

## Quantum resources of quantum and classical variational methods

Spriggs, T.E.; Ahmadi, A.; Chen, B.; Greplová, E.

**DOI**

[10.1088/2632-2153/adaca2](https://doi.org/10.1088/2632-2153/adaca2)

**Publication date**

2025

**Document Version**

Final published version

**Published in**

Machine Learning

**Citation (APA)**

Spriggs, T. E., Ahmadi, A., Chen, B., & Greplová, E. (2025). Quantum resources of quantum and classical variational methods. *Machine Learning*, 6(1), Article 015042. <https://doi.org/10.1088/2632-2153/adaca2>

**Important note**

To cite this publication, please use the final published version (if applicable).  
Please check the document version above.

**Copyright**

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

**Takedown policy**

Please contact us and provide details if you believe this document breaches copyrights.  
We will remove access to the work immediately and investigate your claim.

PAPER • OPEN ACCESS

## Quantum resources of quantum and classical variational methods

To cite this article: Thomas Spriggs *et al* 2025 *Mach. Learn.: Sci. Technol.* **6** 015042

View the [article online](#) for updates and enhancements.

### You may also like

- [Quantitative assessment of PINN inference on experimental data for gravity currents flows](#)  
Mickaël Delcey, Yoann Cheny, Jean Schneider et al.
- [Efficient Monte Carlo simulation of streamer discharges with deep-learning denoising models](#)  
F M Bayo-Muñoz, A Malagón-Romero and A Luque
- [Exact solution of long-range stabilizer Rényi entropy in the dual-unitary XXZ model](#)  
Jordi Arnau Montaña López and Pavel Kos



## PAPER

## OPEN ACCESS

RECEIVED  
4 October 2024REVISED  
17 December 2024ACCEPTED FOR PUBLICATION  
21 January 2025PUBLISHED  
14 February 2025

Original Content from  
this work may be used  
under the terms of the  
[Creative Commons  
Attribution 4.0 licence](#).

Any further distribution  
of this work must  
maintain attribution to  
the author(s) and the title  
of the work, journal  
citation and DOI.



## Quantum resources of quantum and classical variational methods

Thomas Spriggs<sup>1,3,\*</sup> , Arash Ahmadi<sup>1,3,\*</sup> , Bokai Chen<sup>2</sup> and Eliska Greplova<sup>1</sup> <sup>1</sup> QuTech and Kavli Institute of Nanoscience, Delft University of Technology, Delft, The Netherlands<sup>2</sup> Kavli Institute of Nanoscience, Delft University of Technology, Delft, The Netherlands<sup>3</sup> TS and AA contributed equally.

\* Authors to whom any correspondence should be addressed.

E-mail: [t.e.spriggs@tudelft.nl](mailto:t.e.spriggs@tudelft.nl) and [a.ahmadi-1@tudelft.nl](mailto:a.ahmadi-1@tudelft.nl)**Keywords:** quantum information, neural networks, variational methods, quantum circuits, tensor networks, neural quantum states

## Abstract

Variational techniques have long been at the heart of atomic, solid-state, and many-body physics. They have recently extended to quantum and classical machine learning, providing a basis for representing quantum states via neural networks. These methods generally aim to minimize the energy of a given ansatz, though open questions remain about the expressivity of quantum and classical variational ansätze. The connection between variational techniques and quantum computing, through variational quantum algorithms, offers opportunities to explore the quantum complexity of classical methods. We demonstrate how the concept of non-stabilizerness, or magic, can create a bridge between quantum information and variational techniques and we show that energy accuracy is a necessary but not always sufficient condition for accuracy in non-stabilizerness. Through systematic benchmarking of neural network quantum states, matrix product states, and variational quantum methods, we show that while classical techniques are more accurate in non-stabilizerness, not accounting for the symmetries of the system can have a severe impact on this accuracy. Our findings form a basis for a universal expressivity characterization of both quantum and classical variational methods.

## 1. Introduction

Due to the prohibitive scaling of large scale quantum wavefunctions, approximate methods are typically employed to numerically find ground states and time evolve quantum systems. Classical variational methods have a rich track record and span physics intuition driven parametrization of the trial wavefunction [1–5], ansätze capturing specific entanglement properties [6, 7], and, more recently, physics-agnostic wavefunctions based on neural networks [8–13].

In parallel to this progress, quantum variational methods were formulated with the same goal: to capture potentially complicated wavefunctions using a small number of degrees of freedom [14–17]. In this instance, the variational ansatz is formulated as a quantum circuit. Quantum operations in such circuits are then parametrized and the parameters of the quantum gates are updated, using classical optimization techniques, until the energy of the output of the circuit well approximates the energy of the sought-after quantum state.

In search for ground states of specific Hamiltonians, both classical and quantum variational techniques rely on the energy variational principle: the expectation value of the Hamiltonian is evaluated with respect to the trial state and then minimized. The lowest achievable value is then considered a good approximation for the true ground state energy of the system.

Once we optimize the energy of our ansatz, the question arises regarding how well is the quantum state itself represented by our approximate representation. While variational minimization brings the system to the point in Hilbert space with the required energy, it is far from certain that the optimized state will structurally reflect the global properties of the system. In particular, especially long range correlations are known to pose a challenge.

In this work, we assess the expressivity of both quantum and classical variational methods from a quantum computational complexity point of view: we analyze the amount of quantum resources that the variational ansatz expresses when optimized in accordance with the classical energy variational principle.

In quantum computing, one way of assessing quantum resources is to measure how far a given state is from being efficiently and exactly simulated on a classical computer [18]. The quantum operations that allow for efficient, exact classical simulation are elements of the Clifford group. All the other operations are referred to as non-Clifford. To quantify how ‘far’ a given operation is from the Clifford group, a measure called magic, or non-stabilizerness, was introduced [19].

Non-stabilizerness has recently been at the forefront of quantum information literature as more scalable techniques have emerged to quantitatively evaluate it [20–24]. This progress allowed for the first exploration of non-stabilizerness evaluation for tensor networks [22, 25, 26].

At first sight, the notions of non-stabilizerness and classical variational principle are unrelated. One is designed to determine areas of potential quantum-computational advantage, the other to assess performance of classical, few-parameter ansätze. However, a key property we want from a variational ansatz is to faithfully approximate the sought-after state, beyond just its energy. An interesting interplay with quantum information arises: classical variational ansätze are unrestricted by, for example, the Gottesmann–Knill theorem [18], and are allowed to converge to any point in the Hilbert space that minimize the energy regardless of where the exact solution lies with respect to the Clifford group.

In this work, we perform a systematic benchmark for non-stabilizerness expressivity of classical and quantum variational ansätze using the transverse-field Ising model (TFIM), a common testbed model for variational methods. Specifically, we compare non-stabilizerness expressivity of neural quantum states (NQS), matrix product states (MPS), and a variational quantum eigensolver (VQE). We find that high energy precision is a necessary, but not always sufficient, condition for the few-parameter variational ansätze to express the non-stabilizerness of the model correctly. We observe that quantum variational methods generally show worse performance compared to the state-of-the-art classical methods. In the context of this work it means that even though the ansatz itself is quantum, it is worse at predicting the energy and non-stabilizerness of the exact wavefunction. Moreover, the variations in energy and non-stabilizerness accuracy across multiple repeats are generally much larger for the quantum variational method compared to the results from neural quantum states.

The outline of this paper is as follows. Section 2 provides an overview of methods and techniques used in this work: section 2.1 briefly elucidates the specific Hamiltonian with which this work explores, section 2.2 details the three algorithms used for finding the ground state of the given Hamiltonian, and section 2.3 presents the measure of quantum resources that we use to assess the non-stabilizerness of the ground state wavefunctions obtained. Section 3 presents the comparison of the ground state energy and non-stabilizerness found using density matrix renormalization group (DMRG), NQS, and a VQE; this comparison focuses on the respective accuracies of different methods compared to exact diagonalization (ED), the interplay between the energy and non-stabilizerness accuracies, and the fluctuations of their solutions with repeated runs. In section 4 we discuss our findings and possible next steps in aligning classical and quantum variational methods.

## 2. Methods

### 2.1. TFIM

This paper is focused on parameterized ansätze that approximate the ground state of the TFIM. The Hamiltonian of this model is defined as

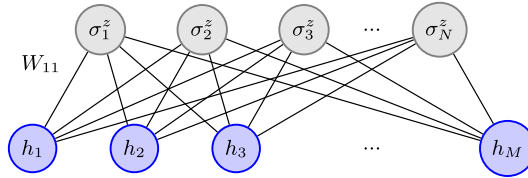
$$H = J \sum_i \sigma_i^z \sigma_{i+1}^z - h \sum_i \sigma_i^x. \quad (1)$$

In this definition,  $\sigma^z$  and  $\sigma^x$  are the Pauli matrices,  $J$  is a constant that dictates the coupling between neighboring spins (we will consider  $J < 0$  and thus neighboring spins being aligned is energetically favorable), and  $h$  is the transverse magnetic field strength; we are considering the case of periodic boundaries. The basis that will be used to represent the wavefunction will be the spin projections along the  $z$ -axis, comprised of either spin up or down relative to this axis. This is also known as the computational basis.

The TFIM begets a phase transition with a critical point in the vicinity of  $|h/J| = 1$  [27]. To explore this phase diagram we will find solutions of the ground state with fixed  $J = -1$  and varying  $h$  over the range of 0–3.

### 2.2. Parameterized quantum states

The energy spectrum, dynamics, and a plethora of other observables related to a quantum system are obtainable through the wavefunction,  $\psi$ . However, Nature does not grant one access to the wavefunction,



**Figure 1.** A pictorial representation of the restricted Boltzmann machine (RBM) with  $M$  hidden neurons and  $N$  visible neurons encoding the projection in the  $\sigma^z$  basis.

nor would it be computationally feasible to store these data for a system larger than a few tens of qubits. The latter complication is due to the exponential increase in the number eigenstates with system size. In this work we employ three different parameterized quantum states that all circumvent these limitations.

Parameterized quantum states can offer two key features. Firstly, by representing the wavefunction as a parameterized function,  $\psi_{\theta}$ , we can learn a representation of the true wavefunction by learning the parameters  $\theta$ . Secondly, by restricting the representation such that the number of parameters grows polynomially (rather than exponentially) with the system size, we achieve an efficient representation that makes simulations feasible for larger systems.

### 2.2.1. NQS

Beginning with the work of Carleo and Troyer, neural-network-based methods were introduced to act as physics-agnostic, parameterized quantum states known as NQS [8]. NQS have shown to be useful in a variety of problems, such as finding the ground state in different quantum systems [13, 28, 29], evolving quantum states through time [8, 30], and simulating quantum circuits [31]. There exist different architectures for the NQS [8–13], for the purpose of this paper, however, we restrict ourselves to a simple case of NQS, namely the restricted Boltzmann machine (RBM) depicted in figure 1.

The RBM has the simple structure of two neural network layers with no intra-layer connections but full inter-layer connectivity: one visible layer consisting of  $N$  units and one hidden layer of  $M$  units.  $N$  must be equal to the number of qubits/particles in the system, whereas there is no formal restriction on  $M$ ; the ratio  $M/N$ , referred to as  $\alpha$ , is often used to convey the size of an RBM. The number of parameters in the RBM is given by  $N + \alpha N + \alpha N^2$ , and thus for fixed  $\alpha$  this scales polynomially with the system size; however, this does not guarantee that the desired state can be represented by this network, in which case the true ground state may still require a number of parameters larger than this. The wavefunction amplitude given by the RBM ansatz is

$$\psi_{\theta}(\mathbf{s}) = \sum_{\mathbf{h}} e^{\sum_j a_j \sigma_j + \sum_i b_i h_i + \sum_{ij} W_{ij} h_i \sigma_j}, \quad (2)$$

where  $\mathbf{s} = (\sigma_1, \sigma_2, \dots, \sigma_N)$  such that  $\sigma_i = \pm 1$  is the spin configuration in  $z$  basis;  $\mathbf{h} = (h_1, h_2, \dots, h_M)$ , with  $h_i = \pm 1$ , denotes the  $M$  hidden spin variables; and the parameters to be optimized are  $\theta = \{\mathbf{a}, \mathbf{b}, \mathbf{W}\}$ . The full wavefunction can be constructed from

$$|\psi_{\theta}\rangle = \sum_{\mathbf{s}} \psi_{\theta}(\mathbf{s}) |\mathbf{s}\rangle. \quad (3)$$

With the specific form of the parameterized quantum state in-hand, what remains is to train this model such that the wavefunction that is represented captures the behavior of the ground state of the Hamiltonian being considered. To do this we will consider the expectation value of the Hamiltonian with respect to the NQS wavefunction,  $\langle H \rangle_{\psi_{\theta}} := \langle \psi_{\theta} | H | \psi_{\theta} \rangle$ . In line with the variational principle [32], this quantity is bounded from below by the ground state energy and will equal that bound when the NQS represents the ground state wavefunction.

One of the benefits of this parameterized quantum state is that, for a fixed  $\alpha$ , the number of free parameters that characterize the wavefunction in equation (2) is polynomial with the number of qubits. This work would be undone if we required the full wavefunction or access to the entire Hilbert space to compute expectation values.

To maintain the efficient, computationally feasible computations of expectation values we use Monte Carlo (MC) sampling alongside so-called *local estimators*. The combination of MC sampling and a variational ansatz is known as variational MC (VMC). An excellent outline of the computational efficiencies included in VMC is outlined in [33]. Following the conventions of that work, we can reduce the calculation

of the expectation value from a sum over an exponentially large Hilbert space to a statistical expectation of local estimators. Thus reducing

$$\langle \hat{O} \rangle = \frac{\langle \psi_{\theta} | \hat{O} | \psi_{\theta} \rangle}{\langle \psi_{\theta} | \psi_{\theta} \rangle}, \quad (4)$$

to the more tractable sum of  $\langle \hat{O} \rangle = \langle O_{\text{loc}} \rangle_P$ . The local estimator and probability distribution are defined as

$$O_{\text{loc}}(\mathbf{s}) = \sum_{\mathbf{s}'} \langle \mathbf{s} | \hat{O} | \mathbf{s}' \rangle \frac{\langle \mathbf{s}' | \psi_{\theta} \rangle}{\langle \mathbf{s} | \psi_{\theta} \rangle}, \quad (5)$$

and

$$P(\mathbf{s}) = \frac{|\langle \psi_{\theta} | \mathbf{s} \rangle|^2}{\sum_{\mathbf{s}'} |\langle \psi_{\theta} | \mathbf{s}' \rangle|^2}, \quad (6)$$

respectively. The VMC procedure generates an ensemble of configurations distributed under  $P(\mathbf{s})$ .

The computation of equation (5) is only efficient if each local estimator contains, at most, a polynomial number (relative to the total number of qubits in the system) of connected states,  $\mathbf{s}'$ : equivalent to only having a polynomial number of distinct matrix elements,  $\langle \mathbf{s} | \hat{O} | \mathbf{s}' \rangle$ . As can be seen from equation (1), the Hamiltonian of the TFIM contains only terms that involve one or two qubits, and thus there are only two or four connected states for each term in the Hamiltonian, regardless of the system size. Therefore, the condition of only a polynomial number of neighboring states is met.

Using this efficient computation for expectation values, we can compute the energy of the NQS for a given set of parameters,  $\langle H \rangle_{\psi_{\theta}}$ . As the NQS is a differentiable function, we can also compute the gradient of this energy with respect to each of the network's parameters. As is the common paradigm for training neural networks, the cost function (in this case the energy) and its derivatives are used to tune the parameters until the cost function is minimized. And, as previously mentioned, this will occur when the NQS represents the ground state wavefunction. For this work we used the Python package NetKet [34, 35] which implements the efficient computations laid out here.

For small systems, the entire Hilbert space is not prohibitively large. Therefore, for comparison, we can compute the exact expectation value,  $\langle H \rangle_{\psi_{\theta}}$ , rather than having to rely on MC sampling. The effects of exact against MC sampling is mentioned in appendix A alongside further details of the RBM architecture. Unless stated otherwise, all RBM data shown in this paper was generated from a network with  $\alpha = 5$ .

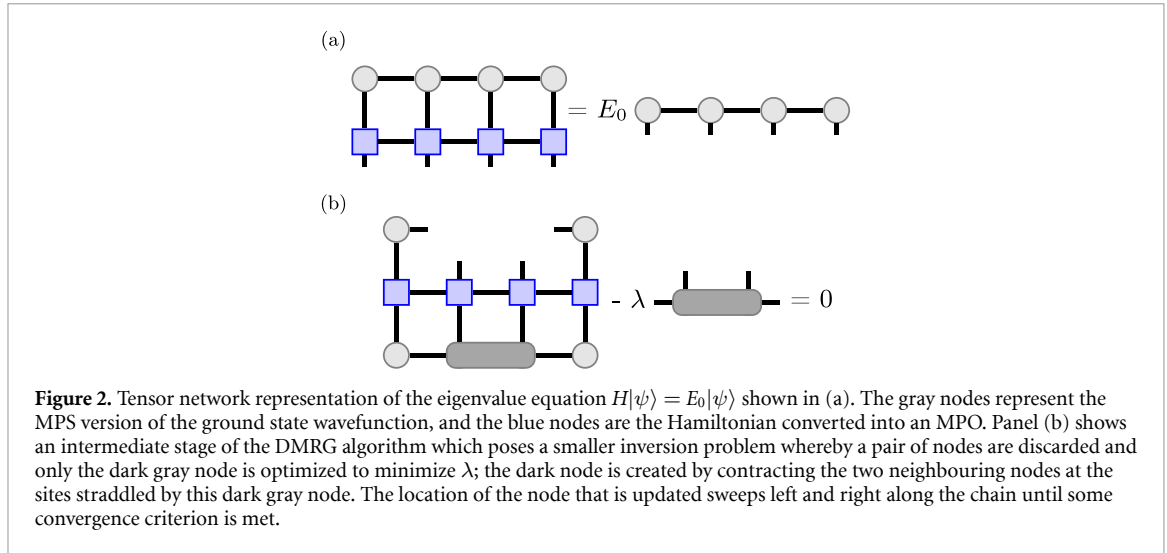
### 2.2.2. DMRG

A well established, classical approach to finding the smallest eigenvalue of a given Hamiltonian was introduced by White in [36], and is known as DMRG. DMRG is an iterative process of solving smaller matrix inversion problems to eventually invert a larger matrix. If the inversion problem can be written in terms of MPSs and matrix product operators (MPOs) then there is an efficient implementation of DMRG [6].

A generic tensor can be represented as an MPS or MPO. However, tensors representing low-entanglement wavefunctions, or local Hamiltonians, can be represented as low-rank MPSs and MPOs respectively. The rank of the MPS is called the bond dimension,  $D$ , and can be adjusted during the DMRG procedure to balance between computational efficiency (small  $D$ ) and allowing for high entanglement (large  $D$ ). It is common practice to limit the bond dimension to achieve a more efficient representation of the wavefunction [37]. For this work, however, we allowed the DMRG algorithm to grow the bond dimension to as large a value as 100 if there was a significant decrease in energy.

The parameters that get updated in the DMRG algorithm are the elements of the MPS at each node. The way that these parameters are optimized is as follows:

1. An initial, random, set of values is chosen for every element in the MPS  $|\psi\rangle$ .
2. The MPO representing the Hamiltonian is contracted from both sides by the MPS. This is equivalent to the equation  $\langle \psi | H | \psi \rangle = E_0$ .
3. A pair of neighboring nodes are chosen and are contracted into a larger tensor; the result is no longer an MPS (this is shown in figure 2(b)). An iterative method is then performed in this subspace (subspace because it is still not the entire Hilbert space) to compute a new tensor that replaces the selected pair but lowering the eigenvalue of the MPS.
4. The altered tensor is then broken back into two tensors using singular value decomposition to return  $|\psi\rangle$  to an MPS whilst maintaining the lower eigenvalue achieved in step 3.



5. The next neighboring pair of links are chosen and steps 3 and 4 are repeated, systematically moving along the chain of the MPS, until a predefined convergence criterion is met.

Upon the termination of the DMRG algorithm, one is left with an MPS that represents the ground state wavefunction, and the ground state energy. For this work we used the implementation provided by ITensor [38, 39].

### 2.2.3. VQE

An alternate approach to classically approximating the wavefunction is to use a parameterized quantum circuit to prepare a trial wavefunction and then use classical optimization to vary the parameters until a desired wavefunction is reproduced. When this process is applied to finding the ground state wavefunction it is known as a VQE [40].

The formulation of VQE spearheaded the development of a rich field of variational quantum algorithms (VQAs) [14–17]. We will first outline the VQE algorithm and then explore some of the challenges related to its implementation.

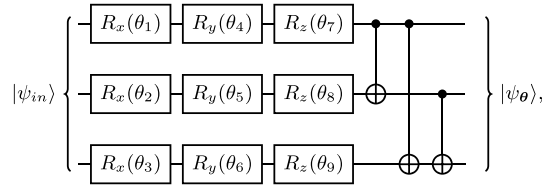
One begins the VQE procedure with an easily prepared initial state wavefunction,  $|\psi_{\text{in}}\rangle$ , to which one applies a sequence of unitary operators, or *gates*, that comprise the parameterized quantum circuit. The output of the circuit will be a parameterized wavefunction  $|\psi_{\theta}\rangle$ . Thus, with the quantum circuit denoted by the unitary matrix  $\mathcal{U}(\theta)$ , the parameterized wavefunction generated by the quantum circuit reads

$$|\psi_{\theta}\rangle = \mathcal{U}(\theta) |\psi_{\text{in}}\rangle. \quad (7)$$

An example of a parameterized quantum circuit is depicted in figure 3; the circuit used for this work is four contiguous layers of the circuit shown, repeated in series. The values of the parameterized gates are varied to find the minimum of the expectation value of the Hamiltonian, equation (1), with respect to  $|\psi_{\theta}\rangle$ , denoted  $\langle H \rangle_{\psi_{\theta}}$  as in section 2.2.1. For this work, given the system sizes simulated, a classical simulation of the quantum circuit was performed using PennyLane [42]. There are, however, implementations on quantum hardware [43–48].

Given that this work presents classical simulations of quantum circuits, access is granted to the full wavefunction rather than just projective measurements. This limits the size of the system that can be simulated but the effects of using the full wavefunction rather than projective measurements to estimate  $\langle H \rangle_{\psi_{\theta}}$  is presented in appendix B.

Considering, still, the case of a classical simulation, free from hardware-related noise and imperfections, there exist fundamental limitations in the VQE algorithm. Firstly, common to all variational algorithms that require a predetermined ansatz, including those in this section, is that there is no guarantee that the ansatz in question (the quantum circuit) is expressive enough to represent the function it hopes to capture (the wavefunction) using a polynomial number of parameters. This is especially true for strongly correlated quantum systems [49]. Broadly speaking there are two approaches to circuit design for VQEs: problem-specific ansätze and hardware-efficient ansätze. In the former regime, one constructs the ansatz from elements that one hopes captures known behaviors of the solution. The latter attempts to create a more general, easier to implement on quantum hardware, ansatz built from simple gates; this is where the



**Figure 3.** A single layer of the VQE ansatz that evolves an initial state,  $|\psi_{in}\rangle$  to a parameterized state  $|\psi_{\theta}\rangle = \mathcal{U}(\theta)|\psi_{in}\rangle$ . The output state is then used as a trial for the ground state and the expectation value of the Hamiltonian is computed. The parameters,  $\theta$ , which are angles of the rotation gates, are varied to minimize this expectation value. Diagram created using [41].

ansatz for this work resides. Neither approach, however, is guaranteed to produce an ansatz that can represent the desired function, and thus the algorithm can converge on a suboptimal solution. Secondly, VQAs have been shown to be especially prone to so-called *barren plateaus* [50–53]. The phenomenon of a barren plateau is when the gradient of the loss function with respect to the variational parameters vanishes exponentially with the number of qubits in the system. This means that training the variational parameters becomes exceptionally difficult for even modest sized systems—even with gradient-free optimizers [54]—as the true minimum becomes exponentially hard to find. Thirdly, and even affecting small systems, the loss landscape for VQEs often contains many local minima, and therefore even without vanishing gradients there can be inherent difficulties training these quantum circuits. The alleviation of these difficulties is ongoing, with a good survey available [55], the references therein, as well as [56–59].

### 2.3. Non-stabilizerness

The non-stabilizerness, also commonly called magic, is known to be the resource for quantum computation [60]. It is known, from the Gottesman–Knill theorem [18], that Clifford circuits (that is, circuits comprised entirely of gates in the Clifford group) can be efficiently simulated on classical hardware. Therefore, in order to achieve a quantum advantage we need to go beyond the Clifford group, which means adding non-Clifford operations. Adding non-Clifford operations means increasing non-stabilizerness. The non-stabilizerness is measured by different methods, such as the robustness of magic [61–63], the min-relative entropy [64] and stabilizer entropies [20, 65].

For the purpose of this work, we use the 2-Renyi stabilizer entropy,  $M_2$ , since it is easily calculable for small systems and it has been shown to be a magic monotone [66, 67]. The 2-Renyi stabilizer entropy for pure states is defined as

$$M_2(|\psi\rangle) = -\log \sum_{P \in \mathcal{P}_N} \frac{\langle \psi | P | \psi \rangle^4}{2^N}, \quad (8)$$

where  $\mathcal{P}_N$  is the set of all  $N$ -qubit Pauli strings and  $|\psi\rangle$  can either be a parameterized representation of the quantum state, or the exact result found through ED.

Recent studies have also explored the non-stabilizerness in random circuits [24, 68], and many-body quantum systems such as the TFIM [21, 69–71], the Potts model [26, 72], chaotic models [73], and the class of models known as generalized Rokhsar–Kivelson systems [72]. It has also been shown that in special cases there can be a significant reduction of equation (8) into a form that scales only polynomially in the system size [74].

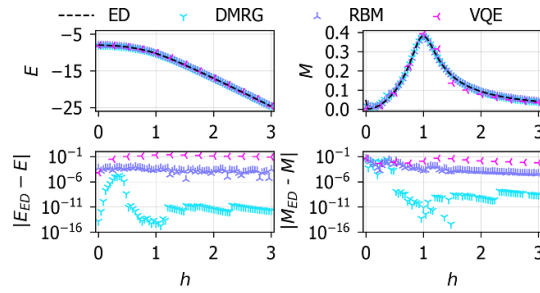
We are motivated by the observation that classical simulatability from the Clifford group point of view does not enter generic formulation of quantum or classical variational problems. We use equation (8) to evaluate the non-stabilizerness of exact transverse-field Ising (TFI) ground states as well as its approximations determined using the methods described above, and compare the performance of these methods from the non-stabilizerness perspective across the phase diagram.

## 3. Results

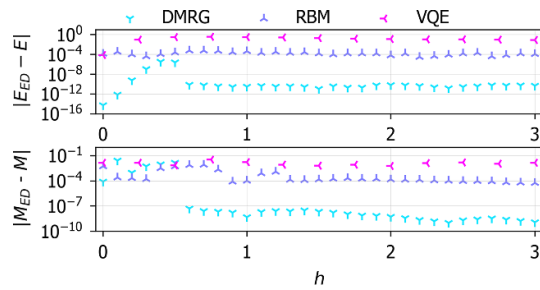
### 3.1. Quantum and classical model accuracy

The simplest comparison between the three aforementioned methods is to compare them each to the results obtained from ED.

Figure 4 shows the energy and magic obtained from ED and all three variational approaches applied to an 8 qubit TFI system. Both the values themselves (top) and the accuracies with respect to the ED result (bottom) are shown. From the left-hand panels of this figure it is clear that there is a hierarchy in terms of energy accuracy: DMRG is more accurate than the RBM, which is in-turn more accurate than the VQE.



**Figure 4.** (top) The energy (left) and magic (right) of the ground state wavefunction for the 8 qubit transverse field Ising model at a range of transverse-field strengths,  $h$ . (bottom) The accuracy of the estimates of energy (left) and magic (right) from the three methods presented in section 2.2 relative to the result from exact diagonalization (ED). Each datum for the RBM and VQE is the mean of 10 repeats and the statistical errors are omitted for clarity; see figure 7 for error estimates.



**Figure 5.** The accuracy of the estimates of energy (top) and magic (bottom) from the three methods presented in section 2.2 relative to the result from exact diagonalization (ED). These results are for a 12 qubit system, shown at a range of transverse-field strengths,  $h$ . Each datum for the RBM and VQE is the mean of 10 repeats and the statistical errors are omitted for clarity; see figure 7 for error estimates.

From the right-hand panels, however, the magic accuracy does not quite follow this hierarchy; beyond  $h = 1$  the accuracies follow that of the energy, but approaching criticality from below there are signs that although the DMRG and RBM converged on states with more accurate energy than the VQE, the magic accuracies of all three are similar. It should be noted that in this region of the phase diagram the ground state of the TFIM is known to be degenerate (or near-degenerate, depending on finite system size effects and the specific value of  $h$ ). The magic of the states within the degenerate eigenspace can differ greatly and thus the accuracy in energy does not constrain the state to have a similar magic to the ED result. One possible way to circumvent this would be to use ansätze that are aware of the symmetry that leads to this degeneracy. We will return to this in section 3.2 as this highlights an important case of when energy accuracy alone is not sufficient for assessing the quality of the approximate ground state.

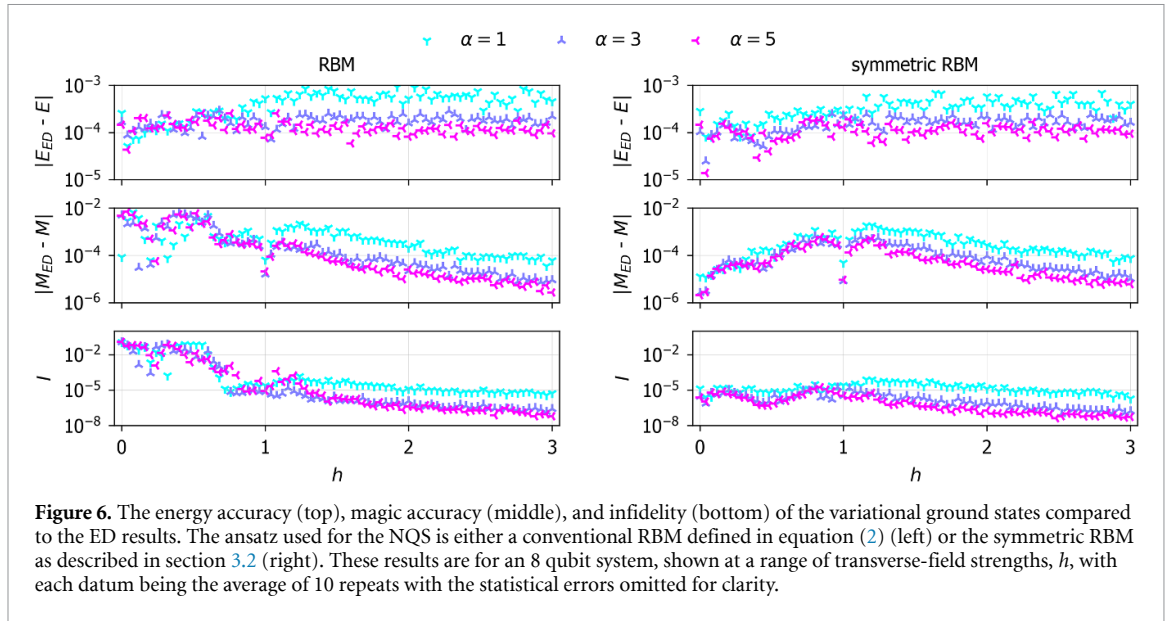
Figure 5 shows the energy and magic accuracy of the same methods for a 12 qubit system. Due to the increased computational cost of simulating a larger system, the  $h$  spacing is increased. Once again the accuracy in energy follows the ordering noted in figure 4. Also in a similar fashion to the 8 qubit simulation, the accuracy of the magic does not follow this trend over the entire phase diagram.

From figures 4 and 5 we can conclude that energy accuracy and magic accuracy appear strongly correlated for most of the phase diagram, we can also say that, for these implementations, DMRG outperforms RBM which outperforms VQE when it comes to energy accuracy. In appendix B we present the energy and magic accuracy computed for a 4 qubit system. In this smaller setting the VQE and RBM perform similarly. Despite the smaller system size, the same pattern in energy and magic accuracies can be seen: they appear correlated except for the low  $h$  region.

For clarity of the figures in this section, the statistical fluctuations of the data were not shown. In section 3.3 will explore the statistical uncertainties to measure the robustness of these methods and further probe the correlations between the energy and magic of the ground state solutions found.

### 3.2. Non-stabilizerness and symmetry

By and large, in the figures from section 3.1, where we selected best performing hyperparameter-configuration for MPS, VQE and RBM, the accuracy of the energy appears to be directly linked to the accuracy of the magic. One may expect that by minimizing the energy towards the ground state then one ends up nearby in the Hilbert space and thus the magic of the solution is likely similar. However, the exception of



**Figure 6.** The energy accuracy (top), magic accuracy (middle), and infidelity (bottom) of the variational ground states compared to the ED results. The ansatz used for the NQS is either a conventional RBM defined in equation (2) (left) or the symmetric RBM as described in section 3.2 (right). These results are for an 8 qubit system, shown at a range of transverse-field strengths,  $h$ , with each datum being the average of 10 repeats with the statistical errors omitted for clarity.

the degenerate ground states at low  $h$  was noted. This section will explore the interplay between the accuracy of energy and magic further, paying careful attention to the degeneracy of the solutions in the small  $h$  region.

The Hamiltonian of the TFIM exhibits a global  $\mathbb{Z}_2$  symmetry, meaning that the energy of two configurations that differ by flipping every single spin is the same. The ground states of this Hamiltonian do not have to respect this symmetry to still achieve the smallest possible energy. This leads to a range of different, but energetically equivalent, variational ground states. Importantly, the magic of these degenerate states can differ greatly. To give an indication of this, consider the  $h = 0$  case and any superposition of the all up and all down states (all of which have the same energy): an equal superposition has zero magic as it can be prepared using only Clifford operations acting on the all up state, however, for other unequal superpositions it is not guaranteed that one could prepare this state using only Clifford operations and thus it can have non-zero magic. To circumvent this, one can explicitly encode the global  $\mathbb{Z}_2$  symmetry into the ansatz. We will explore this in the context of the RBM.

For such a small, discrete symmetry group, one can make a symmetric (or  $\mathbb{Z}_2$ -invariant) RBM simply by feeding both the configuration  $\mathbf{s}$  and  $-\mathbf{s}$  through the RBM defined in equation (2) and taking the mean of each amplitude as the amplitude of the configuration  $\mathbf{s}$  (and thus also of  $-\mathbf{s}$ ). All other details of the training and RBM architecture are unchanged.

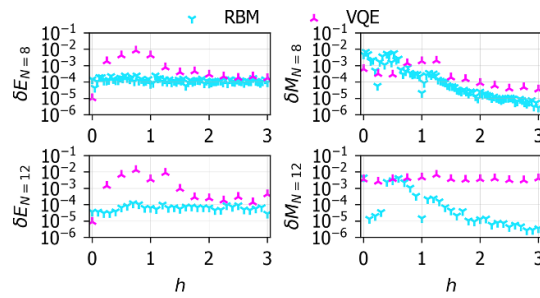
To explore the intuition that energy accuracy alone is a good probe of whether or not the approximate ground state is similar to the true ground state, we compare it with two other figures of merit. Firstly, as in the other sections of this work, we compare with the magic accuracy, but then we also show the infidelity of the approximate ground state with the ground state found through ED. We define the infidelity of the normalized approximate state  $|\psi_\theta\rangle$  with the normalized ED result  $|\psi_{ED}\rangle$  as

$$I = 1 - |\langle\psi_\theta|\psi_{ED}\rangle|^2. \quad (9)$$

Figure 6 shows the energy accuracy, magic accuracy, and infidelity of the RBM and symmetric RBM ground states with respect to the ED results. The data are shown for three sizes of RBM, denoted by  $\alpha = 1, 3$ , and 5, to explore if the increase in the number of parameters allows the network to learn this symmetry without the need for the explicit construction of the symmetric RBM.

From figure 6 it is clear that the energy accuracies of the RBM and the symmetric RBM are similar. However, the effect on the magic accuracy and infidelity in the  $0 \leq h \leq 1$  range is stark: the symmetric RBM outperforms the conventional RBM by up to five orders of magnitude. There does not seem to be much change to the rest of the  $h$ -range considered. Importantly, though, is that this five orders of magnitude improvement, indicating a significantly better ground state solution, is completely imperceptible when considering the energy accuracy alone.

The effect of increasing the size of the RBM, through increasing  $\alpha$ , is one of a mild increase in quality of all metrics, with the exception of the small  $h$  region for the conventional RBM. Interestingly, this is the area that is most plagued by the effect of the degenerate ground states, and thus the extra parameters do not account for the symmetry. This means that just using a larger neural network was not enough to overcome



**Figure 7.** The statistical errors of the energy (left) and magic (right) against the transverse-field strength,  $h$ , across the 10 repeats of the ground state optimization procedure using an RBM and VQE. Data are shown for 8 (top) and 12 (bottom) qubits, with the corresponding mean values reported in figures 4 and 5 respectively.

the fact that energy accuracy alone is insufficient for the ansatz to fully approximate the non-stabilizerness of the ground state.

### 3.3. Fluctuations in quantum and classical solutions

One theme that this work aims to shed some light onto is the representativeness of a ground state found through energy minimization alone. Magic is used as a second axis to ascertain how well the approximate state represents the true ground state. In this section we will explore how similar in magic repeated energy minimizations are, probing the landscape around, what one hopes to be, the energy minimum containing the true ground state.

Figure 7 shows the statistical error from 10 repeated ground state searches, in energy (left) and magic (right), for 8 qubit (top) and 12 qubit (bottom) simulations. The comparison is shown only for the VQE and RBM methods. From this figure it can be seen that the fluctuations in energy are almost always smaller for the RBM compared to the VQE. The exception to this is the  $h = 0$  point; the ground state of the Hamiltonian here is, in fact, the initial state fed into the VQE and thus this is perhaps not surprising. The fact that the VQE can, in very high and low field, reach almost the level of consistency of the RBM yet it cannot minimize the energy as well suggests that perhaps this is a limitation of the expressivity of the ansatz.

In section 3.2 we already noted the difficulties with an RBM that is not aware of the symmetries of the system. This is reflected here in the fluctuations of the magic; interestingly this is less pronounced in the 12 qubit simulation. Those data notwithstanding, the fact that the fluctuations of the magic are very small does suggest that each approximate ground state, when near to the minimum in energy, is also in a region with similar magic.

## 4. Discussion and conclusions

In this work we set up the framework to reconcile classical simulatability notions of quantum resource theory with the formalism of classical variational techniques. We used non-stabilizerness, a quantity that measures how far is a given state from being classically simulatable in a sense of Gottesmann–Knill theorem, as a figure of merit for quality of classical variational approximation of a quantum state. Specifically, we assessed non-stabilizerness expressivity for three qualitatively different types of variational ansätze: tensor networks, neural networks and variational quantum circuits.

We found that when comparing the best model (in terms of energy performance) for each method with ED, on average, there is a general trend for better energy to correspond to better approximation in non-stabilizerness. This is an encouraging observation that suggests that the state reconstructed as a result of classical variational procedure has a complexity structure that represents the exact quantum solution reasonably well. At the same time, we immediately noted that this correspondence is far from straightforward and universal. For example, an RBM constructed to share the symmetry of the Hamiltonian led to ground states with up to four orders of magnitude higher accuracy in non-stabilizerness than a conventional RBM, with no change in the energy accuracy.

We hope this work will be a stepping stone towards starting a closer dialogue between quantum computing and classical variational perspectives on solving the quantum many-body problem, and more specifically, on how the quality of the results is assessed. We set up a system of benchmarks for small system sizes ( $N = 8, 12$ ) and noted a number of similarities and differences between accuracies in non-stabilizerness and in energy. The possible next steps include obtaining large scale benchmarks for each of the methods we tested here. The current limitation of system size is the calculation of the non-stabilizerness, however, the

first impressive step towards this calculation at scale has already been taken in [22] for an MPS representation of a wavefunction. It would be interesting to see how the smaller system observations translate into a large scale benchmark. Further ahead, one could think of how to embed non-stabilizerness optimization iterations into energy-based quantum and classical variational models.

### Data availability statement

The data that support the findings of this study are openly available at the following URL/DOI: [10.5281/zenodo.13759650](https://zenodo.org/record/13759650) [75] and <https://gitlab.com/QMAI/papers/quantumresourcesml> [76].

### Acknowledgments

We are thankful for instructive discussions on DMRG with B M La Rivière, J Soto Garcia, and P S Tarabunga. We also thank A Valenti, J Nys, and S Khaleefah for fruitful conversations regarding NQS. This work is part of the project Engineered Topological Quantum Networks (Project No. VI.Veni.212.278) of the research program NWO Talent Programme Veni Science domain 2021 which is financed by the Dutch Research Council (NWO). This publication is also part of the project Optimal Digital-Analog Quantum Circuits with File Number NGF.1582.22.026 of the research programme NGF-Quantum Delta NL 2022 which is (partly) financed by the Dutch Research Council (NWO) and the Dutch National Growth Fund initiative Quantum Delta NL.

### Author contributions

EG designed the project with input from TS and AA. TS and AA jointly performed neural network and ED simulations. TS and BC performed the VQE simulations. TS performed the tensor network simulations. AA provided the non-stabilizerness calculation. TS, AA, BC, and EG analyzed and interpreted the data. TS, EG, and AA co-wrote the manuscript. EG supervised the project.

### Appendix A. NQS architecture

There are a vast number of hyperparameters and architectural configurations that one can tweak to alter the training of the NQS. To arrive at the specific RBM presented in figures 4 and 5 we explored: the size of the ansatz, using the entire Hilbert space for the computation of expectation values, and the inclusion of the stochastic reconfiguration (SR) preconditioner as part of the parameter optimization routine. Other avenues of exploration could include introducing a representation learning step to map the configurations from projective measurements to continuous vectors [11] or other neural network architectures [8–13], but these lie outside the scope of this work. This appendix will also cover the stopping criterion used to estimate when the ground state is approximately reached.

The effects of a larger RBM were shown in figure 6, from this it can be seen that there is a mild improvement in all metrics with increased model size. The drawback, however, is in the time taken to train the larger model: more parameters leads to increased training time. Given that the improvement is larger from  $\alpha = 1$  to  $\alpha = 3$  than it is between  $\alpha = 3$  and  $\alpha = 5$ , alongside the increased training time, we chose to stop at  $\alpha = 5$ .

In section 2.2.1 we mentioned the approximations made as part of the VMC procedure, notably, as part of the computation of the expectation value, the reduction of the sum over the entire Hilbert space (equation (4)) to the MC estimate sampled from equation (6). We tested the energy and magic accuracy for an 8 qubit system using both the full expectation value and the MC estimate to train the RBMs. We found that, across the entire range of  $h$ , the fluctuations between repeated optimizations of the exact expectation solutions were lower than the MC sampled counterparts, and the energy accuracies were smoother as a function of  $h$  for the exact expectation method, too. Both effects were, however, mild, and as the exact expectation is much more computationally expensive (it scales exponentially with the system size), we chose to remain with the MC sampling.

Finally, for the configuration of the RBM, we tested the inclusion of the SR preconditioner. Introduced to VMC in [77, 78], SR alters the gradient obtained from an optimizer in a way that accounts for the curvature of the optimization landscape that is caused only by the ansatz. This can be seen as an application of what is known as the natural gradient in the wider machine learning community [79, 80]. Put simply, the inclusion of SR lead to better energy and magic accuracies of the ground states found.

During the optimization of the RBM parameters one must establish a point at which to stop. For this work we considered the training to be done when the relative change in energy did not improve by more than

$10^{-7}$  for 500 consecutive epochs. This can be easily implemented using the `callbacks.EarlyStopping` function from `NetKet`.

We used the ADAM [81] optimizer and 1000 (5000) samples for the MC estimates of expectation values for the 8 (12) qubit systems for all of the RBM results presented in this work.

## Appendix B. VQE architecture

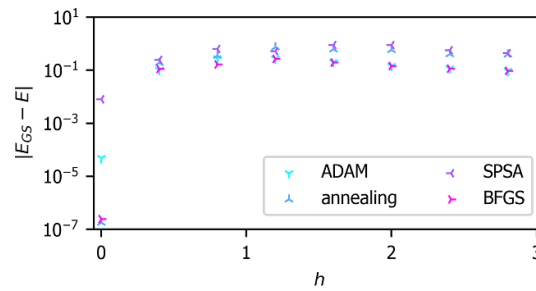
Whilst, in theory, VQE is a Hamiltonian-agnostic and flexible algorithm, there are optimizations that can be made in its implementation. In a similar fashion to appendix A, this section will explore the choices made when exploring the implementation of the VQE: the number of layers in the ansatz, the method for computing the expectation value of the Hamiltonian, and the optimizer. It will also elucidate the stopping criteria used.

Figure 3 shows a single layer of the VQE ansatz used in this work. We explored using up to four layers of this as the ansatz and found that the energy and magic accuracies improved with more layers, however, interestingly the performance of one and two layers were similar, as were three and four layers. The training time increases with each added layer and therefore we chose not to go beyond four layers, using four layers as the final ansatz.

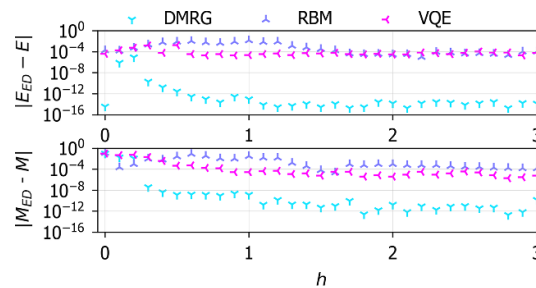
Given that this VQE implementation was performed using a classical simulation of a quantum circuit, we could compute the expectation value of the Hamiltonian two ways: either using projective measurements of the final state or by exact computation (exploiting the full wavefunction of the final state that is only easily accessible in classical simulations and not in real quantum experiments). We tested using both methods to compute the expectation value of the Hamiltonian during training and found no significant difference between the two sets of ground states found, neither in energy nor magic. However, the time taken for the ground state search when using measured expectations was up to three orders of magnitude longer than when using exact expectations (up from  $10^2$  to  $10^5$  s). This could be caused by several things, the two that seem most pertinent to mention are related to the gradient of the expectation of the Hamiltonian with respect to the circuit parameters. Firstly, as the projective measurement step of a quantum circuit is not differentiable, the gradient was estimated using the parameter-shift rule [82, 83] which requires multiple evaluations of the circuit; this is in contrast to the exact case which only required a single execution to compute the value and gradient of the expectation value. Secondly, as an estimation of the expectation value can only be as accurate, or less accurate, than the exact value, the estimations of the gradient of the expectation value with respect to the circuit angles can only be equal to, or less accurate, than the exact case. Given this, then, it is natural to assume that with likely many sub-optimal estimations of the gradient, the optimization procedure would take longer and require more evaluations of the quantum circuit. Both of these factors will contribute to the increased runtime of the VQE procedure. Given the lack of clear improvement by either approach, and as this paper does not extend to system sizes beyond the range of what is classically simulatable, we chose to use the exact expectation value for the VQE procedure for significantly more efficient use of computational resources.

To assess the point at which we stop training the quantum circuit we implemented a two-step stopping criteria. Firstly, to consider a single instance of a ground state to have been found we require that three consecutive epochs give the same energy to within  $10^{-6}$ , but then we also require that if the whole VQE is reinitialised then the next instance must not differ from the previous by more than a relative change of  $10^{-4}$  in the energy. The training of the quantum circuit was less stable and led to larger fluctuations in energy than for the RBM, and thus the stopping criterion for the single instance is less strict, but the problem of local minima in the search space was more common here and thus the second part of the stopping criteria was added.

There is a large community interest in alleviating the problem of barren plateaus in quantum machine learning applications. Specifically to VQEs, the authors in [84] explored the effect of using three different optimizers: Broyden–Fletcher–Goldfarb–Shanno (BFGS) [85–88], ADAM [81], and the natural gradient [79, 80] to avoid the problem of local minima in the variational landscape. Here we consider the energy accuracy of an 8 qubit energy minimization of a VQE for four different optimizers: ADAM, BFGS, SPSA [89], and simulated annealing [90]. This is a mix of gradient- and non-gradient-based optimizers: BFGS and ADAM were chosen to cover those of interest for avoiding local minima, SPSA for its utility with applications on quantum hardware [45], and annealing as a second non-gradient based approach as this avoids the use of the parameter-shift rule. The ansatz chosen for these optimizations is the three-layer version of the ansatz depicted in figure 3. The number of layers was reduced to three because the annealing optimizer consistently failed to converge for a four-layer ansatz. Figure 8 shows the energy accuracies achieved from 10 repeated energy minimizations of the TFI Hamiltonian for each of the chosen optimizers. From this we can see that both BFGS and ADAM perform the best but there is a very quick degradation of performance with increased



**Figure 8.** Energy accuracy against transverse-field strength for a three-layer VQE ansatz, shown for a range of optimizers used in the energy minimization procedure. All data are for an 8 qubit simulation.



**Figure 9.** The energy accuracy (top) and magic accuracy (bottom) against transverse-field strength for a 4 qubit simulation. Each datum is the average over 10 minimizations.

transverse field strength,  $h$ . As concluded by the authors of [84], both ADAM and BFGS struggle to scale to systems with a large number of parameters; for the ansatz used here there are 72 free parameters, much larger than the upper limit of 42 that they propose. BFGS especially struggles to scale to larger systems. For this reason, all other VQE data shown in this work uses ADAM as the optimizer. As previously mentioned, reducing the number of layers in the ansatz led to worse energy accuracies, so we will now explore the other way to reduce the number of parameters: a smaller system size.

To explore a regime where the VQE optimization maintains a flexible and expressive ansatz, but with fewer variational parameters, we will consider a 4 qubit TFI system. Exactly as done in section 3.1, figure 9 shows the accuracies in energy and magic of the VQE, RBM, and DMRG algorithms. Unlike in figures 4 and 5, however, the energy accuracy of the VQE is better than that of the RBM. This shows that the VQE is able to achieve accurate ground state energies, but figures 4 and 5 show that this does not easily scale to larger systems with our current implementation. It should also be noted that in figure 9 we can still see that the accuracy in energy does not always correlate with accuracy in magic.

## ORCID iDs

Thomas Spriggs  <https://orcid.org/0000-0001-7513-6437>

Arash Ahmadi  <https://orcid.org/0009-0001-9655-8345>

Bokai Chen  <https://orcid.org/0009-0002-1726-3950>

Eliska Greplova  <https://orcid.org/0000-0003-3369-2544>

## References

- [1] Jastrow R 1955 Many-body problem with strong forces *Phys. Rev.* **98** 1479
- [2] Gutzwiller M C 1963 Effect of correlation on the ferromagnetism of transition metals *Phys. Rev. Lett.* **10** 159
- [3] Gutzwiller M C 1964 Effect of correlation on the ferromagnetism of transition metals *Phys. Rev.* **134** A923
- [4] Gutzwiller M C 1965 Correlation of electrons in a narrow  $s$  band *Phys. Rev.* **137** A1726
- [5] Bardeen J, Cooper L N and Schrieffer J R 1957 Theory of superconductivity *Phys. Rev.* **108** 1175
- [6] Schollwöck U 2011 The density-matrix renormalization group in the age of matrix product states *Ann. Phys., NY* **326** 96
- [7] Perez-Garcia D, Verstraete F, Cirac J I and Wolf M M 2007 PEPS as unique ground states of local Hamiltonians (arXiv:0707.2260)
- [8] Carleo G and Troyer M 2017 Solving the quantum many-body problem with artificial neural networks *Science* **355** 602
- [9] Choo K, Neupert T and Carleo G 2019 Two-dimensional frustrated  $J_1 - J_2$  model studied with neural network quantum states *Phys. Rev. B* **100** 125124
- [10] Hibat-Allah M, Ganahl M, Hayward L E, Melko R G and Carrasquilla J 2020 Recurrent neural network wave functions *Phys. Rev. Res.* **2** 023358

- [11] Viteritti L L, Rende R and Becca F 2023 Transformer variational wave functions for frustrated quantum spin systems *Phys. Rev. Lett.* **130** 236401
- [12] Luo D, Carleo G, Clark B K and Stokes J 2021 Gauge equivariant neural networks for quantum lattice gauge theories *Phys. Rev. Lett.* **127** 276402
- [13] Roth C and MacDonald A H 2021 Group convolutional neural networks improve quantum state accuracy (arXiv:2104.05085)
- [14] Cerezo M et al 2021 Variational quantum algorithms *Nat. Rev. Phys.* **3** 625
- [15] Kremenetski V, Hogg T, Hadfield S, Cotton S and Tubman N 2021 Quantum alternating operator ansatz (QAOA) phase diagrams and applications for quantum chemistry (arXiv:2108.13056)
- [16] Moll N et al 2018 Quantum optimization using variational algorithms on near-term quantum devices *Quantum Sci. Technol.* **3** 030503
- [17] Yao Y-X, Gomes N, Zhang F, Wang C-Z, Ho K-M, Iadecola T and Orth P P 2021 Adaptive variational quantum dynamics simulations *PRX Quantum* **2** 030307
- [18] Gottesman D 1998 The Heisenberg representation of quantum computers (arXiv:quant-ph/9807006)
- [19] Bravyi S and Kitaev A 2005 Universal quantum computation with ideal Clifford gates and noisy ancillas *Phys. Rev. A* **71** 022316
- [20] Haug T, Lee S and Kim M S 2023 Efficient stabilizer entropies for quantum computers (arXiv:2305.19152)
- [21] Haug T and Piroli L 2023 Quantifying nonstabilizerness of matrix product states *Phys. Rev. B* **107** 035148
- [22] Tarabunga P S, Tirrito E, Banuls M C and Dalmonte M 2024 Nonstabilizerness via matrix product states in the Pauli basis (arXiv:2401.16498)
- [23] Lami G and Collura M 2023 Nonstabilizerness via perfect Pauli sampling of matrix product states *Phys. Rev. Lett.* **131** 180401
- [24] Ahmadi A and Greplova E 2024 Quantifying non-stabilizerness via information scrambling *SciPost Phys.* **16** 043
- [25] Frau M, Tarabunga P, Collura M, Dalmonte M and Tirrito E 2024 Nonstabilizerness versus entanglement in matrix product states *Phys. Rev. B* **110** 045101
- [26] White C D, Cao C and Swingle B 2021 Conformal field theories are magical *Phys. Rev. B* **103** 075145
- [27] Sachdev S 2023 *Quantum Phases of Matter* (Cambridge University Press)
- [28] Nomura Y and Imada M 2021 Dirac-type nodal spin liquid revealed by refined quantum many-body solver using neural-network wave function, correlation ratio and level spectroscopy *Phys. Rev. X* **11** 031034
- [29] Astrakhantsev N, Westerhout T, Tiwari A, Choo K, Chen A, Fischer M H, Carleo G and Neupert T 2021 Broken-symmetry ground states of the Heisenberg model on the pyrochlore lattice *Phys. Rev. X* **11** 041021
- [30] Gutiérrez I L and Mendl C B 2022 Real time evolution with neural-network quantum states *Quantum* **6** 627
- [31] Medvidović M and Carleo G 2021 Classical variational simulation of the quantum approximate optimization algorithm *npj Quantum Inf.* **7** 101
- [32] Becca F and Sorella S 2017 *Quantum Monte Carlo Approaches for Correlated Systems* (Cambridge University Press) (<https://doi.org/10.1017/9781316417041>)
- [33] Medvidović M and Moreno J R 2024 Neural-network quantum states for many-body physics (arXiv:2402.11014)
- [34] Carleo G et al 2019 NetKet: a machine learning toolkit for many-body quantum systems *SoftwareX* **10** 100311
- [35] Vicentini F et al 2022 NetKet 3: machine learning toolbox for many-body quantum systems *SciPost Phys. Codebases* **7**
- [36] White S R 1992 Density matrix formulation for quantum renormalization groups *Phys. Rev. Lett.* **69** 2863
- [37] Cirac J I, Pérez-García D, Schuch N and Verstraete F 2021 Matrix product states and projected entangled pair states: concepts, symmetries, theorems *Rev. Mod. Phys.* **93** 045003
- [38] Fishman M, White S R and Stoudenmire E M 2022 The ITensor software library for tensor network calculations *SciPost Phys. Codebases* **4**
- [39] Fishman M, White S R and Stoudenmire E M 2022 Codebase release 0.3 for ITensor *SciPost Phys. Codebases* **4**–r0.3
- [40] McClean J R, Romero J, Babbush R and Aspuru-Guzik A 2016 The theory of variational hybrid quantum-classical algorithms *New J. Phys.* **18** 023023
- [41] Kay A 2023 Tutorial on the Quantikz package (arXiv:1809.03842)
- [42] Bergholm V et al 2018 PennyLane: automatic differentiation of hybrid quantum-classical computations (arXiv:1811.04968)
- [43] Peruzzo A, McClean J, Shadbolt P, Yung M-H, Zhou X-Q, Love P J, Aspuru-Guzik A and O'Brien J L 2014 A variational eigenvalue solver on a photonic quantum processor *Nat. Commun.* **5** 4213
- [44] O'Malley P J et al 2016 Scalable quantum simulation of molecular energies *Phys. Rev. X* **6** 031007
- [45] Kandala A, Mezzacapo A, Temme K, Takita M, Brink M, Chow J M and Gambetta J M 2017 Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets *Nature* **549** 242
- [46] Bentellis A, Matic-Flierl A, Mendl C B and Lorenz J M 2023 Benchmarking the variational quantum eigensolver using different quantum hardware 2023 *IEEE Int. Conf. on Quantum Computing and Engineering (QCE)* vol 1 pp 518–23 (<https://doi.org/10.1109/QCE57702.2023.00065>)
- [47] Atas Y Y, Zhang J, Lewis R, Jahanpour A, Haase J F and Muschik C A 2021 SU(2) hadrons on a quantum computer via a variational approach *Nat. Commun.* **12** 6499
- [48] Zhang J, Ferguson R, Kühn S, Haase J F, Wilson C M, Jansen K and Muschik C A 2023 Simulating gauge theories with variational quantum eigensolvers in superconducting microwave cavities *Quantum* **7** 1148
- [49] Avella A et al 2012 *Strongly Correlated Systems* (Springer)
- [50] McClean J R, Boixo S, Smelyanskiy V N, Babbush R and Neven H 2018 Barren plateaus in quantum neural network training landscapes *Nat. Commun.* **9** 4812
- [51] Wang S, Fontana E, Cerezo M, Sharma K, Sone A, Cincio L and Coles P J 2021 Noise-induced barren plateaus in variational quantum algorithms *Nat. Commun.* **12** 6961
- [52] Cerezo M, Verdon G, Huang H-Y, Cincio L and Coles P J 2022 Challenges and opportunities in quantum machine learning *Nat. Comput. Sci.* **2** 567
- [53] Holmes Z, Sharma K, Cerezo M and Coles P J 2022 Connecting ansatz expressibility to gradient magnitudes and barren plateaus *PRX Quantum* **3** 010313
- [54] Arrasmith A, Cerezo M, Czarnik P, Cincio L and Coles P J 2021 Effect of barren plateaus on gradient-free optimization *Quantum* **5** 558
- [55] Fedorov D A, Peng B, Govind N and Alexeev Y 2022 VQE method: a short survey and recent developments *Mater. Theory* **6** 2
- [56] You X, Chakrabarti S and Wu X 2022 A convergence theory for over-parameterized variational quantum eigensolvers (arXiv:2205.12481)

- [57] Tang H L, Shkolnikov V, Barron G S, Grimsley H R, Mayhall N J, Barnes E and Economou S E 2021 Qubit-ADAPT-VQE: an adaptive algorithm for constructing hardware-efficient ansätze on a quantum processor *PRX Quantum* **2** 020310
- [58] Matsuo A, Suzuki Y, Hamamura I and Yamashita S 2023 Enhancing VQE convergence for optimization problems with problem-specific parameterized quantum circuits *IEICE Trans. Inf. Syst.* **106** 1772
- [59] Anastasiou P G, Chen Y, Mayhall N J, Barnes E and Economou S E 2024 Tetris-ADAPT-VQE: an adaptive algorithm that yields shallower, denser circuit ansätze *Phys. Rev. Res.* **6** 013254
- [60] Chitambar E and Gour G 2019 Quantum resource theories *Rev. Mod. Phys.* **91** 025001
- [61] Veitch V, Mousavian S A H, Gottesman D and Emerson J 2014 The resource theory of stabilizer quantum computation *New J. Phys.* **16** 013009
- [62] Howard M and Campbell E 2017 Application of a resource theory for magic states to fault-tolerant quantum computing *Phys. Rev. Lett.* **118** 090501
- [63] Hamaguchi H, Hamada K and Yoshioka N 2023 Handbook for efficiently quantifying robustness of magic (arXiv:2311.01362)
- [64] Liu Z-W and Winter A 2022 Many-body quantum magic *PRX Quantum* **3** 020333
- [65] Leone L, Oliviero S F and Hamma A 2022 Stabilizer Rényi entropy *Phys. Rev. Lett.* **128** 050402
- [66] Haug T and Piroli L 2023 Stabilizer entropies and nonstabilizerness monotones *Quantum* **7** 1092
- [67] Leone L and Bittel L 2024 Stabilizer entropies are monotones for magic-state resource theory (arXiv:2404.11652)
- [68] Ahmadi A, Helsen J, Karaca C and Greplova E 2024 Mutual information fluctuations and non-stabilizerness in random circuits (arXiv:2408.03831)
- [69] Oliviero S F E, Leone L and Hamma A 2022 Magic-state resource theory for the ground state of the transverse-field Ising model *Phys. Rev. A* **106** 042426
- [70] Odavić J, Haug T, Torre G, Hamma A, Franchini F and Giampaolo S M 2023 Complexity of frustration: a new source of non-local non-stabilizerness *SciPost Phys.* **15** 131
- [71] Tarabunga P S, Tirrito E, Chanda T and Dalmondo M 2023 Many-body magic via Pauli-Markov chains—from criticality to gauge theories *PRX Quantum* **4** 040317
- [72] Tarabunga P S and Castelnovo C 2024 Magic in generalized Rokhsar-Kivelson wavefunctions *Quantum* **8** 1347
- [73] Passarelli G, Lucignano P, Rossini D and Russomanno A 2024 Chaos and magic in the dissipative quantum kicked top (arXiv:2406.16585)
- [74] Passarelli G, Fazio R and Lucignano P 2024 Nonstabilizerness of permutationally invariant systems *Phys. Rev. A* **110** 022436
- [75] Spriggs T, Ahmadi A, Chen B and Greplova E 2024 Quantum resources of quantum and classical variational methods *Zenodo* <https://doi.org/10.5281/zenodo.13759651>
- [76] Spriggs T, Ahmadi A, Chen B and Greplova E 2024 Quantum resources of quantum and classical variational methods (available at: <https://gitlab.com/QMAI/papers/quantumresourcesml>)
- [77] Sorella S 1998 Green function Monte Carlo with stochastic reconfiguration *Phys. Rev. Lett.* **80** 4558
- [78] Sorella S 2005 Wave function optimization in the variational Monte Carlo method *Phys. Rev. B* **71** 241103
- [79] Amari S-I 1998 Natural gradient works efficiently in learning *Neural Comput.* **10** 251
- [80] Stokes J, Izaac J, Killoran N and Carleo G 2020 Quantum natural gradient *Quantum* **4** 269
- [81] Kingma D P and Ba J 2014 Adam: a method for stochastic optimization (arXiv:1412.6980)
- [82] Mitarai K, Negoro M, Kitagawa M and Fujii K 2018 Quantum circuit learning *Phys. Rev. A* **98** 032309
- [83] Schuld M, Bergholm V, Gogolin C, Izaac J and Killoran N 2019 Evaluating analytic gradients on quantum hardware *Phys. Rev. A* **99** 032331
- [84] Wierichs D, Gogolin C and Kastoryano M 2020 Avoiding local minima in variational quantum eigensolvers with the natural gradient optimizer *Phys. Rev. Res.* **2** 043246
- [85] Broyden C G 1970 The convergence of a class of double-rank minimization algorithms 1. General considerations *IMA J. Appl. Math.* **6** 76
- [86] Fletcher R 1970 A new approach to variable metric algorithms *Comput. J.* **13** 317
- [87] Goldfarb D 1970 A family of variable-metric methods derived by variational means *Math. Comput.* **24** 23
- [88] Shanno D F 1970 Conditioning of quasi-Newton methods for function minimization *Math. Comput.* **24** 647
- [89] Spall J C 1992 Multivariate stochastic approximation using a simultaneous perturbation gradient approximation *IEEE Trans. Autom. Control* **37** 332
- [90] Xiang Y, Sun D, Fan W and Gong X 1997 Generalized simulated annealing algorithm and its application to the Thomson model *Phys. Lett. A* **233** 216