nature physics

PERSPECTIVE

https://doi.org/10.1038/s41567-019-0545-1

Restricted Boltzmann machines in quantum physics

Roger G. Melko[1,2*], Giuseppe Carleo[3], Juan Carrasquilla[4] and J. Ignacio Cirac[5]

A type of stochastic neural network called a restricted Boltzmann machine has been widely used in artificial intelligence applications for decades. They are now finding new life in the simulation of complex wavefunctions in quantum many-body physics.

The study of strongly correlated quantum many-body physics is the last frontier still untamed by the sophisticated, high-performance computational frameworks pervasive in other areas of the sciences. The challenge is fundamental; how does one use classical computers to study the many-body quantum wavefunction, an object inherently fraught with exponentially scaling complexity? Researchers have developed a variety of tools over the years to attempt to tame this complexity, such as the quantum Monte Carlo and density matrix renormalization group (DMRG) methods[1–3]. Despite this, key problems remain unsolved in condensed-matter physics, ranging from a numerical solution to the Hubbard model to the origins of thermalization from quantum dynamics, and more. These problems share deep connections with challenges in the quantum information sciences, where experimentalists are now obtaining sufficient control of individual qubits and their interactions to enable the engineering of coherent many-body devices[4,5]. There, computational tools will be key in the reconstruction and characterization of wavefunctions from experimental measurements[6], a necessary step in the design, construction and validation of quantum computers.

Enter machine learning with its recent high-profile success in a range of real-world problems, such as computer vision and natural language processing, each exhibiting a complexity that scales surprisingly similar to the quantum many-body problem[7,8]. The quantum physics community has taken notice, throwing themselves headfirst into an exploration of the algorithms underlying modern machine learning, with an eye on making progress in quantum physics. Among the various approaches involving supervised, unsupervised and reinforcement learning, restricted Boltzmann machines (RBMs) stand out as objects superbly interpretable under the edifice of well-established statistical physics and quantum information theories. We begin this Perspective with a review of the representational power of RBMs, and their relationship to that most elusive object: the quantum wavefunction. Guided by the extensive frameworks already in existence for the simulation of quantum systems on classical computers, we discuss two different paradigms for the 'learning' of the RBM parameters, variational minimization of the energy[7] and reconstruction from experimental data[9,10]. We close with a perspective on the future of this young field of quantum-state reconstruction with machine learning and its role in the coming age of 'quantum supremacy'[11].

RBMs and efficient representation of quantum states

Boltzmann machines were created as a way of graphically representing a probability distribution $p(\mathbf{v})$, by defining a number of stochastic variables (nodes) $\mathbf{v}$ and an 'energy' E by which the nodes interact (Box 1). Hopfield networks, familiar to physicists as fully connected Ising models, were the first to employ this paradigm to model associative memory[12]. Hopfield's work established a connection between the concept of learning and the optimization of neural-network energy parameters (weights and biases). This in turn led to fruitful cross-pollination between machine learning and core concepts in statistical physics, such as glassiness and inverse Ising problems[13].

Whereas Hopfield networks have a number of nodes equal to the size of the physical system under consideration, Boltzmann machines extend this space to include unphysical nodes and interactions[14]. This hidden or latent space $\mathbf{h}$ can capture higher-order correlations between visible nodes, significantly increasing the expressiveness of the network representing the joint distribution $p(\mathbf{v},\mathbf{h})$. To facilitate a practical algorithm for learning the network parameters from data, Boltzmann machines became restricted; that is, couplings were allowed only between visible and hidden units, but not intralayer (visible–visible or hidden–hidden)[15]. Thus coined, RBMs played a historically important role in generative pre-training, dimensionality reduction, collaborative filtering, topic modelling and more[16–18]. Even today, RBMs offer a gold standard in explicit density models, where a well-understood Monte Carlo strategy is used to estimate the cost function. These features are particularly appealing to physicists, and constitute one of the reasons for the present rise of popularity of RBMs in the field of wavefunction representation.

An RBM is known to be a universal approximator[19]. That is, given a sufficient number of hidden units, the network is capable of approximately representing any desired probability distribution to arbitrary accuracy. This capability immediately raises the prospect of using RBMs as approximate representations of many-body wavefunctions, dubbed neural-network quantum states[7]. However to do so, RBMs must first be modified to accommodate complex phases. This may be achieved in a number of ways, for example, by allowing the weights to be complex numbers[7] or by indirectly parameterizing a phase with additional hidden units[9]. This conceptual leap, from probability distributions to complex-valued quantum amplitudes, sets a distinct shift from classical applications to quantum ones, and often demands the development of novel machine learning tools specific for quantum applications.

Given that sufficiently large RBMs are universal approximators, it is natural to ask a theoretical question: precisely which types of quantum wavefunctions can they represent efficiently, that is, using

[1]Department of Physics and Astronomy, University of Waterloo, Waterloo, Ontario, Canada. [2]Perimeter Institute for Theoretical Physics, Waterloo, Ontario, Canada. [3]Center for Computational Quantum Physics, Flatiron Institute, New York, NY, USA. [4]Vector Institute for Artificial Intelligence, MaRS Centre, Toronto, Ontario, Canada. [5]Max-Planck-Institut für Quantenoptik, Garching, Germany. *e-mail: rgmelko@uwaterloo.ca

Box 1 | A probabilistic graphical model

Classical probability distributions can be parametrized by an RBM. This is constituted by a layer of visible units v_j and a layer of hidden units h_i, stochastic binary variables that can take values 0 or 1. Machine parameters λ are weight and biases, which define an energy E and parametrize the joint probability distribution $p(\mathbf{v},\mathbf{h})$ as a graphical model (see figure). An RBM can be interpreted as a wavefunction $p(\mathbf{v}) = |\langle \mathbf{v}|\Psi\rangle|^2$ if it is generalized to include a complex phase, where $p(\mathbf{v})$ is the marginal distribution over visible units $p(\mathbf{v}) = \sum_{\mathbf{h}} p(\mathbf{v},\mathbf{h})$. A glossary with relevant definition follows.

Biases. For a visible unit v_j and a hidden unit h_i, the respective biases in the RBM energy are b_j and c_i. They act like a local magnetic field in the energy of an Ising model (see equation (1)).

Energy. In analogy to statistical physics the energy of an RBM is defined given the joint configuration $(\mathbf{v},\mathbf{h})$ of visible and hidden units:

$$E_\lambda(\mathbf{v},\mathbf{h}) = -\sum_{i,j} W_{ij} h_i v_j - \sum_{j=1}^{V} b_j v_j - \sum_{i=1}^{H} c_i h_i \quad (1)$$

Generative model. Framed as an unsupervised learning problem, a generative model is a statistical model of a probability distribution that underlies some set of data.

Hidden units. There are H units in the second layer of the RBM, denoted by the vector $\mathbf{h} = (h_1, ..., h_H)$, representing latent variables and referred to as 'hidden'. The number of hidden units H can be adjusted to tune the representational capacity of the RBM.

Joint distribution. The RBM assigns a probability to each joint configuration $(\mathbf{v},\mathbf{h})$ according to the Boltzmann distribution of the energy:

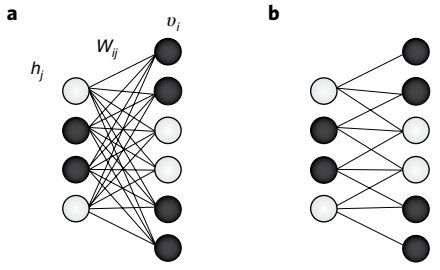
$$p_\lambda(\mathbf{v},\mathbf{h}) = \frac{1}{Z_\lambda} e^{-E_\lambda} \quad (2)$$

where Z_λ is the normalizing constant, or partition function.

Parameters. An RBM's energy is defined via a set of neural-network parameters $\lambda = \{\mathbf{W}, \mathbf{b}, \mathbf{c}\}$, consisting of weights and biases.

Visible units. There are V units in the first layer of the RBM, denoted by the vector $\mathbf{v} = (v_1, ..., v_V)$. The number of visible units V is fixed to the number of physical lattice sites or qubits.

Weights. W_{ij} is the symmetric connection or interaction between the visible unit v_j and the hidden unit h_i. Weights act like an interaction term between Ising variables (see equation (1)).



Restricted Boltzmann machine. a, A standard, fully connected RBM where all visible units are connected to each hidden unit. This RBM is capable of encoding a wavefunction with volume-law entanglement. **b**, A sparse RBM with connections between neighbours. Such a machine reintroduces the idea of spatial locality and is capable of representing quantum states with an area-law entanglement structure.

realistic computational resources? This task is made simpler by theoretical frameworks already in place in physics, such as the field of tensor networks (TNs)[20]. There, entanglement has been identified as an indicator of efficiency for many-body wavefunction representation[21]. Essentially, states fulfilling the entanglement area law are well described by TNs, meaning that the number of parameters required for their accurate representation scales polynomially with respect to the number of qubits. The area law is associated to low-energy states of Hamiltonians with local interactions, and indicates that the amount of entanglement between one subregion and its complement is bounded by a term proportional to surface area of the boundary between them[22]. This explains the success of the DMRG algorithm, which is based on matrix product states (MPS)[23], the simplest TNs that adapt to a one-dimensional geometry. In higher dimensions, while efficient representations exist (such as the analogous projected entangled pair states (PEPS))[24], efficient optimization of parameters is currently the central challenge.

A distinct feature of conventional RBMs used in machine learning is that they can immediately be interpreted under the lens of well-established TN theory[25,26]. In particular, the entanglement that can be captured by an RBM provides a quantification of its expressive power, based on its connectivity. Most intuitively, it has been shown that the full connectivity (between visible and hidden layers) of a conventional RBM bounds the entanglement entropy to scale with a subregion's volume, in contrast to its area[27]. It follows that a conventional RBM may be capable of efficiently encoding wavefunctions outside of the subset defined by MPS or PEPS. In some cases, it is capable of doing so with very few parameters (for example, independent of the system size or scaling polynomially with the number of visible units). This may open up the possibility of addressing quantum many-body problems that are out of reach of TNs; for example, finding the ground states of systems with long-range interactions (for example, Coulomb or dipole–dipole interactions)[28] or excited states of local Hamiltonians. In both cases, the area law is violated and representations amenable to volume-law entanglement are required. In other cases that may fulfill the area law, including the problem of finding the ground states of local Hamiltonians in three dimensions, there is no known way to contract the tensors of a TN representation efficiently. However, here RBMs could operate just as well as they do in lower dimensions[29]. Another set of problems where RBMs appear to be more powerful than TN methods are those related to chiral topological order, as one can encode prototypical states very efficiently with RBMs, including chiral states such as the Laughlin state[26].

Of course, like other ansatze, RBMs will have different advantages and disadvantages over TNs depending on the structure of the wavefunction one is targeting. Indeed, there are clear connections between RBMs and TNs that help make this relationship clear, and which can be exploited to improve the performance of each. For instance, certain classes of PEPS can be written exactly as RBMs, and thus for those states one can take advantage of learning algorithms (see the next section) that obtain an efficient optimization of the parameters, even in three dimensions. Furthermore, one can consider a larger family of the so-called string-bond states (which include MPS), that extend RBMs in a natural way and that are also amenable to fast learning algorithms[25,26]. Alternatively, one could increase the efficiency of Boltzmann machines by further restricting the interlayer connections to be short-ranged or sparse, thus recovering the area law suitable for the study of ground-state wavefunctions of local Hamiltonians.

Indeed, the theoretical connections between TNs and RBMs promise very fruitful applications in solving and understanding quantum many-body systems[27,30]. Finally, a natural strategy to go beyond the RBM paradigm is to consider their deep generalizations[31]. Theoretically, these are known to efficiently represent any quantum state that results from physical evolution[32,33]. However,

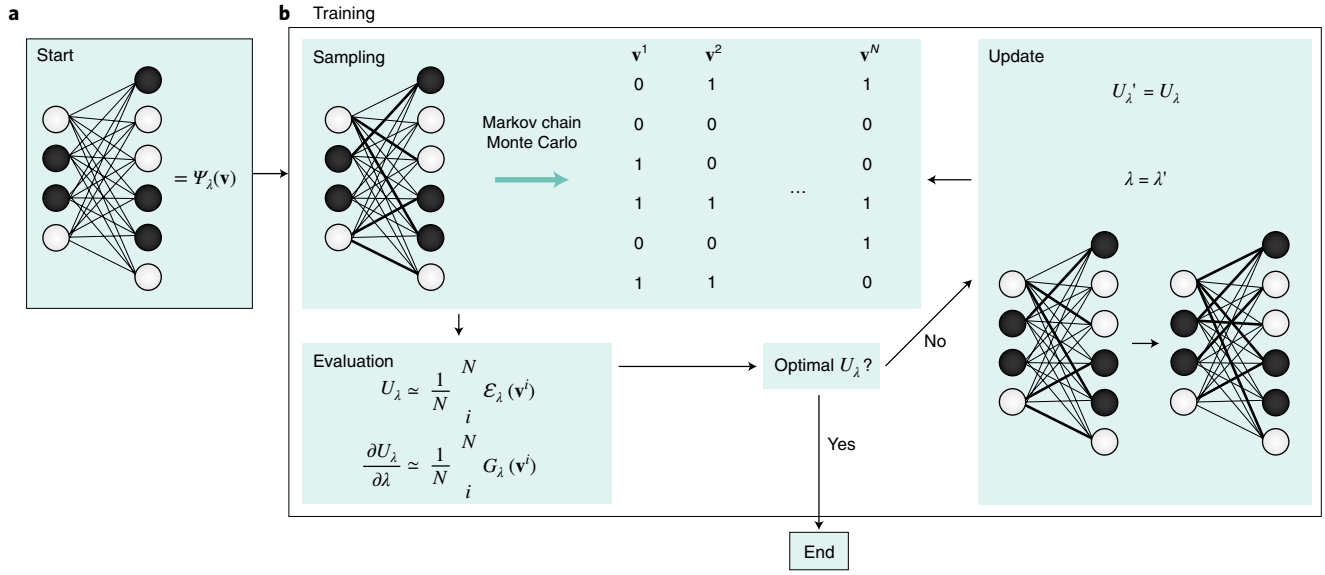


Fig. 1 | Using variational methods to learn the parameters of an RBM. a, An RBM with complex-valued weights is taken to represent the target many-body state in a given computational basis $\langle \mathbf{v}|\Psi\rangle$. The goal is to find RBM parameters such that the energy functional $U_\lambda = \frac{\langle \Psi_\lambda | H | \Psi_\lambda \rangle}{\langle \Psi_\lambda | \Psi_\lambda \rangle}$ attains a minimum. **b**, At each learning step, a large number N of samples $\mathbf{v}^1 \ldots \mathbf{v}^N$ are generated according to the probability distribution defined by $|\Psi_\lambda(\mathbf{v})|^2$, using Markov chain Monte Carlo. These samples are used to estimate the expectation value of the energy functional and its gradient. Their statistical estimators $\varepsilon_\lambda(\mathbf{v}) = \langle \mathbf{v}| H |\Psi_\lambda\rangle / \langle \mathbf{v}|\Psi_\lambda\rangle$ and $G_\lambda(\mathbf{v})$, respectively, are explained in ref. [3]. Many different choices for $G_\lambda(\mathbf{v})$ are possible. One effective choice can simply be traditional stochastic gradient descent, but other improved estimators can be used, for example, stochastic reconfiguration as described in ref. [3]. Gradient information is used to update the network parameters in a direction that decreases the energy functional. The new parameters are used again, and until the optimal energy is found and the training ends.

these models come with a significantly higher computational cost with respect to RBM quantum states, thus limiting their practical use to date.

Learning the parameters of an RBM given a Hamiltonian

Having established the use of RBMs and their natural extensions as approximate representations of quantum states, the question becomes: how does one find the optimal parameters (weights and biases) so that the stochastic neural network best represents the target wavefunction? First, we consider the definition of the target wavefunction as the ground-state eigenvector of some Hamiltonian of interest. This is the usual task in theoretical condensed-matter physics, atomic and molecular optics, and quantum chemistry, and is the purview of conventional numerical methods based on the variational principle, such as variational Monte Carlo[3] and DMRG[1]. Finding the ground state of a given Hamiltonian (H) can be recast as a high-dimensional minimization problem of the energy expectation value $U_\lambda = \langle \Psi_\lambda | H | \Psi_\lambda \rangle / \langle \Psi_\lambda | \Psi_\lambda \rangle$, over the network parameters λ encoding the wavefunction Ψ (ref. [7]). Analogous to machine learning applications, U is a highly nonlinear loss function, whose value and gradient are both estimated stochastically. Modern methods, such as the stochastic reconfiguration traditionally used in standard variational Monte Carlo calculations[3,34], have been used to optimize the variational parameters in an RBM wavefunction with great success (Fig. 1).

Since its recent introduction, this approach and subsequent extensions have been used to study a variety of interacting quantum models, including spins[7,26,28,35,36], bosons[37], fermions[36,38,39] and continuous-space Hamiltonians[39–41]. Generalizations of the static ground-state variational principle have also been used to adapt the RBM ansatz for purposes other than solving for the ground-state wavefunction. For example, RBM representations have proven sufficiently flexible to study excited states in given symmetry sectors and determine energy gaps[42]. Beyond the static case, an important generalization of RBM states is obtained when allowing the network parameters to carry a time dependence. In this case, it is possible to use the time-dependent variational principle of Dirac and Frenkel to find the best variational state solving the time-dependent Schrödinger equation[7,43]. A generalization of this approach has been recently introduced for the simulation of arbitrary unitary operators, leading to an algorithm for the classical simulation of quantum circuits with RBM states[44].

The results obtained so far have been encouraging, and have in some cases improved on existing state-of-the-art methods. For example, highly accurate ground-state wavefunctions for one-dimensional models have been obtained. In two dimensions, RBM ground states are competitive with TN-based states, and cases where they can improve on existing variational results have been demonstrated, both for two-dimensional antiferromagnets and traditionally challenging chiral topological phases[26,45]. Success in fermionic matter, infamous for both inducing the prohibitive 'sign problem' in quantum Monte Carlo simulations as well as the anomalously large entanglement scaling arising from the presence of a Fermi surface, would represent the holy grail for numerical simulation techniques.

While only in its early stages, applications to fermionic matter have shown how RBM states can offer systematic improvement over existing ansatz wavefunctions developed in the past decade[36]. In time-dependent applications, RBM states have been used to study out-of-equilibrium dynamics of strongly interacting spin systems, reaching good agreement with DMRG results on benchmark one-dimensional problems[7,46,47] and setting the scene for calculations in higher dimensions, where several interesting physical regimes are currently beyond the reach of existing many-body techniques.

While many-body variational applications based on RBMs inherit all the powerful ideas of artificial neural networks and machine learning techniques, they also inherit their main drawbacks. Most notably, it is an open challenge to identify which network architectures are the most amenable to an efficient

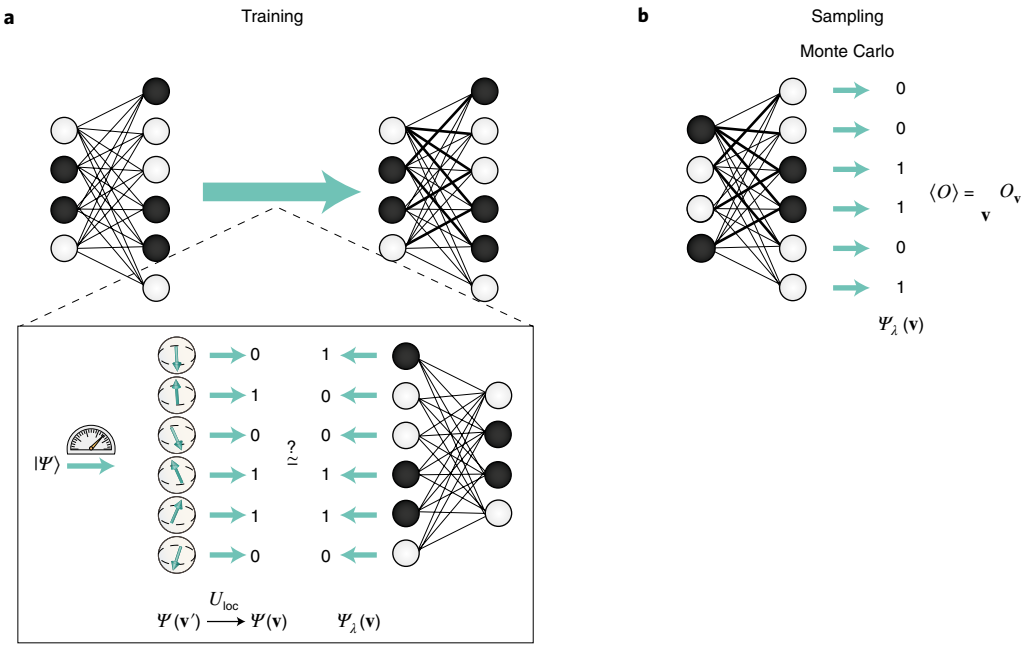


Fig. 2 | Training an RBM with experimental data. a, The objective of the training is to minimize the statistical 'distance' between the experimental data distributions obtained in a number of bases and the probabilities that the RBM state assigns to the data. To do this, experimental data are obtained from projective measurements of a pure quantum state $|\Psi\rangle$ in several bases — depicted as arrows in the Bloch spheres. This measurement set is used to train an RBM that represents the wavefunction in the computational basis. The necessary in-between step is to use a unitary matrix transformation U_{loc} to transform the measurement from the various bases into the computational basis. For a computationally feasible reconstruction, this unitary transformation should be spatially local — namely, it should act only on a small number of neighbouring variables in the system. **b**, After the training, the machine parameters λ (weights and biases) are fixed, as depicted with varying-thickness connections between units. Through a Monte Carlo sampling of the RBM (illustrated as out-coming spin configurations drawn from the RBM), the trained states can then be used to evaluate expectation values $\langle O \rangle$ of static diagonal and off-diagonal operators in the computational basis, as well as quantities such as entanglement entropy. In practice, RBM reconstructions can be used to estimate quantities that are challenging for direct observation in experiments using only simple measurements currently available in cold atoms and quantum annealers[36–38].

optimization of the variational parameters in cases of interest for physics applications. Ultimately, this constitutes a central challenge for future research, where the increased network expressibility and entanglement content will confront the intrinsic hardness of the optimization problem. Research efforts in this direction have only just started. Guidance from general principles developed in both the machine learning and TN communities will be crucial to make further progress[48].

Learning the parameters of an RBM given measurement data

With the growing numbers of highly controlled atoms, molecules and qubits in the era of 'quantum supremacy' comes the necessary breakdown of quantum-state tomography and other conventional tools for state benchmarking and verification[49–52]. Fortunately, the rapid increase in the amount of data being produced by these experiments suggests that RBMs may offer an ideal tool for the practical representation and characterization of their quantum states. In fact, RBMs were originally designed to learn from data in a machine learning context, giving them abilities that nicely complement conventional TN- or Monte Carlo-based techniques. Leveraging a data-driven approach offers a unique opportunity to employ powerful generative modelling techniques — well studied in the machine learning context — to aid in the rapidly accelerating development of quantum experiments and technology.

The foundation of machine learning with RBMs is the ability to learn an unknown probability distribution from a set of data drawn according to this distribution. Consider the Born rule, which dictates that the probability of measuring a quantum state is proportional to the square of its wavefunction. This immediately suggests that, if sufficient experimental measurements are available, standard machine learning techniques may be able to reconstruct this state in an RBM (suitably generalized to represent a complex wavefunction, as discussed above).

The type and amount of experimental data required for this reconstruction depends crucially on the structure of the quantum state. The simplest case occurs when the wavefunction has a positive sign structure in the basis in which measurements are taken. This is the case for many physical systems, such as the transverse field Ising model, effective models of interacting atoms or ions[53], and quantum devices such as the D-Wave computer[54,55]. Then, the RBM has a simple encoding of the wavefunction as $\Psi = \sqrt{p(\mathbf{v})}$ (ref. [9]). Measurements in a single basis (σ^z or orbital occupation, say) are sufficient for the reconstruction using fast learning techniques such as contrastive divergence (an approximate maximum-likelihood learning algorithm[16]). The quality of the reconstruction again depends on the expressiveness of the neural network; however, it also depends on the number of measurements, that is, the size of the training set.

As discussed above, many wavefunctions of interest in the physical sciences are not simply real and positive, and standard machine learning techniques for RBMs have been extended to this case. These have successfully demonstrated reconstruction of physical wavefunctions of a size that is typical for experiments[9]. Of course, to train the machine parameters that encode the phase structure, measurements must be provided in additional bases (Fig. 2) — a significant challenge for experimentalists in many cases.

Once trained, an RBM representation of a wavefunction can take full advantage of generative modelling techniques. This means that a trained RBM wavefunction can produce new measurements in the computational basis, and all corresponding observables, such as diagonal correlation functions. In exactly the same way as wavefunctions optimized by variational techniques, a trained RBM wavefunction can be used to measure off-diagonal observables and quantities that are difficult to measure experimentally, such as entanglement entropies[9,56]. Finally, through the use of a purification, this approach has been extended to mixed states, and the reconstruction of a full density matrix has recently been demonstrated with real experimental data[57].

The future of generative modelling in quantum science

Clearly, machine learning offers exciting prospects for computational research in quantum materials and devices. The stunning advances that we have witnessed in industry applications fuel optimism; however, with it comes the steady reminder that there is much yet to understand about the core advantages that machine learning may hold over more traditional computational approaches. The ability to view RBMs through the lens of established theoretical frameworks such as TNs has been crucial in advancing the field and underlies much of their appeal to physicists.

As the field advances, so does the potential for future research, both in the theoretical underpinnings of machine learning algorithms and their practical implementation for quantum materials and devices in the near term. A key focus will be on the relationship between the expressibility of various architectures, including deep neural networks, and the efficient learning of network parameters. In the case of learning from Hamiltonians, the hope is that machine learning will make new connections between disparate concepts, such as the phase structure of the wavefunction, the scaling of correlations, the quantum Monte Carlo sign problem and glassy optimization landscapes. A central question is: what is the natural measure of complexity for a quantum many-body state described by an artificial neural network? Concepts beyond entanglement entropy increasingly seem necessary, suggesting that a refinement of the core lessons taught by TNs may be in our future.

For data-driven learning, as quantum devices are soon poised to overcome classical computation, it is expected that the exponential scaling problem of quantum-state reconstruction cannot ultimately be overcome in general. However, by exploiting structural information of the states produced by physical devices in combination with RBMs and other neural-network representations of the wavefunction[58], the frontier between states that can be represented classically and those produced by a quantum device can be pushed forward. The ability to reconstruct a quantum state from data, using classical resources, may soon serve as the baseline for defining 'quantum advantage' on near-term devices and noisy intermediate-scale quantum computers. The difference between this and conventional definitions of 'quantum supremacy' is a theoretically interesting question that we will soon be forced to confront.

In artificial intelligence research, RBMs have largely been eclipsed by other generative models[59]. Several other explicit density models are widely used, such as autoencoders that use a variational procedure to approximate the optimization problem[60] and powerful recurrent neural networks that have a tractable density[61]. Explorations into the use of these strategies has already begun in the context of density matrix reconstruction[58,62] and ground-state problems[42,63]. Modern implicit density models, such as generative adversarial neural networks[64] and autoregressive models[61], have achieved remarkable feats in industrial applications. However, the consequences of working with implicit representations of quantum states is still an open research problem.While these and other complementary techniques surely have many advantages and disadvantages over RBMs, the study of generative modelling for wavefunction reconstruction has to date barely scratched the surface. Clearly, the exploration of machine learning in many-body physics has just begun.

Received: 30 September 2018; Accepted: 1 May 2019;
Published online: 24 June 2019

References

1. White, S. R. Density matrix formulation for quantum renormalization groups. *Phys. Rev. Lett.* **69**, 2863–2866 (1992).
2. Sandvik, A. W. Computational studies of quantum spin systems. *AIP Conf. Proc.* **1297**, 135–338 (2010).
3. Becca, F. & Sorella, S. *Quantum Monte Carlo Approaches for Correlated Systems* (Cambridge Univ. Press, 2017).
4. Bernien, H. et al. Probing many-body dynamics on a 51-atom quantum simulator. *Nature* **551**, 579–584 (2017).
5. Zhang, J. et al. Observation of a many-body dynamical phase transition with a 53-qubit quantum simulator. *Nature* **551**, 601–604 (2017).
6. Torlai, G. et al. Integrating neural networks with a quantum simulator for state reconstruction. Preprint at https://arxiv.org/abs/1904.08441 (2019).
7. Carleo, G. & Troyer, M. Solving the quantum many-body problem with artificial neural networks. *Science* **355**, 602–606 (2017).
8. Carrasquilla, J. & Melko, R. G. Machine learning phases of matter. *Nat. Phys.* **13**, 431–434 (2017).
9. Torlai, G. et al. Neural-network quantum state tomography. *Nat. Phys.* **14**, 447–450 (2018).
10. Torlai, G. & Melko, R. G. Learning thermodynamics with Boltzmann machines. *Phys. Rev. B* **94**, 165134 (2016).
11. Terhal, B. M. Quantum supremacy, here we come. *Nat. Phys.* **14**, 530–531 (2018).
12. Hopfield, J. J. Neural networks and physical systems with emergent collective computational abilities. *Proc. Natl Acad. Sci. USA* **79**, 2554–2558 (1982).
13. Nguyen, H. C., Zecchina, R. & Berg, J. Inverse statistical problems: from the inverse ising problem to data science. *Adv. Phys.* **66**, 197–261 (2017).
14. Ackley, D. H., Hinton, G. E. & Sejnowski, T. J. A learning algorithm for Boltzmann machines. *Cogn. Sci.* **9**, 147–169 (1985).
15. Smolensky, P. in *Parallel Distributed Processing: Explorations in the Microstructure of Cognition* Vol. 1. (eds Rumelhart, D. E. & McClelland, J. L.) 194–281 (MIT Press, 1986).
16. Hinton, G. E. Training products of experts by minimizing contrastive divergence. *Neural Comput.* **14**, 1771–1800 (2002).
17. Hinton, G. E. & Salakhutdinov, R. R. Reducing the dimensionality of data with neural networks. *Science* **313**, 504–507 (2006).
18. Salakhutdinov, R., Mnih, A. & Hinton, G. Restricted Boltzmann machines for collaborative filtering. In *Proc. 24th International Conference on Machine Learning* (ed. Ghahramani, Z.) 791–798 (ACM, 2007).
19. Roux, N. L. & Bengio, Y. Representational power of restricted Boltzmann machines and deep belief networks. *Neural Comput.* **20**, 1631–1649 (2008).
20. Orus, R. A practical introduction to tensor networks: matrix product states and projected entangled pair states. *Ann. Phys.* **349**, 117–158 (2014).
21. Vidal, G. Efficient simulation of one-dimensional quantum many-body systems. *Phys. Rev. Lett.* **93**, 040502 (2004).
22. Eisert, J., Cramer, M. & Plenio, M. B. Colloquium: area laws for the entanglement entropy. *Rev. Mod. Phys.* **82**, 277–306 (2010).
23. Östlund, S. & Rommer, S. Thermodynamic limit of density matrix renormalization. *Phys. Rev. Lett.* **75**, 3537–3540 (1995).
24. Verstraete, F. & Cirac, J. I. Renormalization algorithms for quantum-many body systems in two and higher dimensions. Preprint at https://arxiv.org/abs/cond-mat/0407066 (2004).
25. Clark, S. R. Unifying neural-network quantum states and correlator product states via tensor networks. *J. Phys. A* **51**, 135301 (2018).
26. Glasser, I., Pancotti, N., August, M., Rodriguez, I. D. & Cirac, J. I. Neural-network quantum states, string-bond states, and chiral topological states. *Phys. Rev. X* **8**, 011006 (2018).
27. Chen, J., Cheng, S., Xie, H., Wang, L. & Xiang, T. Equivalence of restricted Boltzmann machines and tensor network states. *Phys. Rev. B* **97**, 085104 (2018).
28. Deng, D.-L., Li, X. & Das Sarma, S. Quantum entanglement in neural network states. *Phys. Rev. X* **7**, 021021 (2017).
29. Sfondrini, A., Cerrillo, J., Schuch, N. & Cirac, J. I. Simulating two- and three-dimensional frustrated quantum systems with string-bond states. *Phys. Rev. B* **81**, 214426 (2010).
30. Pastori, L., Kaubruegger, R. & Budich, J. C. Generalized transfer matrix states from artificial neural networks. *Phys. Rev. B* **99**, 165123 (2019).
31. Salakhutdinov, R. Learning deep generative models. *Ann. Rev. Stat. Appl.* **2**, 361–385 (2015).
32. Gao, X. & Duan, L.-M. Efficient representation of quantum many-body states with deep neural networks. *Nat. Commun.* **8**, 662 (2017).

33. Carleo, G., Nomura, Y. & Imada, M. Constructing exact representations of quantum many-body systems with deep neural networks. *Nat. Commun.* **9**, 5322 (2018).
34. Sorella, S. Green function Monte Carlo with stochastic reconfiguration. *Phys. Rev. Lett.* **80**, 4558–4561 (1998).
35. Deng, D.-L., Li, X. & Das Sarma, S. Machine learning topological states. *Phys. Rev. B* **96**, 195145 (2017).
36. Nomura, Y., Darmawan, A. S., Yamaji, Y. & Imada, M. Restricted Boltzmann machine learning for solving strongly correlated quantum systems. *Phys. Rev. B* **96**, 205152 (2017).
37. Saito, H. Solving the Bose–Hubbard model with machine learning. *J. Phys. Soc. Jpn* **86**, 093001 (2017).
38. Luo, D. & Clark, B. K. Backflow transformations via neural networks for quantum many-body wave-functions. Preprint at https://arxiv.org/abs/1807.10770 (2018).
39. Han, J., Zhang, L. & E, W. Solving many-electron Schrodinger equation using deep neural networks. Preprint at https://arxiv.org/abs/1807.07014 (2018).
40. Teng, P. Machine-learning quantum mechanics: solving quantum mechanics problems using radial basis function networks. *Phys. Rev. E* **98**, 033305 (2018).
41. Saito, H. Method to solve quantum few-body problems with artificial neural networks. *J. Phys. Soc. Jpn* **87**, 074002 (2018).
42. Choo, K., Carleo, G., Regnault, N. & Neupert, T. Symmetries and many-body excitations with neural-network quantum states. *Phys. Rev. Lett.* **121**, 167204 (2018).
43. Carleo, G., Becca, F., Schiro, M. & Fabrizio, M. Localization and glassy dynamics of many-body quantum systems. *Sci. Rep.* **2**, 243 (2012).
44. Jónsson, B., Bauer, B. & Carleo, G. Neural-network states for the classical simulation of quantum computing. Preprint at https://arxiv.org/abs/1808.05232 (2018).
45. Kaubruegger, R., Pastori, L. & Budich, J. C. Chiral topological phases from artificial neural networks. *Phys. Rev. B* **97**, 195136 (2018).
46. Schmitt, M. & Heyl, M. Quantum dynamics in transverse-field Ising models from classical networks. *SciPost Phys.* **4**, 013 (2018).
47. Czischek, S., Gärttner, M. & Gasenzer, T. Quenches near Ising quantum criticality as a challenge for artificial neural networks. *Phys. Rev. B* **98**, 024311 (2018).
48. Vieijra, T. et al. Restricted Boltzmann machines for quantum states with nonabelian or anyonic symmetries. Preprint at https://arxiv.org/abs/1905.06034 (2019).
49. Vogel, K. & Risken, H. Determination of quasiprobability distributions in terms of probability distributions for the rotated quadrature phase. *Phys. Rev. A* **40**, 2847–2849 (1989).
50. James, D. F. V., Kwiat, P. G., Munro, W. J. & White, A. G. Measurement of qubits. *Phys. Rev. A* **64**, 052312 (2001).
51. Roos, C. F. et al. Bell states of atoms with ultralong lifetimes and their tomographic state analysis. *Phys. Rev. Lett.* **92**, 220402 (2004).
52. Cramer, M. et al. Efficient quantum state tomography. *Nat. Commun.* **1**, 149 (2009).
53. Kim, K. et al. Quantum simulation of frustrated ising spins with trapped ions. *Nature* **465**, 590–593 (2010).
54. Harris, R. et al. Phase transitions in a programmable quantum spin glass simulator. *Science* **361**, 162–165 (2018).
55. King, A. D. et al. Observation of topological phenomena in a programmable lattice of 1,800 qubits. *Nature* **560**, 456–460 (2018).
56. Hastings, M. B., González, I., Kallin, A. B. & Melko, R. G. Measuring Renyi entanglement entropy in quantum Monte Carlo simulations. *Phys. Rev. Lett.* **104**, 157201 (2010).
57. Torlai, G. & Melko, R. G. Latent space purification via neural density operators. *Phys. Rev. Lett.* **120**, 240503 (2018).
58. Rocchetto, A., Grant, E., Strelchuk, S., Carleo, G. & Severini, S. Learning hard quantum distributions with variational autoencoders. *npj Quantum Inf.* **4**, 28 (2018).
59. Goodfellow, I. NIPS 2016 tutorial: generative adversarial networks. Preprint at https://arxiv.org/abs/1701.00160 (2016).
60. Kingma, D. P. & Welling, M. Auto-encoding variational Bayes. Preprint at https://arxiv.org/abs/1312.6114 (2013).
61. van den Oord, A., Kalchbrenner, N. & Kavukcuoglu, K. Pixel recurrent neural networks. Preprint at https://arxiv.org/abs/1601.06759 (2016).
62. Carrasquilla, J., Torlai, G., Melko, R. G. & Aolita, L. Reconstructing quantum states with generative models. *Nat. Mach. Intell.* **1**, 155–161 (2019).
63. Sharir, O., Levine, Y., Wies, N., Carleo, G. & Shashua, A. Deep autoregressive models for the efficient variational simulation of many-body quantum systems. Preprint at https://arxiv.org/abs/1902.04057 (2019).
64. Goodfellow, I. et al. in *Advances in Neural Information Processing Systems* Vol. 27 (eds Ghahramani, Z., Welling, M., Cortes, C., Lawrence, N. D. & Weinberger, K. Q.) 2672–2680 (Curran Associates, 2014).

Acknowledgements

We thank G. Torlai, A. Rocchetto and M. Albergo for fruitful discussions. This research was supported by NSERC of Canada, the Perimeter Institute for Theoretical Physics, the Shared Hierarchical Academic Research Computing Network (SHARCNET) and the National Science Foundation under grant no. PHY-1748958. Research at Perimeter Institute is supported in part by the Government of Canada through the Department of Innovation, Science and Economic Development Canada and by the Province of Ontario through the Ministry of Economic Development, Job Creation and Trade. R.G.M. acknowledges support from a Canada Research Chair. J.C. acknowledges financial and computational support from the AI grant. J.I.C. was funded from the European Research Council (grant agreement no. 742102).

Competing interests

The authors declare no competing interests.

Additional information

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence should be addressed to R.G.M.