SCIENCE CHINA
Information Sciences

• RESEARCH PAPER •

August 2025, Vol. 68, Iss. 8, 180509:1–180509:25
https://doi.org/10.1007/s11432-024-4488-0

Special Topic: Quantum Information

Towards the ultimate limits of quantum channel discrimination and quantum communication

Kun FANG[1*], Gilad GOUR[2*] & Xin WANG[3*]

[1]*School of Data Science, The Chinese University of Hong Kong, Shenzhen, Shenzhen 518172, China*
[2]*Faculty of Mathematics, Technion-Israel Institute of Technology, Haifa 3200003, Israel*
[3]*Thrust of Artificial Intelligence, The Hong Kong University of Science and Technology (Guangzhou), Guangzhou 511453, China*

Received 31 December 2024/Revised 8 June 2025/Accepted 10 June 2025/Published online 10 July 2025

Abstract Distinguishability is fundamental to information theory and extends naturally to quantum systems. While quantum state discrimination is well understood, quantum channel discrimination remains challenging due to the dynamic nature of channels and the variety of discrimination strategies. This work advances the understanding of quantum channel discrimination and its fundamental limits. We develop new tools for quantum divergences, including sharper bounds on the quantum hypothesis testing relative entropy and additivity results for channel divergences. We establish a quantum Stein's lemma for memoryless channel discrimination, and link the strong converse property to the asymptotic equipartition property and continuity of divergences. Notably, we prove the equivalence of exponentially strong converse properties under coherent and sequential strategies. We further explore the interplay among operational regimes, discrimination strategies, and channel divergences, deriving exponents in various settings and contributing to a unified framework for channel discrimination. Finally, we recast quantum communication tasks as discrimination problems, uncovering deep connections between channel capacities, channel discrimination, and the mathematical structure of channel divergences. These results bridge two core areas of quantum information theory and offer new insights for future exploration.

Keywords quantum information, quantum channel discrimination, quantum communication, quantum capacity, quantum hypothesis testing, quantum divergence

Citation Fang K, Gour G, Wang X. Towards the ultimate limits of quantum channel discrimination and quantum communication. Sci China Inf Sci, 2025, 68(8): 180509, https://doi.org/10.1007/s11432-024-4488-0

1 Introduction

Distinguishability is a central topic in information theory from both theoretical and practical perspectives. A fundamental framework for studying distinguishability is asymmetric hypothesis testing. In this setting, a source generates a sample x from one of two probability distributions $p \equiv \{p(x)\}_{x\in\mathcal{X}}$ or $q \equiv \{q(x)\}_{x\in\mathcal{X}}$. The objective of asymmetric hypothesis testing is to minimize the Type-II error (decides p when the fact is q) while keeping the Type-I error (decides q when the fact is p) within a certain threshold. The celebrated Chernoff-Stein's lemma [1] states that, for any constant bound on the Type-I error, the optimal Type-II error decays exponentially fast in the number of samples, and the decay rate is exactly the relative entropy (Kullback-Leibler divergence),

$$D(p\|q) = \sum_{x\in\mathcal{X}} p(x)\log_2[p(x)/q(x)]. \tag{1}$$

In particular, this lemma also states the "strong converse property", a desirable mathematical property in information theory [2] that delineates a sharp boundary for the tradeoff between the Type-I and Type-II errors in the asymptotic regime: any possible scheme with Type-II error decaying to zero with an exponent larger than the relative entropy will result in the Type-I error converging to one in the asymptotic limit. Therefore, the Chernoff-Stein's lemma provides a rigorous operational interpretation of the relative entropy and establishes a crucial connection between hypothesis testing and information theory [3].

* Corresponding author (email: kunfang@cuhk.edu.cn, giladgour@technion.ac.il, felixxinwang@hkust-gz.edu.cn)

© Science China Press 2025

A natural question is whether the above result generalizes to the quantum case. Substantial efforts have been made to answer this fundamental question in quantum information community (see, e.g., [4–15]). The basic task is quantum state discrimination, in which we are given an independent and identically distributed (i.i.d.) quantum state, which could be either $\rho^{\otimes n}$ or $\sigma^{\otimes n}$. We set that $\rho^{\otimes n}$ is the null hypothesis and $\sigma^{\otimes n}$ is the alternative hypothesis. The goal is to perform a binary measurement $\{\Pi_n, I-\Pi_n\}$ on the state to determine which hypothesis is true. The corresponding error probabilities are defined analogously to the classical case, as follows:

$$\text{(Type-I)}\quad \alpha_n(\Pi_n) := \mathrm{Tr}[(I-\Pi_n)\rho^{\otimes n}], \qquad \text{(Type-II)}\quad \beta_n(\Pi_n) := \mathrm{Tr}[\Pi_n\sigma^{\otimes n}]. \tag{2}$$

The quantum version of the Chernoff-Stein's lemma (also known as quantum Stein's lemma) states that [4, 5]

$$\lim_{n\to\infty}\frac{1}{n}D_H^{\varepsilon}\left(\rho^{\otimes n}\|\sigma^{\otimes n}\right) = D(\rho\|\sigma), \quad \forall\varepsilon\in(0,1), \tag{3}$$

where $D_H^{\varepsilon}(\rho\|\sigma) := -\log\inf\{\mathrm{Tr}[\Pi\sigma] : 0 \leqslant \Pi \leqslant I, \mathrm{Tr}[(I-\Pi)\rho] \leqslant \varepsilon\}$ denotes the quantum hypothesis testing relative entropy that characterizes the decay rate of the optimal Type-II error and $D(\rho\|\sigma) = \mathrm{Tr}[\rho(\log\rho - \log\sigma)]$ denotes the quantum relative entropy. This quantum Stein's lemma delivers a rigorous operational interpretation for the quantum relative entropy. Extended research on quantum Stein's lemma are presented in [6–9, 14–18].

Although research in quantum hypothesis testing has largely focused on quantum states, there are various situations in which quantum channels play the role of the main objects of study. The task of channel discrimination is very similar to that of state discrimination. Here, we are given black-box access to n uses of a channel $\mathcal{G}$ with the aim to identify it from candidates $\mathcal{N}$ and $\mathcal{M}$. However, quantum channel discrimination has more diversities in terms of discrimination strategies (e.g., product strategy, coherent strategy, sequential strategy) due to its nature as dynamic resources [10, 13, 19–25], which leads to several variants of the quantum channel Stein's lemma. In particular, for the coherent strategies (also known as parallel strategies in some literature), the black box can be used n times in parallel to any state with a reference system before performing a measurement at the final output to identify the channel. Based on the recently developed resource theory of asymmetric distinguishability for quantum channels, the state-of-the-art result [13] arrives at

$$\lim_{\varepsilon\to 0}\lim_{n\to\infty}\frac{1}{n}D_H^{\varepsilon}(\mathcal{N}^{\otimes n}\|\mathcal{M}^{\otimes n}) = D^{\mathrm{reg}}(\mathcal{N}\|\mathcal{M}) \tag{4}$$

with $D_H^{\varepsilon}(\mathcal{N}\|\mathcal{M})$ denotes the hypothesis testing relative entropy of quantum channels and $D^{\mathrm{reg}}(\mathcal{N}\|\mathcal{M})$ denotes the regularized quantum relative entropy. That is, for Type-I error bounded by ε, the asymptotic optimal rate of the Type-II error exponent is given by $D^{\mathrm{reg}}(\mathcal{N}\|\mathcal{M})$ when ε goes to 0.

However, the condition of vanishing ε left a notable gap to achieve the quantum channel version of Stein's lemma. Unlike state discrimination, the dynamic feature of quantum channels raises challenging difficulties in determining the optimal discrimination scheme as we have to handle the additional optimization of the input states and the non-i.i.d. structure of the testing states. To fill the gap, it requires a deeper understanding and analysis on the error exponent in hypothesis testing of quantum channels. The solution could promptly advance our understanding of quantum channel discrimination, quantum communication [26–29], and the related field of quantum metrology [30–32]. Beyond the quantum channel Stein's lemma, various channel divergences have emerged to analyze different regimes of quantum channel discrimination. Establishing a unified framework that encompasses these divergences and discrimination regimes is a desirable step toward a deeper understanding of the manipulation and operational characterization of quantum channels.

In this work, we provide a study towards the ultimate limits of quantum channel discrimination and quantum communication. Our contributions are summarized as follows.

• In Section 2, we present several technical advancements in quantum divergences for quantum states and channels. Specifically, we provide a quantitative improvement in lower bounding the quantum hypothesis testing relative entropy using the Petz Rényi divergence, addressing an open question posed by Nuradha and Wilde in [33, Remark 4]. Additionally, we demonstrate that the previously explored amortized and regularized channel divergences are generally additive under the tensor product of distinct quantum channels. These technical results are expected to be of independent interest and provide valuable tools for future research.

• In Section 3, we investigate the limits of the unstabilized quantum channel divergences and prove a quantum channel analog of Stein's lemma without quantum memory assistance. To further strengthen the result, we introduce the (exponentially) strong converse properties for channel discrimination and establish its equivalence to the asymptotic equipartition property (AEP) of various quantum channel divergences as well as the continuity of the quantum channel Rényi divergence. Leveraging these equivalent characterizations, we demonstrate, rather surprisingly, that the exponentially strong converse properties under coherent and sequential strategies are equivalent.

• In Section 4, we study the interplay between the strategies of channel discrimination (e.g., sequential, coherent, product), the operational regimes (e.g., error exponent, Stein exponent, strong converse exponent), and three variants of channel divergences (e.g., Petz, Umegaki, sandwiched). We find a nice correspondence that shows that the proper divergences to use (Petz, Umegaki, sandwiched) are determined by the operational regime of interest, while the types of channel extension (one-shot, regularized, amortized) are determined by the discrimination strategies. We determine the exponents of quantum channel discrimination in various regimes and contribute towards a complete picture of channel discrimination in a unified framework.

• In Section 5, we present a new perspective by framing the study of quantum communication problems as quantum channel discrimination tasks. This offers deeper insights into the intricate relationships between channel capacities, channel discrimination, and the mathematical properties of quantum channel divergences. Leveraging this connection, we demonstrate that the channel coherent information and quantum channel capacity can be precisely characterized as Stein exponent for discriminating between two quantum channels under product and coherent strategies without quantum memory assistance, respectively. Furthermore, we show that the strong converse property of quantum channel capacity, a long-standing open problem in quantum information theory, can be established if the channels being discriminated exhibit the strong converse property.

Our technical results are primarily presented in terms of the unstabilized channel divergence, a versatile yet less explored notion of channel divergence compared to the more commonly studied divergences in the literature. This concept naturally arises in the context of quantum communication problems, offering a broader framework for analysis. Given the extensive applications of quantum divergences and quantum channel discrimination [13,24,26–28,34–42], this work contributes to a more comprehensive understanding of the ultimate limits of quantum channel discrimination in various regimes. Moreover, it provides a novel perspective on quantum communication problems by framing them as tasks of quantum channel discrimination, thereby bridging two fundamental areas of quantum information science and paving the way for addressing the remaining challenges in future studies.

2 Preliminaries

In this section, we introduce the notations to be employed throughout the paper. Subsequently, we investigate the mathematical tool of quantum divergences as applied to both states and channels. Following this, we review the operational task of quantum channel discrimination under different strategies and consider scenarios both with and without quantum memory assistance.

2.1 Notation

In this paper, we only consider finite-dimensional Hilbert spaces, which we denote with capital Latin letters such as C. The dimension of a Hilbert space C is denoted by $|C|$. The set of linear operators on Hilbert space C is denoted by $\mathfrak{L}(C)$ and the set of density matrices acting on it by $\mathfrak{D}(C)$. Density matrices are represented by small Greek letters such as ρ_C, where the subscript indicates that ρ acts on C. For a state $\rho_{AB} \in \mathfrak{D}(AB)$ we will also use the convention that $\rho_A = \mathrm{Tr}_B[\rho_{AB}]$ denotes the marginal on system A. The support of an operator X is denoted by $\mathrm{supp}(X)$. The projector onto the subspace spanned by the positive eigenvalues of X is represented as X_+. The identity operator is denoted by I and the maximally mixed state is denoted by π. Quantum channels will be denoted by calligraphic large Latin letters such as $\mathcal{N}$ and the set of all quantum channels from A to B by $\mathrm{CPTP}(A \to B)$, which stands for completely positive and trace-preserving maps. The identity channel is represented by $\mathcal{I}$, while the replacer channel is denoted as $\mathcal{R}^\sigma$, which maps any input state to the fixed state σ. Throughout the paper, we take the logarithm to be base two unless stated otherwise.

2.2 Quantum divergences

A divergence between two quantum states is defined as a real-valued function $\boldsymbol{D} : \mathfrak{D} \times \mathfrak{D} \to \mathbb{R} \cup \{\infty\}$ subject to the data processing inequality $\boldsymbol{D}(\mathcal{E}(\rho)\|\mathcal{E}(\sigma)) \leqslant \boldsymbol{D}(\rho\|\sigma)$ for all quantum states $\rho, \sigma \in \mathfrak{D}(A)$ and quantum channel $\mathcal{E} \in \mathrm{CPTP}(A \to B)$. Divergences serve as crucial tools for quantifying the distinguishability of quantum states. In our discussion, we will frequently employ the following quantum divergences, which hold particular relevance.

Definition 1 (Umegaki relative entropy)**.** The Umegaki relative entropy (also called quantum relative entropy) between two quantum states $\rho, \sigma \in \mathfrak{D}(A)$ is defined by [43]

$$D(\rho\|\sigma) := \mathrm{Tr}[\rho(\log\rho - \log\sigma)], \tag{5}$$

if $\mathrm{supp}(\rho) \subseteq \mathrm{supp}(\sigma)$ and $+\infty$ otherwise.

Definition 2 (Petz Rényi divergence)**.** The Petz Rényi divergence of order α between two quantum states $\rho, \sigma \in \mathfrak{D}(A)$ is defined by [44]

$$\bar{D}_\alpha(\rho\|\sigma) := \frac{1}{\alpha - 1}\log\mathrm{Tr}\left[\rho^\alpha\sigma^{1-\alpha}\right], \tag{6}$$

if $\alpha \in (0, 1)$ or $\alpha \in (1, +\infty)$ with $\mathrm{supp}(\rho) \subseteq \mathrm{supp}(\sigma)$, and $+\infty$ otherwise.

Definition 3 (Sandwiched Rényi divergence)**.** The sandwiched Rényi divergence of order α between two quantum states $\rho, \sigma \in \mathfrak{D}(A)$ is defined by [45, 46]

$$\widetilde{D}_\alpha(\rho\|\sigma) := \frac{1}{\alpha - 1}\log\mathrm{Tr}\left[\sigma^{\frac{1-\alpha}{2\alpha}}\rho\sigma^{\frac{1-\alpha}{2\alpha}}\right]^\alpha, \tag{7}$$

if $\alpha \in (0, 1)$ or $\alpha \in (1, +\infty)$ with $\mathrm{supp}(\rho) \subseteq \mathrm{supp}(\sigma)$, and $+\infty$ otherwise.

Definition 4 (Max-relative entropy)**.** The max-relative entropy between two quantum states $\rho, \sigma \in \mathfrak{D}(A)$ is defined by [47, 48]

$$D_{\max}(\rho\|\sigma) := \log\inf\left\{t \in \mathbb{R} \ : \ \rho \leqslant t\sigma\right\}, \tag{8}$$

if $\mathrm{supp}(\rho) \subseteq \mathrm{supp}(\sigma)$ and $+\infty$ otherwise. Let $F(\rho, \sigma) := \|\sqrt{\rho}\sqrt{\sigma}\|_1 + \sqrt{(1 - \mathrm{Tr}\,\rho)(1 - \mathrm{Tr}\,\sigma)}$ be the generalized fidelity and $P(\rho, \sigma) := \sqrt{1 - F^2(\rho, \sigma)}$ be the purified distance. Let $\varepsilon \in (0, 1)$. Then the smoothed max-relative entropy is defined by

$$D^\varepsilon_{\max}(\rho\|\sigma) := \inf_{\rho' : P(\rho', \rho) \leqslant \varepsilon} D_{\max}(\rho'\|\sigma), \tag{9}$$

where the infimum is taken over all subnormalized states that are ε-close to the state ρ.

Definition 5 (Quantum hypothesis testing)**.** Let $\varepsilon \in [0, 1]$. The quantum hypothesis testing relative entropy between two quantum state $\rho, \sigma \in \mathfrak{D}(A)$ is defined by

$$D^\varepsilon_H(\rho\|\sigma) := -\log\inf\{\mathrm{Tr}[\Pi\sigma] : 0 \leqslant \Pi \leqslant I, \mathrm{Tr}[\Pi\rho] \geqslant 1 - \varepsilon\}. \tag{10}$$

The following result establishes an inequality relating the quantum hypothesis testing relative entropy and the sandwiched Rényi divergence [10, Lemma 5]. For any $\alpha \in (1, +\infty)$ and $\varepsilon \in (0, 1)$, it holds that

$$D^\varepsilon_H(\rho\|\sigma) \leqslant \widetilde{D}_\alpha(\rho\|\sigma) + \frac{\alpha}{\alpha - 1}\log\frac{1}{1 - \varepsilon}. \tag{11}$$

The quantum hypothesis testing relative entropy can also be lower bounded by the Petz Rényi divergence [49, Proposition 3]. For any $\alpha \in (0, 1)$ and $\varepsilon \in (0, 1)$, it holds that

$$D^\varepsilon_H(\rho\|\sigma) \geqslant \bar{D}_\alpha(\rho\|\sigma) - \frac{\alpha}{1 - \alpha}\log\frac{1}{\varepsilon}. \tag{12}$$

Here we provide a tighter lower bound with a simple proof, addressing the open question posed by Nuradha and Wilde in [33, Remark 4].

Lemma 1. Let $\varepsilon \in (0,1)$ and $\rho, \sigma \in \mathfrak{D}(A)$. For any $\alpha \in (0,1)$,

$$D_H^{\varepsilon}(\rho\|\sigma) \geqslant \bar{D}_\alpha(\rho\|\sigma) + \frac{\alpha}{1-\alpha}\left(\frac{h(\alpha)}{\alpha} - \log\left(\frac{1}{\varepsilon}\right)\right), \tag{13}$$

where $h(\alpha) = -\alpha \log \alpha - (1-\alpha)\log(1-\alpha)$ is the binary entropy.

Proof. Recall a variational expression of the hypothesis testing relative entropy [50, Eq. (2)]

$$2^{-D_H^{\varepsilon}(\rho\|\sigma)} = \max_{t \geqslant 0} \Big\{ t(1-\varepsilon) - \operatorname{Tr}(t\rho - \sigma)_+ \Big\}. \tag{14}$$

To bound the term $\operatorname{Tr}(t\rho - \sigma)_+$ we use the quantum weighted geometric-mean inequality; i.e., for any two positive semidefinite matrices M, N and any $\alpha \in [0,1]$

$$\frac{1}{2}\operatorname{Tr}\Big[M + N - \big|M - N\big|\Big] \leqslant \operatorname{Tr}\big[M^\alpha N^{1-\alpha}\big]. \tag{15}$$

Since the term $|M-N|$ can be expressed as $|M-N| = 2(M-N)_+ - (M-N)$, the above inequality is equivalent to

$$\operatorname{Tr}(M-N)_+ \geqslant \operatorname{Tr}[M] - \operatorname{Tr}\big[M^\alpha N^{1-\alpha}\big]. \tag{16}$$

Taking $M = t\rho$ and $N = \sigma$, we have

$$\operatorname{Tr}(t\rho - \sigma)_+ \geqslant t - t^\alpha \operatorname{Tr}\big[\rho^\alpha \sigma^{1-\alpha}\big] = t - t^\alpha 2^{(\alpha-1)\bar{D}_\alpha(\rho\|\sigma)}. \tag{17}$$

Substituting this into (14) gives

$$2^{-D_H^{\varepsilon}(\rho\|\sigma)} = \max_{t \geqslant 0} \Big\{ t(1-\varepsilon) - \operatorname{Tr}(t\rho - \sigma)_+ \Big\} \leqslant \max_{t \geqslant 0} \Big\{ -t\varepsilon + t^\alpha 2^{(\alpha-1)\bar{D}_\alpha(\rho\|\sigma)} \Big\}. \tag{18}$$

It is straightforward to check that for fixed $\alpha, \rho, \sigma, \varepsilon$, the function $t \mapsto -t\varepsilon + t^\alpha 2^{(\alpha-1)\bar{D}_\alpha(\rho\|\sigma)}$ obtains its maximal value at

$$t = \left(\frac{\alpha}{\varepsilon}\right)^{\frac{1}{1-\alpha}} 2^{-\bar{D}_\alpha(\rho\|\sigma)}. \tag{19}$$

Substituting this value into the optimization in (18) gives

$$2^{-D_H^{\varepsilon}(\rho\|\sigma)} \leqslant (1-\alpha)\left(\frac{\alpha}{\varepsilon}\right)^{\frac{\alpha}{1-\alpha}} 2^{-\bar{D}_\alpha(\rho\|\sigma)}. \tag{20}$$

By taking $-\log$ on both sides, we get (13) and conclude the proof.

2.3 Quantum channel divergences

The divergence between quantum states can be naturally extended to quantum channels. The key idea is to quantify the worst-case divergence among the outputs produced by these channels. Depending on the selection of input states, three distinct variants of quantum channel divergences arise, namely unstabilized, stabilized, and amortized divergences. It is noteworthy that channel divergences have been served as crucial tools in various fundamental areas, including the resource theory of quantum channels [13, 34–37, 51], quantum communication [24, 26–28, 38, 52], quantum coherence [39, 40], fault-tolerant quantum computing [41], and quantum thermodynamics [42]. We review their definitions here and provide several general properties, which will be used in the later discussions and can be of independent interests for future studies as well.

2.3.1 *Unstabilized quantum channel divergence*

Definition 6. Let $\boldsymbol{D}$ be a quantum state divergence. The unstabilized quantum channel divergence between two quantum channels $\mathcal{N}, \mathcal{M} \in \mathrm{CPTP}(A \to B)$ is defined by

$$\boldsymbol{d}(\mathcal{N}\|\mathcal{M}) := \sup_{\rho \in \mathfrak{D}(A)} \boldsymbol{D}(\mathcal{N}_{A\to B}(\rho_A)\|\mathcal{M}_{A\to B}(\rho_A)), \tag{21}$$

where the supremum is taken over all density operators ρ on system A.

The term "unstabilized" arises from the observation that the divergence value typically varies when appending an identity map, as expressed by the inequality

$$\boldsymbol{d}(\mathcal{N}\|\mathcal{M}) \neq \boldsymbol{d}(\mathcal{N}\otimes\mathcal{I}\|\mathcal{M}\otimes\mathcal{I}). \tag{22}$$

This distinguishes it from the conventional channel divergence [53, Definition II.2].

In the following, we use $d, \widetilde{d}, \bar{d}, d^{\varepsilon}_{\max}, d^{\varepsilon}_{H}$ to represent the unstabilized channel divergences induced by $D, \widetilde{D}, \bar{D}, D^{\varepsilon}_{\max}, D^{\varepsilon}_{H}$, respectively.

Many properties of state divergence can be extended to channel divergences. For instance, the following continuity property holds true.

Lemma 2. Let $\bar{d}_\alpha, \tilde{d}_\alpha$ and d be the unstabilized quantum channel divergences induced by the Petz Rényi divergence, the sandwiched Rényi divergence and the Umegaki relative entropy, respectively. Then for any $\mathcal{N},\mathcal{M}\in \mathrm{CPTP}(A\to B)$, it holds that

$$\lim_{\alpha\to 1}\bar{d}_\alpha(\mathcal{N}\|\mathcal{M}) = \lim_{\alpha\to 1}\tilde{d}_\alpha(\mathcal{N}\|\mathcal{M}) = d(\mathcal{N}\|\mathcal{M}). \tag{23}$$

Proof. The proof follows similarly as [10, Lemma 10].

A widely-studied unstabilized channel divergence is the min-output entropy [54]

$$h(\mathcal{N}) := \min_{\rho\in\mathfrak{D}(A)} H(\mathcal{N}(\rho)) = \log|B| - d(\mathcal{N}\|\mathcal{R}^{\pi}_{E}), \tag{24}$$

where $\mathcal{N}\in\mathrm{CPTP}(A\to B)$ and the maximally mixed state $\pi\in\mathfrak{D}(B)$. It is known that this quantity is not additive under the tensor product of quantum channels [54].

Given that an unstabilized quantum channel divergence is generally non-additive, it is natural to introduce its regularized counterpart.

Definition 7. Let $\boldsymbol{D}$ be a quantum state divergence. For any $\mathcal{N},\mathcal{M}\in\mathrm{CPTP}(A\to B)$, the regularized version of the unstabilized channel divergence is defined by

$$\boldsymbol{d}^{\mathrm{reg}}(\mathcal{N}\|\mathcal{M}) := \sup_{n\in\mathbb{N}}\frac{1}{n}\boldsymbol{d}(\mathcal{N}^{\otimes n}\|\mathcal{M}^{\otimes n}). \tag{25}$$

If the quantum state divergence $\boldsymbol{D}$ is superadditive under tensor product, i.e.,

$$\boldsymbol{D}(\rho_1\otimes\rho_2\|\sigma_1\otimes\sigma_2) \geqslant \boldsymbol{D}(\rho_1\|\sigma_1)+\boldsymbol{D}(\rho_2\|\sigma_2), \tag{26}$$

then it is easy to check that its unstabilized channel divergence is also superadditive, i.e.,

$$\boldsymbol{d}(\mathcal{N}_1\otimes\mathcal{N}_2\|\mathcal{M}_1\otimes\mathcal{M}_2) \geqslant \boldsymbol{d}(\mathcal{N}_1\|\mathcal{M}_1)+\boldsymbol{d}(\mathcal{N}_2\|\mathcal{M}_2). \tag{27}$$

Using a standard argument, we also have

$$\boldsymbol{d}^{\mathrm{reg}}(\mathcal{N}\|\mathcal{M}) = \lim_{n\to\infty}\frac{1}{n}\boldsymbol{d}(\mathcal{N}^{\otimes n}\|\mathcal{M}^{\otimes n}). \tag{28}$$

Later, as demonstrated in Theorem 12, we will see that the unstabilized channel divergence can exhibit an extremely non-additive behavior. In other words, an unbounded number of channel uses may be necessary to achieve its regularization.

2.3.2 *Stabilized quantum channel divergence*

The unstabilized quantum channel divergence exhibits deviation when an identity map is appended. To mitigate this, we can consider a stabilized version that allows the inclusion of an identity map.

Definition 8. Let $\boldsymbol{D}$ be a quantum state divergence. The (stabilized) quantum channel divergence between two quantum channels $\mathcal{N},\mathcal{M}\in\mathrm{CPTP}(A\to B)$ is defined by [53]

$$\boldsymbol{D}(\mathcal{N}\|\mathcal{M}) := \sup_{|R|\in\mathbb{N}}\boldsymbol{d}(\mathcal{I}_R\otimes\mathcal{N}\|\mathcal{I}_R\otimes\mathcal{M}), \tag{29}$$

where the supremum is taken over Hilbert space R of arbitrary dimension.

Remark 1. As a consequence of purification, data processing, and the Schmidt decomposition, the supremum can be constrained such that the reference system R is isomorphic to the channel input system A [53]. Thus, $\boldsymbol{D}(\mathcal{N}\|\mathcal{M}) = \boldsymbol{d}(\mathcal{I}_R \otimes \mathcal{N}\|\mathcal{I}_R \otimes \mathcal{M})$, where R is isomorphic to A.

Similar to the unstabilized channel divergence, the stabilized version is non-additive [24] in general. This observation motivates the introduction of their regularization.

Definition 9. Let $\boldsymbol{D}$ be a quantum state divergence. For any $\mathcal{N}, \mathcal{M} \in \mathrm{CPTP}(A \to B)$, the regularized version of the stabilized channel divergence is defined by

$$\boldsymbol{D}^{\mathrm{reg}}(\mathcal{N}\|\mathcal{M}) := \sup_{n\in\mathbb{N}} \frac{1}{n} \boldsymbol{D}\left(\mathcal{N}^{\otimes n} \middle\| \mathcal{M}^{\otimes n}\right). \tag{30}$$

2.3.3 *Amortized quantum channel divergence*

Both the unstabilized and stabilized channel divergences assess the distinguishability of channel outputs using the same input state. Alternatively, a method for inducing channel divergence is amortization, which uses different input states.

Definition 10. Let $\boldsymbol{D}$ be a quantum state divergence. The amortized quantum channel divergence between two quantum channels $\mathcal{N}, \mathcal{M} \in \mathrm{CPTP}(A \to B)$ is defined by [38]

$$\boldsymbol{D}^{A}(\mathcal{N}\|\mathcal{M}) := \sup_{\rho,\sigma\in\mathfrak{D}(RA)} \Big[\boldsymbol{D}\left(\mathcal{I}_R \otimes \mathcal{N}(\rho_{RA}) \| \mathcal{I}_R \otimes \mathcal{M}(\sigma_{RA})\right) - \boldsymbol{D}\left(\rho_{RA}\|\sigma_{RA}\right)\Big], \tag{31}$$

where the supremum is taken over all quantum states $\rho, \sigma \in \mathfrak{D}(RA)$ and R is of arbitrary dimension.

As previously mentioned, both unstabilized and stabilized channel divergences are generally non-additive. In contrast, the amortized channel divergence can inherit the additivity property from the corresponding state divergence.

Lemma 3. Let $\boldsymbol{D}$ be a quantum state divergence. Let $\mathcal{N}_1, \mathcal{M}_1 \in \mathrm{CPTP}(A_1 \to B_1)$ and $\mathcal{N}_2, \mathcal{M}_2 \in \mathrm{CPTP}(A_2 \to B_2)$. If $\boldsymbol{D}$ is additive under the tensor product of quantum states, then

$$\boldsymbol{D}^{A}(\mathcal{N}_1 \otimes \mathcal{N}_2\|\mathcal{M}_1 \otimes \mathcal{M}_2) = \boldsymbol{D}^{A}(\mathcal{N}_1\|\mathcal{M}_1) + \boldsymbol{D}^{A}(\mathcal{M}_1\|\mathcal{M}_2). \tag{32}$$

Proof. For any quantum state $\rho, \sigma \in \mathfrak{D}(RA_1A_2)$, it holds that

$$\boldsymbol{D}(\mathcal{N}_1 \otimes \mathcal{N}_2(\rho)\|\mathcal{M}_1 \otimes \mathcal{M}_2(\sigma)) \leqslant \boldsymbol{D}^{A}(\mathcal{N}_1\|\mathcal{M}_1) + \boldsymbol{D}(\mathcal{N}_2(\rho)\|\mathcal{M}_2(\sigma)) \tag{33}$$

$$\leqslant \boldsymbol{D}^{A}(\mathcal{N}_1\|\mathcal{M}_1) + \boldsymbol{D}^{A}(\mathcal{N}_2\|\mathcal{M}_2) + \boldsymbol{D}(\rho\|\sigma), \tag{34}$$

where the two inequalities follow by using the definition of the amortized channel divergence twice. Then moving the term $\boldsymbol{D}(\rho\|\sigma)$ to the l.h.s. and taking supremum over all input states ρ, σ, we have one direction of the stated result. On the other hand, for any input states $\rho_1, \sigma_1 \in \mathfrak{D}(R_1A_1)$ and $\rho_2, \sigma_2 \in \mathfrak{D}(R_2A_2)$, we have

$$\boldsymbol{D}^{A}(\mathcal{N}_1 \otimes \mathcal{N}_2\|\mathcal{M}_1 \otimes \mathcal{M}_2) \tag{35}$$

$$\geqslant \sup_{\rho_1,\rho_2,\sigma_1,\sigma_2} \Big[\boldsymbol{D}(\mathcal{N}_1 \otimes \mathcal{N}_2(\rho_1 \otimes \rho_2)\|\mathcal{M}_1 \otimes \mathcal{M}_2(\sigma_1 \otimes \sigma_2)) - \boldsymbol{D}(\rho_1 \otimes \rho_2\|\sigma_1 \otimes \sigma_2)\Big] \tag{36}$$

$$= \sup_{\rho_1,\rho_2,\sigma_1,\sigma_2} \Big[\boldsymbol{D}(\mathcal{N}_1(\rho_1)\|\mathcal{M}_1(\sigma_1)) - \boldsymbol{D}(\rho_1\|\sigma_1)\Big] + \Big[\boldsymbol{D}(\mathcal{N}_2(\rho_2)\|\mathcal{M}_2(\sigma_2)) - \boldsymbol{D}(\rho_2\|\sigma_2)\Big] \tag{37}$$

$$= \sup_{\rho_1,\sigma_1} \Big[\boldsymbol{D}(\mathcal{N}_1(\rho_1)\|\mathcal{M}_1(\sigma_1)) - \boldsymbol{D}(\rho_1\|\sigma_1)\Big] + \sup_{\rho_2,\sigma_2} \Big[\boldsymbol{D}(\mathcal{N}_2(\rho_2)\|\mathcal{M}_2(\sigma_2)) - \boldsymbol{D}(\rho_2\|\sigma_2)\Big] \tag{38}$$

$$= \boldsymbol{D}^{A}(\mathcal{N}_1\|\mathcal{M}_1) + \boldsymbol{D}^{A}(\mathcal{N}_2\|\mathcal{M}_2), \tag{39}$$

where the inequality follows as tensor product states are particular choices of input states for $\boldsymbol{D}^{A}(\mathcal{N}_1 \otimes \mathcal{N}_2\|\mathcal{M}_1 \otimes \mathcal{M}_2)$, the first equality follows by the additivity assumption of $\boldsymbol{D}$. This concludes the proof.

By the chain rules of Umegaki relative entropy [24, Corollary 3] and the sandwiched Rényi divergence [55, Theorem 5.4], it follows that $D^{\mathrm{reg}}(\mathcal{N}\|\mathcal{M}) = D^{A}(\mathcal{N}\|\mathcal{M})$ and $\widetilde{D}_{\alpha}^{\mathrm{reg}}(\mathcal{N}\|\mathcal{M}) = \widetilde{D}_{\alpha}^{A}(\mathcal{N}\|\mathcal{M})$ for any quantum channels $\mathcal{N}, \mathcal{M} \in \mathrm{CPTP}(A \to B)$ and $\alpha > 1$. Consequently, from Lemma 3, we can infer that D^{reg} and $\widetilde{D}_{\alpha}^{\mathrm{reg}}$ are also additive under the tensor product of distinct quantum channels. Establishing this directly from their definitions can be challenging.

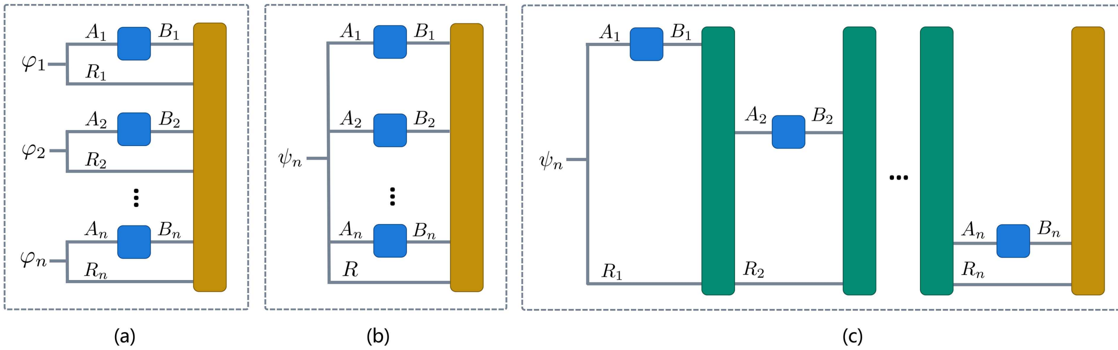

Figure 1 (Color online) Illustration depicting different classes of strategies for quantum channel discrimination. Each blue box represents an unknown quantum channel $\mathcal{G} \in \{\mathcal{N}, \mathcal{M}\}$ to discriminate, each yellow box represents a quantum measurement $\{\Pi_n, I - \Pi_n\}$, and each green box represents an update channel $\mathcal{P}_i$. (a) Product strategy; (b) coherent strategy; (c) sequential strategy.

Lemma 4. Let $\mathcal{N}_1, \mathcal{M}_1 \in \mathrm{CPTP}(A_1 \to B_1)$ and $\mathcal{N}_2, \mathcal{M}_2 \in \mathrm{CPTP}(A_2 \to B_2)$. For any $\alpha \in (1, +\infty)$, the following additivity properties hold

$$D^{\mathrm{reg}}(\mathcal{N}_1 \otimes \mathcal{N}_2 \| \mathcal{M}_1 \otimes \mathcal{M}_2) = D^{\mathrm{reg}}(\mathcal{N}_1 \| \mathcal{M}_1) + D^{\mathrm{reg}}(\mathcal{N}_2 \| \mathcal{M}_2), \tag{40}$$

$$\widetilde{D}_\alpha^{\mathrm{reg}}(\mathcal{N}_1 \otimes \mathcal{N}_2 \| \mathcal{M}_1 \otimes \mathcal{M}_2) = \widetilde{D}_\alpha^{\mathrm{reg}}(\mathcal{N}_1 \| \mathcal{M}_1) + \widetilde{D}_\alpha^{\mathrm{reg}}(\mathcal{N}_2 \| \mathcal{M}_2). \tag{41}$$

The next result establishes the chain relation among different variants of channel divergences.

Lemma 5. Let $\boldsymbol{D}$ be a quantum state divergence that is superadditive under the tensor product of quantum states. Then for any $\mathcal{N}, \mathcal{M} \in \mathrm{CPTP}(A \to B)$, it holds that

$$\boldsymbol{d}(\mathcal{N} \| \mathcal{M}) \leqslant \boldsymbol{D}(\mathcal{N} \| \mathcal{M}) \leqslant \boldsymbol{D}^{\mathrm{reg}}(\mathcal{N} \| \mathcal{M}) \leqslant \boldsymbol{D}^A(\mathcal{N} \| \mathcal{M}). \tag{42}$$

Proof. The first two inequalities follow from their definitions. We also have that

$$\frac{1}{n}\boldsymbol{D}(\mathcal{N}^{\otimes n} \| \mathcal{M}^{\otimes n}) \leqslant \frac{1}{n}\boldsymbol{D}^A(\mathcal{N}^{\otimes n} \| \mathcal{M}^{\otimes n}) \leqslant \boldsymbol{D}^A(\mathcal{N} \| \mathcal{M}), \tag{43}$$

where the first inequality follows by definition and the second inequality follows from (34). Taking the supremum over all integers n, we have $\boldsymbol{D}^{\mathrm{reg}}(\mathcal{N} \| \mathcal{M}) \leqslant \boldsymbol{D}^A(\mathcal{N} \| \mathcal{M})$.

2.4 Quantum channel discrimination

The task of channel discrimination closely parallels that of state discrimination. In the case of an unknown quantum channel $\mathcal{G}$, the goal is to identify it among potential candidates $\mathcal{N}$ or $\mathcal{M}$. A standard approach to discrimination involves hypothesis testing to distinguish between the null hypothesis $\mathcal{G} = \mathcal{N}$ and the alternative hypothesis $\mathcal{G} = \mathcal{M}$. What distinguishes channel discrimination is the varied selection of discrimination strategies and whether the utilization of quantum memories is permitted.

Different classes of available strategies are illustrated in Figure 1. Each strategy class comprises two components, denoted as (S_n, Π_n), where S_n is a method for generating a testing state, and Π_n $(0 \leqslant \Pi_n \leqslant I)$ defines a quantum test, a binary quantum measurement $\{\Pi_n, I - \Pi_n\}$ performed on this state. For a given strategy (S_n, Π_n), let $\rho_n(S_n)$ and $\sigma_n(S_n)$ be the testing states generated by n uses of the channel, depending on whether it is $\mathcal{N}$ or $\mathcal{M}$. Then the Type-I and Type-II errors are defined as

$$(\text{Type-I}) \quad \alpha_n(S_n, \Pi_n) := \mathrm{Tr}[(I - \Pi_n)\rho_n(S_n)], \tag{44}$$

$$(\text{Type-II}) \quad \beta_n(S_n, \Pi_n) := \mathrm{Tr}[\Pi_n \sigma_n(S_n)], \tag{45}$$

respectively. As perfect discrimination (i.e., simultaneous elimination of both errors) is not always possible, the focus shifts to the asymptotic behavior of α_n and β_n for sufficiently large n, expecting a tradeoff between minimizing α_n and minimizing β_n.

Product strategy. Let R_i be the ancillary quantum system of a quantum memory for the i-th use of the quantum channel. In a product strategy (Figure 1(a)), the testing state is created by selecting a sequence of input states $\varphi_i \in \mathfrak{D}(R_i A_i)$ and sending the A_i system to the unknown channel $\mathcal{G}$ individually.

The generated testing state is then given by $\mathcal{G}^{\otimes n}(\bigotimes_{i=1}^n \varphi_i)$. The class of all product strategies is denoted as PRO. It is important to note that the input states considered here are not restricted to having an i.i.d. structure (e.g., $\varphi^{\otimes n}$) but rather general tensor product states. In other words, we allow the choice of different input states for different instances of $\mathcal{G}$, distinguishing it from the product strategy discussed in [10]. If the dimension of the ancillary quantum system reduces to 1, it corresponds to product strategies without quantum memory assistance.

Coherent strategy. Let R be the ancillary quantum system of a quantum memory. In a coherent strategy (Figure 1(b)), the testing state is created by choosing an input state $\psi_n \in \mathfrak{D}(RA^n)$ and sending the corresponding A_i system to each copy of the channel. The generated testing state is then given by $\mathcal{G}^{\otimes n}(\psi_n)$. The class of all coherent strategies is denoted as COH. It is evident that if our choice of ψ_n has a tensor product structure $\bigotimes_{i=1}^n \varphi_i$ with $\varphi_i \in \mathfrak{D}(R_iA_i)$, we effectively obtain a product strategy. Thus, we have the set inclusion PRO $\subseteq$ COH. If the dimension of the reference systems reduces to 1, it corresponds to coherent strategies without quantum memory assistance.

Sequential strategy. Let R_i be the ancillary quantum system of a quantum memory for the i-th use of the quantum channel. In a sequential strategy (Figure 1(c)), the testing state is created adaptively. Initially, we choose an initial state $\psi_n \in \mathfrak{D}(R_1A_1)$ and send it through one copy of the channel $\mathcal{G}$ followed by the application of an update channel $\mathcal{P}_1$. Subsequently, another copy of the channel $\mathcal{G}$ is applied, followed by an update channel $\mathcal{P}_2$. This process is repeated n times, resulting in the final testing state $\mathcal{G} \circ \mathcal{P}_{n-1} \circ \cdots \circ \mathcal{P}_2 \circ \mathcal{G} \circ \mathcal{P}_1 \circ \mathcal{G}(\psi_n)$, where $\mathcal{P}_i \in \mathrm{CPTP}(R_iB_i \to R_{i+1}A_{i+1})$. The class of all sequential strategies is denoted as SEQ. It is evident that if all update channels $\mathcal{P}_i$ are chosen as identity maps, the sequential strategy reduces to a coherent strategy. Thus, we have COH $\subseteq$ SEQ. If the dimension of the ancillary quantum system reduces to 1, it corresponds to sequential strategies without quantum memory assistance.

3 Limits of quantum channel divergence

In this section, we investigate the limits of the unstabilized quantum channel divergences and prove a quantum channel analog of Stein's lemma without quantum memory assistance. To further strengthen the result, we introduce the (exponentially) strong converse properties for channel discrimination and establish its equivalence to the AEP of various quantum channel divergences as well as the continuity of the quantum channel Rényi divergence. Leveraging these equivalent characterizations, we demonstrate, rather surprisingly, that the exponentially strong converse properties under coherent and sequential strategies are equivalent.

Given the widespread applications of quantum Stein's lemma, our channel Stein's lemma is anticipated to have significant implications once its strong converse version is completely solved. Our results contribute to distinct perspectives towards establishing such a result and can serve as building blocks for its applications. This includes facilitating a deeper understanding of the tasks of quantum channel discrimination and quantum communication in subsequent sections.

3.1 A quantum channel Stein's lemma without memory assistance

The following result establishes an analog of the Stein's lemma for quantum channels.

Proposition 1. For any two quantum channels $\mathcal{N}, \mathcal{M} \in \mathrm{CPTP}(A \to B)$, it holds that

$$\lim_{\varepsilon \to 0} \lim_{n \to \infty} \frac{1}{n} d_H^{\varepsilon}(\mathcal{N}^{\otimes n} \| \mathcal{M}^{\otimes n}) = d^{\mathrm{reg}}(\mathcal{N} \| \mathcal{M}). \tag{46}$$

Proof. Recall that for any $\rho, \sigma \in \mathfrak{D}(A)$ and $\varepsilon \in [0, 1)$, it holds that

$$D_H^{\varepsilon}(\rho \| \sigma) \leqslant \frac{1}{1-\varepsilon}[D(\rho \| \sigma) + h_2(\varepsilon)], \tag{47}$$

where $h_2(\cdot)$ is the binary entropy (see e.g., [17]). Applying this to $\mathcal{N}^{\otimes n}(\rho_n)$ and $\mathcal{M}^{\otimes n}(\rho_n)$ and taking supremum over all input states $\rho_n \in \mathfrak{D}(A^n)$, we have

$$d_H^{\varepsilon}(\mathcal{N}^{\otimes n} \| \mathcal{M}^{\otimes n}) \leqslant \frac{1}{1-\varepsilon}\big[d(\mathcal{N}^{\otimes n} \| \mathcal{M}^{\otimes n}) + h_2(\varepsilon)\big]. \tag{48}$$

Taking limits on both sides, we have

$$\lim_{\varepsilon\to 0}\lim_{n\to\infty}\frac{1}{n}d_H^{\varepsilon}(\mathcal{N}^{\otimes n}\|\mathcal{M}^{\otimes n})\leqslant d^{\mathrm{reg}}(\mathcal{N}\|\mathcal{M}). \tag{49}$$

For the other direction, suppose the optimal solution for $d(\mathcal{N}\|\mathcal{M})$ is taken at ρ_A. Then we have

$$\lim_{n\to\infty}\frac{1}{n}d_H^{\varepsilon}(\mathcal{N}^{\otimes n}\|\mathcal{M}^{\otimes n})\geqslant\lim_{n\to\infty}\frac{1}{n}d_H^{\varepsilon}(\mathcal{N}(\rho)^{\otimes n}\|\mathcal{M}(\rho)^{\otimes n})=d(\mathcal{N}(\rho)\|\mathcal{M}(\rho))=d(\mathcal{N}\|\mathcal{M}), \tag{50}$$

where the first inequality follows as $\rho^{\otimes n}$ is a particular choice for the unstabilized divergence, the first equality follows by the quantum Stein's Lemma, and the second equality follows by the optimality assumption of ρ. Then for any fixed m, by replacing $\mathcal{N}$ with $\mathcal{N}^{\otimes m}$ and $\mathcal{M}$ with $\mathcal{M}^{\otimes m}$, we have

$$\lim_{n\to\infty}\frac{1}{mn}d_H^{\varepsilon}(\mathcal{N}^{\otimes mn}\|\mathcal{M}^{\otimes mn})\geqslant\frac{1}{m}d(\mathcal{N}^{\otimes m}\|\mathcal{M}^{\otimes m}). \tag{51}$$

Finally taking $m\to\infty$ and then $\varepsilon\to 0$, we have the achievable part and conclude the proof.

3.2 Towards a strong converse version

Similar to the strong converse property of quantum state discrimination, an analog property can also be defined for quantum channels.

Definition 11 (Strong converse property)**.** Let $\mathcal{N},\mathcal{M}\in\mathrm{CPTP}(A\to B)$ be two quantum channels. These channels exhibit the strong converse property for coherent channel discrimination strategies without quantum memory assistance if, for any sequence of strategies where the Type-II errors β_n satisfy

$$\liminf_{n\to\infty}-\frac{1}{n}\log\beta_n=:r>d^{\mathrm{reg}}(\mathcal{N}\|\mathcal{M}), \tag{52}$$

there necessarily exists a subsequence of Type-I errors α_{n_k} that converges to 1 as $n_k\to\infty$.

If the strong converse property holds, the Type-I error will typically converge to one exponentially fast. Therefore, we introduce a stronger version by requiring exponential convergence and term this condition as an exponentially strong converse property.

Definition 12 (Exponentially strong converse property)**.** Let $\mathcal{N},\mathcal{M}\in\mathrm{CPTP}(A\to B)$ be two quantum channels. These channels exhibit the strong converse property for coherent channel discrimination strategies without quantum memory assistance if, for any sequence of strategies where the Type-II errors β_n satisfy

$$\liminf_{n\to\infty}-\frac{1}{n}\log\beta_n=:r>d^{\mathrm{reg}}(\mathcal{N}\|\mathcal{M}), \tag{53}$$

there necessarily exists a subsequence of Type-I errors α_{n_k} such that $1-\alpha_{n_k}\leqslant 2^{-cn_k}$ for a constant $c>0$ and for sufficiently large n_k.

The strong converse properties require the study of all suitable discrimination strategies, which can be hard to validate in general. In the following, we provide several equivalent characterizations related to the limits of unstabilized channel divergences.

From the proof of Proposition 1, we actually have a stronger statement that

$$\lim_{n\to\infty}\frac{1}{n}d_H^{\varepsilon}\left(\mathcal{N}^{\otimes n}\middle\|\mathcal{M}^{\otimes n}\right)\geqslant d^{\mathrm{reg}}(\mathcal{N}\|\mathcal{M}),\quad\forall\varepsilon\in(0,1). \tag{54}$$

The following result shows that the other direction is equivalent to the strong converse property in Definition 11.

Theorem 1. Let $\mathcal{N},\mathcal{M}\in\mathrm{CPTP}(A\to B)$ be two quantum channels. Then these channels exhibit the strong converse property as defined in Definition 11 if and only if the following holds:

$$\limsup_{n\to\infty}\frac{1}{n}d_H^{\varepsilon}\left(\mathcal{N}^{\otimes n}\middle\|\mathcal{M}^{\otimes n}\right)\leqslant d^{\mathrm{reg}}(\mathcal{N}\|\mathcal{M}),\quad\forall\varepsilon\in(0,1). \tag{55}$$

Proof. Suppose the strong converse property as defined in Definition 11 holds and assume that

$$\limsup_{n\to\infty} \frac{1}{n} d_H^\varepsilon \left(\mathcal{N}^{\otimes n} \middle\| \mathcal{M}^{\otimes n}\right) > d^{\text{reg}}(\mathcal{N}\|\mathcal{M}), \tag{56}$$

then there exists a subsequence n_k such that $\lim_{n_k\to\infty} \frac{1}{n_k} d_H^\varepsilon \left(\mathcal{N}^{\otimes n_k} \middle\| \mathcal{M}^{\otimes n_k}\right) > d^{\text{reg}}(\mathcal{N}\|\mathcal{M})$. This implies a sequence of strategies such that the Type-I error $\alpha_{n_k} \leqslant \varepsilon$ and the Type-II error $\lim_{n_k\to\infty} -\frac{1}{n_k}\log\beta_{n_k} > d^{\text{reg}}(\mathcal{N}\|\mathcal{M})$. By Definition 11, we know that the second condition implies a subsequence of α_{n_k} converges to 1, which contradicts to the first condition $\alpha_{n_k} \leqslant \varepsilon$. So Eq. (55) holds. On the other hand, we prove that Eq. (55) implies Definition 11. For any strategies such that $\liminf_{n\to\infty} -\frac{1}{n}\log\beta_n > d^{\text{reg}}(\mathcal{N}\|\mathcal{M})$. We now show that there exists a subsequence of α_n converges to 1. Assume there exists $0<\varepsilon<1$ such that $\alpha_n \leqslant \varepsilon$. By the definition of d_H^ε, we have $-\frac{1}{n}\log\beta_n \leqslant d_H^\varepsilon(\mathcal{N}^{\otimes n}\|\mathcal{M}^{\otimes n})$. This implies

$$\liminf_{n\to\infty} -\frac{1}{n}\log\beta_n \leqslant \limsup_{n\to\infty} \frac{1}{n} d_H^\varepsilon \left(\mathcal{N}^{\otimes n} \middle\| \mathcal{M}^{\otimes n}\right) \leqslant d^{\text{reg}}(\mathcal{N}\|\mathcal{M}), \tag{57}$$

which forms a contradiction to the assumption that $\liminf_{n\to\infty} -\frac{1}{n}\log\beta_n > d^{\text{reg}}(\mathcal{N}\|\mathcal{M})$.

The following shows that the AEP of max-relative entropy is also equivalent to Definition 11.

Theorem 2. Let $\mathcal{N},\mathcal{M} \in \text{CPTP}(A\to B)$ be two quantum channels. Then these channels exhibit the strong converse property as defined in Definition 11 if and only if the following holds:

$$\limsup_{n\to\infty} \frac{1}{n} d_{\max}^\varepsilon \left(\mathcal{N}^{\otimes n} \middle\| \mathcal{M}^{\otimes n}\right) \leqslant d^{\text{reg}}(\mathcal{N}\|\mathcal{M}), \quad \forall \varepsilon \in (0,1). \tag{58}$$

This is also equivalent to the following:

$$\lim_{\varepsilon\to 0}\limsup_{n\to\infty} \frac{1}{n} d_{\max}^\varepsilon \left(\mathcal{N}^{\otimes n} \middle\| \mathcal{M}^{\otimes n}\right) \leqslant d^{\text{reg}}(\mathcal{N}\|\mathcal{M}). \tag{59}$$

Proof. By Theorem 1, we only need to prove that Eqs. (55), (58) and (59) are equivalent.

(i) (55) $\Longrightarrow$ (58): for any two quantum states $\rho,\sigma \in \mathfrak{D}(A)$, any $\varepsilon \in (0,1)$, it is known that [56, Proposition 4.1],

$$D_{\max}^\varepsilon(\rho\|\sigma) \leqslant D_H^{1-\frac{1}{2}\varepsilon^2}(\rho\|\sigma) + \log\left(\frac{2}{\varepsilon^2}\right). \tag{60}$$

Applying this to channel divergence gives

$$d_{\max}^\varepsilon(\mathcal{N}\|\mathcal{M}) \leqslant d_H^{1-\frac{1}{2}\varepsilon^2}(\mathcal{N}\|\mathcal{M}) + \log\left(\frac{2}{\varepsilon^2}\right). \tag{61}$$

Taking n copies of $\mathcal{N}$ and $\mathcal{M}$, we get

$$\frac{1}{n} d_{\max}^\varepsilon \left(\mathcal{N}^{\otimes n} \middle\| \mathcal{M}^{\otimes n}\right) \leqslant \frac{1}{n} d_H^{1-\frac{1}{2}\varepsilon^2} \left(\mathcal{N}^{\otimes n} \middle\| \mathcal{M}^{\otimes n}\right) + \frac{1}{n}\log\left(\frac{2}{\varepsilon^2}\right). \tag{62}$$

Taking $\limsup_{n\to\infty}$ on both sides, we can see that Eq. (55) implies (58).

(ii) (58) $\Longrightarrow$ (59): trivial.

(iii) (59) $\Longrightarrow$ (55): for any two quantum states $\rho,\sigma \in \mathfrak{D}(A)$, any $\varepsilon \in (0,1)$, and any $\varepsilon' \in (0,1-\varepsilon)$, it is known that [57, Theorem 11],

$$D_H^{\varepsilon'}(\rho\|\sigma) + \log(1-\varepsilon-\varepsilon') \leqslant D_{\max}^\varepsilon(\rho\|\sigma). \tag{63}$$

Applying this to channel divergence gives

$$d_H^{\varepsilon'}(\mathcal{N}\|\mathcal{M}) + \log(1-\varepsilon-\varepsilon') \leqslant d_{\max}^\varepsilon(\mathcal{N}\|\mathcal{M}). \tag{64}$$

Taking n copies of $\mathcal{N}$ and $\mathcal{M}$, we get

$$\frac{1}{n} d_H^{\varepsilon'} \left(\mathcal{N}^{\otimes n} \middle\| \mathcal{M}^{\otimes n}\right) + \frac{1}{n}\log(1-\varepsilon-\varepsilon') \leqslant \frac{1}{n} d_{\max}^\varepsilon \left(\mathcal{N}^{\otimes n} \middle\| \mathcal{M}^{\otimes n}\right). \tag{65}$$

Taking $\limsup_{n\to\infty}$ and $\lim_{\varepsilon\to 0}$, we get

$$\frac{1}{n} d_H^{\varepsilon'} \left(\mathcal{N}^{\otimes n} \middle\| \mathcal{M}^{\otimes n}\right) \leqslant d^{\text{reg}}(\mathcal{N}\|\mathcal{M}), \tag{66}$$

which implies (55).

It is interesting to see that Eqs. (58) and (59) are actually equivalent, despite the latter appearing much weaker than the former. As D_H and $D_{\max}$ are the two extreme cases of one-shot quantum divergences, the above result would also apply to other intermediate divergences such as the information spectrum relative entropies [58, 59] and the recently introduced smoothed sandwiched Rényi divergence [60].

Besides the above AEPs, the strong converse properties also relate to the continuity of the regularized (amortized) sandwiched Rényi channel divergence at $\alpha = 1$.

Theorem 3. Let $\mathcal{N}, \mathcal{M} \in \mathrm{CPTP}(A \to B)$ be two quantum channels. Then the following continuitity:

$$\lim_{\alpha \to 1^+} \widetilde{d}_\alpha^{\mathrm{reg}}(\mathcal{N}\|\mathcal{M}) = d^{\mathrm{reg}}(\mathcal{N}\|\mathcal{M}) \tag{67}$$

implies that these channels exhibit the exponentially strong converse property as defined in Definition 12. Conversely, if the exponentially strong converse property as defined in Definition 12 holds true for channels $\mathcal{I} \otimes \mathcal{N}$ and $\mathcal{I} \otimes \mathcal{M}$ with the identity channel $\mathcal{I} \in \mathrm{CPTP}(A \to A)$, then

$$\lim_{\alpha \to 1^+} \widetilde{d}_\alpha^{\mathrm{reg}}(\mathcal{I} \otimes \mathcal{N}\|\mathcal{I} \otimes \mathcal{M}) = d^{\mathrm{reg}}(\mathcal{I} \otimes \mathcal{N}\|\mathcal{I} \otimes \mathcal{M}). \tag{68}$$

Proof. Note that by the monotonicity of sandwiched Rényi divergence with respect to α, the limits in the above statement can be replaced with $\inf_{\alpha>1}$. Suppose the continuity in (67) holds. Recall that [10, Lemma 5] for any $\alpha \in (1, +\infty)$ and $\varepsilon \in (0, 1)$, it holds that

$$D_H^\varepsilon(\rho\|\sigma) \leqslant \widetilde{D}_\alpha(\rho\|\sigma) + \frac{\alpha}{\alpha - 1} \log \frac{1}{1-\varepsilon}. \tag{69}$$

Applying this to the discrimination of n copies of the channels, it implies

$$-\frac{1}{n} \log \beta_n \leqslant \frac{1}{n} \widetilde{d}_\alpha(\mathcal{N}^{\otimes n}\|\mathcal{M}^{\otimes n}) + \frac{1}{n} \frac{\alpha}{\alpha - 1} \log \frac{1}{1-\alpha_n}. \tag{70}$$

If $\liminf_{n\to\infty} -\frac{1}{n} \log \beta_n := r > d^{\mathrm{reg}}(\mathcal{N}\|\mathcal{M})$, then there exists a subsequence n_k and $\delta > 0$ such that $-\frac{1}{n_k} \log \beta_{n_k} > r - \delta > d^{\mathrm{reg}}(\mathcal{N}\|\mathcal{M})$. Let $r' := r - \delta$. We have

$$r' < \frac{1}{n_k} \widetilde{d}_\alpha(\mathcal{N}^{\otimes n_k}\|\mathcal{M}^{\otimes n_k}) + \frac{1}{n_k} \frac{\alpha}{\alpha - 1} \log \frac{1}{1-\alpha_{n_k}}. \tag{71}$$

Since $\frac{1}{n_k} \widetilde{d}_\alpha(\mathcal{N}^{\otimes n_k}\|\mathcal{M}^{\otimes n_k}) \leqslant \widetilde{d}_\alpha^{\mathrm{reg}}(\mathcal{N}\|\mathcal{M})$, we have

$$r' < \widetilde{d}_\alpha^{\mathrm{reg}}(\mathcal{N}\|\mathcal{M}) + \frac{1}{n_k} \frac{\alpha}{\alpha - 1} \log \frac{1}{1-\alpha_{n_k}}, \tag{72}$$

which is equivalent to

$$1 - \alpha_{n_k} < 2^{-\frac{\alpha-1}{\alpha} n_k \left(r' - \widetilde{d}_\alpha^{\mathrm{reg}}(\mathcal{N}\|\mathcal{M})\right)}. \tag{73}$$

Since $r' > d^{\mathrm{reg}}(\mathcal{N}\|\mathcal{M}) = \inf_{\alpha>1} \widetilde{d}_\alpha^{\mathrm{reg}}(\mathcal{N}\|\mathcal{M})$ by assumption, there exists $\alpha > 1$ such that $r' > \widetilde{d}_\alpha^{\mathrm{reg}}(\mathcal{N}\|\mathcal{M})$. We can choose $c := (\alpha - 1)/\alpha(r' - \widetilde{d}_\alpha^{\mathrm{reg}}(\mathcal{N}\|\mathcal{M}))$. This implies the exponentially strong converse property in Definition 12.

We now prove the second statement. Suppose the exponentially strong converse property in Definition 12 holds true for $\mathcal{I} \otimes \mathcal{N}$ and $\mathcal{I} \otimes \mathcal{M}$. For any $\alpha > 1$, we have $\widetilde{D}_\alpha^{\mathrm{reg}}(\mathcal{N}\|\mathcal{M}) \geqslant D^{\mathrm{reg}}(\mathcal{N}\|\mathcal{M})$. Thus it is clear that $\inf_{\alpha>1} \widetilde{D}_\alpha^{\mathrm{reg}}(\mathcal{N}\|\mathcal{M}) \geqslant D^{\mathrm{reg}}(\mathcal{N}\|\mathcal{M})$. We now prove the other direction. If $\inf_{\alpha>1} \widetilde{D}_\alpha^{\mathrm{reg}}(\mathcal{N}\|\mathcal{M}) > D^{\mathrm{reg}}(\mathcal{N}\|\mathcal{M})$, we can find $r \in \mathbb{R}$ such that $\inf_{\alpha>1} \widetilde{D}_\alpha^{\mathrm{reg}}(\mathcal{N}\|\mathcal{M}) > r > D^{\mathrm{reg}}(\mathcal{N}\|\mathcal{M})$. Consider a sequence of coherent channel discrimination strategies such that the Type-II error converges at an exponential rate r. By the result [55, Theorem 5.5 and Remark 5.6], we know that the strong converse exponent is zero since $r < \inf_{\alpha>1} \widetilde{D}_\alpha^{\mathrm{reg}}(\mathcal{N}\|\mathcal{M})$, which means the Type-I error does not exponentially converge to one. However, by Definition 12, the condition $r > D^{\mathrm{reg}}(\mathcal{N}\|\mathcal{M})$ implies that the Type-I error has to converge exponentially to one, which forms a contradiction and concludes that $\inf_{\alpha>1} \widetilde{D}_\alpha^{\mathrm{reg}}(\mathcal{N}\|\mathcal{M}) \leqslant D^{\mathrm{reg}}(\mathcal{N}\|\mathcal{M})$. This proves (68).

Note that the second statement above holds for the stabilized channel divergence, as its proof relies on the results in [55, Theorem 5.5 and Remark 5.6]. It would be interesting to determine whether this result also holds for the unstabilized channel divergence in general.

Corollary 1. Let $\mathcal{N}, \mathcal{M} \in \mathrm{CPTP}(A \to B)$ be two quantum channels. The exponentially strong converse property as defined in Definition 12 holds true for channels $\mathcal{I} \otimes \mathcal{N}$ and $\mathcal{I} \otimes \mathcal{M}$ with the identity channel $\mathcal{I} \in \mathrm{CPTP}(A \to A)$ if and only if one of the following continuities holds

$$\lim_{\alpha \to 1^+} \widetilde{D}_\alpha^{\mathrm{reg}}(\mathcal{N}\|\mathcal{M}) = D^{\mathrm{reg}}(\mathcal{N}\|\mathcal{M}), \tag{74}$$

$$\lim_{\alpha \to 1^+} \widetilde{D}_\alpha^{A}(\mathcal{N}\|\mathcal{M}) = D^{A}(\mathcal{N}\|\mathcal{M}). \tag{75}$$

Proof. The first equation follows from Theorem 3. The second equation follows from the existing results $\widetilde{D}_\alpha^{\mathrm{reg}}(\mathcal{N}\|\mathcal{M}) = \widetilde{D}_\alpha^{A}(\mathcal{N}\|\mathcal{M})$ [55, Theorem 5.4] and $D^{\mathrm{reg}}(\mathcal{N}\|\mathcal{M}) = D^{A}(\mathcal{N}\|\mathcal{M})$ [24, Corollary 3].

Note that the exponentially strong converse property in Definition 12 is defined for coherent strategies. Here, we demonstrate that it is equivalent to the exponentially strong converse property under sequential strategies. This is quite remarkable, as sequential strategies can be significantly more general than coherent strategies.

Theorem 4. Let $\mathcal{N}, \mathcal{M} \in \mathrm{CPTP}(A \to B)$ be two quantum channels and $\mathcal{I} \in \mathrm{CPTP}(A \to A)$ be the identity channel. The exponentially strong converse property in Definition 12 holds true under coherent strategies for channels $\mathcal{I} \otimes \mathcal{N}$ and $\mathcal{I} \otimes \mathcal{M}$ if and only if it holds true under sequential strategies.

Proof. By the result in [23, Proposition 20], for any sequential strategies and $\alpha > 1$ it holds that

$$-\frac{1}{n}\log(1-\alpha_n) \geqslant \frac{\alpha-1}{\alpha}\left(-\frac{1}{n}\log\beta_n - \widetilde{D}_\alpha^{A}(\mathcal{N}\|\mathcal{M})\right). \tag{76}$$

By the exponentially strong converse property under sequential strategies, we assume $\liminf_{n\to\infty} -\frac{1}{n}\beta_n := r > D^{\mathrm{reg}}(\mathcal{N}\|\mathcal{M}) = D^{A}(\mathcal{N}\|\mathcal{M})$ where the second equality follows by [24, Corollary 3]. This implies that there exists $\delta > 0$ and a subsequence β_{n_k} such that $-\frac{1}{n_k}\log\beta_{n_k} \geqslant r - \delta > D^{A}(\mathcal{N}\|\mathcal{M})$ for sufficiently large n_k. By Corollary 1, the exponentially strong converse property in Definition 12 is equivalent to the continuity of the amortized channel divergence $\lim_{\alpha\to 1^+} \widetilde{D}_\alpha^{A}(\mathcal{N}\|\mathcal{M}) = D^{A}(\mathcal{N}\|\mathcal{M})$. As $r > D^{A}(\mathcal{N}\|\mathcal{M})$, there exists $\alpha_0 > 1$ such that $r - \delta > \widetilde{D}_{\alpha_0}^{A}(\mathcal{N}\|\mathcal{M})$. Then we have

$$-\frac{1}{n_k}\log\beta_{n_k} - \widetilde{D}_{\alpha_0}^{A}(\mathcal{N}\|\mathcal{M}) \geqslant r - \delta - \widetilde{D}_{\alpha_0}^{A}(\mathcal{N}\|\mathcal{M}) =: b > 0. \tag{77}$$

Taking this into (76), we get

$$-\frac{1}{n_k}\log(1-\alpha_{n_k}) \geqslant \frac{\alpha_0 - 1}{\alpha_0} b =: c > 0, \tag{78}$$

which is equivalent to $1-\alpha_{n_k} \leqslant 2^{-cn_k}$. This establishes the exponentially strong converse property under sequential strategies. Conversely, since any coherent strategy is a specific case of a sequential strategy, the strong converse property under sequential strategies also implies the property under coherent strategies.

4 Quantum channel discrimination in different regimes

The task of channel discrimination aims to distinguish a quantum channel from the other under a given type of strategy. A standard approach for discrimination is to perform hypothesis testing and make a decision based on the testing result. However, two types of error (Type-I error and Type-II error) arise. In the same spirit of state discrimination, one can study the asymptotic behavior of these errors in different operational regimes (see Figure 2), particularly, (I) error exponent regime that studies the exponent of the exponential convergence of the Type-I error given that the Type-II error exponentially decays; (II) Stein exponent regime that studies the exponent of the exponential decay of the Type-II error given that the Type-I error is within a constant threshold; (III) strong converse exponent regime that studies the exponent of the exponential convergence of the Type-I error given that the Type-II error exponentially decays.

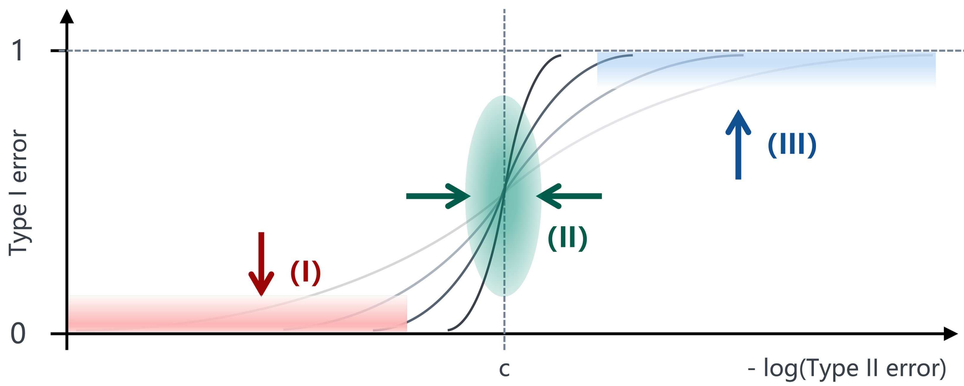

Figure 2 (Color online) Illustration depicting different regimes of quantum channel discrimination. Each curve represents the tradeoff between the Type-I and Type-II errors for varying block lengths, with darker lines corresponding to longer block lengths. (I) represents the error exponent regime, (II) represents the Stein exponent regime, and (III) represents the strong converse exponent regime.

Quantum state discrimination in different operational regimes has been well-studied. In particular, there is a nice correspondence between the regime studied and the quantum divergence to use. More precisely, the Stein exponent is given by the Umegaki relative entropy [4,5], the strong converse exponent is determined by the sandwiched Rényi divergence [11], and the error exponent is determined by the Petz Rényi divergence [7,8,61]. However, when it comes to channel discrimination, the situation becomes much more involved due to the diverse range of discrimination strategies and different extensions of channel divergence.

In this section, we study the interplay between the strategies of channel discrimination (e.g., sequential, coherent, product), the operational regimes (e.g., error exponent, Stein exponent, strong converse exponent), and three variants of channel divergences (e.g., Petz, Umegaki, sandwiched). We find a nice correspondence that shows that the proper divergences to use (Petz, Umegaki, sandwiched) are determined by the operational regime of interest, while the types of channel extension (one-shot, regularized, amortized) are determined by the discrimination strategies. Our results contribute towards a complete picture of channel discrimination in a unified framework.

4.1 Stein exponent

In this subsection, we consider minimizing the Type-II error probability, under the constraint that the Type-I error probability does not exceed a constant threshold $\varepsilon \in (0,1)$. We characterize the exact exponent, named Stein exponent, with which the Type-II error exponentially decays.

Definition 13 (Stein exponent)**.** Let $\mathcal{N},\mathcal{M} \in \mathrm{CPTP}(A \to B)$ be two quantum channels and $\varepsilon \in (0,1)$ be a fixed error. The Stein exponents of quantum channel discrimination by the strategy class $\Omega \in \{\mathrm{PRO},\mathrm{COH},\mathrm{SEQ}\}$ without quantum memory assistance are defined by

$$E^{\mathrm{st}}_{\Omega,\sup}(\varepsilon|\mathcal{N}\|\mathcal{M}) := \limsup_{n\to\infty}\frac{1}{n}d^{\varepsilon}_{H,\Omega}(\mathcal{N}^{\otimes n}\|\mathcal{M}^{\otimes n}), \tag{79}$$

$$E^{\mathrm{st}}_{\Omega,\inf}(\varepsilon|\mathcal{N}\|\mathcal{M}) := \liminf_{n\to\infty}\frac{1}{n}d^{\varepsilon}_{H,\Omega}(\mathcal{N}^{\otimes n}\|\mathcal{M}^{\otimes n}), \tag{80}$$

where

$$d^{\varepsilon}_{H,\Omega}(\mathcal{N}^{\otimes n}\|\mathcal{M}^{\otimes n}) := \sup_{(S_n,\Pi_n)\in\Omega}\left\{-\frac{1}{n}\log\beta_n(S_n,\Pi_n) : \alpha_n(S_n,\Pi_n) \leqslant \varepsilon\right\}. \tag{81}$$

The supremum is taken over all possible strategies $(S_n,\Pi_n)\in\Omega$ satisfying the condition and the type-I and type-II errors are defined in (44) and (45), respectively.

The non-asymptotic quantity in (81) can also be written as a notion of hypothesis testing relative entropy between the testing states,

$$d^{\varepsilon}_{H,\Omega}(\mathcal{N}^{\otimes n}\|\mathcal{M}^{\otimes n}) = \sup_{S_n\in\Omega} D^{\varepsilon}_H(\rho_n(S_n)\|\sigma_n(S_n)), \tag{82}$$

where the hypothesis testing relative entropy on the r.h.s. is between two quantum states which is defined in (10) and the supremum is taken over all strategies $S_n \in \Omega$ that generate the testing states $\rho_n(S_n)$ and $\sigma_n(S_n)$. More explicitly, when $\Omega = \text{PRO}$, we have $\rho_n(S_n) = \mathcal{N}^{\otimes n}(\otimes_{i=1}^n \varphi_i)$, $\sigma_n(S_n) = \mathcal{M}^{\otimes n}(\otimes_{i=1}^n \varphi_i)$ and the supremum is taken over all $\varphi_i \in \mathfrak{D}(R_i A_i)$. When $\Omega = \text{COH}$, we have $\rho_n(S_n) = \mathcal{N}^{\otimes n}(\psi_n)$, $\sigma_n(S_n) = \mathcal{M}^{\otimes n}(\psi_n)$ and the supremum is taken over all $\psi_n \in \mathfrak{D}(RA^n)$. When $\Omega = \text{SEQ}$, we have $\rho_n(S_n) = \mathcal{N} \circ \mathcal{P}_{n-1} \circ \cdots \circ \mathcal{P}_2 \circ \mathcal{N} \circ \mathcal{P}_1 \circ \mathcal{N}(\psi_n)$, $\sigma_n(S_n) = \mathcal{M} \circ \mathcal{P}_{n-1} \circ \cdots \circ \mathcal{P}_2 \circ \mathcal{M} \circ \mathcal{P}_1 \circ \mathcal{M}(\psi_n)$ and the supremum is taken over all $\psi_n \in \mathfrak{D}(R_1 A_1)$ and $\mathcal{P}_i \in \text{CPTP}(R_i B_i \to R_{i+1} A_{i+1})$.

Theorem 5 (Product strategy)**.** Let $\mathcal{N}, \mathcal{M} \in \text{CPTP}(A \to B)$ be two quantum channels and $\varepsilon \in (0,1)$ be a fixed error. Then it holds that

$$E^{\text{st}}_{\text{PRO},\sup}(\varepsilon|\mathcal{N}\|\mathcal{M}) = E^{\text{st}}_{\text{PRO},\inf}(\varepsilon|\mathcal{N}\|\mathcal{M}) = d(\mathcal{N}\|\mathcal{M}). \tag{83}$$

Proof. It suffices to show that

$$\lim_{n\to\infty} \frac{1}{n} d^{\varepsilon}_{H,\text{PRO}}(\mathcal{N}^{\otimes n}\|\mathcal{M}^{\otimes n}) = d(\mathcal{N}\|\mathcal{M}). \tag{84}$$

For the achievable part, let $\varphi \in \mathfrak{D}(A)$ be an optimal input state for $d(\mathcal{N}\|\mathcal{M})$, i.e., $D(\mathcal{N}(\varphi)\|\mathcal{M}(\varphi)) = d(\mathcal{N}\|\mathcal{M})$. Using $\varphi^{\otimes n}$ as the input state in the product strategy, we have

$$E^{\text{st}}_{\text{PRO},\inf}(\varepsilon|\mathcal{N}\|\mathcal{M}) \geqslant \liminf_{n\to\infty} \frac{1}{n} d^{\varepsilon}_H([\mathcal{N}(\varphi)]^{\otimes n}\|[\mathcal{M}(\varphi)]^{\otimes n}) = D(\mathcal{N}(\varphi)\|\mathcal{M}(\varphi)) = d(\mathcal{N}\|\mathcal{M}), \tag{85}$$

where the first equality follows from the quantum Stein's lemma [4,5] and the second equality follows from the optimality assumption of φ. For the converse part, consider any input states $\otimes_{i=1}^n \varphi_i$ with $\varphi_i \in \mathfrak{D}(A_i)$ and $\alpha > 1$ we have

$$\frac{1}{n} D^{\varepsilon}_H \left(\bigotimes_{i=1}^n \mathcal{N}(\varphi_i) \Big\| \bigotimes_{i=1}^n \mathcal{M}(\varphi_i) \right) \leqslant \frac{1}{n} \widetilde{D}_\alpha \left(\bigotimes_{i=1}^n \mathcal{N}(\varphi_i) \Big\| \bigotimes_{i=1}^n \mathcal{M}(\varphi_i) \right) + \frac{1}{n} \frac{\alpha}{\alpha-1} \log \frac{1}{1-\varepsilon} \tag{86}$$

$$= \frac{1}{n} \sum_{i=1}^n \widetilde{D}_\alpha \left(\mathcal{N}(\varphi_i) \| \mathcal{M}(\varphi_i) \right) + \frac{1}{n} \frac{\alpha}{\alpha-1} \log \frac{1}{1-\varepsilon} \tag{87}$$

$$\leqslant \widetilde{d}_\alpha(\mathcal{N}\|\mathcal{M}) + \frac{1}{n} \frac{\alpha}{\alpha-1} \log \frac{1}{1-\varepsilon}, \tag{88}$$

where the first inequality follows from (11), the first equality follows from the additivity of sandwiched Rényi divergence under tensor product states, and the second inequality follows from the definition of channel divergence. Taking the supremum of all input states $\otimes_{i=1}^n \varphi_i$ and taking the limit of $n \to \infty$, we have

$$E^{\text{st}}_{\text{PRO},\sup}(\varepsilon|\mathcal{N}\|\mathcal{M}) = \limsup_{n\to\infty} \frac{1}{n} d^{\varepsilon}_{H,\text{PRO}}(\mathcal{N}^{\otimes n}\|\mathcal{M}^{\otimes n}) \leqslant \widetilde{d}_\alpha(\mathcal{N}\|\mathcal{M}). \tag{89}$$

Finally, taking $\alpha \to 1$ and applying Lemma 2, we have the converse part.

Note that we can actually extend the input choices of product strategy to a convex combination of tensor product states $\sum_{j=1}^m p_j(\otimes_{i=1}^n \varphi_{i,j})$. In this case, Theorem 5 still holds by adding an extra step in the proof of the converse part and using the joint quasi-convexity of the sandwiched Rényi divergence (e.g., [11, Corollary 3.16]). This indicates that shared randomness between the input states for each use of the channel will not help to get a faster convergence rate of the Type-II error for channel discrimination.

Theorem 6 (Coherent strategy)**.** Let $\mathcal{N}, \mathcal{M} \in \text{CPTP}(A \to B)$ be two quantum channels and $\varepsilon \in (0,1)$ be a fixed error. If these channels exhibit the strong converse property as defined in Definition 11, then it implies that

$$E^{\text{st}}_{\text{COH},\sup}(\varepsilon|\mathcal{N}\|\mathcal{M}) = E^{\text{st}}_{\text{COH},\inf}(\varepsilon|\mathcal{N}\|\mathcal{M}) = d^{\text{reg}}(\mathcal{N}\|\mathcal{M}). \tag{90}$$

Proof. The assertion is a combination of (54) (achievability) and a restatement of Theorem 1 (converse) by noting that

$$d^{\varepsilon}_{H,\text{COH}}(\mathcal{N}^{\otimes n}\|\mathcal{M}^{\otimes n}) = d^{\varepsilon}_H(\mathcal{N}^{\otimes n}\|\mathcal{M}^{\otimes n}), \tag{91}$$

where the l.h.s. is the operational definition and the r.h.s. is the mathematical definition.

Theorem 7 (Sequential strategy)**.** Let $\mathcal{N},\mathcal{M}\in\mathrm{CPTP}(A\to B)$ be two quantum channels. Let $\mathcal{I}\in\mathrm{CPTP}(A\to A)$ be the identity channel. Then if the exponentially strong converse property, as defined in Definition 12, holds for the channels $\mathcal{I}\otimes\mathcal{N}$ and $\mathcal{I}\otimes\mathcal{M}$, this implies that

$$E^{\mathrm{st}}_{\mathrm{SEQ,sup}}(\varepsilon|\mathcal{N}\|\mathcal{M})=E^{\mathrm{st}}_{\mathrm{SEQ,inf}}(\varepsilon|\mathcal{N}\|\mathcal{M})=D^{A}(\mathcal{N}\|\mathcal{M}). \tag{92}$$

Proof. By definition it is clear that $E^{\mathrm{st}}_{\mathrm{SEQ,inf}}(\varepsilon|\mathcal{N}\|\mathcal{M})$ is monotone increasing in ε. Thus for any fixed $\varepsilon\in(0,1)$ we have

$$E^{\mathrm{st}}_{\mathrm{SEQ,inf}}(\varepsilon|\mathcal{N}\|\mathcal{M})\geqslant\lim_{\varepsilon\to 0}E^{\mathrm{st}}_{\mathrm{SEQ,inf}}(\varepsilon|\mathcal{N}\|\mathcal{M})=D^{A}(\mathcal{N}\|\mathcal{M}), \tag{93}$$

where the equality follows from [13, Theorem 6].

Next we prove the converse part. For any $\psi_n\in\mathfrak{D}(RA^n)$, $\mathcal{P}_i\in\mathrm{CPTP}(R_iB_i\to R_{i+1}A_{i+1})$, denote

$$\rho_n=\mathcal{N}\circ\mathcal{P}_{n-1}\circ\cdots\circ\mathcal{P}_2\circ\mathcal{N}\circ\mathcal{P}_1\circ\mathcal{N}(\psi_n), \tag{94}$$

$$\sigma_n=\mathcal{M}\circ\mathcal{P}_{n-1}\circ\cdots\circ\mathcal{P}_2\circ\mathcal{M}\circ\mathcal{P}_1\circ\mathcal{M}(\psi_n). \tag{95}$$

Due to (11), it holds for any $\alpha>1$ that

$$\frac{1}{n}D^{\varepsilon}_{H}(\rho_n\|\sigma_n)\leqslant\frac{1}{n}\widetilde{D}_{\alpha}(\rho_n\|\sigma_n)+\frac{1}{n}\frac{\alpha}{\alpha-1}\log\frac{1}{1-\varepsilon}. \tag{96}$$

Note that for any quantum state ρ,σ and quantum channels $\mathcal{E},\mathcal{F}$, we have by definition

$$\widetilde{D}_{\alpha}(\mathcal{E}(\rho)\|\mathcal{F}(\sigma))\leqslant\widetilde{D}^{A}_{\alpha}(\mathcal{E}\|\mathcal{F})+\widetilde{D}_{\alpha}(\rho\|\sigma). \tag{97}$$

By using this relation and the data-processing inequality of $\widetilde{D}_{\alpha}$ iteratively, we have $\widetilde{D}_{\alpha}(\rho_n\|\sigma_n)\leqslant n\widetilde{D}^{A}_{\alpha}(\mathcal{N}\|\mathcal{M})$. This gives

$$\frac{1}{n}D^{\varepsilon}_{H}(\rho_n\|\sigma_n)\leqslant\widetilde{D}^{A}_{\alpha}(\mathcal{N}\|\mathcal{M})+\frac{1}{n}\frac{\alpha}{\alpha-1}\log\frac{1}{1-\varepsilon}. \tag{98}$$

Taking on both sides the supremum over all sequential strategies following by the limit $n\to\infty$ gives

$$E^{\mathrm{st}}_{\mathrm{SEQ,sup}}(\varepsilon|\mathcal{N}\|\mathcal{M})=\limsup_{n\to\infty}\frac{1}{n}D^{\varepsilon}_{H,\mathrm{SEQ}}\left(\mathcal{N}^{\otimes n}\middle\|\mathcal{M}^{\otimes n}\right)\leqslant\widetilde{D}^{A}_{\alpha}(\mathcal{N}\|\mathcal{M}). \tag{99}$$

Since the above inequality holds for all $\alpha>1$, by taking $\alpha\to 1^{+}$ and using Corollary 1, we have

$$E^{\mathrm{st}}_{\mathrm{SEQ,sup}}(\varepsilon|\mathcal{N}\|\mathcal{M})\leqslant D^{A}(\mathcal{N}\|\mathcal{M}). \tag{100}$$

Combining (93) and (100), we have the complete proof.

Note that Theorems 6 and 7 have been proved in [13, Theorems 3 and 6] for vanishing ε. But the above results are stronger as they hold for any fixed $\varepsilon\in(0,1)$ without the need to take $\varepsilon\to 0$.

4.2 Strong converse exponent

In the task of state discrimination, the strong converse exponent is defined by

$$E^{\mathrm{sc}}(r|\rho\|\sigma):=\inf_{\{\Pi_n\}}\left\{-\liminf_{n\to+\infty}\frac{1}{n}\log\mathrm{Tr}\,\rho^{\otimes n}\Pi_n:\limsup_{n\to+\infty}\frac{1}{n}\log\mathrm{Tr}\,\sigma^{\otimes n}\Pi_n\leqslant -r\right\}, \tag{101}$$

where the infimum is taken over all possible sequences of quantum tests $\{\Pi_n\}_{n\in\mathbb{N}}$ satisfying the condition. It has been shown in [11, Theorem 4.10] that this exponent is precisely characterized by

$$E^{\mathrm{sc}}(r|\rho\|\sigma)=\sup_{\alpha>1}\frac{\alpha-1}{\alpha}\left[r-\widetilde{D}_{\alpha}(\rho\|\sigma)\right]. \tag{102}$$

We aim to extend this result to the channel case.

Let us start by defining the strong converse exponent of channel discrimination.

Definition 14 (Strong converse exponent)**.** Let $\mathcal{N}, \mathcal{M} \in \mathrm{CPTP}(A \to B)$ and $r > 0$. The strong converse exponents of channel discrimination by the strategy class $\Omega \in \{\mathrm{PRO}, \mathrm{COH}, \mathrm{SEQ}\}$ without quantum memory assistance are defined by

$$E_{\Omega}^{\mathrm{sc}}(r|\mathcal{N}\|\mathcal{M}) := \inf_{(S_n,\Pi_n)\in\Omega} \left\{-\liminf_{n\to+\infty}\frac{1}{n}\log(1-\alpha_n(S_n,\Pi_n)) : \limsup_{n\to+\infty}\frac{1}{n}\log\beta_n(S_n,\Pi_n) \leqslant -r\right\}, \tag{103}$$

where the infimum is taken over all possible strategies $(S_n, \Pi_n) \in \Omega$ satisfying the condition.

Theorem 8 (Product strategy)**.** Let $\mathcal{N}, \mathcal{M} \in \mathrm{CPTP}(A \to B)$ and $r > 0$. Then it holds that

$$E_{\mathrm{PRO}}^{\mathrm{sc}}(r|\mathcal{N}\|\mathcal{M}) = \sup_{\alpha>1}\frac{\alpha-1}{\alpha}\left[r - \widetilde{d}_\alpha(\mathcal{N}\|\mathcal{M})\right]. \tag{104}$$

Proof. We first prove the converse part which closely follows the proof of its state analog in [11, Lemma 4.7]. For any product strategy $(\{\varphi_i\}_{i=1}^n, \Pi_n)$ with input states $\varphi_i \in \mathfrak{D}(A_i)$ and measurement operator $0 \leqslant \Pi_n \leqslant I$. Let $\rho_n := \mathcal{N}^{\otimes n}(\otimes_{i=1}^n \varphi_i)$, $\sigma_n := \mathcal{M}^{\otimes n}(\otimes_{i=1}^n \varphi_i)$ be the output states and $p_n := (\operatorname{Tr}\rho_n\Pi_n, \operatorname{Tr}\rho_n(I-\Pi_n))$ and $q_n := (\operatorname{Tr}\sigma_n\Pi_n, \operatorname{Tr}\sigma_n(I-\Pi_n))$ be the post-measurement states. Then the Type-I error is $\alpha_n = \operatorname{Tr}[(I-\Pi_n)\rho_n]$ and the Type-II error is $\beta_n = \operatorname{Tr}[\Pi_n\sigma_n]$. By definition it suffices to consider sequences $(\{\varphi_i\}_{i=1}^n, \Pi_n)$ such that $\limsup_{n\to+\infty}\frac{1}{n}\log\beta_n \leqslant -r$. From the data-processing of the sandwiched Rényi divergence, we have for any $\alpha > 1$ that

$$\begin{aligned}\widetilde{D}_\alpha(\rho_n\|\sigma_n) &\geqslant \widetilde{D}_\alpha(p_n\|q_n)\\ &\geqslant \frac{1}{\alpha-1}\log\left[(\operatorname{Tr}\rho_n\Pi_n)^\alpha(\operatorname{Tr}\sigma_n\Pi_n)^{1-\alpha}\right] = \frac{\alpha}{\alpha-1}\log(1-\alpha_n) - \log\beta_n.\end{aligned} \tag{105}$$

This can be equivalently written as

$$-\frac{1}{n}\log(1-\alpha_n) \geqslant \frac{\alpha-1}{\alpha}\left[-\frac{1}{n}\log\beta_n - \frac{1}{n}\widetilde{D}_\alpha(\rho_n\|\sigma_n)\right]. \tag{106}$$

By the assumption of $(\{\varphi_i\}_{i=1}^n, \Pi_n)$ and taking $\limsup_{n\to\infty}$ on both sides, we have

$$-\liminf_{n\to+\infty}\frac{1}{n}\log(1-\alpha_n) \geqslant \frac{\alpha-1}{\alpha}\left[r - \liminf_{n\to+\infty}\frac{1}{n}\widetilde{D}_\alpha(\rho_n\|\sigma_n)\right]. \tag{107}$$

By the additivity of sandwiched Rényi divergence under tensor product states and the definition of channel divergence, we have $\widetilde{D}_\alpha(\rho_n\|\sigma_n) = \sum_{i=1}^n \widetilde{D}_\alpha(\mathcal{N}(\varphi_i)\|\mathcal{M}(\varphi_i)) \leqslant n\widetilde{d}_\alpha(\mathcal{N}\|\mathcal{M})$. Thus

$$-\liminf_{n\to+\infty}\frac{1}{n}\log(1-\alpha_n) \geqslant \frac{\alpha-1}{\alpha}\left[r - \widetilde{d}_\alpha(\mathcal{N}\|\mathcal{M})\right]. \tag{108}$$

Finally taking the infimum over all product strategies and the supremum over all $\alpha > 1$ on both sides, we can conclude the converse part

$$E_{\mathrm{PRO}}^{\mathrm{sc}}(r|\mathcal{N}\|\mathcal{M}) \geqslant \sup_{\alpha>1}\frac{\alpha-1}{\alpha}\left[r - \widetilde{d}_\alpha(\mathcal{N}\|\mathcal{M})\right]. \tag{109}$$

We then proceed to show the achievable part. Let $\varphi \in \mathfrak{D}(A)$ be an optimal quantum state such that $\widetilde{d}_\alpha(\mathcal{N}\|\mathcal{M}) = \widetilde{D}_\alpha(\mathcal{N}(\varphi)\|\mathcal{M}(\varphi))$. Consider the task of distinguishing quantum states $\mathcal{N}(\varphi)$ and $\mathcal{M}(\varphi)$. Suppose the optimal test in $E^{\mathrm{sc}}(r|\mathcal{N}(\varphi)\|\mathcal{M}(\varphi))$ is given by the sequence $\{\Pi_n\}_{n\in\mathbb{N}}$. Then by the quantum converse Hoeffiding theorem (see (102)) we have

$$\limsup_{n\to+\infty}\frac{1}{n}\log\operatorname{Tr}[\mathcal{M}(\varphi)]^{\otimes n}\Pi_n \leqslant -r, \tag{110}$$

$$-\liminf_{n\to+\infty}\frac{1}{n}\log\operatorname{Tr}[\mathcal{N}(\varphi)]^{\otimes n}\Pi_n = \sup_{\alpha>1}\frac{\alpha-1}{\alpha}\left[r - \widetilde{D}_\alpha(\mathcal{N}(\varphi)\|\mathcal{M}(\varphi))\right]. \tag{111}$$

Note that $(\{\varphi\}_{i=1}^n, \Pi_n)$ is a product strategy for the task of channel discrimination between $\mathcal{N}^{\otimes n}$ and $\mathcal{M}^{\otimes n}$. We have

$$E_{\mathrm{PRO}}^{\mathrm{sc}}(r|\mathcal{N}\|\mathcal{M}) \leqslant -\liminf_{n\to+\infty}\frac{1}{n}\log\operatorname{Tr}\mathcal{N}^{\otimes n}(\varphi^{\otimes n})\Pi_n \tag{112}$$

$$= \sup_{\alpha>1} \frac{\alpha-1}{\alpha} \left[r - \widetilde{D}_\alpha(\mathcal{N}(\varphi)\|\mathcal{M}(\varphi)) \right] \tag{113}$$

$$= \sup_{\alpha>1} \frac{\alpha-1}{\alpha} \left[r - \widetilde{d}_\alpha(\mathcal{N}\|\mathcal{M}) \right], \tag{114}$$

where the first equality follows from (111), the second equality follows from the optimality assumption of φ. Combining (109) and (114), we have the complete proof.

Note here that one can extend the input choices of product strategy to a convex combination of tensor product states $\sum_{j=1}^m p_j(\otimes_{i=1}^n \varphi_{i,j})$. In this case, Theorem 8 still holds by adding an additional step in the proof of the converse part and using the joint quasi-convexity of the sandwiched Rényi divergence (e.g., [11, Corollary 3.16]). This indicates that shared randomness between the input states for each use of the channel will provide no advantage in reducing the convergence rate of the Type-I error.

Remark 2. The strong converse exponents under coherent and sequential strategies were established in [55, Theorem 5.5]. However, regarding the exact threshold for exponential convergence, their result only identifies the threshold as $\inf_{\alpha>1} \widetilde{D}_\alpha^{\mathrm{reg}}(\mathcal{N}\|\mathcal{M})$. The continuity result in Theorem 3 could fully determine this threshold as $D^{\mathrm{reg}}(\mathcal{N}\|\mathcal{M})$ if the strong converse property can be proven.

4.3 Error exponent

In the task of state discrimination, the error exponent is defined by

$$E^{\mathrm{er}}(r|\rho\|\sigma) := \sup_{\{\Pi_n\}} \left\{ -\limsup_{n\to+\infty} \frac{1}{n} \log \mathrm{Tr}[(I-\Pi_n)\rho^{\otimes n}] : \limsup_{n\to+\infty} \frac{1}{n} \log \mathrm{Tr}[\Pi_n \sigma^{\otimes n}] \leqslant -r \right\}, \tag{115}$$

where the supremum is taken over all possible sequences of quantum tests $\{\Pi_n\}_{n\in\mathbb{N}}$ satisfying the condition. It has been shown in [7, 8, 61] that the error exponent is precisely given by

$$E^{\mathrm{er}}(r|\rho\|\sigma) = \sup_{0<\alpha<1} \frac{\alpha-1}{\alpha} \left[r - \tilde{D}_\alpha(\rho\|\sigma) \right]. \tag{116}$$

We aim to extend this result to the channel case.

Definition 15 (Error exponent)**.** Let $\mathcal{N}, \mathcal{M} \in \mathrm{CPTP}(A \to B)$ and $r > 0$. The error exponents of quantum channel discrimination by the strategy class $\Omega \in \{\mathrm{PRO}, \mathrm{COH}, \mathrm{SEQ}\}$ without quantum memory assistance are defined by

$$E_\Omega^{\mathrm{er}}(r|\mathcal{N}\|\mathcal{M}) := \sup_{(S_n,\Pi_n)\in\Omega} \left\{ -\limsup_{n\to+\infty} \frac{1}{n} \log \alpha_n(S_n,\Pi_n) : \limsup_{n\to+\infty} \frac{1}{n} \log \beta_n(S_n,\Pi_n) \leqslant -r \right\}, \tag{117}$$

where the supremum is taken over all possible strategies $(S_n, \Pi_n) \in \Omega$ satisfying the condition.

Theorem 9 (Product strategy)**.** Let $\mathcal{N}, \mathcal{M} \in \mathrm{CPTP}(A \to B)$ and $r > 0$. Then it holds that

$$E_{\mathrm{PRO}}^{\mathrm{er}}(r|\mathcal{N}\|\mathcal{M}) \geqslant \sup_{0<\alpha<1} \frac{\alpha-1}{\alpha} \left[r - \tilde{D}_\alpha(\mathcal{N}\|\mathcal{M}) \right]. \tag{118}$$

Proof. Let $\varphi \in \mathfrak{D}(A)$ an optimal input state such that $\bar{d}_\alpha(\mathcal{N}\|\mathcal{M}) = \bar{D}_\alpha(\mathcal{N}(\varphi)\|\mathcal{M}(\varphi))$. Consider the task of distinguishing quantum states $\mathcal{N}(\varphi)$ and $\mathcal{M}(\varphi)$. Suppose the optimal test in $E^{\mathrm{er}}(r|\mathcal{N}(\varphi)\|\mathcal{M}(\varphi))$ is given by the sequence $\{\Pi_n\}_{n\in\mathbb{N}}$. Then by the quantum Hoeffding theorem (116) we have

$$\limsup_{n\to+\infty} \frac{1}{n} \log \mathrm{Tr}[\mathcal{M}(\varphi)]^{\otimes n} \Pi_n \leqslant -r, \tag{119}$$

$$-\limsup_{n\to+\infty} \frac{1}{n} \log(1 - \mathrm{Tr}[\mathcal{N}(\varphi)]^{\otimes n} \Pi_n) = \sup_{0<\alpha<1} \frac{\alpha-1}{\alpha} \left[r - \bar{D}_\alpha(\mathcal{N}(\varphi)\|\mathcal{M}(\varphi)) \right]. \tag{120}$$

Note that $(\{\varphi\}_{i=1}^n, \Pi_n)$ is a product strategy for the task of channel discrimination between $\mathcal{N}^{\otimes n}$ and $\mathcal{M}^{\otimes n}$. Then we have

$$E_{\mathrm{PRO}}^{\mathrm{er}}(r|\mathcal{N}\|\mathcal{M}) \geqslant -\limsup_{n\to+\infty} \frac{1}{n} \log(1 - \mathrm{Tr}\, \mathcal{N}^{\otimes n}(\varphi^{\otimes n})\Pi_n) \tag{121}$$

$$= -\limsup_{n\to+\infty}\frac{1}{n}\log(1-\mathrm{Tr}[\mathcal{N}(\varphi)]^{\otimes n}\Pi_n) \tag{122}$$

$$= \sup_{0<\alpha<1}\frac{\alpha-1}{\alpha}\left[r-\bar{D}_\alpha(\mathcal{N}(\varphi)\|\mathcal{M}(\varphi))\right] \tag{123}$$

$$= \sup_{0<\alpha<1}\frac{\alpha-1}{\alpha}\left[r-\bar{d}_\alpha(\mathcal{N}\|\mathcal{M})\right], \tag{124}$$

where the second equality follows from (120), the third equality follows from the optimality assumption of φ. This completes the proof.

Theorem 10 (Coherent strategy)**.** Let $\mathcal{N},\mathcal{M}\in\mathrm{CPTP}(A\to B)$ and $r>0$. Then it holds that

$$E^{\mathrm{er}}_{\mathrm{COH}}(r|\mathcal{N}\|\mathcal{M}) \geqslant \sup_{0<\alpha<1}\frac{\alpha-1}{\alpha}\left[r-\bar{d}^{\mathrm{reg}}_\alpha(\mathcal{N}\|\mathcal{M})\right]. \tag{125}$$

Proof. For any given $m\in\mathbb{N}$, let $\psi_m\in\mathfrak{D}(A^m)$ an optimal input state such that $\bar{d}_\alpha(\mathcal{N}^{\otimes m}\|\mathcal{M}^{\otimes m}) = \bar{D}_\alpha(\mathcal{N}^{\otimes m}(\psi_m)\|\mathcal{M}^{\otimes m}(\psi_m))$. Denote $\rho_m := \mathcal{N}^{\otimes m}(\psi_m)$ and $\sigma_m := \mathcal{M}^{\otimes m}(\psi_m)$. Consider the task of distinguishing quantum states ρ_m and σ_m. Suppose the optimal test in $E^{\mathrm{er}}(r|\rho_m\|\sigma_m)$ is given by the sequence $\{\Pi_{m,n}\}_{n\in\mathbb{N}}$. Then by the quantum Hoeffding theorem (see (116)) we have

$$\limsup_{n\to+\infty}\frac{1}{n}\log\mathrm{Tr}[\sigma_m]^{\otimes n}\Pi_{m,n}\leqslant -r, \tag{126}$$

$$-\limsup_{n\to+\infty}\frac{1}{n}\log(1-\mathrm{Tr}[\rho_m]^{\otimes n}\Pi_{m,n}) = \sup_{0<\alpha<1}\frac{\alpha-1}{\alpha}\left[r-\bar{D}_\alpha(\rho_m\|\sigma_m)\right]. \tag{127}$$

Note that $(\psi_m^{\otimes n},\Pi_{m,n})$ is a coherent strategy for the task of channel discrimination between $\mathcal{N}^{\otimes mn}$ and $\mathcal{M}^{\otimes mn}$, satisfying

$$\limsup_{n\to+\infty}\frac{1}{mn}\log\mathrm{Tr}\,\mathcal{M}^{\otimes mn}(\psi_m^{\otimes n})\Pi_{m,n}\leqslant -\frac{r}{m}. \tag{128}$$

Then we have

$$E^{\mathrm{er}}_{\mathrm{COH}}\left(\frac{r}{m}|\mathcal{N}\|\mathcal{M}\right) \geqslant -\limsup_{n\to+\infty}\frac{1}{mn}\log(1-\mathrm{Tr}\,\mathcal{N}^{\otimes mn}(\psi_m^{\otimes n})\Pi_{m,n}) \tag{129}$$

$$= -\limsup_{n\to+\infty}\frac{1}{mn}\log(1-\mathrm{Tr}[\rho_m]^{\otimes n}\Pi_{m,n}) \tag{130}$$

$$= \frac{1}{m}\sup_{0<\alpha<1}\frac{\alpha-1}{\alpha}\left[r-\bar{D}_\alpha(\rho_m\|\sigma_m)\right] \tag{131}$$

$$= \frac{1}{m}\sup_{0<\alpha<1}\frac{\alpha-1}{\alpha}\left[r-\bar{d}_\alpha(\mathcal{N}^{\otimes m}\|\mathcal{M}^{\otimes m})\right] \tag{132}$$

$$= \sup_{0<\alpha<1}\frac{\alpha-1}{\alpha}\left[\frac{r}{m}-\frac{1}{m}\bar{d}_\alpha(\mathcal{N}^{\otimes m}\|\mathcal{M}^{\otimes m})\right], \tag{133}$$

where the second equality follows from (127), the third equality follows from the optimality assumption of ψ_m. Replacing r/m as r, we have

$$E^{\mathrm{er}}_{\mathrm{COH}}(r|\mathcal{N}\|\mathcal{M}) \geqslant \sup_{0<\alpha<1}\frac{\alpha-1}{\alpha}\left[r-\frac{1}{m}\bar{d}_\alpha(\mathcal{N}^{\otimes m}\|\mathcal{M}^{\otimes m})\right]. \tag{134}$$

Since Eq. (134) holds for any integer $m\in\mathbb{N}$, we have

$$E^{\mathrm{er}}_{\mathrm{COH}}(r|\mathcal{N}\|\mathcal{M}) \geqslant \sup_{m\in\mathbb{N}}\sup_{0<\alpha<1}\frac{\alpha-1}{\alpha}\left[r-\frac{1}{m}\bar{d}_\alpha(\mathcal{N}^{\otimes m}\|\mathcal{M}^{\otimes m})\right] \tag{135}$$

$$= \sup_{0<\alpha<1}\sup_{m\in\mathbb{N}}\frac{\alpha-1}{\alpha}\left[r-\frac{1}{m}\bar{d}_\alpha(\mathcal{N}^{\otimes m}\|\mathcal{M}^{\otimes m})\right] \tag{136}$$

$$= \sup_{0<\alpha<1}\frac{\alpha-1}{\alpha}\left[r-\sup_{m\in\mathbb{N}}\frac{1}{m}\bar{d}_\alpha(\mathcal{N}^{\otimes m}\|\mathcal{M}^{\otimes m})\right] \tag{137}$$

$$= \sup_{0<\alpha<1}\frac{\alpha-1}{\alpha}\left[r-\bar{d}^{\mathrm{reg}}_\alpha(\mathcal{N}\|\mathcal{M})\right]. \tag{138}$$

This completes the proof.

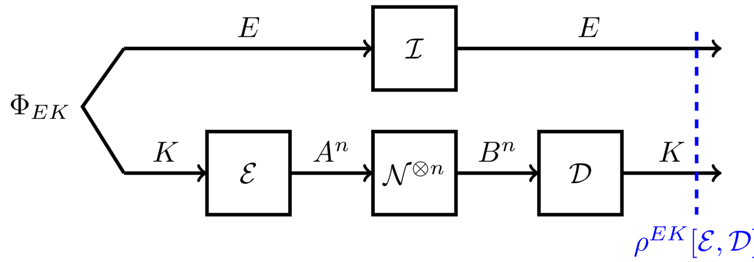

Figure 3 (Color online) Coding scheme for entanglement transmission over n uses of a channel $\mathcal{N} \in \mathrm{CPTP}(A \to B)$. The systems E and K are isomorphic. The encoder $\mathcal{E} \in \mathrm{CPTP}(K \to A^n)$ encodes the part K of the maximally entangled state Φ_{EK} into the channel input systems. Later, the decoder $\mathcal{D} \in \mathrm{CPTP}(B^n \to K)$ recovers the state from the channel output systems. The map $\mathcal{I} \in \mathrm{CPTP}(E \to E)$ is the identity map. The final state after the coding strategy is denoted as $\rho_{EK}[\mathcal{E},\mathcal{D}]$ and the performance of the code is quantified using the fidelity $F(\Phi_{EK}, \rho_{EK}[\mathcal{E},\mathcal{D}])$.

5 Quantum communication as quantum channel discrimination

Quantum communication via quantum channels forms the cornerstone of future quantum networks [62] and the quantum channel capacity is a central question in quantum Shannon theory [26–29]. In this section, we present a perspective by framing the study of quantum communication problems as quantum channel discrimination tasks. This perspective offers deeper insights into the intricate relationships between channel capacities, channel discrimination, and the mathematical properties of quantum channel divergences. On one hand, leveraging this connection, we demonstrate that the channel coherent information and quantum channel capacity can be precisely characterized as Stein exponent for discriminating between two quantum channels under the product and coherent strategies without quantum memory assistance, respectively. Furthermore, we show that the strong converse property of quantum channel capacity can be established if the channels being discriminated exhibit the strong converse property. On the other hand, the extreme non-additivity of quantum channel capacity implies a similar fundamental property for the unstabilized channel divergence, which can be of independent interest for future studies.

5.1 Operational interpretation of quantum channel capacity

In this subsection, we discuss quantum channel communication and its operational interpretation in the context of quantum channel discrimination. The coding scheme for n uses of the channel is depicted in Figure 3. We are given a quantum channel $\mathcal{N} \in \mathrm{CPTP}(A \to B)$ and denote by $\mathcal{N}^{\otimes n}$ the n-fold parallel repetition of this channel. An entanglement transmission code for $\mathcal{N}^{\otimes n}$ is given by a triplet $\{|K|, \mathcal{E}, \mathcal{D}\}$, where $|K|$ is the local dimension of a maximally entangled state $\Phi_{EK} := \frac{1}{|E|}\sum_{i,j=1}^{|E|} |ii\rangle\langle jj|$ that is to be transmitted over $\mathcal{N}^{\otimes n}$. The quantum channels $\mathcal{E} \in \mathrm{CPTP}(K \to A^n)$ and $\mathcal{D} \in \mathrm{CPTP}(B^n \to K)$ are encoding and decoding operations, respectively. Denote the outcome state after the coding strategy by

$$\rho_{EK}[\mathcal{E},\mathcal{D}] := \mathcal{I}_E \otimes \mathcal{D}_{B\to K} \circ \mathcal{N}_{A\to B} \circ \mathcal{E}_{K\to A}(\Phi_{EK}). \tag{139}$$

With this in hand, we now say that a triplet $\{r, n, \varepsilon\}$ is achievable on the channel $\mathcal{N}$ if there exists an entanglement transmission code satisfying

$$\frac{1}{n}\log|K| \geqslant r \qquad \text{and} \qquad F(\Phi_{EK}, \rho_{EK}[\mathcal{E},\mathcal{D}]) \geqslant 1-\varepsilon, \tag{140}$$

where $F(\rho,\sigma) := (\|\sqrt{\rho}\sqrt{\sigma}\|_1)^2$ is the quantum fidelity and $\|\cdot\|_1$ is the trace norm. If one of the states is pure, we have the simplification $F(|\psi\rangle\langle\psi|, \sigma) = \mathrm{Tr}[|\psi\rangle\langle\psi|\sigma]$.

When considering a single use of the channel, the one-shot quantum capacity, which establishes the boundary of all achievable triples $\{r, 1, \varepsilon\}$, is defined as follows.

Definition 16. Let $\mathcal{N} \in \mathrm{CPTP}(A \to B)$ be a quantum channel and $\varepsilon \in (0,1)$ be a fixed error. The one-shot quantum capacity of $\mathcal{N}$ is defined by

$$Q^{(1)}(\mathcal{N},\varepsilon) := \sup_{\substack{|K|=|E|\in\mathbb{N}\\ \mathcal{E}\in\mathrm{CPTP}(K\to A)\\ \mathcal{D}\in\mathrm{CPTP}(B\to K)}} \Big\{\log|K| : \mathrm{Tr}\left(\rho_{EK}[\mathcal{E},\mathcal{D}]\cdot\Phi_{EK}\right) \geqslant 1-\varepsilon\Big\}. \tag{141}$$

Then the quantum capacity is defined as the asymptotic limit

$$Q(\mathcal{N}) := \lim_{\varepsilon\to 0}\lim_{n\to\infty}\frac{1}{n}Q^{(1)}(\mathcal{N}^{\otimes n},\varepsilon). \tag{142}$$

The well-established Lloyd-Shor-Devetak theorem [63–65] states that the quantum capacity of a channel can be expressed in terms of a regularized channel coherent information:

$$Q(\mathcal{N}) = \lim_{n\to\infty}\frac{1}{n}I_c(\mathcal{N}^{\otimes n}) = \sup_{n\in\mathbb{N}}\frac{1}{n}I_c(\mathcal{N}^{\otimes n}), \tag{143}$$

where the channel coherent information is defined by

$$I_c(\mathcal{N}) := \sup_{\rho\in\mathfrak{D}(EA)} -H(E|B)_\rho \quad \text{with} \quad \rho_{EB} = \mathcal{I}_E\otimes\mathcal{N}_{A\to B}(\rho_{EA}), \tag{144}$$

and the supremum is taken over all density matrices ρ on system $E\otimes A$ and E is isomorphic to A.

Based on the notion of unstabilized channel divergence, we can rewrite the channel coherent information and quantum capacity as follows.

Theorem 11. For any quantum channel $\mathcal{N}\in\mathrm{CPTP}(A\to B)$, it holds that

$$I_c(\mathcal{N}) = d\left(\mathcal{I}_E\otimes\mathcal{N}_{A\to B}\middle\|\mathcal{R}_E^\pi\otimes\mathcal{N}_{A\to B}\right) - \log|E|, \tag{145}$$

$$Q(\mathcal{N}) = d^{\mathrm{reg}}\left(\mathcal{I}_E\otimes\mathcal{N}_{A\to B}\middle\|\mathcal{R}_E^\pi\otimes\mathcal{N}_{A\to B}\right) - \log|E|, \tag{146}$$

where E is isomorphic to A and $\mathcal{R}_E^\pi\in\mathrm{CPTP}(E\to E)$ represents a replacer channel that maps any input state to a maximally mixed state $\pi\in\mathfrak{D}(E)$.

Proof. For any given state $\rho\in\mathfrak{D}(EB)$, we have

$$-H(E|B)_\rho = D(\rho_{EB}\|I_E\otimes[\mathrm{Tr}_E\,\rho_{EB}]) = D(\rho_{EB}\|\mathcal{R}_E^\pi\otimes\mathcal{I}_B(\rho_{EB})) - \log|E|. \tag{147}$$

Let $\rho_{EB} = \mathcal{I}_E\otimes\mathcal{N}_{A\to B}(\rho_{EA})$ and take supremum over all $\rho\in\mathfrak{D}(EA)$. We get

$$I_c(\mathcal{N}) = \sup_{\rho\in\mathfrak{D}(EA)} D\left(\mathcal{I}_E\otimes\mathcal{N}_{A\to B}(\rho_{EA})\middle\|(\mathcal{R}_E^\pi\otimes\mathcal{I}_B)\circ(\mathcal{I}_E\otimes\mathcal{N}_{A\to B})(\rho_{EA})\right) - \log|E| \tag{148}$$

$$= d\left(\mathcal{I}_E\otimes\mathcal{N}_{A\to B}\middle\|\mathcal{R}_E^I\otimes\mathcal{N}_{A\to B}\right) - \log|E|. \tag{149}$$

Eq. (146) directly follows from (145) by taking a regularization on both sides. That is,

$$Q(\mathcal{N}) = \lim_{n\to\infty}\frac{1}{n}I_c(\mathcal{N}^{\otimes n}) \tag{150}$$

$$= \lim_{n\to\infty}\frac{1}{n}d\left(\mathcal{I}_{E^n}\otimes[\mathcal{N}_{A\to B}]^{\otimes n}\middle\|\mathcal{R}_{\pi^{\otimes n}}\otimes[\mathcal{N}_{A\to B}]^{\otimes n}\right) - \log|E| \tag{151}$$

$$= \lim_{n\to\infty}\frac{1}{n}d\left(\left[\mathcal{I}_E\otimes\mathcal{N}_{A\to B}\right]^{\otimes n}\middle\|\left[\mathcal{R}_\pi\otimes\mathcal{N}_{A\to B}\right]^{\otimes n}\right) - \log|E| \tag{152}$$

$$= d^{\mathrm{reg}}\left(\mathcal{I}_E\otimes\mathcal{N}_{A\to B}\middle\|\mathcal{R}_E^\pi\otimes\mathcal{N}_{A\to B}\right) - \log|E|, \tag{153}$$

which completes the proof.

Remark 3 (Operational interpretation)**.** From the operational meaning of d^{reg}, we can understand quantum capacity as the Stein exponent of channel discrimination between the ideal case $\mathcal{I}_E\otimes\mathcal{N}_{A\to B}$ and the worst case $\mathcal{R}_E^\pi\otimes\mathcal{N}_{A\to B}$. Noting that $d^{\mathrm{reg}}(\mathcal{I}_E\|\mathcal{R}_E^\pi) = \log|E|$, we can also write

$$Q(\mathcal{N}) = d^{\mathrm{reg}}(\mathcal{I}_E\otimes\mathcal{N}_{A\to B}\|\mathcal{R}_E^\pi\otimes\mathcal{N}_{A\to B}) - d^{\mathrm{reg}}(\mathcal{I}_E\|\mathcal{R}_E^\pi), \tag{154}$$

indicating that the quantum capacity of a channel $\mathcal{N}$ can be understood as the "power" of this channel as a catalyst to discriminate the perfect channel $\mathcal{I}_E$ and the completely useless channel $\mathcal{R}_E^\pi$ for quantum communication.

Drawing upon the correspondence established in Theorem 11 and the extreme non-additivity of channel coherent information as shown in [66] (where an unbounded number of channel uses may be necessary to detect quantum capacity), we can infer that the unstabilized quantum channel divergence can also exhibit extreme non-additivity.

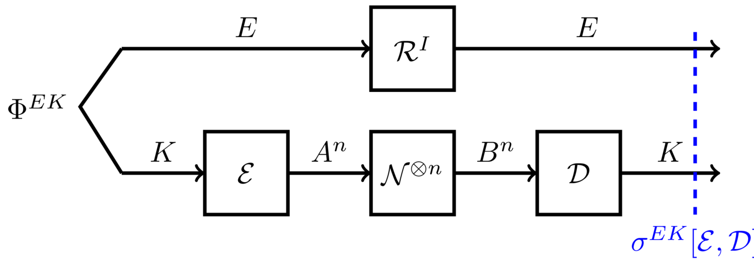

Figure 4 (Color online) Conceptual process that replaces $\mathcal{I}_E$ in Figure 3 with the CP map $\mathcal{R}_E^I$, and gives the final state $\sigma_{EK}[\mathcal{E},\mathcal{D}]$.

Theorem 12. Let d be the unstabilized quantum channel divergence induced by the Umegaki relative entropy. Then d is extremely non-additive. That is, for any $n \in \mathbb{N}$, there exists quantum channels $\mathcal{N},\mathcal{M} \in \mathrm{CPTP}(A \to B)$ such that

$$d(\mathcal{N}^{\otimes n}\|\mathcal{M}^{\otimes n}) < d^{\mathrm{reg}}(\mathcal{N}\|\mathcal{M}). \tag{155}$$

Proof. It has been shown in [66] that for any $n \in \mathbb{N}$, there exists quantum channels $\mathcal{E} \in \mathrm{CPTP}(A \to B)$ such that $I_c(\mathcal{E}^{\otimes n}) = 0 < Q(\mathcal{E})$. Then by the relation in Theorem 11, we can take $\mathcal{N} = \mathcal{I} \otimes \mathcal{E}$ and $\mathcal{M} = \mathcal{R}_E^\pi \otimes \mathcal{E}$. This gives $d(\mathcal{N}^{\otimes n}\|\mathcal{M}^{\otimes n}) = \log|A| < d^{\mathrm{reg}}(\mathcal{N}\|\mathcal{M})$.

5.2 One-shot converse bound for quantum channel capacity

Here we establish a converse bound for one-shot quantum capacity, which can be seen as a smoothed analogue of channel coherent information. The one-shot converse bound is mostly inspired by the channel divergence formula of coherent information (146). That is, the channel coherent information as well as quantum capacity characterize the distinguishability between the channel $\mathcal{I}_E \otimes \mathcal{N}_{A\to B}$ and the CP map $\mathcal{R}_E^I \otimes \mathcal{N}_{A\to B}$ (here we use the identity operator I instead of π to absorb the constant factor $\log|E|$). It is thus convenient for us to consider a conceptual process in Figure 4 which replaces the identity map $\mathcal{I}_E$ in Figure 3 with the CP map $\mathcal{R}_E^I$. Its final state is denoted as

$$\sigma_{EK}[\mathcal{E},\mathcal{D}] := \mathcal{R}_E^I \otimes \mathcal{D}_{B\to K} \circ \mathcal{N}_{A\to B} \circ \mathcal{E}_{K\to A}(\Phi_{EK}) = \mathcal{R}_E^I \otimes \mathcal{I}_K(\rho_{EK}[\mathcal{E},\mathcal{D}]). \tag{156}$$

Theorem 13 (One-shot converse bound)**.** For any $\mathcal{N} \in \mathrm{CPTP}(A \to B)$ and $\varepsilon \in (0,1)$, it holds that

$$Q^{(1)}(\mathcal{N},\varepsilon) \leqslant \sup_{|E|\in\mathbb{N}} d_H^\varepsilon \left(\mathcal{I}_E \otimes \mathcal{N}_{A\to B} \middle\| \mathcal{R}_E^I \otimes \mathcal{N}_{A\to B}\right), \tag{157}$$

where the supremum is taken over E of arbitrary dimension.

Proof. For any entanglement transmission code $\{|K|,\mathcal{E},\mathcal{D}\}$ such that $\mathrm{Tr}\,\Phi_{EK}\rho_{EK}[\mathcal{E},\mathcal{D}] \geqslant 1-\varepsilon$. We have a key observation that

$$\begin{aligned}
\mathrm{Tr}\,\Phi_{EK}\sigma_{EK}[\mathcal{E},\mathcal{D}] &= \mathrm{Tr}\,\Phi_{EK}\left\{\mathcal{R}_E^I \otimes \mathcal{I}_K(\rho_{EK}[\mathcal{E},\mathcal{D}])\right\} &(158)\\
&= \mathrm{Tr}\,\Phi_{EK}\left\{I_E \otimes \mathrm{Tr}_E(\rho_{EK}[\mathcal{E},\mathcal{D}])\right\} &(159)\\
&= \mathrm{Tr}\left\{\mathrm{Tr}_E(\Phi_{EK})\right\}\left\{\mathrm{Tr}_E(\rho_{EK}[\mathcal{E},\mathcal{D}])\right\} &(160)\\
&= \mathrm{Tr}\left\{I_K/|K|\right\}\left\{\mathrm{Tr}_E(\rho_{EK}[\mathcal{E},\mathcal{D}])\right\} &(161)\\
&= 1/|K|, &(162)
\end{aligned}$$

where the first line follows by (156), the second line follows by definition of $\mathcal{R}_E^I$, the third line follows by the identity $\mathrm{Tr}\,X_{AB}(I_A \otimes Y_B) = \mathrm{Tr}\{\mathrm{Tr}_A X_{AB}\}\{Y_B\}$, the last line follows by the fact that $\mathrm{Tr}_E(\rho_{EK}[\mathcal{E},\mathcal{D}])$ is a normalized quantum state. Then we have

$$\begin{aligned}
\log|K| &= -\log \mathrm{Tr}\,\Phi_{EK}\sigma_{EK}[\mathcal{E},\mathcal{D}] &(163)\\
&\leqslant D_H^\varepsilon\left(\rho_{EK}[\mathcal{E},\mathcal{D}]\middle\|\sigma_{EK}[\mathcal{E},\mathcal{D}]\right) &(164)\\
&\leqslant D_H^\varepsilon\left((\mathcal{I}_E \otimes \mathcal{N}_{A\to B}) \circ (\mathcal{E}_{K\to A}(\Phi_{EK}))\middle\|(\mathcal{R}_E^I \otimes \mathcal{N}_{A\to B}) \circ (\mathcal{E}_{K\to A}(\Phi_{EK}))\right) &(165)\\
&\leqslant \sup_{\rho\in\mathfrak{D}(EA)} D_H^\varepsilon\left(\mathcal{I}_E \otimes \mathcal{N}_{A\to B}(\rho_{EA})\middle\|\mathcal{R}_E^I \otimes \mathcal{N}_{A\to B}(\rho_{EA})\right), &(166)
\end{aligned}$$

where the first inequality follows because Φ_{EK} is a particular choice of quantum test for hypothesis testing relative entropy that satisfies $\operatorname{Tr}\Phi_{EK}\rho_{EK}[\mathcal{E},\mathcal{D}] \geqslant 1-\varepsilon$, the second inequality follows by the data-processing inequality of D_H^ε under the action of $\mathcal{I}_E \otimes \mathcal{D}_{B\to K}$, the third inequality follows by relaxing $\mathcal{E}_{K\to A}(\Phi_{EK})$ to all density matrices on system $E\otimes A$. Finally taking supremum over all possible codes $\{|K|,\mathcal{E},\mathcal{D}\}$, we have the desired result.

Corollary 2. For any $\mathcal{N} \in \mathrm{CPTP}(A\to B)$ and $\varepsilon \in (0,1)$, $\delta \in (0,1)$ and $\varepsilon+\delta<1$, it holds that

$$Q^{(1)}(\mathcal{N},\varepsilon) \leqslant d_{\max}^{\delta}\left(\mathcal{I}_E\otimes\mathcal{N}_{A\to B}\middle\|\mathcal{R}_E^{\pi}\otimes\mathcal{N}_{A\to B}\right) + \log\frac{1}{1-\varepsilon-\delta} - \log|E|, \tag{167}$$

where E is isomorphic to A.

Proof. Combining Theorem 13 and (63) we have

$$Q^{(1)}(\mathcal{N},\varepsilon) \leqslant \sup_{\substack{|E|\in\mathbb{N}\\ \rho\in\mathfrak{D}(EA)}} D_{\max}^{\delta}\left(\mathcal{I}_E\otimes\mathcal{N}_{A\to B}(\rho_{EA})\middle\|\mathcal{R}_E^{I}\otimes\mathcal{N}_{A\to B}(\rho_{EA})\right) + \log\frac{1}{1-\varepsilon-\delta}. \tag{168}$$

Since $D_{\max}^{\delta}$ is jointly quasi-convex [57, Lemma 7], the optimization can be restricted to pure states. Furthermore, due to the isometry invariance property of $D_{\max}^{\delta}$, we can, without loss of generality, assume that E is isomorphic to A. Finally, noting that $\mathcal{R}_E^I(\cdot) = |E|\mathcal{R}_E^{\pi}(\cdot)$ and $D_{\max}^{\delta}(\rho\|a\sigma) = D_{\max}^{\delta}(\rho\|\sigma) - \log a$, we have the asserted result.

5.3 Towards the strong converse property of quantum channel capacity

Consider any entanglement transmission code with an achievable triplet $\{r,n,\varepsilon\}$. The strong converse property of channel $\mathcal{N}$ is that if the communication rate $r > Q(\mathcal{N})$, then the communication error ε converges to one as n goes to infinity. Similar to the proof of Theorem 1, this can be equivalently expressed by

$$\limsup_{n\to\infty}\frac{1}{n}Q^{(1)}(\mathcal{N}^{\otimes n},\varepsilon) \leqslant Q(\mathcal{N}), \quad \forall\varepsilon\in(0,1). \tag{169}$$

The strong converse property of quantum capacity is a long-standing open problem in quantum information theory. Upon the connection between quantum communication and channel discrimination, we show that the strong converser property for channel discrimination implies the strong converse property of quantum capacity.

Theorem 14 (Strong converse property)**.** Let $\mathcal{N} \in \mathrm{CPTP}(A\to B)$ be a quantum channel, $\mathcal{I} \in \mathrm{CPTP}(E\to E)$ be the identity channel with E isomorphic to A and $\mathcal{R}_E^{\pi} \in \mathrm{CPTP}(E\to E)$ be the replacer channel. Then if the channels $\mathcal{I}_E\otimes\mathcal{N}_{A\to B}$ and $\mathcal{R}_E^{\pi}\otimes\mathcal{N}_{A\to B}$ exhibit the strong converse property, as defined in Definition 11, this implies the strong converse property of the channel capacity for $\mathcal{N}$.

Proof. For any $\varepsilon\in(0,1)$, let $\delta\in(0,1)$ such that $\varepsilon+\delta<1$. Then

$$\limsup_{n\to\infty}\frac{1}{n}Q^{(1)}(\mathcal{N}^{\otimes n},\varepsilon) \leqslant \limsup_{n\to\infty}\frac{1}{n}d_{\max}^{\delta}\left((\mathcal{I}_E\otimes\mathcal{N}_{A\to B})^{\otimes n}\middle\|(\mathcal{R}_E^{\pi}\otimes\mathcal{N}_{A\to B})^{\otimes n}\right) - \log|E| \tag{170}$$

$$\leqslant d^{\mathrm{reg}}\left(\mathcal{I}_E\otimes\mathcal{N}_{A\to B}\middle\|\mathcal{R}_E^{\pi}\otimes\mathcal{N}_{A\to B}\right) - \log|E| \tag{171}$$

$$= Q(\mathcal{N}), \tag{172}$$

where the first inequality follows from Corollary 2, the second inequality follows from Theorem 2 and the equality follows from Theorem 11. This completes the proof.

6 Discussion

In conclusion, this work advances the understanding of the ultimate limits of quantum channel discrimination and quantum communication by developing versatile tools and frameworks rooted in unstabilized channel divergence. We address key open problems, such as improving bounds on hypothesis testing relative entropy, proving additivity for channel divergences, and establishing a quantum channel analog of Stein's lemma. Our unified approach links channel discrimination strategies with operational regimes and mathematical divergences, providing a comprehensive perspective on quantum channel discrimination across various settings. Furthermore, by framing quantum communication problems as quantum channel discrimination tasks, we uncover connections between channel capacities, channel discrimination, and operational exponents. These results bridge two core areas of quantum information theory and offer new insights for future exploration.

An initial attempt to prove the exponentially strong converse property for two general channels was presented in the first arXiv submission of this work (arXiv:2110.14842v1). However, this effort triggers the discovery of a gap in a technical lemma from [9], which undermines the validity of the original proof and leaves the problem unresolved. Notably, this gap has drawn great attention in the quantum information community since then and the generalized quantum Stein's lemma, originally proposed in [9], has been recently reproved in [67,68]. Given our findings in this work that the strong converse property is a pivotal element for achieving a complete understanding of quantum channel discrimination, we encourage interested readers to give further investigations into this important problem. Several results from our initial analysis remain valid and could hold independent interest. These details are included in Appendix A, and we hope they will inspire and support future efforts to resolve this challenging issue.

Acknowledgements The work of Kun FANG was supported by National Natural Science Foundation of China (Grant Nos. 92470113, 12404569), Shenzhen Science and Technology Program (Grant No. JCYJ20240813113519025), Shenzhen Fundamental Research Program (Grant No. JCYJ20241202124023031), 1+1+1 CUHK-CUHK(SZ)-GDST Joint Collaboration Fund (Grant No. GRDP2025-022), and University Development Fund (Grant No. UDF01003565). The work of Gilad GOUR was supported by Israel Science Foundation (Grant No. 1192/24). The work of Xin WANG was partially supported by National Key R&D Program of China (Grant No. 2024YFE0102500), National Natural Science Foundation of China (Grant No. 12447107), Guangdong Natural Science Foundation (Grant No. 2025A1515012834), Guangdong Provincial Quantum Science Strategic Initiative (Grant Nos. GDZX2403008, GDZX2403001), Guangdong Provincial Key Lab of Integrated Communication, Sensing and Computation for Ubiquitous Internet of Things (Grant No. 2023B1212010007), the Quantum Science Center of Guangdong-Hong Kong-Macao Greater Bay Area, and Education Bureau of Guangzhou Municipality. We thank Marco TOMAMICHEL for pointing out the gap in our initial attempt to prove the exponentially strong converse property in our first submission of this work (arXiv:2110.14842v1).

Supporting information Appendix A. The supporting information is available online at info.scichina.com and link.springer.com. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.

References

1 Chernoff H. A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. Ann Math Statist, 1952, 23: 493–507
2 Wolfowitz J. Coding Theorems of Information Theory. Springer: Berlin, 1978
3 Blahut R. Hypothesis testing and information theory. IEEE Trans Inform Theor, 1974, 20: 405–417
4 Hiai F, Petz D. The proper formula for relative entropy and its asymptotics in quantum probability. Commun Math Phys, 1991, 143: 99–114
5 Nagaoka H, Ogawa T. Strong converse and Stein's lemma in quantum hypothesis testing. IEEE Trans Inform Theor, 2000, 46: 2428–2433
6 Hayashi M. Optimal sequence of quantum measurements in the sense of Steins lemma in quantum hypothesis testing. J Phys A-Math Gen, 2002, 35: 10759–10773
7 Audenaert K M R, Nussbaum M, Szkoła A, et al. Asymptotic error rates in quantum hypothesis testing. Commun Math Phys, 2008, 279: 251–283
8 Hayashi M. Error exponent in asymmetric quantum hypothesis testing and its application to classical-quantum channel coding. Phys Rev A, 2007, 76: 062301
9 Brandão F G S L, Plenio M B. A generalization of quantum Stein's lemma. Commun Math Phys, 2010, 295: 791–828
10 Cooney T, Mosonyi M, Wilde M M. Strong converse exponents for a quantum channel discrimination problem and quantum-feedback-assisted communication. Commun Math Phys, 2016, 344: 797–829
11 Mosonyi M, Ogawa T. Quantum hypothesis testing and the operational interpretation of the quantum Rényi relative entropies. Commun Math Phys, 2015, 334: 1617–1648
12 Wang X, Wilde M M. Resource theory of asymmetric distinguishability. Phys Rev Res, 2019, 1: 033170
13 Wang X, Wilde M M. Resource theory of asymmetric distinguishability for quantum channels. Phys Rev Res, 2019, 1: 033169
14 Fang K, Fawzi H, Fawzi O. Adversarial quantum channel discrimination. 2025. ArXiv:2506.03060
15 Fang K, Fawzi H, Fawzi O. Generalized quantum asymptotic equipartition. 2024. ArXiv:2411.04035
16 Nagaoka H, Hayashi M. An information-spectrum approach to classical and quantum hypothesis testing for simple hypotheses. IEEE Trans Inform Theor, 2007, 53: 534–549
17 Wang L, Renner R. One-shot classical-quantum capacity and hypothesis testing. Phys Rev Lett, 2012, 108: 200501
18 Fang K, Fawzi H, Fawzi O. Efficient approximation of regularized relative entropies and applications. 2025. ArXiv:2502.15659
19 Chiribella G, D'Ariano G M, Perinotti P. Memory effects in quantum channel discrimination. Phys Rev Lett, 2008, 101: 180501
20 Hayashi M. Discrimination of two channels by adaptive methods and its application to quantum system. IEEE Trans Inform Theor, 2009, 55: 3807–3820
21 Piani M, Watrous J. All entangled states are useful for channel discrimination. Phys Rev Lett, 2009, 102: 250501
22 Yuan H, Fung C H F. Fidelity and Fisher information on quantum channels. New J Phys, 2017, 19: 113039
23 Wilde M M, Berta M, Hirche C, et al. Amortized channel divergence for asymptotic quantum channel discrimination. Lett Math Phys, 2020, 110: 2277–2336
24 Fang K, Fawzi O, Renner R, et al. Chain rule for the quantum relative entropy. Phys Rev Lett, 2020, 124: 100501
25 Pirandola S, Laurenza R, Lupo C, et al. Fundamental limits to quantum channel discrimination. npj Quantum Inf, 2019, 5: 1–8
26 Wang X, Fang K, Duan R. Semidefinite programming converse bounds for quantum communication. IEEE Trans Inform Theor, 2019, 65: 2583–2592

27 Wang X, Fang K, Tomamichel M. On converse bounds for classical communication over quantum channels. IEEE Trans Inform Theor, 2019, 65: 4609–4619
28 Fang K, Wang X, Tomamichel M, et al. Quantum channel simulation and the channel's smooth max-information. IEEE Trans Inform Theor, 2020, 66: 2129–2140
29 Fang K, Fawzi H. Geometric Rényi divergence and its applications in quantum channel capacities. Commun Math Phys, 2021, 384: 1615–1677
30 Pirandola S, Bardhan B R, Gehring T, et al. Advances in photonic quantum sensing. Nat Photon, 2018, 12: 724–733
31 Braun D, Adesso G, Benatti F, et al. Quantum-enhanced measurements without entanglement. Rev Mod Phys, 2018, 90: 35006
32 Degen C L, Reinhard F, Cappellaro P. Quantum sensing. Rev Mod Phys, 2017, 89: 35002
33 Nuradha T, Wilde M M. Fidelity-based smooth min-relative entropy: properties and applications. IEEE Trans Inform Theor, 2024, 70: 4170–4196
34 Chitambar E, Gour G. Quantum resource theories. Rev Mod Phys, 2019, 91: 025001
35 Gour G, Winter A. How to quantify a dynamical quantum resource. Phys Rev Lett, 2019, 123: 150401
36 Fang K, Liu Z W. No-go theorems for quantum resource purification: new approach and channel theory. PRX Quantum, 2022, 3: 010337
37 Regula B, Takagi R. Fundamental limitations on distillation of quantum channel resources. Nat Commun, 2021, 12: 1–2
38 Berta M, Hirche C, Kaur E, et al. Amortized channel divergence for asymptotic quantum channel discrimination. 2018. ArXiv:1808.01498
39 Díaz M G, Fang K, Wang X, et al. Using and reusing coherence to realize quantum processes. Quantum, 2018, 2: 100
40 Saxena G, Chitambar E, Gour G. Dynamical resource theory of quantum coherence. Phys Rev Res, 2020, 2: 23298
41 Wang X, Wilde M M, Su Y. Quantifying the magic of quantum channels. New J Phys, 2019, 21: 103002
42 Faist P, Berta M, Brandão F. Thermodynamic capacity of quantum processes. Phys Rev Lett, 2019, 122: 200601
43 Umegaki H. Conditional expectation in an operator algebra. Tohoku Math J Second Series, 1954, 6: 177–181
44 Petz D. Quasi-entropies for finite quantum systems. Rep Math Phys, 1986, 23: 57–65
45 Müller-Lennert M, Dupuis F, Szehr O, et al. On quantum Rényi entropies: a new generalization and some properties. J Math Phys, 2013, 54: 122203
46 Wilde M M, Winter A, Yang D. Strong converse for the classical capacity of entanglement-breaking and Hadamard channels via a sandwiched Rényi relative entropy. Commun Math Phys, 2014, 331: 593–622
47 Datta N. Min- and max-relative entropies and a new entanglement monotone. IEEE Trans Inform Theor, 2009, 55: 2816–2826
48 Renner R. Security of QKD. Dissertation for Ph.D. Degree. Zurich: ETH Zurich, 2005
49 Qi H, Wang Q, Wilde M M. Applications of position-based coding to classical communication over quantum channels. J Phys A-Math Theor, 2018, 51: 444002
50 Buscemi F, Gour G. Quantum relative Lorenz curves. Phys Rev A, 2017, 95: 012110
51 Theurer T, Fang K, Gour G. Single-shot entanglement manipulation of states and channels revisited. IEEE Trans Inform Theor, 2025, 71: 5330–5360
52 George I, Lin J, van Himbeeck T, et al. Finite-key analysis of quantum key distribution with characterized devices using entropy accumulation. 2022. ArXiv:2203.06554
53 Leditzky F, Kaur E, Datta N, et al. Approaches for approximate additivity of the Holevo information of quantum channels. Phys Rev A, 2018, 97: 012332
54 Hastings M B. Superadditivity of communication capacity using entangled inputs. Nat Phys, 2009, 5: 255–257
55 Fawzi H, Fawzi O. Defining quantum divergences via convex optimization. Quantum, 2021, 5: 387
56 Dupuis F, Krämer L, Faist P, et al. Generalized entropies. 2012. ArXiv:1211.3141
57 Datta N, Mosonyi M, Hsieh M H, et al. A smooth entropy approach to quantum hypothesis testing and the classical capacity of quantum channels. IEEE Trans Inform Theor, 2013, 59: 8014–8026
58 Tomamichel M, Hayashi M. A hierarchy of information quantities for finite block length analysis of quantum tasks. IEEE Trans Inform Theor, 2013, 59: 7693–7710
59 Datta N, Leditzky F. Second-order asymptotics for source coding, dense coding, and pure-state entanglement conversions. IEEE Trans Inform Theor, 2014, 61: 582–608
60 Rubboli R, Tomamichel M. Fundamental limits on correlated catalytic state transformations. Phys Rev Lett, 2022, 129: 120506
61 Nagaoka H. The converse part of the theorem for quantum Hoeffding bound. 2006. ArXiv:quant-ph/0611289
62 Fang K, Zhao J T, Li X F, et al. Quantum NETwork: from theory to practice. Sci China Inf Sci, 2023, 66: 180509
63 Lloyd S. Capacity of the noisy quantum channel. Phys Rev A, 1997, 55: 1613–1622
64 Shor P W. The quantum channel capacity and coherent information. In: Proceedings of MSRI Workshop on Quantum Computation, 2002
65 Devetak I. The private classical capacity and quantum capacity of a quantum channel. IEEE Trans Inform Theor, 2005, 51: 44–55
66 Cubitt T, Elkouss D, Matthews W, et al. Unbounded number of channel uses may be required to detect quantum capacity. Nat Commun, 2015, 6: 6739
67 Hayashi M, Yamasaki H. Generalized quantum Stein's lemma and second law of quantum resource theories. 2024. ArXiv:2408.02722
68 Lami L. A solution of the generalised quantum Stein's lemma. 2024. ArXiv:2408.06410