Second-Order Characterizations via Partial Smoothing

Anurag Anshu*, Mario Berta†, Rahul Jain‡, Marco Tomamichel§
* Institute for Quantum Computing, University of Waterloo and Perimeter Institute for Theoretical Physics, Waterloo, Canada
‡ Department of Computing, Imperial College London, London, United Kingdom
‡ Centre for Quantum Technologies and Department of Computer Science, National University of Singapore, Singapore 117543; and MajuLab, CNRS-UNS-NUS-NTU International Joint Research Unit, UMI 3654, Singapore
§ Centre for Quantum Software and Information, University of Technology Sydney, Sydney, Australia

***Abstract*—Smooth entropies are a tool for quantifying resource trade-offs in information theory and cryptography. However, in typical multi-partite problems some of the sub-systems are often left unchanged and this is not reflected by the standard smoothing of information measures over a ball of close states. We propose to smooth instead only over a ball of close states which also have some of the reduced states on the relevant sub-systems fixed. This partial smoothing of information measures naturally allows to give more refined characterizations of various information-theoretic problems in the one-shot setting. As a consequence, we can derive asymptotic second-order characterizations for tasks such as privacy amplification against classical side information or classical state splitting. For quantum problems like state merging the general resource trade-off is tightly characterized by partially smoothed information measures as well.**

I. Introduction

One-shot information theory concerns itself with finding tight bounds on the resource trade-offs for various operational problems in information theory and cryptography (see, e.g., [1] for an introduction). Smooth entropies and smooth mutual informations have in many cases proven to be adequate information measures in this context. On the one hand, smooth min-entropy was first introduced in the context of quantum cryptography [2]. More precisely, the smooth conditional min-entropy was introduced to characterize the amount of uniform and independent randomness that can be extracted from a correlated random variable. On the other hand, the smooth max-information has been introduced to quantify the communication requirements in quantum extensions of Slepian-Wolf coding [3]. Since then smooth entropy measures of various kinds have been used to characterize a plethora of other tasks as well. Our main contributions can be summarized as follows:

- We introduce a notion of partially smoothed mutual max-information and conditional min-entropy and establish some of their basic mathematical properties.
- We show that these new definitions are equivalent to their fully smoothed counterparts, up to terms that vanish in the first-order i.i.d. asymptotics. Moreover, for the fully classical case this equivalence even holds for the asymptotic second-order i.i.d. asymptotics.
- We give several examples of operational problems where the new quantities naturally appear to give tighter bounds for the one-shot problem. In particular, for classical problems this leads to asymptotic second-order i.i.d. expansions.

II. Preliminaries

Before proceeding to our results, we introduce some basic one-shot information theoretic quantities in the quantum domain. For brevity of presentation, we define the following notation for set on operators on a register A. The set of positive semi-definite (psd) operators acting on A is defined as $\mathcal{P}(A)$. We also define two subsets: the set of quantum states (i.e. psd operators with unit trace), denoted $\mathcal{S}_\circ(A)$, and sub-normalized states (i.e. psd operators with trace not exceeding unity), denoted $\mathcal{S}_\bullet(A)$.

Let ρ and σ be two psd operators. If $\rho \ll \sigma$ we define the max-divergence [4], [5] as

$$D_{\max}(\rho\|\sigma) := \inf\left\{\lambda : \rho \leq \exp(\lambda)\sigma\right\},$$

and otherwise it is defined as $+\infty$. Above, $\geq$ denotes the Löwner partial order of operators in $\mathcal{P}(A)$. Max-divergence can be used to define various notions of mutual max-information and conditional min-entropy, respectively. We will concern ourselves with the following two definitions [2], [3]. For any bipartite state $\rho_{AB} \in \mathcal{S}_\bullet(AB)$, we have

$$I_{\max}(A;B)_\rho := \inf_{\sigma_B \in \mathcal{S}_\bullet(B)} D_{\max}(\rho_{AB}\|\rho_A \otimes \sigma_B),$$
$$H_{\min}(A|B)_\rho := -D_{\max}(\rho_{AB}\|1_A \otimes \rho_B).$$

Our goal is to define a smooth max-information and smooth min-entropy based on the above quantities, i.e. quantities for which ρ_{AB} is replaced with a ball of states close to ρ_{AB}. In particular, we want the states in this ball to have the property that the A subsystem is (essentially) left intact. To do this we need to use a metric $\Delta(\cdot,\cdot)$ on (sub-normalized) quantum states, i.e. positive semi-definite operators with trace not exceeding one. In this work, we mainly consider the purified distance and the trace distance. The purified distance [6] based on the generalized fidelity is given as

$$P(\rho,\tau) := \sqrt{1-F^2(\rho,\tau)} \quad \text{with}$$
$$F(\rho,\tau) := \operatorname{tr}\left[\left|\sqrt{\rho}\sqrt{\tau}\right|\right] + \sqrt{(1-\operatorname{tr}[\rho])(1-\operatorname{tr}[\tau])}$$

and the generalized trace distance [1] is given as

$$T(\rho,\tau) := \frac{1}{2}\operatorname{tr}\left[\left|\rho-\tau\right|\right] + \frac{1}{2}\left|\operatorname{tr}[\rho]-\operatorname{tr}[\tau]\right|.$$

Above definitions take into account that ρ,τ may be sub-normalized. The standard measures, which are not a metric for sub-normalized states, are

$$\bar{F}(\rho,\tau) := \operatorname{tr}\left[\left|\sqrt{\rho}\sqrt{\tau}\right|\right] \quad \text{and} \quad \|\rho-\tau\|_1 := \operatorname{tr}\left[\left|\rho-\tau\right|\right].$$

The following two definitions are rather standard:

Definition 1. *Let $\rho_{AB} \in \mathcal{S}_\circ(AB)$ with valid (ε,Δ).*[1] *The (ε,Δ)-smooth max-information of A and B is defined as*[2]

$$\begin{aligned} I_{\max}^{\varepsilon,\Delta}(A;B)_\rho \;:=\; \inf \quad & D_{\max}(\tilde{\rho}_{AB}\|\rho_A\otimes\sigma_B) \\ \text{s.t.} \quad & \tilde{\rho}_{AB}\in\mathcal{S}_\circ(AB), \\ & \Delta(\tilde{\rho}_{AB},\rho_{AB})\leq\varepsilon, \\ & \sigma_B\in\mathcal{S}_\circ(B). \end{aligned}$$

Moreover, the (ε,Δ)-smooth conditional min-entropy of A given B is defined as

$$\begin{aligned} H_{\min}^{\varepsilon,\Delta}(A|B)_\rho \;:=\; \sup \quad & -D_{\max}(\tilde{\rho}_{AB}\|1_A\otimes\rho_B) \\ \text{s.t.} \quad & \tilde{\rho}_{AB}\in\mathcal{S}_\bullet(AB), \\ & \Delta(\tilde{\rho}_{AB},\rho_{AB})\leq\varepsilon. \end{aligned}$$

As for the non-smooth case there are other definitions in use that we will not discuss here specifically, e.g. in the definition of the smooth max-information one can fix σ_B to be ρ_B to arrive at a different quantity, and similarly the min-entropy can be further optimized over $\sigma_B\in\mathcal{S}_\bullet(B)$. Note that for the smooth min-entropy it is necessary to smooth over sub-normalized states as otherwise the quantity will not be invariant under the application of local embedding maps [1, Sec. 6.2.3].

III. Partially smoothed information measures

We now propose the following new definition for the smooth max-information:

Definition 2. *Let $\rho_{AB} \in \mathcal{S}_\circ(AB)$ with valid (ε,Δ). The (ε,Δ)-*smooth max-information with fixed A of ρ_{AB} *is defined as*

$$\begin{aligned} I_{\max}^{\varepsilon,\Delta}(\dot{A};B)_\rho \;:=\; \inf \quad & D_{\max}(\tilde{\rho}_{AB}\|\rho_A\otimes\sigma_B) \\ \text{s.t.} \quad & \tilde{\rho}_{AB}\in\mathcal{S}_\circ(AB), \\ & \Delta(\tilde{\rho}_{AB},\rho_{AB})\leq\varepsilon, \\ & \tilde{\rho}_A=\rho_A, \\ & \sigma_B\in\mathcal{S}_\circ(B). \end{aligned}$$

We also suggest the following new definition of smooth conditional min-entropy:

Definition 3. *Let $\rho_{AB} \in \mathcal{S}_\circ(AB)$ with valid (ε,Δ). Then, the (ε,Δ)-*smooth min-entropy with fixed B of ρ_{AB} *is defined as*

$$\begin{aligned} H_{\min}^{\varepsilon,\Delta}(A|\dot{B})_\rho \;:=\; \sup \quad & -D_{\max}(\tilde{\rho}_{AB}\|1_A\otimes\rho_B) \\ \text{s.t.} \quad & \tilde{\rho}_{AB}\in\mathcal{S}_\bullet(AB), \\ & \Delta(\tilde{\rho}_{AB},\rho_{AB})\leq\varepsilon, \\ & \tilde{\rho}_B\leq\rho_B, . \end{aligned}$$

If the input states are classical in a fixed basis all the definitions apply for this case as well. It is then immediate to see that the respective optimizations over $\tilde{\rho}_{AB}$ and σ_B can without loss of generality be restricted to be diagonal in this fixed basis as well.[3] An important property of our definitions is that they are monotonic under quantum operations.

Lemma 1. *Let $\rho_{AB} \in \mathcal{S}_\circ(AB)$ with valid (ε,Δ). For any two completely positive trace preserving maps $\mathcal{E}:\mathcal{P}(A)\to\mathcal{P}(A')$ and $\mathcal{F}:\mathcal{P}(B)\to\mathcal{P}(B')$, we have*

$$I_{\max}^{\varepsilon}(\dot{A};B)_\rho \geq I_{\max}^{\varepsilon}(\dot{A}';B')_\tau$$

where $\tau_{A'B'}=(\mathcal{E}\otimes\mathcal{F})(\rho_{AB})$. Furthermore, if $\mathcal{E}$ is also sub-unital (i.e. it satisfies $\mathcal{E}(1_A)\leq 1_{A'}$), then

$$H_{\min}^{\varepsilon}(A;\dot{B})_\rho \leq H_{\min}^{\varepsilon}(A';\dot{B}')_\tau.$$

Proof. See full version [7, Lem. 3]. □

Clearly every operational definition should be invariant under isometries as embeddings are essentially just a choice of modelling and should not effect operational quantities. This is exhibited by the following lemma.

Lemma 2. *Let $\rho_{AB} \in \mathcal{S}_\circ(AB)$ with valid (ε,Δ). For any two isometries $U:A\to A'$, $V:B\to B'$, it holds that*

$$\begin{aligned} I_{\max}^{\varepsilon,\Delta}(\dot{A};B)_\rho &= I_{\max}^{\varepsilon,\Delta}(\dot{A}';B')_\rho \quad \textit{and} \\ H_{\min}^{\varepsilon,\Delta}(A;\dot{B})_\rho &= H_{\min}^{\varepsilon,\Delta}(A';\dot{B}')_\rho \end{aligned}$$

where $\rho_{A'B'}=(U\otimes V)\rho_{AB}(U\otimes V)^\dagger$.

Proof. See full version [7, Lem. 4]. □

One could hope to replace $\tilde{\rho}_A\leq\rho_A$ in Definition 2 by an equality, thus forcing the state $\tilde{\rho}_{AB}$ to have the same trace as ρ_{AB}. However, for such a definition one would then need to show a property analogous to the above invariance under isometries, which seems non-trivial. The following argument gives also an indication that sub-normalized states are desirable in this context, although it does not conclusively show that they are necessary for our definition.

For the (unconditional) min-entropy, invariance under isometries can only hold if we allow sub-normalized states. To see this, consider the min-entropy of the state $\rho=1/d$,

[1] We call a tuple (ε,Δ) with $\varepsilon\geq 0$ valid for a state ρ if $\Delta(\rho,0)>\varepsilon$ (with 0 denoting the additive identity).

[2] The original definition of the smooth max-information in [3, Eq. 12] was slightly different and based on $D_{\max}(\tilde{\rho}_{AB}\|\tilde{\rho}_A\otimes\sigma_B)$.

[3] To see this, for example for the smooth max-information, note that if this were not so then the full dephasing map (in the classical basis) could be applied to both sides of the operator inequality

$$\tilde{\rho}_{AB}\leq\rho_A\otimes\sigma_B, \tag{1}$$

yielding a new feasible solution since the distance between ρ_{AB} and $\tilde{\rho}_{AB}$ is also reduced when the dephasing map is applied due to Lem. 1 (with $\varepsilon=0$).

which is maximal for normalized states of dimension d and thus cannot be increased by smoothing over this set. However, if embedded into a larger space smoothing will yield a larger min-entropy. Allowing sub-normalized states introduces an alternative to moving weight out of the support of ρ and it turns out that this is exactly what is needed to ensure the quantity is invariant under isometries.

IV. Relation to other entropy measures

A. Classical Setting

Since the (generalized) trace distance is directly connected to error probabilities it is often natural to stick to this distance measure for classical problems. We will do so in this section. We will also continue using the notations $\mathcal{P}, \mathcal{S}_\circ, \mathcal{S}_\bullet$, although now we restrict to diagonal matrices in some basis, interpreted as (potentially sub-normalized) probability distributions. In order to establish an asymptotic equipartition property for our locally smoothed information measures we relate them to other well-studied entropic quantities such as information spectrum divergences [8]. Note that standard asymptotic equipartition proofs for mutual information and conditional entropy do not leave any of the marginals unchanged.

Definition 4. *For $P_X, Q_X \in \mathcal{P}(X)$ and $\varepsilon \in [0,1]$, the max-information spectrum divergence is defined as*

$$D_s^\varepsilon(P_X\|Q_X) := \inf\left\{a : \Pr_{x\leftarrow p_X}\left\{\frac{P_X(x)}{Q_X(x)} > 2^a\right\} < \varepsilon\right\}.$$

Importantly, the max-information spectrum divergence has the following i.i.d. asymptotic second-order expansion [9]

$$\frac{1}{n}D_s^\varepsilon(P_X^{\times n}\|Q_X^{\times n}) = D(P_X\|Q_X) + \sqrt{\frac{V(P_X\|Q_X)}{n}}\cdot\Phi^{-1}(\varepsilon) + O\left(\frac{\log n}{n}\right)$$

with the Kullback-Leibler divergence and the relative entropy variance

$$D(P_X\|Q_X) := \sum_x P_X(x)\log\left(\frac{P_X(x)}{Q_X(x)}\right)$$
$$V(P_X\|Q_X) := \mathbb{E}\left[(\log P_X - \log Q_X - D(P_X\|Q_X))^2\right]$$

respectively, as well as the cumulative standard Gaussian distribution $\Phi(x) := \int_{-\infty}^x \frac{1}{2\pi}\exp(x^2/2)\,\mathrm{d}x$. We then define the information spectrum max-information and conditional min-entropy as

$$I_s^\varepsilon(X;Y)_P := D_s^\varepsilon(P_{XY}\|P_X\times P_Y) \quad \text{and}$$
$$H_s^\varepsilon(X|Y)_P := -D_s^\varepsilon(P_{XY}\|1_X\times P_Y)$$

respectively. This leads to the following equivalence result:

Theorem 1. *Let $P_{XY} \in \mathcal{S}_\circ(XY)$ and $0 < \varepsilon + \delta \leq 1$. Then, we have*

$$I_s^{\frac{\varepsilon}{1-\delta}+\delta}(X;Y)_P - \log\frac{1}{\delta^2} \leq I_{\max}^{\varepsilon,T}(\dot{X};Y)_P \leq I_s^\varepsilon(X;Y)_P + 1$$
$$H_s^{\frac{\varepsilon}{1-\delta}}(X|Y)_P + \log\frac{1}{\delta} \geq H_{\min}^{\varepsilon,T}(X|\dot{Y})_P \geq H_s^\varepsilon(X|Y)_P - 1\,.$$

Proof. See full version [7, Thm. 1]. □

This implies the asymptotic second-order expansions

$$\frac{1}{n}I_{\max}^{\varepsilon,T}(\dot{X};Y)_P = I(X;Y)_P + \sqrt{\frac{V(X;Y)_P}{n}}\cdot\Phi^{-1}(\varepsilon) + O\left(\frac{\log n}{n}\right)$$
$$\frac{1}{n}H_{\min}^{\varepsilon,T}(X|\dot{Y})_P = H(X|Y)_P + \sqrt{\frac{V(X|Y)_P}{n}}\cdot\Phi^{-1}(\varepsilon) + O\left(\frac{\log n}{n}\right)$$

with the mutual information variance $V(X;Y)_P := V(P_{XY}\|P_X\times P_Y)$ and the conditional information variance $V(X|Y)_P := V(P_{XY}\|1_x\times P_Y)$.

B. Quantum setting

For quantum problems Uhlmann's theorem indicates that it is natural to work with fidelity based distance measures such as the purified distance — which is what we will use in this section. Now, the equivalence proof of Theorem 1 as presented in [7, Thm. 1] crucially uses the idea of conditioning on the classical side information and hence we cannot give a direct quantum analogue. Instead, we find the following inequalities:

Theorem 2. *Let $\rho_{AB} \in \mathcal{S}_\circ(AB)$ and $0 \leq 2\varepsilon + \delta \leq 1$ with $\delta > 0$. Then, we have*

$$I_{\max}^{2\varepsilon+\delta,P}(\dot{A};B)_\rho \leq I_{\max}^{\varepsilon,P}(A;B)_\rho + \log\frac{8+\delta^2}{\delta^2}\,,$$

and by definition we also have the opposite inequality $I_{\max}^{\varepsilon,P}(\dot{A};B)_\rho \geq I_{\max}^{\varepsilon,P}(A;B)_\rho$. Moreover, we have

$$H_{\min}^{2\varepsilon+\delta,P}(A|\dot{B})_\rho \geq H_{\min}^{\varepsilon,P}(A|B)_\rho - \log\frac{8+\delta^2}{\delta^2}\,,$$

and by definition we also have the opposite inequality $H_{\min}^{\varepsilon,P}(A|\dot{B})_\rho \leq H_{\min}^{\varepsilon,P}(A|B)_\rho$.

Proof. See full version [7, Thm. 2-3]. □

V. Operational examples

We find that many existing proofs and protocols readily apply and give tighter bounds when combined with our restricted smoothing. In the following we discuss various basic classical and quantum examples in bipartite settings.

A. Classical state splitting

Let $\varepsilon \in (0,1]$ be the error parameter. There are two parties Alice and Bob. Alice possesses random variable X, taking values over a finite set $\mathcal{X}$ and a random variable Y, taking values over a finite set $\mathcal{Y}$. Alice sends a message to Bob and at the end Bob outputs random variable $\hat{Y}$ such that $T(P_{XY}, P_{X\hat{Y}}) \leq \varepsilon$. They are allowed to use shared randomness between them which is independent of XY at the beginning of the protocol.

We note that a generalization of this task with additional side information was studied in [10, Thm. 1]. These results together with [11] imply that the minimal number $R(P_{XY},\varepsilon)$

of bits communicated from Alice to Bob to achieve classical state splitting with error $\varepsilon \in (0,1]$ in generalized trace distance is bounded as

$$I_s^{\varepsilon/(1-\delta)}(P_{XY}\|P_X \times P_Y) - \log\frac{1}{\delta}$$
$$\leq R(P_{XY},\varepsilon) \leq I_s^{\varepsilon-3\delta}(P_{XY}\|P_X \times P_Y) + \log\frac{1}{\delta^2}$$

for $\delta \in (0,1)$ small enough. We show an even tighter characterization in terms of the smooth max-information.

Theorem 3. *Let $P_{XY} \in \mathcal{S}_\circ(XY)$. Then, for any $\delta \in (0,\varepsilon]$, the minimal number $R(P_{XY},\varepsilon)$ of bits communicated from Alice to Bob to achieve classical state splitting with error $\varepsilon \in (0,1]$ in generalized trace distance is bounded as*

$$I_{\max}^{\varepsilon,T}(\dot{X};Y)_P \leq R(P_{XY},\varepsilon) \leq I_{\max}^{\varepsilon-\delta,T}(\dot{X};Y)_P + \log\log\frac{1}{\delta^2}\,.$$

Proof. See full version [7, Thm. 4]. □

B. Strong privacy amplification against side information

For a set of two-universal hash functions $\{f^s_{X\to Z}\}_{s\in S}$ and classical-quantum states

$$\rho_{XB} = \sum_{X\in\mathcal{X}} |x\rangle\langle x| \otimes \rho_B^x \in \mathcal{S}_\circ(XB)$$

we use the same composable security criterion for ε-random and secret bits as, e.g, in [1, Sect. 7.3],

$$\Delta\Bigg(\underbrace{\sum_{\substack{s\in S\\ z\in\mathcal{Z}}} \frac{|sz\rangle\langle sz|_{SZ}}{|S|} \otimes \left(\sum_{x:f^s(x)=z} \rho_B^x\right)}_{=:\,\omega_{SZB}}, \frac{1_{SZ}}{|S||Z|}\otimes\rho_B\Bigg) \leq \varepsilon\,. \tag{2}$$

Note that in contrast to the setting studied in [12, Sect. III] or [13] we have a composable security definition by putting the reduced state on B on the lhs of Eq. (2). We refer to [14, App. B] for a more detailed discussion. The maximal number of ε-random and secret bits that can be extracted from ρ_{XB} such that (2) is denoted $\ell^\Delta(\rho_{XB},\varepsilon)$, where Δ is either P or T, as usual.

Theorem 4. *Let $\rho_{XB} \in \mathcal{S}_\circ(XB)$ be classical-quantum on XB and $\varepsilon \in (0,1]$. Then, the maximal number of ε-random and secret bits that can be extracted from ρ_{XB} is bounded as*

$$H_{\min}^{\varepsilon-\delta,P}(X|\dot{B})_\rho - \log\frac{1}{\delta^4} \leq \ell^P(\rho_{XB},\varepsilon) \leq H_{\min}^{\varepsilon,P}(X|\dot{B})_\rho \tag{3}$$

for any $\delta \in (0,\varepsilon]$. Moreover, when $B = Y$ is classical then we also have

$$H_{\min}^{\varepsilon-\delta,T}(X|\dot{Y})_P - \log\frac{1}{4\delta^2} \leq \ell^T(P_{XY},\varepsilon) \leq H_{\min}^{\varepsilon,T}(X|\dot{Y})_P\,. \tag{4}$$

This implies the asymptotic second-order expansion

$$\frac{1}{n}\ell^T\left(P_{XY}^{\times n},\varepsilon\right) = H(X|Y)_P + \sqrt{\frac{V(X|Y)_P}{n}}\cdot\Phi^{-1}(\varepsilon) + O\left(\frac{\log n}{n}\right)$$

as first given in [15, Thm. 25] (see also [16, Thm. 3]).

Proof. We first prove the lower bound in Eq. (3). Let $\tilde{\rho}_{XB} \in \mathcal{S}_\bullet(XB)$ be the optimizer in the definition of $H_{\min}^{\varepsilon-\delta,P}(X|\dot{B})_\rho$ and let

$$\tilde{\omega}_{SZB} := \frac{1}{|S|}\sum_{\substack{s\in S\\ z\in\mathcal{Z}}} |s\rangle\langle s|_S \otimes |z\rangle\langle z| \otimes \left(\sum_{x:f^s(x)=z} \tilde{\rho}_B^x\right). \tag{5}$$

Since by definition $\rho_B \geq \tilde{\rho}_B$ and by data-processing $P(\omega_{SZB},\tilde{\omega}_{SZB}) \leq \varepsilon - \delta$ we get that

$$\begin{aligned}
&P\left(\omega_{SZB}, \frac{1_S}{|S|}\otimes\frac{1_Z}{|Z|}\otimes\rho_B\right)\\
&\leq P\left(\omega_{SZB}, \frac{1_S}{|S|}\otimes\frac{1_Z}{|Z|}\otimes\tilde{\rho}_B\right)\\
&\leq P\left(\tilde{\omega}_{SZB}, \frac{1_S}{|S|}\otimes\frac{1_Z}{|Z|}\otimes\tilde{\rho}_B\right) + P(\omega_{SZB},\tilde{\omega}_{SZB})\\
&\leq P\left(\tilde{\omega}_{SZB}, \frac{1_S}{|S|}\otimes\frac{1_Z}{|Z|}\otimes\tilde{\rho}_B\right) + \varepsilon - \delta\\
&\leq \sqrt{2T\left(\tilde{\omega}_{SZB}, \frac{1_S}{|S|}\otimes\frac{1_Z}{|Z|}\otimes\tilde{\rho}_B\right)} + \varepsilon - \delta
\end{aligned}$$

where in the last step we employed the equivalence of generalized trace distance and purified distance [1, Lem. 3.5]. Now, standard achievability proofs such as [13, Thm. 6] applied to $\tilde{\rho}_{XB} \in \mathcal{S}_\bullet(XB)$ give

$$T\left(\tilde{\omega}_{SZB}, \frac{1_S}{|S|}\otimes\frac{1_Z}{|Z|}\otimes\tilde{\rho}_B\right) \leq \frac{1}{2}\sqrt{|Z|\cdot 2^{-H_{\min}^{\varepsilon-\delta,P}(X|\dot{B})_\rho}}$$

and choosing $\log|Z| = H_{\min}^{\varepsilon-\delta,P}(X|\dot{B})_\rho - \log\frac{1}{\delta^4}$ leads to the claim. For the upper bound in Eq. (3) we follow [1, Sect. 7.3.3] but adapted to our partially smoothed conditional min-entropy. Namely, assume by contradiction that there exists a protocol which extracts $\ell > H_{\min}^{\varepsilon,P}(X|\dot{B})_\rho$ bits of ε-random and secret bits. Then, since applying a function on X cannot increase the smooth conditional min-entropy [7, Lem. 6] we have for all $s \in \mathcal{S}$ that

$$\ell > H_{\min}^{\varepsilon,P}(X|\dot{B})_\rho \geq H_{\min}^{\varepsilon,P}(Z|\dot{B})_{\rho^s} \quad \text{with}$$
$$\rho_{ZB}^s := \sum_{z\in\mathcal{Z}} |z\rangle\langle z|_Z \otimes \left(\sum_{x:f^s(x)=z} \rho_B^x\right).$$

Hence, for all $\tilde{\rho}_{ZB} \in \mathcal{S}_\bullet(ZB)$ with $P(\tilde{\rho}_{ZB},\rho_{ZB}^s) \leq \varepsilon$ we have $H_{\min}(Z|B)_{\tilde{\rho}} < \ell$. This in turn implies $P(\rho_{ZB}^s, \frac{1_Z}{|Z|}\otimes\rho_B) > \varepsilon \implies P(\omega_{SZB}, \frac{1_S}{|S|}\otimes\frac{1_Z}{|Z|}\otimes\rho_B) > \varepsilon$, which is in contradiction to Eq. (2).

The upper bound in Eq. (4) follows in the same way as the upper bound in Eq. (3), just by noting that in the classical case the monotonicity under functions also holds for the generalized trace distance [7, Lem. 6]. For the lower bound in Eq. (4), denote in the security criterion Eq. (2) the state ω_{SZB} for $B = Y$ classical by Q_{SZY}, let $\tilde{P}_{XY} \in \mathcal{S}_\bullet(XY)$ be the optimizer in the definition of $H_{\min}^{\varepsilon-\delta,T}(X|\dot{Y})_P$, and let $\tilde{Q}_{SZY}$ be defined

as in Eq. (5). Since by definition $P_Y \geq \tilde{P}_Y$ and by data-processing $T\left(Q_{SZY}, \tilde{Q}_{SZY}\right) \leq \varepsilon - \delta$ we get that

$$\begin{aligned} & T\left(Q_{SZY}, \frac{1_S}{|S|} \times \frac{1_Z}{|Z|} \times P_Y\right) \\ & \leq T\left(\tilde{Q}_{SZY}, \frac{1_S}{|S|} \times \frac{1_Z}{|Z|} \times \tilde{P}_Y\right) + T\left(Q_{SZY}, \tilde{Q}_{SZY}\right) \\ & \leq T\left(\tilde{Q}_{SZY}, \frac{1_S}{|S|} \times \frac{1_Z}{|Z|} \times \tilde{P}_Y\right) + \varepsilon - \delta\,. \end{aligned}$$

Now, standard achievability proofs such as [13, Thm. 6] applied to $\tilde{P}_{XY} \in \mathcal{S}_\bullet(XY)$ lead to the claim for $\log|Z| = H_{\min}^{\varepsilon-\delta,T}(X|\dot{Y})_P - \log\frac{1}{4\delta^2}$. □

C. Quantum state merging

A pure tripartite state ρ_{ABR} is shared between parties Alice (A), Bob (B), and the reference R. The goal is to send the A-marginal from Alice to Bob using classical communication and entanglement assistance while not changing the overall state [17], [18], [19], [20]. More precisely, for $\rho_{ABR} \in \mathcal{S}_\circ(ABR)$ of rank-one and A_0B_0 additional quantum systems, a quantum channel

$$\mathcal{E} : AA_0 \otimes BB_0 \to A_1 \otimes B_1\bar{B}B$$

is a quantum state merging of ρ_{ABR} with error $\varepsilon \in [0,1]$, if it is a local operation with classical forward communication process for the bipartition $AA_0 \to A_1$ versus $BB_0 \to B_1\bar{B}B$, and $P\big((\mathcal{E}\otimes\mathcal{I}_R)(\Phi_{A_0B_0}\otimes\rho_{ABR}), \Phi_{A_1B_1}\otimes\rho_{B\bar{B}R}\big) \leq \varepsilon$ where $\rho_{B\bar{B}R} = (\mathcal{I}_{A\to\bar{B}} \otimes \mathcal{I}_{BR})(\rho_{ABR})$, and $\Phi_{A_0B_0}$, $\Phi_{A_1B_1}$ are maximally entangled states on A_0B_0, A_1B_1, respectively. The difference $\log|A_0| - \log|A_1|$ quantifies the entanglement cost. We find the following theorem.

Theorem 5. *Let $\rho_{ABR} \in \mathcal{S}_\circ(ABR)$ be of rank-one. For free classical communication assistance the minimal entanglement cost $E(\rho_{ABR}, \varepsilon)$ for quantum state merging of ρ_{ABR} with error $\varepsilon \in (0,1]$ in purified distance is bounded as*

$$-H_{\min}^{\varepsilon,P}(A|\dot{R})_\rho \leq E(\rho_{ABR},\varepsilon) \leq -H_{\min}^{\varepsilon-\delta,P}(A|\dot{R})_\rho + \log\frac{1}{\delta^4}$$

for any $\delta \in (0,\varepsilon]$. Alternatively, for unlimited entanglement assistance — not necessarily constraint to the form of maximally entangled states — the minimal classical communication cost $C(\rho_{ABR},\varepsilon)$ for quantum state merging of ρ_{ABR} with error $\varepsilon \in (0,1]$ in purified distance is bounded as

$$I_{\max}^{\varepsilon,P}(\dot{R};A)_\rho \leq C(\rho_{ABR},\varepsilon) \leq I_{\max}^{\varepsilon-\delta,P}(\dot{R};A)_\rho + \log\frac{1}{\delta^4}$$

for any $\delta \in (0,\varepsilon]$.

Proof. See full version [7, Thm. 6] □

The asymptotic first-order expansions then follow from the asymptotic first order expansion for max-information and conditional min-entropy. Thus, we recover the original results on quantum state merging [17], [18]. As shown in these references in first-order asymptotically the entanglement cost and classical communication cost can actually be simultaneously minimised — whereas this becomes unclear in the one-shot setting. The asymptotic second-order expansions are an open problem but are now again reduced to giving the asymptotic second-order expansions of $H_{\min}^{\varepsilon,P}(A|\dot{R})_\rho$ and $I_{\max}^{\varepsilon,P}(\dot{R};A)_\rho$, respectively.

VI. Outlook

As we have seen our locally smoothed information measures naturally appear in a plethora of operational tasks in quantum information theory. The main open problem raised by our work is to give asymptotic i.i.d. second-order expansions of the partially smoothed information measures $H_{\min}^{\varepsilon,\Delta}(A|\dot{B})_\rho$ and $I_{\max}^{\varepsilon,\Delta}(\dot{A};B)_\rho$ for the quantum case.

References

[1] M. Tomamichel, *Quantum Information Processing with Finite Resources — Mathematical Foundations*, vol. 5 of *SpringerBriefs in Mathematical Physics*. 2016.

[2] R. Renner, *Security of Quantum Key Distribution*. PhD thesis, ETH Zurich, 2005.

[3] M. Berta, M. Christandl, and R. Renner, "The Quantum Reverse Shannon Theorem Based on One-Shot Information Theory," *Communications in Mathematical Physics*, vol. 306, no. 3, pp. 579–615, 2011.

[4] R. Jain, J. Radhakrishnan, and P. Sen, "Privacy and Interaction in Quantum Communication Complexity and a Theorem About the Relative Entropy of Quantum States," in *Proceedings IEEE FOCS*, pp. 429–438, 2002.

[5] N. Datta, "Min- and Max- Relative Entropies and a New Entanglement Monotone," *IEEE Transactions on Information Theory*, vol. 55, no. 6, pp. 2816–2826, 2009.

[6] M. Tomamichel, R. Colbeck, and R. Renner, "Duality Between Smooth Min- and Max-Entropies," *IEEE Transactions on Information Theory*, vol. 56, no. 9, pp. 4674–4681, 2010.

[7] A. Anshu, M. Berta, R. Jain, and M. Tomamichel, "Partially Smoothed Information Measures," *arXiv:1807.05630*, 2018.

[8] T. S. Han, *Information-Spectrum Methods in Information Theory*. Applications of Mathematics, Springer, 2002.

[9] V. Strassen, "Asymptotische Abschätzungen in Shannons Informationstheorie," in *Transactions Prague Conference on Information Theory*, pp. 689–723, 1962.

[10] A. Anshu, R. Jain, and N. A. Warsi, "A Unified Approach to Source and Message Compression," *arXiv:1707.03619*, 2017.

[11] M. Braverman and A. Rao, "Information Equals Amortized Communication," in *Proceedings IEEE FOCS*, pp. 748–757, 2011.

[12] M. Tomamichel and M. Hayashi, "A Hierarchy of Information Quantities for Finite Block Length Analysis of Quantum Tasks," *IEEE Transactions on Information Theory*, vol. 59, no. 11, pp. 7693–7710, 2013.

[13] M. Tomamichel, C. Schaffner, A. Smith, and R. Renner, "Leftover Hashing Against Quantum Side Information," *IEEE Transactions on Information Theory*, vol. 57, no. 8, pp. 5524–5535, 2011.

[14] C. Portmann and R. Renner, "Cryptographic Security of Quantum Key Distribution," *arXiv:1409.3525*, 2014.

[15] M. Hayashi, "Security Analysis of ε -Almost Dual Universal2 Hash Functions: Smoothing of Min Entropy Versus Smoothing of Renyi Entropy of Order 2," *IEEE Transactions on Information Theory*, vol. 62, no. 6, pp. 3451–3476, 2016.

[16] S. Watanabe and M. Hayashi, "Non-Asymptotic Analysis of Privacy Amplification via Rényi Entropy and Inf-Spectral Entropy," in *Proceedings IEEE ISIT*, pp. 2715–2719, 2013.

[17] M. Horodecki, J. Oppenheim, and A. Winter, "Partial Quantum Information," *Nature*, vol. 436, no. 7051, pp. 673–6, 2005.

[18] M. Horodecki, J. Oppenheim, and A. Winter, "Quantum State Merging and Negative Information," *Communications in Mathematical Physics*, vol. 269, no. 1, pp. 107–136, 2006.

[19] M. Berta, "Single-Shot Quantum State Merging," Diploma Thesis, ETH Zurich, 2008.

[20] F. Dupuis, M. Berta, J. Wullschleger, and R. Renner, "One-Shot Decoupling," *Communications in Mathematical Physics*, vol. 328, no. 1, p. 251, 2014.