Quantum decoupling via efficient ‘classical’ operations and the entanglement cost of one-shot quantum protocols

Anurag Anshu∗ Rahul Jain†

Abstract

We address the question of efficient implementation of quantum protocols, with short depth circuits and small additional resource such as entanglement. We introduce two new methods in this direction. The first method, inspired by the technique of classical correlated sampling, is to unitarily extend a given quantum state into a quantum state uniform in a subspace. The second method involves two new versions of the convex-split lemma that use exponentially small amount of additional resource in comparison to the previous quantum version. Using these methods, we obtain the following results.

First, we consider the task of quantum decoupling, where the aim is to apply an operation on an $n$-qubit register so as to make it independent of an inaccessible quantum system. Most previous works consider the model where one applies a unitary followed by discarding a quantum system. These works achieve decoupling with the aid of a random unitary. It is known that random unitaries can be replaced by random circuits of size $O(n \log n)$ and depth $\text{poly}(\log n)$, or unitary 2 designs based on Clifford circuits of similar size and depth. An equivalent model of decoupling is to apply a unitary followed by discarding a classical system. These two models are related by a Clifford circuit of depth 1.

We consider the second model and show that given any choice of basis such as the computational basis, decoupling can be achieved by a unitary that takes the basis vectors to basis vectors. Thus, the circuit acts in a ‘classical’ manner and additionally uses $O(n)$ catalytic qubits in maximally mixed quantum state. Our unitary performs addition and multiplication modulo a prime and hence achieve the circuit size of $O(n \log n)$ and logarithmic depth.

Next, we construct a new one-shot entanglement-assisted protocol for quantum channel coding that achieves near-optimal communication through a given channel. The number of qubits of entanglement used in this protocol is proportional to the number of qubits input to the channel. Previous one-shot works were either near-optimal in communication but required exponentially more entanglement, or required small amount of entanglement but did not achieve near-optimal communication. We also achieve similar exponential improvement in the entanglement required for one-shot quantum state redistribution, while keeping the communication similar to the best known achievable communication.

1 Introduction

Decoupling is a fundamental tool for various protocols in classical and quantum information theory. Broadly, it refers to the process of applying some quantum operation on one of the two given systems (which share quantum correlation), so as to make the two systems independent of each other. This idea has been applied in various tasks such as quantum state merging [123456], quantum state redistribution [7891011].

∗Centre for Quantum Technologies, National University of Singapore, Singapore. a0109169@u.nus.edu
†Centre for Quantum Technologies, National University of Singapore; MajuLab, UMI 3654, Singapore and VAJRA Adjunct Faculty, TIFR, Mumbai, India. rahul@comp.nus.edu.sg
quantum channel coding \cite{12, 13, 14, 15, 16}, randomness extraction \cite{17, 18, 19}, quantum thermodynamics \cite{20} and black hole physics \cite{21, 22}. The central approach in many of these works is to perform a random unitary operation \cite{1, 2} and then discard a part of the system. This approach has been expanded upon in various works such as \cite{23, 24, 25}. Due to the importance of decoupling technique and the limitation that random unitaries cannot be implemented with a short quantum circuit, there is a great interest in finding efficient circuits for the purpose of decoupling. Some techniques involve replacing random unitaries with 2-designs \cite{26, 27, 28, 29} which can be simulated by Clifford circuits of small depth and random quantum circuits of small depth \cite{30}.

The classical analogue of decoupling method is the random binning approach of Slepian and Wolf \cite{31}, where a random function is applied to a sample from a source, in order to ‘concentrate’ its information content in as less number of registers as possible. The properties of a random function can also be reproduced by a random permutation of strings (followed by discarding some number of bits), as discussed in details in \cite{24}. The approach of random functions has important applications in the theory of randomness extractor \cite{32, 33, 34}, where the aim is to distill uniform bits from a non-uniform source $C$. While random functions extract maximum possible number of uniform bits from $C$, they require a large number of additional seed randomness and are not efficient. A well known solution is to replace random functions with a family of pairwise independent functions \cite{34}, which is the classical analogue of the aforementioned unitary 2-designs.

More precisely, one chooses positive integers $a, b$ uniformly at random from the set $\{1, 2, \ldots, p\}$ (for some integer $p$ which is usually a sufficiently small prime), and applies a function $f_{a,b}$ on the sample $c$ from $C$ to output $f_{a,b}(c)$. Averaged over $a$ and $b$, this process mimics the action of a random function. For most applications, it suffices to choose $f_{a,b}$ to be a linear function, making its implementation efficient. It has been shown \cite{17, 18} that pairwise independent functions also extract randomness if the source $C$ is correlated with a quantum system $R$, through a classical-quantum state $\Psi_{RC}$. The number of uniform bits that can be extracted is optimally characterized by the conditional min entropy $H_{\min}(C|R)_\Psi$. Further improvements in the seed size have been obtained in \cite{35}, based on Trevisan’s construction \cite{36}.

It is natural to ask how well do random functions/permutations (or their pairwise independent counterparts) behave when applied to a quantum register $C$ having arbitrary quantum correlations with $R$. This action can be implemented, for example, by choosing a basis (preferably the computational basis on register $C$) and applying permutations that take basis vectors to basis vectors. One would expect that such ‘classical’ operations may fail to remove the quantum correlations between registers $C$ and $R$, owing to the quantum coherence introduced by the off-diagonal terms. Indeed, this limitation is observed in the analysis of decoupling capacity of random permutations \cite{24}, Chapter 7, where guarantees on decoupling are obtained only when $H_{\min}(C|R)_\Psi > 0$. This can be taken as an indication that fully quantum techniques, such as random unitaries or unitary 2-designs, might be necessary for decoupling an arbitrary quantum state $\Psi_{RC}$.

This is shown not to be true by the method of convex-split \cite{6}, which states that adding the quantum state $\sigma_{C_1} \otimes \ldots \otimes \sigma_{C_N}$ (for some large enough $N$) and randomly swapping register $C$ with one of the registers $C_1, \ldots, C_N$ makes the register $R$ independent of all the other registers. The process of swapping two registers is a ‘classical’ operation (that is, it takes basis vectors to basis vectors). Unfortunately, the value of $N$ can be as large as $O(|C|)$, and hence swapping register $C$ with a random register $C_i$ requires a circuit of depth $O(|C|)$, which is exponential in the number of qubits of register $C$. Even an alternate implementation of swap operation, by placing the registers on a three dimensional grid, would require $O(|C|^{1/3})$ operations. Thus, it has so far been unknown if one can achieve quantum decoupling by efficient ‘classical’ operations.

Recent works have shown several applications of the convex-split method in one-shot quantum information theory, along with the dual method of position-based decoding \cite{37}. The methods have been used to
obtain near-optimal communication for one-shot entanglement-assisted quantum channel coding [37], near-optimal communication for one-shot quantum state splitting [6] (with slight improvement of the additive $\log \log |C|$ factor over [5], for communicating the register $C$) and smallest known communication for one-shot quantum state redistribution [11]. As mentioned earlier, all these protocols use a large amount of entanglement. Other known protocols ([38, 15, 16] for entanglement-assisted quantum channel coding and [10, 9] for quantum state redistribution) that do not rely on these two methods use exponentially small entanglement, but their communication is not known to be near-optimal. Thus, it is an important problem to find a scheme that achieves the best of both of the lines of work.

Our results

Our main result is the introduction of two new techniques to improve the resources such as entanglement and circuit depth in quantum protocols. These technique can be viewed as generic schemes for exponentially improving the resource required in the convex-split and position-based decoding methods. We summarize the results in Figure 1.

Efficient decoupling procedures: The quantity of interest in a decoupling procedure is the number of bits or qubits that are discarded to achieve the decoupling. There are two models under which decoupling is performed. The first model involves adding a quantum state, applying a global unitary (without involving the register $R$) and then discarding some quantum system. The second model also involves adding a quantum state followed by a unitary, but the system that is discarded is classical and the unitary acts in a classical-quantum manner [39]. The two models can be converted into each other by a Clifford circuit of depth 1 and the number of qubits/bits discarded are the same up to a factor of 2, due to the well known duality between teleportation [40] and super-dense coding [41]. Additional quantum systems that are not discarded act as a catalyst for the decoupling process [5, 6, 42, 43, 44]. For example, the randomness used in the process of decoupling via unitary 2-design acts as a catalyst. This randomness can be fixed by standard derandomization arguments, but it leads to a loss in efficient implementation.

In this work, we consider the second model of decoupling. We construct two new convex-split lemmas (Theorems 1 and 2) which immediately lead to efficient decoupling procedures for a quantum state $\Psi_{RC}$. The second procedure solves the aforementioned problem of decoupling via an efficient classical operation.

- The convex-split lemma introduced in Theorem 1 shows how to achieve decoupling using random mixture of small number of Heisenberg-Weyl (HW) operators, which generalize the Pauli $X$ and $Z$ operators. We can also replace HW operators with tensor products of Pauli $X$ and $Z$ if the register $C$ admits a qubit decomposition. The additional shared randomness used to choose the HW operations is on $4 \log |C|$ bits. The idea for this construction comes from the duality between convex-split and position-based decoding. Rephrasing the entanglement-assisted quantum channel coding result of [38] (and its one-shot analogue from [16]) in terms of position-based decoding, we find that the dual convex-split result is as given in Theorem 1. The HW operators do not act in a classical manner, as they do not permute basis vectors among themselves.

- The decoupling procedure obtained in Theorem 2 and Corollary 1 is as follows. This procedure enlarges the Hilbert space $\mathcal{H}_C \otimes \mathcal{H}_C$ in a manner that the resulting Hilbert space $\mathcal{H}_G$ has prime dimension $|G| \leq 2|C|^2$. It also introduces a register $L$ of size approximately $N \overset{\text{def}}{=} \log |C| - H_{\min}(C|R)_\Psi$. A

\footnote{This is possible due to Bertrand’s postulate [45], which says that there is a prime between any natural number and its twice.}
preferred basis on $\mathcal{H}_C$ (such as the computational basis in the qubit representation of the registers) is chosen, which gives a basis $\{|i\rangle_G\}_{i=0}^{G-1}$ on $\mathcal{H}_G$. Similarly, a preferred basis $\{|\ell\rangle\}_1^N$ is chosen on $\mathcal{H}_L$. Following this, a unitary operation $U = \sum_{\ell=1}^N U_\ell \otimes |\ell\rangle_L$ is applied, where $U_\ell$ (formally defined in Definition 3) acts on two registers $G, G' \equiv G$ as

$$U_\ell |i\rangle_G |j\rangle_{G'} = |i + (j - i) \ell \pmod{|G|}\rangle_G |j + (j - i) \ell \pmod{|G|}\rangle_{G'}.$$  

Upon tracing out register $L$, register $R$ becomes independent of $GG'$ (Theorem 2). Furthermore, the final state on registers $GG'$ is maximally mixed and the register $G'$ is returned in the original state. As can be seen, the unitaries $U_\ell$ are ‘classical’ as they take basis vectors to basis vectors and perform addition and multiplication modulo $|G|$. This makes the construction of $U$ efficient, as discussed in Appendix A. These unitaries have another nice property that they act as a representation of the cyclic group (Lemma 2), reflecting the property of permutation operations in the convex-split method.

- In the language of resource theory of coherence, both the decoupling procedures constructed above belong to the class of Physically Incoherent Operations [45]. Thus, an immediate implication of Theorems 1 and 2 is that quantum decoupling can be performed by incoherent unitaries. The results in Theorems 1 and 2 perform the same as decoupling via random unitary [23, 18, 25], when we consider the size of discarded system. None of the three methods are optimal as the decoupled register $C$ may not be maximally mixed in a general procedure. Indeed, it is known that the optimum number of discarded qubits for decoupling is characterized by the max-mutual information, rather than the conditional min-entropy [5, 6, 42]. We discuss this further below.

**Exponential improvement in entanglement:** A flattening procedure, that extends a quantum state $\sigma_C$ to a quantum state maximally mixed in a subspace (Definition 5), was originally used in the context of classical correlated sampling in several works [47, 48, 49, 50, 51, 52, 53, 54]. We observe that this extension can be constructed in a unitary manner using embezzling states [55] (Section 4). This leads to the following consequences.

- We show how to achieve a near-optimal one-shot communication over a quantum channel $N_{A \rightarrow B}$ with $O(\log |A|)$ qubits of pre-shared entanglement (Theorem 7). We also use the protocol in [28] and its one-shot analogue in [16] as a subroutine. This resolves the open question posed in [37].
- In Theorem 4 we give a new protocol for achieving decoupling in terms of the max-mutual information, using $O(\log |C|)$ additional qubits. As a consequence, we obtain near-optimal communication cost for quantum state splitting (Corollary 5), with small amount of initial entanglement. This slightly improves upon the communication cost for the same task in [5], with similar amount of initial entanglement.
- We also give various analogues of position-based decoding in Theorems 5 and 6. Along with our convex-split results, this leads to new protocols for quantum state redistribution (Corollary 4). Furthermore, it exponentially improves upon the entanglement required in the protocol for quantum state redistribution in [11], without changing the communication cost.

A new property of the embezzling procedure that we use is that the state after the embezzlement is close to the desired state in max-relative entropy (which is stronger than closeness in fidelity, see Claims 1 and 2). This is crucial to all of our applications. A different method of using embezzling states to flatten a quantum state was given in [5] for quantum state splitting. It is not clear if their method can be used to obtain the above applications.
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 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.

2 Preliminaries

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 $\ell_1$ norm of an operator $X$ on $\mathcal{H}$ is $\|X\|_1 := \text{Tr} \sqrt{X^\dagger X}$ and $\ell_2$ norm is $\|X\|_2 := \sqrt{\text{Tr} XX^\dagger}$. 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 $\psi$ to represent the state and also the density matrix $|\psi\rangle \langle \psi|$, associated with $|\psi\rangle$. Given a quantum state $\rho$ on $\mathcal{H}$, support of $\rho$, called $\text{supp}(\rho)$ is the subspace of $\mathcal{H}$ spanned by all eigenvectors of $\rho$ with non-zero eigenvalues.

A quantum register $A$ is associated with some Hilbert space $\mathcal{H}_A$. Define $|A| := \text{dim}(\mathcal{H}_A)$. 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 $\rho$ 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 \( \rho \) and \( \sigma \). The identity operator on \( \mathcal{H}_A \) (and associated register \( A \)) is denoted \( I_A \). The maximally mixed state \( |\lambda_A\rangle \) on register \( A \) is represented by \( \mu_A \).

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

\[
\rho_B := \text{Tr}_A \rho_{AB} := \sum_i (|i\rangle \otimes \mathds{1}_B)\rho_{AB}(|i\rangle \otimes \mathds{1}_B),
\]

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 \( \rho_{AB} \). 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 \( \rho_A \) is a pure state \( \rho_{AB} \in \mathcal{D}(AB) \) such that \( \text{Tr}_B \rho_{AB} = \rho_A \). 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 \( \rho_A \) is \( (\rho_A^+ \otimes \mathds{1}_B)(\sum_i |i\rangle_A|i\rangle_B) \).

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 \( U_A^\dagger U_A = U_A U_A^\dagger = I_A \). An isometry \( V_{A \to B} : \mathcal{H}_A \to \mathcal{H}_B \) is such that \( V^\dagger V = I_A \) and \( V V^\dagger = I_B \). The set of all unitary operations on register \( A \) is denoted by \( \mathcal{U}(A) \). Some standard unitaries are the \( X, Z, H \) (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 \( [56] \). 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** \( [57] \), see also \( [58] \) For \( \rho_A, \sigma_A \in \mathcal{D}(A) \),

\[
F(\rho_A, \sigma_A) \overset{\text{def}}{=} \| \sqrt{\rho_A} \sqrt{\sigma_A} \|_1.
\]

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

\[
F(P, Q) \overset{\text{def}}{=} \sum_i \sqrt{p_i \cdot q_i}.
\]

2. **Purified distance** \( [59] \) For \( \rho_A, \sigma_A \in \mathcal{D}(A) \),

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

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

\[
\mathcal{B}_\varepsilon(\rho_A) \overset{\text{def}}{=} \{ \rho'_A \in \mathcal{D}(A) \mid P(\rho_A, \rho'_A) \leq \varepsilon \}.
\]

4. **Smooth max-relative entropy** \( [60] \), see also \( [61] \) For \( \rho_A, \sigma_A \in \mathcal{D}(A) \) such that \( \text{supp}(\rho_A) \subset \text{supp}(\sigma_A) \),

\[
D_{\text{max}}^\varepsilon(\rho_A \| \sigma_A) \overset{\text{def}}{=} \min_{\rho'_A \in \mathcal{B}_\varepsilon(\rho_A)} \min \{ \lambda \in \mathbb{R} : 2^\lambda \sigma_A \geq \rho'_A \}.
\]
5. Hypothesis testing relative entropy (62), see also (63) For \( \rho_A, \sigma_A \in \mathcal{D}(A) \),

\[
D_H^\varepsilon (\rho_A \| \sigma_A) \overset{\text{def}}{=} \max_{0 < \Pi < I, \Tr(\Pi \rho_A) \geq 1 - \varepsilon} \log \left( \frac{1}{\Tr(\Pi \sigma_A)} \right).
\]

6. Max-information (64) For \( \rho_{AB} \in \mathcal{D}(AB) \),

\[
I_{\max}^\rho (A : B) \overset{\text{def}}{=} D_{\max} (\rho_{AB} \| \rho_A \otimes \rho_B).
\]

7. Smooth max-information (64) For \( \rho_{AB} \in \mathcal{D}(AB) \),

\[
I_{\max}^\varepsilon (A : B) \overset{\text{def}}{=} D_{\max} (\rho_{AB} \| \rho_A \otimes \rho_B) \varepsilon.
\]

8. Conditional min-entropy (17) For \( \rho_{AB} \in \mathcal{D}(AB) \),

\[
H_{\min}^\rho (A | B) \overset{\text{def}}{=} - \min_{\sigma_B \in \mathcal{D}(B)} D_{\max} (\rho_{AB} \| I_A \otimes \sigma_B).
\]

9. Smooth conditional min-entropy (17) For \( \rho_{AB} \in \mathcal{D}(AB) \),

\[
H_{\min}^\varepsilon (A | B) \overset{\text{def}}{=} \max_{\rho' \in B(\rho)} H_{\min}^\rho (A | B).
\]

We will use the following facts.

**Fact 1** (Triangle inequality for purified distance 65). For states \( \rho_A, \sigma_A, \tau_A \in \mathcal{D}(A) \),

\[
P(\rho_A, \sigma_A) \leq P(\rho_A, \tau_A) + P(\tau_A, \sigma_A).
\]

**Fact 2** (Monotonicity under quantum operations 66, 67). For quantum states \( \rho, \sigma \in \mathcal{D}(A) \), and quantum operation \( E(\cdot) : \mathcal{L}(A) \to \mathcal{L}(B) \), it holds that

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

**Fact 3** (Uhlmann’s theorem 58). Let \( \rho_A, \sigma_A \in \mathcal{D}(A) \). Let \( \rho_{AB} \in \mathcal{D}(AB) \) be a purification of \( \rho_A \) and \( \sigma_{AC} \in \mathcal{D}(AC) \) be a purification of \( \sigma_A \). There exists an isometry \( V : C \to B \) such that,

\[
F(\vert \theta \rangle \langle \theta \vert_{AB}, \vert \rho \rangle \langle \rho \vert_{AB}) = F(\rho_A, \sigma_A),
\]

where \( \vert \theta \rangle_{AB} = (I_A \otimes V) \vert \sigma \rangle_{AC} \).

**Fact 4** (Gentle measurement lemma 68, 69). Let \( \rho \) be a quantum state and \( 0 < A < I \) be an operator. Then

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

Following fact implies the Pinsker’s inequality.
Fact 5 (Lemma 5 [70]). For quantum states $\rho_A, \sigma_A \in \mathcal{D}(A)$,
\[
F(\rho, \sigma) \geq 2^{-\frac{1}{2}D(\rho\|\sigma)}.
\]

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

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

Proof. We prove each item as follows. Let $\eta$ be chosen below.

- Given the quantum state $\omega_C$, we construct an operator $O$ by increasing each eigenvalue of $\omega_C$ to the nearest multiple of $\frac{\eta}{|C|}$, and define $\sigma_C \overset{\text{def}}{=} \frac{O}{\text{Tr}(O)}$. We have
  \[
  1 = \text{Tr}(\sigma_C) \leq \text{Tr}(O) \leq \text{Tr}(\sigma_C) + |C| \frac{\eta}{|C|} = 1 + \eta.
  \]
  Define $\eta' \overset{\text{def}}{=} \text{Tr}(O) - 1$ which implies $0 \leq \eta' \leq \eta$. The eigenvalues of $\sigma_C$ are integer multiples of $\frac{\eta'}{|1+\eta'|C|}$. We choose $\eta$ (which determines $\eta'$ as well) such that $\frac{\eta}{1+\eta} = \gamma$. This ensures that $\gamma \leq \eta \leq \frac{1}{\gamma}$. Furthermore, eigenvalues of $\sigma_C$ 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 $\omega_C$ to the nearest multiple of $\frac{\eta}{|C|}$, and define $\sigma_C \overset{\text{def}}{=} \frac{O}{\text{Tr}(O)}$. We have
  \[
  1 = \text{Tr}(\omega_C) \geq \text{Tr}(O) \geq \text{Tr}(\omega_C) - |C| \frac{\eta}{|C|} = 1 - \eta.
  \]
  Define $\eta' \overset{\text{def}}{=} 1 - \text{Tr}(O)$, which implies $0 \leq \eta' \leq \eta$. The eigenvalues of $\sigma_C$ are integer multiples of $\frac{\eta'}{|1-\eta'|C|}$. We choose $\eta$ (which determines $\eta'$ as well) such that $\frac{\eta}{1-\eta} = \gamma$. This ensures that $\frac{1}{1+\gamma} \leq \eta \leq \gamma$. Furthermore, eigenvalues of $\sigma_C$ 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. \qed

Fact 7 (37). Let $\rho_A, \sigma_A \in \mathcal{D}(\mathcal{H}_A)$ be quantum states. Let $\Lambda \in \mathcal{L}(\mathcal{H}_A)$, $0 \leq \Lambda \preceq I_A$ be a positive semidefinite operator. Then it holds that
\[
|\sqrt{\text{Tr}(A\rho_A)} - \sqrt{\text{Tr}(A\sigma_A)}| \leq P(\rho_A, \sigma_A).
\]
Fact 8 ([53]). 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),

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

Fact 9 ([6]). Let $\rho_1, \ldots, \rho_n, \theta$ be quantum states and $\{p_i\}_i$ be a probability distribution. Define $\rho \overset{\text{def}}{=} \sum_i p_i \rho_i$. Then it holds that

$$D \left( \sum_i p_i \rho_i \parallel \theta \right) = \sum_i p_i \left( D(\rho_i \parallel \theta) - D(\rho_i \parallel \rho) \right).$$

Fact 10 (Hayashi-Nagaoka inequality [63]). 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

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

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

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

Fact 11 (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

$$(U \otimes 1_{C'}) |\Phi\rangle_{CC'} = (1_{C} \otimes U^T) |\Phi\rangle_{CC'}.$$
Fact 14. It holds that for $0 < a < C$,
\[
\sum_{b=0}^{C-1} e^{\frac{2\pi i a b}{c}} = 0.
\]

Proof. Let $S \overset{\text{def}}{=} \sum_{b=0}^{C-1} e^{\frac{2\pi i a b}{c}}$. We have
\[
e^{\frac{2\pi i a}{c}} S = \sum_{b=0}^{C-1} e^{\frac{2\pi i a (b+1)}{c}} = \sum_{b=1}^{C-1} e^{\frac{2\pi i a b}{c}} + e^{\frac{2\pi i a C}{c}} = 1 + \sum_{b=1}^{C-1} e^{\frac{2\pi i a b}{c}} = S.
\]
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 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.

3.1 Convex-split using a mixture of HW operators

Definition 1. Given a register $C$ and a basis $\{|c\rangle\}_{c=0}^{|C|-1}$. Define the Heisenberg-Weyl (HW) unitaries $\{V_{a,b}\}_{a,b=0}^{Z|-1}$ with $V_{a,b} : \mathcal{H}_C \rightarrow \mathcal{H}_C$ as $V_{a,b} \overset{\text{def}}{=} \sum_c e^{\frac{2\pi i a b}{|c|}} |c+a\rangle_Z \langle c|_Z$.

Following is a well known lemma.

Lemma 1. For all $|c\rangle_C$, $|c'\rangle_C$, it holds that
\[
\frac{1}{|C|^2} \sum_{a,b} V_{a,b} |c\rangle_C \langle c'|_C V_{a,b}^\dagger = \delta_{c,c'} \mu_C.
\]

In particular, this implies that for any state $\rho_{RC}$,
\[
\frac{1}{|C|^2} \sum_{a,b} V_{a,b} \rho_{RC} V_{a,b}^\dagger = \rho_R \otimes \mu_C.
\]

Proof. Consider
\[
\frac{1}{|C|^2} \sum_{a,b} V_{a,b} |c\rangle_C \langle c'|_C V_{a,b}^\dagger = \left( \frac{1}{|C|^2} \sum_{a,b} e^{\frac{2\pi i (c-c') a b}{|c|}} |c+a\rangle \langle c'+a|_C \right) \\
= \delta_{c,c'} \left( \frac{1}{|C|} \sum_a |c+a\rangle \langle c+a|_C \right) \\
= \delta_{c,c'} \mu_Z.
\]
In the second equation, we have used Fact\textsuperscript{14} Expand \( \rho_{RC} = \sum_{c,c'} \rho_{RC}^{c,c'} \otimes \ket{c} \bra{c'}_C \). Consider

\[
\frac{1}{|C|^2} \sum_{a,b} V_{a,b} \rho_{RC} V_{a,b}^\dagger = \sum_{c,c'} \rho_{RC}^{c,c'} \otimes \left( \frac{1}{|C|^2} \sum_{a,b} V_{a,b} \ket{c} \bra{c'}_C V_{a,b}^\dagger \right) = \sum_{c,c'} \delta_{c,c'} \rho_{RC}^{c,c} \otimes \mu_C = \rho_R \otimes \mu_C.
\]

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

For the ease of notation, we represent \((a, b)\) as \(x\) and let \(X\) be the set of all \(x\).

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

\[
\frac{1}{|X|^2} \sum_{x,x'} \delta_{x,x'} = 1, \quad \forall x,x', \quad \forall j \neq k.
\]

Introduce registers \(X_1 \equiv X_2\) such that \(|X_1| = |X_2| = |X|\). Let \(V^{(j)} : \mathcal{H}_{CX_1X_2} \to \mathcal{H}_{CX_1X_2}\) be defined as

\[
V^{(j)} = \sum_{x_1,x_2} V_{f_j(x_1,x_2)} \otimes \ket{x_1,x_2} \bra{x_1,x_2}_{X_1X_2}.
\]

As discussed in [72, Example 6] or [73], there exists an efficient construction of pairwise independent function family for any \(X\) with \(|X|\) a prime power. In our setting, \(|X| = |C|^2\). 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.

\textbf{Theorem 1.} Suppose \(\log |C|\) is an integer. Let \(\Psi_{RC}\) be a quantum state. Define \(k \overset{\text{def}}{=} D_{\max}(\Psi_{RC} \| \Psi_R \otimes \mu_C)\). Define the quantum state

\[
\tau_j \overset{\text{def}}{=} V^{(j)}(\Psi_{RC} \otimes \mu_{X_1X_2}) V^{(j)\dagger}, \quad \tau \overset{\text{def}}{=} \frac{1}{N} \sum_j \tau_j.
\]

It holds that

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

\textbf{Proof.} We now proceed to the desired inequality. Observe that \(V^{(j)}(\cdot)\) acts controlled on registers \(X_1, X_2\). Thus,

\[
D(\tau \| \Psi_R \otimes \mu_C \otimes \mu_{X_1X_2}) = \frac{1}{|X|^2} \sum_{x_1,x_2} D \left( \frac{1}{N} \sum_j V_{f_j(x_1,x_2)} \Psi_{RC} V_{f_j(x_1,x_2)}^\dagger \left\| \Psi_R \otimes \mu_C \right\| \right).
\]
Using Fact 9 we have

\[
D \left( \frac{1}{N} \sum_j V_{f_j(x_1,x_2)} \Psi \bigg| \bigg. \Psi R \otimes \mu \bigg) \\
= \frac{1}{N} \sum_j \left( D \left( V_{f_j(x_1,x_2)} \Psi \bigg| \bigg. \Psi R \otimes \mu \bigg) \\
- D \left( \Psi \bigg| \bigg. \Psi R \otimes \mu \bigg) \right) \right)
\]

Thus,

\[
D(\tau \parallel \Psi R \otimes \mu \otimes \mu X_1 X_2) = D(\Psi R \parallel \Psi R \otimes \mu)
\]

\[
- \frac{1}{N|x|^2} \sum_{x_1} \sum_{x_2} \left( D \left( \Psi \bigg| \bigg. \Psi R + \frac{1}{N} \sum_k V_{f_k(x_1,x_2)} \Psi \right) \right)
\]

where we have used the convexity of relative entropy. From the pairwise independent property of the family of functions, this simplifies to

\[
D(\tau \parallel \Psi R \otimes \mu \otimes \mu X_1 X_2) \leq D(\Psi R \parallel \Psi R \otimes \mu)
\]

\[
- \frac{1}{N} \sum_j \left( D \left( \Psi \bigg| \bigg. \Psi R + \frac{1}{N} \sum_{k \neq j} V_{f_j(x_1,x_2)} \Psi \right) \right)
\]

\[
= D(\Psi R \parallel \Psi R \otimes \mu) - \frac{1}{N} \sum_j \left( D \left( \Psi \bigg| \bigg. \Psi R + \frac{1}{N} \sum_{k \neq j} V_{f_j(x_1,x_2)} \Psi \right) \right) .
\]
Lemma 1 ensures that
\[
D(\tau \| \Psi_R \otimes \mu_C \otimes \mu_{X_1X_2}) \leq D(\Psi_{RC} \| \Psi_R \otimes \mu_C)
\]
\[
- \frac{1}{N} \sum_j \left( D\left( \Psi_{RC} \left\| \frac{1}{N} \Psi_{RC} + \frac{1}{N} \sum_{k\neq j}^N V_x^j V_x^j \Psi_{RC} V_x^j V_x^j \right\| \right)
\]
\[
= D(\Psi_{RC} \| \Psi_R \otimes \mu_C) - \frac{1}{N} \sum_j \left( D\left( \Psi_{RC} \left\| \frac{1}{N} \Psi_{RC} + \frac{1}{N} \sum_{k\neq j}^N \Psi_R \otimes \mu_C \right\| \right)
\]
\[
= D(\Psi_{RC} \| \Psi_R \otimes \mu_C) - D\left( \Psi_{RC} \left\| \frac{1}{N} \Psi_{RC} + \frac{N-1}{N} \Psi_R \otimes \mu_C \right\| \right).
\]

Using the inequality \( \Psi_{RC} \leq 2^k \Psi_R \otimes \mu_C \) and the operator monotonicity of logarithm [74], we conclude that
\[
D(\tau \| \Psi_R \otimes \mu_C \otimes \mu_{X_1X_2}) \quad \leq \quad D(\Psi_{RC} \| \Psi_R \otimes \mu_C) - D(\Psi_{RC} \| \Psi_R \otimes \mu_C) + \log \left( 1 + \frac{2^k - 1}{N} \right)
\]
\[
= \log \left( 1 + \frac{2^k - 1}{N} \right).
\]

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.

### 3.2 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 \(|G| \geq |C|^2\) 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 [45].

Let \( \{|c\rangle\}_{c=0}^{|G|-1} \) 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 \( |0\rangle|c\rangle|c'\rangle \) as \(|c|C\rangle + |c'\rangle \). This gives \(|C|^2\) basis vectors for \( G \). The remaining \(|G| - |C|^2\) basis vectors are constructed by relabeling \( |1\rangle|c\rangle|c'\rangle \) as \(|C|^2 + c|C\rangle + |c'\rangle \) as long as \(|C|^2 + c|C\rangle + |c'\rangle \leq |G| - 1 \). We note that the constraint \(|C|^2 + c|C\rangle + |c'\rangle \leq |G| - 1 \) 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 3.** For an integer \( \ell \in \{0, 1, \ldots |G| - 1\} \), 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:
\[
U_{\ell} \defeq \sum_{i,j} |i + (j - i)\ell\rangle_{G_1}|j + (j - i)\ell\rangle_{G_2} \langle i|_{G_1} \langle j|_{G_2}.
\]

We choose the convention that the expression in the kets for registers \( G_1, G_2 \) are evaluated modulo \(|G|\).
Lemma 2. For every $m, \ell \in \{0, 1, \ldots |G| - 1\}$, it holds that $U_\ell$ is a unitary. Furthermore

$$U_m U_\ell = U_{m + \ell}, \quad U_\ell^\dagger = U_{-\ell}. \tag{1}$$

Proof. We first show that $U_\ell$ 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 $\ell$, 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, $U_\ell$ is a unitary.

Now, consider

$$U_m U_\ell |i\rangle_{G_1} |j\rangle_{G_2} = U_m |i + (j - i)\ell\rangle_{G_1} |j + (j - i)\ell\rangle_{G_2}$$

$$= |i + (j - i)\ell + (j - i)m\rangle_{G_1} |j + (j - i)\ell + (j - i)m\rangle_{G_2}$$

$$= |i + (j - i)(\ell + m)\rangle_{G_1} |j + (j - i)(\ell + m)\rangle_{G_2}$$

$$= U_{m + \ell} |i\rangle_{G_1} |j\rangle_{G_2}.$$

Thus, $U_m U_\ell = U_{m + \ell}$. Since $U_0 = 1$, we conclude $U_\ell^\dagger = U_{-\ell}$. This completes the proof. \hfill \Box

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

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

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

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

$$\text{Tr}_{G_2} \left(U_m (|i\rangle \langle i'|_{G_1} \otimes \mu_{G_2}) U_m^\dagger \right)$$

$$= \frac{1}{|G|} \sum_{j = 0}^{|G| - 1} \text{Tr}_{G_2} \left(U_m (|i\rangle \langle i'|_{G_1} \otimes |j\rangle \langle j|_{G_2}) U_m^\dagger \right)$$

$$= \frac{1}{|G|} \sum_{j = 0}^{|G| - 1} \text{Tr}_{G_2} \left(|jm + i(1 - m)\rangle \langle jm + i'(1 - m)|_{G_1} \otimes |j(m + 1) - im\rangle \langle j(m + 1) - i'm|_{G_2} \right)$$

$$= \frac{1}{|G|} \sum_{j = 0}^{|G| - 1} |jm + i(1 - m)\rangle \langle jm + i'(1 - m)|_{G_1} \cdot \delta_{i,i'}$$

$$= \frac{1}{|G|} \sum_{j = 0}^{|G| - 1} |jm + i(1 - m)\rangle \langle jm + i'(1 - m)|_{G_1} \cdot \delta_{i,i'}$$

$$= \mu_{G_1} \cdot \delta_{i,i'} \tag{2}.$$
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 $\Psi_R = \text{Tr}_{C_0}(\Psi_{RC_0}) = \sum_c \Psi_R^{(c,c)}$. For any $m > 0$, using Equation 2 we have

$$\text{Tr}_{G_2}(U_m (\Psi_{RC_0} \otimes |0\rangle\langle 0|_Q \otimes \mu_{C_1} \otimes \mu_{G_2}) U_m^\dagger)$$

$$= \sum_{c,c',c_1} \frac{1}{|C|} \Psi_R^{(c,c')} \otimes \text{Tr}_{G_2}(U_m (|c\rangle\langle c'|_{C_0} \otimes |0\rangle\langle 0|_Q \otimes |c_1\rangle\langle c_1|_{C_1} \otimes \mu_{G_2}) U_m^\dagger)$$

$$= \sum_{c,c',c_1} \frac{1}{|C|} \Psi_R^{(c,c')} \otimes \mu_{G_1} \cdot \delta_{c,c'}$$

$$= \sum_c \Psi_R^{(c,c)} \otimes \mu_{G_1} = \Psi_R \otimes \mu_{G_1},$$

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

Now, we are in a position to prove our main result.

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

$$\tau_{RG_1 G_2} \overset{\text{def}}{=} \frac{1}{N} \sum_{\ell \in S} U_{\ell} (\Psi_{RC_0} \otimes |0\rangle\langle 0|_Q \otimes \mu_{C_1} \otimes \mu_{G_2}) U_{\ell}^\dagger.$$

It holds that

$$D(\tau_{RG_1 G_2}||\Psi_R \otimes \mu_{G_1} \otimes \mu_{G_2}) \leq \log \left(1 + \frac{2^{k+1} - 1}{N}\right).$$

From Fact 5 we conclude that

$$F^2(\tau_{RG_1 G_2}, \Psi_R \otimes \mu_{G_1} \otimes \mu_{G_2}) \geq \frac{1}{1 + \frac{2^{k+1} - 1}{N}}.$$

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

$$\Psi_{RC_0} \otimes |0\rangle\langle 0|_Q \otimes \mu_{C_1} \preceq 2^k \Psi_R \otimes \mu_{C_0} \otimes |0\rangle\langle 0|_Q \otimes \mu_{C_1} \preceq \frac{|G_1|}{|C_0||C_1|}2^k \Psi_R \otimes \mu_{G_1} \preceq 2^{k+1} \Psi_R \otimes \mu_{G_1}. \ (3)$$

15
Using Fact 9 we have
\[
D(\tau_{RG_1G_2}) \leq \frac{1}{N} \sum_{m \in S} \left( D\left( U_{\ell} (\Psi_{RC_0} \otimes |0\rangle \langle 0|_Q \otimes \mu_{C_1} \otimes \mu_{G_2}) U_{\ell}^{\dagger} \right) \right)
\]
Since logarithm is operator monotone [74],
\[
D\left( \tau_{RG_1G_2} \right) \leq \frac{2^{k+1}}{N} \Psi_R \otimes \mu_{G_1} + \frac{N-1}{N} \Psi_R \otimes \mu_{G_1} = \left( 1 + \frac{2^{k+1} - 1}{N} \right) \Psi_R \otimes \mu_{G_1}.
\]
Since \( \mu_{G_1} \otimes \mu_{G_2} \) is maximally mixed in the support of \( U_{\ell} \),
\[
U_{\ell}^{\dagger} (\mu_{G_1} \otimes \mu_{G_2}) U_{\ell} = \mu_{G_1} \otimes \mu_{G_2}. \tag{5}
\]
Moreover, from Lemma 2 we have
\[
\text{Tr}_{G_2} \left( U_{\ell}^{\dagger} \tau_{RG_1G_2} U_{\ell} \right) \leq \frac{2^{k+1}}{N} \Psi_R \otimes \mu_{G_1} + \frac{N-1}{N} \Psi_R \otimes \mu_{G_1} = \left( 1 + \frac{2^{k+1} - 1}{N} \right) \Psi_R \otimes \mu_{G_1}.
\]
where in last operator inequality, we have used Equation 3. Using Lemma 3, we conclude that
\[
\text{Tr}_{G_2} \left( U_{\ell}^{\dagger} \tau_{RG_1G_2} U_{\ell} \right) \leq \frac{2^{k+1}}{N} \Psi_R \otimes \mu_{G_1} + \frac{N-1}{N} \Psi_R \otimes \mu_{G_1} = \left( 1 + \frac{2^{k+1} - 1}{N} \right) \Psi_R \otimes \mu_{G_1}.
\]
Since logarithm is operator monotone [74],
\[
D\left( \Psi_{RC_0} \otimes |0\rangle \langle 0|_Q \otimes \mu_{C_1} \right) \leq \text{Tr}_{G_2} \left( U_{\ell}^{\dagger} \tau_{RG_1G_2} U_{\ell} \right) \geq \log \left( 1 + \frac{2^k - 1}{N} \right) + D\left( \Psi_{RC_0} \otimes |0\rangle \langle 0|_Q \otimes \mu_{C_1} \right) \Psi_R \otimes \mu_{G_1}.)
\]
From Equations $4$ and $5$:

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

This completes the proof.

An immediate corollary is the smooth version of above result.

**Corollary 1.** Let $\varepsilon, \delta \in (0, 1)$ and $\Psi_R$ be a quantum state. Let $k \overset{\text{def}}{=} \log |C| - H_{\text{min}}^R(C|R)_\Psi + \log \frac{8}{\varepsilon^2}$ 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

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

It holds that

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

**Proof.** From Fact $13$ we conclude that

$$
\min_{\Psi'_R \in \mathcal{B}(\Psi_R)} D_{\text{max}}(\Psi'_R \parallel \Psi'_R \otimes \mu_C) \leq \log |C| - H_{\text{min}}^R(C|R)_\Psi + \log \frac{8}{\varepsilon^2} = k'.
$$

Let $\Psi'_R$ be the quantum state achieving the infimum above. Define

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

We use Theorem $2$ to conclude that

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

By triangle inequality for purified distance, this implies that

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

This concludes the proof. \(\square\)
Decoupling up to the max-mutual information using a flattening procedure

We introduce a close variant of the embezzling state \[55\].

**Definition 4.** Let \( a, n \) be positive integers such that \( n \geq a \) and let \( D \) be a register satisfying \( |D| \geq n - a \). Define

\[
\xi^{a:n}_D \overset{\text{def}}{=} \frac{1}{S(a, n)} \sum_{j=a}^{n} \frac{1}{j} |j\rangle\langle j|_D,
\]

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

\[
|\xi^{a:n}_D\rangle_D' \overset{\text{def}}{=} \frac{1}{\sqrt{S(a : n)}} \sum_{j=a}^{n} \frac{1}{\sqrt{j}} |j\rangle_D'\langle j|_D
\]

as the canonical purification of \( \xi^{a:n}_D \), where \( D' \equiv D \).

We have the following claim, which is a variant of the property of embezzling states proved in \[55\].

**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

\[
W_b |j\rangle_D |0\rangle_E = |\lfloor j/b \rfloor\rangle_D |j \text{ (mod } b)\rangle_E.
\]

It holds that

\[
W_b (\xi^{a:n}_D \otimes |0\rangle\langle 0|_E) W_b^\dagger \preceq (1 + 15\delta) \xi^{1:n}_D \otimes \frac{1}{b} \sum_{e=0}^{b-1} |e\rangle\langle e|_E.
\]

**Proof.** Consider

\[
W_b (\xi^{a:n}_D \otimes |0\rangle\langle 0|_E) W_b^\dagger
\]

\[
= \frac{1}{S(a, n)} \sum_{j=a}^{n} \frac{1}{j} W_b (|j\rangle\langle j|_D \otimes |0\rangle\langle 0|_E) W_b^\dagger
\]

\[
= \frac{1}{S(a, n)} \sum_{j=a}^{n} \frac{1}{j} |\lfloor j/b \rfloor\rangle (|\lfloor j/b \rfloor\rangle_\text{mod } b) \langle j \text{ (mod } b)|_E
\]

\[
= \frac{1}{S(a, n)} \sum_{j'=\lfloor \frac{j}{b} \rfloor}^{\lfloor \frac{n}{b} \rfloor} \sum_{e=0}^{b-1} \frac{1}{b j'} |j'|_{D} \otimes |e\rangle\langle e|_E
\]

\[
\preceq \frac{1}{S(a, n)} \sum_{j'=\lfloor \frac{n}{b} \rfloor}^{\lfloor \frac{n}{b} \rfloor} \sum_{e=0}^{b-1} \frac{1}{b j'} |j'|_D \otimes |e\rangle\langle e|_E
\]

\[
= \frac{1}{S(a, n)} \sum_{j'=\lfloor \frac{n}{b} \rfloor}^{\lfloor \frac{n}{b} \rfloor} \sum_{e=0}^{b-1} \frac{1}{b j'} |j'|_D \otimes |e\rangle\langle e|_E \leq \frac{S(1, n)}{S(a, n)} \xi^{1:n}_D \otimes \sum_{e=0}^{b-1} \frac{1}{b} |e\rangle\langle e|_E.
\]
Now, as shown in [75], \(|S(a, n) - \log \frac{n}{a}| \leq 4\). 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. \(\blacksquare\)

Following claim shows how to ‘unembezzle’ a state.

**Claim 2.** Fix the integers \(n, b, a\) as given in Claim 7. Let the register \(D\) satisfy \(n^2 \geq |D| \geq (n + 1)b\). Let \(W_b\) be as defined in Claim 7. It holds that

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

**Proof.** We observe that \(W_b \uparrow |j\rangle_D |e\rangle_E = |jb + e\rangle_D |0\rangle_E\) for all \(j \leq n\) and \(e < b\). We leave the action of \(W_b\) unspecified for \(j \geq n, e \geq b\). Consider

\[
W_b \uparrow \left( \frac{\xi_D^{1:n}}{b} \sum_{e=0}^{b-1} |e\rangle \langle e| \right) W_b = \frac{1}{S(1, n)} \sum_{j=1}^{n-1} \sum_{e=0}^{b-1} \frac{1}{b} W_b \uparrow |j\rangle_D |e\rangle_D \langle e|_E W_b
\]

\[
= \frac{1}{S(1, n)} \sum_{j=1}^{n-1} \sum_{e=0}^{b-1} |jb + e\rangle \langle jb + e|_D \otimes |0\rangle \langle 0|_E
\]

\[
\leq \frac{2}{S(1, n)} \sum_{j=1}^{n-1} \sum_{e=0}^{b-1} \frac{1}{b} |jb + e\rangle \langle jb + e|_D \otimes |0\rangle \langle 0|_E
\]

\[
\leq \frac{2}{S(1, n)} \sum_{j'=1}^{n+b} \frac{1}{j'} |j'\rangle \langle j'|_D \otimes |0\rangle \langle 0|_E
\]

\[
\leq \frac{2S(1, |D|)}{S(1, n)} \xi_D^{1:|D|} \otimes |0\rangle \langle 0|_E
\]

\[
\leq 4 \xi_D^{1:|D|} \otimes |0\rangle \langle 0|_E,
\]

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

A ‘purified version’ of above claims is the following restatement of the result in [55].

**Claim 3.** Let \(\delta \in (0, \frac{1}{\sqrt{2}})\). 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 \(|\mu\rangle_{E' E}\) be defined by \(\frac{1}{\sqrt{b}} \sum_{e=0}^{b-1} |e\rangle_{E'} \langle e|_E\). It holds that

\[
P \left( (W_b \otimes W_b) \left( \xi_D^{1:n} \otimes |0\rangle \langle 0|_E \otimes |0\rangle \langle 0|_E \right) (W_b \uparrow \otimes W_b \uparrow) , \xi_D^{1:n} \otimes |0\rangle \langle 0|_{E' E'} \right) \leq 5\sqrt{\delta}.
\]
\textbf{Proof.} We have
\[
(W_b \otimes W_b) |\xi_{a:n}^{D:D'}\rangle_{D' \otimes |0\rangle_{E'} \otimes |0\rangle_E} = \frac{1}{\sqrt{S(a:n)}} \sum_{j=a}^{n} \frac{1}{\sqrt{b}} |j/b\rangle_{D'} |j/b\rangle_{D} |j \pmod{b}\rangle_{E'} |j \pmod{b}\rangle_{E}
\]
\[
= \frac{1}{\sqrt{S(a:n)}} \sum_{j'= [a/b]}^{n/b} \sum_{e=0}^{b-1} \frac{1}{\sqrt{b} j' + e} |j', j'\rangle_{D' \otimes D} |e, e\rangle_{E' \otimes E}.
\]
Since $|\xi_{1:n}^{D:D'}\rangle_{D' \otimes |0\rangle_{E'} \otimes |0\rangle_E}$, we have
\[
F \left( (W_b \otimes W_b)(|\xi_{a:n}^{D:D'}\rangle \otimes |0\rangle_{E'} \otimes |0\rangle_{E}) (W_b \otimes W_b)^\dagger, |\xi_{1:n}^{D:D'}\rangle \otimes \mu \right)
\]
\[
= \frac{1}{\sqrt{S(a:n)S(1:n)}} \sum_{j'= [a/b]}^{n/b} \sum_{e=0}^{b-1} \frac{1}{\sqrt{b} j' + e}
\]
\[
\geq \frac{\sqrt{1 - \delta}}{\sqrt{S(a:n)S(1:n)}} \sum_{j'= [a/b]}^{n/b} \sum_{e=0}^{b-1} \frac{1}{\sqrt{b} j'}
\]
\[
= \sqrt{1 - \delta} \cdot \frac{S([a/b], [n/b])}{\sqrt{S(a:n)S(1:n)}} \geq \sqrt{1 - 25\delta},
\]
where we use the fact that $|S(a, n) - \log \frac{a}{b}| \leq 4$. This completes the proof. \hfill \square

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

\textbf{Definition 5. Flattening a quantum state:} Fix a $\gamma \in (0, 1)$ such that $\frac{|C|}{\gamma}$ is an integer and a quantum state $\sigma_{C} \overset{\text{def}}{=} \sum_{c} q(c) |c\rangle \langle c|_{C}$ 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 $\sigma_{CE}$ as follows:
\[
\sigma_{CE} \overset{\text{def}}{=} \sum_{c} q(c) |c\rangle \langle c|_{C} \otimes \left( \frac{\gamma}{q(c)|C|} \sum_{e=0}^{\frac{|c||C|}{\gamma} - 1} |e\rangle \langle e|_{E} \right) = \frac{\gamma}{|C|} \sum_{c} \sum_{e=0}^{\frac{|c||C|}{\gamma} - 1} |c\rangle \langle c|_{C} \otimes |e\rangle \langle e|_{E}.
\]
Observe that $\sigma_{CE}$ is uniform in its support.

The flattening of $\sigma_{C}$ can be realized in a unitary manner as follows. We define some registers and unitaries required for this process.

\textbf{Definition 6.} Fix $\delta \in (0, \frac{1}{10})$. Let $a = |E| = \frac{|C|}{\gamma} \max_{c} q(c)$ and $n = a^{\frac{1}{\delta}}$. Introduce a register $D$ satisfying $|D| \geq n$ with the quantum state $|\xi_{a:n}^{D:D'}\rangle$ as given in Definition 4. Define the unitary $W : \mathcal{H}_{CED} \to \mathcal{H}_{CED}$ as
\[
W \overset{\text{def}}{=} \sum_{c} |c\rangle \langle c|_{C} \otimes W_{a|c||C|}^{\nu(c)|C|},
\]
where $q(c), \gamma$ are given in Definition 5 and the unitary $W_{a|c||C|}^{\nu(c)|C|}$ is defined in Claim 7.

20
Flattening is ensured via the following relation, which uses Claim 1

\[ W (\sigma_C | 0 \rangle | E \rangle \otimes \xi_D^{a,n}) W^\dagger = \sum_c q(c) | c \rangle \langle c | \otimes W_{\frac{|c|}{\gamma}} (| 0 \rangle | E \rangle \otimes \xi_D^{a,n}) W_{\frac{|c|}{\gamma}}^\dagger \preceq (1 + 15\delta)\sigma_{CE} \otimes \xi_D^{1,n}. \]  

(6)

Given the flattening of a quantum state \(\sigma_C\), Definition 1 gives us \(|\text{supp}(\sigma_{CE})|\) units of \(V_\epsilon : \text{supp}(\sigma_{CE}) \to \text{supp}(\sigma_{CE})\). If \(|\text{supp}(\sigma_{CE})|^2 = \left(\frac{|c|}{\gamma}\right)^2\) is a prime power, Definition 2 gives us the collection of units \(V^\ell : \text{supp}(\sigma_{CE}) \otimes H_{X_1 X_2} \to \text{supp}(\sigma_{CE}) \otimes H_{X_1 X_2}\), with \(\log |X_1| = \log |X_2| = \log \left(\frac{|c|}{\gamma}\right)^2\). This allows us to construct the quantum states

\[ \tau_\ell = V^{(\ell)} \left( W (\Psi_{RC} \otimes |0\rangle \langle 0 | \otimes \xi_D^{a,n}) W^\dagger \otimes \mu_{X_1 X_2} \right) V^{(\ell)^\dagger}, \]  

(7)

Now we prove the following theorem, which is the analogue of Theorem 1 for a flattened quantum state.

**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 \(\Psi_{RC}, \omega_C\). Let \(k \stackrel{\text{def}}{=} D_{\max}(\Psi_{RC} || \sigma_R \otimes \omega_C)\) and \(N\) be an integer. Let \(\sigma_C\) be the quantum state as constructed in the first part of Fact 6 using \(\omega_C\). For quantum states \(\tau_\ell\) as given in Equation 7 define

\[ \tau = \frac{1}{N} \sum_\ell \tau_\ell. \]

It holds that

\[ D(\tau || \sigma_R \otimes \sigma_{CE} \otimes \xi_D^{1,n} \otimes \mu_{X_1 X_2}) \leq 15\delta + \log \left(1 + \frac{2^{k+2} - 1}{N}\right). \]

Since one can choose \(\log |D| = \log n \leq \frac{1}{\delta} \log \frac{|c|}{\gamma}\), the number of qubits of additional registers is \(\log |D| + \log |E| + 2 \log |X_1| \leq (4 + \frac{3}{4}) \log \frac{|c|}{\gamma}\).

**Proof.** From Fact 6 we have that the eigenvalues of \(\sigma_C\) are integer multiples of \(\frac{\gamma}{|c|}\) and

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

Consider,

\[ W (\Psi_{RC} \otimes |0\rangle \langle 0 | \otimes \xi_D^{a,n}) W^\dagger \preceq 2^{k+1} \Psi_R \otimes W (\sigma_C \otimes |0\rangle \langle 0 | \otimes \xi_D^{a,n}) W^\dagger \]

\[ \preceq 2^{k+1} \Psi_R \otimes (1 + 15\delta) \Psi_R \otimes \sigma_{CE} \otimes \xi_D^{1,n} \]

\[ \preceq 2^{k+2} \Psi_R \otimes \sigma_{CE} \otimes \xi_D^{1,n}, \]

where (a) uses Equation 6. 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
\[
\begin{aligned}
\frac{1}{|X|} \sum_x V_x W (\Psi_{RC} \otimes |0\rangle_E \otimes \xi_D) W^\dagger V_x^\dagger \\
= \frac{1}{S(a,n)} \sum_{j=0}^n \frac{1}{j} \sum_{c,c'} \Psi_{RC}^{(c,c')} \otimes \frac{1}{|X|} \sum_x V_x \left( |c\rangle \langle c| \otimes | j (\text{mod } b(c))\rangle \langle j (\text{mod } b(c))|_D \\
\otimes ||j/b(c)|| \langle|j/b(c)||_E \right) V_x^\dagger \\
\Rightarrow (a) \sum_{c,c'} \Psi_{RC}^{(c,c')} \otimes \delta_{c,c'} \sigma_{CE} \otimes \frac{1}{S(a,n)} \sum_{j=0}^n \frac{1}{j} ||j/b(c)|| \langle|j/b(c)||_D \\
= \sum_c \Psi_{RC}^{(c,c)} \otimes \sigma_{CE} \otimes \frac{1}{S(a,n)} \sum_{j=0}^n \frac{1}{j} ||j/b(c)|| \langle|j/b(c)||_D \\
\leq \frac{(1 + 15\delta)}{1} \sum_c \Psi_{RC}^{(c,c)} \otimes \sigma_{CE} \otimes \xi_D^{1,n} \\
= (1 + 15\delta) \Psi_{RC} \otimes \sigma_{CE} \otimes \xi_D^{1,n}.
\end{aligned}
\]

The equality (a) uses Lemma[1]. The operator inequality (b) 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.

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 \(\Psi_{RC}\). Let \(k \stackrel{\text{def}}{=} \min_{\Psi_{RC} \in \mathcal{F}(\Psi_{RC})} \max_{\Psi_{RC} \boxtimes \Psi_C} ||\Psi_{RC} \boxtimes \Psi_C||\) and \(N \stackrel{\text{def}}{=} \frac{2^{2k+2}}{\delta^2} \). Let \(\sigma_{CE}\) be the quantum state as constructed in the first part of Fact[5] using \(\Psi_{RC}\). For quantum states \(\tau_\ell\) as given in Equation[7] define
\[
\tau \stackrel{\text{def}}{=} \frac{1}{N} \sum_{\ell} \tau_\ell.
\]

It holds that
\[
P(\tau, \Psi_{RC} \otimes \sigma_{CE} \otimes \xi_D^{1,n} \otimes \mu_{X_1X_2}) \leq 2\varepsilon + 4\sqrt{\delta}.
\]

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

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 \(\sigma_C\) as given in Definition[8] let \(\mathcal{H}_{CE} \stackrel{\text{def}}{=} \text{supp}(\sigma_{CE}) \subseteq \mathcal{H}_C \otimes \mathcal{H}_E\) denote the support of \(\sigma_{CE}\). 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, \(\mathcal{H}_F \subseteq \mathcal{H}_Q \otimes \mathcal{H}_{CE} \otimes \mathcal{H}_{CE}^\prime \) and \(\text{supp}(|0\rangle_Q \otimes \mathcal{H}_{CE} \otimes \mathcal{H}_{CE}^\prime \subseteq \mathcal{H}_F\). This choice of \(F\) is guaranteed by Bertrand’s postulate[45]. Introduce registers \(F_2, F_1 \equiv F\) such that \(\mathcal{H}_{C_0E_0} \otimes \mathcal{H}_{C_1E_1} \subseteq \mathcal{H}_{F_1}\). We identify the pair \((c,c)\) 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
units constructed in Definition 3 by setting \( C \leftarrow \text{supp}(\sigma_{CE}) \). 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 \( \sigma_C \). Define the quantum states

\[
\tau_\ell \overset{\text{def}}{=} U_\ell \left( W (\Psi_{RC_0} \otimes |0\rangle \langle 0|_{E_0} \otimes \xi_{D,n}^{1:n}) W^\dagger \otimes |0\rangle \langle 0|_Q \otimes \sigma_{C_1E_1} \otimes \mu_{F_2} \right) U_\ell^\dagger,
\]

where \( W \) is as given in Definition 6 and \( |D| \geq n \). We have the following theorem, proof of which is given in Appendix B.

**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 \( \Psi_{RC_1}, \omega_C \). Let \( k \overset{\text{def}}{=} \max_{\Psi_{RC} \in \mathcal{B}(\Psi_{RC})} \max_{\omega_C} \| \Psi_{RC} \| \omega_C \), \( S \subseteq \{0, 1, \ldots, \frac{|C|^2}{\gamma^2} - 1\} \) and \( N \overset{\text{def}}{=} |S| \). Let \( \sigma_C \) be the quantum state as constructed in the first part of Fact 6 using \( \omega_C \). For the quantum states \( \tau_\ell \) as constructed in Equation 9, define

\[
\tau \overset{\text{def}}{=} \frac{1}{N} \sum_{\ell \in S} \tau_\ell.
\]

It holds that

\[
D(\tau \| \Psi_{R} \otimes \mu_{F_1} \otimes \xi_{D,n}^{1:n} \otimes \mu_{F_2}) \leq 15\delta + \log \left( 1 + \frac{2^{k+2} - 1}{N} \right).
\]

Since one can choose \( \log |D| = \log n \leq \frac{1}{8} \log \frac{|C|}{\gamma} \), the number of qubits of additional registers is \( \log |D| + 2 \log |F| \leq (4 + \frac{1}{8}) \log \frac{|C|}{\gamma} \).

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 \( \Psi_{BC} \). Let \( k \overset{\text{def}}{=} \min_{\Psi_{BC} \in \mathcal{B}(\Psi_{BC})} \max_{\sigma_C} \| \Psi_{RC} \| \omega_C \), \( S \subseteq \{0, 1, \ldots, \frac{|C|^2}{\gamma^2} - 1\} \) and \( N \overset{\text{def}}{=} |S| \). Let \( \sigma_C \) be the quantum state as constructed in the first part of Fact 6 using \( \Psi_{BC} \). For the quantum states \( \tau_\ell \) as constructed in Equation 9, define

\[
\tau \overset{\text{def}}{=} \frac{1}{N} \sum_{\ell = 1}^N \tau_\ell.
\]

It holds that

\[
P(\tau, \Psi_{R} \otimes \mu_{F_1} \otimes \xi_{D,n}^{1:n} \otimes \mu_{F_2}) \leq 2\varepsilon + 4\sqrt{\delta}.
\]

Since one can choose \( \log |D| = \log n \leq \frac{1}{8} \log \frac{|C|}{\gamma} \), the number of qubits of additional registers is \( \log |D| + 2 \log |F| \leq (4 + \frac{1}{8}) \log \frac{|C|}{\gamma} \).

### 5 Analogue of position-based decoding

We now show how to perform hypothesis testing as a dual to Theorem 2 in analogy with position-based decoding [37]. We note that similar construction can achieve a dual to Theorem 11 but we do not state it here as it will be constructed in details in Theorem 7. We have the following theorem.

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

\[
|S| \leq \frac{\delta^2}{4\varepsilon} 2D_H(\Psi_{BC} \| \Psi_{B \otimes \mu_C}).
\]
For each $\ell \in S$, let $\tau_\ell$ be the quantum state defined in Theorem 2 with $\Psi_{RC} \leftarrow \Psi_{BC}$. There exists an POVM $\{\Lambda_{-1}, \Lambda_\ell\}_{\ell \in S}$ such that
\[
\text{Tr}(\Lambda_\ell \tau_\ell) \geq 1 - \varepsilon - 4\delta \quad \forall \ell \in S.
\]

**Proof.** Let $\Omega_{BC}$ be the operator such that
\[
\text{Tr}(\Omega_{BC} \Psi_{BC}) \geq 1 - \varepsilon, \quad \text{Tr} (\Omega_{BC} \Psi_B \otimes \mu_C) = 2^{-\mathcal{D}_H^2(\Psi_{BC} \| \Psi_B \otimes \mu_C)}.
\]

We have
\[
\text{Tr}(\Omega_{BC_0} \Psi_B \otimes \mu_{G_1}) \leq \frac{|Q||C|^2}{|G|} \text{Tr} (\Omega_{BC_0} \Psi_B \otimes \mu_{C_0} \otimes \mu_Q \otimes \mu_{C_1}) \\
\leq 2 \cdot 2^{-\mathcal{D}_H^2(\Psi_{BC} \| \Psi_B \otimes \mu_C)} = 2^{-\mathcal{D}_H^2(\Psi_{BC} \| \Psi_B \otimes \mu_C)}.
\]

Let $\{\Lambda_{-1}, \Lambda_\ell\}_{\ell \in S}$ be the POVM as constructed in Fact 10 using the operators $U_\ell \Omega_{BC_0} U_\ell^\dagger$. We have
\[
\text{Tr} ((I - \Lambda_\ell) \tau_\ell) = \text{Tr} \left((I - \Lambda_\ell) U_\ell \Psi_{BC_0} \otimes |0\rangle \langle 0_Q \otimes \mu_{C_1} \otimes \mu_{G_2} U_\ell^\dagger\right) \\
\leq (1 + c) \text{Tr} \left((I - U_\ell \Omega_{BC_0} U_\ell^\dagger) U_\ell \Psi_{BC_0} \otimes |0\rangle \langle 0_Q \otimes \mu_{C_1} \otimes \mu_{G_2} U_\ell^\dagger\right) \\
+ (2 + c + c^{-1}) \sum_{m \in S, m \neq \ell} \text{Tr} \left((U_m \Omega_{BC_0} U_m^\dagger) U_\ell \Psi_{BC_0} \otimes |0\rangle \langle 0_Q \otimes \mu_{C_1} \otimes \mu_{G_2} U_\ell^\dagger\right) \\
= (1 + c) \text{Tr} \left((I - \Omega_{BC_0}) \Psi_{BC_0} \otimes |0\rangle \langle 0_Q \otimes \mu_{C_1} \otimes \mu_{G_2}\right) \\
+ (2 + c + c^{-1}) \sum_{m \in S, m \neq \ell} \text{Tr} \left(\Omega_{BC_0} U_{\ell - m} \Psi_{BC_0} \otimes |0\rangle \langle 0_Q \otimes \mu_{C_1} \otimes \mu_{G_2} U_{\ell - m}^\dagger\right) \\
\leq (1 + c)\varepsilon + (2 + c + c^{-1}) \sum_{m \in S, m \neq \ell} \text{Tr} \left(\Omega_{BC_0} U_{\ell - m} \Psi_{BC_0} \otimes |0\rangle \langle 0_Q \otimes \mu_{C_1} \otimes \mu_{G_2} U_{\ell - m}^\dagger\right).
\]

Above, $(a)$ uses Fact 10. From Lemma 3,
\[
\text{Tr}_{G_2} \left( U_{\ell - m} \Psi_{BC_0} \otimes |0\rangle \langle 0_Q \otimes \mu_{C_1} \otimes \mu_{G_2} U_{\ell - m}^\dagger\right) = \Psi_B \otimes \mu_{G_1}.
\]

Thus choosing $c = \frac{\delta}{\varepsilon}$ and using Equation 10,
\[
\text{Tr} ((I - \Lambda_\ell) \tau_\ell) \leq \varepsilon + \delta + \frac{4\varepsilon}{\delta} |S| \text{Tr} (\Omega_{BC_0} \Psi_B \otimes \mu_{G_1}) \leq \varepsilon + \delta + \frac{4\varepsilon}{\delta} |S| 2^{-\mathcal{D}_H^2(\Psi_{BC} \| \Psi_B \otimes \mu_C)} \leq \varepsilon + 4\delta,
\]
from the choice of $|S|$. This completes the proof. \qed

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.

**Theorem 6.** Let $\varepsilon \in (0, 1), \delta \in (0, \frac{1}{10})$ and $\Psi_{BC}, \omega_C$ be quantum states. Let $S \subseteq \{0, 1, \ldots, |G| - 1\}$ such that
\[
|S| \leq \frac{\delta^2}{4\varepsilon^2} 2^{-\mathcal{D}_H^2(\Psi_{BC} \| \Psi_B \otimes \omega_C)}.
\]

Let $\sigma_C$ be the quantum state as constructed in the second part of Fact 6 using $\omega_C$. Let $\tau_\ell$ be the quantum states as defined in Equation 9 using the quantum states $\Psi_{RC} \leftarrow \Psi_{BC}, \sigma_{CE}$ and by choosing $|D| \leq 2n \cdot |E| \leq 2|E|^{1 + \frac{1}{8}}$. There exists a collection of POVM $\{\Lambda_{-1}, \Lambda_\ell\}_{\ell \in S}$ such that
\[
\text{Tr}(\Lambda_\ell \tau_\ell) \geq 1 - \varepsilon - 64\delta \quad \forall \ell \in S.
\]
Proof. We will outline the main steps of the proof, which closely follow those of Theorem 5. Let $\Omega_{BC}$ be the operator that satisfies

$$\text{Tr}(\Omega_{BC}\Psi_{BC}) \geq 1 - \varepsilon, \quad \text{Tr}(\Omega_{BC}\Psi_B \otimes \omega_C) = 2^{-D_H(\Psi_{BC}\|\Psi_B \otimes \omega_C)}.$$ 

From Fact 6 we have

$$\sigma_C \leq \frac{1}{1 - \gamma} \omega_C \implies \text{Tr}(\Omega_{BC}\Psi_B \otimes \sigma_C) \leq 2 \cdot \text{Tr}(\Omega_{BC}\Psi_B \otimes \omega_C) = 2^{1 - D_H(\Psi_{BC}\|\Psi_B \otimes \omega_C)}.$$

Let $\{\Lambda_n, \Lambda_\ell\}_{\ell \in S}$ be the POVM constructed in Fact 10 using the operators $\{U_\ell W\Omega_{BC_0}W^\dagger U_\ell^\dagger\}_{\ell \in S}$. Rest of the calculation follows using Fact 10. The following claim is similar to Lemma 3 and is proved in Appendix C.

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

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

\[
\text{Tr}\left(U_m W\Omega_{BC_0} W^\dagger U_m^\dagger \tau_\ell\right) = \text{Tr}\left(U_m W\Omega_{BC_0} W^\dagger U_m^\dagger \left(W (\Psi_{BC_0} \otimes |0\rangle\langle 0|_{E_0} \otimes \xi_D^{1n}) W^\dagger \otimes |0\rangle\langle 0|_Q \otimes \sigma_{C_1} E_1 \otimes \mu_{F_2}\right)U_\ell^\dagger\right) = \text{Tr}\left(W\Omega_{BC_0} W^\dagger U_{\ell - m}\left(W (\Psi_{BC_0} \otimes |0\rangle\langle 0|_{E_0} \otimes \xi_D^{1n}) W^\dagger \otimes |0\rangle\langle 0|_Q \otimes \sigma_{C_1} E_1 \otimes \mu_{F_2}\right)U_{\ell - m}^\dagger\right) \\
\leq (1 + 15\delta) \text{Tr}\left(\Omega_{BC_0} W^\dagger (\Psi_R \otimes \mu_{F_1} \otimes \xi_D^{1n}) W\right) \\
\leq 2 \cdot \text{Tr}\left(\Omega_{BC_0} W^\dagger (\Psi_R \otimes \mu_{F_1} \otimes \xi_D^{1n}) W\right). 
\]

Above, (a) follows from Claim 4. Now, we use the fact that

$$\mu_{F_1} \leq \frac{|Q||\text{supp}(\sigma_{CE})|^2}{|F|} \mu_Q \otimes \sigma_{C_0 E_0} \otimes \sigma_{C_1 E_1} \leq 2 \cdot \mu_Q \otimes \sigma_{C_0 E_0} \otimes \sigma_{C_1 E_1}.$$

Thus,

\[
\text{Tr}\left(U_m W\Omega_{BC_0} W^\dagger U_m^\dagger \tau_\ell\right) \\
\leq 2 \cdot \text{Tr}\left(\Omega_{BC_0} W^\dagger (\Psi_R \otimes \mu_{F_1} \otimes \xi_D^{1n}) W\right) \\
\leq 4 \cdot \text{Tr}\left(\Omega_{BC_0} W^\dagger (\Psi_R \otimes \mu_Q \otimes \mu_{C_0 E_0} \otimes \mu_{C_1 E_1} \otimes \xi_D^{1n}) W\right) \\
= 4 \cdot \text{Tr}\left(\Omega_{BC_0} W^\dagger (\Psi_R \otimes \mu_{C_0 E_0} \otimes \xi_D^{1n}) W\right). 
\]

Finally, we use Claim 2 to conclude that

\[
\text{Tr}\left(U_m W\Omega_{BC_0} W^\dagger U_m^\dagger \tau_\ell\right) \\
\leq 4 \cdot \text{Tr}\left(\Omega_{BC_0} W^\dagger (\Psi_R \otimes \mu_{C_0 E_0} \otimes \xi_D^{1n}) W\right) \\
\leq 16 \cdot \text{Tr}\left(\Omega_{BC_0} \Psi_R \otimes \sigma_{C_0} \otimes |0\rangle\langle 0|_{E_0} \otimes \xi_D^{1n}\right) \\
= 16 \cdot \text{Tr}\left(\Omega_{BC_0} \Psi_R \otimes \sigma_{C_0}\right) \leq 2^{5 - D_H(\Psi_{BC}\|\Psi_B \otimes \omega_C)},
\]

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

6.1 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 [37]. We begin by defining an entanglement-assisted code.

**Definition 7.** Fix an $\varepsilon \in (0, 1)$ and a positive integer $R$. Let $M'$ be a register of dimension $|M| = 2^R$. A $(R, \varepsilon)$ entanglement-assisted code for a quantum channel $\mathcal{N}_{C \rightarrow B}$ consists of a shared entanglement $|\Theta\rangle_{E_A E_B}$ between Alice ($E_A$) and Bob ($E_B$) and

- An encoding operation $E_m : \mathcal{L}(E_A) \rightarrow \mathcal{L}(C)$ for each $m \in \{1, 2, \ldots, 2^R\}$,
- A decoding operation $D : \mathcal{L}(E_B) \rightarrow \mathcal{L}(M')$ which leads to a classical distribution on register $M'$ such that
  \[ \Pr [M' \neq m] \leq \varepsilon, \quad \forall m \in \{1, 2, \ldots, 2^R\}. \]

We have the following theorem, near-optimality of which is shown by the converse given in [76]. For the ease of presentation, we will represent the relation $P(|\psi\rangle\langle\psi|, |\phi\rangle) \leq \varepsilon$ between two pure states $|\psi\rangle, |\phi\rangle$ as $|\psi\rangle \approx |\phi\rangle$.

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

\[ R \leq D_H^\varepsilon (\mathcal{N}_{A \rightarrow B}(|\Psi\rangle_{AC})||\mathcal{N}_{A \rightarrow B}(|\Psi\rangle_{A} \otimes |\Psi\rangle_{C}) - 5 - \log \frac{4(\varepsilon + 4\gamma^{1/4})}{\delta'}, \]

there exists a $(R, \varepsilon + 4\gamma^{1/4} + \delta' + 20\sqrt{\delta})$ entanglement-assisted code for a quantum channel $\mathcal{N}_{A \rightarrow B}$. The protocol uses $\frac{1}{3} \log \frac{|A|}{\gamma^2}$ qubits of shared entanglement and $4 \log |A|$ bits of shared randomness. The latter can be fixed by standard derandomization argument.

**Proof.** Without loss of generality, we can assume that $|\Psi\rangle_{AC}$ is the canonical purification of $|\Psi\rangle_{A}$, by applying a local unitary on register $C$ which does not change the hypothesis testing relative entropy. Let $\Psi_{BC} \overset{\text{def}}{=} \mathcal{N}_{A \rightarrow B}(|\Psi\rangle_{AC})$. From Fact 6, there exists a quantum state $\sigma_C$ such that the eigenvalues of $\sigma_C$ are integer multiples of $\frac{1}{|C|}$ and

\[ \sigma_C \leq \frac{1}{1 - \gamma} |\Psi_C\rangle\langle\Psi_C| \implies P(\sigma_C, \sigma_C) \leq \sqrt{\gamma}. \]

Let $|\sigma\rangle_{AC}$ be the canonical purification of $|\sigma\rangle_{C}$ and $\sigma_{BC} \overset{\text{def}}{=} \mathcal{N}_{A \rightarrow B}(|\sigma\rangle_{AC})$. Using Fact 8,

\[ \sigma_A \leq \frac{1}{1 - \gamma} |\Psi_A\rangle, \quad P(\sigma_{AC}, |\Psi\rangle_{AC}) \leq 2\sqrt{P(\sigma_C, |\Psi_C\rangle)} \leq 2\gamma^{1/4} \]

and using Fact 2

\[ \sigma_B = \mathcal{N}_{A \rightarrow B}(|\sigma\rangle_{A}) \leq \frac{1}{1 - \gamma} \mathcal{N}_{A \rightarrow B}(|\Psi\rangle_{A}) = \frac{1}{1 - \gamma} |\Psi_B\rangle, \quad P(\sigma_{BC}, \Psi_{BC}) \leq P(\sigma_{AC}, |\Psi\rangle_{AC}) \leq 2\gamma^{1/4}. \]
Let $\Omega_{BC}$ be the optimum operator in the definition of $D_H^0(\Psi_{BC}\|\Psi_B \otimes \Psi_C)$. From Fact\textsuperscript{7} \[\text{Tr} \left( \Omega_{BC} \sigma_{BC} \right) \geq 1 - \varepsilon - 4\gamma^{1/4}, \quad \text{Tr} \left( \Omega_{BC} \sigma_B \otimes \sigma_C \right) \leq \frac{1}{(1 - \gamma)^2} \text{Tr} \left( \Omega_{BC} \Psi_B \otimes \Psi_C \right) \leq 2^{2-D_H^0(\Psi_{BC}\|\Psi_B \otimes \Psi_C)}.\] (11)

We expand $|\sigma\rangle_{AC} = \sum_c \sqrt{q(c)}|c\rangle_A|c\rangle_C$. Let $E$ be the register and $\sigma_{CE}$ be the quantum state as obtained in Definition\textsuperscript{5}. It holds that $|E| \leq \frac{|A|}{\gamma}$. Consider the following purification of $\sigma_{CE}$, which is maximally entangled.

$$|\sigma'\rangle_{ACE'E} \overset{\text{def}}{=} \sqrt{\frac{\gamma}{|C|}} \sum_{c,e \leq \frac{\log|C|}{\gamma}} |c,e\rangle_{AE'}|c\rangle_{CE}.$$  

Let $a = \frac{|C|}{\gamma}$, $n = a^{\frac{1}{\delta}}$ and register $D$ satisfy $|D| = n \left( \frac{|C|}{\gamma} + 1 \right)$. This ensures that Claims\textsuperscript{1,2,3} apply to register $E$. From Definition\textsuperscript{3} $|\xi^{acn}_{D'D}\rangle$ is the canonical purification of $\xi^{acn}_{D'D}$ with $D' \equiv D$. Given the unitary $W \overset{\text{def}}{=} \sum_c |c\rangle \langle c| \otimes W_{\sigma(c)c}$ from Definition\textsuperscript{3} let $W'^{alice} \overset{\text{def}}{=} W_{AE'D'}$ and $W'^{bob} \overset{\text{def}}{=} W_{CE'D}$. Using Definition\textsuperscript{11} we obtain $|\text{supp}(\sigma_{CE})|^2 = \left( \frac{|C|}{\gamma} \right)^2$ unitaries $V_x : \text{supp}(\sigma_{CE}) \to \text{supp}(\sigma_{CE})$. From Claim\textsuperscript{8} we have

$$(W'^{alice} \otimes W'^{bob}) |\sigma\rangle_{AC} \otimes |\xi^{acn}_{D'D}\rangle \otimes |0,0\rangle_{E'E} \overset{5\gamma}{\approx} |\sigma'\rangle_{ACE'E} \otimes |\xi^{1n}_{D'D}\rangle.$$  

Since $V_x$ acts in $\text{supp}(\sigma_{CE})$, Fact\textsuperscript{11} ensures that there exists a unitary $V_x^T : \text{supp}(\sigma'_{AE'}) \to \text{supp}(\sigma'_{AE'})$ such that $(V_x^T \otimes I)|\sigma'\rangle_{ACE'E} = (I \otimes V_x)|\sigma\rangle_{ACE'E}$. Thus, we obtain

$$\left( V_x^T W'^{alice} \otimes W'^{bob} \right) |\sigma\rangle_{AC} \otimes |\xi^{acn}_{D'D}\rangle \otimes |0,0\rangle_{E'E} \overset{5\gamma}{\approx} (I \otimes V_x)|\sigma'\rangle_{ACE'E} \otimes |\xi^{1n}_{D'D}\rangle.$$  

By triangle inequality for purified distance (Fact\textsuperscript{11}), these equations lead to

$$\left( W'^{alice} \right)^T V_x^T W'^{alice} \otimes W'^{bob} |\sigma\rangle_{AC} \otimes |\xi^{acn}_{D'D}\rangle \otimes |0,0\rangle_{E'E} \overset{10\gamma}{\approx} (I \otimes V_x W'^{bob}) |\sigma\rangle_{AC} \otimes |\xi^{acn}_{D'D}\rangle \otimes |0,0\rangle_{E'E}.\] (12)

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\textsuperscript{2} gives a family of functions $\{f_m : \mathcal{X} \times \mathcal{X} \to \mathcal{X}\}$ and a collection of unitaries

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

Let $\{A_{-1}, A_1, \ldots, A_{2n}\}$ be POVM as defined in Fact\textsuperscript{11} using the operators $\{ (V^{(m)} W'^{bob}) \Omega_{BC} (V^{(m)} W'^{bob})^\dagger \}_{m=1}^{2R}$. 

**Shared resources:** Alice and Bob share the state $|\sigma\rangle_{AC} |\xi^{acn}_{D'D}\rangle |0,0\rangle_{E'E}$. They also possess $\mu_{X_1X_2}$ in shared registers $X_1X_2$. Thus, the number of qubits of shared entanglement is

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

27
**Encoding:** To send the message \( m \in \{1, 2, \ldots, 2^R\} \), Alice applies the unitary
\[
\sum_{x_1, x_2} (W^{alice})^T f_m(x_1, x_2) W^{alice} \otimes |x_1, x_2\rangle \langle x_1, x_2| X_1 X_2
\]
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 \( \{\Lambda_{-1}, \Lambda_1, \ldots, \Lambda_{2^R}\} \) and outputs \( m' \) upon obtaining the outcome \( \Lambda_{m'} \).

**Error analysis:** Let \( \theta_m \) be the quantum state on Bob’s registers just after Alice’s transmission through the channel. Define the following quantum state:
\[
\theta_m \overset{\text{def}}{=} \frac{1}{|X_1|^2} \sum_{x_1, x_2} |x_1, x_2\rangle \langle x_1, x_2| X_1 X_2 \otimes (V_m(x_1, x_2) W^{bob}) (\sigma_{BC} \otimes \xi_D^{m,n} \otimes |0\rangle \langle 0|) (V_m(x_1, x_2) W^{bob})^T.
\]
From Equation[12] we have \( \Pr(\theta_m, \theta_m') \leq 10 \sqrt{\delta} \). Thus, from Fact[7]
\[
\Pr[M' \neq m] = \text{Tr} ((1 - \Lambda_m) \theta_m) \leq \text{Tr} ((1 - \Lambda_m) \theta_m) + 2 \text{Pr}(\theta_m, \theta_m') \leq 20 \sqrt{\delta} + \text{Tr} ((1 - \Lambda_m) \theta_m).
\]
Applying Fact[10] we conclude
\[
\Pr[M' \neq m] = 20 \sqrt{\delta} + \text{Tr} ((1 - \Lambda_m) \theta_m) \leq 20 \sqrt{\delta} + (1 + c) \left( 1 - \text{Tr} \left( (V^{(m)} W^{bob}) \Omega_{BC}(V^{(m)} W^{bob})^T \theta_m \right) \right) + (2 + c + c^{-1}) \sum_{m' \neq m} \text{Tr} \left( (V^{(m') \Gamma} W^{bob}) \Omega_{BC}(V^{(m') \Gamma} W^{bob})^T \theta_m \right) \leq 20 \sqrt{\delta} + (1 + c) \left( 1 - \text{Tr} \left( \Omega_{BC}(V^{(m)} W^{bob})^T \theta_m (V^{(m)} W^{bob}) \right) \right) + (2 + c + c^{-1}) \sum_{m' \neq m} \text{Tr} \left( (V^{(m') \Gamma} W^{bob}) \Omega_{BC}(V^{(m') \Gamma} W^{bob})^T \theta_m \right).
\]
Since
\[
(V^{(m)} W^{bob})^T \theta_m (V^{(m)} W^{bob}) = \frac{1}{|X_1|^2} \sum_{x_1, x_2} |x_1, x_2\rangle \langle x_1, x_2| X_1 X_2 \otimes (\sigma_{BC} \otimes \xi_D^{m,n} \otimes |0\rangle \langle 0|_E) ,
\]
from Equation[11] we have
\[
\text{Tr} \left( \Omega_{BC}(V^{(m)} W^{bob})^T \theta_m (V^{(m)} W^{bob}) \right) = \text{Tr}(\Omega_{BC} \sigma_{BC}) \geq 1 - \varepsilon - 4 \gamma^{1/4}.
\]
For \( m' \neq m \), consider
\[
\text{Tr} \left( (V^{(m')} W^{bob}) \Omega_{BC}(V^{(m')} W^{bob})^T \theta_m \right) = \frac{1}{|X_1|^2} \sum_{x_1, x_2} \text{Tr} \left( \Omega_{BC}(V_{m'(x_1, x_2)} W^{bob}) (\sigma_{BC} \otimes \xi_D^{m,n} \otimes |0\rangle \langle 0|_E) (V_{m'(x_1, x_2)} W^{bob})^T (V_{m'(x_1, x_2)} W^{bob}) \right) = \frac{1}{|X|} \sum_{x} \text{Tr} \left( \Omega_{BC}(V_{x} W^{bob}) (\sigma_{BC} \otimes \xi_D^{m,n} \otimes |0\rangle \langle 0|_E) (V_{x} W^{bob})^T (V_{x} W^{bob}) \right).
\]
where we have used Definition 2 to introduce variables \( x, x' \) in a manner similar to Equation 1. From Equation 8, we have

\[
\frac{1}{|\mathcal{X}|} \sum_{x'} V_{x'} W_{bob}^{\dagger} (\sigma_B \otimes \sigma_{CE} \otimes |0\rangle \langle 0|_E) (V_{x'} W_{bob}^{\dagger})^\dagger \preceq (1 + 15\delta) \sigma_B \otimes \sigma_{CE} \otimes \xi^{1:n}.
\]

Thus,

\[
\text{Tr} \left( (V^{(m')} W_{bob}^{\dagger}) \Omega_{BC} (V^{(m')} W_{bob}^{\dagger})^\dagger \right) \leq (1 + 15\delta) \sum_x \text{Tr} \left( \Omega_{BC} (V_{x'} W_{bob}^{\dagger})^\dagger \left( \sigma_B \otimes \sigma_{CE} \otimes \xi^{1:n} \right) (V_{x'} W_{bob}^{\dagger}) \right) \leq 8 \cdot \text{Tr} \left( \Omega_{BC} \sigma_B \otimes \sigma_{CE} \right) \leq 2^{5-D_{\text{H}}(\Psi_{BC} \| \Psi_B \otimes \Psi_C)}.
\]

where (a) uses the fact that \( V_{x'}^{\dagger} \sigma_{CE} V_{x} = \sigma_{CE} \), (b) uses Claim 2 and (c) uses Equation 11. Using it with Equation 14 and Equation 13, we conclude

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

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

6.2 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 [11]). Since the proof is similar to that given in [11], 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 (A), 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 \( \Phi_{RABC} \) between Alice (A), Bob (BC) and Reference (R) satisfies \( \Phi_{RABC} \in B^{4\varepsilon+65\delta}(\Psi_{RABC}) \). Reference plays no role in the protocol.
The number of qubits of shared entanglement required is at most \( \left( 4 + \frac{1}{\varepsilon} \right) \log \frac{|C|}{\delta} \) and the number of qubits communicated is

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

By the argument in [6] that shows how a convex-split for \( |\Psi\rangle_{RC} \) 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 $\Phi_{RAC}$ between Alice (A), Bob (C) and Reference (R) satisfies $\Phi_{RAC} \in B_{2^{2\varepsilon + 8\sqrt{\delta}}} \left( |\Psi_{RABC}\rangle \right)$. Reference plays no role in the protocol. The number of qubits communicated is

$$\frac{1}{2} I_{\max}(R : C) + 2 + 2 \log \frac{1}{\delta}.$$ 

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

Thus, the result improves upon the number of qubits communicated in [5] by an additive factor of $\log \log |C|$ and at the same time achieves the number of qubits of entanglement required. It achieves the same communication as given in [6], but exponentially improves upon the number of qubits of entanglement.

Conclusion

In this work, we have studied the problem of decoupling quantum registers by means of efficient operations and reduction in the resource required in the applications of convex-split and position-based decoding. We have given an efficient operation that achieves the task and acts ‘classically’ on a preferred basis. Along with our second method, we exponentially improve the amount of entanglement required in the quantum communication protocols based on the convex-split and position-based decoding methods.

An important question is to see if the number of bits of additional randomness used in our decoupling protocol can further be reduced. It is known that seed size in randomness extraction in presence of quantum side information can be very small [35] (based on Trevisan’s construction [36]). Since our construction treats classical side information and quantum side information in similar manner, we can hope to have similar results even in the case of quantum decoupling. Similarly, we would like to understand if the amount of entanglement can further be reduced in our applications to quantum communication tasks.

Acknowledgment

This work is supported by the Singapore Ministry of Education through the Tier 3 Grant “Random numbers from quantum processes” MOE2012-T3-1-009 and VAJRA Grant, Department of Science and Technology, Government of India.

References

[1] M. Horodecki, J. Oppenheim, and A. Winter, “Partial quantum information,” Nature, vol. 436, pp. 673–676, Aug 2005.

[2] M. Horodecki, J. Oppenheim, and A. Winter, “Quantum state merging and negative information,” Communications in Mathematical Physics, vol. 269, no. 1, pp. 107–136, 2007.

[3] A. Abeyesinghe, I. Devetak, P. Hayden, and A. Winter, “The mother of all protocols: restructuring quantum information’s family tree,” Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, vol. 465, no. 2108, pp. 2537–2563, 2009.
[4] M. Berta, “Single-shot quantum state merging.” Master’s thesis, ETH Zurich, http://arxiv.org/abs/0912.4495, 2009.

[5] M. Berta, M. Christandl, and R. Renner, “The Quantum Reverse Shannon Theorem based on one-shot information theory,” Commun. Math. Phys., vol. 306, no. 3, pp. 579–615, 2011.

[6] A. Anshu, V. K. Devabathini, and R. Jain, “Quantum communication using coherent rejection sampling,” Phys. Rev. Lett., vol. 119, p. 120506, Sep 2017.

[7] I. Devetak and J. Yard, “Exact cost of redistributing multipartite quantum states,” Phys. Rev. Lett., vol. 100, no. 230501, 2008.

[8] J. T. Yard and I. Devetak, “Optimal quantum source coding with quantum side information at the encoder and decoder,” IEEE Transactions on Information Theory, vol. 55, pp. 5339–5351, 2009.

[9] N. Datta, M.-H. Hsieh, and J. Oppenheim, “An upper bound on the second order asymptotic expansion for the quantum communication cost of state redistribution,” Journal of Mathematical Physics, vol. 57, no. 5, p. 052203, 2016.

[10] M. Berta, M. Christandl, and D. Touchette, “Smooth entropy bounds on one-shot quantum state redistribution,” IEEE Transactions on Information Theory, vol. 62, pp. 1425–1439, March 2016.

[11] A. Anshu, R. Jain, and N. A. Warsi, “A one-shot achievability result for quantum state redistribution,” IEEE Transactions on Information Theory, vol. 64, pp. 1425–1435, March 2018.

[12] I. Devetak, “The private classical capacity and quantum capacity of a quantum channel,” IEEE Transactions on Information Theory, vol. 51, pp. 44–55, Jan 2005.

[13] P. Hayden, M. Horodecki, A. Winter, and J. Yard, “A decoupling approach to the quantum capacity,” Open Systems & Information Dynamics, vol. 15, no. 01, pp. 7–19, 2008.

[14] F. Dupuis, P. Hayden, and K. Li, “A father protocol for quantum broadcast channels,” IEEE Transactions on Information Theory, vol. 56, pp. 2946–2956, June 2010.

[15] N. Datta and M. H. Hsieh, “One-shot entanglement-assisted quantum and classical communication,” IEEE Transactions on Information Theory, vol. 59, pp. 1929–1939, March 2013.

[16] N. Datta, M. Tomamichel, and M. M. Wilde, “On the second-order asymptotics for entanglement-assisted communication,” Quantum Information Processing, vol. 15, no. 6, pp. 2569–2591, 2016.

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

[18] M. Berta, “Quantum side information: Uncertainty relations, extractors, channel simulations.” PhD Thesis, ETH Zurich, Diss. ETH No. 21180, arXiv:quant-ph/1310.4581, 2005.

[19] M. Berta, O. Fawzi, and S. Wehner, “Quantum to classical randomness extractors,” IEEE Transactions on Information Theory, vol. 60, pp. 1168–1192, Feb 2014.

[20] L. d. Rio, J. Aberg, R. Renner, O. Dahlsten, and V. Vedral, “The thermodynamic meaning of negative entropy,” Nature, vol. 474, 2011.
[21] D. N. Page, “Average entropy of a subsystem,” Phys. Rev. Lett., vol. 71, pp. 1291–1294, Aug 1993.

[22] P. Hayden and J. Preskill, “Black holes as mirrors: quantum information in random subsystems,” Journal of High Energy Physics, vol. 2007, no. 09, p. 120, 2007.

[23] F. Dupuis, “The decoupling approach to quantum information theory,” PhD Thesis, Université de Montréal., http://arxiv.org/abs/1410.0664, 2010.

[24] O. Szehr, “Decoupling theorems.” Master’s theisis, ETH Zurich, arXiv:quant-ph/1207.3927, 2011.

[25] F. Dupuis, M. Berta, J. Wullschleger, and R. Renner, “One-shot decoupling,” Communications in Mathematical Physics, vol. 328, no. 251, 2014.

[26] C. Dankert, R. Cleve, J. Emerson, and E. Livine, “Exact and approximate unitary 2-designs and their application to fidelity estimation,” Phys. Rev. A, vol. 80, p. 012304, Jul 2009.

[27] D. P. DiVincenzo, D. W. Leung, and B. M. Terhal, “Quantum data hiding,” IEEE Transactions on Information Theory, vol. 48, pp. 580–598, March 2002.

[28] H. F. Chau, “Unconditionally secure key distribution in higher dimensions by depolarization,” IEEE Transactions on Information Theory, vol. 51, pp. 1451–1468, April 2005.

[29] R. Cleve, D. Leung, L. Liu, and C. Wang, “Near-linear constructions of exact unitary 2-designs,” Quantum Information and Computation, vol. 16, no. 9-10, 2016.

[30] W. Brown and O. Fawzi, “Decoupling with random quantum circuits,” Commun. Math. Phys., vol. 340, no. 867, 2015.

[31] D. Slepian and J. Wolf, “Noiseless coding of correlated information sources,” IEEE Transactions on Information Theory, vol. 19, pp. 471–480, Jul 1973.

[32] N. Nisan and D. Zuckerman, “Randomness is linear in space,” Journal of Computer and System Sciences, vol. 52, no. 1, pp. 43 – 52, 1996.

[33] J. Radhakrishnan and A. Ta-Shma, “Bounds for dispersers, extractors, and depth-two superconcentrators,” SIAM Journal on Discrete Mathematics, vol. 13, no. 1, pp. 2–24, 2000.

[34] S. Vadhan, “Pseudorandomness.” Lecture notes, Online: http://people.seas.harvard.edu/~salil/pseudorandomness/, 2011.

[35] A. De, C. Portmann, T. Vidick, and R. Renner, “Trevisan’s extractor in the presence of quantum side information,” SIAM Journal on Computing, vol. 41, no. 4, pp. 915–940, 2012.

[36] L. Trevisan, “Extractors and pseudorandom generators,” J. ACM, vol. 48, pp. 860–879, July 2001.

[37] A. Anshu, R. Jain, and N. A. Warsi, “Building blocks for communication over noisy quantum networks,” IEEE Transactions on Information Theory, pp. 1–1, 2018.

[38] C. H. Bennett, P. W. Shor, J. A. Smolin, and A. V. Thapliyal, “Entanglement-assisted capacity of a quantum channel and the reverse shannon theorem,” IEEE Transactions on Information Theory, vol. 48, pp. 2637–2655, Oct 2002.
[39] B. Groisman, S. Popescu, and A. Winter, “Quantum, classical, and total amount of correlations in a quantum state,” Phys. Rev. A, vol. 72, p. 032317, Sep 2005.

[40] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters, “Teleporting an unknown quantum state via dual classical and einstein-podolsky-rosen channels,” Phys. Rev. Lett., vol. 70, pp. 1895–1899, Mar 1993.

[41] C. H. Bennett and S. J. Wiesner, “Communication via one- and two-particle operators on einstein-podolsky-rosen states,” Phys. Rev. Lett., vol. 69, no. 20, pp. 2881–2884, 1992.

[42] C. Majenz, M. Berta, F. Dupuis, R. Renner, and M. Christandl, “Catalytic decoupling of quantum information,” Phys. Rev. Lett., vol. 118, p. 080503, Feb 2017.

[43] A. Anshu, M.-H. Hsieh, and R. Jain, “Quantifying resource in catalytic resource theory.” Physical Review Letters (to appear), 2017.

[44] M. Berta and C. Majenz, “Disentanglement cost of quantum states.” https://arxiv.org/abs/1708.00360, 2017.

[45] P. Tchebychev, “Mémoire sur les nombres premiers,” Journal de mathématiques pures et appliquées, vol. 1, pp. 366–390, 1852.

[46] A. Streltsov, G. Adesso, and M. B. Plenio, “Colloquium: Quantum coherence as a resource,” Rev. Mod. Phys., vol. 89, p. 041003, Oct 2017.

[47] A. Broder, “On the resemblance and containment of documents,” in Proceedings of the Compression and Complexity of Sequences 1997, SEQUENCES ’97, (Washington, DC, USA), pp. 21–, IEEE Computer Society, 1997.

[48] M. S. Charikar, “Similarity estimation techniques from rounding algorithms,” in Proceedings of the Thirty-fourth Annual ACM Symposium on Theory of Computing, STOC ’02, (New York, NY, USA), pp. 380–388, ACM, 2002.

[49] J. Kleinberg and E. Tardos, “Approximation algorithms for classification problems with pairwise relationships: Metric labeling and markov random fields,” J. ACM, vol. 49, pp. 616–639, Sept. 2002.

[50] T. Holenstein, “Parallel repetition: simplifications and the no-signaling case,” in Proceedings of the thirty-ninth annual ACM Symposium on Theory of Computing, STOC ’07, (New York, NY, USA), pp. 411–419, 2007.

[51] B. Barak, M. Hardt, I. Haviv, A. Rao, O. Regev, and D. Steurer, “Rounding parallel repetitions of unique games,” in Proceedings of the 2008 49th Annual IEEE Symposium on Foundations of Computer Science, FOCS ’08, (Washington, DC, USA), pp. 374–383, IEEE Computer Society, 2008.

[52] M. Braverman and A. Rao, “Information equals amortized communication,” in Proceedings of the 52nd Symposium on Foundations of Computer Science, FOCS ’11, (Washington, DC, USA), pp. 748–757, IEEE Computer Society, 2011.
[53] A. Anshu, R. Jain, P. Mukhopadhyay, A. Shayeghi, and P. Yao, “New one-shot quantum protocols with application to communication complexity,” IEEE Transactions on Information Theory, vol. 62, pp. 7566–7577, Dec 2016.

[54] A. Anshu, R. Jain, and N. Warsi, “A unified approach to source and message compression,” https://arxiv.org/abs/1707.03619, 2017.

[55] W. van Dam and P. Hayden, “Universal entanglement transformations without communication,” Phys. Rev. A, vol. 67, p. 060302, Jun 2003.

[56] A. Barenco, C. H. Bennett, R. Cleve, D. P. DiVincenzo, N. Margolus, P. Shor, T. Sleator, J. A. Smolin, and H. Weinfurter, “Elementary gates for quantum computation,” Phys. Rev. A, vol. 52, pp. 3457–3467, Nov 1995.

[57] R. Jozsa, “Fidelity for mixed quantum states,” Journal of Modern Optics, vol. 41, no. 12, pp. 2315–2323, 1994.

[58] A. Uhlmann, “The “transition probability” in the state space of a *-algebra,” Rep. Math. Phys., vol. 9, pp. 273–279, 1976.

[59] A. Gilchrist, N. K. Langford, and M. A. Nielsen, “Distance measures to compare real and ideal quantum processes,” Phys. Rev. A, vol. 71, p. 062310, Jun 2005.

[60] N. Datta, “Min- and max- relative entropies and a new entanglement monotone,” IEEE Transactions on Information Theory, vol. 55, pp. 2816–2826, 2009.

[61] R. Jain, J. Radhakrishnan, and P. Sen, “A property of quantum relative entropy with an application to privacy in quantum communication,” J. ACM, vol. 56, pp. 33:1–33:32, Sept. 2009.

[62] F. Buscemi and N. Datta, “The quantum capacity of channels with arbitrarily correlated noise,” IEEE Transactions on Information Theory, vol. 56, pp. 1447–1460, 2010.

[63] M. Hayashi and H. Nagaoka, “General formulas for capacity of classical-quantum channels,” IEEE Transactions on Information Theory, vol. 49, pp. 1753–1768, July 2003.

[64] N. Ciganović, N. J. Beaudry, and R. Renner, “Smooth max-information as one-shot generalization for mutual information,” IEEE Transactions on Information Theory, vol. 60, pp. 1573–1581, 2014.

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

[66] H. Barnum, C. M. Caves, C. A. Fuch, R. Jozsa, and B. Schumacher, “Noncommuting mixed states cannot be broadcast,” Phys. Rev. Lett., vol. 76, no. 15, pp. 2818–2821, 1996.

[67] G. Lindblad, “Completely positive maps and entropy inequalities,” Commun. Math. Phys., vol. 40, pp. 147–151, 1975.

[68] A. Winter, “Coding theorem and strong converse for quantum channels,” IEEE Transactions on Information Theory, vol. 45, no. 7, pp. 2481–2485, 1999.
The circuit size of the decoupling unitary can be bounded as follows. The relabeling $|c_0, c_1⟩_{G_0G_1} → |c_0|C⟩ + |c_1⟩_{G_1}$ can be performed by the multiplication algorithm of Schönhage and Strassen [77] using a circuit of size $O(\log |C| \log \log |C|)$ and depth $O(\log \log |C|)$. Thus, we focus on unitary transformation over the basis $\{|i⟩\}_{i=0}^{|G|}$ for $\mathcal{H}_{G_1}$. From Definition 3 we have

$$U = \sum_{i,j,\ell} j\ell + i(1 - \ell) |G_1⟩ |j(\ell + 1) - i\ell⟩_{G_2} ⟨i|G_1⟩ |j⟩_{G_2} \otimes |\ell⟩⟨\ell|L.$$
Proof. Fact 6 ensures that
\[
\omega_C \lesssim \frac{1}{1-\gamma} \sigma_C \implies \Psi_{RC} \lesssim \frac{1}{1-\gamma} 2^k \Psi_R \otimes \sigma_C \lesssim 2^{k+1} \Psi_R \otimes \sigma_C.
\]
Using Claim 4 and Fact 9, we proceed similar to Theorem 2

\[
\begin{align*}
D &\left( \tau \big| \Psi_R \otimes \mu_{F_1} \otimes \xi_D^{1:n} \otimes \mu_{F_2} \right) \\
&= \frac{1}{N} \sum_{\ell \in S} \left( D(\tau_\ell \big| \Psi_R \otimes \mu_{F_1} \otimes \xi_D^{1:n} \otimes \mu_{F_2}) - D(\tau_\ell \big| \tau) \right) \\
&= \frac{1}{N} \sum_{\ell \in S} \left( D\left( W \left( \Psi_{RC_0} \otimes |0\rangle \langle 0|_{E_0} \otimes \xi_D^{a:n} \right) W^\dagger \otimes |0\rangle \langle 0|_Q \otimes \sigma_{C_1 E_1} \otimes \mu_{F_2} \big| \Psi_R \otimes \mu_{F_1} \otimes \xi_D^{1:n} \otimes \mu_{F_2} \right) \\
&\quad - D\left( W \left( \Psi_{RC_0} \otimes |0\rangle \langle 0|_{E_0} \otimes \xi_D^{a:n} \right) W^\dagger \otimes |0\rangle \langle 0|_Q \otimes \sigma_{C_1 E_1} \otimes \mu_{F_2} \big| U_\ell^\dagger \tau U_\ell \right) \right) \\
&\leq \frac{1}{N} \sum_{\ell \in S} \left( D\left( W \left( \Psi_{RC_0} \otimes |0\rangle \langle 0|_{E_0} \otimes \xi_D^{a:n} \right) W^\dagger \otimes |0\rangle \langle 0|_Q \otimes \sigma_{C_1 E_1} \right| \Psi_R \otimes \mu_{F_1} \otimes \xi_D^{1:n} \right) \\
&\quad - D\left( W \left( \Psi_{RC_0} \otimes |0\rangle \langle 0|_{E_0} \otimes \xi_D^{a:n} \right) W^\dagger \otimes |0\rangle \langle 0|_Q \otimes \sigma_{C_1 E_1} \big| \Tr_{F_2} \left( U_\ell^\dagger \tau U_\ell \right) \right) \right). (15)
\end{align*}
\]

Now, we have

\[
\begin{align*}
\Tr_{F_2} \left( U_\ell^\dagger \tau U_\ell \right) \\
&= \frac{1}{N} W \left( \Psi_{RC_0} \otimes |0\rangle \langle 0|_{E_0} \otimes \xi_D^{a:n} \right) W^\dagger \otimes |0\rangle \langle 0|_Q \otimes \sigma_{C_1 E_1} \\
&+ \frac{1}{N} \sum_{m \in S, m \neq \ell} \Tr_{F_2} \left( U_{m-\ell} \left( W \left( \Psi_{RC_0} \otimes |0\rangle \langle 0|_{E_0} \otimes \xi_D^{a:n} \right) W^\dagger \otimes |0\rangle \langle 0|_Q \otimes \sigma_{C_1 E_1} \otimes \mu_{F_2} \right) U_{m-\ell}^\dagger \right) \\
&\leq \frac{1}{N} W \left( \Psi_{RC_0} \otimes |0\rangle \langle 0|_{E_0} \otimes \xi_D^{a:n} \right) W^\dagger \otimes |0\rangle \langle 0|_Q \otimes \sigma_{C_1 E_1} + \frac{(1 + 15\delta)(N - 1)}{N} \Psi_R \otimes \mu_{F_1} \otimes \xi_D^{1:n}.
\end{align*}
\]

Moreover, using the relation \( \Psi_{RC_0} \preceq 2^{k+1} \Psi_R \otimes \sigma_{C_0} \), Equation 6 and Claim 1, we conclude

\[
\begin{align*}
W \left( \Psi_{RC_0} \otimes |0\rangle \langle 0|_{E_0} \otimes \xi_D^{a:n} \right) W^\dagger \otimes |0\rangle \langle 0|_Q \otimes \sigma_{C_1 E_1} \\
&\preceq 2^{k+1} \Psi_R \otimes W \left( \sigma_{C_0} \otimes |0\rangle \langle 0|_{E_0} \otimes \xi_D^{a:n} \right) W^\dagger \otimes |0\rangle \langle 0|_Q \otimes \sigma_{C_1 E_1} \\
&\preceq 2^{k+1} (1 + 15\delta) \Psi_R \otimes \sigma_{C_0 E_0} \otimes |0\rangle \langle 0|_Q \otimes \sigma_{C_1 E_1} \otimes \xi_D^{1:n} \\
&\preceq 2^{k+2} (1 + 15\delta) \frac{|F|}{\supp(\sigma_{CE})^2} \Psi_R \otimes \mu_{F_1} \otimes \xi_D^{1:n} \\
&\preceq 2^{k+2} (1 + 15\delta) \Psi_R \otimes \mu_{F_1} \otimes \xi_D^{1:n}.
\end{align*}
\]

Using this in Equation 16 we conclude that

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

Along with Equation 15 this leads to

\[
D \left( \tau \big| \Psi_R \otimes \mu_{F_1} \otimes \xi_D^{1:n} \otimes \mu_{F_2} \right) \preceq 15\delta + \log \left( 1 + \frac{2^{k+2} - 1}{N} \right).
\]

This completes the proof. □
C Proof of Claim 4

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

\[
W|c\rangle_{C_0} |0\rangle_{E_0} |k\rangle_D = |c\rangle_{C_0} |k \mod b(c)\rangle_{E_0} |k/b(c)\rangle_D.
\]

Thus,

\[
\begin{align*}
\text{Tr}_{F_2} \left( U_m \left( W \left( |c\rangle \langle c'|_{C_0} \otimes |0\rangle_{E_0} \otimes \xi_D^{an} \right) W^\dagger \otimes |0\rangle \langle 0|_Q \otimes \sigma_{C_1 E_1} \otimes \mu_{F_2} \right) U_m^\dagger \right) \\
= & \sum_{k=a}^n \frac{1}{k} \text{Tr}_{F_2} \left( U_m \left( W \left( |c\rangle \langle c'|_{C_0} \otimes |0\rangle_{E_0} \otimes |k\rangle \langle k|_D \right) W^\dagger \otimes |0\rangle \langle 0|_Q \otimes \sigma_{C_1 E_1} \otimes \mu_{F_2} \right) U_m^\dagger \right) \\
= & \sum_{k=a}^n \frac{1}{k} \text{Tr}_{F_2} \left( U_m \left( |c\rangle \langle c'|_{C_0} \otimes |k \mod b(c)\rangle \langle k \mod b(c')|_{E_0} \otimes |0\rangle \langle 0|_Q \otimes \sigma_{C_1 E_1} \otimes \mu_{F_2} \right) U_m^\dagger \right) \\
& \otimes |[k/b(c)]\rangle \langle [k/b(c')]|_D \\
= & \mu_{F_1} \cdot \delta_{c,c'}.
\end{align*}
\]

As shown in Lemma 3,

\[
\text{Tr}_{F_2} \left( U_m \left( |c\rangle \langle c'|_{C_0} \otimes |k \mod b(c)\rangle \langle k \mod b(c')|_{E_0} \otimes |0\rangle \langle 0|_Q \otimes \sigma_{C_1 E_1} \otimes \mu_{F_2} \right) U_m^\dagger \right) = \mu_{F_1} \cdot \delta_{c,c'}.
\]

Hence, we conclude that

\[
\begin{align*}
\text{Tr}_{F_2} \left( U_m \left( W \left( |c\rangle \langle c'|_{C_0} \otimes |0\rangle_{E_0} \otimes \xi_D^{an} \right) W^\dagger \otimes |0\rangle \langle 0|_Q \otimes \sigma_{C_1 E_1} \otimes \mu_{F_2} \right) U_m^\dagger \right) \\
= & \mu_{F_1} \otimes \sum_{k=a}^n \frac{1}{k} \langle [k/b(c)] \langle [k/b(c')]|_D \cdot \delta_{c,c'} \\
\leq & \left( 1 + 15\delta \right) \mu_{F_1} \otimes \xi_D^{an} \cdot \delta_{c,c'},
\end{align*}
\]

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

\[
\text{Tr}_{F_2} \tau_m = \sum_{c,c'} \Psi_R^{(c,c')} \otimes \text{Tr}_{F_2} \left( U_m \left( W \left( |c\rangle \langle c'|_{C_0} \otimes |0\rangle_{E_0} \otimes \xi_D^{an} \right) W^\dagger \otimes |0\rangle \langle 0|_Q \otimes \sigma_{C_1 E_1} \otimes \mu_{F_2} \right) U_m^\dagger \right) \\
\leq & \left( 1 + 15\delta \right) \sum_c \Psi_R^{(c,c)} \otimes \mu_{F_1} \otimes \xi_D^{an}.
\]

This completes the proof.

\( \square \)