Supplementary Material for ‘Efficient methods for one-shot quantum communication’

Anurag Anshu [1] Rahul Jain[2]

[1] School of Engineering and Applied Sciences, Harvard University, Cambridge, MA, USA
anuraganshu@fas.harvard.edu
[2] Centre for Quantum Technologies, National University of Singapore; MajuLab, UMI 3654, Singapore and VAJRA Adjunct Faculty, TIFR, Mumbai, India. rahul@comp.nus.edu.sg

Here, we provide complete proofs for all the claims made in the main text. For the ease of navigation, we have discussed all the results and their interconnections in Supplementary Figure 1.

1 Supplementary notes

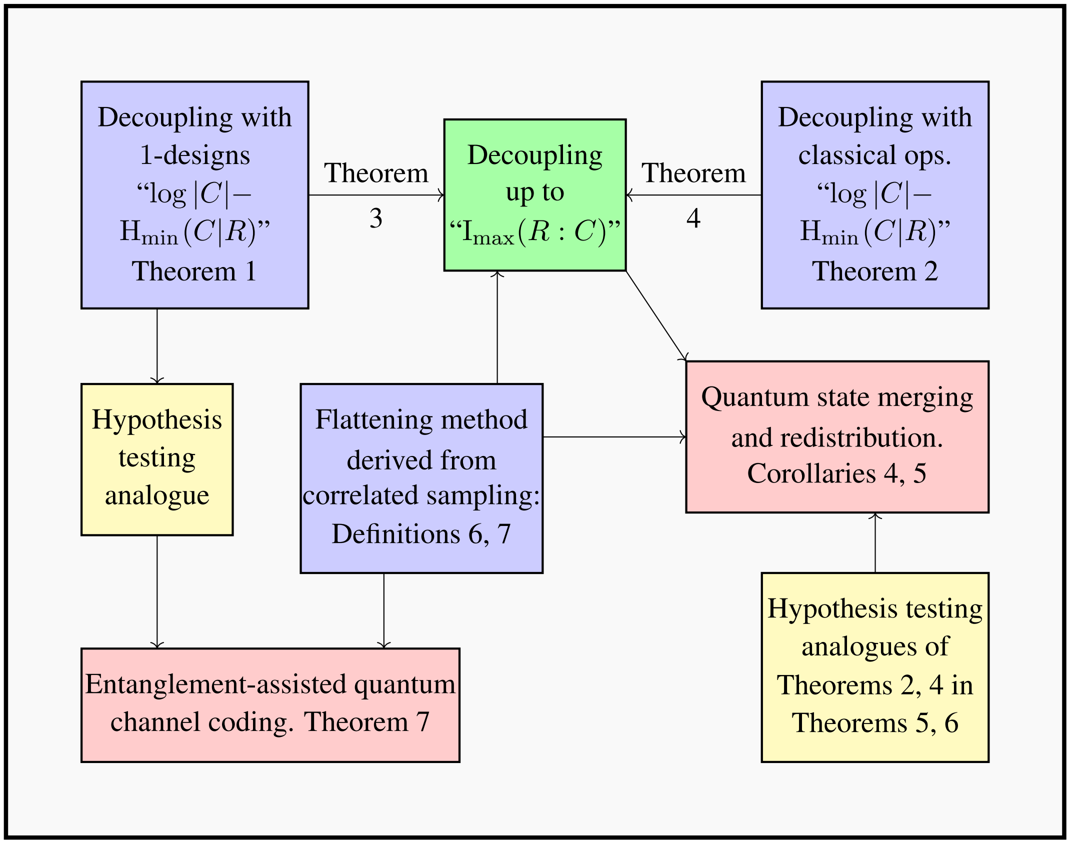
All the logarithms are evaluated to the base 2. Consider a finite dimensional Hilbert space $\mathcal{H}$ endowed with an inner product $\langle \cdot, \cdot \rangle$ (In this paper, we only consider finite dimensional Hilbert-spaces). The ℓ_1 norm of an operator X on $\mathcal{H}$ is $\|X\|_1 := \mathrm{Tr}\sqrt{X^\dagger X}$ and ℓ_2 norm is $\|X\|_2 := \sqrt{\mathrm{Tr} X X^\dagger}$. A quantum state (or a density matrix or a state) is a positive semi-definite matrix on $\mathcal{H}$ with trace equal to 1. It is called *pure* if and only if its rank is 1. A sub-normalized state is a positive semi-definite matrix on $\mathcal{H}$ with trace less than or equal to 1. Let $|\psi\rangle$ be a unit vector on $\mathcal{H}$, that is $\langle \psi, \psi \rangle = 1$. With some abuse of notation, we use ψ to represent the state and also the density matrix $|\psi\rangle\langle\psi|$, associated with $|\psi\rangle$. Given a quantum state ρ on $\mathcal{H}$, *support of* ρ, called $\mathrm{supp}(\rho)$ is the subspace of $\mathcal{H}$ spanned by all eigenvectors of ρ with non-zero eigenvalues.

A *quantum register* A is associated with some Hilbert space $\mathcal{H}_A$. Define $|A| := \dim(\mathcal{H}_A)$. Let $\mathcal{L}(A)$ represent the set of all linear operators on $\mathcal{H}_A$. For operators $O, O' \in \mathcal{L}(A)$, the notation $O \preceq O'$ represents the Löwner order, that is, $O' - O$ is a positive semi-definite matrix. We denote by $\mathcal{D}(A)$, the set of quantum states on the Hilbert space $\mathcal{H}_A$. State ρ with subscript A indicates $\rho_A \in \mathcal{D}(A)$. If two registers A, B are associated with the same Hilbert space, we shall represent the relation by $A \equiv B$. Composition of two registers A and B, denoted AB, is associated with Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$. For two quantum states $\rho \in \mathcal{D}(A)$ and $\sigma \in \mathcal{D}(B)$, $\rho \otimes \sigma \in \mathcal{D}(AB)$ represents the tensor product (Kronecker product) of ρ and σ. The identity operator on $\mathcal{H}_A$ (and associated register A) is denoted I_A. The maximally mixed state $\frac{\mathrm{I}_A}{|A|}$ on register A is represented by μ_A.

Let $\rho_{AB} \in \mathcal{D}(AB)$. We define

$$\rho_B := \mathrm{Tr}_A \rho_{AB} := \sum_i (\langle i| \otimes \mathrm{I}_B) \rho_{AB} (|i\rangle \otimes \mathrm{I}_B),$$

where $\{|i\rangle\}_i$ is an orthonormal basis for the Hilbert space $\mathcal{H}_A$. The state $\rho_B \in \mathcal{D}(B)$ is referred to as the marginal state of ρ_{AB}. Unless otherwise stated, a missing register from subscript in a state will represent partial trace over that register. Given a $\rho_A \in \mathcal{D}(A)$, a *purification* of ρ_A is a pure state $\rho_{AB} \in \mathcal{D}(AB)$ such



Supplementary Figure 1: An outline of our results, which are all derived in the one-shot setting. All the decoupling statements are stated as convex-split theorems. The results in blue rectangles are main tools that may be of independent interest. Method A in the main text corresponds to the top two blue rectangles and Method B in the main text corresponds to the lower blue rectangle. The results in red rectangles are quantum communication tasks for which we obtain entanglement cost proportional to the *number of qubits of register to be communicated*, while maintaining the best known communication bounds. The result in green rectangle is the near optimal decoupling result and those in yellow rectangles are the hypothesis testing/position-based decoding analogues of convex-split theorems.

that $\mathrm{Tr}_B\rho_{AB} = \rho_A$. Purification of a quantum state is not unique. Suppose $A \equiv B$. Given $\{|i\rangle_A\}$ and $\{|i\rangle_B\}$ as orthonormal bases over $\mathcal{H}_A$ and $\mathcal{H}_B$ respectively, the *canonical purification* of a quantum state ρ_A is $(\rho_A^{\frac{1}{2}} \otimes \mathrm{I}_B)\left(\sum_i |i\rangle_A|i\rangle_B\right)$.

A quantum map $\mathcal{E} : \mathcal{L}(A) \to \mathcal{L}(B)$ is a completely positive and trace preserving (CPTP) linear map (mapping states in $\mathcal{D}(A)$ to states in $\mathcal{D}(B)$). A *unitary* operator $U_A : \mathcal{H}_A \to \mathcal{H}_A$ is such that $U_A^{\dagger}U_A = U_AU_A^{\dagger} = \mathrm{I}_A$. An *isometry* $V_{A\to B} : \mathcal{H}_A \to \mathcal{H}_B$ is such that $V^{\dagger}V = \mathrm{I}_A$ and $VV^{\dagger} = \mathrm{I}_B$. The set of all unitary operations on register A is denoted by $\mathcal{U}(A)$. Some standard unitaries are the $\mathsf{X}, \mathsf{Z}, \mathsf{H}$ (Pauli-X, Pauli-Z and Hadamard, respectively) gates on qubits, the CNOT gate on a pair of qubits and the Toffoli gate on three qubits [1]. We will drop the register labels on unitaries unless when it is required. We shall consider the following information theoretic quantities. We consider only normalized states in the definitions below. Let $\varepsilon \in (0, 1)$.

1. **Fidelity** ([2], see also [3]) For $\rho_A, \sigma_A \in \mathcal{D}(A)$,

$$\mathrm{F}(\rho_A, \sigma_A) \stackrel{\text{def}}{=} \|\sqrt{\rho_A}\sqrt{\sigma_A}\|_1.$$

   For classical probability distributions $P = \{p_i\}, Q = \{q_i\}$,

$$\mathrm{F}(P, Q) \stackrel{\text{def}}{=} \sum_i \sqrt{p_i \cdot q_i}.$$

2. **Purified distance** ([4]) For $\rho_A, \sigma_A \in \mathcal{D}(A)$,

$$\mathrm{P}(\rho_A, \sigma_A) = \sqrt{1 - \mathrm{F}^2(\rho_A, \sigma_A)}.$$

3. ε**-ball** For $\rho_A \in \mathcal{D}(A)$,

$$\mathcal{B}^{\varepsilon}(\rho_A) \stackrel{\text{def}}{=} \{\rho'_A \in \mathcal{D}(A) |\ \mathrm{P}(\rho_A, \rho'_A) \leq \varepsilon\}.$$

4. **Smooth max-relative entropy** ([5], see also [6]) For $\rho_A, \sigma_A \in \mathcal{D}(A)$ such that $\text{supp}(\rho_A) \subset \text{supp}(\sigma_A)$,

$$\mathrm{D}^{\varepsilon}_{\max}(\rho_A \| \sigma_A) \stackrel{\text{def}}{=} \min_{\rho'_A \in \mathcal{B}^{\varepsilon}(\rho_A)} \min\{\lambda \in \mathbb{R} : 2^{\lambda}\sigma_A \geq \rho'_A\}.$$

5. **Hypothesis testing relative entropy** ([7], see also [8]) For $\rho_A, \sigma_A \in \mathcal{D}(A)$,

$$\mathrm{D}^{\varepsilon}_{\mathrm{H}}(\rho_A \| \sigma_A) \stackrel{\text{def}}{=} \max_{0 < \Pi < I, \mathrm{Tr}(\Pi\rho_A) \geq 1 - \varepsilon} \log\left(\frac{1}{\mathrm{Tr}(\Pi\sigma_A)}\right).$$

6. **Max-information** ([9]) For $\rho_{AB} \in \mathcal{D}(AB)$,

$$\mathrm{I}_{\max}(A : B)_{\rho} \stackrel{\text{def}}{=} \mathrm{D}_{\max}(\rho_{AB} \| \rho_A \otimes \rho_B).$$

7. **Smooth max-information** ([9]) For $\rho_{AB} \in \mathcal{D}(AB)$,

$$\mathrm{I}^{\varepsilon}_{\max}(A : B)_{\rho} \stackrel{\text{def}}{=} \mathrm{D}_{\max}(\rho_{AB} \| \rho_A \otimes \rho_B)\, \varepsilon.$$

8. **Conditional min-entropy** ([10]) For $\rho_{AB} \in \mathcal{D}(AB)$,

$$\mathrm{H}_{\min}(A|B)_{\rho} \stackrel{\text{def}}{=} - \min_{\sigma_B \in \mathcal{D}(B)} \mathrm{D}_{\max}(\rho_{AB} \| \mathrm{I}_A \otimes \sigma_B).$$

9. **Smooth conditional min-entropy** ([10]) For $\rho_{AB} \in \mathcal{D}(AB)$,

$$\mathrm{H}^{\varepsilon}_{\min}(A|B)_{\rho} \stackrel{\text{def}}{=} \max_{\rho' \in \mathcal{B}^{\varepsilon}(\rho)} \mathrm{H}_{\min}(A|B)_{\rho'}.$$

2 Supplementary discussion

2.1 Convex-split with improved resources: basic constructions

We begin this section by providing a construction of convex-split of a quantum state that uses small amount of additional randomness.

2.1.1 Convex-split using a mixture of unitaries from a 1-design

The unitary 1-design is defined as follows.

Definition 1. *Fix a register C. A collection of unitaries $\{V_x : \mathcal{H}_C \to \mathcal{H}_C\}_{x=1}^{|C|^2}$ form a 1-design if*

$$\frac{1}{|C|^2}\sum_x V_x M V_x^\dagger = \mathrm{Tr}(M)\frac{\mathrm{I}_C}{|C|}, \quad \forall M \in \mathcal{L}(C).$$

These unitaries have an additional property that they are perfect decouplers, that is,

$$\frac{1}{|C|^2}\sum_x V_x \rho_{RC} V_x^\dagger = \rho_R \otimes \frac{\mathrm{I}_C}{|C|}, \tag{1}$$

which is evident from Definition 1. In order to use a small subset of them decoupling, we will require the notion of pairwise independent functions.

Definition 2. *Let $\{f_j : \mathcal{X} \times \mathcal{X} \to \mathcal{X}\}_{j=1}^{|\mathcal{X}|}$ be a family of pairwise independent functions. That is,*

$$|\{(x_1, x_2) : f_j(x_1, x_2) = x, f_k(x_1, x_2) = x'\}| = 1, \quad \forall x, x', \quad \forall j \neq k.$$

Introduce registers $X_1 \equiv X_2$ such that $|X_1| = |X_2| = |\mathcal{X}|$. Let $V^{(j)} : \mathcal{H}_{CX_1X_2} \to \mathcal{H}_{CX_1X_2}$ be defined as $V^{(j)} = \sum_{x_1,x_2} V_{f_j(x_1,x_2)} \otimes |x_1, x_2\rangle\langle x_1, x_2|_{X_1X_2}$.

As discussed in [11, Example 6] or [12], there exists an efficient construction of pairwise independent function family for any $\mathcal{X}$ with $|\mathcal{X}|$ a prime power. In our setting, $|\mathcal{X}| = |C|^2$. Hence, such a construction exists whenever $\log|C|$ is an integer. The following theorem ensures that convex-split can be achieved with small amount of additional resource. Its proof appears in Section 3.2.

Theorem 1. *Suppose $\log|C|$ is an integer. Let Ψ_{RC} be a quantum state. Define $k \stackrel{\mathrm{def}}{=} \mathrm{D}_{\max}(\Psi_{RC}\|\Psi_R \otimes \mu_C)$. For any integer $N > 1$, define the quantum state*

$$\tau_j \stackrel{\mathrm{def}}{=} V^{(j)}(\Psi_{RC} \otimes \mu_{X_1X_2})V^{(j)\dagger}, \tau \stackrel{\mathrm{def}}{=} \frac{1}{N}\sum_{j=1}^{N} \tau_j,$$

where the unitaries $V^{(j)}$ are given in Definition 2. It holds that

$$\mathrm{D}(\tau\|\Psi_R \otimes \mu_C \otimes \mu_{X_1X_2}) \leq \log\left(1 + \frac{2^k - 1}{N}\right). \tag{2}$$

Now we give a canonical example of a unitary 1-design, which we will use in later sections.

Definition 3. *Given a register C and a basis $\{|c\rangle\}_{c=0}^{|C|-1}$. Define the Heisenberg-Weyl (HW) unitaries $\{V_{a,b}\}_{a,b=0}^{|C|-1}$ with $V_{a,b} : \mathcal{H}_C \to \mathcal{H}_C$ as $V_{a,b} \stackrel{\mathrm{def}}{=} \sum_c e^{\frac{2\pi icb}{|C|}}|c + a\rangle_C\langle c|_C$.*

Following is a well known lemma, which shows that HW unitaries are a 1-design.

Lemma 1. *For all $|c\rangle_C, |c'\rangle_C$, it holds that*

$$\frac{1}{|C|^2}\sum_{a,b} V_{a,b}|c\rangle\langle c'|_C V_{a,b}^\dagger = \delta_{c,c'}\mu_C.$$

In particular, this implies that for any state ρ_{RC},

$$\frac{1}{|C|^2}\sum_{a,b} V_{a,b}\rho_{RC} V_{a,b}^\dagger = \rho_R \otimes \mu_C.$$

Proof. Consider

$$\begin{aligned} \frac{1}{|C|^2}\sum_{a,b} V_{a,b}|c\rangle\langle c'|_C V_{a,b}^\dagger &= \left(\frac{1}{|C|^2}\sum_{a,b} e^{\frac{2\pi i(c-c')b}{|C|}}|c+a\rangle\langle c'+a|_C\right) \\ &= \delta_{c,c'}\left(\frac{1}{|C|}\sum_{a}|c+a\rangle\langle c+a|_C\right) \\ &= \delta_{c,c'}\mu_Z. \end{aligned}$$

In the second equation, we have used Fact 15. Expand $\rho_{RC} = \sum_{c,c'} \rho_R^{c,c'} \otimes |c\rangle\langle c'|_C$. Consider

$$\begin{aligned} \frac{1}{|C|^2}\sum_{a,b} V_{a,b}\rho_{RC} V_{a,b}^\dagger &= \sum_{c,c'} \rho_R^{c,c'} \otimes \left(\frac{1}{|C|^2}\sum_{a,b} V_{a,b}|c\rangle\langle c'|_C V_{a,b}^\dagger\right) \\ &= \sum_{c,c'} \delta_{c,c'}\rho_R^{c,c} \otimes \mu_C = \rho_R \otimes \mu_C. \end{aligned}$$

Above, $\delta_{z,z'}$ is the delta function. This completes the proof. $\square$

Above construction uses HW unitaries which also involve a phase. Hence, these unitaries are not classical. Below, we provide a construction that is completely classical, that is, it permutes basis vectors to basis vectors.

2.1.2 Convex-split with classical unitaries

Fix a register C. Let Q be a register with $|Q| = 2$. We denote by G a register such that $2|C|^2 \geq |G| \geq |C|^2$ is a prime and $\mathcal{H}_G$ is a subspace of $\mathcal{H}_Q \otimes \mathcal{H}_C \otimes \mathcal{H}_C$. This choice of G can be made due to Bertrand's postulate [13], which states that there is a prime number between n and $2n$, for any positive integer n. Let $\{|c\rangle\}_{c=0}^{|C|-1}$ be an arbitrary choice of basis in $\mathcal{H}_C$, a natural example of which is the computational basis. This ensures that $\{|q\rangle|c\rangle|c'\rangle\}$ with $q \in \{0,1\}$ is a basis on $\mathcal{H}_Q \otimes \mathcal{H}_C \otimes \mathcal{H}_C$. We construct a basis $\{|i\rangle\}_{i=0}^{|G|-1}$ on $\mathcal{H}_G$ as follows. We relabel the vector $|0\rangle|c,c'\rangle$ as $|c|C|+c'\rangle$. This gives $|C|^2$ basis vectors for G. The remaining $|G|-|C|^2$ basis vectors are constructed by relabeling $|1\rangle|c,c'\rangle$ as $||C|^2+c|C|+c'\rangle$ as long as $|C|^2+c|C|+c' \leq |G|-1$. We note that the constraint $|C|^2+c|C|+c' \leq |G|-1$ is automatically satisfied in our analysis below, as all the additions, subtractions and multiplications appearing below are performed modulo $|G|$, unless explicitly stated. Now, introduce registers $C_0, C_1 \equiv C$ and $G_1, G_2 \equiv G$, where G_1 is chosen such that $\mathcal{H}_{G_1} \subset \mathcal{H}_Q \otimes \mathcal{H}_{C_0} \otimes \mathcal{H}_{C_1}$.

Definition 4. *For an integer $\ell \in \{0, 1, \dots |G|-1\}$, define the operation $U_\ell : \mathcal{H}_{G_1} \otimes \mathcal{H}_{G_2} \to \mathcal{H}_{G_1} \otimes \mathcal{H}_{G_2}$ as follows:*

$$U_\ell \stackrel{\text{def}}{=} \sum_{i,j} |i + (j-i)\ell\rangle_{G_1} |j + (j-i)\ell\rangle_{G_2} \langle i|_{G_1} \langle j|_{G_2}.$$

We choose the convention that the expression in the kets for registers G_1, G_2 are evaluated modulo $|G|$. The following lemma shows that the unitaries in Definition 4 behave in a 'cyclic' manner, analogous to the permutations in the convex-split lemma from [14].

Lemma 2. *For every $m, \ell \in \{0, 1, \dots |G|-1\}$, it holds that U_ℓ is a unitary. Furthermore*

$$U_m U_\ell = U_{m+\ell}, \quad U_\ell^\dagger = U_{-\ell}.$$

Proof. We first show that U_ℓ is a unitary. Let i, i', j, j' be such that

$$i + (j-i)\ell = i' + (j'-i')\ell, \quad j + (j-i)\ell = j' + (j'-i')\ell.$$

This can be rearranged to obtain

$$(j-j')\ell + (i-i')(1-\ell) = 0, \quad (j-j')(1+\ell) - (i-i')\ell = 0.$$

Multiplying the first equation by ℓ, the second by $(1-\ell)$ and adding, we obtain $j - j' = 0$. Thus, $(i-i')(1-\ell) = 0$ and $(i-i')\ell = 0$. Adding, we conclude that $i = i'$. Hence, U_ℓ is a unitary.
Now, consider

$$\begin{aligned} U_m U_\ell |i\rangle_{G_1} |j\rangle_{G_2} &= U_m |i + (j-i)\ell\rangle_{G_1} |j + (j-i)\ell\rangle_{G_2} \\ &= |i + (j-i)\ell + (j-i)m\rangle_{G_1} |j + (j-i)\ell + (j-i)m\rangle_{G_2} \\ &= |i + (j-i)(\ell+m)\rangle_{G_1} |j + (j-i)(\ell+m)\rangle_{G_2} \\ &= U_{m+\ell} |i\rangle_{G_1} |j\rangle_{G_2}. \end{aligned}$$

Thus, $U_m U_\ell = U_{m+\ell}$. Since $U_0 = \mathrm{I}$, we conclude $U_\ell^\dagger = U_{-\ell}$. This completes the proof. $\square$

Following is an important property of our collection of unitaries and is analogous to Lemma 1.

Lemma 3. *For any quantum state Ψ_{RC_0} and any $m \in \{1, \dots |G|-1\}$, it holds that*

$$\mathrm{Tr}_{G_2}\left(U_m \left(\Psi_{RC_0} \otimes |0\rangle\langle 0|_Q \otimes \mu_{C_1} \otimes \mu_{G_2}\right) U_m^\dagger\right) = \Psi_R \otimes \mu_{G_1},$$

where we use the fact that $\mathcal{H}_{G_1} \subseteq \mathcal{H}_Q \otimes \mathcal{H}_{C_0} \otimes \mathcal{H}_{C_1}$ to change the register label.

Proof. Define $\delta_{i,i'} \overset{\text{def}}{=} 1$ if $i = i'$ and 0 otherwise. Consider

$$\begin{aligned}
&\mathrm{Tr}_{G_2}\left(U_m\left(|i\rangle\langle i'|_{G_1} \otimes \mu_{G_2}\right)U_m^\dagger\right)\\
&= \frac{1}{|G|}\sum_{j=0}^{|G|-1} \mathrm{Tr}_{G_2}\left(U_m\left(|i\rangle\langle i'|_{G_1} \otimes |j\rangle\langle j|_{G_2}\right)U_m^\dagger\right)\\
&= \frac{1}{|G|}\sum_{j=0}^{|G|-1} \mathrm{Tr}_{G_2}\Big(|jm + i(1-m)\rangle\langle jm + i'(1-m)|_{G_1} \otimes |j(m+1) - im\rangle\langle j(m+1) - i'm|_{G_2}\Big)\\
&= \frac{1}{|G|}\sum_{j=0}^{|G|-1} |jm + i(1-m)\rangle\langle jm + i'(1-m)|_{G_1} \cdot \delta_{i,i'}\\
&= \frac{1}{|G|}\sum_{j=0}^{|G|-1} |jm + i(1-m)\rangle\langle jm + i(1-m)|_{G_1} \cdot \delta_{i,i'}\\
&= \mu_{G_1} \cdot \delta_{i,i'},
\end{aligned} \tag{3}$$

where we have used the fact that for $0 < m < |G|$ and $|G|$ prime, the quantity $jm + i(1-m)$ takes all possible values in $\{0, 1, \ldots |G| - 1\}$ as j varies in $\{0, 1, \ldots |G| - 1\}$. For this, observe that for two j, j',

$$jm + i(1-m) = j'm + i(1-m) \implies (j - j')m = 0,$$

which implies $j = j'$ as $m \neq 0$. Now, expand $\Psi_{RC_0} = \sum_{c,c'} \Psi_R^{(c,c')} \otimes |c\rangle\langle c'|_{C_0}$, where $\Psi_R^{(c,c')}$ are some matrices. Observe that $\Psi_R = \mathrm{Tr}_{C_0}(\Psi_{RC_0}) = \sum_c \Psi_R^{(c,c)}$. For any $m > 0$, using Equation 3, we have

$$\begin{aligned}
&\mathrm{Tr}_{G_2}\left(U_m\left(\Psi_{RC_0} \otimes |0\rangle\langle 0|_Q \otimes \mu_{C_1} \otimes \mu_{G_2}\right)U_m^\dagger\right)\\
&= \sum_{c,c',c_1} \frac{1}{|C|}\Psi_R^{(c,c')} \otimes \mathrm{Tr}_{G_2}\left(U_m\left(|c\rangle\langle c'|_{C_0} \otimes |0\rangle\langle 0|_Q \otimes |c_1\rangle\langle c_1|_{C_1} \otimes \mu_{G_2}\right)U_m^\dagger\right)\\
&= \sum_{c,c',c_1} \frac{1}{|C|}\Psi_R^{(c,c')} \otimes \mu_{G_1} \cdot \delta_{c,c'}\\
&= \sum_c \Psi_R^{(c,c)} \otimes \mu_{G_1} = \Psi_R \otimes \mu_{G_1},
\end{aligned}$$

where we have used that fact that $c|C| + c_1 = c'|C| + c_1 \iff c = c'$. This completes the proof. $\square$

Now, we are in a position to prove our main result. Its proof appears in Section 3.2.

Theorem 2. *Let Ψ_{RC} be a quantum state and let $k \overset{\text{def}}{=} \mathrm{D}_{\max}(\Psi_{RC} \| \Psi_R \otimes \mu_C)$. For a subset $S \subseteq \{0, 1, \ldots |G| - 1\}$ of size $N \overset{\text{def}}{=} |S|$, define the quantum state*

$$\tau_{RG_1G_2} \overset{\text{def}}{=} \frac{1}{N}\sum_{\ell \in S} U_\ell\left(\Psi_{RC_0} \otimes |0\rangle\langle 0|_Q \otimes \mu_{C_1} \otimes \mu_{G_2}\right)U_\ell^\dagger,$$

where the unitaries U_ℓ are given in Definition 4. It holds that

$$\mathrm{D}(\tau_{RG_1G_2}\|\Psi_R \otimes \mu_{G_1} \otimes \mu_{G_2}) \leq \log\left(1 + \frac{2^{k+1}-1}{N}\right). \tag{4}$$

From Fact 5, we conclude that

$$\mathrm{F}^2(\tau_{RG_1G_2}, \Psi_R \otimes \mu_{G_1} \otimes \mu_{G_2}) \geq \frac{1}{1 + \frac{2^{k+1}-1}{N}}.$$

An immediate corollary is the smooth version of above result.

Corollary 1. *Let $\varepsilon, \delta \in (0,1)$ and Ψ_{RC} be a quantum state. Let $k \stackrel{\mathrm{def}}{=} \log|C| - \mathrm{H}_{\min}^{\frac{\varepsilon}{2}}(C|R)_\Psi + \log\frac{8}{\varepsilon^3}$ and $N \geq \frac{2^{k+1}}{\delta^2}$. For a set $S \subseteq \{0, 1, \ldots |G|-1\}$ of size $|S| = N$, define the quantum state*

$$\tau_{RG_1G_2} \stackrel{\mathrm{def}}{=} \frac{1}{N}\sum_{\ell \in S} U_\ell \left(\Psi_{RC_0} \otimes |0\rangle\langle 0|_Q \otimes \mu_{C_1} \otimes \mu_{G_2}\right) U_\ell^\dagger.$$

It holds that

$$\mathrm{P}(\tau_{RG_1G_2}, \Psi_R \otimes \mu_{G_1} \otimes \mu_{G_2}) \leq 2\varepsilon + \delta.$$

2.1.3 Implementation of the unitaries in Definition 4

The circuit size of the unitaries in Definition 4 can be bounded as follows. The relabeling $|c_0, c_1\rangle_{C_0C_1} \to |c_0|C| + c_1\rangle_{G_1}$ can be performed by the multiplication algorithm of Schönhage and Strassen [15] using a circuit of size $\mathcal{O}(\log|C| \log\log|C|)$ and depth $\mathcal{O}(\log\log|C|)$. Thus, we focus on unitary transformation over the basis $\{|i\rangle\}_{i=0}^{|G|}$ for $\mathcal{H}_{G_1}$. From Definition 4, we have

$$U_\ell = \sum_{i,j,\ell} |j\ell + i(1-\ell)\rangle_{G_1} |j(\ell+1) - i\ell\rangle_{G_2} \langle i|_{G_1} \langle j|_{G_2}.$$

Using McLaughlin's algorithm [16] based on the algorithm of Schönhage and Strassen [15] (see [17, Section 2.4.3] for details), and the standard techniques of reversible computing [18, 19], the following transformation can be achieved with a circuit of size $\mathcal{O}(\log|C| \log\log|C|)$ and depth $\mathcal{O}(\log\log|C|)$:

$$\begin{aligned} W_{1,\ell} \quad : \quad & |i\rangle_{G_1}|j\rangle_{G_2}|0\rangle_{G_1'}|0\rangle_{G_2'} \\ \to \quad & |i\rangle_{G_1}|j\rangle_{G_2}|0 + j\ell + i(1-\ell)\rangle_{G_1'}|0 + j(\ell+1) - i\ell\rangle_{G_2'} \\ = \quad & |i\rangle_{G_1}|j\rangle_{G_2}|j\ell + i(1-\ell)\rangle_{G_1'}|j(\ell+1) - i\ell\rangle_{G_2'}. \end{aligned}$$

$W_{1,\ell}$ also uses $\mathcal{O}(\log|C|)$ ancillary qubits in initial state $|0\rangle$, which are returned in the initial state after the computation. Now, we swap registers G_1, G_1' and G_2, G_2':

$$\begin{aligned} S \quad : \quad & |i\rangle_{G_1}|j\rangle_{G_2}|j\ell + i(1-\ell)\rangle_{G_1'}|j(\ell+1) - i\ell\rangle_{G_2'} \\ \to \quad & |j\ell + i(1-\ell)\rangle_{G_1}|j(\ell+1) - i\ell\rangle_{G_2}|i\rangle_{G_1'}|j\rangle_{G_2'}. \end{aligned}$$

Swapping two qubits requires three CNOT gates. Hence this operation can be done in depth 3. Finally we observe that

$$i = (\ell+1) \cdot (j\ell + i(1-\ell)) - \ell \cdot (j(\ell+1) - i\ell)$$

and

$$j = \ell \cdot (j\ell + i(1-\ell)) + (\ell - 1) \cdot (j(\ell+1) - i\ell) .$$

Thus, using the aforementioned circuit for modular multiplication and addition, we can achieve the transformation:

$$\begin{aligned} W_{2,\ell} \quad : \quad & |j\ell + i(1-\ell)\rangle_{G_1} |j(\ell+1) - i\ell\rangle_{G_2} |i\rangle_{G'_1} |j\rangle_{G'_2} \\ \to \quad & |j\ell + i(1-\ell)\rangle_{G_1} |j(\ell+1) - i\ell\rangle_{G_2} |i - i\rangle_{G'_1} |j - j\rangle_{G'_2} \\ = \quad & |j\ell + i(1-\ell)\rangle_{G_1} |j(\ell+1) - i\ell\rangle_{G_2} |0\rangle_{G'_1} |0\rangle_{G'_2} . \end{aligned}$$

Thus, U_ℓ can be implemented as $U_\ell = W_{2,ell} S W_{1,\ell}$. All of these constructions can be implemented using the Toffoli gate [18]. Hence, the overall size of the circuit is $\mathcal{O}(\log |C| \log\log |C|)$, depth is $\mathcal{O}(\log\log |C|)$ and additional ancillary $\mathcal{O}(\log |C|)$ qubits initialized in $|0\rangle$, that are reset to $|0\rangle$, are used.

2.2 Decoupling up to the max-mutual information using a flattening procedure

We introduce a close variant of the embezzling state [20]. Below, the registers D, E will serve to hold the embezzling state.

Definition 5. *Let a, n be positive integers such that $n \geq a$ and let D be a register satisfying $|D| \geq n - a$. Define*

$$\xi_D^{a:n} \stackrel{\text{def}}{=} \frac{1}{S(a,n)} \sum_{j=a}^{n} \frac{1}{j} |j\rangle\langle j|_D ,$$

where $S(a,n) \stackrel{\text{def}}{=} \sum_{j=a}^{n} \frac{1}{j}$ is the normalization factor. Define

$$|\xi^{a:n}\rangle_{D'D} \stackrel{\text{def}}{=} \frac{1}{\sqrt{S(a:n)}} \sum_{j=a}^{n} \frac{1}{\sqrt{j}} |j\rangle_{D'} |j\rangle_D$$

as the canonical purification of $\xi_D^{a:n}$, where $D' \equiv D$.

2.2.1 Embezzling uniform states with closeness in max-relative entropy distance

We have the following claim, which is a variant of the property of embezzling states proved in [20] (where the closeness was established in fidelity for general states).

Claim 1. *Let $\delta \in (0, \frac{1}{15})$ and a, b, n be integers such that $n \geq a^{\frac{1}{\delta}}, a \geq 2$ and $a \geq b$. Fix registers D, E satisfying $|D| \geq n$ and $|E| \geq b$. Let W_b be the unitary that acts as*

$$W_b |j\rangle_D |0\rangle_E = |\lfloor j/b \rfloor\rangle_D |j \ (mod\, b)\rangle_E .$$

It holds that

$$W_b \left(\xi_D^{a:n} \otimes |0\rangle\langle 0|_E\right) W_b^\dagger \preceq (1 + 15\delta) \xi_D^{1:n} \otimes \frac{1}{b} \sum_{e=0}^{b-1} |e\rangle\langle e|_E .$$

Proof. Consider

$$
\begin{aligned}
& W_b \left(\xi_D^{a:n} \otimes |0\rangle\langle 0|_E\right) W_b^\dagger \\
&= \frac{1}{S(a,n)} \sum_{j=a}^{n} \frac{1}{j} W_b \left(|j\rangle\langle j|_D \otimes |0\rangle\langle 0|_E\right) W_b^\dagger \\
&= \frac{1}{S(a,n)} \sum_{j=a}^{n} \frac{1}{j} |\lfloor j/b \rfloor\rangle\langle \lfloor j/b \rfloor|_D \otimes |j \pmod b\rangle\langle j \pmod b|_E \\
&= \frac{1}{S(a,n)} \sum_{j'=\lfloor \frac{a}{b} \rfloor}^{\lfloor \frac{n}{b} \rfloor} \sum_{e=0}^{b-1} \frac{1}{bj'+e} |j'\rangle\langle j'|_D \otimes |e\rangle\langle e|_E \\
&\preceq \frac{1}{S(a,n)} \sum_{j'=\lfloor \frac{a}{b} \rfloor}^{\lfloor \frac{n}{b} \rfloor} \sum_{e=0}^{b-1} \frac{1}{bj'} |j'\rangle\langle j'|_D \otimes |e\rangle\langle e|_E \\
&= \frac{1}{S(a,n)} \sum_{j'=\lfloor \frac{a}{b} \rfloor}^{\lfloor \frac{n}{b} \rfloor} \frac{1}{j'} |j'\rangle\langle j'|_D \otimes \sum_{e=0}^{b-1} \frac{1}{b} |e\rangle\langle e|_E \preceq \frac{S(1,n)}{S(a,n)} \xi_D^{1:n} \otimes \sum_{e=0}^{b-1} \frac{1}{b} |e\rangle\langle e|_E.
\end{aligned}
$$

Now, as shown in [21], $|S(a,n) - \log \frac{n}{a}| \leq 4$. Thus,

$$
\frac{S(1,n)}{S(a,n)} \leq \frac{\log n + 4}{\log n - \log a - 4} \leq \frac{1+4\delta}{1-5\delta} \leq 1 + 15\delta.
$$

This completes the proof. □

Following claim shows how to 'unembezzle' a state.

Claim 2. *Fix the integers n, b, a as given in Claim 1. Let the register D satisfy $n^2 \geq |D| \geq (n+1)b$. Let W_b be as defined in Claim 1. It holds that*

$$
W_b^\dagger \left(\xi_D^{1:n} \otimes \frac{1}{b} \sum_{e=0}^{b-1} |e\rangle\langle e|_E \right) W_b \preceq 4 \cdot \xi_D^{1:|D|} \otimes |0\rangle\langle 0|_E.
$$

Proof. We observe that $W_b^\dagger |j\rangle_D |e\rangle_E = |jb+e\rangle_D |0\rangle_E$ for all $j \leq n$ and $e < b$. We leave the action of $W_b^\dagger$

unspecified for $j \geq n, e \geq b$. Consider

$$\begin{aligned}
W_b^\dagger \left(\xi_D^{1:n} \otimes \frac{1}{b} \sum_{e=0}^{b-1} |e\rangle\langle e|_E \right) W_b &= \frac{1}{S(1,n)} \sum_{j=1}^{n} \sum_{e=0}^{b-1} \frac{1}{jb} W_b^\dagger |j\rangle\langle j|_D \otimes |e\rangle\langle e|_E W_b \\
&= \frac{1}{S(1,n)} \sum_{j=1}^{n} \sum_{e=0}^{b-1} \frac{1}{jb} |jb+e\rangle\langle jb+e|_D \otimes |0\rangle\langle 0|_E \\
&\preceq \frac{2}{S(1,n)} \sum_{j=1}^{n} \sum_{e=0}^{b-1} \frac{1}{jb+e} |jb+e\rangle\langle jb+e|_D \otimes |0\rangle\langle 0|_E \\
&\preceq \frac{2}{S(1,n)} \sum_{j'=1}^{nb+b} \frac{1}{j'} |j'\rangle\langle j'|_D \otimes |0\rangle\langle 0|_E \\
&\preceq \frac{2}{S(1,n)} \sum_{j'=1}^{|D|} \frac{1}{j'} |j'\rangle\langle j'|_D \otimes |0\rangle\langle 0|_E \\
&= \frac{2S(1,|D|)}{S(1,n)} \xi_D^{1:|D|} \otimes |0\rangle\langle 0|_E \\
&\preceq 4\xi_D^{1:|D|} \otimes |0\rangle\langle 0|_E,
\end{aligned}$$

where in the last operator inequality, we use the fact that $|D| \leq n^2$. This completes the proof. □

A 'purified version' of above claims is the following restatement of the result in [20].

Claim 3. *Let $\delta \in (0, \frac{1}{25})$. Let a, b, n be positive integers such that $n \geq a^{\frac{1}{\delta}}, a \geq b/\delta$ and let D be a register satisfying $|D| \geq n - a$. Let $|\mu\rangle_{E'E} \overset{\text{def}}{=} \frac{1}{\sqrt{b}} \sum_{e=0}^{b-1} |e\rangle_{E'} |e\rangle_E$. It holds that*

$$\mathrm{P}\left((W_b \otimes W_b) \left(\xi_{D'D}^{a:n} \otimes |0\rangle\langle 0|_{E'} \otimes |0\rangle\langle 0|_E \right) \left(W_b^\dagger \otimes W_b^\dagger \right), \xi_{D'D}^{1:n} \otimes \mu_{EE'} \right) \leq 5\sqrt{\delta}.$$

Proof. We have

$$\begin{aligned}
&(W_b \otimes W_b) |\xi^{a:n}\rangle_{D'D} \otimes |0\rangle_{E'} \otimes |0\rangle_E \\
&= \frac{1}{\sqrt{S(a:n)}} \sum_{j=a}^{n} \frac{1}{\sqrt{j}} |\lfloor j/b \rfloor\rangle_{D'} |\lfloor j/b \rfloor\rangle_D |j \pmod{b}\rangle_{E'} |j \pmod{b}\rangle_E \\
&= \frac{1}{\sqrt{S(a:n)}} \sum_{j'=\lfloor a/b \rfloor}^{\lfloor n/b \rfloor} \sum_{e=0}^{b-1} \frac{1}{\sqrt{bj'+e}} |j', j'\rangle_{D'D} |e, e\rangle_{E'E}.
\end{aligned}$$

Since $|\xi^{1:n}\rangle_{D'D}|\mu\rangle_{E'E} = \frac{1}{\sqrt{S(1:n)}}\sum_{j'=1}^{n}\sum_{e=0}^{b-1}\frac{1}{\sqrt{j'b}}|j',j'\rangle_{D'D}|e,e\rangle_{E'E}$, we have

$$\begin{aligned}
&\mathrm{F}\left((W_b \otimes W_b)\left(\xi^{a:n}_{D'D} \otimes |0\rangle\langle 0|_{E'} \otimes |0\rangle\langle 0|_E\right)\left(W_b^\dagger \otimes W_b^\dagger\right), \xi^{1:n}_{D'D} \otimes \mu\right) \\
&= \frac{1}{\sqrt{S(a:n)S(1:n)}} \sum_{j'=\lfloor a/b \rfloor}^{\lfloor n/b \rfloor} \sum_{e=0}^{b-1} \frac{1}{j'b\sqrt{1+\frac{e}{j'b}}} \\
&\geq \frac{\sqrt{1-\delta}}{\sqrt{S(a:n)S(1:n)}} \sum_{j'=\lfloor a/b \rfloor}^{\lfloor n/b \rfloor} \sum_{e=0}^{b-1} \frac{1}{j'b} \\
&= \sqrt{1-\delta} \cdot \frac{S(\lfloor a/b \rfloor, \lfloor n/b \rfloor)}{\sqrt{S(a:n)S(1:n)}} \geq \sqrt{1-25\delta},
\end{aligned}$$

where we use the fact that $|S(a,n) - \log \frac{n}{a}| \leq 4$. This completes the proof. $\square$

2.2.2 The flattening procedure

We now introduce the following definition, which shows how to extend a suitable quantum state to make it uniform in a subspace.

Definition 6. Flattening a quantum state: *Fix a $\gamma \in (0,1)$ such that $\frac{|C|}{\gamma}$ is an integer and a quantum state $\sigma_C \stackrel{\mathrm{def}}{=} \sum_c q(c)|c\rangle\langle c|_C$ with eigenvalues $q(c)$ that are integer multiples of $\frac{\gamma}{|C|}$. For a register E satisfying $|E| = \frac{|C|}{\gamma}\max_c q(c)$, define the quantum state σ_{CE} as follows:*

$$\sigma_{CE} \stackrel{\mathrm{def}}{=} \sum_c q(c)|c\rangle\langle c|_C \otimes \left(\frac{\gamma}{q(c)|C|} \sum_{e=0}^{\frac{q(c)|C|}{\gamma}-1} |e\rangle\langle e|_E\right) = \frac{\gamma}{|C|}\sum_c \sum_{e=0}^{\frac{q(c)|C|}{\gamma}-1} |c\rangle\langle c|_C \otimes |e\rangle\langle e|_E.$$

Observe that σ_{CE} is uniform in its support.

The flattening of σ_C can be realized in a unitary manner as follows. We define some registers and unitaries required for this process.

Definition 7. *Fix $\delta \in (0, \frac{1}{15})$. Let $a \stackrel{\mathrm{def}}{=} |E| = \frac{|C|}{\gamma}\max_c q(c)$ and $n \stackrel{\mathrm{def}}{=} a^{\frac{1}{\delta}}$. Introduce a register D satisfying $|D| \geq n$ with the quantum state $\xi^{a:n}_D$ as given in Definition 5. Define the unitary $W : \mathcal{H}_{CED} \to \mathcal{H}_{CED}$ as*

$$W \stackrel{\mathrm{def}}{=} \sum_c |c\rangle\langle c|_C \otimes W_{\frac{q(c)|C|}{\gamma}},$$

where $q(c), \gamma$ are given in Definition 6 and the unitary $W_{\frac{q(c)|C|}{\gamma}}$ is defined in Claim 1.

Flattening is ensured via the following relation, which uses Claim 1.

$$W\left(\sigma_C \otimes |0\rangle\langle 0|_E \otimes \xi^{a:n}_D\right)W^\dagger = \sum_c q(c)|c\rangle\langle c|_C \otimes W_{\frac{q(c)|C|}{\gamma}}\left(|0\rangle\langle 0|_E \otimes \xi^{a:n}_D\right)W^\dagger_{\frac{q(c)|C|}{\gamma}} \preceq (1+15\delta)\sigma_{CE}\otimes\xi^{1:n}_D. \quad (5)$$

Given the flattening of a quantum state σ_C, Definition 3 gives us $|\mathrm{supp}(\sigma_{CE})|^2$ unitaries $V_x : \mathrm{supp}(\sigma_{CE}) \to \mathrm{supp}(\sigma_{CE})$. If $|\mathrm{supp}(\sigma_{CE})|^2 = \left(\frac{|C|}{\gamma}\right)^2$ is a prime power, Definition 2 gives us the collection of unitaries $V^{(\ell)} : \mathrm{supp}(\sigma_{CE}) \otimes \mathcal{H}_{X_1X_2} \to \mathrm{supp}(\sigma_{CE}) \otimes \mathcal{H}_{X_1X_2}$, with $\log|X_1| = \log|X_2| = \log\left(\frac{|C|}{\gamma}\right)^2$. This allows us to construct the quantum states

$$\tau_\ell \stackrel{\mathrm{def}}{=} V^{(\ell)} \left(W \left(\Psi_{RC} \otimes |0\rangle\langle 0|_E \otimes \xi_D^{a:n} \right) W^\dagger \otimes \mu_{X_1X_2} \right) V^{(\ell)\dagger}. \tag{6}$$

2.2.3 Revisiting the theorems in subsection 2.1

Now we prove the following theorem, which is the analogue of Theorem 1 for a flattened quantum state. Its proof appears in subsection 3.3.

Theorem 3. *Fix $\varepsilon \in (0,1), \gamma \in (0, \frac{1}{2})$, $\delta \in (0, \frac{1}{15})$ such that $\frac{|C|}{\gamma}$ is a prime power and quantum states Ψ_{RC}, ω_C. Let $k \stackrel{\mathrm{def}}{=} \mathrm{D}_{\max}\left(\Psi_{RC} \| \Psi_R \otimes \omega_C\right)$ and $N > 1$ be an integer. Let σ_C be the quantum state as constructed in the first part of Fact 6 using ω_C. For quantum states τ_ℓ as given in Equation 6, define*

$$\tau \stackrel{\mathrm{def}}{=} \frac{1}{N} \sum_\ell \tau_\ell.$$

It holds that

$$\mathrm{D}\left(\tau \middle\| \Psi_R \otimes \sigma_{CE} \otimes \xi_D^{1:n} \otimes \mu_{X_1X_2}\right) \le 15\delta + \log\left(1 + \frac{2^{k+2} - 1}{N}\right). \tag{7}$$

Since one can choose $\log|D| = \log n \le \frac{1}{\delta} \log \frac{|C|}{\gamma}$, the number of qubits of additional registers is $\log|D| + \log|E| + 2\log|X_1| \le \left(4 + \frac{1}{\delta}\right) \log \frac{|C|}{\gamma}$.

For later application, we also state a smooth version of Theorem 3, which is similar to Corollary 1.

Corollary 2. *Fix $\varepsilon \in (0,1), \delta \in (0, \frac{1}{15}), \gamma \in (0,1)$ such that $\frac{|C|}{\gamma}$ is an integer and a quantum state Ψ_{RC}. Let $k \stackrel{\mathrm{def}}{=} \min_{\Psi'_{RC} \in \mathcal{B}^\varepsilon(\Psi_{RC})} \mathrm{D}_{\max}(\Psi'_{RC} \| \Psi_R \otimes \Psi_C)$ and $N \stackrel{\mathrm{def}}{=} \frac{3 \cdot 2^{k+2}}{\delta^3}$. Let σ_C be the quantum state as constructed in the first part of Fact 6 using Ψ_C. For quantum states τ_ℓ as given in Equation 6, define*

$$\tau \stackrel{\mathrm{def}}{=} \frac{1}{N} \sum_\ell \tau_\ell.$$

It holds that

$$\mathrm{P}(\tau, \Psi_R \otimes \sigma_{CE} \otimes \xi_D^{1:n} \otimes \mu_{X_1X_2}) \le 2\varepsilon + 4\sqrt{\delta}.$$

Since one can choose $\log|D| = \log n \le \frac{1}{\delta} \log \frac{|C|}{\gamma}$, the number of qubits of additional registers is $\log|D| + 2\log|E| \le \left(4 + \frac{1}{\delta}\right) \log \frac{|C|}{\gamma}$.

In a similar manner, we obtain an improved version of Theorem 2. We first construct the desired states to be used in the statement of the Theorem. For the flattening of a quantum state σ_C as given in Definition 6, let $\mathcal{H}'_{CE} \stackrel{\mathrm{def}}{=} \mathrm{supp}(\sigma_{CE}) \subset \mathcal{H}_C \otimes \mathcal{H}_E$ denote the support of σ_{CE}. Introduce registers $C_0E_0 \equiv CE$ and $C_1E_1 \equiv CE$. Let Q be a register such that $|Q| = 2$. Let F be a register such that $|F|$ is a prime,

$\mathcal{H}_F \subseteq \mathcal{H}_Q \otimes \mathcal{H}'_{CE} \otimes \mathcal{H}'_{CE}$ and $\mathrm{supp}(|0\rangle\langle 0|_Q) \otimes \mathcal{H}'_{CE} \otimes \mathcal{H}'_{CE} \subseteq \mathcal{H}_F$. This choice of F is guaranteed by Bertrand's postulate [13]. Introduce register $F_2, F_1 \equiv F$ such that $\mathcal{H}'_{C_0E_0} \otimes \mathcal{H}'_{C_1E_1} \subseteq \mathcal{H}_{F_1}$. We identify the pair (c, e) with an element in $\{0, 1, \ldots \frac{|C|}{\gamma} - 1\}$ through some one to one mapping and let $\{U_\ell\}_{\ell=0}^{|F|-1}$ be the unitaries constructed in Definition 4, by setting $C \leftarrow \mathrm{supp}(\sigma_{CE})$. Observe that $U_\ell : \mathcal{H}_{F_1} \otimes \mathcal{H}_{F_2} \to \mathcal{H}_{F_1} \otimes \mathcal{H}_{F_2}$ are 'classical' as long as choice of the preferred basis on $\mathcal{H}_C$ is the eigenbasis of σ_C. Define the quantum states

$$\tau_\ell \stackrel{\mathrm{def}}{=} U_\ell \left(W \left(\Psi_{RC_0} \otimes |0\rangle\langle 0|_{E_0} \otimes \xi_D^{a:n} \right) W^\dagger \otimes |0\rangle\langle 0|_Q \otimes \sigma_{C_1E_1} \otimes \mu_{F_2} \right) U_\ell^\dagger, \tag{8}$$

where W is as given in Definition 7 and $|D| \geq n$. We have the following theorem. Its proof appears in subsection 3.3.

Theorem 4. *Fix $\varepsilon, \gamma \in (0, 1)$, $\delta \in (0, \frac{1}{15})$ such that $\frac{|C|}{\gamma}$ is an integer and quantum states Ψ_{RC}, ω_C . Let $k \stackrel{\mathrm{def}}{=} \mathrm{D}_{\max}\left(\Psi_{RC} \| \Psi_R \otimes \omega_C\right)$, $S \subseteq \{0, 1, \ldots \frac{|C|^2}{\gamma^2} - 1\}$ and $N \stackrel{\mathrm{def}}{=} |S|$. Let σ_C be the quantum state as constructed in the first part of Fact 6 using ω_C. For the quantum states τ_ℓ as constructed in Equation 8, define*

$$\tau \stackrel{\mathrm{def}}{=} \frac{1}{N} \sum_{\ell \in S} \tau_\ell.$$

It holds that

$$\mathrm{D}\left(\tau \middle\| \Psi_R \otimes \mu_{F_1} \otimes \xi_D^{1:n} \otimes \mu_{F_2}\right) \leq 15\delta + \log\left(1 + \frac{2^{k+2} - 1}{N}\right). \tag{9}$$

Since one can choose $\log|D| = \log n \leq \frac{1}{\delta} \log \frac{|C|}{\gamma}$, the number of qubits of additional registers is $\log|D| + 2\log|F| \leq \left(4 + \frac{1}{\delta}\right) \log \frac{|C|}{\gamma}$.

We state a smooth version of Theorem 4 which will be used later.

Corollary 3. *Fix $\varepsilon, \gamma \in (0, 1)$, $\delta \in (0, \frac{1}{15})$ such that $\frac{|C|}{\gamma}$ is an integer and a quantum state Ψ_{RC} . Let $k \stackrel{\mathrm{def}}{=} \min_{\Psi'_{RC} \in \mathcal{B}^\varepsilon(\Psi_{RC})} \mathrm{D}_{\max}(\Psi'_{RC} \| \Psi_R \otimes \Psi_C)$ and $N \stackrel{\mathrm{def}}{=} \frac{3 \cdot 2^{k+2}}{\delta^3}$. Let σ_C be the quantum state as constructed in the first part of Fact 6 using Ψ_C. For the quantum states τ_ℓ as constructed in Equation 8, define*

$$\tau \stackrel{\mathrm{def}}{=} \frac{1}{N} \sum_{\ell=1}^{N} \tau_\ell.$$

It holds that

$$\mathrm{P}(\tau, \Psi_R \otimes \mu_{F_1} \otimes \xi_D^{1:n} \otimes \mu_{F_2}) \leq 2\varepsilon + 4\sqrt{\delta}.$$

Since one can choose $\log|D| = \log n \leq \frac{1}{\delta} \log \frac{|C|}{\gamma}$, the number of qubits of additional registers is $\log|D| + 2\log|F| \leq \left(4 + \frac{1}{\delta}\right) \log \frac{|C|}{\gamma}$.

2.3 Analogues of position-based decoding

We now show how to perform hypothesis testing as a dual to Theorem 2, in analogy with position-based decoding [22]. We note that similar construction can achieve a dual to Theorem 1, but we do not state it here as it will be constructed in details in Theorem 7. We have the following theorem. Its proof appears in subsection 3.4.

Theorem 5. *Let $\varepsilon \in (0,1)$ and Ψ_{BC} be a quantum state. Let $\mathcal{S} \subseteq \{0,1,\dots |G|-1\}$ such that*

$$|\mathcal{S}| \leq \frac{\delta^2}{4\varepsilon} 2^{\mathrm{D}_{\mathrm{H}}^{\varepsilon}(\Psi_{BC} \| \Psi_B \otimes \mu_C)}.$$

For each $\ell \in \mathcal{S}$, let τ_ℓ be the quantum state defined in Theorem 2 with $\Psi_{RC} \leftarrow \Psi_{BC}$. There exists an POVM $\{\Lambda_{-1}, \Lambda_\ell\}_{\ell \in \mathcal{S}}$ such that

$$\mathrm{Tr}\left(\Lambda_\ell \tau_\ell\right) \geq 1 - \varepsilon - 4\delta \quad \forall \ell \in \mathcal{S}. \tag{10}$$

Along the lines similar to Theorem 5, we have the following theorem for position-based decoding. We will directly use the registers and unitaries as introduced in Theorem 4. The proof appears in subsection 3.4.

Theorem 6. *Let $\varepsilon \in (0,1), \delta \in (0, \frac{1}{15})$ and Ψ_{BC}, ω_C be quantum states. Let $\mathcal{S} \subseteq \{0,1,\dots |G|-1\}$ such that*

$$|\mathcal{S}| \leq \frac{\delta^2}{4\varepsilon} 2^{\mathrm{D}_{\mathrm{H}}^{\varepsilon}(\Psi_{BC} \| \Psi_B \otimes \omega_C)}.$$

Let σ_C be the quantum state as constructed in the second part of Fact 6 using ω_C. Let τ_ℓ be the quantum states as defined in Equation 8, using the quantum states $\Psi_{RC} \leftarrow \Psi_{BC}$, σ_{CE} and by choosing $|D| \leq 2n \cdot |E| \leq 2|E|^{1+\frac{1}{\delta}}$. There exists a collection of POVM $\{\Lambda_{-1}, \Lambda_\ell\}_{\ell \in \mathcal{S}}$ such that

$$\mathrm{Tr}\left(\Lambda_\ell \tau_\ell\right) \geq 1 - \varepsilon - 64\delta \quad \forall \ell \in \mathcal{S}. \tag{11}$$

2.4 Applications

2.4.1 Entanglement-assisted quantum channel coding

We show how exponential improvement in entanglement can be obtained for entanglement-assisted quantum channel coding, in comparison to the entanglement required in [22]. We begin by defining an entanglement-assisted code.

Definition 8. *Fix an $\varepsilon \in (0,1)$ and a positive integer R. Let M' be a register of dimension $|M| = 2^R$. A (R, ε) entanglement-assisted code for a quantum channel $\mathcal{N}_{C \to B}$ consists of a shared entanglement $|\Theta\rangle_{E_A E_B}$ between Alice (E_A) and Bob (E_B) and*

- *An encoding operation $\mathcal{E}_m : \mathcal{L}(E_A) \to \mathcal{L}(C)$ for each $m \in \{1, 2, \dots 2^R\}$,*
- *A decoding operation $\mathcal{D} : \mathcal{L}(BE_B) \to \mathcal{L}(M')$ which leads to a classical distribution on register M' such that*

$$\Pr\left[M' \neq m\right] \leq \varepsilon, \quad \forall m \in \{1, 2, \dots 2^R\}.$$

We have the following theorem, near-optimality of which is shown by the converse given in [23]. Its proof appears in subsection 3.5.

Theorem 7. *Let $\varepsilon, \delta' \in (0,1), \delta \in (0, \frac{1}{25}), \gamma \in (0, \frac{1}{2})$. For any pure quantum state $|\Psi\rangle_{AC}$ and*

$$R \leq \mathrm{D}_{\mathrm{H}}^{\varepsilon}\left(\mathcal{N}_{A \to B}(\Psi_{AC}) \| \mathcal{N}_{A \to B}(\Psi_A) \otimes \Psi_C\right) - 5 - \log \frac{4(\varepsilon + 4\gamma^{1/4})}{\delta'}, \tag{12}$$

there exists a $(R, \varepsilon + 4\gamma^{1/4} + \delta' + 20\sqrt{\delta})$ entanglement-assisted code for a quantum channel $\mathcal{N}_{A \to B}$. The protocol uses $\frac{1}{\delta} \log \frac{|A|}{\gamma \cdot \delta}$ qubits of shared entanglement and $4 \log |A|$ bits of shared randomness. The latter can be fixed by standard derandomization argument.

2.4.2 Consequences for quantum state merging and quantum state redistribution

Combining Corollary 3 (which is a smooth version of Theorem 4; alternatively we could use Corollary 2) and Theorem 6, we exponentially improve upon the entanglement cost of the protocol for quantum state redistribution given in [24]. Since the proof is similar to that given in [24], we give the statement of the result.

Corollary 4. *Fix $\varepsilon \in (0,1), \delta \in (0, \frac{1}{15})$ and a pure quantum state $|\Psi\rangle_{RABC}$. There exists an entanglement-assisted one-way protocol in which Alice (AC), Bob (B) and Reference (R) start with the quantum state $|\Psi\rangle_{RABC}$ and Alice communicates a message to Bob such that the final state Φ_{RABC} between Alice (A), Bob (BC) and Reference (R) satisfies $\Phi_{RABC} \in \mathcal{B}^{4\varepsilon+65\delta}(\Psi_{RABC})$. Reference plays no role in the protocol. The number of qubits of shared entanglement required is at most $\left(4+\frac{1}{\delta}\right)\log\frac{|C|}{\delta}$ and the number of qubits communicated is*

$$\min_{\omega_C} \frac{1}{2}\left(\mathrm{D}^{\varepsilon}_{\max}\left(\Psi'_{RBC}\middle\|\Psi'_{RB}\otimes\omega_C\right) - \mathrm{D}^{\varepsilon}_{\mathrm{H}}\left(\Psi_{BC}\middle\|\Psi_B\otimes\omega_C\right) + \log\frac{32}{\varepsilon^2\delta^6}\right).$$

By the argument in [14] that shows how a convex-split for Ψ_{RC} can be used to obtain a protocol for the task of quantum state splitting, we obtain the following corollary using Theorem 4.

Corollary 5. *Fix $\varepsilon, \delta \in (0, \frac{1}{15})$ and a pure quantum state $|\Psi\rangle_{RAC}$. There exists an entanglement-assisted one-way protocol in which Alice (AC) and Reference (R) start with the quantum state $|\Psi\rangle_{RAC}$ and Alice communicates a message to Bob such that the final state Φ_{RAC} between Alice (A), Bob (C) and Reference (R) satisfies $\Phi_{RAC} \in \mathcal{B}^{2\varepsilon+8\sqrt{\delta}}(\Psi_{RABC})$. Reference plays no role in the protocol. The number of qubits communicated is*

$$\frac{1}{2}\mathrm{I}^{\varepsilon}_{\max}(R:C)_{\Psi} + 2 + 2\log\frac{1}{\delta}.$$

The number of qubits of entanglement required is at most $\left(4+\frac{1}{\delta}\right)\log\frac{|C|}{\delta}$.

Thus, the result improves upon the number of qubits communicated in [25] by an additive factor of $\log\log|C|$ and at the same time achieves the same number of qubits of entanglement required. It achieves the same communication as given in [14], but exponentially improves upon the number of qubits of entanglement.

3 Supplementary Methods

3.1 Basic facts used in our proofs

We will use the following facts.

Fact 1 (Triangle inequality for purified distance [26]). *For states $\rho_A, \sigma_A, \tau_A \in \mathcal{D}(A)$,*

$$\mathrm{P}(\rho_A, \sigma_A) \leq \mathrm{P}(\rho_A, \tau_A) + \mathrm{P}(\tau_A, \sigma_A).$$

Fact 2 (Monotonicity under quantum operations [27, 28]). *For quantum states ρ, $\sigma \in \mathcal{D}(A)$, and quantum operation $\mathcal{E}(\cdot) : \mathcal{L}(A) \to \mathcal{L}(B)$, it holds that*

$$\|\mathcal{E}(\rho) - \mathcal{E}(\sigma)\|_1 \leq \|\rho - \sigma\|_1 \quad \text{and} \quad \mathrm{F}(\mathcal{E}(\rho), \mathcal{E}(\sigma)) \geq \mathrm{F}(\rho, \sigma) \quad \text{and} \quad \mathrm{D}(\rho\|\sigma) \geq \mathrm{D}(\mathcal{E}(\rho)\|\mathcal{E}(\sigma)).$$

Fact 3 (Uhlmann's theorem [3]). *Let $\rho_A, \sigma_A \in \mathcal{D}(A)$. Let $\rho_{AB} \in \mathcal{D}(AB)$ be a purification of ρ_A and $\sigma_{AC} \in \mathcal{D}(AC)$ be a purification of σ_A. There exists an isometry $V : C \to B$ such that,*

$$\mathrm{F}(|\theta\rangle\langle\theta|_{AB}, |\rho\rangle\langle\rho|_{AB}) = \mathrm{F}(\rho_A, \sigma_A),$$

where $|\theta\rangle_{AB} = (\mathrm{I}_A \otimes V)|\sigma\rangle_{AC}$.

Fact 4 (Gentle measurement lemma [29, 30]). *Let ρ be a quantum state and $0 < A < I$ be an operator. Then*

$$\mathrm{F}(\rho, \frac{A\rho A}{\mathrm{Tr}(A^2\rho)}) \geq \sqrt{\mathrm{Tr}(A^2\rho)}.$$

Following fact implies the Pinsker's inequality.

Fact 5 (Lemma 5 [31]). *For quantum states $\rho_A, \sigma_A \in \mathcal{D}(A)$,*

$$\mathrm{F}(\rho, \sigma) \geq 2^{-\frac{1}{2}\mathrm{D}(\rho\|\sigma)}.$$

Fact 6. *Fix a $\gamma \in (0, 1)$ and a quantum state ω_C. It holds that*

- *there exists a quantum state σ_C such that $\omega_C \preceq \frac{1}{1-\gamma}\sigma_C$ and the eigenvalues of σ_C are integer multiples of $\frac{\gamma}{|C|}$.*
- *there exists a quantum state σ_C such that $\sigma_C \preceq \frac{1}{1-\gamma}\omega_C$ and the eigenvalues of σ_C are integer multiples of $\frac{\gamma}{|C|}$.*

Proof. We prove each item as follows. Let η be chosen below.

- Given the quantum state ω_C, we construct an operator O by increasing each eigenvalue of ω_C to the nearest multiple of $\frac{\eta}{|C|}$, and define $\sigma_C \stackrel{\mathrm{def}}{=} \frac{O}{\mathrm{Tr}(O)}$. We have

$$1 = \mathrm{Tr}(\sigma_C) \leq \mathrm{Tr}(O) \leq \mathrm{Tr}(\sigma_C) + |C|\frac{\eta}{|C|} = 1 + \eta.$$

  Define $\eta' \stackrel{\mathrm{def}}{=} \mathrm{Tr}(O) - 1$ which implies $0 \leq \eta' \leq \eta$. The eigenvalues of σ_C are integer multiples of $\frac{\eta}{(1+\eta')|C|}$. We choose η (which determines η' as well) such that $\frac{\eta}{1+\eta'} = \gamma$. This ensures that $\gamma \leq \eta \leq \frac{\gamma}{1-\gamma}$. Furthermore, eigenvalues of σ_C are integer multiples of $\frac{\gamma}{|C|}$ and

$$\omega_C \preceq O = (1 + \eta')\sigma_C \preceq (1 + \eta)\sigma_C \preceq \frac{1}{1-\gamma}\sigma_C.$$

- This follows in a similar manner. We construct an operator O by decreasing each eigenvalue of ω_C to the nearest multiple of $\frac{\eta}{|C|}$, and define $\sigma_C \stackrel{\mathrm{def}}{=} \frac{O}{\mathrm{Tr}(O)}$. We have

$$1 = \mathrm{Tr}(\omega_C) \geq \mathrm{Tr}(O) \geq \mathrm{Tr}(\omega_C) - |C|\frac{\eta}{|C|} = 1 - \eta.$$

  Define $\eta' \stackrel{\mathrm{def}}{=} 1 - \mathrm{Tr}(O)$, which implies $0 \leq \eta' \leq \eta$. The eigenvalues of σ_C are integer multiples of $\frac{\eta}{(1-\eta')|C|}$. We choose η (which determines η' as well) such that $\frac{\eta}{1-\eta'} = \gamma$. This ensures that $\frac{\gamma}{1+\gamma} \leq \eta \leq \gamma$. Furthermore, eigenvalues of σ_C are integer multiples of $\frac{\gamma}{|C|}$ and

$$\sigma_C = \frac{1}{1-\eta'}O \preceq \frac{1}{1-\eta'}\sigma_C \preceq \frac{1}{1-\eta}\sigma_C \preceq \frac{1}{1-\gamma}\sigma_C.$$

This completes the proof. $\square$

Fact 7 ([22])**.** *Let $\rho_A, \sigma_A \in \mathcal{D}(\mathcal{H}_A)$ be quantum states. Let $\Lambda \in \mathcal{L}(\mathcal{H}_A)$, $0 \preceq \Lambda \preceq \mathrm{I}_A$ be a positive semidefinite operator. Then it holds that*

$$|\sqrt{\mathrm{Tr}\,(\Lambda\rho_A)} - \sqrt{\mathrm{Tr}\,(\Lambda\sigma_A)}| \leq \mathrm{P}(\rho_A, \sigma_A).$$

Fact 8 ([32])**.** *Given quantum states $\rho_A, \sigma_A \in \mathcal{D}(\mathcal{H}_A)$ and their respective canonical purification $|\rho\rangle_{AB}, |\sigma\rangle_{AB}$ (for $B \equiv A$ and some fixed basis over the registers),*

$$\mathrm{F}(\rho_{AB}, \sigma_{AB}) = \mathrm{Tr}\,(\sqrt{\rho_A}\sqrt{\sigma_A}) \geq 1 - \sqrt{1 - \mathrm{F}(\rho_A, \sigma_A)^2} = 1 - \mathrm{P}(\rho_A, \sigma_A).$$

Fact 9. *Let $\rho_{AB} = \sum_i p_i |i\rangle\langle i|_A \otimes \rho_B^{(i)}$ and $\sigma_{AB} = \sum_i p_i |i\rangle\langle i|_A \otimes \sigma_B^{(i)}$ be two classical quantum states, with $\{p_i\}_i$ a probability distribution. It holds that*

$$\mathrm{D}(\rho_{AB} \| \sigma_{AB}) = \sum_i p_i \mathrm{D}\left(\rho_B^{(i)} \middle\| \sigma_B^{(i)}\right).$$

Fact 10 ([14])**.** *Let $\rho_1, \ldots \rho_n, \theta$ be quantum states and $\{p_i\}_i$ be a probability distribution. Define $\rho \overset{\mathrm{def}}{=} \sum_i p_i \rho_i$. Then it holds that*

$$\mathrm{D}\left(\sum_i p_i \rho_i \middle\| \theta\right) = \sum_i p_i \left(\mathrm{D}(\rho_i \| \theta) - \mathrm{D}(\rho_i \| \rho)\right).$$

Fact 11 (Hayashi-Nagaoka inequality [8])**.** *Fix a $c > 1$ and an integer $N > 0$. Let $\{\Omega_0, \ldots \Omega_{N-1}\}_{i=0}$ be a collection of positive semi-definite operators. Define*

$$\Lambda_i \overset{\mathrm{def}}{=} \left(\sum_{i'} \Omega_{i'}\right)^{-\frac{1}{2}} \Omega_i \left(\sum_{i'} \Omega_{i'}\right)^{-\frac{1}{2}}$$

and Λ_{-1} be the projector orthogonal to the support of $\sum_{i'} \Omega_{i'}$. The operators $\{\Lambda_{-1}, \Lambda_0, \ldots \Lambda_{N-1}\}$ form a POVM. Then

$$\mathrm{I} - \Lambda_i \preceq (1 + c)(\mathrm{I} - \Omega_i) + (1 + c + c^{-1}) \left(\sum_{i' \neq i} \Omega_{i'}\right).$$

Fact 12 (Transpose method)**.** *Let C, C' be registers such that $C \equiv C'$. Let $|\Phi\rangle_{CC'}$ be the maximally entangled state on $\mathcal{H}_C \otimes \mathcal{H}_{C'}$ with $\Phi_C = \mu_C$ and $\Phi_{C'} = \mu_{C'}$. For any unitary $U : \mathcal{H}_C \to \mathcal{H}_C$, there exists a unitary $U^T : \mathcal{H}_{C'} \to \mathcal{H}_{C'}$ such that*

$$(U \otimes \mathrm{I}_{C'})|\Phi\rangle_{CC'} = (\mathrm{I}_C \otimes U^T)|\Phi\rangle_{CC'}.$$

The following fact was stated in [22, Claim 4], with proof adapted from [9].

Fact 13. *Let $\delta \in (0, 1)$. For quantum states $\sigma_A, \sigma_B, \rho_{AB}$, there exists a quantum state $\rho'_{AB} \in \mathcal{B}^{\delta}(\rho_{AB})$ such that*

$$\mathrm{D}_{\max}\left(\rho'_{AB} \middle\| \rho'_A \otimes \sigma_B\right) \leq \mathrm{D}_{\max}(\rho_{AB} \| \sigma_A \otimes \sigma_B) + \log \frac{3}{\delta^2}.$$

Fact 14. *Let $\varepsilon \in (0,1)$ and ρ_{AB} be a quantum state. It holds that*

$$\min_{\rho' \in \mathcal{B}^{2\varepsilon}(\rho)} \mathrm{D}_{\max}\left(\rho'_{AB} \middle\| \rho'_A \otimes \frac{\mathrm{I}_B}{|B|}\right) \leq \log|B| - \mathrm{H}^{\varepsilon}_{\min}(B|A)_\rho + 3\log\frac{2}{\varepsilon}.$$

Proof. From [33, Fact 12] (a corollary of an argument in [9]), for every quantum state σ_A, it holds that

$$\min_{\rho' \in \mathcal{B}^{2\varepsilon}(\rho)} \mathrm{D}_{\max}\left(\rho'_{AB} \middle\| \rho'_A \otimes \frac{\mathrm{I}_B}{|B|}\right) \leq \mathrm{D}^{\varepsilon}_{\max}\left(\rho_{AB} \middle\| \sigma_A \otimes \frac{\mathrm{I}_B}{|B|}\right) + 3\log\frac{2}{\varepsilon}.$$

Minimizing over all σ_A, we have

$$\min_{\rho' \in \mathcal{B}^{2\varepsilon}(\rho)} \mathrm{D}_{\max}\left(\rho'_{AB} \middle\| \rho'_A \otimes \frac{\mathrm{I}_B}{|B|}\right) \leq \log|B| + \min_{\sigma_A} \mathrm{D}^{\varepsilon}_{\max}(\rho_{AB}\|\sigma_A \otimes \mathrm{I}_B) + 3\log\frac{2}{\varepsilon}.$$

The proof now concludes by the definition of $\mathrm{H}^{\varepsilon}_{\min}(B|A)_\rho$. □

Fact 15. *It holds that for $0 < a < C$,*

$$\sum_{b=0}^{C-1} e^{\frac{2\pi i ab}{C}} = 0.$$

Proof. Let $S \stackrel{\mathrm{def}}{=} \sum_{b=0}^{C-1} e^{\frac{2\pi i ab}{C}}$. We have

$$e^{\frac{2\pi i a}{C}} S = \sum_{b=0}^{C-1} e^{\frac{2\pi i a(b+1)}{C}} = \sum_{b=1}^{C-1} e^{\frac{2\pi i ab}{C}} + e^{\frac{2\pi i aC}{C}} = 1 + \sum_{b=1}^{C-1} e^{\frac{2\pi i ab}{C}} = S.$$

Thus, $(1 - e^{\frac{2\pi i a}{C}})S = 0$. Since $e^{\frac{2\pi i a}{C}} \neq 1$ for $0 < a < C$, the proof concludes. □

3.2 Proofs in subsection 2.1

Proof of Theorem 1. Observe that $V^{(j)}$ acts controlled on registers X_1, X_2. Thus, Fact 9 ensures that

$$\mathrm{D}(\tau\|\Psi_R \otimes \mu_C \otimes \mu_{X_1X_2}) = \frac{1}{|\mathcal{X}|^2}\sum_{x_1,x_2} \mathrm{D}\left(\frac{1}{N}\sum_{j=1}^{N} V_{f_j(x_1,x_2)}\Psi_{RC}V^{\dagger}_{f_j(x_1,x_2)} \middle\| \Psi_R \otimes \mu_C\right). \tag{13}$$

Consider the relative entropy term in the summation on the right hand side:

$$
\begin{aligned}
&\mathrm{D}\left(\frac{1}{N}\sum_{j=1}^{N} V_{f_j(x_1,x_2)}\Psi_{RC}V^{\dagger}_{f_j(x_1,x_2)}\middle\|\Psi_R\otimes\mu_C\right)\\
&\overset{(a)}{=}\frac{1}{N}\sum_{j=1}^{N}\Bigg(\mathrm{D}\left(V_{f_j(x_1,x_2)}\Psi_{RC}V^{\dagger}_{f_j(x_1,x_2)}\middle\|\Psi_R\otimes\mu_C\right)\\
&-\mathrm{D}\left(V_{f_j(x_1,x_2)}\Psi_{RC}V^{\dagger}_{f_j(x_1,x_2)}\middle\|\frac{1}{N}\sum_{k=1}^{N}V_{f_k(x_1,x_2)}\Psi_{RC}V^{\dagger}_{f_k(x_1,x_2)}\right)\Bigg)\\
&\overset{(b)}{=}\frac{1}{N}\sum_{j=1}^{N}\Bigg(\mathrm{D}\left(\Psi_{RC}\middle\|\Psi_R\otimes V^{\dagger}_{f_j(x_1,x_2)}\mu_C V_{f_j(x_1,x_2)}\right)\\
&-\mathrm{D}\left(\Psi_{RC}\middle\|\frac{1}{N}\Psi_{RC}+\sum_{k\neq j}\frac{1}{N}V^{\dagger}_{f_j(x_1,x_2)}V_{f_k(x_1,x_2)}\Psi_{RC}V^{\dagger}_{f_k(x_1,x_2)}V_{f_j(x_1,x_2)}\right)\Bigg)\\
&\overset{(c)}{=}\frac{1}{N}\sum_{j=1}^{N}\left(\mathrm{D}(\Psi_{RC}\|\Psi_R\otimes\mu_C)-\mathrm{D}\left(\Psi_{RC}\middle\|\frac{1}{N}\Psi_{RC}+\sum_{k\neq j}\frac{1}{N}V^{\dagger}_{f_j(x_1,x_2)}V_{f_k(x_1,x_2)}\Psi_{RC}V^{\dagger}_{f_k(x_1,x_2)}V_{f_j(x_1,x_2)}\right)\right)\\
&=\mathrm{D}(\Psi_{RC}\|\Psi_R\otimes\mu_C)-\frac{1}{N}\sum_{j=1}^{N}\left(\mathrm{D}\left(\Psi_{RC}\middle\|\frac{1}{N}\Psi_{RC}+\frac{1}{N}\sum_{k\neq j}V^{\dagger}_{f_j(x_1,x_2)}V_{f_k(x_1,x_2)}\Psi_{RC}V^{\dagger}_{f_k(x_1,x_2)}V_{f_j(x_1,x_2)}\right)\right),
\end{aligned}
$$

where (a) uses Fact 10, (b) uses unitarity to move $V_{f_j(x_1,x_2)}$ to the second argument of the quantum relative entropy and (c) uses the fact that μ_C is uniform. Thus, using it in Equation 13,

$$
\begin{aligned}
&\mathrm{D}(\tau\|\Psi_R\otimes\mu_C\otimes\mu_{X_1X_2})=\mathrm{D}(\Psi_{RC}\|\Psi_R\otimes\mu_C)\\
&-\frac{1}{N|\mathcal{X}|^2}\sum_{j=1}^{N}\sum_{x_1,x_2}\left(\mathrm{D}\left(\Psi_{RC}\middle\|\frac{1}{N}\Psi_{RC}+\frac{1}{N}\sum_{k\neq j}V^{\dagger}_{f_j(x_1,x_2)}V_{f_k(x_1,x_2)}\Psi_{RC}V^{\dagger}_{f_k(x_1,x_2)}V_{f_j(x_1,x_2)}\right)\right)\\
&\overset{(a)}{\leq}\mathrm{D}(\Psi_{RC}\|\Psi_R\otimes\mu_C)\\
&-\frac{1}{N}\sum_{j=1}^{N}\left(\mathrm{D}\left(\Psi_{RC}\middle\|\frac{1}{N}\Psi_{RC}+\frac{1}{N|\mathcal{X}|^2}\sum_{k\neq j}\sum_{x_1,x_2}V^{\dagger}_{f_j(x_1,x_2)}V_{f_k(x_1,x_2)}\Psi_{RC}V^{\dagger}_{f_k(x_1,x_2)}V_{f_j(x_1,x_2)}\right)\right),
\end{aligned}
$$

where (a) uses the convexity of relative entropy. From the pairwise independent property of the family of functions (Definition 2), this simplifies to

$$
\begin{aligned}
&\mathrm{D}(\tau\|\Psi_R\otimes\mu_C\otimes\mu_{X_1X_2})\leq\mathrm{D}(\Psi_{RC}\|\Psi_R\otimes\mu_C)\\
&-\frac{1}{N}\sum_{j=1}^{N}\left(\mathrm{D}\left(\Psi_{RC}\middle\|\frac{1}{N}\Psi_{RC}+\frac{1}{N}\sum_{k\neq j}\sum_{x,x'}\frac{|\{(x_1,x_2):f_j(x_1,x_2)=x,f_k(x_1,x_2)=x'\}|}{|\mathcal{X}|^2}V^{\dagger}_{x}V_{x'}\Psi_{RC}V^{\dagger}_{x'}V_{x}\right)\right)\\
&=\mathrm{D}(\Psi_{RC}\|\Psi_R\otimes\mu_C)-\frac{1}{N}\sum_{j=1}^{N}\left(\mathrm{D}\left(\Psi_{RC}\middle\|\frac{1}{N}\Psi_{RC}+\frac{1}{N|\mathcal{X}|^2}\sum_{k\neq j}\sum_{x,x'}V^{\dagger}_{x}V_{x'}\Psi_{RC}V^{\dagger}_{x'}V_{x}\right)\right). \qquad (14)
\end{aligned}
$$

Equation 1 ensures that

$$
\begin{aligned}
&\mathrm{D}(\tau\|\Psi_R\otimes\mu_C\otimes\mu_{X_1X_2})\leq \mathrm{D}(\Psi_{RC}\|\Psi_R\otimes\mu_C)\\
&-\frac{1}{N}\sum_{j=1}^{N}\left(\mathrm{D}\left(\Psi_{RC}\middle\|\frac{1}{N}\Psi_{RC}+\frac{1}{N|\mathcal{X}|^2}\sum_{k\neq j}\sum_{x,x'}V_x^\dagger V_{x'}\Psi_{RC}V_{x'}^\dagger V_x\right)\right)\\
&=\mathrm{D}(\Psi_{RC}\|\Psi_R\otimes\mu_C)-\frac{1}{N}\sum_{j=1}^{N}\left(\mathrm{D}\left(\Psi_{RC}\middle\|\frac{1}{N}\Psi_{RC}+\frac{1}{N}\sum_{k\neq j}\Psi_R\otimes\mu_C\right)\right)\\
&=\mathrm{D}(\Psi_{RC}\|\Psi_R\otimes\mu_C)-\mathrm{D}\left(\Psi_{RC}\middle\|\frac{1}{N}\Psi_{RC}+\frac{N-1}{N}\Psi_R\otimes\mu_C\right).
\end{aligned}
\tag{15}
$$

Using the inequality $\Psi_{RC}\preceq 2^k\Psi_R\otimes\mu_C$, we find

$$
\frac{1}{N}\Psi_{RC}+\frac{N-1}{N}\Psi_R\otimes\mu_C\preceq\left(1+\frac{2^k-1}{N}\right)\Psi_R\otimes\mu_C.
$$

Thus, the operator monotonicity of logarithm [34] ensures that

$$
\mathrm{D}\left(\Psi_{RC}\middle\|\frac{1}{N}\Psi_{RC}+\frac{N-1}{N}\Psi_R\otimes\mu_C\right)\geq \mathrm{D}(\Psi_{RC}\|\Psi_R\otimes\mu_C)-\log\left(1+\frac{2^k-1}{N}\right).
$$

Using it in Equation 15, we conclude that

$$
\begin{aligned}
&\mathrm{D}(\tau\|\Psi_R\otimes\mu_C\otimes\mu_{X_1X_2})\\
&\leq \mathrm{D}(\Psi_{RC}\|\Psi_R\otimes\mu_C)-\mathrm{D}(\Psi_{RC}\|\Psi_R\otimes\mu_C)+\log\left(1+\frac{2^k-1}{N}\right)\\
&=\log\left(1+\frac{2^k-1}{N}\right).
\end{aligned}
$$

This completes the proof. □

Proof of Theorem 2. By definition of k, we have $\Psi_{RC_0}\preceq 2^k\Psi_R\otimes\mu_{C_0}$. This implies

$$
\Psi_{RC_0}\otimes|0\rangle\langle 0|_Q\otimes\mu_{C_1}\preceq 2^k\Psi_R\otimes\mu_{C_0}\otimes|0\rangle\langle 0|_Q\otimes\mu_{C_1}\preceq\frac{|G_1|}{|C_0||C_1|}2^k\Psi_R\otimes\mu_{G_1}\preceq 2^{k+1}\Psi_R\otimes\mu_{G_1}. \tag{16}
$$

We have

$$\begin{aligned}
&\mathrm{D}\left(\tau_{RG_1G_2}\|\Psi_R\otimes\mu_{G_1}\otimes\mu_{G_2}\right)\\
&\overset{(a)}{=}\frac{1}{N}\sum_{\ell\in S}\Bigg(\mathrm{D}\Big(U_\ell\left(\Psi_{RC_0}\otimes|0\rangle\langle 0|_Q\otimes\mu_{C_1}\otimes\mu_{G_2}\right)U_\ell^\dagger\Big\|\Psi_R\otimes\mu_{G_1}\otimes\mu_{G_2}\Big)\\
&\qquad\qquad -\mathrm{D}\Big(U_\ell\left(\Psi_{RC_0}\otimes|0\rangle\langle 0|_Q\otimes\mu_{C_1}\otimes\mu_{G_2}\right)U_\ell^\dagger\Big\|\tau_{RG_1G_2}\Big)\Bigg)\\
&\overset{(b)}{=}\frac{1}{N}\sum_{\ell\in S}\Bigg(\mathrm{D}\Big(\Psi_{RC_0}\otimes|0\rangle\langle 0|_Q\otimes\mu_{C_1}\otimes\mu_{G_2}\Big\|\Psi_R\otimes U_\ell^\dagger\left(\mu_{G_1}\otimes\mu_{G_2}\right)U_\ell\Big)\\
&\qquad\qquad -\mathrm{D}\Big(\Psi_{RC_0}\otimes|0\rangle\langle 0|_Q\otimes\mu_{C_1}\otimes\mu_{G_2}\Big\|U_\ell^\dagger\tau_{RG_1G_2}U_\ell\Big)\Bigg)\\
&\overset{(c)}{\leq}\frac{1}{N}\sum_{\ell\in S}\Bigg(\mathrm{D}\Big(\Psi_{RC_0}\otimes|0\rangle\langle 0|_Q\otimes\mu_{C_1}\otimes\mu_{G_2}\Big\|\Psi_R\otimes U_\ell^\dagger\left(\mu_{G_1}\otimes\mu_{G_2}\right)U_\ell\Big)\\
&\qquad\qquad -\mathrm{D}\Big(\Psi_{RC_0}\otimes|0\rangle\langle 0|_Q\otimes\mu_{C_1}\Big\|\mathrm{Tr}_{G_2}\left(U_\ell^\dagger\tau_{RG_1G_2}U_\ell\right)\Big)\Bigg),
\end{aligned}\tag{17}$$

where (a) uses Fact 10, (b) uses the unitarity to U_ℓ to the second argument of the quantum relative entropy and (c) uses the monotonicity of quantum relative entropy (Fact 2) upon tracing out the register G_2. Since $\mu_{G_1}\otimes\mu_{G_2}$ is maximally mixed in the support of U_ℓ,

$$U_\ell^\dagger\left(\mu_{G_1}\otimes\mu_{G_2}\right)U_\ell=\mu_{G_1}\otimes\mu_{G_2}.\tag{18}$$

Moreover, we have

$$\begin{aligned}
&\mathrm{Tr}_{G_2}\left(U_\ell^\dagger\tau_{RG_1G_2}U_\ell\right)\\
&\overset{(a)}{=}\frac{1}{N}\sum_{m\in S}\mathrm{Tr}_{G_2}\left(U_\ell^\dagger U_m\left(\Psi_{RC_0}\otimes|0\rangle\langle 0|_Q\otimes\mu_{C_1}\otimes\mu_{G_2}\right)U_m^\dagger U_\ell\right)\\
&\overset{(b)}{=}\frac{1}{N}\Psi_{RC_0}\otimes|0\rangle\langle 0|_Q\otimes\mu_{C_1}+\frac{1}{N}\sum_{m\in S,m\neq\ell}\mathrm{Tr}_{G_2}\left(U_{m-\ell}\left(\Psi_{RC_0}\otimes|0\rangle\langle 0|_Q\otimes\mu_{C_1}\otimes\mu_{G_2}\right)U_{m-\ell}^\dagger\right)\\
&\overset{(c)}{\preceq}\frac{2^{k+1}}{N}\Psi_R\otimes\mu_{G_1}+\frac{1}{N}\sum_{m\in S,m\neq\ell}\mathrm{Tr}_{G_2}\left(U_{m-\ell}\left(\Psi_{RC_0}\otimes|0\rangle\langle 0|_Q\otimes\mu_{C_1}\otimes\mu_{G_2}\right)U_{m-\ell}^\dagger\right),
\end{aligned}$$

where (a) uses the definition of $\tau_{RG_1G_2}$ as given in the Theorem statement, (b) uses Lemma 2 and (c) uses Equation 16. Using Lemma 3, we conclude that

$$\mathrm{Tr}_{G_2}\left(U_\ell^\dagger\tau_{RG_1G_2}U_\ell\right)\preceq\frac{2^{k+1}}{N}\Psi_R\otimes\mu_{G_1}+\frac{N-1}{N}\Psi_R\otimes\mu_{G_1}=\left(1+\frac{2^{k+1}-1}{N}\right)\Psi_R\otimes\mu_{G_1}.$$

Since logarithm is operator monotone [34],

$$\begin{aligned}
&\mathrm{D}\Big(\Psi_{RC_0}\otimes|0\rangle\langle 0|_Q\otimes\mu_{C_1}\Big\|\mathrm{Tr}_{G_2}\left(U_\ell^\dagger\tau_{RG_1G_2}U_\ell\right)\Big)\\
&\geq\log\left(1+\frac{2^k-1}{N}\right)+\mathrm{D}\left(\Psi_{RC_0}\otimes|0\rangle\langle 0|_Q\otimes\mu_{C_1}\|\Psi_R\otimes\mu_{G_1}\right).
\end{aligned}$$

From Equations 17 and 18,

$$
\begin{aligned}
&\mathrm{D}(\tau_{RG_1G_2}\|\Psi_R\otimes\mu_{G_1}\otimes\mu_{G_2})\\
&\leq \frac{1}{N}\sum_{\ell\in S}\Big(\mathrm{D}(\Psi_{RC_0}\otimes|0\rangle\langle 0|_Q\otimes\mu_{C_1}\otimes\mu_{G_2}\|\Psi_R\otimes\mu_{G_1}\otimes\mu_{G_2}) - \mathrm{D}(\Psi_{RC_0}\otimes|0\rangle\langle 0|_Q\otimes\mu_{C_1}\|\Psi_R\otimes\mu_{G_1})\Big)\\
&\qquad\qquad\qquad + \log\left(1+\frac{2^{k+1}-1}{N}\right)\\
&= \frac{1}{N}\sum_{\ell\in S}\Big(\mathrm{D}(\Psi_{RC_0}\otimes|0\rangle\langle 0|_Q\otimes\mu_{C_1}\|\Psi_R\otimes\mu_{G_1}) - \mathrm{D}(\Psi_{RC_0}\otimes|0\rangle\langle 0|_Q\otimes\mu_{C_1}\|\Psi_R\otimes\mu_{G_1})\Big)\\
&\qquad\qquad\qquad + \log\left(1+\frac{2^{k+1}-1}{N}\right)\\
&= \log\left(1+\frac{2^{k+1}-1}{N}\right).
\end{aligned}
$$

This completes the proof. $\square$

Proof of Corollary 1. From Fact 14, we conclude that

$$\min_{\Psi'_{RC}\in\mathcal{B}^{\varepsilon}(\Psi_{RC})}\mathrm{D}_{\max}\left(\Psi'_{RC}\middle\|\Psi'_R\otimes\mu_C\right)\leq \log|C| - \mathrm{H}_{\min}^{\frac{\varepsilon}{2}}(C|R)_{\Psi} + \log\frac{8}{\varepsilon^3} = k'.$$

Let Ψ'_{RC} be the quantum state achieving the infimum above. Define

$$\tau'_{RG_1G_2}\stackrel{\mathrm{def}}{=}\frac{1}{N}\sum_{\ell\in S}U_\ell\left(\Psi'_{RC_0}\otimes|0\rangle\langle 0|_Q\otimes\mu_{C_1}\otimes\mu_{G_2}\right)U_\ell^\dagger.$$

We use Theorem 2 to conclude that

$$\mathrm{P}(\tau'_{RG_1G_2},\Psi'_R\otimes\mu_{G_1}\otimes\mu_{G_2})\leq\delta.$$

By triangle inequality for purified distance, this implies that

$$\mathrm{P}(\tau_{RG_1G_2},\Psi_R\otimes\mu_{G_1}\otimes\mu_{G_2})\leq 2\varepsilon+\delta.$$

This concludes the proof. $\square$

3.3 Proofs in subsection 2.2

Proof of Theorem 3. From Fact 6, we have that the eigenvalues of σ_C are integer multiples of $\frac{\gamma}{|C|}$ and

$$\omega_C\preceq\frac{1}{1-\gamma}\sigma_C \implies \Psi_{RC}\preceq\frac{1}{1-\gamma}2^k\Psi_R\otimes\sigma_C\preceq 2^{k+1}\Psi_R\otimes\sigma_C. \tag{19}$$

Consider,

$$
\begin{aligned}
W\left(\Psi_{RC}\otimes|0\rangle\langle 0|_E\otimes\xi_D^{a:n}\right)W^\dagger &\stackrel{(a)}{\preceq} 2^{k+1}\Psi_R\otimes W\left(\sigma_C\otimes|0\rangle\langle 0|_E\otimes\xi_D^{a:n}\right)W^\dagger\\
&\stackrel{(b)}{\preceq} 2^{k+1}(1+15\delta)\Psi_R\otimes\sigma_{CE}\otimes\xi_D^{1:n}\\
&\preceq 2^{k+2}\Psi_R\otimes\sigma_{CE}\otimes\xi_D^{1:n},
\end{aligned}
$$

where (a) uses Equation 19 and (b) uses Equation 5. Expand $\Psi_{RC} = \sum_{c,c'} \Psi_R^{(c,c')} \otimes |c\rangle\langle c'|_C$. For convenience, set $b(c) \stackrel{\text{def}}{=} q(c)|C|/\gamma$. Consider

$$\begin{aligned}
&\frac{1}{|X|} \sum_x V_x W \left(\Psi_{RC} \otimes |0\rangle\langle 0|_E \otimes \xi_D^{a:n}\right) W^\dagger V_x^\dagger \\
&\stackrel{(a)}{=} \frac{1}{S(a,n)} \sum_{j=a}^{n} \frac{1}{j} \sum_{c,c'} \Psi_R^{(c,c')} \otimes \frac{1}{|X|} \sum_x V_x \Big(|c\rangle\langle c'|_C \otimes |j \pmod{b(c)}\rangle\langle j \pmod{b(c')}|_D \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad \otimes |\lfloor j/b(c) \rfloor\rangle\langle \lfloor j/b(c') \rfloor|_E \Big) V_x^\dagger \\
&\stackrel{(b)}{=} \sum_{c,c'} \Psi_R^{(c,c)} \otimes \delta_{c,c'} \sigma_{CE} \otimes \frac{1}{S(a,n)} \sum_{j=a}^{n} \frac{1}{j} |\lfloor j/b(c) \rfloor\rangle\langle \lfloor j/b(c) \rfloor|_D \\
&= \sum_c \Psi_R^{(c,c)} \otimes \sigma_{CE} \otimes \frac{1}{S(a,n)} \sum_{j=a}^{n} \frac{1}{j} |\lfloor j/b(c) \rfloor\rangle\langle \lfloor j/b(c) \rfloor|_D \\
&\stackrel{(c)}{\preceq} (1+15\delta) \sum_c \Psi_R^{(c,c)} \otimes \sigma_{CE} \otimes \xi_D^{1:n} \\
&= (1+15\delta) \Psi_R \otimes \sigma_{CE} \otimes \xi_D^{1:n}.
\end{aligned} \tag{20}$$

Above, (a) uses the definition of the state $\xi_D^{a:n}$ in Definition 5 and that of W in Definition 7. The equality (b) uses Lemma 1. The operator inequality (c) uses the fact that $\frac{S(1,n)}{S(a,n)} \leq (1+15\delta)$, as given in Claim 1. The rest of the argument is identical to Theorem 1, up to the factor of $(1+15\delta)$ induced by above operator inequality. This completes the proof. □

Proof of Theorem 4. Fact 6 ensures that

$$\omega_C \preceq \frac{1}{1-\gamma} \sigma_C \implies \Psi_{RC} \preceq \frac{1}{1-\gamma} 2^k \Psi_R \otimes \sigma_C \preceq 2^{k+1} \Psi_R \otimes \sigma_C. \tag{21}$$

Now, we proceed similar to Theorem 2:

$$\begin{aligned}
&\mathrm{D}\left(\tau \middle\| \Psi_R \otimes \mu_{F_1} \otimes \xi_D^{1:n} \otimes \mu_{F_2}\right) \\
&\stackrel{(a)}{=} \frac{1}{N} \sum_{\ell \in S} \left(\mathrm{D}\left(\tau_\ell \middle\| \Psi_R \otimes \mu_{F_1} \otimes \xi_D^{1:n} \otimes \mu_{F_2}\right) - \mathrm{D}(\tau_\ell \| \tau) \right) \\
&\stackrel{(b)}{=} \frac{1}{N} \sum_{\ell \in S} \Bigg(\mathrm{D}\Big(W \left(\Psi_{RC_0} \otimes |0\rangle\langle 0|_{E_0} \otimes \xi_D^{a:n}\right) W^\dagger \otimes |0\rangle\langle 0|_Q \otimes \sigma_{C_1E_1} \otimes \mu_{F_2} \Big\| \Psi_R \otimes \mu_{F_1} \otimes \xi_D^{1:n} \otimes \mu_{F_2} \Big) \\
&\qquad\qquad - \mathrm{D}\Big(W \left(\Psi_{RC_0} \otimes |0\rangle\langle 0|_{E_0} \otimes \xi_D^{a:n}\right) W^\dagger \otimes |0\rangle\langle 0|_Q \otimes \sigma_{C_1E_1} \otimes \mu_{F_2} \Big\| U_\ell^\dagger \tau U_\ell \Big) \Bigg) \\
&\stackrel{(c)}{\leq} \frac{1}{N} \sum_{\ell \in S} \Bigg(\mathrm{D}\Big(W \left(\Psi_{RC_0} \otimes |0\rangle\langle 0|_{E_0} \otimes \xi_D^{a:n}\right) W^\dagger \otimes |0\rangle\langle 0|_Q \otimes \sigma_{C_1E_1} \Big\| \Psi_R \otimes \mu_{F_1} \otimes \xi_D^{1:n} \Big) \\
&\qquad\qquad - \mathrm{D}\Big(W \left(\Psi_{RC_0} \otimes |0\rangle\langle 0|_{E_0} \otimes \xi_D^{a:n}\right) W^\dagger \otimes |0\rangle\langle 0|_Q \otimes \sigma_{C_1E_1} \Big\| \mathrm{Tr}_{F_2}\left(U_\ell^\dagger \tau U_\ell \right) \Big) \Bigg),
\end{aligned} \tag{22}$$

where (a) uses Fact 10, (b) uses the definition of τ_ℓ in Equation 8 and (c) uses the monotonicity of quantum relative entropy (Fact 2) under tracing out the register F_2. Now, we have

$$\begin{aligned}
&\mathrm{Tr}_{F_2}\left(U_\ell^\dagger \tau U_\ell\right)\\
&\overset{(a)}{=} \frac{1}{N}\mathrm{Tr}_{F_2}\left(U_\ell^\dagger \tau_\ell U_\ell\right) + \frac{1}{N}\sum_{m\in S, m\neq \ell}\mathrm{Tr}_{F_2}\left(U_\ell^\dagger \tau_m U_\ell\right)\\
&\overset{(b)}{=} \frac{1}{N} W\left(\Psi_{RC_0}\otimes |0\rangle\langle 0|_{E_0}\otimes \xi_D^{a:n}\right)W^\dagger \otimes |0\rangle\langle 0|_Q \otimes \sigma_{C_1E_1}\\
&+\frac{1}{N}\sum_{m\in S, m\neq \ell}\mathrm{Tr}_{F_2}\left(U_{m-\ell}\left(W\left(\Psi_{RC_0}\otimes |0\rangle\langle 0|_{E_0}\otimes \xi_D^{a:n}\right)W^\dagger \otimes |0\rangle\langle 0|_Q \otimes \sigma_{C_1E_1}\otimes \mu_{F_2}\right)U_{m-\ell}^\dagger\right)\\
&\overset{(c)}{=} \frac{1}{N} W\left(\Psi_{RC_0}\otimes |0\rangle\langle 0|_{E_0}\otimes \xi_D^{a:n}\right)W^\dagger \otimes |0\rangle\langle 0|_Q \otimes \sigma_{C_1E_1} + \frac{1}{N}\sum_{m\in S, m\neq \ell}\mathrm{Tr}_{F_2}\left(\tau_{m-\ell}\right)\\
&\overset{(d)}{\preceq} \frac{1}{N} W\left(\Psi_{RC_0}\otimes |0\rangle\langle 0|_{E_0}\otimes \xi_D^{a:n}\right)W^\dagger \otimes |0\rangle\langle 0|_Q \otimes \sigma_{C_1E_1} + \frac{(1+15\delta)(N-1)}{N}\Psi_R\otimes \mu_{F_1}\otimes \xi_D^{1:n},
\end{aligned} \tag{23}$$

where (a) uses the definition of τ in the statement of the theorem, (b) uses the definition of τ_ℓ in Equation 8 along with the cyclic property of the unitary U_ℓ in Lemma 2, (c) uses the definition of $\tau_{m-\ell}$ from Equation 8 and (d) uses Claim 4. Next, we evaluate

$$\begin{aligned}
&W\left(\Psi_{RC_0}\otimes |0\rangle\langle 0|_{E_0}\otimes \xi_D^{a:n}\right)W^\dagger \otimes |0\rangle\langle 0|_Q \otimes \sigma_{C_1E_1}\\
&\overset{(a)}{\preceq} 2^{k+1}\Psi_R\otimes W\left(\sigma_{C_0}\otimes |0\rangle\langle 0|_{E_0}\otimes \xi_D^{a:n}\right)W^\dagger \otimes |0\rangle\langle 0|_Q \otimes \sigma_{C_1E_1}\\
&\overset{(b)}{\preceq} 2^{k+1}(1+15\delta)\Psi_R\otimes \sigma_{C_0E_0}\otimes |0\rangle\langle 0|_Q \otimes \sigma_{C_1E_1}\otimes \xi_D^{1:n}\\
&\overset{(c)}{\preceq} 2^{k+2}(1+15\delta)\cdot \frac{|F|}{|\mathrm{supp}(\sigma_{CE})|^2}\Psi_R\otimes \mu_{F_1}\otimes \xi_D^{1:n}\\
&\preceq 2^{k+2}(1+15\delta)\Psi_R\otimes \mu_{F_1}\otimes \xi_D^{1:n},
\end{aligned}$$

where (a) uses Equation 21, (b) uses Equation 5 and (c) uses Claim 1. Using this in Equation 23, we conclude that

$$\mathrm{Tr}_{F_2}\left(U_\ell^\dagger \tau U_\ell\right) \preceq (1+15\delta)\cdot\left(1+\frac{2^{k+2}-1}{N}\right)\Psi_R\otimes \mu_{F_1}\otimes \xi_D^{1:n}.$$

Along with Equation 22, this leads to

$$\mathrm{D}\left(\tau \middle\| \Psi_R\otimes \mu_{F_1}\otimes \xi_D^{1:n}\otimes \mu_{F_2}\right) \leq 15\delta + \log\left(1+\frac{2^{k+2}-1}{N}\right).$$

This completes the proof. $\square$

3.4 Proofs in subsection 2.3

Proof of Theorem 5. Let Ω_{BC} be the operator such that

$$\mathrm{Tr}\left(\Omega_{BC}\Psi_{BC}\right) \geq 1-\varepsilon, \quad \mathrm{Tr}\left(\Omega_{BC}\Psi_B\otimes \mu_C\right) = 2^{-\mathrm{D}_{\mathrm{H}}^{\varepsilon}(\Psi_{BC}\|\Psi_B\otimes\mu_C)}.$$

We have

$$\begin{aligned}\mathrm{Tr}(\Omega_{BC_0}\Psi_B\otimes\mu_{G_1}) &\leq \frac{|Q||C|^2}{|G|}\mathrm{Tr}(\Omega_{BC_0}\Psi_B\otimes\mu_{C_0}\otimes\mu_Q\otimes\mu_{C_1})\\ &\leq 2\cdot 2^{-\mathrm{D}^{\varepsilon}_{\mathrm{H}}(\Psi_{BC}\|\Psi_B\otimes\mu_C)} = 2^{1-\mathrm{D}^{\varepsilon}_{\mathrm{H}}(\Psi_{BC}\|\Psi_B\otimes\mu_C)}. \end{aligned} \tag{24}$$

Let $\{\Lambda_{-1},\Lambda_\ell\}_{\ell\in\mathcal{S}}$ be the POVM as constructed in Fact 11 using the operators $U_\ell\Omega_{BC_0}U_\ell^\dagger$. We have

$$\begin{aligned}\mathrm{Tr}\left((\mathrm{I}-\Lambda_\ell)\tau_\ell\right) &= \mathrm{Tr}\left((\mathrm{I}-\Lambda_\ell)U_\ell\Psi_{BC_0}\otimes|0\rangle\langle 0|_Q\otimes\mu_{C_1}\otimes\mu_{G_2}U_\ell^\dagger\right)\\ &\overset{(a)}{\leq} (1+c)\mathrm{Tr}\left((\mathrm{I}-U_\ell\Omega_{BC_0}U_\ell^\dagger)U_\ell\Psi_{BC_0}\otimes|0\rangle\langle 0|_Q\otimes\mu_{C_1}\otimes\mu_{G_2}U_\ell^\dagger\right)\\ &\quad +(2+c+c^{-1})\sum_{m\in\mathcal{S},m\neq\ell}\mathrm{Tr}\left((U_m\Omega_{BC_0}U_m^\dagger)U_\ell\Psi_{BC_0}\otimes|0\rangle\langle 0|_Q\otimes\mu_{C_1}\otimes\mu_{G_2}U_\ell^\dagger\right)\\ &= (1+c)\mathrm{Tr}\left((\mathrm{I}-\Omega_{BC_0})\Psi_{BC_0}\otimes|0\rangle\langle 0|_Q\otimes\mu_{C_1}\otimes\mu_{G_2}\right)\\ &\quad +(2+c+c^{-1})\sum_{m\in\mathcal{S},m\neq\ell}\mathrm{Tr}\left(\Omega_{BC_0}U_{\ell-m}\Psi_{BC_0}\otimes|0\rangle\langle 0|_Q\otimes\mu_{C_1}\otimes\mu_{G_2}U_{\ell-m}^\dagger\right)\\ &\leq (1+c)\varepsilon+(2+c+c^{-1})\sum_{m\in\mathcal{S},m\neq\ell}\mathrm{Tr}\left(\Omega_{BC_0}U_{\ell-m}\Psi_{BC_0}\otimes|0\rangle\langle 0|_Q\otimes\mu_{C_1}\otimes\mu_{G_2}U_{\ell-m}^\dagger\right).\end{aligned}$$

Above, (a) uses Fact 11. From Lemma 3,

$$\mathrm{Tr}_{G_2}\left(U_{\ell-m}\Psi_{BC_0}\otimes|0\rangle\langle 0|_Q\otimes\mu_{C_1}\otimes\mu_{G_2}U_{\ell-m}^\dagger\right)=\Psi_B\otimes\mu_{G_1}.$$

Thus choosing $c=\frac{\delta}{\varepsilon}$ and using Equation 24,

$$\mathrm{Tr}\left((\mathrm{I}-\Lambda_\ell)\tau_\ell\right)\leq\varepsilon+\delta+\frac{4\varepsilon}{\delta}|\mathcal{S}|\mathrm{Tr}\left(\Omega_{BC_0}\Psi_B\otimes\mu_{G_1}\right)\leq\varepsilon+\delta+\frac{4\varepsilon}{\delta}|\mathcal{S}|2^{1-\mathrm{D}^{\varepsilon}_{\mathrm{H}}(\Psi_{BC}\|\Psi_B\otimes\mu_C)}\leq\varepsilon+4\delta,$$

from the choice of $|\mathcal{S}|$. This completes the proof. □

Proof of Theorem 6. We will outline the main steps of the proof, which closely follow those of Theorem 5. Let Ω_{BC} be the operator that satisfies

$$\mathrm{Tr}\left(\Omega_{BC}\Psi_{BC}\right)\geq 1-\varepsilon,\quad \mathrm{Tr}\left(\Omega_{BC}\Psi_B\otimes\omega_C\right)=2^{-\mathrm{D}^{\varepsilon}_{\mathrm{H}}(\Psi_{BC}\|\Psi_B\otimes\omega_C)}.$$

From Fact 6, we have

$$\sigma_C\preceq\frac{1}{1-\gamma}\omega_C\implies\mathrm{Tr}(\Omega_{BC}\Psi_B\otimes\sigma_C)\leq 2\cdot\mathrm{Tr}(\Omega_{BC}\Psi_B\otimes\omega_C)=2^{1-\mathrm{D}^{\varepsilon}_{\mathrm{H}}(\Psi_{BC}\|\Psi_B\otimes\omega_C)}.$$

Let $\{\Lambda_{-1},\Lambda_\ell\}_{\ell\in\mathcal{S}}$ be the POVM constructed in Fact 11 using the operators $\{U_\ell W\Omega_{BC_0}W^\dagger U_\ell^\dagger\}_{\ell\in\mathcal{S}}$. Rest of the calculation follows using Fact 11. The following claim is similar to Lemma 3.

Claim 4. *For any $m\in\{0,1,\ldots|F|-1\}$, it holds that $\mathrm{Tr}_{F_2}(\tau_m)\preceq(1+15\delta)\Psi_R\otimes\mu_{F_1}\otimes\xi_D^{1:n}$.*

Proof. We expand $\Psi_{RC_0} = \sum_{c,c'} \Psi_R^{(c,c')} \otimes |c\rangle\langle c'|_{C_0}$. For convenience, set $b(c) = q(c)|C|/\gamma$. Recall that

$$W|c\rangle_{C_0}|0\rangle_{E_0}|k\rangle_D = |c\rangle_{C_0}|k \pmod{b(c)}\rangle_{E_0}|\lfloor k/b(c)\rfloor\rangle_D.$$

Thus,

$$\begin{aligned}
&\mathrm{Tr}_{F_2}\left(U_m\left(W\left(|c\rangle\langle c'|_{C_0}\otimes|0\rangle\langle 0|_{E_0}\otimes\xi_D^{a:n}\right)W^\dagger\otimes|0\rangle\langle 0|_Q\otimes\sigma_{C_1E_1}\otimes\mu_{F_2}\right)U_m^\dagger\right)\\
&=\sum_{k=a}^{n}\frac{1}{k}\mathrm{Tr}_{F_2}\left(U_m\left(W\left(|c\rangle\langle c'|_{C_0}\otimes|0\rangle\langle 0|_{E_0}\otimes|k\rangle\langle k|_D\right)W^\dagger\otimes|0\rangle\langle 0|_Q\otimes\sigma_{C_1E_1}\otimes\mu_{F_2}\right)U_m^\dagger\right)\\
&=\sum_{k=a}^{n}\frac{1}{k}\mathrm{Tr}_{F_2}\Big(U_m\left(|c\rangle\langle c'|_{C_0}\otimes|k \pmod{b(c)}\rangle\langle k \pmod{b(c')}|_{E_0}\otimes|0\rangle\langle 0|_Q\otimes\sigma_{C_1E_1}\otimes\mu_{F_2}\right)U_m^\dagger\\
&\qquad\qquad\otimes|\lfloor k/b(c)\rfloor\rangle\langle\lfloor k/b(c')\rfloor|_D\Big)
\end{aligned}$$

As shown in Lemma 3,

$$\mathrm{Tr}_{F_2}\left(U_m\left(|c\rangle\langle c'|_{C_0}\otimes|k \pmod{b(c)}\rangle\langle k \pmod{b(c')}|_{E_0}\otimes|0\rangle\langle 0|_Q\otimes\sigma_{C_1E_1}\otimes\mu_{F_2}\right)U_m^\dagger\right)=\mu_{F_1}\cdot\delta_{c,c'}.$$

Hence, we conclude that

$$\begin{aligned}
&\mathrm{Tr}_{F_2}\left(U_m\left(W\left(|c\rangle\langle c'|_{C_0}\otimes|0\rangle\langle 0|_{E_0}\otimes\xi_D^{a:n}\right)W^\dagger\otimes|0\rangle\langle 0|_Q\otimes\sigma_{C_1E_1}\otimes\mu_{F_2}\right)U_m^\dagger\right)\\
&=\mu_{F_1}\otimes\sum_{k=a}^{n}\frac{1}{k}|\lfloor k/b(c)\rfloor\rangle\langle\lfloor k/b(c)\rfloor|_D\cdot\delta_{c,c'}\\
&\preceq(1+15\delta)\mu_{F_1}\otimes\xi_D^{1:n}\cdot\delta_{c,c'},
\end{aligned}$$

where in the last operator inequality, we have used an argument similar to that used in Claim 1. Thus,

$$\begin{aligned}
\mathrm{Tr}_{F_2}\tau_m &= \sum_{c,c'}\Psi_R^{(c,c')}\otimes\mathrm{Tr}_{F_2}\left(U_m\left(W\left(|c\rangle\langle c'|_{C_0}\otimes|0\rangle\langle 0|_{E_0}\otimes\xi_D^{a:n}\right)W^\dagger\otimes|0\rangle\langle 0|_Q\otimes\sigma_{C_1E_1}\otimes\mu_{F_2}\right)U_m^\dagger\right)\\
&\preceq (1+15\delta)\sum_{c}\Psi_R^{(c,c)}\otimes\mu_{F_1}\otimes\xi_D^{1:n}.
\end{aligned}$$

This completes the proof. $\square$

We require the following inequality for $m,\ell\in S$ with $m\neq\ell$.

$$\begin{aligned}
&\mathrm{Tr}\left(U_mW\Omega_{BC_0}W^\dagger U_m^\dagger\tau_\ell\right)\\
&=\mathrm{Tr}\left(U_mW\Omega_{BC_0}W^\dagger U_m^\dagger U_\ell\left(W\left(\Psi_{BC_0}\otimes|0\rangle\langle 0|_{E_0}\otimes\xi_D^{a:n}\right)W^\dagger\otimes|0\rangle\langle 0|_Q\otimes\sigma_{C_1E_1}\otimes\mu_{F_2}\right)U_\ell^\dagger\right)\\
&=\mathrm{Tr}\left(W\Omega_{BC_0}W^\dagger U_{\ell-m}\left(W\left(\Psi_{BC_0}\otimes|0\rangle\langle 0|_{E_0}\otimes\xi_D^{a:n}\right)W^\dagger\otimes|0\rangle\langle 0|_Q\otimes\sigma_{C_1E_1}\otimes\mu_{F_2}\right)U_{\ell-m}^\dagger\right)\\
&\overset{(a)}{\leq}(1+15\delta)\mathrm{Tr}\left(\Omega_{BC_0}W^\dagger\left(\Psi_R\otimes\mu_{F_1}\otimes\xi_D^{1:n}\right)W\right)\\
&\leq 2\cdot\mathrm{Tr}\left(\Omega_{BC_0}W^\dagger\left(\Psi_R\otimes\mu_{F_1}\otimes\xi_D^{1:n}\right)W\right).
\end{aligned}\tag{25}$$

Above, (a) follows from Claim 4. Now, it holds that

$$\mu_{F_1} \preceq \frac{|Q||\mathrm{supp}(\sigma_{CE})|^2}{|F|}\mu_Q \otimes \sigma_{C_0E_0} \otimes \sigma_{C_1E_1} \preceq 2 \cdot \mu_Q \otimes \sigma_{C_0E_0} \otimes \sigma_{C_1E_1}. \tag{26}$$

Thus,

$$\begin{aligned}
&\mathrm{Tr}\left(U_m W \Omega_{BC_0} W^\dagger U_m^\dagger \tau_\ell\right) \\
&\overset{(a)}{\leq} 2 \cdot \mathrm{Tr}\left(\Omega_{BC_0} W^\dagger \left(\Psi_R \otimes \mu_{F_1} \otimes \xi_D^{1:n}\right) W\right) \\
&\overset{(b)}{\leq} 4 \cdot \mathrm{Tr}\left(\Omega_{BC_0} W^\dagger \left(\Psi_R \otimes \mu_Q \otimes \mu_{C_0E_0} \otimes \mu_{C_1E_1} \otimes \xi_D^{1:n}\right) W\right) \\
&= 4 \cdot \mathrm{Tr}\left(\Omega_{BC_0} W^\dagger \left(\Psi_R \otimes \mu_{C_0E_0} \otimes \xi_D^{1:n}\right) W\right),
\end{aligned}$$

where (a) uses Equation 25 and (b) uses Equation 26. Finally, we use Claim 2 to conclude that

$$\begin{aligned}
&\mathrm{Tr}\left(U_m W \Omega_{BC_0} W^\dagger U_m^\dagger \tau_\ell\right) \\
&\leq 4 \cdot \mathrm{Tr}\left(\Omega_{BC_0} W^\dagger \left(\Psi_R \otimes \mu_{C_0E_0} \otimes \xi_D^{1:n}\right) W\right) \\
&\leq 16 \cdot \mathrm{Tr}\left(\Omega_{BC_0} \Psi_R \otimes \sigma_{C_0} \otimes |0\rangle\langle 0|_{E_0} \otimes \xi_D^{1:|D|}\right) \\
&= 16 \cdot \mathrm{Tr}\left(\Omega_{BC_0} \Psi_R \otimes \sigma_{C_0}\right) \leq 2^{5 - \mathrm{D}_{\mathrm{H}}^{\varepsilon}(\Psi_{BC} \| \Psi_B \otimes \omega_C)},
\end{aligned}$$

where we use the fact that Ω_{BC} only acts in the support of $\Psi_R \otimes \sigma_{C_0}$. This completes the proof. □

3.5 Proofs in subsection 2.4

For the ease of presentation in this section, we will represent the relation $\mathrm{P}(|\psi\rangle\langle\psi|, |\phi\rangle) \leq \varepsilon$ between two pure states $|\psi\rangle, |\phi\rangle$ as $|\psi\rangle \overset{\varepsilon}{\approx} |\phi\rangle$.

Proof of Theorem 7. Without loss of generality, we can assume that $|\Psi\rangle_{AC}$ is the canonical purification of Ψ_A, by applying a local unitary on register C which does not change the hypothesis testing relative entropy. Let $\Psi_{BC} \overset{\mathrm{def}}{=} \mathcal{N}_{A\to B}(\Psi_{AC})$.

Useful quantum states and unitaries: We begin by introducing some new quantum states and unitaries used in the proof. We also discuss their properties.

- *Slightly perturbed quantum states:* From Fact 6, there exists a quantum state σ_C such that the eigenvalues of σ_C are integer multiples of $\frac{\gamma}{|C|}$ and

$$\sigma_C \preceq \frac{1}{1-\gamma}\Psi_C \implies \mathrm{P}(\Psi_C, \sigma_C) \leq \sqrt{\gamma}. \tag{27}$$

  Let $|\sigma\rangle_{AC}$ be the canonical purification of σ_C and $\sigma_{BC} \overset{\mathrm{def}}{=} \mathcal{N}_{A\to B}(\sigma_{AC})$. We obtain

$$\mathrm{P}(\sigma_{AC}, \Psi_{AC}) = \sqrt{1 - \mathrm{F}^2(\sigma_{AC}, \Psi_{AC})} \overset{(a)}{\leq} 2\sqrt{\mathrm{P}(\sigma_C, \Psi_C)} \overset{(b)}{\leq} 2\gamma^{1/4}, \tag{28}$$

where (a) uses Fact 8 and (b) uses Equation 27. From this, we infer that

$$\sigma_B = \mathcal{N}_{A\to B}(\sigma_A) \overset{(a)}{\preceq} \frac{1}{1-\gamma}\mathcal{N}_{A\to B}(\Psi_A) = \frac{1}{1-\gamma}\Psi_B, \quad \mathrm{P}(\sigma_{BC}, \Psi_{BC}) \overset{(b)}{\leq} \mathrm{P}(\sigma_{AC}, \Psi_{AC}) \leq 2\gamma^{1/4}, \tag{29}$$

where (a) uses the inequality $\sigma_A \preceq \frac{1}{1-\gamma}\Psi_A$ which follows from Equation 27 and the property of canonical purifications. The inequality (b) uses monotonicity for the purified distance (Fact 2). Let Ω_{BC} be the optimum operator in the definition of $\mathrm{D}_{\mathrm{H}}^{\varepsilon}(\Psi_{BC}\|\Psi_B \otimes \Psi_C)$. From the property of Ω_{BC}, we have

$$\mathrm{Tr}\left(\Omega_{BC}\sigma_{BC}\right) \overset{(a)}{\geq} 1-\varepsilon-4\gamma^{1/4}, \quad \mathrm{Tr}\left(\Omega_{BC}\sigma_B \otimes \sigma_C\right) \overset{(b)}{\leq} \frac{1}{(1-\gamma)^2}\mathrm{Tr}\left(\Omega_{BC}\Psi_B \otimes \Psi_C\right) \leq 2^{2-\mathrm{D}_{\mathrm{H}}^{\varepsilon}(\Psi_{BC}\|\Psi_B\otimes\Psi_C)}, \tag{30}$$

where (a) uses Equation 29 along with Fact 7 and (b) uses Equations 27 and 29. We expand $|\sigma\rangle_{AC} = \sum_c \sqrt{q(c)}|c\rangle_A|c\rangle_C$.

- *Flattened quantum states:* Let E be the register and σ_{CE} be the quantum state as obtained via the flattening procedure in Definition 6. As stated in Definition 6, we have $|E| \leq \frac{|A|}{\gamma}$. Consider the following purification of σ_{CE}, which is maximally entangled.

$$|\sigma'\rangle_{ACE'E} \overset{\mathrm{def}}{=} \sqrt{\frac{\gamma}{|C|}} \sum_{c,e:e\leq \frac{\gamma q(c)}{|C|}} |c,e\rangle_{AE'}|c,e\rangle_{CE}.$$

Let $a \overset{\mathrm{def}}{=} \frac{|C|}{\gamma\cdot\delta}, n = a^{\frac{1}{\delta}}$ and register D satisfy $|D| = n\left(\frac{|C|}{\gamma}+1\right)$. This choice is made to ensure that Claims 1, 2 and 3 apply to register E. From Definition 5, $|\xi^{a:n}\rangle_{D'D}$ is the canonical purification of $\xi_D^{a:n}$ with $D' \equiv D$. Given the unitary $W \overset{\mathrm{def}}{=} \sum_c |c\rangle\langle c| \otimes W_{\frac{q(c)|C|}{\gamma}}$ from Definition 7, let $W^{alice} \overset{\mathrm{def}}{=} W_{AE'D'}$ and $W^{bob} \overset{\mathrm{def}}{=} W_{CED}$.

- *Heisenberg-Weyl unitaries:* Using Definition 3, we obtain a collection of $\left(\frac{|C|}{\gamma}\right)^2 = |\mathrm{supp}(\sigma_{CE})|^2$ unitaries $V_x : \mathrm{supp}(\sigma_{CE}) \to \mathrm{supp}(\sigma_{CE})$. From Claim 3, we have

$$\left(W^{alice} \otimes W^{bob}\right)|\sigma\rangle_{AC} \otimes |\xi^{a:n}\rangle_{D'D} \otimes |0,0\rangle_{E'E} \overset{5\sqrt{\delta}}{\approx} |\sigma'\rangle_{ACE'E} \otimes |\xi^{1:n}\rangle_{D'D}. \tag{31}$$

Since V_x acts in $\mathrm{supp}(\sigma_{CE})$, Fact 12 ensures that there exists a unitary $V_x^T : \mathrm{supp}(\sigma'_{AE'}) \to \mathrm{supp}(\sigma'_{AE'})$ such that

$$(V_x^T \otimes \mathrm{I})|\sigma'\rangle_{ACE'E} = (\mathrm{I} \otimes V_x)|\sigma'\rangle_{ACE'E}. \tag{32}$$

Combining Equations 31 and 32, we conclude

$$\left(V_x^T W^{alice} \otimes W^{bob}\right)|\sigma\rangle_{AC} \otimes |\xi^{a:n}\rangle_{D'D} \otimes |0,0\rangle_{E'E} \overset{5\sqrt{\delta}}{\approx} (\mathrm{I} \otimes V_x)|\sigma'\rangle_{ACE'E} \otimes |\xi^{1:n}\rangle_{D'D}. \tag{33}$$

Consider

$$\begin{aligned}
&\left((W^{alice})^\dagger V_x^T W^{alice} \otimes W^{bob}\right)|\sigma\rangle_{AC} \otimes |\xi^{a:n}\rangle_{D'D} \otimes |0,0\rangle_{E'E} \\
&\overset{5\sqrt{\delta}}{\approx} \left((W^{alice})^\dagger \otimes \mathrm{I}\right)(\mathrm{I}\otimes V_x)|\sigma'\rangle_{ACE'E} \otimes |\xi^{1:n}\rangle_{D'D} \\
&= (\mathrm{I}\otimes V_x)\left((W^{alice})^\dagger \otimes \mathrm{I}\right)|\sigma'\rangle_{ACE'E} \otimes |\xi^{1:n}\rangle_{D'D} \\
&\overset{5\sqrt{\delta}}{\approx} (\mathrm{I}\otimes V_x)(\mathrm{I}\otimes W^{bob})|\sigma\rangle_{AC} \otimes |\xi^{a:n}\rangle_{D'D} \otimes |0,0\rangle_{E'E} \\
&= (\mathrm{I}\otimes V_x W^{bob})|\sigma\rangle_{AC} \otimes |\xi^{a:n}\rangle_{D'D} \otimes |0,0\rangle_{E'E},
\end{aligned}$$

where the first step uses Equation 33 and the third step uses Equation 31. Using the triangle inequality for purified distance (Fact 1), we hence find

$$\begin{aligned}
&\left((W^{alice})^\dagger V_x^T W^{alice} \otimes W^{bob}\right)|\sigma\rangle_{AC} \otimes |\xi^{a:n}\rangle_{D'D} \otimes |0,0\rangle_{E'E} \\
&\overset{10\sqrt{\delta}}{\approx} (\mathrm{I}\otimes V_x W^{bob})|\sigma\rangle_{AC} \otimes |\xi^{a:n}\rangle_{D'D} \otimes |0,0\rangle_{E'E}.
\end{aligned} \tag{34}$$

- *Controlled unitaries using pairwise-independent functions:* Introduce registers X_1, X_2 where $|X_1| = |X_2| = \left(\frac{|C|}{\gamma}\right)^2$. Since $\left(\frac{|C|}{\gamma}\right)^2$ is a prime power, Definition 2 gives a family of functions $\{f_m : \mathcal{X}\times\mathcal{X} \to \mathcal{X}\}$ and a collection of unitaries

$$V^{(m)} = \sum_{x_1,x_2} V_{f_m(x_1,x_2)} \otimes |x_1,x_2\rangle\langle x_1,x_2|_{X_1X_2}. \tag{35}$$

Let $\{\Lambda_{-1}, \Lambda_1, \dots \Lambda_{2^R}\}$ be POVM as defined in Fact 11 using the operators $\{(V^{(m)}W^{bob})\Omega_{BC}(V^{(m)}W^{bob})^\dagger\}_{m=1}^{2^R}$.

Shared resources: Alice and Bob share the state $|\sigma\rangle_{AC}|\xi^{a:n}\rangle_{D'D}|0,0\rangle_{E'E}$ which appear in Equation 34. They also possess the maximally mixed state $\mu_{X_1X_2}$ in shared registers X_1X_2. Thus, the number of qubits of shared entanglement is

$$\log|C| + \log|D| \le \log n + 2\log\frac{|C|}{\gamma} = \frac{1}{\delta}\log\frac{|C|}{\gamma\cdot\delta} + 2\log\frac{|C|}{\gamma} = \frac{1}{\delta}\log\frac{|A|}{\gamma\cdot\delta} + 2\log\frac{|A|}{\gamma}.$$

Encoding: To send the message $m \in \{1,2,\dots 2^R\}$, Alice applies the unitary

$$\sum_{x_1,x_2}(W^{alice})^\dagger V^T_{f_m(x_1,x_2)} W^{alice} \otimes |x_1,x_2\rangle\langle x_1,x_2|_{X_1X_2}$$

on her registers. She then sends the register A through the channel.

Decoding: Bob applies the unitary W^{bob} on his registers. He applies the POVM $\{\Lambda_{-1}, \Lambda_1, \dots \Lambda_{2^R}\}$ and outputs m' upon obtaining the outcome $\Lambda_{m'}$.

Error analysis: Let θ'_m be the quantum state on Bob's registers just after Alice's transmission through the channel. Define the following quantum state:

$$\theta_m \stackrel{\mathrm{def}}{=} \frac{1}{|X_1|^2}\sum_{x_1,x_2} |x_1,x_2\rangle\langle x_1,x_2|_{X_1X_2} \otimes (V_{f_m(x_1,x_2)}W^{bob})\left(\sigma_{BC}\otimes\xi_D^{a:n}\otimes|0\rangle\langle 0|_E\right)(V_{f_m(x_1,x_2)}W^{bob})^\dagger. \tag{36}$$

From Equation 34, we have $\mathrm{P}(\theta_m,\theta'_m)\le 10\sqrt{\delta}$. Thus, from Fact 7,

$$\Pr[M'\neq m] = \mathrm{Tr}\left((1-\Lambda_m)\theta'_m\right) \le 2\mathrm{P}(\theta_m,\theta'_m) + \mathrm{Tr}\left((1-\Lambda_m)\theta_m\right) \le 20\sqrt{\delta} + \mathrm{Tr}\left((1-\Lambda_m)\theta_m\right). \tag{37}$$

We thus conclude

$$\begin{aligned}
\Pr[M'\neq m] &\stackrel{(a)}{\le} 20\sqrt{\delta} + \mathrm{Tr}\left((1-\Lambda_m)\theta_m\right)\\
&\stackrel{(b)}{\le} 20\sqrt{\delta} + (1+c)\left(1-\mathrm{Tr}\left((V^{(m)}W^{bob})\Omega_{BC}(V^{(m)}W^{bob})^\dagger\theta_m\right)\right)\\
&\quad +(2+c+c^{-1})\sum_{m'\neq m}\mathrm{Tr}\left((V^{(m')}W^{bob})\Omega_{BC}(V^{(m')}W^{bob})^\dagger\theta_m\right)\\
&= 20\sqrt{\delta} + (1+c)\left(1-\mathrm{Tr}\left(\Omega_{BC}(V^{(m)}W^{bob})^\dagger\theta_m(V^{(m)}W^{bob})\right)\right)\\
&\quad +(2+c+c^{-1})\sum_{m'\neq m}\mathrm{Tr}\left((V^{(m')}W^{bob})\Omega_{BC}(V^{(m')}W^{bob})^\dagger\theta_m\right),
\end{aligned} \tag{38}$$

where (a) uses Equation 37 and (b) uses Fact 11. From Equations 36 and 35, we have

$$\begin{aligned}
&(V^{(m)}W^{bob})^\dagger\theta_m(V^{(m)}W^{bob})\\
&= \frac{1}{|X_1|^2}\sum_{x_1,x_2}|x_1,x_2\rangle\langle x_1,x_2|_{X_1X_2}\otimes\left(\sigma_{BC}\otimes\xi_D^{a:n}\otimes|0\rangle\langle 0|_E\right).
\end{aligned}$$

Thus, Equation 30 implies

$$\mathrm{Tr}\left(\Omega_{BC}(V^{(m)}W^{bob})^\dagger\theta_m(V^{(m)}W^{bob})\right) = \mathrm{Tr}(\Omega_{BC}\sigma_{BC}) \ge 1-\varepsilon-4\gamma^{1/4}. \tag{39}$$

For $m'\neq m$ in Equation 38, consider

$$\begin{aligned}
&\mathrm{Tr}\left((V^{(m')}W^{bob})\Omega_{BC}(V^{(m')}W^{bob})^\dagger\theta_m\right) \stackrel{(a)}{=} \frac{1}{|X_1|^2}\sum_{x_1,x_2}\\
&\mathrm{Tr}\left(\Omega_{BC}(V_{f_{m'}(x_1,x_2)}W^{bob})^\dagger(V_{f_m(x_1,x_2)}W^{bob})\left(\sigma_{BC}\otimes\xi_D^{a:n}\otimes|0\rangle\langle 0|_E\right)(V_{f_m(x_1,x_2)}W^{bob})^\dagger(V_{f_{m'}(x_1,x_2)}W^{bob})\right)\\
&\stackrel{(b)}{=} \frac{1}{|\mathcal{X}|}\sum_x \mathrm{Tr}\Bigg(\Omega_{BC}(V_xW^{bob})^\dagger\left(\frac{1}{|\mathcal{X}|}\sum_{x'}V_{x'}W^{bob}\left(\sigma_{BC}\otimes\xi_D^{a:n}\otimes|0\rangle\langle 0|_E\right)(V_{x'}W^{bob})^\dagger\right)(V_xW^{bob})\Bigg),
\end{aligned} \tag{40}$$

where (a) uses Equation 36 and (b) uses Definition 2 to introduce variables x, x' in a manner similar to Equation 14. From Equation 20, the term in the inner bracket can be upper bounded as

$$\frac{1}{|\mathcal{X}|}\sum_{x'}V_{x'}W^{bob}\left(\sigma_{BC}\otimes\xi_D^{a:n}\otimes|0\rangle\langle 0|_E\right)(V_{x'}W^{bob})^\dagger \preceq (1+15\delta)\sigma_B\otimes\sigma_{CE}\otimes\xi^{1:n}.$$

Thus Equation 40 simplifies to,

$$
\begin{aligned}
&\mathrm{Tr}\left((V^{(m')}W^{bob})\Omega_{BC}(V^{(m')}W^{bob})^{\dagger}\theta_m\right)\\
&\leq \frac{(1+15\delta)}{|\mathcal{X}|}\sum_x \mathrm{Tr}\Bigg(\Omega_{BC}(V_xW^{bob})^{\dagger}\Bigg(\sigma_B\otimes\sigma_{CE}\otimes\xi^{1:n}\Bigg)(V_xW^{bob})\Bigg)\\
&\overset{(a)}{=} (1+15\delta)\mathrm{Tr}\Bigg(\Omega_{BC}(W^{bob})^{\dagger}\Bigg(\sigma_B\otimes\sigma_{CE}\otimes\xi^{1:n}\Bigg)(W^{bob})\Bigg)\\
&\overset{(b)}{\leq} 4(1+15\delta)\mathrm{Tr}\Bigg(\Omega_{BC}\Bigg(\sigma_B\otimes\sigma_C\otimes\xi^{1:|D|}\otimes|0\rangle\langle 0|_E\Bigg)\Bigg)\\
&\leq 8\cdot\mathrm{Tr}\Bigg(\Omega_{BC}\sigma_B\otimes\sigma_C\Bigg)\overset{(c)}{\leq} 2^{5-\mathrm{D}_{\mathrm{H}}^{\varepsilon}(\Psi_{BC}\|\Psi_B\otimes\Psi_C)}.
\end{aligned}
$$

where (a) uses the fact that $V_x^{\dagger}\sigma_{CE}V_x = \sigma_{CE}$, (b) uses Claim 2 and (c) uses Equation 30. Using it with Equation 39 and Equation 38, we conclude

$$\Pr[M'\neq m]\leq 20\sqrt{\delta}+(1+c)(\varepsilon+4\gamma^{1/4})+\frac{4}{c}2^{R+5-\mathrm{D}_{\mathrm{H}}^{\varepsilon}(\Psi_{BC}\|\Psi_B\otimes\Psi_C)}.$$

Setting $c=\frac{\delta'}{\varepsilon+4\gamma^{1/4}}$ and from the choice of R, the proof concludes. □

Supplementary Referencess

[1] Barenco, A. *et al.* Elementary gates for quantum computation. *Phys. Rev. A* **52**, 3457–3467 (1995).

[2] Jozsa, R. Fidelity for mixed quantum states. *Journal of Modern Optics* **41**, 2315–2323 (1994).

[3] Uhlmann, A. The ”transition probability” in the state space of a *-algebra. *Rep. Math. Phys.* **9**, 273–279 (1976).

[4] Gilchrist, A., Langford, N. K. & Nielsen, M. A. Distance measures to compare real and ideal quantum processes. *Phys. Rev. A* **71**, 062310 (2005).

[5] Datta, N. Min- and max- relative entropies and a new entanglement monotone. *IEEE Trans. Inf. Theory* **55**, 2816–2826 (2009).

[6] Jain, R., Radhakrishnan, J. & Sen, P. A property of quantum relative entropy with an application to privacy in quantum communication. *J. ACM* **56**, 33:1–33:32 (2009).

[7] Buscemi, F. & Datta, N. The quantum capacity of channels with arbitrarily correlated noise. *IEEE Trans. Inf. Theory* **56**, 1447–1460 (2010).

[8] Hayashi, M. & Nagaoka, H. General formulas for capacity of classical-quantum channels. *IEEE Trans. Inf. Theory* **49**, 1753–1768 (2003).

[9] Ciganović, N., Beaudry, N. J. & Renner, R. Smooth max-information as one-shot generalization for mutual information. *IEEE Trans. Inf. Theory* **60**, 1573–1581 (2014).

[10] Renner, R. Security of quantum key distribution (2005). PhD Thesis, ETH Zurich, Diss. ETH No. 16242, arXiv:quant-ph/0512258.

[11] Kopparty, S., Cheung, Y. K. & Nikolov, A. K-wise independent hashing and applications (2013). Lecture notes, http://sites.math.rutgers.edu/ sk1233/courses/topics-S13/lec5.pdf.

[12] Lovett, S. Pairwise independent hash functions and applications (2015). Lecture notes, http://cseweb.ucsd.edu/ slovett/teaching/SP15-CSE190/.

[13] Tchebychev, P. Mémoire sur les nombres premiers. *Journal de mathématiques pures et appliquées* **1**, 366–390 (1852).

[14] Anshu, A., Devabathini, V. K. & Jain, R. Quantum communication using coherent rejection sampling. *Phys. Rev. Lett.* **119**, 120506 (2017).

[15] Schönhage, A. & Strassen, V. Schnelle multiplikation großer zahlen. *Computing* **7**, 281–292 (1971).

[16] McLaughlin, P. B. New frameworks for montgomery's modular multiplication method. *Math. Comput.* **73**, 899–906 (2004).

[17] Brent, R. P. & Zimmermann, P. *Modern Computer Arithmetic, Cambridge Monographs on Applied and Computational Mathematics* (Cambridge University Press, Cambridge, United Kingdom, 2010).

[18] Toffoli, T. Reversible computing. In de Bakker, J. & van Leeuwen, J. (eds.) *Automata, Languages and Programming*, 632–644 (Springer Berlin Heidelberg, Berlin, Heidelberg, 1980).

[19] Saeedi, M. & Markov, I. L. Synthesis and optimization of reversible circuits; a survey. *ACM Comput. Surv.* **45**, 21:1–21:34 (2013).

[20] van Dam, W. & Hayden, P. Universal entanglement transformations without communication. *Phys. Rev. A* **67**, 060302 (2003).

[21] Mascheroni, L. Adnotationes ad calculum integralem euleri, in quibus nonnulla problemata ab eulero proposita resolvuntur. *Galeati, Ticini* (1790).

[22] Anshu, A., Jain, R. & Warsi, N. A. Building blocks for communication over noisy quantum networks. *IEEE Trans. Inf. Theory* **65**, 1287–1306 (2019).

[23] Matthews, W. & Wehner, S. Finite blocklength converse bounds for quantum channels. *IEEE Trans. Inf. Theory* **60**, 7317–7329 (2014).

[24] Anshu, A., Jain, R. & Warsi, N. A. A one-shot achievability result for quantum state redistribution. *IEEE Trans. Inf. Theory* **64**, 1425–1435 (2018).

[25] Berta, M., Christandl, M. & Renner, R. The Quantum Reverse Shannon Theorem based on one-shot information theory. *Comm. Math. Phys.* **306**, 579–615 (2011).

[26] Tomamichel, M. A framework for non-asymptotic quantum information theory (2012). PhD Thesis, ETH Zurich, http://arXiv.org/abs/1203.2142.

[27] Barnum, H., Caves, C. M., Fuch, C. A., Jozsa, R. & Schumacher, B. Noncommuting mixed states cannot be broadcast. *Phys. Rev. Lett.* **76**, 2818–2821 (1996).

[28] Lindblad, G. Completely positive maps and entropy inequalities. *Comm. Math. Phys.* **40**, 147–151 (1975).

[29] Winter, A. Coding theorem and strong converse for quantum channels. *IEEE Trans. Inf. Theory* **45**, 2481–2485 (1999).

[30] Ogawa, T. & Nagaoka, H. A new proof of the channel coding theorem via hypothesis testing in quantum information theory. In *Information Theory, 2002. Proceedings. 2002 IEEE International Symposium on*, 73– (2002).

[31] Jain, R., Radhakrishnan, J. & P, S. A lower bound for the bounded round quantum communication complexity of set disjointness. In *FOCS*, 220–229 (2003).

[32] Anshu, A., Jain, R., Mukhopadhyay, P., Shayeghi, A. & Yao, P. New one-shot quantum protocols with application to communication complexity. *IEEE Trans. Inf. Theory* **62**, 7566–7577 (2016).

[33] Anshu, A., Jain, R. & Warsi, N. A. Convex-split and hypothesis testing approach to one-shot quantum measurement compression and randomness extraction. *IEEE Trans. Inf. Theory* 1–1 (2019).

[34] Carlen, E. Trace inequalities and quantum entropy: an introductory course. entropy and the quantum. *Contemp. Math.* **529**, 73–140 (2010).