PERSPECTIVE • OPEN ACCESS

Quantum optimization using variational algorithms on near-term quantum devices

To cite this article: Nikolaj Moll *et al* 2018 *Quantum Sci. Technol.* **3** 030503

View the article online for updates and enhancements.

Quantum Science and Technology

PERSPECTIVE

Quantum optimization using variational algorithms on near-term quantum devices

Nikolaj Moll[1,3] , Panagiotis Barkoutsos[1], Lev S Bishop[2], Jerry M Chow[2], Andrew Cross[2], Daniel J Egger[1], Stefan Filipp[1], Andreas Fuhrer[1], Jay M Gambetta[2], Marc Ganzhorn[1], Abhinav Kandala[2], Antonio Mezzacapo[2], Peter Müller[1], Walter Riess[1], Gian Salis[1] , John Smolin[2], Ivano Tavernelli[1] and Kristan Temme[2]

[1] IBM Research—Zurich, Säumerstrasse 4, 8803 Rüschlikon, Switzerland
[2] IBM T.J. Watson Research Center, Yorktown Heights, NY 10598, United States of America
[3] Author to whom any correspondence should be addressed.

Abstract

Universal fault-tolerant quantum computers will require error-free execution of long sequences of quantum gate operations, which is expected to involve millions of physical qubits. Before the full power of such machines will be available, near-term quantum devices will provide several hundred qubits and limited error correction. Still, there is a realistic prospect to run useful algorithms within the limited circuit depth of such devices. Particularly promising are optimization algorithms that follow a hybrid approach: the aim is to steer a highly entangled state on a quantum system to a target state that minimizes a cost function via variation of some gate parameters. This variational approach can be used both for classical optimization problems as well as for problems in quantum chemistry. The challenge is to converge to the target state given the limited coherence time and connectivity of the qubits. In this context, the *quantum volume* as a metric to compare the power of near-term quantum devices is discussed. With focus on chemistry applications, a general description of variational algorithms is provided and the mapping from fermions to qubits is explained. Coupled-cluster and heuristic trial wave-functions are considered for efficiently finding molecular ground states. Furthermore, simple error-mitigation schemes are introduced that could improve the accuracy of determining ground-state energies. Advancing these techniques may lead to near-term demonstrations of useful quantum computation with systems containing several hundred qubits.

1. Introduction

Recent advances in the field of quantum computing have boosted the hope that one day complex problems can be solved efficiently on quantum computers. The ultimate goal is a universal fault-tolerant quantum computer that runs arbitrary algorithms much faster than on a classical computer. However, millions of physical qubits and high-fidelity gate operations are required to implement a universal fault-tolerant quantum computer, a system that currently cannot be built. Yet, quantum devices with a couple of hundred physical qubits with limited or no error correction are likely to become available in the near future. With it comes the question how to exploit these devices for useful calculations. In this paper, we discuss how the variational quantum eigensolver (VQE) can be run on near-term quantum devices to tackle optimization problems that are exponentially hard on classical computers.

We differentiate between two types of optimization problems. The first kind are quantum optimization problems, such as finding the ground state of a complex molecule or the simulation of its dynamics. In this case, optimization typically involves minimization of the total energy as described by the energy expectation value of a non-trivial Hamiltonian as a function of some molecular parameters, such as interatomic distances. The second kind are classical optimization problems which can usually be mapped onto a relatively simple Ising-type
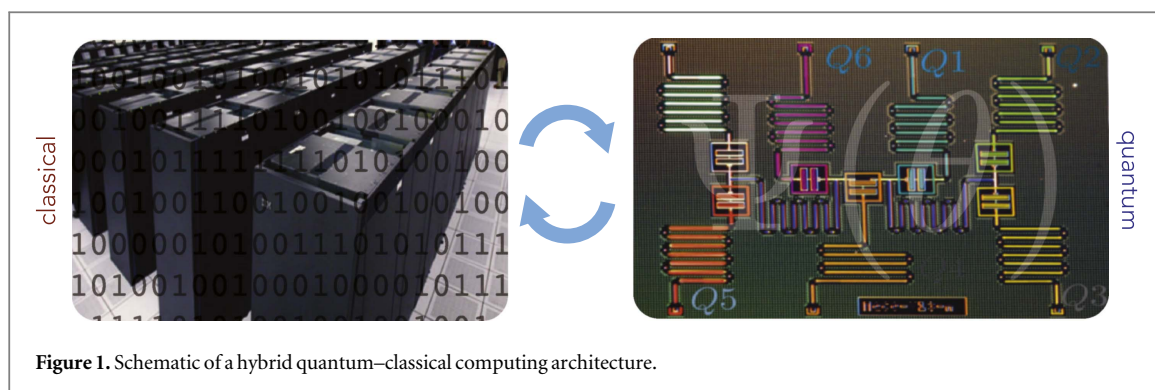
Figure 1. Schematic of a hybrid quantum–classical computing architecture.

Hamiltonian. In both cases, exponential scaling of the required computational resources with the problem size can make the problems hard to solve or even intractable on classical computers.

Generally, optimization problems are solved by finding the extremum of an objective function, such as cost, energy, profit or error. As the cost function typically depends on a large set of parameters, finding a solution involves searching a high-dimensional parameter space, which quickly makes a brute-force approach unfeasible. A quantum computer operates on Hilbert space, which grows exponentially as 2^N with the number of qubits N. The idea is to use this vast state space with the help of quantum entanglement, and thus boost the efficiency in finding the right solution, ideally with exponential speed-up [1–5]. A more careful analysis shows, however, that the speed-up for classical optimization problems is in many cases rather modest [6–8]. In contrast, one can benefit from quantum speed-up in problems that are directly related to the quantum–mechanical description of nature itself. A prominent example is finding the many-electron wavefunction of a molecular system. Classical computers fail to solve such problems exactly for more than a few tens of electrons because of the exponential increase of Hilbert space with the number of electrons. The large state space of a quantum computer can be used to simulate a chemical system and calculate its properties, including correlations and reaction rates, once the challenge of efficiently mapping the fermionic problem to the available qubit hardware is overcome.

In fact, on a quantum device the natural way is to solve the chemical system in second quantization [3, 4, 9–33] formulated in terms of fermionic annihilation and creation operators. Because of the different statistics there is no direct one-to-one mapping: each fermion operator must be represented by a string of qubit operators, which induces long-range qubit-qubit correlations in the system and places demanding requirements on the connectivity and the number of gates (see section 4.1). To compute the quantum evolution of chemical systems on a digital quantum computer, decomposition into discrete time steps is required and accordingly long gate sequences [3, 14, 34].

On current quantum devices, gate errors and decoherence restrict the number of sequential gate operations that can be performed while keeping a meaningful, coherent quantum state. Moreover, connectivity between qubits is limited by the physical routing of the wires on a qubit chip. This is why a new class of hybrid classical quantum algorithms, called the VQE [33–43], holds a lot of prospects for near-term quantum-computing systems (see figure 1). These algorithms work with short-depth circuits and will result in approximate results when the number of qubits, their coherence and the connectivity is large enough. These requirements on the quantum system can be quantified by the *quantum volume* [44], a hardware-independent figure or merit for the power of a quantum computer.

The VQE can be used both for classical optimization problems as well as for fermionic Hamiltonians describing, e.g., quantum chemistry. In quantum chemistry the VQE is used to calculate ground states [33–39] of chemical systems. The high-dimensional trial wavefunctions, which are costly to represent on a classical computer, are generated on the quantum computer using parametrized single-qubit and entangling gates. The optimization of the gate parameters is performed on a classical computer by summing expectation values of the qubit operators measured on the quantum device and thereby calculating the total energy as a cost function. This can in principle lead to very short-depth circuits which ideally run in a time that is shorter than the coherence time of the quantum computer. The same VQE can be applied to other physical systems in condensed matter such as the Fermi–Hubbard model [2, 12, 17, 45–48] and spin systems [49–52].

Hybrid algorithms are, however, not resilient against decoherence and gate errors, which may lead to inaccurate estimates of the expectation values. Currently available error-correction schemes, such as those based on surface codes [53], require a significant number of qubits, rendering quantum simulations of practical systems challenging in the near future. Still, novel schemes that do not require ancillas or code qubits can help mitigate induced errors, enabling longer and bigger quantum computations. Such error mitigation schemes

[54, 55] need to be developed further and tested to improve accuracy without the full overhead of error-correction codes for universal quantum computing.

The paper is structured as follows: the quantum volume is discussed in section 2 before we explain the VQE in section 3 and its application to quantum chemistry problems in section 4. After a brief discussion of the prospects of solving classical optimization problems with near-term quantum devices in section 5, we elaborate on the choice of suitable optimizers for the classical feedback in the VQE in section 6 and discuss the prospects of fighting back decoherence in near-term quantum devices without full error correction in section 7. Finally, we conclude in section 8.

2. Quantum volume, a metric for near-term quantum devices

For current quantum processors, various architectures and physical qubit realizations are being considered. While quantum systems based on superconducting qubits [34, 56–61] at the moment seem to be leading the way, ion-trap-based systems [62, 63] are close competitors. Furthermore, semiconductor-based spin qubits [64–66] and other quantum architectures [67–69] may still become important in the future. Given the different hardware implementations it is often difficult to benchmark the usefulness or power of quantum systems, which is why a hardware-independent measure is required. To define a suitable metric, we first note that a quantum computer's performance depends on five main hardware parameters:

(i) Number of physical qubits N.

(ii) Connectivity between qubits.

(iii) Number of gates that can be applied before errors or decoherence mask the result.

(iv) Available hardware gate set.
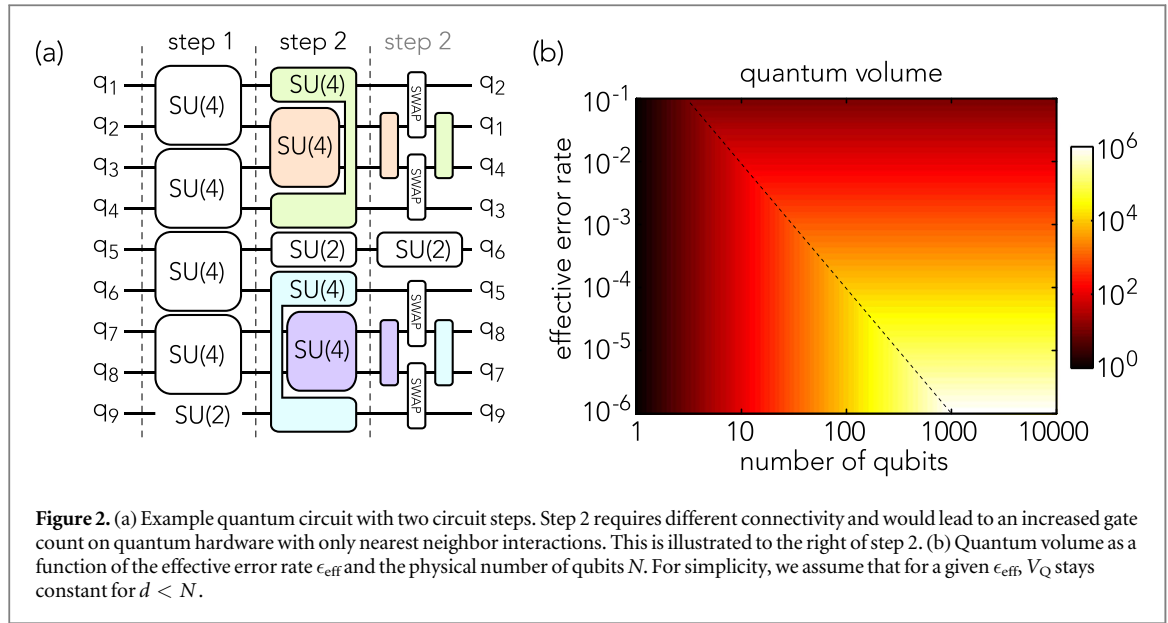
(v) Number of operations that can be run in parallel.

With the goal to quantify a quantum computer's power with a single parameter, we would like to consider a metric based on the question 'can this device run a given algorithm?'. For any given instance of a quantum algorithm, there is a lower bound on the number of qubits N required to run the algorithm, as well as the necessary number of steps (or circuit depth) d. We therefore define a *quantum volume* V_Q [44] that takes into account both the number of qubits N and the allowable depth d of quantum circuits that can be run on a near-term quantum device. In the simplest case, we could just choose the quantum volume to be $d \cdot N$; however, this has some undesirable properties in that it can be gamed in various ways. For example, in many cases the smallest error rates and therefore the largest circuit depth will result from very few qubits, even $N = 2$, as in this case there will be less connectivity and parallelization overhead and fewer issues with crosstalk between qubits. However, clearly $N = 2$ is a completely uninteresting limit. Also the other extreme, where a device has many qubits but little coherence, i.e. $d \approx 1$, is not interesting because such a system cannot use entanglement as a resource and calculations become effectively classical.

We therefore conceptually define the quantum volume as

$$\tilde{V}_Q = \min[N, d(N)]^2. \tag{1}$$

Here, the number of qubits N is an easily accessible hardware parameter; however, the achievable circuit depth $d(N)$ needs further specification in terms of the hardware parameters given in the list above.

We start by considering one step of a quantum algorithm (a depth-one circuit) on a number of N qubits. Such a step is expressed as a unitary operator that can be written as a tensor product of randomly chosen arbitrary two-qubit gates on disjoint pairs of qubits (see step 1 in figure 2(a)). Here, we allow any unitary two-qubit operation in the SU(4) group, which may consist of a combination of one- and two-qubit gates on the actual hardware. Then an effective error rate $\epsilon_{\rm eff}$ is defined as the error rate per two-qubit gate averaged over many realizations of such depth-one circuits. Therefore, $\epsilon_{\rm eff}$ depends on the gate overhead required when all-to-all connectivity, full parallelism and a suitable gate set is not available. Thereby, it also encapsulates both the errors of single- and two-qubit gates. If the hardware supports all possible two-qubit gates directly (requiring an all-to-all connectivity) with identical error rate ϵ, and in addition allows unlimited gate parallelism, then $\epsilon_{\rm eff} = \epsilon$. If the connectivity is limited, then it will be necessary to insert additional SWAP gates to permute the qubits in order to implement the random two-qubit gates, leading to an increase of $\epsilon_{\rm eff} > \epsilon$. A planar nearest-neighbor qubit coupling would lead to an effective error rate of $\epsilon_{\rm eff} \propto \sqrt{N}\epsilon$, and a linear chain of qubits would yield an effective error rate of $\epsilon_{\rm eff} \propto N\epsilon$. On the other hand a hardware which supports more complex gates such as the Tofoli gate directly or the use of a compiler which efficiently compresses the gates of a test circuit could also lead

Figure 2. (a) Example quantum circuit with two circuit steps. Step 2 requires different connectivity and would lead to an increased gate count on quantum hardware with only nearest neighbor interactions. This is illustrated to the right of step 2. (b) Quantum volume as a function of the effective error rate ϵ_{eff} and the physical number of qubits N. For simplicity, we assume that for a given ϵ_{eff}, V_Q stays constant for $d < N$.

to a situation with $\epsilon_{\text{eff}} < \epsilon$. Other special features and limitations of the hardware must be dealt with in a similar manner.

The error rate of a single circuit step scales with the number of simultaneous two-qubit gates $\epsilon_{1\text{step}} \propto N\epsilon_{\text{eff}}$. In other words, we can estimate the circuit depth in which, on average, a single error occurs as $d \simeq 1/(N\epsilon_{\text{eff}})$, linking the effective error ϵ_{eff} to the previous definition of the quantum volume using the circuit depth. As an example, if an effective error rate $\epsilon_{\text{eff}} = 10^{-4}$ is experimentally achievable, depth $d = 10$ algorithms could be run on a 1000-qubit device, and $d = 100$ algorithms on a 100-qubit device.

However, the effective error rate ϵ_{eff} will depend not only on the gate error rates and the connectivity but, more generally, on the complexity of the quantum system which grows with the number of qubits, for example, because of crosstalk. The effective error rate $\epsilon_{\text{eff}}(N)$ will therefore likely be a function of N even if full connectivity is available. Moreover, ϵ_{eff} also depends on the sophistication of the scheduling algorithm responsible for mapping the quantum algorithm considered to the hardware. Both hardware and software improvements will thus impact the effective error rate $\epsilon_{\text{eff}}(N)$.

Finally, we note that with this definition the allowable circuit depth $d \simeq 1/(N\epsilon_{\text{eff}})$ decreases with N at constant effective error ϵ_{eff}, which means that a system's quantum volume decreases if more qubits with the same fidelity are made available on the hardware. However, a given algorithm does not necessarily need all N available qubits. It could even be beneficial for an algorithm that requires $n < N$ qubits to run on an N-qubit machine when selecting a subset of qubits with good connectivity is selected. We therefore further refine the definition of the quantum volume in equation (1):
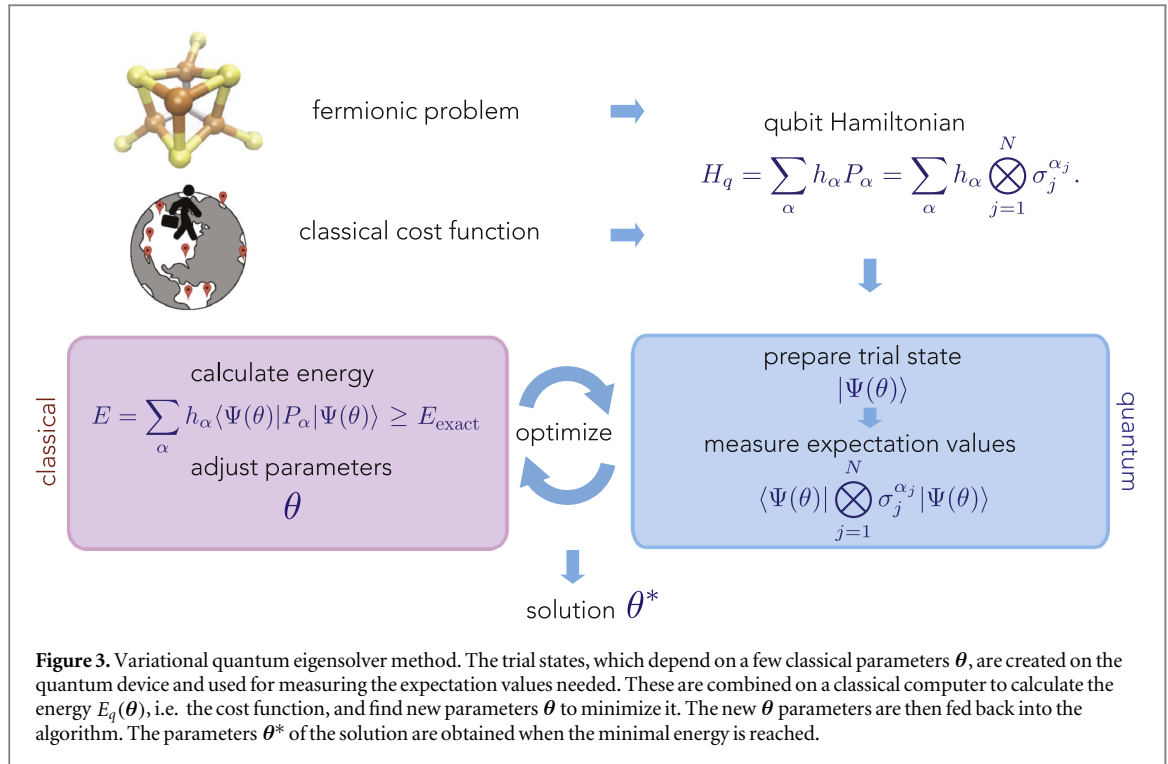
$$V_{\text{Q}} = \max_{n<N}\left(\min\left[n, \frac{1}{n\epsilon_{\text{eff}}(n)}\right]^2\right),\tag{2}$$

where the maximum is taken over an arbitrary choice of n qubits to maximize the quantum volume that can be obtained with such a subset.

We plot an example quantum circuit with two circuit steps and the functional dependence of the quantum volume on the number of qubits and an effective two-qubit error rate in figure 2. The dashed line denotes the tipping point where $d = 1/(N\epsilon_{\text{eff}}) = N$. From any point on this line, a significant increase in V_Q requires improvements in both ϵ_{eff} and N. We also see that the usefulness of current quantum devices is likely limited by the typical effective error rates, which are $\epsilon_{\text{eff}} > 10^{-3}$. To improve ϵ_{eff} we will have to start encoding quantum states in logical qubits with an overhead in the number of physical qubits. This will eventually lead to fault tolerant quantum computing.

The quantum volume is therefore an architecture-neutral metric that characterizes the capability of a chosen quantum computing architecture to run useful quantum circuits. It enables the comparison of hardware with widely different performance characteristics and quantifies the complexity of algorithms that can be run on such a system. An important conclusion that we can draw for the usefulness of near-term quantum devices is that when increasing the number of qubits the power of the quantum device will increase only if the effective error rate is improved at the same time.

Figure 3. Variational quantum eigensolver method. The trial states, which depend on a few classical parameters $\boldsymbol{\theta}$, are created on the quantum device and used for measuring the expectation values needed. These are combined on a classical computer to calculate the energy $E_q(\boldsymbol{\theta})$, i.e. the cost function, and find new parameters $\boldsymbol{\theta}$ to minimize it. The new $\boldsymbol{\theta}$ parameters are then fed back into the algorithm. The parameters $\boldsymbol{\theta}^*$ of the solution are obtained when the minimal energy is reached.

3. Exploring Hilbert space with the VQE

To exploit near-term quantum devices, applications and algorithms have to be tailored to current quantum hardware with only tens or hundreds of qubits and without full quantum error correction. One main constraint is the limited quantum volume that restricts the depth of meaningful quantum circuits. Still, a small-scale quantum computer with hundred qubits can process quantum states that cannot even be stored in any classical memory. A natural way to make use of this quantum advantage is via a hybrid quantum–classical architecture: a quantum co-processor prepares multi-qubit quantum states $|\Psi(\boldsymbol{\theta})\rangle$ parametrized by control parameters $\boldsymbol{\theta}$. The subsequent measurement of a cost function $E_q(\boldsymbol{\theta}) = \langle\Psi(\boldsymbol{\theta})|H_q|\Psi(\boldsymbol{\theta})\rangle$, typically the energy of a problem Hamiltonian H_q, serves a classical computer to find new values $\boldsymbol{\theta}$ in order to minimize $E_q(\boldsymbol{\theta})$ and find the ground-state energy

$$E_q^{\min} = \min_{\boldsymbol{\theta}}(\langle\Psi(\boldsymbol{\theta})|H_q|\Psi(\boldsymbol{\theta})\rangle). \tag{3}$$

This VQE approach to Hamiltonian-problem solving has been recently applied in different contexts [34, 37, 40, 70–72]. In fact, the Hamiltonian H_q can take many forms, the only requirement being that it can be mapped to a system of interacting qubits with a non-exponentially increasing number of terms. Here we distinguish two relevant cases: Hamiltonians that describe fermionic condensed-matter or molecular system (section 4) and Hamiltonians that describe a classical optimization problem (section 5).

3.1. Variational quantum eigensolver method

In detail, the VQE method consists of four main steps as shown in figure 3. First, on the quantum processor a tentative variational eigenstate, a trial state, $|\Psi(\boldsymbol{\theta})\rangle$ is generated by a sequence of gates parameterized by a set of control parameters $\boldsymbol{\theta}$. In the ideal case, this trial state depends on a small number of classical parameters $\boldsymbol{\theta}$, whereas the set of gates is chosen to efficiently explore Hilbert space. In particular, the class of states forming the solution to the minimization problem in equation (3) has to lie within the set of possible trial states. Suitable gate sets which provide a good approximation to the wanted target state, which minimizes the cost function, have been found for both classical optimization problems [41] (section 5) and quantum chemistry problems (section 4). Aside from these considerations, it is also essential that hardware constraints be taken into account. As not all gates are directly realizable in hardware, decomposing them into those available in the quantum hardware adds extra overhead in circuit depth. An alternative is, therefore, to use a heuristic approach based on gates that are readily available in hardware [72] as discussed below.

Second, once the trial state has been prepared and the expectation value of the problem Hamiltonian H_q is determined. The problem Hamiltonian can be decomposed into Pauli strings $P_\alpha = \sigma_1^{\alpha_1} \otimes \sigma_2^{\alpha_2} \otimes \ldots \sigma_N^{\alpha_N}$ with single-qubit Pauli operators and the identity operator $\mathbf{1}$, $\sigma_i^j \in \{\mathbf{1}, \sigma_i^x, \sigma_i^y, \sigma_i^z\}$ such that

$$H_q = \sum_\alpha h_\alpha P_\alpha. \tag{4}$$

N denotes the number of qubits. To determine the expectation value of each Pauli operator in P_α, each single qubit's population is measured repeatedly for a given number of experiments with identical trial state preparation $|\Psi(\boldsymbol{\theta})\rangle$. This corresponds to measuring σ_j^z for each qubit; other Pauli operators can be determined by applying a pre-rotation on the qubit before the measurement that effectively rotates the measurement axis. To determine the expectation value of the Pauli strings, the measurement outcomes are multiplied for each run of the experiment and then averaged.

In a third step, the cost function $E_q(\boldsymbol{\theta}) = \langle\Psi(\boldsymbol{\theta})|H_q|\Psi(\boldsymbol{\theta})\rangle = \sum_\alpha h_\alpha \langle\Psi(\boldsymbol{\theta})|P_\alpha|\Psi(\boldsymbol{\theta})\rangle$ is calculated by summing up the expectation values of P_α with corresponding coefficients h_α.

Finally, the value of $E_q(\boldsymbol{\theta})$ is minimized as a function of the parameters $\boldsymbol{\theta}$. A classical optimization algorithm processes $E_q(\boldsymbol{\theta})$ and provides new parameters $\boldsymbol{\theta}$. For each parameter set, a new set of gates for trial state preparation has to be loaded onto the quantum processor. As this requires rather time-consuming re-programming of the quantum hardware, it is important that only a minimal number of queries should be made to the quantum processor. Moreover, the calculated expectation values will be noisy because of the limited sampling statistics of the qubit state. Therefore, classical robust optimizers have to used that can handle the noise on the measured expectation values and scale favorably with the number of parameters as described in section 6. The procedure ends when the minimum of $E_q(\boldsymbol{\theta})$ in equation (3) is reached within a given accuracy and the optimal parameters $\boldsymbol{\theta}*$ are found.

4. Quantum chemistry with qubits

To demonstrate the potential of a quantum processor with limited quantum volume, one needs to consider quantum algorithms that provide a large scaling advantage compared with their classical counterparts. The solution of the electronic structure problem in quantum chemistry belongs to this class: because of the exponential scaling of the problem, it is impossible to find an exact solution to the Schrödinger equation of systems with more than a few tens of electrons on a classical computer. Several approximations have been introduced to access the properties of large-scale systems with more than 1000 electrons on high-performance computers. The aim is to reach the required accuracy for chemical energies ($\sim$50 meV). One approach is to approximate the many-electron Hamiltonian itself using, for example, density-functional theory [73]. There, the original system of interacting electrons is replaced by a fictitious one of non-interacting electrons moving in a modified external potential that allows, at least *in principle*, the original exact solution to be recovered.

An alternative approach starts from the exact Hamiltonian and attempts to find suitable approximations for the system wavefunction in the many-electron Hilbert space. This calculation can, in principle, be performed either within the first or the second quantization formalism. In first quantization, all spatial integrals have to be evaluated on the quantum computer. For this reason, approaches based on second quantization are more suited for first-generation quantum devices. In this case, all spatial integrals are evaluated beforehand on a classical computer, whereas the sampling of the Hilbert space is performed in the orbital configuration space spanned by molecular Slater determinants. This approach maps naturally to the variational method described above (section 3). It starts from the one-electron basis states that are obtained by solving the Hartree–Fock equation. These Hartree–Fock orbitals are then used to construct an anti-symmetrized product wavefunction, the Slater determinant, which is used as a starting point for a perturbative expansion. In this expansion a controlled series of *excited* configurations is added until a sufficiently accurate approximation of the ground state is found.

4.1. Mapping fermions to qubits
The electronic Hamiltonian in second quantization is given by

$$H_F = \sum_{ij} t_{ij} a_i^\dagger a_j + \sum_{ijkl} u_{ijkl} \, a_i^\dagger a_k^\dagger a_l a_j, \tag{5}$$

where the operators $a_i^\dagger$ and a_i create and annihilate electrons in the ith orbital. The parameters t_{ij} and u_{ijkl} describe the one- and two-electron interactions and can be efficiently computed classically as the overlap integrals of the orbitals in the basis set [74]. The two-electron term scales at most with the number of orbitals to the fourth power [4, 75] and therefore does not grow exponentially, which would prohibit efficient computation even on a quantum computer.

Because a_i and $a_i^\dagger$, unlike the Pauli spin operators, follow fermionic commutation rules $\{a_i, a_j\} = 0$, $\{a_i^\dagger, a_j^\dagger\} = 0$, $\{a_i, a_j^\dagger\} = \delta_{ij}$, a direct implementation of equation (5) on a qubit-based quantum processor is not feasible without a mapping from fermionic to Pauli operators. The fermionic nature of electrons implies that many-electron wavefunctions must be anti-symmetric with respect to particle exchange. This is reflected in the way fermionic creation and annhilation operators act on state vectors:

$$a_i^\dagger |f_1, \, ..., f_{i-1}, f_i, f_{i+1}, \, ..., f_N\rangle = \delta_{f_i,0} \, p_i \, |f_1, \, ..., f_{i-1}, 1, f_{i+1}, \, ..., f_N\rangle, \tag{6}$$

$$a_i |f_1, \, ..., f_{i-1}, f_i, f_{i+1}, \, ..., f_N\rangle = \delta_{f_i,1} \, p_i \, |f_1, \, ..., f_{i-1}, 0, f_{i+1}, \, ..., f_N\rangle. \tag{7}$$

Here $p_i = (-1)^{\sum_{k=0}^{i-1} f_k}$ denotes the parity and $f_i \in \{0, 1\}$ the occupation number of the fermionic orbital i. The naive replacement of the fermionic operators $a_i^{(\dagger)}$ by Pauli ladder operators $\sigma_i^\pm = (\sigma^x \pm i\sigma^y)/2$ does, however, not reproduce equations (7) because $\sigma_i^\pm$ describe distinguishable *particles* with no special symmetries.

A variety of mappings have been developed that guarantee that the fermion statistics are captured on a system of qubits [76–78]. Among those, the Jordan–Wigner mapping [79] is particularly intuitive: it is based on a one-to-one mapping of fermionic to qubit occupations, i.e. the occupancy information is stored *locally*. To take into account the parity information p_i in equations (7), fermionic operators are translated as

$$a_i^\dagger \to \mathbf{1}^{\otimes i-1} \otimes \sigma^- \otimes (\sigma^z)^{\otimes N-i}, \tag{8}$$

$$a_i \to \mathbf{1}^{\otimes i-1} \otimes \sigma^+ \otimes (\sigma^z)^{\otimes N-i}, \tag{9}$$

where N is the total number of qubits considered. It is obvious that calculating the parity when acting on qubit i requires the knowledge of all state occupations $j < i$, which is accomplished by the σ^z terms in equation (9). However, this introduces a non-locality in the mapping and, when inserted into the Hamiltonian in equation (5), gives rise to long sequences of σ^z operators intercalating between $\sigma^\pm$ operators of length k, known as k-local terms This means that a fermionic wavefunction is spread out over $\mathcal{O}(N)$ qubits, posing fidelity issues in the readout process of the expectation value of the Hamiltonian.

Recent schemes for tapering off qubits in mapped fermionic Hamiltonians [78, 80], based on fermionic symmetries, can partially alleviate the hardware requirements necessary for performing simulations of fermionic systems. These second-quantized tapering schemes exploit symmetries in the mapped qubit Hamiltonian to reduce the simulation space needed to host the mapped fermionic system.

The Jordan–Wigner transformation [79] consists of a local occupancy map and a non-local, $\mathcal{O}(N)$, parity function, whereas the binary-tree transformation encodes both operations on maps that scale $\mathcal{O}(\log(N))$ with the number of qubits [76–78], which is a clear advantage compared with the Jordan–Wigner transformation.

4.2. Coupled cluster trial wavefunctions

Once a mapping of fermions to qubit has been chosen, suitable trial states for the VQE have to be prepared on the quantum processor. At best, these trial states incorporate the structure of the problem Hamiltonian and known properties of the solution state, such as the total number N of fermions. While one could aim to find a gate set that allows one to generate all possible excited Slater determinant configurations, which is known as the full configuration interaction approach, the number of states scales factorially with the number of electrons, a clear obstacle for computing larger molecules. One way to improve the efficiency is to use a coupled-cluster approach for creating the trial states, which allows a systematic sampling of all relevant excited Slater determinants up to a given excitation degree. In conventional quantum chemistry, these coupled-cluster expansions are used as a benchmark for all other approaches.

In the unitary coupled-cluster (UCC) approach [81], which is a variational version of the commonly used coupled-cluster method [82], the unitary operator $U(\boldsymbol{\theta})$ that is used to generate a trial wavefunction $|\Psi(\boldsymbol{\theta})\rangle$ from the reference state $|\Phi\rangle$ is given by

$$|\Psi(\boldsymbol{\theta})\rangle = U(\boldsymbol{\theta})|\Phi\rangle = e^{T(\boldsymbol{\theta}) - T^\dagger(\boldsymbol{\theta})}|\Phi\rangle. \tag{10}$$

It is constructed by exponentiation of the cluster operator $T(\boldsymbol{\theta})$ defined as

$$T(\boldsymbol{\theta}) = \sum_k T_{(k)}(\boldsymbol{\theta}), \quad T_{(1)}(\boldsymbol{\theta}) = \sum_{\substack{i \in \mathrm{occ} \\ j \in \mathrm{unocc}}} \theta_{(i)}^{(j)} a_j^\dagger a_i, \quad T_{(2)}(\boldsymbol{\theta}) = \sum_{\substack{i,j \in \mathrm{occ} \\ l,k \in \mathrm{unocc}}} \theta_{(i,j)}^{(k,l)} a_l^\dagger a_k^\dagger a_j a_i, ... \tag{11}$$
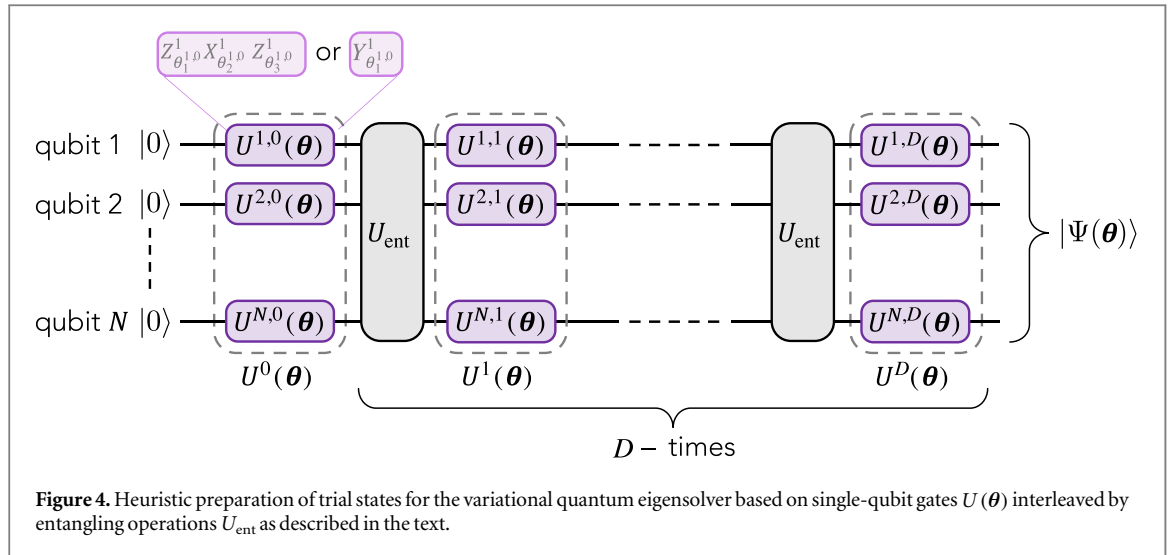
Here, the coefficients $\boldsymbol{\theta}$ describes a vector of parameters that will be optimized using VQE. A common choice for the reference state $|\Phi\rangle$ is the ground-state Slater-determinant made up of the lowest-energy molecular orbitals obtained from the solution of the Hartree–Fock equation.

The coefficients $\boldsymbol{\theta}$ of the cluster operators are not independent and their value decreases with the order of the excitation. Therefore, this expansion is typically truncated at the double (UCCSD) or triple level (UCCSDT) of excitation without significantly reducing the accuracy. In fact, the exponentiation of the cluster operator $T(\boldsymbol{\theta})$ introduces higher uncorrelated excitations at each level of truncation, e.g., for $T(\boldsymbol{\theta}) = T_{(1)}(\boldsymbol{\theta}) + T_{(2)}(\boldsymbol{\theta})$

$$e^{T(\boldsymbol{\theta})} = 1 + T_{(1)}(\boldsymbol{\theta}) + T_{(2)}(\boldsymbol{\theta}) + \frac{T_{(1)}^2(\boldsymbol{\theta})}{2!} + T_{(1)}(\boldsymbol{\theta}) T_{(2)}(\boldsymbol{\theta}) + \frac{T_{(2)}^2(\boldsymbol{\theta})}{2!} + ..., \tag{12}$$

the expansion produced triplet and quadruple excitations in the first few terms of the expansion (fifth and sixth terms, respectively). Despite the compactness of this expansion, the number of coefficients $\boldsymbol{\theta}$ increases already in UCCSD with the number of orbitals to the fourth power, which impacts the efficiency of the classical

Figure 4. Heuristic preparation of trial states for the variational quantum eigensolver based on single-qubit gates $U(\boldsymbol{\theta})$ interleaved by entangling operations U_{ent} as described in the text.

optimization of the trial state $|\Psi(\boldsymbol{\theta})\rangle$. In practice, in the case of large molecular systems the limited achievable circuit depth in current quantum devices requires a further truncation of the series in equation (12). Thus, while the coupled cluster method guarantees in principle an efficient convergence towards the exact ground state, its implementation in state-of-the-art quantum computers requires further studies in terms of how different approximations (truncations) affect the accuracy of the solution.

4.3. Hardware-efficient trial states suitable for near-term quantum hardware

A much simpler approach is, therefore, the heuristic generation of the trial state with unitary operations that are more suited to the available quantum hardware [72]. Independently of the particular problem to be solved, one may choose trial states that can be efficiently generated in current quantum hardware and at the same time allow the generation of highly entangled states that are close to the target state.

This approach is showcased in the examples provided in sections 4.4 and 5.2. As shown in figure 4, the preparation of the heuristic trial states comprises two types of quantum gates, single-qubit Euler rotations $U(\boldsymbol{\theta})$ determined by the rotation angles $\boldsymbol{\theta}$ and an entangling *drift* operation U_{ent} acting on pairs of qubits. The N-qubit trial states are obtained by applying a sequence of D entanglers U_{ent} alternating with the Euler rotations on the N qubits to the initial ground state $|00\ldots0\rangle$,

$$|\Phi(\boldsymbol{\theta})\rangle = \overbrace{U^D(\boldsymbol{\theta})\, U_{\mathrm{ent}} \ldots\, U^1(\boldsymbol{\theta})\, U_{\mathrm{ent}}}^{D\text{-times}}\, U^0(\boldsymbol{\theta})|00\ldots0\rangle. \tag{13}$$

This gate sequence has a total number of $p = N(3D + 2)$ independent angles.

To be more specific, the single-qubit operations are decomposed into rotations about the x- and the z-axes, $U^{q,i}(\boldsymbol{\theta}) = Z^q_{\theta^{q,i}_1} X^q_{\theta^{q,i}_2} Z^q_{\theta^{q,i}_3}$, with $X^q(\theta^{q,i}_j) = \exp[-\mathrm{i}\theta^{q,i}_j \sigma^x_q/2]$ (and similarly for Y and Z) denoting the unitary operation acting on qubit q at the ith position in the gate sequences. The heuristic approach does not rely on the accurate implementation of specific two-qubit gates and can be used with any U_{ent} that generates sufficient entanglement. A natural choice can be the cross-resonance gate [83, 84] as a two-qubit gate suited for the fixed-frequency superconducting qubit architecture as used, for example, for the IBM Q experience [61].

4.4. Small molecules calculated with the VQE

As an application of the method described above, we present the calculation of the ground-state energy of simple molecules such as the hydrogen molecule: the starting point is the Hamiltonian in second quantization in equation (5) with the one-body terms, t_{ij}, representing the kinetic energy of the electrons and the potential energy that they experience in the presence of the nuclei,

$$t_{ij} = \int \mathrm{d}\boldsymbol{x}_1 \phi^*_i(r_1)\left(-\frac{\boldsymbol{\nabla}^2_1}{2} + \sum_{n=1}^{2}\frac{Z_n}{|r_1 - R_n|}\right)\phi_j(\boldsymbol{x}_1) \tag{14}$$

and the Coulomb repulsion terms

$$u_{ijkl} = \iint \mathrm{d}\boldsymbol{r}_1 \mathrm{d}\boldsymbol{r}_2\ \phi^*_i(r_1)\phi_j(r_1)\frac{1}{|r_1 - r_2|}\phi^*_k(r_2)\phi_l(r_2). \tag{15}$$

Z_n are the nuclei charges Z_n $(n = 1, 2)$, and each wavefunction $\phi_i(\boldsymbol{x}_1)$ orbital is a $1s$ orbital centered at the one hydrogen atom. We assume that the system is in its spin singlet state. After reduction [78] a two-qubit
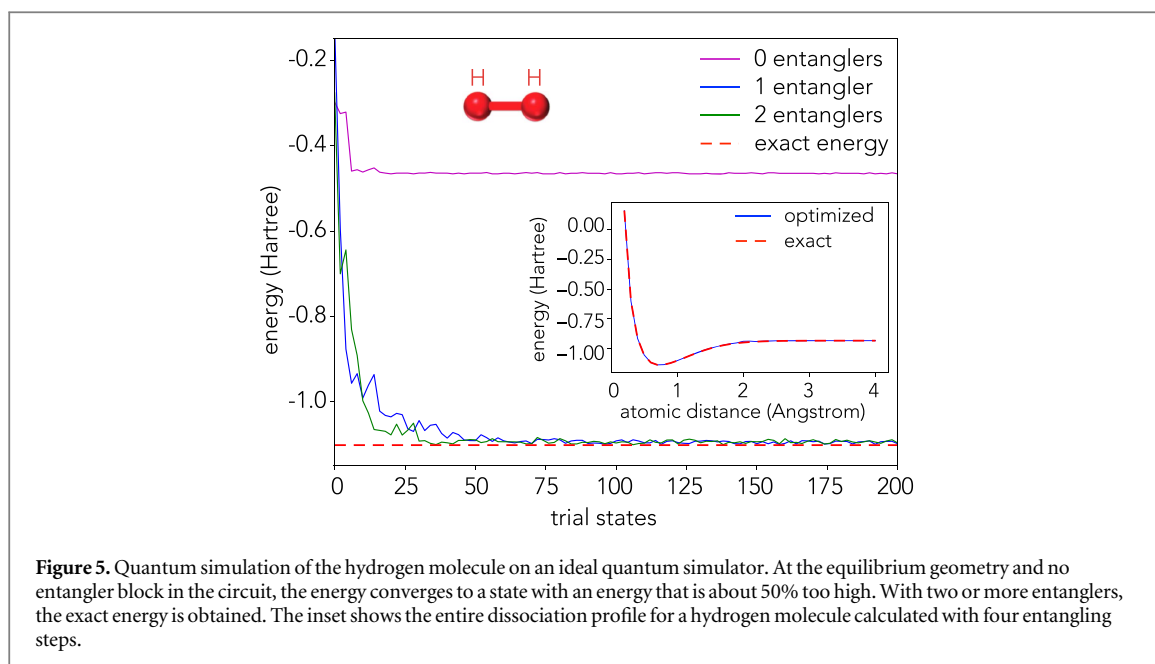
Figure 5. Quantum simulation of the hydrogen molecule on an ideal quantum simulator. At the equilibrium geometry and no entangler block in the circuit, the energy converges to a state with an energy that is about 50% too high. With two or more entanglers, the exact energy is obtained. The inset shows the entire dissociation profile for a hydrogen molecule calculated with four entangling steps.

Hamiltonian is obtained

$$H_{H_2} = f_0 \, \mathbf{1} \otimes \mathbf{1} + f_1 \, \sigma_z \otimes \sigma_z + f_2 \, \sigma_z \otimes \mathbf{1} + f_3 \, \mathbf{1} \otimes \sigma_z + f_4 \, \sigma_x \otimes \sigma_x \tag{16}$$

with $f_0 = -1.0524$, $f_1 = 0.01128$, $f_2 = 0.3979$, $f_3 = 0.3979$, and $f_4 = 0.1809$. These coefficients are calculated at the equilibrium distance of 0.74 Å using equations (14) and (15) and are given in Hartrees.

We evaluate the ground state of the Hamiltonian in (16) on an ideal quantum simulator [61] using a heuristic trial wavefunction approach (section 4.3) with an increasing number of entangling steps (one, two and four). Here, the single qubit rotations of heuristic trial wavefunctions where implemented as $U^i(\theta) = Y(\theta_0^i) Z(\theta_1^i)$ and the entanglement was introduced via control phase gates [85]. Figure 5 shows that a no entangling step is not sufficient to converge towards the correct energy value, whereas one or more entanglers can reproduce the expected results within a few tens of optimization steps in the rotation-angle space $\boldsymbol{\theta}$.

This method can be extended to larger molecules. For lithium hydride (LiH) and beryllium dihydride (BeH$_2$) the second-quantized fermionic Hamiltonian is constructed using a minimal set of atomic orbitals [72] (labeled by the conventional hydrogenic quantum numbers). In BeH$_2$ the basis is composed of the $1s$, $2s$, $2p_x$ orbitals associated to beryllium, and the $1s$ orbital associated to each hydrogen atom. This results in a total of ten spin orbitals. The two innermost $1s$ spin orbitals of beryllium are assumed to be completely filled. The remaining eight spin-orbitals of BeH$_2$ are reduced to six by exploiting spin-parity symmetries [78]. Similarly, the LiH is mapped onto four qubits. It is demonstrated numerically that in the absence of noise, a number of entangling steps $D = 8$ and $D = 28$ are required to achieve chemical accuracy for LiH and BeH$_2$, respectively, for the given experimental connectivity. However, the combined effect of decoherence and finite sampling limits the optimal depth for optimizations on current quantum hardware to between zero and two entanglers, which results in deviations of the simulated bond-dissociation energies from the real values. Decreasing the effective error rates or applying error-mitigation schemes as discussed in section 7 will improve the accuracy of the computations.

5. Classical optimization with qubits

The complex Hamiltonians of quantum chemistry problems give quantum computers an inherent advantage over classical hardware. For classical optimization the advantage is not as obvious because many of the relevant problems can be mapped to a relatively simple Ising-spin Hamiltonian. It is diagonal in the computational basis and can be tackled by a range of classical methods. One of the issues with classical solvers is to avoid solutions in local minima of the cost function. In this context simulated annealing [86] is an approach that makes use of thermal fluctuations to escape such local minima. Quantum annealing [87] additionally exploits quantum tunneling and can potentially reach a ground state faster especially for problems with very corrugated cost functions [6, 7]. The potential for quantum speed-up with this approach is heavily debated in the community; however, because of the tremendous application space even a modest speed-up for a selected number of

problems might have a significant impact. Moreover, understanding the detailed evolution of the optimization process and the potential role of entanglement is critical even for improving algorithms that run on classical hardware. This is why the application of the VQE for solving classical optimization problems on gate-based near-term quantum devices is especially interesting.

To run the VQE we again consider two different ways to create trial wavefunctions. First, the quantum approximate optimization algorithm (QAOA) [41] is discussed, which is a polynomial-time algorithm for finding an approximate solution to a classical optimization problem with a desired accuracy. It is related to the quantum adiabatic algorithm [88], but has shorter circuit-depth requirements. Second, we give a short example how heuristic trial states can be used to solve a *MaxCut* problem on a real quantum device using the VQE.

5.1. QAOA with short depth

Similarly to the approach described in section 4.3 the trial wavefunction in the QAOA is guided towards the solution by repeated unitary evolution according to two Hamiltonians. The first one is the Hamiltonian H_C, which encodes the classical cost function $C(\mathbf{x})$ of a binary constrained optimization problem. The second one is a mixing Hamiltonian H_M, which helps guide the optimization in Hilbert space towards the ground state of H_C. The number of times that both Hamiltonians are applied in the optimization process defines the level D of the circuit and determines the complexity of the algorithm.

Without loss of generality, it is assumed that an optimal solution vertic $\mathbf{x}$ *minimizes* the cost function $C(\mathbf{x})$ which is a polynomial in the binary components $x_i \in \{0, 1\}$ of the variable $\mathbf{x}$. Encoding of the cost function $C(\mathbf{x})$ into a Hamiltonian H_C requires shifting the binary variables $x_i \rightarrow (1 - z_i)/2$ with $z_i \in \{-1, 1\}$ and then substituting $z_i \rightarrow \sigma_i^z$ to obtain an Ising-type Hamiltonian. We chose the same notation as in equation (4) but consider only diagonal terms $\sigma_i^j \in \{\mathbb{1}, \sigma_i^z\}$ which gives

$$H_C = \sum_\alpha h_\alpha P_\alpha = \sum_\alpha h_\alpha \bigotimes_{i_\alpha} \sigma_{i_\alpha}^z. \tag{17}$$

Here the index i_α runs over all $\sigma_{i_\alpha}^z$ in P_α, which constitutes a k-local term (many-body interaction term among $k \leqslant N$ qubits), matching the polynomial terms in the cost function C with corresponding real coefficients h_α. The second Hamiltonian H_M is just a global transverse field, i.e. $H_M = -\sum_i \sigma_i^x$. To find the ground state of the problem Hamiltonian H_C, one proceeds by applying the evolution operator

$$U(\boldsymbol{\beta}, \boldsymbol{\gamma}) = \prod_{l=1}^{D} e^{-\mathrm{i}\beta_l H_M} e^{-\mathrm{i}\gamma_l H_C} \tag{18}$$

to a starting state $|\psi_0\rangle$ that can easily be generated on the quantum computer, e.g. a uniform superposition state. Using the VQE, the parameters of the final state $|\boldsymbol{\beta}, \boldsymbol{\gamma}\rangle = U(\boldsymbol{\beta}, \boldsymbol{\gamma})|\psi_0\rangle$ are then adjusted such as to minimize the expectation value $\langle \boldsymbol{\beta}, \boldsymbol{\gamma}|H_C|\boldsymbol{\beta}, \boldsymbol{\gamma}\rangle$. Measurement of the final state $|\boldsymbol{\beta}, \boldsymbol{\gamma}\rangle$ directly yields the solution of the classical optimization problem with a probability that approaches unity as D increases. However, with increasing D the circuit depth required will reach the decoherence limits of available quantum hardware, and the fidelity of the result will again decrease. Also, the number of classical parameters that need to be optimized for large D will result in a slower convergence. Instead of using the VQE, choosing a fine interpolation $(\beta_l, \gamma_l) = (1 - l/D, l/D)$ with $l = 0,$ $\dots, D$ would be equivalent to first order with a trotterized version of the adiabatic quantum algorithm [1, 88]. By letting the VQE select optimal parameters (γ_l, β_l), a more direct path to the target state becomes possible and the algorithm can reach the ground state with high accuracy even for relatively small values of D. The QAOA has been generalized and successfully applied to MaxCut with analytical and numerical studies [89].

5.2. VQE applied to the MaxCut problem

To give an example of a classical optimization problem, we discuss an instance of the maximum-cut (MaxCut) problem with five qubits. Instead of generating trial states with the QAOA, we again use the hardware-efficient approach explained in section 4.3 to run the algorithm on a real quantum device. The MaxCut problem is an NP-complete binary optimization problem, with applications in clustering, network science, and statistical physics. It aims at grouping the nodes of a graph into two subgroups by cutting across the links between them. The cut is to be made in such a way that the added weights of the links (edges) that were cut are maximized.

The formal definition of this problem is the following: consider an n-node non-directed graph with edge weights $w_{ij} > 0$, $w_{ij} = w_{ji}$, where (i, j) enumerate the nodes linked by the corresponding edge [90]. The profit function to be maximized is therefore the sum of edge weights connecting points in the two different subsets. By assigning a subset label $x_i = 0$ or $x_i = 1$ to each node i, one tries to maximize

$$C(\mathbf{x}) = \sum_{i,j} w_{ij} x_i (1 - x_j). \tag{19}$$
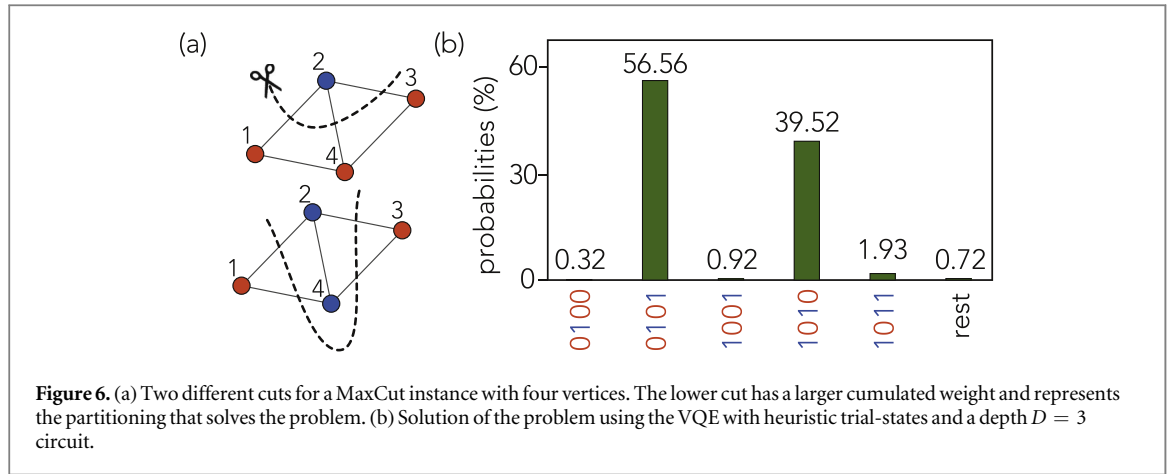
Figure 6. (a) Two different cuts for a MaxCut instance with four vertices. The lower cut has a larger cumulated weight and represents the partitioning that solves the problem. (b) Solution of the problem using the VQE with heuristic trial-states and a depth $D = 3$ circuit.

We can then use the mapping described in section 5.1 to obtain the Ising Hamiltonian

$$H_I = \sum_{i<j} \frac{w_{ij}}{2}(1 - \sigma_i^z)(1 + \sigma_j^z) \tag{20}$$

$$= -\frac{1}{2}\sum_{i<j} w_{ij}\sigma_i^z\sigma_j^z + \text{const.} \tag{21}$$

In other words, the weighted MaxCut problem is equivalent to finding the ground state of the Ising Hamiltonian

$$H_C = \sum_{i<j} w_{ij}\sigma_i^z\sigma_j^z. \tag{22}$$

For exploring the solution space of H_C we use the approach from section 4.3 to define a hardware-efficient heuristic trial wave function

$$|\psi(\theta)\rangle = [U(\boldsymbol{\theta})\,U_{\text{ent}}]^D|\psi_0\rangle, \tag{23}$$

where U_{ent} is a collection of fully entangling gates that are diagonal and the number of entanglers D defines the level of the quantum circuit. The single-qubit rotations are chosen to be $U(\theta) = \prod_{i=1}^{N} Y(\theta_i)$, where N is the number of qubits. For a classical problem this choice allows a search over the space of quantum states with only real coefficients, while still exploiting entanglement to potentially converge faster to the solution. Evaluation of the energy expectation value for a specific trial wavefunction is especially simple in this case as it is sufficient to measure all four qubits and extract the pairwise $\sigma_i^z\sigma_j^z$ correlators. Figure 6(a) shows two different cuts through a problem instance with four nodes (qubits). The lower of the two solves the problem if all non-zero weights in w_{ij} are assumed to be equal. When we implement this on an ideal quantum simulator [61, 85] and use the VQE to optimize the parameters of the trial state in 100 trial steps, we get the state probabilities shown in figure 6(b). For this simple simulation, the solution is found with a probability that is higher than 95%.

6. Classical robust optimizers for measured expectation values

The optimization cycle of the VQE (see section 3) involves evaluation of the cost function on a real quantum device, e.g., a superconducting quantum processor, and adjustment of the variational parameters using classical optimization algorithms (see section 3). In the latter, several important aspects need to be considered for successful application of the VQE.

First, the optimization could get stuck in a local minimum that would correspond to an excited state of the system. Using a suitable optimization routine can prevent finding such false minima. Gradient-descent methods may be combined with simulated annealing steps or strategies that involve starting from multiple initial points. In this context, in [38] a greedy search with multiple starting points is alternated with a Powell search, showing good performances on Hubbard lattices of up to twelve sites.

Second, because of the limited number of samples of the Hamiltonian terms on the quantum computer one only has access to a noisy energy (cost) value. The error in the energy estimation goes as $\mathcal{O}(1/\sqrt{s})$, with s the number of samples taken. Grouping Pauli operators into commuting sets [40, 72] that can be measured with the same state preparation and post-rotations reduces the number of separate measurements and enables more averages and better sampling statistics. Still, the choice of the optimizer must take into account that the cost function is affected by stochastic fluctuations. In fact, while UCC methods and other analytical variational circuits in principle support the use of gradient-based methods that increase the efficiency of the optimization

[91], an imperfect knowledge of the unitary gates implemented in a given quantum device and statistical noise render gradient-based approaches less useful. Derivative-free methods, such as Nelder–Mead and the TOMLAB method, have been tested for optimization of the hydrogen molecule, resulting in a superior performance of the latter method in the presence of stochastic noise [40].

Third, time overheads due to repeated sampling and the number of function evaluations to update the variational parameters will affect the performance of the optimization. In this spirit, the use of a simultaneous perturbation stochastic approximation (SPSA) [92], used in [72] for molecular structure problems, provides both a constant overhead in terms of the number of variational parameters and robustness with respect to stochastic fluctuations. Extensions of the SPSA method that include approximations to the Hessian matrix can be explored to improve the speed of the optimization in the final steps, where estimating second derivatives helps achieve faster convergence [93]. In contrast, additional savings in time overhead in SPSA optimizations that rely on just one evaluation of the cost function per update step [94] could further improve the performance in large-scale quantum problems where sampling is particularly difficult. While simultaneous perturbation methods can be very useful in the optimization of fermionic problems, for classical problems, such as instances of MaxCut, the ease of evaluating the cost function may favor standard gradient-descent or derivative-free routines.

Another critical aspect is the improvement of the classical control hardware for running the VQE on a quantum device: measurement of the cost function with sufficient accuracy requires repeated sampling of the output state and thereby also repeated cycles of qubit initialization, application of the quantum gates and qubit measurement. The speed of the execution of the optimization can be improved on the hardware side by using integrated active reset techniques. In the case of superconducting qubits this is true for both qubits and resonators [95, 96]. Moreover, the costly time overhead in synthesizing and loading control pulses onto the quantum processor for trial-state preparation can be reduced by short-latency field-programmable gate-array-based control and measurement architectures such that time overheads are solely related to the execution of the quantum gates and the readout of the qubits.

7. Prospects of fighting decoherence without full error correction

The hardest challenge for practical near-term quantum devices is their sensitivity to noise. Any computation that has the potential to leverage quantum effects and to provide a quantum speed-up over classical algorithms needs sufficiently coherent qubits. It was realized early on [97] that the coupling to the environment sets both a time and size limit for a quantum computation. Hence, the strength of this coupling determines how large a computation can be performed. This constant limit has to be contrasted to the improvements that are gained from the asymptotic scaling advantages of quantum algorithms. This limitation was, at least in theory, remedied with the advent of quantum error correction [98–100]. However, in spite of rapid experimental progress, the resource requirements for fully fault-tolerant operations with current codes [53] seem prohibitively large [101, 102]. In turn, hopes were raised that non-error-corrected devices will soon become available that reach a regime of reasonably long coherence times and give rise to dynamics too complex to be simulated on a classical computer [43, 103]. In light of these developments, the question arises which computational tasks can be accomplished with quantum devices that have only limited or no error correction. Depending on the form of the actual physical noise, it is expected that the production of entropy in any quantum circuit that is subject to noise will set a limit to this approach [104], and error correction is indispensable for any advanced form of quantum information processing. However, the full computational power of even short-depth circuits is not yet fully understood, and based on complexity-theoretic grounds, it can be argued, that even finite-depth circuits lie beyond the computational power of a classical computer [43, 105].

Recent experiments in which the quantum simulation of small molecules was performed [72] showed that even for very short-depth circuits the effects of decoherence become apparent. For the simulation to be of value, the effect of this error needs to be mitigated, and several proposals have been made to deal with the effects of decoherence in short-depth quantum computation [39, 54, 55, 106].

For a large fraction of applications, the computational task can be abstracted to estimate the expectation value of some observable after the application of a short-depth quantum circuit. This estimation must be accurate enough to achieve a simulation precision that outperforms approximate classical simulation tasks. Techniques to mitigate the error in the estimation of expectation values were introduced in [55]. It is shown that the estimate can be improved in the presence of noise with only a modest time overhead. This approach requires no additional hardware resources such as fresh ancilla or code qubits.

In this scheme, the estimation of an expectation value is improved by an *extrapolation to the limit of zero noise* as originally proposed by Richardson [107]. The method requires no *a priori* knowledge about the noise source, except that the noise is weak and time-independent. To understand this approach it is useful to choose a more physically motivated description of the computation rather than the gate-based quantum circuits. It is more

convenient to consider a time-dependent Hamiltonian dynamics $H(t) = \sum_\alpha J_\alpha(t) P_\alpha$ that implements the circuit, where $J_\alpha(t)$ are coupling coefficients and P_α are N-qubit Pauli operators. In this model the coherent evolution is subject to a noise contribution $\mathcal{L}$ that is effectively constant in time and acts on a time scale much larger than the time-dependent Hamiltonian implementing the quantum circuit. The time evolution up to some time T of the open system with initial state ρ_0 can by described by a Lindblad master equation

$$\frac{\partial}{\partial t}\rho(t) = -\mathrm{i}[H(t), \rho(t)] + \lambda\mathcal{L}(\rho(t)). \tag{24}$$

The expectation value $E(\lambda)$ of some observable A is then obtained from the final state $\rho_\lambda(T)$ and can be written as a power series of the noise rate λ

$$E(\lambda) = E^*(0) + \sum_{i=1}^{n} a_i \lambda^i + O(\lambda^{n+1}), \tag{25}$$

where $E^*(0)$ corresponds to the noise-free expectation value.

Richardson proposed a so-called *deferred approach to the limit* to estimate an expectation value such as $E^*(0)$ with high accuracy [107, 108]. For this purpose, the expectation value $E(\lambda_j)$ is measured for different noise rates $\lambda_j = c_j \lambda$, where c_j is a rescaling factor and λ the actual noise rate in the experiment. The noise-free expectation value can then be estimated by [55]

$$E^*(0) = \sum_{j=0}^{n} \gamma_j E(\lambda_j) + O(\lambda^{n+1}), \tag{26}$$

where $\sum_{j=0}^{n} \gamma_j = 1$ and $\sum_{j=0}^{n} \gamma_j c_j^k = 0$ for $k = 1...n$. In this way the largest terms in the error up to $O(\lambda^n)$ are canceled, thus leading to an estimation of the noise-free expectation value with very high accuracy. In practice however, the noise rate λ is fixed. To still obtain an experimental estimate of the expectation values $E(\lambda_j)$, the following trick can be applied: the quantum circuit $H(t)$ can be run for a time $c_j T$ and with a reduced coupling J_α/c_j. As the noise $\mathcal{L}$ is assumed to be constant in time, it can be shown that the state resulting from a rescaled dynamics is identical to the state obtained from the dynamics with an effectively rescaled noise parameter. Depending on the nature of the noise, relative errors for the noise-free expectation value range from 10^{-6} to 10^{-11} [55].
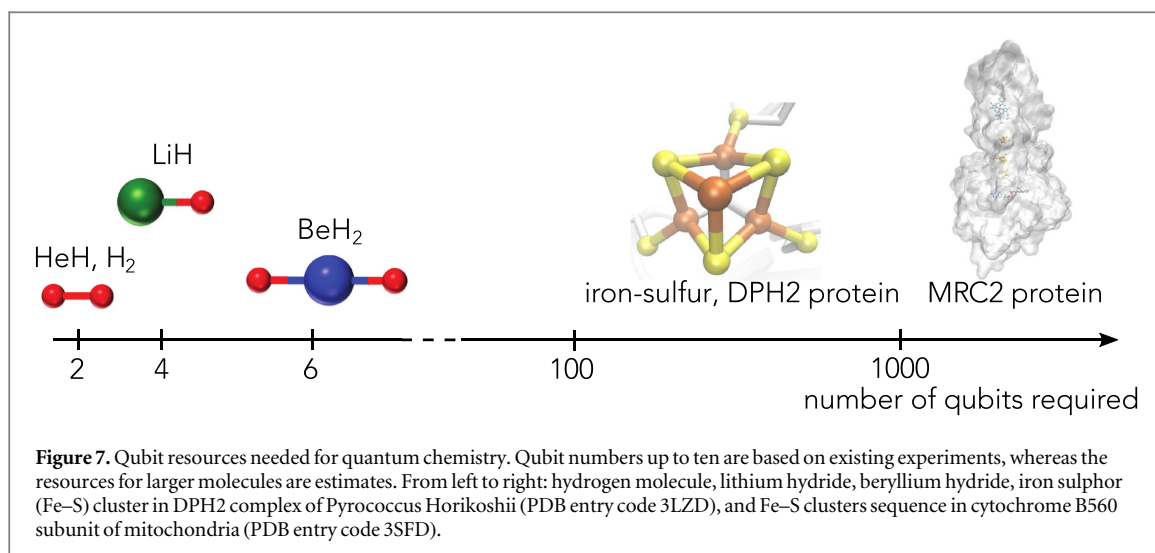
8. Conclusion

Current and near-term quantum processors will most likely be limited to a few hundred, maybe a thousand qubits, and operate without quantum error correction. If the qubits and their control were ideal, the computational power of quantum devices with a couple hundred qubits would already dwarf that of any classical computer and could show *quantum advantage*. However, errors in the quantum operations reduce their computational power.

In this paper it is argued that a proper metric, such as the *quantum volume*, should be used to assess the computing power of a quantum processor and to compare different prototypes on a fair basis. With this metric, it becomes clear that not only the qubit number has to be increased, but also and even more importantly, the effective error rate needs to be significantly reduced before practical applications come within reach. Simple estimates show that to run an algorithm with depth hundred on a hundred-qubit device requires an effective error rate of 0.01%. This number is not completely unrealistic, but shows the necessity to construct algorithms with short depth. Moreover, error-mitigation schemes using no or only a small number of extra ancilla qubits will be important to compensate systematic deviations in the computed result.

Besides enlarging the quantum volume and reducing the effect of errors, it is essential to find suitable methods and algorithms to use quantum effects efficiently. We have discussed that a promising way forward is to consider hybrid quantum–classical architectures in which the quantum processor is used to generate trial quantum states that could not be stored in conventional memory. The VQE method can be used to solve any type of problem that can be cast into a physical Hamiltonian. Constrained binary optimization problems can be described by an Ising-type Hamiltonian, whereas problems from the field of quantum chemistry or material science map into a more general spin Hamiltonian with more than longitudinal interactions among the spins.

For Ising-type Hamiltonian problems, it is not clear how much quantum speed-up can be expected, because many fast classical algorithms have already been developed [41]. In contrast, the Hamiltonian for chemistry and materials-related problems contains so-called non-stoquastic terms, which makes it difficult to solve these problems exactly on a classical computer. It is, therefore, believed that using a quantum processor will lead to exponential speed-up. The current state of the art encompasses proof of concept simulations of small molecules: in the context of superconducting qubits the hydrogen molecule has been simulated with two qubits [33, 34, 72] and larger molecules such as LiH and BeH_2 have been simulated with seven qubits [72]. As the size of the systems

Figure 7. Qubit resources needed for quantum chemistry. Qubit numbers up to ten are based on existing experiments, whereas the resources for larger molecules are estimates. From left to right: hydrogen molecule, lithium hydride, beryllium hydride, iron sulphor (Fe–S) cluster in DPH2 complex of Pyrococcus Horikoshii (PDB entry code 3LZD), and Fe–S clusters sequence in cytochrome B560 subunit of mitochondria (PDB entry code 3SFD).

under study grows in electron number so does the required number of qubits, for example, the simulation of the electronic structure of small organic molecules such as benzene and ethane [13] already requires tens to hundreds of qubits. In the case of strongly correlated electrons, even the simplest systems made of a few atoms, like for instance the chromium dimer [109], quickly become intractable for classical computers when accurate numerical solutions are required. To address strongly correlated problems of practical relevance such as the nitrogen fixation catalytic center in bacteria [27] or the iron–sulphur clusters in the respiratory chain protein complexes [110, 111] (see figure 7) quantum processors with a significantly increased quantum volume are needed. To achieve this, the capabilities of next-generation quantum processors have to improve along several directions:

 (i) Improvement of coherence and qubit control, as well as development of error-mitigation schemes.

 (ii) Hardware-efficient and problem-specific trial state preparation when using VQE.

(iii) Efficient circuit optimization by code optimizers and improved mappings from fermions to qubits.

(iv) Classical parameter optimization methods suited for VQE.

As for (i), current best error rates of $\sim 10^{-4}$ for single and $\sim 10^{-3}$ for two-qubit gate fidelities in the case of superconducting qubit architectures do not provide sufficient accuracy for more complex quantum calculations. The coherence time of qubits has to be improved, e.g., by improving fabrication techniques or chip designs. The control pulses for qubits and their interaction have to be optimized to avoid systematic gate errors. Any remaining errors have to be compensated by error-mitigation strategies.

As for (ii), trial states that require only a variation of a few parameters to prepare the targeted solution state are required. It is an open question how to construct suitable trial states for a general problem set. One may speculate that some combination of heuristic and problem-specific approaches is best suited for the VQE, e.g., hardware-efficient trial wavefunctions which obey certain physical constraints, for example, to conserve the particle number in the quantum chemistry context. Moreover, enlarging the set of available gates, e.g. by exploring coupling primitives that allow different types of interactions between two or more qubits to be realized [112, 113] is considered to create problem-specific trial states and render the VQE efficient.

As for (iii), different fermions-to-qubits maps have been proposed which do not require the creation of entanglement over the entire qubit space. Among the different variants of the Jordan–Wigner and binary-tree methods, one can envisage approaches that perform better in the presence of system-specific noise. Moreover, it may be possible to identify new maps into qubits, which are especially suited for VQEs and that can exploit, for instance, the use of additional ancilla qubits to further reduce the number and the complexity of the gates. Of particular interest is also the possibility to optimize quantum circuits using *post-processing* tools at compilation [27]. The use of high-level languages for the generation and the manipulation of quantum circuits will indeed offer the possibility to rationalize the qubits resources, thus reducing the circuit depth and therefore the time to solution.

As for (iv), specialized classical optimizers that can deal with large stochastic fluctuations resulting from queries to the quantum processor in the VQE are required. The possibility that optimization routines get trapped in false local minima or the effect of high noise render the robustness of optimizers of critical

importance for near-term applications. Even the use of quantum-enhanced optimization schemes may be envisaged.

In conclusion, several promising approaches to make use of near-term devices with hundreds of qubits and limited coherence times have been developed. Overcoming the remaining challenges will allow us to solve tangible problems, most likely in quantum chemistry, material science or classical optimization.

ORCID iDs

Nikolaj Moll https://orcid.org/0000-0001-5645-4667
Gian Salis https://orcid.org/0000-0001-6263-4517

References

[1] Lloyd S 1996 Universal quantum simulators *Science* **273** 1073–8
[2] Abrams D S and Lloyd S 1997 Simulation of many-body Fermi systems on a universal quantum computer *Phys. Rev. Lett.* **79** 2586–9
[3] Abrams D S and Lloyd S 1999 Quantum algorithm providing exponential speed increase for finding eigenvalues and eigenvectors *Phys. Rev. Lett.* **83** 5162–5
[4] Aspuru-Guzik A, Dutoi A D, Love P J and Head-Gordon M 2005 Simulated quantum computation of molecular energies *Science* **309** 1704–7
[5] Harrow A W, Hassidim A and Lloyd S 2009 Quantum algorithm for linear systems of equations *Phys. Rev. Lett.* **103** 150502
[6] Denchev V S, Boixo S, Isakov S V, Ding N, Babbush R, Smelyanskiy V, Martinis J and Neven H 2016 What is the computational value of finite-range tunneling? *Phys. Rev.* X **6** 031015
[7] Albash T and Lidar D A 2017 Evidence for a limited quantum speedup on a quantum annealer arXiv:1705.07452 [quant-ph]
[8] Smolin J A and Smith G 2014 Classical signature of quantum annealing *Frontiers Phys.* **2** 52
[9] Buluta I and Nori F 2009 Quantum simulators *Science* **326** 108–11
[10] Lanyon B P *et al* 2010 Towards quantum chemistry on a quantum computer *Nat. Chem.* **2** 106–11
[11] Brown K L, Munro W J and Kendon V M 2010 Using quantum computers for quantum simulation *Entropy* **12** 2268–307
[12] Temme K, Osborne T J, Vollbrecht K G, Poulin D and Verstraete F 2011 Quantum metropolis sampling *Nature* **471** 87–90
[13] Kassal I, Whitfield J D, Perdomo-Ortiz A, Yung M H and Aspuru-Guzik A 2011 Simulating chemistry using quantum computers *Annu. Rev. Phys. Chem.* **62** 185–207
[14] Whitfield J D, Biamonte J and Aspuru-Guzik A 2011 Simulation of electronic structure Hamiltonians using quantum computers *Mol. Phys.* **109** 735–50
[15] Aspuru-Guzik A and Walther P 2012 Photonic quantum simulators *Nat. Phys.* **8** 285–91
[16] Whitfield J D, Love P J and Aspuru-Guzik A 2012 Computational complexity in electronic structure *Phys. Chem. Chem. Phys.* **15** 397–411
[17] Yung M H and Aspuru-Guzik A 2012 A quantum–quantum metropolis algorithm *Proc. Natl. Acad. Sci.* **109** 754–9
[18] Jones N C, Whitfield J D, McMahon P L, Yung M H, Meter R V, Aspuru-Guzik A and Yamamoto Y 2012 Faster quantum chemistry simulation on fault-tolerant quantum computers *New J. Phys.* **14** 115023
[19] Toloui B and Love P J 2013 Quantum algorithms for quantum chemistry based on the sparsity of the CI-matrix arXiv:1312.2579 [quant-ph]
[20] Wecker D, Bauer B, Clark B K, Hastings M B and Troyer M 2014 Gate-count estimates for performing quantum chemistry on small quantum computers *Phys. Rev.* A **90** 022305
[21] Georgescu I, Ashhab S and Nori F 2014 Quantum simulation *Rev. Mod. Phys.* **86** 153–85
[22] Hastings M B, Wecker D, Bauer B and Troyer M 2015 Improving quantum algorithms for quantum chemistry *Quantum Inf. Comput.* **15** 1–21
[23] Poulin D, Hastings M B, Wecker D, Wiebe N, Doherty A C and Troyer M 2015 The trotter step size required for accurate quantum simulation of quantum chemistry *Quantum Inf. Comput.* **15** 0361
[24] Mueck L 2015 Quantum reform *Nat. Chem.* **7** 361–3
[25] García-Álvarez L, Heras U L, Mezzacapo A, Sanz M, Solano E and Lamata L 2016 Quantum chemistry and charge transport in biomolecules with superconducting circuits *Sci. Rep.* **6** 27836
[26] Whitfield J D 2015 Unified views of quantum simulation algorithms for chemistry arXiv:1502.03771 [physics, physics:quant-ph]
[27] Reiher M, Wiebe N, Svore K M, Wecker D and Troyer M 2017 Elucidating reaction mechanisms on quantum computers *Proc. Natl Acad. Sci.* **114** 7555–60
[28] Babbush R, Berry D W, Kivlichan I D, Wei A Y, Love P J and Aspuru-Guzik A 2016 Exponentially more precise quantum simulation of fermions in second quantization *New J. Phys.* **18** 033032
[29] Whitfield J D, Havlíček V and Troyer M 2016 Local spin operators for fermion simulations *Phys. Rev.* A **94** 030301
[30] Wendin G 2017 Quantum information processing with superconducting circuits: a review *Rep. Prog. Phys.* **80** 106001
[31] Popkin G 2016 Quest for qubits *Science* **354** 1090–3
[32] Shen Y, Zhang X, Zhang S, Zhang J N, Yung M H and Kim K 2017 Quantum implementation of the unitary coupled cluster for simulating molecular electronic structure *Phys. Rev.* A **95** 020501
[33] Colless J I, Ramasesh V V, Dahlen D, Blok M S, McClean J R, Carter J, de Jong W A and Siddiqi I 2018 Robust determination of molecular spectra on a quantum processor *Phys. Rev.* X **8** 011021
[34] O'Malley P *et al* 2016 Scalable quantum simulation of molecular energies *Phys. Rev.* X **6** 031007
[35] Li Z, Yung M H, Chen H, Lu D, Whitfield J D, Peng X, Aspuru-Guzik A and Du J 2011 Solving quantum ground-state problems with nuclear magnetic resonance *Sci. Rep.* **1** 88
[36] Tempel D G and Aspuru-Guzik A 2012 Quantum computing without wavefunctions: time-dependent density functional theory for universal quantum computation *Sci. Rep.* **2** 391
[37] 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

[38] Wecker D, Hastings M B and Troyer M 2015 Progress towards practical quantum variational algorithms *Phys. Rev. A* **92** 042303

[39] Li Y and Benjamin S C 2017 Efficient variational quantum simulator incorporating active error minimization *Phys. Rev. X* **7** 021050

[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] Farhi E, Goldstone J and Gutmann S 2014 A quantum approximate optimization algorithm arXiv:1411.4028 [quant-ph]

[42] Farhi E, Goldstone J and Gutmann S 2014 A quantum approximate optimization algorithm applied to a bounded occurrence constraint problem arXiv:1412.6062 [quant-ph]

[43] Farhi E and Harrow A W 2016 Quantum supremacy through the quantum approximate optimization algorithm arXiv:1602.07674 [quant-ph]

[44] Bishop L S, Bravyi S, Cross A W, Gambetta J M and Smolin J A 2017 Quantum https://dal.objectstorage.open.softlayer.com/v1/AUTH_039c3bf6e6e54d76b8e66152e2f87877/community-documents/quatnum-volumehp08co1vbo0cc8fr.pdf

[45] Verstraete F and Cirac J I 2005 Mapping local Hamiltonians of fermions to local Hamiltonians of spins *J. Stat. Mech.* P09012

[46] Barends R *et al* 2015 Digital quantum simulation of fermionic models with a superconducting circuit *Nat. Commun.* **6** 7654

[47] Reiner J M, Marthaler M, Braumüller J, Weides M and Schön G 2016 Emulating the one-dimensional Fermi–Hubbard model by a double chain of qubits *Phys. Rev. A* **94** 032338

[48] Havlíček V, Troyer M and Whitfield J D 2017 Operator locality in the quantum simulation of fermionic models *Phys. Rev. A* **95** 032332

[49] Heras U L, Mezzacapo A, Lamata L, Filipp S, Wallraff A and Solano E 2014 Digital quantum simulation of spin systems in superconducting circuits *Phys. Rev. Lett.* **112** 200501

[50] You J Q, Shi X F, Hu X and Nori F 2010 Quantum emulation of a spin system with topologically protected ground states using superconducting quantum circuits *Phys. Rev. B* **81** 014505

[51] Biamonte J D 2008 Nonperturbative k-body to two-body commuting conversion Hamiltonians and embedding problem instances into Ising spins *Phys. Rev. A* **77** 052331

[52] Zintchenko I, Hastings M B and Troyer M 2015 From local to global ground states in Ising spin glasses *Phys. Rev. B* **91** 024201

[53] Fowler A G, Mariantoni M, Martinis J M and Cleland A N 2012 Surface codes: towards practical large-scale quantum computation *Phys. Rev. A* **86** 032324

[54] Schwenk I, Reiner J M, Zanker S, Tian L, Leppäkangas J and Marthaler M 2017 Reconstructing the ideal results of a perturbed analog quantum simulator arXiv:1701.02683 [quant-ph]

[55] Temme K, Bravyi S and Gambetta J M 2017 Error mitigation for short depth quantum circuits *Phys. Rev. Lett.* **119** 180509

[56] Córcoles A D, Magesan E, Srinivasan S J, Cross A W, Steffen M, Gambetta J M and Chow J M 2015 Demonstration of a quantum error detection code using a square lattice of four superconducting qubits *Nat. Commun.* **6** 6979

[57] Takita M, Córcoles A D, Magesan E, Abdo B, Brink M, Cross A, Chow J M and Gambetta J M 2016 Demonstration of weight-four parity measurements in the surface code architecture *Phys. Rev. Lett.* **117** 210505

[58] Ristè D, Poletto S, Huang M Z, Bruno A, Vesterinen V, Saira O P and DiCarlo L 2015 Detecting bit-flip errors in a logical qubit using stabilizer measurements *Nat. Commun.* **6** 6983

[59] Barends R *et al* 2014 Superconducting quantum circuits at the surface code threshold for fault tolerance *Nature* **508** 500–3

[60] Ofek N *et al* 2016 Extending the lifetime of a quantum bit with error correction in superconducting circuits *Nature* **536** 441–5

[61] IBM Quantum Experience 2016 http://research.ibm.com/ibm-qx

[62] Debnath S, Linke N M, Figgatt C, Landsman K A, Wright K and Monroe C 2016 Demonstration of a small programmable quantum computer with atomic qubits *Nature* **536** 63–6

[63] Monz T, Nigg D, Martinez E A, Brandl M F, Schindler P, Rines R, Wang S X, Chuang I L and Blatt R 2016 Realization of a scalable shor algorithm *Science* **351** 1068–70

[64] Veldhorst M *et al* 2015 A two-qubit logic gate in silicon *Nature* **526** 410-14

[65] Zajac D M, Sigillito A J, Russ M, Borjans F, Taylor J M, Burkard G and Petta J R 2018 Quantum CNOT gate for spins in silicon *Science* **359** 439-42

[66] Nichol J M, Orona L A, Harvey S P, Fallahi S, Gardner G C, Manfra M J and Yacoby A 2017 High-fidelity entangling gate for double-quantum-dot spin qubits *npj Quantum Inf.* **3** 3

[67] Mourik V, Zuo K, Frolov S M, Plissard S R, Bakkers E P A M and Kouwenhoven L P 2012 Signatures of majorana fermions in hybrid superconductor-semiconductor nanowire devices *Science* **336** 1003–7

[68] Sarma S D, Freedman M and Nayak C 2015 Majorana zero modes and topological quantum computation *npj Quantum Inf.* **1** 15001

[69] O'Brien J L 2007 Optical quantum computing *Science* **318** 1567–70

[70] Barrett S, Hammerer K, Harrison S, Northup T E and Osborne T J 2013 Simulating quantum fields with cavity QED *Phys. Rev. Lett.* **110** 090501

[71] Eichler C, Mlynek J, Butscher J, Kurpiers P, Hammerer K, Osborne T and Wallraff A 2015 Exploring interacting quantum many-body systems by experimentally creating continuous matrix product states in superconducting circuits *Phys. Rev. X* **5** 041044

[72] 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** 242–6

[73] Kohn W and Sham L J 1965 Self-consistent equations including exchange and correlation effects *Phys. Rev.* **140** A1133

[74] Jorgensen P 1981 *Second Quantization-Based Methods in Quantum Chemistry* (New York: Academic)

[75] Babbush R, Wiebe N, McClean J, McClain J, Neven H and Chan G K L 2018 Low depth quantum simulation of electronic structure *Phys. Rev. x* **8** 011044

[76] Bravyi S B and Kitaev A Y 2002 Fermionic quantum computation *Ann. Phys., NY* **298** 210–26

[77] Tranter A, Sofia S, Seeley J, Kaicher M, McClean J, Babbush R, Coveney P V, Mintert F, Wilhelm F and Love P J 2015 The Bravyi-Kitaev transformation: properties and applications *Int. J. Quantum Chem.* **115** 1431–41

[78] Bravyi S, Gambetta J M, Mezzacapo A and Temme K 2017 Tapering off qubits to simulate fermionic Hamiltonians arXiv:1701.08213 [quant-ph]

[79] Jordan P and Wigner E 1928 Über das paulische Äquivalenzverbot *Z. Phys.* **47** 631–51

[80] Moll N, Fuhrer A, Staar P and Tavernelli I 2016 Optimizing qubit resources for quantum chemistry simulations in second quantization on a quantum computer *J. Phys. A: Math. Theor.* **49** 295301

[81] Taube A G and Bartlett R J 2006 New perspectives on unitary coupled-cluster theory *Int. J. Quantum Chem.* **106** 3393–401

[82] Stanton J F and Bartlett R J 1993 *J. Chem. Phys.* **98** 7029–39

[83] Chow J M *et al* 2011 simple all-microwave entangling gate for fixed-frequency superconducting qubits *Phys. Rev. Lett.* **107** 080502

[84] Rigetti C and Devoret M 2010 Fully microwave-tunable universal gates in superconducting qubits with linear couplings and fixed transition frequencies *Phys. Rev.* B **81** 134507

[85] Quantum Information Software Kit 2017 http://qiskit.org

[86] Kirkpatrick S, Gelatt C D and Vecchi M P 1983 Optimization by simulated annealing *Science* **220** 671–80

[87] Kadowaki T and Nishimori H 1998 Quantum annealing in the transverse Ising model *Phys. Rev.* E **58** 5355–63

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

[89] Farhi E, Goldstone J, Gutmann S and Neven H 2017 Quantum algorithms for fixed qubit architectures arXiv:1703.06199 [quant-ph]

[90] Lucas A 2014 Ising formulations of many np problems *Frontiers Phys.* **2** 5

[91] Romero J, Babbush R, McClean J R, Hempel C, Love P and Aspuru-Guzik A 2017 Strategies for quantum computing molecular energies using the unitary coupled cluster ansatz arXiv:1701.02691 [quant-ph]

[92] Spall J C 1992 Multivariate stochastic approximation using a simultaneous perturbation gradient approximation *IEEE Trans. Autom. Control* **37** 332–41

[93] Spall J 2000 Adaptive stochastic approximation by the simultaneous perturbation method *IEEE Trans. Autom. Control* **45** 1839–53

[94] Spall J C 1997 A one-measurement form of simultaneous perturbation stochastic approximation *Automatica* **33** 109–12

[95] McClure D, Paik H, Bishop L, Steffen M, Chow J M and Gambetta J M 2016 Rapid driven reset of a qubit readout resonator *Phys. Rev. Appl.* **5** 011001

[96] Bultink C, Rol M, O'Brien T, Fu X, Dikken B, Dickel C, Vermeulen R, de Sterke J, Bruno A, Schouten R and DiCarlo L 2016 Active resonator reset in the nonlinear dispersive regime of circuit QED *Phys. Rev. Appl.* **6** 034008

[97] Unruh W G 1995 Maintaining coherence in quantum computers *Phys. Rev.* A **51** 992–7

[98] Shor P W 1995 Scheme for reducing decoherence in quantum computer memory *Phys. Rev.* A **52** R2493–6

[99] Steane A M 1996 Error correcting codes in quantum theory *Phys. Rev. Lett.* **77** 793–7

[100] Calderbank A R and Shor P W 1996 Good quantum error-correcting codes exist *Phys. Rev.* A **54** 1098–105

[101] Jones N C, Van Meter R, Fowler A G, McMahon P L, Kim J, Ladd T D and Yamamoto Y 2012 Layered architecture for quantum computing *Phys. Rev.* X **2** 031007

[102] Devitt S J, Stephens A M, Munro W J and Nemoto K 2013 Requirements for fault-tolerant factoring on an atom-optics quantum computer *Nat. Commun.* **4** 2524

[103] Boixo S, Isakov S V, Smelyanskiy V N, Babbush R, Ding N, Jiang Z, Martinis J M and Neven H 2016 Characterizing quantum supremacy in near-term devices arXiv:1608.00263 [quant-ph]

[104] Aharonov D, Ben-Or M, Impagliazzo R and Nisan N 1996 Limitations of noisy reversible computation arXiv:quant-ph/9611028

[105] Terhal B M and DiVincenzo D P 2002 Adaptive quantum computation, constant depth quantum circuits and Arthur-Merlin games *Quantum Inf. Comput.* **4** 134

[106] McClean J R, Kimchi-Schwartz M E, Carter J and de Jong W A 2017 Hybrid quantum–classical hierarchy for mitigation of decoherence and determination of excited states *Phys. Rev.* A **95** 042308

[107] Richardson L F and Gaunt J A 1927 The deferred approach to the limit: I. Single lattice: II. Interpenetrating lattices *Phil. Trans. R. Soc.* A **226** 299–361

[108] Sidi A 2003 *Practical Extrapolation Methods: Theory and Applications* (Cambridge: Cambridge University Press)

[109] Booth G H, Smart S D and Alavi A 2014 Linear-scaling and parallelisable algorithms for stochastic quantum chemistry *Mol. Phys.* **112** 1855–69

[110] Zhang Y *et al* 2010 Diphthamide biosynthesis requires an organic radical generated by an iron–sulphur enzyme *Nature* **465** 891–6

[111] Zhou Q *et al* 2011 Thiabendazole inhibits ubiquinone reduction activity of mitochondrial respiratory complex II via a water molecule mediated binding feature, Thiabendazole inhibits ubiquinone reduction activity of mitochondrial respiratory complex II via a water molecule mediated binding feature *Protein Cell* **2** 531–42

[112] McKay D C, Filipp S, Mezzacapo A, Magesan E, Chow J M and Gambetta J M 2016 Universal gate for fixed-frequency qubits via a tunable bus *Phys. Rev. Appl.* **6** 064007

[113] Roth M, Ganzhorn M, Moll N, Filipp S, Salis G and Schmidt S 2017 Analysis of a parametrically driven exchange-type gate and a two-photon excitation gate between superconducting qubits *Phys. Rev.* A **96** 062323