

COMPARISON OF QUANTUM BINARY EXPERIMENTS

ANNA JENČOVÁ*

Mathematical Institute, Slovak Academy of Sciences,
Štefánikova 49, 814 73 Bratislava, Slovakia
(e-mail: jenca@mat.savba.sk)

(Received April 5, 2012 – Revised June 21, 2012)

A quantum binary experiment consists of a pair of density operators on a finite-dimensional Hilbert space. An experiment \mathcal{E} is called ϵ -deficient with respect to another experiment \mathcal{F} if, up to ϵ , its risk functions are not worse than the risk functions of \mathcal{F} , with respect to all statistical decision problems. It is known in the theory of classical statistical experiments that (1) for pairs of probability distributions, one can restrict oneself to testing problems in the definition of deficiency and (2) that 0-deficiency is a necessary and sufficient condition for existence of a stochastic mapping that maps one pair onto another. We show that in the quantum case, the property (1) holds precisely if \mathcal{E} consist of commuting densities. As for property (2), we show that if \mathcal{E} is 0-deficient with respect to \mathcal{F} , then there exists a completely positive mapping that maps \mathcal{E} onto \mathcal{F} , but it is not necessarily trace preserving.

Keywords: comparison of statistical experiments, quantum binary experiments, deficiency, statistical morphisms.

1. Introduction

In classical statistics, a statistical experiment is a parametrized family of probability distributions on a sample space (X, Σ) . The theory of experiments and their comparison was introduced by Blackwell [2] and further developed by many authors, e.g. Torgersen, [17, 18]. Most of the results needed here can be found in [16].

For our purposes, a classical *statistical experiment* $\mathcal{E} = (X, \{p_\theta, \theta \in \Theta\})$ is a parametrized set of probability distributions $p_\theta, \theta \in \Theta$, over a finite set X , where Θ is a finite set of parameters. This can be interpreted as follows: X is a set of possible outcomes $x \in X$ of some experiment, each occurring with probability $p(x)$, where p is a member of a parametrized family $\{p_\theta\}$, but the value of the parameter is not known. After observing x , a decision d is chosen from a finite set D of possible decisions, with some probability $\mu(x, d)$. The function $\mu : X \times D \rightarrow [0, 1]$ is called the *decision function*. It is clear that a decision function is a Markov kernel (or a stochastic matrix), that is, $d \mapsto \mu(x, d)$ is a probability distribution for all $x \in X$.

*Supported by the grants VEGA 2/0032/09 and meta-QUTE ITMS 26240120022.

A *loss function* $W : \Theta \times D \rightarrow \mathbb{R}^+$ represents the loss suffered if $d \in D$ is chosen and the true value of the parameter is θ . The *risk*, or the average loss of the decision procedure μ when the true value is θ is computed as

$$R_{\mathcal{E}}(\theta, W, \mu) = \sum_{x,d} W_{\theta}(d) \mu(x, d) p_{\theta}(x).$$

The couple (D, W) is called a *decision problem*. If D consists of two points, then the decision problems (D, W) are precisely the problems of hypothesis testing.

Let \mathcal{F} be another experiment with the same set of parameters, then its “informative value” can be compared to that of \mathcal{E} by comparing their risk functions for all decision problems. This leads to the definitions of (k, ϵ) -deficiency and ϵ -deficiency, see Section 2. One of the most important results of the theory is the following *randomization criterion*.

THEOREM 1. *Let $\mathcal{E} = (X, \{p_{\theta}, \theta \in \Theta\})$ and $\mathcal{F} = (Y, \{q_{\theta}, \theta \in \Theta\})$ be two experiments. Then \mathcal{E} is ϵ -deficient with respect to \mathcal{F} if and only if there is a Markov kernel $\lambda : X \times Y \rightarrow [0, 1]$ such that*

$$\|\lambda(p_{\theta}) - q_{\theta}\|_1 \leq 2\epsilon,$$

where $\lambda(p) = \sum_x \lambda(x, y) p(x)$.

For $\epsilon = 0$, this is the Blackwell–Sherman–Stein theorem [2, 13, 15]. For a general ϵ it was proved in [17].

If Θ consists of two points, then the experiment is called *binary*. In this case, ϵ -deficiency is equivalent to $(2, \epsilon)$ -deficiency [17], which means that such experiments can be compared by considering only the risk functions of hypothesis testing problems.

The development of the quantum version of comparison of statistical experiments was started recently by several authors [14, 3, 8]. A quantum statistical experiment is a set of density operators on a Hilbert space, mostly of finite dimension. Some versions of the randomization criterion, resp. the Blackwell–Sherman–Stein theorem were obtained, in particular, conditions were found for the existence of a trace preserving completely positive map that maps one experiment onto the other. It was conjectured in [14] that the existence of such positive (but not necessarily completely positive) trace preserving map is equivalent to 0-deficiency. A weaker form of this was obtained in [3], where the notion of a *statistical morphism* was introduced. The (even weaker) notion of a *k-statistical morphism* was considered in [8].

The present paper reviews some of the results of [3] and [8], with focus on the problem of comparison of binary experiments. As an extension of [8], we prove that $(2, \epsilon)$ -deficiency and ϵ -deficiency of a quantum experiment \mathcal{E} with respect to another quantum experiment \mathcal{F} are equivalent for any \mathcal{F} precisely if the experiment \mathcal{E} is abelian, that is, all density matrices ρ_{θ} commute. Moreover, we use the results in [12] to show that any *k-statistical morphism* can be extended to a map that is completely positive, but not trace preserving in general.

2. Quantum statistical experiments

Let \mathcal{H} be a finite-dimensional Hilbert space and let $\mathcal{A} \subseteq B(\mathcal{H})$ be a C^* -algebra. Let $\mathcal{S}(\mathcal{A})$ denote the set of density operators in \mathcal{A} . A (quantum) statistical experiment \mathcal{E} consists of \mathcal{A} and a family $\{\rho_\theta, \theta \in \Theta\} \subset \mathcal{S}(\mathcal{A})$, which is written as $\mathcal{E} = (\mathcal{A}, \{\rho_\theta, \theta \in \Theta\})$. Throughout the paper, we suppose that Θ is a finite set.

The family $\{\rho_\theta, \theta \in \Theta\}$ represents our knowledge of the state of the quantum system represented by \mathcal{A} : it is known that this family contains the state of the system but the true value of θ is not known.

Let (D, W) be a decision problem. The decision is made by a measurement on \mathcal{A} with values in D . Any such measurement is given by a positive operator-valued measure (POVM) $M : D \rightarrow \mathcal{A}$, that is, a collection of operators $M = \{M_d, d \in D\} \subset \mathcal{A}^+$ such that $\sum_d M_d = I$. If all M_d are projections, we say that M is a projection-valued measure (PVM). We will denote the set of all measurements by $\mathcal{M}(D, \mathcal{E})$.

Note that any POVM defines a positive trace preserving map $M : \mathcal{A} \rightarrow \mathcal{F}(D)$, where $\mathcal{F}(D)$ is the C^* -algebra of all functions $D \rightarrow \mathbb{C}$. The map is given by

$$M(a)(d) = \text{Tr } M_d a, \quad a \in \mathcal{A}, \quad d \in D,$$

and any positive trace preserving map $\mathcal{A} \rightarrow \mathcal{F}(D)$ is obtained in this way. Moreover, we define the map $\hat{M} : \mathcal{F}(D) \rightarrow \mathcal{A}$ by

$$\hat{M}(f) = \sum_d f(d) (\text{Tr } M_d)^{-1} M_d, \quad f \in \mathcal{F}(D).$$

Then \hat{M} is again positive and trace preserving. Since $\mathcal{F}(D)$ is abelian, both M and \hat{M} are also completely positive [10].

As it was pointed out in [3], the set of quantum experiments contains the set of classical experiments and these correspond precisely to abelian experiments, that is, experiments such that all densities in the family $\{\rho_\theta, \theta \in \Theta\}$ commute. Indeed, let \mathcal{E} be abelian and let \mathcal{C} be the subalgebra generated by $\{\rho_\theta, \theta \in \Theta\}$. Then \mathcal{C} is generated by a PVM P concentrated on a finite set X and we have the classical experiment $(X, \{p_\theta := P(\rho_\theta), \theta \in \Theta\})$. Conversely, let $(Y, \{q_\theta, \theta \in \Theta\})$ be any classical experiment with $|Y| \leq \dim(\mathcal{H})$ and let $Q : Y \rightarrow \mathcal{A}$ be any PVM, then $(\mathcal{A}, \{\hat{Q}(q_\theta), \theta \in \Theta\})$ defines an abelian quantum experiment. It is clear that $p_\theta = P(\rho_\theta)$ and $\rho_\theta = \hat{P}(p_\theta)$, $\theta \in \Theta$, so that \mathcal{E} and $(X, \{p_\theta\})$ are mapped onto each other by completely positive trace preserving maps. In particular, the experiments are equivalent in the sense defined below.

3. Deficiency

Let \mathcal{E} be an experiment and let (D, W) be a decision problem. The *risk* of the decision procedure $M \in \mathcal{M}(D, \mathcal{E})$ at θ is computed as [5]

$$R_{\mathcal{E}}(\theta, W, M) = \sum_{d \in D} M(\rho_\theta)(d) W_\theta(d) = \sum_d W_\theta(d) \text{Tr } \rho_\theta M_d.$$

Let now $\mathcal{F} = (\mathcal{B}, \{\sigma_\theta, \theta \in \Theta\})$ be another experiment, with $\mathcal{B} \subset B(\mathcal{K})$ for a finite-dimensional Hilbert space \mathcal{K} and with the same parameter set. Let $k \in \mathbb{N}$, $D_k := \{0, \dots, k-1\}$ and let $\epsilon \geq 0$. We say that \mathcal{E} is (k, ϵ) -deficient with respect to \mathcal{F} , in notation $\mathcal{E} \geq_{k, \epsilon} \mathcal{F}$, if for every decision problem (D_k, W) (equivalently, for all decision problems (D, W) with $|D| = k$) and every $N \in \mathcal{M}(D_k, \mathcal{F})$, there is some $M \in \mathcal{M}(D_k, \mathcal{E})$ such that

$$R_{\mathcal{E}}(\theta, W, M) \leq R_{\mathcal{F}}(\theta, W, N) + \epsilon \|W_\theta\|, \quad \theta \in \Theta,$$

where $\|W_\theta\| = \sup_{x \in D_k} W_\theta(x)$. We say that \mathcal{E} is ϵ -deficient with respect to \mathcal{F} , $\mathcal{E} \geq_\epsilon \mathcal{F}$, if it is (k, ϵ) -deficient for all $k \in \mathbb{N}$.

The relation \leq_0 defines a preorder on the set of all experiments. If we have $\mathcal{E} \geq_0 \mathcal{F}$ and simultaneously $\mathcal{F} \geq_0 \mathcal{E}$, then we say that \mathcal{E} and \mathcal{F} are *equivalent*, $\mathcal{E} \sim \mathcal{F}$. The equivalence relation $\mathcal{E} \sim_k \mathcal{F}$ is defined similarly, and \mathcal{E} and \mathcal{F} are called *k-equivalent*.

The Theorem 2 below (apart from (iii)) was proved in [8, Theorem 5] in a more general setting. We give the proof in our simpler case, just for the convenience of the reader.

The most important ingredient of the proof is the *minimax theorem*, which can be found in [16].

THEOREM 2. *Let $\mathcal{E} = (\mathcal{A}, \{\rho_\theta, \theta \in \Theta\})$ and $\mathcal{F} = (\mathcal{B}, \{\sigma_\theta, \theta \in \Theta\})$ be two experiments with the same parameter set Θ , $|\Theta| < \infty$. Let $k \in \mathbb{N}$, $\epsilon \geq 0$. The following are equivalent.*

(i) $\mathcal{E} \geq_{k, \epsilon} \mathcal{F}$

(ii) For every loss function $W : \Theta \times D_k \rightarrow \mathbb{R}^+$,

$$\min_{M \in \mathcal{M}(D_k, \mathcal{E})} \sum_{\theta} R_{\mathcal{E}}(\theta, W, M) \leq \min_{N \in \mathcal{M}(D_k, \mathcal{F})} \sum_{\theta} R_{\mathcal{F}}(\theta, W, N) + \epsilon \|W\|$$

where $\|W\| = \sum_{\theta} \|W_\theta\|$.

(iii) For every loss function $W : \Theta \times D_k \rightarrow \mathbb{R}^+$,

$$\max_{M \in \mathcal{M}(D_k, \mathcal{E})} \sum_{\theta} R_{\mathcal{E}}(\theta, W, M) \geq \max_{N \in \mathcal{M}(D_k, \mathcal{F})} \sum_{\theta} R_{\mathcal{F}}(\theta, W, N) - \epsilon \|W\|.$$

(iv) For every $N \in \mathcal{M}(D_k, \mathcal{F})$ there is some $M \in \mathcal{M}(D_k, \mathcal{E})$ such that

$$\|M(\rho_\theta) - N(\sigma_\theta)\|_1 \leq 2\epsilon, \quad \forall \theta \in \Theta.$$

Proof: Suppose (i), then for any $N \in \mathcal{M}(D_k, \mathcal{F})$, there is some $M \in \mathcal{M}(D_k, \mathcal{E})$ such that

$$\sum_{\theta} R_{\mathcal{E}}(\theta, W, M) \leq \sum_{\theta} R_{\mathcal{F}}(\theta, W, N) + \epsilon \|W\|,$$

this implies (ii).

Suppose (ii) and let $W : \Theta \times D_k \rightarrow \mathbb{R}^+$ be a loss function. Then $\tilde{W} : \Theta \times D_k \rightarrow \mathbb{R}^+$ given by $\tilde{W}_\theta = \|W_\theta\| - W_\theta$ is a loss function with $\|\tilde{W}\| \leq \|W\|$. Since $R_{\mathcal{E}}(\theta, \tilde{W}, M) = \|W_\theta\| - R_{\mathcal{E}}(\theta, W, M)$ and similarly for $R_{\mathcal{F}}$, we have (ii) implies (iii).

Suppose (iii), and let $N \in \mathcal{M}(D_k, \mathcal{F})$. Then for every loss function W , we have

$$\max_{M \in \mathcal{M}(D_k, \mathcal{E})} \sum_{\theta} R_{\mathcal{E}}(\theta, W, M) \geq \sum_{\theta} R_{\mathcal{F}}(\theta, W, N) - \epsilon \|W\|,$$

and this implies that

$$\sup_{W, \|W\| \leq 1} \min_{M \in \mathcal{M}(D_k, \mathcal{E})} \sum_{\theta} (R_{\mathcal{F}}(\theta, W, N) - R_{\mathcal{E}}(\theta, W, M)) \leq \epsilon.$$

The set $\mathcal{M} = \mathcal{M}(D_k, \mathcal{E})$ is compact and obviously convex and the set \mathcal{W} of all loss functions W with $\|W\| \leq 1$ is convex as well. Moreover, the function $(M, W) \mapsto \sum_{\theta} (R_{\mathcal{F}}(\theta, W, N) - R_{\mathcal{E}}(\theta, W, M))$ is linear in both arguments, hence the minimax theorem applies and we get

$$\begin{aligned} \epsilon &\geq \min_{M \in \mathcal{M}} \sup_{W \in \mathcal{W}} \sum_{\theta} (R_{\mathcal{F}}(\theta, W, N) - R_{\mathcal{E}}(\theta, W, M)) \\ &= \min_{M \in \mathcal{M}} \sup_{W \in \mathcal{W}} \sum_{\theta, d} W_{\theta}(d) (N(\sigma_{\theta})(d) - M(\rho_{\theta})(d)). \end{aligned}$$

Let $\mathcal{P}(\Theta)$ be the set of all probability measures on Θ and let $p \in \mathcal{P}(\Theta)$. For $M \in \mathcal{M}$ fixed, let W be given by

$$W_{\theta}(x) = \begin{cases} p(\theta) & \text{if } N(\sigma_{\theta})(x) - M(\rho_{\theta})(x) > 0, \\ 0 & \text{otherwise.} \end{cases}$$

Then $W \in \mathcal{W}$, so that we get

$$\begin{aligned} \epsilon &\geq \min_{M \in \mathcal{M}} \sum_{\theta} \sum_{x \in D_k} W_{\theta}(x) (N(\sigma_{\theta})(x) - M(\rho_{\theta})(x)) \\ &= \min_{M \in \mathcal{M}} \sum_{\theta} p(\theta) \frac{1}{2} \|N(\sigma_{\theta}) - M(\rho_{\theta})\|_1. \end{aligned}$$

Since this holds for any $p \in \mathcal{P}(\Theta)$, we have obtained

$$\sup_{p \in \mathcal{P}(\Theta)} \min_{M \in \mathcal{M}} \sum_{\theta} p(\theta) \|M(\rho_{\theta}) - N(\sigma_{\theta})\|_1 \leq 2\epsilon.$$

The set $\mathcal{P}(\Theta)$ is convex and the function $\mathcal{M} \times \mathcal{P}(\Theta) \rightarrow \mathbb{R}$, given by $(M, p) \mapsto \sum_{\theta} p(\theta) \|M(\rho_{\theta}) - N(\sigma_{\theta})\|_1$ is convex in M and concave (linear) in p . Hence the minimax theorem applies again and we have

$$\min_M \sup_p \|M(\rho_{\theta}) - N(\sigma_{\theta})\|_1 = \sup_p \min_M \sum_{\theta} p(\theta) \|M(\rho_{\theta}) - N(\sigma_{\theta})\|_1 \leq 2\epsilon$$

which clearly implies (iv), by taking the probability measures concentrated in $\theta \in \Theta$.

Suppose (iv) and let $N \in \mathcal{M}(D_k, \mathcal{F})$. Let $M \in \mathcal{M}(D_k, \mathcal{E})$ be chosen for N by

(iv). Then for any loss function W ,

$$\begin{aligned} R_{\mathcal{E}}(\theta, W, M) - R_{\mathcal{F}}(\theta, W, N) &= \sum_{x \in D_k} W_{\theta}(x)(M(\rho_{\theta})(x) - N(\sigma_{\theta})(x)) \\ &\leq \frac{\|W_{\theta}\|}{2} \|M(\rho_{\theta}) - N(\sigma_{\theta})\|_1 \leq \epsilon \|W_{\theta}\|. \end{aligned}$$

so that $\mathcal{E} \geq_{k, \epsilon} \mathcal{F}$. □

The following corollary is a generalization of the classical randomization criterion to the case when the experiment \mathcal{F} is abelian. In the case when $\epsilon = 0$, it was proved in [3].

COROLLARY 1. *Let $\mathcal{E} = (\mathcal{A}, \{\rho_{\theta}, \theta \in \Theta\})$ and let $\mathcal{F} = (\mathcal{B}, \{\sigma_{\theta}, \theta \in \Theta\})$ be abelian. Then $\mathcal{E} \geq_{\epsilon} \mathcal{F}$ if and only if there is a completely positive trace preserving map $T : \mathcal{A} \rightarrow \mathcal{B}$ such that*

$$\|T(\rho_{\theta}) - \sigma_{\theta}\|_1 \leq 2\epsilon, \quad \theta \in \Theta.$$

Proof: Let $(X, \{p_{\theta}, \theta \in \Theta\})$ be a classical experiment equivalent to \mathcal{F} and let $P = (P_1, \dots, P_m)$ be the PVM such that $P(\sigma_{\theta}) = p_{\theta}$, $\theta \in \Theta$. Suppose $\mathcal{E} \geq_{\epsilon} \mathcal{F}$, then $P \in \mathcal{M}(X, \mathcal{F})$ and by Theorem 2 (iv), there is some $M \in \mathcal{M}(X, \mathcal{E})$ such that

$$\|M(\rho_{\theta}) - P(\sigma_{\theta})\|_1 = \|M(\rho_{\theta}) - p_{\theta}\|_1 \leq 2\epsilon.$$

Put $T = \hat{P} \circ M$, then $T : \mathcal{A} \rightarrow \mathcal{B}_0 \subseteq \mathcal{B}$ is positive and trace preserving, where \mathcal{B}_0 is the abelian subalgebra generated by P . Hence T is also completely positive. Moreover,

$$\|T(\rho_{\theta}) - \sigma_{\theta}\|_1 = \|\hat{P}(M(\rho_{\theta}) - p_{\theta})\|_1 \leq \|M(\rho_{\theta}) - p_{\theta}\|_1 \leq 2\epsilon.$$

For the converse, let $N \in \mathcal{M}(D, \mathcal{F})$ for any finite set D . Put $Q = N \circ T$, then $Q \in \mathcal{M}(D, \mathcal{E})$ and

$$\|Q(\rho_{\theta}) - N(\sigma_{\theta})\|_1 = \|N(T(\rho_{\theta}) - \sigma_{\theta})\|_1 \leq 2\epsilon.$$

By Theorem 2 (iv), this implies $\mathcal{E} \geq_{\epsilon} \mathcal{F}$. □

3.1. Deficiency with respect to testing problems

Let (D_2, W) be a decision problem. Then any $M \in \mathcal{M}(D_2, \mathcal{E})$ has the form $(M_0, I - M_0)$ for some $0 \leq M_0 \leq I$ and the risk of M is

$$R_{\mathcal{E}}(\theta, M, W) = W_{\theta}(1) + (W_{\theta}(0) - W_{\theta}(1))\text{Tr } \rho_{\theta} M_0.$$

By Theorem 2 (iii), $\mathcal{E} \geq_{2, \epsilon} \mathcal{F}$ if and only if

$$\max_{\substack{M_0 \in \mathcal{A}, \\ 0 \leq M_0 \leq I}} \text{Tr} \sum_{\theta} A_{\theta} \rho_{\theta} M_0 \geq \max_{\substack{N_0 \in \mathcal{B}, \\ 0 \leq N_0 \leq I}} \text{Tr} \sum_{\theta} A_{\theta} \sigma_{\theta} N_0 - \epsilon \|W\| \quad (1)$$

for all loss functions W , where we denote $A_\theta := W_\theta(0) - W_\theta(1)$. It is easy to see that

$$\max_{0 \leq M_0 \leq 1} \text{Tr} \sum_{\theta} A_\theta \rho_\theta M_0 = \text{Tr} \left[\sum_{\theta} A_\theta \rho_\theta \right]^+ = \frac{1}{2} \left(\sum_{\theta} A_\theta + \left\| \sum_{\theta} A_\theta \rho_\theta \right\|_1 \right), \quad (2)$$

here we have used the equality $\text{Tr} a^+ = \frac{1}{2}(\text{Tr} a + \text{Tr} |a|)$ for a self-adjoint element $a \in \mathcal{A}$.

THEOREM 3. $\mathcal{E} \geq_{2,\epsilon} \mathcal{F}$ if and only if

$$\left\| \sum_{\theta} A_\theta \rho_\theta \right\|_1 \geq \left\| \sum_{\theta} A_\theta \sigma_\theta \right\|_1 - 2\epsilon \sum_{\theta} |A_\theta|$$

for any coefficients $A_\theta \in \mathbb{R}$.

Proof: Follows from (1) and (2). For the ‘if’ part, put $A_\theta = W_\theta(0) - W_\theta(1)$, we then have $\sum_{\theta} |A_\theta| \leq \|W\|$. For the converse, let $F_+ := \{\theta, A_\theta > 0\}$, $F_- := \{\theta, A_\theta \leq 0\}$ and put

$$W_\theta(0) = \begin{cases} A_\theta & \text{if } \theta \in F_+ \\ 0 & \text{otherwise} \end{cases},$$

$$W_\theta(1) = \begin{cases} -A_\theta & \text{if } \theta \in F_- \\ 0 & \text{otherwise} \end{cases}.$$

Then W is a loss function with $\|W\| = \sum_{\theta} |A_\theta|$. □

3.2. Deficiency and sufficiency

Let $T : \mathcal{A} \rightarrow \mathcal{B}$ be a completely positive trace preserving map. The experiment $\mathcal{F} = (\mathcal{B}, \{T(\rho_\theta), \theta \in \Theta\})$ is called a *randomization* of \mathcal{E} . If $N \in \mathcal{M}(D, \mathcal{F})$, then $T^*(N) \in \mathcal{M}(D, \mathcal{E})$ and it is clear that $T^*(N)$ has the same risks as N , hence \mathcal{E} is 0-deficient with respect to \mathcal{F} .

Suppose that in this setting, \mathcal{F} is $(k, 0)$ -deficient with respect to \mathcal{E} , then we say that T is *k-sufficient* for \mathcal{E} . If also \mathcal{E} is a randomization of \mathcal{F} , then we say that T is *sufficient* for \mathcal{E} , this definition of sufficiency was introduced in [11]. If T is a restriction to a subalgebra $\mathcal{A}_0 \subset \mathcal{A}$, then we say that \mathcal{A}_0 is *k-sufficient*, resp. *sufficient* for \mathcal{E} , if T is. If the experiments are abelian, then it follows by the randomization criterion that T is sufficient if and only if it is *k-sufficient* for every $k \in \mathbb{N}$. Moreover, for abelian binary experiments, T is sufficient if and only if it is 2-sufficient. (In fact, the last statement hold for all classical statistical experiments [16].)

It is not clear if any of the above two statements holds for quantum experiments. The latter condition for binary experiments was investigated in [6], for a subalgebra \mathcal{A}_0 . It was shown that \mathcal{A}_0 is 2-sufficient if and only if it contains all projections

$P_{t,+}$, $t \geq 0$ (see Lemma 1) and that this is equivalent to sufficiency in some cases. In particular:

THEOREM 4. *Let $\mathcal{E} = (\mathcal{A}, \{\rho_1, \rho_2\})$ be an experiment and let $\mathcal{A}_0 \subseteq \mathcal{A}$ be an abelian subalgebra. Then the following are equivalent:*

- (i) \mathcal{A}_0 is 2-sufficient,
- (ii) \mathcal{A}_0 is sufficient,
- (iii) \mathcal{A}_0 is sufficient and \mathcal{E} is abelian.

Proof: The equivalence of (i) and (ii) was proved in [6, Theorem 5(2)], (ii) \implies (iii) follows from [9, Theorem 9.10]. (iii) \implies (i) is obvious. \square

4. Binary experiments

Let $\mathcal{E} = (\mathcal{A}, \{\rho_1, \rho_2\})$ be a binary experiment. Note that we may suppose that $\rho_1 + \rho_2$ is invertible, since \mathcal{E} can be replaced by the experiment $(P\mathcal{A}P, \{\rho_1, \rho_2\})$, where $P = \text{supp}(\rho_1 + \rho_2)$ is the support projection of $\rho_1 + \rho_2$.

Let us denote

$$f_{\mathcal{E}}(t) := \max_{\substack{M \in \mathcal{A}, \\ 0 \leq M \leq I}} \text{Tr}(\rho_1 - t\rho_2)M, \quad t \in \mathbb{R}.$$

Then by (2),

$$f_{\mathcal{E}}(t) = \text{Tr}(\rho_1 - t\rho_2)_+ = \frac{1}{2}(\|\rho_1 - t\rho_2\|_1 + 1 - t). \quad (3)$$

It is easy to see that Theorem 3 for binary experiments has the following form.

THEOREM 5. *Let $\mathcal{E} = (\mathcal{A}, \{\rho_1, \rho_2\})$ and $\mathcal{F} = (\mathcal{B}, \{\sigma_1, \sigma_2\})$. Then the following are equivalent:*

- (i) $\mathcal{E} \geq_{2,\epsilon} \mathcal{F}$,
- (ii) $\|\rho_1 - t\rho_2\|_1 \geq \|\sigma_1 - t\sigma_2\|_1 - 2(1+t)\epsilon$ for all $t \geq 0$,
- (iii) $f_{\mathcal{E}}(t) \geq f_{\mathcal{F}}(t) - (1+t)\epsilon$ for all $t \geq 0$.

We will need some properties of the function $f_{\mathcal{E}}$. First, we state the quantum version of the Neyman–Pearson lemma [4, 5]. For this, let us denote $P_{t,+} := \text{supp}(\rho_1 - t\rho_2)_+$ and $P_{t,0} = \ker(\rho_1 - t\rho_2)$ for $t \geq 0$.

LEMMA 1. *We have $f_{\mathcal{E}}(t) = \text{Tr}(\rho_1 - t\rho_2)M$ for some $M \in \mathcal{A}$, $0 \leq M \leq I$, if and only if*

$$M = P_{t,+} + X, \quad 0 \leq X \leq P_{t,0}.$$

The proof of the following lemma can be found in Appendix.

LEMMA 2.

- (i) $f_{\mathcal{E}}$ is continuous, convex and $f_{\mathcal{E}}(t) \geq \max\{1 - t, 0\}$, $t \in \mathbb{R}$.
- (ii) $f_{\mathcal{E}}$ is nonincreasing in \mathbb{R} . Moreover, $f_{\mathcal{E}}$ is analytic in \mathbb{R} except of some points $0 \leq t_1 < \dots < t_l$, $l \leq \dim(\mathcal{H})$, where $f_{\mathcal{E}}$ is not differentiable. These are exactly the points where $P_{t,0} \neq 0$.

We will denote by $\mathcal{T}_{\mathcal{E}} := \{t_1, \dots, t_l\}$ the set of points defined in (ii).

4.1. Deficiency and 2-deficiency for binary experiments

For classical binary experiments, it was proved in [17] that $\mathcal{E} \geq_{2,\epsilon} \mathcal{F}$ is equivalent with $\mathcal{E} \geq_{\epsilon} \mathcal{F}$, so that for comparison of such experiments it is enough to consider all testing problems. We prove below that this equivalence remains true if only \mathcal{E} is abelian, and that this property characterizes abelian binary experiments.

We will need the following lemma.

LEMMA 3. *Let $s_1, s_2 \notin \mathcal{T}_{\mathcal{E}}$, $0 < s_1 < s_2$. Then there is a classical experiment $\mathcal{F} = (X = \{1, 2, 3\}, \{p, q\})$, such that $f_{\mathcal{E}}(t) \geq f_{\mathcal{F}}(t)$ for all t and $f_{\mathcal{E}}(s_i) = f_{\mathcal{F}}(s_i)$, $i = 1, 2$.*

Proof: Let us define linear functions $g_i(t) := a_i - tb_i$, $i = 0, \dots, 3$, where $a_0 = b_0 = 1$, $a_3 = b_3 = 0$ and $a_i = f_{\mathcal{E}}(s_i) - s_i f'_{\mathcal{E}}(s_i)$, $b_i = -f'_{\mathcal{E}}(s_i)$, $i = 1, 2$, so that

$$g_i(t) = f_{\mathcal{E}}(s_i) + (t - s_i)f'_{\mathcal{E}}(s_i)$$

is tangent to $f_{\mathcal{E}}$ at s_i , $i = 0, 1, 2$, where we put $s_0 = 0$. Since $f_{\mathcal{E}}$ is convex and $f_{\mathcal{E}}(t) \geq \max\{1 - t, 0\}$, $g_i(t) \leq f(t)$, for all i and t . Moreover, since $f_{\mathcal{E}}$ is also nonincreasing, we have for any $t < 0$, $-1 = f'_{\mathcal{E}}(t) \leq f'_{\mathcal{E}}(s_1) \leq f'_{\mathcal{E}}(s_2) \leq 0$ so that $b_0 \geq b_1 \geq b_2 \geq b_3$. Convexity and $f_{\mathcal{E}}(0) = 1$ also imply that

$$\begin{aligned} 1 - a_1 &= 1 - f_{\mathcal{E}}(s_1) + s_1 f'_{\mathcal{E}}(s_1) \geq 0, \\ a_1 - a_2 &= f_{\mathcal{E}}(s_1) - f_{\mathcal{E}}(s_2) - f'_{\mathcal{E}}(s_2)(s_1 - s_2) + s_1(b_1 - b_2) \geq 0, \\ a_2 &= f_{\mathcal{E}}(s_2) + s_2 b_2 \geq 0, \end{aligned}$$

so that $a_0 \geq a_1 \geq a_2 \geq a_3$. Put $p_i := a_{i-1} - a_i$, $q_i := b_{i-1} - b_i$, $i = 1, 2, 3$, then $p = (p_1, p_2, p_3)$ and $q = (q_1, q_2, q_3)$ are probability measures. Let $\mathcal{F} := (\{1, 2, 3\}, \{p, q\})$, then

$$f_{\mathcal{F}}(t) = \sum_{i, p_i - tq_i > 0} p_i - tq_i = \sum_{i, g_{i-1}(t) > g_i(t)} g_{i-1}(t) - g_i(t).$$

Let us now define the points t'_0, \dots, t'_3 as follows. Put $t'_0 := 0$ and for $i = 1, 2, 3$, let $t'_i := t'_{i-1}$ if $g_i = g_{i-1}$, otherwise let t'_i be such that $g_i(t) < g_{i-1}(t)$ for $t < t'_i$ and $g_i(t'_i) = g_{i-1}(t'_i)$. Note that $t'_i \geq 0$, since $g_i(0) \leq g_{i-1}(0)$. Moreover, since $g_i(s_i) = f_{\mathcal{E}}(s_i) \geq g_{i-1}(s_i)$, we have $t'_i \leq s_i$ for $i = 0, 1, 2$. In fact, $t'_i < s_i$ for $i = 1, 2$, since $g_{i-1}(s_i) = g_i(s_i) = f_{\mathcal{E}}(s_i)$ implies $f_{\mathcal{E}} = g_i = g_{i-1}$ in some interval containing s_i , so that $t'_i = t'_{i-1} \leq s_{i-1} < s_i$. Similarly, for $i = 2, 3$, $g_i(s_{i-1}) \leq f_{\mathcal{E}}(s_{i-1}) = g_{i-1}(s_{i-1})$, so that we either have $t'_i = t'_{i-1}$ or $t'_i > s_{i-1}$. In the case that $g_2(t) > 0$ for all t , we put $t'_3 = \infty$. Putting all together, we have $0 = t'_0 \leq t'_1 < s_1 < t'_2 < s_2 < t'_3 \leq \infty$ and

$$\begin{aligned} f_{\mathcal{F}}(t) &= \sum_{j=i}^3 g_{j-1}(t) - g_j(t) = g_{i-1}(t), \quad t \in \langle t'_{i-1}, t'_i \rangle, \quad i = 1, 2, 3, \\ f_{\mathcal{F}}(t) &= 0, \quad t \in \langle t'_3, \infty \rangle. \end{aligned}$$

It follows that $f_{\mathcal{F}}(t) \leq f_{\mathcal{E}}(t)$ for all t and $f_{\mathcal{F}}(s_i) = f_{\mathcal{E}}(s_i)$, $i = 1, 2$. □

We will now state the main result of this section.

THEOREM 6. *Let $\mathcal{E} = (\mathcal{A}, \{\rho_1, \rho_2\})$ be a binary experiment. Then the following are equivalent:*

- (i) $\mathcal{E} \geq_{2,\epsilon} \mathcal{F} \iff \mathcal{E} \geq_\epsilon \mathcal{F}$ for any $\epsilon \geq 0$ and any abelian binary experiment \mathcal{F} ,
- (ii) $\mathcal{E} \geq_{2,\epsilon} \mathcal{F} \iff \mathcal{E} \geq_\epsilon \mathcal{F}$ for any $\epsilon \geq 0$ and any binary experiment \mathcal{F} ,
- (iii) $\mathcal{E} \geq_{2,0} \mathcal{F} \iff \mathcal{E} \geq_0 \mathcal{F}$ for any abelian binary experiment \mathcal{F} ,
- (iv) \mathcal{E} is abelian.

Proof: Suppose (i) and let $\mathcal{F} = (\mathcal{B}, \{\sigma_1, \sigma_2\})$ be any binary experiment such that $\mathcal{E} \geq_{2,\epsilon} \mathcal{F}$. Let D be a finite set and let $N \in \mathcal{M}(D, \mathcal{F})$. Put $p_i := N(\sigma_i)$, $i = 1, 2$, and let $\mathcal{F}_N := (D, \{p_1, p_2\})$. Then by Theorem 5, we have for each $t \geq 0$,

$$\|\rho_1 - t\rho_2\|_1 \geq \|\sigma_1 - t\sigma_2\|_1 - 2(1+t)\epsilon \geq \|p_1 - tp_2\|_1 - 2(1+t)\epsilon.$$

Hence $\mathcal{E} \geq_{2,\epsilon} \mathcal{F}_N$ and (i) implies that $\mathcal{E} \geq_\epsilon \mathcal{F}_N$. By Corollary 1, there is some $M \in \mathcal{M}(D, \mathcal{E})$ such that

$$\|M(\rho_i) - N(\sigma_i)\|_1 = \|M(\rho_i) - p_i\|_1 \leq 2\epsilon, \quad i = 1, 2.$$

By Theorem 2, $\mathcal{E} \geq_\epsilon \mathcal{F}$ and this implies (ii). (ii) trivially implies (iii).

Suppose (iii). Choose any points $s_1, s_2 \notin \mathcal{T}_\mathcal{E}$, $0 < s_1 < s_2$, then by Lemma 3, there is a classical experiment $\mathcal{F} = (\{1, 2, 3\}, \{p_1, p_2\})$ such that $f_\mathcal{E}(t) \geq f_\mathcal{F}(t)$ for $t \geq 0$ and $f_\mathcal{E}(s_i) = f_\mathcal{F}(s_i)$, $i = 1, 2$. By Theorem 5, this implies that $\mathcal{E} \geq_{2,0} \mathcal{F}$ and by (iii), $\mathcal{E} \geq_0 \mathcal{F}$. By Corollary 1, there is a POVM $M : \{1, 2, 3\} \rightarrow \mathcal{A}$ such that $p_k = M(\rho_k)$, $k = 1, 2$. For $i = 1, 2$, put $J_i := \{j \in \{1, 2, 3\}, p_1(j) - s_i p_2(j) > 0\}$, then we have

$$\begin{aligned} f_\mathcal{E}(s_i) &= f_\mathcal{F}(s_i) = \sum_{j \in J_i} p_1(j) - s_i p_2(j) \\ &= \sum_{j \in J_i} \text{Tr}(\rho_1 M_j) - s_i \text{Tr}(\rho_2 M_j) = \text{Tr}(\rho_1 - s_i \rho_2) \sum_{j \in J_i} M_j. \end{aligned}$$

Since $s_i \notin \mathcal{T}_\mathcal{E}$, we have $P_{s_i,0} = 0$ and Lemma 1 implies that $\sum_{j \in J_i} M_j = P_{s_i,+}$. Hence the projection $P_{s_i,+}$ is in the range of M . Since for all $j \in \{1, 2, 3\}$ we either have $M_j \leq P_{s_i,+}$ or $M_j \leq I - P_{s_i,+}$, $P_{s_i,+}$ must commute with all M_j . In particular, $P_{s_1,+}$ and $P_{s_2,+}$ commute.

Since this can be done for any such s_1, s_2 , it follows that all $\{P_{t,+}, t \notin \mathcal{T}_\mathcal{E}\}$ are mutually commuting projections. Since $t \mapsto P_{t,+}$ is right-continuous, it follows that $P_{t_j,+}$ commutes with all $P_{s,+}$ for $s \notin \mathcal{T}_\mathcal{E}$, and by repeating this argument, $P_{t,+}$ are mutually commuting projections for all $t \geq 0$.

Let now \mathcal{A}_0 be the subalgebra generated by $\{P_{t,+}, t \geq 0\}$. Then \mathcal{A}_0 is an abelian subalgebra which is 2-sufficient for \mathcal{E} . Hence \mathcal{E} must be abelian by Theorem 4.

The implication (iv) \implies (i) was proved by Torgersen [17]. \square

REMARK 1. If $\dim(\mathcal{H}) = \dim(\mathcal{K}) = 2$, it was proved in [1] that $\mathcal{E} \geq_{2,0} \mathcal{F}$ if and only if \mathcal{F} is a randomization of \mathcal{E} . The above proof shows that if $\dim(\mathcal{K}) \geq 3$ this is no longer true unless \mathcal{E} is abelian.

5. Statistical morphisms

Let $S_{\mathcal{E}} := \text{span}\{\rho_{\theta}, \theta \in \Theta\}$. A k -statistical morphism [3, 8] is a linear map $L : S_{\mathcal{E}} \rightarrow \mathcal{B}$ such that

- (i) $L(\rho_{\theta}) \in \mathcal{S}(\mathcal{B})$ for all θ ,
- (ii) for each POVM $N : D_k \rightarrow \mathcal{B}$ there is some $M \in \mathcal{M}(D_k, \mathcal{E})$ satisfying

$$\text{Tr } L(\rho)N_i = \text{Tr } \rho M_i, \quad i \in D_k, \quad \rho \in S_{\mathcal{E}}.$$

The map L is a *statistical morphism* if it is a k -statistical morphism for any k . It is clear that any positive trace preserving map $L : \mathcal{A} \rightarrow \mathcal{B}$ defines a statistical morphism. The proof of the following proposition appears also in [8].

PROPOSITION 1. $\mathcal{E} \geq_{k,0} \mathcal{F}$ if and only if there is a k -statistical morphism $L : S_{\mathcal{E}} \rightarrow \mathcal{B}$ such that $L(\rho_{\theta}) = \sigma_{\theta}$.

Proof: Suppose that $\mathcal{E} \geq_{k,0} \mathcal{F}$ for some k , then we also have $\mathcal{E} \geq_{2,0} \mathcal{F}$, and by Theorem 3, this implies $\|\sum_{\theta} A_{\theta} \rho_{\theta}\|_1 \geq \|\sum_{\theta} A_{\theta} \sigma_{\theta}\|_1$ for any $A_{\theta} \in \mathbb{R}$. Put $L : \rho_{\theta} \mapsto \sigma_{\theta}$ and extend to $S_{\mathcal{E}}$ by $L(\sum_{\theta} A_{\theta} \rho_{\theta}) = \sum_{\theta} A_{\theta} L(\rho_{\theta})$, then $\|L(x)\|_1 \leq \|x\|_1$ for $x \in S_{\mathcal{E}}$, so that L is a well-defined linear map on $S_{\mathcal{E}}$. Theorem 2 (iv) now implies that L is a k -statistical morphism. The converse is obvious. \square

In [14] and [3], a question was raised whether 0-deficiency is equivalent with the existence of a trace preserving positive map that maps one experiment onto another. It is clear that this question is equivalent with the question if any statistical morphism can be extended to a trace preserving positive map. We show below that if \mathcal{E} and \mathcal{F} are binary experiments, then any k -statistical morphism such that $L(\rho_i) = \sigma_i$, $i = 1, 2$ can be extended to even a completely positive map, but Theorem 6 implies that such an extension is not trace preserving in general. This shows that the condition that the map preserves trace cannot be omitted.

Let t_1 be as in Lemma 2. Note that

$$t_1 = \max\{t \geq 0, f_{\mathcal{E}}(t) = 1 - t\} = \max\{t \geq 0, \rho_1 - t\rho_2 \geq 0\} \quad (4)$$

and $t_1 = 0$ if and only if $\text{supp } \rho_2 \not\leq \text{supp } \rho_1$. Let us denote

$$t_{\max} := \min\{t \geq 0, f_{\mathcal{E}}(t) = 0\} = \min\{t \geq 0, \rho_1 - t\rho_2 \leq 0\}. \quad (5)$$

Then we have either $t_{\max} = t_1$ or $t_{\max} = \infty$, and the latter happens if and only if $\text{supp } \rho_1 \not\leq \text{supp } \rho_2$. We have

$$t_1 \rho_2 \leq \rho_1 \leq t_{\max} \rho_2 \quad (6)$$

and t_1, t_{\max} are extremal values for which the inequality occurs. Equivalently,

$$t_{\max}^{-1} \rho_1 \leq \rho_2 \leq t_1^{-1} \rho_1 \quad (7)$$

with t_{\max}^{-1} and t_1^{-1} extremal. We also remark that $t_1 = \sup(\rho_1/\rho_2)$ and $t_{\max} = \inf(\rho_1/\rho_2)$ as defined in [12].

THEOREM 7. Let $\mathcal{E} = (\mathcal{A}, \{\rho_1, \rho_2\})$, $\mathcal{F} = (\mathcal{B}, \{\sigma_1, \sigma_2\})$ be binary experiments. If $\mathcal{E} \geq_{2,0} \mathcal{F}$, then there is a completely positive map $T : \mathcal{A} \rightarrow \mathcal{B}$ such that $T(\rho_i) = \sigma_i$, $i = 1, 2$.

Proof: Let $\mathcal{E} \geq_{0,2} \mathcal{F}$, then there is a 2-statistical morphism $L : S_{\mathcal{E}} \rightarrow \mathcal{B}$, $L(\rho_i) = \sigma_i$, $i = 1, 2$. Moreover, $f_{\mathcal{E}}(t) \geq f_{\mathcal{F}}(t)$ for all t . Let t'_1 and t'_{\max} be as in (4) and (5) for \mathcal{F} . Since $f_{\mathcal{F}}(t) \geq \max\{0, 1 - t\}$, we must have $t_1 \leq t'_1$ and $t'_{\max} \leq t_{\max}$. The rest of the proof is the same as the proof of [12, Theorem 21]:

Let $u, v \in S_{\mathcal{E}}$ be positive elements such that $\ker(u) \not\subseteq \ker(v)$ and $\ker(v) \not\subseteq \ker(u)$. Then there are some $\varphi, \psi \in \mathcal{H}$ such that $u\varphi = v\psi = 0$, but $u\psi \neq 0$, $v\varphi \neq 0$. Put

$$T(a) = \frac{\langle \psi, a\psi \rangle}{\langle \psi, u\psi \rangle} L(u) + \frac{\langle \varphi, a\varphi \rangle}{\langle \varphi, v\varphi \rangle} L(v), \quad a \in \mathcal{A},$$

then T is a completely positive extension of L . We show that such u and v exist.

Suppose $t_{\max} < \infty$ so that $\text{supp } \rho_1 \leq \text{supp } \rho_2$, then $u := t_{\max}\rho_2 - \rho_1$, $v := \rho_1 - t_1\rho_2$. Then $u, v \geq 0$ and the condition on the kernels follows by extremality of t_1 and t_{\max} . If $t_{\max} = \infty$ but $t_1 > 0$, then we put $u := t_1^{-1}\rho_1 - \rho_2$ and $v := \rho_2$. Finally, if $t_{\max} = \infty$ and $t_1 = 0$, then we put $u := \rho_1$, $v := \rho_2$. \square

REMARK 1. One can see that the extension obtained in the above proof cannot be trace preserving unless $\dim \mathcal{H} = 2$ and \mathcal{E} is abelian.

Appendix: Proof of Lemma 2

The statement (i) follows easily by definition and (3).

Let $\rho(t) := \rho_1 - t\rho_2$. It can be shown ([7, Chap. II]) that the eigenvalues of $\rho(t)$ are analytic functions $t \mapsto \lambda_i(t)$ for all $t \in \mathbb{R}$. It follows that $\rho(t)$ has a constant number N of distinct eigenvalues $\lambda_1(t), \dots, \lambda_N(t)$, apart from some exceptional points where some of these eigenvalues are equal, and there is a finite number of such points in any finite interval. Moreover, let $P_i(t)$ be the eigenprojection corresponding to $\lambda_i(t)$ for a non-exceptional point t , then $t \mapsto P_i(t)$ can be extended to an analytic function for all t such that, if s is an exceptional point, then the projection corresponding to $\lambda_i(s)$ is given by $\sum_{j, \lambda_j(s) = \lambda_i(s)} P_j(s)$. By continuity, $\text{Tr } P_i(t)$ is a constant, we denote it by m_i . If s is not an exceptional point, m_i is the multiplicity of $\lambda_i(s)$.

By differentiating the equation $\text{Tr } \rho(s)P_i(s) = m_i\lambda_i(s)$ one obtains

$$\lambda'_i(s) = -\frac{1}{m_i} \text{Tr } \rho_2 P_i(s). \quad (8)$$

It follows that $\lambda_i(s)$ is nonincreasing for all s , moreover, $\lambda'_i(s) = 0$ implies that $\rho_2 P_i(s) = 0$, so that $\rho(t)P_i(s) = \rho(s)P_i(s) = \lambda_i(s)P_i(s)$ for all t and $\lambda_i(s)$ is an eigenvalue of $\rho(t)$ for all t . Hence λ_i is either strictly decreasing or a constant, which must be nonzero, since we assumed that $\rho_1 + \rho_2$ is invertible. It follows that each λ_i hits 0 at most once, so that there is only $l \leq N$ points where $\lambda_i(t) = 0$ for some i . Let us denote the points by $0 \leq t_1 < \dots < t_l$, it is clear that these are exactly the points where $P_{t,0} \neq 0$. Let $J_j := \{i, \lambda_i(t_j) > 0\}$, $j = 1, \dots, l$. Then

$J_j \subset J_{j-1}$ and

$$f_{\mathcal{E}}(t) = \sum_{i \in J_j} m_i \lambda_i(t), \quad t \in \langle t_{j-1}, t_j \rangle, \quad j = 1, \dots, l.$$

This implies (ii). □

REFERENCES

- [1] P. Alberti and A. Uhlmann: A problem relating to positive linear map on matrix algebras, *Rep. Math. Phys.* **18**, (1980), 163–176.
- [2] D. Blackwell: Comparison of experiments, *Proc. 2nd Berkeley Symp. on Math. Stat. and Probab.*, 93–102 (1951).
- [3] F. Buscemi: Comparison of quantum statistical models: equivalent conditions for sufficiency, *Commun. Math. Phys.* **310**, (2012), 625–647, arXiv:1004.3794.
- [4] C. E. Helstrom: *Quantum Detection and Estimation Theory*, Academic Press, New York 1976.
- [5] A. Holevo: *Statistical Structure of Quantum Theory*, Springer-Verlag Berlin Heidelberg (2001)
- [6] A. Jenčová: Quantum hypothesis testing and sufficient subalgebras, *Lett. Math. Phys.* **93**, (2010), 15–27.
- [7] T. Kato: *Perturbation Theory for Linear Operators*, Springer 1995.
- [8] K. Matsumoto: A quantum version of randomization condition, arXiv:1012.2650.
- [9] M. Ohya and D. Petz, *Quantum Entropy and its Use*, Springer, Heidelberg (1993).
- [10] V. Paulsen: *Completely Bounded Maps and Operator Algebras*, Cambridge University Press, 2002.
- [11] D. Petz: Sufficiency of channels over von Neumann algebras, *Quart. J. Math. Oxford* **39** (1988), 97–108.
- [12] D. Reeb, M. J. Kastoryano and M. M. Wolf: Hilbert’s projective metric in quantum information theory, *J. Math. Phys.* **52**, (2011) 082201.
- [13] S. Sherman: On a theorem of Hardy, Littlewood, Pólya and Blackwell, *Proc. Nat. Acad. Sciences* **37**, (1951), 826–831.
- [14] E. Shmaya: Comparison of information structures and completely positive maps, *J. Phys. A: Math. and Gen.* **38**, (2005), i9717–9127.
- [15] C. Stein: *Notes on a Seminar on Theoretical Statistics. I. Comparison of Experiments*, University of Chicago 1951.
- [16] H. Strasser: *Mathematical Theory of Statistics*, de Gruyter, Berlin, New York 1985.
- [17] E. Torgersen: Comparison of statistical experiments when the parameter space is finite, *Z. Wahrscheinlichkeitstheorie verw. geb.* **16**, (1970), 219–249.
- [18] E. Torgersen, *Comparison of Statistical Experiments*, Cambridge University Press 1991.