MIT Open Access Articles

*Experimentally efficient methods for estimating
the performance of quantum measurements*

The MIT Faculty has made this article openly available. **Please share**
how this access benefits you. Your story matters.

Citation: Magesan, Easwar, and Paola Cappellaro. "Experimentally efficient methods for estimating the performance of quantum measurements." *Physical Review A* 88, no. 2 (August 2013). © 2013 American Physical Society

As Published: <http://dx.doi.org/10.1103/PhysRevA.88.022127>

Publisher: American Physical Society

Persistent URL: <http://hdl.handle.net/1721.1/81395>

Version: Final published version: final published article, as it appeared in a journal, conference proceedings, or other formally published context

Terms of Use: Article is made available in accordance with the publisher's policy and may be subject to US copyright law. Please refer to the publisher's site for terms of use.

Experimentally efficient methods for estimating the performance of quantum measurements

Easwar Magesan and Paola Cappellaro

Research Laboratory of Electronics and Nuclear Science and Engineering Department, MIT, Cambridge Massachusetts 02139, USA

(Received 5 February 2013; published 29 August 2013)

Efficient methods for characterizing the performance of quantum measurements are important in the experimental quantum sciences. Ideally, one requires both a physically relevant distinguishability measure between measurement operations and a well-defined experimental procedure for estimating the distinguishability measure. Here, we propose the average measurement fidelity and error between quantum measurements as distinguishability measures. We present protocols for obtaining bounds on these quantities that are both estimable using experimentally accessible quantities and scalable in the size of the quantum system. We also explain why the bounds should be valid in large generality and illustrate the method via numerical examples.

DOI: [10.1103/PhysRevA.88.022127](https://doi.org/10.1103/PhysRevA.88.022127)

PACS number(s): 03.65.Wj, 03.67.Lx, 03.67.Ac

I. INTRODUCTION

Measurement plays a fundamental role in the quantum sciences, as it allows one to evolve and extract information from a quantum system. From a practical perspective, measurements are imperative for the successful implementation of a wide variety of protocols in quantum information processing (QIP) and communication [1], as well as high-precision metrology [2]. If large-scale QIP is achieved experimentally, certain tasks that are conjectured to be classically inefficient will become realizable using a scalable number of resources [3,4].

Various models of computation have been proposed for QIP, such as the standard circuit model [5], measurement-based quantum computation [6], and topological quantum computation [7]. Each of these models relies greatly on the ability to perform accurate measurements. A large number of physical systems have been proposed as candidates for implementing QIP. A short, and certainly nonexhaustive, list of systems includes superconducting circuits [8,9], nitrogen-vacancy centers [10], trapped ions [11], NMR [12], quantum dots [13], and optical implementations [14]. Measurement schemes vary greatly across these systems and so it is important to have straightforward protocols for comparing measurements that are both independent of the particular type of implementation and scalable in the size of the system. Providing such protocols is the main goal of this paper.

There has been a significant amount of research aimed at *completely characterizing* the error on the operations used to process and measure quantum information via quantum process [15–17] and measurement tomography [18,19]. In principle, these tomographic methods can be used to characterize the error affecting any quantum operation or measurement realized in an experimental setting. Unfortunately, there are various significant drawbacks to complete process and measurement tomography. Process tomography requires an exponential number of resources in the number n of quantum bits (qubits) that comprise the system (the number of parameters required to just describe the process scales as 16^n), is not robust against state-preparation and measurement errors [20], and requires intensive classical postprocessing of measurement data [21]. Not surprisingly, since measurement tomography is realized by essentially inverting process tomography [18,19], it suffers from drawbacks similar to those of quantum process tomography. In particular, complete measurement tomography

requires the ability to prepare a complicated set of pure input states with extremely high precision (possibly by performing complex unitary gates), scales badly in the size of the system, and can require lengthy postprocessing of the tomographic data.

In many cases, one may only be interested in a subset of parameters characterizing the noise or in determining the *strength* of the noise process rather than the entire process itself. For instance, in fault tolerance [22–25], as long as the strength of the noise affecting state preparation, gates, and measurements is bounded by some value, large-scale computation is possible. As a result, various methods for *partially characterizing* the error of quantum gates [26–35] have recently been proposed which attempt to circumvent many of the problems associated with full process tomography. The error is defined via the “quantum average gate fidelity,” which is derived from the quantum channel fidelity between two quantum processes. Since quantum gates are unitary operations, the average gate fidelity and higher order moments take a simple form and can be calculated analytically [36–38].

Similarly, *partial* estimation of measurement errors by methods more efficient than full measurement tomography would be extremely valuable but have yet to be considered in significant detail. There are a variety of questions to consider in this endeavor, such as how to deal with the abundance of different measurement schemes across the various implementations and the ambiguity regarding what the single parameter characterizing the error should be (see Ref. [39] for two proposals of such distance measures). Since measurements are nonunitary processes, the quantum channel fidelity between the measurement and the noise process is a very complex quantity and cannot be calculated analytically as is the case for quantum gates.

One of the main goals of this paper is to initiate research into scalable verification and characterization of measurement devices. Here, we discuss methods for characterizing the error of a quantum measurement and present protocols for estimating the performance of a measurement with respect to these error measures. The protocols are scalable in the number of qubits comprising the system. In addition, the protocols are very general and should be straightforward to implement in most systems. They require the ability to prepare a random state from the basis that constitutes the ideal measurement and perform the noisy measurement (perhaps consecutively).

As quantum systems continue to scale to larger sizes, and full tomography of the measurement becomes infeasible, the protocols provided here can be used as fast and straightforward methods to verify that the performance of a measuring device achieves some desired threshold.

Specifically, we propose the average measurement fidelity and error rate of a quantum measurement as natural measures of the error and provide scalable, experimentally implementable protocols for estimating these quantities. The output of the protocol is a *lower (upper) bound* on the average fidelity (error rate) of the measurement. Computing these bounds from the measurement data is straightforward, and generic estimates of how many trials need to be performed can be obtained using statistical methods. We provide a direct comparison of the time complexity between our protocol and that of completely reconstructing the noisy positive operator-valued measure (POVM) elements via tomography. Our method scales independently of d (where d is the dimension of the system, e.g., $d = 2^n$ for an n -qubit system), while a complete reconstruction scales as $O(d^3)$. Thus, while our method gives much less information than full tomography, it is scalable, is straightforward to implement, and involves less postprocessing of measurement data. Ideally, the protocols we provide here will be useful as a starting point for further analysis and research into more efficient tomographic methods of quantum measurements.

We focus mainly on the case where we attempt to implement a finite-dimensional projective measurement [also known as a projective-valued measure (PVM)] but, instead, implement a general quantum measurement in the form of a POVM. In particular, we analyze the following two scenarios.

(1) The ideal measurement is a rank 1 PVM and we are only concerned with characterizing measurement probabilities.

(2) The ideal measurement is a rank 1 PVM and we are concerned with simultaneously characterizing measurement probabilities and output states.

Rank 1 PVMs are significant in QIP since each computational model mentioned above can achieve universal quantum computation using rank 1 PVMs. We leave extending our results to arbitrary-rank PVMs as a direction of future research, however, we anticipate that the results presented here can be carried over to these measurements. Moreover, while the most general measurement model is a POVM, any POVM can be realized as a PVM on an extended Hilbert space by attaching ancilla systems to the Hilbert space of interest [40]. Thus, characterizing the quality of PVMs can potentially provide direct information about one's ability to implement POVMs.

The structure of the presentation is as follows. First, in Sec. II we set notation, discuss the framework of PVMs and POVMs, and define the noisy measurement models. Next, in Sec. III, we define the *quantum average measurement fidelity* and *quantum average measurement error*, which are the operationally relevant metrics we use to compare measurement operations. In Sec. IV we present the main result for the first of the two scenarios listed above. We provide an expression for a lower (upper) bound on the fidelity (error rate) and present an experimental protocol for obtaining this bound. In Sec. V we give a derivation of the bound and show that it satisfies certain necessary conditions to be useful in practice. Section VI provides a discussion of the general validity of the bound,

and numerical examples for the case of a single-qubit system are analyzed. Our derived bounds are shown to be valid for *every* example analyzed, confirming the discussion regarding the validity of the method. We discuss the extension of the method to rank 1 PVMs with output states in Sec. VII and provide the experimental protocol for this case in Sec. VIII. The resource analysis for the protocols are provided in Sec. IX and concluding remarks are made in Sec. X. The reader interested in the bounds and protocols without most of the technical details is referred to the text surrounding Eq. (5.1) for rank 1 PVMs and the text surrounding Eq. (7.14) for rank 1 PVMs with output states.

By definition, the average fidelity and error rate of a quantum measurement are equivalent (see Sec. III). Hence, the majority of the presentation is in terms of the average fidelity, however, it is important to keep in mind that analogous results hold for the error rate.

II. MEASUREMENT OF QUANTUM SYSTEMS AND THE ERROR MODEL

Denote the quantum system $\mathcal{S}$ and suppose that it is represented by a Hilbert space $\mathcal{H}$ of dimension $d < \infty$ ($d = 2^n$ for an n -qubit system). The most general type of measurement one can make is a POVM. A POVM consists of a set $\{E_k\}$ of linear operators on $\mathcal{H}$ that satisfy

$$E_k \geq 0, \quad \sum_k E_k = \mathbb{1}. \quad (2.1)$$

If the input state to the measurement is σ , the probability of obtaining outcome “ k ” is given by

$$r_k = \text{tr}(E_k \sigma). \quad (2.2)$$

An important subset of POVM measurements is PVMs, which correspond to the case of each E_k being equal to a projection operator Π_k . Hence, in addition to the conditions listed in Eq. (2.1), $\Pi_k^2 = \Pi_k$ for each k .

PVMs are a very general class of quantum measurements and, as previously mentioned, are sufficient for performing universal quantum computation in various computational models [5,6]. For instance, computational basis measurements allow for universality in the standard circuit model and single-qubit projective measurements provide universality in MBQC. Other important examples of PVMs in quantum information theory are parity measurements, which are used extensively in quantum error correction and fault tolerance [41–44]. It is important to note that, by Naimark's theorem [40], any POVM can be implemented via a PVM on a larger Hilbert space. Hence, methods for determining error rates of PVMs can also give direct information regarding the error associated with performing general POVM measurements.

PVMs are in a 1-1 correspondence with observables (Hermitian operators) $\mathcal{O}$, where $\mathcal{O}$ is nondegenerate if and only if the PVM consists of rank 1 projective elements. We write the observable $\mathcal{O}$ as

$$\mathcal{O} = \sum_{k=1}^b \lambda_k \Pi_k, \quad (2.3)$$

where the eigenvalues λ_k correspond to the measurement values and $\Pi_k := \sum_{j=1}^{d_k} |\psi_j^k\rangle\langle\psi_j^k|$ is the rank d_k projector onto the eigenspace spanned by the eigenvectors $|\psi_j^k\rangle$. Hence, $\sum_{k=1}^b d_k = d$ and the PVM is rank 1 if and only if $d_k = 1$ for every k . If $\mathcal{M}$ represents the ideal PVM without postselection, then $\mathcal{M}$ has the following action on each input state σ ,

$$\mathcal{M}(\sigma) = \sum_{k=1}^b \Pi_k \sigma \Pi_k = \sum_{k=1}^b p_k \left[\frac{\Pi_k \sigma \Pi_k}{\text{tr}(\Pi_k \sigma)} \right], \quad (2.4)$$

where we have defined

$$p_k = p_k(\sigma) = \text{tr}(\Pi_k \sigma). \quad (2.5)$$

In the case of analyzing only measurement probabilities, the PVM is a mapping from the set of quantum states to the probability vectors $(p_1, \dots, p_b)$,

$$\mathcal{M}(\sigma) = (p_1(\sigma), \dots, p_b(\sigma)) = (\text{tr}(\Pi_1 \sigma), \dots, \text{tr}(\Pi_b \sigma)). \quad (2.6)$$

This case is especially interesting since in many scenarios one is mainly interested in the output *value* of the measurement rather than the output state itself. For instance, the output of a quantum algorithm is usually the value of a measurement in the computational basis. Hence, as long as the measurement values are obtained with correct probabilities, the measurement is deemed successful.

In addition, in many current implementations of measurements, the action of performing the measurement operation to obtain the output value destroys or drastically alters the state of the system. For instance in photodetection, which forms the basis of various measurement schemes in optical, atomic, and superconducting systems, the measurement consists of recording the occurrence of a photon. If the state of the system is encoded into a degree of freedom (mode) of the photon, such as polarization or frequency, detection of the photon can record the output of a measurement of this degree of freedom. In common photodetectors, such as avalanche photodiodes, the photon is lost in the process of creating a current through the photoelectric effect. Hence, the system encoding the information to be measured is destroyed, but the measurement output can be accessed.

An example where the system is not destroyed yet the state is not preserved from the measurement is the nitrogen-vacancy center in diamond [10]. In this case, the processes of measurement and ground-state polarization are identical. Hence, while one can obtain the relevant measurement outcomes and statistics, the output state of the measurement is always the ground state. There are situations where preserving the measurement output state is useful. For instance, in any paradigm where quantum information is evolved by performing measurements, such as in MBQC [6], one requires the state of the system to be well preserved under measurements. In MBQC, quantum information is evolved by performing single-qubit measurements on a highly entangled initial state. Any degradation of the state or system caused by the measurement process can be highly detrimental to obtaining a high-fidelity output.

In the general case of Eq. (2.4), the noisy measurement $\mathcal{E}$ is modeled by

$$\mathcal{E}(\sigma) = \sum_{k=1}^b r_k(\sigma) \rho_k(\sigma), \quad (2.7)$$

where we allow both the noisy measurement probabilities r_k and the output states ρ_k to be functions of σ . By Eq. (2.4), $r_k(\sigma)$ and $\rho_k(\sigma)$ are ideally given by $p_k(\sigma) = \text{tr}(\Pi_k \sigma)$ and $\frac{\Pi_k \sigma \Pi_k}{\text{tr}(\Pi_k \sigma)}$, respectively. If we analyze only measurement probabilities, then

$$\mathcal{E}(\sigma) = (r_1(\sigma), \dots, r_b(\sigma)) = (\text{tr}(E_1 \sigma), \dots, \text{tr}(E_b \sigma)), \quad (2.8)$$

where $\mathcal{E}$ is allowed to be of a completely general form by assuming that it is modeled by a POVM $\{E_k\}_{k=1}^b$. The first scenario we are interested in is to compare the probability distributions in Eqs. (2.6) and (2.8). We analyze the more general case of measurement probabilities and output states [i.e., comparing Eqs. (2.4) and (2.7)] in Sec. VII. We first discuss the figure of merits we use to compare the ideal measurement process $\mathcal{M}$ and noisy measurement process $\mathcal{E}$.

III. QUANTUM AVERAGE MEASUREMENT FIDELITY AND ERROR

A natural question is, How should we compare the ideal and actual measurements $\mathcal{M}$ and $\mathcal{E}$? A set of criteria that a distance measure, Δ , for comparing ideal and real quantum processes should satisfy has been given previously [45]. Currently, no known Δ satisfies all of these criteria simultaneously. Thus, one must settle for Δ to satisfy a subset of these criteria, in addition to other criteria that may be useful for the particular task at hand.

As mentioned previously, the average gate fidelity $\overline{\mathcal{F}_{\mathcal{E}, \mathcal{U}}}$ is a useful method for comparing an intended unitary operation $\mathcal{U}$ and actual quantum process $\mathcal{E}$. There are various reasons for the utility of the average gate fidelity; for instance, it satisfies the following properties.

- (1) There is a straightforward method for evaluating $\overline{\mathcal{F}_{\mathcal{E}, \mathcal{U}}}$ (given a description of $\mathcal{U}$ and $\mathcal{E}$).
- (2) $\overline{\mathcal{F}_{\mathcal{E}, \mathcal{U}}}$ has a well-motivated physical interpretation.
- (3) All states are taken into account in an unbiased manner when calculating $\overline{\mathcal{F}_{\mathcal{E}, \mathcal{U}}}$.
- (4) $\overline{\mathcal{F}_{\mathcal{E}, \mathcal{U}}}$ is experimentally accessible via efficient protocols.

An important drawback of the average gate fidelity is that it is not a metric. We would like similar properties to hold for our method of comparing ideal and real measurements. Let us briefly outline how the average gate fidelity is derived from more general quantities, which will provide intuition for how to define our method for comparing measurements.

The average gate fidelity is derived from the state-dependent quantum channel fidelity, which is a standard method for comparing quantum operations. If $\mathcal{E}_1$ and $\mathcal{E}_2$ are quantum operations and σ is a quantum state, the quantum channel fidelity between $\mathcal{E}_1$ and $\mathcal{E}_2$, denoted $F_{\mathcal{E}_1, \mathcal{E}_2}$, is given by the standard state fidelity between $\mathcal{E}_1(\sigma)$ and $\mathcal{E}_2(\sigma)$:

$$F_{\mathcal{E}_1, \mathcal{E}_2}(\sigma) = F(\mathcal{E}_1(\sigma), \mathcal{E}_2(\sigma)) = \left(\text{tr} \sqrt{\sqrt{\mathcal{E}_1(\sigma)} \mathcal{E}_2(\sigma) \sqrt{\mathcal{E}_1(\sigma)}} \right)^2. \quad (3.1)$$

When one of the operations in Eq. (3.1) is unitary (say $\mathcal{E}_2 = \mathcal{U}$), the channel fidelity is called the quantum gate fidelity, and when σ is pure ($\sigma = |\psi\rangle\langle\psi|$), the gate fidelity takes the extremely simple form

$$F_{\mathcal{E}_1, \mathcal{U}}(|\psi\rangle\langle\psi|) = \text{tr}(\mathcal{E}_1(|\psi\rangle\langle\psi|) \mathcal{U}(|\psi\rangle\langle\psi|)) \\ = \langle\psi|\Lambda(|\psi\rangle\langle\psi|)|\psi\rangle, \quad (3.2)$$

where

$$\Lambda = \mathcal{U}^\dagger \circ \mathcal{E}_1. \quad (3.3)$$

The *average quantum gate fidelity*, denoted $\overline{F_{\mathcal{E}_1, \mathcal{U}}}$, is obtained by integrating over all pure input states. The integral is taken over the unitarily invariant Haar measure (also known as the Fubini-Study measure) on the set of pure states [46]. In this paper we denote the Fubini-Study measure by μ . This gives

$$\overline{F_{\mathcal{E}_1, \mathcal{U}}} = \int \text{tr}(\mathcal{E}_1(|\psi\rangle\langle\psi|) \mathcal{U}(|\psi\rangle\langle\psi|)) d\psi \\ = \int \langle\psi|\Lambda(|\psi\rangle\langle\psi|)|\psi\rangle d\psi = \frac{\sum_j \text{tr}(A_j) \text{tr}(A_j^\dagger) + d}{d^2 + d}, \quad (3.4)$$

where $\{A_j\}$ is any set of Kraus operators for Λ [36]. Thus, the average gate fidelity reduces to an extremely simple form because one of the operations is unitary.

Following this intuition, Eqs. (2.4), (2.7), and (3.1) imply that for the ideal ($\mathcal{M}$) and noisy ($\mathcal{E}$) measurements.

$$F(\mathcal{E}(\sigma), \mathcal{M}(\sigma)) \\ = \left(\text{tr} \sqrt{\sqrt{\sum_{k=1}^b r_k \rho_k} \left(\sum_{k=1}^b p_k \left[\frac{\Pi_k \sigma \Pi_k}{\text{tr}(\Pi_k \sigma)} \right] \right) \sqrt{\sum_{k=1}^b r_k \rho_k}} \right)^2, \quad (3.5)$$

where the state dependence in p_k , r_k , ρ_k is omitted for notational convenience. The operational significance of this quantity comes from the fact that it is just the standard fidelity between the two quantum states $\mathcal{E}(\sigma)$ and $\mathcal{M}(\sigma)$. More precisely, Eq. (3.5) is related to the maximum distinguishability between probability distributions one could obtain using any POVM measurement on states $\mathcal{E}(\sigma)$ and $\mathcal{M}(\sigma)$ [47].

Unlike the unitary case in Eq. (3.2), such an expression does not reduce to a simple form in general. However, ideally, the figure of merit we use to distinguish $\mathcal{M}$ and $\mathcal{E}$ will contain evenly weighted information from *all* possible input states and also have direct operational significance. As a result, we define the *quantum average measurement fidelity*, denoted $\overline{F_{\mathcal{E}, \mathcal{M}}}$, to be the Haar integral of Eq. (3.5) over pure input states:

$$\overline{F_{\mathcal{E}, \mathcal{M}}} = \int F(\mathcal{E}(|\psi\rangle\langle\psi|), \mathcal{M}(|\psi\rangle\langle\psi|)) d\psi. \quad (3.6)$$

We also define

$$r_{\mathcal{E}, \mathcal{M}} = 1 - \overline{F_{\mathcal{E}, \mathcal{M}}} \quad (3.7)$$

to be the *quantum average measurement error*. In the case of only analyzing measurement probabilities, that is, $\mathcal{M}$ and $\mathcal{E}$ are given by Eqs. (2.6) and (2.8), respectively, the states

associated with p_k and r_k can be taken to be Π_k , which gives

$$\overline{F_{\mathcal{E}, \mathcal{M}}} = \int \left(\sum_{k=1}^b \sqrt{p_k r_k} \right)^2 d\psi \\ = \int \left(\sum_{k=1}^b \sqrt{\text{tr}(\Pi_k |\psi\rangle\langle\psi|) \text{tr}(E_k |\psi\rangle\langle\psi|)} \right)^2 d\psi, \quad (3.8) \\ r_{\mathcal{E}, \mathcal{M}} = 1 - \int \left(\sum_{k=1}^b \sqrt{\text{tr}(\Pi_k |\psi\rangle\langle\psi|) \text{tr}(E_k |\psi\rangle\langle\psi|)} \right)^2 d\psi. \quad (3.9)$$

The direct relationship between $r_{\mathcal{E}, \mathcal{M}}$ and $\overline{F_{\mathcal{E}, \mathcal{M}}}$ implies that statements and bounds proven about one directly apply to the other. Thus, as mentioned in Sec. I, we phrase the majority of the discussion in terms of $\overline{F_{\mathcal{E}, \mathcal{M}}}$. Since $\overline{F_{\mathcal{E}, \mathcal{M}}}$ does not have a simple form like the average gate fidelity in Eq. (3.4), our goal is to provide efficient methods for estimating $\overline{F_{\mathcal{E}, \mathcal{M}}}$ that can be experimentally implemented in a simple manner.

An important reason for using the average measurement fidelity is that it provides a state-independent distance measure which can be connected to the diamond norm distance [48] between quantum operations [49]. Computing the diamond norm is an exponentially hard task since one needs a complete description of the quantum operations. Thus, the diamond norm is neither straightforward to calculate nor experimentally accessible. However, it is commonly used in fault-tolerant analyses of threshold error rates of physical operations. Thus, information about the diamond norm provided by the average measurement error defined above can potentially provide information regarding the ability to perform fault-tolerant computation.

IV. EXPERIMENTAL PROTOCOL: RANK 1 PVMs

In this section, we look at obtaining a lower bound on $\overline{F_{\mathcal{E}, \mathcal{M}}}$ for measurement probabilities of rank 1 PVMs. Again, we note that upper bounds on $r_{\mathcal{E}, \mathcal{M}}$ are equivalent to lower bounds on $\overline{F_{\mathcal{E}, \mathcal{M}}}$, and so, for both clarity and consistency, we phrase the following discussion in terms of only $\overline{F_{\mathcal{E}, \mathcal{M}}}$.

There are d PVM elements, $\{\Pi_1, \dots, \Pi_d\}$, each of which is a rank 1 projection operator. From Eq. (2.8), the outcome of the measurement of a state $|\psi\rangle\langle\psi|$ can be associated with a member from $\{1, \dots, d\}$ with frequency distribution

$$(r_1, \dots, r_d) = (\text{tr}(E_1 |\psi\rangle\langle\psi|), \dots, \text{tr}(E_d |\psi\rangle\langle\psi|)), \quad (4.1)$$

where the E_j represent the noisy POVM elements.

Our expression for the lower bound is given by

$$lb = \frac{1 + d\overline{X}}{1 + d}, \quad (4.2)$$

where

$$\overline{X} = \frac{1}{d^2} \sum_{(l, m) \in \mathcal{D}} \sqrt{u_l u_m}, \quad (4.3)$$

$$u_l = \text{tr}(\Pi_l E_l), \quad (4.4)$$

and

$$\mathcal{D} = \{0, \dots, d-1\} \times \{0, \dots, d-1\}. \quad (4.5)$$

The u_l measure the overlap between the l th ideal PVM element and the l th noisy POVM element. Thus, lb can intuitively be thought of as a parameter that measures how well ideal PVM elements are preserved when input to the noisy measurements (with the inclusion of dimensional factors).

Goal: Obtain a lower bound, lb , for $\overline{F_{\mathcal{E}, \mathcal{M}}}$ as defined in Eq. (3.9).

The experimental protocol to obtain lb is as follows.

Experimental protocol

Step 1: Choose a pair of indices (l, m) uniformly at random from $\mathcal{D}$.

Step 2: Independently, for both $j = l$ and $j = m$,

(a) prepare the quantum state $|\psi_j\rangle$, perform the noisy measurement $\mathcal{E}$ on $\Pi_j = |\psi_j\rangle\langle\psi_j|$, and record whether outcome “ j ” is obtained; and

(b) repeat step 2a many times and denote the frequency of obtaining “ j ” by $\hat{u}_j$, that is, $\hat{u}_j$ is an estimator of $u_j = \text{tr}(\Pi_j E_j)$ (See Sec. IX A for a discussion of the number of repetitions required to estimate u_j to a desired accuracy and confidence).

Step 3: Repeat steps 1 and 2 K times, where K is dictated by the desired accuracy and confidence in estimating lb . (See Sec. IX B for a discussion of the size of K .)

Step 4: Compute an estimator $\hat{lb}$ for the lower bound lb of $\overline{F_{\mathcal{E}, \mathcal{M}}}$ defined in Eq. (5.1) via the formula

$$\hat{lb} = \frac{1 + d\hat{X}}{1 + d}, \quad (4.6)$$

where

$$\hat{X} = \frac{1}{K} \sum_{(k_1, k_2)} \sqrt{\hat{u}_{k_1} \hat{u}_{k_2}} \quad (4.7)$$

is an estimator of $\overline{X}$ defined in Eq. (4.3) and the $\{(1_1, 1_2), \dots, (K_1, K_2)\}$ are the K trials dictated by step 3.

This concludes the protocol.

There are various important points about the protocol that should be emphasized. First, the number of trials required in steps 2b and 3 is independent of d . Thus, the time complexity of the entire protocol is independent of d and depends only on the desired accuracy and confidence of the estimate $\hat{lb}$ of lb (see Sec. IX). Second, lb can be estimated from the above protocol using only

- (i) applications of the noisy measurement and
- (ii) the ability to prepare the d pure input states $|\psi_j\rangle$.

Finally, it is straightforward to show that the following two properties of lb hold (see Sec. V D).

- (1) In the limit of $F_{\mathcal{E}, \mathcal{M}} \uparrow 1$,

$$lb \uparrow 1. \quad (4.8)$$

(2) lb scales well in d ; that is, if each $u_k = \text{tr}(\Pi_k E_k)$ is of the order of $1 - \delta$, then, as $d \rightarrow \infty$, $lb \rightarrow 1 - \delta$ (and does not converge to 0 or some other small constant).

These are clearly necessary conditions for lb to be a good lower bound on $\overline{F_{\mathcal{E}, \mathcal{M}}}$. Property 1 implies that, loosely

speaking, in the small error limit lb can be taken as an estimate of $\overline{F_{\mathcal{E}, \mathcal{M}}}$. Property 2 implies that as $d \rightarrow \infty$, the lower bound remains of the order of $1 - \delta$. This is important because if the lower bound generically converges to 0 (or some other constant) as $d \rightarrow \infty$, then the lower bound is ineffective at providing useful information about the noise.

V. DERIVATION OF THE LOWER BOUND

We need to show that

$$lb := \frac{1 + \overline{X}d}{1 + d} = \frac{d + \sum_{k=1}^d u_k + \sum_{l \neq m} \sqrt{u_l u_m}}{d(d+1)} \quad (5.1)$$

is a lower bound for $\overline{F_{\mathcal{E}, \mathcal{M}}}$. First, from Eq. (3.8),

$$F(\mathcal{E}(\sigma), \mathcal{M}(\sigma)) = \left(\sum_k \sqrt{r_k p_k} \right)^2, \quad (5.2)$$

where the state dependence is implicit in the p_k and r_k . Taking the integral over all pure states gives

$$\begin{aligned} & \int F(\mathcal{E}(|\psi\rangle\langle\psi|), \mathcal{M}(|\psi\rangle\langle\psi|)) d\psi \\ &= \sum_{k=1}^d \left[\int r_k p_k d\psi \right] + \sum_{l \neq m} \left[\int \sqrt{r_l r_m p_l p_m} d\psi \right]. \end{aligned} \quad (5.3)$$

With these tools in hand, let us look at the sums in Eq. (5.3) separately.

$$\text{A. } \sum_{k=1}^d \int r_k p_k d\psi$$

First,

$$\begin{aligned} \int r_k p_k d\psi &= \int \text{tr}(E_k |\psi\rangle\langle\psi|) \text{tr}(\Pi_k |\psi\rangle\langle\psi|) d\psi \\ &= \int \text{tr}([E_k \otimes \Pi_k] |\psi\rangle\langle\psi| \otimes |\psi\rangle\langle\psi|) d\psi \\ &= \text{tr} \left([E_k \otimes \Pi_k] \int |\psi\rangle\langle\psi|^{\otimes 2} d\psi \right). \end{aligned} \quad (5.4)$$

To compute this integral we use Schur's lemma, which states that, for any positive integer t [38,50],

$$\int |\psi\rangle\langle\psi|^{\otimes t} d\psi = \frac{\Pi_{\text{sym}}(t, d)}{\text{tr}[\Pi_{\text{sym}}(t, d)]}, \quad (5.5)$$

where $\Pi_{\text{sym}}(t, d)$ is the projector onto the symmetric subspace of the t -partite Hilbert space (whose t factor spaces each have dimension d). The symmetric subspace of a t -partite Hilbert space consists of states that are left unchanged under permutations of the factor spaces. Reasoning for why Eq. (5.5) holds is as follows. First, by symmetry of $|\psi\rangle\langle\psi|^{\otimes t}$, $\int |\psi\rangle\langle\psi|^{\otimes t} d\psi$ must only have support on the symmetric subspace. In addition, since this operator is invariant under

multiplication by any unitary operator of the form $U^{\otimes t}$, and has unit trace, Schur's lemma implies that it must be equal to the normalized projector onto the symmetric subspace, which is just the right-hand side of Eq. (5.5). This implies that

$$\int r_k p_k d\psi = \frac{2\text{tr}([E_k \otimes \Pi_k] \Pi_{\text{sym}}(2, d))}{d(d+1)}. \quad (5.6)$$

Now, one can show that $\Pi_{\text{sym}}(t, d)$ is equal to the normalized sum of the $t!$ elements in the group of permutation operators on the t -partite Hilbert space:

$$\Pi_{\text{sym}}(t, d) = \frac{1}{t!} \sum_{\sigma} P_{\sigma}. \quad (5.7)$$

This can be shown by first noting that the set of all permutations is a subgroup of the unitary group. Thus the square of the normalized sum is just equal to the normalized sum itself

$$\left(\frac{1}{t!} \sum_{\sigma} P_{\sigma} \right)^2 = \frac{1}{t!} \sum_{\sigma} P_{\sigma}. \quad (5.8)$$

Hence, since $\frac{1}{t!} \sum_{\sigma} P_{\sigma}$ commutes with any permutation P_{τ} , it must be equal to the projection operator onto the symmetric subspace. As a simple example, in the case of $t = 2$, there are two permutation operators, $\mathbb{1}$, and the SWAP operation which swaps the two factor spaces. Thus

$$\Pi_{\text{sym}}(2, d) = \frac{\mathbb{1} \otimes \mathbb{1} + \text{SWAP}}{2}. \quad (5.9)$$

$$\begin{aligned} & \int \sqrt{\text{tr}(\Pi_l |\psi\rangle\langle\psi|) \text{tr}(E_l |\psi\rangle\langle\psi|)} \sqrt{\text{tr}(\Pi_m |\psi\rangle\langle\psi|) \text{tr}(E_m |\psi\rangle\langle\psi|)} d\psi \\ & \geq \int \sqrt{\text{tr}(\Pi_l |\psi\rangle\langle\psi|) \text{tr}(\Pi_l E_l \Pi_l (|\psi\rangle\langle\psi|))} \sqrt{\text{tr}(\Pi_m |\psi\rangle\langle\psi|) \text{tr}(\Pi_m E_m \Pi_m (|\psi\rangle\langle\psi|))} d\psi \\ & = \int \sqrt{\langle\psi| E_l |\psi\rangle} \text{tr}(\Pi_l (|\psi\rangle\langle\psi|)) \sqrt{\langle\psi| E_m |\psi\rangle} \text{tr}(\Pi_m (|\psi\rangle\langle\psi|)) d\psi = \sqrt{u_l u_m} \int \text{tr}(\Pi_l (|\psi\rangle\langle\psi|)) \text{tr}(\Pi_m (|\psi\rangle\langle\psi|)) d\psi. \end{aligned} \quad (5.13)$$

Note that for each j it is generally not true that $E_j - u_j \Pi_j$ is positive semidefinite (if this were the case, then the above inequality would always hold).

From Eq. (5.5), since Π_l and Π_m are projectors onto orthogonal subspaces,

$$\begin{aligned} & \int \text{tr}(\Pi_l (|\psi\rangle\langle\psi|)) \text{tr}(\Pi_m (|\psi\rangle\langle\psi|)) d\psi \\ & = \int \text{tr}[(\Pi_l \otimes \Pi_m) |\psi\rangle\langle\psi| \otimes |\psi\rangle\langle\psi|] d\psi \\ & = \frac{\text{tr}(\Pi_l \Pi_m) + \text{tr}(\Pi_l) \text{tr}(\Pi_m)}{d(d+1)} = \frac{1}{d(d+1)}. \end{aligned} \quad (5.14)$$

Thus

$$\int \sqrt{r_l r_m p_l p_m} d\psi \geq \frac{\sqrt{\text{tr}(E_l \Pi_l) \text{tr}(E_m \Pi_m)}}{d(d+1)} = \frac{\sqrt{u_l u_m}}{d(d+1)}, \quad (5.15)$$

which gives the last term in the numerator of Eq. (5.1).

Equations (5.6) and (5.9) imply that

$$\begin{aligned} \int r_k p_k d\psi &= \frac{\text{tr}([E_k \otimes \Pi_k] [\mathbb{1} \otimes \mathbb{1} + \text{SWAP}])}{d(d+1)} \\ &= \frac{\text{tr}(E_k \Pi_k) + \text{tr}(E_k) \text{tr}(\Pi_k)}{d(d+1)} \\ &= \frac{\text{tr}(E_k \Pi_k) + d \text{tr}(E_k \frac{\mathbb{1}}{d})}{d(d+1)}, \end{aligned} \quad (5.10)$$

and so

$$\sum_{k=1}^d \int r_k p_k d\psi = \sum_{k=1}^d \frac{\text{tr}(E_k \Pi_k) + d \text{tr}(E_k \frac{\mathbb{1}}{d})}{d(d+1)} = \frac{\sum_{k=1}^d u_k + d}{d(d+1)}, \quad (5.11)$$

since $\sum_{k=1}^d E_k = \mathbb{1}$. This gives the first two terms in the numerator in Eq. (5.1).

$$\text{B. } \sum_{l \neq m} \int \sqrt{r_l r_m p_l p_m} d\psi$$

Next,

$$\begin{aligned} \int \sqrt{r_l r_m p_l p_m} d\psi &= \int \sqrt{\text{tr}(\Pi_l |\psi\rangle\langle\psi|) \text{tr}(E_l |\psi\rangle\langle\psi|)} \\ &\quad \times \sqrt{\text{tr}(\Pi_m |\psi\rangle\langle\psi|) \text{tr}(E_m |\psi\rangle\langle\psi|)} d\psi. \end{aligned} \quad (5.12)$$

Computing the above integral analytically is difficult because of the square root in the argument. Now we assume that the following inequality holds (see Sec. VI):

In total, assuming that the inequality in Eq. (5.13) holds,

$$\frac{1 + d\bar{X}}{1 + d} = \frac{d + \sum_{k=1}^d u_k + \sum_{l \neq m} \sqrt{u_l u_m}}{d(d+1)} \quad (5.16)$$

is a lower bound for $\overline{F_{\mathcal{E}, \mathcal{M}}}$. Before analyzing the validity of the lower bound, we briefly compare its expression with that of the average gate fidelity and show that it satisfies certain necessary conditions to be useful in practice.

C. Comparison with quantum gate fidelity

There is a nice parallel between the lower bound on the average measurement fidelity and the exact expression for the average quantum gate fidelity. As previously mentioned, the quantum average gate fidelity takes a particularly nice form because one of the operations is unitary [36]. Indeed, if one compares the unitary $\mathcal{U}$ and quantum operation $\mathcal{E}$, the average

gate fidelity between $\mathcal{U}$ and $\mathcal{E}$, $\overline{\mathcal{F}_{\mathcal{E},\mathcal{U}}}$, is given by

$$\overline{\mathcal{F}_{\mathcal{E},\mathcal{U}}} = \frac{\overline{A}d + 1}{d + 1}, \quad (5.17)$$

where

$$\overline{A} = \frac{1}{d^2} \sum_k \text{tr}(A_k) \text{tr}(A_k^\dagger), \quad (5.18)$$

and $\{A_k\}$ is any set of Kraus operators for the quantum operation $\Lambda = \mathcal{U}^\dagger \circ \mathcal{E}$.

Our expression for a lower bound on the average measurement fidelity $\overline{F_{\mathcal{E},\mathcal{M}}}$ takes a similar form:

$$lb = \frac{\overline{X}d + 1}{d + 1}, \quad (5.19)$$

where

$$\overline{X} = \frac{1}{d^2} \sum_{l,m} \sqrt{\text{tr}(\Pi_l E_l) \text{tr}(\Pi_m E_m)}. \quad (5.20)$$

In some sense this is not surprising since each quantity is a Haar integral over functions of two copies of a quantum state $|\psi\rangle\langle\psi|$. A direction of further research is to understand properties of the average measurement fidelity in more detail and draw more parallels with well-known distinguishability measures such as the average gate fidelity.

D. Necessary conditions for lb to be a useful lower bound on $\overline{F_{\mathcal{E},\mathcal{M}}}$

In this subsection we discuss two necessary conditions lb must satisfy in order to be a useful lower bound on $\overline{F_{\mathcal{E},\mathcal{M}}}$.

(1) *Limit of no error*: As the measurement error goes to 0, $lb \uparrow 0$.

(2) *Scaling in d* : lb scales well in dimension d of the system.

Here, we show that lb satisfies both of these criteria.

1. Limit of no error

Suppose the measurement error goes to 0 in that the POVM elements converge to the ideal PVM elements:

$$E_j \rightarrow \Pi_j. \quad (5.21)$$

This implies, for each $j \in \{1, \dots, d\}$,

$$u_j = \text{tr}(\Pi_j E_j) \uparrow 1. \quad (5.22)$$

Hence, from Eqs. (4.7) and (5.1),

$$lb \uparrow 1. \quad (5.23)$$

Thus lb satisfies the necessary condition of converging to 1 in the limit of no errors.

2. Scaling in d

Let us now explicitly show that lb scales well in d . More precisely, if each $u_k = \text{tr}(\Pi_k E_k)$ is of the order of $1 - \delta$ for some $\delta > 0$, then, as $d \rightarrow \infty$, $lb \rightarrow 1 - \delta$. Hence, the lower bound does not generically converge to a constant value that is

independent of the E_k . Such an effect would render the lower bound useless as the system size grows large.

Again, by Eqs. (4.7) and (5.1) we see that if each u_k satisfies

$$u_k \sim 1 - \delta \quad (5.24)$$

for some $\delta > 0$, then

$$lb \sim \frac{1 + \overline{X}d}{1 + d} = \frac{1 + (1 - \delta)d}{1 + d}. \quad (5.25)$$

Hence as $d \rightarrow \infty$,

$$lb \rightarrow 1 - \delta, \quad (5.26)$$

which is what we wanted to show.

VI. VALIDITY OF THE LOWER BOUND

The lower bound in Eq. (5.1) is valid provided the inequality in Eq. (5.13) holds. The goal of this section is to show that this inequality holds in very general situations. To set notation, we define

$$\begin{aligned} f_{l,m}(|\psi\rangle) &= \sqrt{\text{tr}(\Pi_l |\psi\rangle\langle\psi|) \text{tr}(E_l |\psi\rangle\langle\psi|)} \\ &\quad \times \sqrt{\text{tr}(\Pi_m |\psi\rangle\langle\psi|) \text{tr}(E_m |\psi\rangle\langle\psi|)}, \\ g_{l,m}(|\psi\rangle) &= \sqrt{\text{tr}(\Pi_l |\psi\rangle\langle\psi|) \text{tr}(\Pi_l E_l \Pi_l |\psi\rangle\langle\psi|)} \\ &\quad \times \sqrt{\text{tr}(\Pi_m |\psi\rangle\langle\psi|) \text{tr}(\Pi_m E_m \Pi_m |\psi\rangle\langle\psi|)} \\ &= \sqrt{u_l u_m} \text{tr}(\Pi_l |\psi\rangle\langle\psi|) \text{tr}(\Pi_m |\psi\rangle\langle\psi|) \end{aligned} \quad (6.1)$$

so that we want to show

$$\int [f_{l,m}(|\psi\rangle) - g_{l,m}(|\psi\rangle)] d\psi \geq 0. \quad (6.2)$$

We demonstrate that Eq. (6.2) should hold by showing that the set of states for which $f_{l,m}(|\psi\rangle) - g_{l,m}(|\psi\rangle) < 0$ is small, and even if $|\psi\rangle$ satisfies $f_{l,m}(|\psi\rangle) - g_{l,m}(|\psi\rangle) < 0$, $|f_{l,m}(|\psi\rangle) - g_{l,m}(|\psi\rangle)|$ will be small. So, let us look at the measure of states $|\psi\rangle$ that could satisfy

$$f_{l,m}(|\psi\rangle) - g_{l,m}(|\psi\rangle) < 0 \quad (6.3)$$

and what $|f_{l,m}(|\psi\rangle) - g_{l,m}(|\psi\rangle)|$ will look like for such a state.

First, note that

$$\int [f_{l,m}(|\psi\rangle) - g_{l,m}(|\psi\rangle)] d\psi = \int k_{l,m}(|\psi\rangle) h_{l,m}(|\psi\rangle) d\psi, \quad (6.4)$$

where

$$k_{l,m}(|\psi\rangle) := \sqrt{\text{tr}(\Pi_l |\psi\rangle\langle\psi|) \text{tr}(\Pi_m |\psi\rangle\langle\psi|)}, \quad (6.5)$$

$$h_{l,m}(|\psi\rangle) := \sqrt{\text{tr}(E_l |\psi\rangle\langle\psi|) \text{tr}(E_m |\psi\rangle\langle\psi|)} \quad (6.6)$$

$$- \sqrt{u_l u_m} \sqrt{\text{tr}(\Pi_l |\psi\rangle\langle\psi|) \text{tr}(\Pi_m |\psi\rangle\langle\psi|)}. \quad (6.7)$$

Since $k_{l,m}(|\psi\rangle) \geq 0$ for all $|\psi\rangle$,

$$f_{l,m}(|\psi\rangle) - g_{l,m}(|\psi\rangle) < 0 \Leftrightarrow h_{l,m}(|\psi\rangle) < 0. \quad (6.8)$$

However, if $h_{l,m}(|\psi\rangle) < 0$, then either

$$\langle\psi|E_l - u_l\Pi_l|\psi\rangle < 0 \quad \text{or} \quad \langle\psi|E_m - u_m\Pi_m|\psi\rangle < 0. \quad (6.9)$$

Without loss of generality, suppose $\langle\psi|E_l - u_l\Pi_l|\psi\rangle < 0$. Thus, if $|\psi\rangle$ satisfies $f_{l,m}(|\psi\rangle) - g_{l,m}(|\psi\rangle) < 0$, then $\langle\psi|E_l - u_l\Pi_l|\psi\rangle < 0$.

As a matrix written in the $\{|\psi_k\rangle\}$ basis, $E_l - u_l\Pi_l$ is equal to E_l except in the (l,l) th entry, which is 0. For instance, if $l = 0$, then $E_0 - u_0\Pi_0$ is given by

$$\begin{bmatrix} 0 & E_0^{0,1} & \dots & \dots & E_0^{0,d-2} & E_0^{0,d-1} \\ E_0^{1,0} & E_0^{1,1} & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ E_0^{d-2,0} & & \ddots & \ddots & \ddots & \vdots \\ E_0^{d-1,0} & \dots & \dots & \dots & \dots & E_0^{d-1,d-1} \end{bmatrix}. \quad (6.10)$$

Hence, $E_l - u_l\Pi_l$ is not positive semidefinite unless the off-diagonals in the l th row (column) are equal to 0. This is because, since E_l is positive semidefinite, the following bound on the off-diagonal elements holds [51]:

$$|E_l^{i,j}| \leq \sqrt{E_l^{i,i}} \sqrt{E_l^{j,j}}. \quad (6.11)$$

Now, note that if $|\psi\rangle = |\psi_l\rangle$, then $\langle\psi|E_l - u_l\Pi_l|\psi\rangle = 0$, and if $|\psi\rangle$ is orthogonal to $|\psi_l\rangle$, $\langle\psi|E_l - u_l\Pi_l|\psi\rangle \geq 0$. Hence, there is some small region in the space of pure states centered around the state $|\psi_l\rangle$ for which one could have

$$\langle\psi|E_l - u_l\Pi_l|\psi\rangle < 0. \quad (6.12)$$

That is, the function on pure states defined by $\langle\psi|E_l - u_l\Pi_l|\psi\rangle$ has a minimum (which can be less than 0) that is achieved for some state close to $|\psi_l\rangle$. As $|\psi\rangle$ moves farther away from $|\psi_l\rangle$, and accumulates more amplitude in the subspace orthogonal to $|\psi_l\rangle$, we see that

$$\langle\psi|E_l - u_l\Pi_l|\psi\rangle \geq 0. \quad (6.13)$$

Thus, Eq. (6.9) can only be satisfied if $|\psi\rangle$ has a high enough amplitude in the subspace defined by $|\psi_l\rangle$. The amount of amplitude that is required depends on the size of the off-diagonal elements (coherence) in the l th column (row) of E_l . If the coherence is not too large, then $|\psi\rangle$ will need to have a high amplitude in $|\psi_l\rangle$ to satisfy Eq. (6.9), and as the coherence grows larger, $|\psi\rangle$ can potentially satisfy Eq. (6.9) without being very close to $|\psi_l\rangle$.

The key point is that the measure of the set of states for which $|\psi\rangle$ is close to $|\psi_l\rangle$ (or $|\psi_m\rangle$) will be small (and, by Levy's lemma [52], decreases exponentially in the dimension d). Moreover, if $|\psi\rangle$ has a high amplitude in $|\psi_l\rangle$, and so $\langle\psi|E_l - u_l\Pi_l|\psi\rangle < 0$ may occur, then $\text{tr}(\Pi_m|\psi\rangle\langle\psi|)$ is small

since $|\psi_l\rangle$ and $|\psi_m\rangle$ are orthogonal. Thus, by definition, if $|\psi\rangle$ is close to either $|\psi_l\rangle$ or $|\psi_m\rangle$, $k_{l,m}(|\psi\rangle)$ will be small. Hence, $k_{l,m}(|\psi\rangle)$ acts like a *modulating factor* in the expression for $f_{l,m}(|\psi\rangle) - g_{l,m}(|\psi\rangle)$ to ensure that $f_{l,m}(|\psi\rangle) - g_{l,m}(|\psi\rangle)$ is close to 0 when $|\psi\rangle$ is close to either $|\psi_l\rangle$ or $|\psi_m\rangle$.

In total, then, the measure of states for which $|\psi\rangle$ is close to $|\psi_l\rangle$ or $|\psi_m\rangle$ is small, and for any such state, $f_{l,m}(|\psi\rangle) - g_{l,m}(|\psi\rangle)$ will be close to 0. Equivalently, the set of states for which $f_{l,m}(|\psi\rangle) - g_{l,m}(|\psi\rangle) < 0$ is small, and for any such state, $|f_{l,m}(|\psi\rangle) - g_{l,m}(|\psi\rangle)|$ will be close to 0. This is what we wanted to show and so we expect Eq. (6.2) to hold for most practical cases.

Note that the bounds are guaranteed to be valid if there is no coherence at all, that is, if the POVM elements are diagonal in the $\{|\psi_k\rangle\}$ basis. In the next section we obtain a sufficient condition for Eq. (6.2) to hold in the case of a single qubit. In Sec. VIB we perform a detailed numerical investigation of the single-qubit case and show that our lower bound always holds; that is, for *all* values of coherence magnitude, the bounds are valid. This provides evidence that the bounds will be valid in large generality. Clearly, a more in-depth investigation of sufficient conditions for the bounds to be valid is desirable, however, the above argument and numerical results in Sec. VIB indicate that the bounds and protocol will be valid in most practical cases.

A. Sufficient condition for the single-qubit case

Using the notation in the previous section, for a single qubit, the lower bound in Eq. (5.1) is valid if

$$\int [f_{0,1}(|\psi\rangle) - g_{0,1}(|\psi\rangle)] d\psi \geq 0. \quad (6.14)$$

Our goal in this section is to obtain a sufficient condition for when Eq. (6.14) holds.

We use the Bloch sphere representation of a single qubit,

$$|\psi\rangle = \cos\left(\frac{\theta}{2}\right)|0\rangle + e^{i\phi}\sin\left(\frac{\theta}{2}\right)|1\rangle, \quad (6.15)$$

where $\theta \in [0, \pi)$, $\phi \in [0, 2\pi)$. Hence, the POVM elements are given by

$$E_0 = \begin{pmatrix} u_0 & \gamma \\ \gamma & \text{tr}(E_0\Pi_1) \end{pmatrix}, \quad (6.16)$$

$$E_1 = \begin{pmatrix} \text{tr}(E_1\Pi_0) & -\gamma \\ -\gamma & u_1 \end{pmatrix}, \quad (6.17)$$

where, since the results will depend on the magnitude of the coherence in the POVM elements, we assume without loss of generality that $\gamma \in \mathbb{R}$ and $\gamma > 0$.

Since our goal is obtaining a lower bound on $\int [f_{0,1}(|\psi\rangle) - g_{0,1}(|\psi\rangle)] d\psi$, we need to understand when one of $\langle\psi|E_0 - u_0\Pi_0|\psi\rangle < 0$ or $\langle\psi|E_1 - u_1\Pi_1|\psi\rangle < 0$ holds since this is a necessary condition for $f_{0,1}(|\psi\rangle) - g_{0,1}(|\psi\rangle) < 0$ [see the discussion surrounding Eq. (6.9)]. Without loss of generality,

let us assume that

$$\langle \psi | E_0 - u_0 \Pi_0 | \psi \rangle < 0. \quad (6.18)$$

Then

$$\begin{aligned} \langle \psi | E_0 - u_0 \Pi_0 | \psi \rangle &= 2\text{Re} \left(\gamma \cos \left(\frac{\theta}{2} \right) e^{i\phi} \sin \left(\frac{\theta}{2} \right) \right) \\ &\quad + \text{tr}(E_0 \Pi_1) \sin^2 \left(\frac{\theta}{2} \right), \end{aligned} \quad (6.19)$$

and it is straightforward to show that $\langle \psi | E_0 - u_0 \Pi_0 | \psi \rangle < 0$ is satisfied if

$$\phi \in \left(\frac{\pi}{2}, \frac{3\pi}{2} \right), \quad \theta \in \left[0, 2 \arccot \left(-\frac{\text{tr}(E_0 \Pi_1)}{2\gamma \cos(\phi)} \right) \right]. \quad (6.20)$$

Hence, denoting the set of states that satisfy $\langle \psi | E_0 - u_0 \Pi_0 | \psi \rangle < 0$ by $A_{0,\gamma}$, we can see that $A_{0,\gamma}$ is contained in the set of all states with ϕ and θ given by Eq. (6.20). As expected, the measure of $A_{0,\gamma}$, $\mu(A_{0,\gamma})$, is extremely small for weak coherence and grows larger as the coherence increases in magnitude.

The uniform measure on the Bloch sphere has density function

$$\frac{1}{4\pi} \sin(\theta). \quad (6.21)$$

Hence, from Eqs. (6.20), $\mu(A_{0,\gamma})$ is at most

$$\begin{aligned} &\frac{1}{4\pi} \int_{\frac{\pi}{2}}^{\frac{3\pi}{2}} \int_0^{2\arccot\left(-\frac{\text{tr}(E_0 \Pi_1)}{2\gamma \cos(\phi)}\right)} \sin(\theta) d\theta d\phi \\ &= \frac{1}{4} - \frac{1}{4\pi} \int_{\frac{\pi}{2}}^{\frac{3\pi}{2}} \cos \left(2 \arccot \left(-\frac{\text{tr}(E_0 \Pi_1)}{2\gamma \cos(\phi)} \right) \right) d\phi. \end{aligned} \quad (6.22)$$

Now, if $|\psi\rangle \in A_{0,\gamma}$,

$$\theta < 2 \arccot \left(-\frac{\text{tr}(E_0 \Pi_1)}{2\gamma \cos(\phi)} \right), \quad (6.23)$$

and so

$$\begin{aligned} \text{tr}(\Pi_1 |\psi\rangle \langle \psi|) &\leq \sin^2 \left(\arccot \left(-\frac{\text{tr}(E_0 \Pi_1)}{2\gamma \cos(\phi)} \right) \right) \\ &= \frac{\cos^2(\phi)}{\cos^2(\phi) + \left(\frac{\text{tr}(E_0 \Pi_1)}{2\gamma} \right)^2}. \end{aligned} \quad (6.24)$$

Thus, for weak coherence γ , the set of states which satisfy Eq. (6.18) will have small measure bounded by Eq. (6.22). Using a symmetric argument for state $|1\rangle$ gives the following set of equations.

$$\begin{aligned} \mu(A_{0,\gamma}) &\leq \frac{1}{4} - \frac{1}{4\pi} \int_{\frac{\pi}{2}}^{\frac{3\pi}{2}} \cos \left(2 \arccot \left(-\frac{\text{tr}(E_0 \Pi_1)}{2\gamma \cos(\phi)} \right) \right) d\phi \\ \mu(A_{1,\gamma}) &\leq \frac{1}{4} - \frac{1}{4\pi} \int_{\frac{\pi}{2}}^{\frac{3\pi}{2}} \cos \left(2 \arccot \left(-\frac{\text{tr}(E_1 \Pi_0)}{2\gamma \cos(\phi)} \right) \right) d\phi. \end{aligned} \quad (6.25)$$

$$\text{If } |\psi\rangle \in A_{0,\gamma}, \quad \text{tr}(\Pi_1 |\psi\rangle \langle \psi|) \leq \frac{\cos^2(\phi)}{\cos^2(\phi) + \left(\frac{\text{tr}(E_0 \Pi_1)}{2\gamma} \right)^2} \quad \text{with } \phi \in \left(\frac{\pi}{2}, \frac{3\pi}{2} \right).$$

$$\text{If } |\psi\rangle \in A_{1,\gamma}, \quad \text{tr}(\Pi_0 |\psi\rangle \langle \psi|) \leq \frac{\cos^2(\phi)}{\cos^2(\phi) + \left(\frac{\text{tr}(E_1 \Pi_0)}{2\gamma} \right)^2} \quad \text{with } \phi \in \left(\frac{\pi}{2}, \frac{3\pi}{2} \right).$$

We can now obtain a sufficient condition for

$$\int [f_{0,1}(|\psi\rangle) - g_{0,1}(|\psi\rangle)] d\psi = \sum_{j=0}^1 \int_{A_{j,\gamma}} [f_{0,1}(|\psi\rangle) - g_{0,1}(|\psi\rangle)] d\psi + \int_{(A_{0,\gamma} \cup A_{1,\gamma})^c} [f_{0,1}(|\psi\rangle) - g_{0,1}(|\psi\rangle)] d\psi \geq 0. \quad (6.26)$$

By the set of equations in Eq. (6.25) one can show that

$$\begin{aligned} \int_{A_{0,\gamma}} [f_{0,1}(|\psi\rangle) - g_{0,1}(|\psi\rangle)] d\psi &\geq \int_{A_{0,\gamma}} 0 - \sqrt{u_0 u_1} \text{tr}(\Pi_0 |\psi\rangle \langle \psi|) \text{tr}(\Pi_1 |\psi\rangle \langle \psi|) d\psi \\ &= -\frac{\sqrt{u_0 u_1}}{2\pi} \int_{\frac{\pi}{2}}^{\frac{3\pi}{2}} \sin^4 \left(\arccot \left(-\frac{\text{tr}(E_1 \Pi_0)}{2\gamma \cos(\phi)} \right) \right) d\phi. \end{aligned} \quad (6.27)$$

By Eq. (6.24),

$$\int_{A_{0,\gamma}} [f_{0,1}(|\psi\rangle) - g_{0,1}(|\psi\rangle)] d\psi \geq -\frac{\sqrt{u_0 u_1}}{2\pi} \int_{\frac{\pi}{2}}^{\frac{3\pi}{2}} \left[\frac{\cos^2(\phi)}{\cos^2(\phi) + \left(\frac{\text{tr}(E_0 \Pi_1)}{2\gamma} \right)^2} \right]^2 d\phi =: -\delta_{0,\gamma}, \quad (6.28)$$

and similarly,

$$\int_{A_{1,\gamma}} [f_{0,1}(|\psi\rangle) - g_{0,1}(|\psi\rangle)] d\psi \geq -\frac{\sqrt{u_0 u_1}}{2\pi} \int_{\frac{\pi}{2}}^{\frac{3\pi}{2}} \left[\frac{\cos^2(\phi)}{\cos^2(\phi) + \left(\frac{\text{tr}(E_1 \Pi_0)}{2\gamma}\right)^2} \right]^2 d\phi =: -\delta_{1,\gamma}. \quad (6.29)$$

Hence

$$\int [f_{0,1}(|\psi\rangle) - g_{0,1}(|\psi\rangle)] d\psi \geq -(\delta_{0,\gamma} + \delta_{1,\gamma}) + \int_{(A_{0,\gamma} \cup A_{1,\gamma})^c} [f_{0,1}(|\psi\rangle) - g_{0,1}(|\psi\rangle)] d\psi, \quad (6.30)$$

where

$$\mu((A_{0,\gamma} \cup A_{1,\gamma})^c) \geq \frac{1}{2} + \frac{1}{4\pi} \int_{\frac{\pi}{2}}^{\frac{3\pi}{2}} \cos \left(2 \arccot \left(-\frac{\text{tr}(E_0 \Pi_1)}{2\gamma \cos(\phi)} \right) \right) + \frac{1}{4\pi} \int_{\frac{\pi}{2}}^{\frac{3\pi}{2}} \cos \left(2 \arccot \left(-\frac{\text{tr}(E_1 \Pi_0)}{2\gamma \cos(\phi)} \right) \right) d\phi \quad (6.31)$$

and if $|\psi\rangle \in (A_{0,\gamma} \cup A_{1,\gamma})^c$,

$$f_{0,1}(|\psi\rangle) - g_{0,1}(|\psi\rangle) \geq 0. \quad (6.32)$$

This gives the desired sufficient condition for $\int [f_{0,1}(|\psi\rangle) - g_{0,1}(|\psi\rangle)] d\psi \geq 0$ and, so, a sufficient condition for the lower bound to be valid:

$$\int_{(A_{0,\gamma} \cup A_{1,\gamma})^c} [f_{0,1}(|\psi\rangle) - g_{0,1}(|\psi\rangle)] d\psi \geq \delta_{0,\gamma} + \delta_{1,\gamma}. \quad (6.33)$$

Note that the right-hand side of Eq. (6.33) is typically small and the left-hand side is non-negative with the integral taken over a set with large measure.

We have shown that, for a single qubit, an analytical sufficient condition for the lower bound to be valid does indeed exist. This sufficient condition explicitly defines a regime where our bounds are provably valid. However, clearly it will also be useful to numerically analyze the exact relationship between the average measurement fidelity and our derived lower bound lb for the single-qubit case. This analysis will further aid in shaping our understanding of the validity of the bounds and is performed in the next section. We find that the lower bound is *always* valid, independent of the magnitude of the coherence.

B. Numerical examples for the single-qubit case

In the previous section we obtained a sufficient condition for lb to be valid in the case of a single qubit. We now numerically analyze various single-qubit examples to observe the *exact* behavior of lb relative to the exact value of $\overline{F_{\mathcal{E},\mathcal{M}}}$, which will also help clarify many of the technical details presented thus far.

In order to perform integration over the Fubini-Study measure on single-qubit states, we again use the Bloch sphere representation of a single qubit state,

$$|\psi\rangle = \cos\left(\frac{\theta}{2}\right) |0\rangle + \sin\left(\frac{\theta}{2}\right) e^{i\phi} |1\rangle, \quad (6.34)$$

where $\theta \in [0, \pi]$, $\phi \in [0, 2\pi]$. Thus any function f on pure states can be written as a function $f(\theta, \phi)$ on the unit sphere $\mathbb{S}^2 \subset \mathbb{R}^3$. In addition, the integral of $f(\theta, \phi)$ with respect to

the Fubini-Study measure is just the usual double-integral,

$$\frac{1}{4\pi} \int_{\theta=0}^{\pi} \int_{\phi=0}^{2\pi} f(\theta, \phi) \sin(\theta) d\theta d\phi, \quad (6.35)$$

where the normalization $\frac{1}{4\pi}$ ensures we are integrating over a probability measure.

Our goal is to provide a direct comparison between $\overline{F_{\mathcal{E},\mathcal{M}}}$ and lb for the case of a single-qubit projective measurement in the computational basis. As before, we write the noisy POVM elements as $\{E_0, E_1\}$. From Eq. (5.3), the exact value of the average measurement fidelity, $\overline{F_{\mathcal{E},\mathcal{M}}}$, is given by

$$\begin{aligned} \overline{F_{\mathcal{E},\mathcal{M}}} &= \int F(\mathcal{E}(|\psi\rangle\langle\psi|), \mathcal{M}(|\psi\rangle\langle\psi|)) d\psi \\ &= \sum_{k=0}^1 \left[\int r_k p_k d\psi \right] + \sum_{l \neq m} \left[\int \sqrt{r_l r_m p_l p_m} d\psi \right] \\ &= \sum_{k=0}^1 \left[\int r_k p_k d\psi \right] + 2 \int \sqrt{r_0 r_1 p_0 p_1} d\psi, \end{aligned} \quad (6.36)$$

where, using the Bloch sphere representation and the elements of the POVM $\{E_0, E_1\}$, we can write

$$\begin{aligned} p_0 &= \cos^2\left(\frac{\theta}{2}\right), \quad p_1 = 1 - p_0 = \sin^2\left(\frac{\theta}{2}\right), \\ r_0 &= E_0^{0,0} \cos^2\left(\frac{\theta}{2}\right) + E_0^{1,1} \sin^2\left(\frac{\theta}{2}\right) + \text{Re}(E_0^{0,1} e^{i\phi}) \sin(\theta), \\ r_1 &= 1 - r_0. \end{aligned} \quad (6.37)$$

For each $k = 0, 1$ we can compute $\int r_k p_k d\psi$ exactly by the techniques introduced in Sec. V A [specifically see Eq. (5.11)]. However, as noted in Sec. V B, calculating $\int \sqrt{r_0 r_1 p_0 p_1} d\psi$ analytically is generally not possible. One can, however, directly input the expressions in Eqs. (6.37) into $\int \sqrt{r_0 r_1 p_0 p_1} d\psi$ so that a numerical analysis can be performed.

The POVM elements $\{E_0, E_1\}$ that model the noise must each be positive semidefinite and $E_0 + E_1 = \mathbb{1}$. Hence, we

TABLE I. Values of lb and ub for three cases of u_0 .

	u_0		
	0.99	0.995	0.999
lb	0.9933	0.9967	0.9993
ub	0.0067	0.0033	0.0007

write

$$E_0 = \begin{pmatrix} u_0 & \gamma \\ \gamma^* & 1 - u_0 \end{pmatrix}, \quad (6.38)$$

$$E_1 = \begin{pmatrix} 1 - u_0 & -\gamma \\ -\gamma^* & u_0 \end{pmatrix}, \quad (6.39)$$

where $\gamma \in \mathbb{C}$ is given by

$$\gamma = r e^{i\phi} \quad (6.40)$$

and $\phi \in [0, 2\pi)$. The linear operator E_0 (E_1) is positive semidefinite if and only if its leading principal minors are non-negative. Since $d = 2$, this reduces to the condition

$$|\gamma| = r \leq \sqrt{u_0(1 - u_0)}. \quad (6.41)$$

For various cases, we observe how the exact measurement fidelity compares to lb as R approaches its maximum value,

$$R_{\max} := \sqrt{u_0(1 - u_0)}. \quad (6.42)$$

Since the results will depend only on R and not ϕ , without loss of generality, one can assume $\gamma \in \mathbb{R}$ so that $\gamma = R$. As a verification of this, the numerics given below were performed for various $\phi \in [0, 2\pi)$, and as expected, the results were independent of ϕ .

Three of the cases we analyzed are $u_0 = 0.99, 0.995$, and 0.999 . The values of lb and ub for each case are listed in Table I. Plots of the average measurement error $\overline{F_{\mathcal{E}, \mathcal{M}}}$ as a function of $|\gamma|$ for each case are given in Fig. 1. The horizontal axis is cut off at $R_{\max}$ for each case. Thus, all admissible values of the coherence are provided.

Various important features are evident from Fig. 1. First, lb is *always* a lower bound for $\overline{F_{\mathcal{E}, \mathcal{M}}}$ (the curves in the plot do not intersect). This agrees with the discussion in Sec. V, where we noted that lb should be a valid lower bound in large generality (even for large coherence) due to both the anticorrelation of the random variables $\text{tr}(\Pi_0|\psi\rangle\langle\psi|)$ and $\text{tr}(\Pi_1|\psi\rangle\langle\psi|)$ and the measure of the set of states for which $g_{0,1} > f_{0,1}$ is small. Second, as expected, the difference between lb and $\overline{F_{\mathcal{E}, \mathcal{M}}}$ grows smaller as the magnitude of the coherence increases. Thus, the bounds become tighter as the coherence in the POVM elements increases.

These results suggest that the bounds are always valid regardless of how large $|\gamma|$ is relative to u_0 . We tested whether there were any violations when u_0 was varied from 0.5 to 0.9999 in increments of 10^{-4} , and the coherence was varied from its minimum to its maximum in each case. We found *no* violations of the lower bound.

FIG. 1. (Color online) Plot of exact average measurement fidelity $\overline{F_{\mathcal{E}, \mathcal{M}}}$ [dashed (blue) curve] versus coherence strength $|\gamma|$ for three values of u_0 : 0.99 (top left), 0.995 (bottom left), and 0.999 (bottom right). The lower bound lb for $\overline{F_{\mathcal{E}, \mathcal{M}}}$ in each case is also given [horizontal (green) line]. The magnitude of the coherence was varied in each case up to its maximum value, $R_{\max}$.

VII. CHARACTERIZING MEASUREMENT PROBABILITIES AND OUTPUT STATES FOR RANK 1 PVMs

Ideally, a quantum measurement produces an output state that can be utilized for further purposes. In this case one would like to characterize the error in both the measurement statistics and the output states of the measurement. In this section we briefly outline such a method for rank 1 PVMs and discuss conditions for the lower bound on the measurement fidelity $\overline{F_{\mathcal{E}, \mathcal{M}}}$ to be valid.

From Eq. (2.4) $\mathcal{M}$ is given by

$$\mathcal{M}(\sigma) = \sum_{k=1}^d \Pi_k \sigma \Pi_k = \sum_{k=1}^b p_k \Pi_k, \quad (7.1)$$

where

$$p_k = \text{tr}(\Pi_k \sigma), \quad (7.2)$$

and the output state Π_k is *independent* of the input state σ . Note that this assumption is only valid because the ideal measurement is a rank 1 PVM. We assume the noisy measurement process $\mathcal{E}$ is of the form

$$\mathcal{E}(\sigma) = \sum_{k=1}^b r_k \rho_k, \quad (7.3)$$

where, as before,

$$r_k = \text{tr}(E_k \sigma) \quad (7.4)$$

is modeled via a POVM $\{E_k\}_{k=1}^d$. We assume the ρ_k are independent of σ since, ideally, the output states Π_k are independent of σ . However, it may be the case that there is some dependence of ρ_k on the input σ . When the dependence

is weak we expect the results presented here to generally still be valid, however, the extent to which this is the case is left as an area for further research. When there is a large dependence, clearly new techniques will have to be applied. There are simple tests one could perform to help detect such state dependence of output states. For instance, suppose we take two input states, $|\psi_1\rangle$ and $|\psi_2\rangle$. We input $|\psi_1\rangle$ into the measurement and observe whether outcome “ k ” is obtained. Whenever it is, we can make a successive measurement on the output state $\rho_k(|\psi_1\rangle)$ and observe whether outcome “ k ” is obtained again. Repeating this many times produces the statistic $\text{tr}(\rho_k(|\psi_1\rangle)E_k)$. Repeating this procedure for $|\psi_2\rangle$ gives $\text{tr}(\rho_k(|\psi_2\rangle)E_k)$. If these quantities differ, then there must be some input state dependence in the output states of the measurement.

The average measurement fidelity is still given by

$$\begin{aligned}\overline{F_{\mathcal{E},\mathcal{M}}} &= \int F_{\mathcal{E},\mathcal{M}}(|\psi\rangle\langle\psi|)d\psi \\ &= \int F\left(\sum_k r_k \rho_k, \sum_k p_k \Pi_k\right)d\psi,\end{aligned}\quad (7.5)$$

and as before,

$$r_{\mathcal{E},\mathcal{M}} = 1 - \int F_{\mathcal{E},\mathcal{M}}(|\psi\rangle\langle\psi|)d\psi. \quad (7.6)$$

In this case, where we are also interested in output states, the expression for $\overline{F_{\mathcal{E},\mathcal{M}}}$ is much more complicated. However, we can use strong concavity of the fidelity [53] to obtain

$$\begin{aligned}\overline{F_{\mathcal{E},\mathcal{M}}} &= \int F(\mathcal{E}(|\psi\rangle), \mathcal{M}(|\psi\rangle)) \\ &\geq \int \left(\sum_k \sqrt{r_k p_k} \sqrt{\text{tr}(\rho_k \Pi_k)}\right)^2 d\psi.\end{aligned}\quad (7.7)$$

Since $\text{tr}(\rho_k \Pi_k)$ is just the fidelity between the k th ideal and the noisy output states of the measurement, we define

$$F_k := F(\rho_k, \Pi_k) = \text{tr}(\rho_k \Pi_k). \quad (7.8)$$

Moreover, since there is only state dependence in p_k and r_k ,

$$\overline{F_{\mathcal{E},\mathcal{M}}} \geq \sum_{(l,m) \in \mathcal{D}} \left[\int \sqrt{r_l r_m p_l p_m} d\psi \right] \sqrt{F_l F_m}. \quad (7.9)$$

From Sec. V, and specifically Eq. (5.15), this gives

$$\overline{F_{\mathcal{E},\mathcal{M}}} \geq \frac{1 + d\overline{Y}}{1 + d}, \quad (7.10)$$

where

$$\overline{Y} := \frac{1}{d^2} \sum_{(l,m) \in \mathcal{D}} \sqrt{u_l u_m} \sqrt{F_l F_m}. \quad (7.11)$$

At this point, we could define $lb = \frac{1+d\overline{Y}}{1+d}$ (and thus $ub = 1 - \frac{1+d\overline{Y}}{1+d}$), except the F_k cannot be obtained directly from the noisy measurement $\mathcal{E}$. The reason for this is that F_k describes the overlap between the output state from the ideal measurement, Π_k , and the output state from the real measurement, ρ_k . Equivalently, this is the same as the result obtained by making an ideal measurement in the Π_k basis on the real measurement output state ρ_k . By definition however,

we cannot make an ideal measurement $\mathcal{M}$. All we can do is make another real measurement $\mathcal{E}$ and look at resulting outcome probabilities, namely, $\text{tr}(\rho_k E_k)$.

Let us denote Q_k to be this probability,

$$\begin{aligned}Q_k &= \mathbb{P}[\text{obtaining outcome } k \text{ from } \mathcal{E} \text{ if the input is } \rho_k] \\ &= \text{tr}(\rho_k E_k)\end{aligned}\quad (7.12)$$

where the notation “ $\mathbb{P}(\cdot)$ ” means “probability of.” Thus, if it is true that for each k

$$F_k = \text{tr}(\rho_k \Pi_k) \geq Q_k = \text{tr}(\rho_k E_k), \quad (7.13)$$

then

$$\overline{F_{\mathcal{E},\mathcal{M}}} \geq lb := \frac{1 + d\overline{Z}}{1 + d}, \quad (7.14)$$

where

$$\overline{Z} := \frac{1}{d^2} \sum_{(l,m) \in \mathcal{D}} \sqrt{u_l u_m} \sqrt{Q_l Q_m}. \quad (7.15)$$

Motivation for why the assumption in Eq. (7.13) will often be true is that ideally, for each k , $E_k = \Pi_k$. Hence, $\text{tr}(\rho_k \Pi_k)$ is just the k th diagonal element of ρ_k , $\rho_k^{k,k}$. Since $E_k \leq 1$, if E_k is diagonal in the $\{|\psi_k\rangle\}$ basis, then Eq. (7.13) is guaranteed to hold. This is because the probability amplitude of Π_k is spread across the diagonal elements of E_k . Even when there is coherence in the E_k (nonzero off-diagonal elements), we expect that the inequality should still hold in realistic situations, however, a more detailed analysis of the conditions for this to occur is required. Note also that, realistically, one does not need Eq. (7.13) to hold for every k , just a large enough number to ensure $\overline{Y} \geq \overline{Z}$. We now provide the experimental protocol for determining lb .

VIII. EXPERIMENTAL PROTOCOL: RANK 1 PVMs WITH OUTPUT STATES

Goal: Obtain lb in the case of imperfect measurement probabilities and output states.

Protocol

Step 1: Choose a pair of indices (l, m) uniformly at random from $\mathcal{D} = \{0, \dots, d-1\} \times \{0, \dots, d-1\}$.

Step 2: For each $j = l$ and $j = m$:

(a) Prepare the quantum state Π_j , perform the noisy measurement $\mathcal{E}$ on Π_j , and record whether outcome j is obtained,

(b) If outcome j is obtained in step a, repeat the measurement on the current state of the system, and record whether outcome j is obtained again,

(c) Repeat steps a and b many times and denote the frequency of obtaining j in each step $\hat{u}_j$ and $\hat{Q}_j$, respectively. (See below for a discussion of the number of repetitions required to estimate u_j and Q_j to the desired accuracies.)

Step 3: Repeat steps 1 and 2 K times, where K is dictated by the desired accuracy and confidence in estimating lb . (See below for a discussion of the size of K).

Step 4: Compute the estimator, $\hat{lb}$, of lb [where lb is given in Eq. (7.14)] to the accuracy and confidence dictated by the

K trials $\{(1,1_2), \dots, (K_1, K_2)\}$ from step 3:

$$\hat{lb} = \frac{1 + d \left[\frac{1}{K} \sum_{(k_1, k_2)} \sqrt{u_{k_1} u_{k_2}} \sqrt{Q_{k_1} Q_{k_2}} \right]}{1 + d}. \quad (8.1)$$

This concludes the protocol.

We should emphasize various points about this protocol. First, similarly to the case of no output states (see Sec. IV), the number of trials required to implement the above protocol is independent of d and depends only on the desired accuracy and confidence of the estimates in each of steps 2c and 3. See Secs. IX A and IX B for respective discussions about the time complexity of each of these steps.

Second, lb can be estimated using only

- (i) sequential applications of the noisy measurement and
- (ii) the ability to prepare each of the d pure input states Π_j .

Finally, it is straightforward to show that the two necessary conditions given previously in Sec. VD for lb to be a useful lower bound also hold here. Indeed,

- (i) in the limit of $F_{\mathcal{E}, \mathcal{M}} \uparrow 1$,

$$lb \uparrow 1, \quad (8.2)$$

and

- (ii) lb scales well in d .

The first condition holds since, as $F_{\mathcal{E}, \mathcal{M}} \uparrow 1$, it must be the case that $E_k \rightarrow \Pi_k$ and $\rho_k \rightarrow \Pi_k$. Therefore, for every k , $u_k \rightarrow 1$ and $Q_k \rightarrow 1$. The second condition holds using an argument analogous to that given in Sec. VD2.

IX. RESOURCE ANALYSIS

In this section we discuss the time complexity and resources required for the protocols and compare the time complexity of the protocol with a full reconstruction of the noisy measurement (i.e., a full reconstruction of the POVM elements E_k). First, we analyze the number of trials required in step 2b of the protocol in Sec. IV and step 2c of the protocol in Sec. VIII. Afterwards, we analyze the number of trials required in step 3 of both protocols.

A. Number of trials required in step 2b in Sec. IV and step 2c in Sec. VIII

We first explicitly analyze the number of trials for step 2b of Sec. IV. The discussion carries over in a straightforward manner to step 2c in Sec. VIII.

1. Step 2b in Sec. IV

For each $k \in \{0, 1, \dots, d-1\}$, we would like to understand how many samples are required to obtain an estimate of u_k . Let us fix k and define $z_k = \text{tr}(\Pi_k(\mathbb{1} - E_k))$, where, since $\{E_k\}_{k=1}^d$ is a POVM,

$$\mathbb{1} - E_k = \sum_{j \neq k} E_j. \quad (9.1)$$

Then

$$z_k \geq 0, \quad u_k + z_k = \text{tr}(\Pi_k) = 1, \quad (9.2)$$

so $\{u_k, z_k\} = \{u_k, 1 - u_k\}$ forms a probability distribution on the binary measurement outcome space $\{“k”, “not k”\}$. Let

us define the random variable V , which takes the values $\{“k”, “not k”\}$ with associated probabilities $\{u_k, 1 - u_k\}$. Then V is an asymmetric Bernoulli random variable [54] (it is asymmetric since, in general, $u_k \neq \frac{1}{2}$), and we can call outcome “ k ” a success and outcome “not k ” a failure.

Let $\hat{u}_k$ be the estimator of u_k obtained in step 2b of the protocol. Since u_k is just the probability of success in a Bernoulli random variable, we can use maximum likelihood estimation (MLE) techniques [54] to determine the number of trials required to estimate u_k to some desired accuracy and confidence. As the trials performed in the protocol are independent, standard MLE gives that $\hat{u}_k$ is just equal to the frequency of obtaining outcome “ k .” Thus, if N is the number of trials and n_k is the number of times outcome “ k ” is observed, standard MLE gives the estimate

$$\hat{u}_k = \frac{n_k}{N}, \quad (9.3)$$

which agrees with intuition.

There can be a problem with this procedure if u_k is very close to 1. In particular, if k is observed in all N trials, then a probability of 1 is assigned to “ k ” and a probability of 0 is assigned to “not k ,” which is not a physically realistic scenario. This is an example of a much more general problem that can arise when using MLE to assign probabilities to rare events that are not observed because the data set is not large enough. One solution to this problem is to utilize additive (Laplace) smoothing methods to augment the MLE procedure [55]. The main idea behind such a smoothing technique is to assign higher (nonzero) weight to low-probability outcomes. In our case, the low-probability outcome is “not k ” and the weight can be controlled by a parameter λ according to the formula

$$\hat{u}_k = \frac{n_k + \lambda}{N + 2\lambda}. \quad (9.4)$$

The factor of 2 multiplying λ in the denominator arises because there are two possible outcomes of the experiment and is required to ensure that $\{\hat{u}_k, 1 - \hat{u}_k\}$ is a probability distribution. Note that when this smoothing technique is used, it will tend to fairly bias the estimate of u_k to be smaller than the actual value of u_k , so the estimation will be honest.

The key point in terms of time complexity, which we now prove, is that the number of trials N required to estimate u_k to accuracy ϵ with confidence $1 - \delta$ is *independent* of the size of the system. Since the set of possible probability distributions $\{u_k, z_k\}$ satisfies certain consistency conditions, the estimator $\hat{u}_k$ converges in distribution to u_k :

$$\hat{u}_k \xrightarrow{D} u_k. \quad (9.5)$$

Moreover, the mean of $\hat{u}_k$ is equal to u_k and the variance of $\hat{u}_k$ scales as

$$\text{Var}(\hat{u}_k) \sim \frac{1}{NI(\{u_k, z_k\})}, \quad (9.6)$$

where $I(\{u_k, z_k\})$ is the Fisher information [54] of one observation of the true probability distribution. The Fisher information contained in one observation is a measure of how much information about u_k is gained on average from observing $\{“k”, “not k”\}$ (with distribution $\{u_k, z_k\}$). If $u_k \sim 1$, then the amount of information gained on average is large.

Since $\{“k,” “not k”\}$ is a Bernoulli random variable, it is possible to explicitly compute the Fisher information of one observation [54]:

$$I(\{u_k, z_k\}) = \frac{1}{u_k z_k} = \frac{1}{u_k(1-u_k)}. \quad (9.7)$$

Hence, from Eqs. (9.6) and (9.7),

$$\text{Var}(\hat{u}_k) \sim \frac{u_k(1-u_k)}{N}. \quad (9.8)$$

Now, suppose we want to estimate u_k to accuracy ϵ with confidence $1 - \delta$, i.e., we want

$$\mathbb{P}(|\hat{u}_k - u_k| \geq \epsilon) \leq \delta. \quad (9.9)$$

By Chebyshev's theorem, for any integer $j \geq 1$,

$$\mathbb{P}(|\hat{u}_k - u_k| \geq j\sigma(\hat{u}_k)) \leq \frac{1}{j^2}. \quad (9.10)$$

Choose j_δ to be the smallest j such that

$$\frac{1}{j_\delta^2} \leq \delta. \quad (9.11)$$

Then

$$\mathbb{P}\left(|\hat{u}_k - u_k| \geq \frac{j_\delta \sqrt{u_k(1-u_k)}}{\sqrt{N}}\right) \leq \frac{1}{j_\delta^2}, \quad (9.12)$$

and so we set

$$\frac{j_\delta \sqrt{u_k(1-u_k)}}{\sqrt{N}} \leq \epsilon. \quad (9.13)$$

This gives

$$N \geq \frac{j_\delta^2 u_k(1-u_k)}{\epsilon^2}. \quad (9.14)$$

In total, if $N \geq \frac{j_\delta^2 u_k(1-u_k)}{\epsilon^2}$ with $j_\delta \geq \frac{1}{\delta}$, then

$$\mathbb{P}(|\hat{u}_k - u_k| \geq \epsilon) \leq \delta. \quad (9.15)$$

Under the assumption that the u_k are independent of d , N is independent of d . Hence, the number of trials required to estimate u_k to accuracy ϵ and confidence $1 - \delta$ is independent of d . If the probabilities u_k do depend on d , then the number of trials will be a function of d . Determining the extent to which the u_k can have dependence on d is an interesting question and likely depends on the particular scenario at hand. This is left as a topic for further investigation.

2. Step 2c in Sec. VIII

In the protocol in Sec. VIII, we estimate both u_k and Q_k , where u_k is the probability of obtaining outcome “ k ” on the first measurement and Q_k is the probability of obtaining outcome “ k ” on the second measurement *given* that the result of the first measurement is “ k .” Note that observing “ k ” in the second measurement is also a Bernoulli random variable, so the discussion from above carries over in an analogous manner to estimating Q_k .

Let N_1 and N_2 be the number of trials required to estimate u_k and Q_k to each of their respective accuracies and confidences. Suppose we perform the first measurement N_1 times (so we have estimated u_k to its desired accuracy and confidence). If the number of times “ k ” is observed over these trials is greater

than N_2 , then we have estimated both u_k and Q_k to their desired accuracies. If the number of times “ k ” is observed in the first measurement is less than N_2 , then we keep repeating until N_2 outcomes of “ k ” are recorded in the first measurement. Let M_2 be the number of times the first measurement has to be performed before N_2 values of “ k ” are recorded. The total number of trials is no more than

$$N = \min\{N_1, M_2\}. \quad (9.16)$$

As N only depends on the accuracies, confidences, and values of u_k and Q_k , it is independent of d .

B. Number of trials required in step 3 in Sec. IV and step 3 in Sec. VIII

Let us now discuss how many trials are required in step 3 of each of the protocols. The argument is the same for each protocol, so without loss of generality, we use the notation from Sec. IV (the discussion in Sec. VIII follows by replacing X with Z and making appropriate changes). The random variable $X : \mathcal{D} \rightarrow [0, 1]$ is defined by

$$X(l, m) = \sqrt{u_l u_m} \quad (9.17)$$

[in Sec. VIII this will be $Z(l, m) = \sqrt{u_l u_m} \sqrt{Q_l Q_m}$] and the expectation value of X is denoted $\bar{X}$. Suppose one wants to estimate $\bar{X}$ to accuracy ϵ and confidence $1 - \delta$. Let $\hat{\bar{X}}$ denote the estimator of $\bar{X}$ obtained from taking the average of K independent samples, $\{X_1, \dots, X_K\}$, of X :

$$\hat{\bar{X}} = \frac{1}{K} \sum_{(k_1, k_2)} \sqrt{u_{k_1} u_{k_2}}. \quad (9.18)$$

By Hoeffding's inequality

$$\mathbb{P}(|\hat{\bar{X}} - \bar{X}| \geq \epsilon) \leq 2e^{\frac{-2(K\epsilon)^2}{K(b-a)^2}} = 2e^{\frac{-2K\epsilon^2}{(b-a)^2}}, \quad (9.19)$$

where $[a, b]$ is the range of X (here $[a, b] \subseteq [0, 1]$). Hence, setting

$$\delta = 2e^{\frac{-2K\epsilon^2}{(b-a)^2}} \quad (9.20)$$

gives

$$K = \frac{\ln\left(\frac{2}{\delta}\right)(b-a)^2}{2\epsilon^2} \leq \frac{\ln\left(\frac{2}{\delta}\right)}{2\epsilon^2}, \quad (9.21)$$

which is independent of d . In practice, $b - a \ll 1$, which will improve this bound on K .

C. Complete reconstruction of POVM elements

For completeness, we provide an explicit protocol and determination of the time complexity required to perform complete tomography of the noisy POVM elements $\{E_k\}_{k=1}^d$. The idea is to input various pure states into the noisy measurement and analyze the frequency of obtaining particular outcomes. If the matrix representation of E_k in the $|\psi_m\rangle$ basis is written as $E_k^{i,j}$, then there are $\frac{d^2+d}{2}$ elements in the main diagonal and upper triangular section of E_k that need to be estimated (since each $E_k \geq 0$, the lower triangular part of E_k is completely determined by the upper triangular part). One

can estimate the elements of E_k by first defining the pure states,

$$|\psi_{i,j}^+\rangle = \frac{|\psi_i\rangle + |\psi_j\rangle}{\sqrt{2}} \quad \text{and} \quad |\psi_{i,j}^-\rangle = \frac{|\psi_i\rangle - |\psi_j\rangle}{\sqrt{2}}.$$

Then, since

$$\begin{aligned} |\psi_i\rangle\langle\psi_j| &= |\psi_{i,j}^+\rangle\langle\psi_{i,j}^+| + i|\psi_{i,j}^-\rangle\langle\psi_{i,j}^-| \\ &\quad - \left(\frac{1+i}{2}\right)|\psi_i\rangle\langle\psi_i| - \left(\frac{1+i}{2}\right)|\psi_j\rangle\langle\psi_j|, \end{aligned} \quad (9.22)$$

we can see that

$$\begin{aligned} E_k^{i,j} &= \text{tr}(E_k \Pi_{i,j}^+) + i \text{tr}(E_k \Pi_{i,j}^-) \\ &\quad - \left(\frac{1+i}{2}\right) \text{tr}(E_k \Pi_i) - \left(\frac{1+i}{2}\right) \text{tr}(E_k \Pi_j). \end{aligned} \quad (9.23)$$

The algorithm to determine the set $\{E_k\}_{k=1}^d$ is as follows.

Step 1: For each state $|\psi_{j,j}\rangle \in \{|\psi_{k,k}\rangle\}_{k=1}^d$, input $|\psi_{j,j}\rangle$ into the noisy measurement $\mathcal{E}$ many times and record the frequency of obtaining *each* of the d different possible outcomes “ k .” For each k this gives $\text{tr}(E_k |\psi_{j,j}\rangle\langle\psi_{j,j}|)$.

Step 2: For each state $|\psi_{i,j}^+\rangle \in \{|\psi_{k,l}^+\rangle\}_{k,l=1}^d$, input $|\psi_{i,j}^+\rangle$ into the noisy measurement $\mathcal{E}$ many times and record the frequency of obtaining *each* of the d different possible outcomes “ k .” For each k this gives $\text{tr}(E_k |\psi_{i,j}^+\rangle\langle\psi_{i,j}^+|) = \text{tr}(E_k \Pi_{i,j}^+)$.

Step 3: For each state $|\psi_{i,j}^-\rangle \in \{|\psi_{k,l}^-\rangle\}_{k,l=1}^d$, input $|\psi_{i,j}^-\rangle$ into the noisy measurement $\mathcal{E}$ many times and record the frequency of obtaining *each* of the d different possible outcomes “ k .” For each k this gives $\text{tr}(E_k |\psi_{i,j}^-\rangle\langle\psi_{i,j}^-|) = \text{tr}(E_k \Pi_{i,j}^-)$.

Step 4: Combine all of the elements estimated in steps 1 through 3 to construct the E_k . Step 1 gives the diagonal elements of the E_k since

$$E_k^{j,j} = \text{tr}(E_k |\psi_{j,j}\rangle\langle\psi_{j,j}|). \quad (9.24)$$

Equation (9.23) and steps 1 through 3 give all of the off-diagonal elements $E_k^{i,j}$. This concludes the protocol.

The number of trials one will have to perform is again dictated by MLE. It is important to note that the MLE procedure in this case is more involved than the simple Bernoulli procedure for our protocol (described in Sec. IX A). This is because, when the noisy measurement is performed, one must keep track of which value of j is obtained (not just whether or not the outcome was j). Thus the number of trials in each step will be greater than that required to estimate the u_k because events with a low probability may rarely be seen (if at all). Smoothing techniques will likely have to be employed to ensure that rare events are not assigned zero probability. Thus, the determination of each $\text{tr}(E_k |\phi\rangle\langle\phi|)$, where $|\phi\rangle$ is one of $|\psi_{j,j}\rangle$, $|\psi_{i,j}^+\rangle$, or $|\psi_{i,j}^-\rangle$, requires greater time complexity than estimating each u_k . Since there are $d + 2\frac{d(d-1)}{2} = d^2$ such $|\phi\rangle$, one will have to estimate d^3 different probabilities over steps 1 through 3 using MLE. Hence, the full reconstruction requires the estimation of d^3 probabilities with more complicated postprocessing of the measurement data (as well as a larger number of trials to estimate each probability). In addition, there is an added complexity in preparing d^3 different input states since more complex rotations may be required.

X. DISCUSSION

We have provided a straightforward, efficient, and experimentally implementable method for obtaining estimates for lower bounds on the average fidelity of projective (rank 1) quantum measurements. As realizations of quantum protocols scale to larger sizes, and full measurement tomography becomes impossible to implement, our protocol can potentially be used as a simple method to benchmark the performance of a measuring device.

We have discussed conditions for the validity of the bounds and explained why they should hold in extremely general situations. The bounds could also potentially be useful as estimates of the average measurement fidelity in the small-error regime. We have presented a set of numerical examples for a single-qubit system. In every instance analyzed, that is, for all possible values of the coherence in the noisy POVM operators, the bounds were found to be valid. This provides further evidence that the bounds should hold in extremely general situations and should be useful in practice. In addition, the bounds became better approximations of the actual measurement fidelity as the magnitude of the coherence increased.

The protocols are scalable and only require the ability to prepare states from a basis set and perform the noisy measurement (sequentially when there are output states) to estimate d probabilities. In addition, postprocessing of the data is completely straightforward and avoids the difficulties in associating large sets of tomographic data to valid mathematical objects. This can be compared with a full reconstruction of the POVM elements of the noisy measurement, which requires the preparation of d^2 input states, the estimation of d^3 probabilities, and more involved postprocessing. In addition, the d^2 input states required for tomography can be highly complex. In many situations, these input states will have to be prepared using complicated unitary rotations. Minimizing the required set of input states, as well as their complexity, is important for obtaining a more faithful characterization of the measurement that is less prone to state-preparation errors.

It is important to note that the experimental protocols provide an *estimate* for a lower bound on the average measurement fidelity. More precisely, one chooses a number of trials K to obtain the estimate, where K depends on the desired accuracy ϵ and confidence $1 - \delta$ of the estimate. Thus, under the assumption that the lower bound is valid, the true value of the average measurement fidelity is no more than ϵ units of distance *below* the estimated lower bound with confidence $1 - \delta$.

There are a number of different questions and avenues for future research. First, we have focused on the case of rank 1 PVMs, however, we expect that our results can be extended to higher rank PVMs, especially low-rank PVM's in large Hilbert spaces. As well, since any POVM can be implemented via a PVM on an extended Hilbert space, the protocol can potentially give information regarding the quality of implementations of POVM measurements. It will also be useful to analyze the extent to which the ideas presented here can be used to characterize no-ideal POVM measurements that are not implemented as PVMs on a larger space.

While we have shown that the bounds derived here should hold in large generality, a deeper understanding of the validity of the bounds will clearly be useful. In most physically relevant cases, where the coherence is not overwhelmingly large, the protocol should produce valid bounds. In any noise estimation or characterization scheme, there is a trade-off between the amount of information one is able to extract and the amount of resources required to implement the scheme. Here, we obtain a single parameter which serves as an upper (lower) bound on the error (fidelity) of the measurement in a scalable amount of time. Ideally, one would like as much information about the physical measurement as possible. Further analysis of schemes that give more information about an imperfect measurement than average fidelities or errors, while retaining properties such as scalability, will clearly be useful.

Finally, as previously mentioned, the algorithms given here require the ability to prepare the basis states $\{|\psi_k\rangle\}$ which constitute the measurement. In practice, these states typically have errors and may actually be created by a measurement procedure, which is exactly what we want to characterize. There are various systems, however, where state preparation is very different from the measurement procedure. For instance, in superconducting qubit systems, state preparation is performed by cooling the system to low temperatures and performing unitary rotations via a circuit-QED setup that couples the qubits to superconducting resonators. Measurements are also performed using the circuit-QED setup, however, in this case,

one drives the cavity at a large detuning with the qubit and observes the transmitted radiation. Typically, unitary rotations have much higher fidelities than measurements in this system and so one expects state preparation to be much more accurate and robust than performing measurements. We note that running the protocol with noisy states can still provide valid bounds. For instance, if the noisy input states are given by a convex combination of elements of $\{|\psi_k\rangle\}$, then the bounds still hold. The same is also true for noisy input states with small coherence in the $\{|\psi_k\rangle\}$ basis. The performance of the bounds under more general noise models on input states is a topic for future research.

The scalable protocols presented here can be useful for determining the quality of experimental quantum measurements. There is still much to investigate with regard to useful metrics for comparing measurements and proposing experimentally efficient methods for characterizing measurement devices. As experimental quantum systems scale to larger sizes, such methods will be useful for characterizing and controlling multiqubit systems.

ACKNOWLEDGMENTS

The authors are grateful for helpful discussions with Alexandre Cooper, Joseph Emerson, Jay Gambetta, Masoud Mohseni, and Marcus Silva. E.M. acknowledges financial support from the National Science Foundation through Grant No. NSF PHY-1125846.

-
- [1] R. Feynman, *Int. J. Theor. Phys.* **21**, 467 (1982).
 - [2] V. Giovannetti, S. Lloyd, and L. Maccone, *Science* **306**, 1330 (2004).
 - [3] P. Shor, *Proceedings of the 35'th Annual Symposium on Foundations of Computer Science (FOCS)* (IEEE Press, Los Alamitos, CA, 1994), pp. 124–134.
 - [4] S. Lloyd, *Science* **273**, 1073 (1996).
 - [5] D. Deutsch, *Proc. R. Soc. London A* **400**, 97 (1985).
 - [6] R. Raussendorf and H. J. Briegel, *Phys. Rev. Lett.* **86**, 5188 (2001).
 - [7] A. Kitaev, *Ann. Phys.* **303**, 2 (1997).
 - [8] V. Bouchiat *et al.*, *Phys. Scripta A* **T76**, 165 (1998).
 - [9] J. Koch, T. M. Yu, J. Gambetta, A. A. Houck, D. I. Schuster, J. Majer, A. Blais, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf, *Phys. Rev. A* **76**, 042319 (2007).
 - [10] J. Wrachtrup and F. Jelezko, *J. Phys.: Condens. Matter* **18**, S807 (2006).
 - [11] J. I. Cirac and P. Zoller, *Phys. Rev. Lett.* **74**, 4091 (1995).
 - [12] D. Cory, A. Fahmy, and T. Havel, *Proceedings of the Fourth Workshop on Physics and Computation* (Elsevier, Amsterdam, The Netherlands, 1996).
 - [13] D. Loss and D. P. DiVincenzo, *Phys. Rev. A* **57**, 120 (1998).
 - [14] E. Knill, R. Laflamme, and G. J. Milburn, *Nature* **409**, 46 (2001).
 - [15] I. Chuang and M. Nielsen, *J. Mod. Opt.* **44**, 2455 (1997).
 - [16] J. B. Altepeter, D. Branning, E. Jeffrey, T. C. Wei, P. G. Kwiat, R. T. Thew, J. L. O'Brien, M. A. Nielsen, and A. G. White, *Phys. Rev. Lett.* **90**, 193601 (2003).
 - [17] M. Mohseni and D. A. Lidar, *Phys. Rev. Lett.* **97**, 170501 (2006).
 - [18] J. S. Lundeen *et al.*, *Nature Phys.* **5**, 27 (2009).
 - [19] L. Zhang *et al.*, *Nature Photon.* **6**, 364 (2012).
 - [20] S. Merkel *et al.*, [arXiv:1211.0322](https://arxiv.org/abs/1211.0322).
 - [21] T. Monz, P. Schindler, J. T. Barreiro, M. Chwalla, D. Nigg, W. A. Coish, M. Harlander, W. Hänsel, M. Hennrich, and R. Blatt, *Phys. Rev. Lett.* **106**, 130506 (2011).
 - [22] P. Shor, *Proceedings of the 37'th Annual Symposium on Foundations of Computer Science (FOCS)* (IEEE Press, Burlington, VT, 1996).
 - [23] D. Aharonov and M. Ben-Or, *Proceedings of the 29th Annual ACM Symposium on Theory of Computing (STOC)* (Association for Computing Machinery, New York, 1997).
 - [24] E. Knill, R. Laflamme, and W. Zurek, *Proc. R. Soc. London A* **454**, 365 (1997).
 - [25] J. Preskill, [arXiv:quant-ph/9712048](https://arxiv.org/abs/quant-ph/9712048).
 - [26] J. Emerson *et al.*, *Science* **317**, 1893 (2007).
 - [27] M. Silva, E. Magesan, D. W. Kribs, and J. Emerson, *Phys. Rev. A* **78**, 012347 (2008).
 - [28] E. Knill, D. Leibfried, R. Reichle, J. Britton, R. B. Blakestad, J. D. Jost, C. Langer, R. Ozeri, S. Seidelin, and D. J. Wineland, *Phys. Rev. A* **77**, 012307 (2008).
 - [29] E. Magesan, J. M. Gambetta, and J. Emerson, *Phys. Rev. Lett.* **106**, 180504 (2011).
 - [30] E. Magesan *et al.*, *Phys. Rev. Lett.* **109**, 080505 (2012).
 - [31] O. Moussa, M. P. da Silva, C. A. Ryan, and R. Laflamme, *Phys. Rev. Lett.* **109**, 070504 (2012).

- [32] A. Bendersky, F. Pastawski, and J. P. Paz, *Phys. Rev. Lett.* **100**, 190403 (2008).
- [33] M. P. da Silva, O. Landon-Cardinal, and D. Poulin, *Phys. Rev. Lett.* **107**, 210404 (2011).
- [34] C. T. Schmiegelow, A. Bendersky, M. A. Larotonda, and J. P. Paz, *Phys. Rev. Lett.* **107**, 100502 (2011).
- [35] S. T. Flammia and Y.-K. Liu, *Phys. Rev. Lett.* **106**, 230501 (2011).
- [36] M. Nielsen, *Phys. Lett. A* **303**, 249 (2002).
- [37] J. Emerson, R. Alicki, and K. Życzkowski, *J. Opt. B* **7**, S347 (2005).
- [38] E. Magesan, R. Blume-Kohout, and J. Emerson, *Phys. Rev. A* **84**, 012309 (2011).
- [39] O. Oreshkov and J. Calsamiglia, *Phys. Rev. A* **79**, 032336 (2009).
- [40] V. Paulsen, *Completely Bounded Maps and Operator Algebras*, Vol. 78 (Cambridge University Press, Cambridge, UK, 2002).
- [41] D. Gottesman, Ph.D. thesis, Cornell University, 1997; [arXiv:quant-ph/9705052](#).
- [42] D. Gottesman, *Chaos Solitons Fractals* **10**, 1749 (1999).
- [43] M. D. Reed *et al.*, *Nature* **482**, 382 (2012).
- [44] D. DiVincenzo and F. Solgun, [arXiv:1205.1910](#).
- [45] A. Gilchrist, N. K. Langford, and M. A. Nielsen, *Phys. Rev. A* **71**, 062310 (2005).
- [46] I. Bengtsson and K. Życzkowski, *Geometry of Quantum States: An Introduction to Quantum Entanglement* (Cambridge University Press, Cambridge, UK, 2006).
- [47] C. Fuchs and J. van de Graaf, *IEEE Trans. Inf. Theory* **45**, 1216 (1999).
- [48] A. Kitaev, *Russ. Math. Surv.* **52**, 1191 (1997).
- [49] S. Beigi and R. Koenig, *New J. Phys.* **13**, 093036 (2011).
- [50] J. Renes *et al.*, *J. Math. Phys.* **45**, 2171 (2004).
- [51] R. Horn and C. Johnson, *Matrix Analysis* (Cambridge University Press, Cambridge, UK, 1990).
- [52] M. Ledoux, *The Concentration of Measure Phenomenon* (American Mathematical Society, Providence, RI, 2001).
- [53] C. A. Fuchs, Ph.D. thesis, University of New Mexico, 2006; [arXiv:quant-ph/9601020](#).
- [54] M. J. Schervish, *Theory of Statistics*, 1st ed. (Springer, New York, 1997).
- [55] C. Manning, P. Raghavan, and H. Schütze, *Introduction to Information Retrieval* (Cambridge University Press, New York, 2008).