Quantum Science and Technology

PAPER

Extending the variational quantum eigensolver to finite temperatures

Johannes Selisko, Maximilian Amsler, Thomas Hammerschmidt, Ralf Drautz, and Thomas Eckl

1 Robert Bosch GmbH, Robert-Bosch-Campus 1, 71272 Renningen, Germany
2 ICAMS, Ruhr-Universität Bochum, Universitätsstr. 150, 44801 Bochum, Germany
* Authors to whom any correspondence should be addressed.
E-mail: johannes.selisko@de.bosch.com, maximilian.amsler@de.bosch.com and thomas.eckl@de.bosch.com

Keywords: quantum computing, NISQ, free energy, variational, thermal

Supplementary material for this article is available online

Abstract
We present a variational quantum thermalizer (VQT), called quantum-VQT (qVQT), which extends the variational quantum eigensolver to finite temperatures. The qVQT makes use of an intermediate measurement between two variational circuits to encode a density matrix on a quantum device. A classical optimization provides the thermal state and, simultaneously, all associated excited states of a quantum mechanical system. We demonstrate the capabilities of the qVQT for two different spin systems. First, we analyze the performance of qVQT as a function of the circuit depth and the temperature for a one-dimensional Heisenberg chain. Second, we use the excited states to map the complete, temperature dependent phase diagram of a two-dimensional $J_1-J_2$ Heisenberg model. Numerical experiments on both quantum simulators and real quantum hardware demonstrate the efficiency of our approach, which can be readily applied to study various quantum many-body systems at finite temperatures on currently available noisy intermediate-scale quantum devices.

1. Introduction

Recent advances in quantum computing have been driving intense research in the development of quantum algorithms that offer significant improvements over their classical counterparts, and for some applications quantum advantage has already been experimentally demonstrated [1, 2]. In particular, quantum algorithms are used for studying interacting many-electron systems that fundamentally govern the properties of materials and molecules [3, 4]. In quantum chemistry and materials modeling, a quantum advantage [5] could be achieved either by offering a significant, potentially exponential acceleration of conventional methods to (approximately) solve the electronic Schrödinger equation, or by improving accuracy by incorporating a better description of the many-body effects of strongly correlated electronic systems [6]. However, universal fault-tolerant quantum hardware [7] is required to harness the full potential of quantum computing, which is expected to be deployed only within the next decade. Currently available experimental devices, so-called noisy intermediate-scale quantum (NISQ) computers, are limited by their inherent circuit noise and their decoherence time, posing strong constraints with respect to the number of qubits and their connectivity, the circuit depth, and the number of gate operations which can be executed within quantum algorithms [8].

Due to these constraints, the execution of algorithms like quantum phase estimation [9] or quantum Fourier transform [10] are impractical on NISQs, and most methods that produce quantum circuits executable on available hardware are centered around hybrid quantum–classical algorithms. For example, the variational quantum eigensolver (VQE) [11–13] uses a quantum computer to store a parametrized wave function and measure its energy, while a classical, external minimization of the energy through variational parameters provides an approximation to the ground-state of the system. Although the classical optimization...
in a VQE is challenging due to the presence of local minima and barren plateaus [13–15], its utility has already been experimentally demonstrated for small molecular systems [11, 16].

In addition to the ground states, many applications require the assessment of the thermal state at finite temperatures (Gibbs state) or the properties of excited states in a system. The Gibbs state minimizes the Helmholtz free energy \( F = E - TS \) at an inverse temperature \( \beta = 1/(k_B T) \) (in units of inverse normalized interaction parameters) with the energy \( E \) and the entropy \( S \). This state is mixed and can be formulated in terms of the eigenstates \( |\varphi_i\rangle \) and eigenenergies \( \epsilon_i \) of the Hamiltonian:

\[
\hat{\rho}_{\text{Gibbs}} = \sum_{i=0}^{\infty} p_i|\varphi_i\rangle \langle \varphi_i|,
\]

with

\[
p_i = e^{-\beta \epsilon_i} / Z, \quad Z = \sum_i e^{-\beta \epsilon_i}.
\]

To date, two main classes of algorithms have been developed to compute the Gibbs state. The first class is based on computing each eigenstate separately and subsequently mixing them according to their probabilities \( p_i \) in equation (1). Such algorithms usually start out by computing the ground-state using a VQE, followed by successively computing the excited (eigen-) states and projecting them out, or penalizing the already computed eigenstates [17, 18]. Other algorithms like the subspace search VQE (SSVQE) [19] construct a smaller subspace of the problem’s Hilbert space and find its largest eigenvalue and eigenvector through optimization in order to construct the \( k \)-th excited state. By increasing the size of the subspace, all excited states can be found. The weighted SSVQE [19] assigns fixed weights to some orthogonal input states and, by minimization, maps the input states to the excited states (i.e. the states are mapped according to their eigenvalues: the lowest eigenvalue is assigned to the state with the highest weight). Other approaches estimate the higher order reduced density matrices [20] or use the quantum computer to construct an entangled subspace problem which can be solved classically [21].

The second class of algorithms prepares the Gibbs state itself on the quantum device, which involves the explicit treatment of the entropy \( S \). The thermofield-double-states method, for example, doubles the system and collapses it in order to introduce entropy into the system [22–24]. However, the measurement of the entropy is far from trivial: a quantum computer can measure the expectation value of an operator \( \hat{O} \) on a state described by a density matrix \( \hat{\rho} \) by taking the trace \( \text{Tr}(\hat{O}\hat{\rho}) \), but the expression for the entropy \( \hat{\rho}\ln(\hat{\rho}) \) is not linear in \( \hat{\rho} \), rendering the corresponding measurement more demanding. On the other hand, using imaginary time evolution to obtain the Gibbs state requires deep quantum circuits which are challenging to implement on NISQ devices [25–27]. Variational imaginary time evolution can circumvent this issue by approximating the imaginary time evolution with a variational quantum circuit (VQC) [25, 28].

Verdon et al [29] recently introduced a combination of a machine-learning algorithm with a VQC, called hybrid variational quantum thermalizer (hVQT), which generalizes the VQE toward finite temperatures and involves a neural network that learns the entropic probability distribution, while a quantum circuit prepares the eigenstates of the Hamiltonian (see also [30]). This approach is an extension of the weighted SSVQE and includes the tuning of the weights of the eigenstates to obtain the probabilities required in the construction of the Gibbs state. It is similar to the \( \beta \)-VQE introduced in [31]. In contrast to imaginary time evolution, the design of this algorithm allows accessing properties of the thermal state and of each eigenstate. However, the hVQT requires an intimate interaction of classical and quantum computation, which typically results in longer runtimes (see section 2.5), and has thus motivated the development of alternative algorithms [32].

To alleviate above issues of the hVQT and to maximally benefit from a possible advantage of quantum machine-learning [33, 34] we develop an algorithm which transfers the generation of the entropic probability distribution directly onto the quantum computer, thereby allowing to fully assess the Gibbs state on the quantum device. This approach, which we call quantum-VQT (qVQT), offers significant advantages over the hVQT as it minimizes the communication between classical and quantum computer, and we demonstrate the improved performance of qVQT over hVQT with numerical experiments on real NISQ hardware (see section 2.5). In the remainder of this manuscript we present in detail the qVQT algorithm and demonstrate its performance by applying it to solve a one-dimensional (1D) Heisenberg chain, and by computing the complete, temperature dependent phase diagram of a two-dimensional (2D) \( J_1-J_2 \) Heisenberg model.

2. Methods

2.1. Principles of the qVQT

The flowchart and the relevant components of the qVQT algorithm are shown in figure 1. The fundamental idea of the qVQT is to use two separate VQCs and an intermediate measurement to obtain a mixed state on a
quantum computer (red block ‘QPU’ in figure 1). A classical optimization determines the parameters for which this mixed state represents the Gibbs state (blue block ‘CPU’ in figure 1). Different flavors of VQC have been proposed in the literature, e.g. hardware efficient VQC, particle number conserving VQC, variational Hamilton ansatz, etc, which can also be used in qVQT [12, 35, 36]. We denote the parameters of the first VQC (VQC₁) with \( \vec{\phi} \) while the parameters of the second VQC (VQC₂) are referred to as \( \vec{\theta} \).

The first variational circuit VQC₁(\( \vec{\phi} \)) and the intermediate measurement generate a classical distribution (see ‘QPU’ in figure 1). Specifically, the superposition of the basis states |\( b_i \rangle \) produced by VQC₁(\( \vec{\phi} \)) collapses to a probability distribution. \( \hat{\rho}_{VQC₁} \) in equation (3) represents the density matrix after the first variational circuit VQC₁(\( \vec{\phi} \)), while \( \hat{\rho}_{im} \) in equation (4) is the density matrix after the intermediate measurement:

\[
\hat{\rho}_{VQC₁} = \left( \sum_i a_i (\vec{\phi}) |b_i \rangle \right) \left( \sum_i a_i^* (\vec{\phi}) \langle b_i | \right)
\]

\[
\hat{\rho}_{im} = \sum_i |a_i (\vec{\phi})|^2 |b_i \rangle \langle b_i |.
\]

The second variational circuit VQC₂(\( \vec{\theta} \)) maps the basis states |\( b_i \rangle \) to a superposition of these basis states while preserving the orthogonality, and prepares the state

\[
\hat{\rho}_{VQC₂} = \sum_i |a_i (\vec{\phi})|^2 |\psi_i (\vec{\theta}) \rangle \langle \psi_i (\vec{\theta}) |
\]

which has the same form as equation (1).

We obtain the free energy \( F \) by minimizing its value over the parameter set (\( \vec{\phi}, \vec{\theta} \)), a task which is performed classically using an arbitrary (local) optimizer (see ‘CPU’ in figure 1). The energy \( E \) is obtained by measuring the expectation value of the Hamilton operator after the second variational circuit VQC₂, and the entropy \( S \) is obtained by the intermediate measurement of VQC₁. From the probabilities \( p_i \) of observing the basis state |\( b_i \rangle \) we obtain the entropy:

\[
S = \sum_i p_i \ln (p_i).
\]
To assess the resource cost and scaling of the qVQT we first discuss the error estimate as a function of the number of measurements in section 2.2, then determine the required memory resources in section 2.3, and analyze the complexity of the qVQT in section 2.4. Finally we discuss the execution times and accuracy on real NISQ hardware in section 2.5.

2.2. Derivation of the measurement precision

Within the qVQT the probabilities are obtained from the intermediate measurement. Its standard error is obtained from summing over all possible outcomes of \( N \) measurements:

\[
(\Delta p_i)^2 = \sum_{j=0}^{N} \binom{N}{j} p_j^i (1 - p_j)^{(N-j)} \cdot \left( \frac{j}{N} - p_i \right)^2 
\]

(7)

\[
= \frac{p_i (1 - p_i)}{N}.
\]

(8)

Drawing \( N \) samples from a random distribution with standard deviation \( \sigma \) yields a standard error of \( \Delta = \frac{\sigma}{\sqrt{N}} \). Hence, an algorithm which measures all eigenvalues \( \epsilon_i \) of a Hamiltonian with \( p_i N \) shots yields standard errors of \( p_i \) and \( \epsilon_i \):

\[
\Delta p_i = \frac{\sqrt{p_i (1 - p_i)}}{\sqrt{N}} 
\]

(9)

\[
\Delta \epsilon_i = \frac{\sigma_i}{\sqrt{p_i N}}.
\]

(10)

With \( k_B \) set to 1, the standard error of the free energy is given by:

\[
\Delta F = \sqrt{\sum_i \left[ \left( \frac{\partial F}{\partial \epsilon_i} \Delta \epsilon_i \right)^2 + \left( \frac{\partial F}{\partial p_i} \Delta p_i \right)^2 \right]}
\]

(11)

\[
= \sqrt{\sum_i \left[ (p_i \Delta \epsilon_i)^2 + ((\epsilon_i + T \ln (p_i) + 1) \Delta p_i)^2 \right]}
\]

(12)

\[
= \sqrt{\frac{1}{N} \sum_i p_i \left[ \sigma_i^2 + (\epsilon_i + T \ln (p_i) + 1)^2 (1 - p_i) \right]}.
\]

(13)

The leading order is, similarly to VQE, proportional to \( 1/\sqrt{N} \).

2.3. Memory requirements

Computing the entropy requires the storage of all probabilities. Since the number of states grows exponentially with the system size the memory requirement \( M \) grows exponentially as well up to the point where the number of measurements \( N \) limits the number of states which are measured. This could be the case if there are excited states with probabilities comparable to \( 1/N \). Since the states and counts are stored in a dictionary, all states which do not occur in the \( N \) measurements do not require any memory, while all other states require memory \( M(N) \) to store an integer smaller than \( N \).

The qVQT can circumvent this issue if we allow the first variational circuit to split the system into \( n/n_s \) independent subsystems of size \( n_s \), which means that there are no entangling gates between these subsystems. In this case, the subsystems that form the total system are independent from each other by construction. Hence, we can write the complexity of the total system as a product of the constituents, and therefore the entropy of the whole system is given by the sum of the entropy of the subsystems. The measurement is done in the same way, but the probabilities are only stored in each subsystem. This reduces the number of probabilities, which need to be stored, from \( 2^n \) to \( n/n_s \cdot 2^{n_s} \).

\[
M = \frac{n}{n_s} \cdot M(N).
\]

(14)

This does not mean that the eigenstates of the Hamiltonian need to be product states of states in these independent subsystems, but rather imposes an approximation to their associated probabilities, which are products of the probabilities of the input states of the subsystems. The simplest example is the case where \( n_s = 1 \), which corresponds to all qubits being independent of each other and the variational circuit \( \text{VQC}_i \) does not contain entangling gates. A single qubit rotation gives the probability for each qubit to be 0 or 1 after the intermediate measurement, and the probability of each input state is the product of the probabilities of each qubit exhibiting the corresponding state.
2.4. Complexity
The goal of the qVQT is to approximate a density matrix of a mixed state, which is a $2^n \times 2^n$-dimensional complex and symmetric matrix with a trace of 1 for a system of dimension $n$. Therefore, the density matrix in qVQT has $2^{2n} - 1$ degrees of freedom. An equivalent VQE only tries to find a pure state (a statevector) and has hence $2^n - 2$ degrees of freedom. We assess whether the computational effort is comparable to a VQE with twice as many qubits or if it adds an exponential pre-factor to the cost of a VQE in section A of the SM. For the rather small examples used in our numerical experiments we find that the necessary number of parameters, and therefore the cost of the optimization, strongly depends on the variational circuits and we do not find evidence for an exponential scaling compared to VQE.

2.5. Performance on real NISQ hardware
A major advantage of the qVQT over the hVQT is the minimization of the communication between classical and quantum computer, which includes the classically required pre (and post) processing when executing a quantum circuit on a quantum computer. The total execution time of a quantum circuit on a quantum hardware includes transpiling the quantum circuit into native gates, translating gates into hardware pulses, the actual execution on the hardware, and post processing of the results.

In figure 2 we show the execution times (excluding any waiting time in the queue or other classical computational overhead) obtained using the Qiskit runtime service. Our quantum circuit uses all states (VQC$_1$ executes Hadamard gates on all $n$ qubits, which corresponds to the high temperature limit) and VQC$_2$ is a hardware efficient variational circuit with linear entanglement and Pauli $y$ and $z$ rotations. The equivalent $2^n$ hVQT circuits prepare the input states using Pauli $x$ gates and use the same VQC$_2$ circuit as the qVQT. For both VQT methods the total number of shots is chosen to be $2^{12} = 4096$, and is distributed evenly among all hVQT circuits. Note that we disable any error mitigation strategies.

Figure 2 shows that, in the case of hVQT (independent of the backend, either statevector simulator, ‘hVQT_SV’, or real hardware, ‘hVQT_oslo’), the runtime for large number of qubits becomes proportional to the number of states $2^n$ taken into account. The reason of this disadvantageous scaling stems from the fact that transpiling and classical pre (and post) processing times are proportional to the number of distinct circuits that need to be executed. This leads to an exponential scaling in the number of qubits since all states are taken into account. In practical applications (at finite temperatures) one only takes $M$ states into account and the runtime scales linearly in $M$ for large $M$. In contrast, the qVQT runtimes are completely independent of the number of states taken into account in VQC$_1$, since they only execute one specific circuit (see ‘qVQT_SV’ for statevector simulator, or real hardware ‘qVQT_oslo’).

In fact, the advantage of qVQT over hVQT is temperature dependent. At 0 K, qVQT and hVQT are both merely evaluating the ground state energy and are thus similar to VQE. When the temperature increases, qVQT gains a significant advantage over hVQT, as demonstrated with our real hardware experiments. Our ultimate performance metric is the overall time-to-solution to reach the minimum of the optimization process, which is essentially the product of the runtime for one evaluation of the free energy and the total number of iterations of the optimizer. To compare the performance between qVQT and hVQT, we argue that hVQT and qVQT optimize the same objective function for the same circuit, and that therefore the number of iterations (for a gradient free optimizer) is expected to be identical. This applies particularly to the system studied in this manuscript, where the first variational circuit with only one rotation layer is equivalent to the simple factorized models used in [29]. Thus, analyzing the runtime for one single free energy evaluation is sufficient to compare the performance between qVQT and hVQT.

Note that when using a gradient-based optimizer, the time-to-solution depends also on the runtime of the gradient evaluation, which can be rather costly. To address this issue, we propose a hybrid qVQT–hVQT approach to further accelerate gradient-based optimizations: for the first variational circuit in qVQT, use a circuit which has an equivalent Boltzmann machine giving the same probability distribution from hVQT (similar to the circuit we use and the equivalent factorized model). This allows to evaluate the gradient of the parameters of VQC$_1$ using back propagation methods, while the gradient evaluation of the parameters of VQC$_2$ benefits from the runtime advantage of qVQT.

We assess the performance of hVQT and qVQT on real systems including the noise of the hardware by evaluating the free energy of the Heisenberg chain from section 3.1 for $\beta = 1.0$. We choose a simple first variational circuit consisting of a layer of single qubit rotation gates and vary the depth of the second variational circuit, i.e. we increased the number of repetitions of the hardware efficient SU2 circuit implemented in Qiskit. In these experiments, we perform the optimization of the variational parameters on a noisless state vector simulator for 16 random starting points and compare the final free energy evaluation for the best over all the 16 runs. While the qVQT is completely executed on the ‘ibm_cairo’ device of IBM with 10,000 shots and a mid circuit measurement, the hVQT calculates the probability distribution by
performing a noiseless state vector simulation of the first circuit and then preparing the input states for the second variational circuit on the hardware. For hVQT, the energy evaluation is executed on the same qubits of the ‘ibm_cairo’ chip of IBM as for qVQT, and the shots are distributed according to the probabilities and also sum up to 10 000.

In order to obtain meaningful results on the hardware, we implement readout error correction using matrix-free measurement mitigation [37] of the mthree library [38] and zero noise extrapolation using the prototype-zne package [39], both for qVQT and hVQT. In figure 3 we show the obtained free energies and, for qVQT, we include approximate error bars which represent the standard deviations $\sigma$ from six distinct runs. We did not gather statistical error bars for hVQT by performing multiple runs due to its significantly larger runtime, but figure 3 shows clearly that the hVQT and state-vector results lie well within the $2\sigma$ error range of qVQT. Thus, we conclude that qVQT performs similarly well to hVQT on the ‘ibm_cairo’ device in terms of accuracy. Note that the main source of error is the measurement of the energy, which is independent of the method to prepare the density matrix on the quantum device.

The residual differences between the results from hVQT and qVQT where the free energies do not lie within the $1\sigma$ error range can be attributed to several factors:
qVQT and hVQT differ by the method chosen to prepare the input bit strings for the second variational circuit. It is possible that the preparation by a measurement process is more accurate than the preparation using rotation gates.

- The errors of the devices might change over time and thus lead to differences in the free energy.
- While hVQT calculates the entropy exactly and agrees with the state vector entropy, the qVQT exhibit errors in the entropy calculation due to the finite number of shots and the noise of the system. However, it appears that this error (in $S$) has the opposite sign of the error of the energy ($E$), leading to a favorable error cancellation in qVQT.

With respect to runtime, hVQT needs $2.7 \times$ more quantum resources than qVQT when excluding the time for the calibration of the M3 error mitigation, and including the transpilation time (for example when performing runs in Qiskit runtime) hVQT requires roughly $18 \times$ more quantum compute resources than qVQT.

### 3. Results and discussion

To demonstrate the utility of the qVQT we investigate two model systems: a 1D Heisenberg chain and a 2D J1–J2 Heisenberg model. For this purpose, we implement the qVQT algorithm using toolchains provided by Qiskit [40] and its associated statevector simulator. In the following the term ‘experiments’ refers to noise-less statevector simulations on classical hardware (performance assessments based on real NISQ hardware are discussed in section 2.5).

For all numerical experiments we consider three performance metrics (similar to [29]). The first one is given by the difference between the numerically computed and the exact free energy $\Delta F$, while the second one is given by the fidelity $f$ as

$$f(\hat{\rho}_1, \hat{\rho}_2) = \left( \text{Tr} \left( \sqrt{\sqrt{\hat{\rho}_1} \hat{\rho}_2 \sqrt{\hat{\rho}_1}} \right) \right)^2.$$  \hspace{1cm} (15)

To obtain a criteria which vanishes as the density matrix $\hat{\rho}$ approaches the Gibbs state $\hat{\rho}_{\text{Gibbs}}$ we use the metric $f_m = 1 - f(\hat{\rho}, \hat{\rho}_{\text{Gibbs}})$. The third metric we use is the trace distance:

$$Td(\hat{\rho}, \hat{\rho}_{\text{Gibbs}}) = \frac{1}{2} \text{Tr} \left( \sqrt{\hat{\rho} \hat{\rho}_{\text{Gibbs}}} \hat{\rho}_{\text{Gibbs}} \hat{\rho} \right).$$  \hspace{1cm} (16)

All three metrics vanish in the limit of ideal performance.

#### 3.1. 1D Heisenberg chain with transverse fields

The first model system we investigate is the 1D Heisenberg chain with transverse fields and nearest neighbor hopping, given by the Hamiltonian:

$$H = \sum_{\langle ij \rangle} J \left[ \sigma_i^x \sigma_j^x + \sigma_i^y \sigma_j^y + \sigma_i^z \sigma_j^z \right] + \sum_i \left[ J_x \sigma_i^x + J_z \sigma_i^z \right],$$  \hspace{1cm} (17)

where $\sigma_i^{\{x,y,z\}}$ denotes a Pauli $\{x,y,z\}$ operator on qubit $i$ and the first sum runs over all pairs $\langle ij \rangle$ of nearest neighbors. Verdon et al [29] analyzed this particular model for four-qubits with $J = -1, J_x = 0.2$ and $J_z = 0.3$, using the hVQT. To allow a direct comparison we employ the same parameters and use the qVQT to calculate the thermal state at the inverse temperature of $\beta = 1.3$, which Verdon et al pointed out to be the most challenging.

#### 3.1.1. Required circuit depth

For a qVQT with a minimal entropy circuit, i.e. where the circuit $\text{VQC}_1$ only contains a Pauli $x$ rotation on each qubit, we already obtain accurate results with fast convergence. Increasing the number of variational parameters for the energy circuit $\text{VQC}_2$, which is chosen to be a hardware efficient ansatz using Pauli $y$ and $z$ rotations with linear entanglement, further improves accuracy. We perform a statistical analysis with different total number of variational parameters by conducting 100 runs starting from random initial parameters, and using the limited-memory Broyden–Fletcher–Goldfarb–Shanno bound optimizer (L_BFGS_B) [41, 42] as implemented in Qiskit with a target gradient tolerance of $1 \times 10^{-3} \text{ rad}^{-1}$. The three performance metrics of these statistical experiments are shown in figure 4, illustrating that the qVQT yields an approximation to the density matrix that can be improved in accuracy by increasing the circuit depth and the associated computational cost. The corresponding scaling with respect to the required number of variational parameters is linear, as suggested by our numerical experiments (see SM, section A).
Figure 4. Results of a four-qubit 1D Heisenberg chain with transverse fields at $\beta = 1.3$, showing the top 20th percentiles of the three performance metrics as a function of the total number of parameters. The difference between the computed and exact values of the free energy $\Delta F$, the $f_m$-fidelity measure between the computed and exact density matrix, and the trace distance $Td$ between the computed and exact density matrix are shown in green, red, and blue, respectively.

Figure 5. The temperature dependence of a four-qubit 1D Heisenberg chain with transverse fields, showing the three performance metrics as a function of the inverse temperature. The color coding corresponds to the one in figure 4. The bottom line within each criteria corresponds to the best result obtained from 100 runs starting from random parameters, using the L_BFGS_B optimizer with a gradient tolerance of $1 \times 10^{-3}$ rad$^{-1}$. The dark shaded region denotes the range of the best 20-percentile of all runs, while the light-shaded region denotes the interval up to the average over all runs. The dashed vertical line denotes the value of $\beta = 1.3$ used in section 3.1.1.

3.1.2. Temperature dependence

In figure 5 we show the dependence of the three performance metrics on the inverse temperature. In the low-temperature limit ($\beta \gg 1$) the ground state is dominant and the accuracy improves as the algorithm does not need to calculate the excited states very precisely. In the high-temperature limit ($\beta \ll 1$), on the other hand, the splitting of all eigenstates becomes less important in comparison to the entropy. Clearly, temperatures around $\beta = 1$ are the most challenging for the algorithm as both classical and quantum-mechanical correlation need to be correctly captured.

Since the qVQT algorithm is similar to a hVQT and mainly differs by the method to produce the classical probability distribution, the temperature dependence above is comparable to the results of Verdon et al [29]. However, and most importantly, by switching from hVQT to qVQT a significantly smaller number of quantum circuits needs to be executed on the quantum device due to the intermediate measurement, a key advantage of the qVQT (see section 2.5).
3.2. 2D J₁–J₂ Heisenberg model

The second model system we investigate is the Heisenberg model with nearest and next-nearest neighbor interactions. Such systems have been extensively used to simulate and better understand the behavior of real magnetic materials that can be mapped to a spin model [43]. The Hamiltonian of this model is given by:

\[ H = \sum_{\langle i, j \rangle} J_1 S_i^z S_j^z + \sum_{\langle\langle i, j \rangle\rangle} J_2 S_i^z S_j^z, \]  

where the first and second summations run over nearest and next-nearest neighbors, respectively. It is well known that this system develops three phases, depending on the relative interaction parameter \( \alpha \), defined as \( J_1 = \sin(\alpha) \) and \( J_2 = \cos(\alpha) \), which introduces a normalization of \( J_1^2 + J_2^2 = 1 \). When \( J_2 > 0 \) and the next nearest neighbor interactions are stronger than the nearest neighbor interactions, the spins form a stripe configuration. When the nearest neighbor interactions are more important or \( J_2 < 0 \), the system turns ferromagnetic (FM) for \( J_1 < 0 \) and antiferromagnetic (AFM) for \( J_1 > 0 \).

To distinguish these three phases we construct two correlation functions that serve as order parameters:

\[ c_0 = \frac{1}{N_{nn}} \sum_{\langle i, j \rangle} \langle \sigma_i^z \sigma_j^z \rangle, \]
\[ c_1 = \frac{1}{N_{nnn}} \sum_{\langle\langle i, j \rangle\rangle} \langle \sigma_i^z \sigma_j^z \rangle, \]

where \( c_0 \) is the nearest neighbor correlation function averaged over all \( N_{nn} \) pairs of nearest neighbors, and, analogously, \( c_1 \) is the next-nearest neighbor correlation function. Note that the Hamiltonian is symmetric for rotations of all spins in the same way, i.e. for every eigenstate also the rotated states are eigenstates of the Hamiltonian with the same energy. For the term used to calculate the correlation functions this translates to \( \langle \sigma_i^z \sigma_j^z \rangle = \frac{1}{3} \sum S_i S_j \). Because the eigenvalues of \( \sum S_i S_j \) range between \(-3\) and \(1\), the correlation functions can assume values in the interval \([-1, 1/3]\). A correlation function which is greater than zero means that the FM contribution is dominating the AFM one, while for a negative correlation function the AFM contribution is dominant.

We perform numerical qVQT experiments using state vector simulators for a four-qubit, 2D J₁–J₂ Heisenberg lattice at a range of parameter angles \( \alpha \) and an inverse temperature of \( \beta = 1 \) to obtain the eigenstates of the Hamiltonian and, from those, the phase diagram. At each value of \( \alpha \) we perform 100 runs, starting with random initial parameters, again using the L_BFGS_B optimizer with a gradient convergence tolerance of \( 1 \times 10^{-3} \text{ rad}^{-1} \). Our qVQT algorithm uses two hardware efficient variational circuits with depths of 2 and 7 for VQC₁ and VQC₂, respectively, which results in a total of 76 variational parameters. VQC₁ includes only Pauli \( y \) rotations, while VQC₂ includes Pauli \( y \) and \( z \) rotations. Both use linear entanglement and include an additional final rotation gate layer.

The exact correlation functions and the results obtained by the numerical qVQT-experiments at \( \beta = 1 \) are shown in figure 6. We clearly see the three phases together with the corresponding phase transitions. For almost all parameter angles the numerical experiments match the exact values very well, except for the region near the phase transition between the FM and the AFM phase at \( J_1 = 0 \) and \( J_2 = -1 \). Interestingly for this model, these increased errors in the qVQT calculation seem to be of similar magnitude on both sides of the transition but exhibit opposite signs and therefore cancel each other out, i.e. do not influence the transition angle where the correlation function \( c_0 \) vanishes, the point that we identify with the phase transition. Two possible explanations for the source of these errors are: (a) either the approximation of the classical probability distribution breaks down in this regime, or (b) the splitting of the energy eigenstates cannot be assessed precisely when the FM and AFM states are in the same energy domain. The exact reasoning remains an open question and is the subject of future analysis.

Next, we compute the temperature dependent phase diagram of the Heisenberg model. For this purpose we use the results from our qVQT-calculation at \( \beta = 1 \) and monitor the energy spectrum to calculate the correlation functions of each eigenstate independently. The correlation functions \( c_0 \) and \( c_1 \) are then obtained by mixing the correlation functions of the eigenstates according to the probabilities \( p_i(\beta) \) given in equation (1):

\[ c_{0/1}(\beta) = \sum_{i=0}^{\beta} p_i(\beta) c_{0/1}^{i}. \]

The notation \( c_{0/1}^{i} \) corresponds to the value of the correlation function \( c_0 \) (or \( c_1 \)) evaluated on the eigenstate \( \varphi_i \). This approach has the advantage that the temperature-dependent results can be estimated
Figure 6. Nearest neighbor $c_0$ and next-nearest neighbor correlation function $c_1$ of a four-qubit 2D $J_1$–$J_2$ Heisenberg model as a function of the parameter angle $\alpha$ at inverse temperature $\beta = 1$. The lines are the results obtained by exact diagonalization, while the crosses denote the qVQT-results.

Figure 7. The phase diagram of a four-qubit 2D Heisenberg model, obtained from a qVQT experiment (dashed red) together with the exact ground truth (solid yellow). The phase transitions are given by the angles where the correlation functions vanish. The colormap indicates the combined deviation of the qVQT correlation functions from the exact results, $\Delta_c = |c_{\text{theo}} - c_{\text{qVQT}}| + |c_{\text{theo}} - c_{\text{qVQT}}|$, with linear interpolation between all data points. The explicit qVQT results are calculated on the white crosses at $T = 1$ and classically extended to temperatures $T \neq 1$.

without explicitly performing an optimization at each temperature, but comes with the drawback that the precision decreases when the temperature differs significantly from the original optimization temperature. For example at higher temperatures the errors on higher excited states get more pronounced as the probability $p_i(\beta)$ increases.
This behavior is reflected in figure 7, which shows the phase boundaries as a function of temperature computed from our numerical experiments and the exact results obtained by diagonalization of the Hamiltonian. Note that the agreement is excellent for $\beta \approx 1$, while the accuracy decreases as we deviate from this particular inverse temperature. Nevertheless, the overall phase behavior is captured correctly, demonstrating that our extension from the qVQT-calculation at $\beta = 1$ reproduces the physics with sufficient accuracy.

4. Conclusion

In summary, we present a new variational quantum algorithm, called qVQT, which is an extension of the VQE to finite temperatures. Our approach expands on the idea of the hVQT, but implements both the entropic and energetic contribution to the free energy on a quantum circuit. In this way we effectively reduce communication between classical and quantum device and the number of executed quantum circuits to compute an accurate Gibbs state.

We demonstrate the utility of the qVQT by performing extensive numerical experiments on quantum simulators for two model systems and show that our algorithm is well suited to calculate finite temperature properties or excited states on a quantum computer. The resource requirements as well as the scaling behavior are comparable to VQEs (see also section B in the SM), and we expect our algorithm to perform equally well for a given problem size. Hence, the qVQT provides a powerful tool to study quantum systems at finite temperature, producing accurate results with resources available on current NISQ devices for a wide range of applications.

Data availability statement

All data that support the findings of this study are included within the article (and any supplementary files).

Acknowledgments

We thank the members of the MANIQU consortium for valuable expert discussions. J S, M A and T E gratefully acknowledge support from the German Federal Ministry of Education and Research (BMBF) under Project No. 13N15574.

Code availability

The code that supports the findings of this study is available from the authors upon reasonable request.

Author contributions

J S developed and implemented the algorithms, M A supervised the research, T E conceived the project. J S and M A took the lead in writing the manuscript. All authors provided critical feedback and helped shape the research, analysis, and manuscript.

Conflict of interest

The authors declare no competing interests.

ORCID iDs

Johannes Selisko  https://orcid.org/0000-0001-6733-0782
Maximilian Amsler  https://orcid.org/0000-0001-8350-2476
Thomas Hammerschmidt  https://orcid.org/0000-0002-2270-4469
Ralf Drautz  https://orcid.org/0000-0001-7101-8804

References

[1] Arute F et al 2019 Quantum supremacy using a programmable superconducting processor Nature 574 505–10
[2] Wu Y et al 2021 Strong quantum computational advantage using a superconducting quantum processor Phys. Rev. Lett. 127 180501
[3] Arute F et al (Google AI Quantum And Collaborators) 2020 Hartree-Fock on a superconducting qubit quantum computer Science 369 1084–9
[4] Huggins W J, O’Gorman B A, Rubin N C, Reichman D R, Babbush R and Lee J 2022 Unbiasing fermionic quantum Monte Carlo with a quantum computer Nature 603 416–20
[5] Lloyd S 1996 Universal quantum simulators Science 273 1073–8
[6] Bauer B, Wecker D, Millis A J, Hastings M B and Troyer M 2016 Hybrid quantum-classical approach to correlated materials Phys. Rev. X 6 031045
[7] Shor P W 1996 Fault-tolerant quantum computation Proc. 37th Conf. on Foundations of Computer Science pp 56–65
[8] Preskill J 2018 Quantum computing in the NISQ era and beyond Quantum 2 79
[9] Kitaev A Y 1995 Quantum measurements and the Abelian Stabilizer problem (arXiv:quant-ph/9511026
[10] Coppersmith D 2002 An approximate Fourier transform useful in quantum factoring (arXiv:quant-ph/0201067)
[11] 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 1–7
[12] Tilly J et al 2022 The variational quantum eigensolver: a review of methods and best practices Phys. Rep. 986 1–128
[13] Cerezo M et al 2021 Variational quantum algorithms Nat. Rev. Phys. 3 625–44
[14] McClean J R, Boixo S, Smelyanskiy V N, Babbush R and Neven H 2019 Subspace-search variational quantum eigensolver for excited states Phys. Rev. X 9 031045
[15] Wang S, Fontana E, Cerezo M, Sharma K, Sone A, Cincio L and Coles P J 2021 Noise-induced barren plateaus in variational quantum eigensolver for small molecules and quantum magnets Nature 549 242–6
[16] Kuroiwa K and Nakagawa Y O 2020 Penalty methods for a variational quantum eigensolver Phys. Rev. Res. 3 013197
[17] Higgott O, Wang D and Briereley S 2019 Variational quantum computation of excited states Quantum 3 156
[18] Nakanishi K M, Mitarai K and Fujii K 2020 Efficient symmetry-preserving state preparation Quantum Sci. Technol. 9 1–6
[19] 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
[20] Kandala A, Mezzacapo A, Temme K, Takita M, Brink M, Chow J M and Gambetta J M 2017 Hardware-efficient variational quantum eigensolver Phys. Rev. Lett. 122 230501
[21] Wu J and Hsieh T H 2019 Variational thermal quantum simulation via thermofield double states Phys. Rev. Lett. 123 220502
[22] Zoufal C, Lucchi A and Woerner S 2021 Hybrid quantum-classical approach to correlated materials Mach. Learn.: Sci. Technol. 2 028011
[23] Fedorov D A, Peng B, Govind N and Alexeev Y 2022 VQE method: a short survey and recent developments Mater. Theory 6 2
[24] Bard T, Zhu L, Barson G, Mayhall N J, Economou S E and Barnes E 2020 Efficient symmetry-preserving state preparation circuits for the variational quantum eigensolver algorithm npj Quantum Inf. 6 1–9
[25] Verdon G, Marks J, Nanda S, Leichenauer S and Hidary J 2019 Quantum Hamiltonian-based models and the variational quantum eigensolver algorithm (arXiv:1910.02071)
[26] Motta M et al 2020 Determining eigenstates and thermal states on a quantum computer using quantum information theory Proc. Nat. Phys. 16 205–10
[27] Nishit H, Kosugi T and Matsushita Y I 2021 Implementation of quantum imaginary-time evolution method on NISQ devices by introducing nonlocal approximation npj Quantum Inf. 7 1–7
[28] Zoufal C, Lucchi A and Woerner S 2021 Variational quantum Boltzmann machines Quantum Mach. Intell. 3 7
[29] Perdomo-Ortiz A, Benedetti M, Reale-Gómez J and Biswas R 2018 Opportunities and challenges for quantum-assisted machine learning in near-term quantum computers Quantum Sci. Technol. 3 035002
[30] Guo X Y et al 2023 Variational quantum simulation of thermal statistical states on a superconducting quantum processor Chinese Phys. B 32 010307
[31] Liu J G, Mao L, Zhang P and Wang L 2021 Solving quantum statistical mechanics with variational autoregressive networks and quantum circuits Mach. Learn.: Sci. Technol. 2 028011
[32] Foldager J, Pesah A and Hansen L K 2022 Noise-assisted variational quantum thermalization Sci. Rep. 12 3862
[33] Alcazar J, Leyton-Ortega V and Perdomo-Ortiz A 2020 Classical versus quantum models in machine learning: insights from a finance application Mach. Learn.: Sci. Technol. 1 035003
[34] Fedorov D A, Peng B, Govind N and Alexeev Y 2022 VQE method: a short survey and recent developments Mater. Theory 6 2
[35] Bard T, Zhu L, Barson G, Mayhall N J, Economou S E and Barnes E 2020 Efficient symmetry-preserving state preparation circuits for the variational quantum eigensolver algorithm npj Quantum Inf. 6 1–9
[36] Nation P D, Kang H, Sundaresan N and Gambetta J M 2021 Scalable mitigation of measurement errors on quantum computers PRX Quantum 2 040326
[37] Nation P D and Treinish M 2023 mthree (GitHub) (available at: https://github.com/qiskit-Extensions/mthree)
[38] Rivero P, Metz F, Hasan A, Branczyk A M and Johnson C 2023 Zero noise extrapolation prototype (GitHub) (available at: https://github.com/qiskit-community/prototype-zne
[39] Anis M S et al 2021 Qiskit: an open-source framework for quantum computing (Zenodo) (arXiv:2106.03446)
[40] Byrd R H, Lu P, Nocedal J and Zhu C 1999 Limited memory algorithm for bound constrained optimization SIAM J. Sci. Comput. 16 1190–208
[41] Zhu C, Byrd R H, Lu P and Nocedal J 1999 Algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound-constrained optimization ACM Trans. Math. Softw. 23 550–60
[42] Mikhailyevsky A V, Shvartsev A V, Valiulin V E and Barabanov A F 2016 Thermodynamic properties of the 2D frustrated Heisenberg model for the entire J1–J2 circle J. Magn. Magn. Mater. 419 131–9