\textbf{I. SUPPLEMENTARY NOTES}

\section*{A. Supplementary Note 1 - Proof of Theorem 1}

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

\textbf{Lemma 1.} Consider a Hermitian matrix $L_{AB} \in \mathcal{B}(\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B})$, 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 $\mathcal{M}_B(Y) = \sum_l \text{tr}(N_l Y)\langle l | l \rangle$, with a POVM $\{N_l\}$ and orthonormal states $\{|l\rangle\}$.

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

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

On the other hand,

$$\max_{\mathcal{M}_B} \|\text{id}_A \otimes \mathcal{M}_B (L_{AB})\|_1$$

$$= \max_{\mathcal{M}_B} \left\| \sum_{i,j=1}^{d_A} |i\rangle \langle j | \otimes \mathcal{M}_B(L_{ij}) \right\|_1$$

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

$$\geq \max_{\mathcal{M}_B} \max_{K_A = K_A^\dagger, \|K_A\| \leq 1} \max_{\|K_B\| \leq 1} \left\| \text{tr} \left( K_A \otimes K_B \left( \sum_{i,j=1}^{d_A} |i\rangle \langle j | \otimes \mathcal{M}_B(L_{ij}) \right) \right) \right\|_1$$

$$\geq \max_i \max_{\mathcal{M}_B} \|M_B(L_{ii})\|_1, \max_{i \neq j} \max_{\mathcal{M}_B} \|M_B(L_{ij} + L_{ji})\|_1, \max_{i \neq j} \max_{\mathcal{M}_B} \|M_B(i(L_{ij} - L_{ji}))\|_1 \tag{2},$$

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

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

and similarly

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

To complete the proof it is enough to observe

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

$\square$
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 cptp map $\Lambda : D(A) \to D(B)$ we define the Choi-Jamiołkowski state of $\Lambda$ as $J(\Lambda) := \text{id}_A \otimes \Lambda_A(\Phi_{AA'})$. For two cptp maps $\Lambda_0$ and $\Lambda_1$ it then holds

$$\frac{1}{d_A} \|\Lambda_0 - \Lambda_1\|_\diamond \leq \|J(\Lambda_0) - J(\Lambda_1)\|_1 \leq \|\Lambda_0 - \Lambda_1\|_\diamond. \quad (6)$$

**Proof.** The second inequality in (6) is trivial, as the diamond norm between two cptp 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.,

$$|\psi\rangle_{AA'} = (\sqrt{d_A} C \otimes \text{id}) |\Phi\rangle_{AA'}$$

for a suitable $C \in B(C^{d_A})$, 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

$$\|\Lambda_0 - \Lambda_1\|_\diamond = \|\text{id}_A \otimes (\Lambda_0 - \Lambda_1)|\psi\rangle\langle\psi|\|_1$$

$$= \|\text{id}_A \otimes (\Lambda_0 - \Lambda_1) \left( (\sqrt{d_A} C \otimes \text{id}) |\Phi\rangle_{AA'} (\sqrt{d_A} C \otimes \text{id})^\dagger \right)\|_1$$

$$= \left\| \left(\sqrt{d_A} C \otimes \text{id}\right) (\text{id}_A \otimes (\Lambda_0 - \Lambda_1) |\Phi_{AA'}\rangle) (\sqrt{d_A} C \otimes \text{id})^\dagger \right\|_1$$

$$\leq d_A \|C\|_\infty^2 \|\text{id}_A \otimes (\Lambda_0 - \Lambda_1) |\Phi_{AA'}\rangle\|_1$$

$$\leq d_A \|J(\Lambda_0) - J(\Lambda_1)\|_1,$$

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. \hfill \Box

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

**Theorem 1 (restatement).** Let $\Lambda : D(A) \to D(B_1 \otimes \ldots \otimes B_n)$ be a cptp map. Define $\Lambda_j := \text{tr}_{B_j} \circ \Lambda$ as the effective dynamics from $D(A)$ to $D(B_j)$ and fix a number $1 > \delta > 0$. Then there exists a measurement $\{M_k\}_k$ ($M_k \geq 0$, $\sum_k M_k = I$) and a set $S \subseteq \{1, \ldots, n\}$ with $|S| \geq (1 - \delta)n$ such that for all $j \in S$,

$$\|\Lambda_j - \mathcal{E}_j\|_\diamond \leq \left( \frac{27 \ln(2) (d_A)^3 \log(d_A)}{n \delta^3} \right)^{1/3}, \quad (7)$$

with

$$\mathcal{E}_j(X) := \sum_k \text{tr}(M_k X) \sigma_{j,k}, \quad (8)$$

for states $\sigma_{j,k} \in D(B_j)$. 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-Jamiolkowski state of $\Lambda$ [2]. Define $\pi := \text{id}_A \otimes M_1 \otimes \ldots \otimes M_n(\rho)$, 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 $\{N_{i,l}\}$.

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 $\rho_{A|B_i}$ is close to a separable state $\sum_{\mu} \rho(\mu,A) \otimes \rho_{B_i|\mu}$ with the ensemble $\{\rho(\mu,A)\}$ independent of $i$. We will conclude showing that by the properties of the Choi-Jamiolkowski isomorphism, this implies that the effective channel from $A$ to $B_i$ is close to a measure-andPrepare channel with a POVM independent of $i$.

Let $\mu$ be the uniform distribution over $[n]$ and define $\mu^{\wedge k}$ as the distribution on $[n]^k$ obtained by sampling $m$ times without replacement according to $\mu$; i.e.

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

Then

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

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

$$
=: f(k).
$$

The inequality comes from the fact that $\pi$ is separable between $A$ and $B_1B_2\ldots B_n$ 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

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

$$
\geq \mathbb{E}_{(j_1,\ldots,j_{k-1}) \sim \mu^{\wedge k-1}} \max_{M_{j_1},\ldots,M_{j_{k-1}}} \mathbb{E}_{j_k \notin J_k} \left( I(A:B_{j_1})_{\pi} + \ldots + \max_{M_{j_{k}}} I(A:B_{j_k}|B_{j_1},\ldots,B_{j_{k-1}})_{\pi} \right) 
$$

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

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

$$
\overset{(iv)}{=} \mathbb{E}_{(j_1,\ldots,j_{k-1}) \sim \mu^{\wedge k-1}} \min_{M_{j_1},\ldots,M_{j_{k-1}}} \mathbb{E}_{j_k \notin J_k} \max_{J_k} I(A:B_{j_k}|B_{j_1},\ldots,B_{j_{k-1}})_{\pi}, 
$$

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

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

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

where $\mu^{\wedge q-1}$ is the distribution on $[n]^{q-1}$ obtained by sampling $(q-1)$ times without replacement according to $\mu$.
and so there exists a \( q \leq k \) such that

\[
\mathbb{E} \min_{(j_1, \ldots, j_q) \sim \mu^{\otimes q-1} \ M_{j_1}, \ldots, M_{j_{q-1}}} \max_{j \notin j_q} \mathbb{E} \max I(A : B_j | B_{j_1}, \ldots, B_{j_{q-1}}) \pi \leq \frac{\log d_A}{k},
\]

where we relabelled \( j_q \to j \). 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 \( \rho^x_{ABj} \) 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 \( \rho^x_A \) 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,

\[
\left\| \text{id}_A \otimes M_j \left( \rho_{ABj} - \mathbb{E}_z \rho^x_A \otimes \rho^z_{B_j} \right) \right\|_1^2 \leq \left\| \text{id}_A \otimes M_j \left( \mathbb{E}_z \rho^x_{ABj} - \mathbb{E}_z \rho^x_A \otimes \rho^z_{B_j} \right) \right\|_1^2 \leq 2 \ln(2) I(A : B_j | B_{j_1}, \ldots, B_{j_{q-1}}) \pi.
\]

By Eq. (14) and convexity of \( x \mapsto x^2 \),

\[
\mathbb{E} \max_{j \notin J} \left\| \text{id}_A \otimes M_j \left( \rho_{ABj} - \mathbb{E}_z \rho^x_A \otimes \rho^z_{B_j} \right) \right\|_1 \leq \sqrt{2 \ln(2) \frac{\log d_A}{k}}.
\]

Now, by Lemma 1, we have.

\[
\left\| \rho_{ABj} - \mathbb{E}_z \rho^x_A \otimes \rho^z_{B_j} \right\|_1 \leq (d_A)^2 \max_{M_j} \left\| \text{id}_A \otimes M_j \left( \rho_{ABj} - \mathbb{E}_z \rho^x_A \otimes \rho^z_{B_j} \right) \right\|_1,
\]

and so

\[
\mathbb{E} \left\| \rho_{ABj} - \mathbb{E}_z \rho^x_A \otimes \rho^z_{B_j} \right\|_1 \leq \sqrt{2 \ln(2) \frac{(d_A)^4 \log d_A}{k}}.
\]

Note that \( \mathbb{E}_z \rho^x_A \otimes \rho^z_{B_j} = \sum_z p(z) \rho^x_A \otimes \rho^z_{B_j} \) is the Choi-Jamiolkowski state of a measure-and-prepare channel \( \mathcal{E}_j [3] \), since \( \mathbb{E}_z \rho^x_A = \rho_A = \mathbb{1} / d_A \). It is explicitly given by

\[
\mathcal{E}_j(X) := d_A \mathbb{E}_z (\rho^x_A)^T X \rho^z_{B_j}.
\]

Note that the POVM \( \{d_A p(z) \rho^x_A\} \) is independent of \( j \).

Thanks to Lemma 2, we can now bound the distance of two maps by the distance of their Choi-Jamiolkowski states

\[
\| \text{tr}_{B_j} \circ \Lambda - \mathcal{E}_j \|_\diamond \leq d_A \| \rho_{ABj} - \mathbb{E}_z \rho^x_A \otimes \rho^z_{B_j} \|_1,
\]

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
\[
\mathbb{E}_j \left\| \text{tr}_{B_j} \circ \Lambda - \mathcal{E}_j \right\|_\infty = \mathbb{E}_{j \notin J} \left\| \text{tr}_{B_j} \circ \Lambda - \mathcal{E}_j \right\|_\infty + \frac{k}{n} \mathbb{E}_{j \in J} \left\| \text{tr}_{B_j} \circ \Lambda - \mathcal{E}_j \right\|_\infty
\leq \sqrt{2 \ln(2) \frac{(d_A)^6 \log(d_A)}{2k} + \frac{2k}{n}},
\]
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 \(^1\)
\[
\mathbb{E}_j \left\| \text{tr}_{B_j} \circ \Lambda - \mathcal{E}_j \right\|_\infty \leq \left( \frac{27 \ln(2)(d_A)^6 \log(d_A)}{n} \right)^{1/3}.
\]

Finally applying Markov’s inequality,
\[
\Pr \left( \left\| \text{tr}_{B_j} \circ \Lambda - \mathcal{E}_j \right\|_\infty \geq \frac{1}{\delta} \left( \frac{27 \ln(2)(d_A)^6 \log(d_A)}{n} \right)^{1/3} \right) \leq \delta.
\]

\[\Box\]

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 \(\Lambda_{S_t} := \text{tr}_{\bigcup_{S_t \in [n]} B_t} \circ \Lambda\) as the effective channel from \(\mathcal{D}(A)\) to \(\mathcal{D}(\bigotimes_{t \in S_t} B_t)\). 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 \(S_t \subseteq [n]\),
\[
\|\Lambda_{S_t} - \mathcal{E}_{S_t}\|_\infty \leq \left( \frac{27 \ln(2)(d_A)^6 \log(d_A)}{nt^3} \right)^{1/3},
\]
\[
\mathcal{E}_{S_t}(X) := \sum_k \text{tr}(M_k X) \sigma_{S_t,k},
\]
for states \(\sigma_{S_t,k} \in \mathcal{D}(\bigotimes_{t \in S_t} B_t)\).

**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 \(M_i(X) := \sum_l \text{tr}(N_{i,l} X) \langle l | l \rangle\), 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,
\[
\log d_A \geq \max_{C_{j_1},\ldots,C_{j_k}} \mathbb{E}_{M_{j_1},\ldots,M_{j_k}} [I(A : B_{C_{j_1}}, \ldots, B_{C_{j_k}}) \pi_C]
\]
\[
= \max_{C_{j_1},\ldots,C_{j_k}} \mathbb{E}_{M_{j_1},\ldots,M_{j_k}} [I(A : B_{C_{j_1}} \pi_C + \ldots + I(A : B_{C_{j_k}} | B_{C_{j_1}}, \ldots, B_{C_{j_{k-1}}} \pi_C)) =: f(t),
\]

\(^1\) The expression \(a/\sqrt{k} + bk\) is minimal for \(k = (\frac{a}{2b})^{2/3}\). We further use that for \(b = 2/n < 1\) it holds \(b^{5/6} \geq b^{5/6}\).
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

\begin{align*}
\log d_A \geq \sum_{q=1}^k & \min_{C_{j_1}, \ldots, C_{j_q-1}, M_{j_1}, \ldots, M_{j_q-1}} \max_{C_j, M_j} \mathbb{E} \max I(A : B_{C_{j_1}}, \ldots, B_{C_{j_q-1}} | B_{C_{j_1}}, \ldots, B_{C_{j_q-1}}) \pi \\
& \leq \frac{\log d_A}{t}, \tag{30}
\end{align*}

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

\begin{align*}
\mathbb{E} & \max_{C_j \not\in C} \max I(A : B_{C_j} | B_{C_{j_1}}, \ldots, B_{C_{j_q-1}}) \pi \leq \frac{\log d_A}{t} \tag{31}
\end{align*}

where we relabelled $j_q \to j$. 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

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

\begin{align*}
\mathbb{E} & \left\| \text{tr}_{C_j} \circ \Lambda - \mathcal{E}_{C_j} \right\|_\diamond \leq \sqrt{2 \ln(2) \left( \frac{d_A}{k} \right)^6 \log d_A \frac{d_A}{k}}, \tag{32}
\end{align*}

Then

\begin{align*}
\mathbb{E} & \left\| \text{tr}_{C_j} \circ \Lambda - \mathcal{E}_{C_j} \right\|_\diamond \leq \sqrt{2 \ln(2) \left( \frac{d_A}{k} \right)^6 \log d_A \frac{d_A}{k}} + \frac{2kt}{n}. \tag{33}
\end{align*}

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. \hfill \Box

\section*{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 $0 \leq N \leq I$ and $\text{tr}(N \rho) \geq 1 - \delta$. Then

$$\|\rho - \sqrt{N} \rho \sqrt{N}\|_1 \leq 2\sqrt{\delta}. \quad (34)$$

Proposition 3 (restatement). Let $E$ be the channel given by Eq. (26). Suppose that for every $i = \{1, \ldots, t\}$ and $1 > \delta > 0$,

$$\min_{\rho \in \mathcal{D}(A)} p_{\text{guess}}(\{\text{tr}(M \rho), \sigma_{B_i,k}\}) \geq 1 - \delta. \quad (35)$$

Then there exists POVMs $\{N_{B_1,k}\}, \ldots, \{N_{B_t,k}\}$ such that

$$\min_{\rho} \sum_k \text{tr}(M \rho) \text{tr} \left( \bigotimes_i N_{B_i,k} \sigma_{B_i \ldots B_t,k} \right) \geq 1 - 6t\delta^{1/4}. \quad (36)$$

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

Since for $j = \{1, 2\}$, $\min_{\rho \in \mathcal{D}(A)} p_{\text{guess}}(\{\text{tr}(M \rho), \sigma_{B_i,k}\}) \geq 1 - \delta$, 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 \mathcal{D}(A),

$$\sum_k \text{tr}(M \rho) \text{tr}(N_{B_j,k} \sigma_{B_j,k}) \geq 1 - \delta. \quad (37)$$

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),

$$\sum_{k \in X_j} \text{tr}(\rho M_k) \leq \sqrt{\delta}. \quad (38)$$

Let $G = X_1^c \cap X_2^c$, with $X_j^c$ the complement of $X_j$. Then

$$\sum_k \text{tr}(M \rho) \text{tr} \left( (N_{B_1,k} \otimes N_{B_2,k}) \sigma_{B_1 B_2,k} \right) \geq \sum_{k \in G} \text{tr}(M \rho) \text{tr} \left( (N_{B_1,k} \otimes N_{B_2,k}) \sigma_{B_1 B_2,k} \right)$$

$$\geq \sum_{k \in G} \text{tr}(M \rho) \text{tr}(N_{B_1,k} \sigma_{B_1,k}) \text{tr}(N_{B_2,k} \sigma_{B_2,k}) - 4\delta^{1/4}$$

$$\geq (1 - \sqrt{\delta}) \sum_{k \in G} \text{tr}(M \rho) \text{tr}(N_{B_1,k} \sigma_{B_1,k}) - 4\delta^{1/4}$$

$$\geq (1 - \sqrt{\delta})(1 - \delta - 2\sqrt{\delta}) - 4\delta^{1/4}$$

$$\geq 1 - 12\delta^{1/4},$$

where in the third line we used Lemma 3. In more detail, we have

$$\text{tr} \left( (N_{B_1,k} \otimes N_{B_2,k}) \sigma_{B_1 B_2,k} \right) = \text{tr}(N_{B_1,k} \sigma_{B_1,k}) \text{tr}(N_{B_2,k} \sigma'_{B_2,k}), \quad (40)$$

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),

$$\text{tr} \left( (N_{B_1,k} \otimes N_{B_2,k}) \sigma_{B_1 B_2,k} \right) \geq \text{tr}(N_{B_1,k} \sigma_{B_1,k}) \text{tr}(N_{B_2,k} \sigma_{B_2,k}) - 4\delta^{1/4}. \quad (41)$$

$\square$
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 $\rho_{AB}$,

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

(42)

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

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

$$|I(\rho) - I(\sigma)| \leq 4\|\rho - \sigma\|_1 \log d_A + 2h_2(\|\rho - \sigma\|_1).$$

(43)

Corollary 4 (restatement). Let $\Lambda : \mathcal{D}(B) \rightarrow \mathcal{D}(B_1 \otimes \ldots \otimes B_n)$ be a cptp 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

$$I(\rho_{AB}) \leq \max_{\Gamma_{QC} \in \mathcal{QC}} I(\rho_{AB}),$$

(44)

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

$$\lim_{n \rightarrow \infty} \left( \max_{\Lambda : \mathcal{D}(B) \rightarrow \mathcal{D}(B_1 \otimes \ldots \otimes B_n)} \mathbb{E} I(\rho_{AB}) \right) = \max_{\Gamma_{QC} \in \mathcal{QC}} I(\rho_{AB}),$$

(45)

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

Proof. By definition, for all cptp 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 $\{M_k\}_k$ 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 $\rho_{A'B'}$ it holds

$$I(\rho_{A'B'}) \leq I(\rho_{A'B'}) + 4\epsilon \log d_A + 2h_2(\epsilon),$$

(46)

with

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

(47)

and

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

(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 \( I(A : B_c) := \max_{\Gamma \in \mathcal{QC}} I(A : B | \text{id} \otimes \Gamma(\rho_{AB})) \) with measurement on \( B \) and stores the result in \( n \) classical registers, one for each \( B_i \): \( \Gamma(X) = \sum_l \tr(N_l X)|l\rangle\langle l|^{\otimes n} \). 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 \( \delta = n^{-1} - \eta^3 \), for any \( 0 < \eta < 1 \). Then one obtains,

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

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 \( \Lambda = \Lambda_{B \rightarrow B_1 B_2 \ldots B_n} \).

\[\blacksquare\]

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
  \[ I(A : B)_{\sigma_{AB}^{sep}} \leq \min\{\log d_A, \log d_B\} \]
  for a separable state \( \sigma_{AB}^{sep} \) [8].

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

- **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