1 Mathematical essentials and terminologies

Like in the main paper, for the reader’s convenience we compile a list of symbols and abbreviations; see Table 1. If interested, a more rigorous approach to this problem can be found in J.-Y Kao’s PhD thesis [1].

1.1 Linear mappings and quantum operations

We are only concerned with finite-dimensional Hilbert spaces, denoted by $\mathcal{H}$. The symbol $L$ is for the vector space formed by linear mappings, e.g. $L(H_1, H_2)$ is the space constituted by linear mappings from $H_1$ to $H_2$, $L(H) := L(H, H)$ is a space of operators, and Choi isomorphism $\mathcal{T}$ is an isomorphism between $L(L(H_1), L(H_2))$ and $L(H_1) \otimes L(H_2)$ [2].

The term “linear mapping” usually refers to one from an operator algebra to another, i.e. they constitute the space $L(L(H_1), L(H_2))$; the context dictates what they actually mean. The symbol $L$ is for a linear mapping, while $S$ and $S_i$ are for quantum operations and sub-operations thereof: $S$ and $S_i$ are CP. In general we use $S$ without a subscript for a TP (deterministic) quantum operation, but sometimes we’ll be a bit loose on this, as in Sec. 5.1.

1.2 Entangling capacity

Entangling capacity of a deterministic operation or sub-operation $S_i$ with respect to an entanglement measure $m$, when aided by an ancilla, is defined as [3, 4, 5, 6]

$$EC_m(S_i) := \max_{\rho} \{ m(I_a \otimes S_i(\rho)) - m(\rho) \},$$

where $m$ gives the same result whether the input state is normalized or not. For a probabilistic operation $S$ with sub-operations $S_i$, it’s

$$EC_m(S) := \max_{\rho} \left\{ \sum_i p_i m(I_a \otimes S_i(\rho)) - m(\rho) \right\}.$$  \(1\)

Here the ancilla space $H_a$ can be any finite-dimensional inner-product space, $I_a$ is the identity mapping on $L(H_a)$, and the maximizations are over all density operators in $L(H_a \otimes \mathcal{H})$.

In Ref. [6], entangling capacity was actually defined to be the supremum, rather than maximum over all states. However, if the measure $m$ is continuous, then since the space is finite-dimensional, physical states in the space of Hermitian operators is a compact set $\mathcal{K}$, and $m(\mathcal{K})$ is also compact [7, 8], so there’s no difference between the supremum and maximum.

1.3 Hilbert-Schmidt inner product

We make use of Hilbert-Schmidt inner product: for $O_i \in L(H_i)$

$$(O_2|O_1) := \text{tr}(O_2^\dagger O_1),$$

linear in the second argument, like bra-ket. It’s an inner product, because it satisfies all the properties for an inner product [9]. For a linear mapping $L : L(H_1) \to L(H_2)$ [2, 10],

$$(O_2|L(O_1)) = (L^\dagger(O_2)|O_1);$$  \(4\)
| Symbol | Definition |
|--------|------------|
| CP     | Completely positive |
| TP     | Trace preserving |
| HP     | Hermiticity preserving |
| PPT    | Positive partial transpose (preserving) |
| $\mathcal{H}$ | Hilbert space |
| EC$_m$ | Entangling capacity with respect to measure $m$ |
| $\mathcal{L}(V, W)$ and $\mathcal{L}(V)$ | Space formed by linear mappings from $V$ to $W$; $\mathcal{L}(V) = \mathcal{L}(V, V)$ |
| $L$   | A linear mapping from operators to operators |
| $\mathcal{F}$ | Choi isomorphism |
| $I$ and $\mathcal{I}$ | Identity operator and identity mapping |
| $T$ and $\Gamma$ | Transposition and partial transposition |
| $P$   | A positive operator |
| $H$   | A Hermitian operator |
| $O$   | A generic operator |
| $\tilde{H}^\pm$ | $H$ as a difference of positive operators |
| $H^\pm$ | Like above, but by eigendecomposition |
| $\tilde{L}^\pm$ | $L$ as the difference of two CP mappings |
| $L^\pm$ | Like above, but by eigendecomposition |
| $S$ and $S_i$ | Quantum operations and sub-operations, unless explicitly specified |
| $L^\Gamma$ | $\Gamma \circ L \circ \Gamma$ |
| $S_i^\Gamma$ | $S_i^\Gamma$ |
| $S_{i^\pm}$ | $S_{i^\pm}$ |
| $d_i$ | Dimension of Hilbert space $\mathcal{H}_i$ |
| $\|\cdots\|$ | Operator norm |
| $\|\cdots\|_p$ | Schatten $p$-norm. When $p = 1$ it’s trace norm. |
| ker | The kernel or null space of a linear mapping |
| ran | The range/image of a linear mapping |

Table 1: List of symbols and acronyms
\( L^\dagger : \mathcal{L}(\mathcal{H}_2) \to \mathcal{L}(\mathcal{H}_1) \) is the adjoint of \( L \). In particular, we’re interested in \( L^\dagger(I_2) \in \mathcal{L}(\mathcal{H}_1) \), because

\[
\text{tr}L(O_1) = (I_2[L(O_1)]) = (L^\dagger(I_2))[O_1].
\]  

(5)

The trace of an operator after a linear mapping \( L \) can be expressed in terms of an inner product between the operator and \( L^\dagger(I) \).

If \( L \) is HP, it always admits an operator-sum representation \([11, 12, 13]\):

\[
L(O_1) = \sum_i c_i V_i O_1 V_i^\dagger, \quad V_i \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2) \text{ and } c_i \in \mathbb{R}.
\]

(6)

Therefore

\[
L^\dagger(I) = \sum_i c_i V_i^\dagger V_i.
\]

(7)

From this it’s apparent \( L^\dagger(I) \) is Hermitian if \( L \) is HP. A simple \( L^\dagger(I) \) is more elegant and concise than the right side of (7), and that’s one of the reasons why we express our results in, for example, \( S^\mathcal{T}^\dagger(I) \), rather than the operators constituting the operator-sum representation of \( S^\mathcal{T} \).

Note \([10]\)

\[
(O_2 \otimes Q_2|O_1 \otimes Q_1) = (O_2|O_1)(Q_2|Q_1).
\]

(8)

We will utilize this relation frequently for proofs.

### 1.4 Orthonormal basis of operators and Choi isomorphism

Consider a Hilbert space \( \mathcal{H}_1 \) with an orthonormal basis \( \{a_i\} \). With Hilbert-Schmidt inner product, we can choose an orthonormal basis of operators for \( \mathcal{L}(\mathcal{H}_1) \) \([10, 14]\), by defining

\[
E_{ij} := |a_i\rangle \langle a_j|.
\]

(9)

Clearly \( E_{ij}^\dagger = E_{ji} \), so \( \{E_{ij}^\dagger\} = \{E_{ij}\} \).

Let \( L : \mathcal{L}(\mathcal{H}_1) \to \mathcal{L}(\mathcal{H}_2) \). Choi isomorphism \( \mathcal{T} \) then is defined as \([12]\)

\[
\mathcal{T}(L) := I \otimes (\sum_{i,j} E_{ij} \otimes E_{ij}) = \sum_{i,j} E_{ij} \otimes L(E_{ij}) \in \mathcal{L}(\mathcal{H}_1) \otimes \mathcal{L}(\mathcal{H}_2).
\]

(10)

It can be found \( L(O) = \text{tr}_1(O^\mathcal{T} \otimes I \mathcal{T}(L)) \) \([15]\):

\[
\text{tr}_1(O^\mathcal{T} \otimes I \mathcal{T}(L)) = \sum_{i,j} \text{tr}(O^\mathcal{T} E_{ij}) L(E_{ij})
\]

\[
= \sum_{i,j} \text{tr}(E_{ij} O L(E_{ij}) \circ \text{transposition is TP}
\]

\[
= \sum_{i,j} (E_{ij} |O) L(E_{ij})
\]

\[
= L \left( \sum_{i,j} E_{ij} |E_{ij}|O \right)
\]

\[
= L(O).
\]

The subscript beside \( \text{tr} \) refers to the party to be traced out; in this case, it’s \( \mathcal{H}_1 \) (or \( \mathcal{L}(\mathcal{H}) \)).

### 1.5 Bipartite system and partial transposition

When there are two parties Alice and Bob, let’s consider a linear mapping from \( \mathcal{L}(\mathcal{H}_{A_1} \otimes \mathcal{H}_{B_1}) \) to \( \mathcal{L}(\mathcal{H}_{A_2} \otimes \mathcal{H}_{B_2}) \). We don’t assume \( \mathcal{H}_{A_1} \) and \( \mathcal{H}_{A_2} \) are isomorphic, similarly for \( \mathcal{H}_{B_1} \) and \( \mathcal{H}_{B_2} \), so the dimensions before and after \( L \) can be different. Assign a basis for \( \mathcal{L}(\mathcal{H}_{A_1}) \) by (9), with \( \{|a_i\}\} \) being an orthonormal basis of \( \mathcal{H}_{A_1} \); then define a basis for \( \mathcal{L}(\mathcal{H}_{B_1}) \):

\[
F_{ij} := |b_i\rangle \langle b_j|,
\]

(11)

where \( \{|b_i\}\} \) is an orthonormal basis of \( B_1 \). Choi isomorphism for this composite system then is

\[
\mathcal{T}(L) = \sum_{i,j,k,l} E_{ij} \otimes F_{kl} \otimes L(E_{ij} \otimes F_{kl}).
\]

(12)
\( \mathcal{F} (L) \) is in \( \mathcal{L}(H_A \otimes H_B) \). From now on \( \{ E_{ij} \} \) and \( \{ F_{ij} \} \) will denote bases of operators.

Transposition \( T \) is a linear mapping: \( T : \mathcal{L}(H) \rightarrow \mathcal{L}(H) \). Similarly partial transposition, \( \Gamma = T_A \otimes I_B \), is also a linear mapping \( \Gamma : \mathcal{L}(H_A \otimes H_B) \rightarrow \mathcal{L}(H_A \otimes H_B) \). They are both the inverses to themselves, i.e. \( T \circ T = I \) and \( \Gamma \circ \Gamma = I_A \otimes I_B = I_{AB} \), so both are isomorphisms.

Partial transposition of a linear mapping \( L \) is defined as [16]

\[
L^\Gamma := \Gamma \circ L \circ \Gamma.
\]

Note for \( L \in \mathcal{L}(H_A \otimes H_B) \) and the second one maps \( \mathcal{L}(H_A \otimes H_B) \) to \( \mathcal{L}(H_A \otimes H_B) \). (13) itself is a linear mapping on linear mappings (on operators): If we define \( f(L) := L^\Gamma \), then

\[
f \in \mathcal{L} \{ \mathcal{L}(H_A \otimes H_B), \mathcal{L}(H_{A2} \otimes H_{B2}) \}.
\]

Furthermore, as \( (L^\Gamma)^\Gamma = L \), \( f \) is its own inverse, so it is also an isomorphism.

### 1.6 Trace, Schatten norm and positive operators

For \( p_1, p_2 \geq 0 \), by the eigendecomposition \( P_1 = \sum_i p_i |\psi_i \rangle \langle \psi_i| \) we can find [6];

\[
\min_i p_i \text{tr} P_2 \leq \text{tr} P_1 P_2 \leq \max_i p_i \text{tr} P_2 = ||P_1|| ||P_2||_1.
\]

where eigenvalue 0 is included here. The inequality on the right can be regarded as an application of Hölder’s inequality [17]. Either side of (14) becomes an equality iff the subspace orthogonal to ker \( P_2 \) is in the eigenspace of \( P_1 \) with the maximum or minimum eigenvalue, e.g. if \( P_1 = pI \), \( \text{tr} P_1 P_2 = p ||P_2||_1 \), ker refers to the kernel or null space.

More generally, for an operator \( O \) its Schatten \( p \)-norm [9, 18, 19] is

\[
||O||_p = \left( \text{tr} |O|^p \right)^{1/p},
\]

where \( |O| = \sqrt{O^\dagger O} \). For \( p = \infty \) it is (defined as) the largest singular value of \( O \), hence the same as the operator norm, \( || \cdots || \) in our notation; for \( p = 1 \) it is exactly the trace norm [18, 19]. Matrix Hölder inequality [20] states that for \( 1 \leq p, q \leq \infty \) satisfying \( 1/p + 1/q = 1 \):

\[
|\text{tr} O_1 O_2| = ||O_1, O_2|| \leq ||O_1||_p ||O_2||_q.
\]

As the upper bounds derived in our work are mainly based on (14), we can actually derive more general bounds using (16) and \( p \)- and \( q \)-norms, with \( 1/p + 1/q = 1 \) and \( 1 \leq p, q \leq \infty \). Further discussions will be given in Sec. 4.6.

From (14), we can easily see why for a sub-operation \( S_i \),

\[
0 \leq S_i^\Gamma (I) \leq I,
\]

otherwise the probability could surpass 1 or be negative, and also why a TP mapping \( L \) should have \( L^\dagger(I) = I \), which can also be regarded as a special case of the fact that \( (O_1 |O_2) = \text{tr} O_2 \) for all \( O_2 \) if and only if \( O_1 = I \). Additionally, the composition of TP mappings is still TP, so for a deterministic \( S \)

\[
S^{\Gamma \dagger} (I) = I.
\]

For any positive operator \( P \), its trace is equal to its trace norm, \( \text{tr} P = ||P||_1 \). In this supplemental material (as in the main paper) we will use them interchangeably. As an example, for any CP mapping \( L \), \( L^\dagger(I) \geq 0 \), so \( \text{tr} L^\dagger(I) = ||L^\dagger(I)||_1 \). More specifically, \( S_i^{\Gamma \dagger}(I) \geq 0 \) and \( \text{tr} S_i^{\Gamma \dagger}(I) = ||S_i^{\Gamma \dagger}(I)||_1 \).

### 2 Hermitian operators: Decomposition and orthogonality

#### 2.1 Decomposition of a Hermitian operator

Any Hermitian operator \( H \) can be decomposed as [21]

\[
H = \tilde{H}^+ - \tilde{H}^-, \ \tilde{H}^\pm \geq 0.
\]
Eigendecomposition $H = H^+ - H^-$ is one of such decompositions. Let ker denote the kernel or null space of an operator and ran the range or image of a mapping. Define

$$H^\pm := (\ker H^\pm)^\perp = \text{ran} H^\pm;$$

(20)
similarly for $\tilde{H}^\pm$. Hence,

$$H = H^+ \oplus H^- \oplus \ker H,$$

(21)
and those three subspaces are mutually orthogonal.

Here is a lemma concerning such a decomposition, which is a generalization of Lemma 2 from Ref. [21]:

**Lemma 1.** For any Hermitian operator $H$ on a finite-dimensional Hilbert space $\mathcal{H}$, among all possible such decompositions: $H = \tilde{H}^+ - \tilde{H}^-$, $\tilde{H}^\pm \geq 0$, the eigendecomposition $H = H^+ - H^-$ is the unique one that minimizes $\text{tr} H^+$, $\text{tr} H^-$, and $\text{tr}(H^+ + H^-)$; minimizing any one of them is the same as minimizing each of them. A decomposition where $\tilde{H}^\pm$ are orthogonal is equivalent to the eigendecomposition.

**Proof:**

**Eigendecomposition minimizes traces**

We find $\mathcal{H}^\pm \subset \tilde{H}^\pm$; otherwise there would exist $|\psi^\pm\rangle \in \mathcal{H}^\pm$ such that $\langle \psi^\pm | H | \psi^\pm \rangle = \langle \psi^\pm | \tilde{H}^+ - \tilde{H}^- | \psi^\pm \rangle$ does not come out positive/negative. Furthermore, since $\tilde{H}^\pm$ may intersect $\tilde{H}^\mp$ (and ker $H$),

$$| \langle \psi^\pm | H | \psi^\pm \rangle | = \langle \psi^\pm | H^\pm | \psi^\pm \rangle \leq \langle \psi^\pm | \tilde{H}^\pm | \psi^\pm \rangle,$$

(22)

because $\langle \psi^\pm | H | \psi^\pm \rangle = \langle \psi^\pm | \tilde{H}^+ | \psi^\pm \rangle - \langle \psi^\pm | \tilde{H}^- | \psi^\pm \rangle$ and $\tilde{H}^\pm \geq 0$. Hence

$$\text{tr} \tilde{H}^\pm \geq \text{tr} H^\pm,$$

(23)
i.e. eigendecomposition minimizes tr$\tilde{H}^\pm$ and tr$\tilde{H}^+ + \text{tr} \tilde{H}^-$.  

**Only eigendecomposition minimizes traces**

Suppose we want to minimize both tr$\tilde{H}^\pm$ simultaneously. Because $\tilde{H}^\pm \geq 0$ and because of (22), for both sides of (23) to be equal, i.e. to minimize both tr$\tilde{H}^\pm$, $\tilde{H}^\pm$ must be the same as $H^\pm$ (as $\mathcal{H}^\pm \subset \tilde{H}^\pm$), implying

$$\pm \langle \psi^\pm | H | \phi^\pm \rangle = \langle \psi^\pm | \tilde{H}^\pm | \phi^\pm \rangle = \langle \psi^\pm | H^\pm | \phi^\pm \rangle$$

(24)
and

$$\langle \psi^\pm | \tilde{H} | \phi^\pm \rangle = 0,$$

(25)
where $|\psi^\pm\rangle$ and $|\phi^\pm\rangle$ are any vectors in $\mathcal{H}^\pm$; the same is true for any vectors in ker $H$. Therefore $\langle \psi | \tilde{H}^\pm | \phi \rangle = \langle \psi | H^\pm | \phi \rangle$ for any $|\psi\rangle, |\phi\rangle \in \mathcal{H}$ as $\mathcal{H} = \mathcal{H}^+ \oplus \mathcal{H}^- \oplus \ker H$, so $\tilde{H}^\pm = H^\pm$. Because both tr$\tilde{H}^+$ and tr$\tilde{H}^-$ are minimized, so is tr$(\tilde{H}^+ + \tilde{H}^-)$.

It is clear that to minimize tr$(\tilde{H}^+ + \tilde{H}^-)$, both tr$\tilde{H}^+$ and tr$\tilde{H}^-$ should be minimized, leading to the same conclusion. If we just want to minimize one of tr$\tilde{H}^+$ and tr$\tilde{H}^-$, say tr$\tilde{H}^-$, then by (23) tr$\tilde{H}^- = \text{tr} H^-$. Because tr$H = \text{tr} H^+ - \text{tr} H^- = \text{tr} \tilde{H}^+ - \text{tr} \tilde{H}^-$, tr$\tilde{H}^+ = \text{tr} H^+$, i.e. both tr$\tilde{H}^+$ and tr$\tilde{H}^-$ are minimized. Hence minimizing any of them is identical to minimizing all of them, and minimization of traces requires eigendecomposition.

**Orthogonality is equivalent to eigendecomposition**

It has been known that if $\tilde{H}^\pm = H^\pm$, $\tilde{H}^\perp \perp \tilde{H}^\perp$, so we only have to show the reverse. If $\tilde{H}^\perp \perp \tilde{H}^\perp$, because $\mathcal{H}^\perp \subset \tilde{H}^\perp$ and $\mathcal{H}^\perp \perp \mathcal{H}^\perp$, $\mathcal{H}^\perp \subset \mathcal{H}^\perp \oplus \ker H$. For any vector $|\psi\rangle \in \ker H$, $H^\perp |\psi\rangle = H^- |\psi\rangle$; however, because $\tilde{H}^\perp \perp \tilde{H}^\perp$, $\tilde{H}^\perp |\psi\rangle = \tilde{H}^- |\psi\rangle = 0$, i.e. ker $H \subset \ker \tilde{H}^\perp$, which, by $\tilde{H}^\perp \subset \mathcal{H}^\perp \oplus \ker H$, implies $\tilde{H}^\perp \subset \mathcal{H}^\perp$, as ker $\tilde{H}^\perp \perp \tilde{H}^\perp$. Since $\mathcal{H}^\perp \subset \tilde{H}^\perp$, we conclude $\mathcal{H}^\perp = \tilde{H}^\perp$, so according to the arguments from the second part of the proof $\tilde{H}^\perp = H^\perp$.

Or alternatively, we can directly apply the spectral theorem [22]: As $\tilde{H}^\pm$ are orthogonal, (after eigendecomposing $\tilde{H}^\pm$) $\tilde{H}^+ - \tilde{H}^-$ is an orthogonal spectral resolution [22] of $H$.  

□
2.1.1 How to construct such a decomposition

In this section we will discuss given a Hermitian operator $H$, how to construct such a decomposition for it. The goal is to find a Hermitian operator $H'$ such that

$$\tilde{H}^\pm = H^\pm + H' \geq 0. \quad (26)$$

From (22) and $\mathcal{H}^\pm \subseteq \tilde{\mathcal{H}}^\pm$ it may seem $\tilde{H}^\pm \geq H^\pm$, but this is not true in general—what (22) implies is actually $\tilde{H}^\pm \geq H^\pm$ to the restriction of $\mathcal{H}^\pm$. To be specific, for a vector $|\psi^\pm\rangle$ where $|\psi^\pm\rangle \in \mathcal{H}^\pm$, the value of

$$(|\psi^+\rangle + |\psi^-\rangle) H' (|\psi^+\rangle + |\psi^-\rangle), \quad (27)$$

even though $\langle\psi^+|H'|\psi^\pm\rangle \geq 0$, is not necessarily non-negative, as it depends on $\langle\psi^\pm|H'|\psi^\mp\rangle$ as well. In terms of matrices, it just reflects the fact that a matrix that has non-negative diagonal entries are not necessarily positive (semi-definite).

Let’s choose a random orthonormal basis $\{|\phi_i\rangle\}$, and use it to construct $H'$: $\sum_i a_i |\phi_i\rangle \langle \phi_i |$. The eigenvalues $a_i$ are to be determined. Defining matrices $M$ and $A$ as

$$M_{ij} := \langle \phi_i | H' | \phi_j \rangle \quad \text{and} \quad A_{ij} := a_i \delta_{ij}, \quad (28)$$

(26) is equivalent to

$$M - A \geq 0. \quad (29)$$

Now we can try different combination of $\{a_i\}$; if (29) is satisfied, then we find a valid decomposition. Sylvester’s theorem may be useful in this process.

The benefit of choosing the eigenvectors of $H'$ as the standard basis is that we have less parameters to fix. Here are some guidelines on how to select the basis and eigenvalues:

1. $\text{tr} H' \geq 0$, as Lemma 1 suggests, i.e. $\sum_i a_i \geq 0$.
2. For $a_i < 0$, the corresponding $|\phi_i\rangle \notin \ker H$ and $|\phi_i\rangle \notin \mathcal{H}^\pm$. The reason is that, if $|\phi_i\rangle \in \mathcal{H}^+$, then $|\phi_i\rangle \bot \mathcal{H}^-$ and $\langle \phi_i | H^- + H' | \phi_i \rangle = a_i < 0$.
3. As a positive operator is also positive to the restriction of any subspace, it is advisable to determine the negative parts of $H'$ first, and check if $H^\pm + H'$ is positive when restricted to the subspace $\mathcal{H}^\mp$; however it should be emphasized even if $H^\pm + H' \geq 0$ to the restriction to this subspace, it does not mean $H^\pm + H' \geq 0$ over the entire space.
4. Similarly, for $M - A \geq 0$ it is necessary that $M - A$ have non-negative diagonal elements, but it is not a sufficient condition, as we have pointed out. The third and fourth points can be regarded as an application of Sylvester’s theorem.

2.2 Orthogonality of ensembles

**Lemma 2.** Suppose $P_1$ and $P_2$ are two positive operators, and they have such ensembles:[23, 24]

$$P_i = \sum_j |\psi^i_j\rangle \langle \psi^i_j |, \quad (30)$$

where each $\{|\psi^i_j\rangle\}$ is a set of nonzero vectors that aren’t necessarily normalized or mutually orthogonal. Then each eigenvector/eigenspace of $P_2$ corresponding to a nonzero eigenvalue is orthogonal to each of $P_1$ if and only if

$$\langle \psi^1_i | \psi^2_j \rangle = 0 \ \forall i, j, \quad (31)$$

in other words, if and only if

$$\text{ran} P_1 \bot \text{ran} P_2.$$ 

**Proof.** From Ref. [23], if the eigenensembles of $P_i$ are

$$P_i = \sum_{j \in i} |e^i_j\rangle \langle e^i_j |, \quad (32)$$

where each $\{|e^i_j\rangle\}$ is a set of nonzero vectors that aren’t necessarily normalized or mutually orthogonal. Then each eigenvector/eigenspace of $P_2$ corresponding to a nonzero eigenvalue is orthogonal to each of $P_1$ if and only if

$$\langle e^1_i | e^2_j \rangle = 0 \ \forall i, j,$$
where \(|e_j^i\rangle\) aren’t normalized, then there exist unitary matrices \(U^i\) such that

\[
\begin{cases}
|e_j^i\rangle = \sum_k U^i_{jk} |\psi_k^i\rangle & j \leq \text{Rank}(P_i) \\
|0\rangle = \sum_k U^i_{jk} |\psi_k^i\rangle & j > \text{Rank}(P_i)
\end{cases}
\] (33)

and

\[|\psi_j^i\rangle = \sum_k (U^i)^{-1}_{jk} |e_j^i\rangle.\] (34)

If \(\langle e_1^i | e_2^j \rangle = 0 \forall i, j\), from (34) we can see \(\langle \psi_1^i | \psi_2^j \rangle = 0 \forall i, j\). Similarly, if \(\langle \psi_1^i | \psi_2^j \rangle = 0\) then from (33) \(\langle e_1^i | e_2^j \rangle = 0 \forall i, j\). In other words, because each \(|\psi_j^i\rangle\) is a linear combination of \(|e_j^i\rangle\), and each \(|e_j^i\rangle\) is also linear combinations of \(|\psi_j^i\rangle\), orthogonality of two ensembles guarantees that of the other two.

Here’s a more elegant proof: The orthogonality of the vectors as stated in the lemma is satisfied if and only if their spans are orthogonal. From (32), the spans of \(|\psi_j^i\rangle\) are identical to \((\ker P_i)^\perp = \text{ran} P_i\), which of course are also identical to the spans of \(|e_j^i\rangle\). Therefore the orthogonality of one pair implies that of the other.

3 Adjoint of a linear mapping and complex conjugation of an operator

3.1 Adjoint

With \(L : \mathcal{L}(\mathcal{H}_1) \rightarrow \mathcal{L}(\mathcal{H}_2), \mathcal{T}(L) \in \mathcal{L}(\mathcal{H}_1) \otimes \mathcal{L}(\mathcal{H}_2)\). Hence \(\text{tr}_2\) in the following lemma traces out \(\mathcal{L}(\mathcal{H}_2)\).

Lemma 3. For a linear mapping \(L\),

\[L^\dagger(I) = \text{tr}_2 \mathcal{T}(L)^*.\]

Proof. As in Sec. 1.4, we can find

\[
\text{tr}_L(O) = \text{tr}(O^\dagger \otimes I \mathcal{T}(L)) = \sum_{i,j} \text{tr}(O^\dagger E_{ij}) \text{tr}(E_{ij}) = \sum_{i,j} (E_{ij} O) \text{tr}(E_{ij}) = \left( \sum_{i,j} E_{ij} \text{tr}(E_{ij})^* |O\rangle \right),
\] (35)

where \(L(E_{ij})^* = (L(E_{ij}))^*\). In addition, since \(E_{ij}^* = E_{ij}\),

\[\text{tr}_2 \mathcal{T}(L)^* = \sum_{i,j} E_{ij} \text{tr}(E_{ij})^*.\] (36)

Because

\[\text{tr}_L(O) = (I | L(O) \rangle) = (L^\dagger(I) | O\rangle),\]

by comparing (35) and (36) we can conclude \(L^\dagger(I) = \text{tr}_2 \mathcal{T}(L)^*\). 

3.2 Complex conjugation and basis

Note the importance of the basis for \(\mathcal{H}_1\) chosen in the proof above, because the transpose and complex conjugation of an operator depend on the basis. As for \(\mathcal{H}_2\), since \(\text{tr}(A^*) = (\text{tr} A)^*\) with whichever (orthonormal) basis, it does not matter.

Now assume the mapping \(L\) is HP, so we can perform a spectral decomposition on \(\mathcal{T}(L)\). Suppose its eigendecomposition is \(\mathcal{T}(L) = \sum_{i} c_i |v_i\rangle \langle v_i|\), with

\[|v_i\rangle = \sum_{j,k} d^i_{jk} |a_j\rangle |b_k\rangle,
\]
where \( \{b_l\} \) is any orthonormal basis of \( \mathcal{H}_2 \), and \( E_{ij} = |a_i\rangle\langle a_j| \) is the basis chosen for Choi isomorphism. Then,

\[
(|v_i^*\rangle\langle v_i|) = \sum_{j,k,l,m} \frac{d_{jk}^*}{d_{lm}^*} |a_l\rangle\langle a_l| \otimes |b_m\rangle\langle b_m|
\]

\[
= \sum_{j,k} \frac{d_{jk}^*}{d_{ik}} |a_j\rangle \langle b_k| \sum_{l,m} d_{lm}^* (a_l) \langle b_m|
\]

\[
= |v_i^*\rangle\langle v_i^*|,
\]

where

\[
|v_i^*\rangle := \sum_{j,k} \frac{d_{jk}^*}{d_{ik}} |a_j\rangle |b_k\rangle.
\]

Because \( \langle u|v^*\rangle = \langle u^*|v\rangle \),

\[
\langle v_i^*|v_j^*\rangle = \langle v_i|v_j\rangle^* = \delta_{ij}.
\]

We thus obtain

\[
\mathcal{F}(L)^* = \sum_i c_i |v_i^*\rangle\langle v_i^*|,
\]

and \( \{\langle v_i^*\rangle\} \) is an orthonormal basis. For a vector \( |v\rangle \in \mathcal{H}_1 \otimes \mathcal{H}_2 \), after fixing the basis for \( \mathcal{H}_1 \), \( |v^*\rangle \) depends on the basis for \( \mathcal{H}_2 \), but \( \text{tr}_2|v^*\rangle\langle v^*| \) would be the same regardless of the basis for \( \mathcal{H}_2 \).

From (38) and (39), \( \mathcal{F}(L)^* \) and \( \mathcal{F}(L) \) have the same eigenvalues, so

**Corollary 1.** For any HP mapping \( L \),

\[
\text{tr}L^\dagger(I) = \text{tr}\mathcal{F}(L).
\]

### 4 Proof of Proposition 1 and 2

Let’s restate Proposition 1 from the main article:

**Proposition 1. 1.** There exist upper and lower bounds for entangling capacities of deterministic operations:

\[
\frac{||S^\dagger(I)||_1}{d_AD_B} \leq EC_N(S) \leq ||\overline{S}^\dagger_-(I)||_1 ||\rho^\dagger||_1,
\]

\[
\log \frac{||\mathcal{F}(S^\dagger)||_1}{d_AD_B} = \log \left(1 + 2 \frac{||S^\dagger(I)||_1}{d_AD_B} \right) \leq EC_L(S) \leq \log(1 + 2 ||\overline{S}^\dagger_-(I)||_1).
\]

2. For a probabilistic operation composed of sub-operations \( S_i \),

\[
\sum_i \frac{||S_i^\dagger(I)||_1}{d_AD_B} \leq EC_N(S) \leq \sum_i ||\overline{S}_i^\dagger_-(I)||_1 ||\rho^\dagger||_1,
\]

\[
\sum_i \frac{\text{tr}\mathcal{F}(S_i)}{d_AD_B} \log \frac{||\mathcal{F}(S_i^\dagger)||_1}{\text{tr}\mathcal{F}(S_i)} \leq EC_L(S) \leq \log(1 + 2 \sum_i ||\overline{S}_i^\dagger_-(I)||_1).
\]

The upper bounds of part 2 can be applied to a deterministic operation \( S = \sum_i S_i \).

3. With an initial negativity \( E_N \), the expected negativity, i.e. probability times \( p_i \) the actual negativity \( E_{N_i} \), after a sub-operation \( S_i \) is bounded by:

\[
p_i E_{N_i} \leq E_N(||\overline{S}_i^\dagger_-(I)|| + ||\overline{S}_i^\dagger_+(I)||) + ||\overline{S}_i^\dagger_-(I)||.
\]

The entangling capacity of a sub-operation is bounded from below by:

\[
EC_N(S_i) \geq \frac{\text{tr}S_i^\dagger(I)}{\text{tr}S_i^\dagger(I)} = \left( \frac{\text{tr}S_i^\dagger(I)}{\text{tr}S_i^\dagger(I)} - 1 \right)^{-1},
\]

\[
EC_L(S_i) \geq \log \left(1 + 2 \frac{\text{tr}S_i^\dagger(I)}{\text{tr}S_i^\dagger(I)} \right) = \log \left[ 1 + 2 \left( \frac{\text{tr}S_i^\dagger(I)}{\text{tr}S_i^\dagger(I)} - 1 \right)^{-1} \right]
\]

All the upper bounds remain the same after the addition of an ancilla: \( ||(I \otimes \overline{S}_i^\dagger_-(I)|| = ||\overline{S}_i^\dagger_-(I)|| \).
4.1 Some equalities

Here are some relevant equalities. Suppose $L: \mathcal{L}(\mathcal{H}_1) \to \mathcal{L}(\mathcal{H}_2)$, and the dimension of $\mathcal{H}_1$ is $d$.

If $L$ is TP,

$$\text{tr} \mathcal{F}(L) = \text{tr} \sum_{i,j} E_{ij} \otimes L(E_{ij}) = \sum_{i,j} \delta_{ij} = d. \tag{42}$$

Therefore for a TP operation $S$, $\text{tr} \mathcal{F}(S) = \text{tr} \mathcal{F}(S^T) = d_A d_B$. Note $S^T$ is also TP.

If $L$ is HPTP, then $L^!(I) = \tilde{L}^!(I) - \tilde{L}^!(I) = I$, so $\tilde{L}^!(I) = I + \tilde{L}^!(I)$. Hence

$$1 + 2||\tilde{L}^!(I)|| = ||I + 2\tilde{L}^!(I)|| = ||\tilde{L}^!(I) + \tilde{L}^!(I)||. \tag{43}$$

because $||I + P|| = 1 + ||P||$ for $P \geq 0$, and

$$d + 2||\tilde{L}^!(I)||_1 = ||I + 2\tilde{L}^!(I)||_1 = ||\tilde{L}^!(I) + \tilde{L}^!(I)||_1. \tag{44}$$

4.2 Upper bounds

We start with

**Lemma 4.** For an HPTP mapping $L$ and Hermitian $H$,

$$\text{tr} L(H)^- - \text{tr} H^- \leq ||\tilde{L}^!(I)|| ||H||_1, \tag{45}$$

$$||L(H)||_1 - ||H||_1 \leq 2||\tilde{L}^!(I)|| ||H||_1. \tag{46}$$

**Proof.** For any Hermitian $H$ [21]

$$||H||_1 = \text{tr} H^+ + \text{tr} H^- = \text{tr} H + 2\text{tr} H^-.$$

With an HP $L$, since $\tilde{L}_\pm$ are CP, by Lemma 1 [6]

$$\text{tr} L(H)^\pm \leq \text{tr} \tilde{L}^+(H^\pm) + \text{tr} \tilde{L}^-(H^\mp). \tag{47}$$

Since $L$ is TP as well,

$$\text{tr} L(H)^- - \text{tr} H^- \leq \text{tr}[\tilde{L}^+(H^-) + \tilde{L}^-(H^+) - L(H^-)]$$

$$= \text{tr}[\tilde{L}^!(I)(H^+ + H^-)]$$

$$\leq ||\tilde{L}^!(I)|| ||H||_1 \text{ by (14).}$$

By (45) and $L$ being TP,

$$||L(H)||_1 - ||H||_1 = 2(\text{tr} L(H)^- - \text{tr} H^-)$$

$$\leq 2||\tilde{L}^!(I)|| ||H||_1.$$

The proof is completed. \hfill \Box

4.2.1 Deterministic operation

An operation $S$ is CP, so it is HP [2, 11, 12]. Because $\Gamma$ is HPTP, $S^\Gamma$ is HPTP. Hence we can apply Lemma 4 to $S^\Gamma(\rho^\Gamma) = S(\rho)^\Gamma$:

$$E_N(S(\rho)) - E_N(\rho) = \text{tr} S^\Gamma(\rho^\Gamma)^- - \text{tr} \rho^\Gamma^-$$

$$\leq ||S^\Gamma^!(I)|| ||\rho^\Gamma||_1, \tag{47}$$

and

$$E_L(S(\rho)) - E_L(\rho) = \log ||S^\Gamma(\rho^\Gamma)||_1 - \log ||\rho^\Gamma||_1$$

$$= \log(1 + \frac{||S^\Gamma(\rho^\Gamma)||_1 - ||\rho^\Gamma||_1}{||\rho^\Gamma||_1})$$

$$\leq \log(1 + 2||\tilde{S}^\Gamma_!(I)||). \tag{48}$$
The upper bounds from part 1 of Proposition 1 are proven.

For $L = \sum_i L_i$, where $L_i$ are HP and $L$ is TP, we can choose $\tilde{L}_\pm = \sum_i \tilde{L}_i \pm$. With

$$(\sum_i \tilde{L}_i \pm) = \sum_i \tilde{L}_i \pm,$$

by Lemma 4 we obtain

$$\text{tr} L (H) - \text{tr} H^- \leq ||\sum_i \tilde{L}_i \pm (I)|| ||H||_1$$

by the triangle inequality. Therefore for a deterministic operation $S = \sum_i S_i$, if we choose $\tilde{S}_\pm = \sum_i (\tilde{S}_i \pm)$,

$$E_N(S(\rho)) - E_N(\rho) \leq ||\sum_i \tilde{S}_i \pm (I)|| ||\rho^F||_1$$

and

$$E_L(S(\rho)) - E_L(\rho) \leq \log (1 + 2 \sum_i \tilde{S}_i \pm ||I||)$$

Hence we have proved part 2 of Proposition 1 can be applied to deterministic $S = \sum_i S_i$.

Let’s emphasize again $\tilde{S}_\pm = \sum_i (\tilde{S}_i \pm)$ are some of the possible decompositions of $S^F$, and they may or may not be the ideal choices that yield the smallest bounds.

### 4.2.2 Probabilistic operation

Let’s prove the upper bounds from part 2 of Proposition 1 for probabilistic operations. The average negativity after $S$ is

$$\sum_i p_i \log ||\rho_i^F||_1 = \sum_i \text{tr} S_i(\rho) F^-$$

Furthermore, by (46)

$$\sum_i \text{tr} S_i(\rho) F^\pm \leq \sum_i [\text{tr} \tilde{S}_i^F(\rho^F_\pm) + \text{tr} \tilde{S}_i^F(\rho^\mp_\pm)]$$

$$= \text{tr} \left( \sum_i \tilde{S}_i^F \pm \right) (\rho^F_\pm) + \text{tr} \left( \sum_i \tilde{S}_i^F \mp \right) (\rho^\mp_\pm).$$

Essentially, we now have $S^F = \sum_i S_i^F + \sum_i \tilde{S}_i^F \pm - \sum_i \tilde{S}_i^F \mp$, with $\tilde{S}_\pm = \sum_i \tilde{S}_i \pm$, and $S$ is TP as in the deterministic case, c.f. (46). Thus Lemma 4 can be used, and we essentially recover the right sides of (49) for the bounds of $\sum_i p_i E_{N_i} - E_N$, where $E_N$ is the initial negativity.

As to the bounds for $E_{C_L}$ of a probabilistic operation, with an initial logarithmic negativity $E_L$:

$$\sum_i p_i \log ||\rho_i^F||_1 = \sum_i \text{tr} S_i(\rho) F^-$$

$$= \sum_i p_i \log ||\tilde{S}_i^F(\rho^F)||_1 - \log ||\rho^F||_1$$

$$\leq \log \sum_i ||\tilde{S}_i^F(\rho^F)||_1 - \log ||\rho^F||_1$$

$$= \log \left( 1 + \sum_i ||\tilde{S}_i^F(\rho^F)||_1 - ||\rho^F||_1 \right),$$

by the concavity of logarithm and Jensen’s inequality [17, 25, 26]. By (45),

$$\sum_i ||\tilde{S}_i^F(\rho^F)||_1 = \sum_i [\text{tr} \tilde{S}_i^F(\rho^F) + 2 \text{tr} \tilde{S}_i^F(\rho^F)^-]$$

$$= \text{tr} \sum_i \tilde{S}_i^F(\rho^F) + 2 \sum_i \text{tr} \tilde{S}_i^F(\rho^F)^- - \text{tr} \rho^F + 2 \sum_i \text{tr} \tilde{S}_i^F(\rho^F)^-,$$

because $S^F = \sum_i S_i^F$ is TP. Now we are back to (52). Following the same steps as before, we can obtain the desired result.
4.2.3 Sub-operation

Let’s move on to upper bounds from part 3 of Proposition 1. By (14), (45) and (46),

$$p_i E_{N_i} = \text{tr} S_F^T(\rho^F) \leq ||\widetilde{S}_{I+}^F(\Gamma)|| ||\rho^F||_1 + ||\widetilde{S}_{I-}^F|| ||\rho^F||_1$$

$$= ||\widetilde{S}_{I+}^F(\Gamma)|| ||\rho^F||_1 + ||\widetilde{S}_{I-}^F(\Gamma)|| (1 + ||\rho^F||_1)$$

$$= E_N \left( ||\widetilde{S}_{I+}^F(\Gamma)|| + ||\widetilde{S}_{I-}^F(\Gamma)|| \right) + ||\widetilde{S}_{I-}^F(\Gamma)||.$$

4.3 Lower bounds

4.3.1 Deterministic operation

Employing the method proposed by Campbell [6], to obtain the lower bounds for entangling capacities, an ancilla is used: We apply the operation $I_{A_n} \otimes S_{A_nB_n}$ on the state $|\Psi\rangle$ defined as:

$$|\Psi\rangle\langle\Psi| := \frac{1}{d_A d_B} \sum_{i,j,k,l} |a_i\rangle\langle a_j| \otimes |b_k\rangle\langle b_l|$$

$$= |\Psi_{A_nA_n}\rangle \langle \Psi_{A_nA_n}| \otimes |\Psi_{B_nB_n}\rangle \langle \Psi_{B_nB_n}|,$$

$$= \frac{1}{d_A d_B} \sum_{i,j,k,l} E_{ij} \otimes F_{kl} \otimes E_{ij} \otimes F_{kl}$$ (55)

where

$$|\Psi_{A_nA_n}\rangle := \frac{1}{\sqrt{d_A}} |a_i\rangle|a_i\rangle$$ (56)

is a maximally entangled state, similarly for $|\Psi_{B_nB_n}\rangle$. Here $A_n$ and $B_n$ are the ancilla, and are clones of the original system $A_s$ and $B_s$ (By clones we mean $\mathcal{H}_{A_n}$ and $\mathcal{H}_{B_n}$ are isomorphic to $\mathcal{H}_{A_s}$ and $\mathcal{H}_{B_s}$. An ancilla in general does not necessarily have to be isomorphic to the system). (55) is essentially a normalized Choi isomorphism.

We will ignore the subscripts of $I_{A_nB_n}$ and $S_{A_nB_n}$ from now on. As A and B are not entangled,

$$\mathcal{I} \otimes S|\Psi\rangle\langle\Psi| = \mathcal{T}(S)/(d_A d_B)$$ (57)

alone can give us a lower bound for entangling capacity. With $\Gamma = T_A \otimes I_B$, we have

$$\mathcal{T}(S)^\Gamma = d_A d_B \langle \mathcal{I} \otimes S|\Psi\rangle\langle\Psi| \rangle^\Gamma$$

$$= \mathcal{I} \otimes S^F \left( \sum_{i,j,k,l} E_{ij} \otimes F_{kl} \otimes E_{ij} \otimes F_{kl} \right)^\Gamma$$

$$= \mathcal{I} \otimes S^F \left( \sum_{i,j,k,l} E_{ji} \otimes F_{kl} \otimes E_{ji} \otimes F_{kl} \right)$$

$$= \mathcal{T}(S^F),$$ (58)

because $(\mathcal{I} \otimes S)^\Gamma = \mathcal{T}^\Gamma \otimes S^F = \mathcal{I} \otimes S^\Gamma$. Therefore

$$[\mathcal{T}(S)^\Gamma]^{\pm} = \mathcal{T}(S^\Gamma)^{\pm} = \mathcal{T}(S^\Gamma_\pm).$$ (59)

By Corollary 1, $\text{tr} \mathcal{T}(S^\Gamma_\pm) = \text{tr} S^\Gamma_\pm(\Gamma)$, so

$$\text{tr} \left\{ [\mathcal{I} \otimes S(|\Psi\rangle\langle\Psi|)]^\Gamma \right\}^{\pm} = \frac{1}{d_A d_B} \text{tr} \mathcal{T}(S^\Gamma_\pm) = \frac{1}{d_A d_B} \text{tr} S^\Gamma_\pm(\Gamma).$$ (60)
Alternatively, we can directly calculate $\text{tr}\mathcal{T}(S^\Gamma_\pm)$ without using Corollary 1:

$$
\text{tr}\mathcal{T}(S^\Gamma_\pm) = \text{tr} \sum_{i,j,k,l} E_{ji} \otimes F_{kl} \otimes S^\Gamma_\pm(E_{ij} \otimes F_{kl}) = \sum_{i,j,k,l} \delta_{ij} \delta_{kl} \text{tr} S^\Gamma_\pm(E_{ij} \otimes F_{kl}) = \text{tr} S^\Gamma_\pm \left( \sum_{i,k} E_{ii} \otimes F_{kk} \right) = \text{tr} S^\Gamma_\pm(I),
$$

(61)

because $\text{tr} L(I) = (L^\dagger(I)|I) = \text{tr}(L(I)^\dagger)$, and $L^\dagger(I)$ is Hermitian is $L$ is HP, (7). Hence

$$
E_N(I \otimes S(|\Psi\rangle\langle\Psi|)) = \frac{1}{d_A d_B} \text{tr}\mathcal{T}(S^\Gamma) = \frac{1}{d_A d_B} \text{tr} S^\Gamma_\pm(I).
$$

We can use (45) to acquire the bound for logarithmic negativity as in Proposition 1. Or, because $\mathcal{T}(S^\Gamma_\pm)$ are orthogonal, by (60)

$$
||I \otimes S(|\Psi\rangle\langle\Psi|)||_1 = \frac{1}{d_A d_B} ||\mathcal{T}(S^\Gamma)||_1,
$$

(62)

leading to one of the expressions of the lower bound. As $||\mathcal{T}(S^\Gamma)||_1 = \text{tr}\mathcal{T}(S^\Gamma_+^\dagger) + \text{tr}\mathcal{T}(S^\Gamma_-^\dagger)$, by Corollary 1,

$$
||\mathcal{T}(S^\Gamma)||_1 = \text{tr} S^\Gamma_+^\dagger(I) + \text{tr} S^\Gamma_-^\dagger(I).
$$

(63)

Since $S^\Gamma_+^\dagger(I) - S^\Gamma_-^\dagger(I) = I$,

$$
\text{tr} S^\Gamma_+^\dagger(I) + \text{tr} S^\Gamma_-^\dagger(I) = \text{tr} I + 2\text{tr} S^\Gamma_-^\dagger(I) = d_A d_B + 2\text{tr} S^\Gamma_-^\dagger(I),
$$

(64)

which gives us another expression of the lower bound. We can also use (44) to show this relation.

### 4.3.2 Sub-operation

To obtain the lower bounds for a sub-operation the procedures are pretty much the same, but now we should take the probability into account:

$$
p_i = \text{tr} I \otimes S_i(|\Psi\rangle\langle\Psi|) = \frac{\text{tr}\mathcal{T}(S_i)}{d_A d_B} = \frac{\text{tr} S^\Gamma_\pm(I)}{d_A d_B},
$$

(65)

c.f. (60) or (61). Use the same method to obtain $\text{tr} I \otimes S^\Gamma_-^\dagger(|\Psi\rangle\langle\Psi|)$. Thus the negativity is

$$
E_{N_i}(|\Psi\rangle\langle\Psi|) = \frac{\text{tr} S^\Gamma_-^\dagger(I)}{\text{tr} S^\Gamma_\pm(I)}.
$$

(66)

Similarly,

$$
||I \otimes S_i(|\Psi\rangle\langle\Psi|)||_1 = \frac{||\mathcal{T}(S^\Gamma_i)||_1}{d_A d_B} = \frac{\text{tr} S^\Gamma_+^\dagger(I) + \text{tr} S^\Gamma_-^\dagger(I)}{d_A d_B},
$$

(67)

c.f. (63), so

$$
E_{L_i}(|\Psi\rangle\langle\Psi|) = \log \frac{||\mathcal{T}(S^\Gamma_i)||_1}{\text{tr} S^\Gamma_\pm(I)} = \log \frac{\text{tr} S^\Gamma_+^\dagger(I) + \text{tr} S^\Gamma_-^\dagger(I)}{\text{tr} S^\Gamma_\pm(I)}
$$

(68)

The lower bounds for deterministic operations can be regarded as a special case of sub-operations—this can be easily shown with Corollary 1 and (12).

Because transposition and partial transposition are TP, $\text{tr}\mathcal{T}(L) = \text{tr}\mathcal{T}(L^\Gamma)$. For an HP $L$, it means

$$
\text{tr}\mathcal{T}(L) = \text{tr}\mathcal{T}(L^\Gamma) = \text{tr}\mathcal{T}(L^\Gamma),
$$

$$
\text{tr}\mathcal{T}(L) = \text{tr}\mathcal{T}(L^\Gamma) = \text{tr}\mathcal{T}(L^\Gamma) - \text{tr}\mathcal{T}(L^\Gamma).
$$

Therefore by Corollary 1

$$
\text{tr} S^\Gamma_\pm(I) = \text{tr} S^\Gamma_\pm(I) = \text{tr} S^\Gamma_\pm(I) - \text{tr} S^\Gamma_\pm(I).
$$

(69)

With (69) we can adjust (66) and (68) to our liking, such as in Proposition 1. Note in general we can’t expect $S^\Gamma_\pm(I) \geq 0$, so its trace may not equate to its trace norm.
4.3.3 Probabilistic operation

For the average negativity of a probabilistic operation, by (65) and (66)

$$\sum_i p_i E_N(I \otimes S_i(\Psi)) = \sum_i \frac{\text{tr} S_i^{\dagger\dagger}(I)} {d_A d_B}. \quad (70)$$

For the average logarithmic negativity, by (65) and (68)

$$\sum_i p_i E_L(I \otimes S_i(\Psi)) = \sum_i \frac{\text{tr} \mathcal{F}(S_i)} {d_A d_B} \log \frac{||\mathcal{F}(S_i^{\dagger\dagger})||_1} {\text{tr} \mathcal{F}(S_i)}, \quad (71)$$

Note for the upper bounds the denominator $p_i$ could be removed due to concavity (see (53)), which can’t be applied here. Also note by Corollary 1 and (69) there are several expressions for these bounds, just like sub-operations.

4.4 Proposition 2

As defined in the main paper, for two operators or linear mappings $X$ and $Y$

$$D_{p, \Gamma}(X, Y) := ||X – Y||_{p, \Gamma} := ||X^\Gamma – Y^\Gamma||_p, \quad (72)$$

where

$$||L||_p := ||\mathcal{F}(L)||_p \text{ for any linear mappings } L. \quad (73)$$

As discussed in Sec. 1.5, partial transpositions on both operators and linear mappings are isomorphisms.

Now let’s prove:

**Proposition 2.** For any density operators $\rho$ and $\rho_i$, and deterministic operations $S$ and $S_i$, we have an equality

$$(S^\Gamma_2 – S^\Gamma_1)^\dagger(I) = (S^\Gamma_2 – S^\Gamma_1)^\dagger(I), \quad (74)$$

and the following inequalities:

$$D_{1, \Gamma}(S_1(\rho), S_2(\rho)) \leq 2|| (S^\Gamma_2 – S^\Gamma_1)^\dagger(I) \| \|_1 \| \|_1 \| \rho^\Gamma \|_1 \quad (75)$$

$$\leq D_{1, \Gamma}(S_1, S_2) \| \rho^\Gamma \|_1, \quad (76)$$

$$D_{1, \Gamma}(S(\rho_1), S(\rho_2)) \leq (1 + 2|| S^\Gamma_1^\dagger(I) \| ) D_{1, \Gamma}(\rho_1, \rho_2). \quad (77)$$

**Proof.** Because $S_i$ are deterministic, $S^\Gamma_1(I) = I$, and

$$S^\Gamma_1^\dagger(I) = 0.$$

Hence we can phrase our problem this way: If an HP mapping $L$ has $L(I) = 0$, given an operator $O$ with a fixed trace norm, how large can $||L(O)||_1$ be? Let’s begin with

$$L^\dagger(I) = L^\dagger_{+}(I) – L^\dagger_{-}(I) = 0 \Rightarrow L^\dagger_{+}(I) = L^\dagger_{-}(I), \quad (78)$$

confirming $(S^\Gamma_2 – S^\Gamma_1)^\dagger(I) = (S^\Gamma_2 – S^\Gamma_1)^\dagger(I)$. Then by (46),

$$||L(O)||_1 = \text{tr} L(O)^\dagger + \text{tr} L(O)^\dagger$$

$$\leq \text{tr} (L^\dagger_+ + L^\dagger_-)(O^\dagger + O^-)$$

$$= 2\text{tr} \left[ L^\dagger_+(I)(O^\dagger + O^-) \right]$$

$$\leq 2||L^\dagger_+(I)|| ||O||_1 \text{ by (14) and } ||O||_1 = \text{tr}(O^\dagger + O^-).$$

(75) is proven. To show (76), we need the inequality $||O||_1 \leq ||O||_1 \ [19, 27]$, so $||(S^\Gamma_2 – S^\Gamma_1)^\dagger(I)|| \leq ||(S^\Gamma_2 – S^\Gamma_1)^\dagger(I)||$. By Corollary 1, $\text{tr}(S^\Gamma_2 – S^\Gamma_1)^\dagger(I) = \text{tr} \mathcal{F}(S^\Gamma_2 – S^\Gamma_1)^\dagger$, so by (78)

$$2\text{tr} \mathcal{F}(S^\Gamma_2 – S^\Gamma_1)^\dagger = \text{tr} \mathcal{F}(S^\Gamma_2 – S^\Gamma_1)^\dagger + \text{tr} \mathcal{F}(S^\Gamma_2 – S^\Gamma_1)^\dagger = ||\mathcal{F}(S^\Gamma_2 – S^\Gamma_1)||_1 = ||S_2 – S_1||_{1, \Gamma}.$$
To show (77) is pretty similar, as
\[ S(\rho_1)^F - S(\rho_2)^F = S^F[(\rho_1 - \rho_2)^F]. \]
Setting \( O = (\rho_1 - \rho_2)^F \), by (46) we can easily find
\[
\|S^F(O)\|_1 \leq \text{tr}(S_+^F(S_+^F + S_-^F)(O^+ + O^-))
\leq \|S_+^F(I) + S_-^F(I)\| \|O\|_1
= \left(1 + 2\|S_+^F(I)\|\right)\|\rho_1 - \rho_2\|_1, \text{ by (43)}.
\]

4.5 Upper bounds with an ancilla

Since \( \tilde{L} \pm \) are CP, so are \( I \otimes \tilde{L} \pm \), and \( I \otimes (\tilde{L} + - \tilde{L} -) = I \otimes \tilde{L} + - I \otimes \tilde{L} - \), which is therefore a decomposition of \( I \otimes L \) as a difference of CP mappings, or in our notation, \( I \otimes \tilde{L} \).

With this in mind, let’s show the upper bounds are valid with an ancilla; in particular
\[
|||I \otimes \tilde{S}_i - \tilde{S}_i \|\|_1 \leq |||I \otimes \tilde{S}_i \|\|_1 \leq ||I \otimes \tilde{S}_i \|\|_1 \leq \|\rho_1 - \rho_2\|_1, \text{ by (43)}.
\]

4.6 Upper bounds using Schatten \( p \)-norm

All the upper bounds come from
\[
\text{tr}P(H^+ + H^-),
\]where \( P \) originates from the adjoint of a CP mapping. By matrix Hölder inequality (16) [20],
\[
\text{tr}P(H^+ + H^-) \leq ||P||_p ||H^+ + H^-||_q,
\]
where \( 1 \leq p, q \leq \infty \) and \( 1/p + 1/q = 1 \). We do not need to take the absolute value as \( \text{tr}P_1P_2 \geq 0 \).

For a Hermitian operator \( H \),
\[
\|H\| = H^+ + H^-,
\]
because $H^+$ and $H^-$ have orthogonal eigenvectors, c.f. Lemma 1. Therefore,
\[ \|H^+ + H^-\|_p = \|H\|_p, \]
and (80) becomes
\[ \text{tr} P(H^+ + H^-) \leq \|P\|_p \|H\|_q. \] (81)

Using (81), we can generalize the upper bounds we’ve obtained. For example, the upper bound for negativity from part 1 of Proposition 1 becomes
\[ EC_N(S) \leq \|\vec{S}_+^T(I)\|_p \|\rho^T\|_q, \]
and the upper bound for logarithmic negativity will be state-dependent when $q \neq 1$. Or for (77), we may want to use another norm for states, then the norm for $S^T_+(I)$ can be adjusted accordingly. The rest can be derived easily.

5 Equivalence of the bounds

By part 2 of Proposition 1, $EC_L(S) \leq \log(1 + 2\|\sum_i S_i^T \|_1)$; be aware that here it is $S_i^T$ instead of $\vec{S}_i^T$. In Ref. [6], it was shown for a deterministic operation $S = \sum_i S_i$, where $S_i(\mathbf{\rho}) = V_i \mathbf{\rho} V_i^\dagger$ and $V_i$ has a Schmidt decomposition [28]
\[ V_i = \sum_{j(i)} \lambda_{ij} A_{ij} \otimes B_{ij}, \] (82)
the entangling capacity has an upper bound:
\[ EC_L(S) \leq \log \left( \sum_i \|\sum_{j(i)} \lambda_{ij} A_{ij}^\dagger A_{ij}\| \|\sum_{k(i)} \lambda_{ik} B_{ik}^\dagger B_{ik}\| \right). \]

It was shown by Campbell for a unitary operator $U = \sum_i \lambda_i A_i \otimes B_i$ the lower bound is [6]
\[ \log \frac{(\sum_i \lambda_i)^2}{d_A d_B}, \] (83)
while our lower bound is
\[ \log \left( 1 + 2 \frac{\|S^T_+(I)\|_1}{d_A d_B} \right). \] (84)

Here we would like to show

Lemma 5. For deterministic $S = \sum_i S_i$, where $S_i = V_i \mathbf{\rho} V_i^\dagger$ and the Schmidt decompositions of $V_i$ are $V_i = \sum_{j(i)} \lambda_{ij} A_{ij} \otimes B_{ij}$,
\[ 1 + 2\|\sum_i S_i^T \|_1 = \|\sum_i \sum_{j(i)} \lambda_{ij} A_{ij}^\dagger A_{ij} \otimes \sum_{k(i)} \lambda_{ik} B_{ik}^\dagger B_{ik}\| \]
\[ \leq \sum_i \|\sum_{j(i)} \lambda_{ij} A_{ij}^\dagger A_{ij}\| \|\sum_{k(i)} \lambda_{ik} B_{ik}^\dagger B_{ik}\|. \]

For a unitary operation $S(\rho) = U \rho U^\dagger$, where the Schmidt decomposition of $U$ is $U = \sum_i \lambda_i A_i \otimes B_i$,
\[ 1 + 2 \frac{\|S(i)\|_1}{d_A d_B} = (\sum_i \lambda_i)^2 d_A d_B. \]

By which, $EC_L(S) \leq \log(1 + 2\|\sum_i S_i^T \|_1)$ \leq ( $\sum_i \|\sum_{j(i)} \lambda_{ij} A_{ij}^\dagger A_{ij}\| \|\sum_{k(i)} \lambda_{ik} B_{ik}^\dagger B_{ik}\|$), and the lower bound obtained in Ref. [6], (83), is the same as that from Proposition 1, (84).

5.1 Upper bound

Utilizing the tricks from Appendix E of Ref. [6], we are able to give the following proof.
5.1.1 Single sub-operation

Let’s first consider just one sub-operation $S_i$, and drop the subscript $i$ for this moment. Suppose $S(O) = VOV^\dagger$, with $V = \sum_i \lambda_i A_i \otimes B_i$ being the Schmidt decomposition. Then

$$S^T(O) = (VOV^\dagger)^\Gamma = \sum_{i,j} \lambda_i \lambda_j A_i^\dagger \otimes B_i O(A_i^\dagger \otimes B_j)^\dagger. \tag{85}$$

Define

$$V_{ij}^{\pm} := (A_i^\dagger \otimes B_j \pm A_j^\dagger \otimes B_i \text{ for } )/\sqrt{2} \text{ for } i \neq j,$$

$$V_{ij}^+ := A_i^\dagger \otimes B_i. \tag{86}$$

We can see for $i \neq j$

$$A_i^\dagger \otimes B_i O(A_i^\dagger \otimes B_i)^\dagger + A_j^\dagger \otimes B_j O(A_j^\dagger \otimes B_j)^\dagger = V_{ij}^+ O V_{ij}^{-\dagger} - V_{ij}^- O V_{ij}^{-\dagger}. \tag{87}$$

Note $V_{ij}^{\pm} = \pm V_{ji}^{\mp}$ and therefore $V_{ij}^+ O V_{ij}^- = V_{ji}^- O V_{ji}^+$. We now have

$$S^T(O) = \sum_i \lambda_i^2 V_{ii}^+ O V_{ii}^{-\dagger} + \frac{1}{2} \sum_{i \neq j} \lambda_i \lambda_j V_{ij}^+ O V_{ij}^{-\dagger} - \frac{1}{2} \sum_{i \neq j} \lambda_i \lambda_j V_{ij}^- O V_{ij}^{-\dagger}. \tag{88}$$

As each summation on the right side of (88) is in the Kraus form, each is a CP mapping, so we have

$$\tilde{S}_{T+}(O) = \sum_i \lambda_i^2 V_{ii}^+ O V_{ii}^{-\dagger} + \frac{1}{2} \sum_{i \neq j} \lambda_i \lambda_j V_{ij}^+ O V_{ij}^{-\dagger} \tag{89}$$

\tilde{S}_{T-}(O) = \frac{1}{2} \sum_{i \neq j} \lambda_i \lambda_j V_{ij}^- O V_{ij}^{-\dagger}.

Now the problems is, is $\tilde{S}_{T\pm} = S_T^\pm$? This is equivalent to $\mathcal{T}(\tilde{S}_{T\pm})$ having orthogonal eigenvectors, by Lemma 1 (and Choi isomorphism, well, being bijective). We can find $\{V_{ij}^+\}$ and $\{V_{kl}^{\mp}\}$ are mutually orthogonal:

$$(V_{ij}^+ | V_{kl}^-) = (A_i^\dagger \otimes B_j + A_j^\dagger \otimes B_i | A_i^\dagger \otimes B_j - A_j^\dagger \otimes B_i) /2 = 0; \tag{90}$$

the orthogonality of the rest can be shown by the orthonormality of $\{A_i\}$ and $\{B_i\}$. Choose an orthonormal basis $\{|a_i\rangle\}$ for the composite system $A \otimes B$ whose dimension is $d = d_A d_B$. Like (56), define

$$|\Psi\rangle := \frac{1}{\sqrt{d}} \sum_i |a_i\rangle |a_i\rangle. \tag{91}$$

Then the Choi isomorphism of an operation by a single $V_{kl}^{\mp}$ is

$$\sum_{ij} |a_i\rangle \langle a_j| V_{kl}^\pm |a_i\rangle \langle a_j| V_{kl}^{\mp \dagger} = d(I \otimes V_{kl}^\pm) |\Psi\rangle \langle \Psi| (I \otimes V_{kl}^{\mp \dagger}); \tag{92}$$

Therefore $\mathcal{T}(\tilde{S}_{T\pm})$ are ensembles of pure states. It was proved in Appendix F of Ref. [6] that if operators $O_1$ and $O_2$ are orthogonal, then $I \otimes O_1 |\Psi\rangle$ and $I \otimes O_2 |\Psi\rangle$ are orthogonal:

$$\langle \Psi| (I \otimes O_2)^\dagger I \otimes O_1 |\Psi\rangle = \frac{1}{d} \sum_{ij} |a_i\rangle \langle a_j| O_2^* O_1 |a_j\rangle = \frac{1}{d} (O_2 |O_1\rangle. \tag{93}$$

By Lemma 2, $\mathcal{T}(\tilde{S}_{T\pm})$ have orthogonal eigenvectors, so $\tilde{S}_{T\pm}$ as defined in (89) are equal to $S_T^\pm$. Therefore,

$$S_{T+}^T(I) = \frac{1}{2} \sum_{i \neq j} \lambda_i \lambda_j V_{ij}^{-\dagger} V_{ij}^-, \tag{94}$$

$$S_{T-}^T(I) = \sum_i \lambda_i^2 V_{ii}^{+\dagger} V_{ii}^- + \frac{1}{2} \sum_{i \neq j} \lambda_i \lambda_j V_{ij}^{+\dagger} V_{ij}^- + \frac{1}{2} \sum_{i \neq j} \lambda_i \lambda_j V_{ij}^{-\dagger} V_{ij}^-. \tag{95}$$
and
\[
S^{\dagger +}_t(I) + S^{\dagger -}_t(I) = \sum_i \lambda_i^2 A_i^\dagger A_i^* \otimes B_i^\dagger B_i + \frac{1}{2} \sum_{i \neq j} \lambda_i \lambda_j \left( A_i^\dagger A_j^* \otimes B_i^\dagger B_j + A_j^\dagger A_i^* \otimes B_i^\dagger B_j \right)
\]
\[
= \sum_i \lambda_i \lambda_j A_i^\dagger A_j^* \otimes B_i^\dagger B_j + \sum_{i \neq j} \lambda_i \lambda_j A_i^\dagger A_j^* \otimes B_i^\dagger B_j
\]
\[
= \sum_i \lambda_i A_i^\dagger A_i^* \otimes \sum_j \lambda_j B_j^\dagger B_j.
\]

5.2 Lower bound

Let’s come back to the operation \( S = \sum_i S_i \), where \( S_i(A) = V_i A V_i^\dagger \) and \( V_i = \sum_j \lambda_j A_{ij} \otimes B_{ij} \) is the Schmidt decomposition. By (43), for a TP \( L = \sum_i L_i \) with each \( L_i \) being HP, we have
\[
1 + 2 \left| \sum_i L_i^\dagger (I) \right| = \left| \sum_i \left( L_i^\dagger (I) + L_i^\dagger (I) \right) \right|
\]
along with (94) we obtain,
\[
1 + 2 \left| \sum_i S_i^\dagger (I) \right| = \left| \sum_i \left( S_i^\dagger (I) + S_i^\dagger (I) \right) \right|
\]
\[
= \left| \sum_{j(i)} \sum_{k(i)} \lambda_{ij} A_{ij}^\dagger A_{ij}^* \otimes \lambda_{ik} B_{ik}^\dagger B_{ik} \right|
\]
\[
\leq \sum_i \left| \sum_{j(i)} \sum_{k(i)} \lambda_{ij} A_{ij}^\dagger A_{ij}^* \right| \left| \sum_{k(i)} \lambda_{ik} B_{ik}^\dagger B_{ik} \right|
\]
\[
= \sum_i \left| \sum_{j(i)} \lambda_{ij} A_{ij}^\dagger A_{ij} \right| \left| \sum_{k(i)} \lambda_{ik} B_{ik}^\dagger B_{ik} \right|
\]

5.1.2 In the entirety

In a finite-dimensional system, the following statements for either a unitary operation

Proposition 3. In a finite-dimensional system, the following statements for either a unitary operation \( S(\rho) = U \rho U^\dagger \) or a pure state \( |\psi\rangle \) are equivalent:

1. The Schmidt rank of \( U \) or \( |\psi\rangle \) is 1.
2. The operation/state is separable.
3. The operation/state is PPT.

That PPT pure states are necessarily separable has already been shown in, e.g. [21, 29], where the negativity of a pure state \( |\psi\rangle \) was explicitly calculated, given its Schmidt decomposition. Below we will provide an alternative proof for pure states.

6 PPT and separable unitary operations

Proposition 3. In a finite-dimensional system, the following statements for either a unitary operation \( S(\rho) = U \rho U^\dagger \) or a pure state \( |\psi\rangle \) are equivalent:

1. The Schmidt rank of \( U \) or \( |\psi\rangle \) is 1.
2. The operation/state is separable.
3. The operation/state is PPT.

because \( \{A_i\} \) and \( \{B_i\} \) are orthonormal sets of operators. (96) and (97) together demonstrate the equivalence.
6.1 Proof for pure states

By definition a pure state is separable if and only if its Schmidt rank is 1. Suppose the Schmidt decomposition of a pure state is

\[ |\psi\rangle = \sum_i \lambda_i |a_i\rangle |b_i\rangle. \]  

(98)

Since partial transpose with respect to different bases are related unitarily, we can do this with respect to any basis, and let’s choose \( \{|a_i\}\). We obtain

\[ \rho^\Gamma = (|\psi\rangle \langle \psi|)^\Gamma = \sum_{i,j} \lambda_i \lambda_j |a_i\rangle \langle a_i| \otimes |b_j\rangle \langle b_j|. \]  

(99)

There’s a decomposition of \( \rho^\Gamma \) as \( \rho^\Gamma = \tilde{\rho}^+ - \tilde{\rho}^- \), where

\[ \tilde{\rho}^+ = \sum_i \lambda_i^2 |\langle a_i|b_i\rangle|^2 + \sum_{i<j} \lambda_i \lambda_j \frac{|a_i\rangle \langle a_j| + |a_j\rangle \langle a_i|}{\sqrt{2}} \frac{|b_i\rangle \langle b_j| + |b_j\rangle \langle b_i|}{\sqrt{2}}, \]  

\[ \tilde{\rho}^- = \sum_{i<j} \lambda_i \lambda_j \frac{|a_i\rangle \langle a_j| - |a_j\rangle \langle a_i|}{\sqrt{2}} \frac{|b_i\rangle \langle b_j| - |b_j\rangle \langle b_i|}{\sqrt{2}}. \]  

(100)

As the vectors are orthonormal, \( (100a) \) and \( (100a) \) together are eigendecomposition of \( \rho^\Gamma \); in other words \( \tilde{\rho}^\pm = \rho^\pm \). Therefore, \( \rho^\Gamma = 0 \), i.e. \( \rho \) is PPT if and only if the Schmidt rank of \( |\psi\rangle \) is 1.

6.2 Proof for unitary operations

First, it’s easy to see a unitary operation \( SO = UOU^\dagger \) is separable if \( U \) has Schmidt rank 1. To show it’s separable only if the Schmidt rank is 1, with \( U = \sum_i \lambda_i A_i \otimes B_i \) and \( (91) \):

\[ I_{AB} \otimes U |\Psi_{AB}\rangle = \frac{1}{\sqrt{A_D A_B}} \sum_{i,j,k} \lambda_i I_{AB} \otimes (A_i \otimes B_j) |a_k\rangle \langle a_k| b_j \rangle \langle b_j| \]

\[ = \frac{1}{\sqrt{A_D A_B}} \sum_i \lambda_i \left( I_A \otimes A_i \sum_j |a_j\rangle \langle a_j| \right) \otimes \left( \sum_k I_B \otimes B_k |a_k\rangle \langle a_k| \right) \]

\[ = \sum_i \lambda_i (I_A \otimes A_i |\Psi_A\rangle) \otimes (I_B \otimes B_i |\Psi_B\rangle). \]  

(101)

If \( U \) is separable, then the state above is separable with respect to A and B. Because of \( (93) \), \( I_A \otimes A_i |\Psi_A\rangle \) are orthogonal with \( I_A \otimes A_i |\Psi_A\rangle \) for \( i \neq j \), likewise for B. Therefore, \( (101) \) is a Schmidt decomposition of \( I_{AB} \otimes U |\Psi_{AB}\rangle \) (after normalizing \( I_A \otimes A_i |\Psi_A\rangle \) and \( I_B \otimes B_j |\Psi_B\rangle \)), and this vector is separable if and only if the Schmidt rank of \( U \) is 1. Therefore, no unitary operators with Schmidt rank higher than 1 are separable, so a unitary operator is separable if and only if its Schmidt rank is 1.

From \( (86) \) and \( (89) \), it’s clear that a unitary operation is PPT if and only if there exist no \( V_{ij}^- \), which happens if and only if the Schmidt rank of the unitary operator is 1. Thus a unitary operation is PPT if and only if it’s separable.

\( (96) \) and \( (97) \) also confirm this: The unitary operation is PPT if and only if \( \text{tr} \mathcal{F} (S^\Gamma)^- = \| S^{\Gamma^\dagger}_\downarrow (I) \|_1 = 0 \), i.e. \( (\sum_i \lambda_i)^2 = d_A d_B \). Because each \( \lambda_i > 0 \), \( (\sum_i \lambda_i)^2 \geq \sum_i \lambda_i^2 \), and they are equal if and only if there’s only one \( \lambda_i \). Since

\[ \text{tr} U^\dagger U = \text{tr} I = (U |U\rangle = \sum_i \lambda_i A_i \otimes B_i \sum_j \lambda_j A_j \otimes B_j) = d_A d_B, \]  

(102)

for \( (\sum_i \lambda_i)^2 \) to become \( d_A d_B \) the Schmidt rank of \( U \) must be 1, i.e. separable. \( \square \)

7 Schmidt decomposition and eigendecomposition

The method from Ref. [6] used to show Lemma 5 and Proposition 3 can be further generalized as follows.

As was shown in Sec. 5.1, for a (sub-)operation \( S(A) = VAV^\dagger \) where the Schmidt decomposition of \( V \) is \( \sum_i \lambda_i A_i \otimes B_i \), \( S^\Gamma_k \) can be determined from the Schmidt decomposition of \( V \), in particular \( (86) \). Let’s express
and because of (93), we can easily see
\[ T_L \]
we can find a decomposition of
Because both \( \{A_i\} \) and \( \{B_i\} \) are composed of orthonormal operators, \((V_{ij}^\pm V_{kl}^\pm) = \delta_{ik}\delta_{jl}\), and thus \( V_{ij}^+ : i \leq j \) and \( V_{ij}^- : i < j \) are sets of orthonormal operators, and \( V_{ij}^+ \perp V_{kl}^- \). Therefore after Choi isomorphism, \( \mathcal{F}(S^T) \) becomes an ensemble of orthogonal pure states, c.f. (92) and (93), giving us an eigendecomposition of it.

To find the eigenvalues and eigenvectors of \( \mathcal{F}(S^T) \), because \( \{V_{ij}^+ : i \leq j\} \) and \( \{V_{ij}^- : i < j\} \) are normalized and because of (93), we can easily see \( \mathcal{F}(S^T) \) has
eigenvectors \( \sqrt{d}I \otimes V_{ij}^\pm |\Psi\rangle \) with corresponding eigenvalues \( \pm \lambda_i \lambda_j \), \( i \leq j \) for \( V_{ij}^+ \) and \( i < j \) for \( V_{ij}^- \). (104)

For example, for a unitary operation \( S(A) = UAU^\dagger \) where \( U = \sum_{i=1}^{2} \lambda_i A_i \otimes B_i \) with Schmidt rank 2, \( \mathcal{F}(S^T) \) has three positive eigenvalues and a negative one: \( \lambda_1^2, \lambda_2^2 \), and \( \pm \lambda_1 \lambda_2 \).

In general, given a linear mapping \( L \) of this form (c.f. (85)):
\[ L(O) = \sum_{i,j} A_i \otimes B_j O(A_i \otimes B_j)^\dagger, \] (105)
we can find a decomposition of \( L \): \( L = \tilde{L}_+ - \tilde{L}_- \). If \( \{A_i\} \) and \( \{B_i\} \) are sets of orthogonal operators, then \( \tilde{L}_+ = \tilde{L}_- \), and we can determine the spectrum of \( \mathcal{F}(L) \) by \( \{A_i\} \) and \( \{B_i\} \) alone.

Similarly, with an operator of the form
\[ H = \sum_{i,j} |v_j\rangle\langle v_i| \otimes |w_i\rangle\langle w_j|, \] (106)
where \( |v_i\rangle \) and \( |w_i\rangle \) aren’t necessarily orthonormal, we can obtain
\[ \tilde{H}^+ = \sum_i (|v_i\rangle|w_i\rangle)(|v_i\rangle\langle w_i|) + \frac{1}{2} \sum_{i<j} (|v_i\rangle|w_j\rangle + |v_j\rangle|w_i\rangle)(|v_i\rangle\langle w_j| + \langle v_j|\langle w_i|) \]
\[ \tilde{H}^- = \frac{1}{2} \sum_{i<j} (|v_i\rangle\langle w_j| - \langle v_j|\langle w_i|)(|v_i\rangle\langle w_j| - \langle v_j|\langle w_i|). \] (107)
If \( \{|v_i\}\) and \( \{|w_i\}\) are orthogonal sets of vectors, this gives us the eigendecomposition of \( H \).

8 Norms

8.1 Upper bound of \( ||S^T(I)|| \)
To prove the inequality
\[ ||S^T(I)|| \leq ||S^\dagger(I)|| = (||S||_{1, r} - d_A d_B)/2 \]
is quite easy: By the inequality \( ||O|| \leq ||O||_1 \) [10, 27], \( ||S^\dagger(I)|| \leq ||S^T(I)|| \) by Corollary 1, \( \text{tr}S^\dagger(I) = \text{tr}\mathcal{F}(S^\dagger) = ||S^\dagger(I)|| \) by (42) and (45), \( ||\mathcal{F}(S^T)||_1 = \text{tr}\mathcal{F}(S^\dagger) + 2\text{tr}\mathcal{F}(S^T) = d_A d_B + 2\text{tr}\mathcal{F}(S^T) \), so
\[ 2\text{tr}S^\dagger(I) = ||S||_{1, r} - d_A d_B. \]

The difference between the two sides of the inequality, \( ||S^\dagger(I)||_1 - ||S^T(I)||_1 = \text{tr}\mathcal{F}(S^\dagger) - ||S^T(I)|| \), is the sum of the eigenvalues of \( S^\dagger(I) \) minus the largest one. When the operation is PPT, the right side of the inequality is zero.

8.2 Equivalence of norms
Here we will show
\[ \frac{||H||_1}{\min(d_A, d_B)} \leq ||H||_{1,r} \leq \min(d_A, d_B)||H||_1. \] (108)

19
By Ref. [21], for a state with Schmidt decomposition \(|\psi\rangle = \sum_i \lambda_i |a_i\rangle |b_i\rangle\)

\[ ||(|\psi\rangle \langle \psi|)^{\Gamma}||_1 = ||(|\psi\rangle \langle \psi|)||_{1,\Gamma} = \left( \sum_i \lambda_i \right)^2. \]  

(109)

This can be proved with the aid of a unitary operator as in Ref. [21], or

\[ [(|\psi\rangle \langle \psi|)^{\Gamma}]^2 = \sum_i \lambda_i^2 |a_i\rangle \langle a_i| \otimes \sum_j \lambda_j^2 |b_j\rangle \langle b_j|, \]

so \(|(|\psi\rangle \langle \psi|)^{\Gamma}|| = \sum_{i,j} \lambda_i \lambda_j |a_i\rangle \langle a_i| \otimes |b_j\rangle \langle b_j|\) and \(|(|\psi\rangle \langle \psi|)^{\Gamma}||_1 = (\sum_i \lambda_i)^2. \)

Because the Schmidt rank is at most \(\min(d_A, d_B)\), by Schwarz inequality

\[ \left( \sum_i \lambda_i \right)^2 \leq \sum_i \lambda^2 \sum_j 1^2 = \min(d_A, d_B). \]  

(111)

Hence, for any Hermitian operator \(H\) with spectral decomposition \(H = \sum_i h_i |\psi_i\rangle \langle \psi_i|\)

\[ ||H^{\Gamma}||_1 = ||H||_{1,\Gamma} \leq \sum_i |h_i| ||(|\psi_i\rangle \langle \psi_i|)||_{1,\Gamma} \leq \min(d_A, d_B) \sum_i |h_i| = \min(d_A, d_B)||H||_1. \]  

(112)

Since \((H^{\Gamma})^\Gamma = H\) (and because \(\Gamma\) is HP), for the same reason \(||H||_1 \leq \min(d_A, d_B)||H^{\Gamma}||_1 = \min(d_A, d_B)||H||_{1,\Gamma}.\)

As for HP mappings, \(\mathcal{F}(L) \in \mathcal{L}(\mathcal{H}_{A_1} \otimes \mathcal{H}_{A_2} \otimes \mathcal{H}_{B_1} \otimes \mathcal{H}_{B_2})\), and \(\Gamma = T_{A_1A_2} \otimes \mathcal{I}_{B_1B_2}\). The Schmidt rank of a bipartite state, with the parties being \((\mathcal{H}_{A_1} \otimes \mathcal{H}_{A_2})\) and \((\mathcal{H}_{B_1} \otimes \mathcal{H}_{B_2})\), is at most \(\min(d_{A_1}, d_{B_1}, d_{B_2})\), so

\[ \frac{||L||_1}{\min(d_{A_1}, d_{A_2}, d_{B_1}, d_{B_2})} \leq ||L||_{1,\Gamma} \leq \min(d_{A_1}, d_{A_2}, d_{B_1}, d_{B_2}) ||L||_1. \]  

(113)

For non-Hermitian operators and non-HP mappings, we except similar relations to exist.

9 Reaching the entangling capacity

9.1 Proposition 4

Let’s show the following proposition:

**Proposition 4.** Suppose \(\mathcal{S}_\pm = S_\pm^{\Gamma}\). Let ran denote the range or image of a mapping, and

\[ S_\pm^{\Gamma}(O) = \sum_i c_i^\pm V_i^\pm O V_i^\pm \dagger, \quad c_i^\pm > 0, \]

(144)

be operator-sum representations of \(S_\pm^{\Gamma}\), and

\[ \rho^{\Gamma\pm} = \sum_i |\psi_i^\pm\rangle \langle \psi_i^\pm|, \]

(115)

be ensembles of \(\rho^{\Gamma\pm}\). The upper bound of the entangling capacity given by Proposition 1 is reached if and only if the two conditions below are both satisfied:

1. ran \([S_\pm^{\Gamma}(\rho^{\Gamma+}) + S_\pm^{\Gamma}(\rho^{\Gamma-})]\) and ran \([S_\pm^{\Gamma}(\rho^{\Gamma+}) + S_\pm^{\Gamma}(\rho^{\Gamma-})]\) are orthogonal, which is equivalent to the orthogonality of the following vectors:

\[ \langle V_i^\pm \psi_i^\mp | V_k^\mp \psi_k^\mp \rangle = 0 \quad \text{and} \quad \langle V_i^\pm \psi_i^\mp | V_j^\mp \psi_j^\mp \rangle = 0 \quad \forall i, j, k, l. \]

(116)

2. ran \(\rho^{\Gamma}\) is a subspace of the eigenspace relative to the largest eigenvalue of \(S_\pm^{\Gamma\dagger}(I)\).

The second condition above is satisfied by any state \(\rho\) when \(S_\pm^{\Gamma\dagger}(I) \preceq I\). Besides, the upper and lower bounds in (41) are the same if and only if \(S_\pm^{\Gamma\dagger}(I) \preceq I\).
Proof. First, for $P \geq 0$, $||P||_1/d = \text{tr}P/d$, where $d$ is the dimension of the space, is the average of its eigenvalues (zero included), whereas $||P||$ is its largest eigenvalue. Hence they’re the same if and only if $P \propto I$, and $||S^+\Gamma I)\|/(d_A d_B) = ||S^+\Gamma I)\|$ if and only if $S^+\Gamma I) \propto I$.

For the rest of the proof please read the proof of Lemma 4 if necessary. For $\text{tr}[S^+\Gamma I)(\rho^+ + \rho^-)]$ to equal $||S^+\Gamma I)|| ||\rho^||_1$, it’s necessary and sufficient that $V$ is a subspace of the eigenspace of $S^+\Gamma I)\$’s largest eigenvalue, i.e. the second condition, and it’s satisfied when $S^+\Gamma I \propto I; see the discussion below (14). Note because $S^\Gamma I) = S^\Gamma I) + I$, their spectra differ by 1 exactly.

For $\text{tr}(S^\Gamma I)\$ to be equal, by Lemma 1 it’s necessary and sufficient that the eigenvectors of $S^\Gamma I)$ and those of $S^\Gamma I)$ are orthogonal with each other, that is, they are the eigendecomposition $S^\Gamma I)$. From (114) and (115) we acquire ensembles of $S^\Gamma I) + S^\Gamma I)$:

$$S^\Gamma I) + S^\Gamma I) = \sum_{i,j} c^\Gamma I V^\dagger_i \tilde{\psi}^+_j \tilde{\psi}^+_j V^\dagger_i + \sum_{k,l} c^\Gamma k V^\dagger_k \tilde{\psi}^+_l \tilde{\psi}^+_l V^\dagger_k.$$  

(117)

By Lemma 2, the eigenvectors of $S^\Gamma I) + S^\Gamma I)$ and those of $S^\Gamma I) + S^\Gamma I)$ are orthogonal with each other if and only if

$$\langle V^\dagger_i \tilde{\psi}^+_j | V^\dagger_k \tilde{\psi}^+_l \rangle = 0 \forall i,j,k,l$$  

(118)

and

$$\langle V^\dagger_i \tilde{\psi}^+_j | V^\dagger_k \tilde{\psi}^+_l \rangle = 0 \forall i,j,k,l.$$  

(119)

the first condition.

\[\square\]

9.2 $2 \otimes 2$ unitary operation

The unitary operator is

$$\begin{pmatrix}
\cos \alpha & \sin \alpha & 0 & 0 \\
-\sin \alpha & \cos \alpha & 0 & 0 \\
0 & 0 & \cos \beta & \sin \beta \\
0 & 0 & -\sin \beta & \cos \beta
\end{pmatrix}.$$  

The eigenvalues of $\mathcal{G}(\sigma^\Gamma I)$ are:

$$\{c^\Gamma \} = \{2 + 2 \cos(\beta - \alpha), 2 |\sin(\beta - \alpha)|\}, c^- = -2 |\sin(\beta - \alpha)|.$$

For pure and separable state $|\psi\rangle = |\psi_1\rangle |\psi_2\rangle$, (116) becomes

$$\langle (\psi_1^\dagger | \psi_2\rangle | V^\dagger_i \tilde{\psi}^+_j (|\psi_1\rangle |\psi_2\rangle) \rangle = 0 \forall i,j.$$  

(120)

With $|\psi_i\rangle = \cos \theta_i |\uparrow\rangle + e^{i\phi_i} \sin \theta_i |\downarrow\rangle$, depending on the sign of $\sin(\beta - \alpha)$,

$$\{\langle \psi_1^\dagger | \psi_2 | V^{-1} \tilde{\psi}^+_1 |\psi_1\rangle |\psi_2\rangle \} = \{ \pm 2i \sec(\beta - \alpha) \cos(2\theta_1) \sin(2\theta_2) \sin \phi_2, 2 \cos(2\theta_1) \pm i \sin(2\theta_2) \sin \phi_2, 2 - \cos(2\theta_1) \mp i \sin(2\theta_2) \sin \phi_2 \},$$

taking the upper ones when $\sin(\beta - \alpha) \geq 0$. For all of them to vanish, $\theta_1 = \pi/4 + n\pi/2$, and $\theta_2 = m\pi/2$ or $\phi_2 = p\pi$, with $m, n, p \in \mathbb{Z}$.

9.3 $2 \otimes 3$ unitary operation

Consider a $2 \otimes 3$ unitary operator

$$\begin{pmatrix}
1_3 & 0_{3 \times 3} \\
0_{3 \times 3} & \cos \beta & -\sin \beta \\
0_{3 \times 3} & \sin \beta & \cos \beta \\
0 & 0 & 1
\end{pmatrix} \cdot \begin{pmatrix}
1_4 & 0_{4 \times 3} \\
0_{4 \times 3} & \cos \alpha & \sin \alpha \\
0_{4 \times 4} & -\sin \alpha & \cos \alpha
\end{pmatrix}.$$  

(121)
For the upper and lower bound to be the same with $\tilde{S}_{\Gamma} = S_{\Gamma}$, we need $S_{\Gamma}^\dagger (I) \propto I$, i.e. the maximum and minimum eigenvalues of $S_{\Gamma}^\dagger (I) \propto I$ should be identical. From Fig. 1 they seem to be the same at 1/2 when $\alpha = 2\pi/3$ and $\beta = 0$, which can indeed be verified analytically, so the exact entangling capacity at $\alpha = 2\pi/3$ and $\beta = 0$ can be acquired.

![Figure 1: The max/min eigenvalues of $S_{\Gamma}^\dagger (I)$, with eigenvalue 0 taken into account.](image)

9.3.1 Degeneracy

When $\alpha = 2\pi/3$ and $\beta = 0$, the spectrum of $\mathcal{T}(S_{\Gamma})$ is $\{-3, 3, 3, 3\}$, which has degeneracy. When trying to find the optimal state for it i.e. to solve (120), the eigenvectors with eigenvalue 3 that came out of the program were not orthogonal but were linearly independent, and we did not try to orthogonalize them for the following reason:

The dimension of an eigenspace with a degenerate eigenvalue is higher than one. Suppose the said eigenspace can be spanned by $\{|v_i\rangle\}$ and $\{|\tilde{v}_i\rangle\}$, both with the same cardinality but the former being an orthonormal set and the latter being normalized but not orthogonal. Hence one is the linear combination of the other:

$$|v_i\rangle = \sum_j c_{ij} |\tilde{v}_j\rangle, \quad c_{ij} \text{ is non-singular.}$$

The Kraus operator corresponding to an eigenvector $|v_i\rangle = \sum_{j,k} d_{ijk}^j |a_j\rangle |b_k\rangle$ of a linear mapping is

$$V_i := V(|v_i\rangle) := \sum_{j,k} d_{ijk}^j |b_k\rangle \langle a_j|,$$

which is a linear process, that is, $V_i = \sum_j c_{ij} \tilde{V}_j$, where $\tilde{V}_j := V(|\tilde{v}_j\rangle)$. Therefore, $V_i |\psi\rangle = \sum_j c_{ij} \tilde{V}_j |\psi\rangle$, and $\{V_i |\psi\rangle\}$ has the same span as $\{\tilde{V}_i |\psi\rangle\}$ does. For two (sub)spaces to be orthogonal, it does not matter which vectors we choose to span them, just as discussed in the proof for Lemma 2. This explains why there is no need to orthogonalize eigenvectors with the same eigenvalue if we just want to find the optimal state.

9.3.2 The solution

Assuming $|\psi\rangle = (\cos \theta_1 |\uparrow\rangle + \sin \theta_1 |\downarrow\rangle) \otimes (\cos \theta_2 \sin \phi |0\rangle + \sin \theta_2 \sin \phi |1\rangle + \cos \phi |2\rangle)$, for this operation at $\alpha = 2\pi/3$ and $\beta = 0$, (120) becomes

$$0 = 4 \sin^2 \theta_1 - 3f,$$
$$0 = 4g \sin^2 \theta_1 - f,$$
$$0 = \cos^2 \theta_1 (1 - 3g),$$

where

$$f = \cos^2 \theta_1 (1 + g), \quad g = \sin^2 \phi \cos^2 \theta_2.$$
From (125), either \( \cos^2 \theta_1 = 0 \) or \( 1 - 3g = 0 \). However, \( \cos^2 \theta_1 \neq 0 \); otherwise \( f \) would vanish, and from (123), \( \sin^2 \theta_1 = 0 \), a contradiction. Hence, \( g = 1/3 \), so \( f = 4 \cos^2 \theta_1/3 \), and (124) is satisfied. (123) now becomes

\[
4 \sin^2 \theta_1 - 4 \cos^2 \theta_1 = 0 = -4 \cos 2\theta_1.
\]

Thus the solution is

\[
\sin^2 \phi \cos^2 \theta_2 = 1/3 \text{ and } \cos 2\theta_1 = 0.
\]

Figure 2: Upper and lower bounds for EC\(_L\) (base 2) of operations of the form \( S(\rho) = pS_1 + (1 - p)S_2 \) with unitary \( S_i \), the lower bound accompanied as a reference. For both the \( 2 \otimes 3 \) and \( 3 \otimes 3 \) operations, \( S_1 \) has \( \alpha = \pi/3 \) and \( \beta = \pi/5 \), and \( S_2 \) has for \( S_2 \), \( \alpha = \pi/4 \) and \( \beta = \pi/3 \). Upper bound 1 refers to that obtained with \( S_T^\Gamma \), and upper bound 2 with \( pS_T^\Gamma - (1 - p)S_T^\Gamma \). Fig. 2a clearly demonstrate that \( S_T^\Gamma(I) \) in general doesn’t have the smallest operator norm among all \( S_T^\Gamma(I) \).

9.4 \( S_T^\Gamma \) does not necessarily have the smallest operator norm

For a Hermitian operator \( H \), because there exist \( \tilde{H} \pm \) such that \( \tilde{H} \pm - H \pm \neq 0 \) (see Sec. 2.1.1), \( S_T^\Gamma - S_T^\Gamma \) is not always CP, and \( S_T^\Gamma(I) - S_T^\Gamma(I) \) can be non-positive. Therefore it can happen that \( ||S_T^\Gamma(I)|| > ||S_T^\Gamma(I)|| \); nevertheless, \( S_T^\Gamma(I) \) has the smallest trace (norm), because of Lemma 1 and Corollary 1. How to minimize the operator norm under such decompositions may be an interesting mathematical problem.

To show \( ||S_T^\Gamma(I)|| \neq ||S_T^\Gamma(I)|| \) in general, we consider a deterministic operation \( S = pS_1 + (1 - p)S_2 \), where \( 0 \leq p \leq 1 \) and \( S_i \) are unitary operations. We choose \( 2 \otimes 3 \) and \( 3 \otimes 3 \) unitary operations, the former being (121)
and the latter being
\[
\begin{pmatrix}
I_6 & 0_{6\times3} \\
0_{3\times6} & \begin{pmatrix}
\cos \beta & \sin \beta & 0 \\
-\sin \beta & \cos \beta & 0 \\
0 & 0 & 1
\end{pmatrix}
\end{pmatrix}
\begin{pmatrix}
I_7 & 0_{7\times2} \\
0_{2\times7} & \begin{pmatrix}
\cos \alpha & \sin \alpha & 0 \\
-\sin \alpha & \cos \alpha & 0
\end{pmatrix}
\end{pmatrix}.
\]

The result is illustrated by Fig. 2, where we can see \( ||S^{r+}_-(I)|| > ||\tilde{S}^{r+}_-(I)|| \) with \( \tilde{S}^{r+}_- = pS^{r+}_1 + (1 - p)S^{r+}_2 \) in some cases.

References

[1] J.-Y. Kao, arXiv:2008.03893 [quant-ph] (2020), URL https://arxiv.org/abs/2008.03893.
[2] E. Størmer, Positive linear maps of operator algebras (Springer, Berlin/Heidelberg, 2013).
[3] B. Kraus and J. I. Cirac, Phys. Rev. A 63, 062309 (2001), URL https://link.aps.org/doi/10.1103/PhysRevA.63.062309.
[4] M. S. Leifer, L. Henderson, and N. Linden, Phys. Rev. A 67, 012306 (2003), URL https://link.aps.org/doi/10.1103/PhysRevA.67.012306.
[5] A. Chefles, Phys. Rev. A 72, 042332 (2005), URL https://link.aps.org/doi/10.1103/PhysRevA.72.042332.
[6] E. T. Campbell, Phys. Rev. A 82, 042314 (2010), URL https://link.aps.org/doi/10.1103/PhysRevA.82.042314.
[7] L. H. Loomis and S. Sternberg, Advanced Calculus (Jones and Bartlett Publishers, Boston, 1990), rev. ed.
[8] L. Tu, An Introduction to Manifolds (Springer, New York, 2010), 2nd ed.
[9] R. A. Horn and C. R. Johnson, Matrix Analysis (Cambridge University Press, Cambridge, 2013), 2nd ed.
[10] A. Jamiołkowski, Rep. Math. Phys. 3, 275 (1972), ISSN 0034-4877, URL http://www.sciencedirect.com/science/article/pii/0034487772900110.
[11] K. Kraus, Ann. Phys. 64, 311 (1971), ISSN 0003-4916, URL http://www.sciencedirect.com/science/article/pii/0003491671901084.
[12] M.-D. Choi, Linear Algebra Its Appl. 10, 285 (1975), ISSN 0024-3795, URL http://www.sciencedirect.com/science/article/pii/0024379575900750.
[13] I. Bengtsson and K. Zyczkowski, Geometry of Quantum States: An Introduction to Quantum Entanglement (Cambridge University Press, Cambridge, 2008).
[14] J. de Pillis, Pacific J. Math. 23, 129 (1967), URL https://projecteuclid.org/euclid.pjm/1102991990.
[15] M. Jiang, S. Luo, and S. Fu, Phys. Rev. A 87, 022310 (2013), URL https://link.aps.org/doi/10.1103/PhysRevA.87.022310.
[16] E. M. Rains, Phys. Rev. A 60, 173 (1999), URL https://link.aps.org/doi/10.1103/PhysRevA.60.173.
[17] W. Rudin, Real and Complex Analysis (McGraw-Hill, New York, 1987), 3rd ed.
[18] R. Bhatia, Matrix Analysis (Springer, New York, 1997).
[19] A. E. Rastegin, J. Stat. Phys. 148, 1040 (2012), URL https://doi.org/10.1007/s10955-012-0569-8.
[20] B. Baumgartner (2011), math-ph/1106.6189.
[21] G. Vidal and R. F. Werner, Phys. Rev. A 65, 032314 (2002), URL http://link.aps.org/doi/10.1103/PhysRevA.65.032314.
[22] S. Roman, *Advanced Linear Algebra* (Springer, New York, 2008), 3rd ed.

[23] L. P. Hughston, R. Jozsa, and W. K. Wootters, Phys. Lett. A 183, 14 (1993), ISSN 0375-9601, URL http://www.sciencedirect.com/science/article/pii/0375960193908809.

[24] P. Horodecki, Phys. Lett. A 232, 333 (1997), ISSN 0375-9601, URL http://www.sciencedirect.com/science/article/pii/0375960197004167.

[25] M. B. Plenio, Phys. Rev. Lett. 95, 090503 (2005), URL http://link.aps.org/doi/10.1103/PhysRevLett.95.090503.

[26] T. M. Cover and J. A. Thomas, *Elements of information theory* (John Wiley & Sons, Inc., Hoboken, 2006), 2nd ed.

[27] D. A. Lidar, P. Zanardi, and K. Khodjasteh, Phys. Rev. A 78, 012308 (2008), URL https://link.aps.org/doi/10.1103/PhysRevA.78.012308.

[28] M. A. Nielsen, C. M. Dawson, J. L. Dodd, A. Gilchrist, D. Mortimer, T. J. Osborne, M. J. Bremner, A. W. Harrow, and A. Hines, Phys. Rev. A 67, 052301 (2003), URL https://link.aps.org/doi/10.1103/PhysRevA.67.052301.

[29] K. Życzkowski and I. Bengtsson, Annals of Physics 295, 115 (2002), ISSN 0003-4916, URL http://www.sciencedirect.com/science/article/pii/S0003491601962013.