A new quantum version of $f$-divergence

Keiji Matsumoto
Quantum Computation Group, National Institute of Informatics,
2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo 101-8430,
e-mail:keiji@nii.ac.jp

September 25, 2014

Abstract
Quantum $f$-Divergence proposed by Petz (see [7] for comprehensive review) is widely studied, especially due to its implication to perfect error correction. In [8], the present author proposed a new quantum version of "generalized fidelity". This is essentially a new version of quantum $f$-divergence ($D^{\max}_f$, below) when $f$ is operator monotone decreasing. In the definition, we supposed two states are strictly positive. In this paper, the definition is generalized to include operator convex function $f$ and, to states which may have null eigenspace. Not only the definition, also the explicit formula for $D^{\max}_f$ when $\rho$ and $\sigma$ have null eigenspaces is given, and several properties of the quantity is studied (convexity, monotonicity by CPTP maps.) Also, the condition $D^{\max}_f(\rho||\sigma) = D^{\max}_f(\Lambda(\rho)||\Lambda(\sigma))$, where $\Lambda$ is CPTP map, is studied. As is well-known, the analogous condition for $D_f$ implied reversibility of $\Lambda$, or "sufficiency" of the map $\Lambda$. $D^{\max}_f$-version of the condition gives another type of "sufficiency" property of $\Lambda$.

1 Introduction
Quantum $f$-Divergence proposed by Petz (see [7] for comprehensive review) is widely studied, especially due to its implication to perfect error correction. In [8], the present author proposed a new quantum version of "generalized fidelity". This is essentially a new version of quantum $f$-divergence ($D^{\max}_f$, below) when $f$ is operator monotone decreasing. In the definition, we supposed two states are strictly positive. In this paper, the definition is generalized to include operator convex function $f$ and, to states which may have null eigenspace. Not only the definition, also the explicit formula for $D^{\max}_f(\rho||\sigma)$ when $\rho$ and $\sigma$ have null eigenspaces is given, and several properties of the quantity is studied (convexity, monotonicity by CPTP maps.) Also, the condition $D^{\max}_f(\rho||\sigma) = D^{\max}_f(\Lambda(\rho)||\Lambda(\sigma))$, where $\Lambda$ is CPTP map, is studied. As is well-known, the analogous condition for $D_f$ implied reversibility of $\Lambda$, or "suf-
ficiency” of the map $\Lambda$. $D_f^{\max}$-version of the condition gives another type of "sufficiency" property of $\Lambda$.

In the following, Hilbert spaces are denoted by symbols such as $\mathcal{H}_A$, $\mathcal{H}_B$ etc., $\mathcal{H}$ with subscript denoting the name of each system. The dimension of $\mathcal{H}_A$, $\mathcal{H}_B$ etc. is denoted by $|A|$, $|B|$, etc. In the paper, it is assumed that dimensions of Hilbert spaces are finite, $|A|, |B| < \infty$. The set of operators and density operators over $\mathcal{H}_A$ will be denoted by $\mathcal{L}_A$ and $\mathcal{S}_A$, respectively. The composite system $AB$ is the system corresponding to the Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$. Operators and density operators over $\mathcal{H}_A \otimes \mathcal{H}_B$ will be denoted by $\mathcal{L}_{AB}$ and $\mathcal{S}_{AB}$, respectively. The identity operator in $\mathcal{H}_A$ and identity transform in $\mathcal{L}_A$ will be denoted by $1_A$ and $I_A$, respectively.

2 Reverse test and maximal $f$-divergence

A reverse test of a pair $\{\rho, \sigma\}$ of positive definite operators over system $A$ is a triplet $(\Gamma, \{p, q\})$ of a trace preserving positive linear map from non-negative measures over some a finite set $X$ (or commutative algebra with dimension $|X|$) to $\mathcal{L}_A$, and non-negative measures $p$ and $q$ over $X$, such that

$$\Gamma (p) = \rho, \Gamma (q) = \sigma.$$ (Note $\Gamma$ is necessarily CPTP.) Consider the following condition of the function $f$.

(FC) $f$ is a proper closed convex function on $[0, \infty)$. Also, $f(0) = 0$.

For a function $f$ satisfying above (FC), we define maximal $f$-divergence

$$D_f^{\max} (\rho || \sigma) = \inf_{(\Gamma, \{p, q\})} D_f (p || q),$$

where the inf is taken over all the reverse tests, and $D_f$ is a $f$-divergence

$$D_f (p || q) := \sum_{x \in X} q(x) f \left( \frac{p(x)}{q(x)} \right).$$

In the definition of $f$ divergence, the convention is

$$0 \cdot f (\gamma/0) := \lim_{\varepsilon \downarrow 0} f (\gamma/\varepsilon).$$ (1)

The name comes from the fact that $D_f^{\max} (\rho || \sigma)$ is the largest quantum version of $D_f (p || q)$; here, quantum version of $D_f (p || q)$ is any $D_f^Q (\rho || \sigma)$ such that

(D1) $D_f^Q (\Lambda (\rho) || \Lambda (\sigma)) \leq D_f^Q (\rho || \sigma)$ holds for any CPTP map $\Lambda : \mathcal{L}_A \rightarrow \mathcal{L}_B$, any $\rho, \sigma \in \mathcal{S}_A$, for any Hilbert spaces $\mathcal{H}_A, \mathcal{H}_B$.

(D2) $D_f^Q (p || q) = D_f (p || q)$ for any probability distributions $p, q$ over any finite sets.

2
Here $D_f^Q (p||q)$ should be understood as a short for
\[
D_f^Q \left( \sum_{x \in X} p(x) |e_x\rangle \langle e_x| \right) \left( \sum_{x \in X} q(x) |e_x\rangle \langle e_x| \right),
\]
where $\{|e_x\rangle : x \in X\}$ is an orthonormal system of vectors. Note this does not depend on the choice of an orthonormal system of vectors, since
\[
D_f^Q (U \rho U^\dagger || U \sigma U^\dagger) = D_f^Q (\rho||\sigma)
\]
for any unitary operators due to (D1).

We also consider the following condition which is stronger than (D1).

(D1') $D_f^Q (\Lambda (\rho) || \Lambda (\sigma)) \leq D_f^Q (\rho||\sigma)$ holds for any trace preserving positive map $\Lambda : \mathcal{L}_A \to \mathcal{L}_B$, any $\rho, \sigma \in S_A$, for any Hilbert spaces $\mathcal{H}_A, \mathcal{H}_B$.

**Lemma 1** If (FC) is satisfied, $D_f^{\max}$ satisfies above (D1), (D1') and (D2). Also, if a two point functional $D_f^Q$ satisfies satisfies both of (D1) and (D2), or both of (D1') and (D2),
\[
D_f^Q (\rho||\sigma) \leq D_f^{\max} (\rho||\sigma).
\]

**Proof.** Let $\Lambda$ be a trace preserving positive map. Then,
\[
D_f^{\max} (\Lambda (\rho) || \Lambda (\sigma)) = \inf_{(\Gamma, \{p, q\})} \{ D_f (p||q) : (\Gamma, \{p, q\}) : a reverse test of $\{\Lambda (\rho), \Lambda (\sigma)\}$ \}
\leq \inf_{(\Gamma, \{p, q\})} \{ D_f (p||q) : \Gamma = \Gamma' \circ \Lambda, (\Gamma', \{p, q\}) : a reverse test of $\{\rho, \sigma\}$ \}
= D_f^{\max} (\rho||\sigma).
\]
Hence, $D_f^{\max}$ satisfies (D1'), and thus (D1) also. Also,
\[
D_f^{\max} (p||q) = \inf \{ D_f (p'||q') : p = \Gamma (p'), q = \Gamma (q'), \Gamma: \text{stochastic map} \}
\geq \inf \{ D_f (\Gamma (p') || \Gamma (q')) : p = \Gamma (p'), q = \Gamma (q'), \Gamma: \text{stochastic map} \}
= D_f (p||q).
\]
The opposite inequality is trivial. so we have $D_f^{\max} (p||q) = D_f (p||q)$. Thus, $D_f^{\max}$ satisfies (D2).

Suppose $D_f^Q$ satisfies (D1) (or (D1')) and (D2), and let $(\Gamma, \{p, q\})$ be a reverse test of $\{\rho, \sigma\}$. Then,
\[
D_f^Q (\rho||\sigma) = D_f^Q (\Gamma (p) || \Gamma (q)) \leq D_f^Q (p||q) = D_f (p||q).
\]
Therefore, taking infimum over all the reverse tests of $\{\rho, \sigma\}$, we have $D_f^Q (\rho||\sigma) \leq D_f^{\max} (\rho||\sigma)$. ■
Lemma 2 Suppose (FC) is satisfied. For any non-negative $\gamma$,

$$\lim_{\varepsilon \downarrow 0} \varepsilon f \left( \frac{\gamma}{\varepsilon} \right) = \lim_{y \to \infty} \frac{f(\gamma y)}{y}$$

exists in $(-\infty, \infty]$. Also, for any non-negative operators $X$ and $Y$, $D_f^{\max}(X||Y)$ has a value in $(-\infty, \infty]$. 

Proof. Since $f$ is proper on $[0, \infty)$, there is non-negative $y_0$ such that $f(y_0)$ is finite. Since $f$ is convex, the function

$$y \to g(y) := \frac{f(y) - f(y_0)}{y - y_0}$$

is non-decreasing. 

Also, there is $y'$ such that $g(y') > -\infty$: otherwise, for all $y > y_0$, $f(y) = -\infty$, contradicting with the assumption that $f$ is proper on $[0, \infty)$. Therefore, since $g(y)$ is non-increasing, when $y$ is large enough, $g(y) > -\infty$. Therefore, if $\gamma > 0$,

$$\lim_{y \to \infty} \frac{f(\gamma y)}{y} = \lim_{y \to \infty} \frac{f(y)}{y} = \lim_{y \to \infty} \frac{f(y_0) + (y - y_0) g(y)}{y} = \lim_{y \to \infty} \left(1 - \frac{y_0}{y}\right) g(y) > -\infty$$

If $\gamma = 0$, obviously,

$$\lim_{\varepsilon \downarrow 0} \varepsilon f \left( \frac{\gamma}{\varepsilon} \right) = \lim_{\varepsilon \downarrow 0} \varepsilon f(0) > -\infty.$$

Also, since $f$ is proper, closed, convex, and finite on $[0, \infty)$, by Propositions 26, 27, and 28 there is a finite number $\alpha_1$ and $\alpha_2$ such that

$$D_f(p||q) \geq \sum_{x \in X} \left( \alpha_1 p(x) + \alpha_2 q(x) \right).$$

$$= \alpha_1 \sum_{x \in X} p(x) + \alpha_2 \sum_{x \in X} q(x)$$

$$= \alpha_1 \text{tr} X + \alpha_2 \text{tr} Y.$$ 

Since the last end does not depends on $p$ and $q$, we have

$$D_f^{\max}(X||Y) \geq \alpha_1 \text{tr} X + \alpha_2 \text{tr} Y > -\infty.$$

3 Commutative Radon-Nikodym derivative and the minimal reverse test

Let $\rho$ and $\sigma \in \mathcal{L}_A$ be positive matrices, and denote by $\pi_{\rho}$ the projection onto $\text{supp}\rho$. In this section, we suppose

$$\pi_{\sigma} \geq \pi_{\rho}.$$
When this is true, the commutative Radon-Nikodym derivative of \( \rho \) with respect to \( \sigma \) is defined by

\[
d (\rho, \sigma) := \sigma^{-1/2} \rho \sigma^{-1/2}.
\]

( \( \sigma^{-1} \) is generalized inverse.)

Based on \( d(\rho, \sigma) \), we define the minimal reverse test by the following composition. Let

\[
d (\rho, \sigma) = \sum x d_x P_x
\]

be the spectrum decomposition of \( d(\rho, \sigma) \), where \( d_x \) is an eigenvalue of \( d(\rho, \sigma) \) and \( P_x \) is the projector onto the corresponding eigenspace. Then we define

\[
q (x) := \text{tr} \sigma P_x, \quad p (x) := d_x q (x),
\]

\[
\Gamma (\delta_x) := \frac{1}{q (x)} \sqrt{\sigma} P_x \sqrt{\sigma}.
\]

(2)

Obviously,

\[
\Gamma (q) = \sum x q (x) \frac{q (x)}{q (x)} \sqrt{\sigma} P_x \sqrt{\sigma} = \sigma.
\]

Also,

\[
\rho = \sqrt{\sigma} d (\rho, \sigma) \sqrt{\sigma}
\]

\[
= \sum x d_x \sqrt{\sigma} P_x \sqrt{\sigma}
\]

\[
= \sum x d_x q (x) \Gamma (\delta_x) = \sum x p (x) \Gamma (\delta_x)
\]

\[
= \Gamma (p).
\]

Therefore, \( (\Gamma, \{p, q\}) \) is a reverse test of \( \{\rho, \sigma\} \).

The minimal reverse test satisfies

\[
D'_f (\rho||\sigma) = D_f (p||q),
\]

(3)

where

\[
D'_f (\rho||\sigma) := \text{tr} \sigma f (d (\rho, \sigma)) = \text{tr} \sigma f \left( \sigma^{-1/2} \rho \sigma^{-1/2} \right)
\]

\[
= \text{tr} \sigma \cdot \sigma^{-1/2} f (\rho \sigma^{-1}) \sigma^{1/2} = \text{tr} \sigma f (\rho \sigma^{-1})
\]

\[
= \text{tr} \sigma f (\sigma^{-1} \rho).
\]
Indeed,

\[
D_f (p||q) = \sum_x q(x) f\left(\frac{p(x)}{q(x)}\right) \\
= \sum_x q(x) f(\sigma_p) = \sum_x \text{tr} \sigma P_x f(\sigma_p) \\
= \text{tr} \sigma \sum_x P_x f(\sigma_p) P_x \\
= \text{tr} \sigma f(\sigma_p) = D_f^\prime (p||q).
\]

4 An expression of \(D_f^{\text{max}}\)

From here, we basically suppose that \(f\) satisfies the following condition:

\(\text{\textbf{(F)}}\) \(f\) is an operator convex function such that \(f(x) < \infty\) and continuous for \(x \in [0, \infty)\). Also, \(f(0) = 0\).

Note such \(f\) has necessarily proper and closed extension to \(\mathbb{R}\), i.e., satisfies \((\text{FC})\). If \(\pi_{\sigma} \not\geq \pi_{\rho}\), we define

\[
D_f^\prime (\rho||\sigma) := \inf_{\rho_1} \left\{ \text{tr} \sigma f(\sigma_{\rho_1}) + \lim_{\varepsilon \downarrow 0} \varepsilon f(1 - \text{tr} \rho_{\rho_1}) : \rho \geq \rho_1 \geq 0, \pi_{\sigma} \geq \pi_{\rho_1} \right\}
\]

For a trace preserving positive map \(\Lambda : \mathcal{L}_A \rightarrow \mathcal{L}_B\), define

\[
\Lambda_\sigma (Z) := \{\Lambda (\sigma)\}^{1/2} \Lambda \left(\sigma^{1/2} Z \sigma^{1/2}\right) \{\Lambda (\sigma)\}^{-1/2},
\]

which satisfies

\[
\Lambda_\sigma (1_A) = \Lambda_\sigma (\pi_\sigma) = \pi_{\Lambda(\sigma)}.\]

Thus, viewed as a map from \(\pi_\sigma \mathcal{L}_A \pi_\sigma\) to \(\pi_{\Lambda(\sigma)} \mathcal{L}_B \pi_{\Lambda(\sigma)}\), \(\Lambda_\sigma\) is unital. \(\Lambda_\sigma\) satisfies

\[
\Lambda_\sigma (d(\rho, \sigma)) = d(\Lambda (\rho), \Lambda (\sigma)).
\]

\textbf{Lemma 3} Suppose that \(\text{\textbf{(F)}}\) is satisfied, and \(\pi_\sigma \geq \pi_\rho\). Suppose also \(\Lambda\) is trace preserving and positive. Then,

\[
\Lambda (\sigma)^{1/2} f(\sigma_{\rho}) \Lambda (\sigma)^{1/2} \\
\leq \Lambda (\sigma)^{1/2} \Lambda_\sigma (f(\sigma_{\rho})) \Lambda (\sigma)^{1/2} \\
= \Lambda \left(\sigma^{1/2} f(\sigma_{\rho}) \sigma^{1/2}\right).
\]

\textbf{Proof.} If \(\pi_\sigma \geq \pi_\rho\), and \(\Lambda\) is a CPTP map from \(\mathcal{L}_A\) to \(\mathcal{L}_B\), then

\[
\pi_\sigma \geq \pi_\rho \implies \pi_{\Lambda(\sigma)} \geq \pi_{\Lambda(\rho)}.
\]

Therefore, \(d(\Lambda (\rho), \Lambda (\sigma))\) exists and finite.
Also,

\[
\Lambda (\sigma)^{1/2} f (d (\Lambda (\rho) , \Lambda (\sigma))) \Lambda (\sigma)^{1/2} \\
= (a) \Lambda (\sigma)^{1/2} f (\Lambda_{\sigma} (d (\rho, \sigma))) \Lambda (\sigma)^{1/2} \\
\leq (b) \Lambda (\sigma)^{1/2} \Lambda_{\sigma} (f (d (\rho, \sigma))) \Lambda (\sigma)^{1/2} \\
= (c) \Lambda (\sigma)^{1/2} f (d (\rho, \sigma)) \Lambda (\sigma)^{1/2} ,
\]

where (a), (b) and (c) is by (6), Proposition 18, and definition of \( \Lambda_{\sigma} \), respectively.

\[\text{Lemma 4} \] Suppose (F) is satisfied. Then \( D_f' (\rho||\sigma) \) satisfies conditions (D1'), (D2).

\[\text{Proof.} \] If \( \pi_{\sigma} \geq \pi_{\rho} \), by (7),

\[
D_f' (\Lambda (\rho) \| \Lambda (\sigma)) = \text{tr} \Lambda (\sigma) f (d (\Lambda (\rho) , \Lambda (\sigma))) \\
\leq \text{tr} \sigma^{1/2} f (d (\rho, \sigma)) \sigma^{1/2} = D_f' (\rho||\sigma) .
\]

Also, \( D_f' (p||q) = D_f (p||q) \), obviously.

Suppose \( \pi_{\rho} \not\geq \pi_{\sigma} \).

\[
D_f' (\Lambda (\rho) \| \Lambda (\sigma)) = \inf_{\rho_1} \left\{ \text{tr} \Lambda (\sigma) f (d (\rho_1, \Lambda (\sigma))) + \lim_{\epsilon \downarrow 0} \epsilon f \left( 1 - \frac{\text{tr} \rho_1}{\epsilon} \right) ; \ \text{\( \Lambda (\rho) \geq \rho_1 \geq 0, \ \pi_{\lambda(\sigma)} \geq \pi_{\rho_1} \)} \right\} \\
\leq \inf_{\rho'_1} \left\{ \text{tr} \Lambda (\sigma) f (d (\Lambda (\rho'_1) , \Lambda (\sigma))) + \lim_{\epsilon \downarrow 0} \epsilon f \left( 1 - \frac{\text{tr} \Lambda (\rho'_1)}{\epsilon} \right) ; \ \rho \geq \rho'_1 \geq 0, \ \pi_{\sigma} \geq \pi_{\rho'_1} \right\} \\
\leq \inf_{\rho'_1} \left\{ \text{tr} \sigma f (d (\rho'_1, \sigma)) + \lim_{\epsilon \downarrow 0} \epsilon f \left( 1 - \frac{\text{tr} \rho'_1}{\epsilon} \right) ; \ \rho \geq \rho'_1 \geq 0, \ \pi_{\sigma} \geq \pi_{\rho_1} \right\} \\
= D_f' (\rho||\sigma) .
\]

\[\text{Theorem 5} \] Suppose (F) is satisfied. Then,

\[
D_f' (\rho||\sigma) = D_f^{\max} (\rho||\sigma) \quad (9)
\]

\[\text{Proof.} \] By Lemmas 4 and 5 we have

\[
D_f' (\rho||\sigma) \leq D_f^{\max} (\rho||\sigma) .
\]

So it remains to show the opposite inequality.
If \( \pi_\sigma \geq \pi_\rho \), by (3), the minimal reverse test achieves equality above. Hence, suppose \( \pi_\sigma \not\geq \pi_\rho \). Let \((\Gamma, \{p, q\})\) be a reverse test of \(\{\rho, \sigma\}\) such that

\[
\mathcal{X} = \mathcal{X}_q \cup \{x_0\},
\]

\[
q(x) \begin{cases} > 0 & \text{for } x \in \mathcal{X}_q, \\ = 0 & \text{for } x = x_0, \\ \end{cases}
\]

\[
\Gamma (\delta_{x_0}) = \frac{1}{1 - \text{tr} \rho_1} (\rho - \rho_1),
\]

\[
p(x_0) = 1 - \text{tr} \rho_1,
\]

\[
\sum_{x \in \mathcal{X}_q} p(x) = \rho_1.
\]

Then,

\[
D_f (p||q) = \sum_{x \in \mathcal{X}_q} q(x) f \left( \frac{p(x)}{q(x)} \right) + \lim_{\epsilon \downarrow 0} \epsilon f \left( \frac{1 - \text{tr} \rho_1}{\epsilon} \right).
\]

Therefore,

\[
D_f^{\text{max}} (\rho||\sigma) \leq \inf_{\rho_1} \left\{ D_f' (\rho_1||\sigma) + \lim_{\epsilon \downarrow 0} \epsilon f \left( \frac{1 - \text{tr} \rho_1}{\epsilon} \right) : \rho \geq \rho_1 \geq 0, \pi_\sigma \geq \pi_\rho \right\}
\]

\[
= D_f' (\rho||\sigma).
\]

Hence, we have the assertion. \(\blacksquare\)

5 Properties of \(D_f^{\text{max}}\) and \(D'_f\)

Theorem 6 When \(f\) satisfies (FC), for all \(c (0 \leq c \leq 1)\),

\[
D_f^{\text{max}} (c\rho_0 + (1-c) \rho_1 || c\sigma_0 + (1-c) \sigma_1)
\]

\[
\leq c D_f^{\text{max}} (\rho_0 || \sigma_0) + (1-c) D_f^{\text{max}} (\rho_1 || \sigma_1),
\]

(10)

and

\[
D_f^{\text{max}} (\rho||X) \leq D_f^{\text{max}} (\rho||\sigma),
\]

(11)

if \(X \geq \sigma\).

Proof.

\[
c D_f^{\text{max}} (\rho_0 || \sigma_0) + (1-c) D_f^{\text{max}} (\rho_1 || \sigma_1)
\]

\[
= \inf \{ c D_f (\rho_0 || q_0) + (1-c) D_f (p_1 || q_1) : (\Gamma_0, \{p_0, q_0\}), (\Gamma_1, \{p_1, q_1\}) \}
\]

\[
\geq \inf \{ D_f (c\rho_0 + (1-c) \rho_1 || c\sigma_0 + (1-c) \sigma_1) : (\Gamma_0, \{p_0, q_0\}), (\Gamma_1, \{p_1, q_1\}) \},
\]

where \((\Gamma_0, \{p_0, q_0\})\) and \((\Gamma_1, \{p_1, q_1\})\) moves over all the reverse tests of \(\{\rho_0, \sigma_0\}\) and \(\{\rho_1, \sigma_1\}\), respectively.
Mixing \((\Gamma_0, \{p_0, q_0\})\) and \((\Gamma_1, \{p_1, q_1\})\) with probability \(c\) and \(1-c\), one can compose a reverse test \((\Gamma, \{\rho, \sigma\})\) of \(\{cp_0 + (1-c)\rho_1, c\sigma_0 + (1-c)\sigma_1\}\): let 
\[
\mathcal{X} = \{0,1\} \times \mathcal{X}, \quad \mathcal{F}(0, x) := cp_0, \quad \mathcal{F}(1, x) := (1-c)\rho_1, \quad \mathcal{G}(0, x) := cq_0, \quad \mathcal{G}(1, x) := (1-c)q_1.
\]
Then,
\[
D_f(\mathcal{F}, \mathcal{G}) = D_f(cp_0 + (1-c)\rho_1 || cq_0 + (1-c)q_1).
\]
Therefore, minimizing over all the reverse tests of \(\{cp_0 + (1-c)\rho_1, c\sigma_0 + (1-c)\sigma_1\}\), we obtain (10).

Let \(X' := X - \sigma \geq 0\). Then
\[
D_f^\max(\rho||X) = D_f^\max(\rho||\sigma + X')
\]
\[
\leq \inf \{D_f(\rho||q + r) : \Gamma(p) = \rho, \Gamma(q) = \sigma, \Gamma(r) = X, \supp r \text{ is disjoint with } \supp q, \supp p\}
\]
\[
= \inf \{D_f(\rho||q) : \Gamma(p) = \rho, \Gamma(q) = \sigma, \Gamma(r) = X, \supp r \text{ is disjoint with } \supp q, \supp p\}
\]
\[
= D_f^\max(\rho||\sigma),
\]
where the identity in the third line is due to the following calculation: let \(\mathcal{X}_1 = \supp q \cup \supp p\), and \(\mathcal{X}_2 = \supp r\). Then,
\[
\sum_{x \in \mathcal{X}_1 \cup \mathcal{X}_2} q(x) + r(x)) f \left( \frac{p(x)}{q(x) + r(x)} \right)
\]
\[
= \sum_{x \in \mathcal{X}_1} q(x) f \left( \frac{p(x)}{q(x)} \right) + \sum_{x \in \mathcal{X}_2} r(x) f \left( \frac{0}{r(x)} \right).
\]
Since \(f(0) = 0\) by the condition (FC), we have the identity. \(\blacksquare\)

Suppose (F) and \(\pi_{\sigma} \geq \pi_{\rho}\) is satisfied. Then, we can prove (10) and (11) using the identity (9). In other words, we have
\[
D_f'(cp_0 + (1-c)\rho_1 || c\sigma_0 + (1-c)\sigma_1)
\]
\[
\leq c D_f'(\rho_0 || \sigma_0) + (1-c) D_f'(\rho_1 || \sigma_1), \quad (12)
\]
\[
D_f'(\rho||\sigma) \leq D_f'(\rho || X), \quad X \geq \sigma. \quad (13)
\]
(12) is due to the known inequality
\[
g(cp_0 + (1-c)\rho_1, c\sigma_0 + (1-c)\sigma_1)
\]
\[
\leq c g(\rho_0, \sigma_0) + (1-c) g(\rho_1, \sigma_1),
\]
where
\[
g(\rho, \sigma) := \sqrt{f\left(\sigma^{-1/2} \rho \sigma^{-1/2}\right)} \sqrt{\sigma}.
\]
(13) is proved as follows. Since \(X \geq \sigma\), there is an operator \(C\) such that
\[
CX^{1/2} = \sigma^{1/2}, \quad ||C|| \leq 1.
\]
Since \(\pi_X \geq \pi_{\sigma}\), it follows that
\[
C = \sigma^{1/2} X^{-1/2}.
\]
where $X^{-1/2}$ stands for the generalized inverse of $X^{1/2}$. Since $\pi_X \geq \pi_\sigma \geq \pi_\rho$,

$$D_f' (\rho||\sigma) = \tr XCX^\dagger f(d(\rho,\sigma)) = \tr X C^\dagger f\left((d(\rho,\sigma))C\right) \geq \tr X f\left(C^\dagger(d(\rho,\sigma))C\right) = \tr X f\left(X^{-1/2}\rho X^{-1/2}\right) = D_f' (\rho||X),$$

where the inequality in the second line is due to Proposition 17.

6 Examples

Throughout this section, we suppose $\pi_\sigma \geq \pi_\rho$.

For an operator monotone function $g$ from $[0, \infty)$ to itself, generalized fidelity $F_g^{\min}$ is defined by

$$F_g^{\min} (\rho,\sigma) := \tr f(\rho f(d(\sigma,\rho))).$$

By Proposition 19,

$$D_{f_{\alpha}}^{\max} (\rho||\sigma) = F_{g_{\alpha}}^{\min} (\sigma,\rho),$$

with $f(y) := g(0) - g(y)$. As an example,

$$g_\alpha (y) = y^\alpha, \quad f_\alpha (y) = -y^\alpha \quad (0 \leq \alpha \leq 1).$$

The corresponding $D_{f_{\alpha}}^{\max}$ is

$$D_{f_{\alpha}}^{\max} (\rho||\sigma) = -\tr \sigma (\sigma^{-1}\rho)^\alpha.$$

If $\pi_\sigma = \pi_\rho$, we have

$$D_{f_{\alpha}}^{\max} (\rho||\sigma) = -\tr \rho (\rho^{-1}\sigma) (\rho^{-1}\sigma)^{-\alpha} = D_{f_{1-\alpha}}^{\max} (\rho||\sigma).$$

Also, with

$$f(y) := y \log y,$$

$$D_{f_{\alpha}}^{\max} (\rho||\sigma) = \tr \sigma \sigma^{-1}\rho (\log \sigma^{-1}\rho) = \tr \rho \log \sigma^{-1}\rho = \tr \rho \log \rho^{-1/2}\sigma^{-1}\rho^{1/2},$$

where the last end equals to $D^{R} (\rho||\sigma)$, the largest quantum relative entropy.

With

$$f_\alpha (y) = y^\alpha, \quad (\alpha > 1),$$

we have

$$D_{f_{\alpha}}^{\max} (\rho||\sigma) = \tr \sigma (\sigma^{-1}\rho)^\alpha.$$

In particular,

$$D_{f_{2}}^{\max} (\rho||\sigma) = \tr \sigma^{-1}\rho^{2}. $$
7 Relation to RLD Fisher metric

Suppose \( f \) is three times differentiable, and
\[
|f'''(x)| < c, \quad \forall x \in (1 - \varepsilon, 1 + \varepsilon).
\]

If \( \rho > 0 \), we have, for any Hermitian operators \( X, Y \),
\[
\frac{d^2}{dt \, ds} \left. D_f' (\rho + sX) \right|_{t=0, s=0} = \frac{d^2}{dt \, ds} \left. D_f' (\rho + tY) \right|_{t=0, s=0}
\]
\[
= \frac{d^2}{dt \, ds} \left. D_f'' (\rho + sX) \right|_{t=0, s=0}
\]
\[
= f''(1) \Re J^R_\rho (X, Y),
\]
where \( J^R_\rho \) is the RLD Fisher metric,
\[
J^R_\rho (X, Y) := \text{tr} X \rho^{-1} Y.
\]

(The result obviously generalizes to \( \rho \geq 0 \) case, if we limit \( X \) and \( Y \) to the Hermitian operators living in the support of \( \rho \).)

The derivation is as follows. Suppose \( \sigma > 0 \). Then,
\[
\text{tr} \, \sigma \, f (\sigma^{-1} \rho)
\]
\[
= \text{tr} \, \sigma \left\{ f (1_A) + f' (1) (\sigma^{-1} \rho - 1A) + \frac{1}{2} f'' (1) (\sigma^{-1} \rho - 1A)^2 + Z \right\}
\]
\[
= f (1) + \frac{1}{2} f'' (1) \text{tr} \, \sigma (\sigma^{-1} \rho - 1A)^2 + \text{tr} \, \sigma Z,
\]
where \( Z \) satisfies
\[
\frac{\|Z\|}{\|\sigma^{-1} \rho - 1A\|} \to 0, \text{ if } \sigma \to \rho.
\]

Therefore,
\[
\text{tr} \, (\rho - tY) \left( (\rho - tY)^{-1} (\rho + sX) \right)
\]
\[
= f (1) + \frac{1}{2} f'' (1) \text{tr} \, \rho \left\{ (\rho - tY)^{-1} (\rho + sX) - 1A \right\}^2 + o (\text{max} \{t, s\})^2
\]
\[
= f (1) + \frac{1}{2} f'' (1) \text{tr} \, \rho \left\{ t \rho^{-1} Y + s\rho^{-1} X \right\}^2 + o (\text{max} \{t, s\})^2
\]
\[
= f (1) + \frac{1}{2} f''(1) \left\{ s^2 J_\rho (X, X) + t^2 J^R_\rho (Y, Y) + 2ts J^R_\rho (X, Y) \right\} + o (\text{max} \{t, s\})^2
\]
where to derive the second identity we used
\[
(\rho + W)^{-1} = \rho^{-1} - t \rho^{-1} W \rho^{-1} + o (\|W\|).
\]
Also,
\[
\text{tr} \left( \rho + sX + tY \right) \left( (\rho + sX + tY)^{-1} \rho \right) \\
= f(1) + \frac{1}{2} f''(1) \text{tr} \rho \left\{ (\rho + sX + tY)^{-1} \rho - 1_A \right\}^2 + o(\max \{t, s\})^2 \\
= f(1) + \frac{1}{2} f''(1) \text{tr} \rho \left\{ -t \rho^{-1}Y - s\rho^{-1}X \right\}^2 + o(\max \{t, s\})^2 \\
= f(1) + \frac{1}{2} f''(1) \left\{ s^2 J_\rho(X, X) + t^2 J_\rho^R(Y, Y) + 2 ts J_\rho^R(X, Y) \right\} + o(\max \{t, s\})^2
\]
and
\[
\text{tr} \rho f \left( \rho^{-1} (\rho + sX + tY) \right) \\
= f(1) + \frac{1}{2} f''(1) \text{tr} \rho \left\{ \rho^{-1} (\rho + sX + tY) - 1_A \right\}^2 + o(\max \{t, s\})^2 \\
= f(1) + \frac{1}{2} f''(1) \text{tr} \rho \left\{ t \rho^{-1}Y + s\rho^{-1}X \right\}^2 + o(\max \{t, s\})^2 \\
= f(1) + \frac{1}{2} f''(1) \left\{ s^2 J_\rho(X, X) + t^2 J_\rho^R(Y, Y) + 2 ts J_\rho^R(X, Y) \right\} + o(\max \{t, s\})^2.
\]

8 When \( \pi_\sigma \not\supset \pi_\rho \)

In this section, we study the relation between \( D^{\text{max}}_f (\rho||\sigma) = D'_f (\rho||\sigma) \) and \( \lim_{n \to \infty} D'_f (\rho||\sigma_n) \) in the case of \( \pi_\sigma \not\supset \pi_\rho \), where \( \sigma_n \) is any positive operator with
\[
\text{supp} \sigma_n \supset \text{supp} \rho, \text{supp} \sigma
\]
and
\[
\lim_{n \to \infty} \sigma_n = \sigma.
\]

In the case of \( \pi_\sigma \not\supset \pi_\rho \), the minimal reverse test \( (\Gamma, \{p, q\}) \) of \( \{\rho, \sigma\} \) is defined by
\[
\mathcal{X} := \tilde{X} \cup \{x_0\} \\
\Gamma (\delta_z) := \begin{cases} 
\frac{1}{\text{tr} \rho} (\rho - \tilde{\rho}) & x = x_0 \\
\frac{1}{\text{tr} \rho} \cdot \tilde{\Gamma} (\delta_z) & \text{otherwise}
\end{cases},
\]
\[
p(x) := \begin{cases} 
1 - \text{tr} \tilde{\rho} & x = x_0 \\
\tilde{p} (x) & \text{otherwise}
\end{cases},
\]
\[
q(x) := \begin{cases} 
0 & x = x_0 \\
\tilde{q} (x) & \text{otherwise}
\end{cases},
\]
where \( \left( \tilde{\Gamma}, \{\tilde{p}, \tilde{q}\} \right) \) is the minimal reverse test of the pair \( \{\tilde{\rho}, \sigma\} \) of positive operators, and \( \tilde{\rho} \) is defined by
\[
\rho_{11} := \pi_\sigma \rho \pi_\sigma, \rho_{12} := \pi_\sigma \rho \pi, \rho_{22} := \pi \rho \pi, \\
\tilde{\rho} := \rho_{11} - \rho_{12} \rho_{22}^{-1} \rho_{21},
\]

\(12\)
where $\rho_{22}^{-1}$ is the generalized inverse of $\rho_{22}$.

Observe, if $\rho$ is invertible,
\[
(\pi_{\sigma} \rho^{-1} \pi_{\sigma})^{-1} = \hat{\rho}.
\]

**Lemma 7** Suppose $\rho_1$ is supported on $\text{supp} \sigma$ and $\rho_1 \leq \rho$. Then,
\[
\hat{\rho} \geq \rho_1.
\]

**Proof.** Since $\rho_1 \leq \rho$,
\[
\begin{bmatrix}
\rho_{11} - \rho_1 & \rho_{12} \\
\rho_{21} & \rho_{22}
\end{bmatrix} \geq 0.
\]
Therefore, by Proposition 22 we should have
\[
\rho_{11} - \rho_1 \geq \rho_{12} \rho_{22}^{-1} \rho_{21},
\]
and we have the assertion. ■

**Lemma 8** Suppose $f$ satisfies (F) and that
\[
\lim_{y \to \infty} \frac{f(y)}{y} = \infty.
\] (17)

Suppose also $\pi_{\sigma} \not\geq \pi_{\rho}$. Then,
\[
D_f^{\max} (\rho || \sigma) = D_f' (\rho || \sigma) = \lim_{n \to \infty} D_f' (\rho || \sigma_n) = \infty,
\]
where $\sigma_n$ is a positive operator with (14) and $\lim_{n \to \infty} \sigma_n = \sigma$.

**Proof.** Since $f$ is convex, there is a finite number $\alpha$ satisfying
\[
f(y) \geq f(1) + \alpha (y - 1).
\]

Therefore,
\[
D_f^{\max} (\rho || \sigma)
= D_f' (\rho || \sigma)
\]
\[
= \inf_{\rho_1} \left\{ \text{tr} \ f \left( \frac{1 - \text{tr} \rho_1}{\varepsilon} \right) ; \ \rho \geq \rho_1 \geq 0, \ \pi_{\sigma} \geq \pi_{\rho_1} \right\}
\]
\[
\geq \inf_{\rho_1} \left\{ f(1) \text{tr} \sigma + \alpha \text{tr} (d(\rho_1, \sigma) - 1) + \lim_{\varepsilon \downarrow 0} \varepsilon f \left( \frac{1 - \text{tr} \rho_1}{\varepsilon} \right) ; \ \rho \geq \rho_1 \geq 0, \ \pi_{\sigma} \geq \pi_{\rho_1} \right\}
\]
\[
= \inf_{\rho_1} \left\{ f(1) \text{tr} \sigma + \alpha \text{tr} \rho_1 - \sigma + \lim_{\varepsilon \downarrow 0} \varepsilon f \left( \frac{1 - \text{tr} \rho_1}{\varepsilon} \right) ; \ \rho \geq \rho_1 \geq 0, \ \pi_{\sigma} \geq \pi_{\rho_1} \right\}
\]
\[
\geq f(1) - |\alpha| + \inf_{\rho_1} \left\{ \lim_{\varepsilon \downarrow 0} \varepsilon f \left( \frac{1 - \text{tr} \rho_1}{\varepsilon} \right) ; \ \rho \geq \rho_1 \geq 0, \ \pi_{\sigma} \geq \pi_{\rho_1} \right\}.
\]
The last end of this is ∞, because by Lemma 7 we have

\[ 1 - \text{tr} \rho_1 = \text{tr} (\rho - \rho_1) \geq \text{tr} (\rho - \tilde{\rho}) = \text{tr} \begin{bmatrix} \rho_{12} \rho_{22}^{-1} \rho_{21} & \rho_{12} \\ \rho_{21} & \rho_{22} \end{bmatrix} > 0. \]

Thus,

\[ D_{\max}^f (\rho||\sigma) = D'_{f} (\rho||\sigma) = \infty. \]

Then, with \( c_n \) be the largest eigenvalue of \( \sigma_n - \sigma \), we have

\[ \sigma_n \leq \sigma + c_n I_A. \]

Also, let \( \pi \) be the smallest projector such that \( \pi \sigma + \pi \geq \pi \rho \). Then \( \pi \pi = 0 \).

Thus, by (11),

\[ D'_{f} (\rho||\sigma_n) \geq D'_{f} (\rho||\sigma + c_n I_A) = D'_{f} (\rho||\sigma + c_n (\pi \sigma + \pi)) \]

\[ = \text{tr} \left( \sigma + c_n (\pi \sigma + \pi) \right) f (d (\rho, \sigma + c_n (\pi \sigma + \pi))) + c_n \text{tr} \pi f (d (\rho, \sigma + c_n (\pi \sigma + \pi))) \]

\[ \geq (i) \text{tr} \left( \sigma + c_n \pi \sigma \right) f (d (\rho, \pi \sigma + c_n (\pi \sigma + \pi))) + c_n \text{tr} \pi f (d (\rho, \pi \sigma + c_n (\pi \sigma + \pi))) \]

\[ \geq (ii) \text{tr} \sigma f (d (\pi \rho \pi \sigma, \sigma)) + c_n \text{tr} \pi f (d (\pi \rho \pi, \pi)) \]

Here, the inequality (i) and (ii) is by Proposition 17 and (11). The second term of the last end goes to ∞ as \( n \to \infty \) by (17). Therefore, we have the assertion.

\[ \text{Lemma 9} \]

Suppose \( (F) \) is satisfied. Let \( \{ \sigma_n; n \in \mathbb{N} \} \) be a sequence of positive definite operators such that (14) and \( \lim_{n \to \infty} \sigma_n = \sigma \) are satisfied. Then,

\[ \lim_{n \to \infty} D'_{f} (\rho||\sigma_n) = \lim_{\varepsilon \downarrow 0} D'_{f} (\rho||\sigma + \varepsilon I_A), \]

(18)

where limits in both ends may be ∞.

**Proof.** By Lemma 8 the assertion is true if \( \lim_{x \to \infty} \frac{f(y)}{y} = \infty \). Hence, in what follows, we suppose

\[ \lim_{y \to \infty} \frac{f(y)}{y} < \infty \]

(19)
is true. Therefore, by Proposition 20, we have

$$f(y) = ay + \int_{(0, \infty)} \psi_t(y) \, d\mu(t), \quad (20)$$

where $a$ is a finite number, $\psi_t(y) = -\frac{y}{y+t}$, and $\mu$ is a non-negative finite measure with

$$\int_{(0, \infty)} \frac{1}{1+t} \, d\mu(t) < \infty. \quad (21)$$

By (13), $D_f'(\rho|\sigma + \varepsilon 1_A)$ is monotone non-increasing in $\varepsilon$. Therefore, the right hand side of (18) exists in $\mathbb{R} \cup \{\infty\}$. Let $c_n := ||\sigma_n - \sigma||$. Since $\sigma_n - \sigma \leq ||\sigma_n - \sigma|| 1_A$ by (13),

$$D_f'(\rho|\sigma_n) \geq D_f'(\rho|\sigma + \varepsilon 1_A).$$

Thus taking the limit of both ends, we have

$$\lim_{n \to \infty} D_f'(\rho|\sigma_n) \geq \lim_{\varepsilon \downarrow 0} D_f'(\rho|\sigma + \varepsilon 1_A).$$

If $n$ is large enough, $\sigma_n$ is arbitrarily close to $\sigma + c_n \pi_{\sigma_n}$, where $c_n > 0$ and $\lim_{n \to \infty} c_n = 0$. Also, $\text{supp} \sigma_n = \text{supp} (\sigma + c_n \pi_{\sigma_n})$. Therefore, for any $\delta$ such that $0 < \delta \leq 1$,

$$(\sigma + c_n \pi_{\sigma_n})^{-1/2} \sigma_n (\sigma + c_n \pi_{\sigma_n})^{-1/2} \geq (1 - \delta) \pi_{\sigma_n}, \; \exists n_0 \forall n \geq n_0,$$

or equivalently,

$$\sigma_n \geq (1 - \delta) (\sigma + c_n \pi_{\sigma_n}), \; \exists n_0 \forall n \geq n_0.$$

By (13),

$$D_f'(\rho|\sigma_n) \leq D_f'(\rho|\sigma + \varepsilon 1_A) \leq D_f'(\rho|\sigma + (1 - \delta) \pi_{\sigma_n})$$

$$= D_f'(\rho|\sigma + (1 - \delta) (\sigma + c_n 1_A))$$

$$= (1 - \delta) \text{tr} (\sigma + c_n 1_A) f \left( \frac{1}{1 - \delta} d (\sigma, \sigma + c_n 1_A) \right).$$

Therefore, using (20) and the fact that $x \to \psi_t(x)$ is operator monotone non-increasing,

$$D_f'(\rho|\sigma_n) \leq a \text{tr} \rho + (1 - \delta) \int_{(0, \infty)} \text{tr} (\sigma + c_n 1_A) \psi_t \left( \frac{1}{1 - \delta} d (\rho, \sigma + c_n 1_A) \right) \, d\mu(t)$$

$$\leq a + (1 - \delta) \int_{(0, \infty)} \text{tr} (\sigma + c_n 1_A) \psi_t (d (\rho, \sigma + c_n 1_A)) \, d\mu(t)$$

$$= a\delta + (1 - \delta) D_f'(\rho|\sigma + c_n 1_A).$$

Letting $n \to \infty$, we have

$$\lim_{n \to \infty} D_f'(\rho|\sigma_n) \leq a\delta + (1 - \delta) \lim_{\varepsilon \downarrow 0} D_f'(\rho|\sigma + \varepsilon 1_A).$$

15
Since $\delta > 0$ is arbitrary, we have
\[
\lim_{n \to \infty} D_{f} (\rho||\sigma_n) \leq \lim_{\varepsilon \downarrow 0} D_{f} (\rho||\sigma + \varepsilon 1_A),
\]
and thus the assertion is proved. \qed

**Theorem 10** Suppose $f$ satisfies (F). Then,
\[
\lim_{n \to \infty} D_{f} (\rho||\sigma_n) = D_{f}^{\max} (\rho||\sigma) = D_{f}^{\prime} (\rho ||\sigma) + (1 - \text{tr} \bar{\rho}) \lim_{y \to \infty} \frac{f(y)}{y}, \quad (22)
\]
where $\{\sigma_n; n \in \mathbb{N}\}$ is a sequence of positive definite operators such that (14) and $\lim_{n \to \infty} \sigma_n = \sigma$ are satisfied. Also, the minimal reverse test (15) achieves
\[
D_{f}^{\max} (\rho||\sigma) = D_{f} (p||q).
\]

**Proof.** By Lemma 8, the assertion is true if $\lim_{y \to \infty} \frac{f(y)}{y} = \infty$. Hence, in what follows, we suppose (19) holds true. This means $f$ is in the form of (20).

Let
\[
\sigma_{\varepsilon} := \sigma + \varepsilon \pi,
\]
where $\pi$ is the smallest projector such that $\pi \rho \leq \pi \sigma + \pi$ and $\pi \sigma \pi = 0$. Then $d (\rho, \sigma_{\varepsilon})$ is well-defined and supported on $\text{supp} \sigma_{\varepsilon}$. Observe that
\[
0 \leq -\psi_{t} (y) \leq 1
\]
for any $t > 0, y \geq 0$. Therefore,
\[
0 \leq -\psi_{t} (d (\rho, \sigma_{\varepsilon})) \leq \pi \sigma_{\varepsilon}.
\]
Therefore, when $\varepsilon$ is small,
\[
0 \leq -\text{tr} \sigma_{\varepsilon} \psi_{t} (d (\rho, \sigma_{\varepsilon})) \leq -\text{tr} \sigma_{\varepsilon} \leq 2.
\]
Hence, by bounded convergence theorem,
\[
- \lim_{\varepsilon \downarrow 0} \int \text{tr} \sigma_{t} \psi_{t} (d (\rho, \sigma_{\varepsilon})) \, d\mu (t) = - \int \left\{ \lim_{\varepsilon \downarrow 0} \text{tr} \sigma_{t} \psi_{t} (d (\rho, \sigma_{\varepsilon})) \right\} \, d\mu (t).
\]
Observe that
\[
- \text{tr} \sigma_{t} \psi_{t} (d (\rho, \sigma_{\varepsilon})) = \text{tr} \sigma_{t} \rho (\rho + t \sigma_{\varepsilon})^{-1}
\]
\[
= \text{tr} \left[ \begin{array}{cc}
\sigma_{\rho_{11}} & \sigma_{\rho_{12}} \\
\varepsilon_{\rho_{21}} & \varepsilon_{\rho_{22}}
\end{array} \right] \left[ \begin{array}{cc}
Z_{11} & Z_{12} \\
Z_{21} & Z_{22}
\end{array} \right],
\]
\[
16
where
\[ Z_{11} := \left( \rho_{11} + t\sigma - \rho_{12} (\rho_{22} + \varepsilon t\pi)^{-1} \rho_{21} \right)^{-1}, \]
\[ Z_{21} := - (\rho_{22} + \varepsilon t\pi)^{-1} \rho_{21} Z_{11}, \]
\[ Z_{12} := - Z_{11} \rho_{12} (\rho_{22} + \varepsilon t\pi)^{-1}, \]
\[ Z_{22} := (\rho_{22} + \varepsilon t\pi)^{-1} + (\rho_{22} + \varepsilon t\pi)^{-1} \rho_{21} Z_{11} \rho_{12} (\rho_{22} + \varepsilon t\pi)^{-1}, \]

Therefore,
\[
- \operatorname{tr} \sigma \psi_t (d (\rho, \sigma)) = \operatorname{tr} \left( \rho_{11} - \rho_{12} (\rho_{22} + \varepsilon t\pi) \right)^{-1} \rho_{21} \left( \rho_{11} + t\sigma - \rho_{12} (\rho_{22} + \varepsilon t\pi)^{-1} \rho_{21} \right)^{-1} \sigma \\
+ \varepsilon \operatorname{tr} (\rho_{21} Z_{12} + \rho_{22} Z_{22}).
\]

Since \( \rho + t\sigma \) is non-singular for any \( t > 0 \) and \( \varepsilon \geq 0 \), so is \( \rho_{11} + t\sigma - \rho_{12} (\rho_{22} + \varepsilon t\pi)^{-1} \rho_{21} \). Therefore, for any \( \varepsilon \geq 0 \), \( \operatorname{tr} (\rho_{21} Z_{12} + \rho_{22} Z_{22}) \) is bounded. Therefore,
\[
- \lim_{\varepsilon \downarrow 0} \operatorname{tr} \sigma \psi_t (d (\rho, \sigma)) = \operatorname{tr} \left( \rho_{11} - \rho_{12} \rho_{21} \right)^{-1} \left( \rho_{11} + t\sigma - \rho_{12} \rho_{21} \right)^{-1} \sigma \\
= \operatorname{tr} \tilde{\rho} (\sigma^{-1} \tilde{\rho} + t)^{-1} = \operatorname{tr} \sigma^{-1/2} \tilde{\rho} \sigma^{-1/2} \left( \sigma^{-1/2} \tilde{\rho} \sigma^{-1/2} + t \right)^{-1} \\
= - \operatorname{tr} \sigma \psi_t (d (\tilde{\rho}, \sigma)).
\]

Therefore, by (30),
\[
\lim_{\varepsilon \downarrow 0} D'_{\tilde{\rho}} (\rho || \sigma) = a + \int D'_{\psi_t} (\rho || \sigma) \, d\mu (t),
\]
which, combined with Lemma 9 leads to
\[
\lim_{n \to \infty} D'_{\rho} (\rho || \sigma_n) = a + \int D'_{\psi_t} (\rho || \sigma) \, d\mu (t).
\]

On the other hand, let \( \rho_1 \) be an arbitrary positive operator supported on \( \text{supp} \sigma \) such that \( \rho \geq \rho_1 \). Then, by (21),
\[
D'_{\rho} (\rho_1 || \sigma) + \lim_{\varepsilon \downarrow 0} \varepsilon f \left( \frac{1 - \operatorname{tr} \rho_1}{\varepsilon} \right) = a \operatorname{tr} \rho_1 + \int (0, \infty) D'_{\psi_t} (\rho_1 || \sigma) \, d\mu (t) \\
+ \lim_{\varepsilon \downarrow 0} \varepsilon \left( a \frac{1 - \operatorname{tr} \rho_1}{\varepsilon} + \int (0, \infty) \psi_t \left( \frac{1 - \operatorname{tr} \rho_1}{\varepsilon} \right) \, d\mu (t) \right) \\
= a + \int D'_{\psi_t} (\rho_1 || \sigma) \, d\mu (t) - \lim_{\varepsilon \downarrow 0} \int (0, \infty) \frac{1 - \operatorname{tr} \rho_1}{\varepsilon} \, d\mu (t).
Since the integrand \( \frac{1}{y} \) is monotone decreasing in \( y \) and is integrable with respect to \( \mu \) (see (21)), by monotone convergence theorem we have

\[
\lim_{\epsilon \downarrow 0} \int_{(0,\infty)} \frac{1}{\epsilon} \frac{1}{1 - \Tr \rho_1} + t \, d\mu(t) = \lim_{y \to \infty} \int_{(0,\infty)} \frac{1}{y + t} \, d\mu(t) = \int_{(0,\infty)} \lim_{y \to \infty} \frac{1}{y + t} \, d\mu(t) = 0.
\]

After all, by Lemma 7 and by the fact that \( \psi_t(x) \) is operator monotone decreasing (Proposition 16),

\[
D_{\text{max}} f(\rho||\sigma) = \sup_{\rho_1} \left\{ D'_f(\rho_1||\sigma) + \lim_{\epsilon \downarrow 0} \epsilon f \left( \frac{1 - \Tr \rho_1}{\epsilon} \right) \mid \text{supp } \rho_1 \subset \text{supp } \sigma, \rho_1 \leq \rho \right\}
\]

\[
= a + \inf_{\rho'} \left\{ \int D'_{\psi_t}(\rho_1||\sigma) \, d\mu(t) \mid \text{supp } \rho_1 \subset \text{supp } \sigma, \rho_1 \leq \rho \right\}
\]

\[
= a + \int D'_{\psi_t}(\tilde{\rho}||\sigma) \, d\mu(t)
\]

\[
= \lim_{n \to \infty} D'_{f}(\rho_n||\sigma_n). \tag{23}
\]

Thus, the first equality of (22) is proved. Also, the supremum in the first line is achieved by \( \rho_1 = \tilde{\rho} \), which corresponds to the minimal reverse test (19).

The second equality of (22) follows by (23) and the identity

\[
\lim_{y \to \infty} \frac{f(y)}{y} = a - \lim_{y \to \infty} \int_{(0,\infty)} \frac{d\mu(t)}{y + t} = a. \tag{24}
\]

For example, if

\[
\rho = |\varphi \rangle \langle \varphi|, \text{ supp } \sigma \neq \text{ supp } |\varphi \rangle \langle \varphi|,
\]

or

\[
\sigma = |\varphi \rangle \langle \varphi|, \text{ supp } \rho \neq \text{ supp } |\varphi \rangle \langle \varphi|,
\]

then \( \tilde{\rho} = 0 \). Therefore, if \( \lim_{y \to \infty} f(y)/y = 0 \) holds (e.g., \( f(y) = -y^\alpha \) \((0 < \alpha < 1)\)), we have

\[
D_{f_{\text{max}}}^\rho(\rho||\sigma) = 0.
\]

9 Implications of equality

A function \( f \) with (F), by Proposition 21, is written as

\[
f(y) = ay + by^2 + \int_{(0,\infty)} \left( \frac{y}{1 + \epsilon} + \psi_t(y) \right) \, d\mu(t), \tag{25}
\]
where \( a \) is a real number, \( b > 0 \), \( \mu \) is a non-negative Borel measure with
\[
\int_{(0, \infty)} \frac{d\mu(t)}{(1 + t)^2} < \infty,
\]
and \( \psi_t(y) = -\frac{y}{t + y} \).

In what follows, in the case that \( \text{supp} \rho \subset \text{supp} \sigma \), we let \( d := d(\rho, \sigma) \) and \( d' := d(\Lambda(\rho), \Lambda(\sigma)) \). Otherwise, we let
\[
d := d(\hat{\rho}, \sigma), \quad d' := d(\Lambda(\hat{\rho}), \Lambda(\sigma)) ,
\]
where \( \hat{\rho} \) is defined by (16). In either case, we suppose
\[
\text{spec} \, d \cup \text{spec} \, d' \subset \text{supp} \, \mu \cup \{0\}, \quad (26)
\]
where \( \text{supp} \, \mu \) is the smallest relatively closed subset of \((0, \infty)\) such that any open subset of \((0, \infty)/\text{supp} \, \mu\) is measure zero.

**Lemma 11** Suppose that \( t \in (0, \infty) \to g(t) \in [0, \infty) \) is continuous and that
\[
\int_{(0, \infty)} g(t) \, d\mu(t) = 0
\]
for a Borel measure \( \mu \). Then, if \( t_0 \in \text{supp} \, \mu \),
\[
g(t_0) = 0.
\]

**Proof.** The proof is much draws upon the proof of Theorem 2.1 of [11]. Let
\[
U_\mu := \bigcup \{U; \mu(U) = 0, \text{U is open subset of } (0, \infty) \}.
\]
Then, \( U_\mu \) is open set. Since \((0, \infty)\) is separable, there is a sequence \( \{U_n; n \in \mathbb{N}\} \) of open subset of \((0, \infty)\) such that \( \mu(U_n) = 0 \) and
\[
U_\mu = \bigcup_{n \in \mathbb{N}} U_n.
\]
Therefore,
\[
\mu(U_\mu) \leq \sum_{n \in \mathbb{N}} \mu(U_n) = 0.
\]
Therefore, \((0, \infty)/U_\mu\), which is relatively closed subset of \((0, \infty)\), contains \( \text{supp} \, \mu \).
On the other hand, if \( U \) is an open set with \( \mu(U) = 0, \text{U is by subset of } U_\mu \), by definition. Therefore, \((0, \infty)/U_\mu\) is contained in \( \text{supp} \, \mu \). Therefore,
\[
\text{supp} \, \mu = (0, \infty)/U_\mu.
\]
Suppose \( t_0 \in \text{supp} \, \mu \) and \( g(t_0) > 0 \). Then, by continuity of \( g \), there is an \( \varepsilon > 0 \) such that \( g(t) > 0 \) for any \( t \in [t_0 - \varepsilon, t_0 + \varepsilon] \). Therefore,
\[
\int_{(0, \infty)} g(t) \, d\mu(t) \geq \int_{[t_0 - \varepsilon, t_0 + \varepsilon]} g(t) \, d\mu(t)
\]
\[
\geq 2\varepsilon \cdot \min_{t \in [t_0 - \varepsilon, t_0 + \varepsilon]} g(t) > 0.
\]
This contradicts with \( \int_{(0, \infty)} g(t) \, d\mu(t) = 0 \). Therefore, we have \( g(t_0) = 0 \). \( \blacksquare \)
Lemma 12 Suppose \( f \) satisfies (F) and \( \{12\} \). Let \( \Lambda \) be a CPTP map from \( \mathcal{L}_A \) to \( \mathcal{L}_B \). Also, let \( (\Gamma, \{p, q\}) \) and \( (\Gamma', \{p', q'\}) \) be the minimal reverse test of \( \{\rho, \sigma\} \) and \( \{\Lambda (\rho), \Lambda (\sigma)\} \), respectively. Then,

\[
D_{\max}^f (\rho||\sigma) = D_{\max}^f (\Lambda (\rho)||\Lambda (\sigma)) < \infty, \tag{27}
\]

only if

\[
\Lambda_\sigma (h (d)) = h (\Lambda_\sigma (d)) \tag{28}
\]

holds. Here, \( \Lambda_\sigma \) is a subunital CP map defined by \( \{4\} \), and \( h \) is an arbitrary function whose domain contains the spectrum of \( d = d(\rho, \sigma) \) and \( \Lambda_\sigma (d) \).

Also, if in addition \( \sigma \) is invertible, \( \{28\} \) is sufficient for

\[
D_{\max}^f (\rho||\sigma) = D_{\max}^f (\Lambda (\rho)||\Lambda (\sigma)). \tag{29}
\]

Proof. First, suppose \( \text{supp} \rho \subset \text{supp} \sigma \) holds. Then, by \( \{7\}, \{27\} \) implies

\[
\Lambda \left( \sigma^{1/2} f (d) \sigma^{1/2} \right) = \sqrt{\Lambda (\sigma)} f (d') \sqrt{\Lambda (\sigma)}. \tag{30}
\]

Observe

\[
\sigma^{1/2} f (d) \sigma^{1/2} = a \rho + b \sigma^{1/2} f_2 (d) \sigma^{1/2} + \int_{(0, \infty)} \left( \frac{\rho}{1 + t} + \sigma^{1/2} \psi_t (d) \sigma^{1/2} \right) d \mu (t),
\]

where \( f_2 (y) := y^2 \). Since \( \psi_t (y) \) and \( f_2 (y) \) satisfies (F), by \( \{7\} \) and Lemma \( \{11\} \), \( \{29\} \) is possible only if

\[
\Lambda \left( \sigma^{1/2} \psi_t (d) \sigma^{1/2} \right) = \sqrt{\Lambda (\sigma)} \psi_t (d') \sqrt{\Lambda (\sigma)}, \quad \forall t \in \text{supp} \mu \cup \{0\}.
\]

This, by the assumption \( \{26\} \) and Proposition \( \{23\} \) implies

\[
\Lambda \left( \sigma^{1/2} h (d) \sigma^{1/2} \right) = \sqrt{\Lambda (\sigma)} h (d') \sqrt{\Lambda (\sigma)},
\]

which, using \( \{6\} \), implies \( \{28\} \).

Next, suppose \( \text{supp} \rho \not\subset \text{supp} \sigma \). Let \( \hat{\rho}' \) be the largest positive operator supported on \( \text{supp} \pi_{\Lambda (\sigma)} \) such that

\[
\hat{\rho}' \leq \Lambda (\rho).
\]

Since \( \Lambda (\hat{\rho}) \) is smaller than \( \Lambda (\rho) \) and supported on \( \pi_{\Lambda (\sigma)} \), we have

\[
\Lambda (\hat{\rho}) \leq \hat{\rho}'.
\]

Therefore, by \( \{7\} \), for each \( t \),

\[
\Lambda \left( \sigma^{1/2} \psi_t (d (\hat{\rho}, \sigma)) \sigma^{1/2} \right) \geq \{\Lambda (\sigma)\}^{1/2} \psi_t (d (\Lambda (\hat{\rho}), \Lambda (\sigma))) \{\Lambda (\sigma)\}^{1/2}
\]

\[
\geq \{\Lambda (\sigma)\}^{1/2} \psi_t (d (\hat{\rho}', \Lambda (\sigma))) \{\Lambda (\sigma)\}^{1/2},
\]

20
where the second inequality is by the fact that $\psi_t(y)$ is operator monotone decreasing and (\textit{??}). Therefore, by $D_{\text{max}}^\max(\rho||\sigma) < \infty$, Theorem 10 and (30), (27) holds only if

$$
\Lambda \left( \sigma^{1/2} \psi_t(d(\tilde{\rho}, \sigma)) \sigma^{1/2} \right) = \{ \Lambda (\sigma) \}^{1/2} \psi_t(d(\Lambda (\tilde{\rho}), \Lambda (\sigma))) \{ \Lambda (\sigma) \}^{1/2} 
= \{ \Lambda (\sigma) \}^{1/2} \psi_t(d(\Lambda (\tilde{\rho}), \Lambda (\sigma))) \{ \Lambda (\sigma) \}^{1/2}, \ \forall t \in \text{supp} \mu \cup \{0 \}.
$$

(31)

The first equality of (31) leads to (28), in the same way as the case of $\text{supp} \rho \subset \text{supp} \sigma$.

The second assertion of the theorem is proved by straightforward computation using (6).

**Theorem 13** Suppose $f$ satisfies (F) and (26). Let $\Lambda$ be a CPTP map from $\mathcal{L}_A$ to $\mathcal{L}_B$. Also, let $(\Gamma, \{ p, q \})$ and $(\Gamma', \{ p', q' \})$ be the minimal reverse test of $\{ \rho, \sigma \}$ and $\{ \Lambda (\rho), \Lambda (\sigma) \}$, respectively. Then, holds (27) if and only if

$$
\Lambda (\Gamma (\delta_x)) = \Gamma' (\delta_x),
$$

(32)

$$
p = p', q = q',
$$

(33)

$$
D_f (p||q) < \infty.
$$

(34)

**Proof.** Since ‘if’ is trivial, we prove ‘only if’.

First suppose $\text{supp} \rho \subset \text{supp} \sigma$ holds. Let $d := \sum_x d_x P_x$, where $d_x$ and $P_x$ is an eigenvalue and projection onto eigenspace, and let $P'_x := \Lambda (P_x)$. Applying (28) to

$$
h_0(y) := \begin{cases} 
1, & (y = d_x), \\
0, & \text{otherwise}.
\end{cases}
$$

we have

$$
P_x = \Lambda (P_x) = \Lambda (h_0 (d)) = h_0 (\Lambda (d)).
$$

Since

$$
(P'_x)^2 = \{ h_0 (\Lambda (d)) \}^2 = h_0 (\Lambda (d)) = P'_x,
$$

$P'_x$ is a projector. Since

$$
d' = \Lambda (d) = \sum_x d_x \Lambda (P_x) = \sum_x d_x P',
$$

$P'_x$s are the projectors onto the eigenspaces of $d'$, and $\text{spec} d = \text{spec} d'$. Therefore, we have

$$
\Lambda (\Gamma (\delta_x)) = \Lambda (\sqrt{\Lambda (\sigma) P_x \sqrt{\sigma}}) = \sqrt{\Lambda (\sigma) P_x \sqrt{\Lambda (\sigma)}}
= \sqrt{\Lambda (\sigma) P'_x \sqrt{\Lambda (\sigma)}} = \Gamma' (\delta_x).
$$

Therefore,

$$
\sum_x q(x) \Gamma' (\delta_x) = \Lambda \left( \sum_x q(x) \Gamma (\delta_x) \right) = \Lambda (\sigma).
$$
If any probability distribution \( q' \) satisfies

\[
\Gamma' (q') = \sum_x q' (x) \Gamma' (\delta_x) = \sum_x q (x) \Gamma' (\delta_x),
\]

we have to have

\[
\sum_x q' (x) \sqrt{\Lambda (\sigma)} P_x \sqrt{\Lambda (\sigma)} = \sum_x q (x) \sqrt{\Lambda (\sigma)} P_x \sqrt{\Lambda (\sigma)}.
\]

Since \( P_x \) is supported on \( \text{supp} d' = \text{supp} \Lambda (\sigma) \), this is equivalent to

\[
\sum_x q' (x) P_x = \sum_x q (x) P_x.
\]

Since \( P_x \) are orthogonal projectors, we have \( q' = q \). In the same way, we can prove that \( p' = p \).

Next, suppose \( \text{supp} \rho \not\subset \text{supp} \sigma \). Let \( \tilde{\rho}' \) be the largest operator supported on \( \text{supp} \Lambda (\sigma) \) such that \( \Lambda (\rho) \geq \tilde{\rho}' \). Then, since \( \Lambda (\rho) \) and \( \tilde{\rho}' \) are supported on \( \text{supp} \Lambda (\sigma) \), the second identity of (31) is true only if

\[
\Lambda (\tilde{\rho}) = \tilde{\rho}'. \tag{35}
\]

Let \( \left( \tilde{\Gamma}, \{\tilde{p}, \tilde{q}\} \right) \) and \( \left( \tilde{\Gamma}', \{\tilde{p}', \tilde{q}'\} \right) \) be the minimal reverse test for \( \{\rho, \sigma\} \) and \( \{\Lambda (\rho), \Lambda (\sigma)\} \). Then, by the first identity of (31) and the argument for the case of \( \text{supp} \rho \subset \text{supp} \sigma \), we have

\[
\Lambda \left( \tilde{\Gamma} (\delta_x) \right) = \tilde{\Gamma}' (\delta_x),
\]

\[
\tilde{p} = \tilde{p}', \tilde{q} = \tilde{q}'.
\]

Therefore, by (35) and the definition of the minimal reverse test, we have the asserted result.

Let \( \{A_i\} \) be the Kraus operators of \( \Lambda_\sigma \),

\[
\Lambda_\sigma (Z) = \sum_i A_i Z A_i^\dagger.
\]

Then, since \( \sum_x A_i P_x A_i^\dagger = P'_x \), \( A_i P_x A_i^\dagger \) is supported on \( \text{supp} P'_x \). Thus

\[
A_i = \bigoplus_x A_{x,i},
\]

where \( A_{x,i} \) is a linear map from \( \text{supp} P_x \) to \( \text{supp} P'_x \) and

\[
\sum_i A_{x,i} (A_{x,i})^\dagger = P'_x.
\]

Observe

\[
\Lambda_\sigma^\dagger (\Lambda (\sigma)) = \sigma^{1/2} \left\{ (\Lambda)^\dagger \left( \pi_{\Lambda (\sigma)} \right) \right\} \sigma^{1/2} = \sigma.
\]

22
With
\[ W_x := \sqrt{\sigma} P_x, \quad W'_x := \sqrt{\Lambda (\sigma)} P'_x, \]
this means
\[
\Lambda (A (\sigma)) = \bigoplus_{x_1, x_2} \sum_i (W'_{x_1, i} A_{x_1, i})^\dagger W_{x_2} A_{x_2, i}
= \sum_{x_1, x_2} P_{x_2} \sigma P_{x_1} = \sum_{x_1, x_2} P_{x_2} W'^\dagger_{x_2} W_{x_1} P_{x_1},
\]
or
\[
\sum_i (W'_{x_1, i} A_{x_1, i})^\dagger W_{x_2} A_{x_2, i} = W'^\dagger_{x_2} W_{x_1}. \quad (36)
\]
In particular, if each eigenspace of \(d\) and \(d'\) are one-dimensional, \(a_{x, i} := A_{x, i}\) are the scalar with
\[
\sum_i |a_{x, i}|^2 = 1,
\]
\(|w_x) := W_x\) and \(|w'_x) := W'_x\) are column vectors. Let
\[
|a_x) := \sum_i a_{x, i} |e_i),
\]
where \(\{|e_i)\}_{i}\) is orthonormal system of vectors. Then, (36) in this case can be written as
\[
\langle a_{x_1} | a_{x_2} \rangle \langle w'_{x_1} | w'_{x_2} \rangle = \langle w_{x_1} | w_{x_2} \rangle.
\]
This is necessary and sufficient condition for the existence of of CPTP map which maps \(|w_x)\) to \(|w'_x)\).

**Corollary 14** Let \((\Gamma, \{p, q\})\) be a reverse test with
\[
D^{max}_f (\rho||\sigma) = D_f (p||q) < \infty \quad (37)
\]
where \(f\) is a function with (F) and
\[
\spec d (\rho, \sigma) \cup \left\{ \frac{p(x)}{q(x)} : x \in \mathcal{X} \right\} \subset \supp \mu \cup \{0\}.
\]
Here \(\mu\) is defined by (26). Then (37) holds for any \(f\) with (F).

**Proof.** Let \((\Gamma', \{p', q'\})\) and \((\Gamma'', \{p'', q''\})\) be the minimal reverse test of \(\{p, q\}\) and \(\{\rho, \sigma\}\), respectively. Then by Theorem 13
\[
p' = p'', q' = q''.
\]
By Theorem 10 for any function \(f_0\) with (F),
\[
D^{max}_{f_0} (p||q) = D^{max}_{f_0} (p'||q') = D^{max}_{f_0} (p''||q'') = D^{max}_{f_0} (\rho||\sigma).
\]
Also, by Lemma 4 \(D_f (p||q) = D^{max}_f (p||q)\). Therefore, we have the assertion. □
Corollary 15 Suppose that there is a measurement $M$ taking values on the finite set $\mathcal{Z}$ such that

$$D_{\max}^{\rho}(\rho||\sigma) = D_f(P^M_\rho || P^M_\sigma)$$

and that $\text{supp} \mu$ contains $(0, \infty)$. Then $\rho$ and $\sigma$ commute with each other.

Proof. Let $\{P^M_\rho, P^M_\sigma\}$ be probability distributions over the finite set $\mathcal{Z}$. Let $(\Gamma, \{p, q\})$ and $(\Gamma', \{p', q'\})$ be the minimal reverse test of $\{\rho, \sigma\}$ and $\{P^M_\rho, P^M_\sigma\}$, respectively. Then by Theorem 13, we have

$$P^M_{\Gamma'(\delta_x)} = \Gamma'(\delta_x),$$

where $x$s are the members of the set

$$\mathcal{X} = \tilde{\mathcal{X}} \cup \{x_0\}.$$

By composition of the minimal reverse test, $\Gamma'(\delta_x)$ is the probability distribution

$$\Gamma'(\delta_x)(z) = \begin{cases} 
\frac{P^M_\sigma(z)}{\sum_{z \in \mathcal{Z}_x} P^M_\sigma(z)}, & (x \in \mathcal{X}), \\
\frac{P^M_\rho(z)}{\sum_{z \in \mathcal{Z}_0} P^M_\rho(z)}, & (x = x_0),
\end{cases}$$

where

$$\mathcal{Z}_x := \left\{ z : \frac{P^M_\rho(z)}{P^M_\sigma(z)} = d'_x, z \in \text{supp} P^M_\sigma \right\},$$

$$\mathcal{Z}_0 := \text{supp} P^M_\rho \setminus \text{supp} P^M_\sigma.$$

(Also,

$$p'(x) = \begin{cases} 
\sum_{z \in \mathcal{Z}_x} P^M_\rho(z), & (x \in \mathcal{X}), \\
\sum_{z \in \mathcal{Z}_0} P^M_\rho(z), & (x = x_0),
\end{cases}$$

$$q'(x) = \begin{cases} 
\sum_{z \in \mathcal{Z}_x} P^M_\sigma(z), & (x \in \mathcal{X}), \\
0, & (x = x_0).
\end{cases}$$

) Therefore, if $x_1 \neq x_2$

$$\text{supp} \Gamma'(\delta_{x_1}) \cap \text{supp} \Gamma'(\delta_{x_2}) = \emptyset.$$

Therefore, by (38),

$$\text{supp} \Gamma(\delta_{x_1}) \cap \text{supp} \Gamma(\delta_{x_2}) = \emptyset.$$
9.1 Other implications

(28) means $h(d(\rho, \sigma))$ is a multiplicative domain of $\Lambda_\sigma$, thus

$$\Lambda_\sigma(Z h(d)) = \Lambda_\sigma(Z) \Lambda_\sigma(h(d)),$$
$$\Lambda_\sigma(h(d)Z) = \Lambda_\sigma(h(d)) \Lambda_\sigma(Z)$$

holds for any $Z$.

Below, $\langle \cdot, \cdot \rangle$ and $\| \cdot \|_{HS}$ denotes Hilbert-Schmidt inner product, $\langle Z_1, Z_2 \rangle = \text{tr}(Z_1^\dagger Z_2)$, and the norm based on it, respectively. Define for a given $\sigma \in \mathcal{L}_A$ and $\Lambda : \mathcal{L}_A \to \mathcal{L}_B$, the operator $V_{\sigma, \Lambda} : \mathcal{L}_B \to \mathcal{L}_A$ by

$$V_{\sigma, \Lambda}(Z) := \Lambda^\dagger \left(Z \{ \Lambda(\sigma) \}^{-1/2} \right)^{1/2}.$$

(Below, so far as no confusion is likely, we suppress the subscripts $\sigma, \Lambda$, and simply write $V(Z)$. ) Then,

$$V^\dagger(Z) = \Lambda \left(Z \sigma^{1/2} \right) \Lambda(\sigma)^{-1/2},$$

and

$$\text{tr} \Lambda(\sigma) = \text{tr} \sigma \Leftrightarrow V \left(\Lambda(\sigma)^{1/2} \right) = \sigma^{1/2}.$$

(The former and the latter is (4.2) and Lemma 4.1 of [7], respectively. ) Also, $V$ is a contraction with respect to Hilbert-Schmidt norm (Lemma 4.2 of [7]),

$$\|V(Z)\|_{HS} \leq \|Z\|_{HS}. \quad (41)$$

Using $V$, (28) can be written as

$$\Lambda(\sigma)^{1/2} h(d') = V^\dagger \sigma^{1/2} h(d). \quad (42)$$

Using (42) for $h(x)$ and $h(x)^2$,

$$\left\| V^\dagger \sigma^{1/2} h(d) \right\|_{HS}^2 = \left\langle \Lambda(\sigma)^{1/2} h(d'), \Lambda(\sigma)^{1/2} h(d') \right\rangle$$
$$= \left\langle \Lambda(\sigma)^{1/2}, \Lambda(\sigma)^{1/2} h(d')^2 \right\rangle$$
$$= \left\langle \Lambda(\sigma)^{1/2}, V^\dagger \sigma^{1/2} h(d)^2 \right\rangle$$
$$= \left\langle V \Lambda(\sigma)^{1/2}, \sigma^{1/2} h(d)^2 \right\rangle$$
$$= (i) \left\langle \sigma^{1/2}, \sigma^{1/2} h(d)^2 \right\rangle$$
$$= \left\| \sigma^{1/2} h(d) \right\|_{HS}^2,$$

where (40) is used to show the equality (i). Therefore,

$$0 \leq \left\| V V^\dagger \sigma^{1/2} h(d) - \sigma^{1/2} h(d) \right\|_{HS}^2$$
$$= \left\| V V^\dagger \sigma^{1/2} h(d) \right\|_{HS}^2 + \left\| \sigma^{1/2} h(d) \right\|_{HS}^2 - 2 \left\| V^\dagger \sigma^{1/2} h(d) \right\|_{HS}^2$$
$$= \left\| V V^\dagger \sigma^{1/2} h(d) \right\|_{HS}^2 - \left\| \sigma^{1/2} h(d) \right\|_{HS}^2 \leq 0.$$
where the last inequality is (41). Therefore, we have

\[ V V^\dagger \sigma^{1/2} h (d) = \sigma^{1/2} h (d), \]

and, combined with (42),

\[ V \Lambda (\sigma)^{1/2} h (d') = \sigma^{1/2} h (d), \tag{43} \]

or

\[ \Lambda^\dagger \left( h \left( \Lambda (\rho) \{ \Lambda (\sigma) \}^{-1} \right) \right) = h (\rho \sigma^{-1}). \]

References

[1] R. Bhatia, Matrix Analysis, Springer (1996)
[2] R. Bhatia, Positive Definite Matrices, Princeton University Press (2007)
[3] A. Chefles, "Deterministic quantum state transformations," Phys. Lett A 270, 14 (2000)
[4] M.D. Choi, "A Schwarz inequality for positive linear maps on \textsc{c}*-algebras," Ill. J. Math., 18, 565-574 (1974)
[5] A. Ebadian, I. Nikoufar, and M. Gordjic, "Perspectives of matrix convex functions," Proc Natl Acad Sci USA, 108(18), 7313–7314, (2011)
[6] D. G. Luenberger, "Optimization by vector space methods", Wiley, New York, (1969)
[7] F. Hiai, M. Mosonyi, D. Petz and C. Beny, "Quantum f-divergences and error corrections," Reviews in Mathematical Physics, 23, 691–747 (2011)
[8] K. Matsumoto, "Reverse test and quantum analogue of classical Fidelity and generalized Fidelity," http://xxx.lanl.gov/abs/1006.0302 (2010)
[9] A. Jencova, "Reversibility conditions for quantum operations," http://lanl.arxiv.org/abs/1107.0453 (2012)
[10] G. G. Magarill-Ilyaev, V. M. Tikhomirov, "Convex Analysis: Theory and Applications," (AMS, 2003)
[11] K. Parthasarathy, "Probability and Measures on Metric Spaces," (Academic Press, 1967)
[12] R. T. Rockafellar, "Convex Analysis" (Princeton, 1970)
[13] A. Uhlmann, "Eine Bemerkung über vollständig positive Abbildungen von Dichteopera-toren", Wiss. Z. KMU Leipzig.Math.-Naturwiss. R. 34(6), 580-582 (1985).
Some backgrounds from matrix analysis

If $\Lambda^\dagger$ is a unital 2-positive map, following Schwartz inequality holds,

$$\Lambda^\dagger (Z^\dagger) \Lambda^\dagger (Z) \leq \Lambda^\dagger (Z^\dagger Z).$$

(44)

A function $f$ from a subset of $\mathbb{R}$ to $\mathbb{R}$ is called operator monotone if

$$Z_1 \geq Z_2 \geq 0 \implies f(Z_1) \geq Z_2,$$

and called operator convex if

$$f(sZ_1 + (1 - s)Z_2) \leq sf(Z_1) + (1 - s)f(Z_2), \quad 0 \leq s \leq 1.$$

Proposition 16 The function $\psi_t(y) := -\frac{y}{y+t}$ is operator monotone decreasing and operator convex on $[0, \infty)$.

Proposition 17 (Theorem V.2.3 of [1]) Let $f$ be a continuous function on $[0, \infty)$ then, the followings are equivalent: (i) $f$ is operator convex and $f(0) \leq 0$. (ii) For any positive operator $Z$ and an operator $C$ such that $\|C\| \leq 1$, $f(C^\dagger AC) \leq C^\dagger f(A)C$.

Proposition 18 (Theorem 2.1 of [4]) Let $f$ be an operator convex function defined on $[0, \infty)$. Then, for any unital positive map $\Lambda^\dagger : \mathcal{L}_A \to \mathcal{L}_B$,

$$f(\Lambda^\dagger (A)) \leq \Lambda^\dagger (f(A))$$

holds for any $A \geq 0$.

Proof. By Proposition 17, the assertion is true if $f(0) \leq 0$ and $\Lambda^\dagger$ is CP and subunital. Since $\Lambda^\dagger$ is unital,

$$f(\Lambda^\dagger (A)) + c1_B = f(\Lambda^\dagger (A)) + c\Lambda^\dagger (1_A) \leq \Lambda^\dagger (f(A)) + c\Lambda^\dagger (1_A)$$

$$= \Lambda^\dagger (f(A) + c1_A).$$

Therefore, we can remove the restriction $f(0) \leq 0$. Since for any positive map $\Lambda^\dagger$, its restriction to the commutative algebra generated by $A$ is CP. Thus we have the assertion.

Proposition 19 (Theorem V.2.5 of [1]) Let $f$ be a (continuous) function mapping the positive half-line $[0, \infty)$ into itself. Then $f$ is operator monotone if and only if it is operator concave.

Proposition 20 (Proposition 8.4 of [7]) Let $f$ be a continuous operator convex function on $[0, \infty)$. Then, if

$$\lim_{x \to \infty} \frac{f(y)}{y} < \infty,$$
there is a real number $a$ and a non-negative Borel measure $\mu$ such that

$$f(y) = f(0) + ay + \int_{(0, \infty)} \psi_t(y) \, d\mu(t),$$

$$\psi_t(y) := -\frac{y}{y+t},$$

and

$$\int_{(0, \infty)} \frac{d\mu(t)}{1+t} < \infty. \quad (45)$$

Proposition 21 (Theorem 8.1 if \cite{7}) A continuous real valued function $f$ on $[0, \infty)$ is operator convex if and only if

$$f(y) = f(0) + ay + by^2 + \int_{(0, \infty)} \left(\frac{y}{1+t} - \frac{y}{y+t}\right) \, d\mu(t),$$

where $a$ is a real number, $b$ is a non-negative real number, and $\mu$ is a finite non-negative measure satisfying

$$\int_{(0, \infty)} \frac{d\mu(t)}{(1+t)^2} < \infty. \quad (46)$$

Proposition 22 (Exercise 1.3.5 of \cite{2}) Let $X$, $Y$ be a positive definite matrices. Then,

$$\begin{bmatrix} X & C \\ C^\dagger & Y \end{bmatrix} \geq 0 \quad (46)$$

implies

$$X \geq CY^{-1}C^\dagger \quad (47)$$

and

$$Y \geq C^\dagger X^{-1}C. \quad (48)$$

Proof. (46) implies that $(1 - \pi_X)C = 0$ or $C(1 - \pi_Y) = 0$. To prove the former statement, suppose $(1 - \pi_X)C \neq 0$. Then, there is a unit vector $|\varphi\rangle$ in the support of $1 - \pi_X$ such that $\langle \varphi | C \neq 0$. Therefore, for a sufficiently large $c > 0$,

$$\begin{bmatrix} -c \langle \varphi | & \langle \varphi | C \end{bmatrix} \begin{bmatrix} X & C \\ C^\dagger & Y \end{bmatrix} \begin{bmatrix} -c |\varphi\rangle \\ C^\dagger |\varphi\rangle \end{bmatrix} = 0 - 2c \langle \varphi | CC^\dagger |\varphi\rangle + \langle \varphi | C^\dagger YC |\varphi\rangle \geq 0.$$

This contradicts with (46). Therefore, we have $(1 - \pi_X)C = 0$. The proof of $C(1 - \pi_Y) = 0$ is almost parallel.

If (46) holds,

$$\begin{bmatrix} 1 & -CY^{-1} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X & C \\ C^\dagger & Y \end{bmatrix} \begin{bmatrix} 1 \\ -Y^{-1}C^\dagger \end{bmatrix} = \begin{bmatrix} X - CY^{-1}C^\dagger & C - C\pi_Y \\ C^\dagger - \pi_YC^\dagger & Y \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \geq 0, \quad (47)$$

which implies (47). (48) is proved almost parallelly. \qed
Proposition 23 (Lemma 5.2 of [7]) If $f$ is a complex-valued function on finitely many points $\{x_i; i \in I\}$, $I \subset [0, \infty)$, then for any pairwise different positive numbers $\{t_i; i \in I\}$ there exist complex numbers $\{c_i; i \in I\}$ such that $f(x_i) = \sum_{j \in I} \frac{c_j}{x_i+t_j} , i \in I$.

B Some backgrounds from convex analysis

The epigraph $\text{epi } f$ of function $f : \mathbb{R}^n \to \mathbb{R} \cup \{-\infty, -\infty\}$ is the set

$$\text{epi } f := \{(y, z); y \in \mathbb{R}^n, z \geq f(y)\} \subset \mathbb{R}^{n+1}.$$ 

By definition,

$$\text{epi } \sup_j f_j = \bigcap_j \text{epi } f_j .$$ (49)

The effective domain $\text{dom } f$ is

$$\text{dom } f := \{y; y \in \mathbb{R}^n, f(y) < \infty\} \subset \mathbb{R}^n .$$

$f$ is said to be convex if $\text{epi } f$ is convex, and closed if $\text{epi } f$ is closed. Also, $f$ is said to be proper if $f(y) \neq -\infty$ for any $y$ and $f(y) \neq \infty$ for some $y$. A sublinear function $f$ is a function which is convex and homogeneous. The closure $\text{cl } f$ of $f$ is the function such that

$$\text{epi } (\text{cl } f) = \text{cl } (\text{epi } f) .$$

The dual $f^* : \mathbb{R}^n \to \mathbb{R} \cup \{-\infty, -\infty\}$ of $f$ is

$$f^*(x) := \sup_{y \in \mathbb{R}^n} x \cdot y - f(y) .$$

The recession cone $0^+ C$ of the convex set $C$ is

$$0^+ C := \{y; y_0 + \lambda y \in C, \forall y_0 \in C, \forall \lambda \geq 0\} ,$$

and the recession function $f^0$ of the convex function $f$ is the function such that

$$\text{epi } (f^0) = 0^+ \text{epi } (f) .$$

Proposition 24 (Theorem 7.4, [12]) If $f$ is proper and convex, so is $\text{cl } f$ .

Proposition 25 (Theorem 12.2, [12]) Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ be a proper convex function. Then, $f^{**} = \text{cl } f$ .

Proposition 26 If $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ is a proper convex function, $f^0$ is a proper sublinear function. Moreover,

$$f^0 (y) = \sup_{y' \in \text{dom } f} \{f(y') + y - f(x)\} = \sup_{\lambda > 0} \frac{f(y_0 + \lambda y) - f(y_0)}{\lambda} = \lim_{\lambda \to \infty} \frac{f(y_0 + \lambda y) - f(y_0)}{\lambda} ,$$
where \( y_0 \in \text{dom } f \). Also, if \( y \in \text{dom } f \),

\[
f^0(y) = \lim_{\varepsilon \downarrow 0} \varepsilon f \left( \frac{y}{\varepsilon} \right).
\]

If \( 0 \in \text{dom } f \), the latter formula actually holds for any \( y \in \mathbb{R}^n \).

**Proposition 27** If \( f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\} \) is a proper convex function, then

\[
g(y, \lambda) := \begin{cases} 
\lambda f \left( \frac{y}{\lambda} \right) & (\lambda > 0), \\
f^0(y) & (\lambda = 0), \\
\infty & (\lambda < 0),
\end{cases}
\]

is closed, proper, and sublinear.

**Proof.** Let

\[
\tilde{g}(y, \lambda) := \begin{cases} 
\lambda f \left( \frac{y}{\lambda} \right) & (\lambda > 0), \\
\delta(x|0) & (\lambda = 0), \\
\infty & (\lambda < 0),
\end{cases}
\]

where

\[
\delta(y|0) = \begin{cases} 
0 & (y = 0), \\
\infty & (y \neq 0).
\end{cases}
\]

Then \( g \) is proper and sublinear by p. 35 of [12]. By p. 67 of [12], \( g = \text{cl } \tilde{g} \). Thus \( g \) is closed, proper, and sublinear. ■

**Proposition 28** (p. 43 of [10]) Let \( f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\} \) is a proper sublinear function. Then it is closed if and only if it is pointwise supremum of linear functions that does not exceed \( f \).

30