Spin qudit tomography and state reconstruction error

Michael A. Perlin ●,* Diego Barberena ●, and Ana Maria Rey ●

*JILA, National Institute of Standards and Technology and University of Colorado, 440 UCB, Boulder, Colorado 80309, USA
and Center for Theory of Quantum Matter, University of Colorado, Boulder, Colorado 80309, USA*

We consider the task of performing quantum state tomography on a d-level spin qudit, using only measurements of spin projection onto different quantization axes. After introducing a basis of operators closely related to the spherical harmonics, which obey the rotational symmetries of spin qudits, we map our quantum tomography task onto the classical problem of signal recovery on the sphere. We then provide algorithms with $O(rd^3)$ serial runtime, parallelizable down to $O(rd^2)$ for: (i) computing *a priori* upper bounds on the expected error with which spin projection measurements along r given axes can reconstruct an unknown qudit state, and (ii) estimating *a posteriori* the statistical error in a reconstructed state. Our algorithms motivate a simple randomized tomography protocol for which we find that using more measurement axes can yield substantial benefits that plateau after $r \approx 3d$.

I. INTRODUCTION

Quantum state tomography, the task of reconstructing a quantum state by collecting and processing measurement data, is an essential primitive for quantum sensing, quantum simulation, and quantum information processing. The central importance of quantum state tomography has led to the development of techniques based on least-squares inversion [1], linear regression [2], maximum likelihood estimation [3,4], Bayesian inference [5–7], compressed sensing [8,9], and neural networks [10], among others. These techniques are typically developed in a general, information-theoretic setting and make minimal assumptions about the physical medium of a quantum state. As a consequence, even well-established techniques can be ill suited for physical platforms with unique or limited capabilities.

Due to advancements in experimental capabilities to address nuclear spin states (i.e., hyperfine levels) in ultracold atomic systems [11–16] as well as developments in the control of ultracold molecular systems [17–27], a particular setting of growing interest is the spin qudit, or a multilevel quantum angular momentum degree of freedom. Spin qudits can provide advantages over their qubit counterparts for quantum sensing [28,29], enable quantum simulations of SU(d) magnetism [16,30–34], and offer unique capabilities for quantum computation and error correction [35–37]. In all cases, quantum state tomography is necessary to take full advantage of a spin qudit.[1]

The problem of qudit tomography is not new with an extensive literature on a variety of techniques [29,38–48]. However, most existing protocols either rely on infinite-dimensional representations of a quantum spin [29,41,49] or require the capability to perform essentially arbitrary operations on a qudit [42–48], generally resulting in tomographic protocols that can be highly inefficient or unachievable in practice. The protocols based on infinite-dimensional representations of a quantum spin have the advantage of reconstructing its state from measurements of spin projection onto different spatial axes, which are generally accessible with any spin qudit. Nonetheless, these protocols obfuscate the minimal requirements for performing full state tomography, provide no straightforward error bounds or guarantees of accuracy, and (with the notable exception of Ref. [41]) generally extract only a small fraction of the information contained in measurement data.

In this paper, we consider the task of performing spin qubit tomography using only measurements of spin projection onto different spatial axes. This sort of task was first considered in Ref. [38] as well as a few later works [39–41]. Specifically, Ref. [38] provided an explicit protocol for reconstructing a d-level spin qudit state from measurements of spin projection along $2d - 1$ axes, the minimum number necessary for full tomography of an arbitrary (possibly mixed) qudit state. However, the protocol in Ref. [38] involves a choice of a single (arbitrary) angle θ and provides no means for comparing different choices of θ, which may result in wildly different statistical errors (i.e., precision) in a reconstructed state. Other works provide insightful discussions into the problem of spin qudit tomography, but either: (i) require making assumptions about the qudit state in question [41] (making the tomographic protocol only valid for a restricted set of possible states), (ii) do not address the question of statistical error [39], or (iii) provide a measure of statistical error that is needlessly conservative and computationally demanding [40]. We address these shortcomings in this paper and identify remaining avenues for refining spin qudit tomography protocols.

---

*mika.perlin@gmail.com

[1]Note that the measurement of collective observables for quantum sensing or simulation can be recast as a single-spin tomography task.

In Sec. II, we introduce a set of qudit operators that are closely related to the spherical harmonics and which play a central role in our paper. We then map the quantum problem of spin qudit tomography onto the classical problem of signal recovery on the sphere in Sec. III, thereby providing an intuitive perspective on spin qudit tomography. In Sec. IV we provide *a priori* upper bounds and a posteriori estimates of the statistical error in a qudit state reconstructed from measurements of spin projection along a given set of r measurement axes. The capability to determine upper bounds on reconstruction error *a priori* motivates a simple randomized tomography protocol that we outline in Sec. V, and for which we numerically find that using more measurement axes yields substantial benefits that plateau after $r \approx 3d$. To facilitate the use of our protocols, we make all of our codes publicly available in Ref. [50], which also contains the best measurement axes we found for $d \leqslant 30$ and $r = 3d$.

II. POLARIZATION OPERATORS

We begin by introducing a set of qudit operators that are closely related to the spherical harmonics (in a sense that will be clarified below) and which play a central role in our paper. Consider a d-state spin qudit with total spin $s \equiv \frac{d-1}{2}$. The defining property of a spin qudit, distinguishing it from other qudits, is that it describes an angular momentum degree of freedom, which has specific implications for how a spin qudit should transform under the group SO(3) of rotations in three-dimensional (3D) space. Due to the central importance of these transformation rules for a spin qudit, we seek a basis of operators that transform nicely under 3D rotations.[2] One such basis is that of the *polarization operators* [51,52], defined by

$$T_{\ell m} \equiv \sqrt{\frac{2\ell+1}{2s+1}} \sum_{\mu,\nu=-s}^{s} \langle s\mu; \ell m | s\nu \rangle |\nu\rangle\langle\mu|, \quad (1)$$

where $|\mu\rangle$ is an eigenstate of the axial spin projection operator $S_z |\mu\rangle = \mu |\mu\rangle$; and $\langle s\mu; \ell m | s\nu \rangle$ is a Clebsch-Gordan coefficient that enforces $\ell \in \{0, 1, \ldots, d-1\}$ and $m \in \{-\ell, -\ell+1, \ldots, \ell\}$, such that there are d^2 polarization operators in total. For brevity, we will generally treat the value of d as constant but arbitrary throughout this paper, and we will suppress any explicit dependence of quantities or operators, such as $T_{\ell m}$ on d. The polarization operators are orthonormal with respect to the trace inner product and transform nicely under conjugation,

$$(T_{\ell m} | T_{\ell' m'}) = \delta_{\ell\ell'} \delta_{mm'}, \quad T_{\ell m}^\dagger = (-1)^m T_{\ell,-m}, \quad (2)$$

where for any $d \times d$ matrix $X = \sum_{\mu,\nu} X_{\mu\nu} |\mu\rangle\langle\nu|$ we define the d^2-component vector $|X\rangle \equiv \sum_{\mu,\nu} X_{\mu\nu} |\mu\nu\rangle$; $(X|$ is the conjugate transpose of $|X\rangle$, such that $(X|Y) = \text{tr}(X^\dagger Y)$; and $\delta_{kk'} \equiv 1$ if $k = k'$ and 0 otherwise. These properties of the polarization operators allow us to expand any density operator

ρ in the polarization operator basis as

$$\rho = \sum_{\ell=0}^{d-1} \sum_{m=-\ell}^{\ell} \rho_{\ell m} T_{\ell m}, \quad \rho_{\ell m} \equiv \langle T_{\ell m}^\dagger \rangle_\rho, \quad (3)$$

where $\langle X \rangle_\rho \equiv \text{tr}(\rho X) = (\rho|X)$ and $\rho^\dagger = \rho$ implies that $\rho_{\ell m}^* = (-1)^m \rho_{\ell,-m}$. The polarization operators can be interpreted in terms of an absorption process whereby $T_{\ell m} |\psi\rangle$ is (up to normalization) the state obtained after a spin-s state $|\psi\rangle$ absorbs a particle with total spin ℓ and spin projection m onto a fixed quantization axis. Similar to the complex spherical harmonics $Y_{\ell m}$, we will refer to ℓ as the *degree* and m as the *order* of $T_{\ell m}$.

The polarization operators are spherical tensor operators, whose degree is preserved under 3D rotations generated by the spin operators S_x, S_y, and S_z. Moreover, the degree-ℓ polarization operators $T_{\ell m}$ transform similarly to spin-ℓ particles and spherical harmonics $Y_{\ell m}$ under 3D rotations (see Appendix A). Specifically, for any triplet of angles $\boldsymbol{\omega} = (\alpha, \beta, \gamma)$, we can define the rotation operator,

$$R(\boldsymbol{\omega}) \equiv e^{-i\alpha S_z} e^{-i\beta S_y} e^{-i\gamma S_z}, \quad (4)$$

and expand rotated polarization operators as

$$T_{\boldsymbol{\omega}\ell m} \equiv R(\boldsymbol{\omega}) T_{\ell m} R(\boldsymbol{\omega})^\dagger = \sum_{n=-\ell}^{\ell} D_{mn}^\ell(\bar{\boldsymbol{\omega}})^* T_{\ell n}, \quad (5)$$

where $\bar{\boldsymbol{\omega}} = (\gamma, \beta, \alpha)$ is the reversal of $\boldsymbol{\omega}$, and

$$D_{mn}^\ell(\bar{\boldsymbol{\omega}}) \equiv \langle \ell m | R(\bar{\boldsymbol{\omega}}) | \ell n \rangle \quad (6)$$

are (Wigner) rotation matrix elements. For reasons that will become clear shortly, throughout this paper we will primarily consider rotations of the sphere that take the north pole to a point $\boldsymbol{v} = (\alpha, \beta)$ at azimuthal angle α and polar angle β. For ease of notation, we therefore define $R(\boldsymbol{v}) \equiv R(\alpha, \beta, 0)$, $T_{\boldsymbol{v}\ell m} \equiv T_{(\alpha,\beta,0),\ell m}$, and $D_{mn}^\ell(\boldsymbol{v}) \equiv D_{mn}^\ell(0, \beta, \alpha)$.

The polarization operators $T_{\ell m}$ share a connection to the spherical harmonics $Y_{\ell m}$ that goes beyond the rules for their transformation under 3D rotations. In fact, the phase-space representation of $T_{\ell m}$ is proportional to $Y_{\ell m}$. The phase-space representation of a spin qudit operator X assigns to each point $\boldsymbol{v}$ on the sphere the complex number,

$$X^{\text{PS}}(\boldsymbol{v}) \equiv \langle s_{\boldsymbol{v}} | X | s_{\boldsymbol{v}} \rangle, \quad (7)$$

where $|s_{\boldsymbol{v}}\rangle \equiv R(\boldsymbol{v})|s\rangle$ is the state of a spin qudit polarized along $\boldsymbol{v}$. This representation is faithful in the sense that X is uniquely determined by the phase-space values $X^{\text{PS}}(\boldsymbol{v})$ at all points $\boldsymbol{v}$ on the sphere. The transformation rules for polarization operators in Eq. (5) together with the fact that $\langle s | T_{\ell m} | s \rangle = 0$ unless $m = 0$ suffice to show that

$$T_{\ell m}^{\text{PS}}(\boldsymbol{v}) = c_\ell Y_{\ell m}(\boldsymbol{v}), \quad (8)$$

where the scalar c_ℓ simply enforces $(T_{\ell m} | T_{\ell m}) = 1$ (see Appendix A). The polarization operators $T_{\ell m}$ are, thus, a quantum analog of the spherical harmonics $Y_{\ell m}$ and play an important role in phase-space formalisms for spin qudits [53].

As a special case, the phase-space representation ρ^{PS} of a spin qudit state ρ is commonly known as its Husimi distribution. Performing tomography on an unknown qudit state ρ is,

---

[2] Technically speaking, we seek a basis of operators that transform as an irreducible representation of SO(3).

therefore, equivalent to reconstructing the unknown distribution ρ^{PS} on the sphere. In principle, the representation ρ^{PS} of a finite-dimensional qudit state ρ can be reconstructed from the values of $\rho^{\text{PS}}(\boldsymbol{v}) = \langle s_{\boldsymbol{v}}|\rho|s_{\boldsymbol{v}}\rangle$ at a finite number of points $\boldsymbol{v}$. In practice, the value $\langle s_{\boldsymbol{v}}|\rho|s_{\boldsymbol{v}}\rangle$ is determined by measuring spin projection along $\boldsymbol{v}$, which also provides measurement data on all spin projections $\langle \mu_{\boldsymbol{v}}|\rho|\mu_{\boldsymbol{v}}\rangle$ with $\mu \in \{s, s-1, \ldots, -s\}$ and $|\mu_{\boldsymbol{v}}\rangle \equiv R(\boldsymbol{v})|\mu\rangle$; one would like to make use of this additional data as well. We clarify the connection between the quantum problem of reconstructing ρ from spin projection measurements and the classical problem of reconstructing ρ^{PS} from its values of $\rho^{\text{PS}}(\boldsymbol{v})$ in the following section.

III. SPIN TOMOGRAPHY AS SIGNAL RECOVERY ON THE SPHERE

Our goal is to reconstruct an arbitrary state ρ of a spin qudit from measurements of spin projection onto different quantization axes. We are, thus, nominally restricted to measuring projectors $\Pi_{\boldsymbol{v}\mu} \equiv |\mu_{\boldsymbol{v}}\rangle\langle\mu_{\boldsymbol{v}}|$, where $|\mu_{\boldsymbol{v}}\rangle \equiv R(\boldsymbol{v})|\mu\rangle$ is a state with spin projection μ onto the measurement axis $\boldsymbol{v}$. For any fixed axis $\boldsymbol{v}$, the sets $\{\Pi_{\boldsymbol{v}\mu}\}$ and $\{T_{\boldsymbol{v}\ell,0}\}$ (i.e., all $T_{\boldsymbol{v}\ell m}$ with $m = 0$) are both complete bases for the space of operators that are diagonal in the basis $\{|\mu_{\boldsymbol{v}}\rangle\}$. Measuring the projectors $\{\Pi_{\boldsymbol{v}\mu}\}$ is, therefore, equivalent to measuring the polarization operators $\{T_{\boldsymbol{v}\ell,0}\}$, and provides data on the expectation values $\langle T_{\boldsymbol{v}\ell,0}\rangle_\rho$.

In order to reconstruct an arbitrary density operator ρ from the expectation values $\langle T_{\boldsymbol{v}\ell,0}\rangle_\rho$, we essentially need to find a set of coefficients $C_{\ell m k}(\boldsymbol{v})$ that would allow us to recover any matrix element $\rho_{\ell m}$ of ρ through

$$\rho_{\ell m}^* = \langle T_{\ell m}\rangle_\rho = \sum_{\boldsymbol{v},k} C_{\ell m k}(\boldsymbol{v})\langle T_{\boldsymbol{v}k,0}\rangle_\rho. \tag{9}$$

Expanding the rotated polarization operators $T_{\boldsymbol{v}k,0}$ into a sum of unrotated polarization operators $T_{\ell n}$ according to Eq. (5), we find that the recovery condition in Eq. (9) is satisfied when

$$T_{\ell m} = \sum_{\boldsymbol{v},k,n} C_{\ell m k}(\boldsymbol{v}) D_{0,n}^k(\boldsymbol{v})^* T_{kn}. \tag{10}$$

Orthogonality of the polarization operators then implies the decomposition $C_{\ell m k}(\boldsymbol{v}) = \delta_{\ell k} C_{\ell m}(\boldsymbol{v})$, and, in turn,

$$\sum_{\boldsymbol{v}} C_{\ell m}(\boldsymbol{v}) D_{0,n}^\ell(\boldsymbol{v})^* = \delta_{mn} \tag{11}$$

for all ℓ's.

In fact, the problem of finding suitable axes V and coefficients $C_{\ell m}(\boldsymbol{v})$ to satisfy Eq. (11) can be mapped onto the well-studied problem of signal recovery on the sphere (see Fig. 1) [54–57]. The signal recovery problem can be stated as follows: given a square-integrable function f on the sphere with the spherical harmonic expansion,

$$f(\boldsymbol{v}) = \sum_{\ell,m} f_{\ell m} Y_{\ell m}(\boldsymbol{v}), \tag{12}$$

where $f_{\ell m}$ are complex coefficients, find a set of points $V = \{\boldsymbol{v}\}$ and associated coefficients $\tilde{C}_{\ell m}(\boldsymbol{v})$ with which we can reconstruct f, or equivalently its coefficients $f_{\ell m}$, from knowl-
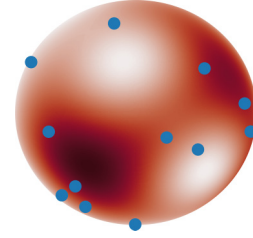

FIG. 1. Signal recovery on the sphere is the problem of reconstructing an unknown function f (red distribution) from its values $f(\boldsymbol{v})$ at specific points $\boldsymbol{v} \in V$ (blue dots) on the sphere. For almost all choices of V, reconstruction of f is possible if there are at least as many points in V as there are degrees of freedom in f.

edge of the function's value $f(\boldsymbol{v})$ at all points $\boldsymbol{v} \in V$; that is

$$f_{\ell m} = \sum_{\boldsymbol{v}} \tilde{C}_{\ell m}(\boldsymbol{v}) f(\boldsymbol{v}) = \sum_{\boldsymbol{v},k,n} \tilde{C}_{\ell m}(\boldsymbol{v}) Y_{kn}(\boldsymbol{v}) f_{kn}. \tag{13}$$

Reconstruction of functions with arbitrary coefficients $f_{\ell m}$ implies that

$$\sum_{\boldsymbol{v}} \tilde{C}_{\ell m}(\boldsymbol{v}) Y_{kn}(\boldsymbol{v}) = \delta_{\ell k}\delta_{mn}, \tag{14}$$

which is a stronger version of the condition that we found for the spin qudit tomography problem in Eq. (11). We will refer to Eq. (14) as the *full recovery problem* and Eq. (11) as the *reduced recovery problem*. Due to the fact that $D_{0,m}^\ell(\boldsymbol{v}) = \sqrt{\frac{4\pi}{2\ell+1}} Y_{\ell m}(\boldsymbol{v})$, any solution to the full recovery problem automatically solves the reduced recovery problem by setting $C_{\ell m}(\boldsymbol{v}) = \sqrt{\frac{2\ell+1}{4\pi}} \tilde{C}_{\ell m}(\boldsymbol{v})^*$. In principle, this mapping allows us to import a host of existing signal recovery algorithms [54–57] for the task of spin qudit tomography. In practice, spin qudits typically have only a modest dimension d, which allows for simpler and optimized tomography protocols that are practical despite worse scaling with d (see Sec. V). A natural avenue to develop better spin qudit tomography protocols would, therefore, be to build on the existing classical signal recovery algorithms, tailoring them to solve the reduced recovery problem in Eq. (11) rather than the full recovery problem in Eq. (14). We leave these developments to future work.

If the function f is *band-limited* at degree L, which is to say that $f_{\ell m} = 0$ for all $\ell \geqslant L$, then the full recovery problem in Eq. (14) is provably solvable with a suitable choice of $|V| = L^2$ points on the sphere [58,59]. The existence of these solutions to the full recovery problem, in turn, implies the existence of d^2 measurement axes that suffice to reconstruct arbitrary states of d-level spin qudit, whose possible states (or rather, phase-space representations) are band-limited at degree d. Moreover, for any fixed degree ℓ, finding solutions to the reduced recovery problem in Eq. (11) is equivalent to the recovery of a degree-ℓ function $f_\ell = \sum_m f_{\ell m} Y_{\ell m}$, which is provably possible with $|V| = 2\ell + 1$ samples [58]. In the case of spin qudit tomography, the degree ℓ takes a maximal value of $\ell_{\max} \equiv d - 1$, so state recovery requires as many measurement axes as there are polarization operators with degree $\ell_{\max}$, namely, $2\ell_{\max} + 1 = 2d - 1$.

IV. STATE RECONSTRUCTION ERROR

For the practically minded, proving the existence of solutions to a problem is less interesting than the exposition of a particular solution. On a high level, a spin qudit tomography protocol consists of: (i) selecting a set of measurement axes, (ii) collecting measurement data on spin projection onto these axes, and then (iii) processing the collected data to reconstruct the state of the spin qudit. Whereas step (ii) can involve a host of platform-dependent technical challenges, in the following sections we discuss the steps to take before and after collecting measurement data.

To this end, we begin by asking a question: what is a "good" choice of measurement axes? Intuitively, a good choice of axes should minimize the error with which one can reconstruct an unknown quantum state from associated measurement data. If we can quantify this intuition, then we can optimize over different choices of measurement axes to find a set that (approximately) minimizes the error in reconstructed states.

A set of measurement axes $V = \{v\}$ nominally induces a set of projectors $\{\Pi_{v\mu}\}$ that will be measured in an experiment. By a simple change of basis, measuring these projectors is equivalent to measuring the polarization operators $\{T_{v\ell,0}\}$. Flattening each $d \times d$ matrix $T_{v\ell,0}$ into the d^2-component column vector $|T_{v\ell,0})$, we construct the *measurement matrix*,

$$M_V \equiv \sum_{v,\ell} |v\ell\rangle (T_{v\ell,0}|. \qquad (15)$$

Here v and ℓ label a row of M_V, or, equivalently, label a standard ("one-hot") basis vector $|v\ell\rangle$ of a $(|V| \times d)$-dimensional vector space, and $(T_{v\ell,0}|$ is the conjugate transpose of $|T_{v\ell,0})$. A necessary and sufficient condition for V to allow for full state tomography is that the measured polarization operators $T_{v\ell,0}$, or, equivalently, the rows of M_V, span the entire (d^2-dimensional) space of operators on a d-level spin qudit. In this case M_V must be full rank with d^2 nonzero singular values. Indexing these singular values M_k^V and the corresponding (normalized) left singular vectors $x_k^V \equiv \sum_j x_{kj}^V |j\rangle$ by an integer $k \in \{1, 2, \ldots, d^2\}$, we can construct the orthonormal qudit operators,

$$Q_k^V \equiv \sum_j \left(q_{kj}^V\right)^* T_j, \quad q_{kj}^V \equiv \frac{x_{kj}^V}{M_k^V}, \qquad (16)$$

where for shorthand we use a combined index $j = (v, \ell)$ to specify both a measurement axis v and a degree ℓ, which identify the polarization operator $T_j \equiv T_{v\ell,0}$. These operators allow us to expand any state ρ of a d-level spin qudit in the form

$$\rho = \sum_{k=1}^{d^2} \rho_k^V Q_k^V, \quad \rho_k^V \equiv \langle Q_k^{V\dagger} \rangle_\rho. \qquad (17)$$

Given empirical estimates $\tilde{T}_j$ of the expectation values $\langle T_j \rangle_\rho$, an empirical estimate $\tilde{\rho}_V$ of ρ is then

$$\tilde{\rho}_V \equiv \sum_k \tilde{\rho}_k^V Q_k^V, \qquad (18)$$

where, using the fact that $T_j = T_j^\dagger$ (because they are diagonal polarization operators with degree $m = 0$),

$$\tilde{\rho}_k^V \equiv \sum_j q_{kj}^V \tilde{T}_j \approx \sum_j q_{kj}^V \langle T_j \rangle_\rho = \langle Q_k^{V\dagger} \rangle_\rho = \rho_k^V. \qquad (19)$$

The measurement matrix M_V allows us to make concrete statements about the statistical error between the empirical estimate $\tilde{\rho}_V$ and the true state ρ. Assume, for example, that the estimates $\tilde{T}_j$ are equal to $\langle T_j \rangle_\rho$ up to uncorrelated noise with variance no greater than ϵ^2,

$$\tilde{T}_j = \langle T_j \rangle_\rho + \epsilon_j, \quad \langle\langle \epsilon_j \epsilon_{j'} \rangle\rangle \leqslant \epsilon^2 \delta_{jj'}. \qquad (20)$$

Here $\{\epsilon_j\}$'s are independent random variables, and we use the double brackets $\langle\langle \cdot \rangle\rangle$ to denote statistical averaging over experimental trials that estimate $\langle T_j \rangle_\rho$. In this case, the mean squared error with which $\tilde{\rho}_k^V$ approximates ρ_k^V is

$$\langle\langle |\tilde{\rho}_k^V - \rho_k^V|^2 \rangle\rangle = \langle\langle (\tilde{\rho}_k^V - \rho_k^V)^* (\tilde{\rho}_k^V - \rho_k^V) \rangle\rangle \qquad (21)$$

$$= \sum_{j,j'} \left(q_{kj}^V\right)^* q_{kj'}^V \langle\langle \epsilon_j \epsilon_{j'} \rangle\rangle \qquad (22)$$

$$\leqslant \sum_j |q_{kj}^V|^2 \epsilon^2 = \left(\frac{\epsilon}{M_k^V}\right)^2. \qquad (23)$$

Using the fact that the operators Q_k^V are orthonormal, we can, therefore, bound the mean squared (Euclidean) distance between $\tilde{\rho}_V$ and ρ as

$$\mathcal{E}_V(\rho)^2 \equiv \langle\langle \|\tilde{\rho}_V - \rho\|^2 \rangle\rangle \leqslant \epsilon^2 \mathcal{S}_V^2, \qquad (24)$$

where $\|X\|^2 \equiv (X|X) = \text{tr}(X^\dagger X)$ is the squared (Euclidean, Frobenius, or Hilbert-Schmidt) norm of X, and the *classical error scale* $\mathcal{S}_V$ is defined by

$$\mathcal{S}_V^2 \equiv \sum_k \left(M_k^V\right)^{-2} = \|M_V^{-1}\|^2, \qquad (25)$$

where M_V^{-1} is the left inverse of M_V, satisfying $M_V^{-1} M_V = 1$. We refer to the error scale $\mathcal{S}_V$ as classical because the bound in Eq. (24) applies in the presence of classical sources of measurement error. Note that the classical error scale $\mathcal{S}_V$ diverges if the measurement matrix M_V is singular, which indicates that measuring spin projections along all axes in V does not provide sufficient information to reconstruct arbitrary quantum states.

Computing the classical error scale $\mathcal{S}_V$ and estimates $\tilde{\rho}_k^V \approx \rho_k^V$ requires building the measurement matrix M_V and computing its singular value decomposition. The complexity of this task can be greatly reduced by the fact that the degree ℓ of a polarization operator $T_{\ell m}$ is preserved under rotations, which implies that the unitary,

$$U \equiv \sum_{\ell=0}^{d-1} \sum_{m=-\ell}^{\ell} |T_{\ell m}\rangle \langle \ell m|, \qquad (26)$$

with vectors $|T_{\ell m}\rangle$ in a column indexed by integers (ℓ, m) block diagonalizes the measurement matrix into d blocks indexed by the degree ℓ,

$$M_V U = \sum_{\ell=0}^{d-1} |\ell\rangle \langle \ell| \otimes M_{V\ell}, \qquad (27)$$
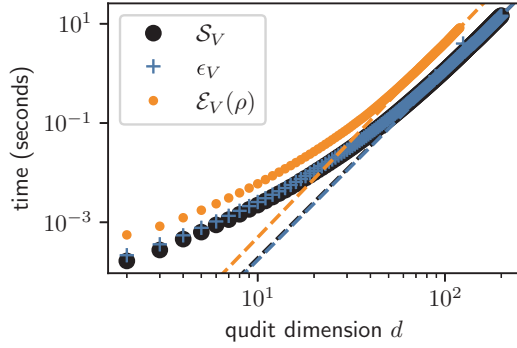
FIG. 2. Serial runtime to compute $\mathcal{S}_V$, ϵ_V, or $\mathcal{E}_V(\rho)$ with $|V| = 2d - 1$ randomly chosen measurement axes and a randomly chosen qudit state ρ. Each point is an average over 10^3 calculations or 5 min of runtime, whichever comes first. These results do not count fixed runtimes to precompute quantities that can be recycled for every new choice of V and ρ. Dashed lines show fits to a runtime $t = cd^\alpha$ for the 20 largest values of d, finding $\alpha \approx 3.8 \pm 0.1$.

where the $|V| \times (2\ell + 1)$-sized blocks are

$$M_{V\ell} \equiv \sum_{\boldsymbol{v},m} |\boldsymbol{v}\rangle\, (T_{\boldsymbol{v}\ell,0}|T_{\ell m})\, \langle m| = \sum_{\boldsymbol{v},m} D^\ell_{0,m}(\boldsymbol{v})|\boldsymbol{v}\rangle\langle m|. \quad (28)$$

Here $D^\ell_{0,m}(\boldsymbol{v})$ is a Wigner rotation matrix element, defined in Eq. (6). As the singular values of M_V are invariant under unitary transformations, it follows that:

$$\mathcal{S}^2_V = \sum_\ell \mathcal{S}^2_{V\ell}, \quad \mathcal{S}^2_{V\ell} \equiv \left\| M^{-1}_{V\ell} \right\|^2, \quad (29)$$

where $M^{-1}_{V\ell}$ is the left inverse of $M_{V\ell}$. Constructing the block $M_{V\ell}$ and computing its singular value decomposition takes at most $O(|V|d^2)$ time. If we assume that $|V| \sim d$, then computing the classical error scale $\mathcal{S}_V$ takes $O(d^4)$ serial or $O(d^3)$ parallel runtime (see Fig. 2).

The assumption that observables can be estimated up to uncorrelated noise with maximal variance ϵ^2, summarized by Eq. (20), is reasonable when the measurement error is dominated by classical sources of experimental noise. However, this assumption breaks down when measurement error is limited by fundamental quantum shot noise (i.e., finite sampling error). We relax the assumption of Eq. (20) in Appendix B, where we instead assume that $\tilde{\rho}_V$ is built from n independent measurements of spin projection along every axis $\boldsymbol{v} \in V$ with shot noise as the dominant source of error. In this case, the constraints that $\mathrm{tr}(\rho) = 1$ and $\langle \Pi_{\boldsymbol{v}\mu} \rangle_\rho \geqslant 0$ allow us to bound the mean-squared distance between $\tilde{\rho}_V$ and ρ as

$$\mathcal{E}_V(\rho)^2 < \frac{\epsilon^2_V}{n}, \quad \epsilon^2_V \equiv \sum_\ell \Gamma^2_\ell \mathcal{S}^2_{V\ell}, \quad (30)$$

where the *quantum error scale* ϵ_V is defined in terms of the spectral range of $T_{\ell,0}$,

$$\Gamma_\ell \equiv \frac{\max_\mu t_{\ell\mu} - \min_\mu t_{\ell\mu}}{2}, \quad t_{\ell\mu} \equiv \langle \mu | T_{\ell,0} | \mu \rangle. \quad (31)$$

If d is even or ℓ is odd, then $\Gamma_\ell = \max_\mu t_{\ell\mu}$. For comparison with the classical error bound in Eq. (24), we note that $\epsilon^2_V < \mathcal{S}^2_V / 2$, so the previous bound still holds with the replacement $\epsilon^2 \to 1/2n$. The factors Γ^2_ℓ are quick to compute and can be

recycled for every new choice of axes V, so the complexity of computing ϵ_V is the same as that of $\mathcal{S}_V$ (see Fig. 2).

Although straightforward to compute, the bound in Eq. (30) is not tight as it is acquired by bounding the statistical error $\epsilon_{\boldsymbol{v}\ell}$ in the empirical estimate $\tilde{T}_{\boldsymbol{v}\ell,0}$ of $\langle T_{\boldsymbol{v}\ell,0} \rangle_\rho$ by $\langle\langle \epsilon^2_{\boldsymbol{v}\ell} \rangle\rangle \leqslant \Gamma^2_\ell$. The individual bounds on $\langle\langle \epsilon^2_{\boldsymbol{v}\ell} \rangle\rangle$ for each axis $\boldsymbol{v}$ and degree ℓ are tight, but these bounds cannot all be achieved simultaneously. There is, therefore, still room for improvement on the bound in Eq. (30) by maximizing $\mathcal{E}_V$ over the set of all physical qudit states ρ. We discuss this maximization problem in Appendix C but leave its full solution to future work. We also note that the reconstruction error bound in Eq. (30) obeys the "standard quantum limit" of $\sim 1/n$ scaling in the number of measurements. In principle, this scaling can be improved to $\sim 1/n^2$ by preparing and measuring entangled copies of many qudits [60].

The error scales $\mathcal{S}_V$ and ϵ_V provide pessimistic upper bounds on statistical error, which can be calculated without prior knowledge of the true qudit state ρ. The actual error in the reconstruction $\tilde{\rho}_V$ of a particular state ρ may be considerably smaller and may depend on ρ itself. Written out in full, the mean-squared distance between $\tilde{\rho}_V$ and ρ is (see Appendix B)

$$\mathcal{E}_V(\rho)^2 = \sum_{\boldsymbol{v},\boldsymbol{w},\ell} \langle \boldsymbol{v} | \left(M^{-1}_{V\ell} \right)^\dagger M^{-1}_{V\ell} | \boldsymbol{w} \rangle \langle\langle \epsilon_{\boldsymbol{v}\ell} \epsilon_{\boldsymbol{w}\ell} \rangle\rangle. \quad (32)$$

The covariances $\langle\langle \epsilon_{\boldsymbol{v}\ell} \epsilon_{\boldsymbol{w}\ell} \rangle\rangle$ are generally determined by the sources of measurement error in any given experiment but will typically satisfy $\langle\langle \epsilon_{\boldsymbol{v}\ell} \epsilon_{\boldsymbol{w}\ell} \rangle\rangle = \delta_{\boldsymbol{v}\boldsymbol{w}} \langle\langle \epsilon^2_{\boldsymbol{v}\ell} \rangle\rangle$ because measurements along $\boldsymbol{v}$ are independent of measurements along $\boldsymbol{w}$. If the measurement error is limited by shot noise, then (see Appendix C)

$$\mathcal{E}_V(\rho)^2 \overset{\mathrm{SNL}}{=} \frac{1}{n} \sum_\ell [\langle \chi_{V\ell} | \rho_\ell \rangle - \langle \rho_\ell | \mathcal{N}_{V\ell} | \rho_\ell \rangle], \quad (33)$$

where $\overset{\mathrm{SNL}}{=}$ indicates equality in the "shot-noise-limited" regime; $|\rho_\ell\rangle \equiv \sum_m \rho_{\ell m} |m\rangle$ is a vector of the polarization operator components $\rho_{\ell m}$ of ρ, defined in Eq. (3); and the matrix $\mathcal{N}_{V\ell}$ and vector $|\chi_{V\ell}\rangle$ are defined below. Although the true shot-noise-limited error in $\tilde{\rho}_V$ cannot be known exactly without knowing ρ, this error can be estimated *a posteriori* by $\mathcal{E}_V(\rho) \approx \mathcal{E}_V(\tilde{\rho}_V)$. After constructing an estimate $\tilde{\rho}_V$ of ρ, the complexity of computing the error $\mathcal{E}_V(\tilde{\rho}_V)$ from Eq. (33) is the same as that of computing $\mathcal{S}_V$ or ϵ_V (see Fig. 2).

We now define $\mathcal{N}_{V\ell}$ and $|\chi_{V\ell}\rangle$ for the sake of completion, but note that these definitions can be skipped without consequence for the remaining discussions in this paper. The matrix $\mathcal{N}_{V\ell}$ is

$$\mathcal{N}_{V\ell} \equiv M^\dagger_{V\ell} \mathrm{diag}\left[\left(M^{-1}_{V\ell} \right)^\dagger M^{-1}_{V\ell} \right] M_{V\ell}, \quad (34)$$

where $\mathrm{diag}[X]$ sets the off-diagonal parts of X to zero. The vector $|\chi_{V\ell}\rangle \equiv \sum_m \chi^V_{\ell m} |m\rangle$ is defined by

$$\chi^V_{LM} \equiv \sum_\ell (\mathcal{N}_{V\ell} | \mathcal{D}_M | \tilde{g}_{L\ell}), \quad (35)$$

$$\tilde{g}_{L\ell} \equiv \sum_{m,m'} (T_{L,m+m'} | T^\dagger_{\ell m} T_{\ell m'}) |m\rangle\langle m'|, \quad (36)$$

$$\mathcal{D}_M \equiv \sum_{m,m'} \delta_{M,m'-m} |mm'\rangle\langle mm'|. \quad (37)$$

Here $\tilde{g}_{L\ell}$ is essentially a matrix of structure constants for the polarization operator algebra (see Appendix D), and $\mathcal{D}_M$ simply picks off the Mth diagonal of the matrix it acts on.

V. TOMOGRAPHY PROTOCOL

The ability to certify a statistical error bound on the empirical estimate $\tilde{\rho}_V$ of an unknown quantum state ρ motivates the following protocol for spin qudit tomography:

(1) Select a random set of measurement axes V by uniformly sampling points on the sphere,[3] and use any standard minimization algorithm to optimize the $2|V|$ parameters in V (two angles for each point $\boldsymbol{v} \in V$) by minimizing the quantum error scale ϵ_V in Eq. (30). If $|V|$ is too large for such optimization, you can simply generate many sets of random measurement axes and then choose the set with the smallest quantum error scale ϵ_V. Note that computing the error scale ϵ_V requires for each $\ell \in \{0, 1, \dots, d-1\}$, constructing the measurement matrix $M_{V\ell}$ in Eq. (28) and computing its singular value decomposition. Save all measurement matrix data associated with the final measurement axes V for later use.

(2) For each axis $\boldsymbol{v} \in V$, make n measurements of spin projection, and set $\tilde{\Pi}_{\boldsymbol{v}\mu} \approx \langle \Pi_{\boldsymbol{v}\mu} \rangle_\rho$ to the fraction of times in which the measurement outcome was μ.

(3) Use the estimates $\tilde{\Pi}_{\boldsymbol{v}\mu}$ of $\langle \Pi_{\boldsymbol{v}\mu} \rangle_\rho$ to compute estimates of $\langle T_{\boldsymbol{v}\ell,0} \rangle_\rho$,

$$\tilde{T}_{\boldsymbol{v}\ell,0} \equiv \sum_\mu \langle \mu | T_{\ell,0} | \mu \rangle \tilde{\Pi}_{\boldsymbol{v}\mu}, \tag{38}$$

where the matrix elements of $T_{\ell,0}$ are provided in Eq. (1).

(4) Denoting the nonzero singular values of $M_{V\ell}$ by $M_{\ell k}^V$ and the corresponding left singular vectors by $\boldsymbol{x}_{\ell k}^V = \sum_{\boldsymbol{v}} x_{\ell k \boldsymbol{v}}^V |\boldsymbol{v}\rangle$, compute the operators and coefficients,

$$Q_{\ell k} \equiv \frac{1}{M_{\ell k}^V} \sum_{\boldsymbol{v}} (x_{\ell k \boldsymbol{v}}^V)^* T_{\boldsymbol{v}\ell,0}, \tag{39}$$

$$\tilde{\rho}_{\ell k}^V \equiv \frac{1}{M_{\ell k}^V} \sum_{\boldsymbol{v}} x_{\ell k \boldsymbol{v}}^V \tilde{T}_{\boldsymbol{v}\ell,0}, \tag{40}$$

and combine them into the estimate,

$$\tilde{\rho}_V = \sum_{\ell, k} \tilde{\rho}_{\ell k}^V Q_{\ell k} \approx \rho. \tag{41}$$

The expected reconstruction error in $\tilde{\rho}_V$, or its root-mean-square distance from ρ, is provided by Eq. (32). If measurement error is shot-noise-limited, then the error in $\tilde{\rho}_V$ is approximately $\mathcal{E}_V(\rho) \approx \mathcal{E}_V(\tilde{\rho}_V)$ and can be computed from Eq. (33). If $\tilde{\rho}_V$ has negative eigenvalues, its distance from ρ can be reduced with maximum-likelihood corrections [4], which will additionally guarantee that $\tilde{\rho}_V$ satisfies all requirements for being a physical state.

The tomography protocol outlined above leaves open the question of *how many* measurement axes to use. Although $2d - 1$ measurement axes may be sufficient to perform full

---

[3]To sample a point (α, β) from the uniform distribution on the sphere (with azimuthal angle α and polar angle β), you can sample a point $(a, b) \in [0, 1] \times [0, 1]$ from the uniform distribution on the unit square and then set $\alpha = 2\pi a$ and $\beta = \arccos(1 - 2b)$.
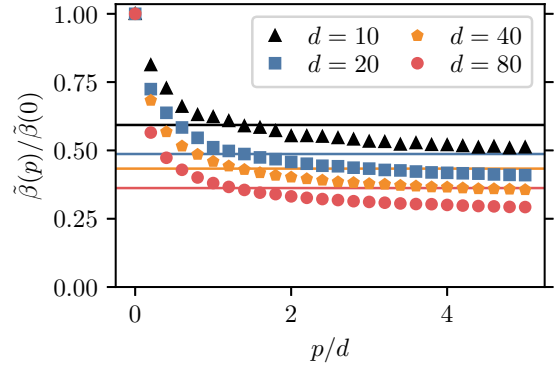

FIG. 3. Empirical measurement-adjusted error scales $\tilde{\beta}(p)$ with p excess measurement axes, determined by minimizing over 10^3 choices of measurement axes V or 5 min. of runtime (for each p), whichever comes first. Marker and color indicate the qudit dimension d. The rapid initial drop in $\tilde{\beta}(p)$ implies that using more measurement axes can substantially lower the upper bound on the reconstruction error provided in Eq. (30) and that these benefits plateau after $p \approx d$. Horizontal reference lines mark the smallest measurement-adjusted error scales $\min_\theta \beta_\theta / \tilde{\beta}(0)$ achievable with the method in Ref. [38], which is parametrized by an arbitrarily chosen angle θ.

state tomography, this is not necessarily the best choice of $|V|$. Increasing the number of measurement axes generally decreases the quantum error scale ϵ_V but comes at the cost of having to estimate more observables. At a fixed total number of measurements, increasing $|V|$ reduces the number of measurements n devoted to each axis $\boldsymbol{v} \in V$. This trade-off begs the question: How should one choose the number of measurement axes $|V|$?

The reconstruction error bound in Eq. (30) nominally provides a straightforward answer: at a fixed total number of measurements $N = n|V|$, the number of measurement axes should be chosen to minimize the (squared) reconstruction error $\mathcal{E}_V(\rho)^2 < \epsilon_V^2 / n \propto \epsilon_V^2 |V|$. We, therefore, consider the measurement-adjusted error scale $\beta(p)$ defined by

$$\beta(p)^2 \equiv \min_V \{ \epsilon_V^2 |V| : |V| = 2d - 1 + p \}, \tag{42}$$

where p is the number of "extra" measurement axes exceeding $2d - 1$. Although we cannot minimize over all suitable choices of measurement axes V to compute $\beta(p)$, we can compute an empirical upper bound $\tilde{\beta}(p) \geqslant \beta(p)$ by minimizing over a large number of randomly chosen V. Figure 3 shows the results of such empirical minimization where we find that $\tilde{\beta}(p)$ drops substantially with p before plateauing at $p \approx d$, after which there are only minor benefits to using more measurement axes. In the interest of reducing experimental complexity as well as the runtime of our randomized tomography protocol, which grow linearly in $|V|$, we, therefore, conclude that this protocol should be performed with $|V| \approx 3d$ measurement axes. We provide the best measurement axes that we found for a randomized tomography protocol with $d \leqslant 30$ and $|V| = 3d$ in Ref. [50].

For reference, Fig. 3 also shows the smallest measurement-adjusted error scales β_θ achievable with the method of Ref. [38], which is comparable to those achieved with our

randomized protocol at $|V| \approx 3d$. The method of Ref. [38] requires choosing an angle θ, namely, the polar angle of all measurement axes, and provides no prescription for making this choice. We, therefore, find the optimal choice of θ by minimizing the error scale β_θ over all θ (see Appendix E) and show $\min \beta_\theta / \tilde{\beta}(0)$ in Fig. 3. Empirically, we find that the optimal angle for the method of Ref. [38] is $\theta_{\text{opt}} \approx \frac{\pi}{2}(1 - \frac{1}{1.34d})$ (see Appendix E), which approaches $\pi/2$ as $d \to \infty$. However, the error scale $\beta_{\pi/2} = \infty$, reflecting the fact that full state tomography is impossible with measurement axes lying in a single plane. The method of Ref. [38], therefore, requires extremely careful fine-tuning of measurement axis orientations for large spin dimensions. For this reason, we expect our randomized tomography protocol to be more robust to errors in axis orientation. We leave a detailed analysis of robustness to errors in axis orientation and the effect of these errors on state reconstruction to future work.

As a final point, we note that any information about an unknown qudit state ρ, obtained from prior knowledge or preliminary measurement data, can be used to construct tailored or adaptive measurement protocols [5,7,61,62] that are more efficient in terms of the number of measurements required to estimate ρ to a fixed precision. We leave the development of tailored and adaptive measurement protocols to future work as well.

APPENDIX A: ROTATING POLARIZATION OPERATORS

Denoting the state of a spin-s particle with spin projection μ onto a quantization axis by $|s\mu\rangle$, we define

$$S_z \equiv \sum_{\mu=-s}^{s} \mu |s\mu\rangle\langle s\mu|,$$

$$S_\pm \equiv \sum_{\mu=-s}^{s} \sqrt{s(s+1) - \mu(\mu \pm 1)} |s, \mu \pm 1\rangle\langle s\mu|, \quad (A1)$$

as well as

$$S_x \equiv \frac{1}{2}(S_+ + S_-), \quad S_y \equiv -\frac{i}{2}(S_+ - S_-), \quad \mathbf{S} \equiv (S_x, S_y, S_z). \quad (A2)$$

The spin vector $\mathbf{S}$ generates rotations of a spin-s system in 3D space. Specifically, the operator $e^{-i\theta \mathbf{S}\cdot\hat{\mathbf{n}}}$ rotates a spin-s system by an angle θ about the unit vector $\hat{\mathbf{n}}$.

Observing that $S_z = T_{1,0}$ and $S_\pm \propto T_{1,\pm 1}$, we can use the operator product expansion of the polarization operators (see Appendix D), the properties of Clebsch-Gordan coefficients, the properties of Wigner 6-j symbols and a computer algebra

system to simplify the commutators,

$$[S_z, T_{\ell m}] = m T_{\ell m},$$
$$[S_\pm, T_{\ell m}] = \sqrt{\ell(\ell+1) - m(m \pm 1)} \, T_{\ell, m \pm 1}, \quad (A3)$$

which implies that $T_{\ell m}$ is a spherical tensor operator, whose degree degree ℓ is preserved under rotations generated by $\mathbf{S}$. Moreover, by comparing Eqs. (A1) and (A3) we see that the polarization operators $T_{\ell m}$ transform identically to spin-ℓ particles under the (adjoint) action of the spin operators S_z and $S_\pm$. For any triplet of angles $\boldsymbol{\omega} = (\alpha, \beta, \gamma)$, we can, therefore, define the rotation operator,

$$R(\boldsymbol{\omega}) \equiv e^{-i\alpha S_z} e^{-i\beta S_y} e^{-i\gamma S_z}, \quad (A4)$$

and expand rotated polarization operators as

$$T_{\boldsymbol{\omega}\ell m} \equiv R(\boldsymbol{\omega}) T_{\ell m} R(\boldsymbol{\omega})^\dagger = \sum_{n=-\ell}^{\ell} D_{mn}^\ell(\bar{\boldsymbol{\omega}})^* T_{\ell n}, \quad (A5)$$

where $\bar{\boldsymbol{\omega}} = (\gamma, \beta, \alpha)$ is the reversal of $\boldsymbol{\omega}$, and

$$D_{mn}^\ell(\bar{\boldsymbol{\omega}}) \equiv \langle \ell m | R(\bar{\boldsymbol{\omega}}) | \ell n \rangle = (T_{\ell n} | R(\boldsymbol{\omega}) \otimes R(\boldsymbol{\omega})^* | T_{\ell m})^*$$
$$= (T_{\ell m} | R(-\bar{\boldsymbol{\omega}}) \otimes R(-\bar{\boldsymbol{\omega}})^* | T_{\ell n}) \quad (A6)$$

are matrix elements of the rotation operator $R(\boldsymbol{\omega})$ for spin-ℓ particles.

For any angle doublet $\boldsymbol{v} = (\alpha, \beta)$, we define $R(\boldsymbol{v}) \equiv R(\alpha, \beta, 0)$ and $D_{mn}^\ell(\boldsymbol{v}) = D_{mn}^\ell(0, \beta, \alpha)$ for shorthand. The transformation rules in Eq. (A5) imply that we can expand the phase-space representation of $T_{\ell m}$ as

$$T_{\ell m}^{\text{PS}}(\boldsymbol{v}) \equiv \langle s_{\boldsymbol{v}} | T_{\ell m} | s_{\boldsymbol{v}} \rangle = \langle s | R(\boldsymbol{v})^\dagger T_{\ell m} R(\boldsymbol{v}) | s \rangle$$
$$= D_{0,m}^\ell(\boldsymbol{v}) \langle s | T_{\ell,0} | s \rangle, \quad (A7)$$

where

$$\langle s | T_{\ell,0} | s \rangle = \sqrt{\frac{2\ell+1}{2s+1}} \, \langle ss; \ell, 0 | ss \rangle$$
$$= \sqrt{\frac{2\ell+1}{2s+\ell+1} \left(\frac{(2s)!}{(2s+\ell)!}\right)\left(\frac{(2s)!}{(2s-\ell)!}\right)}, \quad (A8)$$

and the properties of the rotation matrix elements D_{mn}^ℓ imply that

$$D_{0,m}^\ell(\boldsymbol{v}) = \sqrt{\frac{4\pi}{2\ell+1}} \, Y_{\ell m}(\boldsymbol{v}), \quad (A9)$$

so

$$T_{\ell m}^{\text{PS}}(\boldsymbol{v}) = \sqrt{\frac{4\pi}{2s+\ell+1} \left(\frac{(2s)!}{(2s+\ell)!}\right)\left(\frac{(2s)!}{(2s-\ell)!}\right)} \, Y_{\ell m}(\boldsymbol{v}). \quad (A10)$$

In this way, the polarization operators are a quantum analog of the spherical harmonics.

APPENDIX B: AN IMPROVED RECONSTRUCTION ERROR BOUND

In Sec. IV of the main text, we provided a reconstruction error bound using the assumption of Eq. (20), namely, that expectation values derived from spin projection measurements can be estimated up to uncorrelated errors with

maximal variance ϵ^2. This assumption is reasonable if the measurement error is dominated by experimental sources of noise, and it yields a simple derivation of the reconstruction bound in Eq. (24). Nonetheless, there are two problems with the assumption of Eq. (20): (i) there is no *a priori* guarantee for the value of ϵ, which must be inferred from experimental outcomes, and (ii) the assumption that all measurement errors are uncorrelated is unjustified (and generally false). Here, we relax the assumption of Eq. (20) and derive an explicit error bound in terms of the qudit dimension d and the number of spin projection measurements made along every measurement axis.

To this end, we fix a particular set of measurement axes V and consider performing n measurements of spin projection along every axis $\boldsymbol{v} \in V$ for a total of $N = |V|n$ measurements. Such a procedure is equivalent to making N local measurements of the N-fold product state $\rho^{\otimes N}$. For convenience, we index the tensor factors of $\rho^{\otimes N}$ by the integers (i, j) with $i \in \{1, 2, \ldots, |V|\}$ specifying a measurement axis $\boldsymbol{v}_i \in V$ and $j \in \{1, 2, \ldots, n\}$ specifying the copy of ρ prepared for the jth measurement spin projection along a particular axis. We then define the projectors $\Pi_{i\mu} \equiv |\mu_{\boldsymbol{v}_i}\rangle\langle\mu_{\boldsymbol{v}_i}|$ onto single-qudit states $|\mu_{\boldsymbol{v}_i}\rangle$ with definite spin projection μ along axis $\boldsymbol{v}_i \in V$ and define $\Pi_{i\mu}^j$ to be an N-qudit operator with $\Pi_{i\mu}$ on the (i, j)-th tensor factor and the identity elsewhere. We denote the experimental outcome of measuring $\Pi_{i\mu}$ in the (i, j)-th copy of ρ by $\tilde{\Pi}_{i\mu}^j \in \{0, 1\}$. In other words, $\tilde{\Pi}_{i\mu}^j$ is the "single-shot estimate" of $\Pi_{i\mu}$, with $\tilde{\Pi}_{i\mu}^j = 1$ if outcome μ was observed on the (i, j)-th experimental trial, and $\tilde{\Pi}_{i\mu}^j = 0$ otherwise. An empirical estimate of the expectation value $\langle\Pi_{i\mu}\rangle_\rho$ is provided by the fraction of times that outcome μ was observed when measuring spin projection along axis $\boldsymbol{v}_i$, that is

$$\tilde{\Pi}_{i\mu} \equiv \frac{1}{n}\sum_{j=1}^n \tilde{\Pi}_{i\mu}^j \approx \frac{1}{n}\sum_{j=1}^n \mathrm{tr}\left(\rho^{\otimes N}\Pi_{i\mu}^j\right) = \mathrm{tr}(\rho\Pi_{i\mu}). \quad \text{(B1)}$$

For reasons that will be clarified shortly, it will be useful to think of $\tilde{\Pi}_{i\mu}$ as an empirical estimate of $\langle\bar{\Pi}_{i\mu}\rangle_{\rho^{\otimes N}}$, where

$$\bar{\Pi}_{i\mu} \equiv \frac{1}{n}\sum_{j=1}^n \Pi_{i\mu}^j \quad \text{(B2)}$$

is the average of $\Pi_{i\mu}$ applied to all copies of ρ for which spin projection is measured along the axis $\boldsymbol{v}_i$. Equation (B1) implies that

$$\tilde{\Pi}_{i\mu} \approx \langle\bar{\Pi}_{i\mu}\rangle_{\rho^{\otimes N}} = \langle\Pi_{i\mu}\rangle_\rho. \quad \text{(B3)}$$

1. Errors in the spin-projection basis

Finite sampling error (i.e., shot noise) generally induces statistical error ϵ_O into the empirical estimate $\tilde{O}$ of an observable O,

$$\epsilon_O \equiv \tilde{O} - \langle O\rangle, \quad \text{(B4)}$$

where the single brackets $\langle\cdot\rangle$ denote an expectation value with respect to the measured quantum state. On average, this

statistical error will be zero, which is to say that

$$\langle\!\langle\epsilon_O\rangle\!\rangle = \langle\!\langle\tilde{O} - \langle O\rangle\rangle\!\rangle = \langle O - \langle O\rangle\rangle = 0, \quad \text{(B5)}$$

where the double brackets $\langle\!\langle\cdot\rangle\!\rangle$ denote statistical averaging over experimental trials that estimate $\langle O\rangle$. However, the covariance between statistical errors ϵ_O and ϵ_Q on the empirical estimates $\tilde{O}$ and $\tilde{Q}$ of observables O and Q is

$$\langle\!\langle\epsilon_O\epsilon_Q\rangle\!\rangle = \langle(\tilde{O} - \langle O\rangle)(\tilde{Q} - \langle Q\rangle)\rangle = \langle(O - \langle O\rangle)(Q - \langle Q\rangle)\rangle$$
$$= \langle OQ\rangle - \langle O\rangle\langle Q\rangle. \quad \text{(B6)}$$

In the context of spin qudit tomography, we can, therefore, define the statistical error,

$$\epsilon_{i\mu} \equiv \tilde{\Pi}_{i\mu} - \langle\Pi_{i\mu}\rangle_\rho = \tilde{\Pi}_{i\mu} - \langle\bar{\Pi}_{i\mu}\rangle_{\rho^{\otimes N}}, \quad \text{(B7)}$$

in the empirical estimate of $\langle\Pi_{i\mu}\rangle_\rho$ and use Eq. (B2) to expand

$$\langle\!\langle\epsilon_{i\mu}\epsilon_{i'\mu'}\rangle\!\rangle = \langle\bar{\Pi}_{i\mu}\bar{\Pi}_{i'\mu'}\rangle_{\rho^{\otimes N}} - \langle\bar{\Pi}_{i\mu}\rangle_{\rho^{\otimes N}}\langle\bar{\Pi}_{i'\mu'}\rangle_{\rho^{\otimes N}}$$
$$= \frac{1}{n^2}\sum_{j,j'=1}^n \left[\langle\Pi_{i\mu}^j\Pi_{i'\mu'}^{j'}\rangle_{\rho^{\otimes N}} - \langle\Pi_{i\mu}^j\rangle_{\rho^{\otimes N}}\langle\Pi_{i'\mu'}^{j'}\rangle_{\rho^{\otimes N}}\right]. \quad \text{(B8)}$$

If $(i, j) \neq (i', j')$, then $\Pi_{i\mu}^j$ and $\Pi_{i'\mu'}^{j'}$ address different tensor factors of the product state $\rho^{\otimes N}$, so the expectation value of their product factorizes due to the fact that $\mathrm{tr}[(A \otimes B)(A' \otimes B')] = \mathrm{tr}(AA')\mathrm{tr}(BB')$. This factorization can also be seen as a consequence of the fact that if $(i, j) \neq (i', j')$, then $\Pi_{i\mu}^j$ and $\Pi_{i'\mu'}^{j'}$ are "spatially separated" on $\rho^{\otimes N}$, which means that their expectation values cannot have quantum correlations. The terms in Eq. (B8) with $(i, j) \neq (i', j')$, therefore, vanish, so

$$\langle\!\langle\epsilon_{i\mu}\epsilon_{i'\mu'}\rangle\!\rangle = \delta_{ii'}\frac{1}{n^2}\sum_{j=1}^n \left[\langle\Pi_{i\mu}^j\Pi_{i\mu'}^j\rangle_{\rho^{\otimes N}} - \langle\Pi_{i\mu}^j\rangle_{\rho^{\otimes N}}\langle\Pi_{i\mu'}^j\rangle_{\rho^{\otimes N}}\right] \quad \text{(B9)}$$

$$= \delta_{ii'}\frac{1}{n}[\langle\Pi_{i\mu}\Pi_{i\mu'}\rangle_\rho - \langle\Pi_{i\mu}\rangle_\rho\langle\Pi_{i\mu'}\rangle_\rho] \quad \text{(B10)}$$

$$= \delta_{ii'}\frac{1}{n}\mathrm{cov}_\rho(\Pi_{i\mu}, \Pi_{i\mu'}), \quad \text{(B11)}$$

where $\mathrm{cov}_\rho(X, Y) \equiv \langle XY\rangle_\rho - \langle X\rangle_\rho\langle Y\rangle_\rho$.

2. Errors in the polarization operator basis

Rather than the statistical errors $\epsilon_{i\mu} \equiv \tilde{\Pi}_{i\mu} - \langle\Pi_{i\mu}\rangle_\rho$ in the estimates $\tilde{\Pi}_{i\mu}$ of the projectors $\Pi_{i\mu}$, we now consider the statistical errors $\epsilon_{i\ell} \equiv \tilde{T}_{i\ell} - \langle T_{i\ell}\rangle_\rho$ in the estimates $\tilde{T}_{i\ell}$ of the polarization operators $T_{i\ell} \equiv T_{\boldsymbol{v}_i\ell,0}$. We can expand the polarization operators $T_{i\ell}$ as a sum over projectors $\Pi_{i\mu}$ as

$$T_{i\ell} = \sum_\mu t_{\ell\mu}\Pi_{i\mu}, \quad t_{\ell\mu} \equiv \langle\mu|T_{\ell,0}|\mu\rangle$$
$$= \sqrt{\frac{2\ell+1}{d}}\langle s\mu; \ell, 0|s\mu\rangle, \quad \text{(B12)}$$

and likewise $\tilde{T}_{i\ell} \equiv \sum_\mu t_{\ell\mu} \tilde{\Pi}_{i\mu}$. The covariance between errors in the polarization operator basis is then

$$
\begin{aligned}
\langle\!\langle \epsilon_{i\ell} \epsilon_{i'\ell'} \rangle\!\rangle &= \sum_{\mu,\mu'} t_{\ell\mu} t_{\ell'\mu'} \langle\!\langle \epsilon_{i\mu} \epsilon_{i'\mu'} \rangle\!\rangle \\
&= \delta_{ii'} \frac{1}{n} \sum_{\mu,\mu'} t_{\ell\mu} t_{\ell'\mu'} \mathrm{cov}_\rho(\Pi_{i\mu}, \Pi_{i\mu'}) \\
&= \delta_{ii'} \frac{1}{n} \mathrm{cov}_\rho(T_{i\ell}, T_{i\ell'}),
\end{aligned}
\tag{B13}
$$

where we used the fact that the covariance $\mathrm{cov}_\rho(X, Y)$ is linear in both X and Y. Due to the appearance of $\delta_{ii'}$ above and the orthogonality of polarization operators $T_{i\ell}$ and $T_{i'\ell'}$ with degrees $\ell \neq \ell'$, it turns out that only the variances $\langle\!\langle \epsilon_{i\ell}^2 \rangle\!\rangle$ will ultimately contribute to the reconstruction error (see Appendix B 3). We, therefore, seek to find an upper bound on $\langle\!\langle \epsilon_{i\ell}^2 \rangle\!\rangle$.

To this end, we define the probability $p_\mu^i \equiv \langle \Pi_{i\mu} \rangle_\rho$, collect these probabilities into the classical probability distribution $p^i = \sum_\mu p_\mu^i |\mu\rangle$, and define the vector $t_\ell \equiv \sum_\mu t_{\ell\mu} |\mu\rangle$. We then observe that

$$
\langle\!\langle \epsilon_{i\ell}^2 \rangle\!\rangle = \frac{1}{n} \sigma_{p^i}^2(t_\ell), \quad \sigma_p^2(X) \equiv \sum_\mu p_\mu X_\mu^2 - \left(\sum_\mu p_\mu X_\mu \right)^2,
\tag{B14}
$$

where $\sigma_p^2(X)$ is the weighted variance of X. This variance is maximal when p has equal weight on the largest and smallest values of X, which implies that

$$
\begin{aligned}
\sigma_p^2(t_\ell) &\leq \Gamma_\ell^2, \\
\Gamma_\ell &\equiv \frac{\max_\mu t_{\ell\mu} - \min_\mu t_{\ell\mu}}{2}, \quad \text{so} \quad \langle\!\langle \epsilon_{i\ell}^2 \rangle\!\rangle \leq \frac{1}{n} \Gamma_\ell^2.
\end{aligned}
\tag{B15}
$$

Note that this bound on $\langle\!\langle \epsilon_{i\ell}^2 \rangle\!\rangle$ is tight as equality is achieved by the state,

$$
\rho_i^\star = \tfrac{1}{2}(\Pi_{i\mu_{\max}} + \Pi_{i\mu_{\min}}),
\tag{B16}
$$

where $\mu_{\max}$ ($\mu_{\min}$) is the index that maximizes (minimizes) $t_{\ell\mu}$.

To find an analytical bound on $\langle\!\langle \epsilon_{i\ell}^2 \rangle\!\rangle$ that is easier to interpret, we can use normalization of the polarization operators $(T_{i\ell}|T_{i\ell}) = \sum_\mu t_{\ell\mu}^2 = 1$ and the fact that all probabilities $p_\mu \leq 1$ to bound

$$
\sigma_p^2(t_\ell) \leq \sum_\mu p_\mu t_{\ell\mu}^2 \leq \sum_\mu t_{\ell\mu}^2 = 1, \quad \text{so} \quad \langle\!\langle \epsilon_{i\ell}^2 \rangle\!\rangle < \frac{1}{n}.
\tag{B17}
$$

We can get a tighter bound by considering the fact that $t_{\ell\mu}^2 = t_{\ell,-\mu}^2$ due to the symmetries of the Clebsch-Gordan coefficients. It follows that if $\mu_{\max} \neq 0$ then

$$
\begin{aligned}
\sigma_p^2(t_\ell) &\leq \sum_\mu p_\mu t_{\ell\mu}^2 \leq t_{\ell\mu_{\max}}^2 = \frac{1}{2} \left(t_{\ell\mu_{\max}}^2 + t_{\ell,-\mu_{\max}}^2 \right) \\
&\overset{\mu_{\max} \neq 0}{\leq} \frac{1}{2} \sum_\mu t_{\ell\mu}^2 = \frac{1}{2}.
\end{aligned}
\tag{B18}
$$

If $\mu_{\max} = 0$, then similarly,

$$
\begin{aligned}
t_{\ell\mu_{\max}}^2 + 2t_{\ell\mu_{\min}}^2 &= t_{\ell\mu_{\max}}^2 + t_{\ell\mu_{\min}}^2 + t_{\ell,-\mu_{\min}}^2 \\
&\overset{\mu_{\max}=0}{\leq} \sum_\mu t_{\ell\mu}^2 = 1,
\end{aligned}
$$

$$
\text{so} \quad |t_{\ell\mu_{\min}}| \overset{\mu_{\max}=0}{\leq} \sqrt{\frac{1 - t_{\ell\mu_{\max}}^2}{2}},
\tag{B19}
$$

which lets us bound

$$
\begin{aligned}
\Gamma_\ell &= \frac{1}{2}(t_{\ell\mu_{\max}} - t_{\ell\mu_{\min}}) \\
&\leq \frac{1}{2}(t_{\ell\mu_{\max}} + |t_{\ell\mu_{\min}}|) \\
&\overset{\mu_{\max}=0}{\leq} \frac{1}{2} t_{\ell\mu_{\max}} + \frac{1}{2}\sqrt{\frac{1 - t_{\ell\mu_{\max}}^2}{2}} \\
&\equiv \lambda(t_{\ell\mu_{\max}}).
\end{aligned}
\tag{B20}
$$

It is straightforward to show that $\lambda(x)$ is maximally $\lambda^\star \equiv \max_x \lambda(x) = \sqrt{3/8}$, so

$$
\Gamma_\ell^2 \overset{\mu_{\max}=0}{\leq} (\lambda^\star)^2 = \frac{3}{8} < \frac{1}{2}.
\tag{B21}
$$

Altogether, we, thus, find that in all cases,

$$
\sigma_p^2(t_\ell) \leq \frac{1}{2}, \quad \text{so} \quad \langle\!\langle \epsilon_{i\ell}^2 \rangle\!\rangle \leq \frac{1}{2n}.
\tag{B22}
$$

3. Revisiting the reconstruction error bound

We now revisit the derivation of reconstruction error in Sec. IV to make use of the bounds on variances $\langle\!\langle \epsilon_{i\ell}^2 \rangle\!\rangle$. To recap, for a set of measurement axes $V = \{\boldsymbol{v}\}$ and degrees $\ell \in \{0, 1, \ldots, d-1\}$ we construct the measurement matrix,

$$
M_V \equiv \sum_{\boldsymbol{v},\ell} |\boldsymbol{v}\ell\rangle (T_{\boldsymbol{v}\ell,0}|,
\tag{B23}
$$

which can be block diagonalized as

$$
\begin{aligned}
M_V U &= \sum_\ell |\ell\rangle\langle\ell| \otimes M_{V\ell}, \quad U \equiv \sum_{\ell,m} |T_{\ell m}\rangle \langle\ell m|, \\
M_{V\ell} &= \sum_{m,\boldsymbol{v}} D_{m,0}^\ell(\boldsymbol{v})|\boldsymbol{v}\rangle\langle m|,
\end{aligned}
\tag{B24}
$$

where $D_{mn}^\ell(\boldsymbol{v}) \equiv \langle\ell m|R(\boldsymbol{v})|\ell n\rangle$ is a (Wigner) rotation matrix element for a spin-ℓ particle. The block-diagonal structure of M_V allows us to index its singular values $M_{\ell m}^V$ and corresponding (normalized) left singular vectors $\boldsymbol{x}_{\ell m}^V = \sum_i x_{\ell m i}^V |\boldsymbol{v}_i\rangle$ by the indices (ℓ, m) where the integer $|m| \leq \ell$. These singular vectors and values define the orthonormal operators,

$$
Q_{\ell m}^V \equiv \sum_i \left(q_{\ell m i}^V \right)^* T_{i\ell}, \quad q_{\ell m i}^V \equiv \frac{x_{\ell m i}^V}{M_{\ell m}^V},
\tag{B25}
$$

where $i \in \{1, 2, \ldots, |V|\}$ indexes an axis $\boldsymbol{v}_i \in V$ with $T_{i\ell} \equiv T_{\boldsymbol{v}_i\ell}$. The state ρ can be expanded in the basis of these operators as

$$
\rho = \sum_{\ell,m} \langle Q_{\ell m}^{V\,\dagger} \rangle_\rho Q_{\ell m}^V,
\tag{B26}
$$

and the estimates $\tilde{T}_{i\ell}$ of $\langle T_{i\ell}\rangle_\rho$ can be used to construct the following estimate $\tilde{\rho}_V$ of ρ:

$$
\tilde{\rho}_V \equiv \sum_{\ell,m} \left[\sum_i q^V_{\ell m i} \tilde{T}_{i\ell} \right] Q^V_{\ell m} \approx \sum_{\ell,m} \left[\sum_i q^V_{\ell m i} \langle T_{i\ell}\rangle_\rho \right] Q^V_{\ell m}
$$
$$
= \sum_{\ell,m} \langle Q^V_{\ell m}\rangle_\rho Q^V_{\ell m} = \rho. \tag{B27}
$$

Recalling that $\epsilon_{i\ell} \equiv \tilde{T}_{i\ell} - \langle T_{i\ell}\rangle_\rho$, we can use orthonormality of all $Q^V_{\ell m}$'s to expand the mean-squared distance between $\tilde{\rho}_V$ and ρ as

$$
\mathcal{E}_V(\rho)^2 \equiv \langle\langle \|\tilde{\rho}_V - \rho\|^2 \rangle\rangle = \sum_{\ell,m,i,i'} \left(q^V_{\ell m i} \right)^* q^V_{\ell m i'} \langle\langle \epsilon_{i\ell}\epsilon_{i'\ell} \rangle\rangle
$$
$$
= \sum_{\ell,m,i} |q^V_{\ell m i}|^2 \langle\langle \epsilon^2_{i\ell} \rangle\rangle < \frac{1}{n}\sum_\ell \Gamma^2_\ell \mathcal{S}^2_{V\ell}, \tag{B28}
$$

where we used the fact that $\langle\langle \epsilon^2_{i\ell} \rangle\rangle \leqslant \Gamma^2_\ell/n$, and

$$
\sum_{m,i} |q^V_{\ell m i}|^2 = \sum_m \left(M^V_{\ell m} \right)^{-2} = \|M^{-1}_{V\ell}\| = \mathcal{S}^2_{V\ell}. \tag{B29}
$$

Here $M^{-1}_{V\ell}$ is the left inverse of $M_{V\ell}$. The fact that $\langle\langle \epsilon^2_{i\ell} \rangle\rangle < 1/2n$ also implies that

$$
\mathcal{E}_V(\rho)^2 < \frac{1}{2n}\sum_\ell \mathcal{S}^2_{V\ell} = \frac{\mathcal{S}^2_V}{2n}. \tag{B30}
$$

Note that the bound in Eq. (B28) is not tight as the individual bounds on the variances $\langle\langle \epsilon^2_{i\ell} \rangle\rangle$ cannot all be achieved simultaneously. There is, therefore, still room for improvement on the bound in Eq. (30) by maximizing $\mathcal{E}_V$ over the set of physically achievable qudit states ρ.

APPENDIX C: EXACT RECONSTRUCTION ERROR

Here we find exact expressions for reconstruction error, which can be used to estimate the error in a given reconstruction $\tilde{\rho}_V$ of an unknown state ρ after performing tomography. To this end, we start with Eq. (B28) from Appendix B 3 to write

$$
\mathcal{E}_V(\rho)^2 = \sum_{\ell,m,i,i'} \left(q^V_{\ell m i} \right)^* q^V_{\ell m i'} \langle\langle \epsilon_{i\ell}\epsilon_{i'\ell} \rangle\rangle
$$
$$
= \frac{1}{n}\sum_{\ell,i} |\tilde{q}_{\ell i}|^2 \mathrm{cov}_\rho(T_{i\ell}, T_{i\ell}),
$$
$$
|\tilde{q}_{\ell i}|^2 = \sum_m |q_{\ell m i}|^2, \tag{C1}
$$

where $\tilde{q}_{\ell i} = \sum_m (q_{\ell m i})^* |m\rangle$ and we used the fact that $\langle\langle \epsilon_{i\ell}\epsilon_{i'\ell} \rangle\rangle = \delta_{ii'} \mathrm{cov}_\rho(T_{i\ell}, T_{i\ell})/n$. Identifying the singular value decomposition $M_{V\ell} = U_{V\ell}\Sigma_{V\ell}W^\dagger_{V\ell}$, we then we observe that $\tilde{q}_{\ell i} = \Sigma^{-1}_{V\ell}U^\dagger_{V\ell}|v_i\rangle$, which allows us to simplify

$$
|\tilde{q}_{\ell i}|^2 = \langle v_i | U_{V\ell}\Sigma^{-2}_{V\ell}U^\dagger_{V\ell}|v_i\rangle = \langle v_i | \left(M^{-1}_{V\ell} \right)^\dagger M^{-1}_{V\ell}|v_i\rangle. \tag{C2}
$$

Using the fact that all $T_{i\ell} = T^\dagger_{i\ell}$'s, we can also expand

$$
\mathrm{cov}_\rho(T_{i\ell}, T_{i\ell}) = \mathrm{cov}_\rho(T^\dagger_{i\ell}, T_{i\ell})
$$
$$
= \sum_{m,m'} D^\ell_{0,m}(v_i) D^\ell_{0,m'}(v_i)^* \mathrm{cov}_\rho(T^\dagger_{\ell m}, T_{\ell m'}), \tag{C3}
$$

which implies that

$$
\mathcal{E}_V(\rho)^2 = \frac{1}{n}\sum_{\ell,i,m,m'} D^\ell_{0,m'}(v_i)^* |\tilde{q}_{\ell i}|^2 D^\ell_{0,m}(v_i)\mathrm{cov}_\rho(T^\dagger_{\ell m}, T_{\ell m'}). \tag{C4}
$$

Altogether, this reconstruction error can be expressed more compactly by defining the *covariance matrix*,

$$
\mathcal{C}_\ell[\rho] \equiv \sum_{m,m'} \mathrm{cov}_\rho(T^\dagger_{\ell m}, T_{\ell m'})|m\rangle\langle m'|, \tag{C5}
$$

and the *noise matrix*,

$$
\mathcal{N}_{V\ell} \equiv M^\dagger_{V\ell}\mathrm{diag}\left[\left(M^{-1}_{V\ell} \right)^\dagger M^{-1}_{V\ell} \right] M_{V\ell}, \tag{C6}
$$

where $\mathrm{diag}[X]$ sets all off-diagonal entries of X to zero, in terms of which,

$$
\mathcal{E}_V(\rho)^2 = \frac{1}{n}\sum_\ell \left(\mathcal{N}_{V\ell} | \mathcal{C}_\ell[\rho] \right), \tag{C7}
$$

where $(X|Y) = \mathrm{tr}(X^\dagger Y)$ is a trace inner product.

The result in Eq. (C7) essentially expresses reconstruction error as a weighted sum of the covariances $\mathrm{cov}_\rho(T_{\ell m}, T_{\ell m'})$ where the weights are given by the corresponding matrix elements of the noise matrix $\mathcal{N}_{V\ell}$. This expression is perhaps the most physically meaningful form of the reconstruction error $\mathcal{E}_V(\rho)$ that we will consider in this paper, but in practice it turns out that Eq. (C7) is inconvenient and inefficient to evaluate for any given state ρ. To find a more practical expression of reconstruction error, we use the fact that

$$
\langle T^\dagger_{\ell m}\rangle_\rho = (\rho|T^\dagger_{\ell m}) = \mathrm{tr}(\rho T^\dagger_{\ell m}) = \mathrm{tr}(T^\dagger_{\ell m}\rho) = (T_{\ell m}|\rho), \tag{C8}
$$

to expand the covariance matrix as

$$
\mathcal{C}_\ell[\rho] = \sum_{m,m'} |m\rangle\langle m'|[(\rho|T^\dagger_{\ell m}T_{\ell m'}) - (\rho|T^\dagger_{\ell m})(\rho|T_{\ell m'})] \tag{C9}
$$
$$
= \sum_{m,m'} |m\rangle\langle m'|[(T^\dagger_{\ell m'}T_{\ell m}|\rho) - (T_{\ell m}|\rho)(T^\dagger_{\ell m'}|\rho)] \tag{C10}
$$
$$
= \sum_{m,m'} |m\rangle\langle m'|\mathcal{I}[(T_{\ell m'}T_{\ell m}|\rho) - (T_{\ell m}|\rho)(T_{\ell m'}|\rho)], \tag{C11}
$$

where we define the inversion operator $\mathcal{I} \equiv \sum_m (-1)^m|-m\rangle\langle m|$. We then expand the product $T_{\ell m'}T_{\ell m}$ as

$$
(T_{\ell m'}T_{\ell m}|\rho) = \sum_L g^L_{\ell m'm}(T_{L,m'+m}|\rho),
$$
$$
g^L_{\ell m'm} \equiv (T_{L,m'+m}|T_{\ell m'}T_{\ell m}) = f^{L,m'+m}_{\ell m';\ell m}, \tag{C12}
$$

where the (real) factors $f^{L,m'+m}_{\ell m';\ell m}$ are provided in Appendix D. Substituting the covariance matrix back into Eq. (C7) and

replacing $(T_{\ell m}|\rho) \to \rho_{\ell m}$, we get

$$\mathcal{E}_V(\rho)^2 = \frac{1}{n}\left[\sum_{\ell,m}\left(\chi_{\ell m}^V\right)^*\rho_{\ell m} - \sum_{\ell,m,m'}\langle m'|\mathcal{IN}_{V\ell}|m\rangle\rho_{\ell m}\rho_{\ell m'}\right],$$
(C13)

where

$$\chi_{LM}^V \equiv \sum_{\ell,m,m'}\delta_{M,m'+m}\langle m'|\mathcal{IN}_{V\ell}|m\rangle^* g_{\ell m'm}^L$$
(C14)

$$= \sum_{\ell,m,m'}\delta_{M,-m'+m}\langle m|\mathcal{N}_{V\ell}|m'\rangle(-1)^{m'}g_{\ell,-m',m}^L$$
(C15)

$$= \sum_{\ell}(\mathcal{N}_{V\ell}|\mathcal{D}_M|\mathcal{I}g_{L\ell})$$
(C16)

can be written in terms of the matrices,

$$g_{L\ell} \equiv \sum_{m,m'}g_{\ell m'm}^L|m'\rangle\langle m|, \quad \mathcal{D}_M \equiv \sum_{m,m'}\delta_{M,-m'+m}|m'm\rangle\langle m'm|.$$
(C17)

Here $\mathcal{D}_M$ simply picks off the Mth diagonal of the matrix it acts on such that $(\mathcal{N}_{V\ell}|\mathcal{D}_M|\mathcal{I}g_{L\ell})$ is an inner product of the Mth diagonal of $\mathcal{I}g_{L\ell}$ with the $(-M)$-th diagonal of $\mathcal{N}_{V\ell}$. Defining the $(2\ell + 1)$-component vectors,

$$|\rho_\ell\rangle \equiv \sum_m\rho_{\ell m}|m\rangle, \quad |\chi_{V\ell}\rangle \equiv \sum_{\ell,m}\chi_{\ell m}^V|m\rangle,$$
(C18)

we can write the expansion in Eq. (C13) in the vectorized form

$$\mathcal{E}_V(\rho)^2 = \frac{1}{n}\sum_\ell[\langle\chi_{V\ell}|\rho_\ell\rangle - \langle\rho_\ell|\mathcal{N}_{V\ell}|\rho_\ell\rangle].$$
(C19)

Comments on a tight reconstruction error bound

In principle, maximizing the reconstruction error in Eq. (C19) over all qudit states ρ would provide a tight upper bound on the reconstruction error for any set of axes V. To simplify this task somewhat, we first maximize Eq. (C19) over all ρ's with $\mathrm{tr}(\rho) = 1$: this maximum occurs at a state $\sigma_V^\star$ whose components are given by

$$|\sigma_{V\ell}^\star\rangle \overset{\ell\neq 0}{\equiv} \frac{1}{2}\mathcal{N}_{V\ell}^{-1}|\chi_{V\ell}\rangle, \quad |\sigma_{V,0}^\star\rangle \equiv \frac{1}{\sqrt{d}}|0\rangle.$$
(C20)

The corresponding maximum of $\mathcal{E}_V$ is given by

$$\mathcal{E}_V(\sigma_V^\star)^2 = \frac{1}{n}\sum_{\ell>0}\left[\frac{1}{4}\langle\chi_{V\ell}|\mathcal{N}_{V\ell}^{-1}|\chi_{V\ell}\rangle - \frac{1}{d}\mathrm{tr}(\mathcal{N}_{V\ell})\right],$$
(C21)

where the $\mathrm{tr}(\mathcal{N}_{V\ell})$ terms above come from simplifying the $\ell = 0$ terms of Eq. (C19) with $\rho \to \sigma_V^\star$. Although $\mathcal{E}_V(\sigma_V^\star)$ is a strict upper bound on $\mathcal{E}_V(\rho)$ over all ρ with $\mathrm{tr}(\rho) = 1$, this bound turns out to be useless in practice because $\sigma_V^\star$ will generally be a nonphysical state with negative eigenvalues. To find tight bound on $\mathcal{E}_V(\rho)$ over the space of physical qudit states ρ, we also need to constrain ρ to have no negative eigenvalues. Equipped with $\sigma_V^\star$ and $\mathcal{E}_V(\sigma_V^\star)$, we can expand

$$\mathcal{E}_V(\rho)^2 = \mathcal{E}_V(\sigma_V^\star)^2 - \frac{1}{n}\|\rho - \sigma_V^\star\|_V^2,$$
$$\|X\|_V^2 \equiv \sum_\ell\langle X_\ell|\mathcal{N}_{V\ell}|X_\ell\rangle,$$
(C22)

where $X_\ell \equiv \sum_m(T_{\ell m}|X)|m\rangle$ is a vector of the degree-ℓ components of X in the polarization operator basis, and $\|X\|_V$ is a noise-weighted norm of X. Maximizing $\mathcal{E}_V$ over all qudit states ρ, thus, amounts to finding the closest physical qudit state ρ to $\sigma_V^\star$ with distance measured by the metric
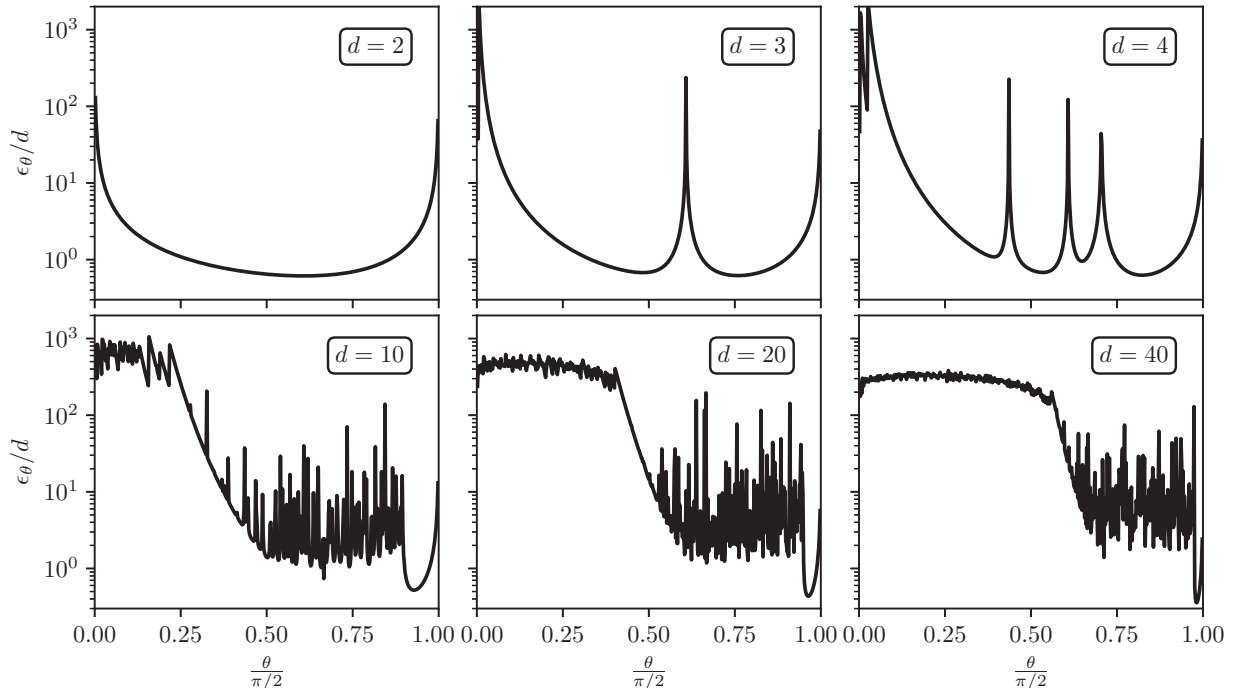
FIG. 4. Quantum error scale ϵ_θ as a function of the polar angle θ in the tomography method of Ref. [38] for a few qudit dimensions d.

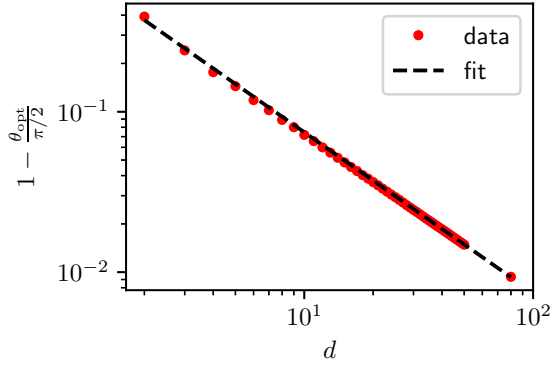FIG. 5. Optimum angle θ_{opt} as a function of qudit dimension d for the tomography method of Ref. [38] and a fit to $\theta_{\text{opt}} = \frac{\pi}{2}(1 - \frac{1}{xd})$ finding $x \approx 1.34$.

$D_V(X, Y) \equiv \|X - Y\|_V$. We leave this minimization problem to future work and note that solving it will likely require making use of the positivity conditions derived in Ref. [51]. A loose lower bound on $\|\rho - \sigma_V^\star\|_V$ can be found by minimization under the constraint $\|\rho\| \leqslant 1$, which may provide a tighter upper bound on $\mathcal{E}_V(\rho)$ than that in Eq. (30) of the main text.

APPENDIX D: POLARIZATION OPERATOR PRODUCT EXPANSION

The polarization operators on the d-dimensional Hilbert space of a spin-s system (with $s \equiv \frac{d-1}{2}$) are defined by

$$T_{\ell m} \equiv \sqrt{\frac{2\ell + 1}{2s + 1}} \sum_{\mu,\nu=-s}^{s} \langle s\mu; \ell m | s\nu \rangle |\nu\rangle\langle\mu|, \qquad \text{(D1)}$$

where $\langle s\mu; \ell m | s\nu \rangle$ is a Clebsch-Gordan coefficient that enforces $\ell \in \{0, 1, \ldots, 2s\}$ and $m \in \{-\ell, -\ell + 1, \ldots, \ell\}$. We wish to compute the coefficients of the operator product expansion,

$$T_{\ell_1 m_1} T_{\ell_2 m_2} = \sum_{L,M} f_{\ell_1 m_1; \ell_2 m_2}^{LM} T_{LM},$$

$$f_{\ell_1 m_1; \ell_2 m_2}^{LM} \equiv (T_{LM} | T_{\ell_1 m_1} T_{\ell_2 m_2}), \qquad \text{(D2)}$$

which allow us to simplify the commutators in Eq. (A3) of Appendix A. Using the symmetry properties of Clebsch-Gordan coefficients, namely,

$$\langle \ell_1 m_1; \ell_2 m_2 | LM \rangle$$
$$= (-1)^{\ell_2 + m_2} \sqrt{\frac{2L + 1}{2\ell_1 + 1}} \langle L, -M; \ell_2 m_2 | \ell_1, -m_1 \rangle \quad \text{(D3)}$$

$$\langle \ell_1 m_1; \ell_2 m_2 | LM \rangle = (-1)^{\ell_1 + \ell_2 - L} \langle \ell_1, -m_1; \ell_2, -m_2 | L, -M \rangle, \qquad \text{(D4)}$$

we can find that the polarization operators transform under conjugation as

$$T_{\ell m}^\dagger = \sqrt{\frac{2\ell + 1}{2s + 1}} \sum_{\mu,\nu} (-1)^m \langle s\nu; \ell, -m | s\mu \rangle |\mu\rangle\langle\nu|$$
$$= (-1)^m T_{\ell,-m}, \qquad \text{(D5)}$$

which implies that

$$f_{\ell_1 m_1; \ell_2 m_2}^{LM} = (-1)^M \sqrt{\frac{(2L+1)(2\ell_1+1)(2\ell_2+1)}{(2s+1)(2s+1)(2s+1)}} \sum_{\mu,\nu,\rho} \langle s\nu; L, -M | s\mu \rangle \langle s\rho; \ell_1 m_1 | s\nu \rangle \langle s\mu; \ell_2 m_2 | s\rho \rangle. \qquad \text{(D6)}$$

Replacing Clebsch-Gordan coefficients by Wigner 3-j symbols with the identity,

$$\langle \ell_1 m_1; \ell_2 m_2 | LM \rangle = (-1)^{2\ell_2}(-1)^{L-M} \sqrt{2L+1} \begin{pmatrix} L & \ell_2 & \ell_1 \\ -M & m_2 & m_1 \end{pmatrix}, \qquad \text{(D7)}$$

we can use the fact that $2\ell_2$ is always even (because ℓ_2 is always an integer) to expand

$$f_{\ell_1 m_1; \ell_2 m_2}^{LM} = (-1)^M \sqrt{(2L+1)(2\ell_1+1)(2\ell_2+1)}$$
$$\times \sum_{\mu,\nu,\rho} (-1)^{3s-\mu-\nu-\rho} \begin{pmatrix} s & L & s \\ -\mu & -M & \nu \end{pmatrix} \begin{pmatrix} s & \ell_1 & s \\ -\nu & m_1 & \rho \end{pmatrix} \begin{pmatrix} s & \ell_2 & s \\ -\rho & m_2 & \mu \end{pmatrix}. \qquad \text{(D8)}$$

This sum can be simplified by the introduction of Wigner 6-j symbols, giving us

$$f_{\ell_1 m_1; \ell_2 m_2}^{LM} = (-1)^{2s+M} \sqrt{(2L+1)(2\ell_1+1)(2\ell_2+1)} \begin{pmatrix} L & \ell_1 & \ell_2 \\ M & -m_1 & -m_2 \end{pmatrix} \begin{Bmatrix} L & \ell_1 & \ell_2 \\ s & s & s \end{Bmatrix} \qquad \text{(D9)}$$

$$= (-1)^{2s+L} \sqrt{(2\ell_1+1)(2\ell_2+1)} \langle \ell_1 m_1; \ell_2 m_2 | LM \rangle \begin{Bmatrix} \ell_1 & \ell_2 & L \\ s & s & s \end{Bmatrix}. \qquad \text{(D10)}$$

APPENDIX E: OPTIMIZING THE METHOD OF NEWTON AND YOUNG

Reference [38] constructs an explicit protocol for spin qudit tomography, which involves measuring spin projection along $2d - 1$ axes equally spaced at a polar angle θ. However,

this method does not provide any prescription for choosing θ. Here, we show the importance of making a good choice of θ and empirically find the optimal value of θ_{opt} that minimizes the corresponding quantum error scale ϵ_θ, which controls state

reconstruction error. To this end, Fig. 4 shows the quantum error scale ϵ_θ as a function of the polar angle θ in the tomography method of Ref. [38] for a few qudit dimensions d. Although a good choice of θ yields an error scale $\epsilon_\theta \approx d$ (for the dimensions shown), this error scale can increase by orders of magnitude for poor choices of θ. In turn, Fig. 5 shows the optimal angle θ_{opt} as a function of the qudit dimension d together with a fit to $\theta_{opt} = \frac{\pi}{2}(1 - \frac{1}{xd})$ finding $x \approx 1.34$.

[1] T. Opatrný, D.-G. Welsch, and W. Vogel, Phys. Rev. A **56**, 1788 (1997).

[2] B. Qi, Z. Hou, L. Li, D. Dong, G. Xiang, and G. Guo, Sci. Rep. **3**, 3496 (2013).

[3] Y. S. Teo, H. Zhu, B.-G. Englert, J. Řeháček, and Z. Hradil, Phys. Rev. Lett. **107**, 020404 (2011).

[4] J. A. Smolin, J. M. Gambetta, and G. Smith, Phys. Rev. Lett. **108**, 070502 (2012).

[5] F. Huszár and N. M. T. Houlsby, Phys. Rev. A **85**, 052120 (2012).

[6] C. Ferrie, New J. Phys. **16**, 093035 (2014).

[7] C. Granade, J. Combes, and D. G. Cory, New J. Phys. **18**, 033024 (2016).

[8] D. Gross, Y.-K. Liu, S. T. Flammia, S. Becker, and J. Eisert, Phys. Rev. Lett. **105**, 150401 (2010).

[9] A. Kalev, R. L. Kosut, and I. H. Deutsch, npj Quantum Inf. **1**, 1 (2015).

[10] G. Torlai, G. Mazzola, J. Carrasquilla, M. Troyer, R. Melko, and G. Carleo, Nat. Phys. **14**, 447 (2018).

[11] A. J. Daley, Quantum Inf. Process. **10**, 865 (2011).

[12] M. Lu, N. Q. Burdick, S. H. Youn, and B. L. Lev, Phys. Rev. Lett. **107**, 190401 (2011).

[13] B. E. Mischuck, S. T. Merkel, and I. H. Deutsch, Phys. Rev. A **85**, 022302 (2012).

[14] K. Aikawa, A. Frisch, M. Mark, S. Baier, A. Rietzler, R. Grimm, and F. Ferlaino, Phys. Rev. Lett. **108**, 210401 (2012).

[15] A. Smith, B. E. Anderson, H. Sosa-Martinez, C. A. Riofrìo, I. H. Deutsch, and P. S. Jessen, Phys. Rev. Lett. **111**, 170502 (2013).

[16] M. A. Cazalilla and A. M. Rey, Rep. Prog. Phys. **77**, 124401 (2014).

[17] J. L. Bohn, A. M. Rey, and J. Ye, Science **357**, 1002 (2017).

[18] T. Takekoshi, L. Reichsöllner, A. Schindewolf, J. M. Hutson, C. R. Le Sueur, O. Dulieu, F. Ferlaino, R. Grimm, and H.-C. Nägerl, Phys. Rev. Lett. **113**, 205301 (2014).

[19] M. Zeppenfeld, B. G. U. Englert, R. Glöckner, A. Prehn, M. Mielenz, C. Sommer, L. D. van Buuren, M. Motsch, and G. Rempe, Nature (London) **491**, 570 (2012).

[20] I. Kozyryev, L. Baum, K. Matsuda, B. L. Augenbraun, L. Anderegg, A. P. Sedlack, and J. M. Doyle, Phys. Rev. Lett. **118**, 173201 (2017).

[21] P. Puri, M. Mills, C. Schneider, I. Simbotin, J. A. Montgomery, R. Côté, A. G. Suits, and E. R. Hudson, Science **357**, 1370 (2017).

[22] X. Wu, T. Gantner, M. Koller, M. Zeppenfeld, S. Chervenkov, and G. Rempe, Science **358**, 645 (2017).

[23] L. D. Marco, G. Valtolina, K. Matsuda, W. G. Tobias, J. P. Covey, and J. Ye, Science **363**, 853 (2019).

[24] L. R. Liu, J. D. Hood, Y. Yu, J. T. Zhang, K. Wang, Y.-W. Lin, T. Rosenband, and K.-K. Ni, Phys. Rev. X **9**, 021039 (2019).

[25] L. Anderegg, L. W. Cheuk, Y. Bao, S. Burchesky, W. Ketterle, K.-K. Ni, and J. M. Doyle, Science **365**, 1156 (2019).

[26] C. W. Chou, A. L. Collopy, C. Kurz, Y. Lin, M. E. Harding, P. N. Plessow, T. Fortier, S. Diddams, D. Leibfried, and D. R. Leibrandt, Science **367**, 1458 (2020).

[27] Y. Lin, D. R. Leibrandt, D. Leibfried, and C.-w. Chou, Nature (London) **581**, 273 (2020).

[28] D. Hemmer, E. Montano, B. Q. Baragiola, L. M. Norris, E. Shojaee, I. H. Deutsch, and P. S. Jessen, Phys. Rev. A **104**, 023710 (2021).

[29] A. Evrard, V. Makhalov, T. Chalopin, L. A. Sidorenkov, J. Dalibard, R. Lopes, and S. Nascimbene, Phys. Rev. Lett. **122**, 173601 (2019).

[30] D. Banerjee, M. Bögli, M. Dalmonte, E. Rico, P. Stebler, U.-J. Wiese, and P. Zoller, Phys. Rev. Lett. **110**, 125303 (2013).

[31] X. Zhang, M. Bishof, S. L. Bromley, C. V. Kraus, M. S. Safronova, P. Zoller, A. M. Rey, and J. Ye, Science **345**, 1467 (2014).

[32] F. Scazza, C. Hofrichter, M. Höfer, P. C. D. Groot, I. Bloch, and S. Fölling, Nat. Phys. **10**, 779 (2014).

[33] A. Goban, R. B. Hutson, G. E. Marti, S. L. Campbell, M. A. Perlin, P. S. Julienne, J. P. D'Incao, A. M. Rey, and J. Ye, Nature (London) **563**, 369 (2018).

[34] M. A. Perlin, D. Barberena, M. Mamaev, B. Sundar, R. J. Lewis-Swan, and A. M. Rey, arXiv:2109.11019.

[35] V. V. Albert, J. P. Covey, and J. Preskill, Phys. Rev. X **10**, 031050 (2020).

[36] J. A. Gross, Phys. Rev. Lett. **127**, 010504 (2021).

[37] K. Barnes, P. Battaglino, B. J. Bloom, K. Cassella, R. Coxe, N. Crisosto, J. P. King, S. S. Kondov, K. Kotru, S. C. Larsen, J. Lauigan, B. J. Lester, M. McDonald, E. Megidish, S. Narayanaswami, C. Nishiguchi, R. Notermans, L. S. Peng, A. Ryou, T.-Y. Wu, and M. Yarwood, arXiv:2108.04790.

[38] R. G. Newton and B.-l. Young, Ann. Phys. (NY) **49**, 393 (1968).

[39] H. F. Hofmann and S. Takeuchi, Phys. Rev. A **69**, 042108 (2004).

[40] S. N. Filippov and V. I. Man'ko, J. Russ. Laser Res. **31**, 32 (2010).

[41] R. Schmied and P. Treutlein, New J. Phys. **13**, 065019 (2011).

[42] S. T. Flammia, A. Silberfarb, and C. M. Caves, Found. Phys. **35**, 1985 (2005).

[43] R. T. Thew, K. Nemoto, A. G. White, and W. J. Munro, Phys. Rev. A **66**, 012303 (2002).

[44] R. Salazar and A. Delgado, Phys. Rev. A **86**, 012118 (2012).

[45] H. Sosa-Martinez, arXiv:1706.06536.

[46] D. Ha and Y. Kwon, Quantum Inf. Process. **17**, 232 (2018).

[47] Q. P. Stefano, L. Rebón, S. Ledesma, and C. Iemmi, Opt. Lett. **44**, 2558 (2019).

[48] A. M. Pălici, T.-A. Isdrailă, S. Ataman, and R. Ionicioiu, Quantum Sci. Technol. **5**, 045004 (2020).

[49] V. I. Man'ko and O. V. Man'ko, J. Exp. Theor. Phys. **85**, 430 (1997).

[50] Spin Qudit Tomography, https://github.com/perlinm/spin_qudit_tomography (2021).

[51] S. Kryszewski and M. Zachciał, J. Phys. A **39**, 5921 (2006).

[52] R. A. Bertlmann and P. Krammer, J. Phys. A: Math. Theor. **41**, 235303 (2008).

[53] F. Li, C. Braun, and A. Garg, Europhys. Lett. **102**, 60006 (2013).

[54] J. D. McEwen and Y. Wiaux, IEEE Trans. Signal Process. **59**, 5876 (2011).

[55] H. Rauhut and R. Ward, arXiv:1102.4097.

[56] Y. F. Alem, D. H. Chae, and R. A. Kennedy, Sparse signal recovery on the sphere: Optimizing the sensing matrix through sampling, in *2012 6th International Conference on Signal Processing and Communication Systems, Gold Coast, Australia, 2012* (IEEE, Piscataway, NJ, 2012), pp. 1–6, doi: 10.1109/ICSPCS.2012.6508014.

[57] Z. Khalid, R. A. Kennedy, and J. D. McEwen, IEEE Trans. Signal Process. **62**, 4597 (2014).

[58] W. Freeden and M. Schreiner, *Spherical Functions of Mathematical Geosciences: A Scalar, Vectorial, and Tensorial Setup* (Springer, Berlin, 2008).

[59] W. Freeden, M. Z. Nashed, and M. Schreiner, in *Spherical Sampling*, edited by W. Freeden, M. Z. Nashed, and M. Schreiner, Geosystems Mathematics (Springer, Berlin, 2018).

[60] V. Giovannetti, S. Lloyd, and L. Maccone, Phys. Rev. Lett. **96**, 010401 (2006).

[61] C. Ferrie, Phys. Rev. Lett. **113**, 190404 (2014).

[62] L. Pereira, L. Zambrano, J. Cortés-Vega, S. Niklitschek, and A. Delgado, Phys. Rev. A **98**, 012339 (2018).