Purifying Deep Boltzmann Machines for Thermal Quantum States

Yusuke Nomura[,1,*] Nobuyuki Yoshioka[,2,3,†] and Franco Nori[3,4,5]

[1]*RIKEN Center for Emergent Matter Science, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan*
[2]*Department of Applied Physics, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan*
[3]*Theoretical Quantum Physics Laboratory, RIKEN Cluster for Pioneering Research (CPR), Wako-shi, Saitama 351-0198, Japan*
[4]*RIKEN Center for Quantum Computing (RQC), Wako-shi, Saitama 351-0198, Japan*
[5]*Physics Department, University of Michigan, Ann Arbor, Michigan 48109-1040, USA*

We develop two cutting-edge approaches to construct deep neural networks representing the purified finite-temperature states of quantum many-body systems. Both methods commonly aim to represent the Gibbs state by a highly expressive neural-network wave function, exemplifying the idea of purification. The first method is an entirely deterministic approach to generate deep Boltzmann machines representing the purified Gibbs state exactly. This strongly assures the remarkable flexibility of the ansatz which can fully exploit the quantum-to-classical mapping. The second method employs stochastic sampling to optimize the network parameters such that the imaginary time evolution is well approximated within the expressibility of neural networks. Numerical demonstrations for transverse-field Ising models and Heisenberg models show that our methods are powerful enough to investigate the finite-temperature properties of strongly correlated quantum many-body systems, even when the problematic effect of frustration is present.

Introduction.—The thermal behavior of quantum many-body systems is one of the most fundamental problems in physics. Statistical mechanics states that the density matrix describing a system in thermal equilibrium, governed by a Hamiltonian $\mathcal{H}$ at an inverse temperature β, is given by the Gibbs state $\rho = e^{-\beta\mathcal{H}}/\text{Tr}[e^{-\beta\mathcal{H}}]$. Computing and extracting physical properties from the Gibbs state is a significant challenge to understand natural phenomena, which in reality all occur at finite β.

One of the most celebrated numerical techniques in lattice systems is the quantum Monte Carlo (QMC) method [1–4], typically based on the path integral formalism of the partition function. The QMC method yields numerically exact solutions when the positive definiteness is assured; otherwise, the infamous negative sign problem arises. Many physically intriguing systems fall into the latter category, and therefore various efforts have been devoted to overcome this difficulty: tensor-network-based algorithms [5–8] mostly applied to one-dimensional (1D) systems, dynamical mean-field theory [9] which becomes exact in the infinite coordination-number limit, diagrammatic Monte Carlo methods [10,11], to name a few [12–14]. In another notable approach [15–18] using the thermal pure quantum (TPQ) states, one can extract the ensemble property from a single pure state that represents the thermal equilibrium. We point out, however, that it remains extremely challenging to establish a methodology that is both reliable and scalable for finite-temperature calculations in two-dimensional (2D) systems—the most exotic and intriguing realm in quantum many-body problems.

Neural networks, initially developed for classical data processing in the context of machine learning, offer a very strong methodology for quantum physics [19–25]. As was first demonstrated by Carleo and Troyer [19], neural networks applied as variational wave functions, commonly dubbed as the neural quantum states, are capable of simulating ground states [19,26–28], excited states [29–31], and even out-of-equilibrium property [19,32–38] of strongly correlated systems up to unprecedently large size. Among the tremendous variety of network structures, Boltzmann machines with restricted connectivity, known to be universal approximators for arbitrary real/complex amplitudes [39,40], are useful for statistical mechanics and quantum information. The Boltzmann machines with the shallowest structure are already powerful enough to compactly express complex quantum states with extensively-growing quantum entanglement [41,42]. Furthermore, deep Boltzmann machines (DBMs), i.e., the ones with multiple hidden layers, are guaranteed to provide an efficient description for an even wider range of quantum states [43]. Strongly motivated by their extremely high representability, the ground states in quantum many-body spin systems have been successfully simulated by DBMs [44].

In the present study, we provide two state-of-the-art methods to construct DBMs that capture the finite-temperature behavior of quantum many-body systems. Both methods share the strategy of employing DBMs to express the purified Gibbs state. Namely, a mixed state under imaginary time evolution is compactly encoded as a

pure DBM wave function in the enlarged Hilbert space. In the first method, we find a completely deterministic way to construct DBMs, realizing the exact purified expression of finite temperature states. This proves the remarkable flexibility and power of DBMs for investigating finite-temperature many-body phenomena. In the second method, we provide a stochastic way to simulate the imaginary time evolution which exploits the versatile expressive power of DBMs as approximators. For demonstration, we apply these methods to the 1D transverse-field Ising (TFI) model and the 2D J_1–J_2 Heisenberg model on the square lattice to find surprisingly high accuracy compared to numerically exact methods. We emphasize that only a polynomial number of auxiliary spins suffices in both approaches, yielding a huge computational advantage even under the presence of the problematic effect of frustration.

DBM for purification.—Throughout this work, we utilize the idea of purification to represent the Gibbs state. Namely, the finite-temperature density matrix ρ of a target system $\mathcal{S}$ is encoded as a pure state in an extended system $\mathcal{S} + \mathcal{A}$, such that ρ is obtained by tracing out the ancillary system $\mathcal{A}$ [45]. For instance, the purification of an infinite-temperature state can be generated by the superposition $\sum_x |x\rangle_{\mathcal{S}} |a_x\rangle_{\mathcal{A}}$, where $\{|x\rangle_{\mathcal{S}}\}$ is the complete orthonormal basis set of the target system, while $\{|a_x\rangle_{\mathcal{A}}\}$ is an orthonormal but not necessarily complete basis set of ancillary system.

For the sake of concreteness, let us consider a quantum many-body spin-1/2 system. We introduce N_{site} binary degrees of freedom $\{\sigma_i\}$ so that $|\sigma\rangle = |\sigma_1, \ldots, \sigma_{N_{\text{site}}}\rangle$ spans the Hilbert space of the target system $\mathcal{S}$. Hereafter we call them physical spins. As the ancillary system $\mathcal{A}$, we introduce an identical number of ancilla spins $\{\sigma_i'\}$ so that an arbitrary mixed state can be purified in principle. While the purification of a mixed state is not unique, here we exclusively take the purified infinite-temperature state as $|\Psi(T = \infty)\rangle = \otimes_{i=1}^{N_{\text{site}}} (|\uparrow \downarrow'\rangle + |\downarrow \uparrow'\rangle)_i$, and perform the imaginary time evolution as $|\Psi(T)\rangle = e^{-\beta \mathcal{H}/2} \otimes \mathbb{1}' |\Psi(T = \infty)\rangle$, with $\beta = 1/T$ to simulate the finite-temperature Gibbs state. Note that the infinite-temperature state $\rho_\infty = \mathbb{1}/2^{N_{\text{site}}}$ and the finite-temperature state $\rho_T = e^{-\beta \mathcal{H}}/\text{Tr}[e^{-\beta \mathcal{H}}]$ are reproduced by tracing out the ancilla spins.

Intriguingly, the purified Gibbs state at each temperature can be efficiently expressed by the DBM. In particular, we use the DBM with two hidden layers [see Fig. 1(a)] to represent a purified wave function whose amplitude $\Psi(\sigma, \sigma') = \langle \sigma, \sigma' | \Psi \rangle$ is parametrized as

$$\Psi(\sigma, \sigma') = \sum_{h,d} \phi(\sigma, \sigma'; h, d), \tag{1}$$

$$\phi(\sigma, \sigma'; h, d) = \exp\left[\sum_j b_j h_j + \sum_{ji} h_j(W_{ji}\sigma_i + W'_{ji}\sigma'_i)\right.$$
$$\left. + \sum_{jk} W''_{jk} h_j d_k\right], \tag{2}$$
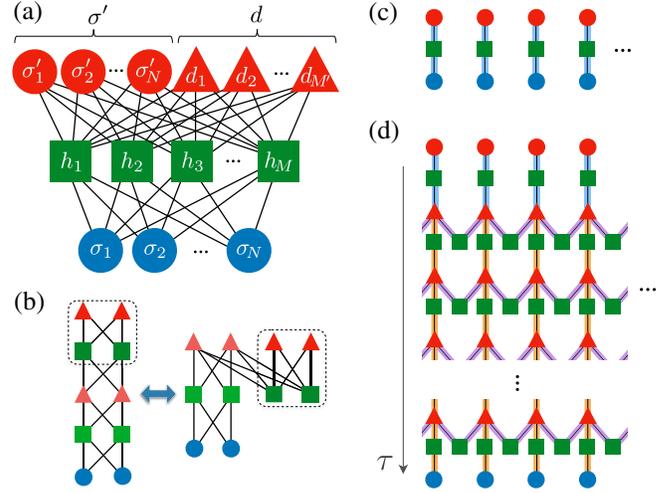
FIG. 1. (a) Structure of the three-layer DBM used in the present study. The visible layer (blue) corresponds to physical spins σ_i. Two hidden layers are distinguished as hidden (green) and deep (red) layers; while the hidden layer simply consists of hidden spins h_j, the deep layer is composed of the deep spins d_k and ancilla spins σ_i', which are introduced to purify the Gibbs-state density matrix. Here, the numbers of physical, hidden, and deep spins are denoted as N, M, M', respectively. The number of ancilla spins is taken as N. (b) Two different ways of depicting the identical DBM structure. A network with seemingly many hidden layers can always be recast into that with only two hidden layers. (c) The DBM representing the infinite-temperature state $|\Psi(T = \infty)\rangle$ for quantum spin-1/2 systems. (d) The DBM construction for the finite-temperature states of the 1D TFI model [$|\Psi(T)\rangle$ in Eq. (3)]. The arrow denotes the growth of the DBM structure along the imaginary-time τ propagation. Light blue, orange, and purple bonds in (c) and (d) have couplings $i(\pi/4)$, $\frac{1}{2}\text{arcosh}[1/\tanh(\Gamma\delta_\tau)]$, and $\frac{1}{2}\text{arcosh}(e^{2J\delta_\tau})$, respectively.

where we have introduced hidden spins $\{h_j\}$ and deep spins $\{d_k\}$, in addition to the physical spins $\{\sigma_i\}$ constituting the visible layer and ancilla spins $\{\sigma_i'\}$ allocated in the second hidden layer (we define all spins to be either $+1$ or -1). This structure is "universal" in terms of the DBM architecture; arbitrary multi-hidden-layer structure can be rearranged to have only two hidden layers as shown in Fig 1(b). The number of complex variational parameters $\mathcal{W} = \{b, W, W'\}$ [46] are directly related to the number of h and d spins, which controls the representability of the DBM wave function. Note that the purification technique for neural networks has been considered in the context of quantum tomography [47] and dissipative quantum physics [34–37].

Method (I): Analytic purification using DBM.—In method (I), we analytically construct DBMs that exactly reproduce the behaviours of Gibbs states. First, to represent the infinite-temperature state, we introduce a DBM with N_{site} hidden and ancilla spins [Fig. 1(c)]. By setting the parameters $\mathcal{W}_\infty$ as $b_j = 0$ and $W_{ji} = W'_{ji} = i(\pi/4)\delta_{ji}$, we find that the analytical expression of the DBM wave

function becomes $\Psi(\sigma, \sigma') = \prod_i 2 \cosh[i(\pi/4)(\sigma_i + \sigma_i')]$. This exactly reproduces the $|\Psi(T = \infty)\rangle$ described above.

Next, to express finite-temperature states $|\Psi(T)\rangle$ analytically, we introduce the Suzuki-Trotter decomposition [1]. Namely, given the Hamiltonian $\mathcal{H} = \sum_\nu \mathcal{H}_\nu$, the purified Gibbs state up to the Trotter error is expressed as

$$|\Psi(T)\rangle := \left(\left[\prod_\nu e^{-\delta_\tau \mathcal{H}_\nu}\right]^{N_\tau} \otimes \mathbb{1}'\right) |\Psi(T = \infty)\rangle, \quad (3)$$

where $\delta_\tau = \beta/2N_\tau$ is the propagation time step. Remarkably, thanks to the flexible representability of the DBM, we can find analytical solutions for Eq. (3): Starting from the initial DBM state $|\Psi(T = \infty)\rangle$ given by the $\mathcal{W}_\infty$ parameters, each short-time propagation $e^{-\delta_\tau \mathcal{H}_\nu}$ can be implemented exactly by modifying the parameters and structure of the DBM [48], whose the explicit expression depends on the form of the Hamiltonian.

As a concrete example, let us consider the TFI model on the N_{site}-spin chain under periodic boundary condition. The Hamiltonian is given by $\mathcal{H} = \mathcal{H}_1 + \mathcal{H}_2$, with $\mathcal{H}_1 = -J \sum_i \sigma_i^z \sigma_{i+1}^z$ and $\mathcal{H}_2 = -\Gamma \sum_i \sigma_i^x$, where $\sigma_i^a (a = x, y, z)$ denotes the Pauli matrix operating on the ith site. We take the Ising-type interaction J as the energy unit ($J = 1$), and Γ as the strength of the transverse magnetic field. Using method (I), we can analytically construct the finite-temperature state using the DBM. Solutions to realize the propagation by updating the DBM parameters from $\mathcal{W}$ to $\bar{\mathcal{W}}$, i.e., $e^{-\delta_\tau \mathcal{H}_\nu} |\Psi_{\mathcal{W}}\rangle = C|\Psi_{\bar{\mathcal{W}}}\rangle$ (C: a constant) can be sketched as follows [44]: For $\nu = 1$ (interaction propagator), we add a single hidden spin for each neighboring visible spins σ_i and σ_{i+1}. For $\nu = 2$ (transverse-field propagator), we add new hidden- and deep-spin layers between the visible and neighboring hidden layers. All the DBM parameters are determined analytically (see Supplemental Material for detailed implementation [50]). As a consequence of propagarions, the DBM architecture grows as in Fig. 1(d), and the number of hidden and deep spins scale as $\mathcal{O}(N_\tau N_{\text{site}})$, which is also common among general local Hamiltonians.

Now that the Gibbs states are represented, let us discuss how the physical quantities are computed using the DBM framework in general. Since the expectation value of a physical observable $\langle \mathcal{O} \rangle = \langle \Psi(T)|\mathcal{O} \otimes \mathbb{1}'|\Psi(T)\rangle / \langle \Psi(T)|\Psi(T)\rangle$ becomes analytically intractable, we use the Monte Carlo (MC) method for its numerical estimation. The sampling weight is based on the expression of the normalization factor of the DBM state [Eq. (1)] given as

$$\langle \Psi|\Psi \rangle = \sum_{\sigma, \sigma'} |\Psi(\sigma, \sigma')|^2$$

$$= \sum_{\sigma, \sigma'} \sum_{h_1, d_1, h_2, d_2} \phi^*(\sigma, \sigma'; h_1, d_1)\phi(\sigma, \sigma'; h_2, d_2). \quad (4)$$
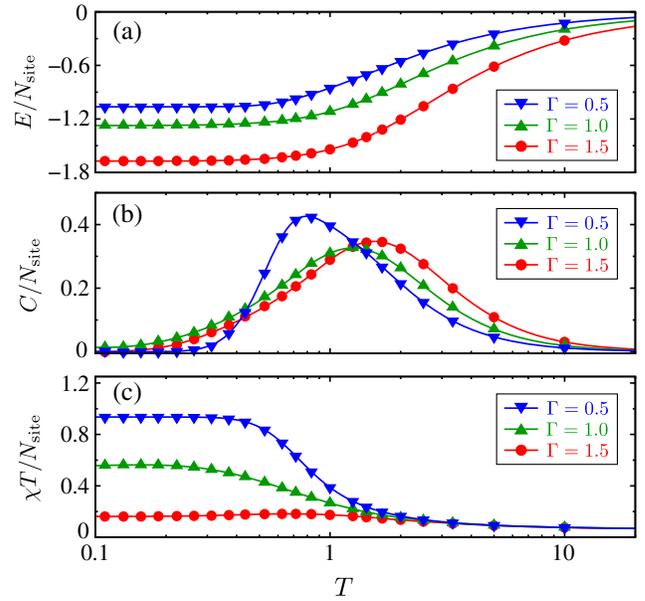
FIG. 2. Finite-temperature calculations for the 1D TFI model on a 16-site chain ($N_{\text{site}} = 16$) with periodic boundary condition: (a) Energy E, (b) specific heat C, and (c) susceptibility $\chi = (1/T) \sum_i \langle \sigma_0 \sigma_i \rangle$. The symbols denote the DBM results [method (I) with $\delta_\tau = 0.05$], which agree well with the exact-diagonalization results (solid curves).

Namely, we sample over the configurations of $(\sigma, \sigma', h_1, h_2, d_1, d_2)$ weighted by the product of amplitudes as $w(\sigma, \sigma'; h, d) \equiv \phi^*(\sigma, \sigma'; h_1, d_1)\phi(\sigma, \sigma'; h_2, d_2)$. Alternatively, it is possible to trace out the hidden spins h analytically, and use $\sum_{h_1, h_2} \phi^*(\sigma, \sigma'; h_1, d_1)\phi(\sigma, \sigma'; h_2, d_2)$ as the MC sampling weight over configurations of $(\sigma, \sigma', d_1, d_2)$ (or trace out d spins and sample over h spins). See Supplemental Material for more details of the method [50].

To verify the construction of the DBM and the proposed MC sampling strategy, we apply the method to the 16-site TFI model. Figure 2 shows the DBM results (symbols) for the temperature dependence of the (a) energy, (b) specific heat, and (c) susceptibility. As expected, the DBM results follow the exact temperature evolution (solid curves). This confirms the remarkable representability of the DBM not only at zero temperature [44] but also at finite temperatures, offering an intriguing quantum-to-classical mapping.

Method (II): Numerical purification using DBM.—When the weight of individual spin configuration can be taken to be always positive, method (I) is quite useful and provides numerically exact finite-temperature results. However, when the frustration exists in the spin Hamiltonian, for instance, we cannot avoid the existence of negative weights as in other finite-temperature calculations based on the QMC method. While it is possible to construct finite-temperature states analytically, the estimation of physical quantities becomes extremely difficult because of the negative sign problem.

To make the application to frustrated models possible, we propose an alternative method which employs DBMs with only ancilla spins σ' in the second hidden layer. By tracing out hidden spins h, such a purified DBM wave function has a simple form:

$$\Psi(\sigma, \sigma') = \prod_j 2\cosh\left[b_j + \sum_i (W_{ji}\sigma_i + W'_{ji}\sigma'_i)\right]. \quad (5)$$

Then, we can avoid negative signs by simply employing $|\Psi(\sigma, \sigma')|^2$ as the weight for sampling over σ and σ' spins. However, in this case, differently from method (I), the imaginary-time evolution cannot be followed analytically. Instead, we need to update parameters numerically at each time so that the DBM obeys the imaginary-time evolution starting from the infinite-temperature state $\Psi_\infty(\sigma, \sigma') = \prod_i 2\cosh[i(\pi/4)(\sigma_i + \sigma'_i)]$ (recall that this infinite-temperature DBM does not require d spins). For this purpose, we employ the stochastic reconfiguration (SR) method [50–52] [53]. The SR optimization is designed to minimize the distance between quantum states following the exact and variational imaginary-time evolution, as much as possible within the expressive power of the DBM wave function in Eq. (5). The expressive power is systematically controlled by the number of h spins; it is ensured that any quantum states can be represented exactly (universal approximation) by an infinitely wide network structure [39,40].

As a demonstration of method (II), we turn to a highly challenging problem: frustrated spin systems. As a representative, here, we consider the 2D antiferromagnetic J_1–J_2 Heisenberg model on the $L \times L$ square lattices with periodic boundary condition. The Hamiltonian reads $\mathcal{H} = J_1 \sum_{\langle i,j \rangle} \mathbf{S}_i \cdot \mathbf{S}_j + J_2 \sum_{\langle\langle i,j \rangle\rangle} \mathbf{S}_i \cdot \mathbf{S}_j$. Here, $\mathbf{S}_i$ is the spin-1/2 operator at site i, and $J_1(= 1)$ and J_2 are the nearest-neighbor and next-nearest-neighbor couplings, respectively. When J_2 is finite, the spin configuration cannot satisfy the energy gain by the J_1 and J_2 interactions simultaneously (frustration). Around $J_2 = 0.5$, where the frustration is strong, an exotic state of matter, quantum spin liquid without any symmetry breaking, might be stabilized as the ground state [21,56–61]. The model also attracts attention because of its possible relevance to the physics of high-T_c cuprates [62–64]. However, because numerically exact QMC results are not available due to the sign problem, the ground-state phase diagram is still under active debate.

For this challenging problem, the wave functions using neural networks have started to be applied to the zero temperature calculation [61,65–70]. However, to detect a hallmark of the possible quantum spin liquid phase experimentally, the finite-temperature behavior needs to be elucidated. Here, we apply method (II) to perform the finite-temperature calculations for $J_2 = 0$ and 0.5. To check its accuracy, we compare the results with numerically exact ones obtained by method (I) for the non-frustrated
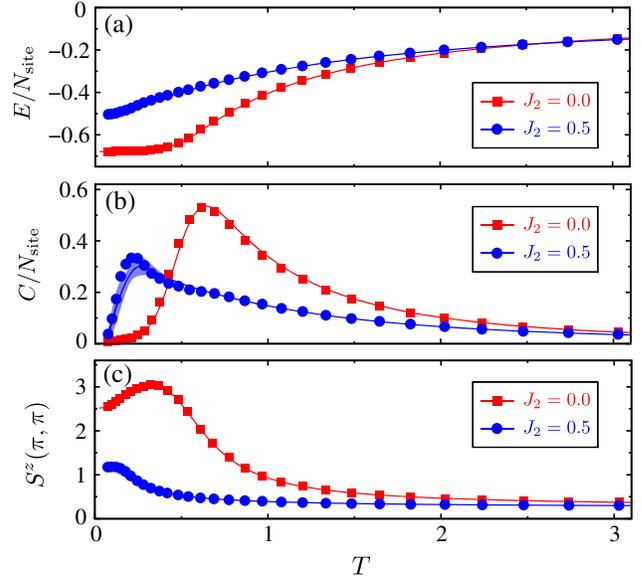
FIG. 3. Finite-temperature calculations for the 2D J_1-J_2 Heisenberg model ($J_1 = 1$) on the 6×6 square lattice with periodic boundary condition: (a) Energy E, (b) specific heat C, and (c) z component of the spin structure factor $S^z(\mathbf{q}) = (1/N_{\text{site}}) \sum_{ij} e^{i\mathbf{q}\cdot(\mathbf{R}_i - \mathbf{R}_j)} \langle S_i^z S_j^z \rangle$ at $\mathbf{q} = (\pi, \pi)$. The symbols denote method (II) results with $8N_{\text{site}}$ hidden spins, which show a good agreement with the numerically exact references (solid curves) obtained by method (I) with $\delta_\tau = 0.005$ ($J_2 = 0$) and the TPQ method ($J_2 = 0.5$). The TPQ calculations are performed with $\mathcal{H}\Phi$ [71]. The shaded regions show the size of the error bars of the TPQ results [in method (I), the size of the error bars is small].

case ($J_2 = 0$). For the frustrated case ($J_2 = 0.5$), method (I) suffers from the sign problem, but, up to a 6×6 lattice, the TPQ results are available, which are also numerically exact. Therefore, we perform calculations using the 6×6 lattice with the total magnetization restricted to be zero ($\sum_i S_i^z = 0$). To further improve the accuracy of the calculation, we utilize the translational and point-group symmetry of the extended system [70]. See Supplemental Material for the practical details [50].

Figure 3 shows method (II) results (symbols) for the temperature dependence of the (a) energy, (b) specific heat, and (c) z component of spin structure factor $S^z(\pi, \pi)$, which quantifies the Néel-type antiferromagnetic correlation. We can see that, by the frustration, the antiferromagnetic correlation is largely suppressed, and the entropy release slows down. Method (II) results accurately reproduce the exact imaginary-time evolution, showing its reliability even in the frustrated regime.

In method (II), by optimizing parameters numerically, we obtain a more compact and dense network to represent finite-temperature states compared to the analytically derived network in method (I): the number of hidden units is $N_h = 8N_{\text{site}}$ in this case, which is in contrast to $\mathcal{O}(N_\tau N_{\text{site}})$ in method (I). The compactness of the network without d spins results in the absence of the negative

weights in the MC sampling. Considering the observation that the N_h scales polynomially with respect to the system size, our results are strongly encouraging to expect a computational advantage in even more dedicated simulations for larger systems.

From the application to the Heisenberg model, we see that both the implementation of symmetrization and increasing N_h contributes to enhancing the accuracy of calculations [50]. It is an important future task to check the reliability of method (II) for other Hamiltonians, since the convergence of the error in the imaginary-time evolution with respect to N_h should be model dependent.

Summary and outlook.—In summary, we have proposed two cutting-edge approaches that utilize DBMs to simulate the finite-temperature properties of quantum many-body systems. In the first approach, we provide a deterministic construction of DBMs that exactly represents Gibbs states, which proves the suitability and flexibility of neural networks for encoding thermal properties. In the second approach, the DBM network parameters are optimized stochastically so that the imaginary time evolution can be approximated efficiently, even for one of the most challenging 2D problems, such as frustrated systems.

Several future directions can be envisioned. It is an interesting question how the neural-network quantum states perform under other schemes of finite-temperature calculation such as TPQ methods. All variational ansatz are, by construction, not powerful enough to express Haar random states, which are taken as the initial states in TPQ calculations. Nonetheless, results by tensor-network-based algorithms [8] imply that using the truncated Hilbert space is sufficient in practical simulations. The major obstacle is considered to be the entanglement growth along the time propagation, which we expect to be simulated well by neural networks, based on previous works on real-time evolution [19,32,33]. Also, it is natural to explore the scalability of our methods in larger and/or more complex systems, or ask whether other network structures (e.g., deep feed-forward networks) are suited for the finite-temperature calculations; the trade-off relationship between the representability and trainability of shallow and deep neural networks remains open.

Y. N. and N. Y. contributed equally to this work.

*yusuke.nomura@riken.jp
†nyoshioka@ap.t.u-tokyo.ac.jp

[1] M. Suzuki, Relationship between d-dimensional quantal spin systems and $(d+1)$-dimensional ising systems: Equivalence, critical exponents and systematic approximants of the partition function and spin correlations, Prog. Theor. Phys. **56**, 1454 (1976).

[2] M. Suzuki, S. Miyashita, and A. Kuroda, Monte Carlo simulation of quantum spin systems. I, Prog. Theor. Phys. **58**, 1377 (1977).

[3] A. W. Sandvik, Stochastic series expansion method with operator-loop update, Phys. Rev. B **59**, R14157 (1999).

[4] J. Gubernatis, N. Kawashima, and P. Werner, *Quantum Monte Carlo Methods* (Cambridge University Press, Cambridge, England, 2016).

[5] F. Verstraete, J. J. García-Ripoll, and J. I. Cirac, Matrix Product Density Operators: Simulation of Finite-Temperature and Dissipative Systems, Phys. Rev. Lett. **93**, 207204 (2004).

[6] M. Zwolak and G. Vidal, Mixed-State Dynamics in One-Dimensional Quantum Lattice Systems: A Time-Dependent Superoperator Renormalization Algorithm, Phys. Rev. Lett. **93**, 207205 (2004).

[7] A. E. Feiguin and S. R. White, Finite-temperature density matrix renormalization using an enlarged Hilbert space, Phys. Rev. B **72**, 220401(R) (2005).

[8] S. R. White, Minimally Entangled Typical Quantum States at Finite Temperature, Phys. Rev. Lett. **102**, 190601 (2009).

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

[10] N. V. Prokof'ev and B. V. Svistunov, Polaron Problem by Diagrammatic Quantum Monte Carlo, Phys. Rev. Lett. **81**, 2514 (1998).

[11] K. V. Houcke, E. Kozik, N. Prokof'ev, and B. Svistunov, Diagrammatic Monte Carlo, Phys. Procedia **6**, 95 (2010).

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

[13] N. Irikura and H. Saito, Neural-network quantum states at finite temperature, Phys. Rev. Research **2**, 013284 (2020).

[14] T. Iitaka, Random phase product sate for canonical ensemble, arXiv:2006.14459.

[15] S. Sugiura and A. Shimizu, Thermal Pure Quantum States at Finite Temperature, Phys. Rev. Lett. **108**, 240401 (2012).

[16] S. Sugiura and A. Shimizu, Canonical Thermal Pure Quantum State, Phys. Rev. Lett. **111**, 010401 (2013).

[17] K. Takai, K. Ido, T. Misawa, Y. Yamaji, and M. Imada, Finite-temperature variational Monte Carlo method for strongly correlated electron systems, J. Phys. Soc. Jpn. **85**, 034601 (2016).

[18] A. Iwaki, A. Shimizu, and C. Hotta, Thermal pure quantum matrix product states recovering a volume law entanglement, Phys. Rev. Research **3**, 022015 (2021).

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

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

[21] L. Wang and A. W. Sandvik, Critical Level Crossings and Gapless Spin Liquid in the Square-Lattice Spin-1/2 J_1–J_2 Heisenberg Antiferromagnet, Phys. Rev. Lett. **121**, 107202 (2018).

[22] G. Carleo, I. Cirac, K. Cranmer, L. Daudet, M. Schuld, N. Tishby, L. Vogt-Maranto, and L. Zdeborová, Machine learning and the physical sciences, Rev. Mod. Phys. **91**, 045002 (2019).

[23] R. G. Melko, G. Carleo, J. Carrasquilla, and J. I. Cirac, Restricted Boltzmann machines in quantum physics, Nat. Phys. **15**, 887 (2019).

[24] A. Melkani, C. Gneiting, and F. Nori, Eigenstate extraction with neural-network tomography, Phys. Rev. A **102**, 022412 (2020).

[25] S. Ahmed, C. S. Muñoz, F. Nori, and A. F. Kockum, Quantum state tomography with conditional generative adversarial networks, arXiv:2008.03240.

[26] Y. Nomura, A. S. Darmawan, Y. Yamaji, and M. Imada, Restricted Boltzmann machine learning for solving strongly correlated quantum systems, Phys. Rev. B **96**, 205152 (2017).

[27] H. Saito, Solving the Bose-Hubbard model with machine learning, J. Phys. Soc. Jpn. **86**, 093001 (2017).

[28] Z. Cai and J. Liu, Approximating quantum many-body wave functions using artificial neural networks, Phys. Rev. B **97**, 035116 (2018).

[29] K. Choo, G. Carleo, N. Regnault, and T. Neupert, Symmetries and Many-Body Excitations with Neural-Network Quantum States, Phys. Rev. Lett. **121**, 167204 (2018).

[30] Y. Nomura, Machine learning quantum states—extensions to fermion–boson coupled systems and excited-state calculations, J. Phys. Soc. Jpn. **89**, 054706 (2020).

[31] N. Yoshioka, W. Mizukami, and F. Nori, Solving quasiparticle band spectra of real solids using neural-network quantum states, Commun. Phys. **4**, 106 (2021).

[32] M. Schmitt and M. Heyl, Quantum Many-Body Dynamics in Two Dimensions with Artificial Neural Networks, Phys. Rev. Lett. **125**, 100503 (2020).

[33] P. Xie and E. Weinan, Coarse-grained spectral projection: A deep learning assisted approach to quantum unitary dynamics, Phys. Rev. B **103**, 024304 (2021).

[34] N. Yoshioka and R. Hamazaki, Constructing neural stationary states for open quantum many-body systems, Phys. Rev. B **99**, 214306 (2019).

[35] A. Nagy and V. Savona, Variational Quantum Monte Carlo Method with a Neural-Network Ansatz for Open Quantum Systems, Phys. Rev. Lett. **122**, 250501 (2019).

[36] M. J. Hartmann and G. Carleo, Neural-Network Approach to Dissipative Quantum Many-Body Dynamics, Phys. Rev. Lett. **122**, 250502 (2019).

[37] F. Vicentini, A. Biella, N. Regnault, and C. Ciuti, Variational Neural-Network Ansatz for Steady States in Open Quantum Systems, Phys. Rev. Lett. **122**, 250503 (2019).

[38] D. Yuan, H. Wang, Z. Wang, and D. L. Deng, Solving the Liouvillian Gap with Artificial Neural Networks, Phys. Rev. Lett. **126**, 160401 (2021).

[39] N. L. Roux and Y. Bengio, Representational power of restricted Boltzmann machines and deep belief networks, Neural Comput. **20**, 1631 (2008).

[40] G. Montufar and N. Ay, Refinements of universal approximation results for deep belief networks and restricted Boltzmann machines, Neural Comput. **23**, 1306 (2011).

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

[42] Y. Levine, Or Sharir, N. Cohen, and A. Shashua, Quantum Entanglement in Deep Learning Architectures, Phys. Rev. Lett. **122**, 065301 (2019).

[43] X. Gao and L.-M. Duan, Efficient representation of quantum many-body states with deep neural networks, Nat. Commun. **8**, 662 (2017).

[44] G. Carleo, Y. Nomura, and M. Imada, Constructing exact representations of quantum many-body systems with deep neural networks, Nat. Commun. **9**, 5322 (2018).

[45] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, England, 2000).

[46] The renormalization factor is omitted in the expression. We ignore bias terms for visible and deep units because they are irrelevant in the present case.

[47] G. Torlai and R. G. Melko, Latent Space Purification via Neural Density Operators, Phys. Rev. Lett. **120**, 240503 (2018).

[48] For spin-1/2 systems, imaginary-time propagators can be analytically represented by the DBM. This has been discussed in the context of classical many-body interacting systems as well [49].

[49] N. Yoshioka, Y. Akagi, and H. Katsura, Transforming generalized Ising models into Boltzmann machines, Phys. Rev. E **99**, 032113 (2019).

[50] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.127.060601 for more details of the proposed two methods and supplemental benchmark data.

[51] S. Sorella, Green Function Monte Carlo with Stochastic Reconfiguration, Phys. Rev. Lett. **80**, 4558 (1998).

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

[53] The same optimization scheme has been developed independently in artificial-intelligence community, called "natural gradient" [54,55].

[54] S.-I. Amari, K. Kurata, and H. Nagaoka, Information geometry of Boltzmann machines, IEEE Trans. Neural Networks **3**, 260 (1992).

[55] S.-I. Amari, Natural Gradient Works Efficiently in Learning, Neural Comput. **10**, 251 (1998).

[56] H.-C. Jiang, H. Yao, and L. Balents, Spin liquid ground state of the spin-$\frac{1}{2}$ square J_1–J_2 Heisenberg model, Phys. Rev. B **86**, 024424 (2012).

[57] W.-J. Hu, F. Becca, A. Parola, and S. Sorella, Direct evidence for a gapless Z_2 spin liquid by frustrating Néel antiferromagnetism, Phys. Rev. B **88**, 060402(R) (2013).

[58] S. Morita, R. Kaneko, and M. Imada, Quantum Spin Liquid in Spin 1/2 J_1–J_2 Heisenberg model on square lattice: Many-variable variational Monte Carlo study combined with quantum-number projections, J. Phys. Soc. Jpn. **84**, 024720 (2015).

[59] F. Ferrari and F. Becca, Gapless spin liquid and valence-bond solid in the J_1–J_2 Heisenberg model on the square lattice: Insights from singlet and triplet excitations, Phys. Rev. B **102**, 014417 (2020).

[60] W.-Y. Liu, S.-S. Gong, Y.-B. Li, D. Poilblanc, W.-Q. Chen, and Z.-C. Gu, Gapless quantum spin liquid and global phase diagram of the spin-1/2 J_1–J_2 square antiferromagnetic Heisenberg model, arXiv:2009.01821.

[61] Y. Nomura and M. Imada, Dirac-type nodal spin liquid revealed by machine learning, arXiv:2005.14142.

[62] F. Nori and G. T. Zimanyi, Comparison of frustration and doping effects in generalized flux phases, Europhys. Lett. **16**, 397 (1991).

[63] F. Nori, E. Gagliano, and S. Bacci, Does Frustration Describe Doping in Models for High-Temperature Superconductors?, Phys. Rev. Lett. **68**, 240 (1992).

[64] F. Nori, R. Merlin, S. Haas, A. W. Sandvik, and E. Dagotto, Magnetic Raman Scattering in Two-Dimensional Spin-1/2 Heisenberg Antiferromagnets: Spectral Shape Anomaly and Magnetostrictive Effects, Phys. Rev. Lett. **75**, 553 (1995).

[65] X. Liang, W.-Y. Liu, P.-Z. Lin, G.-C. Guo, Y.-S. Zhang, and L. He, Solving frustrated quantum many-particle models with convolutional neural networks, Phys. Rev. B **98**, 104426 (2018).

[66] K. Choo, T. Neupert, and G. Carleo, Two-dimensional frustrated $J_1 - J_2$ model studied with neural network quantum states, Phys. Rev. B **100**, 125124 (2019).

[67] F. Ferrari, F. Becca, and J. Carrasquilla, Neural Gutzwiller-projected variational wave functions, Phys. Rev. B **100**, 125131 (2019).

[68] T. Westerhout, N. Astrakhantsev, K. S. Tikhonov, M. I. Katsnelson, and A. A. Bagrov, Generalization properties of neural network approximations to frustrated magnet ground states, Nat. Commun. **11**, 1593 (2020).

[69] A. Szabó and C. Castelnovo, Neural network wave functions and the sign problem, Phys. Rev. Research **2**, 033075 (2020).

[70] Y. Nomura, Helping restricted Boltzmann machines with quantum-state representation by restoring symmetry, J. Phys. Condens. Matter **33**, 174003 (2021).

[71] M. Kawamura, K. Yoshimi, T. Misawa, Y. Yamaji, S. Todo, and N. Kawashima, Quantum lattice model solver $\mathcal{H}\Phi$, Comput. Phys. Commun. **217**, 180 (2017).

[72] G. Carleo, K. Choo, D. Hofmann, J. E. T. Smith, T. Westerhout, F. Alet, E. J. Davis, S. Efthymiou, I. Glasser, S.-H. Lin, M. Mauri, G. Mazzola, C. B. Mendl, E. van Nieuwenburg, O. O'Reilly, H. Theveniaut, G. Torlai, F. Vicentini, and A. Wietek, Netket: A machine learning toolkit for many-body quantum systems, SoftwareX **10**, 100311 (2019).