

## ARTICLE OPEN



# Unitary-coupled restricted Boltzmann machine ansatz for quantum simulations

Chang Yu Hsieh<sup>1✉</sup>, Qiming Sun<sup>2</sup>, Shengyu Zhang<sup>1</sup> and Chee Kong Lee<sup>2✉</sup>

Neural-network quantum state (NQS) has attracted significant interests as a powerful wave-function ansatz to model quantum phenomena. In particular, a variant of NQS based on the restricted Boltzmann machine (RBM) has been adapted to model the ground state of spin lattices and the electronic structures of small molecules in quantum devices. Despite these progresses, significant challenges remain with the RBM-NQS-based quantum simulations. In this work, we present a state-preparation protocol to generate a specific set of complex-valued RBM-NQS, which we name the unitary-coupled RBM-NQS, in quantum circuits. Our proposal expands the applicability of NQS as prior works deal exclusively with real-valued RBM-NQS for quantum algorithms. With this scheme, we achieve (1) modeling complex-valued wave functions, (2) using as few as one ancilla qubit to simulate  $M$  hidden spins in an RBM architecture, and (3) avoiding post-selections to improve scalability.

*npj Quantum Information* (2021)7:19; <https://doi.org/10.1038/s41534-020-00347-1>

## INTRODUCTION

Hybrid quantum-classical (HQC) algorithms<sup>1</sup> offer an exciting avenue to explore the potential of a noisy intermediate-scale quantum<sup>2</sup> device without quantum error corrections. The HQC algorithms run on parametrized quantum circuits aiming to minimize an objective function, such as the average energy. The Variational Quantum Eigensolver<sup>3</sup> (VQE) and Quantum Approximate Optimization Algorithm<sup>4</sup> are two prominent examples leading the current wave of HQC algorithm developments. In particular, VQE has been experimentally demonstrated on several leading platforms of quantum computations<sup>3,5–8</sup>. These encouraging experimental outcomes strengthen our anticipation that quantum simulations<sup>9–13</sup> should be among the first set of applications to benefit from quantum computations. Nevertheless, it is also becoming increasingly clear that further developments<sup>14–19</sup> are required to improve VQE and similar HQC algorithms if the goal is to establish an unambiguous quantum advantage for problems of realistic interests. For instance, many recent developments attempt to address the following aspects: (1) design wave-function ansatz<sup>5,14,20–25</sup> with efficient usage of variational parameters and circuit depth, (2) reduce the number of required measurements<sup>26–33</sup>, and (3) overcome the challenge of high-dimensional optimization<sup>34–43</sup> needed for training parametrized circuits. Any of these technical challenges could potentially become a computational bottleneck for an HQC algorithm beyond the small-scale testings reported in the recent literature.

In this work, our primary focus is to investigate whether neural-network quantum states (NQS)<sup>44–46</sup> can be tailored to better fit the paradigm of HQC algorithms. We focus on a particular form of NQS based on the restricted Boltzmann Machine (RBM) architecture. Within the communities of computational many-body physics, quantum information and condensed matter physics, there is a growing trend of adopting neural networks techniques, such as the RBM architecture, for various applications. Notable examples include identification of different phases of matter<sup>47–52</sup>, scalable quantum state tomography<sup>53,54</sup>, efficient sampler to accelerate Monte Carlo simulations<sup>55–57</sup>, quantum error correction codes<sup>58–61</sup>, and variational ansatz for many-body simulations<sup>62–72</sup>.

Especially, the work by Carleo and Troyer<sup>62</sup> demonstrated that a complex-valued RBM model can efficiently model many-body wave functions with fewer variational parameters than tensor-network methods for some spin-lattice models. Subsequent investigations<sup>73–78</sup> have further clarified and affirmed the usefulness of variants of Boltzmann machines to model many-body quantum states of complex systems. These encouraging results and rapidly accumulated knowledge about RBM-NQS motivated us to investigate whether it is suitable to apply this family of ansatz in the context of quantum simulation algorithms. In most part of this work, we should use RBM-NQS and NQS interchangeably when there is no risk of ambiguity.

While there are already quantum simulation algorithms<sup>79,80</sup> using NQS as variational ansatz with encouraging results, some fundamental obstacles limit their scope of applications. For instance, the existing approaches require the preparation of an extended wave function composed of all visible and hidden spins. As each hidden spin is explicitly modeled with an ancilla qubit, these prior methods consume too many qubits, which are expensive resources for near-term quantum devices. Furthermore, there is no scalable strategy for preparing a general NQS in quantum circuits. This is because many NQS can only be obtained via non-unitary transformations on an input state. Existing NQS-based simulation algorithm<sup>79</sup> relies on a probabilistic post-selection to achieve the non-unitary operations. Finally, the lack of complex parameters severely limits the usefulness of the NQS for quantum simulations. For instance, (1) complex-valued wave functions allow us to simulate fermions in time reversal symmetry breaking systems such as electrons in the presence of a magnetic field. (2) To simulate quantum dynamics, it is a necessity to account for the accumulation of dynamical phase factors.

The aforementioned limitations certainly have cast doubts on whether the NQS should be used in quantum simulations; despite many of their theoretical merits as wave-function ansatz and convincing demonstrations in classical simulations cited above. To address these deficiencies, we propose a state-preparation protocol for creating complex-valued NQS in a quantum circuit. In particular, the state-preparation protocol does not use  $N + M$

<sup>1</sup>Tencent Quantum Lab, 518057 Shenzhen, Guangdong, China. <sup>2</sup>Tencent America, Palo Alto, CA 94306, USA. ✉email: kimhsieh@tencent.com; cheekonglee@tencent.com

qubits to model  $N$  visible spins and  $M$  hidden spins explicitly. This is because every term in a RBM Hamiltonian commutes with each other, we can explicitly arrange the order in which the unitary gates acting on the hidden spins. Hence, a single ancilla qubit (representing one hidden spin) can be recycled upon measurement and be reused to represent another hidden spin in a subsequent stage. This qubit-recycle scheme<sup>81,82</sup> tremendously reduces the number of physical qubits (down to just one extra ancilla at the bare minimum) needed to execute the proposed state-preparation protocol. This advantage cannot be underestimated as typical RBM-NQS might use as many as  $M = \text{poly}(N)$  hidden spins.

In order to avoid a probabilistic state preparation discussed in ref. <sup>79</sup>, we have done two things differently. First, we consider a further restricted subclass of RBM-NQS ansatz, dubbed the unitary-coupled RBM-NQS. For these states, the visible and hidden spins are only coupled by unitary operators, i.e. the coupling parameters for a RBM wave function are purely imaginary-valued. Owing to unitary couplings, a state preparation circuit scales as  $\mathcal{O}(NM)$  in circuit depth. There is no obvious disadvantage to confine an ansatz selection to unitary-coupled RBM-NQS. As thoroughly analyzed in Supplementary Note 2, arbitrary RBM-NQS can be systematically converted into a unitary-coupled form without sacrificing representation accuracy. Furthermore, numerical investigations in this work demonstrate that unitary-coupled RBM can accurately model a wide range of physical systems without excessive number of hidden spins.

The second thing we have done differently is to modify a strict post-selection on hidden spins. As explained later, an NQS is properly generated when all hidden spins are projected onto  $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$ . In this work, we observe that a single NQS can be decomposed into an ensemble of modified NQS, which are essentially outputted by a quantum circuit when a projective measurement on hidden spins yields a state other than  $|+\rangle$ . Under this ensemble scheme, expectation values for any observable of the original NQS can be estimated from these modified NQS within a Monte Carlo framework.

Finally, we modify the imaginary-time-dependent variational algorithm<sup>34</sup> to better suit the specific structure of an NQS. We term this algorithm NQS-imaginary-time evolution (NQS-ITE). Due to the commutative structure of a NQS ansatz, the simulation algorithm based on the imaginary-time-dependent variational principle (ITDVP) can be drastically simplified as one just needs to perform measurement on one quantum circuit instead of working with  $\mathcal{O}(N_{\text{var}}^2)$  different quantum circuits where  $N_{\text{var}}$  is the number of total variational parameters. For large-scale simulations,  $N_{\text{var}} \propto \mathcal{O}(MN)$  could be a huge number.

In summary, the present work has both expanded the scope of application for NQS-based HQC simulations and has lowered the barrier for experimental implementations. In order to frame the significance of this work in a better perspective, we summarize and clarify the challenges we set out to address in prior works in Supplementary Note 1.

## RESULTS

### Preparing a complex-valued NQS in a quantum circuit

A brief introduction of NQS can be found in the “Methods” section. All NQS may be obtained from at least one bipartite Ising Hamiltonian  $\hat{H}_{\text{RBM}}(\theta) = \sum_i b_i \hat{v}_i^z + \sum_j m_j \hat{h}_j^z + \sum_{ij} W_{ij} \hat{v}_i^z \hat{h}_j^z$ , where  $\hat{v}_i^z$  or  $\hat{h}_j^z$  is the Pauli Z operator for the visible or hidden qubit, respectively. We also denote complex-valued RBM parameters as  $\theta = [b_1, \dots, b_N, m_1, \dots, m_M, W_{11}, \dots, W_{NM}]$ , and use superscripts R and I to denote the real and imaginary parts, respectively. The complex-valued NQS can be created with a two-step approach. First, entangle  $N + M$  qubits (including all visible and hidden spins

of an RBM architecture) according to

$$|\Psi_{\text{vh}}(\theta)\rangle = \left[ \frac{e^{\hat{H}_{\text{RBM}}^{\text{R}}(\theta)}}{\sqrt{{}_{\text{vh}}\langle + \dots + | e^{2\hat{H}_{\text{RBM}}^{\text{R}}(\theta)} | + \dots + \rangle_{\text{vh}}}} \right] | + \dots + \rangle_{\text{vh}}, \quad (1)$$

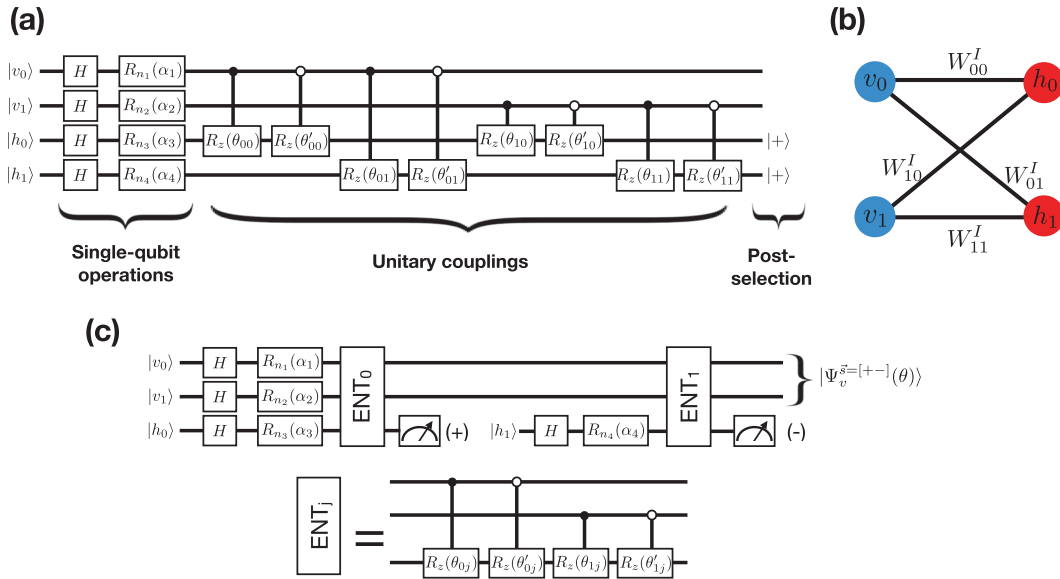
where  $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ ,  $\hat{H}_{\text{RBM}}^{\text{R}}(\theta)$  is the Hermitian part of the RBM Hamiltonian and the subscript vh denotes visible and hidden (ancilla) qubits. Equation (1) gives a conceptually simple wave function that could be generated by first applying single-qubit transformations  $\exp(b_i \hat{v}_i^z)$  and  $\exp(m_j \hat{h}_j^z)$  on individual qubits followed by  $\exp(W_{ij} \hat{v}_i^z \hat{h}_j^z)$  to couple qubits. The quantum operations are non-unitary when RBM parameters take on real parts, i.e.  $b_i^{\text{R}} \neq 0$ ,  $m_j^{\text{R}} \neq 0$ , or  $W_{ij}^{\text{R}} \neq 0$ . In general, the non-unitary two-qubit operation mediating entanglement across the visible-hidden layer are difficult to implement. One could adopt the probabilistic scheme introduced in ref. <sup>79</sup> to generate the inter-layer couplings with an extra ancilla qubit. However, for complex-valued wave function, this approach in ref. <sup>79</sup> is difficult to scale with the number of qubits involved. An alternative to implementing the non-unitary two-qubit operation is the quantum imaginary-time evolution algorithm proposed by ref. <sup>83</sup>; however, such algorithm requires deep circuits for systems with long correlation length.

Once the extended wave function  $|\Psi_{\text{vh}}(\theta)\rangle$  is generated, all ancilla qubits (i.e. hidden spins) are post-selected for  $|+\rangle_{\text{h}}$  and the desired NQS in Eq. (16) is reconstructed in the quantum circuit,

$$\begin{aligned} |\Psi_{\text{v}}(\theta)\rangle &= \frac{{}_{\text{h}}\langle + \dots + | \Psi_{\text{vh}}(\theta) \rangle}{\sqrt{{}_{\text{vh}}\langle \Psi_{\text{vh}}(\theta) | \hat{P}_+^{(\text{h})} | \Psi_{\text{vh}}(\theta) \rangle}} \\ &= \frac{1}{N_{\text{v}}} \sum_{\mathbf{h}} e^{\hat{H}_{\text{RBM}}(\theta, \mathbf{h})} | + \dots + \rangle_{\text{v}}, \end{aligned} \quad (2)$$

where  $\hat{P}_+^{(\text{h})} = | + \dots + \rangle_{\text{h}} \langle + \dots + |$ ,  $N_{\text{v}} = \sqrt{{}_{\text{vh}}\langle + \dots + | e^{\hat{H}_{\text{RBM}}^{\text{R}}(\theta)} \hat{P}_+^{(\text{h})} e^{\hat{H}_{\text{RBM}}^{\text{R}}(\theta)} | + \dots + \rangle_{\text{vh}}}$  and  $e^{\hat{H}_{\text{RBM}}(\theta, \mathbf{h})}$  is an operator acting on the visible spins only as we replace the Pauli operator  $\hat{h}_j^z$  with a binary value ( $\pm 1$ ) of  $\mathbf{h} \in \mathbf{h} = [h_1, \dots, h_M]$ . From the second line of Eq. (2), it is clear that the hidden spins jointly mediate a specific quantum transformation,  $\sum_{\mathbf{h}} e^{\hat{H}_{\text{RBM}}(\theta, \mathbf{h})}$ , on the visible-spin wave function. Since this transformation involves post-selection and non-unitary in general, the amplitude-amplification type of techniques<sup>84,85</sup> and other more advanced techniques such as linear combination of unitaries<sup>86</sup> are not convenient due to the need for extra quantum resources and longer quantum circuits are generally expected. Overall, the challenge to efficiently implement  $\sum_{\mathbf{h}} e^{\hat{H}_{\text{RBM}}(\theta, \mathbf{h})}$  creates another obstacle to prepare complex-valued NQS in quantum circuits.

In the rest of this section, we describe a scalable state-preparation protocol that overcomes these two obstacles. In particular, we implement complex-valued NQS using as few as one ancilla qubit and entirely avoids the post-selection. To simplify presentations, we illustrate how to prepare a subset of NQS that we dub the unitary-coupled RBM-NQS, which only allow purely imaginary-valued inter-layer couplings  $W_{ij} = iW_{ij}^{\text{I}}$ . Generation of unitary-coupled RBM-NQS bypass the inherent challenge to mediate entanglement via non-unitary transformations. Figure 1a gives a circuit diagram of preparing a unitary-coupled RBM-NQS composed of two visible spins with inter-qubit couplings mediated by two hidden spins. The Hadamard gates prepare the  $|+\rangle$  state, and the parametrized single-qubit rotations are not fixed along the z-axis because of the non-unitary operations,  $\exp(b_i^{\text{R}} \hat{v}_i^z)$  and  $\exp(m_j^{\text{R}} \hat{h}_j^z)$ . In the circuit diagram of Fig. 1a, the single-qubit rotations  $R_n(\theta)$  are determined via relations of the form  $\exp(b_i^{\text{R}} \hat{v}_i^z) | + \rangle / c = R_n(\theta) | + \rangle$  with the



**Fig. 1** Quantum circuits for preparing unitary-coupled RBM states. Unitary-coupled RBM-NQS. **a** Quantum circuit to directly prepare a two-qubit unitary-coupled RBM state having two hidden spins. The inter-layer coupling is mediated by unitary gates. Note  $\theta'_{ij} = -\theta_{ij}$ . **b** Schematic of the RBM state generated by the quantum circuit in panel **a**. **c** Scalable preparation of unitary-coupled RBM state described in Eqs. (3)–(5). A sample circuit in which the hidden spins are projected onto  $|+-\rangle$  is shown.

normalization factor  $c = \sqrt{\langle + | \exp(2b_i^R \hat{v}_i^Z) | + \rangle}$ . Figure 1b gives a schematic depicting the RBM state generated by the circuit in Fig. 1a. As explained in Supplementary Note II, the unitary-coupled RBM-NQS does not necessarily suffer loss of expressive power. While we claim it is better to model quantum systems with unitary-coupled RBM-NQS for near-term applications; there is a straightforward extension of the current protocol to generate arbitrary complex-valued NQS in case it is desired. We defer the discussion of this extension to Supplementary Note VI.

### Scalable preparation of a unitary-coupled RBM-NQS in a quantum circuit

To begin, we note that Eq. (2) can be cast in an alternative form

$$|\Psi_v(\theta)\rangle = \frac{1}{N_v} \left[ \langle + | \left[ e^{i\hat{h}_M^Z(m_M + \sum_i iW_{M,i}^I \hat{v}_i^Z)} \right] | + \rangle \right]_M \left[ \langle + | \left[ e^{i\hat{h}_{M-1}^Z(m_{M-1} + \sum_i iW_{M-1,i}^I \hat{v}_i^Z)} \right] | + \rangle \right]_{M-1} \cdots \times \left[ \langle + | \left[ e^{i\hat{h}_1^Z(m_1 + \sum_i iW_{1,i}^I \hat{v}_i^Z)} \right] | + \rangle \right]_1 e^{\sum_i b_i \hat{v}_i^Z} | + \cdots + \rangle_v. \quad (3)$$

Each block,  $[\langle + | \left[ e^{i\hat{h}_j^Z(m_j + \sum_i iW_{j,i}^I \hat{v}_i^Z)} \right] | + \rangle]_j$ , encodes  $j$ th hidden spin's effects on all visible ones. Clearly, one can use a single ancilla qubit to implement these transformations sequentially. As shown in Eq. (3), we specifically consider  $W_{ij} = iW_{ij}^I$  for the unitary-coupled RBM-NQS. Our proposed approach to bypass the post-selection of  $|+\rangle$  on all hidden spins is inspired by the following observation,

$$\begin{aligned} \langle + | \left[ e^{i\hat{h}_j^Z(m_j + \sum_i iW_{j,i}^I \hat{v}_i^Z)} \right] | + \rangle &= \sum_{s=\pm} \langle + | e^{im_j^R \hat{h}_j^Z} | s \rangle \langle s | e^{i(m_j^I + \sum_i iW_{j,i}^I \hat{v}_i^Z) \hat{h}_j^Z} | + \rangle \\ &= \sum_{s=\pm} R_s(m_j^R) \langle s | e^{i(m_j^I + \sum_i iW_{j,i}^I \hat{v}_i^Z) \hat{h}_j^Z} | + \rangle, \end{aligned} \quad (4)$$

where a resolution of identity  $\sum_{s=\pm} |s\rangle \langle s|$  for the ancilla qubit is inserted in the middle and  $R_s(m_j^R) = \langle + | e^{im_j^R \hat{h}_j^Z} | s \rangle$  can be computed classically as it is the transformation matrix element associated with a single qubit. Not obtaining  $R_s(m_j)$  experimentally is the key to avoid post-selection. Note that the decomposition of  $\exp(i\hat{h}_j^Z(m_j + \sum_i iW_{j,i}^I \hat{v}_i^Z))$  introduced in Eq. (4) is exact as these

operators commute. Using Eq. (4), we re-write Eq. (3),

$$\begin{aligned} |\Psi_v(\theta)\rangle &= \sum_{s_M=\pm} \cdots \sum_{s_1=\pm} \frac{1}{N_v} \left( \prod_{j=1}^M R_{s_j}(m_j^R) \right) \langle s_M | \left[ e^{i\hat{h}_M^Z(m_M^I + \sum_i iW_{M,i}^I \hat{v}_i^Z)} \right] | + \rangle \cdots \\ &\quad \times \langle s_1 | \left[ e^{i\hat{h}_1^Z(m_1^I + \sum_i iW_{1,i}^I \hat{v}_i^Z)} \right] | + \rangle e^{\sum_i b_i \hat{v}_i^Z} | + \cdots + \rangle_v \\ &= \sum_{s_M=\pm} \cdots \sum_{s_1=\pm} \frac{N_{\vec{s}}}{N_v} \left( \prod_{j=1}^M R_{s_j}(m_j^R) \right) |\Psi_v^{\vec{s}}(\theta)\rangle, \end{aligned} \quad (5)$$

where  $\vec{s} = [s_1, \dots, s_M]$ .  $|\Psi_v^{\vec{s}}(\theta)\rangle$  is a visible-spin wave function created by projecting hidden spins onto basis states  $|s_1 \cdots s_M\rangle_h$  instead of enforcing the post-selection  $|+\cdots+\rangle_h$ . Due to the decomposition introduced in Eq. (4), only a portion of  $\exp(\hat{H}_{\text{RBM}})$  contributes to the generation of  $|\Psi_v^{\vec{s}}(\theta)\rangle$  and  $N_{\vec{s}}$  is the

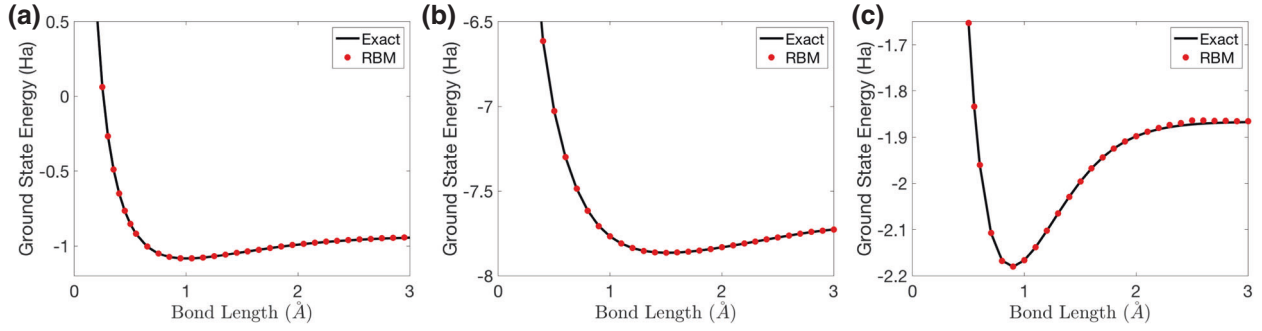
normalization to keep  $\langle \Psi_v^{\vec{s}} | \Psi_v^{\vec{s}} \rangle = 1$ . While there is no post-selection in Eq. (5), there is now a summation over all possible  $\vec{s}$ .

Instead of working directly with Eq. (5), we are primarily interested in the expectation value of an observable  $\hat{O}$ , which can be formulated as

$$\langle \Psi_v(\theta) | \hat{O} | \Psi_v(\theta) \rangle = \int d\mathbf{z} |\langle \mathbf{z} | \Psi_v(\theta) \rangle|^2 \left[ \int d\mathbf{z}' O(\mathbf{z}, \mathbf{z}') \frac{\langle \mathbf{z}' | \Psi_v(\theta) \rangle}{\langle \mathbf{z} | \Psi_v(\theta) \rangle} \right]. \quad (6)$$

The equation above can be interpreted as follows. The expectation value of an observable  $\hat{O}$  can be turned into the average of the expression inside the square bracket if we can efficiently sample  $\mathbf{z}$  according to the probability density  $|\langle \mathbf{z} | \Psi_v(\theta) \rangle|^2$ , i.e. projecting the NQS in some basis. We further analyze this probability density by exposing the details of the RBM architecture,

$$\begin{aligned} |\langle \mathbf{z} | \Psi_v(\theta) \rangle|^2 &= \frac{|\langle \mathbf{z} | \Psi_v(\theta) \rangle|^2}{\langle \Psi_v(\theta) | \Psi_v(\theta) \rangle} \\ &= \frac{\sum_{\vec{s}} \frac{N_{\vec{s}}^2}{N_v^2} \left( \prod_{j=1}^M R_{s_j}^2(m_j^R) \right) |\langle \mathbf{z} | \Psi_v^{\vec{s}}(\theta) \rangle|^2}{\sum_{\vec{s}} \frac{N_{\vec{s}}^2}{N_v^2} \left( \prod_{j=1}^M R_{s_j}^2(m_j^R) \right)}, \end{aligned} \quad (7)$$



**Fig. 2 Numerical results for molecular systems.** The ground state energy of  $H_2$ ,  $LiH$ , and linear  $H_4$  molecules as a function of inter-nuclear distances. The energies computed with the NQS-ITE algorithm and the exact diagonalization are solid red dots and solid black lines, respectively. Results for **a**  $H_2$  ( $N=2$ ,  $M=2$ ), **b**  $LiH$  ( $N=4$ ,  $M=4$ ), and **c** linear  $H_4$  ( $N=8$ ,  $M=24$ ).

where  $\sum_{\vec{s}} = \sum_{s_1} \cdots \sum_{s_M}$  and it is implicitly assumed that  $\langle \Psi_v(\theta) | \Psi_v(\theta) \rangle = 1$ . Instead of performing the exact summation over all  $\vec{s}$  in Eq. (7) (this is essentially the same summation in Eq. (5) mentioned at the end of the last paragraph), we should estimate this sum with the Monte Carlo method. We note that  $N_{\vec{s}}^2$  in Eq. (7) is the probability of observing  $\{s_1, \dots, s_M\}$  upon measuring those  $M$  hidden spins during the construction of the state  $|\Psi_v^{\vec{s}}(\theta)\rangle$ . Hence, samples of  $\vec{s}$  are effectively drawn from the probability density  $N_{\vec{s}}^2$  during the construction of  $|\Psi_v^{\vec{s}}(\theta)\rangle$ . In short, we replace the exact summation according to

$$\sum_{\vec{s}} N_{\vec{s}}^2 f(\vec{s}) \xrightarrow[\text{according to } N_{\vec{s}}^2]{\text{Monte Carlo sampling of } \vec{s}} \frac{1}{N_{\text{exp}}} \sum_{k=1}^{N_{\text{exp}}} f(\vec{s}_k), \quad (8)$$

where  $f(\vec{s}) = (\prod_{j=1}^M R_{s_j}^2(m_j^R)) |\langle \mathbf{z} | \Psi_v^{\vec{s}}(\theta) \rangle|^2$  or  $f(\vec{s}) = (\prod_{j=1}^M R_{s_j}^2(m_j^R))$  for the numerator and the denominator in Eq. (7), respectively, and  $N_{\text{exp}}$  is the number of sampling experiments performed.

The factor  $(\prod_{j=1}^M R_{s_j}^2(m_j^R))$  in Eq. (7) should be calculated classically. This probability  $|\langle \mathbf{z} | \Psi_v^{\vec{s}}(\theta) \rangle|^2$  is again sampled from projective measurements on visible spins. The only thing that is prohibitively expensive to estimate either classically or experimentally is the normalization constant  $N_v$ . This is the reason we introduce the denominator (which is really just 1) in Eq. (7) that carries another  $N_v$  to cancel the one in the numerator. By using Eqs. (6) and (7) together, the challenging post-selection is replaced with a Monte Carlo framework that needs to sample multiple  $|\Psi_v^{\vec{s}}(\theta)\rangle$  according to Eq. (7). It is worth noting that, by using the ensemble state preparation outlined here, only projective measurements in the computational basis are used in estimating the expectation value of an operator (e.g. the system Hamiltonian), whereas projective measurements along all three axes are typically required in the standard VQE. Additional details on the ensemble state preparation method may be found in Supplementary Note III.

### Quantum simulations with NQS-ITE

Next, we discuss the ITDVP to find a ground state of a Hamiltonian  $\hat{H}$ . The idea is to propagate a trial wave function  $|\Psi_v(\theta_\tau)\rangle$  in the imaginary-time domain. If the trial wave function  $|\Psi_v(\theta_0)\rangle$  at time  $\tau=0$  has a non-zero overlap with the ground state  $|\Psi_{gs}\rangle$ , then it should converge to an ansatz closest to  $|\Psi_{gs}\rangle$  when  $\tau \gg 1$ . With a variational ansatz, the time-evolved  $|\Psi_v(\theta_\tau)\rangle$  can be prepared in a quantum circuit if  $\theta_\tau$  is given. In Supplementary Note IV, we summarize the standard ITDVP<sup>34</sup> for parametrized quantum circuits and explain why we need to modify the standard

approach for the NQS ansatz considered in this work. In this section, we present an overview of the NQS-ITE algorithm and details of the derivation may be found in Supplementary Note V.

In short, the equations of motion for  $\theta_\tau$  assume the following form:

$$\dot{\theta}_n = \sum_m A_{nm}^{-1} C_m. \quad (9)$$

The matrix  $A$  and vector  $C$  read,

$$A_{nm} = \text{Re} \left( \langle \hat{O}_n^\dagger \hat{O}_m \rangle_v - \langle \hat{O}_n^\dagger \rangle_v \langle \hat{O}_m \rangle_v \right), \text{ and} \quad (10)$$

$$C_m = \text{Re} \left( \langle \hat{O}_m^\dagger \hat{H} \rangle_v - \langle \hat{O}_m^\dagger \rangle_v \langle \hat{H} \rangle_v \right),$$

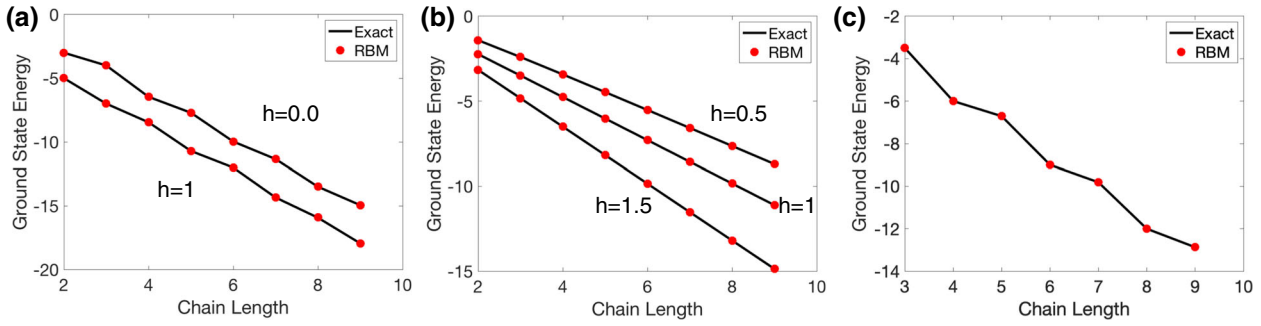
where  $\langle \cdots \rangle_v = \langle \Psi_v(\theta) | \cdots | \Psi_v(\theta) \rangle$ . The  $O_n$  operators are defined as follows,

$$\hat{O}_n = \begin{cases} i^{1-\delta} \hat{V}_i^z, & \text{if } \theta_n = b_i, \\ i^{1-\delta} \tanh(m_j + \sum_i i W_{ij}^I \hat{V}_i^z), & \text{if } \theta_n = m_j, \\ i \hat{V}_i^z \tanh(m_j + \sum_i i W_{ij}^I \hat{V}_i^z), & \text{if } \theta_n = W_{ij}^I, \end{cases} \quad (11)$$

where  $\delta=0$  if  $\theta_n = b_i^I$  or  $\theta_n = m_j^I$  and  $\delta=1$  if  $\theta_n = b_i^R$  or  $\theta_n = m_j^R$ . We note that Eqs. (9)–(11) essentially give the stochastic reconfiguration method for the variational Monte Carlo framework. In the standard ITDVP approach<sup>34</sup>, every matrix element  $A_{mn} = \langle \partial_{\theta_n} \Psi_v(\theta) | \partial_{\theta_m} \Psi_v(\theta) \rangle$  requires measurements with respect to a distinct quantum circuit. Hence, it takes  $\mathcal{O}(N_{\text{var}}^2)$  sets of state preparations to construct  $A$  matrix. By utilizing the commutative property of operators in the RBM Hamiltonian, only one RBM state needs to be prepared for constructing the entire  $A$  matrix. In fact, one can simultaneously estimate all  $\mathcal{O}(N_{\text{var}}^2)$  matrix elements with every given sample of  $\mathbf{z}$  according to the definition of  $A$  matrix element in Eq. (10). This could be a tremendous advantage for large-scale simulations when  $N_{\text{var}} \propto \mathcal{O}(MN)$  could be a huge number.

Next, we point out an interesting observation. The standard gradient-based energy minimization, as done within the VQE approach, updates  $\theta_n$  with the equation of motion  $\dot{\theta}_n = C_n$  where  $C_n$  being the gradient vector for the energy function  $E_\theta = \langle \Psi_v(\theta) | \hat{H} | \Psi_v(\theta) \rangle$ . The comparison of Eq. (9) to the equation of motion above reveals that the NQS-ITE introduces a preconditioner  $A^{-1}$  to the gradient vector in order to adjust the step size to account for the intrinsically curved metric for the NQS manifold. However, the evaluations of the matrix  $A$  requires no more experimental efforts for NQS-ITE than for a standard gradient-descent-based approach as explained. As the imaginary-time method tends to give better result than the gradient descent, one should always adopt the imaginary-time propagation whenever NQS is used as the trial wave function. It is worth noting that a quantum generalization of natural gradient descent has recently been proposed for variational quantum circuits, this algorithm is





**Fig. 3 Numerical results for spin systems.** The ground state energy of spin chains computed with the NQS-ITE algorithm (solid red dots) and exact diagonalization (solid black lines). Results for **a** Heisenberg chain as a function of chain length at different values of  $h$  field, **b** Ising chain as a function of chain length at different values of  $h$  field, **c** J1–J2 chain with  $J_1 = 1$  and  $J_2 = 0.5$ . In panel **a**, the result for  $h = 1$  has been shifted downwards by 2 units for visibility.

also shown to be a better optimizer than the standard gradient-descent approach<sup>87,88</sup>.

### Numerical results

To demonstrate the effectiveness of RBM-NQS ansatz and NQS-ITE algorithm, we report numerical simulations on three different types of systems: molecules, spin chains, and nanostructures (triple quantum dots (TQD)). We first present results with the standard complex-valued NQS ansatz (which requires the extended state preparation protocol described in Supplementary Note VI), then we repeat the simulations for selected examples in the last subsection to demonstrate that the unitary-coupled RBM-NQS could achieve the same level of accuracy with similar number of hidden spins.

**Molecular systems.** We first test the NQS-ITE algorithm on common molecular benchmarks: the dissociation curves of  $H_2$ , LiH, and an  $H_4$  chain. The molecular Hamiltonians are first projected onto a discrete set of molecular orbitals. Here, we use the conventional STO-3G basis set, which constitutes the minimum number of orbitals required to represent a given atomic shell. The resulting fermionic Hamiltonians are subsequently mapped onto qubit Hamiltonians using the Bravyi–Kitaev transformation<sup>89</sup>. The computation of the integrals in second quantization and transformation of the Hamiltonians are done using PySCF<sup>90</sup> and OpenFermion<sup>91</sup>.

Modelling  $H_2$ , LiH, and linear  $H_4$  molecules in the STO-3G basis requires 2, 4, and 8 visible spins, respectively. The ground state energies, estimated from the NQS-ITE algorithm, as a function of inter-nuclear distance are plotted in Fig. 2. Despite using a modest number of hidden spins, the NQS ansatz models the molecular wave functions very well as the NQS-ITE algorithm gives nearly exact numerical results for all cases in Fig. 2. We consider more number of hidden spins for the linear  $H_4$  molecule in order to correctly reproduce the wave function at large bond distance. Our result on linear  $H_4$  molecule is largely consistently with the VMC (variational Monte Carlo) result discussed in ref. <sup>92</sup>, a comprehensive benchmark studies of state-of-the-art simulation methods for quantum chemistry and many-body physics. This consistency can be understood as the RBM-NQS shares many similarities with the Jastrow factor in the variational ansatz adopted for the VMC method in that benchmark test.

**Spin systems.** Next, we consider the problem of finding the ground state of three prototypical spin models, i.e. the transverse-field Ising model, the antiferromagnetic Heisenberg (AFH) model and the J1–J2 model. AFH and J1–J2 models are prominent models to study frustrated systems in condensed matter physics.

The spin Hamiltonians can be written as

$$H_{\text{TFI}} = -h \sum_i \hat{\sigma}_i^x - \sum_{\langle ij \rangle} \hat{\sigma}_i^z \hat{\sigma}_j^z; \quad (12)$$

$$H_{\text{AFH}} = -h \sum_i \hat{\sigma}_i^z + \sum_{\langle ij \rangle} \hat{\sigma}_i^x \hat{\sigma}_j^x + \hat{\sigma}_i^y \hat{\sigma}_j^y + \hat{\sigma}_i^z \hat{\sigma}_j^z; \quad (13)$$

$$H_{\text{J1J2}} = J_1 \sum_{\langle ij \rangle} \hat{\sigma}_i^x \hat{\sigma}_j^x + \hat{\sigma}_i^y \hat{\sigma}_j^y + \hat{\sigma}_i^z \hat{\sigma}_j^z + J_2 \sum_{\langle\langle ij \rangle\rangle} \hat{\sigma}_i^x \hat{\sigma}_j^x + \hat{\sigma}_i^y \hat{\sigma}_j^y + \hat{\sigma}_i^z \hat{\sigma}_j^z, \quad (14)$$

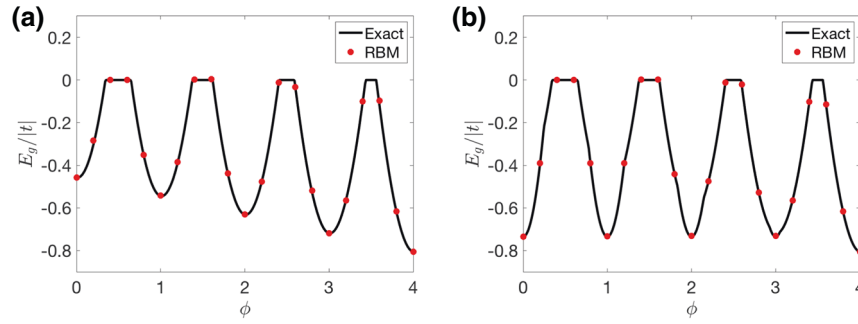
where  $h$  is the uniform field strength,  $\langle ij \rangle$  denotes the nearest-neighbor couplings, and  $\langle\langle ij \rangle\rangle$  denotes the next nearest-neighbor couplings. Open boundary condition is assumed for these linear spin chains. In Fig. 3 we again compute the energies using the NQS-ITE algorithm. Excellent agreement between NQS-ITE and the numerically exact results are obtained as expected.  $M = N$  number of hidden spins are used in all three cases including frustrated AFH and J1–J2 chains.

**Triple quantum dots (TQD).** A lateral TQD is an artificial, fully tunable molecule constructed using metallic gates localizing electrons in a semiconductor field-effect transistor. A TQD allows one to study new phenomena not present in a single or double quantum dot, e.g. topological effects<sup>93,94</sup>. The Hamiltonian of the TQD subject to a uniform perpendicular magnetic field,  $\mathbf{B} = B\hat{z}$ , is given by

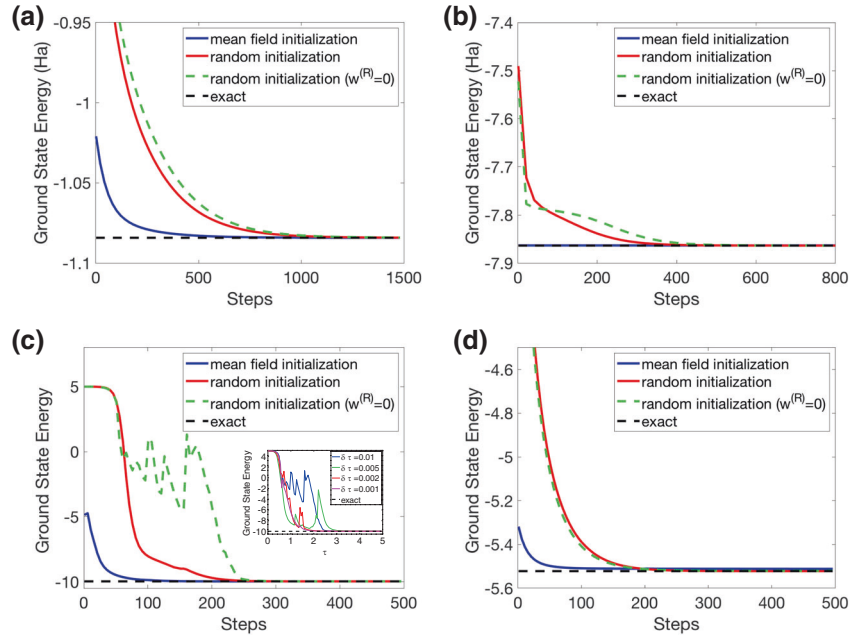
$$H = \sum_{i,\sigma} E_{i\sigma} \hat{d}_{i\sigma}^\dagger \hat{d}_{i\sigma} + \sum_{\sigma,i,j; i \neq j} \tilde{t}_{ij} \hat{d}_{i\sigma}^\dagger \hat{d}_{j\sigma} + \sum_i U_i \hat{n}_{i\downarrow} \hat{n}_{i\uparrow} + \frac{1}{2} \sum_{i,j; i \neq j} V_{ij} Q_i Q_j, \quad (15)$$

where  $\hat{d}_{i\sigma}$  ( $\hat{d}_{i\sigma}^\dagger$ ) is fermionic annihilation (creation) operator with spin  $\sigma = \pm 1/2$  on orbital  $i$ .  $\hat{n}_{i\sigma} = \hat{d}_{i\sigma}^\dagger \hat{d}_{i\sigma}$  and  $\hat{Q}_i = \hat{n}_{i\downarrow} + \hat{n}_{i\uparrow}$  are the spin and charge density on orbital level  $i$ .  $U_i$  and  $V_{ij}$  gives the strength of on-site and off-site Coulomb repulsion, respectively. Each dot is represented by a single orbital with energy  $E_{i\sigma} = E_i + g^* \mu_B B \sigma$ , where  $g^*$  is the effective Landé  $g$ -factor,  $\mu_B$  is the Bohr magneton. The dots are connected by magnetic-field-dependent hopping matrix elements  $\tilde{t}_{ij} = t_{ij} e^{2\pi i \phi_{ij}}$  where  $\phi_{ij} = \phi$  is the magnetic flux. The details of this model can be found in the “Methods” section.

The ground state energy of the TQD as a function of magnetic field is plotted in Fig. 4, we observe excellent agreement between the results from the exact diagonalization and NQS-ITE. It is worth noting that, at non-zero magnetic field, the ground state wave function of a TQD could be complex, thus a complex RBM ansatz is necessary for accurate representation of the ground state wave



**Fig. 4 Numerical results for triple quantum dots.** The ground state energy of a lateral TQD ( $N=M=6$ ) as a function of magnetic field obtained from the NQS-ITE algorithm (solid red dots) and the exact diagonalization (solid black lines). Panel **a**  $U_i = 50|t|$ ,  $V_{ij} = 10|t|$  and  $E_i = t$ , and **b**  $U_i = 0.5|t|$ ,  $V_{ij} = 0.1|t|$ , and  $E_i = t$ .



**Fig. 5 Imaginary-time evolution with RBM states.** Simulations of the ground state energy using NQS-ITE with random initialization (solid red lines), mean-field initialization (solid blue lines), and random initialization for the unitary-coupled RBM ansatz (dashed green lines). The dashed black lines show the ground state energy from exact diagonalization. Results for **a**  $H_2$  at equilibrium bond length (1.05 Å), **b**  $LiH$  at equilibrium bond length (1.5 Å), **c** Heisenberg chain ( $N=6$ ,  $h=0$ ), and **d** Ising chain ( $N=6$ ,  $h=0.5$ ). The inset in **c** displays the results for the unitary-coupled RBM ansatz using different time-step sizes,  $\delta\tau$ , where  $\tau = \text{steps} \cdot \delta\tau$ .

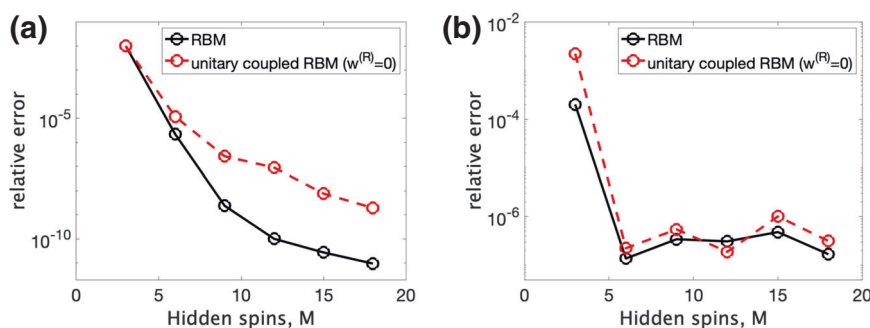
function. Again, RBM-NQS may model this complex system with a modest number of hidden spins with  $M=N$  even in the non-perturbative regime (in which Coulomb repulsion and kinetic energy are of comparable strength) as displayed in Fig. 4b.

**Results with unitary-coupled RBM-NQS.** At last, we repeat calculations of a few models studied in previous sections with the unitary-coupled RBM-NQS ansatz. The key insight revealed in Figs. 5 and 6 is that, in most cases, the unitary-coupled RBM-NQS delivers comparable performance to that of the standard complex RBM-NQS. Although, it is also suggestive that the unitary-coupled RBM simulation could be improved if one either uses a smaller time-step size for the NQS-ITE algorithm or uses more number of hidden spins. The first point is well illustrated by Fig. 5c and the accompanied inset. The non-monotonic trend in the average energy (during the imaginary-time evolution) is a signature that the propagated quantum state is not well approximated by the unitary-coupled RBM ansatz when the imaginary-time-step size is large. This energy fluctuation improves consistently when the

time-step size decreases as shown in the inset of Fig. 5c. Note that the NQS-ITE eventually finds the true ground state wave function even when the time-step size is not taken to be smaller.

The second point is illustrated by comparing the representation accuracy between a standard RBM-NQS and a unitary-coupled RBM-NQS with the same number of hidden spins. Clearly, the standard RBM ansatz achieves better accuracy. Although higher accuracy may be attained when more number of hidden spins are used for a unitary-coupled RBM ansatz. In fact, one can see that both ansatzes achieve roughly the same relative error when there are roughly same number of variational parameters for the AFH model in Fig. 6a. Note a standard RBM has almost twice the number of variational parameters (both real and imaginary coupling parameters for each visible–hidden connections).

Finally, we also present another set of results (solid red lines) in Fig. 5. These results are obtained with RBM ansatz whose parameters are initialized with a mean-field solution to the original problems. As shown, a better initial guess improves the convergence rate in comparison to randomly initialized cases.



**Fig. 6** Comparison between full RBM and unitary-coupled RBM states. Comparing relative errors between RBM-NQS and unitary-coupled RBM-NQS for **a** AFH model with  $h = 0$  and  $M = N = 6$ , and **b** TFI model with  $h = 0.5$  and  $M = N = 6$ . The definition of the relative error is  $|E_{\text{exact}} - E_{\text{rbm}}|/|E_{\text{exact}}|$  where rbm refers to either RBM-NQS or unitary-coupled RBM-NQS.

Particularly, mean-field initial wave function already provides nearly exact ground state wave function for lithium hydride molecule, as it can be seen from Fig. 5b that the ground state energy from the mean-field approximation is nearly identical to the exact ground state energy. The accuracy of the mean-field approximation for LiH molecule is possibly due to the low level of electron correlations among the orbitals in minimal basis set. To test this hypothesis, we solve for the ground state of LiH using the Hartree–Fock method and found that it already provides a very accurate ground state solution. In the “Methods” section, we briefly comment how one can systematically make an intelligent guess on a good initial state under various situations.

## DISCUSSION

In conclusion, we present a practical approach to exploit a popular machine-learning model for quantum simulations on a digital quantum computer. Before fault-tolerant quantum computation becomes readily available, the HQC algorithms will prevail as a popular approach for investigating novel applications of a quantum computer. Successful experimental demonstrations of HQC algorithms have certainly attracted attentions and boosted confidence in quantum computations. Nevertheless, many obstacles still prevent a clear demonstration of unambiguous quantum advantages for these HQC algorithms. One possible path towards this goal is to investigate more powerful wave-function ansatz that can achieve a good tradeoff between expressive power and number of variational parameters. With fewer parameters, potentially, one may deal with a shorter-depth state preparation and deals with a simpler optimization problem.

From this perspective, the NQS certainly seems to be a promising option to investigate. For instance it has been shown that an RBM state could be mapped to the powerful matrix product states commonly used in condensed matter physics, but uses variational parameters more economically<sup>77</sup>. Additionally, it is also known that a fully-connected RBM ansatz satisfies an entanglement volume scaling. Thus, it is very intriguing to further investigate whether one can exploit these properties to minimize number of variational parameters under a realistic setting. While the long-range connectivity between qubits is not necessarily easy to realize in many quantum hardware at the moment; it is at least experimentally feasible with one of the leading hardware architectures, the ion-trap-based quantum computers<sup>95,96</sup> having all-to-all connectivity among qubits. Theoretically, one may also design quantum algorithms based on the deep Boltzmann machines<sup>97</sup> that further elevates the expressive power of Boltzmann machine architectures with only short-range couplings suitable for quantum hardware featuring local connectivity among qubits.

In this work, we set out to improve the existing NQS state-preparation protocol in the quantum circuits. As mentioned in the

introduction, prior quantum algorithms using RBM ansatz suffer from several obstacles that prevents simulations for complex systems with many degrees of freedom. Our proposed state-preparation protocols have significantly expanded the scope of applications for NQS as an ansatz for quantum simulations and lowered the experimental barriers. By including complex-valued parameters into the RBM-NQS, the proposed quantum algorithm is capable of simulating some important quantum materials and quantum dynamics<sup>98</sup>. Our numerical testings manifest encouraging signs that the NQS ansatz performs remarkably well across a variety of systems of practical and theoretical interests. Due to the qubit-recycling scheme, we reduce the number of required qubits from  $\mathcal{O}(N + M)$  in previous works down to  $\mathcal{O}(N)$  with sequential implementations of visible–hidden layer interactions. By imposing  $W_{ij} = iW_{ji}^*$ , we avoid the probabilistic preparation of inter-layer couplings, further improving the practicality of NQS-based simulations. The ensemble state preparation bypasses the post-selection on hidden spins.

Additionally, the circuit depth of preparing a unitary-coupled RBM state scales as  $\mathcal{O}(NM)$  whereas the implementation for a standard RBM state either requires exponentially deep circuit or probabilistic post-selection of ancilla qubits. Finally, it has been previously shown that imaginary-time algorithm offers superior performance compared to VQE<sup>34</sup>, but at the expense of more state preparations and measurements. In this work, we exploit the properties of RBM architecture and show that the number of different quantum states required in the imaginary-time algorithm could be reduced from  $\mathcal{O}(N_{\text{var}}^2) \sim \mathcal{O}((NM)^2)$  down to just one circuit like the standard gradient descent for VQE. However, unlike standard VQE cases, the Monte Carlo approach for observable estimation allows us to perform measurement in one fixed basis for the Pauli operators. Since the variational Monte Carlo requires efficient estimation of ratios of wave function amplitudes which might not be possible for arbitrary wave-function ansatz such as the hardware efficient ones. Hence, the experimental costs for NQS-ITE is comparable to what standard VQE with UCC or hardware efficient ansatz demands if not significantly cheaper. Since Boltzmann machine is a widely used machine-learning model with many applications, we expect the HQC paradigm building on the Monte Carlo framework and the NQS ansatz in this work can be adopted for solving other important problems, such as discrete optimizations and machine learning.

## METHODS

### RBM as trial wave function

Recently, Troyer and Carleo<sup>62</sup> used the RBM neural-network architecture as wave-function ansatz to model many-body physics and attained impressive results. Since then, several other wave-function ansatz inspired by neural-network architecture have been explored. Collectively, we now refer to this set of wave-function ansatzs as the NQS. Although, in this work, we

only consider RBM-NQS, which is particularly convenient to model quantum systems composed of two-level systems (TLSs) such as spin lattice commonly studied in condensed matter physics and quantum chemistry problems formulated in terms of qubits. For these systems, each TLS is directly identified with a visible spin in the corresponding RBM model. The entanglement between these TLSs (or visible spins) is mediated by the pairwise interactions between visible and hidden spins. In short, a many-body wave function in the NQS form reads

$$|\Psi_v(\theta)\rangle = \frac{1}{N_v} \sum_{\mathbf{v}} \left( \sum_{\mathbf{h}} e^{E_\theta(\mathbf{v}, \mathbf{h})} \right) |\mathbf{v}\rangle, \quad (16)$$

with energy function  $E_\theta(\mathbf{v}, \mathbf{h}) = \sum_i b_i v_i + \sum_j m_j h_j + \sum_{ij} W_{ij} v_i h_j$ , its complex conjugate  $\bar{E}_\theta(\mathbf{v}, \mathbf{h})$ , and the normalization constant

$N_v = \sqrt{\sum_{\mathbf{v}} \left( \sum_{\mathbf{h}} e^{\bar{E}_\theta(\mathbf{v}, \mathbf{h})} \right) \left( \sum_{\mathbf{h}} e^{E_\theta(\mathbf{v}, \mathbf{h})} \right)}$ . The RBM parameters are collectively denoted by  $\theta = \{\mathbf{b}, \mathbf{m}, \mathbf{W}\}$ . As shown in Eq. (16), the hidden spins are summed over in the bracket on the right-hand side of Eq. (16) to give a wave function  $|\Psi_v(\theta)\rangle$  for the visible spins, which represent the physical system of interest. In principle,  $\theta$  should possess non-vanishing imaginary components to describe complex-valued wave function.

To prepare an NQS in a quantum circuit, we should take the energy function  $E_\theta(\mathbf{v}, \mathbf{h})$  and promote it to an Hermitian operator by replacing the binary values of  $\mathbf{v}$  and  $\mathbf{h}$  with the corresponding Pauli operators. The quantum-circuit analog of Eq. (16) is decomposed into Eqs. (1) and (2): first entangle the visible and hidden spins and then post-selects the hidden spins to mediate the desired non-unitary transformation on the visible-spin wave functions.

### Initial state preparation for quantum simulations

The NQS ansatz can be used in conjunction with most HQC simulation algorithms in addition to the time-dependent variational method outlined above. All these methods aim to solve highly non-trivial optimization in which the quality of solutions or the convergence rate depends crucially on the overlap of the initial state with the ground state  $|\Psi_{gs}\rangle$ . The two-stage initialization protocols described here gives a systematic approach to guide the preparation of high-quality initial states. In short, the idea is to selectively optimize a subset of parameters to obtain an approximate solution that could be used as the initial state in a subsequent simulation optimizing over all parameters.

In the simplest case, one may consider a mean-field approximation, which restricts the considerations to completely factorized product-state wave function  $|\Psi_v^0(\vec{b})\rangle = |\psi_{v_1}(b_1)\rangle \otimes \cdots \otimes |\psi_{v_N}(b_N)\rangle$  for all visible spins in a quantum simulation. We then subsequently use  $|\Psi_v^0(\vec{b})\rangle$  as the starting point of another simulation in which the hidden spins are introduced along with corresponding parameters  $\{m_1, \dots, m_M, W_{11}, \dots, W_{NM}\}$  that collectively facilitate the formation of entangled NQS,  $|\Psi_v(\theta)\rangle$ . We note the mean-field approximation (single-body physics problem) can be easily done on a classical computer.

Nevertheless, for strongly correlated systems, the product states are not guaranteed to support a high overlap with  $|\Psi_{gs}\rangle$ . Instead of optimizing  $\vec{b}$  (the mean-field approximation) in the first run, it will be beneficial to consider an NQS with specifically designed sparse connectivity. In the second-stage calculation, the fully-connected architecture will be restored as usual, and the total number of variational parameters scale as  $\mathcal{O}(NM)$ . An obvious question is how to decide the connectivity of this sparsely connected RBM architecture for the first-stage simulation, which needs to balance the expressive power of the variational ansatz and the complexity of the optimization tasks. For lattice systems, one may consider short-range RBMs that constitute a special class of the well-established entangled-plaquette states<sup>72</sup>. In this case, the total number of variational parameters scale as  $\mathcal{O}(N)$ .

### Simulation details for numerical studies

In all our simulations, we use a constant learning rate of 0.01. The variational parameters are initialized as Gaussian random numbers with mean zero and variance 0.01, except in cases where the initial conditions are obtained from mean-field solutions (Fig. 5, green dashed lines).

**Hamiltonians of  $H_2$ , LiH and  $H_4$ .** We treat the molecules in the minimal STO-3G basis and use PySCF to compute the integrals in the second quantization. The resulting fermionic Hamiltonians are subsequently

mapped onto qubit Hamiltonians using Bravyi–Kitaev transformation with OpenFermion<sup>91</sup>. Due to the symmetry in  $H_2$ , the final Hamiltonian consists of 2 qubits<sup>99</sup>, whereas the LiH and  $H_4$  Hamiltonians contain 4 and 8 qubits, respectively.

**Hamiltonian for TQD.** For TQD simulations, we use the parameters from ref. <sup>94</sup>, i.e.  $t = -0.23$  meV,  $g^x = -0.44$ ,  $E_i = -|t|$ , and  $\phi/B = 1.257^{-1}$ . We consider two cases in Fig. 4: (a)  $U_i = 50|t|$ ,  $V_{ij} = 10|t|$  (strongly localized case), and (b)  $U_i = 0.5|t|$ ,  $V_{ij} = 0.1|t|$  (non-perturbative regime). We use Bravyi–Kitaev transformation to map the Hubbard Hamiltonian onto a qubit Hamiltonian using OpenFermion.

### DATA AVAILABILITY

The numerical data presented in this study are available from the corresponding author upon reasonable request.

### CODE AVAILABILITY

The codes used to produce the data presented in this study is available from the corresponding author upon reasonable request.

Received: 3 April 2020; Accepted: 20 November 2020;

Published online: 29 January 2021

### REFERENCES

- McClean, J. R., Romero, J., Babbush, R. & Aspuru-Guzik, A. The theory of variational hybrid quantum-classical algorithms. *N. J. Phys.* **18**, 023023 (2016).
- Preskill, J. Quantum computing in the NISQ era and beyond. *Quantum* **2**, 79 (2018).
- Peruzzo, A. et al. A variational eigenvalue solver on a photonic quantum processor. *Nat. Commun.* **5**, 4213 (2014).
- Farhi, E., Goldstone, J. & Gutmann, S. A quantum approximate optimization algorithm. Preprint at <https://arxiv.org/abs/1411.4028> (2014).
- Kandala, A. et al. Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets. *Nature* **549**, 242 (2017).
- Hempel, C. et al. Quantum chemistry calculations on a trapped-ion quantum simulator. *Phys. Rev. X* **8**, 031022 (2018).
- Colless, J. I. et al. Computation of molecular spectra on a quantum processor with an error-resilient algorithm. *Phys. Rev. X* **8**, 011021 (2018).
- Sagstizabal, R. et al. Experimental error mitigation via symmetry verification in a variational quantum eigensolver. *Phys. Rev. A* **100**, 010302 (2019).
- Aspuru-Guzik, A., Dutoi, A. D., Love, P. J. & Head-Gordon, M. Simulated quantum computation of molecular energies. *Science* **309**, 1704–1707 (2005).
- Li, Y., Hu, J., Zhang, X.-M., Song, Z. & Yung, M.-H. Variational quantum simulation for quantum chemistry. *Adv. Theor. Simul.* **2**, 1800182 (2019).
- Cao, Y. et al. Quantum chemistry in the age of quantum computing. *Chem. Rev.* **119**, 10856–10915 (2019).
- McArdle, S., Endo, S., Aspuru-Guzik, A., Benjamin, S. C. & Yuan, X. Quantum computational chemistry. *Rev. Mod. Phys.* **92**, 015003 (2020).
- Childs, A. M., Maslov, D., Nam, Y., Ross, N. J. & Su, Y. Toward the first quantum simulation with quantum speedup. *Proc. Natl Acad. Sci. USA* **115**, 9456–9461 (2018).
- Wecker, D., Hastings, M. B. & Troyer, M. Progress towards practical quantum variational algorithms. *Phys. Rev. A* **92**, 042303 (2015).
- Kühn, M., Zanker, S., Deglmann, P., Marthaler, M. & Weiß, H. Accuracy and resource estimations for quantum chemistry on a near-term quantum computer. *J. Chem. Theory Comput.* **15**, 4764–4780 (2019).
- Li, Z., Li, J., Dattani, N. S., Umrigar, C. & Chan, G. K.-L. The electronic complexity of the ground-state of the fmo cofactor of nitrogenase as relevant to quantum simulations. *J. Chem. Phys.* **150**, 024302 (2019).
- Reiher, M., Wiebe, N., Svore, K. M., Wecker, D. & Troyer, M. Elucidating reaction mechanisms on quantum computers. *Proc. Natl Acad. Sci. USA* **114**, 7555–7560 (2017).
- Kivlichan, I. D. et al. Quantum simulation of electronic structure with linear depth and connectivity. *Phys. Rev. Lett.* **120**, 110501 (2018).
- Babbush, R. et al. Low-depth quantum simulation of materials. *Phys. Rev. X* **8**, 011044 (2018).
- Ryabinkin, I. G., Yen, T.-C., Genin, S. N. & Izmaylov, A. F. Qubit coupled cluster method: a systematic approach to quantum chemistry on a quantum computer. *J. Chem. Theory Comput.* **14**, 6317–6326 (2018).



21. Ryabinkin, I. G., Lang, R. A., Genin, S. N. & Izmaylov, A. F. Iterative qubit coupled cluster approach with efficient screening of generators. *J. Chem. Theory Comput.* **16**, 1055 (2020).
22. Dallaire-Demers, P.-L., Fontalvo, J. R., Veis, L., Sim, S. & Aspuru-Guzik, A. Low-depth circuit ansatz for preparing correlated fermionic states on a quantum computer. *Quantum Sci. Technol.* **4**, 045005 (2019).
23. Romero, J. et al. Strategies for quantum computing molecular energies using the unitary coupled cluster ansatz. *Quantum Sci. Technol.* **4**, 014008 (2018).
24. O'Brien, T. E. et al. Calculating energy derivatives for quantum chemistry on a quantum computer. *npj Quantum Inf.* **5**, 113 (2019).
25. Benfenati, F. et al. Extended wavefunctions for the variational quantum eigensolver. In *Quantum Information and Measurement*, paper F5A-36 (Optical Society of America, 2019).
26. Zhao, A. et al. Measurement reduction in variational quantum algorithms. *Phys. Rev. A* **101**, 062322 (2020).
27. Rubin, N. C., Babbush, R. & McClean, J. Application of fermionic marginal constraints to hybrid quantum algorithms. *N. J. Phys.* **20**, 053020 (2018).
28. Crawford, O. et al. Efficient quantum measurement of pauli operators. Preprint at <https://arxiv.org/abs/1908.06942> (2019).
29. Izmaylov, A. F., Yen, T.-C. & Ryabinkin, I. G. Revising the measurement process in the variational quantum eigensolver: is it possible to reduce the number of separately measured operators? *Chem. Sci.* **10**, 3746–3755 (2019).
30. Izmaylov, A. F., Yen, T.-C., Lang, R. A. & Verteletskyi, V. Unitary partitioning approach to the measurement problem in the variational quantum eigensolver method. *J. Chem. Theory Comput.* **16**, 190 (2020).
31. Verteletskyi, V., Yen, T.-C. & Izmaylov, A. F. Measurement optimization in the variational quantum eigensolver using a minimum clique cover. *J. Chem. Phys.* **152**, 124114 (2020).
32. Huggins, W. J. et al. Efficient and noise resilient measurements for quantum chemistry on near-term quantum computers. Preprint at <https://arxiv.org/abs/1907.13117> (2019).
33. Mitarai, K. & Fujii, K. Methodology for replacing indirect measurements with direct measurements. *Phys. Rev. Res.* **1**, 013006 (2019).
34. McArdle, S. et al. Variational ansatz-based quantum simulation of imaginary time evolution. *npj Quantum Inf.* **5**, 75 (2019).
35. Yang, Z.-C., Rahmani, A., Shabani, A., Neven, H. & Chamon, C. Optimizing variational quantum algorithms using pontryagin's minimum principle. *Phys. Rev. X* **7**, 021027 (2017).
36. Zhu, D. et al. Training of quantum circuits on a hybrid quantum computer. *Sci. Adv.* **5**, eaaw9918 (2019).
37. Shaydulin, R., Safran, I. & Larson, J. Multistart methods for quantum approximate optimization. *IEEE High Performance Extreme Computing Conference (HPEC)*, Walham, MA, USA, 1–8 (2019).
38. Guerreschi, G. G. & Smelyanskiy, M. Practical optimization for hybrid quantum-classical algorithms. Preprint at <https://arxiv.org/abs/1701.01450> (2017).
39. Nakanishi, K. M., Fujii, K. & Todo, S. Sequential minimal optimization for quantum-classical hybrid algorithms. *Phys. Rev. Res.* **2**, 043158 (2020).
40. Parrish, R. M., Iosue, J. T., Ozaeta, A. & McMahon, P. L. A Jacobi diagonalization and Anderson acceleration algorithm for variational quantum algorithm parameter optimization. Preprint at <https://arxiv.org/abs/1904.03206> (2019).
41. Parrish, R. M., Hohenstein, E. G., McMahon, P. L. & Martinez, T. J. Hybrid quantum/classical derivative theory: analytical gradients and excited-state dynamics for the multistate contracted variational quantum eigensolver. Preprint at <https://arxiv.org/abs/1906.08728> (2019).
42. Schuld, M., Bergholm, V., Gogolin, C., Izaac, J. & Killoran, N. Evaluating analytic gradients on quantum hardware. *Phys. Rev. A* **99**, 032331 (2019).
43. Moseley, B., Osborne, M. & Benjamin, S. Bayesian optimisation for variational quantum eigensolvers. *Quantum* **3**, 4 (2018).
44. Sarma, S. D., Deng, D.-L. & Duan, L.-M. Machine learning meets quantum physics. *Phys. Today* **72**, 48 (2019).
45. Melko, R. G., Carleo, G., Carrasquilla, J. & Cirac, J. I. Restricted Boltzmann machines in quantum physics. *Nat. Phys.* **15**, 887 (2019).
46. Jia, Z.-A. et al. Quantum neural network states: a brief review of methods and applications. *Adv. Quantum Technol.* **2**, 1800077 (2019).
47. Torlai, G. & Melko, R. G. Learning thermodynamics with boltzmann machines. *Phys. Rev. B* **94**, 165134 (2016).
48. Carrasquilla, J. & Melko, R. G. Machine learning phases of matter. *Nat. Phys.* **13**, 431 (2017).
49. Kaubregger, R., Pastori, L. & Budich, J. C. Chiral topological phases from artificial neural networks. *Phys. Rev. B* **97**, 195136 (2018).
50. Koch-Janusz, M. & Ringel, Z. Mutual information, neural networks and the renormalization group. *Nat. Phys.* **14**, 578 (2018).
51. Czischek, S., Gärtner, M. & Gasenzer, T. Quenches near ising quantum criticality as a challenge for artificial neural networks. *Phys. Rev. B* **98**, 024311 (2018).
52. Lu, S., Gao, X. & Duan, L.-M. Efficient representation of topologically ordered states with restricted boltzmann machines. *Phys. Rev. B* **99**, 155136 (2019).
53. Xu, Q. & Xu, S. Neural network state estimation for full quantum state tomography. Preprint at <https://arxiv.org/abs/1811.06654> (2018).
54. Torlai, G. et al. Integrating neural networks with a quantum simulator for state reconstruction. *Phys. Rev. Lett.* **123**, 230504 (2019).
55. Huang, L., Yang, Y.-f. & Wang, L. Recommender engine for continuous-time quantum Monte Carlo methods. *Phys. Rev. E* **95**, 031301 (2017).
56. Wang, L. Exploring cluster Monte Carlo updates with Boltzmann machines. *Phys. Rev. E* **96**, 051301 (2017).
57. Inack, E., Santoro, G., Dell'Anna, L. & Pilati, S. Projective quantum monte carlo simulations guided by unrestricted neural network states. *Phys. Rev. B* **98**, 235145 (2018).
58. Torlai, G. & Melko, R. G. Neural decoder for topological codes. *Phys. Rev. Lett.* **119**, 030501 (2017).
59. Bausch, J. & Leditzky, F. Quantum codes from neural networks. *N. J. Phys.* **22**, 023005 (2020).
60. Zhang, Y.-H., Jia, Z.-A., Wu, Y.-C. & Guo, G.-C. An efficient algorithmic way to construct Boltzmann machine representations for arbitrary stabilizer code. Preprint at <https://arxiv.org/abs/1809.08631> (2018).
61. Jia, Z.-A. et al. Efficient machine-learning representations of a surface code with boundaries, defects, domain walls, and twists. *Phys. Rev. A* **99**, 012307 (2019).
62. Carleo, G. & Troyer, M. Solving the quantum many-body problem with artificial neural networks. *Science* **355**, 602–606 (2017).
63. Deng, D.-L., Li, X. & Sarma, S. D. Machine learning topological states. *Phys. Rev. B* **96**, 195145 (2017).
64. Saito, H. Method to solve quantum few-body problems with artificial neural networks. *J. Phys. Soc. Jpn* **87**, 074002 (2018).
65. Hartmann, M. J. & Carleo, G. Neural-network approach to dissipative quantum many-body dynamics. *Phys. Rev. Lett.* **122**, 250502 (2019).
66. Nagy, A. & Savona, V. Variational quantum Monte Carlo method with a neural-network ansatz for open quantum systems. *Phys. Rev. Lett.* **122**, 250501 (2019).
67. Yoshioka, N. & Hamazaki, R. Constructing neural stationary states for open quantum many-body systems. *Phys. Rev. B* **99**, 214306 (2019).
68. Luo, D. & Clark, B. K. Backflow transformations via neural networks for quantum many-body wave functions. *Phys. Rev. Lett.* **122**, 226401 (2019).
69. Torlai, G. & Melko, R. G. Latent space purification via neural density operators. *Phys. Rev. Lett.* **120**, 240503 (2018).
70. 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).
71. 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).
72. 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).
73. Deng, D.-L., Li, X. & Sarma, S. D. Quantum entanglement in neural network states. *Phys. Rev. X* **7**, 021021 (2017).
74. Huang, Y. & Moore, J. E. Neural network representation of tensor network and chiral states. Preprint at <https://arxiv.org/abs/1701.06246> (2017).
75. Gao, X. & Duan, L.-M. Efficient representation of quantum many-body states with deep neural networks. *Nat. Commun.* **8**, 662 (2017).
76. Clark, S. R. Unifying neural-network quantum states and correlator product states via tensor networks. *J. Phys. A Math. Theor.* **51**, 135301 (2018).
77. 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).
78. Choo, K., Mezzacapo, A. & Carleo, G. Fermionic neural-network states for ab-initio electronic structure. *Nat. Commun.* **11**, 2368 (2020).
79. Xia, R. & Kais, S. Quantum machine learning for electronic structure calculations. *Nat. Commun.* **9**, 4195 (2018).
80. Gardas, B., Rams, M. M. & Dziarmaga, J. Quantum neural networks to simulate many-body quantum systems. *Phys. Rev. B* **98**, 184304 (2018).
81. Liu, J.-G., Zhang, Y.-H., Wan, Y. & Wang, L. Variational quantum eigensolver with fewer qubits. *Phys. Rev. Res.* **1**, 023025 (2019).
82. Huggins, W., Patil, P., Mitchell, B., Whaley, K. B. & Stoudenmire, E. M. Towards quantum machine learning with tensor networks. *Quantum Sci. Technol.* **4**, 024001 (2019).
83. Motta, M. et al. Determining eigenstates and thermal states on a quantum computer using quantum imaginary time evolution. *Nat. Phys.* **16**, 205 (2020).
84. Brassard, G., Hoyer, P., Mosca, M. & Tapp, A. Quantum amplitude amplification and estimation. *Contemp. Math.* **305**, 53–74 (2002).
85. Berry, D. W., Childs, A. M., Cleve, R., Kothari, R. & Somma, R. D. Exponential improvement in precision for simulating sparse hamiltonians. In *Proceedings of*

- the Forty-Sixth Annual ACM Symposium on Theory of Computing, STOC'14, 283–292 (ACM, New York, NY, 2014).
86. Childs, A. & Wiebe, N. Hamiltonian simulation using linear combinations of unitary operations. *Quantum Inf. Comput.* **12**, 901 (2012).
  87. Stokes, J., Isaac, J., Killoran, N. & Carleo, G. Quantum natural gradient. *Quantum* **4**, 269 (2020).
  88. Koczor, B. & Benjamin, S. C. Quantum natural gradient generalised to non-unitary circuits. Preprint at <https://arxiv.org/abs/1912.08660> (2019).
  89. Seeley, J. T., Richard, M. J. & Love, P. J. The Bravyi-Kitaev transformation for quantum computation of electronic structure. *J. Chem. Phys.* **137**, 224109 (2012).
  90. Sun, Q. et al. Pyscf: the python-based simulations of chemistry framework. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **8**, e1340 (2018).
  91. McClean, J. R. et al. OpenFermion: the electronic structure package for quantum computers. *Quantum Sci. Technol.* **5**, 034014 (2020).
  92. Motta, M. et al. Towards the solution of the many-electron problem in real materials: equation of state of the hydrogen chain with state-of-the-art many-body methods. *Phys. Rev. X* **7**, 031059 (2017).
  93. Hsieh, C.-Y., Shim, Y.-P., Korkusinski, M. & Hawrylak, P. Physics of lateral triple quantum-dot molecules with controlled electron numbers. *Rep. Prog. Phys.* **75**, 114501 (2012).
  94. Delgado, F. et al. Spin-selective aharonov-bohm oscillations in a lateral triple quantum dot. *Phys. Rev. Lett.* **101**, 226810 (2008).
  95. Brown, K. R., Kim, J. & Monroe, C. Co-designing a scalable quantum computer with trapped atomic ions. *npj Quantum Inf.* **2**, 16034 (2016).
  96. Bruzewicz, C. D., Chiaverini, J., McConnell, R. & Sage, J. M. Trapped-ion quantum computing: progress and challenges. *Appl. Phys. Rev.* **6**, 021314 (2019).
  97. Gao, X. & Duan, L.-M. Efficient representation of quantum many-body states with deep neural networks. *Nat. Commun.* **8**, 662 (2017).
  98. Lee, C.-K., Patil, P., Zhang, S. & Hsieh, C.-Y. A neural-network variational quantum algorithm for many-body dynamics. Preprint at <https://arxiv.org/abs/2008.13329> (2020).
  99. O'Malley, P. J. et al. Scalable quantum simulation of molecular energies. *Phys. Rev. X* **6**, 031007 (2016).

## ACKNOWLEDGEMENTS

C.Y.H., C.K.L., and Q.S. thank G. Carleo for insightful comments on the manuscript.

## AUTHOR CONTRIBUTIONS

C.Y.H. and C.K.L. conceived the project. C.Y.H. designed the state-preparation protocol and performed the theoretical analysis. C.K.L. and Q.S. performed the simulations. All authors participated in analyzing data. C.Y.H. and C.K.L. contributed to the writing.

## COMPETING INTERESTS

The authors declare no competing interests.

## ADDITIONAL INFORMATION

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41534-020-00347-1>.

**Correspondence** and requests for materials should be addressed to C.Y.H. or C.K.L.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021