Tight relations and equivalences between smooth relative entropies

Bartosz Regula
RIKEN Center for Quantum Computing
Wako, Japan
Email: bartosz.regula@gmail.com

Ludovico Lami
Scuola Normale Superiore
Pisa, Italy
Email: ludovico.lami@gmail.com

Nilanjana Datta
University of Cambridge
Cambridge, United Kingdom
Email: n.datta@damtp.cam.ac.uk

***Abstract*—The precise one-shot characterisation of operational tasks in classical and quantum information theory relies on different forms of smooth entropic quantities. A particularly important connection is between the hypothesis testing relative entropy and the smoothed max-relative entropy, which together govern many operational settings. We first strengthen this connection into a type of equivalence: we show that the hypothesis testing relative entropy is equivalent to a variant of the smooth max-relative entropy based on the information spectrum divergence, which can be alternatively understood as a measured smooth max-relative entropy. Furthermore, we improve a fundamental lemma due to Datta and Renner that connects the different variants of the smoothed max-relative entropy, introducing a modified proof technique based on matrix geometric means. We use the unveiled connections and tools to strictly improve on previously known one-shot bounds and duality relations between the smooth max-relative entropy and the hypothesis testing relative entropy, sharpening also bounds that connect the max-relative entropy with Rényi divergences.**

I. INTRODUCTION

Smooth relative entropies [1] were defined to address the need for a precise understanding of the performance of various operational protocols in information theory in settings beyond the asymptotic i.i.d. one. In *one-shot information theory*, several different, inequivalent types of smooth entropies are required to fully characterise the performance of different tasks. Similarly to the broad division of tasks in classical information theory into two types, packing-type and covering-type problems, the more general setting of quantum information can also be categorised in a similar way. On one side, there are problems that can be connected with hypothesis testing, including quantum channel coding [2]–[6], quantum data compression [6], [7], and quantum resource distillation [8]–[10]. Due to the underlying hypothesis testing structure, the one-shot performance of these protocols can be related with a quantity known as the *hypothesis testing relative entropy* D_H^ε [2], [3]. On the other side, covering-type problems include quantum channel simulation [11], [12], privacy amplification [1], [13], [14], decoupling [15], convex split [16], and quantum resource dilution [8], [9]. The one-shot characterisation of such protocols typically relies on a quantity known as the *smooth max-relative entropy* $D_{\max}^\varepsilon$ [1], [17]. Understanding the precise relations between the two smooth relative entropies has thus been an important problem in one-shot quantum information theory [13], [18]–[21].

Despite their many differences, the hypothesis testing relative entropy and the max-relative entropy are known to be quantitatively related, albeit in a complementary fashion. Informally, the value of $D_{\max}^\varepsilon$ closely matches that of $D_H^{1-\varepsilon}$, up to asymptotically negligible factors [13], [18]. This means that the two smooth quantities satisfy a 'weak/strong converse duality': understanding the behaviour of D_H^ε in the small ε regime (weak converse) is equivalent to understanding the large ε (strong converse) behaviour of $D_{\max}^\varepsilon$, and vice versa. This fact was crucial in obtaining results such as asymptotic equipartition theorems for smooth entropic quantities [18], [22]–[27] or in characterising the higher-order corrections of optimal rates in quantum information tasks [13], [28]. However, the precise one-shot bounds that connected $D_{\max}^\varepsilon$ and $D_H^{1-\varepsilon}$ in the literature are often not tight, motivating us to look for alternative approaches and stronger links.

A. Summary of main results

Here we introduce new techniques for the study of the relations between the two quantities, strengthening both the conceptual and the quantitative connections. Our approach employs an intermediary quantity $\widetilde{D}_{\max}^\varepsilon$, closely related to the smooth max-relative entropy and originally introduced in [28] as a type of an information spectrum divergence [13]. This relaxed variant of max-relative entropy was previously used as a technical tool in many proofs, but our closer investigation reveals that it is intrinsically connected to $D_H^{1-\varepsilon}$, and in fact equivalent to it in a precise sense (Section III). This finding allows us to establish new tight bounds between the hypothesis testing relative entropy and $\widetilde{D}_{\max}^\varepsilon$. To relate the latter with the operationally important quantity $D_{\max}^\varepsilon$, previous works relied on an important lemma originally shown by Datta and Renner [29], which we, however, notice is not tight for several types of distance measures employed in the smoothing. We introduce an improved proof technique based on the operator geometric mean that leads to a statement that tightens the previous formulations of this lemma, in particular when the smoothing is over normalised quantum states (Section IV). Our proof method is also easier to generalise, which we exemplify with a multi-partite extension of the lemma to a setting of simultaneous state smoothing [30]. We apply our results to obtain a set of improved bounds that tighten the duality relation between $D_{\max}^\varepsilon$ and $D_H^{1-\varepsilon}$, strengthening both the precise one-shot relations between the two quantities (Section V-A) as well as the bounds connecting the smoothed max-relative entropy with Rényi divergences (Section V-B). Additional extensions of our results, further relations and equivalences, and complete technical details of all proofs can be found in our technical paper [31].

II. SMOOTH DIVERGENCES

A. Hypothesis testing entropy and max-relative entropy

We use Greek letters (ρ, σ) to denote quantum states acting on a finite-dimensional Hilbert space. We will sometimes specialise our results to classical probability distributions, which are denoted by Latin letters (p, q).

Smooth divergences are broadly understood as divergences (relative entropies) between quantum states that incorporate the allowance for some form of error or uncertainty about the states under consideration, quantified by a smoothing parameter ε.

A fundamental divergence that finds use in one-shot quantum information theory is the ***hypothesis testing relative entropy***, defined for two quantum states ρ and σ as [2], [3]

$$D_H^\varepsilon(\rho\|\sigma) := -\log\inf\left\{ \operatorname{Tr} M\sigma \;\middle|\; 0 \le M \le \mathbb{1},\ \operatorname{Tr}(\mathbb{1}-M)\rho \le \varepsilon \right\},$$

where $\varepsilon \in [0,1]$. This can be noticed to be simply the optimal exponent of the type II error of hypothesis testing between ρ and σ when the type I error probability is constrained to be at most ε.

Another fundamental quantity that we focus on here is the ***max-relative entropy*** [17]

$$D_{\max}(\rho\|\sigma) := \log\inf\left\{ \lambda \ge 0 \;\middle|\; \rho \le \lambda\sigma \right\}. \tag{1}$$

This can be understood as the sandwiched Rényi divergence $\widetilde{D}_\alpha$ of order ∞ [32]. The smooth variant of this quantity is defined not by evaluating $D_{\max}(\rho\|\sigma)$ exactly, but instead by optimising over all states $\rho' \approx_\varepsilon \rho$ in an ε-ball around ρ in order to find the minimal value of $D_{\max}(\rho'\|\sigma)$ [1], [17]. This definition crucially depends on how exactly we define the smoothing neighbourhood, and in particular on how we quantify the distance between quantum states.

Two of the most fundamental measures of distance are the trace (total variation) distance $\frac{1}{2}\left\|\rho-\rho'\right\|_1$ and the purified distance $P(\rho,\rho') := \sqrt{1-F(\rho,\rho')}$, where $F(\rho,\rho')$ denotes the fidelity (Bhattacharyya coefficient)

$$F(\rho,\rho') := \left\|\sqrt{\rho}\sqrt{\rho'}\right\|_1^2 = \left(\operatorname{Tr}\sqrt{\sqrt{\rho}\rho'\sqrt{\rho}}\right)^2. \tag{2}$$

We then define the ***smooth max-relative entropy*** as

$$D_{\max}^{\varepsilon,\Delta}(\rho\|\sigma) := \inf_{\rho'\in\mathcal{B}_\Delta^\varepsilon(\rho)} D_{\max}(\rho'\|\sigma), \tag{3}$$

where Δ denotes one of the following choices of smoothing: T for trace distance smoothing with $\mathcal{B}_T^\varepsilon(\rho) := \left\{\rho' \;\middle|\; \frac{1}{2}\left\|\rho-\rho'\right\|_1 \le \varepsilon,\ \rho' \ge 0,\ \operatorname{Tr}\rho' = 1\right\}$, and P for purified distance smoothing.

B. Modified max-relative entropy $\widetilde{D}_{\max}$

An important quantity in our approach will be a variant of the smooth max-relative entropy, formalised by Datta and Leditzky [28] as

$$\widetilde{D}_{\max}^\varepsilon(\rho\|\sigma) := \log\inf\left\{ \lambda \ge 0 \;\middle|\; \operatorname{Tr}(\rho-\lambda\sigma)_+ \le \varepsilon \right\}, \tag{4}$$

where $\operatorname{Tr}(\cdot)_+$ denotes the trace of the positive part of a Hermitian operator. In fact, $\widetilde{D}_{\max}^\varepsilon$ was first introduced as an alternative to the information spectrum divergence D_s^ε [13] and denoted $\overline{D}_s^{\,\varepsilon}$ in [28]. Here we prefer to relate it to the max-relative entropy, the connection with which becomes clearer once we use the variational form of $\operatorname{Tr}(\cdot)_+$ to write

$$\widetilde{D}_{\max}^\varepsilon(\rho\|\sigma) = \log\inf\left\{ \lambda \;\middle|\; \rho \le \lambda\sigma + Q,\ Q \ge 0,\ \operatorname{Tr} Q \le \varepsilon \right\},$$

which more closely resembles the optimisation problem that defines $D_{\max}$. For classical systems, the connection of this quantity with $D_{\max}^\varepsilon$ is much stronger: we have a complete equivalence between the modified and the standard smooth max-relative entropy. This was first noticed in [33], [34], where a classical equivalent of $\widetilde{D}_{\max}^\varepsilon$ was introduced in the context of characterising differential privacy.

The modified max-relative entropy $\widetilde{D}_{\max}^\varepsilon$ made implicit appearances in many early works on quantum hypothesis testing already [23], [35], [36] through its connection with the information spectrum method [29], [36]. It was first formalised as a divergence in [28], where it was used to derive second-order expansions for i.i.d. quantum information tasks. Some of its basic properties, including the data processing inequality, were established there. More recently, it also found use in characterising quantum differential privacy [37]–[39], generalising the earlier classical results [33], [34]. Although it was studied mostly as a technical tool, we will show that the quantity has many precise connections with both the hypothesis testing relative entropy D_H^ε and with the standard smoothed max-relative entropy $D_{\max}^{\varepsilon,T}$ that were not previously known.

III. AN EQUIVALENCE

Although the connection between $D_{\max}^\varepsilon$ and $D_H^{1-\varepsilon}$ is now a well-known duality in one-shot quantum information theory, the precise bounds between the two are often rather loose — they match in the asymptotic regime, but large gaps are expected between them at the single-copy level [13]. Surprisingly, here we show that, as long as one considers the modified max-relative entropy variant $\widetilde{D}_{\max}^\varepsilon$, the hypothesis testing relative entropy and the smooth max-relative entropy are in a precise sense equivalent to each other — either of them can be obtained from the other. This will allow us to obtain much tighter bounds between the two quantities than were previously known.

Theorem 1. *For all quantum states ρ and σ and for all $\varepsilon \in (0,1)$ it holds that*

$$D_H^{1-\varepsilon}(\rho\|\sigma) = \inf_{\delta\in[0,\varepsilon)}\left[\widetilde{D}_{\max}^\delta(\rho\|\sigma) - \log(\varepsilon-\delta)\right], \tag{5}$$

$$\widetilde{D}_{\max}^\varepsilon(\rho\|\sigma) = \sup_{\delta\in(\varepsilon,1]}\left[D_H^{1-\delta}(\rho\|\sigma) + \log(\delta-\varepsilon)\right]. \tag{6}$$

Proof. We first compute the Lagrange dual of D_H^ε as (see [31])

$$\log \inf_{h,y,P} \left\{ h \,\middle|\, y\rho \le h\sigma + P,\ P \ge 0,\ h,y \ge 0,\ 1+\operatorname{Tr} P \le y\varepsilon \right\},$$

where equality follows by strong duality: $z = 1$, $M' = \varepsilon\mathbb{1}$ is a strictly feasible solution, so Slater's criterion applies [40, Sec. 5.9]. Substituting $\mu = \frac{1}{y}$, $Q = \mu P$, and $\lambda = h\mu$ we get

$$\begin{aligned}
&D_H^{1-\varepsilon}(\rho\|\sigma) \\
&= \log \inf_{h,\mu,P} \left\{ h \,\middle|\, \rho \le h\mu\,\sigma + \mu P,\ P \ge 0,\ h, \frac{1}{\mu} \ge 0, \right. \\
&\qquad\qquad \left. 1 + \operatorname{Tr} P \le \frac{1}{\mu}\varepsilon \right\} \\
&= \log \inf_{h,\mu,Q} \left\{ h \,\middle|\, \rho \le h\mu\,\sigma + Q,\ Q \ge 0,\ h \ge 0,\ \mu > 0, \right. \\
&\qquad\qquad \left. \operatorname{Tr} Q \le \varepsilon - \mu \right\} \\
&= \log \inf_{\mu\in(0,\varepsilon]} \frac{1}{\mu} \inf_{\lambda,Q} \left\{ \lambda \mid \rho \le \lambda\sigma + Q,\ Q \ge 0,\ \operatorname{Tr} Q \le \varepsilon - \mu \right\} \\
&= \inf_{\mu\in(0,\varepsilon]} \left[\widetilde{D}_{\max}^{\varepsilon-\mu}(\rho\|\sigma) - \log\mu \right],
\end{aligned} \tag{7}$$

where in the last line we recalled the definition of $\widetilde{D}_{\max}$ in (II-B). This is precisely the expression claimed in (5).

Take now $\varepsilon \in [0,1)$. Evaluating the Lagrange dual of the optimisation problem that defines $\widetilde{D}_{\max}^\varepsilon$ (see also [39]), we have

$$\begin{aligned}
&\widetilde{D}_{\max}^\varepsilon(\rho\|\sigma) \\
&= \log\sup \left\{ \frac{\operatorname{Tr} W'\rho - \varepsilon}{\operatorname{Tr} W'\sigma} \,\middle|\, 0 \le W' \le \mathbb{1} \right\} \\
&\overset{\text{(i)}}{=} \log\sup \left\{ \frac{\operatorname{Tr} W'\rho - \varepsilon}{\operatorname{Tr} W'\sigma} \,\middle|\, 0 \le W' \le \mathbb{1},\ \operatorname{Tr} W'\rho > \varepsilon \right\} \\
&= \log \sup_{\delta\in(\varepsilon,1]} \sup \left\{ \frac{\operatorname{Tr} W'\rho - \varepsilon}{\operatorname{Tr} W'\sigma} \,\middle|\, 0 \le W' \le \mathbb{1},\ \operatorname{Tr} W'\rho = \delta \right\} \\
&= \log \sup_{\delta\in(\varepsilon,1]} \sup \left\{ \frac{\delta - \varepsilon}{\operatorname{Tr} W'\sigma} \,\middle|\, 0 \le W' \le \mathbb{1},\ \operatorname{Tr} W'\rho = \delta \right\} \\
&\overset{\text{(ii)}}{=} \log \sup_{\delta\in(\varepsilon,1]} \sup \left\{ \frac{\delta - \varepsilon}{\operatorname{Tr} W'\sigma} \,\middle|\, 0 \le W' \le \mathbb{1},\ \operatorname{Tr} W'\rho \ge \delta \right\} \\
&\overset{\text{(iii)}}{=} \sup_{\delta\in(\varepsilon,1]} \left[D_H^{1-\delta}(\rho\|\sigma) + \log(\delta - \varepsilon) \right]
\end{aligned} \tag{8}$$

where in (i) we observed that, since the optimal value must be positive (consider that $W' = \mathbb{1}$ is feasible), we can restrict to operators W' such that $\operatorname{Tr} W'\rho > \varepsilon$ without loss of generality; in (ii), we used the fact that the condition $\operatorname{Tr} W'\rho = \delta$ can be relaxed to $\operatorname{Tr} W'\rho \ge \delta$ without loss of generality: for any W'' with $\operatorname{Tr} W''\rho > \delta$ we can scale it down to a solution $W' \le W''$ which has $\operatorname{Tr} W'\rho = \delta$ and for which the optimal value cannot be smaller; finally, (iii) is by definition of $D_H^{1-\delta}$. ■

The equivalence derived in Theorem 1 can be made even stronger in the case of classical distributions (or commuting quantum states), where instead of the quantity $\widetilde{D}_{\max}^\varepsilon$, the smoothed max-relative entropy itself features in the relation [31], providing a close connection between the two types of smoothed entropies.

IV. IMPROVING A USEFUL LEMMA

Although we have already shown a very close relation between D_H^ε and the modified quantity $\widetilde{D}_{\max}^\varepsilon$, this is so far insufficient to tightly connect the hypothesis testing relative entropy with the smooth max-relative entropy $D_{\max}^\varepsilon$ itself. The issue is that, although one can straightforwardly upper bound $\widetilde{D}_{\max}^\varepsilon$ with $D_{\max}^{\varepsilon,T}$, the other direction is much less obvious: $\widetilde{D}_{\max}^\varepsilon$ involves an optimisation over values of λ such that $\rho \le \lambda\sigma + Q$, where Q is a positive semidefinite operator of sufficiently small trace, but how can this operator inequality yield a state $\rho' \approx_\varepsilon \rho$ that can be used as a feasible solution for the smooth max-relative entropy?

A solution to this conundrum was first given in a fundamental lemma by Datta and Renner [29, Lemma 3]. It found use in many results that studied the asymptotics of the max-relative entropy, e.g. [18], [22]–[27]. Here we introduce a modified proof approach, allowing us to give better estimates.

Theorem 2 (Tightened Datta–Renner lemma)**.** *Let ρ be a state, $A, Q \ge 0$ positive semi-definite with $\operatorname{Tr} Q \le \varepsilon < 1$, and assume that $\rho \le A + Q$ holds. Then there exists a subnormalised state ρ' with $\operatorname{Tr}\rho' \le 1$ such that*

$$\rho' \le A, \qquad \frac{\rho'}{\operatorname{Tr}\rho'} \le \frac{A}{1-\varepsilon}\,; \tag{9}$$

moreover, ρ' is close to ρ in the following senses:

$$F\big(\rho,\, \rho'\big) \ge (1-\varepsilon)^2\,, \tag{10}$$

$$F\Big(\rho,\, \frac{\rho'}{\operatorname{Tr}\rho'}\Big) \ge 1 - \varepsilon\,, \tag{11}$$

$$\frac{1}{2}\left\| \rho - \frac{\rho'}{\operatorname{Tr}\rho'} \right\|_1 \le \sqrt{\varepsilon}\,. \tag{12}$$

Here, the fidelity bound with unnormalised states in (10) is the same as the one found in [25], [29], but the other bounds improve on known estimates (e.g. [23], [29]), and in particular on the statement of the lemma with normalised smoothing found in [23, Lemma C.5].

The previous works [23], [25], [29] proved variants of this result by using the operator $T := A^{1/2}(A+Q)^{-1/2}$ to define the ansatz $\rho' := T\rho T^\dagger$, which is easily verified to satisfy $\rho' \le A$. The non-positivity of the operator T, however, prevents us from using well-known tools such as the celebrated gentle measurement lemma [41] that would allow us to bound the distance between ρ' and ρ, leading to suboptimal estimates. Is there, then, a way to instead pick a positive semidefinite operator in this approach?

Notice that the operator $G := A^{1/2}U(A+Q)^{-1/2}$, where U is some unitary, still satisfies $G\rho G^\dagger \le A$. The question then becomes whether one can choose U in a way that makes G positive semidefinite. This is indeed always possible, and if $A > 0$, then such a solution is unique [42, Proposition 4.1.8]:

U equals $(A^{-1/2}(A+Q)^{-1}A^{-1/2})^{1/2}A^{1/2}(A+Q)^{1/2}$, and G becomes the ***geometric mean*** of A and $(A+Q)^{-1}$.

More generally, for $A, B \geq 0$ one defines the geometric mean $A \# B$ as

$$A \# B \coloneqq A^{1/2}\left(A^{-1/2}BA^{-1/2}\right)^{1/2}A^{1/2}. \tag{13}$$

The properties of the geometric mean that will be relevant to us is that it is positive semidefinite, and that it is monotone non-increasing in either argument: in particular, $A \leq C$ implies that $A \# B \leq C \# B$ (see e.g. [42, Ch. 4.1]). We can now prove Theorem 2 using the reasoning outlined above.

Proof of Theorem 2. Up to projecting down onto the support of $A+Q$, we can assume without loss of generality that $A+Q>0$ is invertible. Define

$$G \coloneqq A \# \left((A+Q)^{-1}\right). \tag{14}$$

Due to the monotonicity of the operator geometric mean we have that

$$0 \leq G \leq (A+Q) \# \left((A+Q)^{-1}\right) = \mathbb{1}\,. \tag{15}$$

Conjugating by G, from $\rho \leq A+Q$ we deduce that

$$\rho' \coloneqq G\rho G \leq G(A+Q)G = A\,, \tag{16}$$

as a simple calculation using the formula on the rightmost side of (14) reveals. We now estimate

$$\begin{aligned} 1-\operatorname{Tr}\rho G^2 &= \operatorname{Tr}\rho\left(\mathbb{1}-G^2\right) \leq \operatorname{Tr}(A+Q)\left(\mathbb{1}-G^2\right) \\ &= \operatorname{Tr}(A+Q) - \operatorname{Tr}G(A+Q)G = \operatorname{Tr}Q \leq \varepsilon\,, \end{aligned} \tag{17}$$

where the first inequality holds due to the fact that $\mathbb{1}-G^2 \geq 0$. Applying the gentle measurement lemma [41], [43] gives the claimed bounds. ■

We note that, although the known formulations of the gentle measurement lemma give tight estimates on the distance of the normalised state $\frac{G\rho G}{\operatorname{Tr}G^2\rho}$ to ρ, it turns out that the previous bounds for the *sub*normalised state $G\rho G$ were not tight in trace distance. In the full technical paper [31], we consider variants of the max-relative smoothed over such subnormalised states in more detail, and in particular we provide a *gentler* measurement lemma that improves the distance estimates.

V. TIGHTENED BOUNDS AND RELATIONS

A. Inequalities between smooth max-relative entropy and hypothesis testing relative entropy

We are now ready to tackle the question of relating the two fundamental quantities, $D^{\varepsilon}_{\max}$ and D^{ε}_{H}.

The first key ingredient will be tight bounds between D^{ε}_{H} and the modified smooth max-relative entropy $\widetilde{D}^{\varepsilon}_{\max}$, obtained from the precise connection between the two that we established in Theorem 1.

Lemma 3. *For all $\varepsilon \in (0,1)$ and all $\mu \in (0,\varepsilon]$, it holds that*

$$\begin{aligned} \widetilde{D}^{\varepsilon}_{\max}(\rho\|\sigma) + \log\frac{1}{\varepsilon(1-\varepsilon)} &\leq D^{1-\varepsilon}_{H}(\rho\|\sigma) \\ &\leq \widetilde{D}^{\varepsilon-\mu}_{\max}(\rho\|\sigma) + \log\frac{1}{\mu}. \end{aligned} \tag{18}$$

The first inequality improves over the previously known bound of [28, Proposition 4.7], and indeed also over a stronger bound that was implicit in the proof of [18, Theorem 11]. The second inequality was known [28]. A complete proof of the Lemma can be found in [31].

A point of note here is that the bounds can be verified to be tight in many ways. The tightness of the error term $\log\frac{1}{\varepsilon(1-\varepsilon)}$ is particularly easy to see in the trivial case $\rho=\sigma$, where it holds that $D^{1-\varepsilon}_{H}(\rho\|\rho) = \log\frac{1}{\varepsilon}$ and $\widetilde{D}^{\varepsilon}_{\max}(\rho\|\rho) = \log(1-\varepsilon)$. In light of Theorem 1, we also see that the upper bound is as tight as possible, as it holds with equality by taking the infimum over μ. The lower bound is additionally tight in an i.i.d. asymptotic sense at the level of exponents, as we will shortly see in Sec. V-B.

Applying now the improved Datta–Renner lemma (Theorem 2) with the choice $A=\lambda\sigma$ immediately gives bounds connecting $\widetilde{D}^{\varepsilon}_{\max}(\rho\|\sigma)$ with the smoothed max-relative entropy. Putting our findings together, we obtain upper and lower bounds that directly relate $D^{\varepsilon}_{\max}$ with $D^{1-\varepsilon}_{H}$, recovering the known duality between the two quantities and improving on many of the previously known quantitative bounds between them.

Corollary 4 (Tightened weak/strong converse duality between $D_{\max}$ and D_H)**.** *For all quantum states ρ and σ, all $\varepsilon \in (0,1)$, and all $\mu \in (0,\varepsilon]$, it holds that*

$$D^{\sqrt{\varepsilon},T}_{\max}(\rho\|\sigma) + \log\frac{1}{\varepsilon} \leq D^{1-\varepsilon}_{H}(\rho\|\sigma) \leq D^{\varepsilon-\mu,T}_{\max}(\rho\|\sigma) + \log\frac{1}{\mu}, \tag{19}$$

$$D^{\sqrt{\varepsilon},P}_{\max}(\rho\|\sigma) + \log\frac{1}{\varepsilon} \leq D^{1-\varepsilon}_{H}(\rho\|\sigma) \leq D^{\varepsilon-\mu,P}_{\max}(\rho\|\sigma) + \log\frac{1}{\mu}. \tag{20}$$

For classical systems or commuting quantum states, a stronger trace distance bound holds:

$$D^{\varepsilon,T}_{\max}(p\|q) + \log\frac{1}{\varepsilon} \leq D^{1-\varepsilon}_{H}(p\|q) \leq D^{\varepsilon-\mu,T}_{\max}(p\|q) + \log\frac{1}{\mu}. \tag{21}$$

Here our focus is in particular on the leftmost inequalities in Eqs. (19) and (20), which are a significant improvement over the state-of-the-art bound of [20, Theorem 4]: the left-hand side of our bound is larger by an additive term of $\log\frac{1}{\varepsilon(1-\varepsilon)}$. The result improves also on bounds stated in [19, Proposition 4.1] and [18, Theorem 11], which had tighter error terms but worse smoothing terms than the bound of [20].

Once again, the bounds are in many ways tight. Even in the simplest case $\rho=\sigma$, previous results did not give tight bounds; in contrast, our bound is, to the best of our knowledge, the first statement of the weak/strong converse duality between D_H and $D_{\max}$ that gives tight error terms in this sense. The lower bound also gives a tight constraint on the asymptotic error exponent of $D^{\varepsilon,P}_{\max}$ [44] and on the exponent of $D^{\varepsilon,T}_{\max}$ for classical systems, as we will see in the next section.

We observe that the results have a mismatch of the order of ε in the upper and lower bounds: the lower bounds on $D^{1-\varepsilon}_{H}$ involve smooth max-relative entropy with smoothing of order $\sqrt{\varepsilon}$ (effectively due to the use of the gentle measurement lemma in Theorem 2), while the smoothing parameter in the upper bounds is of order ε. This scaling of the upper bound is tight for the trace distance, as can be seen from the classical case in (21). However, the upper bound for the purified distance in (20) can be tightened [20, Theorem 4] so that the order of the smoothing term matches the term $\sqrt{\varepsilon}$ in the lower bound in (20). This is a crucial property that allowed e.g. for the computation of the second-order expansion of the max-relative entropy for the purified distance [13] and the evaluation of the error and strong converse exponents for $D_{\max}^{\varepsilon,P}$ [44], [45]. The matching scaling of order ε in the classical case in (21) already tells us that the trace distance smoothing behaves differently. However, one could still ask: is it possible that we could instead improve the *lower* bound, e.g. by establishing that $D_{\max}^{\varepsilon,T} + \log\frac{1}{\varepsilon} \le D_H^{1-\varepsilon}$ holds for all quantum states? This is in fact impossible, as we now argue. To this end, consider two pure states $\psi = |\psi\rangle\langle\psi|$ and $\phi = |\phi\rangle\langle\phi|$ with trace distance $\varepsilon = \sqrt{1 - |\langle\psi|\phi\rangle|^2}$. It is not difficult to verify that the hypothesis testing relative entropy $D_H^{\delta}(\psi\|\phi)$ is infinite iff $\delta \ge 1-\varepsilon^2$, while $D_{\max}^{\delta,T}(\psi\|\phi)$ is infinite iff $\delta < \varepsilon$. Hence, any relation of the form

$$D_{\max}^{\delta,T}(\psi\|\phi) + g(\delta) \le D_H^{f(\delta)}(\psi\|\phi) \tag{22}$$

must be such that $f(\delta) \ge 1-\varepsilon^2$ for all $\delta < \varepsilon$. The best case scenario is therefore $f(\varepsilon) = 1-\varepsilon^2$, which precisely corresponds to the bound in (19). The fact that the choice of $g(\varepsilon) = \log\frac{1}{\varepsilon^2}$ is optimal in general can be verified by considering the trivial case $\psi = \phi$, where $D_{\max}^{\varepsilon,T}(\psi\|\psi) = 0$ but $D_H^{1-\varepsilon^2}(\psi\|\psi) = \log\frac{1}{\varepsilon^2}$.

B. Inequalities with Rényi relative entropies

The Petz–Rényi relative entropies D_α [46] and the sandwiched Rényi relative entropies $\widetilde{D}_\alpha$ [32], [47] are defined, respectively, as

$$D_\alpha(\rho\|\sigma) \coloneqq \frac{1}{\alpha-1}\log\operatorname{Tr}\left(\rho^\alpha\sigma^{1-\alpha}\right), \tag{23}$$

$$\widetilde{D}_\alpha(\rho\|\sigma) \coloneqq \frac{1}{\alpha-1}\log\operatorname{Tr}\left(\sigma^{\frac{1-\alpha}{2\alpha}}\rho\,\sigma^{\frac{1-\alpha}{2\alpha}}\right)^\alpha. \tag{24}$$

Both are additive under tensor products: $D_\alpha(\rho^{\otimes n}\|\sigma^{\otimes n}) = nD_\alpha(\rho\|\sigma)$ and $\widetilde{D}_\alpha(\rho^{\otimes n}\|\sigma^{\otimes n}) = n\widetilde{D}_\alpha(\rho\|\sigma)$.

Several bounds were given in the literature that connect smooth entropies with Rényi α divergences. Here we discuss how they can be improved using the relations established in this work.

First, we obtain upper bounds.

Corollary 5. *For all $\varepsilon \in (0,1)$ and all $\alpha > 1$, it holds that*

$$\widetilde{D}_{\max}^{\varepsilon}(\rho\|\sigma) + \log\frac{1}{1-\varepsilon} \le \widetilde{D}_\alpha(\rho\|\sigma) + \frac{1}{\alpha-1}\log\frac{1}{\varepsilon}, \tag{25}$$

As a result,

$$D_{\max}^{\varepsilon,T}(\rho\|\sigma) \le D_{\max}^{\varepsilon,P}(\rho\|\sigma) \le \widetilde{D}_\alpha(\rho\|\sigma) + \frac{1}{\alpha-1}\log\frac{1}{\varepsilon^2}. \tag{26}$$

The bound in (26) improves over previously known bounds in [20, Theorem 3] and [25, Proposition 6.22], losing a superfluous additive factor of $\log\frac{1}{1-\varepsilon^2}$ in the former and tightening the smoothing term in the latter.

The result relies on our Corollary 4 and standard argument based on the data processing of the Rényi divergences [48, Lemma IV.7]. The proof can be found in [31].

To investigate the tightness of the bounds, we will look at the error exponent of the quantity $\widetilde{D}_{\max}^{\varepsilon}$, that is, the largest exponent E such that $\widetilde{D}_{\max}^{2^{-nE+o(n)}}(\rho^{\otimes n}\|\sigma^{\otimes n}) = nR + o(n)$ for some fixed rate $R > 0$. Plugging $\varepsilon = 2^{-nE+o(n)}$ into (25) and dividing by n, we have in the limit $n \to \infty$ that

$$E \ge \sup_{\alpha>1}\,(\alpha-1)\left(R - \widetilde{D}_\alpha(\rho\|\sigma)\right). \tag{27}$$

This asymptotic bound is in fact known to be tight: this was established in [48, Theorem IV.4] as a key step in the derivation of the strong converse exponent of quantum hypothesis testing.

For the smooth max-relative entropy, (26) gives an asymptotically tight bound on the error exponent of $D_{\max}^{\varepsilon,P}$ [44]. We can also show [31] that the bound of (27) gives exactly the error exponent of $D_{\max}^{\varepsilon,T}$ when ρ and σ commute.

We can also give a lower bound.

Corollary 6. *For all $\alpha \in (0,1)$, it holds that*

$$D_{\max}^{\varepsilon,P}(\rho\|\sigma) \ge D_{\max}^{\varepsilon,T}(\rho\|\sigma) \ge \widetilde{D}_{\max}^{\varepsilon}(\rho\|\sigma) \tag{28}$$

$$\ge D_\alpha(\rho\|\sigma) - \frac{1}{1-\alpha}\log\frac{1}{1-\varepsilon}. \tag{29}$$

As a bound on the smooth max-relative entropy, this improves on the bound given in [21, Proposition 4]. The proof can be found in [31].

Once again, to study the tightness of the bound, let us look at exponents — now the strong converse exponent of $\widetilde{D}_{\max}^{\varepsilon}$, namely the least $E_{\rm sc}$ such that $\widetilde{D}_{\max}^{1-2^{-nE_{\rm sc}+o(n)}}(\rho^{\otimes n}\|\sigma^{\otimes n}) = nR + o(n)$ for some fixed $R > 0$. Corollary 6 gives a lower bound on this exponent as

$$E_{\rm sc} \ge \sup_{\alpha\in(0,1)}\,(\alpha-1)\left(R - D_\alpha(\rho\|\sigma)\right). \tag{30}$$

Using the known optimal error exponent of hypothesis testing [49]–[51], we can show that this bound is indeed tight. It is also tight for the strong converse exponent of $D_{\max}^{\varepsilon,T}$ when ρ and σ are commuting states [31] (see also [52]).

ACKNOWLEDGMENT

We acknowledge helpful discussions with Mario Berta and Mark M. Wilde. LL acknowledges support from MIUR (Ministero dell'Istruzione, dell'Università e della Ricerca) through the project 'Dipartimenti di Eccellenza 2023–2027' of the 'Classe di Scienze' department at the Scuola Normale Superiore.

References

[1] R. Renner, "Security of Quantum Key Distribution," 2005.

[2] L. Wang and R. Renner, "One-Shot Classical-Quantum Capacity and Hypothesis Testing," *Phys. Rev. Lett.*, vol. 108, p. 200501, 2012.

[3] F. Buscemi and N. Datta, "The Quantum Capacity of Channels With Arbitrarily Correlated Noise," *IEEE Trans. Inf. Theory*, vol. 56, pp. 1447–1460, 2010.

[4] D. Leung and W. Matthews, "On the Power of PPT-Preserving and Non-Signalling Codes," *IEEE Trans. Inf. Theory*, vol. 61, pp. 4486–4499, 2015.

[5] A. Anshu, R. Jain, and N. A. Warsi, "Building Blocks for Communication Over Noisy Quantum Networks," *IEEE Trans. Inf. Theory*, vol. 65, pp. 1287–1306, 2019.

[6] H.-C. Cheng, "Simple and Tighter Derivation of Achievability for Classical Communication Over Quantum Channels," *PRX Quantum*, vol. 4, p. 040330, 2023.

[7] J. M. Renes and R. Renner, "One-Shot Classical Data Compression with Quantum Side Information and the Distillation of Common Randomness or Secret Keys," *IEEE Trans. Inf. Theory*, vol. 58, pp. 1985–1991, 2012.

[8] F. G. S. L. Brandão and M. B. Plenio, "A Reversible Theory of Entanglement and its Relation to the Second Law," *Commun. Math. Phys.*, vol. 295, pp. 829–851, 2010.

[9] Z.-W. Liu, K. Bu, and R. Takagi, "One-Shot Operational Quantum Resource Theory," *Phys. Rev. Lett.*, vol. 123, p. 020401, 2019.

[10] B. Regula, K. Bu, R. Takagi, and Z.-W. Liu, "Benchmarking one-shot distillation in general quantum resource theories," *Phys. Rev. A*, vol. 101, p. 062315, 2020.

[11] M. Berta, M. Christandl, and R. Renner, "The Quantum Reverse Shannon Theorem Based on One-Shot Information Theory," *Commun. Math. Phys.*, vol. 306, p. 579, 2011.

[12] K. Fang, X. Wang, M. Tomamichel, and M. Berta, "Quantum Channel Simulation and the Channel's Smooth Max-Information," *IEEE Trans. Inf. Theory*, vol. 66, pp. 2129–2140, 2020.

[13] M. Tomamichel and M. Hayashi, "A Hierarchy of Information Quantities for Finite Block Length Analysis of Quantum Tasks," *IEEE Trans. Inf. Theory*, vol. 59, pp. 7693–7710, 2013.

[14] Y.-C. Shen, L. Gao, and H.-C. Cheng, "Optimal Second-Order Rates for Quantum Soft Covering and Privacy Amplification," *IEEE Trans. Inf. Theory*, vol. 70, pp. 5077–5091, 2024.

[15] F. Dupuis, M. Berta, J. Wullschleger, and R. Renner, "One-Shot Decoupling," *Commun. Math. Phys.*, vol. 328, pp. 251–284, 2014.

[16] A. Anshu, V. K. Devabathini, and R. Jain, "Quantum Communication Using Coherent Rejection Sampling," *Phys. Rev. Lett.*, vol. 119, p. 120506, 2017.

[17] N. Datta, "Min- and Max-Relative Entropies and a New Entanglement Monotone," *IEEE Trans. Inf. Theory*, vol. 55, pp. 2816–2826, 2009.

[18] N. Datta, M. Mosonyi, M. H. Hsieh, and F. G. S. L. Brandão, "A Smooth Entropy Approach to Quantum Hypothesis Testing and the Classical Capacity of Quantum Channels," *IEEE Trans. Inf. Theory*, vol. 59, pp. 8014–8026, 2013.

[19] F. Dupuis, L. Krämer, P. Faist, J. M. Renes, and R. Renner, "Generalized entropies," in *XVIIth International Congress on Mathematical Physics*. World Scientific, 2012, pp. 134–153.

[20] A. Anshu, M. Berta, R. Jain, and M. Tomamichel, "A minimax approach to one-shot entropy inequalities," *J. Math. Phys.*, vol. 60, p. 122201, 2019.

[21] X. Wang and M. M. Wilde, "Resource theory of asymmetric distinguishability," *Phys. Rev. Research*, vol. 1, p. 033170, 2019.

[22] M. Tomamichel, R. Colbeck, and R. Renner, "A Fully Quantum Asymptotic Equipartition Property," *IEEE Trans. Inf. Theory*, vol. 55, pp. 5840–5847, 2009.

[23] F. G. S. L. Brandão and M. B. Plenio, "A Generalization of Quantum Stein's Lemma," *Commun. Math. Phys.*, vol. 295, pp. 791–828, 2010.

[24] N. Datta, "Max-relative entropy of entanglement, alias log robustness," *Int. J. Quantum Inform.*, vol. 07, pp. 475–491, 2009.

[25] M. Tomamichel, *Quantum Information Processing with Finite Resources*. Springer, 2016.

[26] L. Lami, "A solution of the generalised quantum Stein's lemma," 2024.

[27] L. Lami, M. Berta, and B. Regula, "Asymptotic quantification of entanglement with a single copy," 2024.

[28] N. Datta and F. Leditzky, "Second-Order Asymptotics for Source Coding, Dense Coding, and Pure-State Entanglement Conversions," *IEEE Trans. Inf. Theory*, vol. 61, pp. 582–608, 2015.

[29] N. Datta and R. Renner, "Smooth Entropies and the Quantum Information Spectrum," *IEEE Trans. Inf. Theory*, vol. 55, pp. 2807–2815, 2009.

[30] L. Drescher and O. Fawzi, "On simultaneous min-entropy smoothing," in *2013 IEEE International Symposium on Information Theory (ISIT)*, 2013, pp. 161–165.

[31] B. Regula, L. Lami, and N. Datta, "Tight relations and equivalences between smooth relative entropies," 2025, arXiv preprint 2501.12447.

[32] M. Müller-Lennert, F. Dupuis, O. Szehr, S. Fehr, and M. Tomamichel, "On quantum Rényi entropies: A new generalization and some properties," *J. Math. Phys.*, vol. 54, p. 122203, 2013.

[33] C. Dwork, G. N. Rothblum, and S. Vadhan, "Boosting and Differential Privacy," in *2010 IEEE 51st Annual Symposium on Foundations of Computer Science*, 2010, pp. 51–60.

[34] C. Dwork and A. Roth, "The Algorithmic Foundations of Differential Privacy," *Found. Trends Theor. Comput. Sci.*, vol. 9, pp. 211–407, 2014.

[35] T. Ogawa and H. Nagaoka, "Strong converse and Stein's lemma in quantum hypothesis testing," *IEEE Trans. Inf. Theory*, vol. 46, pp. 2428–2433, 2000.

[36] H. Nagaoka and M. Hayashi, "An Information-Spectrum Approach to Classical and Quantum Hypothesis Testing for Simple Hypotheses," *IEEE Trans. Inf. Theory*, vol. 53, pp. 534–549, 2007.

[37] L. Zhou and M. Ying, "Differential Privacy in Quantum Computation," in *2017 IEEE 30th Computer Security Foundations Symposium (CSF)*, 2017, pp. 249–262.

[38] C. Hirche, C. Rouzé, and D. S. França, "Quantum Differential Privacy: An Information Theory Perspective," *IEEE Trans. Inf. Theory*, vol. 69, pp. 5771–5787, 2023.

[39] T. Nuradha, Z. Goldfeld, and M. M. Wilde, "Quantum Pufferfish Privacy: A Flexible Privacy Framework for Quantum Systems," *IEEE Trans. Inf. Theory*, vol. 70, pp. 5731–5762, 2024.

[40] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York: Cambridge University Press, 2004.

[41] A. Winter, "Coding theorem and strong converse for quantum channels," *IEEE Trans. Inf. Theory*, vol. 45, pp. 2481–2485, 1999.

[42] R. Bhatia, *Positive Definite Matrices*. Princeton University Press, 2007.

[43] T. Ogawa and H. Nagaoka, "Making Good Codes for Classical-Quantum Channel Coding via Quantum Hypothesis Testing," *IEEE Trans. Inf. Theory*, vol. 53, pp. 2261–2266, 2007.

[44] K. Li, Y. Yao, and M. Hayashi, "Tight Exponential Analysis for Smoothing the Max-Relative Entropy and for Quantum Privacy Amplification," *IEEE Trans. Inf. Theory*, vol. 69, pp. 1680–1694, 2023.

[45] K. Li and Y. Yao, "Operational Interpretation of the Sandwiched Rényi Divergence of Order 1/2 to 1 as Strong Converse Exponents," *Commun. Math. Phys.*, vol. 405, p. 22, 2024.

[46] D. Petz, "Quasi-entropies for finite quantum systems," *Rep. Math. Phys.*, vol. 23, pp. 57–65, 1986.

[47] M. M. Wilde, A. Winter, and D. Yang, "Strong Converse for the Classical Capacity of Entanglement-Breaking and Hadamard Channels via a Sandwiched Rényi Relative Entropy," *Commun. Math. Phys.*, vol. 331, pp. 593–622, 2014.

[48] M. Mosonyi and T. Ogawa, "Quantum Hypothesis Testing and the Operational Interpretation of the Quantum Rényi Relative Entropies," *Commun. Math. Phys.*, vol. 334, pp. 1617–1648, 2015.

[49] K. M. R. Audenaert, J. Calsamiglia, R. Muñoz-Tapia, E. Bagan, Ll. Masanes, A. Acin, and F. Verstraete, "Discriminating States: The Quantum Chernoff Bound," *Phys. Rev. Lett.*, vol. 98, p. 160501, 2007.

[50] M. Hayashi, "Error exponent in asymmetric quantum hypothesis testing and its application to classical-quantum channel coding," *Phys. Rev. A*, vol. 76, p. 062301, 2007.

[51] H. Nagaoka, "The Converse Part of The Theorem for Quantum Hoeffding Bound," 2006.

[52] R. Salzmann and N. Datta, "Total insecurity of communication via strong converse for quantum privacy amplification," 2022.