A quantum version of randomization criteria

Keiji Matsumoto

November 22, 2011

Abstract

In classical statistical decision theory, comparison of experiments plays very important role. Especially, so-called randomization criteria is most important. In this paper, we establish two kinds of quantum analogue of these concepts, and apply to some examples.

1 Introduction

In classical statistical decision theory, comparison of experiments plays very important role. Especially, so-called randomization criteria is most important. In this paper, we establish two kinds of quantum analogue these concepts, and apply to some examples.

2 Review of classical theory

2.1 Decision theory, framework

A statistical experiment $\mathcal{E} = (\mathcal{X}, \mathcal{X}, \{P_\theta; \theta \in \Theta\})$ consists of four parts. First, the data space $\mathcal{X}$, or the totality of all the possible data $x$. Second, a collection $\mathcal{X}$ of subsets of $\mathcal{X}$, indicating minimal unit of the events which the statistician is concerned with.

The third element of an experiment is a parameter set $\Theta$, which indexes all the possible explaining theory for the occurring data. To each $\theta$ corresponds a probability distribution $P_\theta$ of data, which is the fourth element of an experiment. $P_\theta$ has to be $\mathcal{X}$-measurable.

A statistician makes decision based on the data $x \in \mathcal{X}$. The totality of possible decisions made by the statistician is the decision space $\mathcal{D}$, which is a topological space and is equipped with Baire $\sigma$-field $\mathcal{D}$. For example, if the statistician is estimating the value of the parameter $\theta \in \times$ behind the data from the data $x \in \mathcal{X}$, we define $(\mathcal{D}, \mathcal{D}) = (\mathbb{R}\times, \mathcal{B}(\mathbb{R}^l))$. If the statistician is trying to distinguish whether $\theta \in \Theta_0$ or $\theta \in \Theta_1$, $(\mathcal{D}, \mathcal{D}) = (\{0,1\}, 2^{\{0,1\}})$.

The performance is measured by a loss function $l_\theta(t)$, or a function which depends not only on the decision $t \in \mathcal{D}$ but also on the true value of the parameter $\theta \in \Theta$. It is assumed that the function $t \rightarrow l_\theta(t)$ is lower semicontinuous.
and non-negative. For example, in case of estimation of $\theta$, the loss function may be

$$l_\theta (t) = \begin{cases} 
0, & \|t - \theta\| \leq c \\
1, & \text{otherwise}
\end{cases}.$$  

In case of testing $\theta \in \Theta_0$ against $\theta \in \Theta_1$, $l_\theta (t)$ may be chosen so that $l_\theta (0) = 1 - l_\theta (1)$ and

$$l_\theta (0) = \begin{cases} 
0, & (\theta \in \Theta_0) \\
1, & (\theta \in \Theta_1)
\end{cases}.$$  

In general the statistician’s strategy (decision) is described by a bilinear map $D$

$$D : l_\theta \times P_\theta \rightarrow D(l_\theta, P_\theta) \in \mathbb{R},$$

which satisfies

$$|D(f, P)| \leq \|f\| \cdot \|P\|_1,$$

$$D(f, P) \geq 0, \quad \text{if } f \geq 0, P \geq 0,$$

$$D(1, P) = P(D).$$

Here, $P$ is a member of $L$-space of $\mathcal{E}$, or a bounded signed measure such that $P \perp \nu$ for all $\nu$ with $P_\theta \perp \nu, \forall \theta \in \Theta$. The meaning of $D(l_\theta, P_\theta)$ is average of $l_\theta$ when the statistician takes the decision corresponding to $D$, and in many cases,

$$D(l_\theta, P_\theta) = \int_X \int_\mathcal{D} l_\theta (t) R(dt, x) P_\theta (dx),$$

where $R(dt, x)$ is a Markov kernel. $D$ is said to be $k$-decision when $|D| = k$.

Note here that $\Theta$ can be any set, and the function $\theta \rightarrow P_\theta (B)$ can be any function.

2.2 Deficiency

Let $e : \times \rightarrow \mathbb{R}_+$ be a function with $0 \leq e_\theta \leq 1$, $\|l_\theta\| := \sup_{t \in \mathcal{D}} |l_\theta (t)|$, and $\|l\| := \sup_{\theta \in \Theta} \|l_\theta\|$. An experiment $\mathcal{E} = (\mathcal{X}, \mathcal{X}, \{P_\theta; \theta \in \Theta\})$ is said to be $e$-deficient relative to another experiment $\mathcal{F} = (\mathcal{Y}, \mathcal{Y}, \{Q_\theta; \theta \in \Theta\})$ (denoted by $\mathcal{E} \geq_e \mathcal{F}$), if and only if, for any loss function $l$ with $\|l\| \leq 1$, for any finite subset $\Theta_0$ of $\Theta$, and any decision $D$ on the experiment $\mathcal{F}$, there is a decision $D'$ on the experiment $\mathcal{E}$ such that

$$D'(l_\theta, P_\theta) \leq D(l_\theta, Q_\theta) + e_\theta, \forall \theta \in \Theta_0.$$  

(1)

Deficiency $\delta (\mathcal{E}, \mathcal{F})$ is defined by

$$\delta (\mathcal{E}, \mathcal{F}) := \inf_e \left\{ \sup_{\theta} e_\theta : \mathcal{E} \geq_e \mathcal{F} \right\},$$

$\mathcal{E} \geq_0 \mathcal{F}$ is denoted by $\mathcal{E} \geq \mathcal{F}$, and when this holds, $\mathcal{E}$ is said to be more informative than $\mathcal{F}$.
An experiment $\mathcal{E}$ is said to be $e$-deficient for $k$-decision problems relative to another experiment $\mathcal{F}$, if and only if, for any loss function $l$, any finite subset $\Theta_0$ of $\Theta$, and any $k$-decision $D$ on the experiment $\mathcal{F}$, there is a $k$-decision $D'$ on the experiment $\mathcal{E}$ such that (11) holds for any $\theta \in \Theta_0$ (denoted by $\mathcal{E} \geq_{e,k} \mathcal{F}$).

Also, we define deficiency $\delta_k (\mathcal{E}, \mathcal{F})$ for $k$-decision problems by restricting $D'$ and $D$ to the $k$-decisions on $\mathcal{E}$ and $\mathcal{F}$, respectively. $\mathcal{E} \geq 0, k \mathcal{F}$ is denoted by $\mathcal{E} \geq_k \mathcal{F}$, and when this holds, $\mathcal{E}$ is more informative than $\mathcal{F}$ for $k$-decision problems. For notational convenience, we define $\delta_\infty := \delta$, and $e$-deficiency with respect to $\infty$-decision problems as $e$-deficiency.

Finally, we define

$$\Delta (\mathcal{E}, \mathcal{F}) := \max \{ \delta (\mathcal{E}, \mathcal{F}), \delta (\mathcal{F}, \mathcal{E}) \},$$
$$\Delta_k (\mathcal{E}, \mathcal{F}) := \max \{ \delta_k (\mathcal{E}, \mathcal{F}), \delta_k (\mathcal{F}, \mathcal{E}) \},$$
$$\Delta_\infty := \Delta.$$

When $\Delta (\mathcal{E}, \mathcal{F}) = 0$ (, $\Delta_k (\mathcal{E}, \mathcal{F}) = 0$, resp.) we say $\mathcal{E}$ and $\mathcal{F}$ are equivalent (, equivalent for $k$-decision problems, resp.), and represent the situation by the symbol $\mathcal{E} \sim \mathcal{F}$ (,$\mathcal{E} \sim_k \mathcal{F}$, resp.).

One can prove the following necessary and sufficient condition for $\mathcal{E} \geq_{e,k} \mathcal{F}$

(i) For any loss function $l$ with $0 \leq l_\theta (t) \leq 1$, and any $k$-decision $D$ on the experiment $\mathcal{F}$, there is a $k$-decision $D'$ on the experiment $\mathcal{E}$ such that, for any probability distribution $\pi$ with finite support

$$\int_{\Theta} D' (l_\theta, P_\theta) d\pi \leq \int_{\Theta} \{ D (l_\theta, Q_\theta) + e_\theta \} d\pi.$$

(ii) For any loss function $l$, and any $k$-decision $D$ on the experiment $\mathcal{F}$, there is a $k$-decision $D'$ on $\mathcal{E}$ such that

$$\| D' (l_\theta, P_\theta) - D (l_\theta, Q_\theta) \| \leq e_\theta, \forall \theta \in \Theta.$$

Also, $\mathcal{E} \geq_e \mathcal{F}$ is equivalent to :

(iii) (Randomization criterion) There is an affine positive map $\Lambda$ such that

$$\| \Lambda (P_\theta) - Q_\theta \|_1 \leq e_\theta, \forall \theta \in \Theta.$$

3 Notation and mathematical background

$\mathcal{H}$ and $\mathcal{K}$ are separable Hilbert spaces. $\mathcal{B} (\mathcal{H})$, $\mathcal{S}_1 (\mathcal{H})$, and $\mathcal{LC} (\mathcal{H})$ is the space of bounded operators, trace class operators, and compact operators over Hilbert space $\mathcal{H}$, respectively. It is known that $\mathcal{LC} (\mathcal{H})^*$ and $\mathcal{S}_1 (\mathcal{H})^*$ is isometrically isomorphic to $\mathcal{S}_1 (\mathcal{H})$ and $\mathcal{B} (\mathcal{H})$, respectively. Here, continuity is defined in terms of the operator norm $\| \cdot \|$. We furnish $\mathcal{S}_1 (\mathcal{H}) \simeq \mathcal{LC} (\mathcal{H})^*$ with
weak* topology. $B(\mathcal{H})$, $S_1(\mathcal{H})$, and $\mathcal{L}\mathcal{C}(\mathcal{H})$ is also furnished with topology induced by the norm $||\cdot||$, $||\cdot||_1$ and $||\cdot||$, respectively. We can also introduce to the spaces $S_1(\mathcal{H}) \simeq \mathcal{L}\mathcal{C}(\mathcal{H})^*$ and $B(\mathcal{H}) \simeq S_1(\mathcal{H})^*$ the operator norm as a linear functional, but they coincide with $||\cdot||_1$ and $||\cdot||$, respectively. Define ball $B(\mathcal{H}) := \{L; ||L|| \leq 1\}$, etc. Then, by Alaoglu’s theorem, $\text{ball } S_1(\mathcal{H})$ is weak* compact.

Given a set $Y$ and topological spaces $(X_y, \tau_y)$, and we furnish $\times_{y \in Y} X_y$ with the product topology, whose local base is a family of sets in the form of

$$\{x; x_y \in U_y, y \in F\},$$

where $U_y$ is an open set in $X_y$, and $F$ is a finite set in $Y$ \cite{10}. Note that $\times_{y \in Y} U_y$ is not necessarily open. Note also that, under this topology, the projection $P_y$ from $\times_{y \in Y} X_y$ to $X_y$, $P_y(x) = x_y$, is continuous. Also, the product topology is the weakest topology which makes $P_y$ continuous. Tychonoff’s theorem states that $\times_{y \in Y} D_y$ is compact in product topology if $D_y$ is compact.

A map $\Lambda : S_1(\mathcal{H}) \to S_1(\mathcal{K})$ is said to be completely positive if and only if $\Lambda \otimes I_n$ is positive for any $n$, where $I_n$ is the identity map form $\mathbb{C}^n$ to $\mathbb{C}^n$. This is equivalent to

$$\sum_{i,j=1}^{n} \langle \psi_i | \Lambda (|\varphi_i\rangle \langle \varphi_j|) |\psi_j\rangle \geq 0, \quad (2)$$

for any $\{|\varphi_i\rangle\}_{i=1}^{n}$ $(B_i \in \mathcal{H})$, any $\{|\psi_i\rangle\}_{i=1}^{n}$ $(|\psi_i\rangle \in \mathcal{K})$, and for any $n$. $M$ is said to be trace preserving if

$$\text{tr} \Lambda (X) = \text{tr} X, \quad \forall X \in S_1(\mathcal{H}). \quad (3)$$

We put

$$Ch(\mathcal{H}, \mathcal{K}) := \{\Lambda; \Lambda \text{ linear, } (2), (3)\}.$$

**Lemma 1** The set $Ch(\mathcal{H}, \mathcal{K})$, viewed as a subset of $(S_1(\mathcal{K}))^{\text{ball }} S_1(\mathcal{H})$, is compact and convex, in the product topology. Also, $(S_1(\mathcal{K}))^{\text{ball }} S_1(\mathcal{H})$ is locally convex in the product topology. In addition, if $X \in S_1(\mathcal{H})$ and $B \in \mathcal{L}\mathcal{C}(\mathcal{K})$, the linear functional

$$\Lambda \in Ch(\mathcal{H}, \mathcal{K}) \to \text{tr} \Lambda (X) B \in \mathbb{C}$$

is continuous in the product topology.

**Proof.** Since $||\Lambda|| = 1$, $Ch(\mathcal{H}, \mathcal{K}) \subset (\text{ball } S_1(\mathcal{K}))^{\text{ball }} S_1(\mathcal{H})$. Due to Alaoglu’s theorem, ball $S_1(\mathcal{K})$ is compact in the weak* topology. Therefore, by Tychonoff’s theorem, $(\text{ball } S_1(\mathcal{K}))^{\text{ball }} S_1(\mathcal{H})$ is compact in the product topology. Hence, it suffices to show $Ch(\mathcal{H}, \mathcal{K})$ is closed in the product topology. Let $\{\Lambda_\alpha\}$ a net in $Ch(\mathcal{H}, \mathcal{K})$, with $\Lambda_\alpha \to \Lambda$ in the product topology. Then,

$$\Lambda_\alpha (a_1 X_1 + a_2 X_2) - \{a_1 \Lambda_\alpha (X_1) + a_2 \Lambda_\alpha (X_2)\}$$

$$\to \Lambda (a_1 X_1 + a_2 X_2) - \{a_1 \Lambda (X_1) + a_2 \Lambda (X_2)\} = 0,$$
\[
\sum_{i,j=1}^{n} \langle \psi_i | \Lambda_\alpha (|\varphi_i \rangle \langle \varphi_j|) |\psi_j \rangle \rightarrow \sum_{i,j=1}^{n} \langle \psi_i | \Lambda (|\varphi_i \rangle \langle \varphi_j|) |\psi_j \rangle \geq 0,
\]
and
\[
\text{tr} \Lambda_\alpha (X) = \text{tr} X \rightarrow \text{tr} \Lambda (X).
\]
Therefore, \( \Lambda \) is also in \( Ch (\mathcal{H}, \mathcal{K}) \). Thus, \( Ch (\mathcal{H}, \mathcal{K}) \) is closed and compact.

To prove the second assertion, recall that \( \mathcal{S}_1 (\mathcal{K}) \) is locally convex in weak* topology, which has a convex local base. Therefore, in product topology, \( (\mathcal{S}_1 (\mathcal{K}))^{\text{ball} \mathcal{S}_1 (\mathcal{H})} \) is locally convex, with a local base being family of sets of the form
\[
\{ x; x_y \in U_y, y \in F \},
\]
where \( U_y \) is a member of a local base of weak* topology.

To prove the third assertion, let \( \{ \Lambda_\alpha \} \) a net in \( Ch (\mathcal{H}, \mathcal{K}) \), with \( \Lambda_\alpha \rightarrow \Lambda \) in the product topology. Then, \( \Lambda_\alpha (X) \rightarrow \Lambda (X) \) in weak* topology, for any \( X \in \text{ball} \mathcal{S}_1 (\mathcal{H}) \). Therefore, \( \text{tr} \Lambda_\alpha (X) A \rightarrow \text{tr} \Lambda (X) A \), for any \( A \in \mathcal{L} C (\mathcal{H}) \). Hence, we have the assertion. ■

POVM over a measurable space \( (\mathcal{D}, \mathcal{D}) \) in the Hilbert space \( \mathcal{H} \) is denoted by \( Mes (\mathcal{D}, \mathcal{D}; \mathcal{H}) \). The space of signed measures and signed finitely additive measures over \( (\mathcal{D}, \mathcal{D}) \) is denoted by \( ca (\mathcal{D}, \mathcal{D}) \) and \( ba (\mathcal{D}, \mathcal{D}) \), respectively. They are metrized by the total variation norm. The space of bounded measurable function is denoted by \( L^\infty (\mathcal{D}, \mathcal{D}) \). Then, \( ba (\mathcal{D}, \mathcal{D}) \simeq L^\infty (\mathcal{D}, \mathcal{D})^* \). The map
\[
f_M : X \in \mathcal{S}_1 (\mathcal{H}) \rightarrow \text{tr} XM (\cdot) \in ca (\mathcal{D}, \mathcal{D})
\]
is linear, positive, and \( f_M (X) (\mathcal{D}) = \text{tr} X \). Conversely, any linear, bounded, and positive map from \( \mathcal{S}_1 (\mathcal{H}) \) to \( ca (\mathcal{D}, \mathcal{D}) \) with \( f (X) (\mathcal{D}) = \text{tr} X \) is in this form.

The space of elements \( f \) of \( \mathcal{B} (\mathcal{S}_1 (\mathcal{H}), ba (\mathcal{D}, \mathcal{D})) \) with \( f \geq 0 \) and \( f (\rho) (\mathcal{D}) = \text{tr} \rho (\rho \geq 0) \) is denoted by \( \overline{Mes} (\mathcal{D}, \mathcal{D}; \mathcal{H}) \).

**Lemma 2** Suppose \( \mathcal{D} \) is locally compact. The set \( \overline{Mes} (\mathcal{D}, \mathcal{D}; \mathcal{H}) \), viewed as a subset of \( (\text{ball} \ ba (\mathcal{D}, \mathcal{D}))^{\text{ball} \mathcal{S}_1 (\mathcal{H})} \), is compact and convex, in the product topology. Also, \( (\text{ball} \ ba (\mathcal{D}, \mathcal{D}))^{\text{ball} \mathcal{S}_1 (\mathcal{H})} \) is locally convex in the product topology. In addition, if \( X \in \mathcal{S}_1 (\mathcal{H}) \) and \( l \in L^\infty (\mathcal{D}, \mathcal{D}) \), the linear functional
\[
f \in \overline{Mes} (\mathcal{D}, \mathcal{D}; \mathcal{H}) \rightarrow \int_D l (t) f (X) (d t) \in \mathbb{R}
\]
is continuous in the product topology.

**Proof.** Observe that \( \overline{Mes} (\mathcal{D}, \mathcal{D}; \mathcal{H}) \) is a subset of \( (\text{ball} \ ba (\mathcal{D}, \mathcal{D}))^{\text{ball} \mathcal{S}_1 (\mathcal{H})} \), which is compact in the product topology. Hence, to show the first assertion, it suffices to show \( \overline{Mes} (\mathcal{D}, \mathcal{D}; \mathcal{H}) \) is closed. Let \( \{ f_\alpha \} \) be a net in \( \overline{Mes} (\mathcal{D}, \mathcal{D}; \mathcal{H}) \) with \( f_\alpha \rightarrow f \). Then,
\[
f_\alpha \left( a_1 X_1 + a_2 X_2 \right) - \{ a_1 f_\alpha (X_1) + a_2 f_\alpha (X_2) \}
- \rightarrow f \left( a_1 X_1 + a_2 X_2 \right) - \{ a_1 f (X_1) + a_2 f (X_2) \} = 0,
\]

5
∀ρ ≥ 0, ρ ∈ S₁(H), ∀B ∈ D, f_α(ρ)(B) → f(ρ)(B) ≥ 0,

and

∀ρ ≥ 0, ρ ∈ S₁(H), f_α(ρ)(D) → f(ρ)(D) = tr ρ.

Thus, f ∈ \(\overline{Mes}(D, D; H)\), and \(\overline{Mes}(D, D; H)\) is closed, and therefore is compact.

To prove the second assertion, recall that \(ba(D, D)\) is locally convex in weak* topology, which has a convex local base. Therefore, in product topology, \((ba(D, D))^{ballS₁(H)}\) is locally convex, with a local base being family of sets of the form

\[
\{ x; x_y \in U_y, y \in F \},
\]

where \(U_y\) is a member of a local base of weak* topology.

To prove the third assertion, let \(\{f_α\}\) a net in \(\overline{Mes}(D, D; H)\), with \(f_α → Λ\) in the product topology. Then, \(f_α(X) → f(X)\) in weak* topology, for any \(X ∈ ballS₁(H)\). Therefore, \(\int_{\Omega} l(t) f_α(X)(d t) → \int_{\Omega} l(t) f(X)(d t)\), for any \(l ∈ L^∞(D, D)\). Hence, we have the assertion.

### 4 Quantum Theory: framework

A quantum experiment \(E = (H, \{ρ_θ; θ ∈ Ω\})\) consists of Hilbet space \(H\) and the family \(\{ρ_θ; θ ∈ Ω\}\) of states over \(H\). (More generally, though we do not use such setting in this paper, \(E = (H, \{ω_θ; θ ∈ Ω\})\), where \(H\) is a \(C^*\)-algebra and \(ω_θ\) is a state over \(H\).)

From here, \(Ω\) is an arbitrary set.

A quantum decision space \(H_D\) and quantum decision rule \(D\) is a member of \(Ch(H, H_D)\), respectively. Loss function \(L\) is a non-negative function from \(Ω × S₁(H_D)\) to \(R\), such that

\[L_θ(X) ≥ 0 \ (X ≥ 0)\]

and

\[\sup \left\{ \frac{|L_θ(X) - L_θ(Y)|}{\|X - Y\|_1}; X ≥ 0, Y ≥ 0, tr X ≤ 1, tr Y ≤ 1 \right\} ≤ 1. \quad (4)\]

For example:

1. \(L_θ(X_θ) := \|X_θ - ρ_{0,θ}\|_1\), with \(θ → ρ_{0,θ} ∈ S₁(H_D)\) being continuous in \(\|\cdot\|_1\).

2. Suppose \(H_D < ∞\). Then, for a continuous (in trace norm) function \(θ → ρ_{0,θ}\) and \(L_θ(X) := 1 - tr \sqrt{\frac{1}{2}X ρ_{0,θ}^{1/2}}\) satisfies (4), due to

\[1 - tr \sqrt{\frac{1}{2}X ρ_{0,θ}^{1/2}} ≤ \|ρ_{0,θ} - X\|_1.\]
3. \( L_{\theta} (X) = \text{tr} L_{\theta} X \). Here, \( L_{\theta} \geq 0, \| L_{\theta} \| \leq 1, L_{\theta} \in \mathcal{L}C (\mathcal{H}_D) \). \( L_{\theta} \) is called loss operator. and due to [4].

A classical decision \( M \) is a POVM in \( \mathcal{H} \) over a classical decision space \((\mathcal{D}, \mathfrak{D})\). Also, one may consider probabilistic quantum decision. If \(|\mathcal{D}| < \infty\), we can consider such decision as a CPTP map \( D : \mathcal{S}_1 (\mathcal{H}) \to \mathcal{D} \times \mathcal{S}_1 (\mathcal{H}_D) \):

\[
D : X \rightarrow \bigoplus_{t=1}^{\mathcal{D}} p_t X_t,
\]

and a proper loss function would be

\[
L_{\theta} (D (X_\theta)) = \frac{1}{2} \sum_{t \in \mathcal{D}} p_{\theta,t} \| X_{\theta,t} - \rho_{0,\theta,t} \|_1.
\]

It is easy to see, by triangle inequality,

\[
|L_{\theta} (D (X_\theta)) - L_{\theta} (D' (X_\theta))| \leq \frac{1}{2} \sum_{t \in \mathcal{D}} ||p_{\theta,t} X_{\theta,t} - p'_{\theta,t} X'_{\theta,t} - (p_{\theta,t} - p'_{\theta,t}) \rho_{0,\theta,t}||_1
\]

\[
\leq \frac{1}{2} \sum_{t \in \mathcal{D}} \left( ||p_{\theta,t} X_{\theta,t} - p'_{\theta,t} X'_{\theta,t}||_1 + |p_{\theta,t} - p'_{\theta,t}| \right)
\]

\[
\leq ||D (X_\theta) - D' (X_\theta)||_1.
\]

Hence, this case is also satisfies [4].

A quantum experiment \( \mathcal{E} \) is said to be q-\( e \)-deficient relative to \( \mathcal{F} = (\mathcal{K}, \{ \sigma_\theta ; \theta \in \Theta \}) \) for \( k \)-decision problems (denoted by \( \mathcal{E} \geq^q_{e,k} \mathcal{F} \)), if and only if, for any if and only if, for for \( \mathcal{H}_D \) with \( \dim \mathcal{H}_D = k \), any loss function \( L \) with [4], any decision \( D \) on the experiment \( \mathcal{F} \),

\[
\inf_{D' \in \mathcal{C}(\mathcal{H}, \mathcal{H}_D) \, \theta \in \Theta} \sup \{ L_{\theta} (D' (\rho_0)) - L_{\theta} (D (\sigma_\theta)) - e_\theta \} \leq 0. \tag{5}
\]

When [4] is true for \( k = \infty \), we say \( \mathcal{E} \) is q-\( e \)-deficient relative to \( \mathcal{F} \), and denote this situation by \( \mathcal{E} \geq^q \mathcal{F} \). (Thus, \( \mathcal{E} \geq^q \mathcal{F} \) means \( \mathcal{E} \geq^q_{e,\infty} \mathcal{F} \).

Also, we can consider tasks with classical outputs. Different from purely classical setting, decision space is a measurable space \((\mathcal{D}, \mathfrak{D})\), and loss function \( t \to l_\theta (t) \) is a measurable function taking values in \([0, 1]\).

\( \mathcal{E} \) is said to be c-\( e \)-deficient relative to \( \mathcal{F} \) (denoted by \( \mathcal{E} \geq^c_{e} \mathcal{F} \)), if and only if, for any loss function \( l \) with \( 0 \leq l_\theta (t) \leq 1 \), for any decision \( M \) on the experiment \( \mathcal{F} \),

\[
\inf_{M'} \int_{\mathcal{D}} l_\theta (t) \text{tr} \rho_0 M' \, (dt) \leq \int_{\mathcal{D}} l_\theta (t) \text{tr} \sigma_\theta M \, (dt) + e_\theta, \, \forall \theta \in \Theta. \tag{6}
\]

c-\( e \)-deficiency for \( k \)-decision problems is defined by posing the restriction \(|\mathcal{D}| \leq k\) and it is denoted by \( \mathcal{E} \geq^c_{e,k} \mathcal{F} \).

q- and c-deficiency is defined in parallel with deficiency and denoted by \( \delta^q (\mathcal{E}, \mathcal{F}) \) and \( \delta^c (\mathcal{E}, \mathcal{F}) \), respectively. Their \( k \)-decision versions \( \delta^q_k (\mathcal{E}, \mathcal{F}) \) and \( \delta^c_k (\mathcal{E}, \mathcal{F}) \) are also defined analogously.
5 Some backgrounds from convex analysis

**Theorem 3** (Minimax theorem [11]). Let $T$ be a convex, compact subset of locally convex vector space and let $V$ be a convex subset of a vector space. Assume that $f : T \times V \rightarrow \mathbb{R}$ satisfies the following conditions: (1) $t \rightarrow f(t, v)$ is continuous and concave on $T$ for every $v \in V$. (2) $v \rightarrow f(t, v)$ is convex on $V$ for every $t \in T$. Then

$$\inf_{v \in V} \sup_{t \in T} f(t, v) = \sup_{t \in T} \inf_{v \in V} f(t, v) = \inf_{v \in V} f(t_0, v), \quad \exists t_0 \in T.$$

**Corollary 4** (Minimax theorem) Let $T$ be a convex subspace of a vector space, and let $V$ be a convex, compact subset of a locally convex vector space. Assume that $f : T \times V \rightarrow \mathbb{R}$ satisfies the following conditions: (1) $t \rightarrow f(t, v)$ is concave on $T$ for every $v \in V$. (2) $v \rightarrow f(t, v)$ is continuous convex on $V$ for every $t \in T$. Then

$$\inf_{v \in V} \sup_{t \in T} f(t, v) = \sup_{t \in T} \inf_{v \in V} f(t, v) = \sup_{t \in T} f(t, v_0), \quad \exists v_0 \in V.$$

**Proof.** Since

$$\inf_{v \in V} \sup_{t \in T} f(t, v) = -\sup_{v \in V} \inf_{t \in T} (-f(t, v)),$$

letting $f'(t, v) := -f(t, v)$ and and applying Theorem 3, we have the assertion. $\blacksquare$

Below, $P_{\Theta}$ is the set of probability measures over $\Theta$ whose support is a finite set. Furnish $\times$ with discrete topology and Corollary 4.5 of [11], then, we obtain:

**Lemma 5** Let $M_1, M_2$ be a subset of $\mathbb{R}^\Theta$. Suppose that $M_2$ is subconvex and $\alpha (M_2) := \{ f ; f \geq g, \forall g \in M_2 \}$ is closed in the product topology. Then, the following (i) and (ii) are equivalent:

**Theorem 6** (i) For every $f \in M_1$, there is some $g \in M_2$ with $g \leq f$.

(ii) For any $\pi \in P_{\Theta}$,

$$\inf_{g \in M_2} \int_\Theta g \, d\pi \leq \inf_{f \in M_1} \int_\Theta f \, d\pi.$$

6 Quantum randomization criterion

**Theorem 7** $E$ is q-e-deficient relative to $F$ for k-decision problems if and only if each of the following three holds (below, $\dim \mathcal{H}_D = k$):

(i) For any finite subset $\Theta_0 \subset \Theta$, for any family $\{ L_\theta \}_{\theta \in \Theta_0}$ with (4), and for any $D \in Ch(K, \mathcal{H}_D)$ there exists a $D' \in Ch(\mathcal{H}, \mathcal{H}_D)$,

$$\inf_{D' \in Ch(\mathcal{H}, \mathcal{H}_D)} \sup_{\theta \in \Theta_0} \{ L_\theta (D'(\omega_\theta)) - L_\theta (D(\eta_\theta)) - e_\theta \} \leq 0.$$
(ii) For any finite subset $\Theta_0 \subset \Theta$, for any loss operator $\{L_\theta\}$, any decision $D \in Ch(\mathcal{K}, \mathcal{H}_D)$ on the experiment $\mathcal{F}$,
$$\inf_{D' \in Ch(\mathcal{K}, \mathcal{H}_D)} \sup_{\theta \in \Theta_0} \{\operatorname{tr} L_\theta D'(\rho_\theta) - \operatorname{tr} L_\theta D(\sigma_\theta) - e_\theta\} \leq 0.$$ 

(iii) For any loss operator $\{L_\theta\}$, any $k$-decision $D$ on the experiment $\mathcal{F}$, and any $\pi \in \mathcal{P}_\Theta$,
$$\exists D'_\pi \in Ch(\mathcal{H}, \mathcal{H}_D), \quad \int_{\Theta} \{\operatorname{tr} L_\theta D'_\pi(\rho_\theta) - \operatorname{tr} L_\theta D(\sigma_\theta) - e_\theta\} d\pi.$$

(iv) For any $D$ on the experiment $\mathcal{F}$,
$$\exists D'_0 \in Ch(\mathcal{H}, \mathcal{H}_D), \quad \sup_{\theta \in \Theta} \{\|D'_0(\rho_\theta) - D(\sigma_\theta)\|_1 - e_\theta\} \leq 0,$$

Proof. Obviously, $[5] \Rightarrow (i) \Rightarrow (ii) \Rightarrow (iii)$, and $(iv) \Rightarrow [5]$. Hence, we show $(iii) \Rightarrow (iv)$. Since $\pi \in \mathcal{P}_\Theta$, the map
$$D' \rightarrow \int_{\Theta} \{\operatorname{tr} L_\theta D'(\rho_\theta) - \operatorname{tr} L_\theta D(\sigma_\theta) - e_\theta\} d\pi$$
is continuous in the product topology. Since $Ch(\mathcal{H}, \mathcal{H}_D)$ is compact with respect to the product topology due to Lemma[1] Corollary[2] leads to $

\begin{align*}
\sup_{L_\theta : L_\theta \in \mathcal{L}(\mathcal{K}), \|L\| \leq 1} & \inf_{D' \in Ch(\mathcal{H}, \mathcal{H}_D)} \int_{\Theta} \{\operatorname{tr} L_\theta D'(\rho_\theta) - \operatorname{tr} L_\theta D(\sigma_\theta) - e_\theta\} d\pi \\
= & \inf_{D' \in Ch(\mathcal{H}, \mathcal{H}_D)} \sup_{L_\theta : L_\theta \in \mathcal{L}(\mathcal{K}), \|L\| \leq 1} \int_{\Theta} \{\operatorname{tr} L_\theta D'(\rho_\theta) - \operatorname{tr} L_\theta D(\sigma_\theta) - e_\theta\} d\pi \\
= & \inf_{D' \in Ch(\mathcal{H}, \mathcal{H}_D)} \int_{\Theta} \{\|D'(\rho_\theta) - D(\sigma_\theta)\|_1 - e_\theta\} d\pi
\end{align*}

Let $M_2 := \{g; g(\theta) = \|D'(\rho_\theta) - D(\sigma_\theta)\|_1, D' \in Ch(\mathcal{H}, \mathcal{H}_D)\}$. Letting $\pi$ be the one concentrated at $\{\theta\}$, and applying Corollary[4] we have
$$\exists D'_0 \in Ch(\mathcal{H}, \mathcal{H}_D), \quad \inf_{D' \in Ch(\mathcal{H}, \mathcal{H}_D)} \{|\|D'(\rho_\theta) - D(\sigma_\theta)\|_1\| = \|D'_0(\rho_\theta) - D(\sigma_\theta)\|_1\|.$$ 

Therefore,
$$\alpha(M_2) = \{g; g(\theta) \geq \|D'_0(\rho_\theta) - D(\sigma_\theta)\|_1, D' \in Ch(\mathcal{H}, \mathcal{H}_D)\},$$
which is closed under the product topology. Obviously, $\alpha(M_2)$ is convex. Hence, letting $M_1 = \{e\}$ and applying Lemma[5] we obtain
$$\exists D'_0 \in Ch(\mathcal{H}, \mathcal{H}_D), \quad \forall \theta \in \Theta \quad \|D'_0(\rho_\theta) - D(\sigma_\theta)\|_1 \leq e_\theta.$$ 

Hence, we have (iv).

Letting $\mathcal{H}_D = \mathcal{K}$ and $D = I$, we obtain:

**Theorem 8** $\mathcal{E} \supseteq \mathcal{F}$ is equivalent to there is a CPTP map $\Lambda$ with
$$\|\Lambda(\rho_\theta) - \sigma_\theta\|_1 \leq e_\theta, \quad \forall \theta \in \Theta.$$ 

7 Classical decision space

Lemma 9 ([11], Theorem 41.7) There is a positive linear operator \( T : ba(D, \mathcal{D}) \to ca(D, \mathcal{D}) \) such that

(i) \( \|T\| = 1 \),
(ii) \( T(\mu)(D) = \mu(D) \), if \( \mu \geq 0 \)
(iii) \( T|_{ca(D, \mathcal{D})} = \text{id}|_{ca(D, \mathcal{D})} \)

Theorem 10 \( E \geq_c \mathcal{F} \) if and only if one of the following two holds:

(i) For any decision space \((D, \mathcal{D})\), for any measurable loss function \( l \) with \( 0 \leq l(\theta)(t) \leq 1 \), for any decision \( M \) on the experiment \( \mathcal{F} \), and for any \( \pi \in \mathcal{P}_\Theta \), there is some decision \( M' \) on the experiment \( E \) such that

\[
\int_\Theta \int_{t \in D} l(\theta)(t) \text{tr} \rho M'(dt) \, d\pi \leq \int_\Theta \left\{ \int_{t \in D} l(\theta)(t) \text{tr} \sigma M(dt) + e(\theta) \right\} \, d\pi.
\]

(ii) For any decision space \((D, \mathcal{D})\), any decision \( M \) on the experiment \( \mathcal{F} \), there is some decision \( M' \) on the experiment \( E \) such that

\[
\sup_{\theta \in \Theta} \{ \| f_M(\rho) - f_M(\sigma) \|_1 - e(\theta) \} \leq 0.
\]

Proof. \( (i) \Rightarrow (i), (ii) \Rightarrow (i) \) is trivial. Hence, we have to show \( (i) \Rightarrow (ii) \). Suppose \( (i) \) holds true. Then, using the argument parallel to the proof \( (iii) \Rightarrow (iv) \) of Theorem 7, we have, for any \( M \in Mes(D, \mathcal{D}) \),

\[
\exists f_0 \in Mes(D, \mathcal{D}) \quad \sup_{\theta \in \Theta} \{ \| f_0(\rho) - f_M(\sigma) \|_1 - e(\theta) \} \leq 0.
\]

Let \( T \) be as of Lemma 9

\[
\sup_{\theta \in \Theta} \{ \| T \circ f_0(\rho) - f_M(\sigma) \|_1 - e(\theta) \}
= \sup_{\theta \in \Theta} \{ \| T \circ f_0(\rho) - T \circ f_M(\sigma) \|_1 - e(\theta) \}
\leq \sup_{\theta \in \Theta} \{ \| f_0(\rho) - f_M(\sigma) \|_1 - e(\theta) \} \leq 0.
\]

Then, \( f'_0 := T \circ f_0 \), which is a bounded linear map from \( S_1(H) \) to \( ca(D, \mathcal{D}) \), satisfies \( f'_0 \geq 0 \), and

\[
f'_0(\rho)(D) = T(f_0(\rho))(D) = f_0(\rho)(D) = \text{tr} \rho, \quad (\rho \geq 0).
\]

Thus, there is a \( M' \) such that \( f_{M'} = f'_0 \), and \( (ii) \) is proved. ■

Using almost parallel argument, we have:
Theorem 11 \( \mathcal{E} \geq_{\varepsilon, k} \mathcal{F} \) if and only if one of the following two holds:

(i) With \(|\mathcal{D}| = k\), for any measurable loss function \( l \) with \( 0 \leq l_\theta(t) \leq 1 \), for any \( k \)-decision \( M \) on the experiment \( \mathcal{F} \), and for any \( \pi \in \mathcal{P}_\theta \), there is some \( k \)-decision \( M' \) on the experiment \( \mathcal{E} \) such that

\[
\int_{\Theta} \left\{ \sum_{t \in \mathcal{D}} l_\theta(t) \rho \, \text{tr} \, M'(t) \right\} \, d\pi \leq \int_{\Theta} \left\{ \sum_{t \in \mathcal{D}} l_\theta(t) \rho \, \text{tr} \, M(k) + e_\theta \right\} \, d\pi.
\]

(ii) With \(|\mathcal{D}| = k\), any \( k \)-decision \( M \) on the experiment \( \mathcal{F} \), there is some \( k \)-decision \( M' \) on the experiment \( \mathcal{E} \) such that

\[
\sup_{\theta \in \Theta} \left\{ \left\| f_{M'}(\rho_\theta) - f_M(\sigma_\theta) \right\|_1 - e_\theta \right\} \leq 0.
\]

In case \( k = 2 \), c-\( c \)-deficiency has more explicite expression. Since

\[
p(\theta) \sum_{t \in \{0, 1\}} l_\theta(t) \rho \, \text{tr} \, M'(t) = p(\theta) \left\{ (l_\theta(0) - l_\theta(1)) \rho \, \text{tr} \, M'(0) + l_\theta(1) \right\},
\]

we have

\[
\int_{\Theta} \left\{ \sum_{t \in \mathcal{D}} l_\theta(t) \rho \, \text{tr} \, M'(t) - \text{tr} \, M(t) \right\} \, d\mu = \int_{\Theta} p(\theta) \left\{ (l_\theta(0) - l_\theta(1)) \rho \, \text{tr} \, M'(0) - \text{tr} \, M(0) \right\} \, d\mu.
\]

Hence, without loss of generality, we can suppose \( l_\theta(0) - l_\theta(1) = \pm 1 \). Therefore, letting \( \theta \to a_\theta \) be a measurable function with \( |a_\theta| = p(\theta) \),

\[
\inf_{M'} \int_{\Theta} \left\{ \sum_{t \in \mathcal{D}} l_\theta(t) \rho \, \text{tr} \, M'(t) - \text{tr} \, M(t) \right\} \, d\mu = \inf_{M'} \text{tr} \left( \int_{\Theta} a_\theta \rho \, d\mu \right) M'(0) - \text{tr} \left( \int_{\Theta} a_\theta \rho \, d\mu \right) M(0) - \int_{\Theta} |a_\theta| e_\theta \, d\mu \leq 0
\]

Since this holds for any \( M \), we have

\[
- \left\| \int_{\Theta} a_\theta \rho \, d\mu \right\|_1 + \left\| \int_{\Theta} a_\theta \sigma \, d\mu \right\|_1 - \int_{\Theta} |a_\theta| e_\theta \, d\mu \leq 0,
\]

or

\[
\left\| \int_{\Theta} a_\theta \rho \, d\mu \right\|_1 \geq \left\| \int_{\Theta} a_\theta \sigma \, d\mu \right\|_1 - \int_{\Theta} |a_\theta| e_\theta \, d\mu.
\]
where \( \theta \to a_\theta \) is an arbitrary \( L^1(\Theta, \mu) \) with \( \int |a_\theta| \, d\mu = 1 \). Especially when \( \Theta = \{0, 1\} \), this is equivalent to

\[
\|\rho_0 - s\rho_1\|_1 \geq \|\sigma_0 - s\sigma_1\|_1 - e_0 - se_1, \forall s \geq 0.
\] (8)

In case \( \dim \mathcal{H} = \dim \mathcal{K} = 2 \), it is known that

\[
\|\rho_0 - s\rho_1\|_1 \geq \|\sigma_0 - s\sigma_1\|_1 - e_0 - se_1, \forall s \geq 0,
\] (9)

is necessary and sufficient for \( E \geq q_0 F \). In other words, \( E \geq q_0 F \) is equivalent to \( E \geq c_0, 2 F \). However, in case that \( \dim \mathcal{H} = \dim \mathcal{K} = 3 \), (9) fails to be sufficient for \( E \geq q_0 F \).

In case of classical case, more strongly, (8), or \( E \geq c_0, 2 F \), is known to be equivalent to \( E \geq c_0, F \) [12][13]. The following theorem is found independently in [8].

**Theorem 12** Suppose \( \Theta = \{0, 1\} \), and \( [\rho_0, \rho_1] = 0 \). Then, \( E \geq c_0, F \) if and only if (8) holds, or \( E \geq c_0, 2 F \).

**Proof.** Let \( F^M \) be a classical experiment consisted with \( Q^M_\theta \) respectively, where \( Q^M_\theta (dx) = tr \sigma_\theta M (dx) \). Then, by Theorem [10] \( E \geq c_0, F \) if and only if

\[
E \geq c_0 F^M, \forall M.
\]

As noted above, this equivalent to [12]

\[
\|\rho_0 - s\rho_1\|_1 \geq \|Q^M_0 - sQ^M_1\|_1 - e_0 - se_1, \forall M, \forall s \geq 0.
\]

Therefore, since

\[
\max_M \|Q^M_0 - sQ^M_1\|_1 = \|\sigma_0 - s\sigma_1\|_1,
\]

we have the assertion. \( \blacksquare \)

**8 Statistical morphism**

[3] introduced the notion of statistical morphism, which we use here with some non-essential modifications. A map \( \Gamma \) from \( \{\rho_\theta\}_{\theta \in \Theta} \subset \mathcal{S}_1(\mathcal{H}) \) into \( \mathcal{S}_1(\mathcal{K}) \) is said to be \( k \)-statistical morphism if and only if, for any \( \epsilon > 0 \) and for \( k \)-decision \( M \) over \( \mathcal{H} \), there exists a \( k \)-decision \( M' \) over \( \mathcal{K} \) with

\[
\sum_{t \in \mathcal{D}} |tr \Gamma (\rho_\theta) M(t) - tr \rho_\theta M'(t)| < \epsilon, \forall \theta \in \Theta.
\] (10)

Under the assumptions (A’) or (B), \( E \geq c_0, k, F \) is equivalent to the existence of \( k \)-statistical morphism \( \Gamma \) on \( \{\rho_\theta\}_{\theta \in \Theta} \) with \( \Gamma (\rho_\theta) = \sigma_\theta \), \( \forall \theta \in \Theta \), is \( k \)-statistical morphism, for any \( k \). The following lemma has some implications on its converse statement.
Lemma 13 Suppose $\dim \mathcal{H} < \infty$. Then, any $k$-statistical morphism $\Gamma$ on $\{\rho_\theta\}_{\theta \in \Theta}$ has linear, positive, and trace preserving extension $\Gamma'$ to span $\{\rho_\theta\}_{\theta \in \Theta}$.

Proof. Let $\{\rho_\theta\}_{i=1}^n$ be linear independent elements of $\{\rho_\theta\}_{\theta \in \Theta}$ and define $\Gamma'$ by linear combination of $\{\Gamma(\rho_\theta)\}_{\theta}$:

$$\Gamma' \left( \sum_{i=1}^n a_i \rho_\theta \right) = \sum_{i=1}^n a_i \Gamma(\rho_\theta).$$

Obviously, $\Gamma'$ is linear and trace preserving. First, we prove $\Gamma(\rho_\theta) = \Gamma'(\rho_\theta)$:

By definition, for any $M$ and for any $\varepsilon > 0$, there is $M'$ with (10). Let $\rho_\theta = \sum_{i=1}^n a_i \rho_\theta$. Then,

$$\sum_{t \in D} |\operatorname{tr} \Gamma(\rho_\theta) M(t) - \operatorname{tr} \rho_\theta M' (t)|$$

$$= \sum_{t \in D} \left| \operatorname{tr} \Gamma(\rho_\theta) M(t) - \sum_{i=1}^n a_i \operatorname{tr} \rho_\theta M' (t) \right|$$

$$= \sum_{t \in D} \left| \operatorname{tr} \Gamma(\rho_\theta) M(t) - \sum_{i=1}^n a_i \operatorname{tr} \Gamma(\rho_\theta) M (t) \right|$$

$$= \sum_{t \in D} |\operatorname{tr} \Gamma(\rho_\theta) M(t) - \operatorname{tr} \Gamma'(\rho_\theta) M (t)| < \varepsilon.$$ 

Since $\varepsilon$ is arbitrary, and $M$ is arbitrary $k$-valued measurement, we have $\Gamma(\rho_\theta) = \Gamma'(\rho_\theta)$, and $\Gamma'$ is a linear extension of $\Gamma$.

Finally, we prove that $\Gamma'$ is positive on span $\{\rho_\theta\}_{\theta \in \Theta}$. For any positive matrix $M \leq 1$ and any $\rho = \sum_i a_i \rho_\theta \geq 0$,

$$\operatorname{tr} \Gamma'(\rho) M = \operatorname{tr} \sum_{i=1}^n a_i \Gamma(\rho_\theta) M \geq \operatorname{tr} \sum_{i=1}^n a_i \rho_\theta M' - n \varepsilon \geq -n \varepsilon.$$ 

Since $\varepsilon > 0$ and $M \geq 0$ are arbitrary, we have positivity of $\Gamma'$ on span $\{\rho_\theta\}_{\theta \in \Theta}$.

Theorem 14 Suppose $\dim \mathcal{H} < \infty$ and span $\{\rho_\theta\}_{\theta \in \Theta}$ is the totality of Hermitian matrices. Then, $E_{0,k} \mathcal{F}$ holds if and only if there is a positive trace preserving map $\Gamma$ with $\Gamma(\rho_\theta) = \sigma_\theta$, $\forall \theta \in \Theta$. Namely, $E_{0,2} \mathcal{F}$, $E_{0,3} \mathcal{F}$, ..., $E_{0,k} \mathcal{F}$ are all equivalent to $E_{\geq 0} \mathcal{F}$.

Proof. The first statement follows directly from Lemma 13. As for the second statement, it is obvious that $E_{0,k} \mathcal{F}$ implies $E_{0,2} \mathcal{F}$, $E_{0,3} \mathcal{F}$, ..., $E_{0,k} \mathcal{F}$. Conversely, suppose $E_{0,k} \mathcal{F}$. Then by Lemma 13 there is a positive linear, and trace preserving map $\Gamma$ with $\Gamma(\rho_\theta) = \sigma_\theta$, $\forall \theta \in \Theta$, which implies $E_{\geq 0} \mathcal{F}$. ■

Classically, it is known that $E_{0,2} \mathcal{F}$, $E_{0,3} \mathcal{F}$, ..., $E_{0,k} \mathcal{F}$ are all equivalent to $E_{\geq 0} \mathcal{F}$, provided $\Theta$ is a finite set [12, 13]. The above theorem is a quantum version of this statement.
9 Technicalities

9.1 Weak and weak* topology

For the detail of the following statements, see [5], for example. Let $E$ and $E'$ be a normed Banach space and the totality of continuous linear functional, respectively. If we take as the norm of $f \in E'$ the operator norm $\|f\|$ as a functional, then $E'$ become a normed linear space called conjugate space. $E'$ is complete, and thus is a Banach space. The topology introduced by $\|f\|$ is called strong topology.

The weak* topology $\sigma (E',E)$ in $E'$ is introduced as follows. For every $\alpha > 0$ and every finite number of elements $x_i (i=1,\cdots,n)$, we denote by $W (x_1,\cdots x_n,\alpha)$ the set of all $f$ such that $|f (x_i)| \leq \alpha$. The topology for which the sets $W (x_1,\cdots x_n,\alpha)$ form the fundamental system of the neighbours of zero is called weak* topology. In other words, an open set containing 0 is in a union of the sets $W (x_1,\cdots x_n,\alpha)$. The weak topology $\sigma (E,E')$ in $E$ is defined by exchanging the role of $E$ and $E'$ above.

The weak and weak* topologies are locally convex topologies since sets $W (x_1,\cdots x_n,\alpha)$ are convex.

The sequence $\{f_i\}_{i=1}^n$ in $E'$ is called weakly convergent to the functional $f_0$ if it converges to $f_0$ in the weak* topology. In order for $\{f_i\}_{i=1}^\infty$ to be weakly convergent to $f_0$, it is necessary and sufficient that $\lim_{n \to \infty} f_n (x) = f_0 (x)$ for every $x \in E$.

A convex set in a normed linear space $E'$ has the same closure both in the initial topology and in the weak* topology $\sigma (E',E)$. In particular, if the sequence $\{f_i\}_{i=1}^n$ is weakly convergent to $f_0$, there exists a sequence of linear combinations $\{\sum_{i=1}^m \lambda_i f_i\}$ converging in the norm to $f_0$.

Every closed sphere in $E'$ is compact in the weak* topology $\sigma (E',E)$ (Alaoglu’s theorem).

9.2 Measure theory

A finitely additive set function $\nu$ defined on a finitely additive set family $\mathcal{B}$ of a topological space $\Theta$ is said to be regular if and only if for each $E \in \mathcal{B}$ and $\varepsilon > 0$ there are sets $E', E'' \in \mathcal{B}$ such that $\overline{E'} \subset E \subset (E'')^c$ and $|\nu (C)| < \varepsilon$ for any $C \in \mathcal{B}$ with $C \subset E'' - E'$.

Let $\nu$ be a regular and finitely additive set function defined on a $\sigma$-field $\mathcal{B}$ of a compact space $\Theta$. Then, $\nu$ is countably additive (Theorem III.5.13, [7]).

Baire $\sigma$-field of a topological space is the smallest $\sigma$-field which makes every continuous function measurable.
10 Compact covariant experiments

10.1 Compact covariant experiments

Let $\dim \mathcal{H} < \infty$, $\dim \mathcal{K} < \infty$. Let $G$ be a compact group, and $g \rightarrow U_g \in \text{SU}(\mathcal{H})$ and $g \rightarrow V_g \in \text{SU}(\mathcal{K})$ be representations of $G$. Suppose that there is a natural action $\theta \rightarrow g\theta$ of $g \in G$ on $\theta \in \Theta$. Moreover, we suppose that for any $\theta$, there is $g \in G$ with $g\theta = \theta$. Then we consider the covariant experiments, which satisfy

\[ \rho_{g\theta} = U_g \rho_\theta U_g^\dagger, \quad \sigma_{g\theta} = V_g \sigma_\theta V_g^\dagger, \]

or

\[ \rho_{g\theta} = U_g \rho_\theta U_g^\dagger, \quad \sigma_{g\theta} = V_g \sigma_\theta V_g^\dagger. \]

We further suppose that the assumptions of 8, which are conditions (A') and (B), hold true. Then, due to Theorem 8, we have

\[ \delta^q(E, F) = \inf_{\Phi} \sup_{\theta \in \Theta} \| \Phi(\rho_\theta) - \sigma_\theta \|_1 \]

\[ \leq \inf_{g \in G} \| \Phi(U_g \rho_g U_g^\dagger) - V_g \sigma_0 V_g^\dagger \|_1 \]

\[ = \inf_{g \in G} \| V_g^\dagger \Phi(U_g \rho_0 U_g^\dagger) V_g - \sigma_0 \|_1. \]

Denote by $M$ the average with respect to Haar measure of $G$, and define

\[ \Phi_*(\rho) := MV_g^\dagger \Phi(U_g \rho_0 U_g^\dagger) V_g. \]

Then, $\Phi_*$ is covariant,

\[ \Phi_*(U_g \rho U_g^\dagger) = V_g \Phi_*(\rho) V_g^\dagger, \quad (11) \]

and, by convexity of the norm $\| \cdot \|_1$,

\[ \sup_{g \in G} \| V_g^\dagger \Phi(U_g \rho_0 U_g^\dagger) V_g - \sigma_0 \|_1 \]

\[ \geq \| \Phi_*(\rho_0) - \sigma_0 \|_1 \]

\[ = \| V_g^\dagger \Phi_*(U_g \rho_0 U_g^\dagger) V_g - \sigma_0 \|_1, \quad \forall g \in G. \]

Therefore,

\[ \delta^q(E, F) = \inf_{\Phi_*} \| \Phi_*(\rho_0) - \sigma_0 \|_1, \]

where $\Phi_*$ runs over all the CPTP maps with (11).

Let $Ch_{\Phi_*}$ be the Choi’s representation of a channel $\Phi_*$,

\[ Ch_{\Phi_*} := \Phi_* \otimes I \left( \sum_{i,j=1}^{\dim \mathcal{H}} |i\rangle \langle i| \otimes |j\rangle \langle j| \right), \]

where $\{ |i\rangle \}$ is a CONS of $\mathcal{H}$. Then, (11) can be written as

\[ [U_g \otimes V_g, Ch_{\Phi_*}] = 0, \quad (g \in G), \quad (12) \]
10.2 Examples

Example 15 $\mathcal{H} = \mathcal{K} = \mathbb{C}^d$, $G = \text{SU}(d)$, $U_g = g$, and $V_g = V g V^\dagger$. Then, $\Phi_*$ has to be depolarizing channel,

$$
\Phi_*(X) := \frac{(1 - \lambda) \text{tr} X}{d} 1 + \lambda V^\dagger X V, \quad (0 \leq \lambda \leq 1).
$$

Hence,

$$
\mathcal{E}_g \geq h, F \iff \sigma_0 = \frac{\lambda}{d} 1 + (1 - \lambda) V^\dagger \rho_0 V.
$$

Especially, suppose $\rho_0$ and $\sigma_0$ have the same spectrum. Then, although the set $\{U \rho_0 U^\dagger\}_{U \in \text{SU}(d)}$ equals the set $\{U \sigma_0 U^\dagger\}_{U \in \text{SU}(d)}$, $\mathcal{E}_g \geq h, F$ unless $V^\dagger \rho_0 V = \sigma_0$.

Now, let $\mathcal{H} = \mathcal{K} = \mathbb{C}^2$, and

$$
V^\dagger \rho_0 V = \frac{1}{2} \begin{bmatrix}
1 + u & 0 \\
0 & 1 - u
\end{bmatrix} \quad (u \geq 0),
$$

$$
\sigma_0 = \frac{1}{2} \begin{bmatrix}
1 + z & x - \sqrt{-1} y \\
x + \sqrt{-1} y & 1 - z
\end{bmatrix}.
$$

Then,

$$
\delta^g (\mathcal{E}, F) = \inf_{\Phi_*} \| \Phi_* (\rho_0) - \sigma_0 \|_1
$$

$$
= \inf_{0 \leq \lambda \leq 1} \frac{1}{2} \sqrt{(z - u) + x^2 + y^2}
$$

$$
= \begin{cases}
\frac{1}{2} \sqrt{(z - u)^2 + x^2 + y^2}, & (z \geq u), \\
\frac{1}{2} \sqrt{x^2 + y^2}, & (0 \leq z \leq u), \\
\frac{1}{2} \sqrt{z^2 + x^2 + y^2}, & (z \leq 0).
\end{cases}
$$

When $z \leq 0$, the optimal $\Phi_* (\rho_0) = \frac{1}{2} 1$. Thus, best approximate experiment $\mathcal{E}'$ to $F$ with $\mathcal{E}_g \geq h, \mathcal{E}'$ is $\mathcal{E}' = (\mathbb{C}^2, \{\frac{1}{2} 1\}, \Theta)$.

Example 16 $\mathcal{H} = \mathcal{K} = \mathbb{C}^d$, $d$ is prime power, and $G = \{X_d^z Z_d^t\}_{z,t \in \{0, 1, \ldots, d-1\}}$.

where

$$
X_d = \begin{bmatrix}
0 & \cdots & \cdots & \cdots & 0 & 1 \\
1 & 0 & \ddots & \ddots & \ddots & 0 \\
0 & 1 & \ddots & \ddots & \ddots & \vdots \\
\vdots & \ddots & \ddots & \ddots & \ddots & \ddots \\
0 & \cdots & 0 & 1 & 0
\end{bmatrix},
$$

$$
Z_d = \begin{bmatrix}
1 & 0 & \cdots & \cdots & 0 \\
0 & \omega_d & 0 & \ddots & \vdots \\
\vdots & 0 & \omega_d^2 & \ddots & \vdots \\
\vdots & \ddots & \ddots & \ddots & \ddots \\
0 & \cdots & 0 & \omega_d^{d-1}
\end{bmatrix}.
$$

Also, $U_g = V_g = g$. Note that

$$
\text{Ch}_{\Phi_*} = \sum_{t,s,t',s' \in \{0, 1, \ldots, d-1\}} a_{t,s,t',s'} \sum_{d} X_d^{t} Z_d^{s} \otimes X_d^{t'} Z_d^{s'}.
$$
where \( a_{t,s,t',s'} \) are complex numbers. Since

\[
\left( X_d^{t''} Z_d^{s''} \otimes X_d^{t''} Z_d^{s''} \right) \left( X_d^t Z_d^s \otimes X_d^t Z_d^s \right) = \omega_d^{s''(t'-t)-t''(s'+s')}
\]

(12) implies that \( a_{t,s,t',s'} \) takes non-zero value for \( t, s, t', s' \) with \( t' = t \) and \( s' = d - s \). Therefore, considering that \( C h_{\Phi_*} \) is Hermitian, and that \( \Phi_* \) is trace preserving, the space of channels satisfying (11) is (as a real vector space) \( d^2 - 1 \) dimensional. On the other hand, a channel

\[
\Phi_*(\rho) = \sum_{t,s,t',s' \in \{0,1,\ldots,d-1\}} p_{t,s,t',s'} (X_d^t Z_d^s) \rho (X_d^{t'} Z_d^{s'}) \quad (13)
\]
satisfies (11), and the space of channels with (13) is \( d^2 - 1 \). Hence, (11) is equivalent to (13).

Therefore,

\[
\delta^q(\mathcal{E},\mathcal{F}) = \min_{\rho'} \| \rho' - \sigma_0 \|_1
\]

where \( \rho' \) moves all over the convex hull of the set \( \{ (X_d^t Z_d^s) \rho_0 (X_d^{t'} Z_d^{s'}) ; t, s \in \{0,1,\ldots,d-1\} \} \).

Especially, when \( d = 2 \), letting \( \bar{x}_0 = (x_{01}, x_{02}, x_{03}) \) and \( \bar{y}_0 \) be the Bloch representation of \( \rho_0 \) and \( \sigma_0 \), respectively, we have

\[
\delta^q(\mathcal{E},\mathcal{F}) = \min_{\bar{x}} \| \bar{x} - \bar{y}_0 \|,
\]

where and \( \bar{x} \) moves all over the convex hull of \( (x_{01}, x_{02}, x_{03}), (x_{01}, -x_{02}, x_{03}), (-x_{01}, x_{02}, x_{03}), (x_{01}, -x_{02}, -x_{03}) \), and \( (-x_{01}, x_{02}, -x_{03}) \).

11 Translation experiments

11.1 Models and questions

Let \( \dim H = \dim K = \infty \) (countable), and define

\[
\rho_\theta := W_{A\theta} \rho W_{A\theta}^\dagger, \quad \sigma_\theta := W_{B\theta} \sigma W_{B\theta}^\dagger
\]

where

\[
W_\theta := e^{\sqrt{-1}(\theta^1 P - \theta^2 Q)}, \theta \in \mathbb{R}^2,
\]
is a Weyl operator, and \( A \) and \( B \) are real invertible \( 2 \times 2 \) matrices. Applying Theorem 8 we have

\[
\delta^q(\mathcal{E},\mathcal{F}) = \inf_{\Phi, \theta \in \Theta} \| \Phi(\rho_\theta) - \sigma_\theta \|_1.
\]
11.2 Restriction to covariant maps

The argument of this section draws upon [9]. For any $\Phi$, define

$$\Phi_\theta (X) := W_{B\theta}^\dagger \Phi \left(W_A \rho W_{A\theta}^\dagger\right) W_{B\theta}.$$  

Then,

$$\delta^q (\mathcal{E}, \mathcal{F}) = \inf_{\Phi} \sup_{\theta \in \Theta} \|\Phi_\theta (\rho) - \sigma\|_1$$

$$= \inf_{\Phi} \sup_{\theta \in \Theta} \sup_{X \geq 0, \|X\| \leq 1} \text{tr} \left(\Phi_\theta (\rho) - \sigma\right) X$$

$$= \inf_{\Phi} \sup_{\theta \in \Theta} \sup_{X \geq 0, \|X\| \leq 1} \text{tr} \left(\Phi_\theta (\rho) - \sigma\right) X$$

$$\geq \inf_{\Phi} \sup_{\theta \in \Theta} \text{M}_\theta \text{tr} \left(\Phi_\theta (\rho) - \sigma\right) X,$$

where $M_\theta$ is the invariant mean of the translation group in $\mathbb{R}^2$. Note, if $\rho$ is a density operator, the map

$$\Phi_* (\rho) : X \to \text{M}_\theta \text{tr} \Phi_\theta (\rho) X$$

is linear and bounded, and maps 1 to 1. Also, the mapping $\Phi_* : \rho \to \Phi_* (\rho)$ is linear, and covariant:

$$\Phi_* \left(W_A \rho W_{A\theta}^\dagger\right) [X] = \Phi_* (\rho) \left[W_{B\theta}^\dagger X W_{B\theta}\right]. \quad (14)$$

Thus,

$$\sup_{\theta \in \Theta} \sup_{X \geq 0, \|X\| \leq 1} \left(\Phi_* (\rho) [X] - \text{tr} \sigma X\right)$$

$$= \sup_{\theta \in \Theta} \sup_{X \geq 0, \|X\| \leq 1} \left(\Phi_* (\rho) \left[W_{B\theta}^\dagger X W_{B\theta}\right] - \text{tr} \sigma W_{B\theta}^\dagger X W_{B\theta}\right)$$

$$= \sup_{X \geq 0, \|X\| \leq 1} \left(\Phi_* (\rho) [X] - \text{tr} \sigma X\right).$$

Hence, in optimizing $\Phi$, we just have to consider $\Phi_*$ with covariant property [13].

$\Phi_*$ is seemingly difficult to handle, since its output state may not be normal, i.e., may not have the density. However, it turns out that $\Phi_*$ with non-normal output is not optimal. Since $S_1 (\mathcal{H})$ is the dual of the space of compact operators $\mathcal{LC} (\mathcal{H})$, there is a positive $Y_\rho \in S_1 (\mathcal{H})$ with

$$\Phi_* (\rho) [X] = \text{tr} Y_\rho X, \forall X \in \mathcal{LC} (\mathcal{H}).$$

Consider the map $\rho \to \sum_{i=1}^\infty \Phi_* (\rho) \left|i\right> \left<i\right| = \text{tr} Y_\rho$. Since this is linear in $\rho$ and positive, there is a positive operator $T$ with

$$\text{tr} Y_\rho = \text{tr} T \rho.$$
Due to covariant property of $\Phi^*$, $T$ commutes $W_{B\theta}$ for all $\theta \in \mathbb{R}^2$. Therefore, $T = c1$. Thus, $c = \text{tr} Y_{\rho}$ is independent of the input $\rho$. Therefore, $\rho_* := \frac{1}{c} Y_{\rho}$ is a density operator. We denote by $\Phi'$ the CPTP map which sends $\rho$ to $\rho_*$. Letting $\{A_n\}$ be a sequence of compact operators such that $\lim_{n \to \infty} \text{tr} (\sigma - c \rho_*)_X = \text{tr} (\sigma - \rho_*)_X$ (0 $\leq c \leq 1$),

$$
\sup_{X \geq 0, \|X\| \leq 1} (\Phi_*(\rho) [X] - \text{tr} X) = \sup_{X \geq 0, \|X\| \leq 1} |\Phi_*(\rho) [X] - \text{tr} X| \\
\geq \lim_{n \to \infty} |\text{tr} (\sigma - c \rho_*)_X| \\
= \text{tr} [\sigma - c \rho_*]_+ \\
\geq \text{tr} [\sigma - \rho_*]_+ = \|\sigma - \rho_*\|_1 = \|\Phi'_*(\rho) - \sigma\|_1.
$$

(15)

Therefore, $\Phi'_*$ is always better than $\Phi_*$. After all, we have

$$
\inf \sup_{\theta \in \Theta} \|\Phi(\rho_\theta) - \sigma_\theta\|_1 = \inf \|\Phi(\rho) - \sigma\|_1,
$$

where $\Phi$ runs over all the CPTP maps with

$$
\Phi \left(W_{A\theta}^* \rho W_{A\theta}^\dagger\right) = W_{B\theta}^\dagger \Phi(\rho) W_{B\theta},
$$

or

$$
\Phi^* \left(W_{B\theta}^\dagger \sigma W_{B\theta}\right) = W_{A\theta}^\dagger \Phi^*(\sigma) W_{A\theta}.
$$

(16)

### 11.3 Characterization of covariant maps

Inserting $X = W_\xi$, one has

$$
\Phi^* \left(W_{B\theta}^\dagger \sigma W_{B\theta}\right) = W_{A\theta}^\dagger \Phi^*(\sigma) W_{A\theta},
$$

$$
e^{-\sqrt{-1}J^TJ B\theta} W_{A\theta}^\dagger \Phi^*(W_{\xi}) = \Phi^*(W_{\xi}) W_{A\theta},
$$

where

$$
J = \begin{bmatrix}
0 & 1 \\
-1 & 0
\end{bmatrix}.
$$

Since this holds for any $\theta$, we have

$$
\Phi^*(W_{\xi}) = c(\xi) W_{C\xi},
$$

where $C$ satisfies

$$
C^T JA = JB.
$$

Using the identity

$$
A^T JA = (\det A) J,
$$
or
\[ JA = (\det A)A^{T^{-1}}J, \]
we have
\[ C = \frac{\det B}{\det A}AB^{-1}. \]

According to Theorem 2.3 of [6], for \( \Phi^* \) to be trace preserving, if \( \det A \neq \det B \), \( c(\xi) \) has to be a non-commutative characteristic function
\[ c(\xi) = \text{tr} \rho' W_\Omega \xi, \]
where \( \rho' \) is a density matrix and \( \Omega \) is a (non-zero) matrix satisfying
\[ J - C^TJC = \Omega^TJ \Omega, \]
or
\[ \Omega = (1 - \det C)^{1/2} S = \left( 1 - \frac{\det B}{\det A} \right)^{1/2} S, \]
with
\[ \det S = 1. \]

Hence,
\[ \text{tr} \Phi(\rho) W_\xi = \text{tr} \rho \Phi^*(W_\xi) = \text{tr} \rho' W_\Omega \text{tr} \rho W_\xi. \]  \hspace{1cm} (17)

When \( A = B, \Omega = 0. \) In this case, \( c(\xi) \) is a characteristic function of a classical probability distribution \( F \) over \( \mathbb{R}^2 \),
\[ c(\xi) = \int e^{\sqrt{-1}(\xi_1 z^2 - \xi_2 z^1)} \frac{dF(x)}{2\pi}, \]
and we have
\[ \text{tr} \rho \Phi^*(W_\xi) = c(\xi) \text{tr} \rho W_\xi. \]  \hspace{1cm} (18)

Letting \( P_\rho \) be the \( P \)-function of \( \rho \), respectively, we have
\[ \text{tr} \Phi(\rho) W_\xi = \text{tr} \int P_{\Phi(\rho)}(z) W_\xi W_z |0\rangle \langle z| \frac{dz}{2\pi} = \text{tr} \int P_{\Phi(\rho)}(z) e^{\sqrt{-1}(\xi_1 z^2 - \xi_2 z^1)} W_z W_\xi |0\rangle \langle z| \frac{dz}{2\pi} \]
\[ = \langle 0| \xi \rangle \int P_{\Phi(\rho)}(z) e^{\sqrt{-1}(\xi_1 z^2 - \xi_2 z^1)} \frac{dz}{2\pi}, \]
and thus,
\[ \int P_{\Phi(\rho)}(z) e^{\sqrt{-1}(\xi_1 z^2 - \xi_2 z^1)} \frac{dz}{2\pi} = c(\xi) \int P_{\rho}(z) e^{\sqrt{-1}(\xi_1 z^2 - \xi_2 z^1)} \frac{dz}{2\pi}. \]

Therefore, in case \( A = B, \) (10) is equivalent to
\[ P_{\Phi(\rho)}(x) = \int P_{\rho}(x - y) dF(y). \]

Therefore, denoting convolution by \( ' * ' \),
\[ \inf_{\Phi} \sup_{\theta} \| \Phi(\rho_\theta) - \sigma_\theta \|_1 = \inf \left\{ \| \rho' - \sigma \|_1 : \exists F \ P_{\rho'} = P_{\rho} * F \right\} \]
\[ \leq \| P_{\sigma} - P_{\rho} * F \|_1, \ \forall F. \]  \hspace{1cm} (19)
11.4 Gaussian shift models

When \( \rho \) is a Gaussian state with mean value zero, \( \rho \) satisfies

\[
\text{tr} \rho W_\xi = e^{-\xi^T \Sigma_\rho \xi/4},
\]

where

\[
\Sigma_\rho = \begin{bmatrix}
\frac{1}{2} \text{tr} \rho Q^2 & \frac{1}{2} \text{tr} \rho (PQ + QP) \\
\frac{1}{2} \text{tr} \rho (PQ + QP) & \text{tr} \rho P^2
\end{bmatrix}.
\]

The necessary and sufficient condition for \( \rho' \) with \( e^{-\xi^T \Sigma_\rho \xi/4} = \text{tr} \rho' W_\xi \) to exist is

\[
\Sigma + \sqrt{-1} J \geq 0,
\]

which is, being \( 2 \times 2 \) matrices, equivalent to

\[
\det \Sigma \geq 1.
\]

Suppose \( A = B \). Then, due to (18),

\[
\mathcal{E} \geq_0 \mathcal{F} \iff \Sigma_\rho \leq \Sigma_\sigma.
\]

Suppose \( \det A \neq \det B \). Then, if \( \mathcal{E} \geq_0 \mathcal{F} \), due to (17), \( \text{tr} \rho' W_\Omega \xi \) is also Gaussian, \( \text{tr} \rho' W_\xi = e^{-\xi^T \Sigma_\rho \xi/4} \), or

\[
\text{tr} \rho' W_\Omega \xi = e^{-\xi^T \Sigma_\rho \Omega \xi/4}.
\]

Also, by (17), we have

\[
\Sigma_\sigma = \Omega^T \Sigma_\rho \Omega + C^T \Sigma_\rho C,
\]

or, with \( A' = AB^{-1} \)

\[
\Sigma_\rho = \left(1 - (\det A')^{-1}\right) S^T \Sigma_\rho' S + (\det A')^{-2} A'^T \Sigma_\rho A'.
\]

Therefore, for \( \rho' \) with \( \text{tr} \rho' W_\xi = e^{-\xi^T \Sigma_\rho \xi/4} \) to exist, the following is necessary and sufficient:

\[
\Sigma_\sigma - (\det A')^{-2} A'^T \Sigma_\rho A' \geq 0
\]

and

\[
\det \left[ \left(1 - (\det A')^{-1}\right)^{-1} S^{-1} \left( \Sigma_\sigma - (\det A')^{-2} A'^T \Sigma_\rho A' \right) S^{-1} \right]
= \left(1 - (\det A')^{-1}\right)^{-2} \det \left[ \left( \Sigma_\sigma - (\det A')^{-2} A'^T \Sigma_\rho A' \right) \right] \geq 1.
\]

This gives the necessary and sufficient condition for \( \mathcal{E} \geq_0 \mathcal{F} \). With \( \Sigma_\sigma = \Sigma_\rho = a^2 \mathbf{1} \), these conditions can be written as

\[
A'^T A' \geq \mathbf{1},
\]

\[
\det \left[ \left(1 - (\det A')^{-2} A'^T A' \right) \right] \geq a^{-4} \left(1 - (\det A')^{-1}\right)^2.
\]
The first inequality is derived from (21) by elementary but tedious component-wise computation.

In classical case, with $\Sigma_\sigma = \Sigma_\rho = a^2 \mathbf{1}$, $\mathcal{E} \geq_0 \mathcal{F}$ is equivalent to (23) [13]. In quantum case, when $a \gg 1$, (23) implies (24), and thus $\mathcal{E} \geq_0 \mathcal{F}$.

However, when $a$ is not very large, quantum case is very much different from classical case. For example, suppose $a = 1$. Without loss of generality, let

$$A' = O \begin{bmatrix} \alpha & \gamma \\ 0 & \beta \end{bmatrix},$$

where $O$ is an orthogonal matrix. Then, (24) is written as

$$-(\alpha - \beta)^2 - \gamma^2 \geq 0.$$

Hence, $\mathcal{E} \geq_0 \mathcal{F}$ is equivalent to

$$A' = \alpha O,$$

where $\alpha \geq 1$ and $O$ is an orthogonal matrix. This is very much stronger than the classical condition (23).

References

[1] P. Alberti, "On the simultaneous transformation of density operators by means of a completely positive", Unity Preserving Linear Map. Publ. RIMS (Kyoto), 21:617–644, 1985.

[2] P. Alberti and A. Uhlmann, "A problem relating to positive linear maps on matrix algebras", Rep. Math. Phys. 18, 163-176 (1980).

[3] F. Buscemi, Comparison of quantum statistical models: a "Quantum Blackwell theorem", http://arxiv.org/abs/1004.3794 (2010).

[4] A. Chefles, R. Jozsa, and A. Winter, "On the existence of physical transformations between sets of quantum states "arXiv:quant-ph/0307227.

[5] J. B. Conway, "A Course in Functional Analysis", (2nd ed.) Springer-Verlag (1997).

[6] B. Demoen, P. Vanheuverzwijn and A. Verbeure, "Completely positive maps on the CCR-algebra," Lett. Math. Phys. 2, 161 (1977).

[7] N. Dunford, J. Schwartz, "Linear Operators, Part I", Interscience (1952).

[8] A. Jencova, private communication (2011).

[9] O. Kruger, "Quantum Information Theory with Gaussian Systems", doctoral dissertation (2006).
[10] J. L. Kelley, "General Topology", Graduate Texts in Mathematics, Springer-Verlag (1975).

[11] H. Strasser, "Mathematical Theory of Statistics", de Gruyter (1985).

[12] E. Torgersen, "Comparison of statistical experiments when the parameter space is finite", Z. Wahrscheinlichkeitstheorie verw. Geb. 16, 219-249 (1970).

[13] E. Torgersen, "Comparison of Statistical Experiments", Cambridge University Press (1991).