





## Mitigating quantum gate errors for variational eigensolvers using hardware-inspired zero-noise extrapolation

Alexey Uvarov <sup>1</sup>, Daniil Rabinovich <sup>1,2</sup>, Olga Lakhmanskaya <sup>3</sup>, Kirill Lakhmanskiy,<sup>3</sup>  
Jacob Biamonte,<sup>4</sup> and Soumik Adhikary <sup>1</sup>

<sup>1</sup>*Skolkovo Institute of Science and Technology, Moscow 121205, Russia*

<sup>2</sup>*Moscow Institute of Physics and Technology, Dolgoprudny 141700, Russia*

<sup>3</sup>*The Russian Quantum Center, Moscow 121205, Russia*

<sup>4</sup>*Yanqi Lake Beijing Institute of Mathematical Sciences and Applications, Yanqi Island, Huairou District, Beijing 101408, China*



(Received 20 July 2023; revised 3 November 2023; accepted 12 June 2024; published 1 July 2024)

Variational quantum algorithms have emerged as a cornerstone of contemporary quantum algorithms research. Practical implementations of these algorithms, despite offering certain levels of robustness against systematic errors, show a decline in performance due to the presence of stochastic errors and limited coherence time. In this work, we develop a recipe for mitigating quantum gate errors using zero-noise extrapolation. We introduce an experimentally amenable method to control error strength in the circuit. We utilize the fact that gate errors in a physical quantum device are distributed inhomogeneously over different qubits and qubit pairs. As a result, one can achieve different circuit error sums based on the manner in which abstract qubits in the circuit are mapped to a physical device. We apply the proposed protocol to variational quantum algorithms and find that the estimated energy is approximately linear with respect to the circuit error sum (CES). Consequently, a linear fit through the energy-CES data, when extrapolated to zero CES, can approximate the energy estimated by a noiseless variational algorithm. We demonstrate this numerically and investigate the applicability range of the technique.

DOI: [10.1103/PhysRevA.110.012404](https://doi.org/10.1103/PhysRevA.110.012404)

### I. INTRODUCTION

Variational algorithms are designed to operate within the practical limitations of near-term quantum computers which are inherently noisy. Such algorithms are known to partially alleviate certain systematic limitations of near-term devices, such as variability in pulse timing and limited coherence times [1–4]. This is accomplished by the use of a short depth parametrized circuit where the parameters are trained using a quantum-to-classical feedback loop. Equipped with these practical advantages, variational algorithms have found several applications, such as quantum approximate optimization (QAOA) [5–17], variational eigensolvers (VQEs) [18–22], and quantum assisted machine learning [23–27].

Variational algorithms are susceptible to stochastic noise [28]. For example, noise can induce barren plateaus in the training landscape, thereby making optimization difficult [29]. Although small amounts of noise can remove local minima in the cost landscape [14], in general the presence of stochastic noise results in the worsening of the final outcome of the algorithm. This necessitates quantum error mitigation [30]. While there are several methods to mitigate errors in a quantum circuit [31–39], for the purpose of this paper we focus on zero-noise extrapolation (ZNE) [37,40,41]. In ZNE, an algorithm is executed at different levels of noise in order to establish a dependence between the output of the algorithm and the noise strength. The dependence is then extrapolated to the zero-noise limit, giving an approximation of the output of the algorithm in noiseless conditions.

To execute ZNE, one must be able to scale the strength of the noise in a controllable manner. In this paper we introduce an experimentally amenable method to scale the circuit noise strength. The rationale behind our approach comes from the realization that qubits are not made equal; two-qubit gates acting on different pairs of qubits can have different fidelities [42,43]. Therefore the level of noise encountered while executing a circuit is determined by how the abstract qubits in the circuit are mapped to their physical counterparts. Thus, by choosing different abstract-to-physical qubit mappings one can control how noise changes in a circuit. Applying this approach to the VQE algorithm, we show that the energy estimated by a noisy VQE is approximately linear with respect to the total circuit error sum. We extrapolate this linear trend to the zero noise limit and show that ZNE recovers the noise-free energy estimation with high accuracy. In fact, we establish that for certain types of variational circuits it is guaranteed that ZNE would recover the exact energy as estimated by noise-free VQE. In addition, we investigate the behavior of ZNE with increasing strength of the noise and observe that the extrapolation error grows quite modestly, following an approximately linear scaling. Finally, we show that the performance of the proposed ZNE protocol is better or comparable to the existing ZNE techniques, while not suffering from their drawbacks.

The paper is structured as follows: In Sec. II, we briefly recall the variational quantum eigensolver algorithm. In Sec. III A, we describe the behavior of the energy estimated by VQE in presence of small errors. In Sec. III B, we propose

our method of scaling noise for ZNE and apply it to VQE. In Sec. IV we present the results of numerical experiments supporting our proposal. Section V concludes the paper.

## II. VARIATIONAL QUANTUM EIGENSOLVER

The variational quantum eigensolver is purpose built to approximate the ground state and the ground-state energy of a given Hamiltonian  $H$  of  $n$  qubits. The algorithm starts with the preparation of a so-called variational state  $|\psi(\boldsymbol{\theta})\rangle = U(\boldsymbol{\theta})|0\rangle^{\otimes n} = [\prod_{j=1}^p U_j(\theta_j)]|0\rangle^{\otimes n}$ , where  $U(\boldsymbol{\theta})$  is a variational *Ansatz* and  $\boldsymbol{\theta} \in [0, 2\pi)^{\times p}$  are tunable parameters. Subsequently, local measurements are performed on the variational state to recover the expectation values of Pauli strings  $P_\alpha$ , which are then classically processed to construct the energy function of the given Hamiltonian  $\langle\psi(\boldsymbol{\theta})|H|\psi(\boldsymbol{\theta})\rangle = \sum_\alpha h_\alpha \langle\psi(\boldsymbol{\theta})|P_\alpha|\psi(\boldsymbol{\theta})\rangle$ . Here  $h_\alpha \in \mathbb{R}$  and  $P_\alpha = \otimes_{j=1}^n \sigma_{\alpha_j}$  with  $\alpha_j \in \{0, 1, 2, 3\}$ . In this step we make use of the fact that any Hamiltonian admits an expansion in the basis of Pauli strings. Finally the energy function is minimized using a classical coprocessor which outputs

$$\boldsymbol{\theta}^* \in \arg \min_{\boldsymbol{\theta}} \langle\psi(\boldsymbol{\theta})|H|\psi(\boldsymbol{\theta})\rangle, \quad (1)$$

$$E^* = \min_{\boldsymbol{\theta}} \langle\psi(\boldsymbol{\theta})|H|\psi(\boldsymbol{\theta})\rangle, \quad (2)$$

$$|\psi(\boldsymbol{\theta}^*)\rangle = U(\boldsymbol{\theta}^*)|0\rangle^{\otimes n}. \quad (3)$$

Here,  $|\psi(\boldsymbol{\theta}^*)\rangle$  is a  $p$ -depth approximate ground state of  $H$ . The proximity of  $|\psi(\boldsymbol{\theta}^*)\rangle$  to the true ground state of  $H$  typically cannot be determined *a priori* based only on the minimization of the energy function. Nevertheless, one can establish bounds on their overlap following the stability lemma [44].

Over time, several improvements in VQE have been reported. For example, there is a number of techniques that do not fix the *Ansatz* circuit  $U(\boldsymbol{\theta})$  in advance, but instead construct it during the optimization [45–49]. Other techniques include efficient estimation of the gradient of the energy function to enable gradient descent [50–52], grouping terms to lower the number of measurements [53], tailoring the *Ansatz* circuit to the restrictions of the problem [54], and many more. An extensive general review of current usage of VQE can be found in Ref. [55], while Refs. [19,56] discuss the variational algorithms specifically in the context of quantum chemistry.

In addition to the theoretical advances, several recent experiments have demonstrated VQE implementation on physical hardware [57–60]. Nevertheless, the results clearly indicate diminishing algorithmic performances due to the presence of hardware noise. Motivated by such observations, error mitigation has been considered as a mean to boost the VQE performance. See Sec. 8 in Ref. [55] for current error mitigation techniques applied to VQE.

## III. HARDWARE-INSPIRED ZERO-NOISE EXTRAPOLATION

Consider a variational circuit comprising of single-qubit gates and two-qubit gates arranged in  $d$  structurally identical layers. We denote a single-qubit gate belonging to the layer  $l$  applied to the qubit  $j$  as  $G_j^l$ , and a two-qubit gate belonging to the layer  $l$  applied to the qubit pair  $(j, k)$  as  $G_{jk}^l$ . We define

$T$  to be the set of qubit pairs to which the two-qubit gates are applied. The set  $T$  is typically determined by the *Ansatz* that is being considered. We use this circuit to variationally minimize a problem Hamiltonian  $H$ .

Here we assume that the single-qubit gates are noiseless. This is in line with many experimental observations that single-qubit gates have very low infidelities which do not influence the performance of our algorithm [61–65]. The two-qubit gates on the other hand are noisy. In those architectures where the single-qubit gates contribute a significant error (see, e.g., Ref. [66]), it can be treated in the same manner as we treat the two-qubit gate errors below.

We consider a simplistic noise model where the application of any two-qubit gate is followed by a transformation:

$$\rho \rightarrow (1 - q)\rho + q\mathcal{E}(\rho), \quad (4)$$

where  $q$  is the gate error rate and  $\mathcal{E}$  is a completely positive trace-preserving (CPTP) map applied to the pair of qubits on which the gate operates.

We consider inhomogeneous errors associated with two-qubit gates. We denote the gate error rate associated with  $G_{jk}^l$  to be  $q_{jk}^l$ . We further assume that the error rate for a gate depends only on the qubit pair it acts on; that is,  $q_{jk}^l = q_{jk} \forall (j, k)$ . For simplicity our analysis neglects the role of state preparation and measurement (SPAM) errors and crosstalk errors. Thus, the set  $\{q_{jk}\}$  and the channel  $\mathcal{E}$  describe the error model completely.

Recall that the action of the map in (4) on a given state is tantamount to applying  $\mathcal{E}$  with probability  $q$ , and operating trivially with probability  $(1 - q)$ . Following this argument one can show that, if  $\rho_0$  is the state that one expects to be prepared by the noiseless circuit, one would instead obtain the state:

$$\rho = \sum_s \left( \prod_{\substack{(j,k) \in T \\ l \in [1,d]}} (1 - q_{jk})^{1-s_{jk}^l} (q_{jk})^{s_{jk}^l} \right) \rho_s. \quad (5)$$

Here  $s$  is a two-dimensional array with elements  $s_{jk}^l \in \{0, 1\}$ , that indexes the two-qubit gates after which  $\mathcal{E}$  is appended: if  $s_{jk}^l = 1$  the channel  $\mathcal{E}$  is appended after the application of  $G_{jk}^l$  while if  $s_{jk}^l = 0$  the state remains intact after the application of  $G_{jk}^l$ . Therefore  $\rho_s$  is the state obtained from  $\rho_0$  by appending the error channels  $\mathcal{E}$  as determined by  $s$ ;  $\rho_0 = \rho_s$  with  $s_{jk}^l = 0 \forall (j, k) \in T, l \in [1, d]$ .

### A. Perturbative analysis

In this section we consider  $q_{jk}$  to be small perturbative terms, such that  $(\max\{q_{jk}\})|T|d \ll 1$ . Here  $|\cdot|$  represents the cardinality of a set. This allows us to discard all terms that are at least quadratic in  $q_{jk}$ . The Taylor expansion of (5) in  $q_{jk}$  yields the linear approximation to  $\rho$ :

$$\rho = \left( 1 - d \sum_{(j,k) \in T} q_{jk} \right) \rho_0 + \sum_{\substack{(j,k) \in T \\ l \in [1,d]}} q_{jk} \rho_{jk}^l + O(q^2), \quad (6)$$

where  $\rho_{jk}^l = \rho_s$  such that  $s$  has only one nonzero entry  $s_{jk}^l = 1$ .

Let us now consider the behavior of the energy  $E = \text{Tr } \rho H$  with respect to the error rates:

$$\begin{aligned} E - E_0 &= -dE_0 \sum_{(j,k) \in T} q_{jk} + \sum_{\substack{(j,k) \in T \\ l \in [1,d]}} q_{jk} E_{jk}^l + O(q^2) \\ &= d(A - E_0) \sum_{(j,k) \in T} q_{jk} + \sum_{\substack{(j,k) \in T \\ l \in [1,d]}} q_{jk} \epsilon_{jk}^l + O(q^2). \end{aligned} \quad (7)$$

Here,  $E_0 = \text{Tr } \rho_0 H$  is the energy of the *Ansatz* state in the absence of noise,  $E_{jk}^l = \text{Tr } \rho_{jk}^l H$  are the energies obtained by applying an error channel  $\mathcal{E}$  to qubits  $(j, k)$  in the  $l$ th layer of the *Ansatz*,  $A = \frac{1}{|T|d} \sum E_{jk}^l$ , and  $\epsilon_{jk}^l = E_{jk}^l - A$ .

Equation (7) has two  $q_{jk}$ -dependent terms. The first term shows a linear dependence of  $(E - E_0)$  with respect to the circuit error sum  $d \sum q_{jk}$ . The second term, on the other hand, quantifies the deviation from the linear trend. In Appendix A we establish bounds on the relative deviation and show that under specific conditions  $(E - E_0)$  behaves linearly with respect to the circuit error sum.

### B. Zero-noise extrapolation with permutation fit

In this section we lay down the recipe for ZNE assisted by inhomogeneous errors. Consider once again the variational circuit which we now want to implement on a physical device. Here one must make a distinction between the abstract qubits (ones in the circuit) and the physical qubits (ones in the device).

We begin by assuming that it is possible to implement two-qubit gates on all pairs of physical qubits and we denote the corresponding error rates as  $\bar{q}_{jk}$ ;  $j, k \in [1, n]$ . Later in the paper we also consider cases of limited connectivity between physical qubits.

To implement an  $n$ -qubit (abstract) circuit on an  $n$ -qubit (physical) device, one would first require to map the abstract qubits to their physical counterparts—what we call abstract-to-physical qubit mapping. Mathematically the mapping can be specified by a permutation  $\pi \in S_n$  such that an abstract qubit indexed  $j$  is mapped to the physical qubit indexed  $\pi(j)$ . Examples of different mappings are demonstrated in Fig. 1. Under a chosen qubit mapping the error rate associated with the gate  $G_{jk}^l$ , operating on the abstract qubit pair  $(j, k)$  and hence the physical qubit pair  $(\pi(j), \pi(k))$ , will be denoted as  $\bar{q}_{\pi(j)\pi(k)} \equiv \bar{q}_{\pi(jk)}$ . For this choice of qubit mapping, the energy of the VQE *Ansatz* state (7) transforms into the following:

$$E - E_0 = d(A - E_0) \sum_{(j,k) \in T} \bar{q}_{\pi(jk)} + \sum_{\substack{(j,k) \in T \\ l \in [1,d]}} \bar{q}_{\pi(jk)} \epsilon_{jk}^l + O(q^2). \quad (8)$$

Evidently, for a generic *Ansatz* circuit the sum  $\sum \bar{q}_{\pi(jk)}$  will depend on the permutation  $\pi$  and hence on the abstract-to-physical qubit mapping. Thus, by taking a number of permutations, we obtain the energies for different circuit error sums which can be approximated by a linear dependence of the form  $E = a \sum \bar{q}_{\pi(jk)} + b$ , where the slope  $a$  approximates the term  $d(A - E_0)$ , and the interception point  $b$  approximates the noise-free energy  $E_0$ . Consequently, by making a linear fit through the noisy

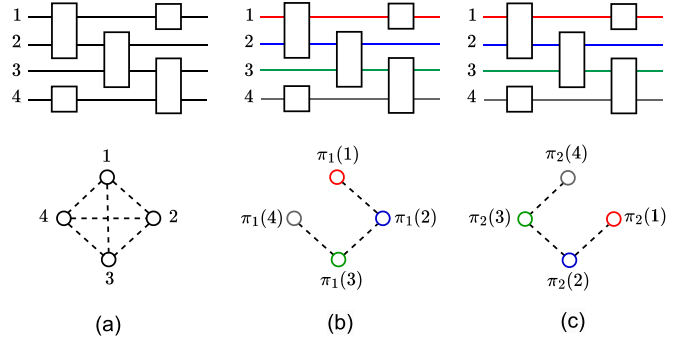


FIG. 1. (a) A four-qubit (abstract) circuit (top) to be executed on a four-qubit (physical) device (bottom) with all-to-all connectivity, depicted with dashed lines. (b), (c) Two different abstract-to-physical qubit mappings  $\pi_1$  and  $\pi_2$  (also indicated by colors). The corresponding qubit connections used for circuit implementation are indicated with dashed lines. Each of these mappings corresponds to a different circuit error sum.

data, one could recover an approximation to the noiseless energy  $E_0$ . In general, one cannot guarantee that the estimate constructed this way is exact. However, in Appendix B we show that this method of ZNE, when performed over all permutations  $\pi \in S_n$ , can recover the exact value of the noise-free energy for *Ansatz* structures corresponding to regular graphs, in the limit of small errors  $\bar{q}_{jk}$ .

## IV. NUMERICAL RESULTS

In this section we demonstrate the performance of ZNE applied to noise-perturbed VQE circuits. To investigate the applicability range of the protocol, we use it for a number of different problem Hamiltonians, noise models and error distributions.

We consider three types of Hamiltonians: The first two are variants of the transverse field Ising model,

$$H = \sum_{j=1}^n J_j Z_j Z_{j+1} + \sum_{j=1}^n h_j X_j, \quad (9)$$

where  $Z_{n+1} \equiv Z_1$ . Two cases are given by (a)  $J_j = h_j = 1$  and (b)  $J_1 = 6, J_j = 1 \forall j \neq 1$ , and  $h_j = 1$ . We refer to these as Ising A and Ising B.

The third type of Hamiltonian is constructed from the electronic structure model of the  $\text{H}_2\text{O}$  molecule. We construct the second-quantized Hamiltonian in the ccpvdz basis using the pycsfdriver module of Qiskit, then choose different sizes of the active space to create problems of different sizes. Finally, we apply the Bravyi-Kitaev transformation and apply a penalty term to ensure the correct particle number:

$$H_p = H + \mu(\hat{n} - n_e)^2. \quad (10)$$

Here  $\hat{n}$  is the particle-number operator  $\sum a_i^\dagger a_i$  transformed with the Bravyi-Kitaev transformation, and  $n_e$  is the desired number of electrons. We choose the weight of the penalty term  $\mu$  according to the “rough” rule from Ref. [67], which means that  $\mu$  is equal to twice the sum of absolute values of the weights of Pauli strings in  $H$ . This choice ensures that the ground state of  $H_p$  has the correct particle number.

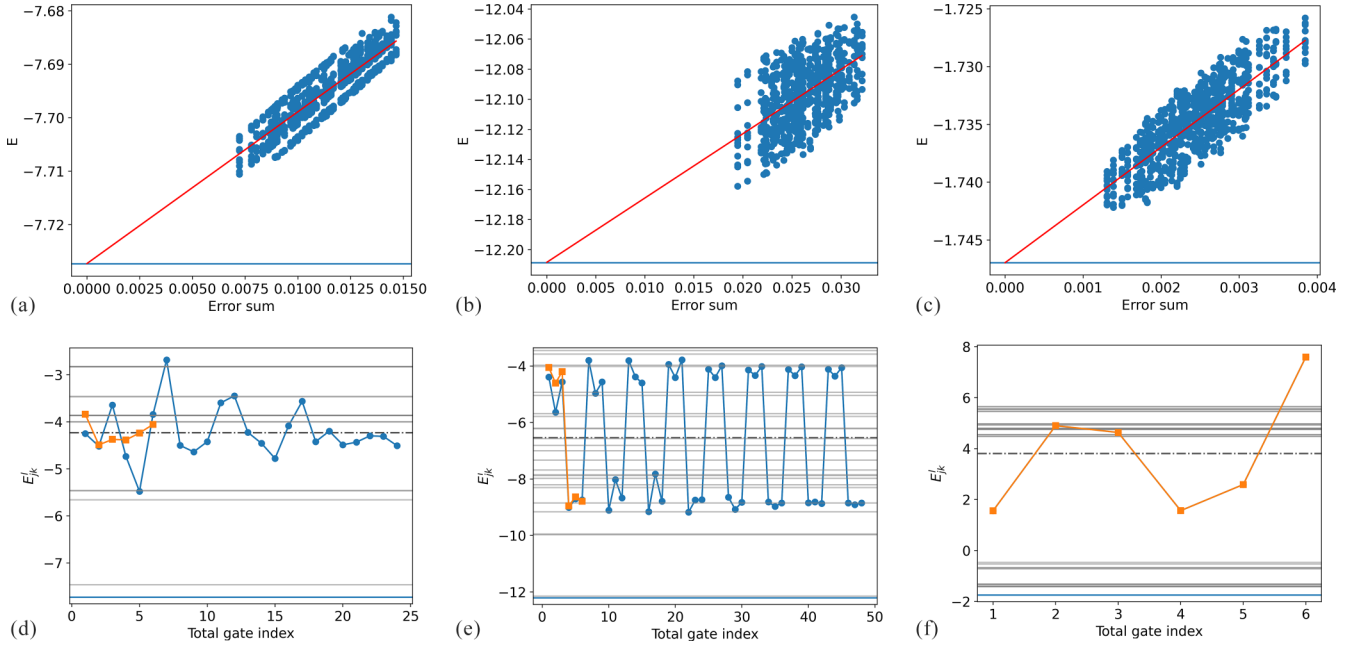


FIG. 2. (a)–(c) Zero-noise extrapolation (ZNE) using the proposed method for  $n = 6$  qubit Ising A, Ising B, and water CAS Hamiltonians, respectively. Blue blue dots on these plots represent energy  $E$  as per (8) for different permutations  $\pi \in S_n$  and the red line is a linear fit taken over energies corresponding to all possible permutations. (d)–(f) The energies  $E_{jk}^l$  [see (7)] for Ising A, Ising B, and water Hamiltonians respectively (shown in circles). The squares show the perturbation energies averaged over the layers of the *Ansatz*  $\frac{1}{d} \sum_l E_{jk}^l$ . The gray lines represent the eigenvalues of the Hamiltonian. The dash-dotted lines depict the average perturbation energy  $A$ . On all the panels blue horizontal lines show the noiseless VQE energy.

We minimize the energy of the Hamiltonians with respect to variational states prepared by a hardware efficient *Ansatz* (HEA) with ring and line topology. For both topologies, in every layer of the *ansatz* we apply single-qubit gates to each qubit ( $R_Y$ , then  $R_X$ ) and then apply two-qubit  $R_{ZZ}$  gates to nearest neighboring qubits. For the ring topology we apply an additional  $R_{ZZ}$  gate to the qubit pair  $(n, 1)$ . In the rest of the paper we mostly focus on the results for the ring topology, as both *Ansätze* gave qualitatively similar results.

The optimization is done in the Qiskit statevector simulator using the limited-memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS-B) algorithm. The starting point for the optimization is chosen by picking every parameter at random from the normal distribution with mean zero and  $\sigma = 10^{-3}$ . To improve convergence, we use a variant of layerwise learning [68,69]: first we start with a one-layered *Ansatz* and optimize it. Then we add another layer, initialized at parameter values close to zero, and optimize the new *Ansatz*, then continue the same procedure with the addition of each subsequent layer. Before every new layer we slightly perturb the solution to avoid the saturation effect [14].

Using an approximately optimal *Ansatz* state, we simulate the ZNE protocol. For that, each two-qubit gate  $g_{jk}$  is appended with a local noise channel with random strength  $\bar{q}_{jk} = \bar{q}_{kj}$ . A number of random qubit permutations is sampled, and for each permutation the CES and the noisy VQE energy are calculated. Finally, these data are used to make a linear fit and obtain the error-mitigated value of the energy.

### A. Zero-noise extrapolation

We begin by considering a case of  $n = 6$  qubits with depolarizing noise of strength  $\{\bar{q}_{jk}\}$  sampled from a uniform distribution on the interval  $[0, 0.001]$ . The depth of the *Ansatz* circuit is chosen such that the noiseless VQE energy differs from the ground energy of the problem by no more than 1% of the spectral gap. In the considered case, the required depths were four, eight, and one layer for the Ising A, Ising B, and water Hamiltonians, respectively.

The results of ZNE for all considered Hamiltonians are demonstrated in Fig. 2. Each point in Figs. 2(a)–2(c) represents energy  $E$  obtained for a specific permutation  $\pi \in S_n$ ; all possible 720 permutations were used for ZNE. A linear fit of energies for the considered perturbations indeed allows one to recover the energy of the noiseless VQE approximation with a high accuracy. In all three cases, the extrapolated energy differs from the true ground energy by at most  $5 \times 10^{-4}$ . Indeed, in Appendix B we show that, in the case of *Ansatz* circuits with a regular graph structure, the ZNE over all permutations approximates noiseless expectation value exactly [up to higher-order terms in (8)].

The energies obtained for the Ising A model and the water Hamiltonian better approximate a linear trend compared with the energies obtained for the Ising B model. To explain this behavior we recall that the deviation from the linear trend is governed by the nonuniformity of the energies  $\epsilon_{jk}^l$ , which is captured by the quantity  $\max_l |\sum_l \epsilon_{jk}^l|$  (see Appendix A). Figures 2(d)–2(f) illustrate that point: for the Ising A model and water Hamiltonians the energies  $E_{jk}^l$  are more concentrated around the mean, and the linear trend with the CES

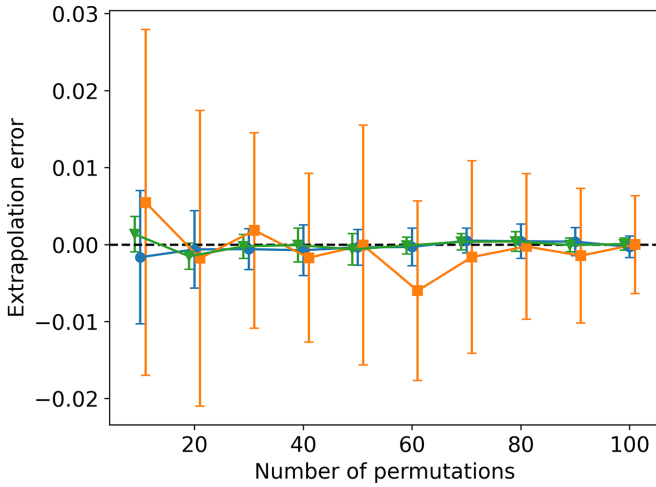


FIG. 3. ZNE error versus the number of permutations used. Blue circles: Ising A, orange squares: Ising B, green triangles: H<sub>2</sub>O. The markers for Ising B and water are offset horizontally for visibility. The error bars indicate standard deviation from considering 20 different noise realizations (sets of  $\{\bar{q}_{jk}\}$  sampled). For every realization a different set of randomly generated permutations was considered.

is more apparent. At the same time, for the Ising B model the distribution of the energies is wider, and consequently, the linear trend is less evident.

Deviation from the linear trend, however, does not necessarily worsen the quality of the approximation, when all permutations are considered. Indeed, notwithstanding any deviation, the ring topology allows us to recover an exact energy approximation (see Appendix B). However, taking all permutations is not feasible except for the smallest problems. In addition, in many quantum computer architectures, implementing an arbitrary permutation of a circuit is impossible without introducing additional swap gates. For example, while ion-based quantum processors have all-to-all qubit connectivity, quantum processors based on superconducting qubits or trapped atoms do not. To address that, we investigated the performance of ZNE when only a limited number of permutations is available. Figure 3 shows the performance of ZNE as a function of the size of the permutation pool. For each of the three Hamiltonians, the uncertainty in the ZNE energy increases for smaller size of the permutation pool. However, the energies are still centered around the noise-free energy, with the standard deviation being much smaller than any energy scale of the problem in each case.

Notwithstanding the success of the protocol with limited number of permutations, we investigate a specific case where available permutations are dictated by connectivity of the hardware. For this we ran an experiment with  $n = 6$  qubits assuming that the connectivity graph is a  $2 \times 3$  square lattice. In this case, we consider an *Ansatz* of the line topology and embedded it in all 16 possible ways into the device. We ran ZNE over 20 random realizations of the noise for the Ising A problem and found that the mean extrapolation error is  $-3.6 \times 10^{-4}$  units with standard deviation  $2.1 \times 10^{-3}$ . This demonstrates that the proposed ZNE protocol can work adequately within the practical limitations of hardware connectivity.

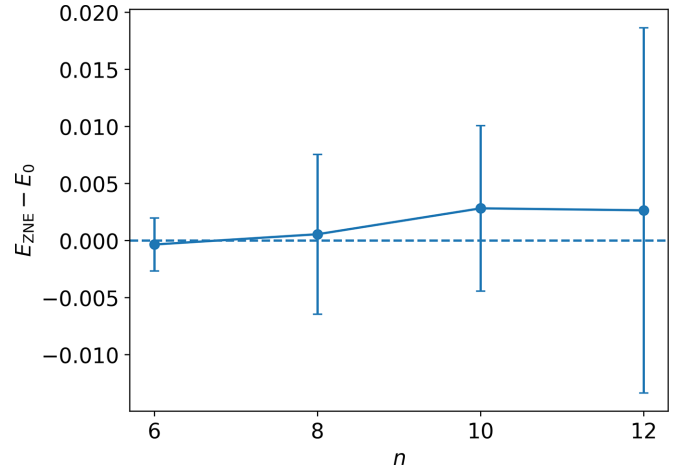


FIG. 4. Scaling of the extrapolation error for depolarizing noise with  $\langle q \rangle = 5 \times 10^{-4}$  with respect to  $n$ . The mean values and deviations are estimated from considering 20 different noise realizations.

### B. Scaling with noise strength and problem size

The proposed method relies on the inhomogeneity of the error rates, essentially taking sums of error rates over particular subsets of qubit pairs. For larger system sizes and circuits, however, the sum of error rates converges towards the mean, leading to the concentration of circuit error sums. This might induce an instability in the ZNE estimated energy. To investigate this, we perform ZNE for system sizes  $n \in \{6, 8, 10, 12\}$  for Ising A Hamiltonian, using 50 random permutations and averaging the results over 20 different noise realizations of average strength  $\langle q \rangle = 5 \times 10^{-4}$  (corresponding to uniform sampling from  $[0, 0.001]$ ). The depth of the noiseless VQE circuit again is taken to approximate the ground energy within 1% of the spectral gap. The resulting distributions of ZNE extrapolation error are shown in Fig. 4. We note that the standard deviation increases gradually with respect to  $n$ . One can offset this increase in the standard deviation by averaging the ZNE energy over at most polynomially (with respect to  $n$ ) many experiments.

In the theoretical analysis, we demanded that  $(\max\{q_{jk}\})|T|d \ll 1$  in order to discard the quadratic terms. However, this requirement can be restrictive for practical purposes. Thus, we investigate the real applicability range of the method. The behavior of the ZNE approximation for the  $n = 12$  qubit Ising A Hamiltonian across a wide range of mean noise strength  $\langle q \rangle$  is shown in Fig. 5. It is seen that the ZNE error is smaller than the spectral gap of the problem up to noise threshold value  $|T|d\langle q \rangle_{\text{thr}} \approx 1$ . Moreover, even for larger values of  $|T|d\langle q \rangle$ , the extrapolation was closer to the true ground-state energy than any of the noisy energy values. In general, the absolute value of the extrapolation error appears to scale polynomially with the amplitude of the noise (Fig. 5, inset).

### C. Zero-noise extrapolation with different noise models

The success of the proposed ZNE protocol may depend on the distribution of error rates, as well as on the considered noise model. To investigate this, we simulate ZNE for (i)

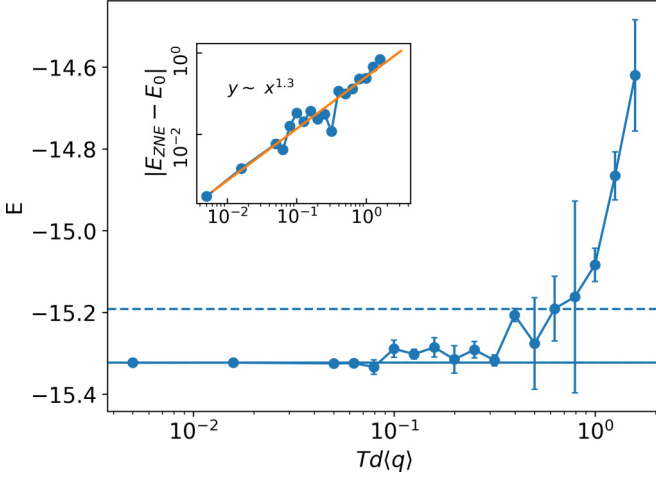


FIG. 5. ZNE estimated energy as a function of the magnitude of circuit errors for  $n = 12$  qubits. The mean values and deviations are estimated from considering three different noise realizations. Solid horizontal line shows the noise-free VQE energy, dashed line is separated from it by the value of the spectral gap.

two-qubit depolarizing noise, as well as for (ii) more realistic two-qubit Pauli error [70]. For each noise model the error rates are sampled from (a) uniform and (b) more realistic lognormal distributions [66,71]. To make comparison possible, the

distributions were normalized to have the same means and variances. For the Pauli error, the means were divided by 15, so that the sum of the error coefficients has the same mean and variance.

For each noise type we simulate the experiments described in Secs. IV A and IV B. For the noiseless VQE we considered circuits of up to 10 layers. As before, for most of the problems this was enough to find energy approximation within 1% of the spectral gap from the true ground state. The exceptions were Ising B for  $n = 10$  and  $n = 12$ , where the error was about 0.5 of the gap, and  $\text{H}_2\text{O}$  for  $n = 12$ , where the error was approximately equal to the gap.

In Table I we present results of ZNE using 50 random permutations and averaged over 20 noise realizations for the discussed noise types and system sizes. Here  $\Delta E$  represents ZNE error for a mean noise strength  $\langle q \rangle = 5 \times 10^{-4}$ . It can be seen that ZNE for different noise models and error distributions demonstrate qualitatively similar results, thus indicating the potential of the proposed protocol. The error growth for larger system sizes is caused by the fixed number of permutations considered and can be countered by its increase, in accordance with Fig. 3.

In addition, in Table I we summarize the threshold strength of the noise  $\langle q \rangle_{\text{thr}}$ , below which ZNE estimated energy differs from noiseless VQE energy by no more than the spectral gap. Those results were obtained by simulating ZNE with

TABLE I. Experimental results for ZNE using 50 permutations with different Hamiltonians and noise models.  $\Delta E$  gives ZNE error produced for mean noise strength  $\langle q \rangle = 5 \times 10^{-4}$  (corresponding to sampling from  $[0, 0.001]$  in the case of uniform distribution). The results are averaged over 20 noise realizations.  $q_{\text{thr}}$  represents the mean noise strength at which ZNE estimation exceeds noiseless VQE energy by the value of spectral gap. The results are averaged over three noise realizations.

Noise type	Hamiltonian					
	Ising A		Ising B		H <sub>2</sub> O	
	$\Delta E$	$ T d\langle q \rangle_{\text{thr}}$	$\Delta E$	$ T d\langle q \rangle_{\text{thr}}$	$\Delta E$	$ T d\langle q \rangle_{\text{thr}}$
$n = 6$						
Depolarizing, uniform	$(-3.4 \pm 23.2) \times 10^{-4}$	$0.83 \pm 0.11$	$(0.0 \pm 1.6) \times 10^{-2}$	$0.14 \pm 0.04$	$(-0.6 \pm 2.0) \times 10^{-3}$	$0.77 \pm 0.15$
Depolarizing, lognormal	$(-1.4 \pm 26.3) \times 10^{-4}$	$0.35 \pm 0.03$	$(-0.2 \pm 1.0) \times 10^{-2}$	$0.09 \pm 0.03$	$(0.2 \pm 1.2) \times 10^{-3}$	$0.50 \pm 0.17$
Pauli, uniform	$(2.2 \pm 5.2) \times 10^{-3}$	$0.65 \pm 0.10$	$(0.2 \pm 2.0) \times 10^{-2}$	$0.10 \pm 0.03$	$(0.1 \pm 5.5) \times 10^{-3}$	$0.41 \pm 0.14$
Pauli, lognormal	$(-7.3 \pm 43.9) \times 10^{-4}$	$0.33 \pm 0.04$	$(0.0 \pm 2.1) \times 10^{-2}$	$0.06 \pm 0.02$	$(0.4 \pm 5.2) \times 10^{-3}$	$0.15 \pm 0.03$
$n = 8$						
Depolarizing, uniform	$(0.5 \pm 7.0) \times 10^{-3}$	$0.81 \pm 0.11$	$(0.2 \pm 3.6) \times 10^{-2}$	$0.18 \pm 0.05$	$(0.6 \pm 2.2) \times 10^{-2}$	$0.18 \pm 0.02$
Depolarizing, lognormal	$(-0.8 \pm 4.2) \times 10^{-3}$	$0.33 \pm 0.03$	$(-0.4 \pm 3.6) \times 10^{-2}$	$0.09 \pm 0.01$	$(-0.2 \pm 2.1) \times 10^{-2}$	$0.13 \pm 0.01$
Pauli, uniform	$(0.2 \pm 7.8) \times 10^{-3}$	$0.53 \pm 0.04$	$(-0.8 \pm 5.9) \times 10^{-2}$	$0.13 \pm 0.03$	$(0.7 \pm 3.6) \times 10^{-2}$	$0.07 \pm 0.01$
Pauli, lognormal	$(0.2 \pm 1.0) \times 10^{-2}$	$0.31 \pm 0.02$	$(-0.1 \pm 5.5) \times 10^{-2}$	$0.06 \pm 0.01$	$(0.2 \pm 5.2) \times 10^{-2}$	$0.07 \pm 0.02$
$n = 10$						
Depolarizing, uniform	$(2.8 \pm 7.2) \times 10^{-3}$	$0.72 \pm 0.06$	$(-0.6 \pm 1.7) \times 10^{-2}$	$0.14 \pm 0.04$	$(0.6 \pm 4.7) \times 10^{-2}$	$0.10 \pm 0.03$
Depolarizing, lognormal	$(0.0 \pm 1.0) \times 10^{-2}$	$0.30 \pm 0.01$	$(-0.3 \pm 1.5) \times 10^{-2}$	$0.09 \pm 0.02$	$(-1.3 \pm 4.7) \times 10^{-2}$	$0.09 \pm 0.02$
Pauli, uniform	$(0.5 \pm 1.5) \times 10^{-2}$	$0.51 \pm 0.04$	$(-0.3 \pm 2.8) \times 10^{-2}$	$0.08 \pm 0.02$	$(-1.6 \pm 8.6) \times 10^{-2}$	$0.09 \pm 0.03$
Pauli, lognormal	$(0.0 \pm 1.4) \times 10^{-2}$	$0.30 \pm 0.03$	$(-0.3 \pm 2.3) \times 10^{-2}$	$0.10 \pm 0.03$	$(1.4 \pm 8.1) \times 10^{-2}$	$0.04 \pm 0.02$
$n = 12$						
Depolarizing, uniform	$(0.2 \pm 1.6) \times 10^{-2}$	$0.57 \pm 0.03$	$(0.2 \pm 1.7) \times 10^{-2}$	$0.18 \pm 0.03$	$0.006 \pm 0.153$	$0.08 \pm 0.01$
Depolarizing, lognormal	$(0.2 \pm 1.4) \times 10^{-2}$	$0.31 \pm 0.03$	$(0.2 \pm 1.7) \times 10^{-2}$	$0.09 \pm 0.02$	$-0.04 \pm 0.11$	$0.06 \pm 0.01$
Pauli, uniform	$(-0.4 \pm 2.0) \times 10^{-2}$	$0.49 \pm 0.04$	$(-0.8 \pm 3.5) \times 10^{-2}$	$0.14 \pm 0.04$	$0.12 \pm 0.48$	$0.09 \pm 0.03$
Pauli, lognormal	$(0.1 \pm 2.7) \times 10^{-2}$	$0.25 \pm 0.03$	$(-0.2 \pm 4.2) \times 10^{-2}$	$0.27 \pm 0.03$	$0.06 \pm 0.43$	$0.04 \pm 0.01$

TABLE II. Extrapolation error for Hamiltonians of  $n = 6$  qubits with different ZNE techniques. The circuit depth was the same as in Fig. 2. The folding extrapolations were obtained from linear extrapolation over four points.

Setup	Pauli twirling	Unitary folding (gates)	Unitary folding (circuit)	This work
Ising A, depolarizing lognormal	$-0.027 \pm 0.030$	$(4.0 \pm 4.4) \times 10^{-7}$	$-0.036 \pm 0.078$	$(-1.4 \pm 26.3) \times 10^{-4}$
Ising B, depolarizing lognormal	$-0.16 \pm 0.19$	$(5.0 \pm 4.6) \times 10^{-6}$	$-0.12 \pm 0.03$	$(-0.2 \pm 1.0) \times 10^{-2}$
H <sub>2</sub> O, depolarizing lognormal	$-0.028 \pm 0.055$	$(7.4 \pm 6.4) \times 10^{-8}$	$-0.015 \pm 0.004$	$(0.2 \pm 1.2) \times 10^{-3}$

50 permutations across a wide range of noise amplitudes, averaging the results over three noise realizations, in the same manner as depicted in Fig. 5. By interpolating the resulting curves we identify the threshold values  $\langle q \rangle_{\text{thr}}$ , at which the curve crosses spectral gap. Again, it is seen that noise strength  $|T|d\langle q \rangle_{\text{thr}} \lesssim 1$  allows one to recover ZNE energy within one spectral gap from the noiseless VQE. Moreover, even for larger values of  $|T|d\langle q \rangle$  the extrapolation is closer to the true ground-state energy than any of the noisy energy values.

#### D. Comparison with existing zero-noise extrapolation techniques

State-of-the-art zero-noise extrapolation techniques rely on increasing the noise by adding extra gates or by implementing the existing ones in a different manner. Up to minor differences, we identified the following techniques in the literature, which can be directly compared with the proposed protocol:

(1) Unitary folding [72]. Parts of the circuit are repeated together with their inverses so that the length of the circuit is increased while preserving the total unitary that would be implemented in the absence of noise. This can be done at the level of the entire circuit or individual gates and layers. In the latter case it is referred to as gate folding [73,74].

(2) Pauli twirling and Pauli gate insertion [40]. With Pauli twirling, any local noise channel is transformed into a stochastic Pauli channel. Afterward, stochastic Pauli noise is amplified by randomly adding Pauli gates.

To compare the techniques, we took the depolarizing noise with lognormal error rate distribution and implemented the different ZNE techniques. As earlier, the ZNE error is averaged over 20 random realizations of noise. The results are shown in Table II. Our technique shows better results than circuit-level unitary folding and Pauli twirling, while looses to the gate-level unitary folding. We note here that unitary folding necessitates longer circuits, which naturally increase gate count and execution time, making it more vulnerable to crosstalk errors and limited coherence time. Therefore, we conclude that our protocol produces at worst comparable results to the state of the art techniques, without suffering from some of their drawbacks.

#### V. CONCLUSIONS

Existing studies on the behavior of VQE-estimated energy with respect to noise primarily focus on homogeneous error models [28,75–77]. While a linear dependence between energy and the noise strength is well established in this case, it does not automatically translate to the case of inhomogeneous noise. Motivated by the fact that gate errors are inhomogeneously distributed across pairs of qubits in a physical hardware, we propose a method for zero-noise extrapolation. In particular, we showed that, by changing the abstract-to-

physical qubit mapping, it is possible to vary the level of noise in a quantum circuit in a controllable way. This enables us to execute ZNE in an experimentally amenable manner.

While the proposed ZNE protocol can be applied to estimate noiseless expectation value of any observable with respect to any quantum circuit, for the purpose of demonstration we apply this method to VQE. We found that the energy approximated by a noisy VQE circuit is approximately linear with respect to the circuit error sum. An analytic bound was derived to quantify the quality of the linear approximation. We found that this bound depends on the nature of the problem Hamiltonian and the error rates. We numerically demonstrated that with the proposed ZNE protocol one can approximate the noiseless VQE energy with high accuracy, reducing the noise-induced error by approximately two orders of magnitude. This was verified for a range of problem Hamiltonians, noise models and error distributions. Moreover, we proved that for a specific class of *Ansätze* the proposed protocol recovers the exact noise free VQE energy.

Next we examined how the performance of our proposed protocol gets impacted when considering permutation pools of limited sizes for the extrapolation; indeed considering all permutations maybe unfeasible especially for larger system sizes. We show that the ZNE extrapolation error is smaller than all energy scales of the problem, even when extrapolating over less than 10% of all possible permutations. Furthermore, we demonstrated that while increasing the system size (while keeping the size of the permutation pool fixed) does induce some instability in the ZNE protocol, it can be removed with at most polynomially many permutations considered for extrapolation.

Finally, we compared the performance of the proposed protocol to the state-of-the-art ZNE techniques. We showed that the performance of our hardware inspired ZNE, in the worst case, is comparable to the state of the art techniques, while not suffering from some of their drawbacks.

The code and the data produced in the work are available upon reasonable request.

#### ACKNOWLEDGMENTS

We thank Lianna Akopyan for stimulating discussions. O.L. and K.L. acknowledge the framework of the Roadmap for Quantum Computing (Contracts No. 868-1.3-15/15-2021 and No. R2163).

#### APPENDIX A: RELATIVE DEVIATION FROM LINEAR TREND IN EQ. (7)

We begin our analysis with a simplified form of (7) for the sake of brevity:

$$y = \Delta x + \delta y, \quad (\text{A1})$$

where  $y = (E - E_0)$ ,  $\Delta = (A - E_0)$ ,  $x = \sum_{(j,k) \in T} dq_{jk}$ , and  $\delta y = \sum_{(j,k) \in T} (\sum_{l \in [1,d]} \epsilon_{jk}^l) q_{jk} = \sum_{(j,k) \in T} \tilde{\epsilon}_{jk} q_{jk}$ . Taking into account that both  $x$  and  $\delta y$  are error-rate-dependent terms, one can consequently conclude that  $\delta y$  represents a deviation from the linear dependence  $y = \Delta x$ . Nevertheless, one can upper bound the deviation  $\delta y$  between the actual function  $y$  as per (A1) and the linear dependence  $y = \Delta x$ . To accomplish this we consider the relative deviation:

$$\frac{|\delta y|}{\Delta x} = \frac{|\sum_{(j,k) \in T} \tilde{\epsilon}_{jk} q_{jk}|}{d(A - E_0) \sum_{(j,k) \in T} q_{jk}}. \quad (\text{A2})$$

One can readily establish a bound on this quantity as

$$\frac{|\delta y|}{\Delta x} \leq \frac{\max\{|\tilde{\epsilon}_{jk}|\}_{(j,k) \in T}}{d(A - E_0)} = B_1. \quad (\text{A3})$$

This bound is not necessarily tight; however, one can improve it by taking into account the distribution of the error rates themselves:

$$\frac{|\delta y|}{\Delta x} \leq \frac{|T| \max\{|\tilde{\epsilon}_{jk}|\}_{(j,k) \in T} \text{Diam}\{q_{jk}\}_{(j,k) \in T}}{d(A - E_0) \sum_{(j,k) \in T} q_{jk}} = B_2. \quad (\text{A4})$$

Here  $\text{Diam}\{A\}$  is the diameter of a set  $A$ . Combining (A5) and (A4) we arrive at the final result:

$$\frac{|\delta y|}{\Delta x} \leq \min\{B_1, B_2\}. \quad (\text{A5})$$

## APPENDIX B: THE ZERO-NOISE EXTRAPOLATION OVER ALL PERMUTATIONS IS EXACT

Consider the function

$$y = \Delta x + \delta y. \quad (\text{B1})$$

We note that (B1) maps to (8) when  $y = (E - E_0)$ ,  $\Delta = (A - E_0)$ ,  $x = \sum_{(j,k) \in T} d\bar{q}_{\pi(jk)}$ , and  $\delta y = \sum_{(j,k) \in T} (\sum_{l \in [1,d]} \epsilon_{jk}^l) \bar{q}_{\pi(jk)} = \sum_{(j,k) \in T} \tilde{\epsilon}_{jk} \bar{q}_{\pi(jk)}$ . In our approach of zero-noise extrapolation we fit a linear approximation of the form  $y = \alpha + \beta x$  to the function in (B1). The constants  $\alpha$  and  $\beta$  are inferred by the least squares method where they have standard expressions in terms of sample statistics:

$$\alpha = \langle y \rangle - \beta \langle x \rangle, \quad (\text{B2})$$

$$\beta = \frac{\text{Cov}(x, y)}{\text{Var } x}. \quad (\text{B3})$$

Here the angular brackets denote averaging over all samples (in our case, over all abstract-to-physical qubit mappings

labeled by  $\pi \in S_n$ ). Substituting (B1) and (B3) in (B2), we obtain

$$\alpha = -\langle x \rangle \frac{\text{Cov}(x, \delta y)}{\text{Var } x}. \quad (\text{B4})$$

Recall that for an exact zero-noise extrapolation we would get  $\alpha = 0$ . In deriving (B4) we have used the result  $\langle \delta y \rangle = 0$ . This follows from two facts:

(1)  $\langle \bar{q}_{\pi(jk)} \rangle = \frac{1}{n!} \sum_{\pi \in S_n} \bar{q}_{\pi(jk)} = \bar{q}_a$ , where  $\bar{q}_a$  is the average error.

(2)  $\sum_{(j,k) \in T} \tilde{\epsilon}_{jk} = 0$  by definition.

Now we consider the covariance  $\text{Cov}(x, \delta y)$  in more detail. Owing to the fact that  $\langle \delta y \rangle = 0$  we get

$$\text{Cov}(x, \delta y) = \langle x \delta y \rangle = \frac{1}{n!} \sum_{\substack{(j_1, k_1) \in T \\ (j_2, k_2) \in T}} \tilde{\epsilon}_{j_1 k_1} \langle \bar{q}_{\pi(j_1 k_1)} \bar{q}_{\pi(j_2 k_2)} \rangle. \quad (\text{B5})$$

We now group the summands in (B5) such that (a)  $j_2 = j_1, k_2 = k_1$ , (b)  $j_2 = j_1, k_2 \neq k_1$  or  $j_2 \neq j_1, k_2 = k_1$ , and (c)  $j_2 \neq j_1, k_2 \neq k_1$ . Denoting the average  $\langle \bar{q}_{\pi(j_1 k_1)} \bar{q}_{\pi(j_2 k_2)} \rangle$  for the aforementioned cases to be  $\kappa_2, \kappa_1$ , and  $\kappa_0$ , respectively, we arrive at the expression

$$\text{Cov}(x, \delta y) = \sum_{(jk) \in T} \tilde{\epsilon}_{jk} [n_2^{(jk)} \kappa_2 + n_1^{(jk)} \kappa_1 + n_0^{(jk)} \kappa_0]. \quad (\text{B6})$$

Here  $n_2^{(jk)}$  is the number of two-qubit gates acting on the qubit pair  $(j, k)$ ,  $n_1^{(jk)}$  is the number of two-qubit gates acting either on the qubit  $j$  or  $k$ , and  $n_0^{(jk)}$  is the number of gates that does not act either on  $j$  or  $k$ . If these three values are independent on  $(j, k)$ , the covariance  $\text{Cov}(x, \delta y)$  vanishes owing to the fact that  $\sum_{(j,k) \in T} \tilde{\epsilon}_{jk} = 0$  implying that  $\alpha = 0$ , which is what we expected for an exact zero-noise extrapolation.

We note here that  $n_2^{(jk)}, n_1^{(jk)}, n_0^{(jk)}$  are independent of  $(j, k)$  only for specific *Ansatz* circuits. To characterize the structure of such *Ansätze* completely we depict the *Ansatz* circuit as a multigraph with vertices corresponding to qubits and edges corresponding to two-qubit gates. Indeed  $n_2^{(jk)}$  is the count of gates acting on  $(j, k)$ ,  $n_1^{(jk)} = \text{deg}(j) + \text{deg}(k) - n_2^{(jk)}$ , and  $n_0^{(jk)} = |T|d - n_1^{(jk)} - n_2^{(jk)}$ . Thus, these values only depend on degrees of vertices and multiplicities of edges. This implies that the gate-independence condition can be satisfied by taking a regular graph (e.g., a cycle) as the structure of the *Ansatz*.

- [1] M. P. Harrigan, K. J. Sung, M. Neeley, K. J. Satzinger, F. Arute, K. Arya, J. Atalaya, J. C. Bardin, R. Barends, S. Boixo *et al.*, Quantum approximate optimization of non-planar graph problems on a planar superconducting processor, *Nat. Phys.* **17**, 332 (2021).
- [2] G. Pagano, A. Bapat, P. Becker, K. S. Collins, A. De, P. W. Hess, H. B. Kaplan, A. Kyprianidis, W. L. Tan, C. Baldwin, L. T. Brady, A. Deshpande, F. Liu, S. Jordan, A. V. Gorshkov, and C. Monroe, Quantum approximate

optimization of the long-range Ising model with a trapped-ion quantum simulator, *Proc. Natl. Acad. Sci. USA* **117**, 25396 (2020).

- [3] G. G. Guerreschi and A. Y. Matsuura, QAOA for Max-Cut requires hundreds of qubits for quantum speed-up, *Sci. Rep.* **9**, 6903 (2019).
- [4] A. Butko, G. Micheliannakis, S. Williams, C. Iancu, D. Donofrio, J. Shalf, J. Carter, and I. Siddiqi, Understanding quantum control processor capabilities, and limitations through



- circuit characterization, in *2020 International Conference on Rebooting Computing (ICRC), Atlanta, GA, USA* (IEEE, Piscataway, NJ, 2020), pp. 66–75.
- [5] M. Y. Niu, S. Lu, and I. L. Chuang, Optimizing QAOA: Success probability, and runtime dependence on circuit depth, [arXiv:1905.12134](#).
- [6] E. Farhi, J. Goldstone, and S. Gutmann, A quantum approximate optimization algorithm, [arXiv:1411.4028](#).
- [7] S. Lloyd, Quantum approximate optimization is computationally universal, [arXiv:1812.11075](#).
- [8] M. E. S. Morales, J. D. Biamonte, and Z. Zimborás, On the universality of the quantum approximate optimization algorithm, *Quantum Inf. Process.* **19**, 291 (2020).
- [9] E. Farhi and A. W. Harrow, Quantum supremacy through the quantum approximate optimization algorithm, [arXiv:1602.07674](#).
- [10] V. Akshay, H. Philathong, M. E. S. Morales, and J. D. Biamonte, Reachability deficits in quantum approximate optimization, *Phys. Rev. Lett.* **124**, 090504 (2020).
- [11] E. Farhi, J. Goldstone, S. Gutmann, and L. Zhou, The quantum approximate optimization algorithm, and the Sherrington-Kirkpatrick model at infinite size, *Quantum* **6**, 759 (2022).
- [12] M. M. Wauters, G. B. Mbeng, and G. E. Santoro, Polynomial scaling of the quantum approximate optimization algorithm for ground-state preparation of the fully connected  $p$ -spin ferromagnet in a transverse field, *Phys. Rev. A* **102**, 062404 (2020).
- [13] V. Akshay, D. Rabinovich, E. Campos, and J. Biamonte, Parameter concentrations in quantum approximate optimization, *Phys. Rev. A* **104**, L010401 (2021).
- [14] E. Campos, D. Rabinovich, V. Akshay, and J. Biamonte, Training saturation in layerwise quantum approximate optimisation, *Phys. Rev. A* **104**, L030401 (2021).
- [15] D. Rabinovich, R. Sengupta, E. Campos, V. Akshay, and J. Biamonte, Progress towards analytically optimal angles in quantum approximate optimisation, *Mathematics* **10**, 2601 (2022).
- [16] V. Akshay, H. Philathong, E. Campos, D. Rabinovich, I. Zacharov, X.-M. Zhang, and J. D. Biamonte, Circuit depth scaling for quantum approximate optimization, *Phys. Rev. A* **106**, 042438 (2022).
- [17] D. Rabinovich, S. Adhikary, E. Campos, V. Akshay, E. Anikin, R. Sengupta, O. Lakhmanskaya, K. Lakhmanskii, and J. Biamonte, Ion-native variational ansatz for quantum approximate optimization, *Phys. Rev. A* **106**, 032418 (2022).
- [18] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O’Brien, A variational eigenvalue solver on a photonic quantum processor, *Nat. Commun.* **5**, 4213 (2014).
- [19] Y. Cao, J. Romero, J. P. Olson, M. Degroote, P. D. Johnson, M. Kieferová, I. D. Kivlichan, T. Menke, B. Peropadre, N. P. D. Sawaya *et al.*, Quantum chemistry in the age of quantum computing, *Chem. Rev. (Washington, DC, U. S.)* **119**, 10856 (2019).
- [20] S. McArdle, S. Endo, A. Aspuru-Guzik, S. C. Benjamin, and X. Yuan, Quantum computational chemistry, *Rev. Mod. Phys.* **92**, 015003 (2020).
- [21] D. Wecker, M. B. Hastings, and M. Troyer, Progress towards practical quantum variational algorithms, *Phys. Rev. A* **92**, 042303 (2015).
- [22] B. Bauer, D. Wecker, A. J. Millis, M. B. Hastings, and M. Troyer, Hybrid quantum-classical approach to correlated materials, *Phys. Rev. X* **6**, 031045 (2016).
- [23] J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, and S. Lloyd, Quantum machine learning, *Nature (London)* **549**, 195 (2017).
- [24] M. Schuld, and N. Killoran, Quantum machine learning in feature Hilbert spaces, *Phys. Rev. Lett.* **122**, 040504 (2019).
- [25] S. Adhikary, S. Dangwal, and D. Bhowmik, Supervised learning with a quantum classifier using multi-level systems, *Quantum Inf. Process.* **19**, 89 (2020).
- [26] S. Adhikary, Entanglement assisted training algorithm for supervised quantum classifiers, *Quantum Inf. Process.* **20**, 254 (2021).
- [27] V. Havlíček, A. D. Córcoles, K. Temme, A. W. Harrow, A. Kandala, J. M. Chow, and J. M. Gambetta, Supervised learning with quantum-enhanced feature spaces, *Nature (London)* **567**, 209 (2019).
- [28] E. Fontana, N. Fitzpatrick, D. Muñoz Ramo, R. Duncan, and I. Rungger, Evaluating the noise resilience of variational quantum algorithms, *Phys. Rev. A* **104**, 022403 (2021).
- [29] S. Wang, E. Fontana, M. Cerezo, K. Sharma, A. Sone, L. Cincio, and P. J. Coles, Noise-induced barren plateaus in variational quantum algorithms, *Nat. Commun.* **12**, 6961 (2021).
- [30] Z. Cai, R. Babbush, S. C. Benjamin, S. Endo, W. J. Huggins, Y. Li, J. R. McClean, and T. E. O’Brien, Quantum error mitigation, *Rev. Mod. Phys.* **95**, 045005 (2023).
- [31] S. Bravyi, S. Sheldon, A. Kandala, D. C. McKay, and J. M. Gambetta, Mitigating measurement errors in multiqubit experiments, *Phys. Rev. A* **103**, 042605 (2021).
- [32] P. Czarnik, A. Arrasmith, P. J. Coles, and L. Cincio, Error mitigation with Clifford quantum-circuit data, *Quantum* **5**, 592 (2021).
- [33] S. Endo, S. C. Benjamin, and Y. Li, Practical quantum error mitigation for near-future applications, *Phys. Rev. X* **8**, 031027 (2018).
- [34] A. Lowe, M. H. Gordon, P. Czarnik, A. Arrasmith, P. J. Coles, and L. Cincio, Unified approach to data-driven quantum error mitigation, *Phys. Rev. Res.* **3**, 033098 (2021).
- [35] F. B. Maciejewski, Z. Zimborás, and M. Oszmaniec, Mitigation of readout noise in near-term quantum devices by classical post-processing based on detector tomography, *Quantum* **4**, 257 (2020).
- [36] B. Pokharel, N. Anand, B. Fortman, and D. A. Lidar, Demonstration of fidelity improvement using dynamical decoupling with superconducting qubits, *Phys. Rev. Lett.* **121**, 220502 (2018).
- [37] K. Temme, S. Bravyi, and J. M. Gambetta, Error mitigation for short-depth quantum circuits, *Phys. Rev. Lett.* **119**, 180509 (2017).
- [38] L. Viola and S. Lloyd, Dynamical suppression of decoherence in two-state quantum systems, *Phys. Rev. A* **58**, 2733 (1998).
- [39] S. Zhang, Y. Lu, K. Zhang, W. Chen, Y. Li, J.-N. Zhang, and K. Kim, Error-mitigated quantum gates exceeding physical fidelities in a trapped-ion system, *Nat. Commun.* **11**, 587 (2020).
- [40] Y. Li, and S. C. Benjamin, Efficient variational quantum simulator incorporating active error minimization, *Phys. Rev. X* **7**, 021050 (2017).
- [41] A. Kandala, K. Temme, A. D. Córcoles, A. Mezzacapo, J. M. Chow, and J. M. Gambetta, Error mitigation extends the com-

- putational reach of a noisy quantum processor, *Nature (London)* **567**, 491 (2019).
- [42] I. Pogorelov, T. Feldker, Ch. D. Marciniak, L. Postler, G. Jacob, O. Kriegelsteiner, V. Podlesnic, M. Meth, V. Negnevitsky, M. Stadler, B. Höfer, C. Wächter, K. Lakhmanskiy, R. Blatt, P. Schindler, and T. Monz, Compact ion-trap quantum computing demonstrator, *PRX Quantum* **2**, 020343 (2021).
- [43] K. Wright, K. M. Beck, S. Debnath, J. M. Amini, Y. Nam, N. Grzesiak, J.-S. Chen, N. C. Pistenti, M. Chmielewski, C. Collins, K. M. Hudek, J. Mizrahi, J. D. Wong-Campos, S. Allen, J. Apisdorf, P. Solomon, M. Williams, A. M. Ducore, A. Blinov, S. M. Kreikemeier, V. Chaplin, M. Keesan, C. Monroe, and J. Kim, Benchmarking an 11-qubit quantum computer, *Nat. Commun.* **10**, 5464 (2019).
- [44] J. Biamonte, Universal variational quantum computation, *Phys. Rev. A* **103**, L030401 (2021).
- [45] H. R. Grimsley, S. E. Economou, E. Barnes, and N. J. Mayhall, An adaptive variational algorithm for exact molecular simulations on a quantum computer, *Nat. Commun.* **10**, 3007 (2019).
- [46] H. L. Tang, V. O. Shkolnikov, G. S. Barron, H. R. Grimsley, N. J. Mayhall, E. Barnes, and S. E. Economou, Qubit-ADAPT-VQE: An adaptive algorithm for constructing hardware-efficient Ansätze on a quantum processor, *PRX Quantum* **2**, 020310 (2021).
- [47] S. Sim, J. Romero, J. F. Gonthier, and A. A. Kunitsa, Adaptive pruning-based optimization of parameterized quantum circuits, *Quantum Sci. Technol.* **6**, 025019 (2021).
- [48] M. Bilkis, M. Cerezo, G. Verdon, P. J. Coles, and L. Cincio, A semi-agnostic ansatz with variable structure for variational quantum algorithms, *Quantum Machine Intelligence* **5**, 43 (2023).
- [49] M. D. Sapova and A. K. Fedorov, Variational quantum eigensolver techniques for simulating carbon monoxide oxidation, *Commun. Phys.* **5**, 199 (2022).
- [50] K. Mitarai, M. Negoro, M. Kitagawa, and K. Fujii, Quantum circuit learning, *Phys. Rev. A* **98**, 032309 (2018).
- [51] M. Schuld, V. Bergholm, C. Gogolin, J. Izaac, and N. Killoran, Evaluating analytic gradients on quantum hardware, *Phys. Rev. A* **99**, 032331 (2019).
- [52] O. Kyriienko and V. E. Elfving, Generalized quantum circuit differentiation rules, *Phys. Rev. A* **104**, 052417 (2021).
- [53] V. Verteletskiy, T.-C. Yen, and A. F. Izmaylov, Measurement optimization in the variational quantum eigensolver using a minimum clique cover, *J. Chem. Phys.* **152**, 124114 (2020).
- [54] P. Kl Barkoutsos, J. F. Gonthier, I. Sokolov, N. Moll, G. Salis, A. Fuhrer, M. Ganzhorn, D. J. Egger, M. Troyer, A. Mezzacapo, S. Filipp, and I. Tavernelli, Quantum algorithms for electronic structure calculations: Particle/hole Hamiltonian, and optimized wavefunction expansions, *Phys. Rev. A* **98**, 022322 (2018).
- [55] J. Tilly, H. Chen, S. Cao, D. Picozzi, K. Setia, Y. Li, E. Grant, L. Wossnig, I. Rungger, G. H. Booth, and J. Tennyson, The variational quantum eigensolver: A review of methods, and best practices, *Phys. Rep.* **986**, 1 (2022).
- [56] V. E. Elfving, B. W. Broer, M. Webber, J. Gavartin, M. D. Halls, K. P. Lorton, and A. Bochevarov, How will quantum computers provide an industrially relevant computational advantage in quantum chemistry? [arXiv:2009.12472](https://arxiv.org/abs/2009.12472).
- [57] A. Kandala, A. Mezzacapo, K. Temme, M. Takita, M. Brink, J. M. Chow, and J. M. Gambetta, Hardware-efficient variational quantum eigensolver for small molecules, and quantum magnets, *Nature (London)* **549**, 242 (2017).
- [58] R. Shaydulin and M. Pistoia, QAOA with  $n \cdot p \geq 200$ , in *2023 IEEE International Conference on Quantum Computing and Engineering (QCE), Bellevue, WA, USA* (IEEE, Piscataway, NJ, 2023), pp. 1074–1077.
- [59] E. Pelofske, A. Bärttschi, J. Golden, and S. Eidenbenz, High-round QAOA for max  $k$ -sat on trapped ion NISQ devices, in *2023 IEEE International Conference on Quantum Computing, and Engineering (QCE) Bellevue, WA, USA* (IEEE, Piscataway, NJ, 2023), Vol. 01, pp. 506–517.
- [60] S. H. Sack and D. J. Egger, Large-scale quantum approximate optimization on nonplanar graphs with machine learning noise mitigation, *Phys. Rev. Res.* **6**, 013223 (2024).
- [61] N. Akerman, N. Navon, S. Kotler, Y. Glickman, and R. Ozeri, Universal gate-set for trapped-ion qubits using a narrow linewidth diode laser, *New J. Phys.* **17**, 113060 (2015).
- [62] C. J. Ballance, T. P. Harty, N. M. Linke, M. A. Sepiol, and D. M. Lucas, High-fidelity quantum logic gates using trapped-ion hyperfine qubits, *Phys. Rev. Lett.* **117**, 060504 (2016).
- [63] A. Bermudez, X. Xu, R. Nigmatullin, J. O’Gorman, V. Negnevitsky, P. Schindler, T. Monz, U. G. Poschinger, C. Hempel, J. Home, F. Schmidt-Kaler, M. Biercuk, R. Blatt, S. Benjamin, and M. Müller, Assessing the progress of trapped-ion processors towards fault-tolerant quantum computation, *Phys. Rev. X* **7**, 041061 (2017).
- [64] H. Levine, A. Keesling, A. Omran, H. Bernien, S. Schwartz, A. S. Zibrov, M. Endres, M. Greiner, V. Vuletić, and M. D. Lukin, High-fidelity control, and entanglement of Rydberg-atom qubits, *Phys. Rev. Lett.* **121**, 123603 (2018).
- [65] C. D. Bruzewicz, J. Chiaverini, R. McConnell, and J. M. Sage, Trapped-ion quantum computing: Progress, and challenges, *Appl. Phys. Rev.* **6**, 021314 (2019).
- [66] A. Morvan, B. Villalonga, X. Mi, S. Mandrà, A. Bengtsson, P. V. Klimov, Z. Chen, S. Hong, C. Erickson, I. K. Drozdov, J. Chau, G. Laun, R. Movassagh, A. Asfaw, L. T. A. N. Brandão, R. Peralta, D. Abanin, R. Acharya, R. Allen *et al.*, Phase transition in random circuit sampling, [arXiv:2304.11119](https://arxiv.org/abs/2304.11119).
- [67] K. Kuroiwa and Y. O. Nakagawa, Penalty methods for a variational quantum eigensolver, *Phys. Rev. Res.* **3**, 013197 (2021).
- [68] L. Zhou, S.-T. Wang, Soonwon Choi, H. Pichler, and M. D. Lukin, Quantum approximate optimization algorithm: Performance, mechanism, and implementation on near-term devices, *Phys. Rev. X* **10**, 021067 (2020).
- [69] A. Skolik, J. R. McClean, M. Mohseni, P. V. Der Smagt, and M. Leib, Layerwise learning for quantum neural networks, *Quantum Mach. Intell.* **3**, 5 (2021).
- [70] J. J. Wallman and J. Emerson, Noise tailoring for scalable quantum computation via randomized compiling, *Phys. Rev. A* **94**, 052325 (2016).
- [71] Y. Kim, A. Eddins, S. Anand, K. X. Wei, E. V. Den Berg, S. Rosenblatt, H. Nayfeh, Y. Wu, M. Zaletel, K. Temme *et al.*, Evidence for the utility of quantum computing before fault tolerance, *Nature (London)* **618**, 500 (2023).
- [72] T. Giurgica-Tiron, Y. Hindy, R. LaRose, A. Mari, and W. J. Zeng, Digital zero noise extrapolation for quantum error mitigation, in *2020 IEEE International Conference on Quantum Computing, and Engineering (QCE), Denver, CO, USA* (IEEE, Piscataway, NJ, 2020), pp. 306–316.

- [73] A. He, B. Nachman, W. A. de Jong, and C. W. Bauer, Resource efficient zero noise extrapolation with identity insertions, *Phys. Rev. A* **102**, 012426 (2020).
- [74] E. F. Dumitrescu, A. J. McCaskey, G. Hagen, G. R. Jansen, T. D. Morris, T. Papenbrock, R. C. Pooser, D. J. Dean, and P. Lougovski, Cloud quantum computing of an atomic nucleus, *Phys. Rev. Lett.* **120**, 210501 (2018).
- [75] K. Dalton, C. K. Long, Y. S. Yordanov, C. G. Smith, C. H. W. Barnes, N. Mertig, and D. R. M. Arvidsson-Shukur, Quantifying the effect of gate errors on variational quantum eigensolvers for quantum chemistry, *npj Quantum Inf.* **10**, 18 (2024).
- [76] J. Kattemölle and G. Burkard, Ability of error correlations to improve the performance of variational quantum algorithms, *Phys. Rev. A* **107**, 042426 (2023).
- [77] D. Rabinovich, E. Campos, S. Adhikary, E. Pankovets, D. Vinichenko, and J. Biamonte, Robustness of variational quantum algorithms against stochastic parameter perturbation, *Phys. Rev. A* **109**, 042426 (2024).