

Blind spots of randomized benchmarking under temporal correlations

Varun Srivastava^{1,*}, Abhinash Kumar Roy^{1,†}, Soumik Mahanti^{1,2}, Jasleen Kaur,¹
Salini Karuvade³, and Alexei Gilchrist^{1,‡}

¹*Department of Physical and Mathematical Sciences, Macquarie University, Sydney, New South Wales, Australia*

²*School of Computer Science, University of Technology Sydney, Ultimo, Sydney, New South Wales 2007, Australia*

³*School of Physics, The University of Sydney, Sydney, New South Wales 2006, Australia*



(Received 22 October 2025; accepted 10 May 2026; published 8 June 2026)

Randomized benchmarking (RB) is a widely adopted protocol for estimating the average gate fidelity in quantum hardware. However, its standard formulation relies on the assumption of temporally uncorrelated noise, an assumption often violated in current devices. In this work, we derive analytic expressions for the average sequence fidelity (ASF) in the presence of temporally correlated (non-Markovian) noise with classical memory, including cases where such correlations originate from interactions with a quantum environment. We show how the ASF can be interpreted to extract meaningful benchmarking parameters under such noise and identify classes of interaction Hamiltonians that render temporal correlations completely invisible to RB. We further provide operational criteria for witnessing temporal correlations due to quantum memory through RB experiments. Importantly, while classical correlations may remain undetectable in the ASF data, they can nonetheless significantly affect worst-case errors quantified by the diamond norm, a metric central to fault-tolerant quantum computing. In particular, we demonstrate that temporal correlations may suppress worst-case errors, highlighting that temporal correlations may not always have detrimental effects on gate performance.

DOI: [10.1103/r1z5-ycqc](https://doi.org/10.1103/r1z5-ycqc)

I. INTRODUCTION

Quantum characterization, verification, and validation (QCVV) is essential for developing reliable quantum technologies [1,2]. Among QCVV tools, randomized benchmarking (RB) has become one of the most widely implemented methods for characterizing the average performance of a gate set via a single decay parameter [3–6]. RB offers two significant practical advantages for characterizing gates, namely, scalability and insensitivity to state-preparation and measurement (SPAM) errors. In contrast, process tomography suffers on both counts [7]. This practicality has driven broad adoption of RB for gate characterization across platforms including trapped-ion [4], superconducting [8,9], and neutral-atom processors [10,11].

RB assumes temporally uncorrelated (i.e., Markovian) and gate-independent noise in its standard formulation. Under these assumptions, the average sequence fidelity (ASF), which is the quantity the RB protocol estimates, follows a simple exponential decay model [12]. Since its inception, numerous variants have been proposed to adapt the protocol to more

general scenarios, namely, simultaneous RB to probe multi-qubit crosstalk [13], interleaved RB to estimate the average error of a targeted gate [14], extensions beyond the Clifford group to more general gate sets [1,15,16], relaxations to weak gate-dependent noise [5,17], direct RB for multiqubit native gates [18], and logical-level RB on encoded qubits [19]. These variants largely inherit the assumption of Markovianity of the noise; however, most of the present quantum hardware routinely exhibits temporally correlated noise [20,21]. Only recently, the RB framework has been extended to accommodate non-Markovian dynamics [22,23]. Despite this progress, a systematic analysis of RB under diverse memory structures in the environment remains unexplored. These memory structures can be broadly categorized into classical or quantum memory depending on the type of correlations mediated by the environment [24,25].

Prior studies have examined RB in the presence of classically correlated noise arising from coupling to stochastic classical fields, revealing departures from the Markovian ASF model [22,26–29]. For instance, Ref. [27] shows that slowly varying $1/f$ noise can produce decay curves better captured by multiexponential fits. In Ref. [22], two specific models, the controlled dephasing noise model and the shallow pocket model, were presented as examples of environment with classical memory, where the ASF curve showed monotonic decay. Crucially, the effective memory in the environment can be classical even when it originates from interactions with nearby quantum environment for certain classes of interaction Hamiltonians [30], which prominently includes couplings naturally present in the present hardware, such as ZZ interactions in superconducting devices [31,32]. These observations motivate a

*Contact author: varun.srivastava@students.mq.edu.au

†Contact author: abhinash.roy@students.mq.edu.au

‡Contact author: alexei.gilchrist@mq.edu.au

systematic study of RB in the most general classical-memory scenarios and what RB can (and cannot) reveal about both average gate errors and worst-case errors across gate sequences. Importantly, the process matrix formalism offers a unified and operational way to model and compare memory structures and assess their impact on RB [24,25,33].

In this work, we present a general investigation of the RB protocol in the presence of classical memory, including memory effects that originate from interactions with a quantum environment. We derive analytical expressions for the ASF, which reveal the emergence of multiple decay parameters, and we outline how these parameters can be reliably extracted and interpreted from RB data. A central message of our study is that classical-memory effects often remain hidden in ASF curves. In particular, we show that the ASF remains monotonically decreasing with sequence length under such memory. Hence, any observed nonmonotonicity provides a strong diagnostic for genuinely quantum-memory effects. We further identify parameter regimes in which the ASF of a classical-memory process is indistinguishable from that of a Markovian model, rendering RB effectively blind to these correlations, and we specify a class of interaction Hamiltonians that lead to such RB blindness. Finally, we demonstrate that even when RB fails to reveal memory effects, these correlations can substantially influence the worst-case performance of gates, a metric that is critical for assessing fault-tolerance thresholds [34,35].

The paper is structured as follows: In Sec. II, we provide a background for the RB protocol. We also present a brief introduction to the process matrix formalism and various memory structures in Sec. III. In Sec. IV, we cast the RB procedure in the process matrix formalism and derive the resulting ASF curves for time-dependent Markovian and classical-memory scenarios. In Sec. V, we study the various blindness criteria and class of RB-blind Hamiltonians. Section VI includes discussion on worst-case error performance of gates under RB blind scenario and relevance of memory effects. We conclude and discuss future prospects in Sec. VII.

II. RANDOMIZED BENCHMARKING PROTOCOL

In this work, our analysis includes the standard randomized benchmarking protocol, where the gate set is sampled from the n -qubit Clifford group $\mathcal{CG}_n$ [3,5,6,36,37]. The standard RB protocol proceeds through the following steps (for other variants, see Refs. [12,38] and references therein):

- (1) Prepare an initial state ρ (ideally $|\psi\rangle\langle\psi|$).
- (2) The prepared state ρ is subject to a sequence of $m+1$ Clifford gates $S_i \equiv C_{m+1}^{(i)} \circ C_m^{(i)} \circ \dots \circ C_2^{(i)} \circ C_1^{(i)}$, where the first m gates are sampled uniformly at random from $\mathcal{CG}_n$ and the final gate is a fixed *motion-reversal* gate chosen to invert the preceding sequence, $C_{m+1}^{(i)} = (C_m^{(i)} \circ \dots \circ C_2^{(i)} \circ C_1^{(i)})^\dagger$. Under ideal gate implementation, $S_i \equiv \mathcal{I}$; i.e., the overall action is an identity map. More generally, each gate is noisy and modeled as $C_t^{(i)} = \mathcal{G}_t^{(i)} \circ \mathcal{N}_t$, where $\mathcal{G}_t^{(i)}$ is the ideal gate implementation at time step t and $\mathcal{N}_t$ is a completely positive trace-preserving (CPTP) map representing the noise present in the implementation of $\mathcal{G}_t^{(i)}$. Here, we adopt the convention that noise acts *before* the ideal gate; equivalently, one can

represent the noisy gate as $C_t^{(i)} = \tilde{\mathcal{N}}_t \circ \mathcal{G}_t^{(i)}$ with appropriate modification to the analysis.

- (3) Measure with a two-outcome positive operator-valued measure (POVM) $\{E, \mathbb{I} - E\}$ (ideally $E = |\psi\rangle\langle\psi|$), and estimate the sequence fidelity $F_i^{(m)} := \text{Tr}[E S_i(\rho)] \in [0, 1]$.

- (4) Repeat over k random sequences of the same length m and average to estimate the ASF, $F^{(m)} := \frac{1}{k} \sum_{i=1}^k F_i^{(m)}$. A careful discussion of finite-shot fluctuations and finite sampling over random sequences can be found in Refs. [39,40]. Our focus here is on the ASF; adapting those statistical analyses to non-Markovian noise is left for future work.

- (5) Repeat steps 1–4 for several values of m and fit $F^{(m)}$ as a function of m to extract decay parameters and an average error rate.

Under the assumption of time- and gate-independent Markovian noise, the ASF takes the exponential form

$$F^{(m)} = Aq^m + B, \quad (1)$$

where $q \in [0, 1]$. The coefficients A and B capture the SPAM errors. Fitting the experimentally obtained ASF data to Eq. (1) yields an estimate of the decay parameter q , which is related to the average gate error by the relation [3]

$$r = (1 - q) \frac{d - 1}{d}, \quad (2)$$

where d is the dimension of the system. A key feature of Eq. (1) is that it separates SPAM errors from gate errors, enabling a robust estimation of q from experimental data. This separation is, in fact, a general feature of Markovian noise. For time-dependent Markovian noise, it has been shown in Ref. [39] that the ASF takes the form

$$F^{(m)} = A \prod_{t=1}^m q_t + B, \quad (3)$$

where $q_t \in [0, 1]$ quantifies the average gate error at time step t . In contrast to non-Markovian noise models, as noted in Refs. [22,23], the SPAM parameters do not separate from the gate errors in general. Moreover, the temporal correlations between noise maps make it nontrivial to define an average error rate per gate and only allow the estimation of an average error over entire gate sequence of length m .

III. PROCESS MATRIX FORMALISM

In the RB protocol, the system is acted on repeatedly by quantum operations (gates) at multiple time steps, making it natural to describe the experiment using the process matrix formalism [22,33,41]. In a multitime quantum process, the system is subjected to interventions at several time steps; each intervention can be viewed as a “lab” implementing a valid quantum operation, i.e., a completely positive trace nonincreasing map, which in case of RB protocol would correspond to implementing Clifford gates followed by a final POVM measurement. Between interventions, the system interacts with an environment via joint system-environment unitaries, which in our case would correspond to the noise associated with each gate implementation. Process matrices, introduced in Ref. [42], provide an operational description of the correlations compatible with locally valid quantum

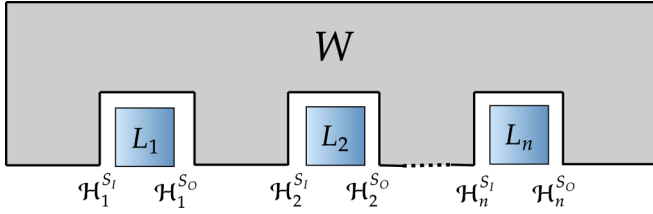


FIG. 1. Schematic diagram of process matrix W with n intervention labs $\{L_1, L_2, \dots, L_n\}$. The process matrix captures the most general correlations compatible with the local validity of quantum mechanics in the labs.

mechanics in the labs (Fig. 1). When the labs are related by a definite causal order, as in the multitime setting like RB, the process matrix satisfies causality constraints [41] and is equivalent to a quantum comb/process tensor [43]. This framework is widely used to model non-Markovian quantum dynamics and provides an operational notion of quantum stochastic processes [33,43]. We will use this formalism to model non-Markovian noise in RB in Sec. IV.

In detail, we consider a system S interacting with an environment E . At each discrete time step t , an instrument acts locally on S , described by CP trace nonincreasing maps $\{\mathcal{T}_{m_t|x_t}\} : \mathcal{L}(\mathcal{H}_t^{S_i}) \rightarrow \mathcal{L}(\mathcal{H}_t^{S_o})$, where x_t labels the instrument setting and m_t the outcome. The corresponding CPTP map is $\mathcal{T}_{x_t} = \sum_{m_t} \mathcal{T}_{m_t|x_t}$ (i.e., outcomes ignored). Between interventions, S and E undergo joint unitary evolution, generating correlations across time that can lead to non-Markovian behavior (see Fig. 2). For calculations, we use the Choi-Jamiołkowski (CJ) isomorphism [44,45]. For a linear map $\mathcal{M} : \mathcal{L}(\mathcal{H}^A) \rightarrow \mathcal{L}(\mathcal{H}^B)$, its CJ operator is $[[\mathcal{M}]] = (\mathcal{I} \otimes \mathcal{M})(|\mathbb{I}\rangle\rangle\langle\langle\mathbb{I}|) \in \mathcal{L}(\mathcal{H}^A \otimes \mathcal{H}^B)$, where $|\mathbb{I}\rangle\rangle = \sum_j |j\rangle \otimes |j\rangle$ is the (unnormalized) maximally entangled state in a fixed basis $\{|j\rangle\}$.

All multitime system-environment correlations are then encoded in a single process matrix $W \geq 0 \in \mathcal{L}(\bigotimes_{t=1}^n (\mathcal{H}_t^{S_i} \otimes \mathcal{H}_t^{S_o}))$. For an n -step experiment, the probability of outcomes $\vec{m}_n = (m_1, \dots, m_n)$ given settings $\vec{x}_n = (x_1, \dots, x_n)$ is given by the generalized Born rule [42]

$$p(\vec{m}_n|\vec{x}_n) = \text{Tr} \left[W^T \left(\bigotimes_{t=1}^n [[\mathcal{T}_{m_t|x_t}]] \right) \right], \quad (4)$$

where the transpose is taken in the CJ reference basis.

W can be viewed as a state up to a normalization factor, which is restricted by linear causality constraints to ensure valid probabilities for all admissible instruments [46]. This “statelike” encoding of multitime correlations is powerful, as it allows one to infer memory effects of the environment directly from the correlation structure of W . Given a specific system-environment model (e.g., a sequence of joint unitaries), W can be constructed by composing CJ operators of the unitaries via the link product; given the CJ form of two linear operators $[[\mathcal{M}]]$ and $[[\mathcal{N}]]$, the link product $[[\mathcal{M}]] \star [[\mathcal{N}]] = \text{Tr}_{M \cap N}[(\mathbb{I}^{N \setminus M} \otimes [[\mathcal{M}]]^{T_{M \cap N}})([[\mathcal{N}]] \otimes \mathbb{I}^{M \setminus N})]$, where $T_{M \cap N}$ is the partial transpose over the shared Hilbert space. In Appendix A, we detail the properties of the link product and how it can be used to evaluate time-ordered process matrices and Eq. (4). In the following, we provide a brief

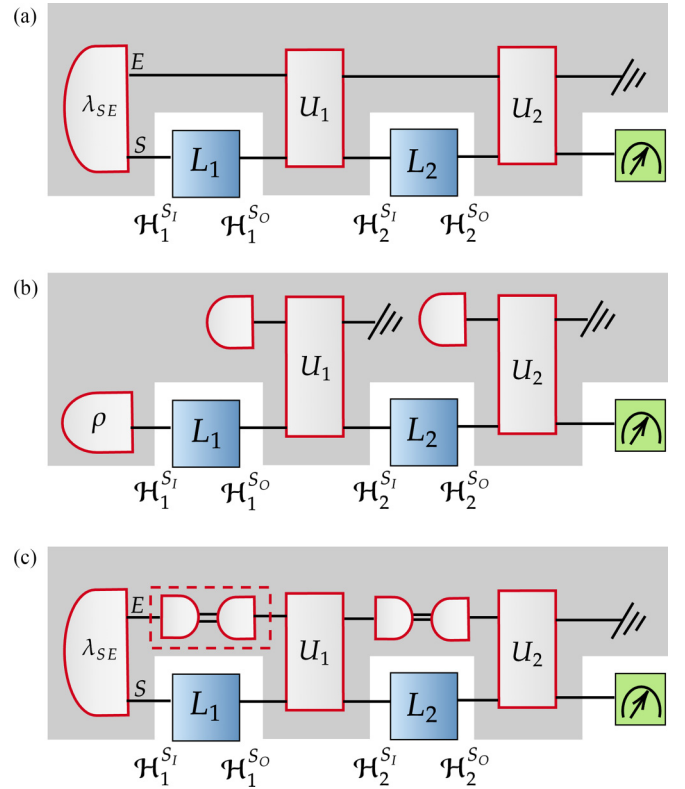


FIG. 2. Schematic of two-time-step multitime process with labs L_1, L_2 and process matrix (in gray) capturing multitime correlations due to system-environment interactions. Panel (a) shows a general multitime process with initial state and system-environment unitaries. Panel (b) shows a Markovian process where the initial state is a product state and environment is traced out and reprepared at each time step. Panel (c) represents a classical-memory process, where an entanglement-breaking channel acts on the environment at each step (red dashed box), preserving only classical correlations across time.

introduction to Markovian and non-Markovian processes with classical memory, which would be useful to model different noise models of RB in Sec. IV. For a detailed derivation of different memory structures, see Ref. [24].

A. Markovian processes

In a Markovian process, the environment does not carry information forward, so no temporal correlations are mediated between intervention times. In the process matrix formalism, this is equivalent to a product structure [24,33,41]

$$W_M = \rho^{S_i} \bigotimes_{t=1}^{n-1} [[\mathcal{N}_t]]^{S_o S_i}, \quad (5)$$

where $[[\mathcal{N}_t]] \in \mathcal{L}(\mathcal{H}_t^{S_o} \otimes \mathcal{H}_{t+1}^{S_i})$ is the CJ operator of the CPTP map from time t to $t+1$ and ρ^{S_i} is the initial system state. Operationally, this corresponds to the picture in which S interacts at each step with a fresh, uncorrelated environment [see Fig. 2 (b)]. Henceforth, we will ignore the superscripts on operators labeling their input/output Hilbert spaces whenever it is clear from the context.

B. Non-Markovian processes

Any process that does not factorize as in Eq. (5) is non-Markovian. Physically, this occurs when a shared environment interacts with the system at multiple times, retains information, and thereby mediates temporal correlations between labs. Depending on the type of correlations transmitted, non-Markovian processes can be broadly classified into *classical-memory* and *quantum-memory* processes [24,25,47]. In this work, we focus on RB scenarios in which the noise is consistent with *classical memory*.

Intuitively, classical memory means that the environment effectively generates and stores a *classical record* that is fed forward to influence later dynamics [see Fig. 2 (c)]. We call this the *classical feedforward (CFF)* mechanism: Between times t and $t+1$, the effective environmental action is a conditional instrument $\{\mathcal{N}_{a_t|x_t}\}_{a_t}$, whose choice of setting x_t may depend on the classical history $(\vec{a}_{t-1}, \vec{x}_{t-1})$ via $p(x_t|\vec{a}_{t-1}, \vec{x}_{t-1})$. The resulting process matrix is

$$W_{\text{CFF}} = \sum_{\vec{x}_n, \vec{a}_n} p(x_0) \rho_{a_0|x_0} \bigotimes_{t=1}^{n-1} p(x_t|\vec{a}_{t-1}, \vec{x}_{t-1}) [\mathcal{N}_{a_t|x_t}], \quad (6)$$

where $\{\rho_{a_0|x_0}\}_{a_0} \in \mathcal{L}(\mathcal{H}_1^{S_I})$ are subnormalized states satisfying $\text{Tr}(\sum_{a_0} \rho_{a_0|x_0}) = 1$; i.e., for each setting x_0 , the state $\rho_{a_0|x_0}$ is selected with probability $\text{Tr}(\rho_{a_0|x_0})$.

A useful special case is *classical common cause (CCC)*, where the environment is characterized by a single classical label x and applies deterministic CPTP maps conditioned on x :

$$W_{\text{CCC}} = \sum_x p(x) \rho_x \otimes [\mathcal{N}_x] \otimes \underbrace{\cdots}_{n-1 \text{ times}} \otimes [\mathcal{N}_x]. \quad (7)$$

CCC processes are therefore convex mixtures of Markovian processes in Eq. (5).

Finally, such classical correlations can arise either from coupling to classical stochastic fields or from genuine interactions with a quantum environment [30,48]; importantly, any classical-memory process can be *simulated* by a classical stochastic model. In Appendix B, we summarize system-environment Hamiltonian mechanisms known to generate these effects. Any process matrix that is neither Markovian nor of classical-memory type constitutes a *quantum-memory* process. In RB, the process matrix provides a compact description of the underlying noise. For standard RB with Markovian noise, W factorizes as in Eq. (5). Here, instead, we focus on classical-memory noise models described by Eqs. (6) and (7), with particular emphasis on the CCC class.

IV. MODELING NON-MARKOVIAN NOISE IN RANDOMIZED BENCHMARKING

In this section, we demonstrate how the process matrix formalism can be applied to model non-Markovian noise within the randomized benchmarking protocol. Consider a model with initially uncorrelated system-environment state $\rho^S \otimes \sigma^E$. The action of the noisy gate is given by $\mathcal{C}_t = (\mathcal{I} \otimes \mathcal{G}_t) \circ \mathcal{U}_t^{SE}$, where $\mathcal{U}_t^{SE}$ is the joint system-environment unitary and $\mathcal{G}_t$ are the ideal gates sampled from the n -qubit Clifford group. In the final time step, we apply the motion-reversal gate and measure with POVM element E . All components except for the control

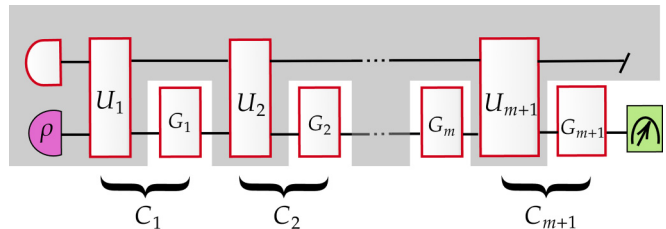


FIG. 3. RB protocol in the presence of non-Markovian noise. The process matrix construction provides a natural framework to study non-Markovian effects in RB. Each noisy gate $\{\mathcal{C}_1, \mathcal{C}_2, \dots, \mathcal{C}_{m+1}\}$ can be divided into ideal Clifford gates $\{G_1, G_2, \dots, G_{m+1}\}$, which act only on the system, and the noise is captured by the process matrix comb (in gray).

operations $\{\mathcal{G}_t\}_{t=1}^{m+1}$ and E can be collected into the process matrix, which acts as a higher-order quantum operation mapping the chosen sequence of gates to corresponding sequence fidelity (see Fig. 3) [49]. More generally, if the initial state is correlated, one can apply a trace-and-replace operation on S to prepare an effective product state. Operationally, this requires the initial S - E coupling (and hence initial correlations) to be weak. In the following subsections, we first formulate the ASF calculation within the process matrix framework, which allows us to model Markovian as well as non-Markovian gate errors with classical memory.

A. ASF calculation using process matrix formalism

The instruments in the RB protocol are unitary gates $\{\mathcal{G}_t\} \in \mathcal{CF}_n$ (see Fig. 3). For a sequence S_α , the operation at time t is the unitary channel $\mathcal{G}_t^{(\alpha)}(\cdot) = G_t^{(\alpha)}(\cdot)G_t^{(\alpha)\dagger}$ represented in the CJ form as $[[\mathcal{G}_t]]$. Using the link product to concatenate the initial system-environment state and joint unitaries, we obtain the process matrix W . The sequence fidelity can be expressed as the link product between W and the instruments as the following:

$$F_\alpha^{(m)} = W \star \mathbb{T}_\alpha, \quad (8)$$

where

$$W = (\rho \otimes \sigma) \star [[\mathcal{U}_1^{SE}]] \star \cdots \star [[\mathcal{U}_{m+1}^{SE}]] \star \mathbb{I}^E, \\ \mathbb{T}_\alpha = [[\mathcal{G}_1^{(\alpha)}]] \star \cdots \star [[\mathcal{G}_m^{(\alpha)}]] \star [[\mathcal{G}_{m+1}^{(\alpha)}]] \star E^T. \quad (9)$$

Moreover, we define a set of new Clifford operations, $\mathcal{K}_1^{(\alpha)} = \mathcal{G}_1^{(\alpha)}$ and $\mathcal{K}_i^{(\alpha)} = \mathcal{G}_i^{(\alpha)} \circ \mathcal{K}_{i-1}^{(\alpha)} \forall i \geq 2$. Since the set $\mathcal{CF}_n$ forms a group, $\mathcal{K}_i^{(\alpha)} \in \mathcal{CF}_n \forall i$, using the cyclicity and the linearity of trace operation, Eq. (8) can be equivalently represented as

$$F_\alpha^{(m)} = \tilde{W}_\alpha \star \Gamma, \quad (10)$$

where $\Gamma = |\mathbb{I}\rangle\rangle\langle\langle\mathbb{I}|^{\otimes(m)} \otimes E^T$. Note that $|\mathbb{I}\rangle\rangle\langle\langle\mathbb{I}| = [[\mathbb{I}]]$ is the Choi operator of the identity channel. Hence, after this rewriting, the new instruments applied to process are simply m identity channel slots together with a final POVM measurement and all dependence on the particular RB sequence α is absorbed into an effective, sequence-dependent process

matrix $\tilde{W}_\alpha$, given by

$$\tilde{W}_\alpha = \mathbb{K}^{(\alpha)} W \mathbb{K}^{\dagger(\alpha)},$$

$$\mathbb{K}^{(\alpha)} = \mathbb{I} \bigotimes_{t=1}^m (K_t^{T(\alpha)} \otimes K_t^{\dagger(\alpha)}), \quad (11)$$

where superscript T denotes the transpose in the specified basis. We summarize the above in the following theorem.

Theorem 1. Given an initial system-environment state $\rho \otimes \sigma \in \mathcal{L}(\mathcal{H}_0^S \otimes \mathcal{H}_0^E)$ and a sequence of noisy gates, the noise structure is captured by the process matrix W defined in Eq. (9). The ASF is

$$\bar{F}^{(m)} = \frac{1}{\Omega^m} \sum_{\alpha=1}^{\Omega^m} \tilde{W}_\alpha \star \Gamma, \quad \Omega = |\mathcal{CF}_n|. \quad (12)$$

The details of the derivation are provided in Appendix C. We note that in Ref. [22], the authors also provided a general expression for ASF for an arbitrary process matrix. However, the ASF is formulated at a level of generality that does not directly translate into experimentally practical procedures for extracting relevant metrics of interest unless we specify details of the noise model. In our work, we focus on a physically well-motivated subclass of non-Markovian processes (classical-memory models) for which the ASF acquires an interpretable structure and the relevant parameters can be estimated from RB data.

We conclude this section by showing that Eq. (12) reproduces the standard ASF for a time-dependent Markovian process whose process matrix has the product form $W_M = \rho \star [[\mathcal{N}_1]] \star \cdots \star [[\mathcal{N}_{m+1}]]$, where ρ is the initial system state and $[[\mathcal{N}_t]]$ denotes the Choi operator of the CPTP noise map at time step t . With our convention that noise acts before the ideal Clifford gate (Fig. 3), we define $\tilde{\rho} := \rho \star [[\mathcal{N}_1]] = \mathcal{N}_1(\rho)$, so that W_M can be written as $W_M = \tilde{\rho} \otimes [[\mathcal{N}_2]] \otimes \cdots \otimes [[\mathcal{N}_{m+1}]]$.

Applying the sequence-induced transformation $W \mapsto \tilde{W}_\alpha$ from Eq. (11) and substituting into Eq. (12), the ASF reduces to

$$\bar{F}_M^{(m)} = (\tilde{\rho} \otimes Z_2 \otimes \cdots \otimes Z_{m+1}) \star \Gamma, \quad (13)$$

where Z_t is the twirl of $[[\mathcal{N}_t]]$ over $\mathcal{CF}_n$, namely, $Z_t := \frac{1}{\Omega} \sum_{K_t \in \mathcal{CF}_n} (K_t^{T(\alpha)} \otimes K_t^{\dagger(\alpha)}) [[\mathcal{N}_t]] (K_t^{T(\alpha)} \otimes K_t^{\dagger(\alpha)})^\dagger$, with $\Omega = |\mathcal{CF}_n|$. Since $\mathcal{CF}_n$ forms a unitary 2-design, this average is equal to the Haar second moment. Using Schur-Weyl duality (see Appendix C for details), the twirl projects onto span $\{|\mathbb{I}\rangle, |\mathbb{I}\rangle\langle\mathbb{I}|\}$, yielding

$$Z_t = c_1(\mathcal{N}_t) \frac{\mathbb{I}}{d} + c_2(\mathcal{N}_t) |\mathbb{I}\rangle\langle\mathbb{I}|,$$

with coefficients

$$c_1(\mathcal{N}_t) = \frac{d \operatorname{Tr}([[\mathcal{N}_t]]) - \operatorname{Tr}(|\mathbb{I}\rangle\langle\mathbb{I}|) \operatorname{Tr}([[\mathcal{N}_t]])}{d^2 - 1},$$

$$c_2(\mathcal{N}_t) = \frac{\operatorname{Tr}(|\mathbb{I}\rangle\langle\mathbb{I}|) \operatorname{Tr}([[\mathcal{N}_t]]) - d^{-1} \operatorname{Tr}([[\mathcal{N}_t]])}{d^2 - 1}. \quad (14)$$

Substituting this form of Z_t into Eq. (13) gives the desired time-dependent Markovian ASF, $\bar{F}_M^{(m)} = A \prod_{t=1}^m q_t + B$, as in

Eq. (3), where $A = \operatorname{Tr}[E(\mathcal{N}_1(\rho) - \mathbb{I}/d)]$, $B = \frac{1}{d} \operatorname{Tr}(E)$, and $q_t = c_2(\mathcal{N}_t)$.

B. Non-Markovian processes with classical memory

We now derive the ASF for non-Markovian noise models whose multi-time correlations exhibit classical memory. In the following, we analyze both cases for CCC- and CFF-type memory models and derive the modified ASF expression from RB protocol.

1. ASF for classical common cause

The process matrix for a CCC-type process is given by Eq. (7) and represents a *convex mixture of Markovian processes*. Accordingly, each run of the RB protocol may be viewed as sampling a noisy Markovian branch from an associated probability distribution. This class of noise models captures several experimentally relevant scenarios, including drifting qubit resonance frequencies [27], stochastic electromagnetic field fluctuations, and ZZ-type Hamiltonian couplings between qubits in superconducting architectures [30]. Therefore, despite its apparent simplicity, this noise model is of considerable practical relevance. Moreover, in Ref. [30], a class of system-environment interaction Hamiltonians was identified that give rise to such memory models and we summarize the results in Appendix B.

We can write Eq. (7) as $W_{\text{CCC}} = \sum_x p_x W_M^{(x)}$, where each $W_M^{(x)}$ corresponds to a Markovian process matrix with associated noise map $\mathcal{N}_x$ and weight factor p_x . Here, we are considering the scenario where each Markovian process matrix $W_M^{(x)}$ corresponds to a time-independent process. It is straightforward to extend the analysis to a convex mixture of time-dependent Markovian processes. From Eq. (10), we get the sequence fidelity $F_\alpha^{(m)} = \sum_x p_x \tilde{W}_{M,(\alpha)}^{(x)} \star \Gamma$, for some sequence of gates S_α . Using Eq. (12), we obtain the ASF

$$\bar{F}^{(m)} = \sum_x p_x A_x q_x^m + B, \quad (15)$$

where $A_x = \operatorname{Tr}(E(\mathcal{N}_x(\rho) - \mathbb{I}/d))$, $B = \operatorname{Tr}(E)/d$, and $q_x = c_2(\mathcal{N}_x)$ as defined in Eq. (14).

We thus obtain an ASF that is a *sum of exponentials* rather than a single exponential; therefore, estimating the decay parameters becomes a multiexponential fitting problem. Such problems are well studied in signal processing (see Refs. [38,50,51] and Sec. VB). We also note that the SPAM coefficient $A_x = \operatorname{Tr}(E(\mathcal{N}_x(\rho) - \mathbb{I}/d))$ is no longer independent of the decay parameters. The origin of this coupling is that, in standard RB, the first noise map is not twirled in the same way as the subsequent steps. In the Markovian setting, this does not cause difficulties because noise at different times is independent, so preparation errors cannot influence later dynamics beyond their effect on the initial state. In the presence of temporal correlations, however, the environment can correlate the initial preparation with subsequent noise, so SPAM contributions can become coupled to the effective decay behavior. Consequently, even for the CCC model, the standard RB analysis must be modified to properly account for correlated SPAM errors.

Randomization of the initial state and final measurement. To decouple SPAM from the decay parameters, we additionally twirl the first noise map by repeating the RB experiment over a family of input states and matched two-outcome measurements. Concretely, we consider preparations $\rho = |\psi\rangle\langle\psi|$ together with POVMs $\{E, \mathbb{I} - E\}$, where $E = |\psi\rangle\langle\psi|$.

We implement this by sampling a unitary $\mathcal{U}$ (ideally from the Haar measure, or equivalently from a unitary 2-design) and preparing and measuring in the rotated basis. Assuming that the SPAM procedures are uncorrelated with the subsequent RB dynamics, we model the implemented preparation and measurement as

$$\rho_{\mathcal{U}} = \mathcal{U} \circ \mathcal{P}(|0\rangle\langle 0|), \quad E_{\mathcal{U}} = \mathcal{U} \circ \mathcal{M}(|0\rangle\langle 0|), \quad (16)$$

where $\mathcal{P}$ and $\mathcal{M}$ are CPTP maps that capture the errors in the preparation and measurement stages, respectively, and $\mathcal{U}$ is the randomizing unitary. This additional randomization effectively twirls the first noise map. Substituting Eq. (16) in the expression for A_x and B , we obtain

$$\begin{aligned} A_x &= q_x \text{Tr} \left(\mathcal{M}(|0\rangle\langle 0|) \left[\mathcal{P}(|0\rangle\langle 0|) - \frac{\mathbb{I}}{d} \right] \right) \\ &= q_x A, \\ B &= \frac{1}{d}. \end{aligned} \quad (17)$$

Thus, we observe that the SPAM errors have been effectively decoupled and we get the ASF for CCC models below.

Theorem 2. For a non-Markovian CCC-type error model, the modified RB protocol under randomization of initial states and final measurements is

$$\bar{F}_{\text{CCC}}^{(m)} = A \sum_x p_x q_x^{m+1} + B, \quad (18)$$

where now A encodes the SPAM errors and B is determined by the dimension of the system.

Note that this procedure would require more resources in terms of the number of runs of the RB experiment; however, this is expected as we are trying to benchmark correlated noise.

Average error rates. Assuming that Eq. (18) can in principle be fitted efficiently to extract the parameters $\{p_x, q_x\}$, the interpretation of the ASF differs significantly from the Markovian case. Although the error maps $\mathcal{N}_x$ are CPTP maps acting only on the system, they are correlated through a common environmental degree of freedom. For Markovian errors, the independence of error maps allows one to assign an average error rate to each gate individually. In contrast, for errors modeled by CCC, it is not meaningful to define a per-gate error rate. Rather, we can provide the average error rate of a sequence of gates of length m . For a given gate sequence S_i of length m , the entire process is described by the CPTP map $\mathcal{R}_{S_i}$ such that $\mathcal{R}_{S_i}(\rho) = \sum_x p_x \mathcal{G}_{m+1} \circ \mathcal{N}_x \cdots \circ \mathcal{G}_1 \circ \mathcal{N}_x(\rho)$. With ideal state preparation and measurement ($\mathcal{P} = \mathcal{M} = \mathcal{I}$), the average fidelity of $\mathcal{R}_{S_i}$ is $\hat{F}_{(\mathcal{R}_{S_i})}^{(m)}$ and its average error rate is $\hat{r}_{(\mathcal{R}_{S_i})} = 1 - \hat{F}_{(\mathcal{R}_{S_i})}^{(m)}$. Averaging over all sequence S_i gives the average error rate of the m length process.

Corollary 1. For general non-Markovian noise, it is not, in general, meaningful to assign a sequence-independent

“average error per gate.” Instead, one should quantify performance at the level of the implemented circuit depth. In particular, for CCC noise models the average error rate for a sequence of length m is

$$r(m) = \frac{d-1}{d} \left(1 - \sum_x p_x q_x^{m+1} \right). \quad (19)$$

Thus, the parameters $\{p_x, q_x\}$ provide a quantitative measure of how far, on average, the system state deviates from the ideal target state after circuit depth m . Even though per-gate error rates are ill defined, the framework still yields a meaningful characterization of overall computational accuracy for sequences of arbitrary length.

An example of multiple exponential fitting. Reference [38] shows that several RB variants under Markovian noise can yield ASFs that are sums of (matrix) exponentials and studies how to extract multiple decay parameters from such data. In particular, the authors use high-resolution spectral estimation methods such as Estimation of Signal Parameters via Rotational Invariance Techniques (ESPRIT) [50] and Multiple Signal Classification (MUSIC) [52], and identify regimes where reliable estimation is possible. Since the CCC ASF derived here is likewise a sum of exponentials, the same techniques apply. Below, we illustrate this by extracting multiple decay parameters for a simple CCC model using ESPRIT.

Consider a qubit process where the system-environment interaction unitary at every time step is the controlled unitary of the form $U^{ES} = |0\rangle\langle 0|^E \otimes U_1 + |1\rangle\langle 1|^E \otimes U_2$, where $U_i = e^{-i\lambda_i |\lambda_i\rangle\langle\lambda_i| + e^{-i\lambda_i^\perp} |\lambda_i^\perp\rangle\langle\lambda_i^\perp|}$ for $i \in \{1, 2\}$, $|\lambda_i\rangle$ and $|\lambda_i^\perp\rangle$ represent an arbitrary vector and the corresponding orthogonal vector, respectively, in the Hilbert space of the system, and $\lambda_i, \lambda_i^\perp$ can take real values. This interaction results in a CCC process, which is a convex sum of two Markovian processes, where the weights depend on the initial environment state (Appendix B). If the initial environment state is an equal superposition of $|0\rangle$ and $|1\rangle$, the weights will be identical and the resulting ASF will be of the form $A(q_1^m + q_2^m)/2 + B$, where the decay parameters $q_i = (4 \cos^2(\Delta\lambda_i/2) - 1)/3$, with $\Delta\lambda_i = \lambda_i - \lambda_i^\perp$ for $i \in \{1, 2\}$. Using this model, we simulate the randomized benchmarking decay curve for parameters $q_1 = 0.9$ and $q_2 = 0.99$. The RB protocol is implemented with 300 random Clifford sequences, each measured over 5000 shots, and sequence lengths ranging from 5 to 200. We first fit the simulated data using a single-exponential decay model, corresponding to the traditional RB fitting approach. Employing the `scipy.optimize` module in Python, we obtain the fitted parameters $q = 0.981$, $A = 0.303$, and $B = 0.538$.

Next, we apply the ESPRIT method (see Appendix D for details on the fitting procedure) to extract two decay components, yielding $q_1 = 0.918$, $q_2 = 0.990$, and corresponding weights $p_1 = 0.451$ and $p_2 = 0.549$. The associated SPAM parameters are found to be $A = 0.435$ and $B = 0.503$ (see Fig. 4). Finally, using Eq. (19), we compare the average error rates derived from the theoretical model with those obtained from the single-exponential and ESPRIT fits, as illustrated in Fig. 5. This example illustrates that the multiple exponent fit parameters provide a physically meaningful interpretation in terms of average errors as defined in Corollary 1 and Eq. (19)

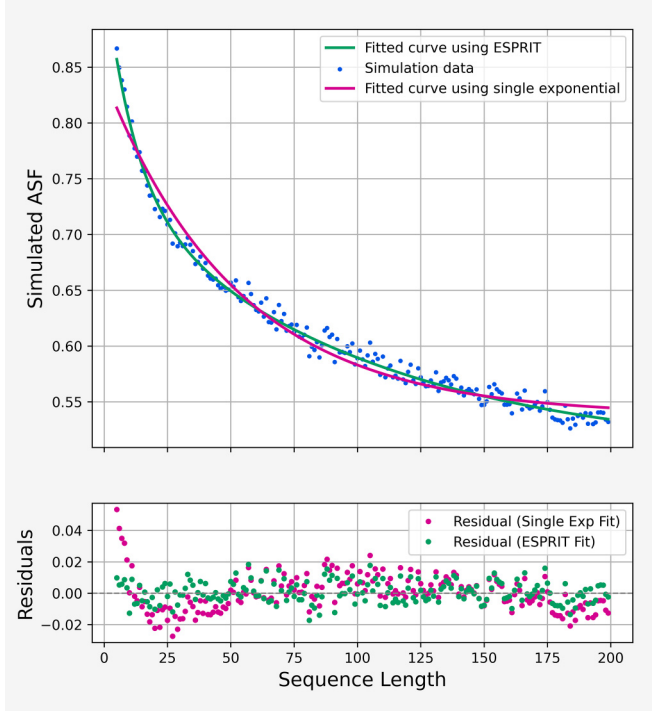


FIG. 4. Comparing the performance of ESPRIT with single exponential fit for the simulated data with exponents 0.9 and 0.99. The root mean square error and adjusted R^2 values for ESPRIT fit are 0.0068 and 0.9908, respectively, while the same parameters take respective values 0.0116 and 0.9735 for single-exponential fitting.

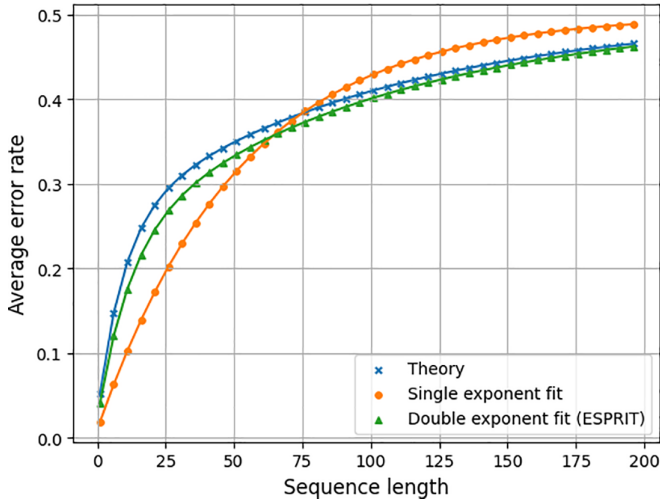


FIG. 5. Average error rate vs sequence length. Comparison of the average error rates predicted by a single-exponential fit with parameter $q = 0.981$ and by a double-exponential model obtained using ESPRIT, with parameters $q_1 = 0.918$ and $q_2 = 0.990$. The single-exponential curve underestimates the average error rate at short sequence lengths, where the ASF decay is dominated by the parameter q_1 . Conversely, at large sequence lengths, the decay is governed primarily by q_2 . The ESPRIT fit (green) also underestimates the theoretical curve (blue), which is due to the sampling of only 300 random Clifford sequences. Nevertheless, the ESPRIT fit provides a clearly improved description of the data.

provides an interpretation of these parameters in presence of CCC-type errors.

2. ASF for classical feedforward

The CFF model generalizes CCC and represents the most general form of *classical* memory in our framework. The central idea is that the environment carries a hidden classical state that can change during the RB sequence. At each time step t , the noise acting on the system is determined by the current value of this classical state. The interaction also produces a classical outcome, which is then used to update the classical state for subsequent steps. As a result, the effective noise at later times may depend on earlier times, but only through this classical record.

Formally, each system-environment interaction at time t is described by a quantum instrument. The instrument applied at time t is selected according to a conditional probability distribution that depends on the past classical record (previous instrument settings and outcomes).

The process matrix for the RB setup in Fig. 3 can be written as

$$W_{\text{CFF}} = \sum_{\vec{x}_n, \vec{a}_n} p(x_1) \rho_{a_1|x_1} \bigotimes_{t=2}^{m+1} p(x_t | \vec{a}_{t-1}, \vec{x}_{t-1}) [\mathcal{N}_{a_t|x_t}],$$

where for the initial state $\rho = \mathcal{U} \circ \mathcal{P}|0\rangle\langle 0|$, we define $\rho_{a_1|x_1} = \mathcal{N}_{a_1|x_1}(\rho)$ and for time step t , $\mathcal{N}_{a_t|x_t}$ is a completely positive linear map such that $\sum_{a_t} \mathcal{N}_{a_t|x_t}$ is a CPTP map. The labels $\vec{x}_t \equiv \{x_1, \dots, x_t\}$ are the instrument settings and $\vec{a}_t \equiv \{a_1, \dots, a_t\}$ are the measurement outcomes of the instruments.

Employing the randomization of initial states and final measurements discussed in previous subsection, with SPAM procedures are described by Eq. (16), the ASF for CFF processes takes the following functional form.

Theorem 3. The ASF for classical feedforward model after applying the randomization of initial states and final measurements is

$$\bar{F}_{\text{CFF}}^{(m)} = A \left(\sum_{\vec{a}, \vec{x}} \prod_{i=0}^m \gamma_{a_i, x_i | \vec{x}_{i-1}, \vec{a}_{i-1}} \right) + B, \quad (20)$$

where $A = \text{Tr}(\mathcal{M}(|0\rangle\langle 0|)[\mathcal{P}(|0\rangle\langle 0|) - \mathbb{I}/d])$, $B = 1/d$, and $\gamma_{a_i, x_i | \vec{a}_{i-1}, \vec{x}_{i-1}} = p(x_i | \vec{a}_{i-1}, \vec{x}_{i-1}) \beta_{a_i | x_i}$ such that

$$\beta_{a_i | x_i} = \frac{1}{d^2 - 1} [\text{Tr}([\mathcal{N}_{a_i | x_i}] | \mathbb{I}) \langle \langle \mathbb{I} | \rangle \rangle - d^{-1} \text{Tr}([\mathcal{N}_{a_i | x_i}])]. \quad (21)$$

Note that here $\sum_{a_i} \text{Tr}([\mathcal{N}_{a_i | x_i}]) = d$.

The decay parameters in Eq. (20) allow one to define an average error rate for length- m sequences, in direct analogy with Eq. (19). In practice, however, reliably extracting these parameters depends on the specifics of the underlying interaction model and on how well separated the decay modes are, since near-degenerate parameters are difficult to resolve with algorithms like ESPRIT. Nevertheless, Eq. (20) provides a principled interpretation of RB outcomes in the presence of CFF-type noise, and a framework for benchmarking devices when such effects are suspected. Moreover, as we will show in the next section, this discussion of the most general

classical-memory noise is important for the witness for quantum non-Markovianity (see Corollary 4).

V. NON-MARKOVIAN ASF AND CONDITIONS FOR BLINDNESS TO RB PROTOCOL

The RB protocol estimates the average error rate of a gate set without incurring the resource costs of full process tomography [53,54]. However, this averaging procedure can also mask important non-Markovian effects. As shown in Sec. IV B, classical-memory error models can produce ASFs with multiple decay parameters, which can in principle reveal signatures of non-Markovian behavior. Nevertheless, there exist scenarios where the resulting ASF is indistinguishable from that of a purely Markovian noise model.

In this section, we establish criteria under which classical-memory errors are effectively *blind* to the RB protocol, rendering them undetectable through standard ASF analysis.

A. Monotonicity of ASF and witness of quantum memory

A deviation in the ASF curve from a single-parameter exponential decay is often interpreted as a signature of non-Markovian dynamics [22,23,26–29,37,55–57]. A definitive indicator is an *increase* in the ASF with sequence length, since under the assumption of Markovian noise, the ASF is $\tilde{F}_M^{(m)} = A \prod_{t=1}^m q_t + B$, where the decay parameters $\{q_t\}$ correspond to CPTP map $\mathcal{N}_t$ and are determined by Eq. (14), i.e.,

$$q_t = c_2(\mathcal{N}_t) = \frac{\text{Tr}(|\mathbb{I}\rangle\langle\mathbb{I}| [[\mathcal{N}_t]]) - 1}{d^2 - 1}, \quad (22)$$

where we have used $\text{Tr}([[\mathcal{N}_t]]) = d$ since $\mathcal{N}_t$ is CPTP. Although the RB decay parameter q_t can in principle take values in the interval $[-1/(d^2 - 1), 1]$, a negative value requires the noise map to be sufficiently far from the identity channel. This notion of distance from identity channel can be quantified by the diamond distance.

Definition 1. The diamond distance between two channels $\mathcal{M}$ and $\mathcal{N}$ is

$$\diamond(\mathcal{M}, \mathcal{N}) = \frac{1}{2} \max_{\chi \in \mathcal{H}_{\text{aux}} \otimes \mathcal{H}_S} \|\mathcal{I} \otimes \mathcal{N}(\chi) - \mathcal{I} \otimes \mathcal{M}(\chi)\|_1, \quad (23)$$

where χ is a quantum state living on the tensor product of system Hilbert space $\mathcal{H}_S$ and an auxiliary Hilbert space $\mathcal{H}_{\text{aux}}$. Since we work in finite-dimensional Hilbert spaces, we can take $\dim(\mathcal{H}_S) = \dim(\mathcal{H}_{\text{aux}})$. The trace norm is defined as $\|A\|_1 = \text{Tr} \sqrt{A^\dagger A}$. The diamond distance gives a well-defined notion of distinguishability between two channels in terms of single-shot discrimination tasks [58,59].

A sufficient condition for non-negativity of decay parameters is provided in the following theorem.

Theorem 4. For a Markovian error model, if the worst-case error, quantified by the diamond distance between the noise map $\mathcal{N}_t$ and identity channel, satisfies the condition

$$\diamond(\mathcal{I}, \mathcal{N}_t) < \frac{d^2 - 1}{d^2}, \quad (24)$$

then the corresponding decay parameters q_t are guaranteed to be non-negative.

Proof. Consider the normalized Choi states $\rho_{\mathcal{I}} = |\mathbb{I}\rangle\langle\mathbb{I}|/d$ and $\rho_{\mathcal{N}_t} = [[\mathcal{N}_t]]/d$, so that $\rho_{\mathcal{I}}$ is a pure state. The entanglement fidelity of the noise map $\mathcal{N}_t$ is given by [60]

$$F_e(\mathcal{N}_t) = \text{Tr}(\rho_{\mathcal{I}} \rho_{\mathcal{N}_t}) = \frac{1}{d^2} \text{Tr}(|\mathbb{I}\rangle\langle\mathbb{I}| [[\mathcal{N}_t]]), \quad (25)$$

and the decay parameter in Eq. (22) is obtained in terms of entanglement fidelity as

$$q_t = \frac{d^2 F_e(\mathcal{N}_t) - 1}{d^2 - 1}. \quad (26)$$

Since $\rho_{\mathcal{I}}$ is pure, the Fuchs-van de Graaf inequality implies that [61]

$$1 - F_e(\mathcal{N}_t) \leq \frac{1}{2} \|\rho_{\mathcal{I}} - \rho_{\mathcal{N}_t}\|_1. \quad (27)$$

Using the relation between the trace norm of Choi matrices and the diamond norm (see Lemma 7 of Ref. [39]),

$$\frac{1}{2} \|\rho_{\mathcal{I}} - \rho_{\mathcal{N}_t}\|_1 = \frac{1}{2d} \|\mathbb{I}\rangle\langle\mathbb{I}| - [[\mathcal{N}_t]]\|_1 \leq \diamond(\mathcal{I}, \mathcal{N}_t), \quad (28)$$

we obtain

$$F_e(\mathcal{N}_t) \geq 1 - \diamond(\mathcal{I}, \mathcal{N}_t). \quad (29)$$

Using Eqs. (26) and (29), we obtain the following lower bound:

$$q_t \geq 1 - \frac{d^2}{d^2 - 1} \diamond(\mathcal{I}, \mathcal{N}_t), \quad (30)$$

leading to the sufficiency condition in Eq. (24) for non-negativity of the decay parameters. ■

In particular, for qubit systems this condition reduces to a threshold of 3/4, so in any well-calibrated experiment, where the diamond error is well below this value, the decay parameters are expected to be positive. Also note that the threshold in Eq. (24) asymptotically approaches unity as system size increases.

For CCC-type errors, whose ASF is given by Eq. (18), we model the noise as a convex combination of Markovian error channels represented by CPTP maps $\mathcal{N}_x$. *A priori*, there is no requirement that the decay parameter q_x associated with a given Markovian branch be strictly positive. For instance, consider the action of a single Clifford gate $\mathcal{C}$ applied to a pure state $|\psi\rangle\langle\psi|$,

$$\mathcal{C}(|\psi\rangle\langle\psi|) = \mathcal{G} \circ \left(\sum_x p_x \mathcal{N}_x(|\psi\rangle\langle\psi|) \right), \quad (31)$$

where $\mathcal{G}$ denotes the ideal gate implementation, p_x are the branch probabilities, and $\mathcal{N}_x$ the associated error channels. The average fidelity of the error channel is given by $\sum_x p_x \int \langle\psi| \mathcal{N}_x(|\psi\rangle\langle\psi|) |\psi\rangle d\psi = \sum_x p_x (q_x + (1 - q_x)/d)$. Since physical gates are expected to have high average fidelity, we impose the condition $\sum_x p_x (q_x + (1 - q_x)/d) = 1 - \epsilon$, for $0 < \epsilon \ll 1$, and therefore $\sum_x p_x q_x = (1 - d\epsilon)/(d - 1)$. Thus, even though the overall gate fidelity may be close to unity, the individual q_x may in principle take negative values for some x . Such unfavorable branches will have correlations along the entire sequence albeit with small probability and may lead to oscillatory behavior in the ASF.

A toy example for such a case could be a convex combination of two Markovian branches $\mathcal{N}_1 = \mathcal{I}$ with $q_1 = 1$ and $\mathcal{N}_2 = \mathcal{X}$ with $q_2 = -1/3$, where $\mathcal{X}$ is a bit-flip operation. If we target the average fidelity of single gate application to be 0.9, the mixture weight on the bit-flip branch is 0.15. The ASF for an m length sequence is $A(0.85 + 0.15(-1/3)^{m+1}) + B$, which exhibits exponentially damped even-odd oscillations. Although this is a contrived example, it illustrates that the ASF for CCC noise models need not be monotonic even though single gate action is taken to be high fidelity. In practice, however, such oscillatory deviations will be strongly suppressed by the exponential damping factor and may be further obscured by statistical fluctuations arising from finite sampling and the use of a limited Clifford gate subset.

Moreover, for a CCC-type model, if we make a reasonable physical assumption that for all noise maps $\mathcal{N}_x$, labeled by a Markovian branch x , the threshold in Eq. (24) holds, i.e., each noise map is sufficiently close to identity channel, then decay parameters q_x are positive for all x . This is justified by observations, for instance, in Ref. [27], where the nonexponential decays observed in silicon quantum-dot qubit were explained by low-frequency drift of qubit resonance frequency, which fits the CCC-type error model [29]. As long as the decay parameters lie in the range $(0, 1)$, it follows that $\bar{F}_{\text{CCC}}^{(m)} > \bar{F}_{\text{CCC}}^{(m+1)} \forall m$, which is evident from

$$\bar{F}_{\text{CCC}}^{(m)} - \bar{F}_{\text{CCC}}^{(m+1)} = A \left(\sum_x p_x q_x^{m+1} [1 - q_x] \right), \quad (32)$$

where parameters $\{p_x\} \in [0, 1]$ are branch weights and $A = \text{Tr}(\mathcal{M}(|0\rangle\langle 0|)[\mathcal{P}(|0\rangle\langle 0|) - \mathbb{I}/d])$. In typical RB settings $A > 0$, reflecting that SPAM errors (captured by $\mathcal{P}$ and $\mathcal{M}$) are small perturbations around the ideal preparation and measurement, and an explicit sufficient condition for $A > 0$ can be derived using the same type of threshold argument as above.

Corollary 2. For a CCC noise model, under the assumption that the noise map corresponding to every Markovian branch x satisfies the sufficiency condition in Eq. (24), the ASF, $\bar{F}_{\text{CCC}}^{(m)}$ is a monotonically decreasing function of the sequence length m .

Analogously, for CFF noise models, the monotonicity condition is

$$\begin{aligned} \bar{F}_{\text{CFF}}^{(m)} - \bar{F}_{\text{CFF}}^{(m+1)} = A \left(\sum_{\vec{a}, \vec{x}} \prod_{i=0}^m \gamma_{a_i, x_i | \vec{a}_{i-1}, \vec{x}_{i-1}} \right) \\ \times \left(1 - \sum_{a_{m+1}, x_{m+1} | \vec{a}_m, \vec{x}_m} \gamma_{a_{m+1}, x_{m+1} | \vec{a}_m, \vec{x}_m} \right), \end{aligned} \quad (33)$$

where $\gamma_{a_i, x_i | \vec{a}_{i-1}, \vec{x}_{i-1}} = p(x_i | \vec{a}_{i-1}, \vec{x}_{i-1}) \beta_{a_i | x_i}$. Since $p(x_i | \vec{a}_{i-1}, \vec{x}_{i-1})$, the relevant decay parameters are $\beta_{a_i | x_i}$. These parameters can take values in the range $[-\text{Tr}([\mathcal{N}_{a_i | x_i}]) / d(d^2 - 1), 1]$, with the constraint $\sum_{a_i} \text{Tr}([\mathcal{N}_{a_i | x_i}]) = d$. Consequently, the ASF can, in principle, also exhibit nonmonotonic behavior under CFF noise. However, as in the CCC case, such deviations are likely to be difficult to observe in experiments.

Corollary 3. For a CFF noise model, assume that for every time step t and every instrument setting/outcome pair (x_t, a_t) , the normalized CP map $\mathcal{N}_{a_t | x_t} / kd$ [where $k = \text{Tr}([\mathcal{N}_{a_t | x_t}])$]

is sufficiently close to the identity in the sense that

$$\diamond \left(\mathcal{I}, \frac{\mathcal{N}_{a_t | x_t}}{kd} \right) < \frac{d^2 - 1}{d^2}. \quad (34)$$

Then, the ASF $\bar{F}_{\text{CFF}}^{(m)}$ is a monotonically decreasing function of the sequence length m .

In Ref. [39], it was observed that any nonmonotonic behavior of the ASF constitutes a witness for non-Markovianity. The following corollary provides a refined version of the claim, namely, a witness for the coherent memory in the environment.

Corollary 4. Under the assumptions of Corollaries 2 and 3, the ASF predicted by CCC and CFF (classical-memory) noise models is a monotonically decreasing function of the sequence length m . Consequently, an experimentally observed *nonmonotonic* ASF is incompatible with these classical-memory models in the above regime, and therefore points to either a violation of the stated assumptions or to additional non-Markovian effects due to quantum memory.

B. Complete blindness to RB

In general, the ASF for classical-memory noise models deviates from a single-exponential decay, making it crucial to resolve and distinguish the different decay parameters present in the RB data. Even in cases where distinct decay parameters can be resolved (with methods like ESPRIT), non-Markovian noise may still mislead optimization strategies. Consider, for instance, a CCC-type model that is a convex combination of two Markovian branches with parameters $\{p, q_1\}$ and $\{1 - p, q_2\}$. Suppose $q_1 > q_2$ and one attempts to tune experimental parameters to bias the system toward the higher-fidelity branch (i.e., $p \approx 1$). While this tuning increases the average sequence fidelity, the physical nature of the corresponding error channels is essential; for instance, consider a scenario where q_1 corresponds to a coherent unitary rotation, whereas q_2 corresponds to a stochastic noise process. It is well known that coherent errors accumulate more adversely than stochastic errors with respect to worst-case error metrics [34]. As a result, simply maximizing average fidelity may not yield the most robust error performance.

On the other hand, consider the case where the decay parameters associated with the two Markovian branches are nearly identical, $q_1 \approx q_2 = q$. In this situation, the ASF reduces to a form that appears indistinguishable from a purely Markovian process with a single decay parameter q . Nevertheless, the underlying noise remains non-Markovian and can lead to nontrivial effects, which we discuss in Sec. VI. This indicates that for non-Markovian errors with classical memory, if all decay parameters are equal, the ASF is *completely blind* to the underlying non-Markovian correlations.

As noted earlier, classical-memory effects can arise from system-environment interactions where the environment itself is quantum mechanical. A natural question, therefore, is whether there exists a class of system-environment interactions that generate classical-memory noise, which is completely blind to the RB protocol. The following theorem provides a sufficient class of RB blind Hamiltonian.

Theorem 5. Consider a system-environment Hamiltonian of the form

$$H^{ES} = \sum_i H_i^E \otimes H_i^S.$$

Assume that the environment operators commute, $[H_i^E, H_j^E] = 0$ for all i, j , with a common eigenbasis $\{|\lambda\rangle\}$ and eigenvalues $\{\lambda_i\}$. Then,

$$H^{ES} = \sum_{\lambda} |\lambda\rangle\langle\lambda| \otimes H_{\lambda}^S, \quad H_{\lambda}^S := \sum_i \lambda_i H_i^S,$$

and the induced multitime process is CCC (a convex mixture of Markovian branches labeled by λ ; see Appendix B). Moreover, if the RB decay parameters of all branches coincide, i.e.,

$$q_{\lambda} = q_{\lambda'} \quad \text{for all } \lambda, \lambda',$$

then the resulting CCC process is *completely RB blind*.

Proof. Since $[H_i^E, H_j^E] = 0$ for all i, j , the operators $\{H_i^E\}_i$ admit a common eigenbasis $\{|\lambda\rangle\}$ with eigenvalues $\{\lambda_i\}$. Expanding H^{ES} in this basis gives

$$H^{ES} = \sum_{\lambda} |\lambda\rangle\langle\lambda| \otimes H_{\lambda}^S,$$

where $H_{\lambda}^S := \sum_i \lambda_i H_i^S$. Tracing over the environment therefore yields a convex mixture of Markovian branches labeled by λ , i.e., a CCC process (see Appendix B).

If the RB decay parameters of all branches coincide, $q_{\lambda} = q_{\lambda'}$ for all λ, λ' , then the ASF collapses to a single exponential with a unique decay parameter. Consequently, the RB signal is identical to that of a Markovian model with the same decay, and RB is completely blind to the underlying CCC correlations.

To make the equal-decay condition explicit, let $\{e_{\lambda,m}\}$ denote the spectrum of H_{λ}^S . For the unitary branch $U_{\lambda}(t) = e^{-itH_{\lambda}^S}$ (with $\hbar = 1$), the corresponding RB decay parameter can be written as

$$q_{\lambda} = \frac{1}{d^2 - 1} (|\text{Tr}(U_{\lambda}(t))|^2 - 1).$$

Equivalently, q_{λ} depends on spectral differences through quantities such as

$$C_{\lambda} := \sum_{m>n} \cos(t(e_{\lambda,m} - e_{\lambda,n})),$$

so a sufficient condition for complete RB blindness is that C_{λ} is invariant over λ . ■

For instance, consider the case when the system and environment both are qubits and the coupling between them is given by $Z \otimes Z$. In this scenario, $H^{ES} = |0\rangle\langle 0| \otimes Z - |1\rangle\langle 1| \otimes Z$ and clearly the above condition is satisfied, resulting in identical decay constants for both the Markovian branches in the CCC process. Clearly, such Hamiltonians lead to non-Markovian noise that is invisible in RB protocol. As we will show in the following section, the memory effects in such processes still have a nontrivial impact on the worst-case error, an important parameter to assess fault tolerance.

VI. WORST-CASE ERRORS

RB protocols are designed to estimate the average error rates of gates, and the protocols are experimentally accessible. However, fault-tolerance thresholds are typically determined by worst-case error metrics, such as the diamond norm, which are not directly measurable in standard experiments [6,35,39]. In particular, for coherent errors, worst-case and average metrics can differ by orders of magnitude [34]. It is therefore essential to analyze worst-case errors in the presence of non-Markovian noise, especially in scenarios where the non-Markovianity is completely hidden from RB, as shown in the previous section.

Consider a qubit system-environment interaction between successive gate applications described by $U^{ES} = \exp(-i\delta Z \otimes Z)$ and $\delta = \pi/100$. Let the initial environment state be

$$|\phi\rangle_E = \sqrt{p}|0\rangle + \sqrt{1-p}|1\rangle. \quad (35)$$

This interaction produces two Markovian branches with identical decay parameters $q_1 = q_2 = q = (|\text{Tr}(\exp(-i\delta Z))|^2 - 1)/3 = 1 - 4\delta^2/3 + \mathcal{O}(\delta^4)$. Consequently, the ASF exhibits a single-exponential decay and the average gate error inferred from such a model is $r_{\text{avg}} \approx 2\delta^2/3 + \mathcal{O}(\delta^4)$.

Given a sequence of gates in the RB procedure $S_{\alpha} \equiv \{\mathcal{G}_1^{\alpha}, \mathcal{G}_2^{\alpha}, \dots, \mathcal{G}_m^{\alpha}, \mathcal{G}_{m+1}^{\alpha}\}$ and redefining the Clifford operations $\mathcal{K}_1^{\alpha} = \mathcal{G}_1^{\alpha}$ and $\mathcal{K}_i^{\alpha} = \mathcal{G}_i^{\alpha} \circ \mathcal{K}_{i-1}^{\alpha} \forall i \geq 2$ as in Sec. IV A, the entire sequence of gates is a CPTP map $\mathcal{R}_{S_{\alpha}}(\rho) = \sum_{x \in \{1,2\}} p_x \mathcal{K}_m^{\alpha \dagger} \circ \mathcal{N}_x \circ \mathcal{K}_m^{\alpha} \dots \mathcal{K}_1^{\alpha \dagger} \circ \mathcal{N}_x \circ \mathcal{K}_1^{\alpha} \circ \mathcal{N}_x(\rho)$ such that

$$\begin{aligned} \mathcal{N}_1(\rho) &= \exp(-i\delta Z)\rho \exp(i\delta Z), \\ \mathcal{N}_2(\rho) &= \exp(i\delta Z)\rho \exp(-i\delta Z), \end{aligned} \quad (36)$$

and $p_1 = p, p_2 = 1 - p$. To quantify the worst-case error, we consider the diamond distance between $\mathcal{R}_{S_{\alpha}}$ and the identity channel $\mathcal{I}$, which represents ideal gate implementations

$$\diamond(\mathcal{R}_{S_{\alpha}}, \mathcal{I}) = \frac{1}{2} \max_{\chi \in \mathcal{H}_{\text{aux}} \otimes \mathcal{H}_S} \|\mathcal{I} \otimes \mathcal{R}_{S_{\alpha}}(\chi) - \mathcal{I} \otimes \mathcal{I}(\chi)\|. \quad (37)$$

The diamond distance in Eq. (37) also depends explicitly on the chosen gate sequence S_{α} . In the presence of temporal correlations, the noise cannot be assigned independently to individual gates, and thus error metrics such as the diamond distance can only be meaningfully defined at the level of an entire sequence rather than per gate. Let $\Lambda_{\chi}^{\alpha} = \mathcal{R}_{S_{\alpha}} \otimes \mathcal{I}(\chi) - \mathcal{I} \otimes \mathcal{I}(\chi)$ and up to second order in δ , we have $\exp(-i\delta Z) = \mathbb{I}(1 - \delta^2/2) - iZ\delta + \mathcal{O}(\delta^3)$. Using this, we get

$$\Lambda_{\chi}^{\alpha} = i(2p - 1)\delta\Lambda_{1,\chi}^{\alpha} - \delta^2\Lambda_{2,\chi}^{\alpha} + \mathcal{O}(\delta^3), \quad (38)$$

such that

$$\begin{aligned} \Lambda_{1,\chi}^{\alpha} &= \sum_{j=1}^m [\chi, \mathbb{I} \otimes P_j^{\alpha}] + [\chi, \mathbb{I} \otimes Z], \\ \Lambda_{2,\chi}^{\alpha} &= (m+1)\chi + \sum_{j=1}^m (\mathbb{I} \otimes P_j^{\alpha} Z)\chi + \sum_{j=1}^m \chi(\mathbb{I} \otimes Z P_j^{\alpha}) \end{aligned}$$

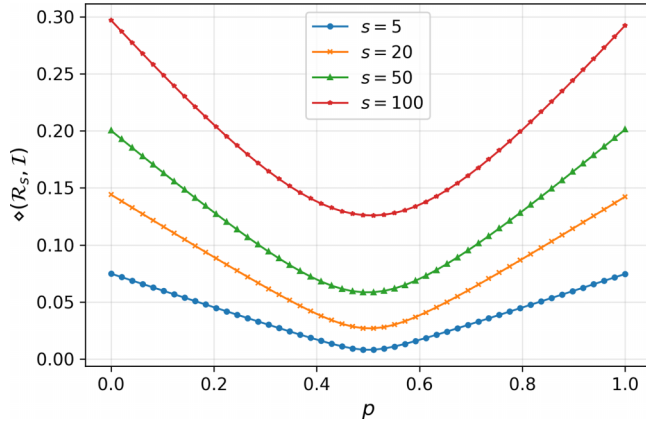


FIG. 6. Diamond norm distance average for various sequence lengths with respect to mixing parameter p in Eq. (35).

$$+ \sum_{j>k} ((\mathbb{I} \otimes P_j^\alpha P_k^\alpha) \chi + \chi (\mathbb{I} \otimes P_j^\alpha P_k^\alpha)) \\ - \left(\sum_{j=1}^m \mathbb{I} \otimes P_j^\alpha + \mathbb{I} \otimes Z \right) \chi \left(\sum_{k=1}^m \mathbb{I} \otimes P_k^\alpha + \mathbb{I} \otimes Z \right), \quad (39)$$

where $P_j^\alpha = K_j^{\alpha\dagger} Z K_j^\alpha$ and $[A, B]$ is the commutator operation. The diamond distance in Eq. (37) now becomes $\max_{\chi \in \mathcal{H}_S \otimes \mathcal{H}_S} \|\sqrt{\Lambda_\chi^{\alpha\dagger} \Lambda_\chi^\alpha}\|_1$. When $p = 1$ or $p = 0$ (purely Markovian coherent errors), $\Lambda_\chi^\alpha = \pm i \delta \Lambda_{1,\chi}^\alpha + \mathcal{O}(\delta^2)$, whereas for $p = 0.5$ (maximal mixing between the two Markovian branches), $\Lambda_\chi^\alpha = -\delta^2 \Lambda_{2,\chi}^\alpha$. Although calculating Eq. (37) exactly requires knowledge of the specific sequence S_α , these results suggest scalings:

$$p = \{0, 1\} \rightarrow \diamond(R_{S_\alpha}, \mathcal{I}) \approx \mathcal{O}(\delta) = \mathcal{O}(\sqrt{r_{\text{avg}}}), \\ p = 0.5 \rightarrow \diamond(R_{S_\alpha}, \mathcal{I}) \approx \mathcal{O}(\delta^2) = \mathcal{O}(r_{\text{avg}}). \quad (40)$$

We verified this numerically. In order to investigate the variation of worst-case error for different sequence lengths with respect to the mixing parameter p , we exploit the fact that the diamond norm distance between two channels can be cast in the form of a semidefinite program [62]. In particular, for a given sequence S_α and mixing parameter p , the diamond norm distance of the resulting effective channel $\mathcal{R}_{S_\alpha}$ from the identity channel is obtained as

$$\max \quad \frac{1}{2} \text{Tr}[X([\mathcal{R}_{S_\alpha}] - [\mathcal{I}])], \\ \text{such that} \quad -2(\mathbb{I} \otimes \rho) \leq X \leq 2(\mathbb{I} \otimes \rho), \\ \rho \geq 0, \quad \text{Tr}[\rho] = 1. \quad (41)$$

For a given sequence length, we sample 50 random Clifford sequences and find the average diamond norm distance corresponding to mixing parameter p . We plot the results for various sequence lengths for mixing parameters ranging from zero to one in Fig. 6. As evident from the graph, the diamond norm distance is minimum for the parameter value $p = 0.5$, i.e., for the maximal CCC process, and maximum for $p = 0, 1$, i.e., when the process is a single Markovian branch with coherent errors. Importantly, this result suggests that, for

present transmon-based quantum hardware, where a natural ZZ coupling is present, it is useful to initiate the environmental qubits in the superposed state of ground and excited states (such as $|+\rangle$) rather than the ground state for reduced worst-case errors. More broadly, our findings indicate that non-Markovian effects are not necessarily detrimental to gate performance. Understanding how non-Markovian memory effects influence the worst-case error remains an interesting direction for further investigation.

VII. DISCUSSION AND CONCLUSION

In this work, we developed a general framework to study RB under non-Markovian noise models that introduce classical temporal correlations, either through stochastic classical fields or interactions with a quantum environment. We derived analytical expressions for the ASF in such settings and demonstrated that, unlike the standard Markovian case, the ASF can be expressed as a sum of multiple exponential decays. We showed how the additional decay parameters can be interpreted to yield useful information for estimating average errors, and how established signal-processing techniques such as ESPRIT and MUSIC can be employed to extract these parameters from RB data.

Our analysis revealed several important features. First, while ASF curves with classical correlations can in principle exhibit nonmonotonic behavior, such deviations are exponentially suppressed, hence difficult to observe in practice. When all decay parameters remain positive, the ASF is guaranteed to decrease monotonically, making non-Markovian effects invisible to monotonicity-based diagnostics. Thus, any observed violation of a monotonically decreasing ASF is a strong indicator of genuinely quantum-memory effects. Moreover, we identified classes of interaction Hamiltonians that generate classical-memory processes completely blind to RB, yielding ASF curves indistinguishable from Markovian noise despite the presence of temporal correlations. Finally, we investigated the implications for worst-case error metrics. While RB provides estimates of average gate errors, fault-tolerance thresholds are governed by worst-case performance quantified by the diamond norm [34,35,37]. We argued that classical-memory processes can alter worst-case error behavior, and in some scenarios, non-Markovian effects may even reduce worst-case errors. These findings underline the importance of going beyond average metrics when assessing the reliability of quantum devices especially in the presence of correlated noise.

Overall, our results clarify both the capabilities and blind spots of randomized benchmarking in the presence of temporal correlations. They also motivate the development of complementary protocols capable of jointly estimating average and worst-case error metrics under realistic noise models [34,39,63]. Moreover, it is important to come up with theoretical bounds on worst-case errors in case of non-Markovian noise, similar to the bounds presented in Ref. [39] for Markovian noise models. It would also be interesting to explore whether the RB protocol can be modified to reliably provide an operational metric on the amount of non-Markovian noise present during gate implementations. Finally, we note that the present analysis has been limited to Clifford operations under

the assumption of gate-independent errors. We anticipate that our framework can be readily extended to more general gate sets, and that established techniques [17,23] can be naturally integrated to treat these broader scenarios.

ACKNOWLEDGMENTS

V.S., A.K.R., S.M., and J.K acknowledge funding from the Sydney Quantum Academy. This work was supported by the Hon-Hai Research Institute through the Australian Quantum Software Network (AQSN) microgrant. We also wish to acknowledge support from the ARC Centre of Excellence for Engineered Quantum Systems, EQUS (CE170100009), and S.K. acknowledges support through an EQUS Deborah Jin Fellowship.

DATA AVAILABILITY

There are no publicly available research data or software supporting this manuscript. Requests for further information or data should be sent to the authors.

APPENDIX A: LINK PRODUCT AND CONSTRUCTING TIME-ORDERED PROCESS MATRIX

Here, we provide a brief summary of the link product and the use of it to construct multitime processes.

The link product ($\star$) of two operators $A \in \mathcal{L}(\bigotimes_{p_j \in p} \mathcal{H}_{p_j})$ and $B \in \mathcal{L}(\bigotimes_{q_j \in q} \mathcal{H}_{q_j})$ is defined as

$$A \star B = \text{Tr}_{p \cap q}[(\mathbb{I}_{q \setminus p} \otimes A^{T_{p \cap q}})(B \otimes \mathbb{I}_{p \setminus q})],$$

$$\in \mathcal{L}(\mathcal{H}_{q \setminus p} \otimes \mathcal{H}_{p \setminus q}), \quad (\text{A1})$$

where the operators $(\cdot)^{T_{p \cap q}}$ and $\text{Tr}_{p \cap q}$ represent the partial transpose and partial trace, respectively, over the common space $p \cap q$. In particular, when the operators do not share common space, link product reduces to the usual tensor product; i.e., $A \star B = A \otimes B$ when $p \cap q = \emptyset$. The link product has the following useful properties. It is (1) Hermiticity preserving, i.e., A, B Hermitian $\Rightarrow A \star B$ is Hermitian, (2) positivity preserving, i.e., $A, B \geq 0 \Rightarrow A \star B \geq 0$, (3) commutative $A \star B = B \star A$, and (4) associative, i.e., $A \star (B \star C) = (A \star B) \star C$, where operators A, B , and C do not have any joint overlap.

Note that the action of a map $\mathcal{M} : S_I \rightarrow S_O$ on state ρ^{S_I} in terms of the link product is

$$\begin{aligned} \mathcal{M}(\rho^{S_I}) &= \text{Tr}_{S_I}([\mathcal{M}]^{S_I S_O}(\rho^{T_{S_I}} \otimes \mathbb{I}^{S_O})) \\ &= [[\mathcal{M}]]^{S_I S_O} \star \rho^{S_I} \\ &= \rho^{S_I} \star [[\mathcal{M}]]^{S_I S_O}. \end{aligned} \quad (\text{A2})$$

Consider a multitime scenario with N interventions depicted by the labs $A_1, A_2, \dots, A_N$. In general, the initial system and environment can be correlated and given by the joint state $\rho^{A_1^I E_1}$, where we use the superscript $I(O)$ to denote the input (output) space of the labs. The local operations in the lab A_i are elements of the instrument $\{\mathcal{M}_{a_i} : A_i^I \rightarrow A_i^O\}$, and between the labs A_i and A_{i+1} , there is a joint system-environment interaction given by the unitary $\mathcal{U}_i : E_i A_i^O \rightarrow E_{i+1} A_{i+1}^I$. With this setting, the probability of obtaining the

outcomes $a_1, a_2, \dots, a_N$ is given by the Born rule,

$$\mathbb{P}(a_1, a_2, \dots, a_N) = \text{Tr}_{E_N A_N^O}[\mathcal{M}_{a_N}(\mathcal{U}_{N-1} \cdots \times \cdots (\mathcal{U}_2(\mathcal{M}_{a_2}(\mathcal{U}_1(\mathcal{M}_{a_1}(\rho)))))). \quad (\text{A3})$$

Using Eq. (A2) recursively, and using the commutativity and associativity properties of the link product, one can rearrange the above to obtain the following:

$$\mathbb{P}(a_1, a_2, \dots, a_N) = \text{Tr}[W^T \cdot M], \quad (\text{A4})$$

where W and M are operators in the space $\bigotimes_i (A_i^I \otimes A_i^O)$ of the following form:

$$W = \rho^{E_1 A_1^I} \star (\star_{i=1}^{N-1} [[\mathcal{U}_i]]^{E_i A_i^O E_{i+1} A_{i+1}^I}) \star \mathbb{I}^{E_N} \quad (\text{A5})$$

and

$$M = \bigotimes_i [[\mathcal{M}_{a_i}]]^{A_i^I A_i^O}. \quad (\text{A6})$$

The object W is known as the process matrix and is a bounded positive operator. This mapping of multitime processes in terms of positive operator allows one to employ tools from many body systems for investigating multitime processes. For example, in the Markovian scenario, where the system is initialized in state $\rho^{A_1^I}$, followed by the application of channels $\mathcal{E}_1, \dots, \mathcal{E}_{n-1}$ is given by

$$W = \rho_0 \star [[\mathcal{E}_1]] \star \cdots \star [[\mathcal{E}_{n-1}]] \star \mathbb{I}. \quad (\text{A7})$$

Since there are no shared spaces, the link product reduces to the usual tensor product, resulting in the process matrix of the following form:

$$W = \rho_0 \otimes [[\mathcal{E}_1]] \otimes \cdots \otimes [[\mathcal{E}_{n-1}]] \otimes \mathbb{I}. \quad (\text{A8})$$

That is, the process matrix of a Markovian process is similar to an uncorrelated state. A process that cannot be factorized as above is called a non-Markovian process.

APPENDIX B: HAMILTONIAN DESCRIPTION OF NON-MARKOVIAN PROCESSES WITH CLASSICAL MEMORY

As mentioned in the main text, the classical-memory process can result from interaction of the system with a quantum environment. For instance, consider the following class of environment-system Hamiltonian:

$$H^{ES} = \sum_x H_x^E \otimes H_x^S, \quad (\text{B1})$$

such that $[H_x^E, H_y^E] = 0 \ \forall x, y$ [30]. The constraint on the environment part of the Hamiltonian implies that there is a common eigenbasis $\{|\lambda\rangle\}$ for the set $\{H_x^E\}$, so one can rearrange H^{ES} to the form $H^{ES} = \sum_\lambda |\lambda\rangle\langle\lambda|^E \otimes H_\lambda^S$, where $H_\lambda^S = \sum_x \lambda^x H_x^S$ and we will replace the Hilbert space notation $\mathcal{H}^{S/E}$ with simply S/E for convenience. If the interaction time is Δt , the resulting unitary in the Choi form is given by $[[U]]^{ESE'S'} = \sum_{\lambda, \lambda'} |\lambda\lambda\rangle\langle\lambda'\lambda'|^{EE'} \otimes \tilde{U}_\lambda \langle\tilde{U}_{\lambda'}|^{SS'}$, where $\tilde{U}_\lambda = \exp(-i\Delta t H_\lambda)$. Now, consider a three intervention scenario with labs A, B , and C , and the interactions between the labs active for t_{AB} and t_{BC} , respectively, and let $A_{I/O}$ denote the input/output Hilbert space for lab A . If the process starts with

an arbitrary correlated environment-system state $\rho^{E_1 A_I}$, the resulting process matrix is given by

$$W = \rho^{E_1 A_I} \star [[U(t_{AB})]]^{E_1 A_O E_2 B_I} \star [[U(t_{BC})]]^{E_2 B_O E_3 C_I} \star \mathbb{I}^{E_3}, \quad (\text{B2})$$

where A_i, B_i, C_i denote the corresponding system Hilbert space. Using the Choi form of the interaction unitaries results in the following:

$$W = \sum_{\lambda\lambda'\sigma\sigma'} (\rho^{E_1 A_I} \star |\lambda\rangle\langle\lambda'|^{E_1 E_2} \star |\sigma\rangle\langle\sigma'|^{E_2 E_3} \star \mathbb{I}^{E_3}) \otimes |\tilde{U}_\lambda\rangle\langle\tilde{U}_{\lambda'}|^{A_O B_I} \otimes |\tilde{U}_\sigma\rangle\langle\tilde{U}_{\sigma'}|^{B_O C_I}, \quad (\text{B3})$$

where we have used the fact that the link product between objects without common spaces is the usual tensor product. A straightforward simplification results in the following form:

$$W = \sum_{\lambda} (\rho^{A_I E_1} \star |\lambda\rangle\langle\lambda|^{E_1}) \otimes [[\tilde{U}_\lambda]]^{A_O B_I} \otimes [[\tilde{U}_\lambda]]^{B_O C_I} = \sum_{\lambda} p_\lambda W_\lambda, \quad (\text{B4})$$

where W_λ is Markovian process with the weight $p_\lambda = \text{Tr}(\rho^{A_I E_1} \star |\lambda\rangle\langle\lambda|^{E_1})$. Therefore, the above class of Hamiltonian results in a convex sum of Markovian processes, i.e., CCC process regardless of the amount of initial correlation in the environment-system state. The proof is straightforward to extend to arbitrary number of time steps and can be generalized to include time dependence in the Hamiltonian acting between the labs (for more details, we refer the readers to Ref. [30]). Importantly, this class of Hamiltonian includes the generator of the CNOT gate, and Ising interactions such as ZZ coupling, which is an always-on interaction between superconducting qubits in present quantum hardware [31,32].

In a similar manner, one can obtain a classical feedforward memory process through interaction with quantum environments. This case, however, is not as straightforward as the CCC case, where the interaction was with a common environment at every time step. For instance, again consider a three intervention scenario, with the interaction during the first time step of the form $H_1 = \sum_m H_m^{E_0} \otimes H^{E_1 S} \otimes \mathbb{I}^{E_2}$ and the interaction during the second time step of the form $H_2 = \sum_m H_m^{E_0} \otimes H_m^{E_1} \otimes H_m^{E_2 S}$. If the following conditions are satisfied,

- (1) $[H_m^{E_0}, H_n^{E_0}] = 0$ for all m, n and for both H_1 and H_2 ,
- (2) $[H_m^{E_1}, H_n^{E_1}] = 0$ for all m, n for H_2 ,
- (3) initial system-environment state is uncorrelated in spaces E_1 and E_2 ,

then it is straightforward to show that the resulting process matrix is of the form

$$W = \sum_{xy} p_x \rho_x^{A_I} \otimes [[\mathcal{E}_{y|x}]]^{A_O B_I} \otimes [[\mathcal{T}_{xy}]]^{B_O C_I}. \quad (\text{B5})$$

Here, $[[\mathcal{E}_{y|x}]]$ is CJ map corresponding to the CP map $\mathcal{E}_{y|x}$, with $\sum_y \mathcal{E}_{y|x}$ being a CPTP map for all x , and $[[\mathcal{T}_{xy}]]$ are CJ maps corresponding to CPTP maps, and $p(x)\rho_x^{A_I}$ has its origin in the reduced states of the initial ES state in some basis $|x\rangle$ of Hilbert space of E_0 . Evidently, the process is not a classical common cause process but a more general classical-memory process with feedforward.

APPENDIX C: ASF CALCULATION IN PROCESS MATRIX FRAMEWORK

1. Derivation of the ASF

Here, we provide a step-by-step derivation of Eq. (12). Consider an RB protocol with $m+1$ gates, where the first m gates are picked randomly from the Clifford group of n qubits, $\mathcal{CF}_n$, and the last gate is the motion-reversal gate. For a given sequence of gates S_α , each noisy gate $C_t^{(\alpha)}$, applied at time step t , is decomposed into the noisy system-environment unitary and the ideal gate as $C_t \equiv (\mathcal{I} \otimes \mathcal{G}_t) \circ \mathcal{U}_t^{SE}$. The process matrix is then evaluated as

$$W = (\rho \otimes \sigma) \star [[\mathcal{U}_1^{SE}]] \star \dots \star [[\mathcal{U}_{m+1}^{SE}]] \star \mathbb{I}^E, \quad (\text{C1})$$

where the identity at the end of the equation is equivalent to tracing out the environmental degrees of freedom.

For this derivation, it would be helpful to keep an account of the relevant Hilbert spaces. Let $(\rho \otimes \sigma) \in \mathcal{L}(\mathcal{H}_0^{S_O} \otimes \mathcal{H}^E)$ and $[[\mathcal{U}_t^{SE}]] \in \mathcal{L}(\mathcal{H}_{t-1}^{S_O} \otimes \mathcal{H}^E \otimes \mathcal{H}_t^{S_I} \otimes \mathcal{H}^E)$, where the labels S_O, S_I refer to the output/input Hilbert spaces of the system of interest at time step t . Note that although we have not labeled the environment with temporal labels, they should also be contracted in the correct order. Now notice that in the RB protocol, the first unitary does not get twirled; therefore, we can define a system-environment state $\lambda = (\rho \otimes \sigma) \star [[\mathcal{U}_1^{SE}]]$. This can lead to SPAM errors influencing the temporal correlations. The process matrix therefore is a positive linear operator $W \in \mathcal{L}(\mathcal{H}_1^{S_I} \otimes \mathcal{H}_1^{S_O} \otimes \dots \otimes \mathcal{H}_m^{S_I} \otimes \mathcal{H}_m^{S_O} \otimes \mathcal{H}_{m+1}^{S_I})$.

Similarly, the instruments in the form of ideal Clifford gates can be composed into a higher-order operator as

$$\mathbb{T}_\alpha = [[\mathcal{G}_1^{(\alpha)}]] \otimes \dots \otimes [[\mathcal{G}_m^{(\alpha)}]] \otimes [[\mathcal{G}_{m+1}^{(\alpha)}]] \star E^T, \quad (\text{C2})$$

where $\mathbb{T}_\alpha \in \mathcal{L}(\mathcal{H}_1^{S_I} \otimes \mathcal{H}_1^{S_O} \otimes \dots \otimes \mathcal{H}_m^{S_I} \otimes \mathcal{H}_m^{S_O} \otimes \mathcal{H}_{m+1}^{S_I})$ is another positive linear operator that encodes the operations that were intended to be applied in the noiseless limit. Note that the link product is a tensor product if the operators do not share common Hilbert spaces.

Each ideal gate can be represented in the Choi form $[[\mathcal{G}_t^{(\alpha)}]] \in \mathcal{L}(\mathcal{H}_t^{S_I} \otimes \mathcal{H}_t^{S_O})$ and the Choi representation of the POVM element is $E^T \in \mathcal{L}(\mathcal{H}_{m+1}^{S_O})$, which can be contracted with the last motion-reversal gate as $[[\mathcal{G}_{m+1}^{(\alpha)}]] \star E^T = (G_{m+1}^{\dagger(\alpha)} E G_{m+1}^{(\alpha)})^T$. We then redefine the Clifford gates as $\mathcal{K}_1^{(\alpha)} = \mathcal{G}_1^{(\alpha)}$ and $\mathcal{K}_i^{(\alpha)} = \mathcal{G}_i^{(\alpha)} \circ \mathcal{K}_{i-1}^{(\alpha)} \forall i \geq 2$. Then, for any time $t \geq 2$ we can write

$$\begin{aligned} [[\mathcal{G}_t^{(\alpha)}]] &= [[\mathcal{K}_t^{(\alpha)} \circ \mathcal{K}_{t-1}^{\dagger(\alpha)}]] \\ &= |K_t^{(\alpha)} K_{t-1}^{\dagger(\alpha)}\rangle\langle K_t^{(\alpha)} K_{t-1}^{\dagger(\alpha)}| \\ &= K_{t-1}^{*(\alpha)} \otimes K_t^{(\alpha)} |\mathbb{I}\rangle\langle\mathbb{I}| K_{t-1}^{T(\alpha)} \otimes K_t^{\dagger(\alpha)}, \end{aligned} \quad (\text{C3})$$

where for an operator A , its complex conjugate is A^* . We can rewrite Eq. (C2) as

$$\mathbb{T}_\alpha = \mathbb{K}^T(|\mathbb{I}\rangle\langle\mathbb{I}| \otimes E^T) \mathbb{K}^*,$$

where

$$\mathbb{K}^{(\alpha)} := \mathbb{I} \bigotimes_{t=1}^m (K_t^{T(\alpha)} \otimes K_t^{\dagger(\alpha)}).$$

The sequence fidelity of a given sequence S^α is then the contraction of the process matrix with the instrument operator

$$\begin{aligned} F_\alpha^{(m)} &= W * \mathbb{T}_\alpha \\ &= \text{Tr}(W^T \mathbb{K}^T (|\mathbb{I}\rangle\rangle\langle\langle\mathbb{I}| \otimes E^T) \mathbb{K}^*) \\ &= \tilde{W}_\alpha \star \Gamma, \end{aligned} \quad (\text{C4})$$

where $\tilde{W}_\alpha = \mathbb{K}^{(\alpha)} W \mathbb{K}^{\dagger(\alpha)}$ and $\Gamma = |\mathbb{I}\rangle\rangle\langle\langle\mathbb{I}| \otimes E^T$.

The average sequence fidelity is the average of sequence fidelity $F_\alpha^{(m)}$ over uniformly selected gates from the Clifford group. This expression is given in Eq. (12):

$$\bar{F}^{(m)} = \frac{1}{\Omega^m} \sum_{\alpha=1}^{\Omega^m} \tilde{W}_\alpha \star \Gamma, \quad \Omega = |\mathcal{CF}_n|. \quad (\text{C5})$$

2. Unitary 2-designs and Haar second moment

Consider an operator $O \in \mathcal{L}(\mathcal{H} \otimes \mathcal{H})$, then the Haar second moment is defined as

$$\begin{aligned} \mathbb{E}_{U \sim \mu_H} [U^{\otimes 2} O (U^\dagger)^{\otimes 2}] &= \int_U d\mu_H [U^{\otimes 2} O (U^\dagger)^{\otimes 2}] \\ &= c_{\mathbb{I},O} \mathbb{I} + c_{\mathbb{F},O} \mathbb{F}, \end{aligned} \quad (\text{C6})$$

where

$$\begin{aligned} c_{\mathbb{I},O} &= \frac{\text{Tr}(O) - d^{-1} \text{Tr}(\mathbb{F}O)}{d^2 - 1}, \\ c_{\mathbb{F},O} &= \frac{\text{Tr}(\mathbb{F}O) - d^{-1} \text{Tr}(O)}{d^2 - 1}, \end{aligned} \quad (\text{C7})$$

where $\mathbb{F}$ is the SWAP operator and d is the dimension of Hilbert space $\mathcal{H}$.

This is a consequence of Schur-Weyl duality that relates the second moment over a Haar measure to the corresponding permutation matrices acting on $\mathcal{H} \otimes \mathcal{H}$, namely, $\{\mathbb{I}, \mathbb{F}\}$. The Clifford group is particularly useful because it forms a unitary 2-design, which means that its second moment under the uniform distribution matches the Haar second moment [64].

In Eq. (13), we saw that the ASF for the time-dependent Markovian model is

$$\bar{F}_M^{(m)} = (\tilde{\rho} \otimes Z_2 \otimes \cdots \otimes Z_{m+1}) \star \Gamma, \quad (\text{C8})$$

where

$$Z_t := \frac{1}{\Omega} \sum_{K_t \in \mathcal{CF}_n} (K_t^{T(\alpha)} \otimes K_t^{\dagger(\alpha)}) [[\mathcal{N}_t]] (K_t^{T(\alpha)} \otimes K_t^{\dagger(\alpha)})^\dagger$$

and Ω is the number of elements in the Clifford group of n qubits. Now, $Z_t \in \mathcal{L}(\mathcal{H}_{t-1}^{S_O} \otimes \mathcal{H}_t^{S_I})$, and then taking a partial transpose over the Hilbert space $\mathcal{H}_{t-1}^{S_O}$ gives

$$Z_t^{T_{S_O}} := \frac{1}{\Omega} \sum_{K_t \in \mathcal{CF}_n} (K_t^{\dagger(\alpha)} \otimes K_t^{\dagger(\alpha)}) [[\mathcal{N}_t]] (K_t^{(\alpha)} \otimes K_t^{(\alpha)}), \quad (\text{C9})$$

which is equal to the Haar second moment. Since partial transpose of $\mathbb{F} = |\mathbb{I}\rangle\rangle\langle\langle\mathbb{I}|$, it is straightforward to verify that

$$Z_t = c_1(\mathcal{N}_t) \frac{\mathbb{I}}{d} + c_2(\mathcal{N}_t) |\mathbb{I}\rangle\rangle\langle\langle\mathbb{I}|, \quad (\text{C10})$$

where the coefficients

$$\begin{aligned} c_1(\mathcal{N}_t) &= \frac{d \text{Tr}([\mathcal{N}_t]) - \text{Tr}(|\mathbb{I}\rangle\rangle\langle\langle\mathbb{I}|) \text{Tr}([\mathcal{N}_t])}{d^2 - 1}, \\ c_2(\mathcal{N}_t) &= \frac{\text{Tr}(|\mathbb{I}\rangle\rangle\langle\langle\mathbb{I}|) \text{Tr}([\mathcal{N}_t]) - d^{-1} \text{Tr}([\mathcal{N}_t])}{d^2 - 1} \end{aligned} \quad (\text{C11})$$

are related as $c_1(\mathcal{N}_t) + c_2(\mathcal{N}_t) = \text{Tr}([\mathcal{N}_t])/d$. For any CPTP map, $\text{Tr}([\mathcal{N}_t]) = d$ and therefore $c_1(\mathcal{N}_t) + c_2(\mathcal{N}_t) = 1$. Equation (C8) can then be evaluated to give the ASF

$$\bar{F}_M^{(m)} = A \prod_{t=1}^m c_2(\mathcal{N}_t) + B,$$

where $A = \text{Tr}[E(\mathcal{N}_1(\rho) - \mathbb{I}/d)]$ and $B = \frac{1}{d} \text{Tr}(E)$.

APPENDIX D: DATA FITTING FOR RANDOMIZED BENCHMARKING

Here, we present our data fitting algorithm and examples of extracting the decay parameters from RB data for different CCC processes, where the origin of the RB data is a model comprising two exponential decay terms and additional noise:

$$y_i = A_1 q_1^{i+1} + A_2 q_2^{i+1} + B + n_i, \quad (\text{D1})$$

where $A_1 = Ap_1$ and $A_2 = Ap_2$. The constant terms A and B capture the information of SPAM errors, while p_1 and p_2 denote the probabilities corresponding to the two different Markovian branches of the CCC process; and n_i is a noise term representing the deviation from the true ASF due to finite sampling or other possible noise. The fitting program aims to extract the probabilities p_1, p_2 , the constants A, B , and the exponents q_1, q_2 from the provided data $\{y_i\}$. For data fitting, we employ ESPRIT technique followed by solving a linear equation to extract the required parameters. The complete workflow comprises the following steps.

(1) The ASF data for a sequence length m $[y_1, y_2, \dots, y_m]$ can be expressed in the following matrix equation:

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} q_1^2 & q_2^2 & 1 \\ q_1^3 & q_2^3 & 1 \\ \vdots & \vdots & \vdots \\ q_1^{m+1} & q_2^{m+1} & 1 \end{pmatrix} \begin{pmatrix} A_1 \\ A_2 \\ B \end{pmatrix} + \begin{pmatrix} n_1 \\ n_2 \\ \vdots \\ n_m \end{pmatrix}.$$

(2) Splitting and writing the data in the form of a Hankel matrix of dimension $l \times (m-l+1)$:

$$\begin{pmatrix} y_1 & \cdots & y_{m-l+1} \\ \vdots & \ddots & \vdots \\ y_l & \cdots & y_m \end{pmatrix} = \begin{pmatrix} q_1^2 & q_2^2 & 1 \\ \vdots & \vdots & \vdots \\ q_1^{l+1} & q_2^{l+1} & 1 \end{pmatrix} \times \begin{pmatrix} A_1 & \cdots & A_1 q_1^{m-l} \\ A_2 & \cdots & A_2 q_2^{m-l} \\ B & \cdots & B \end{pmatrix} + \begin{pmatrix} n_1 & \cdots & n_{m-l+1} \\ \vdots & \ddots & \vdots \\ n_l & \cdots & n_m \end{pmatrix}, \quad (\text{D2})$$

we obtain its singular values. Equation (D2) can be rewritten in shorthand as $Y = VA + N$, where Y is the Hankel matrix, V is called the Vandermonde matrix, and N is the noise matrix.

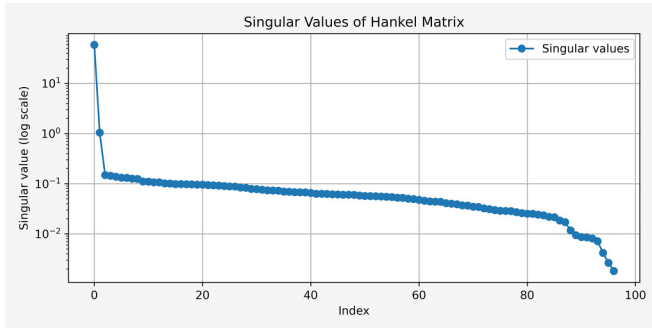


FIG. 7. The plot shows singular values of the Hankel matrix from the simulated data presented in Sec. VB. The singular values are plotted in logarithmic scale on the Y axis, and the two dominant singular values are clearly distinguished from the rest, which correspond to two different exponents present in the data.

We choose $l = \lfloor m/2 \rfloor$, where $\lfloor \cdot \rfloor$ is the floor function. The dominant singular values of the Hankel matrix denote the number of different exponents (or decay parameters) present in the data (see Fig. 7).

(3) The number of dominant singular values (supplied as the “model order”) as well as the data serves as input for the ESPRIT algorithm, which outputs the different exponents.

(4) Using the exponents, we determine A_1 , A_2 , and constant B by solving the linear equation [Eq. (D2)].

(5) Once we obtain the parameter values for A_1 and A_2 , we can find p_1 and p_2 from the following relations: $p_1 = \frac{A_1}{A_1 + A_2}$ and $p_2 = \frac{A_2}{A_1 + A_2}$.

ESPRIT algorithm is based on a rotational invariance property of the Vandermonde matrix. This relation takes the form

$$J_2 V = J_1 V H, \quad (\text{D3})$$

where J_1 and J_2 are the last and first row deletion matrices, respectively, and H is the diagonal matrix comprising the

exponents

$$J_1 = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix}_{l-1 \times l} = [\mathbf{I}_{l-1} \mathbf{0}],$$

$$J_2 = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}_{l-1 \times l} = [\mathbf{0} \mathbf{I}_{l-1}],$$

$$H = \begin{pmatrix} q_1 & 0 & 0 \\ 0 & q_2 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

The singular value decomposition of the Hankel matrix in Eq. (D2) can be broken into two parts, and by selecting the dominant singular values as the signal and the rest as noise, we obtain $Y = U_s \Sigma_s W_s^\dagger + U_n \Sigma_n W_n^\dagger$. Here, Σ_s contain the dominant singular values that are identified as signal and U_s, W_s are the corresponding parts of the isometries U, W that constitute the singular value decomposition. Similarly, U_n, W_n correspond to the noise part. Comparing it to $Y = VA + N$, one can write the signal part as $VA = U_s \Sigma_s W_s^\dagger$ or $V = U_s \Sigma_s W_s^\dagger A^{-1}$. Replacing this relation in Eq. (D3), it follows that

$$J_2 U_s = J_1 U_s \Phi, \quad (\text{D4})$$

where the eigenvalues of Φ correspond to the exponents of the data. The ESPRIT algorithm uses Eq. (D4) and upon solving the linear equation for an unknown Φ , gives out the eigenvalues (exponents q_1, q_2). For the ideal case with no noise and sequence length 200, it can perfectly resolve different exponents with differences as close as 10^{-4} . In the presence of noise, the performance depends both on the noise strength and the separation between the exponents (see Ref. [38] for details).

-
- [1] J. Eisert, D. Hangleiter, N. Walk, I. Roth, D. Markham, R. Parekh, U. Chabaud, and E. Kashefi, Quantum certification and benchmarking, *Nat. Rev. Phys.* **2**, 382 (2020).
 - [2] R. Blume-Kohout, T. Proctor, and K. Young, Quantum characterization, verification, and validation, [arXiv:2503.16383](https://arxiv.org/abs/2503.16383).
 - [3] J. Emerson, R. Alicki, and K. Życzkowski, Scalable noise estimation with random unitary operators, *J. Opt. B: Quantum Semiclassical Opt.* **7**, S347 (2005).
 - [4] E. Knill, D. Leibfried, R. Reichle, J. Britton, R. B. Blakestad, J. D. Jost, C. Langer, R. Ozeri, S. Seidelin, and D. J. Wineland, Randomized benchmarking of quantum gates, *Phys. Rev. A* **77**, 012307 (2008).
 - [5] E. Magesan, J. M. Gambetta, and J. Emerson, Scalable and robust randomized benchmarking of quantum processes, *Phys. Rev. Lett.* **106**, 180504 (2011).
 - [6] E. Magesan, J. M. Gambetta, and J. Emerson, Characterizing quantum gates via randomized benchmarking, *Phys. Rev. A* **85**, 042311 (2012).
 - [7] I. L. Chuang and M. A. Nielsen, Prescription for experimental determination of the dynamics of a quantum black box, *J. Mod. Opt.* **44**, 2455 (1997).
 - [8] J. M. Chow, J. M. Gambetta, L. Tornberg, J. Koch, L. S. Bishop, A. A. Houck, B. R. Johnson, L. Frunzio, S. M. Girvin, and R. J. Schoelkopf, Randomized benchmarking and process tomography for gate errors in a solid-state qubit, *Phys. Rev. Lett.* **102**, 090502 (2009).
 - [9] A. D. Córcoles, J. M. Gambetta, J. M. Chow, J. A. Smolin, M. Ware, J. Strand, B. L. T. Plourde, and M. Steffen, Process verification of two-qubit quantum gates by randomized benchmarking, *Phys. Rev. A* **87**, 030301(R) (2013).
 - [10] S. Olmschenk, R. Chicireanu, K. D. Nelson, and J. V. Porto, Randomized benchmarking of atomic qubits in an optical lattice, *New J. Phys.* **12**, 113007 (2010).
 - [11] T. Xia, M. Lichtman, K. Maller, A. W. Carr, M. J. Piotrowicz, L. Isenhower, and M. Saffman, Randomized benchmarking of single-qubit gates in a 2D array of neutral-atom qubits, *Phys. Rev. Lett.* **114**, 100503 (2015).
 - [12] A. Silva and E. Greplova, Hands-on introduction to randomized benchmarking, *SciPost Phys. Lect. Notes* **97** (2025).
 - [13] J. M. Gambetta, A. D. Córcoles, S. T. Merkel, B. R. Johnson, J. A. Smolin, J. M. Chow, C. A. Ryan, C. Rigetti, S. Poletto, T. A. Ohki, M. B. Ketchen, and M. Steffen, Characterization

- of addressability by simultaneous randomized benchmarking, *Phys. Rev. Lett.* **109**, 240504 (2012).
- [14] E. Magesan, J. M. Gambetta, B. R. Johnson, C. A. Ryan, J. M. Chow, S. T. Merkel, M. P. da Silva, G. A. Keefe, M. B. Rothwell, T. A. Ohki, M. B. Ketchen, and M. Steffen, Efficient measurement of quantum gate error by interleaved randomized benchmarking, *Phys. Rev. Lett.* **109**, 080505 (2012).
- [15] D. S. França and A. K. Hashagen, Approximate randomized benchmarking for finite groups, *J. Phys. A: Math. Theor.* **51**, 395302 (2018).
- [16] J. Helsen, X. Xue, L. M. K. Vandersypen, and S. Wehner, A new class of efficient randomized benchmarking protocols, *npj Quantum Inf.* **5**, 71 (2019).
- [17] J. J. Wallman, Randomized benchmarking with gate-dependent noise, *Quantum* **2**, 47 (2018).
- [18] T. J. Proctor, A. Carignan-Dugas, K. Rudinger, E. Nielsen, R. Blume-Kohout, and K. Young, Direct randomized benchmarking for multiqubit devices, *Phys. Rev. Lett.* **123**, 030503 (2019).
- [19] J. Combes, C. Granade, C. Ferrie, and S. T. Flammia, Logical randomized benchmarking, [arXiv:1702.03688](https://arxiv.org/abs/1702.03688).
- [20] M. McEwen *et al.*, Removing leakage-induced correlated errors in superconducting quantum error correction, *Nat. Commun.* **12**, 1761 (2021).
- [21] J. Morris, F. A. Pollock, and K. Modi, Quantifying non-Markovian memory in a superconducting quantum computer, *Open Syst. Inf. Dyn.* **29**, 2250007 (2022).
- [22] P. Figueroa-Romero, K. Modi, R. J. Harris, T. M. Stace, and M.-H. Hsieh, Randomized benchmarking for non-Markovian noise, *PRX Quantum* **2**, 040351 (2021).
- [23] P. Figueroa-Romero, K. Modi, and M.-H. Hsieh, Towards a general framework of randomized benchmarking incorporating non-Markovian noise, *Quantum* **6**, 868 (2022).
- [24] C. Giarmatzi and F. Costa, Witnessing quantum memory in non-Markovian processes, *Quantum* **5**, 440 (2021).
- [25] P. Taranto, M. T. Quintino, M. Murao, and S. Milz, Characterising the hierarchy of multi-time quantum processes with classical memory, *Quantum* **8**, 1328 (2024).
- [26] H. Ball, T. M. Stace, S. T. Flammia, and M. J. Biercuk, Effect of noise correlations on randomized benchmarking, *Phys. Rev. A* **93**, 022303 (2016).
- [27] M. A. Fogarty, M. Veldhorst, R. Harper, C. H. Yang, S. D. Bartlett, S. T. Flammia, and A. S. Dzurak, Nonexponential fidelity decay in randomized benchmarking with low-frequency noise, *Phys. Rev. A* **92**, 022326 (2015).
- [28] J. Qi and H. K. Ng, Randomized benchmarking in the presence of time-correlated dephasing noise, *Phys. Rev. A* **103**, 022607 (2021).
- [29] M. Veldhorst, J. C. C. Hwang, C. H. Yang, A. W. Leenstra, B. de Ronde, J. P. Dehollain, J. T. Muhonen, F. E. Hudson, K. M. Itoh, A. Morello, and A. S. Dzurak, An addressable quantum dot qubit with fault-tolerant control-fidelity, *Nat. Nanotechnol.* **9**, 981 (2014).
- [30] K. Goswami, A. K. Roy, V. Srivastava, B. Perez, C. Giarmatzi, A. Gilchrist, and F. Costa, Hamiltonian characterization of multi-time processes with classical memory, *New J. Phys.* **27**, 114515 (2025).
- [31] P. Mundada, G. Zhang, T. Hazard, and A. Houck, Suppression of qubit crosstalk in a tunable coupling superconducting circuit, *Phys. Rev. Appl.* **12**, 054023 (2019).
- [32] Z. Ni, S. Li, L. Zhang, J. Chu, J. Niu, T. Yan, X. Deng, L. Hu, J. Li, Y. Zhong, S. Liu, F. Yan, Y. Xu, and D. Yu, Scalable method for eliminating residual ZZ interaction between superconducting qubits, *Phys. Rev. Lett.* **129**, 040502 (2022).
- [33] F. A. Pollock, C. Rodríguez-Rosario, T. Frauenheim, M. Paternostro, and K. Modi, Non-Markovian quantum processes: Complete framework and efficient characterization, *Phys. Rev. A* **97**, 012127 (2018).
- [34] Y. R. Sanders, J. J. Wallman, and B. C. Sanders, Bounding quantum gate error rate based on reported average fidelity, *New J. Phys.* **18**, 012002 (2015).
- [35] R. Kueng, D. M. Long, A. C. Doherty, and S. T. Flammia, Comparing experiments to the fault-tolerance threshold, *Phys. Rev. Lett.* **117**, 170502 (2016).
- [36] C. A. Ryan, M. Laforest, and R. Laflamme, Randomized benchmarking of single- and multi-qubit control in liquid-state NMR quantum information processing, *New J. Phys.* **11**, 013034 (2009).
- [37] J. M. Epstein, A. W. Cross, E. Magesan, and J. M. Gambetta, Investigating the limits of randomized benchmarking protocols, *Phys. Rev. A* **89**, 062321 (2014).
- [38] J. Helsen, I. Roth, E. Onorati, A. H. Werner, and J. Eisert, General framework for randomized benchmarking, *PRX Quantum* **3**, 020357 (2022).
- [39] J. J. Wallman and S. T. Flammia, Randomized benchmarking with confidence, *New J. Phys.* **16**, 103032 (2014).
- [40] C. Granade, C. Ferrie, and D. G. Cory, Accelerated randomized benchmarking, *New J. Phys.* **17**, 013042 (2015).
- [41] F. Costa and S. Shrapnel, Quantum causal modelling, *New J. Phys.* **18**, 063032 (2016).
- [42] O. Oreshkov, F. Costa, and C. Brukner, Quantum correlations with no causal order, *Nat. Commun.* **3**, 1092 (2012).
- [43] F. A. Pollock, C. Rodríguez-Rosario, T. Frauenheim, M. Paternostro, and K. Modi, Operational Markov condition for quantum processes, *Phys. Rev. Lett.* **120**, 040405 (2018).
- [44] M.-D. Choi, Completely positive linear maps on complex matrices, *Linear Algebra Appl.* **10**, 285 (1975).
- [45] A. Jamiolkowski, Linear transformations which preserve trace and positive semidefiniteness of operators, *Rep. Math. Phys.* **3**, 275 (1972).
- [46] M. Araújo, C. Branciard, F. Costa, A. Feix, C. Giarmatzi, and C. Brukner, Witnessing causal nonseparability, *New J. Phys.* **17**, 102001 (2015).
- [47] A. K. Roy, V. Srivastava, S. Mahanti, C. Giarmatzi, and A. Gilchrist, Semi-device-independent certification of quantum non-Markovianity using sequential random access codes, *Phys. Rev. A* **110**, 012608 (2024).
- [48] G. A. L. White, F. A. Pollock, L. C. L. Hollenberg, C. D. Hill, and K. Modi, What can unitary sequences tell us about multi-time physics? *Quantum* **9**, 1695 (2025).
- [49] P. Taranto, S. Milz, M. Murao, M. T. Quintino, and K. Modi, Higher-order quantum operations, [arXiv:2503.09693](https://arxiv.org/abs/2503.09693).
- [50] R. Roy, A. Paulraj, and T. Kailath, *MILCOM 1986-IEEE Military Communications Conference: Communications-Computers: Teamed for the 90's* (IEEE, Monterey, CA, USA, 1986), Vol. 3.
- [51] W. Li and W. Liao, Stable super-resolution limit and smallest singular value of restricted Fourier matrices, [arXiv:1709.03146](https://arxiv.org/abs/1709.03146).

- [52] R. Schmidt, Multiple emitter location and signal parameter estimation, *IEEE Trans. Antennas Propag.* **34**, 276 (1986).
- [53] G. A. L. White, F. A. Pollock, L. C. L. Hollenberg, K. Modi, and C. D. Hill, Non-Markovian quantum process tomography, *PRX Quantum* **3**, 020344 (2022).
- [54] A. K. Roy, V. Srivastava, C. Giarmatzi, and A. Gilchrist, Practical tomography of multi-time processes, [arXiv:2604.01482](#).
- [55] A. Ceasura, P. Iyer, J. J. Wallman, and H. Pashayan, Non-exponential behaviour in logical randomized benchmarking, [arXiv:2212.05488](#).
- [56] S. Gandhari and M. J. Gullans, Quantum non-Markovian noise in randomized benchmarking of spin-boson models, [arXiv:2502.14702](#).
- [57] S. Mavadia, C. L. Edmunds, C. Hempel, H. Ball, F. Roy, T. M. Stace, and M. J. Biercuk, Experimental quantum verification in the presence of temporally correlated noise, *npj Quantum Inf.* **4**, 7 (2018).
- [58] M. F. Sacchi, Minimum error discrimination of Pauli channels, *J. Opt. B: Quantum Semiclassical Opt.* **7**, S333 (2005).
- [59] J. Watrous, *The Theory of Quantum Information* (Cambridge University Press, 2018).
- [60] M. A. Nielsen, A simple formula for the average gate fidelity of a quantum dynamical operation, *Phys. Lett. A* **303**, 249 (2002).
- [61] C. A. Fuchs and J. van de Graaf, Cryptographic distinguishability measures for quantum mechanical states, [arXiv:quant-ph/9712042](#).
- [62] P. Skrzypczyk and D. Cavalcanti, *Semidefinite Programming in Quantum Information Science* (IOP Publishing, 2023).
- [63] X. Zhang, Z. Wu, G. A. L. White, P. Huo, Q. Zhuang, K. Modi, I. Kassal, and X. Wang, Learning and forecasting open quantum dynamics with correlated noise, *Commun. Phys.* **8**, 29 (2025).
- [64] A. A. Mele, Introduction to Haar measure tools in quantum information: A beginner's tutorial, *Quantum* **8**, 1340 (2024).