Complementarity and state estimation

Thomas Baier and Dénes Petz

Department of Mathematics and its Applications,
Central European University
H-1051 Budapest, Nádor utca 9, Hungary

and

Alfréd Rényi Institute of Mathematics,
H-1364 Budapest, POB 127, Hungary

Abstract

The concept of complementarity (or quasi-orthogonality) is extended to POVM's. It is shown in the setting of unconstrained state estimation that the determinant of the mean quadratic error matrix is minimal if the POVM's are complementary (and informationally complete). Several examples of the scheme are given.

Key words and phrases: quadratic error matrix, POVM, unbiased basis, state estimation.

Introduction

Quantum state estimation is a fundamental problem in the field of quantum information, with applications in quantum cryptography, quantum computing, quantum control, quantum measurement theory and it can be considered as foundational issues in quantum mechanics. Only a finite number of identical copies of the unknown quantum system is available for measurements and one can consider the question of efficient estimation of the state. One side of this problem is the adequate experimental techniques, and the other side is the theory based on the adaptation of statistics to the quantum mechanical formalism. The present paper is about the use of classical statistics to the quantum mechanical formalism. There is a huge literature about the optimal positive operator valued measurements (POVMs) for a general quantum state estimation, see, for example [1, 4]. Optimality has several different conditions and formulations.

Most of the work in state estimation has focused on states of a qubit, pure states [6], or mixed states [4, 15]. The estimation procedure for pure states is simpler, partially due to the smaller number of parameters. The subject of the present paper is state estimation for a n-level quantum system. In this case the boundary of the state space is not the set of pure states but the non-invertible density matrices. The accuracy of the estimation can be quantified by the average of the quadratic error matrix (where the average is with respect to the a priory measure of the possible states). When different estimation schemes are compared, the determinant of mean quadratic error matrix can be used, since matrices are typically not comparable by the positive semidefiniteness. This approach is contained also in the papers [13, 14]. It was obtained there that complementary von Neumann measurements are optimal. The same result was obtained earlier by Wootters and Fields [17], but optimality there had a different formulation. In our approach the measurements are not necessarily von Neumann type. The concept of complementarity (or quasi-orthogonality) is extended to positive operator valued measurements. The main result is a rigorous proof of the fact that quasi-orthogonal POVMs are optimal when the determinant of the mean quadratic error matrix gives efficiency.

1 Preliminaries

A finite dimensional quantum system is described by the algebra $B(\mathcal{H})$ of operators on a finite dimensional Hilbert space $\mathcal{H}$. The algebra $B(\mathcal{H})$ can be equipped with the Hilbert-Schmidt inner product

$$\langle A, B\rangle = \frac{1}{n}\mathrm{Tr}\,(A^* B) \qquad (A, B \in B(\mathcal{H})) \tag{1}$$

and with this inner product it is a Hilbert space itself. If $\dim \mathcal{H} = n$, the dimension of $B(\mathcal{H})$ is n^2 and we can choose an orthonormal basis of the form $\{\sigma_i : 0 \le i \le n^2 - 1\}$ in $B(\mathcal{H})$ that consists of self-adjoint operators and where $\sigma_0 \equiv I$ denotes the identity. We have

$$\langle \sigma_i, \sigma_j\rangle = 0 \quad (i \ne j), \qquad\qquad \langle \sigma_i, \sigma_i\rangle = 1. \tag{2}$$

The subspace $\mathcal{S} := \mathit{span}\{\sigma_i : 1 \le i \le n^2-1\}$ is the (n^2-1)-dimensional linear subspace of traceless operators in $B(\mathcal{H})$. For self-adjoint operators the inner product (1) takes always real values and they form a real vector space over such a basis.

The standard example is the case $\dim \mathcal{H} = 2$. Then the Pauli matrices

$$\sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_2 = \begin{bmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{bmatrix}, \quad \sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \tag{3}$$

together with the identity matrix form a basis of $M_2(\mathbb{C})$.

A density matrix ρ has the expansion

$$\rho = \frac{I}{n} + \sum_{i=1}^{n^2-1} \theta_i \sigma_i \equiv \frac{I}{n} + \boldsymbol{\theta} \cdot \boldsymbol{\sigma}, \tag{4}$$

where $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_{n^2-1}) \in \mathbb{R}^{n^2-1}$ is the generalized Bloch vector. Let us denote the set of possible Bloch vectors by $\mathcal{R}$, $\mathcal{R} := \{\boldsymbol{\theta} \in \mathbb{R}^{n^2-1} : I/n + \boldsymbol{\theta} \cdot \boldsymbol{\sigma} > 0\}$.

A measurement is modeled by a positive operator-valued measure (POVM). This is a map from a set $\{a_1, a_2, \ldots, a_d\}$ of outcomes into positive operators, $a_i \mapsto E_i$ such that $M := \{E_i : 1 \leq i \leq d\}$ is a set of operators satisfying the conditions

$$E_i \geq 0 \qquad \text{and} \qquad \sum_{i=1}^{d} E_i = I. \tag{5}$$

The operators E_i have an expansion

$$E_i = \frac{\operatorname{Tr} E_i}{n} \sigma_0 + \sum_{j=1}^{n^2-1} u_{ij} \sigma_j \equiv \frac{\operatorname{Tr} E_i}{n} \sigma_0 + \mathbf{u}_i \cdot \boldsymbol{\sigma}\,. \tag{6}$$

The vector $\mathbf{u}_i = (u_{i1}, u_{i2}, \ldots, u_{i(n^2-1)}) \in \mathbb{R}^{n^2-1}$ determines the operator E_i up to its trace. The following example contains a simple situation:

Example 1.1 In the case of a qubit, i.e. $\dim \mathcal{H} = 2$, consider the operators

$$\begin{aligned} E_1 &= \tfrac{1}{4} I + \tfrac{1}{4\sqrt{3}} \left(\sigma_1 + \sigma_2 + \sigma_3\right) & \qquad E_2 &= \tfrac{1}{4} I + \tfrac{1}{4\sqrt{3}} \left(\sigma_1 - \sigma_2 - \sigma_3\right) \\ E_3 &= \tfrac{1}{4} I + \tfrac{1}{4\sqrt{3}} \left(-\sigma_1 + \sigma_2 - \sigma_3\right) & \qquad E_4 &= \tfrac{1}{4} I + \tfrac{1}{4\sqrt{3}} \left(-\sigma_1 - \sigma_2 + \sigma_3\right) \end{aligned} \tag{7}$$

where the σ_i $(i = 1, 2, 3)$ are the Pauli matrices given in (3) [15]. This operators fulfill $\sum_{i=1}^{4} E_i = I$. With the anticommutator $\{\sigma_i, \sigma_j\} := \sigma_i \sigma_j + \sigma_j \sigma_i = 0$ $(i \neq j)$ of the Pauli matrices it follows that $E_i^2 = {}^1\!/_2\, E_i$ and therefore the E_i are subnormalized projections, thus positive, and the set $M := \{E_i : 1 \leq i \leq 4\}$ forms a POVM. □

If the actual state of the quantum system is (4), then the probability of outcome a_i in a measurement is given by $\operatorname{Tr}(\rho E_i)$. With the expansions (4) of ρ and (6) of E_i we can express this probabilities in terms of the inner product between the vector $\boldsymbol{\theta}$ and $\mathbf{u}_i$ by

$$p_i \equiv \operatorname{Tr}(\rho E_i) = \frac{\operatorname{Tr} E_i}{n} + n\,(\mathbf{u}_i, \boldsymbol{\theta}) \qquad (1 \leq i \leq d). \tag{8}$$

If all operators in a POVM are projections, it is called von Neumann measurement. In this case there is a self-adjoint operator $A = \sum_{i=1}^{d} a_i P_i$ associated with the measurement and A is called an observable. The a_i are the eigenvalues of A and the P_i are the projections on the corresponding eigenspaces.

Two von Neumann measurements $\{P_i : 1 \le i \le n\}$ and $\{Q_j : 1 \le j \le n\}$ consisting of minimal projections are called complementary if

$$\mathrm{Tr}\,(P_i Q_j) = \frac{1}{n} \qquad (1 \le i, j \le n). \tag{9}$$

The above projections have expansions $P_i = I/n + \mathbf{u}_i \cdot \boldsymbol{\sigma}$ and $Q_j = I/n + \mathbf{v}_j \cdot \boldsymbol{\sigma}$. Condition (9) i is equivalent to condition that the vectors $\mathbf{u}_i$ and $\mathbf{v}_j$ are orthogonal for all i, j. The condition of orthogonality in the traceless subspace $\mathcal{S}$ is called quasi-orthogonality.

A von Neumann measurement corresponds to a basis of the Hilbert space if it consists of minimal projections. If $\{e_i : 1 \le i \le n\}$ and $\{f_j : 1 \le j \le n\}$ are two bases of the Hilbert space $\mathcal{H}$, they are called mutually unbiased if

$$|(e_i, f_j)| = \frac{1}{\sqrt{n}} \qquad (1 \le i, j \le n). \tag{10}$$

Condition (9) on the operators in the von Neumann measurement and condition (10) on the corresponding bases of $\mathcal{H}$ are equivalent, hence the mutually unbiased condition corresponds to complementarity or quasi-orthogonality [10, 12]. From the viewpoint of quasi-orthogonality we can extend complementarity of von Neumann measurements to complementarity of POVMs. We call two POVMs quasi-orthogonal, if the linear subspaces spanned by $M_1 = \{E_i : 1 \le i \le d_1\}$ and $M_2 = \{F_j : 1 \le j \le d_2\}$ are quasi-orthogonal. This can be formulated by the following two equivalent conditions:

$$(i) \qquad span\{E_i : 1 \le i \le d_1\} \ominus \mathbb{C}I \;\perp\; span\{F_j : 1 \le j \le d_2\} \ominus \mathbb{C}I\,,$$
$$(ii) \qquad \mathrm{Tr}\,(E_i F_j) = \mathrm{Tr}\,(E_i)\mathrm{Tr}\,(F_j) \quad (1 \le i \le d_1,\, 1 \le j \le d_2)\,.$$

2 Estimation schemes

Suppose we are given an ensemble of N identical copies of a quantum system in the state ρ, which is not known to us. The state of the ensemble is described by a density operator

$$\rho_{\boldsymbol{\theta}}^{\otimes N} := \rho_{\boldsymbol{\theta}}^1 \otimes \rho_{\boldsymbol{\theta}}^2 \otimes \ldots \otimes \rho_{\boldsymbol{\theta}}^N \in B(\mathcal{H}^{\otimes N})\,. \tag{11}$$

where $\mathcal{H}^{\otimes N} := \otimes_{i=1}^N \mathcal{H}$. In this paper we consider only separate measurements M on the ensemble, i.e. all operators $E_i \in M$ are of the form

$$E_i = E_i^1 \otimes E_i^2 \otimes \ldots \otimes E_i^N \tag{12}$$

This correspond to the situation when each copy of the ensemble is measured separately without any correlations between this measurements.

To estimate the state we apply the following strategy: We choose a set $\mathcal{M} := \{M^{(k)} : 1 \le k \le m\}$ of m different measurements $M^{(k)} = \{E_i^{(k)} : 1 \le i \le d^{(k)}\}$ and we will call

$\mathcal{M}$ a measurement scheme. We require that the operators $\{E_i^{(k)} : 1 \leq i \leq d^{(k)}; 1 \leq k \leq m\}$ linearly span $B(\mathcal{H})$, i.e. $\mathcal{M}$ is informationally complete, and

$$\sum_{k=1}^{m}(d^{(k)} - 1) = n^2 - 1. \tag{13}$$

Furthermore we divide the ensemble into m subensembles of size $N^{(k)}$ $(1 \leq k \leq m)$. On the individual copies in kth subensemble we perform the measurement $M^{(k)}$. From the data obtained in this measurements we construct a classical estimate $\boldsymbol{\nu}$ of the measurement probabilities and from this we obtain an estimate $\hat{\boldsymbol{\theta}}$ of the generalized Bloch vector $\boldsymbol{\theta}$ by linear inversion.

If the true state of the system has the Bloch vector $\boldsymbol{\theta}$, the probability of an individual outcome to appear in a particular measurement is

$$p_i^{(k)} = \operatorname{Tr} \rho E_i^{(k)} = \frac{\operatorname{Tr} E_i^{(k)}}{n} + n\,(\mathbf{u}_i^{(k)}, \boldsymbol{\theta}) \qquad (1 \leq k \leq m, 1 \leq i \leq d^{(k)} - 1) \tag{14}$$

and

$$p_{d^{(k)}}^{(k)} = 1 - \sum_{j=1}^{d^{(k)}-1} p_j^{(k)} \qquad (1 \leq k \leq m). \tag{15}$$

Thus the probabilities $p_i^{(k)}$ and the Bloch vector $\boldsymbol{\theta}$ are related by the $(n^2 - 1)$ linear equations (14). Let us, with the abbreviation $r_i^{(k)} := \operatorname{Tr} E_i^{(k)}$, introduce the $(n^2 - 1) \times (n^2 - 1)$ matrix T and the vectors $\mathbf{p}$, $\mathbf{r}$ as

$$T = \begin{bmatrix} T^{(1)} \\ \vdots \\ T^{(m)} \end{bmatrix}, T^{(k)} = \begin{bmatrix} (\mathbf{u}_1^{(k)})^{\,t} \\ \vdots \\ (\mathbf{u}_{d^{(k)}-1}^{(k)})^{\,t} \end{bmatrix}, \mathbf{p} = \begin{bmatrix} \mathbf{p}^{(1)} \\ \vdots \\ \mathbf{p}^{(m)} \end{bmatrix}, \mathbf{p}^{(k)} = \begin{bmatrix} p_1^{(k)} \\ \vdots \\ p_{d^{(k)}-1}^{(k)} \end{bmatrix}, \tag{16}$$

$$\mathbf{r} = \begin{bmatrix} \mathbf{r}^{(1)} \\ \vdots \\ \mathbf{r}^{(m)} \end{bmatrix}, \quad \mathbf{r}^{(k)} = \begin{bmatrix} r_1^{(k)} \\ \vdots \\ r_{d^{(k)}-1}^{(k)} \end{bmatrix}, \tag{17}$$

where $(\mathbf{u}_i^{(k)})^{\,t}$ denotes the transposed of the vector $\mathbf{u}_i^{(k)}$. Using this notation we can represent the system (14) in a matrix form

$$n \cdot T\boldsymbol{\theta} = \mathbf{p} - \frac{1}{n}\mathbf{r}\,. \tag{18}$$

Since we chose the measurement scheme to be informationally complete and we required (13), the rows $\mathbf{u}_i^{(k)}$ of $T^{(k)}$ form a set of linear independent vectors in the subspace $\mathcal{S}^{(k)} = span\{\mathbf{u}_i^{(k)} : 1 \leq i \leq d^{(k)}\}$ and the matrix T is quadratic and of full rank. Thus we can invert equation (18) to

$$\boldsymbol{\theta} = \frac{1}{n} T^{-1}(\mathbf{p} - \frac{1}{n}\mathbf{r})\,. \tag{19}$$

Given an estimate $\boldsymbol{\nu}$ for the classical probabilities in p from the measured data we obtain an unconstrained estimate $\hat{\boldsymbol{\theta}}$ as

$$\hat{\boldsymbol{\theta}} = \frac{1}{n} T^{-1}(\boldsymbol{\nu} - \frac{1}{n}\mathbf{r})\,. \qquad (20)$$

The terminology "unconstrained estimate" was used already in the papers, [13, 14], the motivation comes from the fact that $\hat{\boldsymbol{\theta}} \in \mathcal{R}$ is not certain. The extension of the present situation to "constrained estimate" will be the subject of a forthcoming paper.

The minimum variance unbiased estimate for the measurement probabilities is given by the relative frequencies of the measurement outcomes (see e.g. [8]):

$$\boldsymbol{\nu} = \begin{bmatrix} \boldsymbol{\nu}^{(1)} \\ \vdots \\ \boldsymbol{\nu}^{(m)} \end{bmatrix}, \quad \boldsymbol{\nu}^{(k)} = \begin{bmatrix} n_1^{(k)}/N^{(k)} \\ \vdots \\ n_{d^{(k)}-1}^{(k)}/N^{(k)} \end{bmatrix}, \qquad (21)$$

where $n_i^{(k)}$ is the number of times an outcome appears among the outcomes of the measurement $M^{(k)}$.

Given the true state $\boldsymbol{\theta}$, we describe the quality of an estimate $\hat{\boldsymbol{\theta}}$ by the mean quadratic error matrix

$$V_{\boldsymbol{\theta}}(\hat{\boldsymbol{\theta}}) = \mathrm{E}_{\boldsymbol{\theta}}[(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})^t]. \qquad (22)$$

The unconstrained estimate (20) as a random vector is a linear transformation of the random vector $\boldsymbol{\nu}$ and in this case the matrix $V_{\boldsymbol{\theta}}(\hat{\boldsymbol{\theta}})$ depends on the mean quadratic error matrix $W_{\boldsymbol{\theta}}(\boldsymbol{\nu}) = \mathrm{E}_{\boldsymbol{\theta}}[(\boldsymbol{\nu} - \mathbf{p})(\boldsymbol{\nu} - \mathbf{p})^t]$ of the classical estimate $\boldsymbol{\nu}$ by

$$V_{\boldsymbol{\theta}}(\hat{\boldsymbol{\theta}}) = \frac{1}{n^2} T^{-1} W_{\boldsymbol{\theta}}(\boldsymbol{\nu})(T^{-1})^t. \qquad (23)$$

Here both $V_{\boldsymbol{\theta}}$ and $W_{\boldsymbol{\theta}}$ depend by (14) as well on the state $\boldsymbol{\theta}$ as on the particular measurement scheme $\mathcal{M}$.

Since the individual measurements are statistically independent, $W_{\boldsymbol{\theta}}$ is a block diagonal matrix where the blocks $W_{\boldsymbol{\theta}}^{(k)}$ are related to the measurements $M^{(k)}$:

$$W_{\boldsymbol{\theta}} = \begin{pmatrix} W_{\boldsymbol{\theta}}^{(1)} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & W_{\boldsymbol{\theta}}^{(m)} \end{pmatrix} \qquad (24)$$

If we use the relative frequencies (21) as an estimate, the matrix elements of $W_{\boldsymbol{\theta}}^{(k)}$ are given by

$$W_{(i,j)}^{(k)} = \frac{1}{N^{(k)}} \left(\delta_{ij} p_i^{(k)} - p_i^{(k)} p_j^{(k)}\right) \qquad (1 \le k \le m\,;\ 1 \le i, j \le d^{(k)} - 1) \qquad (25)$$

In this paper we assume that the unknown states are distributed with respect to a prior probability measure μ on the state space $\mathcal{R}$. The main condition we impose on the measure μ is that it is unitarily invariant:

$$\int_{\mathcal{R}} f(\rho) d\mu(\rho) = \int_{\mathcal{R}} f(U\rho U^*) d\mu\,(\rho) \tag{26}$$

for any integrable functions f on $\mathcal{R}$ and for any unitary $U \in U(n)$. We denote the average with respect to μ by $\langle\,\cdot\,\rangle$. Since unitary conjugation corresponds to an orthogonal transformation of the Bloch vectors $\boldsymbol{\theta}$, an example for such a measure is the normalized Lebesgue measure on $\mathcal{R} \subset \mathbb{R}^{(n^2-1)}$.

To quantify the quality of an estimation scheme we apply the average of the mean quadratic error matrix of the estimate $\hat{\boldsymbol{\theta}}$ taken over the set $\mathcal{R}$ of possible true states with respect to the prior μ:

$$\langle V_{\boldsymbol{\theta}} \rangle = \int_{\mathcal{R}} V_{\boldsymbol{\theta}}\, d\mu(\boldsymbol{\theta}) = \frac{1}{n^2}\, T^{-1} \int_{\mathcal{R}} W_{\boldsymbol{\theta}}\, d\mu(\boldsymbol{\theta})\ (T^{-1})^t\,. \tag{27}$$

Since for two different measurement schemes the average mean quadratic error matrices are not necessarily comparable by positive semi-definiteness, we use, analog to [13],

$$\det\langle V_{\boldsymbol{\theta}} \rangle = \frac{1}{n^2}\, \det\langle W_{\boldsymbol{\theta}} \rangle\, (\det T)^{-2} \tag{28}$$

to compare the quality of different measurement schemes.

3 Optimality of quasi-orthogonal measurements

In this section we analyze the so-called unconstrained estimate and compare the estimation schemes by means of the determinant of the average mean quadratic error matrix.

If we want to compare measurement schemes with regard to quasi-orthogonality of the measurements they use, we need to restrict ourself to certain families within this is meaningful. For instance in the case of a qubit, from the viewpoint of complementarity there is no sense in comparing a measurement scheme using a single POVM, as described in Example 1.1, to a scheme where the three complementary spin observables $S_i = {}^1\!/_2\, \sigma_i$ $(i = 1, 2, 3)$ are measured. Since complementarity of two measurements is a geometric property between the subspaces their measurement operators span, we will compare schemes which differ only in the orientation of this subspaces. For this purpose let us consider unitarily conjugated measurements $M^{(k)}$ and $\hat{M}^{(k)}$,

$$\hat{M}^{(k)} = \{U E_i^{(k)} U^* : E_i^{(k)} \in M^{(k)}\}\,, \tag{29}$$

where $U \in U(n)$. Unitary conjugation of the operators $E_i^{(k)}$ results in an orthogonal transformation the vectors $\mathbf{u}_i^{(k)}$ as defined in (6). Thus unitary conjugation of the measurement $M^{(k)}$ results in a rotation of the subspace $\mathcal{S}^{(k)}$, while the the geometrical configuration of the vectors $\mathbf{u}_i^{(k)}$ and $\hat{\mathbf{u}}_i^{(k)} := U\mathbf{u}_i^{(k)}U^*$ $(1 \leq i \leq d^{(k)})$ remains the same.

We will compare measurement schemes within the following families: Given an informational complete measurement scheme $\mathcal{M} = \{M^{(1)}, M^{(2)} \ldots, M^{(m)}\}$ we define the family $\mathcal{U}(\mathcal{M})$ as the set of all measurement schemes whose elements are unitarily conjugated to the elements of $\mathcal{M}$:

$$\begin{aligned}\mathcal{U}(\mathcal{M}) &= \{\mathcal{M}' = \{U_1 M^{(1)} U_1^*, \ldots, U_m M^{(m)} U_m^*\} : \\ &\qquad \mathcal{M}' \text{ is informationally complete and } U_k \in U(n)\}\end{aligned}$$

This means that the subspaces $\mathcal{S}^{(k)}$ spanned by the individual measurements $M^{(k)}$ are rotated in $\mathcal{S}$.

In this setting we can show the following theorem for the unconstrained estimate (recall that the unconstrained estimate of a state was defined in (20)):

Theorem 3.1 *Let the possible true states of the system be distributed according to a unitarily invariant probability measure μ on $\mathcal{R}$ and let $\mathcal{M}_0 = \{M_0^{(1)}, M_0^{(2)} \ldots, M_0^{(m)}\}$ be a measurement scheme consisting of pairwise quasi-orthogonal measurements $M_0^{(k)}$, $1 \leq k \leq m$. Let $\mathcal{M}_1 = \{U_1 M_0^{(1)} U_1^*, U_2 M_0^{(2)} U_2^* \ldots, U_m M_0^{(m)} U_m^*\}$ be another measurement scheme obtained by unitaries from $\mathcal{M}_0$. Then*

$$\det\langle V^{\mathcal{M}_0}(\hat{\boldsymbol{\theta}})\rangle \leq \det\langle V^{\mathcal{M}_1}(\hat{\boldsymbol{\theta}})\rangle.$$

Therefore $\mathcal{M}_0$ is optimal in the family $\mathcal{U}(\mathcal{M}_0)$ of measurements.

Proof. For the proof we first show that the average mean quadratic error matrix $\langle W_{\boldsymbol{\theta}}\rangle$ of the classical estimate is the same for both $\mathcal{M}_0$ and $\mathcal{M}_1$ and only the matrix T, as defined in (16), depends on the particular choice of the measurement scheme. To see this we notice, that the blocks $W_{\boldsymbol{\theta}}^{(k)}$ are functions of the probabilities $p_i^{(k)}$ of the kth measurement only, and as such they are functions of ρ and $M^{(k)}$:

$$W_{\boldsymbol{\theta}}^{(k)} = W^{(k)}(p_1^{(k)}, p_2^{(k)}, \ldots, p_{d^{(k)}-1}^{(k)}) = W^{(k)}(\rho, M^{(k)}). \tag{30}$$

By our assumption every measurement $M^{(k)} := U_k M_0^{(k)} U_k^*$ in $\mathcal{M}_1$ is unitarily conjugated to some $M_0^{(k)}$. Consequently an element $E_i^{(k)} \in M^{(k)}$ is unitarily conjugated to the element $E_{0,i}^{(k)} \in M_0^{(k)}$. From (8) and the cyclic invariance of the trace we get

$$p_i^{(k)} = \mathrm{Tr}\,(\rho E_i^{(k)}) = \mathrm{Tr}\left(\rho\,(U_k E_{0,i}^{(k)} U_k^*)\right) = \mathrm{Tr}\left((U_k^* \rho U_k)\, E_{0,i}^{(k)}\right). \tag{31}$$

This allows us to move the unitarily conjugation between the arguments of (30) and we get the relation

$$W^{(k)}(\rho, M^{(k)}) = W^{(k)}(\rho, U_k M_0^{(k)} U_k^*) = W^{(k)}(U_k^* \rho U_k, M_0^{(k)}). \tag{32}$$

By our assumption on the measure μ substituting $U_k^*\rho U_k \to \rho$ will not change the value of the integrals of the blocks $W^{(k)}$:

$$\begin{aligned}\langle W_{\boldsymbol{\theta}}^{(k)}\rangle &= \int_{\mathcal{R}} W^{(k)}(\rho, M^{(k)})\, d\mu(\rho) = \int_{\mathcal{R}} W^{(k)}(U_k^*\rho U_k, M_0^{(k)})\, d\mu(\rho) \\ &= \int_{\mathcal{R}} W^{(k)}(\rho, M_0^{(k)})\, d\mu(\rho)\,.\end{aligned}$$

Thus for $\mathcal{M}_0$ and $\mathcal{M}_1$ the average mean quadratic error matrix $\langle W_{\boldsymbol{\theta}}\rangle$ of the estimate $\boldsymbol{\nu}$ of the classical probabilities is identical and we get for the average mean quadratic error matrix of $\hat{\boldsymbol{\theta}}$:

$$\langle V_{\boldsymbol{\theta}}^{\mathcal{M}_i}\rangle = \frac{1}{n^2}\, T_{\mathcal{M}_i}^{-1}\langle W_{\boldsymbol{\theta}}\rangle (T_{\mathcal{M}_i}^{-1})^t = \frac{1}{n^2}\, T_{\mathcal{M}_i}^{-1} \begin{pmatrix} A^{(1)} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & A^{(m)} \end{pmatrix} (T_{\mathcal{M}_i}^{-1})^t \tag{33}$$

for $i = 0, 1$ with some constant matrices $A^{(k)} := \langle W_{\boldsymbol{\theta}}^{(k)}\rangle$.

Thus for all measurement schemes from the family $\mathcal{U}(\mathcal{M}_0)$ the mean quadratic error matrix $\langle V_{\boldsymbol{\theta}}\rangle$ depends on the transformation matrix T only. By

$$\det\langle V_{\boldsymbol{\theta}}\rangle = \frac{1}{n^2}\, \det\langle W_{\boldsymbol{\theta}}\rangle\, (\det T)^{-2}$$

it is enough to show that for all measurement schemes $\mathcal{M} \in \mathcal{U}(\mathcal{M}_0)$ the $\det T$ is maximal for $\mathcal{M}_0$. $\det T$ corresponds to the volume of the parallelepiped $\mathrm{Par}\{\mathbf{u}_i^{(k)} : 1 \le i \le (d^{(k)} - 1),\, 1 \le k \le m\}$ spanned by the row vectors in T. To calculate this volume for a given T we apply the Gram-Schmidt orthogonalization method (without normalization) on its row vectors: We can perform the orthogonalization first in the subspaces $\mathcal{S}^{(k)}$ spanned by the vectors $\{\mathbf{u}_i^{(k)} : 1 \le i \le (d^{(k)} - 1)\}$ independently. We start with $\mathbf{u}_1^{(k)} =: \tilde{\mathbf{u}}_1^{(k)}$ and obtain recursively

$$\tilde{\mathbf{u}}_i^{(k)} = \mathbf{u}_i^{(k)} - P_{i-1}\mathbf{u}_i^{(k)} \tag{34}$$

where P_{i-1} is the orthogonal projection on $span\{\tilde{\mathbf{u}}_1^{(k)}, \tilde{\mathbf{u}}_2^{(k)}, \ldots, \tilde{\mathbf{u}}_{i-1}^{(k)}\}$ in $B(\mathcal{H})$. The volume of the parallelepiped spanned by the resulting vectors remains invariant in this procedure since (34) are elementary row operations on T. If the $\mathcal{S}^{(k)}$ are already orthogonal, orthogonalization in the subspaces $\mathcal{S}^{(k)}$ is sufficient and

$$\begin{aligned}\det T &= \mathrm{Vol}\left(\mathrm{Par}\{\mathbf{u}_i^{(k)} : 1 \le i \le (d^{(k)} - 1),\, 1 \le k \le m\}\right) \\ &= \mathrm{Vol}\left(\mathrm{Par}\{\tilde{\mathbf{u}}_i^{(k)} : 1 \le i \le (d^{(k)} - 1),\, 1 \le k \le m\}\right) \\ &= \prod_{k=1}^{m} \prod_{i=1}^{d^{(k)}-1} \|\, \tilde{\mathbf{u}}_i^{(k)}\|\,.\end{aligned} \tag{35}$$

Otherwise we have to continue the orthogonalization procedure. In this case first note that for measurements $M^{(k)}$ and $\hat{M}^{(k)}$ unitarily conjugated by some U the Gram-Schmidt procedure gives a geometrically equal result in the subspaces $\mathcal{S}^{(k)}$ and $\hat{\mathcal{S}}^{(k)}$. Orthogonalization and unitary conjugation can be interchanged and the following diagram commutes:

$$\begin{array}{llcl} M^{(k)}: & \{\mathbf{u}_1^{(k)}, \mathbf{u}_2^{(k)}, \ldots, \mathbf{u}_{d^{(k)}-1}^{(k)}\} & \xrightarrow{G-S} & \{\tilde{\mathbf{u}}_1^{(k)}, \tilde{\mathbf{u}}_2^{(k)}, \ldots, \tilde{\mathbf{u}}_{d^{(k)}-1}^{(k)}\} \\ & \quad\quad {\scriptstyle U}\Big\downarrow & & \quad\quad {\scriptstyle U}\Big\downarrow \\ \hat{M}^{(k)}: & U\{\mathbf{u}_1^{(k)}, \mathbf{u}_2^{(k)}, \ldots, \mathbf{u}_{d^{(k)}-1}^{(k)}\}\,U^* & \xrightarrow{G-S} & U\{\tilde{\mathbf{u}}_1^{(k)}, \tilde{\mathbf{u}}_2^{(k)}, \ldots, \tilde{\mathbf{u}}_{d^{(k)}-1}^{(k)}\}\,U^* \end{array} \tag{36}$$

In particular, since unitary conjugation by U results in a orthogonal rotation of the Bloch vectors $\mathbf{u}_i^{(k)}$, we have $\|\,\tilde{\mathbf{u}}_i^{(k)}\| = \|\,U\tilde{\mathbf{u}}_i^{(k)}U^*\|$.

If we need to continue the orthogonalization procedure, this will decrease the length of the vectors $\tilde{\mathbf{u}}_i^{(k)}$, and by (35) this results in a smaller volume of $\mathrm{Par}\{\mathbf{u}_i^{(k)} : 1 \le i \le (d^{(k)}-1),\, 1 \le k \le m\}$. Thus T has maximal determinant, if (and actually only if) the $\mathcal{S}^{(k)}$ are orthogonal. □

4 Examples

Let us discuss the result by means of examples where the existence of quasi-orthogonal measurements is known.

Example 4.1 If a von Neumann measurement consists of minimal projections it corresponds to an observable A with a non degenerate spectrum of eigenvalues. Complementary of such observables is equivalent to mutual unbiasedness of their eigenbases. Unitary conjugation of the measurement of an observables corresponds to a change of their eigenbasis. If the measurement of $n+1$ observables with non-degenerate spectrum is used for the estimation, by the above result it is optimal to choose them complementary. It is known in the cases where n is a prime power that $m = n+1$ complementary observables with non-degenerate spectrum exist [17]. An experimental realization of such a measurement in the case of $\dim \mathcal{H} = 4$ can be found in [1]. □

Example 4.2 Let us consider a system of m qubits described by $B(\mathcal{H}) \simeq M_2^1(\mathbb{C}) \otimes M_2^2(\mathbb{C}) \otimes \cdots \otimes M_2^m(\mathbb{C})$ with $n = \dim \mathcal{H} = 2^m$. We can construct an orthonormal basis of $M_{2^m}(\mathbb{C})$ from the Pauli matrices given in (3) by

$$\sigma^{(k)} = \sigma_{i_1} \otimes \sigma_{i_2} \otimes \cdots \otimes \sigma_{i_m} \qquad \text{where} \quad (i_j = 0, 1, 2, 3) \tag{37}$$

and $1 \le k \le n$. The operators

$$P_\pm^{(k)} := \frac{1}{2}(I \pm \sigma^{(k)}) \qquad (1 \le k \le n^2 - 1) \tag{38}$$

are projections of rank $n/2$ in $B(\mathcal{H})$. They define measurements $M^{(k)} = \{P_+^{(k)}, P_-^{(k)}\}$ $(1 \le k \le n^2 - 1)$ on the qubits which correspond to observables in $B(\mathcal{H})$ with two distinct eigenvalues of multiplicity $n/2$. By the above result, among measurement schemes using $n^2 - 1$ such observables, the above is optimal. The $M^{(k)}$ are separate POVMs measuring spin observables $S_i = {}^1\!/_2\,\sigma_i$ $(i = 1, 2, 3)$ on the individual qubits. While the measurement scheme defined by the $M^{(k)}$ is not optimal among schemes using separate measurements [5], it is of importance in state estimation for its simple experimental feasibility [7]. □

Example 4.3 A composite system of k different r-level systems on a Hilbert space $\mathcal{H}$, $\dim \mathcal{H} = r$, is described by the algebra

$$B(\mathcal{H}^{\otimes k}) \simeq M_{r^k}(\mathbb{C}) \simeq \bigotimes_{i=1}^{k} M_r(\mathbb{C})$$

If r is assumed to be a prime power, there exist $r + 1$ complementary observables A_i $(1 \le i \le r + 1)$ on $B(\mathcal{H})$. If the true state of the composite system is ρ, the measurement of one of this observables on one of the subsystems, described by the subalgebra $\mathcal{A}_1 \simeq B(\mathcal{H})$, gives information on the reduced density $\rho_1 = \mathrm{Tr}_{(1)}\rho$, where $\mathrm{Tr}_{(1)}$ denotes the partial trace onto $\mathcal{A}_1$. In order to get more information, we can change the density ρ by an interaction. For a Hamiltonian H, the new state after the interaction is

$$e^{-iHt/\hbar}\rho e^{iHt/\hbar} = W\rho W^*, \tag{39}$$

Performing the measurement of an observable A_i $(1 \le i \le r + 1)$ on $\mathcal{A}_1$ after the interaction gives information about the new reduced density $\rho_2 = \mathrm{Tr}_{(1)}W\rho W^*$ and this is equivalent to the measurement of the observable W^*A_iW contained in a subalgebra $\mathcal{A}_2 = W^*\mathcal{A}_1W$. By using further interactions, we can construct an information complete measurement scheme equivalent to the measurement of observables $W_j^*A_iW_j \in \mathcal{A}_j$ $(1 \le i \le r + 1,\ 1 \le j \le (r^{2k} - 1)/(r^2 - 1))$. By the above theorem it is optimal if we choose the interactions such that the subalgebras $\mathcal{A}_j$ $(1 \le j \le (r^{2k} - 1)/(r^2 - 1))$ are quasi-orthogonal. Let us remark that this remains true if we use an arbitrary informationally complete measurement scheme on the subsystem $B(\mathcal{H})$ instead of the measurements of the complementary observables A_i $(1 \le i \le r + 1)$. For the case when r is a prime power it was shown in [9] that $(r^{2k} - 1)/(r^2 - 1)$ quasi-orthogonal subalgebras exist.

5 Discussion

We studied the role of complementary (or quasi-orthogonality) for state estimation schemes using separate POVMs on a finite number of identical states of a quantum system. It was assumed that the true states are distributed according to a unitarily invariant prior distribution on the state space, The main result showed that if a measurement scheme $\mathcal{M}_0 = \{M_0^{(1)}, M_0^{(2)}, \dots, M_0^{(m)}\}$ consists of complementary measurements, it performs better than any measurement scheme of the form $\mathcal{M}' = \{U_1 M^{(1)} U_1^*, U_2 M^{(2)} U_2^*, \dots, U_m M^{(m)} U_m^*\}$ obtained with unitaries $U_k \in U(n)$ $(1 \le k \le m)$. As a figure of merit the determinant of the average of the mean quadratic error matrix was used.

In [17] a similar result was obtained for von Neumann measurements in an asymptotic setting. In [17] the information gain in the estimation procedure was considered. This quantity is asymptotically related to the determinant of the mean quadratic error matrix of the estimate.

If the dimension of the Hilbert space is n, then the existence of informationally complete von Neumann measurements

$$(E_1^{(k)}, E_2^{(k)}, \dots, E_n^{(k)}) \quad (1 \le k \le n+1, \quad E_i^{(k)} \text{ are rank 1 projections})$$

is not known, this is another formulation of the MUB problem [12]. However, if we do not insist on projections of rank one, but arbitrary POVMs are allowed, then the existence holds for arbitrary dimension n.

References

[1] R.B.A. Adamson, A.M. Steinberg, Improving quantum state estimation with mutually unbiased bases, arXiv:0808.0944v3 (quant-ph).

[2] T. Baier, *A state estimation scheme for finite quantum systems*, PhD thesis, Central Europian University, Budapest, 2009.

[3] T. Baier, D. Petz, K. M. Hangos and A. Magyar, Comparison of some methods of quantum state estimation, *Quantum Probability and Infinite Dimensional Analysis, Proceedings of the 26th Conference*, eds. L. Accardi, W. Freudenberg, M. Schürmann, pp. 64–78, World Scientific, 2007.

[4] E. Bagan, M.A. Ballester, R.D. Gill, A. Monras and R. Munoz-Tapia, Optimal full estimation of qubit mixed states, Phys. Rev. A **73**, 032301, 2006.

[5] M. D. de Burgh, N. K. Langford, A. C. Doherty and A. Gilchrist, Choice of Measurement Sets in Qubit Tomography, arXiv:0706.3756 (quant-ph).

[6] R. Gill and S. Massar, State estimation for large ensembles, Phys.Rev. **A61**, 042312, 2002.

[7] D. F. V. James, P. G. Kwiat, W. J. Munro, and A. G. White, Measurement of qubits Phys. Rev. **A64**, 052312, 2001.

[8] E. L. Lehmann and G. Casella, *Theory of point estimation*, Springer-Verlag, New York, second edition, 1998.

[9] H. Ohno, Quasi-orthogonal subalgebras of matrix algebras, Linear Alg. Appl. **429**(2008), 2146-2158.

[10] D. Petz, Complementarity in quantum systems, Rep. Math. Phys. **59**(2007), 209–224.

[11] D. Petz, *Quantum Information Theory and Quantum Statistics*, Springer-Verlag, Heidelberg, 2008.

[12] D. Petz, Complementarity and the algebraic structure of finite quantum systems, J. of Physics: Conference Series **143**(2009), 012011.

[13] D. Petz, K.M. Hangos and A. Magyar, Point estimation of states of finite quantum systems, J. Phys. A: Math. Theor. **40**(2007), 7955–7969.

[14] D. Petz, K.M. Hangos and L. Ruppert, Quantum state tomography with finite sample size, in *Quantum Bio-Informatics*, eds. L. Accardi, W. Freudenberg, M. Ohya, World Scientific, 2008, pp. 247–257.

[15] J. Rehácek, B. Englert and D. Kaszlikowski, Minimal qubit tomography, Physical Review A **70**, 052321, 2004.

[16] K.G.H. Vollbrecht and R.F. Werner, Why two qubits are special, J. Math. Phys. **41** 6772–6782, 2000.

[17] W.K. Wootters and B.D. Fields, Optimal state determination by mutually unbiased measurements, Ann. Physics, **191**, 363–381, 1989.