Dual Map Framework for Noise Characterization of Quantum Computers

James Sud,[1, 2, *] Jeffrey Marshall,[1, 2] Zhihui Wang,[1, 2] Eleanor Rieffel,[1] and Filip A. Wudarski[1, 2, †]

[1]*Quantum Artificial Intelligence Laboratory (QuAIL),*
NASA Ames Research Center, Moffett Field, CA, 94035, USA
[2]*USRA Research Institute for Advanced Computer Science (RIACS), Mountain View, CA, 94043, USA*
(Dated: December 9, 2021)

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 non-local errors that occur during a circuit. Finally, we demonstrate the performance of the method on Rigetti's Aspen-9 quantum computer for QAOA circuits up to six qubits, successfully predicting the observed measurements on a majority of the qubits.

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 algorithms 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 such as [15–17]. So far, benchmarking and characterization methods have suffered substantial shortcomings - either returning limited information (e.g. average fidelity for RB) or restricted to small systems due to exponential scaling (tomographic methods). In this work, we develop a characterization scheme that efficiently returns information about process matrix of the marginal noise channel acting on a single qubit. The method combines ease and efficiency of benchmarking techniques with substan-tially richer information content. Additionally, the introduced protocol operates without additional compilation overhead, as opposed to RB approaches, which require twirling subroutine to cast the noisy channel into a convenient form of a Pauli channel.

Quantum noise, which can lead to computational errors, are inevitable companions 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 [18–20]. For our purposes, the most natural representation is of the following form

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

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 χ matrix via quantum process tomography [11]. In order to represent a valid quantum channel, Eq. (1) has to be CPTP, which happens when $\chi \geq 0$ (CP condition) and

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

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) [21–23], which aims to find approximately optimal solutions to optimization problems.

---
[*] jsud@usra.edu
[†] fwudarski@usra.edu

In this work, we study the application of our characterization method for QAOA run on combinatorial optimization problems. The characterization and effect of local noise in QAOA circuits has been previously studied [24–27], 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 to 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 V we demonstrate the efficacy of the method in characterizing error on classical simulations, as well as on the Aspen-9 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

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

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

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

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, and, as explained above, the simulation of which 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 parameterized unitaries of a mixing Hamiltonian B and a phasing Hamiltonian H, as so

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

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 $(\vec{\gamma}, \vec{\beta})$ [28–31] as well as choosing optimal starting states and mixing Hamiltonians [23, 32, 33] 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

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

Many popular combinatorial optimization problems can be cast into QUBO form [34]. 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 parameterized 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

$$\rho(\vec{\gamma}, \vec{\beta}) = \mathcal{E}\mathcal{U}(\vec{\gamma}, \vec{\beta}) \left[|\psi_0\rangle\langle\psi_0|\right]. \qquad (7)$$
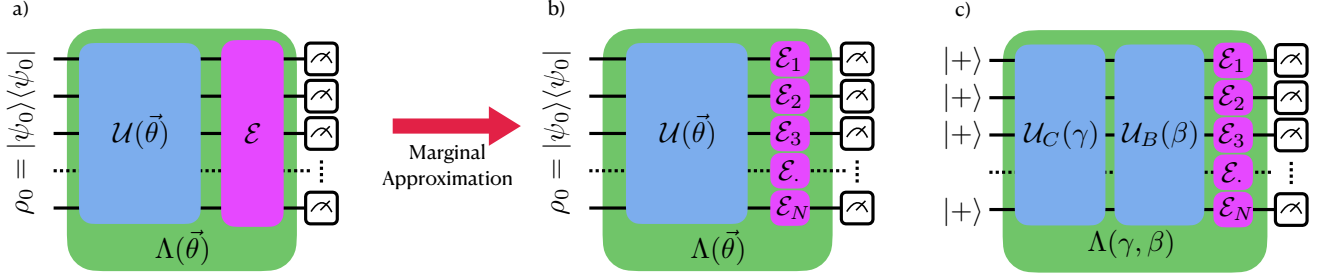
Figure 1: Operational framework for the noisy circuit characterization that is described by a quantum map $\Lambda(\vec{\theta})$, where $\vec{\theta}$ is a collection of circuit parameters. In a) noise is "moved" to the last layer (according to $\mathcal{E} = \Lambda(\vec{\theta}) \circ \mathcal{U}^{\dagger}(\vec{\theta})$, and in general it is of non-local character. The non-locality is neglected in b) by approximating circuit with only local quantum channels $\mathcal{E}_i$ acting on i-th qubit. In c) we depict 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 [24, 25].

In the following subsections, we analyze the effects of the dual map of various common error channels on Hamiltonians of form Eq. (6) wit 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, parameterized by a depolarization rate p, as

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

with $\sigma_1, \sigma_2, \sigma_3$ 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 noise channel acting on each qubit, i.e. $\mathcal{E}_p = \mathcal{E}_p^1 \otimes \mathcal{E}_p^2 \otimes \ldots \otimes \mathcal{E}_p^n$, 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 Hamiltonian, and we assume that p is constant on all qubits, $\mathcal{E}_p$ effectively multiplies this term by p^k. Specifically, we have

$$\mathcal{E}_p^\#(Z_i) = pZ_i, \qquad (9)$$
$$\mathcal{E}_p^\#(Z_i Z_j) = p^2 Z_i Z_j. \qquad (10)$$

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

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

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

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

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

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

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

$$\mathcal{E}_\gamma^\#(Z_i) = (1-\gamma)Z_i + \gamma I, \qquad (14)$$
$$\mathcal{E}_\gamma^\#(Z_i Z_j) = (1-\gamma)^2 Z_i Z_j + \gamma(1-\gamma)(Z_i + Z_j) + \gamma^2 I. \qquad (15)$$

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)

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

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, thus all single qubit Z terms go to zero [35], 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):

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

2. d-regular graph, all h's constant, all nonzero J's are constant (all equal h, J, respectively):

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

3. max-k-colorable-subgraph [33] ($h_i = d_i$ where d_i is degree of vertex i, $J_{ij} = -1$ if edge (i,j) exists in the graph):

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

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 overrotations, etc), please see Appendix A.

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 marginal approximation to the effective noise (MATEN), which is defined as follows

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

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

where $\mathrm{Tr}_{\bar{k}}(\mathcal{E}) = \mathcal{E}_k$ traces out all subsystem except k-th (see Fig. 1 b)).

In order to determine a MATEN, we express a noisy map in terms of 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

$$\tilde{I} = I, \quad (24)$$
$$\tilde{X} = (p_0 + p_1 - p_2 - p_3)X + 4t_{01}I$$
$$+ 2((t_{12} - v_{03})Y + (t_{13} + v_{02}))Z, \quad (25)$$
$$\tilde{Y} = (p_0 + p_2 - p_1 - p_3)Y + 4t_{02}I$$
$$+ 2((t_{23} - v_{01})Z + (t_{12} + v_{03}))X, \quad (26)$$
$$\tilde{Z} = (p_0 + p_3 - p_1 - p_2)Z + 4t_{03}I$$
$$+ 2((t_{13} - v_{02})X + (t_{23} + v_{01}))Y, \quad (27)$$

where $\tilde{I}, \tilde{X}, \tilde{Y}, \tilde{Z}$ represent the noisy transformations of the Pauli operators, and $\tilde{I} = I$ due to the property that the dual of trace preserving maps are unital (i.e.

$\mathcal{E}^{\#}(I) = I$). We can rewrite the coefficients in each equations 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 parameterized circuit of interest[36]:

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

2. Implement the circuit on a quantum device, 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.

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 [37]. Additionally, if the noise channels *mildly* depend on circuits [38], 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, 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 Max-Cut, 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.

IV. LOCAL VS NON-LOCAL 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 [39, 40]) 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 non-local noise channels with the MATEN approach. 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 non-local channels, iii) finally we repeat numerical analysis from ii), but for random Pauli channels and analyze some scaling properties.

Since single-qubit χ matrix (in Pauli basis) is a positive operator of trace one, we can treat it as a 4-dimensional 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 non-local 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 χ matrix is translated to

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

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 non-local process matrix χ as the maximally entangled state, that is defined as a projector onto $|\Psi\rangle = \frac{1}{2}\sum_{i=0}^{3}|ii\ldots i\rangle$, with N 4D subsystems (each corresponding to a qubit), note that this is a GHZ state [41]. 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 non-trivial 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
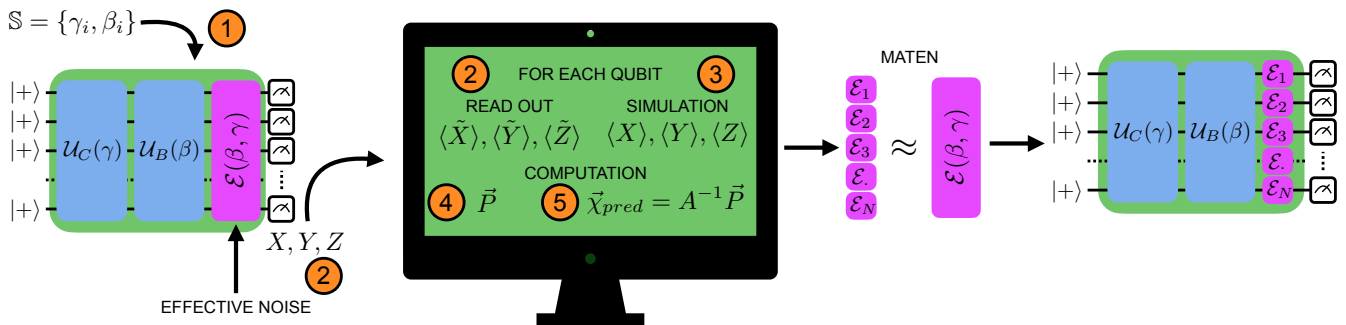
Figure 2: Protocol for characterizing local approximation to effective noise in the case of a single-layered QAOA algorithm. 1) run multiple times a circuit with different parameters from the set $\mathbb{S}$, next 2) measure all qubit registers in $X, Y,$ and Z bases. Use output bitstrings to infer (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 $\vec{\chi}_{pred}$. Perform this procedure for each qubit in order to reconstruct the MATEN that approximates effective noise.

yields the minimum fidelity value between the marginal approximation (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 non-local χ matrix and its MA. Since, the MA gives a trivial state, one can easily compute the fidelity [4, 42]

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

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 multi-qubit gates, and higher for single-qubit ones). Therefore, we can escape this unfavorable scaling by restricting to channels that are close to perfect (noiseless) case, i.e. to the identity channel ($\chi_{00} = 1$ and all other elements equal to zero). This also implies that non-local 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 [43]), since they represent noise regimes that are more relevant for the current hardware technology and help tailor error correcting schemes. In order to properly address this issue, we incorporate numerical methods to find out how well the MA can represent the true non-local noise process. Here, we use random sampling of full χ matrices and random samples of Pauli channels (i.e. χ matrices with a random proba-

bility 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 [44]), and computed all marginals of the multi-qubit 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 non-local 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 [45]. 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 [46, 47]

$$\chi = \text{diag}[\cos^2(\theta_{4^N-1}), \cos^2(\theta_{4^N-2})\sin^2(\theta_{4^N-1}), \ldots,$$
$$\sin^2(\theta_1)\sin^2(\theta_2), \cdots \sin^2(\theta_{4^N-1})] \qquad (30)$$

We employ Sequential Least Squares Programming (SLSQP) [48] 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/circuit fidelity, the MA can provide results with acceptable accuracy. Therefore, the MATEN protocol, identifies a MA that can estimate the leading order of the effective noise channel.

V. RESULTS

In this section we present the success of the method presented in III for noisy simulations.
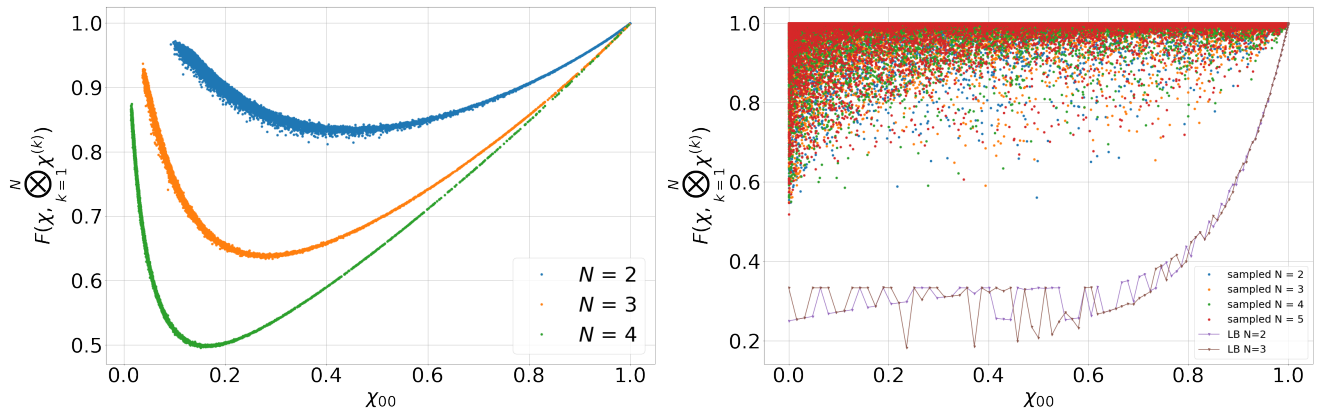
Figure 3: Comparison between random non-local χ matrix and its MA in terms of fidelity. Left plot depicts case of full random channels, while right plot restricts the analysis to random Pauli channels. Additionally, we provide a numerical lower bound (LB) on Pauli channels with two and three qubits (right plot).

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 [49]. In some cases, we additionally generate and apply error channels via QuTiP [44] 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 parameterized 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 $\vec{p}$, and $[-1, 1]$ for $\vec{t}$ and $\vec{v}$ 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 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 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 statevector simulations for perfect evaluation of observables.
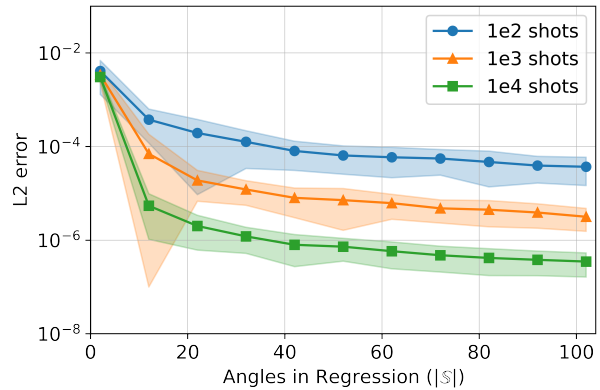

Figure 4: Average L2 distance between the true and predicted χ for classical simulations for randomly chosen two qubit QUBO instances with weights in the range $[0, 1]$, as a function of the number of angles $|\mathbb{S}|$ used in the regression and the number of shots used in the estimation of expectation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

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 $.800 \pm .125$, but we see fidelities much higher than this value for sufficiently large $|\mathbb{S}|$ and number of shots.

2. Non-Local 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 non-local noise, then, we apply an error map of the form

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

where we have a combination of purely local channel $\mathcal{E}_1^{(n)}$ and 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 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 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 non-local 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 there we are not injecting any purely local noise to 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 0 and standard deviation of $1/\sqrt{numshots}$ 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 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 in causes expectation values to fluctuate between measurements, thus essentially introducing a non-constant 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 non-constant between different angles in a similar manner to sampling noise. Errors can be extremely angle-dependent on quantum computers, especially for parameterized two qubit gates such as [50], 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 non-zero 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

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

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

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

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 A 3. This model applies the two-qubit dephasing noise layer (Eq. (32)) on each pair qubits the phase gates act, after
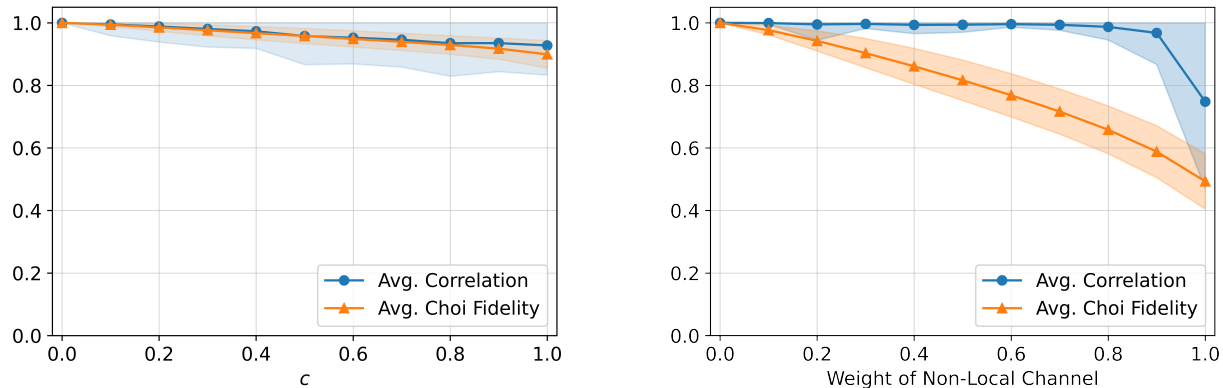
Figure 5: Average testing correlations and full state fidelities between the actual noise model and predicted noise model for two (left) and four (right) qubit fully connected QUBO problems with all $J = 1$ and all $h = 0$, as a function of the weight of the applied non-local channel, or c in Eq. (31). Solid lines depict the average over 100 runs, shading depicts one standard deviation above and below the average.
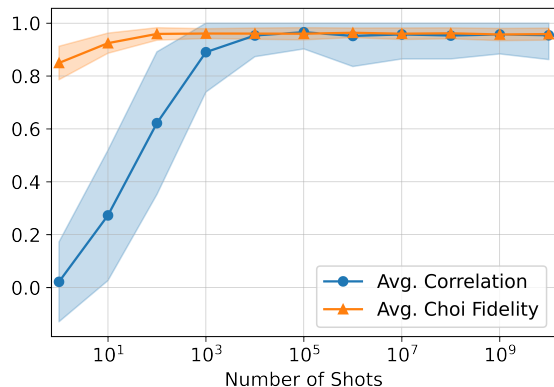

Figure 6: Average testing correlations and Choi fidelities between the actual noise model and predicted noise model for two qubit QUBO problem with $J = 1$ and both $h = 0$, as a function of the number of shots used to estimate expectation values. Solid lines depict the average over 100 runs, shading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

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 10-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 1-qubit correlations, so we chose to add in the average of all 2-qubit correlations as well. Additionally, we report the average fidelity between the actual 10-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 1 and 2-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 $\omega = .5$, and then follow a symmetric pattern once $\omega > .5$. 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 approximation 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'S ASPEN-9 DEVICE

In this section we apply the error characterization method from Sec. III to the Aspen-9 Quantum Processing Unit (QPU) from Rigetti Computing [51]. We run the characterization procedure for QAOA circuits with phase separation given by Hamiltonians of the form in Eq. (6), with all $h_i = 1$ (to break $\mathbb{Z}_2$ symmetry) and $J_{ij} = \delta_{i+1,j}$ (forming a line topology), and implemented using a single CPHASE(γ) gate, and with mixing via the standard X-mixer. These experiments were run at $N = 2$ and $N = 6$ with $|\mathbb{S}| = 100$, where N is the number of qubits and $|\mathbb{S}|$ 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 0). The intention of this is to make sure that our method works
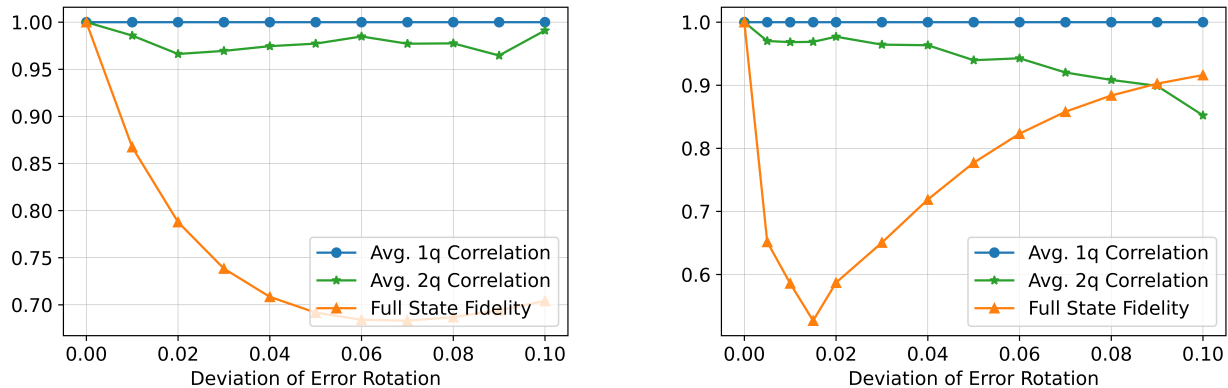
Figure 7: Average testing correlations and Choi fidelities between the actual noise model and predicted noise model for 10-qubit ring (right) and fully-connected (right) QUBO problems with all $J = 1$ and all $h = 0$, as a function of the deviation ω of both the phasing and mixing operators.

when only single qubit gates are present, removing main sources of crosstalk and non-local 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 0 (again equivalent to setting all J_{ij} to 0). 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 0, 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 Pauli expectation values. These are shown for both the two and six qubit cases in table I.

qubit	1q only	2q idle	2q active
34	0.9964	0.5154	0.9622
35	0.9982	0.7758	0.9704

Table I: Correlations for the method performed on QAOA circuits for qubits 34 and 35 on Rigetti Aspen-9 device. The low values for the 2q idle case are likely explained by day-to-day changes in qubit calibration and error sources on the device.

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. We note that there is low fidelity for the "2q idle" case. This is likely explained by the fact that this case was run a few days after the other experiments, as this experiment idea was conceived after running the "1q only" and "2q active" cases. Due to day-to-day changes in calibration,

qubits 34 and 35 may have experienced calibration issues on the day of running. Unfortunately, the Aspen-9 device was de-commissioned before a re-run of the experiment began. This faulty qubit can result in elevated angle-dependent, nonlocal, or inter-circuit noise, which have the potential to reduce the accuracy of a MATEN.

The six-qubit experiments are presented in Table II.

qubit	1q only	2q idle	2q active
30	0.9963	0.9254	0.5477
31	0.9901	0.9041	0.6692
32	0.9923	0.6584	0.5363
33	0.9948	0.0480	0.0449
34	0.9936	0.8674	0.7855
35	0.9985	0.9908	0.9801

Table II: Correlations for the method performed on QAOA circuits for qubits 30-35 on Rigetti Aspen-9 device.

Here, we see that all metrics remain high for the "1q only" case, but for the "2q idle" case for qubits 32 and 33 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, 31, 32, and 34. 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 $.69 \pm .30$. These values are far from the ideal values of 1, but the positive correlation values suggest that the method approximately captures the dominant error channels present on the QPU. The wide variability in performance on various qubits suggest that certain qubits may have more angle-dependent noise, or may have larger sources of crosstalk, as analyzed in Sec. V A. In particular, qubits 32 and 33 experience a sharp decline in correlations in both the "2q idle/active" cases, with the correlations of qubit 33

plummeting to roughly .05. The correlations on qubit 33 of roughly .05 are additionally much lower than we see even on the right 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 to zero, may introduce some significant crosstalk between the two qubits that is far from the intended phasing operation, which the MATEN is not equipped to accurately handle. In the simulations we perform, artificially added errors come in the form of randomly chosen Kraus maps or overrotations, but the error maps on a quantum device may be of a specific, more detrimental for. Additionally, even with a two-qubit gate with angle set to 0, 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.

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 presented a method to compute a marginal approximation to the effective noise (MATEN) of a parameterized quantum circuit, that is efficient in terms of 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. We finally show that the method is effective in computing a MATEN on a few qubits of Rigetti's Aspen-9 quantum computer. Lower values in extracted correlations of expectation values can be inform us that the system exhibits a fair amount of angle dependant (gate) noise, as well as errors that are absent in the theoretical model, e.g. readout or leakage to the non-computational sub-

space. The latter can be modelled 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 dominant sources of error, which can in turn point to particularly 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 dual map framework introduced 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 non-trivial 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 non-correlated. 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.

VIII.   ACKNOWLEDGEMENTS

[1] Christophe Piveteau, David Sutter, Sergey Bravyi, Jay M. Gambetta, and Kristan Temme, "Error mitigation for universal gates on encoded qubits," Phys. Rev. Lett. 127, 200505 (2021).

[2] Kristan Temme, Sergey Bravyi, and Jay M. Gambetta, "Error mitigation for short-depth quantum circuits," Phys. Rev. Lett. 119, 180509 (2017).

[3] Ying Li and Simon C. Benjamin, "Efficient Variational Quantum Simulator Incorporating Active Error Mini-

mization," Phys. Rev. X 7, 021050 (2017).

[4] Michael A Nielsen, "A simple formula for the average gate fidelity of a quantum dynamical operation," Physics Letters A 303, 249–252 (2002).

[5] Filip Wudarski, Jeffrey Marshall, Andre Petukhov, and Eleanor Rieffel, "Augmented fidelities for single-qubit gates," Phys. Rev. A 102, 052612 (2020).

[6] Joseph Emerson, Robert Alicki, and Karol \.Zyczkowski, "Scalable noise estimation with random unitary opera-

tors," **7**, S347–S352 (2005), publisher: IOP Publishing.

[7] Easwar Magesan, J. M. Gambetta, and Joseph Emerson, "Scalable and Robust Randomized Benchmarking of Quantum Processes," Physical Review Letters **106**, 180504 (2011), publisher: American Physical Society.

[8] Jahan Claes, Eleanor Rieffel, and Zhihui Wang, "Character Randomized Benchmarking for Non-Multiplicity-Free Groups With Applications to Subspace, Leakage, and Matchgate Randomized Benchmarking," PRX Quantum **2**, 010351 (2021), publisher: American Physical Society.

[9] Alexander Erhard, Joel J. Wallman, Lukas Postler, Michael Meth, Roman Stricker, Esteban A. Martinez, Philipp Schindler, Thomas Monz, Joseph Emerson, and Rainer Blatt, "Characterizing large-scale quantum computers via cycle benchmarking," Nature Communications **10**, 5347 (2019).

[10] Steven T. Flammia and Yi-Kai Liu, "Direct Fidelity Estimation from Few Pauli Measurements," Physical Review Letters **106**, 230501 (2011), publisher: American Physical Society.

[11] Isaac L. Chuang and M. A. Nielsen, "Prescription for experimental determination of the dynamics of a quantum black box," Journal of Modern Optics **44**, 2455–2467 (1997).

[12] Daniel Greenbaum, "Introduction to Quantum Gate Set Tomography," arXiv:1509.02921 [quant-ph] (2015), arXiv: 1509.02921.

[13] S. G. Schirmer, A. Kolli, and D. K. L. Oi, "Experimental Hamiltonian identification for controlled two-level systems," Physical Review A **69**, 050306 (2004), publisher: American Physical Society.

[14] Shelby Kimmel, Guang Hao Low, and Theodore J. Yoder, "Robust calibration of a universal single-qubit gate set via robust phase estimation," Physical Review A **92**, 062315 (2015), publisher: American Physical Society.

[15] Mingyu Sun and Michael R. Geller, "Efficient characterization of correlated SPAM errors," arXiv:1810.10523 [quant-ph] (2020), arXiv: 1810.10523.

[16] Junan Lin, Joel J. Wallman, Ian Hincks, and Raymond Laflamme, "Independent state and measurement characterization for quantum computers," Physical Review Research **3**, 033285 (2021), publisher: American Physical Society.

[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), publisher: American Physical Society.

[18] Simon Milz, Felix A. Pollock, and Kavan Modi, "An introduction to operational quantum dynamics," Open Systems & Information Dynamics **24**, 1740016 (2017).

[19] Heinz-Peter Breuer, Francesco Petruccione, *et al.*, *The theory of open quantum systems* (Oxford University Press on Demand, 2002).

[20] Ingemar Bengtsson and Karol Życzkowski, *Geometry of quantum states: an introduction to quantum entanglement* (Cambridge university press, 2017).

[21] Edward Farhi, Jeffrey Goldstone, and Sam Gutmann, "A Quantum Approximate Optimization Algorithm," arXiv:1411.4028 [quant-ph] (2014), arXiv: 1411.4028.

[22] Tad Hogg, "Quantum search heuristics," Physical Review A **61**, 052311 (2000), publisher: American Physical Society.

[23] Stuart Hadfield, Zhihui Wang, Bryan O'Gorman, Eleanor G. Rieffel, Davide Venturelli, and Rupak Biswas, "From the Quantum Approximate Optimization Algorithm to a Quantum Alternating Operator Ansatz," Algorithms **12**, 34 (2019), number: 2 Publisher: Multidisciplinary Digital Publishing Institute.

[24] Jeffrey Marshall, Filip Wudarski, Stuart Hadfield, and Tad Hogg, "Characterizing local noise in QAOA circuits," **1**, 025208 (2020), publisher: IOP Publishing.

[25] Cheng Xue, Zhao-Yun Chen, Yu-Chun Wu, and Guo-Ping Guo, "Effects of Quantum Noise on Quantum Approximate Optimization Algorithm," **38**, 030302 (2021), publisher: IOP Publishing.

[26] Samson Wang, Enrico Fontana, M. Cerezo, Kunal Sharma, Akira Sone, Lukasz Cincio, and Patrick J. Coles, "Noise-induced barren plateaus in variational quantum algorithms," (2021), arXiv:2007.14384 [quant-ph].

[27] Michael Streif, Martin Leib, Filip Wudarski, Eleanor Rieffel, and Zhihui Wang, "Quantum algorithms with local particle-number conservation: Noise effects and error correction," Physical Review A **103** (2021), 10.1103/physreva.103.042412.

[28] Michael Streif and Martin Leib, "Training the quantum approximate optimization algorithm without access to a quantum processing unit," **5**, 034008 (2020), publisher: IOP Publishing.

[29] Leo Zhou, Sheng-Tao Wang, Soonwon Choi, Hannes Pichler, and Mikhail D. Lukin, "Quantum Approximate Optimization Algorithm: Performance, Mechanism, and Implementation on Near-Term Devices," Physical Review X **10**, 021067 (2020), publisher: American Physical Society.

[30] Ruslan Shaydulin, Ilya Safro, and Jeffrey Larson, "Multistart Methods for Quantum Approximate optimization," in *2019 IEEE High Performance Extreme Computing Conference (HPEC)* (2019) pp. 1–8, iSSN: 2643-1971.

[31] Fernando G. S. L. Brandao, Michael Broughton, Edward Farhi, Sam Gutmann, and Hartmut Neven, "For Fixed Control Parameters the Quantum Approximate Optimization Algorithm's Objective Function Value Concentrates for Typical Instances," arXiv:1812.04170 [quant-ph] (2018), arXiv: 1812.04170.

[32] Stefan H. Sack and Maksym Serbyn, "Quantum annealing initialization of the quantum approximate optimization algorithm," Quantum **5**, 491 (2021), publisher: Verein zur Förderung des Open Access Publizierens in den Quantenwissenschaften.

[33] Zhihui Wang, Nicholas C. Rubin, Jason M. Dominy, and Eleanor G. Rieffel, "XY mixers: Analytical and numerical results for the quantum alternating operator ansatz," Physical Review A **101**, 012320 (2020), publisher: American Physical Society.

[34] Fred Glover, Gary Kochenberger, and Yu Du, "A Tutorial on Formulating and Using QUBO Models," arXiv:1811.11538 [quant-ph] (2019), arXiv: 1811.11538.

[35] Ruslan Shaydulin, Stuart Hadfield, Tad Hogg, and Ilya Safro, "Classical symmetries and the Quantum Approximate Optimization Algorithm," Quantum Information Processing **20**, 359 (2021), arXiv: 2012.04713.

[36] 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 algo-

rithm, and could serve only as a characterization protocol.

[37] Tianyi Peng, Aram W. Harrow, Maris Ozols, and Xiaodi Wu, "Simulating Large Quantum Circuits on a Small Quantum Computer," Physical Review Letters **125**, 150504 (2020), publisher: American Physical Society.

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

[39] Abdullah Ash-Saki, Mahabubul Alam, and Swaroop Ghosh, "Experimental Characterization, Modeling, and Analysis of Crosstalk in a Quantum Computer," IEEE Transactions on Quantum Engineering **1**, 1–6 (2020), conference Name: IEEE Transactions on Quantum Engineering.

[40] Mohan Sarovar, Timothy Proctor, Kenneth Rudinger, Kevin Young, Erik Nielsen, and Robin Blume-Kohout, "Detecting crosstalk errors in quantum information processors," Quantum **4**, 321 (2020), publisher: Verein zur Förderung des Open Access Publizierens in den Quantenwissenschaften.

[41] Daniel M. Greenberger, Michael A. Horne, and Anton Zeilinger, "Going beyond bell's theorem," (2007), arXiv:0712.0921 [quant-ph].

[42] Karol Życzkowski and Hans-Jürgen Sommers, "Average fidelity between random quantum states," Physical Review A **71** (2005), 10.1103/physreva.71.032313.

[43] Steven T. Flammia and Joel J. Wallman, "Efficient estimation of pauli channels," ACM Transactions on Quantum Computing **1**, 1–32 (2020).

[44] J. R. Johansson, P. D. Nation, and Franco Nori, "QuTiP: An open-source Python framework for the dynamics of open quantum systems," Computer Physics Communications **183**, 1760–1772 (2012), arXiv: 1110.0573.

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

[46] A. Hurwitz, "über die erzeugung der invarianten durch integration," Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse **1897**, 71–2 (1897).

[47] Karol Życzkowski and Hans-Jürgen Sommers, "Induced measures in the space of mixed quantum states," Journal of Physics A: Mathematical and General **34**, 7111–7125 (2001).

[48] Dieter Kraft *et al.*, "A software package for sequential quadratic programming," (1988).

[49] Salvatore Mandrà, Jeffrey Marshall, Eleanor G. Rieffel, and Rupak Biswas, "HybridQ: A Hybrid Simulator for Quantum Circuits," (2021), arXiv:2111.06868 [quant-ph].

[50] Deanna M. Abrams, Nicolas Didier, Blake R. Johnson, Marcus P. da Silva, and Colm A. Ryan, "Implementation of XY entangling gates with a single calibrated pulse," Nature Electronics **3**, 744–750 (2020).

[51] "Rigetti quantum computing," Https://www.rigetti.com/.

Appendix A: 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 section 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 0 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

$$\langle \Psi | \mathcal{O} | \Psi \rangle = \langle \Psi | \prod_{i \in S_Y} Z_i \prod_{j \in S_Y} Y_j \prod_{k \in S_X} X_k | \Psi \rangle . \tag{A1}$$

Then noting that Y = -iZX and write

$$-i \langle \Psi | \prod_{i \in S_Z} Z_i \prod_{j \in S_Y} Z_j X_j \prod_{k \in S_X} X_k | \Psi \rangle = -i \langle \Psi | \prod_{i \in S_Z \cup S_Y} Z_i \prod_{k \in S_X \cup S_Y} X_k | \Psi \rangle . \tag{A2}$$

We can then expand out $| \Psi \rangle$ in terms of bitstrings l

$$-i \langle \Psi | \prod_{i \in S_Z \cup S_Y} Z_i \prod_{k \in S_X \cup S_Y} X_k \sum_l c_l (|l\rangle + |\bar{l}\rangle). \tag{A3}$$

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

$$-i \langle \Psi | \prod_{i \in S_Z \cup S_Y} Z_i \sum_l c_l (|l^*\rangle + |\bar{l^*}\rangle) \tag{A4}$$

$$-i\sum_l c_l(\langle\Psi|\prod_{i\in S_Z\cup S_Y}Z_i\,|l^*\rangle + (\langle\Psi|\prod_{i\in S_Z\cup S_Y}Z_i\,|\overline{l^*}\rangle)) \tag{A5}$$

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

$$-i\sum_l c_l(\langle\Psi|\prod_{i\in S_Z\cup S_Y}Z_i\,|l^*\rangle + (\langle\Psi|\prod_{i\in S_Z\cup S_Y}(-Z_i)\,|\overline{l^*}\rangle\,\langle l^*|\,(-Z)\,|l^*\rangle). \tag{A6}$$

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

$$-i\sum_l c_l(\langle\Psi|\prod_{i\in S_Z\cup S_Y}Z_i\,|l^*\rangle + (-1)^{S_Y+S_Z}\langle\Psi|\prod_{i\in S_Z\cup S_Y}|\overline{l^*}\rangle\,\langle l^*|\,(-Z_i)\,|l^*\rangle) \tag{A7}$$

$$= -i\sum_l c_l(\langle\Psi|\prod_{i\in S_Z\cup S_Y}Z_i\,|l^*\rangle + (-1)^{|S_Y|+|S_Z|}\langle\Psi|\prod_{i\in S_Z\cup S_Y}|\overline{l^*}\rangle\,\langle l^*|\,Z_i\,|l^*\rangle) \tag{A8}$$

$$= -i\sum_l c_l(\langle\Psi|\prod_{i\in S_Z\cup S_Y}Z_i\,|l^*\rangle + (-1)^{|S_Y|+|S_Z|}\langle\Psi|\prod_{i\in S_Z\cup S_Y}|\overline{l^*}\rangle\,\langle l^*|\,Z_i\,|l^*\rangle) \tag{A9}$$

$$= -i\sum_l c_l(\langle\Psi|\prod_{i\in S_Z\cup S_Y}Z_i\,|l^*\rangle + (-1)^{|S_Y|+|S_Z|}\langle\Psi|\prod_{i\in S_Z\cup S_Y}Z_i\,|l^*\rangle) \tag{A10}$$

$$= -i\sum_l c_l(1+(-1)^{|S_Y|+|S_Z|})\,\langle\Psi|\prod_{i\in S_Z\cup S_Y}Z_i\,|l^*\rangle. \tag{A11}$$

Where we use the first property again in the second to last step. Now we can clearly see that if there 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

$$\mathcal{E}^{\#}(Z) = p_Z Z + p_Y Y + p_X X + p_I I, \tag{A12}$$

$$\begin{aligned}\mathcal{E}^{\#}(ZZ) = &\ p_Z^2 ZZ + p_Y^2 YY + p_X^2 XX + p_I^2 I + \\ &+ P_Z P_Y(ZY + YZ) + P_Z P_X(ZX + XZ) \\ &+ P_Z P_I(ZI + IZ) + P_Y P_X(YX + XY) \\ &+ P_Y P_I(YI + IY) + P_X P_I(XI + IX).\end{aligned} \tag{A13}$$

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 form given in Eq. (6) is:

$$\begin{aligned}H' = &\sum_i h_i(p_Z Z_i + p_Y Y_i + p_X X_i + P_I I) \tag{A14}\\ &+ \sum_{i<j} J_{ij}(p_Z^2 Z_i Z_j + p_Y^2 Y_i Y_j + p_X^2 X_i X_j + p_I^2 I) \\ &+ \sum_{i,j} J_{i,j}(P_Z P_Y Z_i Y_j + P_Z P_X Z_i X_j + P_Z P_I Z_i I \\ &+ P_Y P_X Y_i X_j + P_Y P_I Y_i I + P_X P_I X_i I). \tag{A15}\end{aligned}$$

From here various assumptions can be made. If all h's=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 Thm.1 in the 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

$$H' = p_Z^2 H + p_Z p_Y \sum_{i,j} J_{ij} Z_i Y_j. \tag{A16}$$

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

2. Constant Mixing Overrotations

Take a very simple model, where the phase is applied correctly, but the mixer applied as $H_M = \sum_i \tilde{\beta}_i X_i$ where $\beta_i = \beta + \delta\beta_i$, i.e. a small over/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, $H' = U'^\dagger_M H U'_M$. Now use $e^{-i\theta X} Z e^{i\theta X} = \cos(2\theta)Z - \sin(2\theta)Y$.
Assuming the standard form $H = \sum_{i<j} J_{ij} Z_i Z_j$,

$$\begin{aligned}
H' &= \sum_{i<j} J_{ij} \cos(2\delta\beta_i) \cos(2\delta\beta_j) Z_i Z_j + \sum_{i<j} J_{ij} \cos(2\delta\beta_i) \sin(2\delta\beta_j) Z_i Y_j \\
&+ \sum_{i<j} J_{ij} \sin(2\delta\beta_i \cos(2\delta\beta_j) Y_i Z_j + \sum_{i<j} J_{ij} \sin(2\delta\beta_i \sin(2\delta\beta_j) Y_i Y_j.
\end{aligned} \tag{A17}$$

Let's 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

$$H' = \cos^2(2\delta\beta) H + \sin(4\delta\beta) \sum_{i,j} \frac{J_{ij}}{2} Z_i Y_j + \sin^2(2\delta\beta) \sum_{i<j} J_{ij} Y_i Y_j. \tag{A18}$$

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 non-trivial way. Let us look at a perturbation to order $\delta\beta$:

$$H' = H + 2\delta\beta \sum_{i,j} J_{ij} Z_i Y_j + O(\delta\beta^2). \tag{A19}$$

The first order correction to any energy level is

$$\sum_{i,j} J_{ij} \langle E | Z_i Y_j | E \rangle = 0, \tag{A20}$$

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

$$\langle E_1 | Y_j | E_2 \rangle, \tag{A21}$$

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:

$$\langle H' \rangle = \sum_n E_n \left| \langle m | \left(\frac{|n\rangle - i\epsilon \sum_k X_k |n\rangle}{\sqrt{1 + N\epsilon^2}} \right) \right|^2 \tag{A22}$$

$$= \frac{1}{1 + N\epsilon^2} \left(E_m + \epsilon^2 \sum_n E_n | \langle m | \sum_k X_k |n\rangle |^2 \right). \tag{A23}$$

We now note that the inner product is 1 if and only if n is a bitflip 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 w.l.o.g the case where let k be the vertex of highest index.

$$E_m - E_n = \Big(\sum_{\substack{i<j \\ j\neq k}} J_{ij} \langle m|Z_i Z_j|m\rangle + \sum_{i\neq k} J_{ik} \langle m|Z_i Z_k|m\rangle\Big) -$$

$$\Big(\sum_{\substack{i<j \\ j\neq k}} \langle m|X_k Z_i Z_j X_k|m\rangle + \sum_{i\neq k} J_{ik} \langle m|X_k Z_i Z_k X_k|m\rangle\Big). \quad \text{(A24)}$$

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

$$2\sum_{i\neq k} J_{ik} \langle m|Z_i Z_k|m\rangle \quad \text{(A25)}$$

Then we can easily rearranged 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. (A22):

$$\frac{1}{1+N\epsilon^2}\Big(E_m + \epsilon^2\big(\sum_k E_m - 2\sum_{i\neq k} J_{ik}\langle Z_i Z_k\rangle\big)\Big) \quad \text{(A26)}$$

$$= \frac{1}{1+N\epsilon^2}\big(E_m + \epsilon^2 N E_m - 4\sum_{i<k} J_{ik}\langle Z_i Z_k\rangle\big) \quad \text{(A27)}$$

$$= E_m\big(\frac{1+(N-4)\epsilon^2}{1+N\epsilon^2}\big) \quad \text{(A28)}$$

$$\approx E_m(1 - 4\epsilon^2 + 4N\epsilon^4 + \mathcal{O}(\epsilon^6)). \quad \text{(A29)}$$

So for this case overrotations also just scales the old eigenvalues

3. Non-constant Mixing Overrotations

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

$$\mathcal{E}_U^\alpha(\rho) = \int_0^{2\pi} \frac{e^{\kappa\cos(\varepsilon)}}{2\pi I_0(\kappa)} U(\alpha+\varepsilon)\rho U^\dagger(\alpha+\varepsilon)d\varepsilon, \quad \text{(A30)}$$

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

The integral for mixer leads to the following map

$$\mathcal{E}_M^\beta(\rho) = \frac{1}{2}\Big(1 + \frac{I_2(\kappa)}{I_0(\kappa)}\cos(2\beta)\Big)\rho + \frac{1}{2}\Big(1 - \frac{I_2(\kappa)}{I_0(\kappa)}\cos(2\beta)\Big)X\rho X - \frac{iI_2(\kappa)}{I_0(\kappa)}\frac{\sin(2\beta)}{2}[X,\rho], \quad \text{(A31)}$$

while for a single gate of phase (assuming that we have 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

$$\mathcal{E}_P^{\gamma,J}(\rho) = \frac{1}{2}\Big(1 - \frac{I_2(\kappa)}{I_0(\kappa)}\Big)A_1(\gamma,J)\rho A_1^\dagger(\gamma,J) + \frac{1}{2}\Big(1 + \frac{I_2(\kappa)}{I_0(\kappa)}\Big)A_2(\gamma,J)\rho A_2^\dagger(\gamma,J), \quad \text{(A32)}$$

where

$$A_1(\gamma,J) = \text{Diag}\big[(1, -e^{2i\gamma J}, -e^{2i\gamma J}, 1)\big] \quad , \quad A_2(\gamma,J) = \text{Diag}\big[(1, e^{2i\gamma J}, e^{2i\gamma J}, 1)\big]. \quad \text{(A33)}$$

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 form

$$\mathcal{E}_M^\beta(\rho) = \frac{1}{2}\left(1 + \frac{I_2(\kappa)}{I_0(\kappa)}\cos(2\beta)\right)\rho + \frac{1}{2}\left(1 - \frac{I_2(\kappa)}{I_0(\kappa)}\cos(2\beta)\right)X\rho X. \tag{A34}$$

This maps

$$Z \to \frac{I_2(\kappa)}{I_0(\kappa)}Z. \tag{A35}$$

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

4. Pauli Channel

Interesting class of noisy channel is so-called Pauli channels, that have form

$$\mathcal{E}(\rho) = \sum_{k=0}^{3} p_k \sigma_k \rho \sigma_k, \tag{A36}$$

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 (8).

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

$$\mathcal{E}^\#(Z) = (p_0 + p_3 - (p_1 + p_2)) * Z, \tag{A37}$$

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 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 same as for the amplitude damping case analyzed in Sec II A 2 and A_2 given in

$$A_2 = \begin{pmatrix} 0 & 0 \\ 0 & \sqrt{\gamma} \end{pmatrix}. \tag{A38}$$

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