Comparison of quantum channels and statistical experiments

Anna Jenčová
Mathematical Institute, Slovak Academy of Sciences
Štefánikova 49, Bratislava, Slovakia
Email: jenca@mat.savba.sk

***Abstract*—For a pair of quantum channels with the same input space, we show that the possibility of approximation of one channel by post-processings of the other channel can be characterized by comparing the success probabilities for the two ensembles obtained as outputs for any ensemble on the input space coupled with an ancilla. This provides an operational interpretation to a natural extension of Le Cam's deficiency to quantum channels. In particular, we obtain a version of the randomization criterion for quantum statistical experiments. The proofs are based on some properties of the diamond norm and its dual, which are of independent interest.**

I. Introduction

The theory of comparison of statistical experiments started in the work of Blackwell [1], who introduced a natural ordering of experiments in terms of the risks of optimal decision rules. This ordering was extended by Le Cam [2] into a deficiency measure on statistical experiments, expressing how well an experiment $\mathcal{S}$ can be approximated by randomizations of another experiment $\mathcal{T}$. Le Cam's randomization criterion shows that deficiency also gives the maximal loss in the average payoffs of decision procedures, experienced when the experiment $\mathcal{S}$ is replaced by $\mathcal{T}$. For an account on comparison of statistical experiments, see e.g. [3, 4].

An extension of Blackwell's results for quantum experiments was first obtained by Shmaya [5] in the framework of quantum information structures. In [6], a theory of comparison for both classical and quantum experiments is developed in terms of statistical morphisms. In both works, either additional entanglement or composition of the experiment with a complete set of states is required. Quantum versions of Le Cam's randomization criterion were studied in [7, 8]. In particular, Matsumoto in [8] introduced a natural generalization of classical decision problems to quantum ones and proved a quantum randomization criterion in this setting. The main drawback of this approach is the lack of operational interpretation for quantum decision problems.

Comparison of channels can be obtained as an extension of the theory of comparison of experiments. A natural idea is the following: given two channels with the same input space, compare the two experiments emerging as outputs for a single input experiment. If the output experiment of the channel Ψ is always more informative than the output of the channel Φ, we say that Ψ is less noisy than Φ. An ordering of classical channels was first introduced in the work by Shannon [9], where a coding/decoding criterion was applied. Similar orderings were studied in e.g. [10, 11]. For some more recent works see e.g. [12, 13].

In the quantum setting, it is possible to use a stronger ordering, namely to consider experiments on the input space coupled with an ancilla. As it turns out, for quantum channels, Ψ is less noisy in this stronger sense if and only if Φ is a post-processing of Ψ. In fact, it is enough to compare guessing probabilities for ensembles of states. This remarkable result was first obtained by Chefles in [14], based on [5]. It was extended and refined in [6], in particular it was proved that no entanglement in the input ensemble is needed. Some applications were already found in [13, 15–17].

The aim of the present work is to establish an approximate version of these results, which may be called the randomization criterion for quantum channels. More precisely, we study an extension of Le Cam's deficiency for quantum channels, based on the diamond norm. Such definitions appear naturally in quantum information theory, for example the approximate (anti)degradable channels, [18]. We show that deficiency can be characterized by comparing success probabilities for output ensembles, with respect to the success probability of the input ensemble. These results are then applied to statistical experiments and a quantum randomization criterion is proved in terms of success probabilities.

The diamond norm appears as a distinguishability norm for quantum channels [19]. As it was observed in [20], this norm can be defined using the order structure given by the cone of completely positive maps. We also show that the dual norm on positive elements can be expressed as the optimal success probability for a certain ensemble. These properties provide a convenient framework for proving our results and are of independent interest.

II. Notations and preliminaries

If not stated otherwise, the full proofs can be found in [21].

Throughout the paper, all Hilbert spaces are finite dimensional. If $\mathcal{H}$ is a Hilbert space, we fix an orthonormal basis $\{|e_i\rangle, i = 1, \dots, \dim(\mathcal{H})\}$ in $\mathcal{H}$. We will denote the algebra of linear operators on $\mathcal{H}$ by $B(\mathcal{H})$, the set of positive operators by $B(\mathcal{H})^+$ and the real vector space of self-adjoint elements

This work was supported by the grants VEGA 2/0069/16 and by Science and Technology Assistance Agency under the contract no. APVV-0178-11.

978-1-5090-1805-5/16/$31.00 ©2016 IEEE

by $B_h(\mathcal{H})$. The set of states, or density operators, on $\mathcal{H}$ will be denoted by $\mathfrak{S}(\mathcal{H}) := \{\sigma \in B(\mathcal{H})^+,\ \mathrm{Tr}\,\sigma = 1\}$.

Let $\mathcal{L}(\mathcal{H},\mathcal{K})$ denote the real vector space of Hermitian linear maps $B(\mathcal{H}) \to B(\mathcal{K})$. The set $\mathcal{L}(\mathcal{H},\mathcal{K})^+$ of completely positive maps forms a closed convex cone in $\mathcal{L}(\mathcal{H},\mathcal{K})$ which is pointed and generating. With this cone, $\mathcal{L}(\mathcal{H},\mathcal{K})$ becomes an ordered vector space. We will denote the corresponding order by $\le$. An element of $\mathcal{L}(\mathcal{H},\mathcal{K})^+$ that preserves trace is usually called a channel. We will denote the set of all channels by $\mathcal{C}(\mathcal{H},\mathcal{K})$.

For $\phi \in \mathcal{L}(\mathcal{H},\mathcal{H})$, we define

$$s(\phi) = \sum_{i,j} \langle e_i, \phi(|e_i\rangle\langle e_j|)e_j\rangle.$$

It is easy to see that s defines a linear functional $\mathcal{L}(\mathcal{H},\mathcal{H}) \to \mathbb{R}$ and for all $\phi \in \mathcal{L}(\mathcal{H},\mathcal{K})$, $\psi \in \mathcal{L}(\mathcal{K},\mathcal{H})$, $s(\psi \circ \phi) = s(\phi \circ \psi)$.

We now identify the dual space of $\mathcal{L}(\mathcal{H},\mathcal{K})$ with $\mathcal{L}(\mathcal{K},\mathcal{H})$, where duality is given by

$$\langle \psi, \phi\rangle = s(\psi \circ \phi), \qquad \phi \in \mathcal{L}(\mathcal{H},\mathcal{K}),\ \psi \in \mathcal{L}(\mathcal{K},\mathcal{H}).$$

Note that the tracelike property of s implies that we have $\langle \phi, \psi\rangle = \langle \psi, \phi\rangle$ and

$$\langle \phi, \xi \circ \psi\rangle = \langle \phi \circ \xi, \psi\rangle = \langle \psi \circ \phi, \xi\rangle.$$

The dual cone of positive functionals satisfies

$$\begin{aligned}(\mathcal{L}(\mathcal{H},\mathcal{K})^+)^* &:= \{\psi \in \mathcal{L}(\mathcal{K},\mathcal{H}), \langle \psi, \phi\rangle \ge 0, \forall \phi \in \mathcal{L}(\mathcal{H},\mathcal{K})^+\}\\ &= \mathcal{L}(\mathcal{K},\mathcal{H})^+,\end{aligned}$$

so that the cone of completely positive maps is self-dual.

Remark 1. Let us denote

$$X_{\mathcal{H}} := \sum_{i,j} |e_i\rangle\langle e_j| \otimes |e_i\rangle\langle e_j| \in B(\mathcal{H}\otimes\mathcal{H})^+.$$

The Choi representation $C : \phi \mapsto (\phi \otimes id_{\mathcal{H}})(X_{\mathcal{H}})$ provides an order isomorphism of $\mathcal{L}(\mathcal{H},\mathcal{K})$ onto $B_h(\mathcal{K}\otimes\mathcal{H})$ with the cone of positive operators $B(\mathcal{K}\otimes\mathcal{H})^+$. Note also that for any $\phi \in \mathcal{L}(\mathcal{H},\mathcal{H})$, $s(\phi) = \mathrm{Tr}\, C(\phi)X_{\mathcal{H}}$, so that

$$\langle \psi, \phi\rangle = s(\psi\circ\phi) = \mathrm{Tr}\,[C(\psi\circ\phi)X_{\mathcal{H}}] = \mathrm{Tr}\,[C(\phi)C(\psi^*)].$$

It is of course possible to use the Choi representation with this duality, but for our purposes it is mostly more convenient to work with the spaces of mappings.

III. The diamond norm and its dual

The diamond norm in $\mathcal{L}(\mathcal{H},\mathcal{K})$ is defined by

$$\|\phi\|_\diamond = \sup_{\rho\in\mathfrak{S}(\mathcal{H}\otimes\mathcal{H})} \|(\phi\otimes id)(\rho)\|_1, \tag{1}$$

where $\|\cdot\|_1$ denotes the trace norm in $B(\mathcal{K}\otimes\mathcal{H})$. It was proved in [20] that this norm is obtained from the set of channels and the order structure in $\mathcal{L}(\mathcal{H},\mathcal{K})$. Namely, for $\phi \in \mathcal{L}(\mathcal{H},\mathcal{K})$,

$$\|\phi\|_\diamond = \inf_{\alpha\in\mathcal{C}(\mathcal{H},\mathcal{K})} \inf\{\lambda > 0, -\lambda\alpha \le \phi \le \lambda\alpha\}. \tag{2}$$

It was also shown that the dual norm in $\mathcal{L}(\mathcal{K},\mathcal{H})$, which we will denote by $\|\cdot\|^\diamond$, is similarly obtained from the set of erasure channels $\{\phi_\sigma : B(\mathcal{K}) \ni A \mapsto \mathrm{Tr}\,[A]\sigma,\ \sigma \in \mathfrak{S}(\mathcal{H})\}$:

$$\|\psi\|^\diamond = \inf_{\sigma\in\mathfrak{S}(\mathcal{H})} \inf\{\lambda > 0, -\lambda\phi_\sigma \le \psi \le \lambda\phi_\sigma\}. \tag{3}$$

We list some useful properties of these norms.

Proposition 1. (i) *If $\phi \in \mathcal{L}(\mathcal{H},\mathcal{K})^+$, then*

$$\|\phi\|_\diamond = \sup_{\sigma\in\mathfrak{S}(\mathcal{H})} \mathrm{Tr}\,[\phi(\sigma)], \quad \|\phi\|^\diamond = \sup_{\alpha\in\mathcal{C}(\mathcal{K},\mathcal{H})} \langle \alpha, \phi\rangle.$$

(ii) *If $\phi, \psi \in \mathcal{C}(\mathcal{H},\mathcal{K})$, then*

$$\|\phi-\psi\|_\diamond = 2 \sup_{\gamma\ge 0, \|\gamma\|^\diamond \le 1} \langle \gamma, \phi-\psi\rangle$$

(iii) *If $\chi \in \mathcal{C}(\mathcal{K},\mathcal{K}')$ and $\xi \in \mathcal{C}(\mathcal{H}',\mathcal{H})$, then the maps $\phi \mapsto \chi\circ\phi$ and $\phi \mapsto \phi\circ\xi$ are contractions with respect to both $\|\cdot\|_\diamond$ and $\|\cdot\|^\diamond$.*

An important property of the dual norm is its relation to success probabilities for ensembles of quantum states. Let $\mathcal{E} = \{\lambda_i, \sigma_i\}_{i=1}^k$ be and ensemble on $\mathcal{H}$, here $\sigma_i \in \mathfrak{S}(\mathcal{H})$ and $\lambda_1, \dots, \lambda_k$ are prior probabilities. In the setting of multiple hypothesis testing, the task is to guess which one is the true state. Any procedure to obtain such a guess can be identified with some POVM $M = \{M_1, \dots, M_k\}$, $M_i \in B(\mathcal{H})^+$, $\sum_i M_i = I$. Here $\mathrm{Tr}\,\sigma_i M_j$ is interpreted as the probability that σ_j is chosen while the true state is σ_i, so that the average success probability for the procedure M is $\sum_i \lambda_i \mathrm{Tr}\, M_i\sigma_i$. One can show that the maximum probability of a successful guess for this ensemble has the form $P_{succ}(\mathcal{E}) = \|\phi_{\mathcal{E}}\|^\diamond$, where $\phi_{\mathcal{E}} \in \mathcal{L}(\mathbb{C}^n,\mathcal{K})$ is the map $A \mapsto \sum_i A_{ii}\lambda_i\sigma_i$. More generally, we have

Proposition 2. *Let $\gamma \in \mathcal{L}(\mathcal{K},\mathcal{H})^+$. Then there is an (equiprobable) ensemble $\mathcal{E}_\gamma$ on $\mathcal{H}\otimes\mathcal{K}$ such that*

$$\|\gamma\|^\diamond = \dim(\mathcal{K})\mathrm{Tr}\,[\gamma(I)]P_{succ}(\mathcal{E}_\gamma).$$

Moreover, for any $\phi \in \mathcal{L}(\mathcal{H},\mathcal{H}')^+$, we have

$$\mathcal{E}_{\phi\circ\gamma} = (\phi\otimes id)(\mathcal{E}_\gamma).$$

IV. The main result

Let $\Phi \in \mathcal{C}(\mathcal{H},\mathcal{K})$ and $\Psi \in \mathcal{C}(\mathcal{H},\mathcal{K}')$. Similarly to Le Cam's deficiency for statistical experiments, we may define the deficiency of Φ with respect to Ψ by

$$\delta(\Phi,\Psi) = \inf_{\alpha\in\mathcal{C}(\mathcal{K}',\mathcal{K})} \|\Phi - \alpha\circ\Psi\|_\diamond.$$

Since $\mathcal{C}(\mathcal{K}',\mathcal{K})$ is convex and compact, the infimum is attained, in particular, $\delta(\Phi,\Psi) = 0$ if and only if $\Phi = \alpha\circ\Psi$ for some $\alpha \in \mathcal{C}(\mathcal{K}',\mathcal{K})$. In this case, we write $\Phi \preceq \Psi$ We also define Le Cam distance by

$$\Delta(\Phi,\Psi) = \max\{\delta(\Phi,\Psi), \delta(\Psi,\Phi)\}.$$

This defines a preorder on the set of channels with the same input space. The following data processing inequalities for δ are obvious consequences of their definition and Proposition 1 (iii).

Proposition 3. *Let $\Phi_1, \Phi_2, \Phi, \Psi_1, \Psi_2, \Psi$ be channels with the same input space.*

(i) *If $\Phi_1 \preceq \Phi_2$, then $\delta(\Phi_1, \Psi) \leq \delta(\Phi_2, \Psi)$.*
(ii) *If $\Psi_1 \preceq \Psi_2$, then $\delta(\Phi, \Psi_1) \geq \delta(\Phi, \Psi_2)$.*

Let now $\delta(\Phi, \Psi) = 0$. Then for any ensemble $\mathcal{E}$ on the tensor product $\mathcal{H} \otimes \mathcal{H}_0$ with an ancillary Hilbert space $\mathcal{H}_0$, we have

$$P_{succ}((\Phi \otimes id_{\mathcal{H}_0})(\mathcal{E})) \leq P_{succ}((\Psi \otimes id_{\mathcal{H}_0})(\mathcal{E})). \quad (4)$$

The converse was proved in [6, 14]. Our aim is to prove an ϵ-version of this result.

Theorem 1. *Let $\Phi \in \mathcal{C}(\mathcal{H}, \mathcal{K})$, $\Psi \in \mathcal{C}(\mathcal{H}, \mathcal{K}')$, $\epsilon \geq 0$. Then $\delta(\Phi, \Psi) \leq \epsilon$ if and only for any finite dimensional Hilbert space $\mathcal{K}_0$ and any ensemble $\mathcal{E}$ on $\mathcal{H} \otimes \mathcal{K}_0$,*

$$P_{succ}((\Phi \otimes id_{\mathcal{K}_0})(\mathcal{E})) \leq P_{succ}((\Psi \otimes id_{\mathcal{K}_0})(\mathcal{E})) + \frac{\epsilon}{2} P_{succ}(\mathcal{E})$$

Moreover, one can restrict to $\mathcal{K}_0 = \mathcal{K}$ and equiprobable ensembles with $k = \dim(\mathcal{K})^2$ elements.

Proof. Assume that $\alpha \in \mathcal{C}(\mathcal{K}', \mathcal{K})$ is a channel such that

$$\|\Phi - \alpha \circ \Psi\|_\diamond \leq \epsilon.$$

Then for any $\gamma \in \mathcal{L}(\mathcal{K}_0, \mathcal{H})^+$ and $\chi \in \mathcal{C}(\mathcal{K}, \mathcal{K}_0)$, we have by positivity and Proposition 1 (ii), (iii) that

$$\begin{aligned} \langle \gamma, \chi \circ \Phi \rangle &\leq \langle \gamma, \chi \circ \alpha \circ \Psi \rangle + |\langle \gamma, \chi \circ (\Phi - \alpha \circ \Psi) \rangle| \\ &\leq \langle \gamma, \chi \circ \alpha \circ \Psi \rangle + \frac{1}{2} \|\gamma\|^\diamond \|\chi \circ (\alpha \circ \Psi - \Phi)\|_\diamond \\ &\leq \langle \gamma, \chi \circ \alpha \circ \Psi \rangle + \frac{1}{2} \epsilon \|\gamma\|^\diamond \end{aligned}$$

From this, we obtain by Proposition 1 (i) and properties of s that

$$\begin{aligned} \|\Phi \circ \gamma\|^\diamond &= \sup_{\chi \in \mathcal{C}(\mathcal{K}, \mathcal{K}_0)} \langle \chi, \Phi \circ \gamma \rangle = \sup_{\chi \in \mathcal{C}(\mathcal{K}, \mathcal{K}_0)} \langle \gamma, \chi \circ \Phi \rangle \\ &\leq \sup_{\chi \in \mathcal{C}(\mathcal{K}, \mathcal{K}_0)} \langle \gamma, \chi \circ \alpha \circ \Psi \rangle + \frac{1}{2} \epsilon \|\gamma\|^\diamond \\ &= \sup_{\chi \in \mathcal{C}(\mathcal{K}, \mathcal{K}_0)} \langle \chi \circ \alpha, \Psi \circ \gamma \rangle + \frac{1}{2} \epsilon \|\gamma\|^\diamond \\ &\leq \|\Psi \circ \gamma\|^\diamond + \frac{1}{2} \epsilon \|\gamma\|^\diamond. \end{aligned}$$

Hence we have proved that $\delta(\Phi, \Psi) \leq \epsilon$ implies

$$\|\Phi \circ \gamma\|^\diamond \leq \|\Psi \circ \gamma\|^\diamond + \frac{1}{2} \epsilon \|\gamma\|^\diamond, \quad \forall \gamma \in \mathcal{L}(\mathcal{K}_0, \mathcal{H})^+. \quad (5)$$

Since by (1) we have $\|\phi\|_\diamond = \|\phi \otimes id_{\mathcal{K}_0}\|_\diamond$ for any $\mathcal{K}_0$, we also have $\delta(\Phi \otimes id_{\mathcal{K}_0}, \Psi \otimes id_{\mathcal{K}_0}) \leq \epsilon$. Hence we obtain

$$\|(\Phi \otimes id_{\mathcal{K}_0}) \circ \gamma\|^\diamond \leq \|(\Psi \otimes id_{\mathcal{K}_0}) \circ \gamma\|^\diamond + \frac{\epsilon}{2} \|\gamma\|^\diamond \quad (6)$$

for all $\gamma \in \mathcal{L}(\mathcal{K}_1, \mathcal{H} \otimes \mathcal{K}_0)$ and any $\mathcal{K}_1$. If $\mathcal{E}$ is any ensemble on $\mathcal{H} \otimes \mathcal{K}_0$, then

$$P_{succ}((\Phi \otimes id)(\mathcal{E})) = \|\phi_{(\Phi \otimes id)(\mathcal{E})}\|^\diamond = \|(\Phi \otimes id) \circ \phi_{\mathcal{E}}\|^\diamond$$

and similarly for Ψ. Putting $\gamma = \phi_{\mathcal{E}}$ in (6) implies the desired inequality.

For the converse, note that by Proposition 1 (ii), we have

$$\delta(\Phi, \Psi) = 2 \min_{\alpha \in \mathcal{C}(\mathcal{K}', \mathcal{K})} \left\{ \max_{\substack{\gamma \in \mathcal{L}(\mathcal{K}, \mathcal{H})^+, \\ \|\gamma\|^\diamond \leq 1}} \langle \gamma, \Phi - \alpha \circ \Psi \rangle \right\}$$

Since the sets $\mathcal{C}(\mathcal{K}', \mathcal{K})$ and $\{\gamma \in \mathcal{L}(\mathcal{K}, \mathcal{H})^+, \|\gamma\|^\diamond \leq 1\}$ are both convex and compact and the map $(\alpha, \gamma) \mapsto \langle \gamma, \Phi - \alpha \circ \Psi \rangle$ is linear in both variables, we may apply the minimax theorem, see e.g. [3]. It follows that

$$\begin{aligned} \delta(\Phi, \Psi) &= 2 \max_\gamma \min_\alpha \langle \gamma, \Phi - \alpha \circ \Psi \rangle \\ &= 2 \max_\gamma \{\langle \gamma, \Phi \rangle - \|\Psi \circ \gamma\|^\diamond\} \\ &\leq 2 \max_\gamma \{\|\Phi \circ \gamma\|^\diamond - \|\Psi \circ \gamma\|^\diamond\}. \end{aligned}$$

Proposition 2 and the assumption now imply that the last expression is less that ϵ.

□

In the case $\epsilon = 0$, we obtain a stronger condition. Similar results were proved in [6].

Theorem 2. *Let $\Phi \in \mathcal{C}(\mathcal{H}, \mathcal{K})$, $\Psi \in \mathcal{C}(\mathcal{H}, \mathcal{K}')$ and let $\xi \in \mathcal{C}(\mathcal{K}_0, \mathcal{K})$ be a surjective channel. Then $\delta(\Phi, \Psi) = 0$ if and only if for any ensemble $\mathcal{E}$ on $\mathcal{H} \otimes \mathcal{K}_0$,*

$$P_{succ}((\Phi \otimes \xi)(\mathcal{E})) \leq P_{succ}((\Psi \otimes \xi)(\mathcal{E})).$$

In particular, by choosing ξ as a classical-to-quantum channel of the form $A \mapsto \sum_i A_{ii} \sigma_i$ for a set of states $\{\sigma_i\}$ that spans $B(\mathcal{K})$, we see that for $\epsilon = 0$ we may restrict to ensembles of separable states.

V. The randomization criterion for quantum experiments

A quantum statistical experiment is a pair $\mathcal{T} = (\mathcal{H}, \{\rho_\theta, \theta \in \Theta\})$, where $\rho_\theta \in \mathfrak{S}(\mathcal{H})$ for all $\theta \in \Theta$ and Θ is an arbitrary set of parameters. Any experiment can be viewed as the set of possible states of some physical system, determined by some prior information on the true state. Note that this definition contains also classical statistical experiments on finite sample spaces, which can be identified with diagonal density matrices.

Based on the outcome of a measurement on the system, a decision j is chosen from a (finite) set D of decisions. This procedure, or a decision rule, is represented by a POVM $\{M_j, j \in D\}$ on $\mathcal{H}$. The performance of a decision rule is assessed by a payoff function, which in our case is a map $g : \Theta \times D \to \mathbb{R}^+$, representing the payoff obtained if $j \in D$ is chosen while the true state is ρ_θ. The average payoff of the decision rule M at $\theta \in \Theta$ is computed as

$$P_{\mathcal{T}}(\theta, M, g) = \sum_{j \in D} g_{\theta, j} \mathrm{Tr}\, \rho_\theta M_j.$$

The next theorem is the celebrated Le Cam's randomization criterion for classical statistical experiments. Note that our setting contains only experiments on finite sample spaces, but the theorem holds in a much more general case.

Theorem 3. *[2] Let $\mathcal{T}$ and $\mathcal{S}$ be classical statistical experiments. Then the following are equivalent.*

(i) *Tor any decision space (D, g) and any decision rule M for $\mathcal{S}$, there is some decision rule N for $\mathcal{T}$ such that*

$$\sup_{\theta\in\Theta}\left[P_{\mathcal{S}}(\theta, M, g) - P_{\mathcal{T}}(\theta, N, g) - \epsilon \max_d |g(\theta, d)|\right] \leq 0.$$

(ii) *There is some channel α such that*

$$\sup_\theta \|\sigma_\theta - \alpha(\rho_\theta)\|_1 \leq 2\epsilon$$

Remark 2. One can show that the condition (i) of the above theorem is equivalent to

$$P_{succ}(\{\lambda_j, \sum_{\theta\in\Theta_0} \mu_\theta^j \sigma_\theta\}) \leq P_{succ}(\{\lambda_j, \sum_{\theta\in\Theta_0} \mu_\theta^j \rho_\theta\}) + \epsilon P_{succ}(\mathcal{E})$$

for any ensemble of the form $\mathcal{E} = \{\lambda_j, \sum_{\theta\in\Theta_0} \mu_\theta^j |e_\theta\rangle\langle e_\theta|\}$ and any finite subset $\Theta_0 \subseteq \Theta$.

As it was proved in [22], Theorem 3 does not hold for quantum experiments. The quantum randomization criterion proved in [8] is based on an extension of the classical decision spaces to quantum ones, but an operational interpretation of the quantum decision problems is not clear. The aim of the present section is to apply Theorem 1 to prove a quantum randomization criterion, formulated in terms of optimal guessing probabilities of some ensembles. In view of Remark 2, this gives a quantum extension of Le Cam's theorem. In the case $\epsilon = 0$, a similar result was obtained in [6].

Theorem 4. *Let $\mathcal{S} = (\mathcal{K}, \{\sigma_\theta,\ \theta \in \Theta\})$ and $\mathcal{T} = (\mathcal{H}, \{\rho_\theta,\ \theta \in \Theta\})$ be quantum statistical experiments and let $\epsilon \geq 0$. Then the following are equivalent.*

(i) *There is some $\alpha \in \mathcal{C}(\mathcal{H}, \mathcal{K})$ such that*

$$\sup_{\theta\in\Theta} \|\sigma_\theta - \alpha(\rho_\theta)\|_1 \leq 2\epsilon$$

(ii) *Let $\{\theta_1, \dots, \theta_n\}$ be any finite subset of Θ and let $\mathcal{E} = \{\lambda_i, \tau_i\}_{i=1}^k$ be any ensemble on $\mathbb{C}^n \otimes \mathcal{K}$, consisting of block-diagonal states $\tau_i = \sum_{j=1}^n |e_j^n\rangle\langle e_j^n| \otimes \tau_i^j$, $\tau_i^j \in B(\mathcal{K})^+$, $\sum_j \operatorname{Tr} \tau_i^j = 1$. Then*

$$P_{succ}(\{\lambda_i, \sum_{j=1}^n \sigma_{\theta_j} \otimes \tau_i^j\}) \leq$$
$$\leq P_{succ}(\{\lambda_i, \sum_{j=1}^n \rho_{\theta_j} \otimes \tau_i^j\}) + \epsilon P_{succ}(\mathcal{E})$$

Moreover, in (ii) we may restrict to equiprobable ensembles with $k = \dim(\mathcal{K})^2$.

Proof. Let $\{\theta_1, \dots, \theta_n\} \subseteq \Theta$ and let $\phi_{\mathcal{S}} \in \mathcal{L}(\mathbb{C}^n, \mathcal{K})$ be given by $A \mapsto \sum_{j=1}^n A_{jj}\sigma_{\theta_j}$. It is easy to see that $\phi_{\mathcal{S}}$ is a channel. Moreover, by [21, Lemma 2], we have for any $\alpha \in \mathcal{C}(\mathcal{H}, \mathcal{K})$,

$$\|\phi_{\mathcal{S}} - \alpha \circ \phi_{\mathcal{T}}\|_\diamond = \max_j \|\sigma_{\theta_j} - \alpha(\rho_{\theta_j})\|_1,$$

where $\phi_{\mathcal{T}}$ is defined analogically. By Theorem 1, the restriction of (i) to $\{\theta_1, \dots, \theta_n\}$ is equivalent to

$$P_{succ}((\phi_{\mathcal{S}} \otimes id)(\mathcal{E})) \leq P_{succ}((\phi_{\mathcal{T}} \otimes id)(\mathcal{E})) + \epsilon P_{succ}(\mathcal{E})$$

for any ensemble $\mathcal{E}$ on $\mathbb{C}^n \otimes \mathcal{K}$. It is now clear that (i) implies (ii). Since for any state $\rho \in \mathfrak{S}(\mathbb{C}^n \otimes \mathcal{K})$,

$$(\phi_{\mathcal{S}} \otimes id)(\rho) = \sum_j \sigma_{\theta_j} \otimes \tau^j = (\phi_{\mathcal{S}} \otimes id)(\tau)$$
$$(\phi_{\mathcal{T}} \otimes id)(\rho) = \sum_j \rho_{\theta_j} \otimes \tau^j = (\phi_{\mathcal{T}} \otimes id)(\tau)$$

where $\tau = \sum_j |e_j^n\rangle\langle e_j^n| \otimes \tau^j$ is a block diagonal state, we see that (ii) implies that

$$\inf_{\alpha\in\mathcal{C}(\mathcal{H},\mathcal{K})} \sup_{\theta\in\Theta_0} \|\sigma_\theta - \alpha(\rho_\theta)\|_1 \leq 2\epsilon$$

for any finite subset $\Theta_0 \subseteq \Theta$. Let $\mathcal{P}_\Theta$ denote the set of probability measures over Θ with finite support, then we clearly have

$$\sup_{p\in\mathcal{P}_\Theta} \min_{\alpha\in\mathcal{C}(\mathcal{H},\mathcal{K})} \sum_{\theta\in\Theta} p(\theta)\|\sigma_\theta - \phi(\rho_\theta)\|_1 \leq 2\epsilon.$$

Now we use the minimax theorem once more. For this, note that $\mathcal{P}_\Theta$ is a convex set, $\mathcal{C}(\mathcal{H}, \mathcal{K})$ is compact and convex, the map $(p, \alpha) \mapsto \sum_{\theta\in\Theta} p(\theta)\|\sigma_\theta - \alpha(\rho_\theta)\|_1$ is linear in p and continuous and convex in α. The minimax theorem can be applied and we obtain

$$\sup_{p\in\mathcal{P}_\Theta} \min_{\alpha\in\mathcal{C}(\mathcal{H},\mathcal{K})} \sum_{\theta\in\Theta} p(\theta)\|\sigma_\theta - \alpha(\rho_\theta)\|_1$$
$$= \min_{\alpha\in\mathcal{C}(\mathcal{H},\mathcal{K})} \sup_{p\in\mathcal{P}_\Theta} \sum_{\theta\in\Theta} p(\theta)\|\sigma_\theta - \alpha(\rho_\theta)\|_1$$
$$= \min_{\alpha\in\mathcal{C}(\mathcal{H},\mathcal{K})} \sup_{\theta\in\Theta} \|\sigma_\theta - \alpha(\rho_\theta)\|_1.$$

Hence (ii) implies (i).

□

We will say that an experiment $\mathcal{S}_0 = (\mathcal{K}, \{\tau_\theta, \theta \in \Theta_0\})$ is complete if the set $\{\tau_\theta, \theta \in \Theta_0\}$ spans $B(\mathcal{H})$. If Θ_0 is a finite set, then $\phi_{\mathcal{S}_0}$ is a surjective channel in $\mathcal{C}(\mathbb{C}^{|\Theta_0|}, \mathcal{K})$. By an application of Theorem 2 we obtain

Corollary 1. *Let $\mathcal{S} = (\mathcal{K}, \{\sigma_\theta,\ \theta \in \Theta\})$ and $\mathcal{T} = (\mathcal{H}, \{\rho_\theta,\ \theta \in \Theta\})$ be quantum statistical experiments. Let $\mathcal{S}_0 = (\mathcal{K}, \{\tau_1, \dots, \tau_N\})$ be a complete experiment. Then $\sigma_\theta = \alpha(\rho_\theta)$ for some $\alpha(\mathcal{H}, \mathcal{K})$ if and only if for any $\{\theta_1, \dots, \theta_n\} \subseteq \Theta$ and any collection $\{\Lambda_{j,l}^i\}$, $i = 1, \dots, k$, $j = 1, \dots, N$, $l = 1, \dots, n$ of nonnegative numbers such that $\sum_{j,l} \Lambda_{j,l}^i = 1$ for all i, we have*

$$P_{succ}(\{1/k, \sum_{j,l} \Lambda_{j,l}^i \sigma_{\theta_l} \otimes \tau_j\}) \leq P_{succ}(\{1/k, \sum_{j,l} \Lambda_{j,l}^i \rho_{\theta_l} \otimes \tau_j\}).$$

Corollary 2. *Let $\Phi \in \mathcal{C}(\mathcal{H}, \mathcal{K})$, $\Psi \in \mathcal{C}(\mathcal{H}, \mathcal{K}')$ and let $\mathcal{T}_0 = (\mathcal{H}, \{\tau_1^{\mathcal{H}}, \dots, \tau_M^{\mathcal{H}}\})$, $\mathcal{S}_0 = (\mathcal{K}, \{\tau_1^{\mathcal{K}}, \dots, \tau_N^{\mathcal{K}}\})$ be complete experiments. Then $\delta(\Phi, \Psi) = 0$ if and only if*

$$P_{succ}((\Phi \otimes id_{\mathcal{K}_0})(\mathcal{E})) \leq P_{succ}((\Psi \otimes id_{\mathcal{K}_0})(\mathcal{E}))$$

holds for all ensembles of states of the form

$$\mathcal{E} = \{\lambda_i, \sum_{j,l} \Lambda_{j,l}^i \tau_l^{\mathcal{H}} \otimes \tau_j^{\mathcal{K}}\}.$$

VI. Concluding remarks

We proved a version of the randomization criterion for quantum channels and applied it to obtain a randomization criterion for quantum statistical experiments. The deficiency $\delta(\Phi, \Psi)$ in some special cases already appeared in quantum information theory and our results can be further used to obtain an operational definition e.g. for the approximately (anti)degradable channels [18], similarly as it was done for antidegradable channels in [15]. Another possible application is to ϵ-private and ϵ-correctable channels [23].

We used some properties of the diamond norm and its dual that can be obtained solely from the order structure given by completely positive maps and the trace preserving condition. This suggests the possibility to apply similar methods to more general situations. For example, one may assume some structure in the channels, obtaining similar results for more specific quantum protocols, such as quantum combs, [24]. It is also possible to define deficiency in terms of pre-processings. In the special case of POVMs regarded as a special kind of channels, this leads to an approximate version of the ordering of POVMs by cleanness, [25]. More generally, the processing can consist of a combination of pre- and post-processing, also allowing some correlations between input and output systems, either classical or quantum. This would be closer to the original definition by Shannon, [9]. It seems that all these situations can be treated within the suggested framework. Another challenging problem is the extension of these results to infinite dimensional Hilbert spaces. Some partial results in this direction were obtained in [26]. Although the methods used in [20] rely on finite dimensions, it seems plausible that the useful properties of the norms can be extended also to this case. All these problems are left for future work.

References

[1] D. Blackwell, "Comparison of experiments," *Proc. 2nd Berkeley Symp. on Math. Stat. and Probab.*, pp. 93–102, 1951.

[2] L. L. Cam, "Sufficiency and approximate sufficiency," *Ann. Math. Statist.*, vol. 34, pp. 1419–1455, 1964.

[3] H. Strasser, *Mathematical Theory of Statistics*. Berlin, New York: de Gruyter, 1985.

[4] E. Torgersen, *Comparison of Statistical Experiments*. Cambridge University Press, 1991.

[5] E. Shmaya, "Comparison of information structures and completely positive maps," *J. Phys. A-Math. Gen.*, vol. 38, pp. 9717–9727, 2005.

[6] F. Buscemi, "Comparison of Quantum Statistical Models: Equivalent Conditions for Sufficiency," *Commun. Math. Phys.*, vol. 310, pp. 625–647, 2012.

[7] A. Jenčová, "Comparison of quantum binary experiments," *Rep. Math. Phys.*, vol. 70, pp. 237–249, 2012.

[8] K. Matsumoto, "A quantum version of randomization condition," arXiv:1012.2650, 2010.

[9] C. E. Shannon, "A note on a partial ordering for communication channels," *Inform. Control*, vol. 1, pp. 390–397, 1958.

[10] J. Körner and K. Marton, "Comparison of two noisy channels," in *Colloquia Mathematica Societatis Janos Bolyai, Topics in Information Theory*. North Holland, 1977, pp. 411–424.

[11] I. Cziszár and J. Körner, *Information Theory. Coding Theorems for Discrete Memoryless Channels*. New York - San Francisco - London: Academic Press, 1981.

[12] M. Raginsky, "Shannon meets Blackwell and Le Cam: Channels, codes, and statistical experiments," *IEEE ISIT Proceedings*, pp. 1220–1224, 2011.

[13] F. Buscemi, "Degradable channels, less noisy channels, and quantum statistical morphisms: an equivalence relation," arXiv:1511.08893, 2015.

[14] A. Chefles, "The quantum Blackwell theorem and minimum error state discrimination," arXiv:0907.0866, 2009.

[15] F. Buscemi, N. Datta, and S. Strelchuk, "Game-theoretic characterization of antidegradable channels," *J. Math. Phys.*, vol. 55, p. 092202, 2014.

[16] F. Buscemi, "Fully quantum second-lawlike statements from the theory of statistical comparisons," arXiv:1505.00535, 2014.

[17] F. Buscemi and N. Datta, "Equivalence between divisibility and monotonic decrease of information in classical and quantum stochastic processes," *Phys. Rev. A*, vol. 93, p. 012101, 2016.

[18] D. Sutter, V. B. Scholz, and R. Renner, "Approximate degradable quantum channels," *IEEE ISIT Proceedings*, pp. 2767–2771, 2015.

[19] A. Kitaev, "Quantum computations: Algorithms and error correction," *Russian Mathematical Surveys*, vol. 52, pp. 1191–1249, 1997.

[20] A. Jenčová, "Base norms and discrimination of generalized quantum channels," *J. Math. Phys*, vol. 55, p. 022201, 2014.

[21] ——, "Comparison of quantum channels and statistical experiments," arXiv:1512.07016, 2015.

[22] K. Matsumoto, "An example of a quantum statistical model which cannot be mapped to a less informative one by any trace preserving positive map," arXiv:1409.5658, 2014.

[23] D. Kretschmann, D. W. Kribs, and R. W. Spekkens, "Complementarity of private and correctable subsystems in quantum cryptography and error correction," *Phys. Rev. A*, vol. 78, p. 032330, 2008.

[24] G. Chiribella, G. M. D'Ariano, and P. Perinotti, "Theoretical framework for quantum networks," *Phys. Rev. A*, vol. 80, 2009.

[25] F. Buscemi, M. Keyl, G. M. D'Ariano, P. Perinotti, and R. F. Werner, "Clean positive operator valued measures," *J. Math. Phys.*, vol. 46, 2005.

[26] K. Kaniowski, K. Lubnauer, and A. Luczak, "Quantum Blackwell-Sherman-Stein Theorem and Related Results," *Open Systems & Information Dynamics*, vol. 20, p. 1350017, 2013.