Least-bias state estimation with incomplete unbiased measurements

Jaroslav Řeháček,[1] Zdeněk Hradil,[1] Yong Siah Teo,[1] Luis L. Sánchez-Soto,[2,3] Hui Khoon Ng,[4,5,6] Jing Hao Chai,[5] and Berthold-Georg Englert[5,6,7]

[1]*Department of Optics, Palacký University, 17. listopadu 12, 77146 Olomouc, Czech Republic*
[2]*Departamento de Óptica, Facultad de Física, Universidad Complutense, 28040 Madrid, Spain*
[3]*Max-Planck-Institut für die Physik des Lichts, Günther-Scharowsky-Straße 1, Bau 24, 91058 Erlangen, Germany*
[4]*Yale-NUS College, 16 College Avenue West, Singapore 138527, Singapore*
[5]*Centre for Quantum Technologies, National University of Singapore, 3 Science Drive 2, Singapore 117543, Singapore*
[6]*MajuLab, CNRS-UNS-NUS-NTU International Joint Unit, UMI 3654, Singapore*
[7]*Department of Physics, National University of Singapore, 2 Science Drive 3, Singapore 117542, Singapore*

(Received 14 August 2015; published 3 November 2015)

Measuring incomplete sets of mutually unbiased bases constitutes a sensible approach to the tomography of high-dimensional quantum systems. The unbiased nature of these bases optimizes the uncertainty hypervolume. However, imposing unbiasedness on the probabilities for the unmeasured bases does not generally yield the estimator with the largest von Neumann entropy, a popular figure of merit in this context. Furthermore, this imposition typically leads to mock density matrices that are not even positive definite. This provides a strong argument against perfunctory applications of linear estimation strategies. We propose to use instead the physical state estimators that maximize the Shannon entropy of the unmeasured outcomes, which quantifies our lack of knowledge fittingly and gives physically meaningful statistical predictions.

DOI: 10.1103/PhysRevA.92.052303 PACS number(s): 03.67.−a, 03.65.Ud, 03.65.Wj

I. INTRODUCTION

Modern quantum technologies harness characteristic features of quantum systems to gain performance that is otherwise unattainable through classical means. This progress ultimately relies on the ability to create, manipulate, and measure quantum states. All of these tasks require a step-by-step verification in the experimental procedures; this is essentially the scope of quantum tomography [1].

Typically, a tomographic protocol attempts to infer the unknown quantum state from the distinct outcomes of a collection of measurements performed on a finite set of identical copies of the system. With these limited resources, the choice of optimal measurements and the design of efficient reconstruction algorithms are crucial.

When the measurement outcomes form an informationally complete set [2–6], the data obtained contain maximal information and a unique state estimator can be inferred. Unfortunately, as we probe more intricate quantum systems [7,8], such an informationally complete set of measurements becomes extremely difficult to implement. In addition, the complete knowledge of the quantum state of a system with many degrees of freedom is usually not needed, as one could very well be interested in a few parameters only, such as the fidelity with respect a target state or a measure of entanglement.

We are thus inevitably led to the consideration of alternative techniques. A promising class of new protocols are explicitly optimized for particular kinds of states. This includes states with low rank [9–11], such as matrix product states [12,13], or multiscale entanglement renormalization ansatz states [14]. The specific but pertinent example of permutation invariance was also examined [15–18].

In the same spirit, several algorithms for estimating quantum states from incomplete measurement data have been reported recently [19–22]. Here, we revisit the problem in the context of mutually unbiased (MU) measurements, which are known to be optimal for state reconstruction [23] if a complete set of bases is to be measured. At the fundamental level, mutually unbiased bases (MUB) are part of the mathematical framework for an explicit formalism of the complementarity principle. The idea emerged in the pioneering work of Schwinger [24], and has since been integrated into the foundation of quantum theory: apart from their role in tomography, MUB are instrumental in addressing a number of enthralling questions [25].

A first, self-suggesting, if naive, approach could be to assume a uniform distribution for the outcomes of the unmeasured bases from a set of MUB, as this seems to be consistent with the very physical nature of MUB, which minimize the statistical uncertainty hypervolumes. However, as we show here, this is often incompatible with the positivity required by quantum mechanics, and even when the resulting estimator is physical, it is usually not the estimator with the largest von Neumann entropy [19].

The bases of eigenstates of complementary observables are called *unbiased* because we cannot predict at all the outcome of a projective measurement in one basis if the system is prepared in a state from another basis—all outcomes are equally probable, there is no preference in our prediction. This notion of being unbiased must not be confused with other uses of the adjective. In fact, there are many different meanings and connotations associated with "bias": statistical bias, bias in a sample, cultural bias, and media bias are but a few uses of the word.

In the context of state estimation, *statistical bias* is of some importance; an estimator has a statistical bias if its average over all thinkable compositions of a finite sample of measurement results deviates from the true value. This will be of no concern here because we shall take for granted that the measured sample is so large that statistical noise in the data can be safely ignored. Also, we do not have to worry about a *biased sample*, a common problem when polls are taken.

The simple "estimators of unbiased linear inversion," which we shall introduce in Sec. II C, are unbiased in the sense that they assume equal probability for all outcomes of the unmeasured bases from a set of MUB; they are, however, estimators with a statistical bias (for data from a finite sample), a property that cannot be avoided [26,27] if one insists, as we do, that all permissible estimators are physical—they must be *bona fide* density operators. As noted in the preceding paragraph, the statistical bias is irrelevant in the current context, and it is worth recalling that, despite the negative connotation of "bias," a statistical bias is not only harmless, but can be rather beneficial (see Jaynes's discussion in Secs. 17.2 and 17.3 of Ref. [28]).

The plan of this paper is as follows. After introducing background material and notational conventions in Sec. II, we illustrate various aspects of the said self-suggesting approach in Sec. III, with particular emphasis on unphysical properties of the resulting estimators. We conclude that the "estimators of unbiased linear inversion" are often unphysical—they are *not* estimators. Then, in Sec. IV, we introduce a natural alternative that keeps these estimators whenever they are physical and, when they are not, replaces them with physical estimators. Rather than being unbiased about the probabilities of unmeasured bases, the physical estimators are minimally biased. In this alternative approach, we maximize the Shannon entropy with due attention to the physical constraints and so minimize the bias. This entails a simple optimization algorithm over the state space, in which a single equation is iterated. In Sec. V, we consider alternatives to maximizing the Shannon entropy. Some of them minimize the bias with respect to another criterion, others are characterized in a different way. Finally, we offer a summary and conclusions in Sec. VI.

II. BASIC MATTERS

A. Mutually unbiased bases

The state of a d-dimensional quantum system is specified by a positive semidefinite, unit-trace density operator ϱ. Carefully note that a different symbol ρ is reserved for its $d \times d$ matrix representation, which requires $d^2 - 1$ independent real parameters for its complete characterization. If a von Neumann maximal test is chosen to fix $d - 1$ of these parameters, then a total of $d + 1$ tests is necessary to reconstruct the state. This strategy is optimal when the bases in which the measurements are carried out are "as different as possible;" that is, when these bases are MU [25].

Throughout this paper, we take the dimension d to be a prime or a prime-power integer. Then, a maximal number of $d + 1$ MUB [29] exist and can be explicitly constructed [30–39]. We denote by $|\psi_{\alpha k}\rangle$ the kth ket in the αth basis of the set of MUB; here and below, Greek indices $(\alpha, \beta, \ldots)$ label the $d + 1$ distinct MUB, whereas Latin indices $(k, l, \ldots)$ label the d outcomes in each basis.

We define MU projectors as $\Pi_{\alpha k} = |\psi_{\alpha k}\rangle\langle\psi_{\alpha k}|$. Any two MU projectors satisfy the trace relation

$$\mathrm{Tr}(\Pi_{\alpha k}\,\Pi_{\beta\ell}) = \delta_{\alpha\beta}\delta_{k\ell} + \frac{1}{d}(1 - \delta_{\alpha\beta}) \tag{2.1}$$

that states their orthonormality for $\alpha = \beta$ and their mutual unbiasedness for $\alpha \neq \beta$. Besides, these projectors constitute a complete set of projective measurements for each α,

$$\sum_{k=0}^{d-1} \Pi_{\alpha k} = \mathbb{1}. \tag{2.2}$$

To facilitate the discussion of the physics behind incomplete MUB tomography, the number of copies of the system probed by the measurement apparatus is taken sufficiently large, so statistical fluctuations in the measurement data are negligible.

The problem we are studying here is, therefore, not one of state estimation *sensu stricto*, where the estimation of the probabilities from observed relative frequencies is the central theme. Rather, we are dealing with the problem of converting the probabilities into a statistical operator, which requires a deliberate choice when the tomography is incomplete.

B. Complete tomography with MUB

The $d^2 + d$ measured probabilities

$$p_{\alpha k} = \mathrm{Tr}(\Pi_{\alpha k}\,\varrho), \tag{2.3}$$

which obey the $d + 1$ constraints $\sum_k p_{\alpha k} = 1$, fully characterize the density operator ϱ of the system. First, if only one basis is measured—the αth basis, say—our state estimator is

$$\widehat{\varrho}_\alpha = \sum_{k=0}^{d-1} p_{\alpha k}\Pi_{\alpha k}. \tag{2.4}$$

It is the most natural estimator, inasmuch as we estimate the probabilities of the unmeasured bases in the unbiased manner of $\widehat{p}_{\beta l} = \mathrm{Tr}(\Pi_{\beta l}\,\widehat{\varrho}_\alpha) = 1/d$ for $\beta \neq \alpha$; this is most adequate for MUB if we are mindful of Laplace's advice to assign equal probabilities to alternatives about which we have no information [40]. Second, after measuring all $d + 1$ bases, we know the density operator completely,

$$\varrho = \sum_{\alpha=1}^{d+1} \widehat{\varrho}_\alpha - \mathbb{1} = \frac{1}{d}\mathbb{1} + \sum_{\alpha=1}^{d+1}\left(\widehat{\varrho}_\alpha - \frac{1}{d}\mathbb{1}\right); \tag{2.5}$$

alternatively, we write

$$\varrho = \frac{1}{d}\mathbb{1} + \sum_{\alpha=1}^{d+1}\sum_{k=0}^{d-1} w_{\alpha k}\,\Pi_{\alpha k}, \quad \text{with} \quad w_{\alpha k} = p_{\alpha k} - \frac{1}{d}. \tag{2.6}$$

While the mappings $\varrho \mapsto p_{\alpha k}$ and $p_{\alpha k} \mapsto \widehat{\varrho}_\alpha$ are linear, the mappings $p_{\alpha k} \mapsto \varrho$ in Eqs. (2.5) or (2.6) appear to be affine. In fact, these are linear mappings too, since we can make use of identities such as

$$\mathbb{1} = \frac{1}{d+1}\sum_{\alpha=1}^{d+1}\sum_{j,k=0}^{d-1} p_{\alpha j}\,\Pi_{\alpha k}, \qquad w_{\alpha k} = \sum_{j=0}^{d-1} p_{\alpha j}\left(\delta_{jk} - \frac{1}{d}\right), \tag{2.7}$$

if we wish.

Here is a caveat. The linear map $p_{\alpha k} \mapsto \widehat{\varrho}_\alpha$ in (2.5) for a single basis continues to yield physical estimators if one replaces the probabilities $p_{\alpha k}$ by the corresponding relative frequencies from an experiment with a finite sample. The resulting single-basis estimators have no statistical bias. For the full tomography, however, one cannot replace the probabilities in (2.6) by relative frequencies. If one does, one gets mock estimators that have no statistical bias but are unphysical, whereas physical estimators are statistically biased since they require a suitable nonlinear mapping from the relative frequencies to estimated probabilities. See Refs. [41,42] for various aspects of these matters.

The second version of Eq. (2.5) exploits a well-known important geometrical property of density operators: the difference between any ϱ and the completely mixed state $\frac{1}{d}\mathbb{1}$ is a traceless Hermitian operator, and these operators are elements of a d^2-1 real vector space. The sum over α in Eq. (2.5) or (2.6) is a sum over hyperplanes in this vector space; there are $d+1$ of them, each $(d-1)$ dimensional since $\sum_k w_{\alpha k}=0$. The hyperplanes are orthogonal in the sense of

$$\mathrm{Tr}\left(\sum_{k=0}^{d-1} w_{\alpha k}\Pi_{\alpha k}\sum_{l=0}^{d-1} w'_{\beta l}\Pi_{\beta l}\right)=0,\quad \text{for}\quad \alpha\neq\beta, \tag{2.8}$$

and in

$$\begin{aligned}\mathrm{Tr}(\varrho\,\varrho')&=\frac{1}{d}+\sum_{\alpha=1}^{d+1}\sum_{k=0}^{d-1} w_{\alpha k}w'_{\alpha k},\\ \mathrm{Tr}[(\varrho-\varrho')^2]&=\sum_{\alpha=1}^{d+1}\sum_{k=0}^{d-1}(w_{\alpha k}-w'_{\alpha k})^2\end{aligned} \tag{2.9}$$

we recognize the Euclidean metric of the vector space.

C. Linear inversion for incomplete MUB tomography

The single-basis estimator in Eq. (2.4) is obtained from the right-hand side of Eq. (2.5) by setting $\widehat{\varrho}_\beta\mapsto\frac{1}{d}\mathbb{1}$ for the basis indices β that are different from the privileged index α. Now, if an intermediate number of bases from the set of MUB has been measured—say, the first M bases with $1<M<d+1$—the unbiasedness of the bases and the linearity just noted suggest to set $\widehat{\varrho}_\beta\mapsto\frac{1}{d}\mathbb{1}$ for the $d+1-M$ unmeasured bases, which yields the Mth estimator of *unbiased linear inversion* (ULIN),

$$\begin{aligned}\widehat{\varrho}^{(M)}_{\mathrm{ULIN}}&=\frac{1}{d}\mathbb{1}+\sum_{\alpha=1}^{M}\left(\widehat{\varrho}_\alpha-\frac{1}{d}\mathbb{1}\right)=\sum_{\alpha=1}^{M}\widehat{\varrho}_\alpha-\frac{M-1}{d}\mathbb{1}\\ &=\frac{1}{d}\mathbb{1}+\sum_{\alpha=1}^{M}\sum_{k=0}^{d-1}w_{\alpha k}\,\Pi_{\alpha k}.\end{aligned} \tag{2.10}$$

We can also regard the truncation of the α summation as effected by a replacement of the coefficients $w_{\alpha k}$ in Eq. (2.6) in accordance with

$$w_{\alpha k}\mapsto\widehat{w}^{(M)}_{\alpha k}=\begin{cases}p_{\alpha k}-\dfrac{1}{d} & \text{for}\quad 1\leqslant\alpha\leqslant M,\\ 0 & \text{for}\quad \alpha>M.\end{cases} \tag{2.11}$$

Yet, irrespective of how we regard the map of the measured probabilities $p_{\alpha k}$ with $\alpha\leqslant M$ onto the ULIN estimator $\widehat{\varrho}^{(M)}_{\mathrm{ULIN}}$, the map simply amounts to estimating the probabilities of the unmeasured bases in the unbiased manner—faithful to Laplace's advice, so to say:

$$\widehat{p}_{\beta l}=\mathrm{Tr}\left(\Pi_{\beta l}\,\widehat{\varrho}^{(M)}_{\mathrm{ULIN}}\right)=\frac{1}{d}\quad \text{for } \beta>M. \tag{2.12}$$

All other estimators are characterized by the nonzero operator they assign to the unmeasured part in

$$\widehat{\varrho}^{(M)}_{\mathrm{other}}=\widehat{\varrho}^{(M)}_{\mathrm{ULIN}}+\left(\widehat{\varrho}-\frac{1}{d}\mathbb{1}\right)^{(M)}_{\mathrm{unmeasured}}, \tag{2.13}$$

where

$$\left(\widehat{\varrho}-\frac{1}{d}\mathbb{1}\right)^{(M)}_{\mathrm{unmeasured}}=\sum_{\beta=M+1}^{d+1}\sum_{l=0}^{d-1}\widehat{w}_{\beta l,\mathrm{other}}\,\Pi_{\beta l} \tag{2.14}$$

is the estimator for the contribution of the unmeasured bases to the sums in Eqs. (2.5) and (2.6).

Before proceeding, six comments are in order. First, the map $\varrho\mapsto\widehat{\varrho}^{(M)}_{\mathrm{ULIN}}$ projects $\varrho-\frac{1}{d}\mathbb{1}$ onto the span of the first M hyperplanes in the vector space discussed above; applying the map a second time has no effect. This projection property implies

$$\begin{aligned}\mathrm{Tr}\left(\widehat{\varrho}^{(M)}_{\mathrm{ULIN}}\,\varrho'\right)&=\mathrm{Tr}\left(\varrho\,\widehat{\varrho'}^{(M)}_{\mathrm{ULIN}}\right)=\mathrm{Tr}\left(\widehat{\varrho}^{(M)}_{\mathrm{ULIN}}\,\widehat{\varrho'}^{(M)}_{\mathrm{ULIN}}\right)\\ &=\frac{1}{d}+\sum_{\alpha=1}^{M}\sum_{k=0}^{d-1}w_{\alpha k}w'_{\alpha k},\end{aligned} \tag{2.15}$$

for any two density operators ϱ and ϱ'.

Second, we note that, since Eq. (2.11) involves only the measured bases, it is also well defined for systems where M MUB exist, but this set cannot be extended to $d+1$ bases. This is the case, e.g., for $d=6$ and $M=3$ where only sets of at most three MUB have been found thus far [43–48]. There are also nonextendable sets for prime-power dimensions [49], although ULIN estimators for these sets suffer from the same problems as the ULIN estimators for subsets of full sets (see Sec. III): they are often unphysical.

Also, whether the unmeasured bases are pairwise unbiased among themselves, or unbiased with some of the measured bases, is not important for the maximum property of the following third comment, nor for the physical least-bias estimators of Secs. IV B and V A. If one wishes, one can choose the unmeasured bases such that they have largest average distance (see Ref. [50]), so that they resemble a set of MUB as best as they can.

Third, it is also worth noting that, of all estimators consistent with the measured probabilities, the ULIN estimator $\widehat{\varrho}^{(M)}_{\mathrm{ULIN}}$ maximizes the Shannon entropy

$$\mathrm{H}^{(M)}(\varrho)=-\sum_{\beta=M+1}^{d+1}\sum_{l=0}^{d-1}p_{\beta l}\ln p_{\beta l} \tag{2.16}$$

associated with the unmeasured probabilities, which must be calculated according to Eq. (2.3). The ULIN estimator has this property by construction, since it gives the uniform estimates of Eq. (2.12) for the $\widehat{p}_{\beta l}$'s.

For simplicity, we are taking some liberties with the Shannon entropy in (2.16): the probabilities are normalized to unit sum for each β, rather than to total unit sum and we are using the natural rather than the binary logarithm. This betrayal of the pure doctrine is of no consequence, however.

Fourth, irrespective of the unbiased way of estimating the $\widehat{p}_{\beta l}$'s, the ULIN estimator has a statistical bias; recall the pertinent remarks in the Introduction.

Fifth, as we shall see, after measuring M bases, we actually have some information about the $d+1-M$ unmeasured bases. We are not really faithful to Laplace's advice as long as we do not account for this information properly. For $M=1$ we recover the estimator (2.4) with $\alpha=1$, whereas $\widehat{\varrho}^{(d+1)}_{\mathrm{ULIN}}=\varrho$ is the full-tomography case of $M=d+1$ in Eq. (2.5). The reliability of $\widehat{\varrho}^{(M)}_{\mathrm{ULIN}}$ for these limiting values of M is, however, misleading, and so is the self-suggesting probability estimation (2.12). As we shall see—and this is one main objective of this work—the ULIN estimator is not assuredly physical for intermediate M values, except in the qubit situation of $d=2$. That is, $\widehat{\varrho}^{(M)}_{\mathrm{ULIN}}\ngeqslant 0$ is possible. The correct implementation of Laplace's advice is not achieved by the simple-minded estimates in (2.12).

Sixth, there is no other linear map $p_{\alpha k}\mapsto\widehat{\varrho}^{(M)}_{\mathrm{LIN}}$ that could be used instead of $p_{\alpha k}\mapsto\widehat{\varrho}^{(M)}_{\mathrm{ULIN}}$ if we require for consistency, as we must, that $p_{\alpha k}=\mathrm{Tr}(\widehat{\varrho}^{(M)}_{\mathrm{LIN}}\,\Pi_{\alpha k})$ for the M bases that have been measured. There are, of course, legitimate state estimators $\widehat{\varrho}^{(M)}$ that are consistent with the known probabilities for bases $\alpha=1,2,\ldots,M$—namely the physical "other" estimators of Eq. (2.13)—and these estimators make up a convex set; correspondingly, there is a convex set of estimated probabilities $\widehat{p}_{\beta l}$ for $\beta=M+1,\ldots,d+1$. The ULIN estimator of Eq. (2.10) and the unbiasedly estimated probabilities $\widehat{p}_{\beta l}=1/d$ may or may not be in the respective sets. When they are outside, we have to choose the best state estimator in accordance with a suitable optimality criterion. We say more about this in Sec. IV.

III. ASPECTS OF INCOMPLETE MUB TOMOGRAPHY

A. ULIN estimators for a single qubit ($d=2$)

As a first illustration, let us take the elementary example of a single qubit ($d=2$). The three MUB here are just the eigenstates of the standard Pauli operators σ_x, σ_y, and σ_z, so that $\Pi_{10}=\frac{1}{2}(\mathbb{1}+\sigma_x)\equiv\Pi_{x+},\ \ldots,\ \Pi_{31}=\frac{1}{2}(\mathbb{1}-\sigma_z)\equiv\Pi_{z-}$ and

$$\varrho=\widehat{\varrho}^{(3)}_{\mathrm{ULIN}}=\tfrac{1}{2}(\mathbb{1}+s_x\sigma_x+s_y\sigma_y+s_z\sigma_z),\qquad(3.1)$$

where $\mathbf{s}=(s_x,s_y,s_z)$ is the three-dimensional Bloch vector with, e.g., $s_x=\mathrm{Tr}(\varrho\,\sigma_x)=\frac{1}{2}(p_{x+}-p_{x-})$. The eigenvalues of ϱ are $\frac{1}{2}(1\pm|\mathbf{s}|)$, and the von Neumann entropy

$$S(\varrho)=-\,\mathrm{Tr}(\varrho\ln\varrho)\qquad(3.2)$$

grows monotonously with shrinking Bloch-vector length $|s|$.

When a single basis is measured ($M=1$), say that of σ_x, only s_x is known, and we have

$$\widehat{\varrho}^{(1)}_{\mathrm{ULIN}}=\tfrac{1}{2}(\mathbb{1}+s_x\sigma_x).\qquad(3.3)$$

Clearly, this physical estimator maximizes the von Neumann entropy, since $\widehat{p}_{y\pm}=\widehat{p}_{z\pm}=\frac{1}{2}$ give the shortest Bloch vector consistent with the known probabilities $p_{x\pm}=\frac{1}{2}(1\pm s_x)$. Likewise, the ULIN estimator for two measured bases,

$$\widehat{\varrho}^{(2)}_{\mathrm{ULIN}}=\tfrac{1}{2}(\mathbb{1}+s_x\sigma_x+s_y\sigma_y),\qquad(3.4)$$

which gives the estimated probabilities $\widehat{p}_{z_+}=\widehat{p}_{z_-}=\frac{1}{2}$ for the unmeasured σ_z, maximizes the von Neumann entropy regardless of the values for the observed probabilities $p_{x_\pm}$ and $p_{y_\pm}$.

These properties of the ULIN estimators for a qubit are consequences of the simple spherical geometry of the Bloch ball and the orthogonality of the planes corresponding to the distinct single-qubit MU Pauli observables. Such lines of argument were employed for justifying the optimality of MUB tomography in higher dimensions [23]: when picturing the results obtained with a finite number of copies using "fuzzy" hyperplanes, their mutual orthogonality makes the uncertainty hypervolume particularly small.

So, for a qubit, the unbiased unmeasured probabilities of the ULIN estimator, as quantified by the Shannon entropy, optimize (maximize) the von Neumann entropy, much like the unbiased MU observables optimize (minimize) statistical uncertainty. These two properties complement each other. In fact, for any $d\geqslant 2$ and M, if ϱ is equal to a projector $\Pi_{\alpha k}$ of one of the measured bases, then the ULIN estimator is the state itself according to the defining relation (2.1). In view of all this evidence, one might expect this mutual compatibility to extend to higher dimensions—and so arrive at the ULIN estimator of Eq. (2.10). However, we shall see that this somewhat naive approach fails for $d>2$.

B. ULIN estimators for a single qutrit ($d=3$)

For the $d=3$ case of a qutrit, we use this set of four MUB:

$$\mathcal{B}_1\,\widehat{=}\,\frac{1}{\sqrt{3}}\begin{pmatrix}1&1&1\\1&q^2&q\\1&q&q^2\end{pmatrix},\quad \mathcal{B}_2\,\widehat{=}\,\frac{1}{\sqrt{3}}\begin{pmatrix}1&1&1\\1&q^2&q\\q&q^2&1\end{pmatrix},$$
$$\mathcal{B}_3\,\widehat{=}\,\frac{1}{\sqrt{3}}\begin{pmatrix}1&1&1\\1&q^2&q\\q^2&1&q\end{pmatrix},\quad \mathcal{B}_4\,\widehat{=}\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix},\qquad(3.5)$$

where the columns are the probability amplitudes of the basis kets with reference to the fourth basis, and $q=e^{i2\pi/3}$ is the basic cubic root of unity. The 3×3 matrices that represent the respective single-basis estimators of Eq. (2.4) are

$$\widehat{\rho}_1=\frac{1}{3}\begin{pmatrix}1&z_1&z_1^*\\z_1^*&1&z_1\\z_1&z_1^*&1\end{pmatrix},\quad \widehat{\rho}_2=\frac{1}{3}\begin{pmatrix}1&z_2&q^2z_2^*\\z_2^*&1&q^2z_2\\qz_2&qz_2^*&1\end{pmatrix},$$
$$\widehat{\rho}_3=\frac{1}{3}\begin{pmatrix}1&z_3&qz_3^*\\z_3^*&1&qz_3\\q^2z_3&q^2z_3^*&1\end{pmatrix},\quad \widehat{\rho}_4=\begin{pmatrix}p_{40}&0&0\\0&p_{41}&0\\0&0&p_{42}\end{pmatrix},\qquad(3.6)$$

where

$$z_\alpha=\sum_{k=0}^{2}q^k p_{\alpha k}=\mathrm{Tr}(\varrho\,Z_\alpha),\quad\text{with}\quad Z_\alpha=\sum_{k=0}^{2}q^k\Pi_{\alpha k}.\qquad(3.7)$$

We have $p_{\alpha k}=\frac{1}{3}+w_{\alpha k}=\frac{1}{3}(1+q^{-k}z_\alpha+q^kz_\alpha^*)$ for the probabilities in terms of the z_α's, and the unitary 3×3 matrices for the pairwise complementary observables Z_α are

$$Z_\alpha \mathrel{\widehat{=}} \begin{pmatrix} 0 & 0 & q^{1-\alpha} \\ 1 & 0 & 0 \\ 0 & q^{\alpha-1} & 0 \end{pmatrix}, \quad \alpha = 1,2,3, \qquad Z_4 \mathrel{\widehat{=}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & q & 0 \\ 0 & 0 & q^2 \end{pmatrix}. \tag{3.8}$$

We note that $\widehat{\rho}_\alpha = \frac{1}{3}(\mathbb{1} + z_\alpha^* Z_\alpha + z_\alpha Z_\alpha^\dagger)$, and the replacement $\widehat{\varrho}_\beta \mapsto \frac{1}{3}\mathbb{1}$ for an unmeasured basis is here simply implemented by $z_\beta \mapsto 0$.

Consider now a qutrit state for which the density operator is an incoherent mixture of two projectors, one each from the first and the second bases,

$$\varrho = \lambda_1 \Pi_{1j} + \lambda_2 \Pi_{2k}, \tag{3.9}$$

with $0 \leqslant \lambda_1 = 1 - \lambda_2 \leqslant 1$. As an immediate consequence of $\mathrm{Tr}(\Pi_{\alpha k} Z_\beta) = \delta_{\alpha\beta} q^k$, we have here $z_1 = \lambda_1 q^j$, $z_2 = \lambda_2 q^k$, and $z_3 = z_4 = 0$. It follows that $\widehat{\varrho}^{(2)}_{\mathrm{ULIN}} = \widehat{\varrho}^{(3)}_{\mathrm{ULIN}} = \widehat{\varrho}^{(4)}_{\mathrm{ULIN}} = \varrho$, so that all ULIN estimators are genuine density operators for this particular ϱ.

Matters are quite different for

$$\rho = \frac{1}{2}\begin{pmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \tag{3.10}$$

for which $z_1 = z_2 = z_3 = -\frac{1}{2}$ and $z_4 = -\frac{1}{2}q^2$; all $\widehat{\varrho}_\alpha$'s have eigenvalues $0,\frac{1}{2},\frac{1}{2}$. While $\widehat{\varrho}^{(1)}_{\mathrm{ULIN}} = \widehat{\varrho}_1$ and $\widehat{\varrho}^{(4)}_{\mathrm{ULIN}} = \varrho$ are density operators, as they are always, here $\widehat{\varrho}^{(2)}_{\mathrm{ULIN}}$ and $\widehat{\varrho}^{(3)}_{\mathrm{ULIN}}$ are not acceptable estimators; they have negative determinants of $-\frac{1}{27}$ and $-\frac{5}{108}$, respectively.

The last example illustrates the statement in Sec. II C that "the ULIN estimator is not assuredly physical for intermediate M values," here for $1 < M < d + 1 = 4$. Analogous examples can be constructed for higher dimensions; see Sec. III C. Regarding the other assertion in Sec. II C—that there is no consistent "linear map $p_{\alpha k} \mapsto \widehat{\varrho}^{(M)}_{\mathrm{LIN}}$ that could be used instead of $p_{\alpha k} \mapsto \widehat{\varrho}^{(M)}_{\mathrm{ULIN}}$"—we shall now argue that one gets a contradiction when assuming that there is such a map. For this purpose, we look specifically at the case of $M = d = 3$, but the argument can clearly be modified for other dimensions $d > 2$ and other intermediate M values.

Suppose we have measured the probabilities of bases 1, 2, and 3 but lack data for basis 4. Then we know the values of z_1, z_2, and z_3, which determine all off-diagonal elements in $\widehat{\rho}^{(3)}_{\mathrm{LIN}}$, and we need a linear (or affine) map $(z_1,z_2,z_3) \mapsto \hat{z}_4^{(3)}$ for the estimation of the diagonal matrix elements. As an immediate consequence of the linearity, we note that convex sums of (z_1,z_2,z_3) values yield respective convex sums of $\hat{z}_4^{(3)}$ values,

$$\begin{aligned} &\hat{z}_4^{(3)}(\lambda z_1 + \lambda' z_1', \lambda z_2 + \lambda' z_2', \lambda z_3 + \lambda' z_3') \\ &\quad = \lambda \hat{z}_4^{(3)}(z_1,z_2,z_3) + \lambda' \hat{z}_4^{(3)}(z_1',z_2',z_3'), \end{aligned} \tag{3.11}$$

for $0 \leqslant \lambda = 1 - \lambda' \leqslant 1$. Further, for $z_1 = z_2 = z_3 = u$ with $0 < |u| \leqslant \frac{1}{2}$, we have

$$\begin{aligned} \widehat{\rho}^{(3)}_{\mathrm{LIN}} &= \begin{pmatrix} \cdot & u & 0 \\ u^* & \cdot & 0 \\ 0 & 0 & \cdot \end{pmatrix} \\ &= \lambda \begin{pmatrix} \cdot & \frac{1}{2}e^{i\varphi} & 0 \\ \frac{1}{2}e^{-i\varphi} & \cdot & 0 \\ 0 & 0 & \cdot \end{pmatrix} + \lambda' \begin{pmatrix} \cdot & -\frac{1}{2}e^{i\varphi} & 0 \\ -\frac{1}{2}e^{-i\varphi} & \cdot & 0 \\ 0 & 0 & \cdot \end{pmatrix} \end{aligned} \tag{3.12}$$

with $e^{i\varphi} = u/|u|$ and $\lambda = \frac{1}{2} + |u|$, and the dots stand in for the yet-unknown diagonal entries. In the convex sum, both matrices must have $p_{40} = p_{41} = \frac{1}{2}$ and $p_{42} = 0$ on the diagonal, which implies the same values for this $\widehat{\rho}^{(3)}_{\mathrm{LIN}}$. We conclude that

$$\hat{z}_4^{(3)}(u,u,u) = \tfrac{1}{2}(1+q) = -\tfrac{1}{2}q^2, \tag{3.13}$$

which, by convexity, also holds for $u = 0 = \frac{1}{2}u' + \frac{1}{2}(-u')$. Analogous arguments for

$$z_1 = q z_2 = q^2 z_3 = u\ :\quad \widehat{\rho}^{(3)}_{\mathrm{LIN}} = \begin{pmatrix} \frac{1}{2} & 0 & u^* \\ 0 & 0 & 0 \\ u & 0 & \frac{1}{2} \end{pmatrix}, \tag{3.14}$$

and

$$z_1 = q^2 z_2 = q z_3 = u\ :\quad \widehat{\rho}^{(3)}_{\mathrm{LIN}} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \frac{1}{2} & u \\ 0 & u^* & \frac{1}{2} \end{pmatrix}, \tag{3.15}$$

establish

$$\begin{aligned} \hat{z}_4^{(3)}(u,q^2u,qu) &= \tfrac{1}{2}(1+q^2) = -\tfrac{1}{2}q, \\ \hat{z}_4^{(3)}(u,qu,q^2u) &= \tfrac{1}{2}(q+q^2) = -\tfrac{1}{2}. \end{aligned} \tag{3.16}$$

Obviously, there is a contradiction for $z_1 = z_2 = z_3 = 0$, as we cannot have $-\frac{1}{2}$ and $-\frac{1}{2}q$ and also $-\frac{1}{2}q^2$ for $\hat{z}_4^{(3)}(0,0,0)$. Case closed.

C. ULIN estimators for higher dimensions

The $d = 3$ example of Eq. (3.10) is easily generalized to any dimension d. The ULIN estimator for $M = d$,

$$\widehat{\rho}^{(d)}_{\mathrm{ULIN}} = \begin{pmatrix} 1/d & -1/2 & 0 & 0 & \cdots \\ -1/2 & 1/d & 0 & 0 & \cdots \\ 0 & 0 & 1/d & 0 & \cdots \\ 0 & 0 & 0 & 1/d & \\ \vdots & \vdots & \vdots & & \ddots \end{pmatrix} \ngeqslant 0, \tag{3.17}$$

is unphysical for all $d > 2$ because it has the negative eigenvalue $1/d - 1/2$.

Yet another example of unphysical ULIN estimators for $1 < M < d + 1$ is provided by a projector to a superposition of two states from the first basis [the ϱ's of Eqs. (3.10) and (3.17) project on superpositions of two states from the $(d+1)$th basis]; that is,

$$\varrho = (|\psi_{10}\rangle + |\psi_{11}\rangle)\tfrac{1}{2}(\langle\psi_{10}| + \langle\psi_{11}|) = \mathcal{B}_1 \begin{bmatrix} 1/2 & 1/2 & 0 & \cdots \\ 1/2 & 1/2 & 0 & \cdots \\ 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \mathcal{B}_1^\dagger, \tag{3.18}$$

where $\mathcal{B}_1 = (|\psi_{10}\rangle\ |\psi_{11}\rangle \cdots |\psi_{1d-1}\rangle)$ is the row of kets from the first basis, as in Eq. (3.5), and $\mathcal{B}_1^\dagger$ is the adjoint column of bras. Note that we are using square parentheses to denote matrices expressed with the first basis, as opposed to the usual computational basis.

Here,

$$\widehat{\varrho}_1 = \widehat{\varrho}^{(1)}_{\mathrm{ULIN}} = \mathcal{B}_1 \begin{bmatrix} 1/2 & 0 & 0 & \cdots \\ 0 & 1/2 & 0 & \cdots \\ 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \mathcal{B}_1^\dagger \tag{3.19}$$

accounts for all the diagonal elements whereas, for $\alpha = 2,3,\ldots,d+1$, the nonzero matrix elements of

$$\widehat{\varrho}_\alpha - \frac{1}{d}\mathbb{1} = \mathcal{B}_\alpha \begin{bmatrix} w_{\alpha 0} & 0 & \cdots & 0 \\ 0 & w_{\alpha 1} & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & & \cdots\cdot & w_{\alpha\, d-1} \end{bmatrix} \mathcal{B}_\alpha^\dagger = \mathcal{B}_1 \begin{bmatrix} 0 & * & * & \cdots \\ * & 0 & * & \cdots \\ * & * & 0 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \mathcal{B}_1^\dagger. \tag{3.20}$$

Here, the matrix to the αth basis is diagonal with the kth entry equal to $w_{\alpha k} = \mathrm{Re}(\langle\psi_{\alpha k}|\psi_{10}\rangle/\langle\psi_{\alpha k}|\psi_{11}\rangle)$, whereas the matrix to the first basis has null entries on the diagonal and some, if not all, off-diagonal entries are nonzero (indicated by the symbol $*$). Then, for $M = 2,3,\ldots,d$, the $d \times d$ matrix in

$$\widehat{\varrho}^{(M)}_{\mathrm{ULIN}} = \mathcal{B}_1 \begin{bmatrix} 1/2 & * & * & \cdots & * \\ * & 1/2 & * & \cdots & * \\ * & * & 0 & \cdots & * \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ * & * & * & \cdots & 0 \end{bmatrix} \mathcal{B}_1^\dagger \tag{3.21}$$

has Hermitian 2×2 submatrices $\left[\begin{smallmatrix} * & * \\ * & 0 \end{smallmatrix}\right] \ngeqslant 0$ with nonzero off-diagonal elements and one or two vanishing diagonal elements. Such a 2×2 matrix has a negative determinant and cannot be positive semidefinite. It follows that $\widehat{\varrho}^{(M)}_{\mathrm{ULIN}} \ngeqslant 0$ is unphysical for intermediate M values.

For illustration, we take once more the qutrit MUB in Eq. (3.5), for which

$$\widehat{\rho}^{(1)}_{\mathrm{ULIN}} = \frac{1}{2}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \widehat{\rho}^{(2)}_{\mathrm{ULIN}} = \frac{1}{6}\begin{bmatrix} 3 & 1 & q \\ 1 & 3 & q \\ q^2 & q^2 & 0 \end{bmatrix},$$
$$\widehat{\rho}^{(3)}_{\mathrm{ULIN}} = \frac{1}{6}\begin{bmatrix} 3 & 2 & -1 \\ 2 & 3 & -1 \\ -1 & -1 & 0 \end{bmatrix}, \quad \widehat{\rho}^{(4)}_{\mathrm{ULIN}} = \frac{1}{2}\begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \tag{3.22}$$

Indeed, $\widehat{\varrho}^{(2)}_{\mathrm{ULIN}}$ and $\widehat{\varrho}^{(3)}_{\mathrm{ULIN}}$ are unphysical.

In summary, then, we have seen in so many examples that ULIN estimators can be unphysical. The unbiased estimators $\widehat{\varrho}_\beta = \frac{1}{d}\mathbb{1}$ are not permissible for $\beta > M$ and a nonzero "unmeasured" term is needed in Eq. (2.13) for a physical estimator $\widehat{\varrho}^{(M)}_{\mathrm{other}}$. In this situation, the Md measured probabilities $p_{\alpha k}$ with $\alpha \leqslant M$ impose constraints on the $(d+1-M)d$ unmeasured probabilities, which exclude the use of $\widehat{p}_{\beta l} = \frac{1}{d}$ for $\beta > M$. It follows that we have some, if limited, knowledge about the observables associated with the unmeasured bases: our best guess is *not* that all outcomes are equally likely for the future von Neumann test of a basis not measured as yet. Put differently, physical state estimators for $1 < M < d+1$ typically possess some bias in the unmeasured probabilities, such that they deviate from the uninformative uniform distribution. This is the reason why we cannot implement Laplace's advice by the naive estimates of Eq. (2.12): it is simply not true that we have no information at all about the unmeasured probabilities.

Harking back to the qubit estimators in Eqs. (3.3) and (3.4), we observe that these also contain information about the unmeasured bases because $x^2 + y^2 + z^2 \leqslant 1$ must hold for all physical estimators, so a known value for x restricts both y and z, and known values for x and y bound the acceptable z values. Since $\widehat{y} = \widehat{z} = 0$ or $\widehat{z} = 0$, respectively, are always permissible estimates, the intermediate ULIN estimators work for the qubit. This feature, however, has no analog for $d > 2$.

D. Nonpositivity of ULIN estimators

Before moving on to discussing the proper choice for the "unmeasured" contribution to $\widehat{\varrho}^{(M)}_{\mathrm{other}}$ in Eq. (2.13), let us close the subject of negative eigenvalues of intermediate ULIN estimators by a more quantitative study. We ask this question: what is the most negative eigenvalue $\lambda_{\min}$ that $\widehat{\varrho}^{(M)}_{\mathrm{ULIN}}$ can possibly have?

For given ϱ, the smallest eigenvalue of $\widehat{\varrho}^{(M)}_{\mathrm{ULIN}}$ can be found by minimizing $\mathrm{Tr}(\widehat{\varrho}^{(M)}_{\mathrm{ULIN}}\,\sigma)$ over all density operators σ, and further minimization over ϱ establishes

$$\lambda_{\min} = \min_{\varrho,\sigma} \mathrm{Tr}\left(\widehat{\varrho}^{(M)}_{\mathrm{ULIN}}\,\sigma\right) = \min_{\varrho,\sigma} \mathrm{Tr}\left(\varrho\,\widehat{\sigma}^{(M)}_{\mathrm{ULIN}}\right) \geqslant -\frac{M-1}{d}. \tag{3.23}$$

Here, the roles of ϱ and σ are interchangeable thanks to the symmetry noted in Eq. (2.15), and the lower bound follows from Eq. (2.10) in conjunction with $\mathrm{Tr}(\widehat{\varrho}_\alpha\,\sigma) \geqslant 0$ for all α.

The equality holds in Eq. (3.23) only if $\sum_{\alpha=1}^{M}\widehat{\varrho}_\alpha$ is rank deficient, and this is the case for certain (d,M) pairs but not for

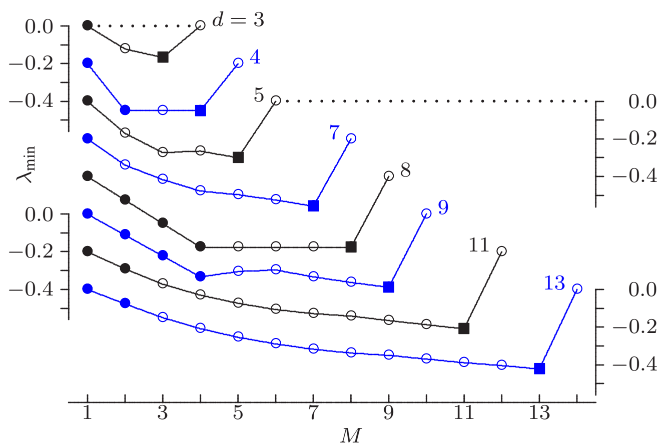

FIG. 1. (Color online) The most negative possible eigenvalue $\lambda_{\min}$ of a ULIN estimator $\widehat{\varrho}^{(M)}_{\mathrm{ULIN}}$ in partial MUB tomography, for prime-power dimensions in the range $3 \leqslant d \leqslant 13$. The successive $\lambda_{\min}$ values of $M = 1,2,\ldots,d+1$ are connected by straight-line segments that guide the eye. The curves are shifted by -0.2 relative to the preceding d value. The vertical-axis labels are shown for $d = 3$ and $d = 9$ on the left, and for $d = 5$ and $d = 13$ on the right. Each plot point represents the optimal value of Eq. (3.23) over 100 repeated numerical searches for solutions of the equation pair (3.24) with the algorithm discussed in the text. Filled circles (•) indicate values with "=" in Eq. (3.23), including $M = 1$ always; empty circles (○) mark values with ">," including $M = d+1$ always. For $M = d$, where ">" is the case, we have the negative eigenvalue $\frac{1}{d} - \frac{1}{2}$ of the example in Eq. (3.17), shown by filled squares (■).

others; see Fig. 1. For example, $\lambda_{\min} = -\frac{1}{4}$ can be established for $(d,M) = (4,2)$ by a simple explicit example. Whether tight or not, the lower bound in Eq. (3.23) tells us that partial MUB tomography with large d and small M is very nearly unbiased. For example, if only two of the MUB are measured, the lowest minimum eigenvalue of the ULIN estimator cannot be smaller than $-1/d$, and would approach zero for large dimensions. Yet, even a small negative eigenvalue renders $\widehat{\varrho}^{(M)}_{\mathrm{ULIN}}$ unphysical if this eigenvalue is different from zero within the numerical accuracy with which it is known.

The extremal density operators ϱ and σ minimizing the traces in Eq. (3.23) are such that σ projects on the smallest eigenvalue of $\widehat{\varrho}^{(M)}_{\mathrm{ULIN}}$ and ϱ projects on the smallest eigenvalue of $\widehat{\sigma}^{(M)}_{\mathrm{ULIN}}$, so that we have the pair of equations

$$\widehat{\varrho}^{(M)}_{\mathrm{ULIN}}\sigma = \mathrm{Tr}\left(\widehat{\varrho}^{(M)}_{\mathrm{ULIN}}\sigma\right)\sigma, \quad \widehat{\sigma}^{(M)}_{\mathrm{ULIN}}\varrho = \mathrm{Tr}\left(\widehat{\sigma}^{(M)}_{\mathrm{ULIN}}\varrho\right)\varrho. \tag{3.24}$$

We solve them by iteration:

S1. For the current ϱ, diagonalize $\widehat{\varrho}^{(M)}_{\mathrm{ULIN}}$ and set σ to the projector onto the smallest eigenvalue.

S2. Diagonalize $\widehat{\sigma}^{(M)}_{\mathrm{ULIN}}$ and update ϱ by setting it to the projector onto the smallest eigenvalue.

S3. Repeat steps S1 and S2 until (3.24) holds with the desired numerical accuracy.

Since we are minimizing a convex function over a convex domain, the possibility of landing at a suboptimal point on the boundary cannot be ruled out. Restarting the numerical search several times with different initial states is necessary.

Figure 1 shows these most negative eigenvalues for the prime-power Hilbert-space dimensions from $d = 3$ to $d = 13$. For $d = 7$ and $M = 2$, for instance, we find that $\lambda_{\min} \approx -0.1394$ while the example of Eq. (3.18) gives a most negative eigenvalue of about -0.1250. Further, we observe the following features:

F1. The most negative eigenvalue is obtained for $M = d$, and is equal to $\frac{1}{d} - \frac{1}{2}$, the negative eigenvalue of $\widehat{\varrho}^{(d)}_{\mathrm{ULIN}}$ in Eq. (3.17). These values are marked by filled squares in Fig. 1.

F2. For $2 \leqslant M \leqslant d$, there are always more M values for which ">" applies in Eq. (3.23) than those for which "=" is the case. These values are marked by empty or filled circles, respectively.

F3. Some deviations from a monotonic decrease of $\lambda_{\min}$ as M increases from 1 to d are observed for some dimensions, notwithstanding the obvious trend.

We leave it as a moot point whether or not these features are also present in dimensions higher than $d = 13$. Should feature F1 be generally true, then

$$\lambda_{\min} \geqslant -\min\left\{\frac{M-1}{d}, \frac{1}{2} - \frac{1}{d}\right\} \tag{3.25}$$

would sharpen the inequality in Eq. (3.23). Currently, this is just a conjecture suggested by the evidence presented in Fig. 1. Whether it holds for all prime-power dimensions d is perhaps of some interest for those who study the properties of MUB. It is, however, of no consequence for quantum state estimation, where the fact that $\lambda_{\min}$ is negative for all intermediate M values matters, while the precise value does not.

IV. LEAST-BIAS MUB INFERENCE

A. Physical unbiased estimators and von Neumann entropy

As we know, if the ULIN estimator $\widehat{\varrho}^{(M)}_{\mathrm{ULIN}}$ is a *bona fide* density operator, it maximizes the Shannon entropy of the unmeasured probabilities in Eq. (2.16). For $d > 2$, only the estimators for $M = 1$ and $M = d+1$ will surely also yield the largest von Neumann entropy of Eq. (3.2). The permissible ULIN estimators for $1 < M < d+1$ do not generally maximize the von Neumann entropy over the convex set of the physical $\widehat{\varrho}^{(M)}_{\mathrm{other}}$'s of Eq. (2.13): they are not the estimators of Ref. [19].

A simple counterexample for $d = 3$ and $M = 2$ is enough to support this statement. We return to Eq. (3.9) and consider the equal-weight case of $\lambda_1 = \lambda_2 = \frac{1}{2}$, so that $z_1 = \frac{1}{2}q^j$, $z_2 = \frac{1}{2}q^k$, and $z_3 = z_4 = 0$. The rank-2 operators $\widehat{\varrho} = \widehat{\varrho}^{(2)}_{\mathrm{ULIN}} = \widehat{\varrho}^{(3)}_{\mathrm{ULIN}} = \widehat{\varrho}^{(4)}_{\mathrm{ULIN}}$ have eigenvalues $0, \frac{1}{6}(3 \pm \sqrt{3})$ and von Neumann entropy $S(\widehat{\varrho}^{(2)}_{\mathrm{ULIN}}) = 0.5157$. The largest value of the von Neumann entropy is 0.6370, however; we obtain it for the $\widehat{\varrho}^{(2)}_{\mathrm{other}}$ with $q^{j+k}\widehat{z}^{(2)}_3 = q^{j-k+1}\widehat{z}^{(2)}_4 = -0.094\,66$.

It is, of course, hardly surprising that the Shannon and the von Neumann entropy are usually maximized by different $\widehat{\varrho}^{(M)}_{\mathrm{other}}$'s. These are two different figures of merit, which serve different purposes: if we intend to measure the remaining bases of the set of MUB, the Shannon entropy is a useful quantity; by contrast, if we want to have a more universally applicable estimator, then maximizing the von Neumann entropy is a time-honored approach in Jaynes's spirit [51–53]. That the two procedures coincide in the qubit case is nothing more than a result of the geometrical simplicity of the problem.

B. Physical estimators with least bias

In the context of incomplete MUB tomography, it is clearly more natural to choose the "unmeasured" contribution in Eq. (2.13) such that the Shannon entropy is maximized rather than the von Neumann entropy. For, the maximization of the Shannon entropy ensures that we accept $\widehat{\varrho}^{(M)}_{\mathrm{ULIN}}$ as the estimator whenever that is permissible—that is, whenever $\widehat{\varrho}^{(M)}_{\mathrm{ULIN}} \geqslant 0$—and when that is not the case, we stay as close to naive unbiasedness as possible (more about this closeness in Sec. V). In other words, it is not necessary to completely discard the notion of unbiasedness; a certain refinement is called for: we opt for the *least-bias* estimator $\widehat{\varrho}^{(M)}_{\mathrm{LB}}$, which is the physical $\widehat{\varrho}^{(M)}_{\mathrm{other}}$ with the largest Shannon entropy for the unmeasured probabilities,

$$\max_{\widehat{\varrho}^{(M)}_{\mathrm{other}}\geqslant 0} \mathrm{H}^{(M)}\big(\widehat{\varrho}^{(M)}_{\mathrm{other}}\big) = \mathrm{H}^{(M)}\big(\widehat{\varrho}^{(M)}_{\mathrm{LB}}\big). \tag{4.1}$$

We repeat, perhaps unnecessarily, that the least-bias estimator is different from the ULIN estimator only if the latter is unphysical.

C. An efficient algorithm

The numerical search for $\widehat{\varrho}^{(M)}_{\mathrm{LB}}$ requires the maximization of $\mathrm{H}^{(M)}(\varrho)$ over the convex set of physical $\widehat{\varrho}^{(M)}_{\mathrm{other}}$'s, that is, over the set of all density operators ϱ with $p_{\alpha k} = \mathrm{Tr}(\varrho\Pi_{\alpha k})$ for $\alpha = 1,2,\ldots,M$. Other than that, ϱ is only constrained by the defining properties of a density operator: $\varrho \geqslant 0$ and $\mathrm{Tr}\,\varrho = 1$. This optimization problem is very similar to the one solved in Ref. [19], where one maximizes the von Neumann entropy over the convex set of maximum-likelihood estimators. Indeed, the algorithm of Ref. [19] can be modified so that it applies to the current problem of finding $\widehat{\varrho}^{(M)}_{\mathrm{LB}}$.

The figure of merit is the following function of ϱ:

$$D_\mu(\varrho) = \sum_{\alpha=1}^{M}\sum_{k=0}^{d-1} p_{\alpha k}\ln[\mathrm{Tr}(\varrho\Pi_{\alpha k})] - \mu\sum_{\beta=M+1}^{d+1}\sum_{l=0}^{d-1}\mathrm{Tr}(\varrho\Pi_{\beta l})\ln[\mathrm{Tr}(\varrho\Pi_{\beta l})]. \tag{4.2}$$

The first term is a mock log-likelihood for relative frequencies equal to the known probabilities $p_{\alpha k}$ for $\alpha \leqslant M$. The second term is the Shannon entropy $\mathrm{H}^{(M)}(\varrho)$ of multiplied by the non-negative parameter μ.

Now, all physical $\widehat{\varrho}^{(M)}_{\mathrm{other}}$'s maximize the $\mu = 0$ function $D_0(\varrho)$, with

$$\max_{\varrho} D_0(\varrho) = D_0\big(\widehat{\varrho}^{(M)}_{\mathrm{other}}\big) = \sum_{\alpha=1}^{M}\sum_{k=0}^{d-1} p_{\alpha k}\ln p_{\alpha k}. \tag{4.3}$$

Because of this extremal property of $D_0(\widehat{\varrho}^{(M)}_{\mathrm{other}})$, the first-order contribution to $D_\mu(\varrho)$ originates solely in the Shannon-entropy term,

$$\max_{\varrho} D_\mu(\varrho) = D_0\big(\widehat{\varrho}^{(M)}_{\mathrm{other}}\big) + \mu\mathrm{H}^{(M)}\big(\widehat{\varrho}^{(M)}_{\mathrm{LB}}\big) + \cdots, \tag{4.4}$$

where the ellipsis stands for contributions of order μ^2 and higher. Therefore, the numerical maximization of $D_\mu(\varrho)$ for sufficiently small μ yields the least-bias estimator $\widehat{\varrho}^{(M)}_{\mathrm{LB}}$. Practical experience shows that $\mu \approx 10^{-4}$ is large enough for a noticeable difference between $D_\mu(\varrho)$ and $D_0(\varrho)$ and small enough to ensure that the density operator that maximizes $D_\mu(\varrho)$ approximates $\widehat{\varrho}^{(M)}_{\mathrm{LB}}$ very well. Usually, it is a good idea to numerically select a μ that minimizes the fluctuations in the extremal ϱ—from the iteration described below—to some prechosen precision.

The first-order response of $D_\mu(\varrho)$ to a variation $\varrho \mapsto \varrho + \delta\varrho$ is

$$\delta D_\mu(\varrho) = \mathrm{Tr}[\delta\varrho\, W(\varrho)], \tag{4.5}$$

with

$$W(\varrho) = \sum_{\alpha=1}^{M}\sum_{k=0}^{d-1}\frac{p_{\alpha k}}{\mathrm{Tr}(\varrho\Pi_{\alpha k})}\Pi_{\alpha k} - \mu\sum_{\beta=M+1}^{d+1}\sum_{l=0}^{d-1}\Pi_{\beta l}\ln[\mathrm{Tr}(\varrho\Pi_{\beta l})]. \tag{4.6}$$

Hence, the positivity constraint for ϱ implies the following extremal equation for $\widehat{\varrho}^{(M)}_{\mathrm{LB}}$:

$$W\big(\widehat{\varrho}^{(M)}_{\mathrm{LB}}\big)\widehat{\varrho}^{(M)}_{\mathrm{LB}} = \mathrm{Tr}\big[W\big(\widehat{\varrho}^{(M)}_{\mathrm{LB}}\big)\widehat{\varrho}^{(M)}_{\mathrm{LB}}\big]\widehat{\varrho}^{(M)}_{\mathrm{LB}}. \tag{4.7}$$

Since $D_\mu(\varrho)$ has no local maxima, we can find $\widehat{\varrho}^{(M)}_{\mathrm{LB}}$ by steepest ascent ("follow the gradient uphill"). The $W(\varrho)$ identifies the gradient in the sense that

$$\delta\varrho \propto W(\varrho)\varrho + \varrho W(\varrho) - 2\varrho\,\mathrm{Tr}[\varrho W(\varrho)] \tag{4.8}$$

gives the largest first-order change in $D_\mu(\varrho)$. We ensure a positive increment of $D_\mu(\varrho)$ by the iteration in accordance with

$$\varrho_n \mapsto \varrho_{n+1} = \frac{(\mathbb{1} + \epsilon\{W(\varrho_n) - \mathrm{Tr}[W(\varrho_n)\varrho_n]\mathbb{1}\})\,\varrho_n(\mathbb{1} + \epsilon\{W(\varrho_n) - \mathrm{Tr}[W(\varrho_n)\varrho_n]\mathbb{1}\})}{1 + \epsilon^2\,\mathrm{Tr}(\{W(\varrho_n) - \mathrm{Tr}[W(\varrho_n)\varrho_n]\mathbb{1}\}^2\varrho_n)}, \tag{4.9}$$

where ϵ is a small positive step size. To obtain the extremal solution $\widehat{\varrho}^{(M)}_{\mathrm{LB}}$, one can start with the maximally mixed state $\varrho_{n=1} = \frac{1}{d}\mathbb{1}$ and iterate Eq. (4.9) for a small μ until the extremal equations (4.7) are obeyed to a satisfactory accuracy.

We note, but do not elaborate this point, that one can speed up the convergence substantially by employing conjugate gradients [54]; each iteration step is then more costly (in CPU time) but that is more than compensated for by the

TABLE I. Comparison between the ULIN and least-bias estimators that maximize the Shannon entropy for ϱ_w. When ϱ_w is the pure state defined by Eq. (3.10) ($w = 0$), the least-bias estimator is the true state ϱ. Also, note that for ϱ_w, $z_1 = z_2 = z_3 = -(1-w)/2$, $z_4 = -(1-w)q^2/2$.

	$w = 0.1$		$w = 0.2$	
	z_3	z_4	z_3	z_4
ϱ_w	-0.450	$0.225 + i\,0.389$	-0.400	$0.200 + i\,0.346$
$\widehat{\varrho}^{(2)}_{\mathrm{ULIN}}$	0	0	0	0
$\widehat{\varrho}^{(2)}_{\mathrm{LB}}$	-0.313	$0.156 + i\,0.271$	-0.126	$0.063 + i\,0.109$
$\widehat{\varrho}^{(3)}_{\mathrm{ULIN}}$	-0.450	0	-0.400	0
$\widehat{\varrho}^{(3)}_{\mathrm{LB}}$	-0.450	$0.174 + i\,0.303$	-0.400	$0.100 + i\,0.173$

much smaller number of steps. Further we note that, in cases where the iteration algorithm proceeds too slowly due to the complexity of the optimization problem for large Hilbert-space dimensions, the sum of the two entropic functions in Eq. (4.2) can alternatively be optimized with gradient-free methods, such as the Nelder-Mead "amoeba" method [55] or simulated annealing [56].

D. Examples

For explicit examples, we look at a class of qutrit states comprising statistical mixtures of ϱ as represented in (3.10) with the maximally mixed state for various admixtures $0 \leqslant w \leqslant 1$:

$$\varrho \mapsto \varrho_w = (1-w)\varrho + w\frac{\mathbb{1}}{3}. \tag{4.10}$$

As always, for $M = 1$ and $M = 4$, the ULIN estimator is always positive semidefinite. Beyond $w = 0.2679$ and $w = \frac{1}{3}$ respectively for $M = 2$ and $M = 3$, the ULIN estimator is precisely the physical least-bias estimator that maximizes the Shannon entropy since the true state would be highly mixed. For other w values, the ULIN estimator possesses at least one negative eigenvalue. The relevant parameters for the unphysical ULIN and physical least-bias estimators for two exemplifying values of w are listed in Table I.

V. COMPARISON WITH OTHER ESTIMATORS

A. Estimators of the least-bias kind

The Shannon entropy quantifies our ignorance about the outcomes of future projective measurements in the yet-unmeasured bases with $\beta = M+1, M+2, \ldots, d+1$ from the set of MUB. Rather than ignorance we can equivalently quantify knowledge by the *predictability* of the future measurements. For an experiment with probabilities $p_0, p_1, \ldots, p_{d-1}$, the entropic measure of predictability is

$$P_{\mathrm{ent}}(p_.) = \sum_{l=0}^{d-1} p_l \log_d(p_l d), \tag{5.1}$$

where the symbol $p_.$ stands for all the d probabilities. We have the extreme values

$$P_{\mathrm{ent}}(p_.) = \begin{cases} 0 & \text{if } p_0 = p_1 = \cdots = p_{d-1} = \frac{1}{d}, \\ 1 & \text{if } p_l = \delta_{l\bar{l}} \quad \text{for a certain } \bar{l}, \end{cases} \tag{5.2}$$

and $0 < P_{\mathrm{ent}}(p_.) < 1$ for all other sets of probabilities, which are among the defining properties of all permissible measures of predictability [57,58]. Convexity

$$P(\lambda p_. + \lambda' p'_.) \leqslant \lambda P(p_.) + \lambda' P(p'_.) \tag{5.3}$$

for $0 \leqslant \lambda = 1 - \lambda' \leqslant 1$ is another important property of all predictability measures.

In terms of $P_{\mathrm{ent}}(p_.)$, the Shannon entropy is

$$\mathrm{H}^{(M)}(\varrho) = \left[d + 1 - M - \sum_{\beta=M+1}^{d+1} P_{\mathrm{ent}}(\mathrm{Tr}(\varrho\Pi_{\beta.})) \right] \ln d, \tag{5.4}$$

so that we could replace the figure of merit in Eq. (4.2) by

$$D_\mu(\varrho) = \sum_{\alpha=1}^{M}\sum_{k=0}^{d-1} p_{\alpha k} \ln\left(\mathrm{Tr}(\varrho\Pi_{\alpha k})\right) - \mu \sum_{\beta=M+1}^{d+1} P_{\mathrm{ent}}(\mathrm{Tr}(\varrho\Pi_{\beta.})), \tag{5.5}$$

with a corresponding minor change in Eq. (4.6), namely $\mu \mapsto \mu/\ln d$.

Other measures of predictability could be employed as well. For example, the purity-based predictability

$$P_{\mathrm{pur}}(p_.) = \frac{d}{d-1}\sum_{l=0}^{d-1}\left(p_l - \frac{1}{d}\right)^2. \tag{5.6}$$

Upon replacing $P_{\mathrm{ent}}(\ )$ in (5.5) by $P_{\mathrm{pur}}(\ )$, the $\widehat{\varrho}^{(M)}_{\mathrm{other}}$ that maximizes $D_\mu(\varrho)$ is simply another least-bias estimator $\widehat{\varrho}^{(M)}_{\mathrm{LB}}$ than the one for $P_{\mathrm{ent}}(\ )$. The two least-bias estimators are different since they refer to different quantifications of the bias regarding the probabilities for the not-yet-measured bases. But neither $\widehat{\varrho}^{(M)}_{\mathrm{LB}}$ is better than the other; they serve different purposes. Any other predictability $P(p_.)$ yields a corresponding least-bias estimator. Our preference, in Sec. IV, for the $\widehat{\varrho}^{(M)}_{\mathrm{LB}}$ associated with $P_{\mathrm{ent}}(p_.)$ is mostly for historical reasons—history of the subject, that is.

Yet another example is provided by the predictability functions that equal the expected gain for various strategies of betting on the future outcome. There is, in particular, the "linear bet" of Sec. 2.3.1 in Ref. [58], which amounts to

$$P_{\mathrm{bet}}(p_.) = \max\{p_0, p_1, p_2\} - \min\{p_0, p_1, p_2\} \tag{5.7}$$

in the qutrit case.

We can regard $P_{\mathrm{pur}}(p_.)$ in (5.6) as the normalized squared Hilbert-Schmidt distance of $p_0, p_1, \ldots, p_{d-1}$ from the uniform distribution $p_l = \frac{1}{d}$. Likewise, $P_{\mathrm{ent}}(p_.)$ is the normalized Kullback-Leibler divergence of the actual probabilities from the uniform distribution. Analogous remarks apply to any other permissible predictability $P(p_.)$. It follows that, irrespective of which $P(p_.)$ we choose, the resulting least-bias estimator equals the ULIN estimator whenever $\widehat{\varrho}^{(M)}_{\mathrm{ULIN}} \geqslant 0$. And when $\widehat{\varrho}^{(M)}_{\mathrm{ULIN}} \ngeqslant 0$, then—as a consequence of (5.3)—each $\widehat{\varrho}^{(M)}_{\mathrm{LB}}$ sits on the border of the convex set of physical $\widehat{\varrho}^{(M)}_{\mathrm{other}}$'s, whatever border point is closest to $\widehat{\varrho}^{(M)}_{\mathrm{ULIN}}$ in the sense specified by the choice of $P(p_.)$.

These matters are illustrated in Fig. 2, which graphically shows the situation for qutrit MUB where $M = 3$ bases are measured. As it turns out, numerical experience shows that

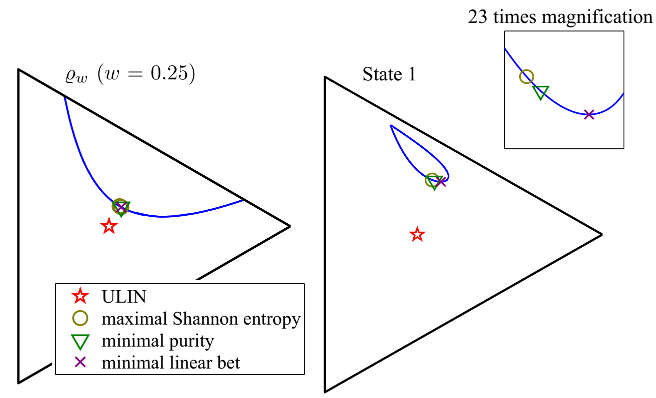

FIG. 2. (Color online) A prototypical showcase of three types of physical least-bias estimators relative to the ULIN estimator for different qutrit true states ($M = 3$). The black triangle is the regular simplex for z_4 in the complex plane, with its corners at $z_4 = 1, q$, and q^2. The blue curve borders the regions of permissible values for z_4. The examples featured here are ϱ_w for $w = \frac{1}{4}$ and state 1 of Table II. The magnified view is for state 1, where the three least-bias estimators are different; by contrast, all three are the same for ϱ_w.

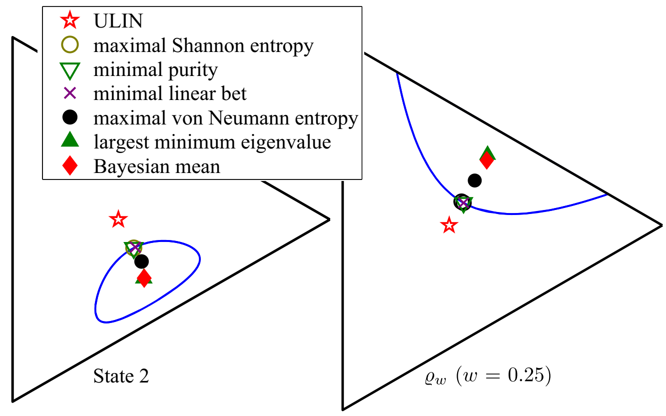

FIG. 4. (Color online) Pictorial comparison of the ULIN estimator with three physical full-rank estimators: the estimators with largest von Neumann entropy and largest minimal eigenvalue, and the Bayesian mean estimator. The rank-deficient three least-bias estimators of Fig. 2 are also indicated; they sit on the blue border of the region of permissible states and are too close to each other for telling them apart. As in Fig. 2, the examples are for qutrit states with $M = 3$, and the triangle is the simplex for z_4. The examples featured here are state 2 and ϱ_w for $w = \frac{1}{4}$ of Table II.

different types of least-bias estimators give estimators that are quite close to each other, hinting that the choice of measure for quantifying bias typically influences the resulting estimator only mildly as long as these nonpathological choices lead to the same extremal solution in the absence of the quantum positivity constraint. Such differences would not matter much in the presence of statistical fluctuations and other experimental error sources. For various qutrit states, the two kinds of least-bias estimators introduced in this section are essentially identical. For completeness, in Fig. 3 we present a comparison of the different figures of merit for the least-bias kind.

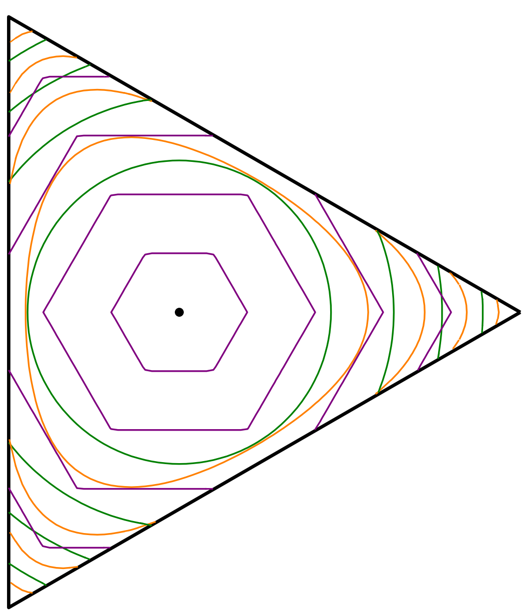

FIG. 3. (Color online) A comparison of the lines of constant "distance" from the ULIN estimator for the three different predictability functions that are used as figure of merit for the least-bias estimators in Figs. 2 and 4. We have $P = 0$ at the center of the triangle and $P = 1$ in each corner, and the contour lines indicate predictability values $P = 0.2, 0.4, 0.6$, and 0.8, respectively. The green circles are for the purity measure of Eq. (5.6); the orange distorted circles are for the Shannon-entropy measure of Eq. (5.1), and the purple hexagons are for the betting measure of Eq. (5.7).

B. Other estimators

The least-bias estimators thus constructed are rank-deficient whenever $\widehat{\varrho}_{\rm ULIN}^{(M)} \ngeq 0$. It may, however, be desirable to use

TABLE II. Values of z_4 for physical estimators to three unphysical ULIN estimators for qutrit states with three measured bases ($d = M = 3$). These data are used in Figs. 2 and 4. The values of z_1, z_2, and z_3 that specify the respective $\widehat{\varrho}_{\rm ULIN}^{(3)}$'s are (i) $z_1 = z_2 = z_3 = -3/8$ for $\varrho_{w=1/4}$; (ii) $z_1 = 0.160 - i\,0.321$, $z_2 = 0.571 - i\,0.192$, and $z_3 = 0.314 + i\,0.165$ for state 1; and (iii) $z_1 = -0.345 + i\,0.0574$, $z_2 = 0.303 + i\,0.328$, and $z_3 = 0.00057 - i\,0.294$ for state 2. The rows report the corresponding three z_4 values for the least-bias estimator $\widehat{\varrho}_{\rm LB}^{(3)}$ of Secs. IV B and IV C for the alternative least-bias estimators of Sec. V A $\widehat{\varrho}_{\rm PUR}^{(3)}$ and $\widehat{\varrho}_{\rm BET}^{(3)}$; for the full-rank estimators with largest von Neumann entropy and largest minimal eigenvalue, $\widehat{\varrho}_{\rm VN}^{(3)}$ and $\widehat{\varrho}_{\rm MINEIG}^{(3)}$; and for the Bayesian mean estimator $\widehat{\varrho}_{\rm BM}^{(3)}$. For information, the last row contains the z_4 values of the actual states that were used to obtain the values of z_1, z_2, z_3 for the $\widehat{\varrho}_{\rm ULIN}^{(3)}$'s.

	$\varrho_{w=\frac{1}{4}}$	State 1	State 2
	z_4	z_4	z_4
$\widehat{\varrho}_{\rm LB}^{(3)}$	$0.067 + i\,0.106$	$0.080 + i\,0.299$	$0.073 - i\,0.136$
$\widehat{\varrho}_{\rm PUR}^{(3)}$	$0.067 + i\,0.106$	$0.093 + i\,0.295$	$0.073 - i\,0.136$
$\widehat{\varrho}_{\rm BET}^{(3)}$	$0.067 + i\,0.106$	$0.128 + i\,0.289$	$0.080 - i\,0.132$
$\widehat{\varrho}_{\rm VN}^{(3)}$	$0.120 + i\,0.208$	$0.090 + i\,0.309$	$0.104 - i\,0.204$
$\widehat{\varrho}_{\rm MINEIG}^{(3)}$	$0.187 + i\,0.325$	$0.003 + i\,0.438$	$0.122 - i\,0.283$
$\widehat{\varrho}_{\rm BM}^{(3)}$	$0.176 + i\,0.305$	$0.021 + i\,0.418$	$0.122 - i\,0.279$
Actual state	$\frac{3}{16} + i\,0.324$	$-0.146 + i\,0.601$	$0.256 - i\,0.337$

estimators that usually have full rank—they are more robust in the sense that a slight perturbation, or inaccuracy in determining them, does not render them unphysical. Other reasons for avoiding rank-deficient estimators have been put forward as well [59].

From the plethora of physical $\widehat{\varrho}_{\text{other}}^{(M)}$'s, we consider three choices: the estimator with the largest von Neumann entropy [19], the Bayesian mean estimator for a uniform prior on the unmeasured probabilities [60], and the estimator having the largest minimal eigenvalue [61,62].

These estimators simply differ in the unmeasured probabilities, as is the case for any estimator consistent with the probabilities of the already-measured bases. To streamline presentations, Fig. 4 and Table II illustrate the estimators for the $M = 3$ case of a qutrit.

VI. CONCLUSIONS

In summary, we have examined whether the ideal Laplace-type linear estimator that assigns equal and unbiased probabilities to all the outcomes of unmeasured bases from the set of MUB is a physical thing to do in incomplete MUB tomography. The answer is definite: such unbiased estimation does not work as a general recipe.

As a natural adjustment to the original inference method that follows a blind application of Laplace's notion of indifference, we recommend the use of the least-bias estimator. It gives the closest-to-uniform unmeasured probability distributions by maximizing the Shannon entropy with due attention to the constraints imposed by the measured probabilities. We also supply a simple iterative algorithm for computing the least-bias estimator.

We compared this least-bias estimator with alternative estimators that quantify the bias differently, and concluded that these alternatives are equally useful for practical purposes. Since all these least-bias estimators are rank-deficient whenever they are different from the linear estimator of ideal Laplace type, we also took a look at three different estimators of full rank. As illustrated by examples, the full-rank estimators are indeed different from the least-bias estimators, but are certainly acceptable as consistent *bona fide* estimators.

Armed with these insights, one could now study questions such as *after measuring in M bases, which basis is optimal for the next von Neumann test?* This matter and others are, however, beyond the scope of the present paper.

ACKNOWLEDGMENTS

Many of the ideas in this paper originated at the Workshop on Mathematical Methods of Quantum Tomography at Fields Institute (Toronto). Z.H., J.R., and Y.S.T. are grateful for the support of the European Social Fund and the state budget of the Czech Republic [Project No. CZ.1.07/2.3.00/30.0004 (POST-UP)], the Grant Agency of the Czech Republic (Grant No. 15-031945), the IGA Project of the Palacký University (Grant No. IGA PrF 2015-002), and the sustainability of post-doc positions at Palacký University. L.L. S.S. acknowledges the support from UCM-Banco Santander Program (Grant No. GR3/14) and helpful discussions with M. Grassl. H.K.N.'s, J.H.C.'s, and B.-G.E.'s work was funded by the Singapore Ministry of Education (partly through the Academic Research Fund Tier 3 MOE2012-T3-1-009) and the National Research Foundation of Singapore. H.K.N. was also funded by a Yale-NUS College start-up grant.

[1] *Quantum State Estimation*, Lecture Notes in Physics Vol. 649, edited by M. G. A. Paris and J. Řeháček (Springer, Berlin, 2004).
[2] E. Prugovečki, Information-theoretical aspects of quantum measurement, Int. J. Theor. Phys. **16**, 321 (1977).
[3] P. Busch and P. J. Lahti, The determination of the past and the future of a physical system in quantum mechanics, Found. Phys. **19**, 633 (1989).
[4] G. M. D'Ariano, P. Perinotti, and M. F. Sacchi, Informationally complete measurements and group representation, J. Opt. B: Quantum Semiclassical Opt. **6**, S487 (2004).
[5] S. T. Flammia, A. Silberfarb, and C. M. Caves, Minimal informationally complete measurements, Found. Phys. **35**, 1985 (2005).
[6] S. Weigert, Simple minimal informationally complete POVMs for qudits, Int. J. Mod. Phys. B **20**, 1942 (2006).
[7] T. Monz, P. Schindler, J. T. Barreiro, M. Chwalla, D. Nigg, W. A. Coish, M. Harlander, W. Hänsel, M. Hennrich, and R. Blatt, 14-Qubit Entanglement: Creation and Coherence, Phys. Rev. Lett. **106**, 130506 (2011).
[8] X.-C. Yao, T.-X. Wang, P. Xu, H. Lu, G.-S. Pan, X.-H. Bao, C.-Z. Peng, C.-Y. Lu, Y.-A. Chen, and J.-W. Pan, Observation of eight-photon entanglement, Nat. Photon. **6**, 225 (2012).
[9] D. Gross, Y. K. Liu, S. T. Flammia, S. Becker, and J. Eisert, Quantum State Tomography via Compressed Sensing, Phys. Rev. Lett. **105**, 150401 (2010).
[10] S. T. Flammia, D. Gross, Y.-K. Liu, and J. Eisert, Quantum tomography via compressed sensing: error bounds, sample complexity and efficient estimators, New J. Phys. **14**, 095022 (2012).
[11] M. Guta, T. Kypraios, and I. Dryden, Rank-based model selection for multiple ions quantum tomography, New J. Phys. **14**, 105002 (2012).
[12] M. Cramer, M. B. Plenio, S. T. Flammia, R. Somma, D. Gross, S. D. Bartlett, O. Landon-Cardinal, D. Poulin, and Y. K. Liu, Efficient quantum state tomography, Nat. Commun. **1**, 149 (2010).
[13] T. Baumgratz, D. Gross, M. Cramer, and M. B. Plenio, Scalable Reconstruction of Density Matrices, Phys. Rev. Lett. **111**, 020401 (2013).
[14] O. Landon-Cardinal and D. Poulin, Practical learning method for multi-scale entangled states, New J. Phys. **14**, 085004 (2012).
[15] G. M. D'Ariano, L. Maccone, and M. Paini, Spin tomography, J. Opt. B: Quantum Semiclassical Opt. **5**, 77 (2003).
[16] G. Tóth, W. Wieczorek, D. Gross, R. Krischek, C. Schwemmer, and H. Weinfurter, Permutationally Invariant Quantum Tomography, Phys. Rev. Lett. **105**, 250403 (2010).
[17] A. B. Klimov, G. Björk, and L. L. Sánchez-Soto, Optimal quantum tomography of permutationally invariant qubits, Phys. Rev. A **87**, 012109 (2013).

[18] T. Moroder, P. Hyllus, G. Tóth, C. Schwemmer, A. Niggebaum, S. Gaile, O. Gühne, and H. Weinfurter, Permutationally invariant state reconstruction, New J. Phys. **14**, 105001 (2012).

[19] Y. S. Teo, H. Zhu, B.-G. Englert, J. Řeháček, and Z. Hradil, Quantum-State Reconstruction by Maximizing Likelihood and Entropy, Phys. Rev. Lett. **107**, 020404 (2011).

[20] Y. S. Teo, B. Stoklasa, B.-G. Englert, J. Řeháček, and Z. Hradil, Incomplete quantum state estimation: A comprehensive study, Phys. Rev. A **85**, 042317 (2012).

[21] Y. S. Teo, B.-G. Englert, J. Řeháček, Z. Hradil, and D. Mogilevtsev, Verification of state and entanglement with incomplete tomography, New J. Phys. **14**, 105020 (2012).

[22] D. S. Gonçalves, C. Lavor, M. A. Gomes-Ruggiero, A. T. Cesário, R. O. Vianna, and T. O. Maciel, Quantum state tomography with incomplete data: Maximum entropy and variational quantum tomography, Phys. Rev. A **87**, 052140 (2013).

[23] W. K. Wootters and B. D. Fields, Optimal state-determination by mutually unbiased measurements, Ann. Phys. **191**, 363 (1989).

[24] J. Schwinger, Unitary operator bases, Proc. Natl. Acad. Sci. U. S. A. **46**, 570 (1960).

[25] T. Durt, B.-G. Englert, I. Bengtsson, and K. Życzkowski, On mutually unbiased bases, Int. J. Quantum Inf. **8**, 535 (2010).

[26] Z. Hradil, J. Summhammer, and H. Rauch, Quantum tomography as normalization of incompatible observations, Phys. Lett. A **261**, 20 (1999).

[27] R. D. Cousins, Why isn't every physicist a Bayesian?, Am. J. Phys. **63**, 398 (1995).

[28] E. T. Jaynes, in *Probability Theory: The Logic of Science*, edited by G. L. Bretthorst (Cambridge University Press, Cambridge, England, 2003).

[29] I. D. Ivanovic, Geometrical description of quantal state determination, J. Phys. A **14**, 3241 (1981).

[30] S. Bandyopadhyay, P. O. Boykin, V. Roychowdhury, and F. Vatan, A new proof for the existence of mutually unbiased bases, Algorithmica **34**, 512 (2002).

[31] A. Klappenecker and M. Rötteler, Constructions of mutually unbiased bases, in *Finite Fields and Applications*, Lecture Notes in Computer Science Vol. 2948, edited by G. Mullen, A. Poli, and H. Stichtenoth (Springer, Berlin, 2003), pp. 137–144.

[32] J. Lawrence, Mutually unbiased bases and trinary operator sets for *N* qutrits, Phys. Rev. A **70**, 012302 (2004).

[33] A. O. Pittenger and M. H. Rubin, Mutually unbiased bases, generalized spin matrices and separability, Linear Algebra Appl. **390**, 255 (2004).

[34] K. R. Parthasarathy, On estimating the state of a finite level quantum system, Infin. Dimens. Anal. Quantum Probab. Relat. Top. **7**, 607 (2004).

[35] P. Wocjan and T. Beth, New construction of mutually unbiased basis in square dimensions, Quantum Inf. Compu. **5**, 93 (2005).

[36] T. Durt, About mutually unbiased bases in even and odd prime power dimensions, J. Phys. A **38**, 5267 (2005).

[37] A. B. Klimov, L. L. Sánchez-Soto, and H. de Guise, Multicomplementary operators via finite Fourier transform, J. Phys. A **38**, 2747 (2005).

[38] A. B. Klimov, J. L. Romero, G. Björk, and L. L. Sánchez-Soto, Geometrical approach to mutually unbiased bases, J. Phys. A **40**, 3987 (2007).

[39] A. B. Klimov, J. L. Romero, G. Björk, and L. L. Sánchez-Soto, Discrete phase-space structure of n-qubit mutually unbiased bases, Ann. Phys. **324**, 53 (2009).

[40] P. S. Laplace, Memoire sur la probabilité des causes par les évènements, Mem. Acad. R. Sci. Paris **6**, 621 (1774).

[41] C. Schwemmer, L. Knips, D. Richart, H. Weinfurter, T. Moroder, M. Kleinmann, and O. Gühne, Systematic Errors in Current Quantum State Tomography Tools, Phys. Rev. Lett. **114**, 080403 (2015).

[42] J. Shang, H. K. Ng, and B.-G. Englert, Quantum state tomography: Mean squared error matters, bias does not, arXiv:1405.5350.

[43] G. Zauner, Quantum designs: Foundations of a noncommutative design theory, Int. J. Quantum Inf. **9**, 445 (2011).

[44] P. Butterley and W. Hall, Numerical evidence for the maximum number of mutually unbiased bases in dimension six, Phys. Lett. **A369**, 5 (2007).

[45] I. Bengtsson, W. Bruzda, Å. Ericsson, J.-Å. Larsson, W. Tadej, and K. Życzkowski, Mutually unbiased bases and Hadamard matrices of order six, J. Math. Phys. **48**, 052106 (2007).

[46] S. Brierley and S. Weigert, Maximal sets of mutually unbiased quantum states in dimension 6, Phys. Rev. A **78**, 042312 (2008).

[47] S. Brierley and S. Weigert, Constructing mutually unbiased bases in dimension six, Phys. Rev. A **79**, 052316 (2009).

[48] P. Jaming, M. Matolcsi, P. Móra, F. Szöllősi, and M. Weiner, A generalized Pauli problem and an infinite family of MUB-triplets in dimension 6, J. Phys. A: Math. Theor. **42**, 245305 (2009).

[49] P. Mandayam, S. Bandyopadhyay, M. Grassl, and W. K. Wootters, Unextendible mutually unbiased bases from Pauli classes, Quantum Inf. Comput. **14**, 0823 (2014).

[50] P. Raynal, X. Lü, and B.-G. Englert, Mutually unbiased bases in six dimensions: The four most distant bases, Phys. Rev. A **83**, 062303 (2011).

[51] E. T. Jaynes, Information theory and statistical mechanics, Phys. Rev. **106**, 620 (1957).

[52] E. T. Jaynes, Information theory and statistical mechanics II, Phys. Rev. **108**, 171 (1957).

[53] V. Bužek, Quantum tomography from incomplete data via *MaxEnt* principle, in *Quantum State Estimation*, Lecture Notes in Physics Vol. 649, edited by M. G. A. Paris and J. Řeháček (Springer, Berlin, 2004), pp. 189–234.

[54] J. R. Shewchuk, An Introduction to the Conjugate Gradient Method Without the Agonizing Pain, Technical Report No. 865018, Carnegie Mellon University, 1994.

[55] J. A. Nelder and R. Mead, A simplex method for function minimization, Comput. J. **7**, 308 (1965).

[56] P. J. M. van Laarhoven and E. H. L. Aarts, *Simulated Annealing: Theory and Applications*, Mathematics and Its Applications (Springer, Berlin, 1985).

[57] S. Dürr, Quantitative wave-particle duality in multibeam interferometers, Phys. Rev. A **64**, 042113 (2001).

[58] B. G. Englert, D. Kaszlikowski, L. C. Kwek, and W. H. Chee, Wave-particle duality in multi-path interferometers: General

concepts and three-path interferometers, Int. J. Quantum Inf. **6**, 129 (2008).

[59] R. Blume-Kohout, Hedged Maximum Likelihood Quantum State Estimation, Phys. Rev. Lett. **105**, 200504 (2010).

[60] R. Blume-Kohout, Optimal, reliable estimation of quantum states, New J. Phys. **12**, 043034 (2010).

[61] A. D. Wilson and T. D. Murphey, Local E-optimality conditions for trajectory design to estimate parameters in nonlinear systems, Proc. Am. Control Conf. 6/2014 (2014), p. 443.

[62] A. S. Hedayat, Study of optimality criteria in design of experiments, in *Statistics and Related Topics*, edited by M. Csörgö, D. A. Dawson, J. N. K. Rao, and A. K. Md. E. Saleh (North-Holland, Amsterdam, 1981), pp. 39–56.