Machine Learning Phase Transitions with a Quantum Processor

A. V. Uvarov,* A. S. Kardashin, and J. D. Biamonte†
Deep Quantum Laboratory, Skolkovo Institute of Science and Technology, 3 Nobel Street, Moscow 143026, Russia

Machine learning has emerged as a promising approach to unveil properties of many-body systems. Recently proposed as a tool to classify phases of matter, the approach relies on classical simulation methods—such as Monte Carlo—which are known to experience an exponential slowdown when simulating certain quantum systems. To overcome this slowdown while still leveraging machine learning, we propose a variational quantum algorithm which merges quantum simulation and quantum machine learning to classify phases of matter. Our classifier is directly fed labeled states recovered by the variational quantum eigensolver algorithm, thereby avoiding the data reading slowdown experienced in many applications of quantum enhanced machine learning. We propose families of variational ansatz states that are inspired directly by tensor networks. This allows us to use tools from tensor network theory to explain properties of the phase diagrams the presented method recovers. Finally, we propose a nearest-neighbour (checkerboard) quantum neural network. This majority vote quantum classifier is successfully trained to recognize phases of matter with 99% accuracy for the transverse field Ising model and 94% accuracy for the XXZ model. These findings suggest that our merger between quantum simulation and quantum enhanced machine learning offers a fertile ground to develop computational insights into quantum systems.

Introduction. The best contemporary algorithms to emulate quantum systems using classical computers suffer from an exponential slowdown in limiting cases. A recent approach is to apply machine learning, which offers new techniques for large-scale data analysis. In particular, machine learning was recently proposed as a tool to recognize phases of matter [1, 2]. These methods still rely on Monte Carlo sampling (or alternative classical simulation methods) which suffers from the an exponential slowdown induced by the so-called sign problem. Independently, quantum algorithms have also been proposed as a platform for machine learning [3–11]. In addition, unlike classical algorithms, quantum simulators are predicted to simulate quantum systems efficiently [12–14]. Here we merge quantum machine learning with quantum simulation, leveraging quantum mechanics to overcome two classical bottlenecks. Namely, we leverage quantum algorithms as a tool to simulate quantum systems by preparing states which are labeled and fed into a quantum classifier. The later removes the data reading slowdown experienced in many applications of quantum enhanced machine learning. While the former utilizes quantum simulators to avoid classical methods with known limitations. We also replace the standard unitary coupled cluster ansatz found in implementations of the variational quantum eigensolver [15] with families of tensor network ansatz states.

The idea of machine learning is to recognize patterns in data. Using machine learning techniques, one can analyze the phase diagrams of strongly interacting quantum systems and thus directly address system properties. In this approach, Monte Carlo samples of such systems are used as input data and classified using supervised [1, 2] or unsupervised [16] learning. This way, spin Hamiltonians of up to a few hundreds of entities can be studied. Nonetheless, for fermionic systems, the use of Monte Carlo methods is drastically restricted by the sign problem, leading to an exponential slowdown.

In the variational quantum circuits approach [17, 18], the quantum computer is required to prepare a sufficiently rich variety of probe states (or circuits, e. g. if the problem is to approximately compile a certain quantum gate [19]). This approach emerged in response to challenges to adopt quantum algorithms for existing hardware. A particular example, the variational quantum eigensolver (VQE), represents an implementation of variational quantum circuits which uses a quantum processor to prepare a family of states characterized by a polynomial number of parameters and minimizes the expectation value of a given Hamiltonian within this family. This approach is widely taken on small- and middle-size quantum computers [15, 20, 21].

In this Letter, we propose a way around the sign problem using a quantum computer. To classify the phases of a given quantum Hamiltonian, we first prepare its approximate ground states variationally, and then feed them as an input to a quantum classifier. In this respect, there is no need in sampling of microscopic configurations with Monte Carlo based methods. Instead, the classifier has direct access to the quantum states [7], yielding thus an effective realization of quantum-enhanced machine learning. To make this algorithm realizable on near-term quantum computers, we propose preparing the quantum states using shallow tensor network based circuits. Numerical tests show that this technique allows the quantum classifier to correctly recognize phase transition in transverse field spin models.

* alexey.uvarov@skoltech.ru
† quantum.skoltech.ru

Tensor network ansatz states for VQE. VQE is a hybrid iterative quantum-classical algorithm used to approximate the ground state of a given Hamiltonian [15]. It relies on preparing an ansatz state $|\psi(\boldsymbol{\theta})\rangle$ by applying a sequence of quantum gates $U(\boldsymbol{\theta})$ and sampling the expectation value of a given Hamiltonian relative to this state, followed by a classical optimizer to minimize the energy, $\langle\psi(\boldsymbol{\theta})|H|\psi(\boldsymbol{\theta})\rangle$. Within the VQE method, we approach the ground state of a given Hamiltonian using tensor networks ansatz states.

We proceed with representing the Hamiltonian as a sum of tensor products of Pauli operators:

$$H = \sum \mathcal{J}_{\alpha_1\alpha_2\ldots\alpha_n} \sigma_{\alpha_1} \otimes \sigma_{\alpha_2} \otimes \cdots \otimes \sigma_{\alpha_n}, \qquad (1)$$

where $\alpha_i \in \{0,1,2,3\}$ enumerate the Pauli matrices $\{\mathbb{1}, X, Y, Z\}$. With the decomposition (1), individual terms of $\langle\psi(\boldsymbol{\theta})|H|\psi(\boldsymbol{\theta})\rangle$ can be estimated and variationally minimized elementwise using a classical-to-quantum process. In each iteration one prepares the state $|\psi(\boldsymbol{\theta})\rangle$ and measures each qubit in the local X, Y, or Z basis, estimates the energy and updates $\boldsymbol{\theta}$. This method can become scalable only if the number of terms in the Hamiltonian is polynomially bounded in the number of spins and the coefficients $\mathcal{J}_{\alpha_1\alpha_2\ldots\alpha_n}$ are defined up to poly(n) digits.

It is therefore not surprising that the performance of VQE crucially depends on the choice of the ansatz state. A common approach is to use the unitary version of the coupled cluster method, the unitary coupled cluster (UCC) ansatz [20, 22–24]. For interacting spin problems, the (non)unitary coupled cluster ansatz can be composed out of spin-flip operators [17, 25]. There is no known classical algorithm to efficiently implement this method, even when the series is truncated to low-order terms [26]. In principle, a quantum computer could efficiently prepare this state, truncated up to some k-th order using the Suzuki–Trotter decomposition [27]. However, for a system of n qubits it requires $\mathcal{O}(n^k)$ unitary gates, making this technique out of reach for the available quantum computers. Still, even if UCC is truncated to single and double interactions (UCCSD), it requires $9n^2$ operations and necessitates applying some optimization strategy that would remedy this problem [28].

Instead, we test a number of shorter ansatz states inspired by tensor network states, namely (i) a rank-one circuit; (ii) a tree tensor network circuit (Fig. 2a); and (iii) a family of checkerboard-shaped circuits (Fig. 2b) with varying depth.

In particular, these states differ in the amount of entanglement they can support. In general, quantum states of n qubits can be represented by n-index tensors, while quantum circuits are embodied by tensor networks. Each quantum gate is seen as a vertex, and each string is an index running through $\{0,1\}$, while the maximum amount of entanglement is determined by the number of strings one needs to cut to separate the subsystems (see Fig. 1). Each string corresponds to at most one ebit of entanglement. An n-qubit state can contain at most $\lfloor n/2 \rfloor$ ebits of entanglement. To formalize the latter, suppose there exists a certain bipartition in the system that brings n_1 qubits to the first subsystem and n_2 qubits to the second. It is then possible to regroup the tensor state into a $2^{n_1} \times 2^{n_2}$ matrix. The rank of this matrix provides an upper bound to the amount of entanglement across this bipartition: a rank-k state can support at most $\log_2 k$ ebits of entanglement, i.e. when it is in the maximally entangled state.

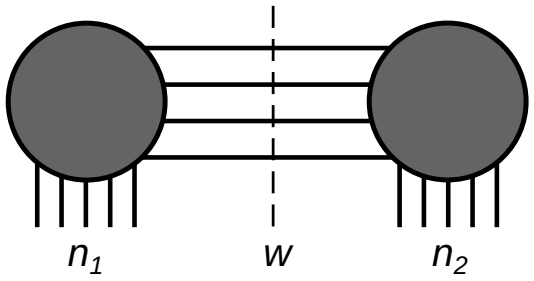

Figure 1. A quantum circuit can be treated as a tensor network state with all bonds having dimension 2. If a bipartition cuts w wires, the total bond dimension is at most 2^w, while the cut separates at most w ebits of entanglement.

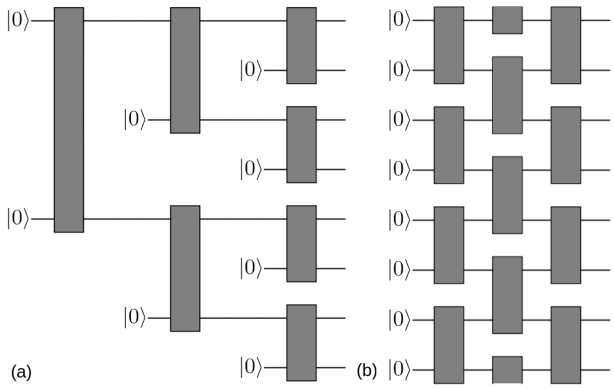

Figure 2. (a) Tree tensor network state. (b) Checkerboard tensor network state. In both cases, the quantum register is instantiated in the $|0\rangle^{\otimes n}$ state and subject to entangling gates. Black boxes indicate two-qubit gates specified by Equation 2 and Figure 3. Each block is parametrized independently.

State preparation. We first approximate the ground state of a Hamiltonian by rank-one states. One can prepare any unentangled state using $2n$ gates by subsequently applying R_y and R_z rotations to each qubit. This ansatz essentially matches the first order truncation of UCC.

Fig. 2a illustrates a circuit implementing a *tree tensor network state*. Such a structure enables long-range correlations but limits the entanglement entropy that a state bipartition can potentially have: any contiguous region can be isolated by $\mathcal{O}(1)$ cuts. It is easy to contract $\langle\psi_{tree}|\, A\, |\psi_{tree}\rangle$ classically with A being a local observable and $|\psi_{tree}\rangle$ being a tree tensor network state. Noteworthy, there is some freedom in choosing the

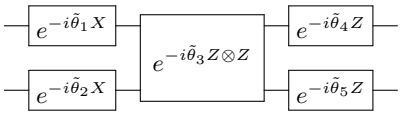

Figure 3. Two-qubit entangler gate used in preparation of the states. The circuit reads left to right.

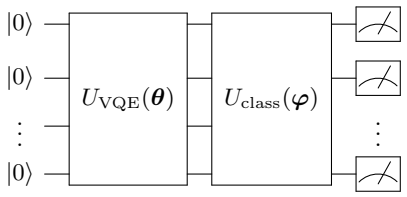

Figure 4. Quantum circuit that implements the classifier. The first part prepares the VQE solution, the second one performs the classification. The assigned label is inferred from the measurements in the Z basis. Both U_{VQE} and U_{class} have the checkerboard structure.

two-qubit blocks that comprise the tree TN. In principle, one can implement any unitary in $SU(4)$ by using three controlled-NOTs and 15 single-qubit rotations [29]. However, throughout this work we used two-qubit gates with fewer parameters $\tilde{\boldsymbol{\theta}}$ to simplify the optimization (Figure 3):

$$U(\tilde{\boldsymbol{\theta}}) = (R_z(\tilde{\theta}_5) \otimes R_z(\tilde{\theta}_4)) \circ \\ \circ R_{zz}(\tilde{\theta}_3) \circ (R_x(\tilde{\theta}_1) \otimes R_x(\tilde{\theta}_2)), \quad (2)$$

where $R_z(\theta) = e^{i\theta Z/2}$, $R_x(\theta) = e^{i\theta X/2}$, and $R_{zz}(\theta) = e^{i\theta Z\otimes Z/2}$. Thus, a complete ansatz would have five free parameters per two-qubit block. Remarkably, the block (2) is inspired by the parametrized Hamiltonian approach [30] and the unitary operators used in QAOA [31]. Of course, the ansatz with such blocks is weaker than it would be if each block implemented the entire $SU(4)$ group. However, such an ansatz would also have some redundancy as the ansatz gates are applied to a fixed n-qubit input state $|0\rangle^{\otimes n}$.

In a *checkerboard tensor network*, the entangling blocks are positioned in a checkerboard pattern as shown in Fig. 2b. In the following, we impose periodic boundary conditions, meaning that the last qubit is linked to the first. For this ansatz, we also use the two-qubit entangling gate shown in Figure 3.

To isolate a region in a checkerboard tensor network state with d layers, one has to cut at least $\mathcal{O}(d)$ bonds regardless of the region size (if the region is very small, one can also make a "horizontal cut" of $\mathcal{O}(L)$ bonds). Therefore, if we set the number of layers to be equal to $\lfloor \log_2(n) \rfloor$, the upper bound on the entanglement scaling is equal to that in critical one-dimensional systems. To implement any maximally entangled state, that is, a state with the maximum possible amount of ebits, one needs to cut $\lfloor \frac{n}{2} \rfloor$ bonds, which gives a lower bound of $2\lfloor \frac{n}{2} \rfloor$ layers for open boundary conditions and $\lfloor \frac{n}{2} \rfloor$ for periodic boundary conditions.

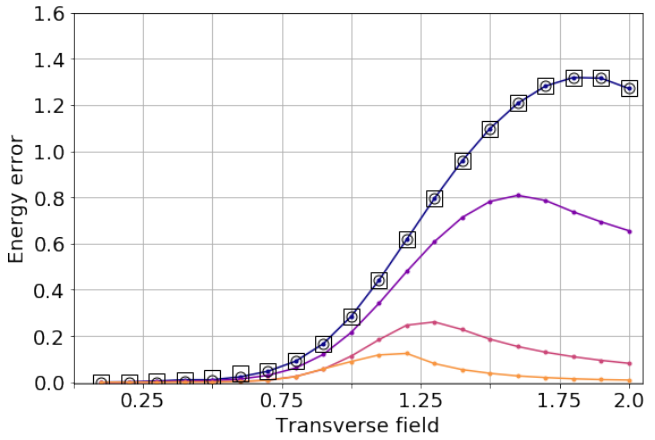

Figure 5. Energy error of VQE solutions for the transverse field Ising model. Hollow squares: rank-1 ansatz, hollow circles: tree tensor network, filled circles: checkerboard states (darkest: 1 layer, brightest: 4 layers).

Quantum classifier. Not only can the checkerboard tensor network be used as a VQE ansatz but it also functions as a quantum majority vote classifier. Each VQE solution is labeled with "0" or "1" depending on whether the magnetic field is above or below the phase transition point. We then prepare a circuit made of two parts: the first part implements the VQE solution $U_{\mathrm{VQE}}(\boldsymbol{\theta})$, the second one is the classifier $U_{\mathrm{class}}(\boldsymbol{\varphi})$ (Figure 4). We measure the output of the circuit in the Z-basis. Let q_0 and q_1 be the total number of measurements in which more than half of the qubits are in "0" or "1" states respectively. Finally, the classifier returns the predicted probability $p = q_1/(q_0 + q_1)$ for the state belonging to the class "1" being equal to the probability that the majority of qubits vote "1", excluding ties.

Let $\{(\boldsymbol{\theta}_i, y_i)\}_{i=1}^{N_{train}}$ be the set of training data points and their labels, $y_i \in \{0, 1\}$. Let $p_i \in [0, 1]$ be the label predicted by the neural network. Then the logarithmic loss function is:

$$L = -\sum_{i=1}^{N_{train}} (y_i \log p_i + (1 - y_i) \log(1 - p_i)) . \quad (3)$$

To minimize L, we used the SPSA algorithm [32]. This algorithm estimates the gradient vector by computing a finite difference in random direction, then performs a gradient descent step. We optimized the log loss over 300 epochs, with both finite differences step size and learning rate starting very coarse and decreasing as $1/\sqrt{n_{epoch}}$, where n_{epoch} is the epoch number.

Numerical results. To compare the performance of various VQE ansatz states, we apply our quantum circuit of $n = 10$ qubits to the transverse field Ising model (TFIM), which being exactly solvable [33, 34] serves for testing purposes. This model is specified by the Hamil-

tonian

$$H_{\mathrm{TFIM}} = J\sum_{i=1}^{n}\sigma_i^z\sigma_{i+1}^z + h\sum_{i=1}^{n}\sigma_i^x,\ J>0,\ h>0, \quad (4)$$

where σ_i^α is the Pauli matrix α acting on the ith spin, and we impose periodic boundary conditions $\sigma_{n+1}^\alpha = \sigma_1^\alpha$.

The ground state of TFIM is determined by the tradeoff between Heisenberg exchange coupling, the first term in (4), favoring collinear orientation of magnetic moments in z direction, and the Zeeman coupling to the transverse magnetic field, the second term. The latter has the tendency to flip the z-component, being thus the source of 'quantum fluctuations' in the system. In a magnetically ordered state, a strong magnetic field $h \geq J$ destroys magnetic order even at zero temperature. This induces quantum fluctuations resulting in ground state restructuring, which is manifested by non-analiticity in the ground state energy of the quantum Hamiltonian (4).

It is therefore intuitively clear that in the absence of magnetic field, $h = 0$, or in the case of high spin polarization, $h = \infty$, the Hamiltonian is dominated by a single interaction, making the ground state disentangled so that even rank-one approximations could provide quantitatively correct results. Meanwhile, this is not the case at criticality, $h = J$, where the competition between two mechanisms in the Hamiltonian (4) favors the formation of highly entangled ground state.

In our numerical implementation, we use Qiskit [35] to simulate quantum circuits and L-BFGS-B method to update the parameters during the classical step of the VQE, while all measurements are assumed ideal. We scan the values of h from 0 to $2J$. For $h = 0$, the optimization process started from a random point, then each additional point started from the previous solution. To eliminate any obviously sub-optimal solutions, we also ran the scanning in the opposite direction and for each value of the field we keep the better result.

All ansatz states show an increase in error near the phase transition point (Fig. 5). With the increasing depth, the checkerboard states show better approximation. It is interesting to point out that the low-depth ansatz states, namely rank-one, tree tensor network and checkerboard states of depth one, all show surprisingly similar results, which is associated with the relative simplicity of the Hamiltonian (4). To feed the dataset into the classifier we prepare 100 data points using VQE with four-layered checkerboard ansatz states. This data was then shuffled and split into the training set (80%) and the test set (20%). Strikingly, the accuracy of the prediction achieves 99%. The outcome of the quantum classifier is presented in Fig. 6 (left).

Another exactly solved model which we use to test our classifier is the antiferromagnetic XXZ spin chain with the Hamiltonian:

$$H = \sum_{i=1}^{n}\left[J_\perp\left(\sigma_i^x\sigma_{i+1}^x + \sigma_i^y\sigma_{i+1}^y\right) + J_z\sigma_i^z\sigma_{i+1}^z\right]. \quad (5)$$

From a physical perspective, (5) corresponds to a uniform exchange coupled system with a uniaxial anisotropy specified by J_z. At $|J_z| < J_\perp$, this model is in the XY, or planar, phase which is characterized by algebraic decay of equal-time spin-spin correlation functions. In the regime $J_z > J_\perp$ the Hamiltonian corresponds to the antiferromagnetic Ising state. The system undergoes a Berezinsky–Kosterlitz–Thouless type phase transition at $J_z = J_\perp$ is [36]. At the phase transition point, the ground state has the highest nearest-neighbour concurrence and a cusp in nearest-neighbour quantum discord [37].

This model has symmetry with respect to rotations in the xy plane, as well as spin-flip symmetry. This fact allows us to augment the training data. Given a VQE approximation, we can create another, equally good approximation by applying a rotation or spin-flip. The structure of the VQE ansatz is conserved in the sense that the new states are produced by the same quantum circuit with different control parameters (see Supplemental Material for more details). In total, we produce 4000 data points.

Despite being more subtle, the phase transition in the XXZ model is also correctly learned by the classifier, yielding correct labels on 94.6 % of test data (Fig. 6, right). In this case, we have added two more layers to the classifier circuit to increase the accuracy.

Conclusions. We proposed a method of classifying phases of matter using a quantum machine learning algorithm paired with quantum simulation. In our numerical tests we achieved 99% accuracy for the transverse field Ising model with a 4-layered classifier and 94% accuracy for the XXZ model with a 6-layered classifier. The quantum classifier works intrinsically with quantum data, providing advantage over the classical methods based on Monte-Carlo sampling.

The proposed classification technique can be applied to any model that can be expressed as a spin model (e.g. fermion problems can be mapped to spin problems by using Jordan–Wigner transformation of Bravyi–Kitaev transformation).

We simulated quantum machine learning for binary classification. Such a protocol can be extended to more classes. For example, for four classes, we could treat "0000000000" as a label of class 1, "0000011111" as a label of class 2, "1111100000" as that of class 3, and finally "1111111111" as that of class 4. Other possibilities would then be labeled according to which of the four strings is closest in terms of Hamming distance.

ACKNOWLEDGMENTS

We thank Dmitry Yudin for fruitful discussions and suggestions. A. U. acknowledges RFBR project number 19-31-90159.

Numerical Methods. The quantum circuits designed in the study were simulated using Qiskit [35]. Figures 5, 6 were prepared using Matplotlib [38]. The data generated and source code developed in this study is open

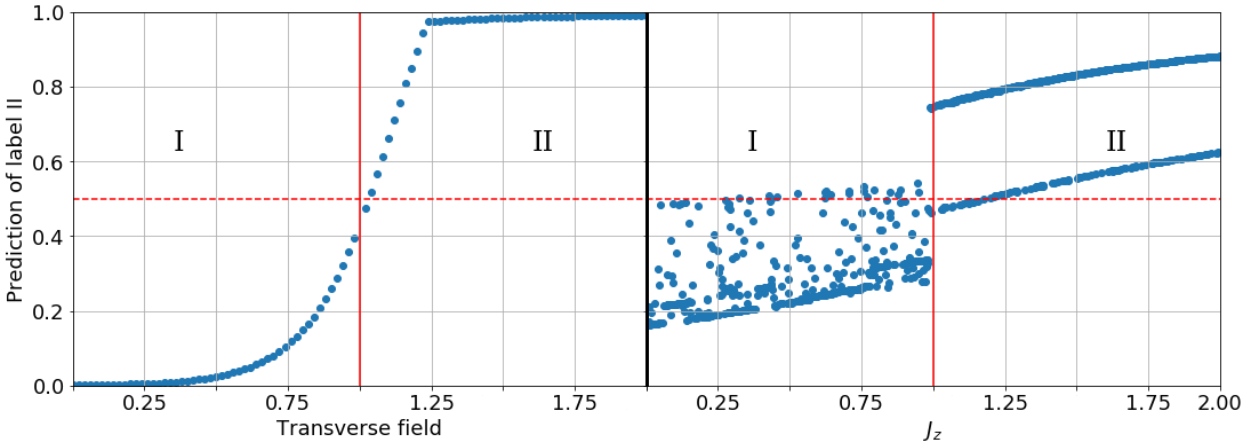

Figure 6. Left: predicted probability of phase II as a function of magnetic field for transverse field Ising model. Right: predicted probability of phase II as a function of J_z for the XXZ model. Roman numbers denote the phases I and II of the models.

source and will be made available online at DeepQuantum's GitHub page after this work is accepted for publication.

[1] J. Carrasquilla and R. G. Melko, Nature Physics **13**, 431 (2017), arXiv: 1605.01735.
[2] E. van Nieuwenburg, Y.-H. Liu, and S. Huber, Nature Physics **13**, 435 (2017).
[3] J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, and S. Lloyd, Nature **549**, 195 (2017), arXiv: 1611.09347.
[4] W. Huggins, P. Patil, B. Mitchell, K. B. Whaley, and E. M. Stoudenmire, Quantum Science and Technology **4**, 024001 (2019).
[5] M. Schuld, A. Bocharov, K. Svore, and N. Wiebe, arXiv:1804.00633 [quant-ph] (2018), arXiv: 1804.00633.
[6] M. Schuld and N. Killoran, Physical Review Letters **122**, 040504 (2019), arXiv: 1803.07128.
[7] V. Havlíček, A. D. Córcoles, K. Temme, A. W. Harrow, A. Kandala, J. M. Chow, and J. M. Gambetta, Nature **567**, 209 (2019).
[8] M. Schuld, I. Sinayskiy, and F. Petruccione, Contemporary Physics **56**, 172 (2015), arXiv: 1409.3097.
[9] B. Duan, J. Yuan, Y. Liu, and D. Li, Physical Review A **96**, 032301 (2017).
[10] B. Duan, J. Yuan, J. Xu, and D. Li, Physical Review A **99**, 032311 (2019).
[11] Y.-B. Sheng and L. Zhou, Science Bulletin **62**, 1025–1029 (2017).
[12] I. Georgescu, S. Ashhab, and F. Nori, Reviews of Modern Physics **86**, 153–185 (2014).
[13] H. Bernien, S. Schwartz, A. Keesling, H. Levine, A. Omran, H. Pichler, S. Choi, A. S. Zibrov, M. Endres, M. Greiner, V. Vuletić, and M. D. Lukin, Nature **551**, 579 (2017).
[14] J. T. Barreiro, M. Müller, P. Schindler, D. Nigg, T. Monz, M. Chwalla, M. Hennrich, C. F. Roos, P. Zoller, and R. Blatt, Nature **470**, 486–491 (2011).
[15] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O'Brien, Nature Communications **5** (2014), 10.1038/ncomms5213.
[16] L. Wang, Physical Review B **94**, 195105 (2016).
[17] J. R. McClean, J. Romero, R. Babbush, and A. Aspuru-Guzik, New Journal of Physics **18**, 023023 (2016), arXiv: 1509.04279.
[18] D. Wecker, M. B. Hastings, and M. Troyer, Physical Review A **92**, 042303 (2015).
[19] S. Khatri, R. LaRose, A. Poremba, L. Cincio, A. T. Sornborger, and P. J. Coles, arXiv:1807.00800 [quant-ph] (2018), arXiv: 1807.00800.
[20] P. J. J. O'Malley, R. Babbush, I. D. Kivlichan, J. Romero, J. R. McClean, R. Barends, J. Kelly, P. Roushan, A. Tranter, N. Ding, B. Campbell, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, A. G. Fowler, E. Jeffrey, A. Megrant, J. Y. Mutus, C. Neill, C. Quintana, D. Sank, A. Vainsencher, J. Wenner, T. C. White, P. V. Coveney, P. J. Love, H. Neven, A. Aspuru-Guzik, and J. M. Martinis, Physical Review X **6** (2016), 10.1103/PhysRevX.6.031007, arXiv: 1512.06860.
[21] J. Colless, V. Ramasesh, D. Dahlen, M. Blok, M. Kimchi-Schwartz, J. McClean, J. Carter, W. de Jong, and I. Siddiqi, Physical Review X **8** (2018), 10.1103/PhysRevX.8.011021.
[22] Y. Shen, X. Zhang, S. Zhang, J.-N. Zhang, M.-H. Yung, and K. Kim, Physical Review A **95** (2017), 10.1103/PhysRevA.95.020501, arXiv: 1506.00443.
[23] C. Hempel, C. Maier, J. Romero, J. McClean, T. Monz, H. Shen, P. Jurcevic, B. P. Lanyon, P. Love, R. Babbush, A. Aspuru-Guzik, R. Blatt, and C. F. Roos, Physical Review X **8** (2018), 10.1103/PhysRevX.8.031022.
[24] E. F. Dumitrescu, A. J. McCaskey, G. Hagen, G. R. Jansen, T. D. Morris, T. Papenbrock, R. C. Pooser, D. J. Dean, and P. Lougovski, Physical Review Letters **120** (2018), 10.1103/PhysRevLett.120.210501, arXiv: 1801.03897.
[25] O. Götze, D. J. J. Farnell, R. F. Bishop, P. H. Y. Li, and

J. Richter, Physical Review B **84** (2011), 10.1103/PhysRevB.84.224428.
[26] A. G. Taube and R. J. Bartlett, International Journal of Quantum Chemistry **106**, 3393 (2006).
[27] B. P. Lanyon, J. D. Whitfield, G. G. Gillet, M. E. Goggin, M. P. Almeida, I. Kassal, J. D. Biamonte, M. Mohseni, B. J. Powell, M. Barbieri, A. Aspuru-Guzik, and A. G. White, Nature Chemistry **2**, 106 (2010), arXiv: 0905.0887.
[28] Y. Cao, J. Romero, J. P. Olson, M. Degroote, P. D. Johnson, M. Kieferová, I. D. Kivlichan, T. Menke, B. Peropadre, N. P. D. Sawaya, S. Sim, L. Veis, and A. Aspuru-Guzik, arXiv:1812.09976 [quant-ph] (2018), arXiv: 1812.09976.
[29] F. Vatan and C. Williams, Physical Review A **69** (2004), 10.1103/PhysRevA.69.032315, arXiv: quant-ph/0308006.
[30] R. Santagati, J. Wang, A. A. Gentile, S. Paesani, N. Wiebe, J. R. McClean, S. Morley-Short, P. J. Shadbolt, D. Bonneau, J. W. Silverstone, D. P. Tew, X. Zhou, J. L. O'Brien, and M. G. Thompson, Science Advances **4**, eaap9646 (2018).
[31] E. Farhi, J. Goldstone, and S. Gutmann, arXiv:1411.4028 [quant-ph] (2014), arXiv: 1411.4028.
[32] J. Spall, IEEE Transactions on Automatic Control **37**, 332–341 (1992).
[33] E. Lieb, T. Schultz, and D. Mattis, Annals of Physics **16**, 407 (1961).
[34] P. Pfeuty, Annals of Physics **57**, 79 (1970).
[35] G. Aleksandrowicz, T. Alexander, P. Barkoutsos, L. Bello, Y. Ben-Haim, D. Bucher, F. J. Cabrera-Hernández, J. Carballo-Franquis, A. Chen, C.-F. Chen, J. M. Chow, A. D. Córcoles-Gonzales, A. J. Cross, A. Cross, J. Cruz-Benito, C. Culver, S. D. L. P. González, E. D. L. Torre, D. Ding, E. Dumitrescu, I. Duran, P. Eendebak, M. Everitt, I. F. Sertage, A. Frisch, A. Fuhrer, J. Gambetta, B. G. Gago, J. Gomez-Mosquera, D. Greenberg, I. Hamamura, V. Havlicek, J. Hellmers, Łukasz Herok, H. Horii, S. Hu, T. Imamichi, T. Itoko, A. Javadi-Abhari, N. Kanazawa, A. Karazeev, K. Krsulich, P. Liu, Y. Luh, Y. Maeng, M. Marques, F. J. Martín-Fernández, D. T. McClure, D. McKay, S. Meesala, A. Mezzacapo, N. Moll, D. M. Rodríguez, G. Nannicini, P. Nation, P. Ollitrault, L. J. O'Riordan, H. Paik, J. Pérez, A. Phan, M. Pistoia, V. Prutyanov, M. Reuter, J. Rice, A. R. Davila, R. H. P. Rudy, M. Ryu, N. Sathaye, C. Schnabel, E. Schoute, K. Setia, Y. Shi, A. Silva, Y. Siraichi, S. Sivarajah, J. A. Smolin, M. Soeken, H. Takahashi, I. Tavernelli, C. Taylor, P. Taylour, K. Trabing, M. Treinish, W. Turner, D. Vogt-Lee, C. Vuillot, J. A. Wildstrom, J. Wilson, E. Winston, C. Wood, S. Wood, S. Wörner, I. Y. Akhalwaya, and C. Zoufal, "Qiskit: An Open-source Framework for Quantum Computing," (2019).
[36] F. Franchini, *An Introduction to Integrable Techniques for One-Dimensional Quantum Systems*, Lecture Notes in Physics, Vol. 940 (Springer International Publishing, Cham, 2017).
[37] R. Dillenschneider, Physical Review B **78**, 224413 (2008).
[38] J. D. Hunter, Computing in Science & Engineering **9**, 90 (2007).

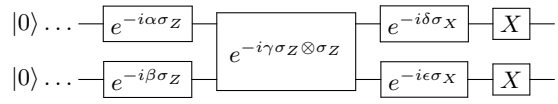

Figure 7. Two-qubit entangler gate with additional X rotations applied.

Appendix A: Supplemental Material

1. Data augmentation for the XXZ model

The XXZ Hamiltonian is symmetric with respect to spin flips and xy plane rotations. Despite the fact that its ground state is non-degenerate, applying these symmetries to a VQE state produces a different state with the same energy, which is an equally valid data point. Thankfully, these actions can be easily performed on the checkerboard states without changing their structure.

Let us start considering rotation symmetry. This rotation is implementing by applying a Z rotation to each qubit:

$$U_{rot} = (e^{i\frac{\varphi}{2}Z})^{\otimes n}. \tag{A1}$$

In the ansatz we deloped, the two-qubit blocks precede Z rotations. So, applying this symmetry amounts to changing the angles in the Z rotations of the last checkerboard layer by φ.

Somewhat more complicated is the application of spin flips. We consider spin flips as applying one of X, Y, or Z operations to all spin simultaneously. The Z spin flip is a special case of the Z rotation(s). The Y flip can be composed out of X and Z flips, so we only need consider the X flip. Consider the quantum circuit in Fig. 7.

Let us use the fact that X and Z Pauli matrices anticommute and push the X gates to the front:

$$Xe^{i\frac{\theta}{2}Z} = \cos\frac{\theta}{2}X + i\sin\frac{\theta}{2}XZ = (\cos\frac{\theta}{2}\mathbb{1} - i\sin\frac{\theta}{2}Z)X = e^{-i\frac{\theta}{2}Z}X. \tag{A2}$$

Thus, if we push the X gates to the left and invert the angles of the Z rotations, the circuit remains invariant. Now, the next gate is the $Z \otimes Z$ rotation. $[X \otimes X, Z \otimes Z] = 0$, therefore the X gates can go through the R_{ZZ} gate without any changes. Finally, the X gates merge with the X rotations by incrementing the angle by $\pi/2$. Thus, to augment the data with the spin-flipped states, one inverts the angles of the Z rotations in the last layer, and increase the angles of the X rotations in the last layer by $\pi/2$.