

## Example Sheet 4 Solutions

**Exercise 1.** Prove that the Holevo capacity is superadditive,

$$\chi^*(\Lambda_1 \otimes \Lambda_2) \geq \chi^*(\Lambda_1) + \chi^*(\Lambda_2),$$

where  $\Lambda_1$  and  $\Lambda_2$  denote quantum channels.

**Solution.**

$$\chi^*(\Lambda_1 \otimes \Lambda_2) := \max_{\{p_x, \rho_x\}} \left\{ S(\Lambda_1 \otimes \Lambda_2(\sum_x p_x \rho_x)) - \sum_x p_x S(\Lambda_1 \otimes \Lambda_2(\rho_x)) \right\}$$

where the maximum is over ensembles  $\{p_x, \rho_x\}$  with states  $\rho_x \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ . The maxima over all such distributions is larger than the maxima over the restriction to product distributions  $\{p_x q_y\}_{x,y}$  and product states  $\{\rho_x^A \otimes \sigma_y^B\}_{x,y}$ . Thus,

$$\chi^*(\Lambda_1 \otimes \Lambda_2) \geq \max_{\{p_x q_y, \rho_x^A \otimes \sigma_y^B\}} \left\{ S((\Lambda_1 \otimes \Lambda_2)(\sum_{xy} p_x q_y \rho_x^A \otimes \sigma_y^B)) - \sum_{xy} p_x q_y S(\Lambda_1 \otimes \Lambda_2(\rho_x^A \otimes \sigma_y^B)) \right\}. \quad (1)$$

Now, using the additivity of the von Neumann entropy under tensor products, we have

$$\begin{aligned} S((\Lambda_1 \otimes \Lambda_2)(\sum_{xy} p_x q_y \rho_x^A \otimes \sigma_y^B)) &= S(\sum_x p_x \Lambda_1(\rho_x^A) \otimes \sum_y q_y \Lambda_2(\sigma_y^B)) \\ &= S(\sum_x p_x \Lambda_1(\rho_x^A)) + S(\sum_y q_y \Lambda_2(\sigma_y^B)), \end{aligned}$$

and

$$\begin{aligned} \sum_{xy} p_x q_y S((\Lambda_1 \otimes \Lambda_2)(\rho_x^A \otimes \sigma_y^B)) &= \sum_{xy} p_x q_y [S(\Lambda_1(\rho_x^A)) + S(\Lambda_2(\sigma_y^B))] \\ &= \sum_{xy} p_x q_y S(\Lambda_1(\rho_x^A)) + \sum_{xy} p_x q_y S(\Lambda_2(\sigma_y^B)) \\ &= \sum_x p_x S(\Lambda_1(\rho_x^A)) + \sum_y q_y S(\Lambda_2(\sigma_y^B)). \end{aligned}$$

So the right-hand side of (1) is equal to

$$\max_{\{p_x q_y, \rho_x^A \otimes \sigma_y^B\}} \left\{ S(\sum_x p_x \Lambda_1(\rho_x^A)) - \sum_x p_x S(\Lambda_1(\rho_x^A)) + S(\sum_y q_y \Lambda_2(\sigma_y^B)) - \sum_y q_y S(\Lambda_2(\sigma_y^B)) \right\}$$

which is the same as

$$\max_{\{p_x, \rho_x^A\}} \left\{ S(\sum_x p_x \Lambda_1(\rho_x^A)) - \sum_x p_x S(\Lambda_1(\rho_x^A)) \right\} + \max_{\{q_y, \sigma_y^B\}} \left\{ S(\sum_y q_y \Lambda_2(\sigma_y^B)) - \sum_y q_y S(\Lambda_2(\sigma_y^B)) \right\}$$

i.e.  $\chi^*(\Lambda_1) + \chi^*(\Lambda_2)$ .

**Exercise 2.** Let  $|\psi\rangle_{ABE}$  be a pure state of a tripartite system  $ABE$ . Define the *coherent information*<sup>1</sup> from  $A$  to  $B$  of  $\psi$  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  $\psi$ .

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.

<sup>1</sup>In the lecture notes, we defined the coherent information of a quantum channel corresponding to an input state  $I_c(\Lambda, \rho) = I_c^{B>Q'}(\rho')$  (c.f. Notes 18).

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.** Show that if  $\rho_{AB}$  is separable then  $I_c^{A>B} \leq 0$ .

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  $\rho_{AB}$  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<sup>2</sup>, the coherent information is convex. Thus

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

But each entropy is the entropy of a pure state and thus vanishes. Thus,  $I_c^{A>B}(\rho) \leq 0$ . 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,  $\Lambda$ , defined by

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

**Solution.**

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

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

<sup>2</sup>Exercise 12 of Example Sheet 3

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

The HSW theorem gives that the product state capacity of  $\Lambda$  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  $\{|\psi_x\rangle, |\psi_x^\perp\rangle\}$ . 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  $\rho$ . Since  $\Lambda$  is unitarily invariant, we can work in the eigenbasis of  $\rho$ , in which case

$\rho = \begin{pmatrix} \lambda & 0 \\ 0 & 1-\lambda \end{pmatrix}$  for some number  $0 \leq \lambda \leq 1$ . 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,

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

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  $\chi$  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  $\theta$  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

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

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  $\chi(\{p_x, \rho_x\}) \leq S(\rho) \leq n$ , we obtain an upper bound of  $n$ .

**Exercise 7. Entropy exchange.** The entropy exchange for a state  $\rho$  and a quantum channel  $\Lambda$  is defined as follows:

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

where  $\rho'_{RQ} = (\text{id}_R \otimes \Lambda)\psi'_{RQ}$ , with  $\psi'_{RQ}$  being a purification of  $\rho$ .

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

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

with  $U_{QE}$  being the Stinespring dilation of the channel  $\Lambda$ .

*Thus the entropy exchange can be interpreted as the amount of entropy introduced by  $\Lambda$  into an initially pure environment.*

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

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

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

**Solution.**

1. We can see that indeed,  $\rho'_{RQ} = \text{tr}_E \rho'_{RQE}$ . Moreover,  $\rho'_{RQE}$  is a pure state as  $\psi'_{RQ} \otimes |0_E\rangle\langle 0_E|$  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

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

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  $\Lambda$ . Now, let  $\rho_Q = \sum_\alpha \lambda_\alpha |\psi_\alpha\rangle\langle \psi_\alpha|$  be a decomposition into pure states. Then

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

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

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

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

where

- $S(\rho, \Lambda)$  is the entropy exchange defined in the previous question,
- $\Lambda$  denotes a quantum operation with Kraus representation

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

with  $\rho$  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  $0 < p < 1$ ,

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

- $F_e(\rho, \Lambda)$  defined by

$$F_e(\rho, \Lambda) := \langle \Psi_{RQ}^\rho | (\text{id} \otimes \Lambda) \Psi_{RQ}^\rho | \Psi_{RQ}^\rho \rangle,$$

is the entanglement fidelity. Here  $|\Psi_{RQ}^\rho\rangle$  is a purification of the state  $\rho$ , 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  $\Lambda$  is a unital CPTP map, then  $S(\Lambda(\rho)) \geq S(\rho)$  for any state  $\rho$ . We can see this simply by using the data-processing inequality:

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

Now, from the previous exercise, we know that  $S(\rho, \Lambda) = S(\rho'_{RQ})$  for  $\rho'_{RQ} = (\text{id}_R \otimes \Lambda)(\psi_{RQ}^\rho)$ . 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  $|f_1\rangle = |\psi_{RQ}^\rho\rangle$ . That is, we take the first vector to be  $|\psi_{RQ}^\rho\rangle$  and then use Gram-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

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

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  $\Gamma$ ). Moreover, we know the first entry is  $\langle f_1 | \rho'_{RQ} | f_1 \rangle = F_e(\rho, \Lambda)$  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<sup>3</sup>,

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

<sup>3</sup>Exercise 8 of Example sheet 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  $\varepsilon$  yields thus result.

By the previous exercise, we can see  $S(\rho, \Lambda)$  as the entropy of  $\text{tr}_E[\rho'_{RQE}]$  (and also of  $\text{tr}_{RQ}[\rho'_{RQE}]$ ). Since  $\rho'_{RQE}$  is a pure state, this quantity is the *entropy of entanglement* of the state  $\rho'_{RQE}$  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  $\rho'_{RQE}$  as induced by the action of  $\Lambda$  (in Stinespring form) on a product state  $\psi_{RQ}^p \otimes |0_E\rangle\langle 0_E|$ , the quantity  $S(\rho, \Lambda)$  is a measurement of the entanglement generated by  $\Lambda$  by its action on this state. On the other hand,  $F_e(\rho, \Lambda)$  is the fidelity between  $\psi_{RQ}^p$  and  $\Lambda(\psi_{RQ}^p)$ , which is close to 1 when  $\psi_{RQ}^p$  is close to  $\Lambda(\psi_{RQ}^p)$  (so  $\varepsilon$  is close to zero in this case). So we can upper bound the entanglement generated by  $\Lambda$  in terms of how much it changes the state in fidelity (and in particular, if  $\Lambda$  does not change the state at all, then it can't generate any entanglement).

**Exercise 9.** Let  $\rho$  be a quantum state on  $\mathcal{H}_Q$ , and  $\Lambda$  a quantum operation on  $Q$ . We say it is possible to *perfectly reverse*  $\Lambda$  on  $\rho$  if there exists a quantum operation  $\Lambda'$  on  $Q$  such that  $F_e(\Lambda' \circ \Lambda, \rho) = 1$ . Show that if it is possible to perfectly reverse  $\Lambda$  on  $\rho$ , then

$$S(\rho) = I_c(\Lambda, \rho)$$

where  $I_c(\Lambda, \rho)$  is the coherent information of the channel  $\Lambda$  when the input state is  $\rho$ .

*Note: there was an error in the stated version of this question;  $\Lambda' \circ \Lambda(\rho) = \rho$  is not enough to prove that  $S(\rho) = I_c(\Lambda, \rho)$ .*

**Solution.**

The quantum data processing inequality gives that

$$I_c(\Lambda' \circ \Lambda, \rho) \leq I_c(\Lambda, \rho) \leq S(\rho_Q). \quad (2)$$

Let  $\Lambda$  have a Stinespring unitary  $U_{QE_1}$  and  $\Lambda'$  have a Stinespring unitary  $U'_{QE_2}$ . Then if  $\psi_{RQ}$  is a purification of  $\rho_Q$ , let

$$\rho''_{RQE_1E_2} = U'_{QE_2} U_{QE_1} (\psi_{RQ} \otimes |0\rangle\langle 0|_{E_1} \otimes |0\rangle\langle 0|_{E_2}) U_{QE_1}^\dagger U_{QE_2}^\dagger$$

denote the joint state of  $RQE_1E_2$  after the action of both channels. Note that identity operators have been omitted in the above expression; e.g.  $U_{QE_1}$  means  $U_{QE_1} \otimes I_{RE_2}$ . Note that  $\rho''_{RQE_1E_2}$  is a pure state. From the definition of coherent information, we have

$$I_c(\Lambda' \circ \Lambda, \rho) = S(\rho''_Q) - S(\rho''_{E_1E_2}). \quad (3)$$

From Exercise 7, we know that

$$S(\rho''_{E_1E_2}) = S(\rho, \Lambda' \circ \Lambda)$$

is the entropy exchange, and by the monotonicity of the fidelity under partial trace, we can see that

$$1 = F_e(\Lambda' \circ \Lambda, \rho) \leq F(\rho''_Q, \rho_Q) \leq 1$$

and hence  $\rho''_Q = \rho_Q$ . Thus, the right-hand side of (3) becomes

$$S(\rho_Q) - S(\rho, \Lambda' \circ \Lambda).$$

Since

$$0 \leq S(\rho, \Lambda' \circ \Lambda) \leq h(F_e(\rho, \Lambda)) + (1 - F_e(\rho, \Lambda)) \log(d^2 - 1) = 0$$

by the quantum Fano inequality, we have  $S(\rho, \Lambda' \circ \Lambda) = 0$  and hence  $I_c(\Lambda' \circ \Lambda, \rho) = S(\rho_Q)$ . Thus the DPI (2) yields the result.