Rényi formulation of the entropic uncertainty principle for POVMs

2010 J. Phys. A: Math. Theor. 43 155302

(http://iopscience.iop.org/1751-8121/43/15/155302)

J. Phys. A: Math. Theor. **43** (2010) 155302 (9pp) doi:10.1088/1751-8113/43/15/155302

Rényi formulation of the entropic uncertainty principle for POVMs

Alexey E Rastegin

Department of Theoretical Physics, Irkutsk State University, Gagarin Boulevard 20, Irkutsk 664003, Russia

E-mail: rast@api.isu.ru

Received 25 August 2009, in final form 16 February 2010
Published 25 March 2010
Online at stacks.iop.org/JPhysA/43/155302

Abstract

Uncertainty relations for a pair of arbitrary measurements are posed in the form of inequalities using the Rényi entropies. The formulation deals with POVM measurements on finite-dimensional Hilbert space. Both the entropic uncertainty relations with state-dependent and state-independent bounds are presented. The results are illustrated on the example of measurements for state discrimination.

PACS numbers: 03.65.Ta, 03.67.−a

1. Introduction

The Heisenberg uncertainty principle [1] is the best known of those results that express distinctions between the quantum world and the classical world. Several ways to pose the uncertainty principle have been developed. According to Robertson's formulation [2], the standard deviations of the observables A and B measured in the quantum state ψ obey $\Delta\mathsf{A}\,\Delta\mathsf{B} \geqslant (1/2)|\langle\psi, [\mathsf{A}, \mathsf{B}]\psi\rangle|$, where $\langle\cdot,\cdot\rangle$ denotes the inner product. Due to various measurement scenarios, many relations have been stated in more detailed terms (see [3–5] and references therein). In particular, explicit noise-disturbance relations were given [6, 7]. Uncertainty relations are also linked with the complementarity [8, 9] and mutually unbiased bases initially studied in [10]. The concept of entropy is widely used in physics [11]. So it is natural to express Heisenberg's uncertainty principle in terms of entropic measures. The first relation for position and momentum in terms of the Shannon entropies was obtained by Hirschman [12]. Hirschman's result has been improved by Beckner [13], and by Bialynicki-Birula and Mycielski [14].

In general, entropic formulation of the uncertainty principle was examined by Deutsch [15]. He stressed that the bound $(1/2)|\langle\psi, [\mathsf{A}, \mathsf{B}]\psi\rangle|$ becomes trivial for any eigenstate of either of the two observables. On the other hand, to each measured observable we can assign the Shannon entropy of generated probability distribution. Deutsch obtained a lower bound on the sum of the Shannon entropies related to observables without degeneracy [15]. The improved lower bound had been conjectured by Kraus [16] and later stated by Maassen and Uffink [17]. The method of proof utilized Riesz's theorem. The authors of [17] have also shown an extension of this method to some canonical variables. A relevant approach to degenerate observables has been developed by the authors of [18]. Using Naimark's theorem, they derived the entropic relation for a pair of generalized measurements as well. An entropic uncertainty relation for $(d+1)$ mutually unbiased orthogonal measurements in d-dimensional Hilbert space was obtained [19]. For the specified classes of observables, the entropic relations were considered in the papers [20, 21].

There exist two considerable extensions of the Shannon entropy. The Rényi entropy [22] is now used in many disciplines. Its properties from the physical viewpoint are examined in [23]. Some bounds in terms of this measure are discussed in [24, 25]. Larsen expressed uncertainty relations via so-called purities [26], which are connected with the Rényi entropy of order 2. Bialynicki-Birula obtained the relations in terms of the Rényi entropies for the position–momentum and angle–angular momentum pairs [27]. Using the Tsallis entropy, some bounds were also given [28–30]. Recently, different kinds of uncertainty principle have been discussed [31–33].

In the present work, we will obtain uncertainty relations for a pair of arbitrary measurements in the form of inequalities using the Rényi entropies. Some remarks on the physical sense of the proved inequalities will be given. The paper is organized as follows. In section 2, the main result and the notation are described. Sections 3 and 4 are devoted to the proof of the claimed result. Due to some complexity, we proceed in several steps. In section 5, a significance of the treated entropic bound in comparison with Robertson's formulation is discussed. We also consider an example in which one measurement is not actually projective.

2. Notation and main result

For real $\alpha > 0$ and $\alpha \neq 1$, the Rényi entropy of order α of the probability distribution $\{p_i\}$ is defined by [22]

$$H_\alpha(p) := (1-\alpha)^{-1} \ln \left\{ \sum_i p_i^\alpha \right\}. \tag{2.1}$$

This quantity is a non-increasing function of α: if $\alpha < \beta$, then $H_\alpha \geqslant H_\beta$ [22]. For $\alpha > 1$, the Rényi entropy $H_\alpha(p)$ is neither purely convex nor purely concave [23]. The limit $\alpha \to 1$ recovers the Shannon entropy $H_1(p) = -\sum_i p_i \ln p_i$. In the following, orders of the Rényi entropy are assumed to be different from 1. The bounds for the Shannon entropies can be obtained by taking the limit $\alpha \to 1$ in the final inequalities.

Let $\mathcal{H}$ be the finite-dimensional Hilbert space. A general quantum measurement is described by 'positive operator-valued measure' (POVM). This is a set $\{\mathsf{M}_i\}$ of positive semidefinite matrices obeying the completeness relation $\sum_i \mathsf{M}_i = \mathbf{I}$, where $\mathbf{I}$ is the identity matrix [34]. For given POVM $\{\mathsf{M}_i\}$ and density operator ρ on $\mathcal{H}$, the probability of ith outcome is equal to $p_i = \mathrm{tr}(\mathsf{M}_i\rho)$ [34]. In mathematics literature, such a set $\{\mathsf{M}_i\}$ is often called 'generalized resolution of the identity' for the space $\mathcal{H}$ (for a discussion, see [35]). In the case of orthogonal projections, we call them 'orthogonal resolution of the identity' or 'projector-valued measure' (PVM).

The Rényi entropy $H_\alpha(\mathsf{M}|\rho)$ of generated probability distribution is then defined by equations (2.1) and $p_i = \mathrm{tr}(\mathsf{M}_i\rho)$. When a quantum state is pure, i.e. $\rho = \psi\psi^\dagger$ and $\|\psi\| = 1$, we write $H_\alpha(\mathsf{M}|\psi)$. In this case, we have $\mathrm{tr}(\mathsf{M}_i\rho) = \langle\psi, \mathsf{M}_i\psi\rangle$. Let $\{\mathsf{M}_i\}$ and $\{\mathsf{N}_j\}$ be two POVMs, and let ψ be a pure state. By definition, we put the function

$$f(\mathsf{M}, \mathsf{N}|\psi) := \max_{ij} \left\|\mathsf{M}_i^{1/2}\psi\right\|^{-1} \left\|\mathsf{N}_j^{1/2}\psi\right\|^{-1} |\langle\mathsf{M}_i\psi, \mathsf{N}_j\psi\rangle|, \tag{2.2}$$

where the maximization is over those values of labels i and j that satisfy $\left\|\mathsf{M}_i^{1/2}\psi\right\| \neq 0$ and $\left\|\mathsf{N}_j^{1/2}\psi\right\| \neq 0$. In the case of a mixed state ρ with the spectral decomposition

$$\rho = \sum_\lambda \lambda\, \psi_\lambda \psi_\lambda^\dagger \tag{2.3}$$

we also define

$$f(\mathsf{M}, \mathsf{N}|\rho) := \max\left\{f(\mathsf{M}, \mathsf{N}|\psi_\lambda) :\ \lambda \in \mathrm{spec}(\rho)\right\}. \tag{2.4}$$

Theorem 1. *Let* $\{\mathsf{M}_i\}$ *and* $\{\mathsf{N}_j\}$ *be two POVM measurements. Then for the arbitrary density operator* ρ*, there holds*

$$H_\alpha(\mathsf{M}|\rho) + H_\beta(\mathsf{N}|\rho) \geqslant -2\ln f(\mathsf{M}, \mathsf{N}|\rho), \tag{2.5}$$

where orders α *and* β *satisfy* $1/\alpha + 1/\beta = 2$*.*

This statement generalizes theorem 2.5 of [18] in the following respects. First, we deal with the Rényi entropies instead of the Shannon entropies. Second, only for pure state, the state-dependent bound is given in [18], whereas our result holds for the arbitrary mixed state. Deriving a relation for the mixed state, the authors of [18] used concavity of the Shannon entropy. But this is a very difficult way to obtain state-dependent bound. We can also give a variety of (2.5) that does not depend on the quantum state. Putting the operator norm $\|\mathsf{A}\| := \max\{\|\mathsf{A}\psi\| : \|\psi\| = 1\}$, we define

$$\bar{f}(\mathsf{M}, \mathsf{N}) := \max_{ij} \left\|\mathsf{M}_i^{1/2}\mathsf{N}_j^{1/2}\right\|. \tag{2.6}$$

Using the relation $|\langle\mathsf{M}_i\psi, \mathsf{N}_j\psi\rangle| \leqslant \left\|\mathsf{M}_i^{1/2}\mathsf{N}_j^{1/2}\right\| \left\|\mathsf{M}_i^{1/2}\psi\right\| \left\|\mathsf{N}_j^{1/2}\psi\right\|$ proven in [18], and definitions (2.2) and (2.4), one satisfies the inequality $f(\mathsf{M}, \mathsf{N}|\rho) \leqslant \bar{f}(\mathsf{M}, \mathsf{N})$. It then follows from (2.5) that

$$H_\alpha(\mathsf{M}|\rho) + H_\beta(\mathsf{N}|\rho) \geqslant -2\ln \bar{f}(\mathsf{M}, \mathsf{N}) \tag{2.7}$$

under the condition $1/\alpha + 1/\beta = 2$. The particular case of (2.7) with the Shannon entropies was stated in [18]. For two one-rank PVMs, it reduces to the relation that had been conjectured by Kraus [16] and later proved by Maassen and Uffink [17]. Note that the authors of [17] have briefly discussed an extension of their relation to canonically conjugate variables including the standard case of position and momentum.

We now recall a version of Riesz's theorem (see theorem 297 in [36]). Let $\mathsf{x} \in \mathbb{C}^n$ be n-tuple of complex numbers x_j and let t_{ij} be the entries of the matrix T of order $m \times n$. Define η to be maximum of $|t_{ij}|$, i.e. $\eta := \max\{|t_{ij}| :\ 1 \leqslant i \leqslant m, 1 \leqslant j \leqslant n\}$. The fixed matrix T describes a linear transformation $\mathbb{C}^n \to \mathbb{C}^m$. That is, to each x we assign m-tuple $\mathsf{y} \in \mathbb{C}^m$ with elements

$$y_i(x) := \sum_{j=1}^{n} t_{ij}\, x_j \qquad (i = 1, \ldots, m). \tag{2.8}$$

For arbitrary $b \geqslant 1$, we also define the function $S_b(x) := \left(\sum_j |x_j|^b\right)^{1/b}$.

Lemma 2. *Suppose the matrix* T *satisfies*

$$\sum_i |y_i|^2 \leqslant \sum_j |x_j|^2 \tag{2.9}$$

for all $\mathsf{x} \in \mathbb{C}^n$*; then*

$$S_a(y) \leqslant \eta^{(2-b)/b}\, S_b(x), \tag{2.10}$$

where $1/a + 1/b = 1$ *and* $1 < b < 2$.

3. Both the measurements are projective

A standard quantum measurement is described by PVM. This is a set $\{\mathsf{P}_i\}$ of Hermitian matrices obeying the property $\mathsf{P}_i\mathsf{P}_k = \delta_{ik}\mathsf{P}_i$ and the completeness relation. The two PVMs $\{\mathsf{P}_i\}$ and $\{\mathsf{Q}_j\}$ generate two probability distributions

$$p_i^{(\psi)} = \langle \mathsf{P}_i\psi, \mathsf{P}_i\psi\rangle, \qquad q_j^{(\psi)} = \langle \mathsf{Q}_j\psi, \mathsf{Q}_j\psi\rangle. \tag{3.1}$$

Proposition 3. *For any two projective measurements* $\{\mathsf{P}_i\}$ *and* $\{\mathsf{Q}_j\}$ *and an arbitrary density matrix* ρ*, there holds*

$$H_\alpha(\mathsf{P}|\rho) + H_\beta(\mathsf{Q}|\rho) \geqslant -2\ln f(\mathsf{P},\mathsf{Q}|\rho), \tag{3.2}$$

where orders α *and* β *satisfy* $1/\alpha + 1/\beta = 2$.

Proof. For those values of labels i and j that fulfil $\|\mathsf{P}_i\psi\| \neq 0$ and $\|\mathsf{Q}_j\psi\| \neq 0$, we define two sets of vectors

$$\mathsf{u}_i := \|\mathsf{P}_i\psi\|^{-1}\,\mathsf{P}_i\psi, \qquad \mathsf{v}_j := \|\mathsf{Q}_j\psi\|^{-1}\,\mathsf{Q}_j\psi. \tag{3.3}$$

With no loss of generality, we can mean that $1 \leqslant i \leqslant m$ and $1 \leqslant j \leqslant n$. The u_i's and the v_j's form two (generally incomplete) orthonormal sets in $\mathcal{H}$. We now define

$$t_{ij} := \langle \mathsf{u}_i, \mathsf{v}_j\rangle. \tag{3.4}$$

In line with (2.8) and (3.4), we have $y_i(x) = \langle \mathsf{u}_i, \mathsf{w}\rangle$, where $\mathsf{w} := \sum_j x_j\mathsf{v}_j$ by definition. It is clear that the vector $\sum_i y_i\mathsf{u}_i$ is an orthogonal projection of the vector w onto the subspace spanned by the u_i's. Hence, we obtain

$$\sum_i |y_i|^2 \leqslant \|\mathsf{w}\|^2 \equiv \sum_j |x_j|^2. \tag{3.5}$$

Therefore, condition (2.9) is satisfied for all $\mathsf{x} \in \mathbb{C}^n$. So result (2.10) can be applied. We shall now utilize this result for the values $y_i' = \|\mathsf{P}_i\psi\|$, $x_j' = \|\mathsf{Q}_j\psi\|$. Using the completeness relation for PVMs $\{\mathsf{P}_i\}$ and $\{\mathsf{Q}_j\}$ and $\langle \mathsf{u}_i, \mathsf{u}_k\rangle = \delta_{ik}$, we obtain $y_i' = \sum_j \langle \mathsf{u}_i, \mathsf{v}_j\rangle\, x_j'$. Thus, the values y_i' and x_j' are really connected by relation (2.8) with the matrix elements (3.4), and $p_i^{(\psi)} = |y_i'|^2$, $q_j^{(\psi)} = |x_j'|^2$ due to (3.1). Let us put $a = 2\alpha$ and $b = 2\beta$. Squaring (2.10), after relevant substitutions we get

$$S_\alpha(p^{(\psi)}) \leqslant f(\mathsf{P},\mathsf{Q}|\psi)^{2(1-\beta)/\beta}\, S_\beta(q^{(\psi)}), \tag{3.6}$$

where $1/\alpha + 1/\beta = 2$ and $1/2 < \beta < 1$. Indeed, due to $\mathsf{P}_i = \mathsf{P}_i^{1/2}$, $\mathsf{Q}_j = \mathsf{Q}_j^{1/2}$ (3.3) and (3.4), the maximum of $|t_{ij}|$ is equal to $f(\mathsf{P},\mathsf{Q}|\psi)$. For mixed state of the form (2.3), the corresponding probabilities are rewritten as

$$p_i = \mathrm{tr}(\mathsf{P}_i\rho) = \sum_\lambda \lambda p_i^{(\lambda)}, \qquad q_j = \mathrm{tr}(\mathsf{Q}_j\rho) = \sum_\lambda \lambda q_j^{(\lambda)}. \tag{3.7}$$

Here, the values $p_i^{(\lambda)}$ and $q_j^{(\lambda)}$ are defined by substituting ψ_λ for ψ into equations (3.1). Inequality (3.6) holds for each of the states ψ_λ. Due to definition (2.4), for all λ and the same conditions on α and β we can write

$$\lambda\, S_\alpha(p^{(\lambda)}) \leqslant f(\mathsf{P},\mathsf{Q}|\rho)^{2(1-\beta)/\beta}\lambda\, S_\beta(q^{(\lambda)}). \tag{3.8}$$

In this stage, we use the Minkowski inequality [36]. By $\alpha > 1$ and $\beta < 1$, there hold

$$S_\alpha\left(\sum_\lambda \lambda p^{(\lambda)}\right) \leqslant \sum_\lambda \lambda S_\alpha(p^{(\lambda)}), \qquad \sum_\lambda \lambda S_\beta(q^{(\lambda)}) \leqslant S_\beta\left(\sum_\lambda \lambda q^{(\lambda)}\right).$$

Summing (3.8) with respect to λ and using these inequalities, we finally get

$$S_\alpha(p) \leqslant f(\mathsf{P},\mathsf{Q}|\rho)^{2(1-\beta)/\beta} S_\beta(q). \tag{3.9}$$

By (2.1), we have $\ln S_\alpha(p) = \alpha^{-1}(1-\alpha)H_\alpha(p)$, $\ln S_\beta(q) = \beta^{-1}(1-\beta)H_\beta(q)$. It then follows from (3.9) and $(1-\beta)/\beta > 0$ that

$$\alpha^{-1}(1-\alpha)\beta(1-\beta)^{-1}H_\alpha(p) \leqslant 2\ln f(\mathsf{P},\mathsf{Q}|\rho) + H_\beta(q). \tag{3.10}$$

If $1/\alpha + 1/\beta = 2$ and $\alpha, \beta \neq 1$, then the factor of $H_\alpha(p)$ in (3.10) is equal to (-1). So inequality (3.10) leads to (3.2). This concludes the proof for $\alpha > \beta$. By permutation of the two PVMs, we recover the remaining case. □

4. The general case

Developing the ideas of [18], we use the Naimark extension. All the necessary details are gathered in the appendix.

Proposition 4. *Let $\{\mathsf{M}_i\}$ be a POVM measurement, and let $\{\mathsf{Q}_j\}$ be a PVM measurement. Then for an arbitrary mixed state ρ*

$$H_\alpha(\mathsf{M}|\rho) + H_\beta(\mathsf{Q}|\rho) \geqslant -2\ln f(\mathsf{M},\mathsf{Q}|\rho), \tag{4.1}$$

where orders α and β satisfy $1/\alpha + 1/\beta = 2$.

Proof. Substituting M_i for E_i and Q_j for G_j in the formulae of the appendix, we consider the measurements $\{\widetilde{\mathsf{M}}_i\}$ and $\{\widetilde{\mathsf{Q}}_j\}$ in the enlarged space $\widetilde{\mathcal{H}}$. The measurement $\{\widetilde{\mathsf{M}}_i\}$ is projective due to the Naimark theorem. The measurement $\{\widetilde{\mathsf{Q}}_j\}$ is projective, since the measurement $\{\mathsf{Q}_j\}$ is projective. By relation (3.2), we have

$$H_\alpha(\widetilde{\mathsf{M}}|\widetilde{\omega}) + H_\beta(\widetilde{\mathsf{Q}}|\widetilde{\omega}) \geqslant -2\ln f(\widetilde{\mathsf{M}},\widetilde{\mathsf{Q}}|\widetilde{\omega}) \tag{4.2}$$

for an arbitrary mixed state $\widetilde{\omega}$ in the enlarged space $\widetilde{\mathcal{H}}$. To each density matrix of the form (2.3) assign the density matrix

$$\widetilde{\rho} = \sum_\lambda \lambda\, \widetilde{\psi}_\lambda \widetilde{\psi}_\lambda^\dagger, \tag{4.3}$$

where the state vector $\widetilde{\psi}_\lambda$ is given by $\widetilde{\psi}_\lambda{}^T := [\psi_\lambda^T\ \mathbf{0}]$. In the particular case of the matrix $\widetilde{\rho}$, relation (4.2) is also valid. Since the vectors ψ_λ form an orthonormal set in $\mathcal{H}$,

$$\widetilde{\psi}_\lambda^\dagger \widetilde{\psi}_\mu = [\psi_\lambda^\dagger \quad \mathbf{0}]\begin{bmatrix}\psi_\mu \\ \mathbf{0}\end{bmatrix} = \psi_\lambda^\dagger \psi_\mu = \delta_{\lambda\mu}. \tag{4.4}$$

So the vectors $\widetilde{\psi}_\lambda$ form the (incomplete) orthonormal set in $\widetilde{\mathcal{H}}$. Due to this fact and the properties of the trace,

$$\mathrm{tr}(\widetilde{\mathsf{M}}_i\widetilde{\rho}) = \sum_\lambda \lambda\langle\widetilde{\psi}_\lambda, \widetilde{\mathsf{M}}_i\widetilde{\psi}_\lambda\rangle = \sum_\lambda \lambda\langle\psi_\lambda, \mathsf{M}_i\psi_\lambda\rangle = \mathrm{tr}(\mathsf{M}_i\rho), \tag{4.5}$$

where we use (A.5). By a similar argument with (A.7), $\mathrm{tr}\{\widetilde{\mathsf{Q}}_j\widetilde{\rho}\} = \mathrm{tr}\{\mathsf{Q}_j\rho\}$. In other words, we have $\widetilde{p}_i = p_i$ and $\widetilde{q}_j = q_j$ for any state of the form (4.3). Therefore, the corresponding Rényi entropies are related by

$$H_\alpha(\widetilde{\mathsf{M}}|\widetilde{\rho}) = H_\alpha(\mathsf{M}|\rho), \qquad H_\beta(\widetilde{\mathsf{Q}}|\widetilde{\rho}) = H_\beta(\mathsf{Q}|\rho). \tag{4.6}$$

By the calculations, we also have

$$f(\widetilde{\mathsf{M}}, \widetilde{\mathsf{Q}}|\widetilde{\rho}) = f(\mathsf{M}, \mathsf{Q}|\rho). \tag{4.7}$$

Formulae (4.6) and (4.7) hold for each pair $\widetilde{\rho}$ and ρ defined by (4.3) and (2.3). Since the latter is a general form of the density matrix on $\mathcal{H}$, relation (4.2) provides (4.1). □

The final step is to prove theorem 1. Following the previous argumentation, let us substitute M_i for E_i and N_j for G_j in the formulae of the appendix. So we will consider the two measurements $\{\widetilde{\mathsf{M}}_i\}$ and $\{\widetilde{\mathsf{N}}_j\}$ in the enlarged space $\widetilde{\mathcal{H}}$. The measurement $\{\widetilde{\mathsf{M}}_i\}$ is now projective due to the Naimark theorem. In general, the measurement $\{\widetilde{\mathsf{N}}_j\}$ is not projective. By relation (4.1), for PVM $\{\widetilde{\mathsf{M}}_i\}$ and POVM $\{\widetilde{\mathsf{N}}_j\}$ we have

$$H_\alpha(\widetilde{\mathsf{M}}|\widetilde{\omega}) + H_\beta(\widetilde{\mathsf{N}}|\widetilde{\omega}) \geqslant -2\ln f(\widetilde{\mathsf{M}}, \widetilde{\mathsf{N}}|\widetilde{\omega}). \tag{4.8}$$

Here, $1/\alpha + 1/\beta = 2$ and $\widetilde{\omega}$ denotes an arbitrary mixed state in the enlarged space $\widetilde{\mathcal{H}}$. Replacing $\widetilde{\mathsf{Q}}_j$ with $\widetilde{\mathsf{N}}_j$ in formulae (4.6) and (4.7), we get $H_\beta(\widetilde{\mathsf{N}}|\widetilde{\rho}) = H_\beta(\mathsf{N}|\rho)$ and $f(\widetilde{\mathsf{M}}, \widetilde{\mathsf{N}}|\widetilde{\rho}) = f(\mathsf{M}, \mathsf{N}|\rho)$ for an arbitrary matrix of the form (4.3). By these equations, from (4.8) we immediately obtain the main result formulated as theorem 1.

5. Discussion

Entropic uncertainty relations provide an alternative way to formulate Heisenberg's uncertainty principle [15]. It is clearly shown in [17] that they can give a more detailed characterization in many respects. Moreover, the method based on Riesz's theorem can be extended to the case of infinite-dimensional Hilbert space. The authors of [17] have also presented an example where a more informative relation is expressed via measures different from the Shannon entropy. We will discuss relations (2.5) and (3.2) in those respects that are not elucidated enough in the literature. Let us consider two observables $\mathsf{A} = \sum_i a_i\mathsf{P}_i$ and $\mathsf{B} = \sum_j b_j\mathsf{Q}_j$. The relation $\Delta\mathsf{A}\Delta\mathsf{B} \geqslant (1/2)|\langle\psi, [\mathsf{A}, \mathsf{B}]\psi\rangle|$ leads to a trivial zero bound, when ψ is an eigenstate of A or B. How does the bound (3.2) run in this case? Take a unit eigenvector φ_0 of A such that $\mathsf{P}_0\varphi_0 = \varphi_0$, $\mathsf{P}_i\varphi_0 = \mathbf{0}$ for $i \neq 0$. Then the function in (3.2) is rewritten as

$$f(\mathsf{P}, \mathsf{Q}|\varphi_0) = \max\{\|\mathsf{Q}_j\varphi_0\|^{-1}\,|\langle\varphi_0, \mathsf{Q}_j\varphi_0\rangle| : \|\mathsf{Q}_j\varphi_0\| \neq 0\}. \tag{5.1}$$

By the Schwarz inequality we have $f(\mathsf{P}, \mathsf{Q}|\varphi_0) \leqslant 1$, where the equality occurs if and only if $\mathsf{Q}_j\varphi_0 = c\,\varphi_0$ for some scalar c and label j, say $j = 1$. This scalar may be unit only, since each Q_j is a projector. If $\mathsf{Q}_1\varphi_0 = \varphi_0$ happens, there is a nonzero subspace $\mathcal{K} \ni \varphi_0$ such that the operators A and B act as commuting in $\mathcal{K}$. The density matrix $\varphi_0\varphi_0^\dagger$ also commutes with both A and B. According to the classical character of this case, $f(\mathsf{P}, \mathsf{Q}|\varphi_0) = 1$ and the entropic bound in (3.2) becomes trivial. Indeed, the eigenvector φ_0 is common for the observables and the outcomes of their measurement are fully certain. But if $\mathsf{Q}_j\varphi_0 \neq \varphi_0$ for all j, then $f(\mathsf{P}, \mathsf{Q}|\varphi_0) < 1$ and the right-hand side of (3.2) is strictly positive. That is, the entropic relation (3.2) actually gives a nontrivial bound. In contrast, we have $\langle\varphi_0, [\mathsf{A}, \mathsf{B}]\varphi_0\rangle = 0$ regardless of how B acts on φ_0. So we found an advantage over the customary formulation of the uncertainty principle.

Another point is that POVMs can directly be put in the entropic formulation. So it is interesting to consider the case where generalized measurements are actual. There exist two well-known approaches to discrimination between the non-identical pure states φ_+ and φ_-. It is handy to parameterize these states in a symmetric way as $\varphi_\pm = \cos\theta\, \mathbf{e}_0 \pm \sin\theta\, \mathbf{e}_1$, where $\{\mathbf{e}_0, \mathbf{e}_1\}$ is an orthonormal set. The overlap between φ_+ and φ_- is equal to $S = \cos 2\theta$ $(0 < 2\theta \leqslant \pi/2)$. In the Helstrom scheme [39], the optimal measurement is reached by two projectors $\mathsf{N}_+ = \mathsf{n}_+ \mathsf{n}_+^\dagger$ and $\mathsf{N}_- = \mathsf{n}_- \mathsf{n}_-^\dagger$ with $\mathsf{n}_\pm = (\mathbf{e}_0 \pm \mathbf{e}_1)/\sqrt{2}$. The method distinguishes between φ_+ and φ_- with the probability of correct answer equal to $(1 + \sin 2\theta)/2$. Another measurement is built for unambiguous discrimination proposed in [40–42]. This approach sometimes gives an inconclusive answer, but never makes an error of misidentification. Putting two normalized vectors

$$\mathsf{m}_\pm = \sin\theta\, \mathbf{e}_0 \pm \cos\theta\, \mathbf{e}_1, \qquad \langle \mathsf{m}_+, \varphi_- \rangle = 0, \qquad \langle \mathsf{m}_-, \varphi_+ \rangle = 0, \tag{5.2}$$

we have $\mathsf{M}_\pm = (1+S)^{-1} \mathsf{m}_\pm \mathsf{m}_\pm^\dagger$ and $\mathsf{M}_0 = 2S(1+S)^{-1} \mathbf{e}_0\, \mathbf{e}_0^\dagger$. The probability of an inconclusive answer is $\langle \varphi_\pm, \mathsf{M}_0 \varphi_\pm \rangle = S$. If both the POVMs have only one-rank elements $\mathsf{M}_i = \mu_i \mathsf{m}_i \mathsf{m}_i^\dagger$ and $\mathsf{N}_j = \nu_j \mathsf{n}_j \mathsf{n}_j^\dagger$, then relation (2.5) is reduced to

$$H_\alpha(\mathsf{M}|\rho) + H_\beta(\mathsf{N}|\rho) \geqslant -\ln\{\max f(i,j)^2\}, \tag{5.3}$$

where $f(i,j)^2 = \mu_i \nu_j\, |\langle \mathsf{m}_i, \mathsf{n}_j \rangle|^2$. In our example, we obtain that the maximum is either $f(+,+)^2 = f(-,-)^2 = (1 + \sin 2\theta)/(2(1+S))$ or $f(0,\pm)^2 = S/(1+S)$. The former holds for $0 \leqslant S \leqslant 4/5$, the latter holds for $4/5 \leqslant S < 1$. For simulation of classical information, orthogonal states are fully sufficient. In the limit $S \to 0$ $(2\theta \to \pi/2)$, the operator M_0 is zero and POVM $\{\mathsf{M}_i\}$ actually coincides with $\{\mathsf{N}_j\}$. Here, relation (5.3) becomes trivial, i.e. $H_\alpha(\mathsf{M}) + H_\beta(\mathsf{N}) \geqslant 0$. For two almost identical states, however, we have $H_\alpha(\mathsf{M}) + H_\beta(\mathsf{N}) \geqslant \ln 2$. In fact, this bound is valid for all values $0 < 2\theta \leqslant \pi/4$. So we observe a principal distinction between minimum error discrimination and unambiguous discrimination that are like in the classical limit solely. It is conventional that an incompleteness of the customary formulation is induced by the dependence of the bound on a quantum state [15, 17]. To sum up we see that such a dependence is not crucial for entropic uncertainty relations.

Acknowledgments

This research was supported in part by the Ministry of Education and Science of the Russian Federation under grant nos 2.2.1.1/1483, 2.1.1/1539. The present author is grateful to anonymous referees for constructive criticism.

Appendix. Naimark's extension and related questions

Naimark proved that each generalized resolution of identity can be realized as an orthogonal resolution of the identity for the enlarged space $\widetilde{\mathcal{H}}$ which contains $\mathcal{H}$ as a subspace [35, 38]. Let $\{\mathsf{E}_i\}$ be a set of positive semidefinite matrices satisfying

$$\sum_i \mathsf{E}_i = \mathbf{I}_{\mathcal{H}}, \tag{A.1}$$

where $\mathbf{I}_{\mathcal{H}}$ is the identity operator in the space $\mathcal{H}$. Let us define a space $\widetilde{\mathcal{H}} := \mathcal{H} \oplus \mathcal{L}$, where $\mathcal{L}$ is a space of proper dimensionality. As is shown in [37], one can build partitioned matrices of the form

$$\widetilde{\mathsf{E}}_i := \begin{bmatrix} \mathsf{E}_i & \mathsf{R}_i \\ \mathsf{R}_i^\dagger & \mathsf{L}_i \end{bmatrix}, \tag{A.2}$$

so that the $\widetilde{\mathsf{E}}_i$'s are orthogonal projections in the enlarged space $\widetilde{\mathcal{H}}$ and

$$\sum_i \widetilde{\mathsf{E}}_i = \widetilde{\mathbf{I}}. \tag{A.3}$$

In (A.2) the orders of submatrices R_i and L_i should be clear from the context. An arbitrary vector in the enlarged space is represented by the column $\widetilde{\mathsf{u}}$ such that $\widetilde{\mathsf{u}}^T = [\mathsf{u}^T\ \mathsf{z}^T]$ with $\mathsf{u} \in \mathcal{H}$ and $\mathsf{z} \in \mathcal{L}$. The entries of this column are components of $\widetilde{\mathsf{u}}$ with respect to the orthonormal basis in $\widetilde{\mathcal{H}}$ obtained by extension of the initial basis in $\mathcal{H}$. To each $\psi \in \mathcal{H}$ assign the vector $\widetilde{\psi} \in \widetilde{\mathcal{H}}$ defined by

$$\widetilde{\psi} := \begin{bmatrix} \psi \\ \mathbf{0} \end{bmatrix}. \tag{A.4}$$

Following the rules of block multiplication, we have

$$\langle \widetilde{\psi}, \widetilde{\mathsf{E}}_j \widetilde{\psi} \rangle \equiv \widetilde{\psi}^\dagger \widetilde{\mathsf{E}}_i \widetilde{\psi} = \begin{bmatrix} \psi^\dagger & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathsf{E}_i \psi \\ \mathsf{R}_i^\dagger \psi \end{bmatrix} = \psi^\dagger \mathsf{E}_i \psi \equiv \langle \psi, \mathsf{E}_j \psi \rangle. \tag{A.5}$$

That is, the probability of getting outcome i is not changed under the extension.

Let $\{\mathsf{G}_j\}$ be another resolution of the identity for the space $\mathcal{H}$. To each G_j assign the operator $\widetilde{\mathsf{G}}_j$ acting on the space $\widetilde{\mathcal{H}}$. In the matrix representation, we define [37]

$$\widetilde{\mathsf{G}}_1 := \begin{bmatrix} \mathsf{G}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{\mathcal{L}} \end{bmatrix}, \qquad \widetilde{\mathsf{G}}_j := \begin{bmatrix} \mathsf{G}_j & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \qquad (j \neq 1). \tag{A.6}$$

Here, the identity matrix $\mathbf{I}_{\mathcal{L}}$ of proper order shows the action of the identity in the subspace $\mathcal{L}$. It is easy to check that the set $\{\widetilde{\mathsf{G}}_j\}$ is a resolution of the identity for the space $\widetilde{\mathcal{H}}$ [37]. In addition, if the resolution $\{\mathsf{G}_j\}$ is orthogonal, then the resolution $\{\widetilde{\mathsf{G}}_j\}$ is also orthogonal. Further, for any state of the form (A.4) we have

$$\langle \widetilde{\psi}, \widetilde{\mathsf{G}}_j \widetilde{\psi} \rangle \equiv \widetilde{\psi}^\dagger \widetilde{\mathsf{G}}_j \widetilde{\psi} = \psi^\dagger \mathsf{G}_j \psi \equiv \langle \psi, \mathsf{G}_j \psi \rangle. \tag{A.7}$$

So, for the second measurement the probability of getting outcome j is also not changed under the extension. To sum up, we can say the following. Starting with the two POVM measurements $\{\mathsf{E}_i\}$ and $\{\mathsf{G}_j\}$, we have built the two measurements $\{\widetilde{\mathsf{E}}_i\}$ and $\{\widetilde{\mathsf{G}}_j\}$ in the enlarged space $\widetilde{\mathcal{H}}$. But the first measurement $\{\widetilde{\mathsf{E}}_i\}$ is now projective.

References

[1] Heisenberg W 1927 *Z. Phys.* **43** 172
[2] Robertson H P 1929 *Phys. Rev.* **34** 163
[3] Hall M J W 1999 *Phys. Rev.* A **59** 2602
[4] Ozawa M 2004 *Ann. Phys.* **311** 350
[5] Busch P, Heinonen T and Lahti P J 2007 *Phys. Rep.* **452** 155
[6] Hall M J W 2004 *Phys. Rev.* A **69** 052113
[7] Ozawa M 2004 *Phys. Lett.* A **320** 367
[8] Garcia Diaz R, Romero J L, Björk G and Bourennane M 2005 *New J. Phys.* **7** 256
[9] Petz D 2007 *Rep. Math. Phys.* **59** 209
[10] Ivanovic I D 1981 *J. Phys. A: Math. Gen.* **14** 3241
[11] Wehrl A 1978 *Rev. Mod. Phys.* **50** 221
[12] Hirschman I I 1957 *Am. J. Math.* **79** 152
[13] Beckner W 1975 *Ann. Math.* **102** 159
[14] Bialynicki-Birula I and Mycielski J 1975 *Commun. Math. Phys.* **44** 129
[15] Deutsch D 1983 *Phys. Rev. Lett.* **50** 631
[16] Kraus K 1987 *Phys. Rev.* D **35** 3070
[17] Maassen H and Uffink J B M 1988 *Phys. Rev. Lett.* **60** 1103
[18] Krishna M and Parthasarathy K R 2002 *Sankhya, Ser.* A **64** 842

[19] Ivanovic I D 1992 *J. Phys. A: Math. Gen.* **25** L363
[20] Wehner S and Winter A 2008 *J. Math. Phys.* **49** 062105
[21] de Vicente J I and Sánchez-Ruiz J 2008 *Phys. Rev.* A **77** 042110
[22] Rényi A 1961 *Proceedings of the 4th Berkeley Symposium on Mathematical Statistics and Probability* (Berkeley, Los Angeles: University of California Press) p 547
[23] Jizba P and Arimitsu T 2004 *Ann. Phys.* **312** 17
[24] Życzkowski K 2003 *Open Sys. Inf. Dyn.* **10** 297
[25] Zachos C K 2007 *J. Phys. A: Math. Theor.* **40** F407
[26] Larsen U 1990 *J. Phys. A: Math. Gen.* **23** 1041
[27] Bialynicki-Birula I 2006 *Phys. Rev.* A **74** 052101
[28] Rajagopal A K 1995 *Phys. Lett.* A **205** 32
[29] Majernik V and Majernikova E 2001 *Rep. Math. Phys.* **47** 381
[30] Wilk G and Włodarczyk Z 2009 *Phys. Rev.* A **79** 062108
[31] Skála L and Kapsa V 2008 *J. Phys. A: Math. Theor.* **41** 265302
[32] Gibilisco P, Hiai F and Petz D 2009 *IEEE Trans. Inf. Theory* **55** 439
[33] Schürmann T and Hoffmann I 2009 *Found. Phys.* **39** 958
[34] Peres A 1993 *Quantum Theory: Concepts and Methods* (Dordrecht: Kluwer)
[35] Holevo A S 1982 *Probabilistic and Statistical Aspects of Quantum Theory (North-Holland Series in Statistics and Probability* vol 1*)* (Amsterdam: North-Holland)
[36] Hardy G H, Littlewood J E and Polya G 1934 *Inequalities* (London: Cambridge University Press)
[37] Parthasarathy K R 1999 *Infin. Dimens. Anal. Quantum Probab. Relat. Top.* **2** 557
[38] Naimark M A 1943 *C R (Dokl.) Acad. Sci.* URSS **41** 359
[39] Helstrom C W 1976 *Quantum Detection and Estimation Theory* (New York: Academic)
[40] Ivanovic I D 1987 *Phys. Lett.* A **123** 257
[41] Dieks D 1988 *Phys. Lett.* A **126** 303
[42] Peres A 1988 *Phys. Lett.* A **128** 19