

ENTROPIC UPPER BOUND FOR BAYES RISK IN THE QUANTUM CASE

BY

RAFAŁ WIECZOREK (ŁÓDŹ) AND HANNA PODSEDKOWSKA (ŁÓDŹ)

Abstract. The entropic upper bound for Bayes risk in a general quantum case is presented. We obtained generalization of the entropic lower bound for probability of detection. Our result indicates upper bound for Bayes risk (in a particular case of loss function – for probability of detection) in a pretty general setting of an arbitrary finite von Neumann algebra. It is also shown under which condition the indicated upper bound is achieved.

2010 AMS Mathematics Subject Classification: Primary: 81P15; Secondary: 81P50.

Key words and phrases: Bayes risk, probability of detection, quantum measurement, entropy.

1. INTRODUCTION

One of the branches of quantum information is the theory of statistics decisions and optimal measurement. It motivates to study the Bayes risk and probability of detection of states of the physical system. Many results in the mentioned field were obtained for the quantum dynamical system represented by algebra of all bounded operators with canonical trace (sometimes even on a finite-dimensional Hilbert space) by [1]–[3], [11], and [12].

In this paper we present more general results received for an arbitrary von Neumann algebra with finite faithful normal trace τ . We employ the definition of the Segal entropy of states from a predual of algebra.

2. BASIC NOTIONS

2.1. Concept of entropy. Let \mathfrak{M} be a semi-finite von Neumann algebra of operators acting on a Hilbert space \mathcal{H} with a normal semi-finite faithful trace τ , identity $\mathbb{1}$, and predual \mathfrak{M}_* . By \mathfrak{M}_*^+ we shall denote the set of positive functionals in \mathfrak{M}_* . These functionals will sometimes be referred to as (non-normalized) states. The set of normalized states, i.e. the elements $\rho \in \mathfrak{M}_*^+$ such that $\rho(\mathbb{1}) = \|\rho\| = 1$, will be denoted by \mathfrak{S} .

The algebra of measurable operators $\tilde{\mathfrak{M}}$ is defined as a topological *-algebra of densely defined closed operators on \mathcal{H} affiliated with \mathfrak{M} with strong addition and strong multiplication.

For each $\rho \in \mathfrak{M}_*$ there is a measurable operator h such that

$$\rho(x) = \tau(xh) = \tau(hx), \quad x \in \mathfrak{M}.$$

The space of all such operators is denoted by $L^1(\mathfrak{M}, \tau)$ and the correspondence above is one-to-one and isometric, where the norm of $L^1(\mathfrak{M}, \tau)$, denoted by $\|\cdot\|_1$, is defined as

$$\|h\|_1 = \tau(|h|), \quad h \in L^1(\mathfrak{M}, \tau).$$

Moreover, self-adjoint operators in $L^1(\mathfrak{M}, \tau)$ correspond to Hermitian functionals in \mathfrak{M}_* , and positive operators in $L^1(\mathfrak{M}, \tau)$ to the states in \mathfrak{M}_* .

For a state ρ the corresponding element in $L^1(\mathfrak{M}, \tau)$ will be denoted by $\hat{\rho}$ and called the *density matrix* of ρ , thus

$$\rho(x) = \tau(x\hat{\rho}) = \tau(\hat{\rho}x), \quad x \in \mathfrak{M}.$$

In particular,

$$\tau(\hat{\rho}) = \rho(\mathbb{1}).$$

Observe that for a finite τ , we have $\mathfrak{M} \subset L^1(\mathfrak{M}, \tau)$.

In the case of the full algebra $\mathbb{B}(\mathcal{H})$, a well-established concept of entropy goes back to J. von Neumann who defined the entropy of a state ρ as

$$S(\rho) = -\operatorname{tr} \hat{\rho} \log \hat{\rho},$$

where $\hat{\rho}$ is a positive trace operator of the trace one.

Unfortunately, when we deal with an arbitrary von Neumann algebra, a satisfactory general definition of entropy is lacking. Thus we employ the Segal entropy (up to the minus sign) of $\rho \in \mathfrak{M}_*$, denoted by $H(\rho)$ and defined as

$$H(\rho) = \tau(\hat{\rho} \log \hat{\rho}),$$

i.e. for the spectral representation of $\hat{\rho}$,

$$\hat{\rho} = \int_0^\infty \lambda e(d\lambda),$$

we have

$$H(\rho) = \int_0^\infty \lambda \log \lambda \tau(e(d\lambda)).$$

Although for a semi-finite algebra \mathfrak{M} this definition is a straightforward generalization of the von Neumann idea, the reasoning which substantiates Segal entropy properties needs a different setup from the one used in the case of $\mathfrak{M} = \mathbb{B}(\mathcal{H})$.

REMARK 2.1. Despite being a seemingly straightforward generalization of von Neumann entropy, the Segal definition exhibits fundamental differences in many respects from that of von Neumann. For example, while the density operator in the von Neumann definition is a trace-class operator, and thus has a discrete spectrum with the eigenvalues summing up to one, this is not the case in the Segal definition. Furthermore, the von Neumann entropy of a state is nonnegative (which is a consequence of the above property of the density operator), while the Segal entropy of a state need not be such. In addition, there are also some technical problems while dealing with a semi-finite trace. For these reasons, we shall consider the case of a finite von Neumann algebra and adopt a definition of entropy more in the spirit of the classical Boltzmann–Gibbs notion, where for a density function f on a probability space $(\Omega, \mathbb{F}, \mu)$, its entropy is defined as

$$H(f) = \int_{\Omega} f(\log f) d\mu.$$

As will be seen, our definition, which is just that of Segal up to a minus sign, assigns a finite nonnegative entropy to a state, and more generally, for each non-normalized state in \mathfrak{M}_*^+ with bounded density, its entropy is finite.

It should be noted that some fundamental investigations concerning entropy and related notions in the above setup were carried out in [9].

REMARK 2.2 (see [6]). *For a finite algebra \mathfrak{M} (this is the case of our interest) with faithful finite normal trace τ , $\tau(\mathbb{1}) = 1$, for each $\rho \in \mathfrak{S}$, $H(\rho) \geq 0$, and for $\hat{\rho} \in \mathfrak{M}$, $H(\rho)$ is also bounded from above.*

Indeed, since $\lambda \log \lambda \geq \lambda - 1$, we have

$$\begin{aligned} (2.1) \quad H(\rho) &= \tau\left(\int_0^\infty \lambda \log \lambda e(d\lambda)\right) = \int_0^\infty \lambda \log \lambda \tau(e(d\lambda)) \geq \int_0^\infty (\lambda - 1) \tau(e(d\lambda)) \\ &= \int_0^\infty \lambda \tau(e(d\lambda)) - \int_0^\infty \tau(e(d\lambda)) = \tau(\hat{\rho}) - \tau(\mathbb{1}) = \rho(\mathbb{1}) - 1, \end{aligned}$$

showing that the entropy is bounded from below, and in particular, it is nonnegative for states. Moreover, since $\hat{\rho}$ is bounded, its spectrum is a bounded set; thus, the function $\lambda \mapsto \lambda \log \lambda$ is bounded on the spectrum, which implies that the entropy is bounded from above.

REMARK 2.3. *In the classical quantum case, that is, for $\mathfrak{M} = \mathbb{B}(\mathcal{H})$, the practical Klein’s inequality holds (see [7] and [10]). The analogue of this inequality is given by the formula*

$$(2.2) \quad \tau(a \log a - a \log b) \geq 0 \quad \text{for } a, b \in \mathfrak{M}_+, \tau(a) = \tau(b) = 1 \text{ and } \text{supp } a \leq \text{supp } b,$$

and it was proved by Umegaki [9] in the case of an arbitrary von Neumann algebra \mathfrak{M} with finite faithful normal trace τ . In addition, it was proved (see [4], Theorem 2.1.2(i)) that the equality in (2.2) holds if and only if $a = b$ for $\mathfrak{M} = \mathbb{B}(\mathbb{C}^d)$.

REMARK 2.4. *Moreover, apart from the very practical Klein's inequality, which holds for the Segal entropy, we have also the following inequality (see [6], Proposition 1). Let $a, b \in \mathfrak{M}$ be such that $0 \leq a \leq b$. Then*

$$(2.3) \quad \tau(a \log b - a \log a) \geq 0,$$

with equality if and only if $ab = ba = a^2$. Moreover, $a \log a$ and $a \log b$ are bounded (belong to \mathfrak{M}), and the numbers $\tau(a \log b)$ and $\tau(a \log a)$ are finite.

(This remark is presented and proved with details in [6], but we remind its main idea to make our reasoning clearer.)

Proof. Since

$$0 \leq a \leq b,$$

we have

$$0 \leq (\log b)a(\log b) \leq (\log b)b(\log b) = b \log^2 b.$$

The operator on the right-hand side of the inequality above is bounded (belongs to \mathfrak{M}), hence $(\log b)a(\log b)$ is also bounded (belongs to \mathfrak{M}). Moreover,

$$(\log b)a(\log b) = (a^{1/2} \log b)^* a^{1/2} \log b,$$

thus $a^{1/2} \log b$ is bounded (belongs to \mathfrak{M}). Consequently, $a^{1/2}(\log b - \log a)$ and $a^{1/2}$ belong to \mathfrak{M} , so from the properties of trace we obtain

$$(2.4) \quad \tau(a(\log b - \log a)) = \tau\left(a^{1/2}\left(a^{1/2}(\log b - \log a)\right)\right) = \tau\left(a^{1/2}(\log b - \log a)a^{1/2}\right).$$

Since the logarithm is an operator monotone function, we have

$$\log b - \log a \geq 0,$$

yielding

$$a^{1/2}(\log b - \log a)a^{1/2} \geq 0,$$

and finally, by equation (2.4),

$$0 \leq \tau\left(a^{1/2}(\log b - \log a)a^{1/2}\right) = \tau(a(\log b - \log a)).$$

The assumption

$$(2.5) \quad \tau(a \log b - a \log a) = 0$$

gives

$$a^{1/2}(\log b - \log a)a^{1/2} = 0,$$

yielding

$$a(\log a - \log b) = 0,$$

i.e.,

$$a \log a = a \log b.$$

Taking adjoints, we get

$$a \log a = (\log b)a.$$

In particular, $\log b$ commutes with a , leaves the range of a invariant and coincides with $\log a$ on the range of a . Thus, on the range of a we have

$$a| \text{Range } a = e^{\log a} | \text{Range } a = e^{\log b} | \text{Range } a = b | \text{Range } a,$$

which is equivalent to the equalities

$$ab = ba = a^2.$$

Conversely, assume that the equality above holds true. Then, a and b commute, so after taking logarithms of both sides, we get

$$2 \log a = \log a + \log b,$$

that is

$$\log a = \log b,$$

which implies the equality

$$a \log a = a \log b,$$

and thus equation (2.5). ■

2.2. Bayes risk in the quantum case. We are given a von Neumann algebra \mathfrak{M} describing the (bounded) observables of a physical system. Let ρ_1, ρ_2, \dots be normal states from \mathfrak{M}_* . We assume that the physical system can be in state ρ_i with *a priori* probability $\pi_i, i = 1, 2, \dots$, where $\pi = (\pi_1, \pi_2, \dots)$ is a probability distribution. On the system we perform a *measurement* (called also a *strategy*) M by which we mean a sequence (M_1, M_2, \dots) of positive operators from \mathfrak{M} such that

$$\sum_{i=1}^{\infty} M_i = \mathbb{1},$$

where the series is convergent in the weak operator topology on \mathfrak{M} .

We want to find, in an optimal way, the state in which the system really is.

If we receive an outcome M_j , we choose the state ρ_i . The probability that the true state is ρ_i when the measurement gives the result M_j is determined by $\rho_i(M_j)$. Thus $\rho_i(M_j)$ is the probability of guessing the state ρ_i correctly. If our guess is ρ_j

while the true one is ρ_i , then we pay a penalty $L(i, j)$. The function $L : \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{R}$ is called a *loss function*. The risk function is defined by the formula

$$R_M(i) = \sum_{j=1}^{\infty} L(i, j) \rho_i(M_j).$$

The expectation of the risk function is called the *Bayes risk* and denoted by $r(M, \pi)$, i.e.

$$r(M, \pi) = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \pi_i L(i, j) \rho_i(M_j).$$

Consider the concrete loss function of the form

$$L(i, j) = 1 - \delta_{ij}.$$

Then we have

$$r(M, \pi) = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \pi_i (1 - \delta_{ij}) \rho_i(M_j) = 1 - \sum_{i=1}^{\infty} \pi_i \rho_i(M_j).$$

In this case, minimizing the Bayes risk is equivalent to maximizing the expression

$$\sum_{i=1}^{\infty} \pi_i \rho_i(M_j)$$

which is the probability of the correct guess while performing the measurement M , and is called the *probability of detection*. We shall denote this probability by $\mathbb{P}_D(M)$.

3. ENTROPIC BOUND

For an arbitrary loss function we have no guarantee of the existence of an optimal measurement, e.g. the one which minimizes the Bayes risk. However, under the assumptions presented in the following theorem, we can consider such an optimal measurement.

THEOREM 3.1 ([5], Theorem 8). *Let L be a loss function that satisfies the following conditions:*

(i) *there are $a_i \geq 0$ such that for each i we have $|L(i, j)| \leq a_i$, $j = 1, 2, \dots$, and $\sum_{i=1}^{\infty} \pi_i a_i < \infty$;*

(ii) *for each i there exists $\lim_{j \rightarrow \infty} L(i, j) = b_i$ such that for some j_0 we have $L(i, j_0) \leq b_i$ for all $i = 1, 2, \dots$.*

Then, there exists an optimal measurement.

In [3] we can find more information on the existence of an optimal measurement. From now on, we will assume that the loss function satisfies the conditions (i) and (ii) of Theorem 3.1. The following theorem will be very helpful for further consideration.

THEOREM 3.2 ([2], Theorem II.2.2). *We have the relation*

$$(3.1) \quad \min_M r(M, \pi) = \max\{\psi(\mathbb{1}) : \psi \in \mathfrak{M}_*, \psi \leq \varphi_j, j = 1, 2, \dots\},$$

where $\varphi_j = \sum_{i=1}^\infty \pi_i L(i, j) \rho_i, j = 1, 2, \dots$. Then the following assertions are equivalent:

(i) *The measurement $M = (M_i)$ is optimal for r , and $\psi \in \mathfrak{M}_*$ maximizes the right-hand side of (3.1).*

(ii) *$\psi \leq \varphi_j, j = 1, 2, \dots$, and $\psi = \sum_{j=1}^\infty \varphi_j M_j = \sum_{j=1}^\infty M_j \varphi_j$.*

Let $c = (c_i)$ be a sequence such that $L(i, j) \leq c_i$ for all i, j and the sum $\sum_{i=1}^\infty \pi_i c_i$ is convergent. Define the functional

$$r_c(M, \pi) := \sum_{j=1}^\infty \varphi_j^c(M_j) = \sum_{i=1}^\infty \pi_i c_i - r(M, \pi),$$

where $\varphi_j^c = \sum_{i=1}^\infty \pi_i (c_i - L(i, j)) \rho_i, j = 1, 2, \dots$. Minimizing the Bayes risk is equivalent to maximizing the functional r_c with positive functionals φ_j^c .

The next result is a simple consequence of Theorem 3.2.

THEOREM 3.3. *We have the relation*

$$(3.2) \quad \max_M r_c(M, \pi) = \min\{\varphi(\mathbb{1}) : \varphi \in \mathfrak{M}_*, \varphi \geq \varphi_j^c, j = 1, 2, \dots\}.$$

Then the following assertions are equivalent:

(i) *The measurement $M = (M_i)$ is optimal for r_c , and $\varphi \in \mathfrak{M}_*$ minimizes the right-hand side of (3.2).*

(ii) *$\varphi \geq \varphi_j^c, j = 1, 2, \dots$, and $\varphi = \sum_{j=1}^\infty \varphi_j^c M_j = \sum_{j=1}^\infty M_j \varphi_j^c$.*

Proof. From (3.1) we obtain

$$(3.3) \quad \min_M r(M, \pi) = \max\{\psi(\mathbb{1}) : \psi \leq \sum_i \pi_i L(i, j) \rho_i, j = 1, 2, \dots\}.$$

Denote by φ the functional $\sum_i \pi_i c_i \rho_i - \psi$. Then the above equality takes the form

$$\min_M r(M, \pi) = \sum_i c_i \pi_i - \min\{\varphi(\mathbb{1}) : \varphi_j^c \leq \varphi, j = 1, 2, \dots\}.$$

Consequently,

$$(3.4) \quad \max_M r_c(M, \pi) = \min\{\varphi(\mathbb{1}) : \varphi_j^c \leq \varphi, j = 1, 2, \dots\}.$$

(i)⇒(ii). Let φ be an optimal functional in (3.2). Then $\varphi \geq \varphi_j^c, j = 1, 2, \dots$. The functional $\psi = \sum_i \pi_i c_i \rho_i - \varphi$ maximizes the right-hand side of (3.1). Then from Theorem 3.2(ii) we have

$$\psi = \sum_j \varphi_j M_j,$$

where $\varphi_j = \sum_i \pi_i L(i, j) \rho_i$ and (M_j) is an optimal measurement. Therefore, the optimal functional φ in (3.4) is of the form $\sum_i \varphi_i^c M_i$.

(ii)⇒(i). Let φ be such that $\varphi \geq \varphi_j^c, j = 1, 2, \dots$, and $\varphi = \sum_{j=1}^\infty \varphi_j^c M_j = \sum_{j=1}^\infty M_j \varphi_j^c$. The functional $\psi = \sum_i \pi_i c_i \rho_i - \varphi$ satisfies the conditions $\psi \leq \varphi_j, j = 1, 2, \dots$, and $\psi = \sum_{j=1}^\infty \varphi_j M_j = \sum_{j=1}^\infty M_j \varphi_j$. Then, by Theorem 3.2, the measurement $M = (M_i)$ is optimal for r (also for r_c) and $\psi \in \mathfrak{M}_*$ maximizes the right-hand side of (3.1). Therefore, $\varphi \in \mathfrak{M}_*$ and φ minimizes the right-hand side of (3.2). ■

In the rest of this article we assume that \mathfrak{M} is a finite von Neumann algebra with faithful finite normal trace $\tau, \tau(\mathbb{1}) = 1$, and $\hat{\rho}_1, \hat{\rho}_2, \dots \in \mathfrak{M}$. Denote by $\|\cdot\|_\infty$ the operator norm in \mathfrak{M} .

THEOREM 3.4 (Main theorem). *Let the series $\sum_{ij} \pi_i (c_i - L(i, j)) \|\hat{\rho}_i\|_\infty$ be convergent and let us put $a_c = \sum_{ij} \pi_i (c_i - L(i, j))$. Then we have the estimate*

$$(3.5) \quad \min_M r(M, \pi) \leq \sum_i \pi_i c_i - 2^{\frac{1}{a_c}(\sum_i H(\varphi_i^c)) - H(\frac{1}{a_c}(\sum_i \varphi_i^c))}.$$

Proof. Note that the convergence of the series $\sum_{ij} \pi_i (c_i - L(i, j)) \|\hat{\rho}_i\|_\infty$ implies that $\hat{\varphi}_1^c, \hat{\varphi}_2^c, \dots \in \mathfrak{M}, \sum_i \hat{\varphi}_i^c \in \mathfrak{M}$ and also the convergence of the series $\sum_{ij} \pi_i (c_i - L(i, j))$. Let φ' be an optimal functional from the right-hand side of (3.2). Then

$$\hat{\varphi}' = \sum_i \hat{\varphi}_i^c M_i,$$

where $M = (M_i)$ is an optimal measurement. The series $\sum_{ij} \pi_i (c_i - L(i, j)) \|\hat{\rho}_i\|_\infty$ is convergent, so the series $\sum_i \hat{\varphi}_i^c M_i$ is the Cauchy series for the norm $\|\cdot\|_\infty$. $\hat{\varphi}_i^c M_i \in \mathfrak{M}$, therefore $\hat{\varphi}' \in \mathfrak{M}$. Observe that $a_c = \tau(\sum_i \hat{\varphi}_i^c)$. By Remark 2.4, the operator $\hat{\varphi}_i^c \log \hat{\varphi}'$ is bounded because $\hat{\varphi}' \geq \hat{\varphi}_i^c$. The operator $(\sum_i \hat{\varphi}_i^c) \log \hat{\varphi}'$ is the pointwise limit of the sequence of operators $(\sum_{i=1}^n \hat{\varphi}_i^c) \log \hat{\varphi}'$, so it is bounded. On the other hand, using the inequality (2.3), we obtain

$$\tau(\hat{\varphi}_i^c \log \hat{\varphi}') \geq \tau(\hat{\varphi}_i^c \log \hat{\varphi}_i^c).$$

In summary, we have the convergence of the series $\sum_i \tau(\hat{\varphi}_i^c \log \hat{\varphi}')$ and

$$(3.6) \quad \sum_i \tau(\hat{\varphi}_i^c (\log \hat{\varphi}' - \log \hat{\varphi}_i^c)) \geq 0.$$

Thus there are two cases. First, the series $\sum_i \tau(\hat{\varphi}_i^c(\log \hat{\varphi}' - \log \hat{\varphi}_i^c))$ is divergent, so

$$\sum_i \tau(\hat{\varphi}_i^c(\log \hat{\varphi}' - \log \hat{\varphi}_i^c)) = \infty.$$

Then $\sum_i \tau(\hat{\varphi}_i^c \log \hat{\varphi}_i^c) = -\infty$ and the inequality (3.5) is true because it takes the form

$$\min_M r(M, \pi) \leq \sum_i \pi_i c_i.$$

Second, since the series $\sum_i \tau(\hat{\varphi}_i^c(\log \hat{\varphi}' - \log \hat{\varphi}_i^c))$ is convergent, so is the series $\sum_i \tau(\hat{\varphi}_i^c \log \hat{\varphi}_i^c)$. Using the inequality (3.6), we have the estimate

$$\begin{aligned} (3.7) \quad \log \max_M r_c(M, \pi) &\geq \log \tau(\hat{\varphi}') - \frac{1}{a_c} \sum_i \tau(\hat{\varphi}_i^c(\log \hat{\varphi}' - \log \hat{\varphi}_i^c)) \\ &= -\frac{1}{a_c} \sum_i \tau \left(\hat{\varphi}_i^c \left(\log \frac{\hat{\varphi}'}{\tau(\hat{\varphi}')} - \log \hat{\varphi}_i^c \right) \right) \\ &= -\frac{1}{a_c} \tau \left(\left(\sum_i \hat{\varphi}_i^c \right) \log \frac{\hat{\varphi}'}{\tau(\hat{\varphi}')} \right) + \frac{1}{a_c} \sum_i \tau(\hat{\varphi}_i^c \log \hat{\varphi}_i^c). \end{aligned}$$

With our notation, it is obvious that the assumptions of Klein's inequality (2.2) hold and

$$(3.8) \quad \tau \left(\frac{\sum_i \hat{\varphi}_i^c}{a_c} \log \frac{\sum_i \hat{\varphi}_i^c}{a_c} \right) \geq \tau \left(\frac{\sum_i \hat{\varphi}_i^c}{a_c} \log \frac{\hat{\varphi}'}{\tau(\hat{\varphi}')} \right),$$

therefore

$$\log \max_M r_c(M, \pi) \geq -\tau \left(\frac{\sum_i \hat{\varphi}_i^c}{a_c} \log \frac{\sum_i \hat{\varphi}_i^c}{a_c} \right) + \frac{1}{a_c} \sum_i \tau(\hat{\varphi}_i^c \log \hat{\varphi}_i^c) =: A.$$

Consequently, we obtain the inequality

$$\min_M r(M, \pi) \leq \sum_i \pi_i c_i - 2^A. \quad \blacksquare$$

In the proof of Theorem 3.4 we used the idea of the proof of Lemma 2 in [8].

COROLLARY 3.1. *Assume that the series $\sum_{i=1}^\infty \pi_i \|\rho_i\|_\infty$ is convergent. Then for the probability of detection we obtain*

$$(3.9) \quad \max_M \mathbb{P}_D(M) \geq 2^{\sum_i \pi_i \log \pi_i - H(\sum_i \pi_i \rho_i) + \sum_i \pi_i H(\rho_i)}.$$

Proof. In Theorem 3.4, let us consider the concrete loss function of the form $L(i, j) = 1 - \delta_{ij}$ and $c = (1, 1, \dots)$. Then we have $r_c(M, \pi) = \sum_i \pi_i \rho_i(M_i)$.

This expression is the probability of detection. From the inequality (3.5) we obtain

$$\max_M \mathbb{P}_D(M) \geq 2^{\sum_i H(\pi_i \rho_i) - H(\sum_i \pi_i \rho_i)}.$$

We have

$$\sum_i H(\pi_i \rho_i) = \sum_i \tau(\pi_i \rho_i \log \pi_i \rho_i) = \sum_i [\pi_i \tau(\rho_i \log \rho_i) + \pi_i \log \pi_i].$$

By Remark 2.2, the inequality $\tau(\rho_i \log \rho_i) \geq 0$ holds. On the other hand, from the inequality $\log x \leq x - 1$ we obtain

$$\tau(\rho_i \log \rho_i) = \int_0^{\|\rho_i\|_\infty} \lambda \log \lambda \tau(e_i(d\lambda)) \leq \log \|\rho_i\|_\infty \leq \|\rho_i\|_\infty - 1.$$

Therefore, $0 \leq \pi_i \tau(\rho_i \log \rho_i) \leq \pi_i \|\rho_i\|_\infty - \pi_i$. By the assumption, the series $\sum_i (\pi_i \|\rho_i\|_\infty - \pi_i)$ is convergent, so the series $\sum_i \pi_i \tau(\rho_i \log \rho_i)$ is also convergent. Consequently,

$$\sum_i [\pi_i \tau(\rho_i \log \rho_i) + \pi_i \log \pi_i] = \sum_i \pi_i \tau(\rho_i \log \rho_i) + \sum_i \pi_i \log \pi_i$$

and

$$\sum_i H(\pi_i \rho_i) - H(\sum_i \pi_i \rho_i) = \sum_i \pi_i H(\rho_i) + \sum_i \pi_i \log \pi_i - H(\sum_i \pi_i \rho_i). \quad \blacksquare$$

In the case of a finite-dimensional Hilbert space and a finite number of states, Corollary 3.1 is the main result in [12].

In the next theorem and corollary we assume that $\mathfrak{M} = \mathbb{B}(\mathbb{C}^d)$ and consider a finite number of states $\rho_1, \rho_2, \dots, \rho_n$.

THEOREM 3.5. *Let $\hat{\varphi}_i^c = \sum_{j=1}^d \lambda_i^j |v_i^j\rangle\langle v_i^j|$ be a spectral decomposition of the operator $\hat{\varphi}_i^c, i = 1, 2, \dots, n$. Write $E = \{v_i^j : j = 1, 2, \dots, d, i = 1, 2, \dots, n\}$. Assume that $\text{Lin } E = \mathbb{C}^d$ and $E \neq A \cup B$, where $A, B \neq \emptyset$ and $\forall v \in A \forall w \in B v \perp w$. We have the equality in (3.5) if and only if $\hat{\varphi}_i^c = aP_i, i = 1, 2, \dots, n$, where a is some positive number, P_i is a projection and $\sum_i P_i = \frac{a_c}{a} \mathbb{1}$.*

Proof. Sufficiency. If the equality holds in (3.5), then it must hold also in (3.7). Thus, we have

$$\sum_i \tau(\hat{\varphi}_i^c(\log \hat{\varphi}' - \log \hat{\varphi}_i^c)) = 0$$

and, by Remark 2.4,

$$\hat{\varphi}' \hat{\varphi}_i^c = \hat{\varphi}_i^c \hat{\varphi}' = (\hat{\varphi}_i^c)^2.$$

The commutation $\hat{\varphi}'\hat{\varphi}_i^c = \hat{\varphi}_i^c\hat{\varphi}'$ means that all eigenvectors from the set E are eigenvectors of $\hat{\varphi}'$. By the assumption on the set E the operator $\hat{\varphi}'$ has only one eigenvalue. Denote it by a . Applying Theorem 3.3, we have the inequalities $\hat{\varphi}' \geq \hat{\varphi}_j^c, j = 1, 2, \dots$, so the operator $\hat{\varphi}'$ is invertible, therefore it is equal to $a\mathbb{1}$. From the equality $\hat{\varphi}_i^c\hat{\varphi}' = (\hat{\varphi}_i^c)^2$ we have $a\hat{\varphi}_i^c = (\hat{\varphi}_i^c)^2$, so all eigenvalues of the operator $\hat{\varphi}_i^c$ are equal to a . Therefore,

$$\hat{\varphi}_i^c = aP_i \quad \text{for some projection } P_i.$$

The equality must also be in formula (3.8), that is,

$$\tau\left(\frac{\sum_i \hat{\varphi}_i^c}{a_c} \log \frac{\sum_i \hat{\varphi}_i^c}{a_c}\right) = \tau\left(\frac{\sum_i \hat{\varphi}_i^c}{a_c} \log \frac{\hat{\varphi}'}{\tau(\hat{\varphi}')}\right).$$

So, by Remark 2.3, $\frac{1}{\tau(\hat{\varphi}')} \hat{\varphi}' = \frac{1}{a_c} \sum_i \hat{\varphi}_i^c$. This gives the condition $\sum_i P_i = \frac{a_c}{a} \mathbb{1}$.

Necessity. Let $M = (M_1, M_2, \dots, M_n), M_i = \frac{a}{a_c} P_i$. We have

$$\sum_i \hat{\varphi}_i^c M_i = \sum_i \frac{a^2}{a_c} P_i = a\mathbb{1} \geq aP_i = \hat{\varphi}_i^c,$$

so Theorem 3.3 implies that M is an optimal measurement and $\max_M r_c(M, \pi) = \tau(a\mathbb{1}) = a$. On the other hand,

$$\begin{aligned} \frac{1}{a_c} \left(\sum_i H(\varphi_i^c) \right) - H\left(\frac{1}{a_c} \left(\sum_i \varphi_i^c \right)\right) &= \frac{1}{a_c} \left(\sum_i H(aP_i) \right) - H(\mathbb{1}) \\ &= \frac{a}{a_c} \sum_i \tau(P_i \log(aP_i)) = \frac{a}{a_c} \sum_i \tau(P_i) \log a = \log a, \end{aligned}$$

therefore

$$\max_M r_c(M, \pi) = 2^{\frac{1}{a_c} (\sum_i H(\varphi_i^c)) - H(\frac{1}{a_c} (\sum_i \varphi_i^c))}. \quad \blacksquare$$

For the probability of detection we obtain

COROLLARY 3.2. Let $\hat{\rho}_i = \sum_{j=1}^d \lambda_i^j |v_i^j\rangle\langle v_i^j|$ be a spectral decomposition of the operator $\hat{\rho}_i, i = 1, 2, \dots, n$. Write $E = \{v_i^j : j = 1, 2, \dots, d, i = 1, 2, \dots, n\}$. Assume that $\text{Lin } E = \mathbb{C}^d$ and $E \neq A \cup B$, where $A, B \neq \emptyset$ and $\forall v \in A \forall w \in B v \perp w$. We have the equality in (3.9) if and only if $\hat{\rho}_i = \frac{1}{m_i} P_i, i = 1, 2, \dots, n$, where $m_i = \tau(P_i), P_i$ is a projection, $\sum_i P_i = m\mathbb{1}, m = \sum_i m_i$ and $\pi_i = m_i/m$.

P r o o f. In Theorem 3.5 consider the loss function of the form $L(i, j) = 1 - \delta_{ij}$ and $c = (1, 1, \dots)$. Then we have $\pi_i \hat{\rho}_i = aP_i$ for some $a > 0$ and the projection P_i . Therefore, $\pi_i = am_i$ and $a = \frac{1}{m}$. Consequently, $\pi_i = m_i/m$ and $\sum_i P_i = m\mathbb{1}$. \blacksquare

REFERENCES

- [1] A. S. Holevo, *Statistical decision theory for quantum systems*, J. Multivariate Anal. 3 (1973), pp. 337–394.
- [2] A. S. Holevo (A. S. Kholevo), *Investigations in the general theory of statistical decisions*, Proc. Steklov Inst. Math. 124 (3) (1978), pp. 1–140.
- [3] A. Łuczak, *Maximizing the probability of detection for pure states*, J. Math. Phys. 50 (5) (2009), 053502.
- [4] S. Neshveyev and E. Størmer, *Dynamical Entropy in Operator Algebras*, Springer, Berlin 2006.
- [5] M. Ozawa, *Optimal measurements for general quantum systems*, Rep. Math. Phys. 18 (1) (1980), pp. 11–28.
- [6] H. Podsiadkowska, *Entropy of quantum measurement*, Entropy 17 (3) (2015), pp. 1181–1196.
- [7] D. Ruelle, *Statistical Mechanics: Rigorous Results*, W. A. Benjamin, Inc., New York–Amsterdam 1969.
- [8] M. Tomamichel, R. Colbeck, and R. Renner, *A fully quantum asymptotic equipartition property*, IEEE Trans. Inform. Theory 55 (12) (2009), pp. 5840–5847.
- [9] H. Umegaki, *Conditional expectation in an operator algebra. IV. Entropy and information*, Kodai Math. Sem. Rep. 14 (1962), pp. 59–85.
- [10] A. Wehrl, *General properties of entropy*, Rev. Modern Phys. 50 (2) (1978), pp. 221–260.
- [11] R. Wieczorek, *On the form of the optimal measurement for the probability of detection*, Internat. J. Theoret. Phys. 54 (12) (2015), pp. 4506–4511.
- [12] S. Yang, J. Lee, and H. Jeong, *Entropic lower bound for distinguishability of quantum states*, Adv. Math. Phys. (2015), article ID 683658.

Rafał Wieczorek
Faculty of Mathematics and Computer Science
Łódź University
ul. S. Banacha 22, 90-238 Łódź, Poland
E-mail: wieczorek@math.uni.lodz.pl

Hanna Podsiadkowska
Faculty of Mathematics and Computer Science
Łódź University
ul. S. Banacha 22, 90-238 Łódź, Poland
E-mail: hpodsedk@math.uni.lodz.pl

Received on 8.2.2017;
revised version on 8.6.2017
