
This version is available at https://strathprints.strath.ac.uk/54109/

Strathprints is designed to allow users to access the research output of the University of Strathclyde. Unless otherwise explicitly stated on the manuscript, Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Please check the manuscript for details of any other licences that may have been applied. You may not engage in further distribution of the material for any profitmaking activities or any commercial gain. You may freely distribute both the url (https://strathprints.strath.ac.uk/) and the content of this paper for research or private study, educational, or not-for-profit purposes without prior permission or charge.

Any correspondence concerning this service should be sent to the Strathprints administrator: strathprints@strath.ac.uk
I. SUPPLEMENTARY NOTES

A. Supplementary Note 1 - Proof of Theorem 1

We will use a variant of Lemma 20 of [1].

Lemma 1. Consider a Hermitian matrix $L_{AB} \in \mathbb{B}(\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B})$, with $d_A \leq d_B$. Then

$$\|L_{AB}\|_1 \leq d_A^2 \max_{\mathcal{M}_B} \|\text{id}_A \otimes \mathcal{M}_B (L_{AB})\|_1,$$

where the maximum is taken over local measurement maps $\mathcal{M}_B(Y) = \sum_i \text{tr}(N_i Y) \langle l | l \rangle$, with a POVM $\{N_i\}$ and orthonormal states $\{|l\rangle\}$.

Proof. Write $L_{AB} = \sum_{i,j=1}^{d_A} |i\rangle \langle j| \otimes L_{ij}$ with $\{|i\rangle\}$ an orthornomal basis for $\mathbb{C}^{d_A}$. On one hand, thanks to the triangle inequality, we have

$$\|L_{AB}\|_1 = \left\| \sum_{i,j=1}^{d_A} |i\rangle \langle j| \otimes L_{ij} \right\|_1 \leq d_A^2 \max_{i,j} \|L_{ij}\|_1. \quad (1)$$

On the other hand,

$$\max_{\mathcal{M}_B} \|\text{id}_A \otimes \mathcal{M}_B (L_{AB})\|_1 \leq \max_{\mathcal{M}_B} \left\| \sum_{i,j=1}^{d_A} |i\rangle \langle j| \otimes \mathcal{M}_B(L_{ij}) \right\|_1 \leq \max_{\mathcal{M}_B} \left\| \text{id}_A \otimes \mathcal{M}_B (L_{AB}) \right\|_1 \leq d_A^2 \max_{i,j} \|L_{ij}\|_1. \quad (2)$$

where we have repeatedly used the expression of the trace norm $\|X\|_1 = \max_{\|K\| \leq 1} |\text{tr}(K X)|$, and the alternative choices $K_A = |i\rangle \langle i|$, $K_A = |i\rangle \langle j| + |j\rangle \langle i|$, or $K_A = i(|i\rangle \langle j| - |j\rangle \langle i|)$ to arrive to the last inequality.

It is clear that

$$\max_{\mathcal{M}_B} \|\mathcal{M}_B(L_{ii})\|_1 = \|L_{ii}\|_1 \quad (3)$$

and similarly

$$\max_{\mathcal{M}_B} \|\mathcal{M}_B(L_{ij} + L_{ji})\|_1 = \|L_{ij} + L_{ji}\|_1, \quad \max_{\mathcal{M}_B} \|\mathcal{M}_B(i(L_{ij} - L_{ji}))\|_1 = \|L_{ij} - L_{ji}\|_1. \quad (4)$$

To complete the proof it is enough to observe

$$\|L_{ij}\|_1 \leq \frac{1}{2} (\|L_{ij} + L_{ji}\|_1 + \|L_{ij} - L_{ji}\|_1) \leq \max \{\|L_{ij} + L_{ji}\|_1, \|L_{ij} - L_{ji}\|_1\}. \quad (5)$$

□
A second lemma bounds the optimal distinguishability of two quantum channels (i.e. their diamond-norm distance) in terms of the distinguishability of their corresponding Choi-Jamiołkowski states.

**Lemma 2.** Let \( \Phi_{AA'} = d_A^{-1} \sum_{k,k'} |k,k\rangle \langle k',k'| \) be a \( d_A \)-dimensional maximally entangled state. For any ctpm map \( \Lambda : \mathcal{D}(A) \to \mathcal{D}(B) \) we define the Choi-Jamiołkowski state of \( \Lambda \) as \( J(\Lambda) := \text{id}_A \otimes \Lambda'_{A|A'}(\Phi_{AA'}) \). For two ctpm maps \( \Lambda_0 \) and \( \Lambda_1 \) it then holds
\[
\frac{1}{d_A} \| \Lambda_0 - \Lambda_1 \|_1 \leq \| J(\Lambda_0) - J(\Lambda_1) \|_1 \leq \| \Lambda_0 - \Lambda_1 \|_1. \tag{6}
\]

**Proof.** The second inequality in (6) is trivial, as the diamond norm between two ctpm maps is defined through a maximization over input states, while \( \| J(\Lambda_0) - J(\Lambda_1) \|_1 \) corresponds to the bias in distinguishing the two operations \( \Lambda_0 \) and \( \Lambda_1 \) by using the maximally entangled state \( \Phi_{AA'} \) as input. The first inequality can be derived as follows.

Any pure state \( |\psi\rangle_{AA'} \) can be obtained by means of a local filtering of the maximally entangled state, i.e.,
\[
|\psi\rangle_{AA'} = (\sqrt{d_A} C \otimes 1)|\Phi\rangle_{AA'}
\]
for a suitable \( C \in \mathbb{B}(\mathcal{C}^{d_A}) \), which, for a normalized \( |\psi\rangle_{AA'} \) satisfies \( \text{tr}(C^\dagger C) = 1 \). From the latter condition, we have that \( \| C \|_\infty \leq 1 \). Let \( |\psi\rangle_{AA'} \) be a normalized pure state optimal for the sake of the diamond norm between \( \Lambda_0 \) and \( \Lambda_1 \). We find
\[
\| \Lambda_0 - \Lambda_1 \|_1 = \| \text{id}_A \otimes (\Lambda_0 - \Lambda_1)(|\psi\rangle\langle \psi|) \|_1
= \left\| \text{id}_A \otimes (\Lambda_0 - \Lambda_1) \left( (\sqrt{d_A} C \otimes 1)\Phi_{AA'}(\sqrt{d_A} C \otimes 1)\right) \right\|_1
= \left\| (\sqrt{d_A} C \otimes 1)(\text{id}_A \otimes (\Lambda_0 - \Lambda_1)|\Phi_{AA'}\rangle\langle \Phi_{AA'}|) \right\|_1
\leq d_A \| C \|_\infty^2 \| \text{id}_A \otimes (\Lambda_0 - \Lambda_1)|\Phi_{AA'}\|_1
\leq d_A \| J(\Lambda_0) - J(\Lambda_1) \|_1,
\]
where we used (twice) Hölder’s inequality \( \| MN \|_1 \leq \min\{\| M \|_\infty \| N \|_1, \| M \|_1 \| N \|_\infty \} \) in the first inequality, and \( \| C \|_\infty \leq 1 \) in the second inequality. \( \square \)

We are in position to prove the main theorem, which we restate for the convenience of the reader.

**Theorem 1 (restatement).** Let \( \Lambda : \mathcal{D}(A) \to \mathcal{D}(B_1 \otimes \ldots \otimes B_n) \) be a ctpm map. Define \( \Lambda_j := \text{tr} \circ B_j \circ \Lambda \) as the effective dynamics from \( \mathcal{D}(A) \) to \( \mathcal{D}(B_j) \) and fix a number \( 1 > \delta > 0 \). Then there exists a measurement \( \{ M_k \}_k \) with \( M_k \geq 0, \sum_k M_k = I \) and a set \( S \subseteq \{ 1, \ldots, n \} \) with \( |S| \geq (1 - \delta)n \) such that for all \( j \in S \),
\[
\| \Lambda_j - \mathcal{E}_j \|_\infty \leq \left( \frac{27 \ln(2)(d_A)^{3} \log(d_A)}{n \delta^3} \right)^{1/3}, \tag{7}
\]
with
\[
\mathcal{E}_j(X) := \sum_k \text{tr}(M_k X)\sigma_{j,k}, \tag{8}
\]
for states \( \sigma_{j,k} \in \mathcal{D}(B_j) \). Here \( d_A \) is the dimension of the space \( A \).
Proof. Let \( \Phi_{AA'} = d_A^{-1} \sum_{k,k'} |k,k\rangle \langle k', k'| \) be a \( d_A \)-dimensional maximally entangled state and 
\( \rho_{AB_1,\ldots,B_n} := \text{id}_A \otimes \Lambda(\Phi_{AA'}) \) be the Choi-Jamiołkowski state of \( \Lambda \) [2]. Define \( \pi := \text{id}_A \otimes M_1 \otimes \cdots \otimes M_n(\rho) \), for quantum-classical channels \( M_1, \ldots, M_n \) defined as \( M_i(X) := \sum_l \text{tr}(N_{i,l}X)|l\rangle \langle l| \), for a POVM \( \{N_{i,l}\}_{l} \).

We will proceed in two steps. In the first we show that conditioned on measuring a few of the \( B_i \)'s of \( \rho_{AB_1,\ldots,B_n} \), the conditional mutual information of \( A \) and \( B_i \) (on average over \( i \)) is small. In the second we show that this implies that the reduced state \( \rho_{AB_i} \) is close to a separable state \( \sum_z p(z)\rho_{z,A} \otimes \rho_{z,B_i} \) with the ensemble \( \{p(z), \rho_{z,A}\} \) independent of \( i \). We will conclude showing that by the properties of the Choi-Jamiołkowski isomorphism, this implies that the effective channel from \( A \) to \( B_i \) is close to a measure-and-prepare channel with a POVM independent of \( i \).

Let \( \mu \) be the uniform distribution over \([n]\) and define \( \mu^\land k \) as the distribution on \([n]^k \) obtained by sampling \( m \) times without replacement according to \( \mu \); i.e.,

\[
\mu^\land k (i_1, \ldots, i_k) = \begin{cases} 
0 & \text{if } i_1, \ldots, i_k \text{ are not all distinct} \\
\frac{\mu(i_1) \cdots \mu(i_k)}{\sum_{j_1, \ldots, j_k \text{ distinct}} \mu(j_1) \cdots \mu(j_k)} & \text{otherwise} 
\end{cases}
\]

Then

\[
\log d_A \geq \mathbb{E}_{(j_1, \ldots, j_k) \sim \mu^\land k} \max_{M_{j_1}, \ldots, M_{j_k}} I(A : B_{j_1}, \ldots, B_{j_k}) \pi 
\]

\[
= \mathbb{E}_{(j_1, \ldots, j_k) \sim \mu^\land k} \max_{M_{j_1}, \ldots, M_{j_k}} \left( I(A : B_{j_1}) + \ldots + I(A : B_{j_k} | B_{j_1}, \ldots, B_{j_{k-1}}) \right) 
\]

\[
=: f(k).
\]

The inequality comes from the fact that \( \pi \) is separable between \( A \) and \( B_1 B_2 \ldots B_n \) because of the action of the quantum-classical channels \( M_1, \ldots, M_n \). The second line follows from the chain rule of mutual information given by Eq. (54) in Supplementary Methods.

Define \( J_k := \{j_1, \ldots, j_{k-1}\} \). We have

\[
f(k) \overset{(i)}{=} \mathbb{E}_{(j_1, \ldots, j_k) \sim \mu^\land k} \max_{M_{j_1}, \ldots, M_{j_{k-1}}} \left( I(A : B_{j_1}) + \ldots + \max_{M_{j_k}} I(A : B_{j_k} | B_{j_1}, \ldots, B_{j_{k-1}}) \right)
\]

\[
\overset{(ii)}{=} \mathbb{E}_{(j_1, \ldots, j_{k-1}) \sim \mu^\land k-1} \max_{M_{j_1}, \ldots, M_{j_{k-1}}} \mathbb{E}_{j_k \notin J_k} \left( I(A : B_{j_1}) + \ldots + \max_{M_{j_k}} I(A : B_{j_k} | B_{j_1}, \ldots, B_{j_{k-1}}) \right)
\]

\[
\overset{(iii)}{=} \mathbb{E}_{(j_1, \ldots, j_{k-1}) \sim \mu^\land k-1} \max_{M_{j_1}, \ldots, M_{j_{k-1}}} \left( I(A : B_{j_1}) + \ldots + \mathbb{E}_{j_k \notin J_k} \max_{M_{j_k}} I(A : B_{j_k} | B_{j_1}, \ldots, B_{j_{k-1}}) \right)
\]

\[
\overset{(iv)}{=} \mathbb{E}_{(j_1, \ldots, j_{k-1}) \sim \mu^\land k-1} \max_{M_{j_1}, \ldots, M_{j_{k-1}}} \left( I(A : B_{j_1}) + \ldots + \mathbb{E}_{j_k \notin J_k} \max_{M_{j_k}} I(A : B_{j_k} | B_{j_1}, \ldots, B_{j_{k-1}}) \right)
\]

\[
+ \mathbb{E}_{(j_1, \ldots, j_{k-1}) \sim \mu^\land k-1} \min_{M_{j_1}, \ldots, M_{j_{k-1}}} \mathbb{E}_{j_k \notin J_k} \max_{M_{j_k}} I(A : B_{j_k} | B_{j_1}, \ldots, B_{j_{k-1}}),
\]

\[
\overset{(v)}{=} f(k-1) + \mathbb{E}_{j_1, \ldots, j_{k-1}} \max_{M_{j_1}, \ldots, M_{j_{k-1}}} \mathbb{E}_{j_k \notin J_k} \max_{M_{j_k}} I(A : B_{j_k} | B_{j_1}, \ldots, B_{j_{k-1}}),
\]

where (i) follows since only \( I(A : B_{j_k} | B_{j_1}, \ldots, B_{j_{k-1}}) \pi \) depends on \( M_{j_k} \); (ii) by convexity of the maximum function; (iii) again because all the other terms in the sum are independent of \( j_k \); (iv) directly by inspection and linearity of expectation; and (v) by the definition of \( f(k) \) in Eq. (10).

From Eqs. (10) and (11), we obtain

\[
\log d_A \geq \sum_{q=1}^{k} \mathbb{E}_{(j_1, \ldots, j_{q-1}) \sim \mu^\land q-1} \min_{M_{j_1}, \ldots, M_{j_{q-1}}} \mathbb{E}_{q \notin J_q} \max_{M_{j_q}} I(A : B_{j_q} | B_{j_1}, \ldots, B_{j_{q-1}}),
\]

where
and so there exists a $q \leq k$ such that
\[
\mathbb{E} \min_{(j_1, \ldots, j_{q-1}) \sim \mu^{q-1}} \mathbb{E} \max_{M_{j_1}, \ldots, M_{j_{q-1}}} I(A : B_j | B_{j_1}, \ldots, B_{j_{q-1}}) \pi \leq \frac{\log d_A}{k},
\]
where we relabelled $j_q \rightarrow j$. Thus there exists a $(q-1)$-tuple $J := (j_1, \ldots, j_{q-1})$ and measurements $M_{j_1}, \ldots, M_{j_{q-1}}$ such that
\[
\mathbb{E} \max_{j \notin J} I(A : B_j | B_{j_1}, \ldots, B_{j_{q-1}}) \pi \leq \frac{\log d_A}{k}.
\]
Let $\rho_{ABj}^*$ be the post-measurement state on $AB_j$ conditioned on obtaining $z$ – a short-hand notation for the ordered collection of the local results – when measuring $M_{j_1}, \ldots, M_{j_{q-1}}$ in the subsystems $B_{j_1}, \ldots, B_{j_{q-1}}$ of $\rho$. Note that $\rho_A^*$ is independent of $B_j$ (for $j \notin J$). By Pinsker’s inequality (55) in Supplementary Methods, convexity of $x \mapsto x^2$, and Eq. (56) in Supplementary Methods,
\[
\left\| \text{id}_A \otimes M_j \left( \rho_{ABj} - \mathbb{E}_z \rho_A^* \otimes \rho_{Bj}^* \right) \right\|_1^2 = \left\| \text{id}_A \otimes M_j \left( \mathbb{E}_z \rho_{ABj}^* - \mathbb{E}_z \rho_A^* \otimes \rho_{Bj}^* \right) \right\|_1^2 \\
\leq \mathbb{E}_z \left\| \text{id}_A \otimes M_j \left( \rho_{ABj}^* - \rho_A^* \otimes \rho_{Bj}^* \right) \right\|_1^2 \\
\leq 2 \ln(2) I(A : B_j | B_{j_1}, \ldots, B_{j_{q-1}}) \pi.
\]
By Eq. (14) and convexity of $x \mapsto x^2$,
\[
\mathbb{E} \max_{j \notin J} \left\| \text{id}_A \otimes M_j \left( \rho_{ABj} - \mathbb{E}_z \rho_A^* \otimes \rho_{Bj}^* \right) \right\|_1 \leq \sqrt{2 \ln(2) \frac{\log d_A}{k}}.
\]
Now, by Lemma 1, we have.
\[
\left\| \rho_{ABj} - \mathbb{E}_z \rho_A^* \otimes \rho_{Bj}^* \right\|_1 \leq (d_A)^2 \max_{M_j} \left\| \text{id}_A \otimes M_j \left( \rho_{ABj} - \mathbb{E}_z \rho_A^* \otimes \rho_{Bj}^* \right) \right\|_1,
\]
and so
\[
\mathbb{E} \left\| \rho_{ABj} - \mathbb{E}_z \rho_A^* \otimes \rho_{Bj}^* \right\|_1 \leq \sqrt{2 \ln(2) \frac{(d_A)^4 \log d_A}{k}}.
\]
Note that $\mathbb{E}_z \rho_A^* \otimes \rho_{Bj}^* = \sum_z p(z) \rho_A^* \otimes \rho_{Bj}^*$ is the Choi-Jamiołkowski state of a measure-and-prepare channel $\mathcal{E}_j$ [3], since $\mathbb{E}_z \rho_A^* = \rho_A = 1/d_A$. It is explicitly given by
\[
\mathcal{E}_j(X) := d_A \mathbb{E}_z \text{tr}(\rho_A^* X) \rho_{Bj}^*.
\]
Note that the POVM $\{d_A p(z) \rho_A^*\}$ is independent of $j$.

Thanks to Lemma 2, we can now bound the distance of two maps by the distance of their Choi-Jamiołkowski states
\[
\left\| \text{tr}_{B_j} \circ \Lambda - \mathcal{E}_j \right\|_\diamond \leq d_A \left\| \rho_{ABj} - \mathbb{E}_z \rho_A^* \otimes \rho_{Bj}^* \right\|_1,
\]
to find
\[
\mathbb{E} \left\| \text{tr}_{B_j} \circ \Lambda - \mathcal{E}_j \right\|_\diamond \leq \sqrt{2 \ln(2) \frac{(d_A)^6 \log d_A}{k}}.
\]
Then
\[ \mathbb{E} \left\| \text{tr}_{B_j} \circ \Lambda - \mathcal{E}_j \right\|_\diamond \leq \mathbb{E} \left\| \text{tr}_{B_j} \circ \Lambda - \mathcal{E}_j \right\|_\diamond + \frac{k}{n} \mathbb{E} \left\| \text{tr}_{B_j} \circ \Lambda - \mathcal{E}_j \right\|_\diamond \leq \sqrt{2 \ln(2) \frac{(d_A)^6 \log d_A}{k}} + \frac{2k}{n}, \] (22)

where we used that the diamond norm between two cptp maps is upper-bounded by 2. Choosing \( k \) to minimize the latter bound we obtain \(^1\)

\[ \mathbb{E} \left\| \text{tr}_{B_j} \circ \Lambda - \mathcal{E}_j \right\|_\diamond \leq \left( \frac{27 \ln(2)(d_A)^6 \log(d_A)}{n} \right)^{1/3}. \] (23)

Finally applying Markov’s inequality,

\[ \Pr \left( \left\| \text{tr}_{B_j} \circ \Lambda - \mathcal{E}_j \right\|_\diamond \geq \frac{1}{\delta} \left( \frac{27 \ln(2)(d_A)^6 \log(d_A)}{n} \right)^{1/3} \right) \leq \delta. \] (24)

B. Supplementary Note 2 - Proof of Theorem 2

The proof of Theorem 2 follows along the same lines as Theorem 1:

**Theorem 2 (restatement).** Let \( \Lambda : \mathcal{D}(A) \to \mathcal{D}(B_1 \otimes \ldots \otimes B_n) \) be a cptp map. For any subset \( S_t \subseteq [n] \) of \( t \) elements, define \( \Lambda_{S_t} := \text{tr}_{\bigcup \in S_t B_i} \circ \Lambda \) as the effective channel from \( \mathcal{D}(A) \) to \( \mathcal{D}(\bigotimes_{l \in S_t} B_l) \). Then for every \( 1 > \delta > 0 \) there exists a measurement \( \{M_k\}_k \) \( (M_k \geq 0, \sum_k M_k = I) \) such that for more than a \( (1 - \delta) \) fraction of the subsets \( S_t \subseteq [n], \)

\[ \left\| \Lambda_{S_t} - \mathcal{E}_{S_t} \right\|_\diamond \leq \left( \frac{27 \ln(2)(d_A)^6 \log(d_A) t}{n \delta^3} \right)^{1/3}, \] (25)

with

\[ \mathcal{E}_{S_t}(X) := \sum_k \text{tr}(M_k X) \sigma_{S_t,k}, \] (26)

for states \( \sigma_{S_t,k} \in \mathcal{D}(\bigotimes_{l \in S_t} B_l) \).

**Proof.** Since the proof is very similar to the proof of Theorem 1, we will only point out the differences.

Let \( \rho_{AB_1,\ldots,B_n} := \text{id}_A \otimes \Lambda(\Phi) \) be the Choi-Jamiolkowski state of \( \Lambda \) and \( C = \{C_1, \ldots, C_{n/t}\} \) be a partition of \( [n] \) into \( n/t \) sets of \( t \) elements each. Define \( \pi_C := \text{id}_A \otimes M_1 \otimes \ldots \otimes M_{n/t}(\rho) \), for quantum-classical channels \( M_1, \ldots, M_{n/t} \) defined as \( M_i(X) := \sum_l \text{tr}(N_{i,l} X) |l\rangle \langle l| \), for a POVM \( \{N_{i,l}\}_l \), with \( M_i \) acting on \( \bigcup_{j \in C_i} B_j \).

As in the proof of Theorem 1, by the chain rule,

\[ \log d_A \geq \sum_{C_1, \ldots, C_{n/t}} \max_{M_1, \ldots, M_{n/t}} \mathbb{E} \left( I(A : B_{C_1}, \ldots, B_{C_{n/t}})_{\pi_C} \right) \]

\[ = \sum_{C_1, \ldots, C_{n/t}} \max_{M_1, \ldots, M_{n/t}} \left( I(A : B_{C_1})_{\pi_C} + \ldots + I(A : B_{C_{n/t}} | B_{C_1}, \ldots, B_{C_{n/t-1}})_{\pi_C} \right) =: f(t), \] (27)

\(^1\) The expression \( a/\sqrt{k} + bk \) is minimal for \( k = (\frac{27}{20})^{2/3} \). We further use that for \( b = 2/n < 1 \) it holds \( b^{1/3} \geq b^{5/6} \).
where the expectation is taken uniformly over the choice of non-overlapping sets $C_{j_1}, \ldots, C_{j_k} \in [n]^t$.

We have

\[
 f(t) = \mathbb{E} \max_{C_{j_1}, \ldots, C_{j_k}} \left( I(A : B_{C_{j_1}} | B_{C_{j_2}, \ldots, B_{C_{j_k}}}) \right)
\]

\[
 \geq \mathbb{E} \max_{C_{j_1}, \ldots, C_{j_k}} \left( I(A : B_{C_{j_1}} | B_{C_{j_2}, \ldots, B_{C_{j_k}}}) \right)
\]

\[
 = \mathbb{E} \max_{C_{j_1}, \ldots, C_{j_k}} \left( I(A : B_{C_{j_1}} | B_{C_{j_2}, \ldots, B_{C_{j_k}}}) \right)
\]

\[
 \geq \mathbb{E} \max_{C_{j_1}, \ldots, C_{j_k}} \left( I(A : B_{C_{j_1}} | B_{C_{j_2}, \ldots, B_{C_{j_k}}}) \right)
\]

\[
 = f(t - 1) + \mathbb{E} \min_{C_{j_1}, \ldots, C_{j_k}} \mathbb{E} \max_{C_{j_1}, \ldots, C_{j_k}} \left( I(A : B_{C_{j_1}} | B_{C_{j_2}, \ldots, B_{C_{j_k}}}) \right)
\]

From Eqs. (27) and (28), we obtain

\[
 \log d_A \geq \sum_{q=1}^{k} \mathbb{E} \min_{C_{j_1}, \ldots, C_{j_q-1}} \mathbb{E} \max_{C_{j_q}} \left( I(A : B_{C_{j_q}} | B_{C_{j_1}, \ldots, B_{C_{j_q-1}}}) \right) \pi, \tag{29}
\]

and so there exists a $q \leq k$ such that

\[
 \mathbb{E} \min_{C_{j_1}, \ldots, C_{j_q-1}} \mathbb{E} \max_{C_{j_q}} \left( I(A : B_{C_{j_q}} | B_{C_{j_1}, \ldots, B_{C_{j_q-1}}}) \right) \leq \frac{\log d_A}{t}, \tag{30}
\]

where we relabelled $j_q \rightarrow j$. Thus there exists a $(q - 1)$-tuple of sets $C := \{C_{j_1}, \ldots, C_{j_q-1}\}$ and measurements $M_{j_1}, \ldots, M_{j_q-1}$ such that

\[
 \mathbb{E} \max_{C_j \notin C} \max_{M_j} \left( I(A : B_{C_j} | B_{C_{j_1}, \ldots, B_{C_{j_q-1}}}) \right) \pi \leq \frac{\log d_A}{t}. \tag{31}
\]

Here we can follow the proof of Theorem 1 without any modifications to obtain that

\[
 \mathbb{E} \left\| \text{tr}_{\Lambda} \circ \Lambda - \mathcal{E}_{C_j} \right\|_\infty \leq \sqrt{2 \ln(2) \left( \frac{d_A}{t} \right)^6 \log d_A}, \tag{32}
\]

Then

\[
 \mathbb{E} \left\| \text{tr}_{\Lambda} \circ \Lambda - \mathcal{E}_{C_j} \right\|_\infty \leq \sqrt{2 \ln(2) \left( \frac{d_A}{t} \right)^6 \log d_A} + \frac{2kt}{n}. \tag{33}
\]

Choosing $k$ to minimize the right-hand side as done in the proof of Theorem 1 and applying Markov’s inequality, we obtain the result. □

C. Supplementary Note 3 - Proof of Proposition 3

We will make use the following well-known lemma:
Lemma 3. (Gentle Measurement [4]) Let \( \rho \) be a density matrix and \( N \) an operator such that \( 0 \leq N \leq \mathbb{I} \) and \( \text{tr}(N\rho) \geq 1 - \delta \). Then
\[
\| \rho - \sqrt{N}\rho\sqrt{N} \|_1 \leq 2\sqrt{\delta}. 
\] (34)

Proposition 3 (restatement). Let \( \mathcal{E} \) be the channel given by Eq. (26). Suppose that for every \( i = \{1, \ldots, t\} \) and \( 1 > \delta > 0 \),
\[
\min_{\rho \in D(A)} p_{\text{guess}}(\{\text{tr}(M_k \rho), \sigma_{B_{j_1}k}\}) \geq 1 - \delta. 
\] (35)
Then there exists POVMs \( \{N_{B_{j_1}k}\}, \ldots, \{N_{B_{j_t}k}\} \) such that
\[
\min_{\rho} \sum_k \text{tr}(M_k \rho) \text{tr} \left( \bigotimes_i N_{B_{j_i}k} \right) \sigma_{B_{j_1} \cdots B_{j_t}k} \geq 1 - 6t\delta^{1/4}. 
\] (36)

Proof. For simplicity we will prove the claim for \( t = 2 \). The general case follows by a similar argument.

Since for \( j = \{1, 2\} \), \( \min_{\rho \in D(A)} p_{\text{guess}}(\{\text{tr}(M_k \rho), \sigma_{B_{j}k}\}) \geq 1 - \delta \), by the minimax theorem [5] it follows that there exists POVMs \( \{N_{B_{1}k}\}, \{N_{B_{2}k}\} \) on \( B_{1} \) and \( B_{2} \), respectively, such that for \( j \in \{1, 2\} \) and all \( \rho \in D(A) \),
\[
\sum_k \text{tr}(M_k \rho) \text{tr}(N_{B_{j}k} \sigma_{B_{j}k}) \geq 1 - \delta. 
\] (37)

Fix \( \rho \) and let \( X_j := \{k : \text{tr}(N_{B_{j}k} \sigma_{B_{j}k}) \leq 1 - \sqrt{\delta} \} \) for \( j = \{1, 2\} \). Then from Eq. (37),
\[
\sum_{k \in X_j} \text{tr}(\rho M_k) \leq \sqrt{\delta}. 
\] (38)

Let \( G = X_1^c \cap X_2^c \), with \( X_j^c \) the complement of \( X_j \). Then
\[
\sum_k \text{tr}(M_k \rho) \text{tr} \left( (N_{B_{1}k} \otimes N_{B_{2}k}) \sigma_{B_{1}B_{2}k} \right) 
\geq \sum_{k \in G} \text{tr}(M_k \rho) \text{tr} \left( (N_{B_{1}k} \otimes N_{B_{2}k}) \sigma_{B_{1}B_{2}k} \right) 
\geq \sum_{k \in G} \text{tr}(M_k \rho) \text{tr}(N_{B_{1}k} \sigma_{B_{1}k}) \text{tr}(N_{B_{2}k} \sigma_{B_{2}k}) - 4\delta^{1/4} 
\geq (1 - \sqrt{\delta}) \sum_{k \in G} \text{tr}(M_k \rho) \text{tr}(N_{B_{1}k} \sigma_{B_{1}k}) - 4\delta^{1/4} 
\geq (1 - \sqrt{\delta})(1 - \delta - 2\sqrt{\delta}) - 4\delta^{1/4} 
\geq 1 - 12\delta^{1/4}, 
\] (39)

where in the third line we used Lemma 3. In more detail, we have
\[
\text{tr} \left( (N_{B_{1}k} \otimes N_{B_{2}k}) \sigma_{B_{1}B_{2}k} \right) = \text{tr}(N_{B_{1}k} \sigma_{B_{1}k}) \text{tr}(N_{B_{2}k} \sigma'_{B_{2}k}), 
\] (40)
with \( \sigma'_{B_{2}k} := \text{tr}_{B_{1}}(N_{B_{1}k} \sigma_{B_{1}B_{2}k})/\text{tr}(N_{B_{1}k} \sigma_{B_{1}k}) \). Since \( \text{tr}(N_{B_{1}k} \sigma_{B_{1}k}) \geq 1 - \delta^{1/2} \), Lemma 3 gives \( \| \sigma'_{B_{2}k} - \sigma_{B_{2}k} \|_1 \leq 4\delta^{1/4} \). Then from Eq. (40),
\[
\text{tr} \left( (N_{B_{1}k} \otimes N_{B_{2}k}) \sigma_{B_{1}B_{2}k} \right) \geq \text{tr}(N_{B_{1}k} \sigma_{B_{1}k}) \text{tr}(N_{B_{2}k} \sigma_{B_{2}k}) - 4\delta^{1/4}. 
\] (41)
D. Supplementary Note 4 - Proof of Corollary 4

Corollary 4 will follow from Theorem 1 and the following well-known continuity relation for mutual information:

**Lemma 4.** (Alicki-Fannes Inequality [6]) For $\rho_{AB}$,

$$|H(A|B)_{\rho} - H(A|B)_{\sigma}| \leq 4\|\rho - \sigma\|_1 \log d_A + 2h_2(\|\rho - \sigma\|_1),$$  \hspace{1cm} (42)

with $H(A|B) = S(AB) - S(B)$ and $h_2$ the binary entropy function.

If $S(A)_\rho = S(A)_\sigma$, then

$$|I(A : B)_{\rho} - I(A : B)_{\sigma}| \leq 4\|\rho - \sigma\|_1 \log d_A + 2h_2(\|\rho - \sigma\|_1).$$  \hspace{1cm} (43)

**Corollary 4 (restatement).** Let $\Lambda: \mathcal{D}(B) \to \mathcal{D}(B_1 \otimes \ldots \otimes B_n)$ be a ctp map. Define $\Lambda_j := \text{tr}_{B_j} \circ \Lambda$ as the effective dynamics from $\mathcal{D}(B)$ to $\mathcal{D}(B_j)$. Then for every $1 > \delta > 0$ there exists a set $\mathcal{S} \subseteq [n]$ with $|\mathcal{S}| \geq (1 - \delta)n$ such that for all $j \in \mathcal{S}$ and all states $\rho_{AB}$ it holds

$$I(A : B_j)_{\text{id}_A \otimes \Lambda_j B(\rho_{AB})} \leq \max_{\Gamma\in QC_{\mathcal{S}}} I(A : B)_{\text{id}_A \otimes \Gamma B(\rho_{AB})} + 4\epsilon \log d_A + 2h_2(\epsilon),$$  \hspace{1cm} (44)

where $\epsilon = \left( \frac{27 \ln(2)(d_B)^6 \log(d_B)}{n^6} \right)^{1/3}$, $h_2$ is the binary entropy function, $h_2(x) = -x \log x - (1 - x) \log(1 - x)$, and the maximum on the right-hand side is over quantum-classical channels $\Gamma_{\mathcal{S}}(X) = \sum_i \text{tr}(N_i X) |i\rangle \langle i|$, with $\{N_i\}$ a POVM and $\{|i\rangle\}$ a set of orthogonal states. As a consequence, for every $\rho_{AB},$

$$\lim_{n \to \infty} \left( \max_{\Lambda_{B \to B_1 B_2 \ldots B_n}} \mathbb{E} I(A : B_j)_{\text{id}_A \otimes \Lambda B(\rho_{AB})} \right) = \max_{\Gamma\in QC_{\mathcal{S}}} I(A : B)_{\text{id}_A \otimes \Gamma B(\rho_{AB})},$$  \hspace{1cm} (45)

with $\mathbb{E}_j X_j = \frac{1}{n^2} \sum_{i=1}^N X_j$, and the maximum on the left-hand side taken over any quantum operation $\Lambda: \mathcal{D}(B) \to \mathcal{D}(B_1 \otimes \ldots \otimes B_n)$.

**Proof.** By definition, for all ctp maps $\Lambda$ and $\mathcal{E}$ acting on $B$, and for any state $\rho_{AB}$, it holds

$$\|\text{id}_A \otimes \Lambda B(\rho) - \text{id}_A \otimes \mathcal{E} B(\rho)\|_1 \leq \|\Lambda - \mathcal{E}\|_\infty.$$  

Combining Theorem 1 and Lemma 4 (specifically, Eq. (43)), we have that for every $1 > \delta > 0$ there exist a measurement $\{M_k\}$ and a set $\mathcal{S} \subseteq [n]$ with $|\mathcal{S}| \geq (1 - \delta)n$ such that for all $j \in \mathcal{S}$ and all states $\rho_{A' A}$ it holds

$$I(A : B)_{\text{id}_A \otimes \Lambda_j (\rho_{AB})} \leq I(A : B)_{\text{id}_A \otimes \mathcal{E}_j (\rho_{AB})} + 4\epsilon \log d_A + 2h_2(\epsilon),$$  \hspace{1cm} (46)

with

$$\mathcal{E}_j(X) = \sum_k \text{tr}(M_k X) |k\rangle \langle k|$$  \hspace{1cm} (47)

and

$$\epsilon = \left( \frac{27 \ln(2)(d_B)^6 \log(d_B)}{n^6} \right)^{1/3}.$$  \hspace{1cm} (48)

The claim is then a simple consequence of substituting $\mathcal{E}_j$ with an optimal quantum-classical channel.
We now turn to the proof of Eq. (45). That the left-hand side of Eq. (45) is larger than the right-hand side is trivial. Indeed one can pick \( \Lambda = \Lambda_{B \rightarrow B_1 B_2 \ldots B_n} \) as the quantum-classical map that uses the POVM \( \{N_l\} \) that achieves the accessible information \( I(A : B_c) := \max_{\Gamma \in QC} I(A : B)_{\text{id} \otimes \Gamma(\rho_{AB})} \) with measurement on \( B \) and stores the result in \( n \) classical registers, one for each \( B_i \); \( \Gamma(X) = \sum_l \text{tr}(N_l X) |l\rangle \langle l|^n \). To prove that the left-hand side of Eq. (45) is smaller than the right-hand side it is sufficient to use Eq. (46) for the choice \( \delta = n^{-\frac{1}{3}} - \eta^3 \), for any \( 0 < \eta < 1 \). Then one obtains,

\[
\frac{1}{n} \sum_{i=1}^n I(A : B_i) \leq \frac{1}{n} \{ (1 - \delta) n [I(A : B_c) + 4 \epsilon \log d_A + 2 h_2 (\epsilon)] + \delta n 2 \log d_A \} \\
= (1 - \delta) [I(A : B_c) + 4 \epsilon \log d_A + 2 h_2 (\epsilon)] + \delta 2 \log d_A \xrightarrow{n \to \infty} I(A : B_c)
\]

where we have used that

\[
\epsilon = \left( \frac{27 \ln(2)(d_B)^6 \log(d_B)}{n \delta^3} \right)^{1/3} \xrightarrow{n \to \infty} 0
\]

for our choice of \( \delta \), independently of the choice of \( \Lambda = \Lambda_{B \rightarrow B_1 B_2 \ldots B_n} \).

\[\square\]

### II. Supplementary Methods

We make use of the following properties of the mutual information:

- **Positivity of conditional mutual information:**
  \[
  I(A : B|C) := I(A : BC) - I(A : C) \geq 0.
  \]
  This is equivalent to strong subadditivity and to monotonicity of mutual information under local operations [7].

- **For a general state \( \rho_{AB} \) it holds [7]**
  \[
  I(A : B)_{\rho_{AB}} \leq 2 \min \{ \log d_A, \log d_B \},
  \]
  with the more stringent bound
  \[
  I(A : B)_{\rho_{AB}}^{\text{sep}} \leq \min \{ \log d_A, \log d_B \}
  \]
  for a separable state \( \rho_{AB}^{\text{sep}} \) [8].

- **Chain rule [7]:**
  \[
  I(A : B_1 B_2 \ldots B_n) = I(A : B_1) + I(A : B_2|B_1) + I(A : B_3|B_1 B_2) + \ldots \\
  \ldots + I(A : B_n|B_1 B_2 \ldots B_{n-1}).
  \]

- **Pinsker’s inequality (for mutual information):**
  \[
  \frac{1}{2 \ln 2} \|\rho_{AB} - \rho_A \otimes \rho_B\|_1^2 \leq I(A : B)_{\rho_{AB}}.
  \]

- **Conditioning on classical information**
  \[
  I(A : B|Z)_{\rho} = \sum_z p(z) I(A : B)_{\rho_z}
  \]
  for a state \( \rho_{ABZ} = \sum_z p(z) \rho_{z,AB} \otimes |z\rangle \langle z|_Z \), with \( \{|z\} \) an orthonormal set.
III. SUPPLEMENTARY REFERENCES