Machine Learning Bell Nonlocality in Quantum Many-body Systems

Dong-Ling Deng

Condensed Matter Theory Center and Joint Quantum Institute,
Department of Physics, University of Maryland, College Park, MD 20742-4111, USA

Machine learning, the core of artificial intelligence and big data science, is one of today's most rapidly growing interdisciplinary fields. Recently, its tools and techniques have been adopted to tackle intricate quantum many-body problems. In this work, we introduce machine learning techniques to the detection of quantum nonlocality in many-body systems, with a focus on the restricted-Boltzmann-machine (RBM) architecture. Using reinforcement learning, we demonstrate that RBM is capable of finding the maximum quantum violations of multipartite Bell inequalities with given measurement settings. Our results build a novel bridge between computer-science-based machine learning and quantum many-body nonlocality, which will benefit future studies in both areas.

Nonlocality is one of the most fascinating and enigmatic features of quantum mechanics that denies any local realistic description of our world [1, 2]. It represents the most profound departure of quantum from classical physics and has been experimentally confirmed in a number of systems through violations of Bell inequalities [3–18]. In addition to its fundamental interest, in practice nonlocality is the key resource for device-independent quantum technologies, such as secure key distribution [19–21] or certifiable random number generators [22–26]. Thus, characterizing and detecting nonlocality is one of the central problems in both quantum information theory and experiment. Here, we introduce machine learning, a branch of computer science [27–29], to the detection of quantum nonlocality (see Fig. 1 for a pictorial illustration).

For quantum many-body systems, whereas entanglement has been extensively studied [30], nonlocality remains rarely explored. Mathematically, it has been proved that the complete characterization of classical correlations for a generic many-body system is an NP-hard problem [31]. Nevertheless, an incomplete list of multipartite Bell inequalities with high-order correlation functions has indeed been discovered for a long time [2]. More recently, Bell inequalities with only two-body correlators were constructed [32–36] and multipartite nonlocality has been demonstrated experimentally in a Bose-Einstein condensate by violating one of these inequalities [37]. This sparks a new wave of interest in the study of nonlocality in many-body systems.

A particular question of both theoretical and experimental relevance is that for a given multipartite Bell inequality, how to obtain its quantum violation in a numerical simulation? To tackle this problem, one has to face at least two challenges. First of all, the Hilbert space of a quantum many-body system grows exponentially with the system size and a complete description of its state requires an exponential amount of information in general, rendering the computation of the quantum expectation value corresponding to the inequality a formidably demanding task. Second, the measurement settings for each party involved in a Bell experiment is arbitrary in principle, making the problem even more complicated. In fact, it has been shown that the computation of the maximum violation of a multipartite Bell inequality is an NP-problem [38]. In this paper, we will not attempt to solve this problem
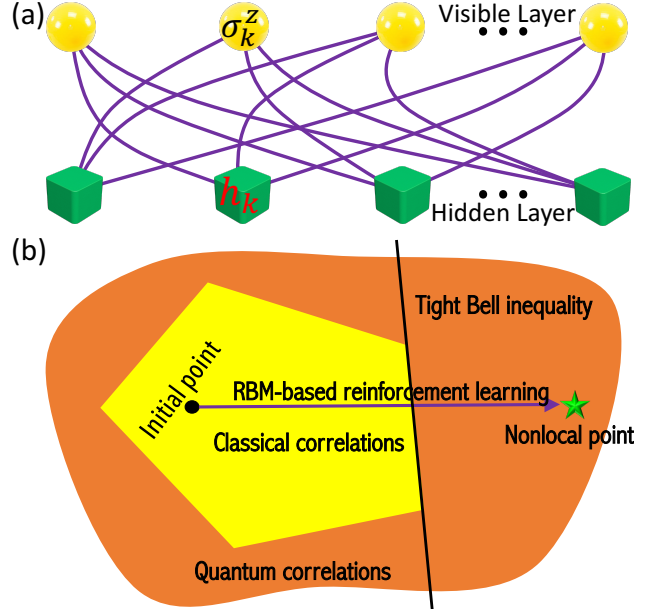

FIG. 1. (a) A sketch of the restricted-Boltzmann-machine (RBM) representation of quantum many-body states. For each spin configuration $\Xi = (\sigma_1^z, \sigma_2^z, \cdots, \sigma_N^z)$, the artificial neural network outputs its corresponding coefficient $\Phi(\Xi)$. (b) A pictorial illustration of the essential idea of machine learning Bell nonlocality in quantum many-body systems. The set of all classical correlations forms a high-dimensional polytope (yellow region), which is a subset of the quantum-correlation set that consists of all possible correlations allowed by quantum mechanics. The black line represents a tight Bell inequality (facet of the polytope). We start with a random RBM, which typically shows only classical correlations (in the sense that it does not violate a given Bell inequality). We then optimize its internal parameters, through reinforcement learning, so as to violate the Bell inequality maximally.

completely, which is implausible due to the NP complexity. Instead, we study a simplified scenario where the given multipartite Bell inequality only involves a polynomial number of correlation functions and the measurement settings for each party are restricted (due to experimental requirements, for instance) and preassigned. We show that machine learning may provide an unprecedented perspective for solving this simpli-

fied, but still sufficiently intricate, quantum many-body problem. Within physics, applications of machine-learning techniques have recently been invoked in various contexts [39–72], such as black hole detection [56], gravitational lenses [57] and wave analysis [58, 59], material design [60], glassy dynamics [61], Monte Carlo simulation [62, 63], topological codes [64], quantum machine learning [72], and topological phases and phase transitions [41–50], etc. Here, we apply machine learning to the detection of quantum many-body nonlocality, focusing on one of the simplest stochastic neural networks for unsupervised learning—the restricted Boltzmann machine (RBM) [73–75] as an example. We demonstrate, through three concrete examples, that RBM-based reinforcement learning is capable of finding the maximum quantum violations of multipartite Bell inequalities with given measurement settings. Our method works for generic Bell inequalities that involve a polynomial number (in system size) of correlation functions, independent of dimensionality, the order of the correlations, or whether the correlation functions are short-range or not. Our results showcase the exceptional power of machine learning in the detection of quantum nonlocality for many-body systems, thus would provide a valuable guide for both theory and experiment.

To begin with, let us first briefly introduce the RBM representation of quantum states [39] and the general recipe for machine learning Bell nonlocality. We consider a quantum system with N spin-$\frac{1}{2}$ particles (qubits) $\Xi = (\sigma_1, \sigma_2, \cdots \sigma_N)$ and use a RBM neural network to describe its many-body wavefunction $\Phi(\Xi)$. A RBM consists of two layers: one called visible layer with N nodes (visible neurons), corresponding to the physical spins; the other called hidden layer with M auxiliary nodes (hidden neurons). The hidden neurons are coupled to the visible ones, but there is no coupling among neurons in the same layer, as schematically illustrated in Fig. 1(a). By tracing out the hidden neurons, we obtain a RBM representation of a quantum many-body state [39]:

$$\Phi_M(\Xi, \Omega) = \sum_{\{h_k\}} e^{\sum_k a_k \sigma_k^z + \sum_{k'} b_{k'} h_{k'} + \sum_{kk'} W_{k'k} h_{k'} \sigma_k^z}, \quad (1)$$

where $\Omega \equiv (a, b, W)$ are internal parameters that fully specify the RBM neural network and $\{h_k\} = \{-1, 1\}^M$ denotes the possible hidden neuron configurations. We mention that any quantum state can be approximated to arbitrary accuracy by the above RBM representation, as long as the number of hidden neurons is large enough [76–78]. It is shown in Ref. [44] that RBM can represent topological states, either symmetry protected or with intrinsic topological order, in an exact and efficient fashion, and the entanglement properties of RBM states are extensively studied in Ref [79].

We consider a standard Bell experiment in which N parties each can freely choose to perform one of K possible measurements $\mathcal{M}_k^{(i)}$ ($i = 1, \cdots, N$ and $k = 0, \cdots, K-1$) with binary outcomes ± 1. We describe the observed correlations by using a collection of expectation values of correlators $\langle \mathcal{M}_{k_1}^{(i_1)} \cdots \mathcal{M}_{k_\alpha}^{(i_\alpha)} \rangle$ and we say that the correlations are

classical when they can be simulated with only shared classical information between parties (or in other words, can be described by a local hidden variable theory [80]). Classical correlations form a high-dimensional (exponential in N) polytope $\mathbb{P}$, which is a bounded convex set with a finite number of extreme points. Each facet of $\mathbb{P}$ corresponds to a tight Bell inequality and correlations that fall outside of $\mathbb{P}$ will violate a Bell inequality and thus manifest nonlocality. We write the Bell inequalities in a generic form: $\mathcal{I} \geq \mathcal{B}^{(c)}$, where $\mathcal{I}$ is a function of the expectation values of the correlators and $\mathcal{B}^{(c)}$ is the classical bound. Within this framework, our general recipe for machine learning of nonlocality through violation of a given Bell inequality is as follows: we begin with a random RBM state, whose observed correlations may or may not fall inside $\mathbb{P}$, but typically do not violate the given inequality; we then use a reinforcement learning scheme recently introduced by Carleo and Troyer [39] to iteratively optimize the internal parameters, such that the minimal expectation value of $\mathcal{I}$ within quantum mechanics will be achieved. If the minimal value is smaller than $\mathcal{B}^{(c)}$, the Bell inequality is maximally violated with a given measurement setting and nonlocality is detected. A pictorial illustration of the classical polytope, a tight Bell inequality, and the essential idea of machine learning Bell nonlocality is shown in Fig. 1(b).

One may also choose another measurement setting and run the same process to obtain the maximal violation for this setting. In order to obtain the maximal violation of the Bell inequality for all measurement settings, one can just scan all possible settings and do the same process repeatedly. We mention that an alternative and more efficient way is to regard all the parameters that specify the measurements as variational parameters as well (on an equal footing as the RBM parameters Ω) and optimize them together with the RBM parameters using a similar reinforcement learning procedure. But this is more technically involved. Here, we will only focus on the former cases with fixed measurement settings (parameters for measurements are preassigned) for simplicity and leave the later approach for future studies.

To show more precisely how this RBM-based reinforcement learning protocol works, we give three concrete examples. The first one concerns Bell inequalities with only short-range two-body correlators in one dimension (1D). This is a case where traditional methods, such as density-matrix renormalization group (DMRG), also work remarkably well [81–83]. We compare our RBM results with that from exact diagonalization (ED) for small system size N and DMRG for larger N, and find that they agree excellently. This validates the effectiveness of our RBM approach. The second and third examples are about Bell inequalities with, either all-to-all but two-body or multipartite, correlators. These examples are beyond the capacity of the DMRG or ED methods for large system sizes and show a striking advantage of RBMs in detecting many-body nonlocality.

Bell inequalities with short-range two-body correlators.—Now, let us consider a 1D system with N (an even integer) qubits. A Bell inequality involving only two-body
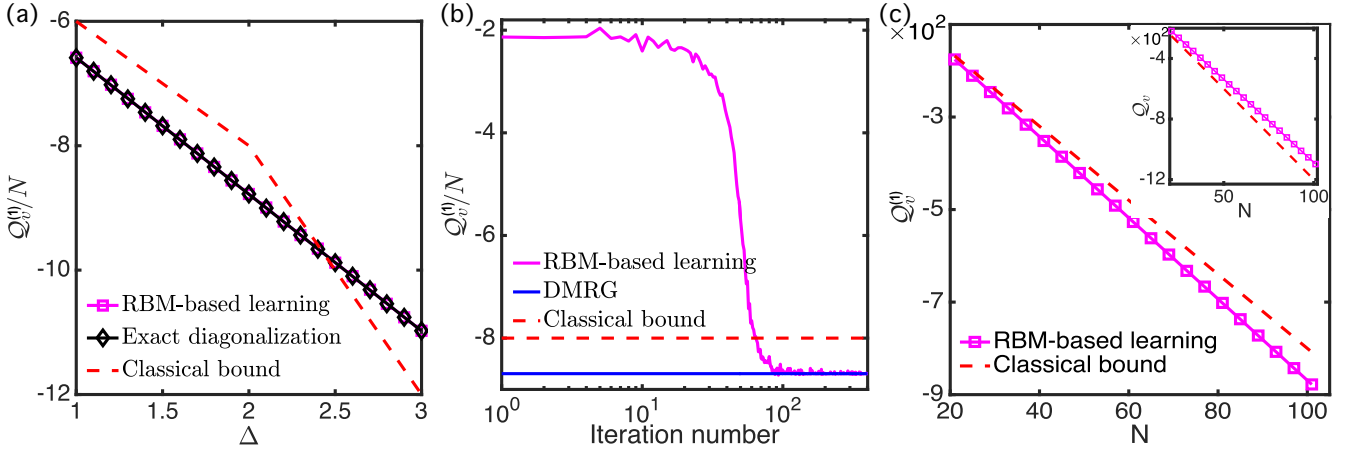
FIG. 2. RBM-based reinforcement learning of many-body Bell nonlocality through quantum violations of the Ineq. (2). The red dashed lines represent the classical bounds $\mathcal{B}_1^{(c)}$, the regions below which show quantum nonlocality and thus are not attainable by any local hidden variable models. We denote the quantum expectation value corresponding to $\mathcal{I}_1$ as $\mathcal{Q}_v^{(1)}$, and without loss of generality, we have fixed $\delta = 0.9$ throughout this figure for simplicity. (a) A comparison between results from RBM and exact diagonalization (ED) for a small system size $N = 20$. The results match each other very well. (b) The obtained quantum expectation value as a function of the iteration number of the learning process, for a larger system size $N = 100$. For this particular learning process, $\mathcal{Q}_v^{(1)}$ begins to cross the classical bound $\mathcal{B}_1^{(c)}$ after 65 iterations, and all the RBM states thereafter violate Ineq. (2) and thus show many-body nonlocality. As the iteration number increases, $\mathcal{Q}_v^{(1)}$ converges quickly to the value computed from DMRG [84]. (c) RBM learned $\mathcal{Q}_v^{(1)}$ as a function of N for $\Delta = 2$. The inset shows the result for $\Delta = 3$, where no quantum violation is observed.

correlators with nearest-neighbor couplings has recently been obtained through a dynamic programming procedure [34]:

$$\mathcal{I}_1 = \sum_{k=0}^{N/2-1} (1+\delta)\mathcal{I}_{\text{even}}^{(k)} + (1-\delta)\mathcal{I}_{\text{odd}}^{(k)} \geq \mathcal{B}_1^{(c)}, \qquad (2)$$

where $\mathcal{I}_{\text{even}}^{(k)} = \sum_{a=0}^{4} \sum_{b=0}^{3} \Lambda_{a,b}(\Delta) \langle \mathcal{M}_a^{(2k)} \mathcal{M}_b^{(2k+1)} \rangle$ and $\mathcal{I}_{\text{odd}}^{(k)} = \mathcal{I}_{\text{even}}^{(k)}(2k \rightarrow 2k+1)$ with $\Lambda(\Delta)$ a four-by-three matrix [85]; $\mathcal{B}_1^{(c)}$ is the classical bound depending on the real parameters δ and Δ: $\mathcal{B}_1^{(c)} = -(4+2|\Delta|)N$ for $|\Delta| \leq 2$ and $|\delta| \leq 1$, and $\mathcal{B}_1^{(c)} = -4|\Delta|N$ for $2 \leq |\Delta| \leq 3$ and $|\delta| \leq 1$. By choosing the measurement settings properly [86], the Bell operator corresponding to the above inequality reduces to the following XXZ-type Hamiltonian:

$$H = \sum_{k=0}^{N-1} g_k(\delta)[\hat{\sigma}_k^x \hat{\sigma}_{k+1}^x + \hat{\sigma}_k^y \hat{\sigma}_{k+1}^y + \Delta \hat{\sigma}_k^z \hat{\sigma}_{k+1}^z],$$

where $g_k(\delta) = 4[1 + (-1)^k]/\sqrt{3}$, and $\hat{\sigma}^x$, $\hat{\sigma}^y$ and $\hat{\sigma}^z$ are the usual Pauli matrices. For this particular measurement settings, the maximal quantum violation of Ineq. (2) corresponds to the ground state energy of H and can be calculated using DMRG, as already discussed in Ref. [34]. Here, we use the above introduced reinforcement learning method to obtain the same quantum violation.

Our results are plotted in Fig. 2. In Fig. 2(a), we compare our results with that from ED for $N = 20$. As shown in this figure, the RBM result matches the ED result very well [84]. We find that the quantum expectation value of $\mathcal{I}_1$ (denoted by $\mathcal{Q}_v^{(1)}$) decreases approximately linearly as we increase Δ.

There is a critical value $\Delta \approx 2.4$, after which no quantum violation will be observed. In Fig. 2(b), we show the convergence of the RBM learning and compare the obtained results with that of DMRG. We find that the initial random RBM states typically do not violate the Ineq. 2, but as the learning process goes on, $\mathcal{Q}_v^{(1)}$ will decrease and begin to violate the inequality after a certain critical iteration number. As the iteration number increases further, $\mathcal{Q}_v^{(1)}$ quickly converges to the DMRG value, validating the effectiveness of the RBM method. Fig. 2(c) shows the converged $\mathcal{Q}_v^{(1)}$ as a function of N. We find that $\mathcal{Q}_v^{(1)}$ decreases linearly with increasing N for the chosen parameters $(\delta, \Delta) = (0.9, 2)$. For $\Delta = 2$, the slope for $\mathcal{Q}_v^{(1)}$ is smaller than that of $\mathcal{B}_1^{(c)}$, thus the larger N the stronger quantum violations. For $\Delta = 3$, no violation is observed for all N, which is consistent with the results in Ref. [34].

Bell inequalities with all-to-all two-body correlators.—As the second example to show the power of RBM in detecting nonlocality, we consider the following Bell inequality introduced by Tura *et al* [32], which involves *all-to-all* two-body correlators and thus are beyond the scope of DMRG:

$$\mathcal{I}_2 = -2S_0 - S_{01} + \frac{1}{2}(S_{00} + S_{11}) \geq \mathcal{B}_2^{(c)}, \qquad (3)$$

where the one- and two- body correlators are defined as: $S_a = \sum_{k=1}^{N} \langle \mathcal{M}_a^{(k)} \rangle$ and $S_{ab} = \sum_{k \neq l}^{N} \langle \mathcal{M}_a^{(k)} \mathcal{M}_b^{(l)} \rangle$ $(a, b = 0, 1)$, and the classical bound $\mathcal{B}_2^{(c)} = -2N$. This inequality has been used in a recent experiment to demonstrate many-body nonlocality of about 480 atoms in a Bose-Einstein condensate [37]. For permutationally-symmetric states, its quantum violations were numerically studied in Ref. [32]. Here, we find
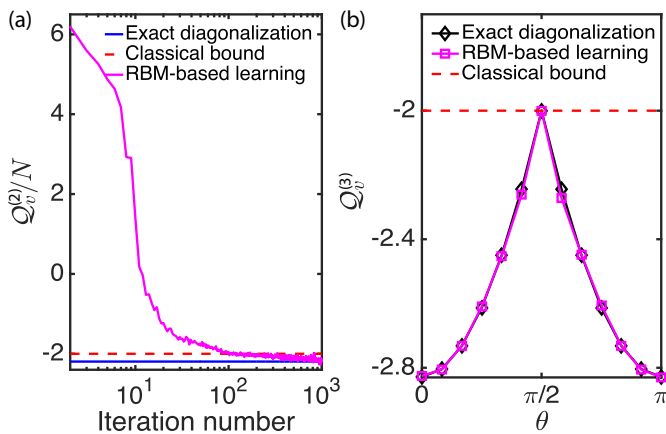
FIG. 3. (a) RBM-learned quantum expectation value ($\mathcal{Q}_v^{(2)}$) as a function of the iteration number, for a typical random sample of measurement angles (see [84]). (b) RBM-learned quantum violations ($\mathcal{Q}_v^{(3)}$) of Ineq. (4) as a function of measurement angle θ. In both (a) and (b), the system size is fixed to be $N = 20$.

that, using the RBM approach, one can obtain the same maximal violations readily if one chooses a permutation-invariant neural network. More interestingly, we find that the RBM approach also works for the cases where the permutation symmetry is released. To this end, we consider a scenario where the measurement settings are chosen as: $\mathcal{M}_0^{(k)} = \sigma^z$ and $\mathcal{M}_1^{(k)} = \cos\theta_k\sigma^z + \sin\theta_k\sigma^x$ ($k = 1, 2, \cdots, N$), where θ_k are random rotation angles drawn from some uniform distributions [84]. We mention that in a real experiment, the measurement angles will never be exact due to various control imperfections or system noises. For instance, in quantum dot spin-qubit experiments, the precision of single qubit rotations is typically limited due to charge fluctuations [87] and Overhauser noise [88, 89]. Thus our consideration of random measurement settings is of both theoretical and experimental relevance. In Fig. 3(a), we show the quantum expectation value corresponding to $\mathcal{I}_2$ (denoted as $\mathcal{Q}_v^{(2)}$) as a function of the iteration number of the learning process for a typical random sample of θ_ks. It is clear that $\mathcal{Q}_v^{(2)}$ decreases as the learning process continues and becomes smaller than the classical bound at a critical iteration number. It converges to the exact minimal value as we increase the iteration number further.

Bell inequalities with multipartite correlators.—To show that RBM is also capable of dealing with Bell inequalities with multipartite correlators, we consider the following Bell inequality introduced in Ref. [90]:

$$\mathcal{I}_3 = -\langle\mathcal{M}_0^{(1)}\mathcal{M}_0^{(2)}\cdots\mathcal{M}_0^{(N)}\rangle - \langle\mathcal{M}_1^{(1)}\mathcal{M}_0^{(2)}\cdots\mathcal{M}_0^{(N)}\rangle$$
$$+ \frac{1}{N-1}\sum_{k=2}^{N}[\langle\mathcal{M}_0^{(1)}\mathcal{M}_1^{(k)}\rangle - \langle\mathcal{M}_1^{(1)}\mathcal{M}_1^{(k)}\rangle] \geq -2. \quad (4)$$

This inequality contains only two dichotomic measurements per party, hence in order to find its maximal quantum violation it is sufficient to consider only traceless real observables [91, 92]. As a result, we consider the following choice of

measurements: $\mathcal{M}_0^{(1)} = \hat{\sigma}^x$, $\mathcal{M}_1^{(1)} = \cos\theta\hat{\sigma}^x + \sin\theta\hat{\sigma}^z$, $\mathcal{M}_0^{(k)} = \hat{\sigma}^z$ and $\mathcal{M}_1^{(k)} = \hat{\sigma}^x$ for all $k = 2, \cdots N$. By using RBM-based reinforcement learning, we have computed the quantum violations of Ineq. (4) and part of our results are plotted in Fig. 3(b) [84]. From this figure, we find that the Ineq. (4) is always violated when $\theta \neq \pi/2$ and the maximal violation is achieved at $\theta = 0$ or π. When $\theta = \pi/2$, $\mathcal{M}_0^{(1)} = \mathcal{M}_1^{(1)} = \hat{\sigma}^z$ and the first party actually has only one measurement hence no quantum violation can be obtained. In addition, from our numerical results we also find that the maximal violation of Ineq. (4) is always $-2\sqrt{2}$, independent of the system size [84]. This can be understood from the observation that the Ineq. (4) is in fact reminiscent of the Clauser-Horne-Shimony-Holt inequality [93], whose maximal quantum violation has proved to be bounded by $2\sqrt{2}$ [94].

We emphasize that in the last two examples, we did not specify the spatial dimensionality of the systems. Unlike DMRG, our RBM approach works for any dimension. In addition, as shown in Ref. [79], entanglement is not a limiting factor for the efficiency of the neural-network representation of quantum many-body states. Thus, we expect that RBM can be used to detect many-body nonlocality for quantum states with massive (e.g., volume-law) entanglement as well. This implies another unparalleled advantage of the RBM approach, when compared with traditional methods, such as DMRG, PEPS [95] (projected entangled pair states), or MERA [96] (multiscale entanglement renormalization ansatz). We also note that one may use other type of neural networks (e.g., deep Boltzmann machine [65] or feedforward neural networks [97], etc.) with different learning algorithms to detect many-body nonlocality. A complete study on detecting nonlocality with different neural network would not only bring new powerful tools for solving intricate problems in the quantum information area, but also provide helpful insight on understanding the internal data structures of the networks themselves. We leave this interesting and important topic for future investigation.

Discussion and conclusion.—Finding out experimentally-friendly Bell inequalities for a given many-body system is a challenging problem, since in general the complexity of characterizing the set of classical correlations scales exponentially with the system size. In the future, it would also be interesting to study how machine learning can provide valuable ideas in designing optimal Bell inequalities for many-body systems. Particularly, recent experiments in cold atomic [98] and trapped ion [99] systems have realized programmable quantum simulators with more than fifty qubits and observed exotic quantum dynamics and phases transitions. It is highly desirable to find appropriate Bell inequalities that can be used in these experiments to demonstrate many-body nonlocality. We believe that machine learning will provide valuable wisdom in tackling this problem as well.

In summary, we have introduced machine learning to the detection of quantum nonlocality in many-body systems. Our discussion is mainly based on the RBM architecture, but its generalizations to other artificial neural networks are possible

and straightforward. Through three concrete examples, we have demonstrated that RBM-based reinforcement learning shows remarkable power in computing quantum violations of certain multipartite Bell inequalities. Our results open a door for machine learning Bell nonlocality, which would benefit future studies across quantum information, machine learning, and artificial intelligence.

[1] J. S. Bell, "On the einstein podolsky rosen paradox," (Long Island City, N.Y.) **1**, 195 (1964).

[2] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, "Bell nonlocality," Rev. Mod. Phys. **86**, 419 (2014).

[3] S. J. Freedman and J. F. Clauser, "Experimental test of local hidden-variable theories," Phys. Rev. Lett. **28**, 938 (1972).

[4] A. Aspect, J. Dalibard, and G. Roger, "Experimental test of bell's inequalities using time-varying analyzers," Phys. Rev. Lett. **49**, 1804 (1982).

[5] G. Weihs, T. Jennewein, C. Simon, H. Weinfurter, and A. Zeilinger, "Violation of bell's inequality under strict einstein locality conditions," Phys. Rev. Lett. **81**, 5039 (1998).

[6] M. A. Rowe, D. Kielpinski, V. Meyer, C. A. Sackett, *et al.*, "Experimental violation of a bell's inequality with efficient detection," Nature **409**, 791 (2001).

[7] M. Giustina, A. Mech, S. Ramelow, B. Wittmann, J. Kofler, J. Beyer, A. Lita, B. Calkins, T. Gerrits, S. W. Nam, *et al.*, "Bell violation using entangled photons without the fair-sampling assumption," Nature **497**, 227 (2013).

[8] B. G. Christensen, K. T. McCusker, J. B. Altepeter, B. Calkins, T. Gerrits, A. E. Lita, A. Miller, L. K. Shalm, Y. Zhang, S. W. Nam, N. Brunner, C. C. W. Lim, N. Gisin, and P. G. Kwiat, "Detection-loophole-free test of quantum nonlocality, and applications," Phys. Rev. Lett. **111**, 130406 (2013).

[9] M. Eibl, S. Gaertner, M. Bourennane, C. Kurtsiefer, M. Żukowski, and H. Weinfurter, "Experimental observation of four-photon entanglement from parametric down-conversion," Phys. Rev. Lett. **90**, 200403 (2003).

[10] Z. Zhao, T. Yang, Y.-A. Chen, A.-N. Zhang, M. Żukowski, and J.-W. Pan, "Experimental violation of local realism by four-photon greenberger-horne-zeilinger entanglement," Phys. Rev. Lett. **91**, 180401 (2003).

[11] B. P. Lanyon, M. Zwerger, P. Jurcevic, C. Hempel, W. Dür, H. J. Briegel, R. Blatt, and C. F. Roos, "Experimental violation of multipartite bell inequalities with trapped ions," Phys. Rev. Lett. **112**, 100403 (2014).

[12] J. Hofmann, M. Krug, N. Ortegel, L. Gérard, M. Weber, W. Rosenfeld, and H. Weinfurter, "Heralded entanglement between widely separated atoms," Science **337**, 72 (2012).

[13] W. Pfaff, T. H. Taminiau, L. Robledo, H. Bernien, M. Markham, D. J. Twitchen, and R. Hanson, "Demonstration of entanglement-by-measurement of solid-state qubits," Nat. Phys. **9** (2013), 10.1038/nphys2444.

[14] M. Ansmann, H. Wang, R. C. Bialczak, M. Hofheinz, E. Lucero, M. Neeley, A. O'connell, D. Sank, M. Weides, J. Wenner, *et al.*, "Violation of bell's inequality in josephson phase qubits," Nature **461**, 504 (2009).

[15] B. Hensen, H. Bernien, A. E. Dréau, A. Reiserer, N. Kalb, M. S. Blok, J. Ruitenberg, R. F. Vermeulen, R. N. Schouten, C. Abellán, *et al.*, "Loophole-free bell inequality violation using electron spins separated by 1.3 kilometres," Nature **526**, 682 (2015).

[16] M. Giustina, M. A. M. Versteegh, S. Wengerowsky, J. Handsteiner, A. Hochrainer, K. Phelan, F. Steinlechner, J. Kofler, J.-A. Larsson, C. Abellán, W. Amaya, V. Pruneri, M. W. Mitchell, J. Beyer, T. Gerrits, A. E. Lita, L. K. Shalm, S. W. Nam, T. Scheidl, R. Ursin, B. Wittmann, and A. Zeilinger, "Significant-loophole-free test of bell's theorem with entangled photons," Phys. Rev. Lett. **115**, 250401 (2015).

[17] L. K. Shalm, E. Meyer-Scott, B. G. Christensen, P. Bierhorst, M. A. Wayne, M. J. Stevens, T. Gerrits, S. Glancy, D. R. Hamel, M. S. Allman, K. J. Coakley, S. D. Dyer, C. Hodge, A. E. Lita, V. B. Verma, C. Lambrocco, E. Tortorici, A. L. Migdall, Y. Zhang, D. R. Kumor, W. H. Farr, F. Marsili, M. D. Shaw, J. A. Stern, C. Abellán, W. Amaya, V. Pruneri, T. Jennewein, M. W. Mitchell, P. G. Kwiat, J. C. Bienfang, R. P. Mirin, E. Knill, and S. W. Nam, "Strong loophole-free test of local realism," Phys. Rev. Lett. **115**, 250402 (2015).

[18] C. Zu, D.-L. Deng, P.-Y. Hou, X.-Y. Chang, F. Wang, and L.-M. Duan, "Experimental distillation of quantum nonlocality," Phys. Rev. Lett. **111**, 050405 (2013).

[19] A. Acín, N. Brunner, N. Gisin, S. Massar, S. Pironio, and V. Scarani, "Device-independent security of quantum cryptography against collective attacks," Phys. Rev. Lett. **98**, 230501 (2007).

[20] S. Pironio and S. Massar, "Security of practical private randomness generation," Phys. Rev. A **87**, 012336 (2013).

[21] U. Vazirani and T. Vidick, "Fully device-independent quantum key distribution," Phys. Rev. Lett. **113**, 140501 (2014).

[22] R. Colbeck, *Quantum And Relativistic Protocols For Secure Multi-Party Computation* (PhD thesis, University of Cambridge, 2007).

[23] S. Pironio, A. Acín, S. Massar, A. B. de La Giroday, D. N. Matsukevich, P. Maunz, S. Olmschenk, D. Hayes, L. Luo, T. A. Manning, *et al.*, "Random numbers certified by bell's theorem," Nature **464**, 1021 (2010).

[24] D.-L. Deng and L.-M. Duan, "Fault-tolerant quantum random-number generator certified by majorana fermions," Phys. Rev. A **88**, 012323 (2013).

[25] M. Herrero-Collantes and J. C. Garcia-Escartin, "Quantum random number generators," Rev. Mod. Phys. **89**, 015004 (2017).

[26] C. A. Miller and Y. Shi, "Robust protocols for securely expanding randomness and distributing keys using untrusted quantum devices," in *Proceedings of the 46th Annual ACM Symposium on Theory of Computing* (ACM, 2014) pp. 417–426.

[27] R. S. Michalski, J. G. Carbonell, and T. M. Mitchell, *Machine learning: An artificial intelligence approach* (Springer Science & Business Media, 2013).

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

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

[30] L. Amico, R. Fazio, A. Osterloh, and V. Vedral, "Entanglement in many-body systems," Rev. Mod. Phys. **80**, 517 (2008).

[31] L. Babai, L. Fortnow, and C. Lund, "Non-deterministic exponential time has two-prover interactive protocols," Comput.

Complex. **1**, 3 (1991).

[32] J. Tura, R. Augusiak, A. B. Sainz, T. Vértesi, M. Lewenstein, and A. Acín, "Detecting nonlocality in many-body quantum states," Science **344**, 1256 (2014).

[33] J. Tura, R. Augusiak, A. B. Sainz, B. Lücke, C. Klempt, M. Lewenstein, and A. Acín, "Nonlocality in many-body quantum systems detected with two-body correlators," Annals of Physics **362**, 370 (2015).

[34] J. Tura, G. De las Cuevas, R. Augusiak, M. Lewenstein, A. Acín, and J. I. Cirac, "Energy as a detector of nonlocality of many-body spin systems," Phys. Rev. X **7**, 021005 (2017).

[35] Z. Wang, S. Singh, and M. Navascués, "Entanglement and nonlocality in infinite 1d systems," Phys. Rev. Lett. **118**, 230401 (2017).

[36] S. Wagner, R. Schmied, M. Fadel, P. Treutlein, N. Sangouard, and J.-D. Bancal, "Bell correlations in a many-body system with finite statistics," arXiv:1702.03088 (2017).

[37] R. Schmied, J.-D. Bancal, B. Allard, M. Fadel, V. Scarani, P. Treutlein, and N. Sangouard, "Bell correlations in a bose-einstein condensate," Science **352**, 441 (2016).

[38] J. Batle, C. R. Ooi, S. Abdalla, and A. Bagdasaryan, "Computing the maximum violation of a bell inequality is an np-problem," Quantum Information Processing **15**, 2649 (2016).

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

[40] L.-F. Arsenault, O. A. von Lilienfeld, and A. J. Millis, "Machine learning for many-body physics: efficient solution of dynamical mean-field theory," arXiv:1506.08858 (2015).

[41] Y. Zhang and E.-A. Kim, "Quantum loop topography for machine learning," Phys. Rev. Lett. **118**, 216401 (2017).

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

[43] E. P. van Nieuwenburg, Y.-H. Liu, and S. D. Huber, "Learning phase transitions by confusion," Nature Physics **13**, 435 (2017).

[44] D.-L. Deng, X. Li, and S. D. Sarma, "Exact machine learning topological states," arXiv:1609.09060 (2016).

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

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

[47] K. Ch'ng, J. Carrasquilla, R. G. Melko, and E. Khatami, "Machine learning phases of strongly correlated fermions," Phys. Rev. X **7**, 031038 (2017).

[48] Y. Zhang, R. G. Melko, and E.-A. Kim, "Machine learning $\mathbb{Z}_2$ quantum spin liquids with quasi-particle statistics," arXiv:1705.01947 .

[49] S. J. Wetzel, "Unsupervised learning of phase transitions: From principal component analysis to variational autoencoders," Phys. Rev. E **96**, 022140 (2017).

[50] W. Hu, R. R. P. Singh, and R. T. Scalettar, "Discovering phases, phase transitions, and crossovers through unsupervised machine learning: A critical examination," Phys. Rev. E **95**, 062122 (2017).

[51] N. Yoshioka, Y. Akagi, and H. Katsura, "Learning disordered topological phases by statistical recovery of symmetry," arXiv:1709.05790 (2017).

[52] G. Torlai and R. G. Melko, "Learning thermodynamics with boltzmann machines," Phys. Rev. B **94**, 165134 (2016).

[53] K.-I. Aoki and T. Kobayashi, "Restricted boltzmann machines for the long range ising models," Mod. Phys. Lett. B , 1650401 (2016).

[54] Y.-Z. You, Z. Yang, and X.-L. Qi, "Machine learning spatial geometry from entanglement features," arXiv:1709.01223 (2017).

[55] G. Torlai, G. Mazzola, J. Carrasquilla, M. Troyer, R. Melko, and G. Carleo, "Many-body quantum state tomography with neural networks," arXiv:1703.05334 (2017).

[56] M. Pasquato, "Detecting intermediate mass black holes in globular clusters with machine learning," arXiv:1606.08548 (2016).

[57] Y. D. Hezaveh, L. Perreault Levasseur, and P. J. Marshall, "Fast automated analysis of strong gravitational lenses with convolutional neural networks," Nature **548**, 555 (2017).

[58] R. Biswas, L. Blackburn, J. Cao, R. Essick, K. A. Hodge, E. Katsavounidis, K. Kim, Y.-M. Kim, E.-O. Le Bigot, C.-H. Lee, J. J. Oh, S. H. Oh, E. J. Son, Y. Tao, R. Vaulin, and X. Wang, "Application of machine learning algorithms to the study of noise artifacts in gravitational-wave data," Phys. Rev. D **88**, 062003 (2013).

[59] B. P. Abbott *et al.* (LIGO Scientific Collaboration and Virgo Collaboration), "Observation of gravitational waves from a binary black hole merger," Phys. Rev. Lett. **116**, 061102 (2016).

[60] S. V. Kalinin, B. G. Sumpter, and R. K. Archibald, "Big-deep-smart data in imaging for guiding materials design," Nat. Mater. **14**, 973 (2015).

[61] S. S. Schoenholz, E. D. Cubuk, D. M. Sussman, E. Kaxiras, and A. J. Liu, "A structural approach to relaxation in glassy liquids," Nat. Phys. **12**, 469 (2016).

[62] J. Liu, Y. Qi, Z. Y. Meng, and L. Fu, "Self-learning monte carlo method," Phys. Rev. B **95**, 041101 (2017).

[63] L. Huang and L. Wang, "Accelerated monte carlo simulations with restricted boltzmann machines," Phys. Rev. B **95**, 035105 (2017).

[64] G. Torlai and R. G. Melko, "Neural decoder for topological codes," Phys. Rev. Lett. **119**, 030501 (2017).

[65] X. Gao and L.-M. Duan, "Efficient representation of quantum many-body states with deep neural networks," Nat. Commu. , 662 (2017).

[66] J. Chen, S. Cheng, H. Xie, L. Wang, and T. Xiang, "On the equivalence of restricted boltzmann machines and tensor network states," arXiv: 1701.04831 (2017).

[67] Y. Huang and J. E. Moore, "Neural network representation of tensor network and chiral states," arXiv:1701.06246 (2017).

[68] F. Schindler, N. Regnault, and T. Neupert, "Probing many-body localization with neural networks," Phys. Rev. B **95**, 245134 (2017).

[69] Z. Cai, "Approximating quantum many-body wave-functions using artificial neural networks," arXiv:1704.05148 (2017).

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

[71] Y. Nomura, A. Darmawan, Y. Yamaji, and M. Imada, "Restricted-boltzmann-machine learning for solving strongly correlated quantum systems," arXiv:1709.06475 (2017).

[72] J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, and S. Lloyd, "Quantum machine learning," Nature , 195 (2017).

[73] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," Science **313**, 504 (2006).

[74] R. Salakhutdinov, A. Mnih, and G. Hinton, "Restricted boltzmann machines for collaborative filtering," in *Proceedings of the 24th international conference on Machine learning* (ACM, 2007) pp. 791–798.

[75] H. Larochelle and Y. Bengio, "Classification using discriminative restricted boltzmann machines," in *Proceedings of the 25th*

international conference on Machine learning (ACM, 2008) pp. 536–543.

[76] A. N. Kolmogorov, "On the representation of continuous functions of many variables by superposition of continuous functions of one variable and addition," Amer. Math. Soc. Transl **28**, 55 (1963).

[77] N. Le Roux and Y. Bengio, "Representational power of restricted boltzmann machines and deep belief networks," Neural Comput. **20**, 1631 (2008).

[78] K. Hornik, "Approximation capabilities of multilayer feedforward networks," Neural networks **4**, 251 (1991).

[79] D.-L. Deng, X. Li, and S. Das Sarma, "Quantum entanglement in neural network states," Phys. Rev. X **7**, 021021 (2017).

[80] A. Einstein, B. Podolsky, and N. Rosen, "Can quantum-mechanical description of physical reality be considered complete?" Phys. Rev. **47**, 777 (1935).

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

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

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

[84] See Supplemental Material at [URL will be inserted by publisher] for details on the structure of the neural networks and the DMRG calculations, and for more numerical data.

[85] The matrix elements of Λ are specified as follows: $\Lambda_{0,0} = \Lambda_{0,1} = \Lambda_{1,0} = \Lambda_{2,1} = -\Lambda_{1,1} = -\Lambda_{2,0} = -\Lambda_{3,0} = -\Lambda_{3,1} = 1$ and $\Lambda_{0,2} = -\Lambda_{1,2} = -\Lambda_{2,2} = \Lambda_{3,2} = \Delta$.

[86] On the even sites, we choose the measurements as $\mathcal{M}_0^{(2k)} = (\hat{\sigma}^x + \hat{\sigma}^y + \hat{\sigma}^z)/3$, $\mathcal{M}_0^{(2k)} = (\hat{\sigma}^x - \hat{\sigma}^y - \hat{\sigma}^z)/3$, $\mathcal{M}_0^{(2k)} = (-\hat{\sigma}^x + \hat{\sigma}^y - \hat{\sigma}^z)/3$, and $\mathcal{M}_0^{(2k)} = (-\hat{\sigma}^x - \hat{\sigma}^y + \hat{\sigma}^z)/3$; whereas on the odd sites, the measurement are $\mathcal{M}_0^{(2k+1)} = \hat{\sigma}^x$, $\mathcal{M}_1^{(2k+1)} = \hat{\sigma}^y$, $\mathcal{M}_2^{(2k+1)} = \hat{\sigma}^z$. Here, $\hat{\sigma}^x$, $\hat{\sigma}^y$, and $\hat{\sigma}^z$ are the usual Pauli matrices.

[87] O. E. Dial, M. D. Shulman, S. P. Harvey, H. Bluhm, V. Umansky, and A. Yacoby, "Charge noise spectroscopy using coherent exchange oscillations in a singlet-triplet qubit," Phys. Rev. Lett. **110**, 146804 (2013).

[88] J. Medford, L. Cywiński, C. Barthel, C. M. Marcus, M. P. Hanson, and A. C. Gossard, "Scaling of dynamical decoupling for spin qubits," Phys. Rev. Lett. **108**, 086802 (2012).

[89] T. Fink and H. Bluhm, "Noise spectroscopy using correlations of single-shot qubit readout," Phys. Rev. Lett. **110**, 010403 (2013).

[90] F. Baccari, D. Cavalcanti, P. Wittek, and A. Acín, "Efficient device-independent entanglement detection for multipartite systems," Phys. Rev. X **7**, 021042 (2017).

[91] L. Masanes, "Extremal quantum correlations for n parties with two dichotomic observables per site," arXiv: 0512100 (2005).

[92] B. Toner and F. Verstraete, "Monogamy of bell correlations and tsirelson's bound," arXiv: 0611001 (2006).

[93] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, "Proposed experiment to test local hidden-variable theories," Phys. Rev. Lett. **23**, 880 (1969).

[94] B. S. Cirel'son, "Quantum generalizations of bell's inequality," Letters in Mathematical Physics **4**, 93 (1980).

[95] F. Verstraete, V. Murg, and J. I. Cirac, "Matrix product states, projected entangled pair states, and variational renormalization group methods for quantum spin systems," Advances in Physics **57**, 143 (2008).

[96] G. Vidal, "Class of quantum many-body states that can be efficiently simulated," Phys. Rev. Lett. **101**, 110501 (2008).

[97] J. Schmidhuber, "Deep learning in neural networks: An overview," Neural Networks **61**, 85 (2015).

[98] H. Bernien, S. Schwartz, A. Keesling, H. Levine, A. Omran, H. Pichler, S. Choi, A. S. Zibrov, M. Endres, M. Greiner, *et al.*, "Probing many-body dynamics on a 51-atom quantum simulator," arXiv:1707.04344, accepted for publication in Nature (2017).

[99] J. Zhang, G. Pagano, P. Hess, A. Kyprianidis, P. Becker, H. Kaplan, A. Gorshkov, Z.-X. Gong, and C. Monroe, "Observation of a many-body dynamical phase transition with a 53-qubit quantum simulator," arXiv:1708.01044 (2017).

Supplementary Material for: Machine Learning Bell Nonlocality in Quantum Many-body Systems

I. BELL INEQUALITIES WITH SHORT-RANGE TWO-BODY CORRELATORS

In the main text, we have shown that one can use the RBM-based reinforcement learning to obtain the quantum violations of Ineq. (2). Here, we give more details on the structure of the neural networks, the DMRG calculations, and the comparison between obtained results from different methods.

As discussed in the main text, with properly chosen measurement settings, the Bell operator corresponding to $\mathcal{I}_1$ reduces to the XXZ-type Hamiltonian H [34]. With the open boundary condition, H does not have other obvious symmetry, except that the total $\Sigma^z = \sum \sigma_k^z$ is conserved. Thus, it is straightforward to choose a RBM without any symmetry. In this case, the number of variational parameters is $N + M + N \times M$, which is large when $N \approx 100$ ($\sim 10^4$). Training this RBM is both time and memory consuming. For this particular example, we find that one can instead use a short-range RBM to reduce the number of parameters, and the accuracy of the final results will not be affected to much. To be more concrete, we consider a RBM with $M = \alpha N$ (α denotes the hidden unit density and we choose it to be an integer number for simplicity) and we rearrange the positions of the hidden neurons, such that at each site there are α hidden neurons coupling only locally (within range R) to the visible neurons. This significantly reduces the number of parameters, from $O(N^2)$ to $O(N)$.

We begin with a random short-range RBM (i.e., all the internal parameters are chosen randomly and independently), which typically does not violate the Ineq. (1) in the main text. We then use a reinforcement learning algorithm introduced in Ref. [39] to optimize the internal parameters of the RBM. The details of this algorithm can be found in the Supplementary Materials of Ref. [39]. In Fig. S1, we plot partial of the weight parameters for the final trained RBM in Fig. 2(b) in the main text. Here, only the wight parameters corresponding to the neurons in the first hidden layer are plotted. The parameters associated to other hidden layers looks similar and thus are omitted for the sake of conciseness.

In order to characterize the accuracy of the trained RBM, one can introduced a quantity called the relative error defined as $\epsilon_{\text{rel}} = |E_0^{(\text{RBM})} - E_0|/E_0$, where E_0 is the true ground-state energy of H and $E_0^{(\text{RBM})}$ denotes the value calculated via the RBM approach. For small system sizes ($N \leq 20$), we find that $\epsilon_{\text{rel}} \sim 10^{-4}$ in our calculations. For larger system sizes, we compare our RBM results with that from DMRG. In our DMRG calculations, we use a MPS representation of the quantum many-body states and variationally optimizes the MPS to minimize the ground state energy (see [82] for details). The maximal bond dimension (where we truncate the MPS) is chosen to be $\chi_{\text{max}} = 100$ and we have checked
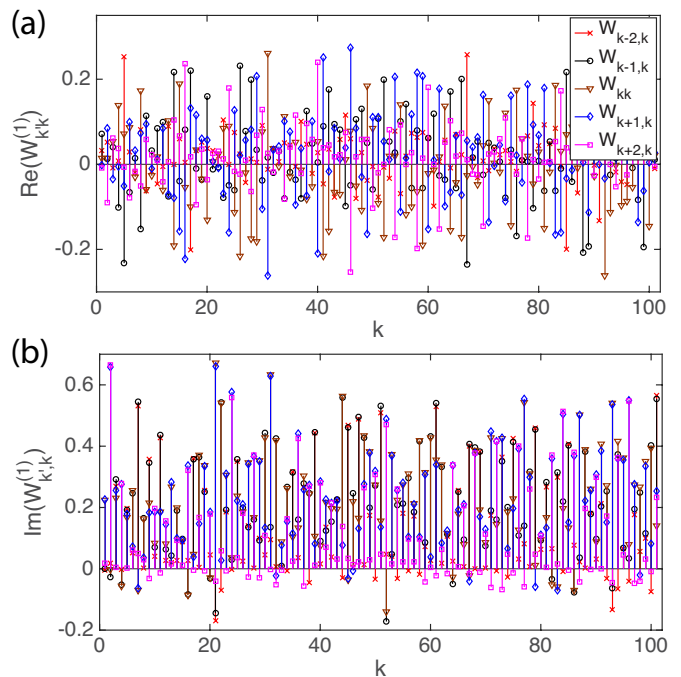

FIG. S1. The learned weight parameters for representing the ground state of H with a short-range restricted Boltzmann machine (here, we have fixed $\alpha = 4$ and $R = 2$). We denote by $W_{k'k}^{(\eta)}$ ($\eta = 1, 2, \cdots, \alpha$, k', $k = 1, 2, \cdots, N$) the coupling strength between the η-th hidden neuron at site k' and the visible neuron at sit k. (a) and (b) show the real and imaginary parts of $W_{k'k}^{(1)}$, respectively. They share the same legend. The parameters specifying H are chosen the same as in Fig. 2(b) in the main text.

that the neglected weight for all the truncations are smaller than 10^{-6}. We have also examined that the typical variances $\sigma^2 = \langle H^2 \rangle - \langle H \rangle^2$ is smaller than 10^{-8}, verifying that the obtained MPS is indeed an eigenstate of H (up to a negligible error rate). In comparing our RBM and DMRG results, we find that the relative error $\epsilon_{\text{rel}} = |E_0^{(\text{RBM})} - E_0^{(\text{DMRG})}|/E_0^{(\text{DMRG})} \sim 10^{-3}$. We mention that the accuracy of the RBM results can be systematically improved by increasing R and α, or the number of iterations in the training process. In this work of detecting many-body nonlocality via RBM, high accuracy is not a major concern, hence $\epsilon \sim 10^{-3}$ is already sufficient for our purpose.

II. BELL INEQUALITIES WITH ALL-TO-ALL TWO-BODY CORRELATORS

The Bell inequality in Ineq. (3) in the main text has a permutation symmetry by construction [32]. Thus, its corresponding Bell operator will also has a permutation symmetry if we choose the same measurement settings for each party. In this case, it is natural to use a permutation-invariant RBM to calculate the quantum violations. This greatly reduces the number of the variational parameters and the same quantum violations as given in Ref. [32] can be readily obtained. What might be more interesting is the case in which
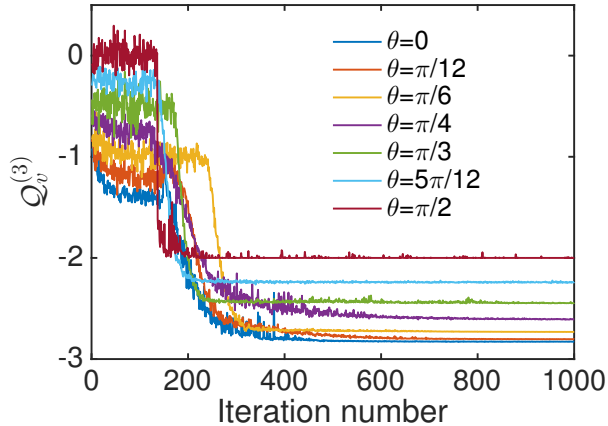
FIG. S2. RBM-learned quantum expectation value $\mathcal{Q}_v^{(3)}$ as a function of the iteration number, for different measurement angle θ.
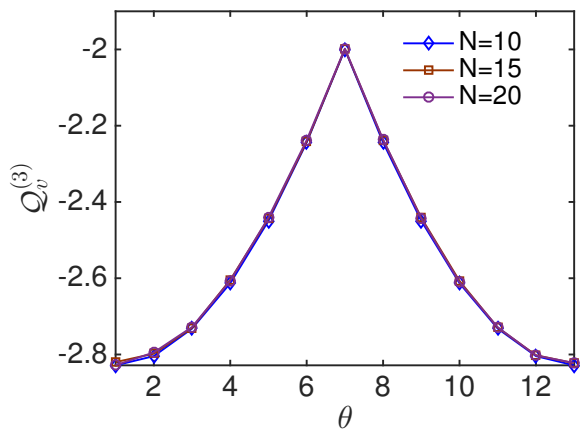

FIG. S3. RBM-learned quantum violations $\mathcal{Q}_v^{(3)}$ as a function of θ, for different system sizes.

different party choose different measurement settings. In this case, the permutation symmetry is violated and it is very challenging to compute the quantum violations of Ineq. (3). In the main text, we have considered a scenario where the mea-

surement setting for each party is random. More specifically, we have chosen the measurement settings to be: $\mathcal{M}_0^{(k)} = \sigma^z$ and $\mathcal{M}_1^{(k)} = \cos\theta_k\sigma^z + \sin\theta_k\sigma^x$ with θ_k being random rotation angles drawn independently from a uniform distribution $[\theta - \varepsilon, \theta + \varepsilon]$. Since there is no obvious symmetry for the corresponding Bell operator and the correlators are all-to-all, we choose the most general RBM with each hidden neuron connected to all the visible ones. In plotting Fig. 3(a) in the main text, we have fixed $\theta = 2\pi/3$ and $\varepsilon = 0.1$.

III. BELL INEQUALITIES WITH MULTIPARTITE CORRELATORS

For the Bell inequality (4) in the main text, it is easy to observe that there is a permutation symmetry between parties indexed from 2 to N. In addition, the considered measurement settings also have this symmetry, and thus so does the corresponding Bell operator. Taking this into consideration, we choose a RBM with the same symmetry: the RBM contains M hidden neurons and each of them connects to all visible neurons, but the baise and weight parameters satisfy $a_2 = a_3 = \cdots = a_N$ and $W_{k',1} = W_{k',2} = \cdots = W_{k',N}$ (for $k' = 2, 3, \cdots, M$), respectively. This reduces the number of parameters from $O(MN)$ to $O(M)$ and significantly simplified the calculations.

In Fig. S2, we show the RBM-learned $\mathcal{Q}_v^{(3)}$ as a function of the iteration number for different measurement angle θ. It is clear from this figure that $\mathcal{Q}_v^{(3)}$ converges rapidly to the corresponding exact values for different θ. We note that in our calculations we have chosen the learning rate to be an exponential decaying function of the iteration number, following Ref. [39]. Thus, at the beginning of the learning process, the learning rate is large. This explains the large fluctuations at the beginning of the learning process. As the iteration number increases, the learning rate becomes small and the curves become smooth.

In Fig. S3, we plot the RBM-learned quantum violations $\mathcal{Q}_v^{(3)}$ as a function of the measurement angle θ for different system sizes. We find that the quantum violations are independent of N (up to negligible numerical errors).