www.sciencemag.org/content/355/6325/602/suppl/DC1

Supplementary Materials for

Solving the quantum many-body problem with artificial neural networks

Giuseppe Carleo* and Matthias Troyer

*Corresponding author. Email: gcarleo@ethz.ch

Published 10 February 2017, *Science* **355**, 602 (2017)
DOI: 10.1126/science.aag2302

This PDF file includes:

Other Supplementary Material for this manuscript includes the following:
(available at www.sciencemag.org/content/355/6325/602/suppl/DC1)

Code and Data Files

Materials and Methods

Stochastic Optimization for the Ground State

In the first part of our Article we have considered the goal of finding the best representation of the ground state of a given quantum Hamiltonian $\mathcal{H}$. The expectation value over our variational states $\mathrm{E}(\mathcal{W}) = \langle\Psi_{\mathrm{M}}|\mathcal{H}|\Psi_{\mathrm{M}}\rangle/\langle\Psi_{\mathrm{M}}|\Psi_{\mathrm{M}}\rangle$ is a function of the network weights $\mathcal{W}$. In order to obtain an optimal solution for which $\nabla\mathrm{E}(\mathcal{W}^{\star}) = 0$, several optimization approaches can be used.

Here, we have found convenient to adopt the Stochastic Reconfiguration (SR) method of Sorella et al. (*32*), which can be interpreted as an effective imaginary-time evolution in the variational subspace. Introducing the variational derivatives with respect to the k-th network parameter,

$$\mathcal{O}_{\mathrm{k}}(\mathcal{S}) \quad = \quad \frac{1}{\Psi_{\mathrm{M}}(\mathcal{S})}\partial_{\mathcal{W}_{\mathrm{k}}}\Psi_{\mathrm{M}}(\mathcal{S}), \qquad \text{(S1)}$$

as well as the so-called local energy

$$\mathrm{E}_{\mathrm{loc}}(\mathcal{S}) \quad = \quad \frac{\langle\mathcal{S}|\mathcal{H}|\Psi_{\mathrm{M}}\rangle}{\Psi_{\mathrm{M}}(\mathcal{S})}, \qquad \text{(S2)}$$

the SR updates at the p-th iteration are of the form

$$\mathcal{W}(p+1) = \mathcal{W}(p) - \gamma(p)S^{-1}(p)F(p), \qquad \text{(S3)}$$

where we have introduced the (Hermitian) covariance matrix

$$S_{kk'}(p) \quad = \quad \langle\mathcal{O}_k^{\star}\mathcal{O}_{k'}\rangle - \langle\mathcal{O}_k^{\star}\rangle\langle\mathcal{O}_{k'}\rangle, \qquad \text{(S4)}$$

the forces

$$F_k(p) \quad = \quad \langle E_{\mathrm{loc}}\mathcal{O}_k^{\star}\rangle - \langle E_{\mathrm{loc}}\rangle\langle\mathcal{O}_k^{\star}\rangle, \qquad \text{(S5)}$$

and a scaling parameter $\gamma(p)$. Since the covariance matrix can be non-invertible, S^{-1} denotes its Moore-Penrose pseudo-inverse. Alternatively, an explicit regularization can be applied, of the form $S_{k,k'}^{\mathrm{reg}} = S_{k,k'} + \lambda(p)\delta_{k,k'}S_{k,k}$. In our work, we have preferred the latter regularization, with a decaying parameter $\lambda(p) = \max(\lambda_0 b^p, \lambda_{\min})$ and typically take $\lambda_0 = 100$, $b = 0.9$ and $\lambda_{\min} = 10^{-4}$.

Initially the network weights $\mathcal{W}$ are set to some small random numbers and then optimized with the procedure outlined above. In Fig. S1 we show the typical behavior of the optimization algorithm, which systematically approaches the exact energy upon increasing the hidden units density α.

An alternative stochastic optimization method that can be readily used also in this context is the Stochastic Gradient Descent (SGD). This approach is particularly popular in the machine-learning community and it is based on noisy (and fast) estimations of the gradient of the function to be optimized. This approach has been used to minimize the ground-state energy of a many-body wave-function in Ref. (*43*). While the SGD approach is particularly appealing in terms of the favorable scaling with the number of parameters to be optimized, in the present work we have found the SR approach to be more robust for our purposes. In particular, a typical implementation of the SGD optimization requires a delicate fine-tuning of the learning rates and other hyper-parameters, with respect to which instead the SR approach is intrinsically less sensible. Nonetheless, it is also possible that more refined implementations of the SGD approach could perform significantly better than what found in our numerical experiments.

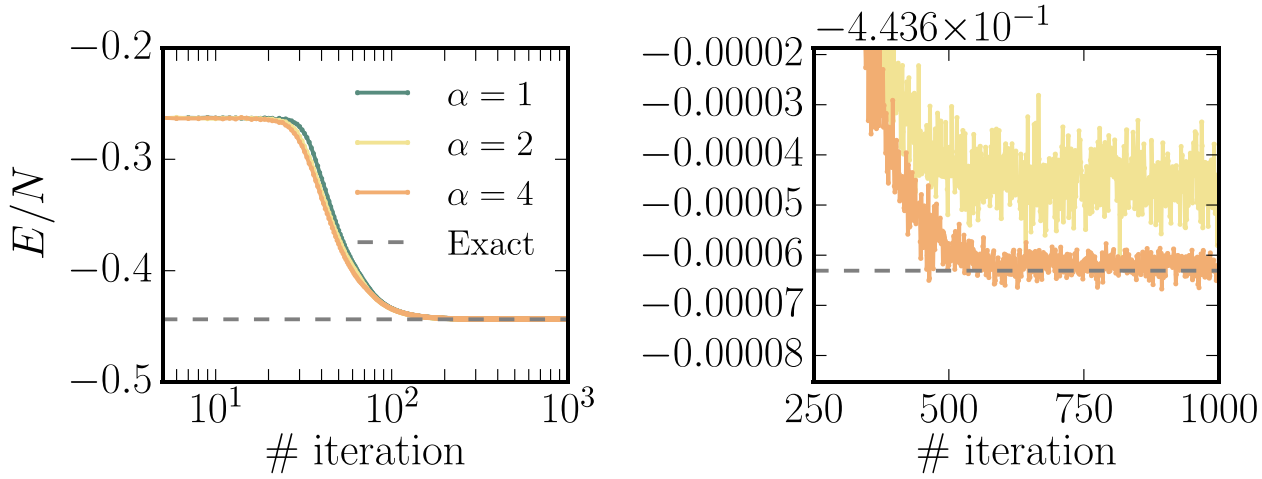

Figure S1.

Convergence properties of the stochastic optimization. Variational energy for the 1D Heisenberg model as a function of the Stochastic Reconfiguration iterates, and for different values of the hidden units density α. The system has PBC over a chain of $N = 40$ spins. The energy converges smoothly to the exact energy (dashed horizontal line) upon increasing α. In the Left panel, we show a complete view of the optimization procedure and on the Right panel a zoom in the neighborhood of the exact energy.

Time-Dependent Variational Monte Carlo

In the second part of our Article we have considered the problem of solving the many-body Schrödinger equation with a variational ansatz of the NQS form. This task can be efficiently accomplished by means of the Time-Dependent Variational Monte Carlo (t-VMC) method of Carleo et al.

In particular, the residuals

$$R(\dot{\mathcal{W}}(t)) = \text{dist}(\partial_t \Psi, -i\mathcal{H}\Psi) \qquad \text{(S6)}$$

are a function of the variational parameters derivatives, $\dot{\mathcal{W}}(t)$, and can be interpreted as the quantum distance between the exactly-evolved state and the variationally evolved one. Since in general we work with unnormalized quantum states, the correct Hilbert-space distance is given by the Fubini-Study metrics

$$\text{dist}_{\text{FS}}(\Phi, \Phi') \quad = \quad \arccos\sqrt{\frac{\langle\Phi'|\Phi\rangle\langle\Phi|\Phi'\rangle}{\langle\Phi'|\Phi'\rangle\langle\Phi|\Phi\rangle}}. \qquad \text{(S7)}$$

The explicit form of the residuals is then obtained considering $\Phi = \Psi + \delta\partial_t\Psi(\mathcal{W}(t))$ and $\Phi' = \Psi - i\delta\mathcal{H}\Psi(\mathcal{W}(t))$. Taking the lowest order in the time-step δ and explicitly minimizing $\text{dist}_{\text{FS}}(\Phi, \Phi')^2$, yields the equations of motion

$$\dot{\mathcal{W}}(t) \quad = \quad -iS^{-1}(t)F(t), \ (2) \qquad \text{(S8)}$$

where the correlation matrix and the forces are defined analogously to the previous section. In this case the diagonal regularization, in general, cannot be applied, and $S^{-1}(t)$ strictly denotes the Moore-Penrose pseudo-inverse.

The outlined procedure is globally stable, as also already proven for other wave functions in past works using the t-VMC approach. In Fig. S2 we show the typical behavior of the time-evolved physical properties of interest, which systematically approach the exact results when increasing α.

Finally notice that, as for the ground-state optimization, other stochastic approaches can be used to study the unitary dynamics. In particular, one can easily devise a SGD approach which, at each time, minimizes $R^2(t)$ with an on-line, noisy estimation of the wave-function gradients.

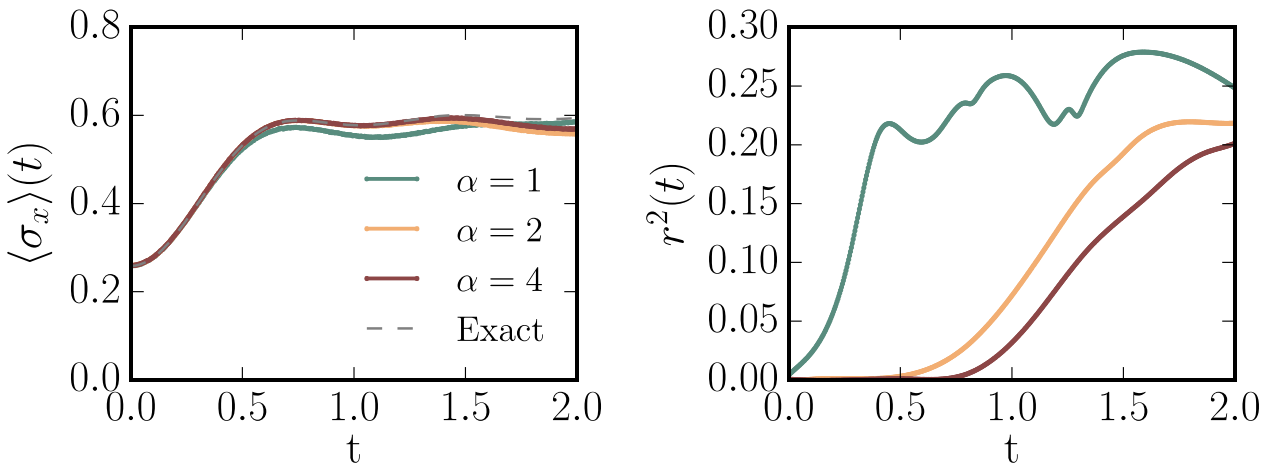

Figure S2

Convergence properties of the stochastic unitary evolution. Time-dependent expectation value of the transverse polarization along the x direction in the TFI model, for a quantum quench from $h_i = 1/2$ to the critical interaction $h_f = 1$. t-VMC results are shown for different values of the hidden units density α. The system has periodic boundary conditions over a chain of $N = 40$ spins. (Left panel) The variational curves for the expectation value of the transverse polarization converge smoothly to the exact solution (dashed line) upon increasing α. (Right panel) The relative residual error $r^2(t) = R^2(t)/D_0^2(t)$, where $D_0^2(t) = dist_{FS}(\Phi, \Phi - i\delta\mathcal{H})^2$ is shown for different values of the hidden unit density, and it is systematically reduced increasing α.

Efficient Stochastic Sampling

We complete the supplementary information giving an explicit expression for the variational derivatives previously introduced and of the overall computational cost of the stochastic sampling. We start rewriting the NQS in the form

$$\Psi_M(\mathcal{S}) \quad = \quad e^{\sum^i a_i \sigma_i^z} \times \Pi_{j=1}^{M} 2\cosh\theta_j(\mathcal{S}), \tag{S9}$$

with the effective angles $\theta_j(\mathcal{S}) = b_j + \sum^i W_{ij}\, \sigma_i^z$. The derivatives then read

$$\begin{aligned} \frac{1}{\Psi_M(\mathcal{S})}\partial_{a_i}\Psi_M(\mathcal{S}) &= \sigma_i^z, \\ \frac{1}{\Psi_M(\mathcal{S})}\partial_{b_j}\Psi_M(\mathcal{S}) &= \tanh\left[\theta_j(\mathcal{S})\right], \\ \frac{1}{\Psi_M(\mathcal{S})}\partial_{W_{ij}}\Psi_M(\mathcal{S}) &= \sigma_i^z \tanh\left[\theta_j(\mathcal{S})\right]. \end{aligned} \tag{S10}$$

In our stochastic procedure, we generate a Markov chain of many-body configurations $\mathcal{S}^{(1)} \rightarrow \mathcal{S}^{(2)} \rightarrow \cdots \mathcal{S}^{(P)}$ sampling the square modulus of the wave function $|\Psi_M(\mathcal{S})|^2$ for a given set of variational parameters. This task can be achieved through a simple Metropolis-Hastings algorithm (*44*), in which at each step of the Markov chain a random spin s is flipped and the new configuration accepted according to the probability

$$A(\mathcal{S}^{(k)} \rightarrow \mathcal{S}^{(k+1)}) \quad = \quad \min\left(1, \left|\frac{\Psi_M(\mathcal{S}^{(k+1)})}{\Psi_M(\mathcal{S}^{(k)})}\right|^2\right). \tag{S11}$$

In order to efficiently compute these acceptances, as well as the variational derivatives, it is useful to keep in memory look-up tables for the effective angles $\theta_j(\mathcal{S}^{(k)})$ and update them when a new configuration is accepted. These are updated according to

$$\theta_j(\mathcal{S}^{(k+1)}) \quad = \quad \theta_j(\mathcal{S}^{(k)}) - 2W_{sj}\sigma_s^z, \tag{S12}$$

when the spin s has been flipped. The overall cost of a Monte Carlo sweep (i.e. of $\mathcal{O}(N)$ single-spin flip moves) is therefore $\mathcal{O}(N \times M) = \mathcal{O}(\alpha N^2)$. Notice that the computation of the variational derivatives, and of the local energies after a Monte Carlo sweep, carry the same computational cost.

Iterative Solver

The most time-consuming part of both the SR optimization and of the t-VMC method is the solution of the linear systems (Eqs. S3 and S8) in the presence of a large number of variational parameters N_{var}. Explicitly forming the correlation matrix S, via stochastic sampling, has a dominant quadratic cost in the number of variational parameters, $\mathcal{O}(N_{\mathrm{var}}^2 \times N_{\mathrm{MC}})$, where N_{MC} denotes the number of Monte Carlo sweeps. However, this cost can be significantly reduced by means of iterative solvers which never form the covariance matrix explicitly. In particular, we adopt the MINRES-QLP method of Choi and Saunders (*45*), which implements a modified conjugate-gradient iteration based on Lanczos tridiagonalization. This method iteratively computes the pseudo-inverse S^{-1} within numerical precision. The backbone of iterative solvers is, in general, the application of the matrix to be inverted to a given (test) vector. This can be efficiently implemented due to the product structure of the covariance matrix, and determines a dominant complexity of $\mathcal{O}(N_{\mathrm{var}} \times N_{\mathrm{MC}})$ operations for the sparse solver. For example, in the most challenging case when translational symmetry is absent, we have $N_{\mathrm{var}} = \alpha N^2$, and the dominant computational cost for solving (Eqs. S3 and S8) is in line with the complexity of the previously described Monte Carlo sampling.

Implementing Symmetries

Very often, physical Hamiltonians exhibit intrinsic symmetries which must be satisfied also by their ground- and dynamically-evolved quantum states. These symmetries can be conveniently used to reduce the number of variational parameters in the NQS.

Let us consider a symmetry group defined by a set of linear transformations T_s, with $s = 1, \dots S$, such that the spin configurations transform according to $T_s \sigma^z = \widetilde{\sigma}^z(s)$. We can enforce the NQS representation to be invariant under the action of T defining

$$\Psi_\alpha(\mathcal{S};\mathcal{W}) = \sum_{\{h_{i,s}\}} \exp\left[\sum_f^\alpha a^{(s)} \sum_s^S \sum_j^N \widetilde{\sigma_j^z}(s) + \right.$$
$$\left. + \sum_f^\alpha b^{(s)} \sum_s^S h_{f,s} + \sum_f^\alpha \sum_s^S h_{f,s} \sum_j^N W_j^{(f)} \widetilde{\sigma_j^z}(s)\right], \tag{S13}$$

where the network weights have now a different dimension with respect to the standard NQS. In particular, $a^{(f)}$ and $b^{(f)}$ are vectors in the feature space with $f = 1, \dots \alpha_s$ and the connectivity matrix $W_j^{(f)}$ contains $\alpha_s \times N$ elements. Notice that this expression corresponds effectively to a standard NQS with $M = S \times \alpha_s$ hidden variables. Tracing out explicitly the hidden variables, we obtain

$$\Psi_\alpha(\mathcal{S};\mathcal{W}) = e^{\sum_{f,s,j} a^{(f)} \widetilde{\sigma_j^z}(s)} \times$$
$$\times \Pi_f \Pi_s 2\cosh\left[b^{(f)} + \sum_j^N W_j^{(f)} \widetilde{\sigma_j^z}(s)\right]. \tag{S14}$$

In the specific case of site-translation invariance, we have that the symmetry group has an orbit of $S = N$ elements. For a given feature f, the matrix $W_j^{(f)}$ can be seen as a filter acting on the N translated copies of a given spin configuration. In other words, each feature has a pool of N associated hidden variables that act with the same filter on the symmetry-transformed images of the spins.

Description of Supplementary Computer Code and Data Files

In the Supplementary Material, we also provide open-source computer programs meant to demonstrate the main algorithms and to allow replication of our results by interested readers.

In particular, we provide a package (Nqs.zip) containing a C++ code capable of sampling from an arbitrary NQS wave-function, measure the expectation value of a given Hamiltonian, and output the generated many-body configuration samples. The sampling is realized implementing the ideas presented in this Supplementary Material (Eqs. S9 to S12).

In the same package, we also provide several data files corresponding to various NQS wave-functions. In particular, we provide the optimized ground-state wave-functions as well as the time-dependent wave-functions corresponding to the results shown in the main Paper. In addition to those, we also provide further NQS wave-functions not explicitly considered in the main Paper.

As detailed in the file 'README' of the package, the interested reader can use one of the many available NQS to compute the expectation value of the energy and/or to generate many-body configurations $\mathcal{S}^{(1)} \rightarrow \mathcal{S}^{(2)} \rightarrow \cdots \mathcal{S}^{(P)}$ sampling the square modulus of the wave function $|\Psi_M(\mathcal{S})|^2$. The latter are particularly useful to reconstruct expectation values of arbitrary observables diagonal in the computational basis.

References and Notes

1. D. Ceperley, B. Alder, Quantum Monte Carlo. *Science* **231**, 555–560 (1986). doi:10.1126/science.231.4738.555 Medline

2. W. M. C. Foulkes, L. Mitas, R. J. Needs, G. Rajagopal, Quantum Monte Carlo simulations of solids. *Rev. Mod. Phys.* **73**, 33–83 (2001). doi:10.1103/RevModPhys.73.33

3. J. Carlson, S. Gandolfi, F. Pederiva, S. C. Pieper, R. Schiavilla, K. E. Schmidt, R. B. Wiringa, Quantum Monte Carlo methods for nuclear physics. *Rev. Mod. Phys.* **87**, 1067–1118 (2015). doi:10.1103/RevModPhys.87.1067

4. S. R. White, Density matrix formulation for quantum renormalization groups. *Phys. Rev. Lett.* **69**, 2863–2866 (1992). doi:10.1103/PhysRevLett.69.2863 Medline

5. S. Rommer, S. Ostlund, Class of ansatz wave functions for one-dimensional spin systems and their relation to the density matrix renormalization group. *Phys. Rev. B* **55**, 2164–2181 (1997). doi:10.1103/PhysRevB.55.2164

6. U. Schollwöck, The density-matrix renormalization group in the age of matrix product states. *Ann. Phys.* **326**, 96–192 (2011). doi:10.1016/j.aop.2010.09.012

7. R. Orús, A practical introduction to tensor networks: Matrix product states and projected entangled pair states. *Ann. Phys.* **349**, 117–158 (2014). doi:10.1016/j.aop.2014.06.013

8. F. Verstraete, V. Murg, J. I. Cirac, Matrix product states, projected entangled pair states, and variational renormalization group methods for quantum spin systems. *Adv. Phys.* **57**, 143–224 (2008). doi:10.1080/14789940801912366

9. K. H. Marti, B. Bauer, M. Reiher, M. Troyer, F. Verstraete, Complete-graph tensor network states: A new fermionic wave function ansatz for molecules. *New J. Phys.* **12**, 103008 (2010). doi:10.1088/1367-2630/12/10/103008

10. M. Troyer, U.-J. Wiese, Computational complexity and fundamental limitations to fermionic quantum Monte Carlo simulations. *Phys. Rev. Lett.* **94**, 170201 (2005). Medline

11. A. Polkovnikov, K. Sengupta, A. Silva, M. Vengalattore, Colloquium: Nonequilibrium dynamics of closed interacting quantum systems. *Rev. Mod. Phys.* **83**, 863–883 (2011). doi:10.1103/RevModPhys.83.863

12. J. Eisert, M. Friesdorf, C. Gogolin, Quantum many-body systems out of equilibrium. *Nat. Phys.* **11**, 124–130 (2015). doi:10.1038/nphys3215

13. A. Montorsi, *The Hubbard Model: A Collection of Reprints* (World Scientific, 1992).

14. D. J. Thouless, *The Quantum Mechanics of Many-Body Systems* (Reprint of the Academic Press, ed. 2, 1972).

15. J. K. Freericks, B. K. Nikolić, O. Frieder, The nonequilibrium quantum many-body problem as a paradigm for extreme data science. *Int. J. Mod. Phys. B* **28**, 1430021 (2014). doi:10.1142/S0217979214300217

16. G. E. Hinton, R. R. Salakhutdinov, Reducing the dimensionality of data with neural networks. *Science* **313**, 504–507 (2006). doi:10.1126/science.1127647 Medline

17. Y. LeCun, Y. Bengio, G. Hinton, Deep learning. *Nature* **521**, 436–444 (2015). doi:10.1038/nature14539 Medline

18. D. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. van den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot, S. Dieleman, D. Grewe, J. Nham, N. Kalchbrenner, I. Sutskever, T. Lillicrap, M. Leach, K. Kavukcuoglu, T. Graepel, D. Hassabis, Mastering the game of Go with deep neural networks and tree search. *Nature* **529**, 484–489 (2016). doi:10.1038/nature16961 Medline

19. S. S. Schoenholz, E. D. Cubuk, D. M. Sussman, E. Kaxiras, A. J. Liu, A structural approach to relaxation in glassy liquids. *Nat. Phys.* **12**, 469–471 (2016). doi:10.1038/nphys3644

20. J. Carrasquilla, R. G. Melko, https://arxiv.org/abs/1605.01735 (2016).

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

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

23. M. H. Amin, E. Andriyash, J. Rolfe, B. Kulchytskyy, R. Melko, https://arxiv.org/abs/1601.02036 (2016).

24. W. L. McMillan, Ground state of liquid He^4. *Phys. Rev.* **138**, A442–A451 (1965). doi:10.1103/PhysRev.138.A442

25. G. Carleo, F. Becca, M. Schiró, M. Fabrizio, Localization and glassy dynamics of many-body quantum systems. *Sci. Rep.* **2**, 243 (2012). doi:10.1038/srep00243 Medline

26. G. Carleo, F. Becca, L. Sanchez-Palencia, S. Sorella, M. Fabrizio, Light-cone effect and supersonic correlations in one- and two-dimensional bosonic superfluids. *Phys. Rev. A* **89**, 031602 (2014). doi:10.1103/PhysRevA.89.031602

27. A. N. Kolmogorov, On the representation of continuous functions of several variables by superpositions of continuous functions of a smaller number of variables. *Dokl. Akad. Nauk SSSR* **108**, 179–182 (1961).

28. K. Hornik, Approximation capabilities of multilayer feedforward networks. *Neural Netw.* **4**, 251–257 (1991). doi:10.1016/0893-6080(91)90009-T

29. N. Le Roux, Y. Bengio, Representational power of restricted Boltzmann machines and deep belief networks. *Neural Comput.* **20**, 1631–1649 (2008). doi:10.1162/neco.2008.04-07-510 Medline

30. K. Sohn, H. Lee, "Learning invariant representations with local transformations," in *Proceedings of the 29th International Conference on Machine Learning*, Edinburgh, Scotland, 26 June to 1 July 2012 (Omnipress, 2012), pp. 1311–1318.

31. M. Norouzi, M. Ranjbar, G. Mori, "Stacks of convolutional restricted Boltzmann machines for shift-invariant feature learning," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 20 to 25 June 2009 (IEEE, 2009), pp. 2735–2742.

32. S. Sorella, M. Casula, D. Rocca, Weak binding between two aromatic rings: Feeling the van der Waals attraction by quantum Monte Carlo methods. *J. Chem. Phys.* **127**, 014105 (2007). doi:10.1063/1.2746035 Medline

33. M. Dolfi, B. Bauer, S. Keller, A. Kosenkov, T. Ewart, A. Kantian, T. Giamarchi, M. Troyer, Matrix product state applications for the ALPS project. *Comput. Phys. Commun.* **185**, 3430–3440 (2014). doi:10.1016/j.cpc.2014.08.019

34. A. W. Sandvik, Finite-size scaling of the ground-state parameters of the two-dimensional Heisenberg model. *Phys. Rev. B* **56**, 11678–11690 (1997). doi:10.1103/PhysRevB.56.11678

35. F. Mezzacapo, N. Schuch, M. Boninsegni, J. I. Cirac, Ground-state properties of quantum many-body systems: Entangled-plaquette states and variational Monte Carlo. *New J. Phys.* **11**, 083026 (2009). doi:10.1088/1367-2630/11/8/083026

36. M. Lubasch, J. I. Cirac, M.-C. Bañuls, Algorithms for finite projected entangled pair states. *Phys. Rev. B* **90**, 064425 (2014). doi:10.1103/PhysRevB.90.064425

37. P. A. M. Dirac, Note on exchange phenomena in the Thomas atom. *Math. Proc. Camb. Philos. Soc.* **26**, 376–385 (1930). doi:10.1017/S0305004100016108

38. Y. I. Frenkel, *Wave Mechanics: Advanced General Theory* (International Series of Monographs on Nuclear Energy: Reactor Design Physics, The Clarendon Press, 1934).

39. S. R. White, A. E. Feiguin, Real-time evolution using the density matrix renormalization group. *Phys. Rev. Lett.* **93**, 076401 (2004). doi:10.1103/PhysRevLett.93.076401 Medline

40. G. Vidal, Efficient simulation of one-dimensional quantum many-body systems. *Phys. Rev. Lett.* **93**, 040502 (2004). doi:10.1103/PhysRevLett.93.040502 Medline

41. A. J. Daley, C. Kollath, U. Schollwock, G. Vidal, Time-dependent density-matrix renormalization-group using adaptive effective Hilbert spaces. *J. Stat. Mech.* **2004**, P04005 (2004).

42. B. Bauer, L. D. Carr, H. G. Evertz, A. Feiguin, J. Freire, S. Fuchs, L. Gamper, J. Gukelberger, E. Gull, S. Guertler, A. Hehn, R. Igarashi, S. V. Isakov, D. Koop, P. N. Ma, P. Mates, H. Matsuo, O. Parcollet, G. Pawłowski, J. D. Picon, L. Pollet, E. Santos, V. W. Scarola, U. Schollwöck, C. Silva, B. Surer, S. Todo, S. Trebst, M. Troyer, M. L. Wall, P. Werner, S. Wessel, The ALPS project release 2.0: Open source software for strongly correlated systems. *J. Stat. Mech.* **2011**, P05001 (2011). doi:10.1088/1742-5468/2011/05/P05001

43. A. Harju, B. Barbiellini, S. Siljamäki, R. M. Nieminen, G. Ortiz, Stochastic gradient approximation: An efficient method to optimize many-body wave functions. *Phys. Rev. Lett.* **79**, 1173–1177 (1997). doi:10.1103/PhysRevLett.79.1173

44. N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, E. Teller, Equation of state calculations by fast computing machines. *J. Chem. Phys.* **21**, 1087–1092 (1953). doi:10.1063/1.1699114

45. S.-C. T. Choi, M. A. Saunders, Algorithm 937: MINRES-QLP for symmetric and Hermitian linear equations and least-squares problems. *ACM Trans. Math. Softw.* **40**, 40 (2014). Medline