to strongly interacting quantum models. Supplemented with state-of-the-art results obtained by a large variety of numerical methods, this metric allows us to clearly identify models, geometries, and regimes in which existing approaches are currently less accurate. With the introduction of future computational techniques and improved computing architectures, the outcomes of this analysis will naturally evolve in time, revealing in a certifiable manner the continuous improvements happening in the field. In this respect, the dataset presented in this work can be a standardized way of taking snapshots of the evolution of quantum many-body techniques with time.

Besides the importance of these benchmarks for fu-

ture developments in computational techniques based on classical computers, it will be especially interesting to use the V-score to directly measure the impact of quantum computing-based approaches. The hardest classical problem instances identified by large V-scores can be good candidates for studies based on quantum algorithms. In that context, the V-score can be used as a metric to assess progress in quantum variational state preparation, in the absence of classical verifiability.

ACKNOWLEDGMENTS

This paper is dedicated to the memory of Sandro Sorella, a dear friend, esteemed colleague, and co-author of this paper, who passed away before its submission. We acknowledge discussions with A. Sandvik and M. Stoudenmire. We thank L. Reading-Ikkanda (Simons Foundation) for preparing Fig. 1. The Flatiron Institute is a division of the Simons Foundation.

A. Funding

G.C. acknowledges support by the NCCR MARVEL, a National Centre of Competence in Research, funded by the Swiss National Science Foundation (grant no. 205602). D.W. and G.C. acknowledge support from the Swiss National Science Foundation under grant no. 200336. R.R. and G.C. acknowledge support from SEFRI through grant no. MB22.00051 (NEQS - Neural Quantum Simulation). N.A. and T.N. acknowledge support from the European Unions Horizon 2020 research and innovation program (ERC-StG-Neupert-757867-PARATOP) and from the Swiss National Science Foundation (grant PP00P2_176877). J.C. acknowledges support from the Natural Sciences and Engineering Research Council (NSERC), the Shared Hierarchical Academic Research Computing Network (SHARCNET), Compute Canada, and the Canadian Institute for Advanced Research (CIFAR) AI chair program. M.I., Y.N., R.P., and M.S. acknowledge support by MEXT as "Program for Promoting Researches on the Supercomputer Fugaku" (Basic Science for Emergence and Functionality in Quantum Matter – Innovative Strongly Correlated Electron Science by Integration of Fugaku and Frontier Experiments, JPMXP1020200104) together with computational resources of supercomputer Fugaku provided by the RIKEN Center for Computational Science (project ID: hp210163 and hp220166). M.I. acknowledges the support of MEXT KAKENHI, Grant-in-Aid for Transformative Research Areas, grant nos. JP22H05111 and JP22H05114 and JSPS KAKENHI grant no. JP19H00658. M.I. and Y.N. acknowledge support from MEXT as "Program for Promoting Researches on the Supercomputer Fugaku" (grant no. JPMXP1020230411). G.M. acknowledges support from the Swiss National Science Foundation (grant PCEFP2_203455). Y.N. acknowledges support from Grant-in-Aids for Scientific Research (JSPS KAKENHI) (grant nos. JP23H04869, JP23H04519, JP23K03307, and JP21H01041) and JST (grant no. JPMJPF2221). J.M.S. and S.R.W. acknowledge the support of the US National Science Foundation under grant DMR-2110041. L.W. and Q.Y. acknowledge the support of the National Natural Science Foundation of China under grant nos. 92270107, 12188101, T2225018, and T2121001. A.W. acknowledges support by the DFG through the Emmy Noether programme (509755282). Y.Y. thanks the Center for Computational Quantum Physics (CCQ) of the Flatiron Institute, Simons Foundation for support and hospitality. Research at the Perimeter Institute is supported in part by the government of Canada through the Department of Innovation, Science and Economic Development and by the Province of Ontario through the Ministry of Colleges and Universities.

B. Author contributions

D.W. implemented the research infrastructure and curated the data. R.R. designed the final form of the V-score. G.C. designed and coordinated the project, with the help of M.I. and R.R. The main text was written by D.W., R.R., F.V., and G.C. with input from all authors. All authors contributed to the supplementary materials. A detailed record of the authors' contributions to the dataset can be found on the GitHub repository.

C. Competing interests

The authors declare no competing interests.

D. Data and materials availability

The VarBench dataset and the code to analyze the benchmark results and reproduce the figures in this manuscript are available at Zenodo [63]. Every benchmark result either has a reference to a paper reporting it or can be reproduced using the code at Zenodo [64]. The software packages used to reproduce the data include ITensor [65], mVMC [66], and NetKet [67]. Future updates to the dataset can be found at `https://github.com/varbench/varbench` .

Supplementary Materials for "Variational Benchmarks for Quantum Many-Body Problems"

S1. OVERVIEW OF THE MANY-BODY HAMILTONIANS

Here we provide the definitions of the many-body quantum Hamiltonians used for benchmarking purposes.

A. Spin models

The transverse-field Ising model (TFIM) is

$$\hat{H} = J \sum_{\langle i,j \rangle} \hat{\sigma}_i^z \hat{\sigma}_j^z + \Gamma \sum_i \hat{\sigma}_i^x, \tag{S1}$$

where $\langle i,j \rangle$ runs over nearest neighbors, Γ is the transverse field strength, and $\hat{\sigma}^x, \hat{\sigma}^y, \hat{\sigma}^z$ are Pauli matrices. In the dataset we use $J = 1$.

The Heisenberg model is

$$\hat{H} = J \sum_{\langle i,j \rangle, a} \hat{\sigma}_i^a \hat{\sigma}_j^a, \tag{S2}$$

where a runs in $\{x, y, z\}$.

The J_1-J_2 model is

$$\hat{H} = J_1 \sum_{\langle i,j \rangle, a} \hat{\sigma}_i^a \hat{\sigma}_j^a + J_2 \sum_{\langle\langle i,j \rangle\rangle, a} \hat{\sigma}_i^a \hat{\sigma}_j^a, \tag{S3}$$

where $\langle\langle i,j \rangle\rangle$ runs over next-nearest neighbors (diagonals if the lattice is 2D square), and J_2 is the next-nearest neighbor interaction. In the dataset we use $J_1 = 1$.

B. Fermions

The t-V model is

$$\hat{H} = -t \sum_{\langle i,j \rangle} \left(\hat{c}_i^\dagger \hat{c}_j + \hat{c}_j^\dagger \hat{c}_i \right) + V \sum_{\langle i,j \rangle} \hat{n}_i \hat{n}_j, \tag{S4}$$

where V is the Coulomb repulsive interaction strength, and the number of fermions is fixed to N_f. In the dataset we consider $t = 1$.

The Hubbard model is

$$\hat{H} = -t \sum_{\langle i,j \rangle, \sigma} \left(\hat{c}_{i\sigma}^\dagger \hat{c}_{j\sigma} + \hat{c}_{j\sigma}^\dagger \hat{c}_{i\sigma} \right) + U \sum_i \hat{n}_{i\uparrow} \hat{n}_{i\downarrow}, \tag{S5}$$

where U is the on-site interaction strength, and the numbers of fermions are fixed to $N_\uparrow$ and $N_\downarrow$. We only consider the case of $N_\uparrow = N_\downarrow$.

Additionally, we include the generalized Hubbard model with second nearest neighbor hopping t_2 and repulse V_2:

$$\begin{aligned} \hat{H} = & -t_1 \sum_{\langle i,j \rangle, \sigma} \left(\hat{c}_{i\sigma}^\dagger \hat{c}_{j\sigma} + \hat{c}_{j\sigma}^\dagger \hat{c}_{i\sigma} \right) - t_2 \sum_{\langle\langle i,j \rangle\rangle, \sigma} \left(\hat{c}_{i\sigma}^\dagger \hat{c}_{j\sigma} + \hat{c}_{j\sigma}^\dagger \hat{c}_{i\sigma} \right) \\ & + U \sum_i \hat{n}_{i\uparrow} \hat{n}_{i\downarrow} + V_1 \sum_{\langle i,j \rangle, \sigma} \hat{n}_{i\sigma} \hat{n}_{j\sigma} + V_2 \sum_{\langle\langle i,j \rangle\rangle, \sigma} \hat{n}_{i\sigma} \hat{n}_{j\sigma}. \end{aligned} \tag{S6}$$

We use $t_1 = 1$, $t_2 = -0.25$, $V_1 = 1$, $V_2 = 0.5$.

C. Impurity models

A typical Anderson impurity Hamiltonian $\hat{H}_\mathrm{A}$ contains two parts

$$\hat{H}_\mathrm{A} := \hat{H}_\mathrm{loc} + \hat{H}_\mathrm{bath}, \tag{S7}$$

$$\begin{aligned} \hat{H}_\mathrm{loc} := & \sum_{\{\alpha\}} \epsilon^0_{\alpha_1\alpha_2} \hat{d}^\dagger_{\alpha_1} \hat{d}_{\alpha_2} \\ & + \sum_{\{\alpha\}} U_{\alpha_1\alpha_2\alpha_3\alpha_4} \hat{d}^\dagger_{\alpha_1} \hat{d}^\dagger_{\alpha_2} \hat{d}_{\alpha_3} \hat{d}_{\alpha_4}, \end{aligned} \tag{S8}$$

$$\begin{aligned} \hat{H}_\mathrm{bath} := & \sum_{\{\alpha\}} \sum_{l=1}^{N_\mathrm{b}} \epsilon^l_{\alpha_1\alpha_2} \hat{c}^\dagger_{l\alpha_1} \hat{c}_{l\alpha_2} \\ & + \sum_{\{\alpha\}} \sum_{l=1}^{N_\mathrm{b}} \left(\nu^l_{\alpha_1\alpha_2} \hat{d}^\dagger_{\alpha_1} \hat{c}_{l\alpha_2} + \mathrm{h.c.} \right), \end{aligned} \tag{S9}$$

where a locally interacting impurity $\hat{H}_\mathrm{loc}$ is coupled to a non-interacting bath $\hat{H}_\mathrm{bath}$. The indices $\{\alpha\}$ are a collection of quantum numbers denoting the impurity (or the l-th bath site) degrees of freedom of the fermionic creation $\hat{d}^\dagger$ (or $\hat{c}^\dagger_l$) and annihilation $\hat{d}$ (or $\hat{c}_l$) operators, and N_b is the number of bath sites per spin-orbital. The bath parameters $\{\epsilon^l, \nu^l\}$ are connected to the hybridization function as $\Delta(\omega) = \sum_{l=1}^{N_\mathrm{b}} \frac{\nu^l \nu^{l\dagger}}{\omega - \epsilon^l}$, and can be obtained by discretizing the hybridization function $\Delta(\omega)$ on the real frequency axis into N_b equidistant intervals $\{I_l\}$ of size $\Delta\omega$ as

$$\epsilon^l_{\alpha\alpha} = \min I_l + \frac{\Delta\omega}{2}, \tag{S10}$$

$$|\nu^l|^2 = \int_{I_l} \mathrm{d}\omega \left(-\frac{\mathrm{Im}\, \Delta(\omega)}{\pi} \right). \tag{S11}$$

We consider two types of interactions that are frequently encountered in DMFT calculations: the single-band Hubbard interaction

$$\hat{H}_\mathrm{loc} = U \hat{n}_\uparrow \hat{n}_\downarrow + \epsilon^0 \left(\hat{n}_\uparrow + \hat{n}_\downarrow \right), \tag{S12}$$

where $\hat{n}_\sigma = \hat{d}^\dagger_\sigma \hat{d}_\sigma$ is the particle number operator, with $\sigma \in \{\uparrow, \downarrow\}$; the three-band rotationally invariant Kanamori interaction [68]

$$\hat{H}_{\mathrm{loc}} = \hat{H}_{\mathrm{DD}} + \hat{H}_{\mathrm{SF}} + \hat{H}_{\mathrm{PH}}, \tag{S13}$$

$$\begin{aligned}\hat{H}_{\mathrm{DD}} := \; & U \sum_m \hat{n}_{m\uparrow}\hat{n}_{m\downarrow} \\ & + (U-2J) \sum_{m'>m,\sigma} \hat{n}_{m\sigma}\hat{n}_{m'\bar{\sigma}} \\ & + (U-3J) \sum_{m'>m,\sigma} \hat{n}_{m\sigma}\hat{n}_{m'\sigma},\end{aligned} \tag{S14}$$

$$\hat{H}_{\mathrm{SF}} := J \sum_{m'm} \left(\hat{d}^\dagger_{m\uparrow}\hat{d}_{m\downarrow}\hat{d}_{m'\uparrow}\hat{d}^\dagger_{m'\downarrow} + \mathrm{h.c.} \right), \tag{S15}$$

$$\hat{H}_{\mathrm{PH}} := -J \sum_{m'>m} \left(\hat{d}^\dagger_{m\uparrow}\hat{d}^\dagger_{m\downarrow}\hat{d}_{m'\uparrow}\hat{d}_{m'\downarrow} + \mathrm{h.c.} \right), \tag{S16}$$

with $m \in \{1,2,3\}$ being the orbital index, and $\hat{H}_{\mathrm{DD}}$, $\hat{H}_{\mathrm{SF}}$, and $\hat{H}_{\mathrm{PH}}$ denoting the density-density, the spin-flip, and the pair-hopping interactions respectively.

The following models representing a collection of typical solutions in practical DMFT calculations are considered: (**SB-Imp**) single-band Anderson impurity model with a semielliptic spectral function, i.e., $-\frac{1}{\pi}\operatorname{Im}\Delta(\omega) = \frac{2}{\pi D}\sqrt{1-\left(\frac{\omega}{D}\right)^2}$ with D being the half-bandwidth and $U = D$; (**SB-DMFT-MT-HF**) DMFT metal solution of the single band Hubbard model on the Bethe lattice with $U = 2D$ at half-filling $n = 1$ and (**SB-DMFT-MT-AHF**) doped case $n = 0.8$; (**SB-DMFT-MI-HF**) DMFT Mott-insulator solution of the single band Hubbard model on the Bethe lattice with $U = 4D$ at half-filling $n = 1$; three-band models with Kanamori interaction $U = 2.3$ eV and $J = 0.4$ eV that are based on the material-realistic DMFT solutions of the archetypal Hund's metal Sr_2RuO_4 in the t_{2g} subspace (**TB-DMFT-SOC**) with and (**TB-DMFT**) without spin-orbit coupling.

S2. V-SCORE

A. Definition and justification of the V-score for lattice models

Given the observed energy expectation $E = \langle\psi|\hat{H}|\psi\rangle$ and variance $\operatorname{Var} E = \langle\psi|\hat{H}^2|\psi\rangle - \langle\psi|\hat{H}|\psi\rangle^2$ of a variational quantum state $|\psi\rangle$, we want to introduce a function of these two quantities, the V-score, as a metric quantifying how close E is to the ground-state energy. As the exact ground-state energy is unknown in general, the mean energy alone is not enough to characterize the quality of a variational optimization. The energy variance is zero for an eigenstate of the Hamiltonian, thus, also for the ground state. Therefore, assuming that the variational optimization does not converge towards an excited state, we can infer that the V-score should be a monotonic function of $\operatorname{Var} E$.

To start with, we investigate how these two observables scale asymptotically with the number of degrees of freedom N. For any well-defined variational state, E scales linearly with N, as the energy is an extensive thermodynamic quantity. To analyze the scaling of $\operatorname{Var} E$, we evoke the cluster property of the variational state. The Hamiltonian $\hat{H}$ is written as a sum of $N_H = O(N)$ local terms, $\hat{H} = \sum_{i=1}^{N_H} \hat{h}_i$. If the correlations of these local terms satisfy the cluster property

$$\left| \langle \hat{h}_i \hat{h}_j \rangle - \langle \hat{h}_i \rangle \langle \hat{h}_j \rangle \right| \leq \frac{A}{d(i,j)^{D+\epsilon}}, \tag{S17}$$

where $A, \epsilon > 0$, d is a distance function on the lattice, and D is the space dimension, then $\operatorname{Var} E$ scales linearly with N. We can therefore construct a dimensionless number,

$$\frac{N \operatorname{Var} E}{E^2}, \tag{S18}$$

that does not scale with the energy unit or with N asymptotically. Moreover, for variational optimizations converging towards the ground state, $\operatorname{Var} E$ can be expected to scale linearly with the energy difference from the ground state [53–55, 69, 70]. Therefore, the dimensionless number in Eq. (S18) linearly quantifies the energy difference as well.

However, the expression in Eq. (S18) is still prone to a shift of energy $\hat{H} \mapsto \hat{H} + C$, where C is an arbitrary constant. In particular, if we fine-tune C such that the ground-state energy is zero, Eq. (S18) can be expected to scale *inversely* with the energy difference from the ground state. Similarly, if C is such that $E = 0$, Eq. (S18) is ill-defined. To solve this issue, we need to fix a zero point of energy E_∞ in the definition of the V-score:

$$\text{V-score} := \frac{N \operatorname{Var} E}{(E - E_\infty)^2}, \tag{S19}$$

such that the average energy E of a variational ground-state wave function $|\psi\rangle$ is always different from E_∞. In this work we choose E_∞ to be the energy expectation of a random state (sampled uniformly on the unit sphere surface) in the Hilbert (sub)space $\mathcal{H}$, because an optimized variational state typically has lower energy than a random state. E_∞ is also the energy expectation of a thermal state at infinite temperature restricted to $\mathcal{H}$, which can be computed from the trace of $\hat{H}$:

$$E_\infty := \frac{\operatorname{Tr} \hat{H}}{\dim \mathcal{H}}, \tag{S20}$$

where $\dim \mathcal{H}$ is the dimension of the Hilbert (sub)space. For the models we consider in this work, the dimension of the Hilbert space is finite for a finite lattice size, and E_∞ is a finite number.

We now discuss our choices for the number of degrees of freedom N. For unconstrained spin-1/2 Hilbert spaces,

we define N to be equal to the number of lattice sites N_s. For the t-V model with fixed particle number, N is defined to be equal to the particle number N_f, and we have $\dim \mathcal{H} = B(N_\mathrm{s}, N_\mathrm{f})$, where B is the binomial coefficient. For the Hubbard model with fixed particle numbers, N is defined to be equal to the sum of the numbers of spin up and down fermions, $N = N_\uparrow + N_\downarrow$, while $\dim \mathcal{H} = B(N_\mathrm{s}, N_\uparrow)\, B(N_\mathrm{s}, N_\downarrow)$. We remark that the V-score can also be applied to estimations of the lowest-energy excited states in symmetry sectors different from the ground state symmetry sector.

This choice of energy shift E_∞ supposes that the model contains only relevant low-energy degrees of freedom. Actually, if we add many high-energy states, they will have the effect of artificially raising E_∞ without contributing much to the ground state. Therefore, for some models we do not consider in this work (e.g. bosonic models or quantum chemical models), a cutoff on the relevant energy scale must be set into place in order to use this definition of E_∞. An alternative strategy to define E_∞ for these models would be to use the mean-field energy, which is not affected by the problem of having high-energy states. However, it is generally not straightforward to compute the mean-field energy as it is in itself an NP-hard problem [71], there exist many variants of mean-field theory, and it could make weak-coupling calculations beyond mean-field theory artificially hard. Impurity models, which we have introduced in Sec. S1 C, require an adapted definition of E_∞, see Sec. S2 E.

B. Calculation of E_∞ for lattice models

1. Analytical formulae for specific models

For quantum spin models, we have $E_\infty = 0$ when the Hamiltonian is written as a sum of Pauli strings with no term proportional to the identity operator, as all Pauli matrices are traceless. For the spinless t-V model with fixed particle number, only the diagonal term $V \sum_{\langle i,j \rangle} \hat{n}_i \hat{n}_j$ contributes, and we have

$$E_\infty = \frac{V|\mathcal{E}|N_\mathrm{f}(N_\mathrm{f}-1)}{N_\mathrm{s}(N_\mathrm{s}-1)}, \tag{S21}$$

where $|\mathcal{E}|$ is the number of nearest neighbor bonds. For the Hubbard model with fixed particle numbers, only the diagonal term $U \sum_i \hat{n}_{i\uparrow}\hat{n}_{i\downarrow}$ contributes, and we have

$$E_\infty = \frac{U N_\uparrow N_\downarrow}{N_\mathrm{s}}. \tag{S22}$$

Apart from fixing the number of fermions, in this work we do not consider symmetries of the Hamiltonians when calculating E_∞.

2. General case

It is generally efficient to get a numerical estimate of E_∞ with stochastic methods. For simplicity, we limit our discussion to spinless fermions and a short-range translation-invariant Hamiltonian $\hat{H}$. Considering a Hilbert space $\mathcal{H}$ with fixed particle number N_f and lattice size N_s, we estimate E_∞ by sampling uniformly a bit string $x = (x_1, \ldots, x_{N_\mathrm{s}})$ with $x_i \in \{0, 1\}$ and the constraint $\sum_i x_i = N_\mathrm{f}$, then taking the average

$$E_\infty = \left\langle \langle x|\hat{H}|x\rangle \right\rangle_{x \sim \mathcal{U}_{N_\mathrm{s},N_\mathrm{f}}(x)}, \tag{S23}$$

where $|x\rangle$ is an element of the Fock basis and $\mathcal{U}_{N_\mathrm{s},N_\mathrm{f}}(x)$ is the aforementioned uniform distribution. The variance of the estimator of E_∞ can be written as

$$\operatorname{Var} E_\infty = \left\langle \langle x|\hat{H}|x\rangle^2 \right\rangle_{x \sim \mathcal{U}_{N_\mathrm{s},N_\mathrm{f}}(x)} - E_\infty^2. \tag{S24}$$

Comparing to the physical variance of the energy at infinite temperature

$$\operatorname{Var}_{\mathcal{H}} \hat{H} := \frac{\operatorname{Tr} \hat{H}^2}{\dim \mathcal{H}} - E_\infty^2, \tag{S25}$$

which can be computed by the same sampling method:

$$\operatorname{Var}_{\mathcal{H}} \hat{H} = \left\langle \langle x|\hat{H}^2|x\rangle \right\rangle_{x \sim \mathcal{U}_{N_\mathrm{s},N_\mathrm{f}}(x)} - E_\infty^2, \tag{S26}$$

we have $\operatorname{Var} E_\infty \leq \operatorname{Var}_{\mathcal{H}} \hat{H}$. As $\hat{H}$ is short-range and translation invariant, and as $|x\rangle$ satisfies the cluster property being a product state, we have the scalings

$$E_\infty = O(N), \quad \operatorname{Var}_{\mathcal{H}} \hat{H} = O(N), \tag{S27}$$

which implies that the one-sample stochastic relative error on E_∞ vanishes with increasing system size:

$$\frac{\sqrt{\operatorname{Var} E_\infty}}{|E_\infty|} \leq \frac{\sqrt{\operatorname{Var}_{\mathcal{H}} \hat{H}}}{|E_\infty|} = O\left(\frac{1}{\sqrt{N}}\right). \tag{S28}$$

This shows that, in the thermodynamic limit $N \to \infty$, even just one sample is enough to estimate E_∞. Moreover, the calculation of $\langle x|\hat{H}|x\rangle$ can be done in a computational time increasing linearly with the number of lattice sites. Therefore, we conclude that the statistical procedure we discussed is efficient.

C. Bounds on the V-score

We now consider bounds on the ratio of the V-score and the energy relative error. The lower bound is obtained when the variational state exactly coincides with an excited state, which is therefore zero. In order to prove an upper bound, we maximize $\operatorname{Var} E$ given a fixed mean energy E. When the spectrum is bounded from above, e.g., for finite systems, $\operatorname{Var} E$ is maximized, at fixed average energy E, when the variational state $|\psi\rangle$ is a linear

combination of the ground state and the maximal energy state

$$|\psi\rangle = \sqrt{1-I}\,|E_0\rangle + \sqrt{I}\,|E_M\rangle\,, \tag{S29}$$

where $E_0, |E_0\rangle$ and $E_M, |E_M\rangle$ are the minimal and the maximal eigenvalue-eigenvector pairs respectively, $I \in [0,1]$ is an interpolation parameter such that $E = (1-I)E_0 + IE_M$, and we have assumed that both $\hat{H}$ and $-\hat{H}$ are non-degenerate for simplicity of notation. The energy variance of $|\psi\rangle$ is equal to $(E_M - E)(E - E_0)$. When we express the variance in terms of the V-score, we reach the following bound for the ratio of the V-score and the relative energy error:

$$0 \leq \frac{\text{V-score}}{(E-E_0)/(E_\infty - E_0)} \leq N\frac{(E_\infty - E_0)(E_M - E)}{(E_\infty - E)^2}. \tag{S30}$$

Eq. (S30) shows that, while the V-score can arbitrarily underestimate the difference from the ground-state energy, it cannot arbitrarily overestimate it. We remark that the upper bound is linear in system size, as the aforementioned variational state does not respect the cluster property for energy correlations.

D. Scaling of the V-score in the limit of vanishing ground state infidelity

In this section, we adapt the argument from Ref. [70] to show the linear scaling of the V-score with the energy relative error in the limit where the ground state infidelity goes to zero. We remark that this result is not directly applicable to large system sizes as the ground state infidelity is expected to grow exponentially with system size for a fixed energy relative error.

We consider a variational state parameterized by a control parameter $c \geq 0$:

$$|\psi(c)\rangle := \sum_i \psi_i(c)\,|E_i\rangle\,, \tag{S31}$$

where $\psi_i(c) \in \mathbb{C}$ and $|E_i\rangle$ are energy eigenstates with eigenvalue E_i. We suppose $E_0 \leq E_i$, and we define the set of ground state indices as $G := \{i \mid E_i = E_0\}$. We introduce the ground state infidelity

$$I(c) := \frac{\sum_{i\notin G}|\psi_i(c)|^2}{\sum_i |\psi_i(c)|^2}. \tag{S32}$$

We suppose that when the control parameter c goes to infinity, the ground state infidelity goes to zero:

$$\lim_{c\to\infty} I(c) = 0. \tag{S33}$$

We also suppose that $I(c) \neq 0$ for every c.

In the following, we determine the scaling of the V-score when $c \to \infty$. For an operator $\hat{O}$, we define

$$\langle \hat{O}\rangle_c := \frac{\langle\psi(c)|\hat{O}|\psi(c)\rangle}{\langle\psi(c)|\psi(c)\rangle}, \tag{S34}$$

$$\langle \hat{O}\rangle_{\text{ex},c} := \frac{\sum_{i\notin G}|\psi_i(c)|^2\,\langle E_i|\hat{O}|E_i\rangle}{\sum_{i\notin G}|\psi_i(c)|^2}. \tag{S35}$$

Then we have

$$\langle \hat{H}\rangle_c - E_0 = (\,\langle \hat{H}\rangle_{\text{ex},c} - E_0)I(c). \tag{S36}$$

We also define

$$\sigma_c^2 := \langle \hat{H}^2\rangle_c - \langle \hat{H}\rangle_c^2 \tag{S37}$$

$$= \langle(\hat{H}-E_0)^2\rangle_c - (\,\langle \hat{H}\rangle_c - E_0)^2 \tag{S38}$$

$$= \langle(\hat{H}-E_0)^2\rangle_{\text{ex},c}\; I(c) - \big((\,\langle \hat{H}\rangle_{\text{ex},c} - E_0)I(c)\big)^2 \tag{S39}$$

$$= \langle(\hat{H}-E_0)^2\rangle_{\text{ex},c}\; I(c)\Big(1 + O\big(I(c)\big)\Big), \tag{S40}$$

where we used $0 < (\,\langle \hat{H}\rangle_{\text{ex},c} - E_0)^2 \leq \langle(\hat{H}-E_0)^2\rangle_{\text{ex},c}$. We then write

$$\lim_{c\to\infty} \frac{\text{V-score}}{(\,\langle \hat{H}\rangle_c - E_0)/(E_\infty - E_0)}$$

$$= \lim_{c\to\infty} \frac{N\sigma_c^2(E_\infty - E_0)}{(\,\langle \hat{H}\rangle_c - E_0)(E_\infty - \langle \hat{H}\rangle_c)^2} \tag{S41}$$

$$= \lim_{c\to\infty} \frac{N\,\langle(\hat{H}-E_0)^2\rangle_{\text{ex},c}(E_\infty - E_0)}{(\,\langle \hat{H}\rangle_{\text{ex},c} - E_0)(E_\infty - \langle \hat{H}\rangle_c)^2} \tag{S42}$$

$$= \lim_{c\to\infty} \frac{N\,\langle(\hat{H}-E_0)^2\rangle_{\text{ex},c}}{(\,\langle \hat{H}\rangle_{\text{ex},c} - E_0)(E_\infty - E_0)}. \tag{S43}$$

If we suppose that the following limits exist:

$$\sigma_{\text{ex}}^2 := \lim_{c\to\infty}\big(\,\langle \hat{H}^2\rangle_{\text{ex},c} - \langle \hat{H}\rangle_{\text{ex},c}^2\big), \tag{S44}$$

$$E_{\text{ex}} := \lim_{c\to\infty}\,\langle \hat{H}\rangle_{\text{ex},c}\,, \tag{S45}$$

then we have

$$\lim_{c\to\infty} \frac{\text{V-score}}{(\,\langle \hat{H}\rangle_c - E_0)/(E_\infty - E_0)} = \frac{N\big(\sigma_{\text{ex}}^2 + (E_{\text{ex}} - E_0)^2\big)}{(E_{\text{ex}} - E_0)(E_\infty - E_0)}. \tag{S46}$$

This limit is non-singular because

$$E_0 + \Delta \leq E_{\text{ex}} \leq E_M, \tag{S47}$$

$$0 \leq \sigma_{\text{ex}}^2 \leq \frac{1}{4}(E_M - E_0 - \Delta)^2, \tag{S48}$$

where Δ is the gap between the ground state subspace and the rest of the spectrum, and E_M is the maximal energy. If we suppose further that $E_{\text{ex}} = E_0 + \Delta$ and $\sigma_{\text{ex}}^2 = 0$, which is valid, e.g., for two-level systems or for a thermal state at temperature $T = 1/c$, then we have

$$\lim_{c\to\infty} \frac{\text{V-score}}{(\,\langle \hat{H}\rangle_c - E_0)/(E_\infty - E_0)} = \frac{N\Delta}{E_\infty - E_0}. \tag{S49}$$

E. E_∞ for impurity models

As mentioned in Sec. S2 A, when applied to impurity models, the definition of E_∞ should be modified such that $|x\rangle$ in Eq. (S23) is restricted to the low-energy subspace. For impurity models, there is a natural way to filter out high-energy states that do not contribute to the ground state. Noticing that the fast convergence of density of bath sites with negative (positive) on-site energy to be fully occupied (empty), $|x\rangle$ can be restricted to the low-energy subspace as $|x\rangle = |\mathrm{o}\rangle \otimes |\mathrm{e}\rangle \otimes |i\rangle$, where $|\mathrm{o}\rangle$ ($|\mathrm{e}\rangle$) denotes the product state of completely occupied (empty) bath sites, $|i\rangle \in \mathcal{H}_{\mathrm{imp}} \otimes \mathcal{H}_{\mathrm{a}}$ is a product state belonging to the Hilbert space composed by the impurity and the active bath site. The active bath site is determined by the one with the smallest absolute on-site energy. Following Eq. (S20), E_∞ is modified for impurity models as

$$E_\infty = \frac{\mathrm{Tr}\,\hat{\rho}_{\mathrm{b}} \otimes \hat{\rho}_i \hat{H}_{\mathrm{A}}}{\dim \mathcal{H}_{\mathrm{imp}}}, \tag{S50}$$

with $\hat{\rho}_{\mathrm{b}} = |\mathrm{e}\rangle\,|\mathrm{o}\rangle\,\langle\mathrm{o}|\,\langle\mathrm{e}|$ and $\hat{\rho}_i = |i\rangle\,\langle i|$. For an impurity model with N_{e} electrons and N_{o} occupied bath sites, we evaluate E_∞ using the stochastic sampling method presented in Sec. S2 B 2 with the constraint that each sampled state $|i\rangle$ has a fixed particle number of $N_{\mathrm{b}} - N_{\mathrm{o}}$.

S3. OVERVIEW OF THE NUMERICAL METHODS

A. Exact diagonalization

The quantum many-body problem can be solved numerically to arbitrary precision on small lattice sizes using exact diagonalization without any approximation. Typically, an iterative algorithm like the Lanczos algorithm [72] is used to solve for the eigenvalues and eigenvectors of the static Schrödinger equation

$$\hat{H}\,|\psi\rangle = E\,|\psi\rangle\,. \tag{S51}$$

The Lanczos algorithm has proven to be a reliable tool for computing ground-state energies up to machine precision. Hence, data retrieved from exact diagonalization can be considered exact. To achieve the currently largest system sizes, both memory and CPU time limitations must be dealt with. To avoid memory bottlenecks, the matrix-vector multiplication operations in the Lanczos algorithm are performed "on-the-fly", i.e., without storing the Hamiltonian matrix, neither in full nor in sparse format, but by implementing a matrix-vector multiplication function. Moreover, using Hamiltonian symmetries allows block-diagonalization to reduce the memory footprint further. Using a symmetry-adapted basis requires efficient algorithms to evaluate matrix elements in this basis, e.g. sublattice coding algorithms. Finally, also large-scale parallelization for distributed memory computers is required, which poses challenges in managing load balancing. The necessary techniques required to achieve the currently largest system sizes are explained in detail in Ref. [73].

B. Tensor networks

1. *Matrix product states and density matrix renormalization group*

The density matrix renormalization group (DMRG) is a variational technique, first introduced by White in 1992 to accurately describe the ground-state properties of one-dimensional (1D) quantum lattices [4]. While exact diagonalization methods operate in an exponentially large basis, DMRG works with the degrees of freedom tied to a few sites at a time. At the heart of the DMRG algorithm is the matrix product state (MPS) ansatz [74], which represents the many-indexed wave function as a chain of tensors, one for each site, with links connecting the sites in a 1D layout. DMRG is an extremely efficient procedure for optimizing the coefficients of the MPS. The required *bond dimension*, i.e. dimension of the indices linking the tensors, is determined by the degree of entanglement in the state being described. In the calculations presented here, which do not utilize parallelization beyond a single node, bond dimensions are limited to about $10,000$. The required bond dimension for the ground state of a model system is thus tied to the area law of entanglement, which states that the entanglement entropy of a bipartition of a system varies as the size of the boundary rather than the volume of either subsystem [75]. DMRG is ideal for gapped 1D systems, where the area law implies that the bond dimension (for a fixed error) is independent of the length of the chain. For a gapless chain, there is a slow logarithmic growth of the bond dimension with length; nevertheless, spin-1/2 chains with lengths in the thousands are still rather easy on a laptop.

For two-dimensional (2D) clusters, the area law implies that the bond dimension is independent of the length but grows exponentially with the width. Despite the exponential, the general efficiency and robustness of DMRG makes it one of the most powerful and versatile methods for studying many 2D lattice models [76]. Success has required developing a variety of specific techniques and "tricks"; for example, the standard two-site DMRG algorithm gives a rough measure of the error associated with using a finite bond dimension, called the truncation error. Even though the truncation error is a crude approximation of the true error, protocols for extrapolating the truncation error to zero can give greatly improved energy estimates and approximate errors on those estimates. Another trick is to rely as much as possible on measurements of local quantities rather than long-distance correlation functions, which are determined by DMRG much less accurately. Similar information to correlation functions can be obtained from a local perturbation of the system followed by the measurement of local

quantities away from the perturbation by following the linear response theory. Note that perturbations, e.g. a global antiferromagnetic field breaking the SU(2) symmetry of an antiferromagnet, can sometimes reduce entanglement, making the calculation easier.

The energy variance and the related V-score presented in this paper deviates from the usual DMRG protocols, but they provide a natural way to compare different algorithms. The calculation of the variance of the energy is straightforward in DMRG, but it is much more costly to compute than the truncation error. Aside from the cost, extrapolations to zero variance are a potential improvement over truncation error extrapolations. To mitigate the cost, a two-site variance has been introduced, which is potentially more robust than the truncation error, and can be used in the one-site DMRG algorithm [77]. The two-site variance has a cost similar to the rest of a two-site DMRG sweep. Here, however, we only report full-variance calculations and forgo extrapolations to allow comparisons with other methods.

A key limitation in calculating the full variance is that one can run out of memory. We use a matrix product operator (MPO) form for the Hamiltonian, and calculating the variance can be done using the square of this MPO — except that this tends to require large amounts of memory. A useful trick is to break the Hamilonian MPO into k pieces, each with smaller bond dimension. Then one needs to sum k^2 terms which may be calculated in parallel.

2. *Fork tensor product states for impurity models*

The many-body wave function of the impurity models is parameterized by the Fork tensor product states (FTPS) [78]. Compared to MPS, which has a chain geometry, FTPS avoids the artificial long-range interaction that is detrimental to MPS by explicitly separating bath degrees of freedom belonging to different bands. Hence, the FTPS is expected to efficiently capture the entanglement structure of multiorbital problems. Furthermore, the bipartite nature of the FTPS makes it possible to extend the efficient DMRG algorithm developed for MPS to FTPS. The ground state is found in our calculations by the single-site DMRG algorithm supplied with a subspace expansion [79]. Except for the three-band model with spin-orbital coupling, which has only U(1) symmetry for the charge sector, all calculations are performed under the global U(1) symmetries for the charge and spin sectors. We first perform 30 DMRG sweeps without symmetries with a relatively low bond dimension to find the correct charge and spin sector. Then, the charge and the spin quantum numbers are fixed to be $\langle \hat{N}_\alpha \rangle$ and $\langle \hat{S}^z_\alpha \rangle$ for each spin-orbital α, respectively. Finally, the ground state is found by another 60 DMRG sweeps in the fixed quantum number sector with a maximum bond dimension $m = 100$ for single-band models and $m = 350$ for three-band models.

C. Variational Monte Carlo

Variational Monte Carlo (VMC) methods are a family of computational methods that do not suffer from the sign problem and whose computational cost is tractable. In particular, VMC combines a variational encoding of the wave function, to reduce the memory complexity associated with *storing* the wave function, with Monte Carlo techniques which lower the computational complexity. This approach was originally introduced to treat models in the continuum such as the helium atom or the electron gas [10, 11], and later adapted to find the ground states of lattice systems as those discussed in the main text of this manuscript [31, 80]. Since then, more sophisticated methods have been proposed on various approximation levels [66, 81–85]. Nowadays, several open-source software implementations of those algorithms are available, such as mVMC [66] and NetKet [67].

In variational approaches such as VMC, a quantum state $|\psi_\theta\rangle$ is encoded into a parameterized function $\psi_\theta(x)$ usually referred to as the *variational ansatz*. This function takes vectors x from a certain basis of the Hilbert space as input and outputs the complex wave function amplitudes such that

$$|\psi_\theta\rangle := \sum_x \psi_\theta(x)\,|x\rangle\,. \tag{S52}$$

If the function ψ_θ is fixed, then a quantum state is uniquely identified by the vector of variational parameters θ. Variational ansatzes are usually chosen such that the number of variational parameters θ increases at most only algebraically with the number of degrees of freedom in the system (in most cases linearly or quadratically). This provides an exponential reduction in the memory complexity with respect to storing the full wave function, which is a vector in an exponentially large Hilbert space.

Even when the wave function is encoded into a small vector of a few variational parameters θ, the expectation value of physical quantities involves two sums over the entire basis of the Hilbert space, leading to exponential computational complexity. To work around this issue, expectation values are computed through statistical averages of stochastic estimators. Instead of summing over the whole Hilbert space, only a few basis elements are selected through rigorous Monte Carlo sampling techniques such as Markov chain Monte Carlo (MCMC). In practice, the expectation value of the energy over the state, $E_\theta = \langle\psi_\theta|\hat{H}|\psi_\theta\rangle$, is approximated by the following expectation value [2]:

$$E_\theta \approx \mathbb{E}_{x\sim P_\theta(x)}\left[E^{\mathrm{loc}}_\theta(x)\right], \tag{S53}$$

where the samples $\{x\}$ are distributed according to the Born probability

$$P_\theta(x) := \frac{|\psi_\theta(x)|^2}{\langle\psi_\theta|\psi_\theta\rangle}, \tag{S54}$$

and

$$E_\theta^{\mathrm{loc}}(x) := \sum_{x'} \langle x|\hat{H}|x'\rangle \, \frac{\psi_\theta(x')}{\psi_\theta(x)} \tag{S55}$$

is called the *local energy*. Assuming that the Hamiltonian $\hat{H}$ has only a polynomial number of non-zero entries $\langle x|\hat{H}|x'\rangle$ for every row $\langle x|$, the computational complexity of this local estimator is algebraic.

In general, when using VMC to determine the ground state of a Hamiltonian $\hat{H}$ one updates the variational parameters θ according to the conjugate gradient of the energy [86, 87]. If $\psi_\theta(x)$ has real-valued parameters or if it is holomorphic, the gradient is estimated using the formula

$$\frac{\partial E_\theta}{\partial \theta^*} \approx \mathbb{E}_{x\sim P_\theta(x)} \left[\left(E_\theta^{\mathrm{loc}}(x) - E_\theta\right) \frac{\partial \log \psi_\theta^*(x)}{\partial \theta^*} \right], \tag{S56}$$

where $\frac{\partial \log \psi_\theta(x)}{\partial \theta}$ is computed by automatic differentiation (AD) [88] in modern softwares. This estimator also has the useful property that $\frac{\partial E_\theta}{\partial \theta^*} = 0$ when the variational state hits an eigenstate of the Hamiltonian such as the ground state, meaning that the optimization will stop if convergence is reached. For non-holomorphic ansatzes with complex-valued parameters, the formula gains a second term $\propto \frac{\partial \log \psi_\theta^*(x)}{\partial \theta}$.

The simplest first-order optimization scheme is the stochastic gradient descent (SGD):

$$\theta \leftarrow \theta - \eta \frac{\partial E_\theta}{\partial \theta^*}, \tag{S57}$$

where η is a sufficiently small positive number called the learning rate. More elaborate optimization schemes that involve accumulating momentums of the gradient, such as Adam [89], can also be employed.

A more advanced optimization method that leverages information about the local geometry of the variational manifold [90, 91] is often used. This method, known as natural gradient [86] in the machine learning literature and as stochastic reconfiguration (SR) [87] in the VMC literature, approximates the imaginary-time evolution by $\exp(-\tau\hat{H})$ for the Hamiltonian $\hat{H}$ with sufficiently large τ to reach the ground state.

After the stochastic optimization of the energy, the variational wave function can be improved further by applying the Lanczos operator $\hat{\mathbb{1}} + \alpha\hat{H}$ to $|\psi_\theta\rangle$ once or twice with the variational parameter α and is employed in some cases of the present benchmark.

1. *Physically motivated ansatzes*

In fermionic systems, which are one of the main focuses of this paper, the variational ansatz takes the form $|\psi\rangle = \hat{\mathcal{C}}\,|\phi\rangle$, where $|\phi\rangle$ is an uncorrelated fermionic state and $\hat{\mathcal{C}}$ denotes a generic correlator. Centering the discussion on systems of spinful electrons, the uncorrelated state $|\phi\rangle$ is given either by a Slater determinant [81, 82, 92] or by a pair-product (PP) wave function (also known as geminal wave function) [83, 84, 93, 94]. The variational PP wave function is defined by

$$\phi(x) := \langle x|\phi\rangle = \left\langle x\right| \left(\sum_{ij\sigma\sigma'} f_{ij}^{\sigma\sigma'} \hat{c}_{i\sigma}^\dagger \hat{c}_{j\sigma'}^\dagger \right)^{\frac{N_\mathrm{e}}{2}} \left|\phi_0\right\rangle, \tag{S58}$$

where $|\phi_0\rangle$ is the vacuum state, N_e is the number of electrons and the pair amplitudes $f_{ij}^{\sigma\sigma'}$ form a matrix of $4N^2$ variational parameters on N sites, which depend on the spatial coordinates i and j, and the spins σ and σ'. The spin-dependence of $f_{ij}^{\sigma\sigma'}$ can be chosen such that either singlet or triplet components appear in the wave function. The PP state, in general, contains Slater determinants as a subset and offers higher flexibility and accuracy. It can accommodate the Hartree–Fock–Bogoliubov type wave function with magnetic, charge, and superconducting orders [84], and paramagnetic metals as well, in a unified and flexible fashion.

The correlator $\hat{\mathcal{C}}$ operates to save the exponentially large number of basis functions, and frequently used examples in the present benchmark are given by introducing and applying various physically adequate operators in $\hat{\mathcal{C}}$, such as

$$\hat{\mathcal{C}} := \hat{\mathcal{L}}^S \hat{\mathcal{L}}^q \hat{\mathcal{P}}^{\mathrm{G}} \hat{\mathcal{P}}^{\mathrm{J_c}} \hat{\mathcal{P}}^{\mathrm{J_s}}. \tag{S59}$$

The correlation factors

$$\hat{\mathcal{P}}^{\mathrm{G}} := \exp\left(\sum_i g_i^{\mathrm{G}} \hat{n}_{i\uparrow}\hat{n}_{i\downarrow} \right), \tag{S60}$$

$$\hat{\mathcal{P}}^{\mathrm{J_c}} := \exp\left(\sum_{i<j} g_{ij}^{\mathrm{J_c}} \hat{n}_i \hat{n}_j \right), \tag{S61}$$

$$\hat{\mathcal{P}}^{\mathrm{J_s}} := \exp\left(\sum_{i<j} g_{ij}^{\mathrm{J_s}} \hat{S}_i^z \hat{S}_j^z \right) \tag{S62}$$

are the Gutzwiller factor [42], the long-range Jastrow correlation factor [95, 96], and the long-range spin Jastrow correlation factor [97], with spatially dependent variational parameters g_i^{G}, $g_{ij}^{\mathrm{J_c}}$, and $g_{ij}^{\mathrm{J_s}}$, respectively; $\hat{S}_i^z = \frac{1}{2}\left(\hat{n}_{i\uparrow} - \hat{n}_{i\downarrow}\right)$ is the z component of the local spin operator. In practice, the translational symmetry is often imposed on the Gutzwiller and the Jastrow factors in order to reduce the number of independent variational parameters.

Furthermore, since on finite sizes the exact ground state possesses all the symmetries of the Hamiltonian, while the variational state $|\phi\rangle$ may break them, the quantum number projections $\hat{\mathcal{L}}$ are also considered to restore the symmetries. In Eq. (S59), $\hat{\mathcal{L}}^S$ and $\hat{\mathcal{L}}^q$ are examples of projection operators which enforce the fixed total spin S and momentum q, respectively [84], and $\hat{\mathcal{L}}^q$ is used in the case of periodic boundary conditions.

In the above example, the variational parameters are contained both in the correlator $\hat{\mathcal{C}}$ (i.e. $g_i^{\mathrm{G}}, g_{ij}^{\mathrm{J_c}}$, and $g_{ij}^{\mathrm{J_s}}$) and in the uncorrelated state $|\phi\rangle$ (i.e. $f_{ij}^{\sigma\sigma'}$). They are optimized to better approximate the ground state by lowering the numerically evaluated energy. By fully optimizing $f_{ij}^{\sigma\sigma'}$ including the long-range part, it is known that not only the symmetry-broken Hartree–Fock–Bogoliubov states and simple single-particle noninteracting Slater determinants, but also correlated metallic states such as Tomonaga–Luttinger liquid can be represented [98, 99]. In order to lower the computational cost by reducing the number of independent variational parameters, one can assume that $f_{ij}^{\sigma\sigma'}$ has a sublattice structure such that it depends only on the relative position vector $\boldsymbol{r}_i - \boldsymbol{r}_j$ and a sublattice index of site j which is denoted as $\eta(j)$. Then one can rewrite it as $f_{\eta(j)}^{\sigma\sigma'}(\boldsymbol{r}_i - \boldsymbol{r}_j)$. In many cases, one can employ variational states which do not break any translational symmetry and assume a fully translational invariance (1×1 sublattice structure, in the case of 2D models), restricting to singlet pairings only (i.e. $f_{ij}^{\sigma\sigma'} = -f_{ij}^{\sigma',\sigma}(1-\delta_{\sigma\sigma'})$). Antiferromagnetic states can be included by extending the sublattice structure to a 2×2 (or larger) unit cell. In most of the present benchmark studies, we do not restrict the sublattice size; namely, the sublattice size is the same as the full lattice. In some cases, $\hat{\mathcal{P}}^{\mathrm{J_s}}$ and $\hat{\mathcal{L}}^S$ are omitted. One can also improve the wave function by implementing dependence on the local density of $|x\rangle$ (see Ref. [100]). The advantage of using this scheme together with Eqs. (S58) and (S59) is that the quantum entanglement beyond the area law can be taken into account.

In an alternative approach, the parameters $f_{ij}^{\sigma\sigma'}$ for $|\phi\rangle$ are obtained starting from an auxiliary non-interacting Hamiltonian, containing hopping and pairing amplitudes [83, 101, 102]. The most general form is given by

$$\hat{H}_{\mathrm{aux}} := \sum_{ij\sigma\sigma'} \left(\chi_{ij}^{\sigma\sigma'} \hat{c}_{i\sigma}^\dagger \hat{c}_{j\sigma'} + \Delta_{ij}^{\sigma\sigma'} \hat{c}_{i\sigma}^\dagger \hat{c}_{j\sigma'}^\dagger + \mathrm{h.c.} \right), \quad \text{(S63)}$$

where $\chi_{ij}^{\sigma\sigma'}$ and $\Delta_{ij}^{\sigma\sigma'}$ are hopping and pairing terms, respectively. Then the ground state of $\hat{H}_{\mathrm{aux}}$ is a PP wave function in Eq. (S58), with the pair amplitudes $f_{ij}^{\sigma\sigma'}$ determined by the parameters $\chi_{ij}^{\sigma\sigma'}$ and $\Delta_{ij}^{\sigma\sigma'}$. The advantage of this approach is two-fold. First, long-range pair amplitudes can be obtained within a short-range parametrization of $\chi_{ij}^{\sigma\sigma'}$ and $\Delta_{ij}^{\sigma\sigma'}$, thus avoiding delicate optimizations of the long-range tails of $f_{ij}^{\sigma\sigma'}$. Second, symmetries may be imposed directly on the auxiliary Hamiltonian, thus avoiding including further projectors, e.g., $\hat{\mathcal{L}}^S$ and $\hat{\mathcal{L}}^q$. For example, by restricting to "singlet" hoppings ($\chi_{ij}^{\sigma\sigma'} = \delta_{\sigma\sigma'}\chi_{ij}$ with $\chi_{ij} = \chi_{ji}^*$) and pairings ($\Delta_{ij}^{\sigma\sigma'} = (1-\delta_{\sigma\sigma'})\Delta_{ij}$ with $\Delta_{ij} = \Delta_{ji}$), the PP wave function is already a singlet, while by imposing translational symmetry in the dependence of these parameters on the lattice sites, the PP state is translationally symmetric automatically. Within this framework, the magnetic order can be described by including on-site hoppings which mimic the presence of a (site-dependent) Zeeman field $\boldsymbol{h}_i = (h_i^x, h_i^y, h_i^z)$, with $h_i^z = \frac{1}{2}(\chi_{ii}^{\uparrow\uparrow} - \chi_{ii}^{\downarrow\downarrow})$, $h_i^x = \chi_{ii}^{\uparrow\downarrow} + \chi_{ii}^{\downarrow\uparrow}$, and $h_i^y = \mathrm{i}(\chi_{ii}^{\uparrow\downarrow} - \chi_{ii}^{\downarrow\uparrow})$ (see for example Refs. [103, 104]).

The general parametrization in Eq. (S63) allows us to include site-dependent terms to describe charge and/or spin inhomogeneities (i.e. stripes) [102, 105]. Jastrow factors, when included, are chosen to have translational and rotational symmetries. Furthermore, backflow correlations can be included, redefining the orbitals of Eq. (S63) on the basis of the many-body electronic configuration [106, 107]. Analogous variational wave functions can be employed to study localized spins systems, e.g., frustrated Heisenberg models on different lattice geometries [83, 101, 103, 108, 109].

Notice that in the context of spin models, the Gutzwiller factor is replaced by the Gutzwiller projector $\hat{\mathcal{P}}_\infty^{\mathrm{G}} := \prod_i \hat{n}_i(2-\hat{n}_i)$, which enforces single occupancy on each lattice site, on top of a fermionic uncorrelated PP state $|\phi\rangle$. In addition, analogously to the electronic case previously discussed, a spin Jastrow factor $\hat{\mathcal{P}}^{\mathrm{J_s}}$ can be included, and quantum number projectors $\hat{\mathcal{L}}^S$, $\hat{\mathcal{L}}^q$ can be applied to enforce lattice symmetries. Also, in this case, the $f_{ij}^{\sigma\sigma'}$ pair amplitudes of the PP state can be assumed to be direct variational parameters (with certain symmetry constraints) or defined through an auxiliary Hamiltonian like the one of Eq. (S63).

In addition, we mention that a possible bias inevitable in the original wave function can be progressively removed by adding a correlator $\hat{\mathcal{M}}$ implemented as a restricted Boltzmann machine (RBM) [12, 85] on top of the variational state, i.e., $|\psi\rangle = \hat{\mathcal{C}}\hat{\mathcal{M}}\,|\phi\rangle$. In this case, Jastrow factors can be omitted to save computational cost. We refer the reader to the following subsection for more details about the RBM. Recently, the VMC approach with the RBM implementation has been successfully applied to reveal the nature of the quantum spin liquids [100, 110, 111].

2. Neural quantum states

Neural networks are structured nonlinear functions that can universally approximate any well-behaved function. Motivated by their recent success of efficiently representing probability distributions in machine learning tasks, they have also been used as variational ansatzes to represent quantum states. They typically have more parameters and higher expressive power than traditional ansatzes. The first proposed neural quantum state (NQS) is a restricted Boltzmann machine (RBM) [12] applied first to quantum spin Hamiltonians and then extended to fermionic systems [85] as is introduced in the last subsection. The RBM is also equivalent to a multilayer perceptron (MLP) with two layers. The finite temperature path integral formalism was shown to exactly

map to a deep Boltzmann machine (DBM), where the Trotter–Suzuki layers correspond to the multiple hidden layers in the DBM [112]. The correlator $\hat{\mathcal{M}}$ represented by NQS can also benefit from techniques developed for traditional ansatzes, such as backflow correlation [113] and Gutzwiller projection [110].

Inspired by the development in computer vision, NQS based on convolutional neural networks (CNN) [114] takes advantage of the locality and the translational symmetry of physical systems on regular lattices. We can also restore symmetries by quantum number projections discussed in the previous subsection. Some of the data in this paper are obtained by RBM combined with quantum number projections [115]. Group convolutions [116, 117] further generalize them to richer symmetries including rotations, reflections, and gauge transformations.

Another direction for developing NQS is perfect sampling from complicated many-body probability distributions, which avoids the issue of autocorrelation time in MCMC. It can be achieved by autoregressive neural networks (ARNN) [118], which are exact likelihood models. They rely on decomposing a joint probability into conditional probabilities that can be sequentially sampled, and they can comprise either dense or convolutional layers. Recurrent neural networks (RNN) [119, 120] maintain a memory of past information during the sampling procedure, which makes them suitable to capture long-range correlations, as suggested by their success in natural language processing.

One interesting feature about the RNN is its capability to encode translation-invariant properties of the bulk of a quantum system [120]. Moreover, RNNs can be extended to multiple spatial dimensions. In particular, one can construct 2D RNNs that were shown to be competitive with DMRG, and also cheaper in terms of computational complexity compared to projected-entangled pair states (PEPS) [119]. RNNs can use much fewer parameters than other architectures to encode information in a large spatial area, thanks to weight sharing between RNN cells at all sites. In a similar fashion to tensor networks, tensorized versions of RNNs have been used to provide higher expressive power [121–123] and more accurate estimations of ground-state energies that outperform DMRG in certain regimes [122, 123]. They were also shown to be able to encode the area law of entanglement [123]. Symmetries, such as U(1) symmetry and spatial symmetries, can be imposed in RNNs to improve the quality of variational calculations [119, 122]. Being the recent state of the art for many machine learning tasks, transformers have also been proposed for NQS as they have a more flexible autoregressive architecture [116, 124]. Additionally, it is worth noting that thermal-like fluctuations can be added to the training of autoregressive models in the hope of escaping local minima that can be encountered when studying disordered or frustrated systems [120–122].

Neural networks can be used as well to simulate fermionic systems within the second quantization formalism [125–127]. The fermionic modes are mapped into an interacting quantum spin model in this formalism. This can be achieved via the Jordan–Wigner [128], the parity, or the Bravyi–Kitaev [129] transformations. The reduction of the fermionic problem into a spin Hamiltonian makes it possible to exploit the success of NQS on spin systems. However, this approach suffers from the disadvantage of producing a spin Hamiltonian with non-local interactions. First quantization is therefore an attractive alternative as it preserves the locality of the physical interactions. In this case, the amplitudes of the variational state $\psi_\theta(x)$ must be antisymmetric with respect to permutations of the particle indices. NQS-based parametrizations of fermionic wave functions borrow inspiration from traditional ansatzes like the Slater–Jastrow state, backflow correlations, and hidden particle representations.

The amplitudes of a Slater–Jastrow-inspired NQS consist of the product of a parameterized antisymmetric reference factor $|\phi\rangle$ (Slater determinant or PP) and a symmetric neural network factor $\hat{\mathcal{M}}$. The neural network is in charge of incorporating correlations on top of the reference wave function. First introduced in Ref. [85], a positive RBM was used as the correlation factor. Later works also implemented correlation factors that can alter the nodal structure of the reference state and respect the translational symmetry via the use of CNN with skip connections [130]. It it noted that if $|\phi\rangle$ is a general Slater determinant of non-orthogonal orbitals, then the NQS is a universal parametrization in the lattice [131]. Alternatively, neural networks have been used to parametrize the N-particle orbitals of an N-particle determinant in order to encode backflow correlations. This variational family has also been shown to be universal in the lattice [113].

Lastly, antisymmetric NQS has also been constructed using "hidden" additional fermionic degrees of freedom [131]. In this case, the variational state is represented by a Slater determinant in the Hilbert space that is augmented by adding the "hidden" fermions. The Slater determinant in the augmented space is then projected into the physical Hilbert space, and this projection is parameterized by a neural network. The neural network parameters are optimized together with the single-particle orbitals of the determinant. This ansatz explicitly contains the above Slater–Jastrow-inspired factorization, as well as a compact representation of configuration-interaction wave functions [131].

3. *Variational auxiliary-field quantum Monte Carlo*

The variational auxiliary-field quantum Monte Carlo (VAFQMC) [132] approach creates a variational ansatz for the ground state wave function of the Hubbard Hamiltonian $\hat{H} = \hat{K} + \hat{V}$, by projections via an optimizable pseudo-Hamiltonian, using the formalism of AFQMC (see Section S3 E 1). In VAFQMC, a single Slater determinant $|\psi_{\mathrm{MF}}\rangle$ is first constructed from an effective

mean-field calculation with a set of variational parameters $\boldsymbol{\alpha}_0$, such that $\hat{H}_{\mathrm{MF}}(\boldsymbol{\alpha}_0)\,|\psi_{\mathrm{MF}}\rangle = E_0(\boldsymbol{\alpha}_0)\,|\psi_{\mathrm{MF}}\rangle$. A variational ansatz is then constructed, by operating onto $|\psi_{\mathrm{MF}}\rangle$ a projection operator:

$$|\psi_\tau\rangle = \exp\left(-\frac{\tau}{2}\Big(\hat{H}_{\mathrm{MF}}(\boldsymbol{\alpha}) + \hat{V}\Big)\right)|\psi_{\mathrm{MF}}\rangle\,, \tag{S64}$$

where $\boldsymbol{\alpha}$ denotes a set of variational parameters.

In Eq. (S64), $\hat{H}_{\mathrm{MF}}(\boldsymbol{\alpha})$ replaces the kinetic part of the Hubbard Hamiltonian with a general quadratic operator of fermionic creation and annihilation operators, which can include, for instance, a d-wave BCS pairing field. In this work, $\hat{H}_{\mathrm{MF}}(\boldsymbol{\alpha})$ was designed to give a $|\psi_{\mathrm{MF}}\rangle$ describing AFM stripes with the expected wavelength [133]. The potential part $\hat{V}$ is kept as the original Hubbard on-site interaction. The parameter τ, the total imaginary time of the projection, is kept fixed and plays the role of an effective inverse temperature.

The projection in Eq. (S64) is further broken up as

$$\begin{aligned}&\exp\left(-\frac{\tau}{2}\left(\hat{H}_{\mathrm{MF}}(\boldsymbol{\alpha}) + \hat{V}\right)\right) = \\ &\prod_{i=1}^{n}\exp\left(-t_i\hat{H}_{\mathrm{MF}}(\boldsymbol{\alpha})\right)\exp\left(-h_i\hat{V}\right) \\ &\times\exp\left(-t_{n+1}\hat{H}_{\mathrm{MF}}(\boldsymbol{\alpha})\right),\end{aligned} \tag{S65}$$

where n is the number of time steps given by $n = \max\left([(U\tau/0.4-1)/5], 1\right)$ depending on the interaction strength U. The variable steps h_i and t_i are treated as additional variational parameters to optimize in order to minimize Trotter errors [134]. However, here we introduce a simple functional form for the non-uniform time steps which depends on a single variational parameter $\Delta\tau$ (see Ref. [132] for details). The variational ansatz $|\psi_\tau\rangle$ in Eq. (S64) can thus be equivalently denoted as $|\psi_n\rangle$, yielding a variational energy $E_n := \langle\psi_n|\hat{H}|\psi_n\rangle\,/\,\langle\psi_n|\psi_n\rangle$.

Now following similar procedures to AFQMC, we recast the variational energy as

$$E_n = \frac{\sum_{\boldsymbol{\sigma}\boldsymbol{\sigma}'}\,\langle\psi_{\mathrm{MF}}|\hat{U}_n^\dagger(\boldsymbol{\sigma}')\hat{H}\hat{U}_n(\boldsymbol{\sigma})|\psi_{\mathrm{MF}}\rangle}{\sum_{\boldsymbol{\sigma}\boldsymbol{\sigma}'}\,\langle\psi_{\mathrm{MF}}|\hat{U}_n^\dagger(\boldsymbol{\sigma}')\hat{U}_n(\boldsymbol{\sigma})|\psi_{\mathrm{MF}}\rangle}, \tag{S66}$$

where

$$\begin{aligned}\hat{U}_n(\boldsymbol{\sigma}) := &\prod_{i=1}^{n}\exp\left(-t_i\hat{H}_{\mathrm{MF}}(\boldsymbol{\alpha})\right)\exp\Big(\lambda_i\sum_j\sigma_{j,i}\hat{m}_j\Big) \\ &\times\exp\left(-t_{n+1}\hat{H}_{\mathrm{MF}}(\boldsymbol{\alpha})\right),\end{aligned} \tag{S67}$$

with the vector $\boldsymbol{\sigma} := \{\sigma_{ji}\} \in \{\pm1\}^{N_{\mathrm{s}}\times n}$ denoting the collection of auxiliary fields arising from the discrete Hubbard–Stratonovich transformation [135]. The spin operator for each site j is defined as $\hat{m}_j = \hat{n}_{j,\uparrow} - \hat{n}_{j,\downarrow}$, and the parameter λ_i is given by $\cosh\lambda_i = \exp(Uh_i)$.

Eq. (S66) is now rewritten in a form suitable for Monte Carlo:

$$E_n = \frac{\sum_{\boldsymbol{\sigma}\boldsymbol{\sigma}'}|W_n(\boldsymbol{\sigma}',\boldsymbol{\sigma})|e_n(\boldsymbol{\sigma}',\boldsymbol{\sigma})S_n(\boldsymbol{\sigma}',\boldsymbol{\sigma})}{\sum_{\boldsymbol{\sigma}\boldsymbol{\sigma}'}|W_n(\boldsymbol{\sigma}',\boldsymbol{\sigma})|S_n(\boldsymbol{\sigma}',\boldsymbol{\sigma})}, \tag{S68}$$

where

$$W_n(\boldsymbol{\sigma}',\boldsymbol{\sigma}) := \langle\psi_{\mathrm{MF}}|\hat{U}_n^\dagger(\boldsymbol{\sigma}')\hat{U}_n(\boldsymbol{\sigma})|\psi_{\mathrm{MF}}\rangle\,, \tag{S69}$$

$$S_n(\boldsymbol{\sigma}',\boldsymbol{\sigma}) := \frac{W_n(\boldsymbol{\sigma}',\boldsymbol{\sigma})}{|W_n(\boldsymbol{\sigma}',\boldsymbol{\sigma})|}, \tag{S70}$$

$$e_n(\boldsymbol{\sigma}',\boldsymbol{\sigma}) := \frac{\langle\psi_{\mathrm{MF}}|\hat{U}_n^\dagger(\boldsymbol{\sigma}')\hat{H}\hat{U}_n(\boldsymbol{\sigma})|\psi_{\mathrm{MF}}\rangle}{W_n(\boldsymbol{\sigma}',\boldsymbol{\sigma})} \tag{S71}$$

are the weight factor, the phase factor, and the local energy respectively. For each set of variational parameters, we estimate the energy by $E_n = \langle S_n e_n\rangle_{W_n}/\langle S_n\rangle_{W_n}$, where $\langle\,\cdot\,\rangle_{W_n}$ indicates the average of a random variable with respect to the Monte Carlo samples from the probability distribution $|W_n|$.

The optimization of the variational parameters is then carried out by generalizing techniques from standard variational Monte Carlo [2, 87] and from machine learning [86]. Minimizing $E_n(\boldsymbol{\alpha})$ with respect to $\boldsymbol{\alpha}$ requires the energy derivatives. Assuming there are $2p$ variational parameters $\alpha_1, \alpha_2, \ldots, \alpha_{2p}$ plus an extra parameter of the minimum time step $\alpha_{2p+1} = \Delta\tau$, we can compute the derivatives by

$$\frac{\partial E_n}{\partial\alpha_j} = \frac{\left\langle S_n\Big(\frac{\partial e_n}{\partial\alpha_j} + (e_n - E_n)O_j\Big)\right\rangle_{W_n}}{\langle S_n\rangle_{W_n}}, \tag{S72}$$

where $O_j := \frac{\partial\ln W_n}{\partial\alpha_j}$. The complex derivatives $\frac{\partial e_n}{\partial\alpha_j}$ and O_j are obtained by automatic differentiation [136]. The final optimized E_n defined within the variational ansatz is dependent on the parameter τ, and provides an upper bound to the true ground-state energy of the given Hamiltonian.

D. Parameterized quantum circuits and variational quantum eigensolver

Variational states can be represented by parameterized quantum circuits (PQC) [18]. A PQC has parameterized single and two-qubit gates, e.g. rotation gates, as well as multi-qubit entangling gates [137]. Different hardware architectures come with different sets of native gates. The set of variational parameters θ is optimized using the variational quantum eigensolver (VQE) algorithm [16]. When using quantum hardware, the expectation value of the Hamiltonian and the energy gradient components must be measured through repeated wave function collapses, thus introducing a sampling noise similar to the one in VMC simulations. We only consider noiseless emulations of the quantum circuits, i.e., neglecting this measurement sampling noise, as well as hardware noise due to decoherence. We do this because our goal is to test whether the relation between the V-score and the relative energy accuracy also holds in the quantum computing setting.

We employ several types of PQC ansatzes and optimization methods. We first tackle the TFIM at the criticality $\Gamma = 1$, defined on a finite chain with L sites.

The circuit is made of a series of d blocks built from single-qubit rotations $\hat{U}_{\rm R}(\theta^k)$, interlayered with entangler blocks $\hat{U}_{\rm ent}$, that spans the required length of the qubit register, with $k = 1, \ldots, d+1$. This is made of a ladder of CNOT (also known as CX) gates with linear connectivity, such that qubit q_i is target of qubit q_{i-1} and controls qubit q_{i+1} for $i = 1, \ldots, L-2$. Since the single-qubit rotations are all local operations, $\hat{U}_{\rm R}(\theta^k)$ can be written as a tensor product of rotations on single qubits:

$$\hat{U}_{\rm R}(\theta^k) = \bigotimes_{i=0}^{L-1} \hat{R}_y(\theta^k_{q_i}), \tag{S73}$$

where $\hat{R}_y(\theta^k_{q_i})$ is a rotation around the y-axis on the Bloch sphere of qubit q_i. The full unitary circuit operation is described by

$$\hat{U}_{\text{R-CX}}(\theta) := \hat{U}_{\rm R}(\theta^{d+1}) \prod_{i=d}^{1} \hat{U}_{\rm ent} \hat{U}_{\rm R}(\theta^i), \tag{S74}$$

and the final parameterized state is

$$|\psi(\theta)\rangle = \hat{U}_{\text{R-CX}}(\theta)\, |0\rangle^{\otimes L}. \tag{S75}$$

The accuracy of the calculation is controlled by the circuit depth d, and the total number of variational parameters is $L(d+1)$.

An alternative is to use physically motivated ansatzes, such as the Hamiltonian variational (HV) ansatz [138]. The unitary operator defining the HV ansatz is made of d blocks, and each block is a product of ℓ operators $\hat{U}_\alpha = \exp(\mathrm{i}\theta^k_\alpha \hat{H}_\alpha)$, with $\alpha = 1, \ldots, \ell$ indexing the non-commuting terms of the Hamiltonian. For TFIM we only need $\ell = 2$. In this case, the full unitary operator is

$$\hat{U}_{\rm HV}(\theta) := \prod_{i=d}^{1} \hat{U}_2(\theta^i_2) \hat{U}_1(\theta^i_1), \tag{S76}$$

which can be efficiently decomposed using one- and two-qubit quantum gates, and the final parameterized state is

$$|\psi(\theta)\rangle = \hat{U}_{\rm HV}(\theta) \left(\frac{|0\rangle + |1\rangle}{\sqrt{2}} \right)^{\otimes L}, \tag{S77}$$

where the initial non-entangled state can be obtained from $|0\rangle^{\otimes L}$ by placing one Hadamard gate on each qubit. The total number of parameters is ℓd.

We use both the heuristic R-CX ansatz and the physically motivated HV ansatz on the TFIM. We have obtained a family of optimized trial states that depend on the circuit depth d, as shown in Fig. S1, confirming the linear scaling of V-score versus energy relative error proposed in the main text.

For the Heisenberg and the J_1-J_2 models, we employ the symmetry-enhanced architecture introduced in Ref. [139]. This PQC is more sophisticated compared to the previous ones, which are included mainly for demonstration purposes, and allows one to enforce translational, point-group, and SU(2) symmetries of the variational wave function in the device noise-free case. This circuit has shown outstanding performance in representing the ground state with only a few variational parameters when applied to 2D frustrated magnets [140].

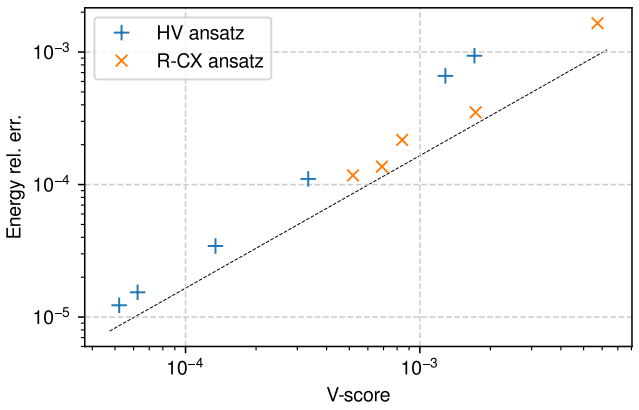


Figure S1. **Comparison of V-scores of VQE ansatzes versus energy relative errors on a 10 sites TFIM.** The black dashed line is the least square fit of all variational methods shown in Fig. 2 in the main text. Each data point corresponds to a full VQE optimization using different circuit depths d. For the HV ansatz, we report $d = 8, 12, 16, 20, 24, 26$, while for R-CX we report $d = 4, 6, 8, 10, 12$. Longer circuits provide systematically better V-scores and energy relative errors.

The initial state to be prepared is the *dimerized state*, splitting the L site indices into $L/2$ arbitrary pairs:

$$|\psi_D\rangle := \bigotimes_{0 \le i < L/2} \frac{1}{\sqrt{2}} \Big(|{\uparrow_{2i}}{\downarrow_{2i+1}}\rangle - |{\downarrow_{2i}}{\uparrow_{2i+1}}\rangle \Big), \tag{S78}$$

which is a product of $L/2$ dimers on the selected pairs. This state is manifestly the SU(2)-singlet. To access other total spin quantum numbers, one should replace one or more singlets in the product Eq. (S78) with triplets $\frac{1}{\sqrt{2}} \Big(|{\uparrow_{2i}}{\downarrow_{2i+1}}\rangle + |{\downarrow_{2i}}{\uparrow_{2i+1}}\rangle \Big)$.

We then introduce the *SWAP operator*

$$\hat{P}_{ij} := \frac{1}{2} \left(\hat{\boldsymbol{S}}_i \cdot \hat{\boldsymbol{S}}_j + \hat{\mathbb{1}} \right), \tag{S79}$$

which exchanges spin states between sites i and j. We note that the SWAP operator commutes with the total spin operator $\hat{S}^2$, therefore the variational ansatz

$$|\psi(\theta)\rangle = \left(\prod_\alpha \mathrm{e}^{\mathrm{i}\theta_\alpha \hat{P}_{i_\alpha j_\alpha}} \right) |\psi_D\rangle \tag{S80}$$

has a well-defined total spin quantum number. The exponential SWAP (eSWAP) operator $\mathrm{e}^{\mathrm{i}\theta_\alpha \hat{P}_{i_\alpha j_\alpha}}$ can be efficiently implemented on a quantum computer, as the

two-qubit SWAP gate can be compiled into single-qubit gates and CNOTs, or is a native gate in architectures alternative to superconductive qubits.

Let us now suppose that, in addition to the SU(2) symmetry, the system also has translational or point-group symmetries, which can be all represented as qubit permutations. In Ref. [139], the authors provide a way that allows one to effectively project the wave function onto any irreducible representation subspace of the spatial symmetry. To this end, let us consider the spatial symmetry projector

$$\hat{P} := \frac{1}{|G|} \sum_{g \in G} \chi_g \hat{g}, \tag{S81}$$

where G is the full spatial symmetry group consisting of the elementary unitary permutations $\hat{g}$, and χ_g are the characters depending on the desired projection quantum number (irreducible representation). The projected wave function $|\psi_P(\theta)\rangle = \frac{\hat{P}}{\sqrt{\mathcal{N}(\theta)}} |\psi(\theta)\rangle$ is normalized with $\mathcal{N}(\theta) := \langle\psi(\theta)|\hat{P}|\psi(\theta)\rangle$, since $\hat{P}^2 = \hat{P}$.

This wave function is optimized using the *natural gradient* approach discussed in Section S3 C. The energy gradient is preconditioned using the *metric tensor*, which mimics imaginary-time evolution in the allowed variational subspace. For this symmetry-enhanced ansatz, the energy gradient reads

$$\partial_i \langle E(\theta)\rangle = 2\,\mathrm{Re}\left(\frac{\langle\psi(\theta)|\hat{H}\hat{P}|\partial_i\psi(\theta)\rangle}{\mathcal{N}(\theta)} - \mathcal{A}_i(\theta)\langle E(\theta)\rangle \right), \tag{S82}$$

where $\mathcal{A}_i(\theta) := \frac{1}{\mathcal{N}(\theta)} \langle\psi(\theta)|\hat{P}|\partial_i\psi(\theta)\rangle$ is the *connection*. The metric tensor is defined as

$$G(\theta)_{ij} := \frac{\langle\partial_i\psi(\theta)|\hat{P}|\partial_j\psi(\theta)\rangle}{\mathcal{N}(\theta)} - \mathcal{A}_i^*(\theta)\mathcal{A}_j(\theta). \tag{S83}$$

The parameter update reads

$$\theta \leftarrow \theta - \eta \sum_j \big(\mathrm{Re}\, G(\theta)\big)^{-1}_{ij} \partial_j \langle E(\theta)\rangle . \tag{S84}$$

The matrix elements required to construct these objects can be measured using the Hadamard test rule [139, 140]. During the simulation of the VQE optimization on a classical computer, we either measure these matrix elements using N_s circuit shots or compute them exactly. In the former case, the metric tensor obtained with sampling should be regularized to make the matrix inversion well defined. Here, we employ the regularization $G_{\mathrm{reg}} = \sqrt{GG} + \beta\mathbb{1}$ as suggested in Ref. [141].

Finally, it is also important to check that the V-score can be efficiently computed with VQE. The current way to estimate the energy with VQE is by calculating the weighted sum of the expectation values of all Pauli operators that compose the Hamiltonian. For instance, the TFIM is made of L local terms, yet most of them can be measured simultaneously [17]. It turns out that only two bases are needed: the computational basis, and the rotated $\{\hat{\sigma}^x\}^{\otimes L}$ basis. A similar argument applies for $\hat{H}^2$: While an upper bound for the number of terms is L^2, it is possible to check that the number of groups of Pauli operators that can be measured simultaneously grows sub-linearly with the system size, in both one and two dimensions.

The 2D Hubbard model introduces the additional complication of the fermion-to-qubit mapping. In more than one dimension, the Jordan–Wigner mapping generally transforms a two-local fermionic operator, i.e. the hopping term, into a non-local qubit operator [16]. However, numerical tests show that the number of bases, thus the number of measurements, grows only sub-quadratically with the system size.

E. Quantum Monte Carlo

In the benchmarks, we include some quantum Monte Carlo (QMC) methods. Although they are not strictly variational and the V-score is not applicable to them, they produce numerically exact results in some cases, while in other cases highly accurate energies and the bias has been certified to be small. These cases are discussed below, with the methods described in detail. The QMC results are used to assess the V-scores of variational methods when ED is not practical.

1. *Auxiliary-field quantum Monte Carlo*

There exist two different auxiliary-field Monte Carlo algorithms, based on one hand the finite temperature grand canonical ensemble [26], and on the other hand the ground state canonical ensemble [27, 28]. The auxiliary-field quantum Monte Carlo (AFQMC) method (for an overview, see e.g. Ref. [3]) used in this work is based on the latter. It filters out the ground state from an initial state by an imaginary-time propagation: $|\psi_G\rangle = \lim_{\tau\to\infty} \mathrm{e}^{-\tau\hat{H}} |\psi_I\rangle$, where $|\psi_I\rangle$ and $|\psi_G\rangle$ are the initial and ground state respectively, τ the imaginary time, and $\hat{H}$ the many-body Hamiltonian. The initial state must satisfy $\langle\psi_I|\psi_G\rangle \neq 0$ but can be otherwise arbitrary. We have typically taken it from a mean-field calculation [142, 143].

The projection is achieved by first discretizing the imaginary time into m small time steps $\Delta\tau$: $\mathrm{e}^{-\tau\hat{H}} = (\mathrm{e}^{-\Delta\tau\hat{H}})^m$, and then applying Trotter-Suzuki breakup [144, 145] to each time step:

$$\mathrm{e}^{-\Delta\tau\hat{H}} = \mathrm{e}^{-\Delta\tau\hat{K}/2}\mathrm{e}^{-\Delta\tau\hat{V}}\mathrm{e}^{-\Delta\tau\hat{K}/2} + O\big(\Delta\tau^3\big). \tag{S85}$$

Here $\hat{K}$ is the kinetic part consisting of one-body operators, and $\hat{V}$ the interacting part containing two-body operators. For all systems computed in this work, we have either extrapolated $\Delta\tau$ to zero or set it to a fixed

value (typically $\Delta\tau = 0.01$) and verified that the Trotter error is well within our statistical error.

Hubbard–Stratonovich (HS) transformation is applied to rewrite the interacting part into a one-body form coupled with auxiliary fields (AFs) $\{x_i\}$. In the Hubbard model, a spin decomposition is applied to the $U > 0$ case:

$$\mathrm{e}^{-\Delta\tau U\hat{n}_{i\uparrow}\hat{n}_{i\downarrow}} = \mathrm{e}^{-\Delta\tau U(\hat{n}_{i\uparrow}+\hat{n}_{i\downarrow})/2}\sum_{x_i=\pm1}\frac{1}{2}\mathrm{e}^{\gamma x_i(\hat{n}_{i\uparrow}-\hat{n}_{i\downarrow})}, \tag{S86}$$

where $\cosh\gamma = \exp(\Delta\tau U/2)$. For the $U < 0$ cases, we apply a charge decomposition form of the HS transformation. A systematic study of the effect of the different HS transformations can be found in Ref. [146].

The projection is carried out by evaluating the propagator as an integral using the Monte Carlo method:

$$\mathrm{e}^{-\Delta\tau\hat{H}} = \int p(\boldsymbol{x})\hat{B}(\boldsymbol{x})\mathrm{d}\boldsymbol{x}, \tag{S87}$$

where $\boldsymbol{x} := \{x_i\} = \{x_1, x_2, \ldots, x_{N_\mathrm{s}}\}$ for a lattice with N_s sites, and $p(\boldsymbol{x})$ is a probability distribution in AF space, which is a uniform function in the discrete HS transformation above. The propagator $\hat{B}(\boldsymbol{x})$ now only contains one-body operators: $\hat{B}(\boldsymbol{x}) := \mathrm{e}^{-\Delta\tau\hat{K}/2}\hat{\boldsymbol{b}}(\boldsymbol{x})\mathrm{e}^{-\Delta\tau\hat{K}/2}$, where $\hat{\boldsymbol{b}}(\boldsymbol{x}) := \prod_{i=1}^{N_\mathrm{s}}\hat{b}_i(x_i)$ is the product of the one-body operators transformed from the interacting part, i.e., the right-hand side of Eq. (S86).

In this work, two different ground state AFQMC methods are used. For sign-problem-free systems, e.g., the repulsive Hubbard model ($U > 0$) at half-filling or the spin-balanced attractive Hubbard model ($U < 0$) [135, 147], numerically exact results are given by AFQMC using a generalized Metropolis algorithm with force bias [30, 148]. On the other hand, for a doped Hubbard model with $U > 0$, AFQMC is applied with a constraint path (CP) approximation [29, 149] to control the sign problem. The two ground state AFQMC methods are described in the following two subsections, respectively.

2. *Sign-problem-free Hubbard model: exact AFQMC with generalized Metropolis algorithm*

For the half-filled repulsive Hubbard model on a square lattice, as well as the spin-balanced attractive Hubbard model, the sign problem is absent in AFQMC because of symmetry. In these cases, exact calculations are performed. (Note that in almost all sign-problem-free calculations, the standard approach has an infinite variance problem which must be properly taken care of [148].) We study these systems using the AFQMC with a generalized Metropolis algorithm [30]. The ground-state energy is given by $\langle\hat{H}\rangle = \langle\psi_L|\hat{H}|\psi_R\rangle / \langle\psi_L|\psi_R\rangle$, where $\langle\psi_L| := \langle\psi_I|\exp\left(-\tau_L\hat{H}\right)$ and $|\psi_R\rangle := \exp\left(-\tau_R\hat{H}\right)|\psi_I\rangle$. Since the Hamiltonian commutes with propagators, energy measurements can be inserted between any two time steps, i.e., with any combination of $\tau_L + \tau_R = \tau$, provided τ is sufficiently larger than the equilibration time, $\tau > \tau_\mathrm{eq}$, to reach the ground state from $|\psi_I\rangle$.

To illustrate the sampling process, we rewrite the energy expectation as a path integral in AF space:

$$\langle\hat{H}\rangle = \frac{\int\frac{\langle\phi_L|\hat{H}|\phi_R\rangle}{\langle\phi_L|\phi_R\rangle}P(\boldsymbol{X})\langle\phi_L|\phi_R\rangle\,\mathrm{d}\boldsymbol{X}}{\int P(\boldsymbol{X})\langle\phi_L|\phi_R\rangle\,\mathrm{d}\boldsymbol{X}}, \tag{S88}$$

where $\langle\phi_L| := \langle\psi_I|\prod_{m=1}^{M_L}\hat{B}(\boldsymbol{x}^{(M-m+1)})$ and $|\phi_R\rangle := \prod_{m=1}^{M_R}\hat{B}(\boldsymbol{x}^{(m)})|\psi_I\rangle$. These are single Slater determinants if the initial state is chosen as a single Slater determinant. In the case of a multi-determinant $|\psi_I\rangle$, the different terms in the linear combination can be sampled. The number $M = \tau/\Delta\tau$ is the total number of time slices, and M_L and M_R correspond to τ_L and τ_R respectively. The variables $\boldsymbol{X} := \{\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \ldots, \boldsymbol{x}^{(M)}\}$ form a $N_\mathrm{s}\times M$-dimensional vector in AF space. The probability distribution $P(\boldsymbol{X}) := \prod_{m=1}^{M}p(\boldsymbol{x}^{(m)})$.

We adopt a force bias method [30] instead of the usual single-site heat bath update, which helps reduce the autocorrelation time in the Monte Carlo sampling. In this method, we update a cluster of AFs at each time slice n simultaneously. The size of the cluster N_c can be as large as N_s and can be tuned according to the acceptance ratio. Below we sketch the algorithm for the spin decomposition; generalization to other HS transformations is straightforward [3, 30]. Each AF in the cluster is proposed a new value according to

$$\mathcal{P}(x_i) := \frac{\mathrm{e}^{\gamma x_i(\bar{n}_{i\uparrow}-\bar{n}_{i\downarrow})}}{\sum_{x_i=\pm1}\mathrm{e}^{\gamma x_i(\bar{n}_{i\uparrow}-\bar{n}_{i\downarrow})}}, \tag{S89}$$

thus giving a probability density for a new candidate cluster $\mathcal{P}(\boldsymbol{x}^{(n)}) := \prod_{i=1}^{N_\mathrm{c}}\mathcal{P}(x_i^{(n)})$. The optimal choice of force bias is given by $\bar{n}_{i\sigma} := \langle\phi_L|\hat{n}_{i\sigma}|\phi_R\rangle / \langle\phi_L|\phi_R\rangle$, where the σ and i are the spin and lattice site indices, and leads to a $\mathcal{P}$ that approximates the target probability to $\mathcal{O}(\sqrt{\Delta\tau})$.

3. *Doped repulsive Hubbard model: constrained path AFQMC*

All the AFQMC results on the repulsive Hubbard model away from half-filling are obtained with CP-AFQMC [3, 29, 150]. CP-AFQMC is built on an open-ended random walk, with an indefinite value of τ. The sign problem is removed by introducing a trial state $|\psi_T\rangle$ to guide and constrain the random walk. In this work, $|\psi_T\rangle$ is obtained from noninteracting calculations for closed-shell systems and Hartree–Fock calculations for open-shell systems. In the latter cases, the CP error is further reduced by applying a self-consistent constraint [151] which couples to a generalized Hartree Fock (GHF) calculation with an effective U_eff determined via the self-consistency [142]. Several results are also provided from constraint release [146] which are essentially exact, as indicated in the main text.

Different from the Metropolis approach discussed above, CP-AFQMC is performed via an open-ended branching random walk, in which a population of walkers $\{|\phi_k^{(n)}\rangle, w^{(n)_k}\}$ are propagated following the time evolution $|\psi^{(n+1)}\rangle = \mathrm{e}^{-\tau\hat{H}}|\psi^{(n)}\rangle$. These walkers sample the many-body wave function in the sense that $|\psi^{(n)}\rangle \propto \sum_k w^{(n)}|\phi_k^{(n)}\rangle / \langle\psi_T|\phi_k^{(n)}\rangle$. The ground-state energy is then given by $\langle\hat{H}\rangle = \langle\psi_T|\hat{H}|\psi_G\rangle / \langle\psi_T|\psi_G\rangle$. After the random walk has equilibrated (i.e., after a sufficient number of steps $n_{\mathrm{eq}} = \tau_{\mathrm{eq}}/\Delta\tau$), the walkers will sample $|\psi_G\rangle$, and all subsequent steps $n > n_{\mathrm{eq}}$ can be used to compute the ground-state energy:

$$\langle\hat{H}\rangle = \frac{\sum_{k,n} \frac{\left\langle\psi_T\middle|\hat{H}\middle|\phi_k^{(n)}\right\rangle}{\left\langle\psi_T\middle|\phi_k^{(n)}\right\rangle} w_k^{(n)}}{\sum_{k,n} w_k^{(n)}}. \tag{S90}$$

In the random walks, importance sampling is introduced which amounts to sampling new AF to advance the walker by a modified probability density:

$$|\phi^{(n+1)}\rangle \leftarrow \int \tilde{p}(\boldsymbol{x})\hat{B}(\boldsymbol{x})\mathrm{d}\boldsymbol{x}\,|\phi^{(n)}\rangle, \tag{S91}$$

where $\tilde{p}(\boldsymbol{x}) \propto \langle\psi_T|\phi^{(n+1)}\rangle / \langle\psi_T|\phi^{(n)}\rangle$ builds in the knowledge from $|\psi_T\rangle$ to improve the sampling efficiency [3]. The actual form of $\tilde{p}(\boldsymbol{x})$ contains force bias terms similar to how we formulated the generalized Metropolis algorithm above.

The importance sampling transformation also automatically imposes a constraint [3]. As $\Delta\tau$ approaches zero, the random walkers will not cross the surface defined by $\langle\psi_T|\phi_k^{(n)}\rangle = 0$, which separates two degenerate regions in AF space, each of which is overcomplete and can fully represent $|\psi_G\rangle$. Constraining the random walks in one region of the Slater determinant space (or equivalently, the AF space) is an exact condition if $|\psi_T\rangle = |\psi_G\rangle$. The CP approximation uses an approximate $|\psi_T\rangle$ to impose this condition, which leads to a systematic error. The ground-state energy computed by the mixed estimate of Eq. (S90) is therefore *not variational* [152]. The CP-AFQMC method has been extensively benchmarked, and the CP error was shown to be small for Hubbard-like systems (see, e.g., Refs. [142, 146, 153]), and the results provided here for the doped Hubbard model are expected to be very accurate, with the relative error below a few tenths of a percent.

As mentioned, the energies for the doped Hubbard model were computed using either (i) constraint release (taken from Ref. [146], essentially exact), (ii) with free-electron $|\psi_T\rangle$ (for closed-shell systems), or (iii) with a self-consistent constraint [151]. In the data included in this paper, we have indicated how each energy was obtained and, in the case of (iii), including the final U_{eff} in the GHF after convergence of the self-consistent iteration between GHF and AFQMC.

4. *Continuous-time quantum Monte Carlo*

The continuous-time quantum Monte Carlo [154] algorithm carries out a diagrammatic expansion of the imaginary-time projection operator $\mathrm{e}^{-\tau\hat{H}}$ and samples interaction expansion terms

$$\begin{aligned}&\langle\psi_I|\mathrm{e}^{-\tau\hat{H}}|\psi_I\rangle = \\ &\sum_{k=0}^{\infty}(-1)^k\int_0^\tau \mathrm{d}\tau_1\left(\prod_{i=1}^{k-1}\int_{\tau_i}^{\tau}\mathrm{d}\tau_{i+1}\right) \\ &\cdot\ \langle\psi_I|\mathrm{e}^{-(\tau-\tau_k)\hat{H}_0}\hat{H}_1\left(\prod_{i=k-1}^{1}\mathrm{e}^{-(\tau_{i+1}-\tau_i)\hat{H}_0}\hat{H}_1\right)\mathrm{e}^{-\tau_1\hat{H}_0}|\psi_I\rangle,\end{aligned} \tag{S92}$$

where $\hat{H} = \hat{H}_0 + \hat{H}_1$. Ground state physical observables are measured as

$$\langle\hat{O}\rangle = \lim_{\tau\to\infty}\frac{\langle\psi_I|\mathrm{e}^{-\tau\hat{H}/2}\,\hat{O}\,\mathrm{e}^{-\tau\hat{H}/2}|\psi_I\rangle}{\langle\psi_I|\mathrm{e}^{-\tau\hat{H}}|\psi_I\rangle}. \tag{S93}$$

For the t-V model, the initial state $|\psi_I\rangle$ is chosen as a single Slater determinant, and $\hat{H}_0$ and $\hat{H}_1$ are the non-interaction and the interaction terms respectively. Here, we write $\hat{H}_1 = \frac{V}{4}\sum_{\langle i,j\rangle}\mathrm{e}^{\mathrm{i}\pi(\hat{n}_i+\hat{n}_j)}$ to ensure that each term in the interaction expansion is evaluated as a determinant. Simulation of the spinless t-V model on the bipartite lattice at half-filling with repulsive interaction is free from the fermion sign problem [155]. Therefore, the results are free from systematic Trotter or constrained path errors. The computational time complexity of this method scales as $O(\tau V N_{\mathrm{s}}^3)$.

S4. COMPARISON OF THE VARIATIONAL METHODS

In addition to characterizing the hardness of Hamiltonians, the V-score can also serve as a metric to compare the accuracy of different variational methods. As shown in Fig. S2, MPS works excellently on 1D problems, while NQS performs better than traditional methods on 2D J_1-J_2 models that are too large for ED. We expect more research in the future to analyze the performances of different variational methods on a wider range of physical models.

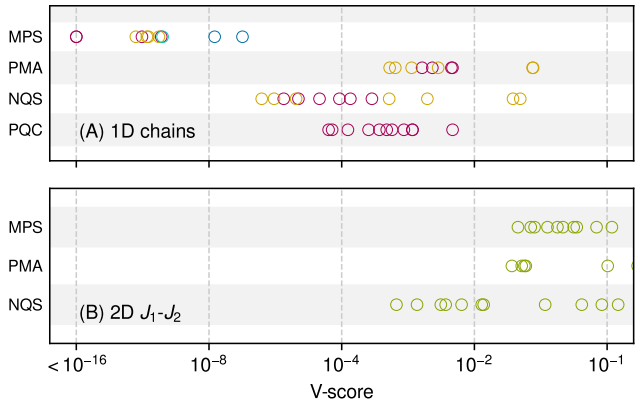


Figure S2. **Comparison of V-scores from different variational methods on given Hamiltonians.** The Hamiltonian is (**A**) 1D chain of TFIM (red), Heisenberg (yellow), t-V (light blue), or Hubbard (dark blue); and (**B**) 2D J_1-J_2 beyond ED size. The methods are matrix product states (MPS), VMC with physically motivated ansatzes (PMA), VMC with neural quantum states (NQS), and parameterized quantum circuits (PQC).

[1] D. Ceperley and B. Alder, Quantum Monte Carlo, Science **231**, 555 (1986).

[2] F. Becca and S. Sorella, *Quantum Monte Carlo Approaches for Correlated Systems* (Cambridge University Press, 2017).

[3] S. Zhang, Auxiliary-field quantum Monte Carlo at zero- and finite-temperature, in *Many-Body Methods for Real Materials*, Vol. 9 (Verlag des Forschungszentrum Jülich, 2019).

[4] S. R. White, Density matrix formulation for quantum renormalization groups, Phys. Rev. Lett. **69**, 2863 (1992).

[5] R. Orús, Tensor networks for complex quantum systems, Nature Rev. Phys. **1**, 538 (2019).

[6] A. Georges and G. Kotliar, Hubbard model in infinite dimensions, Phys. Rev. B **45**, 6479 (1992).

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

[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] J. I. Cirac, D. Pérez-García, N. Schuch, and F. Verstraete, Matrix product states and projected entangled pair states: Concepts, symmetries, theorems, Rev. Mod. Phys. **93**, 045003 (2021).

[10] W. L. McMillan, Ground state of liquid He^4, Phys. Rev. **138**, A442 (1965).

[11] D. Ceperley, G. V. Chester, and M. H. Kalos, Monte Carlo simulation of a many-fermion study, Phys. Rev. B **16**, 3081 (1977).

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

[13] R. P. Feynman, Simulating physics with computers, Int. J. Theor. Phys. **21**, 467 (1982).

[14] S. Lloyd, Universal quantum simulators, Science **273**, 1073 (1996).

[15] A. Y. Kitaev, Quantum measurements and the abelian stabilizer problem (1995), arXiv:quant-ph/9511026.

[16] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O'Brien, A variational eigenvalue solver on a photonic quantum processor, Nat. Commun. **5**, 10.1038/ncomms5213 (2014).

[17] A. Kandala, A. Mezzacapo, K. Temme, M. Takita, M. Brink, J. M. Chow, and J. M. Gambetta, Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets, Nature **549**, 242 (2017).

[18] M. Cerezo, A. Arrasmith, R. Babbush, S. C. Benjamin, S. Endo, K. Fujii, J. R. McClean, K. Mitarai, X. Yuan, L. Cincio, and P. J. Coles, Variational quantum algorithms, Nature Rev. Phys. **3**, 625 (2021).

[19] K. Bharti, A. Cervera-Lierta, T. H. Kyaw, T. Haug, S. Alperin-Lea, A. Anand, M. Degroote, H. Heimonen, J. S. Kottmann, T. Menke, W.-K. Mok, S. Sim, L.-C. Kwek, and A. Aspuru-Guzik, Noisy intermediate-scale quantum algorithms, Rev. Mod. Phys. **94**, 015004 (2022).

[20] E. Farhi, J. Goldstone, S. Gutmann, and M. Sipser, Quantum computation by adiabatic evolution (2000), arXiv:quant-ph/0001106.

[21] M. Motta, C. Sun, A. T. K. Tan, M. J. O'Rourke, E. Ye, A. J. Minnich, F. G. S. L. Brandão, and G. K.-L. Chan, Determining eigenstates and thermal states on a quantum computer using quantum imaginary time evolution, Nat. Phys. **16**, 205 (2020).

[22] R. M. Parrish and P. L. McMahon, Quantum filter diagonalization: Quantum eigendecomposition without full quantum phase estimation (2019), arXiv:1909.08925.

[23] W. Kirby, M. Motta, and A. Mezzacapo, Exact and efficient Lanczos method on a quantum computer, Quantum **7**, 1018 (2023).

[24] S. Bravyi, D. Gosset, and R. König, Quantum advantage with shallow circuits, Science **362**, 308 (2018).

[25] A. Bouland, B. Fefferman, C. Nirkhe, and U. Vazirani, On the complexity and verification of quantum random circuit sampling, Nat. Phys. **15**, 159 (2019).

[26] R. Blankenbecler, D. J. Scalapino, and R. L. Sugar, Monte Carlo calculations of coupled boson-fermion systems. I, Phys. Rev. D **24**, 2278 (1981).

[27] S. Sorella, S. Baroni, R. Car, and M. Parrinello, A novel technique for the simulation of interacting fermion systems, Europhys. Lett. **8**, 663 (1989).

[28] M. Imada and Y. Hatsugai, Numerical studies on the Hubbard model and the *t-J* model in one- and two-dimensions, J. Phys. Soc. Jpn. **58**, 3752 (1989).

[29] S. Zhang, J. Carlson, and J. E. Gubernatis, Constrained path Monte Carlo method for fermion ground states, Phys. Rev. B **55**, 7464 (1997).

[30] H. Shi, S. Chiesa, and S. Zhang, Ground-state properties of strongly interacting Fermi gases in two dimensions, Phys. Rev. A **92**, 033603 (2015).

[31] P. Horsch and T. A. Kaplan, Exact and Monte Carlo studies of Gutzwiller's state for the localised-electron limit in one dimension, J. Phys. C **16**, L1203 (1983).

[32] M. Niss, History of the Lenz–Ising model 1950–1965: From irrelevance to relevance, Arch. Hist. Exact Sci. **63**, 243 (2008).

[33] J. C. Slater and G. F. Koster, Simplified LCAO method for the periodic potential problem, Phys. Rev. **94**, 1498 (1954).

[34] S. Sachdev, *Quantum Phase Transitions* (Cambridge University Press, 2001).

[35] X.-G. Wen, *Quantum Field Theory of Many-Body Systems: From the Origin of Sound to an Origin of Light and Electrons* (Oxford University Press, 2004).

[36] D. Bitko, T. F. Rosenbaum, and G. Aeppli, Quantum critical behavior for a model magnet, Phys. Rev. Lett. **77**, 940 (1996).

[37] P. W. Anderson, Localized magnetic states in metals, Phys. Rev. **124**, 41 (1961).

[38] L. Kouwenhoven and L. Glazman, Revival of the Kondo effect, Phys. World **14**, 33 (2001).

[39] A. C. Hewson, *The Kondo Problem to Heavy Fermions* (Cambridge University Press, 1993).

[40] J. Hubbard, Electron correlations in narrow energy bands, Proc. R. Soc. Lond. A **276**, 238 (1963).

[41] J. Kanamori, Electron correlation and ferromagnetism of transition metals, Prog. Theor. Phys. **30**, 275 (1963).

[42] M. C. Gutzwiller, Effect of correlation on the ferromagnetism of transition metals, Phys. Rev. Lett. **10**, 159 (1963).

[43] P. A. Lee, N. Nagaosa, and X.-G. Wen, Doping a Mott insulator: Physics of high-temperature superconductivity, Rev. Mod. Phys. **78**, 17 (2006).

[44] M. Imada, A. Fujimori, and Y. Tokura, Metal-insulator transitions, Rev. Mod. Phys. **70**, 1039.

[45] E. Manousakis, The spin-1/2 Heisenberg antiferromagnet on a square lattice and its application to the cuprous oxides, Rev. Mod. Phys. **63**, 1 (1991).

[46] L. Balents, Spin liquids in frustrated magnets, Nature **464**, 199 (2010).

[47] L. Savary and L. Balents, Quantum spin liquids: A review, Rep. Prog. Phys. **80**, 016502 (2016).

[48] Y. Li, G. Chen, W. Tong, L. Pi, J. Liu, Z. Yang, X. Wang, and Q. Zhang, Rare-earth triangular lattice spin liquid: A single-crystal study of $YbMgGaO_4$, Phys. Rev. Lett. **115**, 167203 (2015).

[49] M. R. Norman, Colloquium: Herbertsmithite and the search for the quantum spin liquid, Rev. Mod. Phys. **88**, 041002 (2016).

[50] See the supplementary materials, section S1 for the definitions of the Hamiltonians used in this manuscript.

[51] F. Barahona, On the computational complexity of Ising spin glass models, J. Phys. A: Math. Gen. **15**, 3241 (1982).

[52] Y. Kwon, D. M. Ceperley, and R. M. Martin, Effects of three-body and backflow correlations in the two-dimensional electron gas, Phys. Rev. B **48**, 12037 (1993).

[53] M. Imada and T. Kashima, Path-integral renormalization group method for numerical study of strongly correlated electron systems, J. Phys. Soc. Jpn. **69**, 2723 (2000).

[54] T. Kashima and M. Imada, Path-integral renormalization group method for numerical study on ground states of strongly correlated electron systems, J. Phys. Soc. Jpn. **70**, 2287 (2001).

[55] S. Sorella, Generalized Lanczos algorithm for variatonal quantum Monte Carlo, Phys. Rev. B **64**, 024512 (2001).

[56] See the supplementary materials, section S2 for a detailed discussion on the definition of the V-score.

[57] S. Boixo, S. V. Isakov, V. N. Smelyanskiy, R. Babbush, N. Ding, Z. Jiang, M. J. Bremner, J. M. Martinis, and H. Neven, Characterizing quantum supremacy in near-term devices, Nat. Phys. **14**, 595 (2018).

[58] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. S. L. Brandao, D. A. Buell, B. Burkett, Y. Chen, Z. Chen, B. Chiaro, R. Collins, W. Courtney, A. Dunsworth, E. Farhi, B. Foxen, A. Fowler, C. Gidney, M. Giustina, R. Graff, K. Guerin, S. Habegger, M. P. Harrigan, M. J. Hartmann, A. Ho, M. Hoffmann, T. Huang, T. S. Humble, S. V. Isakov, E. Jeffrey, Z. Jiang, D. Kafri, K. Kechedzhi, J. Kelly, P. V. Klimov, S. Knysh, A. Korotkov, F. Kostritsa, D. Landhuis, M. Lindmark, E. Lucero, D. Lyakh, S. Mandrà, J. R. McClean, M. McEwen, A. Megrant, X. Mi, K. Michielsen, M. Mohseni, J. Mutus, O. Naaman, M. Neeley, C. Neill, M. Y. Niu, E. Ostby, A. Petukhov, J. C. Platt, C. Quintana, E. G. Rieffel, P. Roushan, N. C. Rubin, D. Sank, K. J. Satzinger, V. Smelyanskiy, K. J. Sung, M. D. Trevithick, A. Vainsencher, B. Villalonga, T. White, Z. J. Yao, P. Yeh, A. Zalcman, H. Neven, and J. M. Martinis, Quantum supremacy using a programmable superconducting processor, Nature **574**, 505 (2019).

[59] F. Pan, K. Chen, and P. Zhang, Solving the sampling problem of the Sycamore quantum circuits, Phys. Rev. Lett. **129**, 090502 (2022).

[60] X. Gao, M. Kalinowski, C.-N. Chou, M. D. Lukin, B. Barak, and S. Choi, Limitations of linear cross-entropy as a measure for quantum advantage, PRX Quantum **5**, 010334 (2024).

[61] D. Aharonov, X. Gao, Z. Landau, Y. Liu, and U. Vazirani, A polynomial-time classical algorithm for noisy random circuit sampling, in *Proceedings of the 55th Annual ACM Symposium on Theory of Computing* (2023) pp. 945–957.

[62] A. J. Daley, I. Bloch, C. Kokail, S. Flannigan, N. Pearson, M. Troyer, and P. Zoller, Practical quantum advantage in quantum simulation, Nature **607**, 667 (2022).

[63] D. Wu, R. Rossi, F. Vicentini, N. Astrakhantsev, F. Becca, X. Cao, J. Carrasquilla, F. Ferrari, A. Georges, M. Hibat-Allah, M. Imada, A. M. Läuchli, G. Mazzola, A. Mezzacapo, A. Millis, J. R. Moreno, T. Neupert, Y. Nomura, J. Nys, O. Parcollet, R. Pohle, I. Romero, M. Schmid, J. M. Silvester, S. Sorella, L. F. Tocchio, L. Wang, S. R. White, A. Wietek, Q. Yang, Y. Yang, S. Zhang, and G. Carleo, VarBench dataset version 1.1.0 (2024), doi:10.5281/zenodo.13377360.

[64] D. Wu, R. Rossi, F. Vicentini, N. Astrakhantsev, F. Becca, X. Cao, J. Carrasquilla, F. Ferrari, A. Georges, M. Hibat-Allah, M. Imada, A. M. Läuchli, G. Mazzola, A. Mezzacapo, A. Millis, J. R. Moreno, T. Neupert, Y. Nomura, J. Nys, O. Parcollet, R. Pohle, I. Romero, M. Schmid, J. M. Silvester, S. Sorella, L. F. Tocchio, L. Wang, S. R. White, A. Wietek, Q. Yang, Y. Yang, S. Zhang, and G. Carleo, VarBench methods version 1.0.0 (2024), doi:10.5281/zenodo.13377380.

[65] M. Fishman, S. R. White, and E. M. Stoudenmire, The ITensor software library for tensor network calculations (2020), arXiv:2007.14822.

[66] T. Misawa, S. Morita, K. Yoshimi, M. Kawamura, Y. Motoyama, K. Ido, T. Ohgoe, M. Imada, and T. Kato, mVMC – open-source software for many-variable variational Monte Carlo method, Comput. Phys. Commun. **235**, 447 (2019).

[67] F. Vicentini, D. Hofmann, A. Szabó, D. Wu, C. Roth, C. Giuliani, G. Pescia, J. Nys, V. Vargas-Calderón, N. Astrakhantsev, and G. Carleo, NetKet 3: Machine learning toolbox for many-body quantum systems, SciPost Physics Codebases **7** (2022).

[68] A. Georges, L. de' Medici, and J. Mravlje, Strong correlations from Hund's coupling, Annu. Rev. Condens. Matter Phys. **4**, 137 (2013).

[69] T. Mizusaki and M. Imada, Precise estimation of shell model energy by second-order extrapolation method, Phys. Rev. C **67**, 041301 (2003).

[70] M. Taddei, M. Ruggeri, S. Moroni, and M. Holzmann, Iterative backflow renormalization procedure for many-body ground-state wave functions of strongly interacting normal Fermi liquids, Phys. Rev. B **91**, 115106 (2015).

[71] N. Schuch and F. Verstraete, Computational complexity of interacting electrons and fundamental limitations of

density functional theory, Nat. Phys. **5**, 732 (2009).
[72] C. Lanczos, An iteration method for the solution of the eigenvalue problem of linear differential and integral operators, J. Res. Natl. Bur. Stand. **45**, 255 (1950).
[73] A. Wietek and A. M. Läuchli, Sublattice coding algorithm and distributed memory parallelization for large-scale exact diagonalizations of quantum many-body systems, Phys. Rev. E **98**, 033309 (2018).
[74] U. Schollwöck, The density-matrix renormalization group in the age of matrix product states, Ann. Phys. **326**, 96 (2011).
[75] M. C. J. Eisert and M. B. Plenio, Area laws for the entanglement entropy – a review, Rev. Mod. Phys. **82**, 277 (2008).
[76] E. M. Stoudenmire and S. R. White, Studying two-dimensional systems with the density matrix renormalization group, Annu. Rev. Condens. Matter Phys. **3**, 111 (2012).
[77] C. Hubig, J. Haegeman, and U. Schollwöck, Error estimates for extrapolations with matrix-product states, Phys. Rev. B **97**, 045125 (2018).
[78] D. Bauernfeind, M. Zingl, R. Triebl, M. Aichhorn, and H. G. Evertz, Fork tensor-product states: Efficient multiorbital real-time DMFT solver, Phys. Rev. X **7**, 031013 (2017).
[79] C. Hubig, I. P. McCulloch, U. Schollwöck, and F. A. Wolf, Strictly single-site DMRG algorithm with subspace expansion, Phys. Rev. B **91**, 155115 (2015).
[80] H. Shiba, Properties of strongly correlated Fermi liquid in valence fluctuation system – a variational Monte-Carlo study, J. Phys. Soc. Jpn. **55**, 2765 (1986).
[81] H. Yokoyama and H. Shiba, Variational Monte-Carlo studies of Hubbard model. I, J. Phys. Soc. Jpn. **56**, 1490 (1987).
[82] C. Gros, R. Joynt, and T. M. Rice, Antiferromagnetic correlations in almost-localized Fermi liquids, Phys. Rev. B **36**, 381 (1987).
[83] L. Capriotti, F. Becca, A. Parola, and S. Sorella, Resonating valence bond wave functions for strongly frustrated spin systems, Phys. Rev. Lett. **87**, 097201 (2001).
[84] D. Tahara and M. Imada, Variational Monte Carlo method combined with quantum-number projection and multi-variable optimization, J. Phys. Soc. Jpn. **77**, 114701 (2008).
[85] Y. Nomura, A. S. Darmawan, Y. Yamaji, and M. Imada, Restricted Boltzmann machine learning for solving strongly correlated quantum systems, Phys. Rev. B **96**, 205152 (2017).
[86] S.-I. Amari, Natural gradient works efficiently in learning, Neural Comput. **10**, 251 (1998).
[87] S. Sorella, Green function Monte Carlo with stochastic reconfiguration, Phys. Rev. Lett. **80**, 4558 (1998).
[88] A. G. Baydin, B. A. Pearlmutter, A. A. Radul, and J. M. Siskind, Automatic differentiation in machine learning: A survey, J. Mach. Learn. Res. **18**, 1 (2018).
[89] D. P. Kingma and J. Ba, Adam: A method for stochastic optimization, in *3rd International Conference on Learning Representations* (2015).
[90] J. P. Provost and G. Vallee, Riemannian structure on manifolds of quantum states, Commun. Math. Phys. **76**, 289 (1980).
[91] J. Stokes, J. Izaac, N. Killoran, and G. Carleo, Quantum natural gradient, Quantum **4**, 269 (2020).
[92] F. Ferrari, F. Becca, and R. Valentí, Charge-density waves in kagome-lattice extended Hubbard models at the van Hove filling, Phys. Rev. B **106**, L081107 (2022).
[93] C. Gros, Superconductivity in correlated wave functions, Phys. Rev. B **38**, 931 (1988).
[94] N. Astrakhantsev, T. Westerhout, A. Tiwari, K. Choo, A. Chen, M. H. Fischer, G. Carleo, and T. Neupert, Broken-symmetry ground states of the Heisenberg model on the pyrochlore lattice, Phys. Rev. X **11**, 041021 (2021).
[95] R. Jastrow, Many-body problem with strong forces, Phys. Rev. **98**, 1479 (1955).
[96] M. Capello, F. Becca, F. Fabrizio, S. Sorella, and E. Tosatti, Variatinal description of Mott insulators, Phys. Rev. Lett. **94**, 026406 (2005).
[97] D. A. Huse and V. Elser, Simple variational wave functions for two-dimensional Heisenberg spin-1/2 antiferromagnets, Phys. Rev. Lett. **60**, 2531 (1988).
[98] S. Liang, B. Doucot, and P. W. Anderson, Some new variational resonating-valence-bond-type wave functions for the spin-1/2 antiferromagnetic Heisenberg model on a square lattice, Phys. Rev. Lett. **61**, 365 (1988).
[99] R. Kaneko, S. Morita, and M. Imada, Improved multi-variable variational Monte Carlo method examined by high-precision calculations of one-dimensional Hubbard model, J. Phys.: Conf. Ser. **454**, 012046 (2013).
[100] K. Ido, K. Yoshimi, T. Misawa, and M. Imada, Unconventional dual 1D-2D quantum spin liquid revealed by ab initio studies on organic solids family, npj Quantum Mater. **7**, 48 (2022).
[101] W.-J. Hu, F. Becca, A. Parola, and S. Sorella, Direct evidence for a gapless $\mathbb{Z}_2$ spin liquid by frustrating néel antiferromagnetism, Phys. Rev. B **88**, 060402 (2013).
[102] L. F. Tocchio, A. Montorsi, and F. Becca, Metallic and insulating stripes and their relation with superconductivity in the doped Hubbard model, SciPost Phys. **7**, 021 (2019).
[103] Y. Iqbal, W.-J. Hu, R. Thomale, D. Poilblanc, and F. Becca, Spin liquid nature in the Heisenberg J_1-J_2 triangular antiferromagnet, Phys. Rev. B **93**, 144411 (2016).
[104] F. Ferrari and F. Becca, Spectral signatures of fractionalization in the frustrated Heisenberg model on the square lattice, Phys. Rev. B **98**, 100405 (2018).
[105] V. Marino, F. Becca, and L. F. Tocchio, Stripes in the extended t-t' Hubbard model: A variational Monte Carlo analysis, SciPost Phys. **12**, 180 (2022).
[106] L. F. Tocchio, F. Becca, A. Parola, and S. Sorella, Role of backflow correlations for the nonmagnetic phase of the t-t' Hubbard model, Phys. Rev. B **78**, 041101 (2008).
[107] L. F. Tocchio, F. Becca, and C. Gros, Backflow correlations in the Hubbard model: An efficient tool for the study of the metal-insulator transition and the large-U limit, Phys. Rev. B **83**, 195138 (2011).
[108] Y. Iqbal, F. Becca, S. Sorella, and D. Poilblanc, Gapless spin-liquid phase in the kagome spin-1/2 Heisenberg antiferromagnet, Phys. Rev. B **87**, 060405 (2013).
[109] F. Ferrari and F. Becca, Dynamical structure factor of the J_1-J_2 Heisenberg model on the triangular lattice: Magnons, spinons, and gauge fields, Phys. Rev. X **9**, 031026 (2019).
[110] F. Ferrari, F. Becca, and J. Carrasquilla, Neural Gutzwiller-projected variational wave functions, Phys.

Rev. B **100**, 125131 (2019).
[111] Y. Nomura and M. Imada, Dirac-type nodal spin liquid revealed by refined quantum many-body solver using neural-network wave function, correlation ratio, and level spectroscopy, Phys. Rev. X **11**, 031034 (2021).
[112] G. Carleo, Y. Nomura, and M. Imada, Constructing exact representations of quantum many-body systems with deep neural networks, Nat. Commun. **9**, 5322 (2018).
[113] D. Luo and B. K. Clark, Backflow transformations via neural networks for quantum many-body wave functions, Phys. Rev. Lett. **122**, 226401 (2019).
[114] K. Choo, T. Neupert, and G. Carleo, Two-dimensional frustrated J_1-J_2 model studied with neural network quantum states, Phys. Rev. B **100**, 125124 (2019).
[115] Y. Nomura, Helping restricted Boltzmann machines with quantum-state representation by restoring symmetry, J. Phys.: Condens. Matter **33**, 174003 (2021).
[116] D. Luo, G. Carleo, B. K. Clark, and J. Stokes, Gauge equivariant neural networks for quantum lattice gauge theories, Phys. Rev. Lett. **127**, 276402 (2021).
[117] C. Roth and A. H. MacDonald, Group convolutional neural networks improve quantum state accuracy (2021), arXiv:2104.05085.
[118] O. Sharir, Y. Levine, N. Wies, G. Carleo, and A. Shashua, Deep autoregressive models for the efficient variational simulation of many-body quantum systems, Phys. Rev. Lett. **124**, 020503 (2020).
[119] M. Hibat-Allah, M. Ganahl, L. E. Hayward, R. G. Melko, and J. Carrasquilla, Recurrent neural network wave functions, Phys. Rev. Res. **2**, 023358 (2020).
[120] C. Roth, Iterative retraining of quantum spin models using recurrent neural networks (2020), arXiv:2003.06228.
[121] M. Hibat-Allah, E. M. Inack, R. Wiersema, R. G. Melko, and J. Carrasquilla, Variational neural annealing, Nature Mach. Intell. **3**, 952 (2021).
[122] M. Hibat-Allah, R. G. Melko, and J. Carrasquilla, Supplementing recurrent neural network wave functions with symmetry and annealing to improve accuracy, in *NeurIPS 2021 Workshop on Machine Learning and the Physical Sciences* (2021).
[123] D. Wu, R. Rossi, F. Vicentini, and G. Carleo, From tensor-network quantum states to tensorial recurrent neural networks, Phys. Rev. Res. **5**, L032001 (2023).
[124] Y.-H. Zhang and M. Di Ventra, Transformer quantum state: A multipurpose model for quantum many-body problems, Phys. Rev. B **107**, 075147 (2023).
[125] K. Choo, A. Mezzacapo, and G. Carleo, Fermionic neural-network states for ab-initio electronic structure, Nat. Commun. **11**, 2368 (2020).
[126] N. Yoshioka, W. Mizukami, and F. Nori, Solving quasiparticle band spectra of real solids using neural-network quantum states, Commun. Phys. **4**, 106 (2021).
[127] E. R. Bennewitz, F. Hopfmueller, B. Kulchytskyy, J. Carrasquilla, and P. Ronagh, Neural error mitigation of near-term quantum simulations, Nature Mach. Intell. **4**, 618 (2022).
[128] P. Jordan and E. Wigner, Über das Paulische äquivalenzverbot, Z. Angew. Phys. **47**, 631 (1928).
[129] S. B. Bravyi and A. Y. Kitaev, Fermionic quantum computation, Ann. Phys. **298**, 210 (2002).
[130] J. Stokes, J. R. Moreno, E. A. Pnevmatikakis, and G. Carleo, Phases of two-dimensional spinless lattice fermions with first-quantized deep neural-network quantum states, Phys. Rev. B **102**, 205122 (2020).
[131] J. R. Moreno, G. Carleo, A. Georges, and J. Stokes, Fermionic wave functions from neural-network constrained hidden states, Proc. Natl. Acad. Sci. **119**, e2122059119 (2022).
[132] S. Sorella, The phase diagram of the Hubbard model by variational auxiliary field quantum Monte Carlo (2021), arXiv:2101.07045.
[133] H. Xu, H. Shi, E. Vitali, M. Qin, and S. Zhang, Stripes and spin-density waves in the doped two-dimensional Hubbard model: Ground state phase diagram, Phys. Rev. Res. **4**, 013239 (2022).
[134] M. J. S. Beach, R. G. Melko, T. Grover, and T. H. Hsieh, Making Trotters sprint: A variational imaginary time ansatz for quantum many-body systems, Phys. Rev. B **100**, 094434 (2019).
[135] J. E. Hirsch, Two-dimensional Hubbard model: Numerical simulation study, Phys. Rev. B **31**, 4403 (1985).
[136] A. Griewank and A. Walther, *Evaluating Derivatives* (Society for Industrial and Applied Mathematics, 2008).
[137] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, 2010).
[138] D. Wecker, M. B. Hastings, and M. Troyer, Progress towards practical quantum variational algorithms, Phys. Rev. A **92**, 042303 (2015).
[139] K. Seki, T. Shirakawa, and S. Yunoki, Symmetry-adapted variational quantum eigensolver, Phys. Rev. A **101**, 052340 (2020).
[140] N. Astrakhantsev, G. Mazzola, I. Tavernelli, and G. Carleo, Phenomenological theory of variational quantum ground-state preparation, Phys. Rev. Res. **5**, 033225 (2023).
[141] J. Gacon, C. Zoufal, G. Carleo, and S. Woerner, Simultaneous perturbation stochastic approximation of the quantum Fisher information, Quantum **5**, 567 (2021).
[142] M. Qin, H. Shi, and S. Zhang, Benchmark study of the two-dimensional Hubbard model with auxiliary-field quantum Monte Carlo method, Phys. Rev. B **94**, 085103 (2016).
[143] W. Purwanto, W. A. Al-Saidi, H. Krakauer, and S. Zhang, Eliminating spin contamination in auxiliary-field quantum Monte Carlo: Realistic potential energy curve of F_2, J. Chem. Phys. **128**, 114309 (2008).
[144] H. F. Trotter, On the product of semi-groups of operators, Proc. Am. Math. Soc. **10**, 545 (1959).
[145] M. Suzuki, Generalized Trotter's formula and systematic approximants of exponential operators and inner derivations with applications to many-body problems, Commun. Math. Phys. **51**, 183 (1976).
[146] H. Shi and S. Zhang, Symmetry in auxiliary-field quantum Monte Carlo calculations, Phys. Rev. B **88**, 125132 (2013).
[147] 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).
[148] H. Shi and S. Zhang, Infinite variance in fermion quantum Monte Carlo calculations, Phys. Rev. E **93**, 033303 (2016).
[149] H. Nguyen, H. Shi, J. Xu, and S. Zhang, CPMC-Lab: A matlab package for constrained path Monte Carlo calculations, Comput. Phys. Commun. **185**, 3344 (2014).

[150] S. Zhang, J. Carlson, and J. E. Gubernatis, Constrained path quantum Monte Carlo method for fermion ground states, Phys. Rev. Lett. **74**, 3652 (1995).

[151] M. Qin, H. Shi, and S. Zhang, Coupling quantum Monte Carlo and independent-particle calculations: Self-consistent constraint for the sign problem based on the density or the density matrix, Phys. Rev. B **94**, 235119 (2016).

[152] J. Carlson, J. E. Gubernatis, G. Ortiz, and S. Zhang, Issues and observations on applications of the constrained-path Monte Carlo method to many-fermion systems, Phys. Rev. B **59**, 12788 (1999).

[153] 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, S. R. White, S. Zhang, B.-X. Zheng, Z. Zhu, and E. Gull, Solutions of the two-dimensional Hubbard model: Benchmarks and results from a wide range of numerical algorithms, Phys. Rev. X **5**, 041041 (2015).

[154] M. Iazzi and M. Troyer, Efficient continuous-time quantum Monte Carlo algorithm for fermionic lattice models, Phys. Rev. B **91**, 241118 (2015).

[155] L. Wang, M. Iazzi, P. Corboz, and M. Troyer, Efficient continuous-time quantum Monte Carlo method for the ground state of correlated fermions, Phys. Rev. B **91**, 235151 (2015).