

Dual-map framework for noise characterization of quantum computers

James Sud,^{1,2,*} Jeffrey Marshall ,^{1,2} Zihui Wang ,^{1,2} Eleanor Rieffel,¹ and Filip A. Wudarski^{1,2,†}

¹Quantum Artificial Intelligence Laboratory (QuAIL), NASA Ames Research Center, Moffett Field, California 94035, USA

²USRA Research Institute for Advanced Computer Science (RIACS), Mountain View, California 94043, USA



(Received 4 January 2022; revised 18 April 2022; accepted 17 May 2022; published 15 July 2022)

In order to understand the capabilities and limitations of quantum computers, it is necessary to develop methods that efficiently characterize and benchmark error channels present on these devices. In this paper, we present a method that faithfully reconstructs a marginal (local) approximation of the effective noise (MATEN) channel, that acts as a single layer at the end of the circuit. We first introduce a dual-map framework that allows us to analytically derive expectation values of observables with respect to noisy circuits. These findings are supported by numerical simulations of the quantum approximate optimization algorithm (QAOA) that also justify the MATEN, even in the presence of nonlocal errors that occur during a circuit. Finally, we demonstrate the performance of the method on Rigetti's Aspen-11 quantum computer for QAOA circuits up to six qubits, successfully predicting the observed measurements on a majority of the qubits.

DOI: [10.1103/PhysRevA.106.012606](https://doi.org/10.1103/PhysRevA.106.012606)

I. INTRODUCTION

Appropriate and accurate error characterization and benchmarking is vital for many aspects of quantum computation. Understanding dominant forms of error allows for improvements on quantum hardware, bringing these devices closer to the fault-tolerant regime, and possibly allowing for the tailoring of error-correcting codes to specific error channels [1]. On the algorithm side, error characterization opens the possibility for error-aware algorithm design and error mitigation strategies, improving the performance of algorithms on hardware [2,3]. A plethora of protocols have been designed for understanding error. These can be divided into benchmarking protocols, which aim to return numerical values that capture the rate of errors in a process (usually defined as an average fidelity [4,5]), and characterization protocols, which aim to return information about both the level and form of the error channels themselves. Benchmarking protocols include randomized benchmarking [6,7] (along with extensions such as [8]), cycle benchmarking [9], and direct fidelity estimation [10]. Characterization protocols include quantum process tomography [11], gate set tomography [12], Hamiltonian estimation [13], and robust phase estimation [14], as well as state preparation and measurement (SPAM) error characterization methods [15–17]. Moreover, direct spectroscopic methods allow to probe qubit frequency fluctuations and identify phenomena such as $1/f$ noise (see, for example, [18–20]). So far, benchmarking and characterization methods have suffered substantial shortcomings—either returning limited information [e.g., average fidelity for randomized benchmarking (RB)] or restricted to small systems due to exponential scaling (tomographic methods). Furthermore, correlated and

non-Markovian noise is difficult to be faithfully captured [21]; however, some promising techniques have been already proposed (see, for example, [22–26]). In this work, we develop a characterization scheme that efficiently returns information about the process matrix of the marginal noise channel acting on a single qubit. The method combines ease and efficiency of benchmarking techniques with substantially richer information content. Additionally, the introduced protocol operates without additional compilation overhead, as opposed to RB approaches, which require a twirling subroutine to cast the noisy channel into a convenient form of a Pauli channel. Moreover, we demonstrate that the discussed protocol is capable of closely reconstructing nonlocal noise for processes of decent fidelity (above 95%); this, however, does not mean that the protocol is universal and robust against all types of noise (e.g., non-Markovian or $1/f$).

Quantum noise, which can lead to computational errors, is an inevitable companion of quantum evolution. In order to properly describe physically admissible errors, one has to employ the framework of completely positive and trace-preserving (CPTP) maps, which are referred to as error channels. These channels can be represented in numerous ways [27–29]. For our purposes, the most natural representation is of the following form:

$$\mathcal{E}[\rho] = \sum_{k,l=0}^{d^2-1} \chi_{k,l} P_k \rho P_l^\dagger, \quad (1)$$

where P_k are operators, $d = 2^N$ is dimensionality of the Hilbert space for N qubits, and χ is referred to as a process matrix. Setting P_k to orthonormal basis elements (e.g., Pauli matrices), one can determine all elements of the χ matrix via quantum process tomography [11]. In order to represent a valid quantum channel, Eq. (1) has to be CPTP, which happens

*jsud@usra.edu

†fwudarski@usra.edu

when $\chi \geq 0$ (CP condition) and

$$\sum_{k,l=0}^{n^2-1} \chi_{k,l} P_l^\dagger P_k = \mathbb{1} \quad (\text{TP condition}). \quad (2)$$

On noisy intermediate-scale quantum (NISQ) devices, levels of noise are too high for error correction and fault tolerance to occur. Thus, error mitigation, and error-aware algorithm co-design strategies are needed to maximize the performance of algorithms run on these devices. In order to determine these optimal mitigation and co-design strategies, it is imperative to characterize and understand error. A popular and well-studied algorithm for NISQ devices is the quantum approximate optimization algorithm or quantum alternating operator ansatz (QAOA) [30–32], which aims to find approximately optimal solutions to optimization problems. In this work, we study the application of our characterization method for QAOAs run on combinatorial optimization problems. The characterization and effect of local noise in QAOA circuits has been previously studied [33–36]; however, here we provide analytical treatment for popular classes of error channels, as well as for a generic single-qubit noise channel.

Given this understanding of prior work on error characterization and benchmarking, as well as the introduction of error maps and QAOA, we lay out the rest of our paper as follows. In Sec. II, we introduce the framework of the dual map that is essential for analytical derivation of the expectation values of arbitrary observables in the noisy setup. In Sec. III, we use these results to reverse-engineer a method to introduce a marginal approximation of the effective noise (MATEN), the core procedure of this contribution, that allows to estimate local contribution to the error channels. In Sec. IV, we discuss limitations of the MATEN protocol for spatially correlated noise. Finally, in Sec. V we demonstrate the efficacy of the method in characterizing error on classical simulations, as well as on the Aspen-11 quantum computer device from Rigetti Computing.

II. DUAL-MAP FRAMEWORK

Given a quantum state ρ , an error channel \mathcal{E} [in the form of Eq. (1)], and an Hermitian operator O , the noisy expectation value of O with respect to ρ is typically evaluated as $\text{Tr}[O\mathcal{E}(\rho)]$. In this work, however, we consider the dual action of \mathcal{E} , and compute the same expectation as

$$\langle O \rangle = \text{Tr}[\mathcal{E}^\#(O)\rho] = \text{Tr}[O'\rho], \quad (3)$$

where $\mathcal{E}^\#$ is the dual channel of \mathcal{E} , defined as

$$\mathcal{E}^\#[O] = \sum_{k,l=0}^{n^2-1} \chi_{k,l} P_l^\dagger O P_k, \quad (4)$$

which can be derived from the cyclic property of trace.

We can then study properties of the modified operator $O' = \mathcal{E}^\#(O)$, which means that noise affects only the observable and not the state ρ . Therefore, the expectation values with respect to the ideal quantum state ρ (e.g., output of a quantum circuit) can potentially benefit from the local structure of noise of the observables. In particular, if ρ is a pure state (i.e., $\rho^2 = \rho$), we can avoid costly simulations of the density matrix and

focus on unitary simulations and local measurements of a state vector.

Finally, note that it is always possible to decompose a quantum map Λ associated with a noisy quantum circuit as a composition $\Lambda = \mathcal{E} \circ \mathcal{U}$, where \mathcal{U} is the circuit’s (ideal) unitary map, and \mathcal{E} a noisy channel (e.g., with the trivial example $\mathcal{E} = \Lambda \circ \mathcal{U}^\dagger$). Therefore, instead of characterizing the total error map Λ (which includes contribution from the ideal unitary), we focus on determining \mathcal{E} , which one can perceive as an effective noise channel for the circuit (in general dependent on \mathcal{U} , e.g., rotational angles in QAOA). This mathematical trick allows us to “move” effects of noise to the very last layer of the quantum circuit (see Fig. 1) and exploit the dual-map framework. The main advantage of using this formalism is that many observables of interest (e.g., combinatorial or molecular Hamiltonians) can be expressed as a combination of k -local terms, the simulation of which, as explained above, can be significantly more efficient in the dual-map framework. We demonstrate this idea in the subsequent examples.

A. Example: Noisy single-layered QAOA

In this section we analyze the example of QAOA circuits, which are constructed by interleaving layers of parametrized unitaries of a mixing Hamiltonian B and a phasing Hamiltonian H , so

$$|\Psi_{\text{QAOA}}(\vec{\gamma}, \vec{\beta})\rangle = e^{-i\beta_p B} e^{-i\gamma_p H} \dots e^{-i\beta_1 B} e^{-i\gamma_1 H} |\psi_0\rangle, \quad (5)$$

where p is the number of layers in the circuit, $(\vec{\gamma}, \vec{\beta})$ represent length p parameter vectors, and $|\psi_0\rangle$ corresponds to an initial state. Given this form, one can then choose $(\vec{\gamma}, \vec{\beta})$ such that the expectation value of H is optimized (minimized or maximized) when the circuit is applied to a suitably chosen initial state. Strategies for optimizing $(\vec{\gamma}, \vec{\beta})$ [37–40] as well as choosing optimal starting states and mixing Hamiltonians [32,41,42] have been intensively analyzed. In the original formulation, and most applications of QAOA, the cost function is classical, ensuring that its corresponding Hamiltonian consists of Pauli terms only containing the Pauli-Z operators. For this section, we restrict to QAOA applied to the well-studied form of quadratic unconstrained binary optimization (QUBO), with cost functions given by a Hamiltonian of the form

$$H = H_1 + H_2 = \sum_i h_i Z_i + \sum_{ij} J_{ij} Z_i Z_j. \quad (6)$$

Many popular combinatorial optimization problems can be cast into QUBO form [43]. In order to understand the effects of various noise channels on specific problems under certain noise assumptions then, it suffices to compute the action of the dual channel on Pauli terms with limited locality, and analyze how the modified H' Hamiltonians relate to the original cost Hamiltonians.

For parametrized circuits, such as QAOA, in order to perform mathematical analysis, we assume \mathcal{E} is independent of the parameters $(\vec{\gamma}, \vec{\beta})$, and is a product of local channels such that we can write

$$\rho(\vec{\gamma}, \vec{\beta}) = \mathcal{E}\mathcal{U}(\vec{\gamma}, \vec{\beta})[|\psi_0\rangle\langle\psi_0|]. \quad (7)$$

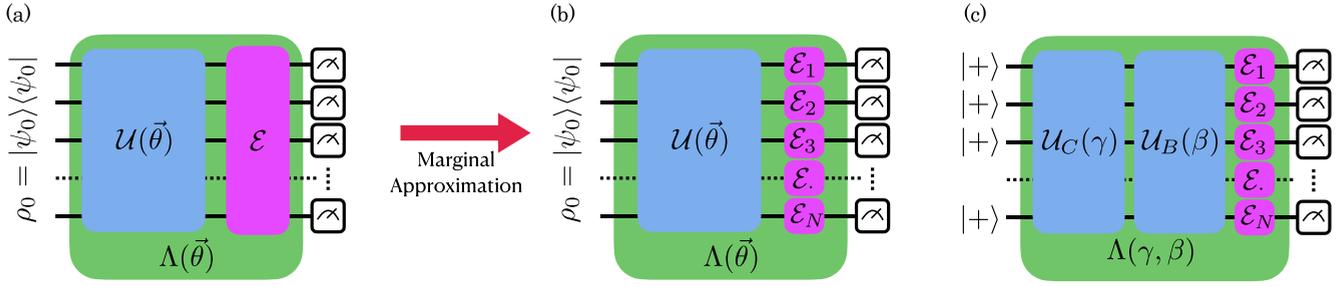


FIG. 1. Operational framework for the noisy circuit characterization that is described by a quantum map $\Lambda(\vec{\theta})$, where $\vec{\theta}$ is a collection of circuit parameters. (a) Noise is “moved” to the last layer [according to $\mathcal{E} = \Lambda(\vec{\theta}) \circ \mathcal{U}^\dagger(\vec{\theta})$], and in general it is of nonlocal character. (b) The nonlocality is neglected by approximating the circuit with only local quantum channels \mathcal{E}_i acting on the i th qubit. (c) The marginal approximation for a single-layered QAOA.

This assumption for QAOA circuits is visualized in Fig. 1(c), with $\mathcal{E} = \bigotimes_{i=1}^N \mathcal{E}_i$ that is equivalent to MATEN (thoroughly described in the next section). A similar noise structure was considered in [33,34].

In the following sections, we analyze the effects of the dual map of various common error channels on Hamiltonians of the form of Eq. (6) with one layer. For these examples, we assume the error channel is identical on each qubit in order to simplify the equations, but this assumption can straightforwardly be relaxed.

1. Single-qubit depolarizing channel

We first define a single-qubit depolarizing channel, parametrized by a depolarization rate p , as

$$\mathcal{E}_p^i(\rho) = \frac{1+3p}{4}\rho + \frac{1-p}{4}\sum_{k=1}^3\sigma_k^i\rho\sigma_k^i, \quad (8)$$

with σ_1, σ_2 , and σ_3 corresponding to Pauli X, Y , and Z respectively, and the indices i corresponding to the qubit that Pauli operators act upon. Note that each Pauli matrix is an eigenmatrix of the depolarizing channel with eigenvalue p , i.e., $\mathcal{E}_p(\sigma_k) = p\sigma_k$, and this map is self-dual ($\mathcal{E} = \mathcal{E}^\#$), so we also have $\mathcal{E}_p^\#(\sigma_k) = p\sigma_k$.

For an N -qubit system we have a noise channel acting on each qubit, i.e., $\mathcal{E}_p = \mathcal{E}_p^1 \otimes \mathcal{E}_p^2 \otimes \dots \otimes \mathcal{E}_p^N$, where \mathcal{E}_p^i corresponds to a one-local channel on qubit i . Therefore, if \mathcal{E}_p acts on a k -local term in the Hamiltonian, and we assume that p is constant on all qubits, \mathcal{E}_p effectively multiplies this term by p^k . Specifically, we have

$$\mathcal{E}_p^\#(Z_i) = pZ_i, \quad (9)$$

$$\mathcal{E}_p^\#(Z_iZ_j) = p^2Z_iZ_j. \quad (10)$$

This allows us to easily identify the action of local depolarizing noise on the QAOA for depth one with the noise channel applied at the end of the circuit, by moving to the dual picture

$$H' = \mathcal{E}_p^\#(H) = pH_1 + p^2H_2. \quad (11)$$

Thus, for single-qubit depolarizing channels, the effect on QAOA cost operators is simply that one-qubit terms are rescaled by p , two-qubit terms are rescaled by p^2 , and k -qubit terms by p^k (although $k > 2$ are not considered for QUBO problems). For a strictly 2-local problem such as MaxCut,

this would mean that the cost is simply rescaled by p^2 . For optimization purposes, this simple rescaling means that the optimal parameter settings stay unchanged.

2. Amplitude damping

Another common error channel is amplitude damping, given by the following map:

$$\mathcal{E}_\gamma\rho = A_1\rho A_1^\dagger + A_2\rho A_2^\dagger, \quad (12)$$

where A_1 and A_2 are Kraus operators parametrized by a damping rate γ and given by

$$A_1 = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-\gamma} \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & \sqrt{\gamma} \\ 0 & 0 \end{pmatrix}. \quad (13)$$

The action of the dual of this error channel on single- and two-qubit Pauli- Z operators are as follows:

$$\mathcal{E}_\gamma^\#(Z_i) = (1-\gamma)Z_i + \gamma I, \quad (14)$$

$$\mathcal{E}_\gamma^\#(Z_iZ_j) = (1-\gamma)^2Z_iZ_j + \gamma(1-\gamma)(Z_i + Z_j) + \gamma^2 I. \quad (15)$$

We can see then that the 1-local terms are simply scaled and shifted. For the 2-local terms, we get not only a scale and a shift from the first and third terms, respectively, but also an extra contribution of 1-local terms from the middle term. We can write out the action of this channel on the general Hamiltonian given in Eq. (6):

$$H' = (1-\gamma)H_1 + \gamma\sum_i h_i + (1-\gamma)^2H_2 + \gamma^2\sum_{i<j} J_{ij} + \gamma(1-\gamma)\sum_{i<j} J_{ij}(Z_i + Z_j). \quad (16)$$

Now the only term that is neither a scale nor a constant shift is the last term. We first note that if $h_i = 0$ for all i and we start in a \mathbb{Z}_2 symmetric state, the resultant QAOA state is \mathbb{Z}_2 symmetric, meaning that the state is invariant under the application of $X^{\otimes N}$; thus all single-qubit Z terms go to zero [44], so this added 1-local term has no effect on the observed cost function value.

Another case where this term has a nice solution can be seen as follows: We can rewrite $\sum_{i<j} J_{ij}(Z_i + Z_j)$ as $\sum_i Z_i(\sum_{j\neq i} J_{ij})$. Next, if $(\sum_{j\neq i} J_{ij}) = ah_i$ for all i and for some constant a , then this term gives us aH_1 , meaning that H_1

is rescaled by $1 - \gamma + a$ instead. This occurs in some cases enumerated below:

(1) If all h 's and J 's are constant (all equal h, J , respectively),

$$\sum_i Z_i \left(\sum_{j \neq i} J_{ij} \right) = \frac{J(N-1)}{h} H_1. \quad (17)$$

(2) For a d -regular graph, all h 's constant, and all nonzero J 's constant (all equal h, J , respectively),

$$\sum_i Z_i \left(\sum_{j \neq i} J_{ij} \right) = \frac{Jd}{h} H_1. \quad (18)$$

(3) For a max- k -colorable subgraph [42] ($h_i = d_i$ where d_i is degree of vertex i , $J_{ij} = -1$ if edge (i, j) exists in the graph),

$$\sum_i Z_i \left(\sum_{j \neq i} J_{ij} \right) = - \sum_i d_i Z_i = -H_1. \quad (19)$$

Notably, if $a = -1$ as in case 3, the Hamiltonian reduces to

$$H' = (1 - \gamma)^2 H + \gamma \sum_i h_i + \gamma^2 \sum_{i < j} J_{ij}, \quad (20)$$

where we see that the entire Hamiltonian is simply scaled and shifted.

For an analysis of the effects of other common error channels on QAOA operators, such as Pauli channels, T1/T2 error, and over-rotations, please see Appendix B.

III. SINGLE-QUBIT NOISE CHARACTERIZATION

So far we have shown how certain local noise channels affect 1-local and 2-local observables, given complete knowledge of the noise. In this section, however, we demonstrate the opposite direction, showing how to exploit the dual-map framework to find a MATEN, which is defined as follows:

Definition 1 (MATEN). For a unitary quantum circuit \mathcal{U} acting on N qubits, and its noisy realization Δ_U , we call $\mathcal{E} = \Delta_U \circ \mathcal{U}^\dagger$ an effective noise channel, that acts as the final CPTP circuit layer. Additionally we define a marginal approximation of the effective noise (MATEN) as

$$\tilde{\mathcal{E}} = \bigotimes_{k=1}^N \text{Tr}_{\bar{k}}(\mathcal{E}) = \bigotimes_{k=1}^N \mathcal{E}_k, \quad (21)$$

where $\text{Tr}_{\bar{k}}(\mathcal{E}) = \mathcal{E}_k$ traces out all subsystems except the k th subsystem [see Fig. 1(b)].

In order to determine a MATEN, we express a noisy map in terms of a so-called process matrix (or χ matrix) that can be in principle measured directly in a set of experiments via quantum process tomography [11]. The map takes the form (for a single qubit)

$$\mathcal{E}_\chi(\rho) = \sum_{k,l=0}^3 \chi_{kl} \sigma_k \rho \sigma_l, \quad (22)$$

where χ_{kl} are elements of the χ matrix, which in general can be expressed as

$$\chi = \begin{pmatrix} p_0 & t_{0,1} + iv_{0,1} & t_{0,2} + iv_{0,2} & t_{0,3} + iv_{0,3} \\ t_{0,1} - iv_{0,1} & p_1 & t_{1,2} - it_{0,3} & t_{1,3} + it_{0,2} \\ t_{0,2} - iv_{0,2} & t_{1,2} + it_{0,3} & p_2 & t_{2,3} - it_{0,1} \\ t_{0,3} - iv_{0,3} & t_{1,3} - it_{0,2} & t_{2,3} + it_{0,1} & p_3 \end{pmatrix}, \quad (23)$$

with t_{kl} and v_{kl} representing real and imaginary parts of χ_{kl} elements, respectively. This form, combined with conditions $\chi \geq 0$ and $\sum_{k=0}^3 p_k = 1$, guarantees that the map is completely positive and trace preserving, and is the most general for the qubit systems. Note that in total we have 12 free parameters, and diagonalizing the χ matrix will lead to a Kraus form (note that the Kraus form is not unique). Given χ we can then evaluate the effect of the dual map on the Pauli observables

$$\tilde{I} = I, \quad (24)$$

$$\begin{aligned} \tilde{X} &= (p_0 + p_1 - p_2 - p_3)X + 4t_{01}I \\ &\quad + 2((t_{12} - v_{03})Y + (t_{13} + v_{02}))Z, \end{aligned} \quad (25)$$

$$\begin{aligned} \tilde{Y} &= (p_0 + p_2 - p_1 - p_3)Y + 4t_{02}I \\ &\quad + 2((t_{23} - v_{01})Z + (t_{12} + v_{03}))X, \end{aligned} \quad (26)$$

$$\begin{aligned} \tilde{Z} &= (p_0 + p_3 - p_1 - p_2)Z + 4t_{03}I \\ &\quad + 2((t_{13} - v_{02})X + (t_{23} + v_{01}))Y, \end{aligned} \quad (27)$$

where $\tilde{I}, \tilde{X}, \tilde{Y}$, and \tilde{Z} represent the noisy transformations of the Pauli operators, and $\tilde{I} = I$ due to the property that the duals of trace-preserving maps are unital [i.e., $\mathcal{E}^\#(I) = I$]. We can rewrite the coefficients in each equation as P_{AB} , forming a vector of coefficients \vec{P} , so, for example, $\tilde{X} = P_{XI}I + P_{XX}X + P_{XY}Y + P_{XZ}Z$. We can also write a simple matrix A that relates coefficients P_{AB} to the χ matrix elements as in $\vec{P} = A\vec{\chi}$, where $\vec{\chi}$ is a 12-dimensional vector having all independent χ matrix elements (i.e., p_k, t_{kl} , and v_{kl}).

Given Eqs. (24)–(27) we can then perform the following procedure for a parametrized circuit of interest:¹

(1) Choose a set \mathbb{S} of parameters to the circuit. For example, for level 1 QAOA this corresponds to choosing $|\mathbb{S}|$ different (γ_1, β_1) pairs.

(2) Implement the circuit on a quantum device, and take many measurements in the X, Y , and Z bases to approximate $\langle \tilde{X} \rangle$, $\langle \tilde{Y} \rangle$, and $\langle \tilde{Z} \rangle$ for each parameter setting in \mathbb{S} on each qubit. Since $\langle \tilde{I} \rangle$ is trivial, the measurement is not needed.

(3) On a classical simulator or via analytic derivation, determine the ideal values of the $\langle X \rangle$, $\langle Y \rangle$, and $\langle Z \rangle$ for each parameter setting in \mathbb{S} for each qubit. $\langle I \rangle$ is trivial to calculate.

(4) Using the ideal and noisy values of all four Pauli observables, determine the coefficients \vec{P} via linear regression on Eqs. (24)–(27) for each qubit.

¹The extension to parameter-free circuit is straightforward, and requires only altering some gates, e.g., $X \rightarrow Z$. However, this procedure would disturb the investigated algorithm, and could serve only as a characterization protocol.

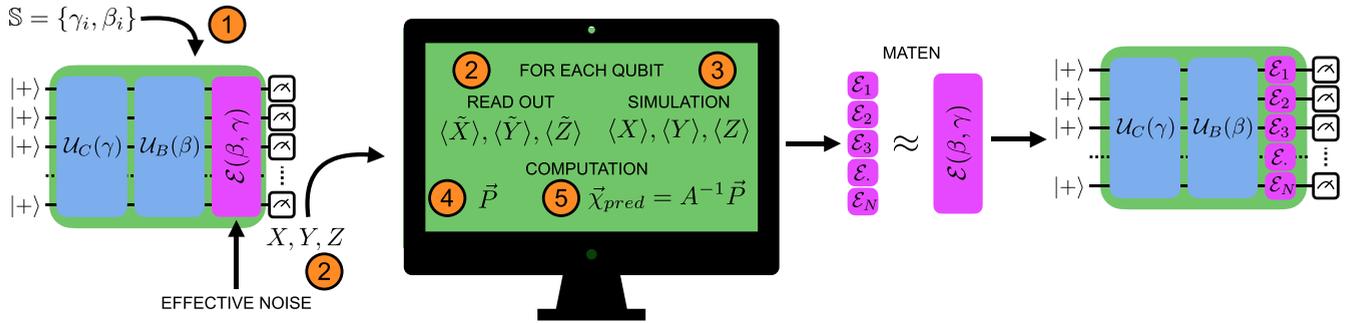


FIG. 2. Protocol for characterizing local approximation to effective noise in the case of a single-layered QAOA algorithm. (1) Run multiple times a circuit with different parameters from the set \mathcal{S} ; (2) next measure all qubit registers in X , Y , and Z bases. Use output bit strings to infer (on a classical computer) noisy expectation values $\langle \tilde{X} \rangle$, $\langle \tilde{Y} \rangle$, and $\langle \tilde{Z} \rangle$. For the same set of parameters, (3) simulate ideal circuits on a classical computer to obtain expectation values. Based on the ideal expectation values, (4) construct vector \vec{P} , and then (5) determine $\vec{\chi}_{pred} = A^{-1} \vec{P}$. Perform this procedure for each qubit in order to reconstruct the MATEN that approximates effective noise.

(5) Given the coefficients \vec{P} , along with the matrix A relating \vec{P} to $\vec{\chi}$ matrix elements, perform $\vec{\chi}_{pred} = A^{-1} \vec{P}$ for each qubit, where $\vec{\chi}_{pred}$ are the elements of the predicted χ matrix.

The above protocol is visualized in the chart in Fig. 2. We note two limitations with the presented procedure. First is that of step 3: in general, it may be prohibitive to determine the ideal values of single-qubit expectation values in simulation. However, for shallow circuits, one can use reverse light cone arguments to calculate local operator expectation values in time and memory growing exponentially with circuit depth, rather than circuit size [45]. Additionally, if the noise channels *mildly* depend on circuits,² one could perform this characterization process on a few sets of qubits individually; this approach would work especially for shallow circuits. Finally, state-of-the-art classical simulators can handle circuits with relatively large depth and qubit number, depending on simulation methods and computational resources.

Second, we note that the QAOA applied to MaxCut, as analyzed in this work, is \mathbb{Z}_2 symmetric. For problems with \mathbb{Z}_2 symmetry, the ideal values of $\langle Y \rangle$ and $\langle Z \rangle$ vanish, so it may be impossible to fully determine \vec{P} , and it remains an open question if we can reliably determine nonzero elements of \vec{P} . If this is the case, one can derive similar equations as Eqs. (24)–(27), but for two-qubit operators, although this becomes much more complicated. For our analysis, we restrict to problems that lack \mathbb{Z}_2 symmetry. For a problem such as MaxCut, this can be achieved by simply adding single-qubit Z terms to the Hamiltonian. Presumably, these single-qubit Z gates do not introduce a significant amount of noise (on Rigetti devices, they are indeed implemented in software), so the χ_{pred} matrix should remain close to that of the original circuit. Thus, the characterized channels for these modified problems should match very well those of the original problems. One can also break this symmetry by starting in a different initial state. For QAOA problems the initial state is usually $|+\rangle^{\otimes N}$, which is \mathbb{Z}_2 itself, but changing $|+\rangle$ to a different non- \mathbb{Z}_2 symmetric state would break that symmetry.

²Here by mildly we mean that noise is static, and parameter (e.g., angle) independent to the leading order.

IV. LOCAL VERSUS NONLOCAL CHANNELS

One of the major challenges in current technology is understanding spatial correlations in noise. Whether or not noise is confined locally to a single qubit, or can be correlated across neighboring (or even distant) qubits (such as in crosstalk [46,47]), determines the efficacy of error mitigation techniques, and quantum error correction (where errors are typically assumed to be independent). Here we aim to find out how well one can approximate nonlocal noise channels with the MATEN approach, leaving temporally correlated noise (e.g., non-Markovian or $1/f$) for future consideration.³ Our strategy is as follows: (i) first we derive a lower bound for the worst-case scenario, (ii) then we numerically compute accuracy of the method for random nonlocal channels, and (iii) finally we repeat numerical analysis from step (ii) but for random Pauli channels, and analyze some scaling properties.

Since the single-qubit χ matrix (in Pauli basis) is a positive operator of trace one, we can treat it as a four-dimensional (4D) quantum state (with some extra constraints imposed by the structure of χ). This enables us to incorporate results from the theory of quantum entanglement for the analysis of nonlocal channels. In particular, all the marginal states for maximally entangled states are maximally mixed states; i.e., they are proportional to the identity matrix. Therefore, the marginal approximation (MA), which on the level of the χ matrix, is translated to

$$\chi = |\Psi\rangle\langle\Psi| \rightarrow \chi_{MA} = \bigotimes_{k=1}^N \left(\frac{1}{4} \mathbb{1} \right) = \frac{1}{4^N} \mathbb{1}_{4^N}, \quad (28)$$

and also yields the maximally mixed state in the full 4^N -dimensional space (where N is the number of considered qubits), which corresponds to the fully depolarizing channel. Above we denote the nonlocal process matrix χ as the maximally entangled state, which is defined as a projector onto $|\Psi\rangle = \frac{1}{\sqrt{2}} \sum_{i=0}^3 |ii \dots i\rangle$, with N 4D subsystems (each corresponding to a qubit); note that this is a GHZ state [48].

³Note that, on the ensemble average level, one expects to describe the effective noise channels with a CPTP map; however, correlation lengths may exceed the averaging time, leading to ill-characterized noise.

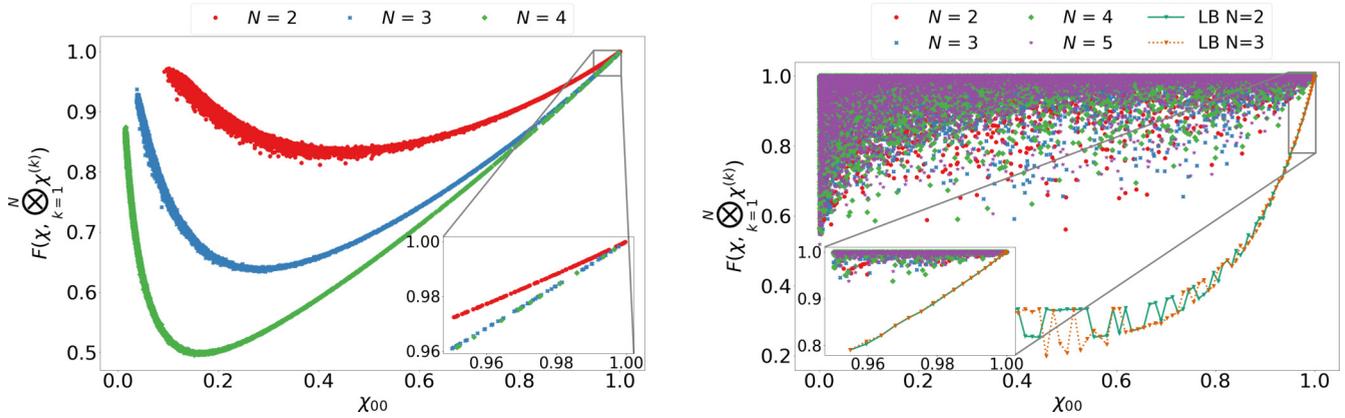


FIG. 3. Comparison between random nonlocal χ matrix and its MA in terms of fidelity. The left plot depicts the case of full random channels, while the right plot restricts the analysis to random Pauli channels. Additionally, we provide a numerical lower bound (LB) on Pauli channels with two and three qubits (right plot). In both cases, insets magnify the region of χ_{00} that is relevant for “high” -fidelity circuits (above 95%), where random nonlocal channels can be reasonably well approximated.

We conjecture that the effective channel with the maximally entangled χ matrix is the worst-case scenario for the proposed MATEN protocol. Since the MATEN approach neglects all nontrivial correlations between different subsystems, and maximally entangled states exhibit the strongest correlations among quantum objects resulting in minimal knowledge of the subsystem’s structure (maximally mixed state), the protocol yields the minimum fidelity value between the MA and the full χ . However, this conjecture requires more rigorous treatment, which we leave as an open problem. Note that maximally entangled χ is a completely valid choice, since $\chi \geq 0$ and the map associated with it is trace preserving.

Having established that the maximally entangled χ matrix is the limiting case for the protocol, now we determine the accuracy of this approximation. For this purpose we incorporate the fidelity of quantum states as a useful figure of merit. We compute it for the nonlocal χ matrix and its MA. Since the MA gives a trivial state, one can easily compute the fidelity [4,49]

$$F(\chi, \chi_{MA}) = \frac{1}{4^N} \text{Tr}(\sqrt{\chi})^2 = \frac{1}{4^N}, \quad (29)$$

where we used the fact that χ is a projector [i.e., $\sqrt{\chi} = \chi$, and $\text{Tr}(\chi) = 1$]. As mentioned before, this result represents the worst-case scenario, and is unlikely to happen in real experiments (especially if one is interested in low-depth circuits), where hardware building blocks operate on fairly high gate fidelity (95–99%, with lower fidelities for multiqubit gates, and higher for single-qubit ones). Therefore, we can escape this unfavorable scaling by restricting to channels that are close to the perfect (noiseless) case, i.e., to the identity channel ($\chi_{00} = 1$ and all other elements equal to zero). This also implies that nonlocal effects are comparably small to the leading order, which is predominantly determined by the χ_{00} element. Similar restrictions are commonly considered in benchmarking literature (see for example [50]), since they represent noise regimes that are more relevant for the current hardware technology and help tailor error-correcting schemes. One may expect high-fidelity quantum circuits ($\chi_{00} > 0.95$) in the future, and therefore milder contribution from spatial correlation (i.e., nonlocality). In order to properly address

this issue, we incorporate numerical methods to find out how well the MA can represent the true nonlocal noise process. Here, we use random sampling of full χ matrices and random samples of Pauli channels (i.e., χ matrices with a random probability vector on the diagonal and all other elements equal to zero). For the case of the full random χ processes, we explore systems composed of $N = 2, 3, 4$ qubits, while for Pauli channels we additionally look at $N = 5$. The results are displayed in Fig. 3, where we took 10 000 samples of random channels (generated with QUTIP [51]) and computed all marginals of the multiqubit χ matrix (i.e., tracing out all but one qubit) and compared fidelity between a tensor product of the marginals (essentially what we call the MA) and the nonlocal one.

For random Pauli channels, we additionally numerically minimize the fidelity between the 4^N -dimensional probability vector representing the Pauli channel, and its MA.⁴ In order to guarantee a genuine probability distribution over our parameters (without having to impose any constraint) we use the modified Hurwitz parametrization for the probability vector [52,53]:

$$\begin{aligned} \chi = & \text{diag}[\cos^2(\theta_{4^N-1}), \cos^2(\theta_{4^N-2}) \sin^2(\theta_{4^N-1}), \dots, \\ & \times \sin^2(\theta_1) \sin^2(\theta_2), \dots, \sin^2(\theta_{4^N-1})] \end{aligned} \quad (30)$$

We employ a Sequential Least Squares Programming (SLSQP) [54] optimization routine to find the lower bound. Surprisingly, two- and three-qubit channels display similar lower bounds (in particular for high-fidelity channels, i.e., χ_{00} close to one). The key observation is that for channels with reasonably large χ_{00} (corresponding to the identity channel), which is directly related to the gate or circuit fidelity, the MA can provide results with acceptable accuracy, which is demonstrated in Fig. 3 insets. Therefore, the MATEN protocol

⁴The marginal approximation for Pauli channels is also done on the level of matrices, and not on the probability vectors.

identifies a MA that can estimate the leading order of the effective noise channel, which also validates approximation for that region of fidelity.

V. RESULTS

In this section we present the success of the method presented in Sec. III for noisy simulations.

A. Classical simulation

For classical simulations, we test our characterization method against a variety of noise sources. Noiseless and noisy classical simulations are performed via pure state and density matrix simulations with HYBRIDQ, an open-source hybrid quantum simulator [55]. In some cases, we additionally generate and apply error channels via QUTIP [51], an open-source toolbox that allows for classical simulation of open quantum systems. With this capability of finding ideal and noisy states and operators, we can easily compute metrics needed to evaluate our method. For all of these experiments, we test the characterization method on parametrized QAOA circuits for QUBO problems.

1. Purely local noise

First we test the efficacy of the characterization method laid out in Sec. III for predicting χ matrices that we manually apply at the end of noiseless classical simulation. To do this, we pick a χ_{in} matrix by iteratively selecting elements uniformly randomly from the interval $[0,1]$ for the elements \bar{p} , and $[-1, 1]$ for \bar{r} and \bar{v} in Eq. (23), and checking if the resultant map is physical (i.e., $\chi \geq 0$) until we succeed. We further choose the same χ_{in} matrix on each qubit, although this is relaxed in the next section. We additionally choose random QUBO problems by randomly drawing J and h from a uniform distribution in the range $[0,1]$. In this experiment, we should expect that for some reasonable number of parameter settings (size of \mathbb{S}) and for a sufficient number of shots (measurements), we should be able to exactly recover the input χ_{in} matrix to arbitrary precision, as the noise is taken to fit perfectly within the MATEN approximation. We quantify the accuracy of determining χ_{in} by taking the L2 distance between the elements of χ_{in} and χ_{pred} , the process matrix our method predicts. The results for various values of shot number and number of regression angles are plotted in Fig. 4. These plots are generated using state-vector simulations for perfect evaluation of observables.

Indeed, we see that for a typical case, increasing the number of angles and the number of shots used in regression allows for more accurate determination of χ_{in} . We further see from this figure that the L2 distance shrinks with added number of shots, by roughly a factor of 10 when the number of shots increases by a factor of 10. We later numerically see this roughly polynomial scaling with the number of shots for various values of $|\mathbb{S}|$. For instance, with $|\mathbb{S}| = 16$ we find the L2 distance goes as $numshots^{-1.39}$. We note that the L2 distance between randomly chosen χ matrices was numerically found to be 0.800 ± 0.125 , but we see fidelities much higher than this value for sufficiently large $|\mathbb{S}|$ and number of shots.

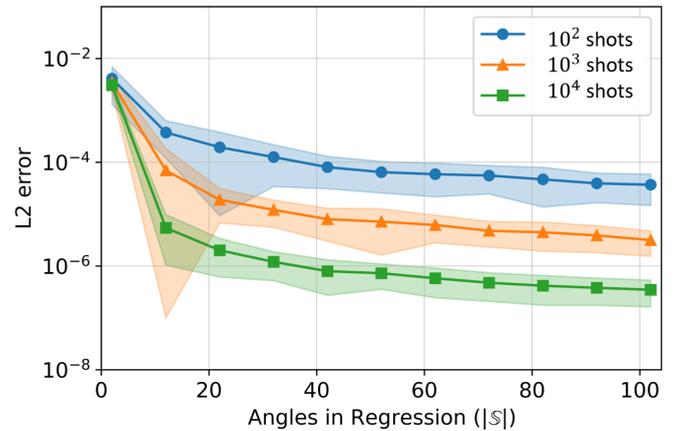


FIG. 4. Average L2 distance between the true and predicted χ for classical simulations for randomly chosen two-qubit QUBO instances with weights in the range $[0,1]$, as a function of the number of angles, $|\mathbb{S}|$, used in the regression and the number of shots used in the estimation of expectation values. Solid lines depict the average over 100 runs, and shading depicts one standard deviation above and below the average. For a large number of shots and angles, the distance is below 10^{-6} .

2. Nonlocal noise at end of circuit

In order to test the resiliency of the noise characterization procedure, we must test the method against noise models that a MATEN is not suited to perfectly capture. For the first of these models, we choose a constant error channel that exists only at the end of a quantum circuit, but is not a simple tensor product of single-qubit channels. In order to apply this combination of local and nonlocal noise, then, we apply an error map of the form

$$\mathcal{E} = (1 - c)\mathcal{E}_1^{(N)} + c\mathcal{E}_n^{(N)}, \quad (31)$$

where we have a combination of a purely local channel $\mathcal{E}_1^{(n)}$ and a nonlocal channel $\mathcal{E}_n^{(N)}$ weighted by a correlation factor $c \in [0, 1]$. In this section we allow the χ_{in} matrices to vary for each qubit (in $\mathcal{E}_1^{(n)}$). Since due to the addition of extra noise ($\mathcal{E}_n^{(n)}$) we no longer expect that $\chi_{pred} \approx \chi_{in}$, we no longer report the fidelity between the two. Instead, we use $(1/3)[r(\langle X \rangle, \langle \tilde{X} \rangle) + r(\langle Y \rangle, \langle \tilde{Y} \rangle) + r(\langle Z \rangle, \langle \tilde{Z} \rangle)]$, the average Pearson correlation coefficient $r(x, y)$ between the measured and predicted noisy expectation values of $\langle X \rangle$, $\langle Y \rangle$, and $\langle Z \rangle$, as these tell us how well our noise model predicts simple observables of interest on the quantum device. However, it is possible to induce overfitting, especially when the number of considered parameter settings ($|\mathbb{S}|$) is small. Thus we additionally look at correlations for an additional “testing set” of parameter settings. For our experiments at around $|\mathbb{S}| = 50$, however, these correlations very closely matched that of the training set, so we only present correlations of the testing set for the following cases. In addition to correlation, we also use Choi fidelity, defined as $F(\Phi_1, \Phi_2)$, the state fidelity between Choi matrices Φ_1, Φ_2 , representing respectively the entire n -qubit maps generated from the chosen error channels, and the predicted MATEN from our method. We present the

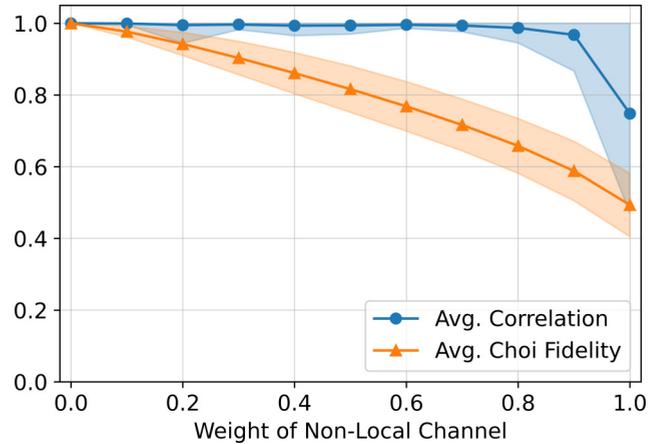
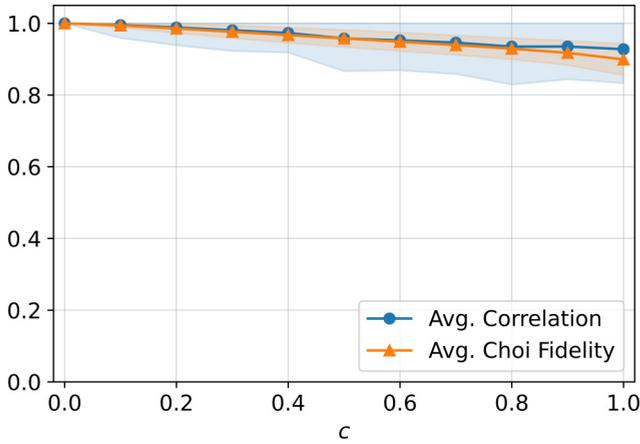


FIG. 5. Average testing correlations and full state fidelities between the actual noise model and predicted noise model for two-qubit (left) and four-qubit (right) fully connected QUBO problems with all $J = 1$ and all $h = 0$, as a function of the weight of the applied nonlocal channel, or c in Eq. (31). Solid lines depict the average over 100 runs; shading depicts one standard deviation above and below the average.

results from the characterization of this noise model in Fig. 5. For these experiments, we fix our problem Hamiltonian to a fully connected QUBO instance with all $J = 1$ and all $h = 0$. Evolution and expectation values are evaluated using density matrix simulation.

From these simulations we see that, as expected, when $c = 0$ and there is only local noise, the model works extremely well. However, as more nonlocal noise is added into the system, the ability to accurately predict the expectation values of Pauli observables begins to falter. At $c = 1$, we typically see a sharp downturn of correlations, as at this point we are not injecting any purely local noise into our system, thus weakening the accuracy of the MATEN.

3. Sampling noise

In addition to the above tests, we also experimented with adding in sampling noise to our noisy simulations. To accomplish this, we choose random Gaussian perturbations with mean zero and standard deviation of $1/\sqrt{\text{numshots}}$ to add to all expectation value measurements, simulating the effect of sampling error on the evaluation of expectation values. Given this form of noise, we repeated the analysis from above, running QAOA with cost Hamiltonian given by Eq. (6) with two qubits and all $h = 0$, $J = 1$. We varied the number of shots on the x axis, and the results of this setup are shown in Fig. 6.

From the simulations we can see that sampling noise diminishes the ability of the method to accurately fit the noisy measurements to ideal measurements, as well as predict the value of noisy measurements. The stochastic noise causes expectation values to fluctuate between measurements, thus essentially introducing a nonconstant noise model. This may cause poor performance as our method depends on having the same error channels for all angles and measurement bases.

Additionally, we note that poor performance may arise if errors are angle dependent, leading to an error model that is nonconstant between different angles in a similar manner to sampling noise. Errors can be extremely angle dependent on quantum computers, especially for parametrized two-qubit

gates such as in [56], so this feature could be an important limitation in the success of the method in the near term.

4. Larger systems: Phasing and mixing rotation error

Finally, in order to scale our simulations to larger system sizes, we performed our characterization routine on 10-qubit instances. For these runs, selecting and applying randomly generated 10-qubit Kraus maps becomes numerically prohibitive, so we switch to a simpler and more realistic noise model. For these experiments, we assume that there is some stochastic error, or deviation in the parameters for both the phasing and mixing operators. In particular, the QAOA angles are assumed to be normally distributed about the desired mean value, with a nonzero standard deviation that defines the total amount of noise. For a given phasing gate $e^{-i\gamma Z_i Z_j}$, this noise is introduced through the Kraus operators

$$A_1 = \sqrt{1 - \omega}I, A_2 = \sqrt{\omega}Z_i Z_j, \quad (32)$$

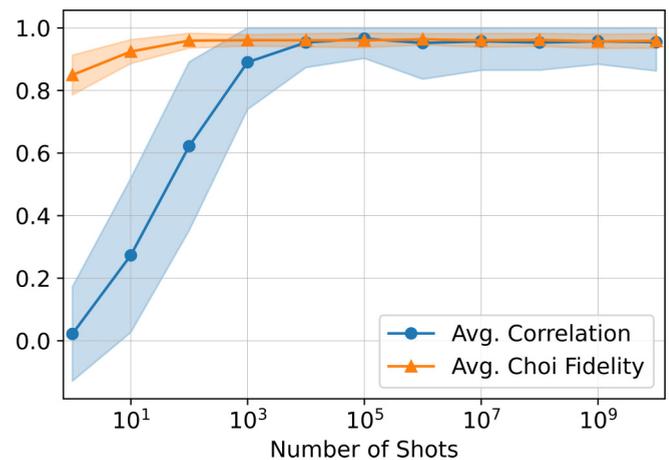


FIG. 6. Average testing correlations and Choi fidelities between the actual noise model and predicted noise model for the two-qubit QUBO problem with $J = 1$ and both $h = 0$, as a function of the number of shots used to estimate expectation values. Solid lines depict the average over 100 runs; shading depicts one standard deviation above and below the average.

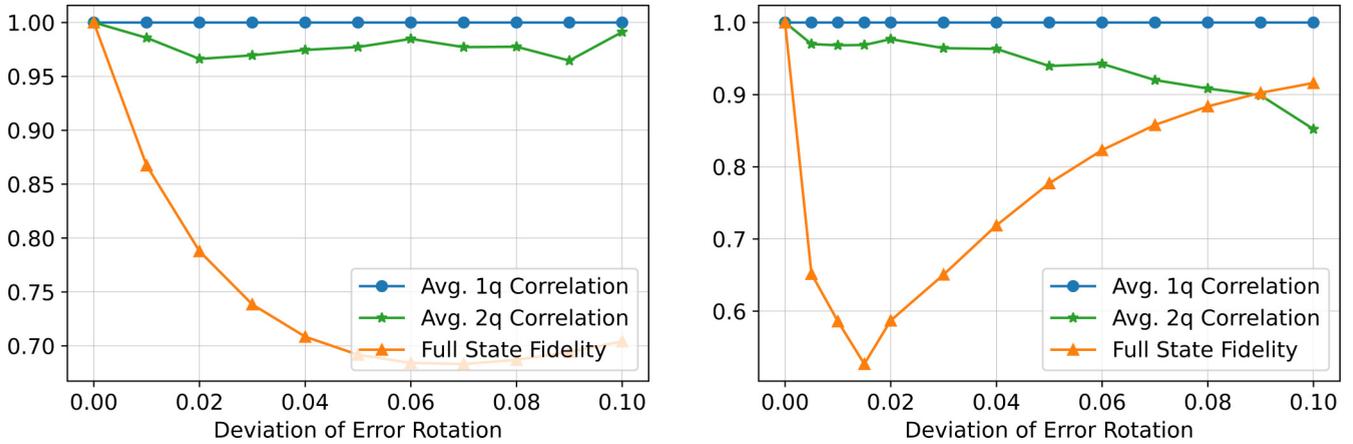


FIG. 7. Average testing correlations and Choi fidelities between the actual noise model and predicted noise model for ten-qubit ring (left) and fully connected (right) QUBO problems with all $J = 1$ and all $h = 0$, as a function of the deviation ω of both the phasing and mixing operators.

and for a mixing gate $e^{-i\beta X_i}$ we apply

$$A_1 = \sqrt{1 - \omega}I, \quad A_2 = \sqrt{\omega}X_i. \quad (33)$$

Here ω defines the amount of noise (related to the standard deviation in the angles' values). A derivation of these noise models is shown in Appendix Sec. B 3. This model applies the two-qubit dephasing noise layer [Eq. (32)] on each pair (i, j) of qubits that the noiseless phase gates act, after the dephasing unitaries and directly before the mixing layer. After the mixing layer the one-qubit X noise is applied on each qubit [Eq. (33)].

Under this noise model we can test our characterization method on larger systems, and test against the assumption that all noise is applied at the end of the circuit. We display the results for the method on ten-qubit-ring and fully connected QUBO problems in Fig. 7. For these plots, no matter the value of ω , we saw that we were able to perfectly reproduce one-qubit correlations, so we chose to add in the average of all two-qubit correlations as well. Additionally, we report the average fidelity between the actual ten-qubit noisy density matrix and the predicted density matrix using the characterized noise model. From these results we find that the fidelity drops rapidly, especially for the fully connected case. Crucially, however, the one- and two-qubit correlations remain very high, even as the ω grows. We note that on the fully connected plot, the fidelity rises after $\omega \approx .02$. This is likely explained by the fact that $\omega = 0.5$ corresponds to the maximally dephasing channel, which our model can capture well. Thus we expect to see the fidelity drop initially as ω grows, then rise back to 1 when $\omega = 0.5$, and then follow a symmetric pattern once $\omega > 0.5$. From these results, however, our main takeaway is that even in the presence of noise which is not local and not strictly at the end of the circuit, the method finds a suitable MATEN that is able to replicate single-qubit expectation values perfectly and two-qubit expectation values very well, even as we scale to large system sizes.

VI. CHARACTERIZATION OF RIGETTI ASPEN-11 DEVICE

In this section we apply the error characterization method from Sec. III to the Aspen-11 Quantum Processing Unit

(QPU) from Rigetti Computing [57]. We run the characterization procedure for QAOA circuits with all qubits initialized in the $|+\rangle$ (using simultaneous Pauli- X rotations on all qubits, decomposed into \sqrt{X} and Pauli- Z rotations). We then apply a phase-separation layer given by Hamiltonians of the form in Eq. (6), with all $h_i = 1$ (to break \mathbb{Z}_2 symmetry, and implemented via single-qubit Pauli- Z rotations) and $J_{ij} = \delta_{i+1,j}$ (forming a line topology, and implemented using parametrized CPHASE(γ) two-qubit gates along with Pauli- Z rotations); with two-qubit gates applied in two separate layers, no single qubit is involved in multiple two-qubit gates simultaneously. Finally, we apply a mixing layer via the standard X mixer implemented by simultaneous parametrized Pauli- X rotations on each qubit. For each experiment we choose p , the number of layers, to be one. These gates were coded into the Rigetti PYQUIL framework using native device gates (Pauli- X and - Z rotations, as well as CPHASE gates), and were transpiled and scheduled using Rigetti's native_quil_to_executable method. We then perform 1000 measurements of each qubit simultaneously in the Pauli X , Y , and Z bases, by applying H , HS^\dagger , and I to the circuit, respectively, and decomposed into native gates.

These experiments were run at $N = 2$ and $N = 6$ with $|\mathcal{S}| = 100$, where N is the number of qubits and $|\mathcal{S}|$ is the number of different parameter settings used. For these experiments, we run under three cases:

(a) 1q only: Remove all two-qubit (CPHASE) gates (equivalent to setting all J_{ij} to zero). The intention of this is to make sure that our method works when only single-qubit gates are present, removing the main sources of crosstalk and nonlocal noise, which could distort the results.

(b) 2q idle: Add back in two-qubit (CPHASE) gates, but set the angles (γ) of all two-qubit gates to zero (again equivalent to setting all J_{ij} to zero). This ideally implements the same circuit as the previous case, but two-qubit gates are physically implemented in the circuit.

(c) 2q active: Lift the restriction of setting two-qubit gate angles to zero, thus performing the method completely as intended.

For these experiments, much like Sec. V A 2, we present statistics on the correlations between predicted and observed

TABLE I. Correlations between single-qubit Pauli observables obtained by experiments vs derived MATEN approximation for QAOA circuits on a linear chain QUBO instance for qubits 10 and 11 on the Rigetti Aspen-11 device

Qubit	1q only	2q idle	2q active
10	0.9961	0.9955	0.9757
11	0.9826	0.9862	0.9778

Pauli expectation values, given by

$$\frac{1}{3}[r(\langle X \rangle, \langle \tilde{X} \rangle) + r(\langle Y \rangle, \langle \tilde{Y} \rangle) + r(\langle Z \rangle, \langle \tilde{Z} \rangle)], \quad (34)$$

with the average Pearson correlation coefficient $r(x, y)$ between the measured and predicted noisy expectation values of $\langle X \rangle$, $\langle Y \rangle$, and $\langle Z \rangle$. This correlation thus tells us how accurately the characterized MATEN predicts single-qubit Pauli observables. These are shown for both the two- and six-qubit cases in Table I.

For the two-qubit experiments, we see that the method is able to predict expectation values of all Pauli observables with a high correlation to the experimental values. This denotes that in all three cases, the derived MATEN is able to accurately predict expectation values of single-qubit Pauli observables, even in the presence of two-qubit gates with arbitrary angles.

The six-qubit experiments are presented in Table II. Here, we see that all metrics remain high for the “1q only” case, but for the “2q idle” case for qubits 30, 36, and 37 we see a significant drop in regression score and average correlation. In the “2q active” case, we see a further decline in the correlations of qubits 30, 32, and 37. For this case, which matches most closely the type of experiments we would like to characterize, our method gives an average expectation value correlation of 0.414 ± 0.337 . These values are far from the ideal values of 1, but the positive correlation values suggest that the method does better than a random attempt at guessing 1-local channels. The wide variability in performance on various qubits suggests that certain qubits may have more angle-dependent noise, or may have larger sources of crosstalk, as analyzed in Sec. V A. In particular, the correlations of qubit 30 plummet from ≈ 0.65 in the “2q idle” case to ≈ 0.07 in the “2q active” case. This correlation is much lower than we see even on the right-hand side of Fig. 5 or anywhere in Fig. 7. This indicates that the errors introduced by two-qubit gates are in a sense worse than both of these cases. We suspect this may be due to the fact that the added two-qubit gate, even with all angles set

TABLE II. Correlations between single-qubit Pauli observables obtained by experiments vs derived MATEN approximation for QAOA circuits on a linear chain QUBO instance for qubits 21, 36, 37, 30, 31, and 32 on the Rigetti Aspen-11 device.

Qubit	1q only	2q idle	2q active
21	0.9905	0.9871	0.9411
36	0.9944	0.6884	0.5983
37	0.7660	0.3097	0.1884
30	0.9455	0.6472	0.0735
31	0.9652	0.8864	0.6518
32	0.9952	0.9272	0.0297

to zero, may introduce some significant noise (e.g., crosstalk between the two qubits) that is far from the intended phasing operation, which the MATEN is not equipped to accurately handle; i.e., the nonlocal channel displays a low value of χ_{00} beyond the protocol’s range. In the simulations we perform, artificially added errors come in the form of randomly chosen Kraus maps or over-rotations, but the error maps on a quantum device may be of a specific, biased, more detrimental form. Additionally, even with a two-qubit gate with angle set to zero, it can be the case that a different unitary is applied from shot to shot, approaching the case of Sec. V A 3, which is the only source of noise we found to reduce correlations to such a low number. Thus we suspect that this error or shot-dependent noise may play a role in the extremely low correlations, as we would not expect to be able to accurately characterize any noise procedure that is changing over time.

We additionally note that for the six qubit experiments, qubit 37 has significantly lower correlation than the other qubits in the “1q only” case. Given the calibration data from the experiment, as presented in Appendix A, it appears that qubit 37 has an unusually low readout fidelity. While the MATEN should in theory correct for arbitrary uncorrelated readout errors, it may be the case that these errors are correlated on the Aspen-11 device. We also note the CPHASE fidelity between qubit 30 and 37 is below 0.9. It may be the case that the relatively poor performance on qubit 37 is due to errors, such as large over-rotations, in the CPHASE gate, which are shown in Fig. 7 to reduce the effectiveness of the MATEN.

VII. DISCUSSION

In this paper we introduce the dual-map framework for computing the effects of error maps on expectation values evaluated on a quantum computer. We then present a method to compute a marginal approximation of the effective noise (MATEN) of a parametrized quantum circuit that is efficient in terms of the number of measurements needed to perform on a quantum computer and is simple to implement. We demonstrate that the method effectively computes a MATEN for local noise at the end of a circuit, and demonstrate that it can be effective even in the presence of nonlocal and inter-circuit noise, especially when the noise is only weakly correlated, i.e., the circuit’s fidelity is reasonably high with $\chi_{00} > 0.95$. We finally show that the method is effective in computing a MATEN on a few qubits of Rigetti’s Aspen-11 quantum computer. Lower values in extracted correlations of expectation values can inform us that the system exhibits a fair amount of angle-dependent (gate) noise, as well as errors that are absent in the theoretical model, e.g., readout or leakage to the noncomputational subspace. The latter can be modeled in a similar fashion as qubits under our scheme, with the difference that the χ process matrix now needs to represent a qudit process. This, however, introduces an extra layer of complexity, which we leave for future analysis.

The error characterization method can additionally be used as a proxy for the fidelity of a gate, layer, or entire circuit, as the values of the computed χ matrices for each qubit (specifically χ_{00}) quantify the difference between the ideal and noisy evolution. Furthermore, the returned χ can inform about dominant sources of error, which can in turn point to particularly

TABLE III. Qubit specification data for the Aspen-11 device from Rigetti Computing as of 2:45 p.m. PST, 1 April 2022. Only information for qubits utilized in Sec. V is presented. RB, randomized benchmarking.

Qubit	Qubit frequency (GHz)	1q RB fidelity	T1 (μ s)	T2 (μ s)	Active reset fidelity	Readout fidelity
10	5.1448	0.9979 \pm 0.0002	19.1550	14.9876	0.9985	0.9610
11	3.7471	0.9986 \pm 0.0001	134.3175	7.6476	0.9840	0.9280
17	3.8316	0.9979 \pm 0.0006	44.0395	18.8874	0.9885	0.9390
21	3.4809	0.9971 \pm 0.0011	63.8754	16.7968	0.9905	0.8865
24	5.2950	0.9985 \pm 0.0002	6.4071	3.4892	0.9925	0.9865
25	3.8223	0.9989 \pm 0.0003	19.2390	10.3975	0.9790	0.8970
26	5.1310	0.9981 \pm 0.0002	22.1224	14.4986	0.9965	0.9595
30	4.2514	0.9982 \pm 0.0006	38.1255	7.6651	0.9900	0.9365
31	3.7286	0.9953 \pm 0.0004	53.8022	8.5207	0.9805	0.9150
32	5.3175	0.9976 \pm 0.0002	29.1685	9.3002	0.9985	0.9565
36	5.1471	0.9985 \pm 0.0002	26.8752	11.5777	0.9865	0.9635
37	3.8340	0.9991 \pm 0.0002	50.3642	34.0851	0.9790	0.8695

effective strategies from error mitigation, leading to algorithmic improvements on NISQ devices. Once dominant sources of error are determined, we leave these error-specific mitigation approaches as open problems for the reader.

The introduced dual-map framework can be used to understand which error channels can be specifically detrimental for a circuit. For instance, with QAOA, we show that depolarizing noise simply flattens the energy landscape; thus it does not affect the location of optimal parameters for the algorithm. However, error sources such as amplitude damping may introduce nontrivial behavior. The characterization procedure we introduce can be used to characterize error in NISQ devices, especially for shallow circuits in which the effective noise channels are expected to be noncorrelated. Overall, the dual-map picture for error channels provides a simple and elegant method for researching the interplay between quantum error and algorithms in the future, and our characterization approach can significantly aid hardware-aware algorithm design on today's devices. Finally, we expect that the future generation of hardware will benefit more from the discussed framework, as we expect to observe a substantial increase in quantum control leading to higher fidelity circuits, which are better approximated by the MATEN protocol.

ACKNOWLEDGMENTS

This material is based upon work supported by the U.S. Department of Energy, Office of Science, National Quantum Information Science Research Centers, Superconducting Quantum Materials and Systems Center (SQMS) under Contract No. DE-AC02-07CH11359 through NASA-DOE interagency agreement SAA2-403602. All authors appreciate support from the NASA Ames Research Center. J.S., J.M., Z.W., and F.A.W. are thankful for support from NASA Academic Mission Services, Contract No. NNA16BD14C.

APPENDIX A: ASPEN-11 DEVICE CALIBRATION DATA

In this section we present supplementary experimental information for the qubits and gates used in the experiments presented in Sec. V, collected at 2:45 p.m. Pacific Standard Time on 1 April 2022 on Rigetti's Aspen-11 device (see Tables III and IV). All data are pulled directly from provided

Rigetti device specifications and calibration data. We use the abbreviation RB for randomized benchmarking. Active reset refers to the fidelity of resetting qubits to the $|0\rangle$ state by first measuring a qubit in the computational basis and applying an X gate if the qubit is measured as 1. No errors in T1, T2, qubit frequency, and active reset/readout fidelities are provided by Rigetti.

APPENDIX B: MORE ERROR MAPS

In this section we derive the effects of various other error channels on the Pauli Z and ZZ terms found in QUBO problems, in the same style as Sec. II A. We first present the following proof to aid in the analysis:

Theorem 1. For \mathbb{Z}_2 -symmetric states, the expectation value of a Pauli string \mathcal{P} is zero if the number of Pauli Z 's plus Pauli Y 's in \mathcal{P} is odd.

Proof. We assume an operator \mathcal{O} on N qubits that is of the form $\prod_{i=1}^N \sigma_{p_i}$, where p_i represents the Pauli that acts on qubit i : either X , Y , Z , or I . We can then define S_X , S_Y , S_Z to be the set of qubits in which our operator is X , Y , and Z . We then look at the expectation value of the most general operator of this form,

$$\langle \Psi | \mathcal{O} | \Psi \rangle = \left\langle \Psi \left| \prod_{i \in S_Y} Z_i \prod_{j \in S_Y} Y_j \prod_{k \in S_X} X_k \right| \Psi \right\rangle. \quad (\text{B1})$$

TABLE IV. Two-qubit gate specification data for the Aspen-11 device from Rigetti Computing as of 2:45 p.m. PST, 1 April 2022. Only information for two-qubit gates utilized in Sec. V is presented.

Edge	CPHASE Fidelity
17 10	0.977337 \pm 0.004702
10 11	0.990532 \pm 0.005010
11 26	0.965611 \pm 0.006795
25 26	0.966236 \pm 0.008327
24 25	0.849931 \pm 0.011106
21 36	0.930675 \pm 0.009189
36 37	0.916244 \pm 0.004991
37 30	0.885947 \pm 0.006471
30 31	0.896984 \pm 0.006994
31 32	0.891021 \pm 0.007039

Then noting that $Y = -iZX$, we write

$$\begin{aligned} & -i \left\langle \Psi \left| \prod_{i \in S_Z} Z_i \prod_{j \in S_Y} Z_j X_j \prod_{k \in S_X} X_k \right| \Psi \right\rangle \\ & = -i \left\langle \Psi \left| \prod_{i \in S_Z \cup S_Y} Z_i \prod_{k \in S_X \cup S_Y} X_k \right| \Psi \right\rangle. \end{aligned} \quad (\text{B2})$$

We can then expand out $|\Psi\rangle$ in terms of bit strings l ,

$$-i \left\langle \Psi \left| \prod_{i \in S_Z \cup S_Y} Z_i \prod_{k \in S_X \cup S_Y} X_k \sum_l c_l (|l\rangle + |\bar{l}\rangle) \right. \right\rangle. \quad (\text{B3})$$

We can then define $|l^*\rangle = \prod_{k \in S_X \cup S_Y} X_k |l\rangle$ and write

$$-i \left\langle \Psi \left| \prod_{i \in S_Z \cup S_Y} Z_i \sum_l c_l (|l^*\rangle + |\bar{l}^*\rangle) \right. \right\rangle, \quad (\text{B4})$$

$$-i \sum_l c_l \left(\left\langle \Psi \left| \prod_{i \in S_Z \cup S_Y} Z_i |l^*\rangle \right. \right\rangle + \left\langle \Psi \left| \prod_{i \in S_Z \cup S_Y} Z_i |\bar{l}^*\rangle \right. \right\rangle \right). \quad (\text{B5})$$

Then we can note the following two properties. First, since Φ is \mathbb{Z}_2 symmetric, $\langle \Phi | l^* \rangle = \langle \Phi | \bar{l}^* \rangle$ by definition. Also, for general state $|l\rangle$, we note that $Z_i |\bar{l}\rangle = |\bar{l}\rangle \langle l | (-Z_i) |l\rangle$ for i on any qubit. Using the second property we have

$$\begin{aligned} & -i \sum_l c_l \left(\left\langle \Psi \left| \prod_{i \in S_Z \cup S_Y} Z_i |l^*\rangle \right. \right\rangle \right. \\ & \left. + \left(\left\langle \Psi \left| \prod_{i \in S_Z \cup S_Y} (-Z_i) |\bar{l}^*\rangle \right. \right\rangle \langle l^* | (-Z) |l^*\rangle \right) \right), \end{aligned} \quad (\text{B6})$$

where $|\bar{x}\rangle$ represents the inverse of $|x\rangle$, obtained by flipping all qubits in the state. Rewriting and then using the first property, we see

$$\begin{aligned} & -i \sum_l c_l \left(\left\langle \Psi \left| \prod_{i \in S_Z \cup S_Y} Z_i |l^*\rangle \right. \right\rangle \right. \\ & \left. + (-1)^{S_Y + S_Z} \left\langle \Psi \left| \prod_{i \in S_Z \cup S_Y} |\bar{l}^*\rangle \right. \right\rangle \langle l^* | (-Z_i) |l^*\rangle \right) \end{aligned} \quad (\text{B7})$$

$$\begin{aligned} & = -i \sum_l c_l \left(\left\langle \Psi \left| \prod_{i \in S_Z \cup S_Y} Z_i |l^*\rangle \right. \right\rangle \right. \\ & \left. + (-1)^{|S_Y| + |S_Z|} \left\langle \Psi \left| \prod_{i \in S_Z \cup S_Y} |\bar{l}^*\rangle \right. \right\rangle \langle l^* | Z_i |l^*\rangle \right) \end{aligned} \quad (\text{B8})$$

$$\begin{aligned} & = -i \sum_l c_l \left(\left\langle \Psi \left| \prod_{i \in S_Z \cup S_Y} Z_i |l^*\rangle \right. \right\rangle \right. \\ & \left. + (-1)^{|S_Y| + |S_Z|} \left\langle \Psi \left| \prod_{i \in S_Z \cup S_Y} |\bar{l}^*\rangle \right. \right\rangle \langle l^* | Z_i |l^*\rangle \right) \end{aligned} \quad (\text{B9})$$

$$\begin{aligned} & = -i \sum_l c_l \left(\left\langle \Psi \left| \prod_{i \in S_Z \cup S_Y} Z_i |l^*\rangle \right. \right\rangle \right. \\ & \left. + (-1)^{|S_Y| + |S_Z|} \left\langle \Psi \left| \prod_{i \in S_Z \cup S_Y} Z_i |l^*\rangle \right. \right\rangle \right) \end{aligned} \quad (\text{B10})$$

$$= -i \sum_l c_l (1 + (-1)^{|S_Y| + |S_Z|}) \left\langle \Psi \left| \prod_{i \in S_Z \cup S_Y} Z_i |l^*\rangle \right. \right\rangle, \quad (\text{B11})$$

where we use the first property again in the second-to-last step. Now we can clearly see that if $|S_Y| + |S_Z|$ is odd this inner product will vanish for all l . ■

We will reference this theorem in the following analyses.

1. Generic single-qubit channel

The action of the dual of a generic single-qubit channel defined by χ in 1 on Z and ZZ is as follows:

$$\mathcal{E}^\#(Z) = p_Z Z + p_Y Y + p_X X + p_I I, \quad (\text{B12})$$

$$\begin{aligned} \mathcal{E}^\#(ZZ) &= p_Z^2 ZZ + p_Y^2 YY + p_X^2 XX + p_I^2 I \\ &+ P_Z P_Y (ZY + YZ) + P_Z P_X (ZX + XZ) \\ &+ P_Z P_I (ZI + IZ) + P_Y P_X (YX + XY) \\ &+ P_Y P_I (YI + IY) + P_X P_I (XI + IX), \end{aligned} \quad (\text{B13})$$

where $p_Z = p_0 + p_3 - p_1 - p_2$, $p_Y = 2(t_{23} + v_{02})$, $p_X = 2(t_{13} - v_{02})$, and $p_I = 4t_{03}$.

Thus, the action on QAOA Hamiltonians of the form given in Eq. (6) is

$$\begin{aligned} H' &= \sum_i h_i (p_Z Z_i + p_Y Y_i + p_X X_i + p_I I) \\ &+ \sum_{i < j} J_{ij} (p_Z^2 Z_i Z_j + p_Y^2 Y_i Y_j + p_X^2 X_i X_j + p_I^2 I) \\ &+ \sum_{i,j} J_{i,j} (P_Z P_Y Z_i Y_j + P_Z P_X Z_i X_j + P_Z P_I Z_i I \\ &+ P_Y P_X Y_i X_j + P_Y P_I Y_i I + P_X P_I X_i I). \end{aligned} \quad (\text{B14})$$

From here various assumptions can be made. If all $h = 0$, which is the case for MaxCut and strictly 2-local QUBO problems, we can eliminate all terms with odd number of $Z + Y$ terms from Theorem 1 in this Appendix. We could also assume that all but p_Z are small, meaning that the noise channel is relatively close to the identity, which is a condition that would be satisfied on quantum hardware with low levels of noise. This would allow us to eliminate all terms quadratic in p_X , p_Y , and p_Z . If we make these two assumptions we reduce to

$$H' = p_Z^2 H + p_Z p_Y \sum_{i,j} J_{ij} Z_i Y_j. \quad (\text{B16})$$

This corresponds to a simple rescaling of the Hamiltonian, plus an additional, nontrivial term, which is examined in Appendix Sec. B 2.

2. Constant mixing over-rotations

Take a very simple model, where the phase is applied correctly but the mixer is applied as $H_M = \sum_i \tilde{\beta}_i X_i$, where $\beta_i = \beta + \delta\beta_i$, i.e., a small over- or under-rotation in the x direction. This means the mixing unitary is of the form $U_M' U_M$, where $U_M = e^{-i\beta \sum_i X_i}$ and $U_M' = e^{-i \sum_i \delta\beta_i X_i}$.

This locally rotates the Hamiltonian, $H' = U_M'^\dagger H U_M'$. Now we use $e^{-i\theta X} Z e^{i\theta X} = \cos(2\theta)Z - \sin(2\theta)Y$.

Assuming the standard form $H = \sum_{i<j} J_{ij} Z_i Z_j$,

$$\begin{aligned} H' &= \sum_{i<j} J_{ij} \cos(2\delta\beta_i) \cos(2\delta\beta_j) Z_i Z_j \\ &+ \sum_{i<j} J_{ij} \cos(2\delta\beta_i) \sin(2\delta\beta_j) Z_i Y_j \\ &+ \sum_{i<j} J_{ij} \sin(2\delta\beta_i) \cos(2\delta\beta_j) Y_i Z_j \\ &+ \sum_{i<j} J_{ij} \sin(2\delta\beta_i) \sin(2\delta\beta_j) Y_i Y_j. \end{aligned} \quad (\text{B17})$$

Let us average the fluctuations $\langle \cos 2\delta\beta_i \cos 2\delta\beta_j \rangle = \cos^2(2\delta\beta)$, etc. This gives the noise-averaged Hamiltonian

$$\begin{aligned} H' &= \cos^2(2\delta\beta) H + \sin(4\delta\beta) \sum_{i,j} \frac{J_{ij}}{2} Z_i Y_j \\ &+ \sin^2(2\delta\beta) \sum_{i<j} J_{ij} Y_i Y_j, \end{aligned} \quad (\text{B18})$$

where in the middle sum, we now sum over all i and all j .

We see that first, the spectrum is flattened by a factor of $\cos^2(2\delta\beta)$, but second, the terms in Y modify it in a nontrivial way. Let us look at a perturbation to order $\delta\beta$:

$$H' = H + 2\delta\beta \sum_{i,j} J_{ij} Z_i Y_j + O(\delta\beta^2). \quad (\text{B19})$$

The first-order correction to any energy level is

$$\sum_{i,j} J_{ij} \langle E | Z_i Y_j | E \rangle = 0, \quad (\text{B20})$$

using that $|E\rangle$ is just some z bit string.

This suggests we must go to second-order perturbation, looking at terms of the form

$$\langle E_1 | Y_j | E_2 \rangle, \quad (\text{B21})$$

where E_2 is a single bit flip from E_1 . Since the magnitude of the change depends on the energy difference $E_2 - E_1$, the correction depends strongly on the spectrum of the original problem.

Since the eigenstates of classical Hamiltonians are computational basis states, just like in the Hamming weight Hamiltonian, the shifted eigenstates are as before. We can then again calculate

$$\langle H' \rangle = \sum_n E_n \left| \langle m | \left(\frac{|n\rangle - i\epsilon \sum_k X_k |n\rangle}{\sqrt{1 + N\epsilon^2}} \right) \right|^2 \quad (\text{B22})$$

$$= \frac{1}{1 + N\epsilon^2} \left(E_m + \epsilon^2 \sum_n E_n \left| \langle m | \sum_k X_k |n\rangle \right|^2 \right). \quad (\text{B23})$$

We now note that the inner product is 1 if and only if n is a bit flip away from m and if k is the index of the bit that is flipped.

We can also calculate the energy difference between E_m and E_n in this case. Here we compute without loss of generality the case where we let k be the vertex of highest index:

$$\begin{aligned} E_m - E_n &= \left(\sum_{\substack{i<j \\ j \neq k}} J_{ij} \langle m | Z_i Z_j | m \rangle + \sum_{i \neq k} J_{ik} \langle m | Z_i Z_k | m \rangle \right) \\ &- \left(\sum_{\substack{i<j \\ j \neq k}} \langle m | X_k Z_i Z_j X_k | m \rangle \right. \\ &\left. + \sum_{i \neq k} J_{ik} \langle m | X_k Z_i Z_k X_k | m \rangle \right). \end{aligned} \quad (\text{B24})$$

Then we use the fact that $X_k Z_k X_k = -Z_k$ and that X_k commutes through $Z_i Z_j$. We can then cancel and add terms, giving us

$$2 \sum_{i \neq k} J_{ik} \langle m | Z_i Z_k | m \rangle. \quad (\text{B25})$$

Then we can easily rearrange to see that $E_n = E_m - 2 \sum_{i \neq k} J_{ik} \langle m | Z_i Z_k | m \rangle$. We can then plug this expression back in for E_n in Eq. (B22):

$$\frac{1}{1 + N\epsilon^2} \left(E_m + \epsilon^2 \left(\sum_k E_m - 2 \sum_{i \neq k} J_{ik} \langle Z_i Z_k \rangle \right) \right) \quad (\text{B26})$$

$$= \frac{1}{1 + N\epsilon^2} \left(E_m + \epsilon^2 N E_m - 4 \sum_{i < k} J_{ik} \langle Z_i Z_k \rangle \right) \quad (\text{B27})$$

$$= E_m \left(\frac{1 + (N - 4)\epsilon^2}{1 + N\epsilon^2} \right) \quad (\text{B28})$$

$$\approx E_m (1 - 4\epsilon^2 + 4N\epsilon^4 + O(\epsilon^6)). \quad (\text{B29})$$

So for this case over-rotations also just scale the old eigenvalues.

3. Nonconstant mixing over-rotations

We may consider that instead of having perfect angles, they demonstrate small stochastic fluctuations. We will exploit a model based on the von Mises distribution of angles (i.e., normal distribution on a circle) and we are looking for the maps

$$\mathcal{E}_U^\alpha(\rho) = \int_0^{2\pi} \frac{e^{\kappa \cos(\epsilon)}}{2\pi I_0(\kappa)} U(\alpha + \epsilon) \rho U^\dagger(\alpha + \epsilon) d\epsilon, \quad (\text{B30})$$

where $1/\kappa$ is variance, $I_0(\kappa)$ is a modified Bessel function, and $U(\alpha + \epsilon)$ is our mixer of phase operator set to angle α with fluctuation ϵ .

The integral for the mixer leads to the following map:

$$\begin{aligned} \mathcal{E}_M^\beta(\rho) &= \frac{1}{2} \left(1 + \frac{I_2(\kappa)}{I_0(\kappa)} \cos(2\beta) \right) \rho \\ &+ \frac{1}{2} \left(1 - \frac{I_2(\kappa)}{I_0(\kappa)} \cos(2\beta) \right) X \rho X \\ &- \frac{i I_2(\kappa) \sin(2\beta)}{I_0(\kappa)} \frac{1}{2} [X, \rho], \end{aligned} \quad (\text{B31})$$

while for a single gate of phase (assuming that we have a cost Hamiltonian $H = \sum J_{ij} Z_i Z_j$), the single ZZ($J_{ij}\gamma$) gate (for $J_{ij} = \{+1, -1\}$) is given by

$$\begin{aligned} \mathcal{E}_p^{\gamma,J}(\rho) &= \frac{1}{2} \left(1 - \frac{I_2(\kappa)}{I_0(\kappa)} \right) A_1(\gamma, J) \rho A_1^\dagger(\gamma, J) \\ &+ \frac{1}{2} \left(1 + \frac{I_2(\kappa)}{I_0(\kappa)} \right) A_2(\gamma, J) \rho A_2^\dagger(\gamma, J), \end{aligned} \quad (\text{B32})$$

where

$$\begin{aligned} A_1(\gamma, J) &= \text{Diag}[(1, -e^{2i\gamma J}, -e^{2i\gamma J}, 1)], \\ A_2(\gamma, J) &= \text{Diag}[(1, e^{2i\gamma J}, e^{2i\gamma J}, 1)]. \end{aligned} \quad (\text{B33})$$

The mixer error map can also be looked at as a composition of a perfect rotation by β , then an application of $\mathcal{E}_M^\beta(\rho)$ with $\beta = 0$. In this case we get a self-dual channel of the form

$$\begin{aligned} \mathcal{E}_M^\beta(\rho) &= \frac{1}{2} \left(1 + \frac{I_2(\kappa)}{I_0(\kappa)} \cos(2\beta) \right) \rho \\ &+ \frac{1}{2} \left(1 - \frac{I_2(\kappa)}{I_0(\kappa)} \cos(2\beta) \right) X \rho X. \end{aligned} \quad (\text{B34})$$

This maps

$$Z \rightarrow \frac{I_2(\kappa)}{I_0(\kappa)} Z. \quad (\text{B35})$$

And the analysis follows local depolarizing noise. We note that in the limit of high variance, $\kappa \rightarrow 0$, and $\frac{I_2(\kappa)}{I_0(\kappa)} \rightarrow 0$, so we have complete disorder in Z ($Z \rightarrow 0$). In the limit of zero variance, $\kappa \rightarrow \infty$, and $\frac{I_2(\kappa)}{I_0(\kappa)} \rightarrow 1$, so there is no effect of the channel ($Z \rightarrow Z$).

4. Pauli channel

An interesting class of noisy channel is the so-called Pauli channel, that has form

$$\mathcal{E}(\rho) = \sum_{k=0}^3 p_k \sigma_k \rho \sigma_k, \quad (\text{B36})$$

where $\sigma_0 = \mathbb{1}$, and for $k = 1, 2, 3$ we get Pauli X, Y , and Z , respectively. $p_k \geq 0$ and $\sum_k p_k = 1$. One may interpret that a given noisy channel (e.g., bit flip corresponding to X , happens with a respective probability). One can set $p_1 = p_2 = p_3 = (1 - p)/4$ and $p_0 = (1 + 3p)/4$ and get depolarizing channel as in Eq. (8).

The action of the dual of a Pauli channel on Z is

$$\mathcal{E}^\#(Z) = (p_0 + p_3 - (p_1 + p_2)) \times Z. \quad (\text{B37})$$

This then reduces to depolarizing noise where we set $p = p_0 + p_3 - (p_1 + p_2)$. The results are similar for other Pauli matrices, with the difference that for the Pauli k matrix ($1 = x, 2 = y, 3 = z$), we will have $p_0 + p_k - (p_i + p_j)$, where i, j are the coefficients standing in front of remaining matrices.

5. Phase damping

Phase damping has also two Krauses, with A_1 the same as for the amplitude damping case analyzed in Sec. II A 2 and A_2 given in

$$A_2 = \begin{pmatrix} 0 & 0 \\ 0 & \sqrt{\gamma} \end{pmatrix}. \quad (\text{B38})$$

In this case, the error maps Z to itself with no scaling or shifting, so the error channel has no effect on classical Hamiltonians.

-
- [1] C. Piveteau, D. Sutter, S. Bravyi, J. M. Gambetta, and K. Temme, Error Mitigation for Universal Gates on Encoded Qubits, *Phys. Rev. Lett.* **127**, 200505 (2021).
 - [2] K. Temme, S. Bravyi, and J. M. Gambetta, Error Mitigation for Short-Depth Quantum Circuits, *Phys. Rev. Lett.* **119**, 180509 (2017).
 - [3] Y. Li and S. C. Benjamin, Efficient Variational Quantum Simulator Incorporating Active Error Minimization, *Phys. Rev. X* **7**, 021050 (2017).
 - [4] M. A. Nielsen, A simple formula for the average gate fidelity of a quantum dynamical operation, *Phys. Lett. A* **303**, 249 (2002).
 - [5] F. Wudarski, J. Marshall, A. Petukhov, and E. Rieffel, Augmented fidelities for single-qubit gates, *Phys. Rev. A* **102**, 052612 (2020).
 - [6] J. Emerson, R. Alicki, and K. Życzkowski, Scalable noise estimation with random unitary operators, *J. Opt. B: Quantum Semiclassical Opt.* **7**, S347 (2005).
 - [7] E. Magesan, J. M. Gambetta, and J. Emerson, Scalable and Robust Randomized Benchmarking of Quantum Processes, *Phys. Rev. Lett.* **106**, 180504 (2011).
 - [8] J. Claes, E. Rieffel, and Z. Wang, Character randomized benchmarking for non-multiplicity-free groups with applications to subspace, leakage, and matchgate randomized benchmarking, *PRX Quantum* **2**, 010351 (2021).
 - [9] A. Erhard, J. J. Wallman, L. Postler, M. Meth, R. Stricker, E. A. Martinez, P. Schindler, T. Monz, J. Emerson, and R. Blatt, Characterizing large-scale quantum computers via cycle benchmarking, *Nat. Commun.* **10**, 5347 (2019).
 - [10] S. T. Flammia and Y.-K. Liu, Direct Fidelity Estimation from Few Pauli Measurements, *Phys. Rev. Lett.* **106**, 230501 (2011).
 - [11] I. L. Chuang and M. A. Nielsen, Prescription for experimental determination of the dynamics of a quantum black box, *J. Mod. Opt.* **44**, 2455 (1997).
 - [12] D. Greenbaum, Introduction to quantum gate set tomography, [arXiv:1509.02921](https://arxiv.org/abs/1509.02921).
 - [13] S. G. Schirmer, A. Kolli, and D. K. L. Oi, Experimental Hamiltonian identification for controlled two-level systems, *Phys. Rev. A* **69**, 050306(R) (2004).
 - [14] S. Kimmel, G. H. Low, and T. J. Yoder, Robust calibration of a universal single-qubit gate set via robust phase estimation, *Phys. Rev. A* **92**, 062315 (2015).
 - [15] M. Sun and M. R. Geller, Efficient characterization of correlated SPAM errors, [arXiv:1810.10523](https://arxiv.org/abs/1810.10523).

- [16] J. Lin, J. J. Wallman, I. Hincks, and R. Laflamme, Independent state and measurement characterization for quantum computers, *Phys. Rev. Research* **3**, 033285 (2021).
- [17] M. Werninghaus, D. J. Egger, and S. Filipp, High-Speed Calibration and Characterization of Superconducting Quantum Processors without Qubit Reset, *PRX Quantum* **2**, 020324 (2021).
- [18] G. A. Álvarez and D. Suter, Measuring the Spectrum of Colored Noise by Dynamical Decoupling, *Phys. Rev. Lett.* **107**, 230501 (2011).
- [19] J. Bylander, S. Gustavsson, F. Yan, F. Yoshihara, K. Harrabi, G. Fitch, D. G. Cory, Y. Nakamura, J.-S. Tsai, and W. D. Oliver, Noise spectroscopy through dynamical decoupling with a superconducting flux qubit, *Nat. Phys.* **7**, 565 (2011).
- [20] C. M. Quintana, Y. Chen, D. Sank, A. G. Petukhov, T. C. White, D. Kafri, B. Chiaro, A. Megrant, R. Barends, B. Campbell, Z. Chen, A. Dunsworth, A. G. Fowler, R. Graff, E. Jeffrey, J. Kelly, E. Lucero, J. Y. Mutus, M. Neeley, C. Neill *et al.*, Observation of Classical-Quantum Crossover of $1/f$ Flux Noise and Its Paramagnetic Temperature Dependence, *Phys. Rev. Lett.* **118**, 057702 (2017).
- [21] E. Onorati, T. Kohler, and T. Cubitt, Fitting quantum noise models to tomography data, [arXiv:2103.17243](https://arxiv.org/abs/2103.17243).
- [22] P. Figueroa-Romero, K. Modi, R. J. Harris, T. M. Stace, and M.-H. Hsieh, Randomized benchmarking for non-Markovian noise, *PRX Quantum* **2**, 040351 (2021).
- [23] Y.-Q. Chen, K.-L. Ma, Y.-C. Zheng, J. Allcock, S. Zhang, and C.-Y. Hsieh, Non-Markovian Noise Characterization with the Transfer Tensor Method, *Phys. Rev. Applied* **13**, 034045 (2020).
- [24] L. Bianchi, E. Grant, A. Rocchetto, and S. Severini, Modelling non-Markovian quantum processes with recurrent neural networks, *New J. Phys.* **20**, 123030 (2018).
- [25] I. A. Luchnikov, S. V. Vintskevich, D. A. Grigoriev, and S. N. Filippov, Machine Learning Non-Markovian Quantum Dynamics, *Phys. Rev. Lett.* **124**, 140502 (2020).
- [26] S. Krastanov, K. Head-Marsden, S. Zhou, S. T. Flammia, L. Jiang, and P. Narang, Unboxing quantum black box models: Learning non-Markovian dynamics, [arXiv:2009.03902](https://arxiv.org/abs/2009.03902).
- [27] S. Milz, F. A. Pollock, and K. Modi, An introduction to operational quantum dynamics, *Open Syst. Inf. Dyn.* **24**, 1740016 (2017).
- [28] H.-P. Breuer, and F. Petruccione, *The Theory of Open Quantum Systems* (Oxford University Press, Oxford, UK, 2002).
- [29] I. Bengtsson and K. Życzkowski, *Geometry of Quantum States: An Introduction to Quantum Entanglement* (Cambridge University Press, Cambridge, UK, 2017).
- [30] E. Farhi, J. Goldstone, and S. Gutmann, A quantum approximate optimization algorithm, [arXiv:1411.4028](https://arxiv.org/abs/1411.4028).
- [31] T. Hogg, Quantum search heuristics, *Phys. Rev. A* **61**, 052311 (2000).
- [32] 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).
- [33] J. Marshall, F. Wudarski, S. Hadfield, and T. Hogg, Characterizing local noise in QAOA circuits, *IOP SciNotes* **1**, 025208 (2020).
- [34] C. Xue, Z.-Y. Chen, Y.-C. Wu, and G.-P. Guo, Effects of Quantum Noise on Quantum Approximate Optimization Algorithm, *Chin. Phys. Lett.* **38**, 030302 (2021).
- [35] 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).
- [36] M. Streif, M. Leib, F. Wudarski, E. Rieffel, and Z. Wang, Quantum algorithms with local particle-number conservation: Noise effects and error correction, *Phys. Rev. A* **103**, 042412 (2021).
- [37] M. Streif and M. Leib, Training the quantum approximate optimization algorithm without access to a quantum processing unit, *Quantum Sci. Technol.* **5**, 034008 (2020).
- [38] 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).
- [39] R. Shaydulin, I. Safro, and J. Larson, Multistart methods for quantum approximate optimization, in *2019 IEEE High Performance Extreme Computing Conference (HPEC)* (IEEE, Piscataway, NJ, 2019), pp. 1–8.
- [40] F. G. S. L. Brandao, M. Broughton, E. Farhi, S. Gutmann, and H. Neven, For fixed control parameters the quantum approximate optimization algorithm’s objective function value concentrates for typical instances, [arXiv:1812.04170](https://arxiv.org/abs/1812.04170).
- [41] S. H. Sack and M. Serbyn, Quantum annealing initialization of the quantum approximate optimization algorithm, *Quantum* **5**, 491 (2021).
- [42] 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).
- [43] F. Glover, G. Kochenberger, and Y. Du, A tutorial on formulating and using QUBO models, [arXiv:1811.11538](https://arxiv.org/abs/1811.11538).
- [44] R. Shaydulin, S. Hadfield, T. Hogg, and I. Safro, Classical symmetries and the quantum approximate optimization algorithm, *Quantum Inf. Process.* **20**, 359 (2021).
- [45] T. Peng, A. W. Harrow, M. Ozols, and X. Wu, Simulating Large Quantum Circuits on a Small Quantum Computer, *Phys. Rev. Lett.* **125**, 150504 (2020).
- [46] A. Ash-Saki, M. Alam, and S. Ghosh, Experimental characterization, modeling, and analysis of crosstalk in a quantum computer, *IEEE Transactions on Quantum Engineering* **1**, 3101406 (2020).
- [47] M. Sarovar, T. Proctor, K. Rudinger, K. Young, E. Nielsen, and R. Blume-Kohout, Detecting crosstalk errors in quantum information processors, *Quantum* **4**, 321 (2020).
- [48] D. M. Greenberger, M. A. Horne, and A. Zeilinger, Going beyond Bell’s theorem, in *Bell’s Theorem, Quantum Theory and Conceptions of the Universe. Fundamental Theories of Physics*, edited by M. Kafatos (Springer, Dordrecht, 1989), Vol. 37.
- [49] K. Życzkowski and H.-J. Sommers, Average fidelity between random quantum states, *Phys. Rev. A* **71**, 032313 (2005).
- [50] S. T. Flammia and J. J. Wallman, Efficient estimation of Pauli channels, *ACM Trans. Quantum Comput.* **1**, 1 (2020).

- [51] J. R. Johansson, P. D. Nation, and F. Nori, QuTiP: An open-source Python framework for the dynamics of open quantum systems, *Comput. Phys. Commun.* **183**, 1760 (2012).
- [52] A. Hurwitz, über die erzeugung der invarianten durch integration, *Nachr. Ges. Wiss. Goettingen, Math.-Phys. Kl.* **1897**, 71 (1897).
- [53] K. Życzkowski and H.-J. Sommers, Induced measures in the space of mixed quantum states, *J. Phys. A: Math. Gen.* **34**, 7111 (2001).
- [54] D. Kraft *et al.*, *A Software Package for Sequential Quadratic Programming* (DFVLR Obersfaffeuhofen, Germany, 1988).
- [55] S. Mandrà, J. Marshall, E. G. Rieffel, and R. Biswas, HybridQ: A hybrid simulator for quantum circuits, *Second International Workshop on Quantum Computing Software (QCS)* (IEEE, Piscataway, NJ, 2021).
- [56] D. M. Abrams, N. Didier, B. R. Johnson, M. P. da Silva, and C. A. Ryan, Implementation of XY entangling gates with a single calibrated pulse, *Nat. Electron.* **3**, 744 (2020).
- [57] Rigetti quantum computing, <https://www.rigetti.com/>.