Commun. Math. Phys. 385, 55–77 (2021)
Digital Object Identifier (DOI) https://doi.org/10.1007/s00220-021-04133-8

Communications in
**Mathematical
Physics**

On Composite Quantum Hypothesis Testing

Mario Berta[1,2,3], **Fernando G. S. L. Brandão**[2,3], **Christoph Hirche**[4]

[1] Department of Computing, Imperial College London, London SW7 2AZ, UK
[2] IQIM, California Institute of Technology, Pasadena, CA 91125, USA
[3] AWS Center for Quantum Computing, Pasadena, CA 91125, USA
[4] QMATH, Department of Mathematical Sciences, University of Copenhagen, Copenhagen, Denmark
E-mail: christoph.hirche@gmail.com

Received: 3 April 2020 / Accepted: 31 May 2021
Published online: 10 June 2021 – © The Author(s) 2021

Abstract: We extend quantum Stein's lemma in asymmetric quantum hypothesis testing to composite null and alternative hypotheses. As our main result, we show that the asymptotic error exponent for testing convex combinations of quantum states $\rho^{\otimes n}$ against convex combinations of quantum states $\sigma^{\otimes n}$ can be written as a regularized quantum relative entropy formula. We prove that in general such a regularization is needed but also discuss various settings where our formula as well as extensions thereof become single-letter. This includes an operational interpretation of the relative entropy of coherence in terms of hypothesis testing. For our proof, we start from the composite Stein's lemma for classical probability distributions and lift the result to the non-commutative setting by using elementary properties of quantum entropy. Finally, our findings also imply an improved recoverability lower bound on the conditional quantum mutual information in terms of the regularized quantum relative entropy—featuring an explicit and universal recovery map.

1. Overview of Results

Hypothesis testing is arguably one of the most fundamental primitives in quantum information theory. As such it has found many applications, e.g., in quantum channel coding [27] and quantum illumination [37,46,56], or for giving an operational interpretation to abstract quantities [13,16,28]. A particular hypothesis testing setting is that of quantum state discrimination where quantum states are assigned to each of the hypotheses and we aim to determine which state is actually given. Several distinct scenarios are of interest, which differ in the priority given to different types of error or in how many copies of a system are given to aid the discrimination. Here, we investigate the setting of asymmetric hypothesis testing where the goal is to discriminate between two n-party quantum states (strategies or hypotheses) ρ_n and σ_n living on the n-fold tensor product of some finite-dimensional inner product space $\mathcal{H}^{\otimes n}$. That is, we are optimizing over all two-outcome positive operator valued measures (POVMs) with $\{M_n, (1 - M_n)\}$ and associate M_n with accepting ρ_n as well as $(1 - M_n)$ with accepting σ_n. This naturally gives rise to the two possible errors

$$\alpha_n(M_n) := \operatorname{Tr}\left[\rho_n(1 - M_n)\right] \qquad \text{Type 1 error,} \tag{1}$$

$$\beta_n(M_n) := \operatorname{Tr}\left[\sigma_n M_n\right] \qquad \text{Type 2 error.} \tag{2}$$

For asymmetric hypothesis testing we minimize the Type 2 error as[1]

$$\beta(n, \varepsilon) := \inf_{0 \ll M_n \ll 1} \left\{ \beta_n(M_n) \middle| \alpha_n(M_n) \leq \varepsilon \right\} \tag{3}$$

while we require the Type 1 error not to exceed a small constant $\varepsilon \in (0, 1)$. We are then interested in finding the optimal error exponent[2]

$$\zeta(n, \varepsilon) := -\frac{\log \beta(n, \varepsilon)}{n}, \tag{4}$$

and its asymptotic limits

$$\zeta(\infty, \varepsilon) := \lim_{n \to \infty} -\frac{\log \beta(n, \varepsilon)}{n}, \quad \zeta(\infty, 0) := \lim_{\varepsilon \to 0} \zeta(\infty, \varepsilon)\,. \tag{5}$$

A well studied discrimination setting is that between fixed independent and identical (iid) states $\rho^{\otimes n}$ and $\sigma^{\otimes n}$, where the asymptotic error exponent is determined by the quantum Stein's lemma [4,30,43] in terms of the quantum relative entropy. Namely, we denote this special case of Eq. (5) by $\zeta_{\rho,\sigma}(\infty, \varepsilon)$ and the Stein's lemma then gives for any $\varepsilon \in (0, 1)$ the formula

$$\zeta_{\rho,\sigma}(\infty, \varepsilon) = D(\rho\|\sigma) := \begin{cases} \operatorname{Tr}\left[\rho\left(\log \rho - \log \sigma\right)\right] & \operatorname{supp}(\rho) \subseteq \operatorname{supp}(\sigma) \\ \infty & \text{otherwise.} \end{cases} \tag{6}$$

In many applications we aim to solve more general discrimination problems and a prominent example of such is that of composite hypotheses—in which we attempt to discriminate between different sets of states. Previously the case of composite iid null hypotheses $\rho^{\otimes n}$ with $\rho \in \mathcal{S}$ and corresponding asymptotic error exponent $\zeta_{\mathcal{S},\sigma}(\infty, \varepsilon)$ was studied in [10,25], leading to the formula

$$\zeta_{\mathcal{S},\sigma}(\infty, \varepsilon) = \inf_{\rho \in S} D(\rho\|\sigma) \quad \forall \varepsilon \in (0, 1)\,. \tag{7}$$

On the other hand, the problem of composite alternative hypotheses is more involved in the non-commutative case. When the set of alternative hypotheses $\mathcal{T}_n$ for $n \in \mathbb{N}$ fulfils certain axioms motivated by the framework of resource theories, it was shown in [13] that the corresponding asymptotic error exponent $\zeta_{\rho,\mathcal{T}}(\infty, \varepsilon)$ is written in terms of the regularized relative entropy distance as

$$\zeta_{\rho,\mathcal{T}}(\infty, \varepsilon) = \lim_{n \to \infty} \frac{1}{n} \inf_{\sigma_n \in \mathcal{T}_n} D\left(\rho^{\otimes n} \|\sigma_n\right) \quad \forall \varepsilon \in (0, 1)\,. \tag{8}$$

[1] Here and henceforth $\ll$ denotes the Loewner order.

[2] Here and henceforth the logarithm is defined with respect to the basis 2.

This regularization is in general needed as we know from the case of the relative entropy of entanglement [54]. Note that this might not be too surprising since the set of alternative hypotheses is not required to be iid in general.

For our main result, we consider the setting where null and alternative hypotheses are both composite and given by convex combinations of n-fold tensor powers of states from given convex, closed sets $\mathcal{S}$ and $\mathcal{T}$. More precisely, for $n \in \mathbb{N}$ we attempt the following discrimination problem.[3]

Null hypothesis: the convex hull of iid states

$$\mathcal{S}_n := \left\{ \int \rho^{\otimes n} \, \mathrm{d}\nu(\rho) \middle| \nu \in \mathcal{S} \right\} \text{ with } \mathcal{S} \subseteq S(\mathcal{H}) \text{ convex and closed} \tag{9}$$

Alternative hypothesis: the convex hull of iid states

$$\mathcal{T}_n := \left\{ \int \sigma^{\otimes n} \, \mathrm{d}\mu(\sigma) \middle| \mu \in \mathcal{T} \right\} \text{ with } \mathcal{T} \subseteq S(\mathcal{H}) \text{ convex and closed} \tag{10}$$

Slightly abusing the notation, $\nu \in \mathcal{S}$ and $\mu \in \mathcal{T}$ stand for probability measures on the Borel σ-algebra of $\mathcal{S}$ and $\mathcal{T}$, respectively. For $\varepsilon \in (0, 1)$ the goal is to determine the optimal error exponent for composite asymmetric hypothesis testing

$$\zeta_{\mathcal{S},\mathcal{T}}(n,\varepsilon) := -\frac{1}{n} \log \inf_{0 \ll M_n \ll 1} \left\{ \sup_{\mu \in \mathcal{T}} \mathrm{Tr}\left[M_n \sigma_n(\mu)\right] \middle| \sup_{\nu \in \mathcal{S}} \mathrm{Tr}\left[(1 - M_n)\rho_n(\nu)\right] \leq \varepsilon \right\} \tag{11}$$

with the abbreviations

$$\rho_n(\nu) := \int \rho^{\otimes n} \mathrm{d}\nu(\rho) \quad \text{and} \quad \sigma_n(\mu) := \int \sigma^{\otimes n} \mathrm{d}\mu(\sigma) \, . \tag{12}$$

It is trivial to see that $\zeta_{\mathcal{S},\mathcal{T}}(n,\varepsilon)$ equivalently gives the error exponent of testing between $\mathcal{S}^{\otimes n} := \{\rho^{\otimes n} | \rho \in \mathcal{S}\}$ and $\mathcal{T}^{\otimes n} := \{\sigma^{\otimes n} | \sigma \in \mathcal{T}\}$. This then explicitly takes the form of an iid problem. The following is our main result, which we prove in Sect. 2 under the support condition

$$\mathrm{supp}(\rho) \subseteq \mathrm{supp}(\sigma) \quad \forall \rho \in \mathcal{S} \quad \forall \sigma \in \mathcal{T} \, . \tag{13}$$

Theorem 1.1. *For the discrimination problem as above, we have*

$$\lim_{\varepsilon \to 0} \liminf_{n \to \infty} \zeta_{\mathcal{S},\mathcal{T}}(n,\varepsilon) = \lim_{\varepsilon \to 0} \limsup_{n \to \infty} \zeta_{\mathcal{S},\mathcal{T}}(n,\varepsilon) \tag{14}$$

$$= \lim_{n \to \infty} \frac{1}{n} \inf_{\substack{\rho \in \mathcal{S} \\ \mu \in \mathcal{T}}} D\left(\rho^{\otimes n} \middle\| \int \sigma^{\otimes n} \, \mathrm{d}\mu(\sigma)\right) \tag{15}$$

[3] Here and henceforth all inner product spaces $\mathcal{H}$ are finite-dimensional and $S(\mathcal{H})$ denotes the set of unit trace positive semi-definite linear operators on $\mathcal{H}$.

Our proof can be found in Sect. 2 and has a clear structure in the sense that we start from the composite Stein's lemma for classical probability distributions and then lift the result to the non-commutative setting by using elementary properties of entropic measures. We emphasise that even in the case of a fixed null hypothesis $\mathcal{S} = \{\rho\}$ our setting is not a special case of the previous results [13], as our sets of alternative hypotheses are not closed under tensor product

$$\sigma_m \in \mathcal{T}_m,\ \sigma_n \in \mathcal{T}_n \not\Rightarrow \sigma_m \otimes \sigma_n \in \mathcal{T}_{mn}\,, \tag{16}$$

which is one of the properties required for the results in [13].

We show that in contrast to the finite classical case [11,34], the regularization in Eq. (15) is needed in general. That is, we provide an explicit example for which the non-regularized relative entropy formula is not an achievable asymptotic error exponent

$$\inf_{\substack{\rho\in\mathcal{S}\\ \sigma\in\mathcal{T}}} D(\rho\|\sigma) > \lim_{\varepsilon\to 0}\limsup_{n\to\infty} \zeta_{\mathcal{S},\mathcal{T}}(n,\varepsilon)\,. \tag{17}$$

In particular, we find that, even for $n \to \infty$, in general

$$\frac{1}{n}\inf_{\mu\in\mathcal{T}} D\Big(\rho^{\otimes n}\Big\|\int \sigma^{\otimes n}\,\mathrm{d}\mu(\sigma)\Big) \neq \inf_{\sigma\in\mathcal{T}} D(\rho\|\sigma)\,, \tag{18}$$

thereby providing a counterexample to this conjectured quantum entropy inequality (see [12, Equation (20)] for a variant) which holds in the finite classical setting (see, e.g., [51, Lemma 3.11]).[4] Note that the $\leq$ direction in Eq. (18) holds trivially.

Nevertheless, there exist non-commutative cases in which the regularization in Eq. (15) is not needed and we discuss several such examples. In particular, we give an operational interpretation of the relative entropy of coherence in terms of hypothesis testing.

Finally, we apply the techniques developed in this work to strengthen previously known quantum relative entropy lower bounds on the conditional quantum mutual information [9,12,22,32,50,51,55]

$$I(A:B|C)_\rho := H(AC)_\rho + H(BC)_\rho - H(ABC)_\rho - H(C)_\rho \tag{19}$$

with $H(C)_\rho := -\operatorname{Tr}\left[\rho_C \log \rho_C\right]$ the von Neumann entropy. We find that

$$I(A:B|C)_\rho \geq \limsup_{n\to\infty}\frac{1}{n} D\Big(\rho_{ABC}^{\otimes n}\Big\|\int \beta_0(t)\,\mathrm{d}t\big(\mathcal{I}_A\otimes\mathcal{R}^{[t]}_{C\to BC}(\rho_{AC})\big)^{\otimes n}\Big) \tag{20}$$

for some universal probability distribution $\beta_0(t)$ and the rotated Petz recovery maps $R^{[t]}_{C\to BC}$ as defined in Sect. 4. In contrast to the previously known bounds in terms of the quantum relative entropy [12,51], the recovery map in Eq. (20) takes a specific form only depending on the reduced state on BC. Note that the regularization in Eq. (20) cannot go away in relative entropy distance, as recently shown in [21]. We end with an overview how all known recoverability lower bounds on the conditional quantum mutual information compare and argue that Eq. (20) represents the last possible strengthening.

The remainder of the paper is structured as follows. In Sect. 2 we prove our main result about composite asymmetric hypothesis testing. This is followed by Sect. 3 where we

[4] After completion of the first version of our work, even simpler examples of composite hypothesis testing problems with no single-letter solution were provided in [38].

discuss several concrete examples including an operational interpretation of the relative entropy of coherence, as well as its Rényi analogues in terms of the Petz divergences [44] and the sandwiched relative entropies [39,57]. In Sect. 4 we prove the refined lower bound on the conditional mutual information from Eq. (20) and use it to show that the regularization in Eq. (15) is needed in general. Finally, we end in Sect. 5 with a discussion of some open questions.

2. Proof of Main Result

In the following we give a proof of our main result Theorem 1.1. We first prove the converse, meaning the $\leq$ direction of Theorem 1.1, which follows from the following proposition.

Proposition 2.1. *For $\rho \in \mathcal{S}$, $\mu \in \mathcal{T}$, and $\varepsilon \in (0,1)$ we have*

$$\zeta_{\mathcal{S},\mathcal{T}}(n,\varepsilon) \leq \inf_{\substack{\rho\in\mathcal{S}\\ \mu\in\mathcal{T}}} \frac{1}{n}\frac{D\left(\rho^{\otimes n}\|\sigma_n(\mu)\right)+1}{1-\varepsilon}\,. \tag{21}$$

Proof. We follow the original converse proof of quantum Stein's lemma [30] for the states $\rho^{\otimes n}$ and $\sigma_n(\mu)$. By the monotonicity of the quantum relative entropy [36], we have for the measurement $\{M_n, (1-M_n)\}$ that

$$\begin{aligned} &D\left(\rho^{\otimes n}\|\sigma_n(\mu)\right)\\ &\geq \operatorname{Tr}\left[M_n\rho^{\otimes n}\right]\log\frac{\operatorname{Tr}\left[M_n\rho^{\otimes n}\right]}{\operatorname{Tr}\left[M_n\sigma_n(\mu)\right]}+\left(1-\operatorname{Tr}\left[M_n\rho^{\otimes n}\right]\right)\log\frac{1-\operatorname{Tr}\left[M_n\rho^{\otimes n}\right]}{1-\operatorname{Tr}\left[M_n\sigma_n(\mu)\right]} &(22)\\ &\geq -\log 2-\operatorname{Tr}\left[M_n\rho^{\otimes n}\right]\log\operatorname{Tr}\left[M_n\sigma_n(\mu)\right] &(23)\\ &\geq -1-\inf_{\rho\in\mathcal{S}}\operatorname{Tr}\left[M_n\rho^{\otimes n}\right]\log\sup_{\mu\in\mathcal{T}}\operatorname{Tr}\left[M_n\sigma_n(\mu)\right] &(24)\\ &\geq -1-(1-\varepsilon)\log\sup_{\mu\in\mathcal{T}}\operatorname{Tr}\left[M_n\sigma_n(\mu)\right] &(25)\end{aligned}$$

leading to

$$-\frac{1}{n}\log\sup_{\mu\in\mathcal{T}}\operatorname{Tr}\left[M_n\sigma_n(\mu)\right] \leq \frac{1}{n}\frac{D\left(\rho^{\otimes n}\|\sigma_n(\mu)\right)+1}{1-\varepsilon} \tag{26}$$

for any $\rho\in\mathcal{S}$, $\mu\in\mathcal{T}$, and $0 \ll M_n \ll 1$ such that $\sup_{\rho\in\mathcal{S}}\operatorname{Tr}\left[(1-M_n)\rho^{\otimes n}\right]\leq\varepsilon$. Taking the supremum over all such M_n and then the infimum over $\rho\in\mathcal{S}$ and $\mu\in\mathcal{T}$ leads to the desired result. □

By taking the appropriate limits in Proposition 2.1, we immediately find the converse statements

$$\lim_{\varepsilon\to 0}\limsup_{n\to\infty}\zeta_{\mathcal{S},\mathcal{T}}(n,\varepsilon) \leq \limsup_{n\to\infty}\frac{1}{n}\inf_{\substack{\rho\in\mathcal{S}\\ \mu\in\mathcal{T}}} D\left(\rho^{\otimes n}\|\sigma_n(\mu)\right) \tag{27}$$

$$\lim_{\varepsilon\to 0}\liminf_{n\to\infty}\zeta_{\mathcal{S},\mathcal{T}}(n,\varepsilon) \leq \liminf_{n\to\infty}\frac{1}{n}\inf_{\substack{\rho\in\mathcal{S}\\ \mu\in\mathcal{T}}} D\left(\rho^{\otimes n}\|\sigma_n(\mu)\right)\,. \tag{28}$$

Remark 2.2. In information-theoretic language Eq. (21) represents a weak converse, i.e. the limit $\varepsilon \to 0$ in Eqs. (27) and (28) is required, and one might be tempted to derive a strong converse that holds for all $\varepsilon \in (0, 1)$ by employing quantum versions of the Rényi relative entropies [39,44,57] or the smooth max-relative entropy [18,31]. However, because of the missing $\sigma_m \in \mathcal{T}_m,\ \sigma_n \in \mathcal{T}_n \not\Rightarrow \sigma_m \otimes \sigma_n \in \mathcal{T}_{mn}$ property, the convergence of aforementioned measures to the quantum relative entropy remains unclear (see, e.g., [3,13,17,19] for corresponding techniques in the context of quantum hypothesis testing). As such, we leave open the question about a strong converse.

For the proof of the achievability, meaning the $\geq$ direction in Theorem 1.1, the basic idea is to start from the corresponding composite Stein's lemma for classical probability distributions and lift the result to the non-commutative setting by solely using properties of quantum entropy. For that we need the measured relative entropy defined as [20,30]

$$D_{\mathcal{M}}(\rho\|\sigma) := \sup_{(\mathcal{X},\mathcal{M})} D\Big(\underbrace{\sum_{x\in\mathcal{X}} \mathrm{Tr}\,[M_x\rho]\,|x\rangle\langle x|}_{=\mathcal{M}(\rho)} \Big\| \underbrace{\sum_{x\in\mathcal{X}} \mathrm{Tr}\,[M_x\sigma]\,|x\rangle\langle x|}_{=\mathcal{M}(\sigma)}\Big)\,, \tag{29}$$

where the optimization is over finite sets $\mathcal{X}$ and measurements $\mathcal{M}$ on $\mathcal{X}$ with $\mathrm{Tr}\,[M_x\rho]$ a measure on $\mathcal{X}$. Henceforth, we write for the classical relative entropy between probability distributions $D(P\|Q)$—defined via the diagonal embedding of P and Q as on the right-hand side of Eq. (29). It is known that we can restrict the a priori unbounded supremum to rank-one projective measurements [7, Theorem 2]. We now prove the achievability direction in Theorem 1.1 in several steps and start with an achievability bound in terms of the measured relative entropy.

Lemma 2.3. *For definitions as above and* $\varepsilon \in (0, 1)$*, we have*

$$\liminf_{n\to\infty} \zeta_{\mathcal{S},\mathcal{T}}(n,\varepsilon) \geq \sup_{k\in\mathbb{N}} \frac{1}{k} \inf_{\substack{\nu\in\mathcal{S}\\ \mu\in\mathcal{T}}} D_{\mathcal{M}}\left(\rho_k(\nu)\|\sigma_k(\mu)\right)\,. \tag{30}$$

Proof. For sets of classical probability distributions $\mathcal{S}$ and $\mathcal{T}$, we get from the corresponding commutative achievability result that for $\delta > 0$ and $\varepsilon \in (0, 1)$, there exists $M_{\varepsilon,\delta} \in \mathbb{N}$ such that for $m \geq M_{\varepsilon,\delta}$ we have

$$\zeta_{\mathcal{S},\mathcal{T}}(m,\varepsilon) \geq \inf_{\substack{P\in\mathcal{S}\\ Q\in\mathcal{T}}} D(P\|Q) - \delta\,. \tag{31}$$

This is a special case of [11, Theorem 2] and we refer to [34] as well as references therein for a general discussion of composite hypothesis testing. Now, the strategy is to first measure the quantum states and then to invoke the classical achievability result from Eq. (31) for the resulting probability distributions.

This argument is made precise as follows. The classical case implies the existence of a sequence of tests $(T_{k,m})_{m\in\mathbb{N}}$ for the discrimination problem between the measured state $\mathcal{M}_k(\mathcal{S}^{\otimes k})^{\otimes m}$ and the measured state $\mathcal{M}_k(\mathcal{T}^{\otimes k})^{\otimes m}$ with $m \in \mathbb{N}$, such that

$$\sup_{\rho\in\mathcal{S}} \mathrm{Tr}\left[(1 - T_{k,m})\mathcal{M}_k(\rho^{\otimes k})^{\otimes m}\right] \leq \varepsilon \tag{32}$$

for all $m \in \mathbb{N}$, and

$$\lim_{m\to\infty} -\frac{1}{m} \log \sup_{\sigma\in\mathcal{T}} \operatorname{Tr}\left[T_{k,m}\mathcal{M}_k(\sigma^{\otimes k})^{\otimes m}\right] \geq \inf_{\substack{\rho\in\mathcal{S}\\ \sigma\in\mathcal{T}}} D\left(\mathcal{M}_k(\rho^{\otimes k})\big\|\mathcal{M}_k(\sigma^{\otimes k}\right) . \tag{33}$$

Hence, for any $\delta > 0$, there exists an m_δ such that for all $m \geq m_\delta$ we have

$$-\frac{1}{m} \log \sup_{\sigma\in\mathcal{T}} \operatorname{Tr}\left[T_{k,m}\mathcal{M}_k(\sigma^{\otimes k})^{\otimes m}\right] \geq \inf_{\substack{\rho\in\mathcal{S}\\ \sigma\in\mathcal{T}}} D\left(\mathcal{M}_k(\rho^{\otimes k})\big\|\mathcal{M}_k(\sigma^{\otimes k}\right) - \delta\,. \tag{34}$$

Defining $T_n := \left(\mathcal{M}_k^\dagger\right)^{\otimes m}(T_{k,m}) \otimes 1_r$ for $n = km + r$, $r \in \{0, \ldots, k-1\}$, we get that

$$\sup_{\rho\in\mathcal{S}} \operatorname{Tr}\left[(1 - T_n)\rho^{\otimes n}\right] = \sup_{\rho\in\mathcal{S}} \operatorname{Tr}\left[(1 - T_{k,m})\mathcal{M}_k(\rho^{\otimes k})^{\otimes m}\right] \leq \varepsilon \tag{35}$$

for all $n \in \mathbb{N}$, and thus

$$\begin{aligned}
\zeta_{\mathcal{S},\mathcal{T}}(n,\varepsilon) &\geq -\frac{1}{n} \log \sup_{\sigma\in\mathcal{T}} \operatorname{Tr}\left[T_n\sigma^{\otimes n}\right] && (36)\\
&= -\frac{1}{km + r} \log \sup_{\sigma\in\mathcal{T}} \operatorname{Tr}\left[T_{k,m}\mathcal{M}_k(\sigma^{\otimes k})^{\otimes m}\right] && (37)\\
&\geq \frac{m}{km + r} \inf_{\substack{\rho\in\mathcal{S}\\ \sigma\in\mathcal{T}}} D\left(\mathcal{M}_k(\rho^{\otimes k})\big\|\mathcal{M}_k(\sigma^{\otimes k})\right) - \frac{m}{km + r}\delta && (38)
\end{aligned}$$

whenever $n \geq km_\delta$. Therefore, we get

$$\liminf_{n\to\infty} \zeta_{\mathcal{S},\mathcal{T}}(n,\varepsilon) \geq \frac{1}{k} \inf_{\substack{\rho\in\mathcal{S}\\ \sigma\in\mathcal{T}}} D\left(\mathcal{M}_k(\rho^{\otimes k})\big\|\mathcal{M}_k(\sigma^{\otimes k})\right) - \frac{1}{k}\delta \tag{39}$$

for any binary POVM $\mathcal{M}_k$ and $\delta > 0$. Taking $\delta \to 0$ and then the supremum over $\mathcal{M}_k$ gives

$$\begin{aligned}
\liminf_{n\to\infty} \zeta_{\mathcal{S},\mathcal{T}}(n,\varepsilon) &\geq \frac{1}{k} \sup_{\mathcal{M}_k} \inf_{\substack{\rho\in\mathcal{S}\\ \sigma\in\mathcal{T}}} D(\mathcal{M}_k(\rho^{\otimes k})\|\mathcal{M}_k(\sigma^{\otimes k})) && (40)\\
&\geq \frac{1}{k} \sup_{\mathcal{M}_k} \inf_{\substack{\nu\in\mathcal{S}\\ \mu\in\mathcal{T}}} D(\mathcal{M}_k(\rho_k(\nu))\|\mathcal{M}_k(\sigma_k(\mu))) && (41)\\
&= \frac{1}{k} \inf_{\substack{\nu\in\mathcal{S}\\ \mu\in\mathcal{T}}} \sup_{\mathcal{M}_k} D(\mathcal{M}_k(\rho_k(\nu))\|\mathcal{M}_k(\sigma_k(\mu)))\,, && (42)
\end{aligned}$$

where the equality follows from Lemma A.2. Since this holds for every $k \in \mathbb{N}$, we find the claimed

$$\liminf_{n\to\infty} \zeta_{\mathcal{S},\mathcal{T}}(n,\varepsilon) \geq \sup_{k\in\mathbb{N}} \frac{1}{k} \inf_{\substack{\nu\in\mathcal{S}\\ \mu\in\mathcal{T}}} D_{\mathcal{M}}(\rho_k(\nu)\|\sigma_k(\mu))\,. \tag{43}$$

□

Next, we argue that the measured relative entropy can in fact be replaced by the quantum relative entropy by only paying an asymptotically vanishing penalty term. For this we need the following lemma, which can be seen as a generalization of the technical argument in the original proof of quantum Stein's lemma [30].

Lemma 2.4. *Let $\rho_n, \sigma_n \in S\left(\mathcal{H}^{\otimes n}\right)$ with σ_n permutation invariant. Then, we have*

$$D\big(\rho_n \big\| \sigma_n\big) - \log \operatorname{poly}(n) \leq D_{\mathcal{M}}\big(\rho_n \big\| \sigma_n\big) \leq D\big(\rho_n \big\| \sigma_n\big)\,, \tag{44}$$

where poly(n) *stands for terms of order at most polynomial in n.*

Proof. We can restrict ourselves to the case where $\operatorname{supp}\left(\rho_n\right) \subseteq \operatorname{supp}\left(\sigma_n\right)$ since otherwise all relative entropy terms evaluate to infinity by definition. The second inequality follows directly from the definition of the measured relative entropy in Eq. (29) together with the fact that the quantum relative entropy is monotone [36]. We now prove the first inequality with the help of asymptotic spectral pinching [25]. The pinching map with respect to $\omega \in S(\mathcal{H})$ is defined as

$$\mathcal{P}_\omega(\cdot) := \sum_{\lambda \in \operatorname{spec}(\omega)} P_\lambda(\cdot) P_\lambda \text{ with the spectral decomposition } \omega = \sum_{\lambda \in \operatorname{spec}(\omega)} \lambda P_\lambda. \tag{45}$$

Crucially, we have the pinching operator inequality [25]

$$\mathcal{P}_\omega[X] \gg \frac{X}{|\operatorname{spec}(\omega)|}\,, \tag{46}$$

where $|\operatorname{spec}(\cdot)|$ denotes the size of the spectrum. From this we can deduce that (see, e.g., [52, Lemma 4.4])

$$D\big(\rho_n \big\| \sigma_n\big) - \log \big|\operatorname{spec}\big(\sigma_n\big)\big| \leq D\big(\mathcal{P}_{\sigma_n}\big(\rho_n\big) \big\| \sigma_n\big) \leq D_{\mathcal{M}}\big(\rho_n \big\| \sigma_n\big)\,, \tag{47}$$

where the second inequality follows since $\mathcal{P}_{\sigma_n}\big(\rho_n\big)$ and σ_n are diagonal in the same basis and the measured relative entropy gives an upper-bound. It remains to show that $\big|\operatorname{spec}\big(\sigma_n\big)\big| \leq \operatorname{poly}(n)$. However, since σ_n is permutation invariant we have by Schur-Weyl duality (see, e.g., [24, Section 5]) that in the Schur basis

$$\sigma_n = \bigoplus_{\lambda \in \Lambda_n} \sigma_{Q_\lambda} \otimes 1_{P_\lambda} \quad \text{with } |\Lambda_n| \leq \operatorname{poly}(n) \text{ and } \dim\left[\sigma^0_{Q_\lambda}\right] \leq \operatorname{poly}(n). \tag{48}$$

where $\sigma^0_{Q_\lambda}$ is the projector onto the support of σ_{Q_λ}. This implies the claim. □

By combining Lemma 2.3 together with Lemma 2.4 we find for $\varepsilon \in (0, 1)$ that

$$\liminf_{n\to\infty} \zeta_{\mathcal{S},\mathcal{T}}(n, \varepsilon) \geq \limsup_{n\to\infty} \frac{1}{n} \inf_{\substack{\nu\in\mathcal{S}\\ \mu\in\mathcal{T}}} D\left(\rho_n(\nu) \| \sigma_n(\mu)\right)\,. \tag{49}$$

The next step is to argue that asymptotically the infimum over states $\rho_n(\nu)$ can without loss of generality be restricted to iid states $\rho^{\otimes n}$ with $\rho \in \mathcal{S}$.

Lemma 2.5. *For definitions as above and $\omega_n \in S\left(\mathcal{H}^{\otimes n}\right)$, we have*

$$\frac{1}{n} \inf_{\nu\in\mathcal{S}} D\left(\rho_n(\nu) \big\| \omega_n\right) \geq \frac{1}{n} \inf_{\rho\in\mathcal{S}} D\left(\rho^{\otimes n} \big\| \omega_n\right) - \frac{2d^2 \log(n+1)}{n}\,, \tag{50}$$

where $d := \dim\left(\mathcal{H}\right)$.

Proof. For $\nu \in \mathcal{S}$ and $H(\rho) := -\operatorname{Tr}\left[\rho \log \rho\right]$ the von Neumann entropy, we observe the following chain of arguments

$$
\begin{aligned}
&\frac{1}{n} D\left(\rho_n(\nu) \middle\| \omega_n\right) \\
&= \frac{1}{n} D\Big(\sum_{i=1}^{N} p_i \rho_i^{\otimes n} \Big\| \omega_n\Big) && (51)\\
&= -\frac{1}{n} H\Big(\sum_{i=1}^{N} p_i \rho_i^{\otimes n}\Big) - \frac{1}{n}\sum_{i=1}^{N} p_i \operatorname{Tr}\left[\rho_i^{\otimes n} \log \omega_n\right] && (52)\\
&\geq -\frac{1}{n}\sum_{i=1}^{N} p_i H\left(\rho_i^{\otimes n}\right) - \frac{\log (n+1)^{2d^2}}{n} - \frac{1}{n}\sum_{i=1}^{N} p_i \operatorname{Tr}\left[\rho_i^{\otimes n} \log \omega_n\right] && (53)\\
&\geq \min_{\rho_i} \frac{1}{n} D\left(\rho_i^{\otimes n} \middle\| \omega_n\right) - \frac{2d^2 \log(n+1)}{n} && (54)\\
&\geq \inf_{\rho \in \mathcal{S}} \frac{1}{n} D\left(\rho^{\otimes n} \middle\| \omega_n\right) - \frac{2d^2 \log(n+1)}{n}\,, && (55)
\end{aligned}
$$

where the first equality holds by an application of Carathédory's theorem with $N \leq (n+1)^{2d^2}$ (Lemma A.3), and the first inequality by an almost-convexity property of the von Neumann entropy (Lemma A.4). All other steps are elementary. Since the above argument holds for all $\nu \in \mathcal{S}$, the claim follows. □

Lemma 2.5 together with Eq. (49) gives for $\varepsilon \to 0$ that

$$
\begin{aligned}
\limsup_{n\to\infty} \frac{1}{n} \inf_{\substack{\rho\in\mathcal{S}\\ \mu\in\mathcal{T}}} D\left(\rho^{\otimes n} \| \sigma_n(\mu)\right) &\leq \sup_{k\in\mathbb{N}} \frac{1}{k} \inf_{\substack{\nu\in\mathcal{S}\\ \mu\in\mathcal{T}}} D_{\mathcal{M}}(\rho_k(\nu)\|\sigma_k(\mu)) && (56)\\
&\leq \lim_{\varepsilon\to 0} \liminf_{n\to\infty} \zeta_{\mathcal{S},\mathcal{T}}(n,\varepsilon) && (57)\\
&\leq \liminf_{n\to\infty} \frac{1}{n} \inf_{\substack{\rho\in\mathcal{S}\\ \mu\in\mathcal{T}}} D\left(\rho^{\otimes n} \| \sigma_n(\mu)\right)\,, && (58)
\end{aligned}
$$

where the last step follows from Eq. (28). This shows that the limit

$$
\lim_{n\to\infty} \frac{1}{n} \inf_{\substack{\rho\in\mathcal{S}\\ \mu\in\mathcal{T}}} D\left(\rho^{\otimes n} \| \sigma_n(\mu)\right) \qquad (59)
$$

exists and all the inequalities above hold as equalities. Furthermore, we have

$$
\begin{aligned}
\limsup_{n\to\infty} \frac{1}{n} \inf_{\substack{\rho\in\mathcal{S}\\ \mu\in\mathcal{T}}} D\left(\rho^{\otimes n} \| \sigma_n(\mu)\right) &\leq \lim_{\varepsilon\to 0} \limsup_{n\to\infty} \zeta_{\mathcal{S},\mathcal{T}}(n,\varepsilon) && (60)\\
&\leq \limsup_{n\to\infty} \frac{1}{n} \inf_{\substack{\rho\in\mathcal{S}\\ \mu\in\mathcal{T}}} D\left(\rho^{\otimes n} \| \sigma_n(\mu)\right)\,, && (61)
\end{aligned}
$$

which concludes the proof of Theorem 1.1. □

3. Examples and Extensions

Here, we discuss several concrete examples of composite discrimination problems—some of which have a single-letter solution.

3.1. Relative entropy of coherence. Following the literature around [5], the set of states diagonal in a fixed basis $\{|c\rangle\}$ is called incoherent and denoted by $\mathcal{C} \subseteq S(\mathcal{H})$. The relative entropy of coherence of $\rho \in S(\mathcal{H})$ is defined as

$$D_{\mathcal{C}}(\rho) := \inf_{\sigma \in \mathcal{C}} D(\rho\|\sigma)\,. \tag{62}$$

Based on our main result (Theorem 1.1), we can characterize the following discrimination problem.

Null hypothesis: the fixed state $\rho^{\otimes n}$
Alternative hypothesis: the convex hull of iid coherent states

$$\bar{\mathcal{C}}_n := \left\{ \int \sigma^{\otimes n}\, \mathrm{d}\mu(\sigma) \middle| \mu \in \mathcal{C} \right\} \tag{63}$$

Namely, Theorem 1.1 gives

$$\zeta_{\bar{\mathcal{C}}}(\infty, 0) := \lim_{\varepsilon\to 0} \lim_{n\to\infty} \zeta_{\bar{\mathcal{C}}}(n, \varepsilon) \tag{64}$$

$$= \lim_{n\to\infty} \frac{1}{n} \inf_{\mu\in\mathcal{C}} D\Big(\rho^{\otimes n} \Big\| \int \sigma^{\otimes n}\, \mathrm{d}\mu(\sigma)\Big) \tag{65}$$

$$= D_{\mathcal{C}}(\rho)\,, \tag{66}$$

where the limit in Eq. (64) exists because the relative entropy of coherence is additive on product states [14], and the last step follows from a general property of the relative entropy of coherence (Lemma A.5) applied to the decohering channel. In fact, there is even a single-letter solution for the following less restricted discrimination problem.

Null hypothesis: the fixed state $\rho^{\otimes n}$
Alternative hypothesis: the convex set of coherent states $\mathcal{C}_n$

It is straightforward to check that this hypothesis testing problem fits the general framework of [13], leading to

$$\zeta_{\mathcal{C}}(\infty, \varepsilon) := \lim_{n\to\infty} \zeta_{\mathcal{C}}(n, \varepsilon) = \lim_{n\to\infty} \frac{1}{n} \inf_{\sigma_n\in\mathcal{C}_n} D\left(\rho^{\otimes n}\|\sigma_n\right) = D_{\mathcal{C}}(\rho) \quad \forall\varepsilon \in (0, 1)\,, \tag{67}$$

where the last step again follows from a general property of the relative entropy of coherence (Lemma A.5). Thus, we have two a priori different hypothesis testing scenarios that both give an operational interpretation to the relative entropy of coherence. In the following we give a simple self-contained proof of Eq. (67) that is different from the rather involved steps in [13] and instead follows ideas from [4,28]. The goal is the quantification of the optimal asymptotic error exponent

$$\zeta_{\mathcal{C}}(n, \varepsilon) := -\frac{1}{n} \log \inf_{\substack{0\ll M_n\ll 1\\ \mathrm{Tr}[M_n\rho^{\otimes n}]\geq 1-\varepsilon}} \sup_{\sigma_n\in\mathcal{C}_n} \mathrm{Tr}\left[M_n\sigma_n\right] \tag{68}$$

$$\text{with}\quad \zeta_{\mathcal{C}}(\infty, \varepsilon) := \lim_{n\to\infty} \zeta_{\mathcal{C}}(n, \varepsilon)\,. \tag{69}$$

Proposition 3.1 *For the discrimination problem as above with* $\varepsilon \in (0,1)$*, we have*

$$\zeta_{\mathcal{C}}(\infty, \varepsilon) = D_{\mathcal{C}}(\rho)\,. \tag{70}$$

Note that Proposition 3.1 is independent of $\operatorname{supp}(\rho)$ as the set $\mathcal{C}$ includes full rank states. A weak converse for $\varepsilon \to 0$ follows exactly as in Lemma 2.1, together with Lemma A.5 to make the expression single-letter. For the strong converse as claimed in Proposition 3.1, we make use of a general family of quantum Rényi entropies: the Petz divergences [44]. For $\rho, \sigma \in S(\mathcal{H})$ and $s \in (0,1) \cup (1,\infty)$ they are defined as

$$D_s\left(\rho\middle\|\sigma\right) := \frac{1}{s-1}\log\operatorname{Tr}\left[\rho^s\sigma^{1-s}\right], \tag{71}$$

whenever either $s < 1$ and ρ is not orthogonal to σ in Hilbert-Schmidt inner product or $s > 1$ and the support of ρ is contained in the support of σ. (Otherwise we set $D_s(\rho\|\sigma) := \infty$.) The corresponding Rényi relative entropies of coherence are given by [14]

$$D_{s,\mathcal{C}}(\rho) := \inf_{\sigma\in\mathcal{C}} D_s(\rho\|\sigma) \text{ with the additivity property } D_{s,\mathcal{C}}\left(\rho^{\otimes n}\right) = nD_{s,\mathcal{C}}(\rho). \tag{72}$$

Using similar standard arguments [40] as in Lemma 2.1 but based on the monotonicity of the Petz divergences, we find for $s \in (1,2]$ that

$$\begin{aligned} &-\frac{1}{n}\log\inf_{0\le M_n\le 1}\left\{\operatorname{Tr}\left[M_n\sigma_n\right]\middle|\operatorname{Tr}\left[(1-M_n)\rho^{\otimes n}\right]\le\varepsilon\right\} \\ &\le \frac{1}{n}\cdot D_s\left(\rho^{\otimes n}\middle\|\sigma_n\right) + \frac{1}{n}\frac{s}{s-1}\frac{1}{\log(1-\varepsilon)}\,. \end{aligned} \tag{73}$$

By taking the infimum over $\sigma_n \in \mathcal{C}_n$, a basic application of Sion's minimax theorem (Lemma A.1), using the additivity from Eq. (72), taking the limit $n \to \infty$ as well as the limit [14]

$$\lim_{s\to 1} D_{s,\mathcal{C}}(\rho) = D_{\mathcal{C}}(\rho)\,, \tag{74}$$

we find the claimed strong converse $\zeta_{\mathcal{C}}(\infty,\varepsilon) \le D_{\mathcal{C}}(\rho)$. The achievability direction of Proposition 3.1 is based on the Petz divergences as well.

Lemma 3.2 *For the discrimination problems as above with* $n \in \mathbb{N}$ *and* $\varepsilon \in (0,1)$*, we have for* $s \in (0,1)$ *that*

$$\zeta_{\mathcal{C}}(n,\varepsilon) \ge D_{s,\mathcal{C}}(\rho) - \frac{1}{n}\frac{s}{1-s}\log\frac{1}{\varepsilon}\,. \tag{75}$$

Taking the limit $n \to \infty$ as well as the limit $s \to 1$ using Eq. (74), we then find the claimed achievability $\zeta_{\mathcal{C}}(\infty,\varepsilon) \ge D_{\mathcal{C}}(\rho)$.

Proof of Lemma 3.2. It is straightforward to check with Sion's minimax theorem (Lemma A.1) that

$$\inf_{\substack{0\le M_n\le 1\\ \operatorname{Tr}[M_n\rho^{\otimes n}]\ge 1-\varepsilon}}\ \sup_{\sigma_n\in\mathcal{C}_n}\operatorname{Tr}\left[M_n\sigma_n\right] = \sup_{\sigma_n\in\mathcal{C}_n}\ \inf_{\substack{0\le M_n\le 1\\ \operatorname{Tr}[M_n\rho^{\otimes n}]\ge 1-\varepsilon}}\operatorname{Tr}\left[M_n\sigma_n\right]. \tag{76}$$

Now, for $\lambda_n \in \mathbb{R}$ with $n \in \mathbb{N}$ we choose $M_n(\lambda_n) := \left\{\rho^{\otimes n} - 2^{\lambda_n}\sigma_n\right\}_+$ where $\{\cdot\}_+$ denotes the projector on the eigenspace of the positive spectrum. We have $0 \ll M_n(\lambda_n) \ll 1$ and by Audenaert's inequality (Lemma A.6) with $s \in (0,1)$ we get

$$\mathrm{Tr}\left[(1 - M_n(\lambda_n))\rho^{\otimes n}\right] \leq 2^{(1-s)\lambda_n}\,\mathrm{Tr}\left[\left(\rho^{\otimes n}\right)^s \sigma_n^{1-s}\right] = 2^{(1-s)\left(\lambda_n - D_s\left(\rho^{\otimes n}\middle\|\sigma_n\right)\right)}. \quad (77)$$

Moreover, again Audenaert's inequality (Lemma A.6) for $s \in (0,1)$ implies

$$\mathrm{Tr}\left[M_n(\lambda_n)\sigma_n\right] \leq 2^{-s\lambda_n}\,\mathrm{Tr}\left[\left(\rho^{\otimes n}\right)^s \sigma_n^{1-s}\right] = 2^{-s\lambda_n - (1-s)D_s\left(\rho^{\otimes n}\middle\|\sigma_n\right)}. \quad (78)$$

Hence, choosing

$$\lambda_n := D_s\left(\rho^{\otimes n}\middle\|\sigma_n\right) + \log \varepsilon^{\frac{1}{1-s}} \text{ with } M_n := M_n(\lambda_n) \quad (79)$$

leads with Eq. (77) to $\mathrm{Tr}\left[M_n\rho^{\otimes n}\right] \geq 1 - \varepsilon$. Finally, Eq. (76) together with Eq. (78) and the additivity property from Eq. (72) leads to the claim.

We note that a more refined analysis of the above calculation allows to determine the Hoeffding bound as well as the strong converse exponent (cf. [4,28]). The former gives an operational interpretation to the Rényi relative entropy of coherence $D_{s,\mathcal{C}}(\rho)$, whereas the latter gives an operational interpretation to the sandwiched Rényi relative entropies of coherence [14]

$$\tilde{D}_{s,\mathcal{C}}(\rho) := \inf_{\sigma\in\mathcal{C}} \tilde{D}_s(\rho\|\sigma) \quad (80)$$

with the sandwiched Rényi entropies

$$\tilde{D}_s(\rho\|\sigma) := \frac{1}{s-1}\log \mathrm{Tr}\left[\left(\sigma^{\frac{1-s}{2s}}\rho\sigma^{\frac{1-s}{2s}}\right)^s\right] \quad (81)$$

whenever either $s < 1$ and ρ is not orthogonal to σ in Hilbert-Schmidt inner product or $s > 1$ and the support of ρ is contained in the support of σ [39,57]. (Otherwise we set $D_s(\rho\|\sigma) := \infty$.) The crucial insight for the proof is again the additivity property $\tilde{D}_{s,\mathcal{C}}\left(\rho^{\otimes n}\right) = n\tilde{D}_{s,\mathcal{C}}(\rho)$, that was already shown in [14].

3.2. Relative entropy of recovery. The relative entropy of recovery of $\rho_{ABC} \in S(\mathcal{H}_{ABC})$ and its regularized version are defined as [7,12,47][5]

$$D(A;B|C)_\rho := \inf_{\mathcal{R}} D\left(\rho_{ABC}\middle\|(\mathcal{I}_A \otimes \mathcal{R}_{C\to BC})(\rho_{AC})\right) \quad (82)$$

$$\text{and} \quad D^\infty(A;B|C)_\rho := \lim_{n\to\infty}\frac{1}{n}D(A;B|C)_{\rho^{\otimes n}}, \quad (83)$$

where the infimum goes over all completely positive and trace preserving maps $\mathcal{R}_{C\to BC}$. It was recently shown that in general [21]

$$D^\infty(A;B|C)_\rho \neq D(A;B|C)_\rho. \quad (84)$$

Using the framework from [13], the following discrimination problem was linked to the regularized relative entropy of recovery [16].

[5] This limit exists and is finite as for $a_n := D(A;B|C)_{\rho^{\otimes n}} \geq 0$ we have the monotonicity property $a_{n+m} \leq a_n + a_m$.

Null hypothesis: the fixed state $\rho_{ABC}^{\otimes n}$
Alternative hypothesis: for any $\mathcal{R}_{C^n\to B^nC^n}$ completely positive and trace preserving, the convex set of states

$$\mathcal{R}^n := \left\{(\mathcal{I}_{A^n}\otimes\mathcal{R}_{C^n\to B^nC^n})\left(\rho_{AC}^{\otimes n}\right)\right\} \tag{85}$$

Namely, for $\varepsilon\in(0,1)$ we have for the corresponding asymptotic error exponent

$$\zeta_{\mathcal{R}}(\infty,\varepsilon) := \lim_{n\to\infty}\zeta_{\mathcal{R}}(n,\varepsilon) = D^{\infty}(A;B|C)_\rho\,. \tag{86}$$

In contrast, our main result (Theorem 1.1) covers the following discrimination problem.

Null hypothesis: the fixed state $\rho_{ABC}^{\otimes n}$
Alternative hypothesis: for any $\mathcal{R}_{C\to BC}$ completely positive and trace preserving, the convex hull of iid states

$$\bar{\mathcal{R}}^n := \left\{\int((\mathcal{I}_A\otimes\mathcal{R}_{C\to BC})(\rho_{AC}))^{\otimes n}\ \mathrm{d}\mu(\mathcal{R})\right\}. \tag{87}$$

Interestingly, we can show that the asymptotic error exponents of the two discrimination problems are actually identical.

Proposition 3.3. *With the definitions as above, we have*

$$\begin{aligned}&\lim_{n\to\infty}\frac{1}{n}\inf_{\mathcal{R}} D\Big(\rho_{ABC}^{\otimes n}\Big\|(\mathcal{I}_A\otimes\mathcal{R}_{C^n\to B^nC^n})\left(\rho_{AC}^{\otimes n}\right)\Big)\\ &\quad=\lim_{n\to\infty}\frac{1}{n}\inf_{\mu\in\mathcal{R}} D\Big(\rho_{ABC}^{\otimes n}\Big\|\int\big((\mathcal{I}_A\otimes\mathcal{R}_{C\to BC})(\rho_{AC})\big)^{\otimes n}\ \mathrm{d}\mu(\mathcal{R})\Big).\end{aligned} \tag{88}$$

Proof. One direction of the inequality is by definition and for the other direction we use a de Finetti reduction for quantum channels [12, Lemma 8] that was first derived in [22]. Namely, we have for $\omega_{C^n}\in S\left(\mathcal{H}_C^{\otimes n}\right)$ and permutation invariant $\mathcal{R}_{C^n\to B^nC^n}$ that

$$\mathcal{R}_{C^n\to B^nC^n}\left(\omega_{C^n}\right) \ll \mathrm{poly}(n)\int\left(\mathcal{R}_{C\to BC}\right)^{\otimes n}\left(\omega_{C^n}\right)\mathrm{d}\nu(\mathcal{R}) \tag{89}$$

for some measure $\mathrm{d}\nu(\mathcal{R})$ over the completely positive and trace preserving maps on $C\to BC$. As explained in the proof of [12, Proposition 9], the joint convexity of the quantum relative entropy together with the operator monotonicity of the logarithm then imply that

$$\begin{aligned}&D\left(\rho_{ABC}^{\otimes n}\middle\|\mathcal{R}_{C^n\to B^nC^n}\left(\rho_{AC}^{\otimes n}\right)\right)\\ &\quad\geq D\Big(\rho_{ABC}^{\otimes n}\Big\|\int\big((\mathcal{I}_A\otimes\mathcal{R}_{C\to BC})(\rho_{AC})\big)^{\otimes n}\ \mathrm{d}\nu(\mathcal{R})\Big)-\log\mathrm{poly}(n)\,.\end{aligned} \tag{90}$$

□

As such, we can conclude that

$$\begin{aligned}\zeta_{\bar{\mathcal{R}}}(\infty,0) &:= \lim_{\varepsilon\to0}\liminf_{n\to\infty}\zeta_{\bar{\mathcal{R}}}(n,\varepsilon)\\ &= \lim_{\varepsilon\to0}\limsup_{n\to\infty}\zeta_{\bar{\mathcal{R}}}(n,\varepsilon) = D^{\infty}(A;B|C)_\rho\,.\end{aligned} \tag{91}$$

3.3. Quantum mutual information. The quantum mutual information of $\rho_{AB} \in S(\mathcal{H}_{AB})$ is defined as

$$I(A:B)_\rho := H(A)_\rho + H(B)_\rho - H(AB)_\rho\,. \tag{92}$$

Our main result from Sect. 2 provides a solution to the following discrimination problem.

Null hypothesis: the fixed state $\rho_{AB}^{\otimes n}$
Alternative hypothesis: the convex hull of iid states

$$\bar{\mathcal{T}}_{A^n:B^n} := \left\{\rho_A^{\otimes n} \otimes \int \sigma_B^{\otimes n}\,\mathrm{d}\mu(\sigma)\middle|\mu \in S(\mathcal{H}_B)\right\}. \tag{93}$$

Namely, we have

$$\bar{\zeta}_{A:B}(\infty,0) := \lim_{\varepsilon\to 0}\lim_{n\to\infty}\bar{\zeta}_{A:B}(n,\varepsilon) \tag{94}$$

$$= \lim_{n\to\infty}\frac{1}{n}\inf_{\mu\in\bar{\mathcal{T}}} D\Big(\rho_{AB}^{\otimes n}\Big\|\rho_A^{\otimes n}\otimes\int\sigma_B^{\otimes n}\,\mathrm{d}\mu(\sigma)\Big) \tag{95}$$

$$=I(A:B)_\rho\,. \tag{96}$$

Here, the last equality follows from the easily checked identity

$$I(A:B)_\rho = \inf_{\sigma_B\in S(\mathcal{H})} D(\rho_{AB}\|\rho_A\otimes\sigma_B)\,. \tag{97}$$

More general composite discrimination problems leading to the quantum mutual information were solved in [28] and in the following we further extend these results (cf. the classical work [53]).

Null hypothesis: the fixed state $\rho_{AB}^{\otimes n}$
Alternative hypothesis: the set of states

$$\mathcal{T}_{A^n:B^n} := \left\{\sigma_{A^n}\otimes\sigma_{B^n} \in S\left(\mathcal{H}_{AB}^{\otimes n}\right)\middle|\sigma_{A^n} \text{ or } \sigma_{B^n} \text{ permutation invariant}\right\}. \tag{98}$$

The goal is again the quantification of the optimal asymptotic error exponent

$$\zeta_{A:B}(n,\varepsilon) := -\frac{1}{n}\log \inf_{\substack{0\ll M_n\ll 1\\ \mathrm{Tr}[M_n\rho^{\otimes n}]\geq 1-\varepsilon}} \sup_{\sigma_{A^n}\otimes\sigma_{B^n}\in\mathcal{T}_n} \mathrm{Tr}\left[M_{A^nB^n}\sigma_{A^n}\otimes\sigma_{B^n}\right] \tag{99}$$

$$\text{with}\quad \zeta_{A:B}(\infty,\varepsilon) := \lim_{n\to\infty}\zeta_{A:B}(n,\varepsilon)\,. \tag{100}$$

Note that the sets $\mathcal{T}_{A^nB^n}$ are not convex and hence the minimax technique used in Sect. 3.1 does not work here. However, following the ideas in [28,53] we can exploit the permutation invariance and use de Finetti reductions of the form [15,26] to find the following.

Proposition 3.4. *For the discrimination problem as above with* $\varepsilon \in (0,1)$*, we have*

$$\zeta_{A:B}(\infty,\varepsilon) = I(A:B)_\rho\,. \tag{101}$$

The achievability direction is based on the following lemma.

Lemma 3.5. *For the discrimination problem as above with* $n \in \mathbb{N}$ *and* $\varepsilon \in (0, 1)$*, we have for* $s \in (0, 1)$ *that*

$$\zeta_{A:B}(n, \varepsilon) \geq \inf_{\sigma \in S(\mathcal{H})} D_s \left(\rho_{AB} \middle\| \sigma_A \otimes \sigma_B\right) - \frac{1}{n} \frac{s}{1-s} \log \frac{1}{\varepsilon} - \frac{\log \operatorname{poly}(n)}{n} . \tag{102}$$

Proof. Without loss of generality assume that σ_{A^n} is permutation invariant. We choose

$$\begin{aligned} M_{A^n B^n}(\lambda_n) &:= \left\{\rho_{AB}^{\otimes n} - 2^{\lambda_n} \omega_{A^n} \otimes \omega_{B^n}\right\}_+ \\ \text{with} \quad \omega_{A^n} &:= \binom{n + |A|^2 - 1}{n}^{-1} \operatorname{Tr}_{\tilde{A}^n} \left[P_{A^n \tilde{A}^n}^{\mathrm{Sym}}\right] , \end{aligned} \tag{103}$$

where $P_{A^n \tilde{A}^n}^{\mathrm{Sym}}$ denotes the projector onto the symmetric subspace of $\mathcal{H}_A^{\otimes n} \otimes \mathcal{H}_{\tilde{A}}^{\otimes n}$ with $|A| = |\tilde{A}|$ (denoting the dimension of $\mathcal{H}_A$ by $|A|$), and similarly for B^n. Audenaert's inequality (Lemma A.6) gives that

$$\begin{aligned} &\operatorname{Tr}\left[(1 - M_{A^n B^n}(\lambda_n))\rho_{AB}^{\otimes n}\right] \leq 2^{(1-s)\lambda_n} \operatorname{Tr}\left[\left(\rho_{AB}^{\otimes n}\right)^s \left(\omega_{A^n} \otimes \omega_{B^n}\right)^{1-s}\right] \\ &\quad \leq 2^{(1-s)\left(\lambda_n - \inf_{\sigma_{A^n} \otimes \sigma_{B^n} \in \mathcal{T}_n} D_s\left(\rho_{AB}^{\otimes n} \middle\| \sigma_{A^n} \otimes \sigma_{B^n}\right)\right)} . \end{aligned} \tag{104}$$

Furthermore, we have by Schur-Weyl duality that $\sigma_{A^n} \leq \binom{n+|A|^2-1}{n} \omega_{A^n}$ for all permutation invariant σ_{A^n} (see, e.g., [28, Lemma 1]) and thus again by Audenaert's inequality (Lemma A.6)

$$\begin{aligned} &\operatorname{Tr}\left[M_{A^n B^n}(\lambda_n) \left(\sigma_{A^n} \otimes \sigma_{B^n}\right)\right] && (105) \\ &= \operatorname{Tr}\left[M_{A^n B^n}(\lambda_n) \left(\sigma_{A^n} \otimes \left(\sum_{\pi \in S_n} U_{B^n}(\pi) \sigma_{B^n} U_{B^n}^\dagger(\pi)\right)\right)\right] \quad (S_n : \text{ symm. group}) \\ &\leq \underbrace{\binom{n + |A|^2 - 1}{n}\binom{n + |B|^2 - 1}{n}}_{=:\, p(n) \,\leq\, \operatorname{poly}(n)} \operatorname{Tr}\left[M_{A^n B^n}(\lambda_n) \left(\omega_{A^n} \otimes \omega_{B^n}\right)\right] \\ &\leq p(n) \cdot 2^{-s\lambda_n} \operatorname{Tr}\left[\left(\rho_{AB}^{\otimes n}\right)^s \left(\omega_{A^n} \otimes \omega_{B^n}\right)^{1-s}\right] \\ &\leq p(n) \cdot 2^{-s\lambda_n - (1-s) \inf_{\sigma_{A^n} \otimes \sigma_{B^n} \in \mathcal{T}_n} D_s\left(\rho_{AB}^{\otimes n} \middle\| \sigma_{A^n} \otimes \sigma_{B^n}\right)} . && (106) \end{aligned}$$

We now choose

$$\lambda_n := \inf_{\sigma_{A^n} \otimes \sigma_{B^n} \in \mathcal{T}_n} D_s \left(\rho_{AB}^{\otimes n} \middle\| \sigma_{A^n} \otimes \sigma_{B^n}\right) + \log \varepsilon^{\frac{1}{1-s}} \text{ with } M_{A^n B^n} := M_{A^n B^n}(\lambda_n), \tag{107}$$

from which we get $\operatorname{Tr}\left[M_{A^n B^n} \rho_{AB}^{\otimes n}\right] \geq 1 - \varepsilon$ and together with Eqs. (99) and (106) that

$$\zeta_{A:B}^n(\varepsilon) \geq \inf_{\sigma_{A^n} \otimes \sigma_{B^n} \in \mathcal{T}_n} D_s \left(\rho_{AB}^{\otimes n} \middle\| \sigma_{A^n} \otimes \sigma_{B^n}\right) - \frac{1}{n} \frac{s}{1-s} \log \frac{1}{\varepsilon} - \frac{\log p(n)}{n} . \tag{108}$$

To deduce the claim it is now sufficient to argue that the Rényi quantum mutual information[6]

$$I_s(A:B)_\rho := \inf_{\sigma_A\otimes\sigma_B\in S(\mathcal{H})} D_s\left(\rho_{AB}\middle\|\sigma_A\otimes\sigma_B\right) \tag{109}$$

is additive on tensor product states. This, however, follows exactly as in the classical case [53, App. A-C] from the (quantum) Sibson identity [48, Lemma 3]

$$D_s\left(\rho_{AB}\middle\|\sigma_A\otimes\sigma_B\right) = D_s\left(\rho_{AB}\middle\|\sigma_A\otimes\bar{\sigma}_B\right) + D_s\left(\bar{\sigma}_B\middle\|\sigma_B\right)$$
$$\text{with}\quad \bar{\sigma}_B := \frac{\left(\mathrm{Tr}_A\left[\rho_{AB}^s\sigma_A^{1-s}\right]\right)^{\frac{1}{s}}}{\mathrm{Tr}\left[\left(\mathrm{Tr}_A\left[\rho_{AB}^s\sigma_A^{1-s}\right]\right)^{\frac{1}{s}}\right]}\,. \tag{110}$$

□

Taking the limit $n\to\infty$ in Lemma 3.5 and then taking the limit $s\to 1$ via the quantum Sibson identity from Eqs. (110) and (97) yields

$$\lim_{s\to 1} I_s(A:B)_\rho = I(A:B)_\rho\,, \tag{111}$$

gives the claimed achievability $\zeta_{A:B}(\infty,0)\geq I(A:B)_\rho$. A weak converse for $\varepsilon\to 0$ follows as in Lemma 2.1 and the strong converse as claimed in Proposition 3.4 is derived similarly as in Proposition 3.1—by noting that it is sufficient to prove a converse for testing

$$\rho_{AB}^{\otimes n} \text{ against } \rho_A^{\otimes n}\otimes\sigma_{B^n}\,. \tag{112}$$

A more refined analysis of the above calculation along the work [28] allows to determine the Hoeffding bound for the product testing discrimination problem as above. However, for the strong converse exponent we are missing the additivity of the sandwiched Rényi quantum mutual information

$$\tilde{I}_s(A:B)_\rho := \inf_{\sigma_A\otimes\sigma_B\in S(\mathcal{H})} \tilde{D}_s\left(\rho_{AB}\middle\|\sigma_A\otimes\sigma_B\right) \tag{113}$$

on product states.

4. Conditional Quantum Mutual Information

Here, we discuss how our results are related to the conditional quantum mutual information. This allows us to show that the regularization in our formula for composite asymmetric hypothesis testing as stated in Theorem 1.1 is needed in general.

[6] This definition is slightly different from the Rényi quantum mutual information discussed in [28].

4.1. Recoverability bounds. The following is a proof of the lower bound on the conditional quantum mutual information from Eq. (20).

Theorem 4.1. *For $\rho_{ABC} \in S(\mathcal{H}_{ABC})$ we have*

$$I(A:B|C)_\rho \geq \limsup_{n\to\infty} \frac{1}{n} D\Big(\rho_{ABC}^{\otimes n}\Big\| \int \beta_0(t) \left(\mathcal{I}_A \otimes \mathcal{R}^{[t]}_{C\to BC}(\rho_{AC})\right)^{\otimes n} \mathrm{d}t\Big), \quad (114)$$

where $\mathcal{R}^{[t]}_{C\to BC}(\cdot) := \rho_{BC}^{\frac{1+it}{2}} \big(\rho_C^{\frac{-1-it}{2}} (\cdot) \rho_C^{\frac{-1+it}{2}}\big) \rho_{BC}^{\frac{1-it}{2}}$ with the inverses understood as generalized inverses and $\beta_0(t) := \frac{\pi}{2}\left(\cosh(\pi t) + 1\right)^{-1}$.

Proof. We start from the lower bound [50, Theorem 4.1] applied to $\rho_{ABC}^{\otimes n}$ (with the support conditions taken care of as in the corresponding proof)

$$I(A:B|C)_\rho = \frac{1}{n} I\left(A^n : B^n \middle| C^n\right)_{\rho^{\otimes n}} \geq \frac{1}{n} D_{\mathcal{M}}\left(\rho_{ABC}^{\otimes n}\middle\|\sigma_{A^nB^nC^n}\right) \quad (115)$$

with

$$\sigma_{A^nB^nC^n} := \int \beta_0(t) \left(\sigma^{[t]}_{ABC}\right)^{\otimes n} \mathrm{d}t \text{ and } \sigma^{[t]}_{ABC} := \left(\mathcal{I}_A \otimes \mathcal{R}^{[t]}_{C\to BC}\right)(\rho_{AC}), \quad (116)$$

where we have used that the conditional quantum mutual information is additive on product states. Now, we simply observe that $\sigma_{A^nB^nC^n}$ is permutation invariant and hence the claim can be deduced from Lemma 2.4 together with taking the limit superior $n \to \infty$. □

Together with previous work we find the following corollary that encompasses all known recoverability lower bounds on the conditional quantum mutual information.

Corollary 4.2. *For $\rho_{ABC} \in S(\mathcal{H}_{ABC})$ the conditional quantum mutual information $I(A:B|C)_\rho$ is lower bounded by*

$$-\int \beta_0(t) \log \left\|\sqrt{\rho_{ABC}}\sqrt{\sigma^{[t]}_{ABC}}\right\|_1^2 \mathrm{d}t \quad (117)$$

$$D_{\mathcal{M}}\Big(\rho_{ABC}\Big\| \int \beta_0(t)\sigma^{[t]}_{ABC}\, \mathrm{d}t\Big) \quad (118)$$

$$\limsup_{n\to\infty} \frac{1}{n} D\Big(\rho_{ABC}^{\otimes n}\Big\| \int \beta_0(t)\big(\sigma^{[t]}_{ABC}\big)^{\otimes n} \mathrm{d}t\Big) \quad (119)$$

with $\sigma^{[t]}_{ABC}$ from Eq. (116).

The first bound was shown in [32, Section 3], the second one in [50, Theorem 4.1], and the third one is Theorem 4.1. We note that the lower bounds are typically strict in the non-commutative case, as can be seen from numerical work (see, e.g., [12]). In contrast to the second and third bound, the first lower bound is not tight in the commutative case but has the advantage that the average over $\beta_0(t)$ stands outside of the distance measure used. Moreover, the distribution $\beta_0(t)$ cannot be taken outside the relative entropy measure in the second and the third bound, since quantum Stein's lemma would then lead to a contradiction to a recent counterexample from [21, Section 5]. Namely, there exists $\theta \in [0, \pi/2]$ such that

$$I(A:B|C)_\rho \not\geq \inf_{\mathcal{R}} D\left(\rho_{ABC}\middle\|(\mathcal{I}_A \otimes \mathcal{R}_{C\to BC})(\rho_{AC})\right) \quad (120)$$

for the pure state $\rho_{ABC} = |\rho\rangle\langle\rho|_{ABC}$ with

$$|\rho\rangle_{ABC} = \frac{1}{\sqrt{2}}\big(\cos(\theta)|0\rangle_A \otimes |1\rangle_C + \sin(\theta)|1\rangle_A \otimes |0\rangle_C\big) \otimes |1\rangle_B + \frac{1}{\sqrt{2}}|0\rangle_A \otimes |0\rangle_B \otimes |0\rangle_C\,. \tag{121}$$

It seems that the only remaining conjectured strengthening is the lower bound in terms of the non-rotated Petz map [8, Section 8]

$$I(A:B|C)_\rho \geq -\log\left\|\sqrt{\rho_{ABC}}\sqrt{\sigma^{[0]}_{ABC}}\right\|_1^2. \tag{122}$$

We refer to [33] for the latest progress in that direction.

The arguments in this section can also be applied to lift the strengthened monotonicity from [50, Corollary 4.2]. For $\rho \in S(\mathcal{H})$, σ a positive semi-definite operator on $\mathcal{H}$, and $\mathcal{N}$ a completely positive trace preserving map on the same space this leads to

$$D(\rho\|\sigma) - D(\mathcal{N}(\rho)\|\mathcal{N}(\sigma)) \geq \limsup_{n\to\infty}\frac{1}{n}D\left(\rho^{\otimes n}\middle\|\int\beta_0(t)\left(\mathcal{R}^{[t]}_{\sigma,\mathcal{N}}(\rho)\right)^{\otimes n}\mathrm{d}t\right), \tag{123}$$

where $\mathcal{R}^{[t]}_{\sigma,\mathcal{N}}(\cdot) := \sigma^{\frac{1+it}{2}}\mathcal{N}^\dagger\left(\mathcal{N}(\sigma)^{\frac{-1-it}{2}}(\cdot)\mathcal{N}(\sigma)^{\frac{-1+it}{2}}\right)\sigma^{\frac{1-it}{2}}$. Together with [32, Section 3] and [50, Corollary 4.2] we then again have the three lower bounds as in Corollary 4.2.

4.2. Regularization necessary. Here, we use our bound on the conditional quantum mutual information (Theorem 4.1) to show that the regularization in Theorem 1.1 is in general needed (see also [10]). That is, we give a proof for Eq. (17). Namely, by Theorem 4.1 we have[7]

$$I(A:B|C)_\rho \geq \limsup_{n\to\infty}\frac{1}{n}D\left(\rho^{\otimes n}_{ABC}\middle\|\int\beta_0(t)\left(\mathcal{I}_A\otimes\mathcal{R}^{[t]}_{C\to BC}(\rho_{AC})\right)^{\otimes n}\mathrm{d}t\right) \tag{124}$$

$$\geq \lim_{n\to\infty}\frac{1}{n}\inf_{\mu\in\mathcal{R}} D\left(\rho^{\otimes n}_{ABC}\middle\|\int\left(\mathcal{I}_A\otimes\mathcal{R}_{C\to BC}\left(\rho_{AC}\right)\right)^{\otimes n}\mathrm{d}\mu(\mathcal{R})\right). \tag{125}$$

From the second composite discrimination problem described in Sect. 3.2 we see that the latter quantity is equal to the asymptotic error exponent $\zeta_{\bar{\mathcal{R}}}(\infty, 0)$ as given in Eq. (91) for testing

$$\rho^{\otimes n}_{ABC} \text{ against } \int\left((\mathcal{I}_A\otimes\mathcal{R}_{C\to BC})(\rho_{AC})\right)^{\otimes n}\ \mathrm{d}\mu(\mathcal{R}). \tag{126}$$

Now, if the regularization in the asymptotic formula for $\zeta_{\bar{\mathcal{R}}}(\infty, 0)$ would actually not be needed this would imply that

$$I(A:B|C)_\rho \geq \inf_{\mathcal{R}} D\left(\rho_{ABC}\|(\mathcal{I}_A\otimes\mathcal{R}_{C\to BC})(\rho_{AC})\right)\,. \tag{127}$$

However, this is in contradiction with the counterexample from [21, Section 5] as discussed in Eq. (120). Hence, we conclude that the regularization for composite asymmetric quantum hypothesis testing is needed in general. □

[7] Alternatively, we could employ the implicitly stated bound [12, Equation 38].

5. Conclusion

We extended quantum Stein's lemma in asymmetric quantum hypothesis testing by showing that the optimal asymptotic error exponent for testing convex combinations of quantum states $\rho^{\otimes n}$ against convex combinations of quantum states $\sigma^{\otimes n}$ is given by a regularized quantum relative entropy formula which does not become single-letter in general. Moreover, we gave various examples when our formula as well as extensions thereof become single-letter. It remains interesting to find more non-commutative settings that allow for single-letter solutions.

Another related problem is that of symmetric hypothesis testing, where it is well-known that in the case of fixed iid states $\rho^{\otimes n}$ against $\sigma^{\otimes n}$ the optimal asymptotic error exponent is given by the quantum Chernoff bound [1,42]

$$C(\rho,\sigma) := \sup_{0\leq s\leq 1} -\log \mathrm{Tr}\left[\rho^s\sigma^{1-s}\right]. \tag{128}$$

For this symmetric setting, it was conjectured in [2] that for finite sets $\mathcal{S}$ and $\mathcal{T}$ the corresponding composite asymptotic error exponent is given by

$$C(\mathcal{S},\mathcal{T}) := \inf_{\substack{\rho\in\mathcal{S}\\ \sigma\in\mathcal{T}}} C(\rho,\sigma), \tag{129}$$

with definitions analogue to those given earlier for the asymmetric setting. However, it was recently shown that already in the setting of a fixed null hypothesis $\mathcal{S}=\{\rho\}$ above conjecture is in general false [38].[8]

Moreover, one can again consider testing convex combinations of iid states $\rho^{\otimes n}$ with $\rho\in\mathcal{S}$ against convex combinations of iid states $\sigma^{\otimes n}$ with $\sigma\in\mathcal{T}$. Similarly to our work for the asymmetric setting, we then have that the following rate for the asymptotic error exponent is achievable (assuming that the limit exists)

$$\sup_{0\leq s\leq 1}\lim_{n\to\infty}\frac{1}{n}\inf_{\substack{\nu\in\mathcal{S}\\ \mu\in\mathcal{T}}} -\log\mathrm{Tr}\left[\left(\int\rho^{\otimes n}\,\mathrm{d}\nu(\rho)\right)^s\left(\int\sigma^{\otimes n}\,\mathrm{d}\mu(\sigma)\right)^{1-s}\right]. \tag{130}$$

However, it was already shown in [29] that this does in general not simplify to the single-letter form in Eq. (129). We refer to [38, Section I] for an excellent overview of the recent progress on composite hypothesis testing.

Finally, we note that finding single-letter achievability results for composite hypothesis testing problems has important applications in network quantum Shannon theory [45, Section 5.2].

Acknowledgements We thank an anonymous referee for extensive feedback and pointing out detailed solutions to multiple errors in previous versions of this manuscript. This work was completed prior to MB and FB joining the AWS Center for Quantum Computing. CH acknowledges support from the VILLUM FONDEN via the QMATH Centre of Excellence (Grant no. 10059), the Spanish MINECO, project FIS2013-40627-P, FIS2016-80681-P (AEI/FEDER, UE) and FPI Grant No. BES-2014-068888, as well as by the Generalitat de Catalunya, CIRIT project no. 2014-SGR-966.

[8] See, however, [35] for a related problem that does allow for an exact single-letter characterisation.

A. Some Lemmas

Here, we give several lemmas that are used in the main part. We start with Sion's minimax theorem [49].

Lemma A.1. *Let X be a compact convex subset of a linear topological space and Y a convex subset of a linear topological space. If a real-valued function on $X \times Y$ is such that*

- ◇ *$f(x, \cdot)$ is upper semi-continuous and quasi-concave on Y for every $x \in X$*
- ◇ *$f(\cdot, y)$ is lower semi-continuous and quasi-convex on X for every $y \in Y$,*

then we have

$$\min_{x\in X} \sup_{y\in Y} f(x, y) = \sup_{y\in Y} \min_{x\in X} f(x, y) . \tag{131}$$

The following is a special case of [11, Lemma 13], which is based on a more involved minimax theorem taking into account the possibility that the relative entropy can be infinite.

Lemma A.2. *Let $\mathcal{S}, \mathcal{T} \subseteq S(\mathcal{H})$ be closed, convex sets. Then, we have*

$$\min_{\substack{\rho\in\mathcal{S}\\ \sigma\in\mathcal{T}}} D_{\mathcal{M}}(\rho\|\sigma) = \sup_{(\mathcal{X},\mathcal{M})} \min_{\substack{\rho\in\mathcal{S}\\ \sigma\in\mathcal{T}}} D\Big(\sum_{x\in\mathcal{X}} \mathrm{Tr}\,[M_x\rho]\,|x\rangle\langle x| \Big\| \sum_{x\in\mathcal{X}} \mathrm{Tr}\,[M_x\rho]\,|x\rangle\langle x| \Big) . \tag{132}$$

We have the following discretization result.

Lemma A.3. *For every probability measure μ on the Borel σ-algebra of $\mathcal{S} \subseteq S(\mathcal{H})$ with the dimension of $\mathcal{H}$ given by d, there exists a probability distribution $\{p_i\}_i^N$ with $N \leq (n+1)^{2d^2}$ and $\rho_i \in \mathcal{S}$ such that*

$$\int \rho^{\otimes n}\, \mathrm{d}\mu(\rho) = \sum_{i=1}^{N} p_i \rho_i^{\otimes n} . \tag{133}$$

Proof. The idea is to use Carathéodory theorem together with the smallness of the symmetric subspace. For pure states the proof from [6, Corollary D.6] applies and the general case follows immediately by considering purifications and taking the partial trace over the purifying system. □

The von Neumann entropy has the following almost-convexity property (besides its well-known concavity).

Lemma A.4. *Let $\rho_i \in S(\mathcal{H})$ for $i = 1, \ldots, N$ and $\{p_i\}$ be a probability distribution. Then, we have*

$$H\Big(\sum_{i=1}^{N} p_i \rho_i\Big) \leq \sum_{i=1}^{N} p_i H(\rho_i) + \log N . \tag{134}$$

Proof. This follows from elementary quantum entropy inequalities (see, e.g., [41, Chapter 11])

$$H\Big(\sum_{i=1}^{N} p_i \rho_i\Big) \leq \sum_{i=1}^{N} p_i H\,(\rho_i) + H(p_i) \leq \sum_{i=1}^{N} p_i H(\rho_i) + \log N . \tag{135}$$

□

The following is a property of the quantum relative entropy [23, Theorem 3].

Lemma A.5. *Let $\mathcal{N}$ be a trace-preserving, completely positive map with $\mathcal{N}(1) = 1$ (unital) and $\mathcal{N}^2 = \mathcal{N}$ (idempotent). Then, the minimum relative entropy distance between $\rho \in S(\mathcal{H})$ and $\sigma \in S(\mathcal{H})$ in the image of $\mathcal{N}$ satisfies*

$$\inf_{\sigma \in \mathrm{Im}(\mathcal{N})} D(\rho\|\sigma) = H(\mathcal{N}(\rho)) - H(\rho) = D(\rho\|\mathcal{N}(\rho)) . \tag{136}$$

In particular, we have for the relative entropy of coherence $D_{\mathcal{C}}(\rho) = D(\rho\|\rho_{\mathrm{diag}})$, where ρ_{diag} denotes the state obtained from ρ by deleting all off-diagonal elements.

Audenaert's inequality originally used to derive the quantum Chernoff bound can be stated as follows [1, Theorem 1].

Lemma A.6. *Let $X, Y \gg 0$ and $s \in (0, 1)$. Then, we have*

$$\mathrm{Tr}\left[X^s Y^{1-s}\right] \geq \mathrm{Tr}\left[X\,(1 - \{X - Y\}_+)\right] + \mathrm{Tr}\left[Y\,\{X - Y\}_+\right] . \tag{137}$$

References

1. Audenaert, K.M.R., Calsamiglia, J., Munoz-Tapia, R., Bagan, E., Masanes, L., Acin, A., Verstraete, F.: Discriminating states: the quantum Chernoff bound. Phys. Rev. Lett. **98**(16), 160501 (2007). https://doi.org/10.1103/PhysRevLett.98.160501
2. Audenaert, K.M.R., Mosonyi, M.: Upper bounds on the error probabilities and asymptotic error exponents in quantum multiple state discrimination. J. Math. Phys. **55**(10), 102201 (2014). https://doi.org/10.1063/1.4898559
3. Audenaert, K.M.R., Mosonyi, M., Verstraete, F.: Quantum state discrimination bounds for finite sample size. J. Math. Phys. **53**(23), 122205 (2012). https://doi.org/10.1063/1.4768252
4. Audenaert, K.M.R., Nussbaum, M., Szkoła, A., Verstraete, F.: Asymptotic error rates in quantum hypothesis testing. Commun. Math. Phys. **279**(1), 251–283 (2008). https://doi.org/10.1007/s00220-008-0417-5
5. Baumgratz, T., Cramer, M., Plenio, M.B.: Quantifying coherence. Phys. Rev. Lett. **113**(14), 140401 (2014). https://doi.org/10.1103/PhysRevLett.113.140401
6. Berta, M., Christandl, M., Renner, R.: The quantum reverse Shannon theorem based on one-shot information theory. Commun. Math. Phys. **306**(3), 579–615 (2011). https://doi.org/10.1007/s00220-011-1309-7
7. Berta, M., Fawzi, O., Tomamichel, M.: On variational expressions for quantum relative entropies. Lett. Math. Phys. **107**(12), 2239–2265 (2017). https://doi.org/10.1007/s11005-017-0990-7
8. Berta, M., Seshadreesan, K., Wilde, M.: Rényi generalizations of the conditional quantum mutual information. J. Math. Phys. **56**(2), 022205 (2015). https://doi.org/10.1063/1.4908102

9. Berta, M., Tomamichel, M.: The fidelity of recovery is multiplicative. IEEE Trans. Inf. Theory **62**(4), 1758–1763 (2016). https://doi.org/10.1109/TIT.2016.2527683
10. Bjelaković, I., Deuschel, J.-D., Krüger, T., Seiler, R., Siegmund-Schultze, R., Szkoła, A.: A quantum version of Sanov's theorem. Commun. Math. Phys. **260**(3), 659–671 (2005). https://doi.org/10.1007/s00220-005-1426-2
11. Brandao, F.G.S.L., Harrow, A.W., Lee, J.R., Peres, Y.: Adversarial hypothesis testing and a quantum Stein's lemma for restricted measurements. IEEE Trans. Inf. Theory **66**(8), 5037–5054 (2020). https://doi.org/10.1109/TIT.2020.2979704
12. Brandao, F.G.S.L., Harrow, A.W., Oppenheim, J., Strelchuk, S.: Quantum conditional mutual information, reconstructed states, and state redistribution. Phys. Rev. Lett. **115**(5), 050501 (2015). https://doi.org/10.1103/PhysRevLett.115.050501
13. Brandao, F.G.S.L., Plenio, M.B.: A generalization of quantum Stein's lemma. Commun. Math. Phys. **295**(3), 791–828 (2010). https://doi.org/10.1007/s00220-010-1005-z
14. Chitambar, E., Gour, G.: Comparison of incoherent operations and measures of coherence. Phys. Rev. A **94**(5), 052336 (2016). https://doi.org/10.1103/PhysRevA.94.052336
15. Christandl, M., König, R., Renner, R.: Postselection technique for quantum channels with applications to quantum cryptography. Phys. Rev. Lett. **102**(2), 020504 (2009). https://doi.org/10.1103/PhysRevLett.102.020504
16. Cooney, T., Hirche, C., Morgan, C., Olson, J.P., Seshadreesan, K.P., Watrous, J., Wilde, M.M.: Operational meaning of quantum measures of recovery. Phys. Rev. A **94**(2), 022310 (2016). https://doi.org/10.1103/PhysRevA.94.022310
17. Datta, N.: Max-relative entropy of entanglement, alias log robustness. Int. J. Quantum Inf. **7**(2), 475–491 (2009)
18. Datta, N.: Min- and max-relative entropies and a new entanglement monotone. IEEE Trans. Inf. Theory **55**(6), 2816–2826 (2009). https://doi.org/10.1109/TIT.2009.2018325
19. Datta, N., Mosonyi, M., Hsieh, M.-H., Brandao, F.G.S.L.: A smooth entropy approach to quantum hypothesis testing and the classical capacity of quantum channels. IEEE Trans. Inf. Theory **59**(12), 8014–8026 (2013). https://doi.org/10.1109/TIT.2013.2282160
20. Donald, M.J.: On the relative entropy. Commun. Math. Phys. **105**(1), 13–34 (1986). https://doi.org/10.1007/BF01212339
21. Fawzi, H., Fawzi, O.: Efficient optimization of the quantum relative entropy. J. Phys. A Math. Theor. **51**(15), 154003 (2018). https://doi.org/10.1088/1751-8121/aab285
22. Fawzi, O., Renner, R.: Quantum conditional mutual information and approximate Markov chains. Commun. Math. Phys. **340**(2), 575–611 (2015). https://doi.org/10.1007/s00220-015-2466-x
23. Gour, G., Marvian, I., Spekkens, R.W.: Measuring the quality of a quantum reference frame: the relative entropy of frameness. Phys. Rev. A **80**(1), 012307 (2009). https://doi.org/10.1103/PhysRevA.80.012307
24. Harrow, A.W.: Applications of coherent classical communication and Schur duality to quantum information theory. PhD thesis, MIT (2005)
25. Hayashi, M.: Optimal sequence of quantum measurements in the sense of Stein's lemma in quantum hypothesis testing. J. Phys. A Math. Gen. **35**(50), 10759 (2002). https://doi.org/10.1088/0305-4470/35/50/307
26. Hayashi, M.: Universal coding for classical-quantum channel. Commun. Math. Phys. **289**(3), 1087–1098 (2009). https://doi.org/10.1007/s00220-009-0825-1
27. Hayashi, M., Nagaoka, H.: General formulas for capacity of classical-quantum channels. IEEE Trans. Inf. Theory **49**(7), 1753–1768 (2003). https://doi.org/10.1109/TIT.2003.813556
28. Hayashi, M., Tomamichel, M.: Correlation detection and an operational interpretation of the Rényi mutual information. J. Math. Phys. **57**(10), 102201 (2016). https://doi.org/10.1063/1.4964755
29. Hiai, F., Mosonyi, M., Hayashi, M.: Quantum hypothesis testing with group symmetry. J. Math. Phys. **50**(10), 103304 (2009). https://doi.org/10.1063/1.3234186
30. Hiai, F., Petz, D.: The proper formula for relative entropy and its asymptotics in quantum probability. Commun. Math. Phys. **143**(1), 99–114 (1991). https://doi.org/10.1007/BF02100287
31. Jain, R., Radhakrishnan, J., Sen, P.: Privacy and interaction in quantum communication complexity and a theorem about the relative entropy of quantum states. Proc. IEEE FOCS **2002**, 429–438 (2002). https://doi.org/10.1109/SFCS.2002.1181967
32. Junge, M., Renner, R., Sutter, D., Wilde, M.M., Winter, A.: Universal recovery maps and approximate sufficiency of quantum relative entropy. Annales Henri Poincaré **19**(10), 2955–2978 (2018). https://doi.org/10.1007/s00023-018-0716-0
33. Lemm, M.: On multivariate trace inequalities of Sutter, Berta and Tomamichel. J. Math. Phys. **59**(1), 012204 (2018). https://doi.org/10.1063/1.5001009
34. Levitan, E., Merhav, N.: A competitive Neyman–Pearson approach to universal hypothesis testing with applications. IEEE Trans. Inf. Theory **48**(8), 2215–2229 (2002). https://doi.org/10.1109/TIT.2002.800478

35. Li, K.: Discriminating quantum states: the multiple Chernoff distance. Ann. Stat. **44**(4), 1661–1679 (2016). https://doi.org/10.1214/16-AOS1436
36. Lindblad, G.: Completely positive maps and entropy inequalities. Commun. Math. Phys. **40**(2), 147–151 (1975). https://doi.org/10.1007/BF01609396
37. Lloyd, S.: Enhanced sensitivity of photodetection via quantum illumination. Science **321**(5895), 1463–1465 (2008). https://doi.org/10.1126/science.1160627
38. Mosonyi, M., Szilágyi, Z., Weiner, M.: On the error exponents of binary quantum state discrimination with composite hypotheses. arXiv:2011.04645 (2020)
39. Müller-Lennert, M., Dupuis, F., Szehr, O., Fehr, S., Tomamichel, M.: On quantum Rényi entropies: a new generalization and some properties. J. Math. Phys. **54**(12), 122203 (2013). https://doi.org/10.1063/1.4838856
40. Nagaoka, H.: Strong Converse Theorems in Quantum Information Theory, pp. 64–65. World Scientific, Singapore (2005). https://doi.org/10.1142/9789812563071_0005
41. Nielsen, M.A., Chuang, I.L.: Quantum Information and Quantum Computation. Cambridge University Press, Cambridge (2000)
42. Nussbaum, M., Szkoła, A.: The Chernoff lower bound for symmetric quantum hypothesis testing. Ann. Stat. **37**(2), 1040–1057 (2009). https://doi.org/10.1214/08-AOS593
43. Ogawa, T., Nagaoka, H.: Strong converse and Stein's lemma in quantum hypothesis testing. IEEE Trans. Inf. Theory **46**(7), 2428–2433 (2000). https://doi.org/10.1109/18.887855
44. Ohya, M., Petz, D.: Quantum Entropy and Its Use. Springer, Berlin (1993)
45. Qi, H., Wang, Q., Wilde, M.M.: Applications of position-based coding to classical communication over quantum channels. J. Phys. A Math. Theor. **51**(44), 444002 (2018). https://doi.org/10.1088/1751-8121/aae290
46. Tan, S.-H., Erkmen, B.I., Giovannetti, V., Guha, S., Lloyd, S., Maccone, L., Pirandola, S., Shapiro, J.H.: Quantum illumination with Gaussian states. Phys. Rev. Lett. **101**(25), 253601 (2008). https://doi.org/10.1103/PhysRevLett.101.253601
47. Seshadreesan, K.P., Wilde, M.M.: Fidelity of recovery, squashed entanglement, and measurement recoverability. Phys. Rev. A **92**(4), 042321 (2015). https://doi.org/10.1103/PhysRevA.92.042321
48. Sharma, N., Warsi, N.A.: Fundamental bound on the reliability of quantum information transmission. Phys. Rev. Lett. **110**(8), 080501 (2013). https://doi.org/10.1103/PhysRevLett.110.080501
49. Sion, M.: On general minimax theorems. Pac. J. Math. **8**, 171–176 (1958). https://doi.org/10.2140/pjm.1958.8.171
50. Sutter, D., Berta, M., Tomamichel, M.: Multivariate trace inequalities. Commun. Math. Phys. **352**(1), 37–58 (2017). https://doi.org/10.1007/s00220-016-2778-5
51. Sutter, D., Tomamichel, M., Harrow, A.W.: Strengthened monotonicity of relative entropy via pinched Petz recovery map. IEEE Trans. Inf. Theory **62**(5), 2907–2913 (2016). https://doi.org/10.1109/TIT.2016.2545680
52. Tomamichel, M.: Quantum Information Processing with Finite Resources: Mathematical Foundations. SpringerBriefs in Mathematical Physics, vol. 5. Springer, Berlin (2015). https://doi.org/10.1007/978-3-319-21891-5
53. Tomamichel, M., Hayashi, M.: Operational interpretation of Rényi information measures via composite hypothesis testing against product and Markov distributions. IEEE Trans. Inf. Theory **64**(2), 1064–1082 (2018). https://doi.org/10.1109/TIT.2017.2776900
54. Vollbrecht, K.G.H., Werner, R.F.: Entanglement measures under symmetry. Phys. Rev. A **64**(6), 062307 (2001). https://doi.org/10.1103/PhysRevA.64.062307
55. Wilde, M.M.: Recoverability in quantum information theory. Proc. R. Soc. Lond. A Math. Phys. Eng. Sci. **471**(2182), 20150338 (2015). https://doi.org/10.1098/rspa.2015.0338
56. Wilde, M.M., Tomamichel, M., Lloyd, S., Berta, M.: Gaussian hypothesis testing and quantum illumination. Phys. Rev. Lett. **119**(12), 120501 (2017). https://doi.org/10.1103/PhysRevLett.119.120501
57. 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. **331**(2), 593–622 (2014). https://doi.org/10.1007/s00220-014-2122-x

Communicated by H-T. Yau