How much secure randomness is in a quantum state?

Kriss Gutierrez Anco, Tristan Nemoz, and Peter Brown

Télécom Paris, Inria, LTCI, Institut Polytechnique de Paris, 91120 Palaiseau, France

How much cryptographically-secure randomness can be extracted from a quantum state? This fundamental question probes the absolute limits of quantum random number generation (QRNG) and yet, despite the technological maturity of QRNGs, it remains unsolved. In this work we consider a general adversarial model that allows for an adversary who has quantum side-information about both the source and the measurement device. Using links between randomness extraction rates and sandwiched Rényi entropies, we provide compact, easy to compute, achievable rates of secure randomness extraction from quantum states. In turn, this provides a simple to evaluate benchmarking tool for the randomness generation rates of QRNG protocols.

1 Introduction

The intrinsic randomness of quantum systems lies at the heart of quantum cryptography, enabling fundamental cryptographic primitives such as quantum random number generation (QRNG) [1] and quantum key distribution (QKD) [2]. Significant progress has been made on from both the theoretical and experimental perspectives and both QRNG and QKD now exist as commercially viable technologies. Despite this progress, a fundamental question remains unsolved: how much secure randomness can you extract from a quantum state?

Security proofs formalize what we mean by secure randomness with precise mathematical definitions [3]. In turn, these definitions enable us to formally prove that, under the various assumptions of the protocol, the randomness generated by our QRNG is secure. Modern security proof frameworks such as the entropy accumulation theorem [4, 5], rely on quantifying the amount of randomness produced during the protocol with respect to an adversary. This typically amounts to an optimization of a conditional entropy over states that are consistent with the observations and assumptions of the protocol [6]. For example, in device-independent QRNG we minimize some conditional entropy over all states that are consistent with a given Bell-violation [7, 8]. Whereas for device-dependent cryptography, we optimize over states that are consistent with the observations made during the protocol and the trusted components, e.g. the measurement operators are fully characterized [9]. There also exist intermediary scenarios, so-called semi-device-independence, where other assumptions are placed on the systems, e.g. energy bounds [10].

Whilst these security proof frameworks allow us to prove that we can extract secure randomness from a source of quantum states, a pertinent question for the design of efficient QRNGs remains: what is the most amount of secure randomness we can extract? Such a quantifier would provide us with the means to assess the resourcefulness of a given source for practical QRNG and would further provide a benchmarking tool for finite-size security proofs of QRNG protocols.

More concretely, consider the setup depicted in Figure 2 (detailed in Section 4), which is an adversarial model introduced in [11] (see also [12, 13] for more recent developments). In this setup, we consider an honest agent Alice who wishes to generate some secret randomness. To do so, she measures a trusted quantum state ρ_A using a trusted POVM $\{M_x\}_x$ producing a classical random variable X. However, unbeknownst to Alice, an eavesdropper Eve has quantum side information about both Alice's source (which produces ρ_A) and Alice's measurement device (which measures $\{M_x\}_x$). We model this general quantum side information by a tripartite state ρ_{ABE} where the

---

Peter Brown: peter.brown@telecom-paris.fr

system A is held by Alice's source, the system B is held by Alice's measurement device (which it can use to help implement the POVM $\{M_x\}_x$ on A) and the system E is held by Eve. The marginal state ρ_A and the POVM $\{M_x\}_x$ are both trusted and known but the systems held by the measurement device and the adversary are unknown and can in principal be selected by Eve to provide her with the most amount of information about X.

We are interested in the maximal amount of cryptographically secure randomness that Alice can extract from her system A in such a general adversarial model. To extract the randomness from her state Alice first measures her system using producing the classical system X. This results in a classical-quantum post-measurement state ρ_{XE} which captures the potential side-information that Eve has about X. To remove this side-information and extract cryptographically secure randomness, Alice performs a randomness extraction (privacy amplification) protocol on the classical system X, producing k bits of ε-secure randomness. In this work we ask the following question: if we optimize the choice of POVM $\{M_x\}_x$, how many bits of ε-secure randomness can we extract? Or, what's an achievable amount of ε-secure randomness that we can extract from a quantum state ρ_A against a general quantum adversary?

Our scheme, as detailed above, involves Alice first measuring her quantum system and then applying a randomness extraction protocol. We note that a more direct method has been considered in [14] wherein so-called quantum-to-classical randomness extractors are introduced. This can be seen as merging of the two steps in our scheme into a single step. In particular, in their scheme they define a family of quantum to classical channels $\mathcal{E}_i : L(A) \to L(X)$ such that $\mathcal{E}_i$ is selected randomly from the family and applied to the source A. They then show that, for several explicit constructions, the resulting X is already ε-secure randomness.

1.1 Contributions

In this work we derive simple, easy to compute expressions for achievable rates of extraction of cryptographically secure randomness from a quantum state. In particular, if we define the Rényi entropy of system A to be $H_\alpha(A) = \frac{1}{1-\alpha} \log \text{Tr}\left[\rho_A^\alpha\right]$ for $\alpha \in (0,1) \cup (1,\infty)$ then we can show that Alice can extract up to

$$2\log(d_A) - \min_{\alpha \in (1,2]} \left(H_{\frac{\alpha}{2\alpha-1}}(A) - \frac{\alpha}{\alpha-1} \log(\varepsilon) \right) \tag{1}$$

bits of ε-secure randomness from her quantum state, where d_A is the dimension of the system A. In addition, if she restricts herself to projective measurements only then she can extract up to

$$\log(d_A) - \min_{\alpha \in (1,2]} \left(H_{\frac{\alpha}{2\alpha-1}}(A) - \frac{\alpha}{\alpha-1} \log(\varepsilon) \right) \tag{2}$$

bits of ε-secure randomness. These results give simple to evaluate benchmarks for quantum random number generators (QRNG). In particular, given a QRNG that uses a source producing states ρ_A, our work shows that such a QRNG could in principle produce randomness at a rate given by (1), setting a benchmark for the performance of QRNGs and their security proofs.

On the technical level, the majority of this work is dedicated to proving that

$$\sup_{\{M_x\}_x} H_\alpha^\uparrow(X|E) = 2\log(d_A) - H_{\frac{\alpha}{2\alpha-1}}(A) \tag{3}$$

where $H_\alpha^\uparrow$ is the conditional sandwiched Rényi entropy [15, 16] and ρ_{XE} is some post-measurement state defined by the adversarial model (see Section 4 and Figure 2 for details). Along the way to proving this, we define what we call the *maximal intrinsic $\mathbb{H}$-randomness problem* which is the optimization problem

$$\sup_{\{M_x\}_x} \mathbb{H}(X|E) \tag{4}$$

defined for any conditional entropy $\mathbb{H}$. This was inspired by a similar problem introduced recently in [17]. However, in that work, $\mathbb{H}$ is restricted to the von Neumann entropy and the min/max

entropies, $\{M_x\}_x$ is restricted to projective measurements and no entanglement with the measurement device was considered. Generalizing this problem to other conditional entropies allows us to connect to the operationally relevant notion of extractable ε-secure randomness [18]. Furthermore, we are able to establish our results for general measurements in a general quantum adversarial model wherein the source, the measurement device and Eve can all be entangled. Note that when we restrict our optimization to projective measurements we are able to recover the von Neumann entropy and min-entropy results of [17], showing that they also hold in this stronger adversarial model.

The work is structured as follows. In Section 2 we introduce the relevant notation and entropic quantities used throughout the work. In Section 3 we solve a simpler version of the maximal intrinsic $\mathbb{H}$-randomness problem where the measurements are restricted to be projective and no entanglement is shared with the measurement device. In Section 4 we consider the fully general scenario giving closed form expressions for the maximal intrinsic $\mathbb{H}$-randomness problem for the operationally relevant conditional sandwiched Rényi entropies. In Section 5 we connect the solutions of Section 3 and Section 4 to the problem of maximal extractable secure randomness and give a simple to compute benchmark for the performance of QRNGs.

2 Preliminaries

We begin by introducing some notation used throughout the article. Quantum systems and their associated Hilbert spaces are denoted by capital Roman letters, e.g. A or B. All Hilbert spaces considered in this work are finite dimensional. The set of positive semidefinite square matrices acting on a system A are denoted $\mathscr{P}(A)$. Given two Hermitian matrices X, Y acting on a Hilbert space A, we write $X \succeq Y$ if $X - Y \in \mathscr{P}(A)$. A *quantum state* of a system A is a an operator $\rho \in \mathscr{P}(A)$ such that $\mathrm{Tr}\,[\rho] = 1$, the set of quantum states for a system A is denoted $\mathscr{D}(A)$. Given $\rho, \sigma \in \mathscr{P}(A)$ we write $\rho \ll \sigma$ if $\ker(\sigma) \subseteq \ker(\rho)$ where $\ker(X) := \{|\psi\rangle \in A : X |\psi\rangle = 0\}$. We say a bipartite state $\rho_{XE} \in \mathscr{D}(XE)$ is *classical-quantum* (cq) if it can be written in the form $\sum_x |x\rangle\langle x| \otimes \rho_{E,x}$ where $\{|x\rangle\}_x$ is some predefined orthonormal (classical) basis and $\rho_{E,x} \in \mathscr{P}(E)$ such that $\sum_x \rho_{E,x} \in \mathscr{D}(E)$. We take the logarithm to be base 2 (denoted log). For $n \in \mathbb{N}$, we use the notation $[n] = \{0, 1, \ldots, n-1\}$.

We say a collection of operators $\{M_x\}_x$ is a POVM on system A if for all x we have $M_x \in \mathscr{P}(A)$ and $\sum_x M_x = I_A$. If in addition $\{M_x\}_x$ are all projectors then we refer to this as a projective measurement or a PVM. A POVM $\{M_x\}_x$ is called *extremal* if for any POVMs $\{F_x\}_x$, $\{G_x\}_x$ and $0 < \lambda < 1$ such that $M_x = \lambda F_x + (1-\lambda)G_x$ for all x, we must have $F_x = G_x = M_x$. Note that extremal POVMs can have at most d^2 outcomes where d is the dimension of the quantum system [19].

In order to define the Rényi entropy families relevant to this work we must first define the sandwiched Rényi divergences, first introduced in [15, 16].

Definition 2.1 (Sandwiched Rényi divergences). Let $\alpha \in [\frac{1}{2}, 1) \cup (1, \infty)$, let $\rho, \sigma \in \mathscr{P}(A)$ with $\rho \neq 0$ and $\rho \ll \sigma$. Then the *sandwiched Rényi divergence of order* α is defined as

$$D_\alpha(\rho\|\sigma) := \frac{1}{\alpha - 1} \log \frac{\mathrm{Tr}\left[\left(\sigma^{\frac{1-\alpha}{2\alpha}} \rho \sigma^{\frac{1-\alpha}{2\alpha}}\right)^\alpha\right]}{\mathrm{Tr}\,[\rho]}. \tag{5}$$

We can use the above divergence to define two families of conditional entropies.

Definition 2.2 (Conditional sandwiched Rényi entropies). Let $\alpha \in [\frac{1}{2}, 1) \cup (1, \infty)$ and let $\rho_{AB} \in \mathscr{D}(AB)$. Then we define

$$H_\alpha^\uparrow(A|B) := \sup_{\sigma_B \in \mathscr{D}(B)} -D_\alpha(\rho_{AB}\|I_A \otimes \sigma_B) \tag{6}$$

and

$$H_\alpha^\downarrow(A|B) := -D_\alpha(\rho_{AB}\|I_A \otimes \rho_B). \tag{7}$$

We also need a notion of a non-conditional Rényi entropy which we define as follows.

Definition 2.3 (Rényi entropy)**.** Let $\alpha \in (0,1) \cup (1,\infty)$ and $\rho_A \in \mathscr{D}(A)$. Then the Rényi entropy of system A of order α is defined as

$$H_\alpha(A) := \frac{1}{1-\alpha} \log(\text{Tr}\,[\rho_A^\alpha])\,. \tag{8}$$

The cases $\alpha \in \{0,1,\infty\}$ are defined by taking the associated limits which are given by

$$\begin{aligned}
H_0(A) &:= \log(\text{rank}(\rho_A)) \\
H_1(A) &:= -\,\text{Tr}\,[\rho_A \log(\rho_A)] \\
H_\infty(A) &:= -\log(\|\rho_A\|_\infty)\,.
\end{aligned} \tag{9}$$

We will use the notation $H(A) := H_1(A)$ and $H_{\min}(A) := H_\infty(A)$.

In addition, for a bipartite state $\rho_{AB} \in \mathscr{D}(AB)$ we also define the *conditional von Neumann entropy* as

$$H(A|B) := H(AB) - H(B) \tag{10}$$

and the *conditional min-entropy* as

$$H_{\min}(A|B) := -\log \inf\{\lambda \in \mathbb{R} \mid \exists\, \sigma_B \in \mathscr{D}(B) \text{ s.t. } \rho_{AB} \succeq I_A \otimes \sigma_B\}\,. \tag{11}$$

Note that these two conditional entropies are also limits of the conditional sandwiched Rényi entropies. That is, we have

$$\lim_{\alpha \to 1} H_\alpha^\uparrow(A|B) = H(A|B) \qquad \text{and} \qquad \lim_{\alpha \to \infty} H_\alpha^\uparrow(A|B) = H_{\min}(A|B)\,. \tag{12}$$

3 The maximal intrinsic $\mathbb{H}$-randomness problem (PVMs)

Let us begin with a simpler problem, which is a special case of the more general problem that we solve in Section 4. We consider a slightly simpler[1] setup than that which is depicted in Figure 2. Alice is given a source producing a state ρ_A which she then measures using a projective measurement $\{M_x\}_x$ producing a classical random variable X. However, suppose an eavesdropper Eve holds side-information about the source. This side information is modeled by a bipartite quantum state ρ_{AE}, which has the constraint that $\text{Tr}_E\,[\rho_{AE}] = \rho_A$. Note that the dimension of E is a priori unbounded.

After the measurement is performed by Alice, the information is represented by the cq-state

$$\rho_{XE} = \sum_x |x\rangle\langle x| \otimes \rho_{E,x} \tag{13}$$

where $\rho_{E,x} = \text{Tr}_A\,[(M_x \otimes I)\rho_{AE}]$. Using this state we can quantify the information Eve has via a conditional entropy $\mathbb{H}(X|E)$. For the moment, we do not explicitly define $\mathbb{H}$ but we note that the particular cases where $\mathbb{H}$ is the von Neumann entropy and the min-entropy have already been studied in [17].

In order to conclude that the outcome of the measurement $\{M_x\}_x$ on the state ρ_A is *intrinsically random*, we must allow for any possible side information held by Eve. That is we look to solve the optimization problem

$$\begin{aligned}
\inf \quad & \mathbb{H}(X|E) \\
\text{s.t.} \quad & \rho_{XE} = \sum_x |x\rangle\langle x| \otimes \text{Tr}_A\,[(M_x \otimes I)\rho_{AE}]\,, \\
& \text{Tr}_E\,[\rho_{AE}] = \rho_A\,, \qquad \rho_{AE} \in \mathscr{D}(AE)\,,
\end{aligned} \tag{14}$$

where we minimize over all possible E systems that could be held by the adversary. The following lemma (see also [20]) shows that it is sufficient to consider any purification ρ_{AE} of ρ_A.

---

[1]In particular, we remove the entanglement with the measurement device and we assume that the measurement $\{M_x\}_x$ is projective.

Lemma 3.1. *Let $\mathbb{H}$ be a conditional entropy satisfying:*

1. *(Strong subadditivity)* $\mathbb{H}(A|BC) \leq \mathbb{H}(A|B)$

2. *(Isometric invariance) For any isometry $V : B \to B'$ we have $\mathbb{H}(A|B)_{\rho_{AB}} = \mathbb{H}(A|B')_{V\rho_{AB}V^\dagger}$*

then the optimal solution to the optimization problem (14) is obtained by any purification ρ_{AE} of ρ_A.

Proof. Take any feasible point of (14) such that ρ_{AE} is a mixed state. Then if $\rho_{AEE'}$ is a purification of ρ_{AE} we have $\mathbb{H}(X|EE') \leq \mathbb{H}(X|E)$ by strong subadditivity. Thus we can always restrict to pure ρ_{AE} when minimizing.

Any two purifications ρ_{AE}, $\rho_{AE'}$ of ρ_A can be related by an isometry on the purifying system. I.e., suppose without loss of generality that $\dim(E) \leq \dim(E')$. Then there exists an isometry $V : E \to E'$ such that $\rho_{AE'} = (I_A \otimes V)\rho_{AE}(I_A \otimes V^\dagger)$. As the measurement occurs only on the A system we then also have $(I_X \otimes V)\rho_{XE}(I_X \otimes V^\dagger) = \rho_{XE'}$. By the isometric invariance property of $\mathbb{H}$ we therefore have $\mathbb{H}(X|E) = \mathbb{H}(X|E')$. Thus all purifications result in the same optimal value. $\qquad\square$

Lemma 3.1 enables us to quantify the intrinsic randomness of a measurement $\{M_x\}_x$ on a state ρ_A as quantified by $\mathbb{H}$ by evaluating $\mathbb{H}(X|E)$ where E is any purification of ρ_A.[2] In this work we are interested in what we call the *maximal intrinsic $\mathbb{H}$-randomness* of the state ρ_A, as such we allow ourselves to maximize $\mathbb{H}(X|E)$ over all projective measurements. This leads to the following definition.

Definition 3.2 (Maximal intrinsic $\mathbb{H}$-randomness (projective))**.** Let $\mathbb{H}$ be a conditional entropy, let ρ_A be a quantum state and let ρ_{AE} be any purification of ρ_A. Then we define the *maximal intrinsic $\mathbb{H}$-randomness for projective measurements* of ρ_A as

$$
\begin{aligned}
\mathcal{R}_{\mathbb{H}}^{\mathrm{PVM}}(\rho_A) := \sup_{\{M_x\}_x} \quad & \mathbb{H}(X|E) \\
\text{s.t.} \quad & \rho_{XE} = \sum_x |x\rangle\langle x| \otimes \mathrm{Tr}_A\left[(M_x \otimes I)\rho_{AE}\right], \\
& \{M_x\}_x \text{ is a projective measurement.}
\end{aligned}
\tag{15}
$$

In [17], closed form expressions for the von Neumann entropy H and the min-entropy $H_{\min}$ were derived. In particular, the authors found that

$$
\mathcal{R}_H^{\mathrm{PVM}}(\rho_A) = \log(d_A) - H(A)
\tag{16}
$$

and

$$
\mathcal{R}_{H_{\min}}^{\mathrm{PVM}}(\rho_A) = \log(d_A) - 2\log \mathrm{Tr}\left[\sqrt{\rho_A}\right].
\tag{17}
$$

In the following subsection we will generalize these results to the sandwiched Rényi entropies $H_\alpha^\uparrow$ and $H_\alpha^\downarrow$. Using this we can recover the results from [17] by taking appropriate limits.

3.1 Maximal intrinsic Rényi-randomness (PVMs)

To begin we have the following result that tells us we can assume, without loss of generality, that our optimal projective measurement is a rank-one projective measurement.

---

[2]This also implies that without loss of generality we can assume that $\dim(A) = \dim(E)$.

Lemma 3.3 (Restriction to rank-one measurements). *Let $\alpha \in [1/2, 1) \cup (1, \infty)$, then for $\mathbb{H}$ either $H_\alpha^\downarrow$ or $H_\alpha^\uparrow$ we have*

$$\mathcal{R}_{\mathbb{H}}^{\text{PVM}}(\rho_A) = \sup_{\{M_x\}_x} \quad \mathbb{H}(X|E)$$

$$\text{s.t.} \quad \rho_{XE} = \sum_x |x\rangle\langle x| \otimes \text{Tr}_A\left[(M_x \otimes I)\rho_{AE}\right], \tag{18}$$

$$\{M_x\}_x \text{ is a rank-one projective measurement.}$$

Proof. Consider any PVM $\{N_x\}_x$ (not necessarily rank-one) and let $r = \max_x \text{rank}(N_x)$. Then there exists a collection of projectors $\{M_{x,y}\}_{x,y}$ such that $\sum_{y=1}^r M_{x,y} = N_x$ such that each $M_{x,y}$ is either a rank-one projector or 0. If we define the two classical-quantum states

$$\rho_{XE} = \sum_x |x\rangle\langle x| \otimes \text{Tr}_A\left[(N_x \otimes I_E)\rho_{AE}\right]$$

$$\rho_{XYE} = \sum_{xy} |xy\rangle\langle xy| \otimes \text{Tr}_A\left[(M_{x,y} \otimes I_E)\rho_{AE}\right] \tag{19}$$

then $\text{Tr}_Y[\rho_{XYE}] = \rho_{XE}$. By Lemma D.1 we know that for all $\alpha \in [1/2, 1) \cup (1, \infty)$ we have

$$H_\alpha^\downarrow(XY|E) \geq H_\alpha^\downarrow(X|E) \quad \text{and} \quad H_\alpha^\uparrow(XY|E) \geq H_\alpha^\uparrow(X|E). \tag{20}$$

Overall this shows that for any non rank-one measurement we can find a rank-one measurement whose conditional entropy is at least as large and hence the result follows. $\square$

Using the above lemma we can restrict the intrinsic randomness to an optimization over rank-one projective measurements. This gives us sufficient structure to prove the following closed form expressions of the intrinsic sandwiched Rényi-randomness, the proof is given in Appendix A.

Theorem 3.4. *Let $\rho_A \in \mathscr{D}(A)$ and let d_A be the dimension of system A. Then we have*

$$\mathcal{R}_{H_\alpha^\uparrow}^{\text{PVM}}(\rho_A) = \log(d_A) - H_{\frac{\alpha}{2\alpha-1}}(A) \tag{21}$$

for all $\alpha > 1$ and we have

$$\mathcal{R}_{H_\alpha^\downarrow}^{\text{PVM}}(\rho_A) = \log(d_A) - H_{\frac{1}{\alpha}}(A). \tag{22}$$

for all $\alpha \in [1/2, 1) \cup (1, \infty)$.

We note that the proof of Theorem 3.4 can also be extended to the limits $\alpha \to 1$ and $\alpha \to \infty$. In particular, by taking these limits we recover the results

$$\mathcal{R}_H^{\text{PVM}}(\rho_A) = \log(d_A) - H(A) \quad \text{and} \quad \mathcal{R}_{H_{\min}}^{\text{PVM}}(\rho_A) = \log(d_A) - 2\log\text{Tr}\left[\sqrt{\rho_A}\right] \tag{23}$$

which were first derived in [17]. Thus our results for $\mathcal{R}_{H_\alpha^\uparrow}^{\text{PVM}}(\rho_A)$ nicely interpolate between the results of [17], can be derived using a single proof technique and give a clear interpretation for the form of $2\log\text{Tr}\left[\sqrt{\rho_A}\right] = H_{1/2}(A)$.

Example 3.5. Consider the qubit state $\rho_A = \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|+\rangle\langle +|$, whose spectrum is equal to $\{\cos^2(\pi/8), \sin^2(\pi/8)\}$. In Figure 1 we plot $\mathcal{R}_{H_\alpha^\downarrow}^{\text{PVM}}(\rho_A)$ and $\mathcal{R}_{H_\alpha^\uparrow}^{\text{PVM}}(\rho_A)$ for a range of α. In particular we can see that $\mathcal{R}_{H_\alpha^\downarrow}^{\text{PVM}}(\rho_A) \leq \mathcal{R}_{H_\alpha^\uparrow}^{\text{PVM}}(\rho_A)$, with an equality in the limit $\alpha \to 1$. However, for other values of α they diverge with

$$\lim_{\alpha \to \infty} \mathcal{R}_{H_\alpha^\uparrow}^{\text{PVM}}(\rho_A) = \mathcal{R}_{H_2^\downarrow}^{\text{PVM}}(\rho_A) \tag{24}$$

and

$$\lim_{\alpha \to \infty} \mathcal{R}_{H_\alpha^\downarrow}^{\text{PVM}}(\rho_A) = \log(d_A) - \log\text{rank}(\rho_A), \tag{25}$$

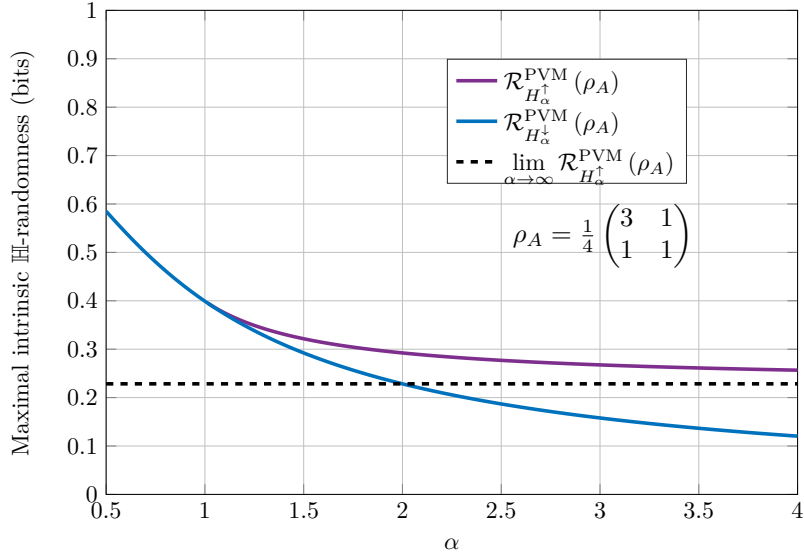which evaluates to 0 in this example as $\text{rank}(\rho_A) = d_A$.

Figure 1: A plot of $\mathcal{R}_{H_\alpha^\downarrow}(\rho_A)$ and $\mathcal{R}_{H_\alpha^\downarrow}(\rho_A)$ for the state $\rho_A = \frac{1}{4}\begin{pmatrix} 3 & 1 \\ 1 & 1 \end{pmatrix}$.

4   The maximal $\mathbb{H}$-randomness problem

We now want to generalize the setup considered in the previous section to the case where we allow Alice to measure any POVM $\{M_x\}_x$. However, this generalization is not as immediate as just replacing the maximization over projective measurement to over POVMs. To see that there are issues consider the trivial measurement $\{I_A/n, I_A/n, \ldots, I_A/n\}$, for any state ρ_{AE} we will have that $H(X|E) = \log(n)$, in particular we can make this as large as possible by making n large despite the measurement outcome carrying no information about the underlying state.

We use a framework introduced in [11] where the adversary may share entanglement with the state used in the Naimark dilation of the POVM, see also [12, 13]. In particular, we now not only allow Eve to be entangled with the source, but we now allow for arbitrary entanglement between the source, the measurement device and Eve (see also Figure 2 for a graphical depiction). That is, we now introduce a quantum system B that is held by the measurement device and can be used to help implement the POVM $\{M_x\}_x$. Thus, the shared state before the measurement is some $\rho_{ABE} \in \mathscr{D}(ABE)$ with the marginal ρ_A still fixed. When the measurement device receives the system A, the most general thing it can do is apply a quantum-to-classical channel $\mathcal{E} : L(AB) \to L(X)$ to produce the measurement outcomes. Such a channel can be shown to take the form

$$\mathcal{E}(\rho_{AB}) = \sum_x \mathrm{Tr}\left[\rho_{AB}N_x\right]|x\rangle\langle x| \tag{26}$$

where N_x is some POVM on AB. Importantly, however, we trust that the measurement device is actually implementing a fixed POVM $\{M_x\}_x$. Thus, we must have that for all $\sigma_A \in \mathscr{D}(A)$,

$$\mathrm{Tr}\left[(\sigma_A \otimes \rho_B)N_x\right] = \mathrm{Tr}\left[\sigma_A M_x\right] \tag{27}$$

or equivalently, $\mathrm{Tr}_B\left[N_x(I_A \otimes \rho_B)\right] = M_x$. By considering a Naimark dilation this is equivalent to there existing a projective measurement $\{P_x\}_x$ that acts on the joint system AB[3] which satisfies what we refer to as the *consistency condition* [13]:

$$\mathrm{Tr}_B\left[P_x(I_A \otimes \rho_B)\right] = M_x. \tag{28}$$

---

[3]The dimension of B may change and hence ρ_B also but as we don't specify its dimension there are no issues here.
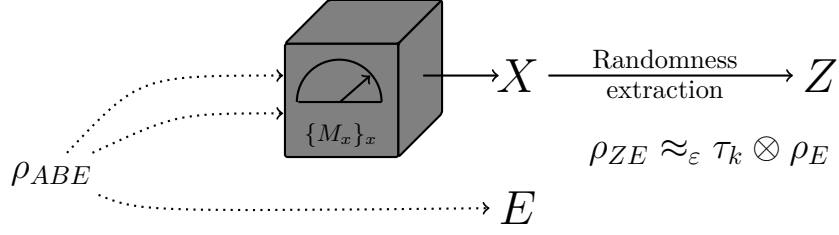
Figure 2: **Schematic of maximal extractable randomness.** We consider a setup where a trusted state $\rho_A \in \mathscr{D}(A)$ is measured by some trusted measurement $\{M_x\}_x$ producing a random variable X. However, we allow for the state to also be entangled with a system B which is held by the measurement device and a system E that is held by the adversary. Thus the joint state ρ_{ABE} is arbitrary apart from the condition that the marginal on A is fixed to ρ_A. Then, after the cq-state ρ_{XE} is produced, a randomness extraction protocol is applied to X to produce a new state ρ_{ZE} where Z is an ε-secure string of k uniformly random bits. If we fix ρ_A, how large can we make k?

Note that such Naimark dilations are not unique and hence we further allow Eve to choose $\{P_x\}_x$ and ρ_B (equivalently ρ_{ABE}) in order to minimize her uncertainty. Therefore, this new setting begins with a state ρ_{ABE} where the system E is still held by Eve. Then the state after Alice's measurement device implements the POVM $\{M_x\}_x$ is given by

$$\rho_{XE} = \sum_x |x\rangle\langle x| \otimes \mathrm{Tr}_{AB}\left[(P_x \otimes I_E)\rho_{ABE}\right] . \tag{29}$$

In analogy to (14), for a fixed state ρ_A, a fixed POVM $\{M_x\}$ and a conditional entropy $\mathbb{H}$, we can define the intrinsic $\mathbb{H}$-randomness by the optimization

$$
\begin{aligned}
\inf \quad & \mathbb{H}(X|E) \\
\text{s.t.} \quad & \rho_{XE} = \sum_x |x\rangle\langle x| \otimes \mathrm{Tr}_{AB}\left[(P_x \otimes I_E)\rho_{ABE}\right], \\
& \mathrm{Tr}_{BE}\left[\rho_{ABE}\right] = \rho_A, \quad \rho_{ABE} \in \mathscr{D}(AE), \\
& \mathrm{Tr}_B\left[P_x(I_A \otimes \rho_B)\right] = M_x, \\
& \{P_x\}_x \text{ is a projective measurement.}
\end{aligned}
\tag{30}
$$

This optimization is evidently more complex that (14) as the optimization is now not only over all extensions of ρ_A, but also over all Naimark dilations of $\{M_x\}_x$. However, importantly, using the same proof technique as Lemma 3.1 we can show that without loss of generality we can take ρ_{ABE} to be a purification of ρ_{AB}. Thus, the remaining optimization is just over all Naimark dilations $(\{P_x\}_x, \rho_B)$ and we can take $\dim(E) = \dim(AB)$. Note that any valid Naimark dilation $(\{P_x\}_x, \rho_B)$ of $\{M_x\}_x$ gives a feasible point as we can always take ρ_{ABE} to be a purification of $\rho_A \otimes \rho_B$. However, the optimization is over all possible ρ_{ABE} and hence there can in principle be entanglement between A and B.

Using the observation regarding purifications, we arrive at our definition of *intrinsic randomness* for a pair $(\rho_A, \{M_x\}_x)$.

Definition 4.1. Let $\rho_A \in \mathscr{D}(A)$ and let $\{M_x\}_x$ be a POVM on the system A. Then the *intrinsic $\mathbb{H}$-randomness* of the pair $(\rho_A, \{M_x\}_x)$ is defined as

$$
\begin{aligned}
\mathcal{I}_{\mathbb{H}}^{\mathrm{POVM}}(\rho_A, \{M_x\}_x) := \inf \quad & \mathbb{H}(X|E) \\
\text{s.t.} \quad & \rho_{XE} = \sum_x |x\rangle\langle x| \otimes \mathrm{Tr}_{AB}\left[(P_x \otimes I_E)|\rho\rangle\langle\rho|_{ABE}\right], \\
& \mathrm{Tr}_B\left[\rho_{AB}\right] = \rho_A, \quad \rho_{AB} \in \mathscr{D}(AB), \\
& \mathrm{Tr}_B\left[P_x(I_A \otimes \rho_B)\right] = M_x, \\
& \{P_x\}_x \text{ is a projective measurement.}
\end{aligned}
\tag{31}
$$

where $|\rho\rangle_{ABE} = \sum_k \sqrt{\rho_{AB}}\,|k\rangle_{AB} \otimes |k\rangle_E$ is a purification of ρ_{AB}.

This simplifies the optimization somewhat by reducing it to an optimization just over Naimark dilations and bipartite states ρ_{AB}. However, the optimization over all Naimark dilations is a priori difficult due to the possibly unbounded dimension of B. Our goal will be to tame this optimization for a special class of POVMs, namely extremal rank-one POVMs. Before that, we show that the adversarial model takes care of the trivial measurement problem highlighted at the beginning of the section.

Example 4.2 (Trivial POVMs produce no intrinsic randomness)**.** Consider the trivial POVM $\{M_x\}_x$ where $M_x = p(x)I_A$ with $p(x)$ being some probability distribution. Such a POVM can be implemented with the Naimark dilation

$$P_x = I_A \otimes |x\rangle\langle x| \qquad \text{and} \qquad \rho_B = \sum_x p(x) |x\rangle\langle x| \ . \tag{32}$$

Define $|\psi\rangle_{BE} = \sum_x \sqrt{p(x)} |xx\rangle$ to be a purification of ρ_B and let the tripartite state be $\rho_A \otimes |\psi\rangle\langle\psi|_{BE}$ then the post-measurement state takes the form

$$\rho_{XE} = \sum_x p(x) |x\rangle\langle x|_X \otimes |x\rangle\langle x|_E \tag{33}$$

which gives $\mathbb{H}(X|E) = 0$. Note in this example we did not require Eve to share entanglement with ρ_A, although this is natural as the cq-state ρ_{XE} is independent of ρ_A.

Using Definition 4.1 we can define the maximal intrinsic randomness of the state ρ_A for POVMs (analogous to Definition 3.2) by maximizing over all POVMs.

Definition 4.3 (Maximal intrinsic $\mathbb{H}$-randomness)**.** Let $\rho_A \in \mathscr{D}(A)$ and let $\mathbb{H}$ be a conditional entropy. Then we define the *maximal intrinsic $\mathbb{H}$-randomness of ρ_A* as

$$\mathcal{R}_{\mathbb{H}}^{\text{POVM}} (\rho_A) = \sup_{\{M_x\}_x} \ \mathcal{I}_{\mathbb{H}}^{\text{POVM}} (\rho_A, \{M_x\}_x) \ . \tag{34}$$

In a similar fashion to Section 3 we now look to find closed form solutions for $\mathcal{R}_{\mathbb{H}}^{\text{POVM}} (\rho_A)$ for the special cases where $\mathbb{H}$ corresponds to a sandwiched Rényi entropy.

4.1 Maximal Intrinsic Rényi randomness

In order to find closed form expressions for $\mathcal{R}_{\mathbb{H}}^{\text{POVM}} (\rho_A)$, we need a way of solving the minimization over the Naimark dilations implicit in the definition. Whilst it is not obvious how to solve such a problem in general, it would be sufficient if we could just solve the problem for the subset of measurements $\{M_x\}_x$ that maximize $\mathcal{I}_{\mathbb{H}}^{\text{POVM}} (\rho_A, \{M_x\}_x)$. The following lemma shows that we can consider rank-one extremal POVMs without loss of generality.

Lemma 4.4 (Reduction to extremal rank-one POVMs)**.** *Let $\alpha \in [\frac{1}{2}, 1) \cup (1, \infty)$, let $\mathbb{H}$ be either $H_\alpha^\downarrow$ or $H_\alpha^\uparrow$ and let be $\rho_A \in \mathscr{D}(A)$. Then*

$$\mathcal{R}_{\mathbb{H}}^{\text{POVM}} (\rho_A) = \sup_{\substack{\text{Extremal rank-one} \\ \{M_x\}_x}} \mathcal{I}_{\mathbb{H}}^{\text{POVM}} (\rho_A, \{M_x\}_x) \ . \tag{35}$$

That is, we can restrict our optimization to extremal rank-one POVMs without loss of generality.

Proof. We begin by proving that $\{M_x\}_x$ can be assumed to be extremal. Suppose that $\{M_x\}_x$ is not extremal, then there exists a convex decomposition

$$M_x = \sum_k \lambda_k N_{x,k} \tag{36}$$

where for each k, $\{N_{x,k}\}_x$ is an extremal POVM and $\{\lambda_k\}_k$ form a probability distribution. By Lemma B.1 we know that the map $2^{\mathcal{I}_{\mathbb{H}}^{\text{POVM}}(\rho_A, \{M_x\}_x)}$ is convex on the set of POVMs and hence we have

$$2^{\mathcal{I}_{\mathbb{H}}^{\text{POVM}}(\rho_A, \{M_x\}_x)} \leq \sum_k \lambda_k \, 2^{\mathcal{I}_{\mathbb{H}}^{\text{POVM}}(\rho_A, \{N_{x,k}\}_x)} \leq \max_k \ 2^{\mathcal{I}_{\mathbb{H}}^{\text{POVM}}(\rho_A, \{N_{x,k}\}_x)} \ . \tag{37}$$

Since log is a monotonically increasing function we therefore have that

$$\mathcal{I}_{\mathbb{H}}^{\mathrm{POVM}}\left(\rho_A, \{M_x\}_x\right) \leq \max_k \ \mathcal{I}_{\mathbb{H}}^{\mathrm{POVM}}\left(\rho_A, \{N_{x,k}\}_x\right). \tag{38}$$

Thus we can always find an extremal POVM that obtains an intrinsic randomness at least as large as any non-extremal POVM and hence we can restrict the optimization to extremal POVMs.

Now suppose $\{M_x\}_x$ is extremal but not rank-one. Consider the spectral decomposition $M_x = \sum_y \lambda_{x,y} |v_{x,y}\rangle\langle v_{x,y}|$ and define a new rank-one POVM $N_{x,y} = \lambda_{x,y} |v_{x,y}\rangle\langle v_{x,y}|$ where the (x,y) range only over indexes where $\lambda_{x,y} > 0$. By Lemma D.3 this is also an extremal POVM. Now consider any dilation $(\{P_{x,y}\}_{x,y}, \sigma_B)$ for $\{N_{x,y}\}_{x,y}$, if we define $Q_x = \sum_y P_{x,y}$ then

$$\mathrm{Tr}_B\left[Q_x(I_A \otimes \sigma_B)\right] = \sum_y \mathrm{Tr}_B\left[P_{x,y}(I_A \otimes \sigma_B)\right] = \sum_y N_{x,y} = M_x \tag{39}$$

thus $(\{Q_x\}_x, \sigma_B)$ is a dilation for M_x. But then we have

$$\inf_{(\{P_{x,y}\}_{x,y}, \sigma_B)} \mathbb{H}(XY|E) \geq \inf_{(\{Q_x\}_x, \sigma_B)} \mathbb{H}(X|E) \tag{40}$$

which follows from Lemma D.1. Overall we have shown that given any POVM we can always find a rank-one extremal POVM that has an intrinsic randomness at least as large. $\square$

Thus we have reduced the optimization to extremal rank-one POVMs. It turns out, as the following lemma implies, that for extremal POVMs the infimum over Naimark dilations is trivial, in the sense that we can without loss of generality assume that B is entangled with neither A nor E. The proof of this lemma is deferred to Appendix B (see Lemma B.3).

Lemma 4.5. *Let $\{M_x\}_x$ be an extremal rank-one POVM on a system A, let $(\{P_x\}_x, \rho_B)$ be any dilation of $\{M_x\}_x$ and let $\rho_{ABE} \in \mathscr{D}(ABE)$ be any pure state with marginals ρ_A and ρ_B. Then the post-measurement state*

$$\rho_{XE} = \sum_x |x\rangle\langle x| \otimes \mathrm{Tr}_{AB}\left[(P_x \otimes I_E)\rho_{ABE}\right] \tag{41}$$

can be written as

$$\rho_{XE} = \sum_x |x\rangle\langle x| \otimes \mathrm{Tr}_A\left[(M_x \otimes I_E)\rho_{AE}\right]. \tag{42}$$

Note that due to the a priori potential for entanglement between ABE the above lemma is not trivial. For instance, if we take Example 4.2 then the considered POVM is not extremal and the above statement does not hold. Explicitly, we find that

$$\mathrm{Tr}_{AB}\left[(P_x \otimes I_E)\rho_{ABE}\right] = |x\rangle\langle x|_E \qquad \neq \qquad \mathrm{Tr}_A\left[(M_x \otimes I_E)\right] = \tfrac{1}{n}\rho_E. \tag{43}$$

Overall, the above lemma implies that for extremal rank-one POVMs we can completely ignore the optimization over Naimark dilations. Combined with Lemma 4.4, this reduces the problem of computing $\mathcal{R}_{\mathbb{H}}^{\mathrm{POVM}}\left(\rho_A\right)$ significantly. In effect, we have shown that

$$
\begin{aligned}
\mathcal{R}_{\mathbb{H}}^{\mathrm{POVM}}\left(\rho_A\right) = \sup \quad & \mathbb{H}(X|E) \\
\mathrm{s.t.} \quad & \rho_{XE} = \sum_x |x\rangle\langle x| \otimes \mathrm{Tr}_A\left[(M_x \otimes I)\,|\rho\rangle\langle\rho|_{AE}\right] \\
& \{M_x\}_x \text{ is an extremal rank-one POVM}
\end{aligned}
\tag{44}
$$

where $|\rho\rangle_{AE}$ is any purification of ρ_A. This is effectively the same problem as was solved in Section 3 but we optimize over all extremal rank-one POVMs rather than just rank-one PVMs. This gives us enough structure to finally solve the optimization analytically. This results in the following theorem, the proof of which is deferred to Appendix B.

Theorem 4.6. *Let $\rho_A \in \mathscr{D}(A)$ and let d_A be the dimension of system A. Then we have*

$$\mathcal{R}^{\mathrm{POVM}}_{H^\uparrow_\alpha}(\rho_A) = 2\log(d_A) - H_{\frac{\alpha}{2\alpha-1}}(A) \tag{45}$$

for all $\alpha > 1$ and we have

$$\mathcal{R}^{\mathrm{POVM}}_{H^\downarrow_\alpha}(\rho_A) = 2\log(d_A) - H_{\frac{1}{\alpha}}(A). \tag{46}$$

for all $\alpha \in [1/2, 1) \cup (1, \infty)$.

It is worth making a few remarks about this result. Firstly, we have that

$$\mathcal{R}^{\mathrm{POVM}}_{\mathbb{H}}(\rho_A) = \log(d_A) + \mathcal{R}^{\mathrm{PVM}}_{\mathbb{H}}(\rho_A). \tag{47}$$

This increase in the maximal amount of intrinsic randomness is due to the fact that extremal POVMs can have up to d_A^2 outcomes whereas projective measurements have at most d_A outcomes. Furthermore, whilst we did not consider Eve to have side information on the measurement device when deriving the result for $\mathcal{R}^{\mathrm{PVM}}_{\mathbb{H}}(\rho_A)$, our construction shows that even if we were to include this side information in the analysis of Section 3 then this would not change the value of $\mathcal{R}^{\mathrm{PVM}}_{\mathbb{H}}(\rho_A)$. This is because rank-one projective measurements are extremal and hence the above reductions apply. This shows that the results of [17] also hold in the presence of side-information with the measurement device.

Another interesting point is that whilst $\mathcal{R}^{\mathrm{PVM}}_{\mathbb{H}}(\rho_A)$ can be equal to 0 (e.g. for the maximally mixed state) we instead have that $\mathcal{R}^{\mathrm{POVM}}_{\mathbb{H}}(\rho_A) \geq \log(d_A)$. Thus for all quantum states, it is possible to extract a significant amount of randomness from them. However, it should be noted that achieving this theoretical maximum may be extremely challenging in practice this is because, whilst the maximal intrinsic randomness $\mathcal{R}^{\mathrm{POVM}}_{\mathbb{H}}(\rho_A)$ is continuous in ρ_A, the intrinsic randomness $\mathcal{I}^{\mathrm{POVM}}_{\mathbb{H}}(\rho_A, \{M_x\}_x)$ is not continuous on the set of POVMs for a fixed state ρ_A.

The possibility for discontinuities is evident from the proof of Theorem 4.6 (see Theorem B.4 for the proof). In the proof we establish the existence of a sequence of extremal rank-one POVMs that achieve an intrinsic randomness rate approaching $\mathcal{R}^{\mathrm{POVM}}_{\mathbb{H}}(\rho_A)$. However, the limiting POVM for this sequence is not extremal and actually has an intrinsic randomness rate far lower than $\mathcal{R}^{\mathrm{POVM}}_{\mathbb{H}}(\rho_A)$. Thus our proof says nothing about the achievability of $\mathcal{R}^{\mathrm{POVM}}_{\mathbb{H}}(\rho_A)$. Nevertheless we can create examples in which it is achievable, for instance consider a qubit state $\rho_A \in \mathscr{D}(\mathbb{C}^2)$ and suppose it has a spectral decomposition $\rho_A = \lambda_0 |0\rangle\langle 0| + \lambda_1 |1\rangle\langle 1|$ with $\lambda_0 \geq 1/2$. Then the POVM $\{|\psi_x\rangle\langle\psi_x|\}_x$ with

$$
\begin{aligned}
|\psi_0\rangle &= \frac{1}{2\sqrt{\lambda_0}} |0\rangle \\
|\psi_1\rangle &= \sqrt{\frac{4\lambda_0 - 1}{12\lambda_0}} |0\rangle + \sqrt{\frac{1}{3}} |1\rangle \\
|\psi_2\rangle &= \sqrt{\frac{4\lambda_0 - 1}{12\lambda_0}} |0\rangle + \sqrt{\frac{1}{3}} e^{\frac{2\mathrm{i}\pi}{3}} |1\rangle \\
|\psi_3\rangle &= \sqrt{\frac{4\lambda_0 - 1}{12\lambda_0}} |0\rangle + \sqrt{\frac{1}{3}} e^{\frac{4\mathrm{i}\pi}{3}} |1\rangle
\end{aligned}
\tag{48}
$$

is an extremal rank-one POVM and moreover, $\mathbb{H}(X|E) = \mathcal{R}^{\mathrm{POVM}}_{\mathbb{H}}(\rho_A)$. Thus, for all qubit states the maximal intrinsic $\mathbb{H}$-randomness is achievable.

Discontinuities in $\mathcal{I}^{\mathrm{POVM}}_{\mathbb{H}}(\rho_A, \{M_x\}_x)$ implies an interesting physical problem. In particular, the POVMs used in the proof of Theorem 4.6 converge to a discontinuity and hence to get high intrinsic randomness rates with these POVMs may require extreme precision in order to not encounter the discontinuity. It would be interesting to understand two-fold:

1. Are there POVMs for all states ρ_A that achieve $\mathcal{R}^{\mathrm{POVM}}_{\mathbb{H}}(\rho_A)$?

2. Are there POVMs that achieve a high intrinsic randomness rate $\mathcal{I}^{\mathrm{POVM}}_{\mathbb{H}}(\rho_A, \{M_x\}_x)$ and are far from any discontinuities?

In particular a positive and constructive answer to the second question would provide a much more experimentally-friendly method of producing high randomness rates as it would be able to tolerate natural imprecisions in the measurement device.

5 Application: How much cryptographically-secure randomness can you extract from ρ_A?

Conditional entropies allow us to quantify the amount of information a potential eavesdropper may have about our measurement outcome. However, if we consider randomness generation as a cryptographic protocol then we look to satisfy some formal security definition (see [3] for details). If ρ_{ZE} represents the final state of our randomness generation protocol, where Z is our random string and E is a potential adversary's system, then we say the protocol is ε-secure (for some $\varepsilon > 0$) if

$$\frac{1}{2}\|\rho_{ZE} - \frac{I_Z}{d_Z} \otimes \rho_E\|_1 \leq \varepsilon \tag{49}$$

where d_Z is the dimension of the Z register.[4] I.e., the final state of the protocol ρ_{ZE} should be basically indistinguishable from secret, uniform randomness.

For a general quantum state ρ_A, the ρ_{XE} corresponding to the cq-state after measuring a POVM $\{M_x\}_x$ on A will not satisfy the above definition. This is because the adversary will still, in general, have some partial knowledge about X. However, a procedure known as randomness extraction (or privacy amplification) allows us to extract the randomness from the large partially random X system into a new Z system that is ε-close to uniformly random for any $\varepsilon > 0$ of our choosing (i.e. satisfies the security definition in (49)). The amount of uniform randomness we can extract, i.e. the number of bits in the system Z, is quantified by certain conditional entropies. In particular we have the following result of [18] which shows that the sandwiched Rényi entropies capture an achievable rate of randomness extraction.

Lemma 5.1 (Dupuis [18])**.** *Let $\mathcal{F}$ be a family of 2-universal hash functions from X to Z. Then for any $\alpha \in (1,2]$ and cq-state $\rho_{XE} \in \mathscr{D}(XE)$*

$$\frac{1}{2}\|\rho_{\mathcal{F}(XS)E} - \frac{I_Z}{d_Z} \otimes \rho_{SE}\|_1 \leq 2^{2/\alpha - 2} 2^{\frac{\alpha-1}{\alpha}(\log(d_Z) - H_\alpha^\uparrow(X|E))} \tag{50}$$

where S is the seed used to sample the family.

This lemma indicates that we can extract approximately $H_\alpha^\uparrow(X|E)$ uniformly random bits from X. In particular if we fix α and choose the output length of the extractor to be $H_\alpha^\uparrow(X|E) + \frac{\alpha}{\alpha-1}\log(\varepsilon)$ then the output of the extractor is ε-secure according to the above definition. This leads to the main result of this work.

Theorem 5.2 (Main result)**.** *Let $\varepsilon > 0$, $\rho_A \in \mathscr{D}(A)$ then it is possible to extract l bits of ε-secure randomness from ρ_A for any $l \in \mathbb{N}$ satisfying*

$$l < 2\log(d_A) - \min_{\alpha \in (1,2]} \left(H_{\frac{\alpha}{2\alpha-1}}(A) - \frac{\alpha}{\alpha-1}\log(\varepsilon) \right). \tag{51}$$

Moreover, given $n \in \mathbb{N}$ copies of ρ_A it is possible to extract l_n bits of ε-secure randomness from $\rho_A^{\otimes n}$ for any $l_n \in \mathbb{N}$ satisfying

$$\frac{l_n}{n} < 2\log(d_A) - \min_{\alpha \in (1,2]} \left(H_{\frac{\alpha}{2\alpha-1}}(A) - \frac{\alpha}{n(\alpha-1)}\log(\varepsilon) \right). \tag{52}$$

Proof. By Theorem 4.6 for any $\delta > 0$ and $\alpha \in (1,2]$ there exists a POVM $\{M_x\}_x$ such that the post-measurement state ρ_{XE} satisfies

$$H_\alpha^\uparrow(X|E) \geq 2\log(d_A) - H_{\frac{\alpha}{2\alpha-1}}(A) - \delta. \tag{53}$$

---

[4]Here we are assuming the protocol never aborts.

Thus, if we choose the output length of the extractor to be any $l \leq 2\log(d_A) - H_{\frac{\alpha}{2\alpha-1}}(A) - \delta + \frac{\alpha}{\alpha-1}\log(\varepsilon)$, then by Lemma 5.1 we have

$$
\begin{aligned}
\frac{1}{2}\|\rho_{\mathcal{F}(XS)E} - \frac{I_Z}{d_Z} \otimes \rho_{SE}\|_1 &\leq 2^{2/\alpha-2}2^{\frac{\alpha-1}{\alpha}(\log(d_Z)-H_\alpha^\uparrow(X|E))} \\
&\leq 2^{2/\alpha-2}2^{\frac{\alpha-1}{\alpha}(2\log(d_A)-H_{\frac{\alpha}{2\alpha-1}}(A)-\delta+\frac{\alpha}{\alpha-1}\log(\varepsilon)-H_\alpha^\uparrow(X|E))} \\
&\leq 2^{2/\alpha-2}2^{\frac{\alpha-1}{\alpha}(\frac{\alpha}{\alpha-1}\log(\varepsilon))} \\
&\leq \varepsilon\,,
\end{aligned}
\tag{54}
$$

where on the second line we use the assumption $\log(d_Z) \leq 2\log(d_A) - H_{\frac{\alpha}{2\alpha-1}}(A) - \delta + \frac{\alpha}{\alpha-1}\log(\varepsilon)$, on the third line we use (53) and on the final line we note that $2^{2/\alpha-2} \leq 1$ for all $\alpha \in (1,2]$. As this holds for all $\delta > 0$ the result follows. The second statement of the corollary follows from the same argument noting that $\mathcal{R}_\mathbb{H}^{\mathrm{POVM}}(\rho_A^{\otimes n}) = n\mathcal{R}_\mathbb{H}^{\mathrm{POVM}}(\rho_A)$. $\qquad\square$

Remark 5.3 (Extraction with projective measurements). The above theorem could also be stated with the restriction of extracting randomness using projective measurements. In that case one just replaces $2\log(d_A)$ with $\log(d_A)$. In particular we find that using projective measurements we can extract up to l-bits of ε-secure randomness for any $l \in \mathbb{N}$ satisfying

$$
l \leq \log(d_A) - \min_{\alpha \in (1,2]} \left(H_{\frac{\alpha}{2\alpha-1}}(A) - \frac{\alpha}{\alpha-1}\log(\varepsilon) \right).
\tag{55}
$$

The statement in Theorem 5.2 regarding extracting randomness from n copies shows that we can approach the asymptotic rates of randomness extraction as $n \to \infty$. In particular, taking $\alpha = 1 + \frac{1}{\sqrt{n}}$ we see that for any $\varepsilon > 0$ we have

$$
\frac{l_n}{n} < 2\log(d_A) - H_{\frac{1+\sqrt{n}}{2+\sqrt{n}}}^\uparrow(A) + \frac{1+\sqrt{n}}{n}\log(\varepsilon)
\tag{56}
$$

and the right-hand-side tends to $\mathcal{R}_H^{\mathrm{POVM}}(\rho_A) = 2\log(d_A) - H(A)$ as $n \to \infty$. Note this also gives an operational meaning to the bounds obtained in [17], they can be seen as quantifying the maximal amount of randomness that can be extracted from a quantum state using projective measurements in the asymptotic limit.

In Figure 3 we exhibit the speed of convergence to the asymptotic rates of randomness extraction. We consider the state

$$
\rho_A = \frac{1}{7}\begin{pmatrix} 3 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 2 & 1 \\ \frac{1}{\sqrt{2}} & 1 & 2 \end{pmatrix}
\tag{57}
$$

which has a spectrum $\{4/7, 2/7, 1/7\}$. We plot the rate of ε-secure extractable randomness for n-uses of the source, for various values of ε. This corresponds to computing (52) by optimizing over $\alpha \in (1,2)$. For comparison, we also plot the following lower bound on the ε-secure extractable randomness rates

$$
2\log(d_A) - \min_{\alpha \in (1,2]} \left(H_{\frac{1}{\alpha}}(A) - \frac{\alpha}{n(\alpha-1)}\log(\varepsilon) \right)
\tag{58}
$$

which corresponds to applying [18] with $H_\alpha^\downarrow(X|E)$. This captures the loss in rates that occurs when we use $H_\alpha^\downarrow$ instead of $H_\alpha^\uparrow$, which in practical security proofs is a common technique due to the difficulty of computing $H_\alpha^\uparrow$ [20].

In the plot, for each value of ε considered, we observe the convergence to the asymptotic rate as $n \to \infty$ (including for the lower bounds). In fact for $n > 10^4$ we see that there is little difference between (52) and its lower bound (58), as for this value of n we have $\alpha \approx 1$ which implies that $H_\alpha^\uparrow \approx H_\alpha^\downarrow$. However, for smaller n, which is also of practical relevance, we see a significant difference between using $H_\alpha^\uparrow$ or $H_\alpha^\downarrow$. This highlights the importance of developing general methods to compute and optimize $H_\alpha^\uparrow(X|E)$ for the purpose of practical and tight finite-size security proofs. We also

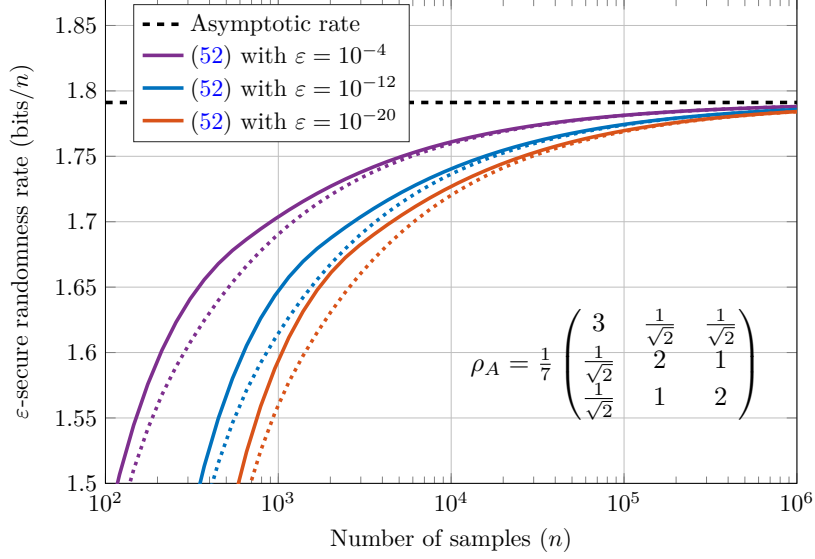Figure 3: A plot of the rate of ε-secure randomness extractable from a memoryless source producing a state ρ_A for various values of ε using (52). The dotted lines represent a lower bound on the rate of ε-secure randomness extractable.

note that for small n, at some point the optimal α hits the barrier of $\alpha = 2$ in the minimization, in the plot this is evident with the kinks in the various curves. This suggests that potentially there is a possible improvement in the result of [18] which would allow for the optimization to be over all $\alpha > 1$.

6 Conclusion

In this work we derived simple to compute expressions for the quantity of ε-secure randomness that is extractable from a given quantum state using both projective measurements and POVMs. Our adversarial model was general allowing for the adversary to have quantum side-information on both the source and the measurement device. Overall this provides a simple metric for the usefulness of a given source for QRNG protocols and acts as an easy to compute benchmarking tool for their performance and security proofs.

To derive these results we defined the maximal intrinsic $\mathbb{H}$-randomness of a quantum state, generalizing the intrinsic randomness problem found in [17]. We were able to derive closed form solutions to this problem for all quantum states and the entire families of operationally relevant sandwiched Rényi entropies. However our proofs of these results raise some interesting questions. In particular, our proof of the maximal randomness extractable using POVMs relies on a limiting argument. That is, we prove the existence of a sequence of POVMs that approach the maximal intrinsic $\mathbb{H}$-randomness of the given state. However, the limiting POVM in our construction does not achieve the maximal intrinsic $\mathbb{H}$-randomness rate, demonstrating a discontinuity in the intrinsic randomness measure. It would be interesting to understand if the maximal intrinsic randomness is always achievable with some POVM. In particular, we showed that they are achievable for all qubit states, however for states of higher dimensions we do not have explicit constructions.

Additionally, the discontinuity encountered in the limiting argument prompts a pertinent practical question. Do there exist POVMs that can produce high rates of extractable randomness in a robust manner? In particular, physical implementations will naturally implement a noisy version of the chosen POVM. Due to the existence of discontinuities, it is unclear a priori if a noisy version of a given POVM will still have a high rate of extractable randomness. Searching for robust POVM constructions that enable the extraction of large amounts of randomness would be of significant interest for optimizing the design of practical QRNGs.

References

[1] Miguel Herrero-Collantes and Juan Carlos Garcia-Escartin. "Quantum random number generators". Rev. Mod. Phys. **89**, 015004 (2017). arXiv:1604.03304.

[2] Charles H. Bennett and Gilles Brassard. "Quantum cryptography: Public key distribution and coin tossing". In Proceedings of IEEE International Conference on Computers, Systems, and Signal Processing. Pages 175–179. IEEENew York (1984). arXiv:2003.06557.

[3] Christopher Portmann and Renato Renner. "Security in quantum cryptography". Reviews of Modern Physics**94** (2022). arXiv:2102.00021.

[4] Frédéric Dupuis, Omar Fawzi, and Renato Renner. "Entropy accumulation". Communications in Mathematical Physics **379**, 867–913 (2020). arXiv:1607.01796.

[5] Tony Metger, Omar Fawzi, David Sutter, and Renato Renner. "Generalised entropy accumulation". In 2022 IEEE 63rd Annual Symposium on Foundations of Computer Science (FOCS). Pages 844–850. IEEE (2022). arXiv:2203.04989.

[6] Igor Devetak and Andreas Winter. "Distillation of secret key and entanglement from quantum states". Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences **461**, 207–235 (2005). arXiv:quant-ph/0306078.

[7] Rotem Arnon-Friedman, Renato Renner, and Thomas Vidick. "Simple and tight device-independent security proofs". SIAM Journal on Computing **48**, 181–225 (2019). arXiv:1607.01797.

[8] Stefano Pironio, Antonio Acín, Nicolas Brunner, Nicolas Gisin, Serge Massar, and Valerio Scarani. "Device-independent quantum key distribution secure against collective attacks". New Journal of Physics **11**, 045021 (2009). arXiv:0903.4460.

[9] Patrick J. Coles, Eric M. Metodiev, and Norbert Lütkenhaus. "Numerical approach for unstructured quantum key distribution". Nature Communications **7**, 11712 (2016). arXiv:1510.01294.

[10] Thomas Van Himbeeck, Erik Woodhead, Nicolas J. Cerf, Raúl García-Patrón, and Stefano Pironio. "Semi-device-independent framework based on natural physical assumptions". Quantum **1**, 33 (2017). arXiv:1612.06828.

[11] Daniela Frauchiger, Renato Renner, and Matthias Troyer. "True randomness from realistic quantum devices" (2013). arXiv:1311.4547.

[12] Gabriel Senno, Thomas Strohm, and Antonio Acín. "Quantifying the intrinsic randomness of quantum measurements". Physical Review Letters**131** (2023). arXiv:2211.03581.

[13] Hao Dai, Boyang Chen, Xingjian Zhang, and Xiongfeng Ma. "Intrinsic randomness under general quantum measurements". Phys. Rev. Res. **5**, 033081 (2023). arXiv:2203.08624.

[14] Mario Berta, Omar Fawzi, and Stephanie Wehner. "Quantum to classical randomness extractors". IEEE Transactions on Information Theory **60**, 1168–1192 (2013). arXiv:1111.2026.

[15] Martin Müller-Lennert, Frédéric Dupuis, Oleg Szehr, Serge Fehr, and Marco Tomamichel. "On quantum Rényi entropies: A new generalization and some properties". Journal of Mathematical Physics**54** (2013). arXiv:1306.3142.

[16] Mark M. Wilde, Andreas Winter, and Dong Yang. "Strong converse for the classical capacity of entanglement-breaking and Hadamard channels via a sandwiched rényi relative entropy". Communications in Mathematical Physics **331**, 593–622 (2014). arXiv:1306.1586.

[17] Shuyang Meng, Fionnuala Curran, Gabriel Senno, Victoria J. Wright, Máté Farkas, Valerio Scarani, and Antonio Acín. "Maximal intrinsic randomness of a quantum state". Phys. Rev. A **110**, L010403 (2024). arXiv:2307.15708.

[18] Frédéric Dupuis. "Privacy amplification and decoupling without smoothing". IEEE Transactions on Information Theory **69**, 7784–7792 (2023). arXiv:2105.05342.

[19] Giacomo Mauro D'Ariano, Paoloplacido Lo Presti, and Paolo Perinotti. "Classical randomness in quantum measurements". Journal of Physics A: Mathematical and General **38**, 5979–5991 (2005). arXiv:quant-ph/0408115.

[20] Thomas Van Himbeeck and Peter Brown. "A tight and general finite-size security proof for quantum key distribution". In preparation.

[21] Marco Tomamichel, Mario Berta, and Masahito Hayashi. "Relating different quantum generalizations of the conditional Rényi entropy". Journal of Mathematical Physics**55** (2014). arXiv:1311.3887.

[22] Dénes Petz. "Quasi-entropies for finite quantum systems". Reports on Mathematical Physics **23**, 57–65 (1986).

[23] Keiji Matsumoto. "A new quantum version of f-divergence". In Masanao Ozawa, Jeremy Butterfield, Hans Halvorson, Miklós Rédei, Yuichiro Kitajima, and Francesco Buscemi, editors, Reality and Measurement in Algebraic Quantum Theory. Pages 229–273. Singapore (2018). Springer Singapore. arXiv:1311.4722.

[24] Andreas Bluhm, Angela Capel, Paul Gondolf, and Tim Möbus. "Unified framework for continuity of sandwiched Rényi divergences" (2023). arXiv:2308.12425.

[25] Marco Tomamichel. "Quantum information processing with finite resources". Springer International Publishing. (2016). arXiv:1504.00233.

[26] John Watrous. "The theory of quantum information". Cambridge University Press. (2018).

A Proof of Theorem 3.4

In this section we prove the closed form expressions for $\mathcal{R}_{H_\alpha^\uparrow}^{\mathrm{PVM}}(\rho_A)$ and $\mathcal{R}_{H_\alpha^\downarrow}^{\mathrm{PVM}}(\rho_A)$. We begin with the following lemma.

Lemma A.1. *Let $\alpha \in [1/2, 1) \cup (1, \infty)$ and let $\beta > 0$. Let ρ_{AE} be a pure state, $\{M_x\}_x$ be a rank-one measurement and define the cq-state $\rho_{XE} = \sum_x |x\rangle\langle x| \otimes \rho_{E,x}$ where $\rho_{E,x} = \mathrm{Tr}_A[(M_x \otimes I_E)\rho_{AE}]$. Then*

$$D_\alpha\left(\rho_{XE} \,\middle\|\, I_A \otimes \frac{\rho_E^\beta}{\mathrm{Tr}\left[\rho_E^\beta\right]}\right) = \frac{1}{\alpha-1} \log\left(\frac{\sum_x \mathrm{Tr}\left[\rho_A^{\frac{\beta(1-\alpha)}{\alpha}+1} M_x\right]^\alpha}{\mathrm{Tr}\left[\rho_A^\beta\right]^{1-\alpha}}\right). \tag{59}$$

Proof. As ρ_{XE} is a classical-quantum state, for any density matrix σ_E satisfying $\mathrm{supp}(\rho_{E,x}) \subseteq \mathrm{supp}(\sigma_E)$ for all x we have

$$D_\alpha(\rho_{XE} \| I_X \otimes \sigma_E) = \frac{1}{\alpha-1} \log\left(\sum_x \mathrm{Tr}\left[\left(\sigma_E^{\frac{1-\alpha}{2\alpha}} \rho_{E,x} \sigma_E^{\frac{1-\alpha}{2\alpha}}\right)^\alpha\right]\right). \tag{60}$$

Now by the Schmidt decomposition we can write $\rho_{AE} = |\psi\rangle\langle\psi|$ with $|\psi\rangle = \sum_i \sqrt{\lambda_i} |v_i\rangle_A \otimes |w_i\rangle_E$. As $M_x = |m_x\rangle\langle m_x|$ is a rank-one measurement operator we see that

$$\begin{aligned}
\rho_{E,x} &= \mathrm{Tr}_A[(M_x \otimes I)\rho_{AE}] \\
&= \sum_{ij} \sqrt{\lambda_i \lambda_j} \langle m_x|v_i\rangle \langle v_j|m_x\rangle |w_i\rangle\langle w_j| \\
&= |\psi_x\rangle\langle\psi_x|_E
\end{aligned} \tag{61}$$

where we have defined the (not necessarily normalized) vector $|\psi_x\rangle_E = \sum_i \sqrt{\lambda_i} \langle m_x|v_i\rangle |w_i\rangle$. Thus, as $\rho_{E,x}$ is always rank-one, (60) simplifies to

$$D_\alpha(\rho_{XE} \| I_X \otimes \sigma_E) = \frac{1}{\alpha-1} \log\left(\sum_x \left(\langle\psi_x| \sigma_E^{\frac{1-\alpha}{\alpha}} |\psi_x\rangle\right)^\alpha\right). \tag{62}$$

Now substituting in $\sigma_E = \frac{\rho_E^\beta}{\text{Tr}[\rho_E^\beta]}$ the right-hand-side becomes

$$\frac{1}{\alpha-1}\log\left(\frac{\sum_x\left(\langle\psi_x|\rho_E^{\frac{\beta(1-\alpha)}{\alpha}}|\psi_x\rangle\right)^\alpha}{\text{Tr}\left[\rho_E^\beta\right]^{1-\alpha}}\right) = \frac{1}{\alpha-1}\log\left(\frac{\sum_x\left(\langle\psi_x|\left(\sum_i\lambda_i^{\frac{\beta(1-\alpha)}{\alpha}}|w_i\rangle\langle w_i|\right)|\psi_x\rangle\right)^\alpha}{\text{Tr}\left[\rho_E^\beta\right]^{1-\alpha}}\right)$$

$$= \frac{1}{\alpha-1}\log\left(\frac{\sum_x\left(\sum_i\lambda_i^{\frac{\beta(1-\alpha)}{\alpha}+1}|\langle v_i|m_x\rangle|^2\right)^\alpha}{\text{Tr}\left[\rho_E^\beta\right]^{1-\alpha}}\right)$$

$$= \frac{1}{\alpha-1}\log\left(\frac{\sum_x\text{Tr}\left[\rho_A^{\frac{\beta(1-\alpha)}{\alpha}+1}M_x\right]^\alpha}{\text{Tr}\left[\rho_A^\beta\right]^{1-\alpha}}\right)$$

$$(63)$$

where on the final line we have noted that $\text{Tr}\left[\rho_E^\beta\right] = \text{Tr}\left[\rho_A^\beta\right]$ as ρ_{AE} is pure. $\qquad\square$

The above lemma allows us to simplify the expressions of the sandwiched-Rényi entropies.

Lemma A.2. *Let ρ_{AE} be a pure state, $\{M_x\}_x$ be a rank-one measurement and define the cq-state $\rho_{XE} = \sum_x |x\rangle\langle x|\otimes\rho_{E,x}$ where $\rho_{E,x} = \text{Tr}_A[(M_x\otimes I)\rho_{AE}]$. Then for $\alpha\in[1/2,1)\cup(1,\infty)$ we have*

$$H_\alpha^\downarrow(X|E) = H_\alpha(X') - H_{1/\alpha}(A) \qquad (64)$$

where X' is a random variable with a distribution $p(x) = \text{Tr}\left[\rho_A^{1/\alpha}M_x\right]/\text{Tr}\left[\rho_A^{1/\alpha}\right]$. Additionally, for $\alpha > 1$ we have

$$H_\alpha^\uparrow(X|E) = H_\alpha(X'') - H_{\frac{\alpha}{2\alpha-1}}(A) \qquad (65)$$

where X'' is a random variable with a distribution $p(x) = \text{Tr}\left[\rho_A^{\frac{\alpha}{2\alpha-1}}M_x\right]/\text{Tr}\left[\rho_A^{\frac{\alpha}{2\alpha-1}}\right]$.

Proof. Let us begin with $H_\alpha^\downarrow(X|E)$. Taking $\beta = 1$ in Lemma A.1, we have

$$H_\alpha^\downarrow(X|E)_{\rho_{XE}} = \frac{1}{1-\alpha}\log\left(\sum_x\left(\text{Tr}\left[\rho_A^{\frac{1}{\alpha}}M_x\right]\right)^\alpha\right)$$

$$= \frac{1}{1-\alpha}\log\left(\text{Tr}\left[\rho_A^{\frac{1}{\alpha}}\right]^\alpha\sum_x\left(\frac{\text{Tr}\left[\rho_A^{\frac{1}{\alpha}}M_x\right]}{\text{Tr}\left[\rho_A^{\frac{1}{\alpha}}\right]}\right)^\alpha\right)$$

$$= \frac{\alpha}{1-\alpha}\log\left(\text{Tr}\left[\rho_A^{\frac{1}{\alpha}}\right]\right) + \frac{1}{1-\alpha}\log\left(\sum_x\left(\frac{\text{Tr}\left[\rho_A^{\frac{1}{\alpha}}M_x\right]}{\text{Tr}\left[\rho_A^{\frac{1}{\alpha}}\right]}\right)^\alpha\right) \qquad (66)$$

$$= -H_{1/\alpha}(A) + H_\alpha(X')$$

where on the fourth line we introduced the classical random variable X' with distribution $p(x) = \text{Tr}\left[\rho_A^{1/\alpha}M_x\right]/\text{Tr}\left[\rho_A^{1/\alpha}\right]$ and noted that the term was a classical Rényi entropy.

Now we prove the case of $H_\alpha^\uparrow(X|E)$. Using [21, Corollary 4] and the above result for $H_\alpha^\downarrow(X|E)$, we have for any $\alpha > 1$

$$H_\alpha^\uparrow(X|E)_{\rho_{XE}} \leq H_{2-\frac{1}{\alpha}}^\downarrow(X|E)_{\rho_{XE}}$$

$$= H_\alpha(X'') - H_{\frac{\alpha}{2\alpha-1}}(A), \qquad (67)$$

where X'' is a classical random variable with a distribution

$$p(x) = \text{Tr}\left[\rho_A^{\frac{\alpha}{2\alpha-1}} M_x\right] / \text{Tr}\left[\rho_A^{\frac{\alpha}{2\alpha-1}}\right] . \tag{68}$$

However, if we apply Lemma A.1 with $\beta = \alpha/(2\alpha - 1)$ then we find also the lower bound

$$
\begin{aligned}
H_\alpha^\uparrow(X|E) &\geq \frac{1}{1-\alpha} \log\left(\sum_x \frac{\text{Tr}\left[\rho_A^{\frac{\alpha}{2\alpha-1}} M_x\right]^\alpha}{\text{Tr}\left[\rho_A^{\frac{\alpha}{2\alpha-1}}\right]^{1-\alpha}}\right) \\
&= \frac{1}{1-\alpha} \log\left(\sum_x \frac{\text{Tr}\left[\rho_A^{\frac{\alpha}{2\alpha-1}} M_x\right]^\alpha}{\text{Tr}\left[\rho_A^{\frac{\alpha}{2\alpha-1}}\right]^\alpha}\right) + \frac{2\alpha-1}{1-\alpha} \log\left(\text{Tr}\left[\rho_A^{\frac{\alpha}{2\alpha-1}}\right]\right) \\
&= H_\alpha(X'') - H_{\frac{\alpha}{2\alpha-1}}(A) ,
\end{aligned}
\tag{69}
$$

which completes the proof. $\qquad\square$

Using the above lemma we can now find our closed form expressions.

Lemma A.3. *Let $\alpha \in [1/2, 1) \cup (1, \infty)$ and let $\rho_A \in \mathscr{D}(A)$. Then*

$$\mathcal{R}_{H_\alpha^\downarrow}^{\text{PVM}}(\rho_A) = \log(d_A) - H_{1/\alpha}(A) \tag{70}$$

where d_A is the dimension of system A. Furthermore, this quantity is achieved if and only if the rank-one PVM $\{M_x\}_x$ satisfies

$$\text{Tr}\left[\rho_A^{\frac{1}{\alpha}} M_x\right] = \frac{\text{Tr}\left[\rho_A^{\frac{1}{\alpha}}\right]}{d_A} \tag{71}$$

for all $x \in [d_A]$.

Proof. By Lemma A.2 and Lemma 3.3 we know that

$$\mathcal{R}_{H_\alpha^\downarrow}^{\text{PVM}}(\rho_A) = \sup_{\text{rank}-\text{one PVMs}} H_\alpha^\downarrow(X|E) = \sup_{\text{rank}-\text{one PVMs}} H_\alpha(X') - H_{1/\alpha}(A) \tag{72}$$

where X' is a random variable with distribution $p(x) = \text{Tr}\left[\rho_A^{1/\alpha} M_x\right] / \text{Tr}\left[\rho_A^{1/\alpha}\right]$. Since the only $H_\alpha(X')$ depends on the PVM $\{M_x\}_x$, we look to maximize this term. We know that $H_\alpha(X') = \log(d_A)$ is maximal iff there exists a rank-one projective measurement $\{M_x\}_x$ such that

$$\frac{\text{Tr}\left[\rho_A^{\frac{1}{\alpha}} M_x\right]}{\text{Tr}\left[\rho_A^{\frac{1}{\alpha}}\right]} = \frac{1}{d_A} \tag{73}$$

for all $x \in [d_A]$. To complete the proof, we note that this condition is achieved by taking $\{M_x\}_x$ to be the projectors onto a basis which is mutually unbiased with the eigenbasis of ρ_A. $\qquad\square$

We can now use the result for $\mathcal{R}_{H_\alpha^\downarrow}^{\text{PVM}}(\rho_A)$ to derive a result for $\mathcal{R}_{H_\alpha^\uparrow}^{\text{PVM}}(\rho_A)$.

Lemma A.4. *Let $\alpha > 1$ and let $\rho_A \in D(A)$. Then*

$$\mathcal{R}_{H_\alpha^\uparrow}^{\text{PVM}}(\rho_A) = \log(d_A) - H_{\frac{\alpha}{2\alpha-1}}(A) \tag{74}$$

where d_A is the dimension of system A. Furthermore, this quantity is achieved if and only if the rank-one measurement $\{M_x\}_x$ satisfies

$$\text{Tr}\left[\rho_A^{\frac{\alpha}{2\alpha-1}} M_x\right] = \frac{\text{Tr}\left[\rho_A^{\frac{\alpha}{2\alpha-1}}\right]}{d_A} \tag{75}$$

for all $x \in [d_A]$.

Proof. By the same logic as the previous lemma, we can use Lemma A.2 and Lemma 3.3 to show that

$$\mathcal{R}_{H_\alpha^\uparrow}^{\mathrm{PVM}}(\rho_A) = \sup_{\mathrm{rank-one\ PVMs}} H_\alpha^\uparrow(X|E) = \sup_{\mathrm{rank-one\ PVMs}} H_\alpha(X') - H_{\frac{\alpha}{2\alpha-1}}(A) \tag{76}$$

where X' is a random variable with distribution $p(x) = \mathrm{Tr}\left[\rho_A^{\frac{\alpha}{2\alpha-1}} M_x\right] / \mathrm{Tr}\left[\rho_A^{\frac{\alpha}{2\alpha-1}}\right]$. Since the only $H_\alpha(X')$ depends on the PVM $\{M_x\}_x$, we look to maximize this term. We know that $H_\alpha(X') = \log(d_A)$ is maximal iff there exists a rank-one projective measurement $\{M_x\}_x$ such that

$$\frac{\mathrm{Tr}\left[\rho_A^{\frac{\alpha}{2\alpha-1}} M_x\right]}{\mathrm{Tr}\left[\rho_A^{\frac{\alpha}{2\alpha-1}}\right]} = \frac{1}{d_A} \tag{77}$$

for all $x \in [d_A]$. To complete the proof, we note that this condition is achieved by taking $\{M_x\}_x$ to be the projectors onto a basis which is mutually unbiased with the eigenbasis of ρ_A. $\square$

Remark A.5 (Intrinsic randomness for H and $H_{\min}$)**.** The proof of Lemma A.4 also extends to the limit $\alpha \to 1$, recovering

$$\mathcal{R}_H^{\mathrm{PVM}}(\rho_A) = \log(d_A) - H(A), \tag{78}$$

and to the limit $\alpha \to \infty$, recovering

$$\mathcal{R}_{H_{\min}}^{\mathrm{PVM}}(\rho_A) = \log(d_A) - H_{1/2}(A), \tag{79}$$

which were proven in [17] using different techniques.

B   Proof of Theorem 4.6

In this section we will develop all of the necessary results to prove Theorem 4.6. We begin with some preliminary results.

Lemma B.1 (Convexity of exponential intrinsic randomness)**.** *Let $\alpha \in [\frac{1}{2}, 1) \cup (1, \infty)$, let $\mathbb{H}$ be either $H_\alpha^\downarrow$ or $H_\alpha^\uparrow$ and let $\rho_A \in \mathscr{D}(A)$. Then for any two POVMs $\{M_x\}_x$, $\{N_x\}_x$ and $\lambda \in [0,1]$ we have*

$$2^{\mathcal{I}_{\mathbb{H}}^{\mathrm{POVM}}(\rho_A, \{\lambda M_x + (1-\lambda) N_x\}_x)} \le \lambda\, 2^{\mathcal{I}_{\mathbb{H}}^{\mathrm{POVM}}(\rho_A, \{M_x\}_x)} + (1-\lambda)\, 2^{\mathcal{I}_{\mathbb{H}}^{\mathrm{POVM}}(\rho_A, \{N_x\}_x)}. \tag{80}$$

Proof. Let $\varepsilon > 0$. Let $(\{P_x\}_x, \sigma_B)$ be a dilation for $\mathcal{I}_{\mathbb{H}}^{\mathrm{POVM}}(\rho_A, \{M_x\}_x)$ and let $(\{Q_x\}_x, \tau_{B'})$ be a dilation for $\mathcal{I}_{\mathbb{H}}^{\mathrm{POVM}}(\rho_A, \{N_x\}_x)$ satisfying

$$\mathbb{H}(X|E)_{\sigma_{XE}} \le \mathcal{I}_{\mathbb{H}}^{\mathrm{POVM}}(\rho_A, \{M_x\}_x) + \varepsilon \tag{81}$$

and

$$\mathbb{H}(X|E)_{\tau_{XE}} \le \mathcal{I}_{\mathbb{H}}^{\mathrm{POVM}}(\rho_A, \{N_x\}_x) + \varepsilon. \tag{82}$$

Note that we may assume $\dim(B) = \dim(B')$ without loss of generality. To see this note that if $\dim(B) < \dim(B')$ then we can embed B into B' using an isometry $V : B \to B'$ which transforms the dilation to $(\{VP_xV^\dagger\}_x, V\sigma V^\dagger)$ where to complete the projective measurement we add a new outcome $P_\perp = I_{AB'} - I_A \otimes VV^\dagger$. However by construction this new measurement operator satisfies $\mathrm{Tr}\left[P_\perp(I_A \otimes V\sigma V^\dagger)\right] = 0$. This construction yields a new dilation for $\{M_x\}_x$ but it does not change $\mathbb{H}(X|E)$ as the post-measurement state remains unchanged. Hence we may assume the dilating systems are the same for both POVMs.

Hence we have two states σ_{AB} and τ_{AB} such that $\mathrm{Tr}_B[\sigma_{AB}] = \rho_A$ and $\mathrm{Tr}_B[\tau_{AB}] = \rho_A$ and $(\{P_x\}_x, \sigma_B)$ is a dilation of $\{M_x\}_x$ and $(\{Q_x\}_x, \tau_B)$ is a dilation of $\{N_x\}_x$. Now for any $\lambda \in [0,1]$, consider a new state

$$\omega_{ABF} = \lambda\, \sigma_{AB} \otimes |0\rangle\langle 0|_F + (1-\lambda)\, \tau_{AB}\, |1\rangle\langle 1|_F \tag{83}$$

and projective measurement on ABF defined as

$$R_x = P_x \otimes |0\rangle\langle 0|_F + Q_x \otimes |1\rangle\langle 1|_F. \tag{84}$$

Then a direct calculation shows that this pair forms a dilation of the mixed measurement $\{\lambda M_x + (1-\lambda)N_x\}_x$. So consider the state

$$\rho_{ABEFF'} = \lambda\sigma_{ABE} \otimes |00\rangle\langle 00|_{FF'} + (1-\lambda)\tau_{ABE} \otimes |11\rangle\langle 11|_{FF'} \tag{85}$$

where σ_{ABE} and τ_{ABE} are pure and let $\rho_{ABEFF'G}$ be a purification of $\rho_{ABEFF'}$ we assume that $EF'G$ are held by Eve and BF are held by the measurement device.

Now by definition and strong subadditivity we have

$$2^{\mathcal{I}_{\mathbb{H}}^{\text{POVM}}(\rho_A,\{\lambda M_x+(1-\lambda)N_x\}_x)} \le 2^{\mathbb{H}(X|EF'G)_{\rho_{XEF'G}}} \le 2^{\mathbb{H}(X|EF')_{\rho_{XEF'}}}. \tag{86}$$

Thus, as F' is a classical flag system we have from Lemma D.2

$$\begin{aligned}
2^{\mathcal{I}_{\mathbb{H}}^{\text{POVM}}(\rho_A,\{\lambda M_x+(1-\lambda)N_x\}_x)} &\le \lambda\, 2^{\mathbb{H}(X|E)_{\sigma_{XE}}} + (1-\lambda)\, 2^{\mathbb{H}(X|E)_{\tau_{XE}}} \\
&\le \lambda\, 2^{\mathcal{I}_{\mathbb{H}}^{\text{POVM}}(\rho_A,\{M_x\}_x)+\varepsilon} + (1-\lambda)\, 2^{\mathcal{I}_{\mathbb{H}}^{\text{POVM}}(\rho_A,\{N_x\}_x)+\varepsilon}.
\end{aligned} \tag{87}$$

The result then follows from the fact that this holds for all $\varepsilon > 0$. $\qquad\square$

We now show that if the probe state of the dilation ρ_B is not rank-one then we can always find a rank-one probe state that implements the same extremal POVM.

Lemma B.2 (Adapted from [13, Proposition 1]). *Let $\{M_x\}_x$ be an extremal POVM on A and let $(\{P_x\}_x, \sigma_B)$ be a corresponding dilation. Let $\sum_y p_y\, |v_y\rangle\langle v_y|_B$ be the spectral decomposition of σ_B with $\{|v_y\rangle\}_y$ forming an orthonormal basis. Then for all y such that $p_y > 0$ we have $(\{P_x\}_x, |v_y\rangle\langle v_y|_B)$ is also a dilation of $\{M_x\}_x$.*

Proof. From the consistency condition for dilations, we have that for all x

$$\begin{aligned}
M_x &= \text{Tr}_B\left[P_x\left(I_A \otimes \sigma_B\right)\right] \\
&= \sum_y p_y \text{Tr}_B\left[P_x\left(I_A \otimes |v_y\rangle\langle v_y|_B\right)\right].
\end{aligned} \tag{88}$$

Let us denote $N_x^y = \text{Tr}_B\left[P_x\left(I_A \otimes |v_y\rangle\langle v_y|_B\right)\right]$. Note that we have for all y

$$\begin{aligned}
\sum_x N_x^y &= \sum_x \text{Tr}_B\left[P_x\left(I_A \otimes |v_y\rangle\langle v_y|_B\right)\right] \\
&= \text{Tr}_B\left[\left(\sum_x P_x\right)\left(I_A \otimes |v_y\rangle\langle v_y|_B\right)\right] \\
&= I_A.
\end{aligned} \tag{89}$$

Furthermore, by positivity of the partial trace we also have that N_x^y is positive semidefinite and therefore it is a POVM on A. In particular, since $\{M_x\}_x$ is extremal and since we have

$$M_x = \sum_y p_y N_x^y \tag{90}$$

it means that $N_x^y = M_x$ for all y such that $p_y > 0$. Thus, for all y, the consistency condition is satisfied for $(\{P_x\}_x, |v_y\rangle\langle v_y|_B)$, which means that it is a valid dilation of $\{M_x\}_x$. $\qquad\square$

Using this result we can prove Lemma 4.5 from the main text which we repeat here for clarity. Loosely, it shows that for extremal rank-one POVMs, the entanglement with the measurement device carries no useful information for Eve.

Lemma B.3. *Let $\{M_x\}_x$ be an extremal rank-one POVM on a system A, let $(\{P_x\}_x, \rho_B)$ be any dilation of $\{M_x\}_x$ and let $\rho_{ABE} \in \mathscr{D}(ABE)$ be any pure state with marginals ρ_A and ρ_B. Then the post-measurement state*

$$\rho_{XE} = \sum_x |x\rangle\langle x| \otimes \text{Tr}_{AB}\left[(P_x \otimes I_E)\rho_{ABE}\right] \tag{91}$$

can be written as

$$\rho_{XE} = \sum_x |x\rangle\langle x| \otimes \text{Tr}_A\left[(M_x \otimes I_E)\rho_{AE}\right] . \tag{92}$$

Proof. Consider the spectral decomposition $\rho_B = \sum_k \lambda_k |\tau_k\rangle\langle\tau_k|$. As $\{|\tau_k\rangle\}_k$ is an orthonormal basis, we can write

$$P_x = \sum_{ijkl} p_{ijkl} |i\rangle\langle j|_A \otimes |\tau_k\rangle\langle\tau_l|_B . \tag{93}$$

for some coefficients $p_{ijkl} \in \mathbb{C}$. By Lemma B.2 we know that $(\{P_x\}_x, |\tau_k\rangle\langle\tau_k|)$ is also a dilation for $\{M_x\}_x$ whenever $\lambda_k > 0$ and so

$$M_x = (I_A \otimes \langle\tau_k|)P_x(I_A \otimes |\tau_k\rangle) = \sum_{ij} p_{ijkk} |i\rangle\langle j| , \tag{94}$$

for any k such that $\lambda_k > 0$. However, $M_x = m_x |v\rangle\langle v|$ is also rank-one, thus we can extend the vector $|v\rangle$ to an orthonormal basis $\{|v_i\rangle\}_i$ for A (where we define $|v_0\rangle = |v\rangle$). This means that we can write

$$P_x = m_x |v_0\rangle\langle v_0| \otimes \sum_{k:\lambda_k>0} |\tau_k\rangle\langle\tau_k| + \sum_{ijkl:k\neq l, \lambda_k>0, \lambda_l>0} q_{ijkl} |v_i\rangle\langle v_j| \otimes |\tau_k\rangle\langle\tau_l| + R \tag{95}$$

for some coefficients $q_{ijkl} \in \mathbb{C}$ and R is a matrix such that $(I \otimes \langle\tau_k|)R(I \otimes |\tau_l\rangle) = 0$ whenever $\lambda_k > 0$ and $\lambda_l > 0$. We will show that if $i \neq 0$ and $j \neq 0$ then $q_{ijkl} = 0$. Suppose not and without loss of generality take $i = 0$ and $j \neq 0$ (a similar argument holds for other cases) then consider the vector $|w\rangle = \alpha |v_0\rangle \otimes |\tau_k\rangle + \beta |v_j\rangle \otimes |\tau_l\rangle$ where $\lambda_k, \lambda_l > 0$, then as $P_x \succeq 0$ we should have

$$0 \leq \langle w| P_x |w\rangle = |\alpha|^2 m_x + 2\Re[\overline{\alpha}\beta q_{0jkl}] \tag{96}$$

it is clear that there is a choice of β for which the RHS is negative if $q_{0jkl} \neq 0$. Thus we must have $q_{0jkl} = 0$. Overall this implies that

$$P_x = M_x \otimes \sum_{k,l,\lambda_k,\lambda_l>0} c_{klx} |\tau_k\rangle\langle\tau_l| + R \tag{97}$$

for some coefficients $c_{klx} \in \mathbb{C}$ which may a priori depend on x. We can now go one step further and show that $c_{klx} = 0$ when $k \neq l$. Note that $\sum_x P_x = I$, thus if $k \neq l$ we have

$$\begin{aligned}
0 &= \sum_x (I_A \otimes \langle\tau_k|)P_x(I_A \otimes |\tau_l\rangle) \\
&= \sum_x c_{klx} M_x
\end{aligned} \tag{98}$$

but as $\{M_x\}_x$ is a rank-one extremal POVM, its elements are all linearly independent and hence we must have $c_{klx} = 0$ for all $k \neq l$. This further simplifies P_x to

$$P_x = M_x \otimes \Pi_\sigma + R \tag{99}$$

where $\Pi_\sigma = \sum_{k:\lambda_k>0} |\tau_k\rangle\langle\tau_k|$ is the projector onto the support of σ.

Finally, consider

$$\begin{aligned}
\text{Tr}_{AB}\left[(P_x \otimes I_E)\rho_{ABE}\right] &= \text{Tr}_{AB}\left[((M_x \otimes \Pi_\sigma + R) \otimes I_E)\rho_{ABE}\right] \\
&= \text{Tr}_{AB}\left[(M_x \otimes \Pi_\sigma \otimes I_E)\rho_{ABE}\right] \\
&= \text{Tr}_{AB}\left[(M_x \otimes I_B \otimes I_E)\rho_{ABE}\right] \\
&= \text{Tr}_A\left[M_x \rho_{AE}\right] .
\end{aligned} \tag{100}$$

$\square$

By Lemma 4.4 we know that

$$\mathcal{R}_{\mathbb{H}}^{\text{POVM}}(\rho_A) = \sup_{\substack{\text{Extremal rank-one} \\ \{M_x\}_x}} \mathcal{I}_{\mathbb{H}}^{\text{POVM}}(\rho_A, \{M_x\}_x). \tag{101}$$

By the previous lemma we can further simplify this to

$$\begin{aligned} \mathcal{R}_{\mathbb{H}}^{\text{POVM}}(\rho_A) = \sup \quad & \mathbb{H}(X|E) \\ \text{s.t.} \quad & \rho_{XE} = \sum_x |x\rangle\langle x| \otimes \text{Tr}_A \left[(M_x \otimes I) |\rho\rangle\langle\rho|_{AE} \right] \\ & \{M_x\}_x \text{ is an extremal rank-one POVM} \end{aligned} \tag{102}$$

The following theorem establishes a closed form expression for (102).

Theorem B.4. *For $\alpha \in [\frac{1}{2}, 1) \cup (1, \infty)$ we have*

$$\mathcal{R}_{H_\alpha^\downarrow}^{\text{POVM}}(\rho_A) = 2\log(d_A) - H_{1/\alpha}(A). \tag{103}$$

Furthermore, this quantity is achieved by an extremal rank-one POVM $\{M_x\}_x$ if and only if

$$\text{Tr}\left[\rho_A^{\frac{1}{\alpha}} M_x\right] = \frac{\text{Tr}\left[\rho_A^{\frac{1}{\alpha}}\right]}{d_A^2} \tag{104}$$

for all $x \in [d_A^2]$. Similarly, for $\alpha > 1$ we have

$$\mathcal{R}_{H_\alpha^\uparrow}^{\text{POVM}}(\rho_A) = 2\log(d_A) - H_{\frac{\alpha}{2\alpha-1}}(A) \tag{105}$$

with this quantity being achieved by an extremal rank-one POVM $\{M_x\}_x$ if and only if

$$\text{Tr}\left[\rho_A^{\frac{\alpha}{2\alpha-1}} M_x\right] = \frac{\text{Tr}\left[\rho_A^{\frac{\alpha}{2\alpha-1}}\right]}{d_A^2} \tag{106}$$

for all $x \in [d_A^2]$.

Proof. Let $\mathbb{H}$ be either $H_\alpha^\downarrow$ or $H_\alpha^\uparrow$ for α in the respective ranges as defined in the statement of the theorem. By the previous arguments we have that

$$\mathcal{R}_{\mathbb{H}}^{\text{POVM}}(\rho_A) = \sup_{\substack{\text{Extremal rank-one} \\ \{M_x\}_x}} \mathbb{H}(X|E). \tag{107}$$

As the cq-state ρ_{XE} is a generated by a rank-one measurement $\{M_x\}_x$ on a pure state, we can apply Lemma A.2 to get, for $\alpha \in (1/2, 1) \cup (1, \infty)$,

$$H_\alpha^\downarrow(X|E) = H(X') - H_{1/\alpha}(A) \qquad \text{with} \qquad p(x') = \frac{\text{Tr}\left[\rho_A^{1/\alpha} M_x\right]}{\text{Tr}\left[\rho_A^{1/\alpha}\right]} \tag{108}$$

and for $\alpha > 1$

$$H_\alpha^\uparrow(X|E) = H(X'') - H_{\frac{\alpha}{2\alpha-1}}(A) \qquad \text{with} \qquad p(x'') = \text{Tr}\left[\rho_A^{\frac{\alpha}{2\alpha-1}} M_x\right] / \text{Tr}\left[\rho_A^{\frac{\alpha}{2\alpha-1}}\right]. \tag{109}$$

Thus the remainder of the proof entails maximizing $H(X')$ and $H(X'')$ over the set of extremal rank-one POVMs. In particular we look for an extremal rank-one POVM which generates a uniform distribution for X' and X'' which would establish that $H(X') = H(X'') = 2\log(d_A)$.

Consider for the moment, the d_A^2-outcome rank-one POVM defined by

$$M_{d_A x + y} = \frac{1}{d_A^2} |y\rangle\langle y| \tag{110}$$

for $x, y \in [d_A]$, where $\{|y\rangle\}_y$ is an orthonormal basis that is mutually unbiased with the eigenbasis of ρ_A. This measurement is comprised of d_A copies of the $\{|y\rangle\}_y$ basis measurement. Evidently, this measurement isn't extremal, but it is rank-one and it achieves the desired uniform distribution. By Lemma D.4, we know that for any rank-one POVM $\{M_x\}_x$ with no more than d_A^2 outcomes we can always find an extremal rank-one POVM $\{N_x\}_x$ that is ε-close to $\{M_x\}_x$ for any $\varepsilon > 0$. Thus, by continuity of the mappings $\rho_A \to X'$, $\rho_A \to X''$ and $X \to H_\alpha(X)$ we can find a sequence of extremal rank-one POVMs $\{N_x^{(k)}\}_x$ that tends to (110) and which establishes that

$$\sup_{\substack{\text{Extremal rank-one} \\ \{M_x\}_x}} H_\alpha^\uparrow(X|E) = 2\log(d_A) - H_{\frac{\alpha}{2\alpha-1}}(A) \tag{111}$$

and

$$\sup_{\substack{\text{Extremal rank-one} \\ \{M_x\}_x}} H_\alpha^\downarrow(X|E) = 2\log(d_A) - H_{1/\alpha}(A) \,. \tag{112}$$

$\square$

C Extension to other conditional entropies

The maximal intrinsic randomness problem formulated in Section 3 and Section 4 can be expressed for any conditional entropy. The choice of sandwiched Rényi entropies is natural due to their operational connections to randomness extraction which we demonstrate an application of in Section 5. Nevertheless there is nothing to stop one attempting to solve it for other conditional entropies such as the Petz-Rényi conditional entropies [22] or the Geometric entropies [23].

However, despite the apparent simpler appearance of these entropies, the problem appears to be significantly more difficult. In particular, it is not a priori clear how to perform a reduction to rank-one measurements, which was a crucial step in our proof, due to the fact that these entropies do not satisfy a result akin to Lemma D.1. As a consequence they do not agree with our intuitions regarding which measurements should produce maximal randomness. For example take the Petz-Rényi entropy

$$\overline{H}_\alpha^\downarrow(X|E)_{\rho_{XE}} = \frac{1}{1-\alpha} \log \operatorname{Tr}\left[\rho_{XE}^\alpha \rho_E^{1-\alpha}\right] \,, \tag{113}$$

which we define for $\alpha \in (1, 2]$. Consider the setting of projective measurements only, we may assume that measuring in a basis mutually unbiased with the eigenbasis of ρ_A would produce maximal randomness as it always produces a uniform distribution. However, for example let $\alpha = 3/2$ and consider the qutrit state $\rho = \frac{1}{4}|0\rangle\langle0| + \frac{3}{4}|1\rangle\langle1|$, then the unbiased basis measurement will give $\overline{H}_{3/2}^\downarrow(X|E) = \log(3) - 2\log((1+\sqrt{3})/2) \approx 0.685$. Whereas if we take the measurement defined by the orthonormal basis

$$|m_0\rangle = |0\rangle \qquad |m_1\rangle = \frac{|1\rangle + |2\rangle}{\sqrt{2}} \qquad |m_2\rangle = \frac{|1\rangle - |2\rangle}{\sqrt{2}} \tag{114}$$

then we find $\overline{H}_{3/2}^\downarrow(X|E) = 2\log(\frac{4\sqrt{2}}{\sqrt{2}+3}) \approx 0.716$. This shows that solving even $\mathcal{R}_{\mathbb{H}}^{\text{PVM}}(\rho_A)$ may be very difficult for these other conditional entropies.

D Additional lemmas

Lemma D.1. *Let ρ_{XYE} be a state classical on XY. Then for $\alpha \in [1/2, 1) \cup (1, \infty)$ we have*

$$\begin{aligned} H_\alpha^\uparrow(XY|E) &\geq H_\alpha^\uparrow(X|E) \\ H_\alpha^\downarrow(XY|E) &\geq H_\alpha^\downarrow(X|E) \,. \end{aligned} \tag{115}$$

Proof. Writing $\rho_{XYE} = \sum_{xy} |xy\rangle\langle xy| \otimes \rho_{E,xy}$, if we take any density matrix σ_E such that for all x, y we have $\operatorname{supp}(\rho_{E,xy}) \subseteq \sigma_E$, then

$$D_\alpha(\rho_{XYE} \| I_{XY} \otimes \sigma_E) = \frac{1}{\alpha - 1} \log \left(\sum_{xy} Q_\alpha(\rho_{E,xy} \| \sigma_E) \right) \tag{116}$$

where $Q_\alpha(\rho\|\sigma) := \mathrm{Tr}\left[\left(\sigma^{\frac{1-\alpha}{2\alpha}}\rho\sigma^{\frac{1-\alpha}{2\alpha}}\right)^\alpha\right]$. For $\alpha > 1$, Q_α is superadditive in the first argument [24], in the sense that

$$Q_\alpha(\rho_0 + \rho_1\|\sigma) \geq Q_\alpha(\rho_0\|\sigma) + Q_\alpha(\rho_1\|\sigma) \tag{117}$$

for any $\rho_0, \rho_1 \succeq 0$. Note for $\alpha \in [1/2, 1)$ the inequality is reversed and Q_α is subadditive. Thus for $\alpha > 1$, using the monotonicity of the logarithm we have

$$D_\alpha(\rho_{XYE}\|I_{XY} \otimes \sigma_E) \leq \frac{1}{\alpha - 1} \log\left(\sum_x Q_\alpha\left(\sum_y \rho_{E,xy}\Big\|\sigma_E\right)\right) \tag{118}$$

$$= D_\alpha(\rho_{XE}\|I_X \otimes \sigma_E).$$

Taking $\sigma_E = \rho_E$ this immediately implies $H_\alpha^\downarrow(XY|E) \geq H_\alpha^\downarrow(X|E)$. For the optimized variant we have

$$\begin{aligned}
H_\alpha^\uparrow(X|E) &= \sup_{\substack{\sigma_E \\ \sum_y \rho_{E,xy} \ll \sigma_E}} -D_\alpha(\rho_{XE}\|I_X \otimes \sigma_E) \\
&\leq \sup_{\substack{\sigma_E \\ \sum_y \rho_{E,xy} \ll \sigma_E}} -D_\alpha(\rho_{XYE}\|I_{XY} \otimes \sigma_E) \\
&\leq \sup_{\substack{\sigma_E \\ \rho_{E,xy} \ll \sigma_E}} -D_\alpha(\rho_{XYE}\|I_{XY} \otimes \sigma_E) \\
&= H_\alpha^\uparrow(XY|E).
\end{aligned} \tag{119}$$

where on the third line we have noted that

$$\left\{\sigma_E \in \mathscr{D}(E) \mid \sum_y \rho_{E,xy} \ll \sigma_E \quad \forall x\right\} \subseteq \left\{\sigma_E \in \mathscr{D}(E) \mid \rho_{E,xy} \ll \sigma_E \quad \forall x, y\right\}. \tag{120}$$

The proof for $\alpha < 1$ is the same but uses subadditivity of Q_α instead. $\qquad\square$

Lemma D.2. *Let be $\alpha \in [\frac{1}{2}, 1) \cup (1, \infty)$. Let $\mathbb{H}$ be either $H_\alpha^\downarrow$ or $H_\alpha^\uparrow$. Let $\rho_{XE} \in \mathscr{D}(XE)$, $\sigma_{XE} \in \mathscr{D}(XE)$, $\lambda \in [0, 1]$ and define*

$$\rho_{XEF} = \lambda\rho_{XE} \otimes |0\rangle\langle 0|_F + (1 - \lambda)\sigma_{XE} \otimes |1\rangle\langle 1|_F . \tag{121}$$

Then

$$2^{\mathbb{H}(X|EF)_{\rho_{XEF}}} \leq \lambda 2^{\mathbb{H}(X|E)_{\rho_{XE}}} + (1 - \lambda)2^{\mathbb{H}(X|E)_{\sigma_{XE}}}. \tag{122}$$

Proof. We have from [25, Proposition 5.4]

$$2^{(1-\alpha)H_\alpha^\downarrow(X|EF)} = \lambda\, 2^{(1-\alpha)H_\alpha^\downarrow(X|E)_{\rho_{XE}}} + (1 - \lambda)\, 2^{(1-\alpha)H_\alpha^\downarrow(X|E)_{\sigma_{XE}}} . \tag{123}$$

Now, note that $x \mapsto x^{\frac{1}{1-\alpha}}$ is convex for $\alpha \in [\frac{1}{2}, 1) \cup (1, \infty)$. Thus, by Jensen's inequality we have the desired result for $H_\alpha^\downarrow$.

Similarly, from the same proposition, we have

$$2^{\frac{1-\alpha}{\alpha}H_\alpha^\uparrow(X|EF)} = \lambda\, 2^{\frac{1-\alpha}{\alpha}H_\alpha^\uparrow(X|E)_{\rho_{XE}}} + (1 - \lambda)\, 2^{\frac{1-\alpha}{\alpha}H_\alpha^\uparrow(X|E)_{\sigma_{XE}}} . \tag{124}$$

Since $x \mapsto x^{\frac{\alpha}{1-\alpha}}$ is convex for $\alpha \in [\frac{1}{2}, 1) \cup (1, \infty)$, we once again have the desired result by Jensen's inequality. $\qquad\square$

Lemma D.3. *Let $\{M_x\}_x$ be an extremal POVM and let*

$$M_x = \sum_y \lambda_{x,y}\, |\varphi_{x,y}\rangle\langle\varphi_{x,y}| . \tag{125}$$

be its spectral decomposition. Then $\{\lambda_{x,y}\, |\varphi_{x,y}\rangle\langle\varphi_{x,y}|\}_{xy}$ is an extremal rank-one POVM where the indexes range over (x, y) such that $\lambda_{x,y} > 0$.

Proof. It is clear that $\{\lambda_{x,y} |\varphi_{x,y}\rangle\langle\varphi_{x,y}|\}_{x,y}$ is a POVM, since all its elements are positive and we have

$$\sum_{x,y} \lambda_{x,y} |\varphi_{x,y}\rangle\langle\varphi_{x,y}| = \sum_x M_x = I . \tag{126}$$

Thus it is a rank-one POVM. We now show that it inherits extremality from $\{M_x\}_x$. Since a rank-one POVM is extremal if and only if its elements are linearly independent [26, Corollary 2.50], let us assume for contradiction that there exist complex coefficients $\{\alpha_{x,y}\}_{x,y}$ (some of which are non-zero) such that

$$\sum_{x,y} \alpha_{x,y}\lambda_{x,y} |\varphi_{x,y}\rangle\langle\varphi_{x,y}| = 0 , \tag{127}$$

i.e., the POVM elements $N_{x,y} = \lambda_{x,y} |\varphi_{x,y}\rangle\langle\varphi_{x,y}|$ are linearly dependent. For each x, define θ_x as

$$\theta_x := \sum_y \alpha_{x,y}\lambda_{x,y} |\varphi_{x,y}\rangle\langle\varphi_{x,y}| \tag{128}$$

and define

$$\tau_x := \theta_x + \theta_x^\dagger . \tag{129}$$

Clearly, for all x, τ_x is Hermitian. Furthermore, by assumption we have

$$\sum_x \tau_x = 0 \tag{130}$$

and moreoever by construction we have for each x that

$$\text{supp}(\tau_x) \subseteq \text{supp}(M_x) . \tag{131}$$

Since $\{M_x\}_x$ is extremal, it follows from [26, Theorem 2.47] that we must have $\tau_x = 0$ for each x. Thus, $i\theta_x$ is Hermitian and also satisfies the aforementioned condition, which gives us $\theta_x = 0$ for all x, i.e.,

$$\sum_y \alpha_{x,y}\lambda_{x,y} |\varphi_{x,y}\rangle\langle\varphi_{x,y}| = 0 . \tag{132}$$

Due to orthogonality, this implies that $\alpha_{x,y}\lambda_{x,y} = 0$ for all x,y, which implies $\alpha_{x,y} = 0$. This is in contradiction with the original assumption of linear dependence and hence the measurement $\{N_{x,y}\}_{x,y}$ is extremal. $\square$

Lemma D.4 (Density of extremal rank-one POVMs)**.** *Let A be a quantum system of dimension d and let $N \leq d^2$. Let $\{|m_x\rangle\langle m_x|\}_x$ be rank-one POVM on A with N outcomes. Then for all $\varepsilon > 0$, there exists an extremal rank-one POVM $\{|e_x\rangle\langle e_x|\}_x$ with N outcomes such that*

$$\||m_x\rangle\langle m_x| - |e_x\rangle\langle e_x|\| \leq \varepsilon \tag{133}$$

for all x.

Proof. If $\{|m_x\rangle\langle m_x|\}_x$ is extremal then the statement is immediately true. Suppose thus that the measurement is not extremal. This means that its elements are linearly dependent when viewed as elements of the $\mathbb{R}$-vector space $\text{Herm}(\mathbb{C}^d)$ [26, Corollary 2.50]. That is, there exist some real coefficients $\alpha_1, \cdots, \alpha_N$ that are not all zero such that

$$\sum_{x=0}^{N-1} \alpha_x |m_x\rangle\langle m_x| = 0 . \tag{134}$$

Let S denote the real span of $\{|m_x\rangle\langle m_x|\}_x$. Since $\{|m_x\rangle\langle m_x|\}_x$ is linearly dependent, we have $\dim(S) < N \leq d^2$, with d^2 being the dimension of $\text{Herm}(\mathbb{C}^d)$ viewed as a $\mathbb{R}$-vector space. In particular, this means we can find a subspace E orthogonal to S with non-zero dimension such that

$$\text{Herm}(\mathbb{C}^d) = S \oplus E \tag{135}$$

where the orthogonality is defined using the Hilbert-Schmidt inner product $\langle X, Y \rangle = \mathrm{Tr}\,[XY]$. Now consider any $X \in E$, by the spectral theorem $X = \sum_i \lambda_i \, |w_i\rangle\langle w_i|$ with $\lambda_i \in \mathbb{R}$. Thus at least one of the $|w_i\rangle\langle w_i| \notin S$ otherwise $X \in S$. Let us denote by $|w\rangle\langle w|$ such a projector and let us decompose it as

$$|w\rangle\langle w| = Y_S + Y_E \tag{136}$$

with $Y_S \in S$ and $Y_E \in E$, by assumption $Y_E \neq 0$.

Now, without loss of generality, let us assume that $\alpha_{N-1} \neq 0$, which means that $|m_{N-1}\rangle\langle m_{N-1}| \in S$. For any $\varepsilon \in (0, 1)$, let us define

$$|\tau_\varepsilon\rangle := \sqrt{1 - \varepsilon}\, |m_{N-1}\rangle + \sqrt{\varepsilon}\, |w\rangle \ . \tag{137}$$

We want to show that $|\tau_\varepsilon\rangle\langle\tau_\varepsilon| \notin S$, so that we can replace $|m_{N-1}\rangle\langle m_{N-1}|$ by $|\tau_\varepsilon\rangle\langle\tau_\varepsilon|$ to increase the dimension of the span of the measurement. For the sake of the contradiction, let us assume that $|\tau_\varepsilon\rangle\langle\tau_\varepsilon| \in S$. We then have

$$0 = \mathrm{Tr}\,[|\tau_\varepsilon\rangle\langle\tau_\varepsilon|\, Y_E] = \varepsilon\,\mathrm{Tr}\,[Y_E^2] + \sqrt{\varepsilon(1-\varepsilon)}\mathrm{Tr}\,[Y_E\,(|w\rangle\langle m_{N-1}| + |m_{N-1}\rangle\langle w|)] \ . \tag{138}$$

Since $\mathrm{Tr}\,[Y_E^2] > 0$, there is at most a single $\varepsilon \in (0, 1)$ such that this equation is satisfied. Hence, it is always possible to find an arbitrary small $\varepsilon \in (0, 1)$ such that $|\tau_\varepsilon\rangle\langle\tau_\varepsilon| \notin S$. Thus, by replacing $|m_{N-1}\rangle\langle m_{N-1}|$ by $|\tau_\varepsilon\rangle\langle\tau_\varepsilon|$, the dimension of S increases by 1.

To make sure the new collection of operators forms a POVM we may need to renormalize. In particular, define

$$G_\varepsilon = |\tau_\varepsilon\rangle\langle\tau_\varepsilon| + \sum_{x=0}^{N-2} |m_x\rangle\langle m_x| \tag{139}$$

and then consider, for $x \in [N-1]$

$$|m_x'\rangle\langle m_x'| := G_\varepsilon^{-\frac{1}{2}} |m_x\rangle\langle m_x|\, G_\varepsilon^{-\frac{1}{2}} \tag{140}$$

and $|m_{N-1}'\rangle\langle m_{N-1}'| = G_\varepsilon^{-\frac{1}{2}} |\tau_\varepsilon\rangle\langle\tau_\varepsilon|\, G_\varepsilon^{-\frac{1}{2}}$. Such a construction yields a rank-one POVM whose span has dimension $\dim(S) + 1$. We now show that we can make this new POVM arbitrarily close to the original POVM, first note that

$$\begin{aligned}
G_\varepsilon - I_d &= |\tau_\varepsilon\rangle\langle\tau_\varepsilon| - |m_{N-1}\rangle\langle m_{N-1}| \\
&= \varepsilon\left(|w\rangle\langle w| - |m_{N-1}\rangle\langle m_{N-1}|\right) + \sqrt{\varepsilon(1-\varepsilon)}\left(|m_{N-1}\rangle\langle w| + |w\rangle\langle m_{N-1}|\right)
\end{aligned} \tag{141}$$

and thus $\lim_{\varepsilon\to 0} G_\varepsilon = I_d$. Now, by continuity of $M \mapsto M^{-\frac{1}{2}}$ on some open ball with center I_d, we also have that $G_\varepsilon^{-\frac{1}{2}}$ converges to I_d as ε goes to 0. In particular, for all $x \in [N]$, we have that $G_\varepsilon^{-\frac{1}{2}} |m_x'\rangle\langle m_x'|\, G_\varepsilon^{-\frac{1}{2}}$ converges to $|m_x\rangle\langle m_x|$ as ε goes to 0.

Thus, we managed to find a rank-one POVM such that the dimension of the span of its elements is strictly larger than the previous one, while being arbitrarily close to the previous one. By iterating this process, we can finally find an extremal rank-one POVM arbitrarily close to the original one. $\qquad\square$