Variational quantum anomaly detection: Unsupervised mapping of phase diagrams on a physical quantum computer

Korbinian Kottmann [1], Friederike Metz [2], Joana Fraxanet,[1] and Niccolò Baldelli [1]

[1]*ICFO - Institut de Ciencies Fotoniques, The Barcelona Institute of Science and Technology, Av. Carl Friedrich Gauss 3, 08860 Castelldefels (Barcelona), Spain*

[2]*Quantum Systems Unit, Okinawa Institute of Science and Technology Graduate University, 1919-1 Tancha, Onna, Okinawa 904-0495, Japan*

One of the most promising applications of quantum computing is simulating quantum many-body systems. However, there is still a need for methods to efficiently investigate these systems in a native way, capturing their full complexity. Here we propose variational quantum anomaly detection, an unsupervised quantum machine learning algorithm to analyze quantum data from quantum simulation. The algorithm is used to extract the phase diagram of a system with no prior physical knowledge and can be performed end-to-end on the same quantum device where the system is simulated on. We showcase its capabilities by mapping out the phase diagram of the one-dimensional extended Bose–Hubbard model with dimerized hoppings, which exhibit a symmetry protected topological phase. Further, we show that it can be used with readily accessible devices today by performing the algorithm on a real quantum computer.

I. INTRODUCTION

Since the discovery of Shor's algorithm [1], there have been many attempts to leverage the power of quantum computers to outperform classical computers [2]. One of the most promising near- and far-term applications is simulating quantum many-body systems. Universal quantum simulation algorithms such as phase estimation [2,3] or adiabatic quantum computation [4] require quantum fault tolerance [5], that is, the ability to reliably correct errors that occur during quantum computation. With recent experimental advancements in quantum error correction [6,7], there is hope that quantum fault tolerance can one day be reached. Currently we only have access to noisy intermediate-scale quantum (NISQ) devices [8]. There are several proposals for algorithms on these devices [9,10] such as the variational quantum eigensolver (VQE) [11], the quantum approximate optimization algorithm [12], or the quantum autoencoder [13,14], which employ parameterized circuits that are optimized through a classical feedback loop typically with gradient-based methods [15–17]. These approaches can suffer from so-called barren plateaus, the phenomenon of an exponentially vanishing gradient of the loss function [18]. In practice, this issue occurs for large systems but can be avoided by using shallow circuits and local cost functions [19]. While gradient-free optimization does not solve the barren plateau problem [20], other proposals give hope for large-scale and deep variational quantum circuits [21,22].

With the rise of deep learning in the 2010s, the term *quantum machine learning* was mostly used to refer to leveraging quantum computers for linear algebra tasks such as matrix inversion in subpolynomial time via the Harrow–Hassidim–Lloyd algorithm [23,24]. One famous use case was the quantum recommendation system algorithm with an exponential quantum speedup at the time [25], which inspired classical analogs of the algorithm with the same subpolynomial, complexity (termed machine learning algorithms) [26]. Today, quantum machine learning refers to using quantum circuits as neural networks [27] or kernel functions [28] to perform classical machine learning tasks such as supervised learning [29,30]. There are cases where quantum models have provable advantages over classical models [31], but it has been argued that these instances are special cases and no quantum speedup is to be expected for quantum machine learning with classical data [32].

On the other hand, applying classical machine learning to quantum physics has been a great success story [33], most prominently for the classification and mapping of phase diagrams [34–36]. These methods rely on classical data and are, therefore, restricted by the available classical simulation methods. With physical devices surpassing system sizes that are classically tractable [37], there is need for methods to investigate physical quantum states with quantum computers.

In this paper, we propose a quantum machine learning algorithm for *quantum data*. The data are ground states of quantum many-body systems that are prepared by a quantum simulation subroutine and serve as the input for *Variational Quantum Anomaly Detection* (VQAD). Our quantum anomaly detection (QAD) scheme belongs to the category of variational quantum algorithms where the circuit *learns* characteristic features of the input state [38]. This can in principle be leveraged for obtaining physical insights of the system from training [39] and is in con-
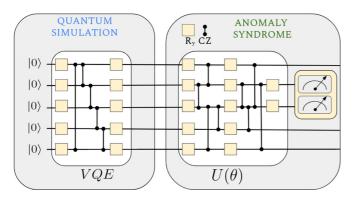
FIG. 1. Overview of our proposal. First, the quantum states are prepared via VQE. Then, they are processed through the anomaly syndrome consisting of a parameterized unitary $U(\theta)$ and a measurement of a subset of qubits, referred to as trash qubits. R_y indicates a parameterized y-axis rotation and CZ a (fixed) controlled-z gate.

trast to previous proposals that are based on kernel methods (one-class support vector machines) [40,41]. In the present study, we use it to map out an unknown phase diagram of a system without requiring knowledge about the order parameter or the number and location of the different phases.

In anomaly detection, the task is to differentiate *normal* data from *anomalous* data, opposed to supervised learning tasks where a fixed set of classes with labels for training are differentiated. On the other hand, the task of anomaly detection requires an *anomaly syndrome*, i.e., an observable that is trained to be of a certain value (typically 0) when *normal* data are input and be significantly larger for *anomalous* data it is tested on. In classical machine learning, anomaly detection has already been used to extract phase diagrams in an unsupervised fashion from simulated and experimental data [42–44]. VQAD allows us to perform anomaly detection directly *on* a quantum computer, and, with programmable devices readily available, we demonstrate it experimentally on a real device.

II. PROPOSAL

The task of detecting anomalies in ground states of quantum many-body Hamiltonians can be loosely divided into two subtasks: Preparing the ground state for specific Hamiltonian parameters and computing an anomaly syndrome indicating whether the state corresponds to a *normal* example or an anomaly. An overview of our proposed algorithm is shown in Fig. 1. The problem of state preparation on quantum computers is one of ongoing research, and in principle, one can use any state preparation subroutine for preparing the ground state. Here we choose the VQE as it has the lowest hardware requirements while achieving reliable results on current devices [11,45]. The VQE algorithm iteratively minimizes the expectation value of a Hamiltonian with the ansatz circuit to find the ground state by optimizing the parameters of the circuit via a quantum-classical feedback loop. We choose a minimal ansatz as depicted in Fig. 1 that is sufficient for simulating the Ising Hamiltonian discussed in Sec. III B. A shallow ansatz allows us to run both the quantum simulation and the QAD on real noisy devices. For more complex systems, the problem of finding a suitable hardware-efficient ansatz can be

addressed, for example, by the adaptive VQE algorithm [46]. In this work we employed the VQE implementation provided by the Qiskit library [47] and optimized it using simultaneous perturbation stochastic approximation (SPSA) [48]. For all technical details, we refer to Appendix A.

Once the ground state is prepared on the quantum device, a subsequent circuit serves as the anomaly syndrome. Our circuit ansatz is inspired by the recently proposed quantum autoencoder, which, similar to its classical counterpart, can be used for compression of classical and quantum data [13,14]. It is composed of several layers, each consisting of parameterized single qubit y rotations and controlled-z gates. After the final layer, a predefined number n_t of *trash* qubits is measured in the computational basis. The objective is to decouple the trash qubits from the rest of the system, effectively compressing the original ground state into a smaller number of qubits. The circuit parameters are then optimized to faithfully compress states that are considered *normal*. However, when the optimized circuit is tested on anomalous states not seen during training, it is expected that the circuit fails to decouple the trash qubits from the rest of the system. To quantify the degree of decoupling we use the Hamming distance d_H of the trash qubit measurement outcomes to the $|0\rangle^{\otimes n_t}$ state, i.e., the number of 1s in a bit-string of measurement outcomes [14]. The cost function C can then be defined as the Hamming distance averaged over several circuit evaluations $C = 1/N \sum_i^N d_{Hi}$, where N is the number of performed measurements or shots. The cost function can also be rewritten in terms of expectation values of local Pauli-z operators Z_j,

$$C = \frac{1}{N} \sum_{i=1}^{N} d_{Hi} = \frac{1}{2} \sum_{j=1}^{n_t} (1 - \langle Z_j \rangle). \qquad (1)$$

The VQAD circuit achieves perfect compression if the trash qubits are fully disentangled from the remaining qubits and mapped into the pure $|0\rangle^{\otimes n_t}$ state resulting in a cost equal to zero.

The specific circuit ansatz for the anomaly syndrome is shown in Fig. 1 for the case of $n_t = 2$ trash qubits. Each layer of the circuit starts with parameterized single-qubit y rotations applied to every qubit followed by a sequence of entangling controlled-z gates. The currently available NISQ devices are inherently noisy and the computations are subject to gate errors. To minimize the number of two-qubit gates we apply the controlled-z gates only between trash qubits and nontrash qubits as well as between trash qubits themselves instead of an all-to-all entangling map [14]. This entangling map is physically motivated as the goal of the circuit is to disentangle the trash qubits from the rest with the trash qubits resulting in the $|0\rangle^{\otimes n_t}$ state. In a single layer each nontrash qubit will be coupled to exactly one trash qubit. This entangling scheme is repeated in the subsequent layers until every nontrash qubit has been coupled to each trash qubit exactly once, i.e., the number of layers of the circuit is equal to n_t. After the final layer, additional single-qubit y rotations act on the trash qubits before they are measured.

Barren plateaus are the fundamental obstacle prohibiting training of variational circuits with increasing numbers of qubits [18]. It was previously shown that using local cost functions and circuits featuring a number of layers scaling at most
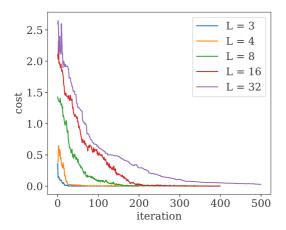
FIG. 2. Scaling of the training cost of the anomaly syndrome ansatz. Successful training of the proposed anomaly syndrome ansatz for $L \in \{3, 4, 8, 16, 32\}$ corresponding to $n_t \in \{1, 2, 3, 4, 5\}$ trash qubits (and therefore n_t circuit layers). The result for $L = 32$ was obtained through MPS simulations with a maximal bond dimension of BD = 100. We used 1000 shots per evaluation and achieve a perfect cost value of 0.00 for all system sizes; however, the run for $L = 32$ shown here finished at 0.03.

logarithmically in the system size can prevent the occurrence of barren plateaus [19]. Additionally for realistic devices, gate errors lead to decoherence, making quantum simulation on real devices a challenging task even for small systems and low depths [49]. The former calls for a minimal number of layers while the latter calls for a minimal number of gates overall. Therefore we seek a minimal solution for our variational circuit that we want to implement on a readily available NISQ-era quantum computer. On the other hand, it is desirable to have an ansatz as general as possible to be able to capture a wide range of problems (see *circuit complexity* [50,51]).

For the anomaly syndrome in this paper, we propose an ansatz that aims at compromising between being general enough to compress the ground states of the investigated systems and still being trainable. One way to make our circuit scalable for larger systems is to choose the number of trash qubits $n_t = \lfloor * \rfloor \log_2 L$, where L is the total number of qubits. Together with the fact that our cost function is composed of only local operators, the training is expected to not suffer from barren plateaus. We empirically confirm successful trainability, i.e., achieving a cost of 0.0 for ground states of the systems discussed later in the paper for $L \in \{3, 4, 8, 16, 32\}$ corresponding to $n_t \in \{1, 2, 3, 4, 5\}$, respectively. In Fig. 2, a ground state of the Ising model Eq. (5) at $g_x, g_z, J = (0.3, 0, 1)$ is taken as a realistic example and we can confirm successful trainability in all cases.

Note that in principle the trash qubits can be placed anywhere in the circuit; however, when performing computations on a real quantum device, it proved advantageous to explicitly take the qubit connectivity structure of the device into account in order to reduce the number of required SWAP operations. Specifically here we placed the trash qubits in the middle of the IBMQ devices.

The training and inference procedure is identical to the classical anomaly detection schemes for mapping out phase diagrams [42]. In the first step, one randomly chooses a train-

ing region in the phase diagram that represents *normal* data, which is an arbitrary definition. Note that no prior knowledge about the phase diagram is therefore required. The circuit representing the anomaly syndrome is then trained on ground states of the training region and tested on the whole phase diagram. States in the same phase as the training data are *normal* and can be disentangled, leading to a low cost. *Anomalous* states can be inferred through an increase in the cost function, signaling that the corresponding ground state cannot be disentangled by the optimized circuit. From the resultant cost profile, we can deduce the phase boundary between the phase the circuit has been trained on and any other phases in the diagram. This procedure is then repeated by training in the anomalous region from the previous iteration until all phase boundaries are found. An example is provided in Fig. 3.

Anomaly detection is a semisupervised learning task. The setting is typically that one is provided with one class of data that is well known, *normal* data, and aims at finding outliers of that distribution, *anomalous* data. An archetypical example is credit card fraud where a big database of normal transactions is provided and one aims at finding fraudulent ones. We consider anomaly detection semisupervised as labeled, normal data is provided while anomalous data is absent for training and to be inferred. Here, however, we arbitrarily define (x, normal) and iteratively find the different classes (phases of matter). The definition of (x, normal) is arbitrary and does not necessitate prior knowledge. Furthermore, it is merely a means to an end to find the different classes. In that sense, the way anomaly detection is used to map out the phase diagram can be regarded as an unsupervised learning method.

Note that in previous works, where the same task has been tackled with classical machine learning techniques, it has been shown that a single ground state was sufficient to successfully train the model [44]. This feature stems from the fact that ground states within the same phase share similar properties and there is very little variance when changing the physical parameters inside one phase. We observe this feature also in the training of the VQAD.

III. RESULTS

A. Simulations with ideal quantum data

In order to test the performance of VQAD, we first study the one-dimensional extended Bose–Hubbard model with dimerized hoppings (DEBHM) [52],

$$H = -\sum_{i=1}^{L-1} (J + \delta J(-1)^i)(b_i^\dagger b_{i+1} + \text{H.c.})$$
$$+ \frac{U}{2} \sum_i^L n_i(n_i - 1) + V \sum_i^{L-1} n_i n_{i+1}, \qquad (2)$$

where $b_i^\dagger (b_i)$ is the bosonic operator representing the creation(annihilation) of a particle at site i of a lattice of length L. The tunneling amplitudes $J - \delta J(J + \delta J)$ indicate hopping processes on odd (even) links connecting nearest-neighbor (NN) sites, while V represents the NN repulsion. Here we take the hardcore boson limit, i.e., the on-site repulsion $U/J \to \infty$, such that the local Hilbert space is two-dimensional and each
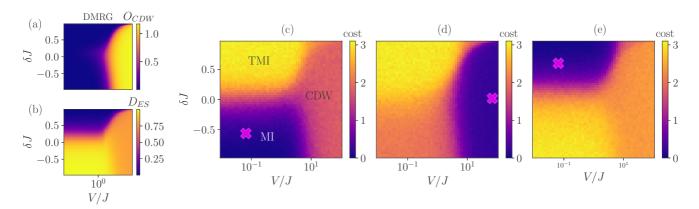
FIG. 3. (a), (b) Phase diagram of the DEBHM from Eq. (2) using (a) the order-parameter O_{CDW} defined in Eq. (3) and (b) the degeneracy of the ES, D_{ES}, defined in Eq. (4). The results were obtained from DMRG simulations for a system of length $L = 12$ at half filling $\bar{n} = 0.5$. We fix the maximum bond dimension $BD = 50$ and the maximum number of bosons per site to $n_0 = 1$. (c)–(e) Cost/anomaly syndrome of a VQAD trained on a single ground state (indicated by a cross) of the $L = 12$ DEBHM using $n_t = 6$ trash qubits in the (c) MI phase, (d) CDW phase, and (e) TMI phase. The cost at each data point is the Hamming distance averaged over 1000 measurement shots using an ideal quantum device simulator.

site can only accommodate 0 or 1 bosons. This model can be effectively mapped into a spin-1/2 system [53].

Previous studies of the DEBHM model at half filling ($\bar{n} = 0.5$) have demonstrated the existence of three distinct phases [52]. For small and intermediate values of V/J and $\delta J > 0$, we find a topological Mott insulator (TMI) displaying features analogous to a symmetry protected topological phase appearing in the dimerized spin-1/2 bond-alternating Heisenberg model [53]. On the other hand, for negative values of δJ we expect a trivial Mott insulator (MI), while in the regime where the NN repulsion dominates, a charge-density wave (CDW) appears.

In Fig. 3(a) and 3(b), we study the phase diagram of the model in Eq. (2) in terms of the parameters δJ and V/J using the density matrix renormalization group algorithm (DMRG) [54–56]. In order to differentiate between the MI phases and the CDW, one can compute the CDW order-parameter

$$O_{\text{CDW}} = \sum_{i=1}^{L/2} (-1)^i \delta n_i, \tag{3}$$

which detects staggered patterns in the density. In Fig. 3(a) we report a vanishing value of O_{CDW} everywhere but in the region with large values of V/J, which corresponds to the CDW [57]. To characterize the TMI we study the entanglement spectrum (ES), which is expected to be doubly degenerate in a topologically nontrivial phase [58] due to the existence of edge states. The ES $\{\lambda_i\}$ is defined in terms of the positive real-valued Schmidt coefficients $\{\alpha_i\}$ of a bipartite decomposition of the system by $\alpha_i^2 = \exp(-\lambda_i)$. We determine its degeneracy using

$$D_{ES} = \sum_i (-1)^i e^{-\lambda_i}. \tag{4}$$

In Fig. 3(b), we show that the quantity D_{ES} vanishes only for small NN interaction strengths V and positive values of δJ, which corresponds to the TMI. The trivial MI and CDW phases do not show a degeneracy and hence do not host topological edge states.

In the following, we test the capabilities of the VQAD with ideal states obtained from DMRG simulations. The anomaly syndrome is trained using a single representative ground state within one of the phases such that the cost measured at the trash qubits is minimized and the states of this phase can be efficiently compressed by the circuit. Afterwards, the trained circuit processes all states from the full phase diagram, ideally with similarly low cost in the same phase and significantly higher cost in other phases.

In Fig. 3(c)–3(e) we show the resultant cost diagram for three circuits, each optimized at a different point in the phase diagram. Indeed, ground states outside of the training phase give rise to a large cost and hence are correctly classified by the VQAD as anomalous. Surprisingly, a single ground-state example (indicated by the cross) was sufficient to successfully train the VQAD and infer all three phases. Similar results were recently reported for the case of classical anomaly detection using neural network autoencoders [44].

To demonstrate the robustness of the VQAD against noise present in currently available NISQ devices we apply a depolarizing noise channel after each gate with error probabilities $p_{\text{err}} = 0.001$ (single-qubit gates) and $p_{\text{err}} = 0.01, 0.07$ (two-qubit gates) and show two exemplary cost profiles of the trained anomaly detector in Fig. 4. Since the noise becomes more prominent with larger circuit depths, we used the two-layer VQAD circuit ansatz with only two trash qubits in this case. While it is not possible to reach a cost of zero in the training phase, the optimization still converges and all three phases can be successfully inferred. Hence this suggests that even if the VQAD is not able to fully disentangle the trash qubits, the phase diagram can still be recovered from the resultant cost profile.

B. Experiments on a real quantum computer

We have seen that with ideal quantum data, VQAD can map out nontrivial phase diagrams including topologically nontrivial phases with and without noise in the anomaly
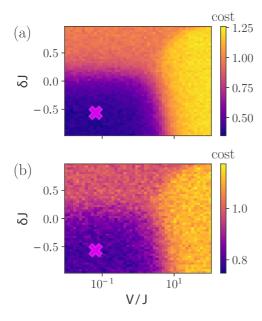
FIG. 4. Cost of a VQAD trained on a single ground state in the MI phase (marked by the cross) of the DEBHM with $L = 12$ sites and $n_t = 2$ trash qubits. The gates of the VQAD circuit are subject to depolarizing noise with $p_{err} = 0.001$ (single-qubit gates) and (a) $p_{err} = 0.01$, (b) $p_{err} = 0.07$ (two-qubit gates). The chosen values are motivated by the error probabilities of real devices.
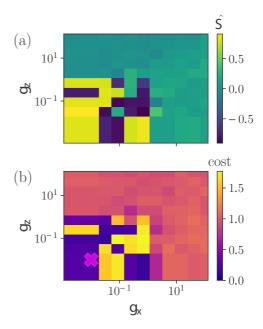

FIG. 5. Real-noise simulations of the staggered magnetization $\hat{S}$ (a) and the anomaly syndrome (b) for the TLFI model. We trained the anomaly syndrome in the ordered phase on a state with positive $\hat{S}$, indicated by the purple cross. Inside the ordered phase, there is a perfect correlation between low cost states for positive $\hat{S}$, and very high cost where VQE converged to a negative $\hat{S}$. The paramagnetic phase is detected by a plateau in the anomaly syndrome.

syndrome. Next we discuss its performance in real-noise simulations, that is, with noise profiles and qubit connectivities from a real quantum device. Furthermore, we perform the quantum simulation subroutine, i.e., the ground-state preparation via VQE, on the same circuit. For this task, we consider the paradigmatic transverse longitudinal field Ising (TLFI) model [59],

$$H = J \sum_{i=1}^{L} Z_i Z_{i+1} - g_x \sum_{i=1}^{L} X_i - g_z \sum_{i=1}^{L} Z_i, \qquad (5)$$

where X_i, Z_i are the Pauli matrices on site i, J is the coupling strength, and g_x, g_z are the transverse and longitudinal fields, respectively. For $g_z = 0$ the model is exactly solvable and shows a quantum phase transition from a ferromagnetic (antiferromagnetic) phase for $g_x/J < 1$ and J negative (positive) to a paramagnetic one for $g_x/J > 1$ [60]. In the following we set $J = 1$ and vary the longitudinal and transverse fields. In this regime the model is not exactly solvable and the phase diagram has been extensively studied numerically [61,62]. The antiferromagnet-paramagnet quantum phase transition is best characterized by the order parameter which in this case is the staggered magnetization

$$\hat{S} = \sum_{i=1}^{L} (-1)^i \frac{Z_i}{L}. \qquad (6)$$

We simulate the ground states of the Hamiltonian in Eq. (5) using VQE for $L = 5$. On a noisy device, long-range entangling gates are performed by consecutive local two-qubit gates (SWAP operation), increasing the actual circuit depth. A large

number of consecutive gates leads to decoherence due to gate errors and destroys the results. With the circuit presented in Fig. 1 for the VQE subroutine, we found a trade-off between expressibility and noise tolerance with a circular entanglement distribution and only one layer. Additionally, we performed measurement error mitigation [63], which can further improve the results of the cost function as seen in Fig. 7. For small values of g_x and g_z, in the ferromagnetic ordered phase, the ground-states $\psi \simeq |10101\rangle$ ($\langle \hat{S} \rangle = 1$) and $\psi \simeq |01010\rangle$ ($\langle \hat{S} \rangle = -1$) have a similar energy, which is why the optimization can get stuck in local minima. Hence in the ordered phase, VQE can converge to either a state with positive or negative staggered magnetization or an equal superposition of the two as can be seen in Fig. 5(a). The VQAD simulation results in Fig. 5(b) show a perfect correlation between positive $\langle \hat{S} \rangle$ and low cost and vice versa negative $\langle \hat{S} \rangle$ and high cost, which, intuitively, can be expected [64]. The disordered phase is detected from the plateau of high cost (~ 1).

We see that VQAD also performs well under realistic conditions so we next test the algorithm on a physical device. For this task, we use the $L = 5$ qubits on `ibmq_jakarta` [63]. To avoid jumps in the staggered magnetization in the ordered phase and improve convergence of the VQE optimization, we reuse already optimized parameters at neighboring points in the phase diagram as a good initial guess. Due to a large computation time overhead per execution on the real device, we additionally prepared preoptimized parameters for both subroutines from a realistic noisy simulation and use these as initial guesses for the optimization on the device. We found
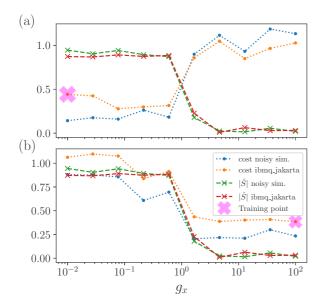
(a)

(b)

FIG. 6. Real device VQAD experiments: We show the order-parameter $\hat{S}$ compared to the VQAD results both for execution on `ibmq_jakarta` and for noisy simulators with the same noise profile. We trained on a single ground state in the ordered (a) and paramagnetic (b) phases. For sampling $\hat{S}$, we use the same parameters for the VQE circuit in simulation and experiment. All values for $\hat{S}$ in the paramagnetic phase are negative and hence for better visualization we plot its absolute value $|\hat{S}|$. For training the anomaly syndrome, the optimized parameters from the simulation are taken as an initial guess.

that for computing the staggered magnetization it is actually not necessary to rerun the VQE optimization on the physical device, and we can achieve faithful results by directly using the optimized parameters from the simulation as seen in Fig. 6. The resulting cost values for the optimized circuit, plotted in Fig. 6, clearly distinguish the two phases, with the cost from the experiment showing solely an almost constant offset compared with the noisy simulation.

IV. OUTLOOK

We showed that our proposed algorithm is capable of mapping out complex phase diagrams, including topologically nontrivial phases. We further demonstrated that the algorithm also works in realistic scenarios for both real-noise simulations and on a real quantum computer. Hence we provide a tool to experimentally explore phase diagrams in future quantum devices, which will be especially useful when physical devices surpass the limit of what can be classically computed.

Currently, the main bottleneck of VQAD is the presence of noise in real devices. We were able to improve our anomaly detection scheme by employing measurement error mitigation and adopting the circuits according to the physical device. These results are promising, and with current efforts on enhancing device performances, error mitigation, and circuit optimization strategies in the community, we are hopeful to see even further improvements soon.

In this work we focused on using VQAD to extract the phase diagram of quantum many-body systems. A possible future extension would be to apply it to the problem of entanglement witnessing and certification in many-body scenarios without tomography. Furthermore, the use of an autoencoder-like architecture has the advantage over kernel-based schemes in that there exist tools of interpreting the feature space in classical autoencoders to gain physical insights [39], which can be a possible future extension for the quantum case discussed here.

APPENDIX: TECHNICAL DETAILS

The code to run the simulations and experiments discussed in the main text can be found in our repository on GitHub [65]. The optimization of the circuit parameters was performed using SPSA [45,48]. To obtain the results presented

in Fig. 3, a VQAD circuit ansatz composed of six layers (6 trash qubits) was employed resulting in $6L + 6$ parameters. For the noisy simulations and real-device execution discussed in Sec. III B, we used the ansatz in Fig. 1, counting $2L$ and $2L + 2$ parameters for the quantum simulation and anomaly syndrome, respectively. In classical real-noise simulations, we used 500 VQE optimization iterations for the initial ground-state optimization, and 200 iterations for all subsequent optimizations where the previously optimized parameters were taken as initial guesses. For the anomaly detection circuit, we found converged results with less than 100 optimization iterations. As an example, calculating the expectation value of the magnetization takes roughly $2-10$ s on a commercial laptop (here: i7-4712HQ), while the real-device execution takes about 30 seconds. Furthermore, we used measurement error mitigation [63] provided by the Qiskit library to improve the results of the VQAD simulations in the presence of noise as illustrated in Fig. 7.
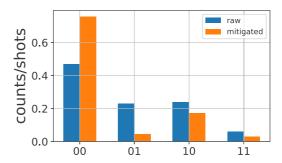

FIG. 7. Comparison of the trash qubit measurement outcomes with and without measurement error mitigation. The anomaly syndrome circuit has been trained with and without error mitigation on a ground state of the TLFI model in the ordered phase in real-noise simulations. Ideally, all of the 1000 shots would result in the 00 bit string. By mitigating the measurement errors we improve the results toward this desired outcome.

[1] P. W. Shor, Algorithms for quantum computation: discrete logarithms and factoring, in *Proceedings of the 35th Annual Symposium on Foundations of Computer Science, NW Washington, DC, US* (IEEE, 1994), pp. 124–134.

[2] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition* (Cambridge University Press, Cambridge, 2010).

[3] A. Yu. Kitaev, Quantum measurements and the Abelian Stabilizer Problem, arXiv:quant-ph/9511026 (1995).

[4] E. Farhi, J. Goldstone, S. Gutmann, and M. Sipser, Quantum computation by adiabatic evolution, arXiv:quant-ph/0001106 (2000).

[5] D. Aharonov and M. Ben-Or, Fault-tolerant quantum computation with constant error rate, SIAM J. Comput. **38**, 1207 (2008).

[6] G. Semeghini, H. Levine, A. Keesling, S. Ebadi, Tout T. Wang, D. Bluvstein, R. Verresen, H. Pichler, M. Kalinowski, R. Samajdar, A. Omran, S. Sachdev, A. Vishwanath, M. Greiner, V. Vuletic, and M. D. Lukin, Probing topological spin liquids on a programmable quantum simulator, arXiv:2104.04119 (2021).

[7] K. J. Satzinger, Y. Liu, A. Smith, C. Knapp, M. Newman, C. Jones, Z. Chen, C. Quintana, X. Mi, A. Dunsworth, C. Gidney, I. Aleiner, F. Arute, K. Arya, J. Atalaya *et al.*, Realizing topologically ordered states on a quantum processor, arXiv:2104.01180 (2021).

[8] J. Preskill, Quantum Computing in the NISQ era and beyond, Quantum **2**, 79 (2018).

[9] M. Benedetti, E. Lloyd, S. Sack, and M. Fiorentini, Parameterized quantum circuits as machine learning models, Quantum Sci. Technol. **4**, 043001 (2019).

[10] K. Bharti, Alba Cervera-Lierta, Thi H. Kyaw, T. Haug, Sumner Alperin-Lea, A. Anand, M. Degroote, H. Heimonen, J. S. Kottmann, T. Menke, W.-K. Mok, S. Sim, L.-C. Kwek, and A. Aspuru-Guzik, Noisy intermediate-scale quantum (NISQ) algorithms, arXiv:2101.08448 (2021).

[11] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O'Brien, A variational eigenvalue solver on a photonic quantum processor, Nat. Commun. **5**, 4213 (2014).

[12] E. Farhi, J. Goldstone, and S. Gutmann, A quantum approximate optimization algorithm, arXiv:1411.4028 (2014).

[13] J. Romero, J. P. Olson, and A. Aspuru-Guzik, Quantum autoencoders for efficient compression of quantum data, Quantum Sci. Technol. **2**, 045001 (2017).

[14] C. Bravo-Prieto, Quantum autoencoders with enhanced data encoding, Mach. Learn.: Sci. Technol. **2**, 035028 (2021).

[15] J. R. McClean, J. Romero, R. Babbush, and A. Aspuru-Guzik, The theory of variational hybrid quantum-classical algorithms, New J. Phys. **18**, 023023 (2015).

[16] J. Stokes, J. Izaac, N. Killoran, and G. Carleo, Quantum natural gradient, Quantum **4**, 269 (2019).

[17] M. Cerezo, A. Arrasmith, R. Babbush, Simon C. Benjamin, S. Endo, K. Fujii, J. R. McClean, K. Mitarai, X. Yuan, L. Cincio, and P. J. Coles, Variational quantum algorithms, arXiv:2012.09265 (2020).

[18] J. R. McClean, S. Boixo, V. N. Smelyanskiy, R. Babbush, and H. Neven, Barren plateaus in quantum neural network training landscapes, Nat. Commun. **9**, 4812 (2018).

[19] M. Cerezo, A. Sone, T. Volkoff, L. Cincio, and P. J. Coles, Cost function dependent barren plateaus in shallow parametrized quantum circuits, Nat. Commun. **12**, 1791 (2021).

[20] A. Arrasmith, M. Cerezo, P. Czarnik, L. Cincio, and Patrick J. Coles, Effect of barren plateaus on gradient-free optimization, Quantum **5**, 558 (2021).

[21] A. Pesah, M. Cerezo, S. Wang, T. Volkoff, Andrew T. Sornborger, and Patrick J. Coles, Absence of Barren Plateaus in Quantum Convolutional Neural Networks, Phys. Rev. X **11**, 041011 (2021).

[22] T. Volkoff and P. J. Coles, Large gradients via correlation in random parameterized quantum circuits, Quantum Sci. Technol. **6**, 025008 (2021).

[23] A. W. Harrow, A. Hassidim, and S. Lloyd, Quantum Algorithm for Solving Linear Systems of Equations, Phys. Rev. Lett. **103**, 150502 (2009).

[24] J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, Nathan Wiebe, and S. Lloyd, Quantum machine learning, Nature **549**, 195 (2017).

[25] I. Kerenidis and A. Prakash, Quantum recommendation systems, arXiv:1603.08675.

[26] E. Tang, A quantum-inspired classical algorithm for recommendation systems, in *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing* (ACM, New York, 2018), pp. 217–228.

[27] A. Pérez-Salinas, A. Cervera-Lierta, E. Gil-Fuster, and J. I. Latorre, Data re-uploading for a universal quantum classifier, Quantum **4**, 226 (2020).

[28] M. Schuld, Supervised quantum machine learning models are kernel methods, arXiv:2101.11020 (2021).

[29] E. Farhi and H. Neven, Classification with quantum neural networks on near term processors, arXiv:1802.06002 (2018).

[30] P. Rebentrost, M. Mohseni, and S. Lloyd, Quantum Support Vector Machine for Big Data Classification, Phys. Rev. Lett. **113**, 130503 (2013).

[31] Y. Liu, S. Arunachalam, and K. Temme, A rigorous and robust quantum speed-up in supervised machine learning, arXiv:2010.02174 (2020).

[32] J. M. Kübler, S. Buchholz, and B. Schölkopf, The inductive bias of quantum kernels, arXiv:2106.03747 (2021).

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

[34] J. Carrasquilla and R. G. Melko, Machine learning phases of matter, Na. Phys. **13**, 431 (2016).

[35] E. P. L. van Nieuwenburg, Y.-H. Liu, and S. D. Huber, Learning phase transitions by confusion, Nat. Phys. **13**, 435 (2016), arXiv:1610.02048.

[36] P. Huembeli, A. Dauphin, P. Wittek, and C. Gogolin, Automated discovery of characteristic features of phase transitions in many-body localization, Phys. Rev. B **99**, 104106 (2019).

[37] F. Arute, K. Arya, R. Babbush, D. Bacon, Joseph C. Bardin, Rami Barends, R. Biswas, S. Boixo, F. G. S. L. Brandao, D. A. Buell, B. Burkett, Yu. Chen, Z. Chen, B. Chiaro, R. Collins *et al.*, Quantum supremacy using a programmable superconducting processor, Nature **574**, 505 (2019).

[38] The term *learning* is commonly used in (quantum) machine learning and data-driven problem solving to refer to data-specific optimization.

[39] R. Iten, T. Metger, H. Wilming, Lidia del Rio, and R. Renner, Discovering Physical Concepts with Neural Networks, Phys. Rev. Lett. **124**, 010508 (2020).

[40] N. Liu and P. Rebentrost, Quantum machine learning for quantum anomaly detection, Phys. Rev. A, **97**, 042315 (2018).

[41] J.-M. Liang, S.-Q. Shen, M. Li, and L. Li, Quantum anomaly detection with density estimation and multivariate Gaussian distribution, Phys. Rev. A **99**, 052310 (2019).

[42] K. Kottmann, P. Huembeli, M. Lewenstein, and A. Acín, Unsupervised Phase Discovery with Deep Anomaly Detection, Phys. Rev. Lett. **125**, 170603 (2020).

[43] N. Käming, A. Dawid, K. Kottmann, M. Lewenstein, Klaus Sengstock, A. Dauphin, and C. Weitenberg, Unsupervised machine learning of topological phase transitions from experimental data, Mach. Learn.: Sci. Technol. **2**, 035037 (2021).

[44] K. Kottmann, P. Corboz, M. Lewenstein, and A. Acín, Unsupervised mapping of phase diagrams of 2D systems from infinite projected entangled-pair states via deep anomaly detection, SciPost Phys. **11**, 025 (2021).

[45] A. Kandala, A. Mezzacapo, K. Temme, M. Takita, M. Brink, Jerry M. Chow, and Jay M. Gambetta, Hardware-efficient variational quantum eigensolver for small molecules and q. magnets, Nature (London) **549**, 242 (2017),.

[46] Harper R. Grimsley, Sophia E. Economou, E. Barnes, and N. J. Mayhall, An adaptive variational algorithm for exact molecular simulations on a quantum computer, Nat. Commun. **10**, 3007 (2019).

[47] H. Abraham, AduOffei, R. Agarwal, I. Y. Akhalwaya, G. Aleksandrowicz, T. Alexander, M. Amy, and E. Arbel, Arijit02, A. Asfaw, A. Avkhadiev, C. Azaustre, AzizNgoueya, A. Banerjee, Aman Bansal *et al.*, Qiskit: An open-source framework for quantum computing (2019).

[48] J. Spall, An overview of the simultaneous perturbation method for efficient optimization, Johns Hopkins APL Tech. Dig. **19**, 482 (1998).

[49] A. Cervera-Lierta, Exact Ising model simulation on a quantum computer, Quantum **2**, 114 (2018).

[50] E. Bernstein and U. Vazirani, Quantum complexity theory, SIAM J. Comput. **26**, 1411 (1997).

[51] F. G. S. L. Brandão, W. Chemissany, N. Hunter-Jones, R. Kueng, and J. Preskill, Models of quantum complexity growth, arXiv:1912.04297 (2019).

[52] K. Sugimoto, S. Ejima, F. Lange, and H. Fehske, Quantum phase transitions in the dimerized extended Bose-Hubbard model, Phys. Rev. A **99**, 012122 (2019).

[53] H. T. Wang, Bo Li, and S. Y. Cho, Topological quantum phase transition in bond-alternating spin-$\frac{1}{2}$ Heisenberg chains, Phys. Rev. B **87**, 054402 (2013).

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

[55] S. R. White, Density Matrix Formulation for Quantum Renormalization Groups, Phys. Rev. Lett. **69**, 2863 (1992).

[56] J. Hauschild and F. Pollmann, Efficient numerical simulations with Tensor Networks: Tensor Network Python (TeNPy), Sci. Post. Phys. Lect. Notes 5 (2018), Code available from https://github.com/tenpy/tenpy.

[57] In the definition of O_{CDW}, we consider only half of the sites of the system because the DMRG algorithm outputs a symmetric state, which is a superposition of the two degenerate ground states.

[58] F. Pollmann, Ari M. Turner, E. Berg, and M. Oshikawa, Entanglement spectrum of a topological phase in one dimension, Phys. Rev. B **81**, 064439 (2010).

[59] H. C. Fogedby, The ising chain in a skew magnetic field, J. Phys. C: Solid State Phys. **11**, 2801 (1978).

[60] S. Sachdev, *Quantum Phase Transitions*, American Cancer Society, 2007, ISBN 9780470022184, https://onlinelibrary.wiley.com/doi/abs/10.1002/9780470022184.hmm108.

[61] P. Sen, Quantum phase transitions in the ising model in a spatially modulated field, Phys. Rev. E **63**, 016112 (2000).

[62] A. A. Ovchinnikov, D. V. Dmitriev, V. Ya. Krivnov, and V. O. Cheranovskii, Antiferromagnetic ising chain in a mixed transverse and longitudinal magnetic field, Phys. Rev. B **68**, 214406 (2003).

[63] S. Bravyi, S. Sheldon, A. Kandala, D. C. McKay, and J. M. Gambetta, Mitigating measurement errors in multiqubit experiments, Phys. Rev. A **103**, 042605 (2021).

[64] In a very hand-wavy way, we can understand this as we train the circuit U to perform $U|01010\rangle = |\Psi\rangle \otimes |00\rangle \text{trash}$ such that $U|01010\rangle = |\Psi\rangle \otimes |11\rangle_{\text{trash}}$ if we input a state with opposite ordering.

[65] K. Kottmann, F. Metz, J. Fraxanet, and N. Baldelli, Variational quantum anomaly detection code on github (2021b), https://github.com/Qottmann/Quantum-anomaly-detection.