Choice of measurement sets in qubit tomography

Mark D. de Burgh,[1] Nathan K. Langford,[1] Andrew C. Doherty,[1] and Alexei Gilchrist[2]

[1]*School of Physical Sciences, University of Queensland, St Lucia, Queensland 4072, Australia*
[2]*Physics Department, Macquarie University, Sydney NSW 2109, Australia*

Optimal generalized measurements for state estimation are well understood. However, practical quantum state tomography is typically performed using a fixed set of projective measurements, and the question of how to choose these measurements has been largely unexplored in the literature. In this work, we develop theoretical asymptotic bounds for the average fidelity of pure qubit tomography using measurement sets whose axes correspond to faces of Platonic solids. We also present comprehensive simulations of maximum likelihood tomography for mixed qubit states using the Platonic solid measurements. We show that overcomplete measurement sets can be used to improve the accuracy of tomographic reconstructions.

PACS number(s): 03.65.Ta, 03.67.−a

I. INTRODUCTION AND BACKGROUND

Quantum tomography [1], the practical estimation of quantum states through the measurement of large numbers of copies, is of fundamental importance in the study of quantum mechanics. With the emergence of quantum-information science, the tomographic reconstruction of finite-dimensional systems [2] has also become an essential technology for characterizing the experimental performance of practical quantum gates and state preparation. Examples include tomography of the polarization states of light [3,4] and electronic states of trapped ions [5,6]. It is important for these experiments that a systematic study is carried out to determine efficient practical measurement and state estimation strategies.

There has already been much theoretical work in this area, and optimal bounds on state estimation as well as constructions for measurements that achieve these bounds are known [7–14]. However, these bounds require all copies of the state to be collected and a combined measurement performed across all the copies. While these collective measurements are known to be more powerful than performing independent measurements on each copy [8], they are totally infeasible for any existing experimental implementation.

Experimental tomography in discrete quantum systems is currently performed using independent measurements in fixed sets of projectors [15,16], but it remains an important open question as to which fixed sets are best to use (although there are several relevant theoretical discussions [13,17–25]). For mainly historical reasons, measurement sets such as those described in [15] have become popular, particularly in optical experiments (except, e.g., [4,26]). These measurement sets are attractive since they have the minimal number of settings, and in these experiments a significant amount of time can be spent changing the measurement settings. However, such decisions are generally made without a quantitative understanding of how the resulting tomographic performance is affected. For example, in Fig. 1 we show that both the average and worst-case fidelities of the tomographic reconstruction are greatly improved by replacing the popular minimal measurement set of [15] with an overcomplete set of six measurements corresponding to the six faces of a cube.

In this paper, we investigate how the choice of measurements affects the quality of tomographic reconstruction for qubit systems. We follow Jones [17] and investigate a class of measurement sets based on Platonic solids. This class gives close-to-optimal performance for tomography when using independent measurements of fixed projectors by spreading the projectors uniformly over the surface of the Bloch sphere. The different Platonic solids allow us to investigate how the number of measurement settings affects the tomographic reconstruction quality, and we show that the more overcomplete sets produce the best results. We begin by deriving analytical bounds for the performance of Platonic solid measurements in the tomography of pure states. We then extend this analysis to systems of one- and two-qubit mixed states by performing comprehensive numerical simulations of the tomographic procedure for the two important cases of photonic and atomic experiments. We present our results as a function of the total number of copies of the quantum state used in the experiment. This work provides a detailed study of how information is acquired during a tomography experiment. To support these results, it would also be very interesting to perform such a study in some common experimental systems, because it would determine whether the quality of the reconstructions was limited by the expected statistics or by other experimental imperfections, such as drifting sources. To the best of our knowledge, such a study has still not been performed.

Platonic solid measurements were first proposed for tomography in [17], where analytical asymptotic bounds on the reconstruction performance were calculated for the special case of pure states using mutual information as the figure of merit. It is difficult to compare these results with more recent work, however, where the most commonly used figure of merit is the average fidelity. We therefore derive new asymptotic bounds for the pure-state case based on the average fidelity. Much of the original interest in fidelity arises from the fact that it bounds the distinguishability of quantum states; see [27]. Recently, a true quantum Chernoff bound on the distinguishability of states has been introduced [28], and this quantity is a tighter bound than the fidelity. We therefore also consider this quantity as a figure of merit and show that the results do not depend greatly on this choice.
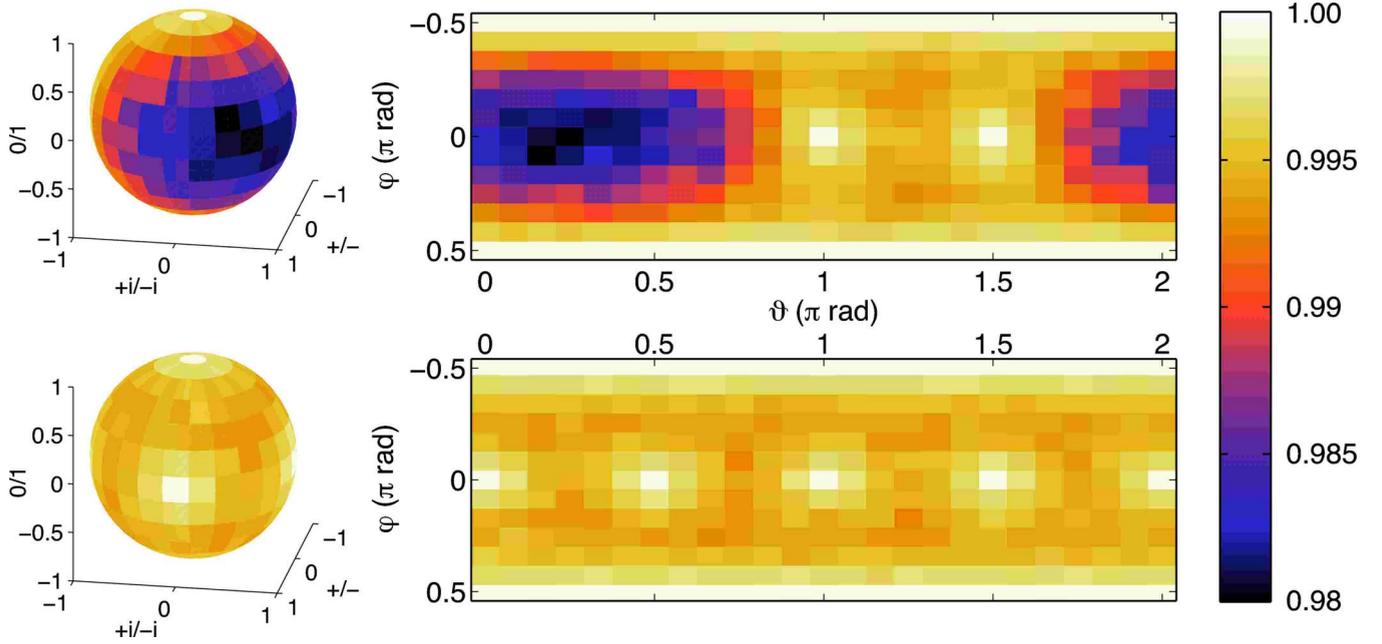
FIG. 1. (Color online) Performance of tomographically reconstructed states on the Bloch sphere for two measurement sets. The top set is a popular set of measurements (James4) used in the literature [15], the bottom set is a more isotropic set composed of measurements along the three spatial axes (a cube measurement set); see text for details. For each target state, the average fidelity with 400 reconstructed states is plotted by color. The ensemble of reconstructed states was generated by adding Possonian noise to a simulated experiment with an average total of 4000 counts, and then performing maximum-likelihood tomography.

Very recently, Roy and Scott [21] were able to derive results about mixed-state reconstruction performance by using another figure of merit, the mean-squared Hilbert-Schmidt distance, choosing this measure specifically because it allowed tractable analytical solutions. In this way, they identified a class of measurements, including all the Platonic solid measurements, that provide optimal reconstruction performance for qubits. In contrast, by using the physically well-motivated average fidelity and quantum Chernoff bound figures of merit for numerical simulations, we show that there is a range of performance within the class of Platonic solids, with the higher-order solids performing better. Perhaps surprisingly, this indicates that the choice of figure of merit can have a qualitative effect on the comparison between tomographic procedures. We will briefly discuss the ramifications and possible causes of this result.

The structure of the paper is as follows. In Sec. II, we introduce the tomography problem, discussing how to assess the performance of tomographic reconstructions. We then define various relevant one- and two-qubit measurement sets in Sec. III, including Platonic solid measurements and some others that have been studied in the literature for comparison purposes. In Sec. IV, we present some mostly analytical results that motivate our more complete numerical investigations and illustrate some interesting effects that should be studied experimentally. In particular, Sec. IV B provides some important intuition about the complications that arise when reconstructing mixed states. We then describe in detail our numerical mixed-state simulations for atomic and photonic qubit systems in Sec. V and present the results for both the fidelity and the quantum Chernoff bound in Sec. VI. This includes a comparison with recent analytical results based on the mean-squared Hilbert-Schmidt distance in Sec. VID.

II. TOMOGRAPHIC RECONSTRUCTION AND RECONSTRUCTION QUALITY

A d-dimensional quantum state ρ is represented by a $d \times d$ positive semidefinite density matrix, with trace 1. There are d^2-1 real parameters to be estimated. In the case of optical experiments such as [3,4], the flux must also be estimated giving d^2 real parameters to be estimated.

Any setting l of an experimental apparatus designed to measure the quantum state may be described by a positive operator value measure (POVM). Each of the k outcomes of a measurement setting is represented by a positive semidefinite operator O_{lk}. The operators satisfy $\Sigma_k O_{lk}=I$. The probability of observing outcome k is given by the Born rule $p_{lk}=\text{tr}(\rho O_{lk})$. In state tomography, each of the l measurements is performed on a large number of copies and estimates obtained for each of the p_{lk}. We denote these estimates $\hat{p}_{lk}$. When the number of linearly independent O_{lk} equals the number of parameters to be estimated, our measurements are known as informationally complete [29]. If these probability estimates were perfect, so that $\hat{p}_{lk}=p_{lk}$, it would be possible to reconstruct the state exactly. In this case, there is a set of operators R_{lk}, known as a dual basis or dual frame, such that $\rho=\Sigma_{lk}p_{lk}R_{lk}$. (When the number of measurements O_{lk} in an informationally complete POVM exceeds the number of parameters to be estimated, the dual frame is not unique.) Since our probability estimates are not perfect due to the finite sample size, they may well be inconsistent with any such

reconstruction. In this case, it is common practice to resort to a maximum likelihood estimate of the state [15,16]. We discuss maximum likelihood reconstruction in detail in Sec. V C.

A. Measures for comparing mixed states

To judge the quality of reconstruction, we need a measure of similarity between the true state ρ and the reconstructed state $\hat{\rho}$. In this paper, we will concentrate on the following two measures.

The most commonly used is the *fidelity*,

$$f(\rho,\hat{\rho}) = [\mathrm{tr}\sqrt{\sqrt{\hat{\rho}}\rho\sqrt{\hat{\rho}}}]^2, \qquad (1)$$

where $0 \leq f(\rho,\hat{\rho}) \leq 1$ and $f=1$ implies $\rho = \hat{\rho}$. The advantages of the fidelity are that it is always easy to calculate, and when one of the states is pure ($\rho = |\psi\rangle\langle\psi|$), it reduces to $f(\psi,\hat{\rho}) = \langle\psi|\hat{\rho}|\psi\rangle$, which has a good physical interpretation. That is, if an experimentalist trying to prepare the pure state ψ actually creates the mixed state $\hat{\rho}$, then the single-shot probability of error (that the wrong state has been prepared) is simply $P_e = 1 - f(\psi,\hat{\rho})$. In the literature, the square root of this quantity is often also termed the fidelity and can serve as an alternative measure of distinguishability, but it no longer has the same nice interpretation. All of the theoretical results we mention are derived using the definition in Eq. (1). As yet, however, no physical interpretation is known for either definition if both states are mixed.

The second quantity we use is the nonlogarithmic version of the recently derived *quantum Chernoff bound* [28],

$$\lambda_{cb}(\rho,\hat{\rho}) = \min_{0 \leq s \leq 1} \mathrm{tr}[\rho^s\hat{\rho}^{1-s}], \qquad (2)$$

where as for the fidelity, $0 \leq \lambda_{cb}(\rho,\hat{\rho}) \leq 1$ and $\lambda_{cb} = 1$ also implies $\rho = \hat{\rho}$. Although this minimization is a convex optimization problem in only one variable and therefore comparatively straightforward, the quantum Chernoff bound is obviously more difficult to calculate than the fidelity. This is particularly significant when trying to derive quantitative analytical results (as in Sec. IV A, where we use the fidelity). However, its main advantage is that it has a clear physical interpretation *even when both states are mixed*. Suppose an experimentalist wishes to determine if her state preparation device creates the state ρ or the state σ. Time is limited and she only has N identically prepared copies of the state to work with. She is assumed to be able to perform any measurement, including a collective measurement, in her efforts to distinguish the two states. The result of [28] is that the probability of her making an error is asymptotically $P_e \approx \lambda_{cb}(\rho,\sigma)^N$ (N large). When one of the states is pure, the quantum Chernoff bound is the same as the fidelity. Also, the square root of the fidelity is always an upper bound on the quantum Chernoff bound [27,28], but this is not always a tight bound, since when one state is pure, $\lambda_{cb}(\psi,\hat{\rho}) = f(\psi,\hat{\rho}) \leq \sqrt{f(\psi,\hat{\rho})}$ with equality only if $\hat{\rho} = \psi$. We suggest that the quantum Chernoff bound is particularly well-suited to characterizing the quality of tomographic reconstruction schemes, because it is defined specifically in the regime most relevant to tomography, where many copies of the quantum system are being assessed.

Next, we define quantities based on the fidelity and the quantum Chernoff bound that we will use to rate the performance of a tomography scheme. The figure of merit we adopt is based on how similar the reconstructed state is to the true state, when averaged over measurement outcomes and an appropriate ensemble of true states. (An interesting alternative approach was advocated by Blume-Kohout and Hayden [30].) A tomography experiment produces a list of outcomes of measurements performed on N copies of the true state ρ. These outcomes are used to obtain an estimate of the state $\hat{\rho}$, and the fidelity (or the quantum Chernoff bound) between the true and estimated states is calculated. Averaging over the outcomes of many such tomography experiments, we define the *point fidelity*,

$$F_{pt}(\rho,N) = \sum_{\hat{\rho}} p_\rho(\hat{\rho}) f(\rho,\hat{\rho}), \qquad (3)$$

and the *point Chernoff bound*,

$$\lambda_{pt}(\rho,N) = \sum_{\hat{\rho}} p_\rho(\hat{\rho}) \lambda_{cb}(\rho,\hat{\rho}), \qquad (4)$$

where $p_\rho(\hat{\rho})$ is the probability of estimating state $\hat{\rho}$ given the true state ρ. Finally, averaging over all possible true states, we obtain the *average fidelity*,

$$F_{av}(N) = \int d\rho F_{pt}(\rho,N), \qquad (5)$$

and the *average Chernoff bound*,

$$\lambda_{av}(N) = \int d\rho \lambda_{pt}(\rho,N). \qquad (6)$$

In other works, the mean-square Hilbert-Schmidt distance and the mean-squared Euclidean distance between Bloch vectors have been used as figures of merit for tomography [20,31]. For single qubits, these measures are equivalent up to constant factors. The Hilbert-Schmidt (HS) distance is defined as $D_{HS}(\rho,\sigma) = \sqrt{\mathrm{tr}(A^\dagger A)}$, where $A = \rho - \sigma$, and is related to the probability of error when an experimentalist must distinguish between the two states ρ and σ with only a single available copy to be measured. Provided both states are equally likely before the measurement, the probability of error in distinguishing single-qubit states ρ and σ is $P_e = \frac{1}{2} - \frac{1}{2\sqrt{2}} D_{HS}(\rho,\sigma)$. In order to obtain an analytically tractable measure of distinguishability in [20,31], the authors square D_{HS} before averaging this quantity over measurement outcomes and states as we have done above. This figure of merit makes it possible to find the optimal measurement sets analytically, but, after squaring and averaging, this quantity is not as well motivated. We will examine the effect of using this figure of merit in Sec. VI D and show that it can lead to qualitatively different behavior.

B. Generating random mixed states for numerical simulations

As described above, in order to assess the success of different tomographic schemes in a way that does not depend on the specific state being reconstructed, we use figures of merit that are formed by averaging over all possible states

[Eqs. (5) and (6)]. To perform these averages numerically, we therefore need to choose states randomly according to a probability density (or *prior*), which will produce a uniform spread throughout the allowed state space. For pure states, there is only one clear choice—the Haar measure. For mixed states, however, the choice is not so obvious (see, e.g., [32]), because the notion of uniformity is directly related to the chosen measure of distance between states (or metric). In this paper, we will use two different priors, one based on the fidelity and the other on the quantum Chernoff bound. The main idea is that given any quantity that can be used to define a metric, it is possible to define a corresponding measure of volume in the state space. The uniform prior is then the probability density, which ensures that two different regions in state space that have the same volume must be occupied with the same probability.

The first prior we use is based on the Bures metric, which is related to the fidelity via $1/2 \arccos f(\rho, \rho + d\rho)$. It is a natural choice when using a fidelity-based figure of merit, and it has also been argued that the Bures prior corresponds to maximal randomness of the input states [33]. The Bures prior is unitarily invariant, and it gives a radial probability density on the Bloch sphere,

$$p_B(r) = \frac{4}{\pi} \frac{r^2}{\sqrt{1 - r^2}}. \tag{7}$$

By contrast, when using a figure of merit based on the quantum Chernoff bound, it is natural to use a prior over states based on this measure. The natural metric induced by the quantum Chernoff bound was derived in [28]. Some of the authors of [28] have also derived the corresponding volume element in the space of states that can be used to generate a prior on qubit mixed states [34]. In Appendix B, we present our general derivation for states of arbitrary dimension. The resulting Chernoff prior is unitarily invariant and the probability density over eigenvalues is given by

$$p(\lambda_1 \cdots \lambda_M) = \frac{C}{\sqrt{\lambda_1 \cdots \lambda_M}} \prod_{k<j} \frac{(\lambda_k - \lambda_j)^2}{(\sqrt{\lambda_j} + \sqrt{\lambda_k})^2}, \tag{8}$$

where the eigenvalues are also constrained to satisfy $\Sigma_{j=1}^M \lambda_j = 1$, and C is a normalization constant. For single-qubit states, this corresponds to a radial probability density on the Bloch sphere given by

$$p_{CB}(r) = \frac{2}{\pi - 2} \frac{1 - \sqrt{1 - r^2}}{\sqrt{1 - r^2}}. \tag{9}$$

A comparison of the density of the Bures and Chernoff priors is given in Fig. 2. Both of these priors have increasing density with state purity, but the Chernoff prior is skewed slightly more toward pure states than the Bures prior.

We will present our mixed-state results first in terms of average fidelity with states drawn according to the Bures prior, which is currently the standard way of assessing mixed-state tomography, and then in terms of the average Chernoff bound with states drawn according to the Chernoff prior. Although the Chernoff bound is more difficult to calculate than the fidelity, these results provide an important
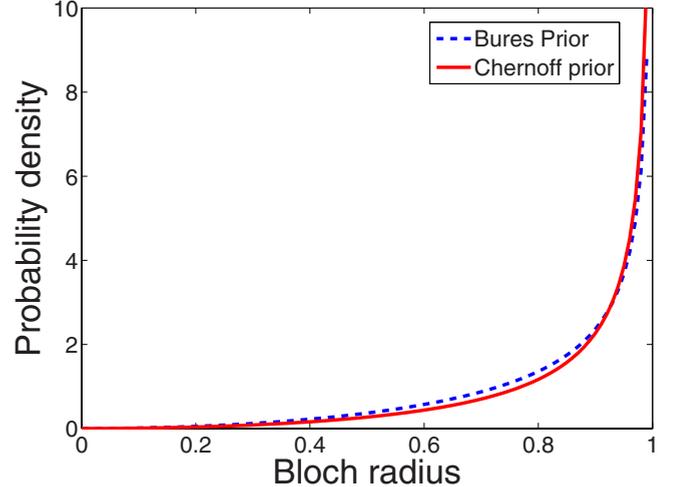
FIG. 2. (Color online) Plot of probability density against Bloch radius for the Bures and Chernoff priors over states. Both distributions have densities that increase with state purity. The Chernoff prior is slightly more skewed toward the pure states than the Bures prior.

confirmation for the qualitative conclusions drawn from the average fidelity results. We note also that the classical version of the Chernoff bound has been used previously in the evaluation of tomography schemes [35].

III. MEASUREMENT SETS

The state density matrix of a qubit can be written as

$$\rho = \frac{I + \vec{r} \cdot \vec{\sigma}}{2}, \tag{10}$$

where $\vec{r}$ is a real three-dimensional vector known as the Bloch vector (with $|\vec{r}| \leq 1$), and $\vec{\sigma}$ is a vector of the Pauli matrices σ_x, σ_y, and σ_z. In the special case of pure states, the Bloch vector is unit length. The pure states therefore define the surface of a unit sphere called the Bloch sphere. A two-outcome projective measurement with the lth measurement setting can be represented by two projectors,

$$O(\pm \vec{m}_l) = \frac{I \pm \vec{m}_l \cdot \vec{\sigma}}{2}, \tag{11}$$

where $\vec{m}_l$ is a real three-dimensional unit Bloch vector. The orthogonal outcomes correspond to opposite directions on the Bloch sphere.

For any given tomographic measurement set described by the unit Bloch vectors $\{\vec{m}_l\}$, there are two important properties that one might intuitively expect to influence the quality of a tomographic reconstruction. First, as states are arranged isotropically inside the Bloch sphere, it is natural to assume that, on average, a reconstruction would perform best if the measurements were also spaced isotropically around the surface of the sphere. This idea was supported (and made precise) by the asymptotic pure state results of [17]. See also [7,36]. It is not clear, *a priori*, whether it is advantageous to have larger measurement sets. On the one hand, increasing

the number of measurements clearly allows more sampling of the Bloch sphere and this is advantageous. On the other hand, given a fixed number of available copies of the state, increasing the number of measurements requires estimating the mean value of each measurement with fewer copies of the state. Consequently, one might expect there to be a tradeoff that would result in an optimal size for the measurement set, however the above-mentioned theoretical results suggest otherwise. Moreover, we have not found any evidence for this tradeoff, at least for the relatively small measurement sets we have considered.

The popular measurement set described in [15] has the minimal number of measurements and is not at all isotropic, and Fig. 1 shows that it is far outperformed by a symmetrical, overcomplete six-measurement set for pure states. In this paper, we will concentrate on the measurement sets whose Bloch vectors form Platonic solids. These are the five convex regular polyhedra: tetrahedron, cube, octahedron, dodecahedron, and icosahedron. These shapes are highly isotropic and thus would be expected to give good tomographic performance. Moreover, they also allow us to explore the effect of increasing the number of measurements somewhat separately from the effects of isotropy.

A. Platonic solid measurements

There are two possible ways of defining Platonic measurement sets, using measurement vectors corresponding to either the faces or the vertices of the regular polyhedra. Throughout this paper, we define a *Platonic solid measurement set* to be a set of L different measurement settings whose $2L$ Bloch vectors $\pm\vec{m}_1\cdots\pm\vec{m}_L$ match the $2L$ centers of the faces of a Platonic solid. This definition has the appealing property that an octahedron has eight measurement directions. Using the alternative definition would result in the same set of shapes because the centers of the faces of one Platonic solid match the vertices of another Platonic solid. Such solids are said to be duals of each other. The dual pairs are the tetrahedron with a rotated version of itself, the cube with the octahedron and the dodecahedron with the icosahedron.

The experiments we will be modeling are described in detail in Sec. V. In the atomic qubit experiment and the dual-detector photonic experiment each measurement has two orthogonal outcomes. These correspond to opposite directions on the Bloch sphere. Thus two opposite directions are measured simultaneously by a single measurement. All the Platonic solids except the tetrahedron have a face opposite every face, and thus these Platonic solid measurements can be measured in this way while the tetrahedron cannot. (In this case, attempting to measure the tetrahedron directions would result in us measuring the octahedron instead.) Direct measurement of the tetrahedron is possible in principle and schemes have been proposed [37,38]. However, because this is quite difficult in practice, this is not the standard way to implement photonic measurements (although see [39,40]) and will not be considered here. For the single-detector photonic qubit experimental model, we will require L measurement settings to obtain the L outcomes, as the orthogonal outcome is not

detected. This enables a tetrahedron measurement to be meaningfully measured in this case.

We will also be interested in two-qubit states. Experimentally, the simplest measurements are projective measurements performed locally on both subsystems (recording joint outcome probabilities). While we would expect measurement sets that include measurements entangled across both subsystems to give improved results, these measurements are difficult to perform. We will examine the performance of Platonic solid measurements on each subsystem, that is, measurements of the form

$$O(\pm\vec{m}_i) \otimes O(\pm\vec{m}_j), \tag{12}$$

where both i and j range over the L measurements of the same Platonic solid.

B. Other relevant measurement sets

For simulations of one- and two-qubit optical experiments (single-detector configuration), we will compare our results to the popular measurement sets described in [15]. The one-qubit state measurement set consists of measuring polarizations $\{H,V,D,R\}$, that is, the horizontal and vertical linear polarizations, the diagonal linear polarization $[|D\rangle=(|H\rangle+|V\rangle)/\sqrt{2}]$, and the right circular polarization $[|R\rangle=(|H\rangle+i|V\rangle)/\sqrt{2}]$. We refer to this measurement set as James4. The two-qubit state measurement set is $\{HH,HV,VV, VH,RH,RV,DV, DH,DR,DD, RD, HD,VD,VL, HL,RL\}$, where, for example, the measurement setting HL means measuring horizontal polarization on the first qubit subsystem and left circular polarization $[|L\rangle=(|H\rangle-i|V\rangle)/\sqrt{2}]$ on the second qubit subsystems [52]. We refer to this measurement set as James16. (Notice that this measurement set does not form complete POVMs. The implementation we have in mind is explained fully in Sec. V.)

In the case of two qubits, we also consider a modified version of the SicPOVM (symmetric informationally complete POVM) of [41] for comparison. We call this measurement a projective SicPOVM. The original SicPOVM for two-qubit states is an entangled measurement in which a single measurement produces one of 16 outcomes. Experimentally this measurement is difficult to realize. Our projective SicPOVM is a measurement set of 16 settings, in which each setting projects onto one of the directions of the SicPOVM. While still requiring measuring entangled projectors, this should be simpler to implement experimentally than the SicPOVM.

IV. ANALYTICAL RESULTS FOR PURE AND MIXED STATES

A. Estimating pure qubit states using Platonic solid measurements

The bound on average fidelity for tomographic reconstruction using collective measurements is known to be $F_{av} \leq (N+1)/(N+2) \approx 1-1/N$, where the approximation is good for large N. In fact, this bound is achievable with collective measurements [8]. It also has been shown that asymptotically this collective bound can also be achieved by an independent

measurement that is perfectly isotropic. This may be implemented by randomly guessing the measurement direction for each copy of the state [18,36]. (Experimentally, however, the time required to switch measurement settings between every count may make such a measurement unattractive.) Measuring the three orthogonal axes of the Bloch sphere (a cube measurement) is known to give an average fidelity bounded by $F_{\mathrm{av}} \leq 1 - 13/(12N)$, and again this bound is achievable [42]. We have derived average fidelity bounds for the other Platonic solids. The derivation only slightly generalizes Appendix B of [18]. We give an overview of the derivation here, leaving the details to Appendix A.

These bounds on average tomography performance are a consequence of the *Cramér-Rao bound* of classical statistics. The Cramér-Rao bound concerns the problem of sampling from a probability distribution, where the probability distribution is determined by some parameters. The goal is to estimate these parameters based on the results of sampling the distribution. The Cramér-Rao bound states that the variance of an unbiased estimator is asymptotically lower bounded by $O(1/N)$, and the coefficient of this scaling is given by the *Fisher information* (the definition may be found in Appendix A).

In a tomography experiment, with a fixed set of measurements, the outcome probability distribution is fixed by the parameters of the quantum state we are trying to estimate, and we can directly apply the Cramér-Rao bound. The procedure is as follows:

(i) Consider a tomography experiment on a quantum state with parameters η, and its fidelity to the state estimated from the tomographic reconstruction with parameters $\hat{\eta}$. In an asymptotic regime, our estimates will be close to the true state, so we can expand the fidelity in a Taylor series about the true state.

(ii) Calculating the expected value of the fidelity over the different measurement outcomes, we obtain an asymptotic expression for the point fidelity that is a function of the covariance matrix of the estimated parameters.

(iii) By applying the Cramér-Rao bound to the covariance matrix of the estimated parameters, we can obtain an upper bound on the point fidelity that is a function only of the number of each kind of measurement we have made, and the state parameters.

(iv) Finally, by averaging this point fidelity expression uniformly over the state parameters (using the Haar measure) we obtain the average fidelity bound.

Our expressions are analytical up to the final step, and then we use numerical integration to average over the state parameters using the Haar measure and obtain our final asymptotic results.

We find $F_{\mathrm{av}} \leq 1 - y/N$, where y is 1.083 for the cube (in agreement with the value of $13/12$ in [42]), 1.049 for the octahedron, 1.018 for the dodecahedron, and 1.008 for the icosahedron. It is well known in classical statistics that the Cramér-Rao bound can be achieved using a maximum likelihood estimator. Hence the above bounds are also tight. These results show that for pure states, the performance of the Platonic solids measurements in tomography is close to the collective measurement bound of $y=1$.

Recall that the quantum Chernoff bound is equal to the fidelity if one of the states involved is pure. Alternatively,

when, as here, both states are pure, the Hilbert-Schmidt distance is also an equivalent measure as $D_{\mathrm{HS}}^2 = (1-f)/2$, so that the choice of figure of merit does not affect the results of this section at all.

B. Mixed-state qubit tomography

When the state to be reconstructed can be mixed, an analytical treatment is more complicated since the requirement that the density matrix be positive semidefinite leads to constraints on the allowed set of parameters. The situation is also complicated by the way in which the fidelity function depends on the mixedness of the state to be reconstructed. We partially address both these issues in this section, with the main aim to develop some qualitative intuition. For quantitative results, we will resort to simulation.

For mixed-state qubit tomography, the state Bloch vector is no longer unit length. The state vector is now

$$\vec{r} = [r_x, r_y, r_z]^T, \qquad (13)$$

where $r_x = r \cos(\phi)\sin(\theta)$, $r_y = r \sin(\phi)\sin(\theta)$, $r_z = r \cos(\theta)$, and r satisfies $r \leq 1$. This constraint on the length of the Bloch vector reflects the requirement that the density matrix is positive semidefinite.

Mixed-state tomography introduces two new complications, which we discuss in this section. Both concern behavior close to the boundary of the Bloch sphere. The first is due solely to the definition of the mixed-state fidelity function,

$$f(\vec{r}, \hat{\vec{r}}) = \frac{1 + \vec{r} \cdot \hat{\vec{r}} + \sqrt{1-r^2}\sqrt{1-\hat{r}^2}}{2}, \qquad (14)$$

where $\hat{\vec{r}}$ is the estimated state vector. It can be understood by considering only errors in the Bloch vector length r [13]. Suppose $r = 1 - \delta$ and our estimate is $\hat{r} = r - \epsilon$. Then provided $\epsilon \ll \delta$ and both are small, we find $F_{\mathrm{pt}} \approx 1 - \epsilon^2/8$. From the statistical arguments of the previous section, we would expect ϵ to scale asymptotically like $1/\sqrt{N}$, so we expect to see all point fidelities (except for perfectly pure states) to scale like $F_{\mathrm{pt}} \approx 1 - 1/8N$ once N is large enough that our errors are well inside the Bloch ball. Alternatively, if $\delta \ll \epsilon$ and both are small, we find that the point fidelities scale like $F_{\mathrm{pt}} \approx 1 - \epsilon/2$. Thus the general behavior of point fidelities is a transition from $1/\sqrt{N}$ scaling at low N to a scaling of $1/N$ at large N. The transition takes place when the error associated with the state reconstruction ϵ becomes small enough that $\hat{r} + \epsilon$ lies inside the Bloch ball. Thus the location of the transition happens at higher N for higher-purity states. This is illustrated in Fig. 3, where numerical simulations of point fidelities for states with different radial parameters are shown. The Bures prior has a radial distribution that is singular at $r=1$, and it turns out that when point fidelities in Fig. 3 are averaged over this prior, the initial $1/\sqrt{N}$ dependence of the point fidelity persists long enough that the asymptotic dependence of the average fidelity on N is not N^{-1} but rather $N^{-3/4}$ [43].

The second complication arises since our estimated physical state has to be positive semidefinite. For qubits this is equivalent to $\hat{r} \leq 1$. An estimator that predicts states lying
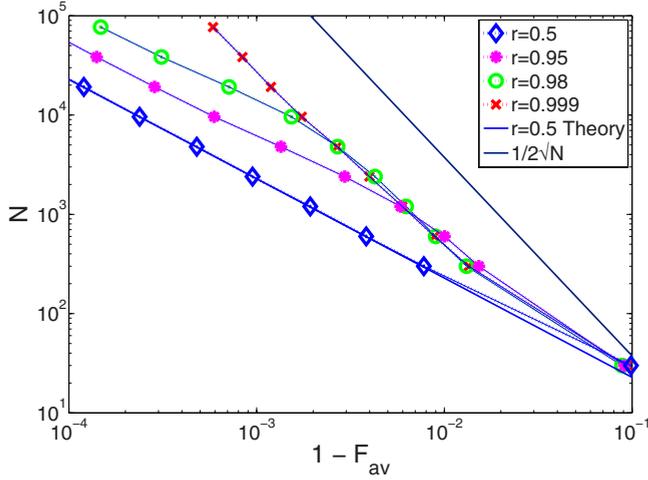
FIG. 3. (Color online) Numerical simulations of atomic point fidelities for states rotated 45 degrees about the x axis from the positive z axis. Each graph corresponds to a different radius on the Bloch sphere. The vertical axis is the number of copies of the state used in the tomography. The horizontal axis is one minus the average point fidelity. The cube measurement set was used for these simulations. Observe mixed states always have a $1/N$ scaling. States that are close to pure start (after we get to a sufficiently large N to obtain a sensible state estimate) with a $1/\sqrt{N}$ dependence and transition to $1/N$ scaling with larger N. This knee occurs at larger N for purer states. The statistical theory of Eq. (15) is plotted for $r = 0.5$.

outside the boundary of the Bloch ball can be improved by mapping these states onto the boundary of the Bloch ball, which always moves the estimate closer to the true state. Maximum likelihood estimation is a specific example of an estimator that does exactly this. This procedure creates a biased estimator that nevertheless performs better than any unbiased one. The naïve Cramér-Rao estimation is therefore not appropriate.

For states well away from the boundary, we can ignore this positivity constraint, and use the unbiased statistical methods of the previous section extended to three parameters to calculate the cube measurement point fidelity,

$$F_{\text{pt}}(\vec{r}, N) \leqslant 1 - \frac{3}{4N} \frac{3 - 3r^2 + 2(r_y^2 r_x^2 + r_z^2 r_x^2 + r_z^2 r_y^2)}{1 - r^2}. \quad (15)$$

The r_x, r_y, and r_z variables appear in the equation as an indication of the state orientation relative to the cubic measurement directions. This angular dependence of the point fidelity is more complicated for the other platonic solids. This theoretical curve is included in Fig. 3 for the case in which $r = 0.5$, and agrees well with the simulations (once N is large enough to obtain a sensible state estimate).

V. NUMERICAL SIMULATIONS FOR MIXED STATES

The main result of this paper is the numerical simulation of tomography using Platonic solid measurements. The simulations are performed in the following way.

(i) Choose N_s states ρ_i at random according to the Bures (or Chernoff) prior. For our simulations, N_s was chosen to be 10 000.

(ii) For each ρ_i, simulate measurements on N_c copies of the state (divided equally between the measurement sets) generating N_c measurement outcomes χ_i^j. For our simulations, we chose $N_c = 100\,000$.

(iii) Perform maximum-likelihood state estimation based on increasing numbers N of these $\chi_i^1 \cdots \chi_i^N$ obtaining state estimates $\hat{\rho}_i(N)$.

(iv) Between each state estimate and true state, calculate the fidelity $f(\hat{\rho}_i(N), \rho_i)$ via Eq. (1) or the quantum Chernoff bound $\lambda_{\text{cb}}(\hat{\rho}_i(N), \rho_i)$ via Eq. (2).

(v) Calculate the average fidelity,

$$F_{\text{av}}(N) = \sum_{i=1}^{N_s} f(\hat{\rho}_i(N), \rho_i)/N_s,$$

or the average Chernoff bound,

$$\lambda_{\text{av}}(N) = \sum_{i=1}^{N_s} \lambda_{\text{cb}}(\hat{\rho}_i(N), \rho_i)/N_s.$$

The Monte Carlo averaging over states and measurement outcomes is performed simultaneously. For the Chernoff bound simulations, the minimization in Eq. (2) was performed using the MATLAB function fminbnd.

We present our results on a log-log plot of number of copies required against $1 - F_{\text{av}}(N)$. To find the number of copies required to achieve a target average fidelity, one can select the target average fidelity on the horizontal axis, and tracing vertically upward, find the number of copies required to achieve that target average fidelity. The insets on the graphs are enlargements of the largest N section of the graph. The width of the lines on these insets is approximately equal to the one standard deviation statistical error in average fidelity due to our choice of N_s.

We simulate two main scenarios: the tomography of atomic qubits where the main source of error results from measuring a finite number of copies of the system, and the tomography of photonic qubits where the main statistical fluctuations arise from using sources that create photons at Poissonian-distributed times.

A. Atomic qubits

The kind of experiment we model here is an atomic qubit as in experiments on trapped ions, such as [5,6] represented by two metastable energy levels $|0\rangle$ and $|1\rangle$, and an auxiliary level $|r\rangle$ used for performing the measurement with a laser driving the transition $|0\rangle \rightarrow |r\rangle$ leading to fluorescence if the state was initially in $|0\rangle$ [44]. Arbitrary single-qubit projective measurements are made available by a first laser pulse tuned to the $|0\rangle \rightarrow |1\rangle$ transition followed by turning a laser on the $|0\rangle \rightarrow |r\rangle$ transition and looking for this fluorescence. We assume the dominant source of error is statistical fluctuation due to the finite number of copies of the state, and neglect all other sources of error. Many of these could be taken into account by replacing our projective measurements with suitable POVMs. However, the effects such as drifts in either
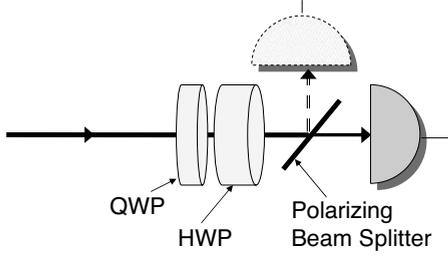
FIG. 4. The single photonic qubit experimental configuration consists of a quarter-wave plate and a half-wave plate to set the basis, and a polarizing beam splitter followed by a photon detector to analyze in that basis. We call this a single-detector configuration. An additional detector can be placed on the reflected port in a dual-detector configuration. This will produce the same count statistics as in the atomic case after waiting for a fixed number of counts for each basis setting.

state preparation or measurement settings would necessitate a different approach.

In this limit, these single-qubit system measurements are well described by Eqs. (10) and (11) and count for each outcome k of a measurement setting l drawn from a multinomial distribution,

$$p(n_{l1} \cdots n_{lK}) = \frac{N_l!}{n_{l1}! \cdots n_{lK}!} p_{l1}^{n_{l1}} p_{l2}^{n_{l2}} \cdots p_{lK}^{n_{lK}}, \qquad (16)$$

where N_l copies of the system are measured with measurement setting l, $p_{lk} = \mathrm{tr}(\rho O_{lk})$, and K is number of possible outcomes for the measurement (two for qubit systems and four for two-qubit systems).

For two-qubit systems, the measurement operators O_{lk} take the form of Eq. (12) and now the index l ranges over all combinations of single-qubit measurement settings on the subsystems. Experimentally this corresponds to noting the joint probability of the four possible outcomes of fluorescence and no fluorescence in each subsystem, for all combinations of measurement settings on the subsystems.

B. Photonic qubits

Our model for experiments on single photonic qubits is shown in Fig. 4. In this model, our qubits are represented by the polarization state of photons issuing from a source with Poissonian arrival times. There are two configurations to consider: a single-detector configuration and a dual-detector configuration.

In the first configuration, the source is directed through a quarter-wave plate, a half-wave plate and polarizing beam splitter (or vertical polarizer), and finally a single photon counter [15]. The angles of the fast axes of both of the wave plates can be rotated, allowing the projection onto vertical polarization fixed by the final polarizer to be rotated into any polarization state that the experimenter requires. The total time available to the experimentalist t is divided equally among the L measurement settings. Photon counts are accumulated over a time $t_l = t/L$ for each measurement setting m_l. The total photons incident on the measurement apparatus during this time N_l are drawn from the Poisson distribution,

$$p(N_l) = \frac{e^{-\mathcal{N}_f t_l}(\mathcal{N}_f t_l)^{N_l}}{N_l!}, \qquad (17)$$

where $\mathcal{N}_f$ is the mean photon flux. The number of these photons reaching the detector n_l is then drawn from a binomial distribution with probability $p_l = \mathrm{tr}(\rho O_l)$. The flux is estimated from a set of r measurements satisfying

$$\sum_{i=1}^{r} O_i = s I_d \qquad (18)$$

via

$$\mathcal{N}_f = \sum_{i=1}^{r} \frac{n_i}{s t_i}. \qquad (19)$$

For the Platonic solid measurements, the four measurements of the tetrahedron satisfy Eq. (18) with $s=2$, and for higher-order Platonic solids each axis has two measurements satisfying Eq. (18). In the latter case, the average is made over all the axes to obtain an improved flux estimate.

If a detector is added at the reflected port of the polarizing beam splitter, both orthogonal polarizations are measured simultaneously and a potentially better estimation of the total flux can be obtained by summing the two detector counts. Indeed, by waiting until a set number of total counts are obtained for each measurement, the two-detector configuration will produce exactly the same binomially distributed statistics as in the atomic case. When the detectors are perfectly efficient, this conclusion is fairly obvious. However, the same will even be true in a real experiment with inefficient detectors, because although the detector loss will affect the overall probability (and hence rate) of detecting a photon, it will not affect the conditional probabilities that determine the measurement outcomes given that a photon is detected. Mathematically, this can be seen by modeling the detector losses as simple beam splitters in front of the detectors. Provided the two detectors have the same efficiency, this is just the same as including a loss element in the source which will change its amplitude but not its statistics. We therefore do not simulate the two-detector configuration separately.

We also study two-photon state tomography. A source produces photon pairs whose arrivals are again Poisson-distributed according to Eq. (17). The single-photon one- and two-detector configurations are duplicated, requiring two and four detectors, respectively. Now only coincidences are counted between the two photonic qubits. The number of coincidences are drawn from the multinomial distribution of Eq. (16). For the single-detector configuration for each qubit, we estimate the flux via

$$\mathcal{N}_f = \sum_{j=1}^{r} \sum_{i=1}^{r} \frac{n_{ij}}{s t_{ij}}, \qquad (20)$$

where $\Sigma_{i=1}^{r} \Sigma_{j=1}^{r} O_i \otimes O_j = s I_d$.

C. Maximum likelihood state reconstruction

After simulating measurements on all copies of the state, we have n_{lk} detections corresponding to each measurement

operator O_{lk}. Maximum likelihood reconstruction is the problem of finding the state ρ that maximizes the likelihood $p(\cdots n_{lk}\cdots|\rho)$ or equivalently minimizes $-\ln p(\cdots n_{lk}\cdots|\rho)$. To make our inversions efficiently solvable, we convert (an approximate version of) this problem into a semidefinite program.

For atomic (single and multiple) qubit systems, and for each measurement setting l, the outcomes follow the multinomial distribution of Eq. (16). We will approximate these statistics with a Gaussian distribution with mean $\mu_{lk}=p_{lk}(\rho)N_l$ and a covariance matrix V_l with diagonal entries $(V_l)_{ii}=N_l p_{li}^e(1-p_{li}^e)$ and off-diagonal entries $(V_l)_{ij}=-N_l p_{li}^e p_{lj}^e$, where $p_{lk}^e=n_{lk}/N_l$. Since the outcomes of different measurement settings are statistically independent, the combined likelihood function is $p(\cdots n_{lk}\cdots|\rho)=\Pi_l p(n_{l1}\cdots n_{lK})$. Defining $\vec{n}_l$ as the vector with kth element n_{lk}, and $\vec{\mu}_l$ as the vector with kth element μ_{lk}, our approximate problem is to find a state ρ that minimizes

$$-\ln p(\cdots n_{lk}\cdots|\rho) \propto \sum_l (\vec{n}_l - \vec{\mu}_l)^T V_l^{-1}(\vec{n}_l - \vec{\mu}_l). \quad (21)$$

We have assumed that the term in the exponential of the Gaussian dominates the likelihood and used experimental probabilities p_{lk}^e instead of $p_{lk}(\rho)$ in the covariance matrices V_l to make the problem convex. This enables the problem to be converted into a semidefinite program and solved with the SEDUMI semidefinite program solver software [45].

For single-detector simulations of single-photon and two-photon experiments, the counts are Poisson-distributed, and as we only measure one outcome per measurement setting, all outcomes are statistically independent. We make a Gaussian approximation with the expected number of counts for measurement outcome O_l equal to the variance in those counts, i.e., $\mu_l=\sigma_l^2=\mathcal{N}_l=p_l\mathcal{N}_f t_l$, where $p_l=\text{tr}(\rho O_l)$. Our approximate problem is to find the state ρ that minimizes the log likelihood

$$-\ln[p(n_1, n_2, \ldots|\rho)] \propto \sum_l \frac{(n_l-\mathcal{N}_l)^2}{2\mathcal{N}_l} = \frac{\mathcal{N}_f t_l}{2}\sum_l \frac{[p_l^e - p_l(\rho)]^2}{p_l(\rho)},$$

where $p_l^e=n_l/(\mathcal{N}_f t_l)$, and we have assumed times spent on each measurement setting t_l are equal. In this case, we were able to use $\mathcal{N}_l$ for mean and variance without losing the convexity of the problem.

VI. NUMERICAL RESULTS

A. Atomic qubit results based on fidelity

The results for our numerical simulations of single-qubit atomic tomography are shown in Fig. 5(a). As noted above, atomic experiments measure both orthogonal outcomes simultaneously, so there are no results for the tetrahedron. The key point from the graph is that, for sufficiently large N, the higher-order Platonic solids provide improved tomography performance. Moreover, because N is the *total* number of system copies measured, this result holds despite the fact that each outcome probability is estimated from a smaller number of measurements for the higher-order solids. The improve-

ment of the icosahedron over the dodecahedron is only marginal, but the differences between the other cases are statistically significant.

A slightly different trend, however, is displayed by the results for the two-qubit atomic tomography, which are shown in Fig. 5(b). In this case, the performance is effectively the same for all the Platonic solids, and experimentally one should choose the cube tensor measurement to minimize the time consumed by changing the measurement settings.

B. Photonic qubit results based on fidelity

The results for a single photonic qubit using the single-detector configuration are shown in Fig. 5(c). In addition to the Platonic solid measurement sets, we have also included the popular minimal measurement sets proposed by James *et al.* in [15] (labeled James4), and the most striking result is how poorly it performs in comparison to the Platonic solids. The significant improvement that can be achieved by using a tetrahedral measurement set (which is also a minimal set) demonstrates the value of using isotropically distributed measurements. However, the results also show that moving to overcomplete measurements sets provides still more improvement, and the reconstruction quality improves with the number of measurements. Interestingly, we note that similar performance is achieved by the dual Platonic solid pairs (the cube and octahedron; the dodecahedron and icosahedron). The clear conclusions from these results are that experimentalists should switch from the James4 set to Platonic solids and that the cube and dodecahedron measurement sets provide the best value improvement in performance for the number of measurement settings.

The photonic two-qubit simulation results for the single-detector configuration are shown in Fig. 5(d). In addition to the Platonic solid measurement sets, we have also included the popular measurement sets proposed by James *et al.* in [15] (labeled James16) and the projective SicPOVM measurements.

Again the most striking result is how poorly the James16 minimal measurement set performs compared to tensor products of any Platonic solid, even the tensor-product tetrahedral set, which has the same (minimal) number of measurements. Furthermore, the tensor-product cube set outperforms the tetrahedral set, again demonstrating the advantage of using overcomplete measurement sets over minimal sets. However, for two photonic qubits, as in the atomic case, moving to more overcomplete measurement sets seems to produce negligible improvement. The projective SicPOVM measurement set discussed in Sec. III B contains entangled measurements that are isotropically distributed over the space of two-qubit pure states. With the same number of measurements, we see that the SicPOVM set performs significantly better than the tetrahedral set. Slightly surprising, however, is that it only shows a negligible improvement over the other platonic solid measurement sets.

Once again, the clear conclusion from these results is that experimentalists should switch immediately from the James16 settings to a tensor-product cube measurement set, or at least, if changing measurement settings consumes too
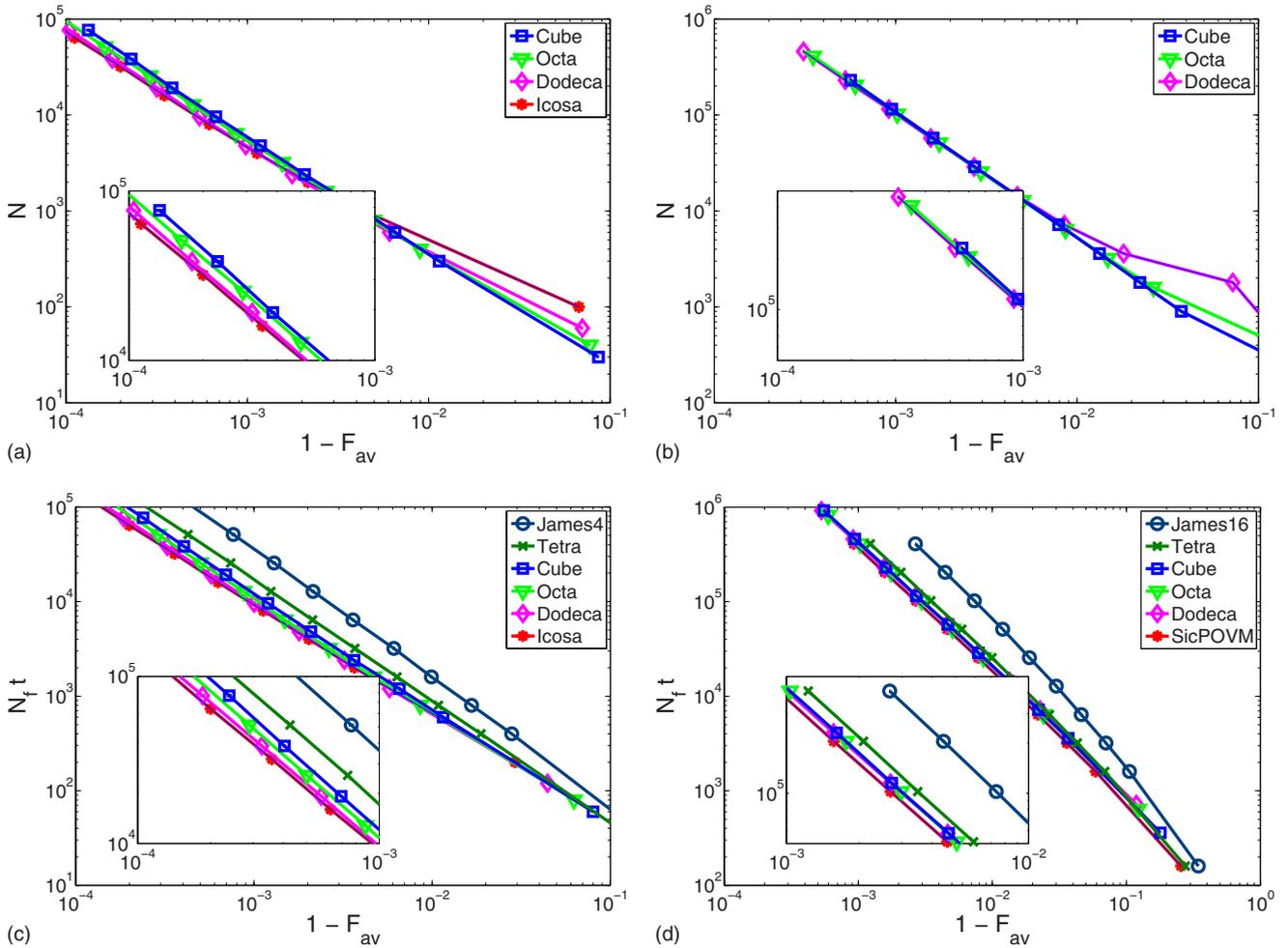
FIG. 5. (Color online) Tomography simulation results with a Bures prior. In (a) and (b), the number of copies of the state are plotted against $1-F_{av}$, where F_{av} is the average fidelity. To read, select the desired average fidelity on the horizontal axis, and trace vertically to find the number of copies of the state required to achieve that fidelity. Similarly, (c) and (d) plot the photon flux N_f multiplied by the total time t against $1-F_{av}$. N.B. the total time t is equally divided between the different measurement settings. Each plot shows the tomography performance of various Platonic solids. The individual figures are tomography for (a) single atomic qubit states (the tetrahedron cannot be measured in this configuration), (b) two atomic qubit states, (c) single photonic qubit with single-detector configuration, and (d) two photonic qubit with single-detector configuration.

much time, to a tensor-product tetrahedral set. Also, we note that there would be little advantage gained in trying to implement the much more difficult SicPOVM measurements, when virtually the same reconstruction quality can be achieved by using an overcomplete set of separable measurements.

C. Results based on the quantum Chernoff bound

As discussed in Sec. II, the average quantum Chernoff bound has a well-motivated physical interpretation that is particularly appealing in the context of tomography. It is thus important to see if our conclusions hold when the quantum Chernoff bound is used instead of the more easily calculated fidelity.

We repeated our tomography simulations using the average quantum Chernoff bound as the figure of merit, and the Chernoff bound prior over states. The results for one and two atomic qubits are shown in Figs. 6(a) and 6(b), respectively.

The results for one and two photonic qubits with single-detector configurations are shown in Figs. 6(c) and 6(d), respectively.

In all cases, the graphs are almost identical to those from Fig. 5. Therefore, we see that all of the conclusions that we drew from the fidelity-based results can also be inferred from the results based on the quantum Chernoff bound, and this in turn provides some extra validation for the fidelity-based results and even for the average fidelity itself in assessing tomographic performance.

D. Comparison with analytical results based on the mean-squared Hilbert-Schmidt distance

While we have focused on the average fidelity and the average quantum Chernoff bound as our figures of merit for characterizing tomography schemes, another commonly used figure of merit, the mean-squared Hilbert-Schmidt distance, produces quite different results.
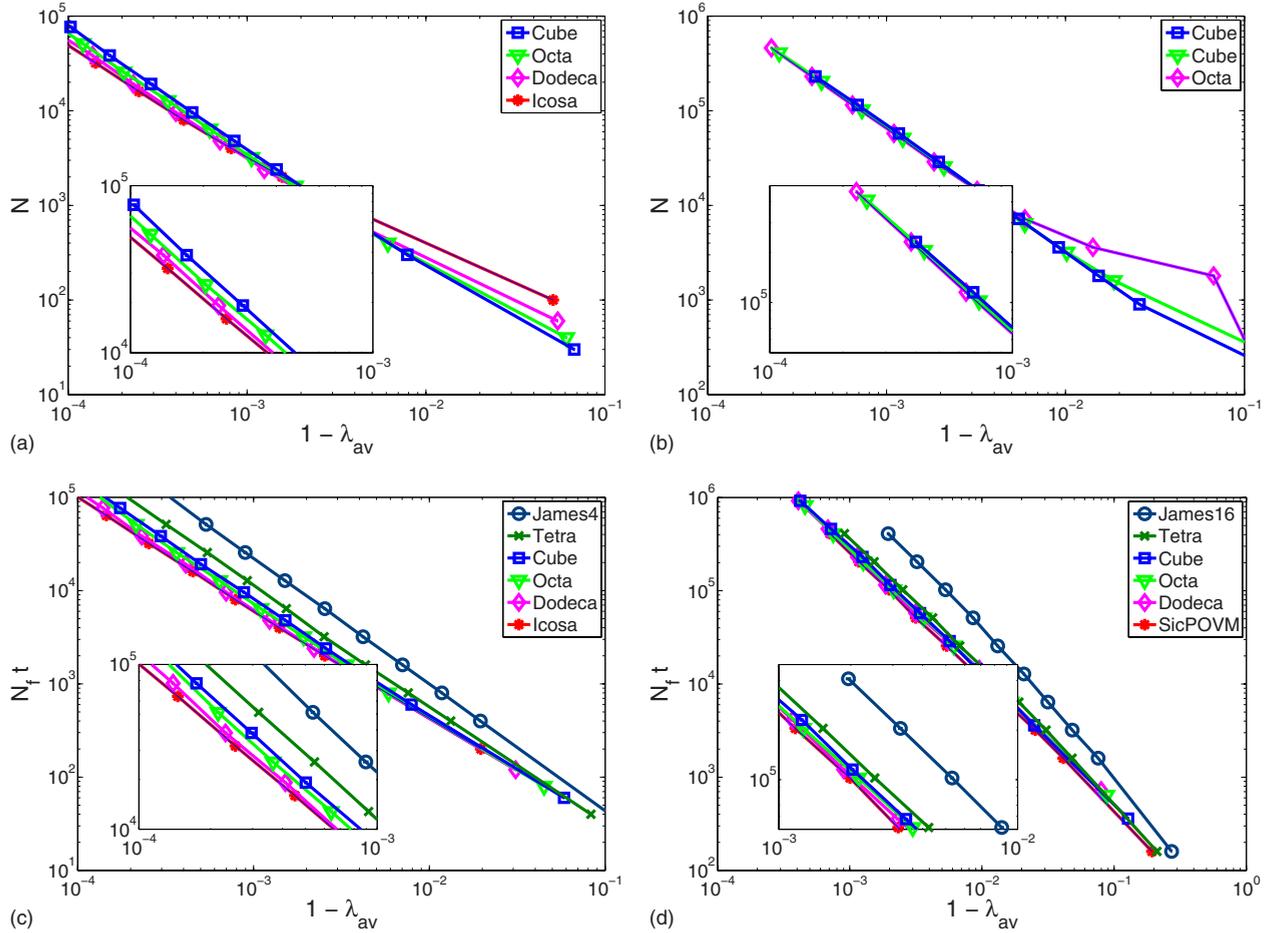
FIG. 6. (Color online) Tomography simulation results with the prior induced by the quantum Chernoff bound. In (a) and (b), the number of copies of the quantum state N is plotted against $1-\lambda_{av}$, where λ_{av} is the average quantum Chernoff bound. In (c) and (d), the photon flux N_f multiplied by the total time t is plotted against $1-\lambda_{av}$. The individual figures are tomography for (a) a single atomic qubit, (b) two atomic qubits, (c) a single photonic qubit with a single-detector configuration, and (d) two photonic qubits using single-detector configurations.

In a recent paper, Roy and Scott [21] found that a generalized class of mutually unbiased bases (e.g., the cube measurement for a single qubit in our terminology) were an optimal measurement set for single qubit tomography. They chose to use the mean-squared Hilbert-Schmidt distance as a figure of merit specifically because it allowed analytical solutions to the complicated problem of characterizing mixed-state tomography. The somewhat surprising feature of their result is that it also holds for a generalized class of mutually unbiased bases, which includes all of the Platonic solid measurement sets, implying that the different Platonic sets should all perform equally well. This contrasts with the results we have obtained using the average fidelity and average quantum Chernoff bound.

To explore this discrepancy, we repeated the calculation of Eq. (15) (for the cube measurement set), replacing the fidelity with squared Hilbert-Schmidt distance to give a point squared Hilbert-Schmidt (HS) distance of

$$\langle D_{HS}^2(r) \rangle \geqslant \frac{3}{2N}(3-r^2), \qquad (22)$$

where r is the radius of the state on the Bloch sphere. This equation was derived in [20] by other means and agrees with

the results of [21]. Note that the point squared Hilbert-Schmidt distance is independent of the orientation of the true state relative to the measurement directions on the Bloch sphere, unlike the point fidelity, which depends explicitly on the true state coordinates r_x, r_y, and r_z [see Eq. (15)]. In other words, if the results of Fig. 1 for the cube measurement set were recalculated using the point squared Hilbert-Schmidt distance, the plot would show no variation other than some statistical fluctuations. This calculation confirms that the choice of figure of merit qualitatively affects the apparent performance of different tomography schemes.

Both the above calculation and the results from [21] neglect the effects of positivity constraints on the reconstructed density matrix. We have therefore also repeated our numerical simulations for atomic qubits using mean-squared Hilbert-Schmidt distance as a figure of merit, a Bures prior, and full maximum likelihood estimation. Our results show that the Platonic solids all performed equivalently. Indeed, our results were within statistical errors of the theoretical value obtained by integrating Eq. (22) over a Bures prior $D_B \geqslant 0.8438/N$.

The above results confirm the conclusion of [21] that all of the Platonic solids appear to be optimal (and equivalent) when the mean-squared Hilbert-Schmidt distance is used as a

figure of merit even when the physical effects of positivity are included. This is qualitatively different from the conclusions obtained by using the physically well-motivated average quantum Chernoff bound.

VII. DISCUSSION AND CONCLUSIONS

A. Experimental recommendations

The most significant implications of the work in this paper apply to tomography of photonic systems performed using single-detector configurations. We have conclusively shown that experimentalists should immediately switch from the minimal four-measurement set of [15] to the overcomplete six-measurement cube set, or at the very least switch to the tetrahedron measurement set if the number of measurement settings is an issue. In particular, we showed that for both one- and two-qubit systems, the minimal sets of [15] are far outperformed by all of the Platonic solid measurement sets. This strongly suggests that the use of these symmetrically arranged measurements will be advantageous for experiments on systems having larger numbers of qubits. Here, however, we note that the number of measurement settings grows very rapidly and the time to change between detector settings is likely to become an issue. If this is the case, then the multiqubit tetrahedral sets should also provide much better performance than the multiqubit sets outlined in [15] without requiring any more measurements.

In both one- and two-qubit cases, the overcomplete Platonic sets (cube and above) also showed further improvement over the minimal tetrahedral measurement set. Indeed for two qubits, these overcomplete tensor-product Platonic sets performed comparably to entangled measurement sets, although only a negligible improvement was achieved by going higher than the six-measurement cube set. For single-qubit tomography, experimentalists would have to judge whether the continued improvement in asymptotic performance for the higher-order Platonic solids would result in a net gain after the time required to switch between measurement settings is taken into account.

As noted earlier, when using a dual-detector configuration, the conditional statistics of the counts across the possible measurement outcomes follow binomial (or multinomial) distributions as with atomic qubits. In all cases we investigated, the atomic tomography converges slightly faster than the single-detector photonic tomography. Also, moving to dual detectors allows some purely practical benefits by allowing some compensation for, or at least the identification of, stability problems in an experimental situation. We note, however, that the standard maximum-likelihood penalty function introduced in [15] is only valid for the single-detector configuration and should be replaced by the multinomial penalty function of Eq. (21), although in practice the difference may often be negligible.

For tomography in atomic systems (such as measurements on trapped ions), experimentalists are generally already performing the cube measurement set. For single-qubit tomography, our results indicate that further improvement could be obtained by moving to higher-order Platonic solids. However, for two-qubit systems, as in the photonic case, the higher-order solids demonstrate only a negligible improvement over the cube set. This suggests that, in atomic systems, experimentalists are already generally using the optimal tomographic measurement schemes.

B. Tomography and the spherical *t*-designs

Mathematically, the idea of isotropically spacing points on a sphere is also captured by the concept of a *spherical t-design*. A description of *t*-designs is beyond the scope of this paper, but note that *t*-designs are sets of points that allow numerical estimation of integrals over the sphere, with the quality of the approximation improving as *t* is increased. This suggests that good measurement sets for tomography may well be associated with *t*-designs. Indeed it has been shown [46] that *t*-designs are optimal for pure state estimation when collective measurements are available.

Popular measurement sets for tomography of general quantum systems include the so-called symmetric informationally complete POVMs (SicPOVMs) [41], which are spherical 2-designs. In fact, the tetrahedron measurement set is the SicPOVM for a single qubit and the other Platonic solids are also 2-designs. Moreover, it is known [47] that the cube and octahedron are 3-designs and the dodecahedron and icosahedron are 5-designs (a *t*-design is also a *t'*-design for all $t' < t$). We suggest that the increasing *t* of the *t*-design for the higher-order Platonic solids may be related to their better performance in the tomographic reconstruction of single-qubit states. In particular, we observe that the dual Platonic solids, which have the same *t*-design, also perform similarly well in the tomography of single qubits.

We believe that the connection between *t*-designs and the performance of different measurement sets in mixed-state tomography deserves to be explored in more detail in future theoretical work. From a more practical perspective, it may also be worth investigating a tomographic measurement set based on McLaren's "improved Snub Cube" [47], a 24-point 7-design, which therefore provides a fairly substantial increase in the value of *t* for a relatively small increase in the number of measurements.

C. Future work

The results in this paper provide definite, useful recommendations for how an experimentalist can improve their tomographic reconstructions at a practical level, but there are still many remaining avenues for extending this work. For example, one could investigate measurement sets made from polyhedra other than the Platonic solids (like the semiregular polyhedra). However, this paper has shown that the Platonic solids provide a good spread in the number of measurements and in performance, and going to even higher-order shapes appears to be of decreasing value. We thus consider exploring other single-qubit shapes to be unfruitful, with the one possible exception of the improved Snub Cube design mentioned above.

We have focused on qubit tomography in this paper, but it would also be interesting to broaden our investigation to include tomography of arbitrary *d*-dimensional systems, since these systems exhibit some advantages for various quantum

computing and quantum communication protocols. Moreover, for certain dimensions, this would also provide some insight into tomographic performance in multiqubit systems in situations where measuring entangled projections are sufficiently feasible experimentally. In higher dimensions, it would be interesting to investigate the generalizations of the platonic solids, the simplex, the hypercube, and the cross-polytope. These are defined in arbitrary dimensions and may straightforwardly be embedded into the Bloch-hypersphere (indeed, the simplex is just the well-known SicPOVM [41]). However, since not all the points on the hypersphere correspond to pure states, finding a simplex, hypercube, or cross-polytope of pure states is nontrivial. In [41], this was done by numerical search for the simplex. The results of this paper suggest that overcomplete measurement sets based on the hypercube and cross-polytope would outperform the SicPOVM for state tomography.

Taking a different approach, we also suggest that tomographic performance should be investigated in the context of real experiments. While experiments in tomography usually report an uncertainty in the reconstruction based on a fixed number of measured systems, it would be interesting to perform an experimental study of how this uncertainty decreases over the time of the tomography. In particular, such a study should provide information about the noise sources affecting the experiment. For example, drifts in either the states or the measurement over time might be expected to lead to a saturation of the average fidelity with N. We have seen that when the experiment is limited by statistical noise, then the quality of the tomography depends systematically on N and on the choice of measurement. This could be tested experimentally and should give detailed information about the true errors in real experiments. In the case of optical experiments, this simply requires rotating regularly among the different measurement settings and keeping track of the order in which counts arrive. State inversions can then be done at various time intervals, and a graph of the number of qubits N against point fidelity similar to our Fig. 3 could be straightforwardly obtained. One should, for example, be able to observe the general behavior of a scaling like $1/\sqrt{N}$ at low numbers of copies transitioning to a scaling of $1/N$ at large numbers of copies. Likewise the improvement of the point fidelity with the choice of measurement could be checked. Comparison to detailed point fidelity simulations should aid in the characterization of experimental error sources. To date, there has been almost no experimental work of this type, although there are very recent results in [48] exploring different measurement sets for two-qubit tomography.

While our results have been stated purely in terms of quantum state tomography, they also have some direct implications for quantum process tomography. In process tomography, probe states are prepared and allowed to evolve under the process to be characterized. The process is then inferred from measurements of the resulting states. There are several variations in how one can obtain the process description, including standard quantum process tomography (SQPT) [49], ancilla-assisted quantum process tomography (AAPT) [50,51], and direct characterization of quantum dynamics (DCQD) [35]. AAPT, where the process is applied to one part of a *single* input state, has the most direct link to our

results as they directly apply to the quality of the state tomography of the single output state. Our results suggest that using overcomplete measurement sets would improve the reconstruction accuracy.

In SQPT, a set of input states are used to probe the quantum process, and the process is obtained from state tomography of the output states. There is a direct connection between AAPT and SQPT if we restrict the measurements to tensor products of system measurements with ancilla measurements. Then, one can interpret the measurements on the ancilla as a remote preparation of the input state to the system, and we can expect the same performance between these tensor product measurements in AAPT and the corresponding input states and measurements in SQPT. In either case, however, the situation for process tomography is complicated by the requirements of complete positivity, as opposed to just positivity for states, and our observations only provide indications of the expected behavior. It would therefore be of great interest to comprehensively compare the various quantum process tomography strategies under a framework similar to the one we have presented, to determine the optimal strategies to adopt.

APPENDIX A: DERIVATION OF PURE-STATE CRAMÉR-RAO BOUNDS FOR SINGLE-QUBIT TOMOGRAPHY

Let $\vec{\eta}=[\theta,\phi]^T$ be the state Bloch vector in spherical polar coordinates, and let $\hat{\vec{\eta}}$ be the state estimate. The similarity between $\vec{\eta}$ and $\hat{\vec{\eta}}$ is given by the fidelity $f_{\vec{\eta}}(\hat{\vec{\eta}})=(1+\vec{r}\cdot\hat{r})/2$, where $\vec{r}=[\cos(\phi)\sin(\theta),\sin(\phi)\sin(\theta),\cos(\theta)]^T$ and similarly $\hat{r}=[\cos(\hat{\phi})\sin(\hat{\theta}),\sin(\hat{\phi})\sin(\hat{\theta}),\cos(\hat{\theta})]^T$.

Since we assume $\vec{\eta}$ and $\hat{\vec{\eta}}$ are close, we expand the fidelity in a Taylor series to second order, giving

$$f_{\vec{\eta}}(\hat{\vec{\eta}}) \approx f_{\vec{\eta}}(\vec{\eta}) + Df_{\vec{\eta}}(\vec{\eta})\cdot(\hat{\vec{\eta}}-\vec{\eta}) + \frac{1}{2}(\hat{\vec{\eta}}-\vec{\eta})H(\vec{\eta})(\hat{\vec{\eta}}-\vec{\eta})^T,$$

(A1)

where $Df_{\vec{\eta}}(\vec{\eta})_i = \frac{\partial f_{\vec{\eta}}(\hat{\vec{\eta}})}{\partial\hat{\eta}_i}\big|_{\hat{\vec{\eta}}=\vec{\eta}}$ and $H(\vec{\eta})$ is the Hessian matrix defined by

$$H(\vec{\eta})_{i,j} = \frac{\partial^2 f_{\vec{\eta}}(\hat{\vec{\eta}})}{\partial\hat{\eta}_i\partial\hat{\eta}_j}\bigg|_{\hat{\vec{\eta}}=\vec{\eta}}.$$

(A2)

Now $f_{\vec{\eta}}(\vec{\eta})=1$ and $Df_{\vec{\eta}}(\vec{\eta})=\vec{0}$, so

$$f_{\vec{\eta}}(\hat{\vec{\eta}}) \approx 1 + \frac{1}{2}\sum_{i,j}\frac{\partial^2 f_{\vec{\eta}}(\hat{\vec{\eta}})}{\partial\hat{\eta}_i\partial\hat{\eta}_j}\bigg|_{\hat{\vec{\eta}}=\vec{\eta}}(\hat{\eta}_i-\eta_i)(\hat{\eta}_j-\eta_j).$$

(A3)

During tomography, each of N identical copies of the qubit state is each measured exactly once. Our measurement set consists of l different measurement settings. So $\mathcal{N}=N/l$ mea-

surements are performed with each measurement setting.

Each measurement setting is represented by Bloch vector $\vec{m}_i$, with outcomes $\chi_i = \pm 1$. One set of measurements has 2^l possible outcomes represented by the string $\chi = \chi_1 \cdots \chi_l$. The probability of obtaining outcome χ is given by

$$p_{\vec{\eta}}(\chi) = \prod_i (1 + \chi_i \vec{r} \cdot \vec{m}_i)/2. \tag{A4}$$

The point fidelity obtained by averaging over the possible outcomes χ of the set of measurements is given by

$$F_{pt}(\vec{\eta}) = \sum_{\chi} p_{\vec{\eta}}(\chi) f_{\vec{\eta}}(\hat{\vec{\eta}}(\chi)) \tag{A5}$$

$$\approx 1 + \frac{1}{2} \text{tr}[H(\vec{\eta})V(\vec{\eta})], \tag{A6}$$

where $V(\vec{\eta})$ is the covariance matrix defined by

$$V(\vec{\eta}) = \sum_{\chi} p(\chi)[\hat{\vec{\eta}}(\chi) - \vec{\eta}][\hat{\vec{\eta}}(\chi) - \vec{\eta}]^T. \tag{A7}$$

The Cramér-Rao bound for unbiased estimators states

$$V(\vec{\eta}) \geq \frac{1}{\mathcal{N}I(\vec{\eta})}, \tag{A8}$$

where $I(\vec{\eta})$ is the Fisher Information matrix defined by

$$I_{ij}(\vec{\eta}) = \sum_{\chi} p_{\vec{\eta}}(\chi) \frac{\partial \ln p_{\vec{\eta}}(\chi)}{\partial \eta_i} \frac{\partial \ln p_{\vec{\eta}}(\chi)}{\partial \eta_j}. \tag{A9}$$

Noting that $H(\vec{\eta})$ is negative definite, the point fidelity bound is then given by

$$F_{pt}(\vec{\eta}) \leq 1 + \frac{1}{2\mathcal{N}} \text{tr}[H(\vec{\eta})\Gamma^{-1}(\vec{\eta})]. \tag{A10}$$

Finally averaging over $\vec{\eta}$ uniformly according to the Haar measure,

$$d\rho = \frac{\sin \theta}{4\pi} d\theta d\phi, \tag{A11}$$

gives the average fidelity Cramér-Rao bound.

APPENDIX B: EIGENVALUE DISTRIBUTION FOR THE QUANTUM CHERNOFF BOUND METRIC

In [28] it was shown that the infinitesimal distance between states ρ and $\rho + d\rho$ according to the quantum Chernoff bound is

$$ds^2 = \frac{1}{2} \sum_{jk} \frac{|\langle j|d\rho|k\rangle|^2}{(\sqrt{\lambda_j} + \sqrt{\lambda_k})^2}, \tag{B1}$$

where $\rho = \Sigma_j \lambda_j |j\rangle\langle j|$ is the eigenvalue decomposition of ρ. We will follow the same procedure as [33] in his calculation of

the Bures volume element to derive our quantum Chernoff bound eigenvalue distribution.

First decompose this into an infinitesimal shift in eigenvalues followed by an infinitesimal unitary rotation,

$$\rho + \delta\rho = (I + \delta U)(\rho + \delta\Lambda)(I + \delta U)^{\dagger} \tag{B2}$$

$$= \rho + \delta\Lambda + [\delta U, \rho], \tag{B3}$$

where we have ignored any terms that are a product of multiple infinitesimal quantities and defined $\langle j|\delta\Lambda|k\rangle = \delta_{jk} d\lambda_j$. In general, we can rewrite the infinitesimal unitary operator as

$$\delta U = \sum_{j \leq k} (dx_{jk} + idy_{jk})|j\rangle\langle k| + \text{H.c.} \tag{B4}$$

Next we substitute Eqs. (B4) and (B3) into Eq. (B1), yielding

$$ds^2 = \frac{1}{8} \sum_j \frac{(d\lambda_j)^2}{\lambda_j} + \sum_{j<k} \frac{(\lambda_k - \lambda_j)^2}{(\sqrt{\lambda_j} + \sqrt{\lambda_k})^2}(dx_{jk}^2 + dy_{jk}^2). \tag{B5}$$

We convert this into a differential volume element by multiplying together all the ds_i formed by a shift in each parameter λ_i assuming the shift in the other parameters is zero. Since the volume element will require normalization, we combine any constant factors. The final differential volume element is given by

$$dV = C' \frac{d\lambda_1 \cdots d\lambda_M}{\sqrt{\lambda_1 \cdots \lambda_M}} \prod_{k<j} \frac{(\lambda_k - \lambda_j)^2}{(\sqrt{\lambda_j} + \sqrt{\lambda_k})^2} dx_{jk} dy_{jk}, \tag{B5}$$

where C' is a normalization constant.

Since the eigenvalue distribution is unitarily invariant, we can concern ourselves with the marginal probability distribution over the eigenvalues of the density operators. That is,

$$p(\lambda_1 \cdots \lambda_M) = \frac{C}{\sqrt{\lambda_1 \cdots \lambda_M}} \prod_{k<j} \frac{(\lambda_k - \lambda_j)^2}{(\sqrt{\lambda_j} + \sqrt{\lambda_k})^2}, \tag{B6}$$

where the eigenvalues are also constrained to satisfy $\Sigma_{j=1}^M \lambda_j = 1$ and C is a normalization constant. Next we examine the single-qubit case.

Single-qubit case

For the qubit case, we have

$$p(\lambda_1 \lambda_2) = \frac{C}{\sqrt{\lambda_1 \lambda_2}} \frac{(\lambda_1 - \lambda_2)^2}{(\sqrt{\lambda_1} + \sqrt{\lambda_2})^2}. \tag{B7}$$

The eigenvalues for the qubit case are related to r (the radius on the Bloch sphere) according to $\lambda_1 = (1+r)/2$ and $\lambda_2 = (1-r)/2$. This yields

$$p(r) = \frac{C}{2} \frac{1 - \sqrt{1 - r^2}}{\sqrt{1 - r^2}}. \tag{B8}$$

By requiring $\int_0^1 p(r) dr = 1$, we calculate $C = 4/(\pi - 2)$.

[1] K. Vogel and H. Risken, Phys. Rev. A **40**, 2847 (1989).

[2] U. Leonhardt, Phys. Rev. Lett. **74**, 4101 (1995).

[3] A. G. White, D. F. V. James, P. H. Eberhard, and P. G. Kwiat, Phys. Rev. Lett. **83**, 3103 (1999).

[4] N. K. Langford, R. B. Dalton, M. D. Harvey, J. L. O'Brien, G. J. Pryde, A. Gilchrist, S. D. Bartlett, and A. G. White, Phys. Rev. Lett. **93**, 053601 (2004).

[5] H. Häeffner, W. Hänsel, C. F. Roos, J. Benhelm, D. C. al kar, M. Chwalla, T. Köerber, U. D. Rapol, M. Riebe, P. O. Schmidt, C. Becher, O. Gühne, W. Dür, and R. Blatt, Nature **438**, 643 (2005).

[6] R. Reichle, D. Leibfried, E. Knill, J. Britton, R. B. Blakestad, J. D. Jost, C. Langer, R. Ozeri, S. Seidelin, and D. J. Wineland, Nature **443**, 838 (2006).

[7] A. S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory* (North-Holland, Amsterdam, 1982).

[8] S. Massar and S. Popescu, Phys. Rev. Lett. **74**, 1259 (1995).

[9] R. Derka, V. Buzek, and A. K. Ekert, Phys. Rev. Lett. **80**, 1571 (1998).

[10] G. Vidal, J. I. Latorre, P. Pascual, and R. Tarrach, Phys. Rev. A **60**, 126 (1999).

[11] J. I. Latorre, P. Pascual, and R. Tarrach, Phys. Rev. Lett. **81**, 1351 (1998).

[12] M. Hayashi and K. Matsumoto, e-print arXiv:quant-ph/0411073.

[13] E. Bagan, M. A. Ballester, R. D. Gill, R. Munoz-Tapia, and O. Romero-Isart, Phys. Rev. Lett. **97**, 130501 (2006).

[14] M. Guţă and J. Kahn, Phys. Rev. A **73**, 052108 (2006).

[15] D. F. V. James, P. G. Kwiat, W. J. Munro, and A. G. White, Phys. Rev. A **64**, 052312 (2001).

[16] *Quantum State Estimation*, edited by M. Paris and J. Rehacek (Springer, Berlin, 2004).

[17] K. R. W. Jones, Ann. Phys. (N.Y.) **207**, 140 (1991).

[18] E. Bagan, A. Monras, and R. Munoz-Tapia, Phys. Rev. A **71**, 062318 (2005).

[19] R. D. Gill and S. Massar, Phys. Rev. A **61**, 042312 (2000).

[20] J. Rehacek, B.-G. Englert, and D. Kaszlikowski, Phys. Rev. A **70**, 052321 (2004).

[21] A. Roy and A. J. Scott, J. Math. Phys. **48**, 072110 (2007).

[22] A. Fujiwara and H. Nagaoka, Phys. Lett. A **201**, 119 (1995).

[23] M. Hayashi, in *Quantum Communication, Computing and Measurement*, edited by O. Hirota, A. S. Holevo, and C. A. Caves (Plenum, New York, 1997), pp. 99–108.

[24] A. Fujiwara and H. Nagaoka, J. Math. Phys. **40**, 4227 (1999).

[25] K. Matsumoto, J. Phys. A **35**, 3111 (2002).

[26] R. Prevedel, P. Walther, F. Tiefenbacher, P. Böhi, R. Kaltenbaek, T. Jennewein, and A. Zeilinger, Nature **445**, 65 (2007).

[27] C. A. Fuchs, e-print arXiv:quant-ph/9601020.

[28] K. M. R. Audenaert, J. Calsamiglia, R. Munoz-Tapia, E. Bagan, L. Masanes, A. Acin, and F. Verstraete, Phys. Rev. Lett. **98**, 160501 (2007).

[29] C. M. Caves, C. A. Fuchs, and R. Schack, J. Math. Phys. **43**, 4537 (2002).

[30] R. Blume-Kohout and P. Hayden, e-print arXiv:quant-ph/0603116.

[31] A. J. Scott, J. Phys. A **39**, 13507 (2006).

[32] K. Zyczkowski and H.-J. Sommers, J. Phys. A **34**, 7111 (2001).

[33] M. J. W. Hall, Phys. Lett. A **242**, 123 (1998).

[34] R. Munoz-Tapia (private communication).

[35] M. Mohseni, A. T. Rezakhani, and D. A. Lidar, Phys. Rev. A **77**, 032322 (2008).

[36] M. Hayashi, J. Phys. A **31**, 4633 (1998).

[37] T. Decker, D. Janzing, and T. Beth, Int. J. Quantum Inf. **2**, 353 (2004).

[38] B.-G. Englert, T. K. Ming, G. C. Guan, and N. H. Khoon, Laser Phys. **15**, 7 (2005).

[39] A. Ling, K. P. Soh, A. Lamas-Linares, and C. Kurtsiefer, Phys. Rev. A **74**, 022309 (2006).

[40] A. Ling, A. Lamas-Linares, and C. Kurtsiefer, e-print arXiv:0807.0991.

[41] J. M. Renes, R. Blume-Kohout, A. J. Scott, and C. M. Caves, J. Math. Phys. **45**, 2171 (2004).

[42] E. Bagan, M. Baig, and R. Munoz-Tapia, Phys. Rev. Lett. **89**, 277904 (2002).

[43] E. Bagan, M. Baig, R. Munoz-Tapia, and A. Rodriguez, Phys. Rev. A **69**, 010304(R) (2004).

[44] J. I. Cirac, L. M. Duan, and P. Zoller, Proceedings of the International School of Physics "Enrico Fermi," Course CXLVIII (OS Press, Amsterdam, 2002), p. 263.

[45] J. F. Sturn, SEDUMI: A MATLAB Toolbox for Optimization over Symmetric Cones.

[46] A. Hayashi, T. Hashimoto, and M. Horibe, Phys. Rev. A **72**, 032325 (2005).

[47] R. H. Hardin and N. J. A. Sloane, Discrete Comput. Geom. **15**, 429 (1996).

[48] R. B. A. Adamson and A. M. Steinberg, e-print arXiv:0808.0944.

[49] I. L. Chuang and M. A. Nielsen, J. Mod. Opt. **44**, 2455 (1997).

[50] G. M. D'Ariano and P. Lo Presti, Phys. Rev. Lett. **86**, 4195 (2001).

[51] 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).

[52] The reason for measuring L instead of R seems purely a historical accident. A measurement set consisting of the tensor product of two James4 sets would also suffice. Nevertheless, we have maintained the L in the James16 measurement set for the simulations in this paper.