IOPscience iopscience.iop.org

Home Search Collections Journals About Contact us My IOPscience

Classical randomness in quantum measurements

This article has been downloaded from IOPscience. Please scroll down to see the full text article.

2005 J. Phys. A: Math. Gen. 38 5979

(http://iopscience.iop.org/0305-4470/38/26/010)

View the table of contents for this issue, or go to the journal homepage for more

Download details:
IP Address: 171.66.16.92
The article was downloaded on 03/06/2010 at 03:49

Please note that terms and conditions apply.

INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND GENERAL

J. Phys. A: Math. Gen. **38** (2005) 5979–5991 doi:10.1088/0305-4470/38/26/010

Classical randomness in quantum measurements

Giacomo Mauro D'Ariano[1,2], **Paoloplacido Lo Presti**[1] **and Paolo Perinotti**[1,3]

[1] QUIT—Quantum Information Theory Group, Dipartimento di Fisica 'A Volta', via Bassi 6, I-27100 Pavia, Italy
[2] Center for Photonic Communication and Computing, Department of Electrical and Computer Engineering, Northwestern University, Evanston, IL 60208, USA
[3] Istituto Nazionale di Fisica della Materia Unità di Pavia, Pavia, Italy

E-mail: dariano@unipv.it, lopresti@unipv.it and perinotti@fisicavolta.unipv.it

Received 11 April 2005, in final form 11 May 2005
Published 15 June 2005
Online at stacks.iop.org/JPhysA/38/5979

Abstract
Similarly to quantum states, also quantum measurements can be 'mixed', corresponding to a random choice within an ensemble of measuring apparatuses. Such mixing is equivalent to a sort of hidden variable, which produces a noise of purely classical nature. It is then natural to ask which apparatuses are *indecomposable*, i.e. do not correspond to any random choice of apparatuses. This problem is interesting not only for foundations, but also for applications, since most optimization strategies give optimal apparatuses that are indecomposable. Mathematically the problem is posed describing each measuring apparatus by a positive operator-valued measure (POVM), which gives the statistics of the outcomes for any input state. The POVMs form a convex set, and in this language the indecomposable apparatuses are represented by extremal points—the analogous of 'pure states' in the convex set of states. Differently from the case of states, however, indecomposable POVMs are not necessarily rank-one, e.g. von Neumann measurements. In this paper we give a complete classification of indecomposable apparatuses (for discrete spectrum), by providing different necessary and sufficient conditions for extremality of POVMs, along with a simple general algorithm for the decomposition of a POVM into extremals. As an interesting application, 'informationally complete' measurements are analysed in this respect. The convex set of POVMs is fully characterized by determining its border in terms of simple algebraic properties of the corresponding POVMs.

PACS numbers: 03.65.Ta, 03.67.−a

 © 2005 IOP Publishing Ltd Printed in the UK

1. Introduction

Measurements are the essence of any experimental science. At extreme sensitivities and precisions they become the true core of quantum mechanics. For any practical need, a measurement can always be regarded as the retrieval of information about the 'state' of the measured system. However, due to the no-cloning theorem [1, 2],[4] even for an elementary system—e.g. a single harmonic oscillator or a spin—it is impossible to recover a complete knowledge of the state of the system from a single measurement [3] without prior knowledge. Then, since in quantum mechanics different incompatible measurements can be performed in principle, one is faced with the problem of which measurement should be adopted for accomplishing a specific task, and which strategy of repeated measurements would be the most statistically efficient. These are the basic issues of the operational viewpoint of quantum estimation theory [4].

A measurement on a quantum system [5] returns a random result e from a set of possible outcomes $\mathsf{E} = \{e = 1, \ldots, N\}$, with probability distribution $p(e|\rho)$ depending on the state ρ of the system in a way which is distinctive of the measuring apparatus according to the Born rule

$$p(e|\rho) = \mathrm{Tr}[\rho P_e]. \quad (1)$$

In equation (1) P_e denote positive operators on the Hilbert space H of the system, representing our knowledge of the measuring apparatus from which we infer information on the state ρ from the probability distribution $p(e|\rho)$. Positivity of P_e is needed for positivity of $p(e|\rho)$, whereas normalization is guaranteed by the completeness $\sum_{e\in\mathsf{E}} P_e = I$. In the present paper we will only consider the simple case of the finite discrete set E. More generally, one has an infinite probability space E (generally continuous), and in this context the set of positive operators $\{P_e\}$ becomes actually a positive operator valued measure (POVM). Every apparatus is described by a POVM, and, conversely, every POVM can be realized in principle by an apparatus[5] [5–10].

The linearity of the Born rule (1) in both arguments ρ and P_e is consistent with the intrinsically statistical nature of the measurement, in which our partial knowledge of both the system and the apparatus reflects in 'convex' structures for both states and POVMs. This means that not only states, but also POVMs can be 'mixed', namely there are POVMs that give probability distributions $p(e|\rho)$ that are equivalent to randomly choosing among different apparatuses. Note that mixed POVMs can also correspond to a single measuring apparatus, not only when the apparatus itself is prepared in a mixed state, but also for pure preparation, as a result of discarding (tracing out) the apparatus after a unitary interaction with the system. Clearly, such mixing is itself a source of 'classical' noise, which can be in principle removed by adopting an indecomposable apparatus in the ensemble corresponding to the mixed POVM. It is then natural to ask which apparatuses are *indecomposable*, i.e. 'pure' in the above sense, or, mathematically, which POVMs correspond to extremal points of the convex set. The classification of such apparatuses is certainly very useful in applications, since most

[4] In [1] it is shown that the cloning machine violates the superposition principle, which applies to a minimum total number of *three* states, and hence does not rule out the possibility of cloning *two* nonorthogonal states. It is violation of unitarity that makes cloning any *two* nonorthogonal states impossible [2].

[5] The first proof of realizability of POVMs in terms of usual projection-valued measurements (PVM) is the Naimark extension theorem [6], which, however, remains at an abstract level. The problem of actual realizability of POVMs cannot be disconnected from that of realizing *instruments*, which provide a more thorough description of the apparatus in terms also of the state-transformation that it affects. In this sense, the most general realization theorem is due to Ozawa [7]. The history of this subject is also complicated by many other aspects of the problem, and for a selection of literature on modern measurement theory we suggest the reviews [5, 8–10].

optimization strategies in quantum estimation theory [4, 5] correspond to minimize a concave (actually linear) function on the POVMs convex set—the so-called 'cost-function'—whence leading to an optimal POVM which is extremal.

Surprisingly, extremality for POVMs generally does not mean being rank-one, as for 'pure' states. In other words, indecomposable POVMs are not necessarily realized by von Neumann measurements. Indeed, as we will see in this paper, there are rank-one POVMs that are not extremal, whereas, on the opposite side, there are higher-rank POVMs which are extremal. Moreover, whenever the optimization problems have additional linear constraints—e.g. for covariant POVMs, or for fixed probability distribution on a given state—the corresponding subset of POVMs is a lower dimensional convex set corresponding to a section by a hyperplane of the complete POVMs convex set, with boundary equal to the section of the original boundary, and whence with extremal points that belong to the boundary of the convex set of all POVMs. For this reason, also the boundary of the POVMs convex set is interesting in practice, since POVMs that are optimal (for a concave cost function) with an additional linear constraint generally are non-extremal, but still belong to the boundary. One can also argue that POVMs which lie inside faces of the convex, physically exhibit a different degree of 'classical' noise in relation with the dimensionality of the face. Note that one should not imagine the POVMs set as a polytope, since, in contrast the set is 'strongly convex', namely the extremal points are not isolated, but lie on continuous manifolds.

In the present paper we address the problem of apparatus decomposability along three lines of attack: (a) by providing simple necessary and sufficient conditions for extremality of POVMs; (b) by establishing the complete structure of the POVMs convex set via the characterization of its border in terms of algebraic properties of the POVMs; (c) by providing a simple general algorithm for the decomposition of POVMs into extremals. For simplicity, the whole paper is restricted to the case of discrete spectrum. In section 2, after clarifying the general features of convex combinations of POVMs, using the method of perturbations we derive three different if-and-only-if conditions for extremality, along with some corollaries giving easy useful conditions, only necessary or sufficient, that will be used in the following. Section 3 exemplifies the results of section 2 in the case of a single qubit. Section 4 presents the characterization of the border of the convex set in terms of algebraic properties of POVMs. Section 5 shows that for every dimension $d = \dim(\mathsf{H})$ there is always an extremal POVM with maximal number N of elements $N = d^2$, corresponding to a so-called informationally complete POVM [11]. After summarizing the results in the concluding section 6, the appendix reports the algorithm for decomposing a POVM into extremals.

2. Convexity and extremality of POVMs

Let us denote by $\mathscr{P}_N$ the convex set of POVMs on a finite-dimensional Hilbert space H, with a number N of outcomes $\mathsf{E} = \{1, \ldots, N\}$. We will represent a POVM in the set as the vector $\mathbf{P} \in \mathscr{P}_N$, $\mathbf{P} = \{P_1, \ldots, P_N\}$ of the N positive operators P_e. The fact that the set $\mathscr{P}_N$ is convex means that it is closed under convex linear combinations, namely for any $\mathbf{P}', \mathbf{P}'' \in \mathscr{P}_N$ also $\mathbf{P} = p\mathbf{P}' + (1 - p)\mathbf{P}'' \in \mathscr{P}_N$ with $0 \leqslant p \leqslant 1$—i.e. p is a probability. Then, $\mathbf{P}$ can be also equivalently achieved by randomly choosing between two different apparatuses corresponding to $\mathbf{P}'$ and $\mathbf{P}''$, respectively, with probability p and $1-p$, since, the overall statistical distribution $p(e|\rho)$ will be the convex combination of the statistics coming from $\mathbf{P}'$ and $\mathbf{P}''$. Note that $\mathscr{P}_N$ contains also the set of POVMs with a strictly smaller number of outcomes, i.e. with $\mathsf{E}' \subset \mathsf{E}$. For such POVMs the elements corresponding to outcomes in $\mathsf{E} \setminus \mathsf{E}'$ will be zero, corresponding to zero probability of occurrence for all states. Clearly, for $N \leqslant M$ one has $\mathscr{P}_N \subseteq \mathscr{P}_M \subseteq \mathscr{P}$, where by $\mathscr{P}$ we denote the convex set of all POVMs with any (generally infinite) discrete spectrum. The extremal points of $\mathscr{P}_N$ represent the 'indecomposable' measurements, which cannot be achieved by mixing different measurements. Obviously, a POVM which is extremal for $\mathscr{P}_N$ is also extremal for $\mathscr{P}_M$ with $M \geqslant N$, whence it is actually extremal for $\mathscr{P}$, and we will simply name it *extremal* without further specifications.

Let us start with a simple example. Consider the following two-outcome POVM for a qubit

$$\mathbf{P} = \left(\tfrac{1}{2}|0\rangle\langle 0|, \tfrac{1}{2}|0\rangle\langle 0| + |1\rangle\langle 1|\right). \tag{2}$$

By defining $\mathbf{D} = \frac{1}{2}(|0\rangle\langle 0|, -|0\rangle\langle 0|)$, the two vectors $\mathbf{P}_\pm = \mathbf{P} \pm \mathbf{D}$ correspond to the following different POVMs,

$$\mathbf{P}_+ = (|0\rangle\langle 0|, |1\rangle\langle 1|), \qquad \mathbf{P}_- = (0, I), \tag{3}$$

which, intuitively are extremal, whereas $\mathbf{P}$ is not, since $\mathbf{P} = \frac{1}{2}\mathbf{P}_+ + \frac{1}{2}\mathbf{P}_-$ (we will see in the following that $\mathbf{P}_\pm$ are indeed extremal).

Now the problem is how to assess when a POVM is extremal. Looking at the above example, one notes that the non-extremality of $\mathbf{P}$ is equivalent to the existence of a vector of operators $\mathbf{D} \neq \mathbf{0}$ such that $\mathbf{P}_\pm = \mathbf{P} \pm \mathbf{D}$ are POVMs, because in this case $\mathbf{P}$ can be written as a convex combination of $\mathbf{P}_+$ and $\mathbf{P}_-$. The non-existence of such vector of operators $\mathbf{D} \neq \mathbf{0}$ is also a necessary condition for extremality of $\mathbf{P}$, since for a non-extremal $\mathbf{P}$ there exist two POVMs $\mathbf{P}_1 \neq \mathbf{P}_2$ such that $\mathbf{P} = \frac{1}{2}\mathbf{P}_1 + \frac{1}{2}\mathbf{P}_2$, whence $\mathbf{D} = \mathbf{P}_1 - \mathbf{P}_2 \neq \mathbf{0}$.

This leads to the method of perturbations for establishing extremality of a point in a convex set, which in the present context will be the following.

Method of perturbations. We call a nonvanishing $\mathbf{D}$ a *perturbation* for the POVM $\mathbf{P}$ if there exists an $\epsilon > 0$ such that $\mathbf{P} \pm \epsilon\mathbf{D}$ are both POVMs. Then a POVM is extremal if and only if it does not admit perturbations.

From the definition it follows that perturbations for POVMs are represented by vectors $\mathbf{D}$ of Hermitian operators D_e (for positivity of $\mathbf{P} \pm \epsilon\mathbf{D}$) and with zero sum $\sum_{e\in\mathsf{E}} D_e = 0$ (for normalization of $\mathbf{P} \pm \epsilon\mathbf{D}$). Specifically, $\mathbf{D}$ is a perturbation for $\mathbf{P}$ if for some $\epsilon > 0$ one has

$$P_e \pm \epsilon D_e \geqslant 0 \qquad \forall e \in \mathsf{E}.$$

Note that the condition $P_e \pm \epsilon D_e \geqslant 0$ is equivalent to $\epsilon|D_e| \leqslant P_e$, and a necessary condition for $\epsilon|D_e| \leqslant P_e$ is that $\mathsf{Ker}(P_e) \subseteq \mathsf{Ker}(D_e)$, or equivalently $\mathsf{Supp}(D_e) \subseteq \mathsf{Supp}(P_e)$ (the support $\mathsf{Supp}(X)$ of an operator X is defined as the orthogonal complement of its kernel $\mathsf{Ker}(X)$). In fact $\mathsf{Ker}(D_e) = \mathsf{Ker}(|D_e|)$, and for a vector $|\psi\rangle \in \mathsf{Ker}(P_e)$ with $|\psi\rangle \notin \mathsf{Ker}(|D_e|)$, one would have

$$\langle\psi|(P_e - \epsilon|D_e|)|\psi\rangle = -\epsilon\langle\psi||D_e||\psi\rangle \leqslant 0,$$

contradicting the hypothesis. Clearly, the Hermitian operators D_e can be taken simply as linearly dependent—instead of having zero sum—i.e. $\sum_e \lambda_e D_e = 0$ for non-vanishing λ_e, and, moreover, one can consider more generally complex operators D_e with

$$\mathsf{Supp}(D_e) \cup \mathsf{Rng}(D_e) \subseteq \mathsf{Supp}(P_e) \qquad \forall e \in \mathsf{E},$$

and satisfying $\epsilon|D_e| \leqslant P_e$. In fact $P_e \pm \epsilon D'_e \geqslant 0$ is satisfied $\forall e \in \mathsf{E}$ by the set of Hermitian operators

$$D'_e = \lambda_e D_e + \lambda_e^* D_e^\dagger,$$

for which $\sum_e D'_e = 0$.

The above considerations show that what really matters in assessing the extremality of the POVM $\mathbf{P} = \{P_e\}$ is just a condition on the supports $\mathsf{Supp}(P_e)$, corresponding to the following theorem.

Theorem 1. *The extremality of the POVM* $\mathbf{P}$ *is equivalent to the nonexistence of non-trivial solutions* $\mathbf{D}$ *of the equation*

$$\sum_{e\in\mathsf{E}} D_e = 0, \qquad \mathsf{Supp}(D_e) \cup \mathsf{Rng}(D_e) \subseteq \mathsf{Supp}(P_e) \qquad \forall e \in \mathsf{E}. \tag{4}$$

The above condition can be made explicit on the P_e eigenvectors $\{|v_n^{(e)}\rangle\}$ corresponding to a nonzero eigenvalue, which therefore span $\mathsf{Supp}(P_e)$. Then, equation (4) becomes the linear homogeneous system of equations in the variables $D_{nm}^{(e)} = \langle v_n^{(e)}|D_e|v_m^{(e)}\rangle$

$$\sum_{e\in\mathsf{E}} \sum_{nm=1}^{\mathrm{rank}(P_e)} D_{nm}^{(e)} |v_n^{(e)}\rangle\langle v_m^{(e)}| = 0 \iff D_{nm}^{(e)} = 0 \qquad \forall n, m, \quad \forall e \in \mathsf{E}, \tag{5}$$

namely the following version of theorem 1.

Theorem 2. *A POVM* $\mathbf{P} = \{P_e\}_{e\in\mathsf{E}}$ *is extremal iff the operators* $|v_n^{(e)}\rangle\langle v_m^{(e)}|$ *made with the eigenvectors of* P_e *are linearly independent* $\forall e \in \mathsf{E}$ *and* $\forall n, m = 1, \ldots, \mathrm{rank}(P_e)$.

Theorem 2 is the characterization of extremal POVMs given by Parthasaraty in [12] in a C^*-algebraic setting. Note that instead of the eigenvectors $\{|v_n^{(e)}\rangle\}$ one can more generally consider a non-orthonormal basis, which can be useful in numerical algorithms.

Another interesting way to state theorem 1 is in terms of *weak independence* of orthogonal projectors Z_e on $\mathsf{Supp}(P_e)$, where we call a generic set of orthogonal projections $\{Z_e\}_{e\in\mathsf{E}}$ weakly independent [13] if for any set of operators $\{T_e\}_{e\in\mathsf{E}}$ on H one has

$$\sum_{e\in\mathsf{E}} Z_e T_e Z_e = 0 \quad \Rightarrow \quad Z_e T_e Z_e = 0, \qquad \forall e \in \mathsf{E}. \tag{6}$$

Note that since the orthogonal projectors are in one-to-one correspondence with their supporting spaces, the notion of weak independence can equivalently be attached to the supporting spaces $\mathsf{Supp}(P_e)$. This leads us to the alternative characterization of extremal POVMs.

Theorem 3. *A POVM* $\mathbf{P} = \{P_e\}_{e\in\mathsf{E}}$ *is extremal iff the supports* $\mathsf{Supp}(P_e)$ *are weakly independent for all* $e \in \mathsf{E}$.

The proof is straightforward if one considers that any perturbation $\mathbf{D}$ for $\mathbf{P}$ can be written as $D_e = Z_e D_e Z_e$ with the constraint

$$\sum_e Z_e D_e Z_e = 0. \tag{7}$$

Then the theorem simply says that the only allowed perturbation for an extremal POVM is the trivial one.

Some corollaries relevant for applications follow immediately from the main theorem 1 or its equivalent versions.

Corollary 1. *For* $\sum_{e\in\mathsf{E}} \dim[\mathsf{Supp}(P_e)]^2 > d^2$ *the POVM* $\mathbf{P}$ *is not extremal.*

This means that a POVM with more than d^2 nonvanishing elements is always decomposable into POVMs with less than d^2 elements. For the case of d^2 elements theorem 1 also implies that

Corollary 2. *An extremal POVM with* d^2 *outcomes must be necessarily rank-one.*

In fact, for larger rank we would have more than d^2 eigenvectors $\left|v_n^{(e)}\right\rangle$, and the operators $\left|v_n^{(e)}\right\rangle\left\langle v_m^{(e)}\right|$ cannot be linearly independent. From theorem 1 it also follows that if some elements of the POVM $\mathbf{P}$ have non-disjoint supports, then $\mathbf{P}$ is not extremal, or, equivalently

Corollary 3. *An extremal POVM* $\mathbf{P} = \{P_e\}$ *must have all supports* $\mathsf{Supp}(P_e)$ *mutually disjoint.*

Precisely, we call two linear spaces A and B *disjoint* when $\mathsf{A} \cap \mathsf{B} = \{0\}$ is the null vector. It is worth noting that two linear spaces that are disjoint are not necessarily orthogonal, whereas, reversely, two orthogonal spaces are clearly disjoint. We emphasize that the condition of corollary 3 is only necessary. Indeed, it is easy to envisage a POVM that satisfies the above condition without being extremal, e.g. a rank-one POVM for $d = 2$ with five elements corresponding to the vertices of a pentagon in the Bloch sphere (see section 3 for extremal POVMs for qubits).

Other obvious consequences of theorem 2 are the following.

Corollary 4. *Orthogonal POVMs are extremal.*

Corollary 5. *A rank-one POVM is extremal if and only if its elements* P_e *are linearly independent.*

Corollary 1 states that for dimension d an extremal POVM can have at most d^2 non-null elements (in section 5 we will show that such an extremal POVM always exists). Here we can immediately conclude that

Corollary 6. *A POVM with* d^2 *elements is necessarily a rank-one 'informationally complete' POVM.*

By definition, an informationally complete POVM $\mathbf{P} = \{P_e\}$ [11] has elements P_e which span the space of all operators on H, thus allowing the estimation of any ensemble average using the same fixed apparatus. The fact that a POVM with d^2 elements necessarily is rank-one is stated in corollary 2, whereas corollary 5 assures that all POVMs elements are linearly independent, whence the set of d^2 linearly independent operators P_e is obviously complete for dimension d. In section 5 we will give an explicit example of such an extremal informationally complete POVM.

For rank greater than one we have only the necessary condition

Corollary 7. *If the POVM is extremal, then its non-vanishing elements are necessarily linearly independent.*

In fact, according to theorem 2 the projectors on the eigenvectors must be linearly independent, whence also the operators P_e. Indeed, for linearly dependent elements there exist coefficients λ_e not all vanishing such that $\sum_{e\in\mathsf{E}} \lambda_e P_e = 0$, and without loss of generality we can take $-1 \leqslant \lambda_e \leqslant 1$. Therefore, the POVM can be written as convex combination $\mathbf{P} = \frac{1}{2}\mathbf{P}^- + \frac{1}{2}\mathbf{P}^+$, with $P_e^{\pm} = (1 \pm \lambda_e)P_e$, and $\mathbf{P}^- \neq \mathbf{P}^+$ (since the λ_e are not all vanishing).

In the appendix we report an algorithm for decomposing a given POVM into extremal.

3. Extremal POVMs for qubits

Using the above results we will give a classification of extremal POVMs for qubits. In this case, corollary 1 implies that the extremal POVMs cannot have more than four elements, and that, apart from the trivial POVM $\mathbf{P} = I$, they must be made of rank-one projectors (otherwise

$\mathsf{Supp}(P_e)$ for different $e \in \mathsf{E}$ would not be mutually disjoint). Now, upon writing the POVM elements in the Bloch form

$$P_e = \alpha_e\,(I + \vec{n}_e \cdot \vec{\sigma})\,, \tag{8}$$

the constraints for normalization and positivity read

$$\alpha_e > 0, \qquad \sum_e \alpha_e = 1, \qquad \sum_e \alpha_e \vec{n}_e = 0. \tag{9}$$

The case of two outcomes corresponds simply to the usual observable $P_e = |e\rangle_{\vec{n}\vec{n}}\langle e|$, $e = 0, 1$, with $|e\rangle_{\vec{n}}$ eigenvector of $\vec{n}\cdot\vec{\sigma}$ corresponding to the eigenvalues $+1, -1$, respectively. In fact, for two outcomes one has $\alpha_0\vec{n}_0 + \alpha_1\vec{n}_1 = 0$, namely $\vec{n}_0 = -\vec{n}_1 \doteq \vec{n}$, and necessarily $\alpha_0 = \alpha_1 = \frac{1}{2}$. We now consider the cases of three and four elements. By theorem 2 a necessary and sufficient condition for extremality is

$$\sum_{e\in\mathsf{E}} \gamma_e \alpha_e (I + \vec{n}_e \cdot \vec{\sigma}) = 0 \quad \Longleftrightarrow \quad \gamma_e = 0 \qquad \forall e \in \mathsf{E}, \tag{10}$$

or, equivalently,

$$\sum_{e\in\mathsf{E}} \gamma_e \alpha_e = 0, \qquad \sum_{e\in\mathsf{E}} \gamma_e \alpha_e \vec{n}_e = 0 \quad \Longleftrightarrow \quad \gamma_e = 0 \qquad \forall e \in \mathsf{E}. \tag{11}$$

For three outcomes, equation (9) implies that $\{\alpha_e\vec{n}_e\}_{e\in\mathsf{E}}$ represent the edges of a triangle, and thus the second condition in equation (11) is satisfied iff $\gamma_e \equiv \gamma$ is independent of e. Then the first condition is satisfied iff $\gamma \equiv 0$. Therefore, all three outcomes rank-one POVMs with pairwise non-proportional elements are extremal. For four outcomes we can see that for an extremal POVM the corresponding unit vectors $\{\vec{n}_e\}_{e\in\mathsf{E}}$ cannot lie on a common plane. Indeed, divide the four vectors $\{\alpha_e\vec{n}_e\}_{e\in\mathsf{E}}$ into two couples, which identify two intersecting planes. Then, the third condition in equation (9) implies that the sums of the couples lie on the intersection of the planes and have the same length and opposite direction. If we multiply by independent scalars γ_e the two elements of a couple, their sum changes direction and lies no longer in the intersection of the two planes, and the second condition in equation (11) cannot be satisfied. Therefore, the two elements of the same couple must be multiplied by the same scalar, which then just rescales their sum. Now, when the rescaling factors are different for the two couples, the two partial sums no longer do sum to the null vector. Then necessarily $\gamma_e \equiv \gamma$ independently of e. In order to satisfy also the first condition in equation (11) we must have $\gamma = 0$. On the other hand, if the four unit vectors lie in the same plane, a non-trivial linear combination can always be found that equals the null operator, hence the POVM is not extremal. By the first condition in equation (11) we have indeed

$$\gamma_0\alpha_0 = -\sum_{e=1,2,3} \gamma_e\alpha_e, \tag{12}$$

then the second condition can be written as

$$\sum_{e=1,2,3} \gamma_e\alpha_e(\vec{n}_e - \vec{n}_0) = 0. \tag{13}$$

Now, either we have a couple of equal vectors $\vec{n}_e$, or the three vectors $\vec{n}_e - \vec{n}_0$ in a two-dimensional plane are linearly dependent. However, in both cases the POVM is not extremal, because in the former case two elements are proportional, while in the latter a non-trivial triple of coefficients γ_e satisfying equation (13) exists.

Note that the three- and four-outcomes POVMs are necessarily unsharp, i.e. there is no state with probability distribution $p(e|\rho) = \delta_{e,\bar{e}}$ for a fixed $\bar{e}$. They provide examples of un-sharp POVMs with purely intrinsical quantum noise.

4. The boundary of the convex set of POVMs

In this section we generalize the results about extremality, and give a full characterization of the elements on the boundary of the convex set of n-outcomes POVMs on the Hilbert space H. Let start from an intuitive geometrical definition of the boundary of a convex set. Consider for example a point lying on some face of a polyhedron. Then there exists a direction (e.g. normal to the face) such that any shift of the point along that direction will bring it inside the convex set, while in the opposite direction it will bring the point outside the convex. In mathematical terms, consider a convex set $\mathscr{C}$ and an element $p \in \mathscr{C}$. Then, p belongs to the boundary $\partial\mathscr{C}$ of $\mathscr{C}$ if and only if there exists $q \in \mathscr{C}$ such that

$$p + \epsilon(q - p) \in \mathscr{C}, \qquad p - \epsilon(q - p) \notin \mathscr{C}, \qquad \forall\epsilon \in [0, 1]. \tag{14}$$

This definition leads to the following characterization of the boundary of the convex set of N-outcomes POVMs.

Theorem 4. *A POVM* $\mathbf{P} \in \mathscr{P}_N$ *belongs to the boundary of* $\mathscr{P}_N$ *iff at least one element* P_f *of* $\mathbf{P}$ *has a non-trivial kernel.*

Let us first prove necessity. Consider two different POVMs $\mathbf{P}$ and $\mathbf{Q}$, and suppose that $\forall\epsilon \in [0, 1]$ $\mathbf{P} + \epsilon\mathbf{D}$ is still a POVM while $\mathbf{P} - \epsilon\mathbf{D}$ is not. This happens only if $\forall\epsilon \in [0, 1] P_f - \epsilon D_f \not\geqslant 0$ for some f. Then some vector ψ must exist such that

$$\langle\psi|P_f|\psi\rangle < \epsilon\langle\psi|D_f|\psi\rangle, \qquad \forall\epsilon \in [0, 1], \tag{15}$$

namely $\langle\psi|P_f|\psi\rangle = 0$. Since by hypothesis P_f is positive semidefinite, then necessarily $\psi \in \mathsf{Ker}(P_f)$. To prove that the condition is also sufficient, consider a POVM element P_f with non-trivial kernel, and take $\psi \in \mathsf{Ker}(P_f)$. Then consider an event g such that $\langle\psi|P_g|\psi\rangle > 0$ (such event must exist for normalization of the POVM), and take $D_f = \kappa|\psi\rangle\langle\psi|$, $D_g = -\kappa|\psi\rangle\langle\psi|$, and $D_e = 0$ otherwise, with κ smaller than the minimum eigenvalue of P_g. Clearly $\forall\epsilon \in [0, 1]\mathbf{P} + \epsilon\mathbf{D}$ is a POVM while $\mathbf{P} - \epsilon\mathbf{D}$ is not, since the element $P_f - \epsilon D_f$ is not positive semidefinite.

We now proceed to study the structure of the faces of $\mathscr{P}_N$. For such a purpose it is convenient to regard a convex set as a subset of an affine space, whose dimension is the number of linearly independent directions along which any internal point can symmetrically be shifted. Clearly, also the faces of the convex set are themselves convex. For example, moving from a point inside a cube one can explore three dimensions while remaining inside, whereas, for the cube faces the number of independent symmetric perturbations is two, and for the sides this number reduces to one. We will keep in mind the above geometrical picture for the classification of the border of the convex set of POVMs using the perturbation method.

According to the results of section 2, a perturbation for a POVM $\mathbf{P}$ is a set of Hermitian operators $\mathbf{D} = \{D_e\}$ with $\sum_{e\in\mathsf{E}} D_e = 0$, and with $\mathsf{Supp}(D_e) \subseteq \mathsf{Supp}(P_e)$. Expressed in the orthonormal basis of eigenvectors of the POVM elements as in equation (5), the operators D_e read

$$D_e = \sum_{mn=1}^{\mathrm{rank}(P_e)} D_{mn}^{(e)}\big|v_m^{(e)}\big\rangle\big\langle v_n^{(e)}\big|. \tag{16}$$

We recall that, according to theorem 2, non-trivial perturbations for $\mathbf{P}$ exist only if the outer products $\big|v_m^{(e)}\big\rangle\big\langle v_n^{(e)}\big|$ are linearly dependent with $e \in \mathsf{E}$ and $1 \leqslant n, m \leqslant \mathrm{rank}(P_e)$. The total number of such outer products is

$$r(\mathbf{P}) \doteq \sum_{e\in\mathsf{E}} \mathrm{rank}(P_e)^2. \tag{17}$$

The number of linearly independent elements in the set of outer products $\left|v_m^{(e)}\right\rangle\left\langle v_n^{(e)}\right|$ is given by

$$l(\mathbf{P}) \doteq \dim\left[\mathsf{Span}\left(\left\{\left|v_m^{(e)}\right\rangle\left\langle v_n^{(e)}\right|\right\}\right)\right]. \tag{18}$$

Now, as in theorem 4, we see that the number $b(\mathbf{P})$ of independent perturbations for $\mathbf{P}$ is given by

$$b(\mathbf{P}) = r(\mathbf{P}) - l(\mathbf{P}). \tag{19}$$

In fact, in equation (5) we have $r(\mathbf{P}) = \sum_{e\in\mathsf{E}} \operatorname{rank}(P_e)^2$ variables $D_{nm}^{(e)}$, whereas the number of linearly independent equations is $l(\mathbf{P}) = \dim\left(\mathsf{Span}\left(\left\{\left|v_m^{(e)}\right\rangle\left\langle v_n^{(e)}\right|\right\}\right)\right)$, whence the number of variables which can be written as linear combination of a linearly independent set is $r(\mathbf{P}) - l(\mathbf{P})$. On the other hand, the dimension of the affine space of the convex set $\mathscr{P}_N$ is given by $d^2(N-1)$, since the POVM normalization constraint corresponds to d^2 independent linear equations, with Nd^2 variables. We have then proved the following characterization of the border of $\mathscr{P}_N$.

Theorem 5. *A POVM* $\mathbf{P} \in \mathscr{P}_N$ *belongs to the boundary* $\partial\mathscr{P}_N$ *of* $\mathscr{P}_N$ *iff* $b(\mathbf{P}) < d^2(N-1)$, $b(\mathbf{P})$ *defined in equations (17)–(19) being the dimension of the face in which* $\mathbf{P}$ *lies.*

From the above theorem it also follows that a POVM $\mathbf{P} \in \mathscr{P}_N$ on the boundary $\partial\mathscr{P}_N$ of $\mathscr{P}_N$ also belongs to $\partial\mathscr{P}_M$ with $M \geqslant N$, whence it belongs to the boundary $\partial\mathscr{P}$ of $\mathscr{P}$. This also implies that $\partial\mathscr{P}_N \subseteq \partial\mathscr{P}_M \subseteq \partial\mathscr{P}$.

5. Extremal informationally complete POVMs

In this section we will give an explicit construction of a rank-one informationally complete POVM as in corollary 6, in this way also proving the existence of extremal POVMs with d^2 elements.

Consider the *shift-and-multiply* finite group of unitary operators

$$U_{pq} = Z^p W^q, \qquad p, q \in \mathbb{Z}_d \tag{20}$$

where $\mathbb{Z}_d = \{0, 1, \ldots, d-1\}$, and Z and W are defined as follows,

$$Z = \sum_j |j \oplus 1\rangle\langle j|, \qquad W = \sum_j \omega^j |j\rangle\langle j|, \tag{21}$$

with $\oplus$ denoting the sum modulo d, $\{|j\rangle\}_0^{d-1}$ an orthonormal basis, $\omega = \mathrm{e}^{\frac{2\pi\mathrm{i}}{d}}$, and the sums are extended to $\mathbb{Z}_d$. We now prove that the following POVM with d^2 outcomes is extremal

$$P_{pq} \doteq \frac{1}{d} U_{pq} \nu U_{pq}^\dagger, \tag{22}$$

for any pure state $\nu = |\psi\rangle\langle\psi|$ on H satisfying the constraints

$$\operatorname{Tr}\left[U_{pq}^\dagger \nu\right] \neq 0, \qquad \forall p, q \in \mathbb{Z}_d. \tag{23}$$

In order to prove the statement, first we note that the operators $d^{-\frac{1}{2}} U_{pq}$ form a complete orthonormal set of unitary operators, i.e. they satisfy

$$\operatorname{Tr}[U_{pq} U_{p',q'}] = d\delta_{pp'}\delta_{qq'}, \tag{24}$$

$$\sum_{pq} U_{pq} \Xi U_{pq}^\dagger = d \operatorname{Tr}[\Xi], \tag{25}$$

for any operator Ξ. From equation (25) it immediately follows that $\sum_{pq} P_{pq} = I$, whence $\{P_{pq}\}$ is a POVM. Then, in order to prove extremality, according to corollary 5 it is sufficient to prove that the d^2 operators P_{pq} are linearly independent, which in turn can be proved by showing that $\{P_{pq}\}$ is itself a complete set in the space of operators (completeness along with the fact that the elements P_{pq} are d^2 implies indeed that they are linearly independent). As mentioned after corollary 6, such a kind of completeness for the POVM corresponds to a so-called *informationally complete measurement* [11]. The completeness of the set $\{P_{pq}\}$ is equivalent to the invertibility of the following operator on $\mathsf{H}^{\otimes 2}$,

$$F = \sum_{pq} |P_{pq}\rangle\!\rangle\langle\!\langle P_{pq}|, \tag{26}$$

where the double-ket notation [14] is used to recall the equivalence between (Hilbert–Schmidt) operators A on H and vectors $|A\rangle\!\rangle = A \otimes I|I\rangle\!\rangle$ of $\mathsf{H}^{\otimes 2}$, $|I\rangle\!\rangle \in \mathsf{H}^{\otimes 2}$ denoting the reference vector $|I\rangle\!\rangle = \sum_{j\in\mathbb{Z}_d} |j\rangle \otimes |j\rangle$ defined in terms of the chosen orthonormal basis $\{|j\rangle\}$. By expanding $\nu = |\psi\rangle\langle\psi|$ over the basis $\{U_{pq}\}$ and using the multiplication rules of the group, one obtains

$$P_{pq} = \frac{1}{d^2} \sum_{rs} \mathrm{e}^{\frac{2\pi \mathrm{i}}{d}(qr-ps)} \operatorname{Tr}\left[U_{rs}^\dagger \nu\right] U_{rs}, \tag{27}$$

which allows us to rewrite equation (26) as follows:

$$F = \frac{1}{d^2} \sum_{rs} \left|\operatorname{Tr}\left[U_{rs}^\dagger \nu\right]\right|^2 |U_{rs}\rangle\!\rangle\langle\!\langle U_{rs}|. \tag{28}$$

Since $\{U_{pq}\rangle\!\rangle\}$ is an orthogonal basis in $\mathsf{H}^{\otimes 2}$, the invertibility of F is equivalent to condition (23), which is clearly satisfied by most density operators $\nu = |\psi\rangle\langle\psi|$ (condition (23) is satisfied by a set of states $|\psi\rangle$ that is dense in H). As an example of state satisfying condition (23), one can consider $|\psi\rangle \propto \sum_j \alpha^j |j\rangle$ for $0 < \alpha < 1$.

6. Conclusions

In this paper we have completely characterized the convex set of POVMs with discrete spectrum. Using the method of perturbations, we have determined the extremal points—the 'indecomposable apparatuses'—by three alternative characterizations corresponding to theorems 1–3, and with easier necessary or sufficient conditions in corollaries 1–7. In particular, we have shown that for finite dimension d an extremal POVMs can have at most d^2 outcomes, and an extremal POVM with d^2 outcomes always exists and is necessarily informationally complete. An explicit realization of such extremal informationally complete POVM has been given in section 3.

The characterization of the convex set $\mathscr{P}_N$ of POVMs with N outcomes has been obtained by determining its boundary $\partial\mathscr{P}_N$, which, in turn, has been characterized in terms of the number $b(\mathbf{P})$ of independent perturbations for the POVM $\mathbf{P}$ in equation (19). This has led to a simple characterization of the boundary in terms of simple algebraic properties of a POVM lying on it. Since $\partial\mathscr{P}_N$ is also a subset of the boundary $\partial\mathscr{P}$ of the full convex set $\mathscr{P}$ of POVMs with discrete spectrum, our result also provides a complete characterization of $\mathscr{P}$.

Finally, in the appendix we reported an algorithm for decomposing a point in a convex set into a minimum number of extremal elements, specializing the algorithm to the case of the convex set $\mathscr{P}_N$ of POVMs with N outcomes.

Acknowledgments

We are grateful to R Mecozzi for his participation at the preliminary stage of this research during his thesis work. This work has been co-founded by the EC under the program ATESIT (Contract No. IST-2000-29681) and the MIUR cofinanziamento 2002. PP acknowledges support from the INFM under project PRA-2002-CLON.

Appendix. Algorithm for decomposition of internal points into extremals

In this appendix we provide an algorithm to decompose a POVM $\mathbf{P} \in \mathscr{P}_N$ into extremal ones. We first present the algorithm in the general case of an abstract convex set, and then we specialize it to the case of POVMs.

Consider a convex set $\mathscr{C}$ and a point p inside it. We want to decompose p into extremal points. We first need two ingredients, which depend on the specific convex set $\mathscr{C}$ under consideration: the *affine space* $\mathscr{A}$ in which $\mathscr{C}$ is embedded (this is just the space of legitimate 'perturbations'), and an 'indicator' $\iota(p)$ which is positive for $p \in \mathscr{C}$, zero on the boundary, and negative outside $\mathscr{C}$. Now, starting from p inside $\mathscr{C}$, we move in some direction d in $\mathscr{A}$ until a face of $\mathscr{C}$ is encountered at $\lambda_+ \doteq \max\{\lambda : p + \lambda d \in \mathscr{C}\}$ (using our indicator λ_+ is given by the value of λ where $\iota(p + \lambda d)$ changes sign). Similarly in the opposite direction one hits the boundary at $\lambda_- \doteq \max\{\lambda : p - \lambda d \in \mathscr{C}\}$. The point p can now be split into the convex combination

$$p = \frac{\lambda_-}{\lambda_+ + \lambda_-} p_+ + \frac{\lambda_+}{\lambda_+ + \lambda_-} p_-, \qquad p_\pm \doteq p \pm \lambda_\pm d. \tag{A.1}$$

If p was in the interior of $\mathscr{C}$ any face of the boundary can be encountered, while if p was already on the boundary of $\mathscr{C}$ the perturbation d brings p on a 'face of a face', e.g. moving on a face of a cube towards an edge. In any case, the dimension $b(p)$ of the face to which p belongs is decreased at least by one.

By applying the same splitting scheme to both $p_\pm$ recursively, we obtain a weighted binary 'tree' of points rooted in p, with the property that the point p' at each node can be written as convex combination of its descendants, and with a depth bounded by the dimension $b(p)$ of the face to which p belongs. Of course the 'leaves' of the tree are extremal points p_i, and one can combine them to obtain the original point as $p = \sum_i \alpha_i p_i$ weighting each leaf p_i with the product α_i of all weights found along the path from the root p to the leaf p_i. Unfortunately, this raw algorithm can produce up to 2^r extremal points p_i, for dimension r of the affine space of $\mathscr{C}$, each leaf being addressed by the vector $d_i = p - p_i$, with $\sum_i \alpha_i d_i = 0$. However, by the Carathéodory's theorem [15], we know that at most $r + 1$ extremal points are needed to decompose p. Indeed, if the number of d_i is larger than $r + 1$, then they must be linearly dependent, and there must exist λ_i not all vanishing and not all positive such that $\sum_i \lambda_i d_i = 0$. Since $\sum_i \alpha_i d_i - \mu\big(\sum_i \lambda_i d_i\big) = 0$, by choosing the greatest μ such that $\alpha_i - \mu\lambda_i \geqslant 0 \quad \forall i$, one finds that 0 can be written as a convex combination of a smaller number of d_i. This procedure can be applied repeatedly to the remaining d_i (the α_i must also be upgraded) until their only combination giving 0 is the one whose coefficients are all positive: at this point their number is for sure not larger than $r + 1$. Therefore, from an initial decomposition with many elements, we end up with a decomposition of p into at most $r + 1$ vectors d_i and probabilities α_i such that $\sum_{i=1}^{r+1} \alpha_i d_i = 0$, whence $p = \sum_{i=1}^{r+1} \alpha_i (p + d_i) = \sum_{i=1}^{r+1} \alpha_i p_i$. Note that the evaluation of $\lambda_\pm$ at each step involves an eigenvalue evaluation, whence the algorithm generally does not provide analytical decompositions.

In order to specialize the algorithm to the case of the POVMs convex set $\mathscr{P}_N$, we need to specify both the corresponding affine space $\mathscr{A}_N$ and the indicator ι of the border.

The affine space is the real $d^2(N-1)$-dimensional linear space of vectors $\mathbf{D} = \{D_e\}$ of N Hermitian operators D_e, with $\sum_{e\in\mathsf{E}} D_e = 0$. This can be obtained as the real span of projectors $|n,e\rangle\langle n,e|$ along with $\mathrm{Re}|n,e\rangle\langle m,e|$ and $\mathrm{Im}|n,e\rangle\langle m,e|$ for $n = 1,\ldots,d$ and $n < m$, where $X = \mathrm{Re}\, X + i\,\mathrm{Im}\, X$ is the Cartesian decomposition of the operator X, $\{|n,e\rangle\}$ denotes any orthonormal basis for the eth copy of the Hilbert space H of the quantum system with $d = \dim(\mathsf{H})$, and $e \in \mathsf{E}' \doteq \mathsf{E}/\{1\}$, and $D_1 = -\sum_{e\in\mathsf{E}'} \mathrm{diag} D_e$. However, if such global basis is used, then when the search algorithm starts from a POVM which is already on a face of the convex one has the problem that generally the basis is not aligned with the face itself, and for a generic direction the perturbed POVM either exits from the convex, or it moves inside it. For that reason it is convenient to consider a 'local basis' of perturbations for a given $\mathbf{P}$. This can be constructed by considering the set $\{X^{(e)}_{m,n}\}$ of Hermitian operators defined as follows:

$$X^{(e)}_{mn} = \mathrm{Re}\,\big|v^{(e)}_m\big\rangle\big\langle v^{(e)}_n\big|, \qquad X^{(e)}_{nm} = \mathrm{Im}\big|v^{(e)}_m\big\rangle\big\langle v^{(e)}_n\big|, \qquad n < m$$
$$X^{(e)}_{nn} = \big|v^{(e)}_n\big\rangle\big\langle v^{(e)}_n\big|. \tag{A.2}$$

Then pick up $l(\mathbf{P})$ linearly independent elements, which we will denote by $V^{(e)}_{mn}$, and call the remaining ones $W^{(f)}_{mn}$, so that we can write

$$W^{(f)}_{mn} = \sum_{epq} c^{fmn}_{epq} V^{(e)}_{pq} \equiv \sum_{e\neq f}\sum_{pq} c^{fmn}_{epq} V^{(e)}_{pq}, \tag{A.3}$$

where the second identity is a consequence of linear independence of operators $\big\{\big|v^{(e)}_m\big\rangle\big\langle v^{(e)}_n\big|\big\}$ for fixed e. Then we can construct the following basis $\{D_e(fmn)\}$ for $\mathbf{P}$ perturbations

$$D_f(fmn) = \sum_{pq} c^{fmn}_{fpq} V^{(f)}_{pq} - W^{(f)}_{mn},$$
$$D_e(fmn) = \sum_{pq} c^{fmn}_{epq} V^{(e)}_{pq}. \tag{A.4}$$

Clearly one has $\sum_{e\in\mathsf{S}} D_e(fmn) = 0\ \forall fmn$, and modulo a suitable rescaling one has $P_e \pm D_e(fmn) \geqslant 0\ \forall fmn$. Notice that, by construction, the operators $D_e(fnm)$ are linearly independent. In fact, using equations (A.4), a generic linear combination of $D_e(fmn)$ for each fixed $e \in \mathsf{E}$ will result in a linear combination of $\big\{V^{(e)}_{mn}\big\} \cup \big\{W^{(e)}_{pq}\big\}$ for the same e. But the set $\big\{V^{(e)}_{mn}\big\} \cup \big\{W^{(e)}_{pq}\big\}$ is linearly independent for fixed $e \in \mathsf{E}$, due to equation (A.4).

As regards the indicator of the boundary, this is simply the minimum of the eigenvalues of the operators P_e, which changes from positive to negative when $\mathbf{P}$ crosses the border.

References

[1] Wootters W K and Zurek W H 1982 *Nature* **299** 802
[2] Yuen H P 1986 *Phys. Lett.* A **113** 405
[3] D'Ariano G M and Yuen H P 1996 *Phys. Rev. Lett.* **76** 2832
[4] Helstrom C W 1976 *Quantum Detection and Estimation Theory* (New York: Academic)
[5] Holevo A S 1982 *Probabilistic and Statistical Aspects of Quantum Theory* (Amsterdam: North-Holland)
[6] Naimark M A 1940 *Iza. Akad. Nauk USSR, Ser. Mat.* **4** 277
[7] Ozawa M 1984 *J. Math. Phys.* **25** 79
[8] Davies E B 1976 *Quantum Theory of Open Systems* (London: Academic)
[9] Kraus K 1983 *States, Effects, and Operations: Fundamental Notions of Quantum Theory (Lecture Notes in Physics* vol 190*)* (Berlin: Springer)
[10] Busch P, Lahti P J and Mittelstaedt P 1991 *The Quantum Theory of Measurement (Lecture Notes in Physics* vol 2*)* (Berlin: Springer)

[11] D'Ariano G M, Perinotti P and Sacchi M F 2004 *J. Opt. B: Quantum Semiclass. Opt.* **6** S487
[12] Parthasaraty K R 1999 *Inf. Dim. Anal.* **2** 557
[13] Størmer E 1974 *Foundations of Quantum Mechanics and Ordered Linear Spaces* ed A Hartkämper and H Neumann (Berlin: Springer) p 85
[14] D'Ariano G M, Lo Presti P and Sacchi M 2000 *Phys. Lett.* A **272** 32
[15] Rockafellar R T 1972 *Convex Analysis* (Princeton, NJ: Princeton University Press)