QIT Lent 2018: Example sheet 4 solutions

Eric P. Hanson

May 14 2018

Exercise 1. The other Fuchs and van de Graaf inequality.

1. Let $\{E_m\}_{m\in M}$ be a POVM, where M is a finite set. Given two states ρ_A and σ_A, use the Cauchy-Schwarz inequality to show that

$$F(\rho_A,\sigma_A) \leq \sum_{m\in M} \sqrt{\operatorname{tr}[E_m\rho]\operatorname{tr}[E_m\sigma]} = F(\tilde\rho_M,\tilde\sigma_M) \tag{1}$$

where

$$\tilde\rho_M = \sum_{m\in M} \operatorname{tr}[E_m\rho]|m\rangle\langle m|, \qquad \tilde\sigma_M = \sum_{m\in M}\operatorname{tr}[E_m\sigma]|m\rangle\langle m| \tag{2}$$

are diagonal states encoding probabilities of the measurement outcomes m, in a Hilbert space $\mathcal{H}_M$ of dimension M with orthonormal basis $\{|m\rangle\}_{m\in M}$.

Hint: use the polar decomposition $\sqrt{\rho}\sqrt{\sigma} = |\sqrt{\rho}\sqrt{\sigma}|U$.

2. Let ρ and σ be mixed states and assume ρ is invertible.

Show that if $\hat M := \rho^{-1/2}|\rho^{1/2}\sigma^{1/2}|\rho^{-1/2}$ has spectral decomposition $\hat M = \sum_m \lambda_m |m\rangle\langle m|$, then

$$|m\rangle\langle m|\sqrt{\rho} = \frac{1}{\lambda_m}|m\rangle\langle m|\sigma^{1/2}U^\dagger,$$

where U is unitary and defined by the polar decomposition $\sqrt{\rho}\sqrt{\sigma} = |\sqrt{\rho}\sqrt{\sigma}|U$.

Hint: start by showing that $|m\rangle\langle m|\hat M = \lambda_m|m\rangle\langle m|$.

3. Show that projective measurement $\{E_m\}_{m\in M}$ defined by $E_m = |m\rangle\langle m|$, achieves equality in equation (1).

Conclude that when ρ is invertible,

$$F(\rho,\sigma) = \min_{\{E_m\}_{m\in M}} F(\tilde\rho_M,\tilde\sigma_M) \tag{3}$$

where the minimum is over POVMs $\{E_m\}_{m\in M}$ and $\tilde\rho_M$ and $\tilde\sigma_M$ are defined in (2).

4. Given two quantum states ρ and σ such that ρ is invertible, use equation (3) to show that

$$D(\rho,\sigma) \geq 1 - F(\rho,\sigma).$$

Solution.

1. Since $\sum_m E_m = I$, we have that

$$\begin{aligned} F(\rho,\sigma) &= \operatorname{tr}[|\sqrt{\rho}\sqrt{\sigma}|] = \operatorname{tr}[\sqrt{\rho}\sqrt{\sigma}U^*] \\ &= \sum_m \operatorname{tr}[\sqrt{\rho}E_m\sqrt{\sigma}U^*] = \sum_m \operatorname{tr}[\sqrt{\rho}\sqrt{E_m}\sqrt{E_m}\sqrt{\sigma}U^*] \\ &= \sum_m \langle\sqrt{E_m}\sqrt{\rho}, \sqrt{E_m}\sqrt{\sigma}U^*\rangle_{\mathrm{HS}} \\ &\leq \sum_m \sqrt{\langle\sqrt{E_m}\sqrt{\rho},\sqrt{E_m}\sqrt{\rho}\rangle_{\mathrm{HS}}}\sqrt{\langle\sqrt{E_m}\sqrt{\sigma}U^*,\sqrt{E_m}\sqrt{\sigma}U^*\rangle_{\mathrm{HS}}} \end{aligned}$$

by the Cauchy-Schwarz inequality, where $\langle A,B\rangle_{\mathrm{HS}} = \operatorname{tr}[A^*B]$ is the Hilbert-Schmidt inner product. Thus,

$$\begin{aligned} F(\rho,\sigma) &\leq \sum_m \sqrt{\operatorname{tr}[\sqrt{\rho}E_m\sqrt{\rho}]}\sqrt{\operatorname{tr}[U\sqrt{\sigma}E_m\sqrt{\sigma}U^*]} \\ &= \sum_m \sqrt{\operatorname{tr}[E_m\rho]}\sqrt{\operatorname{tr}[E_m\sigma]} \end{aligned}$$

by the cyclicity of the trace and that $U^*U = I$.

2. We have that

$$\begin{aligned} |m\rangle\langle m|\hat M &= |m\rangle\langle m|\sum_{m'}\lambda_{m'}|m'\rangle\langle m'| \\ &= \sum_{m'}\lambda_{m'}\delta_{m,m'}|m\rangle\langle m'| \\ &= \lambda_m|m\rangle\langle m|. \end{aligned} \tag{4}$$

Subsituting the definition of $\hat M$,

$$\begin{aligned} |m\rangle\langle m|\hat M &= |m\rangle\langle m|\rho^{-1/2}|\rho^{1/2}\sigma^{1/2}|\rho^{-1/2} \\ &= |m\rangle\langle m|\rho^{-1/2}\sqrt{\rho}\sqrt{\sigma}U^\dagger\rho^{-1/2} \\ &= |m\rangle\langle m|\sqrt{\sigma}U^\dagger\rho^{-1/2}. \end{aligned}$$

Thus, by (4), $|m\rangle\langle m|\sqrt{\sigma}U^\dagger\rho^{-1/2} = \lambda_m|m\rangle\langle m|$. Thus,

$$|m\rangle\langle m|\sqrt{\rho} = \lambda_m^{-1}|m\rangle\langle m|\sqrt{\sigma}U^\dagger$$

as desired.

3. We have equality is Cauchy-Schwarz if and only if the two vectors are linearly dependent. Since that is the only inequality we used, we have that $F(\rho,\sigma) = F(\tilde\rho,\tilde\sigma)$ if and only if there exists constants $\alpha_m \in \mathbb{C}$ such that

$$\sqrt{E_m}\sqrt{\rho} = \alpha_m\sqrt{E_m}\sqrt{\sigma}U^*$$

for each m. Choosing $E_m = |m\rangle\langle m|$, this condition becomes

$$|m\rangle\langle m|\sqrt{\rho} = \alpha_m|m\rangle\langle m|\sqrt{\sigma}U^*.$$

By choosing $\alpha_m = \lambda_m^{-1}$ we see that indeed, $F(\rho,\sigma) = F(\tilde\rho,\tilde\sigma)$. Since for any measurement $\{E_m\}$ we have $F(\rho,\sigma) \leq F(\tilde\rho,\tilde\sigma)$, and that equality is achieved for a specific measurement, (3) follows.

4. Let $\{E_m\}$ be a measurement which achieves $F(\rho,\sigma) = F(\tilde{\rho},\tilde{\sigma})$. Moreover, $D(\tilde{\rho},\tilde{\sigma}) \le D(\rho,\sigma)$ by the data-processing inequality, since the ''measure-and-prepare" map $\Lambda : \rho \mapsto \sum_m \mathrm{tr}[E_m\rho]|m\rangle\langle m|$ is CPTP. Thus, it remains to prove that

$$1 - F(\tilde{\rho},\tilde{\sigma}) \le D(\tilde{\rho},\tilde{\sigma}). \tag{5}$$

These states are both diagonal in the basis $\{|m\rangle\}$ and thus we've reduced to the classical case. We define $p_m = \mathrm{tr}[\rho E_m]$ and $q_m = \mathrm{tr}[\sigma E_m]$. Then $F(\tilde{\rho},\tilde{\sigma}) = \sum_m \sqrt{p_m q_m}$, and $\{p_m\}_m$ and $\{q_m\}_m$ are each probability distributions. Then

$$\begin{aligned}\sum_m (\sqrt{p_m} - \sqrt{q_m})^2 &= \sum_m p_m + \sum_m q_m - 2\sum_m \sqrt{p_m q_m} \\ &= 2(1 - F(\tilde{\rho},\tilde{\sigma})).\end{aligned}$$

On the other hand, we can use the inequality $|\sqrt{p_m} - \sqrt{q_m}| \le |\sqrt{p_m} + \sqrt{q}_m|$ to see

$$\begin{aligned}\sum_m (\sqrt{p_m} - \sqrt{q_m})^2 = \sum_m |\sqrt{p_m} - \sqrt{q_m}|^2 &\le \sum_m |\sqrt{p_m} - \sqrt{q_m}||\sqrt{p_m} + \sqrt{q}_m| \\ &= \sum_m |p_m - q_m| = 2D(\tilde{\rho},\tilde{\sigma})\end{aligned}$$

which completes the proof[1].

Exercise 2. Let $|\psi\rangle_{ABE}$ be a pure state of a tripartite system ABE. define the *coherent information* from A to B of ψ to be

$$I_c^{A>B}(\psi) = -S(A|B)_\psi.$$

Here $S(A|B)_\psi$ denotes the conditional entropy of the subsystem A with respect to subsystem B, given that the composite system ABE is in the pure state $|\psi\rangle_{ABE}$. Henceforth we shall omit ψ.

Prove the following identities:

1. $\frac{1}{2}I(A:B) + \frac{1}{2}I(A:E) = S(A)$
2. $\frac{1}{2}I(A:B) - \frac{1}{2}I(A:E) = I_c^{A>B}$

Solution.

1. By definition of the quantum mutual information,

$$\begin{aligned}\frac{1}{2}I(A:B) + \frac{1}{2}I(A:E) &= \frac{1}{2}[S(A) + S(B) - S(AB) + S(A) + S(E) - S(AE)] \\ &= S(A) + \frac{1}{2}[S(B) - S(AB) + S(E) - S(AE)].\end{aligned}$$

Since $|\psi\rangle_{ABE}$ is pure, $S(B) = S(AE)$ and $S(E) = S(AB)$ (since for any bipartition of the systems ABE, the entropies of both reduced density matrices are equal). Thus, the term in brackets vanishes and we obtain $S(A)$ as desired.

[1] The trick of considering $\sum_m(\sqrt{p_m} - \sqrt{q_m})^2$ and the subsequent bounds are from Nielsen and Chuang, p. 415.

2. Likewise,

$$\begin{aligned}\frac{1}{2}I(A:B) - \frac{1}{2}I(A:E) &= \frac{1}{2}[S(A) + S(B) - S(AB) - S(A) - S(E) + S(AE)] \\ &= \frac{1}{2}[S(B) - S(AB) - S(E) + S(AE)] \\ &= \frac{1}{2}[S(B) - S(AB) - S(AB) + S(B)] \\ &= S(B) - S(AB) = -S(A|B) = I_c^{A>B}.\end{aligned}$$

Exercise 3. A bipartite quantum state ρ_{AB} is said to be *separable* if it can be written as a convex combination of product states, i.e., if there exists an ensemble $\{p_i, \sigma_A^{(i)} \otimes \tau_B^{(i)}\}$, with $\sigma_A^{(i)} \in \mathcal{B}(\mathcal{H}_A)$ and $\tau_B^{(i)} \in \mathcal{B}(\mathcal{H}_B)$, such that

$$\rho_{AB} = \sum_i p_i \sigma_A^{(i)} \otimes \tau_B^{(i)}.$$

This allows us to extend the definition of entanglement to mixed states: *a mixed state is entangled if it is not separable.*

Show that if ρ_{AB} is separable then $I_c^{A>B} \le 0$.

What implication does this have on the conditional entropy $S(A|B)$?

Solution.

Since any separable state can be written as a convex combination of product pure states, any separable state ρ_{AB} may be written as $\rho_{AB} = \sum_i p_i \psi_A^{(i)} \otimes \phi_B^{(i)}$ for some pure states $\psi_A^{(i)}$ and $\phi_B^{(i)}$. Next, because the quantum conditional entropy is concave[2], the coherent information is convex. Thus

$$I_c^{A>B}(\rho) \le \sum_i p_i I_c^{A>B}(\psi_A^{(i)} \otimes \phi_B^{(i)}) = \sum_i p_i [S(\phi_B^{(i)}) - S(\psi_A^{(i)} \otimes \phi_B^{(i)})].$$

But each entropy is the entropy of a pure state and thus vanishes. Thus, $I_c^{A>B}(\rho) \le 0$. Thus conditional entropy is therefore non-negative on separable states.

Exercise 4. Use the HSW theorem to find the product state capacity of the depolarizing channel, Λ, defined by

$$\Lambda(\rho) = p\rho + (1-p)\frac{I}{2}.$$

Solution.

Note we are only dealing with qubits (otherwise Λ would not be trace-preserving). First, note that for any unitary U, we have that

$$\Lambda(U\rho U^*) = pU\rho U^* + (1-p)\frac{I}{2}$$

[2] Execise 9 of Example Sheet 3

and therefore $\Lambda(U\rho U^*) = U\Lambda(\rho)U^*$.

The HSW theorem gives that the product state capacity of Λ is given by

$$\chi^*(\Lambda) = \max_{\{p_x,\rho_x\}} S(\Lambda(\sum_x p_x\rho_x)) - \sum_x p_x S(\Lambda(\rho_x)).$$

We can reduce to pure state ensembles $\{p_x, \psi_x\}$. Then defining the average state $\rho := \sum_x p_x\psi_x$, we consider

$$S(\Lambda(\rho)) - \sum_x p_x S(\Lambda(\psi_x))$$

Since all pure states can be related by unitaries, we have that for each x and y, there is a unitary U such that $\psi_x = U\psi_y U^*$. Thus,

$$S(\Lambda(\psi_x)) = S(\Lambda(U\psi_y U^*)) = S(U\Lambda(\psi_y)U^*) = S(\Lambda(\psi_y))$$

using that the entropy is unitarily invariant. Thus, $S(\Lambda(\psi_x))$ does not depend on x. To calculate its value, we can extend $|\psi_x\rangle$ to an orthonormal basis given by $\{|\psi_x\rangle, |\psi_x\rangle^\perp\}$. Then in this basis,

$$\Lambda(\psi_x) = p\psi_x + (1-p)\frac{I}{2} = \begin{pmatrix} p + \frac{1-p}{2} & 0 \\ 0 & \frac{1-p}{2} \end{pmatrix}.$$

Thus, $S(\Lambda(\psi_x)) = h\left(\frac{1-p}{2}\right)$ where h is the binary entropy. Since $\sum_x p_x = 1$, we have

$$S(\Lambda(\rho)) - \sum_x p_x S(\Lambda(\psi_x)) = S(\Lambda(\rho)) - h\left(\frac{1-p}{2}\right).$$

It remains to maximize the output entropy $S(\Lambda(\rho))$ over all states ρ. Since Λ is unitarily invariant, we can work in the eigenbasis of ρ, in which case $\rho = \begin{pmatrix} \lambda & 0 \\ 0 & 1-\lambda \end{pmatrix}$ for some number $0 \le \lambda \le 1$. In this basis, we have

$$\Lambda(\rho) = \begin{pmatrix} p\lambda + \frac{1-p}{2} & 0 \\ 0 & p(1-\lambda) + \frac{1-p}{2} \end{pmatrix}.$$

Thus,

$$S(\Lambda(\rho)) = h\left(p\lambda + \frac{1-p}{2}\right) = h\left(\frac{1}{2} + p(\lambda - \frac{1}{2})\right)$$

Since the binary entropy is maximized at $\frac{1}{2}$ at which point it takes the value $\log_2 2 = 1$ (since, e.g. the entropy is maximized on a uniform distribution), $S(\Lambda(\rho))$ is maximized when $\lambda = \frac{1}{2}$. Thus,

$$\chi^*(\Lambda) = 1 - h\left(\frac{1-p}{2}\right)$$

gives the product state capacity of the depolarizing (qubit) channel.

Exercise 5. Alice prepares a photon in one of two polarization states, given by the kets $|a\rangle := |1\rangle$, and $|b\rangle := \sin\theta|0\rangle + \cos\theta|1\rangle$, depending on the outcome of a fair coin toss. If the outcome is heads, she prepares the state $|a\rangle$. Otherwise she prepares the state $|b\rangle$. Evaluate the Holevo χ quantity for her ensemble of states. (Use the convention that $|0\rangle = (1\ 0)^T$ and $|1\rangle = (0\ \ 1)^T$). For what value of θ is the Holevo bound achieved? Explain why.

Solution.

The average state is given by

$$\begin{aligned}\rho &= \frac{1}{2}|1\rangle\langle 1| + \frac{1}{2}|b\rangle\langle b| \\ &= \frac{1}{2}[|1\rangle\langle 1| + \sin^2\theta|0\rangle\langle 0| + \cos^2\theta|1\rangle\langle 1| + \sin\theta\cos\theta[|0\rangle\langle 1| + |1\rangle\langle 0|]] \\ &= \frac{1}{2}\begin{pmatrix} \sin^2\theta & \sin\theta\cos\theta \\ \sin\theta\cos\theta & 1+\cos^2\theta \end{pmatrix}\end{aligned}$$

which has eigenvalues $\frac{1}{2}(1 \pm \cos\theta)$. Since the two states are pure, each has zero entropy. The Holevo quantity is therefore simply $h(\frac{1}{2}(1 \pm \cos\theta))$ where h is the binary entropy. Since the binary entropy is maximized at $\frac{1}{2}$, this quantity is maximized when $\cos\theta = 0$, i.e. $\theta = \frac{\pi}{2} + \pi z$ for any $z \in \mathbb{Z}$.

Exercise 6. Alice encodes classical information into n photons which she sends to Bob through a quantum channel. What is the maximum number of bits of information that Bob can infer from the output of the channel by doing measurements on it?

Hint: Use the Holevo bound.

Solution.

Let X denote Alice's classical information, and Y the random variable obtained by Bob from measurements on what he obtains from the quantum channel. The amount of bits of information Bob can infer about Alice's information is exactly $I(X:Y)$. The Holevo bound gives that

$$I(X:Y) \le \chi(\{p_x, \rho_x\})$$

where $\{p_x, \rho_x\}$ is the ensemble used by Alice to encode her information. We don't know how the encoding is done or if the channel is noisy, but we know only n photons were used, which we assume are qubits (information encoded only in the polarization). Thus, the output state $\rho = \sum_x p_x\rho_x$ is 2^n-dimensional, and its entropy is at most n. Since $\chi(\{p_x, \rho_x\}) \le S(\rho) \le n$, we obtain an upper bound of n.

Exercise 7. Entropy exchange. The entropy exchange for a state ρ and a quantum channel Λ is defined as follows:

$$S(\rho, \Lambda) := S(\rho'_{RQ}),$$

where $\rho'_{RQ} = (\mathrm{id}_R \otimes \Lambda)\psi^\rho_{RQ}$, with ψ^ρ_{RQ} being a purification of ρ.

1. Prove that $S(\rho, \Lambda) = S(\rho'_E)$, where $\rho'_E = \mathrm{tr}_{RQ}(\rho'_{RQE})$ with

$$\rho'_{RQE} = (I_R \otimes U_{QE})(\psi^\rho_{RQ} \otimes |0_E\rangle\langle 0_E|)(I_R \otimes U^\dagger_{QE}),$$

   with U_{QE} being the Stinespring dilation of the channel Λ.

   Thus the entropy exchange can be interpreted as the amount of entropy introduced by Λ into an initially pure environment.

2. Prove that the entropy exchange can also be written in the form

$$S(\rho, \Lambda) = S(W) = -\operatorname{tr} W \log W,$$

where W denotes a matrix with elements $W_{ij} = \operatorname{tr}(A_i \rho A_j^\dagger)$, where $\{A_i\}$ denote a set of Kraus operators of Λ.

Solution.

1. We can see that indeed, $\rho'_{RQ} = \operatorname{tr}_E \rho'_{RQE}$. Moreover, ρ'_{RQE} is a pure state as $\psi^\rho_{RQ} \otimes |0_E\rangle\langle 0_E|$ is pure, and we conjugate by a unitary, $I_R \otimes U_{QE}$. Thus, the Schmidt decomposition tells us that across any bipartite partition, the reduced density matrices have the same non-zero eigenvalues, and thus the same entropy. That is, $S(\rho'_{RQ}) = S(\rho'_E)$ as desired.
2. We can trace out R to find

$$\rho'_E = \operatorname{tr}_Q[U_{QE}(\rho_Q \otimes |0_E\rangle\langle 0_E|)U_{QE}^\dagger].$$

Now, we can choose our Strinespring representation to act as $U|\psi_A\rangle \otimes |0_E\rangle = \sum_i A_i|\psi_A\rangle \otimes |i\rangle$ for any pure state $|\psi_A\rangle$ (as shown in the notes where we construct U), where $\{A_i\}$ is a Kraus representation of Λ. Now, let $\rho_Q = \sum_\alpha \lambda_\alpha |\psi_\alpha\rangle\langle\psi_\alpha|$ be a decomposition into pure states. Then

$$\begin{aligned}\rho'_E &= \sum_\alpha \lambda_\alpha \operatorname{tr}_Q[U_{QE}(|\psi_\alpha\rangle\langle\psi_\alpha| \otimes |0_E\rangle\langle 0_E|)U_{QE}^\dagger] \\ &= \sum_\alpha \lambda_\alpha \sum_{i,j} \operatorname{tr}_Q[A_i|\psi_\alpha\rangle\langle\psi_\alpha|A_j^\dagger \otimes |i_E\rangle\langle j_E|] \\ &= \sum_\alpha \lambda_\alpha \sum_{i,j} \operatorname{tr}[A_i|\psi_\alpha\rangle\langle\psi_\alpha|A_j^\dagger]|i_E\rangle\langle j_E| \\ &= \sum_{i,j} \operatorname{tr}[A_i \rho_Q A_j^\dagger]|i_E\rangle\langle j_E|.\end{aligned}$$

Therefore, $W := \rho'_E$ has a matrix representation where the matrix elements are given by $\operatorname{tr}[A_i \rho_Q A_j^\dagger]$. By Step (1), this completes the proof.

Exercise 8. Quantum Fano inequality. Prove the quantum Fano inequality:

$$S(\rho, \Lambda) \le h(F_e(\rho, \Lambda)) + (1 - F_e(\rho, \Lambda)) \log(d^2 - 1)$$

where

- Λ denotes a quantum operation with Kraus representation

$$\Lambda(\rho) = \sum_{i=0}^{d^2} V_i \rho V_i^\dagger,$$

with ρ being the state of a quantum system Q with Hilbert space of dimension d.
- $h(\cdot)$ denotes the binary Shannon entropy, i.e., for any $0 < p < 1$,

$$h(p) = -p \log p - (1 - p) \log(1 - p).$$

- $S(\rho, \Lambda) = -\operatorname{tr}(W \log W)$, with W being a matrix with elements $W_{ij} = \operatorname{tr}(V_i \rho V_j)$. The quantity $S(\rho, \Lambda)$ is called *entropy exchange.*
- $$F_e(\rho, \Lambda) := \langle\psi^\rho_{RQ}|(\mathrm{id} \otimes \Lambda)\psi^\rho_{RQ}|\psi^\rho_{RQ}\rangle,$$

  is the entanglement fidelity. Here $|\psi^\rho_{RQ}\rangle$ is a purification of the state ρ, with R denoting the reference system used for the purification.

What is the implication of the quantum Fano inequality as regards entanglement?

Solution.

First, we have the useful result that the entropy is increasing under unital maps. That is, if Λ is a unital CPTP map, then $S(\Lambda(\rho)) \ge S(\rho)$ for any state ρ. We can see this simply by using the data-processing inequality:

$$S(\rho) = -D(\rho\|I) \le -D(\Lambda(\rho)\|\Lambda(I)) = -D(\Lambda(\rho)\|I) = S(\Lambda(\rho)).$$

Now, from the previous exercise, we know that $S(\rho, \Lambda) = S(\rho'_{RQ})$ for $\rho'_{RQ} = (\mathrm{id}_R \otimes \Lambda)(\psi^\rho_{RQ})$. Let us consider an orthonormal basis $\{|f_i\rangle\}_{i=1}^{d^2}$ for the joint space $\mathcal{H}_R \otimes \mathcal{H}_Q$ such that $|f_1\rangle = |\psi^\rho_{RQ}\rangle$. That is, we take the first vector to be $|\psi^\rho_{RQ}\rangle$ and then use Grahm-Schmidt to complete to an orthonormal basis. Then we define the unital CPTP map

$$\Gamma(\omega_{RQ}) = \sum_i \langle f_i|\omega_{RQ}|f_i\rangle\, |f_i\rangle\langle f_i|.$$

This is sometimes called a *pinching map* for the basis $\{|f_i\rangle\}$. We can check it is CPTP as it has Kraus operators $\{|f_i\rangle\langle f_i|\}_{i=1}^{d^2}$, and unital by substituting $\omega_{RQ} = I_{RQ}$. By our "useful result", we have

$$S(\rho, \Lambda) = S(\rho'_{RQ}) \le S(\Gamma(\rho'_{RQ})) = H(\{\langle f_i|\rho'_{RQ}|f_i\rangle\}_{i=1}^{d^2})$$

which is the Shannon entropy of the probability distribution $p := \{\langle f_i|\rho'_{RQ}|f_i\rangle\}_{i=1}^{d^2}$ (as these are the eigenvalues of $\Gamma(\rho'_{RQ})$ which can see by the definition of Γ). Moreover, we know the first entry is $\langle f_i|\rho'_{RQ}|f_i\rangle = F_e(\rho, \Lambda)$ by construction. Thus, if we define $\varepsilon := 1 - F_e(\rho, \Lambda)$, the distribution p has probability $1 - \varepsilon$ for the first entry. Therefore, by classical Fano's inequality[3],

$$H(p) \le h(\varepsilon) + \varepsilon \log(d^2 - 1) = h(1 - \varepsilon) + \varepsilon \log(d^2 - 1)$$

where to obtain the equality we use that the binary entropy is symmetric (in the sense that $h(x) = h(1 - x)$ for $x \in [0, 1]$). Substituting ε yields thus result.

By the previous exercise, we can see $S(\rho, \Lambda)$ as the entropy of $\operatorname{tr}_E[\rho'_{RQE}]$ (and also of $\operatorname{tr}_{RQ}[\rho'_{RQE}]$). Since ρ'_{RQE} is a pure state, this quantity is the *entropy of entanglement* of the state ρ'_{RQE} across the partition $RQ|E$. This is an *entanglement monotone* (i.e. non-increasing under local operations and classical communication, so-called LOCC operations), and you can check that it is zero for product states and non-zero for entangled states. Since we can see ρ'_{RQE} as induced by the action of Λ (in Stinespring form) on a product state $\psi^\rho_{RQ} \otimes |0_E\rangle\langle 0_E|$, the quantity $S(\rho, \Lambda)$ is a measurement of the entanglement generated by Λ by its action on this state. On the other hand, $F_e(\rho, \Lambda)$ is the fidelity between ψ^ρ_{RQ} and $\Lambda(\psi^\rho_{RQ})$, which

[3] Exercise 8 of Example sheet 1

is close to 1 when ψ^ρ_{RQ} is close to $\Lambda(\psi^\rho_{RQ})$ (so ε is close to zero in this case). So we can upper bound the entanglement generated by Λ in terms of how much it changes the state in fidelity (and in particular, if Λ does not change the state at all, then it can't generate any entanglement).

Exercise 9. Continuity of the von Neumann entropy (Fannes' inequality): Suppose $\rho, \sigma \in \mathcal{D}(\mathcal{H})$ are states such that their trace distance $D(\rho,\sigma)$ satisfies the bound

$$2D(\rho,\sigma) = \|\rho - \sigma\|_1 \leq 1/e.$$

Then

$$|S(\rho) - S(\sigma)| \leq \|\rho-\sigma\|_1 \log d + \eta\left(\|\rho - \sigma\|_1\right), \tag{6}$$

where $d = \dim \mathcal{H}$, and $\eta(x) := -x \log x$.

Let us prove this theorem in steps:

1. Let $r_1 \geq r_2 \geq \ldots \geq r_d$ be the eigenvalues of ρ arranged in non-increasing order, and let $s_1 \geq s_2 \geq \ldots \geq s_d$ be the eigenvalues of σ arranged in non-increasing order. Then prove that:
$$\|\rho - \sigma\|_1 \geq \sum_{j=1}^d |r_j - s_j| \tag{7}$$
2. Check (using elementary calculus) that if $|r-s| \leq 1/2$, then
$$|\eta(r) - \eta(s)| \leq \eta(|r-s|),$$
where $\eta(x) := -x\log x$.
3. Use Step 2 and the triangle inequality to prove that
$$|S(\rho) - S(\sigma)| \leq \sum_j \eta(|r_j - s_j|).$$
4. Let $\varepsilon_j := |r_j - s_j|$, $\forall\, j = 1, 2, \ldots, d$, and $\varepsilon := \sum_j \varepsilon_j$. Let $\lambda_j := \varepsilon_j/\varepsilon$ and note that $\{\lambda_j\}$ forms a probability distribution. Use this fact and Step 3 to prove that
$$|S(\rho) - S(\sigma)| \leq \varepsilon \log d + \eta(\varepsilon).$$
5. Note that $\eta(\varepsilon)$ is a monotonically increasing function of ε for $0 \leq \varepsilon \leq 1/e$. Use this to finally arrive at the statement (6).

Solution.

1. I could not find a short proof of this fact (without assuming something that directly implies it). This subquestion is then nonexaminable; see e.g. the excellent book *Matrix Analysis* by Bhatia[4] for several (more involved) proofs.

[4]DOI: 10.1007/978-1-4612-0653-8. In particular, equation IV.62 of that book is a restatement of this question.

2. Without loss of generality, take $r > s$ and set $\delta := r - s$. Then
$$\begin{aligned}\eta(s+\delta) - \eta(s) - \eta(\delta) &= -(s+\delta)\log(s+\delta) + s\log s + \delta\log(\delta) \\ &= -s\log(s+\delta) - \delta\log(s+\delta) + s\log s + \delta \log(\delta) \\ &= s\log\frac{s}{s+\delta} + \delta\log\frac{\delta}{s+\delta} \\ &\leq 0\end{aligned}$$
since both $\frac{s}{s+\delta} \leq 1$ and $\frac{\delta}{s+\delta} \leq 1$. Thus,
$$\eta(r) - \eta(s) \leq \eta(r-s).$$
On the other hand, set
$$f(\delta) := \eta(s) - \eta(s+\delta) - \eta(\delta).$$
It remains to show that $f(\delta) \leq 0$ for $\delta \leq \frac{1}{2}$ (and $r = s+\delta \leq 1$), which implies
$$\eta(s) - \eta(s+\delta) \leq \eta(\delta)$$
meaning that we have proven
$$|\eta(r) - \eta(s)| \leq \eta(|r-s|).$$
First, $f(0) = 0$, and
$$f'(\delta) = -\eta'(s+\delta) - \eta'(\delta)$$
while
$$\eta'(x) = \frac{\mathrm{d}}{\mathrm{d}x}(-x\log x) = -\log x - \frac{1}{\ln 2},$$
so
$$f'(\delta) = \log(s+\delta) + \frac{1}{\ln 2} + \log(\delta) + \frac{1}{\ln 2}$$
and
$$f''(\delta) = \frac{1}{\ln 2}\left[\frac{1}{s+\delta} + \frac{1}{\delta}\right] > 0$$
for $\delta > 0$. Thus, f is convex. We wish to prove $f(\delta) \leq 0$ for $\delta \in [0, \min(1-s, \frac{1}{2})]$, and by convexity, we only need to check the boundaries (since f will be less than either boundary point between the two). We have immediately $f(0) = 0$. Next, consider $s \geq \frac{1}{2}$. Then let's check $g(s) := f(1-s) = \eta(s) - \eta(1-s) \leq 0$ for $s \geq \frac{1}{2}$. We can see at $s = \frac{1}{2}$, $g(s) = 0$, and $g'(s) = -\log(s) + \log(1-s) \leq 0$ since $1 - s \leq s$. Thus, g is decreasing for all $s \in [\frac{1}{2}, 1]$, proving indeed, $g(s) \leq 0$ on that interval.

Lastly, we need to check when $s < \frac{1}{2}$ that $f(\frac{1}{2}) \leq 0$. In this case,
$$h(s) := f(\frac{1}{2}) = \eta(s) - \eta(s + \tfrac{1}{2}) - \eta(\tfrac{1}{2}).$$
Then $h(0) = -2\eta(\frac{1}{2}) \leq 0$, and $h'(s) = -\log(s) + \log(s+\frac{1}{2}) > 0$, so h is increasing. Thus, h takes a maximum at the end of the interval we wish to test, at $s = \frac{1}{2}$. But here, $h(\frac{1}{2}) = \eta(\frac{1}{2}) - \eta(1) - \eta(\frac{1}{2}) = 0$. Thus, $f(\frac{1}{2}) \leq 0$ for $s \in [0, \frac{1}{2}]$.

3. We have that $S(\rho) = \sum_j \eta(r_j)$ and $S(\sigma) = \sum_j \eta(s_j)$. Therefore,

$$\begin{aligned} |S(\rho) - S(\sigma)| &= \left| \sum_j \eta(r_j) - \eta(s_j) \right| \\ &\leq \sum_j |\eta(r_j) - \eta(s_j)| \\ &\leq \sum_j \eta(|r_j - s_j|). \end{aligned}$$

4. Note that

$$\eta(xy) = -xy \log(xy) = -xy(\log x + \log y) = -xy \log x - xy \log y = y\eta(x) + x\eta(y)$$

Thus, $\eta(\varepsilon_j) = \eta(\varepsilon\lambda_j) = \varepsilon\eta(\lambda_j) + \lambda_j \eta(\varepsilon)$, and

$$\begin{aligned} |S(\rho) - S(\sigma)| &\leq \sum_j \eta(\varepsilon_j) \\ &= \varepsilon \sum_j \eta(\lambda_j) + \eta(\varepsilon) \\ &= \eta(\varepsilon) + H(\{\lambda_j\}_j) \\ &\leq \eta(\varepsilon) + \varepsilon \log d \end{aligned}$$

using that $\{\lambda_j\}_j$ is a probability distribution, and the entropy of any probability distribution with d elements is bounded by $\log d$.

5. By step 1, $\varepsilon \leq \|\rho - \sigma\|_1$. Since $\varepsilon \mapsto \eta(\varepsilon)$ is monotonically increasing on the range $[0, 1/e]$, if $\|\rho - \sigma\|_1 \leq 1/e$, then

$$\begin{aligned} |S(\rho) - S(\sigma)| &\leq \eta(\varepsilon) + \varepsilon \log d \\ &\leq \eta(\|\rho - \sigma\|_1) + \|\rho - \sigma\|_1 \log d \end{aligned}$$

as desired.

Exercise 10. An interesting class of quantum channels are the *entanglement-breaking (EB) channels*. An EB channel Λ is one for which $(\mathrm{id} \otimes \Lambda)(\omega)$ is separable, even for entangled ω[5]. The Holevo capacity has been proved to be additive for EB channels.

1. Prove that any channel of the following form is EB:

$$\Lambda(\rho) = \sum_k \sigma_k \operatorname{tr}(E_k \rho), \tag{8}$$

where σ_k are density matrices and $\{E_k\}$ is a POVM. The above form has the following physical interpretation. Alice does a measurement (POVM) on the input state ρ and communicates the outcomes k to Bob via a classical channel; Bob then prepares an agreed upon state σ_k. Hence, EB channels are also called "measure-and-prepare channels".

[5] It derives its name from the fact that the channel outputs a separable state whenever half of an entangled state is input to it.

2. Prove that if the Choi state $(I \otimes \Lambda)|\Omega\rangle\langle\Omega|$ (where $|\Omega\rangle$ denotes the unnormalized maximally entangled state) is separable, then Λ has the form (8).

Solution.

1. First, let us consider a pure state $|\psi\rangle_{AB}$ with Schmidt decomposition $|\psi\rangle_{AB} = \sum_i \sqrt{\lambda_i}|e_i\rangle|f_i\rangle$. Then

$$\begin{aligned} \mathrm{id}_A \otimes \Lambda(\psi_{AB}) &= \sum_{i,j} \sqrt{\lambda_i \lambda_j} \sum_k \operatorname{tr}(E_k |f_i\rangle\langle f_j|)|e_i\rangle\langle e_j| \otimes \sigma_k. \\ &= \sum_k \sum_{i,j} \sqrt{\lambda_i \lambda_j} \langle f_j|E_k|f_i\rangle |e_i\rangle\langle e_j| \otimes \sigma_k. \end{aligned}$$

Now, let $p_k = \operatorname{tr}[\sum_{i,j} \sqrt{\lambda_i \lambda_j}\langle f_j|E_k|f_i\rangle|e_i\rangle\langle e_j|] = \sum_i \lambda_i \langle f_i|E_k|f_i\rangle$. Then since $E_k \geq 0$ and $\lambda_i \geq 0$ we have $p_k \geq 0$, and moreover

$$\begin{aligned} \sum_k p_k &= \sum_i \sum_k \lambda_i \langle f_i|E_k|f_i\rangle \\ &= \sum_i \lambda_i \langle f_i| \sum_k E_k |f_i\rangle \\ &= \sum_i \lambda_i \langle f_i|I|f_i\rangle \\ &= \sum_i \lambda_i = 1. \end{aligned}$$

Define $\omega_k = \frac{1}{p_k} \sum_{i,j} \sqrt{\lambda_i \lambda_j}\langle f_j|E_k|f_i\rangle|e_i\rangle\langle e_j|$. Let $|\psi\rangle$ be any vector. Then

$$\begin{aligned} \langle\psi|\omega_k\psi\rangle &= \frac{1}{p_k} \sum_{i,j} \sqrt{\lambda_i \lambda_j}\langle f_j|E_k|f_i\rangle\langle\psi|e_i\rangle\langle e_j|\psi\rangle \\ &= \frac{1}{p_k} \sum_\ell \sum_{i,j} \sqrt{\lambda_i \lambda_j}\langle f_j|E_k^{1/2}|\ell\rangle\langle\ell|E_k^{1/2}|f_i\rangle\langle\psi|e_i\rangle\langle e_j|\psi\rangle \\ &= \frac{1}{p_k} \sum_\ell \left| \sum_i \sqrt{\lambda_i}\langle\ell|E_k^{1/2}|f_i\rangle \, \langle\psi|e_i\rangle \right|^2 \geq 0. \end{aligned}$$

Thus, $\omega_k \geq 0$. Since we chose p_k so that $\operatorname{tr} \omega_k = 1$, we have that ω_k is a density matrix. Then

$$\mathrm{id}_A \otimes \Lambda(\psi_{AB}) = \sum_k p_k \omega_k \otimes \sigma_k$$

is a separable state. To treat a mixed state ρ_{AB} we simply consider a convex decomposition into pure states, $\rho_{AB} = \sum_\alpha q_\alpha \psi_\alpha$, and use linearity to see that ρ_{AB} must be separable as well (i.e. using that the set of separable states is convex).

2. Assume $(I \otimes \Lambda)|\Omega\rangle\langle\Omega|$ is separable. Then

$$\begin{aligned} (I \otimes \Lambda)|\Omega\rangle\langle\Omega| &= \sum_{i,j} |i\rangle\langle j| \otimes \Lambda(|i\rangle\langle j|) \\ &= \sum_k p_k \psi_k \otimes \phi_k \end{aligned}$$

for pure states ψ_k, ϕ_k and a probability distribution $\{p_k\}$. We can multiply each side by $|\ell\rangle\langle\ell'| \otimes I$ and partial trace over the first system to find

$$
\begin{aligned}
\Lambda(|\ell'\rangle\langle\ell|) &= \sum_k p_k \operatorname{tr}[|\ell\rangle\langle\ell'|\psi_k]\phi_k \\
&= \sum_k p_k \operatorname{tr}[(|\ell'\rangle\langle\ell|)^T \psi_k]\phi_k \\
&= \sum_k p_k \operatorname{tr}[(|\ell'\rangle\langle\ell|)^T (\bar{\psi}_k)^T]\phi_k \\
&= \sum_k p_k \operatorname{tr}[(\bar{\psi}_k|\ell'\rangle\langle\ell|)^T]\phi_k \\
&= \sum_k p_k \operatorname{tr}[\bar{\psi}_k|\ell'\rangle\langle\ell|]\phi_k \\
&= \sum_k p_k \operatorname{tr}[|\ell'\rangle\langle\ell|\bar{\psi}_k]\phi_k
\end{aligned}
$$

using $\psi_k = \psi_k^\dagger = (\bar{\psi}_k)^T$, where A^T is the transpose of a matrix A in the basis $\{|i\rangle\}$ and $\bar{A}$ is the complex conjugate of A in the same basis. In the second-to-last line we used that the trace is invariant under transpose, and in the last line, the cyclicity of the trace. Note the complex conjugate matrix $\bar{\psi}_k$ is still a pure state (one can check it is a self-adjoint, rank-one projector).

By linearity, we therefore have

$$
\Lambda(\rho) = \sum_k p_k \operatorname{tr}[\rho\bar{\psi}_k]\phi_k
$$

Let $E_k = p_k\bar{\psi}_k$ so that $\Lambda(\rho) = \sum_k \operatorname{tr}[\rho E_k]\phi_k$. It remains to prove $\{E_k\}$ is a POVM. We have immediately that $E_k \geq 0$. To show $\sum_k E_k = I$, we note that we can write

$$
\Lambda(\rho) = \sum_k p_k \langle\bar{\psi}_k|\rho|\bar{\psi}_k\rangle|\phi_k\rangle\langle\phi_k| = \sum_k p_k|\phi_k\rangle\langle\bar{\psi}_k|\rho|\bar{\psi}_k\rangle\langle\phi_k|.
$$

We can therefore define $A_k := \sqrt{p_k}|\phi_k\rangle\langle\bar{\psi}_k|$ yielding $\{A_k\}$ as a set of Kraus operators for Λ. Thus, since Λ is TP,

$$
\begin{aligned}
I &= \sum_k A_k^\dagger A_k \\
&= \sum_k p_k|\bar{\psi}_k\rangle\langle\phi_k||\phi_k\rangle\langle\bar{\psi}_k| \\
&= \sum_k p_k|\bar{\psi}_k\rangle\langle\bar{\psi}_k| \\
&= \sum_k E_k
\end{aligned}
$$

as desired. Thus, $\{E_k\}$ is a POVM, and indeed Λ has the form (8).