





**Quantum algorithms with local particle-number conservation: Noise effects and error correction**Michael Streif <sup>1,2,3,4</sup> Martin Leib <sup>3</sup> Filip Wudarski <sup>1,2</sup> Eleanor Rieffel,<sup>1</sup> and Zhihui Wang <sup>1,2</sup><sup>1</sup>*Quantum Artificial Intelligence Laboratory (QuAIL), NASA Ames Research Center, Moffett Field, California 94035, USA*<sup>2</sup>*USRA Research Institute for Advanced Computer Science (RIACS), Mountain View, California 94043, USA*<sup>3</sup>*Data:Lab, Volkswagen Group, Ungererstrasse 69, 80805 München, Germany*<sup>4</sup>*University Erlangen-Nürnberg (FAU), Institute of Theoretical Physics, Staudtstrasse 7, 91058 Erlangen, Germany*

(Received 20 November 2020; accepted 8 March 2021; published 13 April 2021)

Quantum circuits with local particle-number conservation restrict the quantum computation to a subspace of the Hilbert space of the qubit register. In a noiseless or fault-tolerant quantum computation, such quantities are preserved. In the presence of noise, however, the evolution's symmetry could be broken and nonvalid states could be sampled at the end of the computation. On the other hand, the restriction to a subspace in the ideal case suggests the possibility of more resource-efficient error-mitigation techniques for circuits preserving symmetries that are not possible for general circuits. Here, we analyze the probability of staying in such symmetry-preserved subspaces under noise, providing an exact formula for local depolarizing noise. We apply our findings to benchmark, under depolarizing noise, the symmetry robustness of the  $XY$  quantum alternating operator Ansatz, which has local particle-number conserving symmetries and is a special case of the quantum alternating operator Ansatz. We also analyze the influence of the choice of encoding the problem on the symmetry robustness of the algorithm and discuss a simple adaptation of the bit-flip code to correct for symmetry-breaking errors with reduced resources.

DOI: [10.1103/PhysRevA.103.042412](https://doi.org/10.1103/PhysRevA.103.042412)**I. INTRODUCTION**

Beginning with Shor's algorithm [1] and Grover's algorithm [2] in the 1990s, dozens of algorithms [3] have been found with provable speedups against the best classical algorithms known, and in some cases against any possible classical algorithm. For many problems, however, it is as yet unclear whether quantum computing provides an advantage and, if so, how significant an advantage and by what means. Even for those for which quantum algorithms are established, the resources required to realize these gains are far greater than current or near-term quantum hardware support. For example, for problem sizes of cryptographic interest, Shor's algorithm requires millions of qubits [4], in stark contrast to today's state-of-the-art devices with tens of qubits [5–7]. The recent demonstration [8] of a noisy intermediate-scale quantum (NISQ) device solving computational tasks—though artificial ones—in a matter of minutes, which state-of-the-art supercomputers would require hours, days, or even years to solve, provides evidence that even the small noisy devices of today can be faster than the largest classical computing systems on some problems. An open question is whether NISQ devices will be able to outperform classical computations on problems of practical interest, particularly in areas of broad application such as optimization and quantum simulation. Whether they can or not, they present an unprecedented opportunity to explore quantum algorithms empirically, providing insights into algorithms to be run on larger-scale, fault-tolerant quantum computers of the future. The practical usefulness of NISQ devices for any of these goals depends on better understanding

the effect of noise on various classes of quantum algorithms, and of techniques to mitigate errors on NISQ devices.

Algorithms that preserve symmetries have potential advantages with respect to performance, robustness, ease of analysis, and use of symmetry-aware error-mitigation techniques. Such algorithms may, however, be particularly susceptible to noise, which in general will not respect such symmetries. Here, we consider one such symmetry, the preservation of particle number (or Hamming weight), which appears, for example, in certain quantum alternating operator Ansatz (QAOA) algorithms, and in quantum variational eigensolver algorithms [9–12].

We examine the effect of noise, and potential error-mitigation approaches, on general quantum circuits that preserve particle number and look specifically at the effects of noise and error-mitigation approaches for  $XY$ -QAOA, a special case of the quantum alternating operator Ansatz (QAOA) [13], itself a generalization of the framework used in the quantum approximate optimization algorithm [14]. The original version of the quantum approximate optimization algorithm, which uses a transverse-field mixer, and therefore we dub  $X$ -QAOA, was designed to find approximate solutions to unconstrained optimization problems. Most industry-relevant optimization problems, however, include constraints.  $XY$ -QAOA uses an  $XY$  mixer to preserve the particle number. Recent numerical simulations [15] confirmed intuitions that at least in the noiseless case,  $XY$ -QAOA substantially outperforms  $X$ -QAOA for problems with the appropriate symmetries.

The main contributions of this work are the following:

(a) An analytically derived exact formula for the probability of preserving particle number under locally homogeneous noise such as that of a depolarizing channel for any circuit that preserves particle number in the noiseless case.

(b) Calculation of the probability of preserving particle number as a function of the depth and error rate for XY-QAOA, with specific applications to QAOA applied to the max- $\kappa$ -colorable-subgraph problem, with the coloring of each vertex mapped to a particle-number 1 subspace.

(c) Demonstration of the influence of routing the quantum circuits on realistic two-dimensional (2D) grids on the probability to preserve the local particle-number conservation (LPNC).

(d) Comparison of alternative mappings of the max- $\kappa$ -colorable-subgraph problem to higher particle-number subspaces and their relative robustness.

(e) Construction of a bit-flip code for quantum algorithms with LPNC utilizing the additional symmetries to reduce the computational resources compared to a standard bit-flip code.

This paper is structured as follows. In Sec. II, we define a class of quantum circuits with LPNC and calculate the probability of upholding the symmetry during a computation in the presence of local depolarizing noise. In Sec. III, we benchmark the robustness of XY-QAOA applied to the max- $\kappa$ -colorable-subgraph problem under the influence of noise. In Sec. IV, we analyze how the choice of encoding the problem affects the robustness to noise and benchmark an adaption of the bit-flip code to correct for symmetry-breaking errors. In Sec. V, we conclude and give an outlook.

## II. QUANTUM ALGORITHMS WITH PARTICLE-NUMBER SYMMETRIES UNDER NOISE

In this section, we analyze quantum algorithms with conserved particle numbers under noise and give an analytical expression for the probably of leaving a fixed particle-number subspace.

We assume a system composed of  $n$  subsystems with  $\kappa$  qubits each. We initialize each system with a fixed particle number  $N$ , that is, each system has  $N$  qubits in the  $|1\rangle$  state and  $\kappa - N$  qubits in the  $|0\rangle$  state. For example, for the  $N = 1$  particle-number subspace, each system has only a single particle.

The subsystem states with particle number  $N$  span a subspace, which we refer to as *particle-number subspace*  $\mathcal{H}_N$ . The tensor product of all  $n$  subspaces defines the *feasible subspace* on all  $n$  subsystems,  $\mathcal{H}_{\text{feas}} = \mathcal{H}_N^{(1)} \otimes \mathcal{H}_N^{(2)} \otimes \dots \otimes \mathcal{H}_N^{(n)}$ . We, moreover, define a set of one- and two-qubit unitaries  $G$  such that each element  $g \in G$  upholds the local particle-number conservation (LPNC) in each subsystem, that is,

$$gN_i = N_i g \quad \forall g \in G \quad \forall i, \quad (1)$$

with the particle-number operator

$$N_i = \sum_k^{\dim \mathcal{H}_N} \frac{1}{2} (\mathbb{1}^k - \sigma_z^{(i_k)}), \quad (2)$$

with  $i_k$  the  $k$ th qubit of subsystem  $i$  counting how many qubits are in the  $|1\rangle$  state in subsystem  $i$ , i.e.,  $g|_{\mathcal{H}_{\text{feas}}} : \mathcal{H}_{\text{feas}} \rightarrow$

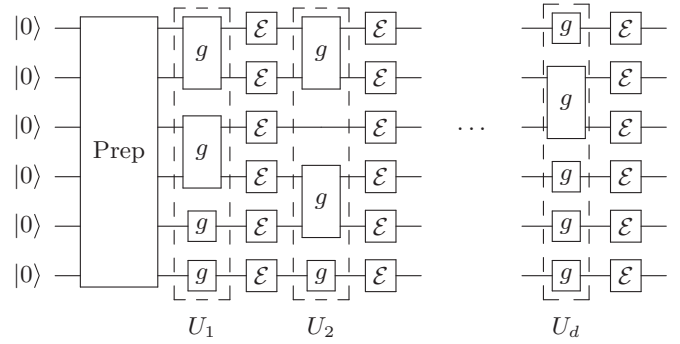


FIG. 1. A circuit illustrating our setup. After initializing the system in a state of  $\mathcal{H}_{\text{feas}}$ , randomly drawn one- and two-qubit unitaries from  $G$  are applied (note that each  $g$  can be a different operation, and we omitted distinct labels for brevity of the notation), followed by local depolarizing channels  $\mathcal{E}$ .

$\mathcal{H}_{\text{feas}}$ . We also introduce a local depolarizing noise channel  $\mathcal{E}(\rho) : \rho \rightarrow \sum_i K_i \rho K_i^\dagger$  on each qubit by defining the Kraus operators,

$$\begin{aligned} K_1 &= \sqrt{1-\eta} \mathbb{1}, & K_2 &= \sqrt{\frac{\eta}{3}} \sigma_x, \\ K_3 &= \sqrt{\frac{\eta}{3}} \sigma_y, & K_4 &= \sqrt{\frac{\eta}{3}} \sigma_z, \end{aligned} \quad (3)$$

where  $\mathbb{1}$  and  $\sigma_{x,y,z}$  are the identity and the three Pauli matrices, respectively, and  $\eta$  is the probability that one of the errors occurs. This channel describes the symmetric decay of a single qubit into the center of the Bloch sphere. Note that when  $\eta = 0.75$ , the output of this channel is a fully mixed state. In the following, without causing confusion, the expression  $\mathcal{E}(\rho)$  on a multiqubit state  $\rho$  is also used to describe the action of the local noise channel on all qubits in  $\rho$ .

To mimic a quantum circuit with LPNC under noise, we first initialize the system in a state  $\rho_0 \in \mathcal{B}(\mathcal{H}_{\text{feas}})$ , where  $\mathcal{B}(\mathcal{H}_{\text{feas}})$  defines the space of bounded linear operators on  $\mathcal{H}_{\text{feas}}$ , and subsequently apply a set of one- and two-qubit unitaries drawn from  $G$ . After each layer of gates, or maximal set of consecutive gates which act on distinct sets of qubits, denoted by  $U_i$ , we apply the depolarizing channel  $\mathcal{E}$  to all qubits, as pictured in Fig. 1. We thus assume that all operations, including idle times of qubits, are subject to the same amount of noise.

To calculate the probability to escape the feasible subspace  $\mathcal{H}_{\text{feas}}$ , we define the projection operator onto the feasible subspace  $\mathcal{P}_{\text{feas}} = \mathcal{P}_N^{(1)} \otimes \mathcal{P}_N^{(2)} \otimes \dots \otimes \mathcal{P}_N^{(n)}$ , with  $\mathcal{P}_N^{(i)}$  the projector on the particle-number subspace with  $N$  particles on subsystem  $i$ . The expectation value  $p_{\text{feas}} = \langle \mathcal{P}_{\text{feas}} \rangle$  with respect to the output state of the quantum computation then is the probability of sampling a valid bit string.

By construction, gates from  $G$  map states with particle number  $N$  to states with the same particle number. Thus, in the absence of noise, the whole circuit preserves the particle number at any point of the computation, and there is no population transfer between different subspaces, i.e.,  $p_{\text{feas}} = 1$ . This changes in the presence of noise, where the expectation

value includes noise channels after each layer of the circuit,

$$\begin{aligned} p_{\text{feas}} &= \text{Tr}[\mathcal{P}_{\text{feas}} \mathcal{E}(U_d \mathcal{E}(\dots \mathcal{E}(U_1 \mathcal{E}(\rho_0) U_1^\dagger) \dots) U_d^\dagger)] \\ &= \text{Tr}[\rho_0 U_1^\dagger \mathcal{E}(\dots \mathcal{E}(U_d^\dagger \mathcal{E}(\mathcal{P}_{\text{feas}}) U_d) \dots) U_1], \end{aligned} \quad (4)$$

where we used that all Kraus operators of the depolarizing channel are Hermitian. Without loss of generality, we assume in the following that all subsystems have the same size and are affected by the same amount of noise, i.e., have equal value of  $\eta$  for all qubits. As a starting point, we exploit the locality of the defined noise channel to write

$$\begin{aligned} \mathcal{E}(\mathcal{P}_{\text{feas}}) &= \mathcal{E}(\mathcal{P}_N^{(1)} \otimes \mathcal{P}_N^{(2)} \otimes \dots \otimes \mathcal{P}_N^{(n)}) \\ &= \mathcal{E}(\mathcal{P}_N^{(1)}) \otimes \mathcal{E}(\mathcal{P}_N^{(2)}) \otimes \dots \otimes \mathcal{E}(\mathcal{P}_N^{(n)}). \end{aligned} \quad (5)$$

We note that the Kraus operator  $K_4$  [cf. Eq. (3)] does not change the population distribution of a qubit and thus commutes with all particle-number projectors and thus can be neglected. The Kraus operators  $K_2$  and  $K_3$  cause local bit flips and thus generate projection operators onto different particle numbers. As all states with a particle number  $j$  require the same number of spin flips to reach a state with particle number  $j'$ , we can write the equation above as a weighted sum of projection operators,

$$\mathcal{E}(\mathcal{P}_{\text{feas}}) = \sum_{j=0}^{\kappa} c_j \mathcal{P}_j^{(1)} \otimes \sum_{j=0}^{\kappa} c_j \mathcal{P}_j^{(2)} \otimes \dots \otimes \sum_{j=0}^{\kappa} c_j \mathcal{P}_j^{(n)}, \quad (6)$$

with  $c_j$  the weight of the projectors  $\mathcal{P}_j^{(i)} \forall i$ . The next step in our calculation is the application of the first layer of the circuit onto this sum of projection operators, i.e.,  $U_1^\dagger \mathcal{E}(\mathcal{P}_{\text{feas}}) U_1$ . As we designed the unitaries to respect the LPNC, we have  $U_1^\dagger \mathcal{E}(\mathcal{P}_{\text{feas}}) U_1 = \mathcal{E}(\mathcal{P}_{\text{feas}})$ , that is,  $\mathcal{E}(\mathcal{P}_{\text{feas}})$  stays invariant under the unitaries in  $G$ . This behavior remains the same for a cascade of  $k$  applications of noise channels and unitaries,

$$\begin{aligned} &U_1^\dagger \mathcal{E}(U_2^\dagger \mathcal{E}(\dots \mathcal{E}(U_d^\dagger \mathcal{E}(\mathcal{P}_{\text{feas}}^i) U_d) \dots) U_2) U_1 \\ &= \sum_{j=0}^{\kappa} c_j^{(k)} \mathcal{P}_j^{(1)} \otimes \sum_{j=0}^{\kappa} c_j^{(k)} \mathcal{P}_j^{(2)} \otimes \dots \otimes \sum_{j=0}^{\kappa} c_j^{(k)} \mathcal{P}_j^{(N)}. \end{aligned} \quad (7)$$

Thus, for  $k = d$  layers, the expectation value, given by Eq. (4), reads

$$\begin{aligned} \langle \mathcal{P}_{\text{feas}} \rangle &= \text{Tr}[(c_N^{(d)} \mathcal{P}_N^{(1)} \otimes c_N^{(d)} \mathcal{P}_N^{(2)} \otimes \dots \otimes c_N^{(d)} \mathcal{P}_N^{(N)}) \rho_0] \\ &= \prod_{i=1}^n c_N^{(d)} = (c_N^{(d)})^n, \end{aligned} \quad (8)$$

where we used that  $\rho_0 \in \mathcal{B}(\mathcal{H}_{\text{feas}})$  with particle number  $N$  and that states with other particle number are mutually orthogonal. We observe that this expression is independent of the unitaries  $U_i$  and that the exact coefficient  $c_N^{(d)}$  is determined solely by the noise level  $\eta$ , the circuit depth  $d$ , and the particle number  $N$ . The calculation of the coefficient  $c_N^{(d)}$  is a combinatorial task and can be reformulated as the simple question: starting with a classical bit string of length  $\kappa$  with  $N$  bits in 1 and  $\kappa - N$  in 0, how likely is it to find a bit string with  $N$  bits in 1 after  $d$  independent symmetric random processes on each bit? In this picture, each random process is the classical representation of the local quantum error channel and has certain probability to flip the bit. To answer this question, we first give

the probability that a single bit is found in its initial state after  $d$  random processes,

$$p_{0/1 \rightarrow 0/1}(d) = \sum_{k=0}^{\lfloor d/2 \rfloor} \binom{d}{d-2k} \tilde{\eta}^{d-2k} (1-\tilde{\eta})^{2k}. \quad (9)$$

In this expression,  $\tilde{\eta} = (1 - 2/3\eta)$  only depends on the error rates from Kraus operators  $K_2$  and  $K_3$ . With this probability, we count all possibilities to find a bit string with  $N$  bits showing up,

$$c_N^{(d)} = \sum_{j=0}^M \binom{N}{j} \binom{\kappa - N}{j} p_{0/1 \rightarrow 0/1}^{\kappa-2j} (1 - p_{0/1 \rightarrow 0/1})^{2j}, \quad (10)$$

where  $M = \min\{N, \kappa - N\}$ . For a system composed of  $n$  subsystems, the probability to stay in the feasible subspace,  $p_{\text{feas}} = (c_N^{(d)})^n$ , only depends on the number of subsystems  $n$ , the particle number  $N$ , the noise level  $\eta$ , and the circuit depth  $d$ . Evidently, if  $d > 1$  and  $\eta \neq 0$ ,  $c_N^{(d)} < 1$  and  $p_{\text{feas}}$  decays exponentially in the number of subsystems.

Figure 2 shows  $c_N^{(d)}$  as a function of the circuit depth  $d$  for  $N = 1$  and  $\kappa = 3$  for various noise levels. The horizontal line shows the expectation value with respect to a completely mixed state,  $\text{Tr}(\mathbb{1}/2^\kappa \mathcal{P}_N)$ . The probability of staying in the feasible subspace monotonically decreases with the noise level  $\eta$ , as shown in Fig. 2(b).

In the following section, we apply the above theory and study the behavior of  $p_{\text{feas}}$  for an example of applications in the NISQ era.

### III. QUANTIFYING THE ROBUSTNESS OF XY-QAOA UNDER DEPOLARIZING NOISE

In this section, we use our findings to analyze the robustness of the quantum alternating operator Ansatz, [13], in particular XY-QAOA [15], an adaption of the quantum approximate optimization algorithm [14], to solve classical optimization problems subject to hard constraints, under the influence of noise.

#### A. Review of the quantum alternating operator Ansatz (QAOA)

The quantum approximate optimization algorithm [14], and the extension of its framework to the general class of quantum alternating operator Ansätze, are leading quantum metaheuristics for exact or approximate solutions of classical optimization problems. Low-depth versions of these algorithms are suitable for NISQ devices. While, in some cases, good parameters can be found analytically [16] or through decomposition techniques that make use of light or causal cones to obtain analytical formulas for the parameters that can be optimized numerically [14,17–20], variational approaches remain popular.

Early experiments on real quantum devices have been realized [21,22]. Despite all this progress, only little is known about its performance in comparison to classical algorithms or its robustness against noise [19,23–28].

In QAOA, the first step is to map the cost function to a classical spin Hamiltonian, commonly referred to as problem Hamiltonian  $H_P$ . This mapping has to be chosen such that

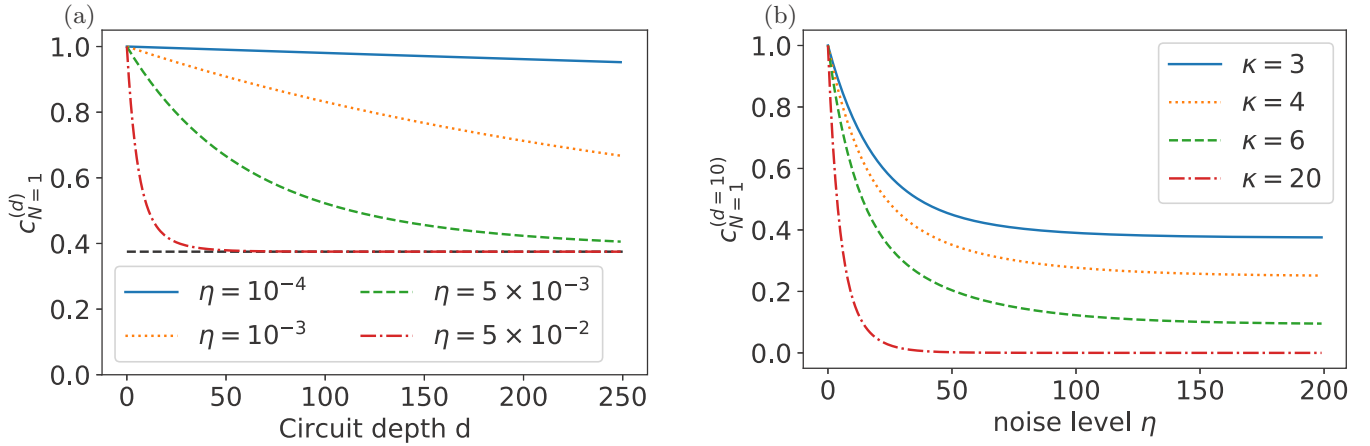


FIG. 2. (a) The probability of staying in the  $N = 1$  subspace of  $\kappa = 3$  qubits for various noise levels as a function of the circuit depth. The dashed line shows the probability with respect to a completely mixed state,  $\rho_{\text{mixed}} = \mathbb{1}/2^3$ . (b) The monotonically decreasing probability of staying in the  $N = 1$  subspace as a function of the level of noise  $\eta$ , shown for four different numbers of colors  $\kappa$  for a circuit of depth  $d = 10$ .

the eigenstates of the Hamiltonian and their corresponding energies are in direct relationship to the solutions and their classical costs of the classical problem, respectively.

The problem Hamiltonian generates, together with a second Hamiltonian, often  $H_X = \sum_i \sigma_x^{(i)}$ , the mixing and phase-separation unitaries,  $U_M(\beta_i) = \exp[-i\beta_i H_M]$  and  $U_{PS}(\gamma_i) = \exp[-i\gamma_i H_P]$ , respectively, which define the QAOA state,

$$|\Psi(\{\beta_i, \gamma_i\})\rangle = \prod_i^p U_M(\beta_i) U_P(\gamma_i) |+\rangle^{\otimes n}, \quad (11)$$

with  $p$  the number of iterations of both unitaries, or QAOA blocks, and the initial state,  $|+\rangle^{\otimes n} = \bigotimes_i^n 1/\sqrt{2}(|0\rangle_i + |1\rangle_i)$ , the superposition of all computational states. The variational parameters  $\{\beta_i, \gamma_i\}$  are then optimized, aiming to minimize the expected energy of the problem Hamiltonian,

$$E_g = \min_{\{\beta_i, \gamma_i\}} \langle \Psi(\{\beta_i, \gamma_i\}) | H_P | \Psi(\{\beta_i, \gamma_i\}) \rangle. \quad (12)$$

Most optimization problems are, however, subject to constraints, which shrink the state space of valid solutions. A common approach to include constraints into quantum algorithms, e.g., quantum annealing or QAOA [29], is to penalize all states not fulfilling the constraints by adding an energy term to the problem Hamiltonian. In quantum annealing, utilizing the driver Hamiltonian for constrained optimization was suggested [30–32].

In this setup, however, the algorithm still accesses the whole state space, including nonvalid solutions.

In [13, 15], an alternative approach was introduced. Instead of penalizing the constraint-violating states in the cost Hamiltonian, the mixing Hamiltonian  $H_M$  in the QAOA state was adapted to respect the underlying symmetry of the constraints and thereby restricts the algorithm's evolution to valid states only. Generally, the feasible subspace, i.e., the subspace of valid states, while still exponentially large, is exponentially smaller than the full search space. For example, to preserve

the particle number on qubits in a subset  $S$ , a possible choice of the mixing Hamiltonian is

$$H_{XY} = \sum_{i,j \in S} H_{XY}^{i,j}, \quad (13)$$

$$H_{XY}^{i,j} = \frac{1}{2} (\sigma_x^{(i)} \sigma_x^{(j)} + \sigma_y^{(i)} \sigma_y^{(j)}). \quad (14)$$

The unitary generated by this mixing Hamiltonian,  $U_M(\beta) \rightarrow U_{XY}(\beta) = \exp[-i\beta H_{XY}]$ , then drives only transitions between states with the same particle number. A direct advantage is that the computation is restricted to valid states only, which significantly shrinks the state space that the algorithm has access to. In [15], it was numerically shown that this approach outperforms the approach with energy penalties.

However, in the presence of noise, leaving the feasible subspace could have drastic consequences for the performance of the algorithm.

## B. Quantifying robustness under depolarizing noise

Here we use the example from [15] and analyze the robustness of XY-QAOA applied to the max- $\kappa$ -colorable-subgraph problem of a graph  $G = (V, E)$ , with vertices  $V$  and edges  $E$ , in the presence of noise. For simplicity of analysis, we consider 3-regular graphs  $G$ . This problem asks to maximize the number of edges in a correctly vertex-colored subgraph. For a precise definition of the problem, we refer to [15]. For  $\kappa = 2$  colors, the internal qubit states  $|0\rangle$  and  $|1\rangle$  can be used to represent the different colors. For more colors, however, more than a single qubit is required. Thus, a logical vertex  $v \in V$  of the graph has to be encoded nonlocally in several qubits of the qubit register. A common encoding strategy is the one-hot encoding [33], where each color is represented by one of the particle-number  $N = 1$  states of  $\kappa$  qubits. For the example of three colors, the three states  $|100\rangle$ ,  $|010\rangle$ , and  $|001\rangle$  could encode the three different colors. The classical cost function encoded in the space spanned by the above states counts the number of edges that connects two vertices of the



same color,

$$f_C = \sum_j \sum_{(v,v') \in E} x_{v,j} x_{v',j}, \quad (15)$$

with  $m = |E|$  the number of edges in the graph  $G$ . The corresponding problem Hamiltonian can be formulated by promoting the binary variables to spin operators,  $x_{a,b} \rightarrow (\mathbb{1} - \sigma_z^{(a,b)})/2$ ,

$$H_P = \frac{1}{4} \sum_{j=1}^{\kappa} \sum_{(v,v') \in E} \sigma_z^{(v,j)} \sigma_z^{(v',j)}. \quad (16)$$

Here, we used that the sum over all 1-local terms,  $\sum_i \sigma_z^{(v,i)}$ , is proportional to the particle-number operator  $N_v$  and hence constant in the feasible subspace.

The mixing Hamiltonian given by Eq. (14) is used to drive transitions between all states with a single particle of each logical vertex. By construction, both the mixing unitaries  $U_{XY}(\beta_i)$  and the phase-separation unitaries  $U_P(\gamma_i)$  preserve the particle number of each vertex subsystem and satisfy the requirements that we made for the unitaries in  $G$ ; cf. Sec. II. As this is true for any choice of variational parameters  $\{\gamma_i, \beta_i\}$ , the probability of staying in the feasible subspace does not depend on the value of the parameters. Thus, we ignore the variational character of QAOA for the following analyses.

We, moreover, note that the approximation ratio, i.e., the normalized ratio between the expectation value with respect to  $H_P$  and the ground-state energy, and the ground-state population are trivially upper bounded by the probability of staying in the feasible subspace, as all nonfeasible states have ground-state probability and approximation ratio equal to zero. While the actual performance measured by the approximation ratio or ground-state probability relies heavily on the parameter updating techniques, we focus on the symmetry aspect in the operators, which is parameter agnostic. In the rest of the paper, we evaluate only the probability of an output state to be in the feasible subspace and will refer this quantity as the robustness, unless otherwise stated.

To apply theories developed in Sec. II and use Eq. (8), we have to quantify the circuit depth required to implement the mixing and problem unitaries,  $U_M$  and  $U_P$ . The circuit depth, however, heavily depends on the topology of the hardware graph. As different technologies, such as ion traps or superconducting qubits, differ in their topology design in the NISQ era, we study two different qubit connectivity scenarios in the following sections: (1) qubits are fully connected and (2) qubits form a 2D grid. Moreover, we assume that all gates that we define in our further analysis are available as native gates and do not require further compilation into different gate sets, which would further increase the circuit depth.

### 1. Quantifying robustness on fully connected hardware

On fully connected quantum computers, we can carry out all interactions without any overhead of routing the qubits on the hardware. The mixing Hamiltonian consists of interactions between all pairs of qubits representing a logical vertex. The corresponding unitary can be approximated by any ordered product of two-qubit unitaries on these pairs,  $\exp[-i\beta H_{XY}^{i,j}]$  [15].

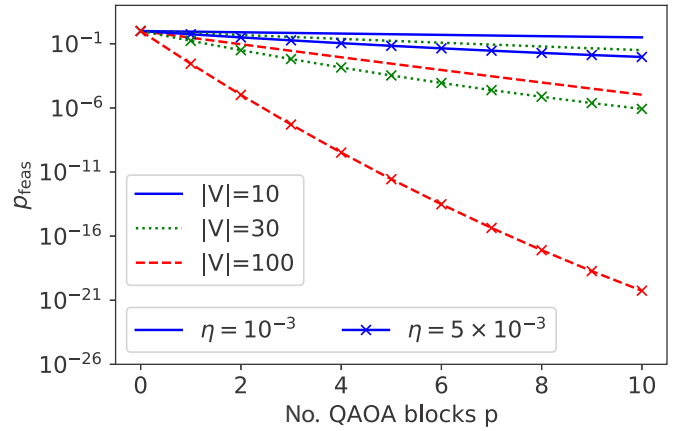


FIG. 3. The probability to stay in the feasible subspace for XY-QAOA applied to the max- $\kappa$ -colorable-subgraph problem with one-hot encoding and  $\kappa = 3$  colors for a 3-regular graph with  $|V| = \{10, 30, 100\}$  logical vertices as a function of the number of QAOA blocks  $p$ , assuming qubits are fully connected.

For  $\kappa$  colors, the unitary  $U_{XY}(\beta)$  can be carried out with circuit depth  $d_{XY} = \kappa - 1$  if  $\kappa$  is even and  $d_{XY} = \kappa$  if  $\kappa$  is odd [34]. The problem unitaries, generated by the Hamiltonian in Eq. (16), include two-qubit unitaries between all  $\kappa$  qubits representing the same color for each connection  $e \in E$  in the logical graph. Thus, we have to carry out a single two-qubit unitary between all color pairs for each connection. On a fully connected quantum computer, this can be parallelized, resulting in circuit depth  $d_P = (k + 1)$  for a  $k$ -regular graph if  $k$  is odd and  $d_P = k$  if  $k$  is even [34]. Thus, in total, we can realize the circuit in depth  $d_{QAOA} = p(d_P + d_{XY})$  for  $p$  QAOA blocks on a fully connected quantum computer. As in the previous section, we apply noise channels after each layer in the circuit and calculate the probability of staying in the feasible subspace with Eqs. (10) and (8). In Fig. 3, we plot the probability of staying in the feasible subspace for  $\kappa = 3$  colors and  $k = 3$ -regular graphs with  $|V| = \{10, 30, 100\}$  vertices for various number of QAOA blocks and noise levels.

### 2. Quantifying robustness on 2D grids

Most available quantum hardware, however, has limited connectivity, thus routing of the quantum circuit becomes necessary. This unavoidably will increase the circuit depth, which in turn increases the amount of noise in the system. To benchmark this effect, here we assume the hardware graph resembling a 2D grid, typical for the prevalent superconducting qubit devices [8]. If the quantum circuit graph cannot be mapped on the hardware graph, SWAP gates have to be used to route the qubits such that all necessary gates can be carried out. With a SWAP network of depth equal to the number of qubits [35], it would be possible to allow for all-to-all connections required only for a linearly connected hardware graph. However, in the NISQ era where the circuit depth is a valuable and expensive asset, for circuit graphs with sparse connectivity, it is possible to find SWAP networks with smaller depth [36].

As compiling strongly depends on the problem instance, we generate 100 random 3-regular graphs with  $|V| = 30$

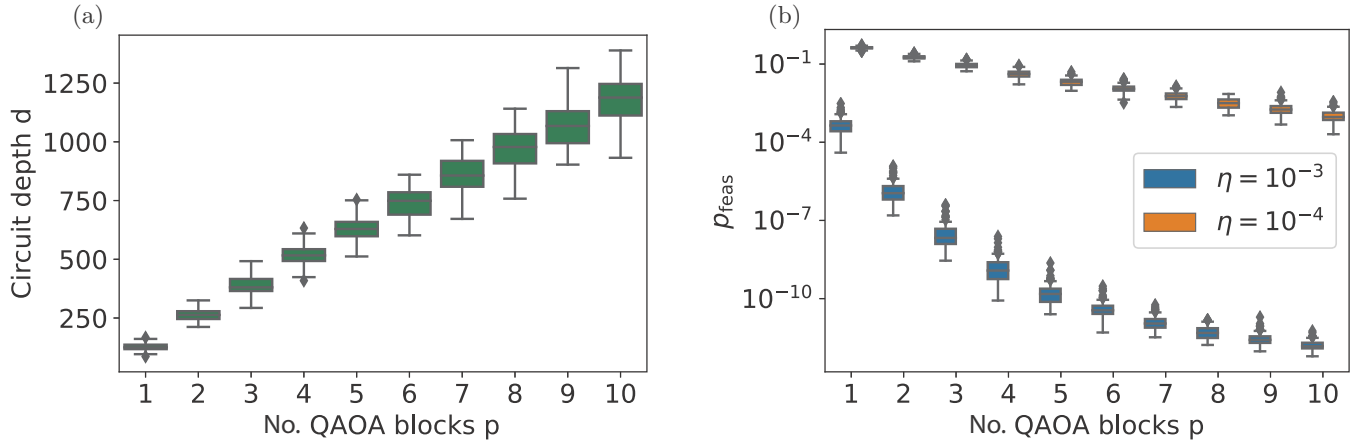


FIG. 4. Results for the max-3-colorable-subgraph problem with one-hot encoding for 3-regular graphs with  $|V| = 30$  logical vertices. (a) The circuit depth of the  $XY$ -QAOA circuit for various number of blocks  $p$  when routing the circuit onto a 2D grid of size  $10 \times 10$ . The variance in the data shows how the circuit depths differ for different instances. (b) The probability of staying in the feasible subspace for two different noise levels plotted against the number of QAOA blocks when taking into account the circuit depth from (a).

vertices, which require 90 physical qubits to be implemented. As the hardware graph, we assume a square grid of size  $10 \times 10$ . To route the quantum circuit, we use `t|ket` [36,37] and calculate the circuit depth of the routed circuit for various numbers of QAOA blocks, shown in Fig. 4(a). We use the found depth to calculate the probability of staying in the feasible subspace for the noise levels  $\eta = \{10^{-3}, 10^{-4}\}$ , resulting in Fig. 4(b). In comparison to Fig. 3, we see that the routing drastically worsens the results.

For both hardware assumptions, we see that already for such problems of rather moderate size, for realistic error rates  $10^{-2}$ – $10^{-3}$  [8,38,39], the probabilities are below reasonable experimental sampling requirements of  $10^4$ – $10^6$  shots [8]. Keeping in mind that QAOA is a hybrid algorithm, independently from the hardware topology many samples will be needed to estimate the cost function up to a precision required for making reasonable parameter updates [40].

### 3. Comparison to QAOA with energy penalties

To stay in the feasible subspace,  $XY$ -QAOA uses the mixer in Eq. (14). It was proposed as an alternative to  $X$ -QAOA that is expected to perform better under no noise, thanks to the design which allows  $XY$ -QAOA to restrict exploration to the subspace of valid solutions of the full Hilbert space [15]. The alternative,  $X$ -QAOA, in analogy with a technique commonly used in quantum annealing, could use penalty terms to encourage the evolution to remain in the feasible subspace while using the standard transverse-field mixing operator of  $X$ -QAOA,  $H_X = \sum_v \sum_j \sigma_x^{(v,j)}$ . For the max- $\kappa$ -colorable subgraph, the following could serve as a penalty term:

$$H_{\text{pen}} = \alpha \frac{1}{2} \sum_{v \in E} \left( (2 - \kappa) \sum_j \sigma_z^{(v,j)} + \sum_{j,j'} \sigma_z^{(v,j)} \sigma_z^{(v,j')} \right). \quad (17)$$

At first glance, it seems possible that the transverse-field mixer could improve the probability of remaining in the feasible subspace under noise because, even though it takes elements in

the feasible subspace to the nonfeasible subspace, it can also take elements from the nonfeasible subspace to the feasible subspace, which the  $XY$  mixer cannot. On the other hand, the transverse field behaves like collective bit-flip noise on all qubits, which, in principle, could be corrected, but because of its simultaneous action on all qubits with the same  $\beta$  angle, it is very unlikely to rotate all qubits back to the feasible subspace. Hence the bit-flip analysis of Sec. II, particularly Eq. (10) and Fig. 2(b) showing that the probability of staying in the feasible subspace decreases monotonically with noise, suggests that the transverse-field mixer may cause more harm than good even in the presence of noise. We now look at the relative effect of noise on  $XY$ -QAOA and  $X$ -QAOA.

Here, we give two complementary arguments for why we would expect that the advantage of  $XY$ -QAOA over  $X$ -QAOA still holds in the presence of noise. In the first section (a), we show that the action of the  $X$  mixer does not lead to an increase of the probability to stay in the feasible subspace, at least under depolarizing noise, while in the second section (b), we show that implementing the penalty term in the  $X$ -QAOA approach requires a higher circuit depth than  $XY$ -QAOA.

*a. Effect of noise.* Let us first compare how different mixers are affected by a local depolarizing channel applied to the feasible subspace projector  $\mathcal{P}_{\text{feas}}$ . We select a projector to the entire feasible subspace, instead of a projector to a particular state, because we aim to find an “average” behavior and not restrict the analysis to a special algorithm initialization, since the feasible subspace projector is a linear combination of product-state projectors (e.g., for a three-qubit system  $\mathcal{P}_{\text{feas}} = P_{|001\rangle} + P_{|010\rangle} + P_{|100\rangle}$ , where  $P_{|\psi\rangle} = |\psi\rangle\langle\psi|$ ) and as such is invariant under the action of the  $XY$  mixer. Therefore, the action of any local noise model will preserve the linear combination product structure. Local depolarizing noise will additionally preserve the diagonal structure in the computational basis (it suffices to consider the action of the noise channel on projectors  $|0\rangle\langle 0|$  and  $|1\rangle\langle 1|$  which leads to  $\rho = \frac{1}{2}\mathbb{1} \pm \frac{\mu}{2}Z$ , respectively, with  $\mu = 1 - \frac{4\eta}{3}$  and  $\eta$  being the depolarizing rate). This behavior means that any probability loss can only be associated with the action of the depolarizing

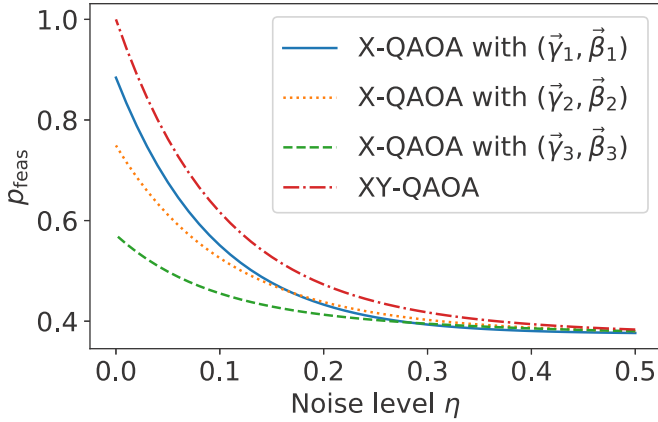


FIG. 5. The probability to stay in the  $N = 1$  subspace of  $\kappa = 3$  qubits for four layers of  $XY$ -QAOA and  $X$ -QAOA. For  $XY$ -QAOA (red), the choice of the parameters does not affect the probability, while for  $X$ -QAOA, the following parameter combinations were used:  $\vec{\beta}_1 = (1.5, 0.2, 0.9, 3.8)$  and  $\vec{\gamma}_1 = (1.5, 3.2, 0.9)$  (blue),  $\vec{\beta}_2 = (1.5, 1.7, 1.9, 0.8)$  and  $\vec{\gamma}_2 = (0.4, 1.2, 1.7)$  (orange),  $\vec{\beta}_3 = (2.2, 3.7, 2.9, 2.8)$  and  $\vec{\gamma}_3 = (1.4, 0.9, 3.1)$  (green). Parameters are given in the form  $\vec{\beta} = (\beta_1, \beta_2, \beta_3, \beta_4)$  and  $\vec{\gamma} = (\gamma_2, \gamma_3, \gamma_4)$ . The angles were selected such that the sum of all  $\beta$  is close to a multiple of  $2\pi$ ; if selected differently (e.g., the sum close to the odd multiples of  $\pi$ ), one would observe even worse performance. Since the first application of the penalty Hamiltonian [see Eq. (17)] does not affect the probability,  $\gamma_1$  is omitted. Our investigation suggests that the selected angles represent the general behavior that is in accordance with Eq. (18).

channel. However, for a standard  $X$  mixer, we additionally have the dependence on the  $\beta$  angle, which can lower the probability of staying in feasible subspace. The  $XY$  mixer provides an upper bound for the  $X$  mixer’s probability of staying in the feasible subspace for level-1 QAOA,

$$\text{Tr}[P_{\text{feas}} \mathcal{E}(U_X P_{\text{feas}} U_X^\dagger)] \leq \text{Tr}[P_{\text{feas}} \mathcal{E}(U_{XY} P_{\text{feas}} U_{XY}^\dagger)], \quad (18)$$

where  $U_X \equiv U_X(\beta_x) = \exp[-i\beta_x \sum_i \sigma_x^{(i)}]$ , and  $U_{XY} \equiv U_{XY}(\beta_{xy}) = \exp[-i\beta_{xy} H_{XY}]$ , with  $H_{XY}$  given in Eq. (14), and the inequality holds for an arbitrary pair of  $\beta_x$  and  $\beta_{xy}$ .

Any rotation out of computational basis plane will decrease the probability of staying in the feasible subspace. Equation (18) technically only holds for one level of  $X$  and  $XY$  mixers. However, the numerical results of four layers of mixers shown in Fig. 5 strengthen our argument that this bound also holds for repeated action of the mixers. In the numerical analysis, we selected a set of angles that is representative of the overall performance, and therefore we expect similar behavior for any other set of angles.

*b. Circuit depth.* Furthermore, both approaches require the implementation of the phase-separation unitary  $U_p$ . The unitary generated by the penalty Hamiltonian [see Eq. (17)] requires two-qubit gates between all pairs of qubits representing a logical vertex. The same holds for the action of the mixing unitary; see Sec. III B 1. Thus, on a fully connected quantum computer, the mixing unitary of  $XY$ -QAOA requires the same depth as the implementation of the penalty term of  $X$ -QAOA. However,  $X$ -QAOA requires an additional layer of local  $X$  rotation yielding a higher total circuit depth for

$X$ -QAOA. The depth on specific hardware will depend on which two-qubit gate or gates are natively implemented.

Moreover, classical methods to postselect the measurement outcomes could be useful in the NISQ. How such methods influence the performance of  $X$ -QAOA or  $XY$ -QAOA is an open question and not discussed here.

#### IV. MITIGATING THE EFFECT OF NOISE

In the previous section, we have used states with a single particle of  $\kappa$  qubits to encode  $\kappa$  colors and found that the effect of noise makes it intractable to push forward to problems with hundreds of variables with the noise levels that are present in currently available quantum hardware. To make  $XY$ -QAOA ready for applications in the NISQ era, one either has to substantially lower the noise levels, find noise-resilient ways to encode the problems, or develop error-mitigation techniques. In the following, we discuss the two latter options.

In Sec. IV A, we consider encoding the feasible states into a subspace corresponding to a larger particle number, which typically is a larger subspace than the  $N = 1$  subspace, and examine how this encoding affects the probability of staying in the feasible subspace. In Sec. IV B, we introduce a bit-flip quantum error correction code adapted to specifically preserve a particle-number preserving subspace and evaluate its efficiency. This code demonstrates the potential gains in resource efficiency through symmetry-aware quantum error correction techniques.

##### A. Encoding in larger particle-number subspaces

Utilizing a larger fraction of the full Hilbert space as feasible subspace increases the number of valid solutions over unwanted solutions. On the other hand, encoding in a different subspace will also change the expression of the cost function, incurring a different circuit depth and therewith changing the amount of noise in the computation.

In this section, we explore if such an approach can be helpful to uphold the probability of sampling valid solutions. We again focus on the max- $\kappa$ -colorable-subgraph problem. The cost function given in Eq. (15) can be easily adapted to other particle-number subspaces by including higher-order interactions. The particle-number  $N$  states of  $n_q$  qubits can be used to implement the max- $\binom{n_q}{N}$ -colorable-subgraph problem with cost function

$$f_C^N = \sum_{(v,v') \in E} \frac{1}{N!} \sum_{\substack{i_1, \dots, i_N \\ i_1 \neq \dots \neq i_N}}^{n_q} x_{v,i_1} x_{v',i_1} \dots x_{v,i_N} x_{v',i_N}, \quad (19)$$

which includes up to  $2N$ -local terms.

As an example, for  $\kappa = 6$  colors, we could, as before, use the six states with a single particle of six qubits. Alternatively, the six states with two particles of four qubits could be used. While the first approach uses a fraction of  $6/2^6 \approx 9\%$  of the respective full Hilbert space, the latter approach uses  $6/2^4 \approx 38\%$  of all states. To see whether this gain in fraction also yields a gain in robustness against noise, we have to calculate the circuit depth that is required to implement the different cost function.

The cost function with  $N = 2$  particles reads

$$f_C^{N=2} = \sum_{(v,v') \in E} \frac{1}{2} \sum_{\substack{i,j \\ i \neq j}}^{n_q=4} x_{v,i} x_{v,j} x_{v',i} x_{v',j}, \quad (20)$$

and includes a product of four binary variables. To construct the corresponding problem Hamiltonian, we again replace the binary variables  $x_{a,b}$  with Pauli operators, which results in a Hamiltonian  $H_P^{N=2}$  with up to 4-local terms. We recognize that not all terms in the Hamiltonian are needed to represent the problem faithfully. For example, as in Sec. III B, the sum over the 1-local terms on each vertex is proportional to the particle-number operator  $N_v$  of that vertex in the feasible subspace,

$$\tilde{N}_v = \sum_i \sigma_z^{(v,i)} = 2N_v - 4 = 0, \quad (21)$$

and thus is not needed to encode the problem faithfully. To make the circuit as short as possible, we next show how to reduce the number of terms in the Hamiltonian when restricted to the feasible subspace. For example, the 2-local terms of the form  $\sum \sigma_z^{(v,i)} \sigma_z^{(v',j)}$  can be simplified in the following way:

$$\begin{aligned} \sum_{\substack{i,j \\ i \neq j}} \sigma_z^{(v,i)} \sigma_z^{(v',j)} &= \sum_i \sigma_z^{(v,i)} \sum_j \sigma_z^{(v',j)} - \sum_i \sigma_z^{(v,i)} \sigma_z^{(v',i)} \\ &= \tilde{N}_v \tilde{N}_{v'} - \sum_i \sigma_z^{(v,i)} \sigma_z^{(v',i)} \\ &= - \sum_i \sigma_z^{(v,i)} \sigma_z^{(v',i)}. \end{aligned} \quad (22)$$

We thus only have to consider the second term, which is a sum of 2-local interactions between the same qubits of vertex  $v$  and  $v'$ . For the other 2-local terms, the calculation is analogous. Similarly, we can show that all 3-local terms are vanishing,

$$\begin{aligned} \sum_{\substack{i,j \\ i \neq j}} \sigma_z^{(v,j)} \sigma_z^{(v,j)} \sigma_z^{(v',i)} \\ &= \sum_i \sigma_z^{(v,j)} \sum_i \sigma_z^{(v,i)} \sigma_z^{(v',i)} - \sum_i \sigma_z^{(v,i)} \sigma_z^{(v,i)} \sigma_z^{(v',i)} \\ &= \tilde{N}_v \sum_i \sigma_z^{(v,i)} \sigma_z^{(v',i)} - \sum_i \sigma_z^{(v',i)} \\ &= \tilde{N}_v \sum_i \sigma_z^{(v,i)} \sigma_z^{(v',i)} - \tilde{N}_{v'} = 0. \end{aligned} \quad (23)$$

For the 4-local terms, we cannot easily extract the particle-number operator, but when restricted to the feasible subspace, we recognize that only four terms are necessary to represent the sum distinctly. A careful collection of all the remaining terms yields the Hamiltonian

$$\begin{aligned} H_P^{N=2} &= \sum_{(v,v') \in E} H_{(v,v')} \quad \text{with} \\ H_{(v,v')} &= \frac{1}{4} \sum_{j=1}^4 \sigma_z^{(v,j)} \sigma_z^{(v',j)} + \frac{1}{8} (\sigma_z^{(v,1)} \sigma_z^{(v,2)} \sigma_z^{(v',1)} \sigma_z^{(v',2)} \\ &\quad + \sigma_z^{(v,3)} \sigma_z^{(v,4)} \sigma_z^{(v',3)} \sigma_z^{(v',4)} + 2\sigma_z^{(v,1)} \sigma_z^{(v,4)} \sigma_z^{(v',1)} \\ &\quad \times \sigma_z^{(v',4)} + 2\sigma_z^{(v,2)} \sigma_z^{(v,4)} \sigma_z^{(v',2)} \sigma_z^{(v',4)}), \end{aligned} \quad (24)$$

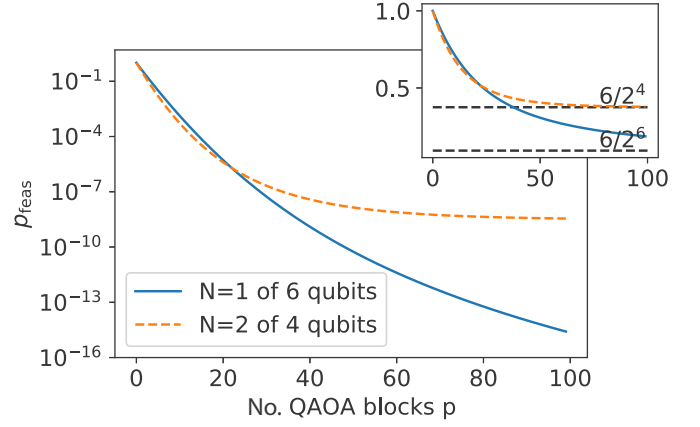


FIG. 6. Comparison of the probability to stay in the feasible subspace of  $N = 1$  states of six qubits and  $N = 2$  states of four qubits for a system size of  $|V| = 20$  (main plot) and  $|V| = 1$  (inset) logical vertices with a noise level of  $\eta = 10^{-3}$ .

where the last expression for four-qubit interactions is invariant under the qubit permutation for each node. For each connection  $(v, v') \in E$  in the problem graph  $G$ , we thus have four two-qubit and four four-qubit interactions. The two-qubit interactions can be executed in one layer. The four-qubit interactions [cf. Eq. (24)] can be executed in three layers, assuming that we can carry out any two- or four-qubit gates as a single gate, i.e., without decomposing them into native one- and two-qubit gates. For a  $k$ -regular graph, the problem Hamiltonian then requires depth  $d_P^{N=2} = 4(k+1)$  to be implemented. Independent from the cost function, the same version of the mixing Hamiltonian [cf. Eq. (14)] can be used as it preserves any particle number and, in analogy to before, results in a circuit of depth  $d_{XY}^{N=2} = 4$ .

In Fig. 6, we show the probability of staying in the feasible subspaces for both encodings as a function of the number of QAOA blocks. As clearly visible in the plot, in the regime of interest, i.e., where  $p_{\text{feas}}$  is fairly high, the overhead of implementing the more complex problem Hamiltonian obliterates any possible advantage of the larger subspace. For very deep circuits, the higher particle number shows an advantage. However, at this level, due to the low value of  $p_{\text{feas}}$ , we would mostly see nonvalid samples, which makes this theoretical advantage futile for practical applications. We note again that these numbers would look even worse in a real experiment, as we would have to account for decomposing the gates into a native gate set and for the routing of the circuit to a restricted hardware graph. As the encoding into the higher particle-number subspace results in a higher gate count with more complex gates, taking these effects into account would make the difference in both encodings even more visible. However, this first result highlights that the way of encoding the problem has a significant influence on the robustness of the quantum algorithm. Whether schemes with higher-order terms pose an alternative to the one-hot encoding also depends



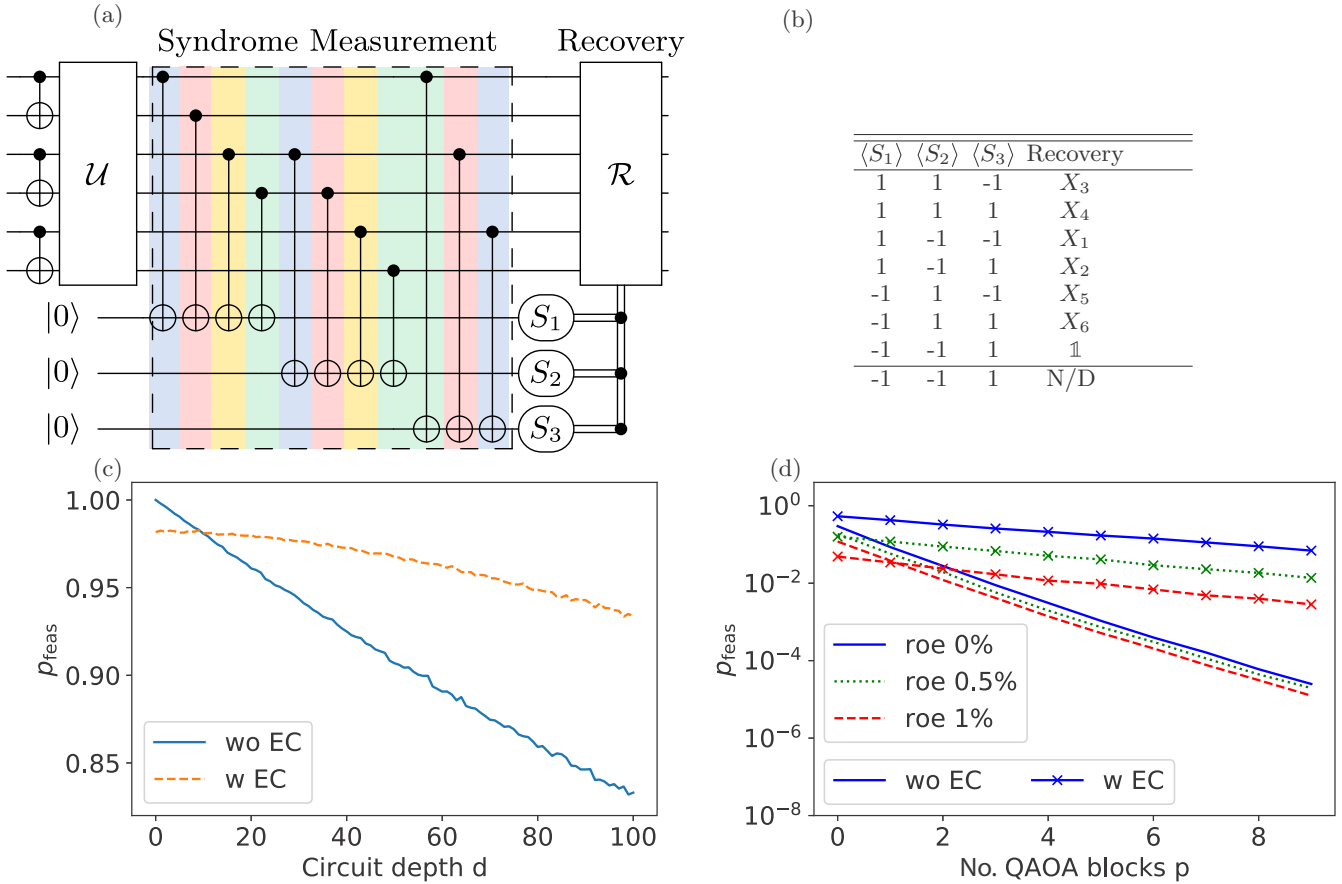


FIG. 7. (a) Error correction circuit correcting for a single bit-flip error on the data qubits with the syndromes given in Eq. (25). The colors highlight the gates which can be carried out in one layer on a fully connected quantum computer. According to the syndrome measurements, one of the recovery actions given in (b) is carried out where  $X_i$  defines the local bit-flip operation  $\sigma_x^{(i)}$  on qubit  $i$ . (c) The probability to sample a valid state for a noise level of  $\eta = 10^{-3}$  when applying the error correction circuit after  $d$  layers at the end of the computation—in comparison to no error correction. For this plot, we simulated the circuit execution and sampled  $10^5$  times from the output state for each circuit depth. (d) Comparison of the error correction scheme for XY-QAOA applied to the max-3-colorable-subgraph problem of a 3-regular graph with  $|V| = 30$  vertices. For this plot, a noise level of  $\eta = 10^{-3}$  and readout errors (roe)  $\{0\%, 0.5\%, 1\%\}$  were used. As in (c),  $10^5$  samples were used. The fluctuations in the data in (c) and (d) stem from the fact that a finite number of measurements was used.

on the availability of complex many-qubit gates in the next generations of NISQ devices.

**B. Correction of symmetry-breaking errors**

In the previous section, we analyzed the probability of sampling valid solutions when encoding the max- $\kappa$ -colorable-subgraph problem in a higher particle-number subspace than the  $N = 1$  particle-number subspace. Due to the overhead in the circuit depth when implementing the different cost function, we did not encounter an advantage over the common one-hot encoding. In this section, we discuss a possibility to correct the symmetry-breaking errors for the example of encoding  $\kappa = 3$  colors into the  $N = 1$  subspace. We design an adaption of the bit-flip code [41,42] that requires fewer ancillary qubits by exploiting the system’s particle-number symmetry, as shown in Fig. 7(a), and analyze its effect. The reduction in resources is possible because we aim only to correct states in the  $N = 1$  particle-number subspace.

In the original bit-flip code, each logical qubit is encoded in three physical qubits, and bit-flip errors are identified through

a majority vote on the three physical qubits. The present adaption benefits from the additional information on the LPNC of the data qubits, which reduces the number of physical qubits from three to two. The syndromes

$$\begin{aligned}
 S_1 &= Z_1 Z_2 Z_3 Z_4 \mathbb{1}_5 \mathbb{1}_6, \\
 S_2 &= \mathbb{1}_1 \mathbb{1}_2 Z_3 Z_4 Z_5 Z_6, \\
 S_3 &= Z_1 \mathbb{1}_2 Z_3 \mathbb{1}_4 Z_5 \mathbb{1}_6,
 \end{aligned}
 \tag{25}$$

with the subscripts  $\{1, 2, 3, 4, 5, 6\}$  labeling the top six qubits shown in Fig. 7(a), are able to detect a single bit-flip error on the data qubits. To read out the syndromes noninvasively, i.e., without destroying the quantum state, three additional ancilla qubits are needed. As in all quantum error correction schemes, measurements of the ancillas do not provide us with any additional information about the data qubits’ quantum state. After measuring the ancillas, the recovery actions, summarized in Fig. 7(b), correct the error by applying a local bit-flip operation. The effectiveness of this error correction procedure depends on the probability of errors in the input state, the noise in the correction circuit itself, and the readout error of

the ancillas. We note that this kind of error correction scheme does not correct for errors inside the feasible subspace.

To discuss how this scheme performs under realistic noise levels, we have to decide how often we want to apply it during the quantum computation. Under the assumption that the error correction circuit is perfect and noiseless, we could apply it after each layer of the circuit and would always see an improvement of the results. However, in realistic scenarios, the error correction circuit itself is subject to noise. To analyze how often we can apply the scheme without worsening the results, we look at the following setup: we initialize the system in a feasible state, subsequently apply  $d$  layers of QAOA with noise, and afterwards apply the syndrome circuit with the ancilla qubits initialized in the  $|0\rangle$  state. We use the same noise model as before and apply noise channels on all nine qubits after each layer. On a fully connected quantum computer, the correction circuit can be carried out in depth 4, highlighted by the colors in Fig. 7(a). Based on the measurement outcomes of the ancillas, we then apply the recovery action on the data qubits and sample from the final state. In Fig. 7(c), we show the probability to sample a feasible state in dependence of the circuit depth  $d$  prior to the onset of correction, in comparison with the situation of no error correction involving only three qubits. As expected, if the input state has no or little noise, the correction scheme even worsens the result. However, for noisier input data, we see a clear advantage in applying the error correction code.

In Sec. III B 1, we applied XY-QAOA to the max-3-colorable-subgraph problem of a 3-regular graph and saw that the results are quickly beyond reasonable sampling assumptions. Here we analyze how these results change when we apply the error correction scheme after every three QAOA blocks (i.e., 21 layers; cf. Sec. III B 1) of the XY-QAOA circuit. To make the situation even more realistic to current hardware prototypes, we include readout errors (roe) on all qubits by defining the symmetric probabilities  $P(0|1) = P(1|0) = \text{roe}$  that a qubit's state is read out falsely. For current technologies, readout errors are between 0.1% and 5% [8,43]. Such errors lead to incorrect recovery actions, tempering the effectiveness of the correction. In Fig. 7(d), we show the probability to find a feasible state for a graph with  $|V| = 30$  vertices in dependence of the number of QAOA blocks for the situations with and without error correction, a noise level of  $\eta = 10^{-3}$ , and readout errors  $\{0\%, 0.5\%, 1\%\}$ . As clearly visible, beyond the crossover, error correction improves the results by several orders of magnitude. As expected, the readout errors have a stronger effect on the computation when error correction is applied. However, we still see a clear advantage in applying this simple error correction scheme for all studied readout error strengths.

In this example, we have assumed a fully connected quantum computer. On devices with limited qubit connectivity, the overhead of routing the error correction has to be taken into account. Moreover, the recovery operation requires precise measurements and fast classical control loops, which, while challenging on certain hardware prototypes where measurement takes a long time, on some other hardware prototypes is expected to be readily achievable [44].

## V. CONCLUSION AND OUTLOOK

In this paper, we have studied the robustness of quantum algorithms with local particle-number conservation under the influence of noise. We found an exact combinatorial expression to calculate the probability of staying in particle-number subspaces under the influence of local depolarizing noise. With these findings, we benchmarked the robustness of XY-QAOA applied to the max- $\kappa$ -colorable-subgraph problem and analyzed the influence of the choice of the problem encoding on the robustness of the algorithm. We, moreover, discussed the possibility of correcting for symmetry-breaking errors. Our results highlight the importance of finding clever encoding techniques or error-mitigation strategies to make noisy small-scale devices applicable to solving optimization problems with constraints.

In this work, our main emphasis was on the question of how likely it is to see a feasible sample at the end of a XY-QAOA, which also upper bounds the performance of the algorithm. However, the present analysis could also be applied to other quantum algorithms with particle-number symmetry, e.g., quantum chemistry applications. In variational quantum eigensolver (VQE) or state preparation algorithms applied to many-body electronic systems, among other symmetries, often it is desired that the number of electrons is preserved. Ways to enforce such symmetries, just like in QAOA, include adding terms in the VQE energy function that penalize symmetry violations [9,10] or designing quantum circuits to limit the quantum evolution within the subspace spanned by states with the desired symmetries, regardless of the values of the variational parameters [11,12]. The latter in a noiseless case often predicts superior results compared with the former. Our noise analysis is based on preserving the total  $S^z$  of a spin (qubit) system. While the encoding of a fermionic system into qubits is impeding the direct application of the analysis that we performed to electronic systems, we believe the conclusion, at a high level, applies to all such designed algorithms.

Moreover, these results could be generalized from particle-number symmetries to more general symmetries. Another application beyond benchmarking quantum algorithms lies in using particle-number subspaces to benchmark quantum hardware.

## ACKNOWLEDGMENTS

The authors would like to thank Andrea Skolik, Michael J. Hartmann, Jason M. Dominy, Sheir Yarkoni, and all members of the QuAIL team for helpful discussions. M.S. was supported by the USRA Feynman Quantum Academy funded by the NAMS R&D Student Program at NASA Ames Research Center and by the Air Force Research Laboratory (AFRL), NYSTECUSRA Contract No. FA8750-19-3-6101. E.R., M.S., F.W., and Z.W. are grateful for financial support from NASA Ames Research Center, the AFRL Information Directorate under Grant No. F4HBKC4162G001, and DARPA under Grant No. IAA 8839, Annex 114. This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement No. 828826.

- [1] P. W. Shor, Algorithms for quantum computation: Discrete logarithms and factoring, in *Proceedings of the 35th Annual Symposium on Foundations of Computer Science, Santa Fe, NM, USA* (IEEE, Piscataway, NJ, 1994), pp. 124–134.
- [2] L. K. Grover, A fast quantum mechanical algorithm for database search, in *Proceedings of the Twenty-Eighth Annual ACM Symposium on Theory of Computing* (ACM, New York, 1996), pp. 212–219.
- [3] S. Jordan, Quantum algorithm zoo, <https://quantumalgorithmzoo.org/> (unpublished).
- [4] C. Gidney and M. Ekerå, How to factor 2048 bit rsa integers in 8 hours using 20 million noisy qubits, [arXiv:1905.09749](https://arxiv.org/abs/1905.09749).
- [5] R. Barends, A. Shabani, L. Lamata, J. Kelly, A. Mezzacapo, U. L. Heras, R. Babbush, A. G. Fowler, B. Campbell, Y. Chen *et al.*, Digitized adiabatic quantum computing with a superconducting circuit, *Nature (London)* **534**, 222 (2016).
- [6] L. DiCarlo, J. M. Chow, J. M. Gambetta, L. S. Bishop, B. R. Johnson, D. I. Schuster, J. Majer, A. Blais, L. Frunzio, S. M. Girvin *et al.*, Demonstration of two-qubit algorithms with a superconducting quantum processor, *Nature (London)* **460**, 240 (2009).
- [7] S. Debnath, N. M. Linke, C. Figgatt, K. A. Landsman, K. Wright, and C. Monroe, Demonstration of a small programmable quantum computer with atomic qubits, *Nature (London)* **536**, 63 (2016).
- [8] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. S. L. Brandao, D. A. Buell *et al.*, Quantum supremacy using a programmable superconducting processor, *Nature (London)* **574**, 505 (2019).
- [9] J. R. McClean, J. Romero, R. Babbush, and A. Aspuru-Guzik, The theory of variational hybrid quantum-classical algorithms, *New J. Phys.* **18**, 023023 (2016).
- [10] I. G. Ryabinkin, S. N. Genin, and A. F. Izmaylov, Constrained variational quantum eigensolver: Quantum computer search engine in the Fock space, *J. Chem. Theory Comput.* **15**, 249 (2019).
- [11] K. Seki, T. Shirakawa, and S. Yunoki, Symmetry-adapted variational quantum eigensolver, *Phys. Rev. A* **101**, 052340 (2020).
- [12] B. T. Gard, L. Zhu, G. S. Barron, N. J. Mayhall, S. E. Economou, and E. Barnes, Efficient symmetry-preserving state preparation circuits for the variational quantum eigensolver algorithm, *NPJ Quantum Inf.* **6**, 10 (2020).
- [13] S. Hadfield, Z. Wang, B. O’Gorman, E. G. Rieffel, D. Venturelli, and R. Biswas, From the quantum approximate optimization algorithm to a quantum alternating operator ansatz, *Algorithms* **12**, 34 (2019).
- [14] E. Farhi, J. Goldstone, and S. Gutmann, A quantum approximate optimization algorithm, [arXiv:1411.4028](https://arxiv.org/abs/1411.4028).
- [15] Z. Wang, N. C. Rubin, J. M. Dominy, and E. G. Rieffel, XY mixers: Analytical and numerical results for the quantum alternating operator ansatz, *Phys. Rev. A* **101**, 012320 (2020).
- [16] Z. Jiang, E. G. Rieffel, and Z. Wang, Near-optimal quantum circuit for Grover’s unstructured search using a transverse field, *Phys. Rev. A* **95**, 062317 (2017).
- [17] Z. Wang, S. Hadfield, Z. Jiang, and E. G. Rieffel, Quantum approximate optimization algorithm for MaxCut: A fermionic view, *Phys. Rev. A* **97**, 022304 (2018).
- [18] M. Streif and M. Leib, Training the quantum approximate optimization algorithm without access to a quantum processing unit, *Quantum Sci. Technol.* **5**, 034008 (2020).
- [19] L. Zhou, S.-T. Wang, S. 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).
- [20] S. Bravyi, D. Gosset, and R. Movassagh, Classical algorithms for quantum mean values, *Nat. Phys.* **17**, 337 (2021).
- [21] J. S. Otterbach, R. Manenti, N. Alidoust, A. Bestwick, M. Block, B. Bloom, S. Caldwell, N. Didier, E. Schuyler Fried, S. Hong *et al.*, Unsupervised machine learning on a hybrid quantum computer, [arXiv:1712.05771](https://arxiv.org/abs/1712.05771).
- [22] M. P. Harrigan, K. J. Sung, M. Neeley *et al.*, Quantum approximate optimization of non-planar graph problems on a planar superconducting processor, *Nat. Phys.* **17**, 332 (2021).
- [23] M. Streif and M. Leib, Comparison of QAOA with quantum and simulated annealing, [arXiv:1901.01903](https://arxiv.org/abs/1901.01903).
- [24] M. B. Hastings, Classical and quantum bounded depth approximation algorithms, [arXiv:1905.07047](https://arxiv.org/abs/1905.07047).
- [25] J. Marshall, F. Wudarski, S. Hadfield, and T. Hogg, Characterizing local noise in QAOA circuits, *IOP SciNotes* **1**, 025208 (2020).
- [26] C. Xue, Z.-Y. Chen, Y.-C. Wu, and G.-P. Guo, Effects of quantum noise on quantum approximate optimization algorithm, *Chinese Phys. Lett.* **38**, 030302 (2021).
- [27] M. Streif and M. Leib, Forbidden subspaces for level-1 quantum approximate optimization algorithm and instantaneous quantum polynomial circuits, *Phys. Rev. A* **102**, 042416 (2020).
- [28] M. Streif, S. Yarkoni, A. Skolik, F. Neukart, and M. Leib, Beating classical heuristics for the binary paint shop problem with the quantum approximate optimization algorithm, [arXiv:2011.03403](https://arxiv.org/abs/2011.03403).
- [29] A. Perdomo-Ortiz, N. Dickson, M. Drew-Brook, G. Rose, and A. Aspuru-Guzik, Finding low-energy conformations of lattice protein models by quantum annealing, *Sci. Rep.* **2**, 571 (2012).
- [30] I. Hen and F. M. Spedalieri, Quantum Annealing for Constrained Optimization, *Phys. Rev. Appl.* **5**, 034007 (2016).
- [31] I. Hen and M. S. Sarandy, Driver Hamiltonians for constrained optimization in quantum annealing, *Phys. Rev. A* **93**, 062312 (2016).
- [32] H. Leipold and F. M. Spedalieri, Constructing driver Hamiltonians for several linear constraints, [arXiv:2006.12028](https://arxiv.org/abs/2006.12028).
- [33] S. Okada, M. Ohzeki, and S. Taguchi, Efficient partition of integer optimization problems with one-hot encoding, *Sci. Rep.* **9**, 1 (2019).
- [34] A. Ehrenfeucht, V. Faber, and H. A. Kierstead, A new method of proving theorems on chromatic index, *Discrete Math.* **52**, 159 (1984).
- [35] I. D. Kivlichan, J. McClean, N. Wiebe, C. Gidney, A. Aspuru-Guzik, G. Kin-Lic Chan, and R. Babbush, Quantum Simulation of Electronic Structure with Linear Depth and Connectivity, *Phys. Rev. Lett.* **120**, 110501 (2018).
- [36] A. Cowtan, S. Dilkes, R. Duncan, A. Krajenbrink, W. Simmons, and S. Sivarajah, On the qubit routing problem, [arXiv:1902.08091](https://arxiv.org/abs/1902.08091).
- [37] S. Sivarajah, S. Dilkes, A. Cowtan, W. Simmons, A. Edgington, and R. Duncan, tket): A retargetable compiler for NISQ devices, *Quantum Sci. Technol.* **6**, 014003 (2020).
- [38] K. Wright, K. M. Beck, S. Debnath, J. M. Amini, Y. Nam, N. Grzesiak, J.-S. Chen, N. C. Panti, M. Chmielewski, C. Collins *et al.*, Benchmarking an 11-qubit quantum computer, *Nat. Commun.* **10**, 5464 (2019).

- [39] R. Barends, J. Kelly, A. Megrant, A. Veitia, D. Sank, E. Jeffrey, T. C. White, J. Mutus, A. G. Fowler, B. Campbell *et al.*, Superconducting quantum circuits at the surface code threshold for fault tolerance, *Nature (London)* **508**, 500 (2014).
- [40] J. R. McClean, S. Boixo, V. N. Smelyanskiy, R. Babbush, and H. Neven, Barren plateaus in quantum neural network training landscapes, *Nat. Commun.* **9**, 4812 (2018).
- [41] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition (10th. ed.)* (Cambridge University Press, Cambridge, 2011).
- [42] E. G. Rieffel and W. H. Polak, *Quantum Computing: A Gentle Introduction* (MIT Press, Cambridge, 2011).
- [43] A. H. Myerson, D. J. Szwer, S. C. Webster, D. T. C. Allcock, M. J. Curtis, G. Imreh, J. A. Sherman, D. N. Stacey, A. M. Steane, and D. M. Lucas, High-Fidelity Readout of Trapped-Ion Qubits, *Phys. Rev. Lett.* **100**, 200502 (2008).
- [44] C. K. Andersen, A. Remm, S. Lazar, S. Krinner, N. Lacroix, G. J. Norris, M. Gabureac, C. Eichler, and A. Wallraff, Repeated quantum error detection in a surface code, *Nat. Phys.* **16**, 875 (2020).