

“© 2013 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.”

# Chain Rules for Smooth Min- and Max-Entropies

Alexander Vitanov, Frédéric Dupuis, Marco Tomamichel, and Renato Renner

**Abstract**—The chain rule for the Shannon and von Neumann entropy, which relates the total entropy of a system to the entropies of its parts, is of central importance to information theory. Here we consider the chain rule for the more general smooth min- and max-entropy, used in one-shot information theory. For these entropy measures, the chain rule no longer holds as an equality. However, the standard chain rule for the von Neumann entropy is retrieved asymptotically when evaluating them for many identical and independently distributed states.

## I. INTRODUCTION

IN classical and quantum information theory, entropy measures are often used to characterize fundamental information processing tasks. For example, in his groundbreaking work on information and communication theory [14], Shannon showed that entropies can be used to quantify the memory needed to store the (compressed) output of an information source or the capacity of a communication channel. It follows immediately from the basic properties of the Shannon entropy that the equality

$$H(AB) = H(A|B) + H(B),$$

which we call the *chain rule*, must hold. Here,  $H(B)$  denotes the entropy of the random variable  $B$  and  $H(A|B)$  is the entropy of the random variable  $A$  averaged over *side information* in  $B$ . The chain rule therefore asserts that the entropy of two (possibly correlated) random variables,  $A$  and  $B$ , can be decomposed into the entropy of  $B$  alone plus the entropy of  $A$  conditioned on knowing  $B$ . More generally, one may average over additional side information,  $C$ , in which case the chain rule takes the more general form

$$H(AB|C) = H(A|BC) + H(B|C). \quad (1)$$

The chain rule forms an integral part of the entropy calculus. The other basic ingredient is strong sub-additivity, which can be written as  $H(A|BC) \leq H(A|C)$ , i.e. additional side information can only decrease the entropy.

The quantum generalization of Shannon's entropy, the *von Neumann entropy*, inherits these fundamental properties. For a quantum state<sup>1</sup>  $\rho_A$  on  $A$ , the von Neumann entropy is

A. Vitanov is with the Department of Mathematics, ETH Zurich, Rämistrasse 101, 8092 Zürich. This work was produced while F. Dupuis was with the Institute for Theoretical Physics, ETH Zurich, 8093 Zürich, Switzerland. Since January 2012 he is with the Department of Computer Science at University of Aarhus, Åbogade 34, 8200 Aarhus N, Denmark. R. Renner is with the Institute for Theoretical Physics, ETH Zurich, 8093 Zürich, Switzerland. M. Tomamichel is with the Center for Quantum Technologies, National University of Singapore, 3 Science Drive 2, Singapore 117543. (e-mail: alexander.vitanov@math.ethz.ch; dupuis@cs.au.dk; cqtmarco@nus.edu.sg; renner@phys.ethz.ch)

Copyright (c) 2012 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to pubs-permissions@ieee.org.

<sup>1</sup>Formal definitions follow in Section II.

defined as  $H(A)_\rho := -\text{tr}(\rho_A \log \rho_A)$ , where  $\text{tr}$  denotes the trace and  $\log$  is taken in base 2 throughout this paper. The conditional von Neumann entropy with classical side information can again be defined by an average, however, this intuitive definition fails if the side information is quantum. Pointing to its fundamental importance, the conditional von Neumann entropy is thus defined by the chain rule itself, i.e.  $H(A|B)_\rho := H(AB)_\rho - H(B)_\rho$ . In addition to the chain rule and strong sub-additivity, it also satisfies a duality relation: For any pure tripartite state  $\rho_{ABC}$ , we have  $H(A|B)_\rho = -H(A|C)_\rho$ .

Shannon and von Neumann entropies have been successfully employed to characterize an enormous variety of information theoretic tasks, many of which are of high practical relevance (examples include the aforementioned tasks of data compression or channel coding). However, a basic assumption usually made in this context is that the underlying random processes (e.g., those relevant for the generation of data, or the occurrence of noise in a communication channel) are modeled asymptotically by an arbitrarily long sequence of random variables that are *independent and identically distributed (i.i.d.)*. In the absence of this assumption (e.g., if a channel is only invoked a small number of times or if its noise model is not i.i.d.), the use of the von Neumann entropy is generally no longer justified. The formalism of smooth min- and max-entropy, introduced in [11]–[13] and further developed in [5], [9], [16], [17], overcomes this limitation and enables the analysis of general situations beyond the i.i.d. scenario. This level of generality turned out to be crucial in various areas, e.g., in physics (where entropies are employed for the analysis of problems in thermodynamics [6]) or in cryptography (where entropies are used to quantify an adversary's uncertainty).

Smooth min- and max-entropy, denoted  $H_{\min}^\varepsilon$  and  $H_{\max}^\varepsilon$ , respectively, depend on a positive real value  $\varepsilon$ , called *smoothing parameter*  $\varepsilon$  (see Section II for formal definitions). When the entropies are used to characterize operational tasks, the smoothing parameter determines the desired accuracy. For example, the smooth min-entropy,  $H_{\min}^\varepsilon(A|B)$ , characterizes the number of fully mixed qubits, independent (i.e. *decoupled*) from side information  $B$ , that can be extracted from a quantum source  $A$  [7], [8]. Furthermore, the smooth max-entropy,  $H_{\max}^\varepsilon(A|B)$ , characterizes the amount of entanglement needed between two parties,  $A$  and  $B$ , to merge a state  $\rho_{AB}$ , where  $\rho_A$  is initially held by  $A$ , to  $B$  [3], [8]. In both cases, the smoothing parameter  $\varepsilon$  corresponds to the maximum distance between the desired final state and the one that can be achieved.

Smooth entropy can be seen as strict generalization of Shannon or von Neumann entropy. In particular, the latter can be recovered by evaluating the smooth min- or max-entropy for i.i.d. states [11], [16]. Accordingly, smooth entropy inherits

many of the basic features of von Neumann entropy, such as strong sub-additivity. In light of this, it should not come as a surprise that smooth entropy also obeys inequalities that generalize the chain rule (1). Deriving these is the main aim of this work.

Specifically, one can obtain four pairs of generalized chain inequalities. For any small smoothing parameters  $\varepsilon', \varepsilon'', \varepsilon''' \geq 0$  and  $\varepsilon > \varepsilon' + 2\varepsilon''$ , we have

$$\begin{aligned} H_{\min}^{\varepsilon}(AB|C)_{\rho} &\geq H_{\min}^{\varepsilon''}(A|BC)_{\rho} + H_{\min}^{\varepsilon'}(B|C)_{\rho} - f, \\ H_{\max}^{\varepsilon}(AB|C)_{\rho} &\leq H_{\max}^{\varepsilon''}(A|BC)_{\rho} + H_{\max}^{\varepsilon'}(B|C)_{\rho} + f, \\ H_{\min}^{\varepsilon'}(AB|C)_{\rho} &\leq H_{\min}^{\varepsilon}(A|BC)_{\rho} + H_{\max}^{\varepsilon''}(B|C)_{\rho} + 2f, \\ H_{\max}^{\varepsilon'}(AB|C)_{\rho} &\geq H_{\min}^{\varepsilon''}(A|BC)_{\rho} + H_{\max}^{\varepsilon}(B|C)_{\rho} - 2f, \\ H_{\min}^{\varepsilon'}(AB|C)_{\rho} &\leq H_{\max}^{\varepsilon''}(A|BC)_{\rho} + H_{\min}^{\varepsilon}(B|C)_{\rho} + 3f, \\ H_{\max}^{\varepsilon'}(AB|C)_{\rho} &\geq H_{\max}^{\varepsilon}(A|BC)_{\rho} + H_{\min}^{\varepsilon''}(B|C)_{\rho} - 3f, \\ H_{\min}^{\varepsilon'}(AB|C)_{\rho} &\leq H_{\max}^{\varepsilon'''}(A|BC)_{\rho} + H_{\max}^{\varepsilon''}(B|C)_{\rho} + g, \\ H_{\max}^{\varepsilon'}(AB|C)_{\rho} &\geq H_{\min}^{\varepsilon'''}(A|BC)_{\rho} + H_{\min}^{\varepsilon''}(B|C)_{\rho} - g, \end{aligned}$$

where  $f$  does not grow more than of the order  $\log 1/e$  when  $e = \varepsilon - \varepsilon' - 2\varepsilon''$  is small, and  $g$  is smaller than 6 for  $\varepsilon' + 2\varepsilon'' + \varepsilon''' < 1/5$ . We note that, in typical applications, we would choose the smoothing parameters so that the correction terms  $f$  and  $g$  are small compared to the typical values of the smooth entropies.

The fact that generalized chain inequalities hold for smooth min- and max-entropy is not only important for establishing a complete entropy calculus, analogous to that for the von Neumann entropy. They are also crucial for applications, as the following example shows.

In quantum key distribution, after the quantum signals have been exchanged and measured, two honest parties, Alice and Bob, are left with two correlated raw keys, about which a potential eavesdropper is guaranteed to have only limited information. This limit on the eavesdropper's knowledge is best expressed [11] by a bound on the smooth min-entropy of Alice's raw key,  $X_A$ , conditioned on the eavesdropper's quantum information,  $E$ , i.e.,  $H_{\min}^{\varepsilon'}(X_A|E)$ . However, to ensure that Bob's final key agrees with her own, Alice will have to send a syndrome,  $S = s(X_A)$ , over an insecure channel. A fundamental question in quantum key distribution is thus to bound  $H_{\min}^{\varepsilon}(X_A|ES)$ , i.e., the smooth min-entropy of  $X_A$  conditioned on the eavesdropper's information after learning  $S$ . The third chain rule above states that

$$\begin{aligned} H_{\min}^{\varepsilon}(X_A|ES) &\geq H_{\min}^{\varepsilon'}(X_A S|E) - H_{\max}^{\varepsilon''}(S|E) - 2f \\ &= H_{\min}^{\varepsilon'}(X_A|E) - H_{\max}^{\varepsilon''}(S|E) - 2f. \end{aligned}$$

Here, we used that  $S = s(X_A)$  and thus  $X_A \rightarrow X_A S$  is an isometry under which the smooth entropies are invariant [17]. Roughly speaking, our chain rule thus implies that the eavesdropper gains at most  $H_{\max}^{\varepsilon''}(S|E)$  bits of information about  $X_A$ , where we assumed that  $f$  is negligible. This is strictly tighter than previous results (see, e.g., [21]), where the gain was bounded by  $\log |S| \geq H_{\max}^{\varepsilon''}(S|E)$ , where  $|S|$  is the number of different syndromes that can be stored in  $S$ . This

leads to strictly tighter bounds, for instance, when  $S$  contains information that has been communicated previously over the public channel and is therefore already included in  $E$ .

Until now, only special cases of the above inequalities have been known, except for the first pair, which has been derived in [8]. In the present paper we provide proofs for the remaining relations. In fact, since smooth min- and max-entropy obey a duality relation similar to that of von Neumann entropy [17],  $H_{\min}^{\varepsilon}(A|B) = -H_{\max}^{\varepsilon}(A|C)$ , the paired inequalities above imply each other. It will therefore suffice to prove only one inequality of each pair.

The paper is organized as follows. In the next section we introduce the notation, terminology, and basic definitions. In particular, we define the (smooth) min- and max-entropy measures and outline some of their basic features. In Section III we derive alternative expressions for the max-entropy based on semidefinite programming duality. While these expressions may be of independent interest, they will be used in Section IV, which is devoted to the statement and proofs of the generalized chain rules.

## II. MATHEMATICAL PRELIMINARIES

### A. Notation and basic definitions

Throughout this paper we focus on finite dimensional Hilbert spaces. Hilbert spaces corresponding to different physical systems are distinguished by different capital Latin letters as subscript  $\mathcal{H}_A, \mathcal{H}_B$  etc. The tensor product of  $\mathcal{H}_A$  and  $\mathcal{H}_B$  is designated in short by  $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$ .

The set of linear operators from  $\mathcal{H}_A$  to  $\mathcal{H}_B$  is denoted by  $\mathcal{L}(\mathcal{H}_A, \mathcal{H}_B)$ . The space of linear operators acting on the Hilbert space  $\mathcal{H}$  is denoted by  $\mathcal{L}(\mathcal{H})$  and the subset of  $\mathcal{L}(\mathcal{H})$  containing the Hermitian operators on  $\mathcal{H}$  is denoted by  $\text{Herm}(\mathcal{H})$ . Note that  $\text{Herm}(\mathcal{H})$  endowed with the Hilbert-Schmidt inner product  $\langle X, Y \rangle := \text{tr}(X^\dagger Y)$ ,  $X, Y \in \text{Herm}(\mathcal{H})$ , is a Hilbert space. Given an operator  $R \in \text{Herm}(\mathcal{H})$ , we write  $R \geq 0$  if and only if  $R$  is positive semi-definite and  $R > 0$  if and only if it is positive definite. Furthermore, let  $\mathcal{S}_{\leq}(\mathcal{H})$  and  $\mathcal{S}_{=}(\mathcal{H})$  denote the sets of sub-normalized and normalized positive semi-definite *density operators* with  $\text{tr} \rho \leq 1$  and  $\text{tr} \rho = 1$ , respectively.

Inequalities between Hermitian operators are defined in the following sense: Let  $R, S \in \text{Herm}(\mathcal{H})$ , then we write  $R \geq S$ , respectively  $R > S$  if and only if  $R - S$  is positive semi-definite, respectively positive definite.

Given an operator  $R$ , the operator norm of  $R$  is denoted by  $\|R\|_{\infty}$  and is equal to the highest singular value of  $R$ . The trace norm of  $R$  is given by  $\|R\|_1 := \text{tr}[\sqrt{R^\dagger R}]$ . The fidelity between two states  $\rho, \sigma \in \mathcal{S}_{\leq}(\mathcal{H})$  is defined as  $F(\rho, \sigma) := \|\sqrt{\rho}\sqrt{\sigma}\|_1$ .

For multipartite operators on product spaces  $\mathcal{H}_{AB}$  we will use subscripts to denote the space on which they act (e.g.  $S_{AB}$  for an operator on  $\mathcal{H}_{AB}$ ). Given a multipartite operator  $S_{AB} \in \mathcal{L}(\mathcal{H}_{AB})$ , the corresponding reduced operator on  $\mathcal{H}_A$  is defined by  $S_A := \text{tr}_B[S_{AB}]$  where  $\text{tr}_B$  denotes the partial trace operator on the subsystem  $\mathcal{H}_B$ . Given a multipartite operator  $S_{AB}$  and the corresponding marginal operator  $S_A$ , we call  $S_{AB}$  an *extension* of  $S_A$ . We omit identities from

expressions which involve multipartite operators whenever mathematically meaningful expressions can be obtained by tensoring the corresponding identities to the operators.

### B. Smooth Min- and Max-Entropies

In the following we successively give the definitions of the non-smooth min- and max-entropies and their smooth versions [11], [9].

**Definition 1.** Let  $\rho_{AB} \in \mathcal{S}_{\leq}(\mathcal{H}_{AB})$ , then the min-entropy of  $A$  conditioned on  $B$  of  $\rho_{AB}$  is defined as

$$H_{\min}(A|B)_{\rho} := \max_{\sigma_B \in \mathcal{S}_{\leq}(\mathcal{H}_B)} H_{\min}(A|B)_{\rho|\sigma}, \quad \text{where}$$

$$H_{\min}(A|B)_{\rho|\sigma} := \sup\{\lambda \in \mathbb{R} : \rho_{AB} \leq 2^{-\lambda} \mathbb{I}_A \otimes \sigma_B\}. \quad (2)$$

Note that  $H_{\min}(A|B)_{\rho|\sigma}$  is finite if and only if  $\text{supp}(\rho_B) \subseteq \text{supp}(\sigma_B)$  and divergent otherwise.

**Definition 2.** Let  $\rho_{AB} \in \mathcal{S}_{\leq}(\mathcal{H}_{AB})$ , then the max-entropy of  $A$  conditioned on  $B$  of  $\rho_{AB}$  is defined as

$$H_{\max}(A|B)_{\rho} := \max_{\sigma_B \in \mathcal{S}_{\leq}(\mathcal{H}_B)} H_{\max}(A|B)_{\rho|\sigma}, \quad \text{where}$$

$$H_{\max}(A|B)_{\rho|\sigma} := \log F(\rho_{AB}, \mathbb{I}_A \otimes \sigma_B)^2. \quad (3)$$

The maximum in (2) and (3) is achieved at  $\mathcal{S}_{=}(\mathcal{H}_B)$ . The  $\varepsilon$ -smooth min- and max-entropies of a state  $\rho$  can be understood as an optimization of the corresponding non-smooth quantities over a set of states  $\varepsilon$ -close to  $\rho$ . We use the *purified distance* to quantify the  $\varepsilon$ -closeness of states.

**Definition 3.** Let  $\rho, \sigma \in \mathcal{S}_{\leq}(\mathcal{H})$ . Then the purified distance between  $\rho$  and  $\sigma$  is defined by

$$P(\rho, \sigma) := \sqrt{1 - \bar{F}(\rho, \sigma)^2}, \quad \text{where} \quad (4)$$

$$\bar{F}(\rho, \sigma) := F(\rho, \sigma) + \sqrt{(1 - \text{tr } \rho)(1 - \text{tr } \sigma)} \quad (5)$$

is the generalized fidelity.

Hereafter, when two states  $\rho, \sigma \in \mathcal{S}_{\leq}(\mathcal{H})$  are said to be  $\varepsilon$ -close we mean  $P(\rho, \sigma) \leq \varepsilon$  and denote this by  $\rho \approx_{\varepsilon} \sigma$ . Some of the basic properties of the purified distance are reviewed in Appendix B, but for a more comprehensive treatment we refer to [17]. With that convention we are ready to introduce a smoothed version of the min- and max-entropies [11].

**Definition 4.** Let  $\varepsilon \geq 0$ ,  $\rho_{AB} \in \mathcal{S}_{\leq}(\mathcal{H}_{AB})$ . Then the  $\varepsilon$ -smooth min-entropy of  $A$  conditioned on  $B$  of  $\rho_{AB}$  is defined as

$$H_{\min}^{\varepsilon}(A|B)_{\rho} := \max_{\tilde{\rho}} H_{\min}(A|B)_{\tilde{\rho}} \quad (6)$$

and the  $\varepsilon$ -smooth max-entropy of  $A$  conditioned on  $B$  of  $\rho_{AB}$  is defined as

$$H_{\max}^{\varepsilon}(A|B)_{\rho} := \min_{\tilde{\rho}} H_{\max}(A|B)_{\tilde{\rho}} \quad (7)$$

where the maximum and the minimum range over all sub-normalized states  $\tilde{\rho}_{AB} \approx_{\varepsilon} \rho_{AB}$ .

The smooth entropies are dual to each other in the following sense. When  $\rho_{ABC} \in \mathcal{S}_{\leq}(\mathcal{H}_{ABC})$  is pure, we have [17]

$$H_{\max}^{\varepsilon}(A|B)_{\rho} = -H_{\min}^{\varepsilon}(A|C)_{\rho}. \quad (8)$$

Finally, the smooth min-entropy is upper-bounded by the smooth max-entropy as shown by the following lemma whose proof is deferred to Appendix A:

**Lemma 5.** Let  $\varepsilon, \varepsilon' \geq 0$  and let  $\rho_{AB} \in \mathcal{S}_{\leq}(\mathcal{H}_{AB})$  be such that  $\varepsilon + \varepsilon' + 2\sqrt{1 - \text{tr } \rho_{AB}} < 1$ . Then,

$$H_{\min}^{\varepsilon'}(A|B)_{\rho} \leq H_{\max}^{\varepsilon}(A|B)_{\rho} + \log \left( \frac{1}{1 - (\varepsilon + \varepsilon' + 2\sqrt{1 - \text{tr } \rho})^2} \right). \quad (9)$$

### C. Semidefinite Programming

This subsection is devoted to the duality theory of semidefinite programs (SDPs). We will present the subject as given in [2] and especially in [20] but will restrict the discussion to the special case which is of interest in this work.

A semidefinite program over the Hilbert spaces  $\mathcal{H}_A$  and  $\mathcal{H}_B$  is a triple  $(\mathcal{F}, R_A, S_B)$ ,  $\mathcal{F} \in \mathcal{L}(\text{Herm}(\mathcal{H}_A), \text{Herm}(\mathcal{H}_B))$ ,  $R_A \in \text{Herm}(\mathcal{H}_A)$  and  $S_B \in \text{Herm}(\mathcal{H}_B)$ , which is associated with the following two optimization problems:

PRIMAL PROBLEM:

$$\begin{aligned} \text{minimize: } & \text{tr}[R_A X_A] \\ \text{subject to: } & \mathcal{F}(X_A) \geq S_B \\ & X_A \geq 0 \end{aligned}$$

DUAL PROBLEM:

$$\begin{aligned} \text{maximize: } & \text{tr}[S_B Y_B] \\ \text{subject to: } & \mathcal{F}^{\dagger}(Y_B) \leq R_A \\ & Y_B \geq 0 \end{aligned}$$

where  $X_A \in \text{Herm}(\mathcal{H}_A)$  and  $Y_B \in \text{Herm}(\mathcal{H}_B)$  are variables.  $X_A \geq 0$  and  $Y_B \geq 0$  such that  $\mathcal{F}(X_A) \geq S_B$  and  $\mathcal{F}^{\dagger}(Y_B) \leq R_A$ , respectively, are called *primal feasible plan* and *dual feasible plan*, respectively. We also denote the solutions to the primal and dual problems by

$$\begin{aligned} \gamma &:= \inf\{\text{tr}[R_A X_A] : X_A \text{ is a primal feasible plan}\}, \\ \delta &:= \sup\{\text{tr}[S_B Y_B] : Y_B \text{ is a dual feasible plan}\}. \end{aligned}$$

The values  $X_A \geq 0$  and  $Y_B \geq 0$  satisfying  $\text{tr}[R_A X_A] = \gamma$  and  $\text{tr}[S_B Y_B] = \delta$  are called *primal optimal plan*, respectively *dual optimal plan*.

According to the *weak duality theorem*  $\gamma \geq \delta$ . The difference  $\gamma - \delta$  is called *duality gap*. The following theorem called *Slater's condition* establishes an easy-to-check condition under which the duality gap vanishes, that is,  $\gamma = \delta$ .

**Theorem 6.** Let  $\gamma$  and  $\delta$  be defined as above and  $(\mathcal{F}, R_A, S_A)$  with  $R_A \in \text{Herm}(\mathcal{H}_A)$  and  $S_B \in \text{Herm}(\mathcal{H}_B)$  a semi-definite program. Then the following two implications hold:

(i) [Strict dual feasibility] Suppose  $\gamma$  is finite and that there exists an operator  $Y_B > 0$  such that  $\mathcal{F}^{\dagger}(Y_B) < R_A$ . Then  $\gamma = \delta$ .

(ii) [Strict primal feasibility] Suppose that  $\delta$  is finite and that there exists an operator  $X_A > 0$  such that  $\mathcal{F}(X_A) > S_B$ . Then  $\gamma = \delta$ .

### III. NEW EXPRESSIONS AND BOUNDS FOR THE SMOOTH MAX-ENTROPY

In the following, we give alternative expressions for  $H_{\max}(A|B)_{\rho|\sigma}$  and  $H_{\max}(A|B)_{\rho}$  based on the analysis of SDPs. Then, we prove inequalities relating these entropies with a new entropic measure that turns out to be a useful tool for proving the chain rules.

### A. New Expressions via SDP Duality

**Lemma 7.** Let  $\rho_{AB} \in \mathcal{S}_{\leq}(\mathcal{H}_{AB})$ ,  $\sigma_B \in \mathcal{S}_{\leq}(\mathcal{H}_B)$  and let  $\rho_{ABC}$  be a purification of  $\rho_{AB}$  on an auxiliary Hilbert space  $\mathcal{H}_C$ . Then the max-entropy of  $A$  conditioned on  $B$  of  $\rho_{AB}$  relative to  $\sigma_B$  is given by

$$H_{\max}(A|B)_{\rho|\sigma} = \log \min_{Z_{AB}} \text{tr}[(\mathbb{I}_A \otimes \sigma_B)Z_{AB}], \quad (10)$$

where the minimum ranges over all  $Z_{AB} \in \mathcal{P}(\mathcal{H}_{AB})$  with  $\rho_{ABC} \leq Z_{AB} \otimes \mathbb{I}_C$ .

*Proof:* Uhlmann's theorem [19] tells us that the fidelity can be expressed as a maximization of the overlap of purifications in which the optimization goes over one purification only. In particular, if  $\rho_{ABC}$  is any purification of  $\rho_{AB}$ , then by Uhlmann's theorem

$$\begin{aligned} 2^{H_{\max}(A|B)_{\rho|\sigma}} &= F(\rho_{AB}, \mathbb{I}_A \otimes \sigma_B)^2 \\ &= \max_{\substack{X_{ABC} \geq 0 \\ \text{tr}_C[X_{ABC}] = \mathbb{I}_A \otimes \sigma_B \\ \text{rank}[X_{ABC}] = 1}} F(\rho_{ABC}, X_{ABC})^2 \\ &\leq \max_{\substack{X_{ABC} \geq 0 \\ \text{tr}_C[X_{ABC}] = \mathbb{I}_A \otimes \sigma_B}} \text{tr}[\rho_{ABC} X_{ABC}] \\ &= \max_{\substack{X_{ABC} \geq 0 \\ \text{tr}_C[X_{ABC}] = \mathbb{I}_A \otimes \sigma_B}} F(\rho_{ABC}, X_{ABC})^2 \\ &\leq F(\rho_{AB}, \mathbb{I}_A \otimes \sigma_B)^2 \\ &= 2^{H_{\max}(A|B)_{\rho|\sigma}}, \end{aligned}$$

where the first inequality follows from the fact that the set over which we optimize becomes larger and the last inequality follows from the fact that the fidelity is monotonously increasing under the partial trace. The above calculation implies that instead of optimizing over rank one operators  $X_{ABC}$  as Uhlmann's theorem demands, one can maximize over all positive semidefinite extensions  $X_{ABC}$  of  $\mathbb{I}_A \otimes \sigma_B$ , that is,

$$2^{H_{\max}(A|B)_{\rho|\sigma}} = \max_{\substack{X_{ABC} \geq 0 \\ \text{tr}_C[X_{ABC}] = \mathbb{I}_A \otimes \sigma_B}} \text{tr}[\rho_{ABC} X_{ABC}]. \quad (11)$$

Moreover, for any positive semidefinite operator  $X_{ABC}$  with  $\text{tr}_C[X_{ABC}] \leq \mathbb{I}_A \otimes \sigma_B$  we can define an operator

$$\bar{X}_{ABC} := X_{ABC} + Y_C \otimes (\mathbb{I}_A \otimes \sigma_B - \text{tr}_C X_{ABC}),$$

with  $Y_C$  an arbitrary element of  $\mathcal{S}_{=}(\mathcal{H}_C)$ . By construction it is constrained by  $\text{tr}_C \bar{X}_{ABC} = \mathbb{I}_A \otimes \sigma_B$  and also satisfies

$$\text{tr}[\bar{X}_{ABC} \rho_{ABC}] \geq \text{tr}[X_{ABC} \rho_{ABC}].$$

Hence, in (11) it is permissible to take the maximum over the set of all nonnegative operators  $X_{ABC}$  whose partial trace  $\text{tr}_C X_{ABC}$  is bounded by  $\mathbb{I}_A \otimes \sigma_B$  (in spite of being equal to  $\mathbb{I}_A \otimes \sigma_B$ ), that is,

$$2^{H_{\max}(A|B)_{\rho|\sigma}} = \max_{\substack{X_{ABC} \geq 0 \\ \text{tr}_C[X_{ABC}] \leq \mathbb{I}_A \otimes \sigma_B}} \text{tr}[\rho_{ABC} X_{ABC}]. \quad (12)$$

Based on (12) we can express  $2^{H_{\max}(A|B)_{\rho|\sigma}}$  in terms of the following SDP:

	PRIMAL PROBLEM:	DUAL PROBLEM:	
minimum:	$\text{tr}[(\mathbb{I}_A \otimes \sigma_B)Z_{AB}]$	maximum:	$\text{tr}[X_{ABC} \rho_{ABC}]$
subject to:	$Z_{AB} \otimes \mathbb{I}_C \geq \rho_{ABC}$ $Z_{AB} \geq 0.$	subject to:	$\text{tr}_C[X_{ABC}] \leq \mathbb{I}_A \otimes \sigma_B$ $X_{ABC} \geq 0$

where  $Z_{AB}$  is a primal variable and  $X_{ABC}$  a dual variable, respectively. Since the space in the dual problem over which one is optimizing, is closed and bounded, it is compact by the Weierstrass theorem. Hence, the dual optimal plan is finite. Furthermore, the operator  $\bar{Z}_{AB} = 2\|\rho_{ABC}\|_{\infty} \mathbb{I}_{AB} > 0$  satisfies Slater's strict primal feasibility condition  $2\|\rho_{ABC}\|_{\infty} \mathbb{I}_{ABC} - \rho_{ABC} > 0$  and thus the duality gap between the primal and dual optimization problems vanishes. ■

Next, we write out the SDP for  $2^{H_{\max}(A|B)_{\rho}}$  and explore the duality gap between the optimization problems.

**Lemma 8.** Let  $\rho_{AB} \in \mathcal{S}_{\leq}(\mathcal{H}_{AB})$  and let  $\rho_{ABC}$  be a purification of  $\rho_{AB}$  on an auxiliary Hilbert space  $\mathcal{H}_C$ . Then the max-entropy of  $A$  conditioned on  $B$  of  $\rho_{AB}$  is given by

$$H_{\max}(A|B)_{\rho} := \log \min_{Z_{AB}} \|Z_B\|_{\infty}, \quad (13)$$

where the minimum ranges over all  $Z_{AB} \in \mathcal{P}(\mathcal{H}_{AB})$  with  $\rho_{ABC} \leq Z_{AB} \otimes \mathbb{I}_C$ .

*Proof:* The only thing that changes with respect to the SDP in Lemma 7 is that  $\sigma_B$  is no longer fixed but it becomes a dual variable. Thus the SDP for  $2^{H_{\max}(A|B)_{\rho}}$  reads:

	PRIMAL PROBLEM:	DUAL PROBLEM:	
minimum:	$\lambda$	maximum:	$\text{tr}[X_{ABC} \rho_{ABC}]$
subject to:	$Z_{AB} \otimes \mathbb{I}_C \geq \rho_{ABC}$ $\lambda \mathbb{I}_B \geq \text{tr}_A[Z_{AB}]$ $Z_{AB} \geq 0, \lambda \geq 0$	subject to:	$\text{tr}_C[X_{ABC}] \leq \mathbb{I}_A \otimes \sigma_B$ $\text{tr}[\sigma_B] \leq 1$ $X_{ABC} \geq 0, \sigma_B \geq 0$

where  $\lambda$  and  $Z_{AB}$  are primal variables and  $\sigma_B$  and  $X_{ABC}$  dual variables. Obviously, the optimal  $\lambda$  is equal to the largest eigenvalue of  $Z_B$ . Hence, the above program may be rewritten in the form:

	PRIMAL PROBLEM:	DUAL PROBLEM:	
minimum:	$\ Z_B\ _{\infty}$	maximum:	$\text{tr}[X_{ABC} \rho_{ABC}]$
subject to:	$Z_{AB} \otimes \mathbb{I}_C \geq \rho_{ABC}$ $Z_{AB} \geq 0$	subject to:	$\text{tr}_C[X_{ABC}] \leq \mathbb{I}_A \otimes \sigma_B$ $\text{tr}[\sigma_B] \leq 1$ $X_{ABC} \geq 0, \sigma_B \geq 0$

In the dual problem we are optimizing over compact sets, thus there exists a finite dual optimal plan. Furthermore,  $\bar{Z}_{AB} = 2\|\rho_{ABC}\|_{\infty} \mathbb{I}_{AB} > 0$  and  $\bar{\lambda} = 2\|\bar{Z}_B\|_{\infty} > 0$  satisfy Slater's strict primal feasibility condition  $\bar{Z}_{AB} \otimes \mathbb{I}_C > \rho_{ABC}$  and  $\bar{\lambda} \mathbb{I}_B > \text{tr}_A[\bar{Z}_{AB}]$  which implies a zero duality gap. ■

Note that one can always write the operator norm of  $Z_B$  as

$$\|Z_B\|_{\infty} = \max_{\sigma_B} \text{tr}[\sigma_B Z_B] = \max_{\sigma_B} \text{tr}[(\mathbb{I}_A \otimes \sigma_B)Z_{AB}],$$

where the maximum ranges over all  $\sigma_B \in \mathcal{S}_{\leq}(\mathcal{H}_B)$ . Expression (13) then acquires the form

$$H_{\max}(A|B)_{\rho} = \log \min_{\rho_{ABC} \leq Z_{AB} \otimes \mathbb{I}_C} \max_{\sigma_B} \text{tr}[(\mathbb{I}_A \otimes \sigma_B)Z_{AB}]. \quad (14)$$

On the other hand from the vanishing of the duality gap in the SDP of  $H_{\max}(A|B)_{\rho\sigma}$  it follows that

$$\log F(\rho_{AB}, \mathbb{I}_A \otimes \sigma_B)^2 = \log \min_{Z_{AB}} \text{tr}[(\mathbb{I}_A \otimes \sigma_B) Z_{AB}]$$

which after maximization of the left- and the right-hand sides over  $\sigma_B \in \mathcal{S}_{\leq}(\mathcal{H}_B)$  implies

$$H_{\max}(A|B)_{\rho} = \log \max_{\sigma_B} \min_{Z_{AB}} \text{tr}[(\mathbb{I}_A \otimes \sigma_B) Z_{AB}].$$

Therefore, the operations  $\min$  and  $\max$  in (14) commute. Since the function  $\text{tr}[(\mathbb{I}_A \otimes \sigma_B) Z_{AB}]$  is bilinear and the sets over which one optimizes are convex, the commutativity of  $\min$  and  $\max$  can alternatively be seen as a consequence of the minimax theorem.

Henceforth, we will use (3), (10) and (13) and (14) as interchangeable expressions for the conditional max-entropy and the conditional relative max-entropy, respectively.

### B. A Bound on the Relative Conditional Entropy

Here we provide two lemmas which give tight upper bounds of the max- and min-entropy in terms of the relative max- and min-entropy, respectively. The first lemma is a new result whereas the latter one is an improved version of Lemma 21 from [18]. Both of the following statements are important for the derivation of chain rules.

**Lemma 9.** *Let  $\varepsilon > 0$ ,  $\rho_{AB} \in \mathcal{S}_{\leq}(\mathcal{H}_{AB})$  and  $\rho'_{AB} \approx_{\varepsilon'} \rho_{AB}$ . Then there exists a state  $\tilde{\rho}_{AB} \approx_{\varepsilon+\varepsilon'} \rho'_{AB}$  such that*

$$H_{\max}(A|B)_{\tilde{\rho}} \leq H_{\max}(A|B)_{\rho|\rho'} + \log \left( \frac{1}{1 - \sqrt{1 - \varepsilon^2}} \right). \quad (15)$$

*Proof:* Let  $\tilde{Z}_{AB}$  be an optimal primal plan for the semidefinite program for  $H_{\max}(A|B)_{\rho|\rho'}$  and  $\Pi_B$  be the minimum rank projector onto the smallest eigenvalues of the reduced operator  $\tilde{Z}_B$  such that  $\text{tr}[\Pi_B^{\perp} \rho'_B] \leq 1 - \sqrt{1 - \varepsilon^2}$  where  $\Pi_B^{\perp}$  is the orthogonal complement of  $\Pi_B$  and let  $\tilde{\rho}_{AB} := \Pi_B \rho_{AB} \Pi_B$ . By Equation (13), we can write

$$\begin{aligned} 2^{H_{\max}(A|B)_{\tilde{\rho}}} &= \min_{\tilde{\rho}_{ABC} \leq \tilde{Z}_{AB} \otimes \mathbb{I}_C} \|\tilde{Z}_B\|_{\infty} \\ &\leq \|\Pi_B \tilde{Z}_B \Pi_B\|_{\infty}, \end{aligned}$$

where we used the fact that  $\rho_{ABC} \leq \tilde{Z}_{AB} \otimes \mathbb{I}_C$  implies  $\tilde{\rho}_{ABC} \leq \Pi_B \tilde{Z}_{AB} \Pi_B \otimes \mathbb{I}_C$ . Let  $\Pi'_B$  be the projector onto the largest eigenvalue of  $\Pi_B \tilde{Z}_B \Pi_B$ . Then the definition of  $\Pi_B$  implies that

$$\text{tr}[(\Pi_B^{\perp} + \Pi'_B) \rho'_B] \geq 1 - \sqrt{1 - \varepsilon^2}. \quad (16)$$

Moreover, by construction  $\Pi_B^{\perp}$  and  $\Pi'_B$  project onto orthogonal eigenspaces of  $\tilde{Z}_B$ , that is,  $\Pi_B^{\perp} \Pi'_B = 0$ . Hence the sum  $\Pi_B^{\perp} + \Pi'_B$  is itself a projector which commutes with  $\tilde{Z}_B$ . We use the last two facts to find an upper bound for

$$\begin{aligned} \|\Pi_B \tilde{Z}_B \Pi_B\|_{\infty} &= \text{tr}[\Pi'_B \tilde{Z}_B] \\ &= \min_{\mu_B} \frac{\text{tr}[\mu_B \tilde{Z}_B]}{\text{tr}[\mu_B]}, \end{aligned} \quad (17)$$

where the minimization is over all positive operators in the support of  $\Pi_B^{\perp} + \Pi'_B$ . Fixing  $\mu_B = (\Pi_B^{\perp} + \Pi'_B) \rho'_B (\Pi_B^{\perp} + \Pi'_B)$ ,

we obtain the following upper bound for (17):

$$\begin{aligned} \|\Pi_B \tilde{Z}_B \Pi_B\|_{\infty} &\leq \frac{\text{tr}[(\Pi_B^{\perp} + \Pi'_B) \rho'_B (\Pi_B^{\perp} + \Pi'_B) \tilde{Z}_B]}{\text{tr}[(\Pi_B^{\perp} + \Pi'_B) \rho'_B]} \\ &= \frac{\text{tr}[(\Pi_B^{\perp} + \Pi'_B) \tilde{Z}_B^{1/2} \rho'_B \tilde{Z}_B^{1/2}]}{\text{tr}[(\Pi_B^{\perp} + \Pi'_B) \rho'_B]} \\ &\leq \frac{\text{tr}[\rho'_B \tilde{Z}_B]}{\text{tr}[(\Pi_B^{\perp} + \Pi'_B) \rho'_B]} \\ &\leq 2^{H_{\max}(A|B)_{\rho|\rho'}} \frac{1}{1 - \sqrt{1 - \varepsilon^2}}, \end{aligned}$$

where in the last line we used Equation (10) and Inequality (16). Finally, taking the logarithm on both sides yields (15).

The proof is concluded by the upper bound

$$\begin{aligned} P(\tilde{\rho}_{AB}, \rho'_{AB}) &= P(\Pi_B \rho_{AB} \Pi_B, \rho'_{AB}) \\ &\leq P(\Pi_B \rho_{AB} \Pi_B, \Pi_B \rho'_{AB} \Pi_B) + P(\Pi_B \rho'_{AB} \Pi_B, \rho'_{AB}) \\ &\leq P(\rho_{AB}, \rho'_{AB}) + \sqrt{2 \text{tr}[\Pi_B^{\perp} \rho'_{AB}] - (\text{tr}[\Pi_B^{\perp} \rho'_{AB}])^2} \\ &\leq \varepsilon' + \varepsilon \end{aligned}$$

where we use Inequality (37) and the fact that the function  $\sqrt{2t - t^2}$  is monotonously increasing in the interval  $[0, 1]$ . ■

**Lemma 10.** *Let  $\varepsilon > 0$  and  $\rho_{ABC} \in \mathcal{S}_{\leq}(\mathcal{H}_{ABC})$  be pure. Then there exist a projector  $\Pi_{AC}$  on  $\mathcal{H}_{AC}$  and a state  $\tilde{\rho}_{ABC} = \Pi_{AC} \rho_{ABC} \Pi_{AC}$  such that  $\tilde{\rho}_{ABC} \approx_{\varepsilon} \rho_{ABC}$  and*

$$H_{\min}(A|B)_{\rho} \leq H_{\min}(A|B)_{\rho|\tilde{\rho}} + \log \left( \frac{1}{1 - \sqrt{1 - \varepsilon^2}} \right).$$

As already remarked, the proof of this lemma follows exactly the one of Lemma 21 in [18], up to the following modification. Instead of defining the dual projector  $\Pi_B$  of  $\Pi_{AC}$  with regard to the pure state  $\rho_{ABC}$  such that it satisfies  $\text{tr}[\Pi_B^{\perp} \rho_B] \leq \varepsilon^2/2$ , we demand

$$\text{tr}[\Pi_B^{\perp} \rho_B] \leq 1 - \sqrt{1 - \varepsilon^2}.$$

In this way on the one hand the tighter bound (37) yields

$$\begin{aligned} P(\tilde{\rho}_{ABC}, \rho_{ABC}) &\leq \sqrt{2 \text{tr}[\Pi_{AC}^{\perp} \rho_{ABC}] - (\text{tr}[\Pi_{AC}^{\perp} \rho_{ABC}])^2} \\ &= \sqrt{2 \text{tr}[\Pi_B^{\perp} \rho_B] - (\text{tr}[\Pi_B^{\perp} \rho_B])^2} \\ &\leq \varepsilon \end{aligned}$$

which is the same as in Lemma 21 and on the other hand the correction term  $\log(2/\varepsilon^2)$  in Lemma 21 is replaced by the tighter expression  $\log(1/1 - \sqrt{1 - \varepsilon^2})$ .

### C. The $\varepsilon$ -Smooth $S$ -Entropy

For the proof of the chain rules we define an auxiliary entropy measure called  $\varepsilon$ -smooth  $S$ -entropy<sup>2</sup>.

We assume that  $\rho_{AB} \in \mathcal{S}_{\leq}(\mathcal{H}_{AB})$  and  $\sigma_B \in \mathcal{S}_{\leq}(\mathcal{H}_B)$  with  $\text{supp}(\rho_B) \subseteq \text{supp}(\sigma_B)$  and denote for every  $\lambda \in \mathbb{R}$  the projector onto the eigenspace corresponding to the strictly negative eigenvalues of the operator  $2^{\lambda} \rho_{AB} - \sigma_B$  by  $P_{AB}^{\lambda}$ .

<sup>2</sup>The idea for this entropy measure was proposed by Robert König.

**Definition 11.** Let  $\varepsilon > 0$ . Then the  $\varepsilon$ -smooth  $S$ -entropy of  $A$  conditioned on  $B$  of  $\rho_{AB}$  relative to  $\sigma_B$  is defined as

$$S^\varepsilon(A|B)_{\rho|\sigma} := \inf\{\lambda \in \mathbb{R} : \text{tr}[P_{AB}^\lambda \rho_{AB}] \leq \varepsilon\}. \quad (18)$$

Intuitively, this evaluates in a  $\varepsilon$ -smoothed way the smallest  $\lambda$  for which  $\rho_{AB} \geq 2^{-\lambda}\sigma_B$  holds. This should be contrasted with the min-entropy, which evaluates to the largest  $\lambda$  such that  $\rho_{AB} \leq 2^{-\lambda}\sigma_B$ . The  $S$ -entropy is a technical tool only, and our results are expressed in terms of the max-entropy instead. In this spirit, the next lemma gives the upper bound of the  $\varepsilon$ -smooth  $S$ -entropy in terms of the max-entropy.

**Lemma 12.** Let  $\varepsilon > 0$ ,  $\rho_{AB} \in \mathcal{S}_{\leq}(\mathcal{H}_{AB})$  and  $\sigma_B \in \mathcal{S}_{\leq}(\mathcal{H}_B)$ . Then,

$$S^\varepsilon(A|B)_{\rho|\sigma} \leq H_{\max}(A|B)_{\rho|\sigma} + \log\left(\frac{1}{\varepsilon^2}\right). \quad (19)$$

*Proof:* Let  $\lambda_{\inf} \in \mathbb{R}$  be the infimum in Definition 11, that is,  $\lambda_{\inf} = S^\varepsilon(A|B)_{\rho|\sigma}$ , let  $\lambda = \lambda_{\inf} - \delta$  where  $\delta > 0$  and let  $P_{AB}^\pm$  denote the projector onto the nonnegative and strictly negative eigenvalues of  $\rho_{AB} - 2^{-\lambda}\sigma_B$ , respectively. Then, a straightforward computation yields

$$\begin{aligned} 2^{\frac{1}{2}H_{\max}(A|B)_{\rho|\sigma} - \frac{1}{2}S^\varepsilon(A|B)_{\rho|\sigma} + \frac{1}{2}\delta} &= \|\sqrt{\rho_{AB}}\sqrt{\sigma_B}\|_1 2^{-\frac{1}{2}\lambda} \\ &\geq \text{tr}[\sqrt{\rho_{AB}}\sqrt{\sigma_B}] 2^{-\frac{1}{2}\lambda} \\ &= \text{tr}[\sqrt{\rho_{AB}}\sqrt{2^{-\lambda}\sigma_B}] \\ &\geq \text{tr}[P_{AB}^+ 2^{-\lambda}\sigma_B + P_{AB}^- \rho_{AB}] \\ &\geq \text{tr}[P_{AB}^- \rho_{AB}] \\ &\geq \varepsilon. \end{aligned} \quad (20)$$

The first inequality follows from Lemma 9.5 in [10]. In the fourth line we have applied Corollary 18 and in the last line we have used the fact that  $P_{AB}^-$  is identical with the projector  $P_{AB}^\lambda$  and  $\text{tr}[P_{AB}^\lambda \rho_{AB}] \geq \varepsilon$  by definition of  $\lambda$  for any  $\delta > 0$ . Finally, taking the logarithm on both sides of (20) and subsequently taking the limit  $\delta \rightarrow 0$  we obtain (19). ■

#### IV. MAIN RESULTS

This section contains the main result of this paper: a derivation of the previously unknown chain rules for smooth min- and max-entropies. To simplify presentation hereafter, we introduce the function

$$f : \varepsilon \mapsto \log \frac{1}{1 - \sqrt{1 - \varepsilon^2}}$$

that appears as an error term in the chain rules. It vanishes as  $\varepsilon \rightarrow 1$  and grows logarithmically in  $\frac{1}{\varepsilon}$  when  $\varepsilon \rightarrow 0$ .

As remarked in the introduction, the explicit form of one of the chain rules has already been derived in Lemma A.6 in [8]. Following the steps of the original proof and using the improved bound from Lemma 10 we can tighten the chain rule inequality presented in Lemma A.6 of [8] as follows:

**Theorem 13.** Let  $\varepsilon > 0$ ,  $\varepsilon', \varepsilon'' \geq 0$  and  $\rho_{ABC} \in \mathcal{S}_{\leq}(\mathcal{H}_{ABC})$ . Then,

$$H_{\min}^{\varepsilon+\varepsilon'+2\varepsilon''}(AB|C)_\rho \geq H_{\min}^{\varepsilon''}(A|BC)_\rho + H_{\min}^{\varepsilon'}(B|C)_\rho - f(\varepsilon).$$

In the remainder of that section we provide proofs for the remaining three pairs of chain rules. Due to the smooth duality relation (8) it is enough to prove only one of each pair.

**Theorem 14.** Let  $\varepsilon > 0$ ,  $\varepsilon', \varepsilon'' \geq 0$  and  $\rho_{ABC} \in \mathcal{S}_{\leq}(\mathcal{H}_{ABC})$ . Then,

$$H_{\min}^{\varepsilon'}(AB|C)_\rho \leq H_{\min}^{\varepsilon+\varepsilon'+2\varepsilon''}(A|BC)_\rho + H_{\max}^{\varepsilon''}(B|C)_\rho + 2f(\varepsilon). \quad (21)$$

*Proof:* Let  $\rho'_{ABC} \approx_{\varepsilon'} \rho_{ABC}$ ,  $\rho''_{BC} \approx_{\varepsilon''} \rho_{BC}$  be such that

$$\begin{aligned} H_{\min}(AB|C)_{\rho'} &= H_{\min}^{\varepsilon'}(AB|C)_\rho, \quad \text{and} \\ H_{\max}(B|C)_{\rho''} &= H_{\max}^{\varepsilon''}(B|C)_\rho, \end{aligned}$$

and let  $\sigma_C \in \mathcal{S}_{\leq}(\mathcal{H}_C)$  be such that

$$\rho'_{ABC} \leq 2^{-H_{\min}(AB|C)_{\rho'}} \sigma_C = 2^{-H_{\min}^{\varepsilon'}(AB|C)_\rho} \sigma_C. \quad (22)$$

For every  $\delta > 0$  and  $\tilde{\varepsilon} > 0$  there is a  $\delta' \in (0, \delta]$  such that the projector  $P_{BC}^\lambda$  onto the strictly negative eigenvalues of the operator  $2^\lambda \rho''_{BC} - \sigma_C$  with  $\lambda := S^{\tilde{\varepsilon}}(B|C)_{\rho''|\sigma} + \delta'$ , satisfies the constraint  $\text{tr}[P_{BC}^\lambda \rho''_{BC}] \leq \tilde{\varepsilon}$  in Definition 11. If  $P_{BC}^{\lambda\perp}$  is the orthogonal complement of  $P_{BC}^\lambda$ , we have

$$P_{BC}^{\lambda\perp} \sigma_C P_{BC}^{\lambda\perp} \leq 2^\lambda P_{BC}^{\lambda\perp} \rho''_{BC} P_{BC}^{\lambda\perp}. \quad (23)$$

A conjugation of (22) with  $P_{BC}^{\lambda\perp}$  together with (23) yields

$$P_{BC}^{\lambda\perp} \rho'_{ABC} P_{BC}^{\lambda\perp} \leq 2^{-H_{\min}^{\varepsilon'}(AB|C)_\rho + \lambda} P_{BC}^{\lambda\perp} \rho''_{BC} P_{BC}^{\lambda\perp},$$

which is equivalent to

$$H_{\min}(A|BC)_{P^{\lambda\perp} \rho'_{ABC} P^{\lambda\perp}} \geq H_{\min}^{\varepsilon'}(AB|C)_\rho - \lambda.$$

A subsequent optimization of the left-hand side over all  $\mathcal{S}_{\leq}(\mathcal{H}_{BC})$  yields

$$H_{\min}(A|BC)_{P^{\lambda\perp} \rho'_{ABC} P^{\lambda\perp}} \geq H_{\min}^{\varepsilon'}(AB|C)_\rho - \lambda \quad (24)$$

Since  $\rho_{ABC}$  is an extension of  $\rho_{BC}$ , by Corollary 22 there exists an extension  $\rho'_{ABC}$  of  $\rho''_{BC}$  such that  $P(\rho'_{ABC}, \rho_{ABC}) = P(\rho''_{BC}, \rho_{BC})$ . Then the triangle inequality as well as (36) and (37) give us the following upper bound for the purified distance between  $P_{BC}^{\lambda\perp} \rho'_{ABC} P_{BC}^{\lambda\perp}$  and  $\rho_{ABC}$ :

$$\begin{aligned} P(P_{BC}^{\lambda\perp} \rho'_{ABC} P_{BC}^{\lambda\perp}, \rho_{ABC}) &\leq P(P_{BC}^{\lambda\perp} \rho'_{ABC} P_{BC}^{\lambda\perp}, P_{BC}^{\lambda\perp} \rho_{ABC} P_{BC}^{\lambda\perp}) \\ &\quad + P(P_{BC}^{\lambda\perp} \rho_{ABC} P_{BC}^{\lambda\perp}, P_{BC}^{\lambda\perp} \rho''_{ABC} P_{BC}^{\lambda\perp}) \\ &\quad + P(P_{BC}^{\lambda\perp} \rho''_{ABC} P_{BC}^{\lambda\perp}, \rho''_{ABC}) \\ &\quad + P(\rho''_{ABC}, \rho_{ABC}) \\ &\leq \sqrt{2\tilde{\varepsilon} - \tilde{\varepsilon}^2} + \varepsilon' + 2\varepsilon''. \end{aligned}$$

After smoothing the left-hand side of (24) and upper-bounding the term  $S^{\tilde{\varepsilon}}(B|C)_{\rho''|\sigma}$  on the right-hand side of (24) by  $H_{\max}(B|C)_{\rho''|\sigma}$  in accordance with Lemma 12 and subsequently optimizing it over  $\mathcal{S}_{\leq}(\mathcal{H}_C)$ , we obtain

$$\begin{aligned} H_{\min}^{\varepsilon'}(AB|C)_\rho &\leq H_{\min}^{\sqrt{2\tilde{\varepsilon} - \tilde{\varepsilon}^2} + \varepsilon' + 2\varepsilon''}(A|BC)_\rho + H_{\max}^{\varepsilon''}(B|C)_\rho \\ &\quad + \log \frac{1}{\tilde{\varepsilon}^2} + \delta'. \end{aligned}$$

Finally, the substitution  $\tilde{\varepsilon} := 1 - \sqrt{1 - \varepsilon^2}$  leads to the chain rule (21) in the limit  $\delta \rightarrow 0$ . ■

**Theorem 15.** Let  $\varepsilon > 0$ ,  $\varepsilon'$ ,  $\varepsilon'' \geq 0$  and  $\rho_{ABC} \in \mathcal{S}_{\leq}(\mathcal{H}_{ABC})$ . Then,

$$H_{\min}^{\varepsilon'}(AB|C)_{\rho} \leq H_{\max}^{\varepsilon''}(A|BC)_{\rho} + H_{\min}^{2\varepsilon+\varepsilon'+2\varepsilon''}(B|C)_{\rho} + 3f(\varepsilon). \quad (25)$$

*Proof:* Let  $\rho_{ABCD}$  be a purification of  $\rho_{ABC}$ . If

$$H_{\max}^{\varepsilon'}(AB|D)_{\rho} \geq H_{\max}^{2\varepsilon+\varepsilon'+2\varepsilon''}(B|AD)_{\rho} + H_{\min}^{\varepsilon''}(A|D)_{\rho} - 3f(\varepsilon)$$

holds, then the chain rule follows by the duality relation (8).

Let  $\rho'_{ABD} \approx_{\varepsilon'} \rho_{ABD}$ ,  $\rho''_{AD} \approx_{\varepsilon''} \rho_{AD}$  be such that

$$\begin{aligned} H_{\max}(AB|D)_{\rho'} &= H_{\max}^{\varepsilon'}(AB|D)_{\rho}, \quad \text{and} \\ H_{\min}(A|D)_{\rho''} &= H_{\min}^{\varepsilon''}(A|D)_{\rho}, \end{aligned}$$

and let  $\sigma_D \in \mathcal{S}_{\leq}(\mathcal{H}_D)$  be such that

$$\rho''_{AD} \leq 2^{-H_{\min}(A|D)_{\rho''}} \sigma_D = 2^{-H_{\min}^{\varepsilon''}(A|D)_{\rho}} \sigma_D. \quad (26)$$

Again we use the fact that for every  $\delta > 0$  there exists a  $\delta' \in (0, \delta]$  such that for  $\lambda := S^{\varepsilon}(AB|D)_{\rho'}|_{\sigma} + \delta'$ ,  $\tilde{\varepsilon} > 0$ , the projector  $P_{ABD}^{\lambda}$  onto the strictly negative eigenvalues of the operator  $2^{\lambda} \rho'_{ABD} - \sigma_D$  satisfies the constraint  $\text{tr}[P_{ABD}^{\lambda} \rho'_{ABD}] \leq \tilde{\varepsilon}$  in Definition 11. If  $P_{ABD}^{\lambda \perp}$  denotes the orthogonal complement of  $P_{ABD}^{\lambda}$ , then

$$2^{\lambda} P_{ABD}^{\lambda \perp} \rho'_{ABD} P_{ABD}^{\lambda \perp} \geq P_{ABD}^{\lambda \perp} \sigma_D P_{ABD}^{\lambda \perp}. \quad (27)$$

A conjugation of (26) with  $P_{ABD}^{\lambda \perp}$  and a subsequent combination with (27) yields

$$2^{\lambda - H_{\min}^{\varepsilon''}(A|D)_{\rho}} P_{ABD}^{\lambda \perp} \rho'_{ABD} P_{ABD}^{\lambda \perp} \geq P_{ABD}^{\lambda \perp} \rho''_{AD} P_{ABD}^{\lambda \perp}. \quad (28)$$

Consider now the max-entropy

$$\begin{aligned} 2^{H_{\max}(B|AD)_{P^{\lambda \perp} \rho'_{ABD} P^{\lambda \perp} | \rho''}} &= \min_{Z_{ABD} \geq 0} \text{tr}[(\mathbb{I}_B \otimes \rho''_{AD}) Z_{ABD}] \\ &\quad P_{ABD}^{\lambda \perp} \rho'_{ABD} P_{ABD}^{\lambda \perp} \leq Z_{ABD} \otimes \mathbb{I}_C \end{aligned} \quad (29)$$

where  $\rho'_{ABCD}$  is a purification of  $\rho'_{ABD}$ . Making use of (28) and the inequality

$$P_{ABD}^{\lambda \perp} \rho'_{ABD} P_{ABD}^{\lambda \perp} \leq P_{ABD}^{\lambda \perp} \otimes \mathbb{I}_C$$

and omitting the identity operator, we can upper-bound the right-hand side of (29) in the following way:

$$\begin{aligned} &\leq \text{tr}[\rho''_{AD} P_{ABD}^{\lambda \perp}] \\ &\leq 2^{\lambda - H_{\min}^{\varepsilon''}(A|D)_{\rho}} \text{tr}[P_{ABD}^{\lambda \perp} \rho'_{ABD} P_{ABD}^{\lambda \perp}] \\ &\leq 2^{\lambda - H_{\min}^{\varepsilon''}(A|D)_{\rho}}, \end{aligned}$$

where we use that the term  $\text{tr}[P_{ABD}^{\lambda \perp} \rho'_{ABD} P_{ABD}^{\lambda \perp}]$  is upper bounded by one. Taking the logarithm and substituting  $\lambda$  yields

$$H_{\max}(B|AD)_{P^{\lambda \perp} \rho'_{ABD} P^{\lambda \perp} | \rho''} \leq S^{\varepsilon}(AB|D)_{\rho'}|_{\sigma} + \delta' - H_{\min}^{\varepsilon''}(A|D)_{\rho}.$$

A subsequent application of Lemma 12 implies

$$\begin{aligned} H_{\max}(B|AD)_{P^{\lambda \perp} \rho'_{ABD} P^{\lambda \perp} | \rho''} &\leq H_{\max}(AB|D)_{\rho'} - H_{\min}^{\varepsilon''}(A|D)_{\rho} \\ &\quad + \delta' + \log \frac{1}{\tilde{\varepsilon}^2}, \end{aligned} \quad (30)$$

where the max-entropy term on the right-hand side has been optimized on  $\mathcal{S}_{\leq}(\mathcal{H}_D)$ . Consider now the left-hand side of (30). Corollary 22 guarantees the existence of an extension

$\rho''_{ABD}$  such that  $P(\rho''_{AD}, \rho_{AD}) = P(\rho''_{ABD}, \rho_{ABD})$ . Then, it follows that

$$\begin{aligned} P(P_{ABD}^{\lambda \perp} \rho'_{ABD} P_{ABD}^{\lambda \perp}, \rho''_{ABD}) &\leq P(P_{ABD}^{\lambda \perp} \rho'_{ABD} P_{ABD}^{\lambda \perp}, \rho'_{ABD}) \\ &\quad + P(\rho'_{ABD}, \rho''_{ABD}) \\ &\leq \sqrt{2\tilde{\varepsilon} - \tilde{\varepsilon}^2} + \varepsilon' + \varepsilon''. \end{aligned}$$

Thus, according to Lemma 9, there exists a state  $\tilde{\rho}_{ABD} \approx_{\varepsilon + \sqrt{2\tilde{\varepsilon} - \tilde{\varepsilon}^2} + \varepsilon' + 2\varepsilon''} \rho_{ABD}$  such that

$$\begin{aligned} H_{\max}(B|AD)_{\tilde{\rho}} &\leq H_{\max}^{\varepsilon'}(AB|D)_{\rho} - H_{\min}^{\varepsilon''}(A|D)_{\rho} \\ &\quad + \delta' + \log \frac{1}{\tilde{\varepsilon}^2} + f(\varepsilon). \end{aligned}$$

Smoothing of the left-hand side and regrouping the terms in the last inequality yields

$$\begin{aligned} H_{\max}^{\varepsilon'}(AB|D)_{\rho} &\geq H_{\max}^{\varepsilon + \sqrt{2\tilde{\varepsilon} - \tilde{\varepsilon}^2} + \varepsilon' + 2\varepsilon''}(B|AD)_{\rho} + H_{\min}^{\varepsilon''}(A|D)_{\rho} \\ &\quad - \delta' - \log \frac{1}{\tilde{\varepsilon}^2} - f(\varepsilon). \end{aligned}$$

Finally, setting  $\tilde{\varepsilon} := 1 - \sqrt{1 - \varepsilon^2}$ , taking the limit  $\delta \rightarrow 0$ , and applying the duality relation for smooth entropies (8), we obtain chain rule (25).  $\blacksquare$

The last chain rule follows from chain rule (21) together with Lemma 5.

**Corollary 16.** Let  $\varepsilon'$ ,  $\varepsilon''$ ,  $\varepsilon''' \geq 0$  and  $\rho_{ABC} \in \mathcal{S}_{\leq}(\mathcal{H}_{ABC})$  such that  $\varepsilon' + 2\varepsilon'' + \varepsilon''' < 1 - 2\sqrt{1 - \text{tr} \rho}$ . Then,

$$\begin{aligned} H_{\min}^{\varepsilon'}(AB|C)_{\rho} &\leq H_{\max}^{\varepsilon''}(A|BC)_{\rho} + H_{\max}^{\varepsilon''}(B|C)_{\rho} \\ &\quad + g(\varepsilon', \varepsilon'', \varepsilon''', \text{tr} \rho), \end{aligned} \quad (31)$$

where  $g(\varepsilon', \varepsilon'', \varepsilon''', \text{tr} \rho) :=$

$$\inf_{\varepsilon} \left\{ 2f(\varepsilon) + \log \left( \frac{1}{1 - (\varepsilon + \varepsilon' + 2\varepsilon'' + \varepsilon''' + 2\sqrt{1 - \text{tr} \rho})^2} \right) \right\},$$

and the infimum is taken in the range  $0 < \varepsilon < 1 - \varepsilon' - 2\varepsilon'' - \varepsilon''' - 2\sqrt{1 - \text{tr} \rho}$ .

*Proof:* Let  $\varepsilon > 0$  be any smoothing parameter such that  $\varepsilon < 1 - \varepsilon' - 2\varepsilon'' - \varepsilon''' - 2\sqrt{1 - \text{tr} \rho}$ . Then, by Lemma 5, the smooth min-entropy term on the right-hand side of (21) is upper bounded by

$$H_{\max}^{\varepsilon''}(A|BC)_{\rho} + \log \left( \frac{1}{1 - (\varepsilon + \varepsilon' + 2\varepsilon'' + \varepsilon''' + 2\sqrt{1 - \text{tr} \rho})^2} \right)$$

which immediately gives (31).  $\blacksquare$

In contrast to the previous chain rules, the last one leads to non-trivial results even if we apply it to non-smooth entropies. For example, for a normalized state  $\rho_{ABC}$ , we find

$$H_{\min}(AB|C)_{\rho} \leq H_{\max}(A|BC)_{\rho} + H_{\max}(B|C)_{\rho} + 4.$$

## V. CONCLUSION

We derived four pairs of chain rules for the smooth entropy, and every combination of min- and max-entropies is considered. Counter-examples suggest that the inequalities cannot be reversed, and thus that this list is complete. In particular, we do not expect a chain rule of the form

$$H_{\min}^{\varepsilon}(AB|C) \leq H_{\min}^{\varepsilon'}(A|BC) + H_{\min}^{\varepsilon''}(B|C) + h, \quad (32)$$



for small smoothing parameters  $\varepsilon, \varepsilon'$  and  $\varepsilon''$  and error term  $h(\varepsilon, \varepsilon', \varepsilon'')$  due to the following counter-example. Let us consider the state  $\rho_{ABCC'} = \frac{1}{2} \sum_{i \in \{0,1\}} \rho_{ABC}^i \otimes |i\rangle\langle i|_{C'}$  with

$$\rho_{ABC}^0 = |\phi\rangle\langle\phi|_{AB} \otimes \pi_C \quad \text{and} \quad \rho_{ABC}^1 = \pi_A \otimes |\phi\rangle\langle\phi|_{BC},$$

where  $|\phi\rangle$  is a maximally entangled state,  $\pi$  is a fully mixed state, we take  $A, B$  and  $C$  to be  $d$ -dimensional quantum systems and  $C'$  is an auxiliary register with basis  $\{|0\rangle, |1\rangle\}$ . Any min-entropy conditioned on the classical register  $C'$  can be expressed as [15]

$$H_{\min}(\cdot|C')_\rho = -\log \frac{\sum_{i=0}^1 2^{-H_{\min}(\cdot|C')_{\rho^i}}}{2} \approx \min_i H_{\min}(\cdot|C')_{\rho^i},$$

where we approximate up to  $\pm 1$ . Thus,  $H_{\min}(AB|CC') = 0$  and  $H_{\min}(A|BCC') = H_{\min}(B|CC') \approx -\log d$  and it is easy to verify that (32) is violated for moderate smoothing  $\varepsilon', \varepsilon'' < \frac{1}{2}$  and  $d$  such that  $\log d \gg h$ .

### Acknowledgments

This work was supported by the Swiss National Science Foundation (SNF) through the National Centre of Competence in Research Quantum Science and Technology and project No. 200020-135048, and by the European Research Council (ERC) via grant No. 258932. MT acknowledges support from the National Research Foundation (Singapore), and the Ministry of Education (Singapore).

### APPENDIX A PROOF OF LEMMA 5

In the following we restate Lemma 5 and prove it using the derived SDPs for the non-smooth max-entropy.

**Restatement of Lemma 5.** *Let  $\varepsilon, \varepsilon' \geq 0$  such that  $\varepsilon + \varepsilon' + 2\sqrt{1 - \text{tr} \rho_{AB}} < 1$  and let  $\rho_{AB} \in \mathcal{S}_{\leq}(\mathcal{H}_{AB})$ . Then,*

$$H_{\min}^{\varepsilon'}(A|B)_\rho \leq H_{\max}^\varepsilon(A|B)_\rho + \log \left( \frac{1}{1 - (\varepsilon + \varepsilon' + 2\sqrt{1 - \text{tr} \rho})^2} \right).$$

*Proof.* Define  $\hat{\rho}_{AB} = \rho_{AB} / \text{tr}(\rho_{AB})$ . According to Lemma 5.2 in [15] there are embeddings  $U : \mathcal{H}_A \rightarrow \mathcal{H}_{A'}$  and  $V : \mathcal{H}_B \rightarrow \mathcal{H}_{B'}$  such that there exists a normalized state  $\bar{\rho}_{A'B'} \approx_\varepsilon \hat{\rho}_{A'B'}$ , where  $\hat{\rho}_{A'B'} = (U \otimes V) \hat{\rho}_{AB} (U^\dagger \otimes V^\dagger)$ , which minimizes the smooth max-entropy  $H_{\max}^{\tilde{\varepsilon}}(A'|B')_{\hat{\rho}} = H_{\max}^{\tilde{\varepsilon}}(A|B)_{\hat{\rho}}$ .

Consider now the quantity  $2^{-H_{\min}^{\tilde{\varepsilon}+\tilde{\varepsilon}'}(A'|B')_{\hat{\rho}}}$ . We are simultaneously minimizing over all  $\sigma_{B'} \in \mathcal{S}_{\leq}(\mathcal{H}_{B'})$  and all states  $\tilde{\rho}_{A'B'}$ , that are  $\tilde{\varepsilon} + \tilde{\varepsilon}'$ -close to the normalized state  $\bar{\rho}_{A'B'}$ . By Uhlmann's theorem the latter constraint translates into  $\text{tr}[\tilde{\rho}_{A'B'} \bar{\rho}_{A'B'}] \geq 1 - (\tilde{\varepsilon} + \tilde{\varepsilon}')^2$  where  $\mathcal{H}_C$  is a purifying system. We can formulate  $2^{-H_{\min}^{\tilde{\varepsilon}+\tilde{\varepsilon}'}(A'|B')_{\hat{\rho}}}$  as the following semidefinite program:

$$\begin{aligned} & \text{PRIMAL PROBLEM:} \\ \text{minimum:} & \quad \text{tr}[\mathbb{I}_{B'} \sigma_{B'}] \\ \text{subject to:} & \quad \mathbb{I}_{A'} \otimes \sigma_{B'} \geq \text{tr}_C[\bar{\rho}_{A'B'C}] \\ & \quad \text{tr}[\tilde{\rho}_{A'B'} \bar{\rho}_{A'B'C}] \geq 1 - (\tilde{\varepsilon} + \tilde{\varepsilon}')^2 \\ & \quad \text{tr}[\tilde{\rho}_{A'B'}] \leq 1 \\ & \quad \sigma_{B'} \geq 0, \tilde{\rho}_{A'B'} \geq 0 \end{aligned}$$

$$\begin{aligned} & \text{DUAL PROBLEM:} \\ \text{maximum:} & \quad (1 - (\tilde{\varepsilon} + \tilde{\varepsilon}')^2) \lambda - \mu \\ \text{subject to:} & \quad \text{tr}_A[E_{A'B'}] \leq \mathbb{I}_B \\ & \quad \lambda \bar{\rho}_{A'B'} \leq E_{A'B'} \otimes \mathbb{I}_C + \mu \mathbb{I}_{A'B'C} \\ & \quad E_{A'B'} \geq 0, \lambda, \mu \geq 0, \end{aligned}$$

where  $\sigma_{B'}$  and  $\tilde{\rho}_{A'B'C}$  are the primal variables and  $E_{A'B'}$ ,  $\lambda$  and  $\mu$  are the dual variables, respectively. Let  $Z_{A'B'}$  be a primal optimal plan for the semidefinite program of  $H_{\max}(A'|B')_{\hat{\rho}}$ , that is  $Z_{A'B'} \otimes \mathbb{I}_C \geq \bar{\rho}_{A'B'C}$  and  $\text{tr}_{A'}[Z_{A'B'}] \leq 2^{H_{\max}(A'|B')_{\hat{\rho}}} \mathbb{I}_{B'}$ . Then the variables  $E_{A'B'} = 2^{-H_{\max}(A'|B')_{\hat{\rho}}} Z_{A'B'}$ ,  $\lambda = 2^{-H_{\max}(A'|B')_{\hat{\rho}}}$  and  $\mu = 0$  are a dual feasible plan for the above semidefinite program. By the weak duality theorem we have then

$$(1 - (\tilde{\varepsilon} + \tilde{\varepsilon}')^2) 2^{-H_{\max}(A'|B')_{\hat{\rho}}} \leq 2^{-H_{\min}^{\tilde{\varepsilon}+\tilde{\varepsilon}'}(A'|B')_{\hat{\rho}}}.$$

Taking the logarithm and considering the fact that all states which are  $\tilde{\varepsilon}'$ -close to  $\hat{\rho}_{A'B'}$  are contained in the  $(\tilde{\varepsilon} + \tilde{\varepsilon}')$ -neighborhood of  $\bar{\rho}_{A'B'}$ , we get

$$H_{\min}^{\tilde{\varepsilon}'}(A'|B')_{\hat{\rho}} \leq H_{\min}^{\tilde{\varepsilon}+\tilde{\varepsilon}'}(A'|B')_{\hat{\rho}} \leq H_{\max}^{\tilde{\varepsilon}}(A'|B')_{\hat{\rho}} + \log \left( \frac{1}{1 - (\tilde{\varepsilon} + \tilde{\varepsilon}')^2} \right). \quad (33)$$

By Proposition 5.3 in [15] we have  $H_{\min}^{\tilde{\varepsilon}'}(A'|B')_{\hat{\rho}} = H_{\min}^{\tilde{\varepsilon}'}(A|B)_{\hat{\rho}}$  and  $H_{\max}^{\tilde{\varepsilon}}(A'|B')_{\hat{\rho}} = H_{\max}^{\tilde{\varepsilon}}(A|B)_{\hat{\rho}}$ . Finally, substituting in (33)  $\tilde{\varepsilon} = \varepsilon + \sqrt{1 - \text{tr}(\rho_{AB})}$  and  $\tilde{\varepsilon}' = \varepsilon' + \sqrt{1 - \text{tr}(\rho_{AB})}$  and considering that  $H_{\min}^\varepsilon(A|B)_\rho \leq H_{\min}^{\varepsilon'+\sqrt{1-\text{tr}\rho}}(A|B)_\rho$  as well as  $H_{\max}^{\varepsilon+\sqrt{1-\text{tr}\rho}}(A|B)_\rho \leq H_{\max}^\varepsilon(A|B)_\rho$  we conclude the proof. ■

### APPENDIX B TECHNICAL LEMMAS

#### A. Operator inequalities

**Theorem 17** ([1], Theorem 1). *Let  $Q$  and  $R$  be positive semidefinite operators on a Hilbert space  $\mathcal{H}$  and let  $0 \leq s \leq 1$ . Then,*

$$\text{tr}[Q^s R^{1-s}] \geq \frac{1}{2} \text{tr}[Q + R - |Q - R|] \quad (34)$$

From this theorem we can draw the following useful corollary.

**Corollary 18.** *Let  $R$  and  $Q$  be positive semidefinite operators on a Hilbert space  $\mathcal{H}$ , let  $0 \leq s \leq 1$  and let  $P_+$  and  $P_-$  denote the orthogonal projectors onto the eigenspaces corresponding to nonnegative and strictly negative eigenvalues of the operator  $Q - R$ , respectively. Then,*

$$\text{tr}[Q^s R^{1-s}] \geq \text{tr}[P_+ R + P_- Q]$$

*Proof.* We make the following decomposition of  $|Q - R|$

$$|Q - R| = P_+ (Q - R) P_+ - P_- (Q - R) P_-, \quad (35)$$

where  $P_{\pm}$  are the projectors onto the nonnegative and strictly negative eigenvalues of  $Q - R$ , respectively. Substituting (35) in (34) and using the fact that  $P_+ + P_- = \mathbb{I}$ , we obtain

$$\begin{aligned} \operatorname{tr} [Q^s R^{1-s}] &\geq \frac{1}{2} \operatorname{tr} [Q + R - |Q - R|] \\ &= \operatorname{tr} [P_- Q + (\mathbb{I} - P_-) R] \\ &= \operatorname{tr} [P_- Q + P_+ R]. \end{aligned}$$

■

### B. Purified Distance: Properties

**Lemma 19** ([17], Lemma 7). *If  $\rho, \sigma \in \mathcal{S}_{\leq}(\mathcal{H})$  and  $\mathcal{E}$  is a trace non-increasing CPM on  $\mathcal{L}(\mathcal{H})$ , then*

$$P(\mathcal{E}(\rho), \mathcal{E}(\sigma)) \leq P(\rho, \sigma).$$

Evidently, for any  $0 \leq \Pi \leq 1$  the map defined by  $\rho \mapsto \Pi\rho\Pi$ ,  $\rho \in \mathcal{S}_{\leq}(\mathcal{H})$  is a trace non-increasing CPM. Thus, in particular, by the above lemma we have

$$P(\Pi\rho\Pi, \Pi\sigma\Pi) \leq P(\rho, \sigma) \quad (36)$$

for  $\rho, \sigma \in \mathcal{S}_{\leq}(\mathcal{H})$ .

**Lemma 20** ([4], Lemma 7). *Let  $\rho \in \mathcal{S}_{\leq}(\mathcal{H})$  and  $0 \leq \Pi \leq \mathbb{I}$ . Then,*

$$P(\Pi\rho\Pi, \rho) \leq \frac{1}{\sqrt{\operatorname{tr} \rho}} \sqrt{(\operatorname{tr} \rho)^2 - (\operatorname{tr} [\Pi^2 \rho])^2}.$$

When  $\Pi$  is a projector, that is  $\Pi^2 = \Pi$ , then a straightforward computation yields

$$P(\Pi\rho\Pi, \rho) \leq \sqrt{2 \operatorname{tr} [\Pi^{\perp} \rho] - (\operatorname{tr} [\Pi^{\perp} \rho])^2} \quad (37)$$

where  $\Pi^{\perp} = \mathbb{I} - \Pi$  is the orthogonal complement of  $\Pi$ .

**Lemma 21** ([17], Lemma 8). *Let  $\rho, \sigma \in \mathcal{S}_{\leq}(\mathcal{H})$ ,  $\mathcal{H}' \cong \mathcal{H}$  and  $\bar{\rho} \in \mathcal{S}_{\leq}(\mathcal{H} \otimes \mathcal{H}')$  be a purification of  $\rho$ . Then, there exists a purification  $\bar{\sigma} \in \mathcal{S}_{\leq}(\mathcal{H} \otimes \mathcal{H}')$  of  $\sigma$  such that  $P(\bar{\rho}, \bar{\sigma}) = P(\rho, \sigma)$ .*

From that lemma one infers the following corollary:

**Corollary 22.** *Let  $\rho, \sigma \in \mathcal{S}_{\leq}(\mathcal{H})$ ,  $\mathcal{H}' \cong \mathcal{H}$  and  $\bar{\rho} \in \mathcal{S}_{\leq}(\mathcal{H} \otimes \mathcal{H}')$  be an extension of  $\rho$ . Then, there exists an extension  $\bar{\sigma} \in \mathcal{S}_{\leq}(\mathcal{H} \otimes \mathcal{H}')$  of  $\sigma$  such that  $P(\bar{\rho}, \bar{\sigma}) = P(\rho, \sigma)$ .*

### REFERENCES

- [1] K. Audenaert, J. Calsamiglia, R. Muñoz-Tapia, E. Bagen, L. Masanes, A. Acín, and F. Verstraete. Discriminating states: The quantum Chernoff bound. *Physical Letters Review*, 98:160501–4, 2007.
- [2] A. Barvinok. *A Course in Convexity*, volume 54 of *Graduate Studies in Mathematics*. American Mathematical Society, 2002.
- [3] M. Berta. Single-shot quantum state merging. Master's thesis, ETH Zurich, 2008. [arXiv: 0912.4495](https://arxiv.org/abs/0912.4495).
- [4] M. Berta, M. Christandl, R. Colbeck, J. Renner, and R. Renner. The uncertainty principle in the presence of quantum memory. *Nature Physics*, 1734, 2010.
- [5] N. Datta. Min- and max-relative entropies and a new entanglement monotone. *IEEE Transactions on Information Theory*, 55(6):2816, 2009.
- [6] L. del Rio, J. Aberg, R. Renner, O. C. O. Dahlsten, and V. Vedral. The thermodynamic meaning of negative entropy. *Nature*, 474(7349):61–63, 2011.
- [7] F. Dupuis. *The Decoupling Approach to Quantum Information Theory*. PhD thesis, Université de Montréal, Apr. 2009. [arXiv: 1004.1641](https://arxiv.org/abs/1004.1641).
- [8] F. Dupuis, M. Berta, J. Wullschlegler, and R. Renner. The decoupling theorem. Dec. 2010. [arXiv: 1012.6044](https://arxiv.org/abs/1012.6044).
- [9] R. König, R. Renner, and C. Schaffner. The operational meaning of min- and max-entropy. *IEEE Transactions on Information Theory*, 55(9):4674–4681, 2009.
- [10] M. A. Nielsen and I. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2000.
- [11] R. Renner. *Security of Quantum Key Distribution*. PhD thesis, ETH Zürich, 2005. Available online: <http://arxiv.org/abs/quant-ph/0512258v2>.
- [12] R. Renner and R. König. Universally composable privacy amplification against quantum adversaries. *Springer Lecture Notes in Computer Science*, 3378(9):407–425, 2005.
- [13] R. Renner and S. Wolf. Smooth Rényi entropy and applications. *Proc. IEEE Int. Symp. Info. Theory*, page 233, 2004.
- [14] C. Shannon. A mathematical theory of communication. *Bell Syst. Tech. J.*, 27:379–423, 1948.
- [15] M. Tomamichel. *A Framework for Non-Asymptotic Quantum Information Theory*. PhD thesis, ETH Zürich, 2012. Available online: [http://arxiv.org/abs/1203.2142](https://arxiv.org/abs/1203.2142).
- [16] M. Tomamichel, R. Colbeck, and R. Renner. A fully quantum asymptotic equipartition property. *IEEE Transactions on Information Theory*, 55:5840–5847, 2009.
- [17] M. Tomamichel, R. Colbeck, and R. Renner. Duality between smooth min- and max-entropies. *IEEE Transactions on Information Theory*, 56:4674–4681, 2010.
- [18] M. Tomamichel, R. Renner, C. Schaffner, and A. Smith. Leftover hashing against quantum side information. *Proc. IEEE Int. Symp. Info. Theory*, pages 2703–2707, 2010.
- [19] A. Uhlmann. The transition probability in the state space of a \*-algebra. *Rep. Math. Phys.*, 9(273), 1976.
- [20] J. Watrous. Theory of quantum information, Fall 2011. Available online: <http://www.cs.uwaterloo.ca/~watrous/CS766/>. Lecture notes.
- [21] S. Winkler, M. Tomamichel, S. Hengli, and R. Renner. Impossibility of growing quantum bit commitments. *Phys. Rev. Lett.*, 107:090502, 2011.