Sufficiency in quantum statistical inference

September 30, 2018

Anna Jenčová[1]
Mathematical Institute of the
Slovak Academy of Sciences
Stefanikova 49, Bratislava, Slovakia

Dénes Petz[2]
Alfreéd Rényi Institute of Mathematics
Hungarian Academy of Sciences
POB 127, H-1364 Budapest, Hungary

This paper attempts to develop a theory of sufficiency in the setting of non-commutative algebras parallel to the ideas in classical mathematical statistics. Sufficiency of a coarse-graining means that all information is extracted about the mutual relation of a given family of states. In the paper sufficient coarse-grainings are characterized in several equivalent ways and the non-commutative analogue of the factorization theorem is obtained. Among the applications the equality case for the strong subadditivity of the von Neumann entropy, the Imoto-Koashi theorem and exponential families are treated. The setting of the paper allows the underlying Hilbert space to be infinite dimensional.

MSC: 46L53, 81R15, 62B05.

Key words: Quantum statistics, coarse-graining, factorization theorem, exponential family, strong subadditivity of entropy, sufficient subalgebra

1 Introduction and preliminaries

A quantum mechanical system is described by a C*-algebra, the dynamical variables (or observables) correspond to the self-adjoint elements and the physical state of the system are modelled by the normalized positive functionals of the algebra, see [4, 5]. The evolution of the system $\mathcal{M}$ can be described in the **Heisenberg picture** in which an observable $A \in \mathcal{M}$ moves into $\alpha(A)$, where α is a linear transformation. α is an automorphism in case of the time evolution of a closed system but it could be the irreversible evolution of an open system. The **Schrödinger picture** is dual, it gives the transformation of the states, the state $\varphi \in \mathcal{M}^*$ moves into $\varphi \circ \alpha$. The algebra of a quantum system is typically non-commutative but the mathematical formalism brings commutative algebras as well. A simple **measurement** is usually modelled by a family of pairwise orthogonal projections, or more generally, by a partition of unity, $(E_i)_{i=1}^n$. Since all E_i are supposed to be positive and $\sum_i E_i = I$, $\beta : \mathbb{C}^n \to \mathcal{M}$, $(z_1, z_2, \ldots, z_n) \mapsto \sum_i z_i E_i$ gives a positive unital mapping from the commutative C*-algebra $\mathbb{C}^n$ to the non-commutative algebra $\mathcal{M}$. Every positive unital mapping occurs in this way. The essential concept in quantum information theory is the state transformation which is affine and the dual of a positive unital mapping. All these and several other situations justify to study of positive unital mappings between C*-algebras from a quantum statistical viewpoint.

If the algebra $\mathcal{M}$ is "small" and $\mathcal{N}$ is "large", and the mapping $\alpha : \mathcal{M} \to \mathcal{N}$ sends the state φ of the system of interest to the state $\varphi \circ \alpha$ at our disposal, then loss of information takes place and the problem of statistical inference is to reconstruct the real state from partial information. In this paper we mostly consider parametric statistical models, a parametric family $\mathcal{S} := \{\varphi_\theta : \theta \in \Theta\}$ of states are given and on the basis of the partial information the correct value of the parameter should be decided. If the partial information is the outcome of a measurement, then we have statistical inference in the very strong sense. However, there are "more" quantum situations, to decide between quantum states on the basis of quantum data, see Example 3 below. The problem we discuss is not the procedure of the decision about the true state of the system but we want to describe the circumstances under which this is perfectly possible.

The paper is organized as follows. In the rest of this section we summarize the relevant basic concepts both in classical statistics in the non-commutative framework. Section 2 is about sufficient subalgebras, or subsystems of a quantum system. Most of the result of this section has been known but we give a complete presentation and in our proof the operator algebraic methods are minimized. Section 3 is devoted to sufficient coarse-grainings. The importance of the multiplicative domain of a completely positive mapping is emphasized here. The factorization theorem of Section 4 is the main result of the paper. Section 5 connects the exponential families of the quantum setting to sufficiency problem. In Section 6 the equality case in the strong subadditivity of the von Neumann entropy is discussed in a possibly infinite dimensional framework and the factorization result is applied.

In this paper C*-algebras always have a unit I. Given a C*-algebra $\mathcal{M}$, a state φ of $\mathcal{M}$ is a linear function $\mathcal{M} \to \mathbb{C}$ such that $\varphi(I) = 1 = \|\varphi\|$. (Note that the second condition is equivalent to the positivity of φ.) The books [4, 5] – among many others – explain the basic facts about C*-algebras. The class of finite dimensional full matrix algebras form a small and algebraically rather trivial subclass of C*-algebras, but from the view-point of non-commutative statistics, almost all ideas and concepts appear in this setting. A matrix algebra $M_n(\mathbb{C})$ admits a canonical trace Tr and all states are described by their densities with respect to Tr. The correspondence is given by $\varphi(A) = \mathrm{Tr} D_\varphi A$ $(A \in M_n(\mathbb{C}))$ and we can simply identify the functional φ by the density D_φ. Note that the density is a positive (semi-definite) matrix of trace 1.

Let $\mathcal{M}$ and $\mathcal{N}$ be C*-algebras. Recall that **2-positivity** of $\alpha : \mathcal{M} \to \mathcal{N}$ means that

$$\left[\begin{array}{cc} \alpha(A) & \alpha(B) \\ \alpha(C) & \alpha(D) \end{array}\right] \geq 0 \qquad \text{if} \qquad \left[\begin{array}{cc} A & B \\ C & D \end{array}\right] \geq 0$$

for 2×2 matrices with operator entries. It is well-known that a 2-positive unit-preserving mapping α satisfies the **Schwarz inequality**

$$\alpha(A^*A) \geq \alpha(A)^*\alpha(A). \tag{1}$$

A 2-positive unital mapping between C*-algebras will be **coarse-graining**. Here are two fundamental examples.

Example 1 Let $\mathcal{X}$ be a finite set and $\mathcal{N}$ be a C*-algebra. Assume that for each $x \in \mathcal{X}$ a positive operator $E(x) \in \mathcal{N}$ is given and $\sum_x E(x) = I$. In quantum mechanics such a setting is a model for a measurement with values in $\mathcal{X}$.

The space $C(\mathcal{X})$ of function on $\mathcal{X}$ is a C*-algebra and the partition of unity E induces a coarse-graining $\alpha : C(\mathcal{X}) \to \mathcal{N}$ given by $\alpha(f) = \sum_x f(x)E(x)$. Therefore a coarse-graining defined on a commutative algebra is an equivalent way to give a measurement. (Note that the condition of 2-positivity is automatically fulfilled on a commutative algebra.) $\square$

Example 2 Let $\mathcal{M}$ be the algebra of all bounded operators acting on a Hilbert space $\mathcal{H}$ and let $\mathcal{N}$ be the infinite tensor product $\mathcal{M} \otimes \mathcal{M} \otimes \ldots$. (To understand the essence of the example one does not need the very formal definition of the infinite tensor product.) If γ denotes the right shift on $\mathcal{N}$, then we can define a sequence α_n of coarse-grainings $\mathcal{M} \to \mathcal{N}$:
$$\alpha_n(A) := \frac{1}{n}(A + \gamma(A) + \ldots + \gamma^{n-1}(A)).$$
α_n is the quantum analogue of the **sample mean**. $\square$

Let $(X_i, \mathcal{A}_i, \mu_i)$ be a measure space $(i = 1, 2)$. Recall that a positive linear map $M : L^\infty(X_1, \mathcal{A}_1, \mu_1) \to L^\infty(X_2, \mathcal{A}_2, \mu_2)$ is called a **Markov operator** if it satisfies $M1 = 1$ and $f_n \searrow 0$ implies $Mf_n \searrow 0$. For mappings defined between von Neumann

algebras, the monotone continuity is called **normality**. In case that $\mathcal{M}$ and $\mathcal{N}$ are von Neumann algebras, a coarse-graining $\mathcal{M} \to \mathcal{N}$ will be always supposed to be normal. Our concept of coarse-graining is the analogue of the Markov operator.

We mostly mean that a coarse-graining transforms observables to observables corresponding to the **Heisenberg picture** and in this case we assume that it is unit preserving. The dual of such a mapping acts on states or on density matrices and it will be called coarse-graining as well.

We recall some well-known results from mathematical statistics, see [23] for details.

Let $(X, \mathcal{A})$ be a measurable space and let $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$ be a set of probability measures on $(X, \mathcal{A})$. A sub-σ-algebra $\mathcal{A}_0 \subset \mathcal{A}$ is **sufficient** for $\mathcal{P}$ if for all $A \in \mathcal{A}$, there is an $\mathcal{A}_0$-measurable function f_A such that for all θ,

$$f_A = P_\theta(A|\mathcal{A}_0) \quad P_\theta - \text{almost everywhere,}$$

that is,

$$P_\theta(A \cap A_0) = \int_{A_0} f_A dP_\theta \tag{2}$$

for all $A_0 \in \mathcal{A}_0$ and for all θ. It is clear from this definition that if $\mathcal{A}_0$ is sufficient then for all P_θ there is a common version of the conditional expectation $E_\theta[g|\mathcal{A}_0]$ for any measurable step function g, or, more generally, for any function $g \in \cap_{\theta \in \Theta} L^1(X, \mathcal{A}, P_\theta)$.

In the most important case, the family $\mathcal{P}$ is **dominated**, that is there is a σ-finite measure μ such that $\mathcal{P} \ll \mu$. The following lemma is a useful tool in examining sufficiency.

Lemma 1 *If $\mathcal{P}$ is dominated, then there is a countable subset $\{P_1, P_2, \ldots\} \subseteq \mathcal{P}$ such that $P_\theta(A) = 0$ holds for all $\theta \in \Theta$ if and only if $P_n(A) = 0$ holds for all $n \in \mathbb{N}$.*

It follows that if $\mathcal{P}$ is dominated then there is a (possibly infinite) convex combination $P_0 = \sum_n c_n P_n$, $P_n \in \mathcal{P}$, such that $\mathcal{P} \equiv P_0$.

For our purposes, it is more suitable to use the following characterization of sufficiency in terms of randomization.

Let $\mathcal{P}_i = \{P_{i,\theta} : \theta \in \Theta\}$ be dominated families of probability measures on $(X_i, \mathcal{A}_i)$, such that $\mathcal{P}_i \equiv \mu_i$, $i = 1, 2$. We say that $(X_2, \mathcal{A}_2, \mathcal{P}_2)$ is a **randomization** of $(X_1, \mathcal{A}_1, \mathcal{P}_1)$, if there exists a Markov operator $M : L^\infty(X_2, \mathcal{A}_2, \mu_2) \to L^\infty(X_1, \mathcal{A}_1, \mu_1)$, satisfying

$$\int (Mf) dP_{\theta,1} = \int f dP_{\theta,2} \qquad (\theta \in \Theta, \ f \in L^\infty(X_2, \mathcal{A}_2, \mathcal{P}_2)).$$

If also $(X_1, \mathcal{A}_1, \mathcal{P}_1)$ is a randomization of $(X_2, \mathcal{A}_2, \mathcal{P}_2)$, then $(X_1, \mathcal{A}_1, \mathcal{P}_1)$ and $(X_2, \mathcal{A}_2, \mathcal{P}_2)$ are **stochastically equivalent**.

For example, let $\mathcal{P} \equiv P_0$ and let $\mathcal{A}_0 \subseteq \mathcal{A}$ be a subalgebra. Then $(X, \mathcal{A}_0, \mathcal{P}|\mathcal{A}_0)$ is obviously a randomization of $(X, \mathcal{A}, \mathcal{P})$, where the Markov operator is the inclusion

$L^\infty(X, \mathcal{A}_0, P_0|_{\mathcal{A}_0}) \to L^\infty(X, \mathcal{A}, P_0)$. On the other hand, if $\mathcal{A}_0$ is sufficient, then the map

$$f \mapsto E[f|\mathcal{A}_0], \qquad E[f|\mathcal{A}_0] = E_\theta[f|\mathcal{A}_0], \quad P_\theta - \text{almost everywhere,}$$

is a Markov operator $L^\infty(X, \mathcal{A}, P_0) \to L^\infty(X, \mathcal{A}_0, P_0|_{\mathcal{A}_0})$ and

$$\int E[f|\mathcal{A}_0] dP_\theta|_{\mathcal{A}_0} = \int f dP_\theta \qquad (f \in L^\infty(X, \mathcal{A}, P_0), \ \theta \in \Theta).$$

We have the following characterizations of sufficient subalgebras.

Proposition 1 *Let $\mathcal{P}$ be a dominated family and let $\mathcal{A}_0 \subseteq \mathcal{A}$ be a sub-σ-algebra. The following are equivalent.*

(i) *$\mathcal{A}_0$ is sufficient for $\mathcal{P}$*

(ii) *There exists a measure P_0 such that $P_0 \equiv \mathcal{P}$ and dP_θ/dP_0 is $\mathcal{A}_0$-measurable for all θ.*

(iii) *$(X, \mathcal{A}, \mathcal{P})$ and $(X, \mathcal{A}_0, \mathcal{P}|_{\mathcal{A}_0})$ are stochastically equivalent*

It follows that if $\mathcal{P} \equiv \mu$, then the sub-σ-algebra generated by the functions $\{dP_\theta/d\mu : \theta \in \Theta\}$ is sufficient for $\mathcal{P}$, moreover, it is contained in any other sufficient subalgebra in $\mathcal{A}$. Such subalgebras are called **minimal sufficient**.

Next we formulate a non-commutative setting. Let $\mathcal{M}$ be a von Neummann algebra and $\mathcal{M}_0$ be its von Neumann subalgebra. Assume that a family $\mathcal{S} := \{\varphi_\theta : \theta \in \Theta\}$ of normal states are given. $(\mathcal{M}, \mathcal{S})$ is called **statistical experiment**. The subalgebra $\mathcal{M}_0 \subset \mathcal{M}$ is **sufficient** for $(\mathcal{M}, \mathcal{S})$ if for every $a \in \mathcal{M}$, there is $\alpha(a) \in \mathcal{M}_0$ such that

$$\varphi_\theta(a) = \varphi_\theta(\alpha(a)) \qquad (\theta \in \Theta) \tag{3}$$

and the correspondence $a \mapsto \alpha(a)$ is a coarse-graining. (Note that a positive mapping is automatically completely positive if it is defined on a commutative algebra.)

Example 3 Consider a bipartite system $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$ and a family $\{\varphi_\theta : \theta \in \Theta\}$ of states on $\mathcal{H}$. Assume that the expectation value of all observables localized at A is known to us, that is, we know the restriction of φ_θ's to $B(\mathcal{H}_A)$ (or the reduced density matrices). This information is not sufficient in general to decide about θ. We impose the further condition that $\mathcal{H}_A = \mathcal{H}_L \otimes \mathcal{H}_R$ and the factorization

$$\varphi_\theta = \varphi_\theta^0 \otimes \varphi_{RB},$$

where φ_θ^0 is a state on $B(\mathcal{H}_L)$ and the state φ_{RB} of $B(\mathcal{H}_R) \otimes B(\mathcal{H}_B)$ is independent of the parameter θ. In this case the restriction of the unknown state to $B(\mathcal{H}_L)$ determines the true value of the parameter θ and φ_θ is recovered uniquely.

The subalgebra $B(\mathcal{H}_L)$ is sufficient and the example is close to typical. In the general case, however, the relation of the subalgebras $B(\mathcal{H}_L)$ and $B(\mathcal{H}_A)$ is more subtle. $\square$

The following lemma is a quantum version of Lemma 1.

Lemma 2 *Assume that the von Neumann algebra $\mathcal{M}$ admits a faithful normal state ψ. Let $\mathcal{S} = \{\varphi_\theta : \theta \in \Theta\}$ be a family of normal states on $\mathcal{M}$. Then there is a sequence (φ_n) of states in $\mathcal{S}$ and a normal state*

$$\omega = \sum_{n=1}^{\infty} \lambda_n \varphi_n$$

such that $\operatorname{supp} \varphi_\theta \leq \operatorname{supp} \omega$ for all $\theta \in \Theta$.

Proof: Let $\{p_i : i \in I\}$ be a set of pairwise orthogonal projections in $\mathcal{M}$, then $\psi(p_i) > 0$ and $\psi(\sum_i p_i) \leq 1$, therefore any such set must be at most countable.

We set

$$\mathcal{P} = \{p_\theta := \operatorname{supp} \varphi_\theta : \theta \in \Theta\}$$

and show that there is a countable subset $\{p_1, p_2, \ldots\} \subset \mathcal{P}$, such that $\sup_\theta p_\theta = \sup_n p_n$

Let $\mathcal{C}$ be a set of at most countable subsets in $\mathcal{P}$, ordered by inclusion. Consider all chains in $\mathcal{C}$, such that if $C \subset D$ in the chain, then $\sup C \neq \sup D$. It is clear that each such chain has at most countably many elements. Let $\{C_1, C_2, \ldots\}$ be a maximal such chain and let $C = \cup_n C_n = \{p_1, p_2, \ldots\}$. Then $\sup_n p_n = \sup_\theta p_\theta$. Indeed, if $\sup_n p_n \neq \sup_\theta p_\theta$, then there is an element $p \in \mathcal{P}$, such that $\sup C \neq \sup C \cup \{p\}$, which contradicts the maximality of $\{C_1, C_2, \ldots\}$.

Let now $\varphi_1, \varphi_2, \ldots$ be elements is $\mathcal{S}$ such that $\operatorname{supp} \varphi_n = p_n$. Choose a sequence $\lambda_1, \lambda_2, \ldots$ such that $\lambda_n > 0$ for all n and $\sum_n \lambda_n = 1$ and put $\omega = \sum \lambda_n \varphi_n$. Then it is clear that $\operatorname{supp} \omega = \sup_n p_n$ and $\operatorname{supp} \varphi_\theta \leq \operatorname{supp} \omega$ for all θ. $\square$

Throughout the paper, we suppose that the hypothesis of the above lemma is satisfied, that is, the von Neumann algebras considered admit a faithful normal state. The algebra $B(\mathcal{H})$ satisfies this condition if and only if the Hilbert space $\mathcal{H}$ is separable.

When the states φ_n belong to $\mathcal{S}$ and for

$$\omega := \sum_{n=1}^{\infty} \lambda_n \varphi_n$$

the condition $\operatorname{supp} \varphi_\theta \leq \operatorname{supp} \omega$ holds for all $\theta \in \Theta$, we say that $\mathcal{S}$ **is dominated by** ω.

2 Sufficient subalgebras

In the study of sufficient subalgebras monotone quasi-entropy quantities could be useful. The **relative entropy** and the **transition probability** are examples of those [16, 13].

Let φ and ω be normal states of a von Neumann algebra and let ξ_φ and ξ_ω be the representing vectors of these states from the natural positive cone. Then the **transition probability** is defined as

$$P_A(\varphi, \omega) = \langle \xi_\varphi, \xi_\omega \rangle.$$

In case of density matrices this reduces to $P_A(D_1, D_2) = \mathrm{Tr}(D_1^{1/2} D_2^{1/2})$.

Theorem 1 *Let $(\mathcal{M}, \{\varphi_\theta : \theta \in \Theta\})$ be a statistical experiment and let $\mathcal{M}_0 \subset \mathcal{M}$ be von Neumann algebras. Assume that $\{\varphi_\theta : \theta \in \Theta\}$ is dominated by a faithful normal state ω. Then the following conditions are equivalent.*

(i) *$\mathcal{M}_0$ is sufficient for (φ_θ).*

(ii) *$P_A(\varphi_\theta, \omega) = P_A(\varphi_\theta | \mathcal{M}_0, \omega | \mathcal{M}_0)$ for all θ.*

(iii) *$[D\varphi_\theta, D\omega]_t = [D(\varphi_\theta | \mathcal{M}_0), D(\omega | \mathcal{M}_0)]_t$ for every real t and for every θ.*

(iv) *$[D\varphi_\theta, D\omega]_t \in \mathcal{M}_0$ for all real t and every θ.*

(v) *The generalized conditional expectation $E_\omega : \mathcal{M} \to \mathcal{M}_0$ leaves all the states φ_θ invariant.*

Note that condition (iii) is formulated in terms of Connes' Radon-Nikodym cocycle and the generalized conditional expectation appearing in (iv) is discussed in the appendix.

The theorem is essentially Thm 9.5 from [13] and we give the detailed proof in the finite dimensional situation. The following two lemmas will be used.

Lemma 3 *Let $T : B(\mathcal{H}) \to B(\mathcal{K})$ be a coarse-graining sending density matrices to densities. Let D_1 and D_2 be density matrices acting on the Hilbert space $\mathcal{H}$. Then*

$$P_A(D_1, D_2) \leq P_A(T(D_1), T(D_2))$$

Proof: On the Hilbert space $B(\mathcal{H})$ one can define an operator Δ as

$$\Delta a = D_2 a D_1^{-1} \qquad (a \in B(\mathcal{H})),$$

where the generalized inverse D_1^{-1} is determined by the relation $D_1 D_1^{-1} = D_1^{-1} D_1 = \mathrm{supp}\, D_1$. This is the so-called **relative modular operator** and it is the product of two commuting positive operators: $\Delta = LR$, where

$$La = D_2 a \quad \text{and} \quad Ra = a D_1^{-1} \qquad (a \in B(\mathcal{H})).$$

We have

$$P_A(D_1, D_2) = \langle D_1^{1/2}, \Delta^{1/2} D_1^{1/2} \rangle.$$

Set

$$\Delta a = D_2 a D_1^{-1} \quad (a \in B(\mathcal{H})) \quad \text{and} \quad \Delta_0 x = T(D_2) x T(D_1)^{-1} \quad (x \in B(\mathcal{K})).$$

Δ and Δ_0 are operators on the spaces $B(\mathcal{H})$ and $B(\mathcal{K})$. (They become Hilbert space with the Hilbert-Schmidt inner product.) The transition probabilities are expressed by the resolvent of relative modular operators:

$$\begin{aligned}
P_A(D_1, D_2) &= \langle D_1^{1/2}, \Delta^{1/2} D_1^{1/2} \rangle \\
&= \frac{1}{\pi} \int_0^\infty t^{-1/2} - t^{1/2} \langle D_1^{1/2}, (\Delta + t)^{-1} D_1^{1/2} \rangle \, dt \\
P_A(T(D_1), T(D_2)) &= \langle T(D_1)^{1/2}, \Delta_0^{1/2} T(D_1)^{1/2} \rangle \\
&= \frac{1}{\pi} \int_0^\infty t^{-1/2} - t^{1/2} \langle T(D_1)^{1/2}, (\Delta_0 + t)^{-1} T(D_1)^{1/2} \rangle \, dt
\end{aligned}$$

where the identity

$$x^{1/2} = \frac{1}{\pi} \int_0^\infty t^{-1/2} - t^{1/2} (x + t)^{-1} \, dt$$

is used. Let us define the operator

$$V(x T(D_1)^{1/2} + \xi) = T^*(x) D_1^{1/2} \tag{4}$$

where $\xi \in [\mathcal{B}(\mathcal{K}) T(D_1)^{1/2}]^\perp$. Then V is a contraction:

$$\begin{aligned}
\|T^*(x) D_1^{1/2}\|^2 &= \mathrm{Tr} D_1 T^*(x^*) T^*(x) \leq \mathrm{Tr} D_1 T^*(x^* x) = \mathrm{Tr} T(D_1) x^* x = \\
&= \|x T(D_1)^{1/2}\|^2 \leq \|x T(D_1)^{1/2} + \xi\|^2
\end{aligned}$$

since the Schwarz inequality is applicable to T^*. Let now $p_1 = \mathrm{supp}\, D_1$ and $q_1 = \mathrm{supp}\, T(D_1)$. Since T^* is unital, $0 = \mathrm{Tr} T(D_1)(1 - q_1) = \mathrm{Tr} D_1(1 - T^*(q_1))$ and therefore $p_1 \leq T^*(q_1)$. The Schwarz inequality (1) now implies

$$\begin{aligned}
\langle V(x T(D_1)^{1/2} + \xi), \Delta V(x T(D_1)^{1/2} + \xi) \rangle &= \mathrm{Tr} D_2 T^*(x) p_1 T^*(x^*) \leq \mathrm{Tr} D_2 T^*(x q_1 x^*) \\
= \langle x T(D_1)^{1/2}, \Delta_0 x T(D_1)^{1/2} \rangle &\leq \langle x T(D_1)^{1/2} + \xi, \Delta_0 (x T(D_1)^{1/2} + \xi) \rangle
\end{aligned}$$

where the last inequality follows from

$$\langle \Delta_0 x T(D1)^{1/2}, \xi \rangle = \langle T(D_2) x T(D_1)^{-1} T(D_1)^{1/2}, \xi \rangle = 0$$

It follows that

$$V^* \Delta V \leq \Delta_0. \tag{5}$$

The function $y \mapsto (y + t)^{-1}$ is operator monotone (decreasing) and operator convex, hence

$$(\Delta_0 + t)^{-1} \leq (V^* \Delta V + t)^{-1} \leq V^* (\Delta + t)^{-1} V \tag{6}$$

(see [6]). Since $V T(D_1)^{1/2} = D_1^{1/2}$, this implies

$$\langle D_1^{1/2}, (\Delta + t)^{-1} D_1^{1/2} \rangle \geq \langle T(D_1)^{1/2}, (\Delta_0 + t)^{-1} T(D_1)^{1/2} \rangle. \tag{7}$$

8

By integrating this inequality we have the monotonicity theorem from the above integral formulas. □

Condition (i) implies that $P_A(\varphi_\theta, \omega) = P_A(\varphi_\theta|\mathcal{M}_0, \omega|\mathcal{M}_0)$ due to the monotonicity of the transition probability under completely positive mappings. Indeed, if α leaves all φ_θ invariant, then $\omega \circ \alpha = \omega$.

Now we are in the position to analyze the case of equality.

Lemma 4 *If*
$$P_A(D_1, D_2) = P_A(T(D_1), T(D_2)),$$
then
$$T^*(T(D_2)^{it}T(D_1)^{-it})p_1 = D_2^{it}D_1^{-it}p_1,$$
where $p_1 = supp\, D_1$.

Proof: From the integral formula for the transition probability we have

$$\langle T(D_1)^{1/2}, V^*(\Delta + t)^{-1}VT(D_1)^{1/2}\rangle = \langle T(D_1)^{1/2}, (\Delta_0 + t)^{-1}T(D_1)^{1/2}\rangle. \qquad (8)$$

for all $t > 0$. This equality together with the operator inequality (6) gives

$$V^*(\Delta + t)^{-1}D_1^{1/2} = (\Delta_0 + t)^{-1}T(D_1)^{1/2} \qquad (9)$$

for all $t > 0$. Differentiating by t we have

$$V^*(\Delta + t)^{-2}D_1^{1/2} = (\Delta_0 + t)^{-2}T(D_1)^{1/2} \qquad (10)$$

and we infer

$$
\begin{aligned}
\|V^*(\Delta + t)^{-1}D_1^{1/2}\|^2 &= \langle(\Delta_0 + t)^{-2}T(D_1)^{1/2}, T(D_1)^{1/2}\rangle \\
&= \langle V^*(\Delta + t)^{-2}D_1^{1/2}, T(D_1)^{1/2}\rangle \\
&= \|(\Delta + t)^{-1}D_1^{1/2}\|^2
\end{aligned}
$$

When $\|V^*\xi\| = \|\xi\|$ holds for a contraction V, it follows that $VV^*\xi = \xi$. In the light of this remark we arrive at the condition

$$VV^*(\Delta + t)^{-1}D_1^{1/2} = (\Delta + t)^{-1}D_1^{1/2}$$

and

$$
\begin{aligned}
V(\Delta_0 + t)^{-1}T(D_1)^{1/2} &= VV^*(\Delta + t)^{-1}D_1^{1/2} \\
&= (\Delta + t)^{-1}D_1^{1/2}
\end{aligned}
$$

By Stone-Weierstrass approximation we have

$$Vf(\Delta_0)T(D_1)^{1/2} = f(\Delta)D_1^{1/2} \qquad (11)$$

9

for continuous functions. In particular for $f(x) = x^{it}$ we have

$$T^*(T(D_2)^{it}T(D_1)^{-it})p_1 = D_2^{it}D_1^{-it}p_1 \,. \tag{12}$$

This condition is necessary and sufficient for the equality. □

The previous lemma shows that condition (ii) implies (iii) and it is clear that (iii) implies (iv). We prove that (iv) implies (i).

Let $\mathcal{M}_1$ be the subalgebra generated by $\{[D\varphi_\theta, D\omega]_t, t \in \mathbb{R}\}$ and let ω_1, φ_1 be the restrictions of ω, φ_θ to $\mathcal{M}_1$. Then $[D\varphi_\theta, D\omega]_t$ satisfies the cocycle condition for $\sigma_t^{\omega_1}$ and therefore there is a weight ψ on $\mathcal{M}_1$, such that $[D\psi, D\omega_1]_t = [D\varphi_\theta, D\omega]_t$.

On the other hand, $\mathcal{M}_1$ is invariant under the modular group σ_ω^t, hence there exists a conditional expectation $F : \mathcal{M} \to \mathcal{M}_1$ preserving ω and

$$[D\psi \circ F, D\omega]_t = [D\psi, D\omega_1]_t = [D\varphi_\theta, D\omega]_t, \quad \forall t$$

It follows that $\psi \circ F = \varphi_\theta$, therefore $\psi = \varphi_1$ and F preserves also φ_θ.

Let now (iv) be satisfied, then $\mathcal{M}_1 \subseteq \mathcal{M}_0$. The conditional expectation F is a coarse-graining $\mathcal{M} \to \mathcal{M}_0$ preserving all φ_θ and (i) follows.

Next, we want to show that (iii) implies (v). Let $E : \mathcal{M} \to \mathcal{M}_0$ be the trace preserving conditional expectation. Then the generalized conditional expectation $E_\omega : \mathcal{M} \to \mathcal{M}_0$ acts as

$$E_\omega(a) = E(D)^{-1/2}E(D^{1/2}aD^{1/2})E(D)^{-1/2}$$

We have to show that

$$\mathrm{Tr}E(D_\theta)E_\omega(a) = \mathrm{Tr}D_\theta a$$

which is equivalently written as

$$\mathrm{Tr}E(D_\theta)^{1/2}E(D)^{-1/2}E(D^{1/2}aD^{1/2})E(D)^{-1/2}E(D_\theta)^{1/2} = \mathrm{Tr}D_\theta a$$

By analytic continuation from condition (iii), we have

$$E(D_\theta)^{1/2}E(D)^{-1/2} = D_\theta^{1/2}D^{-1/2}$$

It follows that

$$\begin{aligned}
\mathrm{Tr}E(D_\theta)E_\omega(a) &= \mathrm{Tr}E(D_\theta)^{1/2}E(D)^{-1/2}D^{1/2}aD^{1/2}E(D)^{-1/2}E(D_\theta)^{1/2} = \\
&= \mathrm{Tr}D_\theta D^{-1/2}D^{1/2}aD^{1/2}D^{-1/2}D_\theta^{1/2} = \mathrm{Tr}D_\theta a \,.
\end{aligned}$$

The implication (v) → (i) is trivial. □

3 Sufficient statistic and coarse-graining

A classical **sufficient statistic** for the family $\mathcal{P}$ is a measurable mapping $T : (X, \mathcal{A}) \to (X_1, \mathcal{A}_1)$ such that the generated sub-σ-algebra $T^{-1}(\mathcal{A}_1) \subset \mathcal{A}$ is sufficient for $\mathcal{P}$. To any statistic T, we associate a Markov operator

$$\tilde{T} : \ L^\infty(X_1, \mathcal{A}_1, P_0^T) \to L^\infty(X, \mathcal{A}, P_0), \quad (\tilde{T}g)(x) = g(T(x))$$

Obviously, $(X_1, \mathcal{A}_1, \mathcal{P}^T)$ is a randomization of $(X, \mathcal{A}, \mathcal{P})$. As in the case of subalgebras, we have

Proposition 2 *The statistic $T : (X, \mathcal{A}) \to (X_1, \mathcal{A}_1)$ is sufficient for $\mathcal{P}$ if and only if $(X, \mathcal{A}, \mathcal{P})$ and $(X_1, \mathcal{A}_1, \mathcal{P}^T)$ are stochastically equivalent.*

Proposition 3 *(Factorization criterion) Let $\mathcal{P} \ll \mu$. The statistic $T : (X, \mathcal{A}) \to (X_1, \mathcal{A}_1)$ is sufficient for $\mathcal{P}$ if and only if there is an $\mathcal{A}_1$-measurable function g_θ for all θ and an $\mathcal{A}$-measurable function h such that*

$$\frac{dP_\theta}{d\mu}(x) = g_\theta(T(x))h(x) \quad P_\theta - almost\ everywhere$$

Let $\mathcal{N}$, $\mathcal{M}$ be C*-algebras and let $\sigma : \mathcal{N} \to \mathcal{M}$ be a coarse-graining. We say that σ is sufficient for the statistical experiment $(\mathcal{M}, \varphi_\theta)$ if there exists a coarse-graining $\beta : \mathcal{M} \to \mathcal{N}$ such that $\varphi_\theta \circ \sigma \circ \beta = \varphi_\theta$ for every θ.

Let $\omega = \sum_n \lambda_n \varphi_n$ be the normal state obtained in Lemma 2 and let $p = \mathrm{supp}\,\omega$, $q = \mathrm{supp}\,\omega \circ \sigma$. Let us define the map $\alpha : q\mathcal{N}q \to p\mathcal{M}p$ by $\alpha(a) = p\sigma(a)p$, then α is a coarse-graining such that $\varphi_\theta \circ \sigma(a) = \varphi_\theta \circ \alpha(qaq)$ for all θ and $\alpha_\omega^* = \sigma_\omega^*$, where the dual σ_ω^* is defined in the Appendix. We check that α is sufficient for $(p\mathcal{M}p, \varphi_\theta|_{p\mathcal{M}p})$ if and only if σ is sufficient for $(\mathcal{M}, \varphi_\theta)$. Indeed, let $\tilde{\beta} : p\mathcal{M}p \to q\mathcal{N}q$ be a coarse-graining such that $\varphi_\theta|_{p\mathcal{M}p} \circ \alpha \circ \tilde{\beta} = \varphi_\theta|_{p\mathcal{M}p}$ and let $\beta : \mathcal{M} \to \mathcal{N}$ be defined by

$$\beta(a) = \tilde{\beta}(pap) + \omega(a)(1 - q)$$

Then β is a coarse-graining and

$$\varphi_\theta \circ \sigma \circ \beta(a) = \varphi_\theta \circ \sigma(q\beta(a)q)) = \varphi_\theta \circ \alpha \circ \tilde{\beta}(pap) = \varphi_\theta(pap) = \varphi_\theta(a)$$

The converse is proved similarly, taking $\tilde{\beta}(a) = q\beta(a)q$ for $a \in p\mathcal{M}p$. Therefore we may, and will, suppose that both ω and $\omega \circ \sigma$ are faithful.

Let us recall the following property of coarse-grainings.

Lemma 5 *Let $\mathcal{M}$ and $\mathcal{N}$ be C*-algebras and let $\sigma : \mathcal{N} \to \mathcal{M}$ be a coarse-graining. Then*

$$\mathcal{N}_\sigma := \{a \in \mathcal{N} : \sigma(a^*a) = \sigma(a)\sigma(a)^* \ and \ \sigma(aa^*) = \sigma(a)^*\sigma(a)\} \tag{13}$$

is a subalgebra of $\mathcal{N}$ and

$$\sigma(ab) = \sigma(a)\sigma(b) \quad and \quad \sigma(ba) = \sigma(b)\sigma(a) \tag{14}$$

holds for all $a \in \mathcal{N}_\sigma$ and $b \in \mathcal{N}$.

Proof: The proof is based only on the Schwarz inequality

$$\sigma(x^*x) \geq \sigma(x)^*\sigma(x).$$

From this we have

$$
\begin{aligned}
t(\sigma(a)\sigma(b) + \sigma(b)^*\sigma(a)^*) &= \sigma(ta^* + b)^*\sigma(ta^* + b) - t^2\sigma(a)\sigma(a)^* - \sigma(b)^*\sigma(b) \\
&\leq \sigma(ta^* + b)^*\sigma(ta^* + b) - t^2\sigma(aa^*) - \sigma(b)^*\sigma(b) \\
&= t\sigma(ab + b^*a^*) + \sigma(b^*b) - \sigma(b)^*\sigma(b)
\end{aligned}
$$

for a real t and $a \in \mathcal{N}_\sigma$. Divide the inequality by t and let $t \to \pm\infty$. Then

$$\sigma(a)\sigma(b) + \sigma(b)^*\sigma(a^*) = \sigma(ab + b^*a^*)$$

and similarly

$$\sigma(a)\sigma(b) - \sigma(b)^*\sigma(a)^* = \sigma(ab - b^*a^*).$$

Adding these two inequalities we have

$$\sigma(ab) = \sigma(a)\sigma(b).$$

$\square$

We call the subalgebra $\mathcal{N}_\sigma$ the **multiplicative domain** of σ.

Let now $\mathcal{N}$ and $\mathcal{M}$ be von Neumann algebras and let ω be a faithful normal state on $\mathcal{M}$ such that $\omega \circ \sigma$ is also faithful. Let

$$\mathcal{N}_1 = \{a \in \mathcal{N}, \ \sigma_\omega^* \circ \sigma(a) = a\}$$

It was proved in [17] that $\mathcal{N}_1$ is a subalgebra of $\mathcal{N}_\sigma$, moreover, $a \in \mathcal{N}_1$ if and only if $\sigma(a^*a) = \sigma(a)^*\sigma(a)$ and $\sigma(\sigma_t^{\omega\circ\sigma}(a)) = \sigma_t^\omega(\sigma(a))$. The restriction of σ to $\mathcal{N}_1$ is an isomorphism onto

$$\mathcal{M}_1 = \{b \in \mathcal{M}, \ \sigma \circ \sigma_\omega^*(b) = b\}$$

The following Theorem was proved in [17] in the case when φ_θ are faithful states.

Theorem 2 *Let $\mathcal{M}$ and $\mathcal{N}$ be von Neumann algebras and let $\sigma : \mathcal{N} \to \mathcal{M}$ be a coarse-graining. Suppose that $(\mathcal{M}, \varphi_\theta)$ is a statistical experiment dominated by a state ω such that both ω and $\omega \circ \sigma$ are faithful and normal.*

Then following properties are equivalent:

(i) $\sigma(\mathcal{N}_\sigma)$ is a sufficient subalgebra for $(\mathcal{M}, \varphi_\theta)$.

(ii) σ is a sufficient coarse-graining for $(\mathcal{M}, \varphi_\theta)$.

(iii) $P_A(\varphi_\theta, \omega) = P_A(\varphi_\theta \circ \sigma, \omega \circ \sigma)$

(iv) $\sigma([D\varphi_\theta \circ \sigma, D\omega \circ \sigma]_t) = [D\varphi_\theta, D\omega]_t$

(v) $\mathcal{M}_1$ *is a sufficient subalgebra for* $(\mathcal{M}, \varphi_\theta)$.

(vi) $\varphi_\theta \circ \sigma \circ \sigma_\omega^* = \varphi_\theta$.

Proof. Suppose (i), then there is a coarse-graining $\gamma : \mathcal{M} \to \sigma(\mathcal{N}_\sigma)$, preserving φ_θ. It is easy to see that the restriction of σ to $\mathcal{N}_\sigma$ is invertible. Let α be the inverse of this restriction and put

$$\beta = \alpha \circ \gamma$$

Then $\beta : \mathcal{M} \to \mathcal{N}$ is a coarse-graining such that $\varphi_\theta \circ \sigma \circ \beta = \varphi_\theta$ and (ii) is proved.

The implications (ii) $\to$ (iii) and (iii) $\to$ (iv) follow from Lemmas 3 and 4.

Suppose (iv) and denote $u_t = [D\varphi_\theta \circ \sigma, D\omega \circ \sigma]_t$, $v_t = [D\varphi_\theta, D\omega]_t$. Then we have $\sigma(u_t) = v_t$ for all t. Let $p_\theta = \operatorname{supp} \varphi_\theta$, $q_\theta = \operatorname{supp} \varphi_\theta \circ \sigma$. Putting $t = 0$ in the condition (iv), we get $\sigma(q_\theta) = p_\theta$ and

$$\sigma(u_t u_t^*) = \sigma(q_\theta) = p_\theta = v_t v_t^* = \sigma(u_t)\sigma(u_t)^*$$

On the other hand, $\sigma(u_t)^*\sigma(u_t) \leq \sigma(u_t^* u_t)$ by Schwartz inequality and from

$$
\begin{aligned}
\omega(\sigma(u_t)^*\sigma(u_t)) &= \omega(v_t^* v_t) = \omega(\sigma_t^\omega(p_\theta)) = \omega(p_\theta) \\
\omega(\sigma(u_t^* u_t)) &= \omega \circ \sigma(u_t^* u_t) = \omega \circ \sigma(\sigma_t^{\omega\circ\sigma}(q_\theta)) = \omega(p_\theta)
\end{aligned}
$$

we get $\sigma(u_t^* u_t) = \sigma(u_t)^*\sigma(u_t)$. Hence $u_t \in \sigma(\mathcal{N}_\sigma)$ for all t. Further, by the cocycle condition and Lemma 5,

$$\sigma(\sigma_t^{\omega\circ\sigma}(u_t)) = \sigma(u_s^* u_{s+t}) = v_s^* v_{t+s} = \sigma_t^\omega(\sigma(u_t))$$

therefore $v_t \in \mathcal{M}_1$ and by Theorem 1, $\mathcal{M}_1$ is sufficient and (v) is proved. As $\mathcal{M}_1$ is a subalgebra in $\sigma(\mathcal{N}_\sigma)$, this implies (i).

Finally, we prove that (ii) is equivalent to (vi). First, note that a coarse-graining is sufficient for $(\mathcal{M}, \varphi_\theta)$ if and only if it is sufficient for $(\mathcal{M}, \psi_\theta)$, where

$$\psi_\theta = \varepsilon\varphi_\theta + (1 - \varepsilon)\omega$$

for some $0 < \varepsilon < 1$.

As the states ψ_θ are faithful and $\omega = \sum_n \lambda_n \psi_n$, it follows from the results in [17] that σ is sufficient if and only if $\psi_\theta \circ \sigma \circ \sigma_\omega^* = \psi_\theta$ for all θ. Since, by definition, $\omega \circ \sigma \circ \sigma_\omega^* = \omega$, this is equivalent to (vi). $\qquad\square$

Let $\mathcal{M}_0 \subset \mathcal{M}$ be a subalgebra. From the above theorem, together with the remarks preceding Lemma 5, we have a generalization of Theorem 1 to the case that $\operatorname{supp} \omega = p$ and $\operatorname{supp} \omega|_{\mathcal{M}_0} = q$. Namely, $\mathcal{M}_0$ is sufficient for $(\mathcal{M}, \varphi_\theta)$ if and only if the coarse-graining $\alpha : sq\mathcal{M}_0 q \to p\mathcal{M}p$, $\alpha(qaq) = pap$ is sufficient for the restricted experiment.

Remark. Let $S(\varphi, \omega)$ be the relative entropy and suppose that $S(\varphi_\theta, \omega)$ is finite for all θ. Then the condition (iii) can be replaced by

$$S(\varphi_\theta, \omega) = S(\varphi_\theta \circ \sigma, \omega \circ \sigma)$$

This can be proved similarly as for the transition probability, using the formula

$$\log x = \int_0^\infty (1+t)^{-1} - (x+t)^{-1} dt. \qquad (15)$$

The equality in inequalities for entropy quantities was studied also in [20]. $\qquad \square$

The previous theorem applies to a measurement which is essentially a positive mapping $\mathcal{N} \to \mathcal{M}$ from a commutative algebra. The concept of sufficient measurement appeared also in [3]. For a non-commuting family of states, there is no sufficient measurement.

4   Factorization

Let $\mathcal{M}$ be a von Neumann algebra and let ω be a faithful state on $\mathcal{M}$. Let $\mathcal{M}_0 \subset \mathcal{M}$ be a subalgebra and assume that it is invariant under the modular group σ_t^ω of ω. Let $\mathcal{M}_1 = \mathcal{M}_0' \cap \mathcal{M}$ be the relative commutant. We show that $\mathcal{M}_1$ is invariant under σ_t^ω as well. If $a \in \mathcal{M}_0$ and $b \in \mathcal{M}_1$, then for $t \in \mathbb{R}$, we have

$$a\sigma_t^\omega(b) = \sigma_t^\omega(\sigma_{-t}^\omega(a)b) = \sigma_t^\omega(b\sigma_{-t}^\omega(a)) = \sigma_t^\omega(b)a$$

Hence $\mathcal{M}_1$ is invariant under σ_t^ω. Let ω_0, ω_1 be the restrictions of ω to $\mathcal{M}_0$ and $\mathcal{M}_1$. Then $\sigma_t^\omega|_{\mathcal{M}_0} = \sigma_t^{\omega_0}$ and $\sigma_t^\omega|_{\mathcal{M}_1} = \sigma_t^{\omega_1}$ are known facts in modular theory.

Recall that the **entropy** of a state φ of a C*-algebra is defined as

$$S(\varphi) := \sup \left\{ \sum_i \lambda_i S(\varphi_i \| \varphi) : \sum_i \varphi_i = \varphi \right\},$$

see (6.9) in [13]. For the sake of simplicity, we will suppose in the rest of this section that the state ω has finite von Neumann entropy $S(\omega)$. Then $\mathcal{M}$ must be a countable direct sum of type I factors, see Theorem 6.10. in [13]. Let τ be the canonical normal semifinite trace on $\mathcal{M}$ and let D_ω be the density of ω with respect to τ, then

$$\sigma_t^\omega(a) = D_\omega^{it} a D_\omega^{-it}, a \in \mathcal{M}.$$

As the subalgebras $\mathcal{M}_0$ and $\mathcal{M}_1$ are invariant under σ_t^ω, we have by Proposition 6.7. in [13] that $S(\omega_0), S(\omega_1) \leq S(\omega) < \infty$. It follows that both $\mathcal{M}_0$ and $\mathcal{M}_1$ must be countable direct sums of type I factors as well.

Let $D_{\omega_0} \in \mathcal{M}_0$ and $D_{\omega_1} \in \mathcal{M}_1$ be the densities of ω_0 and ω_1 with respect to the canonical traces $\tau_0 := \tau|\mathcal{M}_0$ and $\tau_1 := \tau|\mathcal{M}_1$. Then for $a \in \mathcal{M}_0$,

$$D_\omega^{it} a D_\omega^{-it} = \sigma_t^\omega(a) = \sigma_t^{\omega_0}(a) = D_{\omega_0}^{it} a D_{\omega_0}^{-it}.$$

It follows that $w_t := D_{\omega_0}^{-it} D_\omega^{it}$ is a unitary operator in $\mathcal{M}_1$ and the operators $D_{\omega_0}^{it}$ and D_ω^{is} commute for all $t, s \in \mathbb{R}$. It is easy to see that w_t is a strongly continuous one-parameter group. Moreover, we have for $a \in \mathcal{M}_1$,

$$w_t a w_t^* = D_\omega^{it} a D_\omega^{-it} = \sigma_t^{\omega_1}(a) = D_{\omega_1}^{it} a D_{\omega_1}^{-it}$$

Therefore, the unitary $z_t = D_{\omega_1}^{-it} w_t$ is in the center of $\mathcal{M}_1$. Again, w_t and $D_{\omega_1}^{is}$ commute for all t, s and it is easy to see that $z_t = z^{it}$ for some positive element z in the center of $\mathcal{M}_1$. Putting all together, we get

$$D_\omega = D_{\omega_0} D_{\omega_1} z \tag{16}$$

The following theorem is a generalization of the classical factorization theorem.

Theorem 3 *Let $(\mathcal{M}, \mathcal{S})$ be a statistical experiment dominated by a faithful normal state ω such that $S(\omega) < \infty$. Let $\mathcal{M}_0 \subset \mathcal{M}$ be a von Neumann subalgebra invariant with respect to the modular group σ_t^ω. Then $\mathcal{M}_0$ is sufficient for $\mathcal{S}$ if and only if*

$$D_\theta = D_{\theta,0} D_{\omega_1} z, \tag{17}$$

where D_θ, $D_{\theta,0}$ and D_{ω_1} are the densities of φ_θ, $\varphi_\theta|\mathcal{M}_0$ and $\omega|\mathcal{M}_0' \cap \mathcal{M}$, respectively and z is a positive operator from the centre of $\mathcal{M}_0' \cap \mathcal{M}$.

Proof. By the assumptions and (16), we have $D_\omega^{it} = D_{\omega_0}^{it} D_{\omega_1}^{it} z^{it}$. If $\mathcal{M}_0$ is sufficient, then

$$u_t := D_\theta^{it} D_\omega^{-it} = [D\varphi_\theta, D\omega]_t = [D\varphi_\theta|\mathcal{M}_0, D\omega_0]_t = D_{\theta,0}^{it} D_{\omega_0}^{-it},$$

hence $D_\theta^{it} = u_t D_\omega^{it} = D_{\theta,0}^{it} D_{\omega_1}^{it} z^{it}$ and (17) follows.

Conversely, let (17) be true, then $u_t = D_{\theta,0}^{it} D_{\omega_0}^{-it}$ and $\mathcal{M}_0$ is sufficient. $\square$

The essence of the factorization (16) is that the first factor depends on θ while the others do not.

From Theorem 1 (iv), it follows that the subalgebra generated by the partial isometries $\{[D\varphi_\theta, D\omega]_t : t \in \mathbb{R}\}$ is **minimal sufficient**, that is, it is sufficient and contained in any sufficient subalgebra. Moreover, it is invariant under σ_t^ω. We will denote this subalgebra by $\mathcal{M}_\mathcal{S}$. By Theorem 3, we have the decompositions:

$$D_\theta = D_{\mathcal{S},\theta} D_R z_\mathcal{S}, \qquad D_\omega = D_{\mathcal{S},\omega} D_R z_\mathcal{S} \tag{18}$$

where $D_{\mathcal{S},\theta}$, $D_{\mathcal{S},\omega}$ are the densities of the restrictions $\varphi_\theta|\mathcal{M}_\mathcal{S}$ and $\omega|\mathcal{M}_\mathcal{S}$ with respect to the canonical trace $\tau_\mathcal{S}$, it will be called the **$\mathcal{S}$-decomposition**. The next Theorem shows that each decomposition of the form (17) is given by an invariant sufficient subalgebra and (18) is the maximal one.

15

Theorem 4 *Let us suppose that there is a decomposition $D_\theta = L_\theta R$, with some positive operators L_θ, R in $\mathcal{M}$, such that $\operatorname{supp} R = I$ and R commutes with all L_θ. Let $\mathcal{M}_L$ be the von Neumann algebra generated by $\{L_\theta : \theta \in \Theta\}$. Then $\mathcal{M}_L$ is sufficient and invariant under σ_t^ω. Moreover,*

$$L_\theta = D_{\mathcal{S},\theta} R_0\,,$$

where $D_{\mathcal{S},\theta}$ is given by (18) and $R_0 \in \mathcal{M}_L$ is a positive element commuting with all $D_{\mathcal{S},\theta}$.

Proof. We have $D_\omega = \sum_n \lambda_n D_{\theta_n} = \sum_n \lambda_n L_{\theta_n} R$, hence $\sum_n \lambda_n L_{\theta_n}$ converges strongly to some positive operator $L_\omega \in \mathcal{M}_L$, such that $D_\omega = L_\omega R$. For $a \in \mathcal{M}_L$, we get

$$D_\omega^{it} a D_\omega^{-it} = L_\omega^{it} a L_\omega^{-it} \in \mathcal{M}_L$$

and $\mathcal{M}_L$ is invariant under σ_t^ω. It follows also that there is a density operator $D_{\omega_L} \in \mathcal{M}_L$ of the restriction $\omega_L := \omega|_{\mathcal{M}_L}$, such that $D_{\omega_L} c = L_\omega$ for some $c \in \mathcal{M}_L' \cap \mathcal{M}_L$. Moreover, it is easy to see that $\mathcal{M}_{\mathcal{S}} \subset \mathcal{M}_L$, so that $\mathcal{M}_L$ is sufficient and the densities of $\varphi_\theta|_{\mathcal{M}_L}$ satisfy

$$D_{\theta,L}^{it} c^{it} = [D\varphi_\theta|_{\mathcal{M}_L}, D\omega_L]_t D_{\omega_L}^{it} c^{it} = [D\varphi_\theta, D\omega]_t L_\omega^{it} = L_\theta^{it}$$

By Theorem 3, there is a decomposition $D_{\theta,L} = D_{\mathcal{S},\theta} D_{R,L} z_L$, such that $D_{R,L} z_L \in \mathcal{M}_{\mathcal{S}}' \cap \mathcal{M}_L$. Putting all together, we get

$$L_\theta = D_{\theta,L} c = D_{\mathcal{S},\theta} R_0$$

where $R_0 = D_{R,L} z_L c \in \mathcal{M}_{\mathcal{S}}' \cap \mathcal{M}_L$. $\qquad\qquad\square$

It is easy to see that the $\mathcal{S}$-decomposition is, up to a central element in $\mathcal{M}_{\mathcal{S}}$, the unique decomposition having the property described in the previous theorem.

Keeping the assumptions of Theorem 3, let us suppose that $\mathcal{M}$ acts on some Hilbert space $\mathcal{H}$. The relative commutant $\mathcal{M}_{\mathcal{S}}^c := \mathcal{M}_{\mathcal{S}}' \cap \mathcal{M}$ is a countable direct sum of factors of type I, hence there is an orthogonal family of minimal central projections p_n such that $\sum_n p_n = 1$. Therefore, $z_{\mathcal{S}} = \sum_n z_n p_n$, with some $z_n > 0$. Moreover, there is a decomposition

$$\mathcal{H} = \bigoplus_n \mathcal{H}_n^L \otimes \mathcal{H}_n^R, \qquad p_n : \mathcal{H} \to \mathcal{H}_n^L \otimes \mathcal{H}_n^R \qquad (19)$$

such that, up to isomorphism,

$$\mathcal{M}_{\mathcal{S}}^c = \bigoplus_n \mathbb{C} I_{\mathcal{H}_n^L} \otimes B(\mathcal{H}_n^L)$$

$$(\mathcal{M}_{\mathcal{S}}^c)' = \bigoplus_n B(\mathcal{H}_n^L) \otimes \mathbb{C} I_{\mathcal{H}_n^R}$$

From $D_R \in \mathcal{M}_{\mathcal{S}}^c$ and $D_{\mathcal{S},\theta} \in \mathcal{M}_{\mathcal{S}} \subseteq (\mathcal{M}_{\mathcal{S}}^c)'$, we have

$$p_n D_R = c_n^R (1_{\mathcal{H}_n^L} \otimes D_n^R), \qquad p_n D_{\mathcal{S},\theta} = c_n^L(\theta)(D_n(\theta) \otimes 1_{\mathcal{H}_n^R}),$$

16

where D_n^R is a density operator in $B(\mathcal{H}_n^R)$, $D_n(\theta)$ is a density operator in $B(\mathcal{H}_n^L)$ and $c_n^R, c_n^L(\theta) > 0$. From this and (18), we get the following form of the $\mathcal{S}$-decomposition

$$D_\theta = D_{\mathcal{S},\theta} D_R z_{\mathcal{S}} = \sum_n z_n p_n D_{\mathcal{S},\theta} p_n D_R = \sum_n s_n(\theta) D_n(\theta) \otimes D_n^R, \qquad (20)$$

where $s_n(\theta) \geq 0$ for all θ, n. Clearly, this decomposition is unique, up to isomorphisms. It is also clear that $s_n(\theta) = \tau(D_\theta p_n) = \varphi_\theta(p_n)$.

In particular, each statistical experiment $(B(\mathcal{H}), \mathcal{S})$, dominated by a faithful state with finite entropy, defines a decomposition of the form (19) of the Hilbert space $\mathcal{H}$, which is up to isomorphisms unique. Note also that if the dimension of $\mathcal{H}$ is finite, then it can be shown from Theorem 4 that (20) gives the maximal decomposition, obtained by Koashi and Imoto in [8].

Theorem 5 *Let $\mathcal{K}$ and $\mathcal{H}$ be Hilbert spaces and let $(B(\mathcal{H}), \mathcal{S})$ be a statistical experiment, dominated by a faithful state ω with $S(\omega) < \infty$. Let $\alpha : B(\mathcal{K}) \to B(\mathcal{H})$ be a coarse-graining and let $(B(\mathcal{K}), \mathcal{S}_0)$ be the experiment induced by α. Then the following are equivalent.*

(i) *α is sufficient for $(B(\mathcal{H}), \mathcal{S})$.*

(ii) *Let (19) be the decomposition of $\mathcal{H}$ given by $(B(\mathcal{H}), \mathcal{S})$. There is a decomposition $\mathcal{K} = \bigoplus_n \mathcal{K}_n^L \otimes \mathcal{K}_n^R$ such that if $q_n : \mathcal{K} \to \mathcal{K}_n^L \otimes \mathcal{K}_n^R$ is the orthogonal projection, then $\alpha(q_n) = p_n$. Moreover, there are unitaries $U_n : \mathcal{K}_n^L \to \mathcal{H}_n^L$ and coarse-grainings $\alpha_{n,2} : B(\mathcal{K}_n^R) \to B(\mathcal{H}_n^R)$ such that the restriction $\alpha_n := \alpha|q_n \mathcal{B}(\mathcal{K}) q_n$ has the form*

$$\alpha_n = \alpha_{1,n} \otimes \alpha_{2,n}, \qquad \alpha_{1,n}(a) = U_n a U_n^*, \ a \in B(\mathcal{K}_n^L)$$

(iii) *Let $D_\theta = D_{\mathcal{S},\theta} D_R z_R$ be the $\mathcal{S}$-decomposition. The density $D_{\theta,0}$ of $\varphi_\theta \circ \alpha$ has the form*

$$D_{\theta,0} = L_{\theta,0} \alpha^*(D_R z_R).$$

where $L_{\theta,0} \in B(\mathcal{K})$ is a positive operator satisfying $\alpha(L_{\theta,0}) = D_{\mathcal{S},\theta}$.

If any of the above conditions is satisfied, then the $\mathcal{S}_0$-decomposition of the densities $D_{\theta,0}$ is

$$D_{\theta,0} = \alpha_\omega^*(D_{\mathcal{S},\theta}) \alpha^*(D_R z_R) = \sum_n \varphi_\theta(p_n) U_n^* D_n(\theta) U_n \otimes \alpha_{2,n}^*(D_n^R) \qquad (21)$$

Proof. Let $\mathcal{M}_{\mathcal{S}} \subset B(\mathcal{H})$ be generated by $\{[D\varphi_\theta, D\omega]_t : t \in \mathbb{R}\}$, note that in this case $\mathcal{M}_{\mathcal{S}} = (\mathcal{M}_{\mathcal{S}}^c)'$ and $\mathcal{M}_{\mathcal{S}}^c = \mathcal{M}_{\mathcal{S}}'$.

Let us denote by $\mathcal{N}_{\mathcal{S}_0} \subset B(\mathcal{K})$ the subalgebra generated by $\{[D(\varphi_\theta \circ \alpha), D(\omega \circ \alpha)]_t, \ t \in \mathbb{R}\}$. If α is sufficient, then by Theorem 2, $\mathcal{N}_{\mathcal{S}_0}$ is in the multiplicative domain of α and the restriction $\alpha|\mathcal{N}_{\mathcal{S}}$ is a *-isomorphism $\mathcal{N}_{\mathcal{S}_0}$ onto $\mathcal{M}_{\mathcal{S}}$. Hence, $\mathcal{N}_{\mathcal{S}_0}$ has

the same structure as $\mathcal{M}_\mathcal{S}$. Namely, there is an orthogonal family of minimal central projections $\{q_n\}$ in $\mathcal{N}_{\mathcal{S}_0}$ such that $q_n\mathcal{K} = \mathcal{K}_n^L \otimes \mathcal{K}_n^R$,

$$\mathcal{N}_{\mathcal{S}_0} = \bigoplus_n B(\mathcal{K}_n^L) \otimes \mathbb{C}I_{\mathcal{K}_n^R}, \qquad \mathcal{N}'_{\mathcal{S}_0} = \bigoplus_n \mathbb{C}I_{\mathcal{K}_n^L} \otimes B(\mathcal{K}_n^R)$$

and $\alpha(q_n) = p_n$. Moreover, there are unitaries $U_n : \mathcal{K}_n^L \to \mathcal{H}_n^L$, such that if $a \in \mathcal{N}_{\mathcal{S}_0}$, $a = \sum_n a_n \otimes I_{\mathcal{K}_n^R}$ for some $a_n \in B(\mathcal{K}_n^L)$, then $\alpha(a) = \sum_n U_n a_n U_n^* \otimes I_{\mathcal{H}_n^R}$.

Let $b \in \mathcal{N}'_{\mathcal{S}_0}$, then for $a \in \mathcal{M}_\mathcal{S}$,

$$\alpha(b)a = \alpha(b)\alpha(\alpha^{-1}(a)) = \alpha(b\alpha^{-1}(a)) = \alpha(\alpha^{-1}(a)b) = a\alpha(b)$$

so that $\alpha(b) \in \mathcal{M}'_\mathcal{S}$. Consequently, $\alpha(bq_n) = \alpha(b)p_n \in \mathcal{M}'_\mathcal{S} p_n$ and if $b_n \in B(\mathcal{K}_n^R)$, then $\alpha_n(I_{\mathcal{K}_n^L} \otimes b_n) = I_{\mathcal{H}_n^L} \otimes b'_n$ for some $b'_n \in B(\mathcal{H}_n^R)$. It is clear that the map $\alpha_{2,n} : b_n \mapsto b'_n$ is a coarse-graining $B(\mathcal{K}_n^R) \to B(\mathcal{H}_n^R)$. We also have

$$\alpha_n(a_n \otimes b_n) = \alpha_n((a_n \otimes I_{\mathcal{K}_n^R})(I_{\mathcal{K}_n^L} \otimes b_n)) = \alpha_n(a_n \otimes I_{\mathcal{K}_n^R})\alpha_n(I_{\mathcal{K}_n^L} \otimes b_n) = U_n a_n U_n^* \otimes \alpha_{2,n}(b_n),$$

hence $\alpha_n = \alpha_{1,n} \otimes \alpha_{2,n}$ and (ii) is proved.

Conversely, let (ii) be satisfied and let $a \in \mathcal{M}_\mathcal{S}$. Then $a = \sum_n a_n \otimes 1_{\mathcal{H}_n^R}$ and $a = \alpha(b)$ with $b = \sum_n b_n \otimes I_{\mathcal{K}_n^R}$, $b_n = U_n^* a_n U_n$. Clearly, $\alpha(b^*b) = \alpha(b)^*\alpha(b)$, $\alpha(bb^*) = \alpha(b)\alpha(b)^*$ and therefore b is in the multiplicative domain. By Theorem 2 (i), α is sufficient for $\mathcal{S}$ and (i) is proved.

To prove (i) $\to$ (iii), suppose that α is sufficient, then by the first part of the proof of (ii), $\mathcal{N}_{\mathcal{S}_0}$ is a countable direct sum of type I factors and, moreover, if $\tau_{\mathcal{S}_0}$ is the canonical trace on $\mathcal{N}_{\mathcal{S}_0}$, then $\tau_{\mathcal{S}_0} = \tau_\mathcal{S} \circ \alpha$. We have the $\mathcal{S}_0$-decomposition

$$D_{\theta,0} = D_{\mathcal{S}_0,\theta}D_{R,0}z_{R,0}$$

where $D_{\mathcal{S}_0,\theta}$ is the density of $\varphi_\theta \circ \alpha|\mathcal{N}_{\mathcal{S}_0}$ with respect to $\tau_{\mathcal{S}_0}$. For $a \in \mathcal{N}_{\mathcal{S}_0}$, $\alpha(a) \in \mathcal{M}_\mathcal{S}$ and

$$\tau_{\mathcal{S}_0}(D_{\mathcal{S}_0,\theta}a) = \varphi_\theta(\alpha(a)) = \tau_\mathcal{S}(D_{\mathcal{S},\theta}\alpha(a)) = \tau_\mathcal{S}(\alpha(\alpha_\omega^*(D_{\mathcal{S},\theta}))\alpha(a)) = \tau_{\mathcal{S}_0}(\alpha_\omega^*(D_{\mathcal{S},\theta})a),$$

hence $D_{\mathcal{S}_0,\theta} = \alpha_\omega^*(D_{\mathcal{S},\theta})$ and $\alpha(D_{\mathcal{S}_0,\theta}) = D_{\mathcal{S},\theta}$. Further, let $a \in B(\mathcal{K})$, then

$$\mathrm{Tr}D_{\theta,0}a = \mathrm{Tr}D_\theta\alpha(a) = \mathrm{Tr}\alpha(D_{\mathcal{S}_0,\theta})D_R z_R\alpha(a) = \mathrm{Tr}D_{\mathcal{S}_0,\theta}\alpha^*(D_R z_R)a$$

and (iii) follows, with $L_{\theta,0} = D_{\mathcal{S}_0,\theta}$.

Conversely, suppose (iii) and let $a \in B(\mathcal{K})$, then

$$\mathrm{Tr}D_{\theta,0}a = \mathrm{Tr}L_{\theta,0}\alpha^*(D_R z_R)a = \mathrm{Tr}\alpha(aL_{\theta,0})D_R z_R$$

On the other hand

$$\mathrm{Tr}D_{\theta,0}a = \mathrm{Tr}D_\theta\alpha(a) = \mathrm{Tr}\alpha(a)\alpha(L_{\theta,0})D_R z_R$$

18

In particular, by putting $a = L_{\theta,0}$, we get

$$Tr(\alpha(L_{\theta,0}^2) - \alpha(L_{\theta,0})^2)D_R z_R = 0$$

From this and Schwarz inequality, we get

$$W^{1/2}(\alpha(L_{\theta,0}^2) - \alpha(L_{\theta,0})^2)W^{1/2} = 0$$

where $W = D_R z_R$ is positive, $\operatorname{supp} W = 1$. Consequently, $\alpha(L_{\theta,0}^2) = \alpha(L_{\theta,0})^2$, hence $L_{\theta,0}$ is in the multiplicative domain. Since $D_{\mathcal{S},\theta} = \alpha(L_{\theta,0})$ generates $\mathcal{M}_{\mathcal{S}}$, this implies that α is sufficient.

It remains only to prove the second half of (21), which follows easily from (ii). $\square$

Corollary 1 *Let $\mathcal{H}$ and $\mathcal{K}$ be finite dimensional Hilbert spaces. Let $(B(\mathcal{H}), \mathcal{S})$ be a statistical experiment dominated by a faithful state ω and let (19) be the corresponding decomposition of $\mathcal{H}$. Suppose that $\alpha : B(\mathcal{K}) \to B(\mathcal{H})$ is a completely positive map, with the Kraus representation $\alpha(a) = \sum_i V_i a V_i^*$. Then α is sufficient for $(B(\mathcal{H}), \mathcal{S})$ if and only if there is a decomposition $\mathcal{K} = \bigoplus_n \mathcal{K}_n^L \otimes \mathcal{K}_n^R$ and*

$$V_i = \sum_n U_n \otimes L_{i,n}$$

where $U_n : \mathcal{K}_n^L \to \mathcal{H}_n^L$ are unitary and $L_{i,n} : \mathcal{K}_n^R \to \mathcal{H}_n^R$ are linear maps such that $\sum_i L_{i,n} L_{i,n}^* = 1_{H_n^R}$.

Proof. Note first that $S(\omega) < \infty$, so that the conditions of Theorem 5 are satisfied.

It is clear that if V_i have the above form, then the restrictions

$$\alpha_n = \alpha | B(\mathcal{K}_n^L \otimes \mathcal{K}_n^R) = \alpha_{1,n} \otimes \alpha_{2,n},$$

with $\alpha_{1,n}(a) = U_n a U_n^*$ and $\alpha_{2,n}(a) = \sum_i L_{i,n} a L_{i,n}^*$. By Theorem 5 (ii), α is sufficient.

Conversely, if α is sufficient, then there is a decomposition $\mathcal{K} = \bigoplus_n \mathcal{K}_n^L \otimes \mathcal{K}_n^R$ and the corresponding projections q_n satisfy $\alpha(q_n a q_m) = p_n \alpha(a) p_m$. Consequently

$$\alpha(a) = \sum_{n,m} p_n \alpha\Big(\sum_{k,l} q_k a q_l\Big) p_m = \sum_{n,m} p_n \alpha(q_n a q_m) p_m = \sum_i \Big(\sum_n p_n V_i q_n\Big) a \Big(\sum_m q_m V_i^* p_m\Big)$$

Let $V_{i,n} := p_n V_i q_n$, then $V_{i,n} : B(\mathcal{K}_n^L) \otimes B(\mathcal{K}_n^R) \to B(\mathcal{H}_n^L) \otimes B(\mathcal{H}_n^R)$ and

$$\sum_i V_{i,n} a V_{i,n}^* = \alpha_n(a), \qquad a \in B(\mathcal{K}_n^L) \otimes B(\mathcal{K}_n^R).$$

By Theorem 5, there are unitaries $U_n : \mathcal{K}_n^L \to \mathcal{H}_n^L$ and coarse-grainings $\alpha_{2,n} : B(\mathcal{K}_n^R) \to B(\mathcal{H}_n^R)$ such that $\alpha_n = \alpha_{1,n} \otimes \alpha_{2,n}$, in fact, it is easy to see that $\alpha_{2,n}$ have to be completely

positive. This implies that there are linear maps $K_{i,n}: \mathcal{K}_n^R \to \mathcal{H}_n^R$, $\sum_i K_{i,n} K_{i,n}^* = 1_{\mathcal{H}_n^R}$, such that

$$\alpha_n(a) = \sum_i (U_n \otimes K_{i,n}) a (U_n \otimes K_{i,n})^*$$

is another Kraus representation of α_n. Hence there are $\{\mu_{i,j}^n\}$, $\sum_i \mu_{i,j}^n \bar{\mu}_{i,k}^n = \delta_{j,k}$, such that $V_{i,n} = U_n \otimes \sum_j \mu_{i,j}^n K_{j,n}$. Similarly, there are $\nu_{i,j}$, $\sum_i \nu_{i,j} \bar{\nu}_{i,k} = \delta_{j,k}$, such that

$$V_i = \sum_j \nu_{i,j} \left(\sum_n V_{j,n} \right) = \sum_n U_n \otimes L_{i,n}$$

where $L_{i,n} = \sum_{j,k} \nu_{i,j} \mu_{j,k}^n K_{k,n}$. $\qquad\square$

As another corollary, we obtain a result previously proved in [8].

Corollary 2 *Under the assumptions of Corollary 1, suppose that $\mathcal{K} = \mathcal{H}$. Let $D_\theta = \sum_n \varphi_\theta(p_n) D_n(\theta) \otimes D_n^R$ be the $\mathcal{S}$-decomposition. Then $\varphi_\theta \circ \alpha = \varphi_\theta$ for all $\varphi_\theta \in \mathcal{S}$ if and only if*

$$V_i = \sum_n 1_{\mathcal{H}_n^L} \otimes L_{i,n}$$

where $\sum_i L_{i,n} L_{i,n}^* = 1_{\mathcal{H}_n^R}$ and $L_{i,n}$ commutes with D_n^R for all i, n.

Proof. Let α satisfy $\varphi_\theta \circ \alpha = \varphi_\theta$ for all θ, then α is obviously sufficient and by Corollary 1, $V_i = \sum_n U_n \otimes L_{i,n}$. On the other hand, by (21),

$$D_\theta = D_{\theta,0} = \sum_n \varphi_\theta(p_n) U_n^* D_n(\theta) U_n \otimes \alpha_{2,n}^*(D_n^R)$$

and therefore $U_n D_n(\theta) U_n^* = D_n(\theta)$ and $\alpha_{2,n}^*(D_n^R) = \sum_i L_{i,n}^* D_n^R L_{i,n} = D_n^R$ for all θ and n. By construction of the $\mathcal{S}$-decomposition (20), the operators $D_n(\theta)$ generate $B(\mathcal{H}_n^L)$, hence $U_n = 1_{\mathcal{H}_n^L}$. Moreover, the operator D_n^R is in the fixed point space of $\alpha_{2,n}^*$ if and only if it commutes with the Kraus operators $L_{i,n}$ for all i, [7].

The converse statement is obvious. $\qquad\square$

5 Exponential families

A set of measures $\mathcal{P} = \{P_\theta,\ \theta \in \Theta\} \ll \mu$ is an **exponential family** if there are functions $\xi_1, \ldots, \xi_m : \Theta \to \mathbb{R}$ and measurable functions $T_1, \ldots, T_m : X \to \mathbb{R}$ such that for all $\theta \in \Theta$

$$\frac{dP_\theta}{d\mu}(x) = \frac{1}{Z(\theta)} \exp\left(\sum_{i=1}^m \xi_i(\theta) T_i(x) \right) h(x).$$

In this case, all elements in $\mathcal{P}$ are mutually equivalent.

It is immediate from the factorization criterion that the statistic $T = (T_1, \ldots, T_m)$ is sufficient for $\mathcal{P}$. Moreover, it is minimal sufficient if the functions $\{1_{\mathcal{P}}, \xi_1, \ldots, \xi_m\}$ are linearly independent.

Let $\mathcal{P}$ be a family of measures such that the elements are mutually equivalent. Then $\mathcal{P}$ is an exponential family if and only if the linear space spanned by the functions $\{\log \frac{dP}{d\mu}, \ P \in \mathcal{P}\}$, is finite dimensional.

In the non-commutative case, let us assume that ω is a state of the finite dimensional algebra $\mathcal{M}$ and assume that the density of ω is written in the form $\exp H$, $H = H^* \in \mathcal{M}$. Determine the states φ_θ by their density

$$D_\theta := \frac{\exp\left(H + \sum_i \xi_i(\theta) a_i\right)}{Z(\theta)}, \tag{22}$$

where $\xi_1, \ldots, \xi_m : \ \Theta \to \mathbb{R}$ are functions, $a_1, a_2, \ldots, a_m$ are self-adjoint operators from $\mathcal{M}$ and $Z(\theta)$ is for normalization. We call (22) **quantum exponential family** around ω. One can always assume that $\omega(a_i) = 0$ in (22).

The next example tells us how the exponential family arises.

Example 4 Let $a_1, a_2, \ldots, a_m$ be self-adjoint operators from an algebra $\mathcal{M}$ and assume that the density of a state ω is written in the form $\exp H$, $H = H^* \in \mathcal{M}$, moreover $\omega(a_i) = 0$. If Θ is a small neighborhood of $0 \in \mathbb{R}^n$, then minimization of $S(\psi, \omega)$ under the constraints $\psi(a_i) = \theta_i$ ($\theta = (\theta_1, \theta_2, \ldots, \theta_n) \in \Theta, 1 \leq i \leq n$) gives a state D_θ which is of the form (22) and we arrive at an exponential family. The functions $\xi_i(\theta)$ are determined by the constraints

$$\frac{1}{Z(\theta)} \mathrm{Tr} \exp\left(H + \sum_i \xi_i(\theta) a_i\right) a_j = \theta_j,$$

which has a unique solution if θ_j are small enough. $\qquad\square$

Let $\mathcal{M}$ be a von Neumann algebra and ω be a normal state. For $a \in \mathcal{M}^{sa}$ define the state $[\omega^a]$ as the minimizer of

$$\psi \mapsto S(\psi, \omega) - \psi(a). \tag{23}$$

If the density of ω is e^H, then the density of $[\omega^a]$ is nothing else but

$$\frac{\exp\left(H + a\right)}{\mathrm{Tr} \exp\left(H + a\right)},$$

therefore we can extend the above concept of exponential family as

$$\theta \mapsto \varphi_\theta := [\omega^{\sum_i \theta_i a_i}], \tag{24}$$

where $a_1, a_2, \ldots, a_n$ are self-adjoint operators from $\mathcal{M}$. Note that the support of the above states is $\mathrm{supp}\, \omega$. For more details about perturbation of states, see Chap. 12 of

[13] but here we recall the analogue of (22) in the general case. We assume that the von Neumann algebra is in a standard form and the representative of ω is Ω from the positive cone. Let Δ_ω be the modular operator of ω then φ_θ of (24) is the vector state induced by the unit vector

$$\Phi_\theta := \frac{\exp\frac{1}{2}\Big(\log\Delta_\omega + \sum_i \theta_i a_i\Big)\Omega}{\left\|\exp\frac{1}{2}\Big(\log\Delta_\omega + \sum_i \theta_i a_i\Big)\Omega\right\|}. \tag{25}$$

(This formula holds in the strict sense if ω is faithful, since Δ_ω is invertible in this case. For non-faithful ω the formula is modified by the support projection.)

In the next theorem σ_t^ω denotes the modular automorphism group of ω, $\sigma_t^\omega(a) = \Delta_\omega^{it} a \Delta_\omega^{-it}$.

Theorem 6 [15] *Let $\mathcal{M}$ be a von Neumann algebra with a faithful normal state ω and $\mathcal{M}_0$ be a subalgebra. For $a_1, a_2, \ldots, a_n \in \mathcal{A}^{sa}$ the following conditions are equivalent.*

(i) *$\mathcal{M}_0$ is sufficient for the exponential family (24)*

(ii) *$\sigma_t^\omega(a_i) \in \mathcal{M}_0$ for all $t \in \mathbb{R}$ and $1 \le i \le n$.*

(iii) *For the generalized conditional expectation $E_\omega : \mathcal{M} \to \mathcal{M}_0$ $E_\omega(a_i) = a_i$ holds, $1 \le i \le n$.*

Let us denote by $c(\omega, a)$ the minimum in (23), that is, $c(\omega, a) = S([\omega^a], \omega) - [\omega^a](a)$. Then

$$c(\omega, a) = -\log \omega^a(1),$$

where ω^a is the positive functional induced by the vector $\exp\frac{1}{2}\Big(\log\Delta_\omega + a\Big)\Omega$. The function $\theta \mapsto c(\omega, \sum \theta_i a_i)$ is analytic and

$$-\frac{\partial}{\partial\theta_j} c(\omega, \textstyle\sum_i \theta_i a_i) = \varphi_\theta(a_j), \qquad \text{for all } \theta \text{ and } j.$$

Theorem 7 *Let $\mathcal{N}$, $\mathcal{M}$ be von Neumann algebras and let $\alpha : \mathcal{N} \to \mathcal{M}$ be a coarse-graining. Let ω be a faithful normal state on $\mathcal{M}$ and suppose that $\omega_0 := \omega \circ \alpha$ is also faithful. Let φ_θ, $\theta \in \Theta$ be the exponential family $\varphi_\theta = [\omega^{\sum_i \theta_i b_i}]$ for $b_1, \ldots b_k \in \mathcal{M}^{sa}$. Then α is sufficient for $(\mathcal{M}, \varphi_\theta)$ if and only if $b_i = \alpha(a_i)$, $i = 1, \ldots, n$ for some $a_i \in \mathcal{N}^{sa}$ and*

$$\varphi_\theta \circ \alpha = \Big[\omega_0^{\sum_i \theta_i a_i}\Big]. \tag{26}$$

Proof. Let α be sufficient for $(\mathcal{M}, \varphi_\theta)$ and let

$$\mathcal{N}_1 = \{a \in \mathcal{N}, \alpha_\omega^* \circ \alpha(a) = a\} = \{a \in \mathcal{N}_\alpha, \alpha(\sigma_t^{\omega_0}(a)) = \sigma_t^\omega(\alpha(a))\}.$$

22

Then $\alpha(\mathcal{N}_1)$ is a sufficient subalgebra and therefore $\sigma_t^\omega(b_j) \in \alpha(\mathcal{N}_1)$ for all t, $j = 1, \ldots, k$, in particular, $b_i = \alpha(a_i)$, $a_i \in \mathcal{N}_1$. Let $a(\theta) = \sum_j \theta_j a_j$ and consider the expansion

$$
\begin{aligned}
[D\varphi_\theta, D\omega]_t &= [D\omega^{\alpha(a(\theta))}, D\omega]_t \\
&= \sum_{n=0}^{\infty} i^n \int_0^t dt_1 \ldots \int_0^{t_{n-1}} dt_n \sigma_{t_n}^\omega(\alpha(a(\theta)))\ldots\sigma_{t_1}^\omega(\alpha(a(\theta))) \\
&= \sum_{n=0}^{\infty} i^n \int_0^t dt_1 \ldots \int_0^{t_{n-1}} dt_n \alpha(\sigma_{t_n}^{\omega_0}(a(\theta)))\ldots\alpha(\sigma_{t_1}^{\omega_0}(a(\theta))) \\
&= \alpha([D\omega_0^{a(\theta)}, D\omega_0]_t)
\end{aligned}
$$

On the other hand, α is sufficient, therefore $[D\varphi_\theta, \omega]_t \in \alpha(\mathcal{N}_\alpha)$ and

$$\alpha([D\varphi_\theta \circ \alpha, D\omega_0]_t) = [D\varphi_\theta, D\omega]_t$$

As α is invertible on $\mathcal{N}_\alpha$, it follows that $[D\varphi_\theta \circ \alpha, D\omega_0]_t = [D[\omega_0^{a(\theta)}], D\omega_0]_t$ and (26) follows.

Conversely, let $b_i = \alpha(a_i)$ for some $a_i \in \mathcal{N}$ and suppose (26), then

$$\frac{\partial}{\partial\theta_j} c(\omega_0, a(\theta)) = -[\omega_0^{a(\theta)}](a_j) = -\varphi_\theta(\alpha(a_j)) = \frac{\partial}{\partial\theta_j} c(\omega, \alpha(a(\theta)))$$

for all θ and j. Putting $\theta = 0$, it follows that $c(\omega_0, a(\theta)) = c(\omega, \alpha(a(\theta)))$ for all θ. Hence

$$S(\varphi_\theta, \omega) = c(\omega, \alpha(a(\theta))) + \varphi_\theta(\alpha(a(\theta))) = c(\omega_0, a(\theta)) + \varphi_\theta \circ \alpha(a(\theta)) = S(\varphi_\theta \circ \alpha, \omega \circ \alpha)$$

and α is sufficient. □

Remark. Note that in case $\mathcal{M} = B(\mathcal{H})$, $\dim \mathcal{H} = n$, the condition (26) reads

$$\alpha(\log \alpha^*(D_\theta) - \log \alpha^*(D_{\omega_0})) = \log D_\theta - \log D_\omega,$$

where α^* is the dual of α with respect to $\langle A, B \rangle = \mathrm{Tr} A^* B$. This condition is known to be equivalent to sufficiency of α. □

Corollary 3 *Let $\mathcal{M}$ be a von Neumann algebra with a faithful normal state ω, $\mathcal{M}_0$ a commutative subalgebra and (24) the exponential family for $a_1, a_2, \ldots, a_n \in \mathcal{M}^{sa}$. Then $\mathcal{M}_0$ is sufficient for the exponential family if and only if $a_1, \ldots, a_n \in \mathcal{M}_0$ and*

$$\varphi_\theta(a) = \omega(\exp(\textstyle\sum_i \theta_i a_i)a) \quad a \in \mathcal{M}$$

Proof. Let $\mathcal{M}_1$ be the subalgebra generated by $\sigma_t^\omega(a_i)$, $t \in R$, $i = 1, \ldots, n$. Then $\mathcal{M}_1$ is sufficient, by Theorem 6. Let $E : \mathcal{M} \to \mathcal{M}_1$ be the ω preserving conditional expectation, then E preserves all φ_θ, by sufficiency (Theorem 1 (iv)). If $\mathcal{M}_0$ is sufficient,

then $\mathcal{M}_1 \subseteq \mathcal{M}_0$, hence $\mathcal{M}_1$ is commutative. Let $\omega_0, \varphi_{\theta,0}$ be the restriction of ω, φ_θ to $\mathcal{M}_1$, then by Theorem 7,

$$\varphi_{\theta,0} = [\omega_0^{\sum \theta_i a_i}] = \omega_0(\exp(\sum \theta_i a_i) \cdot)$$

It follows that for $a \in \mathcal{M}$,

$$\varphi_\theta(a) = \varphi_{\theta,0}(E(a)) = \omega_0(\exp(\sum \theta_i a_i)E(a)) = \omega(\exp(\sum \theta_i a_i)a)$$

Conversely, let $a_1, \dots, a_n \in \mathcal{M}_0$ and let $\varphi_\theta = \omega(\exp \sum \theta_i a_i) \cdot)$, then the restriction of φ_θ to $\mathcal{M}_0$ is the exponential family $[\omega_0^{\sum \theta_i a_i}]$ and $\mathcal{M}_0$ is sufficient, by Theorem 7. $\square$

6 Strong subadditivity of entropy

Let $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C$ and let ω_{ABC} be a normal state on $B(\mathcal{H})$ with restrictions ω_B, ω_{AB} and ω_{BC}. The von Neumann entropies satisfies the **strong subadditivity**

$$S(\omega_{ABC}) + S(\omega_B) \le S(\omega_{AB}) + S(\omega_{BC}), \tag{27}$$

which was obtained by Lieb and Ruskai [9]. A concise proof using the Jensen operator inequality is contained in [16] and [12] is a didactical presentation of the same ideas. As we want to investigate the case of equality mostly, we suppose below that all the involved entropies are finite. The case of equality was studied in several papers recently but always restricted to finite dimensional Hilbert spaces [7, 10]. Our aim now is to allow infinite dimensional spaces.

The strong subadditivity is equivalent to

$$S(\omega_{AB}, \omega_A \otimes \omega_B) \le S(\omega_{ABC}, \omega_A \otimes \omega_{BC}), \tag{28}$$

which is a consequence of monotonicity of the relative entropy. Clearly, the equality in (27) is equivalent to equality in (28) which means that $B(\mathcal{H}_A) \otimes B(\mathcal{H}_B)$ is a sufficient subalgebra for the states ω_{ABC} and $\omega_A \otimes \omega_{BC}$. Our results on factorization apply.

Theorem 8 *Let ω_{ABC} be a faithful normal state on $B(\mathcal{H})$ such that the von Neumann entropy $S(\omega_{ABC})$ is finite and*

$$S(\omega_{ABC}) + S(\omega_B) = S(\omega_{AB}) + S(\omega_{BC}).$$

Then there is a decomposition $\mathcal{H}_B = \bigoplus_n \mathcal{H}_{nB}^L \otimes \mathcal{H}_{nB}^R$ such that

$$\omega_{ABC} = \sum_n \omega_B(p_n) D_n^L \otimes D_n^R, \tag{29}$$

where $D_n^L \in B(\mathcal{H}_A) \otimes B(\mathcal{H}_{nB}^L)$ and $D_n^R \in B(\mathcal{H}_{nB}^R) \otimes B(\mathcal{H}_C)$ are density operators and $p_n \in B(\mathcal{H}_B)$ are the orthogonal projections $\mathcal{H}_B \to \mathcal{H}_{nB}^L \otimes \mathcal{H}_{nB}^R$.

Proof. Equality in the strong subadditivity is equivalent to sufficiency of the subalgebra $B(\mathcal{H}_A \otimes \mathcal{H}_B) \otimes \mathbb{C}1_C$ for $(B(\mathcal{H}), \mathcal{S})$ where $\mathcal{S} := \{\omega_{ABC}, \omega_A \otimes \omega_{BC}\}$, and the latter is equivalent to

$$[D\omega_{ABC}, D\omega_A \otimes \omega_{BC}]_t = [D\omega_{AB}, D\omega_A \otimes \omega_B]_t \otimes 1_C$$

for all t. Let $\mathcal{N}_B \subset B(\mathcal{H}_B)$ be the subalgebra

$$\mathcal{N}_B = \{b \in B(\mathcal{H}_B) : \ \sigma_t^{\omega_{BC}}(b \otimes I_C) = \sigma_t^{\omega_B}(b) \otimes I_C \text{ for every } t \in \mathbb{R}\} \,.$$

Then in follows from the above equality and the cocycle condition that

$$[D\omega_{AB}, D(\omega_A \otimes \omega_B)]_t \in B(\mathcal{H}_A) \otimes \mathcal{N}_B \quad \text{for all } t$$

and therefore $B(\mathcal{H}_A) \otimes \mathcal{N}_B \otimes \mathbb{C}1_C$ is sufficient for $\mathcal{S}$. Since $\omega_A \otimes \omega_{BC}$ is faithful, $S(\omega_A \otimes \omega_{BC}) < \infty$ and clearly dominates $\mathcal{S}$, moreover, the subalgebra is invariant under $\sigma_t^{\omega_A \otimes \omega_{BC}}$, we have by Theorem 3 that there is a decomposition

$$D_{ABC} = (D_L \otimes 1_C)(1_A \otimes D_R),$$

where $D_L \in B(\mathcal{H}_A) \otimes \mathcal{N}_B$, $D_R \in \mathcal{N}_B' \otimes B(\mathcal{H}_C)$ are density operators.

On the other hand, $\mathcal{N}_B$ is invariant under $\sigma_t^{\omega_B}$, therefore $S(\omega_B|_{\mathcal{N}_B}) \leq S(\omega_B) < \infty$. Similarly as in Section 4, we obtain a decomposition $\mathcal{H}_B = \bigoplus_n \mathcal{H}_{nB}^L \otimes \mathcal{H}_{nB}^R$ such that

$$\mathcal{N}_B = \bigoplus_n B(\mathcal{H}_{nB}^L) \otimes \mathbb{C}1_{\mathcal{H}_{nB}^R} \qquad \mathcal{N}_B' = \bigoplus \mathbb{C}1_{\mathcal{H}_{nB}^L} \otimes B(\mathcal{H}_{nB}^R)$$

and (29) follows. $\qquad\square$

The structure (29) of the density matrix ω_{ABC} is similar to the finite dimensional situation discussed in [7, 10], however the direct sum decomposition may be infinite.

The theorem is stated under the condition of faithfulness of ω_{ABC}. It would be worthwhile to weaken this condition. When ω_{ABC} is pure the strong subadditivity reduces to

$$S(\omega_{AC}) \leq S(\omega_A) + S(\omega_C),$$

which is simply the subadditivity. The equality holds here if $\omega_{AC} = \omega_A \otimes \omega_C$. Since the purification of a product state is a product vector, we have the product structure (29) (without the summation over n). Note that this kind of states were discussed in [21].

The decomposition (29) has a continuous version formulated in terms of direct integrals (see [14] for references about the direct integral of fields of Hilbert spaces and operators or [22]). Let (X, μ) be a measure space. Assume that for $x \in X$ density matrices $D^L(x) \in B(\mathcal{H}_A) \otimes B(\mathcal{H}^L(x))$ and $D^R(x) \in B(\mathcal{H}^R(x)) \otimes B(\mathcal{H}_C)$ such that $\mathcal{H}^L(x)$ and $\mathcal{H}^R(x)$ are measurable fields of Hilbert spaces and the operator fields

$D^L(x)$ and $D^R(x)$ are measurable as well, $x \in X$. Given a probability density $p(x)$ on X

$$\omega_{ABC} := \int_\oplus p(x) D^L(x) \otimes D^R(x) \, d\mu(x) \tag{30}$$

is a density on the Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C$, where

$$\mathcal{H}_B := \int_\oplus \mathcal{H}^L(x) \otimes \mathcal{H}^R(x) \, d\mu(x) \, .$$

Then $B(\mathcal{H}_A) \otimes B(\mathcal{H}_B)$ is a sufficient subalgebra for the states ω_{ABC} and $\omega_A \otimes \omega_{BC}$. If the measure μ is not atomic, then $S(\omega_{ABC}) = \infty$.

7 Appendix

Dual mapping

Let $\mathcal{M}_1$ and $\mathcal{M}_2$ be von Neumann algebras and let $\sigma : \mathcal{M}_1 \to \mathcal{M}_2$ be a coarse-graining. Suppose that a normal state φ_2 is given and $\varphi_1 := \varphi_2 \circ \sigma$ is normal as well.

We assume that both von Neumann algebras are in a standard form and the representative of φ_i is Φ_i from the positive cone. From the modular theory we know that

$$p_i := \overline{J_i \mathcal{M}_i \Phi_i}$$

is the support projection of φ_i (i=1,2).

The dual $\alpha : p_2 \mathcal{M}_2 p_2 \to p_1 \mathcal{M}_1 p_1$ of σ is is characterized by the property

$$\langle A_1, J_1 \alpha(A_2) \rangle = \langle \sigma(A_1), J_2 A_2 \rangle \tag{31}$$

(see Prop. 8.3 in [13]). The dual of the embedding of a subalgebra into an algebra is called generalized conditional expectation [1].

References

[1] L. ACCARDI AND C. CECCHINI, Conditional expectations in von Neumann algebras and a theorem of Takesaki, J. Functional. Anal. **45**(1982), 245–273.

[2] O. BARNDORFF-NIELSEN, *Information and exponential families in statistical theory*, Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons, Ltd., Chichester, 1978.

[3] O.E. BARNDORFF-NIELSEN, R. GILL AND P.E. JUPP, On quantum statistical inference, J. R. Stat. Soc. Ser. B Stat. Methodol. **65** (2003), 775–816.

[4] J. Blank, P. Exner and M. Havlíček, *Hilbert space operators in quantum physics*, American Institute of Physics, 1994.

[5] O. Bratteli and D. W. Robinson, *Operator algebras and quantum statistical mechanics. 1. C*- and W*–algebras, symmetry groups, decomposition of states*, 2nd ed., Texts and Monographs in Physics, Springer Verlag, New York, 1987.

[6] F. Hansen and G.K. Pedersen, Jensen's inequality for operator and Löwner's theorem, Math. Anal. **258**, 229–241 (1982).

[7] P. Hayden, R. Jozsa, D. Petz and A. Winter, Structure of states which satisfy strong subadditivity of quantum entropy with equality, Commun. Math. Phys. **246**(2004), 359–374.

[8] M. Koashi and N. Imoto, Operations that do not disturb partially known quantum states, Phys. Rev. A, **66**(2002), 022318.

[9] E. H. Lieb and M.B. Ruskai, Proof of the strong subaddititivy of quantum mechanical entropy, J. Math. Phys. **14**(1973), 1938–1941.

[10] M. Mosonyi and D. Petz, Structure of sufficient quantum coarse-grainings, Lett. Math. Phys. **68**(2004), 19–30.

[11] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*, Cambridge University Press, 2000.

[12] M. A. Nielsen and D. Petz, A simple proof of the strong subadditivity inequality, quant-ph/0408130, 2004.

[13] M. Ohya and D. Petz, *Quantum Entropy and Its Use*, Springer-Verlag, Heidelberg, 1993, 2nd edition 2004.

[14] D. Petz, Direct integral of multifunctions into von Neumann algebras, Studia Sci. Math. Hungar. **18**(1978), 239–245.

[15] D. Petz, Sufficient subalgebras and the relative entropy of states of a von Neumann algebra, Commun. Math. Phys. **105**(1986), 123–131.

[16] D. Petz, Quasi-entropies for finite quantum systems, Rep. Math. Phys. **21**(1986), pp. 57–65

[17] D. Petz, Sufficiency of channels over von Neumann algebras, Quart. J. Math. Oxford, **39**(1988), 907–1008.

[18] D. Petz, Geometry of canonical correlation on the state space of a quantum system, J. Math. Phys. **35**(1994), 780–795.

[19] D. Petz, Discrimination between states of a quantum system by observations, J Funct. Anal. **120**(1994), 82–97.

[20] M. B. Ruskai, Inequalities for quantum entropy: A review with conditions with equality, J. Math. Phys. **43**(2002), 4358–4375.

[21] B. W. Schumacher and M. A. Nielsen, Quantum Data Processing and Error Correction, arXiv:quant-ph/9604022.

[22] J.T. Schwartz, *W*-algebras*, Gordon and Breach Science Publishers, New York-London-Paris, 1967.

[23] H. Strasser, *Mathematical theory of statistics. Statistical experiments and asymptotic decision theory*, Walter de Gruyter, Berlin, 1985.