Optimization of quantum state tomography in the presence of experimental constraints

Mohammadreza Mohammadi\textsuperscript{1} and Agata M. Brańczyk\textsuperscript{1,2,3,*}

\textsuperscript{1}Department of Physics, University of Toronto, 60 Saint George St., Toronto, Ontario M5S 1A7, Canada
\textsuperscript{2}Department of Chemistry, University of Toronto, 80 Saint George St., Toronto, Ontario M5S 3H6, Canada
\textsuperscript{3}Perimeter Institute for Theoretical Physics, Waterloo, Ontario N2L 2Y5, Canada

(Dated: January 22, 2014)

In the absence of experimental constraints, optimal measurement schemes for quantum state tomography are well understood. We consider the scenario where the experimenter doesn’t have arbitrary freedom to construct their measurement set, and may therefore not be able to implement a known optimal scheme. We introduce a simple procedure for minimizing the uncertainty in the reconstructed quantum state for an arbitrary tomographic scheme. We do this by defining a figure of merit based on the equally-weighted variance of the measurement statistics. This figure of merit is straightforwardly based on the singular value decomposition of the measurement matrix, making it well-suited for optimization.

PACS numbers: 03.65.Wj, 42.50.Ex

The characterization of quantum states is a key step in quantum technologies such as quantum computing \cite{1}, quantum information \cite{2} and quantum cryptography \cite{3}. This can be achieved by performing an informationally complete set of measurements on multiple identically-prepared copies of the state, using the results to reconstruct a representation of the state. Such a procedure is known as \textit{quantum state tomography} (QST) \cite{4–6}.

In principle, as long as the measurement set is informationally complete, the density matrix representing the state can always be reconstructed (Sec. \textsection I). However, in practice, certain measurement schemes are more prone to error than others.

The optimality of various tomographic schemes has been explored in works such as those by Scott \cite{7}, Roy and Scott \cite{8}, and de Burgh \textit{et al.} \cite{9}. Roughly speaking, for two-state systems, measurement schemes whose probability operators are distributed symmetrically on the Bloch sphere—such as those based on mutually unbiased bases and platonic solids—minimize the error in the reconstructed density matrix.

In this paper, we consider a scenario where the experimenter lacks the resources to implement an arbitrary measurement set, and may therefore not be able to implement a known optimal scheme. We show how one can determine the measurement settings which minimize the uncertainty in the reconstructed quantum state.

To do so, we introduce a figure of merit based on the equally-weighted variance (EWV) of the measurement statistics, that can be used to quantify the uncertainty in the reconstructed density matrix (Sec \textsection II). This figure of merit is based on the singular-value decomposition of the measurement matrix and is trivial to compute on a modern computer, making it ideal for the task of optimization. We demonstrate this figure of merit’s utility (Sec \textsection III), by optimizing wave-plate parameters in two non-standard optical tomographic scenarios.

I. TOMOGRAPHIC PROTOCOL

In a typical tomography experiment, multiple copies of the same quantum state interact with a measurement apparatus. The apparatus determines the number of states that registered the outcome \(j\), known as ‘counts’.

This can be modelled with an informationally complete set of probability operators \(\hat{E}_j\). For our purposes, it is more natural to work directly with measurement statistics rather than probabilities. We therefore consider \textit{unnormalized} probability operators that form an \textit{unnormalized} POVM such that \(\sum_j \hat{E}_j = \mathbb{N}I\), where \(\mathbb{N}\) is the total number of counts for the POVM (for a nice introduction to the POVM formalism, refer to Ch 2.2.6 of \cite{2}). In the absence of an informationally complete POVM, multiple POVMs can be combined to form an informationally complete set.

The number of counts for a given outcome \(j\) is given by the expectation value of the corresponding probability operator

\[ n_j = \langle \hat{E}_j \rangle = \text{Tr}[\hat{\rho}\hat{E}_j], \quad (1) \]

where \(\hat{\rho}\) is the density matrix of a \(d\)-level quantum system. The total number of counts for a single POVM is given by \(\mathbb{N} = \sum_j n_j\).

Given a set of measurement outcomes \(n_j\), one is interested in inferring the density matrix of the measured quantum state. This may be done by first decomposing the unknown density matrix \(\hat{\rho}\) in terms of an orthogonal basis:

\[ \hat{\rho} = \frac{1}{\sqrt{2d}} \sum_{i=0}^{d^2-1} S_i \hat{\sigma}_i, \quad (2) \]
where \( \hat{\sigma} \) are the generalized Pauli matrices—a set of \( d \)-dimensional traceless, Hermitian matrices that are orthogonal under the Hilbert-Schmidt inner product, given by \( \text{Tr}[\hat{\sigma}_i \hat{\sigma}_j] = 2 \delta_{ij} \). The parameters \( S_i \) completely characterize the state \( \hat{\rho} \) and the condition \( S_0 = 1 \) ensures that the density matrix is normalized.

Inserting Eq. (2) into Eq. (1), we find a linear relationship between the parameters \( S_i \) and the measurement outcomes \( n_j \):

\[
n_j = \frac{1}{\sqrt{2d}} \sum_{i=0}^{d^2-1} S_i \text{Tr}[\hat{\sigma}_i \hat{E}_j].
\]

This can be written in compact matrix form as

\[
n = M a,
\]

where the vectors

\[
n = (n_1, \ldots, n_N)^T;
\]

\[
a = (S_0, S_1, \ldots, S_{d^2-1})^T,
\]

are related by the measurement matrix

\[
M_{ij} = \frac{1}{\sqrt{2d}} \text{Tr}[\hat{\sigma}_{j-1} \hat{E}_i].
\]

If the tomography scheme consists of \( N \geq d^2 \) operators and is informationally complete (i.e. \( M \) has \( d^2 \) singular values), the density matrix of an unknown state \( \hat{\rho} \) can be reconstructed by calculating the pseudo-inverse \( M^+ \) of the measurement matrix \( M \) to give the parameters \( a \) in terms of the measurement outcomes \( n \):

\[
a = M^+ n.
\]

If \( N = d^2, M^+ = M^{-1} \) is simply the inverse of \( M \) and if \( N > d^2 \), the tomography scheme may lead to an overcomplete set of Eqs. for \( n_j \).

A specific example of how the above formalism can be used to implement the familiar six-outcome tomographic protocol for qubits can be found in Appendix [A].

We note that the inversion technique described above amounts to performing least-squares estimation which does not always produce physical density matrices. Despite this, we argue in Appendix [B] that linear inversion provides a reasonable starting point for developing a figure of merit that characterizes the uncertainty in the reconstructed density matrix.

When it comes to actual state reconstruction in practice, one often resorts to using the popular maximum likelihood estimation method [10]. Alternatively, one can look to a growing number of exciting new techniques such as the forced purity routine [11], Bayesian mean estimation [12], hedged maximum likelihood estimation [13], compressed sensing [14], von Neumann entropy maximization [15], minimax estimation [16] and likelihood-free quantum inference [17], as well as techniques that focus on reconstructing the state with reliable error bars [18] and confidence regions [19].

### II. DERIVING A FIGURE OF MERIT

Our goal is to identify the set of operators \( \{\hat{E}_i\} \) that minimize the uncertainty in the reconstructed density matrix. To do so, we must first define a figure of merit that quantifies this uncertainty.

In our analysis, we consider statistical uncertainties in measurement outcomes rather than systematic errors, e.g. due to misalignment of optical elements. We therefore neglect error in \( M \). We note that a recent study comparing systematic and statistical errors in QST was performed by Langford [20].

Under these assumptions, we can relate uncertainties in the detected number of counts to uncertainties in the density matrix by

\[
\Delta a = M^+ \Delta n,
\]

where

\[
\Delta n = (\Delta n_1, \ldots, \Delta n_N)^T;
\]

\[
\Delta a = (\Delta S_0, \Delta S_1, \ldots, \Delta S_{d^2-1})^T.
\]

To assess a tomography scheme’s robustness to noise, we calculate the equally-weighted variance \( \text{EWV} \), defined in terms of the equal sum of the variances in the parameters \( a \) [21]:

\[
\text{EWV} = \sum_{i=0}^{d^2-1} \Delta S_i^2 = \Delta a \cdot \Delta a.
\]

From Eq. (9), it follows that

\[
\text{EWV} = (M^+ \Delta n) \cdot (M^+ \Delta n)
\]

\[
= \sum_{i=1}^{d^2} \left( \sum_{j=1}^{N} M_{ij}^2 \Delta n_j \right)^2.
\]

Statistical noise can take the form of signal-independent background noise such as additive Gaussian noise, or signal-dependent noise such as Poisson noise. The latter is endemic to experiments involving particles that are spontaneously generated, such as single photons or neutrons.

Here, we consider Poisson noise, which leads to a figure of merit that is input-state-dependent. An average over all input states gives an expression that corresponds to the result for signal-independent noise.

Fluctuations due to Poisson noise are statistically independent from one measurement to the other [22], simplifying the expression to

\[
\text{EWV} = \sum_{i=1}^{d^2} \sum_{j=1}^{N} (M_{ij}^2)^2 (\Delta n_j)^2.
\]

For Poisson processes, the variance is given by \( (\Delta n_j)^2 = n_j \), thus

\[
\text{EWV} = \sum_{i=1}^{d^2} \sum_{j=1}^{N} (M_{ij}^2)^2 n_j.
\]
From Eq. [4], we see that \( n_j = \sum_i M_{jk} a_k \), giving

\[
\text{EWV} = \sum_{i,k=1}^{d^2} \left( M_{ij}^+ \right)^2 M_{jk} a_k.
\] (17)

In Appendix [C] we make use of the singular value decomposition (SVD) \( M^+ = \left( V^+ \right)^T S^+ U^+ \) (i.e. \( M = USV^T \)) to arrive at the final expression for the equally-weighted variance:

\[
\text{EWV} = \sum_{i,k=1}^{d^2} \frac{1}{\mu_i^2} \sum_{k=1}^{d^2} r_{i,k} a_k,
\] (18)

where \( \mu_i \) are the singular values of \( M \) and

\[
r_{i,k} = \sum_{j=1}^{d^2} (U_{ij}^{-1})^2 M_{jk} = \sum_{j=1}^{d^2} U_{jk}^2 M_{jk}.
\] (19)

Eq. (18) is an expression for the equally-weighted variance of a tomography scheme undergoing Poisson noise statistics. As expected for signal-dependent noise processes, it is dependent on the target state given by the vector \( a \). In the next section, we calculate the average equally-weighted variance.

We emphasize that our figure of merit is applicable for tomographic schemes described by the above formalism. Namely, those where the unknown state is parametrized by a Bloch vector; all Bloch vector elements are of equal interest; and that the probability operators form a POVM. Take for counter-example the scheme for measuring the diagonal elements of the density matrix of a single-mode state of an electromagnetic field introduced by Mogilevtsev [23]. For this scheme, using our notation, we have \( \hat{\sigma}_i = |i\rangle \langle i| \) and \( \hat{E}_i = \sum_k (1 - \eta_j) |k\rangle \langle k| \), in the Fock basis. The measurement matrix is then given by the Vandermonde matrix. Naive calculation of the EWV from this measurement matrix may not lead to reasonable assessment of the uncertainty in this scheme, since \( \hat{\sigma}_i \) do not form a complete basis and \( \hat{E}_j \) do not satisfy the condition \( \sum_j \hat{E}_j = \mathbb{I} \).

## A. Average EWV

In many cases, one is interested in how well a tomography scheme performs overall, rather than for one particular state. In this section, we address this question by calculating the average equally-weighted variance for all possible input states:

\[
\langle \text{EWV} \rangle = \sum_{i=1}^{d^2} \frac{1}{\mu_i^2} \sum_{k=1}^{d^2} r_{i,k} \langle a_k \rangle.
\] (20)

The average over the first element of \( a \) is trivial, i.e. \( \langle a_1 \rangle = \langle S_0 \rangle = 1 \). To calculate the average for other values of \( k \), i.e. \( \langle S_k \rangle \) \( \forall k \neq 0 \), we follow the approach of Życzkowski et al. [24] and assume an average over an ensemble of random states for which the probability measure may be factorized. The distribution of eigenvalues and eigenvectors are therefore independent and, hence, can be averaged independently. While the average over eigenvectors is relatively straightforward, performing the average over eigenvalues can be quite tricky as there is no single probability measure for mixed states. Fortunately, as we show below, for any set of eigenvalues, the average over all possible eigenvectors gives \( \langle S_k \rangle = 0 \) \( \forall k \neq 0 \), saving us from the difficulty of averaging over eigenvalues.

To calculate \( \langle S_k \rangle \) we invert Eq. (2) such that

\[
a_{k+1} = S_k = \sqrt{\frac{d}{2}} \text{Tr} \left[ \sum_{i=1}^{d} \lambda_i |\psi_i\rangle \langle \psi_i| \hat{\rho}_{S_k} \right],
\] (21)

where we used \( \text{Tr}[\hat{\sigma}_i \hat{\sigma}_j] = 2 \delta_{ij} \) and made use of the eigenvalue decomposition \( \hat{\rho} = \sum_{i=1}^{d} \lambda_i |\psi_i\rangle \langle \psi_i| \) where \( \lambda_i \geq 0 \) and \( \sum_{i=1}^{d} \lambda_i = 1 \). We can write the average of Eq. (21) over all possible eigenbases \( \langle \gamma \rangle_b \) for a given set of eigenvalues \( \{ \lambda_i \} \) as

\[
\langle S_k \rangle_b = \sqrt{\frac{d}{2}} \left( \text{Tr} \left[ \sum_{i=1}^{d} \lambda_i |\psi_i\rangle \langle \psi_i| \hat{\sigma}_{S_b} \right] \right).
\] (22)

As pointed out in [24], it is natural to assume that the eigenvectors are distributed according to the unique, unitarily invariant, Haar measure on unitary \( d \times d \) matrices. Different assignments of eigenvectors to eigenvalues are all equally likely since any rearrangement of the eigenvectors could be done by a unitary transformation, and the Haar measure is invariant under unitary transformations. For a given set of eigenvalues \( \{ \lambda_i \} \), we can therefore decompose the sum over eigenbases into a sum over permutations of \( \{ 1, 2, \ldots, d \} \) as follows

\[
\langle S_k \rangle_b = \sqrt{\frac{d}{2d!}} \left( \text{Tr} \left[ \sum_{i=1}^{d} \lambda_i |\psi_{\pi(i)}\rangle \langle \psi_{\pi(i)}| \hat{\sigma}_{S_b} \right] \right),
\] (23)

where \( P_d \) is the set of all possible maps that permute \( \{ 1, 2, \ldots, d \} \). We can rewrite the sum over permutations of eigenbases in Eq. (25) as the sum over permutations of eigenvalues as follows

\[
\langle S_k \rangle_b = \sqrt{\frac{d}{2d!}} \left( \text{Tr} \left[ \sum_{i=1}^{d} \left( \sum_{\pi \in P_d} \lambda_{\pi(i)} |\psi_i\rangle \langle \psi_i| \hat{\sigma}_{S_k} \right) \right] \right),
\] (24)

where \( \pi^{-1} \) is the inverse of the map \( \pi \). We note that \( \sum_{\pi \in P_d} \lambda_{\pi^{-1}(i)} = \sum_{\pi \in P_d} \lambda_{\pi(i)} \) and that this is independent of the index \( i \). We can therefore write

\[
\langle S_k \rangle_b = \sqrt{\frac{d}{2d!}} \left( \sum_{\pi \in P_d} \lambda_{\pi(i)} \left( \text{Tr} \left[ \sum_{i=1}^{d} |\psi_i\rangle \langle \psi_i| \hat{\sigma}_{S_k} \right] \right) \right),
\] (25)
Using the fact that $\sum_{i=1}^{d} |\psi_i\rangle\langle\psi_i| = I$, and that the generalized Pauli matrices are traceless, we find that $\langle S_k \rangle_b = 0 \forall k \neq 0$. We therefore conclude that $\langle S_k \rangle = 0 \forall k \neq 0$, and Eq. (20) simplifies to

$$\langle \text{EWV} \rangle = \frac{1}{d^2} \sum_{i=1}^{d^2} r_{i,1}. \tag{26}$$

Inserting the expression for $M_{j,i}$ in Eq. (7) into the expression for $r_{i,k}$ in (19), we evaluate the expression for $r_{i,1}$:

$$r_{i,1} = \frac{1}{\mu_i^2} \sum_{j=1}^{d^2} U_{j,i}^2 M_{j,1}$$

$$= \frac{1}{\mu_i^2} \sum_{j=1}^{d^2} \frac{1}{2d} U_{j,i}^2 \text{Tr}[\sigma_0 \hat{E}_j]. \tag{27b}$$

The operators $\hat{E}_j$ can be thought of as scaled projectors such that $\text{Tr}[\hat{E}_j] = \alpha_j$. We use this, as well as the definition $\sigma_0 = \sqrt{2/d} I$, to give

$$r_{i,1} = \frac{1}{d} \sum_{j=1}^{d^2} \alpha_j U_{j,i}^2, \tag{28}$$

which combined with Eq. (26) gives

$$\langle \text{EWV} \rangle = \frac{1}{d} \sum_{j=1}^{d^2} \alpha_j U_{j,i}^2. \tag{29}$$

Eq. (29) is an expression for the average EWV of a tomography scheme undergoing Poisson noise statistics. This expression is trivial to compute using a modern computer and does not require computational averaging over input states—in contrast to, say, the fidelity [9]—making it ideal for optimization over tuneable parameters.

If all probability operators are equally weighted, i.e. $\alpha_j = \alpha$, Eq. (29) simplifies to

$$\langle \text{EWV} \rangle = \frac{\alpha}{d} \sum_{i=1}^{d^2} 1 \mu_i^2, \tag{30}$$

since each column of the unitary matrix $U$ has unit length.

We note that the $\langle \text{EWV} \rangle$ corresponds to the EWV for signal-independent noise, such as additive Gaussian noise.

**B. Lower bound for $\langle \text{EWV} \rangle$**

When evaluating the robustness of a specific tomography scheme to noise, it is useful to make comparisons with the best possible scheme, i.e. one that minimizes $\langle \text{EWV} \rangle$.

In this section we derive a lower bound for the average EWV. The approach we use is to find a relationship between the diagonal elements and singular values of the Hermitian matrix $M^T M$. This puts constraints on the parameters which determine (EWV). We then use Lagrange multipliers, which provide a strategy for optimizing a function subject to equality constraints, to minimize $\langle \text{EWV} \rangle$.

First, consider the singular value decomposition of the matrix $M^T M$, where $M$ is defined in Eq. (4).

$$M^T M = (V S^T U^T) (U S V^T) \tag{31a}$$

$$= V S^T S V^T \tag{31b}$$

$$= W W^T, \tag{31c}$$

where $V$ is a $d \times d$ orthogonal matrix. Comparing Eqs. (31a) and (31c), we find that the eigenvalues $\{\lambda_i\}_{i=1}^d$ of the matrix $M^T M$ are related to the singular values of the matrix $M$ by $\lambda_i = \mu_i^2$.

To calculate a lower bound for the uncertainty, we begin by considering the sum of the diagonal elements $\{d_j\}_{j=1}^d$ of the matrix $M^T M$:

$$\sum_{j=1}^{d^2} d_j = \sum_{i=1}^{N} \sum_{j=1}^{d^2} M_{j,i} M_{i,j}$$

$$= \frac{1}{2d} \sum_{i=1}^{N} \sum_{j=0}^{d^2-1} \text{Tr}[\hat{\sigma}_j \hat{E}_i]^2. \tag{33}$$

The set of operators

$$B = \left\{ \frac{\hat{\sigma}_j}{\sqrt{2}} \right\} \text{ for } j = 0, \ldots, d^2 - 1 \tag{34}$$

forms an orthonormal basis, in the sense of Hilbert-Schmidt inner product, for the vector space of all $d \times d$ Hermitian matrices over real numbers. Recall that $\hat{E}_i$ are scaled projectors with a trace $\lambda_i$. We can therefore write the “Hilbert-Schmidt length” of the vector $\hat{E}_i$ in this vector space as

$$\text{Tr}[\hat{E}_i \hat{E}_i] = \frac{1}{2} \sum_{j=0}^{d^2-1} \text{Tr}[\hat{\sigma}_j \hat{E}_i]^2 = \lambda_i^2, \tag{35}$$

which simplifies Eq. (33) to

$$\sum_{j=1}^{d^2} d_j = \frac{\sum_{i=1}^{N} \lambda_i^2}{d}. \tag{36}$$

We now make use of the Schur-Horn theorem [25, 26], which states that there exists an $m \times m$ Hermitian matrix with diagonal values $\{d_i\}_{i=1}^m$ and eigenvalues $\{\lambda_i\}_{i=1}^m$ that are both ordered non-increasingly, if and only if,

$$\sum_{i=1}^{l} d_i \leq \sum_{i=1}^{l} \lambda_i, \forall l \in \{1, 2, ..., m\} \tag{37}$$
and
\[
\sum_{i=1}^{m} d_i = \sum_{i=1}^{m} \lambda_i .
\] (38)

Bearing in mind that the matrix \( M^T M \) is Hermitian, we find that
\[
\sum_{i=1}^{d^2} \mu_i = \sum_{j=1}^{d^2} d_j = \frac{\sum_{i} N}{d} \alpha_i^2 ;
\] (39a)
and
\[
\mu_1^2 = \lambda_1 \geq \frac{\sum_{i} N \alpha_i^2}{d^2} .
\] (39b)

Given the above constraints and the fixed sum \( \sum_i N \alpha_i \), which corresponds to having a fixed total number of counts for a given POVM, we used Lagrange multipliers to show that for a given \( \mu_1 \), the expression \( \sum_{i=1}^{d^2} 1/\mu_i^2 \) is minimal when \( \mu_2 = \cdots = \mu_d \), and \( \alpha := \alpha_1 = \cdots = \alpha_N \). This reduces the problem to a single-variable optimization problem, which yields
\[
\mu_1^2 = \frac{N \alpha^2}{d^2} ;
\] (40a)
and
\[
\mu_2^2 = \cdots = \mu_d^2 = \frac{N \alpha^2}{d^2(d+1)} .
\] (40b)

Under these conditions, we can find a lower bound for the average EWV by inserting the expressions for \( \mu_i \) in Eqs. (40) into Eq. (30). The final expression for the lower bound for the average EWV is
\[
\langle \text{EWV} \rangle_{LB} = \frac{d^2(d^2 + d - 1)}{N \alpha} .
\] (41a)

Given a completely unknown state, one is interested in a tomographic scheme that performs best “on average”. Symmetric informationally-complete positive operator valued measures (SIC-POVMs) are known to be optimal for quantum state tomography under signal-independent noise [7]. We find that they are also optimal given Poisson noise processes, by showing that the \( \langle \text{EWV} \rangle \) for SIC-POVMs is equal to \( \langle \text{EWV} \rangle_{LB} \) (see Appendix D).

III. EXAMPLES

The \( \langle \text{EWV} \rangle \) can be used to determine the optimal tomography scheme under a set of practical constraints. We demonstrate this with two examples, using two-state systems, i.e. qubits, encoded in the polarization degree of freedom of single photons. In such systems, the measurement is typically made in the horizontal/vertical basis by counting photons at the output ports of a polarizing beam splitter (PBS). Measurements in different bases are implemented by placing wave plates at different orientations in front of the beam splitter.

In Scenario A, we are restricted to one wave plate per qubit mode, but have the freedom to choose the retardance of each wave plate (Fig. 1a). We use \( \langle \text{EWV} \rangle \) to optimize over the retardance of the wave plate.

In Scenario B, we consider two wave plates with which we can perform QST on a single qubit mode, however, the wave plates are optimized for a different optical frequency to the one used in the experiment (Fig. 3a). We use \( \langle \text{EWV} \rangle \) to optimize over the orientation of the wave plates about the optical axis.

A. Optimization of wave plate retardance

A single wave plate in front of a PBS can generate a POVM consisting of two operators associated with the horizontal and vertical output modes of each PBS, given by \( \{ \hat{E}_H(\beta, \phi), \hat{E}_V(\beta, \phi) \} \) where
\[
\hat{E}_a(\beta, \phi) = N \hat{U}^\dagger(\beta, \phi)|a\rangle \langle a| \hat{U}(\beta, \phi) ,
\] (42)
for \( a = H, V \). In fact, these operators correspond to the special case of a projective-value measure (PVM), where \( \hat{E}_i \hat{E}_j = \hat{E}_i \delta_{i,j} \). The unitary operator associated with the wave plate,
\[
\hat{U}(\beta, \phi) = \cos \left( \frac{\beta}{2} \right) \hat{\sigma}_0 - i \sin \left( \frac{\beta}{2} \right) \hat{v}(\phi) \cdot \hat{\sigma} ,
\] (43)
rotates the operators \( |a\rangle \langle a| \) on the Bloch sphere by an angle \( \beta \), about the vector
\[
\hat{v}(\phi) = \cos(\phi) \hat{k} + \sin(\phi) \hat{i} ,
\] (44)
where $\vec{k}$ and $\vec{v}$ are unit vectors in Euclidean space (defined by the axes in Fig. 1) and $\vec{v} \cdot \vec{r} = v_1 \sigma_1 + v_2 \sigma_2 + v_3 \sigma_3$. As the wave plate is rotated about the optical axis, the projector $\hat{E}_H(t)$ (after appropriate normalization) traces out a figure-eight path on the Bloch sphere, as shown in Fig. 1. The retardance of the wave plate $\beta$ determines the size of the figure eight.

Although $\{\hat{E}_H(\beta, \phi), \hat{E}_V(\beta, \phi)\}$ is not informationally complete, an informationally-complete set of operators can be constructed with different orientations of the wave plate about the optical axis, given by the angle $\phi$, as long as the retardance is not equal to an integer multiple of $\pi$. Here, the total number of projectors $N$ is twice the number of PVMs used to construct the complete tomographic protocol. This scenario was considered in the Fourier Transform Tomography (FTT) scheme introduced in [27].

The task is to identify the optimal retardance $\beta$ that minimizes the average EWV. We consider a tomographic protocol which consists of a set of PVMs whose elements are distributed along the figure eight with equally-spaced values of $\phi$. From these operators, we calculate the coefficients $\alpha_i = \text{Tr}[\hat{E}_i] = \mathcal{N}$. We also construct a measurement matrix $M$ according to Eq. (7) and calculate its singular values $\mu_i$. For $N = 12$ (i.e. six different values of the retardance which corresponds to six different PVMs), the measurement matrix is given by:

$$M^{(12)} = \frac{N}{2} \begin{pmatrix}
1 & -\sqrt{3}(\cos \beta - 1) & -\sin \beta & \frac{(\cos \beta + 3)}{4} \\
1 & \sqrt{3}(\cos \beta - 1) & -\sin \beta & \frac{\cos \beta}{4} \\
1 & \sqrt{3}(\cos \beta - 1) & \sin \beta & \frac{(\cos \beta + 3)}{4} \\
1 & -\sqrt{3}(\cos \beta - 1) & \sin \beta & \frac{\cos \beta}{4} \\
1 & 0 & 2\sin \beta & \frac{(\cos \beta + 3)}{4} \\
1 & 0 & -2\sin \beta & \frac{\cos \beta + 3}{4} \\
1 & -\sqrt{3}(\cos \beta - 1) & -\sin \beta & \frac{\cos \beta}{4} \\
1 & \sqrt{3}(\cos \beta - 1) & -\sin \beta & \frac{\cos \beta + 3}{4} \\
1 & 0 & 2\sin \beta & \frac{\cos \beta}{4} \\
1 & 0 & -2\sin \beta & \frac{(\cos \beta + 3)}{4} \\
1 & -\sqrt{3}(\cos \beta - 1) & \sin \beta & \frac{\cos \beta}{4} \\
1 & \sqrt{3}(\cos \beta - 1) & \sin \beta & \frac{(\cos \beta + 3)}{4} \\
1 & 0 & 2\sin \beta & \frac{(\cos \beta + 3)}{4} \\
1 & 0 & -2\sin \beta & \frac{(\cos \beta + 3)}{4}
\end{pmatrix} (45)$$

Using Mathematica [31], the singular values were found to be:

$$\mu_1 = \frac{N\sqrt{N}}{\sqrt{2}} ; (46a)$$

$$\mu_2 = \frac{N\sqrt{N}}{4\sqrt{2}} \sqrt{9 + 4\cos(\beta) + 3\cos(2\beta)} ; (46b)$$

$$\mu_3 = \frac{N\sqrt{N}}{2} |\sin(\beta)| ; (46c)$$

$$\mu_4 = \frac{N\sqrt{N}}{2} \sin^2 \left(\frac{\beta}{2}\right) , (46d)$$

for $N \geq 5$. Knowing $\mu_i$ and $\alpha_i$, we can calculate the average EWV according to Eq. (29), which is given by

$$\langle \text{EWV} \rangle = \frac{2}{N_T} \left( 1 + 4 \csc^4 \left( \frac{\beta}{2} \right) + \csc^2(\beta) \right)$$

$$+ \frac{9 + 4\cos(\beta) + 3\cos(2\beta)}{8} , (47)$$

where $N_T = N/2$ is the total number of counts. Fig. 2 shows $\langle \text{EWV} \rangle$ as a function of the retardance $\beta$ for 6 equally-spaced time bins for a single qubit (i.e. $N = 12$). This function reaches a minimum of $\langle \text{EWV} \rangle \approx 10.03/N_T$ when $\beta \approx 7\pi/10$. For comparison, a protocol that consists of three PVMs (see Appendix A)—where the measurement operators of each PVM are the eigenstates of the Pauli operators—gives $\langle \text{EWV} \rangle = 10/N_T$, where $N_T = 3N$. The four-outcome SIC-POVM also gives $\langle \text{EWV} \rangle = 10/N_T$, where here $N_T = N$.

**B. Optimization of wave plate orientation**

The second scenario we consider consists of optimization over wave plate orientation. Given two wave plates—a half-wave plate ($\beta = \pi/2$) and a quarter-wave plate ($\beta = \pi/4$)—one is able to access the entire surface of the Bloch sphere. A wave plate’s retardance, however, is dependent on the wavelength of the input light; using the wrong wave plate results in a non-standard retardance. Here we consider the situation where the experimenter only has access to two identical wave plates, designed for a different optical wavelength. In such a scenario, access to the entire Bloch sphere may not be possible.

For two wave plates of retardance $\beta$, the measurement operators associated with the horizontal and vertical output modes of each PBS are given by
constructed a measurement matrix $\mathbf{M}$ to find its singular values $\mu_i$ as well as calculating $\alpha_i$ to compute $\langle \text{EWV} \rangle$.

For $\beta = 3\pi/8$, optimization over $\phi_1$ and $\phi_2$ yields $\langle \text{EWV} \rangle \approx 10.01/N_T$, which occurs at $\phi_1 \approx 7\pi/10$ and $\phi_2 \approx 7\pi/10$ for the second PVM, and $\phi_1 \approx 7\pi/10$ and $\phi_2 \approx \pi/5$ for the third PVM. The resulting six projectors are shown in Fig. 3 b). Fig. 4 shows the optimal $\langle \text{EWV} \rangle$ for different $\beta$; showing that there exists a large retardance window that can lead to effectively optimal state reconstruction.

IV. CONCLUDING REMARKS

We have shown how one can use a simple figure of merit to optimize a tomographic protocol when practical constraints prohibit the implementation of known optimal protocols. In doing so, we considered statistical errors in the form of Poisson noise and assumed no systematic errors. One could in principle follow a similar procedure for noise models appropriate to other experimental scenarios, although whether such a simple figure of merit arises is to be determined. Systematic errors could also be included by considering uncertainties in the measurement matrix, e.g. $\Delta \alpha = \mathbf{M} + \Delta \mathbf{n} + \Delta \mathbf{M} + \mathbf{n}$.

This procedure could be straightforwardly extended to quantum process tomography [32,34] where the measurement statistics relate linearly to the unknown density matrix parameters. Extensions to tomographic schemes that consider unknown parameters in both state and process parameters may also be possible [35–39]. However, we anticipate the nonlinear relationship between measurement statistics and unknown parameters to provide additional challenges.

Our figure of merit is closely related to the trace distance, introduced by Scott [7]. However, the trace distance is only applicable for signal-independent noise, while our figure of merit accounts for signal-dependent noise. Our figure of merit does not possess the same operational interpretation that might make other figures of merit—such as the fidelity [9]—more appealing, however it is vastly simpler to work with as it does not require computational averaging over input states.

We anticipate that the method introduced in this paper will help the design of tomographic experiments performed in laboratories with minimal available resources. Furthermore, our figure of merit can be used to compare the performance of different tomographic protocols, in complement to other figures of merit such as the fidelity.

V. ACKNOWLEDGEMENTS

We thank Daniel James for helpful discussions. We also thank Joe Altepeter, Hubert de Guise and Alessandro Fedrizzi for helpful comments on the manuscript. MM acknowledges support from NSERC (USRA) and
The Fields Institute. AMB acknowledges support from DARPA (QuBE).

Appendix A: Six-measurement example for qubits

Consider a qubit encoded in the polarization of a single photon. One of the simplest measurements one can implement is a two-outcome POVM with elements
\[ \hat{E}_1 = \mathcal{N}|H\rangle\langle H|; \quad \hat{E}_2 = \mathcal{N}|V\rangle\langle V|. \] (A1)
In fact, these operators correspond to the special case of a projective-value measure (PVM), where \( \hat{E}_i\hat{E}_j = \delta_{ij}\). This measurement is implemented by sending multiple copies of the state through a polarizing beam splitter (which splits the state into horizontal and vertical modes) and counting the number of photons in each output port.

However, this is not enough to reconstruct the quantum state because this PVM isn’t informationally complete. Two additional PVMs can be constructed by inserting a half-wave plate (HWP) followed by a quarter-wave plate (QWP) in front of the PBS. Setting the HWP and QWP at 22.5° and 45° to the optical axis, respectively, generates the following PVM:
\[ \hat{E}_3 = \mathcal{N}|D\rangle\langle D|; \quad \hat{E}_4 = \mathcal{N}|A\rangle\langle A|. \] (A2)
where \( |D/A\rangle = |H\rangle \pm |V\rangle \sqrt{2} \). Setting them to 0° and 45° gives
\[ \hat{E}_5 = \mathcal{N}|R\rangle\langle R|; \quad \hat{E}_6 = \mathcal{N}|L\rangle\langle L|, \] (A3)
where \( |R/L\rangle = (|H\rangle \pm i|V\rangle) / \sqrt{2} \). For all three PVMs, we have assumed equal data collections times and therefore the total number of counts is \( \mathcal{N} \) in all three cases.

Performing the experiment, we get a set of six numbers which correspond to the counts at both output ports of the PBS for each of the three PVMs. These can be written as
\[ \mathbf{n} = (n_1, n_2, n_3, n_4, n_5, n_6)^T. \] (A4)
We are interested in finding the vector
\[ \mathbf{a} = (S_0, S_1, S_2, S_3)^T, \] (A5)
which characterizes the state of our qubit.

The measurement matrix \( \mathbf{M} \), which relates \( \mathbf{n} \) to \( \mathbf{a} \) via Eq. (A4), is given by
\[ \mathbf{M} = \frac{\mathcal{N}}{2} \begin{pmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & -1 \\ 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & -1 & 0 \end{pmatrix}, \] (A6)
where
\[ M_{i,j} = \frac{1}{2} \text{Tr}[\hat{\sigma}_{j-1}\hat{E}_i], \] (A7)
and \( \hat{\sigma}_0 = |H\rangle\langle H| + |V\rangle\langle V|, \hat{\sigma}_1 = |H\rangle\langle V| + |V\rangle\langle H|, \hat{\sigma}_2 = i|V\rangle\langle H| - i|H\rangle\langle V| \) and \( \hat{\sigma}_3 = |H\rangle\langle H| - |V\rangle\langle V| \). The pseudoinverse of \( \mathbf{M} \) is given by
\[ \mathbf{M}^+ = \frac{1}{\mathcal{N}} \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & 0 & 1 & -1 & 0 \end{pmatrix}, \] (A8)
which can be used to determine the vector \( \mathbf{a} \) as follows:
\[ \mathbf{a} = \mathbf{M}^+ \mathbf{n} = \left( \frac{1}{3} \sum_{i=1}^{6} n_i \right) \begin{pmatrix} n_3 - n_4 \\ n_5 - n_6 \\ n_1 - n_2 \end{pmatrix}. \] (A9)
The vector \( \mathbf{a} \) then gives the density matrix according to
\[ \hat{\rho} = \frac{1}{2} \sum_{j=1}^{4} a_j \hat{\sigma}_{j-1}. \] (A10)

Appendix B: Discussion of linear inversion and positive semi-definiteness

Use of linear inversion (LI) for state reconstruction yields coefficients \( S_i \) that may not necessarily correspond to a physical state, i.e., the estimated density matrices may be unnormalized or negative definite.

We can denote the estimated coefficients obtained from this method by \( S_{LI} = (S_{0,LI}, S_{1,LI}, ..., S_{d-1,LI}) \). We also denote coefficients corresponding to the actual state (A) of the system by \( S_{A} = (S_{0,A}, S_{1,A}, ..., S_{d-1,A}) \).

The proposed figure of merit, the EWV, is the expectation value of \( \sum_{i=0}^{d-1} (S_{i,LI} - S_{i,A})^2 \). As \( S_{LI} \) may not correspond to a physical state, the experimenter could solve this problem by, say, looking for a physical state (Ph) that is closest to the one obtained by linear inversion; we denote this state by \( S_{Ph} \) and characterize the closeness of states by the distance function \( d(S_{Ph}, S_{LI}) = (\sum_{i=0}^{d-1} (S_{i,Ph} - S_{i,LI})^2)^{1/2} \). In this scenario, we are interested in the expectation value of \( d^2(S_{Ph}, S_{A}) \), as a measure for the error. Since \( S_{Ph} \) is the closest physical state to \( S_{LI} \), we have \( d(S_{Ph}, S_{LI}) \leq d(S_{Ph}, S_{A}) \).

The error computed while taking the positive semi-definiteness of the density matrix into account would be bounded by a coefficient of the old figure of merit. Therefore, minimizing the proposed figure of merit, EWV, also minimizes an upper bound for the error that takes positive semi-definiteness into account.

We add that the geometry of physical states is especially complicated and there are many open problems—e.g., the existence of SIC-POVMs—associated with it. We therefore consider our approach justified in that it provides a simple procedure that leads to reasonable results.
Appendix C: Simplify expression for equally-weighted variance

In this section, we detail the steps between Eqs. (17) and (18). For reference, we repeat Eq. (17) here:

$$\text{EWV} = \sum_{i,k=1}^{d^2} \sum_{j=1}^{N} (M_{ij}^+)^2 M_{jk} a_k.$$  (C1)

We define a vector

$$g_k = \sum_{i,k=1}^{d^2} \sum_{j=1}^{N} (M_{ij}^+)^2 M_{jk}.$$  (C2)

such that Eq. (C1) becomes

$$\text{EWV} = \sum_{k=1}^{d^2} g_k a_k.$$  (C3)

We rewrite the term $g_k$ as

$$g_k \equiv \text{Tr} \left[ M^+ Y^{(k)} (M^+)^T \right]$$  (C4)

where $Y^{(k)}$ is an $N \times N$ diagonal matrix whose diagonal elements are the elements of the $k$th column of $M$, i.e. $Y_{jj}^{(k)} = M_{jk}$. We perform a singular value decomposition $M^+ = (V^+)^T S^+ U^+$ (i.e. $M = USV^T$), where $U$ and $V$ are $d^2 \times d^2$ and $N \times N$ matrices respectively, to give

$$g_k = \text{Tr} \left[ (U^+)^T (S^+)^T V^+ (V^+)^T S^+ U^+ Y^{(k)} \right]$$

$$= \text{Tr} \left[ (S^+)^T S^+ X \right]$$  (C6)

where $X = U^+ Y^{(k)} (U^+)^T$. Since $S^+$ is a diagonal matrix of inverse singular values of $M$, i.e. $S^2_i = 1/\mu_i$, we can write

$$g_k = \sum_{i=1}^{d^2} \frac{1}{\mu_i^2} X_{i,i}$$  (C7)

$$= \sum_{i=1}^{d^2} \sum_{j=1}^{d^2} \left( \sum_{i=1}^{d^2} \frac{1}{\mu_i^2} (U_{ij}^+)^2 Y_{jj}^{(k)} \right)$$  (C8)

$$= \sum_{i=1}^{d^2} \sum_{j=1}^{d^2} (U_{ij}^+)^2 M_{jk}$$  (C9)

Inserting Eq. (C9) into Eq. (C3), we now have

$$\text{EWV} = \sum_{i=1}^{d^2} \sum_{j=1}^{d^2} \sum_{k=1}^{d^2} r_{i,k} a_k.$$  (C10)

where

$$r_{i,k} = \sum_{j=1}^{d^2} (U_{ij}^+)^2 M_{jk}.$$  (C11)

Eqs. (C10) and (C11) correspond to Eqs. (18) and (19) in the main text.

Appendix D: SIC-POVM

In this section, we show that the average equally-weighted variance (EWV) for any symmetric informationally-complete POVM (SIC-POVM) necessarily reaches the lower bound (EWV)$_\text{LB}$. SIC-POVMs are therefore optimal for QST in the sense that they are the least sensitive POVM to Poisson noise. This is consistent with the result that SIC-POVMs are the most immune minimal informationally complete POVM for signal-independent noise [7]. We note, however, that the existence of SIC-POVMs in arbitrary dimensions has not yet been proven or disproven.

To show that (EWV) for SIC-POVMs equals (EWV)$_\text{LB}$, we show that the eigenvalues of the matrix $M^T M$ (constructed using a SIC-POVM) are $\mu_1^2 = \alpha^2$ and $\mu_2^2 = \cdots = \mu_d^2 = \alpha^2/(d+1)$, i.e. the eigenvalues for the lower bound.

First, consider an unnormalized SIC-POVM given by $\{\hat{E}_i\}_{i=1}^d$. By definition

$$\text{Tr} [\hat{E}_i \hat{E}_j] = N^2 \text{Tr} [\hat{F}_i \hat{F}_j];$$  (D1a)

$$\{ \begin{array}{c}
\frac{N^2}{2(d^2+1)} & \text{for } i \neq j \\
\frac{N^2}{2d^2} & \text{for } i = j
\end{array} \}$$  (D1b)

where $\hat{F}_i = \hat{E}_i/N$ are normalized probability operators. One can also see that:

$$\alpha = \text{Tr} [\hat{E}_i] = N \text{Tr} [\hat{F}_i] = \frac{N}{d}.$$  (D2)

Using the fact that the set of operators $B$, defined in Eq. (34), form an orthonormal basis in the sense of the Hilbert-Schmidt inner product, we write

$$\text{Tr} [\hat{E}_i \hat{E}_j] = \frac{1}{2} \sum_{k=0}^{d^2-1} \text{Tr} [\hat{\sigma}_k \hat{E}_i] \text{Tr} [\hat{\sigma}_k \hat{E}_j].$$  (D3)

We can equate Eqs. (D1) and (D3) to give

$$\sum_{k=0}^{d^2-1} \text{Tr} [\hat{\sigma}_k \hat{E}_i] \text{Tr} [\hat{\sigma}_k \hat{E}_j] = \{ \begin{array}{c}
\frac{2N^2}{d(d+1)} & \text{for } i \neq j \\
\frac{2N^2}{d^2} & \text{for } i = j
\end{array} \}.$$  (D4)

Now consider the inner product between two rows of the matrix $M$. For reference, we rewrite Eq. (7) here:

$$M_{i,j} = \frac{1}{\sqrt{2d}} \text{Tr} [\hat{\sigma}_{j-1} \hat{E}_i].$$  (D5)

From Eqs. (D4) and (D5), we can see that

$$\text{row}_i (M) \cdot \text{row}_j (M) = \{ \begin{array}{c}
\frac{N^2}{2(d+1)} & \text{for } i \neq j \\
\frac{N^2}{d^2} & \text{for } i = j
\end{array} \}.$$  (D6)

where $\cdot$ is the inner product in the usual sense.
The next piece of the puzzle is to prove that $M^T M$ is diagonal; the diagonal entries will therefore be the eigenvalues of $M^T M$. To do so, we use Eqs. (D5) and (D6) to construct the following matrix:

$$O = \frac{d \sqrt{d+1}}{N} L,$$

where the matrix $L$ is equal to $M$ with its first column $(N/d^2, \ldots, N/d^2)$ divided by $\sqrt{d+1}$. It is not difficult to confirm that the rows of $O$ are orthonormal. Thus, $O$ is an orthogonal matrix and consequently the columns of $O$ are also orthonormal. We can therefore conclude that the columns of $M$ are orthogonal, i.e.

$$\text{col}_i (M) \cdot \text{col}_j (M) = 0 \quad \text{for} \quad i \neq j,$$

and that $M^T M$ is a diagonal matrix.

We now want to calculate expressions for $d_i$. From Eq. (D5) we see that

$$d_1 = \text{col}_1 (M) \cdot \text{col}_1 (M) = \frac{N^2}{d^2}.$$  

From the orthogonality of $O$ we also see that

$$d_i = \text{col}_i (M) \cdot \text{col}_i (M) = \frac{N^2}{d^2(d+1)} \quad \text{for} \quad i > 1.$$  

By showing that $M^T M$ is a diagonal matrix with the diagonal entries $d_1 = N^2/d^2$ and $d_2 = \cdots = d_P = N^2/(d^3(d+1))$, we show that the eigenvalues of $M^T M$ are $\mu_1^3 = \alpha^2$ and $\mu_2^3 = \cdots = \mu_P^3 = \alpha^2/(d+1)$, i.e. the eigenvalues for the lower bound for $N = d^2$ in Eqs. (10).

We therefore prove that SIC-POVMs do reach the lower bound. We now need to prove that only SIC-POVMs reach the lower bound, when $N = d^2$.

We do so by showing that any set of operators $\{\tilde{\Pi}_i\}_{i=1}^N$, where $N = d^2$, that reaches the lower bound will necessarily correspond to a SIC-POVM. From the definition of the SVD in Eq. (31a) and the expressions for the singular values in Eqs. (40a) and (40b), we can write

$$M^T M = \frac{N^2}{d^2} V \begin{pmatrix} 1 & 0 & \cdots & 0 \\
0 & \frac{1}{d+1} & \cdots & 0 \\
0 & 0 & \cdots & 0 \\
0 & 0 & \cdots & \frac{1}{d+1} \end{pmatrix} V^T,$$

where $V$ is a $d^2 \times d^2$ orthogonal matrix. Eq. (D11) takes this form for any POVM that minimized the noise.

Since the highest eigenvalue of $M^T M$ is $N^2/d^2$, we can write the following for any $d^2$-dimensional vector of unit length, $v$:

$$v^T M^T M v \leq \frac{N^2}{d^2}.$$  

One can check that the equality holds for the vector $v_1 = (1,0,\ldots,0)^T$ and, therefore, $v_1$ is the eigenvector of $M^T M$ with eigenvalue $\frac{N^2}{d^2}$. It follows that $M^T M$ has the form

$$M^T M = \frac{N^2}{d^2} \left( v_1 v_1^T + \frac{1}{d+1} (I_d - v_1 v_1^T) \right).$$  

and that

$$M^T M = \frac{N^2}{d^2} \begin{pmatrix} 1 & 0 & \cdots & 0 \\
0 & \frac{1}{d+1} & \cdots & 0 \\
0 & 0 & \cdots & 0 \\
0 & 0 & \cdots & \frac{1}{d+1} \end{pmatrix}.$$  

From the diagonal entries of $M^T M$, we see that Eq. (D10) holds. Since $M^T M$ is diagonal, Eq. (D5) is also true.

Working backwards, we can see that matrix $O$, defined in Eq. (D7), is an orthogonal matrix, as it consists of orthonormal columns. From this, one can verify Eq. (D6) and consequently (D1). The measurement matrix $M$ that reaches the lower bound for the set of measurements $\{\tilde{\Pi}_i\}_{i=1}^d$, will necessarily correspond to a SIC-POVM.

To calculate the uncertainty due to a Poisson source, we worked with the number of counts, rather than the corresponding probabilities (as was done in [7]). Experimentally, probabilities are obtained by dividing the number of counts by the total number of counts, which is the summation of number counts in a POVM. Thus, probabilities are dependent parameters. This would make calculation of uncertainties a difficult task when dealing with the probabilities. In the presence of signal-dependent noise, all probabilities obtained from a POVM have the same uncertainty and results obtained when working with counts reduce to those obtained when working with probabilities.

[1] P. Kok et al., “Linear optical quantum computing with photonic qubits,” Rev. Mod. Phys. 79, 135 (2007).
[2] M. A. Nielsen and I. L. Chuang, Quantum computation and quantum information (Cambridge university press, 2010).
[3] C. H. Bennett et al., “Quantum cryptography: Public key distribution and coin tossing,” in Proceedings of IEEE International Conference on Computers, Systems and Signal Processing, 175 (1984).
[4] K. Vogel and H. Risken, “Determination of quasiprobability distributions in terms of probability distributions for the rotated quadrature phase,” Phys. Rev. A 40, 2847 (1989).
[5] U. Leonhardt, “Quantum-State Tomography and Discrete Wigner Function,” Phys. Rev. Lett. 74, 4101 (1995).
[6] D. F. V. James et al., “Measurement of qubits,” Phys. Rev. A 64, 052312 (2001).
[7] A. J. Scott, “Tight informationally complete quantum measurements,” Journal of Physics A: Mathematical and General 39, 13507 (2006).
[8] A. Roy and A. J. Scott, “Weighted complex projective 2-designs from bases: Optimal state determination by orthogonal measurements,” Journal of Mathematical Physics 48, 072110 (2007).
[9] M. D. de Burgh et al., “Choice of measurement sets in qubit tomography,” Phys. Rev. A 78, 052122 (2008).
[10] Z. Hradil, “Quantum-state estimation,” Phys. Rev. A 55, R1561 (1997).
[11] M. S. Kaznady and D. F. V. James, “Numerical strategies for quantum tomography: Alternatives to full optimization,” Phys. Rev. A 79, 022109 (2009).
[12] R. Blume-Kohout, “Optimal, reliable estimation of quantum states,” New Journal of Physics 12, 043034 (2010).
[13] R. Blume-Kohout, “Hedged Maximum Likelihood Quantum State Estimation,” Phys. Rev. Lett. 105, 200504 (2010).
[14] D. Gross et al., “Quantum State Tomography via Compressed Sensing,” Phys. Rev. Lett. 105, 150401 (2010).
[15] Y. S. Teo et al., “Quantum-State Reconstruction by Maximizing Likelihood and Entropy,” Phys. Rev. Lett. 107, 020404 (2011).
[16] H. K. Ng and B.-G. Englert, “A simple minimax estimator for quantum states,” arXiv:1202.5136 [quant-ph] (2012).
[17] C. Ferrie and C. E. Granade, “Likelihood-free quantum inference: tomography without the Born Rule,” arXiv:1304.5828 [quant-ph] (2013).
[18] M. Christandl and R. Renner, “Reliable Quantum State Tomography,” arXiv:1108.5329v1 [quant-ph] (2011).
[19] R. Blume-Kohout, “Robust error bars for quantum tomography,” arXiv:1202.5270v1 [quant-ph] (2012).
[20] N. K. Langford, “Errors in quantum tomography: diagnosing systematic versus statistical errors,” New Journal of Physics 15, 035003 (2013).
[21] D. S. Sabatke et al., “Optimization of retardance for a complete Stokes polarimeter,” Opt. Lett. 25, 802 (2000).
[22] F. Goudail, “Noise minimization and equalization for Stokes polarimeters in the presence of signal-dependent Poisson shot noise,” Opt. Lett. 34, 647 (2009).
[23] D. Mogilevtsev, “Diagonal element inference by direct detection,” Optics Communications 156, 307 (1998).
[24] K. Zyczkowski et al., “Generating random density matrices,” Journal of Mathematical Physics 52, 06220 (2011).
[25] I. Schur, “Über eine Klasse von Mittelbildungen mit Anwendungen auf die Determinantentheorie,” Sitzungsber. Berl. Math. Ges. 22, 9 (1923).
[26] A. Horn, “ Doubly Stochastic Matrices and the Diagonal of a Rotation Matrix,” American Journal of Mathematics 76 (1954).
[27] M. Mohammadi, A. M. Brańczyk, and D. F. V. James, “Fourier-transform quantum state tomography,” Phys. Rev. A 87, 012117 (2013).
[28] J. Reháček, B.-G. Englert, and D. Kaszlikowski, “Minimal qubit tomography,” Phys. Rev. A 70, 052321 (2004).
[29] A. Ling et al., “Experimental polarization state tomography using optimal polarimeters,” Phys. Rev. A 74, 022309 (2006).
[30] A. Kalev, J. Shang, and B.-G. Englert, “Symmetric minimal quantum tomography by successive measurements,” Phys. Rev. A 85, 052116 (2012).
[31] I. Wolfram Research, Mathematica Edition: Version 8.0 (Wolfram Research, Inc., Champaign, Illinois, 2010).
[32] I. L. Chuang and M. A. Nielsen, “Prescription for experimental determination of the dynamics of a quantum black box,” J. Mod. Opt. 44, 2455 (1997).
[33] J. B. Altepeter et al., “Ancilla-Assisted Quantum Process Tomography,” Phys. Rev. Lett. 90, 193601 (2003).
[34] J. L. O’Brien et al., “Quantum Process Tomography of a Controlled-NOT Gate,” Phys. Rev. Lett. 93, 080502 (2004).
[35] A. M. Brańczyk et al., “Self-calibrating quantum state tomography,” New Journal of Physics 14, 085003 (2012).
[36] D. Mogilevtsev, J. Reháček, and Z. Hradil, “Self-calibration for self-consistent tomography,” New Journal of Physics 14, 095001 (2012).
[37] S. T. Merkel et al., “Self-Consistent Quantum Process Tomography,” arXiv:1211.0322 [quant-ph] (2012).
[38] R. Blume-Kohout et al., “Robust, self-consistent, closed-form tomography of quantum logic gates on a trapped ion qubit,” ArXiv e-prints (2013).
[39] M. Takahashi, S. D. Bartlett, and A. C. Doherty, “Tomography of a spin qubit in a double quantum dot,” Phys. Rev. A 88, 022120 (2013).