This article has been accepted for publication in IEEE Transactions on Information Theory. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TIT.2026.3661603

Efficient Approximation of Regularized Relative Entropies and Applications

Kun Fang, *Member, IEEE,* Hamza Fawzi, and Omar Fawzi

***Abstract*—The quantum relative entropy is a fundamental quantity in quantum information science, characterizing the distinguishability between two quantum states. However, this quantity is not additive in general for correlated quantum states, necessitating regularization for precise characterization of the operational tasks of interest. Recently, we proposed the study of the regularized relative entropy between two sequences of sets of quantum states in [arXiv: 2411.04035], which captures a general framework for a wide range of quantum information tasks. Here, we show that given suitable structural assumptions and efficient descriptions of the sets, the regularized relative entropy can be efficiently approximated within an additive error by a quantum relative entropy program of polynomial size. This applies in particular to the regularized relative entropy in adversarial quantum channel discrimination. Moreover, we apply the idea of efficient approximation to quantum resource theories. In particular, when the set of interest does not directly satisfy the required structural assumptions, it can be relaxed to one that does. This provides improved and efficient bounds for the entanglement cost of quantum states and channels, entanglement distillation and magic state distillation. Numerical results demonstrate improvements even for the first level of approximation.**

***Index Terms*—Quantum relative entropy, regularization, semidefinite programming, symmetry reduction, entanglement theory, quantum resource theories.**

I. Introduction

The classical relative entropy, also known as the Kullback-Leibler (KL) divergence [1], is a measure of how much a model probability distribution is different from a true probability distribution. It plays a pivotal role in classical information processing and finds applications in diverse domains including machine learning [2], data compression [3], [4] and statistical mechanics [5], [6]. With the development of quantum information science, the quantum relative entropy has been proposed as a quantum generalization of the KL divergence [7],

$$D(\rho\|\sigma) := \operatorname{Tr}[\rho(\log\rho - \log\sigma)], \tag{1}$$

quantifying the distinguishability between quantum states ρ and σ [8]. It has found widespread applications in various fields [9], [10], including quantum machine learning [11], quantum channel coding [12], quantum error correction [13], quantum resource theories [14], and quantum cryptography [15].

A useful property for quantum relative entropy is the additivity between two tensor product states, i.e., $D(\rho_1\otimes\rho_2\|\sigma_1\otimes\sigma_2) = D(\rho_1\|\sigma_1)+D(\rho_2\|\sigma_2)$. However, this property does not hold for general correlated quantum states, i.e., $D(\rho_{12}\|\sigma_{12}) \neq D(\rho_1\|\sigma_1)+D(\rho_2\|\sigma_2)$, necessitating regularization for precise characterization of operational tasks of interest. Recently, we proposed in [16] the study of the regularized quantum relative entropy between two sequences of sets of quantum states $\mathscr{A}_n$, $\mathscr{B}_n$ acting on $\mathcal{H}^{\otimes n}$ for some Hilbert space $\mathcal{H}$:

$$D^\infty(\mathscr{A}\|\mathscr{B}) := \lim_{n\to\infty}\frac{1}{n}D(\mathscr{A}_n\|\mathscr{B}_n), \tag{2}$$

where $D(\mathscr{A}_n\|\mathscr{B}_n) = \inf_{\rho\in\mathscr{A}_n,\sigma\in\mathscr{B}_n} D(\rho\|\sigma)$. This quantity captures a general framework for a wide range of quantum information tasks. This includes quantum hypothesis testing [8], [17], [18], [19], [20], [21], [22], [23] and quantum channel discrimination [19], [24], [25], [26] from a foundational perspective, entropy accumulation theorems for quantum cryptography [27], [28], and quantum resource distillation [18], [29], [30], [31], [32], [33], [34], [35], [36], [37], [38] and preparation [39], [40], [41], [42], which are crucial for quantum computing and quantum networking [43].

In general, computing $D^\infty(\mathscr{A}\|\mathscr{B})$ is challenging due to the limit, which we refer to as "regularization". One notable example is given by the regularized relative entropy of entanglement

$$D^\infty(\rho_{AB}\|\mathrm{SEP}) := \lim_{n\to\infty}\frac{1}{n}D(\rho_{AB}^{\otimes n}\|\mathrm{SEP}(A^n:B^n)), \tag{3}$$

where $\mathrm{SEP}(A^n:B^n)$ denotes the set of all separable states between Hilbert spaces $\mathcal{H}_A^{\otimes n}$ and $\mathcal{H}_B^{\otimes n}$. This quantity uniquely determines the ultimate limits of entanglement manipulation and serves as the key quantity in understanding the second law of quantum entanglement [18], [22], [23]. However, evaluating this quantity is extremely hard in general, as it involves the regularization as well as the separability problem [44].

In this work, we show in Theorem 10 that given suitable structural assumptions (in particular the stability of the polar

K.F. is supported in part by the National Natural Science Foundation of China (Grant No. 92470113 and 12404569), the Shenzhen Science and Technology Program (Grant No. JCYJ20240813113519025), the Shenzhen Fundamental Research Program (Grant No. JCYJ20241202124023031), the General R&D Projects of 1+1+1 CUHK-CUHK(SZ)-GDST Joint Collaboration Fund (Grant No. GRDP2025-022), and the University Development Fund (Grant No. UDF01003565). H.F. was partially funded by UK Research and Innovation (UKRI) under the UK government's Horizon Europe funding guarantee EP/X032051/1. O.F. acknowledges support by the European Research Council (ERC Grant AlgoQIP, Agreement No. 851716), by the European Union's Horizon 2020 research and innovation programme under Grant Agreement No 101017733 (VERIqTAS) and by the Agence Nationale de la Recherche under the Plan France 2030 with the reference ANR-22-PETQ-0009. *(Corresponding author: Kun Fang.)*

Kun Fang is with the School of Data Science, The Chinese University of Hong Kong, Shenzhen, Guangdong 518172, China (e-mail: kunfang@cuhk.edu.cn).

Hamza Fawzi is with the Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge CB3 0WA, United Kingdom (e-mail: h.fawzi@damtp.cam.ac.uk).

Omar Fawzi is with the Univ Lyon, Inria, ENS Lyon, UCBL, LIP, Lyon, France (e-mail: omar.fawzi@ens-lyon.fr).

set under tensor product) and efficient descriptions of the sets, the regularized relative entropy $D^\infty(\mathscr{A}\|\mathscr{B})$ can be efficiently approximated within an additive error by a quantum relative entropy program of polynomial size. We then apply this result in Section IV to several problems in quantum information theory. The first application (Section IV-A) is to compute the regularized relative entropy between the image sets of two quantum channels, which characterizes the optimal exponent in adversarial channel discrimination [16]. For this problem, the relevant sets satisfy the structural assumptions. In the following applications, this will not be the case, but the sets can be relaxed to sets that do satisfy the requirements. For instance, in entanglement theory, the set of separable states can be relaxed to the Rains set [29], [45], which satisfies all necessary assumptions. We illustrate this by obtaining bounds (Section IV-B) on the entanglement cost of quantum states and channels improving on [39], [40], [41], [42]. Numerical results demonstrate improvements even for the first level of approximation. This approach can also be applied to obtain improved bounds on entanglement distillation [29], [45], [18], [22], [23] as discussed in Section IV-C. Similarly, in fault-tolerant quantum computing, the set of stabilizer states can be relaxed to the set of states with non-positive mana [33], which also fulfills the required conditions. As described in Section IV-D, this can be used to obtain improved bounds for magic state distillation.

Generally, our result can be applied by verifying the conditions of the relevant theory and performing necessary relaxations when required. Therefore, we anticipate that this approach has the potential for other applications beyond the specific cases discussed here.

II. Preliminaries

A. Notation

In this section we set the notation and define several quantities that will be used throughout this work. Some frequently used notation are summarized in Table I. Note that we label different physical systems by capital Latin letters and use these labels as subscripts to guide the reader by indicating which system a mathematical object belongs to. We drop the subscripts when they are evident in the context of an expression (or if we are not talking about a specific system).

Notation	Descriptions
$\mathcal{H}_A$	Hilbert space on system A
$\mathscr{L}(A)$	Linear operators on $\mathcal{H}_A$
$\mathscr{H}(A)$	Hermitian operators on $\mathcal{H}_A$
$\mathscr{H}_+(A)$	Positive semidefinite operators on $\mathcal{H}_A$
$\mathscr{H}_{++}(A)$	Positive definite operators on $\mathcal{H}_A$
$\mathscr{D}(A)$	Density matrices on $\mathcal{H}_A$
$\mathscr{A},\mathscr{B},\mathscr{C}$	Set of linear operators
$\mathscr{C}^\circ$	Polar set $\mathscr{C}^\circ := \{X : \mathrm{Tr}[XY] \leq 1, \forall Y \in \mathscr{C}\}$ of $\mathscr{C}$
$\mathscr{C}^\circ_+$	Polar set restricted to positive semidefinite cone $\mathscr{C}^\circ \cap \mathscr{H}_+$
$\mathscr{C}^\circ_{++}$	Polar set restricted to positive definite operators $\mathscr{C}^\circ \cap \mathscr{H}_{++}$
CPTP	Completely positive and trace preserving maps
CP	Completely positive maps
$\log(x)$	Logarithm of x in base two

TABLE I: Overview of notational conventions.

B. Polar set and support function

In the following, we introduce the definitions of the polar set and support function, along with a fundamental result that will be used in our discussions.

Definition 1 (Polar set and support function)**.** *Let $\mathscr{C} \subseteq \mathscr{E}$ be a convex set in some Euclidean space $\mathscr{E}$ with inner product $\langle\cdot,\cdot\rangle$. The polar set of $\mathscr{C}$ is defined by*

$$\mathscr{C}^\circ := \{X \in \mathscr{E} : \langle X, Y\rangle \leq 1, \forall Y \in \mathscr{C}\}. \tag{4}$$

The support function of $\mathscr{C}$ at ω is defined by $h_{\mathscr{C}}(\omega) = \sup_{\sigma\in\mathscr{C}}\langle\sigma,\omega\rangle$.

When the Euclidean space $\mathscr{E}$ is the space of Hermitian matrices endowed with the trace inner product, we will use the notation $\mathscr{C}^\circ_+ := \mathscr{C}^\circ \cap \mathscr{H}_+$ and $\mathscr{C}^\circ_{++} := \mathscr{C}^\circ \cap \mathscr{H}_{++}$ for the intersections of $\mathscr{C}^\circ$ with positive semidefinite operators and positive definite operators, respectively.

It is clear from the definitions that $\omega \in \mathscr{C}^\circ$ if and only if $h_{\mathscr{C}}(\omega) \leq 1$.

Definition 2. *Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be finite-dimensional Hilbert spaces. Consider three sets $\mathscr{A}_1 \subseteq \mathscr{H}_+(\mathcal{H}_1)$, $\mathscr{A}_2 \subseteq \mathscr{H}_+(\mathcal{H}_2)$, and $\mathscr{A}_{12} \subseteq \mathscr{H}_+(\mathcal{H}_1 \otimes \mathcal{H}_2)$. We say $\{\mathscr{A}_1, \mathscr{A}_2, \mathscr{A}_{12}\}$ is closed under tensor product if for any $X_1 \in \mathscr{A}_1$, $X_2 \in \mathscr{A}_2$, we have $X_1 \otimes X_2 \in \mathscr{A}_{12}$. In short, we write $\mathscr{A}_1 \otimes \mathscr{A}_2 \subseteq \mathscr{A}_{12}$.*

The following lemma [16, Lemma 8] provides an equivalent condition for determining if the polar sets of interest are closed under tensor product, which can be easier to validate for specific examples.

Lemma 3. *Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be finite-dimensional Hilbert spaces. Consider three sets $\mathscr{A}_1 \subseteq \mathscr{H}_+(\mathcal{H}_1)$, $\mathscr{A}_2 \subseteq \mathscr{H}_+(\mathcal{H}_2)$, and $\mathscr{A}_{12} \subseteq \mathscr{H}_+(\mathcal{H}_1 \otimes \mathcal{H}_2)$. Their polar sets are closed under tensor product, i.e., $(\mathscr{A}_1)^\circ_+ \otimes (\mathscr{A}_2)^\circ_+ \subseteq (\mathscr{A}_{12})^\circ_+$ if and only if their support functions are sub-multiplicative, i.e., $h_{\mathscr{A}_{12}}(X_1 \otimes X_2) \leq h_{\mathscr{A}_1}(X_1) h_{\mathscr{A}_2}(X_2)$, $\forall X_i \in \mathscr{H}_+(\mathcal{H}_i)$.*

C. Quantum divergences

A functional $\mathbb{D} : \mathscr{D} \times \mathscr{H}_+ \to \mathbb{R}$ is a quantum divergence if it satisfies the data-processing inequality $\mathbb{D}(\mathcal{E}(\rho)\|\mathcal{E}(\sigma)) \leq \mathbb{D}(\rho\|\sigma)$ for any CPTP map $\mathcal{E}$. In the following, we will introduce several quantum divergences and their fundamental properties, which will be used throughout this work. Additionally, we will define quantum divergences between two sets of quantum states, which will be the main quantity of interest in this work.

Definition 4 (Umegaki relative entropy [46])**.** *For any $\rho \in \mathscr{D}$ and $\sigma \in \mathscr{H}_+$, the Umegaki relative entropy is defined by*

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

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

The min-relative entropy is defined by

$$D_{\min}(\rho\|\sigma) := -\log \mathrm{Tr}[\Pi_\rho\sigma], \tag{6}$$

with Π_ρ the projection on the support of ρ. The max-relative entropy is defined by [47],

$$D_{\max}(\rho\|\sigma) := \log\inf\left\{t \in \mathbb{R} \;:\; \rho \leq t\sigma\right\}, \tag{7}$$

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

Definition 5 (Measured relative entropy [48], [8])**.** *For any $\rho \in \mathscr{D}$, $\sigma \in \mathscr{H}_+$, the measured relative entropy is defined by*

$$D_{\mathrm{M}}(\rho\|\sigma) := \sup_{(\mathcal{X},M)} D(P_{\rho,M}\|P_{\sigma,M}), \tag{8}$$

where D is the Kullback–Leibler divergence and the optimization is over finite sets $\mathcal{X}$ and positive operator valued measures M on $\mathcal{X}$ such that $M_x \geq 0$ and $\sum_{x\in\mathcal{X}} M_x = I$, $P_{\rho,M}$ is a measure on $\mathcal{X}$ defined via the relation $P_{\rho,M}(x) = \operatorname{Tr}[M_x\rho]$ for any $x \in \mathcal{X}$.

A variational expression for D_{M} is that [49, Lemma 1],

$$D_{\mathrm{M}}(\rho\|\sigma) = \sup_{\omega\in\mathscr{H}_{++}} \operatorname{Tr}[\rho\log\omega] + 1 - \operatorname{Tr}[\sigma\omega]. \tag{9}$$

Definition 6 (Measured Rényi divergence [49])**.** *Let $\alpha \in (0,1)\cup(1,\infty)$. For any $\rho \in \mathscr{D}$ and $\sigma \in \mathscr{H}_+$, the measured Rényi divergence is defined as*

$$D_{\mathrm{M},\alpha}(\rho\|\sigma) := \sup_{(\mathcal{X},M)} D_\alpha(P_{\rho,M}\|P_{\sigma,M}), \tag{10}$$

where D_α is the classical Rényi divergence.

The following result shows the ordering relation among different relative entropies.

Lemma 7. *Let $\alpha \in [1/2,1)$. For any $\rho \in \mathscr{D}$ and $\sigma \in \mathscr{H}_+$,*

$$D_{\min}(\rho\|\sigma) \leq D_{\mathrm{M},\alpha}(\rho\|\sigma) \leq D_{\mathrm{M}}(\rho\|\sigma) \leq D(\rho\|\sigma). \tag{11}$$

Proof. The last two inequalities follow from the monotonicity in α of the classical Rényi divergences and the data processing inequality for D. As $D_{\mathrm{M},\alpha}$ is monotone increasing in α, it remains to show the first inequality for $\alpha = \frac{1}{2}$. By the variational formula in [49, Eq. (21)], we have

$$D_{\mathrm{M},1/2}(\rho\|\sigma) = -\log\inf_{\omega\in\mathscr{H}_{++}} \operatorname{Tr}[\rho\omega^{-1}]\operatorname{Tr}[\sigma\omega], \tag{12}$$

which is the same as the Alberti's theorem for quantum fidelity (see e.g. [50, Corollary 3.20]). Consider a feasible solution $\omega_\varepsilon = \Pi_\rho + \varepsilon(I-\Pi_\rho) \in \mathscr{H}_{++}$ with $\varepsilon > 0$. It gives $D_{\mathrm{M},1/2}(\rho\|\sigma) \geq -\log\operatorname{Tr}[\rho\omega_\varepsilon^{-1}]\operatorname{Tr}[\sigma\omega_\varepsilon]$. Since ρ has trace one, it gives $\operatorname{Tr}[\rho\omega_\varepsilon^{-1}] = \operatorname{Tr}[\rho] = 1$. Then we have

$$\begin{aligned} D_{\mathrm{M},1/2}(\rho\|\sigma) &\geq -\log\operatorname{Tr}[\sigma\omega_\varepsilon] &(13)\\ &= -\log[(1-\varepsilon)\operatorname{Tr}\Pi_\rho\sigma + \varepsilon]. &(14)\end{aligned}$$

As the above holds for any $\varepsilon > 0$, we take $\varepsilon \to 0^+$ and get $D_{\mathrm{M},1/2}(\rho\|\sigma) \geq -\log\operatorname{Tr}[\Pi_\rho\sigma] = D_{\min}(\rho\|\sigma)$, which completes the proof. □

In this work, we will focus on the study of quantum divergences between two sets of quantum states.

Definition 8 (Quantum divergence between two sets of states)**.** *Let $\mathbb{D}$ be a quantum divergence between two quantum states. For any sets $\mathscr{A} \subseteq \mathscr{D}$ and $\mathscr{B} \subseteq \mathscr{H}_+$, the quantum divergence between these sets is defined by*

$$\mathbb{D}(\mathscr{A}\|\mathscr{B}) := \inf_{\substack{\rho\in\mathscr{A}\\ \sigma\in\mathscr{B}}} \mathbb{D}(\rho\|\sigma). \tag{15}$$

From the geometric perspective, this quantity characterizes the distance between two sets $\mathscr{A}$ and $\mathscr{B}$ under the "distance metric" $\mathbb{D}$. In particular, if $\mathscr{A} = \{\rho\}$ is a singleton, we write $\mathbb{D}(\rho\|\mathscr{B}) := \mathbb{D}(\{\rho\}\|\mathscr{B})$. For two sequences of sets $\{\mathscr{A}_n\}_{n\in\mathbb{N}}$ and $\{\mathscr{B}_n\}_{n\in\mathbb{N}}$, the regularized divergence is defined by

$$\mathbb{D}^\infty(\mathscr{A}\|\mathscr{B}) := \lim_{n\to\infty}\frac{1}{n}\mathbb{D}(\mathscr{A}_n\|\mathscr{B}_n), \tag{16}$$

whenever the limit on the right-hand side exists.

In particular, we will focus on the sequences of sets satisfying the following assumptions.

Assumption 9. *Let $\mathcal{H}$ be a finite-dimensional Hilbert space. Consider a family of sets $\{\mathscr{A}_n\}_{n\in\mathbb{N}}$, with each $\mathscr{A}_n \subseteq \mathscr{H}_+(\mathcal{H}^{\otimes n})$, satisfying the following properties:*

- *(A.1) Each $\mathscr{A}_n$ is convex and compact;*
- *(A.2) Each $\mathscr{A}_n$ is permutation-invariant, i.e., $P_nX_nP_n^\dagger \in \mathscr{A}_n$ for any $X_n \in \mathscr{A}_n$ and any permutation operator P_n on $\mathcal{H}^{\otimes n}$ that permutes the tensor factors;*
- *(A.3) $\mathscr{A}_m \otimes \mathscr{A}_k \subseteq \mathscr{A}_{m+k}$, for all $m,k \in \mathbb{N}$;*
- *(A.4) $(\mathscr{A}_m)_+^\circ \otimes (\mathscr{A}_k)_+^\circ \subseteq (\mathscr{A}_{m+k})_+^\circ$, for all $m,k \in \mathbb{N}$.*

III. Efficient approximation of regularized relative entropies

In this section, we demonstrate that the regularized relative entropy $D^\infty(\mathscr{A}\|\mathscr{B})$ can be efficiently approximated given efficient descriptions of $\{\mathscr{A}_n\}_{n\in\mathbb{N}}$ and $\{\mathscr{B}_n\}_{n\in\mathbb{N}}$. The main result is presented below, with some technical definitions provided later.

Theorem 10 (Efficient approximation of regularized relative entropies)**.** *Let $\mathcal{H}$ be a Hilbert space of finite dimension d. Let $\{\mathscr{A}_n\}_{n\in\mathbb{N}}$ and $\{\mathscr{B}_n\}_{n\in\mathbb{N}}$ be two sequences of sets satisfying Assumption 9 and $\mathscr{A}_n \subseteq \mathscr{D}(\mathcal{H}^{\otimes n})$, $\mathscr{B}_n \subseteq \mathscr{H}_+(\mathcal{H}^{\otimes n})$ and $D_{\max}(\mathscr{A}_n\|\mathscr{B}_n) \leq cn$, for all $n \in \mathbb{N}$ and a constant $c \in \mathbb{R}_+$. If each $\mathscr{A}_m$ and $\mathscr{B}_m$ have semidefinite (SDP) representations of size s^{2m} satisfying the symmetry conditions in Lemma 19, then both $D(\mathscr{A}_m\|\mathscr{B}_m)$ and $D_{\mathrm{M}}(\mathscr{A}_m\|\mathscr{B}_m)$ can be computed by quantum relative entropy programs of size $O((m+1)^{k^2})$ with $k = \max\{s,d\}$.*

As a result, $D^\infty(\mathscr{A}\|\mathscr{B})$ can be approximated within additive error δ by a quantum relative entropy program of size $O((m_0+1)^{k^2})$ with $m_0 = \lceil\frac{8d^2}{\delta}\log\frac{d^2}{\delta}\rceil$.

Note that quantum relative entropy programs can be efficiently solved using interior point methods, as shown in [51], [52]. Moreover, a numerical toolkit named QICS (Quantum Information Conic Solver) has been developed in [53] to solve these programs efficiently in practice.

The remaining part of this section will provide a proof of the above result. We begin by defining quantum relative entropy programs and then show that the converging bounds for $D^\infty(\mathscr{A}\|\mathscr{B})$, derived from [16] under structural assumptions on the sets $\mathscr{A}_n$ and $\mathscr{B}_n$, allow for an approximation of $D^\infty(\mathscr{A}\|\mathscr{B})$ using quantum relative entropy programs. We then exploit the symmetry in these sets to reduce the size of the convex programs.

A. Quantum relative entropy programs

Definition 11 (Conic program)**.** *A conic program over a convex cone $K \subseteq \mathscr{E}$ is an optimization problem of the form $\min\{\langle c, x\rangle : x \in K,\ F(x) = g\}$ where $F : \mathscr{E} \to \mathscr{F}$ is a linear map and $g \in \mathscr{F}$ [54, Chapter 2]. The size of the program is defined as $\dim(K)$.*

Motivated by the conic program, we introduce the conic representation of a convex set [55].

Definition 12 (Conic representation)**.** *A convex set $\mathscr{C} \subseteq \mathscr{E}$ has a conic representation over K if it can be written as $\mathscr{C} = \{\Pi(x) : x \in K,\ F(x) = g\}$ where $\Pi : \operatorname{linspan}(K) \to \mathscr{E}$ and $F : \operatorname{linspan}(K) \to \mathscr{F}$ are linear maps and $g \in \mathscr{F}$. The size of the representation is $\dim(K)$.*

Clearly, if a convex set $\mathscr{C}$ has a conic representation over K, then any linear optimization problem over $\mathscr{C}$ can be expressed as a conic program over K. We recall below the standard definition of a semidefinite representation of a convex set [55, Definition 1].

Definition 13 (SDP representation)**.** *A convex set $\mathscr{C} \subseteq \mathscr{E}$ has a SDP representation of size[1] s^2 if it can be written as $\mathscr{C} = \{\Pi(x) : x \in \mathscr{H}_+(\mathbb{C}^s),\ F(x) = g\}$, where $\Pi : \mathscr{H}(\mathbb{C}^s) \to \mathscr{E}$ and $F : \mathscr{H}(\mathbb{C}^s) \to \mathscr{F}$ are linear maps and $g \in \mathscr{F}$.*

Definition 14 (Quantum relative entropy program)**.** *A quantum relative entropy program is a conic program over a Cartesian product of positive semidefinite cones, quantum relative entropy cones, and operator relative entropy cones. The quantum relative entropy cone is defined as [56], [52], [53]: $\mathcal{K}_{qre}(\mathbb{C}^n) := cl\{(A, B, t) \in \mathscr{H}_+(\mathbb{C}^n) \times \mathscr{H}_+(\mathbb{C}^n) \times \mathbb{R} : A \ll B,\ D(A\|B) \leq t\}$, where $A \ll B$ indicates that the support of A is included in the support of B. The operator relative entropy cone is defined as: $\mathcal{K}_{ore}(\mathbb{C}^n) := cl\{(A, B, T) \in \mathscr{H}_{++}(\mathbb{C}^n) \times \mathscr{H}_{++}(\mathbb{C}^n) \times \mathscr{H}(\mathbb{C}^n) : A \ll B,\ D_{\rm op}(A\|B) \leq T\}$, with $D_{\rm op}(A\|B) := A^{1/2}\log(A^{1/2}B^{-1}A^{1/2})A^{1/2}$ being the operator relative entropy.*

The following well-known proposition will be useful later, see e.g., [57, Lemma 4.1.8]. We include a proof sketch for convenience.

Lemma 15. *Let $\mathscr{C}$ be a convex set in some Euclidean space $\mathscr{E}$. If $\mathscr{C}$ has an SDP representation of size s^2, then the epigraph of its support function*

$$\mathbf{epi}(h_{\mathscr{C}}) := \{(w, t) \in \mathscr{E} \times \mathbb{R} : h_{\mathscr{C}}(w) \leq t\}. \tag{17}$$

has an SDP representation of size $s^2 + 1$.

[1] We adopt this convention for consistency with Definition 12. This convention is different from existing conventions in the literature where the size of the representation would be s instead of s^2.

Proof sketch. Assume $\mathscr{C}$ has an SDP representation as in Definition 13 of size s^2. Given $w \in \mathscr{E}$ we have

$$\begin{aligned} h_{\mathscr{C}}(w) &:= \sup_{z\in\mathscr{C}} \langle w, z\rangle & (18)\\ &= \sup\{\langle w, \Pi(Y)\rangle : Y \in \mathscr{H}_+(\mathbb{C}^s), F(Y) = g\} & (19)\\ &= \inf_{\lambda\in\mathscr{F}} \{\langle \lambda, g\rangle : F^*(\lambda) - \Pi^*(w) \geq 0\}, & (20)\end{aligned}$$

where F^* and Π^* denote respectively the adjoint maps of $F : \mathscr{H}(\mathbb{C}^s) \to \mathscr{F}$ and $\Pi : \mathscr{H}(\mathbb{C}^s) \to \mathscr{E}$, and the last step follows from SDP duality (assuming that the SDP in (19) is strictly feasible). This shows that $\mathbf{epi}(h_{\mathscr{C}}) = \{(w, t) \in \mathscr{E} \times \mathbb{R} : \exists \lambda \in \mathscr{F}$ s.t. $\langle \lambda, g\rangle \leq t$ and $F^*(\lambda) - \Pi^*(w) \in \mathscr{H}_+(\mathbb{C}^s)\}$, which is a semidefinite representation of $\mathbf{epi}(h_{\mathscr{C}})$ of size $s^2 + 1$. ☐

B. Efficient approximation by symmetry reduction

With the above definitions, we now show that the converging bounds for $D^\infty(\mathscr{A}\|\mathscr{B})$, derived from [16] under structural assumptions on the sets $\mathscr{A}_n$ and $\mathscr{B}_n$, allow for an efficient approximation of $D^\infty(\mathscr{A}\|\mathscr{B})$ using quantum relative entropy programs. We first recall the following lemma from [16].

Lemma 16. *[16, Lemma 28, 29] Let $\mathcal{H}$ be a Hilbert space of finite dimension d. Let $\{\mathscr{A}_n\}_{n\in\mathbb{N}}$ and $\{\mathscr{B}_n\}_{n\in\mathbb{N}}$ be two sequences of sets satisfying Assumption 9 and $\mathscr{A}_n \subseteq \mathscr{D}(\mathcal{H}^{\otimes n})$, $\mathscr{B}_n \subseteq \mathscr{H}_+(\mathcal{H}^{\otimes n})$ and $D_{\max}(\mathscr{A}_n\|\mathscr{B}_n) \leq cn$, for all $n \in \mathbb{N}$ and a constant $c \in \mathbb{R}_+$. Then the regularized relative entropy $D^\infty(\mathscr{A}\|\mathscr{B})$ can be estimated using the following bounds:*

$$\frac{1}{m} D_{\rm M}(\mathscr{A}_m\|\mathscr{B}_m) \leq D^\infty(\mathscr{A}\|\mathscr{B}) \leq \frac{1}{m} D(\mathscr{A}_m\|\mathscr{B}_m), \tag{21}$$

for all $m \geq 1$, with explicit convergence guarantees

$$\begin{aligned} &\frac{1}{m} D(\mathscr{A}_m\|\mathscr{B}_m) - \frac{1}{m} D_{\rm M}(\mathscr{A}_m\|\mathscr{B}_m)\\ &\qquad\qquad \leq \frac{1}{m} 2(d^2 + d)\log(m + d). \end{aligned} \tag{22}$$

Proposition 17 (Approximation of regularized relative entropies)**.** *Assume the same conditions for $\{\mathscr{A}_n\}_{n\in\mathbb{N}}$ and $\{\mathscr{B}_n\}_{n\in\mathbb{N}}$ as in Lemma 16. If each $\mathscr{A}_m$ and $\mathscr{B}_m$ have SDP representations of size s^{2m}, then $D(\mathscr{A}_m\|\mathscr{B}_m)$ and $D_{\rm M}(\mathscr{A}_m\|\mathscr{B}_m)$ can be computed by quantum relative entropy programs of size $O(k^{2m})$ with $k = \max\{s, d\}$.*

As a result, we can approximate $D^\infty(\mathscr{A}\|\mathscr{B})$ within additive error δ by a quantum relative entropy program of size $O(k^{2m_0})$ with $m_0 = \lceil \frac{8d^2}{\delta} \log \frac{d^2}{\delta} \rceil$ using Eq. (22).

Proof. Suppose $\mathscr{A}_m$ and $\mathscr{B}_m$ have SDP representations as

$$\mathscr{A}_m = \left\{\Pi_m(X) : X \in \mathscr{H}_+(\mathbb{C}^{s^m}), F_m(X) = g_m\right\}, \tag{23}$$

$$\mathscr{B}_m = \left\{\Pi'_m(X') : X' \in \mathscr{H}_+(\mathbb{C}^{s^m}), F'_m(X') = g'_m\right\}. \tag{24}$$

Then for the quantum relative entropy, we can write

$$
\begin{aligned}
&D(\mathscr{A}_m\|\mathscr{B}_m)\\
&= \inf\{t : \rho_m \in \mathscr{A}_m, \sigma_m \in \mathscr{B}_m, t \geq D(\rho_m\|\sigma_m)\} \quad (25)\\
&= \inf\{t : \rho_m \in \mathscr{A}_m, \sigma_m \in \mathscr{B}_m, (\rho_m, \sigma_m, t) \in \mathcal{K}_{\text{qre}}\} \quad (26)\\
&= \inf\Big\{t : \rho_m = \Pi_m(X), \sigma_m = \Pi'_m(X'),\\
&\qquad F_m(X) = g_m, F'_m(X') = g'_m \quad (27)\\
&\qquad ((\rho_m, \sigma_m, t), X, X') \in \mathcal{K}_{\text{qre}} \times \mathscr{H}_+(\mathbb{C}^{s^m}) \times \mathscr{H}_+(\mathbb{C}^{s^m})\Big\}
\end{aligned}
$$

which is a quantum relative entropy program of size $2d^{2m} + 2s^{2m} + 1 = O(k^{2m})$. For the measured relative entropy, we have from [16, Lemma 19] that

$$
D_{\text{M}}(\mathscr{A}_m\|\mathscr{B}_m) = \sup_{W_m \in (\mathscr{B}_m)^\circ_{++}} -h_{\mathscr{A}_m}(-\log W_m). \quad (28)
$$

We now proceed to write the convex program more explicitly. Since $\mathscr{A}_m \subseteq \mathscr{H}_+$, we know that $h_{\mathscr{A}_m}$ is operator monotone (i.e., $h_{\mathscr{A}_m}(X) \leq h_{\mathscr{A}_m}(Y)$ if $X \leq Y$). This gives

$$
D_{\text{M}}(\mathscr{A}_m\|\mathscr{B}_m) = \sup_{W_m, V_m} \big\{ -h_{\mathscr{A}_m}(V_m) : W_m \in (\mathscr{B}_m)^\circ_{++}, V_m \geq -\log W_m \big\}. \quad (29)
$$

Noting that $h_{\mathscr{A}_m}(V_m + t'I_m) = h_{\mathscr{A}_m}(V_m) + t'$ as $\mathscr{A}_m \subseteq \mathscr{D}$, we have

$$
\begin{aligned}
&h_{\mathscr{A}_m}(V_m)\\
&= \inf_{t'} \{h_{\mathscr{A}_m}(V_m + t'I_m) - t' : V_m + t'I_m \geq 0\} \quad (30)\\
&= \inf_{t,t'} \{t - t' : t > 0, V_m + t'I_m \in t\mathscr{A}^\circ_m, V_m + t'I_m \geq 0\}. \quad (31)
\end{aligned}
$$

Taking this into Eq. (29), we have

$$
\begin{aligned}
&D_{\text{M}}(\mathscr{A}_m\|\mathscr{B}_m)\\
&= \sup_{W_m, V_m, t, t'} \Big\{t' - t : t > 0, (W_m, 1) \in \mathbf{epi}(h_{\mathscr{B}_m}), \quad (32)\\
&\qquad (V_m + t'I_m, t) \in \mathbf{epi}(h_{\mathscr{A}_m}), V_m + t'I_m \geq 0,\\
&\qquad (I_m, W_m, V_m) \in \mathcal{K}_{\text{ore}}\Big\},
\end{aligned}
$$

which is also a quantum relative entropy program of size $2s^{2m} + 2 + 4d^{2m} = O(k^{2m})$ by Lemma 15. □

The above result shows that the regularized relative entropy $D^\infty(\mathscr{A}\|\mathscr{B})$ can be estimated using quantum relative entropy programs but with size exponential in the dimension $\max\{d, s\}$. In the following, we aim to exploit the permutation invariance of $\mathscr{A}_m$ and $\mathscr{B}_m$ to reduce the complexity and make the estimation more efficient. For this, we first show that the optimal solution of the relative entropy programs can be restricted to permutation invariant states.

Let $k \in \mathbb{N}$ be a fixed integer and $\mathcal{H}$ be a finite-dimensional vector space with $\dim(\mathcal{H}) = d$. Then the natural action of the symmetric group $\mathfrak{S}_k$ on $\mathcal{H}^{\otimes k}$ by permuting the indices is given by

$$
\pi \cdot (h_1 \otimes \cdots \otimes h_k) = h_{\pi^{-1}(1)} \otimes \cdots \otimes h_{\pi^{-1}(k)}, \quad (33)
$$

for any $h_i \in \mathcal{H}$ and $\pi \in \mathfrak{S}_k$. Let P_π be the permutation operator corresponding to the action of π on the suitable space. Denote the twirling operation as $\mathcal{T}_k(X) := \frac{1}{|\mathfrak{S}_k|}\sum_{\pi \in \mathfrak{S}_k} P_\pi X P_\pi^\dagger$. The algebra of $\mathfrak{S}_k$-invariant operators on $\mathcal{H}^{\otimes k}$ is denoted by

$$
\mathscr{I}_k := \left\{X \in \mathscr{L}(\mathcal{H}^{\otimes k}) : P_\pi X P_\pi^\dagger = X, \forall \pi \in \mathfrak{S}_k\right\}. \quad (34)
$$

Lemma 18. *Let $\mathscr{A}_m \subseteq \mathscr{D}(\mathcal{H}^{\otimes m})$ and $\mathscr{B}_m \subseteq \mathscr{H}_+(\mathcal{H}^{\otimes m})$ be convex, compact, permutation invariant sets. Then*

$$
\begin{aligned}
D(\mathscr{A}_m\|\mathscr{B}_m) &= D(\mathscr{A}_m \cap \mathscr{I}_m\|\mathscr{B}_m \cap \mathscr{I}_m), \quad (35)\\
D_{\text{M}}(\mathscr{A}_m\|\mathscr{B}_m) &= D_{\text{M}}(\mathscr{A}_m \cap \mathscr{I}_m\|\mathscr{B}_m \cap \mathscr{I}_m). \quad (36)
\end{aligned}
$$

Proof. The direction of "$\leq$" is clear for both equations. We now show the other direction. For the case of quantum relative entropy, we have that for any $\rho_m \in \mathscr{A}_m$ and $\sigma_m \in \mathscr{B}_m$, $D(\rho_m\|\sigma_m) \geq D(\mathcal{T}_m(\rho_m)\|\mathcal{T}_m(\sigma_m)) \geq D(\mathscr{A}_m \cap \mathscr{I}_m\|\mathscr{B}_m \cap \mathscr{I}_m)$, where the first inequality follows by the data-processing inequality of quantum relative entropy and the second inequality follows because $\mathcal{T}_m(\rho_m) \in \mathscr{A}_m \cap \mathscr{I}_m$ and $\mathcal{T}_m(\sigma_m) \in \mathscr{B}_m \cap \mathscr{I}_m$ by the permutation invariance of $\mathscr{A}_m$ and $\mathscr{B}_m$. Optimizing $\rho_m \in \mathscr{A}_m$ and $\sigma_m \in \mathscr{B}_m$ on both sides, we get $D(\mathscr{A}_m\|\mathscr{B}_m) \geq D(\mathscr{A}_m \cap \mathscr{I}_m\|\mathscr{B}_m \cap \mathscr{I}_m)$. The case of measured relative entropy follows by the same argument. □

The following result gives the SDP representation of the intersection of $\mathscr{A}_m$ and $\mathscr{I}_m$.

Lemma 19. *If the SDP representation of $\mathscr{A}_m$ in Eq. (23) satisfies the symmetry conditions: (1) $\Pi_m(P_\pi X P_\pi^\dagger) = P_\pi \Pi_m(X) P_\pi^\dagger$; (2) $F_m(P_\pi X P_\pi^\dagger) = P_\pi F_m(X) P_\pi^\dagger$; and (3) $P_\pi g_m P_\pi^\dagger = g_m$ for any $X \in \mathscr{H}_+((\mathbb{C}^s)^{\otimes m})$ and any permutation operator P_π on the suitable spaces, then*

$$
\begin{aligned}
&\mathscr{A}_m \cap \mathscr{I}_m \quad (37)\\
&= \left\{\Pi_m(X) : X \in \mathscr{H}_+((\mathbb{C}^s)^{\otimes m}) \cap \mathscr{I}_m, F_m(X) = g_m\right\},
\end{aligned}
$$

where $\mathscr{I}_m$ on both sides denotes the algebra of $\mathfrak{S}_m$-invariant operators on the suitable spaces.

Proof. We consider the inclusion "$\supseteq$" first. For any $\Pi_m(X)$ with $X \in \mathscr{H}_+((\mathbb{C}^s)^{\otimes m}) \cap \mathscr{I}_m$ and $F_m(X) = g_m$, it is clear that $\Pi_m(X) \in \mathscr{A}_m$ and moreover, we have

$$
P_\pi \Pi_m(X) P_\pi^\dagger = \Pi_m(P_\pi X P_\pi^\dagger) = \Pi_m(X), \quad (38)
$$

where the first equality follows from the assumption of Π_m and the second equality follows as $X \in \mathscr{I}_m$. As this holds for any $\pi \in \mathfrak{S}_m$, we have $\Pi_m(X) \in \mathscr{I}_m$ and therefore $\Pi_m(X) \in \mathscr{A}_m \cap \mathscr{I}_m$. Now we prove the inclusion "$\subseteq$". For any $\Pi_m(X) \in \mathscr{I}_m$ with $X \in \mathscr{H}_+((\mathbb{C}^s)^{\otimes m}), F_m(X) = g_m$, we have

$$
\Pi_m(\mathcal{T}_m(X)) = \mathcal{T}_m(\Pi_m(X)) = \Pi_m(X), \quad (39)
$$

where the first equality follows from the linearity and permutation invariant assumption of Π_m and the second equality follows as $\Pi_m(X) \in \mathscr{I}_m$. Similarly, we can argue that $F_m(\mathcal{T}_m(X)) = \mathcal{T}_m(g_m) = g_m$. This implies that $\Pi_m(X)$ belongs to the set on the right hand side of Eq. (37) because $\mathcal{T}(X) \in \mathscr{H}_+((\mathbb{C}^s)^{\otimes m}) \cap \mathscr{I}_m$. This concludes the proof. □

Next, we apply representation theory to reduce the size of the SDP representation for $\mathscr{A}_m \cap \mathscr{I}_m$. For that, we follow the notation in [58]. The set $\mathscr{I}_m$ of operators acting on $\mathcal{H}$

that are invariant under permutation is isomorphic to a direct sum of matrix algebras. In order to describe the optimization problems, we have to describe this isomorphism explicitly. For that, let $\mathrm{Par}(d,m)$ be the set of partitions λ of m of height d (i.e., $\lambda_1 \geq \cdots \geq \lambda_d > 0$ with $\lambda_1 + \cdots + \lambda_d = m$), $T_{\lambda,d}$ be the set of semistandard λ-tableaux with entries in $[d]$ and $m_\lambda^{\mathcal{H}} = |T_{\lambda,d_{\mathcal{H}}}|$ (see [58] for more details on these concepts). We then define the map

$$\begin{aligned}\phi_{\mathcal{H}} : \mathscr{I}_m(\mathcal{H}^{\otimes m}) &\to \bigoplus_{\lambda\in\mathrm{Par}(d_{\mathcal{H}},m)} \mathbb{C}^{m_\lambda^{\mathcal{H}}\times m_\lambda^{\mathcal{H}}} \qquad (40)\\ X &\mapsto \bigoplus_{\lambda\in\mathrm{Par}(d_{\mathcal{H}},m)} \left(\langle X u_\gamma, u_\tau\rangle\right)_{\tau,\gamma\in T_{\lambda,d_{\mathcal{H}}}},\end{aligned}$$

where $\{u_\tau\}_{\tau\in T_{\lambda,d_{\mathcal{H}}}}$ are vectors in $\mathcal{H}^{\otimes m}$ the exact definition of which can be found in [58]; see also [59, Section 2.1] for more details.

Note that in this decomposition, the number of blocks and the size of the blocks are bounded by a polynomial in m. In particular, we have

$$t^{\mathcal{H}} := |\mathrm{Par}\,(d_{\mathcal{H}},m)| \leq (m+1)^{d_{\mathcal{H}}}, \qquad (41)$$

$$m_\lambda^{\mathcal{H}} := |T_{\lambda,d_{\mathcal{H}}}| \leq (m+1)^{d_{\mathcal{H}}(d_{\mathcal{H}}-1)/2}, \qquad (42)$$

for any $\lambda \in \mathrm{Par}\,(d_{\mathcal{H}},m)$. Therefore, we get the dimension of the permutation-invariant subspace as

$$m^{\mathcal{H}} := \dim[\mathscr{I}_m] \leq (m+1)^{d_{\mathcal{H}}^2}. \qquad (43)$$

With the SDP representation of $\mathscr{A}_m\cap\mathscr{I}_m$ in Lemma 19, we can now apply the linear map $\phi_{\mathcal{H}}$ in Eq. (40) to decompose the operator X on the exponentially large space into a block-diagonal form. Specifically, let $\mathcal{H}_1 = \mathbb{C}^s$, $\mathcal{H}_2 = \mathbb{C}^d$ and $\mathcal{H}_3 = \mathbb{C}^f$, and define

$$\begin{aligned}\Pi'_m : \bigoplus_{\lambda\in\mathrm{Par}(s,m)} \mathbb{C}^{m_\lambda^{\mathcal{H}_1}\times m_\lambda^{\mathcal{H}_1}} &\to \bigoplus_{\lambda\in\mathrm{Par}(d,m)} \mathbb{C}^{m_\lambda^{\mathcal{H}_2}\times m_\lambda^{\mathcal{H}_2}} \qquad (44)\\ X &\mapsto \phi_{\mathcal{H}_2}(\Pi_m(\phi_{\mathcal{H}_1}^{-1}(X))),\end{aligned}$$

and

$$\begin{aligned}F'_m : \bigoplus_{\lambda\in\mathrm{Par}(s,m)} \mathbb{C}^{m_\lambda^{\mathcal{H}_1}\times m_\lambda^{\mathcal{H}_1}} &\to \bigoplus_{\lambda\in\mathrm{Par}(f,m)} \mathbb{C}^{m_\lambda^{\mathcal{H}_3}\times m_\lambda^{\mathcal{H}_3}} \qquad (45)\\ X &\mapsto \phi_{\mathcal{H}_3}(F_m(\phi_{\mathcal{H}_1}^{-1}(X))),\end{aligned}$$

and the linear operator $g'_m = \phi_{\mathcal{H}_3}(g_m)$. With these notation, we have the following SDP representation.

Lemma 20. *Let $\mathcal{H}_1 = \mathbb{C}^s$ and $\mathcal{H}_2 = \mathbb{C}^d$. Then the SDP representation in Eq. (37) gives*

$$\begin{aligned}\phi_{\mathcal{H}_2}(\mathscr{A}_m\cap\mathscr{I}_m) = \Bigg\{\Pi'_m\Bigg(\bigoplus_{\lambda\in\mathrm{Par}(s,m)} X_\lambda\Bigg) : \forall\lambda\in\mathrm{Par}(s,m),\\ X_\lambda \in \mathscr{H}_+\left(\mathbb{C}^{m_\lambda^{\mathcal{H}_1}}\right),\ F'_m\Bigg(\bigoplus_{\lambda\in\mathrm{Par}(s,m)} X_\lambda\Bigg) = g'_m\Bigg\}, \quad (46)\end{aligned}$$

where the SDP representation on the right hand side is of size at most $(m+1)^{s^2}$.

Proof. We show the inclusion "$\subseteq$" first. For any element $\phi_{\mathcal{H}_2}(\Pi_m(X)) \in \phi_{\mathcal{H}_2}(\mathscr{A}_m\cap\mathscr{I}_m)$, we have $X \in \mathscr{H}_+((\mathbb{C}^s)^{\otimes m})\cap\mathscr{I}_m$ and $F_m(X) = g_m$. Let $X' = \phi_{\mathcal{H}_1}(X)$. Then $X' \in \mathscr{H}_+(\mathcal{H}_1^{\otimes m})$ and $X' \in \bigoplus_{\lambda\in\mathrm{Par}(s,m)} \mathbb{C}^{m_\lambda^{\mathcal{H}_1}\times m_\lambda^{\mathcal{H}_1}}$. Moreover, $\phi_{\mathcal{H}_2}(\Pi_m(X)) = \Pi'_m(X')$ and $F'_m(X') = g'_m$. This implies that $\phi_{\mathcal{H}_2}(\Pi_m(X))$ is included in the right-hand side of Eq. (46).

Now we show the other inclusion "$\supseteq$". For any element $\Pi'_m(X)$ with $X \in \mathscr{H}_+(\mathcal{H}_1^{\otimes m})$, $X \in \bigoplus_{\lambda\in\mathrm{Par}(s,m)} \mathbb{C}^{m_\lambda^{\mathcal{H}_1}\times m_\lambda^{\mathcal{H}_1}}$ and $F'_m(X) = g'_m$, let $X'' = \phi_{\mathcal{H}_1}^{-1}(X)$. Then we have $X'' \in \mathscr{H}_+((\mathbb{C}^d)^{\otimes m})\cap\mathscr{I}_m$, $\Pi'_m(X) = \phi_{\mathcal{H}_2}(\Pi_m(X''))$ and $F_m(X'') = g_m$. This implies $\Pi'_m(X) \in \phi_{\mathcal{H}_2}(\mathscr{A}_m\cap\mathscr{I}_m)$ and concludes the proof. □

Combining Lemmas 18, 19 20 and [58, Lemma 3.3], we have $D(\mathscr{A}_m\|\mathscr{B}_m) = D(\phi_{\mathcal{H}_2}(\mathscr{A}_m\cap\mathscr{I}_m)\|\phi_{\mathcal{H}_2}(\mathscr{B}_m\cap\mathscr{I}_m))$ and $D_{\mathrm{M}}(\mathscr{A}_m\|\mathscr{B}_m) = D_{\mathrm{M}}(\phi_{\mathcal{H}_2}(\mathscr{A}_m\cap\mathscr{I}_m)\|\phi_{\mathcal{H}_2}(\mathscr{B}_m\cap\mathscr{I}_m))$. Together with Proposition 17, we get the efficient approximation of regularized relative entropies presented in Theorem 10.

IV. Applications

In this section, we apply the idea of efficient approximation to several quantum information processing tasks. Note that the applicability of our approximation relies on the structural assumptions of the sets in Assumption 9, which holds directly for many cases such as the singleton set, the set of incoherent states used in coherence theory and the image set of a channel used in adversarial channel discrimination. We exemplify the last case in Section IV-A. In cases where the task of interest does not directly satisfy Assumption 9, one can follow a general methodology of relaxing the set of interest to one that does, particularly regarding the polar assumption in (A.4). For instance, in entanglement theory, the set of separable states can be relaxed to the Rains set, which satisfies all necessary assumptions. Similarly, in fault-tolerant quantum computing, the set of stabilizer states can be relaxed to the set of states with non-positive mana, which also fulfills the required conditions. We provide several examples in Sections IV-B, IV-C and IV-D to illustrate this idea, which gives improvement to the state-of-the-art results in the literature.

A. Adversarial quantum channel discrimination

We now apply the general theory in Theorem 10 to compute the minimum output channel divergence, which serves as the key quantity in adversarial quantum channel discrimination [16].

Definition 21 (Minimum output quantum channel divergence)**.** *Let $\mathbb{D}$ be a quantum divergence between quantum states. Let $\mathcal{N} \in \mathrm{CPTP}(A:B)$ and $\mathcal{M} \in \mathrm{CP}(A:B)$. Define the corresponding minimum output channel divergence by*

$$\mathbb{D}^{\inf}(\mathcal{N}\|\mathcal{M}) := \inf_{\substack{\rho\in\mathscr{D}(A)\\ \sigma\in\mathscr{D}(A)}} \mathbb{D}(\mathcal{N}_{A\to B}(\rho_A)\|\mathcal{M}_{A\to B}(\sigma_A)). \quad (47)$$

Define its regularized channel divergence by

$$\mathbb{D}^{\inf,\infty}(\mathcal{N}\|\mathcal{M}) := \lim_{n\to\infty} \frac{1}{n}\mathbb{D}^{\inf}(\mathcal{N}^{\otimes n}\|\mathcal{M}^{\otimes n}). \qquad (48)$$

The following result shows that the regularized channel divergence can be efficiently approximated by quantum relative entropy programs.

Corollary 22 (Efficient approximation of the regularized minimum output channel divergences)**.** *Let $\mathcal{N} \in \mathrm{CPTP}(A:B)$ and $\mathcal{M} \in \mathrm{CP}(A:B)$ with $\dim A = s$ and $\dim B = d$. The minimum output quantum channel divergences $D(\mathcal{N}^{\otimes m}\|\mathcal{M}^{\otimes m})$ and $D_{\mathrm{M}}(\mathcal{N}^{\otimes m}\|\mathcal{M}^{\otimes m})$ can both be computed by quantum relative entropy programs of size $O((m+1)^{k^2})$ with $k = \max\{s,d\}$.*

As a result, $D^{\inf,\infty}(\mathcal{N}\|\mathcal{M})$ can be approximated within additive error δ by a quantum relative entropy program of size $O((m_0+1)^{k^2})$ with $m_0 = \lceil \frac{8d^2}{\delta}\log\frac{d^2}{\delta}\rceil$.

Proof. It is clear that the minimum output channel divergence is the divergence between the image sets of the channels,

$$\mathscr{A}_m = \{\mathcal{N}^{\otimes m}(\rho) : \rho \in \mathscr{H}_+\big((\mathbb{C}^s)^{\otimes m}\big), \operatorname{Tr}\rho = 1\}, \tag{49}$$
$$\mathscr{B}_m = \{\mathcal{M}^{\otimes m}(\rho) : \rho \in \mathscr{H}_+\big((\mathbb{C}^s)^{\otimes m}\big), \operatorname{Tr}\rho = 1\}. \tag{50}$$

These sets satisfy all the required assumptions in Assumption 9. First, the set of all density matrices $\mathscr{D}$ is convex and compact, so $\mathscr{A}_m$ is also convex and compact. Since $\mathcal{N}^{\otimes m}$ and $\mathscr{D}$ are permutation invariant, we know that $\mathscr{A}_m$ is also permutation invariant. For any $\mathcal{N}^{\otimes m}(\rho_m) \in \mathscr{A}_m$ and $\mathcal{N}^{\otimes k}(\rho_k) \in \mathscr{A}_k$, we have $\mathcal{N}^{\otimes m}(\rho_m) \otimes \mathcal{N}^{\otimes k}(\rho_k) = \mathcal{N}^{\otimes(m+k)}(\rho_m \otimes \rho_k) \in \mathscr{A}_{m+k}$. This implies $\mathscr{A}_m \otimes \mathscr{A}_k \subseteq \mathscr{A}_{m+k}$. The support function of $\mathscr{A}_n$ is given by

$$\begin{aligned} h_{\mathscr{A}_n}(X_n) &= \sup_{\rho_n \in \mathscr{D}} \operatorname{Tr}\left[X_n \mathcal{N}^{\otimes n}(\rho_n)\right] && (51)\\ &= \sup_{\rho_n \in \mathscr{D}} \operatorname{Tr}\left[(\mathcal{N}^{\otimes n})^\dagger(X_n)\rho_n\right] && (52)\\ &= \lambda_{\max}\left((\mathcal{N}^{\otimes n})^\dagger(X_n)\right). && (53)\end{aligned}$$

So for any $X_m \in \mathscr{H}_+$ and $X_k \in \mathscr{H}_+$, we have

$$\begin{aligned} & h_{\mathscr{A}_{m+k}}(X_m \otimes X_k) \\ &= \lambda_{\max}\left((\mathcal{N}^{\otimes(m+k)})^\dagger(X_m \otimes X_k)\right) && (54)\\ &= \lambda_{\max}\left((\mathcal{N}^{\otimes m})^\dagger(X_m) \otimes (\mathcal{N}^{\otimes k})^\dagger(X_k)\right) && (55)\\ &= \lambda_{\max}\left((\mathcal{N}^{\otimes m})^\dagger(X_m)\right)\lambda_{\max}\left((\mathcal{N}^{\otimes k})^\dagger(X_k)\right) && (56)\\ &= h_{\mathscr{A}_m}(X_m)h_{\mathscr{A}_k}(X_k), && (57)\end{aligned}$$

where the third line follows by the multiplicativity of the maximum eigenvalue of tensor product operators. Finally, using Lemma 3, we know that $(\mathscr{A}_m)_+^\circ \otimes (\mathscr{A}_k)_+^\circ \subseteq (\mathscr{A}_{m+k})_+^\circ$. This proves that $\{\mathscr{A}_n\}_{n\in\mathbb{N}}$ satisfy all assumptions in Assumption 9. Moreover, $\mathscr{A}_m = \{\mathcal{N}^{\otimes m}(\rho) : \rho \in \mathscr{H}_+((\mathbb{C}^s)^{\otimes m}), \operatorname{Tr}\rho = 1\}$ is a SDP representation of size s^{2m} and $\Pi_m = \mathcal{N}^{\otimes m}$, $F_m = \operatorname{Tr}$, $g_m = 1$ satisfy the symmetry conditions in Lemma 19. The same holds for $\mathscr{B}_m$. Applying Theorem 10, we have the asserted statement. □

In the following, we provide an explicit example to show that the regularized minimum output channel divergence $D^{\infty,\inf}(\mathcal{N}\|\mathcal{M})$ can be approximated by $D^{\inf}(\mathcal{N}^{\otimes m}\|\mathcal{M}^{\otimes m})/m$ from above and $D_{\mathrm{M}}^{\inf}(\mathcal{N}^{\otimes m}\|\mathcal{M}^{\otimes m})/m$ from below, with the approximation improving as m increases.

This example is given by two qutrit quantum channels. Let $\mathcal{N}(\cdot) = \operatorname{Tr}[\cdot]|\rho\rangle\langle\rho|$ to be the replacer channel with $|\rho\rangle = (2|0\rangle + |1\rangle + 2|2\rangle)/3$. Let $\mathcal{M}$ be the platypus channel [60, Eq. (170)], $\mathcal{M}(X) = M_0 X M_0^\dagger + M_1 X M_1^\dagger$ with Kraus operators

$$M_0 = \begin{bmatrix} \sqrt{p} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \qquad M_1 = \begin{bmatrix} 0 & 0 & 0 \\ \sqrt{1-p} & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{58}$$

Since $\mathcal{M}(I/3) = (p|0\rangle\langle 0| + (1-p)|1\rangle\langle 1| + 2|2\rangle\langle 2|)/3$ is a full rank state for all $p \in (0,1)$, the following quantities are finite and can be evaluated by the QICS package [53],

$$D^{\inf}(\mathcal{N}^{\otimes m}\|\mathcal{M}^{\otimes m}) = \inf_{\sigma_m \in \mathscr{D}} D(\rho^{\otimes m}\|\mathcal{M}^{\otimes m}(\sigma_m)), \tag{59}$$
$$D_{\mathrm{M}}^{\inf}(\mathcal{N}^{\otimes m}\|\mathcal{M}^{\otimes m}) = \inf_{\sigma_m \in \mathscr{D}} D_{\mathrm{M}}(\rho^{\otimes m}\|\mathcal{M}^{\otimes m}(\sigma_m)). \tag{60}$$

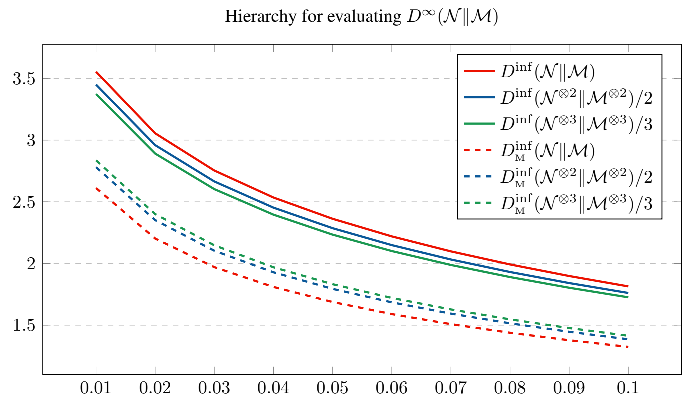

Fig. 1: Estimating the regularized minimum output channel divergences, where the horizontal axis refers to the channel parameter p of $\mathcal{M}$ and the vertical axis refers to the divergence values.

The numerical result is given in Figure 1. It shows a clear separation between upper bounds with $m = 1,2,3$ and $p \in [0.01, 0.1]$ and also lower bounds with the same parameter range, confirming the strict subadditivity of the minimum output Umegaki channel divergence and the strict superadditivity of the minimum output measured channel divergence. Moreover, as we increase the number of m, the lower and upper bounds provide better approximation to $D^{\infty,\inf}(\mathcal{N}\|\mathcal{M})$.

B. Entanglement cost for quantum states and channels

The *entanglement cost* of a quantum state, denoted as $E_{C,\Omega}$, is the minimum number of Bell states required to prepare one copy of this state under a class of operations Ω. Of particular interest is the local operation and classical communication (LOCC) operations. It is known that computing $E_{C,\mathrm{LOCC}}$ is NP-hard in general [61, Theorem 1]. Therefore, finding efficiently computable lower and upper bounds to estimate $E_{C,\mathrm{LOCC}}$ is of fundamental importance. Here, we focus on deriving lower bounds, which indicate that no matter what preparation strategies are used, the amount of entanglement consumed cannot be smaller than this value.

There are several lower bounds in the literature, but they are unsatisfactory for different reasons. One such lower bound is given by the regularized PPT-relative entropy of entanglement [12, Eq. (8.235)],

$$E_{C,\mathrm{LOCC}}(\rho) \geq D^{\infty}(\rho\|\mathrm{PPT}) \quad (61)$$
$$:= \lim_{n\to\infty} \frac{1}{n} D(\rho^{\otimes n}\|\mathrm{PPT}(A^n : B^n)), \quad (62)$$

which is difficult to evaluate due to its regularization. Note that the PPT set does not satisfy the assumption (A.4) and therefore cannot directly apply Theorem 10. A counter-example can be given by the projector on the $\mathbb{C}^3 \otimes \mathbb{C}^3$ antisymmetric subspace, denoted as ρ^a. Since the support function $h_{\mathscr{C}_{\mathrm{PPT}}}(\cdot)$ is given by a semidefinite program, we find numerically that $h_{\mathrm{PPT}}((\rho^a)^{\otimes 2}) - h_{\mathrm{PPT}}(\rho^a)^2 \approx 0.0093 > 0$.

A single-letter lower bound for entanglement cost is provided by the quantum squashed entanglement [62], but its computability remains uncertain due to the unbounded dimension of the extension system. Wang and Duan have proposed two single-letter SDP lower bounds in separate works:

$$E_{\mathrm{WD},1}(\rho_{AB}) := -\log \max \operatorname{Tr} \Pi_\rho V_{AB} \quad \text{s.t. } \|V_{AB}^{\mathsf{T}_B}\|_1 = 1,\ V_{AB} \geq 0, \quad [39] \quad (63)$$
$$E_{\mathrm{WD},2}(\rho_{AB}) := -\log \min \ \|Y_{AB}^{\mathsf{T}_B}\|_\infty \quad \text{s.t. } -Y_{AB} \leq \Pi_\rho^{\mathsf{T}_B} \leq Y_{AB}. \quad [40] \quad (64)$$

It has been shown that for any $\rho \in \mathscr{D}(AB)$,

$$\max\left\{E_{\mathrm{WD},1}(\rho), E_{\mathrm{WD},2}(\rho)\right\} \leq D^{\infty}(\rho\|\mathrm{PPT}) \leq E_{C,\mathrm{LOCC}}(\rho). \quad (65)$$

By the definition of min-relative entropy and the Rains set, we can write [2]

$$E_{\mathrm{WD},1}(\rho_{AB}) = D_{\min}(\rho_{AB}\|\mathrm{Rains}(A:B)). \quad (66)$$

After a detailed examination of the dual program for $E_{\mathrm{WD},2}$, we can also reformulate it in a comparable structure by introducing an appropriate set of operators.

Lemma 23. *For any $\rho \in \mathscr{D}(AB)$, it holds that*

$$E_{\mathrm{WD},2}(\rho_{AB}) = D_{\min}(\rho_{AB}\|\mathrm{WD}(A:B)), \quad (67)$$

with the set of Hermitian operators defined by

$$\mathrm{WD}(A:B) := \left\{\sigma \in \mathscr{H}(AB) : \exists Y \in \mathscr{H}(AB), \ \text{s.t. } -Y \leq \sigma^{\mathsf{T}_B} \leq Y, \|Y^{\mathsf{T}_B}\|_1 \leq 1\right\}. \quad (68)$$

Proof. Using the Lagrangian method, we have the dual SDP of $E_{\mathrm{WD},2}$ as

$$E_{\mathrm{WD},2}(\rho) = -\log \max \ \operatorname{Tr} \Pi_\rho (V - F)^{\mathsf{T}_B} \quad \text{s.t.} \quad (69)$$
$$V + F = (W - X)^{\mathsf{T}_B},\ \operatorname{Tr}(W + X) \leq 1,\ V, F, W, X \geq 0.$$

Let $(V - F)^{\mathsf{T}_B} = \sigma$, $V + F = Y$. By the definition of $D_{\min}$ and $\mathscr{C}_{\mathrm{WD}}$, we have the desired result. □

As discussed above, both SDP bounds $E_{\mathrm{WD},1}$ and $E_{\mathrm{WD},2}$ are essentially entanglement measures induced by the min-relative entropy. However, a significant limitation of these bounds is that they vanish for any full-rank state.

[2] The original definition of $E_{\mathrm{WD},1}$ imposes the condition $\|V_{AB}^{\mathsf{T}_B}\|_1 = 1$. However, it is equivalent to optimize over the condition $\|V_{AB}^{\mathsf{T}_B}\|_1 \leq 1$, as the optimal solution can always be chosen at the boundary.

Proposition 24. *For any full rank state $\rho \in \mathscr{D}(AB)$,*

$$E_{\mathrm{WD},1}(\rho) = E_{\mathrm{WD},2}(\rho) = 0. \quad (70)$$

Proof. If ρ_{AB} is full rank, then $\Pi_\rho = I_{AB}$ and thus

$$E_{\mathrm{WD},1}(\rho_{AB}) = -\log \max \left\{ \operatorname{Tr} \sigma_{AB} : \sigma_{AB} \geq 0, \|\sigma_{AB}^{\mathsf{T}_B}\|_1 \leq 1\right\}. \quad (71)$$

It is clear that for any feasible solution σ_{AB} it holds that $\operatorname{Tr} \sigma_{AB} = \operatorname{Tr} \sigma_{AB}^{\mathsf{T}_B} \leq \|\sigma_{AB}^{\mathsf{T}_B}\|_1 \leq 1$. On the other hand, there is a feasible solution $\sigma_{AB} = I_{AB}/|AB|$ such that $\operatorname{Tr} \sigma_{AB} = 1$. Thus the maximization is taken at $\operatorname{Tr} \sigma_{AB} = 1$ and thus $E_{\mathrm{WD},1}(\rho_{AB}) = 0$. Similarly, given full rank ρ_{AB}, it holds that

$$E_{\mathrm{WD},2}(\rho_{AB}) = -\log \max \left\{ \operatorname{Tr} \sigma_{AB} : \sigma, Y \in \mathscr{H}(AB), -Y_{AB} \leq \sigma_{AB}^{\mathsf{T}_B} \leq Y_{AB}, \|Y^{\mathsf{T}_B}\|_1 \leq 1\right\}. \quad (72)$$

Then for any feasible solutions σ_{AB}, Y_{AB}, it holds that $\operatorname{Tr} \sigma_{AB} = \operatorname{Tr} \sigma_{AB}^{\mathsf{T}_B} \leq \operatorname{Tr} Y_{AB} = \operatorname{Tr} Y_{AB}^{\mathsf{T}_B} \leq \|Y^{\mathsf{T}_B}\|_1 \leq 1$. On the other hand, considering the feasible solution $\sigma_{AB} = Y_{AB} = I_{AB}/|AB|$, we have $\operatorname{Tr} \sigma_{AB} = 1$. Thus the maximization is taken at $\operatorname{Tr} \sigma_{AB} = 1$ and $E_{\mathrm{WD},2}(\rho_{AB}) = 0$. □

Recently, Wang et al. introduced the PPT_k set [41] as

$$\mathrm{PPT}_k(A:B) := \left\{\omega_1 \geq 0 : \exists \{\omega_i\}_{i=2}^k, \ \text{s.t. } \|\omega_i^{\mathsf{T}_B}\|_* \leq \omega_{i+1}, \forall i \in [1:k-1], \|\omega_k^{\mathsf{T}_B}\|_1 \leq 1\right\}, \quad (73)$$

where $k \in \mathbb{N}_+$ and $|X|_* \leq Y$ denotes $-Y \leq X \leq Y$. This set turns out to be related to the quantity χ_p developed in [63]. Building on this set, the authors introduced an efficiently computable lower bound for entanglement cost [41],

$$E_{C,\mathrm{LOCC}}(\rho_{AB}) \geq E_{\mathrm{WJZ}}(\rho_{AB}) \quad (74)$$
$$:= D_{\mathrm{M},1/2}(\rho_{AB}\|\mathrm{PPT}_k(A:B)). \quad (75)$$

In the following, we show that PPT_k satisfies Assumption 9 and thus we can apply our Theorem 10 to get an improved bound.

Lemma 25. *Let $k \geq 2$. The PPT_k set defined in Eq. (73) satisfies Assumption 9.*

Proof. It is clear to check that PPT_k satisfies (A.1) and (A.2). Moreover, suppose $-X_1 \leq Y_1 \leq X_1$ and $-X_2 \leq Y_2 \leq X_2$, we have $-I \leq X_1^{-1/2} Y_1 X_1^{-1/2} \leq I$ and $-I \leq X_2^{-1/2} Y_2 X_2^{-1/2} \leq I$ where the inverses are taken on the supports. This gives $-I \leq X_1^{-1/2} Y_1 X_1^{-1/2} \otimes X_2^{-1/2} Y_2 X_2^{-1/2} \leq I$ which is equivalent to $-X_1 \otimes X_2 \leq Y_1 \otimes Y_2 \leq X_1 \otimes X_2$. Using this result, we can check that PPT_k satisfies (A.3). As for the assumption (A.4), recall the variational characterization of the trace norm $\|X\|_1 = \min\{\operatorname{Tr} Y : -Y \leq X \leq Y\}$ [63, Lemma S1]. We have that

$$\mathrm{PPT}_k(A:B) = \left\{\omega_1 \geq 0 : \exists \{\omega_i\}_{i=2}^k, \ \text{s.t. } \|\omega_i^{\mathsf{T}_B}\|_* \leq \omega_{i+1}, \forall i \in [1:k], \operatorname{Tr} \omega_{k+1} \leq 1\right\}. \quad (76)$$

The support function is $h_{\mathrm{PPT}_k}(\omega) = \sup_{\omega_1 \in \mathrm{PPT}_k} \mathrm{Tr}[\omega\omega_1]$. The Lagrange multiplier is given by

$$\begin{aligned}
&\mathrm{Tr}[\omega\omega_1] + \sum_{i=1}^{k} \mathrm{Tr}[\alpha_i(\omega_i^{\mathsf{T}_B} + \omega_{i+1})] \\
&\quad + \sum_{i=1}^{k} \mathrm{Tr}[\beta_i(\omega_{i+1} - \omega_i^{\mathsf{T}_B})] + t(1 - \mathrm{Tr}[\omega_{k+1}]) \\
&= \mathrm{Tr}[\omega_1(\omega + \alpha_1^{\mathsf{T}_B} - \beta_1^{\mathsf{T}_B})] \\
&\quad + \sum_{i=2}^{k} \mathrm{Tr}[\omega_i(\alpha_{i-1} + \beta_{i-1} + \alpha_i^{\mathsf{T}_B} - \beta_i^{\mathsf{T}_B})] \\
&\quad + \mathrm{Tr}[\omega_{k+1}(\alpha_k + \beta_k - tI)] + t. \qquad (77)
\end{aligned}$$

By the SDP duality, we have the support function

$$\begin{aligned}
h_{\mathrm{PPT}_k}(\omega) = \inf \Big\{ & t \geq 0 : \ \omega \leq \beta_1^{\mathsf{T}_B} - \alpha_1^{\mathsf{T}_B}, \qquad (78) \\
& \alpha_{i-1} + \beta_{i-1} \leq \beta_i^{\mathsf{T}_B} - \alpha_i^{\mathsf{T}_B}, \forall i \in [2:k], \\
& \alpha_k + \beta_k \leq tI, \alpha_i \geq 0, \beta_i \geq 0, \forall i \in [1:k] \Big\}.
\end{aligned}$$

Then for any ω_1, ω_2, assume their optimal solutions in the dual program are respectively given by $\{t^1, \alpha_i^1, \beta_i^1\}$ and $\{t^2, \alpha_i^2, \beta_i^2\}$. Then we construct

$$t^3 := t^1 t^2, \qquad (79)$$
$$\alpha_i^3 := \alpha_i^1 \otimes \beta_i^2 + \beta_i^1 \otimes \alpha_i^2, \qquad (80)$$
$$\beta_i^3 := \alpha_i^1 \otimes \alpha_i^2 + \beta_i^1 \otimes \beta_i^2. \qquad (81)$$

Note that

$$(\alpha_i^1 + \beta_i^1) \otimes (\alpha_i^2 + \beta_i^2) = \alpha_i^3 + \beta_i^3, \qquad (82)$$
$$(\beta_i^1 - \alpha_i^1)^{\mathsf{T}_B} \otimes (\beta_i^2 - \alpha_i^2)^{\mathsf{T}_B} = (\beta_i^3 - \alpha_i^3)^{\mathsf{T}_B}. \qquad (83)$$

Then it is clear that $\{t^3, \alpha_i^3, \beta_i^3\}$ is a feasible solution for $h_{\mathrm{PPT}_k}(\omega_1 \otimes \omega_2)$. This implies that $h_{\mathrm{PPT}_k}(\omega_1 \otimes \omega_2) \leq t^3 = t^1 t^2 = h_{\mathrm{PPT}_k}(\omega_1) h_{\mathrm{PPT}_k}(\omega_2)$, and proves that PPT_k satisfies assumption (A.4) by Lemma 3. This completes the proof. □

Theorem 26. *Let $\rho \in \mathscr{D}(AB)$ and $k \geq 2$. Let $(*) = \max\{E_{\mathrm{WD},1}(\rho), E_{\mathrm{WD},2}(\rho), E_{\mathrm{WJZ}}(\rho)\}$ represent the previously known bounds. Then it holds that*

$$(*) \leq D_{\mathrm{M}}(\rho \| \mathrm{PPT}_k) \leq D^{\infty}(\rho \| \mathrm{PPT}_k) \qquad (84)$$
$$\leq D^{\infty}(\rho \| \mathrm{PPT}) \leq E_{C,\mathrm{LOCC}}(\rho). \qquad (85)$$

Moreover, $D_{\mathrm{M}}(\rho \| \mathrm{PPT}_k)$ can be seen as the first level of approximation to $D^{\infty}(\rho \| \mathrm{PPT}_k)$ and both quantities can be efficiently estimated.

Proof. It has been shown that [41, Proposition S4],

$$\mathrm{PPT}(A:B) \subseteq \mathrm{PPT}_k(A:B) \subseteq \cdots \subseteq \mathrm{Rains}(A:B). \qquad (86)$$

It is also clear from their definitions that $\mathrm{PPT}_2(A:B) = \mathrm{WD}(A:B)$. This implies

$$\begin{aligned}
\mathrm{PPT}(A:B) &\subseteq \mathrm{PPT}_k(A:B) \subseteq \qquad (87) \\
&\cdots \mathrm{PPT}_2(A:B) \subseteq \mathrm{WD}(A:B) \cap \mathrm{Rains}(A:B).
\end{aligned}$$

Therefore, the first two inequalities of the asserted result follow from the relation of divergences in Lemma 7 and the relation of sets in Eq. (87). The second inequality follows from the superadditivity in [16, Lemma 21] and the asymptotic equivalence in [16, Lemma 28]. The third inequality follows from the relation in Eq. (87). The last equality is known from Eq. (65). The computability of $D_{\mathrm{M}}(\rho \| \mathrm{PPT}_k)$ and $D^{\infty}(\rho \| \mathrm{PPT}_k)$ follows from Lemma 25 and Theorem 10. □

Following similar arguments in [41], we can also show that the new measures $D_{\mathrm{M}}(\rho \| \mathrm{PPT}_k)$ and $D^{\infty}(\rho \| \mathrm{PPT}_k)$ satisfy the desired properties such as normalization, faithfulness and (super-) additivity.

Besides the bounds mentioned, which are established on the PPT set, there is another efficiently computable lower bound on $E_{C,\mathrm{LOCC}}$ given by Lami and Regula in [42],

$$\begin{aligned}
E_{\mathrm{LR}}(\rho) := \log \sup \big\{ & \mathrm{Tr}\, X\rho : \\
& \|X^{\mathsf{T}_B}\|_\infty \leq 1, \|X\|_\infty = \mathrm{Tr}[X\rho] \big\}. \qquad (88)
\end{aligned}$$

However, this bound also vanishes for full rank states [64, Eq. (43)].

In the following, we compare our new bounds with previously established ones through several examples, including Isotropic states and Werner states. To the best of our knowledge, the entanglement costs for these states under LOCC operations remain unresolved. Additionally, we use randomly generated quantum states to showcase the broad applicability and improvement of our bound across unstructured quantum states. In all cases, our experiments clearly demonstrate the superiority of our bound (even for the first level of approximation) over existing ones.

Example 1 (Isotropic states and Werner states.) The Isotropic state is defined by a convex mixture of the maximally entangled state and its orthogonal complement,

$$\rho_{I,p} := p|\Phi\rangle\langle\Phi| + \frac{1-p}{d^2-1}(I - |\Phi\rangle\langle\Phi|), \qquad (89)$$

where $|\Phi\rangle = \frac{1}{\sqrt{d}}\sum_{i=1}^{e} |ii\rangle$ is the d-dimensional maximally entangled state. The PPT-relative entropy of entanglement for $\rho_{I,p}$ and its regularization are given by [65, Theorem 7]

$$\begin{aligned}
& D^{\infty}(\rho_{I,p} \| \mathrm{PPT}) = D(\rho_{I,p} \| \mathrm{PPT}) \qquad (90) \\
& = \begin{cases} 0 & \text{if} \quad 0 \leq p \leq \frac{1}{d}, \\ \log d + p \log p + (1-p) \log \frac{1-p}{d-1} & \text{if} \quad \frac{1}{d} \leq p \leq 1. \end{cases}
\end{aligned}$$

The Werner state is defined by a convex mixture of the normalized projectors on the symmetric (ρ^s) and anti-symmetric (ρ^a) subspaces,

$$\rho_{W,p} := (1-p)\rho^s + p\rho^a \qquad (91)$$
$$= \frac{1-p}{d(d+1)}(I+S) + \frac{p}{d(d-1)}(I-S), \qquad (92)$$

where $S = \sum_{i,j=1}^{e} |ij\rangle\langle ji|$ is the SWAP operator of dimension d. The regularized PPT-relative entropy of entanglement for $\rho_{W,p}$ is given by [66]

$$\begin{aligned}
& D^{\infty}(\rho_{W,p} \| \mathrm{PPT}) \\
& = \begin{cases} 0, & \text{if} \quad 0 \leq p \leq \frac{1}{2}, \\ 1 - h(p), & \text{if} \quad \frac{1}{2} < p \leq \frac{d+2}{2d}, \\ \log \frac{d+2}{d} + (1-p) \log \frac{d-2}{d+2}, & \text{if} \quad p > \frac{d+2}{2d}. \end{cases} \qquad (93)
\end{aligned}$$

Note that both the Isotropic states and the Werner states are full-rank states for any $p \in (0,1)$. The bounds $E_{\mathrm{WD},1}$, $E_{\mathrm{WD},2}$, and E_{LR} all reduce to zero. We then compare our new bound $D_{\mathrm{M}}(\rho\|\mathrm{PPT}_2)$ with the previously established bounds E_{WJZ} [41] and the analytical bound $D^{\infty}(\rho\|\mathrm{PPT})$ in Figure 2. It turns out that the first level of approximation $D_{\mathrm{M}}(\rho\|\mathrm{PPT}_2)$ already coincides with the analytical bound $D^{\infty}(\rho\|\mathrm{PPT})$ for both cases, improving the numerical bounds E_{WJZ}.

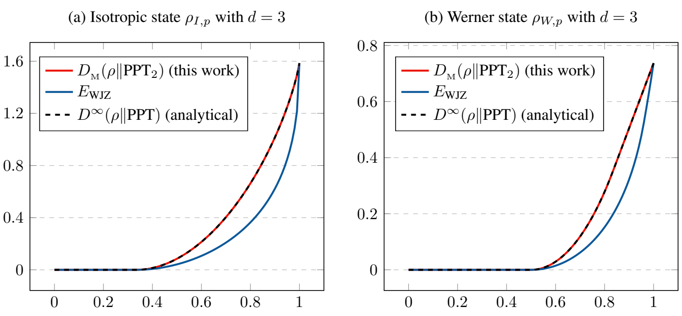

Fig. 2: Comparison of the new lower bound $D_{\mathrm{M}}(\rho\|\mathrm{PPT}_2)$ with previous bounds E_{WJZ} [41] and $D^{\infty}(\rho\|\mathrm{PPT})$ [65], [66] for (a) Isotropic states and (b) Werner states. The horizontal axis is the state parameter p and the vertical axis is the value of the entanglement measure.

Example 2 (Random quantum states.) Since $D_{\mathrm{M}}(\rho\|\mathrm{PPT}_2)$ has been proved to be better than E_{WJZ} in general, we focus our comparison here with E_{LR} [42] by generating random bipartite states according to the Hilbert-Schmidt measure, with varying ranks. For each rank, we generate 500 quantum states of dimension $3 \otimes 3$. The comparison is presented in Figure 3. It is evident that $D_{\mathrm{M}}(\rho\|\mathrm{PPT}_2)$ outperforms E_{LR} in most cases, particularly for higher-rank states.

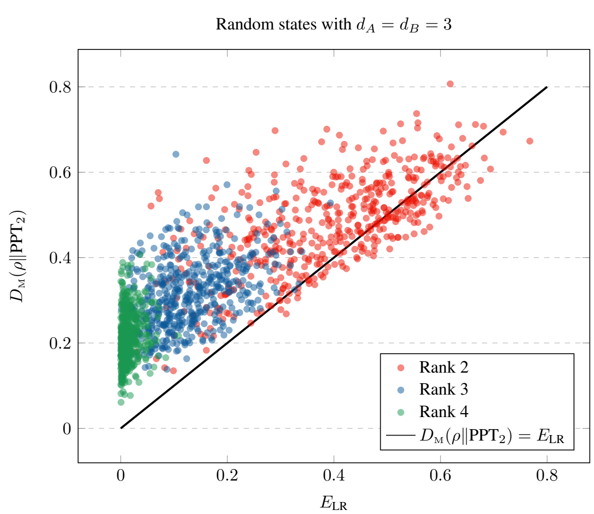

Fig. 3: Comparison of the new bound $D_{\mathrm{M}}(\rho\|\mathrm{PPT}_2)$ with the previously known bound E_{LR} [42] for randomly generated quantum states with different ranks.

Similar to the entanglement cost for quantum states, the entanglement cost of a quantum channel, denoted by $E_{C,\mathrm{LOCC}}(\mathcal{N})$, represents the minimal rate at which entanglement (between the sender and receiver) is required to simulate multiple copies of the channel, given the availability of free classical communication. It is known that [67]

$$E_{C,\mathrm{LOCC}}(\mathcal{N}) \geq \sup_{\rho\in\mathscr{D}(AA')} E_{C,\mathrm{LOCC}}(\mathcal{N}_{A\to B}(\rho_{AA'})), \quad (94)$$

where system $\mathcal{H}_{A'}$ is isomorphic to system $\mathcal{H}_A$. Wang et al. introduced a lower bound for the entanglement cost of a quantum channel in [41],

$$E_{C,\mathrm{LOCC}}(\mathcal{N}) \geq E_{\mathrm{WJZ}}(\mathcal{N}) \geq E_{\mathrm{WJZ}}(\mathcal{N}_{A\to B}(\Phi_{AA'})), \quad (95)$$

where $E_{\mathrm{WJZ}}(\mathcal{N}) := \sup_{\rho\in\mathscr{D}(AA')} E_{\mathrm{WJZ}}(\mathcal{N}_{A\to B}(\rho_{AA'}))$ and $\Phi_{AA'}$ is the maximally entangled state. This lower bound has been used to demonstrate that the resource theory of entanglement is irreversible for amplitude damping channels.

Specifically, the amplitude damping channel is defined by

$$\mathcal{N}_{\mathrm{ad}}(\rho) = E_0\rho E_0^{\dagger} + E_1\rho E_1^{\dagger}, \quad (96)$$

with Kraus operators $E_0 = |0\rangle\langle 0| + \sqrt{1-\gamma}|1\rangle\langle 1|$ and $E_1 = \sqrt{\gamma}|0\rangle\langle 1|$. Its quantum capacity, i.e., the maximal rate at which entanglement can be generated from the channel, is known as [68]

$$Q(\mathcal{N}_{\mathrm{ad}}) = \max_{p\in[0,1]} h_2((1-\gamma)p) - h_2(\gamma p), \quad (97)$$

where h_2 is the binary entropy. It has been shown in [41] that for $0.25 \lesssim \gamma < 1$,

$$E_{C,\mathrm{LOCC}}(\mathcal{N}_{\mathrm{ad}}) \geq E_{\mathrm{WJZ}}(\mathcal{N}_{\mathrm{ad}}(\Phi_{AA'})) > Q(\mathcal{N}_{\mathrm{ad}}), \quad (98)$$

Here, we can improve this bound by

$$E_{C,\mathrm{LOCC}}(\mathcal{N}_{\mathrm{ad}}) \geq D_{\mathrm{M}}(\mathcal{N}_{\mathrm{ad}}(\Phi_{AA'})\|\mathrm{PPT}_2) > Q(\mathcal{N}_{\mathrm{ad}}), \quad (99)$$

and show that the gap exists across the entire parameter region $0 < \gamma < 1$ in Figure 4.

C. Quantum entanglement distillation

Entanglement distillation is an essential quantum information processing task in quantum networks that involves converting multiple copies of noisy entangled states into a smaller number of Bell states. The *distillable entanglement* of a bipartite state ρ_{AB}, denoted by $E_{D,\Omega}(\rho_{AB})$, represents the maximum number of Bell states that can be extracted from the given state with asymptotically vanishing error under the operation class Ω. It has been established that the distillable entanglement under asymptotically non-entanglement generating operations, denoted by $E_{D,\mathrm{ANE}}$, is given by the *regularized relative entropy of entanglement* [18], [22], [23],

$$E_{D,\mathrm{ANE}}(\rho_{AB}) = D^{\infty}(\rho_{AB}\|\mathrm{SEP}) \quad (100)$$

$$:= \lim_{n\to\infty} \frac{1}{n} D(\rho_{AB}^{\otimes n}\|\mathrm{SEP}(A^n : B^n)), \quad (101)$$

where $\mathrm{SEP}(A : B)$ denotes the set of all separable states between $\mathcal{H}_A$ and $\mathcal{H}_B$. Evaluating this quantity is hard in general, as it involves a limit as well as the separability problem, which is known to be computationally hard [44].

As $D^{\infty}(\rho_{AB}\|\mathrm{SEP})$ is a minimization problem, any feasible solution gives an upper bound. Here, we can derive an efficient

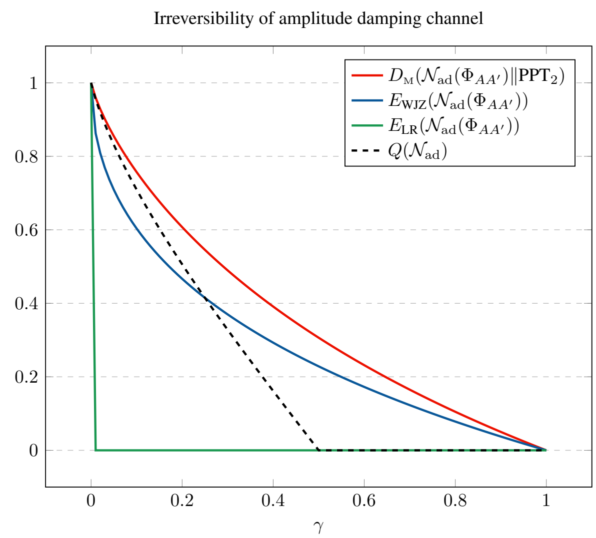

Fig. 4: Comparison of the new lower bound $D_{\mathrm{M}}(\mathcal{N}_{\mathrm{ad}}(\Phi_{AA'})\|\mathrm{PPT}_2)$ with the previously known bounds E_{WJZ} [41], E_{LR} [42] and the quantum capacity Q for amplitude damping channel. The horizontal axis is the channel parameter γ and the vertical axis is the value of the entanglement measure.

lower bound for $D^{\infty}(\rho_{AB}\|\mathrm{SEP})$ by relaxing SEP to the Rains set [29], [45] [3]

$$\mathrm{Rains}(A:B) := \left\{\sigma \in \mathscr{H}_+(AB) : \|\sigma^{\mathsf{T}_B}\|_1 \leq 1\right\}. \qquad (102)$$

We can show that it satisfies Assumption 9. This is because we have, by the SDP duality, that

$$h_{\mathrm{Rains}}(\omega) = \sup_{\sigma\in\mathrm{Rains}(A:B)} \mathrm{Tr}[\omega\sigma] = \inf_{\gamma\geq\omega} \|\gamma^{\mathsf{T}_B}\|_{\infty}, \qquad (103)$$

where $\|\cdot\|_{\infty}$ is the spectral norm. By the multiplicativity of $\|\cdot\|_{\infty}$ we can easily check that h_{Rains} is sub-multiplicative under tensor product, which is equivalent to the polar assumption (A.4) by Lemma 3. The rest of assumptions can also be easily verified.

Therefore, we have the relaxation

$$D^{\infty}(\rho_{AB}\|\mathrm{SEP}) \geq D^{\infty}(\rho_{AB}\|\mathrm{Rains}), \qquad (104)$$

where the right-hand side known as the regularized Rains bound can be efficiently estimated using Theorem 10 by considering $\mathscr{A}_n = \{\rho^{\otimes n}\}$ and $\mathscr{B}_n = \mathrm{Rains}(A^n : B^n)$ which has efficient SDP representations [30].

Remark 1 Similar to the regularized relative entropy of entanglement, the regularized Rains bound has the operational meaning [31] that it is the distillable entanglement under Rains-preserving operations $E_{D,\mathrm{Rains}}$, that is,

$$E_{D,\mathrm{Rains}}(\rho_{AB}) = D^{\infty}(\rho_{AB}\|\mathrm{Rains}). \qquad (105)$$

This marks the first time that an operational regularized entanglement measure has been shown to be efficiently computable, even when expressed as a regularized formula and beyond the zero-error setting. Previous work in [63] studied the zero-error entanglement cost under PPT operations and proved that it is efficiently computable despite the absence of a closed-form formula.

As the Rains-preserving operations is a superset of LOCC operations, the regularized Rains bound also gives an upper bound on the distillable entanglement under LOCC operations $E_{D,\mathrm{LOCC}}$, that is,

$$E_{D,\mathrm{LOCC}}(\rho_{AB}) \leq D^{\infty}(\rho_{AB}\|\mathrm{Rains}) \leq E_{D,\mathrm{ANE}}(\rho_{AB}). \qquad (106)$$

This improves the best known efficiently computable bound for $E_{D,\mathrm{LOCC}}$ as well.

D. Magic state distillation

The above argument for entanglement distillation also applies to magic states, which is a key resource for fault-tolerant quantum computing [32], [33], [34]. The task of magic state distillation aims to extract as many copies of the target magic state as possible with asymptotically vanishing error. The distillable magic is denoted by $M_{D,\Omega}$, where Ω represents the set of allowed operations. Typically, the most natural choice of operations involves stabilizer operations, and the corresponding distillable magic is denoted by $M_{D,\mathrm{STAB}}$. However, characterizing this set of operations is challenging.

Motivated by the idea of the Rains bound from entanglement theory, Wang et al. [36] relaxed the set of all stabilizer states to a set of sub-normalized states with non-positive mana, $\mathcal{W}(\mathcal{H}) := \{\sigma \in \mathscr{H}_+(\mathcal{H}) : \|\sigma\|_{W,1} \leq 1\}$, where $\|\cdot\|_{W,1}$ denotes the Wigner trace norm. Based on the set $\mathcal{W}$, Wang et al. introduced a magic measure $D(\rho\|\mathcal{W})$, called thauma, and proved that it serves as an upper bound for the distillable magic under stabilizer operations, that is,

$$M_{D,\mathrm{STAB}}(\rho) \leq D(\rho\|\mathcal{W})c(T), \qquad (107)$$

where $c(T)$ is a constant depending on the magic state T.

Here, we can verify that $\mathcal{W}$ satisfies our Assumption 9 and apply Theorem 10 to obtain an improved bound through regularization while keeping the computational efficiency.

To see this, it is straightforward to show that the support function of $\mathcal{W}$ is given by

$$h_{\mathcal{W}}(\omega) = \sup_{\sigma\in\mathcal{W}} \mathrm{Tr}[\omega\sigma] = \inf_{\gamma\geq\omega} \|\gamma\|_{W,\infty}, \qquad (108)$$

where the $\|\cdot\|_{W,\infty}$ is the Wigner spectral norm and the second equality follows from the SDP duality. Since $\|\cdot\|_{W,\infty}$ is multiplicative under tensor product, we can verify that the support function $h_{\mathcal{W}}$ is sub-multiplicative under tensor product as well. Hence, the polar set $\mathcal{W}^{\circ}$ is closed under tensor product by Lemma 3, and the remaining assumptions in Assumption 9 can also be verified. Then, we can consider the regularization and get

$$M_{D,\mathrm{STAB}}(\rho) \leq D^{\infty}(\rho\|\mathcal{W})c(T), \qquad (109)$$

where $D^{\infty}(\rho\|\mathcal{W})$ is the reguarlized thauma which remains efficiently computable by applying Theorem 10 with $\mathscr{A}_n = \{\rho^{\otimes n}\}$ and $\mathscr{B}_n = \mathcal{W}(\mathcal{H}^{\otimes n})$. This improves the best-known estimation of magic state distillation under stablizer opeations.

[3] The set of PPT states does not satisfy Assumption 9, see discussion after Eq. (61).

V. Conclusion

We showed that regularized relative entropy between two sets of quantum states can be efficiently computed using convex optimization. This result has broad implications for quantum information theory, including the study of adversarial quantum channel discrimination, the estimation of the entanglement cost of quantum states and channels, entanglement distillation under LOCC operations and magic state distillation under stabilizer operations. Numerical experiments demonstrated that our new bounds outperform existing ones in various scenarios, even for the first level of approximation. Generally, our result can be applied by verifying the conditions of the relevant theory and performing necessary relaxations when required. Therefore, we anticipate that this approach has the potential for far-reaching applications beyond the specific cases discussed here.

Many problems remain open for future investigation. For example, while we have demonstrated that regularized relative entropies can be efficiently computed using convex optimization techniques, developing a more explicit algorithm and its implementation remains an area for further exploration. Additionally, designing a general algorithm to construct the smallest superset of a given set that satisfies the polar assumption presents an intriguing challenge. The solution of this would extend the applicability of our results to broader areas.

Acknowledgment

K.F. thanks Xin Wang for the updates on the PPT_k sets, and Ludovico Lami and Bartosz Regula for pointing out that E_{LR} vanishes for any full-rank state.

References

[1] S. Kullback and R. A. Leibler, "On information and sufficiency," *The Annals of Mathematical Statistics*, vol. 22, no. 1, pp. 79–86, 1951.
[2] K. P. Murphy, *Machine learning: a probabilistic perspective*. MIT Press, 2012.
[3] T. M. Cover, *Elements of information theory*. John Wiley & Sons, 1999.
[4] R. M. Gray, *Entropy and information theory*. Springer Science & Business Media, 2011.
[5] E. T. Jaynes, "Information theory and statistical mechanics," *Physical Review*, vol. 106, no. 4, p. 620, 1957.
[6] G. E. Crooks, "Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences," *Physical Review E*, vol. 60, no. 3, p. 2721, 1999.
[7] H. Umegaki, "Conditional expectation in an operator algebra, IV (entropy and information)," in *Kodai Mathematical Seminar Reports*, vol. 14, no. 2. Department of Mathematics, Tokyo Institute of Technology, 1962, pp. 59–85.
[8] F. Hiai and D. Petz, "The proper formula for relative entropy and its asymptotics in quantum probability," *Communications in Mathematical Physics*, vol. 143, pp. 99–114, 1991.
[9] B. Schumacher and M. D. Westmoreland, "Relative entropy in quantum information theory," *Contemporary Mathematics*, vol. 305, pp. 265–290, 2002.
[10] V. Vedral, "The role of relative entropy in quantum information theory," *Reviews of Modern Physics*, vol. 74, no. 1, p. 197, 2002.
[11] J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, and S. Lloyd, "Quantum machine learning," *Nature*, vol. 549, no. 7671, pp. 195–202, 2017.
[12] M. Hayashi, "Quantum information theory," *Graduate Texts in Physics, Springer*, 2017.
[13] N. J. Cerf and R. Cleve, "Information-theoretic interpretation of quantum error-correcting codes," *Physical Review A*, vol. 56, no. 3, p. 1721, 1997.
[14] E. Chitambar and G. Gour, "Quantum resource theories," *Reviews of Modern Physics*, vol. 91, no. 2, p. 025001, 2019.
[15] S. Pirandola, U. L. Andersen, L. Banchi, M. Berta, D. Bunandar, R. Colbeck, D. Englund, T. Gehring, C. Lupo, C. Ottaviani *et al.*, "Advances in quantum cryptography," *Advances in Optics and Photonics*, vol. 12, no. 4, pp. 1012–1236, 2020.
[16] K. Fang, H. Fawzi, and O. Fawzi, "Generalized quantum asymptotic equipartition," *arXiv preprint arXiv:2411.04035*, 2024.
[17] H. Nagaoka and T. Ogawa, "Strong converse and Stein's lemma in quantum hypothesis testing," *IEEE Transactions on Information Theory*, vol. 46, no. 7, pp. 2428–2433, feb 2000. [Online]. Available: http://ieeexplore.ieee.org/document/887855/
[18] F. G. S. L. Brandão and M. B. Plenio, "A generalization of quantum stein's lemma," *Communications in Mathematical Physics*, vol. 295, no. 3, p. 791–828, Feb. 2010. [Online]. Available: http://dx.doi.org/10.1007/s00220-010-1005-z
[19] F. G. Brandão, A. W. Harrow, J. R. Lee, and Y. Peres, "Adversarial hypothesis testing and a quantum Stein's lemma for restricted measurements," *IEEE Transactions on Information Theory*, vol. 66, no. 8, pp. 5037–5054, 2020.
[20] M. Berta, F. G. Brandão, and C. Hirche, "On composite quantum hypothesis testing," *Communications in Mathematical Physics*, vol. 385, no. 1, pp. 55–77, 2021.
[21] M. Mosonyi, Z. Szilagyi, and M. Weiner, "On the error exponents of binary state discrimination with composite hypotheses," *IEEE Transactions on Information Theory*, vol. 68, no. 2, p. 1032–1067, Feb. 2022. [Online]. Available: http://dx.doi.org/10.1109/TIT.2021.3125683
[22] M. Hayashi and H. Yamasaki, "The generalized quantum Stein's lemma and the second law of quantum resource theories," *Nature Physics*, vol. 21, no. 12, pp. 1988–1993, 2025. [Online]. Available: https://doi.org/10.1038/s41567-025-03047-9
[23] L. Lami, "A solution of the generalized quantum Stein's lemma," *IEEE Transactions on Information Theory*, vol. 71, no. 6, p. 4454–4484, Jun. 2025. [Online]. Available: https://doi.org/10.1109/TIT.2025.3543610
[24] K. Fang, O. Fawzi, R. Renner, and D. Sutter, "Chain rule for the quantum relative entropy," *Physical Review Letters*, vol. 124, no. 10, p. 100501, 2020.
[25] K. Fang, G. Gour, and X. Wang, "Towards the ultimate limits of quantum channel discrimination and quantum communication," *SCIENCE CHINA Information Sciences*, vol. 68, no. 8, pp. 180 509:1–180 509:25, 2025. [Online]. Available: https://link.springer.com/article/10.1007/s11432-024-4488-0
[26] K. Fang, H. Fawzi, and O. Fawzi, "Adversarial quantum channel discrimination," *Physical Review Letters*, vol. 135, p. 260201, Dec 2025. [Online]. Available: https://link.aps.org/doi/10.1103/mtbj-grbk
[27] F. Dupuis, O. Fawzi, and R. Renner, "Entropy accumulation," *Communications in Mathematical Physics*, vol. 379, no. 3, pp. 867–913, 2020.
[28] T. Metger, O. Fawzi, D. Sutter, and R. Renner, "Generalised entropy accumulation," in *2022 IEEE 63rd Annual Symposium on Foundations of Computer Science (FOCS)*. IEEE, 2022, pp. 844–850.
[29] E. M. Rains, "A semidefinite program for distillable entanglement," *IEEE Transactions on Information Theory*, vol. 47, no. 7, pp. 2921–2933, 2001.
[30] K. Fang, X. Wang, M. Tomamichel, and R. Duan, "Non-asymptotic entanglement distillation," *IEEE Transactions on Information Theory*, vol. 65, no. 10, pp. 6454–6465, 2019.
[31] B. Regula, K. Fang, X. Wang, and M. Gu, "One-shot entanglement distillation beyond local operations and classical communication," *New Journal of Physics*, vol. 21, no. 10, p. 103017, 2019.
[32] S. Bravyi and A. Kitaev, "Universal quantum computation with ideal clifford gates and noisy ancillas," *Physical Review A*, vol. 71, no. 2, Feb. 2005. [Online]. Available: http://dx.doi.org/10.1103/PhysRevA.71.022316
[33] V. Veitch, S. H. Mousavian, D. Gottesman, and J. Emerson, "The resource theory of stabilizer quantum computation," *New Journal of Physics*, vol. 16, no. 1, p. 013009, 2014.
[34] K. Fang and Z.-W. Liu, "No-go theorems for quantum resource purification," *Physical Review Letters*, vol. 125, no. 6, p. 060405, 2020.
[35] ——, "No-go theorems for quantum resource purification: New approach and channel theory," *PRX Quantum*, vol. 3, no. 1, p. 010337, 2022.
[36] X. Wang, M. M. Wilde, and Y. Su, "Efficiently computable bounds for magic state distillation," *Physical Review Letters*, vol. 124, no. 9, p. 090505, 2020.
[37] M. Hayashi, K. Fang, and K. Wang, "Finite block length analysis on quantum coherence distillation and incoherent randomness extraction,"

IEEE Transactions on Information Theory, vol. 67, no. 6, pp. 3926–3944, 2021.

[38] M. G. Díaz, K. Fang, X. Wang, M. Rosati, M. Skotiniotis, J. Calsamiglia, and A. Winter, "Using and reusing coherence to realize quantum processes," *Quantum*, vol. 2, p. 100, 2018.

[39] X. Wang and R. Duan, "Nonadditivity of Rains' bound for distillable entanglement," *Physical Review A*, vol. 95, no. 6, pp. 1–5, 2017.

[40] ——, "Irreversibility of Asymptotic Entanglement Manipulation under Quantum Operations Completely Preserving Positivity of Partial Transpose," *Physical Review Letters*, vol. 119, no. 18, pp. 1–7, 2017.

[41] X. Wang, M. Jing, and C. Zhu, "Computable and faithful lower bound on entanglement cost," *Physical Review Letters*, vol. 134, p. 190202, May 2025. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.134.190202

[42] L. Lami and B. Regula, "No second law of entanglement manipulation after all," *Nature Physics*, vol. 19, no. 2, pp. 184–189, 2023.

[43] K. Fang, J. Zhao, X. Li, Y. Li, and R. Duan, "Quantum NETwork: from theory to practice," *Science China Information Sciences*, vol. 66, no. 8, p. 180509, 2023.

[44] L. Gurvits, "Classical deterministic complexity of edmonds' problem and quantum entanglement," in *Proceedings of the thirty-fifth annual ACM symposium on Theory of computing*, 2003, pp. 10–19.

[45] K. Audenaert, B. De Moor, K. G. H. Vollbrecht, and R. F. Werner, "Asymptotic relative entropy of entanglement for orthogonally invariant states," *Physical Review A*, vol. 66, no. 3, p. 032310, 2002.

[46] H. Umegaki, "Conditional expectation in an operator algebra," *Tohoku Mathematical Journal, Second Series*, vol. 6, no. 2-3, pp. 177–181, 1954.

[47] N. Datta, "Min-and max-relative entropies and a new entanglement monotone," *IEEE Transactions on Information Theory*, vol. 55, no. 6, pp. 2816–2826, 2009.

[48] M. J. Donald, "On the relative entropy," *Communications in Mathematical Physics*, vol. 105, pp. 13–34, 1986.

[49] M. Berta, O. Fawzi, and M. Tomamichel, "On variational expressions for quantum relative entropies," *Letters in Mathematical Physics*, vol. 107, no. 12, pp. 2239–2265, 2017.

[50] J. Watrous, *The theory of quantum information*. Cambridge University Press, 2018.

[51] H. Fawzi and J. Saunderson, "Optimal self-concordant barriers for quantum relative entropies," *SIAM Journal on Optimization*, vol. 33, no. 4, pp. 2858–2884, 2023.

[52] K. He, J. Saunderson, and H. Fawzi, "Interior point methods for structured quantum relative entropy optimization problems," *Mathematical Programming Computation*, 2025. [Online]. Available: https://doi.org/10.1007/s12532-025-00296-w

[53] ——, "QICS: Quantum Information Conic Solver," *arXiv preprint arXiv:2410.17803*, 2024.

[54] A. Ben-Tal and A. Nemirovski, *Lectures on modern convex optimization: analysis, algorithms, and engineering applications*. SIAM, 2001.

[55] J. Gouveia, P. A. Parrilo, and R. R. Thomas, "Lifts of convex sets and cone factorizations," *Mathematics of Operations Research*, vol. 38, no. 2, pp. 248–264, 2013.

[56] V. Chandrasekaran and P. Shah, "Relative entropy optimization and its applications," *Mathematical Programming*, vol. 161, pp. 1–32, 2017.

[57] T. Netzer, "Spectrahedra and their shadows," 2012, habilitation thesis, available at https://www.uibk.ac.at/mathematik/algebra/media/tim-netzer/habilitationsschrift.pdf.

[58] O. Fawzi, A. Shayeghi, and H. Ta, "A hierarchy of efficient bounds on quantum capacities exploiting symmetry," *IEEE Transactions on Information Theory*, vol. 68, no. 11, p. 7346–7360, Nov. 2022. [Online]. Available: http://dx.doi.org/10.1109/TIT.2022.3182101

[59] B. Litjens, S. Polak, and A. Schrijver, "Semidefinite bounds for non-binary codes based on quadruples," *Designs, Codes and Cryptography*, vol. 84, pp. 87–100, 2017.

[60] F. Leditzky, D. Leung, V. Siddhu, G. Smith, and J. A. Smolin, "The platypus of the quantum channel zoo," *IEEE Transactions on Information Theory*, vol. 69, no. 6, p. 3825–3849, Jun. 2023. [Online]. Available: http://dx.doi.org/10.1109/TIT.2023.3245985

[61] Y. Huang, "Computing quantum discord is NP-complete," *New Journal of Physics*, vol. 16, no. 3, p. 33027, 2014.

[62] M. Christandl and A. Winter, "“Squashed entanglement”: an additive entanglement measure," *Journal of Mathematical Physics*, vol. 45, no. 3, pp. 829–840, 2004.

[63] L. Lami, F. A. Mele, and B. Regula, "Computable entanglement cost under positive partial transpose operations," *Physical Review Letters*, vol. 134, p. 090202, Mar 2025. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.134.090202

[64] B. Regula and L. Lami, "Functional analytic insights into irreversibility of quantum resources," *arXiv preprint arXiv:2211.15678*, 2022.

[65] E. Rains, "An improved bound on distillable entanglement," *arXiv preprint quant-ph/9809082*, 1998.

[66] K. Audenaert, J. Eisert, E. Jane, M. B. Plenio, S. Virmani, and B. De Moor, "Asymptotic relative entropy of entanglement," *Physical Review Letters*, vol. 87, no. 21, p. 217902, 2001.

[67] M. Berta, F. G. Brandão, M. Christandl, and S. Wehner, "Entanglement cost of quantum channels," *IEEE Transactions on Information Theory*, vol. 59, no. 10, pp. 6779–6795, 2013.

[68] V. Giovannetti and R. Fazio, "Information-capacity description of spin-chain correlations," *Physical Review A—Atomic, Molecular, and Optical Physics*, vol. 71, no. 3, p. 032314, 2005.

Kun Fang received the BS degree in Mathematics from Wuhan University in 2015 and PhD degree in Quantum Information from the University of Technology Sydney in 2018. After that, he worked as a postdoctoral researcher at the University of Cambridge and the University of Waterloo from 2018 to 2020. He served as a senior researcher and tech lead at the Institute for Quantum Computing, Baidu, from 2020 to 2023. Currently, he is an assistant professor in the School of Data Science at The Chinese University of Hong Kong, Shenzhen. His research interests lie in understanding the capabilities and limitations of quantum resources, as well as their usage in quantum computation and quantum communication.

Hamza Fawzi received his Ph.D. degree from MIT in 2016. Since then he has been at the University of Cambridge, where is currently a Professor of Applied Mathematics. His research interests include convex optimization and its applications to quantum information and many-body theory.

Omar Fawzi received the Ph.D. degree in computer science from McGill University in 2012. From 2012 to 2014, he was a postdoctoral researcher at ETH Zürich. From 2014 to 2020, he was an Associate Professor of computer science with ENS de Lyon, where he is currently a Research Director with Inria. His research interests include classical and quantum information with an emphasis on algorithmic questions.