ARTICLE OPEN

Check for updates

Quantifying the effect of gate errors on variational quantum eigensolvers for quantum chemistry

Kieran Dalton[1,2,3]✉, Christopher K. Long[1,2], Yordan S. Yordanov[1,2], Charles G. Smith[1,2], Crispin H. W. Barnes[2], Normann Mertig[1] and David R. M. Arvidsson-Shukur[1]

Variational quantum eigensolvers (VQEs) are leading candidates to demonstrate near-term quantum advantage. Here, we conduct density-matrix simulations of leading gate-based VQEs for a range of molecules. We numerically quantify their level of tolerable depolarizing gate-errors. We find that: (i) The best-performing VQEs require gate-error probabilities between 10^{-6} and 10^{-4} (10^{-4} and 10^{-2} with error mitigation) to predict, within chemical accuracy, ground-state energies of small molecules with $4-14$ orbitals. (ii) ADAPT-VQEs that construct ansatz circuits iteratively outperform fixed-circuit VQEs. (iii) ADAPT-VQEs perform better with circuits constructed from gate-efficient rather than physically-motivated elements. (iv) The maximally-allowed gate-error probability, p_c, for any VQE to achieve chemical accuracy decreases with the number $N_{||}$ of noisy two-qubit gates as $p_c \propto N_{||}^{-1}$. Additionally, p_c decreases with system size, even with error mitigation, implying that larger molecules require even lower gate-errors. Thus, quantum advantage via gate-based VQEs is unlikely unless gate-error probabilities are decreased by orders of magnitude.

npj Quantum Information (2024)10:18 ; https://doi.org/10.1038/s41534-024-00808-x

INTRODUCTION

Calculating the ground-state energy of a molecular Hamiltonian is an important but hard task in computational chemistry[1]. For strongly correlated systems, exact classical approaches quickly become infeasible as system sizes exceed 100 spin-orbitals. Other, approximate, methods often lack accuracy[1–5]. This makes quantum computers an attractive alternative. A potential route to quantum-chemistry simulations relies on the quantum-phase-estimation algorithm (QPEA)[1,6]. However, the QPEA requires executing millions of gates on error-corrected hardware[7]. Realizing such hardware requires significant resource overheads and gate-error probabilities below a minimal threshold[8]. For example, the surface code[9,10], requires thousands of physical qubits to implement a single logical qubit at a gate-error probability of 10^{-4} [11]. In view of these requirements, the QPEA is not yet feasible.

To reduce the qubit number and gate-error requirements, the variational quantum eigensolver (VQE) was proposed[12]. The VQE is a hybrid quantum-classical algorithm that uses a classical optimizer and a parameterized quantum circuit, the 'ansatz', to estimate ground-state energies. Combined with significant hardware developments[13], VQEs have facilitated successful demonstrations of quantum computational chemistry for small systems[12,14–18]. These demonstrations have been aided by VQE algorithms' abilities to correct for certain errors[15,19,20]. Despite these achievements, there are still significant hurdles to overcome for VQEs to become useful. First, the short, gate-efficient ansätze used in small-scale experimental demonstrations[12,14–18] face optimization difficulties for larger systems. This is related to the emergence of barren plateaus (vanishing gradients), which are more likely when the ansatz is unrelated to the Hamiltonian[21,22]. Current research, for example on growing the ansatz circuit *iteratively* (the ADAPT-VQE algorithm[23]) is aimed at avoiding or mitigating the issue of barren plateaus[21]. Another significant hurdle comes from gate-error rates in hardware. Although current noisy intermediate-scale quantum devices[24–26] have sufficiently many qubits to run VQEs for molecules with more than 100 spin-orbitals[13], their gate-error rates are too high.

At present, efforts to ameliorate the gate-error issue aim to either reduce ansatz circuit depths[23,27–30] or implement elaborate error-mitigation schemes[31–33]. However, VQEs are often benchmarked in the absence of gate errors, with circuit depths and CNOT counts used as proxies of their noise resilience[30]. It has been argued that the maximum viable circuit depth for a VQE ansatz circuit is given by the reciprocal of gate-error probability $1/p$[1]. More rigorously, given a gate-error probability p, the maximum VQE circuit depth which cannot be simulated classically scales as $\mathcal{O}(p^{-1})$[34,35]. A research question, which remains under-explored, is to quantify the gate-error probabilities that VQEs can tolerate. Specifically, considering the analogy of a surface code, which has a well-defined fault-tolerance threshold[11], we aim to find the maximally allowed gate-error probability, below which a certain VQE estimates a certain molecule's energies within chemical accuracy. Quantifying the maximally allowed gate-error probability allows the noise resilience of leading VQEs to be ranked, and provides useful goals for the hardware community.

In this article, we numerically quantify under how high gate-error probabilities VQEs can operate successfully. More specifically, using density-matrix simulations, we simulate the ground-state search of leading, gate-based VQEs for a range of molecules. In the presence of depolarizing noise, we show that: (i) Even the best performing VQEs require gate error probabilities p_c on the order of 10^{-6} to 10^{-4} (without error mitigation) in order to predict molecular ground-state energies within chemical accuracy of 1.6×10^{-3} Hartree. This is significantly below the fault-tolerance threshold of the surface code[11]. For small systems, error mitigation can be employed such that the required p_c values can be improved to 10^{-4} to 10^{-2}. (ii) ADAPT-VQEs tend to tolerate higher gate-error probabilities than VQEs that use fixed ansätze, such as UCCSD and k-UpCCGSD. (iii) ADAPT-VQEs tolerate higher gate-

[1]Hitachi Cambridge Laboratory, J. J. Thomson Ave., Cambridge CB3 0HE, UK. [2]Cavendish Laboratory, Department of Physics, University of Cambridge, Cambridge CB3 0HE, UK. [3]Present address: Department of Physics, ETH Zürich, CH-8093 Zürich, Switzerland. ✉email: kd437@cantab.ac.uk

error probabilities when circuits are synthesized from gate-efficient[27–29,36], rather than physically-motivated[23], elements. We support these claims by estimating, in the presence of depolarizing noise, the scaling relation between the maximally tolerable gate-error probability p_c and the number $N_{||}$ of noisy (two-qubit) gates. Our results indicate that $p_c \propto N_{||}^{-1}$ for any gate-based VQE. (iv) We find that the maximally allowed gate-error probability, p_c, decreases with system size, with and without error mitigation. This shows that larger molecules would likely require even lower gate-errors. We conclude that substantial quantum advantage in VQE-based quantum chemistry is unlikely, unless gate-errors are significantly reduced, or error-corrected hardware is realized, or error-mitigation protocols are improved and made scalable.

RESULTS

ADAPT-VQEs

In this work, we investigate several classes of VQE algorithms. Our study prioritizes VQEs with short ansatz circuits, as these are expected to be more noise resilient[29,30]. Specifically, we consider ADAPT-VQEs, which have comparatively short ansatz circuits[23,29] and the ability to mitigate rough parameter landscapes[21]. We further consider UCCSD[37] and k-UpCCGSD[38] as prototypes of fixed ansatz VQEs – the latter for its comparatively shallow ansatz circuits[30]. Before we outline the results of our noise-resilience investigation, we describe the workings of the (ADAPT-)VQE.

The main idea of VQEs is to use shallow ansatz circuits, defined by a set of parameters $\boldsymbol{\theta}$, to generate entangled trial states $\rho(\boldsymbol{\theta})$. A classical optimizer is then used to vary $\boldsymbol{\theta}$ and minimize the energy-expectation value of H. Provided that the ansatz is sufficiently expressive, the Rayleigh-Ritz variational principle,

$$E(\boldsymbol{\theta}) = \mathrm{Tr}[H\rho(\boldsymbol{\theta})] \geq \mathcal{E}_0, \tag{1}$$

allows $\min_{\boldsymbol{\theta}}(E(\boldsymbol{\theta}))$ to approach the molecular ground-state energy $\mathcal{E}_0$[1]. ADAPT-VQEs use a classical optimizer in two ways[23]: to conduct the Rayleigh-Ritz minimization with respect to a parameterized quantum state; and to iteratively construct the ansatz that generates the parameterized state itself. A quantum computer is used to calculate the energy-expectation value of the parameterized state.

Consider the state generated by the ansatz U_n:

$$\rho_n(\theta_1, \ldots, \theta_n) = U_n(\theta_1, \ldots, \theta_n)\rho_0 U_n^\dagger(\theta_1, \ldots, \theta_n). \tag{2}$$

$\rho_n(\theta_1, \ldots, \theta_n)$ is parameterized by n parameters. In ADAPT-VQE, the classical optimizer and the quantum computer work to find a minimum-energy expectation value:

$$E_n \equiv \min_{\theta_1,\ldots,\theta_n} \mathrm{Tr}[H\rho_n(\theta_1, \ldots, \theta_n)]. \tag{3}$$

An ADAPT-VQE iteratively adds parameterized elements to its ansatz to construct $\rho_n(\theta_1, \ldots, \theta_n)$ such that $E_1 > \ldots > E_n$ and E_n approaches $\mathcal{E}_0$.

The iterative ansatz construction proceeds as follows. First, the ADAPT-VQE algorithm initializes a state ρ_0, usually the Hartree-Fock state[2]. Then, the algorithm generates a sequence of trial states by successively adding elements of the form

$$A_\alpha(\theta) = e^{\theta T_\alpha}, \tag{4}$$

picked from a finite pool $\mathcal{P}$ of operators (see below). Here, T_α, for $\alpha \in [1, \ldots, |\mathcal{P}|]$, are anti-Hermitian operators. Thus, the unitary ansatz grows as

$$U_0 = \mathbb{I}, \tag{5}$$

$$U_n(\theta_1, \ldots, \theta_n) = A_n(\theta_n)U_{n-1}(\theta_1, \ldots, \theta_{n-1}). \tag{6}$$

The ansatz element $A_n(\theta_n) \in \mathcal{P}$ is typically picked to yield the steepest energy gradient. For each value of α, a quantum computer evaluates the energy expectation value after adding $A_\alpha(\theta_n)$ in the nth step:

$$E_{n,\alpha}(\theta_n) \equiv \mathrm{Tr}\left[HA_\alpha(\theta_n)\rho_{n-1}A_\alpha^\dagger(\theta_n)\right]. \tag{7}$$

The ADAPT-VQE algorithm then picks the element $A_n \equiv A_{\alpha=\alpha_n}$ with

$$\begin{aligned} \alpha_n &= \underset{\alpha:A_\alpha\in\mathcal{P}}{\mathrm{argmax}} \left|\frac{\partial E_{n,\alpha}(\theta_n)}{\partial\theta_n}\right|_{\theta_n=0} \\ &= \underset{\alpha:A_\alpha\in\mathcal{P}}{\mathrm{argmax}} \left|\mathrm{Tr}\{[H, T_\alpha]\rho_{n-1}\}\right|. \end{aligned} \tag{8}$$

Alternatively, one may define a sub-pool $\mathcal{S} \subset \mathcal{P}$ of operators with the largest gradients and let the algorithm pick the element with the largest energy difference:

$$\alpha_n = \underset{\alpha:A_\alpha\in\mathcal{S}}{\mathrm{argmin}} \left[\min_{\theta_n} E_{n,\alpha}(\theta_n)\right]. \tag{9}$$

After choosing the nth ansatz element $A_n(\theta_n)$, a classical computer optimizes and updates the parameters $\theta_1, \ldots, \theta_n$ to minimize the energy expectation value E_n in Eq. (3). Provided that $E_n - E_{n-1} > \epsilon$, for some energy precision ϵ, the iterative algorithm continues. When $E_n - E_{n-1} \leq \epsilon$ the algorithm halts at some final length $n = N$, and outputs $E_N \equiv E_n$ as the estimate of $\mathcal{E}_0$.

In this work, we focus on the three main types of ADAPT-VQEs: fermionic-ADAPT-VQE, QEB-ADAPT-VQE and qubit-ADAPT-VQE. (Efficient gate-representations for their relevant ansatz elements can be found in refs. [28,29]). These algorithms differ in their ansatz-element pools $\mathcal{P}$.

First, we consider the fermionic-ADAPT-VQE[23]. As the name suggests, this algorithm uses a pool of operators that closely simulate the physics of fermionic excitations. The pool is formed from

$$T_{ik} = a_i^\dagger a_k - a_k^\dagger a_i, \quad \text{and} \tag{10}$$

$$T_{ijkl} = a_i^\dagger a_j^\dagger a_k a_l - a_k^\dagger a_l^\dagger a_i a_j. \tag{11}$$

Here, $a_i^\dagger$ and a_i are fermionic creation and annihilation operators acting on the ith orbital. Throughout this work, we represent these operators using the Jordan-Wigner transformation[39], where

$$a_i^\dagger \mapsto \frac{1}{2}(X_i - \mathrm{i}Y_i)\prod_{r=0}^{i-1} Z_r, \quad \text{and} \tag{12}$$

$$a_i \mapsto \frac{1}{2}(X_i + \mathrm{i}Y_i)\prod_{r=0}^{i-1} Z_r. \tag{13}$$

X_i, Y_i, Z_i are the Pauli operators acting on the ith qubit. The fermionic-ADAPT-VQE leads to shallower and more gate efficient circuits than UCCSD[23]. Further, choosing fermionic excitations along the gradient of minimum energy produces H-tailored circuits, which can potentially avoid barren plateaus[21].

Second, we consider the QEB-ADAPT-VQE[29,36]. This algorithm uses a pool of operators that nearly (up to a $\pm$ sign) simulate the physics of fermionic excitations. The pool is formed from

$$T_{ik} = Q_i^\dagger Q_k - Q_k^\dagger Q_i, \quad \text{and} \tag{14}$$

$$T_{ijkl} = Q_i^\dagger Q_j^\dagger Q_k Q_l - Q_k^\dagger Q_l^\dagger Q_i Q_j, \tag{15}$$

where

$$Q_i^\dagger := \tfrac{1}{2}(X_i - \mathrm{i}Y_i) \text{ and } Q_i := \tfrac{1}{2}(X_i + \mathrm{i}Y_i). \tag{16}$$

$Q_i^\dagger$ and Q_i are known as qubit creation and annihilation operators, respectively. Due to the CNOT efficiency of its pool, the QEB-ADAPT-VQE can find ground-state and excited-state energies with fewer CNOT gates than the fermionic-ADAPT-VQE[28,29,36].

Finally, we consider the qubit-ADAPT-VQE[27]. This algorithm uses a pool of gate-efficient elements without physical motivation.

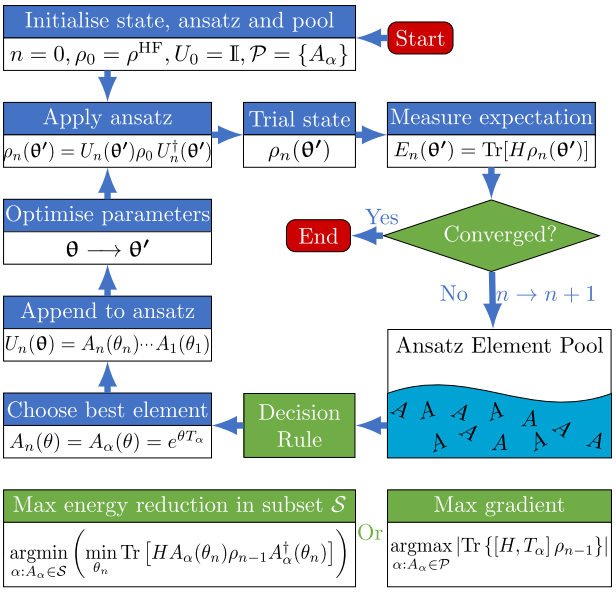

Fig. 1 Flowchart describing the ADAPT-VQE algorithm. At each iteration, an ansatz element is chosen according to one of the two decision rules defined in green below the chart. This element is appended to the ansatz, the parameters are optimized, and the energy expectation value is estimated. The algorithm halts when the change in energy between iterations is below a given threshold.

The pool is formed from segments of Pauli-operator strings:

$$T_{ij} = i\sigma_i\sigma_j, \quad \text{and} \tag{17}$$

$$T_{ijkl} = i\sigma_i\sigma_j\sigma_k\sigma_l, \tag{18}$$

where σ_i denotes Pauli operators X_i, Y_i, Z_i acting on the ith qubit. In previous works, this pool has been found to generate the most shallow and CNOT efficient circuits for ADAPT-VQEs[27]. In our simulations, we use a pool formed from XY-Pauli strings of length two and four with an odd number of Y's. It is possible to use reduced pools[27], but at the expense of reduced circuit efficiency of the final ansatz[29].

Typically, the fermionic-ADAPT-VQE and the qubit-ADAPT-VQE use the gradient-based decision rule expressed in Eq. 8. On the other hand, the original QEB-ADAPT-VQE uses the energy-based decision rule, shown in Eq. 9. These algorithms are summarized in a flow-chart summary in Fig. 1, and in the pseudocode of Supplementary Note 1.

To demonstrate the benefits of iteratively-grown ansätze, we compare them to a typical fixed-ansatz-VQE method: the UCCSD-VQE[17,37]. In Supplementary Note 4, we extend this comparison to the k-UpCCGSD algorithm[38]. Owing to its linear scaling of circuit depth with qubit number, this algorithm was recently put forward as the leading fixed-ansatz VQE[30]. We simulate the workings of the fixed-ansatz methods using the aforementioned fermionic and QEB elements.

Given the breadth of work on VQEs[30], it is not possible to perform an exhaustive analysis of all existing algorithms. Nevertheless, the analytical results in Sec. II E, and the low circuit depths provided by ADAPT-VQE[30], suggest that our results provide a lower bound on the requirements for gate-based VQE algorithms to operate successfully. However, there exist algorithms that differ greatly from typical VQEs, and could deserve future attention. We discuss some of these, and the reasons for our exclusion of them, below. We will not consider iterative qubit coupled cluster (iQCC)[40] and ClusterVQE[41] algorithms. We do not anticipate these algorithms to be feasible options to study strongly correlated systems, whose simulation using quantum algorithms provides the most benefit over classical algorithms. We also omit the DISCO-VQE[42]. Due to its large jumps in Hilbert space during the discrete optimizations of the ansatz, we expect DISCO-VQE to lack tolerance to barren plateaus. These problems may be overcome in future improvements of these VQE algorithms. We leave the design of improved algorithms, and the noise-evaluation of them, to future articles. Finally, we omit the ctrl-VQE algorithm[43]. Although highly interesting, this Hamiltonian algorithm operates with device-tailored pulses, rather than quantum gates, and thus lies outside the scope of this work.

Density-matrix simulations

To investigate the effect of noise on gate-based VQE, we constructed a VQE-tailored density-matrix simulator, expanding the state-vector circuit simulator of ref. [29]. We represent molecular orbitals in the Slater type orbital-3 Gaussians (STO-3G) spin-orbital basis set[44], with the option of frozen orbitals. The openfermion-Psi4 package[45,46] is used to generate the second-quantized Hamiltonian and to perform the Jordan-Wigner transformation[39]. Ansatz parameters are optimized using Nelder-Mead[47] or gradient-descent-based (BFGS)[48] methods in SciPy[49].

We note that, due to the wide array of quantum-computing platforms and their contrasting qubit-control implementations, no noise model can be simultaneously realistic and platform-agnostic. In this work, we model noise by applying single-qubit depolarizing noise to the target qubit i after the application of each two-qubit CNOT gate. Our noise channel can be represented by

$$\mathcal{D}(i,p)[\rho] := (1-p)\rho + \frac{p}{3}\sum_{\sigma_i} \sigma_i \rho \sigma_i, \tag{19}$$

where $p \in [0, 1]$ is the gate-error probability.

In real devices, noise from two-qubit gates completely dominates the noise from single-qubit gates[25,50–52]. Thus, we ignore the latter. Additionally, we exclude state preparation and measurement errors, which are often lower in magnitude than the accumulated two-qubit gate errors[50], and can be mitigated efficiently in experiments[53–58]. (We note that ADAPT-VQE algorithms have high measurement requirements, such that measurement errors may prevent the algorithm from reaching the global minimum energy. This topic requires further investigation). Depolarizing noise is commonly used to represent local and Markovian gate errors when assessing both NISQ[59–61] and quantum-error-correction[11,62–64] algorithms. More realistic models can include thermal-relaxation noise (dephasing and amplitude damping)[65] and device-specific gate errors derived from gate-set-tomography data[66]. When $T_1 \approx T_2$ thermal relaxation noise can be approximated using our depolarizing noise model[62]. This is a reasonable model for superconducting hardware[50]. On the other hand, when $T_2 \ll T_1$ dephasing noise dominates and our depolarizing noise model is less accurate. This is common in trapped-ion devices[51,52] and spin qubits[67] and has recently been investigated in ref. [68]. Moreover, existing VQE algorithms require unrealistically low error rates to give chemically accurate energies. Any attempt to scale down the error rates in realistic noise models to these low levels must be theoretically-justified. This is challenging for a complex, multi-parameter model. Hence, we exclude noise models based on gate-set tomography. Finally, we do not consider coherent errors, since their effect can be suppressed by randomized compiling[69] and dynamical decoupling[70,71]. Randomized compiling[72] can also be used to convert coherent errors to stochastic errors. Additionally, VQE algorithms are somewhat resilient to coherent errors[12,73]. Thus, in this work we focus on incoherent errors. Note that, if VQE algorithms were studied with a coherent noise model, their perceived performance may be greater.

When simulating the smallest molecules (H_2 and H_4) we apply our noise channel [Eq. (19)] after each application of a CNOT gate. This gate-by-gate method is computationally expensive. To facilitate feasible simulations of molecules larger than H_4, we approximate each noisy ansatz element by a corresponding noiseless ansatz

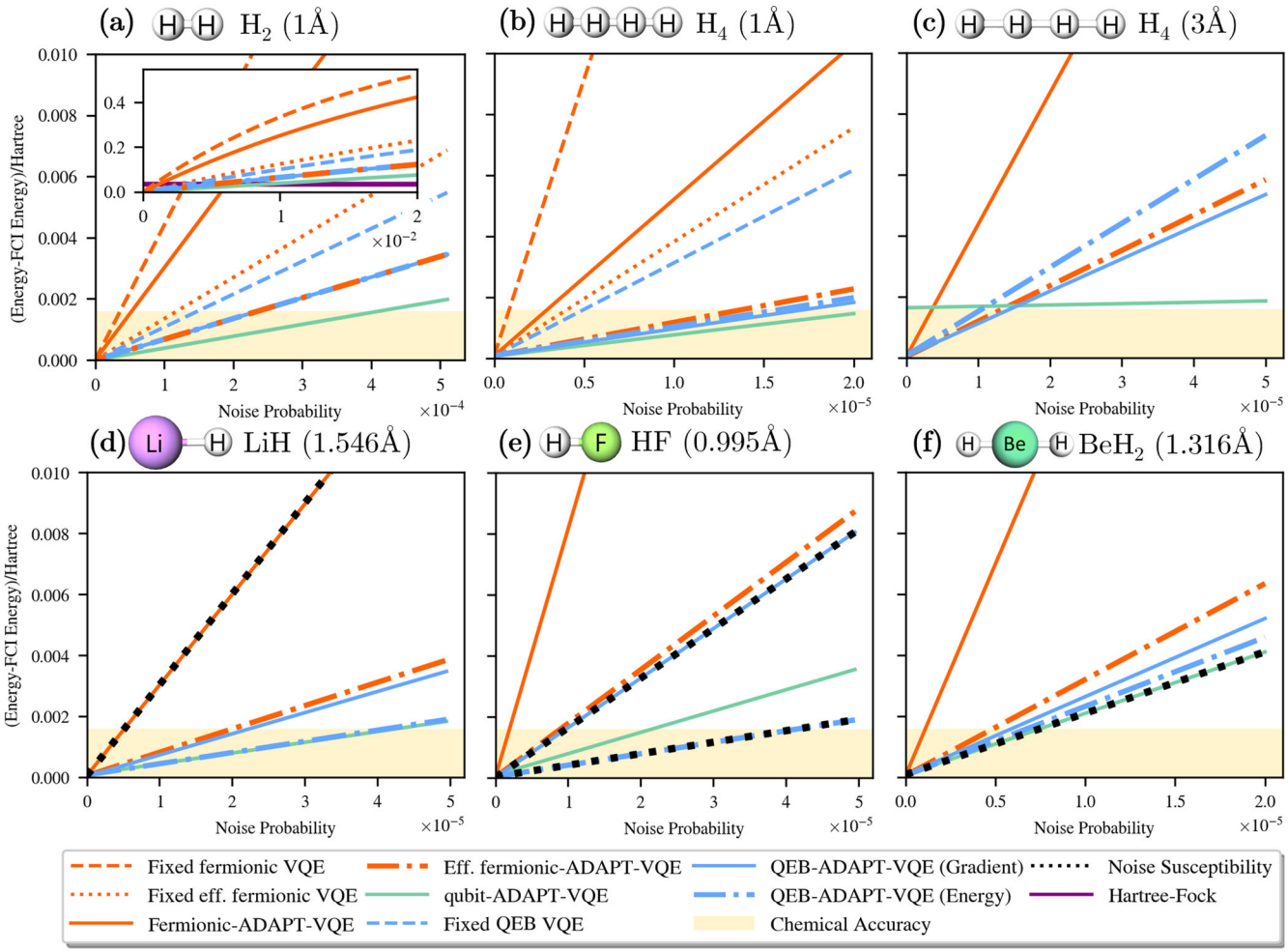

Fig. 2 Energy accuracy as a function of gate-error probability. Plotted for H_2 (**a**), H_4 (at 1 Å (**b**) and 3 Å (**c**) interatomic separation), LiH (**d**), HF (**e**) and BeH_2 (**f**). Ansätze using fermionic, qubit, and Pauli string elements are plotted in red, blue and green, respectively. All curves labeled as "fixed" VQE ansätze use the UCCSD ansatz[17,37]. Energy accuracies lower than chemical accuracy are highlighted by the yellow region. The purple line in the H_2 inset is the energy calculated using the Hartree-Fock[2] state. Extrapolated noise-susceptibility calculations are shown in black for the fermionic-ADAPT-VQE (**d**), the QEB-ADAPT-VQE (**e**) and the qubit-ADAPT-VQE (**f**).

evolution and a noise-inducing evolution. The noise-inducing evolution corresponds to depolarizing noise applied to each qubit in accordance with the number of times that qubit was a CNOT target in the ansatz element. We observe that this lower-bounds the effect of noise. For example, for H_4, applying noise after each CNOT with gate-error probability $p = p'$, gives approximately the same energy accuracy as applying total noise after each element with $p \approx 1.3p'$. Consequently, our simulations of larger molecules should not be compared directly with those for H_2 and H_4. A detailed illustration of our noise approximation is given in Supplementary Note 3.

Energy accuracy is the key metric of VQE performance. It is defined as

$$\Delta E(p,n) := E_n(p) - E_{\mathrm{FCI}}. \tag{20}$$

Here, $E_n(p)$ is the VQE-calculated energy with gate-error probability p in the nth iteration and E_{FCI} is the energy given by the full-configuration-interaction[5] calculation of the true ground-state energy $\mathcal{E}_0$. A key objective of our study is to find the maximally allowed gate-error probability p_c for which $\Delta E(p,n) < 1.6$ milli-Hartree.

Classical optimizers are used to tune $\theta_1, \ldots, \theta_n$. The parameters are optimized until the gradient norm, $|\nabla_{\boldsymbol{\theta}} E| \leq \epsilon_O$, for some precision ϵ_O. In our simulations of H_2 we calculate converged values of $E_n(p)$ using our density-matrix simulator. To keep larger-molecule simulations tractable, we estimate ΔE as follows. We first grow the ansatz circuit $\mathcal{C}_n$ in noiseless, unitary simulations until the nth iteration, for which the energy accuracy $\Delta E(0,n)$ first drops below a cut-off energy precision ϵ_t: $\Delta E(0,n) < \epsilon_t$. Then, we approximate $\Delta E(p,n)$ by simulating the implementation of $\mathcal{C}_n$ with noise on our density-matrix simulator. Thus, $\Delta E(p,n)$ may depend on the iteration n. As demonstrated in Supplementary Note 2, ansatz growth and optimization in the presence of noise have little effect on the noise probability required for chemical accuracy.

Comparison between ADAPT-VQEs with noise

In this section, we benchmark the noise resilience of ADAPT-VQEs using our density-matrix simulator. We study H_2, H_4, LiH, HF and BeH_2. Our simulations were conducted using a parameter optimization cut-off of $|\nabla_{\boldsymbol{\theta}} E| \leq \epsilon_O = 10^{-6}$ Hartree and ansatz growth cut-off of $E_n - E_{n-1} \leq \epsilon = 10^{-12}$ Hartree. In our simulations of the larger molecules, we used an ansatz-truncation cut-off of $\Delta E(0,n) < \epsilon_t = 10^{-4}$ Hartree. Below, we use $\Delta E(p) = \Delta E(p, n_{\mathrm{final}})$ to refer to the energy accuracy at the final ansatz length $n = n_{\mathrm{final}}$. Because of the significant skepticism towards error-mitigation strategies[35,74,75], we omit such strategies from the analyses presented in this section and investigate error mitigation separately in Sec. II F.

The inset of Fig. 2a shows how $\Delta E(p)$ varies with p for H_2. The values of $p \in [0, 0.02]$ include the well-known surface-code fault-

Table 1. Maximum gate-error probabilities p_c [$\times 10^{-5}$] for which chemical accuracy is achieved.

Molecule (Separation)	H_2	H_4 (1 Å)	H_4 (3 Å)	LiH	HF	BeH_2	H_4 (1 Å)	H_4 (3 Å)	LiH	H_2O	NH_2^-
Fermionic-ADAPT-VQE	5.30	0.30	0.36	0.52	0.19	0.11	0.30	0.39	0.96	0.04	0.06
Efficient fermionic-ADAPT-VQE	5.30	1.38	1.34	2.02	0.89	0.49	1.51	1.52	3.47		
QEB-ADAPT-VQE (Energy)	23.59	1.58	1.05	4.14	4.15	0.68	1.72	2.53	7.69	0.35	
QEB-ADAPT-VQE (Gradient)	23.59	1.73	1.45	2.25	0.97	0.59	1.73	1.64	3.92	0.32	
qubit-ADAPT-VQE	41.20	2.19	N/A	4.28	2.15	0.75	2.23	N/A	8.66	0.30	

The first six columns present data from Fig. 2, Sec. II C. Columns seven to nine present data from the top row of Fig. 3, Sec. II D. The data in the final two columns are derived using noise susceptibility, in Sec. II E.

tolerance threshold[8,11] as well as the gate-error probability of currently available quantum hardware[25,50–52]. All tested VQE algorithms require extremely small gate-error probabilities if they are to improve on the Hartree-Fock energy approximation, even for the simple H_2 molecule. The region of chemical accuracy is too small to show in the inset. In real implementations of ADAPT-VQE algorithms, energies exceeding the Hartree-Fock approximation would not be achieved, since, in this case, adding elements to the ansatz does not improve the initial energy accuracy. Here, we add noise to a noiselessly-grown ansatz such that these energies are shown. These observations motivated us to reduce significantly the range of p used in the rest of this study.

The rest of Fig. 2 shows our calculations of $\Delta E(p)$ as a function of p for all considered molecules. The region of chemical accuracy is highlighted by yellow shading. We emphasize six general trends supported by our data. First, the maximally allowed gate-error probabilities p_c for computing ground-state energies within chemical accuracy are extremely small. For all molecules investigated in this study, the value of p_c is on the order of 10^{-6} to 10^{-4} (see Table 1 for details). These values are significantly below the fault-tolerance thresholds of leading error-correction protocols. Second, our simulations of H_2 and H_4 (1 Å) suggest that ADAPT-VQEs outperform fixed ansatz methods. For a given pool of ansatz elements, the corresponding ADAPT-VQE algorithm leads to better energy accuracies than the corresponding fixed ansatz VQE algorithm. Third, the efficient representation of fermionic excitations[28] improves the performance of the fermionic-ADAPT-VQE significantly. This representation reduces CNOT depth, but its scaling of CNOT depth with molecule size is still worse than the scaling of QEB and Pauli string elements. The second and third observations support the claim[29] that the CNOT count is a useful estimator of VQE's noise vulnerability. Fourth, the more gate-efficient (Pauli string and QEB) pools outperform the most physically-motivated (fermionic) pool. The fermionic-ADAPT-VQE is consistently outperformed by either the qubit-ADAPT-VQE or the QEB-ADAPT-VQE. Fifth, sometimes the QEB-ADAPT-VQE outperforms the qubit-ADAPT-VQE and vice versa. For H_2, H_4 (1 Å) and BeH_2, the qubit-ADAPT-VQE outperforms the QEB-ADAPT-VQE. On the other hand, for HF, the QEB-ADAPT-VQE (energy-based decision rule) outperforms the qubit-ADAPT-VQE. For LiH, the QEB-ADAPT-VQE and the qubit-ADAPT-VQE perform similarly. Notably, for H_4 (3 Å), the qubit-ADAPT-VQE fails to add more than two elements to the ansatz. Hence, it never surpasses chemical accuracy. This gives some indication that the qubit-ADAPT-VQE is worse than the QEB-APAPT-VQE at simulating strongly correlated molecules. Sixth, different decision rules for QEB-ADAPT-VQEs yield different performances. For HF, LiH and BeH_2, the energy-reduction decision rule gives a better energy accuracy than the maximum-gradient rule. Conversely, for H_4 (1 Å and 3 Å) the gradient-based decision rule performs better. A study of the optimal decision rules for various molecular-energy landscapes is left for the future.

We close this subsection with a comment on benchmarks of fixed-ansatz VQEs. Both UCCSD and k-UpCCGSD ansätze have been investigated with numerical simulations. However, their energy accuracies significantly worse than those obtained using ADAPT-VQEs. In particular, the energy accuracies for the k-UpCCGSD algorithm do not fit the scale of Fig. 2. Hence, these results are presented separately, in Supplementary Note 4.

Optimal truncation of iteratively-grown ansätze with noise

In the above analyses of ADAPT-VQEs for larger molecules, we truncated noiselessly-grown ansätze in the nth iteration when an energy precision of ϵ_t was reached. Thus, we established a performance hierarchy between different VQEs. However, deeper circuits are generally more vulnerable to noise, and fixing n in noiseless simulations could generate artificially deep circuits. Therefore, to showcase the full benefit of ADAPT-VQEs in the presence of noise, one must vary the truncation length. This would happen automatically for an ansatz grown with noise, as any truncation criterion would be met with shorter ansätze. Here, we simulate this to conduct an alternative, more exact and more computationally expensive, comparison between ADAPT-VQEs. We optimize energy accuracy with respect to ansatz length:

$$\Delta E(p, n_{\text{opt}}) \equiv \min_n \{\Delta E(p, n)\}. \tag{21}$$

Below, we present numerical results for H_4 (1 Å), H_4 (3 Å) and LiH in the left, middle and right columns of Fig. 3, respectively. The color plots show ADAPT-VQE simulations of $\Delta E(p, n)$ as a function of n and p. The top row of Fig. 3 shows the optimized $\Delta E(p, n_{\text{opt}})$ as a function of p. Additional color plots for the VQE methods not shown in Fig. 3 are provided in Supplementary Note 5.

The data in Fig. 3 warrant four comments. First, in the absence of noise ($p = 0$), the energy accuracy decreases monotonically as the ansatz length (n) increases, eventually surpassing chemical accuracy. This is expected from the nature of the ADAPT-VQE methods. The exception is the qubit-ADAPT-VQE simulation of H_4 (3 Å), which simply fails after $n = 2$ and does not achieve chemical accuracy. Second, at small but finite values of p, $\Delta E(p, n)$ initially decreases with larger n. However, after an optimal length $n_{\text{opt}}(p)$, the improvement from appending additional ansatz elements is outweighed by the detrimental increase of noise. At this point, $\Delta E(p, n)$ starts to increase with n. Thus, the ideal ansatz truncation happens at $n = n_{\text{opt}}(p)$. The $n_{\text{opt}}(p)$ values are shown as yellow lines on the color plots of Fig. 3. These curves show $n_{\text{opt}}(p)$ decreasing monotonically as p increases. Third, plots of $\Delta E(p, n_{\text{opt}})$ as a function of p (top Fig. 3), show the same relative performance between the fermionic-ADAPT-VQE, the QEB-ADAPT-VQE and the qubit-ADAPT-VQE as in Fig. 2. Fourth, the values for p_c increase when the ansätze are truncated at an optimal value of n, as compared to the arbitrary truncation used when producing Fig. 2. The values of p_c are presented in Table 1 and Fig. 6. After this optimal-truncation analysis, our overall conclusion remains unchanged: Even for the best-performing ADAPT-VQEs, the maximally allowed gate-error probability is on the order 10^{-6} to 10^{-4} Hartree.

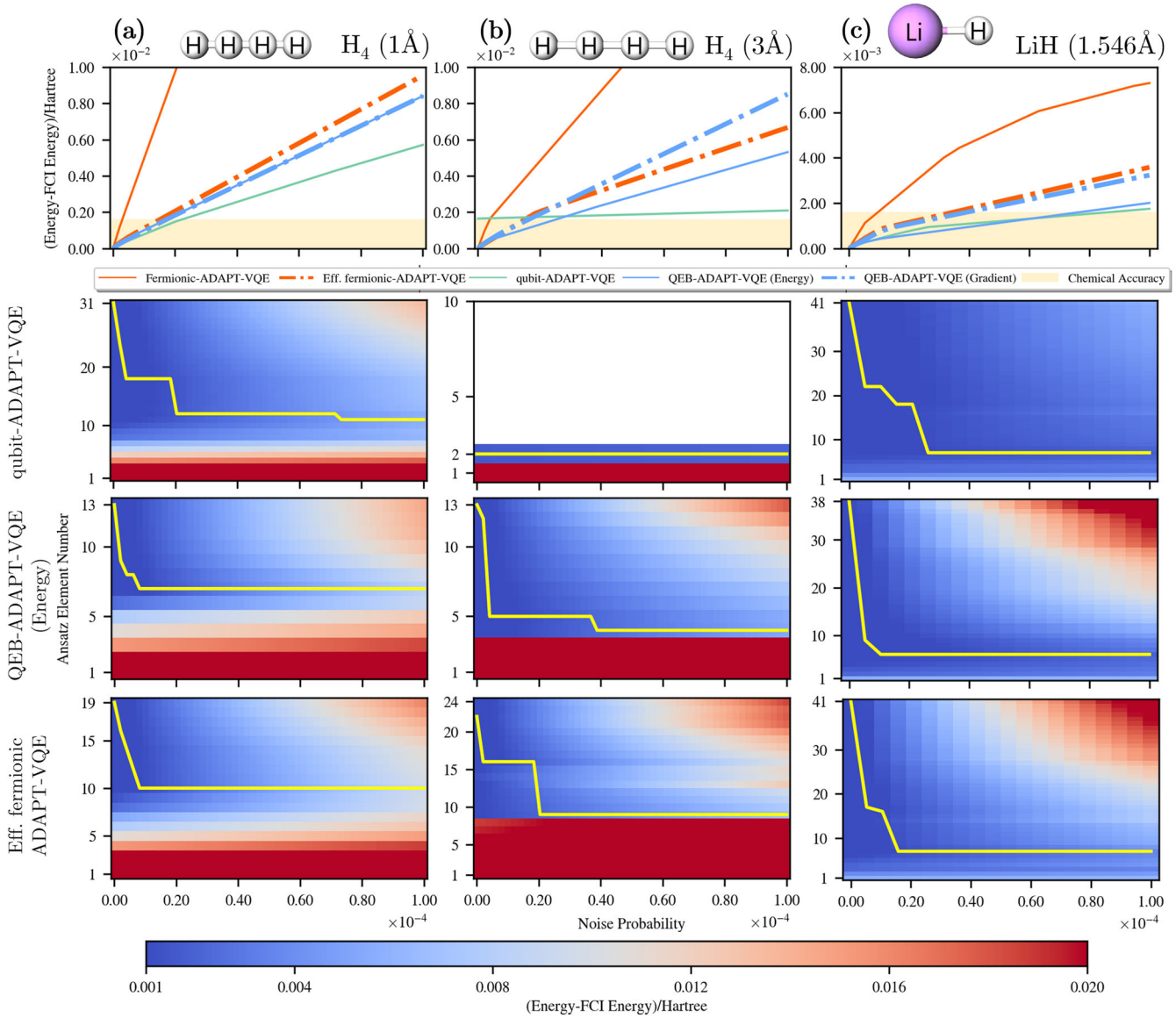

Fig. 3 Color plots representing the energy accuracy (from FCI) at different gate-error probabilities and ansatz lengths. Plotted for three different molecules: H_4 (at 1 Å (**a**) and 3 Å (**b**) interatomic separation) and LiH (**c**). Three iterative-growth methods are included: the fermionic-ADAPT-VQE with efficient elements, the QEB-ADAPT-VQE with the energy decision rule and the qubit-ADAPT-VQE (the plots for the remaining two methods are given in Supplementary Note 2). The yellow lines on the color plot highlight the ansatz lengths that minimize the energy accuracies for each gate-error probability. The top-column figures are extracted from the color plots by plotting the energy accuracy along these curves.

Analytical noise-susceptibility analysis

To analytically support our numerical results, we study the linear response of energy accuracy $\Delta E(p)$ to noisy perturbations of the unitary ansatz circuits. Then, we use our results to show that p_c is roughly inversely proportional to the number N_{II} of noisy (two-qubit) gates.

Noise susceptibility. From Fig. 2 we see that $\Delta E(p) \approx \chi' p$, for some constant χ'. Inspired by this observation, we define the *noise-susceptibility parameter*:

$$\chi = \left.\frac{\partial \Delta E(p)}{\partial p}\right|_{p=0}. \tag{22}$$

Now, we show that $\chi \propto N_{\text{II}}$ (details are given in Supplementary Note 6). If $p = 0$, an ansatz circuit $\mathcal{C}$ can be expressed as a product of R unitary gates: $U = G_R \cdots G_1$. We use $\mathcal{R}_{\text{CX}}$ to denote the set of indices for which G_r is a noisy (CNOT) gate, and we use i_r to denote the qubit which noise acts on. Further, we define a perturbed version of the target unitary U as

$$U_{\text{p}}(\sigma, r, i_r) \quad = G_R \cdots G_{r+1}\sigma_{i_r} G_r \cdots G_1, \tag{23}$$

where the Pauli gate σ acts on qubit i_r after the rth gate. The corresponding energy expectation values are

$$E_U = \text{Tr}\left[HU\rho_0 U^\dagger\right], \tag{24a}$$

$$E_{U_{\text{p}}}(\sigma, r, i_r) = \text{Tr}\left[HU_{\text{p}}(\sigma, r, i_r)\rho_0 U_{\text{p}}^\dagger(\sigma, r, i_r)\right]. \tag{24b}$$

Usually, E_U is close to $\mathcal{E}_0$. Thus, we interpret $E_{U_{\text{p}}}(\sigma, r, i_r)$ as a noise-induced excitation. We call $E_{U_{\text{p}}}(\sigma, r, i_r) - E_U$ the noise-induced fluctuation. The average noise-induced fluctuation of the ansatz is

$$\delta E \equiv \frac{1}{N_{\text{II}}} \sum_{r \in \mathcal{R}_{\text{CX}}} \frac{1}{3} \sum_{\sigma \in \{X,Y,Z\}} \left[E_{U_{\text{p}}}(\sigma, r, i_r) - E_U\right], \tag{25}$$

In Supplementary Note 6, we show that

$$\chi = \delta E \times N_{\text{II}}. \tag{26}$$

Below, we analyze this expression for the noise-susceptibility parameter.

Simplified computations. The energy expectation values underlying χ can be simulated with unitary operations on a state vector. Such simulations are significantly simpler to perform than density-matrix simulations. Thus, we can more easily estimate the energy accuracy for small values of p:

$$\Delta E(p) \approx \chi p. \tag{27}$$

To test our method we compare Eq. (27) (black dotted lines) with some curves in Fig. 2. Eq. (27) estimates the simulated data remarkably well. Next, we use our method to estimate p_c for molecules too large to study with our density-matrix simulator. The estimates of p_c for H_20 and NH_2^- are listed in the final two columns of Table 1. We stress that Eq. (27) is an excellent predictor of $\Delta E(p)$ for the gate-error probabilities $p \in [0, p_c]$ which allow for chemically-accurate simulations.

Scaling. The energy fluctuations are bounded by the spectral range of H: $\delta E \leq \mathcal{E}_{\max} - \mathcal{E}_0$. Thus, Eq. (26) suggests that noise susceptibility grows linearly with N_{II}, as δE is constant. Figure 4 supports this claim. The curves indicate that $\chi \propto N_{\text{II}}$ and $\delta E \approx \mathcal{O}(1)$, for a variety of molecules, ADAPT-VQE algorithms and circuit depths. Combining these observations with Eq. (27), we estimate that

$$p_c \approx \frac{\Delta E_C}{\delta E} \times \frac{1}{N_{\text{II}}} \propto \frac{1}{N_{\text{II}}}. \tag{28}$$

where $\Delta E_C = 1.6 \times 10^{-3}$ Hartree (chemical accuracy). This result is supported by recent results in condensed matter systems[76]. The inverse proportionality between p_c and N_{II} suggests that gate-error probabilities will have to reach extremely small values for useful chemistry calculations with VQE algorithms to be viable.

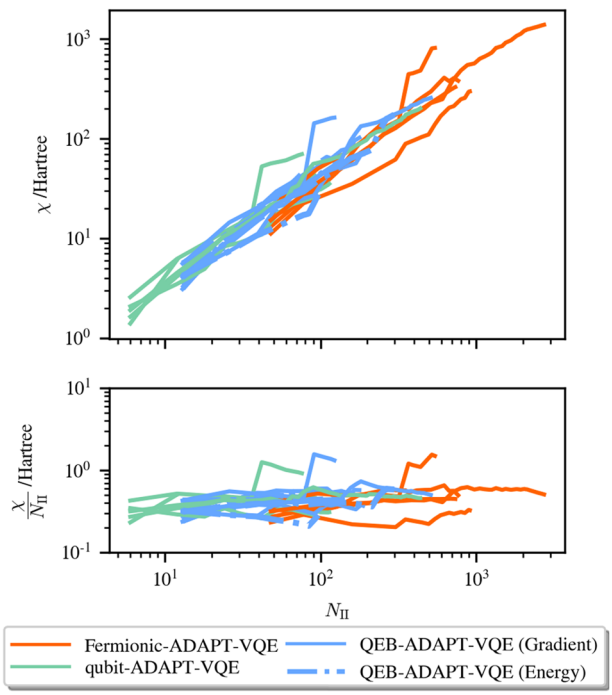

Fig. 4 Numerical noise susceptibility scaling. Noise susceptibility (top) and average energy fluctuation (bottom) as functions of the number N_{II} of CNOT gates for all molecules and algorithms reported in Sec. II C, at all circuit depths.

Alternatively, we require improved VQE algorithms with shallower circuits and fewer noisy (two-qubit) gates.

Quantum error mitigation

In the absence of error-corrected hardware, several strategies to mitigate the effect of noise have been suggested[31–33,77]. Quantum error mitigation is a family of strategies which generally rely on knowledge of a circuit, noise model, or both to generate a set of modified circuits. Sampling from these circuits can generate a better estimate of the noiseless circuit's output[77]. While these strategies have been demonstrated in simple VQE implementations[14,78,79], they suffer, in general, from exponential scaling of sample requirements with qubit number[35,74,75], potentially preventing their viability in useful NISQ VQE implementations. Indeed, leading reviews on quantum computational chemistry[1], state that 'it seems unlikely that error-mitigation methods alone would enable more than a small multiplicative increase in the circuit depth.' This unfavorable scaling has also been observed experimentally, where it has prevented the use of all but the most simple mitigation strategies[26].

The main goal of this work is to assess the required error rates for useful VQE implementations of molecules with more than 100 spin-orbitals. Due to the uncertainty around their scalability, as well as the unclear performance in the presence of time-dependent noise (particularly two-level-system defects[80] which drift in frequency), a study of this type should not include quantum error mitigation in its current form. Despite this, we believe it is relevant to extend our study to ascertain the maximally allowed gate-error probability p_c to calculate molecular energies within chemical accuracy for an error-mitigation protocol with polynomial sampling overhead. To partially address this question, we repeat our numerical simulations using linear zero-noise extrapolation[31,32] with a noise multiplication factor of 3. Despite being biased and heuristic[81], we choose linear zero-noise extrapolation for its modest sampling overhead and numerical stability, which proved useful in recent large-scale demonstrations of error mitigation[26].

The results for H_2, H_4 and LiH are depicted in Fig. 5. Compared to their counterparts in Fig. 3(a), we note the following: (i) The maximally allowed error probability increases by one or two orders of magnitude. This demonstrates the utility of error mitigation to make VQE more viable, especially for smaller molecules. (ii) The resulting energy error displays a roughly parabolic behavior. This is expected from the series expansion of the depolarizing noise and indicates that further improvement (at an increased sampling overhead) may be possible by using higher-order extrapolations. (iii) We note that while all VQE algorithms display an increased noise resilience from error mitigation, their relative p_c-ranking does not change. This suggests that a VQE algorithm with higher noise resilience in the absence of error mitigation would remain more noise resilient when error mitigation is applied. Finally, we put the improved gate-error probabilities p_c into context by plotting them as crosses in Fig. 6. Given the sharp decrease of p_c with the problem size N, in the presence and absence of error mitigation, it is unlikely that error mitigation will improve p_c sufficiently for useful system sizes, $N > 100$. Ultimately, it remains an open question whether the unfavorable scaling of error mitigation prevents its use in realistic quantum-chemistry applications.

DISCUSSION

Any quantum algorithm aimed at near-term NISQ devices must be designed to tolerate some level of noise. In this work, we numerically quantify the maximally allowed depolarizing gate-error probabilities, p_c, required by leading gate-based VQEs to achieve chemically accurate energy estimates. Based on numerical

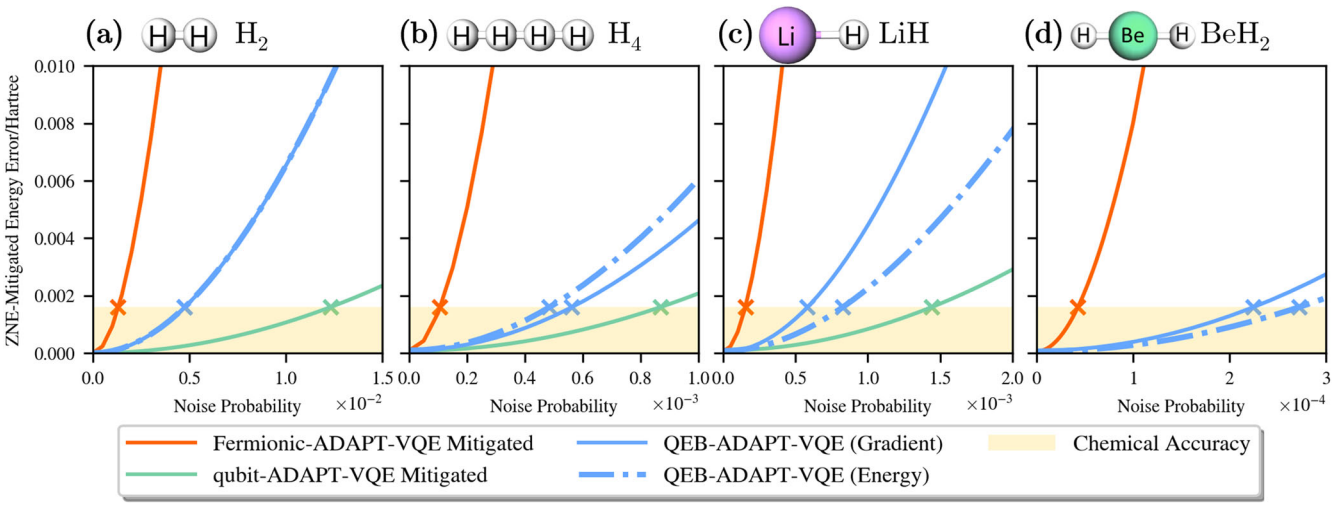

Fig. 5 Plot representing the energy error (from FCI) using different noise probabilities and VQE algorithms. Plotted for H_2 (**a**), H_4 (at 1 Å) (**b**), LiH (**c**), and BeH_2 (**d**), using linear zero-noise extrapolation[31,32]. Energy accuracies lower than chemical accuracy are highlighted by the yellow region. The crosses represent the intercepts of the curves with this region.

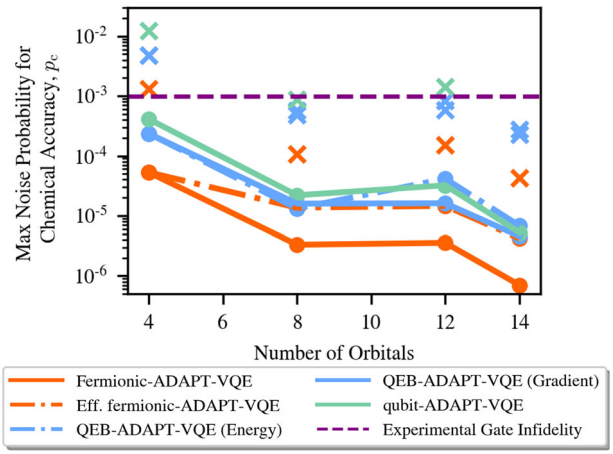

Fig. 6 Plot representing the noise probability required to reach chemical accuracy, p_c, for different ansätze and molecule sizes (number of orbitals). For molecules with the same number of orbitals, the mean probability is taken. The crosses and circles represent the noise probabilities required to reach chemical accuracy with and without error mitigation, respectively. The data without error mitigation is taken from Table 1, and the data with error mitigation is taken from the crosses in Fig. 5. Additionally, a recent state-of-the-art two-qubit gate-error rate with superconducting qubits is shown in purple[93].

simulations, we reach five conclusions. First, even the best-performing VQE algorithms require gate-error probabilities between 10^{-6} and 10^{-4}, for the small molecules we assess. Such errors are at least an order of magnitude below state-of-the-art experiments[25,50] and the surface-code threshold[8,11]. If error mitigation is viable, the p_c values can be improved to 10^{-4} to 10^{-2} with linear zero-noise extrapolation. Second, larger molecules tend to require longer ansatz-circuits and thus, lower gate-error probabilities, see Fig. 6. This is the case both with and without error mitigation. Third, in the presence of noise, ADAPT-VQEs can tolerate approximately an order of magnitude greater gate-errors p_c than equivalent fixed-ansatz VQEs, including those with the shortest ansatz circuits[30,38]. Fourth, the more gate-efficient the ADAPT-VQE ansatz pool, the more noise resilient the algorithm. From a noise-resilience perspective, qubit excitations and Pauli-string excitations outperform fermionic excitations. Fifth, the maximum gate-error probability allowed to reach chemical accuracy is roughly inversely proportional to the number of CNOT gates: $p_c \propto N_{\text{II}}^{-1}$. We now conclude this work with a couple of comments.

In this work, we quantify the maximally allowed gate-error probability of ADAPT-VQEs, UCCSD VQE and the leading fixed-ansatz VQE, k-UpCCGSD[38]. The latter is chosen as due to its favorable circuit depth scaling with molecule size[30]. This ignores plenty of other VQE algorithms which would benefit from similar studies in the future, as discussed in the main text[40–43,82–87].

As opposed to a fault-tolerance threshold in error correction, the maximally allowed gate-error probability p_c crucially depends on the size of the input problem, see Fig. 6. More specifically, p_c tends to shrink as the number of spin orbitals N increases. A key question for future research is to elucidate how fast p_c decreases with N. Our numerical data in Fig. 6 suggests an exponential scaling, both with and without error mitigation. Meanwhile, assuming VQEs achieve molecular ground-state energies with polynomially shallow circuits ($N_{\text{II}} = \text{poly}(N)$), Eq. (28) suggests a polynomial scaling. Having analytical expressions of the decrease of p_c with the number of spin orbitals N, would inform us whether quantum advantage is at all feasible for input problems beyond 100 spin orbitals.

While this study is entirely focused on gate-errors, other sources of noise may also be relevant. These include errors from state preparation and measurement as well as statistical noise due to sampling of expectation values from a limited number of shots. As mentioned when justifying the noise model, errors due to state preparation and measurement tend to be smaller than the accumulated gate errors, and there are widely-implemented methods to compensate for them[53–58]. However, in principle, measurement errors may lead to sub-optimal parameter values or operator choices during ansatz growth of ADAPT-VQE, which may prevent the algorithm from reaching the global minimum energy. A detailed analysis of such effects is left for future work.

While it is possible to sample any expectation value with ϵ accuracy in polynomially few shots[1,30], the scaling prefactor may lead to prohibitively large run-times[30,88–90]. This issue is particularly acute for ADAPT-VQE algorithms, where each growth step requires shots for both parameter optimization and element selection. In this case, the number of necessary gradient measurements for each ansatz growth step is greater than that for VQE parameter optimization, by a factor which scales linearly in the number of qubits[91]. Holistic studies of VQE

run-times[30,88–90] provide predictions which vary greatly depending on the estimation methodology. The estimated run-times are often intractable without significant parallelization. The number of necessary measurements for parameter optimization can potentially be reduced via alternate groupings of Pauli operators[83–85], or tensor contraction of the Hamiltonian (such as by double factorization[86,87,92]). Despite this progress, run-time scaling remains a significant obstacle to overcome before ADAPT-VQEs can perform useful computations on real hardware. A balance must be found between run-time and the acceptable level of statistical noise. This is complicated by the combination of gate errors, measurement errors and statistical errors, which may affect VQEs adversely in a non-trivial way. We leave this as an open problem for the community as, in this work, we focus on the noise resilience of ADAPT-VQEs.

This work numerically investigated the maximally allowed gate-error probability p_c required to achieve chemically accurate predictions as a core metric of VQE performance. Similar to a fault-tolerance threshold in error correction, p_c should provide a transparent metric to compare the noise resilience of VQEs as well as provide useful guidance for the experimental community. Having demonstrated that p_c is between 10^{-4} and 10^{-6} for very small molecules (and worse for larger molecules), we conclude that quantum advantage in VQE-based quantum chemistry requires: (i) Substantially improved error mitigation, (ii) error correction, and/or (iii) significantly improved hardware in which gate errors are reduced by orders of magnitude.

DATA AVAILABILITY

Data generated during the study is available upon request (E-mail: kd437@cantab.ac.uk or ckl45@cam.ac.uk).

CODE AVAILABILITY

The code used for the simulations and analysis carried out for this work is available upon request from the authors (E-mail: kd437@cantab.ac.uk or ckl45@cam.ac.uk).

Received: 14 January 2023; Accepted: 4 January 2024;
Published online: 27 January 2024

REFERENCES

1. McArdle, S., Endo, S., Aspuru-Guzik, A., Benjamin, S. C. & Yuan, X. Quantum computational chemistry. *Rev. Mod. Phys.* **92**, 015003 (2020).
2. Hartree, D. R. & Hartree, W. Self-consistent field, with exchange, for beryllium. *Proc. Math. Phys. Eng. Sci.* **150**, 9–33 (1935).
3. Kohn, W. & Sham, L. J. Self-consistent equations including exchange and correlation effects. *Phys. Rev.* **140**, A1133–A1138 (1965).
4. Hohenberg, P. & Kohn, W. Inhomogeneous electron gas. *Phys. Rev.* **136**, B864–B871 (1964).
5. Rossi, E., Bendazzoli, G. L., Evangelisti, S. & Maynau, D. A full-configuration benchmark for the n2 molecule. *Chem. Phys. Lett.* **310**, 530–536 (1999).
6. Abrams, D. S. & Lloyd, S. Quantum algorithm providing exponential speed increase for finding eigenvalues and eigenvectors. *Phys. Rev. Lett.* **83**, 5162–5165 (1999).
7. Blunt, N. S. et al. Perspective on the current state-of-the-art of quantum computing for drug discovery applications. *J. Chem. Theory Comput.* **18**, 7001–7023 (2022).
8. Campbell, E. T., Terhal, B. M. & Vuillot, C. Roads towards fault-tolerant universal quantum computation. *Nature* **549**, 172–179 (2017).
9. Bravyi, S. B. & Kitaev, A. Y. Quantum codes on a lattice with boundary (1998). Preprint at https://arxiv.org/abs/quant-ph/9811052.
10. Freedman, M. H. & Meyer, D. A. Projective plane and planar quantum codes. *Found. Comput. Math.* **1**, 325–332 (2001).
11. Fowler, A. G., Mariantoni, M., Martinis, J. M. & Cleland, A. N. Surface codes: Towards practical large-scale quantum computation. *Phys. Rev. A* **86**, 032324 (2012).
12. Peruzzo, A. et al. A variational eigenvalue solver on a photonic quantum processor. *Nat. Commun.* **5**, 4213 (2014).
13. Fedorov, A. K., Gisin, N., Beloussov, S. M. & Lvovsky, A. I. Quantum computing at the quantum advantage threshold: a down-to-business review (2022). Preprint at https://arxiv.org/abs/2203.17181.
14. Arute, F. et al. Hartree-fock on a superconducting qubit quantum computer. *Science* **369**, 1084–1089 (2020).
15. O'Malley, P. J. J. et al. Scalable quantum simulation of molecular energies. *Phys. Rev. X* **6**, 031007 (2016).
16. Kandala, A. et al. Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets. *Nature* **549**, 242–246 (2017).
17. Hempel, C. et al. Quantum chemistry calculations on a trapped-ion quantum simulator. *Phys. Rev. X* **8**, 031022 (2018).
18. Xue, X. et al. Quantum logic with spin qubits crossing the surface code threshold. *Nature* **601**, 343–347 (2022).
19. McClean, J. R., Romero, J., Babbush, R. & Aspuru-Guzik, A. The theory of variational hybrid quantum-classical algorithms. *N. J. Phys.* **18**, 023023 (2016).
20. McClean, J. R., Kimchi-Schwartz, M. E., Carter, J. & de Jong, W. A. Hybrid quantum-classical hierarchy for mitigation of decoherence and determination of excited states. *Phys. Rev. A* **95**, 042308 (2017).
21. Grimsley, H. R., Barron, G. S., Barnes, E., Economou, S. E. & Mayhall, N. J. Adaptive, problem-tailored variational quantum eigensolver mitigates rough parameter landscapes and barren plateaus. *npj Quantum Inf.* **9**, 19 (2023).
22. Anschuetz, E. R. & Kiani, B. T. Quantum variational algorithms are swamped with traps. *Nat. Commun.* **13**, 7760 (2022).
23. Grimsley, H. R., Economou, S. E., Barnes, E. & Mayhall, N. J. An adaptive variational algorithm for exact molecular simulations on a quantum computer. *Nat. Commun.* **10**, 3007 (2019).
24. Preskill, J. Quantum Computing in the NISQ era and beyond. *Quantum* **2**, 79 (2018).
25. Ibm quantum systems compute resources. https://quantum-computing.ibm.com/services/resources. Accessed: 2022-09-30.
26. Kim, Y. et al. Evidence for the utility of quantum computing before fault tolerance. *Nature* **618**, 500–505 (2023).
27. Tang, H. L. et al. Qubit-adapt-vqe: An adaptive algorithm for constructing hardware-efficient ansätze on a quantum processor. *PRX Quantum* **2**, 020310 (2021).
28. Yordanov, Y. S., Arvidsson-Shukur, D. R. M. & Barnes, C. H. W. Efficient quantum circuits for quantum computational chemistry. *Phys. Rev. A* **102**, 062612 (2020).
29. Yordanov, Y. S., Armaos, V., Barnes, C. H. W. & Arvidsson-Shukur, D. R. M. Qubit-excitation-based adaptive variational quantum eigensolver. *Commun. Phys.* **4**, 228 (2021).
30. Tilly, J. et al. The variational quantum eigensolver: A review of methods and best practices. *Phys. Rep.* **986**, 1–128 (2022).
31. Temme, K., Bravyi, S. & Gambetta, J. M. Error mitigation for short-depth quantum circuits. *Phys. Rev. Lett.* **119**, 180509 (2017).
32. Li, Y. & Benjamin, S. C. Efficient variational quantum simulator incorporating active error minimization. *Phys. Rev. X* **7**, 021050 (2017).
33. Strikis, A., Qin, D., Chen, Y., Benjamin, S. C. & Li, Y. Learning-based quantum error mitigation. *PRX Quantum* **2**, 040330 (2021).
34. Stilck Franca, D. & García-Patrón, R. Limitations of optimization algorithms on noisy quantum devices. *Nat. Phys.* **17**, 1221–1227 (2021).
35. De Palma, G., Marvian, M., Rouzé, C. & França, D. S. Limitations of variational quantum algorithms: A quantum optimal transport approach. *PRX Quantum* **4**, 010309 (2023).
36. Yordanov, Y. S., Barnes, C. H. W. & Arvidsson-Shukur, D. R. M. Molecular-excited-state calculations with the qubit-excitation-based adaptive variational quantum eigensolver protocol. *Phys. Rev. A* **106**, 032434 (2022).
37. Romero, J. et al. Strategies for quantum computing molecular energies using the unitary coupled cluster ansatz. *Quantum Sci. Technol.* **4**, 014008 (2018).
38. Lee, J., Huggins, W. J., Head-Gordon, M. & Whaley, K. B. Generalized unitary coupled cluster wave functions for quantum computation. *J. Chem. Theory Comput.* **15**, 311–324 (2018).
39. Jordan, P. & Wigner, E. Über das paulische äquivalenzverbot. *Z. Phys.* **47**, 631–651 (1928).
40. Ryabinkin, I. G., Lang, R. A., Genin, S. N. & Izmaylov, A. F. Iterative qubit coupled cluster approach with efficient screening of generators. *J. Chem. Theory Comput.* **16**, 1055–1063 (2020).
41. Zhang, Y. et al. Variational quantum eigensolver with reduced circuit complexity. *npj Quantum Inf.* **8**, 96 (2022).
42. Burton, H. G. A., Marti-Dafcik, D., Tew, D. P. & Wales, D. J. Exact electronic states with shallow quantum circuits from global optimisation. *npj Quantum Inf.* **9**, 75 (2023).
43. Meitei, O. R. et al. Gate-free state preparation for fast variational quantum eigensolver simulations. *npj Quantum Inf.* **7**, 155 (2021).

44. Ditchfield, R., Hehre, W. J. & Pople, J. A. Self-consistent molecular-orbital methods. ix. an extended gaussian-type basis for molecular-orbital studies of organic molecules. *J. Chem. Phys.* **54**, 724–728 (1971).
45. McClean, J. R. et al. OpenFermion: the electronic structure package for quantum computers. *Quantum Sci. Technol.* **5**, 034014 (2020).
46. Turney, J. M. et al. Psi4: an open-source ab initio electronic structure program. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2**, 556–565 (2012).
47. Nelder, J. A. & Mead, R. A simplex method for function minimization. *Comput. J.* **7**, 308–313 (1965).
48. Fletcher, R. *Newton-Like Methods*, chap. 3, 44–79 (John Wiley & Sons, Ltd, 2000).
49. Virtanen, P. et al. Scipy 1.0: fundamental algorithms for scientific computing in python. *Nat. Methods* **17**, 261–272 (2020).
50. Kjaergaard, M. et al. Superconducting qubits: Current state of play. *Annu. Rev. Condens. Matter Phys.* **11**, 369–395 (2020).
51. Wang, Y. et al. High-fidelity two-qubit gates using a microelectromechanical-system-based beam steering system for individual qubit addressing. *Phys. Rev. Lett.* **125**, 150505 (2020).
52. Kang, M. et al. Batch optimization of frequency-modulated pulses for robust two-qubit gates in ion chains. *Phys. Rev. Appl.* **16**, 024039 (2021).
53. Nation, P. D., Kang, H., Sundaresan, N. & Gambetta, J. M. Scalable mitigation of measurement errors on quantum computers. *PRX Quantum* **2**, 040326 (2021).
54. Maciejewski, F. B., Zimborás, Z. & Oszmaniec, M. Mitigation of readout noise in near-term quantum devices by classical post-processing based on detector tomography. *Quantum* **4**, 257 (2020).
55. Bravyi, S., Sheldon, S., Kandala, A., Mckay, D. C. & Gambetta, J. M. Mitigating measurement errors in multiqubit experiments. *Phys. Rev. A* **103**, 042605 (2021).
56. Funcke, L. et al. Measurement error mitigation in quantum computers through classical bit-flip correction. *Phys. Rev. A* **105**, 062404 (2022).
57. McWeeny, R. Some recent advances in density matrix theory. *Rev. Mod. Phys.* **32**, 335–369 (1960).
58. McCaskey, A. J. et al. Quantum chemistry as a benchmark for near-term quantum computers. *npj Quantum Inf.* **5**, 99 (2019).
59. Lee, D. et al. Error-mitigated photonic variational quantum eigensolver using a single-photon ququart. *Optica* **9**, 88–95 (2022).
60. Urbanek, M. et al. Mitigating depolarizing noise on quantum computers with noise-estimation circuits. *Phys. Rev. Lett.* **127**, 270502 (2021).
61. Takagi, R., Endo, S., Minagawa, S. & Gu, M. Fundamental limits of quantum error mitigation. *npj Quantum Inf.* **8**, 114 (2022).
62. Ghosh, J., Fowler, A. G. & Geller, M. R. Surface code with decoherence: An analysis of three superconducting architectures. *Phys. Rev. A* **86**, 062318 (2012).
63. Cross, A. W., Divincenzo, D. P. & Terhal, B. M. A comparative code study for quantum fault tolerance. *Quantum Info Comput.* **9**, 541–572 (2009).
64. Buhrman, H. et al. New limits on fault-tolerant quantum computation. *2006 47th Annual IEEE Symposium on Foundations of Computer Science (FOCS'06)*, 411–419 (Berkeley, CA, USA, 2006).
65. Zeng, J. et al. Simulating noisy variational quantum eigensolver with local noise models. *Quantum Eng.* **3**, e77 (2021).
66. Nielsen, E. et al. Gate Set Tomography. *Quantum* **5**, 557 (2021).
67. Burkard, G., Ladd, T. D., Pan, A., Nichol, J. M. & Petta, J. R. Semiconductor spin qubits. *Rev. Mod. Phys.* **95**, 025003 (2023).
68. Long, C. K., Dalton, K., Barnes, C. H. W., Arvidsson-Shukur, D. R. M. & Mertig, N. Layering and subpool exploration for adaptive variational quantum eigensolvers: Reducing circuit depth, runtime, and susceptibility to noise (2023). Preprint at https://arxiv.org/abs/2308.11708.
69. Hashim, A. et al. Randomized compiling for scalable quantum computing on a noisy superconducting quantum processor. *Phys. Rev. X* **11**, 041039 (2021).
70. Viola, L. & Lloyd, S. Dynamical suppression of decoherence in two-state quantum systems. *Phys. Rev. A* **58**, 2733–2744 (1998).
71. Viola, L., Knill, E. & Lloyd, S. Dynamical decoupling of open quantum systems. *Phys. Rev. Lett.* **82**, 2417–2421 (1999).
72. Wallman, J. J. & Emerson, J. Noise tailoring for scalable quantum computation via randomized compiling. *Phys. Rev. A* **94**, 052325 (2016).
73. Rabinovich, D. et al. On the gate-error robustness of variational quantum algorithms (2023). Preprint at https://arxiv.org/abs/2301.00048.
74. Tsubouchi, K., Sagawa, T. & Yoshioka, N. Universal cost bound of quantum error mitigation based on quantum estimation theory. *Phys. Rev. Lett.* **131**, 210601 (2023).
75. Quek, Y., França, D. S., Khatri, S., Meyer, J. J. & Eisert, J. Exponentially tighter bounds on limitations of quantum error mitigation (2022). Preprint at https://arxiv.org/abs/2210.11505.
76. Kattemölle, J. & van Wezel, J. Variational quantum eigensolver for the heisenberg antiferromagnet on the kagome lattice. *Phys. Rev. B* **106**, 214429 (2022).
77. Cai, Z. et al. Quantum error mitigation. *Rev. Mod. Phys.* **95**, 045005 (2023).
78. Leyton-Ortega, V., Majumder, S. & Pooser, R. C. Quantum error mitigation by hidden inverses protocol in superconducting quantum devices. *Quantum Sci. Technol.* **8**, 014008 (2022).
79. Sagastizabal, R. et al. Experimental error mitigation via symmetry verification in a variational quantum eigensolver. *Phys. Rev. A* **100**, 010302 (2019).
80. Simmonds, R. W. et al. Decoherence in josephson phase qubits from junction resonators. *Phys. Rev. Lett.* **93**, 077003 (2004).
81. Bravyi, S., Dial, O., Gambetta, J. M., Gil, D. & Nazario, Z. The future of quantum computing with superconducting qubits. *J. Appl. Phys.* **132**, 160902 (2022).
82. Feniou, C. et al. Overlap-ADAPT-VQE: practical quantum chemistry on quantum computers via overlap-guided compact Ansätze. *Commun. Phys.* **6**, 192 (2023).
83. Verteletskyi, V., Yen, T.-C. & Izmaylov, A. F. Measurement optimization in the variational quantum eigensolver using a minimum clique cover. *J. Chem. Phys.* **152**, 124114 (2020).
84. Fischer, L. E. et al. Ancilla-free implementation of generalized measurements for qubits embedded in a qudit space. *Phys. Rev. Res.* **4**, 033027 (2022).
85. Miller, D., Fischer, L. E., Sokolov, I. O., Barkoutsos, P. K. & Tavernelli, I. Hardware-tailored diagonalization circuits (2022). Preprint at https://arxiv.org/abs/2203.03646.
86. Oumarou, O., Scheurer, M., Parrish, R. M., Hohenstein, E. G. & Gogolin, C. Accelerating quantum computations of chemistry through regularized compressed double factorization (2023). Preprint at https://arxiv.org/abs/2212.07957.
87. Cohn, J., Motta, M. & Parrish, R. M. Quantum filter diagonalization with compressed double-factorized hamiltonians. *PRX Quantum* **2**, 040352 (2021).
88. Wecker, D., Hastings, M. B. & Troyer, M. Progress towards practical quantum variational algorithms. *Phys. Rev. A* **92**, 042303 (2015).
89. Kühn, M., Zanker, S., Deglmann, P., Marthaler, M. & Weiß, H. Accuracy and resource estimations for quantum chemistry on a near-term quantum computer. *J. Chem. Theory Comput.* **15**, 4764–4780 (2019).
90. Gonthier, J. F. et al. Measurements as a roadblock to near-term practical quantum advantage in chemistry: Resource analysis. *Phys. Rev. Res.* **4**, 033154 (2022).
91. Anastasiou, P. G., Mayhall, N. J., Barnes, E. & Economou, S. E. How to really measure operator gradients in adapt-vqe (2023). Preprint at https://arxiv.org/abs/2306.03227.
92. Hohenstein, E. G. et al. Efficient quantum analytic nuclear gradients with double factorization. *J. Chem. Phys.* **158**, 114119 (2023).
93. Ding, L. et al. High-fidelity, frequency-flexible two-qubit fluxonium gates with a transmon coupler. *Phys. Rev. X* **13**, 031035 (2023).

ACKNOWLEDGEMENTS

We thank Hugo V. Lepage, Flavio Salvati, Frederico Martins, Joseph G. Smith, Wilfred Salmon, and the Hitachi QI team for useful discussions. NM and DRMAS acknowledge useful discussions with Tatsuya Tomaru, Saki Tanaka and Ryo Nagai.

AUTHOR CONTRIBUTIONS

D.R.M.A.-S. and N.M. designed and supervised the project. Y.S.Y., C.K.L. and K.D. wrote the simulation software. K.D. implemented the project and all numerical calculations in close collaboration with C.K.L.. C.K.L. and N.M. developed the noise sensitivity analysis. All authors engaged in technical discussion and the writing of the manuscript.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41534-024-00808-x.

Correspondence and requests for materials should be addressed to Kieran Dalton.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.