Tight informationally complete quantum measurements

A J Scott

Institute for Quantum Information Science, University of Calgary, Calgary, AB T2N 1N4, Canada

E-mail: ascott@qis.ucalgary.ca

Abstract
We introduce a class of informationally complete positive-operator-valued measures which are, in analogy with a tight frame, 'as close as possible' to orthonormal bases for the space of quantum states. These measures are distinguished by an exceptionally simple state-reconstruction formula which allows 'painless' quantum state tomography. Complete sets of mutually unbiased bases and symmetric informationally complete positive-operator-valued measures are both members of this class, the latter being the unique minimal rank-one members. Recast as ensembles of pure quantum states, the rank-one members are in fact equivalent to weighted 2-designs in complex projective space. These measures are shown to be optimal for quantum cloning and linear quantum state tomography.

PACS numbers: 03.65.Wj, 03.67.−a, 02.10.Ud

1. Introduction

The retrieval of classical data from quantum systems, a task described by quantum measurement theory, is an overlooked—though important—component of quantum information processing [1]. The ability to precisely determine a quantum state is paramount to tests of quantum information processing devices such as quantum teleporters, key distributers, cloners, gates and, indeed, quantum computers. Quality assurance requires a complete characterization of the device, which is gained through knowledge of the output states for a judicious choice of input states.

The outcome statistics of a quantum measurement are described by a positive-operator-valued measure (POVM) [2–5]. An *informationally complete POVM* (*IC-POVM*) [6–13] is one with the property that every quantum state is uniquely determined by its measurement statistics. A sequence of measurements on copies of a system in an unknown state, enabling an estimate of the statistics, will then reveal the state. This process is called *quantum state*

tomography [14]. Besides this practical purpose, IC-POVMs with special properties are used for quantum cryptography [15], quantum fingerprinting [16] and are relevant to foundational studies of quantum mechanics [17–19].

This paper introduces a special class of IC-POVMs which are, in analogy with a tight frame [20–22], 'as close as possible' to orthonormal bases for the space of quantum states. These IC-POVMs will be called *tight IC-POVMs*. They allow 'painless' [23] quantum state tomography through a particularly simple state-reconstruction formula. The unique minimal rank-one members are the symmetric IC-POVMs (SIC-POVMs) [24]. Complete sets of mutually unbiased bases (MUBs) [25, 26] also form tight IC-POVMs and, in fact, recast as ensembles of pure quantum states, the tight rank-one IC-POVMs are equivalent to weighted 2-designs in complex projective space. These IC-POVMs are shown to be optimal for linear quantum state tomography and measurement-based quantum cloning.

The paper is organized as follows. In the following section we will introduce the notion of a *t*-design in complex projective space. Such combinatorial designs have recently aroused interest from the perspective of quantum information theory [24, 27–31]. In sections 3 and 4 we will revise the concepts of operator frames and informational completeness, respectively, and then in section 5, introduce the tight IC-POVMs. We will show in what sense the entire class of tight rank-one IC-POVMs can be considered optimal in sections 6 and 7, where, respectively, linear quantum state tomography and measurement-based cloning are investigated. Finally, in section 8 we summarize our results. Finite-dimensional Hilbert spaces are assumed throughout the paper.

2. Complex projective designs

The extension of spherical *t*-designs [32] to projective spaces was first considered by Neumaier [33], but for the most part studied by Hoggar [34–37] and Bannai and Hoggar [38, 39]. For a unified treatment of designs in terms of metric spaces consult the work of Levenshtein [40–42]. Our interest lies with the complex projective space $\mathbb{C}P^{d-1}$ of lines passing through the origin in $\mathbb{C}^d$. In this case each $x \in \mathbb{C}P^{d-1}$ may be represented by a unit vector $|x\rangle \in \mathbb{C}^d$ (modulo a phase), or more appropriately by the rank-one projector $\pi(x) \equiv |x\rangle\langle x|$. We will use both representations in this paper. Roughly speaking, a complex projective *t*-design is then a finite subset of $\mathbb{C}P^{d-1}$ with the property that the discrete average of a polynomial of degree *t* or less over the design equals the uniform average. Many equivalent definitions can be made in these terms (see e.g. [33, 34, 40, 43]). In the general context of compact metric spaces, for example, Levenshtein [41, 42] calls a finite set $\mathcal{D} \subset \mathbb{C}P^{d-1}$ a complex projective *t*-design if

$$\frac{1}{|\mathcal{D}|^2} \sum_{x,y \in \mathcal{D}} f(|\langle x|y\rangle|^2) = \int\int_{\mathbb{C}P^{d-1}} \mathrm{d}\mu_{\mathrm{H}}(x)\,\mathrm{d}\mu_{\mathrm{H}}(y)\, f(|\langle x|y\rangle|^2) \tag{1}$$

for any real polynomial *f* of degree *t* or less, where μ_{H} denotes the unique unitarily-invariant probability measure on $\mathbb{C}P^{d-1}$ induced by the Haar measure on U(*d*). In the current context we deem it appropriate to make a more explicit definition of a *t*-design which is specialized to complex projective spaces. With this in mind, let $\Pi_{\mathrm{sym}}^{(t)}$ denote the projector onto the totally symmetric subspace of $(\mathbb{C}^d)^{\otimes t}$ and consider the following simple fact.

Lemma 1.

$$\int_{\mathbb{C}P^{d-1}} \mathrm{d}\mu_{\mathrm{H}}(x)\pi(x)^{\otimes t} = \binom{d+t-1}{t}^{-1} \Pi_{\mathrm{sym}}^{(t)}. \tag{2}$$

Proof. Use Schur's lemma. The LHS of equation (2) is invariant under all unitaries $U^{\otimes t}$ which act irreducibly on the totally symmetric subspace of $(\mathbb{C}^d)^{\otimes t}$. □

By considering the monomial $|\langle x|y\rangle|^{2t} = \mathrm{tr}[\pi(x)^{\otimes t}\pi(y)^{\otimes t}]$ in equation (1), it can be easily shown that lemma 1 and theorem 6 (below) allow the following equivalent definition of a complex projective t-design.

Definition 2. *A finite set $\mathscr{D} \subset \mathbb{C}P^{d-1}$ is called a t-design (of dimension d) if*

$$\frac{1}{|\mathscr{D}|}\sum_{x\in\mathscr{D}}\pi(x)^{\otimes t} = \binom{d+t-1}{t}^{-1}\Pi_{\mathrm{sym}}^{(t)}. \tag{3}$$

Seymour and Zaslavsky have shown that t-designs in $\mathbb{C}P^{d-1}$ exist for any t and d [44]. It is necessary, however, that the number of design points satisfy [34, 38, 40, 45]

$$|\mathscr{D}| \geqslant \binom{d+\lceil t/2\rceil -1}{\lceil t/2\rceil}\binom{d+\lfloor t/2\rfloor -1}{\lfloor t/2\rfloor}. \tag{4}$$

A design which achieves this bound is called *tight*. Tight t-designs in $\mathbb{C}P^1$ are equivalent to tight spherical t-designs on the Euclidean 2-sphere (via equation (46) in section 5). Such designs exist only for $t = 1, 2, 3, 5$ (see e.g. [46]). When $d \geqslant 3$ it is known that tight t-designs in $\mathbb{C}P^{d-1}$ exist only for $t = 1, 2, 3$ [36, 38, 39]. It is trivial that tight 1-designs exist in all dimensions. Tight 2-designs have been conjectured to also exist for all d [24, 27]. Analytical constructions, however, are known only for $d \leqslant 10$ and $d = 12, 13, 19$ [24, 27, 47–50]. Examples of tight 3-designs are known only for $d = 2, 4, 6$ [34]. When $d \geqslant 3$ and $t \geqslant 5$ the above bound can be improved by more than 1 [51–53].

The concept of t-designs has been generalized to that of *weighted t-designs* [41, 42]. Each design point $x \in \mathscr{D}$ is then appointed a positive weight $w(x)$ under the normalization constraint $\sum_{x\in\mathscr{D}} w(x) = 1$. A countable set $\mathscr{S}$ endowed with a normalized weight function $w : \mathscr{S} \to [0, 1]$ will be called a *weighted set* and denoted by the pair $(\mathscr{S}, w)$.

Definition 3. *A finite weighted set $(\mathscr{D}, w)$, $\mathscr{D} \subset \mathbb{C}P^{d-1}$, is called a weighted t-design (of dimension d) if*

$$\sum_{x\in\mathscr{D}} w(x)\pi(x)^{\otimes t} = \binom{d+t-1}{t}^{-1}\Pi_{\mathrm{sym}}^{(t)}. \tag{5}$$

The weighted t-designs obviously incorporate the 'unweighted' t-designs as the special case $w \equiv 1/|\mathscr{D}|$. Note that the normalization of w is implied by the trace of equation (5). If we instead 'trace out' only one subsystem of these t-partite operators, we can immediately deduce that every weighted t-design is also a weighted $(t-1)$-design. A weighted 1-design is known as a *tight (vector) frame* in the context of frame theory [20–22], in which case the unnormalized states $|\tilde{x}\rangle \equiv \sqrt{w(x)d}|x\rangle$ are the frame vectors, and equation (5) is the tight frame condition: $\sum_{x\in\mathscr{D}} |\tilde{x}\rangle\langle\tilde{x}| = I$. In this form it is immediately apparent that we must have $|\mathscr{D}| \geqslant d$ for a weighted 1-design, with equality only if the frame vectors $|\tilde{x}\rangle$ form an orthonormal basis for $\mathbb{C}^d$. The 2-design case is treated in the following theorem.

Theorem 4. *Let $(\mathscr{D}, w)$ be a weighted 2-design of dimension d. Then $|\mathscr{D}| \geqslant d^2$ with equality only if $w \equiv 1/|\mathscr{D}|$ and $|\langle x|y\rangle|^2 = 1/(d+1)$ for all $x, y \in \mathscr{D}$ with $x \neq y$.*

Proof. By the definition of a weighted 2-design,

$$\sum_{x\in\mathscr{D}} w(x)\pi(x)\otimes\pi(x) = \frac{2}{d(d+1)}\Pi_{\mathrm{sym}}^{(2)}$$

$$= \frac{1}{d(d+1)}\sum_{j,k}|e_j\rangle\langle e_j|\otimes|e_k\rangle\langle e_k| + |e_j\rangle\langle e_k|\otimes|e_k\rangle\langle e_j|, \tag{6}$$

where $\{|e_k\rangle\}_{k=1}^d$ is an orthonormal basis for $\mathbb{C}^d$. Now if we multiply both sides of this equation by $A \otimes I$, where A is an arbitrary linear operator, and then trace out the first subsystem, we find that

$$\sum_{x \in \mathscr{D}} w(x) \operatorname{tr}[\pi(x)A]\pi(x) = \frac{1}{d(d+1)} \sum_{j,k} \langle e_j|A|e_j\rangle|e_k\rangle\langle e_k| + |e_k\rangle\langle e_k|A|e_j\rangle\langle e_j| \tag{7}$$

$$= \frac{1}{d(d+1)}(\operatorname{tr}(A)I + A) \tag{8}$$

and thus any $A \in \operatorname{End}(\mathbb{C}^d)$ can be rewritten as a linear combination of the design projectors:

$$A = d \sum_{x \in \mathscr{D}} w(x)((d+1)\operatorname{tr}[\pi(x)A] - \operatorname{tr}(A))\pi(x), \tag{9}$$

where we have used the fact that a 2-design is also a 1-design, i.e. $I = d \sum_{x \in \mathscr{D}} w(x)\pi(x)$. Consequently, the design projectors $\pi(x)$ span $\operatorname{End}(\mathbb{C}^d) \cong \mathbb{C}^{d^2}$, and thus, there must be at least d^2 many. Furthermore, when $|\mathscr{D}| = d^2$ these operators must be linearly independent. Assuming this to be the case and choosing $A = \pi(y)$ in equation (9), for some fixed $y \in \mathscr{D}$, we find that

$$(w(y)d^2 - 1)\pi(y) + d \sum_{x \neq y} w(x)((d+1)\operatorname{tr}[\pi(x)\pi(y)] - 1)\pi(x) = 0, \tag{10}$$

which, given the linear independence of the design projectors, can be satisfied only if $w(y) = 1/d^2 = 1/|\mathscr{D}|$ and $\operatorname{tr}[\pi(x)\pi(y)] = |\langle x|y\rangle|^2 = 1/(d+1)$ for all $x \neq y$. The same is true for all $y \in \mathscr{D}$. $\qquad\square$

Theorem 4 is essentially a special case of the results of Levenshtein [41, 42]. In fact, the above lower bound (equation (4)) also holds for weighted t-designs, with equality occurring only if the design has uniform weight, i.e. $w \equiv 1/|\mathscr{D}|$. The current proof, however, takes a form which incorporates the theme of this paper. Like in the specific 2-design case, more can be said about the structure of t-designs when equation (4) is satisfied with equality. Our interest lies only with the 2-designs, however, and thus we defer further results in this direction to the work of Bannai and Hoggar [34–39].

We have introduced complex projective t-designs as a special type of weighted subset of $\mathbb{C}P^{d-1}$. Note that the weight function of an arbitrary weighted set $(\mathscr{S}, w)$ may be trivially extended to a countably additive measure on the power set $2^\mathscr{S}$. We will use this observation to generalize the concept of t-designs one step further. Let $\mathfrak{B}(\mathscr{S})$ denote the Borel σ-algebra of $\mathscr{S}$. In the following situation, a set $\mathscr{S}$ endowed with a probability measure $\omega : \mathfrak{B}(\mathscr{S}) \to [0, 1]$, i.e. a (Borel) probability space, will be called a *distribution* and denoted by the pair $(\mathscr{S}, \omega)$.

Definition 5. *A distribution $(\mathscr{D}, \omega)$, $\mathscr{D} \subseteq \mathbb{C}P^{d-1}$, is called a generalized t-design (of dimension d) if*

$$\int_\mathscr{D} d\omega(x)\pi(x)^{\otimes t} = \binom{d+t-1}{t}^{-1} \Pi_{\text{sym}}^{(t)}. \tag{11}$$

In this definition the Lebesgue–Stieltjes integral is used, which reduces to a discrete sum when $\mathscr{D}$ is countable. A generalized t-design is thus a weighted t-design when $\mathscr{D}$ is a finite set. Again, every generalized t-design is also a generalized $(t-1)$-design and, by lemma 1, $(\mathbb{C}P^{d-1}, \mu_H)$ is a generalized t-design for all t.

By allowing any distribution of points in $\mathbb{C}P^{d-1}$ which satisfies equation (11) to be called a 'generalized' t-design, we have in fact contradicted an important purpose of designs, which is to convert integrals into finite sums. In this paper, however, we will allow this discrepancy and henceforth refer to both weighted and 'generalized' complex projective t-designs as simply t-designs. The task of finding t-designs is facilitated by the following theorem (see e.g. [41, 43]).

Theorem 6. *Let $(\mathscr{S}, \omega)$, $\mathscr{S} \subseteq \mathbb{C}P^{d-1}$, be a distribution. Then for any $t \geqslant 1$,*

$$\int\!\!\int_{\mathscr{S}} d\omega(x)\, d\omega(y) |\langle x|y \rangle|^{2t} \geqslant \binom{d+t-1}{t}^{-1}, \tag{12}$$

with equality iff $(\mathscr{S}, \omega)$ is a t-design.

Proof. Consider an arbitrary distribution $(\mathscr{S}, \omega)$ and define

$$S \equiv \int_{\mathscr{S}} d\omega(x) \pi(x)^{\otimes t}, \tag{13}$$

which has support only on the totally symmetric subspace of $(\mathbb{C}^d)^{\otimes t}$. This positive operator can thus have at most $d_{\mathrm{sym}} = \binom{d+t-1}{t}$ nonzero eigenvalues $\lambda_1, \ldots, \lambda_{d_{\mathrm{sym}}}$, which satisfy the equations

$$\mathrm{tr}(S) = \int_{\mathscr{S}} d\omega(x) = 1 = \sum_{k=1}^{d_{\mathrm{sym}}} \lambda_k, \qquad \text{and}$$

$$\mathrm{tr}(S^2) = \int\!\!\int_{\mathscr{S}} d\omega(x)\, d\omega(y) |\langle x|y \rangle|^{2t} = \sum_{k=1}^{d_{\mathrm{sym}}} \lambda_k^2. \tag{14}$$

The lower bound (equation (12)) is apparent from the RHS of these equations. Under the normalization constraint expressed by the first, the second is bounded below: $\mathrm{tr}(S^2) \geqslant 1/d_{\mathrm{sym}}$. Equality can occur if and only if $\lambda_k = 1/d_{\mathrm{sym}}$ for all k, or equivalently $S = \Pi_{\mathrm{sym}}^{(t)}/d_{\mathrm{sym}}$, which is the defining property of a t-design. $\square$

This theorem allows us to check whether a distribution of points in $\mathbb{C}P^{d-1}$ forms a t-design by considering only the angles between the supposed design elements. It also shows that t-designs can be found numerically by parametrizing a distribution and minimizing the LHS of equation (12). The lower bound is in fact a straightforward generalization of the Welch bound [54].

3. Operator frames

Frame theory [20–22] provides a natural setting for the study of informationally complete quantum measurements [10]. In this section we will introduce some of the important concepts of this theory that are relevant to the current investigation. Before beginning, however, we will need to introduce the concept of a superoperator.

Following Caves [55] we will write a linear operator A in vector notation as $|A)$. The vector space of all such operators, $\mathrm{End}(\mathbb{C}^d) \cong \mathbb{C}^{d^2}$, equipped with the Hilbert–Schmidt inner product $(A|B) \equiv \mathrm{tr}(A^\dagger B)$, is a Hilbert space, where we think of $(A|$ as an operator 'bra' and $|B)$ as an operator 'ket'. Addition and scalar multiplication of operator kets then follows that for operators, e.g., $a|A) + b|B) = |aA + bB)$. The usefulness of this notation becomes apparent when we consider linear maps on operators, i.e. superoperators. Given an orthonormal operator

basis $\{E_k\}_{k=1}^{d^2} \subset \mathrm{End}(\mathbb{C}^d)$, $(E_j|E_k) = \delta(j, k)$, a superoperator $\mathcal{S} \in \mathrm{End}(\mathrm{End}(\mathbb{C}^d)) \cong \mathbb{C}^{d^4}$ may be written in two different ways:

$$\mathcal{S} = \sum_{j,k} s_{jk} E_j \odot E_k^\dagger = \sum_{j,k} s_{jk}|E_j)(E_k| \qquad (s_{jk} \in \mathbb{C}). \tag{15}$$

The first representation illustrates the *ordinary* action of the superoperator,

$$\mathcal{S}(A) \equiv \sum_{j,k} s_{jk} E_j A E_k^\dagger, \tag{16}$$

which amounts to inserting A into the location of the '$\odot$' symbol. The second reflects the *left–right* action

$$\mathcal{S}|A) \equiv \sum_{j,k} s_{jk}|E_j)(E_k|A) = \sum_{j,k} s_{jk} E_j \, \mathrm{tr}\left(E_k^\dagger A\right), \tag{17}$$

where the superoperator acts on operators just like an operator on vectors. It is this second 'non-standard' action which will be useful in the current context. The identity superoperators relative to the ordinary and left–right actions are, respectively, $\mathcal{I} \equiv I \odot I$ and $\mathbf{I} \equiv \sum_k |E_k)(E_k|$. Further results on superoperators in the current notation can be found in [56, 57].

The notion of an informationally complete POVM is naturally related to that of a frame, or more specifically, an 'operator' frame. Frames generalize the notion of bases. We call a countable family of operators $\{A(x)\}_{x \in \mathscr{X}} \subset \mathrm{End}(\mathbb{C}^d)$ an *operator frame* if there exist constants $0 < a \leqslant b < \infty$ such that

$$a(C|C) \leqslant \sum_{x \in \mathscr{X}} |(A(x)|C)|^2 \leqslant b(C|C) \tag{18}$$

for all $C \in \mathrm{End}(\mathbb{C}^d)$. For example, all finite linearly spanning subsets of $\mathrm{End}(\mathbb{C}^d)$ are operator frames. When $a = b$ the frame is called *tight* [23]. Tight frames are those frames which are most like orthonormal bases (see e.g. [58]). An operator frame with cardinality $|\mathscr{X}| = d^2$, i.e. an operator basis, is tight if and only if it is an orthonormal basis. For every frame $\{A(x)\}_{x \in \mathscr{X}}$ there is a *dual frame* $\{B(x)\}_{x \in \mathscr{X}}$, such that

$$\sum_{x \in \mathscr{X}} |B(x))(A(x)| = \mathbf{I}, \tag{19}$$

and hence

$$C = \sum_{x \in \mathscr{X}} (A(x)|C)B(x) = \sum_{x \in \mathscr{X}} (B(x)|C)A(x) \tag{20}$$

for all $C \in \mathrm{End}(\mathbb{C}^d)$. Although when $|\mathscr{X}| > d^2$ there are different choices for the dual frame [59], the most 'economical' choice (see proposition 3.2.4 of [21]) is the *canonical dual frame* $\{\tilde{A}(x)\}_{x \in \mathscr{X}}$,

$$|\tilde{A}(x)) \equiv \mathcal{A}^{-1}|A(x)), \tag{21}$$

where the *frame superoperator*

$$\mathcal{A} \equiv \sum_{x \in \mathscr{X}} |A(x))(A(x)|, \tag{22}$$

so that

$$\sum_{x \in \mathscr{X}} |\tilde{A}(x))(A(x)| = \sum_{x \in \mathscr{X}} \mathcal{A}^{-1}|A(x))(A(x)| = \mathcal{A}^{-1}\mathcal{A} = \mathbf{I} \tag{23}$$

as required. Note that the inverse of $\mathcal{A}$ is taken with respect to left–right action and exists whenever $\{A(x)\}_{x \in \mathscr{X}}$ is an operator frame. A tight operator frame is one with $\mathcal{A} = a\mathbf{I}$ and

thus trivially $|\tilde{A}(x)) = |A(x))/a$. In general, however, inverting the frame superoperator will be a difficult analytical task.

In this paper we prefer the concept of generalized (or 'continuous') frames [20, 60, 61] over the preceding more common notion. Suppose now that the set $\mathscr{X}$ (which need no longer be countable) is endowed with a positive measure $\alpha : \mathfrak{B}(\mathscr{X}) \to [0, \infty]$. We call a family of operators $\{A(x)\}_{x \in \mathscr{X}} \subseteq \text{End}(\mathbb{C}^d)$ a *generalized operator frame* (*with respect to* α) if there exist constants $0 < a \leqslant b < \infty$ such that

$$a(C|C) \leqslant \int_{\mathscr{X}} d\alpha(x)|(A(x)|C)|^2 \leqslant b(C|C) \tag{24}$$

for all $C \in \text{End}(\mathbb{C}^d)$. This definition reduces to the above discrete case when $\mathscr{X}$ is countable and α is the counting measure. Again, for every frame $\{A(x)\}_{x \in \mathscr{X}}$ there is a dual frame $\{B(x)\}_{x \in \mathscr{X}}$ such that

$$\int_{\mathscr{X}} d\alpha(x)|B(x))(A(x)| = \mathbf{I}, \tag{25}$$

and the canonical dual frame $\{\tilde{A}(x)\}_{x \in \mathscr{X}}$ is defined through equation (21), where now the frame superoperator

$$\mathcal{A} \equiv \int_{\mathscr{X}} d\alpha(x)|A(x))(A(x)|. \tag{26}$$

A generalized tight operator frame is also defined in analogy to the discrete case.

Definition 7. *An operator frame* $\{A(x)\}_{x \in \mathscr{X}} \subseteq \text{End}(\mathbb{C}^d)$ *with respect to the measure* α *is called tight if*

$$\int_{\mathscr{X}} d\alpha(x)|A(x))(A(x)| = a\mathbf{I}, \tag{27}$$

for some constant $a > 0$, *i.e.* $\mathcal{A} = a\mathbf{I}$.

The argument that tight frames are 'as close as possible' to orthonormal bases comes from this resolution of unity (equation (27)) and the following inequality (see e.g. [58]), which is called the frame bound. Let 'Tr' denote the superoperator trace.

Theorem 8. *Let* $\{A(x)\}_{x \in \mathscr{X}} \subseteq \text{End}(\mathbb{C}^d)$ *be an operator frame with respect to the measure* α. *Then*

$$\int\int_{\mathscr{X}} d\alpha(x) \, d\alpha(y)|(A(x)|A(y))|^2 \geqslant \frac{(\text{Tr}(\mathcal{A}))^2}{d^2}, \tag{28}$$

with equality if and only if $\{A(x)\}_{x \in \mathscr{X}}$ *is a tight operator frame.*

Proof. Let $\lambda_1, \ldots, \lambda_{d^2} > 0$ denote the left–right eigenvalues of $\mathcal{A}$. The LHS of equation (28) can be rewritten as

$$\text{Tr}(\mathcal{A}^2) = \sum_{k=1}^{d^2} \lambda_k{}^2, \tag{29}$$

which under the constraint $\sum_{k=1}^{d^2} \lambda_k = \text{Tr}(\mathcal{A})$ takes its minimum value if and only if $\lambda_1 = \cdots = \lambda_{d^2} = \text{Tr}(\mathcal{A})/d^2$, i.e. $\mathcal{A} = (\text{Tr}(\mathcal{A})/d^2)\mathbf{I}$, which means $\{A(x)\}_{x \in \mathscr{X}}$ is a tight operator frame with $a = \text{Tr}(\mathcal{A})/d^2$. The minimum is $a^2 d^2$. $\qquad\square$

The frame bound can be considered a variant of the Welch bound (theorem 6) with $t = 1$. Theorem 8 shows that tight frames are those which minimize the average correlation amongst

the frame elements. It is straightforward to show that tight operator frames exist for all d and $|\mathscr{X}| \geqslant d^2$. Given any operator frame $\{A(x)\}_{x \in \mathscr{X}}$ we can construct a tight frame through the use of the frame superoperator, e.g. $\{\mathcal{A}^{-1/2}|A(x))\}_{x \in \mathscr{X}}$. In fact, explicit examples of tight *unitary* operator frames are known for all d and $|\mathscr{X}| \geqslant d^2$ [62]. Unitary 2-designs have also recently been considered [31] (unitary 1-designs are equivalent to tight unitary operator frames). These examples are important for the implementation of certain types of quantum processes, which are defined by rewriting equation (27) (and its t-design generalization) in terms of the ordinary action of a superoperator. For example, tight unitary operator frames implement the depolarizing channel. To relate the concept of tight frames to informationally complete POVMs we need to instead consider frames on a subspace of End($\mathbb{C}^d$). This will be done in section 5.

4. Informationally complete quantum measurements

The outcome statistics of a quantum measurement are described by a positive-operator-valued measure (POVM) [2–5]. That is, an operator-valued function defined on a σ-algebra over the set $\mathscr{X}$ of outcomes, $F : \mathfrak{B}(\mathscr{X}) \to \text{End}(\mathbb{C}^d)$, which satisfies (1) $F(\mathscr{S}) \geqslant 0$ for all $\mathscr{S} \in \mathfrak{B}(\mathscr{X})$ with equality if $\mathscr{S} = \emptyset$, (2) $F\left(\bigcup_{k=1}^{\infty} \mathscr{S}_k\right) = \sum_{k=1}^{\infty} F(\mathscr{S}_k)$ for any sequence of disjoint sets $\mathscr{S}_k \in \mathfrak{B}(\mathscr{X})$ and (3) the normalization constraint $F(\mathscr{X}) = I$. In this paper we always take $\mathfrak{B}(\mathscr{X})$ to be the Borel σ-algebra.

An informationally complete quantum measurement [6] is one with the property that each quantum state $\rho \in Q(\mathbb{C}^d) \equiv \left\{A \in \text{End}(\mathbb{C}^d) | A \geqslant 0, \text{tr}(A) = 1\right\}$ is uniquely determined by its measurement statistics $p(\mathscr{S}) \equiv \text{tr}[F(\mathscr{S})\rho]$. Consequently, given multiple copies of a system in an unknown state, a sequence of measurements will give an estimate of the statistics and, hence, identify the state. The measure F is then called an *informationally complete POVM* (*IC-POVM*).

Definition 9. *A POVM $F : \mathfrak{B}(\mathscr{X}) \to \text{End}(\mathbb{C}^d)$ is called informationally complete if for each pair of distinct quantum states $\rho \neq \sigma \in Q(\mathbb{C}^d)$ there exists an event $\mathscr{S} \in \mathfrak{B}(\mathscr{X})$ such that $\text{tr}[F(\mathscr{S})\rho] \neq \text{tr}[F(\mathscr{S})\sigma]$.*

When a quantum measurement has a countable number of outcomes, the indexed set of POVM elements $\{F(x)\}_{x \in \mathscr{X}}$ completely characterizes F and is thus often referred to as the 'POVM'. We will call such measurements *discrete*, or *finite* if we additionally have $|\mathscr{X}| < \infty$. A discrete POVM is informationally complete if and only if for each pair of distinct quantum states $\rho \neq \sigma \in Q(\mathbb{C}^d)$ there exists an outcome $x \in \mathscr{X}$ such that $\text{tr}[F(x)\rho] \neq \text{tr}[F(x)\sigma]$.

To show how a quantum state can be reconstructed from its measurement statistics, we will first need to express F in a standard form. Consider an arbitrary quantum measurement. The POVM defines a natural real-valued *trace measure* [63], $\tau(\mathscr{S}) \equiv \text{tr}[F(\mathscr{S})]$, which inherits the normalization $\tau(\mathscr{X}) = d$. Since each matrix element of F is a complex valued measure which is absolutely continuous with respect to the non-negative finite measure τ, the POVM can be expressed as

$$F(\mathscr{S}) = \int_{\mathscr{S}} \text{d}\tau(x) F'_\tau(x) \equiv \int_{\mathscr{S}} \text{d}\tau(x) P(x), \tag{30}$$

where the Radon–Nikodym derivative $F'_\tau : \mathscr{X} \to \text{End}(\mathbb{C}^d)$ is a positive-operator-valued density (POVD) which is uniquely defined up to a set of zero τ-measure. We will set $F'_\tau \equiv P$. Note that our choice of scalar measure implies that $\text{tr}(P) = 1$, τ-almost everywhere. When P also has unit rank we call F a *rank-one POVM*, in which case it is natural to have $\mathscr{X} \subseteq \mathbb{C}P^{d-1}$ and then $P \equiv \pi$. A rank-one POVM is of course equivalent to a 1-design, i.e. a tight vector

frame [64]. In the special case of a discrete quantum measurement the Radon–Nikodym derivative is simply $P(x) \equiv F_\tau'(x) = F(x)/\operatorname{tr}[F(x)]$.

For an arbitrary POVM F, define the superoperator

$$\mathcal{F} \equiv \int_{\mathscr{X}} \mathrm{d}\tau(x)|P(x))(P(x)|, \tag{31}$$

which is positive and bounded under the left–right action:

$$0 \leqslant (A|\mathcal{F}|A) = \int_{\mathscr{X}} \mathrm{d}\tau(x)\big|(A|P(x))\big|^2 \leqslant \int_{\mathscr{X}} \mathrm{d}\tau(x)(P(x)|P(x))(A|A)$$

$$\leqslant \int_{\mathscr{X}} \mathrm{d}\tau(x)(A|A) = d(A|A) \tag{32}$$

for all $A \in \operatorname{End}(\mathbb{C}^d)$, where we have used the Cauchy–Schwarz inequality and then the fact that $\operatorname{tr}(P^2) \leqslant 1$. Now consider the following straightforward result.

Proposition 10. *Let $F : \mathfrak{B}(\mathscr{X}) \to \operatorname{End}(\mathbb{C}^d)$ be a POVM. Then F is informationally complete iff there exists a constant $a > 0$ such that $(A|\mathcal{F}|A) \geqslant a(A|A)$ for all $A \in \operatorname{End}(\mathbb{C}^d)$.*

Proof. Suppose F is informationally complete. If there existed an operator $A \neq 0$ such that

$$(A|\mathcal{F}|A) = \int_{\mathscr{X}} \mathrm{d}\tau(x)|\operatorname{tr}[P(x)A]|^2 = 0, \tag{33}$$

then we must have $\operatorname{tr}(PA) = 0$, τ-almost everywhere. This operator must therefore be traceless:

$$\operatorname{tr}(A) = \operatorname{tr}[F(\mathscr{X})A] = \int_{\mathscr{X}} \mathrm{d}\tau(x)\operatorname{tr}[P(x)A] = 0. \tag{34}$$

Now for any state $\rho \in \mathrm{Q}(\mathbb{C}^d)$ we can define the state $\sigma = \rho + \epsilon(A + A^\dagger)$, where $\epsilon > 0$ is chosen small enough such that $\sigma \geqslant 0$. Then

$$\operatorname{tr}[F(\mathscr{S})\sigma] = \operatorname{tr}[F(\mathscr{S})\rho] + \epsilon \int_{\mathscr{S}} \mathrm{d}\tau(x)(\operatorname{tr}[P(x)A] + \operatorname{tr}[P(x)A]^*) = \operatorname{tr}[F(\mathscr{S})\rho] \tag{35}$$

for all $\mathscr{S} \in \mathfrak{B}(\mathscr{X})$, with $\sigma \neq \rho$. This means F could not have been informationally complete. Thus for IC-POVMs, $\mathcal{F}$ will always be strictly positive relative to the left–right action. The converse is also true. If for the distinct quantum states $\rho \neq \sigma \in \mathrm{Q}(\mathbb{C}^d)$ we have

$$(\rho - \sigma|\mathcal{F}|\rho - \sigma) = \int_{\mathscr{X}} \mathrm{d}\tau(x)|\operatorname{tr}[P(x)(\rho - \sigma)]|^2 > 0, \tag{36}$$

then there must exist an event $\mathscr{S} \in \mathfrak{B}(\mathscr{X})$, such that

$$\int_{\mathscr{S}} \mathrm{d}\tau(x)\operatorname{tr}[P(x)(\rho - \sigma)] \neq 0, \tag{37}$$

or equivalently, $\operatorname{tr}[F(\mathscr{S})\rho] \neq \operatorname{tr}[F(\mathscr{S})\sigma]$, which means F is informationally complete. □

Note that the proof of proposition 10 made no reference to our particular choice of scalar measure. We could also express the POVM in terms of another. However the trace measure guarantees the boundedness of the superoperator $\mathcal{F}$ and was found to be the best choice for a canonical scalar measure in the current context.

When a POVM F is informationally complete, in which case we have just shown that the corresponding superoperator $\mathcal{F}$ has full rank relative to the left-right action, the POVD P can be considered a generalized operator frame with respect to τ. The canonical dual frame then defines a *reconstruction operator-valued density*

$$|R) \equiv \mathcal{F}^{-1}|P), \tag{38}$$

where the inverse of $\mathcal{F}$, which we now call the *POVM superoperator*, is taken with respect to the left–right action. The identity

$$\int_{\mathscr{X}} d\tau(x)|R(x))(P(x)| = \int_{\mathscr{X}} d\tau(x)\mathcal{F}^{-1}|P(x))(P(x)| = \mathcal{F}^{-1}\mathcal{F} = \mathbf{I}, \quad (39)$$

then allows state reconstruction in terms of the measurement statistics:

$$\rho = \int_{\mathscr{X}} d\tau(x)\,\mathrm{tr}[P(x)\rho]R(x) = \int_{\mathscr{X}} \mathrm{tr}[dF(x)\rho]R(x) = \int_{\mathscr{X}} dp(x)R(x), \quad (40)$$

where $p(\mathscr{S}) \equiv \mathrm{tr}[F(\mathscr{S})\rho] = \int_{\mathscr{S}} d\tau(x)\,\mathrm{tr}[P(x)\rho]$. This *state-reconstruction formula* is an immediate consequence of the left–right action of equation (39) on $|\rho)$.

We will now give some useful properties of the reconstruction operator-valued density (OVD) which will be needed later in the paper. Although R is generally not positive, it inherits all other properties of P. For example, we know that R is Hermitian since $\mathcal{F}$, and thus $\mathcal{F}^{-1}$, maps Hermitian operators to Hermitian operators. Additionally, the left–right action of equation (39) on $|I)$ shows that

$$\int_{\mathscr{X}} d\tau(x)R(x) = I. \quad (41)$$

Note that for an arbitrary POVM, the identity operator is always a left–right eigenvector of the POVM superoperator:

$$\mathcal{F}|I) = \int_{\mathscr{X}} d\tau(x)|P(x))(P(x)|I) = \int_{\mathscr{X}} d\tau(x)|P(x)) = \int_{\mathscr{X}} |dF(x)) = |I), \quad (42)$$

using $\mathrm{tr}(P) = 1$ and the normalization of the POVM. Thus $|I)$ is also an eigenvector of $\mathcal{F}^{-1}$, and we obtain

$$\mathrm{tr}(R) = (I|R) = (I|\mathcal{F}^{-1}|P) = (I|P) = \mathrm{tr}(P) = 1. \quad (43)$$

Finally, it is straightforward to confirm that

$$\mathcal{F}^{-1} = \int_{\mathscr{X}} d\tau(x)|R(x))(R(x)|. \quad (44)$$

Note that we need $|\mathscr{X}| \geqslant d^2$ for F to be informationally complete. If this were not the case then $\mathcal{F}$ could not have full rank. An IC-POVM with $|\mathscr{X}| = d^2$ is called *minimal*. In this case the reconstruction OVD is unique. In general, however, there will be many different choices. When F is a discrete IC-POVM the trace measure can be replaced by the counting measure [10] in equation (30). Then $P'(x) = F(x)$, the POVM superoperator is $\mathcal{F}' = \sum_{x \in \mathscr{X}} |F(x))(F(x)|$ and $|R'(x)) = \mathcal{F}'^{-1}|F(x))$ say. In this case we also have $\rho = \sum_{x \in \mathscr{X}} p(x)R'(x)$. If it were not already obvious, it is now clear from the superoperator $\mathcal{F}'$ that F is informationally complete if and only if $\{F(x)\}_{x \in \mathscr{X}}$ spans $\mathrm{End}(\mathbb{C}^d)$. Although the counting measure might seem more convenient, in section 6 we will show that the canonical dual frame with respect to the trace measure is the optimal choice for quantum state tomography.

5. Tight IC-POVMs

The notion of an informationally complete POVM is naturally related to that of an operator frame. In the previous section we showed how to reconstruct a quantum state from its measurement statistics for an arbitrary IC-POVM. The procedure required inverting a superoperator, however, which may not be a straightforward analytical task. In this section we will investigate a class of IC-POVMs which share a particularly simple state-reconstruction formula. In analogy with a tight frame, these IC-POVMS will be called *tight IC-POVMs*.

Although pure states correspond to rays in a complex vector space, the most natural setting in which to study a general quantum state is Euclidean space. The set of all quantum states $Q(\mathbb{C}^d)$ is embedded in $\mathbb{R}^{d^2-1}$ as follows. Note that each $\rho \in Q(\mathbb{C}^d)$ may be associated with a traceless Hermitian operator under the mapping $\rho \rightarrow \rho - I/d$. Equipped with the Hilbert–Schmidt inner product $(A|B) \equiv \mathrm{tr}(A^\dagger B)$, which induces the Frobenius norm $\|A\| \equiv \sqrt{(A|A)}$, the set of all traceless Hermitian operators, $H_0(\mathbb{C}^d) \equiv \{A \in \mathrm{End}(\mathbb{C}^d) | A^\dagger = A, \mathrm{tr}(A) = 0\} \cong \mathbb{R}^{d^2-1}$, forms a real inner product space in which the images of pure states lie on a sphere, $\|\pi(\psi) - I/d\| = \sqrt{(d-1)/d}$, and the images of mixed states within. In the special case $d = 2$, this isometric embedding maps quantum states surjectively onto a ball in $H_0(\mathbb{C}^2) \cong \mathbb{R}^3$, realizing the Bloch-sphere representation of a qubit, but is otherwise only injective.

Let us now reconsider the POVM superoperator of an arbitrary POVM (equation (31)) in this setting. It is straightforward to confirm that we have the decomposition

$$\mathcal{F} = \frac{\mathcal{I}}{d} + \int_{\mathscr{X}} \mathrm{d}\tau(x) |P(x) - I/d)(P(x) - I/d|. \tag{45}$$

The superoperator $\mathcal{I}/d = |I)(I|/d$ is in fact an eigenprojector (equation (42)). It left–right projects onto the subspace spanned by the identity, whose orthogonal complement, the $(d^2 - 1)$-dimensional subspace of traceless operators, is $\mathcal{F}$-invariant. Define $\boldsymbol{\Pi}_0 \equiv \mathbf{I} - \mathcal{I}/d$, which left–right projects onto this latter subspace. The action of $\boldsymbol{\Pi}_0$ on a quantum state then realizes the above embedding into $H_0(\mathbb{C}^d)$:

$$\boldsymbol{\Pi}_0|\rho) = |\rho - I/d). \tag{46}$$

Let $\mathbf{I}_{H_0}$ denote the identity superoperator for $H_0(\mathbb{C}^d)$ under the left–right action. Noting that $P - I/d$ is a traceless Hermitian OVD, i.e. $P(x) - I/d \in H_0(\mathbb{C}^d)$ for all $x \in \mathscr{X}$, we are now ready to define a tight IC-POVM.

Definition 11. *Let $F : \mathfrak{B}(\mathscr{X}) \rightarrow \mathrm{End}(\mathbb{C}^d)$ be a POVM. Then F is called a tight IC-POVM if the OVD $P - I/d$ forms a tight operator frame (with respect to τ) in $H_0(\mathbb{C}^d)$, i.e.*

$$\int_{\mathscr{X}} \mathrm{d}\tau(x) |P(x) - I/d)(P(x) - I/d| = a\mathbf{I}_{H_0}, \tag{47}$$

or equivalently, $\boldsymbol{\Pi}_0 \mathcal{F} \boldsymbol{\Pi}_0 = a\boldsymbol{\Pi}_0$, for some constant $a > 0$.

Tight IC-POVMs are precisely those POVMs whose images under $\boldsymbol{\Pi}_0$ form tight operator frames in $H_0(\mathbb{C}^d)$. It is in this sense that they are claimed 'as close as possible' to orthonormal bases for the space of quantum states. The constant a can be found by taking the superoperator trace of equation (47):

$$a = a(F) = \frac{1}{d^2 - 1} \int_{\mathscr{X}} \mathrm{d}\tau(x)(P(x) - I/d | P(x) - I/d) \tag{48}$$

$$= \frac{1}{d^2 - 1}\left(\int_{\mathscr{X}} \mathrm{d}\tau(x)(P(x)|P(x)) - 1\right). \tag{49}$$

The POVM superoperator of a tight IC-POVM satisfies the identity

$$\mathcal{F} = \frac{1}{d}\mathcal{I} + a\boldsymbol{\Pi}_0 = a\mathbf{I} + \frac{1-a}{d}\mathcal{I}. \tag{50}$$

Since $a > 0$ by definition, this superoperator obviously has full rank. Its inverse is

$$\mathcal{F}^{-1} = \frac{1}{a}\mathbf{I} - \frac{1-a}{ad}\mathcal{I}, \tag{51}$$

and thus the reconstruction OVD (equation (38)) takes the form

$$R = \frac{1}{a}P - \frac{1-a}{ad}I, \tag{52}$$

where we have used the fact that $\mathrm{tr}(P) = 1$. A tight IC-POVM then has a particularly simple state-reconstruction formula (equation (40)):

$$\rho = \frac{1}{a}\int_{\mathscr{X}} \mathrm{d}p(x)P(x) - \frac{1-a}{ad}I. \tag{53}$$

This formula may also be derived without taking the inverse of the POVM superoperator, but by simply inspecting the left–right action of $\mathcal{F}$ on a quantum state under its definition (equation (31)), and then under the above identity (equation (50)).

The above formulae simplify further in the important special case of a tight *rank-one* IC-POVM. The frame constant then takes its maximum possible value,

$$a = a(F) = \frac{1}{d+1}. \tag{54}$$

Since this is in fact *only* possible for rank-one POVMs, by noting that equation (50) can be taken as an alternative definition in the general case, we obtain the following elegant alternative definition of a tight rank-one IC-POVM.

Proposition 12. *Let* $F : \mathfrak{B}(\mathscr{X}) \to \mathrm{End}(\mathbb{C}^d)$ *be a POVM. Then F is a tight rank-one IC-POVM iff*

$$\mathcal{F} = \frac{\mathbf{I}+\mathcal{I}}{d+1}. \tag{55}$$

The state-reconstruction formula for a tight rank-one IC-POVM also takes an elegant form:

$$\rho = (d+1)\int_{\mathscr{X}} \mathrm{d}p(x)\pi(x) - I, \tag{56}$$

where we have set the POVD to a rank-one projector, $P \equiv \pi$, to emphasize the fact that we are now dealing exclusively with rank-one POVMs. It is then appropriate to consider the measurement outcomes as points in complex projective space, $\mathscr{X} \subseteq \mathbb{C}P^{d-1}$.

We can say some more about the structure of tight rank-one IC-POVMs. Note that $\mathrm{End}(\mathrm{End}(\mathbb{C}^d)) \cong \mathrm{End}(\mathbb{C}^d) \otimes \mathrm{End}(\mathbb{C}^d)$. The natural isomorphism which enables this relationship amounts to replacing each '$\odot$' by '$\otimes$' for a superoperator written in terms of its ordinary action. Rewriting equation (55) in terms of the ordinary action

$$\int_{\mathscr{X}} \mathrm{d}\tau(x)\pi(x) \odot \pi(x) = \frac{1}{d+1}\left(\sum_k E_k \odot E_k^\dagger + I \odot I\right), \tag{57}$$

we see that the condition for a tight rank-one IC-POVM is equivalent to

$$\int_{\mathscr{X}} \mathrm{d}\tau(x)\pi(x) \otimes \pi(x) = \frac{1}{d+1}\left(\sum_k E_k \otimes E_k^\dagger + I \otimes I\right) \tag{58}$$

$$= \frac{1}{d+1}(T + I \otimes I) \tag{59}$$

$$= \frac{2}{d+1}\Pi_{\mathrm{sym}}^{(2)}, \tag{60}$$

where the swap, $T \equiv \sum_{j,k} |e_j\rangle\langle e_k| \otimes |e_k\rangle\langle e_j| = \sum_k E_k \otimes E_k^\dagger$, for any orthonormal operator basis. With $\omega = \tau/d$ in definition 5, we see that tight rank-one IC-POVMs are equivalent to

complex projective 2-designs. A diligent reader might have predicted this outcome from the proof of theorem 4.

Proposition 13. *A rank-one POVM, $F : \mathfrak{B}(\mathscr{X}) \to \mathrm{End}(\mathbb{C}^d)$, $\mathscr{X} \subseteq \mathbb{C}P^{d-1}$, $P \equiv \pi$, is a tight IC-POVM iff the outcome distribution $(\mathscr{X}, \tau/d)$ is a 2-design, i.e.*

$$\int_{\mathscr{X}} \mathrm{d}\tau(x)\pi(x) \otimes \pi(x) = \frac{2}{d+1}\Pi_{\mathrm{sym}}^{(2)}. \tag{61}$$

By theorem 4, there is essentially a unique *minimal* tight rank-one IC-POVM for each dimension, i.e. one with $|\mathscr{X}| = d^2$. This IC-POVM corresponds to a tight 2-design, which in the context of quantum measurements, is called a *symmetric IC-POVM* (*SIC-POVM*) [24]. The defining properties are $\tau(x) \equiv 1/d$, and

$$(\pi(x)|\pi(y)) = |\langle x|y\rangle|^2 = \frac{d\delta(x, y) + 1}{d+1}. \tag{62}$$

Although analytical constructions are known only for $d \leqslant 10$ and $d = 12, 13, 19$ [24, 27, 47–50], SIC-POVMs are conjectured to exist in all dimensions [27, 24] (see also [65–69]). Embedded in $\mathrm{H}_0(\mathbb{C}^d) \cong \mathbb{R}^{d^2-1}$, the elements of a SIC-POVM correspond to the vertices of a regular simplex:

$$\frac{d}{d-1}(\pi(x) - I/d|\pi(y) - I/d) = \frac{d^2\delta(x, y) - 1}{d^2 - 1}. \tag{63}$$

However not all simplices will correspond to a POVM. The factor of $d/(d-1)$ is the result of embedding $\mathrm{Q}(\mathbb{C}^d)$ into the sphere of radius $\sqrt{(d-1)/d}$ in $\mathrm{H}_0(\mathbb{C}^d)$ rather than the unit sphere.

Following the terminology of frame theory, a finite tight rank-one IC-POVM will be called *uniform* when $\tau(x) \equiv d/|\mathscr{X}|$, or *equiangular* [70] if we additionally have $(\pi(x)|\pi(y)) = |\langle x|y\rangle|^2 = c$ for all $x \neq y \in \mathscr{X}$ and some constant c. SIC-POVMs are examples of equiangular tight rank-one IC-POVMs. In fact, these are the only POVMs of this type. To show this, first note that the Welch bound (equation (12)) is saturated for both $t = 1$ and $t = 2$ in the case of a tight rank-one IC-POVM. Equiangularity then implies that, respectively,

$$c = \frac{n - d}{d(n - 1)} \qquad \text{and} \qquad c^2 = \frac{2n - d(d + 1)}{d(d + 1)(n - 1)}, \tag{64}$$

where we have set $|\mathscr{X}| = n$. The only solution to these equations is $n = d^2$ and $c = 1/(d+1)$.

Another important example of a tight rank-one IC-POVM is a complete set of mutually unbiased bases (MUBs) [25, 26]. That is, a set of $d + 1$ orthonormal bases for $\mathbb{C}^d$ with a constant overlap of $1/d$ between elements of different bases:

$$\left(\pi\left(e_j^l\right)\middle|\pi\left(e_k^m\right)\right) = \left|\left\langle e_j^l\middle|e_k^m\right\rangle\right|^2 = \begin{cases} \delta(j, k), & l = m \\ 1/d, & l \neq m. \end{cases} \tag{65}$$

Using theorem 6 it is straightforward to check that the union of $d + 1$ MUBs $\mathscr{D} = \{e_k^m | 1 \leqslant k \leqslant d, 1 \leqslant m \leqslant d + 1\}$ forms a 2-design with uniform weight $w \equiv 1/|\mathscr{D}| = 1/d(d + 1)$ [28, 29]. Thus with $\tau(x) \equiv 1/(d + 1)$ and $\mathscr{X} = \mathscr{D}$ we have a uniform tight rank-one IC-POVM. Embedded in $\mathrm{H}_0(\mathbb{C}^d) \cong \mathbb{R}^{d^2-1}$, the elements of a basis correspond to the vertices of a regular simplex in the $(d - 1)$-dimensional subspace which they span. A complete set of MUBs corresponds to a maximal set of $d + 1$ mutually orthogonal subspaces:

$$\frac{d}{d-1}\left(\pi\left(e_j^l\right) - I/d\middle|\pi\left(e_k^m\right) - I/d\right) = \begin{cases} \dfrac{d\delta(j, k) - 1}{d - 1}, & l = m \\ 0, & l \neq m. \end{cases} \tag{66}$$

Such IC-POVMs allow state determination via orthogonal measurements. The reconstruction formula is given by equation (56). Although constructions are known for prime-power dimensions [25, 26] (see also [71–73]), a complete set of MUBs is unlikely to exist in all dimensions.

Finally, let us rewrite the frame bound (theorem 8) for the context of quantum measurements.

Corollary 14. *Let* $F : \mathfrak{B}(\mathscr{X}) \to \mathrm{End}(\mathbb{C}^d)$ *be a POVM. Then*

$$\iint_{\mathscr{X}} d\tau(x) \, d\tau(y) (P(x)|P(y))^2 \geqslant 1 + \frac{(\mathrm{Tr}(\mathcal{F}) - 1)^2}{d^2 - 1}, \tag{67}$$

with equality iff F is a tight IC-POVM.

Proof. The frame bound (equation (28)) takes the general form $\mathrm{Tr}(\mathcal{A}^2) \geqslant (\mathrm{Tr}(\mathcal{A}))^2/D$, where D is the dimension of the operator space. Setting $\mathcal{A} = \mathcal{F} - \mathcal{I}/d$ and $D = d^2 - 1$ for $\mathrm{H}_0(\mathbb{C}^d)$ then gives equation (67) (using equation (42)).   □

A POVM is a rank-one POVM if and only if $\mathrm{Tr}(\mathcal{F}) = d$. With this value in the RHS of equation (67) we recover the Welch bound (theorem 6) for $t = 2$.

Corollary 15. *Let* $F : \mathfrak{B}(\mathscr{X}) \to \mathrm{End}(\mathbb{C}^d)$, $\mathscr{X} \subseteq \mathbb{C}P^{d-1}$, $P \equiv \pi$, *be a rank-one POVM. Then*

$$\iint_{\mathscr{X}} d\tau(x) \, d\tau(y) (\pi(x)|\pi(y))^2 \geqslant \frac{2d}{d+1}, \tag{68}$$

with equality iff F is a tight IC-POVM.

These corollaries tell us that tight IC-POVMs are those which minimize the average pairwise correlation in the POVD. An operational interpretation of this fact will be given in section 7. It is interesting to note that the above two examples of uniform tight rank-one IC-POVMs, SIC-POVMs and complete sets of MUBs, also minimize the maximal pairwise correlation. As spherical codes [74] on the sphere of radius $\sqrt{(d-1)/d}$ in $\mathrm{H}_0(\mathbb{C}^d)$, SIC-POVMs saturate the simplex bound whilst complete sets of MUBs saturate the orthoplex bound (see e.g. [16]).

6. Optimal linear quantum state tomography

Informationally complete quantum measurements are precisely those measurements which can be used for quantum state tomography. In this section we will show that, amongst all IC-POVMs, the tight rank-one IC-POVMs are the most robust against statistical error in the quantum tomographic process. We will also find that, for an arbitrary IC-POVM, the canonical dual frame with respect to the trace measure is the optimal dual frame for state reconstruction, thus confirming the approach of section 4. These results, however, are shown only for the special case of *linear* quantum state tomography, which will be described later in this section.

Consider a state-reconstruction formula of the form

$$|\rho) = \int_{\mathscr{X}} dp(x)|Q(x)) = \int_{\mathscr{X}} (dF(x)|\rho)|Q(x)), \tag{69}$$

where $Q : \mathscr{X} \to \mathrm{End}(\mathbb{C}^d)$ is an OVD. If this formula is to remain valid for all ρ, then we must have

$$\int_{\mathscr{X}} |Q(x))(dF(x)| = \mathbf{I}, \tag{70}$$

which without loss of generality, can be rewritten as

$$\int_{\mathscr{X}} \mathrm{d}\tau(x)|Q(x))(P(x)| = \mathbf{I}, \tag{71}$$

where the POVD P and trace measure τ are defined in section 4 (equation (30)). Equation (71) restricts $\{Q(x)\}_{x \in \mathscr{X}}$ to a dual frame of $\{P(x)\}_{x \in \mathscr{X}}$ with respect to the trace measure. Our first goal is to find the optimal dual frame.

It will be instructive to start with the special case of a discrete IC-POVM. Suppose that we take N random samples, $y_1, \ldots, y_N$, from a countable set $\mathscr{X}$, where the outcome x occurs with some unknown probability $p(x)$. Our estimate for this probability is

$$\hat{p}(x) = \hat{p}(x; y_1, \ldots, y_N) \equiv \frac{1}{N} \sum_{k=1}^{N} \delta(x, y_k), \tag{72}$$

which of course obeys the expectation $\mathrm{E}[\hat{p}(x)] = p(x)$. An elementary calculation shows that the expected covariance for N samples is

$$\mathrm{E}[(p(x) - \hat{p}(x))(p(y) - \hat{p}(y))] = \frac{1}{N}(p(x)\delta(x, y) - p(x)p(y)). \tag{73}$$

Now suppose that $p(x)$ are outcome probabilities for an informationally complete quantum measurement of the state $\rho \in \mathrm{Q}(\mathbb{C}^d)$. That is, $p(x) = \mathrm{tr}[F(x)\rho]$, where $\{F(x)\}_{x \in \mathscr{X}} \subset \mathrm{End}(\mathbb{C}^d)$ is a discrete IC-POVM. The error in our estimate of ρ,

$$\hat{\rho} = \hat{\rho}(y_1, \ldots, y_N) \equiv \sum_{x \in \mathscr{X}} \hat{p}(x; y_1, \ldots, y_N)Q(x), \tag{74}$$

as measured by the squared Hilbert–Schmidt (or Frobenius) distance, is

$$\|\rho - \hat{\rho}\|^2 = (\rho - \hat{\rho}|\rho - \hat{\rho}) = \sum_{x,y \in \mathscr{X}} (p(x) - \hat{p}(x))(p(y) - \hat{p}(y))(Q(x)|Q(y)), \tag{75}$$

which has the expectation

$$\mathrm{E}[\|\rho - \hat{\rho}\|^2] = \frac{1}{N} \sum_{x,y \in \mathscr{X}} (p(x)\delta(x, y) - p(x)p(y))(Q(x)|Q(y)) \tag{76}$$

$$= \frac{1}{N} \left(\sum_{x \in \mathscr{X}} p(x)(Q(x)|Q(x)) - \mathrm{tr}(\rho^2) \right) \tag{77}$$

$$\equiv \frac{1}{N}(\Delta_p(Q) - \mathrm{tr}(\rho^2)), \tag{78}$$

using equation (73) and then (69). This expression is also a fitting description of the error from an IC-POVM with a continuum of measurement outcomes if we define

$$\Delta_p(Q) \equiv \int_{\mathscr{X}} \mathrm{d}p(x)(Q(x)|Q(x)) \tag{79}$$

in general. This follows from the fact that a countable partition of the outcome set $\mathscr{X}$ allows any continuous IC-POVM to be approximated by a discrete IC-POVM. Our estimate $\hat{p}$ remains a good approximation for the probability measure p, except now with x and $y_1, \ldots, y_N$ in equation (72) indicating members of the partition. In the limit of finer approximating partitions we again arrive at equation (78) for the average error, but now with equation (79) for $\Delta_p(Q)$.

Since we have no control over the purity of ρ, it is the quantity $\Delta_p(Q)$ in equation (78) which is now of interest. The IC-POVM which minimizes $\Delta_p(Q)$, and hence the error, will in

general depend on the quantum state under examination. We thus set $\rho = \rho(\sigma, U) \equiv U\sigma U^\dagger$, and now remove this dependence by taking the (Haar) average over all $U \in \mathrm{U}(d)$:

$$\int_{\mathrm{U}(d)} \mathrm{d}\mu_{\mathrm{H}}(U) \Delta_p(Q) = \int_{\mathrm{U}(d)} \mathrm{d}\mu_{\mathrm{H}}(U) \int_{\mathscr{X}} \mathrm{tr}[\mathrm{d}F(x) U\sigma U^\dagger](Q(x)|Q(x)) \tag{80}$$

$$= \frac{1}{d} \int_{\mathscr{X}} \mathrm{tr}[\mathrm{d}F(x)] \, \mathrm{tr}(\sigma)(Q(x)|Q(x)) \tag{81}$$

$$= \frac{1}{d} \int_{\mathscr{X}} \mathrm{d}\tau(x)(Q(x)|Q(x)) \tag{82}$$

$$\equiv \frac{1}{d} \Delta_\tau(Q), \tag{83}$$

using Shur's lemma for the integral and then setting $\tau \equiv \mathrm{tr}(F)$. The quantity $\Delta_\tau(Q)/d$ is the average value of $\Delta_p(Q)$ when ρ is chosen randomly from an isotropic distribution in Euclidean space (via equation (46)).

We will now minimize $\Delta_\tau(Q)$ over all choices for Q, while keeping the IC-POVM F fixed. Our only constraint is that $\{Q(x)\}_{x\in\mathscr{X}}$ remains a dual frame to $\{P(x)\}_{x\in\mathscr{X}}$, so that the reconstruction formula (equation (69)) remains valid for all ρ. The following lemma shows that the reconstruction OVD defined in section 4, $\{R(x)\}_{x\in\mathscr{X}}$ (equation (38)), is the optimal choice for the dual frame.

Lemma 16. *Let $\{A(x)\}_{x\in\mathscr{X}} \subseteq \mathrm{End}(\mathbb{C}^d)$ be an operator frame with respect to the measure α. Then for all dual frames $\{B(x)\}_{x\in\mathscr{X}}$,*

$$\Delta_\alpha(B) \equiv \int_{\mathscr{X}} \mathrm{d}\alpha(x)(B(x)|B(x)) \geqslant \int_{\mathscr{X}} \mathrm{d}\alpha(x)(\tilde{A}(x)|\tilde{A}(x)) \equiv \Delta_\alpha(\tilde{A}), \tag{84}$$

with equality only if $B \equiv \tilde{A}$, α-almost everywhere, where $\{\tilde{A}(x)\}_{x\in\mathscr{X}}$ is the canonical dual frame.

Proof. Define $D \equiv B - \tilde{A}$ which satisfies

$$\int_{\mathscr{X}} \mathrm{d}\alpha(x)|\tilde{A}(x))(D(x)| = \int_{\mathscr{X}} \mathrm{d}\alpha(x)|\tilde{A}(x))(B(x)| - \int_{\mathscr{X}} \mathrm{d}\alpha(x)|\tilde{A}(x))(\tilde{A}(x)| \tag{85}$$

$$= \int_{\mathscr{X}} \mathrm{d}\alpha(x)\mathcal{A}^{-1}|A(x))(B(x)| - \int_{\mathscr{X}} \mathrm{d}\alpha(x)\mathcal{A}^{-1}|A(x))(A(x)|\mathcal{A}^{-1} \tag{86}$$

$$= \mathcal{A}^{-1}\mathbf{I} - \mathcal{A}^{-1}\mathcal{A}\mathcal{A}^{-1} \tag{87}$$

$$= 0, \tag{88}$$

when $\{B(x)\}_{x\in\mathscr{X}}$ is a dual frame to $\{A(x)\}_{x\in\mathscr{X}}$ and $\{\tilde{A}(x)\}_{x\in\mathscr{X}}$ is the canonical dual frame, using equations (21), (25) and (26). Thus

$$\int_{\mathscr{X}} \mathrm{d}\alpha(x)(D(x)|\tilde{A}(x)) = 0, \tag{89}$$

and

$$\int_{\mathscr{X}} \mathrm{d}\alpha(x)(B(x)|B(x)) = \int_{\mathscr{X}} \mathrm{d}\alpha(x)(\tilde{A}(x)|\tilde{A}(x)) + \int_{\mathscr{X}} \mathrm{d}\alpha(x)(\tilde{A}(x)|D(x)) \tag{90}$$

$$+ \int_{\mathscr{X}} \mathrm{d}\alpha(x)(D(x)|\tilde{A}(x)) + \int_{\mathscr{X}} \mathrm{d}\alpha(x)(D(x)|D(x)) \tag{91}$$

$$= \int_{\mathscr{X}} d\alpha(x)(\tilde{A}(x)|\tilde{A}(x)) + \int_{\mathscr{X}} d\alpha(x)(D(x)|D(x)) \tag{92}$$

$$\geqslant \int_{\mathscr{X}} d\alpha(x)(\tilde{A}(x)|\tilde{A}(x)), \tag{93}$$

with equality if and only if $D \equiv 0$, α-almost everywhere. $\qquad\square$

Setting $A \equiv P$ and $\alpha \equiv \tau$ in lemma 16 confirms the reconstruction method presented in section 4 (equations (31), (38) and (40)). Note that we can retain the dependence on ρ by simply replacing τ by pd in these formulae. An adaptive reconstruction method might make use of this fact. Equation (44) shows that $\Delta_\tau(R) = \text{Tr}(\mathcal{F}^{-1})$. This quantity will now be minimized over all IC-POVMs.

Lemma 17. *Let $F : \mathfrak{B}(\mathscr{X}) \to \text{End}(\mathbb{C}^d)$ be an IC-POVM. Then*

$$\text{Tr}(\mathcal{F}^{-1}) \geqslant d\big(d(d+1) - 1\big), \tag{94}$$

with equality iff F is a tight rank-one IC-POVM.

Proof. We will minimize the quantity

$$\text{Tr}(\mathcal{F}^{-1}) = \sum_{k=1}^{d^2} \frac{1}{\lambda_k}, \tag{95}$$

where $\lambda_1, \ldots, \lambda_{d^2} > 0$ denote the left–right eigenvalues of $\mathcal{F}$. These eigenvalues satisfy the constraint

$$\sum_{k=1}^{d^2} \lambda_k = \text{Tr}(\mathcal{F}) = \int_{\mathscr{X}} d\tau(x)(P(x)|P(x)) \leqslant \int_{\mathscr{X}} d\tau(x) = d, \tag{96}$$

since $\text{tr}(P^2) \leqslant 1$, τ-almost everywhere. We know, however, that the identity operator is always a left–right eigenvector of $\mathcal{F}$ with unit eigenvalue (equation (42)). Thus we in fact have $\lambda_1 = 1$, say, and then $\sum_{k=2}^{d^2} \lambda_k \leqslant d - 1$. Under this latter constraint it is straightforward to show that the RHS of equation (95) takes its minimum value if and only if $\lambda_2 = \cdots = \lambda_{d^2} = (d-1)/(d^2-1) = 1/(d+1)$, or equivalently,

$$\mathcal{F} = 1 \cdot \frac{\mathcal{I}}{d} + \frac{1}{d+1} \cdot \mathbf{\Pi}_0 = \frac{\mathbf{I} + \mathcal{I}}{d+1}, \tag{97}$$

since the subspace of traceless operators is $\mathcal{F}$-invariant (equation (45)). Therefore, by proposition 12, $\text{Tr}(\mathcal{F}^{-1})$ takes its minimum value if and only if F is a tight rank-one IC-POVM. The minimum is $\text{Tr}(\mathcal{F}^{-1}) = 1 \cdot 1 + (d+1) \cdot (d^2 - 1) = d(d(d+1) - 1)$. $\qquad\square$

We have thus confirmed that it is optimal to use a tight rank-one IC-POVM for quantum state tomography. The optimal state-reconstruction formula is then given by equation (56). Before stating these results in a theorem, let us first fully clarify the assumptions that have allowed us to draw this conclusion. First of all, we have chosen the Hilbert–Schmidt metric to measure distances in $\text{Q}(\mathbb{C}^d)$ (see equation (75)). There are other choices to consider and some of these are no doubt more appropriate in the context of quantum states. For example, we could instead quantify the error in $\hat{\rho}$ with the Bures metric [75, 76] $d_{\text{B}}(\rho, \hat{\rho})^2 \equiv 2 - 2\,\text{tr}\sqrt{\sqrt{\rho}\hat{\rho}\sqrt{\rho}}$, or, although not strictly a metric, the relative entropy $S(\rho||\hat{\rho}) \equiv \text{tr}(\rho \log \rho - \rho \log \hat{\rho})$. These choices, however, proved to cumbersome to warrant a detailed investigation in the present paper.

We have also made assumptions about the procedure for state reconstruction. This can be explained as follows. For an informationally complete POVM, F say, every quantum state

is uniquely identified by its measurement statistics. This does not mean, however, that all points on the probability simplex, $\int_{\mathscr{X}} \mathrm{d}p(x) = 1$, describe valid outcome statistics for a measurement (with IC-POVM F) of a quantum state. Due to the possible overcompleteness of a POVM, there can be many choices for the estimate statistics $\hat{p}$ (just as there were many choices for the reconstruction OVD Q) which satisfy $\int_{\mathscr{X}} \mathrm{d}\hat{p}(x)Q(x) = \rho$ (equation (69)) for some fixed $\rho \in \mathrm{Q}(\mathbb{C}^d)$. The state's actual measurement statistics, $p \equiv \mathrm{tr}(F\rho)$, are only but one of these choices. Additionally, for some choices of the estimate statistics we might not even have $\int_{\mathscr{X}} \mathrm{d}\hat{p}(x)Q(x) \in \mathrm{Q}(\mathbb{C}^d)$.

Suppose, for example, that we have a finite IC-POVM with $|\mathscr{X}| = d^2 + K$ possible measurement outcomes. We know that every POVM satisfies the normalization constraint, $\sum_{x \in \mathscr{X}} F(x) = I$, which implies normalization of the statistics: $\sum_{x \in \mathscr{X}} p(x) = 1$. Our previous estimate (equation (72)) satisfies this constraint. It does not, however, incorporate any additional constraints specific to the particular choice of IC-POVM. Embedding the POVM elements in $\mathrm{H}_0(\mathbb{C}^d)$ shows that there will be a further K linear constraints of the form

$$\sum_{x \in \mathscr{X}} c_k(x)F(x) = 0, \quad \text{which imply that} \quad \sum_{x \in \mathscr{X}} c_k(x)p(x) = 0 \tag{98}$$
$$(c_k(x) \in \mathbb{R}, k = 1, \ldots, K).$$

The intersection of the probability simplex in $\mathbb{R}^{d^2+K}$ with the subspace perpendicular to the K vectors $\{c_k(x)\}_{x \in \mathscr{X}}$ forms the subset of statistics which are isomorphic, under the mapping $p = \mathrm{tr}(FA) \to A$, to the normalized Hermitian operators in $\mathrm{End}(\mathbb{C}^d)$. We can thus excise all unphysical estimate statistics which duplicate valid measurement statistics by taking these extra constraints into account. After N measurements, with results $y_1, \ldots, y_N$, the most appropriate choice for $\hat{p}$ will be the maximum-likelihood estimate under these constraints, i.e. that which maximizes $\mathrm{Prob}(p) = \prod_{x \in \mathscr{X}} p(x)^{n(x)}$, where $n(x) \equiv \sum_{k=1}^{N} \delta(x, y_k)$. Under the normalization constraint only, it is straightforward to recover $\hat{p}(x) = n(x)/N$ (equation (72)); under both the normalization and additional constraints, however, this nonlinear optimization problem becomes difficult to solve analytically. One exception is an IC-POVM consisting of $d + 1$ MUBs (equation (65)), in which case the $K = d$ additional constraints [$c_k(e_j^l) = (d + 1)\delta(k, l) - 1$ in equation (98)] single out $\hat{p}(e_j^l) = n(e_j^l)/[(d + 1)\sum_{k=1}^{d} n(e_k^l)]$ for the maximum-likelihood estimate, as one should expect. This means we should treat the outcome probabilities as if they came from $d + 1$ separate orthogonal measurements, each corresponding to one of the bases. In the special case of a *minimal* IC-POVM (i.e $|\mathscr{X}| = d^2$) there are no additional constraints and equation (72) is the best estimate for the outcome statistics. For this reason minimal IC-POVMs should be preferred over other IC-POVMs. Lemma 16 is then redundant since the canonical dual frame is the unique dual frame, namely the dual basis. In general, only when all $K + 1$ linear constraints are taken into account is lemma 16 unnecessary and the particular choice of reconstruction formula unimportant.

By taking the maximum-likelihood estimate under the normalization and all $|\mathscr{X}| - d^2$ additional linear constraints we can remove the redundancy in the estimate statistics. There may still remain unphysical statistics however. If ρ is pure, or if N is not large enough, then $\int_{\mathscr{X}} \mathrm{d}\hat{p}(x)Q(x)$ may not be a positive operator under the linear constraints, and thus, not a quantum state. To overcome this problem we must instead apply the single nonlinear constraint that $\hat{p} \in \{\mathrm{tr}(F\rho)|\rho \in \mathrm{Q}(\mathbb{C}^d)\}$, and again take the maximum-likelihood estimate.

To show that the tight rank-one IC-POVMs are optimal for quantum state tomography we have ignored all additional linear and nonlinear constraints on the estimate statistics, and simply taken equation (72) for $\hat{p}$, with a reconstruction formula in the form of equation (69). Although this simplification will likely lead to less than optimal estimates of the quantum state, the inclusion of all possible constraints on $\hat{p}$ for the maximum-likelihood estimation, or only

the linear constraints, makes any generalization of our results considerably more difficult. In this paper we will thus only claim that tight rank-one IC-POVMs are optimal for *linear* quantum state tomography, with the term 'linear' referring to the previous simplified state-reconstruction procedure, i.e. without the nonlinear optimization needed for maximum-likelihood estimation under the additional constraints. This result is summarized in the following theorem.

Theorem 18. *Let $F : \mathfrak{B}(\mathscr{X}) \to \mathrm{End}(\mathbb{C}^d)$ be an IC-POVM and let $\rho = \rho(\sigma, U) \equiv U\sigma U^\dagger$ for some fixed quantum state $\sigma \in \mathrm{Q}(\mathbb{C}^d)$. Then*

$$e_{\mathrm{av}}^{(F,Q)}(\sigma) \equiv \int_{\mathrm{U}(d)} \mathrm{d}\mu_{\mathrm{H}}(U)\, \mathrm{E}[\|\rho - \hat\rho\|^2] \geqslant \frac{1}{N}\left(\frac{1}{d}\,\mathrm{Tr}(\mathcal{F}^{-1}) - \mathrm{tr}(\sigma^2)\right) \quad (99)$$

$$\geqslant \frac{1}{N}(d(d+1) - 1 - \mathrm{tr}(\sigma^2)) \quad (100)$$

for all reconstruction OVDs $Q : \mathscr{X} \to \mathrm{End}(\mathbb{C}^d)$ which are dual frames to P, where $\hat\rho = \hat\rho(\sigma, U; y_1, \dots, y_N)$ is a linear tomographic estimate of ρ given N measurement outcomes $y_1, \dots, y_N$ (equations (72) and (74)) and the expectation is over these outcomes. Furthermore, equality in equation (99) occurs iff $Q \equiv R$, τ-almost everywhere, and equality in equation (100) occurs iff F is a tight rank-one IC-POVM.

We can also consider the worst-case expectation in the error. The average then provides a lower bound:

$$e_{\mathrm{wc}}(\sigma) \equiv \sup_{U \in \mathrm{U}(d)} \mathrm{E}[\|\rho - \hat\rho\|^2] \geqslant e_{\mathrm{av}}(\sigma) \geqslant \frac{1}{N}(d(d+1) - 1 - \mathrm{tr}(\sigma^2)). \quad (101)$$

Note, however, that if $R = (d+1)P - I = (d+1)\pi - I$, as defined for a tight rank-one IC-POVM (equations (52) and (54) with $P = \pi$), then $(R|R) = \mathrm{tr}(R^2) = d(d+1) - 1$, τ-almost everywhere. Consequently, returning to equation (78) but now with $Q = R$ and $\rho = \rho(\sigma, U) \equiv U\sigma U^\dagger$, we see that regardless of the choice of $U \in \mathrm{U}(d)$, for a tight rank-one IC-POVM we always have

$$e(\sigma, U) \equiv \mathrm{E}[\|\rho - \hat\rho\|^2] = \frac{1}{N}\left(\int_{\mathscr{X}} \mathrm{d}p(x)(R(x)|R(x)) - \mathrm{tr}(\sigma^2)\right) \quad (102)$$

$$= \frac{1}{N}\left((d(d+1) - 1)\int_{\mathscr{X}} \mathrm{d}p(x) - \mathrm{tr}(\sigma^2)\right) \quad (103)$$

$$= \frac{1}{N}(d(d+1) - 1 - \mathrm{tr}(\sigma^2)) \quad (104)$$

when equation (56) is used for state reconstruction. The above inequality (equation (101)) and this last fact imply the following corollary to theorem 18.

Corollary 19. *Let $F : \mathfrak{B}(\mathscr{X}) \to \mathrm{End}(\mathbb{C}^d)$ be an IC-POVM and let $\rho = \rho(\sigma, U) \equiv U\sigma U^\dagger$ for some fixed quantum state $\sigma \in \mathrm{Q}(\mathbb{C}^d)$. Then*

$$e_{\mathrm{wc}}^{(F,Q)}(\sigma) \equiv \sup_{U \in \mathrm{U}(d)} \mathrm{E}[\|\rho - \hat\rho\|^2] \geqslant \frac{1}{N}(d(d+1) - 1 - \mathrm{tr}(\sigma^2)) \quad (105)$$

for all reconstruction OVDs $Q : \mathscr{X} \to \mathrm{End}(\mathbb{C}^d)$ which are dual frames to P, where $\hat\rho = \hat\rho(\sigma, U; y_1, \dots, y_N)$ is a linear tomographic estimate of ρ given N measurement outcomes $y_1, \dots, y_N$ (equations (72) and (74)) and the expectation is over these outcomes. Furthermore, equality in equation (105) occurs iff $Q \equiv R$, τ-almost everywhere, and F is a tight rank-one IC-POVM.

Tight rank-one IC-POVMs are thus optimal for linear quantum state tomography in both an average and worst-case sense. In fact, they form the unique class of POVMs capable of achieving

$$e_{\text{wc}}(\sigma) = e_{\text{av}}(\sigma) = e(\sigma, U) = \frac{1}{N}(d(d+1) - 1 - \text{tr}(\sigma^2)). \tag{106}$$

The type of quantum state tomography considered in this section was based on nonadaptive sequential measurements on copies of the quantum state. This restriction is detrimental to the tomographic process. Given multiple copies of a state, there exist joint measurements on these copies which will outperform any of the measurements considered above (see e.g. [77]). In the following section, however, we will show that the tight rank-one IC-POVMs form the unique class of POVMs which are optimal for state estimation, if given only a single copy of a pure quantum state.

7. Optimal measurement-based cloners

A natural way of assessing the capability of a measuring instrument for state estimation is to consider it in the role of a cloning machine [30, 78–82]. A single copy of an unknown pure quantum state $\psi \in \mathbb{C}P^{d-1}$ is the input to this device, while the output is a finite number of approximate copies of ψ, or in the case of a measurement, an infinite supply of approximate copies described by a single mixed quantum state. This estimate will in general depend on the measurement result. For outcome x we will denote the device's output state by $\hat{\rho}(x) \in Q(\mathbb{C}^d)$. The probability of confirming $\hat{\rho}(x)$ to be $\pi(\psi)$ is then given by the fidelity, $f(\psi, x) \equiv \langle\psi|\hat{\rho}(x)|\psi\rangle$. The average fidelity over all measurement outcomes,

$$f(\psi) \equiv \int_{\mathscr{X}} \text{tr}[dF(x)\pi(\psi)]f(\psi, x) = \int_{\mathscr{X}} d\tau(x)\langle\psi|P(x)|\psi\rangle\langle\psi|\hat{\rho}(x)|\psi\rangle, \tag{107}$$

is the probability that the POVM F, together with the estimate state $\hat{\rho}$, successfully clones ψ. Maximized over all choices for $\hat{\rho}$, this quantity might be interpreted as an operational measure of knowledge (about ψ) gained from the measurement. For the purposes of this section we will call the pair $(F, \hat{\rho})$ a *measurement-based cloning strategy*.

Consider the average success probability for such strategies:

$$f_{\text{av}} \equiv \int_{\mathbb{C}P^{d-1}} d\mu_{\text{H}}(\psi)f(\psi) \tag{108}$$

$$= \int_{\mathbb{C}P^{d-1}} d\mu_{\text{H}}(\psi) \int_{\mathscr{X}} d\tau(x) \, \text{tr}[\pi(\psi)^{\otimes 2} \cdot P(x) \otimes \hat{\rho}(x)] \tag{109}$$

$$= \frac{2}{d(d+1)} \int_{\mathscr{X}} d\tau(x) \, \text{tr}\left[\Pi_{\text{sym}}^{(2)} \cdot P(x) \otimes \hat{\rho}(x)\right] \tag{110}$$

$$= \frac{1}{d(d+1)} \int_{\mathscr{X}} d\tau(x)(1 + \text{tr}[P(x)\hat{\rho}(x)]) \tag{111}$$

$$\leqslant \frac{2}{d+1}. \tag{112}$$

Here we have used lemma 1 and then the identity $2 \, \text{tr}\left(\Pi_{\text{sym}}^{(2)} \cdot A \otimes B\right) = \text{tr}(A)\,\text{tr}(B) + \text{tr}(AB)$. Equality will occur if and only if $\text{tr}(P\hat{\rho}) = 1$, τ-almost everywhere, in which case we must have $\hat{\rho} = P = \pi$, where we now consider $\mathscr{X} \subseteq \mathbb{C}P^{d-1}$. Thus F is capable of achieving the maximum possible average success probability if and only if it is a rank-one POVM. It is no surprise that the best choice for the estimate state is then given by the POVD.

But can we ask for more from the measuring instrument? Let us instead maximize the worst-case success probability. This quantity may be thought of as a guarantee on the success rate. The average success probability provides an upper bound:

$$f_{\text{wc}} \equiv \inf_{\psi \in \mathbb{C}P^{d-1}} f(\psi) \leqslant f_{\text{av}} \leqslant \frac{2}{d+1}. \tag{113}$$

Now consider the conditions upon which equality is achieved. First of all we need $f_{\text{av}} = 2/(d+1)$ and, thus, we require a rank-one POVM $P = \pi$ with the estimate state $\hat{\rho} = \pi$. If additionally we have $f_{\text{wc}} = f_{\text{av}}$ then the variance in the success probability must necessarily vanish, or equivalently, $\int_{\mathbb{C}P^{d-1}} \mathrm{d}\mu_{\text{H}}(\psi) f(\psi)^2 = f_{\text{av}}^2 = 4/(d+1)^2$. The second moment may be calculated in a similar manner to the first:

$$\int_{\mathbb{C}P^{d-1}} \mathrm{d}\mu_{\text{H}}(\psi) f(\psi)^2 = \int_{\mathbb{C}P^{d-1}} \mathrm{d}\mu_{\text{H}}(\psi) \int\int_{\mathscr{X}} \mathrm{d}\tau(x)\,\mathrm{d}\tau(y)\, \mathrm{tr}[\pi(\psi)^{\otimes 4} \cdot \pi(x)^{\otimes 2} \otimes \pi(y)^{\otimes 2}] \tag{114}$$

$$= \frac{24}{d(d+1)(d+2)(d+3)} \int\int_{\mathscr{X}} \mathrm{d}\tau(x)\,\mathrm{d}\tau(y)\, \mathrm{tr}\left[\Pi_{\text{sym}}^{(4)} \cdot \pi(x)^{\otimes 2} \otimes \pi(y)^{\otimes 2}\right] \tag{115}$$

$$= \frac{4}{d(d+1)(d+2)(d+3)} \int\int_{\mathscr{X}} \mathrm{d}\tau(x)\,\mathrm{d}\tau(y)(1 + 4\,\mathrm{tr}[\pi(x)\pi(y)] + \mathrm{tr}[\pi(x)\pi(y)]^2) \tag{116}$$

$$= \frac{4}{d(d+1)(d+2)(d+3)} \left(d^2 + 4d + \int\int_{\mathscr{X}} \mathrm{d}\tau(x)\,\mathrm{d}\tau(y)|\langle x|y\rangle|^4\right). \tag{117}$$

Here we have again used lemma 1 and then a similar identity to the above, except this time with 4! terms. Given the second moment, one can easily check that the condition for zero variance is equivalent to

$$\int\int_{\mathscr{X}} \mathrm{d}\tau(x)\,\mathrm{d}\tau(y)(\pi(x)|\pi(y))^2 = \int\int_{\mathscr{X}} \mathrm{d}\tau(x)\,\mathrm{d}\tau(y)|\langle x|y\rangle|^4 = \frac{2d}{d+1}, \tag{118}$$

which by corollary 15, implies that F is a tight rank-one IC-POVM. This condition is also sufficient. It is straightforward to confirm that for tight rank-one IC-POVMs, $f(\psi) = 2/(d+1)$ independent of ψ.

We have just shown that the worst-case success probability, for a measuring instrument in the role of a cloning machine, can take its maximum value if and only if the corresponding POVM is a tight rank-one IC-POVM. In fact, the tight rank-one IC-POVMs form the unique class of POVMs capable of achieving $f_{\text{wc}} = f_{\text{av}} = f(\psi) = 2/(d+1)$. It is in this sense that a tight rank-one IC-POVM can be claimed optimal for state estimation. Note that, unlike a generic rank-one POVM, a strategy based on a tight rank-one IC-POVM outputs on average an isotropically unbiased estimate of the input state:

$$\mathrm{E}[\hat{\rho}(x)] = \int_{\mathscr{X}} \mathrm{tr}[\mathrm{d}F(x)\pi(\psi)]\hat{\rho}(x) = \int_{\mathscr{X}} \mathrm{d}\tau(x)\langle\psi|\pi(x)|\psi\rangle\pi(x) \tag{119}$$

$$= \int_{\mathscr{X}} \mathrm{d}\tau(x)|\pi(x))(\pi(x)|\pi(\psi)) \tag{120}$$

$$= \mathcal{F}|\pi(\psi)) \tag{121}$$

$$= \frac{I + |\psi\rangle\langle\psi|}{d+1}, \tag{122}$$

where we have used proposition 12. Only the tight rank-one IC-POVMs satisfy equation (122), which could be taken as a defining property. Let us now restate the above facts formally in a theorem.

Theorem 20. *Let $(F, \hat{\rho})$ be a measurement-based cloning strategy with POVM $F : \mathfrak{B}(\mathcal{X}) \to$ End$(\mathbb{C}^d)$. Then*

$$f_{\text{wc}}^{(F,\hat{\rho})} \equiv \inf_{\psi \in \mathbb{C}P^{d-1}} \int_{\mathcal{X}} \text{tr}[\mathrm{d}F(x)\pi(\psi)] \, \text{tr}[\hat{\rho}(x)\pi(\psi)] \leqslant \frac{2}{d+1}, \tag{123}$$

with equality if and only if F is a tight rank-one IC-POVM and $\hat{\rho} = P$.

This theorem is in fact a special case of the results of Hayashi *et al* [30] (see also [81]). If instead N copies of ψ are given, then the optimal joint measurement on these copies that maximizes the average success probability is defined by an N-design. The success probability then increases to $f_{\text{av}} = (N + 1)/(N + d)$ [82]. The measurement that maximizes the worst-case success probability is instead defined by an $(N + 1)$-design, in which case we have $f_{\text{wc}} = f_{\text{av}} = f(\psi) = (N+1)/(N+d)$. The extension to mixed states for this type of quantum state estimation has been considered by Vidal *et al* [77].

8. Conclusion

In this paper we have introduced a special class of informationally complete POVMs which, in analogy to a similar concept in frame theory, are named *tight IC-POVMs*. Embedded as a tight frame in the vector space of all traceless Hermitian operators, which is the natural place to study a quantum state, a tight IC-POVM is as close as possible to an orthonormal basis. It is in this sense that the tight IC-POVMs can be promoted as being special amongst all IC-POVMs. They allow painless quantum state tomography through a particularly simple state-reconstruction formula (equation (53)). The rank-one members of this class have the minimum average pairwise correlation in the POVD for any rank-one POVM (corollary 15) and thus form the family of optimal measurement-based cloners (theorem 20). They are also the best choice for linear quantum state tomography (theorem 18 and corollary 19). The outstanding choice amongst all tight rank-one IC-POVMs are the unique minimal members, the SIC-POVMs [24]. These POVMs are the only equiangular tight rank-one IC-POVMs, minimize the maximal pairwise correlation in the POVD, and can thus be considered the closest, now amongst all tight rank-one IC-POVMs, to an orthonormal basis.

Acknowledgments

References

[1] Nielsen M A and Chuang I L 2000 *Quantum Computation and Quantum Information* (Cambridge: Cambridge University Press)
[2] Davies E B 1976 *Quantum Theory of Open Systems* (London: Academic)
[3] Holevo A S 1982 *Probabilistic and Statistical Aspects of Quantum Theory* (Amsterdam: North-Holland)
[4] Kraus K 1983 *States, Effects and Operations* (Berlin: Springer)
[5] Busch P, Lahti P J and Mittelstaedt P 1996 *The Quantum Theory of Measurement* 2nd edn (Berlin: Springer)
[6] Prugovečki E 1977 Information-theoretic aspects of quantum measurement *Int. J. Theor. Phys.* **16** 321
[7] Busch P and Lahti P J 1989 The determination of the past and the future of a physical system in quantum mechanics *Found. Phys.* **19** 633
[8] Busch P 1991 Informationally complete sets of physical quantities *Int. J. Theor. Phys.* **30** 1217
[9] Hellwig K-E 1993 Quantum measurements and information theory *Int. J. Theor. Phys.* **32** 2401
[10] D'Ariano G M, Perinotti P and Sacchi M F 2004 Informationally complete measurements and groups representation *J. Opt.* B **6** S487

[11] D'Ariano G M, Perinotti P and Sacchi M F 2005 Informationally complete measurements on bipartite quantum systems: comparing local with global measurements *Phys. Rev.* A **72** 042108

[12] Flammia S T, Silberfarb A and Caves C M 2005 Minimal informationally complete measurements for pure states *Found. Phys.* **35** 1985

[13] Weigert S 2006 Simple minimal informationally complete measurements for qudits *Int. J. Mod. Phys.* B **20** 1942

[14] Paris M and Řeháček J (ed) 2004 *Quantum State Estimation* (Berlin: Springer)

[15] Renes J M 2005 Equiangular spherical codes in quantum cryptography *Quantum Inf. Comput.* **5** 81

[16] Scott A J, Walgate J and Sanders B C 2005 Optimal fingerprinting strategies with one-sided error *Preprint* quant-ph/0507048

[17] Fuchs C A 2002 Quantum mechanics as quantum information (and only a little more) *Preprint* quant-ph/0205039

[18] Caves C M, Fuchs C A and Schack R 2002 Unknown quantum states: the quantum de Finetti representation *J. Math. Phys.* **43** 4537

[19] König R and Renner R 2005 A de Finetti representation for finite symmetric quantum states *J. Math. Phys.* **46** 122108

[20] Christensen O 2003 *An Introduction to Frames and Riesz Bases* (Boston, MA: Birkhäuser)

[21] Daubechies I 1992 *Ten Lectures on Wavelets* (Philadelphia, PA: SIAM)

[22] Casazza P G 2000 The art of frame theory *Taiwan. J. Math.* **4** 129

[23] Daubechies I, Grossmann A and Meyer Y 1986 Painless nonorthogonal expansions *J. Math. Phys.* **27** 1271

[24] Renes J M, Blume-Kohout R, Scott A J and Caves C M 2004 Symmetric informationally complete quantum measurements *J. Math. Phys.* **45** 2171

[25] Ivanović I D 1981 Geometrical description of quantal state determination *J. Phys. A: Math. Gen.* **14** 3241

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

[27] Zauner G 1999 Quantendesigns—Grundzüge einer nichtkommutativen designtheorie *PhD Thesis* University of Vienna

[28] Barnum H 2000 Information-disturbance tradeoff in quantum measurement on the uniform ensemble *Department of Computer Science Technical Report* CSTR-00-013 (University of Bristol)

[29] Klappenecker A and Rötteler M 2005 Mutually unbiased bases are complex projective 2-designs *Proc. IEEE Int. Symp. on Information Theory (Adelaide, Australia, Sept. 2005)* p 1740

[30] Hayashi A, Hashimoto T and Horibe M 2005 Reexamination of optimal quantum state estimation of pure states *Phys. Rev.* A **72** 032325

[31] Dankert C, Cleve R, Emerson J and Livine E 2006 Exact and approximate unitary 2-designs: constructions and applications *Preprint* quant-ph/0606161

[32] Delsarte P, Goethals J M and Seidel J J 1977 Spherical codes and designs *Geom. Dedicata* **6** 363

[33] Neumaier A 1981 Combinatorial configurations in terms of distances *Department of Mathematics Memorandum* 81-09 Eindhoven University of Technology

[34] Hoggar S G 1982 *t*-designs in projective spaces *Eur. J. Comb.* **3** 233

[35] Hoggar S G 1984 Parameters of *t*-designs in $\mathbb{F}P^{d-1}$ *Eur. J. Comb.* **5** 29

[36] Hoggar S G 1989 Tight 4 and 5-designs in projective spaces *Graphs Comb.* **5** 87

[37] Hoggar S G 1992 *t*-designs with general angle set *Eur. J. Comb.* **13** 257

[38] Bannai E and Hoggar S G 1985 On tight *t*-designs in compact symmetric spaces of rank one *Proc. Japan Acad.* **61A** 78

[39] Bannai E and Hoggar S G 1989 Tight *t*-designs and squarefree integers *Eur. J. Comb.* **10** 113

[40] Levenshtein V 1992 Designs as maximum codes in polynomial metric spaces *Acta Appl. Math.* **29** 1

[41] Levenshtein V 1998 On designs in compact metric spaces and a universal bound on their size *Discrete Math.* **192** 251

[42] Levenshtein V 1998 Universal bounds for codes and designs *Handbook of Coding Theory* ed V Pless and C W Huffman (Amsterdam: Elsevier) p 499

[43] König H 1999 Cubature formulas on spheres *Advances in Multivariate Approximation* ed W Haußmann, K Jetter and M Reimer (Berlin: Wiley) p 201

[44] Seymour P D and Zaslavsky T 1984 Averaging sets: a generalization of mean values and spherical designs *Adv. Math.* **52** 213

[45] Dunkl C F 1979 Discrete quadrature and bounds on *t*-designs *Michigan Math. J.* **26** 81

[46] Hardin R H and Sloane N J A 1996 McLaren's improved snub cube and other new spherical designs in three dimensions *Discrete Comput. Geom.* **15** 429

[47] Hoggar S G 1998 64 lines from a quaternionic polytope *Geom. Dedic.* **69** 287

[48] Grassl M 2004 On SIC-POVMs and MUBs in dimension 6 *Proc. ERATO Conf. on Quantum Information Science (Tokyo, Sept. 2004)* p 60

[49] Appleby D M 2005 Symmetric informationally complete-positive operator valued measures and the extended Clifford group *J. Math. Phys.* **46** 052107

[50] Grassl M 2005 Tomography of quantum states in small dimensions *Electron. Notes Discrete Math.* **20** 151

[51] Nikova S 1995 On bounds on the size of designs in complex projective spaces *Proc. Int. Workshop on Optimal Codes and Related Topics (Sozopol, Bulgaria, May 1995)* p 121

[52] Boyvalenkov P and Nikova S 1997 On lower bounds on the size of designs in compact symmetric spaces of rank 1 *Arch. Math.* **68** 81

[53] Boyvalenkov P, Boumova S and Danev D 1999 Necessary conditions for existence of some designs in polynomial metric spaces *Eur. J. Comb.* **20** 213

[54] Welch L R 1974 Lower bounds on the maximum cross correlation of signals *IEEE Trans. Inf. Theory* **20** 397

[55] Caves C M 1999 Quantum error correction and reversible operations *J. Supercond.* **12** 707

[56] Rungta P, Munro W J, Nemoto K, Deuar P, Milburn G J and Caves C M 2000 Qudit entanglement *Directions in Quantum Optics: A Collection of Papers Dedicated to the Memory of Dan Walls* ed H J Carmichael, R J Glauber and M O Scully (Berlin: Springer) p 149

[57] Rungta P, Buzek V, Caves C M, Hillery M and Milburn G J 2001 Universal state inversion and concurrence in arbitrary dimensions *Phys. Rev.* A **64** 042315

[58] Casazza P G, Fickus M, Kovačdević J, Leon M T and Tremain J C 2006 A physical interpretation for finite tight frames *Harmonic Analysis and Applications: In Honor of John J Benedetto* ed C Heil (Boston, MA: Birkhäuser)

[59] Li S D 1995 On general frame decompositions *Numer. Funct. Anal. Optim.* **16** 1181

[60] Ali S T, Antoine J-P and Gazeau J-P 1993 Continuous frames in Hilbert space *Ann. Phys.* **222** 1

[61] Kaiser G 1994 *A Friendly Guide to Wavelets* (Boston, MA: Birkhäuser)

[62] Horn R T, Scott A J, Walgate J, Cleve R, Lvovsky A I and Sanders B C 2005 Classical and quantum fingerprinting with shared randomness and one-sided error *Quantum Inf. Comput.* **5** 258

[63] Rosenberg M 1964 The square-integrability of matrix-valued functions with respect to a non-negative Hermitian measure *Duke Math. J.* **31** 291

[64] Eldar Y C and Forney G D Jr 2002 Optimal tight frames and quantum measurement *IEEE Trans. Inf. Theory* **48** 599

[65] Klappenecker A, Rötteler M, Shparlinski I E and Winterhof A 2005 On approximately symmetric informationally complete positive operator-valued measures and related systems of quantum states *J. Math. Phys.* **46** 082104

[66] Colin S, Corbett J, Durt T and Gross D 2005 About SIC POVMs and discrete Wigner distributions *J. Opt.* B **7** S778

[67] Wootters W K 2006 Quantum measurements and finite geometry *Found. Phys.* **36** 112

[68] Godsil C and Roy A 2005 Equiangular lines, mutually unbiased bases, and spin models *Preprint* quant-ph/0511004

[69] Flammia S T 2006 On SIC-POVMs in prime dimensions *Preprint* quant-ph/0605050

[70] Strohmer T and Heath R W Jr 2003 Grassmannian frames with applications to coding and communication *Appl. Comput. Harmon. Anal.* **14** 257

[71] Alltop W 1980 Complex sequences with low periodic correlations *IEEE Trans. Inf. Theory* **26** 350

[72] Calderbank A R, Cameron P J, Kantor W M and Seidel J J 1997 $\mathbb{Z}_4$-Kerdock codes, orthogonal spreads, and extremal Euclidean line-sets *Proc. Lond. Math. Soc.* **75** 436

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

[74] Conway J H and Sloane N J A 1988 *Sphere Packings, Lattices and Groups* (New York: Springer)

[75] Bures D 1969 An extension of Kakutani's theorem on infinite product measures to the tensor product of semifinite w^*-algebras *Trans. Am. Math. Soc.* **135** 199

[76] Uhlmann A 1976 The 'transition probability' in the state space of a ∗-algebra *Rep. Math. Phys.* **9** 273

[77] Vidal G, Latorre J I, Pascual P and Tarrach R 1999 Optimal minimal measurements of mixed states *Phys. Rev.* A **60** 126

[78] Gisin N and Massar S 1997 Optimal quantum cloning machines *Phys. Rev. Lett.* **79** 2153

[79] Massar S and Popescu S 1995 Optimal extraction of information from finite quantum ensembles *Phys. Rev. Lett.* **74** 1259

[80] Derka R, Bužek V and Ekert A K 1998 Universal algorithm for optimal estimation of quantum states from finite ensembles via realizable generalized measurement *Phys. Rev. Lett.* **80** 1571

[81] Latorre J I, Pascual P and Tarrach R 1998 Minimal optimal generalized quantum measurements *Phys. Rev. Lett.* **81** 1351

[82] Bruß D and Macchiavello C 1999 Optimal state estimation for d-dimensional quantum systems *Phys. Lett.* A **253** 249