Reverse test and quantum analogue of classical fidelity and generalized fidelity

Keiji Matsumoto
National Institute of Informatics, 2-1-2, Hitotsubashi, Chiyoda-ku, Tokyo
Quantum Computation and Information Project, SORST, JST, 5-28-3, Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

August 24, 2010

1 Introduction

In doing hypothesis test in quantum mechanical setting, key part is choice of measurement which maps given pair \{\rho, \sigma\} of quantum states to a pair \{p, q\} of probability distributions. Its inverse operation, or a CPTP map from \{p, q\} to \{\rho, \sigma\} is called reverse test, and plays an essential role in characterizing largest monotone quantum analogue of relative entropy \cite{7,8}. In this paper, we exploit the same line of argument in studying quantum analogues of affinity, or classical fidelity \(F(p, q) = \sum_x \sqrt{p(x)} \sqrt{q(x)}\), and more generalized fidelity \(F_f(p, q) := \sum_x p(x) f(q(x)/p(x))\), where \(f\) is an operator monotone function on \([0, \infty)\). (For example, \(f(t) = t^\alpha (0 < \alpha < 1)\).

In the paper, based on reverse test, we define \(F_{\min}(\rho, \sigma)\), which turns out to equal \(\text{tr} \rho \sqrt{\rho^{-1/2}\sigma\rho^{-1/2}}\). This quantity is monotone increasing by the application of TPCP maps, and in fact is the smallest one among the numbers satisfying these properties, while \(F(\rho, \sigma)\) is the largest. It is also proved that \(F_{\min}\) satisfies strong joint concavity using reverse test.

For generalized fidelity, we also introduce \(F_{f,\min}(\rho, \sigma)\) in the similar manner, which turns out to equal \(\text{tr} \rho f\left(\rho^{-1/2}\sigma\rho^{-1/2}\right)\). Again, this quantity is monotone increasing and is the smallest one among the numbers satisfying these properties. Joint concavity of \(F_{f,\min}(\rho, \sigma)\) is also proved using reverse test.

It is known that fidelity between infinitesimally different states gives rise to SLD Fisher information metric \(J^S\), or the smallest monotone metric, and that \(\cos^{-1} F(\rho, \sigma)\) equals the integral of \(J^S\) along the geodesic, or the curve which minimize the integral, connecting \(\rho\) and \(\sigma\).

Correspondingly, \(F_{\min}(\rho, \sigma)\) gives rise to RLD Fisher information metric \(J^R\), or the largest monotone metric. However, the integral of RLD Fisher information metric along the geodesic does not equal \(\cos^{-1} F_{\min}(\rho, \sigma)\). In fact, cosine of the integral, denoted by \(F_R(\rho, \sigma)\), is another monotone quantum analogue of classical fidelity, and is the smallest one among those which satisfy triangle
inequality. On the other hand, $F_{\text{min}}(\rho, \sigma)$ is the integral of RLD Fisher information along the curve which minimize the integral for all the curves with commutative RLD.

An upper and a lower bound of the quantum statistical distance $\Delta(\rho, \sigma) = \frac{1}{2} \| \rho - \sigma \|_1$ using $F(\rho, \sigma)$ is one of notable feature of this quantity. It turns out that $F_{\text{min}}(\rho, \sigma)$ gives analogous bounds of $\Delta_{\text{max}}(\rho, \sigma)$, which is another quantum version of statistical distance $\Delta(p, q) = \frac{1}{2} \| p - q \|_1$, defined using reverse test.

2 Classical fidelity and fidelity

We consider probability distributions over finite set $\mathcal{X}$ with $|\mathcal{X}| = d'$, and quantum states $S(\mathcal{H})$ over $d$-dimensional Hilbert space $\mathcal{H}$. Define classical and quantum fidelity $F$ by

$$F(p, q) : = \sum_{x \in \mathcal{X}} \sqrt{p(x)} \sqrt{q(x)};$$
$$F(\rho, \sigma) : = \max_{U: \text{unitary}} \text{tr} \sqrt{\rho} \sqrt{\sigma} U \sqrt{\sigma} \rho \sqrt{\sigma},$$

Known facts about them are:

- $F(\rho, \sigma) = \min_{M: \text{measurement}} F(M(\rho), M(\sigma))$ (1)
  where $M(\rho)$ is the probability distribution of measurement $M$ applied to $\rho$.

- $F(\rho, \sigma) = \text{tr} W_{\rho}^\dagger W_{\sigma}$, (2)
  where $W_{\rho}$ and $W_{\sigma}$ are $d \times d$ matrices with
  $$W_{\rho} W_{\rho}^\dagger = \rho, W_{\sigma} W_{\sigma}^\dagger = \sigma,$$
  $$W_{\rho}^\dagger W_{\sigma} = W_{\sigma}^\dagger W_{\rho} \geq 0.$$

- Given parameterized family $\{p_t\}$ and $\{\rho_t\}$, define Fisher information $J_t$ and symmetric logarithmic derivative (SLD) Fisher information $J_t^S$ by

  $$J_t : = \sum_{x \in \mathcal{X}} (l_t(x))^2 p_t(x),$$
  $$J_t^S : = \text{tr} (L_t^S)^2 \rho_t,$$
where \( l_t \), called logarithmic derivative, is defined by \( l_t (x) := \frac{d}{dt} \log p_t (x) \), and \( L^S_t \), called symmetric logarithmic derivative (SLD), is a solution to a linear equation
\[
\frac{d \rho_t}{dt} = \frac{1}{2} \left\{ L^S_t \rho_t + \rho_t L^S_t \right\}.
\]
Then,
\[
F (p_t, p_{t+\varepsilon}) = 1 - \frac{1}{8} J_t \varepsilon^2 + o (\varepsilon^2),
\]
\[
F (\rho_t, \rho_{t+\varepsilon}) = 1 - \frac{1}{8} J^S_t \varepsilon^2 + o (\varepsilon^2),
\]
and
\[
F (p, q) = \cos \min_C \frac{1}{2} \int_C \sqrt{J_t} \, dt,
\]
\[
F (\rho, \sigma) = \cos \min_C \frac{1}{2} \int_C \sqrt{J^S_t} \, dt,
\]
where minimum is taken for all the paths with \( p_0 = p, p_1 = q \), and \( \rho_0 = \rho, \rho_1 = \sigma \), respectively.

- (symmetry) \( F (\rho, \sigma) = F (\sigma, \rho) \).
- (Monotonicity) \( F (\rho, \sigma) \leq F (\Lambda (\rho), \Lambda (\sigma)) \).
- (triangle inequality)
  \[
  \cos^{-1} F (\rho, \sigma) \leq \cos^{-1} F (\rho, \tau) + \cos^{-1} F (\tau, \sigma),
  \]
  \[
  F (\rho, \sigma) \geq 2 F (\rho, \tau) F (\tau, \sigma) - 1.
  \]
- (strong joint concavity)
  \[
  F \left( \sum_{y \in Y} \lambda_y \rho_y, \sum_{y \in Y} \mu_y \sigma_y \right) \geq \sum_{y \in Y} \sqrt{\lambda_y \mu_y} F (\rho_y, \sigma_y)
  \]
- (Multiplicativity) \( F (\rho^{\otimes n}, \sigma^{\otimes n}) = F (\rho, \sigma)^n \)

In the paper, we consider quantities satisfying:

\( \text{(N)} \) \( F^Q (p, q) = F (p, q) \), for all the probability distributions \( p, q \).

\( \text{(M)} \) \( F^Q (\rho, \sigma) \leq F^Q (\Lambda (\rho), \Lambda (\sigma)) \)
3 Another quantum analogue of classical fidelity

A triplet \((\Phi, \{p, q\})\) of a CPTP map \(\Phi\) and probability distributions \(p, q\) over the set \(X\) \((|X| = d')\) with

\[\Phi(p) = \rho, \Phi(q) = \sigma,\]  \hspace{1cm} (9)

is called reverse test of \(\{\rho, \sigma\}\). A reverse test \((\Phi, \{p, q\})\) with (9) is said to be minimal satisfying when \(|X| = d = \dim \mathcal{H}\). Define

\[F_{\text{min}}(\rho, \sigma) := \max_{(\Phi, \{p, q\}) : (9)} F(p, q).\]  \hspace{1cm} (10)

**Theorem 1** Suppose \(F^Q(p, q)\) (N) and (M). Then, \(F_{\text{min}}(\rho, \sigma) \leq F^Q(\rho, \sigma) \leq F(\rho, \sigma)\). Also, \(F_{\text{min}}(\rho, \sigma)\) satisfies (N) and (M).

**Proof.** Let \(M\) be a measurement achieving the minimum of (11). Then,

\[F^Q(\rho, \sigma) \leq F^Q(M(\rho), M(\sigma)) = F(M(\rho), M(\sigma)) = F(\rho, \sigma).\]

Let \(\Phi\) be a CPTP map achieving the maximum of (10). Then

\[F^Q(\rho, \sigma) = F^Q(\Phi(p), \Phi(q)) \geq F^Q(p, q) = F(p, q)\]

\[= F_{\text{min}}(\rho, \sigma).\]

Obviously, \(F_{\text{min}}(p, q) \geq F(p, q)\). Also, for any \(p', q'\) with \(p = \Phi(p'), q = \Phi(q')\), \(F(p', q') \leq F(p, q)\), by (11). Therefore, \(F_{\text{min}}(p, q) \leq F(p, q)\), and we have (N).

That \(F_{\text{min}}\) satisfies (M) is proved as follows.

\[F_{\text{min}}(\Lambda(\rho), \Lambda(\sigma)) = \max\{F(p, q) : \Phi: \text{CPTP}, \Phi(p) = \Lambda(\rho), \Phi(q) = \Lambda(\sigma)\}\]

\[\geq \max\{F(p, q) : \Phi = \Phi' \circ \Lambda, \Phi'(p) = \rho, \Phi'(q) = \sigma\}\]

\[= \max\{F(p, q) : \Phi'(p) = \rho, \Phi'(q) = \sigma\}\]

\[= F_{\text{min}}(\rho, \sigma).\]

**Theorem 2** Suppose \(\rho\) and \(\sigma\) are strictly positive. Then,

\[F_{\text{min}}(\rho, \sigma) = \tr \rho \sqrt{\rho^{-1/2} \sigma \rho^{-1/2}},\]

and the maximum of (10) is achieved by any minimal reverse test \((\Phi, \{p, q\})\).

The proof will be given later.
Remark 3 Using geometric mean $A \# B = \sqrt[\sqrt{A} - 1/2 B A^{-1/2} \sqrt{A}}$, 

$$F_{\min}(\rho, \sigma) = \text{tr} (\rho \# \sigma).$$

Hence, the well-known property of $\#$

$$\Lambda (\rho) \# \Lambda (\sigma) \geq \Lambda (\rho \# \sigma)$$

immediately implies that $F_{\min}$ satisfies (M).

Remark 4 In [7][8], minimization of $D (\rho || \sigma) := \sum_{x \in \mathcal{X}} p (x) \ln \frac{p (x)}{q (x)}$ over all the reverse tests $(\Phi, \{p, q\})$ is considered, and it is shown that the minimum is achieved also by any minimal reverse test.

4 Listing all reverse tests

In this section, to solve maximization (10), we give full characterization of all reverse tests $(\Phi, \{p, q\})$ of $\{\rho, \sigma\}$ with $\Phi (\delta_x)$ being a pure state, where $\delta_{x_0} (x)$ denotes a probability distribution concentrated at $x = x_0 \ (x, x_0 \in \mathcal{X})$.

The reason for such a restriction to be made is as follows. If $\Phi (\delta_x) = \rho x = \sum_{y \in \mathcal{Y}} s (y|x) \langle \varphi_{xy} | \varphi_{xy} \rangle$, let $\Phi' : r (x) \rightarrow r (x) s (y|x)$.

Let $\Phi'$ a CPTP map from probability distributions over $\mathcal{X} \times \mathcal{Y}$ to $S (\mathcal{H})$ such that

$$\Phi' (\delta_{(x,y)}) = | \varphi_{xy} \rangle \langle \varphi_{xy} |.$$ 

Then if $\Phi (p) = \rho$ and $\Phi (q) = \sigma$, $\Phi' (p') = \rho$ and $\Phi' (q') = \sigma$, where $p' (x, y) = p (x) s (y|x), q' (x, y) = q (y) s (y|x)$.

Hence, $(\Phi', \{p', q'\})$ is a reverse test of $\{\rho, \sigma\}$ with $\Phi (\delta_x)$ being a pure state, and $p = \Psi_1 (p'), q = \Psi_1 (q')$, where $\Psi_1$ is taking marginal over $\mathcal{Y}$. Hence, $F (p,q) \geq F (p',q')$. Also, observe $p' = \Psi_2 (p), q' = \Psi_2 (q)$, where $\Psi_2 : r (x) \rightarrow r (x) s (y|x)$.

Therefore, after all,

$$F (p,q) = F (p', q').$$

Therefore, we can replace $\Phi$ by $\Phi'$.

Below, we indicate $\Phi$ by a matrix $N$, whose $x$th column vector is $| \varphi_x \rangle$ with $| \varphi_x \rangle \langle \varphi_x | = \Phi (\delta_x)$. Then the condition (9) is rewritten as

$$\sum_{x \in \mathcal{X}} p (x) | \varphi_x \rangle \langle \varphi_x | = \rho, \sum_{x \in \mathcal{X}} q (x) | \varphi_x \rangle \langle \varphi_x | = \sigma$$

(12)

Here note in general, $d'$ can be larger than $d = \dim \mathcal{H}$.
Lemma 5 $WW^\dagger = \rho$ if and only if

$$W = \sqrt{\rho}U,$$

with $U$ being isometry, $UU^\dagger = 1$.

Proof. We only have to show ‘only if’. Suppose $WW^\dagger = \rho$. Then, letting $\sqrt{\rho}^{-1}$ be the (Moore-Penrose) generalized inverse of $\sqrt{\rho}$, we have

$$\left(\sqrt{\rho}^{-1}W\right)\left(\sqrt{\rho}^{-1}W\right)^\dagger = P,$$

where $P$ is the projector on the support of $\rho$. Observe

$$(1 - P)W(1 - P)^\dagger = (1 - P)\rho(1 - P) = 0$$

implies

$$(1 - P)W = 0$$

Therefore,

$$\sqrt{\rho}^{-1}\left(\sqrt{\rho}^{-1}W\right) = PW = W.$$

Let $U'$ be a partial isometry from ker $\left(\sqrt{\rho}^{-1}W\right)^\dagger$ to ker $\rho$,

$$U = \sqrt{\rho}^{-1}W \oplus U'$$

satisfies the requirement. Hence, we have the assertion. ■

In the reminder of the section, we suppose $\rho > 0$. Let $D_p$ and $D_q$ be

$$D_p = \sqrt{\text{diag} \,(p_1, \cdots, p_{d'})}, \quad D_q = \sqrt{\text{diag} \,(q_1, \cdots, q_{d'})}. \quad (13)$$

Then, by (12), $ND_p (ND_p)^\dagger = \rho$ and $ND_q (ND_q)^\dagger = \sigma$. Also,

$$T := \sqrt{\rho^{-1/2} \sigma \rho^{-1/2}} \quad (> 0) \quad (14)$$

satisfies $\left(\sqrt{\rho T}\right) \left(\sqrt{\rho T}\right)^\dagger = \sigma$. Therefore,

$$[\sqrt{\rho} \, 0] V = ND_p, \quad [\sqrt{\rho T} \, 0] U = ND_p, \quad (15)$$

for some $U, V \in U (\mathcal{H}')$, dim $\mathcal{H}' = d'$. Then,

$$ND_q D_p N^\dagger = [\sqrt{\rho} \, 0] UV^\dagger \left[\begin{array}{c} \sqrt{\rho} \\ 0 \end{array}\right]$$

$$= \sqrt{\rho}TA\sqrt{\rho} \geq 0, \quad (16)$$

where, with $P$ being the projector onto $\mathcal{H}$, $A = PUV^\dagger P$. Therefore,

$$TA = A^\dagger T \geq 0. \quad (17)$$
Also, by (15),

\[
[\sqrt{\rho} 0] V (D^{-1} q) V^\dagger = [\sqrt{\rho} T 0] U V^\dagger = [\sqrt{\rho} 0] \left[ \begin{array}{cc} T A & T A' \\ A'^\dagger T & C \end{array} \right],
\]

where \( A' := PU (1 - P) \) and \( C > 0 \). If \( \dot{C} > 0 \) satisfy

\[
\ker TA' \subset \ker C, \quad (TA)^{-1/2} (TA') C^{-1/2} \leq 1,
\]

we have

\[
\tilde{T} := \left[ \begin{array}{cc} T A & T A' \\ A'^\dagger T & C \end{array} \right] \geq 0. \tag{19}
\]

Note \( \{ A; \| A \| \leq 1 \} \) is identical to the totality of matrices with the form

\( A = PU \), where \( U \in U (\mathcal{H}'), \mathcal{H} \subset \mathcal{H}' \), \( \dim \mathcal{H}' = d' \) and \( P \) is the projector onto \( \mathcal{H} \). Then, a reverse test can be composed as indicated in the following (i)-(v):

(i) Choose \( A \in \{ A; \| A \| \leq 1 \} \) with (17).

(ii) Compose \( A' \) such that \( [AA'] \) is isometry.

(iii) Let \( \tilde{T} \) as of (19), \( V \) is diagonalize it : \( \tilde{T} = VDV^\dagger \).

(vi) Define \( |\varphi_x\rangle (x \in \mathcal{X}) \) as the normalized column vectors of \( [\sqrt{\rho}0]V \).

Finally, \( p(x) \) and \( q(x) \) is the square of the magnitude of the \( x \)th column vector of \( [\sqrt{\rho}0]V \) and \( [\sqrt{\rho}TA\sqrt{\rho}TA'] \), respectively. \( p \) and \( q \) are obtained also as follows. Define \( \rho', \sigma' \in S (\mathcal{H}') \) by

\[
\rho' = \begin{bmatrix} \rho & 0 \\ 0 & 0 \end{bmatrix}, \quad \sigma' = \tilde{T}\rho\tilde{T},
\]

and let the measurement \( M \) be projectors onto eigenspaces of \( \tilde{T} \). Then, \( p = M (\rho') \) and \( M = (\sigma') \). So, the last step of the composition is:

(v) Let \( p = M (\rho') \) and \( q = M (\sigma') \), where \( M \) is the projectors onto eigenspaces of \( \tilde{T} \).

5 Proof of Theorem 2

Proof. of Theorem 2. By (16),

\[
\text{tr} \sqrt{\rho} TA \sqrt{\rho} = \text{tr} ND_q D_p N^\dagger = \text{tr} D_q N^\dagger N D_p = \text{tr} D_q D_p,
\]
where the last identity is due to \((N^\dagger N)_{ii} = 1\). On the other hand, (17) implies,

\[
T^2 - (TA)^2 = T^2 - TAA^\dagger T = T(1 - AA^\dagger) T \geq 0.
\]  

(20)

where the last inequality is due to \(\|A\| \leq 1\). Since \(\sqrt{T}\) is operator monotone,

\[
TA \leq T.
\]

Therefore,

\[
F_{\text{min}}(\rho, \sigma) = \max_{p, q \geq 0} \sum_{x \in \mathcal{X}} \sqrt{p(x) q(x)}
\]

\[
= \text{tr} D_q D_p
\]

\[
= \text{tr} \sqrt{\rho} T A \sqrt{\rho}
\]

\[
\leq \text{tr} \sqrt{\rho} T \sqrt{\rho}
\]

\[
= \rho \sqrt{\rho}^{-1/2} \sigma \rho^{-1/2}.
\]

The inequality is achieved when \(A = 1\), which corresponds to minimal reverse tests.

6 Seeing from ‘behind’

By Theorem(2),

\[
F_{\text{min}}(\rho, \sigma) = \text{tr} W_\rho W_\sigma^\dagger,
\]

where \(W_\rho := \sqrt{\rho}\) and \(W_\sigma := \sqrt{\rho} T\), with \(T\) being as of (14). Observe

\[
W_\rho W_\sigma^\dagger = W_\sigma W_\rho^\dagger.
\]

Therefore, by (2), we have

\[
F_{\text{min}}(\rho, \sigma) = F(\rho', \sigma'),
\]  

(21)

where

\[
\rho' := W_\rho^\dagger W_\rho = \rho, \quad \sigma' := W_\sigma^\dagger W_\sigma = T \rho T.
\]

A meaning of (21) is given in the sequel. Letting

\[
W_\rho = \sum_{i=1}^d \sum_{x \in \mathcal{X}} w_{\rho, i} x |i\rangle \langle e_x|, \quad W_\sigma = \sum_{i=1}^d \sum_{x \in \mathcal{X}} w_{\sigma, i} x |i\rangle \langle e_x|
\]

define

\[
|W_\rho\rangle := \sum_{i=1}^d \sum_{x \in \mathcal{X}} w_{\rho, i} x |i\rangle |e_x\rangle, \quad |W_\sigma\rangle := \sum_{i=1}^d \sum_{x \in \mathcal{X}} w_{\sigma, i} x |i\rangle |e_x\rangle.
\]
Then, one can easily check

\[ \rho = \text{tr}_{\mathcal{H}'} |W_\rho\rangle \langle W_\rho|, \quad \sigma := \text{tr}_{\mathcal{H}'} |W_\sigma\rangle \langle W_\sigma| \]

\[ \rho' = \text{tr}_{\mathcal{H}} |W_\rho\rangle \langle W_\rho|, \quad \sigma' = \text{tr}_{\mathcal{H}} |W_\sigma\rangle \langle W_\sigma| \]

hold. Hence, \( F_{\text{min}} (\rho, \sigma) \) equals fidelity of ‘hidden’ part of the purification of \( \rho \) and \( \sigma \).

7 \( F_{\text{min}} \) for pure states

Any reverse test \( (\Phi, \{ p, q \}) \) of \( \{ \rho, |\varphi\rangle \} \) is in the following form:

\[ \Phi (\delta_x) = |\varphi\rangle \langle \varphi|, \quad x \in \text{supp} \; q, \]

\[ \rho = \Phi (p) = c |\varphi\rangle \langle \varphi| + \sum_{x \notin \text{supp} \; q} p (x) \Phi (\delta_x), \quad (22) \]

where \( c := \sum_{x \in \text{supp} \; q} p (x) \). Therefore, by monotonicity of Fidelity by CPTP maps,

\[ \sum_{x \in X} \sqrt{p(x)q(x)} \leq \sqrt{c \cdot 1} + \sqrt{c \cdot 0} = \sqrt{c}, \]

and the inequality is achieved by the following \( q (x) \) and \( p (x) \):

\[ q (x) = \delta_{x_0}, \quad p (x) = c \delta_{x_0} (x \in \text{supp} \; q), \]

where \( x_0 \) is a point in \( \text{supp} \; q \).

Therefore, we maximize \( c \) with \( \rho - c |\varphi\rangle \langle \varphi| \geq 0 \), or equivalently, if \( |\varphi\rangle \in \text{supp} \; \rho \),

\[ 1 - c \rho^{-1/2} |\varphi\rangle \langle \varphi| \rho^{-1/2} \geq 0. \]

Therefore, if \( |\varphi\rangle \in \text{supp} \; \rho \),

\[ F_{\text{min}} (\rho, |\varphi\rangle) = \sqrt{(\text{tr} \; \rho^{-1/2} |\varphi\rangle \langle \varphi| \rho^{-1/2})^{-1}} \]

\[ = \left\| \sqrt{\rho^{-1}} |\varphi\rangle \right\| ^{-1}. \]

In case \( |\varphi\rangle \notin \text{supp} \; \rho \), the maximum of \( c \) with \( \rho - c |\varphi\rangle \langle \varphi| \geq 0 \) is zero, and

\[ F_{\text{min}} (\rho, |\varphi\rangle) = 0. \]

In particular,

\[ F_{\text{min}} (|\psi\rangle, |\varphi\rangle) = 0. \quad (23) \]
8 Properties of $F_{\min}$

Proposition 6 $F_{\min}(\rho, \sigma)$ does not satisfy triangle inequalities: there is $\rho$, $\sigma$, and $\tau$ with

$$\cos^{-1} F_{\min}(\rho, \sigma) > \cos^{-1} F_{\min}(\rho, \tau) + \cos^{-1} F_{\min}(\tau, \sigma),$$

$$F_{\min}(\rho, \sigma) < 2F_{\min}(\rho, \tau) F_{\min}(\tau, \sigma) - 1.$$ 

Proof. We let

$$\rho = |\psi\rangle \langle \psi|, \sigma = |\varphi\rangle \langle \varphi|,$$

$$|\psi\rangle = \begin{bmatrix} \cos \frac{\theta}{2} \\ \sin \frac{\theta}{2} \end{bmatrix}, |\varphi\rangle = \begin{bmatrix} \cos \frac{\theta}{2} \\ -\sin \frac{\theta}{2} \end{bmatrix},$$

$$\tau = \frac{1}{|\cos \frac{\theta}{2}| + |\sin \frac{\theta}{2}|} \begin{bmatrix} |\cos \frac{\theta}{2}| & 0 \\ 0 & |\sin \frac{\theta}{2}| \end{bmatrix}.$$ 

Then,

$$F_{\min}(|\psi\rangle, |\varphi\rangle) = 0,$$

$$F_{\min}(|\psi\rangle, \tau) = F_{\min}(|\varphi\rangle, \tau) = \frac{1}{|\cos \frac{\theta}{2}| + |\sin \frac{\theta}{2}|}.$$

Hence, letting $\theta$ be small enough, we have asserted inequalities. (In fact, the inequality is satisfied all $\theta$ lying between $0$ and $\frac{\pi}{2}$.)

Proposition 7 $F_{\min}(\rho, \sigma) = F_{\min}(\sigma, \rho)$, $F_{\min}(\rho^\otimes n, \sigma^\otimes n) = F_{\min}(\rho, \sigma)^n$.

Proof. Trivial by definition.

Theorem 8 (Strong concavity)

$$F_{\min} \left( \sum_{y \in Y} \lambda_y p_y, \sum_{y \in Y} \mu_y \sigma_y \right) \geq \sum_{y \in Y} \sqrt{\lambda_y \mu_y} F_{\min}(\rho_y, \sigma_y)$$

Proof. Let $(\Phi_y, \{p_y, q_y\})$ be a reverse test of $\{\rho_y, \sigma_y\}$ with

$$F(p_y, q_y) = F_{\min}(\rho_y, \sigma_y),$$

where $p_y$, $q_y$ are probability distributions over $\mathcal{X}$ with $|\mathcal{X}| = d' = d$ (minimal). Define $\tilde{p}_{yo}(x, y) := p_{yo}(x) \delta_{yo}(y)$, $\tilde{q}_{yo}(x, y) := q_{yo}(x) \delta_{yo}(y)$, and $\tilde{\Phi}(\delta_{yo}) := \Phi_y(\delta_x)$. Then,

$$\tilde{\Phi}(\tilde{p}_{yo}) = \sum_{x \in \mathcal{X}, y \in Y} p_{yo}(x) \delta_{yo}(y) \Lambda_y(\delta_x) = p_{yo},$$

$$\tilde{\Phi}(\tilde{q}_{yo}) = \sigma_{yo},$$

$$F(\tilde{p}_{yo}, \tilde{q}_{yo}) = \sqrt{\sum_{x \in \mathcal{X}, y \in Y} p_{yo}(x) q_{yo}(x) \delta_{yo}(y)} = F(p_{yo}, q_{yo}) = F_{\min}(\rho_{yo}, \sigma_{yo}).$$
Therefore,

\[
\sum_{y \in \mathcal{Y}} \sqrt{\lambda_y \mu_y} F_{\min} (\rho_y, \sigma_y) = \sum_{y \in \mathcal{Y}} \sqrt{\lambda_y \mu_y} F (\tilde{\rho}_y, \tilde{\sigma}_y) \\
\leq F \left( \sum_{y \in \mathcal{Y}} \lambda_y \tilde{\rho}_y, \sum_{y \in \mathcal{Y}} \mu_y \tilde{\sigma}_y \right) \leq F_{\min} \left( \hat{\Phi} \left( \sum_{y \in \mathcal{Y}} \lambda_y \tilde{\rho}_y \right); \hat{\Phi} \left( \sum_{y \in \mathcal{Y}} \mu_y \tilde{\sigma}_y \right) \right) \\
= F_{\min} \left( \sum_{y \in \mathcal{Y}} \lambda_y \rho_y, \sum_{y \in \mathcal{Y}} \mu_y \sigma_y \right),
\]

where the second line is due to (8) for classical fidelity, and the third and the fourth line is due to (N) and (M), respectively. □

**Remark 9** Alternative proof is given using (11) and the following property of \# [11]:

\[ A \# C + B \# D \leq (A + B) \# (C + D). \]

### 9 Generalization to \( F_f \)

As noted before, both of the minimum of \( D (p||q) \) and the maximum of \( F (p, q) \) are achieved by minimal ones. This section tries generalization of

Let \( f \) be operator monotone function on \([0, \infty)\). Then one can define

\[ F_f (p, q) := \sum_{x \in \mathcal{X}} p (x) \ f \left( \frac{q (x)}{p (x)} \right). \]

An example is \( f (t) = t^\alpha \) (0 < \( \alpha < 1 \)),

\[ F_\alpha (p, q) = \sum_{x \in \mathcal{X}} p^{1-\alpha} (x) q^\alpha (x). \]

Note

\[ D (p||q) = \lim_{\alpha \to 1} \frac{1}{1 - \alpha} \ln F_{1-\alpha} (p, q). \]

Thus, this family interpolates between \( D (p||q) \) and \( F (p, q) \).

Their quantum analogue is defined as follows.

\[ F_f^{\min} (\rho, \sigma) = \max F_f (p, q) \]

Here the maximum is taken over all the reverse tests \((\Phi, \{p, q\})\) of \{\( \rho, \sigma \)\}.

**Theorem 10** If \( F_f^Q (\rho, \sigma) \) satisfies \( F_f^Q (p, q) = F_f (p, q) \) for all probability distributions \( p, q \), and monotone increasing by any CPTP map,

\[ F_f^{\min} (\rho, \sigma) \leq F_f^Q (\rho, \sigma). \]
Proof. Almost parallel with the proof of Theorem 11, thus omitted. ■

Theorem 11

\[ F_{f}^{\min} (\lambda \rho_0 + (1 - \lambda) \rho_1, \lambda \sigma_0 + (1 - \lambda) \sigma_1) \geq \lambda F_{f}^{\min} (\rho_0, \sigma_0) + (1 - \lambda) F_{f}^{\min} (\rho_1, \sigma_1). \]

Proof. Almost parallel with the proof of Theorem 8, thus omitted. ■

Also, we define

\[ F'_{f} (\rho, \sigma) := \text{tr} \rho^{\sharp} f (T^2) \rho^{\sharp}, \]

where \( T \) is as of [14].

Due to the main result (the equation (3.8) of [1]), one finds an operator connection \( \natural \) with

\[ \rho^{\sharp} f (T^2) \rho^{\sharp} = \rho \natural \sigma \quad (\rho > 0), \]

thus

\[ F'_{f} (\rho, \sigma) = \text{tr} \rho \natural \sigma \quad (\rho > 0). \]

Any operator connection \( \natural \) satisfies

\[ S (A \natural B) S^\dagger \leq (SAS^\dagger) \natural (SBS^\dagger), \]
\[ A \natural B + C \natural D \leq (A + C) \natural (B + D), \]

(see Theorem 3.5 of [1]) which implies

\[ \Lambda (A \natural B) \leq \Lambda (A) \natural \Lambda (B), \]

for any TPCP map \( \Lambda \). This implies

\[ F'_{f} (\Lambda (\rho), \Lambda (\sigma)) \geq F'_{f} (\rho, \sigma). \]

Since \( F'_{f} (p, q) = F (p, q) \), by Theorem 10

\[ F_{f}^{\min} (\rho, \sigma) \leq F'_{f} (\rho, \sigma). \tag{24} \]

Theorem 12 Suppose \( \rho > 0 \). Then,

\[ F_{f}^{\min} (\rho, \sigma) = F'_{f} (\rho, \sigma). \]

Moreover, this number is achieved by any minimal reverse test.

Proof. By (24), we only have to show ‘\( \geq \)’. Let \( \{ \Phi, \{ p, q \} \} \) be a minimal reverse test. Then, by the argument in Section [14] \( p = M (\rho) \) and \( q = M (T \rho T) \), where \( M \) is the projectors onto the eigenvectors \( \{ |e_x\rangle \} \) of \( T = T = \sum_{x \in X} \lambda_x |e_x\rangle \langle e_x| \),

\[ p (x) = \langle e_x| \rho |e_x\rangle, \quad q (x) = \lambda_x^2 (\langle e_x| \rho |e_x\rangle). \]

Therefore,

\[ F_{f}^{\min} (\rho, \sigma) = \sum_{x \in X} \langle e_x| \rho |e_x\rangle f (\lambda_x^2) = \sum_{x \in X} \langle e_x| \rho f (T^2) |e_x\rangle = F'_{f} (\rho, \sigma). \]

Therefore, we have \( F_{f} (\rho, \sigma) \geq F'_{f} (\rho, \sigma) \), and the proof is complete. ■
10 RLD Fisher information and tangent reverse estimation

Given a parameterized family \( \{ \rho_t \} \), we define right logarithmic derivative (RLD) Fisher information \( J^R_t \) by
\[
J^R_t := \text{tr} \left( L^R_t \right)^\dagger L^R_t \rho_t,
\]
where \( L^R_t \) is called right logarithmic derivative and is the unique solution to the linear equation
\[
\frac{d\rho_t}{dt} = L^R_t \rho_t.
\]
(\( L^R_t \) exists if and only if supp \( \frac{d\rho_t}{dt} \subset \text{supp} \rho_t \).)

A triplet \( (\Lambda, p_t, \frac{dp_t}{dt}) \) is said to be tangent reverse estimation of \( \rho_t \) at \( t \) if it satisfies
\[
\Lambda (p_t) = \rho_t, \quad \Lambda \left( \frac{dp_t}{dt} \right) = \frac{dp_t}{dt}.
\]

With \( \Lambda (\delta_x) = |\varphi_x\rangle \langle \varphi_x| \), we have
\[
\sum_{x \in X'} p_t (x) |\varphi_x\rangle \langle \varphi_x| = NP_t N^\dagger = \rho_t,
\]
\[
\sum_{x \in X'} \frac{dp_t (x)}{dt} |\varphi_x\rangle \langle \varphi_x| = N \frac{dP_t}{dt} N^\dagger = \frac{dp_t}{dt},
\]
where \( N = [|\varphi_1\rangle \cdots |\varphi_{d'}\rangle] \) and \( P_t = \text{diag} (p_t (1), \cdots , p_t (d')) \). When \( d' = d = \text{dim} \mathcal{H} \), we say the tangent reverse estimation is minimal. It is known that
\[
J^R_t = \min J_{p_t},
\]
where the minimum is taken for all the tangent reverse estimation of \( \rho_t \) at \( t \).

It is also known that the minimum is achieved by any minimal tangent reverse estimation, and RLD satisfies
\[
L^R_t = NL_t N^{-1}
\]
with \( L_t = \text{diag} (l_t (1), \cdots , l_t (d)) \) and \( l_t = \frac{dp_t/dt}{p_t} \).

11 \( F_{\text{min}} \), RLD Fisher information and the shortest distance

Consider a smooth curve \( \{ \rho_t \} \). Suppose \( \{ \rho_t \} \) is lying interior of \( \mathcal{S} (\mathcal{H}) \) we have
Since $B_t$ then the second and third equality determines $\rho_t$ and thus $f$ where $X$.

They are determined by comparing both ends of $\rho_{t,s}$, and define $A_{t,s}$, $B_{t,s}$, and $C_{t,s}$ by

$$\sqrt{X_{t,s} + \varepsilon A_{t,s} + \varepsilon^2 B_{t,s} + \varepsilon^3 C_{t,s} + o(\varepsilon^3)}.$$

They are determined by comparing both ends of

$$X_{t,s} + \frac{dX_{t,s}}{ds} + \frac{1}{2} \frac{d^2 X_{t,s}}{ds^2} \varepsilon^2 + \frac{1}{3!} \frac{d^3 X_{t,s}}{ds^3} \varepsilon^3 + o(\varepsilon^3) = \left\{ \sqrt{X_{t,s} + \varepsilon A_{t,s} + \varepsilon^2 B_{t,s} + \varepsilon^3 C_{t,s} + o(\varepsilon^3)} \right\}^2,$$

and thus

$$A_{t,s} \sqrt{X_{t,s}} + \sqrt{X_{t,s} A_{t,s}} = \frac{dX_{t,s}}{ds},$$

$$B_{t,s} \sqrt{X_{t,s}} + \sqrt{X_{t,s} B_{t,s}} = \frac{1}{2} \frac{d^2 X_{t,s}}{ds^2} - (A_{t,s})^2,$$

$$C_{t,s} \sqrt{X_{t,s}} + \sqrt{X_{t,s} C_{t,s}} = \frac{1}{3!} \frac{d^3 X_{t,s}}{ds^3} - (A_{t,s} B_{t,s} + B_{t,s} A_{t,s}).$$

Since $\sqrt{X_{t,s}} > 0$, the first equation determines $A_{t,s}$ uniquely, and by Theorem VII.2.12 of [2],

$$\sup_{t \leq s \leq t+\varepsilon} \|A_{t,s}\| \leq \|X_{t,s}\|^{-1/2} \left\| \frac{dX_{t,s}}{ds} \right\|.$$

Then the second and third equality determines $B_{t,s}$ and $C_{t,s}$ uniquely, and

$$\sup_{t \leq s \leq t+\varepsilon} \|B_{t,s}\| \leq f_1 \left( \|\rho_t\|, \left\| \frac{d\rho_t}{ds} \right\|, \left\| \frac{d^2 \rho_t}{ds^2} \right\| \right),$$

$$\sup_{t \leq s \leq t+\varepsilon} \|C_{t,s}\| \leq f_2 \left( \|\rho_t\|, \left\| \frac{d\rho_t}{ds} \right\|, \left\| \frac{d^2 \rho_t}{ds^2} \right\| \right),$$

where $f_1$, $f_2$ is a continuous function. Therefore,
Let us define

\[ F_R(\rho, \sigma) := \cos \left( \min_C \int_C \sqrt{J_t^R} \, dt \right), \]  

where minimization is taken over all the smooth paths connecting \( \rho \) and \( \sigma \). Obviously, \( F_R(\rho, \sigma) \) satisfies (N) and (M), and a triangle inequality (6). Therefore we have:

**Proposition 13**

\[ F_{\min}(\rho, \sigma) \leq F_R(\rho, \sigma) \leq F(\rho, \sigma), \]

and \( F_{\min}(\rho, \sigma) \) does not coincide with \( F_R(\rho, \sigma) \).

**Proof.** Since \( F_R(\rho, \sigma) \) satisfies (N) and (M), the first assertion is obtained by Theorem 1. Since \( F_{\min} \) does not satisfy (6) while \( F_R(\rho, \sigma) \) does, we have the second assertion. \( \blacksquare \)

**Theorem 14**

\[ F_{\min}(\rho, \sigma) = \cos \left( \min_C \frac{1}{2} \int_C \sqrt{J_t^R} \, dt \right), \]

where minimization is taken over all the smooth paths \( C = \{\rho_t\} \) with \( \rho_0 = \rho \) and \( \rho_1 = \sigma \), with \([L_s^R, L_t^R] = 0\), \( 0 \leq s, t \leq 1 \).

**Proof.** \([L_s^R, L_t^R] = 0\) and (25) imply \( L_t^R = NL_tN^{-1}, \forall t \), where \( L_t \) is a diagonal real matrix and the column vectors \( |\varphi_x\rangle (x = 1, \cdots, d') \) of \( N \) are normalized. Define \( P_0 := N^{-1}P_0(N^\dagger)^{-1} \), then

\[ L_0^R P_0 = N L_0 P_0 N^\dagger. \]
Therefore, $L_0P_0$ is Hermitian, and $P_0$ is also diagonal.

Therefore,

$$\rho_t = NP_tN^\dagger, \quad 0 \leq t \leq 1,$$

where $P_t$ is a real diagonal matrix. Writing the $x$th diagonal element of $P_t$ as $p_t(x)$, $\sum_{x=1}^d p_t(x) = 1$, and thus $p_t(x)$ is a probability distribution over \{1, \cdots, d\}. One can check that $L_t$ is the logarithmic derivative of $p_t$, and that $J_{\rho_t} = J_{p_t}$. Therefore,

$$\cos \left( \frac{1}{C} \frac{1}{2} \int_C \sqrt{J_{\rho_t}^R} dt \right) = \cos \left( \frac{1}{C} \frac{1}{2} \int_C \sqrt{J_{\rho_t}} dt \right),$$

where the minimum is taken over all the smooth curves $C = \{p_t\}$ in probability distributions connecting $p_0$ and $p_1$. By (3), the LHS equals $\cos^{-1} F(p_0, p_1)$.

Define $\Phi(\delta_x) = |\langle \phi_x \rangle \langle \psi_x \rangle|$, then $(\Phi, \{p_0, p_1\})$ is a minimal reverse test of $\{\rho, \sigma\}$. Therefore, $\cos^{-1} F(p_0, p_1)$ equals $\cos^{-1} F_{\min}(\rho, \sigma)$, and we have the assertion. ■

**Theorem 15** Suppose that $F^Q(\rho, \sigma)$ satisfies $(M)$, $(N)$, and

$$\cos^{-1} F(\rho, \sigma) \leq \cos^{-1} F(\rho, \tau) + \cos^{-1} F(\tau, \sigma).$$

Then,

$$F_R(\rho, \sigma) \leq F^Q(\rho, \sigma) \leq F(\rho, \sigma).$$

**Proof.** We only have to show the lowerbound. Since any bounded and closed subset of interior of $S(H)$ is compact, there is the unique shortest Riemannian geodesic $C = \{\rho_t\}$ with respect to RLD Fisher information metric connecting any $\rho_0 = \rho > 0$ and $\rho_1 = \sigma > 0$ (see Theorem 1.7.1 of [4]). Then, by (6),

$$\cos^{-1} F^Q(\rho, \sigma) \leq \sum_{k=0}^{1/\varepsilon-1} \cos^{-1} F^Q(\rho_{k\varepsilon}, \rho_{(k+1)\varepsilon})$$

$$\leq \sum_{k=0}^{1/\varepsilon-1} \cos^{-1} F_{\min}(\rho_{k\varepsilon}, \rho_{(k+1)\varepsilon})$$

where the second line is due to Theorem 1. By elementary calculus, one can verify

$$\cos^{-1} F \leq \sqrt{2 \left(1 - F \right) \left(1 + \frac{1}{6} (1 - F) \right)}, \quad (0 \leq F \leq 1). \quad (29)$$

Due to the smoothness of geodesic, (26), and (29), letting

$$f_{i, t, s} := f_i \left( \|\rho_t\|, \left\| \frac{d\rho_t}{ds} \right\|, \left\| \frac{d^2\rho_t}{ds^2} \right\| \right), \quad (i = 1, 2),$$

16
we have

\[
\sum_{k=0}^{1/\varepsilon-1} \cos^{-1} F_{\min} \left( \rho_{k\varepsilon}, \rho_{(k+1)\varepsilon} \right)
\leq \sum_{k=0}^{1/\varepsilon-1} \sqrt{2 \left( 1 - F_{\min} \left( \rho_{k\varepsilon}, \rho_{(k+1)\varepsilon} \right) \right)} \sqrt{1 + \frac{1}{6} \left( 1 - F_{\min} \left( \rho_{k\varepsilon}, \rho_{(k+1)\varepsilon} \right) \right)}
\leq \sum_{k=0}^{1/\varepsilon-1} \sqrt{\frac{\varepsilon^2}{4} J_{k\varepsilon}^R + \frac{1}{3} \sup_{k\varepsilon \leq s \leq (k+1)\varepsilon} f_{1,t,s} \varepsilon^3} \sqrt{1 + \frac{\varepsilon^2}{12} \sup_{k\varepsilon \leq s \leq (k+1)\varepsilon} f_{2,t,s}}
\]

\[
= \sum_{k=0}^{1/\varepsilon-1} \frac{\varepsilon}{2} \left( J_{k\varepsilon}^R \right)^{1/2} \sqrt{1 + \frac{4}{3} \left( J_{k\varepsilon}^R \right)^{1/2} \sup_{k\varepsilon \leq s \leq (k+1)\varepsilon} f_{1,t,s} \varepsilon^3} \left( 1 + \frac{\varepsilon^2}{12} \sup_{k\varepsilon \leq s \leq (k+1)\varepsilon} f_{2,t,s} \right)
\]
\[
\leq \sum_{k=0}^{1/\varepsilon-1} \frac{\varepsilon}{2} \left( J_{k\varepsilon}^R \right)^{1/2} \sqrt{1 + \varepsilon f_3},
\]

where

\[
f_3 := \sup_{t,s,t',s',u \in [0,1]} \left( 1 + \frac{4}{3} \left( J_{k\varepsilon}^R \right)^{1/2} f_{1,t,s} \varepsilon^3 \right) \left( 1 + \frac{1}{12} f_{2,t',s'} \right) - 1 < \infty.
\]

Here, taking \( \varepsilon \to 0 \), we have the assertion. \( \blacksquare \)

### 12 Differential equation for shortest paths

In this section, a differential equation satisfied by the geodesic, or the path achieving the minimum in \( (27) \) is derived, supposing that \( \rho_t \) is an invertible matrix. Let

\[
\mathcal{L} \left( \rho_t, \frac{d\rho_t}{dt}, \lambda_t \right) := \int_C \left\{ J^R_t - \lambda_t \left( \text{tr} \rho_t - 1 \right) \right\} dt = \int_C \left\{ \text{tr} \left( \frac{d\rho_t}{dt} \right)^2 \rho_t^{-1} - \lambda_t \left( \text{tr} \rho_t - 1 \right) \right\} dt.
\]

Taking \( t \) proportional to arc length, this is equivalent to finding the extremal of \( \int_C \left\{ \sqrt{J^R_t} - \lambda_t \left( \text{tr} \rho_t - 1 \right) \right\} dt \). Then, letting \( \{X_t\} \) be an arbitrary smooth
curve with $X_0 = X_1 = 0$,

$$
\frac{d}{d\varepsilon} \mathcal{L} \left( \rho_t + \varepsilon X_t, \frac{d\rho_t}{dt} + \varepsilon \frac{dX_t}{dt}, \lambda_t \right) \bigg|_{\varepsilon = 0} = \int_C \left[ \text{tr} \left( L_t^R + L_t^{R_i} \right) \right] \, dt
$$

$$
= \left[ \text{tr} \rho_t \left( L_t^R + L_t^{R_i} \right) \right]_{\varepsilon = 0} + \int_C \text{tr} X_t \left\{ -L_t^{R_i} L_t^R + \lambda_t \right\} \, dt.
$$

Hence, we have

$$
- \frac{d}{dt} \left( L_t^R + L_t^{R_i} \right) - L_t^{R_i} L_t^R + \lambda_t = 0, \tag{30}
$$

where the first identity is by $J^R_t = 1$. In the sequel, the following identity is used frequently.

$$
\rho_t L_t^{R_i} = (L_t^R \rho_t)^\dagger = \left( \frac{d\rho_t}{dt} \right)^\dagger = L_t^R \rho_t. \tag{31}
$$

Multiplying $\rho$ and taking trace of both ends of (30),

$$
- \text{tr} \rho_t \frac{d}{dt} \left( L_t^R + L_t^{R_i} \right) - 1 + \lambda_t = 0,
$$

where we used $J^R_t = 1$, (31), and $\text{tr} \rho_t L_t^R = \text{tr} \rho_t L_t^{R_i} = \text{tr} \frac{d\rho_t}{dt} = 0$. Observe

$$
\text{tr} \rho_t \frac{d}{dt} \left( L_t^R + L_t^{R_i} \right) = \frac{d}{dt} \text{tr} \rho_t \left( L_t^R + L_t^{R_i} \right) - \text{tr} \frac{d\rho_t}{dt} \left( L_t^R + L_t^{R_i} \right)
$$

$$
= - \text{tr} \left( \rho_t L_t^{R_i} L_t^R + L_t^R \rho_t L_t^{R_i} \right) = -2.
$$

Therefore, $\lambda_t = -1$, and we obtain

$$
\frac{d}{dt} \left( L_t^R + L_t^{R_i} \right) + L_t^{R_i} L_t^R + 1 = 0, \tag{32}
$$

which gives only determines time derivative of only Hermitian part of $L_t^R$. Obviously, we need another equation. By (31), we have

$$
\frac{d\rho_t}{dt} L_t^{R_i} + \rho_t \frac{dL_t^{R_i}}{dt} = L_t^R \frac{d\rho_t}{dt} + \frac{dL_t^R}{dt} \rho_t.
$$

Observing

$$
\frac{d\rho_t}{dt} L_t^{R_i} = L_t^R \rho_t L_t^{R_i} = L_t^R \frac{d\rho_t}{dt},
$$

we obtain

$$
\rho_t \frac{dL_t^{R_i}}{dt} = \frac{dL_t^R}{dt} \rho_t.
$$
and
\[ \rho_t \frac{dL_t^R}{dt} + \frac{dL_t^R}{dt} \rho_t + \rho_t L_t^R \rho_t + \rho_t = 0. \]  
(33)

and
\[ \frac{d\rho_t}{dt} = L_t^R \rho_t, \]
determine time evolution of $\rho_t$.

The differential equation satisfied by the curve achieving minimum in (3) is derived by applying these to commutative case,
\[ 2 \frac{dl_t}{dt} + (l_t)^2 + 1 = 0, \quad \frac{dp_t}{dt} = l_t p_t. \]  
(34)

From these, the differential equation for the curve achieving minimum in (28) is derived as follows. Along the curve, we should have
\[ L_t^R = NL_t N^{-1}, \quad \rho_t = N (D_t)^2 N^\dagger, \]
where $(D_t)^2 = \text{diag}(p_t(1), \ldots, p_t(d))$, $L_t = \text{diag}(l_t(1), \ldots, l_t(d))$. Since $F_{\min}(\rho_0, \rho_1) = F(p_0, p_1)$, $(p_t, l_t)$ should satisfy (34). Therefore, dynamics of $(\rho_t, L_t^R)$ is determined by
\[ 2 \frac{dL_t^R}{dt} + (L_t^R)^2 + 1 = 0, \quad \frac{d\rho_t}{dt} = L_t^R \rho_t. \]  
(35)

13 Another quantum analogy of statistical distance

Statistical distance $\Delta(p, q)$ is defined by
\[ \Delta(p, q) := \frac{1}{2} \| p - q \|_1 = \frac{1}{2} \sum_{x \in X} |p(x) - q(x)|. \]

Its frequently used quantum analogy is
\[ \Delta(\rho, \sigma) := \frac{1}{2} \| \rho - \sigma \|_1, \]
and also called statistical distance. Known facts about them are [3][5][6][9][11]:

- $\Delta(\rho, \sigma) = \max_{M\text{ measurement}} \Delta(M(\rho), M(\sigma))$.

- (Monotonicity by CPTP maps) If $\Lambda$ is a CPTP map,
\[ \Delta(\Lambda(\rho), \Lambda(\sigma)) \leq \Delta(\rho, \sigma) \]

- (Joint convexity)
\[ \Delta(\lambda \rho_0 + (1 - \lambda) \rho_1, \lambda \sigma_0 + (1 - \lambda) \sigma_1) \leq \lambda \Delta(\rho_0, \sigma_0) + (1 - \lambda) \Delta(\rho_1, \sigma_1) \]
• (Triangle inequality)

\[ \Delta(p, \sigma) \leq \Delta(p, \tau) + \Delta(\tau, \sigma) \]

• With \( D(\rho|\sigma) := \text{tr} \rho (\ln \rho - \ln \sigma) \) being relative entropy,

\[ 1 - F(p, \sigma) \leq \Delta(p, \sigma) \leq \sqrt{1 - F(p, \sigma)^2}, \quad (36) \]

\[ D(\rho|\sigma) \geq \frac{1}{2} \Delta(p, \sigma)^2. \quad (37) \]

Here we introduce a new quantum analogue of statistical distance is:

\[ \Delta_{\text{max}}(\rho, \sigma) := \min \{ \Phi, \{p, q\} \} \text{; reverse test of } \{p, q\} \Delta(p, q). \]

**Theorem 16** Suppose \( \Delta^Q(p, \sigma) \) satisfies monotonicity by CPTP maps and \( \Delta^Q(p, q) = \Delta(p, q) \) for any probability distributions \( p, q \). Then

\[ \Delta(p, \sigma) \leq \Delta^Q(p, \sigma) \leq \Delta_{\text{max}}(p, \sigma). \]

Also, \( \Delta_{\text{max}} \) is monotone by CPTP maps and \( \Delta_{\text{min}}(p, q) = \Delta(p, q) \).

**Proof.** Almost parallel with the proof of Theorem 1, thus omitted. □

**Theorem 17** Defining \( D^R(\rho|\sigma) := \text{tr} \rho \ln \sqrt{\rho \sigma^{-1}} \sqrt{\rho} \),

\[ \Delta_{\text{max}}(\lambda \rho_0 + (1-\lambda) \rho_1, \lambda \sigma_0 + (1-\lambda) \sigma_1) \leq \lambda \Delta_{\text{max}}(\rho_0, \sigma_0) + (1-\lambda) \Delta_{\text{max}}(\rho_1, \sigma_1), \quad (38) \]

\[ 1 - F_{\text{min}}(p, \sigma) \leq \Delta_{\text{max}}(p, \sigma) \leq \sqrt{1 - F_{\text{min}}(p, \sigma)^2}, \quad (39) \]

\[ D^R(\rho|\sigma) \geq \frac{1}{2} \Delta_{\text{max}}(p, \sigma)^2. \quad (40) \]

**Proof.** The proof of (38) is almost parallel with the one of Theorem 8 thus omitted. To prove the first inequality of (39), consider the optimal reverse test with \( \Delta(p, q) = \Delta_{\text{max}}(p, \sigma) \). Then, by (36), we have

\[ \Delta_{\text{max}}(p, \sigma) = \Delta(p, q) \geq 1 - F(p, q), \]

On the other hand, by definition of \( F_{\text{min}}, 1 - F(p, q) \geq 1 - F_{\text{max}}(p, \sigma) \). After all, we have \( \Delta_{\text{max}}(p, \sigma) \geq 1 - F_{\text{min}}(p, \sigma) \). The second inequality of (39) and (40) are proved almost parallelly, recalling the following characterization of \( D^R(\rho|\sigma) \) (Theorem 2.4 in [8]):

\[ D^R(\rho|\sigma) = \min D(\rho|q), \]

where minimum is taken over all the reverse test \( \{\Phi, \{p, q\}\} \) of \( \{\rho, \sigma\} \). □

It follows from (39) and (40) that

\[ \Delta_{\text{max}}(|\varphi\rangle, |\psi\rangle) = 1, \]

for all \( |\varphi\rangle, |\psi\rangle \). Another consequence of (39) is:
Proposition 18 The triangle inequality for \( \Delta_{\text{max}} \) does not hold, i.e., there is \( \rho, \sigma, \) and \( \tau \) with
\[
\Delta_{\text{max}}(\rho, \sigma) > \Delta_{\text{max}}(\rho, \tau) + \Delta_{\text{max}}(\tau, \sigma).
\]

Proof. Let \( \rho \) and \( \sigma \) be as of the proof of Proposition 6. Then,
\[
\Delta_{\text{max}}(\rho, \sigma) = 1,
\]
\[
\Delta_{\text{max}}(\rho, \tau) + \Delta_{\text{max}}(\tau, \sigma) \leq \sqrt{1 - F_{\text{min}}(\rho, \tau)^2} + \sqrt{1 - F_{\text{min}}(\tau, \sigma)^2}
\]
\[
= \frac{2}{\sqrt{1 - (|\cos \frac{\theta}{2}| + |\sin \frac{\theta}{2}|)^2}}.
\]

Hence, making \( \theta \) close enough to 0, we have the assertion.

In general, \( \Delta_{\text{max}}(\rho, \sigma) \) is hard to compute. But when \( \sigma = |\varphi\rangle \langle \varphi| \), one can compute the number as follows. Any reverse test \( (\Phi, \{p, q\}) \) of \( \{\rho, |\varphi\rangle\} \) is in the form of (22). Then, with \( c := \sum_{x \in \text{supp } q} p(x) \),
\[
\|p - q\|_1 \geq \frac{1}{2}|c - 1| + \frac{1}{2}|(1 - c) - 0| = 1 - c,
\]
where the inequality is due to monotonicity of \( \|\cdot\|_1 \), and is achieved by \( p(x) = c q(x)(x \in \text{supp } q) \). Therefore, to minimize \( \|p - q\|_1 \), one has to maximize \( c \), which can be any positive number satisfying \( \rho - c |\varphi\rangle \langle \varphi| \geq 0 \), or equivalently,
\[
1 - c \rho^{-1/2} |\varphi\rangle \langle \varphi| \rho^{-1/2} \geq 0.
\]
Therefore,
\[
\Delta_{\text{max}}(\rho, |\varphi\rangle) = 1 - \frac{1}{\|\rho^{-1/2} |\varphi\rangle\|^2} = 1 - F(\rho, |\varphi\rangle)^2.
\]

Also, by the argument in Section , we have the following upperbound to \( \Delta_{\text{max}}(\rho, \sigma) \):
\[
\Delta_{\text{max}}(\rho, \sigma) \leq \Delta(M(\rho), M(T\rho T)) \leq \Delta(\rho, T\rho T),
\]
where \( T \) is as of (14), and \( M \) is the projectors onto eigenspaces of \( T \). (Note the left most end is upperbounded by
\[
\sqrt{1 - F(\rho, T\rho T)^2} = \sqrt{1 - F_{\text{min}}(\rho, \sigma)^2},
\]
and gives a better bound than (39).)

14 Discussions

\( F_{\text{fmin}} \) introduced in this paper resembles \( g \)-quasi relative entropy [10]
\[
S_g(\rho, \sigma) := \text{tr} \rho^{1/2} g \left(L_{\sigma} R_{\rho}^{-1}\right) \left(\rho^{1/2}\right),
\]
where $g$ is operator convex function (thus, $-g$ is operator monotone) on $[0, \infty)$, and

$$L_\sigma(A) = \sigma A, \quad R_\rho(A) = A \rho.$$  

For example, with $g = 1 - x^{1/2}$, it gives rise to $S_\alpha(\rho, \sigma) := 1 - \text{tr} \rho^{1-\alpha} \sigma^\alpha$. This quantity satisfies

$$S_\alpha(\rho, \sigma) \geq 1 - F_{\alpha}^{\text{min}}(\rho, \sigma),$$

where

$$F_{\alpha}^{\text{min}}(\rho, \sigma) := \text{tr} \rho^{1/2} T^{2(1-\alpha)} \rho^{1/2} = \text{tr} \rho^{1/2} \left( \rho^{-1/2} \sigma \rho^{-1/2} \right)^{1-\alpha} \rho^{1/2},$$

and the equality does not hold in general.

It is interesting that both of the maximum of $F_f(p, q)$ are achieved by minimal reverse tests. Some numerics suggests that this is not the case for the minimum of $\Delta(p, q)$.

An open question is whether $F_R$ satisfies strong joint convexity or not. Also, more explicit formula for $F_R$ and $\Delta_{\text{max}}$ would be, even for some special cases, of importance.

References

[1] F. Kubo and T. Ando, “Means of positive linear operators,” Mathematische Annalen, vol. 246, pp. 205–224 (1980).

[2] R. Bhatia, ”Matrix Analysis”, Springer-Verlag (1996).

[3] C. A. Fuchs, ”Distinguishability and Accessible Information in Quantum Theory”, Doctoral dissertation, University of New Mexico (1996).

[4] Jurgen Jost,” Riemannian Geometry and Geometric Analysis” Fifth Edition, Springer-Verlag (2008).

[5] K. Matsumoto, A Geometry of Quantum States, Master Thesis, University of Tokyo, Master’s thesis (1995).

[6] K. Matsumoto, A Geometrical Approach to Quantum Estimation Theory, Doctoral dissertation, University of Tokyo (1998).

[7] K. Matsumoto, ”Reverse Estimation Theory, Complementarity between RLD and SLD, and Monotone Distances”, [http://xxx.lanl.gov/abs/quant-ph/0511170](http://xxx.lanl.gov/abs/quant-ph/0511170).

[8] K. Matsumoto, ”Reverse test and Characterization of Quantum Relative Entropy”, [http://xxx.lanl.gov/abs/1006.0302](http://xxx.lanl.gov/abs/1006.0302).

[9] M. A. Nielsen, I. L. Chuang,” Quantum Computation and Quantum Information”, Cambridge University Press (2000).
[10] D. Petz, ”Quasi-Entropies for Finite Quantum Systems”, Reports on Mathematical Physics, 23, 57 (1986)

[11] A. Uhlmann, “Density operators as an arena for differential geometry,” Rep. Math. Phys., vol.33, pp.253-263(1993).