PROBABILITY
AND
MATHEMATICAL STATISTICS
Vol. 35, Fasc. 2 (2015), pp. 313–324

ON A GENERAL CONCEPT OF SUFFICIENCY IN VON NEUMANN ALGEBRAS*

BY

ANDRZEJ ŁUCZAK (ŁÓDŹ)

Abstract. A general question about the sufficiency of a subalgebra of some bigger algebra in the general operator algebra framework, under the assumption that the subalgebra in question is complete with respect to a family of states, is considered. Two particular cases are dealt with: sufficiency for Bayesian discrimination and sufficiency for unbiased estimation with minimal variance. It turns out that in both cases sufficiency is equivalent to the existence of a map from the bigger algebra into the smaller one having some specific properties.

2010 AMS Mathematics Subject Classification: Primary: 46L53; Secondary: 81P45, 62B15.

Key words and phrases: Sufficient subalgebras, normal states, Bayesian discrimination, unbiased estimation.

1. INTRODUCTION

In this paper, we want to address some aspects of the fundamental question: in what sense can a von Neumann subalgebra of a von Neumann algebra be considered *sufficient*? From the general point of view it seems reasonable to treat the smaller algebra as such if it is able to serve the same purposes as the bigger one, however, these purposes should be specified more concretely. For example, if we are given a family of normal states $\{\rho_\theta : \theta \in \Theta\}$ on the von Neumann algebra $\mathcal{M}$, a von Neumann subalgebra $\mathcal{N}$ of $\mathcal{M}$ can be called *sufficient* if the values of ρ_θ on $\mathcal{N}$ determine the values of ρ_θ on the whole of $\mathcal{M}$. This can be achieved by assuming the existence of a normal map $\alpha\colon \mathcal{M} \to \mathcal{N}$ such that $\rho_\theta = \rho_\theta \circ \alpha$ for all θ, and indeed this is the idea of the most popular definitions of sufficiency (for technical reasons, in these definitions it is also assumed that α is unital and positive – simple sufficiency; α is unital and completely, or at least two-, positive – Petz's sufficiency; or that α is a conditional expectation – Umegaki's sufficiency). But suppose that we have a finite number of states $\rho_1, \ldots, \rho_r$, and we want to

* Research supported by NCN grant No. 2011/01/B/ST1/03994.

discriminate them. Then the subalgebra $\mathcal{N}$ will be *sufficient* if we can find there a measurement which discriminates these states as efficiently as a measurement in $\mathcal{M}$. This idea of sufficiency was introduced by Jenčová in [8] under a somewhat unfortunate name of *two-sufficiency* (the reason for the name was considering only two states). It seems natural in this case to treat $\mathcal{N}$ as sufficient for discrimination. Another problem consists in finding for an arbitrary unbiased estimator $T \in \mathcal{M}$ of a function g of parameter θ an unbiased estimator $S \in \mathcal{N}$ of g whose variation is not greater than that of T. Again, in this case it seems natural to treat $\mathcal{N}$ as sufficient for unbiased estimation with minimal variation. As it will turn out, these two modes of sufficiency also lead to the existence of a map α (as mentioned earlier) with some specific properties. In the course of our analysis, we assume that the subalgebra $\mathcal{N}$ is complete with respect to the family $\{\rho_\theta : \theta \in \Theta\}$. It would be a challenging task to obtain similar results without this somewhat restrictive assumption.

2. PRELIMINARIES AND NOTATION

Let $\mathcal{M}$ be an arbitrary von Neumann algebra with identity $\mathbf{1}$ and predual $\mathcal{M}_*$. By $\mathcal{M}^h$ we denote the selfadjoint part of $\mathcal{M}$, and by $\mathcal{M}_*^h$ – the space of Hermitian functionals in $\mathcal{M}_*$. $\mathcal{M}^h$ and $\mathcal{M}_*^h$ are real Banach spaces, and we have $\left(\mathcal{M}_*^h\right)^* = \mathcal{M}^h$.

Let $\{\rho_\theta : \theta \in \Theta\}$ be a family of normal states on $\mathcal{M}$. The couple $(\mathcal{M}, \{\rho_\theta : \theta \in \Theta\})$ is usually called a *quantum statistical model* or a *quantum statistical experiment*. If we agree that $\mathcal{M}$ represents the bounded observables of a physical system, and ρ_θ are the possible states of this system, then our task consists in finding in an optimal way the state in which the system really is. One of possible approaches to this problem can be taken *via* Quantum Statistical Decision Theory. To briefly describe this approach, assume that we are given a finite number of states $\rho_1, \ldots, \rho_r$ which can occur with probabilities $\pi_1, \ldots, \pi_r$. To make a decision, we perform a measurement $\boldsymbol{M} = (M_1, \ldots, M_r)$ by which we mean positive operators in $\mathcal{M}$ such that $\sum_{j=1}^r M_j = \mathbf{1}$, and after obtaining the outcome M_j we decide that the true state is ρ_j. If the true state is ρ_i, then the probability of obtaining M_j equals $\rho_i(M_j)$. If our guess is j (i.e., the state ρ_j) while the genuine state is ρ_i, then we pay a penalty $L(i,j)$. Thus we have, with fixed *a priori* probability distribution $\pi = (\pi_1, \ldots, \pi_r)$ and *loss function* L, the following *risk function* $R_{\boldsymbol{M}} \colon \{1, \ldots, r\} \to \boldsymbol{R}$ defined as

$$R_{\boldsymbol{M}}(i) = \sum_{j=1}^{r} L(i,j)\rho_i(M_j). \tag{2.1}$$

(Note that $R_{\boldsymbol{M}}(i)$ is our expected loss when the true state is ρ_i.) The expectation of the risk function with respect to the *a priori* probability distribution π is called the *Bayes risk*, and denoted by $\mathbf{r}(\boldsymbol{M})$, i.e.,

$$\mathbf{r}(\boldsymbol{M}) = \sum_{i=1}^{r} \sum_{j=1}^{r} \pi_i L(i,j)\rho_i(M_j). \tag{2.2}$$

Now our aim is to find a measurement $\boldsymbol{M}$ that minimizes the Bayes risk. (An alternative approach, taken for instance in [3], is to treat L not as a loss function but as a *payoff function* in which case formulas (2.1) and (2.2) define expected gains rather than expected losses, and we want to maximize $\mathbf{r}(\boldsymbol{M})$ over all possible measurements $\boldsymbol{M}$.)

The most important particular case of the above general scheme is when we consider the loss function $L(i,j) = 1 - \delta_{ij}$, the penalty is zero when we guess correctly, and one when not. Then

$$\mathbf{r}(\boldsymbol{M}) = \sum_{i \neq j} \pi_i \rho_i(M_j) = 1 - \sum_{i=1}^{r} \pi_i \rho_i(M_i).$$

The term

$$\boldsymbol{P}_D(\boldsymbol{M}) = \sum_{i=1}^{r} \pi_i \rho_i(M_i)$$

represents the probability of correct detection of the true state while performing measurement $\boldsymbol{M} = (M_1, \ldots, M_r)$, so minimizing the Bayes risk is the same as maximizing the probability of correct detection. The quantity

$$\boldsymbol{P}_D = \sup_{\boldsymbol{M}} \sum_{i=1}^{r} \pi_i \rho_i(M_i)$$

is often regarded as the measure of *distinguishability* of the states $\rho_1, \ldots, \rho_r$ occurring with *a priori* probabilities $\pi_1, \ldots, \pi_r$ (see [1], [2], [4], [6], [7] for a more thorough description of these and other problems of Quantum Statistical Decision Theory).

REMARK 2.1. *From the point of view of comparing statistical models, the above setup was slightly generalized so as to allow measurements $(M_1, \ldots, M_k)$ of arbitrary length, and accordingly loss (payoff) functions $L(i,j), i = 1, \ldots, r$, $j = 1, \ldots, k$, in which case we have*

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

However, considering Bayesian discrimination, it seems more appropriate to have the number of states and the number of elements of the underlying measurement equal (we choose the state ρ_j when the outcome of the measurement is M_j), so we adopt this more natural setup.

REMARK 2.2. *It can easily be shown that the set of all measurements of fixed length is compact in an appropriate topology, thus all the "sup" and "inf" above and in the sequel can be replaced by "max" and "min", respectively. We shall not make use of this fact.*

Now, let $(\mathcal{M}, \{\rho_\theta : \theta \in \Theta\})$ be a quantum statistical model, and let $\mathcal{N}$ be a von Neumann subalgebra of $\mathcal{M}$. We shall say that $\mathcal{N}$ is *sufficient for (Bayesian) discrimination* if for each finite collection $\{\rho_{\theta_1}, \ldots, \rho_{\theta_r}\}$ of states, every *a priori* probability distribution $\pi = (\pi_1, \ldots, \pi_r)$ on these states, and every loss function L, we have

$$\inf_{\boldsymbol{N}} \mathbf{r}(\boldsymbol{N}) = \inf_{\boldsymbol{M}} \mathbf{r}(\boldsymbol{M}),$$

where $\boldsymbol{N} = (N_1, \ldots, N_r)$ is a measurement in $\mathcal{N}$. Considering the relation above for arbitrary functions L, it is clear that we can as well incorporate the π_i into L; consequently, for arbitrary real numbers $L(i,j)$, we shall be dealing with the relation

$$\inf_{\boldsymbol{N}} \sum_{i=1}^{r} \sum_{j=1}^{r} L(i,j)\rho_i(N_j) = \inf_{\boldsymbol{M}} \sum_{i=1}^{r} \sum_{j=1}^{r} L(i,j)\rho_i(M_j), \tag{2.3}$$

where $\boldsymbol{M} = (M_1, \ldots, M_r)$ and $\boldsymbol{N} = (N_1, \ldots, N_r)$ are measurements in $\mathcal{M}$ and $\mathcal{N}$, respectively.

Let now $g\colon \Theta \to \boldsymbol{R}$ be any function of θ. A selfadjoint element $T \in \mathcal{M}$ is said to be an *unbiased estimator of* g if for all θ we have $\rho_\theta(T) = g(\theta)$. The variance of T in the state ρ_θ is defined as

$$\boldsymbol{D}_\theta^2 T = \rho_\theta\Big(\big(T - \rho_\theta(T)\mathbf{1}\big)^2\Big) = \rho_\theta(T^2) - \rho_\theta(T)^2.$$

We shall say that $\mathcal{N}$ is *sufficient for unbiased estimation with minimal variance* if for every function g and each unbiased estimator $T \in \mathcal{M}$ of g there is an unbiased estimator $S \in \mathcal{N}$ of g such that $\boldsymbol{D}_\theta^2 S \leqslant \boldsymbol{D}_\theta^2 T$. It is immediately seen that $\mathcal{N}$ is sufficient for unbiased estimation with minimal variance if and only if for each selfadjoint $T \in \mathcal{M}$ there is a selfadjoint $S \in \mathcal{N}$ such that

$$\rho_\theta(S) = \rho_\theta(T) \quad \text{and} \quad \rho_\theta(S^2) \leqslant \rho_\theta(T^2) \quad \text{for all } \theta \in \Theta.$$

The following notion is well known in classical statistics. A subalgebra $\mathcal{N}$ of $\mathcal{M}$ is said to be *complete* with respect to the family of states $\{\rho_\theta : \theta \in \Theta\}$ if for any $S \in \mathcal{N}$ the equality $\rho_\theta(S) = 0$ for all θ yields $S = 0$. It is clear that completeness is equivalent to the separation of the points of $\mathcal{N}$ by the family $\{\rho_\theta : \theta \in \Theta\}$. From the Hahn–Banach theorem we infer also that the linear space $\mathrm{Lin}(\{\rho_\theta|\mathcal{N} : \theta \in \Theta\})$ of all linear combinations of the $\rho_\theta|\mathcal{N}$ is dense in $\mathcal{N}_*$, and the *real* linear space $\mathrm{Lin}_{\boldsymbol{R}}(\{\rho_\theta|\mathcal{N} : \theta \in \Theta\})$ of all linear combinations of the $\rho_\theta|\mathcal{N}$ with *real* coefficients is dense in $\mathcal{N}_*^h$.

3. COMPARISON OF CHANNELS

Let $\mathcal{A}$, $\mathcal{M}$, $\mathcal{N}$ be arbitrary von Neumann algebras, and let $\mathcal{E}$ and $\mathcal{F}$ be quantum channels transferring states of $\mathcal{A}$ to states of $\mathcal{M}$ and $\mathcal{N}$, respectively, i.e., $\mathcal{E}$ and $\mathcal{F}$

are linear maps such that

$$\mathcal{E}\colon \mathcal{A}_* \to \mathcal{M}_*, \qquad \mathcal{F}\colon \mathcal{A}_* \to \mathcal{N}_*,$$

and for each normal state φ on $\mathcal{A}$, $\mathcal{E}\varphi$ is a normal state on $\mathcal{M}$, and $\mathcal{F}\varphi$ is a normal state on $\mathcal{N}$. Following [5], the channel $\mathcal{E}$ is said to be *less noisy* than $\mathcal{F}$, denoted by $\mathcal{F} \prec \mathcal{E}$, if for an arbitrary collection of states $\varphi_1, \ldots, \varphi_r$ in $\mathcal{A}_*$ the states $\mathcal{E}\varphi_1, \ldots, \mathcal{E}\varphi_r$ are more distinguishable than the states $\mathcal{F}\varphi_1, \ldots, \mathcal{F}\varphi_r$ for any *a priori* probability distribution $(\pi_1, \ldots, \pi_r)$. Thus, $\mathcal{F} \prec \mathcal{E}$ if

$$\sup_{\boldsymbol{N}} \sum_{i=1}^{r} \pi_i(\mathcal{F}\varphi_i)(N_i) \leqslant \sup_{\boldsymbol{M}} \sum_{i=1}^{r} \pi_i(\mathcal{E}\varphi_i)(M_i) \tag{3.1}$$

for arbitrary states $\varphi_1, \ldots, \varphi_r$ in $\mathcal{A}_*$ and arbitrary *a priori* probability distribution $(\pi_1, \ldots, \pi_r)$.

LEMMA 3.1. *$\mathcal{F} \prec \mathcal{E}$ if and only if for arbitrary functionals $\varphi_1, \ldots, \varphi_r$ in $\mathcal{A}_*^h$ we have*

$$\sup_{\boldsymbol{N}} \sum_{i=1}^{r} (\mathcal{F}\varphi_i)(N_i) \leqslant \sup_{\boldsymbol{M}} \sum_{i=1}^{r} (\mathcal{E}\varphi_i)(M_i). \tag{3.2}$$

P r o o f. Assume that $\mathcal{F} \prec \mathcal{E}$. First we shall show that relation (3.2) holds for arbitrary positive functionals $\varphi_1, \ldots, \varphi_r$ in $\mathcal{A}_*$. We may assume that $\varphi_i \neq 0$. Then $\varphi_i/\|\varphi_i\|$ are states. Put

$$c = \frac{1}{\|\varphi_1\| + \ldots + \|\varphi_r\|}, \qquad \pi_i = c\|\varphi_i\|, \qquad i = 1, \ldots, r.$$

Then $(\pi_1, \ldots, \pi_r)$ is a probability distribution, and we have

$$\sum_{i=1}^{r} (\mathcal{E}\varphi_i)(M_i) = \frac{1}{c} \sum_{i=1}^{r} \pi_i \Big(\mathcal{E}\Big(\frac{\varphi_i}{\|\varphi_i\|}\Big)\Big)(M_i),$$
$$\sum_{i=1}^{r} (\mathcal{F}\varphi_i)(N_i) = \frac{1}{c} \sum_{i=1}^{r} \pi_i \Big(\mathcal{F}\Big(\frac{\varphi_i}{\|\varphi_i\|}\Big)\Big)(N_i),$$

showing that relation (3.2) follows from (3.1).

Let now $\varphi_1, \ldots, \varphi_r$ be arbitrary functionals in $\mathcal{A}_*^h$. Let φ_0 be a positive functional in $\mathcal{A}_*$ such that $\varphi_0 + \varphi_i \geqslant 0$ for all $i = 1, \ldots, r$. We have

$$\begin{aligned}\sum_{i=1}^{r} \big(\mathcal{E}(\varphi_0 + \varphi_i)\big)(M_i) &= (\mathcal{E}\varphi_0)(\mathbf{1}) + \sum_{i=1}^{r} (\mathcal{E}\varphi_i)(M_i)\\ &= \varphi_0\big(\mathcal{E}^*(\mathbf{1})\big) + \sum_{i=1}^{r} (\mathcal{E}\varphi_i)(M_i) = \varphi_0(\mathbf{1}) + \sum_{i=1}^{r} (\mathcal{E}\varphi_i)(M_i)\end{aligned}$$

since $\mathcal{E}^*(\mathbf{1}) = \mathbf{1}$, and by the same token

$$\sum_{i=1}^{r} \big(\mathcal{F}(\varphi_0 + \varphi_i)\big)(N_i) = \varphi_0(\mathbf{1}) + \sum_{i=1}^{r} (\mathcal{F}\varphi_i)(N_i).$$

Since the functionals $\varphi_0 + \varphi_i$ are positive, we infer that relation (3.2) for positive functionals yields the same relation for Hermitian ones.

It is obvious that relation (3.2) for arbitrary functionals $\varphi_1, \ldots, \varphi_r$ in $\mathcal{A}_*^h$ implies $\mathcal{F} \prec \mathcal{E}$. ■

It turns out that (3.1) is a quite strong relation even if it holds only for two states.

PROPOSITION 3.1. *Assume that for arbitrary states φ_1, φ_2 in $\mathcal{A}_*$ and arbitrary* a priori *probability distribution (π_1, π_2) the inequality*

$$\sup_{\boldsymbol{N}} \sum_{i=1}^{2} \pi_i (\mathcal{F}\varphi_i)(N_i) \leqslant \sup_{\boldsymbol{M}} \sum_{i=1}^{2} \pi_i (\mathcal{E}\varphi_i)(M_i)$$

holds. Then for each functional φ in $\mathcal{A}_*^h$ we have

$$\|\mathcal{F}\varphi\| \leqslant \|\mathcal{E}\varphi\|.$$

P r o o f. As in Lemma 3.1, we may assume that for arbitrary functionals φ_1, φ_2 in $\mathcal{A}_*^h$ we have

$$\sup_{\boldsymbol{N}} \sum_{i=1}^{2} (\mathcal{F}\varphi_i)(N_i) \leqslant \sup_{\boldsymbol{M}} \sum_{i=1}^{2} (\mathcal{E}\varphi_i)(M_i). \tag{3.3}$$

Let $\varphi \in \mathcal{A}_*^h$, and put $\varphi_1 = \varphi$, $\varphi_2 = -\varphi$. Then for any measurement (M_1, M_2) we have

$$(\mathcal{E}\varphi_1)(M_1) + (\mathcal{E}\varphi_2)(M_2) = (\mathcal{E}\varphi)(M_1 - M_2) \leqslant \|\mathcal{E}\varphi\| \tag{3.4}$$

since $\|M_1 - M_2\| \leqslant 1$. On the other hand, for the Jordan decomposition

$$\mathcal{E}\varphi = (\mathcal{E}\varphi)^+ - (\mathcal{E}\varphi)^-,$$

we have

$$\|\mathcal{E}\varphi\| = \|(\mathcal{E}\varphi)^+\| + \|(\mathcal{E}\varphi)^-\|;$$

moreover, the positive functionals $(\mathcal{E}\varphi)^+$ and $(\mathcal{E}\varphi)^-$ have orthogonal supports. Let E denote the support of $(\mathcal{E}\varphi)^+$. Then

$$\|(\mathcal{E}\varphi)^+\| = (\mathcal{E}\varphi)^+(E), \qquad \|(\mathcal{E}\varphi)^-\| = (\mathcal{E}\varphi)^-(\mathbf{1} - E).$$

Consequently,

$$(\mathcal{E}\varphi)(E) = (\mathcal{E}\varphi)^+(E) - (\mathcal{E}\varphi)^-(E) = \|(\mathcal{E}\varphi)^+\|,$$

and

$$\big(\mathcal{E}(-\varphi)\big)(\mathbf{1} - E) = (\mathcal{E}\varphi)^-(\mathbf{1} - E) - (\mathcal{E}\varphi)^+(\mathbf{1} - E) = \|(\mathcal{E}\varphi)^-\|.$$

Thus for the measurement $(E, \mathbf{1} - E)$ we obtain

$$(\mathcal{E}\varphi)(E) + \big(\mathcal{E}(-\varphi)\big)(\mathbf{1} - E) = \|(\mathcal{E}\varphi)^+\| + \|(\mathcal{E}\varphi)^-\| = \|\mathcal{E}\varphi\|,$$

which together with inequality (3.4) shows that

$$\sup_{\boldsymbol{M}} \sum_{i=1}^{2} (\mathcal{E}\varphi_i)(M_i) = \|\mathcal{E}\varphi\|.$$

In the same way we get

$$\sup_{\boldsymbol{N}} \sum_{i=1}^{2} (\mathcal{F}\varphi_i)(N_i) = \|\mathcal{F}\varphi\|,$$

so relation (3.3) yields the inequality $\|\mathcal{F}\varphi\| \leqslant \|\mathcal{E}\varphi\|$. ■

4. CHARACTERIZATION OF SUFFICIENCY FOR BAYESIAN DISCRIMINATION

In this section, we characterize the sufficiency of a von Neumann subalgebra for Bayesian discrimination.

THEOREM 4.1. *Let $\{\rho_\theta : \theta \in \Theta\}$ be a family of normal states on a von Neumann algebra $\mathcal{M}$, and let $\mathcal{N}$ be a von Neumann subalgebra of $\mathcal{M}$ complete with respect to $\{\rho_\theta : \theta \in \Theta\}$. Then $\mathcal{N}$ is sufficient for Bayesian discrimination if and only if there is a normal conditional expectation α from $\mathcal{M}$ onto $\mathcal{N}$ such that*

$$\rho_\theta \circ \alpha = \rho_\theta \quad \textit{for all } \theta \in \Theta. \tag{4.1}$$

P r o o f. Assume first that $\mathcal{N}$ is sufficient for Bayesian discrimination. Take an arbitrary fixed collection $\{\rho_{\theta_1}, \ldots, \rho_{\theta_r}\}$ of states. Let $\mathcal{A}$ be the abelian von Neumann algebra generated by a sequence $\{P_1, \ldots, P_r\}$ of pairwise orthogonal projections summing up to $\mathbf{1}$:

$$\mathcal{A} = \Big\{ \sum_{i=1}^{r} c_i P_i : c_i \in \boldsymbol{C},\ i = 1, \ldots, r \Big\}.$$

Each sequence $(a_1, \ldots, a_r)$ of complex numbers can obviously be identified with a linear functional φ on $\mathcal{A}$, and *vice versa*, by the equality $\varphi(P_i) = a_i$, i.e.,

$$\varphi\Big(\sum_{i=1}^{r} c_i P_i \Big) = \sum_{i=1}^{r} a_i c_i.$$

Define maps $\mathcal{E}$ and $\mathcal{F}$ from $\mathcal{A}_*$ to $\mathcal{M}_*$ and $\mathcal{N}_*$, respectively, as

$$\mathcal{E}\varphi = \sum_{i=1}^{r} a_i \rho_{\theta_i}, \qquad \mathcal{F}\varphi = \sum_{i=1}^{r} a_i (\rho_{\theta_i}|\mathcal{N})$$

for $\varphi = (a_1, \dots, a_r)$. It is easily seen that $\mathcal{E}$ and $\mathcal{F}$ are quantum channels.

Let us take arbitrary functionals $\varphi_1, \dots, \varphi_r$ in $\mathcal{A}_*^h$, and to conform to our previous notation denote them by $\varphi_j = \big(L(1,j), \dots, L(r,j)\big)$, $j = 1, \dots, r$, with $L(i,j) \in \boldsymbol{R}$. Then for every measurement $\boldsymbol{M} = (M_1, \dots, M_r)$ in $\mathcal{M}$ we have

$$\sum_{j=1}^{r} (\mathcal{E}\varphi_j)(M_j) = \sum_{j=1}^{r} \sum_{i=1}^{r} L(i,j) \rho_{\theta_i}(M_j),$$

and similarly for $\mathcal{F}$. Since $\mathcal{N}$ is sufficient, we have the relation

$$\inf_{\boldsymbol{N}} \sum_{j=1}^{r} \sum_{i=1}^{r} L(i,j) \rho_{\theta_i}(N_j) = \inf_{\boldsymbol{M}} \sum_{j=1}^{r} \sum_{i=1}^{r} L(i,j) \rho_{\theta_i}(M_j),$$

which means that

$$\inf_{\boldsymbol{N}} \sum_{j=1}^{r} (\mathcal{F}\varphi_j)(N_j) = \inf_{\boldsymbol{M}} \sum_{j=1}^{r} (\mathcal{E}\varphi_j)(M_j).$$

Proposition 3.1 yields the equality

$$\|\mathcal{F}\varphi\| = \|\mathcal{E}\varphi\|$$

for each $\varphi \in \mathcal{A}_*^h$, i.e.,

$$\Big\| \sum_{i=1}^{r} a_i (\rho_{\theta_i}|\mathcal{N}) \Big\| = \Big\| \sum_{i=1}^{r} a_i \rho_{\theta_i} \Big\|$$

for any $a_1, \dots, a_r \in \boldsymbol{R}$. Since $\rho_{\theta_1}, \dots, \rho_{\theta_r}$ were arbitrary, the relation above allows us to define an isometry

$$\mathcal{L}\colon\ \mathrm{Lin}_{\boldsymbol{R}}(\{\rho_\theta|\mathcal{N} : \theta \in \Theta\}) \to \mathcal{M}_*^h$$

by the formula

$$\mathcal{L}\Big(\sum_{i=1}^{r} a_i (\rho_{\theta_i}|\mathcal{N})\Big) = \sum_{i=1}^{r} a_i \rho_{\theta_i},$$

and this isometry is extended to the whole of $\mathcal{N}_*^h$. Put $\alpha = \mathcal{L}^*$. Then α is a linear normal map from $\mathcal{M}^h$ to $\mathcal{N}^h$ of norm one. For $T \in \mathcal{M}^h$, we have

$$\rho_\theta\big(\alpha(T)\big) = (\rho_\theta|\mathcal{N})\big(\alpha(T)\big) = \big(\mathcal{L}(\rho_\theta|\mathcal{N})\big)(T) = \rho_\theta(T),$$

showing that relation (4.1) holds on $\mathcal{M}^h$. For $S \in \mathcal{N}^h$, we obtain

$$\rho_\theta\big(\alpha(S)\big) = \rho_\theta(S),$$

and the completeness of $\mathcal{N}$ yields $\alpha(S) = S$. Thus α is a projection onto $\mathcal{N}^h$. In particular, α is unital, and the same reasoning as for maps on the whole algebra (not only on its selfadjoint part) shows that α is positive. Now extend α to the whole of $\mathcal{M}$. Then this extended α is a positive projection onto $\mathcal{N}$, and since $\alpha(\mathbf{1}) = \mathbf{1}$, the positivity yields $\|\alpha\| = 1$. Thus α is a (positive) projection onto $\mathcal{N}$ of norm one, i.e., a conditional expectation. It is obvious that relation (4.1) is satisfied.

Now assume that there is a conditional expectation α from $\mathcal{M}$ onto $\mathcal{N}$ satisfying condition (4.1). For each measurement (M_j) in $\mathcal{M}$, put $N_j = \alpha(M_j)$. Then (N_j) is a measurement in $\mathcal{N}$, and for any $\rho_{\theta_1}, \ldots, \rho_{\theta_r}$ and any loss function $L(i,j)$ we have

$$\sum_{i=1}^{r}\sum_{j=1}^{r} L(i,j)\rho_{\theta_i}(N_j) = \sum_{i=1}^{r}\sum_{j=1}^{r} L(i,j)\rho_{\theta_i}(M_j),$$

which shows that $\mathcal{N}$ is sufficient for Bayesian discrimination. ■

5. CHARACTERIZATION OF SUFFICIENCY FOR UNBIASED ESTIMATION WITH MINIMAL VARIANCE

Now we characterize the sufficiency of a von Neumann subalgebra for unbiased estimation with minimal variation.

THEOREM 5.1. *Let $\{\rho_\theta : \theta \in \Theta\}$ be a family of normal states on a von Neumann algebra $\mathcal{M}$, and let $\mathcal{N}$ be a von Neumann subalgebra of $\mathcal{M}$ complete with respect to $\{\rho_\theta : \theta \in \Theta\}$. Then $\mathcal{N}$ is sufficient for unbiased estimation with minimal variance if and only if there is a linear bounded normal Hermitian projection α from $\mathcal{M}$ onto $\mathcal{N}$ such that*

(5.1) $$\rho_\theta \circ \alpha = \rho_\theta \quad \textit{for all } \theta \in \Theta,$$

and

(5.2) $$\alpha(T \circ S) = \alpha(T) \circ S \quad \textit{for all } T \in \mathcal{M},\ S \in \mathcal{N},$$

where "$\circ$" stands for the Jordan product defined as

$$T \circ S = \frac{TS + ST}{2}.$$

P r o o f. Assume that $\mathcal{N}$ is sufficient. Take an arbitrary $T \in \mathcal{M}^h$. There exists uniquely determined $S \in \mathcal{N}^h$ such that $\rho_\theta(S) = \rho_\theta(T)$. Indeed, if we had $\rho_\theta(T) =$

$\rho_\theta(S_1) = \rho_\theta(S_2)$, then from the completeness of $\mathcal{N}$ it would follow that $S_1 = S_2$. Thus we may define a map $\alpha\colon \mathcal{M}^h \to \mathcal{N}^h$ by the formula

$$\alpha(T) = S.$$

We have $\rho_\theta\big(\alpha(T)\big) = \rho_\theta(S) = \rho_\theta(T)$, and thus $\rho_\theta \circ \alpha = \rho_\theta$ for all $\theta \in \Theta$.

Let us take arbitrary $T_1, T_2 \in \mathcal{M}^h$ and put $\alpha(T_1) = S_1$, $\alpha(T_2) = S_2$, $\alpha(T_1 + T_2) = S_3$. Then

$$\begin{aligned}\rho_\theta(S_1 + S_2) &= \rho_\theta\big(\alpha(T_1)\big) + \rho_\theta\big(\alpha(T_2)\big)\\ &= \rho_\theta(T_1 + T_2) = \rho_\theta\big(\alpha(T_1 + T_2)\big) = \rho_\theta(S_3),\end{aligned}$$

showing that $S_1 + S_2 = S_3$, i.e., α is additive. In a similar way we show homogeneity, thus α is linear.

Now, let $\mathcal{N}^h \ni T_n \to T \in \mathcal{N}^h$ and $\alpha(T_n) \to S$. We have

$$\rho_\theta\big(\alpha(T_n)\big) \to \rho_\theta(S),$$

and, on the other hand,

$$\rho_\theta\big(\alpha(T_n)\big) = \rho_\theta(T_n) \to \rho_\theta(T) = \rho_\theta\big(\alpha(T)\big),$$

showing that $\rho_\theta(S) = \rho_\theta\big(\alpha(T)\big)$, and, consequently, $S = \alpha(T)$. This means that α is closed, and thus bounded.

For each $S \in \mathcal{N}$, we have $\rho_\theta\big(\alpha(S)\big) = \rho_\theta(S)$, thus $\alpha(S) = S$, which shows that α is a projection. We extend α in an obvious way from $\mathcal{N}^h$ to $\mathcal{N}$.

For a bounded linear map $\alpha^*\colon \mathcal{N}^* \to \mathcal{M}^*$, we have $\alpha^*(\rho_\theta|\mathcal{N}) = \rho_\theta$, hence $\alpha^*(\mathrm{Lin}\{\rho_\theta|\mathcal{N} : \theta \in \Theta\}) \subset \mathcal{M}_*$. Since $\mathrm{Lin}\{\rho_\theta|\mathcal{N} : \theta \in \Theta\}$ is a dense subspace of $\mathcal{N}_*$, we obtain $\alpha^*(\mathcal{N}_*) \subset \mathcal{M}_*$. Now, putting $\alpha^*|\mathcal{N}_* = \alpha_*$, it is easily seen that $\alpha = (\alpha_*)^*$, which means that α is normal.

Observe that the condition of decreasing variance has now the form

$$\rho_\theta(T^2) \geqslant \rho_\theta(S^2) = \rho_\theta\big(\alpha(T)^2\big).$$

Define, for each $\theta \in \Theta$, a bilinear form on $\mathcal{M}^h$ by the formula

$$[T, S]_\theta = \rho_\theta\big(\alpha(T \circ S) - \alpha(T) \circ \alpha(S)\big), \qquad T, S \in \mathcal{M}^h.$$

We have, for each $T \in \mathcal{M}^h$,

$$[T, T]_\theta = \rho_\theta\big(\alpha(T^2) - \alpha(T)^2\big) = \rho_\theta(T^2) - \rho_\theta\big(\alpha(T)^2\big) \geqslant 0,$$

so the form is positive. The Schwarz inequality yields

$$|[T, S]_\theta|^2 \leqslant \Big(\rho_\theta(T^2) - \rho_\theta\big(\alpha(T)^2\big)\Big)\Big(\rho_\theta(S^2) - \rho_\theta\big(\alpha(S)^2\big)\Big);$$

in particular, for $S \in \mathcal{N}^h$, we have $\big(\rho_\theta(S^2) - \rho_\theta\big(\alpha(S)^2\big)\big) = 0$, thus $[T, S]_\theta = 0$ for any $T \in \mathcal{M}^h$, $S \in \mathcal{N}^h$. This means that

$$0 = [T, S]_\theta = \rho_\theta\big(\alpha(T \circ S) - \alpha(T) \circ \alpha(S)\big) = \rho_\theta\big(\alpha(T \circ S) - \alpha(T) \circ S\big),$$

showing the equality

$$\alpha(T \circ S) = \alpha(T) \circ S \quad \text{for } T \in \mathcal{M}^h,\ S \in \mathcal{N}^h.$$

It follows that the above equality holds for $T \in \mathcal{M}$ and $S \in \mathcal{N}$, so α has all the desired properties.

Assume now that there is a linear bounded normal Hermitian projection α from $\mathcal{M}$ onto $\mathcal{N}$ satisfying (5.1) and (5.2). For each θ, define on $\mathcal{M}^h$ a positive bilinear form

$$\langle T, S\rangle_\theta = \rho_\theta(T \circ S).$$

For an arbitrary $T \in \mathcal{M}^h$, put $S = \alpha(T)$. Then $S \in \mathcal{N}^h$, $\rho_\theta(S) = \rho_\theta(T)$, and from the Schwarz inequality we obtain

$$\begin{aligned}\rho_\theta\big(\alpha(T)^2\big) &= \rho_\theta\Big(\alpha\big(T \circ \alpha(T)\big)\Big) = \rho_\theta\big(T \circ \alpha(T)\big)\\ &= \langle T, \alpha(T)\rangle_\theta \leqslant \sqrt{\rho_\theta(T^2)}\sqrt{\rho_\theta\big(\alpha(T)^2\big)},\end{aligned}$$

showing that

$$\rho_\theta(S^2) = \rho_\theta\big(\alpha(T)^2\big) \leqslant \rho_\theta(T^2),$$

which means that S is the desired estimator having variance not greater than that of T. ∎

REMARK 5.1. *While it is obvious that α in the theorem above is unital, apparently it need not be of norm one or, equivalently, positive. If it were, then being a projection of norm one onto a von Neumann subalgebra, α would be a conditional expectation – a condition which seems to be considerably stronger than that required of α in formula* (5.2).

6. CONCLUDING REMARKS

Theorems 4.1 and 5.1 give nice descriptions of sufficient subalgebras for Bayesian discrimination and unbiased estimation with minimal variance in the case when the subalgebra in question is complete. This assumption, although being a counterpart of a classical notion, seems quite restrictive, and it would be desirable to obtain similar descriptions of sufficiency without it. On the other hand, the notion of completeness turned out to be fruitful in the investigations of sufficiency performed in [9] where it was proved, for example, that under this condition

Umegaki's sufficiency, Petz's sufficiency and simple sufficiency are all equivalent ([9], Theorem 3) or that completeness plus sufficiency yield the minimality of the subalgebra ([9], Theorem 4). Also in this context, a quantum version of Basu's theorem was obtained ([9], Theorem 5). It is hoped that further investigations in the field of sufficiency will clarify the possibility of giving up, or at least weakening, the completeness assumption.

REFERENCES

[1] S. M. Barnett and S. Croke, *Quantum state discrimination*, Adv. Opt. Photon. 1 (2009), pp. 238–278; preprint, arXiv: 0810.1970.

[2] J. Bergou, U. Herzog, and M. Hillery, *Discrimination of quantum states*, in: *Quantum State Estimation*, M. G. A. Paris and J. Řeháček (Eds.), Lecture Notes in Phys., Vol. 649, Springer, Berlin 2004, pp. 417–465.

[3] F. Buscemi, *Comparison of quantum statistical models: equivalent conditions for sufficiency*, Comm. Math. Phys. 310 (2012), pp. 625–647.

[4] A. Chefles, *Quantum state discrimination*, Contemp. Phys. 41 (2000), pp. 401–430.

[5] A. Chefles, *The quantum Blackwell theorem and minimum error state discrimination*, preprint, arXiv: 0907.0866.

[6] C. W. Helstrom, *Quantum Detection and Estimation Theory*, Academic Press, New York–San Francisco–London 1976.

[7] A. S. Holevo, *Statistical decision theory for quantum systems*, J. Multivariate Anal. 3 (1973), pp. 337–394.

[8] A. Jenčová, *Quantum hypothesis testing and sufficient subalgebras*, Lett. Math. Phys. 93 (2010), pp. 15–27; preprint, arXiv: 0810.4045.

[9] A. Łuczak, *Quantum sufficiency in the operator algebra framework*, Internat. J. Theoret. Phys. 53 (2014), pp. 3423–3433.

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

Received on 28.3.2014;
revised version on 3.1.2015