Quantum Machine Intelligence (2025) 7:14
https://doi.org/10.1007/s42484-025-00242-y

RESEARCH ARTICLE

Exploiting symmetries in nuclear Hamiltonians for ground state preparation

Joe Gibbs[1,2] · **Zoë Holmes**[3] · **Paul Stevenson**[1]

Received: 25 September 2024 / Accepted: 20 January 2025

Abstract
The Lipkin and Agassi models are simplified nuclear models that provide natural test beds for quantum simulation methods. Prior work has investigated the suitability of the variational quantum eigensolver (VQE) to find the ground state of these models. There is a growing awareness that if VQE is to prove viable, we will need problem inspired ansätze that take into account the symmetry properties of the problem and use clever initialisation strategies. Here, by focusing on the Lipkin and Agassi models, we investigate how to do this in the context of nuclear physics ground state problems. We further use our observations to discus the potential of new classical, but quantum-inspired, approaches to learning ground states in nuclear problems.

1 Introduction

Nuclei, as interacting systems of spin-1/2 and isospin-1/2 fermions, are natural candidates for simulation on quantum hardware. Moreover, since they are composed of an intermediate number of particles, too many for exact diagonalisation and too few for the thermodynamic approximation, they often prove challenging to simulate accurately by other methods. Nuclei are further characterised by the presence of the strong nuclear interaction which in principle includes many-body forces. In contrast to molecules, nuclei have a single centre but can exhibit deformation and some quasi-molecular structure.

The configuration interaction form of the nuclear shell model features the use of single particle orbitals as basis states, truncated suitably to accommodate only low energy excitations. The nuclear Hamiltonian, renormalised for the truncated basis, is then diagonalised to obtain the nuclear spectrum. Depending on the truncation details, Hamiltonian dimensions can exceed 10^{20} (Dean et al. 2008), and instead methods which target the interesting extremal eigenvalues, such as the Lanczos method, are usually employed. Nevertheless, the method is still prohibitive for many cases of interest, and alternative algorithms with more efficient scaling are called for. Exact solutions are known for simplified versions of the shell model, such as the Lipkin model (Lipkin et al. 1965) (also known as the Lipkin-Meshkov-Glick model) and the slightly more complex Agassi model (Agassi et al. 1966), which we detail later in the present manuscript, and so form natural test beds for new methods.

To overcome the exponential cost of ab-initio calculations of quantum many-body systems, there is growing interest in the use of quantum computers to simulate nuclear structure (Dumitrescu et al. 2018) and reactions (Roggero et al. 2020). One recently proposed method for finding ground states for nuclear problems is the variational quantum eigensolver (VQE) (Peruzzo et al. 2014). The VQE algorithm uses a hybrid quantum classical optimisation loop (Cerezo et al. 2021; Bharti et al. 2022), where the energy of a trial quantum state with respect to a given target Hamiltonian is evaluated on a quantum computer, while a classical optimiser trains a parameterised quantum circuit to minimise this energy. If successfully trained, the state that minimises the energy of the Hamiltonian will be an estimate of its ground state, and the corresponding energy will be an estimate of the ground state energy. Core to the success of such methods will be the choice of *ansatz* for the parameterised quantum circuit.

Prior work has explored finding the ground state of the Lipkin model via VQE with a Bethe solution-inspired circuit ansatz (Robbins and Love 2021), VQE with uni-

✉ Joe Gibbs
j.r.gibbs@surrey.ac.uk

1 School of Mathematics and Physics, University of Surrey, Guildford GU2 7XH, UK

2 AWE, Aldermaston, Reading RG7 4PR, UK

3 Institute of Physics, Ecole Polytechnique Fédéderale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland

Published online: 30 January 2025

tary coupled cluster (UCC) ansatz (Chikaoka and Liang 2022), ADAPT-VQE (Romero et al. 2022), a Hamiltonian-learning VQE method (Robin and Savage 2023), and hybrid classical-quantum algorithms targeting excited states (Hobday et al. 2023; Hlatshwayo et al. 2022; Grossi et al. 2023). Exploratory quantum computing calculations of realistic shell model Hamiltonians using VQE have been undertaken for ^{6}Li (Kiss et al. 2022), isotopes of He, Be, and O (Stetcu et al. 2022), and ^{58}Ni (Bhoy and Stevenson 2024). A method tailored to two valence particles has been applied to ^{6}He, ^{18}O, and ^{42}Ca (Yoshida et al. 2024), and a series of oxygen isotopes calculated out to the neutron drip line (Sarma et al. 2023). Other work has explored the components of the VQE process for the Agassi model (Pérez-Fernández et al. 2022), a more compact encoding of the Agassi model using qudits (Illa et al. 2023), and future larger quantum shell model calculations (Pérez-Obiol et al. 2023) with the ADAPT-VQE ansatz. Here, we investigate the Hamiltonian variational ansatz (Wecker et al. 2015) (HVA) as applied to nuclear problems, beginning with the highly simplified and very symmetric Lipkin model, then moving on to the Agassi model, where we deliberately add a perturbation to explore breaking symmetries of the Hamiltonian.

A major obstacle for executing variational quantum algorithms at scale is the problem of barren plateaus (BP) and exponential concentration (McClean et al. 2018; Cerezo et al. 2021; Holmes et al. 2021; Arrasmith et al. 2022; Marrero et al. 2021; Wang et al. 2021), which exponentially increases the resources required to train a parameterized quantum circuit. These phenomena were initially studied for generic hardware-efficient ansätze (McClean et al. 2018; Holmes et al. 2022). Subsequent work studied more specific circuit structures (Schatzki et al. 2024) and found that encoding symmetries into the ansätze (Meyer et al. 2023; Le et al. 2023) has a positive impact of trainability by increasing the variance of gradients (Larocca et al. 2022; West et al. 2023). Recently, this relationship has been formally proven (Ragone et al. 2024; Fontana et al. 2023). Specifically, it was shown that if the Dynamical Lie Algebra (DLA) is polynomially scaling, the parameterised quantum circuit has guaranteed trainability as the cost function variances will decay at worst polynomially with system size.

Techniques for encoding symmetries into the ansätze of quantum neural networks, also known as "geometric quantum machine learning" (Larocca et al. 2022), have been explored for encoding classical data (Meyer et al. 2023; Chang et al. 2023; Umeano et al. 2023; Tüysüz et al. 2024), condensed matter physics (Sauvage et al. 2024; Lyu et al. 2023), and circuit synthesis (Perrier et al. 2020). Here, we study the effect of encoding symmetries into the ansätze for the VQE applied to nuclear Hamiltonians. In particular, we study two simplified nuclear models, the Lipkin model and Agassi model, and develop ansätze tailored to these models' symmetries. We numerically demonstrate that these symmetrised models, in contrast to a non-structured VQE ansatz, exhibit non-exponentially vanishing gradients and further show the resulting improvements to training. Figure 1 shows a cartoon representation of the VQE algorithm applied to the Agassi model, with its Hamiltonian variational ansatz design made explicit.

While our work here shows that incorporating symmetries provides a simple solution to the BP problem in the context of nuclear problems, this comes at the expense of also making the problem easier to classically simulate. Given the link established in Refs (Ragone et al. 2024; Fontana et al. 2023) between cost variance and DLA scalings, our work indirectly implies that the Lipkin models and Agassi models have poly DLAs (something which would be challenging to compute directly by other means). It is then known that systems with a small DLA are efficient to simulate classically (Goh et al. 2023; Cerezo et al. 2023; Kökcü et al. 2022). Thus, our work can also be viewed as opening up new quantum-inspired avenues for classically simulating these models, where DLA scaling can be determined from computing loss function variances on the quantum computer, as an alternative to the brute-force and exponentially scaling classical calculation. Alternatively, one can note that more realistic nuclear models will be less symmetric and thus not possible to treat exactly with the fully symmetrised ansätze advocated here. We suggest in such contexts our approach will provide a valuable (either classical or quantum) pre-training strategy to obtain a good initialisations for more complex models. We present a demonstration of using symmetrised training to warm-start the VQE of a non-solvable model, by successfully finding an initialisation with the narrow gorge of strong gradients, despite the landscape suffering from a barren plateau.

2 Lipkin model

The Lipkin model was introduced in Lipkin et al. (1965) as a simplified model of a nuclear two-level system, in which n spinless fermions are arranged in two n-fold degenerate levels. The particles interact with a two-body interaction which either scatters pairs up or down or promotes one particle while lowering another. As a simplified nuclear shell model, it largely serves as a test case for approximation methods. This model can also be mapped onto a system of n spin-1/2 particles. When expressed in this way (Vidal et al. 2004), the Hamiltonian is written as

$$H = -\frac{\lambda}{n}\sum_{i<j} X_i X_j - h\sum_i Z_i, \tag{1}$$

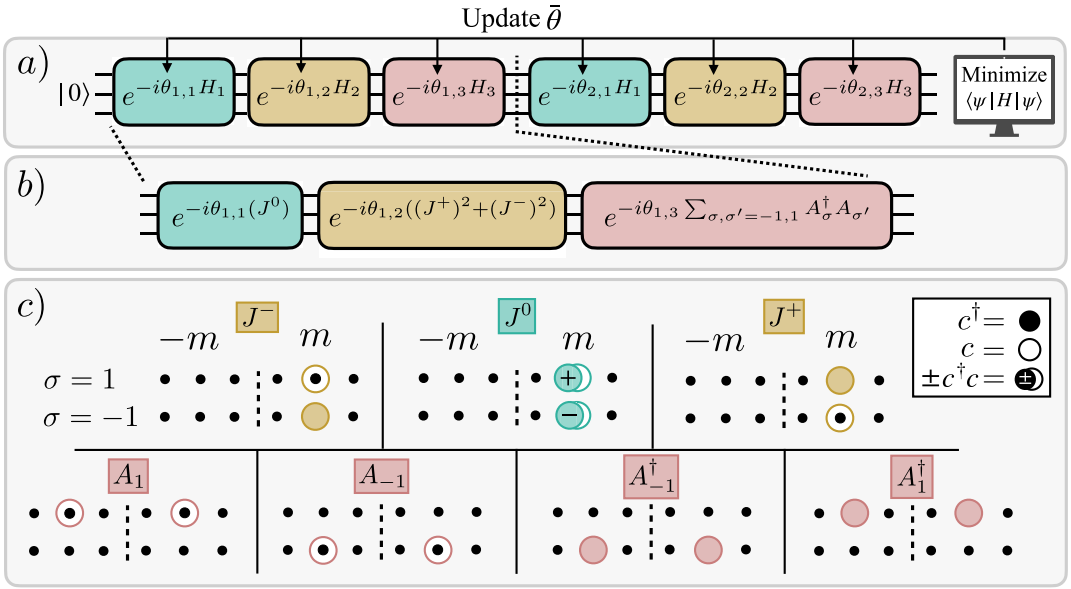

Fig. 1 The Hamiltonian variational ansatz for the Agassi model. **a** Cartoon representation of the VQE algorithm using a Hamiltonian variational ansatz, shown with 3 gate blocks and 2 ansatz layers. **b** The unitaries in each layer of the ansatz are rotations with the hermitian generators corresponding to the three terms in the Agassi model Hamiltonian Eq. 5. (We note that in practise, the gates in **b** would need to be Trotterised to be implemented with a standard gate set, but we do not consider this additional source of error here.) **c** A pictoral representation of the operators in the Hamiltonian. Each term is a sum over magnetisation values, and we show the creation/annihilation operators for $m = 2$

where X_i and Z_i are the Pauli X and Z spin matrices acting on the i$^{\text{th}}$ particle, and λ and h are strengths of the two terms in the Hamiltonian. The inverse scaling of the first term with respect to the number of particles, n, was introduced in Vidal et al. (2004) to ensure a finite free energy per spin in the thermodynamic limit. Throughout our study, we set $\lambda = 1$ and $h = 0.5$.

The HVA was introduced in Wecker et al. (2015) as a problem-tailored ansatz template for the variational quantum eigensolver. For a Hamiltonian expressed as a sum of non-commuting terms, $H = \sum_k H_k$, where an L layered HVA ansatz here is $U(\theta) = \prod_{l=1}^{L}\prod_k e^{-i\theta_{l,k}H_k}$

For close to identity initializations, the HVA has substantial loss variances (Park and Killoran 2024); however, for Hamiltonians with an exponentially growing dynamical Lie algebra, the full landscape has a barren plateau (Fontana et al. 2023; Ragone et al. 2024).

An L-layered HVA for the Lipkin model is given by

$$U(\theta) = \prod_{i=1}^{L}\left(\exp(-i\sum_j \gamma_{i,j}Z_j)\exp(-i\sum_{j<k}\theta_{i,j,k}X_jX_k)\right) \quad (2)$$

$$= \prod_{i=1}^{L}\left(\prod_j \exp(-i\gamma_{i,j}Z_j)\prod_{j<k}\exp(-i\theta_{i,j,k}X_jX_k)\right). \quad (3)$$

The Lipkin model, Eq. 1, is manifestly symmetric with respect to qubit relabelling and is therefore permutationally symmetric. To exploit this property, we can correlate the parameters of an HVA layer such that $\gamma_{i,j} = \gamma_i$ and $\theta_{i,j,k} = \theta_i$. For clarity, this condition fixes the rotation angles for each generator to be all the same within one layer, and the ansatz is therefore permutationally invariant.

We investigate the trainability of performing VQE using this ansatz by computing the average cost function variance across the parameterised landscape. Figure 2 shows a comparison of computing cost function variances, where the cost function for VQE is the energy expectation value. We consider this ansatz in two forms: the first being the permutationally symmetric ansatz with the correlated parameters and the same ansatz structure but un-symmetrised, where all the gate parameters are independent and uncoupled. For both, we use a number of layers equal to the number of qubits. For each system size, we evaluate the cost function 32 times, with randomly chosen each parameters from a uniform distribution in the interval $[-2\pi, 2\pi]$, which we use to calculate the variance in log-space (i.e. the geometric variance).

The unsymmeterised ansatz exhibits a stereotypical barren plateau, signified by the exponential suppression of energy expectation variances as a function of system size. The magnitude of cost function variances is significantly different for the fully permutationally symmetric ansatz. As opposed to

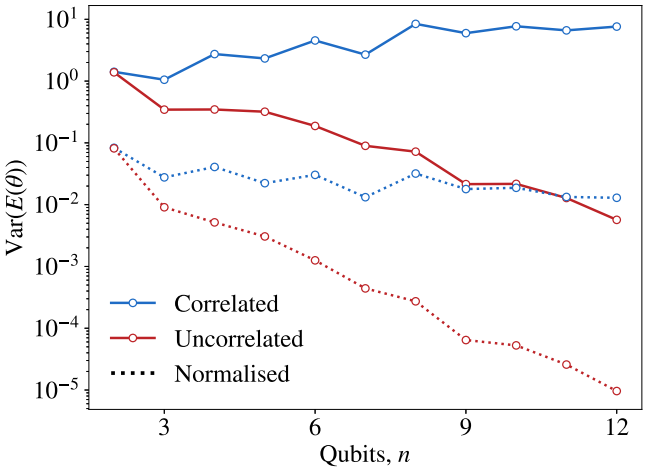

Fig. 2 Variance of VQE cost function landscape for the Lipkin model. Here, we show the scaling of the VQE cost function ($E(\theta) = \langle\psi(\theta)|H|\psi(\theta)\rangle$) variances for random initialisations in the parameter space of the ansatz specified in Eq. 3. The "uncorrelated" ansatz is the parameterised circuit given in Eq. 3 with all rotation gates taking independent parameters; for the "correlated" ansatz, in each layer, the parameters for each gate type are constrained to be equal to preserve the permutation symmetry. We also plot the normalised cost function $\hat{E}(\theta)$ given in Eq. 4 to ensure the cost function is bounded

exponentially decaying with system size, we find they in fact increase with system size.

These increasing variances could be criticised as unphysical and an artifact of the increasing size of the eigenspectum of the Hamiltonian as the number of qubits increases. Similarly, doubling the numerical coefficients of the Hamiltonian would cause the calculated variances to increase by a factor of 4. To most comprehensively remove the effects of the scale of the Hamiltonian on our numerical results, we also use the following normalised and bounded cost function:

$$\hat{E}(\theta) = \frac{\langle\psi(\theta)|H|\psi(\theta)\rangle - E_0}{E_{\max} - E_0} \tag{4}$$

where E_0 and $E_{\max}$ are the minimum and maximum eigenvalues of the Hamiltonian, respectively, to ensure $\hat{E}$ is bounded via $0 \leq \hat{E}(\theta) \leq 1$. We still find that for the unsymmeterised ansatz, the cost function variances exponentially decay with system size, while for the symmetrised ansatz rather than increasing with system size due to the fixed upper bound of 1, they now do not significantly decrease with system size.

While these results give optimism that VQE on the Lipkin model could be performed at scale, it has recently been shown that the ground state energy of permutationally symmetric Hamiltonians can be computed efficiently classically (Anschuetz et al. 2023), meaning this is not a path to near-term quantum advantage over classical approaches. It is possible that the ground state outputted by the symmeterised VQE could be used as an input for another quantum algorithm; however, this cannot involve evolving the ground state by a permutationally symmetric Hamiltonian; otherwise, this will also be classically simulable (Anschuetz et al. 2023). Although the Lipkin model is potentially too symmetric for its study to gain a quantum advantage, this motivates the search for other Hamiltonians with weaker symmetries that can be also exploited for improved trainability.

3 The Agassi model

The Agassi model is an extension of the Lipkin model, with an extra pairing interaction giving more real physics of a nuclear two-level system. It is typically expressed as

$$H = \epsilon J^0 - \frac{V}{2}[(J^+)^2 + (J^-)^2] - g \sum_{\sigma,\sigma'=-1,1} A^\dagger_\sigma A_{\sigma'} \,. \tag{5}$$

Here, J^0 and $J^\pm$ are defined as

$$J^0 = \frac{1}{2} \sum_{m=-j}^{j} c^\dagger_{1,m} c_{1,m} - c^\dagger_{-1,m} c_{-1,m} \tag{6}$$

$$J^+ = \sum_{m=-j}^{j} c^\dagger_{1,m} c_{-1,m} = (J^-)^\dagger \,. \tag{7}$$

Similarly, A_0 and A_1 take the form

$$A_0 = \sum_{m=1}^{j} c_{1,-m} c_{-1,m} - c_{1,m} c_{-1,-m} \tag{8}$$

$$A_1 = \sum_{m=1}^{j} c_{1,-m} c_{1,m} \tag{9}$$

$$A_{-1} = \sum_{m=1}^{j} c_{-1,-m} c_{-1,m} \,. \tag{10}$$

To map the fermionic creation/annihilation operators to Pauli spin operators, we use the Jordan-Wigner transform (Jordan and Wigner 1928). Throughout our study, we set $\epsilon = 0.5$, $V = 1.5$, and $g = 1$.

For performing VQE to produce the ground state of the Agassi model, as done for the Lipkin model, we use the HVA design, which we find to be efficient in the number of trainable parameters required to prepare the ground state for increasing system sizes. For convenience, let $H_1 = J^0$, $H_2 = (J^+)^2 + (J^-)^2$, and $H_3 = \sum_{\sigma,\sigma'=-1,1} A^\dagger_\sigma A_{\sigma'}$. Then, our ansatz with L-layers is defined by

$$U(\theta) = \prod_{i=1}^{L} \prod_{j=1}^{3} \exp(-i\theta_{i,j} H_j) \,. \tag{11}$$

We note that each gate $\exp(-i\theta_{i,j}H_j)$ is in general an n-qubit operator and will require decomposing into 2-qubit gates for execution on a quantum computer, e.g. via a Trotter decomposition. For our numerical results, we do not consider this step and exactly apply the operator via sparse matrix exponentiation.

We choose to target the ground state in the half-filling sector, where the system of $4j$ orbitals has a total occupation of $2j$ fermions. To do this, we add an extra penalty term to the Hamiltonian to target the correct particle number sector, e.g.

$$H_{\text{Half-Filling}} = H + \beta(\hat{N}_1 + \hat{N}_{-1} - 2j)^2 \tag{12}$$

for a sufficiently large β, where $N_\sigma = \sum_{m=-j}^{j} c^\dagger_{\sigma,m} c_{\sigma,m}$.

The β value must be greater than the energy gap between the half-filling ground state and ground states of other filling numbers, and we set this to 100 which satisfies this condition for the systems considered. As the HVA is inspired by adiabatic computing (Tilly et al. 2022), the initial state is chosen to be the ground state of one of the terms in the Hamiltonian. A natural starting initial state is $|\psi_0\rangle = |1\rangle^{\otimes 2j}|0\rangle^{\otimes 2j}$, which is the ground state of the J^0 operator and is easily preparable. We use the same qubit ordering as in Ref. (Pérez-Fernández et al. 2022), where this state corresponds to the $2j$ particles filling the bottom level of the 2-level system.

Compared to the Lipkin model, the Agassi model has a weaker form of permutation symmetry. The Lipkin model has a full permutation symmetry, whereas the Agassi model requires the same permutation applied to the two subsets of qubits with $\sigma = +1/-1$. Also, due to the fermionic nature of the Hamiltonian, exchanging particles introduces wavefunction phases. As a result, the results for provable trainability of permutation symmetric ansätze (Schatzki et al. 2024) do not directly apply to the Agassi model.

We first study the trainability of this ansatz and Hamiltonian combination by quantifying the severity of the barren plateau phenomenon. We note the ansatz design described in Eq. 11 requires $3L$ parameters for a maximum spin of j. For increasing system sizes of the Agassi model, we randomly sample the cost function landscape 1024 times and compute the variance of the measured energy expectation value. We set the number of layers to equal to j (and therefore linearly increasing with system size). Often parameterised gates in variational ansätze are exponentiations of single qubit Pauli matrices, and therefore, their action with respect to the rotation angle has a period of 4π, e.g. if $G^2 = 1$ and $R(\theta) = e^{-i\frac{\theta}{2}G} = \cos(\frac{\theta}{2})I - i\sin(\frac{\theta}{2})G$, then $R(\theta+4\pi) = R(\theta)$. The gate generators used in Eq. 11 do not square to the identity, so it does not have the same 4π periodicity. Therefore, we increase the parameter range uniformly sampled to $[-10, 10]$.

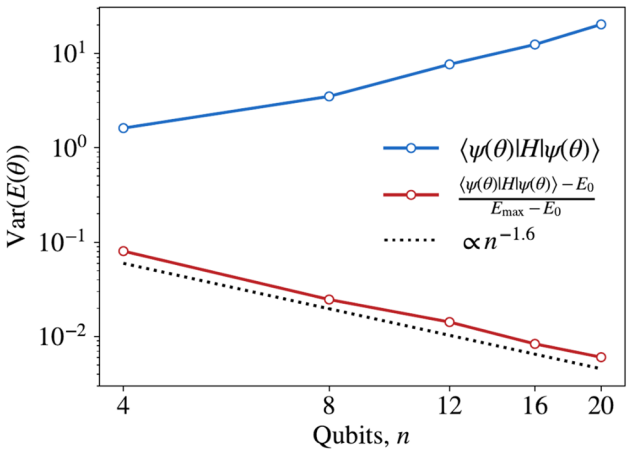

Fig. 3 Variance of VQE cost function landscape for the Agassi model. We plot the variance for both the standard energy expectation value and the normalised expectation value. We find that the variance of the standard energy expectation value grows super-polynomially with system size. The normalised cost function variance decreases polynomially with system size, proportional to $n^{-1.6}$

Figure 3 shows these results. Similar to the Lipkin results presented in Sect. 2, we again find cost function variances that increase with system size. Plotting the normalised cost function [1] to remove the effect of the increasing size of the spectrum, the variances do now decrease with system size but polynomially scaling as $n^{-1.6}$.

We now demonstrate the ability of this ansatz to produce the ground state of the Agassi model. The top row of Fig. 4 shows the results of 100 training runs with random initialisations for increasing system sizes. Plotted in blue are the traces of individual optimisations, where the percentage error from the exact ground state energy is plotted as a function of optimisation steps. The red curve and shading represents the mean and standard deviation of this data. The ansatz is compact in the number of parameters, as empirically we find we need only a linear circuit depth and number of parameters as a function of system size to prepare the ground state to within a 1% error.

We also investigate the benefit gained from warm-starting the training. A feature of the ansatz is that the number of parameters in one layer is independent of system size, as each of the gates is n-qubit operator. We can hypothesise there is a similarity between the solutions in parameter space for neighbouring values of j, particularly for larger system sizes as the thermodynamic limit is approached. Specifically, for a maximum spin j, we randomly sample from the j-1 warm-started solutions that achieved a percentage error less

[1] We stress that this normalisation would not be implemented in practise when performing the algorithm on real quantum hardware. Indeed, it would require already knowing the ground state energy. Therefore, the growing cost function variances indicate this ansatz is highly trainable, even at scale.

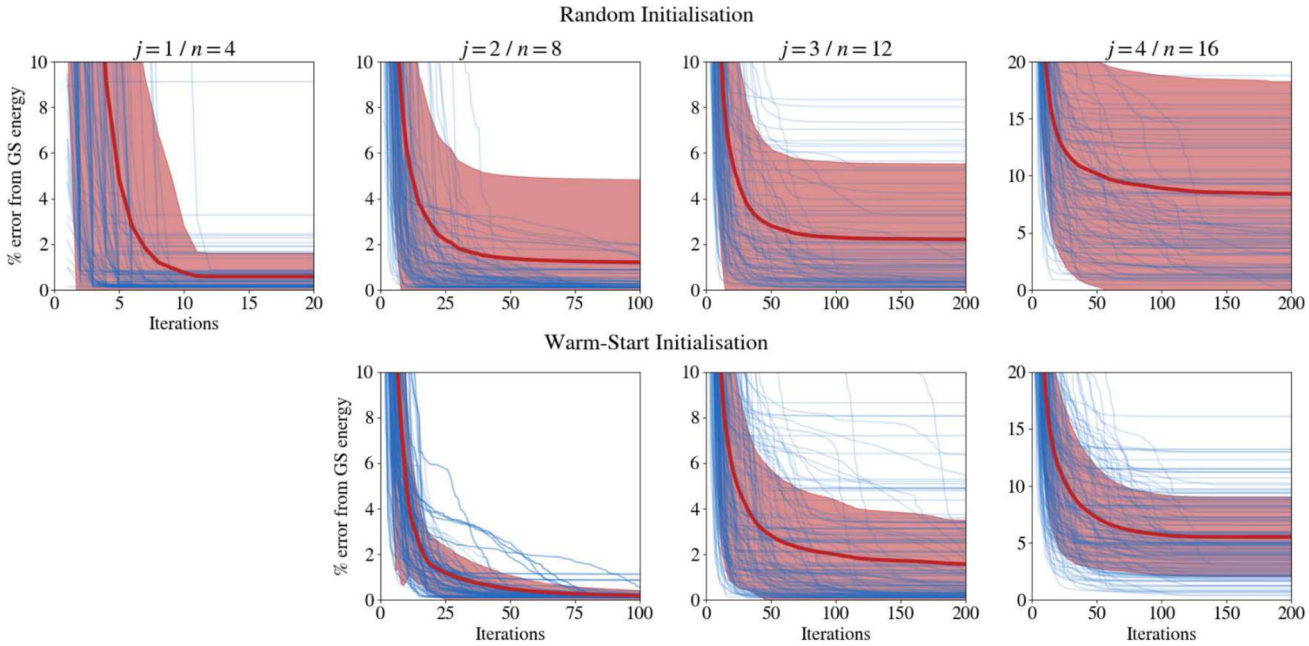

Fig. 4 VQE training of the Agassi model. (Top) The ansatz described in Eq. 11 is optimised to prepare the ground state of the Agassi Hamiltonian, for increasing maximum spin j, requiring a number of qubits $n = 4j$. Plotted in blue are the traces of 100 training runs with random initialisations, and the percentage error from the exact ground state energy is plotted as a function of optimisation steps. The red curve and shading represents the mean and standard deviation of this data. (Bottom) Here, we use a warm-start initialisation, where we randomly choose one of the solutions from the $j-1^{th}$ warm-started training, with an extra layer of identity gates added

than 2% and use the output parameter vector for the initialisation. To keep the linearly increasing circuit depth, we also append one layer of gates initialised close to the identity. The rotation angle is uniformly sampled from a small range $[-10^{-4}, 10^{-4}]$ to avoid potentially initialising at a saddle point and aid escaping local minima. For the optimisation of the parameterised quantum circuit, we use the gradient-based ADAM optimiser (Kingma and Ba 2015). We note that the parameterised-shift rule (Schuld et al. 2019) commonly used for evaluating gradients in parameterised quantum circuits only applies to gates with generators that have at most two unique eigenvalues, which is not true for the gates in this Hamiltonian variational ansatz. More general rules for exact gradient evaluation have been explored in the literature (Izmaylov et al. 2021; Wierichs et al. 2022; Kyriienko and Elfving 2021) that could be used here instead.

It is clear that the warm-starting aids the training process, both in reducing the mean and standard deviation of the termination of the optimisation. Specifically, for $j = 2, 3, 4$, upon termination of the optimisation, we find the mean energy value is reduced by 85.12%, 30.08%, and 34.29%, respectively, and the standard deviations are reduced by 94.05%, 42.05%, and 63.85%. We note that in these implementations, we did not include the effect of shot noise. Thus, in contrast to our loss variance analysis in Fig. 3 where we demonstrate the loss gradients do not exponentially vanish with system size, here we are predominantly exploring the effect of local minima.

4 Warm-starting non-solvable models via parameter transfer

As well as studying our ansatz design to warm-start VQEs of increasing size systems, we also investigate warm-starting non-solvable models. The extension Hamiltonian with a broken symmetry we target is

$$H_{\text{Broken}} = \epsilon J^0_{\text{Broken}} - \frac{V}{2}[(J^+)^2 + (J^-)^2] - g \sum_{\sigma,\sigma'=-1,1} A^\dagger_\sigma A_{\sigma'} \tag{13}$$

where J^0_{Broken} is defined as

$$J^0_{\text{Broken}} = \frac{1}{2}\sum_{m=-j}^{j}(1+\alpha m)(c^\dagger_{1,m}c_{1,m} - c^\dagger_{-1,m}c_{-1,m}). \tag{14}$$

The α parameter adds a linear m dependence to the single particle energies and breaks the permutation invariance of the original Agassi model. For our results, we fix the value of $\alpha =$

2. We note that a constant α causes the breaking of the permutation symmetry to be increasingly strong with system size. To prepare the ground state for H_{Broken}, it is no longer possible to use $G = \{J^0, (J^+)^2 + (J^-)^2, \sum_{\sigma,\sigma'=-1,1} A^\dagger_\sigma A_{\sigma'}\}$ as the gate generators as they can not express the new broken symmetry of the non-interacting ground state. We instead replace the J^0 generator with independent gates acting on each m value, $\{J^0_m\}^j_{m=-j}$ where $J^0_m = c^\dagger_{1,m}c_{1,m} - c^\dagger_{-1,m}c_{-1,m}$. This new "Broken" ansatz, with gate generators G_{Broken}, is now able to fully express the ground state for H_{Broken}.

To take advantage of our previous results demonstrating the strong VQE gradients found for the standard Agassi Hamiltonian, we investigate the potential for warm-starting VQE of H_{Broken} using parameter transfer from symmetrised training. By parameter transfer, we first perform energy minimisation against H_{Broken} using the gate generators G, which finds some reduction in energy but converges prematurely due to limited expressibility. Then, we transfer these parameters over to the ansatz design using generators G_{Broken}, and the optimization is continued, where the initial symmetrised training may find a region in parameter space with strong gradients.

Figure 5 shows the results of this parameter transfer via warm-starting. The warm-starting results are compared against the global variance found by computing energy expectation values of the Hamiltonian H_{Broken} when randomly sampling the ansatz using the gate generators G_{Broken}, and we find these variances to decay exponentially with system size. Then, running VQE against the same Hamiltonian but using the symmetrised ansatz of gate generators G, we copy the parameter vector over to the Broken ansatz. For our example, only the parameter for the J^0 generator needs transforming. Since the individual terms in J^0 mutually commute, the gate in the ansatz is transformed from $\exp(-iJ^0\theta_0)$ to $\prod^j_{m=-j}\exp(-iJ^0_m\theta_0)$, i.e. the $\exp(-iJ^0_m\theta)$ gates are all initialised to the same value found after symmetrised training, but then are allowed to vary independently.

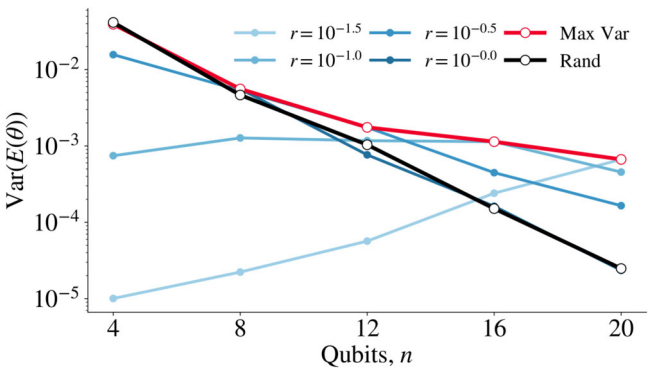

Fig. 5 Warm-starting the Broken Hamiltonian via parameter transfer. Data showing the warm-starting with a symmetrised ansatz can find regions in the parameterised landscape with strong gradients, even when the average properties of the landscape exhibit a barren plateau. The black line shows random sampling of the loss function $\langle\psi(\theta)|H_{\text{Broken}}|\psi(\theta)\rangle$ using the Broken ansatz, with an exponential decay in loss function variances. When initialising the ansatz via parameter transfer from symmetrised training, we randomly sample 1024 loss function values in a hypercube of varying radii r, shown by the thin blue lines. The maximum of these variances for a value of n is given by the thick red line, leading to non-exponentially decaying loss function variances

We then uniformly sample a hypercube around this point found by symmetrised training to study the parameterised landscape. We note that for a very large hypercube radius, the samples will be far from the warm-start, and the loss function variances will see the same decay rate as the global landscape. Also, for a very small hypercube radius, all samples will be very close in the landscape, so again find a small loss function variances. Therefore, we expect an intermediate size hypercube radius where the largest loss function variances are found. We will also expect the optimal hypercube radius to decrease with system size, as the hypercube volume will increase with the number of parameters in the ansatz. To test this, for each system size, we sample the loss function around the warm-started point with differing radii, similar to that studied in other work on warm-starting (Puig et al. 2024).

In Fig. 5, the thin blue lines respectively show the loss function variances when uniformly sampling a hypercube with increasing radius. We note that for the largest radius, $r = 1$, the variances closely match that of a global sampling across the full landscape, shown by the thick black line. The thick red line shows the maximum variances across the range of widths and shows increased and non-exponentially decaying loss function variances. This demonstrated that the symmetrised training was able to warm-start VQE of H_{Broken} by finding an initialisation close to the narrow gorge of strong gradients. As the system size increased, the optimal hypercube radius giving the largest loss function variances decreased. In practise, this optimal hypercube radius could be determined heuristically; however, this relationship has been studied in more detail analytically (Puig et al. 2024). It was found that when warm-starting an iterative variational Trotter compression, for an ansatz with M parameters, a hypercube of radius $\mathcal{O}(\frac{1}{\sqrt{M}})$ around the warm-start supported polynomially large gradients.

5 Discussion

Here, we presented a practical study of the trainability of symmetrised ansätze for two nuclear Hamiltonians: the Lipkin model and the Agassi model. For both Hamiltonians, we employ a variant of the Hamiltonian variational ansatz which has been modified to fully respect the symmetries of the target Hamiltonian. Rather than observing decaying cost variances

(as would usually be expected), or indeed the exponential decay of a barren plateau, we instead find that the variances increase with system size. This effect can be attributed to the fact that the models' Hamiltonian eigenspectrum range grows with system size. On artificially normalising the cost to account for this, we observe that the normalised cost variance decays at worst polynomially. Thus, we see that the symmeterised ansatz can indeed be used to avoid the barren plateau phenomenon.

For both the Lipkin and Agassi models, we empirically find that a linearly increasing number of trainable gates is sufficient to prepare the ground state, and hence our ansatz design seems to be efficient in circuit depth and trainable parameter, requiring a number of parameters scaling linearly with system size. Furthermore, we stress that the number of parameters in each layer of the ansatz does not scale with the number of qubits. This observation allowed us to motivate a simple parameter transfer strategy for using solutions from smaller system sizes to provide a warm-start initialisation. Our numerics indicate that this strategy has a positive effect of the success of the optimisation.

To motivate utility beyond the ground state preparations of Hamiltonians with a polynomially scaling DLA, we study the potential for symmetrised training to warm-start VQEs of more complex Hamiltonians. We consider a variant of the Agassi model with a broken symmetry resulting in a global exponential decay in loss function variances for the HVA. Pre-training with a symmetrised ansatz is able to find regions of the strong gradients, allowing a parameter transfer to initialise within the narrow gorge of strong gradients. These results are in line with other works that have investigated alternate approaches for efficient pre-training which allow an initialisation in landscape regions with strong gradients (Goh et al. 2023; Rudolph et al. 2023).

Determining the scaling of the dimension of the DLA is a critical parameter for understanding both the trainability of the parameterised quantum circuit and the classical simulability of the target problem; however, in practise, this is a challenging task. While the DLA can be studied numerically, the brute-force calculation is costly and scales exponentially with system size, limiting its application .[2] In some cases, the DLA scaling has been determined analytically (Wiersema et al. 2023; Schatzki et al. 2024); however, this is a difficult task for more general Hamiltonians. As a result, although there is a clear relationship between the DLA and both trainability and classical simulability, in reality, determining the DLA scaling has been performed for few systems.

If the recently derived relationship between the size of the DLA subspace and cost function variances is inverted, our results indicate that the Agassi model has a polynomially scaling DLA, implying the system will be efficiently classically simulable by techniques that exploit this property (Goh et al. 2023). This provokes the use of new quantum-inspired classical algorithms for studying nuclear Hamiltonians.

Moving forward, it would be interesting to study how symmeterized ansätze could be applied to ground state computations of more complex nuclear models. In particular, it is natural to investigate whether a symmeterised solution to one of the simpler models studied here might provide a good initialisation for future larger nuclear shell model calculations (Pérez-Obiol et al. 2023). A natural test case could be along an isotopic chain for a magic number of protons, e.g. in oxygen as neutrons are added across and beyond the *sd* shell (Brown 2017). For such models, it would be valuable to further explore whether one could not only upper bound the ground state energy but also lower bound it via Dual-VQE (Westerheim et al. 2023), a semidefinite programming inspired variational quantum algorithm using slack variables (Chen et al. 2023).

Acknowledgements We thank Martin Larocca for interesting discussions of Dynamical Lie Algebra. ZH acknowledges support from the Sandoz Family Foundation-Monique de Meuron program for Academic Promotion. PS acknowledges support from the UK STFC under grants ST/W006472/1 and ST/V001108/1 and support from AWE under the William Penney Fellowship scheme. UK Ministry of Defence ©Crown owned copyright 2024/AWE

Author Contributions J.G. produced the numerical results and created the figures. All authors contributed to the project conception, technical discussion, and the writing of the manuscript text.

Data Availability Data that supports the findings of this study is available under reasonable request.

Declarations

Conflict of Interest The authors declare no competing interests.

References

Agassi D, Lipkin H, Meshkov N (1966) Validity of manybody approximation methods for a solvable model:(iv). the deformed Hartree-Fock solution. Nucl Phys 86:321

[2] see Appendix E of Larocca et al. (2022) for a discussion of the algorithm.

Anschuetz ER, Bauer A, Kiani BT, Lloyd S (2023) Efficient classical algorithms for simulating symmetric quantum systems. Quantum 7:1189

Arrasmith A, Holmes Z, Cerezo M, Coles PJ (2022) Equivalence of quantum barren plateaus to cost concentration and narrow gorges. Quantum Sci Technol 7:045015

Bharti K, Haug T, Vedral V, Kwek L-C (2022) Noisy intermediate-scale quantum algorithm for semidefinite programming. Phys Rev A 105:052445

Bhoy B, Stevenson P (2024) Shell-model study of 58Ni using quantum computing algorithm. New J Phys 26:075001. publisher: IOP Publishing

Brown BA (2017) The oxygen isotopes. Int J Modern Phys E 26:1

Cerezo M, Sone A, Volkoff T, Cincio L, Coles PJ (2021) Cost function dependent barren plateaus in shallow parametrized quantum circuits. Nat Commun 12:1

Cerezo M, Arrasmith A, Babbush R, Benjamin SC, Endo S, Fujii K, McClean JR, Mitarai K, Yuan X, Cincio L, Coles PJ (2021) Variational quantum algorithms. Nat Rev Phys 3:625–644

Cerezo M, Larocca M, García-Martín D, Diaz NL, Braccia P, Fontana E, Rudolph MS, Bermejo P, Ijaz A, Thanasilp S et al (2023) Does provable absence of barren plateaus imply classical simulability? Or, why we need to rethink variational quantum computing. arXiv preprint arXiv:2312.09121

Chang SY, Grossi M, Le Saux B, Vallecorsa S (2023) Approximately equivariant quantum neural network for p4m group symmetries in images. In 2023 IEEE International conference on quantum computing and engineering (QCE), Vol 01, pp 229–235

Chen J, Westerheim H, Holmes Z, Luo I, Nuradha T, Patel D, Rethinasamy S, Wang K, Wilde MM (2023) QSlack: a slack-variable approach for variational quantum semi-definite programming. arXiv preprint arXiv:2312.03830 https://doi.org/10.48550/arXiv.2312.03830

Chikaoka A, Liang H (2022) Quantum computing for the Lipkin model with unitary coupled cluster and structure learning ansatz. Chin Phys C 46:024106

Dean DJ, Hagen G, Hjorth-Jensen M, Papenbrock T (2008) Computational aspects of nuclear coupled-cluster theory. Comput Sci Discov 1:015008

Dumitrescu EF, McCaskey AJ, Hagen G, Jansen GR, Morris TD, Papenbrock T, Pooser RC, Dean DJ, Lougovski P (2018) Cloud quantum computing of an atomic nucleus. Phys Rev Lett 120:210501

Fontana E, Herman D, Chakrabarti S, Kumar N, Yalovetzky R, Heredge J, Sureshbabu SH, Pistoia M (2023) The adjoint is all you need: characterizing barren plateaus in quantum ansatze. arXiv preprint arXiv:2309.07902

Goh ML, Larocca M, Cincio L, Cerezo M, Sauvage F (2023) Lie-algebraic classical simulations for variational quantum computing. arXiv preprint arXiv:2308.01432

Grossi M, Kiss O, De Luca F, Zollo C, Gremese I, Mandarino A (2023) Finite-size criticality in fully connected spin models on superconducting quantum hardware. Phys Rev E 107:024113

Hlatshwayo MQ, Zhang Y, Wibowo H, LaRose R, Lacroix D, Litvinova E (2022) Simulating excited states of the Lipkin model on a quantum computer. Phys Rev C 106:024319

Hobday I, Stevenson P, Benstead J (2023) Variance minimisation on a quantum computer of the Lipkin-Meshkov-Glick model with three particles. EPJ Web Conf 284:16002

Holmes Z, Arrasmith A, Yan B, Coles PJ, Albrecht A, Sornborger AT (2021) Barren plateaus preclude learning scramblers. Phys Rev Lett 126:190501

Holmes Z, Sharma K, Cerezo M, Coles PJ (2022) Connecting ansatz expressibility to gradient magnitudes and barren plateaus. PRX Quantum 3:010313

Illa M, Robin CEP, Savage MJ (2023) Quantum simulations of SO(5) many-fermion systems using qudits. Phys Rev C 108:064306

Izmaylov AF, Lang RA, Yen T-C (2021) Analytic gradients in variational quantum algorithms: algebraic extensions of the parameter-shift rule to general unitary transformations. Phys Rev A 104:062443

Jordan P, Wigner E (1928) Über das Paulische Äquivalenzverbot. Zeitschrift für Physik 47:631

Kingma DP, Ba J (2015) Adam: a method for stochastic optimization. In Proceedings of the 3rd international conference on learning representations (ICLR)

Kiss O, Grossi M, Lougovski P, Sanchez F, Vallecorsa S, Papenbrock T (2022) Quantum computing of the ^{6}Li nucleus via ordered unitary coupled clusters. Phys Rev C 106:034325

Kökcü E, Steckmann T, Wang Y, Freericks J, Dumitrescu EF, Kemper AF (2022) Fixed depth Hamiltonian simulation via Cartan decomposition. Phys Rev Lett 129:070501

Kyriienko O, Elfving VE (2021) Generalized quantum circuit differentiation rules. Phys Rev A 104:052417

Larocca M, Czarnik P, Sharma K, Muraleedharan G, Coles PJ, Cerezo M (2022) Diagnosing barren plateaus with tools from quantum optimal control. Quantum 6:824

Larocca M, Sauvage F, Sbahi FM, Verdon G, Coles PJ, Cerezo M (2022) Group-invariant quantum machine learning. PRX. Quantum 3:030341

Le INM, Kiss O, Schuhmacher J, Tavernelli I, Tacchino F (2023) Symmetry-invariant quantum machine learning force fields. arXiv preprint arXiv:2311.11362

Lipkin HJ, Meshkov N, Glick A (1965) Validity of manybody approximation methods for a solvable model:(i). exact solutions and perturbation theory. Nucl Phys 62:188

Lyu C, Xu X, Yung M-H, Bayat A (2023) Symmetry enhanced variational quantum spin eigensolver. Quantum 7:899

Marrero CO, Kieferová M, Wiebe N (2021) Entanglement-induced barren plateaus. PRX. Quantum 2:040316

McClean JR, Boixo S, Smelyanskiy VN, Babbush R, Neven H (2018) Barren plateaus in quantum neural network training landscapes. Nat Commun 9:1

Meyer JJ, Mularski M, Gil-Fuster E, Mele AA, Arzani F, Wilms A, Eisert J (2023) Exploiting symmetry in variational quantum machine learning. PRX Quantum 4:010328

Park C-Y, Killoran N (2024) Hamiltonian variational ansatz without barren plateaus. Quantum 8:1239

Pérez-Fernández P, Arias J-M, García-Ramos J-E, Lamata L (2022) A digital quantum simulation of the Agassi model. Phys Lett B 829:137133

Pérez-Obiol A, Romero AM, Menéndez J, Rios A, García-Sáez A, Juliá-Díaz B (2023) Nuclear shell-model simulation in digital quantum computers. Sci Rep 13:12291

Perrier E, Tao D, Ferrie C (2020) Quantum geometric machine learning for quantum circuits and control. New J Phys 22:103056

Peruzzo A, McClean J, Shadbolt P, Yung M-H, Zhou X-Q, Love PJ, Aspuru-Guzik A, O'brien JL, (2014) A variational eigenvalue solver on a photonic quantum processor. Nat Commun 5:1

Puig R, Drudis M, Thanasilp S, Holmes Z (2024) Variational quantum simulation: a case study for understanding warm starts. arXiv preprint arXiv:2404.10044 https://doi.org/10.48550/arXiv.2404.10044

Ragone M, Bakalov BN, Sauvage F, Kemper AF, Ortiz Marrero C, Larocca M, Cerezo M (2024) A lie algebraic theory of barren plateaus for deep parameterized quantum circuits. Nat Commun 15:7172

Robbins K, Love PJ (2021) Benchmarking near-term quantum devices with the variational quantum eigensolver and the Lipkin-Meshkov-Glick model. Phys Rev A 104:022412

Robin CEP, Savage MJ (2023) Quantum simulations in effective model spaces: Hamiltonian-learning variational quantum eigen-

solver using digital quantum computers and application to the Lipkin-Meshkov-Glick model. Phys Rev C 108:024313
Roggero A, Li AC, Carlson J, Gupta R, Perdue GN (2020) Quantum computing for neutrino-nucleus scattering. Phys Rev D 101:074038
Romero AM, Engel J, Tang HL, Economou SE (2022) Solving nuclear structure problems with the adaptive variational quantum algorithm. Phys Rev C 105:064317
Rudolph MS, Miller J, Motlagh D, Chen J, Acharya A, Perdomo-Ortiz A (2023) Synergistic pretraining of parametrized quantum circuits via tensor networks. Nat Commun 14:8367
Sarma C, Di Matteo O, Abhishek A, Srivastava PC (2023) Prediction of the neutron drip line in oxygen isotopes using quantum computation. arXiv:2306.06432
Sauvage F, Larocca M, Coles PJ, Cerezo M (2024) Building spatial symmetries into parameterized quantum circuits for faster training. Quantum Sci Technol 9:015029
Schatzki L, Larocca M, Nguyen QT, Sauvage F, Cerezo M (2024) Theoretical guarantees for permutationequivariant quantum neural networks. npj Quantum Inf 10:1
Schuld M, Bergholm V, Gogolin C, Izaac J, Killoran N (2019) Evaluating analytic gradients on quantum hardware. Phys Rev A 99:032331
Stetcu I, Baroni A, Carlson J (2022) Variational approaches to constructing the many-body nuclear ground state for quantum computing. Phys Rev C 105:064308
Tilly J, Chen H, Cao S, Picozzi D, Setia K, Li Y, Grant E, Wossnig L, Rungger I, Booth GH et al (2022) The variational quantum eigensolver: a review of methods and best practices. Phys Rep 986:1
Tüysüz C, Chang SY, Demidik M, Jansen K, Vallecorsa S, Grossi M (2024) Symmetry breaking in geometric quantum machine learning in the presence of noise. arXiv preprint arXiv:2401.10293
Umeano C, Paine AE, Elfving VE, Kyriienko O (2023) What can we learn from quantum convolutional neural networks?. arXiv preprint arXiv:2308.16664
Vidal J, Mosseri R, Dukelsky J (2004) Entanglement in a first-order quantum phase transition. Phys Rev A At Mol Opt Phys 69:054101
Wang S, Fontana E, Cerezo M, Sharma K, Sone A, Cincio L, Coles PJ (2021) Noise-induced barren plateaus in variational quantum algorithms. Nat Commun 12:1
Wecker D, Hastings MB, Troyer M (2015) Progress towards practical quantum variational algorithms. Phys Rev A 92:042303
West MT, Heredge J, Sevior M, Usman M (2023) Provably trainable rotationally equivariant quantum machine learning. arXiv preprint arXiv:2311.05873
Westerheim H, Chen J, Holmes Z, Luo I, Nuradha T, Patel D, Rethinasamy S, Wang S, Wilde MM (2023) Dual-VQE: a quantum algorithm to lower bound the ground-state energy. arXiv preprint arXiv:2312.03083 https://doi.org/10.48550/arXiv.2312.03083
Wierichs D, Izaac J, Wang C, Lin CY-Y (2022) General parameter-shift rules for quantum gradients. Quantum
Wiersema R, Kökcü E, Kemper AF, Bakalov BN (2023) Classification of dynamical lie algebras for translation-invariant 2-local spin systems in one dimension. arXiv preprint arXiv:2309.05690
Yoshida S, Sato T, Ogata T, Naito T, Kimura M (2024) Accurate and precise quantum computation of valence two-neutron systems. Phys Rev C 109:064305