Abstract—We prove an achievability result for privacy amplification and decoupling in terms of the sandwiched Rényi entropy of order $\alpha \in (1, 2]$; this extends previous results which worked for $\alpha = 2$. The fact that this proof works for $\alpha$ close to 1 means that we can bypass the smooth min-entropy in the many applications where the bound comes from the fully quantum AEP or entropy accumulation, and carry out the whole proof using the Rényi entropy, thereby easily obtaining an error exponent for the final task. This effectively replaces smoothing, which is a difficult high-dimensional optimization problem, by an optimization problem over a single real parameter $\alpha$.

Index Terms—Quantum information, privacy amplification, randomness extraction.

I. INTRODUCTION

Consider the following scenario: we have a bipartite quantum state $\rho_{AE}$, with $A$ held by Alice and $E$ by an adversary Eve. Alice would like to apply an operation to $A$ to obtain a $C$ system that is uniform and independent from $E$. Here an “operation” means choosing a channel $\mathcal{R}^h$ at random from a family $\{\mathcal{R}^h_{A-C} : h \in \mathcal{H}\}$, announcing $h$ publicly and applying it to $A$. We are interested both in the case where $A$, $\mathcal{R}^h$ and $C$ are classical and in the case where they are quantum; the classical case corresponds to privacy amplification, and the quantum case to decoupling. In both cases, we would like to maximize the size of $C$ (in number of bits or qubits) that we get as a function of the input state $\rho_{AE}$. We have achievability results in the classical [1], [2] and quantum [3], [4], [5], [6] cases of the form

$$\mathbb{E}_h \left\| \mathcal{R}^{h} (\rho_{AE}) - \frac{1}{|C|} \otimes \rho_E \right\|_1 \leq 2^{\frac{1}{\alpha} (\log |C| - H_\alpha (A|E)_\rho)}, \quad (1)$$

where $H_\alpha (A|E)_\rho$ denotes the sandwiched Rényi conditional entropy of order $\alpha \in \left[\frac{1}{\alpha}, 1\right) \cup (1, \infty)$ of $A$ given $E$ (see below for the definition). However, while the Rényi 2-entropy is what naturally comes out of these proofs, it has mostly been used as an auxiliary quantity, due to one major flaw: it is very sensitive to small variations in the state $\rho_{AE}$ over which it is computed, and as a result we can rarely get good bounds on it in most settings of interest. To remedy this, we instead use the min-entropy $H_{\min} (A|E)_\rho$ (which is a lower bound on the 2-entropy and has additional desirable properties—again, see below for the definition), and smooth it, which means that we compute it on the best state in an $\varepsilon$-ball around $\rho$:

$$H_{\min}^\varepsilon (A|E)_\rho := \sup_{\theta_{AB} : D(\rho, \theta) \leq \varepsilon} H_{\min} (A|B)_{\theta}.$$  

This yields an achievability result of the form:

$$\mathbb{E}_h \left\| \mathcal{R}^{h} (\rho_{AE}) - \frac{1}{|C|} \otimes \rho_E \right\|_1 \leq 2\varepsilon + 2^{\frac{1}{\alpha} (\log |C| - H_{\min}^\varepsilon (A|E)_\rho)}.$$  

(2)

How can we get good lower bounds on $H_{\min}^\varepsilon$ in applications of interest? One simple case involves states where $A$ and $E$ both consist of $n$ independent, identically distributed systems $A_1^n$ and $E_1^n$. This is a major “use case”, since many problems of interest (such as quantum key distribution protocols) can be reduced to it via the de Finetti theorem [2], [8]. In [9], the fully quantum asymptotic equipartition property is proven, which states a lower bound in terms of the von Neumann entropy of a single system:

$$H_{\min}^\varepsilon (A_1^n | E_1^n)_\rho \geq n H(A_1 | E_1)_\rho - O(\sqrt{n}).$$

The proof works by first lower-bounding $H_{\min}^\varepsilon (A_1^n | E_1^n)$ by $H_\alpha (A_1^n | E_1^n) = n H_\alpha (A_1 | E_1)$, and then lower-bounding $H_\alpha$ by $H$. Choosing $\alpha \approx 1 + \frac{1}{\sqrt{n}}$ then yields the bound we want. The same blueprint also works in more complicated settings: in [10], the entropy accumulation theorem (EAT), which can be used to bound the smooth min-entropy of states that are produced by step-by-step processes such as those arising in device-independent QKD protocols, also involves bounding the Rényi entropy for $\alpha$ just above 1.

In this paper, we show how to bypass the smooth min-entropy entirely in this type of proof and use a bound on the Rényi entropy of order $\alpha \in (1, 2]$ directly. In other words, we show that:

$$\mathbb{E}_h \left\| \mathcal{R}^{h} (\rho_{AE}) - \frac{1}{|C|} \otimes \rho_E \right\|_1 \leq 2^{\frac{1}{\alpha} - 1} 2^{\frac{1}{\alpha} (\log |C| - H_\alpha (A|E)_\rho)}.$$  

(3)

This simplifies the overall analysis considerably: one gets much cleaner expressions, we can get error exponents quite

0018-9448 © 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.
easily on the final quantity of interest that are better than those that can be obtained via smoothing,\(^2\) and we replace smoothing (which is a difficult high-dimensional optimization problem) by an optimization over a single real parameter \(\alpha\). This can then be applied to a wide variety of problems, ranging from privacy amplification in QKD, to decoupling [5], [6] in the quantum case, which in turn yields results in quantum channel coding. Furthermore, this approach raises an interesting conceptual question: can we do all of one-shot information theory using Rényi-type quantities instead of min/max quantities and smoothing?

The proof of the result is based on norm interpolation methods reminiscent of the Riesz-Thorin theorem. Specifically, we interpolate between the \(\alpha = 2\) case, which we know how to deal with using previously known bounds, and the \(\alpha = 1\) case, where a trivial bound is sufficient for our purposes.

The rest of this paper is structured as follows. In the next subsection, we give an overview of the state of the art and related results. Then, in Section II, we present the notation we use as well as the necessary background on entropic quantities and norm interpolation, and give a definition of the class of randomizing families of channels to which our theorem applies. In Section III, we prove the main result (i.e. Equation (3)), and Section IV contains a sample application which shows what happens when we combine this result with the entropy accumulation theorem. Finally, Section V concludes with a discussion of the result and of open questions.

A. State of the Art and Related Work

The Rényi entropy has been used in information theory for several decades, see e.g. [12], [13], [14], [15].

For the problem at hand, in the purely classical case (i.e. when \(E\) is a classical random variable), an expression very similar to (3) was derived by Hayashi [11, Equation (67)], and tightness results are provided in [11] and [16]. The results of [11] were then extended to almost 2-universal hash functions in [17]. These results can then be used to obtain error exponents in the iid case.

In the CQ case (i.e. classical \(A\) but quantum \(E\)), the situation is more complicated. In [18], Hayashi gives two versions of (3) that include additional terms involving either the number of distinct eigenvalues of an operator, or the ratio between the minimal and maximal eigenvalues of the same operator. In the iid case, this additional term is not problematic as it grows only polynomially in \(n\) and is thus negligible as \(n \to \infty\). In fact, [19] uses a similar bound to pin down the exact error exponent for the iid version of this task in the high rate regime (i.e. close to the limit). In the general one-shot case, however, this term is much harder to deal with. In particular, such a result cannot be directly applied to a setting like DIQKD due to this issue. Another drawback of [18] is that the bound obtained is in terms of the Petz-Rényi entropy rather than the sandwiched Rényi entropy, which is a worse bound in general.

\(^2\)See [11, Section III.D] for a comparison between error exponents obtained via the Rényi approach and the smooth min-entropy approach in the iid setting.

| Symbol | Definition |
|--------|------------|
| \(A, B, C, \ldots\) | Quantum systems |
| \(A, B, \ldots\) | Hilbert spaces corresponding to systems \(A, B, \ldots\) |
| \(|A|\) | Dimension of \(A\) |
| \(L(A, B)\) | Set of linear operators from \(A\) to \(B\) |
| \(L(A)\) | Set of linear operators on \(A\) |
| \(X_{AB}\) | Operator in \(L(A \otimes B)\) |
| \(X_{A \to B}\) | Operator in \(L(A, B)\) |
| \(N(A)\) | Set of normal operators on \(A\) |
| \(\text{Pos}(A)\) | Set of positive semidefinite operators on \(A\) |
| \(\text{Di}(A)\) | Set of positive semidefinite operators on \(A\) with unit trace |
| \(\mathbb{I}_A\) | Identity operator on \(A\) |
| \(X^\dagger\) | Adjoint of \(X\) |
| \(X_A \geq Y_A\) | \(X - Y \in \text{Pos}(A)\) |
| \(|X|_p\) | Schatten \(p\)-norm of \(X\): \(\text{Tr}((X^\dagger X)^{\frac{p}{2}})^{\frac{2}{p}}\) |
| \(\text{supp}_X\) | Support of \(X\) |
| \(H(A|B)_\alpha\) | Conditional von Neumann entropy: \(-\text{Tr}[(\rho_{AB} \log \rho_{AB}) + \text{Tr}(\rho_B \log \rho_B)]\) |

In the fully quantum case, with both \(A\) and \(E\) quantum, Sharma [20] obtains a result very similar to those of Hayashi in the CQ case, with an additional term involving the number of distinct eigenvalues of an operator. Sharma then goes on to apply this decoupling theorem to a panoply of quantum Shannon theory problems, obtaining error exponents for the iid case for all of them. More recently, [21] derives the exact error exponent in the iid case in the high-rate regime, also using one-shot bounds with the extra term. However, the same caveat applies as in the CQ case: while the additional term poses no problem in the iid case, it is difficult to deal with in general.

II. Preliminaries

A. Notation

In Table I, we summarize the notation used throughout the paper. Note that all Hilbert spaces are finite-dimensional and endowed with a standard computational basis that we denote by \(\{|a\rangle: 1 \leq a \leq |A|\}\) for system \(A\). Furthermore, by “CQ operator”, we mean a bipartite operator \(X_{AB}\) of the form:

\[
X_{AB} = \sum_a |a\rangle\langle a| \otimes X_B(a)
\]

where \(X_B(a) \in L(B)\) for all \(a\).

B. Entropic Quantities

Our main theorem involves the so-called sandwiched conditional quantum Rényi entropy of a bipartite quantum state. This information measure was introduced relatively recently [22], [23] as an alternative way to generalize the classical Rényi entropy to the quantum case. It is defined as follows:

**Definition 1 (Sandwiched Rényi Entropy):** Let \(\rho_{AB} \in D(A \otimes B)\) and \(\sigma_B \in D(B)\), and let \(\alpha \in \left(\frac{1}{2}, 1\right) \cup (1, \infty)\) be a real parameter. Then, the sandwiched Rényi entropy of order \(\alpha\) of \(A\) given \(\sigma_B\) is defined as

\[
H_{\alpha}(A|B)_{\rho|\sigma} := \frac{1}{1 - \alpha} \log \text{Tr}\left(\sigma_B^{\frac{1}{2\alpha}} \rho_{AB} \sigma_B^{\frac{1}{2\alpha}}\right)^\alpha
\]
if \( \alpha < 1 \) and \( \text{Tr}[\rho \sigma] \neq 0 \), or \( \rho \in A \otimes \text{supp}(\sigma) \), and \( -\infty \) otherwise, and the sandwiched Rényi entropy of \( A \) given \( B \) is defined as:

\[
H_{\alpha}(A|B)_{\rho} := \max_{\omega_B \in \text{D}(B)} H_{\alpha}(A|B)_{\rho|\omega}.
\]

This quantity is monotone nonincreasing in \( \alpha \) and satisfies the data processing inequality \([24], [25]\), and if we take the limit as \( \alpha \to 1 \), we get the von Neumann entropy \( H(A|B)_{\rho} \). Furthermore, the limit of the Rényi entropy when \( \alpha \to \infty \) makes sense and is called the min-entropy. It can be shown to be equal to

\[
H_{\text{min}}(A|B)_{\rho} = H_{\infty}(A|B)_{\rho} = \sup \{ \lambda : \exists \sigma_B \in \text{D}(B), \rho_{AB} \leq 2^{-\lambda} \mathbb{1}_A \otimes \sigma_B \}.
\]

Note also that the case \( \alpha = 2 \) corresponds to the collision entropy, which is the quantity that shows up in the original proof of security of privacy amplification \([1], [2], \text{Chapter 5}\), as well as in proofs of the “decoupling theorem” in the fully quantum case \([4], [5], [6]\). This is the main reason why the sandwiched Rényi entropy is relevant for us here: it offers a natural way to interpolate between the \( \alpha = 2 \) case, which we know how to deal with, and the \( \alpha = 1 \) case where a trivial bound suffices for our purposes.

While the Rényi 2-entropy is what naturally comes out of the above proofs, it has mostly been used as an auxiliary quantity, due to its sensitivity to small variations in the state \( \rho_{AB} \) over which it is computed, and as a result we can rarely get good bounds on it in most settings of interest. To remedy this, we instead use the min-entropy (which is a lower bound on the 2-entropy by monotonicity in \( \alpha \) and has additional desirable properties), and smooth it, which means that we compute it on the best state in an \( \varepsilon \)-ball around \( \rho \):

\[
H_{\text{min}}^{\varepsilon}(A|B)_{\rho} := \max_{\theta_{AB} \in B(\rho, \theta) \leq \varepsilon} H_{\text{min}}(A|B)_{\theta}.
\]

C. Schatten \( p \)-Norms and Norm Interpolation

The main technical tool that we will need is a technique from complex analysis that allows us to interpolate between two norms. The relevant norms here are the Schatten \( p \)-norms:

**Definition 2 (Schatten \( p \)-norms):** Let \( N \in \text{L}(A, B) \) be a linear operator, and let \( p \geq 1 \). Then, the Schatten \( p \)-norm of \( N \) is defined as

\[
\|N\|_p := \text{Tr} \left[ (N^\dagger N)^{\frac{p}{2}} \right]^{\frac{1}{p}}.
\]

These norms satisfy the Hölder inequality: for \( M \in \text{L}(B, C) \) and \( N \in \text{L}(A, B) \) and \( p > 1 \),

\[
\|MN\|_1 \leq \|M\|_p \|N\|_p.
\]

It can also be shown that the following version holds for three operators \( M, N, R \):

\[
\|MNR\|_1 \leq \|M\|_{p_1} \|N\|_{p_2} \|R\|_{p_3}.
\]

where \( 1 = \frac{1}{p_1} + \frac{1}{p_2} + \frac{1}{p_3} \). Furthermore, these norms can be expressed as the following optimization problem:

\[
\|M\|_p = \sup_{Y \in \text{L}(C,B) : \|Y\|_{\sup} = 1} \|\text{Tr}[YM]\|,
\]

and if \( Q \in \text{Herm}(A) \), then,

\[
\|Q\|_p = \sup_{Y \in \text{Herm}(A) : \|Y\|_{\sup} = 1} \|\text{Tr}[YQ]\|.
\]

These norms are also unitarily invariant: for any unitaries \( U_B \) and \( V_C \),

\[
\|M\|_p = \|MU\|_p = \|VM\|_p.
\]

The main reason for which these norms are relevant for us is that the sandwiched Rényi entropy can be expressed as:

\[
2^{-\frac{1}{\alpha-1}} H_{\alpha}(A|B)_{\rho} = \|\frac{1}{\rho_B} \sigma_B - \frac{1}{\rho_{AB}} \sigma_B\|_{\alpha}.
\]

Our proof will be based on a Riesz-Thorin-like method to interpolate between the 2-norm and the 1-norm, in order to handle Rényi entropies with \( \alpha \in (1,2] \). These methods have already been applied to quantum information theory before \([26], [27]\), and more specifically for proving properties of the sandwiched Rényi divergence \([25], [28], [29]\). Unfortunately, standard interpolation theorems such as Riesz-Thorin do not seem to apply directly to our case, and we will need to directly use the Hadamard three-line theorem, which is the main ingredient used to prove these standard results:

**Theorem 3 (Hadamard Three-Line Theorem):** Let \( S := \{ z \in \mathbb{C} : 0 \leq \Re(z) \leq 1 \} \), and let \( f : S \to \mathbb{C} \) be a function that is holomorphic on the interior of \( S \) and continuous on the border. Let \( 1 \leq p_0 \leq p_1 \) and \( 0 < \theta < 1 \), and define \( p_0 \leq p_\theta \leq p_1 \) via:

\[
\frac{1}{p_\theta} = \frac{1}{p_0} + \frac{\theta}{p_1}.
\]

Furthermore, for \( k \in (0,1) \), let \( M_k := \sup_{z \in S} |f(k - iz)| \). Then, for any \( 0 \leq \theta \leq 1 \),

\[
|f(\theta)| \leq M_0^{1-\theta} M_1^{\theta}.
\]

A proof can be found in \([30], \text{page 33}\).
functions. In this case, each $h$ is a function from the set $A$ to $C$, and we define $R^h_{A \rightarrow C}$ as:

$$R^h_{A \rightarrow C}(\theta_A) = \sum_{a \in A} \langle a | \theta_A | a \rangle \cdot |h(a)\rangle \langle h(a)|_C,$$

where the Hilbert spaces $A$ and $C$ have computational bases that are indexed by the sets $A$ and $C$ respectively.

We represent the setup in the following manner. Let $\mathcal{U}_{A \rightarrow C}$ be the perfectly randomizing channel

$$\mathcal{U}(\theta_A) = \frac{1}{|C|} \cdot Tr[\theta].$$

Our goal is then to ensure that, for any state $\rho_{AE} \in D(A \otimes E)$,

$$E_{h \sim p}[\|(R^h - \mathcal{U})(\rho_{AE})\|_2 \leq \varepsilon],$$

for some suitably small $\varepsilon$. (Of course, in the CQ case, this condition only needs to be fulfilled for a CQ state $\rho_{AE}$.)

Our proof will be able to deal with families of channels that satisfy the following definition:

**Definition 4 (Randomizing family of channels):** A family of channels $\{R^h_{A \rightarrow C} : h \in \mathcal{H}\}$ with distribution $p$ is $\lambda$-randomizing if, for any $\rho_{AE} \in L(A \otimes E)$,

$$E_{h \sim p}[\|(R^h - \mathcal{U})(\rho_{AE})\|_2 \leq \lambda |\rho_{AE}|_2].$$

In the CQ-case, a randomizing family of hash functions is $\lambda$-randomizing if the above condition holds for any CQ $\rho_{AE}$. One way to prove this condition for a given family of channels is to use the following lemma, which makes no reference to a purifying system $E$ (see also [31]):

**Lemma 5:** Let $\{R^h_{A \rightarrow C} : h \in \mathcal{H}\}$ with distribution $p$ be a family of channels that satisfies

$$E_{h \sim p}[\|(R^h - \mathcal{U})(\rho_{AE})\|_2 \leq \lambda |\rho_{AE}|_2],$$

for any $\rho_{AE} \in L(A)$. Then, it is $\lambda$-randomizing.

**Proof:** First, note that for any $\sigma_{AE} \in L(A \otimes E)$, we can write $\sigma_{AE} = \sigma_A(e, e') \otimes |e\rangle \langle e'|$, and we have that:

$$|\sigma_{AE}|_2^2 = \text{Tr}[\sigma_{AE}^\dagger \sigma_{AE}^\dagger] = \text{Tr}[\sum_{e} \sigma_A(e) \sigma_A(e)^\dagger] = \sum_{e} |\sigma_A(e)|_2^2.$$

Let $\rho_{AE} \in L(A \otimes E)$, and let $N^h := R^h - \mathcal{U}$. We then have that

$$E_{h \sim p}[\|(R^h - \mathcal{U})(\rho_{AE})\|_2] = E_{h \sim p}\left[\sum_{e} |N^h(\rho_{AE}(e))|_2^2\right] = \lambda |\rho_{AE}|_2^2.$$

Substituting this into (5) yields the lemma.

**III. Main Result**

As stated in the introduction, our main technical tool will be norm interpolation: we need a way to interpolate between the cases $\alpha = 1$ and $\alpha = 2$. Unfortunately, standard interpolation results cannot be immediately applied due to the expectation over $h$, and we need the following custom result that we prove directly from the three-line theorem:

**Lemma 8:** Let $h$ be a random variable taking values in a set $\mathcal{H}$, and for each $h \in \mathcal{H}$, let $N^h : L(A) \rightarrow L(B)$ be any superoperator, and let $\alpha \in (1, 2]$. Then, for any
\( \rho_{AE} \in \mathcal{D}(A \otimes E) \) and any \( \sigma_{E} \in \mathcal{D}(E) \) such that
\[
\left\| \frac{1}{\bar{\alpha}} \rho_{AE} \sigma_{E} \right\|_{\alpha} = 1,
\]
we have:
\[
\mathbb{E}_{h} \left\| \mathcal{N}^{h} \left( \frac{1}{\bar{\alpha}} \rho_{AE} \sigma_{E} \right) \right\|_{\alpha} \leq \max_{\eta \in \mathcal{N}(A \otimes E) : \left\| \eta \right\|_{1} = 1} \mathbb{E}_{h} \left\| \mathcal{N}^{h} (\eta_{AE}) \right\|_{1}^{2} \left( \max_{\eta \in \mathcal{N}(A \otimes E) : \left\| \eta \right\|_{2} = 1} \mathbb{E}_{h} \left\| \mathcal{N}^{h} (\eta_{AE}) \right\|_{2} \right)^{2 \left( \frac{\alpha - 1}{\bar{\alpha}} \right)}.
\]
Furthermore, if \( \rho_{AE} \) is a CQ state, then the maximizations in the two terms on the right-hand side can be restricted to CQ operators.

**Proof:** For any \( h \in \mathcal{H} \), let \( Y_{h} \in \text{Herm}(B \otimes E) \) such that
\[
\left\| Y_{h} \right\|_{\frac{1}{\bar{\alpha}}} = 1
\]
and
\[
\text{Tr} \left[ Y_{h} \mathcal{N}^{h} \left( \frac{1}{\bar{\alpha}} \rho_{AE} \sigma_{E} \right) \right] = \left\| \mathcal{N}^{h} \left( \frac{1}{\bar{\alpha}} \rho_{AE} \sigma_{E} \right) \right\|_{\alpha}.
\]
Now, define \( f : S \to \mathbb{C} \) as:
\[
f(z) := \mathbb{E}_{h} \text{Tr} \left[ Y_{h}^{-z \left( \frac{1}{\bar{\alpha}} \right)} \mathcal{N}^{h} \left( \left( \frac{1}{\bar{\alpha}} \rho_{AE} \sigma_{E} \right) \right)^{(1-z)\alpha + z(\alpha/2)} \right].
\]
This function fulfills the conditions of the three-line theorem (Theorem 3), which we now apply with the choices
\[
p_{0} = 1,
p_{1} = 2,
\theta = 2 \left( \frac{\alpha - 1}{\bar{\alpha}} \right) \in (0, 1].
\]
Note that these choices imply that \( p_{0} = \alpha \) and that \( (1 - \theta)\alpha + \theta(\alpha/2) = 1 \). We therefore have that
\[
|f(\theta)| = \mathbb{E}_{h} \text{Tr} \left[ Y_{h} \mathcal{N}^{h} \left( \frac{1}{\bar{\alpha}} \rho_{AE} \sigma_{E} \right) \right] = \mathbb{E}_{h} \left\| \mathcal{N}^{h} \left( \frac{1}{\bar{\alpha}} \rho_{AE} \sigma_{E} \right) \right\|_{\alpha} \leq M_{0}^{2} \mathbb{E}_{h} \left\| \mathcal{N}^{h} (\eta_{AE}) \right\|_{1}^{\frac{2(\alpha - 1)}{\bar{\alpha}}},
\]
where \( M_{0} \) and \( M_{1} \) are defined as in Theorem 3. We now bound these two quantities, starting with \( M_{0} \):
\[
M_{0} \equiv \sup_{t \in \mathbb{R}} |f(it)| = \sup_{t \in \mathbb{R}} \left| \mathbb{E}_{h} \text{Tr} \left[ Y_{h}^{it} \left( \frac{1}{\bar{\alpha}} \right) \mathcal{N}^{h} \left( \frac{1}{\bar{\alpha}} \rho_{AE} \sigma_{E} \right) \left( \frac{1}{\bar{\alpha}} \right) \right] \right| = \mathbb{E}_{h} \left\| \mathcal{N}^{h} \left( \frac{1}{\bar{\alpha}} \rho_{AE} \sigma_{E} \right) \left( \frac{1}{\bar{\alpha}} \right) \right\|_{\alpha} = 1,
\]
and hence we can continue as:
\[
\mathbb{E}_{h} \left\| \mathcal{N}^{h} \left( \frac{1}{\bar{\alpha}} \rho_{AE} \sigma_{E} \right) \right\|_{\alpha} \leq \sup_{t \in \mathbb{R}} \mathbb{E}_{h} \left\| \mathcal{N}^{h} \left( \frac{1}{\bar{\alpha}} \rho_{AE} \sigma_{E} \right) \left( \frac{1}{\bar{\alpha}} \right) \right\|_{\alpha} = 1.
\]
We are now ready to state and prove the main theorem:

**Theorem 9 (Main Theorem):** Let \( \{R^h_{A-C} : h \in \mathcal{H} \} \) be a \( \lambda \)-randomizing family of channels, and let \( \rho_{AE} \in D(A \otimes E) \) and \( \sigma_E \in D(E) \) such that \( \rho_{AE} \) is supported on \( A \otimes \text{supp}(\sigma_E) \).

Then,

\[
\mathbb{E}_h \| (R^h - U)(\rho_{AE}) \|_1 \leq 2^{\frac{1}{2} - 1} \cdot 2^{\frac{1}{2} \log(\text{IC} - H_n(A|E)_{\rho_{AE}}) + 2 \log \lambda}.
\]

Furthermore, if \( \rho_{AE} \) is CQ and \( \{R^h : h \in \mathcal{H} \} \) is a \( \lambda \)-randomizing family of hash functions, the same bound holds.

**Proof:** For any \( h \in \mathcal{H} \), let us define the superoperator \( N^h := R^h - U \). We can now begin as follows:

\[
\mathbb{E}_h \| (R^h - U)(\rho_{AE}) \|_1 = \mathbb{E}_h \| N^h (\rho_{AE}) \|_1 \leq \mathbb{E}_h \| \sigma_E^{-\frac{1}{2}} \rho_{AE} \sigma_E^{-\frac{1}{2}} \|_\alpha \\
\times \mathbb{E}_h \| N^h (\sigma_E^{-\frac{1}{2}} \rho_{AE} \sigma_E^{-\frac{1}{2}} \sigma_E^{-\frac{1}{2}} \|_1 \\
\leq \mathbb{E}_h \| N^h (\rho_{AE}) \|_1 \\
= 2^{\frac{1}{2} \log(\text{IC} - H_n(A|E)_{\rho_{AE}})} \max_{\omega} \mathbb{E}_h \| N^h (\omega_{AE}) \|_1 \\
= 2^{\frac{1}{2} \log(\text{IC} - H_n(A|E)_{\rho_{AE}})} \max_{\omega} \mathbb{E}_h \| N^h (\omega_{AE}) \|_1.
\]

Lemma 8 \( \leq 2^{\frac{1}{2} \log(\text{IC} - H_n(A|E)_{\rho_{AE}})} \max_{\omega} \mathbb{E}_h \| N^h (\omega_{AE}) \|_1 \\
\leq 2^{\frac{1}{2} \log(\text{IC} - H_n(A|E)_{\rho_{AE}})} \max_{\omega} \mathbb{E}_h \| N^h (\omega_{AE}) \|_1.
\]

where \( (a) \) follows from Hölder’s inequality (Equation (4)), and the last inequality follows from a trivial bound of 2 on the 1-norm term that we can obtain using the triangle inequality and Lemma 11. To finish the proof, fix \( \eta_{AE} \) to any normal operator such that \( \| \eta \|_1 \leq 1 \) (and choose a CQ \( \eta_{AE} \) if we are in the CQ case). Then,

\[
\mathbb{E}_h \| N^h (\eta_{AE}) \|_1 = \mathbb{E}_h \| (R^h - U)(\eta_{AE}) \|_1 \leq \lambda \| \eta_{AE} \|_2
\]

where the first inequality follows from Definition 4. This concludes the proof.

**IV. SAMPLE APPLICATION: ENTROPY ACCUMULATION**

In this section, we show how our main theorem can be applied to the entropy accumulation theorem (EAT) [10], which can (among other things) be used to provide security proofs for device-independent QKD protocols with nearly optimal rates. This theorem is used to lower-bound the smooth min-entropy of \( n \) random variables that are produced by a sequential process. Its proof follows the outline given in the introduction: we first lower-bound the smooth min-entropy by a Rényi entropy of order \( \alpha \), which is then lower-bounded by the von Neumann entropy. Choosing \( \alpha \approx 1 + \frac{1}{\sqrt{n}} \) then yields the theorem.

In applications where the bound on the smooth min-entropy is then used for privacy amplification (which is the case for QKD), we can use our main theorem to bypass the smooth min-entropy entirely and directly obtain an error exponent for the trace distance to a perfect key. Furthermore, the resulting proof is overall simpler as it does not involve any smoothing parameter.

To show how this works, we quickly explain a simplified version of the EAT, and then show what happens when we plug in the intermediate statement about Rényi entropies into our main theorem. The EAT provides bounds for states that are generated by step-by-step processes of the form:

\[
\rho_0 \xrightarrow{A_1} \rho_1 \xrightarrow{A_2} \rho_2 \xrightarrow{A_n} \rho_n.
\]

In the above, \( A_i \) are classical systems and that \( T_i \) are classical bits. Suppose we are interested in lower-bounding \( H_{\text{min}}(A_i^0 | X_i^n, E, \text{wt}(T_i^0) = w) \), where \( \text{wt}(T_i^0) \) denotes the Hamming weight of the bitstring \( T_i^0 \) (i.e. the number of positions that are one). The EAT provides such a lower bound of the form:

\[
H_{\text{min}}^\varepsilon(A_i^0 | X_i^n, E, \text{wt}(T_i^0) = w) \geq n f(w) - \sqrt{n} V \sqrt{1 - 2 \log(\varepsilon \Pr(\text{wt}(T_i^0) = w))},
\]

where \( f(\cdot) \) is a tradeoff function that tells us how much entropy we can expect to get the probability of seeing a one on \( T_i \), and \( V \) is a parameter that depends on the size of the A systems and how sensitive \( f \) is to small variations in \( w \); it is given by:

\[
V = 2\| \nabla f \|_{\infty} + 2 \log(1 + 2dA).
\]

We could then use this bound in the usual statement for privacy amplification and get a bound on the trace distance to an ideal key when we apply a randomizing hash function to \( A_i^0 \).

To see how this picture changes when we use our theorem instead, we reproduce here a simplified version of Proposition 4.5 from [10], which is the intermediate statement of the EAT that lower-bounds the Rényi entropy instead of the smooth

\[
\text{in a standard CHS-based DIQKD protocol, } A_i^0 \text{ would be Alice’s output, }
\]

\( X_i^n \) would be Alice’s questions, and \( T_i \) would indicate whether the CHS game at position \( i \) was won; we then want a bound that depends on how many games were won. The specific details of how this works are not relevant here, but see [34, 35] for a detailed analysis of DIQKD using the EAT.

Authorized licensed use limited to the terms of the applicable license agreement with IEEE. Restrictions apply.
The next step is to use a randomizing family of hash functions to extract randomness at rate $R$ (i.e., have the hash function produce $nR$ bits at the output). We expect the error to decay as a function of the difference between $f(w)$ (the first-order term in the amount of randomness produced) and the rate $R$, and of $V$ which should provide a scale for this difference. Indeed, this is what happens if we substitute this bound into our main theorem:

**Theorem 10**: If $\{R^h : h \in \mathcal{H}\}$ is a 1-randomizing family of hash functions that produce $nR$ bits at the output, and if $0 < f(w) - R \leq \frac{V^2}{2}$, we have that:

$$\Pr[\text{wt}(T^p_1) = w] \leq 2 \cdot 2^{nE(R)}$$

where the error exponent $E(R)$ is given by:

$$E(R) = \frac{1}{2} \left( \frac{f(w) - R}{V} \right)^2.$$  

This statement is much nicer to work with than what we get by going through the smooth min-entropy. Note in particular that the probability of the event $\text{wt}(T^p_1) = w$ appears in a much more natural way, and we do not have to assume a particular bound for it, as we normally do with the EAT.

**Proof**: We can immediately substitute the bound from (14) into Theorem 9 to get:

$$\mathbb{E}_{h} \left[ R^h \left( \rho A^p_1 X^p_1 E|\text{wt}(T^p_i) = w) \right) \right] \leq 2^{\frac{n-1}{2}} \cdot 2^{\frac{n}{n-1} \left( R - f(w) + \frac{\alpha}{\alpha - 1} V^2 + \frac{\alpha}{\alpha - 1} \log \frac{1}{\Pr[\text{wt}(T^p_1) = w]} \right)} \leq \frac{2}{\Pr[\text{wt}(T^p_1) = w]} \cdot 2^{n \frac{n}{n-1} \left( R - f(w) + \frac{\alpha}{\alpha - 1} V^2 \right)}.$$  

All that remains to do is to optimize the exponent over $\alpha \in (1, 2]$. To do this, let $\beta := \frac{\alpha}{\alpha - 1}$, and note that the exponent becomes:

$$\beta \left( R - f(w) + \frac{V^2}{2} \right) = \beta^2 \frac{V^2}{2} - \beta (f(w) - R).$$

It is easy to see that this expression is minimized whenever $\beta = \frac{f(w) - R}{V^2}$. Substituting this into (16) yields the bound advertised in the theorem statement. Furthermore, $\alpha \in (1, 2]$ if $0 < \beta \leq \frac{1}{2}$, which means that this bound applies if $0 < f(w) - R \leq \frac{V^2}{2}$ (i.e., only for rates close enough to the first-order term $f(w)$).

---

4Note that one could also use the bound given in [36], which is tighter but a bit more complex. Here we stick to the simpler bound of [10] for ease of exposition.
considered (see [38] for a systematic study of this problem), but so far none has proven to be fully satisfactory.

It would also be interesting to see to what extent our results can produce better results numerically in realistic DIQKD protocols. In particular, it could be naturally combined with a recent result [39] which defines a new family of Rényi entropies that can be computed by semidefinite programs, and which can be used to compute numerical bounds on tradeoff functions for the EAT. This would provide an “all-Rényi” approach to DIQKD security proofs that would likely yield very tight bounds. Note that another route to an all-Rényi approach would be to combine this result with the proof from [40] and [41], which gives bounds on the smooth min-entropy generated by device-independent randomness expansion in terms of Rényi entropies—it would thus also be natural there to cut out the min-entropy and work only in terms of Rényi entropies.

Finally, one might try to extend the main theorem to $\alpha > 2$. However, it might be tempting to conjecture that this requires stronger conditions on the randomizing family of channels: the fact that we have an achievability result at $\alpha = 2$ is tied to the fact that the channels can randomize the second moment of the state. Going higher might therefore require a family that randomizes higher moments.

APPENDIX

Lemma 11: Let $N_{A \rightarrow B}$ be a trace non-increasing, completely positive map, and let $M \in N(A)$ be a normal operator. Then,

$$\|N(M)\|_1 \leq \|M\|_1.$$  

Proof: First note that $M$ can be decomposed as $M = UD$ for a unitary $U$ and positive semidefinite $D$ such that $U$ and $D$ commute. To see this, write the spectral decomposition of $M$ as

$$M = \sum_{i} \lambda_i e^{i\varphi_i} |\psi_i\rangle\langle \psi_i|,$$

where the $\lambda_i$ and $\varphi_i$ are positive reals. Then,

$$D := \sum_{i} \lambda_i |\psi_i\rangle\langle \psi_i|,$$

$$U := \sum_{i} e^{i\varphi_i} |\psi_i\rangle\langle \psi_i|.$$  

Now, let $\Phi_{AA'} := \sum_{ij} |ij\rangle\langle j|$. We then have that

$$\|N(M)\|_1 = \|Tr_{AA'}[N(UD^{\frac{1}{2}}\Phi D^{\frac{1}{2}})]\|_1$$

$$= \sup_{V_{AA'}, ||V_{AA'}||_1 \leq 1} \|Tr[V_{AA'}N(UD^{\frac{1}{2}}\Phi D^{\frac{1}{2}})]\|_1$$

$$\leq \sup_{V_{AA'}, ||V_{AA'}||_1 \leq 1} \|Tr[V_{AA'}N(UD^{\frac{1}{2}}\Phi D^{\frac{1}{2}})]\|_1$$

$$= \|N(D^{\frac{1}{2}}\Phi D^{\frac{1}{2}})\|_1 = \|N[D]\|_1 \leq \|M\|_1.$$  

ACKNOWLEDGMENT

The author would like to thank Joseph M. Renes, Andreas Winter, Marco Tomamichel, and Masahito Hayashi, and the anonymous referees for their helpful discussions, comments, and suggestions.

REFERENCES

[1] R. König, U. Maurer, and R. Renner, “On the power of quantum memory,” IEEE Trans. Inf. Theory, vol. 51, no. 7, pp. 2391–2401, Jul. 2005.

[2] R. Renner, “Security of quantum key distribution,” Ph.D. dissertation, ETH Zürich, Zürich, Switzerland, 2005.

[3] M. Horodecki, J. Oppenheim, and A. Winter, “Quantum state merging and negative information,” Commun. Math. Phys., vol. 269, no. 1, pp. 107–136, 2007.

[4] A. Abeyesinghe, I. Devetak, P. Hayden, and A. Winter, “The mother of all protocols: Restructuring quantum information’s family tree,” Proc. Roy. Soc. A, Math., Phys. Eng. Sci., vol. 465, no. 2108, pp. 2537–2563, Aug. 2009.

[5] F. Dupuis, M. Berta, J. Wullschleger, and R. Renner, “One-shot decoupling,” Commun. Math. Phys., vol. 328, no. 1, pp. 251–284, May 2014.

[6] F. Dupuis, “The decoupling approach to quantum information theory,” Ph.D. dissertation, Département d’informatique et de recherche opérationnelle, Univ. Montréal, Montréal, QC, Canada, 2009.

[7] M. Tomamichel, Quantum Information Processing With Finite Resources: Mathematical Foundations, Springer Verlag, 2016. [Online]. Available: https://link.springer.com/book/10.1007/978-3-319-21891-5

[8] R. Renner, “Symmetry of large physical systems implies independence of subsystems,” Nature Phys., vol. 3, no. 9, pp. 645–649, Sep. 2007.

[9] M. Tomamichel, R. Colbeck, and R. Renner, “A fully quantum asymptotic equipartition property,” IEEE Trans. Inf. Theory, vol. 55, no. 12, pp. 5840–5847, Dec. 2009.

[10] F. Dupuis, O. Fawzi, and R. Renner, “Entropy accumulation,” Commun. Math. Phys., vol. 379, no. 3, pp. 913–967, 2020.

[11] M. Hayashi, “ Tight exponential analysis of universally composable privacy amplification and its applications,” IEEE Trans. Inf. Theory, vol. 59, no. 11, pp. 7728–7746, Nov. 2013.

[12] A. Rényi, “On measures of entropy and information,” in Proc. 4th Berkeley Symp. Math. Statist. Probab., 1961, pp. 547–561. [Online]. Available: http://projecteuclid.org/euclid.bsmsp/1200512181

[13] T. S. Han, Information-Spectrum Methods in Information Theory, Springer-Verlag, 2002. [Online]. Available: https://link.springer.com/book/10.1007/978-3-662-12066-8

[14] R. Gallager, “A simple derivation of the coding theorem and some applications,” IEEE Trans. Inf. Theory, vol. IT-11, no. 1, pp. 3–18, Jan. 1965.

[15] S. Arimoto, “On the converse to the coding theorem for discrete memoryless channels (Corresp.),” IEEE Trans. Inf. Theory, vol. IT-19, no. 3, pp. 357–359, May 1973.

[16] M. Hayashi and S. Watanabe, “Uniform random number generation from Markov chains: Non-asymptotic and asymptotic analyses,” IEEE Trans. Inf. Theory, vol. 62, no. 4, pp. 1795–1822, Apr. 2016.

[17] M. Hayashi, “Security analysis of e-almost dual universal hash functions: Smoothing of min entropy versus smoothing of Rényi entropy of order 2,” IEEE Trans. Inf. Theory, vol. 62, no. 6, pp. 3451–3476, Jun. 2016.

[18] M. Hayashi, “Large deviation analysis for quantum security via smoothing of Rényi entropy of order 2,” 2012, arXiv:1202.0322.

[19] K. Li, Y. Yao, and M. Hayashi, “ Tight exponential analysis for smoothing the max-relative entropy and for quantum privacy amplification,” 2021, arXiv:2111.01075.

[20] N. Sharma, “Random coding exponents galore via decoupling,” 2015, arXiv:1504.07075.

[21] K. Li and Y. Yao, “Reliability function of quantum information decoupling via the sandwiched Rényi divergence,” 2021, arXiv:2111.05674.

[22] 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, no. 12, pp. 1–26, Dec. 2013.

[23] 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, no. 2, pp. 593–622, Oct. 2014.

[24] R. L. Frank and R. Lieb, “Monotonicity of a relative Rényi entropy,” J. Math. Phys., vol. 54, no. 12, pp. 1–6, 2013.

[25] S. Beigi, “Sandwiched Rényi divergence satisfies data processing inequality,” J. Math. Phys., vol. 54, no. 12, Dec. 2013, Art. no. 122202.
[26] D. Pérez-García, M. M. Wolf, D. Petz, and M. B. Ruskai, “Contractivity of positive and trace-preserving maps under $l_p$ norms,” *J. Math. Phys.*, vol. 47, no. 8, Aug. 2006, Art. no. 083506.

[27] P. Delgosha and S. Beigi, “Impossibility of local state transformation via hypercontractivity,” *Commun. Math. Phys.*, vol. 332, no. 1, pp. 449–476, Nov. 2014.

[28] F. Dupuis, “Chain rules for quantum Rényi entropies,” *J. Math. Phys.*, vol. 56, no. 2, Feb. 2015, Art. no. 022203.

[29] A. McKinlay and M. Tomamichel, “Decomposition rules for quantum Rényi mutual information with an application to information exclusion relations,” *J. Math. Phys.*, vol. 61, no. 7, Jul. 2020, Art. no. 072202.

[30] M. Reed and B. Simon, *Methods of Modern Mathematical Physics: Fourier Analysis and Self-Adjointness*, vol. 2. New York, NY, USA: Academic, 1975.

[31] M. Berta, O. Fawzi, V. Scholz, and O. Szehr, “Variations on classical and quantum extractors,” in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2014, pp. 1474–1478.

[32] A. Ben-Aroya, O. Schwartz, and A. Ta-Shma, “Quantum expanders: Motivation and constructions,” *Theory Comput.*, vol. 6, pp. 47–79, 2010, Art. no. 3. [Online]. Available: https://theoryofcomputing.org/articles/v006a003/

[33] M. B. Hastings, “Random unitaries give quantum expanders,” *Phys. Rev. A, Gen. Phys.*, vol. 76, no. 3, Sep. 2007, Art. no. 032315.

[34] R. Arnon-Friedman, R. Renner, and T. Vidick, “Simple and tight device-independent security proofs,” *SIAM J. Comput.*, vol. 48, no. 1, pp. 181–225, Jan. 2019.

[35] R. Arnon-Friedman, F. Dupuis, O. Fawzi, R. Renner, and T. Vidick, “Practical device-independent quantum cryptography via entropy accumulation,” *Nature Commun.*, vol. 9, no. 1, p. 459, Jan. 2018.

[36] F. Dupuis and O. Fawzi, “Entropy accumulation with improved second-order term,” *IEEE Trans. Inf. Theory*, vol. 65, no. 11, pp. 7596–7612, Nov. 2019.

[37] P. Sen, “A one-shot quantum joint typicality lemma,” 2018, arXiv:1806.07278.

[38] M. Berta, K. P. Seshadreesan, and M. M. Wilde, “Rényi generalizations of the conditional quantum mutual information,” *J. Math. Phys.*, vol. 56, no. 2, 2015, Art. no. 022205.

[39] P. Brown, H. Fawzi, and O. Fawzi, “Computing conditional entropies for quantum correlations,” *Nature Commun.*, vol. 12, no. 1, p. 575, Jan. 2021.

[40] E. Knill, Y. Zhang, and H. Fu, “Quantum probability estimation for randomness with quantum side information,” 2018, arXiv:1806.04553.

[41] Y. Zhang, H. Fu, and E. Knill, “Efficient randomness certification by quantum probability estimation,” *Phys. Rev. Res.*, vol. 2, no. 1, Jan. 2020, Art. no. 013016.