Quantifying the Impact of Precision Errors on Quantum Approximate Optimization Algorithms

Gregory Quiroz,
Paraj Titum,
Phillip Lotshaw,
Pavel Lougovski,
Kevin Schultz,
Eugene Dumitrescu,
and Itay Hen

The Johns Hopkins University Applied Physics Laboratory, Laurel, Maryland, 20723, USA
Quantum Computational Sciences Group, Computational Sciences and Engineering Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831, USA
Department of Physics and Astronomy, University of Southern California, Los Angeles, California 90089, USA
Center for Quantum Information Science & Technology, University of Southern California, Los Angeles, California 90089, USA
Information Sciences Institute, University of Southern California, Marina del Rey, California 90292, USA

The quantum approximate optimization algorithm (QAOA) is a hybrid quantum-classical algorithm that seeks to achieve approximate solutions to optimization problems by iteratively alternating between intervals of controlled quantum evolution. Here, we examine the effect of analog precision errors on QAOA performance both from the perspective of algorithmic training and canonical state- and observable-dependent QAOA-relevant metrics. Leveraging cumulant expansions, we recast the faulty QAOA as a control problem in which precision errors are expressed as multiplicative control noise and derive bounds on the performance of QAOA. We show using both analytical techniques and numerical simulations that errors in the analog implementation of QAOA circuits hinder its performance as an optimization algorithm. In particular, we find that any fixed precision implementation of QAOA will be subject to an exponential degradation in performance dependent upon the number of optimal QAOA layers and magnitude of the precision error. Despite this significant reduction, we show that it is possible to mitigate precision errors in QAOA via digitization of the variational parameters, therefore at the cost of increasing circuit depth. We illustrate our results via numerical simulations and analytic and empirical error bounds as a comparison. While focused on precision errors, our approach naturally lends itself to more general noise scenarios and the calculation of error bounds on QAOA performance and broader classes of variational quantum algorithms.

I. INTRODUCTION

The Quantum Approximate Optimization Algorithm (QAOA) [1] provides a hybrid classical-quantum approach to solving combinatorial optimization problems. QAOA has attracted a great deal of attention due to the advent of near- and intermediate-term quantum computing. Furthermore, because of its simplicity, QAOA has been amenable to both analytical study [1–10] and experimental implementation on a variety of platforms [11–14].

While the idealized QAOA provides performance guarantees [1, 5, 10], questions remain regarding its practical implementation, namely its robustness to noise. Previous studies have examined QAOA subject to local decoherence [15–17] and readout error [18] thus providing some insight into the ramifications of systematic and environment-induced noise on QAOA performance. These studies primarily leveraged numerical investigations to capture dependence on noise parameters. Analytical bounds for QAOA performance and training have been developed for a class of noise models that give a simple rescaling to Pauli terms in the density operator [19], e.g., in depolarizing channels. However, to the best of our knowledge, analytical bounds on QAOA success probability or training error in the presence of control noise have yet to be assessed. In this work, we develop bounds particularly focusing on errors in the control parameters. The techniques we develop are amenable to noise generated by imperfect control Hamiltonians, yet extendable to more generic noise models. Without loss of generality we illustrate their use within a simple yet representative error model.

QAOA is well-defined in the gate-model setting, however, its practical implementation involves the alternating application of unitary operators that depend on continuous parameters that must be optimized. Cases exist where the number $p$ of such alternations, commonly referred to as the order of the algorithm, can remain small while achieving an approximately satisfactory solution [1]. Unfortunately, a majority of such cases do not coincidence with problem classes of practical importance. As such, typically the QAOA order must increase commensurately with the problem size (and clause density when applicable) to generate near-optimal configurations [20–24].

In practical settings, the increasingly large sets of continuous variables upon which QAOA is founded are prone to misspecification in which the implemented parameters of the QAOA differ from those intended. We designate these errors as precision errors in the variational parameters and model these as imperfect control in the QAOA Hamiltonians. While static precision errors can be absorbed as a constant shift in the optimized parameters, temporally drifting errors varying on timescales shorter than the total algorithm runtime pose a
more significant threat to the success of the algorithm. This is most notably true if all parameters cannot be retrained more rapidly than the characteristic timescale of the error. Time-varying precision errors with short correlation times can quickly accumulate in optimization problems requiring large $p$, ultimately spoiling the computation and resulting in errors in parameter training and unfavorable reductions in performance guarantees.

The deleterious consequences of precision errors were recognized early on in gate-model quantum computing [25] and more recently in Hamiltonian quantum computing, in particular quantum annealing (QA) [26, 27]. Although the former may be addressed theoretically via quantum error correction [28], questions remain regarding how detrimental such errors are to NISQ-era algorithms and how to effectively combat them without error correction. Moreover, while QA has shown improvement in performance through quantum annealing correction, the presence of unmitigated precision errors has proven to be severely harmful [27]. Misspecification of programmable, continuous parameters within the cost function leads to exponential decay in success probability with problem size and error magnitude [26]. To date, it is unclear whether this behavior observed for QA is indeed a bad omen for QAOA, given the intimate relationship between the two paradigms [1].

In this study, we address these concerns via an analytical and numerical investigation of QAOA in the presence of precision errors. Providing insight into the extent of their harm on parameter training and performance guarantees, we enable the development of more robust estimates of QAOA complexity. Our approach relies on concepts from quantum control theory to analytically estimate the contribution of precision errors. This approach paves the way for the development of bounds on approximation ratios and training error that focus on precision errors and also enable further insight into broader classes of noise models. Through our analysis, we argue that any fixed precision implementation of QAOA is destined to detrimentally affect the success of QAOA, which represents a fundamental limitation to the scalability of the algorithm, and more broadly, Ising machines. In particular, we find an exponential reduction in success probability with increasing problem size and error magnitude. This observation is displayed through numerical investigations of the QAOA variant of Grover’s search and the one-dimensional transverse-field Ising model. Despite the apparent “doomsday scenario” for QAOA, we show that it is possible to mitigate precision errors in QAOA via digitization of the variational parameters.

The manuscript is organized as follows. In Sec. II we provide a concise summary of the main results, and bounds on the performance of QAOA for optimization. In Sec. III we provide a general overview of the QAOA algorithm and the model of precision errors considered in this paper. In Sec. IV we derive the general approach based on the cumulant expansion to analyze the error in QAOA evolution due to precision errors. In Sec V, we derive expressions for the cumulant for the specific case of coherent and stochastic errors before deriving bounds on the errors in terms of difference in expectation values, approximation ratio, gradients and unitaries in Sec. VI. In Sec. VII we show numerical evidence of the effects of precision errors in two example QAOA algorithms implementing (i) Grover search and (ii) 2-SAT (Ising) on a ring. Finally in Sec. VIII we discuss an approach to mitigating the effects of precision errors in the implementation of QAOA.

II. SUMMARY OF MAIN RESULTS

Variational training parameters in QAOA are susceptible to precision errors that potentially threaten the success of the algorithm. We model precision errors via a control noise contribution to the QAOA Hamiltonians. This results in a misspecification of the QAOA control amplitudes, as depicted in the bottom-left panel of Fig. 1. Mathematically, parameter misspecifications are captured by time-dependent error functions that are specified by white noise processes or constant errors to model precision errors in two specific regimes. Upon subjecting QAOA to precision errors, we observed an exponential decrease in the achievable approximate optimal solution for QAOA with increasing noise strength and optimal QAOA order; see Fig. 1(bottom-right) for an illustration of this effect.

Using concepts from quantum control theory, we quantify the effect of precision errors on expectation values and unitary operators. Recent work in quantum control has sought to utilize cumulant expansions to assess interplay between control schemes and temporally-correlated noise processes [29–31]. Inspired by this work, we leverage a cumulant expansion to examine the dynamics of a system subject to a control designated by the noiseless QAOA protocol and noise produced by a control error Hamiltonian representing precision errors. The cumulant-based approach affords a novel perspective on faulty QAOA dynamics that facilitates the development of bounds on various QAOA-relevant quantities. In particular, our approach allows for the development of bounds on approximation ratios and parameter training error. Fig. 1(bottom-right) illustrates the utility of our approach for assessing QAOA susceptibility to precision errors. The cumulant expansion yields expressions for the approximate dynamics of the noisy QAOA. Upper bounds on QAOA-relevant quantities (e.g., the error in average energy) are determined via the cumulant approach. Together, these expressions capture the relative behavior of QAOA as a function of noise strength and QAOA order.

To benchmark our analysis we numerically study the performance of QAOA in the presence of precision errors in two specific implementations:

(i) **Grover Search**: We implement the analytically obtained optimal Grover QAOA for identifying a marked state $|0\rangle$ from a database of $N = 2^n$ entries [32]. The optimal circuit has $p^* \sim \sqrt{N}$ layers.

(ii) **Ising instances**: We utilize optimal circuits to prepare $n$ qubit GHZ states using QAOA with a Ising Hamiltonian on a ring geometry [33]. The optimal circuit has $p^* = n/2$ layers.

In both cases, $p^*$ is the optimal number of QAOA layers and $n$ indicates the number of qubits. We numerically simulate
the QAOA circuits for these two problem instances and calculate the measured errors in the cost observable as well as the average distance between the noisy and perfect unitary operators. The precision errors are drawn from a normal distribution, $\mathcal{N}(\eta, \sqrt{T})$ with mean $\eta$ and variance $\Gamma$. Specifically, we consider two representative limiting cases: (a) Stochastic errors: $\eta = 0$, drawn from $\mathcal{N}(0, \sqrt{T})$, and (b) Coherent errors: $\Gamma = 0$, drawn from $\mathcal{N}(\eta, 0)$.

Finally, let us summarize the main numerical results as well as the analytical bounds obtained on the faulty QAOA evolution as a result of precision errors. Let us consider a QAOA circuit with a total runtime $T$ (and layers $p$) and perfect implementation leading to the state $\rho_0(T)$.

1. **Expectation values**– The noise-averaged error in the measured expectation value of any observable, $O$ has the following upper bound,

$$\|\Delta \langle O(T) \rangle \| \leq e^{\|C_O(T)\|_\infty} - 1 \|O \rho_0(T)\|_1,$$

where $C_O(T)$ [defined in Eq. (30)] denotes the cumulant expansion for the error operator $\Lambda(T)$ that depends on the noise as well as the observable, $O$; see Eq. (27). The operator $C_O(T)$ is dependent upon the specifications of the problem; however, it can be shown to scale as $\|C_O(T)\|_\infty \lesssim O(\eta p^*)$ and $\|C_O(T)\|_\infty \lesssim O(\Gamma p^*)$ for constant and stochastic precision errors, respectively. The numerical investigations of QAOA circuits for both examples considered, indicate the following scaling for the error in expectation value of the cost function Hamiltonian, $H_C$ for weak noise,

$$|\Delta H_C| \approx \langle H_C \rangle_0 (1 - \exp [\chi_H(p^*)]) ,$$

with, $\chi_H(p^*) \equiv \begin{cases} \chi (\Gamma p^*), & \text{stochastic error} \\ \chi (\eta p^*), & \text{coherent error} \end{cases}$

where $\langle H_C \rangle_0$ is the noiseless expectation value and the exact functional form of $\chi_H(p^*)$ depends on the details of the QAOA algorithm. The effects of the coherent error turns out to be dependent on the algorithm, with the $a = 1$ for Grover search and $a = 2$ for the Ising problem.

2. **Unitary operators**– We obtain a bound on the operator norm of the difference between the perfect and faulty unitary operators for the QAOA evolution via the Frobenius norm,

$$\|U - U_0\|_\infty^2 \leq (\Delta U)^2 \leq 2 \left( e^{\|C(T)\|_\infty} - 1 \right),$$

where $\Delta U = \|U - U_0\|$ and $C$ denotes the cumulant expansion for the error operator. [Note that the difference between unitaries has a direct relationship with the error in the expectation value; see Eq. (C5).] Numerically, we evaluate the $\infty$-norm difference between the QAOA evolution operators for the Grover search algorithm. For weak noise and small number of layers $p$ the average difference scales as,

$$\|U - U_0\|_\infty \approx \sqrt{\Gamma p},$$

with,

- [stochastic error]
- [coherent error]
3. **Training Error**—We derive bounds on the error in training variational parameters for gradient-based and closed-loop optimization routines. The former follows from the absolute error in the gradient of the expectation value, which can be bounded by expressions proportional to the operator norm (largest singular value) of the error operator $\Lambda(T)$ and its gradient. We estimate training error in closed-loop optimization via bounds on the mean-squared error (MSE) in the expectation value of $H_C$. In the case of constant errors, the training error is bounded by

$$\text{MSE}(H_C(T)) \leq 2C_{\text{max}}^2 (1 + 2h_E^2), \quad (6)$$

where $C_{\text{max}}$ is the operator norm of $H_C$ (or equivalently, the maximum value of the cost function) and $h_E$ effectively represents the accumulated precision error; see Eq. (63). Similar expressions that depend more generically on the operator norm of $\Lambda(T)$ are obtained for stochastic precision errors.

4. **Approximation ratio**—The bound on the noise-averaged error in the expectation value naturally leads to the bound on the error in the optimal approximation ratio

$$|\Delta \epsilon^*| \leq \frac{1}{C_{\text{max}}} \left( \epsilon \|C_0(T^*)\|_\infty - 1 \right) \|H_C \rho_0(T^*)\|_1. \quad (7)$$

The algorithmic runtime $T^*$ denotes the time required to achieve the noiseless optimal approximation ratio $\epsilon^*$. The Hamiltonian $H_C$ denotes the cost function Hamiltonian.

Through our analytical and numerical analysis, we find that stochastic precision errors yield exponentially increasing error in expectation values and distance between unitaries. This observation subsequently propagates to parameter training and approximation ratios. Despite the detrimental nature of precision errors, we devise a strategy for effectively addressing them by digitizing the variational parameters in a binary representation and implementing each QAOA evolution operator as a composite operator. This strategy relies on each constituent operator possessing a greater precision than the desired precision of the QAOA. More precisely, for a $p$-layer QAOA to achieve a desired accuracy of $2^{-\epsilon}$, the precision error can be at most $\sim \epsilon/\sqrt{p}$. Provided that such stipulations can be met, the digitization strategy ultimately requires an increase in the circuit depth by a factor $\sim \log(\epsilon/\sqrt{p})$. Digitization is not required for constant coherent errors, where the over- or under-rotation can be accounted for (1) by updating the optimized angles exactly by the coherent error or (2) intrinsically by a closed-loop classical optimization routine used for parameter training.

## III. **Quantum Approximate Optimization Algorithm**

In combinatorial optimization, approximation algorithms seek solutions with provable guarantees on the distance between the value of the returned solution and the global optimum. More concretely, consider an objective function $C : \{0, 1\}^n \rightarrow \mathbb{R}$ which is to be optimized. An $\epsilon^*$-approximation algorithm with high probability achieves a solution $x^*$ such that

$$\frac{C(x^*)}{C_{\text{max}}} \geq \epsilon^*. \quad (8)$$

Under these conditions, the algorithm insures with high probability that all solutions are within the approximation ratio $\epsilon^*$ of the maximum $C_{\text{max}}$.

The QAOA generates approximate solutions by stroboscopically alternating between two intervals of controlled evolution. In QAOA, these evolutions are generated by a mixing Hamiltonian $H_M$ and a problem Hamiltonian $H_C$. The mixing Hamiltonian can take various forms; however, typically, it assumes the form of a global transverse field

$$H_M = \sum_{j=1}^n X_j. \quad (9)$$

Other mixing Hamiltonians that account for problem constraints have also been examined [9].

The problem Hamiltonian is constructed by translating the classical cost function of $n$-binary variables into a Hamiltonian of $n$ qubits. This is accomplished by introducing the variable transformation $x_i = (1 - z_i)/2$, where $z_i \in \{-1, 1\}$. The cost function $C(z)$ is then transformed into $H_C$ by replacing $z_i$ with the Pauli operator $\sigma_i^z$. The resulting problem Hamiltonian is an Ising Hamiltonian defined on the underlying graph of the cost function $C(z)$. Note that $H_C$ acts diagonally on computational basis states $|z\rangle$, i.e.

$$H_C |z\rangle = C(z) |z\rangle. \quad (10)$$

The QAOA is formally implemented via a time-dependent Hamiltonian

$$H_0(t) = |1 - s(t)|H_M + s(t)H_C, \quad (11)$$

where

$$s(t) = \begin{cases} 1 & : \quad t \in \Delta t^C_j \\ 0 & : \quad t \in \Delta t^M_j \end{cases}. \quad (12)$$

defines a stroboscopic function that toggles between weighted time-evolutions generated by $H_M$ and $H_C$. The time intervals $\Delta t^C_j = [t_{2j-2}, t_{2j-1}]$ and $\Delta t^M_j = [t_{2j-1}, t_{2j}]$, $j = 1, \ldots, p$ define the periods during with $H_C$ and $H_M$, respectively, govern the evolution. The dynamics generated by $H_0$ can be expressed as

$$U_0(T, 0) = T_+ e^{-i \int_0^T H_0(t) dt} = U_M(\beta_p) U_C(\gamma_p) \cdots U_M(\beta_1) U_C(\gamma_1),$$

$$= U_0^{(p)}(\gamma, \beta). \quad (13)$$
where $\mathcal{T}_+$ denotes the time-ordering operator. The variational parameters $\gamma = (\gamma_1, \ldots, \gamma_p)$ and $\beta = (\beta_1, \ldots, \beta_p)$ control the time-intervals over which $H_M$ and $H_C$ are applied. Each $(\gamma_j, \beta_j)$ are real numbers that parameterize the evolution operators $U_M(\beta) = e^{-i\beta H_M}$ and $U_C(\gamma) = e^{-i\gamma H_C}$ and characterize the total algorithmic runtime $T = \sum_{j=1}^p (|\gamma_j| + |\beta_j|)$.

QAOA evolution is applied to a state initialized in an equal computational basis, is highly concentrated on solutions that are approximately optimal [1]. The expectation value is the average energy with respect to the problem Hamiltonian. The QAOA evolution is then given by

$$
\rho_0(T) = U_C^{(p)}(\gamma, \beta) \rho_0(0) U_C^{(p)}(\gamma, \beta)^\dagger.
$$

The variational parameters are optimized using a classical optimization routine. While there are many options for this procedure [34–37], all of them seek to obtain an approximate solution to the optimization problem by maximizing

$$
F(\gamma, \beta) = \langle H_C \rangle_{\gamma, \beta} = \text{Tr} \left[ \rho_0(T) H_C \right],
$$

the average energy with respect to the problem Hamiltonian. One can define the approximation ratio for the QAOA with respect to Eq. (16) as

$$
\epsilon \equiv F(\gamma, \beta)/C\text{max}.
$$

The QAOA is now said to be an $e^\epsilon$-approximation algorithm if $\epsilon \geq \epsilon^*$. The approximation condition is guaranteed if the resulting quantum state $\rho_0(T)$, when measured in the computational basis, is highly concentrated on solutions that are $\epsilon^*$-approximately optimal [1].

### A. Precision Errors

Control Hamiltonians are designed to implement a particular quantum operation. However, physical realizations of these Hamiltonians are often accompanied by undesired control-dependent additive or multiplicative noise [31, 38, 39]. As a result, unwanted system dynamics generated by control parameter inaccuracies lead to control errors. Precision errors constitute a particular type of control error that while being associated with the misspecifications of variational parameters, also can be treated as control amplitude errors in the QAOA evolution. It is this quantum control point of view that we utilize to examine precision errors in the QAOA setting.

Precision errors can be modeled in QAOA by unintended deviations in the variational parameters. This is equivalent to introducing control noise into the QAOA evolution, such that the faulty dynamics are governed by

$$
H(t) = H_0(t) + H_E(t).
$$

The dynamics are now governed by the ideal QAOA and an additional error Hamiltonian

$$
H_E(t) = \eta_M(t)[1 - s(t)]H_M + \eta_C(t)s(t)H_C.
$$

The error functions $\eta_M(t)$ and $\eta_C(t)$ capture the time-dependent control errors for the mixer and problem Hamiltonians, respectively. The evolution generated by Eq. (18) is that of a faulty QAOA evolution $U_0^{(p)}(\tilde{\gamma}, \tilde{\beta})$, where $\tilde{\gamma}_j = \gamma_j(1 + \eta_{C,j})$ and $\tilde{\beta}_j = \beta_j(1 + \eta_{M,j})$ are implemented rather than the ideal evolution of Eq. (13). So far our model describes a general class of multiplicative control errors, i.e., errors that are generated by a faulty control Hamiltonian where the error magnitude scales proportionally with the control.

The statistical properties of the error functions further specify the precision error characteristics. Namely, we model the error functions as stationary, Gaussian random variables with mean, $\bar{\eta}_\mu(t) = \eta_\mu$, and time-dependent two-point correlation functions

$$
\bar{\eta}_\mu(t_1)\bar{\eta}_\nu(t_2) = \delta_{\mu \nu} \Gamma_\mu(t_1 - t_2).
$$

where $\Gamma_\mu(t_1 - t_2)$ describes the autocorrelation function for $\mu, \nu \in \{M, C\}$. Note that $\cdots \cdots$ denotes classical ensemble averaging. Below, we will focus on two cases: (1) constant coherent errors, where $\Gamma_\mu \equiv \eta_\mu^2$, and (2) stochastic errors, where $\Gamma_\mu(t_1 - t_2) = \Gamma_\mu \delta(t_1 - t_2)$. Together, these scenarios capture two extremes of precision errors, each uniquely detrimental to QAOA with distinct scaling behavior in QAOA performance as a function of $p$.

### IV. ERROR DYNAMICS

We investigate the effect of precision errors on QAOA performance via the average energy with respect to the problem Hamiltonian. The faulty QAOA dynamics generated by Eq. (18) leads to errors in the final state of the system. Ultimately, these errors propagate to the objective function given in Eq. (16). In order to understand the consequences of precision errors on $F(\gamma, \beta)$, we leverage concepts from control theory to isolate the error dynamics dictated by $H_E(t)$. We examine the error dynamics through a cumulant expansion approach that has been previously employed for general studies of open quantum system dynamics in the presence of temporally correlated noise [30].

#### A. Rotated-Frame Dynamics

In quantum control, time-dependent perturbation theory (TDPT) has proven to be an invaluable tool for examining weak noise processes in the presence of strong control. TDPT expansions are typically performed by moving into a rotating reference frame with respect to the control – commonly referred to as the toggling-frame. The noise contributions are then studied within this rotating frame using a particular expansion technique, e.g., Dyson or Magnus expansion [40, 41].
Utilizing the framework of quantum control, the faulty QAOA is analyzed through the lens of TDPT. We work under the assumption that the dynamics are dominated by the ideal QAOA evolution and thus impose a strong control condition \( \|H_E(t)\| \ll \|H_0(t)\| \) \forall t. Furthermore, we enforce a weak noise assumption \( \|H_E(t)\| \ll 1 \). Together, these conditions enable the effects of \( H_E(t) \) on the QAOA evolution to be analyzed via TDPT.

The first step towards isolating the effects of \( H_E(t) \) is to move into an interaction picture co-rotating with the ideal evolution \( H_0(t) \). The erroneous Hamiltonian transforms into

\[
\tilde{H}_E(t) = U_0(T, t)H_E(t)U_0^\dagger(T, t) = \eta_M(t)[1 - s(t)]\tilde{H}_M(t) + \eta_C(t)s(t)\tilde{H}_C(t). \tag{21}
\]

The mixer and problem Hamiltonian in the rotated frame are given by \( \tilde{H}_M(t) = U_0(T, t)H_MU_0^\dagger(T, t) \) and \( \tilde{H}_C(t) = U_0(T, t)H_CU_0^\dagger(T, t) \), respectively. Note that the rotated error Hamiltonian is expressed in terms of a “reverse” interaction picture with respect to the ideal QAOA evolution. As a result, the total dynamics evolution operator

\[
U(T) = T_+e^{-i\int_0^T \tilde{H}(t)dt} \tag{22}
\]

is factorized into the product

\[
U(T) = \tilde{U}_E(T)U_0(T), \tag{23}
\]

where

\[
\tilde{U}_E(T) = T_+e^{-i\int_0^T \tilde{H}_E(t)dt}. \tag{24}
\]

and \( U_0(t) \) is the ideal evolution generated by Eq. (11). The operator \( T_+ \) denotes time-ordering. The advantage of this formulation will become clear in the next section and the equivalence between this approach and the canonical rotated frame is discussed in Appendix A.

**B. Time-dependent Dynamics of Observables**

Expectation values are commonly utilized in variational quantum algorithms to evaluate algorithmic performance and carry out variational parameter training. In QAOA, the average energy with respect to the problem Hamiltonian is the expectation value of interest as in Eq. (16). This metric is utilized to train variational parameters, to maximize the objective function defined by \( H_C \), which in turn defines the approximation ratio for a particular problem and depth. Therefore, the effect of noise on this expectation value will ultimately influence QAOA parameter training and bounds on performance guarantees.

In order to investigate the effect of precision errors on QAOA performance, we first consider the case of a general observable \( O \). The noise-averaged, time-dependent dynamics of \( O \) subject to the faulty QAOA are described by

\[
\langle O(T) \rangle = \langle \text{Tr}[\rho(T)O] \rangle \tag{25}
\]

where it is assumed that we are particularly interested in the expectation value of \( O \) after the total algorithmic runtime \( T \). Using

\[
\rho(t) = \tilde{U}_E(t)U_0(t)\rho(0)U_0^\dagger(t)\tilde{U}_E^\dagger(t), \tag{26}
\]

the expectation value can be rewritten as

\[
\langle O(T) \rangle = \langle \text{Tr}\left[\tilde{U}_E(T)\rho_0(T)\tilde{U}_E^\dagger(T)O\right]\rangle = \text{Tr}\left[\Lambda(T)\rho_0(T)O\right]. \tag{27}
\]

The reason for the particular rotating frame choice is now evident: it allows for the dynamics to be partitioned into an error operator \( \Lambda(T) = O^{-1}\tilde{U}_E^\dagger(T)O\tilde{U}_E(T) \) and the ideal time-evolved state \( \rho_0(T) \). The error operator encapsulates all of the error dynamics, i.e., all dynamics generated by the error Hamiltonian. Eq. (27) expresses the error dynamics in an intuitive way, where an absence of \( H_E(t) \) results in \( \Lambda(T) = I \). This perspective can be quite useful, especially when one seeks to mitigate the effects of the noise, and thus, minimize the distance between \( \Lambda(T) \) and the identity operator. It is important to note that \( \Lambda(T) \) demands \( O \) to be invertible, however, as we discuss in Appendix B.1, if one can write a general observable \( O \) as a sum of invertible operators then similar expression to Eq. (27) can be obtain.

Formally, the error operator can be expressed in terms of a cumulant expansion. This is observed by rewriting the error operator as

\[
\Lambda(T) = T_+e^{-i\int_0^T \tilde{H}_{E,O}(t)dt}, \tag{28}
\]

where the observable-dependent effective Hamiltonian is given by

\[
\tilde{H}_{E,O}(t)=\begin{cases} 
\tilde{H}_E(T + t) & t \in [-T, 0], \\
-O^{-1}\tilde{H}_E(T - t)O & t \in [0, T].
\end{cases} \tag{29}
\]

Eq. (28) lends itself to the cumulant expansion in which the cumulant expressions can be obtained from the moment-generating equation

\[
\Lambda(T) = T_+e^{-i\int_0^T \tilde{H}_{E,O}(t)dt} = e^{\mathcal{C}_O(T)}, \tag{30}
\]

where \( \mathcal{C}_O(T) = \sum_{k=1}^{\infty}(-i)^k\frac{C^{(k)}_{O}(T)}{k!} \). In general, the expansion of \( \Lambda(T) \) involves an infinite number of terms. While in certain scenarios the above cumulant expansion can be truncated exactly [30], this will not be possible for the faulty QAOA problem. As a result, we will truncate the expansion to capture the approximate error dynamics. Leveraging the weak noise condition discussed above, the error dynamics are assumed to be well-approximated by a second order truncation on \( \Lambda(T) \). More concretely, it is assumed that

\[
\|\mathcal{C}_O(T)\|_\infty \approx \|\mathcal{C}^{(1)}_O(T) + \frac{i}{2} \mathcal{C}^{(2)}_O(T)\|_\infty, \tag{31}
\]

where \( \| \cdot \|_\infty \) is the operator norm, or largest singular value. Note that through the triangle inequality, it is trivial to show
that the truncated norm of the cumulant expansion can be up-
per bounded as $\|C_O(T)\|_\infty \lesssim \|H_E(T)\|_\infty + \|H_E(T)\|_2^2$; 
thus, explicitly conveying a connection between the cumulant 
expansion and the weak noise condition.

Under the above condition, the error dynamics are charac-
terized by the first and second cumulants. The first order term

$$C^{(1)}_O(T) = \int_0^T dt_1 \left[ C^{(1)}(\tilde{H}_E(t_1)) - C^{(1)}(\tilde{H}_E(t_1)) \right]$$

is dependent upon the rotated-frame error Hamiltonian, its 
observable-conjugated counterpart $\tilde{H}_E(t) = O^{-1}H_E(t)O$, 
and the cumulant expression $C^{(1)}(A) = \bar{A}$. The second order term

$$\frac{C^{(2)}_O}{2!} = I_1(T) + I_2(T) - I_3(T) - I_4(T),$$

can be decomposed into four integrals:

$$I_1(T) = \int_0^T dt_1 \int_0^{t_1} dt_2 \ C^{(2)}(\tilde{H}_E(t_1), \tilde{H}_E(t_2))$$

$$I_2(T) = \int_0^T dt_1 \int_0^{t_1} dt_2 \ C^{(2)}(\tilde{H}_E(t_2), \tilde{H}_E(t_1))$$

$$I_3(T) = \int_0^T dt_1 \int_0^{t_1} dt_2 \ C^{(2)}(\tilde{H}_E(t_1), \tilde{H}_E(t_1))$$

$$I_4(T) = \int_0^T dt_1 \int_0^{t_1} dt_2 \ C^{(2)}(\tilde{H}_E(t_2), \tilde{H}_E(t_1)).$$

Each integral includes the second order cumulant expression

$$C^{(2)}(A, B) = \bar{A}\bar{B} - \frac{1}{2} (\bar{A}\bar{B} + \bar{B}\bar{A}).$$

Together, these equations form the quantum control frame-
work that will be exploited to calculate the approximate error 
dynamics generated by $H_E(t)$ and subsequently, the expecta-
tion value of $H_E$ in the presence of precision errors. Further-
more, below it will be shown that the cumulant expressions 
facilitate the development of error bounds on various QAOA-
relevant metrics.

V. CUMULANT ANALYSIS

The truncated cumulant expansion can be utilized to calcu-
late the approximate error dynamics of QAOA subject to pre-
cision errors. Below, we investigate the dynamics of QAOA 
for coherent and stochastic errors on the variational param-
ters.

A. Constant Coherent Errors

Constant coherent errors are trivially captured by the cu-
mulant expansion, most notably when the error mean is small, 
i.e., $\eta_\mu \ll 1$. In this case, the dynamics are well-characterized

by the first cumulant, with the second order term providing 
an additional correction that becomes more relevant with in-
creasing error mean. Assuming the QAOA coherent error is 
weak, we find the error dynamics to be governed by

$$C^{(1)}_O(T) = \sum_{\mu=M,C} \eta_\mu \int_0^T dt \left( \tilde{H}_\mu(t) - O^{-1}\tilde{H}_\mu(t)O \right).$$

The QAOA evolution is piece-wise constant which allows for 
the dynamics to be partitioned into a sum of integrals in which 
the control propagator can be expressed as

$$U_0(T, t) = \left\{ \begin{array}{ll}
Q_{p,j+1}e^{-it_2\tilde{H}_M} ; & t \in \Delta t^M_j \\
Q_{p,j}U_C(\gamma_j)e^{-it_2\tilde{H}_C} ; & t \in \Delta t^C_j
\end{array} \right.,$$

where

$$Q_{k;j} = U_M(\beta_k)U_C(\gamma_k)\cdots U_M(\beta_j)U_C(\gamma_j).$$

Examining the error Hamiltonian and the control propagator, 
one finds that certain commutations naturally arise between 
terms that comprise $H_E(t)$ and $U_C(T, t)$. As a result, the first 
cumulant reduces to integrals over constants and thus,

$$C^{(1)}_O(T) = \sum_{\mu=M,C} \eta_\mu \sum_{j=1}^p g^j_\mu \left( Q_{p:k(\mu)}H_\mu Q^\dagger_{p:k(\mu)} \right)$$

$$- O^{-1}Q_{p:k(\mu)}H_\mu Q^\dagger_{p:k(\mu)}O \right),$$

where $g^M_j = \gamma_j$ and $g^C_j = \beta_j$. The index $k(\mu, j)$ captures 
the distinctions in the control evolution of Eq. (40), where 
k($j, M$) = $j + 1$ and $k(J, C) = j$.

B. Stochastic Errors

We now consider the case of stochastic errors in the vari-
ational parameters. Incorporating the zero-mean assumption 
of the noise model results in $C^{(1)}_O(T) = 0$; hence, the dynamics 
are governed by $C^{(2)}_O(T)$. As will be shown below, the four 
integrals that comprise the second order term can be exactly 
calculated for this precision error model.

There are a number of features of our noise model that al-
low the second cumulant expressions to be conveniently sim-
plicated. (1) The statistical properties of the noise, namely, 
the lack of cross-correlations, allows for the second cumu-
lant integrals to be partitioned into terms solely propor-
tional to $H_M(t)$ or $H_C(t)$, i.e., $I_2(T) = I_{2,M}(T) + I_{2,C}(T)$. (2)
When combined with the piecewise-constant nature of the er-
or Hamiltonian, the integral expressions reduce to a sum of 
nested integrals only over the domain in which $t_{j-1} \leq s_1 \leq s_2 \leq t_j$. (3) Leveraging the definition of the ideal QAOA evo-
olution given in Eq. (40) and again exploiting commutations 
between the error Hamiltonian and the ideal propagator, the 
above expressions once more become integrals over constants.
Together, these features result in the following expressions for $I_3(T)$:

$$I_{1,M}(T) = \frac{\Gamma_M}{2} \sum_{j=1}^{p} \beta_j Q_{p,j+1}^\dagger H_M Q_{p,j+1}$$ \hfill (43)

$$I_{1,C}(T) = \frac{\Gamma_C}{2} \sum_{j=1}^{p} \gamma_j Q_{p,j} H_C Q_{p,j}^\dagger$$ \hfill (44)

Following similar procedures, it can be shown that $I_2(T) = O^{-1} I_1(T) O$, while $I_3(T) = I_{3,M}(T) + I_{3,C}(T)$ with

$$I_{3,M}(T) = \frac{\Gamma_M}{2} \sum_{j=1}^{p} \beta_j O^{-1} \left( Q_{p,j+1}^\dagger H_M Q_{p,j+1}^\dagger \right) O \left( Q_{p,j+1}^\dagger H_M Q_{p,j+1}^\dagger \right)$$ \hfill (45)

$$I_{3,C}(T) = \frac{\Gamma_C}{2} \sum_{j=1}^{p} \gamma_j O^{-1} \left( Q_{p,j} H_C Q_{p,j}^\dagger \right) O \left( Q_{p,j} H_C Q_{p,j}^\dagger \right)$$ \hfill (46)

It is straightforward to show the remaining term satisfies $I_1(T) = I_3(T)$ as a result of the noise correlation function.

VI. BOUNDS

Perturbative expansions, like Dyson and Magnus, have proven to be useful tools for developing bounds on open quantum system dynamics [42]. Such bounds have been utilized to investigate and evaluate control schemes designed to mitigate unwanted environmental interactions [43–45]. Naturally, the cumulant expansion affords a similar capability that we will exploit here to bound various metrics relevant to QAOA.

A. Error in Expectation Values

The expectation value of an observable is a key metric of QAOA that is utilized for both parameter training and calculating performance guarantees. Using the cumulant expansion, we bound the expectation value of an observable $O$ in the presence of a faulty QAOA evolution as follows: first, we find that

$$\|O(T)\| = |\text{Tr}[\Lambda(T) \rho_0(T) O]| \leq \|\Lambda(T)\|_\infty \|O \rho_0(T)\|_1 .$$ \hfill (47)

This follows directly from $|\text{Tr}(AB)| \leq \|A\|_\infty \|B\|_1$, where $\|B\|_1 = |\text{Tr}(B)| = \sum s_i(B)$ is the trace norm of $B$, or equivalently, the sum of singular values $s_i(B)$ [42]. Using sub-multiplicativity and the triangle inequality, it is straightforward to show that the operator norm of the error operator can be bounded in terms of the cumulants as

$$\|\Lambda(T)\|_\infty = \|e^{\mathcal{C}_O(T)}\|_\infty \leq e^{\|\mathcal{C}_O(T)\|_\infty};$$ \hfill (48)

see Appendix C for further details. As a result, we obtain the upper bound

$$|\langle O(T) \rangle_0| \leq e^{\|\mathcal{C}_O(T)\|_\infty} \|O \rho_0(T)\|_1$$ \hfill (49)

on any observable subject to a faulty QAOA. Note that this bound is independent of the noise model and therefore, holds for both precision errors and more generic spatio-temporally correlated noise models.

Employing a similar analysis, one can derive a bound on the absolute error between the ideal and faulty expectation values. Consider the quantity

$$|\Delta \langle O(T) \rangle_0| = |\langle O(T) \rangle - \langle O(T) \rangle_0|$$ \hfill (50)

containing the expectation value of $O$ subject to $U(T)$, $\langle O(T) \rangle_0$, and the ideal case $\langle O(T) \rangle_0$ time evolved by $U_0(T)$. In the latter, the error dynamics are equivalent to the identity operator, i.e., $\Lambda(T) \equiv I$. Using this fact, along with properties of the trace, it can be shown that

$$|\Delta \langle O(T) \rangle_0| = |\text{Tr}[(\Lambda(T) - I) \rho_0(T) O]| \leq \left(e^{\|\mathcal{C}_O(T)\|_\infty} - 1\right) \|O \rho_0(T)\|_1 .$$ \hfill (51)

Additionally, if $\|\mathcal{C}_O(T)\|_\infty \leq 1$, then the inequality $e^x - 1 \leq (e - 1)x$ leads to $|\Delta \langle O(T) \rangle_0| \leq \|\mathcal{C}_O(T)\|_\infty \|O \rho_0(T)\|_1$. Note that in both cases, the error between the ideal and faulty expectation values is dependent upon the distance between $\|\mathcal{C}_O(T)\|_\infty$ and zero.

In addition to the absolute error, we provide bounds on the mean squared error (MSE). Representing the second moment in the error, the MSE measures the quality of an estimator in terms of both the variance and bias. We define the estimator as $\langle O(T) \rangle$ and the MSE as

$$\text{MSE}(O(T)) = \langle (O(T) - \langle O(T) \rangle_0)^2 \rangle$$

$$= \text{Var}(O(T)) + \left(\Delta \langle O(T) \rangle\right)^2$$ \hfill (52)
where \( \text{Var}(O(T)) = \langle O^2(T) \rangle − \langle O(T) \rangle^2 \) is the variance in the faulty observable expectation value. Note that defining the estimator in this particular way will result in a bound on the MSE in the asymptotic regime when the number of samples \( N_e \to \infty \). Finite sampling effects are further discussed in the supplement.

We derive two bounds on the MSE: one specific to constant coherent errors and another, more general bound characterized by the operator norm of \( \Lambda(T) \). The former utilizes formal integration of \( \tilde{U}_E(t) \) – similar to typical Trotter error bounding – to upper bound the MSE in terms of \( \|H_E(t)\|_\infty \). The latter employs the techniques used in Eq. (47) and applies to any spatio-temporally correlated, weak-noise model. The second approach inherently provides a bound for constant coherent errors as well, however, we find that the first approach yields a tighter bound than its more general counterpart for this particular error model.

In the case of coherent errors, the MSE is bounded as

\[
\text{MSE}(O(T)) \leq 2\|O\|^2_\infty (1 + 2h^2_E),
\]

where

\[
h_E = \theta_M(T)\|H_M\|_\infty + \theta_C(T)\|H_C\|_\infty
\]

with \( \theta_M(T) = \int_0^T dt |\eta_M(t)|[1 − s(t)] \) and \( \theta_C(T) = \int_0^T dt |\eta_C(t)|s(t) \). This result follows from a bound on the variance \( \text{Var}(O(T)) \leq 2\|O\|^2_\infty \) and on the bias \( \Delta \langle O(T) \rangle \leq \int_0^T \|H_E(t)\|_\infty dt \leq h_E \).

More generally, the variance can be bounded in terms of the error operator according to

\[
\text{Var}(O(T)) \leq \|\Lambda(T)\|_\infty (\|\Lambda(T)\|_\infty \|O_{\rho_0}(T)\|^2_1 + \|O^2_{\rho_0}(T)\|_1).
\]

This follows directly from an application of the triangle inequality and the variant of Hölder’s inequality used in Eq. (47). A bound on the bias [the second term in Eq. (52)] is obtained in a similar manner to that of Eq. (51), namely,

\[
\left(\Delta \langle O(T) \rangle \right)^2 \leq \|\Lambda(T)\|_\infty 1^{2}_\infty \|O_{\rho_0}(T)\|^2_1.
\]

The MSE achieves a quadratic scaling in the norm of the error operator and therefore, the cumulant sum. This can be observed by incorporating Eq. (48) to explicitly express the MSE in terms of the norm of the cumulant sum. If \( \|\mathcal{C}_o(T)\|_\infty \leq 1 \) is assumed, the MSE scales as \( O(\|\mathcal{C}_o(T)\|^2_\infty) \), explicitly conveying a quadratic scaling.

### B. Approximation Ratio

The bound derived in Eq. (51) naturally leads to a bound on the absolute error in the approximation ratio. Specifically, we aim to bound the difference between the faulty and ideal approximation ratios for evolution dictated by the optimal ideal variational parameters. We show that the absolute error can be expressed in terms of the operator norm of the cumulant series \( \mathcal{C}_o(T) \). As a result, one may utilize this bound to quantify the effect of various types of precision errors on the approximation ratio. While we focus on a general bound here, subsequent sections consider specific canonical QAOA problems.

In the ideal scenario, an \( \epsilon \)-approximate QAOA achieves a performance guarantee with the variational parameters \((\gamma^*, \beta^*)\), or equivalently a total algorithmic runtime \( T^* \). The optimal approximation ratio \( \epsilon^* = F(\gamma^*, \beta^*)/C_{\text{max}} \) is expressed in terms of the average energy with respect to the problem Hamiltonian, \( F(\gamma^*, \beta^*) \), i.e., the typical QAOA objective function. This quantity is synonymous with Eq. (25) rescaled by \( C_{\text{max}} \), where \( O = H_C \) and the evolution has been evaluated at the optimal variational parameters.

In order to evaluate the error induced by \( H_E(t) \), and more specifically precision errors, consider the average energy resulting from the faulty dynamics \( \tilde{F}(\gamma^*, \beta^*) \). As can be anticipated, the average energy can be associated with a faulty approximation ratio \( \tilde{\epsilon}^* = \tilde{F}(\gamma^*, \beta^*)/C_{\text{max}} \). The upper bound on the absolute error \( |\Delta \epsilon^*| = |\tilde{\epsilon}^* − \epsilon^*| \) is

\[
|\Delta \epsilon^*| = \frac{1}{C_{\text{max}}} |\tilde{F}(\gamma^*, \beta^*) − F(\gamma^*, \beta^*)|
\]

\[
\leq \frac{1}{C_{\text{max}}} \|\Lambda(T^*) − 1\|_\infty \|H_C \rho_0(T^*)\|_1
\]

\[
\leq \frac{1}{C_{\text{max}}} \left(e\|\mathcal{C}_o(T^*)\|_\infty − 1\right) \|H_C \rho_0(T^*)\|_1,
\]

which follows directly from Eq. (51). In Sec. B 2, we further bound \( \|\mathcal{C}_o(T)\|_\infty \) for the error models considered in Sec. V, i.e., constant coherent errors and stochastic errors. These bounds are incorporated into Eq. (57) in Sec. VII, where specific QAOA problems are investigated.

### C. Error in Training

QAOA variational parameter training routinely involves a classical gradient descent algorithm. Gradients are either analytically determined or approximated via a finite difference approximation. In either case, parameters are updated based on estimates of the average energy using states time evolved by the QAOA with variational parameters acquired from previous iterations. Thus, the expectation value plays an important role in the effective training of the QAOA. When presented with a faulty QAOA, subject to systematic or environment noise sources, errors induced in the expectation value propagate to the gradient, and ultimately, impact parameter training.

Here, we aim to evaluate the impact of errors on the gradient by bounding the absolute error in the gradient of the expectation value. First, we focus on open-loop (offline) optimization protocols that utilize exact gradient expressions for parameter training. We define the error between the faulty and ideal gradients for a single variational parameter

\[
|\partial_{\gamma_j} \Delta \langle O(T) \rangle| = |\partial_{\gamma_j} \langle O(T) \rangle − \partial_{\gamma_j} \langle O(T) \rangle_0|.
\]

Applying the triangle inequality, the error can be separated...
into two terms as follows:

\[
|\partial_{\gamma_j} \Delta(O(T))| \leq |\text{Tr}[\{\partial_{\gamma_j} \Lambda(T)\} \rho_0(T)O]| \\
+ |\text{Tr}[\{\Lambda(T) - 1\} \partial_{\gamma_j} \rho_0(T)O]|. \tag{59}
\]

The first term is characterized by the differentiation of the error operator, while the latter includes a differentiation of the ideal time-evolved density operator. We bound each term individually, starting with

\[
|\text{Tr}[\{\partial_{\gamma_j} \Lambda(T)\} \rho_0(T)O]| \leq \|\partial_{\gamma_j} \Lambda(T)\|_\infty \|O\rho_0(T)\|_1 \\
\leq \lambda(T) \|O\rho_0(T)\|_1. \tag{60}
\]

The first inequality follows from an application of H"older’s inequality, while the second incorporates

\[
\|\partial_{\gamma_j} \Lambda(T)\|_\infty \leq e^{\|C(T)\|_\infty} \frac{e^2\|C(T)\|_\infty - 1}{2\|C(T)\|_\infty} \|\partial_{\gamma_j} C(T)\|_\infty \\
= \lambda(T) \tag{61}
\]

Briefly, this bound is obtained by expressing \(\Lambda(T)\) as an exponential of the cumulant series and using the general definition of a derivative of an exponentiated operator.

The second term in Eq. (59) can be bounded via H"older’s inequality and further simplified using the Liouville-Von Neumann equation. Namely, we find

\[
|\text{Tr}[\{\Lambda(T) - 1\} \partial_{\gamma_j} \rho_0(T)O]| \leq \|\Lambda(T) - 1\|_\infty \\
\times \|\hat{O}(T_j, T) [H_C, \rho_0(T_j)]\|_1, \tag{62}
\]

where \(\hat{O}(T_j, T) = Q_{p,j}OQ_{p,j}^\dagger\). Note that the second term in Eq. (62) follows from observing the connection between \(\|O^\dagger \partial_{\gamma_j} \rho_0(T)\|_1\) and the dynamical equation for the density matrix up to time \(T_j\), i.e., \(\partial_{\gamma_j} \rho_0(T) = -iQ_{p,j} [H_C, \rho_0(T_j)] Q_{p,j}^\dagger\). The supplement contains further elaboration on this bound and the others shown above.

Commonly, QAOA training leverages approximate gradient expressions in place of exact gradients. This most notably occurs when closed loop optimization protocols employ some degree of hardware-in-the-loop querying to train variational parameters. First order gradients are calculated via finite difference approximations and subsequently incorporated into stochastic gradient descent protocols, such as simultaneous perturbation stochastic approximation (SPSA) [46].

The training error in stochastic optimization protocols can be associated with the quality of the estimator, and thus, the MSE. In the asymptotic limit, this training error can be bounded by the expressions given in Sec. VI.A. Coherent errors yield an upper bound of

\[
\text{MSE}(O(T)) \leq 2\|H_C\|_\infty^2 \left(1 + 2h_E^2\right), \tag{63}
\]

on the training error, while the sum of Eqs. (55) and (56) with \(O = H_C\) specifies a general bound in terms of the error operator. Finite sampling effects are pertinent to approximate gradient-based algorithms. Contributing additional corrections to the above bounds, such effects are elaborated upon in the supplement.

### D. Distance Between Unitaries

Error measures expressed in terms of expectation values provide insight into the deleterious effects of noisy QAOA evolution with respect to the measurement of an observable. Although such measures are of practical importance for QAOA, from a theoretical perspective, it can be constructive to eliminate state- and observable-dependence and focus solely on the unitary propagator. Such distance measures are commonly used in optimized control [47] and Trotter error analysis [48] to achieve a holistic understanding of error propagation.

In an effort to obtain such an understanding for the precision error problem in QAOA, we examine the distance between the faulty QAOA and ideal QAOA evolution

\[
\Delta U = \|U(T) - U_0(T)\|_2. \tag{64}
\]

Note that \(\|A\|_2 = \sqrt{\text{Tr}A^\dagger A}\) defines the Frobenius norm. We bound this quantity, focusing particularly on the ensemble average-squared distance, i.e.,

\[
(\Delta U)^2 = \text{Tr}\left[2I - U_E(T) - U_E^\dagger(T)\right]. \tag{65}
\]

The above expression includes a noise averaged unitary that differs from that introduced in Eq. (28). Namely, the unitary is not generated by the effective Hamiltonian \(\hat{H}_{E,O}(t)\), but rather \(\hat{H}_E(t)\). This does not create any issues, however, as a cumulant expansion

\[
\overline{U_E(T)} = \text{Tr}_e e^{-i \int_0^T \hat{H}_E(t) dt} = e^{C(T)} \tag{66}
\]

can still be obtained, with \(C(T) = \sum_{k=1}^\infty (-i)^k C^{(k)}(T)/k!\). Denoting the cumulant terms rewritten in terms of \(\hat{H}_E(t)\). Incorporating the expansion into Eq. (64) leads to

\[
(\Delta U)^2 = \text{Tr}\left[2I - e^{C(T)} - e^{C(T)}\right] \leq 2 \text{Tr}\left(I - e^{C(T)}\right) \leq 2\|I - e^{C(T)}\|_\infty \leq 2(e^{\|C(T)\|_\infty} - 1). \tag{67}
\]

This bound possesses similar qualities to the bound obtained for the absolute error in expectation values, with the primary distinction being the presence of an observable-independent cumulant expansion.

### VII. NUMERICAL SIMULATIONS AND EMPIRICAL BOUNDS

In this section, we discuss the effect of precision errors on two specific examples of QAOA circuit implementations. In the first example, we examine the Grover’s Search Algorithm using QAOA [32] and second we consider obtaining the ground state of the nearest-neighbor Ising model on a ring [33], which also can be mapped to the 2-SAT problem.
on a ring [49]. While the Grover algorithm provides an example of a QAOA algorithm with an analytical solution for the optimal circuit depth and angles, the circuit depth increases exponentially with the number of qubits. In the second example, we obtain the ground state of the Ising problem using a number of QAOA layers that scales linearly in the number of qubits.

We numerically simulate precision errors and investigate its effects on the accuracy of the implementation of the algorithm. The errors \( \eta_{C,M} \) modify the QAOA phases, \( \gamma \rightarrow \gamma(1 + \eta_C) \) and \( \beta \rightarrow \beta(1 + \eta_M) \); see Eq. (19) for definitions of \( \eta_n \). These errors are drawn from a normal distribution: \( \eta_n \in N(\eta, \sqrt{\Gamma}) \) with \( \eta \) being the mean, \( \Gamma \) the variance of the normal distribution, and \( \nu \equiv M, C \). The stochastic error case reduces to \( \eta_n \in N(\eta, \sqrt{\Gamma}) \) \([\Gamma_n(t - t') = \Gamma \delta(t - t')\), i.e., no correlation between subsequent errors in \( \gamma_k \) and \( \beta_k \)\] and the coherent error cases reduce to \( \eta_n \in N(\eta, 0) \) \([\Gamma_n(t - t') = \eta^2\), i.e., constant errors in all \( \gamma_k \) and \( \beta_k \)\]. We average our results over a 1000 noise realizations for the case of the stochastic error, and 1 realization for the coherent case.

In order to numerically extract the loss in accuracy of the computation with increasing precision error, we model the noise-averaged expectation value using a decay function dependent on the number of layers and noise parameters,

\[
\langle H_C \rangle \approx \langle H_C \rangle_0 S_H(p, \mu, \sigma) \tag{68}
\]

where \( \langle H_C \rangle_0 \) corresponds to the noiseless expectation value. The analytical form of the decay function, \( S_H(p) \) depends on the nature of the precision errors (stochastic vs coherent) and the number of layers. In the following, we show that the expectation value decays exponentially away from the noiseless value, and this behavior is generic.

### A. Grover’s Search Algorithm

The Grover search problem involves finding a marked element \( |m\rangle \) in an unsorted database consisting of \( N = 2^n \) elements, where \( n \) is the number of qubits required to represent the database. For a marked state of \( |m\rangle = |0\cdots0\rangle \), we have the following phase and mixer Hamiltonians,

\[
H_C = |0\cdots0\rangle \langle 0\cdots0| = \bigotimes_{i=1}^{n} \frac{I + Z_i}{2} \tag{69}
\]

\[
H_M = \sum_{i=1}^{n} X_i \tag{70}
\]

Clearly, the phase Hamiltonian corresponds to a fully connected Ising Hamiltonian. Unlike the cost function for typical optimization problems, the cost function has all possible tensor product of \( Z_i \) operators. However, note that \( ||H_C|| = 1 \). Also, note that the Hamiltonian encoding the cost function, \( H_C \), while clearly hermitian, is non-invertible.

Typically, for the Grover search problem, each layer of the QAOA ansatz has the following structure [32],

\[
|\psi_{out}\rangle = [W(\gamma, \beta)]^p |+\rangle^n \tag{71}
\]

\[
= \prod_{k=1}^{2p} \left[ e^{-i \beta} \sum_i J_{ij} e^{-(1)^{k+1} i \gamma H_C} \right] |+\rangle^n \tag{72}
\]

with \( \gamma = \pi \) and \( \beta = \frac{\pi}{n} \). The QAOA algorithm applies the unitary \( W(\gamma, \beta) \) \( p \) times. The quantum circuit for a single layer, is shown in Fig. 2(a). We simulate the Grover QAOA circuits using the built-in simulator in Cirq [50]. The expectation value of the cost function, which corresponds to the probability of measuring the marked state, oscillates as a function of the number of layers \( p \). In Fig. 2(b) we show this probability as a function of \( p \). We obtain an expression for the optimal number of layers for Grover search as a function of the \( n \) by numerically fitting the first maximum in \( \langle H_C \rangle \).

\[
p^* (n) \approx a_1 2^{n/2} + a_2 \tag{73}
\]

\( a_1 = 0.32, a_2 = 0.24 \); see Fig. 2 (b). This scaling of the optimal number of layers for Grover is in line with the theoretical analysis in Ref. [32], namely \( p^* \sim O(\sqrt{N}) \).

The Grover cost function, \( H_C \) is not invertible, which poses an issue for the cumulant expansion and the definition of the error operator \( \Lambda(T) \) [Eq. (27)]. However, we can be rewrite \( H_C \) as,

\[
H_C = \frac{1}{N} I + O, \tag{74}
\]

where, \( O = \frac{1}{N} \sum_{\{A_i\}} A_1 \otimes \cdots A_n \tag{75} \)

where, \( A_i \in \{I, Z_i\} \) and \( O \) is invertible and traceless. Using this representation of \( H_C \), we can employ the bounds derived
in the previous sections to bound the expectation value of $H_C$. We consider the effects of stochastic and coherent noise on the measured expectation value of the cost Hamiltonian. We expect a noise-averaged expectation value of,

$$\langle H_C \rangle = \frac{1}{N} I + \langle O \rangle = \frac{1}{N} I + \langle O \rangle_0 S_O(p^*)$$

$$= \frac{1}{N} \left[ 1 - S_O(p^*) \right] + \langle H_C \rangle_0 S_O(p^*)$$

where $S_O(p^*)$ is the decay function specific to the operator $O$.

In the limit of weak noise, we fit the measured expectation value to the following phenomenological decay function,

$$\text{Weak noise limit:} \langle H_C \rangle \approx \langle H_C \rangle_0 S_O(p^*)$$

In Figs. 3 and 4, we numerically study the effects of precision error. Fig. 3 (a) and (b), we study the role of stochastic ($\eta = 0$, $\Gamma \neq 0$) and coherent errors ($\eta \neq 0$, $\Gamma = 0$) as a function of number of layers $p$ with the weak noise limit corresponding to $\sigma \ll 1$ and $\mu \ll 1$ respectively. We numerically obtain a fit for the decay function in Fig. 4 (a) and (b) in the weak noise limit (shown as dashed lines) for the optimal Grover QAOA algorithm,

$$S_O(p^*) \approx \langle H_C \rangle_0 \exp[-\chi_O(p^*)]$$

$$= \left\{ \begin{array}{ll}
-\Gamma p^*/\Gamma_s, & \text{[stochastic error]}

-\eta^2 p^2/\eta_c^2, & \text{[coherent error]}
\end{array} \right.$$ 

where $\sqrt{\Gamma_s} \approx 0.47$ and $\eta_c \approx 3.46$. Clearly, the calculated expectation value decays exponentially with precision error scale. Note that the implementation of the optimal Grover algorithm requires $p_{\text{opt}}$ that grows with the problem size. Conversely, our results imply that for an accurate implementation of the optimal Grover QAOA circuit, the precision errors need to be bounded from above, $\Gamma \ll \Gamma_s/p^*$ and $\eta \ll \eta_c/p^*$.

We compare the numerically obtained $S_O(p^*)$ to the decay functions derived via the cumulant expansion and its associated bounds. Specifically, we focus on the absolute error in the expectation value, and compare the numerically obtained absolute error to three different cumulant calculations for the case of stochastic precision errors. First, we study the absolute error using the second cumulant approximation to the error operator. We compare this approximation to both numerically and analytically obtained bounds to assess their efficacy in the Grover’s search setting. A summary of the analysis is shown in Fig. 5.

The absolute error is estimated by approximating the expectation value dynamics using the second-cumulant-truncated error operator. The cost Hamiltonian is expressed as a sum of terms [see Eq. (74)], which naturally leads to an expectation value of the form

$$\langle H_C \rangle = \sum_{O \in \{I, N, O \}} \text{Tr} \left[ \Lambda_O(T) \rho_S(T) O \right].$$

By truncating each error operator as $\Lambda_O(T) \approx e^{-C_O^2(T)/2}$, we obtain an expression for the approximate dynamics of the expectation value of $H_C$. Subsequently, an estimate of the absolute error is obtained by substituting Eq. (81) into Eq. (50).

The approximate dynamics are compared against bounds on the absolute error derived in Section VI A. Specifically, we...
consider a variant of the upper bound given in Eq. (51) for non-invertible observables:

\[
|\Delta \langle H_C \rangle| \leq \| \sum_{\mathcal{O} \in \{I/N, O\}} \mathcal{O} \Lambda_0 (T) - H_C \|_{\infty};
\]  

(82)

see Appendix B 1 for further details. We consider a numerical evaluation of this upper bound by truncating the error operator to second order and calculating the bound numerically. In addition, we estimate the bound analytically via

\[
|\Delta \langle H_C \rangle| \leq \sum_{\mathcal{O} \in \{I/N, O\}} \| \mathcal{O} \|_\infty \left( e^{\| C^{(2)} (T) \|_{\infty} / 2} - 1 \right)
\]

\[
\leq \min \left[ 2 \left( e^{4p^* \Gamma} - 1 \right), \| H_C \|_{\infty} \right],
\]  

(83)

where the upper bound on the second cumulant is

\[
\| C^{(2)} (T) \|_{\infty} \leq 8p^* \Gamma \sim 8\pi \sqrt{NT}.
\]  

(84)

The bound on the \( \| C^{(2)} (T) \|_{\infty} \) holds for each constituent observable \( \mathcal{O} \) defined for Grover’s search. Note that the upper bound expression is clipped at a maximum value of \( \| H_C \|_{\infty} \) as this represents the physical bound on possible QAOA solutions.

The results of the cumulant comparison are summarized in Fig. 5(a) for Grover’s Search. The second cumulant approximation (purple circles) agrees well with the numerically obtained fit in the weak noise regime. Analytical bounds (green dashed lines) provide a modest representation of precision error-dependence, deviating more substantially as the variance increases. Numerical bounds (blue triangles) and approximate dynamics follow similar trends, indicating that as the variance increases, higher order cumulants likely contribute more significantly to the dynamics.

In addition, we evaluate the effect of precision errors by calculating the average distance between the faulty and ideal unitary. Our results are summarized in in Fig. 3 (c) and (d) which show the stochastic and coherent error case, respectively. In the limit of weak noise, the average distance for

\[
\| U - U_0 \|_{\infty} \propto \begin{cases} \sqrt{\Gamma p}, & \text{[stochastic error]} \\ \eta p, & \text{[coherent error]} \end{cases}
\]  

(85)

The average distance between the noisy and noiseless unitary grows with the number of layers of the QAOA circuit.

Finally, let us discuss the limit of large stochastic noise, \( \Gamma p^* \gg 1 \). Assuming that the output of a noisy QAOA evolution is a completely random state, one would expect that \( \langle H_C \rangle \sim 1/N = 1/2^n/2 \). However, in this limit, the noisy-QAOA does not give rise to uniformly random states. Instead, we observe saturation in the expectation value that scales with problem size according to

\[
\langle H \rangle_{\text{sat}} \propto 2^{\xi n},
\]  

(86)

where \( \xi \approx 0.67 \). The approach to the saturation value is governed by a power-law behavior,

\[
\langle H_C \rangle - \langle H_C \rangle_{\text{sat}} \propto \frac{1}{(\Gamma p^*)^\alpha},
\]  

(87)

where \( \alpha \approx 1.61 \). The details for the fit to the various parameters, \( \sigma_s, \mu_c, \xi \) and \( \alpha \) are provided in Appendix D.

### B. Ising Instances

Essentially all combinatorial optimization problems may be cast as two-body Ising Hamiltonians [51, 52] of the form

\[
H = \sum_{ij} J_{ij} Z_i Z_j + \sum_i h_i Z_i.
\]

It is therefore natural to study the effects of finite precision errors on problems of this type. A paradigmatic model of the above form, which is therefore convenient to analyze, is the one-dimensional, nearest-neighbor Ising model. The problem Hamiltonian of this model for \( n \) qubits, is given by

\[
H_C = \sum_{i=1}^{n} Z_i Z_{i+1},
\]  

(88)

where the summation is taken over neighboring spins on a 1D chain with periodic boundary conditions, i.e., \( Z_{n+1} \equiv Z_1 \). For instance, this particular cost function can be used to find the solutions for 2-SAT on a ring [49].

The cost function Hamiltonian for this Ising model is maximized by a state that is any superposition of all qubits in the \( |0 \rangle \) or \( |1 \rangle \) state. In fact, by restricting to a given parity sector satisfying \( \prod_{i=1}^{n} X_i = 1 \), this cost function can be used in the QAOA setting as an algorithm for preparing a GHZ state [12, 33],

\[
GHZ = \frac{1}{\sqrt{2}} \left( |0 \rangle^\otimes n + |1 \rangle^\otimes n \right).
\]  

(89)

We follow Ref. [33] to use the QAOA ansatz for preparing the GHZ state. Note that we have flipped the sign of the cost Hamiltonian compared with Ref. [33] as we have recast the objective from minimization to maximization. A similar procedure is utilized for the mixer Hamiltonian as well; thus, \( H_M \) is defined by Eq. (70). It was shown that the optimal number of layers to prepare a GHZ state scales polynomially with the number of qubits. In particular, \( p^* = n/2 \), with the following
structure for the variational circuit ansatz:

\[
U_k(\gamma_k, \beta_k) = e^{-i\beta_k \sum_i X_i} e^{-i\gamma_k H_C},
\]

\[
|GHZ\rangle = \prod_{k=1}^{n/2} U_k(\gamma_k, \beta_k) |+\rangle^\otimes n.
\]

The optimal parameters to generate GHZ states for different system sizes are taken from Ref. [33] (Appendix A) to test the susceptibility of the state preparation to precision errors.

We numerically model the noise-averaged expectation value as a function of the noise strength and determine scaling fits for the GHZ state preparation problem. A summary of the results are shown in Fig. 6 for both stochastic and constant precision errors. Numerical simulations are performed using 1000 realizations for stochastic errors and one realization for the coherent error case, using code adapted from Refs. [53, 54]. Data collected over various system sizes are utilized to extract an effective scaling of the average expectation value as a function of the noise parameter and optimal QAOA order, which is a function of system size. In the weak noise limit, the expectation value decays exponentially according to

\[
\langle H_C \rangle \approx \langle H_C \rangle_0 \exp \left[-\chi_{H_C}(p^*)\right]
\]

where,

\[
\chi_{H_C}(p^*) \approx \begin{cases} -\Gamma p^*/\sqrt{\Gamma_\sigma}, & \text{[stochastic error]} \\ -\eta^2 p^*/\eta, & \text{[coherent error]} \end{cases}
\]

with \(\sqrt{\Gamma_\sigma} = 0.51\) and \(\eta = 0.46\). Note that in the case of large stochastic noise, \(\Gamma p^* \gg 1\), we see that the noise-averaged cost function approaches zero in a power-law, \(\langle H_C \rangle \propto 1/(\Gamma p^*)^\alpha\), with \(\alpha = 1.46\).

The scaling fit is compared against the cumulant expansion and its associated bounds in Fig. 5(b). In general, this comparison will require a similar approach to that used in the Grover’s search analysis. Namely, we must calculate the approximate dynamics and bounds by summing the contribution of each constituent \(Z_i Z_{i+1}\) to the faulty expectation value of \(H_C\). However, for the case considered here, \(n = 10\), the inverse of \(H_C\) does exist. Thus, no summation is required in the calculation of \(\langle H_C \rangle\) and the bounding expression given in Eq. (51) directly applies.

Overall, we find good agreement between the second cumulant dynamics and the fit in the weak noise regime. Numerically calculated bounds (blue triangles) and second cumulant dynamics (purple circles) track closely, indicating that the bound is tight for this particular problem class. The analytical bound (green dashed line) is determined by utilizing

\[
\|C(2)_{H_C}(T)\|_\infty \leq n\Gamma \left(1 + \frac{3n}{\sigma_{\min}^2}\right) \sum_{i=1}^{n/2} (|\gamma_i| + |\beta_i|),
\]

where \(\sigma_{\min}\) is the minimum singular value of \(H_C\), to further bound Eq. (51). As expected, this additional bound allows for general behavior of the dynamics to be captured, but does not provide a tight error bound, most notably for increasing variance.

VIII. DIGITIZATION: AVOIDING PRECISION-INDUCED ERRORS IN QAOA

Let us now discuss some general strategies to mitigate precision-induced errors in QAOA. It is quite clear that precision errors in QAOA variational parameters (whether coherent or stochastic) can be quite detrimental to the performance of the algorithm. Constant coherent over- or under-rotations can be straightforwardly accounted for by updating the optimized angles exactly by the coherent error. However, stochastic errors are harder to deal with using this approach as they are random in nature. Below, we propose a method for mitigating the effects of stochastic precision errors that involves digitizing the angles in a binary representation.

From the analysis of the previous section, it is clear that for the implementation of a \(p\)-layer QAOA algorithm to achieve a desired accuracy threshold for the expectation value of the cost function, \(|\Delta \langle H_C \rangle|/\langle H_C \rangle_0 = 2^{-\epsilon} (\epsilon > 0)\), the maximum precision error permitted in the variational parameters is \(\eta H, \eta_C \propto \sqrt{\epsilon/p}\), where the constant pre-factor depends on the details of the problem. Clearly, the precision requirements depend on the problem type. For implementing the Grover search algorithm, the optimal number of layers necessary scales exponentially with the number of qubits, and thus the required precision must increase at least as \(2^{-n/2}\).

We expect this scenario to be fairly generic for QAOA circuits. While short finite-depth QAOA circuits are currently being investigated for NISQ applications in optimization, it is expected that for computationally hard optimization problems it is not unreasonable to expect the number of layer to increase at least super-polynomially with problem size [55].

The digitization of the QAOA angles may allow us to avoid errors induced by precision errors (at the cost of increasing circuit depth). Let us start by discussing how to implement the QAOA cost function and mixer unitaries up to a certain precision error in the variational parameters. Consider an \(N\gamma\),

FIG. 6. Preparation of the GHZ state using the QAOA algorithm with the optimal number of layers \((p = p^*)\) as a function of noise strength \((a) \sigma\) or \((b) \mu\) corresponding to noise generated from a normal distribution \(N(\mu, \sigma)\). (a) Average expectation value (open circles) of the cost function for the case of stochastic error \((\mu = 0)\), the lines (dashed,dash-dot) indicate the scaling fit to the data. (b) Expectation value (open triangles) as a function of the strength of coherent error \((\sigma = 0)\) and the corresponding scaling fit shown as a dashed line.
bit implementation of the angle $\gamma$ in the unitary $\exp(-i\gamma H_C)$,
\[ \gamma \approx \frac{2\pi}{2^{N_2}} \sum_{j=1}^{N_2} A_j 2^j \]
(95)
where, $A_j \in \{0, 1\}$ is the $j$-th bit of the integer closest to $\frac{\gamma}{2\pi}$. Note that the $N_2$ bit implementation has a precision error $\sim 2^{-N_2}$. The unitary may be implemented as,
\[ U_C(\gamma) = \exp(-i\gamma H_C) = \prod_{j=1}^{N_2} \left[ \exp\left(\frac{2\pi}{2^{2N_2}} 2^j H_C\right) \right] A_j \]
(96)
A similar analysis can also be extended to the mixer unitary, $U_M(\beta)$. Now, instead of viewing QAOA as an analog quantum algorithm with continuously variable angles, one only needs to implement certain below-threshold building block unitaries, $U_{Cj} = \exp\left(\frac{i\gamma}{2^{2N_2}} 2^j H_C\right)$ and $U_{Mj} = \exp\left(\frac{i\gamma}{2^{2N_2}} 2^j M_H\right)$, with $j = 0, 1 \cdots, N_\gamma$ or $N_\beta$. (The term “below-threshold” denotes unitaries that can be implemented with a precision error of at most $\sim \sqrt{\epsilon/pN_\gamma \beta}$.) Combined with Eq. (96), both the unitary operators can be implemented within the desired precision at the cost of increasing the circuit depth $\sim pN_\gamma C_\gamma + pN_\beta C_\beta$, where $C_\gamma, C_\beta$ are the maximum circuit depth necessary to realize one of the unitary operations $U_{Cj}, U_{Mj}$.

The implementation of a $p$-layer QAOA circuit with the desired cost-function evaluation accuracy of $2^{-\epsilon}$ requires a precision error of at most $\sim \sqrt{\epsilon/p}$. From the analysis above, an $N_\gamma, \beta$ bit implementation of the variational parameters introduces a precision error $\sim 2^{-N_\gamma, \beta}$ and thus, we conclude that the effects of precision errors larger than $\sqrt{\epsilon/p}$ may be mitigated by $N_\gamma, \beta \propto \log_2(\sqrt{\epsilon/p})$. As a result, mitigating precision errors requires an increase in the circuit depth that is logarithmic with the desired $\epsilon$ and $p$. This would imply that for problems that are hardest for QAOA, where $p \propto 2^n$ (such as Grover’s Search), one can avoid precision-errors at a cost polynomial ($\propto n$) in the number of qubits.

\section{IX. CONCLUSIONS}

In this study, we investigated the effect of precision errors, or the misspecification of variational parameters, on QAOA performance. We provided insight into the extent of their harm on parameter training and performance guarantees. Utilizing concepts from quantum control theory, we analytically estimated the contribution of precision errors to QAOA dynamics. This approach enabled the development of bounds on expectation values, approximation ratios, and parameter training error. Through our analysis, we found that any fixed precision implementation of QAOA will realize performance guarantees that greatly differ from the idealized setting. In particular, we found an exponential reduction in success probability with increasing QAOA order and error magnitude. Numerical studies of the QAOA variant of Grover’s search and the one-dimensional transverse-field Ising model conveyed these observations and provided verification for our analytical estimates.

Despite the detrimental nature of precision errors on QAOA performance, we showed that it is possible to mitigate such errors by digitizing the variational parameters. Each constituent QAOA evolution operator was then expanded into a product of operators whose circuit depth is determined by the desired precision. Provided that the decomposition is expressed in terms of evolution operators that can be implemented with a greater precision than that required by the algorithm, one can successfully achieve a desired precision accuracy for QAOA.

While our study focused on precision errors, the analytical approach inherently extends to a far greater class of noise models. Specifically, one can utilize this approach to study the effect of spatial and temporally correlated noise on QAOA and variational algorithms more generally. In fact, the bounds we developed in this study naturally extend to these more general scenarios without further manipulation. In summary, we view this framework as providing a unique perspective on variational quantum algorithms and as a constructive tool for predicting faulty algorithm performance and facilitating the development of noise-robust variants.

\section{X. ACKNOWLEDGEMENTS}

Gregory Quiroz, Paraj Titum, Pavel Lougovski, Kevin Schultz, Eugene Dumitrescu, and Itay Hen acknowledge funding from the U.S. Department of Energy (DOE), Office of Science, Office of Advanced Scientific Computing Research (ASCR) Quantum Computing Application Teams program, under fieldwork proposal number ERKJ347. Phillip Lotshaw was supported at ORNL by the Defense Advanced Research Projects Agency, Defense Science Office under contract HR001120C0046 with Georgia Tech Research Institute.

\section{Appendix A: Alternative Toggling Frame}

Faulty QAOA evolution dictated by Eq. (18) can be factorized according to Eq.(23), or equivalently, $U(T) = U_0(T) \tilde{U}_E(T)$. The latter expression follows a very typical interaction picture representation in which the time evolution operator
\[ \tilde{U}_E(T) = T_+ e^{-i \int_0^T \tilde{H}_E(t) \, dt} \]  
(A1)
is governed by $\tilde{H}_E(t) = U_0^\dagger(t) H_E(t) U_0(t)$. Equivalence, between the two representations can be shown in a very straightforward manner by observing that
\[ U(T) = U_0(T) \tilde{U}_E(T) = U_0(T) \tilde{U}_E(T) U_0^0(T) U_0(T) = U_E(T) U_0(T); \]  
(A2)
thus, matching the expressions in Eq. (22)-(24).
Both representations lend themselves to cumulant-based expansions of expectation values with the latter inducing a time-dependence on the observable. Following the approach utilized in the main text, the expectation value of an observable $O$ can be expressed in terms of $U(T) = U_0(T)U_E^\dagger(T)$ as

$$
\langle O \rangle = \text{Tr}[O\rho(T)] = \text{Tr}(OU_0(t)U_E(t)\rho(0)(U_E^\dagger(t))^\dagger U_0^\dagger(t))] = \text{Tr}[\Lambda(t)\rho(0)\tilde{O}(t)].
$$

(A3)

The error dynamics are determined by the operator $\Lambda(t) = \tilde{O}_{-1}(t)U_E^\dagger(t)\tilde{O}(t)U_E(t)$, which includes a conjugation by $\tilde{O}(t) = U_0^\dagger(t)OU_0(t)$, the observable in the interaction picture with respect to the noiseless QAOA evolution. As can be seen from comparing the above expressions to those surrounding Eq. (27), the distinction lies in the definition of the observable. Note that while both representations are analogous, the additional time-dependence can result in unnecessary complexity in analytical or numerical investigations. Hence, we select the former for this study.

Appendix B: Additional Bounds

1. Absolute Error Bound: Non-Invertible Observable

Consider the case where an observable $O$ is non-invertible, however, it can be expressed as a sum of invertible operators $O_i$: $O = \sum_i O_i$. A trivial example of an operator expansion of this type is the $n$-qubit Pauli basis for an operator $O \in SU(n)$. Of course, this particular example is relevant to the Grover problem and Ising-type problem considered in the main text and, more generally, in the context of variational quantum algorithms.

Utilizing this expansion for $O$, the expectation value of the observable can be written as

$$
\langle O(T) \rangle = \sum_i \text{Tr}[\rho(T)O_i] = \sum_i \text{Tr}[\Lambda_i(T)\rho_S(T)O_i],
$$

(B1)

where $\Lambda_i(T) = O_i^{-1}U_E^\dagger(T)O_iU_E(T)$. Going a bit further and leveraging the cumulant expansion,

$$
\langle O(T) \rangle = \sum_i \text{Tr}[e^{C_O(T)}\rho_S(T)O_i],
$$

(B2)

with $C_O(T)$ representing the cumulant expansion containing the $j$th observable.

The above expressions can be employed to derive bounds on the various quantities examined in Sec. VI of the main text. For example, consider the absolute error in the expectation value $\Delta(O)$.

An upper bound on this quantity can be derived following a similar approach to that of Sec. VI A to yield

$$
|\Delta(O(T))| = |\text{Tr} \left( \sum_i O_i \Lambda_i(T) - O \right) \rho_S(T)|
\leq \sum_i \| O_i \Lambda_i(T) - O \|_1 \left\| \rho_S(T) \right\|_1
\leq \sum_i \left\| O_i \right\|_\infty \left\| \Lambda_i(T) - I \right\|_\infty
\leq \sum_i \left\| O_i \right\|_\infty \left(e^{\|\Lambda_i(T)\|_\infty} - 1 \right).
$$

(B3)

Here, the variant of Hölder’s inequality from Sec. VI A is used to obtain (1). The trace-norm of a density matrix is unity and thus, the resulting bound in (2) is solely dependent upon the infinity norm of a term proportional to the accumulated error resulting from $\Lambda_i(T)$. In addition to the property of density matrix, we apply the triangle inequality and sub-multiplicativity to obtain (2). By expressing the error operator as a cumulant expansion and bounding the error term as in Sec. C1, we obtain the final bound in (3).

2. Bounds on Cumulants

Many of the bounds derived above share a common feature: dependence on the operator norm of the error operator. Through Eq. (48), we find that the error dynamics can be bounded in terms of the operator norm of the cumulant series $C_O(T)$. When the noise is sufficiently weak, the cumulant series is well-approximated by the leading order cumulant expression, i.e., $C_O^{(1)}(T)$ or $C_O^{(2)}(T)$ in the case of constant coherent errors or zero-mean, Gaussian stochastic errors, respectively. Here, we provide bounds on the first and second cumulant which can be used to further upper bound the expressions presented above.

Using the expressions derived in Sec. V for the first and second cumulants, bounds are obtained for each noise scenario. In the case of constant coherent errors, the first cumulant [Eq. (32)] can be shown to be bounded according to

$$
\left\| C_O^{(1)}(T) \right\| \leq \sum_{\mu=M,C} \eta_\mu \sum_j g_{\mu j}^p \left\| H_\mu \right\| (1 + \left\| O^{-1} \| \| O \|) .
$$

(B4)

Utilizing similar techniques, such as the triangle inequality, sub-multiplicativity, and unitary invariance, a similar expression can be obtained for the second cumulant [Eq. (33)]:

$$
\left\| C_O^{(2)}(T) \right\| \leq \left\| I_3(T) \right\| + 2 \left\| I_5(T) \right\|
\leq \sum_{\mu=M,C} \Gamma_\mu \sum_j g_{\mu j}^p \left\| H_\mu \right\| (1 + 3 \left\| O^{-1} \| \| O \|) .
$$

(B5)

For brevity and without loss of generality, the subscript denoting the operator norm has been dropped and bounds are displayed in terms of a generic unitary invariant norm.
Intuitively, bounds on the first and second cumulant scale with mean and variance, respectively. Both bounds are further characterized by the accumulative time over which each QAOA constituent Hamiltonian contributes to the evolution. Additional dependence on the norm of the of the constituent Hamiltonians and the observable $O$ and its inverse are also observed. The latter emerging via integral expressions conjugated by $O$.

### 3. Bounds on Gradients of Cumulants

Here, we include bounding expressions for $\|\partial_{\gamma_k} C_O^{(n)}(T)\|$ given their relevance to the gradient bound derived in Eq. (61). While we focus of the derivative with respect to $\gamma_k$, we note that equivalent expressions for the derivative with respect to $\beta_k$ follow directly. The primary difference being a change of QAOA Hamiltonian from $H_C$ to $H_M$ at the appropriate locations.

In order to bound the derivatives of the cumulants, we begin by noting that

$$\partial_{\gamma_k} Q_{p,j+1} = \begin{cases} -iQ_{p,k} HC Q_{k-1,j+1} & : k \geq j + 1 \\ 0 & : \text{otherwise} \end{cases}.$$  \hfill (B6)

This expression can be utilized to bound on the derivative of the rotated-frame Hamiltonians as

$$\|\partial_{\gamma_k} (Q_{p,j+1} HC Q_{p,j+1}^\dagger)\| \leq 2 \|H_C\|^2,$$  \hfill (B7)

for $k \geq j + 1$, which in turn can be employed to bound both cumulants. The first cumulant is bounded according to

$$\|\partial_{\gamma_k} C_O^{(1)}(T)\| \leq (1 + \|O^{-1}\|\|O\|) \|H_C\| \times \left[ 2\eta_C \|H_M\| \sum_{j=1}^{k-1} |\beta_j| \right. \right.$$  

$$\left. + \eta_M \left( 1 + 2 \|H_C\| \sum_{j=1}^{k-1} |\gamma_j| \right) \right],$$  \hfill (B8)

where the sums over $j \leq k - 1$ arises from Eq. (B6). The derivative of the second cumulant is bounded as

$$\|\partial_{\gamma_k} C_O^{(2)}(T)\| \leq \|\partial_{\gamma_k} I_1(T)\| + \|\partial_{\gamma_k} I_2(T)\| + 2|\partial_{\gamma_k} I_3(T)\| \leq \|\partial_{\gamma_k} I_1(T)\| \left( 1 + \|O^{-1}\|\|O\| \right) + 2\|\partial_{\gamma_k} I_3(T)\|,$$  \hfill (B9)

in terms of its components integrals. In turn, these integrals can be further bounded in terms of the variational parameters and the norm of the mixer and problem Hamiltonian. The first integral bound is

$$\|\partial_{\gamma_k} I_1(T)\| \leq \Gamma_M \|H_C\|^2 \left( 1 + 2 \|H_C\| \sum_{j=1}^{k-1} |\gamma_j| \right) + 2\Gamma_C \|H_M\|^2 \|H_C\| \sum_{j=1}^{k-1} |\beta_j|,$$  \hfill (B10)

while the derivative of the third integral is bounded by

$$\|\partial_{\gamma_k} I_3(T)\| \leq \|O^{-1}\|\|O\| \|H_C\| \times \left[ \Gamma_C \|H_C\| \left( 1 + 4 \|H_C\| \sum_{j=1}^{k} |\gamma_j| \right) \right. \right.$$  

$$\left. + 4\Gamma_M \|H_M\|^2 \sum_{j=1}^{k-1} |\beta_j| \right].$$  \hfill (B11)

### 4. Finite Sampling Effects

Finite sampling of expectation values will result in estimates that deviate from the asymptotic, infinite sample mean. As such, we account for finite sampling effects in the bounds derived in Sec. VI by modifying the estimator in accordance with the Central Limit Theorem to

$$\langle O \rangle_{\text{est}} = \langle O \rangle \pm \frac{1}{\sqrt{N_s}} \sqrt{\text{Var}(O)}.$$  \hfill (B12)

By definition, this assumes that the sampling statistics of the expectation value are Gaussian. We make use of Eq. (B12) below to derive finite sampling corrections to the bounds on absolute error and mean-squared error.

**a. Absolute Error**

Here, we attain a bound on the absolute error between the finitely sampled noisy expectation value $\langle O \rangle_{\text{est}}$ and the asymptotic noiseless expectation value $\langle O \rangle$. Following Eq. (50), this quantity is defined as

$$|\Delta \langle O(T) \rangle_{\text{est}}| = \left| \langle O(T) \rangle_{\text{est}} - \langle O(T) \rangle \right|.$$  \hfill (B13)

Using Eq. (B12), this expression can be bounded by the sum of the asymptotic bound for $|\Delta \langle O(T) \rangle|$ [Eq. (51)] and a correction term proportional to $1/\sqrt{N_s}$. More concretely, it can be shown that

$$|\Delta \langle O(T) \rangle_{\text{est}}| \leq \frac{1}{\sqrt{N_s}} \sqrt{\text{Var}(O)},$$  \hfill (B14)

where

$$\kappa(O, T) = \|\Lambda(T)\|_{\infty} \|O^2 \rho_0(T)\|_{1} + \|\Lambda(T)\|_{\infty} \|O \rho_0(T)\|_{2}^2$$  \hfill (B15)

results from upper bounding the variance of $O$.

**b. MSE**

Following Eq. (52), the MSE in the finite sampling regime is defined as

$$\text{MSE}(O(T))_{\text{est}} = \text{Var}_{\text{est}}(O(T)) + \left( \Delta \langle O(T) \rangle_{\text{est}} \right)^2,$$  \hfill (B16)
where the estimator is now designated as the finitely sampled expectation value given in Eq. (B12). Similar to the bound on the bias developed in Eq. (B14), the variance can be shown to be bounded by the sum of the asymptotic variance and $N_s$-dependent corrections, i.e.,

$$\text{Var}_{\text{est}}(O) = \langle O^2 \rangle_{\text{est}} - \langle (O)_{\text{est}} \rangle^2$$

$$\leq \text{Var}(O) + \frac{1}{N_s} \text{Var}(O)$$

$$\quad + \frac{1}{\sqrt{N_s}} \left[ \sqrt{\text{Var}(O^2)} + \langle O \rangle \sqrt{\text{Var}(O)} \right]. \quad (B17)$$

The relative difference between the finitely sampled and asymptotic variance can be further bounded according to

$$\text{Var}_{\text{est}}(O) - \text{Var}(O) \leq \frac{1}{N_s} \kappa(O, T) + \frac{1}{\sqrt{N_s}} \left[ \kappa(O^2, T) + 2 \|A(T)\|_\infty \|O\rho_0(T)\|_1 \sqrt{\kappa(O, T)} \right]. \quad (B18)$$

As expected by the definition of Eq. (B12), the dominant term scales according the $1/\sqrt{N_s}$.

Appendix C: Derivations

Here, we provide additional detailed regarding the bounds presented in Sec. VI.

1. Bound on Error Operator

In Sec. VI A, the operator norm of the operator $\Lambda(T)$ was said to be bounded by the exponentiated norm of the cumulant series. We prove this bound as follows:

$$\|\Lambda(T)\| = \|e^{C_O(T)}\|$$

$$= \| \sum_{m=0}^{\infty} C^m_O(T)/m! \|$$

$$\leq \sum_{m=0}^{\infty} \|C^m_O(T)/m! \|$$

$$\leq (1) \sum_{m=0}^{\infty} \|C^m_O(T)/m! \|$$

$$\leq (2) \sum_{m=0}^{\infty} \|C_O(T)\|^m/m!$$

$$= e^{\|C_O(T)\|}. \quad (C1)$$

Here, (1) triangle inequality and (2) sub-multiplicativity have been used to obtain the upper-bounding expressions.

2. Bound on MSE: Coherent Error

MSE can be expressed as a sum of the variance and the bias of the estimator. Here, we derive bounds on each term individually for the case of constant coherent errors. The resulting bounds are combined to obtain Eq. (53). Although the bound derived in the subsequent subsection encompasses constant coherent errors, we take an alternative approach that leverages results from Trotter error analysis to obtain a tighter bound for a case where ensemble averaging (i.e., a cumulant expansion) is not required.

A bound on the variance can be obtained by first noting that

$$\text{Var}(O(T)) = \text{Tr}[\rho(T)O^2] - \text{Tr}[\rho(T)O] \text{Tr}[\rho(T)O]^2$$

$$\leq |\text{Tr}[\rho(T)O]^2| + |\text{Tr}[\rho(T)O]|^2, \quad (C2)$$

where we have used $\text{Var}(O) = |\text{Var}(O)|$ and the triangle inequality. Employing Hölder’s inequality and sub-multiplicativity, this bound is further reduced to

$$\text{Var}(O(T)) \leq \|O^2\|_\infty + \|O^2\|^2_\infty$$

$$\leq 2\|O\|^2_\infty. \quad (C3)$$

Note that the first inequality makes use of $\|\rho(T)\|_1 = 1$.

The square-bias is bounded by first bounding the difference between the faulty and noiseless expectation values. Following a similar procedure to that utilized above,

$$|\Delta \langle O(T) \rangle| = |\text{Tr}[\rho(T) - \rho_0(T)]\rangle|$$

$$= |\text{Tr}[(U(T)O(U(T) - U_0(T))O_0(T))\rho(0)]|$$

$$\leq (1) \|U(T)O(U(T) - U_0(T))O_0(T)\|_\infty. \quad (C4)$$

We now introduce an additional term, $U_1(T)O_1(T) - U_1(T)O_0(T)$, within the operator norm and group terms associated with the difference $U_1(T) - U_0(T)$ and its Hermitian conjugate. Using unitary invariance of the norm and sub-multiplicativity, we obtain

$$|\Delta \langle O(T) \rangle| \leq 2\|O\|_\infty \|U(T) - U_0(T)\|_\infty, \quad (C5)$$

The norm of the difference between the faulty and noiseless propagators is further bounded through the definition of the time evolution dynamical equation. Namely,

$$\|U(T) - U_0(T)\|_\infty = \|\tilde{U}_{E}(T) - 1\|_\infty$$

$$\leq \| - i \int_0^T \hat{H}_E(t)U(t) dt\|_\infty$$

$$\leq \int_0^T \|H_E(t)\|_\infty dt$$

$$\leq h_E, \quad (C6)$$

where the triangle inequality and sub-multiplicativity are used to achieve the final bound in terms Eq. (54). Including Eq. (C6) into the Eq. (C5) leads to $|\Delta \langle O(T) \rangle|^2 \leq 4\|O\|^2_\infty h_E^2$, which in turn yields Eq. (53) when combined with Eq. (C3).

3. Bound on Exact Gradient

In Sec. VI C, we briefly derived a bound on the absolute error in the gradient of the expectation value for a single variational parameter. We elaborate on that derivation here, focusing specifically on the expressions related to norms comprised of the derivative of the error operator and the derivative
of the ideal QAOA time-evolved state. The former appears in the bound derived in Eq. (61), while the latter is a part of Eq. (62).

A bound on the norm of the derivative of the error operator is shown in Eq. (61). We elucidate the details on the derivation of this bound as follows:

\[ \| \partial_{\gamma_k} \Lambda(T) \| = \left\| e^{\mathcal{L}_0(T)} \int_0^1 d\lambda \ e^{-\lambda \mathcal{L}_0(T)} \partial_{\gamma_k} \mathcal{L}_0(T) e^{\lambda \mathcal{L}_0(T)} \right\| \]

\[ \leq \left\| e^{\mathcal{L}_0(T)} \right\| \int_0^1 d\lambda \ e^{-\lambda \| \mathcal{L}_0(T) \|} \partial_{\gamma_k} \mathcal{L}_0(T) \]

\[ \leq e^{\| \mathcal{L}_0(T) \|} \int_0^1 d\lambda \ 2^{\| \mathcal{L}_0(T) \|} \partial_{\gamma_k} \mathcal{L}_0(T) \]

\[ = \frac{e^{\| \mathcal{L}_0(T) \|}}{2^{\| \mathcal{L}_0(T) \|}} \left( e^{2\| \mathcal{L}_0(T) \|} - 1 \right) \| \partial_{\gamma_k} \mathcal{L}_0(T) \|. \quad (C7) \]

The first equality is obtained by expressing the error operator as a cumulant expansion, i.e., \( \Lambda(T) = e^{\mathcal{L}_0(T)} \) and then using an identity derived by Snider [56] to formally calculate the derivative of an exponential operator. Subsequent inequalities follow from applications of sub-multiplicativity and Eq. (48), while the final equality results from an evaluation of the integral.

While the first term in the bound given in Eq. (59) is proportional to the derivative of the error operator, the second is proportional to \( \partial_{\gamma_k} \rho_0(T) \). The subsequent bound on this term [Eq. (62)] is first computed by noting that

\[ \partial_{\gamma_k} \rho_0(T) = (\partial_{\gamma_k} U) \rho_0() U^\dagger + U \rho() (\partial_{\gamma_k} U)^\dagger \]

\[ = -iQ_{p,k+1} [H_C, Q_{k,1} \rho() Q_{k,1}^\dagger ] Q_{p,k+1}^{\dagger} \]

\[ = -iQ_{p,k+1} [H_C, \rho(T_k)] Q_{p,k+1}^{\dagger}. \quad (C8) \]

Using Hölder’s inequality to bound the second term of Eq. (59), we obtain an expression that contains a norm of the product of \( O \) and \( \partial_{\gamma_k} \rho_0(T) \). Employing the result from Eq. (C8), one can show

\[ \| O \partial_{\gamma_k} \rho_0(T) \| = \| O Q_{p,k+1} [H_C, \rho(T_k)] Q_{p,k+1}^{\dagger} \| \]

\[ = \| Q_{p,k+1}^{\dagger} O Q_{p,k+1} [H_C, \rho(T_k)] \| \]

\[ = \| \partial_{T}(T_k) [H_C, \rho(T_k)] \|. \quad (C9) \]

Note that the second equality follows from the unitary invariance of the norm.

Appendix D: Additional Results for Numerical simulations.

In this section, we provide additional details regarding the numerical simulations discussed in Sec. VII. In Sec. D.1 we discuss additional results for the Grover problem and in Sec. D.2 we provide additional information on the Ising problem.

1. Grover Problem

In this section we provide additional details for the fitting procedure used to extract the exponents discussed in Sec. VII A. We discuss the behavior of the noise-averaged expectation value in the presence of stochastic errors and coherent errors in Figs. 7 and 8 respectively. Finally, in Fig. 9 we discuss the fits to the distance between noisy and noisless QAOA evolution.

Let us start by discussing the case of stochastic precision errors for the optimal Grover QAOA algorithm, the data for which is in Fig. 4(a) in the main text. The results of the fitting procedure are discussed in Fig. 10. We fit the numerically evaluated noisy expectation value to the following phenomenological functional form that models both weak and strong noise behavior,

\[ \langle H_C \rangle = \langle H_C \rangle_0 S_H(p^*) + \langle H_C \rangle_{sat} [1 - S_H(p^*)] \quad (D1) \]

This functional form allows us to extract the decay function from the expression

\[ S_H(p^*) = \frac{\langle H_C \rangle - \langle H_C \rangle_{sat}}{\langle H_C \rangle_0 - \langle H_C \rangle_{sat}} \quad (D2) \]

with the condition, \( S_H(p^* = 0) = 1 \). In Fig 7(a), we extract the functional form of the saturation value by examining the large-noise behavior (\( \sigma = 1 \)). We find that the saturation value decays exponentially with the number of qubits \( n \) (or linearly with problem size \( N \)),

\[ \langle H_C \rangle_{sat} \approx 0.69 \times 2^{-0.67n} \quad (D3) \]

Utilizing this expression for the large-noise saturation value, we obtain a fit for the decay function, \( S_H(T, p^*) \) in Fig 7(b) and (c). We find two distinct functional behaviors for the weak noise [(b)] and large noise [(c)] respectively,

\[ S_H(p^*) = \begin{cases} 0.97 e^{-4.55 \Gamma p^*}, & \Gamma p^* < 0.5 \text{ [weak noise]} \\ 0.05 e^{-1.2 \Gamma p^*}, & \Gamma p^* > 1.0 \text{ [strong noise]} \end{cases} \quad (D4) \]

Note that the weak and strong noise behavior is closely related to the behavior of the scaling variable \( p^* \).

Next, we extract the scaling behavior of the noisy expectation value in the presence of coherent errors, the data for which is shown in Fig. 4(b) in the main text. The scaling fit is discussed in Fig. 8. We note that measured expectation value in the presence of constant coherent error does not show a monotonic behavior with increase errors. This is shown in Fig. 8(a) where we see that after an initial decay in the expectation value, the expectation value exhibits oscillations. We extract the initial exponential decay by examining the cases where \( \eta p^* \lesssim 1 \). The initial decay is well described by an exponential decay,

\[ \langle H_C \rangle = \langle H_C \rangle_0 S_H(p^*) \quad \Gamma p^* < 0.5 \text{ [weak noise]} \quad (D5) \]

\[ S_H(p^*) = e^{-11.54(\eta p^*)^{2.08}} \quad (D6) \]

Note that the expectation value decays as a squared-exponential.

Finally, let us discuss the scaling of the difference between the noisy and noisless QAOA evolution operator. We
numerically evaluate the $\infty$-norm of the average difference between the noisy unitary $U$ and the noiseless one, $U_0$ in Fig. 9, the data for which corresponds to Fig. 3 (c) and (d). We have the following behavior for weak noise or small $p$ ($p\sigma^2 \ll 1, \mu p \ll 1$),

$$\|U - U_0\|_\infty \approx \begin{cases} 4.57(\Gamma p)^{0.47}, & \text{[stochastic error]} \\ 5.35(\eta p)^{0.95}, & \text{[coherent error]} \end{cases} \tag{D7}$$

This scaling is consistent with the norm-difference increasing in a diffusive manner for stochastic error ($\propto \sqrt{\Gamma p^*}$) and linearly for coherent error ($\propto \eta p^*$).

2. Ising Problem

In this section, we elaborate on the scaling fits for the Ising problem on a ring, that is discussed in Sec. VII B. The data corresponds to Fig. 6 in the main text. We fit the decay in the expectation value to the following phenomenological function,

$$\langle H_C \rangle \approx \langle H_C \rangle_0 S_H(p^*) \tag{D8}$$

FIG. 7. Scaling fits for the expectation value of the Grover cost Hamiltonian in the presence of stochastic errors [data corresponding to Fig. 4(a)]. (a) shows the fit to the saturation value for large noise. The saturation value decreases linearly with problem size, or exponentially with the number of qubits. (b) Weak noise ($\Gamma p^* < 0.5$) behavior of the decay function (See Eq. D2). For weak noise, the expectation value decays exponentially as evident from the linear fit on the log-linear plot. (c) Large noise ($\Gamma p^* > 1$) behavior of the decay function which behaves as a power-law.

FIG. 8. Scaling fits for the expectation value of the Grover cost function in the presence of coherent error [data corresponding to Fig. 4(b)]. (a) shows the non-monotonic behavior of the noisy expectation value for increasing magnitude of the coherent error. (b) The decay of the expectation value for small coherent error $\eta p^* \lesssim 1$ behaves approximately as a squared exponential decay.

FIG. 9. Scaling fits for the norm difference with increasing number of layers and noise, data corresponding to Figs. 3 (c) and (d). (a) Diffusive growth in the distance for stochastic precision error. (b) Linear growth in the distance for coherent precision error.

FIG. 10. Scaling fits for GHZ Hamiltonian expectation value for stochastic errors. (a) For weak noise ($\Gamma p^* \lesssim 1$) the expectation value of the Hamiltonian decays exponentially. (b) For strong noise ($\Gamma p^* \gtrsim 1$), the expectation value decays as a power law.

For stochastic error, as discussed in Fig. 10, the Hamiltonian expectation value decays as,

$$S_H(p^*) = \begin{cases} 1.00e^{-3.80\Gamma p^*}, & \Gamma p^* < 1.0 \text{ [weak noise]} \\ 0.04(\Gamma p^*)^{-0.46}, & \Gamma p^* > 1.0 \text{ [strong noise]} \end{cases} \tag{D9}$$

Again, for stochastic errors, the expectation value decays exponentially for weak strength of the noise then changing to a power-law for larger strength of noise.

Next we discuss the effect of coherent errors. In this case, the decay function for weak noise is again consistent with an
Note that the exponential decay is consistent with a scaling variable $\eta^2 p^*$. 

\begin{equation}
S_H(p^*) = e^{-4.64\eta^2 p^*}.
\end{equation}
[36] M. Streif and M. Leib, Quantum Science and Technology 5, 034008 (2020).
[37] M. M. Wauters, E. Panizon, G. B. Mbeng, and G. E. Santoro, Phys. Rev. Research 2, 033446 (2020).
[38] H. Ball and M. J. Biercuk, EPJ Quantum Technology 2, 1 (2015).
[39] V. Frey, S. Mavadia, L. Norris, W. De Ferranti, D. Lucarelli, L. Viola, and M. Biercuk, Nature communications 8, 1 (2017).
[40] L. Viola, E. Knill, and S. Lloyd, Physical Review Letters 82, 2417 (1999).
[41] K. Khodjasteh and D. A. Lidar, Physical review letters 95, 180501 (2005).
[42] D. A. Lidar, P. Zanardi, and K. Khodjasteh, Phys. Rev. A 78, 012308 (2008).
[43] K. Khodjasteh and D. A. Lidar, Phys. Rev. A 78, 012355 (2008).
[44] G. S. Uhrig and D. A. Lidar, Phys. Rev. A 82, 012301 (2010).
[45] H. K. Ng, D. A. Lidar, and J. Preskill, Phys. Rev. A 84, 012305 (2011).
[46] J. C. Spall et al., IEEE transactions on automatic control 37, 332 (1992).
[47] F. K. Wilhelm, S. Kirchhoff, S. Machnes, N. Wittler, and D. Sugny, arXiv preprint arXiv:2003.10132 (2020).
[48] A. M. Childs, Y. Su, M. C. Tran, N. Wiebe, and S. Zhu, Phys. Rev. X 11, 011020 (2021).
[49] E. Farhi, J. Goldstone, S. Gutmann, and M. Sipser, arXiv e-prints, quant-ph/0001106 (2000), arXiv:quant-ph/0001106 [quant-ph].
[50] C. Developers, “quantumlib/cirq: Cirq v0.9.1,” (2020), See full list of authors on Github: https://github.com/quantumlib/Cirq/graphs/contributors.
[51] F. Barahona, J. Phys. A: Math. Gen 15, 3241 (1982).
[52] A. Lucas, Front. Phys. 2, 5 (2014).
[53] P. C. Lotshaw and T. S. Humble, “QAOA dataset.”
[54] P. C. Lotshaw, T. S. Humble, R. Herman, J. Ostrowski, and G. Siopsis, arXiv:2102.06813 (2021).
[55] C. Dalyac, L. Henriet, E. Jeandel, W. Lechner, S. Perdrix, M. Porcheron, and M. Veshchezerova, EPJ Quantum Technology 8, 12 (2021).
[56] R. F. Snider, Journal of Mathematical Physics 5, 1580 (1964).