PHYSICAL REVIEW A **93**, 012110 (2016)

Trace-distance measure of coherence

Swapan Rana,[1,*] Preeti Parashar,[2] and Maciej Lewenstein[1,3]
[1]ICFO – Institut de Ciències Fot'oniques, The Barcelona Institute of Science and Technology, 08860 Castelldefels (Barcelona), Spain
[2]Physics and Applied Mathematics Unit, Indian Statistical Institute, 203 BT Road, Kolkata, India
[3]ICREA – Institució Catalana de Recerca i Estudis Avançats, 08010 Barcelona, Spain
(Received 10 November 2015; published 12 January 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 semidefinite 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 $p \in (1,\infty)$.

DOI: 10.1103/PhysRevA.93.012110

I. INTRODUCTION

It is an established fact that quantum mechanical systems differ in many counterintuitive 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 recently, a resource-theoretic framework for coherence has been put forward in Refs. [1,2], and has been subsequently developed [3], and advanced further in Refs. [4–7].

Quantification and interrelations between quantifiers are important aspects in any resource theory. In general, distance based functions are expected to be good quantifiers, subject to the restriction imposed by the theory. In the formalism presented in Ref. [2], a nonnegative convex function C defined on the space of states ρ, acting on d dimensional Hilbert space, is called a coherence measure if it satisfies the following two conditions:

(1) Monotonicity under incoherent channel Λ^I:

$$C(\Lambda^I[\rho]) \leqslant C(\rho),$$

(2) Strong monotonicity under incoherent channel Λ^I:

$$\sum_n p_n C(\rho_n) \leqslant C(\rho), \tag{1}$$

where $\rho_n := (K_n\rho K_n^\dagger)/p_n$, $p_n := \mathrm{Tr}(K_n\rho K_n^\dagger)$, K_n's are $d_n \times d$ incoherent Kraus operators satisfying $\sum_n K_n^\dagger K_n = \mathbb{1}_d$. In this work we will study mainly Eq. (1) for some functions which have been proposed as possible coherence measures.

In addition to its defining property that the resource should not increase on average under the *free* operations, strong monotonicity has important consequences for the additivity question of convex entanglement measures [8,9]; additivity property in turn simplifies further similar questions for many related information-theoretic quantities. In entanglement theory 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 analogs 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

$$C_{\mathrm{tr}}(\rho) := \min_{\delta\in\mathcal{I}} \|\rho-\delta\|_1, \tag{2}$$

where $\mathcal{I}$ is the set of incoherent states. The question is whether $C_{\mathrm{tr}}(\rho)$ satisfies Eq. (1) for all ρ.

Despite its omnipresence in quantum information theory, it is not yet known whether the trace distance measure of entanglement $E_{\mathrm{tr}}(\rho)$ [which is defined like Eq. (2) with $\mathcal{I}$ 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 ρ, $E_{\mathrm{tr}}(\rho)$ has been shown to satisfy strong monotonicity [10]. The difficulty of this problem is reflected by the fact that in general the closest state δ cannot be determined explicitly, and even if the dimension of ρ is small, the dimension of ρ_n may be arbitrarily high. Nonetheless, as $\mathcal{I}$ has a trivial structure compared to $\mathcal{S}$, it seems that answering this question could be easier for $C_{\mathrm{tr}}(\rho)$. Here we show that $C_{\mathrm{tr}}(\rho)$ satisfies Eq. (1) for all single qubit states ρ.

Apart from some measures defined through the convex-roof construction [4,11], the l_1-norm-based measure C_{l_1} and the relative entropy based measure C_r [defined later in Eqs. (22) and (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_r has a clear physical meaning and is the cornerstone of resource theory of coherence [3]. In contrast, C_{l_1} 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_1} 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.

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 semidefinite program for general states. The interrelation between C_{l_1} and C_r is described in Sec. III. It is shown that C_{l_1}

*swapanqic@gmail.com

is an upper bound for C_r for all pure states and qubit states. In Sec. IV we show that for all $p \in (1,\infty)$ neither C_{l_p} nor C_p 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 ρ_{diag} [13]. This can be seen easily: if $\delta = \text{diag}\{x, 1-x\}$ is the nearest incoherent state to $\rho = \left(\begin{smallmatrix} p & q \\ \bar{q} & 1-p \end{smallmatrix}\right)$, then $\rho - \delta$, being Hermitian and traceless, will have eigenvalues $\pm\lambda$. So, the required trace distance is 2λ, and to minimize it we have to minimize the determinant of $\rho - \delta$ (of course with a negative sign, since the roots are $\pm\lambda$), which is simply $|q|^2 + (p-x)^2$, with respect to x. Hence the result follows.

However, if we consider qutrits, then the expression for eigenvalues of $\rho - \delta$, as well as the optimization becomes very messy. As we will show later, finding an *analytic form* even for pure qutrits is almost intractable. But, we could still find analytic expression for C_{tr} (and test the strong monotonicity) for some classes of high-dimensional states. In doing so, we will sometimes optimize over larger sets of general matrices. For this purpose, we extend the definitions of C_{tr} and C_{l_1} to a square matrix X [see also Eq. (22)], i.e.,

$$C_{\text{tr}}(X) := \min_{D\in\Delta} \|X - D\|_1, \tag{3a}$$

$$C_{l_1}(X) := \min_{D\in\Delta} \|X - D\|_{l_1} = \sum_{i\neq j} |X_{ij}|, \tag{3b}$$

Δ being the set of diagonal matrices. The following result will be used extensively.

Proposition 1. Let A be a 2×2 matrix with complex entries and D be its closest diagonal matrix in trace norm. Then $D = \text{diag}(A)$ and hence $C_{\text{tr}}(A) = C_{l_1}(A)$.

Note that the trace norm of $A - D$ is nonnegative and always bounded in finite dimension, e.g., by sum of individual trace norms. Hence we do not require any further restriction (like normality, positivity) on A, D.

As usual, let $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$, $D = \left(\begin{smallmatrix} x & 0 \\ 0 & y \end{smallmatrix}\right)$. If the singular values of $A - D$ are σ_1 and σ_2, then

$$\begin{aligned}
\|A - D\|_1^2 &= (\sigma_1 + \sigma_2)^2 \\
&= \text{Tr}[(A-D)(A-D)^\dagger] + 2|\det(A-D)| \\
&= |a-x|^2 + |d-y|^2 + |b|^2 + |c|^2 \\
&\quad + 2|(a-x)(d-y) - bc| \\
&\geqslant |a-x|^2 + |d-y|^2 + |b|^2 + |c|^2 \\
&\quad + 2||(a-x)||(d-y)| - |b||c|| \\
&\geqslant (|b| + |c|)^2,
\end{aligned} \tag{4}$$

with equality iff $D = \text{diag}(A)$. In the last inequality we have used the fact that $p^2 + q^2 + 2|pq - r| \geqslant 2|r|$, and equality holds iff $p = q = 0$. ■

The purpose of considering general matrices instead of states is to get rid of the positivity and the unit trace conditions. Thus we are minimizing over a larger set. This proposition immediately implies the following facts:

Corollary 2. Let A_i be 2×2 complex matrices and x_i, y_j be complex numbers. Then

$$C_{\text{tr}}\Big(\bigoplus_i x_i A_i \bigoplus_j y_j\Big) = \sum_i |x_i| C_{\text{tr}}(A_i) = C_{l_1}\Big(\bigoplus_i x_i A_i \bigoplus_j y_j\Big). \tag{5}$$

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 $C_{\text{tr}}(\rho)$ for any 2×2 matrix ρ.

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 $\|X\|_p \leqslant \|X\|_{l_p}$ for any matrix X and $1 \leqslant p \leqslant 2$ [14, p. 50], we have

$$\begin{aligned}
C_{\text{tr}}(\rho_n) &= \|\rho_n - \delta_n^\star\|_1 \\
&\leqslant \|\rho_n - \text{diag}(\rho_n)\|_1 \\
&\leqslant \|\rho_n - \text{diag}(\rho_n)\|_{l_1} = C_{l_1}(\rho_n).
\end{aligned} \tag{6}$$

So multiplying by p_n, summing over n, using the fact that C_{l_1} satisfies the strong monotonicity condition [2], and finally $C_{l_1}(\rho) = C_{\text{tr}}(\rho)$, proves the result. By the same reasoning, strong monotonicity is satisfied for all matrices A with $C_{\text{tr}}(A) = C_{l_1}(A)$, 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_{tr} has an analytic expression, also satisfies strong monotonicity.

Proposition 4. Let X be an $n\times n$ complex matrix with nonzero elements only along its diagonal and antidiagonal, $x_{ij} = 0$ for $j \neq i, n+1-i$. The nearest diagonal matrix to X in trace norm is given by $\text{diag}(X)$. Therefore $C_{\text{tr}}(X) = C_{l_1}(X)$ and hence $C_{\text{tr}}(X)$ 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 $|\psi\rangle = 2/3|0\rangle + 2/3|1\rangle + 1/3|2\rangle$, $\text{diag}\{1/2, 1/2, 0\}$ 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 $|\psi\rangle = \sum_i \sqrt{\lambda_i}|i\rangle$ 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, p. 62] $\lambda^\downarrow_{i+j-1}(A - B) \leqslant$

$\lambda_i^\downarrow(A) - \lambda_{n-j+1}^\downarrow(B)$, the matrix $H = |\psi\rangle\langle\psi| - \delta$ has exactly one positive eigenvalue. Let it be α. 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 δ_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

$$xI - H = \begin{pmatrix} x+\delta_1 & & & \\ & x+\delta_2 & & \\ & & \ddots & \\ & & & x+\delta_d \end{pmatrix} - |\psi\rangle\langle\psi|,$$

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

$$\det(\mathbf{A} + \mathbf{u}\mathbf{v}^t) = (1 + \mathbf{v}^t\mathbf{A}^{-1}\mathbf{u})\det\mathbf{A}.$$

Therefore the required determinant is

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

For positive roots, the outmost factors are nonzero and we obtain the equation

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

Equation (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 δ_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 $x^\star$, and hence in general, no simple way to get a general explicit expression for $C_{\rm tr}$. 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

$$C_{\rm tr}(|\psi\rangle) = 2 \min_{\delta_i \geqslant 0,\, \sum\delta_i = 1} \max_{x \geqslant 0} \left\{\sum_{i=1}^{d} \frac{\lambda_i}{x+\delta_i} = 1\right\}. \tag{8}$$

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

$$\begin{aligned} &\text{Minimize } \frac{2}{d}\left(\sum_{i=1}^{d}\frac{\lambda_i}{\delta_i} - 1\right) \\ &\text{subject to } \begin{cases} \frac{\lambda_i}{\delta_i} \geqslant \frac{1}{d}\left(\sum_{i=1}^{d}\frac{\lambda_i}{\delta_i} - 1\right), \forall i = 1,2,\ldots,d, \\ \sum_{i=1}^{d}\delta_i \leqslant 1, \\ \delta_i \geqslant 0, \forall i = 1,2,\ldots,d. \end{cases} \end{aligned} \tag{9}$$

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

$$C_{\rm tr}(|\psi\rangle) = 2 \min_{\delta_i \geqslant 0,\, \sum\delta_i = 1,\, x>0} \left\{\sum_{i=1}^{d} \frac{\lambda_i}{x+\delta_i} \leqslant 1\right\}. \tag{10}$$

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

$$\delta_i := \frac{\lambda_i}{(x^\star + \delta_i^\star)}, \quad i = 1,2,\ldots,d, \tag{11a}$$

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

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

$$x^\star = \frac{1}{d}\left(\sum_{i=1}^{d}\frac{\lambda_i}{\delta_i} - 1\right),$$

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

$$\frac{\lambda_i}{\delta_i} \geqslant x^\star = \frac{1}{d}\left(\sum_{i=1}^{d}\frac{\lambda_i}{\delta_i} - 1\right) (>0), \quad i = 1,2,\ldots,d.$$

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

$$x^\star := \frac{1}{d}\left(\sum_{i=1}^{d}\frac{\lambda_i}{\delta_i^\star} - 1\right), \quad \delta_i := \frac{\lambda_i}{\delta_i^\star} - x^\star, \quad i = 1,2,\ldots,d$$

correspond to the optimal x and δ in Eq. (10).

C. Arbitrary states

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

$$\begin{aligned} &\text{Minimize } \mathrm{Tr}(P+N) \\ &\text{subject to } \begin{cases} P - N = \rho - \delta, \\ \mathrm{Tr}\,\delta = 1, \\ \quad \delta \text{ is diagonal}, \\ P,N,\delta \geqslant 0. \end{cases} \end{aligned} \tag{12}$$

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 $C_{\rm tr}$ is satisfied by all states.

III. RELATION BETWEEN C_{l_1} AND C_r

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

$$C_r(\rho) := \min_{\delta \in \mathcal{I}} S(\rho\|\delta). \tag{13}$$

The minimization could be solved analytically [2], leading to $C_r(\rho) = S(\rho_{\rm diag}) - S(\rho)$, where $S(\rho) = -\mathrm{Tr}(\rho\log_2\rho)$ is the von Neuman entropy. Note that for pure states C_{l_1} is somewhat like the negativity $\mathcal{N}$ [21] of a bipartite state

$$C_{l_1}\Big(|\psi\rangle := \sum \sqrt{\lambda_i}|i\rangle\Big) = \Big(\sum \sqrt{\lambda_i}\Big)^2 - 1$$
$$= 2\mathcal{N}\Big(|\phi\rangle := \sum \sqrt{\lambda_i}|ii\rangle\Big).$$

In entanglement theory, relations between E_r and $\mathcal{N}$ 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_{l_1}.

A. Pure states

For pure qubit states, the relation is $C_{l_1} \geqslant C_r$, which is exactly the well known upper bound for the binary entropy function

$$2\sqrt{x(1-x)} \geqslant H_b(x) := -x \log_2 x - (1-x)\log_2(1-x).$$

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 $E_{\mathcal{N}} := \log_2(1 + 2\mathcal{N})$ is an upper bound on distillable entanglement which coincides with E_r for pure states,

$$\begin{aligned} \log_2 [1 + 2\mathcal{N}(|\phi\rangle)] &\geqslant E_r(|\phi\rangle) \\ \Rightarrow \log_2 \left[1 + C_{l_1}(|\psi\rangle)\right] &\geqslant C_r(|\psi\rangle) \\ \Rightarrow C_{l_1}(|\psi\rangle) &\geqslant 2^{C_r(|\psi\rangle)} - 1. \end{aligned} \tag{14}$$

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_{l_1} and C_r, namely $C_{l_1} \geqslant C_r$ 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,

$$\begin{aligned} \frac{1}{2}H(\lambda) \leqslant \sum_{i=1}^{d-1} \sqrt{\lambda_i \sum_{j=i+1}^{d} \lambda_j} \leqslant \sum_{i=1}^{d-1} \sqrt{\lambda_i}\left(\sum_{j=i+1}^{d} \sqrt{\lambda_j}\right) \\ \Rightarrow C_r(|\psi\rangle) \leqslant C_{l_1}(|\psi\rangle). \end{aligned} \tag{15}$$

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

Proposition 5. For all pure states $|\psi\rangle$,

$$C_{l_1}(|\psi\rangle) \geqslant \max\{C_r(|\psi\rangle),\, 2^{C_r(|\psi\rangle)} - 1\}. \tag{16}$$

The variation of these bounds could be visualized for arbitrary qutrits. In this case, the λ_i's can be taken as $x,(1-x)y,(1-x)(1-y)$ and Fig. 1 shows the plot of C_{l_1} and C_r as a function of $x,y \in [0,1]^2$. Note that $C_{l_1} \geqslant C_r$ gives independent bound than that of Eq. (14). For example, let $x = 1/500$, $y = 1/5$, then $C_{l_1}(|\psi\rangle) = 0.9182$, $C_r = 0.7413$, while the bound in Eq. (14) gives $C_{l_1} \geqslant 0.6717$. On the other hand Eq. (14) gives equality for all maximally coherent states.

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

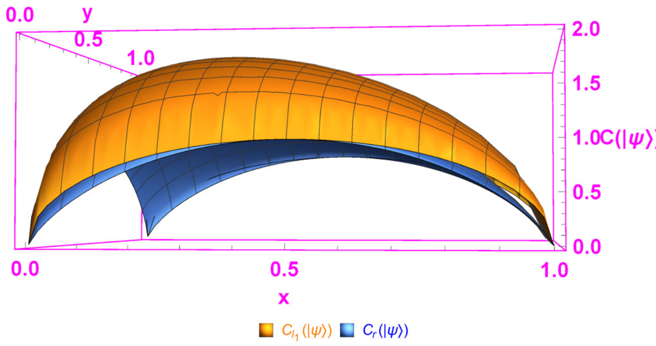

FIG. 1. $C_{l_1}(|\psi\rangle)$ and $C_r(|\psi\rangle)$ for general qutrit $|\psi\rangle$.

B. Arbitrary states

As usual, finding a better bound is more difficult for mixed states. Since the C_{l_1} measure does not have any role in entanglement theory so far, we could use the inequality $C_{l_1} \geqslant C_{\text{tr}}$, 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:

$$\begin{aligned} C_r(\rho) &= S(\rho\|\rho_{\text{diag}}) \\ &\leqslant \|\rho - \rho_{\text{diag}}\|_1 \log_2 d + \frac{1}{e\ln 2} \\ &\leqslant C_{l_1}(\rho)\log_2 d + \frac{1}{e\ln 2}. \end{aligned} \tag{17}$$

The relation between C_{tr} 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_{tr} and hence not applicable to C_{l_1}.

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

$$\chi(\mathcal{E}) := S\left(\sum_i p_i\rho_i\right) - \sum_i p_i S(\rho_i),$$

and it satisfies [28]

$$\chi(\mathcal{E}) \leqslant H(p)t, \quad t := \max_{i,j} \|\rho_i - \rho_j\|_1/2. \tag{18}$$

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

$$\text{diag}(X) = \frac{1}{r}\sum_{k=0}^{r-1} U_k X U_k^\dagger. \tag{19}$$

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 $\text{diag}\{\pm 1, \pm 1, \ldots, \pm 1\}$. Since the second choice involves more terms, in general it leads to inferior bounds. Employing these tools, it follows that

$$\begin{aligned} C_r(\rho) &= S[\text{diag}(\rho)] - S(\rho) \\ &= \chi\left(\left\{\frac{1}{d}, U_k\rho U_k^\dagger\right\}\right) \\ &\leqslant t\log_2 d. \end{aligned} \tag{20}$$

Now let the maximum in the definition of t occur for the pair ρ_i,ρ_j. Since $\{U_k\}$ forms a multiplicative group, we will have $U_i^\dagger U_j = U_k$ for some k. Then

$$\begin{aligned}2t &= \|U_i\rho U_i^\dagger - U_j\rho U_j^\dagger\|_1 = \|\rho - U_k\rho U_k^\dagger\|_1\\ &\leqslant \|\rho - U_k\rho U_k^\dagger\|_{l_1}\\ &\leqslant 2C_{l_1}(\rho).\end{aligned}$$

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

Proposition 6. For any d-dimensional state ρ,

$$C_r(\rho) \leqslant \log_2 d\ \ C_{l_1}(\rho). \tag{21}$$

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 $C_{l_1} \geqslant C_r$, 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 ρ,

$$C_{l_1}(\rho) \geqslant C_r(\rho).$$

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

$$\|X\|_{l_p} := \left(\sum_{i,j}|x_{ij}|^p\right)^{1/p},$$

$$\|X\|_p := (\mathrm{Tr}\,|X|^p)^{1/p} = \left(\sum_i^r \sigma_i^p\right)^{1/p},$$

where σ_i's are the nonzero 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

$$C_{l_p}(\rho) := \min_{\delta\in\mathcal{I}} \|\rho-\delta\|_{l_p}, \tag{22a}$$

$$C_p(\rho) := \min_{\delta\in\mathcal{I}} \|\rho-\delta\|_p. \tag{22b}$$

In Ref. [2] the authors have shown that C_{l_1} satisfies strong monotonicity (and $C_1 = C_{\mathrm{tr}}$ 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 $\|\cdot\|_2^2$ (in general $\|\cdot\|_{l_p}^p$ for $1<p<\infty$) need not be a norm, as it does not satisfy the triangular inequality

$$\|a+b\|_{l_2}^2 \leqslant \|a\|_{l_2}^2 + \|b\|_{l_2}^2.$$

(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 $\|\cdot\|_{l_2}^2$. 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

$$\begin{aligned}\sum_{i=1}^2 p_i C_{l_2}(\rho_i) &= p_2 C_{l_2}(\rho_2)\\ &= \sqrt{2}(|\beta||c| + |\alpha||e|)\\ &\leqslant \sqrt{2}\sqrt{|b|^2+|c|^2+|e|^2}\\ &= C_{l_2}\left(\rho = \begin{bmatrix} a & b & c\\ \bar b & q & e\\ \bar c & \bar e & f\end{bmatrix}\right).\end{aligned}$$

However, with this judicious choice of $\{K_n\}$ with $\alpha=\beta$, and ρ with $b=0$, $c=e$, the strong monotonicity inequality for l_p-norm becomes $2^p \leqslant 2^2$, 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_{l_2} to be a good measure of coherence. To end this discussion, we give the following counterexample:

$$\rho = \frac14\begin{pmatrix}1&0&a&0\\0&1&0&b\\ \bar a&0&1&0\\0&\bar b&0&1\end{pmatrix},\quad K_1 = \begin{pmatrix}0&0&0&0\\1&0&0&0\\0&0&0&0\\0&0&1&0\end{pmatrix},$$

$$K_2 = \begin{pmatrix}0&1&0&0\\0&0&0&0\\0&0&0&1\\0&0&0&0\end{pmatrix}.$$

Then

$$\rho_1 = \frac{1}{p_1}\begin{pmatrix}0&0&0&0\\0&\frac14&0&\frac a4\\0&0&0&0\\0&\frac{\bar a}{4}&0&\frac14\end{pmatrix},\quad \rho_2 = \frac{1}{p_2}\begin{pmatrix}\frac14&0&\frac b4&0\\0&0&0&0\\ \frac{\bar b}{4}&0&\frac14&0\\0&0&0&0\end{pmatrix},$$

$$p_1 = p_2 = \frac12.$$

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

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

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 $C_{l_p}(A) = C_p(A)$ for all p and a Hermitian 2×2 matrix A. Since ρ_i's are effectively in 2×2, we have $C_{l_p}(\rho_i) = C_p(\rho_i)$. Similarly, $C_{l_p}(\rho) = C_p(\rho)$, as ρ 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 monotonicty of C_p and C_{l_p}. 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_{l_p} 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 cannot be a strong monotone.

It also appears that we can give a general method to construct counterexample from any coherent state [33]. Before doing so, we note that it follows from the result of Ref. [31], that C_p is a 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 $\mathbb{1}/d$ to $|0\rangle\langle 0|$. For instance, consider the Kraus operators $K_i = |0\rangle\langle i-1|$, $i = 1,2,\ldots,d$. Now let ρ be a given coherent state and Λ^I be the incoherent channel with Kraus operators $\tilde{K}_i = \mathbb{1} \otimes K_i$. Then we have

$$\begin{aligned} C_p(\Lambda^I[\rho \otimes \mathbb{1}/d]) &= C_p(\rho \otimes |0\rangle\langle 0|) \\ &= C_p(\rho) \\ &> C_p(\rho \otimes \mathbb{1}/d). \end{aligned}$$

In the last line we have used

$$\begin{aligned} C_p(\rho \otimes \mathbb{1}/d) &\leqslant \|\rho \otimes \mathbb{1}/d - \delta^\star \otimes \mathbb{1}/d\|_p \\ &= C_p(\rho)\|\mathbb{1}/d\|_p < C_p(\rho). \end{aligned}$$

Noticing that $C_{l_p}(\rho \otimes \mathbb{1}/d) = d^{1/p-1} C_{l_p}(\rho) < C_{l_p}(\rho)$, C_{l_p} 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_{tr}, then it will also satisfy the flag condition:

$$C_{\text{tr}}\left(\sum_i p_i \rho_i \otimes |i\rangle\langle i|\right) \leqslant C_{\text{tr}}(\rho).$$

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_{tr}) remains quite a frustrating open problem.

Before concluding, let us mention some relevance of our Conjecture 7. As was mentioned earlier, C_{l_1} does not have any physical interpretation yet. In some recent works [35,36], C_{l_1} has been shown to be connected with the success probability of unambiguous state discrimination in interference experiments. If the conjectured relation $C_{l_1}(\rho) \geqslant C_r(\rho)$ holds for all states, then it would probably be the best physical interpretation for C_{l_1}. 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 ρ [3]).

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. S.R. and M.L. acknowledge financial support from the John Templeton Foundation, the EU grants OSYRIS (ERC-2013-AdG Grant No. 339106), QUIC (H2020-FETPROACT-2014 Grant No. 641122), and SIQS (FP7-ICT-2011-9 Grant 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:quant-ph/0612146.
[2] T. Baumgratz, M. Cramer, and M. B. Plenio, Phys. Rev. Lett. **113**, 140401 (2014).
[3] A. Winter and D. Yang, arXiv:1506.07975.
[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:1507.08171.
[6] A. Streltsov, S. Rana, M. N. Bera, and M. Lewenstein, arXiv:1509.07456.
[7] E. Chitambar and M.-H. Hsieh, arXiv:1509.07458.
[8] F. G. S. L. Brandão, M. Horodecki, M. B. Plenio, and S. Virmani, Open Sys. Inf. 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:1505.07387.
[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* (Springer, Berlin, 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:1507.01600.

[17] R. Bhatia, *Matrix Analysis* (Springer, New York, 1997).
[18] J. Anderson, Linear Algebra Appl. **246**, 49 (1996).
[19] R. A. Horn and C. R. Johnson, *Matrix Analysis*, 2nd ed. (Cambridge University Press, Cambridge, 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**, 145 (1991).
[26] M. Fannes, Commun. Math. Phys. **31**, 291 (1973).
[27] K. M. R. Audenaert, J. Phys. A: Math. Theor. **40**, 8127 (2007).
[28] K. M. R. Audenaert, J. Math. Phys. **55**, 112202 (2014).
[29] R. Bhatia, Am. 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 $x,y,\epsilon > 0$, $(x+y)^{1+\epsilon} = x(x+y)^\epsilon + y(x+y)^\epsilon > x \cdot x^\epsilon + y \cdot y^\epsilon = x^{1+\epsilon} + y^{1+\epsilon}$.
[33] We thank Alexander Streltsov for sharing this idea with us.
[34] M. Horodecki, Open Sys. Inf. 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, and M. Hillery, arXiv:1509.04592.