

# Quantum Circuits for Exact Unitary $t$ -Designs and Applications to Higher-Order Randomized Benchmarking

Yoshifumi Nakata<sup>1,2,\*</sup>, Da Zhao<sup>3</sup>, Takayuki Okuda,<sup>4</sup> Eiichi Bannai,<sup>5,6</sup> Yasunari Suzuki,<sup>7,2</sup> Shiro Tamiya,<sup>8</sup> Kentaro Heya,<sup>9</sup> Zhiguang Yan,<sup>10</sup> Kun Zuo,<sup>10</sup> Shuhei Tamate,<sup>10</sup> Yutaka Tabuchi<sup>10</sup> and Yasunobu Nakamura<sup>9,10</sup>

<sup>1</sup>Photon Science Center, Graduate School of Engineering, The University of Tokyo, Bunkyo-ku, Tokyo 113–8656, Japan

<sup>2</sup>JST, PRESTO, 4–1–8 Honcho, Kawaguchi, Saitama 332–0012, Japan

<sup>3</sup>School of Mathematical Sciences, Shanghai Jiao Tong University, Shanghai 400240, China

<sup>4</sup>Department of Mathematics, Hiroshima University, 1-3-1, Kagamiyama, Higashihiroshima 739-8562, Japan

<sup>5</sup>Faculty of Mathematics, Kyushu University (emeritus), Fukuoka 819-0385, Japan

<sup>6</sup>Mathematics Division, National Center for Theoretical Sciences, National Taiwan University, Taipei, 10617, Taiwan

<sup>7</sup>NTT Computer and Data Science Laboratories, NTT Corporation, Musashino 180-8585, Japan

<sup>8</sup>Department of Applied Physics, Graduate School of Engineering, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan

<sup>9</sup>Research Center for Advanced Science and Technology (RCAST), The University of Tokyo, Meguro-ku, Tokyo 153-8904, Japan

<sup>10</sup>RIKEN Center for Quantum Computing (RQC), Wako, Saitama 351-0198, Japan



(Received 25 March 2021; accepted 13 August 2021; published 3 September 2021)

A unitary  $t$ -design is a powerful tool in quantum information science and fundamental physics. Despite its usefulness, only approximate implementations were known for general  $t$ . In this paper, we provide quantum circuits that generate exact unitary  $t$ -designs for any  $t$  on an arbitrary number of qubits. Our construction is inductive and is of practical use in small systems. We then introduce a  $t$ th-order generalization of randomized benchmarking ( $t$ -RB) as an application of exact  $2t$ -designs. We particularly study the 2-RB in detail and show that it reveals *self-adjointness* of quantum noise, a metric related to the feasibility of quantum error correction (QEC). We numerically demonstrate that the 2-RB in one- and two-qubit systems is feasible, and experimentally characterize background noise of a superconducting qubit by the 2-RB. It is shown from the experiment that interactions with adjacent qubits induce the noise that may result in an obstacle toward a realization of QEC.

DOI: [10.1103/PRXQuantum.2.030339](https://doi.org/10.1103/PRXQuantum.2.030339)

## I. INTRODUCTION

Randomness in quantum systems has been driving recent progress of quantum information science [1–34] as well as fundamental physics [35–46]. Theoretically, quantum randomness is often formulated by a unitary drawn uniformly at random, also known as a *Haar* random unitary. However, the Haar randomness is physically unfeasible in large quantum systems. From the viewpoint of

applications, the unitaries that have similar properties of a Haar random unitary are of great importance because they can be used instead of the Haar one. When a random unitary has the same  $t$ th-order statistics as a Haar random unitary on average, it is called a *unitary  $t$ -design*. For instance, when a protocol exploits the  $t$ th power of the measurement probability after applying a Haar random unitary on any state, the protocol also works even if the Haar random unitary is replaced with a unitary  $t$ -design.

A unitary  $t$ -design can be regarded as a quantum generalization of  $t$ -wise independence [1], and have many applications, ranging from communication [2–9], cryptography [10,11], algorithms [12,13], sensing [14–17], to potentially quantum supremacy [18–20]. A unitary  $t$ -design is also related to another important concept in quantum information science, *epsilon-net* [21], implying more

\*nakata@qi.t.u-tokyo.ac.jp

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

applications yet to be discovered. Furthermore, the concept of unitary designs has opened a novel research field over quantum information science, quantum thermodynamics, strongly correlated physics, and quantum gravity [35–44]. Experimentally, unitary designs and related methods have been exploited for benchmarking noisy quantum devices [18,22–34], realizing quantum supremacy [19], demonstrating quantum chaos and quantum holography [45,46]. It is also worthwhile to mention that unitary designs have been studied in combinatorial mathematics [47–52]. Hence, developing the theory of unitary designs is of substantial interest in a wide range of science, both theoretically and experimentally.

An important question about unitary  $t$ -designs is how to implement them by quantum circuits. Many implementations of unitary 2-designs, both approximate and exact, have been proposed [53–62]. In contrast, only *approximate* implementations of unitary  $t$ -designs for general  $t$  were known [63–67]. Explicit constructions of *exact* unitary designs were left open except special cases [49,52,68]. Approximations typically suffice in applications, but exact designs are more preferable in certain protocols especially when they are used multiple times in a single run of the protocol. If this is the case, the error from each approximate implementation accumulates and eventually spoils the protocol.

One of such protocols is a randomized benchmarking (RB) protocol [22–24], a standard method for experimentally estimating quantum noise, where unitary 2-designs are used multiple times. Although the RB is experimentally friendly and is widely used in various experimental systems, it reveals only the average gate fidelity. To obtain more information about the noise, a number of variants have been proposed and experimentally implemented (see, e.g., Ref. [69] and the references therein), which are all based on 2-designs. It is highly expected that, by using higher-order designs, much more information about the noise in quantum systems can be extracted. To this end, explicit constructions of exact unitary  $t$ -designs are important.

Constructing exact designs is, however, by far nontrivial. The difficulty is illustrated by a spherical  $t$ -design, a random real unit vector analogous to a unitary  $t$ -design. The existence of exact spherical  $t$ -designs was proven in a nonconstructive manner more than three decades ago [70]. Since then, more concise proofs and explicit constructions have been under intense investigation in combinatorial mathematics (see, e.g., Refs. [71–75] and the references therein). In particular, it was only recently that constructions in general cases [76] and explicit constructions, in the sense that all the algorithms are given in a computable form [77], were proposed. Finding explicit constructions of exact unitary designs, because they are more complicated than spherical designs, is a rather nontrivial problem.

In this paper, we provide an explicit quantum circuit that generates an exact unitary  $t$ -design for any  $t$  on the arbitrary number  $N$  of qubits. More specifically, we show that an exact unitary  $t$ -design on  $d$ -dimensional Hilbert space, that is, a *qudit*, can be generated from those on smaller spaces, which is obtained based on the recent mathematical results by some of the present authors [77]. Using this result, we provide an inductive construction of quantum circuits for exact unitary  $t$ -designs on  $N$  qubits: we first construct a unitary  $t$ -design on a single qubit and then extend it to  $N$  qubits. Unfortunately, the circuit fails to be efficient, but is still of practical use when the system is small.

As an application of exact unitary designs, we introduce the  $t$ th-order RB, or the  $t$ -RB for short, that harnesses the power of exact unitary 2t-designs. The standard RB corresponds to the 1-RB. The  $t$ -RB enables us to experimentally characterize the higher-order properties of quantum noises in the manner free from state-preparation and measurement (SPAM) errors. We investigate in particular the 2-RB in detail and show that it reveals *self-adjointness* of the noise in the system. The self-adjointness is a metric of the noise related to the feasibility of quantum error correction (QEC): small self-adjointness implies that the noise cannot be approximated by any stochastic Pauli noise. The noise on the system being stochastic Pauli is desirable both in theory and in practice. Stochastic Pauli noises are the commonly used noise models in theoretical studies of QEC, because they are easy to handle numerically, and the properties of QEC, such as error thresholds and logical error rates, for stochastic Pauli noises are well-understood. In addition, there is a practical advantage if the noise on the system is stochastic Pauli because it can be corrected simply by applying Pauli operators, making the error-correcting scheme easier in general.

After numerically demonstrating the feasibility of the 2-RB, we perform the 2-RB in a superconducting system and estimate the self-adjointness of background noise, showing that the 2-RB experiments are feasible. From the experiment, we find that the interactions with adjacent qubits especially decrease the self-adjointness, which may lead to degradation of the performance of QEC with standard decoders. Hence, either improving the system or extending the noise model in theoretical studies of QEC, or both, is important for further experimental developments of quantum information technology.

This paper is organized as follows. In Sec. II, we provide a general introduction of unitary  $t$ -designs. Our main results are summarized in Sec. III for the quantum-circuit construction of exact unitary  $t$ -designs, and in Sec. IV for the  $t$ -RB protocols. A summary of the experiment of the 2-RB is provided in Sec. V. After we explain the structure of the remaining paper in Sec. VI, we provide a proof of the explicit construction in Sec. VII and the theory of the  $t$ -RB in Sec. VIII. The details of the experiment are provided in

Sec. IX. We conclude our paper with summary and discussions in Sec. X. Technical statements are provided in the Appendices.

## II. UNITARY $t$ -DESIGNS

Let  $\mathbf{U}(d)$  be the unitary group of degree  $d < \infty$ . The Haar measure  $\mathbf{H}$  on  $\mathbf{U}(d)$  is the unique unitarily invariant measure on the unitary group, that is, it satisfies

$$\begin{aligned} \forall \mathcal{W} \subset \mathbf{U}(d), \quad \forall V \in \mathbf{U}(d), \\ \mathbf{H}(V\mathcal{W}) = \mathbf{H}(\mathcal{W}V) = \mathbf{H}(\mathcal{W}). \end{aligned} \quad (1)$$

When it is needed to clarify the degree of the unitary group, we denote the Haar measure by  $\mathbf{H}(d)$ .

A unitary  $t$ -design  $\mathbf{U}_t(d)$  is defined by a finite set of unitaries that mimics the  $t$ th-order statistical moment of the Haar measure  $\mathbf{H}$ . Amongst several equivalent definitions [78], we adopt the following definition here.

**Definition 1 (Unitary  $t$ -design):** For  $t \in \mathbb{Z}^+$ , a finite set  $\mathbf{U}_t(d) \subset \mathbf{U}(d)$  of unitaries is a unitary  $t$ -design if

$$\mathbb{E}_{U \sim \mathbf{U}_t(d)}[U^{\otimes t} \otimes U^{\dagger \otimes t}] = \mathbb{E}_{U \sim \mathbf{H}(d)}[U^{\otimes t} \otimes U^{\dagger \otimes t}], \quad (2)$$

where  $\mathbb{E}_{U \sim \mathbf{U}_t(d)}$  is a uniform average over  $\mathbf{U}_t(d)$ , and  $\mathbb{E}_{U \sim \mathbf{H}(d)}$  is the average over the Haar measure.

From an operational viewpoint, this definition implies that a unitary  $t$ -design cannot be distinguished from a Haar random unitary on average even when  $t$  copies of the unitary are given. To classify this, let us define a quantum operation, that is, a completely positive and trace-preserving (CPTP) map,  $\mathcal{G}_\mu^{(t)}$  by

$$\mathcal{G}_\mu^{(t)}(\varrho) := \mathbb{E}_{U \sim \mu}[U^{\otimes t} \varrho U^{\otimes t \dagger}], \quad (3)$$

for any quantum state  $\varrho$  on  $t$  qudits, where  $\mu$  is either the Haar measure  $\mathbf{H}(d)$  on a qudit or a uniform measure over a unitary  $t$ -design  $\mathbf{U}_t(d)$ . Then, we can show that Definition 1 is equivalent to that (see, e.g., [78])

$$\mathcal{G}_{\mathbf{H}}^{(t)} = \mathcal{G}_{\mathbf{U}_t}^{(t)}. \quad (4)$$

This implies that, in *any* experiments that use  $t$  copies of a random unitary, no difference will be observed on average when a  $t$ -design is used instead of the Haar one.

For instance, let us consider the probability distribution  $\{p_i(U) := \text{Tr}[P_i U \rho U^\dagger]\}$  when a one-qudit state is measured by a given positive operator-valued measure (POVM)  $\{P_i\}_i$  after the application of a unitary  $U$ . By setting the  $t$ -qudit state  $\varrho$  in Eq. (3) to  $\rho^{\otimes t}$  and using Eq. (4), it follows that, for any  $s = 1, \dots, t$ ,

$$\mathbb{E}_{U \sim \mathbf{H}} \left[ \prod_{r=1}^s p_{i_r}(U) \right] = \mathbb{E}_{U \sim \mathbf{U}_t} \left[ \prod_{r=1}^s p_{i_r}(U) \right]. \quad (5)$$

Thus, the distribution of the measurement outcomes for a Haar random unitary and that for an unitary  $t$ -design

exactly coincide up to the  $t$ th order on average. Note that this is merely an example, and Eq. (4) implies much more: a Haar random unitary cannot be differentiated from a unitary  $t$ -design even by more complicated experiments over  $t$  qudits.

The existence of an exact unitary  $t$ -design on  $\mathbf{U}(d)$  for any  $t$  and  $d$  follows from the Carathéodotý's theorem and the fact that the dimension of the space, on which  $U^{\otimes t} \otimes U^{\dagger \otimes t}$  is defined, is finite. Note, however, that the proof indicates only the existence of an exact unitary  $t$ -design. How to explicitly construct an exact unitary  $t$ -design has been a highly nontrivial problem.

In our construction, it is convenient to introduce a *strong* unitary  $t$ -design.

**Definition 2 (Strong unitary  $t$ -design):** For  $t \in \mathbb{Z}^+$ , a finite set  $\mathbf{U}_{\leq t}(d)$  of unitaries on  $\mathbf{U}(d)$  is called a strong unitary  $t$ -design if

$$\mathbb{E}_{U \sim \mathbf{U}_{\leq t}(d)}[U^{\otimes r} \otimes U^{\dagger \otimes s}] = \mathbb{E}_{U \sim \mathbf{H}}[U^{\otimes r} \otimes U^{\dagger \otimes s}], \quad (6)$$

for all  $r, 0 \leq r \leq t$ , and all  $s, 0 \leq s \leq t$ .

Clearly, a strong unitary  $t$ -design is a unitary  $t$ -design. Unlike standard unitary designs, strong unitary designs do not have a clear operational interpretation in quantum information processing, but we use it in the intermediate step of our construction.

## III. MAIN RESULT 1: QUANTUM CIRCUITS FOR EXACT UNITARY DESIGNS

In this section, we provide explicit constructions of strong unitary  $t$ -designs for any  $t$ . In particular, a quantum circuit for a strong unitary  $t$ -design on  $N$  qubits is provided. We start with preliminaries in Sec. III A, and provide the construction in Sec. III B. We then comment on the circuit complexity of the construction in Sec. III C.

### A. Preliminaries

Unitary designs have been studied in terms of representation theory [49,51] because the operator  $U^{\otimes t} \otimes U^{\dagger \otimes t}$  in the definition can be regarded as a representation of the unitary group. Our construction is based on representation theory, where irreducible decomposition of the operator plays an important role. A brief introduction of irreducible representations (irreps) is provided in Sec. VII A. Here, we mention a couple of well-known facts that are necessary to state our main result.

Any irrep of the unitary group can be indexed by a non-increasing integer sequence  $\lambda := (\lambda_1, \lambda_2, \dots, \lambda_d)$  of length  $d$ , that is,  $\lambda_i \in \mathbb{Z}$  for  $i = 1, \dots, d$ , and  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_d$  [79]. In particular, *spherical representations* of  $\mathbf{U}(d)$  with respect to  $\mathbf{K} := \mathbf{U}(d_1) \times \mathbf{U}(d - d_1)$  are of great importance in the construction. Let  $\Lambda_{\text{sph}}(d_1, d, t)$  be a set of all

nonincreasing integer sequences  $\lambda$  in the form of

$$\lambda = (\lambda_1, \dots, \lambda_{d_1}, 0, \dots, 0, -\lambda_{d_1}, \dots, -\lambda_1), \quad (7)$$

where  $d_1 \leq d/2$  and  $t \geq \lambda_1 \geq \dots \geq \lambda_{d_1} \geq 0$ . The spherical representation is the irrep indexed by  $\lambda \in \Lambda_{\text{sph}}(d_1, d, t)$  [80]. For a spherical representation  $\lambda \in \Lambda_{\text{sph}}(d_1, d, t)$ , a zonal spherical function  $Z_\lambda^{(d_1)}(z_1, \dots, z_{d_1})$  is defined by the unique bi-K-invariant function [50,77,81]. The zonal spherical functions are a certain type of symmetric polynomials, and can be explicitly written down (see, e.g., Appendix A of Ref. [77]).

### B. Inductive constructions

Our main technical result is to construct a strong unitary  $t$ -design on  $\mathbf{U}(d)$  from those on  $\mathbf{U}(d_1)$  and on  $\mathbf{U}(d - d_1)$ .

**Theorem 1:** *Let  $d_1$  be a positive integer such that  $d_1 \leq d/2$ . Define a set of unitaries  $\mathbf{W}_{d_1 \oplus d - d_1}$  in  $\mathbf{U}(d)$  by*

$$\mathbf{W}_{d_1 \oplus d - d_1} := \{U \oplus V | U \in \mathbf{U}_{\leq t}(d_1), V \in \mathbf{U}_{\leq t}(d - d_1)\}, \quad (8)$$

where  $\mathbf{U}_{\leq t}(d_1)$  and  $\mathbf{U}_{\leq t}(d - d_1)$  are strong unitary  $t$ -designs on  $\mathbf{U}(d_1)$  and  $\mathbf{U}(d - d_1)$ , respectively. Let  $\theta_\lambda := (\theta_\lambda^{(0)}, \dots, \theta_\lambda^{(d_1-1)})$  ( $\theta_\lambda^{(i)} \in [0, \pi/2]$ ) be such that

$$Z_\lambda^{(d_1)}(\cos^2 \theta_\lambda^{(0)}, \dots, \cos^2 \theta_\lambda^{(d_1-1)}) = 0, \quad (9)$$

where  $Z_\lambda^{(d_1)}$  is the zonal spherical function for  $\lambda \in \Lambda_{\text{sph}}(d_1, d, t)$ . Let  $R_\lambda$  be a unitary defined by

$$R_\lambda = \begin{pmatrix} C(\theta_\lambda) & iS(\theta_\lambda) & 0 \\ iS(\theta_\lambda) & C(\theta_\lambda) & 0 \\ 0 & 0 & I_{d-2d_1} \end{pmatrix}, \quad (10)$$

where  $C(\theta_\lambda) = \text{diag}(\cos \theta_\lambda^{(0)}, \dots, \cos \theta_\lambda^{(d_1-1)})$  and  $S(\theta_\lambda) = \text{diag}(\sin \theta_\lambda^{(0)}, \dots, \sin \theta_\lambda^{(d_1-1)})$ , and  $I_{d-2d_1}$  is the identity matrix of size  $d - 2d_1$ . Then,

$$\mathbf{W}_d := \mathbf{W}_{d_1 \oplus d - d_1} \prod_{\lambda \in \Lambda_{\text{sph}}(d_1, d, t)} (R_\lambda \mathbf{W}_{d_1 \oplus d - d_1}) \quad (11)$$

is a strong unitary  $t$ -design on  $\mathbf{U}(d)$ .

Theorem 1 follows from a more general result [77] shown by some of the authors, which works not only for the unitary group but also for a broader class of compact groups. For the sake of completeness, we provide a direct proof of Theorem 1 in Sec. VII.

We then claim that

$$\mathbf{W}_1 = \{1, \omega, \omega^2, \dots, \omega^t\}, \quad (12)$$

where  $\omega = \exp[2\pi/(t+1)]$  is the  $(t+1)$ th root of unity, is a strong unitary  $t$ -design on  $\mathbf{U}(1)$  for any  $t$ . This is easily checked by direct calculations:

$$\mathbb{E}_{U \sim \mathbf{W}_1}[U^{\otimes r} \otimes U^{\dagger \otimes s}] = \frac{1}{t+1} \sum_{z \in \mathbf{W}_1} z^r \bar{z}^s = \delta_{rs}, \quad (13)$$

$$\mathbb{E}_{U \sim \mathbf{H}(1)}[U^{\otimes r} \otimes U^{\dagger \otimes s}] = \int_{\mathbf{U}(1)} z^r \bar{z}^s dz = \delta_{rs}, \quad (14)$$

where  $\delta_{rs}$  is the Kronecker delta. Hence, we have  $\mathbb{E}_{U \sim \mathbf{W}_1}[U^{\otimes r} \otimes U^{\dagger \otimes s}] = \mathbb{E}_{U \sim \mathbf{H}(1)}[U^{\otimes r} \otimes U^{\dagger \otimes s}]$  for all  $s, r$ ,  $0 \leq s, r \leq t$ , implying that  $\mathbf{W}_1$  is a strong unitary  $t$ -design.

From Theorem 1 and  $\mathbf{W}_1$ , a strong unitary  $t$ -design on a qudit can be constructed inductively.

**Corollary 1:** *For  $d \geq 1$ , let  $\mathbf{W}_{1 \oplus d - 1}$  be a set of unitaries given by*

$$\mathbf{W}_{1 \oplus d - 1} = \{z \oplus V | z \in \mathbf{W}_1, V \in \mathbf{U}_{\leq t}(d - 1)\}, \quad (15)$$

where  $\mathbf{W}_1 = \{1, \omega, \dots, \omega^t\}$  with  $\omega$  being the  $(t+1)$ th root of unity, and  $\theta_\lambda \in [0, \pi/2]$  be such that

$$Z_\lambda^{(1)}(\cos^2 \theta_\lambda) = 0. \quad (16)$$

Using a unitary  $R_\lambda = e^{i\theta_\lambda X} \oplus I_{d-2}$ , where  $X$  is the Pauli- $X$  operator, we obtain that

$$\mathbf{W}_{1 \oplus d - 1} \prod_{\lambda \in \Lambda_{\text{sph}}(1, d, t)} (R_\lambda \mathbf{W}_{1 \oplus d - 1}) \quad (17)$$

is a strong unitary  $t$ -design on a qudit.

In this construction, it is important to obtain zeros for the zonal spherical functions  $Z_\lambda^{(1)}$ . This is computationally feasible because they are polynomials of one variable and are explicitly given (see Appendix A of Ref. [77]). Furthermore,  $\Lambda_{\text{sph}}(1, d, t)$  contains only  $t$  elements. Hence, we need to solve  $t$  polynomials with one variable, which is tractable as long as  $t$  is not too large.

We now consider a strong unitary  $t$ -design on  $N$  qubits. Again using Theorem 1, we obtain the quantum circuit on  $(N+1)$  qubits based on that on  $N$  qubits. See also Fig. 1.

**Corollary 2:** *Let  $\mathbf{Q}_N$  be a strong unitary  $t$ -design on  $N$  qubits, and  $\text{Ctrl-Q}_N$  be a set of controlled-unitaries on*



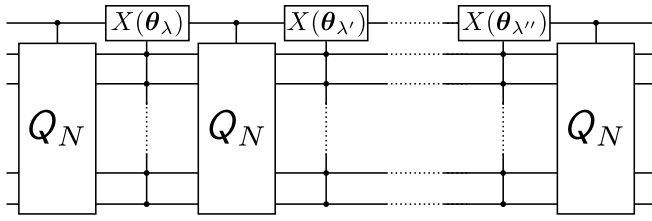


FIG. 1. The quantum circuit that generates an exact unitary  $t$ -design on  $N + 1$  qubits from those on  $N$  qubits. The unitary  $Q_N$  is a quantum circuit for an exact unitary  $t$ -design on  $N$  qubits. The gate  $X(\theta_\lambda)$  is the single-qubit  $X$ -rotation controlled by the other  $N$  qubits, which corresponds to  $R_X(\theta_\lambda)$  in the main text. Note that this gate can be decomposed into a sequence of two-qubit gates using a classical oracle that provides  $\theta_\lambda^{(j)}$  from  $\mathbf{j}$ . The rotation angles  $\theta_\lambda$  are obtained by solving  $Z_\lambda^{(D)} = 0$ , where  $Z_\lambda^{(D)}$  is the zonal spherical function for the spherical representation  $\lambda \in \Lambda_{\text{sph}}(D, 2D, t)$  with  $D = 2^N$ . The number of the controlled- $X$  rotations is  $|\Lambda_{\text{sph}}(D, 2D, t)|$ . By the inductive use of this quantum circuit in terms of  $N$ , we can decompose the circuit to that consisting only of two-qubit gates.

$N + 1$  qubits, defined by

$$\text{Ctrl-}Q_N := \{|0\rangle\langle 0| \otimes U_0 + |1\rangle\langle 1| \otimes U_1 : U_0, U_1 \in Q_N\}. \quad (18)$$

For  $\lambda \in \Lambda_{\text{sph}}(D, 2D, t)$ , where  $D = 2^N$ , let  $\theta_\lambda := (\theta_\lambda^{(0)}, \dots, \theta_\lambda^{(D-1)})$  be such that

$$Z_\lambda^{(D)}(\cos^2 \theta_\lambda^{(0)}, \dots, \cos^2 \theta_\lambda^{(D-1)}) = 0. \quad (19)$$

Representing  $\{0, \dots, D-1\}$  in binary form such as  $\{\mathbf{j}\}_{\mathbf{j} \in \{0,1\}^N}$ , we write  $\theta_\lambda^{(j)}$  as  $\theta_\lambda^{(\mathbf{j})}$ . Let  $R_X(\theta_\lambda)$  be a single-qubit  $X$ -rotation controlled by  $N$  qubits, defined by

$$R_X(\theta_\lambda) = \sum_{\mathbf{j} \in \{0,1\}^N} e^{i\theta_\lambda^{(\mathbf{j})} X} \otimes |\mathbf{j}\rangle\langle \mathbf{j}|. \quad (20)$$

Then,

$$\text{Ctrl-}Q_N \prod_{\lambda \in \Lambda_{\text{sph}}(D, 2D, t)} (R_X(\theta_\lambda) \text{Ctrl-}Q_N) \quad (21)$$

is a strong unitary  $t$ -design on  $N + 1$  qubits.

Corollary 2 implies that a quantum circuit for an exact unitary  $t$ -design can be inductively constructed from a strong unitary  $t$ -design on one qubit, that is, that on  $U(2)$ . Furthermore, a strong unitary  $t$ -design on  $U(2)$  can be constructed using Corollary 1. Thus, combining Corollaries 1 and 2, we obtain a quantum circuit for an exact unitary  $t$ -design for any  $t$  and on an arbitrary number of qubits.

Note that the circuit, constructed in this way, can be explicitly decomposed into two-qubit gates. The controlled

unitary  $\text{Ctrl-}Q_N$  part contains up to three-qubit gates, if the circuit  $Q_N$  on  $N$  qubit is already decomposed into two-qubit gates. The three-qubit gates can be easily rewritten as a series of two-qubit gates. In addition, the  $X$ -rotation controlled by  $N$  qubits,  $R_X(\theta_\lambda)$ , can be decomposed into a sequence of two-qubit gates of polynomial length using a sufficient number of ancillary qubits, which is based on a classical oracle that computes the angle  $\theta_\lambda^{(j)}$  from  $\mathbf{j}$  (see Appendix A).

In special cases, we can find a much more concise construction based on a similar technique.

**Proposition 1:** *Let  $C(4)$  be the Clifford group on 2 qubits. There exists a fixed two-qubit unitary  $U_c$ , such that  $C(4)U_cC(4)$  is an exact unitary 4-design on 2 qubits.*

Analytically, we can prove that there exist unitaries  $V_1$  and  $V_2$  such that  $C(4)V_1C(4)V_2C(4)$  is an exact unitary 4-design on 2 qubits [77]. In addition, an algorithm for computing the unitaries  $V_1$  and  $V_2$  is given. It, however, turns out from numerics that it is not necessary to apply two extra unitaries if we choose a proper unitary  $U_c$ , which leads to Proposition 1. An explicit form of the unitary  $U_c$  is numerically obtained and is provided in Appendix B. Note that the existence of  $U_c$  is confirmed numerically, so the statement holds up to numerical precision.

This construction is only for a 4-design on 2 qubits, but the number of unitaries in the 4-design is much smaller than in Corollary 2. It is an open problem whether a similar construction works for higher-order designs on a larger number of qubits.

### C. Efficiency and comparison with a Haar unitary

To quantitatively evaluate the complexity of the quantum circuit for an exact unitary  $t$ -design obtained in Corollary 2, we provide an order estimate of the number  $G(N)$  of two-qubit gates in the circuit. Assuming  $2^N \gg t$  and using the fact that  $|\Lambda_{\text{sph}}(D, 2D, t)| = O(e^{\pi\sqrt{2t/3}})$  owing to the Hardy and Ramanujan formula for the asymptotics of the number of partitions, we obtain

$$G(N) \approx \exp\left[\pi\sqrt{\frac{2t}{3}}(N-1)\right], \quad (22)$$

to the leading order of  $N$ . Hence, it is necessary to use exponentially many two-qubit gates as the number of qubits increases. This inefficiency of the quantum circuit may be intrinsic because the construction is inductive.

There is another source of inefficiency. In Corollary 2, it is necessary to find zeros of zonal spherical functions (see Eq. (19)) for all  $\lambda \in \Lambda_{\text{sph}}(D, 2D, t)$ . The zonal spherical function is given in terms of the summation of the (normalized) Schur polynomials (see Appendix A in Ref. [77]). It is unlikely that the Schur polynomials have polynomial size algebraic formulas in general [82]. Moreover, the

number of variables for each  $Z_\lambda^{(D)}$  is  $D = 2^N$ . Hence, finding zeros of zonal spherical functions is computationally intractable.

In total, the construction for an exact unitary  $t$ -design on a large number of qubits is inefficient from both quantum-circuit and classical-computation viewpoints. We, however, think that our construction and our proof technique will form a solid basis of searching more efficient constructions of exact, as well as approximate, unitary designs. We also emphasize that, despite its inefficiency, our construction is of practical use on a few-qubit system, as we seek in the following sections.

Before we conclude this section, we comment on advantages of our construction of an exact unitary  $t$ -design over a direct implementation of a Haar random unitary. A naive way of implementing a Haar random unitary by a quantum circuit consists of three steps. First, we sample a Haar random unitary as a matrix by a classical computer. A classical algorithm for this is known [83], but it is trivially inefficient because the size of the matrix is exponentially large. We then classically compute a decomposition of the unitary matrix into a sequence of two-qubit unitaries, providing a classical description of a quantum circuit for the unitary. This step is also inefficient, and the resulting quantum circuit is almost surely composed of the exponentially many number of two-qubit gates. Finally, we implement the circuit in practice.

This quantum circuit for a Haar random unitary is inefficient in terms of the number of qubits and, thus, cannot be of practical use in a large system. Even in a small system, this naive implementation has a crucial difficulty that, every time a unitary is sampled, the protocol described previously outputs a quantum circuit with a rather different sequence of various two-qubit gates. This implies that, in each sampling, one needs to significantly modify the quantum circuit. This is in a sharp contrast to our quantum circuit for a unitary  $t$ -design based on Corollary 2 because it has a *fixed* structure. In each sampling, only what one needs to do is to randomly choose single-qubit gates, or more precisely elements of  $U(1)$  from  $\mathbb{W}_1$  [see Eq. (12)], and to plug them into the quantum circuit with a fixed structure. This will help practical implementations of the circuit in small systems.

It should be also noted that the single-qubit gates in our construction can be sampled from a discrete set, though sampling from a continuous gate set is necessary in the direct implementation of a Haar random unitary. This is another advantage of our construction.

## IV. MAIN RESULT 2: HIGHER-ORDER RANDOMIZED BENCHMARKING

We here introduce a higher-order generalization of the standard RB that uses exact unitary  $2t$ -designs. We call it the  $t$ th-order RB, or simply  $t$ -RB. The standard RB

corresponds to 1-RB. From the higher-order RB, more information about the noise can be extracted. In particular, we show that an alternative characterization of the noise, which we call *self-adjointness*, can be estimated from the 2-RB.

Before we proceed, we emphasize that exact unitary designs, not approximations, are of crucial importance in the RB-type protocols. This is because the protocol uses unitary designs multiple times. Hence, if each unitary has an error due to the approximation, it accumulates in the whole process and results in a large error at the end. As the goal of the RB-type protocol is typically very high, such as benchmarking the fidelity at more than 95%, the error originated from the approximate designs would spoil the protocol. Hence, the use of exact unitary designs is of key importance. This point is more elaborated on in Sec. IV D.

In Sec. IV A, we overview a couple of metrics of the noise, i.e., the average fidelity and unitarity, and introduce the self-adjointness. The importance of the self-adjointness in QEC is argued in Sec. IV B. We then introduce the  $t$ -RB in Sec. IV C. We argue the importance of exact designs in more detail in Sec. IV D. We focus on the 2-RB in Sec. IV E and show that the self-adjointness and the unitarity of the noise can be estimated from the 2-RB at the same time. We briefly comment on the scalability of the  $t$ -RB in Sec. IV F.

### A. Characterizing noises

A noise  $\mathcal{E}$  acting on a  $q$ -qubit system is formulated by a CPTP map. Let  $d$  be defined as  $d := 2^q$ . The average fidelity and the unitarity are defined by

$$F(\mathcal{E}) := \int d\varphi \langle \varphi | \mathcal{E}(|\varphi\rangle\langle\varphi|) | \varphi \rangle, \quad (23)$$

$$u(\mathcal{E}) := \frac{d}{d-1} \int d\varphi \|\mathcal{E}'(|\varphi\rangle\langle\varphi|)\|_2^2, \quad (24)$$

respectively, where  $\mathcal{E}'(\rho) := \mathcal{E}(\rho - I/d)$ , and  $\|A\|_2 = (\text{Tr}[A^\dagger A])^{1/2}$  is the Schatten 2-norm. The average fidelity satisfies  $1/(d+1) \leq F(\mathcal{E}) \leq 1$ , and  $F(\mathcal{E}) = 1$  if and only if the system is noiseless, that is,  $\mathcal{E}$  is the identity channel, whereas the unitarity satisfies  $0 < u(\mathcal{E}) \leq 1$ , and  $u(\mathcal{E}) = 1$  if and only if the noise is coherent, that is,  $\mathcal{E}$  is a unitary channel. The unitarity is an important metric in the context of QEC because coherent noise is known to be hard to correct in general [84–86].

In the RB-type protocols, it is more natural to use a *fidelity parameter*  $f(\mathcal{E})$  rather than the average fidelity itself. It is defined by

$$f(\mathcal{E}) = \frac{dF(\mathcal{E}) - 1}{d - 1}, \quad (25)$$

and satisfies  $-1/(d^2 - 1) \leq f(\mathcal{E}) \leq 1$ .

We next introduce a self-adjointness of the noise. For any linear map  $\mathcal{E}$ , an *adjoint* map  $\mathcal{E}^\dagger$  is defined by  $\text{Tr}[A\mathcal{E}(B)] = \text{Tr}[\mathcal{E}^\dagger(A)B]$ . A noise  $\mathcal{E}$  is called self-adjoint if  $\mathcal{E} = \mathcal{E}^\dagger$ , which is equivalent to that all the Kraus operators of  $\mathcal{E}$  are self-adjoint.

The *self-adjointness*  $H(\mathcal{E})$  of the noise  $\mathcal{E}$  is defined by

$$H(\mathcal{E}) := 1 - \frac{d+1}{2d} \int d\varphi \|\mathcal{E}(|\varphi\rangle\langle\varphi|) - \mathcal{E}^\dagger(|\varphi\rangle\langle\varphi|)\|_2^2. \quad (26)$$

The normalization constant is chosen such that  $0 \leq H(\mathcal{E}) \leq 1$ . Obviously,  $H(\mathcal{E}) = 1$  if and only if  $\mathcal{E}$  is self-adjoint, that is,  $\mathcal{E} = \mathcal{E}^\dagger$ . Note that the self-adjointness has two contributions from the noisy map  $\mathcal{E}$ , one is from the unital part and the other from the nonunital part. The nonunital part of the noise makes the self-adjointness less than one because, if  $\mathcal{E}$  is not unital, then  $\mathcal{E}^\dagger$  is not trace-preserving, which implies that  $\mathcal{E} \neq \mathcal{E}^\dagger$ .

To clearly separate the two contributions, we introduce a *self-adjointness parameter*  $h(\mathcal{E})$ . Using  $\mathcal{E}'(\rho) = \mathcal{E}(\rho - I/d)$ , we defined it by

$$h(\mathcal{E}) := \frac{d}{d-1} \int d\varphi \text{Tr}[\mathcal{E}'(\varphi)\mathcal{E}'^\dagger(\varphi)]. \quad (27)$$

The self-adjointness parameter  $h(\mathcal{E})$  is related to the self-adjointness  $H(\mathcal{E})$  and the unitarity  $u(\mathcal{E})$  by

$$H(\mathcal{E}) = 1 - \frac{d^2-1}{d^2} (u(\mathcal{E}) - h(\mathcal{E})) - \frac{d+1}{2d^2} |\alpha_\mathcal{E}|^2, \quad (28)$$

where  $|\alpha_\mathcal{E}|$  is a measure of the nonunital part of the noise (see Sec. VIII A for the definition). We can clearly observe that  $H(\mathcal{E})$  consists of two factors, the unital part  $h(\mathcal{E})$  and the nonunital part  $|\alpha_\mathcal{E}|$ .

The three metrics of noises, namely, fidelity, unitarity, and self-adjointness, all capture different properties of the noises. The fidelity reveals the first-order property of the noises, whereas the unitarity and the self-adjointness, which are independent to each other, reveal the second-order. In order to improve noisy quantum devices, it is of crucial importance to obtain the information of noise as much as possible. Hence, it is certainly of practical use to introduce the self-adjointness as a metric of noise. In addition, we argue in the next subsection that the self-adjointness has important implications for QEC.

### B. Importance of self-adjointness in QEC

The most important family of self-adjoint noises is stochastic Pauli noises, whose Kraus operators are all proportional to Pauli matrices. In QEC, Pauli noises are the standard yet most important class of noises both in theory and in practice. From a theoretical perspective, Pauli noises are easy to numerically handle. Hence, most

numerical calculations have been carried out by assuming Pauli noises, and it has been confirmed that QEC has preferable features, such as exponential decreases and threshold behaviors of logical error rates, if the noise is Pauli.

The noise being Pauli is also practically preferable in experimental realizations of QEC because it typically simplifies the decoding tasks. This is especially the case for stabilizer codes, such as surface and color codes, whose standard decoders are to estimate what types of Pauli operators should be applied on which physical qubits during recovery operations. For stochastic Pauli noises, if the estimation goes well, the state is fully retrieved with high probability by applying Pauli operators to the suitable physical qubits. In contrast, it is not possible to fully correct non-Pauli noises by applying Pauli operators because they generate undesired coherence between different code spaces. Thus, QEC of non-Pauli noises generally suffers from degradation of logical error rates when the standard decoders are used [86,87] or requires more complicated algorithms for retrieving the performance of QEC. Neither of them is preferable in practice because it induces additional experimental difficulties.

For these reasons, it is desirable to check that the noise on an experimental system is stochastic Pauli. To this end, the self-adjointness provides useful information because, if  $H(\mathcal{E}) \ll 1$ , then the noise is far from self-adjoint and cannot be approximated by Pauli noises. This implies that the practical situation differs from the standard assumption in theoretical studies of QEC and incurs additional difficulties on decoding procedure. Thus, the self-adjointness provides practical information about the feasibility of QEC using Pauli-based decoders.

Note that the difficulty of QEC for non-Pauli noises, captured by the self-adjointness, is highly dependent on the assumptions in QEC schemes. When *any* decoding procedure is available, it would not be so important whether the noise is Pauli or non-Pauli. When this is the case, the unitarity will be a more suitable metric of noise relevant to the feasibility of QEC [84–86]. Note also that non-Pauli noises can be always transformed to a Pauli noise by Pauli twirling. However, Pauli twirling induces additional noise onto the system and, as a result, the performance of QEC will degrade. Thus, it is practically desirable to manufacture the system so that the noise is stochastic Pauli.

We also provide a pedagogical example of noise, where performance of QEC can be directly captured by the self-adjointness but not by fidelity nor unitarity. Consider a  $\theta$ -rotation error around the  $X$ -axis on one qubit, that is,  $\exp[i\theta X/2]$ , where  $X$  is the Pauli- $X$  operator. The average fidelity  $F_\theta$  and the self-adjointness  $H_\theta$  can be obtained as

$$1/3 = F_\pi < F_{\pi/2} = 2/3, \quad (29)$$

$$0 = H_{\pi/2} < H_{\pi} = 1. \quad (30)$$

The unitarity is 1 for any  $\theta$ .

One may expect that the  $\pi/2$ -rotation error is easier to correct than the  $\pi$ -rotation because the former has higher fidelity than the latter. However, this is not the case because  $\pi$ -rotation is simply a perfect bit-flip that can be trivially corrected, whereas the  $\pi/2$ -rotation error is known to be particularly hard to correct [88]. Thus, neither the average fidelity nor the unitarity, which is 1 for both errors, is a good metric of the error correctability. In contrast, the self-adjointness clearly captures whether the error can be corrected, at least in this case, because  $H_{\pi/2}$  and  $H_{\pi}$  are the minimum and the maximum values of the self-adjointness, respectively.

### C. General description of the $t$ th-order RB

We now introduce the  $t$ -RB using an exact unitary  $2t$ -design  $\mathbf{U}_{2t} := \{U_i\}_i$ . As is the case for the standard RB, we assume that the noise is gate- and time-independent, so that the noisy implementation of  $\mathbf{U}_{2t}$  is given by  $\{\mathcal{G}_i := \mathcal{E} \circ \mathcal{U}_i\}_i$ , where  $\mathcal{E}$  is the CPTP map that represents the noise, and we used the notation that  $\mathcal{U}(\rho) := U\rho U^\dagger$ .

Let  $O_{\text{ini}}$  and  $O_{\text{meas}}$  be the initial and measurement operators, respectively, which we assume to be Hermitian. We first apply a sequence of unitaries  $U_i = U_{i_m} \dots U_{i_1}$  onto the initial operator  $O_{\text{ini}}$ . Each  $U_{i_n}$  is chosen uniformly at random from  $\mathbf{U}_{2t}$ , which we denote by  $U_i \sim \mathbf{U}_{2t}^{\times m}$ . We then apply its inverse  $U_{i_{m+1}} := U_i^{-1}$ , and measure  $O_{\text{meas}}$ .

If the system is noiseless,  $\mathcal{E} = \text{id}$ , this protocol results in a trivial expectation value that

$$\text{Tr}[O_{\text{meas}} \mathcal{U}_{i_{m+1}} \circ \mathcal{U}_{i_m} \circ \dots \circ \mathcal{U}_{i_1}(O_{\text{ini}})] = \text{Tr}[O_{\text{meas}} O_{\text{ini}}] \quad (31)$$

owing to the inverse unitary  $U_{i_{m+1}}$ . However, when the system is noisy, the expectation value becomes

$$\langle O_{\text{meas}} \rangle_{O_{\text{ini}}, i} := \text{Tr}[O_{\text{meas}} \mathcal{G}_{i_{m+1}} \circ \mathcal{G}_{i_m} \circ \dots \circ \mathcal{G}_{i_1}(O_{\text{ini}})], \quad (32)$$

which, in general, differs from  $\text{Tr}[O_{\text{meas}} O_{\text{ini}}]$ . The basic idea of the RB-type protocol is to extract some information about the noise  $\mathcal{E}$  from the difference.

In the  $t$ -RB, we especially focus on the average of the  $t$ th power of the expectation value over all choices of the unitary sequence. That is,

$$V^{(t)}(m, \mathcal{E} | O_{\text{ini}}, O_{\text{meas}}) := \mathbb{E}_{U_i \sim \mathbf{U}_{2t}^{\times m}} [(\langle O_{\text{meas}} \rangle_{O_{\text{ini}}, i})^t]. \quad (33)$$

Using the representation-theoretic technique, it can be shown that  $V^{(t)}(m, \mathcal{E} | O_{\text{ini}}, O_{\text{meas}})$  is generally given in the

following form:

$$V^{(t)}(m, \mathcal{E} | O_{\text{ini}}, O_{\text{meas}}) = \sum_{\lambda} \text{Tr}[\hat{A}_{\lambda}^{(t)} (\hat{C}_{\lambda}^{(t)}(\mathcal{E}))^m], \quad (34)$$

where  $\lambda$  labels the irreps of the unitary group,  $\hat{A}_{\lambda}^{(t)}$  and  $\hat{C}_{\lambda}^{(t)}(\mathcal{E})$  are  $m_{\lambda} \times m_{\lambda}$  matrices with  $m_{\lambda}$  being the multiplicity of the irrep  $\lambda$ . This is well-known in the literature of RB-type protocols, but we provide a proof in Sec. C for completeness.

Despite its abstract expression, Eq. (34) has an important implication that the matrix  $\hat{C}_{\lambda}^{(t)}(\mathcal{E})^m$  depends only on  $\mathcal{E}$  and  $m$ , but not on  $O_{\text{ini}}$  and  $O_{\text{meas}}$ . Hence, from the experimental data of  $V^{(t)}(m, \mathcal{E} | O_{\text{ini}}, O_{\text{meas}})$  for various  $m$ , it is, in principle, possible to estimate the matrix  $\hat{C}_{\lambda}^{(t)}(\mathcal{E})$ , which contains certain information of the noise  $\mathcal{E}$ , in the way independent of  $O_{\text{ini}}$  and  $O_{\text{meas}}$ .

In practice, the most important situation is when the representation is multiplicity-free, that is,  $m_{\lambda} = 1$  for any  $\lambda$ . In this case,  $V^{(t)}$  reduces to a much simpler form:

$$V^{(t)}(m, \mathcal{E} | O_{\text{ini}}, O_{\text{meas}}) = \sum_{\lambda} A_{\lambda}^{(t)} (C_{\lambda}^{(t)}(\mathcal{E}))^m, \quad (35)$$

where  $A_{\lambda}^{(t)}, C_{\lambda}^{(t)}(\mathcal{E}) \in \mathbb{R}$ . Note that  $|C_{\lambda}^{(t)}(\mathcal{E})| \leq 1$  because  $V^{(t)}$  is a bounded function. Hence, in this case,  $V^{(t)}$  becomes a sum of some exponentially decreasing functions with respect to  $m$ .

To be more concrete, let us recall the standard RB, corresponding to the 1-RB. As shown in Ref. [22],  $V^{(1)}$  is given by

$$V^{(1)}(m, \mathcal{E} | O_{\text{ini}}, O_{\text{meas}}) = A_0^{(1)} + A_1^{(1)} f(\mathcal{E})^m, \quad (36)$$

where  $A_0^{(1)}$  and  $A_1^{(1)}$  depend only on  $O_{\text{ini}}$  and  $\mathcal{E}(O_{\text{meas}})$ , and  $f(\mathcal{E})$  is the fidelity parameter of the noise  $\mathcal{E}$ . Thus, by fitting experimentally obtained data of  $V^{(1)}$  for different  $m$  with the fitting function  $A + B\alpha^m$ , we can estimate the fidelity parameter  $f(\mathcal{E})$ .

### D. Importance of exact designs in RB

In the RB protocol, it is important to use *exact* unitary designs because designs are used many times, sometimes a few hundreds to a thousand, in a single run of the protocol. To illustrate this, let us consider the 1-RB when the unitary 2-design in the protocol is  $\epsilon$ -approximate.

Let  $m$  be the length of the unitary sequence as previously. It is straightforward to show that

$$V^{(1)}(m, \mathcal{E} | O_{\text{ini}}, O_{\text{meas}}) \approx A'_0 + A_1 f^m + \epsilon(m-2) \times (E_2 f^2 + E_1 f + E_0) + O(m^2 \epsilon^2), \quad (37)$$

where  $E_i$  are some constants that depend on  $O_{\text{ini}}, \mathcal{E}(O_{\text{meas}})$ ,  $f$ , and how the design differs from the exact one. See



Sec. VIII D for the derivation. Compared with the 1-RB with exact ones, that is, Eq. (36), fitting this function with respect to  $m$  is much harder because it is not a simple exponential decay.

The fitting may go well if  $\epsilon \ll f^m/m$ . This requires a very high precision of the design because  $m$  can be a few hundreds in actual experiments. For instance, when  $f = 0.95$ , the degree  $\epsilon$  of approximation of the unitary design should be order  $10^{-5}$  or so. Although it is possible to achieve this degree of approximation by a sufficiently long quantum circuit [64,65,67], the RB becomes unpractical if we use such a long circuit at every use of a unitary design in the protocol and repeat it a few hundreds times.

There might be a possibility to improve Eq. (37) by using different constructions of approximate unitary designs at every step, by which the differences from the exact design may become random so that they cancel out in total. This will be an interesting question, but at this point, it is not clear if such a technique works. In addition, even if it works, we need to assume additional structures of approximate constructions.

The higher-order RB with approximate designs will incur more difficulty in practice. As it uses higher moment of the outcomes, the fitting function becomes more complicated than Eq. (37) when one uses approximate designs. Similarly to the 1-RB with approximate 2-designs, much better degree of approximation, that is, longer quantum circuits, will be needed, which is not practical. Thus, we conclude that exact unitary designs are of crucial importance in a practical implementation of the  $t$ -RB.

### E. Second-order RB

We next focus on the 2-RB using exact unitary 4-designs, and show that the 2-RB reveals the self-adjointness of the noise. To this end, we set the initial operator  $O_{\text{ini}}$  to a traceless one, that is,  $\text{Tr}[O_{\text{ini}}] = 0$ . This setting, together with the fact that the noise is trace-preserving, makes the representation multiplicity-free (see Appendix C). Hence, the expectation value  $V^{(2)}(m, \mathcal{E}|O_{\text{ini}}, O_{\text{meas}})$  for the 2-RB is given by a sum of exponentially decaying functions as shown in Eq. (35).

Note that the expectation value for a traceless initial operator can be obtained by performing the same experiment for two different quantum states  $\rho$  and  $\rho'$ , and by taking the difference of the expectation values before they are squared. That is,

$$V^{(2)}(m, \mathcal{E}|\Delta, O_{\text{meas}}) = \mathbb{E}_{U_i} [(\langle O_{\text{meas}} \rangle_{\rho, i} - \langle O_{\text{meas}} \rangle_{\rho', i})^2] \quad (38)$$

where  $\Delta = \rho - \rho'$  is a traceless operator.

Our second main result in this paper is about  $V^{(2)}(m, \mathcal{E}|\Delta, O_{\text{meas}})$  as summarized in Theorem 2.

**Theorem 2:** *In the setting described previously,  $V^{(2)}(m, \mathcal{E}|\Delta, O_{\text{meas}})$  is given as follows. For single-qubit systems,*

$$V^{(2)}(m, \mathcal{E}|\Delta, O_{\text{meas}}) = A_0 u(\mathcal{E})^m + A_1 \left( \frac{9}{10} f(\mathcal{E})^2 - \frac{1}{5} u(\mathcal{E}) + \frac{3}{10} h(\mathcal{E}) \right)^m, \quad (39)$$

where  $f(\mathcal{E})$ ,  $u(\mathcal{E})$ , and  $h(\mathcal{E})$  are the fidelity parameter, the unitarity, and the self-adjointness parameter of the noise  $\mathcal{E}$ , respectively. For multiqubit systems,

$$V^{(2)}(m, \mathcal{E}|\Delta, O_{\text{meas}}) = A_0 u(\mathcal{E})^m + \sum_{\lambda=I, II, III} A_\lambda C_\lambda(\mathcal{E})^m, \quad (40)$$

where  $0 \leq C_\lambda(\mathcal{E}) \leq 1$  depend only on the noise  $\mathcal{E}$ . Moreover, they satisfy

$$\sum_{\lambda=I, II, III} D_\lambda C_\lambda(\mathcal{E}) = \frac{(d^2 - 1)^2}{2} f(\mathcal{E})^2 - u(\mathcal{E}) + \frac{d^2 - 1}{2} h(\mathcal{E}), \quad (41)$$

where

$$D_I = \frac{d^2(d-1)(d+3)}{4}, \quad (42)$$

$$D_{II} = \frac{d^2(d+1)(d-3)}{4}, \quad (43)$$

$$D_{III} = d^2 - 1. \quad (44)$$

See Sec. VIII E for the proof. In the single-qubit case,  $V^{(2)}$  is a sum of two exponentially decaying functions with respect to  $m$ . Hence, from the double-exponential fitting of the experimental data of  $V^{(2)}$ , we can simultaneously estimate  $u(\mathcal{E})$  and  $\frac{9}{10} f(\mathcal{E})^2 - \frac{1}{5} u(\mathcal{E}) + \frac{3}{10} h(\mathcal{E})$ . As it can be shown that the former is not less than the latter, we can estimate which of the two decaying rates corresponds to which quantity without any ambiguity. It is also possible to estimate the fidelity parameter  $f(\mathcal{E})$  from the same data set by computing  $V^{(1)}(m, \mathcal{E}|\Delta, O_{\text{meas}})$  because a unitary 2-design is also a unitary 1-design. Thus, from the experiment of the 2-RB on a single qubit, all of  $f(\mathcal{E})$ ,  $u(\mathcal{E})$ , and  $h(\mathcal{E})$  can be estimated simultaneously.

In multiqubit systems,  $V^{(2)}$  has a little more complicated form and consists of four exponentially decaying functions. In addition, the decaying rates do not directly correspond to neither the unitarity nor the self-adjointness parameter. We observe from Eq. (41) that  $h(\mathcal{E})$  can be obtained from a linear combination of the decaying rates  $C_\lambda(\mathcal{E})$ , the value of  $u(\mathcal{E})$ , and  $f(\mathcal{E})$ .

TABLE I. Coefficients  $A_0, A_I, A_{II}$ , and  $A_{III}$  appearing in Eq. (40), for the 2-qubit case. The first row provides a pair of the initial and measurement operators  $(\Delta, O_{\text{meas}})$ . We have assumed that the average fidelity of the noise  $\mathcal{E}$  is close to 1, so that the inverse unitary  $U_{i_{m+1}}$  can be applied nearly noiseless (see Sec. VIII E for details). The operator  $ZZ$  is  $Z \otimes Z$ , and  $\rho_- := |00\rangle\langle 00| - |11\rangle\langle 11|$ . By choosing proper operators, we can set some coefficients zero, so that experimental estimations of  $C_\lambda(\mathcal{E})$  become easy.

$(\Delta, O_{\text{meas}})$	$(ZZ,  00\rangle\langle 00 )$	$(ZZ, ZZ)$	$(\rho_-, \rho_-)$
$A_0$	1/5	16/15	4/15
$A_I$	4/5	48/5	41/15
$A_{II}$	0	16/3	1/3
$A_{III}$	0	0	2/3

One may think that, in the case of multiple qubits, it is practically intractable to accurately fit four exponentially decaying functions from experimental data because each data point has an error. This difficulty can be circumvented by choosing appropriate initial and measurement operators. By doing so, we can set some of  $A_\lambda$  as zero in the ideal situation (see Table I). This allows us to estimate the decaying rates one by one. Note that the initial and measurement operators in Table I are all diagonal in the computational basis. Hence, it suffices to perform the experiments for the four initial operators  $|00\rangle, |01\rangle, |10\rangle$ , and  $|11\rangle$ , with the measurement in the computational basis. From the data of these experiments, it is possible to reproduce all cases listed in Table I by postprocessing.

In the multiqubit case, the ambiguity remains to decide which of the decaying rates corresponds to which quantity. This is the case even when we use the step-by-step estimation of the rates because, for instance, it is not clear if the unitarity  $u(\mathcal{E})$  is larger or smaller than  $C_I$ . In this case, we need to additionally perform the unitarity benchmarking (UB) [69,89] to separately estimate  $u(\mathcal{E})$ . If we have an estimated value of  $u(\mathcal{E})$ , the step-by-step estimation allows us to decide all decaying rates without any ambiguity.

See Sec. V A and Sec. IX A for the performance of 2-RB in concrete cases.

### F. Scalability

The  $t$ -RB for  $t \geq 2$  inherits most of the desired properties of the RB-type protocols. For instance, it is experimentally friendly because, apart from using higher-order designs, the difference of the  $t$ -RB from the standard RB (1-RB) is only taking the  $t$ th power of the expectation value before the average. It is also true that the  $t$ -RB is free from SPAM errors (see Eqs. (34) and (35)).

The property that the standard RB does have and the  $t$ -RB does not in general is the scalability. This is for two reasons. First, no efficient construction of exact unitary  $2t$ -designs has been found for  $t \geq 2$  so far. Second, in the  $t$ -RB protocol, it is necessary to apply the inverse unitary

at the end of the unitary sequence. Hence, we need to compute the inverse of each sequence beforehand. When the system is large, the task is intractable in general. This difficulty is avoided in the standard RB by using the Clifford group, which is an exact unitary 2-design. As the inverse is contained in the group, we can find the inverse relatively easily. One may expect that the difficulty of finding the inverse could be also avoided in the  $t$ -RB by using the  $2t$ -design that is also a group, which is called a unitary  $2t$ -group [52]. However, it is known that unitary  $2t$ -groups do not exist for  $t \geq 2$  if the number of qubits is at least three. Thus, in the  $t$ -RB for  $t \geq 2$ , the hardness of finding the inverse in a large system is inevitable.

Nonetheless, we emphasize that, in the current experimental situations, the RB-type protocols for more than three qubits are practically intractable due to the limitation of the coherent time. Thus, the experimental use of the RB-type protocols is currently aiming to characterize the noise on one- or two-qubit systems in a concise manner. Considering this fact, even if the  $t$ -RB is not scalable, it is practically useful and beneficial: it is as concise as the standard RB and provides more information about the noise, such as self-adjointness.

## V. MAIN RESULT 3: 2-RB IN A SUPERCONDUCTING SYSTEM

We finally implement the 2-RB in a superconducting system and estimate the self-adjointness of background noise. Unlike the analytical studies, the expectation values and the average over a unitary 4-design cannot be taken with arbitrary precision in experiments because the number of repetitions of the experiment is practically limited. To check that this limitation does not cause any problem in the evaluation of the self-adjointness, we start with numerically investigating the feasibility of the 2-RB in Sec. V A. We then provide a summary of experimental results in Sec. V B.

In recent years, a number of experiments have been performed to characterize various noises on superconducting quantum systems in detail [90–93]. From our experiments, we show that the interactions with the adjacent qubits particularly decrease the self-adjointness and may cause problems toward realizations of QEC. In particular, our result implies that there exists a gap between the superconducting system and the common noise model used in theoretical studies of QEC, and also that the standard decoders of stabilizer codes may suffer from degradation of logical errors. Hence, toward the realization of QEC, it is desired to further improve the system or to develop the theory of QEC.

### A. Numerical evaluation

When the 2-RB is practically implemented, there are two additional concerns. One is originated from the fact

that the expectation value  $\langle O_{\text{meas}} \rangle_{O_{\text{ini}}, i}$  is obtained from a limited number of measurements in the basis of  $O_{\text{meas}}$ , resulting in an error owing to a finite number of measurements. The other originates from the evaluation of the average  $\mathbb{E}_{U_i \sim \mathcal{U}_4^{\times m}}$  over the sequence of unitaries in the 4-design. Ideally, all sequences in  $\mathcal{U}_4^{\times m}$  should be taken, but practically, the average is often evaluated from a small subset in  $\mathcal{U}_4^{\times m}$  of randomly chosen sequences, leading to an additional error of estimation.

Taking sufficiently many measurements and samplings of unitary sequences will reproduce the analytical results with high accuracy. However, it is complicated to analytically derive the numbers sufficient for achieving a desired accuracy. We, hence, perform numerical experiments and show that experimentally tractable numbers of samplings are sufficient for a reliable 2-RB.

### 1. One-qubit cases

In the case of single-qubit systems, we consider a specific noisy map given by

$$\mathcal{E}_1(\rho) = qe^{i\theta X} \rho e^{-i\theta X} + (1-q)((1-p)\rho + pX\rho X), \tag{45}$$

which is characterized by three parameters  $p, q, \theta$ . The first term of the right-hand side represents a unitary part and the second term represents a stochastic part of the noise. A parameter  $q$  determines a ratio between them. Hence, we can consider  $q$  as a coherent parameter of noise, e.g., noise is unitary when  $q = 1$  and is a probabilistic Pauli noise when  $q = 0$ . The parameters  $\theta$  and  $p$  represent the rotation angle of the unitary part and the error probability of

the stochastic part, respectively. For simplicity, we choose  $\theta$  such that the fidelity parameters of unitary and stochastic parts are equal, that is,  $p = \sin^2 \theta$ . Then, the fidelity parameter  $f(\mathcal{E}_1)$  becomes independent of the coherent parameter  $q$ .

To perform the 2-RB for this noise, we may use the exact 4-design constructed in Corollary 2. However, it is known that the icosahedral group, which we denote by  $I$ , forms an exact 4-design on one qubit [49]. As the icosahedral group has less cardinality than our inductive construction, we use it in the following analysis.

The numerical results for the 2-RB on a single qubit are shown in Fig. 2. For each sequence length  $m$ , we have taken 5000 random unitary sequences from  $I^{\times m}$  and have had 5000 measurements to obtain a single data point of  $V^{(1)}$ . A detailed fitting procedure is provided in Sec. IX A.

To check the accuracy of the 2-RB, we consider the relative errors  $|y - \tilde{y}|/(1 - y)$ , where  $y$  and  $\tilde{y}$  are the theoretical value and the fitting value, respectively. Note that  $1 - y \sim 0$  for all the fitting values when a fidelity close to unity is achieved. For almost all data points of  $F(\mathcal{E})$ ,  $u(\mathcal{E})$ , and  $H(\mathcal{E})$ , we find that the relative errors are less than 5.0%, except the case when  $p$  is large or, equivalently, when the fidelity is small. The relative error becomes moderately large, such as 35%, when  $p = 0.4$  and  $q \geq 0.1$ , corresponding to  $F(\mathcal{E}_1) = 0.7$ . This is because the decaying rate of the second term in Eq. (39) is rather small, making the fitting difficult. However, such a case is not practically relevant because the fidelity is typically higher than 90%. Thus, we conclude that the 2-RB on 1-qubit systems works well in practice.

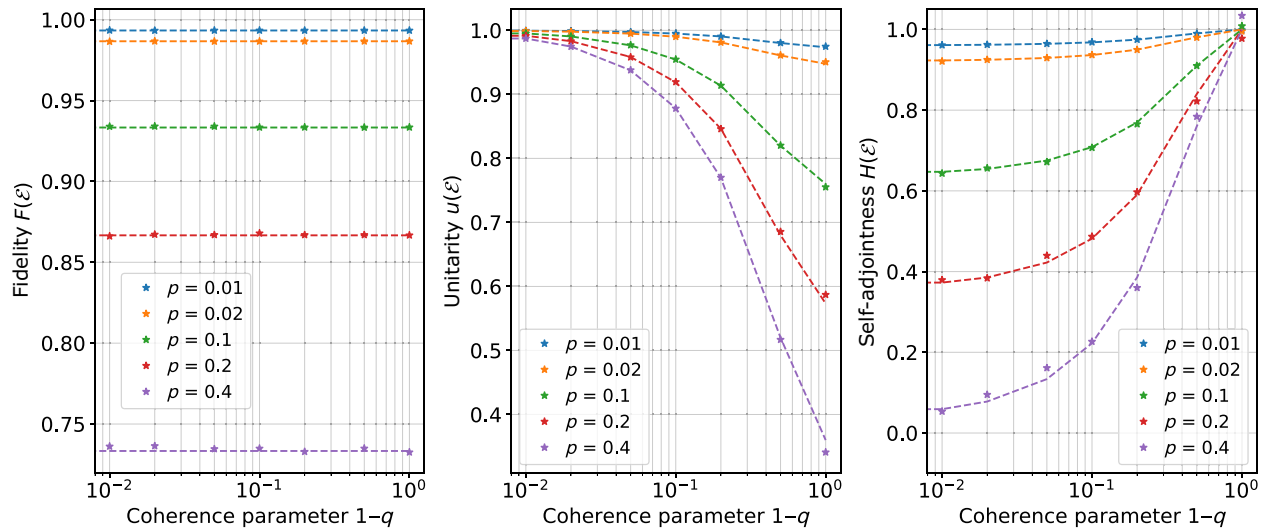


FIG. 2. The estimated values of  $F(\mathcal{E}_1)$ ,  $u(\mathcal{E}_1)$ , and  $H(\mathcal{E}_1)$  obtained by 2-RB on one-qubit system for various parameters  $p$  and  $q$ , where we have taken 5000 samplings both for measurement and for unitary sequences. The dots show the fitting results, and the dashed lines represent theoretical values.

TABLE II. Numerically estimated values of the fidelity, unitarity, and self-adjointness from the single-qubit 2-RB with the finite numbers of measurements and samplings of unitary sequences. We set  $p = q = 0.02$ . The theoretical values are listed at the bottom of the table.

Measurements	Sequences	$F$	$u$	$H$
10	100	0.986(6)	0.9985(5)	0.92(6)
10	500	0.986(1)	0.9984(1)	0.93(1)
10	1000	0.986(6)	0.9984(3)	0.92(3)
100	100	0.986(3)	0.9980(8)	0.92(2)
100	500	0.986(5)	0.9980(4)	0.92(3)
100	1000	0.986(6)	0.9979(4)	0.92(1)
1000	100	0.986(3)	0.9980(1)	0.92(3)
1000	500	0.986(4)	0.9979(5)	0.92(3)
1000	1000	0.986(6)	0.9978(9)	0.92(2)
		0.9866	0.99793	0.9247

To analyze the dependence of the accuracy of the 2-RB on the number of measurements and samplings of random unitary sequences, we additionally perform the 2-RB on one qubit with the various numbers of measurements and samplings. The results are summarized in Table II, where we set the noise parameters to  $p = 0.02$  and  $q = 0.02$ . From these results, it appears that setting the numbers of measurements and samplings of random sequences to a few hundreds is sufficient for a good estimate. These results further indicate that increasing the number of random sequences rather than the number of measurements is preferable to improve the accuracy. See Sec. IX A for the details.

## 2. Two-qubit cases

For two-qubit systems, we consider the noise given by

$$\mathcal{E}_2(\rho) = qe^{i\theta(X \otimes X)} \rho e^{-i\theta(X \otimes X)} + (1-q)((1-p)\rho + p(X \otimes X)\rho(X \otimes X)), \quad (46)$$

which is similar to the one-qubit case. We choose  $\theta$  as  $p = \sin^2 \theta$ , so that  $f(\mathcal{E}_2)$  is independent of the coherent parameter  $q$ . In this case, we use the construction of exact unitary 4-designs given in Proposition 1.

In the two-qubit case, it is needed to fit the experimental data by a sum of four exponentially decaying functions, which is not, in general, easy especially when each data point has errors caused by the finite number of measurements and samplings of unitary sequences. To avoid this difficulty, we use the method explained in Sec. IV E, and determine the four decaying rates, i.e.,  $u(\mathcal{E}_2)$ ,  $C_I(\mathcal{E}_2)$ ,  $C_{II}(\mathcal{E}_2)$ , and  $C_{III}(\mathcal{E}_2)$  in Eq. (40), one by one.

The results are shown in Fig. 3. We have taken  $10^4$  random unitary sequences for each sequence length  $m$  and the parameters  $p, q$ . A detailed process of fittings are explained step by step in Sec. IX A. In the figure, fitted values are shown as data points. Dashed lines are drawn with theoretically calculated values.

Similarly to the case of the single-qubit 2-RB, we have checked the relative errors of the fitting results to the theoretical values. The errors are all less than 3% for all the points except the case when the theoretical value is exactly zero. As in the case of single-qubit 2-RB, when we calculate  $F(\mathcal{E})$ ,  $u(\mathcal{E})$ , and  $H(\mathcal{E})$  from the fitting values, the relative values of almost all the data points are less

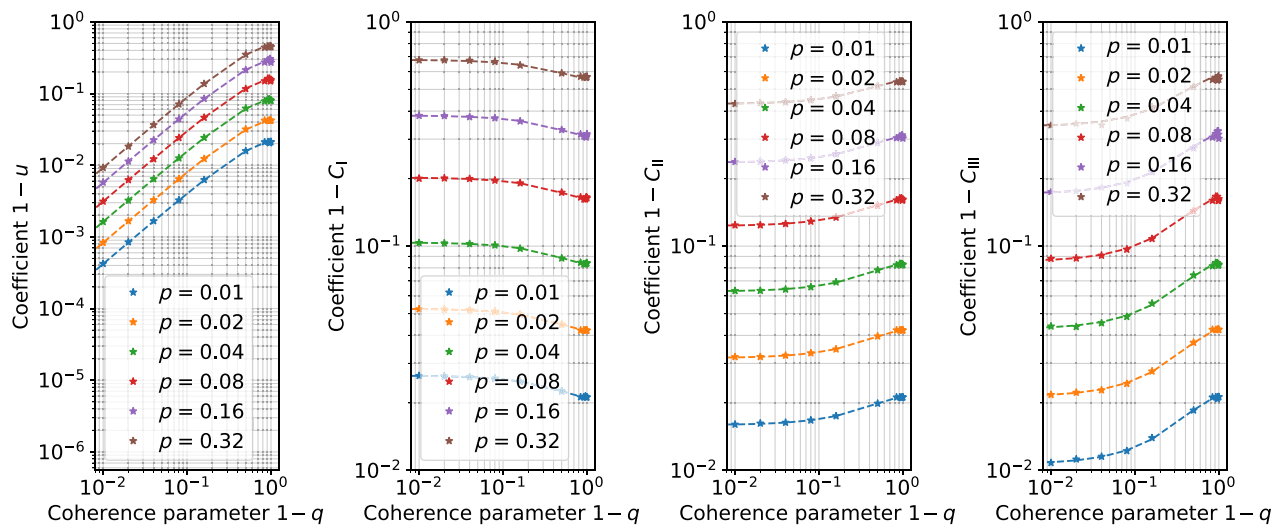


FIG. 3. Four coefficients for exponentially decaying functions in the 2-RB are plotted according to coherence parameter  $q$ , where we had  $10^4$  measurement and  $10^4$  samplings of unitary sequences. Dashed lines are theoretical values. Each color corresponds to each value of the parameter  $p$ .



than 4.0%. Although the relative errors become large when  $p = 0.4$ , such a case is not a problem in typical calibration scenario. Thus, the 2-RB works in actual situations also in the case of two-qubit systems.

### B. Experimental implementations of the 2-RB

We demonstrate the 2-RB in a superconducting-qubit system. We first explain the setup of our experiments, and then verify the feasibility of the 2-RB experiment by comparing the unitarity obtained from the 2-RB with that from the UB [69,89]. We finally characterize background noise of the system. As the background noise is gate- and time-independent, it satisfies the assumptions of the 2-RB (see Sec. VIII B for the detail).

#### 1. Experimental setup

We use two superconducting qubits ( $Q_1$  and  $Q_2$ ) coupled with each other via an electric dipole interaction, which are a part of our 16-qubit device [94]. In all the following experiments, we use the qubit  $Q_1$  as a target qubit of the single-qubit 2-RB and, in some experiments,  $Q_2$  as an environmental qubit that induces additional error onto  $Q_1$ .

The simplified system Hamiltonian  $H$  is formulated as follows:

$$\frac{H}{\hbar} = \frac{\omega_1}{2} Z \otimes I + \frac{\omega_2}{2} I \otimes Z + \frac{\chi_{ge}}{2} Z \otimes Z, \quad (47)$$

where  $\omega_i/2\pi$  is the eigenfrequency of the  $i$ th qubit and  $\chi_{ge}/2\pi = -0.760$  MHz is an effective interaction strength between the qubits [95]. It can be interpreted that the eigenfrequency of  $Q_1$  switches depending on the quantum state of  $Q_2$ . When  $Q_2$  is in the  $|0\rangle$  ( $|1\rangle$ ) state,  $Q_1$  has the eigenfrequency  $(\omega_1 + \chi_{ge})/2\pi$  ( $(\omega_1 - \chi_{ge})/2\pi$ ). In the Bloch sphere representation, the state vector of the qubit rotates around the  $Z$ -axis with its eigenfrequency as the angular velocity.

We use a local oscillator synchronized with the eigenfrequency of the qubit for observation. The state vector is stationary in a rotating frame of the local oscillator because the  $Z$ -axis rotation speed of the Bloch vector matches with that of the measurement basis. The rotation frame picture also holds when the qubit  $Q_1$  couples to the adjacent qubit  $Q_2$  when the qubit  $Q_2$  is in the  $|0\rangle$  or  $|1\rangle$  state. For instance, when the qubit  $Q_2$  is always in the  $|0\rangle$  state, the eigenfrequency of  $Q_1$  is  $(\omega_1 + \chi_{ge})/2\pi$ . We can detune the frequency of the local oscillator from the qubit frequency  $\omega_1$  by  $\chi_{ge}$  to make the state vector of  $Q_1$  stationary.

It is, however, impossible to keep track of the eigenfrequency of the qubit when the state of the adjacent qubit varies. This results in an inevitable  $Z$ -rotation occurring in the quantum state. In an actual experiment involving multiple qubits, the frequency of the local oscillator is usually set to  $\omega_1/2\pi$  to minimize the average  $Z$ -rotation angle. See Sec. VIII B for the detail.

### 2. Comparison with the UB

In the experiment aiming to compare the 2-RB and the UB on a single qubit, we use only  $Q_1$  and add an artificial noise after applying each gate. The isolation of the qubit  $Q_1$  from the qubit  $Q_2$  can be done by keeping the qubit  $Q_2$  in the state  $|0\rangle$  and by setting the frequency of the local oscillator to  $(\omega_1 + \chi_{ge})/2\pi$ , which effectively cancel the coupling between  $Q_1$  and  $Q_2$ . Regarding the noise, we especially choose a single-qubit  $Z$ -rotation by angle  $0.2\pi$ , denoted by  $R_Z(0.2\pi)$ .

Both in the case of the 2-RB and the UB, we use the icosahedral group I and the Clifford group on a single qubits, respectively. Note that the former is an exact 4-design on a single qubit, and the latter is an exact 2-design.

We have taken 100 and 1000 random sequences for the 2-RB and the UB, respectively. This is because the UB with the Clifford group converges slower than the 2-RB with the icosahedral group, which is likely due to the fact that the former and the latter are based on unitary 2- and 4-designs, respectively. A higher-order design typically leads to a quick convergence because it is more concentrating around the average [96]. A faster convergence of the UB with 4-design is expected, which highlights the potential use of a higher-design also for the UB. We have taken  $10^4$  measurements for each random sequence to obtain a data point of  $V^{(2)}$ . The results are summarized in Table III.

From the results, we observe that the unitarity characterized by the 2-RB matches with that by the UB. This indicates that the 2-RB on our single-qubit system works to characterize the gate performance.

Note that the difference between the unitarity from the 2-RB and that from the UB is slightly beyond the standard deviation. This is likely because the noise property varies in the UB experiment. As mentioned, we have taken 1000 random sequences in the UB to ensure the convergence of the statistical average, which has taken more than 10 h in total. As the noise in the experimental system drifts in such a long timescale, the situation of the experiment deviates from the ideal situation, where time-independence of the noise is assumed. Indeed, unlike the theoretical prediction of the UB, the data is slightly different from a single-exponential decay. This deviation is expected to be the origin of a less-precise value of the unitarity estimated from the UB.

TABLE III. The estimated values of the fidelity, unitarity, and self-adjointness from the experiment of the 2-RB and that of the UB in the superconducting qubit system.

Experiment	Group	$F$	$u$	$H$
2-RB	Icosahedral	0.926(6)	0.970(1)	0.6(1)
UB	Clifford	...	0.977(1)	...
Theoretical	...	0.936	1	0.655

### 3. Characterizing background noise

We next perform the single-qubit 2-RB, aiming to characterize background noise of the qubit  $Q_1$  in the experimental system. We intentionally insert a delay time  $t$  after each application of a gate to extract the information of the background noise.

In the following experiments, we have taken 100 random unitary sequences from  $I^{\times m}$ , where  $I$  is the icosahedral group, for each sequence length  $m$  and have had  $10^4$  measurements for each random sequence to get a data point of  $V^{(2)}$ .

In the first experiment, we set the frequency of the local oscillator to  $(\omega_1 + \chi_{ge})/2\pi$  and treat the qubit  $Q_1$  as a target qubit isolated from the qubit  $Q_2$ . The background noise of the isolated qubit is often phenomenologically modeled by the Lindblad master equation given by

$$\frac{d\rho}{dt} = \sum_{k \in [1,2]} L_k^\dagger \rho L_k - \frac{1}{2} \left\{ L_k^\dagger L_k, \rho \right\}, \quad (48)$$

where  $L_1 = \hat{a}/\sqrt{T_1}$  represents the energy dissipation with the relaxation time  $T_1$ ,  $\hat{a} = (X + iY)/2$  is an annihilation operator of the qubit, and  $L_2 = Z/\sqrt{2T_\phi}$  represents the phase dissipation with the relaxation time  $T_\phi = 1/(1/T_2 - 1/2T_1)$ . By solving the Eq. (48), we can obtain phenomenological predictions about the background noise  $\mathcal{E}_t$  corresponding to the delay time  $t$ .

We sweep the delay time  $t$  from 100 to 500 ns. The value  $V^{(2)}$  obtained from the experiments is shown in Fig. 4(a). We estimate the unitarity and the self-adjointness from  $V^{(2)}$  through a fitting based on a sum of *two* exponentially-decaying curves given in Eq. (39). However, we observe single-exponential decays from the results. This indicates two possibilities. One is that the two decaying rates are nearly the same. The other is that one of the two decaying rates is much smaller than the other, so that one exponentially decaying curve becomes quickly negligible as  $m$  increases.

In our experiment, the former is the case because the average fidelity is high, which is confirmed from the 1-RB. We can analytically show that the two decaying rates typically coincide when the fidelity is sufficiently high. More specifically, we have (see Eq. (101) in Sec. VIII B)

$$1 - 4\epsilon \lesssim \frac{h(\mathcal{E}) + u(\mathcal{E})}{2}, \quad (49)$$

to the first order of  $\epsilon$ , where  $\epsilon$  is the infidelity  $1 - F(\mathcal{E})$ . This implies that the two decaying rates in Theorem 2 are approximately greater than  $1 - 4\epsilon$  and  $1 - 6\epsilon$ , respectively. Thus, if  $\epsilon \ll 1$ , which is indeed the case in our system, the two decaying rates are hard to distinguish, making the curve of  $V^{(2)}$  a single-exponential decay.

We, hence, estimate the single-exponential decay rate from the experimental data of the 2-RB and derive  $u(\mathcal{E}_t)$

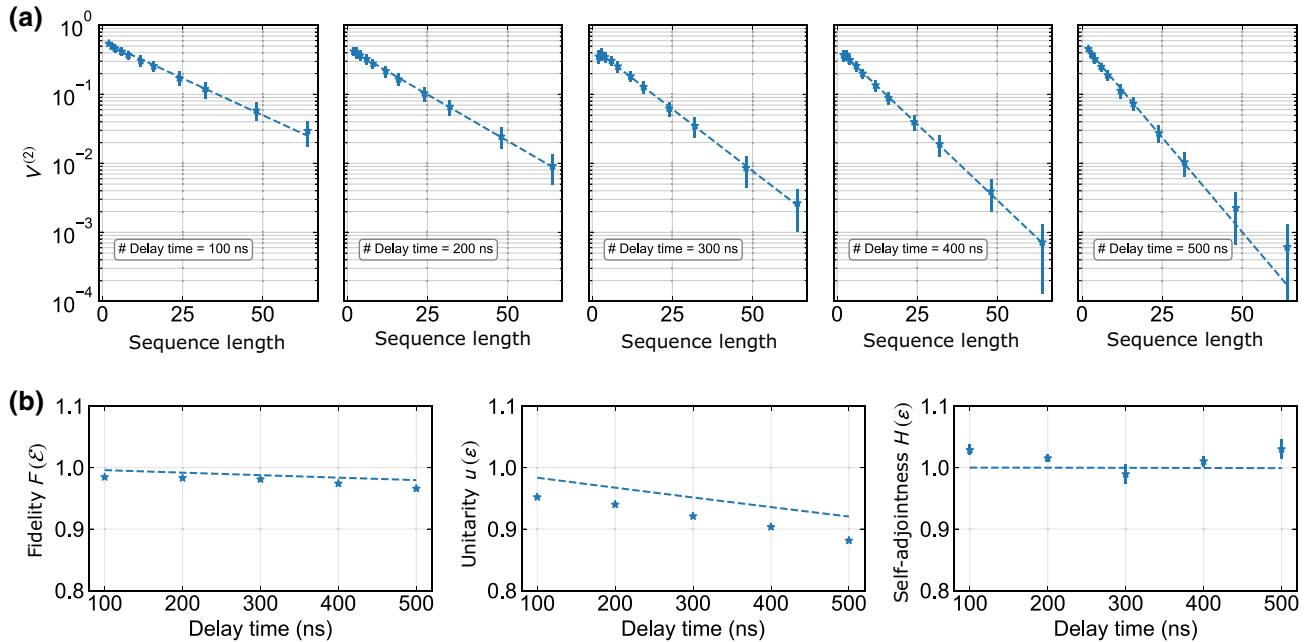


FIG. 4. (a) Experimental results for the single-qubit 2-RB on an isolated qubit when the interleaved delay times  $t$  are swept. We have taken 100 random unitary sequences from  $I^{\times m}$  for each sequence length  $m$  and have had  $10^4$  measurements for each sequence to obtain a data point of  $V^{(2)}$ . The stars represent the values of  $V^{(2)}$  and the dashed lines are fitting results. The error bars represent the standard deviation of  $V^{(2)}$ . As the vertical axis of the figure is logarithmic notation, the error bars at the bottom of the figure are displayed larger. (b) Estimated values of  $F(\mathcal{E}_t)$ ,  $u(\mathcal{E}_t)$ , and  $H(\mathcal{E}_t)$  obtained by the 2-RB. The stars show the fitting results, and the dashed lines are the predicted values from the phenomenological model.

and  $h(\mathcal{E}_t)$  from

$$u(\mathcal{E}_t) = \frac{9}{10}f(\mathcal{E}_t)^2 - \frac{1}{5}u(\mathcal{E}_t) + \frac{3}{10}h(\mathcal{E}_t). \quad (50)$$

Here,  $u(\mathcal{E}_t)$  is obtained from the estimated decaying rate, and  $f(\mathcal{E}_t)$  from the 1-RB (see Eq. (39)).

The obtained fidelity, unitarity, and self-adjointness are summarized in Fig. 4(b). In calculating self-adjointness, we solved Eq. (28), where we substituted  $\alpha_{\mathcal{E}}$  of the phenomenological prediction. They reveal that the background noise of the isolated qubit has the unitarity  $u(\mathcal{E}_t)$  that slowly decreases as the delay increases, whereas its self-adjointness  $H(\mathcal{E}_t)$  is nearly independent of the delay. As we have explained in Sec. IV A, a problem may occur when the unitarity is high and the self-adjointness is low, which is not observed in this experiment. Hence, we conclude that the background noise in this case would not cause any problem toward the realization of QEC.

In the second experiment, we set the frequency of the local oscillator to  $\omega_1/2\pi$  and treated  $Q_1$  as a target qubit exposed to the noise induced by the adjacent qubit  $Q_2$ . In this experiment, no control pulses are applied to  $Q_2$ , so that  $Q_2$  is expected to remain in the  $|0\rangle$  state. This leads to a continuous rotation of the state vector of  $Q_1$  by the interaction Hamiltonian term of  $\chi_{ge}Z/2$ .

In this case, the background noise with the interaction Hamiltonian is modeled by the Lindblad master equation

written as follows:

$$\frac{d\rho}{dt} = \left[ \frac{\chi_{ge}}{2}Z, \rho \right] + \sum_{k \in [1,2]} L_k^\dagger \rho L_k - \frac{1}{2} \left\{ L_k^\dagger L_k, \rho \right\}, \quad (51)$$

providing a phenomenological model.

Similarly to the first experiment, we sweep the delay time  $t$  from 60 to 180 ns. The delay time is set to a shorter time than the first experiment because the fidelity deteriorates due to the Z-rotation error. As the Z-rotation error does not affect the unitarity, we conclude that the decay rate, which is less sensitive to the delay time than the other, corresponds to the unitarity. The results of the experiment are shown in Fig. 5(a). As seen from the results, the curves  $V^{(2)}$  obey double-exponential decay. From the two decaying rates, we obtain the unitarity  $u(\mathcal{E}_t)$  and the self-adjointness  $H(\mathcal{E}_t)$  as a function of the delay time as depicted in Fig. 5(b). Note that, although the unitarity may seem different from the former experiment, it is merely due to the different time scale of the horizontal axis. The unitarities in the two experiments indeed coincide within the standard deviation (see, e.g. the delay time 100 (ns)).

The experimental results qualitatively coincide with the phenomenological predictions obtained from Eq. (51) [see Fig. 5(b)]. However, the experimental values tend to be smaller. This indicates that there exist noise sources not included in the phenomenological model. The candidates

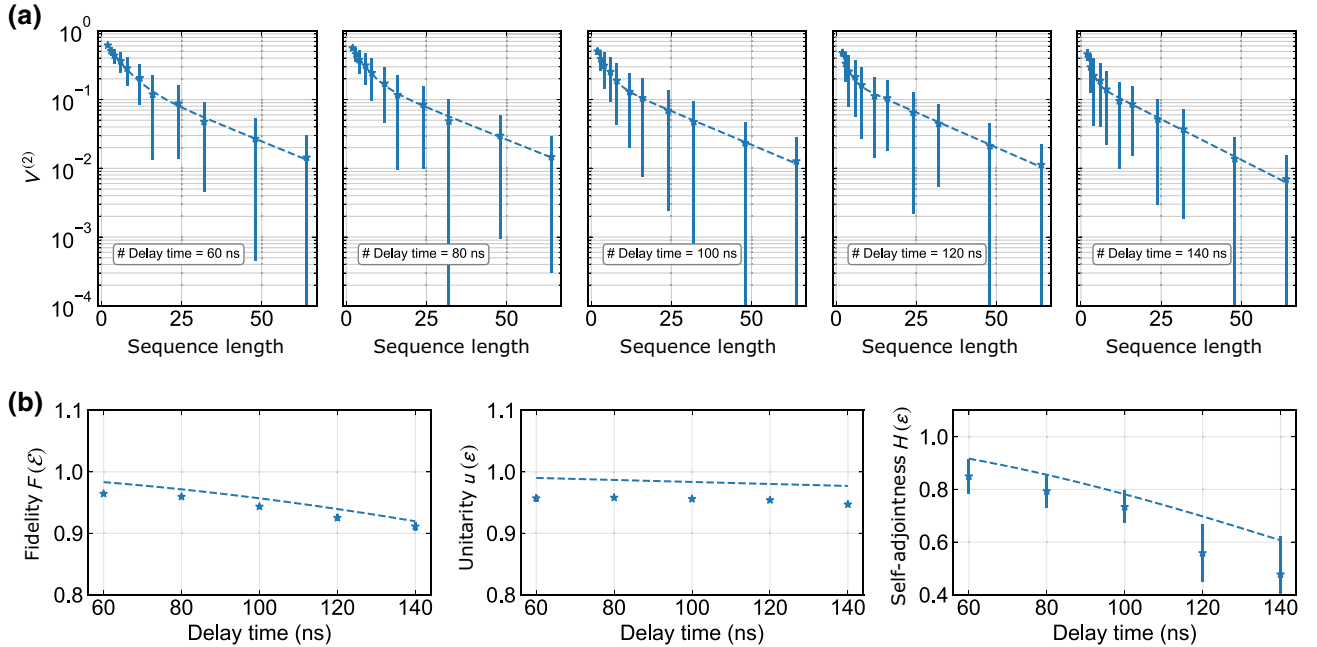


FIG. 5. (a) Experimental results for the single-qubit 2-RB on the qubit coupled to another qubit, where we sweep the interleaved delay time  $t$ . We have taken 100 random unitary sequences from  $I^{\times m}$  for each sequence length  $m$  and have had  $10^4$  measurements for each random sequence to take a data point of  $V^{(2)}$ . The stars correspond to the values of  $V^{(2)}$  and the dashed lines are fitting results. (b) Estimated values of  $F(\mathcal{E}_t)$ ,  $u(\mathcal{E}_t)$ , and  $H(\mathcal{E}_t)$  obtained by the 2-RB. The stars show the fitting results, and the dashed lines are the predicted values from the phenomenological model.

of the additional noise sources are calibration errors in the  $R_X(\pi/2)$  gates, the initial thermal excitation rate of  $Q_2$  (7.2%), and the interaction of  $Q_1$  with adjacent qubits other than  $Q_2$ . Note that the initial thermal excitation of  $Q_2$  makes the noise time-dependent owing to the relaxation and, hence, makes the result different from the theoretical prediction of the 2-RB.

Compared with the first experiment (Fig. 4), we observe from Fig. 5 that the fidelity  $F(\mathcal{E}_t)$  and the self-adjointness  $H(\mathcal{E}_t)$  quickly decrease as the delay time  $t$  increases. The latter decreases especially quickly:  $H(\mathcal{E}_t) \approx 0.48$  at  $t = 140$  (ns). This implies that, even if the fidelity is moderately high ( $F(\mathcal{E}_t) \approx 0.91$  at  $t = 140$  (ns)), the extra  $Z$ -rotation induced by the interaction with another qubit radically changes the property of the noise and makes the noise far from self-adjoint. Consequently, as the delay increases, the noise quickly becomes the one that cannot be approximated by any stochastic Pauli noise.

This result has an important implication toward a realization of QEC. As mentioned in Sec. IV A, theoretical studies of QEC commonly assume stochastic Pauli noises to numerically compute error thresholds and error rates. Our result implies that, when the interaction with another qubit is nonnegligible, we cannot directly apply the theoretical predictions based on Pauli noises. This problem will be more prominent when the system size grows because, in a large system, a qubit interacts with more qubits in an uncontrolled manner, making the noise much less self-adjoint and much far from Pauli noises. To circumvent this, effective cancellation of the dipole interaction is of great importance in the further improvement because the dominant interaction between qubits should be originated from the electric dipole interaction.

This feature of the noise, that is, interactions with other qubits inducing small self-adjointness and the difficulty of approximating the noise by a Pauli noise, is expected to be common in any experimental systems. The 2-RB experiment and the self-adjointness offer a useful method and measure, respectively, to experimentally evaluate the noise in the system from this perspective.

## VI. STRUCTURE OF THE REMAINDER OF THIS PAPER

The remainder of this paper is organized as follows. In Sec. VII, a proof of Theorem 1 is provided. A brief introduction of representations of the unitary group is also provided before the proof. We then explain the higher-order RB in Sec. VIII, including the proof of Theorem 2. The methods used in the numerical analysis, and the experimental demonstrations are provided in Sec. IX. After we summarize the paper in Sec. X, we prove technical statements in Appendices.

## VII. CONSTRUCTIONS OF EXACT DESIGNS

In this section, we provide a proof of Theorem 1. We start with a brief introduction of representations of the unitary group in Sec. VII A, and prove Theorem 1 in Sec. VII B.

### A. Unitary $t$ -designs and representation theory

Unitary  $t$ -designs are closely related to representations of the unitary group because the operator  $U^{\otimes t} \otimes U^{\dagger \otimes t}$  in the definition can be regarded as a representation  $\rho$  of  $U \in \mathbf{U}(d)$  on  $\mathcal{H}_d^{\otimes 2t}$  with  $\mathcal{H}_d$  being the Hilbert space with dimension  $d$ , that is,  $\rho(U) = U^{\otimes t} \otimes U^{\dagger \otimes t}$ . It is natural to consider irreps of the unitary group.

A well-known fact is that each irrep can be indexed by a nonincreasing integer sequence  $\lambda := (\lambda_1, \lambda_2, \dots, \lambda_d)$ , that is,  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_d$ , of length  $d$ . In particular, each irrep in  $U^{\otimes t} \otimes U^{\dagger \otimes t}$  can be indexed by an element of a set  $\Lambda(d, t)$  defined by

$$\Lambda(d, t) := \{\lambda = (\lambda_1, \lambda_2, \dots, \lambda_d) | \lambda_1 \geq \dots \geq \lambda_d, \lambda^+ = \lambda^- \leq t\}, \quad (52)$$

where  $\lambda^+$  and  $\lambda^-$  are the absolute value of sum of positive and negative  $\lambda_i$ , respectively. Using this notation, the representation space  $\mathcal{H}_d^{\otimes 2t}$  is irreducibly decomposed into

$$\mathcal{H}_d^{\otimes 2t} = \bigoplus_{\lambda \in \Lambda(d, t)} V_\lambda^{\oplus m_\lambda}, \quad (53)$$

where  $m_\lambda$  is the multiplicity of the irrep  $\lambda$ . Accordingly, the map  $\rho$  is also decomposed into the irreducible ones  $\rho_\lambda$ .

Based on the irrep  $(\rho_\lambda, V_\lambda)$  of the unitary group, a unitary  $t$ -design  $\mathbf{U}_t(d)$  can be characterized in a representation-theoretic manner: for any  $\lambda \in \Lambda(d, t)$ ,

$$\mathbb{E}_{U \sim \mathbf{U}_t(d)}[\rho_\lambda(U)] = \mathbb{E}_{U \sim \mathbf{H}(d)}[\rho_\lambda(U)]. \quad (54)$$

The strong unitary  $t$ -designs are similarly characterized in terms of irreps [49]. To this end, let  $\Lambda_{\leq}(d, t)$  be

$$\Lambda_{\leq}(d, t) := \{\lambda = (\lambda_1, \lambda_2, \dots, \lambda_d) | \lambda_1 \geq \dots \geq \lambda_d, \lambda^\pm \leq t\}, \quad (55)$$

where  $\lambda^+$  is not necessarily equal to  $\lambda^-$ . Then, a strong unitary  $t$ -design  $\mathbf{U}_{\leq t}(d)$  satisfies

$$\mathbb{E}_{U \sim \mathbf{U}_{\leq t}(d)}[\rho_\lambda(U)] = \mathbb{E}_{U \sim \mathbf{H}(d)}[\rho_\lambda(U)], \quad (56)$$

for any  $\lambda \in \Lambda_{\leq}(d, t)$ .

One of the merits in this characterization is that the right-hand sides of Eqs. (54) and (56) are zero for all nontrivial irreps owing to Schur's orthogonality relation,



which states that, for any unitarily inequivalent irreps  $\lambda$  and  $\lambda'$ ,

$$\mathbb{E}_{U \sim \mathbb{H}}[(\rho_\lambda(U))_{ij}(\rho_{\lambda'}(U))_{i'j'}] = 0, \quad (57)$$

for any  $i, j, i', j'$ , where  $(\rho_\lambda(U))_{ij}$  is the  $(i, j)$  element of the matrix. By setting the irrep  $\rho_{\lambda'}$  to a trivial irrep, that is,  $\rho_{\lambda'}(U) = 1$  for any  $U \in \mathbb{U}(d)$ , we have

$$\mathbb{E}_{U \sim \mathbb{H}}[\rho_\lambda(U)] = 0, \quad (58)$$

for any nontrivial irrep  $\lambda$ . On the other hand, for any trivial irrep  $\lambda$ , it is trivial that

$$\mathbb{E}_{U \sim \mathbb{H}}[\rho_\lambda(U)] = 1. \quad (59)$$

From these facts, (strong) unitary  $t$ -designs can be defined in terms of representation as follows.

**Definition 3 (Unitary designs in representation theory):** An ensemble  $\mathbb{U}_t(d)$  of unitaries is an exact unitary  $t$ -design if it holds for any irrep  $\rho_\lambda$  with  $\lambda \in \Lambda(d, t)$  that

$$\mathbb{E}_{U \sim \mathbb{U}_t(d)}[\rho_\lambda(U)] = \begin{cases} 1 & \text{if the irrep is trivial,} \\ 0 & \text{otherwise.} \end{cases} \quad (60)$$

An ensemble  $\mathbb{U}_{\leq t}(d)$  is a strong unitary  $t$ -design if Eq. (60) holds for any irrep  $\rho_\lambda$  with  $\lambda \in \Lambda_{\leq}(d, t)$ .

### B. Proof of theorem 1

We now prove Theorem 1, which states that  $\mathbb{W}_d$  defined by

$$\mathbb{W}_d := \mathbb{W}_{d_1 \oplus d - d_1} \prod_{\lambda \in \Lambda_{\text{sph}}(d_1, d, t)} (R_\lambda \mathbb{W}_{d_1 \oplus d - d_1}), \quad (61)$$

is a strong unitary  $t$ -design on  $\mathbb{U}(d)$ . Here,

$$\mathbb{W}_{d_1 \oplus d - d_1} = \{U \oplus V \mid U \in \mathbb{U}_{\leq t}(d_1), V \in \mathbb{U}_{\leq t}(d - d_1)\}, \quad (62)$$

where  $\mathbb{U}_{\leq t}(d)$  and  $\mathbb{U}_{\leq t}(d - d_1)$  are strong unitary  $t$ -designs on  $\mathbb{U}(d)$  and  $\mathbb{U}(d - d_1)$ , respectively, and  $R_\lambda$  is constructed by solving the zonal spherical function  $Z_\lambda$ .

It suffices to show

$$\mathbb{E}_{U \sim \mathbb{W}_d}[\rho_\lambda(U)] = 0, \quad (63)$$

for all nontrivial irreps indexed by  $\lambda \in \Lambda_{\leq}(d, t)$ . Note that the average over  $U \sim \mathbb{W}_d$  consists of the independent averages over all  $\mathbb{W}_{d_1 \oplus d - d_1}$ , further consisting of those over the strong unitary  $t$ -designs  $\mathbb{U}_{\leq t}(d_1)$  and  $\mathbb{U}_{\leq t}(d - d_1)$ .

Let us first fix a nontrivial irrep  $\lambda \in \Lambda_{\leq}(d, t)$  and consider  $\mathbb{W}_\lambda$  defined by

$$\mathbb{W}_\lambda := \mathbb{E}_{U \sim \mathbb{W}_{d_1 \oplus d - d_1}}[\rho_\lambda(U)]. \quad (64)$$

As we consider only irreps  $\lambda \in \Lambda_{\leq}(d, t)$ , this average can be replaced with the averages over the product  $\mathbb{H}(d_1) \times \mathbb{H}(d - d_1)$  of the Haar measures on  $\mathbb{K} := \mathbb{U}(d_1) \times \mathbb{U}(d - d_1)$ . That is,

$$\mathbb{W}_\lambda = \mathbb{E}_{U \sim \mathbb{H}(d_1) \times \mathbb{H}(d - d_1)}[\rho_\lambda(U)]. \quad (65)$$

To investigate  $\mathbb{W}_\lambda$ , we consider the irreps of  $\mathbb{K}$ . As  $\mathbb{K}$  is a subgroup of  $\mathbb{U}(d)$ , each irreducible space  $V_\lambda$  of  $\mathbb{U}(d)$  is decomposed into a direct sum of those of irreps of  $\mathbb{K}$ . For the same reason as in Definition 3, every nontrivial irrep of  $\mathbb{K}$  becomes zero by taking the average over  $\mathbb{H}(d_1) \times \mathbb{H}(d - d_1)$ . Hence, if the nontrivial irrep space  $V_\lambda$  of  $\mathbb{U}(d)$  does not contain trivial irreps of  $\mathbb{K}$ ,  $\mathbb{W}_\lambda = 0$ . In contrast, if a nontrivial irrep  $\lambda$  of  $\mathbb{U}(d)$  contains trivial irreps of  $\mathbb{K}$ , then the matrix elements of  $\mathbb{W}_\lambda$  corresponding to the trivial irreps of  $\mathbb{K}$  are one, and the others are zero.

Trivial irreps of  $\mathbb{K}$  in a nontrivial irrep  $\lambda$  of  $\mathbb{U}(d)$  were studied in a great detail because  $(\mathbb{U}(d), \mathbb{K})$  is an example of a Gelfand pair [97,98]. It is known that the irreps of  $\mathbb{U}(d)$  indexed by  $\lambda \in \Lambda_{\text{sph}}(d_1, d, t)$  contains only one trivial irrep of  $\mathbb{K}$ , and that other irreps of  $\mathbb{U}(d)$  contain no trivial irrep of  $\mathbb{K}$  [80]. As trivial irreps are one-dimensional, we denote by  $|w_\lambda\rangle \in V_\lambda$  a unit vector that spans the trivial irrep of  $\mathbb{K}$  in the spherical representation  $\lambda \in \Lambda_{\text{sph}}(d_1, d, t)$  of  $\mathbb{U}(d)$ . Then, we have

$$\mathbb{W}_\lambda = \begin{cases} 0 & \text{if } \lambda \notin \Lambda_{\text{sph}}(d_1, d, t), \\ |w_\lambda\rangle\langle w_\lambda| & \text{if } \lambda \in \Lambda_{\text{sph}}(d_1, d, t). \end{cases} \quad (66)$$

If  $\lambda \notin \Lambda_{\text{sph}}(d_1, d, t)$ , we immediately obtain  $\mathbb{E}_{U \sim \mathbb{W}_{d_1 \oplus d - d_1}}[\rho_\lambda(U)] = 0$  from the definition of  $\mathbb{W}_\lambda$ , which implies Eq. (63).

If  $\lambda \in \Lambda_{\text{sph}}(d_1, d, t)$ , we define a matrix  $M_\lambda(U)$  on  $V_\lambda$  ( $U \in \mathbb{U}(d)$ ) by

$$M_\lambda(U) := \mathbb{E}_{V_1, V_2 \sim \mathbb{H}(d_1) \times \mathbb{H}(d - d_1)}[\rho_\lambda(V_1 U V_2)]. \quad (67)$$

Importantly, for any  $\lambda \in \Lambda_{\text{sph}}(d_1, d, t)$ , there exists at least one  $R_\lambda \in \mathbb{U}(d)$  such that  $\langle w_\lambda | M_\lambda(R_\lambda) | w_\lambda \rangle = 0$ . This follows from the fact that

$$\begin{aligned} \mathbb{E}_{U \sim \mathbb{H}(d)}[\langle w_\lambda | M_\lambda(U) | w_\lambda \rangle] &= \langle w_\lambda | \mathbb{E}_{U \sim \mathbb{H}(d)}[\rho_\lambda(U)] | w_\lambda \rangle \\ &= 0, \end{aligned} \quad (68)$$

$$= 0, \quad (69)$$

where we have used the unitary invariance of  $\mathbb{H}(d)$  and that the irrep  $\rho_\lambda$  is nontrivial, so that  $\mathbb{E}_{U \sim \mathbb{H}(d)}[\rho_\lambda(U)] = 0$ . Owing to the intermediate value theorem,

there always exists at least one unitary  $R_\lambda \in \mathbf{U}(d)$  such that  $\langle w_\lambda | M_\lambda(R_\lambda) | w_\lambda \rangle = 0$ .

Using such  $R_\lambda \in \mathbf{U}(d)$  and Eq. (66), it is straightforward to observe that  $W_\lambda M_\lambda(R_\lambda) W_\lambda = 0$ . Furthermore, it follows that

$$W_\lambda M_\lambda(R_\lambda) W_\lambda = \mathbb{E}_{U, U', V_1, V_2 \sim \mathbf{W}_{d_1 \oplus d - d_1}} [\rho_\lambda(UV_1 R_\lambda V_2 U')] \quad (70)$$

$$= \mathbb{E}_{W_1, W_2 \sim \mathbf{W}_{d_1 \oplus d - d_1}} [\rho_\lambda(W_1 R_\lambda W_2)] \quad (71)$$

$$= \mathbb{E}_{U \sim \mathbf{W}_{d_1 \oplus d - d_1} R_\lambda \mathbf{W}_{d_1 \oplus d - d_1}} [\rho_\lambda(U)]. \quad (72)$$

We, hence, obtain

$$\mathbb{E}_{U \sim \mathbf{W}_{d_1 \oplus d - d_1} R_\lambda \mathbf{W}_{d_1 \oplus d - d_1}} [\rho_\lambda(U)] = 0. \quad (73)$$

Thus, the finite set of unitaries  $\mathbf{W}_{d_1 \oplus d - d_1} R_\lambda \mathbf{W}_{d_1 \oplus d - d_1}$  satisfies the condition for the design, that is, Eq. (63) for any nontrivial irrep  $\lambda \in \Lambda_{\leq}(d, t)$ , which leads to the statement that the set of unitaries defined by

$$\mathbf{W}_d = \mathbf{W}_{d_1 \oplus d - d_1} \prod_{\lambda \in \Lambda_{\text{sph}}(d_1, d, t)} (R_\lambda \mathbf{W}_{d_1 \oplus d - d_1}), \quad (74)$$

is a strong  $t$ -design on  $\mathbf{U}(d)$ .

Finally, let us clarify the relation between  $R_\lambda$  and the zero of the zonal spherical function. To this end, we first observe that the matrix element  $\langle w_\lambda | M_\lambda(U) | w_\lambda \rangle$  of  $M_\lambda$  is the zonal spherical function  $Z_\lambda(U)$ . This can be checked by a simple calculation: for any  $W_1, W_2 \in K$  and  $U \in \mathbf{U}(d)$ , we have

$$Z_\lambda(W_1 U W_2) = \langle w_\lambda | \mathbb{E}[\rho_\lambda(V_1 W_1 U W_2 V_2)] | w_\lambda \rangle \quad (75)$$

$$= \langle w_\lambda | \mathbb{E}[\rho_\lambda(V_1 U V_2)] | w_\lambda \rangle \quad (76)$$

$$= Z_\lambda(U), \quad (77)$$

where the averages are all taken over  $V_1, V_2 \sim \mathbf{H}(d_1) \times \mathbf{H}(d - d_1)$ . Thus,  $Z_\lambda(U)$  is bi- $K$ -invariant and, thus, it is the zonal spherical function. This implies that  $R_\lambda$  is indeed a zero of the zonal spherical function.

Based on this fact, we can provide a matrix form of  $R_\lambda$  in the fixed basis in which a unitary in  $\mathbf{W}_{d_1 \oplus d - d_1}$  is represented as  $U \oplus V$ . To this end, it is important to note that the bi- $K$ -invariance of the zonal spherical function implies that it is characterized by the cosets of  $\mathbf{K} = \mathbf{U}(d_1) \times \mathbf{U}(d - d_1)$  in  $\mathbf{U}(d)$ . The cosets can be further identified with  $d_1$ -dimensional subspaces corresponding to the support on which  $\mathbf{U}(d_1)$  acts. For instance, the identity element in the coset of  $\mathbf{K}$  corresponds to the subspace  $V_0$  spanned by the first  $d_1$  vectors of the fixed basis. The matrix form of  $R_\lambda$  is obtained by specifying the relation between  $V_0$  and the subspace corresponding to another representative of the coset.

To characterize the relation between two subspaces, we use the *principal* angles. For two subspaces  $X$  and  $Y$ , let us refer to  $\theta = \min \arg \cos |\langle x | y \rangle|$ , where the minimum is taken over all unit vectors  $|x\rangle \in X, |y\rangle \in Y$ , as the minimum angle between  $X$  and  $Y$ . The principal angles  $(\theta_0, \dots, \theta_{m-1})$  between two  $m$ -dimensional subspaces  $X$  and  $Y$  are then defined as follows:  $\theta_0$  is the minimum angle between  $X$  and  $Y$ , and  $\theta_{i+1}$  is the minimum angle between  $X \cap \text{span}\{|x_0\rangle, \dots, |x_i\rangle\}$  and  $Y \cap \text{span}\{|y_0\rangle, \dots, |y_i\rangle\}$ , where  $(|x_j\rangle, |y_j\rangle)$  is a pair of the unit vectors that leads to  $\theta_j$ .

The cosine of the principal angles between  $V_0$  and the subspace corresponding to another representative in the coset determines the value of the zonal spherical function  $Z_\lambda$ , and so,  $Z_\lambda$  can be written as  $Z_\lambda(\cos^2 \theta_0, \dots, \cos^2 \theta_{d_1-1})$  [50,77]. See, for example, Refs. [50,77,81] for the explicit form of  $Z_\lambda$  as a polynomial of  $(\cos^2 \theta_1, \dots, \cos^2 \theta_{d_1})$ .

By solving the polynomial, we obtain the principal angles  $(\theta_\lambda^{(0)}, \dots, \theta_\lambda^{(d_1-1)})$  between  $V_0$  and the subspace corresponding to the zero of  $Z_\lambda$ . Recalling the definition of the principal angles and using the left- and right-invariance of the coset by any unitary in  $\mathbf{K}$ , we can take a matrix form of  $R_\lambda$  as follows:

$$R_\lambda = \begin{pmatrix} C(\boldsymbol{\theta}_\lambda) & iS(\boldsymbol{\theta}_\lambda) & 0 \\ iS(\boldsymbol{\theta}_\lambda) & C(\boldsymbol{\theta}_\lambda) & 0 \\ 0 & 0 & I_{d-2d_1} \end{pmatrix}, \quad (78)$$

where  $C(\boldsymbol{\theta}_\lambda) = \text{diag}(\cos \theta_\lambda^{(0)}, \dots, \cos \theta_\lambda^{(d_1-1)})$  and  $S(\boldsymbol{\theta}_\lambda) = \text{diag}(\sin \theta_\lambda^{(0)}, \dots, \sin \theta_\lambda^{(d_1-1)})$ , and  $I_{d-2d_1}$  is the identity matrix of size  $d - 2d_1$ . Note that  $R_\lambda$  is not necessarily in this form because the coset is invariant under the action of  $\mathbf{K}$ . ■

## VIII. HIGHER-ORDER RB

In this section, we investigate the higher-order RB in detail. We begin with a preliminary in Sec. VIII A and explain several basic properties of the self-adjointness in Sec. VIII B. We consider the  $t$ -RB for general  $t$  and the 2-RB in Secs. C and VIII E, respectively.

### A. Liouville representation

Let  $\sigma_0 = I/\sqrt{2}$ ,  $\sigma_1 = X/\sqrt{2}$ ,  $\sigma_2 = Y/\sqrt{2}$ , and  $\sigma_3 = Z/\sqrt{2}$  be normalized Pauli operators on one qubit, where normalization is in terms of the Hilbert-Schmidt inner product. For  $q$  qubits, we introduce a vector  $\vec{n} = (n_1, n_2, \dots, n_q)$  ( $n_i \in \{0, 1, 2, 3\}$ ) and use the notation that

$$\sigma_{\vec{n}} := \sigma_{n_1} \otimes \dots \otimes \sigma_{n_q}. \quad (79)$$

We also denote  $2^q$  by  $d$  in this section.

The Liouville representation is a matrix representation of quantum channels, also known as the Pauli transfer matrix. See, for example, Refs. [84,89,99]. Let  $|\cdot\rangle\rangle$

be a linear map from a set of all linear operators on a  $d$ -dimensional Hilbert space to a  $d^2$ -dimensional vector space that specifically maps  $\sigma_{\vec{n}}$  to a canonical orthonormal basis vector  $e_{\vec{n}}$ . As the map is linear, we have

$$|A\rangle := \sum_{\vec{n}} \text{Tr}[\sigma_{\vec{n}} A] |\sigma_{\vec{n}}\rangle, \quad (80)$$

for any linear operator  $A$ . Note that  $\langle\langle A|B\rangle\rangle = \text{Tr}[A^\dagger B]$ .

Based on this vector representation of linear operators, a linear supermap  $\mathcal{E}$  can be represented by a matrix. The Liouville representation of a linear supermap  $\mathcal{E}$  is defined by

$$L_{\mathcal{E}} := \sum_{\vec{n}} |\mathcal{E}(\sigma_{\vec{n}})\rangle\langle\langle\sigma_{\vec{n}}|, \quad (81)$$

which is a regular matrix of size  $d^2$ . The matrix element in the canonical basis of  $\{|\sigma_{\vec{n}}\rangle\}_{\vec{n}}$  is given by

$$(L_{\mathcal{E}})_{\vec{n}\vec{m}} = \langle\langle\sigma_{\vec{n}}|\mathcal{E}(\sigma_{\vec{m}})\rangle\rangle = \text{Tr}[\sigma_{\vec{n}} \mathcal{E}(\sigma_{\vec{m}})]. \quad (82)$$

The vector and Liouville representations satisfy the following properties:

1.  $L_{\mathcal{E}}|\rho\rangle\rangle = |\mathcal{E}(\rho)\rangle\rangle$ ;
2.  $L_{\mathcal{E}_2 \circ \mathcal{E}_1} = L_{\mathcal{E}_2} L_{\mathcal{E}_1}$ ;
3.  $L_{\alpha \mathcal{E}_1 + \beta \mathcal{E}_2} = \alpha L_{\mathcal{E}_1} + \beta L_{\mathcal{E}_2}$  ( $\alpha, \beta \in \mathbb{C}$ );
4.  $L_{\mathcal{E}_1 \otimes \mathcal{E}_2} = L_{\mathcal{E}_1} \otimes L_{\mathcal{E}_2}$ ;
5.  $L_{\mathcal{E}^\dagger} = L_{\mathcal{E}}^\dagger$ .

Properties of a linear supermap  $\mathcal{E}$  can be also expressed in terms of the Liouville representation. For instance, the linear map  $\mathcal{E}$  is trace-preserving if and only if  $(L_{\mathcal{E}})_{\vec{0}\vec{0}} = 1$  and  $(L_{\mathcal{E}})_{\vec{0}\vec{n}} = 0$  for any  $\vec{n} \neq \vec{0}$ . As we are interested in the CPTP map  $\mathcal{E}$  that represents a noise, its Liouville representation is always in the form of

$$L_{\mathcal{E}} = \begin{pmatrix} 1 & \mathbf{0} \\ \alpha_{\mathcal{E}} & \tilde{L}_{\mathcal{E}} \end{pmatrix}, \quad (83)$$

where  $\mathbf{0}$  is a row vector of length  $d^2 - 1$  with all elements being zero,  $\alpha_{\mathcal{E}}$  is a column vector of length  $d^2 - 1$ , called a *nonunital part* of the noise, and  $\tilde{L}_{\mathcal{E}}$  is a  $(d^2 - 1) \times (d^2 - 1)$  matrix. The nonunital part  $\alpha_{\mathcal{E}}$  of the noise is the zero vector if and only if the map  $\mathcal{E}$  is unital, that is,  $\mathcal{E}(I) = I$  with  $I$  being the identity operator.

In the Liouville representation, the fidelity parameter  $f(\mathcal{E})$  and the unitarity  $u(\mathcal{E})$  of a noisy CPTP map  $\mathcal{E}$  are given by

$$f(\mathcal{E}) = \frac{1}{d^2 - 1} \sum_{\vec{n} \neq \vec{0}} \langle\langle\sigma_{\vec{n}}|L_{\mathcal{E}}|\sigma_{\vec{n}}\rangle\rangle \quad (84)$$

$$= \frac{1}{d^2 - 1} \text{Tr}[\tilde{L}_{\mathcal{E}}], \quad (85)$$

$$u(\mathcal{E}) = \frac{1}{d^2 - 1} \sum_{\vec{n} \neq \vec{0}} \langle\langle\sigma_{\vec{n}}|L_{\mathcal{E}}^\dagger L_{\mathcal{E}}|\sigma_{\vec{n}}\rangle\rangle, \quad (86)$$

$$= \frac{1}{d^2 - 1} \text{Tr}[\tilde{L}_{\mathcal{E}}^\dagger \tilde{L}_{\mathcal{E}}], \quad (87)$$

respectively.

### B. Properties of the self-adjointness

For a CPTP map  $\mathcal{E}$ , the self-adjointness  $H(\mathcal{E})$  and the self-adjointness parameter  $h(\mathcal{E})$  are defined by

$$H(\mathcal{E}) := 1 - \frac{d+1}{2d} \int d\varphi \|\mathcal{E}(|\varphi\rangle\langle\varphi|) - \mathcal{E}^\dagger(|\varphi\rangle\langle\varphi|)\|_2^2, \quad (88)$$

$$h(\mathcal{E}) := \frac{d}{d-1} \int d\varphi \text{Tr}[\mathcal{E}'(\varphi)\mathcal{E}'^\dagger(\varphi)], \quad (89)$$

$$= \frac{1}{d^2 - 1} \text{Tr}[\tilde{L}_{\mathcal{E}}^2], \quad (90)$$

where  $\mathcal{E}'(\rho) = \mathcal{E}(\rho - I/d)$ , and the last line is shown in Appendix D.

We first show the relation between  $H(\mathcal{E})$  and  $h(\mathcal{E})$ , that is, Eq. (28) in Sec. IV A:

$$H(\mathcal{E}) = 1 - \frac{d^2 - 1}{d^2} (u(\mathcal{E}) - h(\mathcal{E})) - \frac{d+1}{2d^2} |\alpha_{\mathcal{E}}|^2. \quad (91)$$

From the definition of  $H(\mathcal{E})$ , we have

$$\begin{aligned} & \frac{2d}{d+1} (1 - H(\mathcal{E})) \\ &= \int d\varphi \left[ \text{Tr}[\mathcal{E}(|\varphi\rangle\langle\varphi|)^2] + \text{Tr}[\mathcal{E}^\dagger(|\varphi\rangle\langle\varphi|)^2] \right. \\ & \quad \left. - 2 \text{Tr}[\mathcal{E}(|\varphi\rangle\langle\varphi|)\mathcal{E}^\dagger(|\varphi\rangle\langle\varphi|)] \right]. \quad (92) \end{aligned}$$

By rewriting  $\mathcal{E}$  with  $\mathcal{E}'$ , the first term on the right-hand side is expressed in terms of the unitarity  $u(\mathcal{E})$ , such as

$$\int d\varphi \text{Tr}[\mathcal{E}(|\varphi\rangle\langle\varphi|)^2] = \frac{d-1}{d} u(\mathcal{E}) + \text{Tr}[\mathcal{E}(I/d)^2]. \quad (93)$$

By using the swap operator  $\mathbb{F} := \sum_{\vec{n}} \sigma_{\vec{n}} \otimes \sigma_{\vec{n}}$ , and the property that  $\text{Tr}[MN] = \text{Tr}[\mathbb{F}(M \otimes N)]$  for any matrices

$M$  and  $N$ , which is called a *swap trick*, it follows that

$$\mathrm{Tr}[\mathcal{E}(I/d)^2] = \mathrm{Tr}[\mathbb{F}\mathcal{E}(I/d)^{\otimes 2}] \quad (94)$$

$$= \frac{1}{d} \sum_{\bar{n}} \langle\langle \sigma_{\bar{n}} | L_{\mathcal{E}} | \sigma_{\bar{0}} \rangle\rangle^2 \quad (95)$$

$$= \frac{1}{d} |\alpha_{\mathcal{E}}|^2 + \frac{1}{d}, \quad (96)$$

which leads to

$$\int d\varphi \mathrm{Tr}[\mathcal{E}(|\varphi\rangle\langle\varphi|)^2] = \frac{d-1}{d} u(\mathcal{E}) + \frac{1}{d} |\alpha_{\mathcal{E}}|^2 + \frac{1}{d}. \quad (97)$$

Similarly, we obtain

$$\int d\varphi \mathrm{Tr}[\mathcal{E}^\dagger(|\varphi\rangle\langle\varphi|)^2] = \frac{d-1}{d} u(\mathcal{E}) + \frac{1}{d}, \quad (98)$$

from the facts that  $L_{\mathcal{E}^\dagger} = L_{\mathcal{E}}^\dagger$  and that  $|\alpha_{\mathcal{E}^\dagger}| = 0$  for any trace-preserving map  $\mathcal{E}$ .

From the definition of  $\mathcal{E}'$ , it is straightforward to show that the self-adjointness parameter  $h(\mathcal{E})$  is given by

$$h(\mathcal{E}) = \frac{1}{d-1} \left[ d \int \mathrm{Tr}[\mathcal{E}(\varphi)\mathcal{E}^\dagger(\varphi)] d\varphi - 1 \right]. \quad (99)$$

Combining these altogether, we arrive at

$$\frac{2d}{d+1} (1 - H(\mathcal{E})) = \frac{2(d-1)}{d} (u(\mathcal{E}) - h(\mathcal{E})) + \frac{1}{d} |\alpha_{\mathcal{E}}|^2, \quad (100)$$

implying Eq. (91).

The self-adjointness parameter also satisfies the following properties (they are all shown in Appendix D):

1.  $-\frac{1}{d^2-1} \leq h(\mathcal{E}) \leq u(\mathcal{E})$ ;
2.  $h(\mathcal{E}) = u(\mathcal{E})$  if and only if  $\tilde{L}_{\mathcal{E}} = \tilde{L}_{\mathcal{E}}^\dagger$ ; for a unital noise  $\mathcal{E}$ ,  $h(\mathcal{E}) = u(\mathcal{E})$  if and only if the noise is self-adjoint ( $\mathcal{E} = \mathcal{E}^\dagger$ );
3.  $h(\mathcal{E}) = -\frac{1}{d^2-1}$  if and only if  $\mathrm{Tr}[K_i K_j] = 0$  for any  $i, j$ , where  $\{K_i\}$  are the Kraus operators of  $\mathcal{E}$ ;
4. the average gate fidelity  $F(\mathcal{E})$  is bounded from above by  $u(\mathcal{E})$  and  $h(\mathcal{E})$ ,

$$F(\mathcal{E}) \leq \frac{d-1}{d} \sqrt{\frac{h(\mathcal{E}) + u(\mathcal{E})}{2}} + \frac{1}{d}. \quad (101)$$

### C. A general expression for the $t$ -RB

We show here that the expectation value  $V^{(t)}(m, \mathcal{E} | O_{\mathrm{ini}}, O_{\mathrm{meas}})$  in the  $t$ -RB has a general form of

$$V^{(t)}(m, \mathcal{E} | O_{\mathrm{ini}}, O_{\mathrm{meas}}) = \sum_{\lambda} \mathrm{Tr}[\hat{A}_{\lambda}(\hat{C}_{\lambda}(\mathcal{E}))^m], \quad (102)$$

where  $\hat{A}_{\lambda}$  is a regular matrix depending on  $O_{\mathrm{ini}}$  and  $\mathcal{E}(O_{\mathrm{meas}})$ , and  $\hat{C}_{\lambda}(\mathcal{E})$  is a regular matrix depending only

on  $\mathcal{E}$ . As we show in the following,  $\lambda$  labels the irreps of a  $t$ -copy representation of the unitary group, and the size of the matrices is equal to the multiplicity of each irrep.

The expectation value is defined by

$$V^{(t)}(m, \mathcal{E} | O_{\mathrm{ini}}, O_{\mathrm{meas}}) := \mathbb{E}_{U_i} \left[ \left( \mathrm{Tr} [O_{\mathrm{meas}} \mathcal{G}_{i_{m+1}} \circ \mathcal{G}_{i_m} \circ \cdots \circ \mathcal{G}_{i_1} (O_{\mathrm{ini}})] \right)^t \right], \quad (103)$$

where  $\mathbb{E}_{U_i}$  is the average over all unitary sequences  $U_i \sim \mathbf{U}_{2t}^{\times m}$ . Note that  $\mathcal{G}_i = \mathcal{E} \circ U_i$  and that  $U_i$  is the unitary channel defined by  $U_i(\rho) = U_i \rho U_i^\dagger$ . In terms of the Liouville representation, we have

$$\mathrm{Tr} [O_{\mathrm{meas}} \mathcal{G}_{i_{m+1}} \circ \mathcal{G}_{i_m} \circ \cdots \circ \mathcal{G}_{i_1} (O_{\mathrm{ini}})] = \langle\langle O'_{\mathrm{meas}} | L_{U'_m \circ \mathcal{E} \circ U'_m} \cdots L_{U'_1 \circ \mathcal{E} \circ U'_1} | O_{\mathrm{ini}} \rangle\rangle, \quad (104)$$

where we have used that  $\mathcal{G}_i = \mathcal{E} \circ U_i$ ,  $U'_n = U_n \circ U_{n-1} \circ \cdots \circ U_2 \circ U_1$ , and  $O'_{\mathrm{meas}} = \mathcal{E}(O_{\mathrm{meas}})$ .

Noting the  $t$ th power and the fact that each unitary is independently chosen from a unitary  $2t$ -design  $\mathbf{U}_{2t}$ , we obtain

$$V^{(t)}(m, \mathcal{E} | O_{\mathrm{ini}}, O_{\mathrm{meas}}) = \left( \langle\langle O'_{\mathrm{meas}} |^{\otimes t} \rangle\rangle (L_{\mathrm{av}})^m (|O_{\mathrm{ini}}\rangle\rangle^{\otimes t} \right), \quad (105)$$

where  $L_{\mathrm{av}}$  is defined by

$$L_{\mathrm{av}} := \mathbb{E}_{U \sim \mathbf{U}_{2t}} [(L_{U \circ \mathcal{E} \circ U^\dagger})^{\otimes t}], \quad (106)$$

$$= \mathbb{E}_{U \sim \mathbf{U}_{2t}} [(L_U L_{\mathcal{E}} L_U^\dagger)^{\otimes t}] \quad (107)$$

$$= \mathbb{E}_{U \sim \mathbf{H}} [(L_U L_{\mathcal{E}} L_U^\dagger)^{\otimes t}]. \quad (108)$$

The last line follows because  $\mathbf{U}_{2t}$  is an exact unitary  $2t$ -design.

To write down  $L_{\mathrm{av}}$  explicitly, let us consider the tensor- $t$  Liouville representation given by

$$\mathbf{U}(d) \ni U \rightarrow L_U^{\otimes t} \in GL(\mathcal{K}), \quad (109)$$

where  $GL(\mathcal{K})$  is the general linear group acting on the  $d^q$ -dimensional vector space  $\mathcal{K}$  defined by

$$\mathcal{K} := \mathrm{span} \left\{ \bigotimes_{s=1}^t |\sigma_{\vec{n}_s}\rangle\rangle : \vec{n}_s \in \{0, 1, 2, 3\}^q, s \in [1, t] \right\}. \quad (110)$$

We denote the irreducible decomposition by

$$\mathcal{K} = \bigoplus_{\lambda} \mathcal{K}_{\lambda}^{\oplus m_{\lambda}}, \quad (111)$$

where  $\lambda$  labels the irreps and  $m_{\lambda}$  is the multiplicity of the irrep labeled by  $\lambda$ .



The key observation is that

$$\forall V \in \mathbf{U}(d), \quad [L_{\text{av}}, L_{V^{\otimes t}}] = 0, \quad (112)$$

which simply follows from the unitary invariance of the Haar measure. This implies that  $L_{\text{av}} \in \text{End}_{\mathbf{U}}(\mathcal{K})$ , where  $\text{End}_{\mathbf{U}}(\mathcal{K})$  is a set of all endomorphisms of  $\mathcal{K}$  that commute with the tensor- $t$  Liouville action of  $\mathbf{U}(d)$ . It is well-known that  $\text{End}_{\mathbf{U}}(\mathcal{K})$  is isomorphic to the direct sum of matrix algebras:

$$\text{End}_{\mathbf{U}}(\mathcal{K}) \simeq \bigoplus_{\lambda} M(m_{\lambda}, \mathbb{C}), \quad (113)$$

where  $M(m_{\lambda}, \mathbb{C})$  is a set of all  $m_{\lambda} \times m_{\lambda}$  matrices over  $\mathbb{C}$ . Thus, the operator  $L_{\text{av}} \in \text{End}_{\mathbf{U}}(\mathcal{K})$  can be represented by a direct sum of matrices.

To obtain the explicit form of  $L_{\text{av}}$ , let  $\mathcal{K}_{\lambda}^{(1)} \oplus \dots \oplus \mathcal{K}_{\lambda}^{(m_{\lambda})}$  be a fixed decomposition of  $\mathcal{K}_{\lambda}^{\oplus m_{\lambda}}$ , and denote  $\eta_{\lambda}^{p \rightarrow q}$  be the isomorphism from  $\mathcal{K}_{\lambda}^{(p)}$  to  $\mathcal{K}_{\lambda}^{(q)}$ . We also denote by  $\Pi_{\lambda}^{(p)}$  the projection onto  $\mathcal{K}_{\lambda}^{(p)}$ . Then, from the explicit form of the isomorphism, we have

$$L_{\text{av}} = \sum_{\lambda} \sum_{p,q=1}^{m_{\lambda}} (\hat{C}_{\lambda}(\mathcal{E}))_{pq} \eta_{\lambda}^{p \rightarrow q} \Pi_{\lambda}^{(p)}, \quad (114)$$

where  $\hat{C}_{\lambda}(\mathcal{E}) \in M(m_{\lambda}, \mathbb{C})$ . Each element  $\hat{C}_{\lambda}(\mathcal{E})$  is given by

$$(\hat{C}_{\lambda}(\mathcal{E}))_{pq} = \text{Tr}[L_{\text{av}} \eta_{\lambda}^{q \rightarrow p} \Pi_{\lambda}^{(q)}], \quad (115)$$

$$= \text{Tr}[\mathbb{E}_{U \sim \mathbf{H}}[(L_{\mathcal{U}} L_{\mathcal{E}} L_{\mathcal{U}^{\dagger}})^{\otimes t}] \eta_{\lambda}^{q \rightarrow p} \Pi_{\lambda}^{(q)}], \quad (116)$$

$$= \mathbb{E}_{U \sim \mathbf{H}} \text{Tr}[L_{\mathcal{E}}^{\otimes t} L_{\mathcal{U}^{\dagger}}^{\otimes t} \eta_{\lambda}^{q \rightarrow p} \Pi_{\lambda}^{(q)} L_{\mathcal{U}}^{\otimes t}], \quad (117)$$

$$= \mathbb{E}_{U \sim \mathbf{H}} \text{Tr}[L_{\mathcal{E}}^{\otimes t} \eta_{\lambda}^{q \rightarrow p} \Pi_{\lambda}^{(q)}], \quad (118)$$

$$= \text{Tr}[L_{\mathcal{E}}^{\otimes t} \eta_{\lambda}^{q \rightarrow p} \Pi_{\lambda}^{(q)}], \quad (119)$$

where we have used the irreducibility in the fourth line.

Consequently, it follows that

$$(L_{\text{av}})^m = \sum_{\lambda} \sum_{p,q=1}^{m_{\lambda}} (\hat{C}_{\lambda}(\mathcal{E}))_{pq}^m \eta_{\lambda}^{p \rightarrow q} \Pi_{\lambda}^{(p)}. \quad (120)$$

Substituting this into Eq. (105), we obtain

$$\mathcal{V}^{(t)}(m, \mathcal{E} | \mathcal{O}_{\text{ini}}, \mathcal{O}_{\text{meas}}) = \sum_{\lambda} \text{Tr}[\hat{A}_{\lambda} (\hat{C}_{\lambda}(\mathcal{E}))^m], \quad (121)$$

where the  $m_{\lambda} \times m_{\lambda}$  matrices  $\hat{A}_{\lambda}$  are given by

$$(\hat{A}_{\lambda})_{pq} = \langle \langle \mathcal{E}(\mathcal{O}_{\text{meas}})^{\otimes t} | \eta_{\lambda}^{q \rightarrow p} \Pi_{\lambda}^{(q)} | \mathcal{O}_{\text{ini}}^{\otimes t} \rangle \rangle. \quad (122)$$

This completes the proof. ■

## D. The first-order RB

Let us briefly overview the 1-RB using an exact unitary 2-design, namely, the standard RB. We also explain how the result changes when the 2-design is an approximate one rather than the exact one.

In the 1-RB, the representation space is given by

$$\mathcal{K} = \text{span}\{|\sigma_{\vec{n}}\rangle\} : \vec{n} \in \{0, 1, 2, 3\}^q. \quad (123)$$

We need to find a irreducible decomposition of  $\mathcal{K}$  under the action of a unitary group  $\mathbf{U}(d)$  as  $U \rightarrow L_{\mathcal{U}}$ . The Liouville representation  $L_{\mathcal{U}}$  is defined by  $L_{\mathcal{U}}|\rho\rangle\rangle = |U\rho U^{\dagger}\rangle\rangle$ . Hence,  $\mathcal{K}$  is irreducibly decomposed to

$$\mathcal{K} = \mathcal{K}_0 \oplus \mathcal{K}_1, \quad (124)$$

where

$$\mathcal{K}_0 = \text{span}\{|\sigma_{\vec{0}}\rangle\rangle\}, \quad (125)$$

$$\mathcal{K}_1 = \text{span}\{|\sigma_{\vec{n}}\rangle\rangle : \vec{n} \in \{0, 1, 2, 3\}^q, \vec{n} \neq \vec{0}\}. \quad (126)$$

Denoting by  $\Pi_0$  and  $\Pi_1$  projectors onto  $\mathcal{K}_0$  and  $\mathcal{K}_1$ , respectively, we have

$$L_{\text{av}} := \mathbb{E}_{U \sim \mathbf{U}_2}[(L_{\mathcal{U} \circ \mathcal{E} \circ \mathcal{U}^{\dagger}})], \quad (127)$$

$$= \Pi_0 + f(\mathcal{E})\Pi_1, \quad (128)$$

where  $f(\mathcal{E})$  is the fidelity parameter. Note that  $\mathbf{U}_2$  is an exact unitary 2-design. We thus obtain that

$$\mathcal{V}^{(1)}(m, \mathcal{E} | \Delta, \mathcal{O}_{\text{meas}}) = A_0 + A_1 f(\mathcal{E})^m, \quad (129)$$

where  $A_i = \langle \langle \mathcal{E}(\mathcal{O}_{\text{meas}}) | \Pi_i | \mathcal{O}_{\text{ini}} \rangle \rangle$  for  $i = 0, 1$ .

When the 2-design is an approximate one  $\mathbf{U}_2^{(\epsilon)}$ , Eq. (128) holds only approximately. The degree of approximation depends on how we measure it, but we here assume that the design is  $\epsilon$ -approximate when Eq. (128) holds up to  $\epsilon$ -approximation. That is, we assume that

$$L_{\text{av}}^{(\epsilon)} := \mathbb{E}_{U \sim \mathbf{U}_2^{(\epsilon)}}[(L_{\mathcal{U} \circ \mathcal{E} \circ \mathcal{U}^{\dagger}})], \quad (130)$$

$$= L_{\text{av}} + \epsilon \Delta, \quad (131)$$

where  $\Delta$  is some operator of  $O(1)$ . Note that standard definitions of approximate designs require harder criteria (see,

e.g., Ref. [78]). In this case, instead of Eq. (129), we have

$$\begin{aligned} V^{(1)}(m, \mathcal{E}|\Delta, O_{\text{meas}}) \\ = A_0 + E + A_1 f^m + \epsilon(m-2)(E_2 f^2 + E_1 f + E_0) \\ + O(m^2 \epsilon^2), \end{aligned} \quad (132)$$

where

$$E_2 = \langle\langle \mathcal{E}(O_{\text{meas}}) | \Pi_0 \Delta \Pi_0 | O_{\text{ini}} \rangle\rangle, \quad (133)$$

$$E_1 = \langle\langle \mathcal{E}(O_{\text{meas}}) | (\Pi_0 \Delta \Pi_1 + \Pi_1 \Delta \Pi_0) | O_{\text{ini}} \rangle\rangle, \quad (134)$$

$$E_0 = \langle\langle \mathcal{E}(O_{\text{meas}}) | \Pi_1 \Delta \Pi_1 | O_{\text{ini}} \rangle\rangle, \quad (135)$$

$$E = \langle\langle \mathcal{E}(O_{\text{meas}}) | \{\Pi_0, \Delta\} + f \{\Pi_1, \Delta\} | O_{\text{ini}} \rangle\rangle. \quad (136)$$

Comparing Eqs. (129) and (132), we observe that using approximate unitary 2-designs result in more complicated form or the fitting function.

### E. The second-order RB

We now focus on the 2-RB. Although the representation space in this case is

$$\mathcal{K} = \text{span}\{|\sigma_{\vec{n}_1 \otimes \vec{n}_2}\rangle : \vec{n}_1, \vec{n}_2 \in \{0, 1, 2, 3\}^q\}, \quad (137)$$

where we have used the notation that  $\sigma_{\vec{n}_1 \otimes \vec{n}_2} = \sigma_{\vec{n}_1} \otimes \sigma_{\vec{n}_2}$ , it is not necessary to consider the whole space because we assume that the initial operator  $\Delta$  is traceless. This, together with the fact that the noise map is trace-preserving, implies that the operator remains traceless during the whole process. We also observe that the whole process is symmetric under the exchange of the first and the second spaces, each labeled by  $\vec{n}_1$  and  $\vec{n}_2$  in Eq. (137). Hence, in the analysis of the 2-RB, the relevant space is only the *traceless symmetric* subspace defined by

$$\begin{aligned} \mathcal{K}_{\text{TS}} := \text{span}\{|\sigma_{\vec{n}_1 \otimes \vec{n}_2} + \sigma_{\vec{n}_2 \otimes \vec{n}_1}\rangle \\ : \vec{n}_1, \vec{n}_2 \in \{0, 1, 2, 3\}^q, (\vec{n}_1, \vec{n}_2) \neq (\vec{0}, \vec{0})\}, \end{aligned} \quad (138)$$

where  $\vec{0} = (0, \dots, 0)$ . The irreducible decomposition of  $\mathcal{K}_{\text{TS}}$  can be obtained by an extensive use of the result in Ref. [100] (see Appendix C), based on which we explicitly compute  $V^{(2)}(m, \mathcal{E}|\Delta, O_{\text{meas}})$ .

It turns out that the situation differs depending on whether  $q = 1$  or  $q \geq 2$ . We, hence, deal with the two cases separately.

#### 1. 2-RB in a single-qubit system

When  $q = 1$ , the irreducible decomposition of  $\mathcal{K}_{\text{TS}}$  is given by

$$\mathcal{K}_{\text{TS}} = \mathcal{K}_0 \oplus \mathcal{K}_1, \quad (139)$$

which is multiplicity-free. Here,  $\mathcal{K}_0$  and  $\mathcal{K}_1$  are

$$\mathcal{K}_0 := \text{span}\{|\sigma_1^{\otimes 2} + \sigma_2^{\otimes 2} + \sigma_3^{\otimes 2}\rangle\}, \quad (140)$$

$$\begin{aligned} \mathcal{K}_1 := \text{span}\{&|S_{1,2}\rangle, |S_{1,3}\rangle, |S_{2,3}\rangle, \\ &|\sigma_1^{\otimes 2} - 2\sigma_2^{\otimes 2} + \sigma_3^{\otimes 2}\rangle, |\sigma_1^{\otimes 2} - \sigma_3^{\otimes 2}\rangle\}, \end{aligned} \quad (141)$$

respectively, with  $S_{n,m} := (\sigma_n \otimes \sigma_m + \sigma_m \otimes \sigma_n)/\sqrt{2}$ . It is obvious that  $d_0 := \dim \mathcal{K}_0 = 1$  and  $d_1 := \dim \mathcal{K}_1 = 5$ .

This decomposition implies that the expectation  $V^{(2)}(m, \mathcal{E}|\Delta, O_{\text{meas}})$  is in the form of

$$V^{(2)}(m, \mathcal{E}|\Delta, O_{\text{meas}}) = A_0 C_0(\mathcal{E})^m + A_1 C_1(\mathcal{E})^m, \quad (142)$$

where both  $A_\lambda$  and  $C_\lambda$  are given by

$$A_\lambda = \langle\langle \mathcal{E}(O_{\text{meas}})^{\otimes 2} | \Pi_\lambda | \Delta^{\otimes 2} \rangle\rangle, \quad (143)$$

$$C_\lambda(\mathcal{E}) = \frac{\text{Tr}[\Pi_\lambda L_\mathcal{E}^{\otimes 2}]}{\text{Tr}[\Pi_\lambda]}, \quad (144)$$

with  $\Pi_\lambda$  being the projections onto  $\mathcal{K}_\lambda$ . As the projections can be constructed explicitly from Eqs. (140) and (141), we can compute  $C_\lambda(\mathcal{E})$ .

First, we have

$$C_0(\mathcal{E}) = \frac{1}{3} \sum_{n,m=1}^3 (\langle\langle \sigma_n | L_\mathcal{E} | \sigma_m \rangle\rangle)^2, \quad (145)$$

$$= \frac{1}{3} \sum_{n=1}^3 \langle\langle \sigma_n | L_\mathcal{E}^\dagger L_\mathcal{E} | \sigma_n \rangle\rangle, \quad (146)$$

$$= u(\mathcal{E}). \quad (147)$$

For  $C_1(\mathcal{E})$ , we start from the relation that

$$\Pi_1 = \Pi_{\text{sym}} - \Pi_0 - |\sigma_0^{\otimes 2}\rangle\langle\sigma_0^{\otimes 2}| - \sum_{n=1}^3 |S_{0,n}\rangle\langle S_{0,n}|, \quad (148)$$

where  $\Pi_{\text{sym}}$  is the projection onto the symmetric subspace of  $\mathcal{K}^{\otimes 2}$ . The projection  $\Pi_{\text{sym}}$  is also expressed by  $(\mathbb{I} + \mathbb{F})/2$ . Here,  $\mathbb{I}$  is the identity operator on  $\mathcal{K}^{\otimes 2}$  and  $\mathbb{F}$  is the swap operator on  $\mathcal{K}^{\otimes 2}$  defined by  $\sum_{n,m=0}^3 |\sigma_n\rangle\langle\sigma_m| \otimes |\sigma_m\rangle\langle\sigma_n|$ . Using the swap trick, we have

$$\text{Tr}[\Pi_{\text{sym}} L_\mathcal{E}^{\otimes 2}] = \frac{1}{2} (\text{Tr}[L_\mathcal{E}]^2 + \text{Tr}[L_\mathcal{E}^2]). \quad (149)$$

Moreover, from the direct calculations, we obtain

$$\text{Tr}[|\sigma_0^{\otimes 2}\rangle\langle\sigma_0^{\otimes 2}| L_\mathcal{E}^{\otimes 2}] = L_{00}^2, \quad (150)$$

$$\text{Tr}[|S_{0,n}\rangle\langle S_{0,n}| L_\mathcal{E}^{\otimes 2}] = L_{00} L_{nn} + L_{0n} L_{n0}, \quad (151)$$

where  $L_{nm} = \langle\langle \sigma_n | L_{\mathcal{E}} | \sigma_m \rangle\rangle$ . Further using the relations

$$\text{Tr}[L_{\mathcal{E}}] = L_{00} + \text{Tr}[\tilde{L}_{\mathcal{E}}], \quad (152)$$

$$\text{Tr}[L_{\mathcal{E}}^2] = L_{00}^2 + 2 \sum_{n=1}^3 L_{0n} L_{n0} + \text{Tr}[\tilde{L}_{\mathcal{E}}^2], \quad (153)$$

we obtain from Eq. (148) that

$$\text{Tr}[\Pi_I L_{\mathcal{E}}^{\otimes 2}] = \frac{1}{2} \text{Tr}[\tilde{L}_{\mathcal{E}}^2] - u(\mathcal{E}) + \frac{1}{2} \text{Tr}[\tilde{L}_{\mathcal{E}}]^2, \quad (154)$$

$$= \frac{3}{2} h(\mathcal{E}) - u(\mathcal{E}) + \frac{9}{2} f(\mathcal{E})^2. \quad (155)$$

Altogether, we obtain

$$\begin{aligned} V(m, \mathcal{E} | \Delta, O_{\text{meas}}) \\ = A_0 u(\mathcal{E})^m + A_1 \left( \frac{9}{10} f(\mathcal{E})^2 - \frac{1}{5} u(\mathcal{E}) + \frac{3}{10} h(\mathcal{E}) \right)^m. \end{aligned} \quad (156)$$

## 2. 2-RB in a multiqubit system

For a multiqubit system ( $q \geq 2$ ), the traceless symmetric subspace is decomposed into four irreducible subspaces:

$$\mathcal{K}_{\text{TS}} = \mathcal{K}_0 \oplus \mathcal{K}_I \oplus \mathcal{K}_{II} \oplus \mathcal{K}_{III}, \quad (157)$$

where  $\mathcal{K}_0 = \text{span}\{\sum_{\vec{n} \neq \vec{0}} |\sigma_{\vec{n} \otimes \vec{n}}\rangle\}$  and the others are given in Appendix C. Each irrep is multiplicity-free. We denote by  $D_\lambda$  the dimension of each subspace, which are

$$D_0 = 1, \quad (158)$$

$$D_I = \frac{d^2(d-1)(d+3)}{4}, \quad (159)$$

$$D_{II} = \frac{d^2(d+1)(d-3)}{4}, \quad (160)$$

$$D_{III} = d^2 - 1. \quad (161)$$

As the decomposition is multiplicity-free,  $V^{(2)}(m, \mathcal{E} | \Delta, O_{\text{meas}})$  is a sum of four exponentially decaying functions. Furthermore, from the fact that  $\mathcal{K}_0 = \text{span}\{\sum_{\vec{n} \neq \vec{0}} |\sigma_{\vec{n} \otimes \vec{n}}\rangle\}$ , we obtain that  $C_0(\mathcal{E}) = u(\mathcal{E})$ . Hence, we have

$$V^{(2)}(m, \mathcal{E} | \Delta, O_{\text{meas}}) = A_0 u(\mathcal{E})^m + \sum_{\lambda=I,II,III} A_\lambda C_\lambda(\mathcal{E})^m, \quad (162)$$

where

$$A_\lambda = \langle\langle \mathcal{E}(O_{\text{meas}})^{\otimes 2} | \Pi_\lambda | O_{\text{ini}}^{\otimes 2} \rangle\rangle, \quad (163)$$

$$C_\lambda(\mathcal{E}) = \frac{\text{Tr}[\Pi_\lambda L_{\mathcal{E}}^{\otimes 2}]}{\text{Tr}[\Pi_\lambda]}, \quad (164)$$

with  $\Pi_\lambda$  being the projections onto the irrep  $\mathcal{K}_\lambda$ .

It is not clear whether each  $C_\lambda(\mathcal{E})$  ( $\lambda = I, II, III$ ) has a clear physical meaning, such as  $C_0(\mathcal{E}) = u(\mathcal{E})$  being the unitarity. However, a linear combination of them does. To see this, we use the relation that

$$\begin{aligned} \Pi_I + \Pi_{II} + \Pi_{III} \\ = \Pi_{\text{sym}} - \Pi_0 - |\sigma_{\vec{0} \otimes \vec{0}}\rangle\langle\sigma_{\vec{0} \otimes \vec{0}}| - \sum_{\vec{n} \neq \vec{0}} |S_{\vec{0}\vec{n}}\rangle\langle S_{\vec{0}\vec{n}}|, \end{aligned} \quad (165)$$

where  $S_{\vec{0}\vec{n}} := (\sigma_{\vec{0} \otimes \vec{n}} + \sigma_{\vec{n} \otimes \vec{0}})/\sqrt{2}$ . From this relation, we can show, by a similar calculation to the one-qubit case, that

$$\begin{aligned} \text{Tr}[(\Pi_I + \Pi_{II} + \Pi_{III}) L_{\mathcal{E}}^{\otimes 2}] \\ = \frac{(d^2 - 1)^2}{2} f(\mathcal{E})^2 - u(\mathcal{E}) + \frac{d^2 - 1}{2} h(\mathcal{E}). \end{aligned} \quad (166)$$

As  $\text{Tr}[\Pi_\lambda L_{\mathcal{E}}^{\otimes 2}] = \text{Tr}[\Pi_\lambda] C_\lambda(\mathcal{E}) = D_\lambda C_\lambda(\mathcal{E})$ , we obtain

$$\sum_{\lambda=I,II,III} D_\lambda C_\lambda(\mathcal{E}) = \frac{(d^2 - 1)^2}{2} f(\mathcal{E})^2 - u(\mathcal{E}) + \frac{d^2 - 1}{2} h(\mathcal{E}). \quad (167)$$

We finally note that Table I in Sec. IV E is obtained by constructing the orthonormal basis in each subspace  $\mathcal{K}_I$ ,  $\mathcal{K}_{II}$ , and  $\mathcal{K}_{III}$  (see Appendix C). We also assume that the noise  $\mathcal{E}$  is weak, so that  $\mathcal{E}(O_{\text{meas}}) \approx O_{\text{meas}}$ . Based on this assumption, we have

$$A_\lambda \approx \langle\langle O_{\text{meas}}^{\otimes 2} | \Pi_\lambda | O_{\text{ini}}^{\otimes 2} \rangle\rangle, \quad (168)$$

enabling us to compute  $A_\lambda$  for given initial and measurement operators.

## IX. EXPERIMENTAL REALIZATION OF 2-RB

Based on the former sections, we explain in detail how we have experimentally implemented the 2-RB and estimated the self-adjointness of the noise in the system. In Sec. IX A, we provide the details of the numerical evaluation of the one- and two-qubit 2-RB discussed in Sec. V A. The details of experiment is given in Sec. VIII B.

### A. Numerical analysis

#### 1. Single-qubit systems

We explain the fitting procedure of the 2-RB on one-qubit systems in detail. The noise we consider is given by

the following CPTP map:

$$\mathcal{E}_1(\rho) = qe^{i\theta X}\rho e^{-i\theta X} + (1-q)((1-p)\rho + pX\rho X), \quad (169)$$

which is characterized by three parameters  $p$ ,  $q$ , and  $\theta$ . We particularly choose  $\theta$  as  $p = \sin^2 \theta$  for the fidelity parameter  $f(\mathcal{E}_1)$  to be independent of the coherence parameter  $q$ . Using the Liouville representation, it is straightforward to compute the fidelity parameter, the unitarity, and the self-adjointness parameter of this noise. They are, respectively, given by

$$f(\mathcal{E}) = 1 - \frac{4}{3}p, \quad (170)$$

$$u(\mathcal{E}) = 1 - \frac{8}{3}p(1-p)(1-q^2), \quad (171)$$

$$h(\mathcal{E}) = 1 - \frac{8}{3}p(1-p)(1+q^2). \quad (172)$$

As shown in Theorem 2,  $V^{(2)}$  for one qubit is

$$\begin{aligned} V^{(2)}(m, \mathcal{E}_1 | \Delta, O_{\text{meas}}) \\ = A_0 u(\mathcal{E}_1)^m + A_1 \left( \frac{9}{10} f(\mathcal{E}_1)^2 - \frac{1}{5} u(\mathcal{E}_1) + \frac{3}{10} h(\mathcal{E}_1) \right)^m. \end{aligned} \quad (173)$$

To obtain Fig. 2 in Sec. V A, we first estimate the fidelity parameter  $f(\mathcal{E}_1)$  from the 1-RB, that is, by fitting  $V^{(1)}(m, \mathcal{E}_1 | |0\rangle\langle 0|, |0\rangle\langle 0|)$  and then obtain  $u(\mathcal{E}_1)$  and  $h(\mathcal{E}_1)$  from the fitting results of  $V^{(2)}(m, \mathcal{E}_1 | Z, |0\rangle\langle 0|)$  and the estimated value of  $f$ . The fitting of  $V^{(2)}$  is first performed based on Eq. (173) by regarding the coefficients  $A_0, A_1$  and the two exponential decaying rates as free parameters. Then, we subtract the first exponential curve of Eq. (173) from the data and carry out the fitting of the second exponential curve again where we consider  $A_1$  and the base of the second exponential curve as free parameters. This procedure is redundant, but turns out to improve the accuracy of the fitting because, in most cases, the first exponential decaying rate is larger than the second, and thus the second

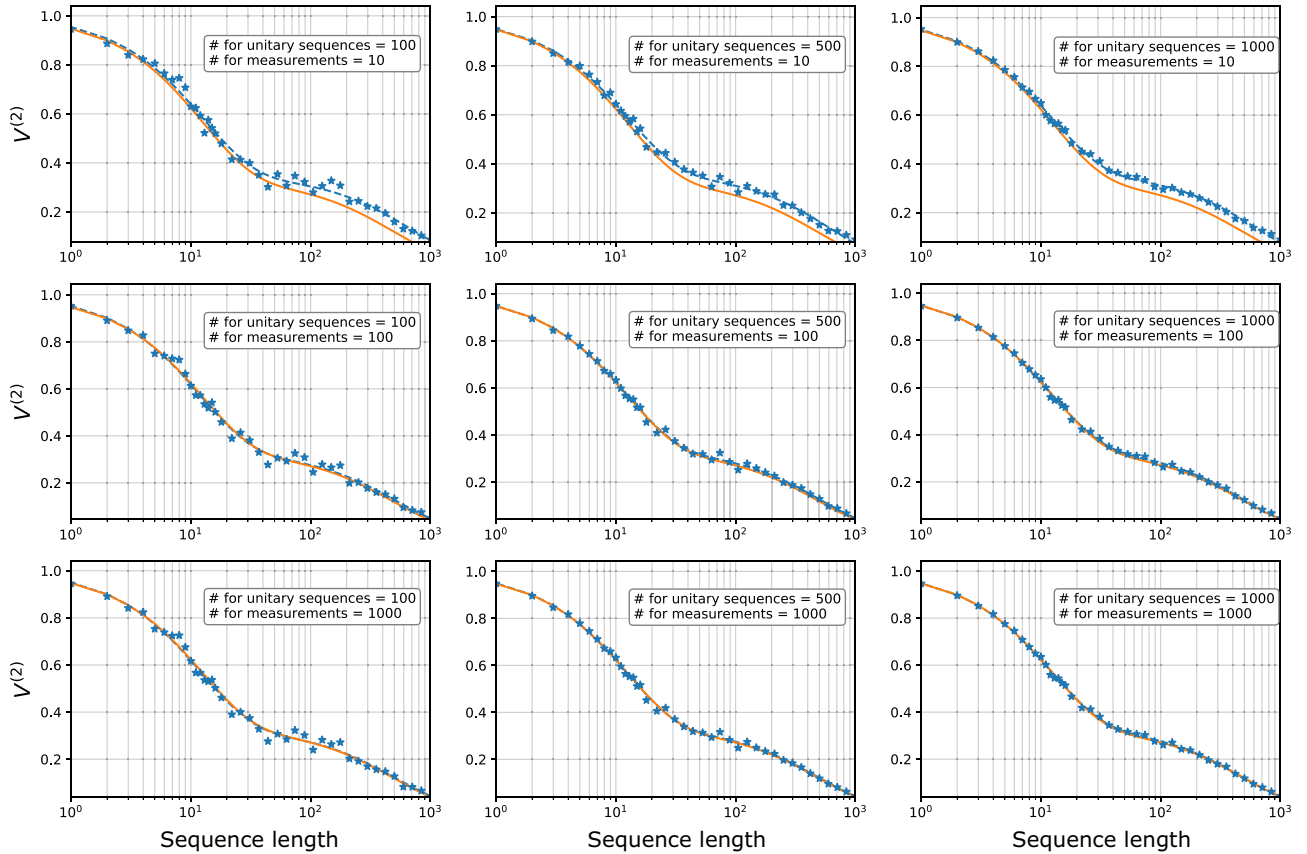


FIG. 6. Numerical results for the single-qubit 2-RB with various numbers of samplings for measurement and those for unitary sequences when the noise parameters are set to  $p = q = 0.02$ . The dots correspond to the values of  $V^{(2)}$  obtained from the given numbers of samplings for measurement and for unitary sequences, and the dashed lines are fitting results. Theoretical values are shown in the solid lines.



exponential curve is clearly visible in the region of long sequence length.

We also estimate how many measurements and unitary sequences suffice to obtain a good estimate of the noise parameters. In Fig. 6, we provide  $V^{(2)}$  for various numbers of measurement and samplings of unitary sequences. We observe that 1000 for both suffice to obtain a good estimate when  $p = q = 0.02$ , which corresponds to the gate fidelity of 97.3%.

When the number of measurements is small, experimental values are positively biased from theoretical values with the infinite number of samplings (see the figures in the top line of Fig. 6). This difference can be understood as follows. Let  $\langle O_{\text{meas}} \rangle_{O_{\text{ini}}, i, n}$  be the expectation value for a random sequence described by  $i$  with the finite number  $n$  of measurements. We describe the expectation value and variance of this random variable averaged over all choices of the unitary sequence as  $\mu$  and  $\sigma^2/n$ , respectively. Note that this mean value is independent of  $n$ , and this variance is inversely proportional to  $n$  because  $\langle O_{\text{meas}} \rangle_{O_{\text{ini}}, i, n}$  is a linear combination of binomial distribution.  $V^{(2)}$  is the expectation value of squared random variable  $\langle O_{\text{meas}} \rangle_{O_{\text{ini}}, i}^2$ , and its expectation value over random sequences is derived as

$$\mathbb{E}[\langle O_{\text{meas}} \rangle_{O_{\text{ini}}, i}^2] = \mathbb{E}[(\langle O_{\text{meas}} \rangle_{O_{\text{ini}}, i} - \mu)^2] + \mu^2 \quad (174)$$

$$= \sigma^2/n + \mu^2. \quad (175)$$

Therefore, an experimentally obtained value with a finite number of measurements is positively biased by  $\sigma^2/n$ . In practice, we can remove the effect of this bias by increasing the number of sampling  $n$  and using the region satisfying  $\mu^2 \ll \sigma^2/n$  for fitting.

We finally check the robustness of the 2-RB on single-qubit systems against SPAM errors. Although the 2-RB is ideally SPAM-error free, the fitting may become harder with the existence of SPAM errors. Our analysis, however, reveals that this is unlikely the case. Here, we model the state-preparation error  $\eta_{\text{prep}}$  as  $\rho = \eta_{\text{prep}}|0\rangle\langle 0| + (1 - \eta_{\text{prep}})|1\rangle\langle 1|$  and  $\rho' = (1 - \eta_{\text{prep}})|0\rangle\langle 0| + \eta_{\text{prep}}|1\rangle\langle 1|$ , and measurement error  $\eta_{\text{meas}}$  as readout bit-flip error, that is, the POVM is  $\{\Pi_{x'}|x' \in \{0, 1\}\}$  where  $\Pi_{x'} = \sum_{x \in \{0, 1\}} \eta_{\text{meas}}(x'|x) \Pi_x$  and  $\eta_{\text{meas}}(x'|x)$  is conditional probability. In this numerical experiments, we assume that  $\eta = \eta_{\text{prep}} = \eta_{\text{meas}}(0|1) = \eta_{\text{meas}}(1|0)$  and obtain 1000 samples and 1000 random sequences. We set the parameters  $p = q = 0.02$ .

The results are provided in Fig. 7 and Table IV. The relative errors of estimates for  $F, u$ , and  $H$  are within 5% except for the estimate for  $u$  when  $\eta = 0.3$ , thus the 2-RB is likely to work well even in realistic situations with SPAM errors as expected from the analytical studies.

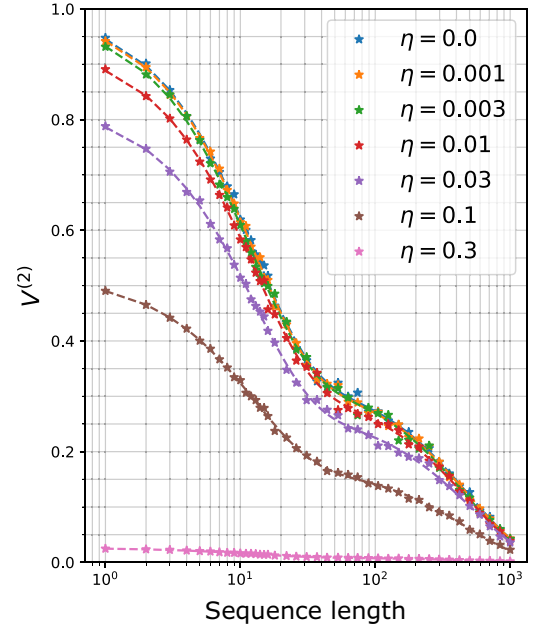


FIG. 7. Numerical results for the single-qubit 2-RB with SPAM errors for  $p = q = 0.02$ . We have taken 1000 measurements and 1000 random unitary sequences. The dots represent the numerically obtained data and the dashed lines are the fitting curves.

## 2. Two-qubit systems

For two-qubit systems, we investigate the noise given by

$$\begin{aligned} \mathcal{E}_2(\rho) &= qe^{i\theta(X \otimes X)} \rho e^{-i\theta(X \otimes X)} \\ &+ (1 - q)((1 - p)\rho + p(X \otimes X)\rho(X \otimes X)), \end{aligned} \quad (176)$$

where we set  $\theta$  to  $p = \sin^2 \theta$ . From Theorem 2,  $V^{(2)}$  in this case shall be in the form of

$$V^{(2)}(m, \mathcal{E}_2|\Delta, O_{\text{meas}}) = A_0 u (\mathcal{E}_2)^m + \sum_{\lambda=I, II, III} A_\lambda C_\lambda (\mathcal{E}_2)^m, \quad (177)$$

TABLE IV. The estimates of  $F, u$ , and  $H$  from the numerics shown in Fig. 7. The parameter  $\eta$  is for SPAM errors such as  $\eta = \eta_{\text{prep}} = \eta_{\text{meas}}(0|1) = \eta_{\text{meas}}(1|0)$ . The theoretical values are listed at the bottom of the table.

SPAM error $\eta$	$F$	$u$	$H$
0.0	0.986(6)	0.9979(5)	0.92(5)
0.001	0.986(5)	0.9979(6)	0.92(7)
0.003	0.986(5)	0.9978(5)	0.92(1)
0.01	0.987(0)	0.9979(2)	0.92(2)
0.03	0.986(5)	0.9980(1)	0.92(8)
0.1	0.986(5)	0.9978(9)	0.92(7)
0.3	0.986(6)	0.9981(8)	0.92(7)
	0.9866	0.99793	0.9247

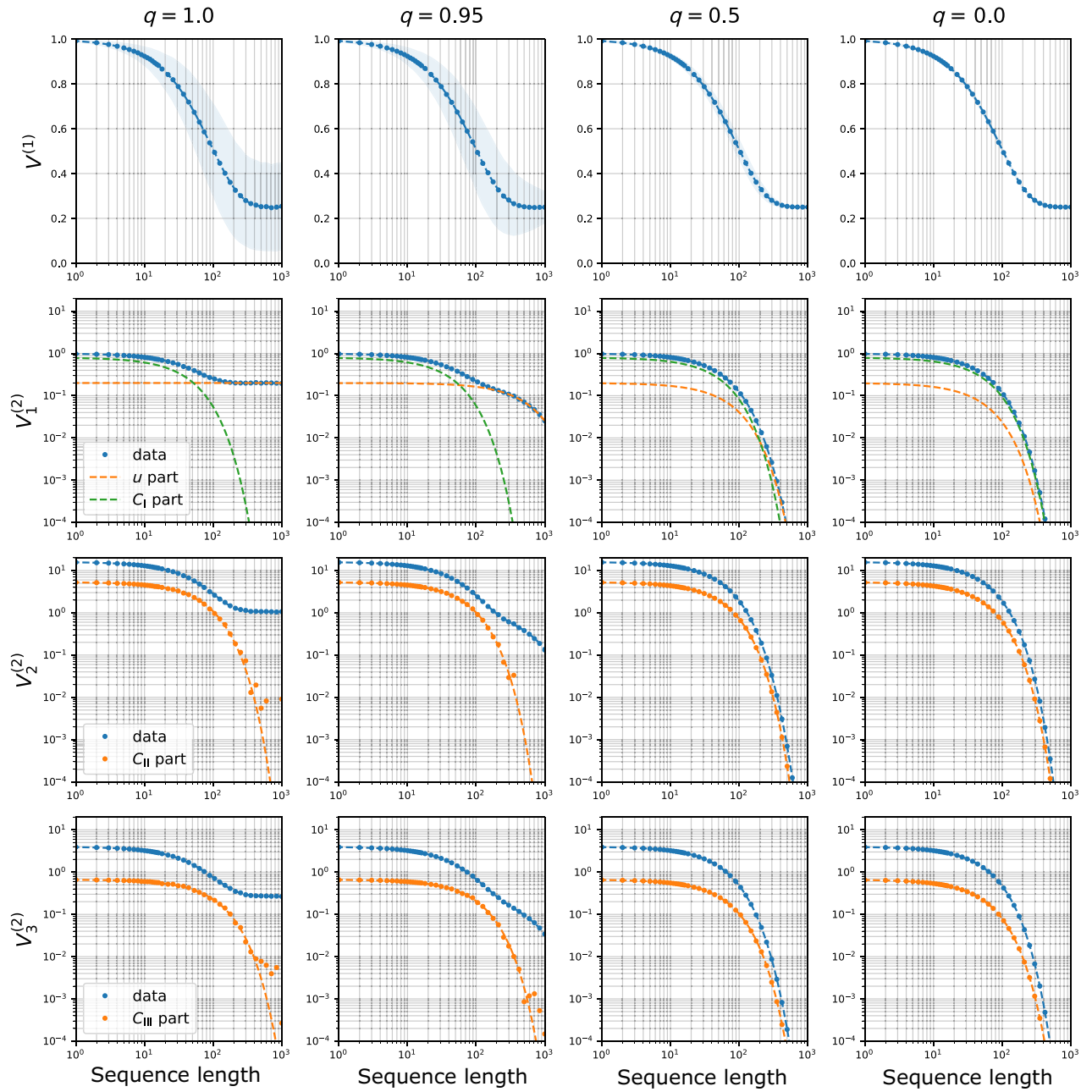


FIG. 8. A step-by-step fitting process of 2-RB for two-qubit systems are shown. The first line of the matrix are about the standard RB (1-RB). The second, third, and fourth lines show 2-RB with two, three, and four exponentially decaying functions, respectively. The initial and measurement operators for 2-RB are chosen according to Table I. See the main text for details.

and  $C_\lambda(\mathcal{E}_2)$  satisfy

$$\begin{aligned} & 84C_I(\mathcal{E}) + 20C_{II}(\mathcal{E}) + 15C_{III}(\mathcal{E}) \\ &= \frac{225}{2}f(\mathcal{E})^2 + \frac{15}{2}h(\mathcal{E}) - u(\mathcal{E}). \end{aligned} \quad (178)$$

With this setting, theoretical values are derived as follows.

$$f(\mathcal{E}_2) = 1 - \frac{16}{15}p, \quad (179)$$

$$u(\mathcal{E}_2) = 1 - \frac{32}{15}p(1-p)(1-q^2), \quad (180)$$

$$C_I(\mathcal{E}_2) = 1 - \frac{4}{105}p(56 - 31p + 14(1-p)q^2), \quad (181)$$

$$C_{II}(\mathcal{E}_2) = 1 - \frac{4}{15}p(8 - 5p - 2(1-p)q^2), \quad (182)$$

$$C_{III}(\mathcal{E}_2) = 1 - \frac{16}{15}p(2 - p - (1-p)q^2), \quad (183)$$

from which the theoretical value of the self-adjointness parameter  $h(\mathcal{E})$  can be computed from Eq. (178).

A step-by-step fitting process of two-qubit systems is shown in Fig. 8 for  $p = 0.01$ . The numbers of measurement and samplings of unitary sequences are both set to  $10^4$ . The fidelity parameters  $f(\mathcal{E}_2)$  can be obtained from  $V^{(1)}(m, \mathcal{E}_1| |00\rangle\langle 00|, |00\rangle\langle 00|)$ , which are plotted in the top line of the figure. In the figure, dashed lines are theoretical values, and the shaded area represents a standard deviation of each data. When  $q$  is near to unity, standard deviations are large even when the sequence length increased. This is because the final quantum state is nearly pure state when  $q \sim 1$ , and probability distributions fluctuate randomly according to chosen random sequences. On the other hand, when  $q \sim 0$ , the final quantum state quickly converges to the maximally mixed state, and thus probability distribution becomes independent of the chosen random sequences.

We then fit four values  $u, C_I, C_{II}, C_{III}$  step by step where initial and measurement operators  $\Delta$  and  $O_{\text{meas}}$  are chosen according to Table I. The fitting results for several  $q$  are listed in Table V.

First, we obtain  $u$  and  $C_I$  from  $V_1^{(2)} := V^{(2)}(m, \mathcal{E}_2| ZZ, |00\rangle\langle 00|)$ . The obtained results are shown in the second line of Fig. 8. The sampled data points and fitting results are shown as blue points and dashed lines, respectively. These lines are linear combinations of two exponentially decaying functions. Exponential decays with coefficient  $u$  and  $C_I$  are shown as orange and green lines, respectively. Although we can clearly see two exponential decays for coherent noise, that is, in the case of  $q \sim 1$ , an exponential decay of  $C_I$  part becomes dominant when noise becomes probabilistic. Even in this case, the fitting results are still reliable as listed in Table V.

Then, we estimate  $C_{II}$  from  $V_2^{(2)} := V^{(2)}(m, \mathcal{E}_2| ZZ, ZZ)$  and  $C_{III}$  from  $V_3^{(2)} := V^{(2)}(m, \mathcal{E}_2| \rho_-, \rho_-)$ , where these initial and measurement operators are chosen from the third and fourth columns of Table I. The obtained results are shown in the third and fourth lines of Fig. 8. In each figure, numerical data is plotted as blue circles and fitting

TABLE V. Fitting results for  $p = 0.01$  in the two-qubit 2-RB. In each cell, fitting results are written in the first line and theoretical values are written in the second line.

$q$	$u$	$C_I$	$C_{II}$	$C_{III}$
0.00	0.9792(7)	0.9787(2)	0.978869(7)	0.97896(1)
	0.97888	0.97879	0.9788000	0.978773
0.50	0.98416(9)	0.97744(3)	0.98015(1)	0.98150(2)
	0.984160	0.977465	0.980120	0.981413
0.95	0.997942(9)	0.97401(3)	0.98354(3)	0.98830(5)
	0.9979410	0.974020	0.983565	0.988304
1.00	0.999994(6)	0.97349(6)	0.98414(5)	0.98930(9)
	1.0000000	0.973505	0.984080	0.989333

TABLE VI. Processed data for  $p = 0.01$  in the two-qubit 2-RB. In each cell, fitting results are written in the first line and theoretical values are written in the second line.

$q$	$F$	$u$	$H$
0.00	0.9920004(5)	0.9792(7)	0.999(2)
	0.99200000	0.97888	1.0000
0.50	0.991998(3)	0.98416(9)	0.9902(3)
	0.9920000	0.984160	0.99010
0.95	0.99203(1)	0.997942(9)	0.9630(5)
	0.992000	0.9979408	0.96426
1.00	0.99198(1)	0.999994(6)	0.9610(8)
	0.992000	1.0000000	0.96040

results are shown as dashed lines. In each fitting process, only a single exponentially decaying term is unknown in advance. We showed unfitted exponential decay as orange circles and fitting results as dashed lines. Although accuracy of orange data becomes not reliable when its value becomes much smaller than the others, we can fit  $C_{II}$  and  $C_{III}$  reliably.

We calculate the averaged fidelity  $F(\mathcal{E})$ , unitarity  $u(\mathcal{E})$ , and self-adjointness  $H(\mathcal{E})$  from the fitting results. The processed values are listed in Table VI. We evaluate relative errors for all the plots with the same method as the single-qubit 2-RB, and confirm that the relative errors are less than 4% for all the cases of  $p = 0.01, 0.02, 0.1, 0.2$  except the case when the theoretical value is exactly zero. Although the relative errors of the self-adjointness become a few tens of percent in the case of  $p = 0.4$ , we can say that reliable values can be obtained when the fidelities of operations are sufficiently high. Note that, although standard deviations of the fitting results become large when  $u$  is almost equal to  $C_I$ , we confirmed that the fitted results are close to theoretical values even in such cases. Thus, we conclude that 2-RB works also for two-qubit systems.

## B. Details of the experiments

In this section, we provide the details of the experiments in Sec. VB. A superconducting qubit can be regarded as a sort of the LC resonant circuit, where a Josephson junction is an effective inductance, and has the Hamiltonian equivalent to that of the one-dimensional free particle trapped in anharmonic potential. The parameter fields of the qubits are summarized in Table VII.

All unitary gates required in the 2-RB for a single-qubit case were implemented by two  $R_X(\pi/2)$  gates and three

TABLE VII. The parameter fields of the qubits.

	$\omega_q/2\pi$	$\alpha/2\pi$	$T_1$	$T_2$ echo
Q1	9.077 GHz	-328.9 MHz	9.724 $\mu\text{s}$	13.670 $\mu\text{s}$
Q2	8.927 GHz	-419.9 MHz	12.634 $\mu\text{s}$	15.763 $\mu\text{s}$

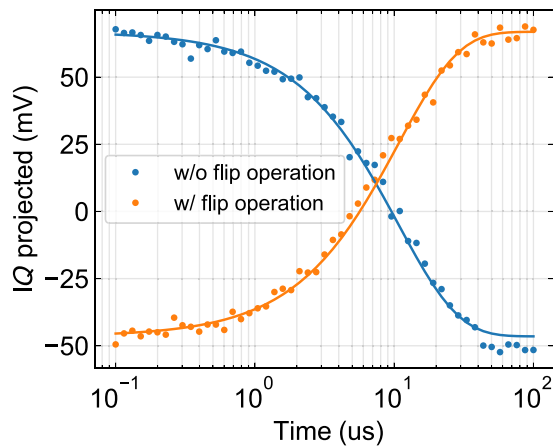


FIG. 9. T1 decay experiment. The horizontal axis represents the delay time, the vertical axis represents the projected value of the IQ readout signal. The blue and orange dots represent the experimental results of the T1 decay experiments with and without flip operation just before measurement, respectively. The lines are fitting curves.

$R_Z$  gates with an arbitrary rotation angle, which are implemented by the shaped microwave pulse (Half-DRAG) with the length 11.70 ns [101] and the Virtual-Z gates [102], respectively. The premeasured averaged gate fidelity of the single-qubit Clifford gate is 0.991.

The single-shot qubit readout was performed via the impedance-matched Josephson parametric amplifier [103], and the assignment fidelity of the readout is 0.943.

As supplemental experiments, T1 decay and Ramsey oscillation were observed to clarify the background noise source of Q1. The experimental results of the T1 decay is shown in Fig. 9, where the horizontal and the vertical axes represent the delay time and the projected value of the IQ readout signal, respectively. The blue and orange dots represent the experimental results with and without a flip operation just before measurement, respectively. The lines provide the fitting curves. As seen from the result, the T1 decay of Q1 follows exponential behavior, which is consistent with the expected behavior in isolated qubits.

The experimental results of the Ramsey oscillation is also given in Fig. 10, where the horizontal and vertical axes represent the detuning from the qubit eigenfrequency and the power spectrum of the Ramsey oscillation, respectively. The points connected by a blue line represent the experimental data, and the orange line provides a fitting curve. In the fitting, we did not take the data in the small power spectrum region ( $< 1 \text{ V}^2$ ) into account. This is because the noise floor derived from the white noise is dominant there. As seen from the result, the power spectrum of the Ramsey oscillation has no peaks other than the qubit eigenfrequency. From the fitting curve, we find that the detuning from the qubit eigenfrequency  $\Delta$  (MHz) and the power spectrum of the Ramsey oscillation  $PS(\Delta)$  ( $\text{V}^2$ )

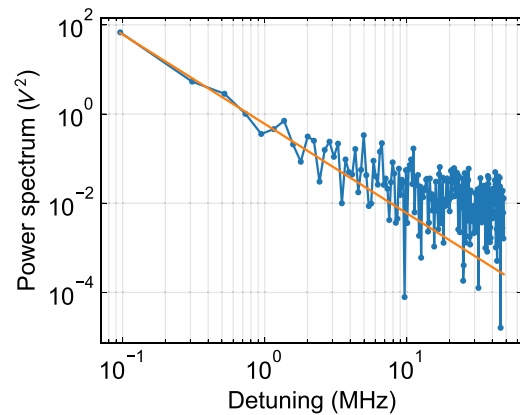


FIG. 10. Ramsey oscillation experiment. The horizontal axis represents the detuning from the qubit eigenfrequency, and the vertical axis represents the power spectrum of the Ramsey oscillation. The points connected by the blue line represent the experimental data, and the orange line represents the fitting curve.

are related as

$$PS(\Delta) \propto \Delta^{-2.004}. \quad (184)$$

This is consistent with the expected behavior when the transmon qubits are isolated well.

From the results of these supplemental experiments, we conclude that Q1 is not in the strong coupling regime with any noise source, which implies that the background noise of Q1 is time-independent. Thus, the requirements for 2-RB are met in our experimental system.

## X. SUMMARY AND DISCUSSION

In this paper, we have provided an explicit constructions of *exact* unitary  $t$ -designs for any  $t$ . In particular, quantum circuits for exact unitary  $t$ -designs on  $N$  qubits have been provided. Our construction is inductive with respect to the number of qubits. Hence, all constructions obtained in this paper are inefficient when the number of qubits is large, implying that it is of practical use only when the size of the system is small.

As an application of exact unitary  $2t$ -designs on a small system, we have proposed the  $t$ -RB, which enables us to experimentally estimate higher-order properties of the noise on a quantum system. As the unitary designs are used in multiple times in a single run of the protocol, it is important for the design to be exact. After providing a general scheme of the  $t$ -RB, we have studied the 2-RB in detail. It has been shown that the 2-RB reveals the self-adjointness of the noise, a characterization of the noise that we argue to play an important role in QEC especially when decoders are based on applications of Pauli operators. Our results have been demonstrated numerically, which shows that the 2-RB is experimentally tractable. We have



then experimentally implemented the single-qubit 2-RB on the superconducting qubit system. From the experimental results, we found that the characteristics of the background noise of a qubit changes depending on the presence of the interaction with the adjacent qubits.

Our results open a number of future problems. Regarding the implementation of  $t$ -designs, it is important to improve the efficiency. Despite that the inefficiency in our construction is likely to be intrinsic owing to an inductive nature of the construction, the representation-theoretic method provides a way to search more efficient ones. More specifically, the key in the construction is the relation between the representation of the whole unitary group and that of a certain subgroup of the unitary group. This indicates that finding the construction of exact unitary designs may be reduced to the problem of searching for a subgroup whose representation has a good relation to that of the whole unitary group.

It is also important to further develop the theory of the  $t$ -RB protocol. In this paper, we have analyzed only the 2-RB in detail. It then turns out that self-adjointness of the noise can be revealed. It is of great interest to concretely investigate what characterization of the noise can be generally obtained from the  $t$ -RB. In the context of QEC, it is also important to comprehensively analyze quantitative relations between the self-adjointness and the feasibility of QEC. Another promising future problem is to use exact higher-designs in the other RB-type protocols. The  $t$ -RB is a straightforward generalization of the standard RB. However, there are numerous variant protocols [69], most of which, if not all, are based on the Clifford group that is an exact unitary 2-design. By extending such protocols to those with higher-designs, the noise on the system can be characterized in more detail.

## ACKNOWLEDGMENTS

The authors are grateful to H. Yamasaki and A. Darmawan for helpful discussions. Y. Nakata is supported by JST PRESTO (No. JPMJPR1865). Y. Suzuki is supported by JST PRESTO (No. JPMJPR1916). Y. Suzuki, K. Heya, Z. Yan, K. Zuo, S. Tamate., Y. Tabuchi, and Y. Nakamura are supported by JST ERATO (No. JPMJER1601) and by MEXT Q-LEAP (No. JPMXS0118068682).

## APPENDIX A: DECOMPOSING $R_X(\theta_\lambda)$ INTO TWO-QUBIT GATES WITH ARBITRARY ACCURACY

We here explicitly show how to decompose  $R_X(\theta_\lambda)$  defined by

$$R_X(\theta_\lambda) = \sum_{\mathbf{j} \in \{0,1\}^N} e^{i\theta_\lambda^{(\mathbf{j})} X} \otimes |\mathbf{j}\rangle\langle\mathbf{j}| \quad (\text{A1})$$

into two-qubit gates with arbitrary precision under the assumption that ancillary qubits and an oracle that approximately compute  $\theta_\lambda^{(\mathbf{j})}$  from  $\mathbf{j}$  can be used.

More precisely, let  $\tilde{\phi}_\lambda^{(\mathbf{j})}$  be a  $m$ -digit binary representation of  $\theta_\lambda^{(\mathbf{j})}/(2\pi) \in [0, 1)$  with an accuracy  $2^{-m}$ , and suppose that an oracle  $Q$  works as  $|\mathbf{j}\rangle|\mathbf{x}\rangle \mapsto |\mathbf{j}\rangle|\mathbf{x} \oplus \tilde{\phi}_\lambda^{(\mathbf{j})}\rangle$  for an arbitrary  $m$ -qubit computational basis  $|\mathbf{x}\rangle$ . From the oracle  $Q$  and an  $m$ -qubit working register, we can construct an  $(N+1)$ -qubit unitary gate  $R_X(\theta_\lambda)$  which approximates  $R_X(\theta_\lambda)$  as

$$\tilde{R}_X(\theta_\lambda)R_X(\theta_\lambda)^\dagger = \sum_{\mathbf{j}} e^{i2\pi\epsilon_{\mathbf{j}}X} \otimes |\mathbf{j}\rangle\langle\mathbf{j}|, \quad (\text{A2})$$

where  $\epsilon_{\mathbf{j}} < 2^{-m}$  for all  $\mathbf{j}$  with two queries to the oracle  $Q$  and with two-qubit quantum gates whose number grows polynomially to  $m$ .

This is done as follows. Suppose that an initial state is  $|\psi_0\rangle := \sum_{\mathbf{j}} |\mathbf{j}\rangle|0\rangle^{\otimes m}|\psi_{\mathbf{j}}\rangle$  without loss of generality. Use the oracle  $Q$  to obtain  $\sum_{\mathbf{j}} |\mathbf{j}\rangle|\tilde{\phi}_\lambda^{(\mathbf{j})}\rangle|\psi_{\mathbf{j}}\rangle$ . Apply  $N$  two-qubit gates  $\Lambda_k = |0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes \exp(i2\pi 2^{-(k+1)}X)$  for  $0 \leq k < N$  to the quantum state where the first part of the tensor product in  $\Lambda_k$  acts on the  $k$ th qubit of the second register and the latter part acts on the last register, then we obtain  $\sum_{\mathbf{j}} |\mathbf{j}\rangle|\tilde{\phi}_\lambda^{(\mathbf{j})}\rangle e^{i2\pi\tilde{\phi}_\lambda^{(\mathbf{j})}X} |\psi_{\mathbf{j}}\rangle$ . We finally undo the second register with an oracle access to obtain  $\sum_{\mathbf{j}} |\mathbf{j}\rangle|0\rangle^{\otimes m} e^{i2\pi\tilde{\phi}_\lambda^{(\mathbf{j})}X} |\psi_{\mathbf{j}}\rangle = (\sum_{\mathbf{j}} |\mathbf{j}\rangle\langle\mathbf{j}| \otimes I^{\otimes m} \otimes e^{i2\pi\tilde{\phi}_\lambda^{(\mathbf{j})}X}) |\psi_0\rangle$ .

As  $2\pi\tilde{\phi}_\lambda^{(\mathbf{j})} - \theta_\lambda^{(\mathbf{j})} < 2^{-m}$ , this process approximates  $R_X(\theta_\lambda)$  within arbitrary accuracy by using sufficiently many number of ancillary qubits  $m$ .

## APPENDIX B: A CLIFFORD-BASED 4-DESIGN ON TWO QUBITS

Denoting by  $R_W(\theta) = \exp[i\theta W]$  ( $W = X, Y, Z$ ) a single-qubit rotation around the  $W$ -axis, we define a single-qubit rotation  $R_1(\theta_1, \theta_2, \theta_3)$  by

$$R_1(\theta_1, \theta_2, \theta_3) = R_Z(\theta_1)R_Y(\theta_2)R_Z(\theta_3). \quad (\text{B1})$$

We also define a two-qubit rotation  $R_2(\varphi_X, \varphi_Y, \varphi_Z)$  by

$$R_2(\varphi_X, \varphi_Y, \varphi_Z) = \exp\left[-i \sum_{W=X,Y,Z} \varphi_W W \otimes W\right]. \quad (\text{B2})$$

We also let  $U_c$  be a fixed two-qubit unitary given by

$$U_c = (R_1(\theta_1, \theta_2, \theta_3) \otimes R_1(\theta'_1, \theta'_2, \theta'_3)) R_2(\varphi_X, \varphi_Y, \varphi_Z) (R_1(\theta_4, \theta_5, \theta_6) \otimes R_1(\theta'_4, \theta'_5, \theta'_6)), \quad (\text{B3})$$

TABLE VIII. Irreducible representations for a single-qubit system. The definitions of the irreducible spaces are based on the irreps of the Clifford group given in Ref. [100], and are explained in the main text.

Highest Weight	Dimension	Multiplicity	Irreducible spaces
$(2, -2)$	5	1	$\mathcal{K}_1$
$(1, -1)$	3	3	$\mathcal{K}_i, \mathcal{K}_r, \mathcal{K}_{\{A\}}$
$(0, 0)$	1	2	$\mathcal{K}_0, \mathcal{K}_{\text{id}}$

where

$$(\theta_1, \theta_2, \theta_3) = (1.50097, 5.69898, 2.53181) \quad (\text{B4})$$

$$(\theta'_1, \theta'_2, \theta'_3) = (1.25383, 0.01700, 6.21127) \quad (\text{B5})$$

$$(\varphi_X, \varphi_Y, \varphi_Z) = (0.376407, 0.368786, 3.69014) \quad (\text{B6})$$

$$(\theta_4, \theta_5, \theta_6) = (4.66335, 3.04854, 1.45524) \quad (\text{B7})$$

$$(\theta'_4, \theta'_5, \theta'_6) = (0.337423, 3.38137, 3.82503). \quad (\text{B8})$$

Then,  $\mathbf{C}(4)U_c\mathbf{C}(4)$  is an exact unitary 4-design on 2 qubits, up to the numerical precision.

These numbers are obtained by numerically searching a zero of a function that is related to the  $\mathbf{C}(2)$ -invariant functions. See Sec. 9.3 of Ref. [77] for more details.

### APPENDIX C: IRREDUCIBLE REPRESENTATIONS

We here provide the irreps of a unitary group  $\mathbf{U}(d)$  on the vector space

$$\mathcal{K} := \text{span}\{|\sigma_{\vec{n}_1} \otimes \vec{n}_2\rangle : \vec{n}_1, \vec{n}_2 \in \{0, 1, 2, 3\}^q\}, \quad (\text{C1})$$

under the action of  $L_{\mathcal{V}^{\otimes 2}}$  with  $\mathcal{V}(\rho) := V\rho V^\dagger$  ( $V \in \mathbf{U}(d)$ ).

To this end, we index the irreps of unitary group  $\mathbf{U}(d)$  by nonincreasing integer sequences:  $(\lambda_1, \lambda_2, \dots, \lambda_d)$ , where  $\lambda_i \in \mathbb{Z}$ , and  $\lambda_i \geq \lambda_j$  ( $i \geq j$ ). The above representation contains the irreps indexed by  $\lambda$ , where  $\lambda^+ \leq 2$ , with  $\lambda^+$  being the sum of positive integers in  $\lambda$ , and  $\sum_i \lambda_i = 0$ .

For a given index  $\lambda$ , the dimension of the corresponding representation space is given by Weyl's dimension

formula:

$$\frac{\prod_{1 \leq i \leq j \leq d} (\lambda_i - \lambda_j + j - i)}{\prod_{k=1}^{d-1} k!}. \quad (\text{C2})$$

The multiplicity can be obtained from the Littlewood-Richardson rule. As a single-qubit case ( $q = 1$ ) is special, in the following we consider the single-qubit case and the multiqubit case separately.

The dimension and the multiplicity of irreps are summarized in Table VIII for one-qubit systems and Table IX for multiqubit systems.

#### 1. Single-qubit systems

To explicitly obtain all the irreps for one qubit, we start with the irreps of the Clifford group  $\mathbf{C}(2)$ , which are provided in Ref. [100]. As the Clifford group is a subgroup of the unitary group  $\mathbf{U}(2)$ , irreps of the unitary group are obtained by taking the union of some irreps of the Clifford group. As the dimensions of the irreps, both of the Clifford and the unitary groups, are known, we can check which irreps of the Clifford group should be combined by dimension counting.

First, from Theorem 1 in Ref. [100], the irreducible decomposition of  $\mathcal{K}$  in terms of  $\mathbf{C}(2)$  is given by

$$\mathcal{K}_{\text{id}} \oplus \mathcal{K}_0 \oplus \mathcal{K}_1 \oplus \mathcal{K}_r \oplus \mathcal{K}_l \oplus \mathcal{K}_{\{S\}} \oplus \mathcal{K}_{\{A\}}. \quad (\text{C3})$$

Here, the important subspaces in our analysis are

$$\mathcal{K}_0 := \text{span}\{|\sigma_1^{\otimes 2} + \sigma_2^{\otimes 2} + \sigma_3^{\otimes 2}\rangle\rangle\}, \quad (\text{C4})$$

$$\mathcal{K}_1 := \text{span}\{|\mathcal{S}_{1,2}\rangle\rangle, |\mathcal{S}_{1,3}\rangle\rangle, |\mathcal{S}_{2,3}\rangle\rangle\}, \quad (\text{C5})$$

$$\mathcal{K}_{\{S\}} := \text{span}\{|\sigma_1^{\otimes 2} - 2\sigma_2^{\otimes 2} + \sigma_3^{\otimes 2}\rangle\rangle, |\sigma_1^{\otimes 2} - \sigma_3^{\otimes 2}\rangle\rangle\}. \quad (\text{C6})$$

See Ref. [100] for the definitions of the other subspaces, where the vector space is denoted by  $V$  instead of  $\mathcal{K}$ . For instance,  $\mathcal{K}_{\text{id}}$  in our notation corresponds to  $V_{\text{id}}$  in the notation of Ref. [100].

TABLE IX. Irreducible representations for a multiqubit system, where  $d = 2^q$ . The definitions of the irreducible spaces are based on the irreps of the Clifford group given in Ref. [100], and are explained in the main text.

Highest Weight	Dimension	Multiplicity	Irreducible spaces
$(2, 0, \dots, 0, -2)$	$(d^2(d-1)(d+3))/4$	1	$\mathcal{K}_1$
$(2, 0, \dots, 0, -1, -1)$	$((d^2-1)(d^2-4))/4$	1	$\mathcal{K}_{[A]}$
$(1, 1, 0, \dots, 0, -2)$	$((d^2-1)(d^2-4))/4$	1	$\mathcal{K}_{[\text{adj}]}$
$(1, 1, 0, \dots, 0, -1, -1)$	$(d^2(d+1)(d-3))/4$	1	$\mathcal{K}_{\Pi}$
$(1, 0, \dots, 0, -1)$	$d^2 - 1$	4	$\mathcal{K}_l, \mathcal{K}_r, \mathcal{K}_{[\text{adj}]}, \mathcal{K}_{\{\text{adj}\}}$
$(0, \dots, 0)$	1	2	$\mathcal{K}_0, \mathcal{K}_{\text{id}}$

The dimension of each subspace is also given in Ref. [100] as

$$\dim \mathcal{K}_{\text{id}} = \dim \mathcal{K}_0 = 1, \quad (\text{C7})$$

$$\dim \mathcal{K}_1 = 2, \quad (\text{C8})$$

$$\dim \mathcal{K}_r = \dim \mathcal{K}_l = \dim \mathcal{K}_{\{S\}} = \dim \mathcal{K}_{\{A\}} = 3. \quad (\text{C9})$$

Comparing these with the dimensions of irreps of the unitary group, given in Table VIII, it is clear that  $\mathcal{K}_{\text{id}}, \mathcal{K}_0$  are also irreps of  $\text{U}(2)$ .

We can also show that  $\mathcal{K}_l := \mathcal{K}_{\{S\}} \oplus \mathcal{K}_1$  is an irrep of  $\text{U}(2)$ . To this end, we show that  $\exists U \in \text{U}(2), \exists |v\rangle \in \mathcal{K}_1$  such that  $L_{U^{\otimes 2}}|v\rangle$  has a support on  $\mathcal{K}_{\{S\}}$ . Together with the dimension counting, it immediately leads to that  $\mathcal{K}_l$  is irreducible.

The statement is shown by construction. Let  $T := \text{diag}(1, e^{i\pi/4})$  be a diagonal unitary matrix in  $\text{U}(2)$ . Using the simple relation that

$$TXT^\dagger = \frac{1}{\sqrt{2}}(X + Y), \quad (\text{C10})$$

$$TZT^\dagger = Z, \quad (\text{C11})$$

it is straightforward to show that

$$L_{T^{\otimes 2}}|\sigma_1^{\otimes 2} - \sigma_3^{\otimes 2}\rangle \propto -|\sigma_1^{\otimes 2} - 2\sigma_2^{\otimes 2} + \sigma_3^{\otimes 2}\rangle + 3|S_{1,2}\rangle. \quad (\text{C12})$$

As  $|\sigma_1^{\otimes 2} - \sigma_3^{\otimes 2}\rangle \in \mathcal{K}_1$  and  $|S_{1,2}\rangle \in \mathcal{K}_{\{S\}}$ , we obtain the desired statement.

The fact that  $\mathcal{K}_1, \mathcal{K}_0$ , and  $\mathcal{K}_{\text{id}}$  are irreducible with respect to  $\text{U}(2)$  implies that so are the rest, that is,  $\mathcal{K}_l, \mathcal{K}_r$ , and  $\mathcal{K}_{\{A\}}$ , owing to the dimension condition. We, hence, obtain that the irreducible decomposition of  $\mathcal{K}$ :

$$\mathcal{K}_{\text{id}} \oplus \mathcal{K}_0 \oplus \mathcal{K}_r \oplus \mathcal{K}_l \oplus \mathcal{K}_{\{A\}} \oplus \mathcal{K}_1, \quad (\text{C13})$$

where  $\mathcal{K}_{\text{id}}$  and  $\mathcal{K}_0$  are equivalent representations, and  $\mathcal{K}_r, \mathcal{K}_l$ , and  $\mathcal{K}_{\{A\}}$  are equivalent.

We finally mention the fact that the traceless symmetric ones are only  $\mathcal{K}_0$  and  $\mathcal{K}_1$ , which can be directly confirmed from their definitions. Thus, the traceless symmetric space  $\mathcal{K}_{\text{TS}}$  is decomposed into irreducible subspaces as

$$\mathcal{K}_{\text{TS}} = \mathcal{K}_0 \oplus \mathcal{K}_1, \quad (\text{C14})$$

which is multiplicity-free.

## 2. Multiqubit systems

For  $q \geq 2$ , we also start with the irreps of the Clifford group  $\mathbf{C}(d)$ . From Theorem 1 in Ref. [100], the irreducible

decomposition by  $\mathbf{C}(d)$  is given by

$$\begin{aligned} \mathcal{K} = & \mathcal{K}_{\text{id}} \oplus \mathcal{K}_0 \oplus \mathcal{K}_1 \oplus \mathcal{K}_2 \oplus \mathcal{K}_{\{\text{adj}\}} \oplus \mathcal{K}_{\{1\}} \oplus \mathcal{K}_{\{2\}} \\ & \oplus \mathcal{K}_{\{\text{adj}\}} \oplus \mathcal{K}_{\{1\}} \oplus \mathcal{K}_{\{2\}} \oplus \mathcal{K}_{\{A\}} \oplus \mathcal{K}_{\{\text{adj}\}}^\perp. \end{aligned} \quad (\text{C15})$$

Similarly to the single-qubit case, by comparing the dimension of each subspace with those in Table IX, we obtain that  $\mathcal{K}_{\text{id}}$  and  $\mathcal{K}_0$  are the trivial irreps, and that  $\mathcal{K}_l, \mathcal{K}_r, \mathcal{K}_{\{\text{adj}\}}$ , and  $\mathcal{K}_{\{\text{adj}\}}$  are those corresponding to  $(1, 0, \dots, 0, -1)$ . We can also observe that  $\mathcal{K}_{\{A\}}$  and  $\mathcal{K}_{\{\text{adj}\}}^\perp$  are also irreps, corresponding to  $(2, 0, \dots, 0, -1, -1)$  and  $(1, 1, \dots, 0, -2)$ , respectively.

We, hence, need to identify which of  $\mathcal{K}_1, \mathcal{K}_2, \mathcal{K}_{\{1\}}, \mathcal{K}_{\{2\}}, \mathcal{K}_{\{1\}}$ , and  $\mathcal{K}_{\{2\}}$  consist of the irrep  $\mathcal{K}_I$  with  $(2, 0, \dots, 0, -2)$  and the irrep  $\mathcal{K}_{II}$  with  $(1, 1, 0, \dots, 0, -1, -1)$ . This can be done again by dimension counting. From [100], the dimension of each subspace is given by

$$\dim \mathcal{K}_1 = \frac{d(d+1)}{2} - 1, \quad (\text{C16})$$

$$\dim \mathcal{K}_2 = \frac{d(d-1)}{2} - 1, \quad (\text{C17})$$

$$\dim \mathcal{K}_{\{1\}} = (d^2 - 1) \left[ \frac{d(d+2)}{8} - 1 \right], \quad (\text{C18})$$

$$\dim \mathcal{K}_{\{2\}} = (d^2 - 1) \left[ \frac{d(d-2)}{8} - 1 \right], \quad (\text{C19})$$

$$\dim \mathcal{K}_{\{1\}} = \frac{d(d^2 - 1)(d+2)}{8}, \quad (\text{C20})$$

$$\dim \mathcal{K}_{\{2\}} = \frac{d(d^2 - 1)(d-2)}{8}. \quad (\text{C21})$$

By a straightforward calculation, it can be shown that the only possible combination to satisfy the dimension condition that  $\dim \mathcal{K}_I = (d^2(d-1)(d+3))/4$  and  $\dim \mathcal{K}_{II} = (d^2(d+1)(d-3))/4$  is

$$\mathcal{K}_I := \mathcal{K}_1 \oplus \mathcal{K}_{\{1\}} \oplus \mathcal{K}_{\{1\}}, \quad (\text{C22})$$

$$\mathcal{K}_{II} := \mathcal{K}_2 \oplus \mathcal{K}_{\{2\}} \oplus \mathcal{K}_{\{2\}}. \quad (\text{C23})$$

We finally consider the irreducible decomposition of the traceless symmetric subspace  $\mathcal{K}_{\text{TS}}$ . From the definitions of each subspace [100], we can easily check that the irreducible decomposition of  $\mathcal{K}_{\text{TS}}$  by  $\text{U}(d)$  is given by

$$\mathcal{K}_{\text{TS}} = \mathcal{K}_0 \oplus \mathcal{K}_1 \oplus \mathcal{K}_{II} \oplus \mathcal{K}_{\{\text{adj}\}}, \quad (\text{C24})$$

which is multiplicity-free. In the main text, we have denoted  $\mathcal{K}_{\{\text{adj}\}}$  by  $\mathcal{K}_{III}$  for the simplicity of notation.

## APPENDIX D: CHARACTERIZATIONS OF A NOISE

We here show the following properties of the self-adjointness  $H(\mathcal{E})$  and the self-adjointness parameter  $h(\mathcal{E})$ .

1. In the Liouville representation, the self-adjointness parameter  $h(\mathcal{E})$  is given by

$$h(\mathcal{E}) = \frac{1}{d^2 - 1} \sum_{\vec{n} \neq \vec{0}} \langle\langle \sigma_{\vec{n}} | L_{\mathcal{E}}^2 | \sigma_{\vec{n}} \rangle\rangle \quad (\text{D1})$$

$$= \frac{1}{d^2 - 1} \text{Tr}[\tilde{L}_{\mathcal{E}}^2]. \quad (\text{D2})$$

2. The self-adjointness parameter  $h(\mathcal{E})$  satisfies

$$-\frac{1}{d^2 - 1} \leq h(\mathcal{E}) \leq u(\mathcal{E}), \quad (\text{D3})$$

which immediately implies that  $0 \leq H(\mathcal{E}) \leq 1$ .

3. We have  $h(\mathcal{E}) = u(\mathcal{E})$  if and only if  $\tilde{L}_{\mathcal{E}} = \tilde{L}_{\mathcal{E}}^\dagger$ . For a unital noise,  $h(\mathcal{E}) = u(\mathcal{E})$  if and only if the noise is self-adjoint.
4. We have  $h(\mathcal{E}) = -1/(d^2 - 1)$  if and only if  $\text{Tr}[K_i K_j] = 0$  for any  $i, j$ , where  $\{K_i\}$  are the Kraus operators of  $\mathcal{E}$ .
5. The average gate fidelity  $F(\mathcal{E})$  is bounded from above by  $u(\mathcal{E})$  and  $h(\mathcal{E})$ :

$$F(\mathcal{E}) \leq \frac{d-1}{d} \sqrt{\frac{h(\mathcal{E}) + u(\mathcal{E})}{2}} + \frac{1}{d}. \quad (\text{D4})$$

We first show Eq. (D1). Recalling that  $(L_{\mathcal{E}})_{\vec{0}\vec{n}} = 0$  for any  $\vec{n} \neq \vec{0}$  because  $\mathcal{E}$  is trace-preserving, we have

$$\sum_{\vec{n}, \vec{m} \neq \vec{0}} \text{Tr}[\sigma_{\vec{n}} \mathcal{E}(\sigma_{\vec{m}})] \text{Tr}[\sigma_{\vec{m}} \mathcal{E}(\sigma_{\vec{n}})] \quad (\text{D5})$$

$$= \sum_{\vec{n}, \vec{m} \neq \vec{0}} \text{Tr}[\sigma_{\vec{n}}^{\otimes 2} (\mathcal{E} \otimes \mathcal{E}^\dagger)(\sigma_{\vec{m}}^{\otimes 2})] \quad (\text{D6})$$

$$= \text{Tr}[(\mathbb{F} - \sigma_0^{\otimes 2})(\mathcal{E} \otimes \mathcal{E}^\dagger)(\mathbb{F} - \sigma_0^{\otimes 2})] \quad (\text{D7})$$

$$= \text{Tr}[\mathbb{F}(\mathcal{E} \otimes \mathcal{E}^\dagger)(\mathbb{F})] - 1, \quad (\text{D8})$$

where  $\mathbb{F} := \sum_{\vec{n}} \sigma_{\vec{n}}^{\otimes 2}$  is the swap operator on  $\mathcal{K}^{\otimes 2}$ . Note that, in the last line, we have used the fact that  $\mathcal{E}$  is trace-preserving, also implying that  $\mathcal{E}^\dagger$  is unital. We now use another expression of the swap operator, which is

$$\mathbb{F} = d \left[ (d+1) \int \varphi^{\otimes 2} d\varphi - \sigma_0^{\otimes 2} \right], \quad (\text{D9})$$

which simply follows from Schur's lemma. Substituting this and using the relation that  $\text{Tr}[\mathbb{F}(A \otimes B)] = \text{Tr}[AB]$  for

any operators  $A$  and  $B$ , we obtain

$$\begin{aligned} & \frac{1}{d^2 - 1} \sum_{\vec{n} \neq \vec{0}} \langle\langle \sigma_{\vec{n}} | L_{\mathcal{E}}^2 | \sigma_{\vec{n}} \rangle\rangle \\ &= \frac{1}{d-1} \left[ d \int \text{Tr}[\mathcal{E}(\varphi) \mathcal{E}^\dagger(\varphi)] d\varphi - 1 \right]. \end{aligned} \quad (\text{D10})$$

On the other hand, it is straightforward to show that

$$h(\mathcal{E}) := \frac{d}{d-1} \int \text{Tr}[\mathcal{E}'(\varphi) \mathcal{E}^{\dagger'}(\varphi)] d\varphi, \quad (\text{D11})$$

$$= \frac{1}{d-1} \left[ d \int \text{Tr}[\mathcal{E}(\varphi) \mathcal{E}^\dagger(\varphi)] d\varphi - 1 \right], \quad (\text{D12})$$

which follows from the definition of  $\mathcal{E}'$ , that is,  $\mathcal{E}'(\rho) := \mathcal{E}(\rho - I/d)$ . Hence, we have Eq. (D1). Note that Eq. (D2) follows simply from the definition of  $\tilde{L}_{\mathcal{E}}$ .

We next show Eq. (D3) and properties 3 and 4. The lower bound of  $h(\mathcal{E})$  in Eq. (D3) is obtained from Eqs. (D1) and (D8), which lead to

$$h(\mathcal{E}) = \frac{1}{d^2 - 1} (\text{Tr}[\mathbb{F}(\mathcal{E} \otimes \mathcal{E}^\dagger)(\mathbb{F})] - 1). \quad (\text{D13})$$

Using the Kraus operators  $\{K_i\}$  for  $\mathcal{E}$  and the swap trick, this can be rewritten as

$$h(\mathcal{E}) = \frac{1}{d^2 - 1} \left( \sum_{ij} |\text{Tr}[K_i K_j]|^2 - 1 \right), \quad (\text{D14})$$

which is not smaller than  $-1/(d^2 - 1)$ . The equality holds if and only if  $\text{Tr}[K_i K_j] = 0$  for any  $i, j$ , which is property 4. A simple instance of such a noise is a unitary noise that maps  $\rho$  to  $\sqrt{X} \rho \sqrt{X}^\dagger$ .

The upper bound of  $h(\mathcal{E})$  in Eq. (D3) is obtained from the relation that

$$\begin{aligned} & \int d\psi \text{Tr} \left[ (\mathcal{E}(|\psi\rangle\langle\psi| - I/d) - \mathcal{E}^\dagger(|\psi\rangle\langle\psi| - I/d))^2 \right] \\ &= \frac{2(d-1)}{d} (u(\mathcal{E}) - h(\mathcal{E})), \end{aligned} \quad (\text{D15})$$

which can be checked by a direct calculation. As the left-hand side is nonnegative, we have  $u(\mathcal{E}) \geq h(\mathcal{E})$ . Regarding property 3, it is obvious that  $h(\mathcal{E}) = u(\mathcal{E})$  if  $\mathcal{E} = \mathcal{E}^\dagger$ . The converse is shown as follows:

$$h(\mathcal{E}) = u(\mathcal{E}), \quad (\text{D16})$$

$$\Leftrightarrow \text{Tr}[\tilde{L}_{\mathcal{E}}^2] = \text{Tr}[\tilde{L}_{\mathcal{E}}^\dagger \tilde{L}_{\mathcal{E}}], \quad (\text{D17})$$

$$\Leftrightarrow \sum_{\vec{n} > \vec{m}} ((L_{\mathcal{E}})_{\vec{n}\vec{m}} - (L_{\mathcal{E}})_{\vec{m}\vec{n}}) = 0, \quad (\text{D18})$$



$$\Leftrightarrow (L_{\mathcal{E}})_{\vec{n}\vec{m}} = (L_{\mathcal{E}})_{\vec{m}\vec{n}}, \quad \forall \vec{n}, \vec{m}, \quad (\text{D19})$$

$$\Leftrightarrow L_{\mathcal{E}} = L_{\mathcal{E}}^{\dagger}. \quad (\text{D20})$$

Note that the last line follows from the fact that  $L_{\mathcal{E}}$  is a real matrix in the Pauli basis.

We finally show the relation between the average fidelity  $F(\mathcal{E})$ , the self-adjointness parameter  $h(\mathcal{E})$ , and the unitarity  $u(\mathcal{E})$ , that is, Eq. (D4). To this end, we again use the fact that  $L_{\mathcal{E}}$  is a real matrix in the Pauli basis, leading to  $\text{Tr}[\tilde{L}_{\mathcal{E}}] = \text{Tr}[L_{\mathcal{E}}]$ . We, hence, have

$$(\text{Tr}[\tilde{L}_{\mathcal{E}}])^2 = \frac{1}{4}(\text{Tr}[\tilde{L}_{\mathcal{E}}] + \text{Tr}[\tilde{L}_{\mathcal{E}}^{\dagger}])^2, \quad (\text{D21})$$

$$\leq \frac{1}{4}\|\text{Tr}[\tilde{L}_{\mathcal{E}}] + \text{Tr}[\tilde{L}_{\mathcal{E}}^{\dagger}]\|_1^2, \quad (\text{D22})$$

$$\leq \frac{d^2 - 1}{4}\|\text{Tr}[\tilde{L}_{\mathcal{E}}] + \text{Tr}[\tilde{L}_{\mathcal{E}}^{\dagger}]\|_2^2, \quad (\text{D23})$$

$$= \frac{d^2 - 1}{4}\text{Tr}[\tilde{L}_{\mathcal{E}}^2 + (\tilde{L}_{\mathcal{E}}^{\dagger})^2 + 2\tilde{L}_{\mathcal{E}}\tilde{L}_{\mathcal{E}}^{\dagger}], \quad (\text{D24})$$

$$= \frac{(d^2 - 1)^2}{2}(h(\mathcal{E}) + u(\mathcal{E})), \quad (\text{D25})$$

where the second inequality follows from the relation that  $\|M\|_1 \leq \sqrt{D}\|M\|_2$  for any  $D \times D$  matrix  $M$ , and the last line from that  $L_{\mathcal{E}}$  is a real matrix, which implies that  $\text{Tr}[\tilde{L}_{\mathcal{E}}^2] = \text{Tr}[(\tilde{L}_{\mathcal{E}}^{\dagger})^2]$ . As  $F(\mathcal{E}) = (\text{Tr}[L_{\mathcal{E}}] + d)/d(d + 1)$  and  $\text{Tr}[L_{\mathcal{E}}] = \text{Tr}[\tilde{L}_{\mathcal{E}}] + 1$ , we obtain Eq. (D4).

---

[1] A. Ambainis and J. Emerson, in *22nd Annual IEEE Conference on Computational Complexity*, San Diego, CA (2007), p. 129.

[2] I. Devetak, The private classical capacity and quantum capacity of a quantum channel, *IEEE Trans. Inf. Theory* **51**, 44 (2005).

[3] I. Devetak and A. Winter, Relating Quantum Privacy and Quantum Coherence: An Operational Approach, *Phys. Rev. Lett.* **93**, 080501 (2004).

[4] B. Groisman, S. Popescu, and A. Winter, Quantum, classical, and total amount of correlations in a quantum state, *Phys. Rev. A* **72**, 032317 (2005).

[5] A. Abeyesinghe, I. Devetak, P. Hayden, and A. Winter, The mother of all protocols: Restructuring quantum information's family tree, *Proc. R. Soc. A* **465**, 2537 (2009).

[6] F. Dupuis, M. Berta, J. Wullschlegler, and R. Renner, One-shot decoupling, *Commun. Math. Phys.* **328**, 251 (2014).

[7] O. Szehr, F. Dupuis, M. Tomamichel, and R. Renner, Decoupling with unitary approximate two-designs, *New J. Phys.* **15**, 053022 (2013).

[8] M. Horodecki, J. Oppenheim, and A. Winter, Partial quantum information, *Nature* **436**, 673 (2005).

[9] M. Horodecki, J. Oppenheim, and A. Winter, Quantum state merging and negative information, *Comms. Math. Phys.* **269**, 107 (2007).

[10] A. Ambainis and A. Smith, in *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques. RANDOM 2004, APPROX 2004*, edited by K. Jansen, S. Khanna, J.D.P. Rolim and D. Ron, Lecture Notes in Computer Science (Springer, Berlin, Heidelberg, 2004), Vol. 3122, p. 249.

[11] P. Hayden, D. Leung, P. W. Shor, and A. Winter, Randomizing quantum states: Constructions and applications, *Commun. Math. Phys.* **250**, 371 (2004).

[12] P. Sen, in *21st Annual IEEE Conference on Computational Complexity (CCC'06)* (2006), p. 14.

[13] F. G. S. L. Brandão and M. Horodecki, Exponential quantum speed-ups are generic, *Q. Inf. Comp.* **13**, 0901 (2013).

[14] R. Kueng, H. Rauhut, and U. Terstiege, Low rank matrix recovery from rank one measurements, *Appl. Comput. Harmon. Anal.* **42**, 88 (2017).

[15] S. Kimmel and Y. Liu, in *2017 International Conference on Sampling Theory and Applications (SampTA)* (2017), p. 345.

[16] R. Kueng, H. Zhu, and D. Gross, Distinguishing quantum states using Clifford orbits, [arXiv:1609.08595](https://arxiv.org/abs/1609.08595) (2016).

[17] M. Oszmaniec, R. Augusiak, C. Gogolin, J. Kołodyński, A. Acín, and M. Lewenstein, Random Bosonic States for Robust Quantum Metrology, *Phys. Rev. X* **6**, 041044 (2016).

[18] S. Boixo *et al.*, Characterizing quantum supremacy in near-term devices, *Nat. Phys.* **14**, 595 (2018).

[19] F. Arute *et al.*, Quantum supremacy using a programmable superconducting processor, *Nature* **574**, 505 (2019).

[20] A. Bouland, B. Fefferman, C. Nirkhe, and U. Vazirani, On the complexity and verification of quantum random circuit sampling, *Nat. Phys.* **15**, 159 (2019).

[21] M. Oszmaniec, A. Sawicki, and M. Horodecki, Epsilon-nets, unitary designs and random quantum circuits, [arXiv:2007.10885](https://arxiv.org/abs/2007.10885) (2020).

[22] J. Emerson, R. Alicki, and K. Życzkowski, Scalable noise estimation with random unitary operators, *J. Opt. B: Quantum Semiclass. opt.* **7**, S347 (2005).

[23] E. Knill *et al.*, Randomized benchmarking of quantum gates, *Phys. Rev. A* **77**, 012307 (2008).

[24] E. Magesan, J. M. Gambetta, and J. Emerson, Scalable and Robust Randomized Benchmarking of Quantum Processes, *Phys. Rev. Lett.* **106**, 180504 (2011).

[25] E. Magesan, J. M. Gambetta, and J. Emerson, Characterizing quantum gates via randomized benchmarking, *Phys. Rev. A* **85**, 042311 (2012).

[26] J. Kelly *et al.*, Optimal Quantum Control Using Randomized Benchmarking, *Phys. Rev. Lett.* **112**, 240504 (2014).

[27] S. Sheldon *et al.*, Characterizing errors on qubit operations via iterative randomized benchmarking, *Phys. Rev. A* **93**, 012301 (2016).

[28] S. Garion *et al.*, Experimental implementation of non-clifford interleaved randomized benchmarking with a controlled-s gate, *Phys. Rev. Res.* **3**, 013204 (2021).

[29] David C. McKay *et al.*, Three-Qubit Randomized Benchmarking, *Phys. Rev. Lett.* **122**, 200502 (2019).

[30] A. D. Córcoles *et al.*, Process verification of two-qubit quantum gates by randomized benchmarking, *Phys. Rev. A* **87**, 030301(R) (2013).

- [31] Jay M. Gambetta *et al.*, Characterization of Addressability by Simultaneous Randomized Benchmarking, *Phys. Rev. Lett.* **109**, 240504 (2012).
- [32] E. Magesan *et al.*, Efficient Measurement of Quantum Gate Error by Interleaved Randomized Benchmarking, *Phys. Rev. Lett.* **109**, 080505 (2012).
- [33] E. Onorati, A. H. Werner, and J. Eisert, Randomized Benchmarking for Individual Quantum Gates, *Phys. Rev. Lett.* **123**, 060501 (2019).
- [34] J. Helsen, I. Roth, E. Onorati, A. H. Werner, and J. Eisert, A general framework for randomized benchmarking, [arXiv:2010.07974](https://arxiv.org/abs/2010.07974) (2020).
- [35] S. Popescu, A. J. Short, and A. Winter, Entanglement and the foundations of statistical mechanics, *Nat. Phys.* **2**, 754 (2006).
- [36] L. del Rio, J. Aberg, R. Renner, O. Dahlsten, and V. Vedral, The thermodynamic meaning of negative entropy, *Nature* **474**, 61 (2011).
- [37] L. del Rio, A. Hutter, R. Renner, and S. Wehner, Relative thermalization, *Phys. Rev. E* **94**, 022104 (2016).
- [38] P. Hayden and J. Preskill, Black holes as mirrors: Quantum information in random subsystems, *J. High Energy Phys.* **2007**, 120 (2007).
- [39] Y. Sekino and L. Susskind, Fast scramblers, *J. High Energy Phys.* **2008**, 065 (2008).
- [40] L. Susskind, Addendum to Fast Scramblers, [arXiv:1101.6048](https://arxiv.org/abs/1101.6048) (2011).
- [41] N. Lashkari, D. Stanford, M. Hastings, T. Osborne, and P. Hayden, Towards the fast scrambling conjecture, *J. High Energy Phys.* **2013**, 22 (2013).
- [42] P. Hosur, X.-L. Qi, D. A. Roberts, and B. Yoshida, Chaos in quantum channels, *J. High Energy Phys.* **2016**, 4 (2016).
- [43] D. A. Roberts and B. Yoshida, Chaos and complexity by design, *J. High Energy Phys.* **2017**, 121 (2017).
- [44] Y. Nakata, E. Wakakuwa, and M. Koashi, Information leakage from quantum black holes with symmetry, [arXiv:2007.00895](https://arxiv.org/abs/2007.00895) (2020).
- [45] K. A. Landsman *et al.*, Verified quantum information scrambling, *Nature* **567**, 61 (2019).
- [46] X. Mi *et al.*, Information scrambling in computationally complex quantum circuits, [arXiv:2101.08870](https://arxiv.org/abs/2101.08870) (2021).
- [47] P. Delsarte, J. M. Goethals, and J. J. Seidel, Bounds for systems of lines, and Jacobi polynomials, *Philips Res. Rep.* **30**, 91 (1975).
- [48] P. Delsarte, J. M. Goethals, and J. J. Seidel, Spherical codes and designs, *Geom. Dedicata* **6**, 363 (1977).
- [49] A. Roy and A. J. Scott, Unitary designs and codes, *Des. Code Cryptogr.* **53**, 13 (2009).
- [50] A. Roy, Bounds for codes and designs in complex subspaces, *J. Algebraic Combin.* **31**, 1 (2010).
- [51] A. Roy and S. Suda, Complex spherical designs and codes, *J. Comb. Des.* **22**, 105 (2011).
- [52] E. Bannai, G. Navarro, N. Rizo, and P. H. Tiep, Unitary  $t$ -groups, *J. Math. Soc. Japan* **72**, 909 (2020).
- [53] D. P. DiVincenzo, D. W. Leung, and B. M. Terhal, Quantum data hiding, *IEEE Trans. Inf. Theory* **48**, 580 (2002).
- [54] W. G. Brown, Y. S. Weinstein, and L. Viola, Quantum pseudorandomness from cluster-state quantum computation, *Phys. Rev. A* **77**, 040303(R) (2008).
- [55] Y. S. Weinstein, W. G. Brown, and L. Viola, Parameters of pseudorandom quantum circuits, *Phys. Rev. A* **78**, 052332 (2008).
- [56] D. Gross, K. Audenaert, and J. Eisert, Evenly distributed unitaries: On the structure of unitary designs, *J. Math. Phys.* **48**, 052104 (2007).
- [57] G. Tóth and J. J. García-Ripoll, Efficient algorithm for multiqubit twirling for ensemble quantum computation, *Phys. Rev. A* **75**, 042311 (2007).
- [58] C. Dankert, R. Cleve, J. Emerson, and E. Livine, Exact and approximate unitary 2-designs and their application to fidelity estimation, *Phys. Rev. A* **80**, 012304 (2009).
- [59] A. W. Harrow and R. A. Low, Random quantum circuits are approximate 2-designs, *Commun. Math. Phys.* **291**, 257 (2009).
- [60] I. T. Diniz and D. Jonathan, Comment on “random quantum circuits are approximate 2-designs”, *Commun. Math. Phys.* **304**, 281 (2011).
- [61] R. Cleve, D. Leung, L. Liu, and C. Wang, Near-linear constructions of exact unitary 2-designs, *Quant. Info. Comp.* **16**, 0721 (2016).
- [62] Y. Nakata, C. Hirche, C. Morgan, and A. Winter, Unitary 2-designs from random X- and Z-diagonal unitaries, *J. Math. Phys.* **58**, 052203 (2017).
- [63] A. W. Harrow and R. A. Low, in *Proc. RANDOM'09*, Lecture Notes in Computer Science (Springer, Berlin Heidelberg, 2009), p. 548.
- [64] F. G. S. L. Brandão, A. W. Harrow, and M. Horodecki, Local random quantum circuits are approximate polynomial-designs, *Commun. Math. Phys.* **346**, 397 (2016).
- [65] Y. Nakata, C. Hirche, M. Koashi, and A. Winter, Efficient Quantum Pseudorandomness with Nearly Time-Independent Hamiltonian Dynamics, *Phys. Rev. X* **7**, 021006 (2017).
- [66] A. Harrow and S. Mehraban, Approximate unitary  $t$ -designs by short random quantum circuits using nearest-neighbor and long-range gates, [arXiv:1809.06957](https://arxiv.org/abs/1809.06957) (2018).
- [67] J. Haferkamp *et al.*, Quantum homeopathy works: Efficient unitary designs with a system-size independent number of non-Clifford gates, [arXiv:2002.09524](https://arxiv.org/abs/2002.09524) (2020).
- [68] E. Bannai, M. Nakahara, D. Zhao, and Y. Zhu, On the explicit constructions of certain unitary  $t$ -designs, *J. Phys. A: Math. Theor.* **52**, 495301 (2019).
- [69] R. Harper, I. Hincks, C. Ferrie, S. T. Flammia, and J. J. Wallman, Statistical analysis of randomized benchmarking, *Phys. Rev. A* **99**, 052350 (2019).
- [70] P. D. Seymour and T. Zaslavsky, Averaging sets: A generalization of mean values and spherical designs, *Adv. Math.* **52**, 213 (1984).
- [71] P. Rabau and B. Bajnok, Bounds for the number of nodes in chebyshev type quadrature formulas, *J. Approx. Theory* **67**, 199 (1991).
- [72] G. Wagner and B. Volkmann, On averaging sets, *Mon. Hefte. Math.* **111**, 69 (1991).
- [73] E. Bannai and E. Bannai, A survey on spherical designs and algebraic combinatorics on spheres, *European J. Combin.* **30**, 1392 (2009).

- [74] A. Bondarenko, D. Radchenko, and M. Viazovska, Optimal asymptotic bounds for spherical designs, *Ann. Math.* **178**, 443 (2013).
- [75] Z. Cui, J. Xia, and Z. Xiang, Rational designs, *Adv. Math.* **352**, 541 (2019).
- [76] Z. Xiang, Explicit spherical designs (to be published, 2020).
- [77] E. Bannai, Y. Nakata, T. Okuda, and D. Zhao, Explicit construction of exact unitary designs, [arXiv:2009.11170](https://arxiv.org/abs/2009.11170) (2020).
- [78] R. A. Low, PhD thesis, University of Bristol, 2010. [arXiv:1006.5227](https://arxiv.org/abs/1006.5227).
- [79] D. Bump, *Lie Groups* (Springer-Verlag, New York, 2004).
- [80] R. Goodman and N. R. Wallach, *Symmetry, Representations, and Invariants* (Springer, Dordrecht, 2009).
- [81] A. T. James and A. G. Constantine, Generalized Jacobi polynomials as spherical functions of the grassmann manifold, *Proc. London Math. Soc.* **s3–29**, 174 (1974).
- [82] P. Chaugule *et al.*, Schur polynomials do not have small formulas if the determinant doesn't!, [arXiv:1911.12520](https://arxiv.org/abs/1911.12520) (2019).
- [83] F. Mezzadri, How to generate random matrices from the classical compact groups, *Not. Am. Math. Soc.* **54**, 592 (2007).
- [84] 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).
- [85] 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).
- [86] Y. Suzuki, K. Fujii, and M. Koashi, Efficient Simulation of Quantum Error Correction under Coherent Error Based on the Nonunitary Free-Fermionic Formalism, *Phys. Rev. Lett.* **119**, 190503 (2017).
- [87] S. Bravyi, M. Englbrecht, R. König, and N. Peard, Correcting coherent errors with surface codes, *npj Quantum Inf.* **4**, 55 (2018).
- [88] A. S. Darmawan and D. Poulin, Tensor-Network Simulations of the Surface Code under Realistic Noise, *Phys. Rev. Lett.* **119**, 040502 (2017).
- [89] J. Wallman, C. Granade, R. Harper, and S. T. Flammia, Estimating the coherence of noise, *New J. Phys.* **17**, 113020 (2015).
- [90] C. D. Wilen *et al.*, Correlated charge noise and relaxation errors in superconducting qubits, *Nature* **594**, 369 (2021).
- [91] J. Han *et al.*, Characterizing noise correlation and enhancing coherence via qubit motion, *Fundam. Res.* **1**, 10 (2021).
- [92] M. McEwen *et al.*, Removing leakage-induced correlated errors in superconducting quantum error correction, *Nat. Commun.* **12**, 1761 (2021).
- [93] Matt McEwen *et al.*, Resolving catastrophic error bursts from cosmic rays in large arrays of superconducting qubits, [arXiv:2104.05219](https://arxiv.org/abs/2104.05219) (2021).
- [94] S. Tamate *et al.*, Scalable packaging design for large-scale superconducting quantum circuit, *Bull. Am. Phys. Soc.* (2021).
- [95] J. Gambetta *et al.*, Qubit-photon interactions in a cavity: Measurement-induced dephasing and number splitting, *Phys. Rev. A* **74**, 042318 (2006).
- [96] R. A. Low, Large deviation bounds for  $k$ -designs, *Proc. R. Soc. A* **465**, 3289 (2009).
- [97] M. Takeuchi, *Modern Spherical Functions* (American Mathematical Society, Providence, R.I., 1994).
- [98] J. A. Wolf, *Harmonic Analysis on Commutative Spaces* (American Mathematical Society, Providence, R.I., 2007).
- [99] B. Dirkse, J. Helsen, and S. Wehner, Efficient unitarity randomized benchmarking of few-qubit clifford gates, *Phys. Rev. A* **99**, 012315 (2019).
- [100] J. Helsen, J. J. Wallman, and S. Wehner, Representations of the multi-qubit clifford group, *J. Math. Phys.* **59**, 072201 (2018).
- [101] E. Lucero *et al.*, Reduced phase error through optimized control of a superconducting qubit, *Phys. Rev. A* **82**, 042339 (2010).
- [102] D. C. McKay, C. J. Wood, S. Sheldon, J. M. Chow, and J. M. Gambetta, Efficient z gates for quantum computing, *Phys. Rev. A* **96**, 022330 (2017).
- [103] J. Y. Mutus *et al.*, Strong environmental coupling in a Josephson parametric amplifier, *Appl. Phys. Lett.* **104**, 263513 (2014).