Trace-distance measure of coherence

Swapan Rana,1,* Preeti Parashar,2 and Maciej Lewenstein1,3

1ICFO – Institut de Ciències Fotòniques, The Barcelona Institute of Science and Technology, 08860 Castelldefels (Barcelona), Spain
2Physics and Applied Mathematics Unit, Indian Statistical Institute, 203 BT Road, Kolkata, India
3ICREA – Institució Catalana de Recerca i Estudis Avançats, 08010 Barcelona, Spain

(Dated: January 22, 2016)

We show that trace distance measure of coherence is a strong monotone for all qubit and, so called, $X$ states. An expression for the trace distance coherence for all pure states and a semi definite program for arbitrary states is provided. We also explore the relation between $l_1$-norm and relative entropy based measures of coherence, and give a sharp inequality connecting the two. In addition, it is shown that both $l_p$-norm- and Schatten-$p$-norm-based measures violate the (strong) monotonicity for all $p \in (1, \infty)$.

PACS numbers: 03.65.Aa, 03.67.Mn, 03.65.Ud

I. INTRODUCTION

It is an established fact that quantum mechanical systems differ in many counter intuitive ways from classical systems. The figure of merit is generally attributed to coherence, i.e., the possibility of quantum mechanical superpositions, which on the level of density matrix description of quantum mechanical states correspond to off-diagonal density matrix elements in the computational or measurement selected basis. Many approaches have been proposed to encompass this important feature since the inception of quantum mechanics. Only very few measures (e.g., negativity, relative entropy of entanglement, Bures’ distance) are known to obey Eq. (1). Similarly, very few functions (mostly exact analogues of those entanglement measures) are known to be coherence measures.

One of the widely used distinguishability measures, the trace distance, has been proposed as a possible candidate for coherence measure in Ref. [2]. It is formally defined as

$$C_u(\rho) := \min_{\delta \in I} \| \rho - \delta \|_1,$$

where $I$ is the set of incoherent states. The question is whether $C_u(\rho)$ satisfies Eq. (1) for all $\rho$.

Despite its omnipresence in quantum information theory, it is not yet known whether the trace distance measure of entanglement $E_u(\rho)$ [which is defined like Eq. (2) with $I$ replaced by $\mathcal{S}$, the set of separable states] satisfies strong monotonicity, even for the simple case of 2-qubits. Under the extra assumption that the closest separable states share the same marginal states with $\rho$, $E_u(\rho)$ has been shown to satisfy strong monotonicity [10]. The difficulty of this problem is reflected by the fact that in general the closest state $\delta$ can not be determined explicitly, and even if the dimension of $\rho$ is small, the dimension of $\rho_n$ may be arbitrarily high. Nonetheless, as $I$ has a trivial structure compared to $\mathcal{S}$, it seems that answering this question could be easier for $C_u(\rho)$. Here we show that $C_u(\rho)$ satisfies Eq. (1) for all single qubit states $\rho$.

Apart from some measures defined through the convex-roof construction [4, 11], the $l_1$-norm-based measure $C_l$ and the relative entropy based measure $C_r$ [defined later in Eq. (22) and Eq. (13), respectively] are the only known coherence measures satisfying the strong monotonicity for all states. Due to its close similarity with relative entropy of entanglement $E_r$, $C_l$ has a clear physical meaning and is the cornerstone of resource theory of coherence [3]. In contrast, $C_r$ has neither an exact analog with an entanglement measure, nor any physical interpretation yet. It is thus desirable and interesting to find any interrelation between them, which hopefully would give some bound on $C_l$ in terms of $C_r$. Recently, also the Hilbert-Schmidt distance has been conjectured [12] to be a coherence measure—we show that this is not the case.

* swapanqc@gmail.com
The organization of this article is as follows: In Sec. II we study the properties of trace distance coherence and prove its strong monotonicity for qubits. We present here also a semi definite program for general states. The interrelation between $C_t$ and $C_e$ is described in Sec. III. It is shown that $C_t$ is an upper bound for $C_e$ for all pure states and qubit states. In Sec. IV we show that for all $p \in (1, \infty)$ neither $C_t$, nor $C_e$ satisfies (strong) monotonicity. We conclude with a short discussion of our results and outlook in Sec. V.

II. TRACE DISTANCE COHERENCE

A. Qubit and X-states

To find the analytic form of trace distance coherence we have to find the (not necessarily unique) closest incoherent state. For a qubit the nearest incoherent state is just $\rho_{\text{diag}}$ [13]. This can be seen easily: if $\rho = \frac{1}{2} \left( \begin{array}{cc} 1 & \delta \\ \delta & 1 \end{array} \right)$, then $\rho - \delta$, being Hermitian and traceless, will have eigenvalues $\pm \delta$. So, the required trace distance is $2\delta$, and to minimize it, we have to minimize the determinant of $\rho - \delta$, i.e.,

$$\min_{\rho - \delta} |\rho - \delta| = \min_{\rho} (\text{det} \rho) = \lambda_1 \lambda_2.$$

Thus we are minimizing over a larger set. This proposition immediately implies the following facts:

**Corollary 2.** Let $A_1, A_2$ be $2 \times 2$ complex matrices and $x_i, y_i$ be complex numbers. Then

$$C_{tr} \left( \bigoplus_{i,j} x_i A_{i} \mp x_j y_j \right) = \sum_{i} |x_i| C_{tr}(A_i) = C_t \left( \bigoplus_{i,j} x_i A_{i} \mp y_j \right).$$

This improves the theorem of Ref. [13], in the sense that it is readily applicable to direct sum of qubits. States of the form $x \oplus A$ were considered therein.

**Corollary 3.** The strong monotonicity is satisfied by $C_{tr}(\rho)$ for any $2 \times 2$ matrix $\rho$.

In Ref. [13], the authors showed that if the dimensions of the Kraus operators are restricted to three, then the strong monotonicity is satisfied by all qubit states. We show here that this is always true irrespective of the dimensions of the Kraus operators involved.

Using the fact that $||A||_{p} \leq ||A||_{1}$ for any matrix $A$ and $1 \leq p \leq 2$ [14, pp. 50], we have

$$C_{tr}(\rho_{n}) = ||\rho_{n} - \delta_{*}^{*}||_{1} \leq ||\rho_{n} - \text{diag}(\rho_{n})||_{1} \leq ||\rho_{n} - \text{diag}(\rho_{n})||_{1} = C_{t}(\rho_{n}).$$

So multiplying by $\rho_{n}$, summing over $n$, using the fact that $C_t$ satisfies the strong monotonicity condition [2], and finally $C_t(\rho) = C_{tr}(\rho)$, proves the result. By the same reasoning, strong monotonicity is satisfied for all matrices $A$ with $C_{tr}(A) = C_{t}(A)$, in particular the matrices in Eq. (5). □

Let us now mention an interesting class of states, the so-called X-states, albeit we do not assume anything (not even normality) except its shape, for which $C_t$ has an analytic expression, also satisfies strong monotonicity.

**Proposition 4.** Let $X$ be an $n \times n$ complex matrix with non-zero elements only along its diagonal and anti-diagonal, $x_{ij} = 0$ for $j \neq i, n + 1 - i$. The nearest diagonal matrix to $X$ in trace norm is given by $\text{diag}(X)$. Therefore $C_{tr}(X) = C_{t}(X)$ and hence $C_{tr}(X)$ satisfies strong monotonicity.
While calculating trace norm, the matrix \( X - \delta \) is a special class of the matrices appearing in (5), and hence the result follows from Cor. 2 and Cor. 3.

We should mention that calculation of trace distance coherence for a very specific class of \( X \)-states (with only three real parameters) has been considered in recent literature [15, 16].

B. Pure states

Finding the closest incoherent state becomes intractable just beyond qubits. For a pure state \( |\psi\rangle \), the intuitively expected nearest incoherent state is \( \delta = \text{diag}(|\psi\rangle \langle \psi|) \). Unfortunately, this is not necessarily true for dimension higher than 2. As an example, for \( |\psi\rangle = \delta = 2/3|0\rangle + 2/3|1\rangle + 1/3|2\rangle \), \( \text{diag}(1/2, 1/2, 0) \) is closer than \( \text{diag}(|\psi\rangle \langle \psi|) \). We will now show why it is difficult to have an analytic formula, even for the simple case of pure qutrits.

Let \( |\psi\rangle = \sum_i \sqrt{x_i}|i\rangle \) be given (if required, we remove any phase by a diagonal unitary, which is an incoherent operation) and let \( \delta = \sum_i \delta_i |i\rangle \langle i| \) be its nearest diagonal state. Then by Weyl’s inequality [17, pp. 62] \( \lambda_j(A - B) \leq \lambda_j^+(A) - \lambda_j^-(B) \), the matrix \( H = |\psi\rangle \langle \psi| - \delta \) has exactly one positive eigenvalue. Let it be \( \lambda \). Since \( H \) is traceless, the sum of the rest of its eigenvalues must be \(-\alpha\), and hence \( \|H\|_1 = 2\alpha \). The problem is thus to find the maximum (as only one is positive) eigenvalue of \( H \) and minimize it with respect to \( \delta_i \)’s.

As usual, we have to solve the characteristic equation for \( H \), namely, \( \det(xI - H) = 0 \). So, let us first calculate the determinant (see [18] for more general case). Writing

\[
xI - H = \begin{pmatrix}
    x + \delta_1 & \sum_{j=2}^d \delta_j & \cdots & \delta_d \\
    \sum_{j=2}^d \delta_j & x + \delta_2 & \cdots & \delta_d \\
    \vdots & \vdots & \ddots & \vdots \\
    \delta_d & \sum_{j=2}^d \delta_j & \cdots & x + \delta_d
\end{pmatrix} - |\psi\rangle \langle \psi|,
\]

we use the Sherman-Morrison-Woodbury formula for determinants [19, pp. 19]:

\[
\det(A + uv^\top) = (1 + v^\top A^{-1}u) \det A.
\]

Therefore the required determinant is

\[
\det(xI - H) = \prod_{i=1}^d \left(1 - \sum_{j=1}^d \frac{\lambda_j}{x + \delta_i}\right).\]

For positive roots, the out-most factors are nonzero and we obtain the equation

\[
\sum_{i=1}^d \frac{\lambda_i}{x + \delta_i} = 1. \tag{7}
\]

Eq. (7) can be viewed as a (monic) polynomial equation in \( x \) of degree \( d \). We have to find its largest root (all roots are real) and then minimize that with respect to \( \delta_i \). Unless \( d = 2 \) (where the roots are of the form \( b \pm \sqrt{b^2 - 4c} \)/2, thereby the largest root is the one with the + sign), there is no simple way to characterize the largest root \( x^* \), and hence in general, no simple way to get a general explicit expression for \( C_\alpha \). Note also that we can consider \( \lambda_i \neq 0 \), if some \( \lambda_i = 0 \), then the problem is reduced to the case with \( d = \# \{\lambda_i \neq 0\} \).

Nonetheless, the above analysis is quite useful since we have

\[
C_\alpha(|\psi\rangle) = 2 \min_{\delta_i \geq 0, \sum_\delta_i = 1} \max_x \left\{ \sum_{i=1}^d \frac{\lambda_i}{x + \delta_i} = 1 \right\}. \tag{8}
\]

The right hand side (RHS) of Eq. (8) can be written as the following optimization problem:

Minimize \( 2 \left( \sum_{i=1}^d \frac{\lambda_i}{\delta_i} - 1 \right) \)

subject to \( \sum_{i=1}^d \delta_i \leq 1 \), \( \forall i = 1, 2, \ldots, d \), \( \delta_i \geq 0 \), \( \forall i = 1, 2, \ldots, d \). \tag{9}

To see this equivalence, first note that Eq. (8) could be rewritten as

\[
C_\alpha(|\psi\rangle) = 2 \min_{\delta_i \geq 0, \sum_\delta_i = 1, x \geq 0} \left\{ \sum_{i=1}^d \frac{\lambda_i}{x + \delta_i} \leq 1 \right\}. \tag{10}
\]

Let \( x^* > 0 \) and \( \delta^* = (\delta_1^*, \delta_2^*, \ldots, \delta_d^*) \) be the optimal value of \( x \), and \( \delta \) in the RHS of (10). Define

\[
\delta_i := \frac{\lambda_i}{(x^* + \delta_i^*)}, \quad i = 1, 2, \ldots, d. \tag{11a}
\]

Then \( x^* + \delta_i^* = \frac{\lambda_i}{\delta_i}, \quad i = 1, 2, \ldots, d \). \tag{11b}

Summing Eq. (11b) and using \( \sum \delta_i^* = 1 \), we have

\[
x^* = \frac{1}{d} \left( \sum_{i=1}^d \lambda_i / \delta_i - 1 \right),
\]

and hence from Eq. (11b) and Eq. (11a),

\[
\frac{\lambda_i}{\delta_i} \geq x^* = \frac{1}{d} \left( \sum_{i=1}^d \frac{\lambda_i}{\delta_i} - 1 \right) (> 0), \quad i = 1, 2, \ldots, d.
\]

Thus the solution of Eq. (10) corresponds to the solution of Eq. (9). Conversely, if \( \delta^* = (\delta_1^*, \delta_2^*, \ldots, \delta_d^*) \) is the optimal values of \( \delta \) in Eq. (9), then one easily verifies that

\[
x^* := \frac{1}{d} \left( \sum_{i=1}^d \lambda_i / \delta_i^* - 1 \right), \quad \delta_i := \frac{\lambda_i}{\delta_i^*} - x^*, \quad i = 1, 2, \ldots, d
\]

correspond to the optimal \( x \) and \( \delta \) in Eq. (10).
C. Arbitrary states

In contrast to trace distance entanglement, we could formulate a semi-definite program to calculate \( C_t(\rho) \) for any arbitrary state \( \rho \). The main idea is that any Hermitian matrix \( \rho \) can be written as a difference of two positive semi definite matrices, \( \rho = \rho^+ - \rho^- \), with \( \rho^\pm \geq 0 \). Then \( \| \rho \|_1 = \text{Tr}(\rho^+ + \rho^-) \) minimized over all such decompositions of \( \rho \). Thus \( C_t(\rho) \) is the optimal value of the following semi definite problem (SDP) [20]:

\[
\begin{align*}
\text{Minimize } \quad & \text{Tr}(P + N) \\
\text{subject to } \quad & P - N = \rho - \delta, \\
& \text{Tr} \delta = 1, \\
& \delta \quad \text{is diagonal}, \\
& P, N, \delta \geq 0.
\end{align*}
\]

We have used this SDP to check the strong monotonicity for random states (however, we were not able to generate the incoherent channels uniformly). Despite our numerical and analytic attempts, no examples violating strong monotonicity were found. This leads us to conjecture that strong monotonicity of \( C_t \) is satisfied by all states.

III. RELATION BETWEEN \( C_t \) AND \( C_r \)

Analogously to the relative entropy of entanglement, the relative entropy of coherence is defined [2] as

\[
C_r(\rho) := \min_{\delta \in I} S(\rho \| \delta).
\]

The minimization could be solved analytically [2], leading to \( C_r(\rho) = S(\rho_{\text{diag}}) - S(\rho) \), where \( S(\rho) = -\text{Tr}(\rho \log_2 \rho) \) is the von Neumann entropy. Note that for pure states \( C_t \) is somewhat like the negativity \( N \) [21] of a bipartite state,

\[
C_t(\rho) := \sum_i \sqrt{\lambda_i} |i\rangle = \left( \sum_i \sqrt{\lambda_i} \right)^2 - 1 = 2N(\rho) := \sum_i \sqrt{\lambda_i} |i\rangle.
\]

In entanglement theory, relations between \( E_r \) and \( N \) have been studied extensively (albeit mainly for two qubits, see, e.g., [22–24]). The aim of this section is to derive interrelations between \( C_r \) and \( C_t \).

A. Pure states

For pure qubit states, the relation is \( C_t \geq C_r \), which is exactly the well known upper bound for the binary entropy function

\[
2 \sqrt{x(1-x)} \geq H_b(x) := -x \log_2 x - (1-x) \log_2(1-x).
\]

For higher dimensional pure states, we will exploit two known results — one from entanglement theory and the other from information theory. It is well known [21] that the logarithmic negativity \( E_N := \log_2(1+2N) \) is an upper bound on distillable entanglement which coincides with \( E_r \) for pure states,

\[
\begin{align*}
\log_2 (1 + 2N(\rho)) & \geq E_r(\rho) \\
& \Rightarrow \log_2 (1 + C_t(\rho)) \geq C_r(\rho) \\
& \Rightarrow C_t(\rho) \geq 2^{C_r(\rho)} - 1.
\end{align*}
\]

Note that this bound is tight in the sense that equality holds for maximally coherent states in any dimension.

There is another simple inequality between \( C_t \) and \( C_r \), namely \( C_t \geq C_r \) for all pure states. Although generally not sharp, this inequality is independent from that in Eq. (14). To prove it, note that it follows [25] from the recursive property of entropy function,

\[
\begin{align*}
\frac{1}{2} H(A) & \leq \sum_{i=1}^{d-1} \sqrt{\lambda_i} \sum_{j=i+1}^{d} \lambda_j \leq \sum_{i=1}^{d-1} \sqrt{\lambda_i} \left( \sum_{j=i+1}^{d} \sqrt{\lambda_j} \right) \\
& \Rightarrow C_r(\rho) \leq C_t(\rho).
\end{align*}
\]

Combining Eqs. (14)-(15) we have the following result.

**Proposition 5.** For all pure states \( \rho \),

\[
C_t(\rho) \geq \max \left\{ C_r(\rho), \ 2^{C_r(\rho)} - 1 \right\}.
\]

The variation of these bounds could be visualized for arbitrary qutrits. In this case, the \( \lambda_i \)'s can be taken as \( x, (1-x)y, (1-x)(1-y) \) and Fig. 1 shows the plot of \( C_t \) and \( C_r \) as a function of \( x, y \in [0, 1]^2 \). Note that \( C_t \geq C_r \) gives independent bound than that of Eq. (14). For example, let \( x = 1/500, \ y = 1/5 \), then \( C_t(\rho) \approx 0.9182, C_r \approx 0.7413 \), while the bound in Eq. (14) gives \( C_t \approx 0.6717 \). On the other hand Eq. (14) gives equality for all maximally coherent states.

![FIG. 1. (Color online) \( C_t(\rho) \) and \( C_r(\rho) \) for general qutrit \( \rho \).](image)

Note also that, Eq. (16) improves the known bound on distillable entanglement (in terms of logarithmic negativity), for all pure states.

B. Arbitrary states

As usual, finding a better bound is more difficult for mixed states. Since the \( C_t \) measure does not have any role in entanglement theory so far, we could use the inequality \( C_t \geq C_r \),
resulting in some rough bounds, due to the proportionality constant already introduced in this step. It is then tempting to use Fannes’s inequality [26], but unfortunately it gives nothing useful:

\[ C_r(\rho) = S(\rho|\rho_{\text{diag}}) \]
\[ \leq |\rho - \rho_{\text{diag}}|_1 \log_2 d + \frac{1}{e \ln 2} \]
\[ \leq C_t(\rho) \log_2 d + \frac{1}{e \ln 2} \]  

(17)

The relation between \( C_t \) and \( C_r \) could be drastically sharpened [than the one mentioned in Eq. (17)] using Fannes-Audenaert bound [27]; unfortunately this bound is not monotonic in \( C_t \) and hence not applicable to \( C_t \).

It turns out that we can use an inequality between Holevo information \( \chi \) and trace norm. For an ensemble \( E := \{p_i, \rho_i\} \), the Holevo information is defined as

\[ \chi(E) := S\left( \sum_i p_i \rho_i \right) - \sum_i p_i S(\rho_i), \]

and it satisfies [28]

\[ \chi(E) \leq H(\rho) t, \quad t := \max_{i,j} |\rho_i - \rho_j|_1 / 2. \]  

(18)

The next ingredient we will use is the fact that for any square matrix \( X \), there are sets of diagonal unitary matrices \( \{U_k\} \) such that

\[ \text{diag}(X) = \frac{1}{r} \sum_{k=0}^{r-1} U_k X U_k^\dagger. \]  

(19)

At least two such sets of unitaries are known [29], one with \( r = d \) = order of \( X \), and \( U_k = U^k, U = \text{diag}(1, \omega, \omega^2, \ldots, \omega^{d-1}), \omega := e^{2\pi i/d} \). The other one is with \( r = 2^d \) and \( U_k \)'s are \( \text{diag}(\pm1, \pm1, \ldots, \pm1) \). Since the second choice involves more terms, in general it leads to inferior bounds. Employing these tools, it follows that

\[ C_r(\rho) = S(\text{diag}(\rho)) - S(\rho) \]
\[ = \chi\left( \frac{1}{d}, U_k \rho U_k^\dagger \right) \]
\[ \leq t \log_2 d. \]  

(20)

Now let the maximum in the definition of \( t \) occur for the pair \( \rho_i, \rho_j \). Since \( \{U_k\} \) forms a multiplicative group, we will have \( U_i^j U_j = U_k \) for some \( k \). Then

\[ 2t = ||U_i \rho U_i^j - U_j \rho U_j^i||_1 = ||\rho - U_j \rho U_j^i||_1 \]
\[ \leq ||\rho - U_k \rho U_k^i||_1 \]
\[ \leq 2C_t(\rho). \]

Plugging into Eq. (20), we get the following result.

**Proposition 6.** For any \( d \)-dimensional state \( \rho \),

\[ C_r(\rho) \leq \log_2 d \ C_t(\rho). \]  

(21)

Note that for qubits it is already sharp, coinciding with the bound for pure states. Our numerical study suggests that the inequality could be sharpened to just \( C_t \geq C_r \), but we could not manage to get rid of this rather annoying multiplicative factor. We thus make the following conjecture.

**Conjecture 7.** For all states \( \rho \),

\[ C_t(\rho) \geq C_r(\rho). \]

**IV. ALL OTHER \( l_p \)-NORM AND SCHATTEN-\( p \)-NORM**

For an \( m \times n \) matrix \( X = (x_{ij}) \) and \( p \in [1, \infty) \), the \( l_p \)-norm and Schatten-\( p \) norms are usually defined as

\[ ||X||_p := \left( \sum_{i,j} |x_{ij}|^p \right)^{1/p}, \]
\[ ||X||_p := (\text{Tr}|X|^p)^{1/p} = \left( \sum_i \sigma_i^p \right)^{1/p}, \]

where \( \sigma_i \)'s are the non-zero singular values of \( X \), i.e., eigenvalues of \( |X| := \sqrt{X^\dagger X} \), and \( r \) is the rank of \( X \). The coherence measure based on the distance induced by these norms are defined as

\[ C_{l_p}(\rho) := \min_{\delta \in \mathcal{F}} ||\rho - \delta||_{l_p}, \]
\[ C_p(\rho) := \min_{\delta \in \mathcal{F}} ||\rho - \delta||_p. \]  

(22a)

(22b)

In Ref. [2], the authors have shown that \( C_t \) satisfies strong monotonicity (and \( C_1 = C_t \) is the subject of this paper). They have also considered coherence measure based on the distance induced by the square of \( l_2 \)-norm and gave an example to show that it does not satisfy strong monotonicity. Although a coherence measure need not be induced by a norm (e.g., \( C_r \) is based on relative entropy which is neither a distance for being asymmetric in its arguments, nor a metric for violating triangular inequality), the counterexample provided in Ref. [2] does not violate strong monotonicity if we take just the \( l_2 \)-norm, instead of its square. Based on this observation it has been conjectured in Ref. [12] that \( l_2 \)-norm induces a legitimate coherence measure. In this section we will show that it is not the case. We will prove the following result:

**Proposition 8.** For all \( p \in (1, \infty) \), there are states violating strong monotonicity for both the measures \( C_{l_p} \) and \( C_p \), thereby neither is a good measure of coherence.

Before presenting our counterexample, let us mention that \( ||.||_{l_2}^2 \) (in general \( ||.||_{l_p}^p \) for \( 1 < p < \infty \)) need not be a norm, as it does not satisfy the triangular inequality

\[ ||a + b||_{l_2} \leq ||a||_{l_2}^2 + ||b||_{l_2}^2. \]

(It is not necessarily true when \( a, b \) are tensors, matrices, vectors, complex numbers, or even real numbers). The homogeneity condition of a norm is violated by \( ||.||_{l_2}^2 \). This is the
reason for the apparent violation of monotonicity by the counterexample provided in Ref. [2]. Indeed the combined state and channel provided in the example satisfies strong monotonicity inequality for any $l_p$-norm. In particular, with those $\{K_n\}$, all states (qutrit, for the dimensions of $K_n$’s) satisfies the strong monotonicity in $l_2$-norm, as

$$\sum_{i=1}^{2} p_i C_{i_l}(\rho_i) = p_2 C_{i_l}(\rho_2) = \frac{\sqrt{2}}{2} |\beta| |c| + |a| |e|$$

$$\leq \sqrt{2} \sqrt{|\beta|^2 + |c|^2 + |e|^2}$$

$$= C_{i_l}(\rho) = \rho = \begin{pmatrix} a & b & c \\ \bar{b} & \bar{q} & \bar{e} \\ \bar{c} & \bar{f} \end{pmatrix}.$$

However, with this judicious choice of $\{K_n\}$ with $a = \beta$, and $p$ with $b = 0$, $c = e$, the strong monotonicity inequality for $l_2$-norm becomes $2^2 \leq 2^2$, which is violated by all $p \in (2, \infty)$.

It is well known [30] that the distance induced by $l_2$-norm (see also [31] for Schatten-$p$ norms) is not contractive under CPTP maps. Since a coherence measure has to be contractive under (incoherent) CPTP maps, there is no reason to think of $C_{i_l}$ to be a good measure of coherence. To end this discussion, we give the following counterexample:

$$\rho = \frac{1}{2} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, K_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, K_2 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Then

$$\rho_1 = \frac{1}{p_1} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \rho_2 = \frac{1}{p_2} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, p_1 = p_2 = \frac{1}{2}.$$

In order for $\rho$ to be a state, we must have $|a|, |b| \leq 1$. The strong monotonicity for the $C_{i_l}$ measure reads

$$(|a| + |b|)^p \leq |a|^p + |b|^p,$$

which is violated [32] by the entire class with $ab \neq 0$ and for all $p \in (1, \infty)$.

Now we move to the calculation for $C_p$. It turns out that we do not need to calculate anything further. Note that if we assume that the matrices in Proposition 1 are Hermitian (to ensure $\sigma_i = |\lambda_i|$, then we have $C_{i_l}(A) = C_p(A)$ for all $p$ and a Hermitian $2 \times 2$ matrix $A$. Since $\rho_i$’s are effectively in $2 \times 2$, we have $C_{i_l}(\rho_i) = C_p(\rho_i)$. Similarly, $C_{i_l}(\rho) = C_p(\rho)$, as $\rho$ is a matrix of the form given in Eq. (5). Thus strong monotonicity for $C_p$ is also violated for all $ab \neq 0$ and $p \in (1, \infty)$.

Up to now we were concerned about only strong monotonicity of $C_p$ and $C_{i_l}$. It appears that for $p \in (1, \infty)$ none is a monotone in the first place.

**Proposition 9.** For $p \in (1, \infty)$, neither $C_p$ nor $C_{i_l}$ is a monotone.

Note that this result is stronger than Proposition 8, because, convexity together with strong monotonicity implies monotonicity. So, if a convex function is not a monotone, it can not be a strong monotone.

It also appears that we can give a general method to construct counterexample from any coherent state [31]. Before doing so, we note that it follows from the result of Ref. [31], that $C_p$ is monotone for all qubit states and for all $p \in [1, \infty)$. So the counterexample should be in dimension higher than 2.

The states themselves being incoherent, there is an incoherent channel transforming $I/d$ to $|0\rangle\langle 0|$. For instance, consider the Kraus operators $K_i = |0\rangle\langle i|$, $i = 1, 2, \ldots, d$. Now let $\rho$ be a given coherent state and $\Lambda^l$ be the incoherent channel with Kraus operators $K_i = 1 \otimes K_i$. Then we have

$$C_p(\Lambda^l(\rho \otimes I/d)) = C_p(\rho \otimes |0\rangle\langle 0|)$$

$$= C_p(\rho) > C_p(\rho \otimes I/d).$$

In the last line we have used

$$C_p(\rho \otimes I/d) \leq ||\rho \otimes I/d - \delta^* \otimes I/d||_p = C_p(\rho)||I/d||_p < C_p(\rho).$$

Noticing that $C_{i_l}(\rho \otimes I/d) = d^{1/p-1} C_{i_l}(\rho) < C_{i_l}(\rho)$, $C_{i_l}$ also violates monotonicity.

**V. DISCUSSION AND CONCLUSION**

Although originated in entanglement theory, strong monotonicity is not a necessary requirement for entanglement measures, but rather an extra feature. In contrast, every coherence measure has to satisfy strong monotonicity. It would be interesting to study the effect of relaxing this constraint. Also restricting the Kraus operators to have same dimension as that of the original state would be worth looking at.

The strong monotonicity of a convex entanglement measure is known to be equivalent to its local unitary invariance and flag condition [34]. In Ref. [2] a quite different flag condition has been mentioned as an extra feature of a coherence measure. Since trace norm is factorizable under tensor products, it follows that if the strong monotonicity holds for $C_{i_l}$, then it will also satisfy the flag condition:

$$C_{i_l} \left( \sum_{l} p_l \rho_l \otimes |i\rangle\langle i| \right) \leq C_{i_l}(\rho).$$

However, this does not help to resolve the main question, and despite the frequent appearance of trace distance in literature, it (at least for $E_{il}$) remains quite a frustrating open problem.

Before concluding, let us mention some relevance of our Conjecture 7. As was mentioned earlier, $C_{i_l}$ does not have any physical interpretation yet. In some recent works [35, 36], $C_{i_l}$ has been shown to be connected with the success probability of unambiguous state discrimination in interference experiments. If the conjectured relation $C_{i_l}(\rho) \geq C_{i_l}(\rho)$ holds for all states, then it would probably be the best physical interpretation for $C_{i_l}$. It will then be analogous to (logarithmic)
negativity in entanglement theory, providing an upper bound for distillable coherence (which coincides with $C_r(\rho)$ for all $\rho$ [3]).

\section*{ACKNOWLEDGMENTS}

We would like to thank Andreas Winter and Koenraad M. R. Audenaert for helpful discussion and correspondence. We also thank Remigiusz Augusiak, Alexander Streltsov, and Manabendra Nath Bera for insightful comments, and Shuming Cheng for bringing Ref. [25] to our attention. SR and ML acknowledge financial support from the John Templeton Foundation, the EU grants OSYRIS (ERC-2013-AdG Grant No. 339106), QUIC (H2020-FETPROACT-2014 No. 641122), and SIQS (FP7-ICT-2011-9 No. 600645), the Spanish MINECO grants FOQUS (FIS2013-46768-P) and “Severo Ochoa” Programme (SEV-2015-0522), and the Generalitat de Catalunya grant 2014 SGR 874.

[1] J. Åberg, arXiv:0612146v1 (2006).
[2] T. Baumgratz, M. Cramer, and M. B. Plenio, Phys. Rev. Lett. 113, 140401 (2014).
[3] A. Winter and D. Yang, arXiv:1506.07975v2 (2015).
[4] A. Streltsov, U. Singh, H. S. Dhar, M. N. Bera, and G. Adesso, Phys. Rev. Lett. 115, 020403 (2015).
[5] E. Chitambar, A. Streltsov, S. Rana, M. N. Bera, G. Adesso, and M. Lewenstein, arXiv:1507.08171v1 (2015).
[6] A. Streltsov, S. Rana, M. N. Bera, and M. Lewenstein, arXiv:1509.07456v2 (2015).
[7] E. Chitambar and M.-H. Hsieh, arXiv:1509.07458v1 (2015).
[8] F. G. S. L. Brandão, M. Horodecki, M. B. Plenio, and S. Virmani, Open Sys. & Information Dyn. 14, 333 (2007).
[9] M. Fukuda and M. M. Wolf, J. Math. Phys. 48, 072101 (2007).
[10] J. Eisert, K. Audenaert, and M. B. Plenio, J. Phys. A: Math. Theor. 36, 5605 (2003).
[11] X. Qi, Z. Bai, and S. Du, arXiv:1505.07387v1 (2015).
[12] S. Cheng and M. J. W. Hall, Phys. Rev. A 92, 042101 (2015).
[13] L.-H. Shao, Z. Xi, H. Fan, and Y. Li, Phys. Rev. A 91, 042120 (2015).
[14] X. Zhan, Matrix Inequalities, ISBN: 3-540-43798-3, Springer-Verlag Berlin, Germany (2002).
[15] T. R. Bromley, M. Cianciaruso, and G. Adesso, Phys. Rev. Lett. 114, 210401 (2015).
[16] M. Cianciaruso, T. R. Bromley, and G. Adesso, arXiv:1507.01600v1 (2015).
[17] R. Bhatia, Matrix Analysis, Springer-Verlag, New York, USA, (1997).
[18] J. Anderson, Linear Algebra Appl. 246, 49 (1996).
[19] R. A. Horn and C. R. Johnson, Matrix Analysis, Second Ed., Cambridge University Press, Cambridge, USA, (2012).
[20] A. Winter, Private communication.
[21] G. Vidal and R. F. Werner, Phys. Rev. A 65, 032314 (2002).
[22] F. Verstraete, K. Audenaert, J. Dehaene, and B. De Moor, J. Phys. A: 34, 10327 (2001).
[23] A. Miranowicz and A. Grudka, J. Opt. B: Quantum Semiclass. Opt. 6, 542 (2004).
[24] A. Miranowicz, S. Ishizaka, B. Horst, and A. Grudka, Phys. Rev. A 78, 052308 (2008).
[25] J. Lin, IEEE Trans. Inf. Theory, 37(1), 145 (1991).
[26] M. Fannes, Commun. Math. Phys. 31, 291 (1973).
[27] K. M. R. Audenaert, J. Math. A: Math. Theor. 40, 8127 (2007).
[28] K. M. R. Audenaert, J. Math. Phys. 55, 112202 (2014).
[29] R. Bhatia, Amer. Math. Monthly 107, 602 (2000).
[30] M. Ozawa, Phys. Lett. A 268, 158 (2000).
[31] D. Perez-Garcia, M. M. Wolf, D. Petz, and M. B. Ruskai, J. Math. Phys. 47, 083506 (2006).
[32] A well-known simple trick to prove the inequality: for $x, y, \epsilon > 0$, $(x + y)^{1+\epsilon} = x(x + y)^\epsilon + y(x + y)^\epsilon > x \cdot x^{1+\epsilon} + y \cdot y^{1+\epsilon}$.
[33] We thank Alexander Streltsov for sharing this idea with us.
[34] M. Horodecki, Open Sys. & Information Dyn. 12, 231 (2005).
[35] M. N. Bera, T. Qureshi, M. A. Siddiqui, and A. K. Pati, Phys. Rev. A 92, 012118 (2015).
[36] E. Bagan, J. A. Bergou, M. Hillery, arXiv:1509.04592v1 (2015).