Final Project on Chain Rules of Smooth Min- and Max- Entropies

Zhong Sheng Hu
zs2hu@uwaterloo.ca

1 Goal

During the study of [10](and its preliminaries [3], [8]), despite the already complex expressions of the four pairs of chain rules, the derivations are not all obvious, and some steps are clearer in other references, which make them more difficult to follow. There is also a theorem which uses some SDP technique, which I think is worth made explicit. Hence the goal of this project is to expand selected derivations and add more explanations, such that it can be more approachable for the hypothetically intended audience sharing the same level as the author's , i.e. beginners of quantum information theory.

This writeup first lists out some necessary definitions, and some properties which are considered immediate. Then we proceed on further discussion on one chain rule.

2 Definitions

Following is a brief summary of smooth min- max- entropies and their properties.

2.1 Smooth Min- and Max- Entropies

Definition 1. $D_{\leq}(\mathcal{X}) = \{\rho \in Pos(\mathcal{X}) : Tr(\rho) \leq 1\}$

Let's call its member ***substate****.*

Definition 2. *min- max- entropies of $\rho_{XY} \in D_{\leq}(\mathcal{X} \otimes \mathcal{Y})$ relative to $\sigma_Y \in D_{\leq}(\mathcal{Y})$ is defined to be*

$$H_{\min}(\rho_{XY} \mid \sigma_Y) = \sup\{\lambda : \rho_{XY} \leq 2^{-\lambda} \mathbb{1}_X \otimes \sigma_Y\}$$
$$H_{\max}(\rho_{XY} \mid \sigma_Y) = 2 \log F(\rho_{XY}, \mathbb{1}_X \otimes \sigma_Y)$$

where F is fidelity.

The subscripts of substates denote the spaces they live in. E.g. for $\rho \in D_{\leq}(\mathcal{X} \otimes \mathcal{Y})$, $\rho_X = Tr_{\mathcal{Y}}(\rho)$, and ρ_{XYZ} is its purification.

Definition 3. *min- max- entropies of ρ_{XY} is the min- max- entropies relative to some optimal choice of σ_Y*

$$H_{\min}(X \mid Y)_\rho = \max_{\sigma_Y} H_{\min}(\rho_{XY} \mid \sigma_Y)$$
$$H_{\max}(X \mid Y)_\rho = \max_{\sigma_Y} H_{\max}(\rho_{XY} \mid \sigma_Y)$$

it's feasible because we assume $\mathcal{X}$ and $\mathcal{Y}$ are finite spaces.

Alternatively[3],

$$H_{\max}(X \mid Y)_\rho = -H_{\min}(X \mid Z)_\rho$$

where ρ_{XY} is purified in $\mathcal{X} \otimes \mathcal{Y} \otimes \mathcal{Z}$, which indicates the duality relation between these two.

Definition 4. *The purified distance between $\rho, \sigma \in D_{\leq}(\mathcal{X})$ is based on fidelity*

$$P(\rho, \sigma) = \sqrt{1 - \sqrt{(1 - Tr\rho)(1 - Tr\sigma)} - F(\rho, \sigma)}$$

Theorem 5. *Purified distance is a metric on $D_{\leq}(\mathcal{X})$[8, Lemma 5].*

This property can be seen directly from fidelity's properties. Identity can be seen from trace norm's property. Commutativity can be seen from commutativity of fidelity and multiplication. The triangular inequality can be seen from assignment 2.

Definition 6. *For $\epsilon \geq 0$, ϵ-ball around $\rho \in D_{\leq}(\mathcal{X}), Tr(\rho) > \epsilon$, is*

$$\mathcal{B}^{\epsilon}(\rho) = \{\sigma \in D_{\leq}(\mathcal{X}) : P(\rho, \sigma) \leq \epsilon\}$$

Definition 7. *Smooth min- max- entropies are optimization of min- max- entropies within the ϵ-ball, respectively,*

$$\rho_{XY} \in D_{\leq}(\mathcal{X} \otimes \mathcal{Y})$$

$$H^{\epsilon}_{\min}(X \mid Y)_{\rho} = \max_{\tilde{\rho} \in \mathcal{B}^{\epsilon}(\mathcal{X})} H_{\min}(X \mid Y)_{\tilde{\rho}}$$

$$H^{\epsilon}_{\max}(X \mid Y)_{\rho} = \min_{\tilde{\rho} \in \mathcal{B}^{\epsilon}(\mathcal{X})} H_{\max}(X \mid Y)_{\tilde{\rho}}$$

It's feasible because ϵ-ball is compact[8, Definition 10].

Or in the form of duality[8],

$$H^{\epsilon}_{\max}(X \mid Y)_{\rho} = -H^{\epsilon}_{\min}(X \mid Z)_{\rho}$$

where ρ_{XYZ} is purification of ρ_{XY} in $\mathcal{X} \otimes \mathcal{Y} \otimes \mathcal{Z}$.

2.2 Some properties

Following are some basic and relevant properties which the subsequent proof is based on.

Lemma 8. *[8, Lemma 7]*

$$\rho, \sigma \in D_{\leq}(\mathcal{X}), \Phi \in CP(\mathcal{X}, \mathcal{Y}), \Phi \text{ is trace non-decreasing}, P(\rho, \sigma) \geq P(\Phi(\rho), \Phi(\sigma)).$$

Since purified distance is fidelity based, following can be drawn from fidelity's properties immediately[Theorem 3.28, Theorem 3.22 in the book],

Lemma 9. *[8, Lemma 8, Corollary 9]*

$$\rho, \sigma \in D_{\leq}(\mathcal{X}), \exists \rho', \sigma' \in D_{\leq}(\mathcal{X} \otimes \mathcal{Y}), s.t.\ Tr_{\mathcal{Y}}(\rho') = \rho, Tr_{\mathcal{Y}}(\sigma') = \sigma,$$

$$P(\rho, \sigma) = P(\rho', \sigma')$$

A special case will be when ρ', σ' are pure.

This lemma then implies the existence of extension of any substates staying in a larger ϵ-ball with the same purified distance.

3 Main Discussion

The main discussion follows. Considering the main steps of all four pairs of chain rules share the same spirit, it might just suffice to discuss the most difficult one, in order to avoid redundancy, together with other concepts accompanied to it.

Following theorem will be the main focus[10, Theorem 15],

Theorem 10. *[10, Theorem 15] Let $\epsilon > 0, \epsilon', \epsilon \geq 0$, and $\rho_{XYZ} \in D_{\leq}(\mathcal{X} \otimes \mathcal{Y} \otimes \mathcal{Z})$, then*

$$H^{\epsilon'}_{\min}(XY \mid Z) \leq H^{\epsilon''}_{\max}(X \mid YZ) + H^{2\epsilon+\epsilon'+\epsilon''}_{\min}(Y \mid Z)^{1} + 3f(\epsilon)$$

where

$$f(\epsilon) = \log \frac{1}{1 - \sqrt{1 - \epsilon^2}}$$

[1] In the paper, the purified distance of min-entropy on the right hand side is $2\epsilon + \epsilon' + 2\epsilon''$, which has an extra factor of 2 before ϵ''. I suspect it's not needed and the discussion will be in the proof below.

3.1 A Theorem Using SDP

Following theorem uses a form of SDP which is not seen from the lectures, is quite puzzling from the first glance. Hence it might be worth verbosifying it.

The theorem is based on the following lemma.

Lemma 11. *[10, Lemma 7]* $\rho_{XY} \in D_{\leq}(\mathcal{X} \otimes \mathcal{Y}), \sigma_Y \in D_{\leq}(\mathcal{Y})$, ρ_{XYZ} *is purification of* ρ_{XY}*, then* $H_{\max}(\rho_{XY} \mid \sigma_Y) = \log \gamma$*, where* γ *is the optimal value of following SDP,*

primal problem:

$$\begin{aligned} \max\ & \langle A_{XYZ}, \rho_{XYZ} \rangle \\ s.t.\ & Tr_Z(A_{XYZ}) \leq \mathbb{1}_X \otimes \sigma_Y \\ & A_{XYZ} \geq 0 \end{aligned}$$

dual problem:

$$\begin{aligned} \min\ & \langle \mathbb{1}_X \otimes \sigma_Y, B_{XY} \rangle \\ s.t.\ & B_{XY} \otimes \mathbb{1}_Z \geq \rho_{XYZ} \\ & B_{XY} \geq 0 \end{aligned}$$

The duality gap for this SDP is closed by using Slater's theorem, and the primal optimal value and dual optimal value equals. Therefore γ *is well defined.*

Following theorem optimizes above over all σ_Y.

Theorem 12. *[10, Lemma 8]* $\rho_{XY} \in D_{\leq}(\mathcal{X} \otimes \mathcal{Y}), \sigma_Y \in D_{\leq}(\mathcal{Y})$, ρ_{XYZ} *is purification of* ρ_{XY}*, then*

$$H_{\max}(X \mid Y)_\rho = \log \gamma'$$

where γ' *is the optimal value of following*

$$\begin{aligned} \min\ & \|B_Y\|_\infty \\ s.t.\ & B_{XY} \otimes \mathbb{1}_Z \geq \rho_{XYZ} \\ & B_{XY} \geq 0 \end{aligned}$$

Proof. We have known by definition

$$H_{\max}(X \mid Y)_\rho = \max_{\sigma_Y} H_{\max}(\rho_{XY} \mid \sigma_Y)$$

Considering the primal problem of the SDP above, if we further optimize that over σ_Y, we indeed generate following SDP,

$$\begin{aligned} \max\ & \langle A_{XYZ}, \rho_{XYZ} \rangle \\ \text{s.t.}\ & Tr_Z(A_{XYZ}) \leq \mathbb{1}_X \otimes \sigma_Y \\ & Tr(\sigma_Y) \leq 1 \\ & A_{XYZ} \geq 0 \\ & \sigma_Y \geq 0 \end{aligned}$$

Therefore our goal is to show the steps to transform this primal problem to the matching dual problem shown in the paper.

Consider transforming the constraints into following, together with slack variables $0 \leq K_1 \in \mathbb{C}$, $K_2 \in Pos(\mathcal{X} \otimes \mathcal{Y})$,

$$\begin{aligned} Tr(\sigma_Y) + K_1 &= 1 \\ Tr_Z(A_{XYZ}) - \mathbb{1}_X \otimes \sigma_Y + K_2 &= 0 \end{aligned}$$

Let's consider any $A' \in Pos(\mathcal{Y} \oplus \mathbb{C} \oplus (\mathcal{X} \otimes \mathcal{Y} \otimes \mathcal{Z}) \oplus (\mathcal{X} \otimes \mathcal{Y}))$, then we can definitely decompose the A' into following

$$A' = \begin{pmatrix} \sigma_Y & \cdot & \cdot & \cdot \\ \cdot & K_1 & \cdot & \cdot \\ \cdot & \cdot & A_{XYZ} & \cdot \\ \cdot & \cdot & \cdot & K_2 \end{pmatrix}$$

where · denotes "don't care". We can then try to construct the Hermitian preserving map, as following

$$\begin{aligned}
\Phi_1(A') &= Tr(\sigma_Y) + K_1 \\
&\quad \Phi_1 \in T(\mathcal{Y} \oplus \mathbb{C} \oplus (\mathcal{X} \otimes \mathcal{Y} \otimes \mathcal{Z}) \oplus (\mathcal{X} \otimes \mathcal{Y}), \mathbb{C}) \\
\Phi_2(A') &= Tr_Z(A_{XYZ}) - \mathbb{1}_X \otimes \sigma_Y + K_2 \\
&\quad \Phi_2 \in T(\mathcal{Y} \oplus \mathbb{C} \oplus (\mathcal{X} \otimes \mathcal{Y} \otimes \mathcal{Z}) \oplus (\mathcal{X} \otimes \mathcal{Y}), \mathcal{X} \otimes \mathcal{Y}) \\
&\quad \Phi = \begin{pmatrix} \Phi_1 & 0 \\ 0 & \Phi_2 \end{pmatrix} \in T(\mathcal{Y} \oplus \mathbb{C} \oplus (\mathcal{X} \otimes \mathcal{Y} \otimes \mathcal{Z}) \oplus (\mathcal{X} \otimes \mathcal{Y}), \mathbb{C} \oplus (\mathcal{X} \otimes \mathcal{Y}))
\end{aligned}$$

All components from Φ_1, Φ_2 are Hermitian preserving, and they only deal with diagonal elements, so Φ is also Hermitian preserving. Hence the constraint becomes

$$\Phi(A') = \begin{pmatrix} Tr(\sigma_Y) + K_1 & \mathbf{0} \\ \mathbf{0} & Tr_Z(A_{XYZ}) - \mathbb{1}_X \otimes \sigma_Y + K_2 \end{pmatrix} = \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} = C \in Pos(\mathbb{C} \oplus (\mathcal{X} \otimes \mathcal{Y}))$$

Under this setting, it wouldn't be difficult already to come up with the target function,

$$\langle A_{XYZ}, \rho_{XYZ} \rangle = \langle A', D \rangle$$

where

$$D = \begin{pmatrix} \mathbf{0} & & & \mathbf{0} \\ & 0 & & \\ & & \rho_{XYZ} & \\ \mathbf{0} & & & \mathbf{0} \end{pmatrix} \in Pos(\mathcal{Y} \oplus \mathbb{C} \oplus (\mathcal{X} \otimes \mathcal{Y} \otimes \mathcal{Z}) \oplus (\mathcal{X} \otimes \mathcal{Y}))$$

Then now we can express the SDP as following

primal problem

$$\begin{aligned}
\max\ & \langle A', D \rangle \\
\text{s.t. } & \Phi(A') = C \\
& A' \geq 0
\end{aligned}$$

dual problem

$$\begin{aligned}
\min\ & \langle B', C \rangle \\
\text{s.t. } & \Phi^*(B') \geq D \\
& B' \in Herm(\mathbb{C} \oplus (\mathcal{X} \otimes \mathcal{Y}))
\end{aligned}$$

Then we will need to compute Φ^*. Consider any $B' = \begin{pmatrix} \lambda & \cdot \\ \cdot & B_{XY} \end{pmatrix} \in Herm(\mathbb{C} \oplus (\mathcal{X} \otimes \mathcal{Y}))$, which is decomposed in similar form to above, we have

$$\begin{aligned}
\langle \Phi(A'), B' \rangle &= \langle \begin{pmatrix} \Phi_1(A') & 0 \\ 0 & \Phi_2(A') \end{pmatrix}, \begin{pmatrix} \lambda & \cdot \\ \cdot & B_{XY} \end{pmatrix} \rangle \\
&= \langle \Phi_1(A'), \lambda \rangle + \langle \Phi_2(A'), B_{XY} \rangle \\
&= \langle A', \Phi_1^*(\lambda) \rangle + \langle A', \Phi_2^*(B_{XY}) \rangle \\
&= \langle A', \Phi_1^*(\lambda) + \Phi_2^*(B_{XY}) \rangle \\
&= \langle A', \Phi^*(B') \rangle
\end{aligned}$$

Compute Φ_1, Φ_2 respectively,

$$\Phi_1^*(\lambda) = \begin{pmatrix} \lambda \mathbb{1}_Y & & & \mathbf{0} \\ & \lambda & & \\ & & \mathbf{0} & \\ \mathbf{0} & & & \mathbf{0} \end{pmatrix}$$

$$\Phi_2^*(B_{XY}) = \begin{pmatrix} -Tr_{\mathcal{X}}(B_{XY}) & & & \mathbf{0} \\ & 0 & & \\ & & B_{XY} \otimes \mathbb{1}_Z & \\ \mathbf{0} & & & B_{XY} \end{pmatrix}$$

Therefore the constraint is

$$\begin{aligned} \Phi^*(B') &= \Phi^*(\begin{pmatrix} \lambda & \cdot \\ \cdot & B_{XY} \end{pmatrix}) \\ &= \Phi_1^*(\lambda) + \Phi_2^*(B_{XY}) \\ &= \begin{pmatrix} \lambda \mathbb{1}_Y - Tr_{\mathcal{X}}(B_{XY}) & & & \mathbf{0} \\ & \lambda & & \\ & & B_{XY} \otimes \mathbb{1}_Z & \\ \mathbf{0} & & & B_{XY} \end{pmatrix} \\ &\geq D \end{aligned}$$

Expanding D and knowing that the only non-zero entry of C is $C_{11} = 1$, we obtained what the paper described,

$$\begin{aligned} \min\ & \langle B', C \rangle = \lambda \\ \text{s.t. } & \lambda \mathbb{1}_Y - Tr_{\mathcal{X}}(B_{XY}) \geq 0 \\ & \lambda \geq 0 \\ & B_{XY} \otimes \mathbb{1}_Z \geq \rho_{XYZ} \\ & B_{XY} \geq 0 \end{aligned}$$

The reason why $\lambda = \|B_Y\|_\infty$ can be obtained from the constraint $\lambda \mathbb{1}_Y - Tr_{\mathcal{X}}(B_{XY}) \geq 0 \Rightarrow \lambda \mathbb{1}_Y \geq B_Y$, and trying to minimize λ. Hence the same final dual SDP as the paper is generated,

$$\begin{aligned} \min\ & \|B_Y\|_\infty \\ \text{s.t. } & B_{XY} \otimes \mathbb{1}_Z \geq \rho_{XYZ} \\ & B_{XY} \geq 0 \end{aligned}$$

Since we know the solution of primal problem is $2^{H_{\max}(X \mid Y)_\rho}$, we will just need to show there is no duality gap and the optimal values are the same. By Slater's theorem, we can definitely find an instance of $A' > 0$, such that $\Phi(A') = C$. Consider following,

For any $G \in Pd(\mathcal{Y}), Tr(G) < 1$

$$A' = \begin{pmatrix} G & & & \mathbf{0} \\ & 1 - Tr(G) & & \\ & & \frac{1}{2dim(\mathcal{Z})} \mathbb{1}_X \otimes G \otimes \mathbb{1}_Z & \\ \mathbf{0} & & & \frac{1}{2} \mathbb{1}_X \otimes G \end{pmatrix} > 0$$

Easily verified that $\Phi(A') = C$. The dual feasible set is definitely not empty, as we can, for example, let $B_{XY} = \|\rho_{XYZ}\|_\infty \mathbb{1}_{XY} \geq \rho_{XYZ}$. Therefore the duality gap vanishes, and also the dual optimal value is feasible. □

3.2 S-entropy

Following is a discussion regarding S-entropy, which is a tool concept to assist proof steps of the chain rules.

Definition 13. *S-entropy of $\rho_{XY} \in D_{\leq}(\mathcal{X} \otimes \mathcal{Y})$ relative to $\sigma_Y \in D_{\leq}(\mathcal{Y})$ is defined as*

$$S^{\epsilon}(\rho_{XY} \mid \sigma_Y) = \inf\{\lambda : Tr(\Pi^{\lambda}_{XY}\rho_{XY}) \leq \epsilon\}$$

where

Π^{λ}_{XY} is the projection matrix onto $(2^{\lambda}\rho_{XY} - \mathbb{1}_X \otimes \sigma_Y)$'s eigenspace with negative eigenvalues.

This definition is quite confusing.

Since we know ρ_{XY}, σ_Y are positive, we know $(2^{\lambda}\rho_{XY} - \mathbb{1}_X \otimes \sigma_Y)$ is Hermitian. Then we can look at its Jordan-Hahn decomposition,

$$2^{\lambda}\rho_{XY} - \mathbb{1}_X \otimes \sigma_Y = A - B$$

where

$$A, B \in Pos(\mathcal{X} \otimes \mathcal{Y}), AB = 0$$

Then we know $\Pi^{\lambda}_{XY} = B^0$ can be found is such a way.

To see how it tries to minimize λ s.t. $2^{\lambda}\rho_{XY} \geq \mathbb{1}_X \otimes \sigma_Y$, we have following

$$2^{\lambda}\rho_{XY} \geq \mathbb{1}_X \otimes \sigma_Y$$
$$2^{\lambda}\rho_{XY} - \mathbb{1}_X \otimes \sigma_Y \geq 0$$
$$A - B \geq 0$$

Then B needs to vanish. Therefore

$$Tr(B^0(\rho_{XY} - 2^{-\lambda}\mathbb{1}_X \otimes \sigma_Y)) \leq Tr(B^0\rho_{XY}) \leq \epsilon$$

which does seem to agree with the intuition described in the paper. [2]

Following lemma is also proved in the paper,

Lemma 14. *[10, Lemma 12] $\epsilon > 0, \rho_{XY} \in D_{\leq}(\mathcal{X} \otimes \mathcal{Y}), \sigma_Y \in D_{\leq}(\mathcal{Y})$,*

$$S^{\epsilon}(\rho_{XY} \mid \sigma_Y) \leq H_{\max}(\rho_{XY} \mid \sigma_Y) - \log \epsilon^2$$

3.3 Other theorems

Following theorems are used during the main proof.

Theorem 15. *[10, Lemma 20] $\rho \in D_{\leq}(\mathcal{X}), \Pi$ is a projection matrix in $\mathcal{X}$, then*

$$P(\Pi\rho\Pi, \rho) \leq \sqrt{2Tr(\Pi^{\perp}\rho) - (Tr(\Pi^{\perp}\rho))^2}$$

where

$$\Pi^{\perp} = \mathbb{1} - \Pi$$

Theorem 16. *[10, Lemma 9] $\epsilon > 0, \rho \in D_{\leq}(\mathcal{X} \otimes \mathcal{Y}), \rho' \in \mathcal{B}^{\epsilon'}(\rho), \exists \tilde{\rho} \in \mathcal{B}^{\epsilon+\epsilon'}(\rho')$ such that*

$$H_{\max}(X \mid Y)_{\tilde{\rho}} \leq H_{\max}(\rho_{XY} \mid \rho'_Y) - log(1 - \sqrt{1 - \epsilon^2})$$

[2] From [2, Proposition 2], it turns out that S-entropy has a very close form to spectral inf-divergence rate, and from [2, Theorem 3], this similarity does not seem a coincident. But I can't get something concrete out of it.

3.4 Main Proof

Following we proceed to the main proof of theorem 10.

Proof. There is a general proof idea shared among many proofs of theorems in this paper: relate several substates by projection matricies, and use triangular inequality and theorem 15 to give a bound on purified distance to some other intermediate state. As long as the distance from ρ is known, an optimization within ϵ-ball can bring ρ back to "centre".

For inequalities that involves both min- max- entropies, like this one, S-entropy is used as a guide to create a proper projection matrix, which helps generating bound between min- and max- entropies by lemma 14. One thing worth noticing, is by giving a bound on $Tr(\Pi^{\lambda}_{XY}\rho_{XY}) \leq \epsilon$, S-entropy subsequently enforces $P(\Pi^{\lambda\perp}_{XY}\rho_{XY}\Pi^{\lambda\perp}_{XY}, \rho_{XY}) \leq \sqrt{2Tr(\Pi^{\lambda}_{XY}\rho_{XY}) - (Tr(\Pi^{\lambda}_{XY}\rho_{XY}))^2} \leq \sqrt{2\epsilon - \epsilon^2}$ via theorem 15. That implies a form of potential candidates of auxiliary matrices to work with.

Then theorem 16 is used to switch to another substate with known distance, which will subsequently be used to bring ρ back into scope.

According to the duality of smooth min- max- entropies, for a substate $\rho_{XYZ} \in D_{\leq}(\mathcal{X} \otimes \mathcal{Y} \otimes \mathcal{Z})$ and its purification $u \in \mathcal{X} \otimes \mathcal{Y} \otimes \mathcal{Z} \otimes \mathcal{W}$, $\rho_{XYZW} = uu^*$, the original problem becomes

$$H^{\epsilon''}_{\max}(XY \mid W) \geq H^{\epsilon''}_{\min}(X \mid W) + H^{2\epsilon+\epsilon'+\epsilon''}_{\min}(Y \mid XW) - 3f(\epsilon)$$

Then we let $\rho'_{XYW} \in \mathcal{B}^{\epsilon'}(\rho_{XYW}), \rho''_{XW} \in \mathcal{B}^{\epsilon''}(\rho_{XW})$, such that following optimizations are achieved,

$$\begin{aligned} H^{\epsilon'}_{\max}(XY \mid W)_{\rho} &= H_{\max}(XY \mid W)_{\rho'} \\ H^{\epsilon''}_{\min}(X \mid W)_{\rho} &= H_{\min}(X \mid W)_{\rho''} \end{aligned}$$

Then further let an optimal choice $\sigma_W \in D_{\leq}(\mathcal{W})$, such that

$$H_{\min}(X \mid W)_{\rho''} = H_{\min}(\rho''_{XW} \mid \sigma_W)$$

Following will use S-entropy to draw some relation involving ρ' and ρ''. By definition of $H_{\min}$ [3]

$$\rho''_{XW} \leq 2^{-H_{\min}(X \mid W)_{\rho''}}(\mathbb{1}_X \otimes \sigma_W) \tag{1}$$
$$\Rightarrow 2^{H_{\min}(X \mid W)_{\rho''}}\rho''_{XW} \otimes \mathbb{1}_Y \leq (\mathbb{1}_{XY} \otimes \sigma_W) \tag{2}$$

Then consider following S-entropy $S^{\tilde{\epsilon}}(\rho'_{XYW} \mid \sigma_W), \tilde{\epsilon} > 0$. Since it's an infimum, then we know by definition, there are choices of $\delta > 0$ that are infinitely close to 0, for $\lambda = S^{\tilde{\epsilon}}(\rho'_{XYW} \mid \sigma_W) + \delta$, so that $Tr(\Pi^{\lambda}_{XYW}\rho'_{XYW}) \leq \tilde{\epsilon}$ holds.

Since Π^{λ}_{XYW} projects to the negative eigenspace of $(2^{\lambda}\rho'_{XYW} - \mathbb{1}_{XY} \otimes \sigma_W)$, we can let $\Pi^{\lambda\perp}_{XYW} = \mathbb{1}_{XYW} - \Pi^{\lambda}_{XYW}$, which projects the matrix to its positive eigenspace instead. We then are sure following trivially holds,

$$\Pi^{\lambda\perp}_{XYW}(2^{\lambda}\rho'_{XYW} - \mathbb{1}_{XY} \otimes \sigma_W)\Pi^{\lambda\perp}_{XYW} \geq 0 \tag{3}$$
$$\Pi^{\lambda\perp}_{XYW}(\mathbb{1}_{XY} \otimes \sigma_W)\Pi^{\lambda\perp}_{XYW} \leq \Pi^{\lambda\perp}_{XYW}2^{\lambda}\rho'_{XYW}\Pi^{\lambda\perp}_{XYW} \tag{4}$$

Now we notice $(\mathbb{1}_{XY} \otimes \sigma_W)$ gives some lower bound and upper bound. Form this one, we can more or less make sense why S-entropy is interested in an inequality in a different direction from min-entropy requires. Conjugating inequality 2 by $\Pi^{\lambda\perp}_{XYW}$, then by transitivity of $\leq$ with 4, we have

[3] The order of tensoring is not right but it shouldn't matter in this case as the matrices with different tensoring orders are isomorphic to each other.

$$2^{H_{\min}(X\mid W)_{\rho''}}\Pi^{\lambda\perp}_{XYW}(\rho''_{XW}\otimes \mathbb{1}_Y)\Pi^{\lambda\perp}_{XYW} \le 2^{\lambda}\Pi^{\lambda\perp}_{XYW}\rho'_{XYW}\Pi^{\lambda\perp}_{XYW}$$

Moving the factor to the right, we have

$$\Pi^{\lambda\perp}_{XYW}(\rho''_{XW}\otimes \mathbb{1}_Y)\Pi^{\lambda\perp}_{XYW} \le 2^{\lambda-H_{\min}(X\mid W)_{\rho''}}\Pi^{\lambda\perp}_{XYW}\rho'_{XYW}\Pi^{\lambda\perp}_{XYW} \tag{5}$$

Then we are looking to use lemma 14 to remove S-entropy from the expression. Consider following conditional max-entropy next

$$\begin{aligned}
2^{H_{\max}(\Pi^{\lambda\perp}_{XYW}\rho'_{XYW}\Pi^{\lambda\perp}_{XYW}\mid \rho''_{XW})} = & \qquad\qquad \textit{lemma } 11\\
& \min\ \langle \rho''_{XW}\otimes \mathbb{1}_Y, B_{XYW}\rangle\\
& \text{s.t. } B_{XYW}\otimes \mathbb{1}_V \ge (\Pi^{\lambda\perp}_{XYW}\otimes \mathbb{1}_V)\tau_{XYWV}(\Pi^{\lambda\perp}_{XYW}\otimes \mathbb{1}_V)\\
& \qquad B_{XYW}\ge 0\\
& \text{where}\\
& \tau_{XYWV} \text{ purifies } \rho'_{XYW} \text{ for some space } \mathcal{V}^{4}
\end{aligned}$$

Notice that $(\Pi^{\lambda\perp}_{XYW}\otimes \mathbb{1}_V)\tau_{XYWV}(\Pi^{\lambda\perp}_{XYW}\otimes \mathbb{1}_V)$ purifies $\Pi^{\lambda\perp}_{XYW}\rho'_{XYW}\Pi^{\lambda\perp}_{XYW}$ as it's a rank one matrix. Therefore following holds

$$(\Pi^{\lambda\perp}_{XYW}\otimes \mathbb{1}_V)\tau_{XYWV}(\Pi^{\lambda\perp}_{XYW}\otimes \mathbb{1}_V) \le \Pi^{\lambda\perp}_{XYW}\otimes \mathbb{1}_V$$

So $\Pi^{\lambda\perp}_{XYW}$ is a candidate of the dual SDP. Therefore we know

$$\begin{aligned}
2^{H_{\max}(\Pi^{\lambda\perp}_{XYW}\rho'_{XYW}\Pi^{\lambda\perp}_{XYW}\mid \rho''_{XW})} &\le \langle \rho''_{XW}\otimes \mathbb{1}_Y, \Pi^{\lambda\perp}_{XYW}\rangle && \Pi^{\lambda\perp}_{XYW} \text{ is a suboptimal choice.}\\
&= Tr(\Pi^{\lambda\perp}_{XYW}(\rho''_{XW}\otimes \mathbb{1}_Y)\Pi^{\lambda\perp}_{XYW})\\
&\le 2^{\lambda-H_{\min}(X\mid W)_{\rho''}}Tr(\Pi^{\lambda\perp}_{XYW}(\rho'_{XYW})\Pi^{\lambda\perp}_{XYW}) && \text{inequality } 5\\
&\le 2^{\lambda-H_{\min}(X\mid W)_{\rho''}}\\
&\text{because the matrix above is a substate}
\end{aligned}$$

Just look at the exponent,

$$\begin{aligned}
H_{\max}(\Pi^{\lambda\perp}_{XYW}\rho'_{XYW}\Pi^{\lambda\perp}_{XYW}\mid \rho''_{XW}) &\le \lambda - H_{\min}(X\mid W)_{\rho''}\\
&= S^{\tilde\epsilon}(\rho'_{XYW}\mid \sigma_W) + \delta - H_{\min}(X\mid W)_{\rho''}\\
&\le H_{\max}(\rho'_{XYW}\mid \sigma_W) - H_{\min}(X\mid W)_{\rho''} + \delta - 2\log\tilde\epsilon && \textit{lemma } 14\\
&\le H_{\max}(XY\mid W)_{\rho'} - H_{\min}(X\mid W)_{\rho''} + \delta - 2\log\tilde\epsilon\\
&\text{due to optimization over } \sigma_W.\\
&= H^{\epsilon'}_{\max}(XY\mid W)_{\rho} - H^{\epsilon''}_{\min}(X\mid W)_{\rho} + \delta - 2\log\tilde\epsilon && (*)
\end{aligned}$$

The form of right hand side suggests that the remaining work is to relax the $H_{\max}$ on left hand side further to find a max-entropy related to ρ. According to lemma 9, we can then find extensions of ρ'_{XYW}, ρ''_{XW}, respectively $\rho'_{XYZW}, \rho''_{XYZW}$, such that $P(\rho'_{XYZW}, \rho_{XYZW}) = P(\rho'_{XYW}, \rho_{XYW}) \le \epsilon'$, $P(\rho''_{XYZW}, \rho_{XYZW}) =$

[4] The paper picks the purification in $\mathcal{X}\otimes\mathcal{Y}\otimes\mathcal{Z}\otimes\mathcal{W}$, which I find strange, as $\mathcal{W}$ is picked to purify ρ_{XYZ} so it can be very large, and we don't know the rank of ρ'_{XYW}, so it's not straightforward to me how ρ'_{XYW} can be purified in $\mathcal{Z}$. However, it does not seem to matter as the purification is not used anywhere in the concrete argument.

$P(\rho''_{XW}, \rho_{XW}) \le \epsilon''$, and then we have following,

$$
\begin{aligned}
&P(\Pi^{\lambda\perp}_{XYW}\rho'_{XYW}\Pi^{\lambda\perp}_{XYW}, \rho_{XYW}) \\
\le &P((\Pi^{\lambda\perp}_{XYW}\otimes \mathbb{1}_Z)(\rho'_{XYZW})(\Pi^{\lambda\perp}_{XYW}\otimes \mathbb{1}_Z), \rho_{XYZW}) \\
&\textit{lemma } 8, \text{and } Tr_{\mathcal{Z}} \text{ is trace-preserving.} \\
\le &P((\Pi^{\lambda\perp}_{XYW}\otimes \mathbb{1}_Z)(\rho'_{XYZW})(\Pi^{\lambda\perp}_{XYW}\otimes \mathbb{1}_Z), \rho'_{XYZW}) + P(\rho'_{XYZW}, \rho_{XYZW}) + P(\rho_{XYZW}, \rho''_{XYZW}) \\
&\text{triangular inequality because purified distance is a metric.} \\
\le &\sqrt{2\widetilde{\epsilon} - \widetilde{\epsilon}^2} + \epsilon' + \epsilon''
\end{aligned}
$$

The last step is due to theorem 15, and the assumption of Π^{λ}_{XYW} made by the S-entropy, then we have $Tr(\Pi^{\lambda}_{XYW}\rho'_{XYW}) \le \widetilde{\epsilon}$.

$$
\begin{aligned}
&P((\Pi^{\lambda\perp}_{XYW}\otimes \mathbb{1}_Z)\rho'_{XYZW}(\Pi^{\lambda\perp}_{XYW}\otimes \mathbb{1}_Z), \rho'_{XYZW}) \\
\le &\sqrt{2Tr(\Pi^{\lambda}_{XYW}\rho'_{XYW}) - (Tr(\Pi^{\lambda}_{XYW}\rho'_{XYW}))^2} \\
&\textit{theorem } 15 \text{ and } Tr_{\mathcal{Z}} \text{ is trace-preserving.} \\
\le &\sqrt{2\widetilde{\epsilon} - \widetilde{\epsilon}^2} \\
&\text{because } x \mapsto \sqrt{2x - x^2} = \sqrt{1-(x-1)^2} \text{ monotonically increases when } x \le 1
\end{aligned}
$$

Then by theorem 16, we know there exists $\widetilde{\rho}_{XYW} \in \mathcal{B}^{\sqrt{2\widetilde{\epsilon}-\widetilde{\epsilon}^2}+\epsilon'+\epsilon''+\epsilon}(\rho''_{XYW})$, such that [5]

$$
H_{\max}(Y \mid XW)_{\widetilde{\rho}} \le H_{\max}(\Pi^{\lambda\perp}_{XYW}\rho'_{XYW}\Pi^{\lambda\perp}_{XYW} \mid \rho''_{XW}) + f(\epsilon)
$$

Substitute that in inequality *, we have

$$
\begin{aligned}
H_{\max}^{\sqrt{2\widetilde{\epsilon}-\widetilde{\epsilon}^2}+\epsilon'+\epsilon''+\epsilon}(Y \mid XW)_{\rho} \le\ & H_{\max}(Y \mid XW)_{\widetilde{\rho}} \\
& \text{optmization within the } \epsilon\text{-ball} \\
\le\ & H_{\max}(\Pi^{\lambda\perp}_{XYW}\rho'_{XYW}\Pi^{\lambda\perp}_{XYW} \mid \rho''_{XW}) + f(\epsilon) \\
\le\ & H^{\epsilon'}_{\max}(XY \mid W)_{\rho} - H^{\epsilon''}_{\min}(X \mid W)_{\rho} + \delta - 2\log\widetilde{\epsilon} + f(\epsilon) \quad \text{inequality } *
\end{aligned}
$$

Since $\widetilde{\epsilon}$ is arbitrary, we can let $\widetilde{\epsilon} = 1 - \sqrt{1-\epsilon^2}$; δ can be infinitesimal, so let $\delta \to 0$ to approximate the infimum, we eventually will be able to obtain from above,

$$
\begin{aligned}
H^{\epsilon'+\epsilon''+2\epsilon}_{\max}(Y \mid XW)_{\rho} &\le H^{\epsilon'}_{\max}(XY \mid W)_{\rho} - H^{\epsilon''}_{\min}(X \mid W)_{\rho} + 3f(\epsilon) \\
H^{\epsilon'}_{\max}(XY \mid W)_{\rho} &\ge H^{\epsilon''}_{\min}(X \mid W)_{\rho} + H^{\epsilon'+\epsilon''+2\epsilon}_{\max}(Y \mid XW)_{\rho} - 3f(\epsilon)
\end{aligned}
$$

as claimed at the beginning. □

[5]This step in the paper sets the bound to be $\sqrt{2\widetilde{\epsilon}-\widetilde{\epsilon}^2} + \epsilon' + \mathbf{2\epsilon''} + \epsilon$, which generates one extra ϵ'' in the distance bound. I don't seem to find that necessary, and it's also suggested on a subsequent bound $f(\epsilon)$, which has no ϵ'''s presence.

References

[1] M. Berta. “Single-shot Quantum State Merging”. In: *ArXiv e-prints* (Dec. 2009). arXiv: 0912.4495 [quant-ph].

[2] N. Datta. “Min- and Max-Relative Entropies and a New Entanglement Monotone”. In: *IEEE Transactions on Information Theory* 55.6 (June 2009), pp. 2816–2826. ISSN: 0018-9448. DOI: 10.1109/TIT.2009.2018325.

[3] R. Koenig, R. Renner, and C. Schaffner. “The operational meaning of min- and max-entropy”. In: *ArXiv e-prints* (July 2008). arXiv: 0807.1338 [quant-ph].

[4] M. Müller-Lennert et al. “On quantum Rényi entropies: A new generalization and some properties”. In: *Journal of Mathematical Physics* 54.12 (Dec. 2013), pp. 122203–122203. DOI: 10.1063/1.4838856. arXiv: 1306.3142 [quant-ph].

[5] Nilanjana Natta. “One-Shot Quantum Information Theory I: Entropic Quantities”. University Lecture. URL: https://www.newton.ac.uk/files/seminar/20130904140015001-153776.pdf.

[6] R. Renner. “Security of Quantum Key Distribution”. PhD thesis. PhD Thesis, 2005, 2005.

[7] R. Renner and S. Wolf. “Smooth Renyi entropy and applications”. In: *International Symposium onInformation Theory, 2004. ISIT 2004. Proceedings.* June 2004, pp. 233–. DOI: 10.1109/ISIT.2004.1365269.

[8] M. Tomamichel, R. Colbeck, and R. Renner. “Duality Between Smooth Min- and Max-Entropies”. In: *ArXiv e-prints* (July 2009). arXiv: 0907.5238 [quant-ph].

[9] M. Tomamichel et al. “Leftover Hashing Against Quantum Side Information”. In: *ArXiv e-prints* (Feb. 2010). arXiv: 1002.2436 [quant-ph].

[10] A. Vitanov et al. “Chain Rules for Smooth Min- and Max-Entropies”. In: *ArXiv e-prints* (May 2012). arXiv: 1205.5231 [quant-ph].

[11] John Watrous. “Optimizing over Interactions”. Lecture Note. URL: https://cs.uwaterloo.ca/~watrous/CS867.Winter2017/Notes/07.pdf.

Some Relevant Things That are Irrelevant

During study of [3], [8], [10], I was confused by how max-entropy is defined and subsequently did some investigation. I suppose it might serve an interesting reading and hence write it down here.

Reading [3], defining max-entropy via duality instead of relative Rényi entropy like min-entropy puzzled me. Without noticing the fidelity is already the case, I tried to look up alternative definition, and it was found in [6], [1], [2] as following

$$H'_{\max}(X \mid Y)_\rho = \log \sup_{\sigma_Y} Tr(\rho^0_{XY}(\mathbb{1}_X \otimes \sigma_Y))$$

where ρ^0_{XY} projects onto eigenspace of ρ_{XY}.

I couldn't understand it, since if $H'_{\max} = H_{\max}$ were true, then it indicated a too easy way of computing fidelity. More confusingly, another form of duality was described as[1, Proposition 3.11],

For a pure state $\rho \in D(\mathcal{X} \otimes \mathcal{Y} \otimes \mathcal{Z})$,

$$H_{\min}(\rho_{XZ} \mid \rho_Z) = -H'_{\max}(X \mid Y)_\rho$$

which seemingly implies $H_{\min}(\rho_{XZ} \mid \rho_Z)$ must be optimal, which does not seem the case.

My confusion eventually went away after I saw these slides [5], which put relative Rényi entropy and min- max- entropies side by side. In the slides, $H'_{\max}$ is defined to be H_0, and $H_{\max}$ to be $D_{\frac{1}{2}}$.

Curiously, from papers around '08, '09, it's obvious that the concrete meaning of $H_{\max}$ was still alternating: [1], [2] defined it as H_0, while [3], [8] defined it in the current way, and interestingly they cite each other.