Communications in Mathematical Physics

Reliability Function of Quantum Information Decoupling via the Sandwiched Rényi Divergence

--Manuscript Draft--

Manuscript Number:	CIMP-D-22-00372
Full Title:	Reliability Function of Quantum Information Decoupling via the Sandwiched Rényi Divergence
Article Type:	Original research
Abstract:	Quantum information decoupling is a fundamental quantum information processing task, which also serves as a crucial tool in a diversity of topics in quantum physics. In this paper, we characterize the reliability function of catalytic quantum information decoupling, that is, the best exponential rate under which perfect decoupling is asymptotically approached. We have obtained the exact formula when the decoupling cost is below a critical value. In the situation of high cost, we provide upper and lower bounds. This result is then applied to quantum state merging, exploiting its inherent connection to decoupling. In addition, as technical tools, we derive the exact exponents for the smoothing of the conditional minentropy and max-information, and we prove a novel bound for the convex-split lemma. Our results are given in terms of the sandwiched Rényi divergence, providing it with a new type of operational meaning in characterizing how fast the performance of quantum information tasks approaches the perfect.

Reliability Function of Quantum Information Decoupling via the Sandwiched Rényi Divergence

Ke Li[1,*] and Yongsheng Yao[1,2,†]

[1]*Institute for Advanced Study in Mathematics, Harbin Institute of Technology, Harbin 150001, China*

[2]*School of Mathematics, Harbin Institute of Technology, Harbin 150001, China*

Quantum information decoupling is a fundamental quantum information processing task, which also serves as a crucial tool in a diversity of topics in quantum physics. In this paper, we characterize the reliability function of catalytic quantum information decoupling, that is, the best exponential rate under which perfect decoupling is asymptotically approached. We have obtained the exact formula when the decoupling cost is below a critical value. In the situation of high cost, we provide upper and lower bounds. This result is then applied to quantum state merging, exploiting its inherent connection to decoupling. In addition, as technical tools, we derive the exact exponents for the smoothing of the conditional min-entropy and max-information, and we prove a novel bound for the convex-split lemma.

Our results are given in terms of the sandwiched Rényi divergence, providing it with a new type of operational meaning in characterizing how fast the performance of quantum information tasks approaches the perfect.

I. INTRODUCTION

Quantum information decoupling [1–3] is the procedure of removing the information of a reference system from the system under control, via physically permitted operations. It is a fundamental quantum information processing task, which has found broad applications, ranging from quantum Shannon theory [1–9] to quantum thermodynamics [10–14] to black-hole physics [15–18]. Since being introduced in [1–3], the problem of quantum information decoupling has attracted continued interest of study from the community. This includes the study of decoupling in the one-shot setting [7, 19–23], the search for more specific and more efficient decoupling operations [24–26], and the investigation of the speed of asymptotic convergence of the decoupling performance [23, 27, 28]. In particular, by introducing an independent system as catalyst, tight one-shot characterization has been derived in [20, 21], which is able to provide the exact second-order asymptotics.

The *reliability function* was introduced by Shannon in information theory [29]. Defined as the rate of exponential decay of the error with the increasing of blocklength, the reliability function provides the desired precise characterization of how rapidly an information processing task approaches the perfect in the asymptotic setting [30]. Study of reliability functions in quantum information dates back to the work of Holevo and Winter [31–33] more than two decades ago. In recent years, there has been a growing body of research in this topic from the quantum community [23, 34–38]. However, complete characterization of the reliability functions in the quantum regime is not known, even for classical-quantum

* carl.ke.lee@gmail.com

† yongsh.yao@gmail.com

channels. Nevertheless, see References [38–44] for a partial list of the fruitful results on the strong converse exponent in the quantum setting, which characterizes how fast a quantum information task becomes the useless.

In this paper, we investigate the reliability function for the task of quantum information decoupling in the catalytic setting. We have obtained the exact formula when the decoupling cost is below a critical value. Specifically, for a bipartite quantum state ρ_{RA}, we consider three different types of decoupling operations on the A system: (a) decoupling via removing a subsystem, (b) decoupling via projective measurement, (c) decoupling via random unitary operation. We show that under any of these three types of decoupling operations, the reliability function is given by the Legendre transformation of the *sandwiched Rényi mutual information*

$$I_\alpha(R:A)_\rho := \min_{\sigma_A \in \mathcal{S}(A)} D_\alpha(\rho_{RA} \| \rho_R \otimes \sigma_A)$$

of order $\alpha \in (1, 2]$. Here $\mathcal{S}(A)$ is the set of all quantum states on system A, and

$$D_\alpha(M\|N) := \frac{1}{\alpha - 1} \log \operatorname{Tr} \left(N^{\frac{1-\alpha}{2\alpha}} M N^{\frac{1-\alpha}{2\alpha}} \right)^\alpha$$

is the *sandwiched Rényi divergence* [41, 45]. This result is obtained by deriving respective upper and lower bounds, and we show that the two bounds coincide when the rate of the decoupling cost is below the critical value. On the one hand, we analyse the convex-split lemma of [20] and derive for it a novel bound in terms of the sandwiched Rényi divergence of order $\alpha \in (1, 2]$, and this constitutes the main technical tool for proving the lower bound. On the other hand, the upper bound is obtained based on an asymptotic analysis of the smoothing quantity of the max-information, for which we show that the exact exponent is given by a formula in terms of the sandwiched Rényi mutual information of order $\alpha \in (1, \infty)$. Furthermore, as application, we provide similar characterization for the reliability function of quantum state merging by exploiting the inherent connection between quantum state merging and decoupling.

Our results, along with the concurrent work of [46] which addresses different problems, have provided the sandwiched Rényi divergence [41, 45] with a new type of operational interpretation by showing that it characterizes the exact exponents under which certain quantum information tasks approach the perfect. This is in stark contrast to what was previously known that the sandwiched Rényi divergence characterizes the strong converse exponents—the optimal exponential rates under which the underlying errors go to 1 [38, 43, 44, 47–49]. Therefore, we conclude that the meaning of this fundamental entropic quantity can be more fruitful than what was previously understood.

Relation to previous works. In References [27] and [23], exponential achievability bounds for the decoupling error were given, which are in terms of the sandwiched Rényi divergence of order $\alpha \in (1, 2]$, too. However, these bounds do not seem to be able to yield the optimal exponent in the asymptotic setting, and thus do not provide much information on the reliability function. Besides, there is no discuss on the converse bound for the exponent of the decoupling error there. Another difference between these two works and the present one is that, in [27] and [23] the decoupling error is measured using the trace distance, while in the present paper we employ the purified distance, or equivalently, the fidelity function.

The reminder of this paper is organized as follows. In Section II we introduce the necessary notation, definitions and some basic properties. In Section III we present the problem formulation, the main results, and the application to quantum state merging. The proofs are given in Section IV and Section V, where in Section IV we prove the characterization of the reliability functions, and in Section V we prove the relation between different types of decoupling as well as the relation between decoupling and quantum state merging. At last, in Section VI we conclude the paper with some discussion and open questions.

II. PRELIMINARIES

A. Notation and basic properties

Let $\mathcal{H}$ be a Hilbert space, and $\mathcal{H}_A$ be the Hilbert space associated with system A. $\mathcal{H}_{AB}$, denoting the Hilbert space of the composite system AB, is the tensor product of $\mathcal{H}_A$ and $\mathcal{H}_B$. We restrict ourselves to finite-dimensional Hilbert spaces throughout this paper. The notation $|A|$ stands for the dimension of $\mathcal{H}_A$. We use $\mathbb{1}_A$ to denote the identity operator on $\mathcal{H}_A$. the notation $\operatorname{supp}(X)$ for an operator X is used for the support of X. The set of unitary operators on $\mathcal{H}$ is denoted as $\mathcal{U}(\mathcal{H})$, and the set of positive semidefinite operators on $\mathcal{H}$ is denoted as $\mathcal{P}(\mathcal{H})$. The set of normalized quantum states and subnormalized quantum states on $\mathcal{H}$ are denoted as $\mathcal{S}(\mathcal{H})$ and $\mathcal{S}_{\leq}(\mathcal{H})$, respectively. That is,

$$\mathcal{S}(\mathcal{H}) = \{\rho \in \mathcal{P}(\mathcal{H}) | \operatorname{Tr}\rho = 1\},$$
$$\mathcal{S}_{\leq}(\mathcal{H}) = \{\rho \in \mathcal{P}(\mathcal{H}) | \operatorname{Tr}\rho \leq 1\}.$$

If the Hilbert space $\mathcal{H}$ is associated with system A, then the above notations $\mathcal{U}(\mathcal{H})$, $\mathcal{P}(\mathcal{H})$, $\mathcal{S}(\mathcal{H})$ and $\mathcal{S}_{\leq}(\mathcal{H})$ are also written as $\mathcal{U}(A)$, $\mathcal{P}(A)$, $\mathcal{S}(A)$ and $\mathcal{S}_{\leq}(A)$, respectively. The discrete Weyl operators on a d-dimensional Hilbert space $\mathcal{H}$ with an orthonormal basis $\{|a\rangle\}_{a=0}^{d-1}$ are a collection of unitary operators

$$W_{a,b} = \sum_{c=0}^{d-1} \mathrm{e}^{\frac{2\pi i b c}{d}} |(a+c) \mod d\rangle\langle c|,$$

where $a, b \in \{0, 1, \ldots, d-1\}$.

For $X, Y \in \mathcal{P}(\mathcal{H})$, we write $X \geq Y$ if $X - Y \in \mathcal{P}(\mathcal{H})$ and $X \leq Y$ if $Y - X \in \mathcal{P}(\mathcal{H})$. $\{X \geq Y\}$ is the spectral projection of $X - Y$ corresponding to all non-negative eigenvalues. $\{X > Y\}$, $\{X \leq Y\}$ and $\{X < Y\}$ are similarly defined.

We use the purified distance [50, 51] to measure the closeness of a pair of states $\rho, \sigma \in \mathcal{S}_{\leq}(\mathcal{H})$. The purified distance is defined as $P(\rho,\sigma) := \sqrt{1 - F^2(\rho,\sigma)}$, where

$$F(\rho,\sigma) := \operatorname{Tr}\sqrt{\sqrt{\sigma}\rho\sqrt{\sigma}} + \sqrt{(1-\operatorname{Tr}\rho)(1-\operatorname{Tr}\sigma)}$$

is the fidelity. The Uhlmann's theorem [52], stated as follows, will play a key role in later proofs. Let $\rho_{AB} \in \mathcal{S}_{\leq}(\mathcal{H}_{\mathcal{AB}})$ be a bipartite state, and let $\sigma_A \in \mathcal{S}_{\leq}(\mathcal{H}_{\mathcal{A}})$. Then there exists an extension σ_{AB} of σ_A such that $P(\rho_{AB}, \sigma_{AB}) = P(\rho_A, \sigma_A)$.

A quantum operation or quantum channel Φ is a linear, completely positive, and trace-preserving (CPTP) map acting on quantum states. We denote by $\Phi_{A\to B}$ a quantum operation from system A to system B. The Stinespring representation theorem [53] states that there is an ancillary system C in a pure state $|0\rangle\langle 0|_C$, a system E and a unitary $U_{AC\to BE}$ such that $\Phi_{A\to B}(\rho_A) = \operatorname{Tr}_E\left(U(\rho_A \otimes |0\rangle\langle 0|_C)U^*\right)$. A quantum measurement is described by a set of positive semidefinite operators $\{M_x\}_x$ such that $\sum_x M_x = \mathbb{1}$. It outputs x with probability $\mathrm{Tr}(\rho M_x)$ when the underlying state is ρ. If a measurement $\{Q_x\}_x$ is such that all the Q_x are projections onto mutually orthogonal subspaces, then it is called a projective measurement. We associate each quantum measurement $\mathcal{M} = \{M_x\}_x$ with a measurement channel $\Phi_{\mathcal{M}} : \rho \mapsto \sum_x (\mathrm{Tr}\, \rho M_x)|x\rangle\langle x|$, where $\{|x\rangle\}$ is an orthonormal basis.

Let σ be a self-adjoint operator on $\mathcal{H}$ with spectral projections $Q_1, \dots, Q_{v(\sigma)}$, where $v(\sigma)$ is the number of different eigenvalues of σ. The associated pinching map $\mathcal{E}_\sigma$ is defined as

$$\mathcal{E}_\sigma : X \to \sum_i Q_i X Q_i.$$

The pinching inequality [54] states that for any $X \in \mathcal{P}(\mathcal{H})$,

$$X \leq v(\sigma)\mathcal{E}_\sigma(X). \tag{1}$$

For $n \in \mathbb{N}$, let S_n be the symmetric group of the permutations of n elements. The set of symmetric states and subnormalized symmetric states on $\mathcal{H}_{A^n}$ are defined, respectively, as

$$\begin{aligned}
\mathcal{S}^{\mathrm{sym}}(A^n) &:= \big\{\rho_{A^n} | \rho_{A^n} \in \mathcal{S}(A^n),\ W_\pi \rho_{A^n} W_\pi^* = \rho_{A^n},\ \forall \pi \in S_n\big\},\\
\mathcal{S}^{\mathrm{sym}}_{\leq}(A^n) &:= \big\{\rho_{A^n} | \rho_{A^n} \in \mathcal{S}_{\leq}(A^n),\ W_\pi \rho_{A^n} W_\pi^* = \rho_{A^n},\ \forall \pi \in S_n\big\},
\end{aligned}$$

where $W_\pi : |\psi_1\rangle \otimes \cdots \otimes |\psi_n\rangle \mapsto |\psi_{\pi^{-1}(1)}\rangle \otimes \cdots \otimes |\psi_{\pi^{-1}(n)}\rangle$ is the natural representation of $\pi \in S_n$. The set of the symmetric states can be dominated by a single symmetric state, in the sense of the following Lemma 1, and two different constructions are given in [55] and [56], respectively. See [44, Appendix A] for a detailed proof.

Lemma 1 *For every Hilbert space $\mathcal{H}_A$ and $n \in \mathbb{N}$, there exists a universal symmetric state $\omega^{(n)}_{A^n} \in \mathcal{S}^{\mathrm{sym}}(A^n)$, such that for any $\rho_{A^n} \in \mathcal{S}^{\mathrm{sym}}(A^n)$ we have*

$$\begin{aligned}
\rho_{A^n} &\leq g_{n,|A|}\omega^{(n)}_{A^n},\\
v(\omega^{(n)}_{A^n}) &\leq (n+1)^{|A|-1},
\end{aligned}$$

where $g_{n,|A|} \leq (n+1)^{\frac{(|A|+2)(|A|-1)}{2}}$ and $v(\omega^{(n)}_{A^n})$ denotes the number of different eigenvalues of $\omega^{(n)}_{A^n}$.

Throughout this paper, log is with base 2 and ln is with base e.

B. Quantum entropies and information divergences

The sandwiched Rényi divergence has been introduced in [41, 45] and is a quantum generalization of the classical Rényi information divergence. For quantum states $\rho, \sigma \in \mathcal{S}(\mathcal{H})$ and a parameter $\alpha \in (0,1) \cup (1,\infty)$, it is defined as

$$D_\alpha(\rho\|\sigma) := \frac{1}{\alpha-1} \log \mathrm{Tr}\left(\sigma^{\frac{1-\alpha}{2\alpha}} \rho \sigma^{\frac{1-\alpha}{2\alpha}}\right)^\alpha$$

if either $\alpha > 1$ and $\mathrm{supp}(\rho) \subseteq \mathrm{supp}(\sigma)$ or $\alpha < 1$ and $\mathrm{supp}(\rho) \not\perp \mathrm{supp}(\sigma)$, otherwise we set $D_\alpha(\rho\|\sigma) = +\infty$.

For a bipartite quantum state $\rho_{AB} \in \mathcal{S}(AB)$ and $\alpha \in (0,1) \cup (1,\infty)$, the sandwiched Rényi mutual information of order α is defined as [41, 57]

$$I_\alpha(A:B)_\rho := \min_{\sigma_B \in \mathcal{S}(B)} D_\alpha(\rho_{AB} \| \rho_A \otimes \sigma_B),$$

and we consider a version of the sandwiched Rényi conditional entropy [45]

$$H_\alpha(A|B)_\rho := -\min_{\sigma_B \in \mathcal{S}(B)} D_\alpha(\rho_{AB} \| \mathbb{1}_A \otimes \sigma_B).$$

when the system B is of dimension 1, we recover from the sandwiched Rényi conditional entropy the Rényi entropy $H_\alpha(A)_\rho := -D_\alpha(\rho_A \| \mathbb{1}_A) = \frac{1}{1-\alpha} \log \operatorname{Tr} \rho_A^\alpha$.

The quantum relative entropy [58]

$$D(\rho\|\sigma) := \begin{cases} \operatorname{Tr}(\rho(\log\rho - \log\sigma)) & \text{if } \operatorname{supp}(\rho) \subseteq \operatorname{supp}(\sigma), \\ +\infty & \text{otherwise} \end{cases}$$

of states ρ and σ is the limit of the sandwiched Rényi divergence when $\alpha \to 1$. In the case $\alpha \to \infty$, we get the max-relative entropy [59]

$$D_{\max}(\rho\|\sigma) := \inf\{\lambda \mid \rho \le 2^\lambda \sigma\}.$$

The limits of $I_\alpha(A:B)_\rho$ and $H_\alpha(A|B)_\rho$ when $\alpha \to 1$ are the quantum mutual information and quantum conditional entropy, respectively: for $\rho_{AB} \in \mathcal{S}(AB)$,

$$\begin{aligned} I(A:B)_\rho &:= \min_{\sigma_B \in \mathcal{S}(B)} D(\rho_{AB} \| \rho_A \otimes \sigma_B) = D(\rho_{AB} \| \rho_A \otimes \rho_B), \\ H(A|B)_\rho &:= -\min_{\sigma_B \in \mathcal{S}(B)} D(\rho_{AB} \| \mathbb{1}_A \otimes \sigma_B) = -D(\rho_{AB} \| \mathbb{1}_A \otimes \rho_B). \end{aligned}$$

The max-information [60] for $\rho_{AB} \in \mathcal{S}(AB)$,

$$I_{\max}(A:B)_\rho := \max_{\sigma_B \in \mathcal{S}(B)} D_{\max}(\rho_{AB} \| \rho_A \otimes \sigma_B),$$

emerges as the limit of $I_\alpha(A:B)_\rho$ when $\alpha \to \infty$.

III. PROBLEM STATEMENT AND MAIN RESULTS

A. Catalytic quantum information decoupling

Let a bipartite quantum state $\rho_{RA} \in \mathcal{S}(RA)$ be given. Quantum information decoupling is the procedure of removing the information about system R from the A system, by performing a quantum operation on the A system. In catalytic quantum information decoupling, an auxiliary system A' in a sate that is independent of the state ρ_{RA} can be added as a catalyst during the decoupling operation. Readers are referred to [21] for a detailed description of standard and catalytic decoupling. We consider three different types of decoupling operations: (a) decoupling via removing a subsystem [3], (b) decoupling via projective measurement [2], (c) decoupling via random unitary operation [61].

(a) decoupling via removing a subsystem. For quantum state ρ_{RA}, a catalytic decoupling scheme via removing a subsystem consists of a catalytic system A' in the state $\sigma_{A'}$ and an isometry operation $U : \mathcal{H}_{AA'} \to \mathcal{H}_{A_1A_2}$, where A_2 is the system to be removed and A_1 is the remaining system. Without loss of generality, we require that $|AA'| = |A_1A_2|$ and hence $U_{AA' \to A_1A_2}$ becomes unitary. The cost of the decoupling is given by the number of qubits that is removed, $\log|A_2|$. The performance is measured by the purified distance between the remaining state $\mathrm{Tr}_{A_2} U(\rho_{RA} \otimes \sigma_{A'})U^*$ and the nearest product state of the form $\rho_R \otimes \omega_{A_1}$. We are interested in the optimal performance when the cost, namely, the number of removed qubits, is bounded.

Definition 2 *Let $\rho_{RA} \in \mathcal{S}(RA)$ be a bipartite quantum state. For a given size of removed system $k \geq 0$ (in qubits), the optimal performance of catalytic decoupling via removing a subsystem is given by*

$$P^{\mathrm{dec}}_{R:A}(\rho_{RA}, k) := \min P\big(\mathrm{Tr}_{A_2} U(\rho_{RA} \otimes \sigma_{A'})U^*, \rho_R \otimes \omega_{A_1}\big), \tag{2}$$

where the minimization is over all system dimensions $|A'|$, $|A_1|$, $|A_2|$ such that $|AA'| = |A_1A_2|$ and $\log|A_2| \leq k$, all states $\sigma_{A'} \in \mathcal{S}(A')$, $\omega_{A_1} \in \mathcal{S}(A_1)$, and all unitary operations $U : \mathcal{H}_{AA'} \to \mathcal{H}_{A_1A_2}$.

(b) decoupling via projective measurement. For quantum state ρ_{RA}, a catalytic decoupling scheme via projective measurement consists of a catalytic system A' in the state $\sigma_{A'}$ and an projective measurement $\mathcal{Q} = \{Q^x\}_{x=1}^m$ on the composite system AA'. The cost of the decoupling is accounted by the number (in bits) of measurement outcomes, $\log m$. The performance is measured by the purified distance between the post-measurement state $\sum_{x=1}^m Q^x_{AA'}(\rho_{RA} \otimes \sigma_{A'})Q^x_{AA'}$ and the nearest product state of the form $\rho_R \otimes \omega_{AA'}$. We are interested in the optimal performance when the cost, namely, the number of measurement outcomes, is bounded.

Definition 3 *Let $\rho_{RA} \in \mathcal{S}(RA)$ be a bipartite quantum state. For a given number of measurement outcomes $k \geq 0$ (in bits), the optimal performance of catalytic decoupling via projective measurement is given by*

$$P^{\mathrm{dec\text{-}m}}_{R:A}(\rho_{RA}, k) := \min P\Big(\sum_{x=1}^m Q^x_{AA'}(\rho_{RA} \otimes \sigma_{A'})Q^x_{AA'}, \rho_R \otimes \omega_{AA'}\Big), \tag{3}$$

where the minimization is over all system dimensions $|A'|$, all states $\sigma_{A'} \in \mathcal{S}(A')$, $\omega_{AA'} \in \mathcal{S}(AA')$, and all projective measurements $\{Q^x_{AA'}\}_{x=1}^m$ such that $\log m \leq k$.

(c) decoupling via random unitary operation. For quantum state ρ_{RA}, a catalytic decoupling scheme via random unitary operation consists of a catalytic system A' in the state $\sigma_{A'}$ and a random unitary operation $\Lambda_{AA'} : X \mapsto \frac{1}{m}\sum_{i=1}^m U_i X U_i^*$ acting on the composite system AA'. The cost of the decoupling is accounted by the number (in bits) of unitary operators in $\Lambda_{AA'}$, $\log m$. The performance is measured by the purified distance between the resulting state $\Lambda_{AA'}(\rho_{RA} \otimes \sigma_{A'})$ and the nearest product state of the form $\rho_R \otimes \omega_{AA'}$. We are interested in the optimal performance when the cost, namely, the number of unitary operators, is bounded.

Definition 4 *Let $\rho_{RA} \in \mathcal{S}(RA)$ be a bipartite quantum state. For a given number of unitary operators $k \geq 0$ (in bits), the optimal performance of catalytic decoupling via random unitary operation is given by*

$$P^{\text{dec-u}}_{R:A}(\rho_{RA}, k) := \min P\Big(\Lambda_{AA'}(\rho_{RA} \otimes \sigma_{A'}), \rho_R \otimes \omega_{AA'}\Big), \tag{4}$$

where the minimization is over all system dimensions $|A'|$, all states $\sigma_{A'} \in \mathcal{S}(A')$, $\omega_{AA'} \in \mathcal{S}(AA')$ and all random unitary operations $\Lambda_{AA'}(\cdot) = \frac{1}{m}\sum_{i=1}^{m} U_i(\cdot)U_i^*$ with $U_i \in \mathcal{U}(AA')$ such that $\log m \leq k$.

The reliability function of quantum information decoupling characterizes the the speed at which perfect decoupling can be approached in the asymptotic setting, in which the underlying bipartite quantum state is in the form of tensor product of n identical copies. Specifically, it is the rate of exponential decreasing of the optimal performance, as a function of the cost.

Definition 5 *Let $\rho_{RA} \in \mathcal{S}(RA)$ be a bipartite quantum state, and $r \geq 0$. The reliability functions of catalytic quantum information decoupling for the state ρ_{RA}, via the three different types of decoupling operations described above, are defined respectively as*

$$E^{\text{dec}}_{R:A}(\rho_{RA}, r) := \limsup_{n\to\infty} \frac{-1}{n} \log P^{\text{dec}}_{R^n:A^n}\big(\rho_{RA}^{\otimes n}, nr\big), \tag{5}$$

$$E^{\text{dec-m}}_{R:A}(\rho_{RA}, r) := \limsup_{n\to\infty} \frac{-1}{n} \log P^{\text{dec-m}}_{R^n:A^n}\big(\rho_{RA}^{\otimes n}, nr\big), \tag{6}$$

$$E^{\text{dec-u}}_{R:A}(\rho_{RA}, r) := \limsup_{n\to\infty} \frac{-1}{n} \log P^{\text{dec-u}}_{R^n:A^n}\big(\rho_{RA}^{\otimes n}, nr\big). \tag{7}$$

B. Main results

At first, we show in the following Proposition 6 equalities that relate the optimal performances or the reliability functions for the three different decoupling operations. With this, we are able to deal with them in a unified way.

Proposition 6 *For $\rho_{RA} \in \mathcal{S}(RA)$ and $k, r \geq 0$, we have*

$$P^{\text{dec}}_{R:A}(\rho_{RA}, k) = P^{\text{dec-m}}_{R:A}(\rho_{RA}, 2k) = P^{\text{dec-u}}_{R:A}(\rho_{RA}, 2k), \tag{8}$$

$$E^{\text{dec}}_{R:A}(\rho_{RA}, r) = E^{\text{dec-m}}_{R:A}(\rho_{RA}, 2r) = E^{\text{dec-u}}_{R:A}(\rho_{RA}, 2r). \tag{9}$$

Then, we derive a one-shot achievability bound for the performance of catalytic quantum information decoupling. The bound is given in terms of the sandwiched Rényi information divergence. In doing so, we have employed the convex-split lemma [20] as a key technical tool. While we present the bound only for the case of decoupling via removing a subsystem in the following Theorem 7, similar results hold for the other two cases in light of Proposition 6.

Theorem 7 *Let $\rho_{RA} \in \mathcal{S}(RA)$. For any $m \in \mathbb{N}$, $0 < s \leq 1$ and $\sigma_A \in \mathcal{S}(A)$, the optimal performance of decoupling A from R is bounded as*

$$P^{\text{dec}}_{R:A}(\rho_{RA}, \log m) \leq \sqrt{\frac{v^s}{s}} \exp\Big\{ -(\ln 2)\ s\Big(\log m - \frac{1}{2} D_{1+s}(\rho_{RA} \| \rho_R \otimes \sigma_A)\Big)\Big\},$$

where v is the number of distinct eigenvalues of $\rho_R \otimes \sigma_A$.

Our main result is the characterization of reliability functions. This is given in Theorem 8 for the case of decoupling via removing a subsystem. Thanks to Proposition 6, similar results follow directly for the other two cases. We mention that we have completely determined the reliability functions when the respective cost is below a critical value.

Theorem 8 *Let $\rho_{RA} \in \mathcal{S}(RA)$ be a bipartite quantum state, and consider the problem of decoupling quantum information in A^n from the reference system R^n for the quantum state $\rho_{RA}^{\otimes n}$. When $r \leq R_{\text{critical}} := \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}s} sI_{1+s}(R:A)_\rho\big|_{s=1}$, we have*

$$E^{\text{dec}}_{R:A}(\rho_{RA}, r) = \max_{0\leq s\leq 1}\left\{s\Big(r - \frac{1}{2}I_{1+s}(R:A)_\rho\Big)\right\}. \tag{10}$$

In general, we have

$$E^{\text{dec}}_{R:A}(\rho_{RA}, r) \geq \max_{0\leq s\leq 1}\left\{s\Big(r - \frac{1}{2}I_{1+s}(R:A)_\rho\Big)\right\}, \tag{11}$$

$$E^{\text{dec}}_{R:A}(\rho_{RA}, r) \leq \sup_{s\geq 0}\left\{s\Big(r - \frac{1}{2}I_{1+s}(R:A)_\rho\Big)\right\}. \tag{12}$$

The lower bound of Eq. (11) is a consequence of Theorem 7. For the upper bound of Eq. (12), we first bound the optimal decoupling performance using the smoothing quantity associated with the max-information, and then we derive the exact exponent in smoothing the max-information. Eq. (10) is derived from the combination of Eqs. (11) and (12).

The results presented in Theorem 8 are depicted in Figure 1. Above the critical value, we are unable to determine the formula for the reliability function. This is indeed a common hard problem in the topic of reliability functions (see, e.g., References [30, 34, 46, 62]). More comments on Theorem 8 can be found in Remark 20.

C. Applications to quantum state merging

The inherent connection between quantum information decoupling and quantum state merging has been established since the invention of these two tasks [1–3]. It was further explored later on, e.g., in [19–21, 28]. Exploiting this connection, we are able to extend the results on the reliability function of quantum information decoupling to the scenario of quantum state merging.

Let ρ_{RAB} be a tripartite pure state. Alice, Bob and a referee hold system A, B and R respectively. Quantum state merging is the task of transmitting the quantum information stored in the A system from Alice to Bob. There are two different ways to achieve this. One is by classical communication and is introduced in [1]. The other one is by quantum communication, firstly considered in [3]. In both cases, free pre-shared entanglement between Alice and Bob is allowed. A formal description is as follows.

A quantum state merging protocol via quantum communication, $\mathcal{M}_1$, consists of using a shared bipartite entangled pure state $\phi_{A'B'}$, Alice applying local unitary $U_{AA'\to A_1A_2}$ and sending the system A_2 to Bob, Bob applying local unitary $V_{A_2BB'\to ABB_1}$ and they discarding the systems A_1 and B_1. A CPTP map $\mathcal{M}_2$ is a quantum state merging protocol via classical communication if it consists of using a shared bipartite entangled pure state $\psi_{A'B'}$, applying local operation at Alice's side, sending k classical bits from Alice to Bob and applying local operation to reproduce systems A and B at Bob's side. The performances of both protocols

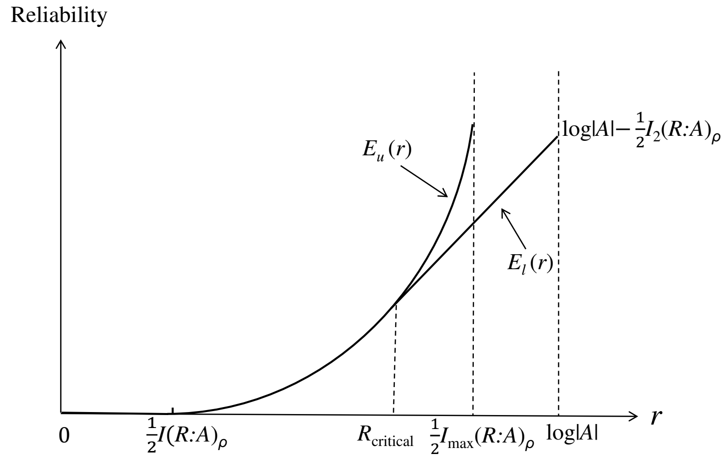

FIG. 1. Reliability function of quantum information decoupling. $E_u(r) := \sup_{s\geq 0}\left\{s\big(r - \frac{1}{2}I_{1+s}(R:A)_\rho\big)\right\}$ is the upper bound of Eq. (12). $E_l(r) := \sup_{0\leq s\leq 1}\left\{s\big(r - \frac{1}{2}I_{1+s}(R:A)_\rho\big)\right\}$ is the lower bound of Eq. (11). The two bounds are equal in the interval $[0, R_{\text{critical}}]$, giving the exact reliability function. The reliability function equals 0 when $r < \frac{1}{2}I(R:A)_\rho$ and it is strictly positive when $r > \frac{1}{2}I(R:A)_\rho$. Above the critical value R_{critical}, the upper bound $E_u(r)$ becomes larger than the lower bound and it is ∞ when $r > \frac{1}{2}I_{\max}(R:A)_\rho$, while the lower bound $E_l(r)$ becomes linear and reaches $\log|A| - \frac{1}{2}I_2(R:A)_\rho$ at $r = \log|A|$.

are given by the purified distance between ρ_{RAB} and the final state on the referee's system R and Bob's system A and B. The cost of state merging that we are concerned with, is the number of qubits ($\log|A_2|$ in $\mathcal{M}_1$) or classical bits (k in $\mathcal{M}_2$) that Alice sends to Bob.

Definition 9 *Let $\rho_{RAB} \in \mathcal{S}(RAB)$ be a tripartite pure state and $r \geq 0$. Let $P^{\text{merg}}_{A\Rightarrow B}(\rho_{RAB}, k)$ denote the optimal performance of quantum state merging via quantum communication of at most k qubits, and let $P^{\text{merg}}_{A\to B}(\rho_{RAB}, k)$ denote the optimal performance of quantum state merging via classical communication of at most k bits. They are defined respectively as*

$$P^{\text{merg}}_{A\Rightarrow B}(\rho_{RAB}, k) := \min_{\mathcal{M}_1} P(\mathcal{M}_1(\rho_{RAB}), \rho_{RAB}), \tag{13}$$

$$P^{\text{merg}}_{A\to B}(\rho_{RAB}, k) := \min_{\mathcal{M}_2} P(\mathcal{M}_2(\rho_{RAB}), \rho_{RAB}), \tag{14}$$

where $\mathcal{M}_1$ is the protocol via quantum communication and the minimization in Eq. (13) is over all such protocols whose communication cost is bounded by k qubits, and $\mathcal{M}_2$ is the protocol via classical communication and the minimization in Eq. (14) is over all possible $\mathcal{M}_2$ whose communication cost is bound by k bits.

The reliability function of quantum state merging characterizes the rate of exponential decreasing of the optimal performance in the asymptotic limit.

Definition 10 *Let $\rho_{RAB} \in \mathcal{S}(RAB)$ be a tripartite pure state and $r \geq 0$. The reliability functions $E^{\text{merg}}_{A\Rightarrow B}(\rho_{RAB}, r)$ and $E^{\text{merg}}_{A\to B}(\rho_{RAB}, r)$ of quantum state merging via quantum communication and classical communication respectively, are defined as*

$$E^{\text{merg}}_{A\Rightarrow B}(\rho_{RAB}, r) := \limsup_{n\to\infty} \frac{-1}{n} \log P^{\text{merg}}_{A^n\Rightarrow B^n}(\rho^{\otimes n}_{RAB}, nr), \tag{15}$$

$$E^{\text{merg}}_{A\to B}(\rho_{RAB}, r) := \limsup_{n\to\infty} \frac{-1}{n} \log P^{\text{merg}}_{A^n\to B^n}(\rho^{\otimes n}_{RAB}, nr). \tag{16}$$

To apply our results on decoupling to the problem of quantum state merging, we show exact equalities relating the optimal performances (or reliability functions) of catalytic decoupling and quantum state merging. Eq. (17) in Proposition 11 is essentially due to Uhlmann's theorem and is implicitly used in previous works [1–3, 19–21, 28]. However, our definition of the optimal decoupling performance (cf. Definition 2) is subtle to enable such an equality relation.

Proposition 11 *For a tripartite pure state $\rho_{RAB} \in \mathcal{S}(RAB)$ and $k \geq 0$, $r \geq 0$, we have*

$$P^{\mathrm{merg}}_{A\Rightarrow B}(\rho_{RAB}, k) = P^{\mathrm{merg}}_{A\to B}(\rho_{RAB}, 2k) = P^{\mathrm{dec}}_{R:A}(\rho_{RA}, k), \tag{17}$$
$$E^{\mathrm{merg}}_{A\Rightarrow B}(\rho_{RAB}, r) = E^{\mathrm{merg}}_{A\to B}(\rho_{RAB}, 2r) = E^{\mathrm{dec}}_{R:A}(\rho_{RA}, r). \tag{18}$$

With Proposition 11, we immediately obtain results in analogy to Theorem 7 and Theorem 8 for quantum state merging. We do not lay them out entirely but only exhibit the following.

Corollary 12 *Let $\rho_{RAB} \in \mathcal{S}(RAB)$ be a tripartite pure state. When the rate of qubits transmission r is such that $0 \leq r \leq R_{\mathrm{critical}} \equiv \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}s} sI_{1+s}(R:A)_\rho\big|_{s=1}$, the reliability function of quantum state merging via quantum communication is given by*

$$E^{\mathrm{merg}}_{A\Rightarrow B}(\rho_{RAB}, r) = \max_{0\leq s\leq 1}\left\{ s\Big(r - \frac{1}{2}I_{1+s}(R:A)_\rho\Big)\right\}; \tag{19}$$

when the rate of classical bits transmission r is such that $0 \leq \frac{r}{2} \leq R_{\mathrm{critical}} \equiv \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}s} sI_{1+s}(R:A)_\rho\big|_{s=1}$, the reliability function of quantum state merging via classical communication is given by

$$E^{\mathrm{merg}}_{A\to B}(\rho_{RAB}, r) = \max_{0\leq s\leq 1}\left\{ \frac{s}{2}\Big(r - I_{1+s}(R:A)_\rho\Big)\right\}. \tag{20}$$

A special case where Bob holds no side information is worth looking at. It can be understood as entanglement-assisted quantum source coding. In this case, Alice and a referee share a pure state ρ_{RA} and Alice wants to send the information in the A system to Bob with the assistance of unlimited entanglement and using noiseless quantum or classical communication. It has been shown in [63, Lemma 8] (see also [64, Proposition 4.86]) that

$$I_{1+s}(R:A)_\rho = \frac{2s+1}{s}\log \operatorname{Tr} \rho_A^{\frac{1}{2s+1}} = 2H_{\frac{1}{2s+1}}(\rho_A),$$

which can also be verified by employing the dual relation of [49, Lemma 6]. So the reliability function in this case has a simpler formula involving only the Rényi entropy of one single system.

IV. PROOF OF THE CHARACTERIZATION OF RELIABILITY FUNCTIONS

In this section, we prove Theorem 8 on the characterization of the reliability functions. As an intermediate step, we also derive the one-shot achievability bound of Theorem 7. The proof is organized as follows. In Section IV A, we analyse the convex-split lemma [20], obtaining a new bound that employs the sandwiched Rényi divergence. This will be a crucial tool for proving Theorem 8, regarding the achievability bound of Eq. (11). Then in

Section IV B, we derive the exact exponent for the asymptotic decreasing of the smoothing quantity for the max-information. This will serve as another key technical tool for proving Theorem 8, for the converse bound of Eq. (12). At last, in Section IV C, we accomplish the proof of Theorem 8 as well as Theorem 7, by employing the established tools mentioned above.

A. A convex-split lemma

The convex-split lemma was introduced in [20] and has broad applications in topics such as one-shot quantum Shannon theory [20, 21, 65–67], entanglement and general resource theories [68, 69], and quantum thermodynamics [70, 71]. Roughly speaking, it quantifies how well the information of a distinct object located among many other identical ones, can be erased by randomly mixing all of them. Originally in [20] and in all of the previous applications, the effect of this erasure is bounded using the max-relative entropy, which, after being smoothed, is sufficiently tight for those purposes. However, for our purpose of deriving the reliability function, this bound does not work any more. Instead, we prove a version of the convex-split lemma with a new bound, employing directly the sandwiched Rényi divergence.

Lemma 13 *Let $\rho_{RA} \in \mathcal{S}(RA)$ and $\sigma_A \in \mathcal{S}(\mathcal{H}_A)$ be quantum states such that* $\operatorname{supp}(\rho_A) \subseteq \operatorname{supp}(\sigma_A)$. *Consider the following state*

$$\tau_{RA_1A_2\cdots A_m} := \frac{1}{m}\sum_{i=1}^{m} \rho_{RA_i} \otimes \left[\sigma^{\otimes(m-1)}\right]_{A^m/A_i},$$

where A^m/A_i denotes the composite system consisting of $A_1, A_2, \cdots, A_{i-1}, A_{i+1}, \cdots, A_m$ and $\left[\sigma^{\otimes(m-1)}\right]_{A^m/A_i}$ is the product state $\sigma^{\otimes(m-1)}$ on these $m-1$ systems. Let $v = v(\rho_R \otimes \sigma_A)$ denote the number of distinct eigenvalues of $\rho_R \otimes \sigma_A$. Then for any $0 < s \leq 1$,

$$D\left(\tau_{RA_1A_2\cdots A_m} \middle\| \rho_R \otimes (\sigma^{\otimes m})_{A^m}\right) \leq \frac{v^s}{(\ln 2)s} \exp\Big\{ -(\ln 2)\ s\big(\log m - D_{1+s}(\rho_{RA}\|\rho_R\otimes\sigma_A)\big)\Big\}.$$

Proof. We use the shorthand $\xi_i \equiv \rho_{RA_i} \otimes \left[\sigma^{\otimes(m-1)}\right]_{A^m/A_i}$ for simplicity and start with

$$\begin{aligned}
& D\Big(\tau_{RA_1A_2\cdots A_m} \big\| \rho_R \otimes (\sigma^{\otimes m})_{A^m}\Big) \\
=& \operatorname{Tr}\left[\Big(\frac{1}{m}\sum_i \xi_i\Big)\Big(\log\big(\frac{1}{m}\sum_i \xi_i\big) - \log\big(\rho_R \otimes (\sigma^{\otimes m})_{A^m}\big)\Big)\right] \\
=& \operatorname{Tr}\left[\xi_1\Big(\log\big(\frac{1}{m}\sum_i \xi_i\big) - \log\big(\rho_R \otimes (\sigma^{\otimes m})_{A^m}\big)\Big)\right] \\
=& \operatorname{Tr}\left[\xi_1\Big(\log \xi_1 - \log\big(\rho_R \otimes (\sigma^{\otimes m})_{A^m}\big)\Big)\right] - \operatorname{Tr}\left[\xi_1\Big(\log \xi_1 - \log\big(\frac{1}{m}\sum_i \xi_i\big)\Big)\right] \\
=& D\Big(\rho_{RA_1} \otimes (\sigma^{\otimes(m-1)})_{A_2\cdots A_m} \big\| \rho_R \otimes (\sigma^{\otimes m})_{A^m}\Big) - D\Big(\rho_{RA_1} \otimes (\sigma^{\otimes(m-1)})_{A_2\cdots A_m} \big\| \frac{1}{m}\sum_i \xi_i\Big) \\
\leq& D\big(\rho_{RA_1}\|\rho_R\otimes\sigma_{A_1}\big) - D\Big(\rho_{RA_1}\big\|\frac{1}{m}\rho_{RA_1} + \frac{m-1}{m}\rho_R\otimes\sigma_{A_1}\Big) \\
=& \operatorname{Tr}\left[\rho_{RA}\Big(\log\big(\frac{1}{m}\rho_{RA} + \frac{m-1}{m}\rho_R\otimes\sigma_A\big) - \log(\rho_R\otimes\sigma_A)\Big)\right],
\end{aligned}$$

where the third line is due to the symmetry of the states $\frac{1}{m}\sum_i \xi_i$ and $\rho_R \otimes (\sigma^{\otimes m})_{A^m}$ over systems $A_1, A_2, \cdots, A_m$, and for the inequality we have used the data processing inequality for relative entropy under partial trace. Now employ the pinching map $\mathcal{E}_{\rho_R \otimes \sigma_A}$ and write $\bar{\rho}_{RA} := \mathcal{E}_{\rho_R \otimes \sigma_A}(\rho_{RA})$. The pinching inequality together with the operator monotonicity of the logarithm gives

$$\log\Big(\frac{1}{m}\rho_{RA} + \frac{m-1}{m}\rho_R \otimes \sigma_A\Big) \leq \log\Big(\frac{v}{m}\bar{\rho}_{RA} + \rho_R \otimes \sigma_A\Big).$$

Making use of this, we proceed as follows.

$$\begin{aligned}
& D\Big(\tau_{RA_1A_2\cdots A_m}\Big\|\rho_R \otimes (\sigma^{\otimes m})_{A^m}\Big) \\
\leq & \operatorname{Tr}\Big[\rho_{RA}\Big(\log\big(\frac{v}{m}\bar{\rho}_{RA} + \rho_R \otimes \sigma_A\big) - \log(\rho_R \otimes \sigma_A)\Big)\Big] \\
= & \operatorname{Tr}\Big[\bar{\rho}_{RA}\log\Big(\frac{v}{m}\bar{\rho}_{RA}(\rho_R \otimes \sigma_A)^{-1} + \mathbb{1}_{RA}\Big)\Big] \\
\leq & \frac{1}{(\ln 2)s}\operatorname{Tr}\Big[\bar{\rho}_{RA}\Big(\frac{v^s}{m^s}(\bar{\rho}_{RA})^s(\rho_R \otimes \sigma_A)^{-s}\Big)\Big] \\
= & \frac{v^s}{(\ln 2)s}\exp\Big\{-(\ln 2)\ s\big(\log m - D_{1+s}(\bar{\rho}_{RA}\|\rho_R \otimes \sigma_A)\big)\Big\} \\
\leq & \frac{v^s}{(\ln 2)s}\exp\Big\{-(\ln 2)\ s\big(\log m - D_{1+s}(\rho_{RA}\|\rho_R \otimes \sigma_A)\big)\Big\},
\end{aligned}$$

where for the third line note that the density matrices $\bar{\rho}_{RA}$ and $\rho_R \otimes \sigma_A$ commute, for the fourth line we have used the inequality $\ln(1+x) \leq \frac{1}{s}x^s$ for $x \geq 0$ and $0 < s \leq 1$, and the last line is by the data processing inequality for the sandwiched Rényi divergence [41, 45, 57, 72]. □

B. Smoothing of max-information and conditional min-entropy

Recall that the smooth max-relative entropy is defined for $\rho \in \mathcal{S}(\mathcal{H})$, $\sigma \in \mathcal{P}(\mathcal{H})$, and $0 \leq \delta < 1$ as [59]

$$D^{\delta}_{\max}(\rho\|\sigma) := \min\left\{\lambda : \exists\, \tilde{\rho} \in \mathcal{S}_{\leq}(\mathcal{H}) \text{ s.t. } P(\tilde{\rho}, \rho) \leq \delta,\, \tilde{\rho} \leq 2^{\lambda}\sigma\right\}.$$

In [46] we have introduced the smoothing quantity for the max-relative entropy

$$\delta(\rho\|\sigma, \lambda) := \min\left\{P(\tilde{\rho}, \rho) : \tilde{\rho} \in \mathcal{S}_{\leq}(\mathcal{H}),\, \tilde{\rho} \leq 2^{\lambda}\sigma\right\}$$

and obtained

$$\lim_{n\to\infty} \frac{-1}{n}\log\delta\big(\rho^{\otimes n}\big\|\sigma^{\otimes n}, nr\big) = \frac{1}{2}\sup_{s\geq 0}\left\{s\big(r - D_{1+s}(\rho\|\sigma)\big)\right\}. \tag{21}$$

In this section, we are interested in the smooth max-information and the smooth conditional min-entropy defined for $\rho_{AB} \in \mathcal{S}(AB)$ and $0 \leq \delta < 1$, respectively as [7, 28, 73]

$$\begin{aligned}
I^{\delta}_{\max}(A:B)_{\rho} &:= \min_{\sigma_B \in \mathcal{S}(B)} D^{\delta}_{\max}(\rho_{AB}\|\rho_A \otimes \sigma_B), \\
H^{\delta}_{\min}(A|B)_{\rho} &:= -\min_{\sigma_B \in \mathcal{S}(B)} D^{\delta}_{\max}(\rho_{AB}\|\mathbb{1}_A \otimes \sigma_B).
\end{aligned}$$

Our purpose is to derive the asymptotic exponents for the smoothing of these two one-shot entropies. Note that in the classical case where ρ_{AB} is a probability distribution, this kind of exponential analysis was done for the conditional min-entropy based on the trace distance [74]. In doing so, we define the corresponding smoothing quantities.

Definition 14 *Let $\rho_{AB} \in \mathcal{S}(AB)$ and $\lambda \in \mathbb{R}$. The smoothing quantity for the max-information and for the conditional min-entropy is defined, respectively, as*

$$\delta_{A:B}(\rho_{AB}, \lambda) := \min_{\sigma_B \in \mathcal{S}(B)} \delta(\rho_{AB} \| \rho_A \otimes \sigma_B, \lambda), \tag{22}$$

$$\delta_{A|B}(\rho_{AB}, \lambda) := \min_{\sigma_B \in \mathcal{S}(B)} \delta(\rho_{AB} \| \mathbb{1}_A \otimes \sigma_B, -\lambda). \tag{23}$$

Note that alternative expressions for the smoothing quantities are

$$\delta_{A:B}(\rho_{AB}, \lambda) = \min \left\{ \delta \in \mathbb{R} : I_{\max}^{\delta}(A:B)_\rho \leq \lambda \right\}, \tag{24}$$

$$\delta_{A|B}(\rho_{AB}, \lambda) = \min \left\{ \delta \in \mathbb{R} : H_{\min}^{\delta}(A|B)_\rho \geq \lambda \right\}, \tag{25}$$

from which the relation to the smooth max-information and smooth conditional min-entropy is easily seen. The main result is stated in the following.

Theorem 15 *For $\rho_{AB} \in \mathcal{S}(AB)$ and $r \in \mathbb{R}$, we have*

$$\lim_{n\to\infty} \frac{-1}{n} \log \delta_{A^n:B^n}\left(\rho_{AB}^{\otimes n}, nr\right) = \frac{1}{2} \sup_{s\geq 0} \left\{ s\big(r - I_{1+s}(A:B)_\rho\big) \right\}, \tag{26}$$

$$\lim_{n\to\infty} \frac{-1}{n} \log \delta_{A^n|B^n}\left(\rho_{AB}^{\otimes n}, nr\right) = \frac{1}{2} \sup_{s\geq 0} \left\{ s\big(H_{1+s}(A|B)_\rho - r\big) \right\}. \tag{27}$$

Proof. Since the r.h.s. of Eqs. (22) and (23) are similar, it suffices to prove, for any $M \in \mathcal{P}(A)$ and $r \in \mathbb{R}$,

$$\lim_{n\to\infty} \frac{-1}{n} \log \min_{\sigma_{B^n}^{(n)} \in \mathcal{S}(B^n)} \delta\left(\rho_{AB}^{\otimes n} \middle\| M_A^{\otimes n} \otimes \sigma_{B^n}^{(n)}, nr\right) = \frac{1}{2} \sup_{s\geq 0} \left\{ s\Big(r - \min_{\sigma_B \in \mathcal{S}(B)} D_{1+s}(\rho_{AB} \| M_A \otimes \sigma_B)\Big) \right\}. \tag{28}$$

Substituting ρ_A for M_A results in Eq. (26), and the substitution of $M_A \leftarrow \mathbb{1}$ and $r \leftarrow -r$ recovers Eq. (27).

Observing that

$$\min_{\sigma_{B^n}^{(n)} \in \mathcal{S}(B^n)} \delta\left(\rho_{AB}^{\otimes n} \middle\| M_A^{\otimes n} \otimes \sigma_{B^n}^{(n)}, nr\right) \leq \min_{\sigma_B^{\otimes n} \in \mathcal{S}(B^n)} \delta\left(\rho_{AB}^{\otimes n} \middle\| M_A^{\otimes n} \otimes \sigma_B^{\otimes n}, nr\right),$$

we obtain the "$\geq$" part of Eq. (28) by invoking Eq. (21). That is,

$$\liminf_{n\to\infty} \frac{-1}{n} \log \min_{\sigma_{B^n}^{(n)} \in \mathcal{S}(B^n)} \delta\left(\rho_{AB}^{\otimes n} \middle\| M_A^{\otimes n} \otimes \sigma_{B^n}^{(n)}, nr\right) \geq \frac{1}{2} \sup_{s\geq 0} \left\{ s\Big(r - \min_{\sigma_B \in \mathcal{S}(B)} D_{1+s}(\rho_{AB} \| M_A \otimes \sigma_B)\Big) \right\}. \tag{29}$$

Now we turn to the proof of the opposite direction. At first, we have

$$\begin{aligned}
&\min_{\sigma_{B^n}^{(n)}\in\mathcal{S}(B^n)} \delta\left(\rho_{AB}^{\otimes n}\middle\| M_A^{\otimes n}\otimes\sigma_{B^n}^{(n)}, nr\right)\\
=&\min\left\{P\big(\rho_{AB}^{\otimes n},\gamma_{A^nB^n}^{(n)}\big) : \gamma_{A^nB^n}^{(n)}\in\mathcal{S}_{\leq}(A^nB^n), \big(\exists\ \sigma_{B^n}^{(n)}\in\mathcal{S}(B^n)\big)\ \gamma_{A^nB^n}^{(n)}\leq 2^{nr}M_A^{\otimes n}\otimes\sigma_{B^n}^{(n)}\right\}\\
=&\min\left\{P\big(\rho_{AB}^{\otimes n},\gamma_{A^nB^n}^{(n)}\big) : \gamma_{A^nB^n}^{(n)}\in\mathcal{S}_{\leq}^{\mathrm{sym}}(A^nB^n), \big(\exists\ \sigma_{B^n}^{(n)}\in\mathcal{S}^{\mathrm{sym}}(B^n)\big)\ \gamma_{A^nB^n}^{(n)}\leq 2^{nr}M_A^{\otimes n}\otimes\sigma_{B^n}^{(n)}\right\}\\
\geq&\min\left\{P\big(\rho_{AB}^{\otimes n},\gamma_{A^nB^n}^{(n)}\big) : \gamma_{A^nB^n}^{(n)}\in\mathcal{S}_{\leq}(A^nB^n), \gamma_{A^nB^n}^{(n)}\leq 2^{nr}g_{n,|B|}M_A^{\otimes n}\otimes\omega_{B^n}^{(n)}\right\},
\end{aligned}\tag{30}$$

where the second line is by definition, in the third line we restrict the minimization to over the symmetric states by making a random permutation and this makes no difference because random permutation operation as a CPTP map keeps operator inequality and does not increase the purified distance, in the fourth line we have employed the universal symmetric state $\omega_{B^n}^{(n)}$ and made use of Lemma 1. Let $\gamma_{A^nB^n}^{(n)*}$ be the optimal state that makes the last line of Eq. (30) achieves the minimum. Let $\mathcal{E}^n\equiv\mathcal{E}_{M_A^{\otimes n}\otimes\omega_{B^n}^{(n)}}$ be the pinching map associated with $M_A^{\otimes n}\otimes\omega_{B^n}^{(n)}$. Then we have

$$\begin{aligned}
\min_{\sigma_{B^n}^{(n)}\in\mathcal{S}(B^n)} \delta\left(\rho_{AB}^{\otimes n}\middle\| M_A^{\otimes n}\otimes\sigma_{B^n}^{(n)}, nr\right) \geq& P\left(\rho_{AB}^{\otimes n},\gamma_{A^nB^n}^{(n)*}\right)\\
\geq& P\left(\mathcal{E}^n\big(\rho_{AB}^{\otimes n}\big),\mathcal{E}^n\big(\gamma_{A^nB^n}^{(n)*}\big)\right),
\end{aligned}\tag{31}$$

and in addition, $\gamma_{A^nB^n}^{(n)*}\leq 2^{nr}g_{n,|B|}M_A^{\otimes n}\otimes\omega_{B^n}^{(n)}$, which after the CPTP map $\mathcal{E}^n$ being applied to both sides yields

$$\mathcal{E}^n\big(\gamma_{A^nB^n}^{(n)*}\big)\leq 2^{nr}g_{n,|B|}M_A^{\otimes n}\otimes\omega_{B^n}^{(n)}.\tag{32}$$

To proceed, we construct a projective measurement $\{\Pi_n, \mathbb{1}-\Pi_n\}$ with

$$\Pi_n:=\left\{\mathcal{E}^n\big(\rho_{AB}^{\otimes n}\big)\geq 9\cdot 2^{nr}g_{n,|B|}M_A^{\otimes n}\otimes\omega_{B^n}^{(n)}\right\}$$

and set

$$p_n=\mathrm{Tr}\left(\mathcal{E}^n(\rho_{AB}^{\otimes n})\Pi_n\right),\tag{33}$$

$$q_n=\mathrm{Tr}\left(\mathcal{E}^n(\gamma_{A^nB^n}^{(n)*})\Pi_n\right).\tag{34}$$

By Eq. (32) and the construction of Π_n, it can be verified that

$$q_n\leq\frac{1}{9}p_n.$$

So letting $\mathcal{M}^n: X\mapsto \mathrm{Tr}(X\Pi_n)|0\rangle\langle 0|+\mathrm{Tr}(X(\mathbb{1}-\Pi_n))|1\rangle\langle 1|$ be the measurement map associated with $\{\Pi_n, \mathbb{1} - \Pi_n\}$, we are able to obtain

$$\begin{aligned}
&\min_{\sigma_{B^n}^{(n)} \in \mathcal{S}(B^n)} \delta\left(\rho_{AB}^{\otimes n} \middle\| M_A^{\otimes n} \otimes \sigma_{B^n}^{(n)}, nr\right) \\
\geq & P\left(\mathcal{M}^n \circ \mathcal{E}^n(\rho_{AB}^{\otimes n}), \mathcal{M}^n \circ \mathcal{E}^n(\gamma_{A^nB^n}^{(n)*})\right) \\
= & \sqrt{1 - \left(\sqrt{p_n}\sqrt{q_n} + \sqrt{1-p_n}\sqrt{\operatorname{Tr}(\gamma_{A^nB^n}^{(n)*}) - q_n}\right)^2} \\
\geq & \sqrt{p_n\left(\frac{1}{3} - \frac{p_n}{9}\right)} \\
\geq & \frac{\sqrt{2p_n}}{3},
\end{aligned} \tag{35}$$

where the second line follows from Eq. (31) and the data processing inequality for purified distance, for the fourth line we have used $q_n \leq \frac{1}{9}p_n$ and $\operatorname{Tr}(\gamma_{A^nB^n}^{(n)*}) - q_n \leq 1$, and for the last line note that $p_n \leq 1$.

We prove later in Proposition 16 the asymptotics for p_n (cf. Eq. (33) for its expression), that is,

$$\lim_{n\to\infty} \frac{-1}{n} \log p_n = \sup_{s\geq 0} \left\{ s\big(r - \min_{\sigma_B \in \mathcal{S}(B)} D_{1+s}(\rho_{AB} \| M_A \otimes \sigma_B)\big) \right\}. \tag{36}$$

Then it follows from Eq. (36) and Eq. (35) that

$$\limsup_{n\to\infty} \frac{-1}{n} \log \min_{\sigma_{B^n}^{(n)} \in \mathcal{S}(B^n)} \delta\left(\rho_{AB}^{\otimes n} \middle\| M_A^{\otimes n} \otimes \sigma_{B^n}^{(n)}, nr\right) \leq \frac{1}{2} \sup_{s\geq 0} \left\{ s\big(r - \min_{\sigma_B \in \mathcal{S}(B)} D_{1+s}(\rho_{AB} \| M_A \otimes \sigma_B)\big) \right\}. \tag{37}$$

Eventually, Eq. (37) and Eq. (29) together lead to Eq. (28), and we are done. □

Proposition 16 *Let $\rho_{AB} \in \mathcal{S}(AB)$ and $M_A \in \mathcal{P}(A)$ be such that $\operatorname{supp}(\rho_A) \subseteq \operatorname{supp}(M_A)$. Let $\omega_{B^n}^{(n)}$ be the universal symmetric state. Let $\mathcal{E}^n \equiv \mathcal{E}_{M_A^{\otimes n} \otimes \omega_{B^n}^{(n)}}$ be the pinching map, and let $f(n) \geq 1$ be any sub-exponential function of n. For given $r \in \mathbb{R}$, consider the sequence*

$$p_n := \operatorname{Tr}\left[\mathcal{E}^n(\rho_{AB}^{\otimes n})\left\{\mathcal{E}^n(\rho_{AB}^{\otimes n}) \geq f(n) 2^{nr} M_A^{\otimes n} \otimes \omega_{B^n}^{(n)}\right\}\right]$$

for $n \in \mathbb{N}$. We have

$$\lim_{n\to\infty} \frac{-1}{n} \log p_n = \sup_{s\geq 0} \left\{ s\big(r - \min_{\sigma_B \in \mathcal{S}(B)} D_{1+s}(\rho_{AB} \| M_A \otimes \sigma_B)\big) \right\}. \tag{38}$$

Proof. For simplicity, we use the shorthands

$$P_n := \mathcal{E}^n(\rho_{AB}^{\otimes n}), \quad Q_n := M_A^{\otimes n} \otimes \omega_{B^n}^{(n)}.$$

We will make frequent use of the fact that P_n commutes with Q_n as well as that, by Lemma 17,

$$\frac{1}{n} D_{1+s}(P_n \| Q_n) \longrightarrow \min_{\sigma_B \in \mathcal{S}(B)} D_{1+s}(\rho_{AB} \| M_A \otimes \sigma_B), \quad \text{as } n \to \infty. \tag{39}$$

In what follows we will first show that the r.h.s. of Eq. (38) is an achievable rate for the exponential decreasing of p_n, and then we prove that this is indeed the optimal rate.

Now we prove the achievability part. For any $s>0$,

$$\begin{aligned}
\frac{-1}{n}\log p_n &= \frac{-1}{n}\log \operatorname{Tr}\left[P_n\left\{P_n \geq f(n)2^{nr}Q_n\right\}\right] \\
&\geq \frac{-1}{n}\log \operatorname{Tr}\left[P_n\Big(\frac{P_n}{f(n)2^{nr}Q_n}\Big)^s\right] \\
&= sr - \frac{1}{n}\log \operatorname{Tr}\left[P_n^{1+s}Q_n^{-s}\right] + \frac{s\log f(n)}{n} \\
&= sr - \frac{s}{n}D_{1+s}\big(P_n\|Q_n\big) + \frac{s\log f(n)}{n} \\
&\overset{n\to\infty}{\longrightarrow} s\big(r - \min_{\sigma_B\in\mathcal{S}(B)} D_{1+s}(\rho_{AB}\|M_A\otimes\sigma_B)\big).
\end{aligned}$$

Noticing that $s>0$ is arbitrary, we conclude from the above estimation that

$$\liminf_{n\to\infty}\frac{-1}{n}\log p_n \geq \sup_{s\geq 0}\left\{s\big(r-\min_{\sigma_B\in\mathcal{S}(B)} D_{1+s}(\rho_{AB}\|M_A\otimes\sigma_B)\big)\right\}. \tag{40}$$

For the other direction, we follow the method of [49] and employ the Gärtner-Ellis theorem of large deviation theory. Let

$$Z_n := \frac{1}{n}\big(\log P_n - \log Q_n - nr - \log f(n)\big).$$

We have

$$\begin{aligned}
p_n &= \operatorname{Tr}\left[P_n\left\{P_n \geq f(n)2^{nr}Q_n\right\}\right] \\
&= \operatorname{Tr}\left[P_n\left\{Z_n \geq 0\right\}\right].
\end{aligned} \tag{41}$$

We see that the observable Z_n commutes with the state P_n. So the above Eq. (41) has a classical-probability-theoretic explanation in which Z_n is regarded as a random variable with P_n being its distribution. Specifically, let the set of orthonormal vectors $\{|a_x^n\rangle\}_x$ be the common eigenvectors of Z_n and P_n. Then Z_n takes value $\langle a_x^n|Z_n|a_x^n\rangle$ with probability $\langle a_x^n|P_n|a_x^n\rangle$. Now, Eq. (41) translates to

$$p_n = \Pr\big\{Z_n \geq 0\big\},$$

with respect to the probability distribution $\{\langle a_x^n|P_n|a_x^n\rangle\}_x$. To apply the Gärtner-Ellis theorem (in the form of Lemma 18), we calculate the asymptotic cumulant generating function of the sequence $\mathcal{Z}=\{Z_n\}_{n\in\mathbb{N}}$:

$$\begin{aligned}
\Lambda_{\mathcal{Z}}(s) &= \lim_{n\to\infty}\frac{1}{n}\log\mathbb{E}\big[2^{nsZ_n}\big] \\
&= \lim_{n\to\infty}\frac{1}{n}\log\mathbb{E}\Big[\big(\frac{P_n}{Q_n}\big)^s(f(n))^{-s}2^{-nsr}\Big] \\
&= \lim_{n\to\infty}\frac{1}{n}\log \operatorname{Tr}\Big[P_n\big(\frac{P_n}{Q_n}\big)^s(f(n))^{-s}2^{-nsr}\Big] \\
&= \lim_{n\to\infty}\Big(\frac{s}{n}D_{1+s}(P_n\|Q_n) - sr - \frac{s\log f(n)}{n}\Big) \\
&= s\big(\min_{\sigma_B\in\mathcal{S}(B)} D_{1+s}(\rho_{AB}\|M_A\otimes\sigma_B) - r\big).
\end{aligned} \tag{42}$$

Set $F(s) := \min_{\sigma_B\in\mathcal{S}(B)} D_{1+s}(\rho_{AB}\|M_A\otimes\sigma_B)$ for later convenience. To proceed, we restrict our attention to the case where $F(0) < r < F(+\infty)$ at the moment. It has been proven in [49] that the function $s\mapsto F(s)$ is monotonically increasing, continuously differentiable, and $s\mapsto sF(s)$ is convex. As a result, we have that $s\mapsto\Lambda_{\mathcal{Z}}(s)$ is differentialbe in $(0,+\infty)$, and

$$\begin{aligned}\lim_{s\to 0}\Lambda'_{\mathcal{Z}}(s) &= D(\rho_{AB}\|M_A\otimes\rho_B) - r = F(0) - r < 0,\\ (\exists s_0>0)\ \lim_{s\to s_0}\Lambda'_{\mathcal{Z}}(s) &\geq \min_{\sigma_B\in\mathcal{S}(B)} D_{1+s_0}(\rho_{AB}\|M_A\otimes\sigma_B) - r = F(s_0) - r > 0.\end{aligned}$$

Hence, Lemma 18 applies, yielding for $r\in(F(0),F(\infty))$,

$$\begin{aligned}\limsup_{n\to\infty}\frac{-1}{n}\log p_n &= \limsup_{n\to\infty}\frac{-1}{n}\log\Pr\big\{Z_n\geq 0\big\}\\ &\leq \sup_{0<s<s_0}\left\{s\big(r - \min_{\sigma_B\in\mathcal{S}(B)} D_{1+s}(\rho_{AB}\|M_A\otimes\sigma_B)\big)\right\} \qquad (43)\\ &\leq \sup_{s>0}\left\{s\big(r - \min_{\sigma_B\in\mathcal{S}(B)} D_{1+s}(\rho_{AB}\|M_A\otimes\sigma_B)\big)\right\}.\end{aligned}$$

Now combining Eq. (40) and Eq. (43) together lets us arrive at

$$\lim_{n\to\infty}\frac{-1}{n}\log p_n = \sup_{s\geq 0}\left\{s\big(r - \min_{\sigma_B\in\mathcal{S}(B)} D_{1+s}(\rho_{AB}\|M_A\otimes\sigma_B)\big)\right\} \qquad (44)$$

for $r\in(F(0),F(\infty))$. To complete the proof, we show that the equality of Eq. (44) can be extended to the whole range $r\in\mathbb{R}$. At first, we observe by definition that, the l.h.s. of Eq. (44) is nonnegative and monotonically increasing with r. Moreover, the equality of Eq. (44) established for $r\in(F(0),F(\infty))$ shows that the l.h.s. goes to 0 when $r\searrow F(0) = D(\rho_{AB}\|M_A\otimes\rho_B)$. So we conclude that the l.h.s. of Eq. (44) equals 0 for $r\leq F(0)$, coinciding with the right hand side. Next, we consider the case $r > F(\infty) = \min_{\sigma_B\in\mathcal{S}(B)} D_{\max}(\rho_{AB}\|M_A\otimes\sigma_B)$. Letting $s\to\infty$ in Lemma 17 we get

$$D_{\max}\Big(\mathcal{E}^n(\rho_{AB}^{\otimes n})\big\|M_A^{\otimes n}\otimes\omega_{B^n}^{(n)}\Big) = n\left(F(\infty) + O\Big(\frac{\log n}{n}\Big)\right).$$

This implies that, for $r > F(\infty)$ and n big enough,

$$\Big\{\mathcal{E}^n\big(\rho_{AB}^{\otimes n}\big)\geq f(n)2^{nr}M_A^{\otimes n}\otimes\omega_{B^n}^{(n)}\Big\} = 0$$

and hence $p_n = 0$. As a result, we conclude that the l.h.s. of Eq. (44) is $+\infty$ for $r > F(\infty)$, coinciding with the right hand side, too. □

The following lemma is due to Hayashi and Tomamichel [49].

Lemma 17 *Let $\rho_{AB}\in\mathcal{S}(AB)$ and $M_A\in\mathcal{P}(A)$ such that $\mathrm{supp}(\rho_A)\subseteq\mathrm{supp}(M_A)$. Let $\omega_{B^n}^{(n)}$ be the universal symmetric state. Let $\mathcal{E}^n\equiv\mathcal{E}_{M_A^{\otimes n}\otimes\omega_{B^n}^{(n)}}$ be the pinching map. For any $s\geq\frac{1}{2}$ we have*

$$\frac{1}{n}D_s\Big(\mathcal{E}^n(\rho_{AB}^{\otimes n})\big\|M_A^{\otimes n}\otimes\omega_{B^n}^{(n)}\Big) = \min_{\sigma_B\in\mathcal{S}(B)} D_s(\rho_{AB}\|M_A\otimes\sigma_B) + O\Big(\frac{\log n}{n}\Big).$$

The implicit constants of the last term are independent of s.

The following lemma is due to Chen [75] and reformulated in [49].

Lemma 18 *Let $\{Z_n\}_{n\in\mathbb{N}} =: \mathcal{Z}$ be a sequence of random variables, with the asymptotic cumulant generating function*

$$\Lambda_{\mathcal{Z}}(t) := \lim_{n\to\infty} \frac{1}{n} \log \mathbb{E}\big[2^{ntZ_n}\big].$$

If the above limit exists and the function $t \mapsto \Lambda_{\mathcal{Z}}(t)$ is differentiable in some interval (a,b), then for any $z \in \big(\lim_{t\searrow a} \Lambda'_{\mathcal{Z}}(t), \lim_{t\nearrow b} \Lambda'_{\mathcal{Z}}(t)\big)$, it holds that

$$\limsup_{n\to\infty} \frac{-1}{n} \log \Pr\{Z_n \geq z\} \leq \sup_{t\in(a,b)} \big\{zt - \Lambda_{\mathcal{Z}}(t)\big\}.$$

C. Proof of Theorem 8

Based on the results obtained in the above two subsections, we are now ready to complete the proof of Theorem 8. At first, we prove the one-shot achievability bound of Theorem 7.

Theorem 7 (restatement) *Let $\rho_{RA} \in \mathcal{S}(RA)$. For any $m \in \mathbb{N}$, $0 < s \leq 1$ and $\sigma_A \in \mathcal{S}(A)$, the optimal performance of decoupling A from R is bounded as*

$$P^{\text{dec}}_{R:A}(\rho_{RA}, \log m) \leq \sqrt{\frac{v^s}{s}} \exp\Big\{ - (\ln 2)\ s\big(\log m - \frac{1}{2} D_{1+s}(\rho_{RA} \| \rho_R \otimes \sigma_A)\big)\Big\},$$

where v is the number of distinct eigenvalues of $\rho_R \otimes \sigma_A$.

Proof. We will employ the convex-split lemma. The convex-split lemma provides directly a catalytic decoupling strategy via random unitary operation, which in turn can be converted into a way for catalytic decoupling via removing a subsystem (cf. Section V A). Identify A with A_1, and let $A' = A_2 A_3 \cdots A_m$, where all the A_i systems have equal dimension. Let A' in the state $\bar{\sigma}_{A'} = \sigma_{A_2} \otimes \sigma_{A_3} \otimes \cdots \otimes \sigma_{A_m}$ be the catalytic system. We construct a random unitary channel

$$\Lambda_{AA'} : X \mapsto \frac{1}{m} \sum_{i=1}^{m} U_i X U_i^*,$$

where $U_i = W_{(1,i)}$ is the swapping between A_1 and A_i (we set $W_{(1,1)} = \mathbb{1}$). Then

$$\Lambda_{AA'}\big(\rho_{RA} \otimes \bar{\sigma}_{A'}\big) = \frac{1}{m} \sum_{i=1}^{m} \rho_{RA_i} \otimes \big[\sigma^{\otimes m-1}\big]_{A^m/A_i},$$

which is in the form of the state in Lemma 13. So we have

$$\begin{aligned}
P^{\text{dec-u}}_{R:A}(\rho_{RA}, \log m) \leq & P\Big(\Lambda_{AA'}\big(\rho_{RA} \otimes \bar{\sigma}_{A'}\big), \rho_R \otimes \big[\sigma^{\otimes m}\big]_{A^m}\Big) \\
\leq & \Big[(\ln 2)\ D\Big(\Lambda_{AA'}\big(\rho_{RA} \otimes \bar{\sigma}_{A'}\big) \big\| \rho_R \otimes \big[\sigma^{\otimes m}\big]_{A^m}\Big)\Big]^{\frac{1}{2}} \\
\leq & \sqrt{\frac{v^s}{s}} \exp\Big\{ - (\ln 2)\ s\big(\frac{1}{2}\log m - \frac{1}{2} D_{1+s}(\rho_{RA} \| \rho_R \otimes \sigma_A)\big)\Big\},
\end{aligned}$$

where for the second line we have used the relation $P(\rho,\sigma) \leq \sqrt{(\ln 2)D(\rho\|\sigma)}$, and for the third line we have used Lemma 13. At last, invoking Proposition 6 lets us confirm the claim. □

The following one-shot converse bound is proved using standard techniques (see, e.g. [19] and [21]).

Proposition 19 *Let $\rho_{RA} \in \mathcal{S}(RA)$. For $k \geq 0$, the optimal performance of decoupling A from R is bounded by the smoothing quantity for the max-information (cf. Eq. (22) in Definition 14):*

$$P^{\mathrm{dec}}_{R:A}(\rho_{RA}, k) \geq \delta_{R:A}(\rho_{RA}, 2k).$$

Proof. Consider an arbitrary decoupling scheme that discards not more than than k qubits. Let the catalytic state be $\sigma_{A'}$ and the unitary decoupling operation be $U : \mathcal{H}_{AA'} \to \mathcal{H}_{A_1A_2}$. Then $\log|A_2| \leq k$ and the performance for this scheme is

$$\epsilon := \min_{\omega_{A_1}\in\mathcal{S}(A_1)} P\big(\operatorname{Tr}_{A_2} U(\rho_{RA}\otimes\sigma_{A'})U^*, \rho_R\otimes\omega_{A_1}\big).$$

Let $\omega^*_{A_1}$ be the optimal state in the above minimization. By Uhlmann's theorem [52], there is a state $\tilde{\rho}_{RA_1A_2}$ such that

$$P\big(U(\rho_{RA}\otimes\sigma_{A'})U^*, \tilde{\rho}_{RA_1A_2}\big) = \epsilon, \tag{45}$$
$$\operatorname{Tr}_{A_2}\tilde{\rho}_{RA_1A_2} = \rho_R\otimes\omega^*_{A_1}. \tag{46}$$

From Eq. (45) we get

$$P\big(\rho_{RA}, \operatorname{Tr}_{A'} U^*\tilde{\rho}_{RA_1A_2}U\big) \leq \epsilon. \tag{47}$$

Eq. (46) implies that

$$\tilde{\rho}_{RA_1A_2} \leq |A_2|^2\rho_R\otimes\omega^*_{A_1}\otimes\frac{\mathbb{1}_{A_2}}{|A_2|} \leq 2^{2k}\rho_R\otimes\omega^*_{A_1}\otimes\frac{\mathbb{1}_{A_2}}{|A_2|},$$

which further yields

$$\operatorname{Tr}_{A'} U^*\tilde{\rho}_{RA_1A_2}U \leq 2^{2k}\rho_R\otimes\operatorname{Tr}_{A'} U^*\big(\omega^*_{A_1}\otimes\frac{\mathbb{1}_{A_2}}{|A_2|}\big)U. \tag{48}$$

Inspecting the definition in Eq. (22), we obtain from Eq. (47) and Eq. (48) that

$$\delta_{R:A}(\rho_{RA}, 2k) \leq \epsilon.$$

Since by assumption ϵ is the performance of an arbitrary decoupling scheme with cost not larger than k, the optimal performance satisfies the above relation as well. That is,

$$\delta_{R:A}(\rho_{RA}, 2k) \leq P^{\mathrm{dec}}_{R:A}(\rho_{RA}, k).$$

□

Proof of Theorem 8. By applying Theorem 7 and making the substitution $\rho_{RA} \leftarrow \rho^{\otimes n}_{RA}$, $\log m \leftarrow nr$, $\sigma_A \leftarrow \sigma^{\otimes n}_A$, we get

$$\frac{-1}{n}\log P^{\mathrm{dec}}_{R^n:A^n}\big(\rho^{\otimes n}_{RA}, nr\big) \geq \frac{-1}{2n}\log\frac{v^s_n}{s} + s\big(r - \frac{1}{2}D_{1+s}(\rho_{RA}\|\rho_R\otimes\sigma_A)\big),$$

where $v_n \leq (n+1)^{|R|+|A|}$ is the number of distinct eigenvalues of $\rho_R^{\otimes n} \otimes \sigma_A^{\otimes n}$. Letting $n \to \infty$, and then maximizing the r.h.s. over $s \in (0,1]$ and $\sigma_A \in \mathcal{S}(A)$, we arrive at

$$E_{R:A}^{\mathrm{dec}}(\rho_{RA}, r) \geq \liminf_{n\to\infty} \frac{-1}{n} \log P_{R^n:A^n}^{\mathrm{dec}}\big(\rho_{RA}^{\otimes n}, nr\big) \geq \sup_{0<s\leq 1} \Big\{ s\big(r - \frac{1}{2} I_{1+s}(R:A)_\rho\big)\Big\},$$

which is Eq. (11). For Eq. (12), we apply Proposition 19 and Theorem 15 to obtain

$$\begin{aligned} E_{R:A}^{\mathrm{dec}}(\rho_{RA}, r) &= \limsup_{n\to\infty} \frac{-1}{n} \log P_{R^n:A^n}^{\mathrm{dec}}\big(\rho_{RA}^{\otimes n}, nr\big) \\ &\leq \lim_{n\to\infty} \frac{-1}{n} \log \delta_{R^n:A^n}\big(\rho_{RA}^{\otimes n}, 2nr\big) \\ &= \sup_{s\geq 0} \Big\{ s\big(r - \frac{1}{2} I_{1+s}(R:A)_\rho\big)\Big\}. \end{aligned}$$

At last, Eq. (10) follows from Eq. (11) and Eq. (12). To see this, we consider the function

$$f(s) = s\big(r - \frac{1}{2} I_{1+s}(R:A)_\rho\big) \tag{49}$$

for $s \in [0, \infty)$. $f(s)$ is concave because $s \mapsto sI_{1+s}(R:A)_\rho$ is convex. So, we have that $f(s)$ reaches the maximum in $[0,1]$ and therefore

$$\max_{s\geq 0} f(s) = \max_{0\leq s\leq 1} f(s), \tag{50}$$

if $f'(1) \leq 0$. This condition is equivalent to $r \leq R_{\mathrm{critical}} = \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}s} sI_{1+s}(R:A)_\rho\big|_{s=1}$. □

Remark 20 *Some comments on the main results are as follows.*

(a) Although the reliability function $E_{R:A}^{\mathrm{dec}}(\rho_{RA}, r)$ is defined as a limit superior, we see in the above proof that when we are able to determine it for $r \leq R_{\mathrm{critical}}$, the limit of $\frac{-1}{n} \log P_{R^n:A^n}^{\mathrm{dec}}\big(\rho_{RA}^{\otimes n}, nr\big)$ actually exists.

(b) When we employ the convex-split lemma (Lemma 13) to derive the achievability results of Theorem 7 and Eq. (11), we can keep using the relative entropy as the measure of closeness between states. Meanwhile, Eq. (12) can be converted into an upper bound for the rate of exponential decay of the performance measured by relative entropy, making use of the relation $P(\rho,\sigma) \leq \sqrt{(\ln 2) D(\rho\|\sigma)}$. This actually enables us to derive results similar to Theorem 8, with the purified distance being replaced by the relative entropy in the definition of the reliability function.

(c) The upper bound of Eq. (12) diverges when $r > \frac{1}{2} I_{\max}(R:A)_\rho$, which can be easily seen. One may guess that it goes to ∞ when $r \nearrow \frac{1}{2} I_{\max}(R:A)_\rho$. However, we show in Proposition 21 in the Appendix that it is actually bounded when $r < \frac{1}{2} I_{\max}(R:A)_\rho$.

V. PROOF OF THE RELATIONS BETWEEN QUANTUM TASKS

In this section, we prove Proposition 6 on the relation between the three types of catalytic quantum information decoupling, and Proposition 11 on the relation between catalytic quantum information decoupling and quantum state merging.

A. Proof of the relation between decouplings

Proposition 6 (restatement) *For $\rho_{RA} \in \mathcal{S}(RA)$ and $k \geq 0$, $r \geq 0$, we have*

$$P^{\text{dec}}_{R:A}(\rho_{RA}, k) = P^{\text{dec-m}}_{R:A}(\rho_{RA}, 2k) = P^{\text{dec-u}}_{R:A}(\rho_{RA}, 2k), \tag{51}$$
$$E^{\text{dec}}_{R:A}(\rho_{RA}, r) = E^{\text{dec-m}}_{R:A}(\rho_{RA}, 2r) = E^{\text{dec-u}}_{R:A}(\rho_{RA}, 2r). \tag{52}$$

Proof. Eq. (52) follows from Eq. (51) directly. So it suffices to prove Eq. (51). In the following, we show the identities $P^{\text{dec}}_{R:A}(\rho_{RA}, k) = P^{\text{dec-m}}_{R:A}(\rho_{RA}, 2k)$ and $P^{\text{dec}}_{R:A}(\rho_{RA}, k) = P^{\text{dec-u}}_{R:A}(\rho_{RA}, 2k)$ separately.

We first prove the former identity. Consider an arbitrary decoupling scheme via removing a subsystem of size not more than k qubits. Let $\sigma_{A'}$ be the catalytic state and $U : \mathcal{H}_{AA'} \to \mathcal{H}_{A_1A_2}$ be the unitary operation, where A_2 is the system to be removed and $\log|A_2| \leq k$. Introduce a system A'' in the maximally mixed state $\frac{\mathbb{1}_{A''}}{|A''|}$ such that $|A''| = |A_2|$. We construct a decoupling scheme via projective measurement: the catalytic state is chosen to be $\sigma_{A'} \otimes \frac{\mathbb{1}_{A''}}{|A''|}$, and the projective measurement is given by $\{Q^x_{AA'A''}\}_{x=1}^m$, with $m = |A_2|^2$ and

$$Q^x_{AA'A''} = (U \otimes \mathbb{1}_{A''})^* \big((W^x_{A_2} \otimes \mathbb{1}_{A''}) \Psi_{A_2A''} (W^x_{A_2} \otimes \mathbb{1}_{A''})^* \otimes \mathbb{1}_{A_1} \big) (U \otimes \mathbb{1}_{A''}),$$

where $\{W^x_{A_2}\}_{x=1}^m$ is the set of discrete Weyl operators on $\mathcal{H}_{A_2}$ and $\Psi_{A_2A''}$ is the maximally entangled state. So, the cost is $\log m = \log|A_2|^2 \leq 2k$. By the definition of $P^{\text{dec-m}}_{R:A}$, for arbitrary state ω_{A_1}, we have

$$\begin{aligned}
&P^{\text{dec-m}}_{R:A}(\rho_{RA}, 2k) \\
\leq& P\Big(\sum_{x=1}^m Q^x_{AA'A''} \big(\rho_{RA} \otimes \sigma_{A'} \otimes \frac{\mathbb{1}_{A''}}{|A''|}\big) Q^x_{AA'A''},\ \rho_R \otimes U^* \big(\omega_{A_1} \otimes \frac{\mathbb{1}_{A_2}}{|A_2|}\big) U \otimes \frac{\mathbb{1}_{A''}}{|A''|} \Big) \\
=& P\Big(U^* \big(\operatorname{Tr}_{A_2} \left[U(\rho_{RA} \otimes \sigma_{A'}) U^* \right] \otimes \frac{\mathbb{1}_{A_2}}{|A_2|} \big) U \otimes \frac{\mathbb{1}_{A''}}{|A''|},\ \rho_R \otimes U^* \big(\omega_{A_1} \otimes \frac{\mathbb{1}_{A_2}}{|A_2|}\big) U \otimes \frac{\mathbb{1}_{A''}}{|A''|} \Big) \\
=& P\big(\operatorname{Tr}_{A_2} U(\rho_{RA} \otimes \sigma_{A'}) U^*, \rho_R \otimes \omega_{A_1} \big),
\end{aligned} \tag{53}$$

where the third line can be verified by direct calculation. Minimizing the last line of Eq. (53) over $\sigma_{A'}$, ω_{A_1} and $U : \mathcal{H}_{AA'} \to \mathcal{H}_{A_1A_2}$, and by Definition 2, we get from Eq. (53) that

$$P^{\text{dec-m}}_{R:A}(\rho_{RA}, 2k) \leq P^{\text{dec}}_{R:A}(\rho_{RA}, k). \tag{54}$$

Next, we deal with the other direction. Let $\sigma_{A'}$ and $\{Q^x_{AA'}\}_{x=1}^m$ be the catalytic state and the projectors of a decoupling scheme via projective measurement with cost $\log m \leq 2k$. Introduce systems C and C' such that $|C| = |C'| = \sqrt{m}$ and denote by $\Psi_{CC'}$ the maximally entangled state on C and C'. We construct a decoupling scheme via removing a subsystem: the catalytic state is given by $\sigma_{A'} \otimes \frac{\mathbb{1}_{CC'}}{|C| \cdot |C'|}$, and the unitary operation is given by

$$U_{AA'CC'} = \sum_{x=1}^m \Big(\sum_{y=1}^m \mathrm{e}^{\frac{2\pi \mathrm{i} xy}{m}} Q^y_{AA'} \Big) \otimes \big(W^x_C \otimes \mathbb{1}_{C'} \big) \Psi_{CC'} \big(W^x_C \otimes \mathbb{1}_{C'} \big)^*,$$

where $\{W^x_C\}_{x=1}^m$ is the set of discrete Weyl operators on $\mathcal{H}_C$. Set $A_1 \equiv AA'C$ be the remaining system, and $A_2 \equiv C'$ be the removed system. So, the cost is $\log|A_2| = \log|C'| \leq k$. By the definition of $P^{\text{dec}}_{R:A}$, for arbitrary state $\omega_{AA'}$, we have

$$\begin{aligned}
&P^{\text{dec}}_{R:A}(\rho_{RA}, k)\\
\le &P\Big(\operatorname{Tr}_{C'}\big[U_{AA'CC'}\big(\rho_{RA}\otimes\sigma_{A'}\otimes\frac{\mathbb{1}_{CC'}}{|C|\cdot|C'|}\big)U^{*}_{AA'CC'}\big],\, \rho_R\otimes\omega_{AA'}\otimes\frac{\mathbb{1}_C}{|C|}\Big)\\
= &P\Big(\sum_x Q^x_{AA'}\big(\rho_{RA}\otimes\sigma_{A'}\big)Q^x_{AA'}\otimes\frac{\mathbb{1}_C}{|C|},\, \rho_R\otimes\omega_{AA'}\otimes\frac{\mathbb{1}_C}{|C|}\Big)\\
= &P\Big(\sum_x Q^x_{AA'}\big(\rho_{RA}\otimes\sigma_{A'}\big)Q^x_{AA'},\, \rho_R\otimes\omega_{AA'}\Big),
\end{aligned} \tag{55}$$

where the third line can be verified by direct calculation. Minimizing the last line of Eq. (55) over $\sigma_{A'}$, $\omega_{AA'}$ and $\{Q^x_{AA'}\}_{x=1}^m$, and by Definition 3, we get from Eq. (55) that

$$P^{\text{dec}}_{R:A}(\rho_{RA}, k) \le P^{\text{dec-m}}_{R:A}(\rho_{RA}, 2k). \tag{56}$$

The combination of Eq. (54) and Eq. (56) leads to the identity $P^{\text{dec}}_{R:A}(\rho_{RA}, k) = P^{\text{dec-m}}_{R:A}(\rho_{RA}, 2k)$.

Now we prove the latter identity. Consider an arbitrary decoupling scheme via removing a subsystem of size not more than k qubits. Let $\sigma_{A'}$ be the catalytic state and $U : \mathcal{H}_{AA'} \to \mathcal{H}_{A_1A_2}$ be the unitary operation, where A_2 is the system to be removed and $\log|A_2| \le k$. We construct a decoupling scheme via random unitary operation: the catalytic state is chosen to be $\sigma_{A'}$ and the random unitary operation is given by

$$\Lambda_{AA'}(\cdot) = \frac{1}{|A_2|^2}\sum_{i=1}^{|A_2|^2}(\mathbb{1}_{A_1}\otimes W^i_{A_2})U(\cdot)U^*(\mathbb{1}_{A_1}\otimes W^i_{A_2})^*, \tag{57}$$

where $\{W^i_{A_2}\}_{i=1}^{|A_2|^2}$ are all the discrete Weyl operators on $\mathcal{H}_{A_2}$. By the definition of $P^{\text{dec-u}}_{R:A}$, for arbitrary state ω_{A_1}, we have

$$\begin{aligned}
&P^{\text{dec-u}}_{R:A}(\rho_{RA}, 2k)\\
\le &P\Big(\Lambda_{AA'}(\rho_{RA}\otimes\sigma_{A'}), \rho_R\otimes\omega_{A_1}\otimes\frac{\mathbb{1}_{A_2}}{|A_2|}\Big)\\
= &P\Big(\operatorname{Tr}_{A_2}\big[U(\rho_{RA}\otimes\sigma_{A'})U^*\big]\otimes\frac{\mathbb{1}_{A_2}}{|A_2|}, \rho_R\otimes\omega_{A_1}\otimes\frac{\mathbb{1}_{A_2}}{|A_2|}\Big)\\
= &P\Big(\operatorname{Tr}_{A_2}\big[U(\rho_{RA}\otimes\sigma_{A'})U^*\big], \rho_R\otimes\omega_{A_1}\Big).
\end{aligned} \tag{58}$$

Minimizing the last line of Eq. (58) over $\sigma_{A'}$, ω_{A_1} and $U : \mathcal{H}_{AA'} \to \mathcal{H}_{A_1A_2}$, and by Definition 2, we get from Eq. (58) that

$$P^{\text{dec-u}}_{R:A}(\rho_{RA}, 2k) \le P^{\text{dec}}_{R:A}(\rho_{RA}, k). \tag{59}$$

Next, we deal with the other direction. Let $\sigma_{A'}$ and $\Lambda_{AA'}$ be the catalytic state and the random unitary channel of a decoupling scheme via random unitary operation with cost bounded by $2k$ bits. $\Lambda_{AA'}$ can be written as

$$\Lambda_{AA'}(\cdot) = \frac{1}{m}\sum_{i=1}^{m} V^i_{AA'}(\cdot)(V^i_{AA'})^*, \tag{60}$$

where $V^i_{AA'} \in \mathcal{U}(AA')$ and $\log m \leq 2k$. Introduce systems C and C' such that $|C| = |C'| = \sqrt{m}$ and denote by $\Psi_{CC'}$ the maximally entangled state on C and C'. We construct a decoupling scheme via removing a subsystem: the catalytic state is given by

$$\sigma_{A'CC'} = \frac{1}{m}\sum_{i=1}^{m} \sigma_{A'} \otimes (W^i_C \otimes \mathbb{1}_{C'})\Psi_{CC'}(W^i_C \otimes \mathbb{1}_{C'})^*$$

with $\{W^i_C\}_{i=1}^m$ be all the discrete Weyl operators on $\mathcal{H}_C$, and the unitary operation is given by

$$U_{AA'CC'} = \sum_{i=1}^{m} V^i_{AA'} \otimes (W^i_C \otimes \mathbb{1}_{C'})\Psi_{CC'}(W^i_C \otimes \mathbb{1}_{C'})^*.$$

Set $A_1 \equiv AA'C$ be the remaining system, and $A_2 \equiv C'$ be the removed system. So, the cost is $\log|A_2| = \log|C'| \leq k$. By the definition of $P^{\text{dec}}_{R:A}$, for arbitrary state $\omega_{AA'}$, we have

$$\begin{aligned}
&P^{\text{dec}}_{R:A}(\rho_{RA}, k)\\
\leq &P\Big(\operatorname{Tr}_{C'}\left[U_{AA'CC'}(\rho_{RA} \otimes \sigma_{A'CC'})U^*_{AA'CC'}\right], \rho_R \otimes \omega_{AA'} \otimes \frac{\mathbb{1}_C}{|C|}\Big)\\
= &P\Big(\Lambda_{AA'}(\rho_{RA} \otimes \sigma_{A'}) \otimes \frac{\mathbb{1}_C}{|C|}, \rho_R \otimes \omega_{AA'} \otimes \frac{\mathbb{1}_C}{|C|}\Big)\\
= &P\big(\Lambda_{AA'}(\rho_{RA} \otimes \sigma_{A'}), \rho_R \otimes \omega_{AA'}\big).
\end{aligned} \tag{61}$$

Minimizing the last line of Eq. (61) over $\sigma_{A'}$, $\Lambda_{AA'}$ and $\omega_{AA'}$, and by Definition 4, we get from Eq. (61) that

$$P^{\text{dec}}_{R:A}(\rho_{RA}, k) \leq P^{\text{dec-u}}_{R:A}(\rho_{RA}, 2k). \tag{62}$$

The combination of Eq. (59) and Eq. (62) completes the proof. □

B. Proof of the relation between decoupling and quantum state merging

Proposition 11 (restatement) *For a tripartite pure state $\rho_{RAB} \in \mathcal{S}(RAB)$ and $k \geq 0$, $r \geq 0$, we have*

$$P^{\text{merg}}_{A\Rightarrow B}(\rho_{RAB}, k) = P^{\text{merg}}_{A\to B}(\rho_{RAB}, 2k) = P^{\text{dec}}_{R:A}(\rho_{RA}, k), \tag{63}$$

$$E^{\text{merg}}_{A\Rightarrow B}(\rho_{RAB}, r) = E^{\text{merg}}_{A\to B}(\rho_{RAB}, 2r) = E^{\text{dec}}_{R:A}(\rho_{RA}, r). \tag{64}$$

Proof. Eq. (64) follows from Eq. (63) directly. So it suffices to prove Eq. (63). Assisted by free entanglement, a merging protocol via quantum communication of k qubits, can simulate a merging protocol via classical communication of $2k$ bits, by using teleportation. Conversely, a merging protocol via classical communication of $2k$ bits, can simulate a merging protocol via quantum communication of k qubits, by using dense coding. So, we easily obtain the relation $P^{\text{merg}}_{A\Rightarrow B}(\rho_{RAB}, k) = P^{\text{merg}}_{A\to B}(\rho_{RAB}, 2k)$. In the following, we prove $P^{\text{merg}}_{A\Rightarrow B}(\rho_{RAB}, k) = P^{\text{dec}}_{R:A}(\rho_{RA}, k)$.

Consider a decoupling scheme with catalytic state $\sigma_{A'}$, in which the system A_2 of size $\log|A_2| \leq k$ is removed after a unitary operator $U_{AA'\to A_1A_2}$ being applied. Let ω_{A_1} be an arbitrary state on A_1. Let $\sigma_{A'B'}$ and $\omega_{A_1B_1}$ be the purification of $\sigma_{A'}$ and ω_{A_1}, respectively. By Uhlmann's theorem [52], there exists an isometry $V_{A_2BB'\to ABB_1}$ such that

$$\begin{aligned}&P\big(\operatorname{Tr}_{A_2}\left[U_{AA'\to A_1A_2}(\rho_{RA}\otimes\sigma_{A'})U^*_{AA'\to A_1A_2}\right],\rho_R\otimes\omega_{A_1}\big)\\=&P\big(V_{A_2BB'\to ABB_1}U_{AA'\to A_1A_2}(\rho_{RAB}\otimes\sigma_{A'B'})U^*_{AA'\to A_1A_2}V^*_{A_2BB'\to ABB_1},\rho_{RAB}\otimes\omega_{A_1B_1}\big).\end{aligned}\tag{65}$$

We construct a merging protocol $\mathcal{M}_1$ which consists of using shared entanglement $\sigma_{A'B'}$, Alice applying $U_{AA'\to A_1A_2}$ and sending A_2 to Bob, and Bob applying isometry $V_{A_2BB'\to ABB_1}$. The cost of quantum communication is $\log|A_2|\le k$. We have

$$\mathcal{M}_1(\rho_{RAB})=\operatorname{Tr}_{A_1B_1}\left[V_{A_2BB'\to ABB_1}U_{AA'\to A_1A_2}(\rho_{RAB}\otimes\sigma_{A'B'})U^*_{AA'\to A_1A_2}V^*_{A_2BB'\to ABB_1}\right].\tag{66}$$

By the monotonicity of purified distance under partial trace, we get from Eq. (65) and Eq. (66) that

$$P\big(\mathcal{M}_1(\rho_{RAB}),\rho_{RAB}\big)\le P\big(\operatorname{Tr}_{A_2}\left[U_{AA'\to A_1A_2}(\rho_{RA}\otimes\sigma_{A'})U^*_{AA'\to A_1A_2}\right],\rho_R\otimes\omega_{A_1}\big)\tag{67}$$

Eq. (67) together with the definitions of $P^{\rm merg}_{A\Rightarrow B}$ and $P^{\rm dec}_{R:A}$ implies

$$P^{\rm merg}_{A\Rightarrow B}(\rho_{RAB},k)\le P^{\rm dec}_{R:A}(\rho_{RA},k).\tag{68}$$

Now, we turn to the proof of the other direction. Let $\mathcal{M}_1$ be a merging protocol via quantum communication of not more than k qubits. We write $\mathcal{M}_1$ as

$$\mathcal{M}_1(\cdot)=\operatorname{Tr}_{A_1B_1}\left[V_{A_2BB'\to ABB_1}U_{AA'\to A_1A_2}\big((\cdot)\otimes\phi_{A'B'}\big)U^*_{AA'\to A_1A_2}V^*_{A_2BB'\to ABB_1}\right],$$

where $\phi_{A'B'}$ is the shared entanglement, $U_{AA'\to A_1A_2}$ is performed by Alice, $V_{A_2BB'\to ABB_1}$ is performed by Bob, and A_2 with $\log|A_2|\le k$ is sent from Alice to Bob. Using Uhlmann's theorem again, we know that there exists a pure state $\varphi_{A_1B_1}$ such that

$$\begin{aligned}&P\big(\mathcal{M}_1(\rho_{RAB}),\rho_{RAB}\big)\\=&P\big(V_{A_2BB'\to ABB_1}U_{AA'\to A_1A_2}\big(\rho_{RAB}\otimes\phi_{A'B'}\big)U^*_{AA'\to A_1A_2}V^*_{A_2BB'\to ABB_1},\rho_{RAB}\otimes\varphi_{A_1B_1}\big).\end{aligned}\tag{69}$$

Because the purified distance decreases under the action of partial trace, we get

$$P(\mathcal{M}_1(\rho_{RAB}),\rho_{RAB})\ge P\big(\operatorname{Tr}_{A_2}\left[U_{AA'\to A_1A_2}\big(\rho_{RA}\otimes\phi_{A'}\big)U^*_{AA'\to A_1A_2}\right],\rho_R\otimes\varphi_{A_1}\big).\tag{70}$$

Since $\operatorname{Tr}_{A_2}\left[U_{AA'\to A_1A_2}\big(\rho_{RA}\otimes\phi_{A'}\big)U^*_{AA'\to A_1A_2}\right]$ in the right hand side of Eq. (70) describes a decoupling scheme via removing $\log|A_2|\le k$ qubits, by the definitions of $P^{\rm merg}_{A\Rightarrow B}$ and $P^{\rm dec}_{R:A}$, we conclude that

$$P^{\rm merg}_{A\Rightarrow B}(\rho_{RAB},k)\ge P^{\rm dec}_{R:A}(\rho_{RA},k).\tag{71}$$

□

VI. DISCUSSION

In this work, we have characterized the reliability function of catalytic quantum information decoupling using the sandwiched Rényi divergence, and have applied it to quantum state merging. Our results add new operational meanings to the sandwiched Rényi information quantities.

The availability of a catalytic system is crucial in our derivations. This is not only for the characterization of the reliability function (Theorem 8), but also for the interplay between different decoupling operations (Proposition 6) and the interplay between decoupling and state merging (Proposition 11). The introduction of the catalyst in decoupling [20, 21] helps us solve the problem.

Before ending the paper, we list a few open problems.

1. Firstly, to characterize the reliability function of quantum information decoupling without involving any catalytic system is an interesting open problem. It is known that when only an asymptotically vanishing error is concerned, the best rate of cost is the same no matter a catalytic system is allowed or not [1–3, 20, 21]. However, for the reliability function we do not know whether catalysis makes a difference.

2. Another open problem is to derive the reliability function when the rate of cost is above the critical value. The upper and lower bounds obtained in the present work do not match in the high-rate case. However, note that the existence of critical points in the study of reliability functions is a common phenomenon, where at the unsolved side the problem becomes more of a combinatorial feature and is hard to tackle [30, 34, 46, 62].

3. Moreover, it is interesting to consider the reliability function when the performance is measured by the trace distance. Our method works for the purified distance, and equivalently, for the fidelity. But for trace distance, quite different techniques may be needed. It is worth mentioning that, as was pointed out in [23], for trace distance there is no obvious way to transfer the reliability function from quantum information decoupling to other quantum tasks (such as quantum state merging). This is in contrast to what we have done in Proposition 11 and Corollary 12 which make the transference by employing the Uhlmann's theorem.

4. Finally, we hope that our work will be stimulating and find more applications, in deriving the reliability functions for more quantum information tasks. Indeed, in a later work [76] we have applied the present results to quantum channel simulation. Can we extend this line of research to entanglement-assisted communication of noisy channels [77], and to the treatment of the amount of comsumed/generated entanglement in quantum tasks? These problems are closely related to the size of the remaining system in quantum information decoupling.

Appendix A: Boundedness of the Upper Bound $E_u(r)$

Proposition 21 *Let $\rho_{RA} \in \mathcal{S}(RA)$ be given. The function $E_u(r) = \sup_{s\geq 0}\big\{s\big(r - \frac{1}{2}I_{1+s}(R:A)_\rho\big)\big\}$ is bounded in the interval $(-\infty, \frac{1}{2}I_{\max}(R:A)_\rho)$.*

Proof. We show that there is a constant C, such that for any $\epsilon > 0$, $E_u\big(\frac{1}{2}I_{\max}(R:A)_\rho - \epsilon\big) \leq C$. Proposition 16 establishes that

$$E_u(r) = \lim_{n\to\infty} \frac{-1}{2n} \log \operatorname{Tr}\left[\mathcal{E}^n\big(\rho_{AB}^{\otimes n}\big)\Big\{\mathcal{E}^n\big(\rho_{AB}^{\otimes n}\big) \geq 2^{n\cdot 2r}\rho_A^{\otimes n} \otimes \omega_{B^n}^{(n)}\Big\}\right], \tag{A1}$$

where $\omega_{B^n}^{(n)}$ is the universal symmetric state of Lemma 1 and $\mathcal{E}^n$ is the pinching map associated with $\rho_A^{\otimes n}\otimes\omega_{B^n}^{(n)}$. Setting $M_A=\rho_A$ and letting $s\to\infty$ in Lemma 17, we get

$$\frac{1}{n}D_{\max}\Big(\mathcal{E}^n(\rho_{AB}^{\otimes n})\big\|\rho_A^{\otimes n}\otimes\omega_{B^n}^{(n)}\Big)=I_{\max}(A:B)_\rho+O\Big(\frac{\log n}{n}\Big).$$

This implies that, for arbitrary $\epsilon>0$, there exists a common eigenvector $|\varphi_n\rangle$ of $\mathcal{E}^n(\rho_{AB}^{\otimes n})$ and $\rho_A^{\otimes n}\otimes\omega_{B^n}^{(n)}$, such that for n big enough,

$$\Big\{\mathcal{E}^n\big(\rho_{AB}^{\otimes n}\big)\geq 2^{n(I_{\max}(A:B)_\rho-2\epsilon)}\rho_A^{\otimes n}\otimes\omega_{B^n}^{(n)}\Big\}\geq|\varphi_n\rangle\langle\varphi_n| \tag{A2}$$

and

$$\begin{aligned}\langle\varphi_n|\mathcal{E}^n\big(\rho_{AB}^{\otimes n}\big)|\varphi_n\rangle&\geq 2^{n(I_{\max}(A:B)_\rho-2\epsilon)}\langle\varphi_n|\rho_A^{\otimes n}\otimes\omega_{B^n}^{(n)}|\varphi_n\rangle\\&\geq 2^{n(I_{\max}(A:B)_\rho-2\epsilon)}\lambda_{\min}\big(\rho_A^{\otimes n}\otimes\omega_{B^n}^{(n)}\big)\\&\geq 2^{n(I_{\max}(A:B)_\rho-2\epsilon)}\big(\lambda_{\min}(\rho_A)\big)^n\frac{1}{g_{n,|B|}\,|B|^n}.\end{aligned} \tag{A3}$$

In Eq. (A3), we have used the fact $g_{n,|B|}\omega_{B^n}^{(n)}\geq(\frac{\mathbb{1}_B}{|B|})^{\otimes n}$ and $\lambda_{\min}(X)$ denotes the minimal eigenvalue of X. Combining Eq. (A1), Eq. (A2) and Eq. (A3), we obtain

$$\begin{aligned}E_u\Big(\frac{1}{2}I_{\max}(A:B)_\rho-\epsilon\Big)&\leq\lim_{n\to\infty}\frac{-1}{2n}\log\langle\varphi_n|\mathcal{E}^n\big(\rho_{AB}^{\otimes n}\big)|\varphi_n\rangle\\&\leq\frac{1}{2}\log\frac{|B|}{\lambda_{\min}(\rho_A)}-\frac{1}{2}I_{\max}(A:B)_\rho+\epsilon.\end{aligned} \tag{A4}$$

At last, since $E_u(r)$ is monotonically increasing, we can choose $C=\frac{1}{2}\log\frac{|B|}{\lambda_{\min}(\rho_A)}-\frac{1}{2}I_{\max}(A:B)_\rho$. $\square$

ACKNOWLEDGMENTS

The authors are grateful to Mark Wilde for bringing to their attention References [27, 63, 64] and to Masahito Hayashi for pointing out Reference [74]. They further thank Nilanjana Datta, Masahito Hayashi, Marco Tomamichel, Mark Wilde and Xiao Xiong for comments or discussions on related topics. The research of KL was supported by NSFC (No. 61871156, No. 12031004), and the research of YY was supported by NSFC (No. 61871156, No. 12071099).

[1] Michał Horodecki, Jonathan Oppenheim, and Andreas Winter. Partial quantum information. *Nature*, 436(7051):673–676, 2005.

[2] Michał Horodecki, Jonathan Oppenheim, and Andreas Winter. Quantum state merging and negative information. *Communications in Mathematical Physics*, 269(1):107–136, 2007.

[3] Anura Abeyesinghe, Igor Devetak, Patrick Hayden, and Andreas Winter. The mother of all protocols: restructuring quantum information's family tree. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 465(2108):2537–2563, 2009.

[4] Igor Devetak and Jon Yard. Exact cost of redistributing multipartite quantum states. *Physical Review Letters*, 100(23):230501, 2008.

[5] Jon T Yard and Igor Devetak. Optimal quantum source coding with quantum side information at the encoder and decoder. *IEEE Transactions on Information Theory*, 55(11):5339–5351, 2009.

[6] Charles H Bennett, Igor Devetak, Aram W Harrow, Peter W Shor, and Andreas Winter. The quantum reverse Shannon theorem and resource tradeoffs for simulating quantum channels. *IEEE Transactions on Information Theory*, 60(5):2926–2959, 2014.

[7] Mario Berta, Matthias Christandl, and Renato Renner. The quantum reverse Shannon theorem based on one-shot information theory. *Communications in Mathematical Physics*, 306(3):579–615, 2011.

[8] Mario Berta, Fernando GSL Brandão, Christian Majenz, and Mark M Wilde. Conditional decoupling of quantum information. *Physical Review Letters*, 121(4):040504, 2018.

[9] Mario Berta, Omar Fawzi, and Stephanie Wehner. Quantum to classical randomness extractors. *IEEE Transactions on Information Theory*, 60(2):1168–1192, 2013.

[10] Lidia Del Rio, Johan Åberg, Renato Renner, Oscar Dahlsten, and Vlatko Vedral. The thermodynamic meaning of negative entropy. *Nature*, 474(7349):61–63, 2011.

[11] Fernando GSL Brandão and Michał Horodecki. An area law for entanglement from exponential decay of correlations. *Nature Physics*, 9(11):721–726, 2013.

[12] Fernando GSL Brandão and Michał Horodecki. Exponential decay of correlations implies area law. *Communications in Mathematical Physics*, 333(2):761–798, 2015.

[13] Johan Åberg. Truly work-like work extraction via a single-shot analysis. *Nature Communications*, 4(1):1–5, 2013.

[14] Lídia Del Rio, Adrian Hutter, Renato Renner, and Stephanie Wehner. Relative thermalization. *Physical Review E*, 94(2):022104, 2016.

[15] Patrick Hayden and John Preskill. Black holes as mirrors: quantum information in random subsystems. *Journal of High Energy Physics*, 2007(09):120, 2007.

[16] Samuel L Braunstein and Arun K Pati. Quantum information cannot be completely hidden in correlations: implications for the black-hole information paradox. *Physical Review Letters*, 98(8):080502, 2007.

[17] Samuel L Braunstein, Stefano Pirandola, and Karol Życzkowski. Better late than never: information retrieval from black holes. *Physical Review Letters*, 110(10):101301, 2013.

[18] Kamil Brádler and Christoph Adami. One-shot decoupling and page curves from a dynamical model for black hole evaporation. *Physical Review Letters*, 116:101301, 2016.

[19] Frédéric Dupuis, Mario Berta, Jürg Wullschleger, and Renato Renner. One-shot decoupling. *Communications in Mathematical Physics*, 328(1):251–284, 2014.

[20] Anurag Anshu, Vamsi Krishna Devabathini, and Rahul Jain. Quantum communication using coherent rejection sampling. *Physical Review Letters*, 119(12):120506, 2017.

[21] Christian Majenz, Mario Berta, Frédéric Dupuis, Renato Renner, and Matthias Christandl. Catalytic decoupling of quantum information. *Physical Review Letters*, 118(8):080503, 2017.

[22] Eyuri Wakakuwa and Yoshifumi Nakata. One-shot randomized and nonrandomized partial decoupling. *Communications in Mathematical Physics*, 386(2):589–649, 2021.

[23] Frédéric Dupuis. Privacy amplification and decoupling without smoothing. *arXiv preprint arXiv:2105.05342*, 2021.

[24] Oleg Szehr, Frédéric Dupuis, Marco Tomamichel, and Renato Renner. Decoupling with unitary approximate two-designs. *New Journal of Physics*, 15(5):053022, 2013.

[25] Winton Brown and Omar Fawzi. Decoupling with random quantum circuits. *Communications in Mathematical Physics*, 340(3):867–900, 2015.

[26] Yoshifumi Nakata, Christoph Hirche, Ciara Morgan, and Andreas Winter. Decoupling with random diagonal unitaries. *Quantum*, 1:18, 2017.

[27] Naresh Sharma. Random coding exponents galore via decoupling. *arXiv preprint arXiv:1504.07075*, 2015.

[28] Anurag Anshu, Mario Berta, Rahul Jain, and Marco Tomamichel. Partially smoothed information measures. *IEEE Transactions on Information Theory*, 66(8):5022–5036, 2020.

[29] Claude E Shannon. Probability of error for optimal codes in a Gaussian channel. *Bell System Technical Journal*, 38(3):611–656, 1959.

[30] R Gallager. *Information Theory and Reliable Communication.* John Wiley & Sons, 1968.

[31] M Burnashev and Alexander S Holevo. On the reliability function for a quantum communication channel. *Problems of Information Transmission*, 34(2):97–107, 1998.

[32] Andreas Winter. Coding theorems of quantum information theory. *PhD Thesis, Universität Bielefeld*, 1999.

[33] Alexander S Holevo. Reliability function of general classical-quantum channel. *IEEE Transactions on Information Theory*, 46(6):2256–2261, 2000.

[34] Marco Dalai. Lower bounds on the probability of error for classical and classical-quantum channels. *IEEE Transactions on Information Theory*, 59(12):8027–8056, 2013.

[35] Masahito Hayashi. Precise evaluation of leaked information with secure randomness extraction in the presence of quantum attacker. *Communications in Mathematical Physics*, 333(1):335–350, 2015.

[36] Marco Dalai and Andreas Winter. Constant compositions in the sphere packing bound for classical-quantum channels. *IEEE Transactions on Information Theory*, 63(9):5603–5617, 2017.

[37] Hao-Chung Cheng, Min-Hsiu Hsieh, and Marco Tomamichel. Quantum sphere-packing bounds with polynomial prefactors. *IEEE Transactions on Information Theory*, 65(5):2872–2898, 2019.

[38] Hao-Chung Cheng, Eric P Hanson, Nilanjana Datta, and Min-Hsiu Hsieh. Non-asymptotic classical data compression with quantum side information. *IEEE Transactions on Information Theory*, 67(2):902–930, 2020.

[39] Robert Koenig and Stephanie Wehner. A strong converse for classical channel coding using entangled inputs. *Physical Review Letters*, 103(7):070504, 2009.

[40] Naresh Sharma and Naqueeb Ahmad Warsi. Fundamental bound on the reliability of quantum information transmission. *Physical Review Letters*, 110(8):080501, 2013.

[41] Mark M Wilde, Andreas Winter, and Dong Yang. Strong converse for the classical capacity of entanglement-breaking and Hadamard channels via a sandwiched Rényi relative entropy. *Communications in Mathematical Physics*, 331(2):593–622, 2014.

[42] Manish K Gupta and Mark M Wilde. Multiplicativity of completely bounded p-norms implies a strong converse for entanglement-assisted capacity. *Communications in Mathematical Physics*, 334(2):867–887, 2015.

[43] Tom Cooney, Milán Mosonyi, and Mark M Wilde. Strong converse exponents for a quantum channel discrimination problem and quantum-feedback-assisted communication. *Communications in Mathematical Physics*, 344(3):797–829, 2016.

[44] Milán Mosonyi and Tomohiro Ogawa. Strong converse exponent for classical-quantum channel coding. *Communications in Mathematical Physics*, 355(1):373–426, 2017.

[45] Martin Müller-Lennert, Frédéric Dupuis, Oleg Szehr, Serge Fehr, and Marco Tomamichel. On quantum Rényi entropies: a new generalization and some properties. *Journal of Mathematical Physics*, 54(12):122203, 2013.

[46] Ke Li, Yongsheng Yao, and Masahito Hayashi. Tight exponential analysis for smoothing the max-relative entropy and for quantum privacy amplification. *arXiv preprint arXiv:2111.01075*, 2021.

[47] Milán Mosonyi and Tomohiro Ogawa. Quantum hypothesis testing and the operational interpretation of the quantum Rényi relative entropies. *Communications in Mathematical Physics*, 334(3):1617–1648, 2015.

[48] Milán Mosonyi and Tomohiro Ogawa. Two approaches to obtain the strong converse exponent of quantum hypothesis testing for general sequences of quantum states. *IEEE Transactions on Information Theory*, 61(12):6975–6994, 2015.

[49] Masahito Hayashi and Marco Tomamichel. Correlation detection and an operational interpretation of the Rényi mutual information. *Journal of Mathematical Physics*, 57(10):102201, 2016.

[50] Alexei Gilchrist, Nathan K Langford, and Michael A Nielsen. Distance measures to compare real and ideal quantum processes. *Physical Review A*, 71(6):062310, 2005.

[51] Marco Tomamichel, Roger Colbeck, and Renato Renner. A fully quantum asymptotic equipartition property. *IEEE Transactions on Information Theory*, 55(12):5840–5847, 2009.

[52] Armin Uhlmann. The "transition probability" in the state space of a $*$-algebra. *Reports on Mathematical Physics*, 9(2):273–279, 1976.

[53] W Forrest Stinespring. Positive functions on C*-algebras. *Proceedings of the American Mathematical Society*, 6(2):211–216, 1955.

[54] Masahito Hayashi. Optimal sequence of quantum measurements in the sense of Stein's lemma in quantum hypothesis testing. *Journal of Physics A: Mathematical and General*, 35(50):10759, 2002.

[55] Matthias Christandl, Robert König, and Renato Renner. Postselection technique for quantum channels with applications to quantum cryptography. *Physical Review Letters*, 102(2):020504, 2009.

[56] Masahito Hayashi. Universal coding for classical-quantum channel. *Communications in Mathematical Physics*, 289(3):1087–1098, 2009.

[57] Salman Beigi. Sandwiched Rényi divergence satisfies data processing inequality. *Journal of Mathematical Physics*, 54(12):122202, 2013.

[58] Hisaharu Umegaki. Conditional expectation in an operator algebra. *Tohoku Mathematical Journal, Second Series*, 6(2-3):177–181, 1954.

[59] Nilanjana Datta. Min-and max-relative entropies and a new entanglement monotone. *IEEE Transactions on Information Theory*, 55(6):2816–2826, 2009.

[60] Nikola Ciganović, Normand J Beaudry, and Renato Renner. Smooth max-information as one-shot generalization for mutual information. *IEEE Transactions on Information Theory*, 60(3):1573–1581, 2013.

[61] Berry Groisman, Sandu Popescu, and Andreas Winter. Quantum, classical, and total amount of correlations in a quantum state. *Physical Review A*, 72(3):032317, 2005.

[62] Masahito Hayashi and Vincent YF Tan. Equivocations, exponents, and second-order coding rates under various Rényi information measures. *IEEE Transactions on Information Theory*, 63(2):975–1005, 2016.

[63] Kun Wang, Xin Wang, and Mark M Wilde. Quantifying the unextendibility of entanglement. *arXiv preprint arXiv:1911.07433*, 2019.

[64] Sumeet Khatri and Mark M Wilde. Principles of quantum communication theory: a modern approach. *arXiv preprint arXiv:2011.04672*, 2020.

[65] Anurag Anshu, Rahul Jain, and Naqueeb Ahmad Warsi. A generalized quantum slepian–wolf. *IEEE Transactions on Information Theory*, 64(3):1436–1453, 2017.

[66] Anurag Anshu, Rahul Jain, and Naqueeb Ahmad Warsi. Building blocks for communication over noisy quantum networks. *IEEE Transactions on Information Theory*, 65(2):1287–1306, 2018.

[67] Anurag Anshu, Rahul Jain, and Naqueeb Ahmad Warsi. Convex-split and hypothesis testing approach to one-shot quantum measurement compression and randomness extraction. *IEEE Transactions on Information Theory*, 65(9):5905–5924, 2019.

[68] Anurag Anshu, Min-Hsiu Hsieh, and Rahul Jain. Quantifying resources in general resource theory with catalysts. *Physical Review Letters*, 121(19):190504, 2018.

[69] Mario Berta and Christian Majenz. Disentanglement cost of quantum states. *Physical Review Letters*, 121(19):190503, 2018.

[70] Philippe Faist, Mario Berta, and Fernando GSL Brandao. Thermodynamic implementations of quantum processes. *Communications in Mathematical Physics*, 384(3):1709–1750, 2021.

[71] Patryk Lipka-Bartosik and Paul Skrzypczyk. All states are universal catalysts in quantum thermodynamics. *Physical Review X*, 11(1):011061, 2021.

[72] Rupert L Frank and Elliott H Lieb. Monotonicity of a relative Rényi entropy. *Journal of Mathematical Physics*, 54(12):122201, 2013.

[73] Renato Renner. Security of quantum key distribution. *PhD Thesis*, 2005.

[74] Masahito Hayashi. Security analysis of ε-almost dual universal2 hash functions: smoothing of min entropy versus smoothing of Rényi entropy of order 2. *IEEE Transactions on Information Theory*, 62(6):3451–3476, 2016.

[75] Po-Ning Chen. Generalization of Gartner-Ellis theorem. *IEEE Transactions on Information Theory*, 46(7):2752–2760, 2000.

[76] Ke Li and Yongsheng Yao. Reliable simulation of quantum channels. *arXiv preprint arXiv:2112.04475*, 2021.

[77] Charles H Bennett, Peter W Shor, John A Smolin, and Ashish V Thapliyal. Entanglement-assisted capacity of a quantum channel and the reverse Shannon theorem. *IEEE Transactions on Information Theory*, 48(10):2637–2655, 2002.