Smooth Rényi Entropy of Ergodic Quantum Information Sources

Schoenmakers, Berry; Tjoelker, Jilles; Tuyls, Pim; Verbitskiy, Evgeny

Published in:
EPRINTS-BOOK-TITLE

Publication date:
2007

Smooth Rényi Entropy of Ergodic Quantum Information Sources

Berry Schoenmakers Jilles Tjoelker
Dept. of Mathematics and Computer Science
Technical University Eindhoven
The Netherlands
berry@win.tue.nl j.tjoelker@student.tue.nl

Pim Tuyls
Information Security Systems
Philips Research Eindhoven
The Netherlands
pim.tuyls@philips.com

Evgeny Verbitskiy
Digital Signal Processing
Philips Research Eindhoven
The Netherlands
evgeny.verbitskiy@philips.com

***Abstract*— We investigate the recently introduced notion of smooth Rényi entropy for the case of *ergodic* information sources, thereby generalizing previous work which concentrated mainly on i.i.d. information sources. We will actually consider ergodic *quantum* information sources, of which ergodic classical information sources are a special case. We prove that the average smooth Rényi entropy rate will approach the entropy rate of a stationary, ergodic source, which is equal to the Shannon entropy rate for a classical source and the von Neumann entropy rate for a quantum source.**

I. Introduction

The elegant notion of smooth Rényi entropy was introduced recently by Renner and Wolf in [6] for classical information sources, and the natural extension to quantum information sources was defined by Renner and König in [5]. In these two papers and further work by Renner and Wolf [7], [4], many properties of smooth Rényi entropy—and smooth min-entropy and smooth max-entropy in particular—have been studied in detail.

A central property of smooth Rényi entropy proved in these works is that for memoryless (i.i.d.) information sources, the average smooth Rényi entropy rate will approach the entropy rate of the source, which is equal to the Shannon entropy for a classical source and the von Neumann entropy for a quantum source. Whereas, in general, the average (conventional) Rényi entropy rate of a memoryless source does not converge to the source's entropy rate.

In this paper we extend the study of smooth Rényi entropy to the more general class of stationary, ergodic sources rather than memoryless sources. We will prove that for both the classical and the quantum case that the average smooth Rényi entropy rate will approach the Shannon and the von Neumann entropy rate, respectively. We will do so by first treating the classical case and then reducing the quantum case to the classical one without losing generality.

In general, smooth Rényi entropy of order $\alpha > 1$, and $\alpha = \infty$ (min-entropy) in particular, is of cryptographic relevance (e.g., for randomness-extraction), and smooth Rényi entropy of order $\alpha < 1$, and $\alpha = 0$ (max-entropy) in particular, are relevant to data compression (minimum encoding length). In these contexts, the importance of *smooth* Rényi entropy is that its rate is basically equal to the Shannon/von Neumann entropy rate for an i.i.d. source (and for ergodic sources as well, as we show in this paper). This is not the case for conventional Rényi entropy. More generally, as shown in the papers by Renner *et al.* mentioned above, smooth Rényi entropy behaves much as Shannon/von Neumann entropy does.

In this paper we focus on the unconditional case, whereas much of the abovementioned work by Renner *et al.* treats the more general conditional case. We leave the extension to the conditional case for future work. However, we do consider two notions of ϵ-closeness, one based on trace distance (also known as variational or statistical distance) and one based on non-normalized density matrices (or probability distributions), where the latter is more suitable to handle the conditional case.[1] Thus, we believe that our results can be extended to the conditional case as well.

We also note that Renner [4] presents a different kind of generalization of i.i.d. quantum sources, namely by analyzing the smooth min-entropy of symmetric (permutation-invariant) quantum states. Or, more precisely, states in a symmetric subspace of $\mathcal{H}^{\otimes n}$ are considered, for $n \in \mathbb{N}$. See [4, Chapter 4] for details, which also covers the conditional case.

II. Preliminaries

Throughout this paper we use $\mathbb{P}$ and $\mathbb{Q}$ to denote probability distributions with over the same finite or countably infinite range $\mathcal{Z}$. Similarly, we use ρ and σ to denote density matrices on the same Hilbert space of a finite or countably infinite dimension. These probability distributions and density matrices are not necessarily normalized (e.g., $\sum_z \mathbb{P}(z) < 1$ if $\mathbb{P}$ is non-normalized and $\mathrm{tr}(\rho) < 1$ if ρ is non-normalized).

For ease of comparison we state all the preliminaries explicitly for the classical case as well as for the quantum case.

Definition 1 (Classical Rényi entropy): The Rényi entropy of order $\alpha \in [0, \infty]$ of probability distribution $\mathbb{P}$ is

$$H_\alpha(\mathbb{P}) = \frac{1}{1-\alpha} \log \sum_{z \in \mathcal{Z}} \mathbb{P}(z)^\alpha,$$

[1] The trace distance was originally used in [6], [5]. The use of non-normalized probability distributions was also shown in the full version of [6] and used in [7]. In this paper, we extend this to the use of non-normalized density matrices in the quantum case.

1-4244-1429-6/07/$25.00 ©2007 IEEE

for $0 < \alpha < \infty$, $\alpha \neq 1$, and $H_\alpha(\mathbb{P}) = \lim_{\beta\to\alpha} H_\beta(\mathbb{P})$ otherwise.
Hence, $H_0(\mathbb{P}) = \log|\{z \in \mathcal{Z} : \mathbb{P}(z) > 0\}|$, $H_1(\mathbb{P}) = H(\mathbb{P})$ (Shannon entropy) and $H_\infty(\mathbb{P}) = -\log\max_{z\in\mathcal{Z}} \mathbb{P}(z)$.

For a random variable Z we use $H_\alpha(Z)$ as a shorthand for $H_\alpha(\mathbb{P}_Z)$, where $\mathbb{P}_Z$ is the probability distribution of Z.

Smooth Rényi entropy was introduced in [6] for the classical case. For $\epsilon \geq 0$, let $\mathcal{B}^\epsilon(\mathbb{P})$ denote either the set of probability distributions which are ϵ-close to $\mathbb{P}$, $\mathcal{B}^\epsilon(\mathbb{P}) = \{\mathbb{Q} : \delta(\mathbb{P},\mathbb{Q}) \leq \epsilon\}$, or the set of non-normalized probability distributions which are ϵ-close to $\mathbb{P}$, $\mathcal{B}^\epsilon(\mathbb{P}) = \{\mathbb{Q} : \sum_{z\in\mathcal{Z}} \mathbb{Q}(z) \geq 1-\epsilon, \forall_{z\in\mathcal{Z}} 0 \leq \mathbb{Q}(z) \leq \mathbb{P}(z)\}$. The first notion of ϵ-closeness, based on the statistical distance $\delta(\mathbb{P},\mathbb{Q}) = \frac{1}{2}\sum_{z\in\mathcal{Z}} |\mathbb{P}(z) - \mathbb{Q}(z)|$, was used in [6]. The second notion was mentioned in the full version of [6], and used in [7].

Definition 2 (Classical smooth Rényi entropy, [6]): The ϵ-smooth Rényi entropy of order $\alpha \in [0,1) \cup (1,\infty]$ of a probability distribution $\mathbb{P}$ is

$$H_\alpha^\epsilon(\mathbb{P}) = \begin{cases} \inf_{\mathbb{Q}\in\mathcal{B}^\epsilon(\mathbb{P})} H_\alpha(\mathbb{Q}), & 0 \leq \alpha < 1, \\ \sup_{\mathbb{Q}\in\mathcal{B}^\epsilon(\mathbb{P})} H_\alpha(\mathbb{Q}), & 1 < \alpha \leq \infty. \end{cases}$$

At the end of this paper, we point out that $H_\alpha^\epsilon(\mathbb{P})$ will actually vary, depending on which notion of ϵ-closeness is used, leading to a maximum difference of $\frac{\alpha}{\alpha-1}\log(1-\epsilon)$.

For a probability distribution $\mathbb{P}$ on, e.g., $\mathcal{Z} = \{0,1\}^{\mathbb{N}}$, we define $\mathbb{P}^n$ as the probability distribution corresponding to the restriction of the "infinite volume" distribution $\mathbb{P}$ to the finite volume $\{0,\ldots,n-1\}$.

Definition 3 (Entropy rate of a classical source): For a stationary source given by its probability measure $\mathbb{P}$, we define

$$\begin{aligned} h(\mathbb{P}) &= \lim_{n\to\infty} \frac{1}{n} H(\mathbb{P}^n), \\ h_\alpha^\epsilon(\mathbb{P}) &= \lim_{n\to\infty} \frac{1}{n} H_\alpha^\epsilon(\mathbb{P}^n). \end{aligned}$$

We will actually prove that $h_\alpha^\epsilon(\mathbb{P}) = h(\mathbb{P})$ as $\epsilon \to 0$.

We use the standard notion of typical sequences and typical sets, which are defined for any information source (not necessarily i.i.d.). See, for instance, [2] or [3].

Definition 4 (Typical sequences, typical set): A sequence $z^n \in \{0,1\}^n$, $n \in \mathbb{N}$, is called ϵ-typical if

$$e^{-n(h(\mathbb{P})+\epsilon)} \leq \mathbb{P}(z^n) \leq e^{-n(h(\mathbb{P})-\epsilon)}.$$

The typical set $T_\epsilon^{(n)}$ is the set of all ϵ-typical sequences from $\{0,1\}^n$.

In this paper we need the following consequence of the AEP, where we refer to [2, Section 16.8] for the AEP for ergodic sources (known as the Shannon-McMillan-Breiman theorem).

Theorem 1 (Classical AEP bounds): Let $\mathbb{P}$ be a stationary, ergodic probability distribution on $\mathcal{Z} = \{0,1\}^{\mathbb{N}}$. Let $\epsilon > 0$. Then, for sufficiently large n,

$$\mathbb{P}(T_\epsilon^{(n)}) \geq 1-\epsilon,$$

and

$$|T_\epsilon^{(n)}| \leq e^{n(h(\mathbb{P})+\epsilon)}.$$

Definition 5 (Quantum Rényi entropy): The Rényi entropy of order $\alpha \in [0,\infty]$ of a density matrix ρ is

$$S_\alpha(\rho) = \frac{1}{1-\alpha}\log\mathrm{tr}(\rho^\alpha)$$

for $0 < \alpha < \infty$, $\alpha \neq 1$, and $S_\alpha(\rho) = \lim_{\beta\to\alpha} S_\beta(\rho)$ otherwise.
Hence, $S_0(\rho) = \log\mathrm{rank}(\rho)$, $S_1(\rho) = S(\rho) = -\mathrm{tr}(\rho\log\rho)$ (von Neumann entropy) and $S_\infty(\rho) = -\log\lambda_{max}(\rho)$.

Analogous to the classical case, smooth Rényi entropy is defined in the quantum case (see [5]). We use either the set of density matrices which are ϵ-close to ρ, $\mathcal{B}^\epsilon(\rho) = \{\sigma : \delta(\rho,\sigma) \leq \epsilon\}$ or the set of non-normalized density matrices which are ϵ-close to ρ, $\mathcal{B}^\epsilon(\rho) = \{\sigma : \mathrm{tr}(\sigma) \geq 1-\epsilon, 0 \leq \sigma \leq \rho\}$. The first notion of ϵ-closeness, based on the trace distance $\delta(\rho,\sigma) = \frac{1}{2}\mathrm{tr}(|\rho-\sigma|)$, was used in [5]. The second notion is introduced here, and will actually be used in the next section.

Definition 6 (Quantum smooth Rényi entropy, [5]): The ϵ-smooth Rényi entropy of order $\alpha \in [0,1)\cup(1,\infty]$ of a density matrix ρ is

$$S_\alpha^\epsilon(\rho) = \begin{cases} \inf_{\sigma\in\mathcal{B}^\epsilon(\rho)} S_\alpha(\sigma), & 0 \leq \alpha < 1, \\ \sup_{\sigma\in\mathcal{B}^\epsilon(\rho)} S_\alpha(\sigma), & 1 < \alpha \leq \infty. \end{cases}$$

Definition 7 (Entropy rates of a quantum source): For a stationary quantum source ρ, given by its local densities $\rho^{(n)} = \rho_{0,\ldots,n-1}$, for $n \in \mathbb{N}$, we define:

$$\begin{aligned} s(\rho) &= \lim_{n\to\infty} \frac{1}{n} S(\rho^{(n)}), \\ s_\alpha^\epsilon(\rho) &= \lim_{n\to\infty} \frac{1}{n} S_\alpha^\epsilon(\rho^{(n)}). \end{aligned}$$

We use the following notion of typical states and typical subspaces, as can be found in [1] (also see [3]).

Definition 8 (Typical state, typical subspace): A pure state $|e_i^{(n)}\rangle$, where $e_i^{(n)}$ is an eigenvector of $\rho^{(n)}$ is called ϵ-typical if the corresponding eigenvalue $\lambda_i^{(n)}$ satisfies

$$e^{-n(s(\rho)+\epsilon)} \leq \lambda_i^{(n)} \leq e^{-n(s(\rho)-\epsilon)}.$$

The typical subspace $T_\epsilon^{(n)}$ is the subspace spanned by all ϵ-typical states.

We will need the following consequences of the quantum AEP for ergodic sources, which has been studied in [1] (see [3] for the quantum AEP for i.i.d. sources).

Theorem 2 (Quantum AEP bounds): Let ρ be a stationary, ergodic quantum source with local densities $\rho^{(n)}$. Let $\epsilon > 0$. Then, for sufficiently large n,

$$\mathrm{tr}(\rho^{(n)} P_{T_\epsilon^{(n)}}) \geq 1-\epsilon,$$

where $P_{T_\epsilon^{(n)}}$ is the projector onto the subspace $T_\epsilon^{(n)}$. Furthermore,

$$\mathrm{tr}(P_{T_\epsilon^{(n)}}) \leq e^{n(s(\rho)+\epsilon)}.$$

Clearly, the quantum AEP for ergodic sources implies the classical AEP for ergodic sources.

The following theorem by Renner and Wolf states that smooth Rényi entropy approaches Shannon entropy in the case of a classical i.i.d. source.

Theorem 3 ([7, Lemma I.2]): Let Z^n denote an n-tuple of i.i.d. random variables with probability distribution $\mathbb{P}_Z$. Then,

$$\lim_{\epsilon\to 0}\lim_{n\to\infty}\frac{1}{n}H_\alpha^\epsilon(Z^n)=H(Z),$$

for any $\alpha\in[0,\infty]$.

The analogous theorem by Renner and König for a quantum i.i.d. source is as follows.

Theorem 4 ([5, Lemma 3]): Let ρ be a density matrix. Then,

$$\lim_{\epsilon\to 0}\lim_{n\to\infty}\frac{1}{n}S_\alpha^\epsilon(\rho^{\otimes n})=S(\rho),$$

for any $\alpha\in[0,\infty]$.

III. Main Result

We extend the results by Renner and Wolf (Theorem 3 above) and by Renner and König (Theorem 4 above) to the case of ergodic sources. Throughout this section, we use the notion of ϵ-closeness based on non-normalized probability distributions and density matrices, so $\mathcal{B}^\epsilon(\mathbb{P})=\{\mathbb{Q}:\sum_{z\in\mathcal{Z}}\mathbb{Q}(z)\geq 1-\epsilon,\forall_{z\in\mathcal{Z}}0\leq\mathbb{Q}(z)\leq\mathbb{P}(z)\}$ and $\mathcal{B}^\epsilon(\rho)=\{\sigma:\mathrm{tr}(\sigma)\geq 1-\epsilon,0\leq\sigma\leq\rho\}$, respectively. In the next section, we will argue that the results are independent on which notion of ϵ-closeness is used.

A. Classical Case

We start with our main result for the classical case. The known result for an i.i.d. source is by Renner and Wolf, Theorem 3 above. We will extend this to a stationary, ergodic source in Theorem 5 below.

Lemma 1: Let $\mathbb{P}$ be a stationary, ergodic information source given by its probability measure and let $0<\epsilon<1/2$. Then we have,

$$h(\mathbb{P})-\epsilon\leq h_\infty^\epsilon(\mathbb{P})\leq h(\mathbb{P})+2\epsilon.$$

Proof: Let $0<\epsilon<1/2$. To prove the lower bound, we show that, for sufficiently large n, $H_\infty^\epsilon(\mathbb{P}^n)\geq n(h(\mathbb{P})-\epsilon)$. Define non-normalized probability distribution $\mathbb{Q}$ for all $z^n\in\{0,1\}^n$ by

$$\mathbb{Q}(z^n)=\begin{cases}\mathbb{P}(z^n), & \text{if } z^n\in\mathcal{T}_\epsilon^{(n)}\\ 0, & \text{if } z^n\notin\mathcal{T}_\epsilon^{(n)}.\end{cases}\tag{1}$$

Clearly, $0\leq\mathbb{Q}(z^n)\leq\mathbb{P}(z^n)$ and, by the AEP, $\mathbb{Q}(\mathcal{T}_\epsilon^{(n)})=\mathbb{P}(\mathcal{T}_\epsilon^{(n)})\geq 1-\epsilon$ for sufficiently large n. So, $\mathbb{Q}\in\mathcal{B}^\epsilon(\mathbb{P}^n)$. Furthermore, for $z^n\in\mathcal{T}_\epsilon^{(n)}$, we have that $-\log\mathbb{P}(z^n)\geq n(h(\mathbb{P})-\epsilon)$, and hence that for any z^n that $-\log\mathbb{Q}(z^n)\geq n(h(\mathbb{P})-\epsilon)$. This implies that $H_\infty(\mathbb{Q})=-\log\max_{z^n}\mathbb{Q}(z^n)\geq n(h(\mathbb{P})-\epsilon)$ and the lower bound follows.

Next, to prove the upper bound, we show that, for sufficiently large n, one has that for all $\mathbb{Q}\in\mathcal{B}^\epsilon(\mathbb{P}^n)$,

$$H_\infty(\mathbb{Q})=-\log\max_{z^n}\mathbb{Q}(z^n)\leq n(h(\mathbb{P})+2\epsilon).$$

This follows from $\max_{z^n\in\mathcal{T}_\epsilon^{(n)}}\mathbb{Q}(z^n)\geq e^{-n(h(\mathbb{P})+2\epsilon)}$, which in turn follows from $\sum_{z^n\in\mathcal{T}_\epsilon^{(n)}}\mathbb{Q}(z^n)\geq|\mathcal{T}_\epsilon^{(n)}|e^{-n(h(\mathbb{P})+2\epsilon)}$. From the AEP we get $|\mathcal{T}_\epsilon^{(n)}|\leq e^{n(h(\mathbb{P})+\epsilon)}$, hence it suffices to prove that, for sufficiently large n,

$$\sum_{z^n\in\mathcal{T}_\epsilon^{(n)}}\mathbb{Q}(z^n)\geq e^{-n\epsilon}.\tag{2}$$

As $\sum_{z^n}\mathbb{Q}(z^n)\geq 1-\epsilon$, for $\mathbb{Q}\in\mathcal{B}^\epsilon(\mathbb{P}^n)$, and also $\sum_{z^n\notin\mathcal{T}_\epsilon^{(n)}}\mathbb{Q}(z^n)\leq\epsilon$ (because $\mathbb{Q}(z^n)\leq\mathbb{P}(z^n)$ and $P(\mathcal{T}_\epsilon^{(n)})\geq 1-\epsilon$ from the AEP), we only need to observe that

$$1-2\epsilon>e^{-n\epsilon}$$

holds for sufficiently large n, using that $\epsilon<1/2$. ∎

We now state an analogous lemma for the *max*-entropy.

Lemma 2: Let $\mathbb{P}$ be a stationary, ergodic information source given by its probability measure and let $0<\epsilon<1/2$. Then we have,

$$h(\mathbb{P})-2\epsilon\leq h_0^\epsilon(\mathbb{P})\leq h(\mathbb{P})+\epsilon.$$

Proof: Let $0<\epsilon<1/2$. To prove the upper bound, we show that, for sufficiently large n, $H_0^\epsilon(\mathbb{P}^n)\leq n(h(\mathbb{P})+\epsilon)$. We do so by showing that $H_0(\mathbb{Q})=\log|\{z^n:\mathbb{Q}(z^n)>0\}|\leq n(h(\mathbb{P})+\epsilon)$ for the non-normalized probability distribution $\mathbb{Q}$, defined by (1) in the proof of Lemma 1. As

$$|\{z^n:\mathbb{Q}(z^n)>0\}|=|\{z^n\in\mathcal{T}_\epsilon^{(n)}:\mathbb{Q}(z^n)>0\}|\leq|\mathcal{T}_\epsilon^{(n)}|,$$

the result follows directly from the AEP.

Next, to prove the lower bound, we show that, for sufficiently large n, one has that for all $\mathbb{Q}\in\mathcal{B}^\epsilon(\mathbb{P}^n)$,

$$H_0(\mathbb{Q})=\log|\{z^n:\mathbb{Q}(z^n)>0\}|\geq n(h(\mathbb{P})-2\epsilon).$$

This is implied by $|\{z^n\in\mathcal{T}_\epsilon^{(n)}:\mathbb{Q}(z^n)>0\}|\geq e^{n(h(\mathbb{P})-2\epsilon)}$, which is in turn implied by

$$\sum_{z^n\in\mathcal{T}_\epsilon^{(n)}}\mathbb{Q}(z^n)\geq\max_{z^n\in\mathcal{T}_\epsilon^{(n)},\mathbb{Q}(z^n)>0}\mathbb{Q}(z^n)e^{n(h(\mathbb{P})-2\epsilon)}.$$

Using inequality (2) from the proof of Lemma 1, it suffices to show that

$$\max_{z^n\in\mathcal{T}_\epsilon^{(n)},\mathbb{Q}(z^n)>0}\mathbb{Q}(z^n)\leq e^{-n\epsilon}e^{-n(h(\mathbb{P})-2\epsilon)}=e^{-n(h(\mathbb{P})-\epsilon)}.$$

This is a direct consequence of the definition of ϵ-typical sequences, as $\mathbb{Q}(z^n)\leq\mathbb{P}(z^n)$ for $\mathbb{Q}\in\mathcal{B}^\epsilon(\mathbb{P}^n)$, using that $\mathbb{Q}(z^n)>0$ holds for at least one $z^n\in\mathcal{T}_\epsilon^{(n)}$ on account of inequality (2). ∎

Theorem 5: For $\alpha\in[0,\infty]$, the ϵ-smooth entropy of a stationary, ergodic information source $\mathbb{P}$ given by its probability measure on $\mathcal{Z}=\{0,1\}^{\mathbb{N}}$ is close to the mean Shannon entropy:

$$\lim_{\epsilon\to 0}h_\alpha^\epsilon(\mathbb{P})=h(\mathbb{P}).$$

Proof: For $\alpha<1$, the monotonicity of smooth Rényi entropy (see, e.g., [7, Lemma 1]) yields $H_\alpha^\epsilon(\mathbb{P}^n)\leq H_0^\epsilon(\mathbb{P}^n)$, and hence $h_\alpha^\epsilon(\mathbb{P})\leq h(\mathbb{P})+\epsilon$ by Lemma 2.

To get a lower bound for $h_\alpha^\epsilon(\mathbb{P})$, we note that

$$H_\alpha^\epsilon(\mathbb{P}^n)\geq H_0^{2\epsilon}(\mathbb{P}^n)-\frac{\log(1/\epsilon)}{1-\alpha},$$

using [7, Lemma 2]. So, $h_\alpha^\epsilon(\mathbb{P}) \geq h_0^{2\epsilon}(\mathbb{P})$ as the constant term on the right-hand side vanishes for $n \to \infty$. Using Lemma 2, we thus get $h_\alpha^\epsilon(\mathbb{P}) \geq h(\mathbb{P}) - 4\epsilon$.

This proves that $\lim_{\epsilon \to 0} h_\alpha^\epsilon(\mathbb{P}) = h(\mathbb{P})$. The proof for $\alpha > 1$ is completely symmetrical, hence omitted. ■

Note that the term 2ϵ in the upper and lower bounds of Lemmas 1 and 2, respectively, can be improved to $(1+\delta)\epsilon$ for any constant $\delta > 0$. Similarly, the term 4ϵ in the proof of Theorem 5 for the lower bound for h_α^ϵ can be improved to $(1+\delta)\epsilon$ for any constant $\delta > 0$.

B. Quantum Case

Although it is possible to prove the quantum case directly, along the same lines as in the classical case, we treat the quantum case indirectly, by reducing it to the classical case. This leads to a more compact proof. To this end, we will first prove Lemma 3 below, which captures the correspondence between $\mathcal{B}^\epsilon(\rho^{(n)})$ and $\mathcal{B}^\epsilon(\lambda^{(n)})$. We only consider the case of ϵ-closeness for non-normalized density matrices and probability distributions (but the lemma also holds for the case of ϵ-closeness based on trace distance).

To prove our lemma, we need Weyl's monotonicity principle which we recall first.

Theorem 6 (Weyl monotonicity): If A, B are m by m Hermitian matrices and B is positive, then $\lambda_i(A) \leq \lambda_i(A+B)$ for all $i = 1, \ldots, m$, where $\lambda_i(M)$ is the i-th eigenvalue of M (ordered from largest to smallest).

Lemma 3: Let ρ be a density matrix with eigenvalues $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_m$.

1) For any density matrix σ with eigenvalues $\mu_1 \geq \mu_2 \geq \ldots \geq \mu_m$,
$$\sigma \in \mathcal{B}^\epsilon(\rho) \Rightarrow \mu \in \mathcal{B}^\epsilon(\lambda).$$
2) Given real numbers $\mu_1, \ldots, \mu_m$ such that $\mu \in \mathcal{B}^\epsilon(\lambda)$, there exists a matrix σ with eigenvalues $\mu_1, \ldots, \mu_m$ such that $\sigma \in \mathcal{B}^\epsilon(\rho)$.

Proof: We prove the result for
$$\begin{aligned} \mathcal{B}^\epsilon(\lambda) &= \{\mu : \sum_i \mu_i \geq 1-\epsilon, \forall_i 0 \leq \mu_i \leq \lambda_i\}, \\ \mathcal{B}^\epsilon(\rho) &= \{\sigma : \mathrm{tr}(\sigma) \geq 1-\epsilon, 0 \leq \sigma \leq \rho\}. \end{aligned}$$
For the first part, let σ be a (possibly non-normalized) density matrix with eigenvalues $\mu_1 \geq \mu_2 \geq \ldots \geq \mu_m$ and suppose $\sigma \in \mathcal{B}^\epsilon(\rho)$. Since σ is positive we have $\mu_i \geq 0$ for all i. And since $\sigma \leq \rho$, we have that $\rho - \sigma$ is positive as well, so $\lambda_i \geq \mu_i$ for all i (using Weyl's monotonicity principle, Theorem 6 above). Finally, note that $\mathrm{tr}(\sigma) \geq 1-\epsilon$ is equivalent to $\sum_i \mu_i \geq 1-\epsilon$, so we conclude that $\mu \in \mathcal{B}^\epsilon(\lambda)$.

For the second part, let $\mu \in \mathcal{B}^\epsilon(\lambda)$ be given. We write the Hermitian matrix ρ in diagonal form,
$$\rho = \sum_i \lambda_i |v_i\rangle\langle v_i|.$$
for eigenvectors v_i ($i = 1, \ldots, m$), and we show that the Hermitian matrix σ, defined by
$$\sigma = \sum_i \mu_i |v_i\rangle\langle v_i|,$$
is in $\mathcal{B}^\epsilon(\rho)$.

Since $\mu \in \mathcal{B}^\epsilon(\lambda)$, we have that $0 \leq \mu_i \leq \lambda_i$, and because ρ and σ commute (eigenvalues of $\rho - \sigma$ are $\lambda_i - \mu_i$), we have $0 \leq \sigma \leq \rho$. Clearly, $\sum_i \mu_i \geq 1-\epsilon$ so $\mathrm{tr}(\sigma) \geq 1-\epsilon$ as well, and therefore $\sigma \in \mathcal{B}^\epsilon(\rho)$. ■

We now proceed to prove the main result for the quantum case.

Theorem 7: For $\alpha \in [0, \infty]$, the ϵ-smooth entropy of a stationary, ergodic quantum source ρ given by its local densities $\rho^{(n)}$, for $n \in \mathbb{N}$, is close to the mean von Neumann entropy:
$$\lim_{\epsilon \to 0} s_\alpha^\epsilon(\rho) = s(\rho).$$
Proof: We will apply Theorem 5 as follows.

First note that for the local densities $\rho^{(n)}$ for a quantum information source ρ, we have that $S(\rho^{(n)}) = H(\lambda^{(n)})$, where $\lambda^{(n)}$ denotes the probability distribution corresponding to the eigenvalues of $\rho^{(n)}$. Consequently, $s(\rho) = h(\lambda)$ as well, where λ denotes the probability distribution corresponding to the eigenvalues of ρ.

Next, we recall the definitions of smooth Rényi entropy in the classical and quantum case, resp.:
$$\begin{aligned} H_\alpha^\epsilon(\mathbb{P}) &= \begin{cases} \inf_{\mathbb{Q} \in \mathcal{B}^\epsilon(\mathbb{P})} H_\alpha(\mathbb{Q}), & 0 \leq \alpha < 1, \\ \sup_{\mathbb{Q} \in \mathcal{B}^\epsilon(\mathbb{P})} H_\alpha(\mathbb{Q}), & 1 < \alpha \leq \infty. \end{cases} \\ S_\alpha^\epsilon(\rho) &= \begin{cases} \inf_{\sigma \in \mathcal{B}^\epsilon(\rho)} S_\alpha(\sigma), & 0 \leq \alpha < 1, \\ \sup_{\sigma \in \mathcal{B}^\epsilon(\rho)} S_\alpha(\sigma), & 1 < \alpha \leq \infty. \end{cases} \end{aligned}$$
We only consider the case $\alpha < 1$, as the other case follows by symmetry. We have that
$$\begin{aligned} S_\alpha^\epsilon(\rho^{(n)}) &= \inf_{\sigma \in \mathcal{B}^\epsilon(\rho^{(n)})} S_\alpha(\sigma) \\ &= \inf_{\mu \in \mathcal{B}^\epsilon(\lambda^{(n)})} H_\alpha(\mu) \\ &= H_\alpha^\epsilon(\lambda^{(n)}), \end{aligned}$$
using that Lemma 3 implies that the infimum over $\mathcal{B}^\epsilon(\rho^{(n)})$ is equal to the infimum over $\mathcal{B}^\epsilon(\lambda^{(n)})$.

As a consequence, we have that $s_\alpha^\epsilon(\rho) = h_\alpha^\epsilon(\lambda)$ and the result follows from Theorem 5. Here, we use the fact that quantum AEP implies classical AEP. ■

We note that the actual convergence rate (as a function of ϵ) is the same as in the classical case, which follows by considering the analogons of Lemmas 1 and 2.

IV. Notions of ϵ-Closeness

As mentioned in the introduction, two notions of ϵ-closeness were originally introduced by Renner and Wolf [6], [7], which can both be used in the definition of classical smooth Rényi entropy. For the quantum case, the paper by Renner and König [5] only considers the notion of ϵ-closeness based on the trace distance. As the natural quantum analogon of the notion of ϵ-closeness based on non-normalized probability distributions, we have used the set of non-normalized density matrices which are ϵ-close to a given density matrix ρ:
$$\mathcal{B}^\epsilon(\rho) = \{\sigma : \mathrm{tr}(\sigma) \geq 1-\epsilon, 0 \leq \sigma \leq \rho\}.$$

The entropy rates (Definitions 3 and 7), and consequently the results for these entropy rates (Theorems 3, 4, 5, and 7) do not depend on which of these notions of ϵ-closeness is used.

Furthermore, if the corresponding notions of ϵ-closeness are used, the quantum case and the classical case are in general connected as follows:

$$S_\alpha^\epsilon(\rho) = \inf_{\sigma \in \mathcal{B}^\epsilon(\rho)} S_\alpha(\sigma) = \inf_{\mu \in \mathcal{B}^\epsilon(\lambda)} H_\alpha(\mu) = H_\alpha^\epsilon(\lambda),$$

where λ denotes the probability distribution corresponding to the eigenvalues of ρ.

We note, however, that the smooth Rényi entropy H_α^ϵ may depend on which notion of ϵ-closeness is used, contrary to what was stated before (see, e.g., Section 3.3 of the full version of [6]). In general, one can show that

$$0 \leq \inf_{\delta(\mathbb{P},\mathbb{Q}) \leq \epsilon} H_\alpha(\mathbb{Q}) - \inf_{\substack{\sum_z \mathbb{Q}(z) \geq 1-\epsilon \\ \forall_z\, 0 \leq \mathbb{Q}(z) \leq \mathbb{P}(z)}} H_\alpha(\mathbb{Q}) \leq \frac{\alpha}{\alpha - 1} \log(1-\epsilon),$$

for $0 \leq \alpha < 1$, and that

$$\frac{\alpha}{\alpha - 1} \log(1-\epsilon) \leq \sup_{\delta(\mathbb{P},\mathbb{Q}) \leq \epsilon} H_\alpha(\mathbb{Q}) - \sup_{\substack{\sum_z \mathbb{Q}(z) \geq 1-\epsilon \\ \forall_z\, 0 \leq \mathbb{Q}(z) \leq \mathbb{P}(z)}} H_\alpha(\mathbb{Q}) \leq 0,$$

for $1 < \alpha \leq \infty$. So, only for $\alpha = 0$ either notion of ϵ-closeness yields the same value for the smooth Rényi entropy H_α^ϵ. But for all other values of α, the difference may be as large as $\frac{\alpha}{\alpha-1} \log(1-\epsilon)$. The maximum difference is attained for the uniform distribution $\mathbb{P}(z) = 1/m$ on a finite range $\mathcal{Z}$ of size m, assuming that ϵ is sufficiently small (i.e., $\epsilon < 1/m$).

Acknowledgment

Boris Škorić is gratefully acknowledged for discussions in the early stage of this work.

References

[1] I. Bjelakovic and A. Szkola, *"The Data Compression Theorem for Ergodic Quantum Information Sources"*, 2003. Available as quant-ph/0301043.

[2] T. M. Cover and J.A. Thomas, *"Elements of Information Theory"*, 2nd edition, Wiley-Interscience, 2006.

[3] M. A. Nielsen and I. L. Chuang, *"Quantum Computation and Quantum Information"*, Cambridge University Press, 2000.

[4] R. Renner, *"Security of Quantum Key Distribution"*, Diss. ETH No. 16242, PhD Thesis, ETH Zürich, September 2005. Also available as quant-ph/0512258.

[5] R. Renner and R. König, LNCS 3378, TCC 2005, pp. 407-425.

[6] R. Renner and S. Wolf, *"Smooth Rényi Entropy and Applications*, ISIT 2004, p. 232. Full version available as http://qi.ethz.ch/pub/publications/smooth.ps.

[7] R. Renner and S. Wolf, *"Simple and Tight Bounds for Information Reconciliation and Privacy Amplification"*, LNCS 3788, Asiacrypt 2005, pp. 199–216.