Corollary 8.7. *As function* $\mathcal{D}(AE) \to \mathbb{R}$*, the guessing probability* $\text{Guess}(A|E)$ *is* convex.

We conclude the subsection by showing that the **chain rule** (point 3. of Lemma A.4) still holds.

Proposition 8.8. *For all states* $\rho_{ABE} \in \mathcal{D}(ABE)$

$$\text{H}_\infty(A|BE) \geq \text{H}_\infty(AB|E) - \log \text{rank}(\rho_B) \geq \text{H}_\infty(AB|E) - \log \dim(\mathcal{H}_B)\,.$$

Proof. Let $\lambda > 0$ and $\sigma_E \in \mathcal{D}(E)$ with $\text{supp}(\rho_E) \subseteq \text{supp}(\sigma_E)$ so that $\lambda \cdot \mathbb{I}_A \otimes \mathbb{I}_B \otimes \sigma_E \geq \rho_{ABE}$. Then also $\lambda \cdot \mathbb{I}_A \otimes \rho_B^0 \otimes \sigma_E \geq \rho_B^0\, \rho_{ABE}\, \rho_B^0 = \rho_{ABE}$, where the equality is due to Corollary 8.2 and Remark 0.2. So, for $\sigma_B := \rho_B^0/\text{rank}(B) \in \mathcal{D}(\mathcal{H}_B)$ we have $\lambda \cdot \text{rank}(\rho_B) \cdot \mathbb{I}_A \otimes \sigma_B \otimes \sigma_E \geq \rho_{ABE}$. □

8.2 Min-Entropy of Superpositions

We start with the following technical observation, which relates a *superposition* of states to the corresponding *mixture* of the states.

Lemma 8.9. *Consider* $|\varphi\rangle \in \mathcal{S}(\mathcal{H})$ *and* $\tilde{\rho} \in \mathcal{D}(\mathcal{H})$ *of the form*

$$|\varphi\rangle = \sum_{x \in \mathcal{X}_\circ} \alpha_x |x\rangle \qquad \textit{and} \qquad \tilde{\rho} = \sum_{x \in \mathcal{X}_\circ} |\alpha_x|^2 |x\rangle\langle x|\,,$$

where $\{|x\rangle\}_{x\in\mathcal{X}}$ *is an orthonormal basis of* $\mathcal{H}$ *and* $\mathcal{X}_\circ \subseteq \mathcal{X}$*. Then,* $|\varphi\rangle\langle\varphi| \leq |\mathcal{X}_\circ| \cdot \tilde{\rho}$*.*

Proof. Let $|\psi\rangle \in \mathcal{H}$. Then

$$\begin{aligned}\langle\psi||\varphi\rangle\langle\varphi||\psi\rangle = |\langle\psi|\varphi\rangle|^2 &= \Big|\sum_x \alpha_x \langle\psi|x\rangle\Big|^2 \\ &\leq |\mathcal{X}_\circ| \sum_x |\alpha_x\langle\psi|x\rangle|^2 = |\mathcal{X}_\circ| \sum_x |\alpha_x|^2 \langle\psi|x\rangle\langle x|\psi\rangle = |\mathcal{X}_\circ| \cdot \langle\psi|\tilde{\rho}|\psi\rangle\end{aligned}$$

where the inequality is obtained by viewing $\sum_x \alpha_x\langle\psi|x\rangle$ as inner product of the length-$|\mathcal{X}_\circ|$ vectors with entries $\alpha_x\langle\psi|x\rangle$ and 1, and applying Cauchy-Schwarz inequality. □

This allows us to related the min-entropy of a superposition to the min-entropy of its mixture.

Proposition 8.10. *Let* $|\Omega\rangle \in \mathcal{S}(AE)$ *be a superposition*

$$|\Omega\rangle = \sum_{x \in \mathcal{X}_\circ} \alpha_x |x\rangle|\varphi^x\rangle$$

where $\{|x\rangle\}_{x\in\mathcal{X}}$ *is an orthonormal basis of* $\mathcal{H}_A$ *and* $\mathcal{X}_\circ \subseteq \mathcal{X}$*, and let* $\rho_{AE}, \tilde{\rho}_{AE} \in \mathcal{D}(AE)$ *be given by*

$$\rho_{AE} = |\Omega\rangle\langle\Omega| \qquad \textit{and} \qquad \tilde{\rho}_{AE} = \sum_{x \in \mathcal{X}_\circ} |\alpha_x|^2 |x\rangle\langle x| \otimes |\varphi^x\rangle\langle\varphi^x|\,.$$

Then, for any $\sigma_E \in \mathcal{D}(E)$

$$\text{H}_\infty(\rho_{AE}|\sigma_E) \geq \text{H}_\infty(\tilde{\rho}_{AE}|\sigma_E) - \log|\mathcal{X}_\circ|\,.$$

Furthermore, if $\mathfrak{T}$ *is a CPTP map that acts on either of* A *and* E*, or on both, then*

$$\text{H}_\infty(\mathfrak{T}(\rho_{AE})|\sigma_E) \geq \text{H}_\infty(\mathfrak{T}(\tilde{\rho}_{AE})|\sigma_E) - \log|\mathcal{X}_\circ|\,.$$

Proof. Let $\lambda > 0$ be so that $\lambda \cdot \mathbb{I} \otimes \sigma_E \geq \tilde{\rho}_{AE}$. Then, by Lemma 8.9 above, it follows immediately that $|\mathcal{X}_\circ|\lambda \cdot \mathbb{I} \otimes \sigma_E \geq \rho_{AE}$, which proves the first inequality. The extension involving the CPTP map follows by observing that if $|\mathcal{X}_\circ| \cdot \tilde{\rho}_{AE} \geq \rho_{AE}$ then also $|\mathcal{X}_\circ| \cdot \mathfrak{T}(\tilde{\rho}_{AE}) \geq \mathfrak{T}(\rho_{AE})$, because a CPTP map preserves positivity, and thus the same argument still applies. □

8.3 Rényi Divergence and Rényi Entropy

We now introduce and discuss a *parameterized* entropy notion $\mathrm{H}_\alpha(A)$, respectively $\mathrm{H}_\alpha(A|E)$ for the conditional version, referred to as (conditional) **Rényi entropy**, which recovers the definition of the min-entropy H_∞ when taking the limit $\alpha \to \infty$. In the limit $\alpha \to 1$, it recovers the well-known **Von Neumann entropy**, the quantum extension of the **Shannon entropy** (Definitions A.2 and A.3). The case $\alpha = 2$ is referred to as **collision entropy**.

Like for the min-entropy, the definition for the unconditional version is quite harmless.

Definition 8.4. *Let* $0 < \alpha < 1$ *or* $1 < \alpha < \infty$. *For a given* $\rho_A \in \mathcal{D}(A)$*, the* **Rény entropy** *of order* α *is defined as*

$$\mathrm{H}_\alpha(A) := \mathrm{H}_\alpha(\rho_A) := \frac{1}{1-\alpha} \log \mathrm{tr}\big(\rho_A^\alpha\big) = \frac{1}{1-\alpha} \log \|\rho_A\|_\alpha^\alpha .$$

When bringing ρ_A into diagonal form, the classical results (see Appendix A.3) apply, and thus $\mathrm{H}_\alpha(A) \to \mathrm{H}_\infty(A)$ for $\alpha \to \infty$, $\mathrm{H}_\alpha(A) \to \mathrm{H}_0(A) := \log \mathrm{rank}(\rho_A)$ for $\alpha \to 0$, and $\mathrm{H}_\alpha(A) \to \mathrm{H}_1(A)$ for $\alpha \to 1$, where

$$\mathrm{H}_1(A) := \mathrm{H}(A) := -\mathrm{tr}(\rho_A \log \rho_A)$$

is known as **Von Neumann entropy**.

The conditional version is again trickier. It will be useful to first define the following. We remark that it is convenient to think of D_α, when applied to density operators, as a distance measure, even though it is not a metric.

Definition 8.5. *Let* $0 < \alpha \neq 1$. *For* $\rho \in \mathcal{D}(\mathcal{H})$ *and* $\sigma \in \mathcal{P}(\mathcal{H})$*, the* **Rény divergence** *of order* α *is defined as*

$$\mathrm{D}_\alpha(\rho\|\sigma) := \frac{1}{\alpha-1} \log \mathrm{tr}\Big(\big(\sigma^{\frac{1-\alpha}{2\alpha}} \rho\, \sigma^{\frac{1-\alpha}{2\alpha}}\big)^\alpha\Big) = \frac{\alpha}{\alpha-1} \log\big\|\sigma^{\frac{1-\alpha}{2\alpha}} \rho\, \sigma^{\frac{1-\alpha}{2\alpha}}\big\|_\alpha$$

whenever $\mathrm{supp}(\rho) \subseteq \mathrm{supp}(\sigma)$*, or* $\alpha < 1$ *and* $\mathrm{supp}(\rho) \not\perp \mathrm{supp}(\sigma)$*, and* $\mathrm{D}_\alpha(\rho, \sigma) := \infty$ *otherwise.*

By taking the limit "component-wise", one expects — and this is indeed the case — that

$$\mathrm{D}_\infty(\rho_{AE} \,\|\, \mathbb{I}_A \otimes \sigma_E) := \lim_{\alpha\to\infty} \mathrm{D}_\alpha(\rho_{AE} \,\|\, \mathbb{I}_A \otimes \sigma_E) = \log\big\|\sigma_E^{-1/2} \rho_{AE}\, \sigma_E^{-1/2}\big\|_\infty = -\mathrm{H}_\infty(\rho_{AE}|\sigma_E) .$$

This indicates that we are on the right track and motivates the following definition.

Definition 8.6. *Let* $0 < \alpha \neq 1$. *For a given* $\rho_{AE} \in \mathcal{D}(AE)$ *we define*

$$\mathrm{H}_\alpha(A|E) := -\min_{\sigma_E} \mathrm{D}_\alpha(\rho_{AE} \,\|\, \mathbb{I}_A \otimes \sigma_E) ,$$

where the min *is over all* $\sigma_E \in \mathcal{D}(E)$.

For an "empty" system $E = \emptyset$, i.e., $\mathcal{H}_E = \mathbb{C}$, we recover $\mathrm{H}_\alpha(A|E) = \mathrm{H}_\alpha(A)$.

Remark 8.4. Using similar reasoning as in the case of the min-entropy, it is good enough to quantify over σ_E with $\mathrm{supp}(\sigma_E) \subseteq \mathrm{supp}(\rho_E)$, and so we may assume without loss of generality that $\mathcal{H}_E = \mathrm{supp}(\rho_E)$. This then also ensures that the minimum is indeed attained, given that the objective function is continuous and the optimization is over a compact set.

Compared to $\alpha \to \infty$, the limit $\alpha \to 1$ is less clear. We state here without proof nor intuition that $\mathrm{D}_\alpha(\rho_{AE}, \mathbb{I}_A \otimes \sigma_E) \to \mathrm{tr}\big(\rho_{AE} \log(\rho_{AE}) - \rho_{AE} \log(\sigma_E)\big)$. Furthermore, one can show that here the optimal choice for σ_E is ρ_E, so that

$$\mathrm{H}_1(A|E) = -\mathrm{tr}\big(\rho_{AE} \log(\rho_{AE}) - \rho_{AE} \log(\rho_E)\big) = \mathrm{H}(AE) - \mathrm{H}(E) ,$$

which is the *conditional* Van Neumann entropy $\mathrm{H}(A|E)$.

D_α is *monotonically increasing* in α, and thus $\mathrm{H}_\alpha(A|E)$ is *monotonically decreasing* in α. We only show the following special case here.

Proposition 8.11. *For any $\rho \in \mathcal{D}(\mathcal{H})$ and $\sigma \in \mathcal{P}(\mathcal{H})$*

$$\mathrm{D}_2(\rho\|\sigma) \leq \mathrm{D}_\infty(\rho\|\sigma)\,.$$

Proof. Let $\rho = \sum_i \lambda_i |i\rangle\langle i|$ be the spectral decomposition of ρ. Then, we have

$$\begin{aligned}\mathrm{D}_2(\rho\|\sigma) &= \log\operatorname{tr}\big((\sigma^{-1/4}\rho\,\sigma^{-1/4})^2\big) = \log\operatorname{tr}\big(\rho\,\sigma^{-1/2}\rho\,\sigma^{-1/2}\big) = \log\sum_i \lambda_i\langle i|\sigma^{-1/2}\rho\,\sigma^{-1/2}|i\rangle\\ &\leq \log\sum_i \lambda_i\,\lambda_{\max}\big(\sigma^{-1/2}\rho\,\sigma^{-1/2}\big) = \log\lambda_{\max}\big(\sigma^{-1/2}\rho\,\sigma^{-1/2}\big) = \mathrm{D}_\infty(\rho\|\sigma)\,,\end{aligned}$$

which proves the claim. □

The Rényi entropy satisfies similar properties than the min-entropy, like monotonicity for classical subsystems and data-processing inequality (for $\alpha \geq \frac{1}{2}$) and chain rule, similar to Propositions 8.4, 8.5 and 8.8. The chair rule is comparably simple.

Proposition 8.12. *For $0 < \alpha \leq \infty$ and any $\rho_{ABE} \in \mathcal{D}(ABE)$, we have*

$$\mathrm{H}_\alpha(A|BE) \geq \mathrm{H}_\alpha(AB|E) - \mathrm{H}_0(B)$$

Proof. Using Corollary 8.2, which ensures $\operatorname{supp}(\rho_{ABE}) \subseteq \operatorname{supp}(\mathbb{I}_A \otimes \rho_B \otimes \mathbb{I}_E)$, and Remark 0.2, we observe that $\rho_B^0\,\rho_{ABE}\,\rho_B^0 = \rho_{ABE}$. So, setting $\sigma_B := \rho_B^0/\operatorname{rank}(\rho_B)$, for $\frac{1}{2} \neq \alpha < \infty$ we obtain

$$\begin{aligned}\mathrm{H}_\alpha(A|BE) &\geq -\mathrm{D}_\alpha(\rho_{ABE}\|\mathbb{I}_A\otimes\sigma_B\otimes\sigma_E) = \frac{\alpha}{1-\alpha}\log\Big\|\big(\sigma_B\otimes\sigma_E\big)^{\frac{1-\alpha}{2\alpha}}\rho_{ABE}\,\big(\sigma_B\otimes\sigma_E\big)^{\frac{1-\alpha}{2\alpha}}\Big\|_\alpha\\ &= \frac{\alpha}{1-\alpha}\log\Big\|\sigma_E^{\frac{1-\alpha}{2\alpha}}\rho_B^0\,\rho_{ABE}\,\rho_B^0\,\sigma_E^{\frac{1-\alpha}{2\alpha}}\Big\|_\alpha - \log\operatorname{rank}(\rho_B) = -\mathrm{D}_\alpha(\rho_{ABE}\|\mathbb{I}_{AB}\otimes\sigma_E) - \mathrm{H}_0(B)\end{aligned}$$

for any $\sigma_E \in \mathcal{D}(E)$, and so the claim follows by maximizing over σ_E. The cases $\alpha = 1$ and ∞ follow by taking limits. □

Arguing the monotonicity for classical subsystems and the data-processing inequality is substantially more involved, and will be obtained in the upcoming sections.

8.4 Data-Processing Inequality

The goal of this section is to prove the data-processing inequality for the divergence D_α.

Theorem 8.13. *Let $\frac{1}{2} \leq \alpha \leq \infty$. For any $\rho \in \mathcal{D}(\mathcal{H})$, $\sigma \geq 0$ and CPTP map $\mathfrak{T} \in \mathcal{L}\big(\mathcal{L}(\mathcal{H}),\mathcal{L}(\mathcal{H}')\big)$:*

$$\mathrm{D}_\alpha(\mathfrak{T}(\rho),\mathfrak{T}(\sigma)) \leq \mathrm{D}_\alpha(\rho,\sigma)\,.$$

The data-processing inequality for the conditional entropy H_α then follows immediately.

Corollary 8.14. *For α as above, $\rho_{AE} \in \mathcal{D}(AE)$ and $\mathfrak{T} \in \mathcal{L}\big(\mathcal{L}(E),\mathcal{L}(E')\big)$ a CPTP map:*

$$\mathrm{H}_\alpha\big(A|\mathfrak{T}(E)\big) \geq \mathrm{H}_\alpha(A|E)\,,$$

where $\mathrm{H}_\alpha\big(A|\mathfrak{T}(E)\big)$ is naturally understood as $\mathrm{H}_\alpha(A|E')$ for $\rho_{AE'} = \mathfrak{T}_{E\to E'}(\rho_{AE})$.

Remark 8.5. Another immediate consequence is that $\mathrm{D}_\alpha(\rho,\sigma) \geq 0$ for all $\rho,\sigma \in \mathcal{D}(\mathcal{H})$, and that $\mathrm{H}_\alpha(A|E) \leq \mathrm{H}_\alpha(A) \leq \log\dim(\mathcal{H}_A)$; indeed, this follows by setting $\mathfrak{T} := \mathrm{tr}$.

For the proof of the data-processing inequality for the divergence, it will be convenient to consider a variant of D_α that is without the $\frac{1}{\alpha-1}$-scaling and without the log. Furthermore, for simplicity and since this is good enough to conclude Corollary 8.14, we restrict to where we restrict to *full-rank*, i.e., *invertible*, $\rho \in \mathcal{D}(\mathcal{H})$ and $\sigma \in \mathcal{P}(\mathcal{H})$, denoted respectively as $\mathcal{D}^\star(\mathcal{H})$ and $\mathcal{P}^\star(\mathcal{H})$. Thus, we define

$$\mathrm{d}_\alpha : \mathcal{D}^\star(\mathcal{H}) \times \mathcal{P}^\star(\mathcal{H}) \to \mathbb{R}\,,\; (\rho,\sigma) \mapsto \mathrm{tr}\Big(\big(\sigma^{\frac{1-\alpha}{2\alpha}}\rho\,\sigma^{\frac{1-\alpha}{2\alpha}}\big)^\alpha\Big)\,.$$

Our goal is to show that d_α is decreasing under the partial trace when $\alpha > 1$, respectively increasing when $\alpha < 1$. By the monotonicity of the log, this implies the data-processing inequality for $\mathfrak{T}$ being the partial trace — for full-rank ρ and σ, and in general due to continuity. By the Stinespring representation, and observing that D_α is invariant under isometries, this then implies Theorem 8.13 for arbitrary $\mathfrak{T}$. The cases $\alpha = 1$ and ∞ follow by taking limits.

We start with the following convexity/concavity property.

Lemma 8.15. *For $\alpha > 1$ ($\frac{1}{2} \leq \alpha < 1$), d_α is jointly convex (concave).*

The strategy of the proof is to reduce the claimed convexity statement to the celebrated **Lieb's Concavity Theorem**, as given in Corollary B.13 in Appendix B.4.

Proof. First, consider $\alpha > 1$. Applying Corollary 7.4 with $p = \alpha$ and optimizing over the choice of R, we obtain

$$\begin{aligned}\mathrm{tr}\Big(\big(\sigma^{\frac{1-\alpha}{2\alpha}}\rho\,\sigma^{\frac{1-\alpha}{2\alpha}}\big)^\alpha\Big) &= p \cdot \max_{R\geq 0}\Big\{\mathrm{tr}\big(\sigma^{\frac{1-\alpha}{2\alpha}}\rho\,\sigma^{\frac{1-\alpha}{2\alpha}}R\big) - \tfrac{1}{q}\mathrm{tr}(R^q)\Big\}\\ &= p \cdot \max_{R\geq 0}\Big\{\mathrm{tr}\big(\rho\,\sigma^{-\frac{1}{2q}}R\,\sigma^{-\frac{1}{2q}}\big) - \tfrac{1}{q}\mathrm{tr}(R^q)\Big\}\,,\end{aligned}$$

where the second equality uses $\frac{1}{q} = 1 - \frac{1}{p} = \frac{\alpha-1}{\alpha}$. Writing $W = \sigma^{-\frac{1}{2q}}R\,\sigma^{-\frac{1}{2q}}$, we then obtain

$$\begin{aligned}\mathrm{tr}\Big(\big(\sigma^{\frac{1-\alpha}{2\alpha}}\rho\,\sigma^{\frac{1-\alpha}{2\alpha}}\big)^\alpha\Big) &= p \cdot \max_{W\geq 0}\Big\{\mathrm{tr}(\rho W) - \tfrac{1}{q}\mathrm{tr}\Big(\big(\sigma^{\frac{1}{2q}}W\sigma^{\frac{1}{2q}}\big)^q\Big)\Big\}\\ &= p \cdot \max_{W\geq 0}\Big\{\mathrm{tr}(\rho W) - \tfrac{1}{q}\mathrm{tr}\Big(\big(W^{1/2}\sigma^{1/q}W^{1/2}\big)^q\Big)\Big\}\,,\end{aligned}$$

where the second equality is due to Lemma 0.4. The goal is now to show that the term $\frac{1}{q}\mathrm{tr}(\cdot)$ is concave as a function of σ for every W. This then ensures that the entire objective function is (jointly) convex as a function of (ρ,σ) for every W, since $\mathrm{tr}(\rho W)$ is linear in ρ and thus trivially convex. It then follows that the max over W is convex as well (see Remark B.2), which then proves the claim.

In order to argue concavity, we again apply Corollary 7.4, now to $R = W^{1/2}\sigma^{1/q}W^{1/2}$ and optimize over L. Writing L as $L = Z^{1/p}$, we then obtain

$$\frac{1}{q}\mathrm{tr}\Big(\big(W^{1/2}\sigma^{1/q}W^{1/2}\big)^q\Big) = \max_{Z\geq 0}\Big\{\mathrm{tr}\big(Z^{1/p}W^{1/2}\sigma^{1/q}W^{1/2}\big) - \tfrac{1}{p}\mathrm{tr}(Z)\Big\}\,.$$

Here, $\mathrm{tr}\big(Z^{1/p}W^{1/2}\sigma^{1/q}W^{1/2}\big)$ is *jointly* concave in Z and σ by Corollary B.13 for any W. Thus, this holds for the entire objective function, and thus the concavity claim we were aiming for holds (see Remark B.2).

For $\alpha < 1$, we can show as above, but using the reverse version of Corollary 7.4, that

$$\begin{aligned}\operatorname{tr}\Big(\big(\sigma^{\frac{1-\alpha}{2\alpha}}\rho\,\sigma^{\frac{1-\alpha}{2\alpha}}\big)^{\alpha}\Big) &= p\cdot\min_{W\geq 0}\Big\{\operatorname{tr}(\rho W) - \tfrac{1}{q}\operatorname{tr}\Big(\big(W^{1/2}\sigma^{1/q}W^{1/2}\big)^{q}\Big)\Big\}\\ &= p\cdot\min_{W\geq 0}\Big\{\operatorname{tr}(\rho W) + \tfrac{1}{q'}\operatorname{tr}\Big(\big(W^{-1/2}\sigma^{1/q'}W^{-1/2}\big)^{q'}\Big)\Big\}\,,\end{aligned}$$

where the second equality is by setting $q' := -q$. It now remains to argue that $\frac{1}{q'}\operatorname{tr}(\cdot)$ is concave, which is done exactly as above, noting that $\frac{1}{q'} = \frac{1-\alpha}{\alpha} \leq 1$ for $\alpha \geq 1/2$, so that Corollary B.13 again applies. □

In order to get from convexity/concavity of d_α to the de-/increasingness of d_α under the partial trace, we first introduce the following.

Definition 8.7. *For a given orthornormal basis $\{|0\rangle,\ldots,|d-1\rangle\}$ of $\mathcal{H}$, the* **generalized Pauli operators** *X and Z are respectively defined as*

$$X := \sum_{\ell=0}^{d-1}|\ell{+}1\rangle\langle\ell| \qquad \text{and} \qquad Z := \sum_{\ell=0}^{d-1}e^{2\pi i\ell/d}|\ell\rangle\langle\ell|\,,$$

where the addition $\ell+1$ is understood to be modulo d.

The generalized Pauli operators X and Z satisfy the following basic properties, all easy to verify. They are both unitaries, $X^d = \mathbb{I} = Z^d$, and $ZX = e^{2\pi i/d}XZ$. Furthermore, $\operatorname{tr}(X) = 0 = \operatorname{tr}(Z)$.

Considering a fixed orthonormal basis, we introduce the superoperator

$$\mathfrak{E}:\mathcal{L}(\mathcal{H})\to\mathcal{L}(\mathcal{H})\,,\ R\mapsto\frac{1}{d^2}\sum_{x,z}X^xZ^zR\,(Z^\dagger)^z(X^\dagger)^x = \frac{1}{d^2}\sum_{x,z}Z^zX^xR\,(X^\dagger)^x(Z^\dagger)^z\,,$$

where the sum is over all x and y in $\mathbb{Z}/d\mathbb{Z}$, and where the equality holds because of the above commutativity property of X and Z.

Lemma 8.16. *As a superoperator,* $\mathfrak{E} = \frac{\mathbb{I}}{\dim(\mathcal{H})}\operatorname{tr} : R\mapsto\frac{\operatorname{tr}(R)}{\dim(\mathcal{H})}\,\mathbb{I}\,.$

Proof. We observe that $\mathfrak{E}(R)$ commutes with X and Z. Indeed,

$$X\mathfrak{E}(R) = \frac{1}{d^2}\sum_{x,z}X^{x+1}Z^zR\,(Z^\dagger)^z(X^\dagger)^x = \frac{1}{d^2}\sum_{x,z}X^xZ^zR\,(Z^\dagger)^z(X^\dagger)^{x-1} = \mathfrak{E}(R)X$$

and similarly for Z. From this it follows that $\mathfrak{E}(R) = \lambda(R)\cdot\mathbb{I}$ for some scalar $\lambda(R)$. Indeed, writing $\mathfrak{E}(R) = \sum_{k,\ell}\varepsilon_{k\ell}|k\rangle\langle\ell|$ and applying $\langle k|\cdot|\ell\rangle$ to both sides of $Z\mathfrak{E}(R) = \mathfrak{E}(R)Z$ shows that $\varepsilon_{k\ell} = 0$ for $k\neq\ell$. Furthermore, applying $\langle\ell{+}1|\cdot|\ell\rangle$ then to $X\mathfrak{E}(R) = \mathfrak{E}(R)X$ shows that $\varepsilon_{\ell\ell} = \varepsilon_{\ell+1,\ell+1}$. The claim now follows from observing that $\lambda(R)\cdot\operatorname{tr}(\mathbb{I}) = \operatorname{tr}\big(\mathfrak{E}(R)\big) = \operatorname{tr}(R)$. □

We are now ready to prove the decreasingness of d_α under the partial trace for $\alpha > 1$, and correspondingly for $\alpha < 1$, from which Theorem 8.13 then follows. Noting that

$$\operatorname{tr}_E(\rho_{AE})\otimes\frac{1}{d}\mathbb{I}_E = \Big(id_A\otimes\frac{1}{d}\mathbb{I}_E\operatorname{tr}\Big)(\rho_{AE}) = (id_A\otimes\mathfrak{E}_E)(\rho_{AE}) = \frac{1}{d^2}\sum_{x,y}X_E^xZ_E^z\rho_{AE}\,(Z_E^\dagger)^z(X_E^\dagger)^x\,,$$

and correspondingly for σ_{AE}, we obtain that

$$\begin{aligned}\mathrm{d}_\alpha\big(\operatorname{tr}_E(\rho_{AE}),\operatorname{tr}_E(\sigma_{AE})\big) &= \mathrm{d}_\alpha\Big(\tfrac{1}{d^2}\textstyle\sum_{x,y}X_E^xZ_E^z\rho_{AE}\,(Z_E^\dagger)^z(X_E^\dagger)^x,\tfrac{1}{d^2}\sum_{x,y}X_E^xZ_E^z\sigma_{AE}\,(Z_E^\dagger)^z(X_E^\dagger)^x\Big)\\ &\leq\frac{1}{d^2}\sum_{x,y}\mathrm{d}_\alpha\Big(X_E^xZ_E^z\rho_{AE}\,(Z_E^\dagger)^z(X_E^\dagger)^x, X_E^xZ_E^z\sigma_{AE}\,(Z_E^\dagger)^z(X_E^\dagger)^x\Big) = \mathrm{d}_\alpha\big(\rho_{AE},\sigma_{AE}\big)\end{aligned}$$

where the inequality is Jensen's inequality, and the final equality is because of the unitary invariance of d_α.

8.5 Duality Property of Rényi Entropy

We contionue the discussion of the general Rényi entropy by showing the following duality relation. With this relation, we can now for instance easily verify that the conditional min-entropy of an EPR pair AE is indeed $\mathrm{H}_\infty(A|E) = -\mathrm{H}_{1/2}(A) = -1$, as we claimed earlier. The same for H_α. Furthermore, it follows that $\mathrm{H}_\alpha(A|E) \geq -\log\dim(\mathcal{H}_A)$ for any state ρ_{AE}.

Theorem 8.17. *Let $\frac{1}{2} \leq \alpha, \beta \leq \infty$ with $\frac{1}{\alpha}+\frac{1}{\beta} = 2$. Then, for any* pure $\rho_{ABE} = |\psi\rangle\langle\psi| \in \mathcal{D}(ABE)$*:*

$$\mathrm{H}_\alpha(A|B) + \mathrm{H}_\beta(A|E) = 0\,.$$

For non-pure states, the inequality ≥ 0 holds.

For the proof, we introduce the following technical tool.

Lemma 8.18. *Let $\rho_{AE} = |\varphi\rangle\langle\varphi| \in \mathcal{D}(AE)$, and set $\rho_A := \mathrm{tr}_E(\rho_{AE})$ and $\rho_E := \mathrm{tr}_A(\rho_{AE})$. Then, for every Hermitian $R_A \in \mathcal{L}(A)$ with $\mathrm{supp}(R_A) \subseteq \mathrm{supp}(\rho_A)$ there exists a Hermitian $R_E \in \mathcal{L}(E)$ with $\mathrm{supp}(R_E) \subseteq \mathrm{supp}(\rho_E)$ — and vice versa — with the same list of non-zero eigenvalues and so that for every $L_A \in \mathcal{L}(A)$*

$$\mathrm{tr}\big(\sqrt{\rho_A}\, L_A \sqrt{\rho_A}\, R_A\big) = \langle\varphi| L_A \otimes R_E |\varphi\rangle\,.$$

Proof. Let $|\varphi\rangle = \sum_i \mu_i |e_i\rangle|f_i\rangle \in \mathcal{S}(AE)$ be the Schmidt decomposition. We may assume that $\mathcal{H}_A = \mathrm{supp}(\rho_A)$ and $\mathcal{H}_E = \mathrm{supp}(\rho_E)$ (because if not, we restrict to the respective subspaces $\mathrm{supp}(\rho_A)$ and $\mathrm{supp}(\rho_E)$). Also, replacing R_E by $U^\dagger R_E U$ for a suitable unitary U, we may assume that $|e_i\rangle = |f_i\rangle$. But then, given that $\sqrt{\rho_A} = \sum_i \sqrt{\mu_i}|e_i\rangle\langle e_i|$ and invoking the notation from Section 0.7, we have $|\sqrt{\rho_A}\rangle = |\varphi\rangle$. Hence, applying Corollary 0.5, we obtain

$$\mathrm{tr}\big(\sqrt{\rho_A}\, L_A \sqrt{\rho_A}\, R_A\big) = \langle\varphi| L_A \otimes R_A^T |\varphi\rangle\,,$$

proving the claim. □

Proof. We may assume that $\alpha < 1$ and $\beta > 1$, and it will be useful to set $0 < \alpha' := \frac{1-\alpha}{\alpha} = \frac{1}{\alpha} - 1 = 1 - \frac{1}{\beta} = -\frac{1-\beta}{\beta} =: -\beta'$; with $\alpha' = 1$ and $\beta' = -1$ in case $\alpha = \frac{1}{2}$ and $\beta = \infty$. Note that, by definition and using Lemma 0.4,

$$\begin{aligned}\mathrm{H}_\alpha(A|B) &= -\min_{\sigma_B} \mathrm{D}_\alpha(\rho_{AB}\,\|\,\mathbb{I}_A \otimes \sigma_B) = -\min_{\sigma_B} \frac{\alpha}{\alpha-1} \log\Big\|\sigma_B^{\frac{1-\alpha}{2\alpha}}\, \rho_{AB}\, \sigma_B^{\frac{1-\alpha}{2\alpha}}\Big\|_\alpha \\ &= \frac{1}{\alpha'} \log\max_{\sigma_B} \big\|\sigma_B^{\alpha'/2} \rho_{AB}\, \sigma_B^{\alpha'/2}\big\|_\alpha = \frac{1}{\alpha'} \log\max_{\sigma_B} \big\|\rho_{AB}^{1/2} \sigma_B^{\alpha'} \rho_{AB}^{1/2}\big\|_\alpha,\end{aligned}$$

with the understanding that the optimization is over all $\sigma_B \in \mathcal{D}(\mathcal{H}_B)$ with $\mathrm{supp}(\sigma_B) \subseteq \mathrm{supp}(\rho_B)$. Set $p := -\frac{1}{\beta'}$ and $q := -\frac{1}{\alpha'}$ so that $\frac{1}{\alpha} + \frac{1}{q} = \frac{1}{\alpha} + \frac{\alpha-1}{\alpha} = 1$ and correspondingly for β and p. Then, by the reverse Hölder inequality, we get that for any $\tau_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ with $\mathrm{supp}(\tau_{AB}) \subseteq \mathrm{supp}(\rho_{AB})$

$$\big\|\rho_{AB}^{1/2} \sigma_B^{\alpha'} \rho_{AB}^{1/2}\big\|_\alpha = \big\|\rho_{AB}^{1/2} \sigma_B^{\alpha'} \rho_{AB}^{1/2}\big\|_\alpha \big\|\tau_{AB}^{-\alpha'}\big\|_q \leq \mathrm{tr}\big(\rho_{AB}^{1/2} \sigma_B^{\alpha'} \rho_{AB}^{1/2} \tau_{AB}^{-\alpha'}\big)$$

where the equality is because $\big\|\tau_{AB}^{-\alpha'}\big\|_q = \mathrm{tr}(\tau_{AB})^{-\alpha'} = 1$.

Similarly, but noting that $\frac{1}{\beta'} < 0$,

$$\mathrm{H}_\beta(A|E) = \frac{1}{\beta'} \log\min_{\tau_E} \big\|\rho_{AE}^{1/2} \tau_E^{\beta'} \rho_{AE}^{1/2}\big\|_\beta = -\frac{1}{\alpha'} \log\min_{\tau_E} \big\|\rho_{AE}^{1/2} \tau_E^{\beta'} \rho_{AE}^{1/2}\big\|_\beta\,,$$

and, by the (ordinary) Hölder inequality now,

$$\|\rho_{AE}^{1/2}\tau_E^{\beta'}\rho_{AE}^{1/2}\|_\beta = \|\sigma_{AE}^{\alpha'}\|_p \|\rho_{AE}^{1/2}\tau_E^{\beta'}\rho_{AE}^{1/2}\|_\beta \geq \mathrm{tr}\big(\rho_{AE}^{1/2}\tau_E^{\beta'}\rho_{AE}^{1/2}\sigma_{AE}^{\alpha'}\big)$$

for any $\sigma_{AE} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_E)$ with $\mathrm{supp}(\sigma_{AE}) \subseteq \mathrm{supp}(\rho_{AE})$.

Finally, by invoking Lemma 8.18 twice, once with AB versus E, and once with AE versus B,

$$\mathrm{tr}\big(\rho_{AB}^{1/2}\sigma_B^{\alpha'}\rho_{AB}^{1/2}\tau_{AB}^{\beta'}\big) = \langle\psi|(\mathbb{I}_A \otimes \sigma_B^{\alpha'} \otimes \tau_E^{\beta'})|\psi\rangle = \mathrm{tr}\big(\rho_{AE}^{1/2}\tau_E^{\beta'}\rho_{AE}^{1/2}\sigma_{AE}^{\alpha'}\big)$$

for suitable choices of τ_{AB} and σ_{AE} as considered above. Putting things together, this proves $\mathrm{H}_\alpha(A|B) + \mathrm{H}_\beta(A|E) \leq 0$.

In order to argue equality, we observe that the two inequalities in the reasoning above become equalities when we minimize/maximize over τ_{AB} and σ_{AE}, and thus over τ_E and σ_B, respectively. Showing $\mathrm{H}_\alpha(A|B) + \mathrm{H}_\beta(A|E) = 0$ then reduces to swapping, say, the max over σ_B and the min over τ_E in front of the objective function $\langle\psi|(\mathbb{I}_A \otimes \sigma_B^{\alpha'} \otimes \tau_E^{\beta'})|\psi\rangle$. This can be done by means of **Von Neumann's minimax theorem** (Theorem B.6), given that the function is concave in σ_B and convex in τ_E. These latter properties follow from the results in Appendix B.4.

The claim for a non-pure state ρ_{ABE} follows by purifying the state to, say, ρ_{ABCE}, and using data-processing inequality to argue that $\mathrm{H}_\alpha(A|B) \geq \mathrm{H}_\alpha(A|BC)$. □

8.6 Monotonicity for Classical Subsystems

Finally, we show that, in line with Proposition 8.4 for the min-entropy, also the Rényi entropy H_α is monotone for classical subsystems.

Proposition 8.19. *For $\frac{1}{2} \leq \alpha \leq \infty$ and $\rho_{XAE} = \sum_x P_X(x)|x\rangle\langle x| \otimes \rho_{AE}^x \in \mathcal{D}(\mathcal{X} \otimes \mathcal{H}_A \otimes \mathcal{H}_E)$:*

$$\mathrm{H}_\alpha(XA|E) \geq \mathrm{H}_\alpha(A|E)\,.$$

For the proof, we introduce the following technical lemma, which can be appreciated as a variant of the data-processing inequality.

Lemma 8.20. *Let $\frac{1}{2} \leq \alpha \leq \infty$. Then, for any $\rho \in \mathcal{D}(\mathcal{H})$, $\sigma \in \mathcal{P}(\mathcal{H})$ and projection $\Pi \in \mathcal{L}(H)$:*

$$\mathrm{D}_\alpha(\Pi\rho\Pi \,\|\, \Pi\sigma\Pi) \leq \mathrm{D}_\alpha(\Pi\rho\Pi \,\|\, \sigma)\,.$$

Proof. We show the claim for $\frac{1}{2} \leq \alpha < 1$. The case $\alpha > 1$ works along the same lines, with the inequalities turned around, and one has to take some care with the support of σ, and, as usual, the cases $\alpha = 1$ and ∞ follow by taking limits.

Given that $0 < \frac{1-\alpha}{\alpha} \leq 1$, Jensen's operator inequality (Theorem B.7) implies that

$$\Pi\,\sigma^{\frac{1-\alpha}{\alpha}}\Pi = \Pi\,\sigma^{\frac{1-\alpha}{\alpha}}\Pi + (\mathbb{I}-\Pi)0^{\frac{1-\alpha}{\alpha}}(\mathbb{I}-\Pi) \leq \big(\Pi\sigma\Pi + (\mathbb{I}-\Pi)0(\mathbb{I}-\Pi)\big)^{\frac{1-\alpha}{\alpha}} = \big(\Pi\sigma\Pi\big)^{\frac{1-\alpha}{\alpha}}\,.$$

Hence,

$$\mathrm{d}_\alpha(\Pi\rho\Pi \,\|\, \sigma) = \mathrm{tr}\Big(\Big(\sigma^{\frac{1-\alpha}{2\alpha}}\,\Pi\rho\Pi\,\sigma^{\frac{1-\alpha}{2\alpha}}\Big)^\alpha\Big) = \mathrm{tr}\Big(\Big(\sqrt{\Pi\rho\Pi}\,\sigma^{\frac{1-\alpha}{\alpha}}\sqrt{\Pi\rho\Pi}\Big)^\alpha\Big)$$
$$= \mathrm{tr}\Big(\Big(\sqrt{\Pi\rho\Pi}\,\Pi\sigma^{\frac{1-\alpha}{\alpha}}\Pi\sqrt{\Pi\rho\Pi}\Big)^\alpha\Big) \leq \mathrm{tr}\Big(\Big(\sqrt{\Pi\rho\Pi}\,(\Pi\sigma\Pi)^{\frac{1-\alpha}{\alpha}}\sqrt{\Pi\rho\Pi}\Big)^\alpha\Big) = \mathrm{d}_\alpha(\Pi\rho\Pi \,\|\, \Pi\sigma\Pi)\,,$$

where the inequality is by Corollary 7.3. The claim follows by observing that $\mathrm{D}_\alpha = \frac{1}{\alpha-1}\log \mathrm{d}_\alpha$, turning the inequality around for $\alpha < 1$. □

Proof of Proposition 8.4. We extend ρ_{XAE} to

$$\rho_{XX'AE} = \sum_{x,x'} \sqrt{P_X(x)P_X(x')}|x\rangle\langle x'| \otimes |x\rangle\langle x'| \otimes \rho^x_{AE} \in \mathcal{D}(\mathcal{H}_X \otimes \mathcal{H}_{X'} \otimes \mathcal{H}_E)\,,$$

where $\mathcal{H}_X = \mathcal{H}_{X'} = \mathbb{C}^{|\mathcal{X}|}$, and we define the projection $\Pi_{XX'} := \sum_x |x\rangle\langle x| \otimes |x\rangle\langle x|$, for which it obviously holds that $\Pi_{XX'}\rho_{XX'AEF}\Pi_{XX'} = \rho_{XX'AEF}$. Therefore, applying the above lemma,

$$\begin{aligned} \mathrm{D}_\alpha(\rho_{XX'AE} \,\|\, \mathbb{I}_{XA} \otimes \sigma_{X'E}) &\geq \mathrm{D}_\alpha(\rho_{XX'AE} \,\|\, \Pi_{XX'}(\mathbb{I}_{XA} \otimes \sigma_{X'E})\Pi_{XX'}) \\ &= \mathrm{D}_\alpha(\rho_{XX'AE} \,\|\, \mathbb{I}_A \otimes \Pi_{XX'}(\mathbb{I}_X \otimes \sigma_{X'E})\Pi_{XX'}) \end{aligned}$$

for any $\sigma_{X'E} \in \mathcal{D}(\mathcal{H}_{X'} \otimes \mathcal{H}_E)$. Expoiting that

$$\mathrm{tr}\big(\Pi_{XX'}(\mathbb{I}_X \otimes \sigma_{X'E})\Pi_{XX'}\big) = \sum_x \mathrm{tr}\big(|x\rangle\langle x| \otimes (|x\rangle\langle x| \otimes \mathbb{I}_E)\sigma_{X'E}(|x\rangle\langle x| \otimes \mathbb{I}_E)\big) = \mathrm{tr}(\sigma_{X'E}) = 1\,,$$

we obtain that $\mathrm{H}_\alpha(XA|X'E) \leq \mathrm{H}_\alpha(A|XX'E)$. Finally, to conclude, we consider the state $\rho_{XX'AEF}$ obtained by purifying each ρ^x_{AE}. We observe that $\rho_{XX'AEF}$ is pure, and the above reasoning applies equally to $\rho_{XX'AF}$. Using the duality property (Theorem 8.17), we then get that

$$\mathrm{H}_\alpha(XA|E) = -\mathrm{H}_\beta(XA|X'F) \geq -\mathrm{H}_\beta(A|XX'F) = \mathrm{H}_\alpha(A|E)\,,$$

as is claimed. □