A proof of additivity violation of minimum output entropy of quantum channels [1]

Motohisa Fukuda

[1] Base on [Fukuda].

Quantum states and (random) channels in finite dimension

1. A quantum state ρ is a positive semi-definite Hermitian operator of trace one on a Hilbert space $\mathbb{C}^n$.
2. A physical picture of (complementary [2]) channels is

$$\Phi(\rho) = \mathrm{Tr}_{\mathbb{C}^n}\left[U(ee^* \otimes \rho)U^*\right]$$

for state ρ.
- The above $e \in S_k$ and $U \in \mathcal{U}(kn)$ define the channel $\Phi(\cdot)$ [3], where S_k is the set of unit vectors in $\mathbb{C}^k$.
- The dimensions of input and output spaces are n and k, respectively.

3. We can define random channels by choosing $U \in \mathcal{U}(kn)$ with respect to the Haar measure on the unitary group.

[2] [Holevo], [King, Matsumoto, Nathanson, Ruskai].
[3] $e = |e\rangle$ and $e^* = \langle e|$ in the conventional bra-ket notation.

Quantum channels as subspaces

1. Again, we have a channel: for $x \in \mathbb{C}^n$,

$$\Phi(xx^*) = \mathrm{Tr}_{\mathbb{C}^n}\left[U(ee^* \otimes xx^*)U^*\right] = \mathrm{Tr}_{\mathbb{C}^n}\left[U(e \otimes x)(e^* \otimes x^*)U^*\right]$$

2. Importantly, $U(e \otimes \mathbb{C}^n)$ defines an n-dimensional subspace in $\mathbb{C}^k \otimes \mathbb{C}^n$. So, identify it as the input space and denote it by E.
3. Then, for $x \in \tilde{E} = E \cap S_{kn}$,

$$\Phi(xx^*) = \mathrm{Tr}_{\mathbb{C}^n}\left[xx^*\right] = XX^*$$

Here, we have the following canonical identification:

$$\mathbb{C}^k \otimes \mathbb{C}^n = \mathcal{M}_{k,n}(\mathbb{C})$$
$$\mathbb{C}^n \cong E \ni x = X$$

Minimum output entropy

1. The von Neumann entropy $S(\cdot)$ of state ρ is:

$$S(\rho) = -\mathrm{Tr}[\rho \log \rho] = -\sum_{i=1}^{d} \lambda_i \log \lambda_i$$

   where λ_i are eigenvalues of ρ.

2. The minimal output entropy of channel Φ is defined by

$$S_{\min}(\Phi) = \min_{\rho} S(\Phi(\rho))$$

   where ρ are input states. [4]

3. Since the entropy function $S(\cdot)$ is concave, we only care about pure input states.

[4] [King and Ruskai]

Additivity violation

1. Define the complex conjugate of (random) channel Φ by

$$\bar{\Phi}(\rho) = \mathrm{Tr}_{\mathbb{C}^n}\left[\bar{U}(ee^* \otimes \rho)U^T\right]$$

2. Then, we have additivity violation [5]

$$S_{\min}(\Phi \otimes \bar{\Phi}) < S_{\min}(\Phi) + S_{\min}(\bar{\Phi}) \quad (= 2S_{\min}(\Phi))$$

with high probability when $1 \ll k \ll n$ ($n \sim k^2$ is possible).

3. Note that

$$\min_{\rho\otimes\sigma} S((\Phi \otimes \Omega)(\rho \otimes \sigma)) = \min_{\rho} S(\Phi(\rho)) + \min_{\sigma} S(\Omega(\sigma))$$

[5] [Hastings]: more precisely, random unitary channels are used.

Basic plot of proving additivity violation

1. Easy part: for any channel Φ

$$S_{\min}(\Phi \otimes \bar{\Phi}) \lessapprox 2 \log k - \frac{\log k}{k}$$

2. Difficult part: There exists a constant $C > 0$ such that

$$S_{\min}(\Phi) > \log k - \frac{C}{2k}.$$

with high probability whenever $1 \ll k$ and $k^2 \lesssim n$.

3. As a result,

$$\begin{aligned} S_{\min}(\Phi \otimes \bar{\Phi}) &\lessapprox 2 \log k - \frac{\log k}{k} \\ &< 2\left(\log k - \frac{C}{2k}\right) < S_{\min}(\Phi) + S_{\min}(\bar{\Phi}) \end{aligned}$$

Note that $S_{\min}(\Phi) = S_{\min}(\bar{\Phi})$.

Hayden-Winter trick

1. First, we have for a Bell state b,

$$\left\|\Phi \otimes \bar{\Phi}(bb^*)\right\|_\infty \geq \frac{(\text{input dimension})}{(\text{output dim.}) \times (\text{environment dim.})}^6$$

2. This means that the largest eigenvalue in our setting is larger than

$$\frac{1}{k} = \frac{n}{k \times n}$$

3. This ratio sort of shows the trade-off in proving violation of additivity. If this ratio is larger, the product channel $\Phi \otimes \bar{\Phi}$ has a larger eigenvalue but the single channel Φ is less noisy.

[6] [Hayden and Winter].

H-W trick with random unitary channels

1. Let's get convinced with the trick by random unitary channels:

$$\Psi(\rho) = \sum_{i=1}^{k} p_i U_i \rho U_i^*$$

where $U_i \in \mathcal{U}(n)$ and $\{p_i\}$ is a probability distribution.

2. Note that

$$(U \otimes \bar{U})b = b$$

3. This enables us to calculate as follows. [7]

$$\Psi \otimes \bar{\Psi}(bb^*) = \sum_{i=1}^{k} p_i^2 bb^* + \sum_{i \neq j} \underbrace{p_i p_j (U_i \otimes \bar{U}_j) bb^* (U_i^* \otimes U_j^T)}_{k^2-k \text{ positive terms}}$$

Therefore,

$$\langle b, \Psi \otimes \bar{\Psi}(bb^*)b \rangle \geq \sum_{i=1}^{k} p_i^2 \geq \frac{1}{k}$$

[7] [Hastings].

Easy part with random unitary channels

1. The largest entropy under the condition is given by the following eigenvalue distribution:

$$\left(\frac{1}{k}, \underbrace{\frac{1}{k^2}, \ldots, \frac{1}{k^2}}_{k^2-k}, 0, \ldots, 0\right)$$

2. We can get the following upper bound

$$\begin{aligned} S\left(\Psi \otimes \bar{\Psi}(bb^*)\right) &\leq -\frac{1}{k}\log\frac{1}{k} - \left(1-\frac{1}{k}\right)\log\left(\frac{1}{k^2}\right) \\ &= 2\log k - \frac{\log k}{k} \end{aligned}$$

3. Almost the same for our case.

How can the difficult part be proven?

1. Wishart matrix and "TUBE"-argument. [8]
   [Hastings]; [Fukuda, King and Moser].
2. Lévy's lemma and "TUBE"-argument.
   [Brandão and Horodecki].
3. Dvoretzky's theorem with Schechtman's improvement.
   [Aubrun, Szarek and Werner].
4. Lévy's lemma and ϵ-net argument. [9]
   [Fukuda].

[8] The name "TUBE" was used in [Fukuda, King and Moser].

[9] This pair of techniques was used to show existence of strongly entangled subspace in [Hayden, Leung and Winter], which resulted in violation of minimum output p-Reni entropy for $p > 1$, proven in [Hayden and Winter].

What we prove

1. We want to show that generic random channels are very noisy:

$$S_{\min}(\Phi) = \min_{\rho} S(\Phi(\rho)) > \log k - \frac{C}{k}.$$

2. By Taylor expansion around $\tilde{I}_k = I_k/k$,

$$S(\Phi(\rho)) \approx \log k - \frac{k}{2}\|\Phi(\rho) - \tilde{I}_k\|_2^2$$

3. In fact, we have

$$\log k - S(\Phi(\rho)) \leq k\|\Phi(\rho) - \tilde{I}_k\|_2^2$$

So, let's prove

$$\max_{\rho} \|\Phi(\rho) - \tilde{I}_k\|_2 \leq \frac{\sqrt{C}}{k}$$

for some $C > 0$.

Idea of proving the difficult part

1. We want to show existence of noisy channels; there exists a constant $C > 0$ such that with some probability

$$\max_{x \in S_n} \|\Phi(xx^*) - \tilde{I}_k\|_2 \leq \frac{C}{k} \tag{1}$$

   where $S_n \subset \mathbb{C}^n$ is the subset of unit vectors.
2. In fact, for fixed $x \in S_n$, $\Phi(xx^*)$ is likely to be very mixed, i.e.,

$$\|\Phi(xx^*) - \tilde{I}_k\|_2$$

   is likely to be small for random channels Φ.
3. If this happens with high probability (measure concentration), (1) may hold for some channels (ϵ-net argument).

Generically outputs are noisy for fixed inputs

1. Fix $x_0 \in \tilde{E} = E \cap S_{kn}$, then random unitary generate the uniformly random vector in the unit sphere $S_{kn} \subset \mathbb{C}^{kn}$:
$$x = Ux_0 \in U\tilde{E} \qquad U \in \mathcal{U}(kn)$$
2. Then, the random outputs are
$$\mathrm{Tr}_{\mathbb{C}^n}[xx^*] = XX^*$$
where XX^* is the Wishart matrix.
3. Our interested function:
$$f(x) = \|XX^* - \mathrm{Tr}[XX^*]\tilde{I}_k\|_2$$
can be easily proven to have the following bound for the median: $\exists C > 0$, for large enough $k, n \in \mathbb{N}$ with $n \gtrsim k^2$
$$\mathrm{med}(f) \leq \frac{C}{k}$$
The large deviation is considered in the next several slides.

Lévy's lemma "in a closer view"

1. Take $A \subset S_{kn}$ with $\Pr(A) \geq \frac{1}{2}$ then

$$\Pr\{x \in S_{kn} : \mathrm{dist}(x, A) > \epsilon\} < C \exp\{-c\epsilon^2 kn\}$$

where $c, C > 0$ are universal constants.

2. Take a real-valued continuous function f on S_{kn} and suppose its Lipschitz constant is $L > 0$ on the following set:

$$G = \{x \in S_{kn} : \mathrm{dist}(x, A) \leq \epsilon\},$$

where $A = \{x \in S_{kn} : f(x) \leq \mathrm{med}(f)\}$.

3. Then, we argue in the following way:

$$x \in G \Rightarrow f(x) \leq \mathrm{med}(f) + \epsilon L$$

Hence we get

$$\Pr\{x \in S_{kn} : f(x) > \mathrm{med}(f) + \epsilon L\} < C \exp\{-c\epsilon^2 kn\}$$

Application of Lévy's lemma (1)

1. Let $\epsilon = \sqrt{t/k}$ where $t > 0$ will be chosen wisely large later:

$$\Pr\left\{x \in S_{kn} : f(x) \geq \mathrm{med}(f) + L\sqrt{t/k}\right\} \leq C\exp\{-ctn\}$$

2. What is L on G? First, for $x, y \in S_{kn}$ we have

$$|f(x) - f(y)| \leq \|XX^* - YY^*\|_2 \leq (\|X\|_\infty + \|Y\|_\infty)\|X - Y\|_2$$

3. Next, for $x \in A \quad (\Leftrightarrow f(x) \leq \mathrm{med}(f) \leq C/k)$,

$$\|XX^*\|_\infty \leq f(x) + \|\tilde{I}_k\|_\infty \leq C/k + 1/k \leq C'/k$$

This implies that for $x \in G$,

$$\|X\|_\infty \leq \sqrt{C'/k} + \sqrt{t/k} \leq 2\sqrt{t/k}$$

4. Finally, we have

$$L \leq 4\sqrt{t/k} \qquad \text{on } G$$

Application of Lévy's lemma (2)

1. Then we have

$$\Pr\{x \in S_{kn} : f(x) > \mathrm{med}(f) + 4C/k\} \leq C\exp\{-(ct)n\}$$

2. Therefore,

$$\Pr\left\{x \in S_{kn} : f(x) > \frac{5t}{k}\right\} \leq C\exp\{-(ct)n\}$$

   Here, it is crucial to have the large deviation bound proportional to e^n and moreover we can adjust the speed of decay via the constant t.

3. In the previous proofs of additivity violation, the function $f(\cdot)$ was essentially modified somehow to have nice Lipschitz constants. This idea was nicely written in [Aubrun, Szarek and Werner]. However, we do not need to do it for our problem, as we have seen.

Setting a net on the domain

1. Our problem is to get an upper-bound for

$$\max_{x \in \tilde{E}} f(x)$$

where $S_n \cong \tilde{E} \hookrightarrow S_{kn}$.

2. Set up an ϵ-net on S_n, denoted by N_ϵ, such that

$$\forall x \in S_n \quad \exists y \in N_\epsilon \subset S_n \quad \text{dist}(x, y) \leq \epsilon$$

3. We can choose N_ϵ such that

$$|N_\epsilon| \leq \left(1 + \frac{2}{\epsilon}\right)^{2n}$$

Indeed, choose N_ϵ as a minimal net and then we have

$$|N_\epsilon| \left(\frac{\epsilon}{2}\right)^{2n} \leq \left(1 + \frac{\epsilon}{2}\right)^{2n}$$

where we compare volumes in $\mathbb{R}^{2n}$.

Estimate on the net

1. Fix $N_\epsilon \subset \tilde{E} \subset S_{kn}$ and we have by union bound method that

$$\begin{aligned}\Pr_{U\in\mathcal{U}(kn)}\left\{\max_{x\in UN_\epsilon} f(x) > \frac{5t}{k}\right\} &\leq |N_\epsilon| \Pr_{x\in S_{kn}}\left\{f(x) > \frac{5t}{k}\right\} \\ &\leq \exp\left\{\left[2\log\left(1+\frac{2}{\epsilon}\right)\right]n\right\} \times C\exp\{-(ct)n\} \qquad (2)\end{aligned}$$

2. If we choose $t > 0$ large enough, the bound is smaller than one for large enough n. This means that there is U such that

$$\max_{x\in UN_\epsilon} f(x) \leq \frac{5t}{k}$$

3. We extend the above estimate, which is just on the net $UN_\epsilon \subset U\tilde{E}$, to that on the whole domain $U\tilde{E}$.

Estimate on the net is almost enough - finishing the proof

1. For all $z \in \tilde{E}$, there exist $x \in N_\epsilon$, $y \in \tilde{E}$ and $0 \leq \delta \leq \epsilon$, and
$$z = x + \delta y$$
2. Hence, by using the triangle inequality,
$$\begin{aligned} f(z) &= \|XX^* + \delta(XY^* + YX^*) + \delta^2 YY^* - \tilde{I}_k\|_2 \\ &\leq f(x) + \epsilon^2 f(y) + \epsilon(|\alpha| + |\beta|)(f(a) + f(b)) \end{aligned}$$
Here, $xy^* + yx^* = \alpha aa^* + \beta bb^*$ where $a, b \in \tilde{E}$, and
$$|\alpha| + |\beta| = \|xy^* + yx^*\|_1 \leq \|xy^*\|_1 + \|yx^*\|_1 = 2$$
3. Therefore,
$$\max_{z \in \tilde{E}} f(z) \leq \frac{1}{1 - \epsilon^2 - 2\epsilon} \cdot \max_{x \in UN_\epsilon} f(x) \leq \frac{1}{1 - \epsilon^2 - 2\epsilon} \cdot \frac{5t}{k}$$
and setting, for example, $\epsilon = \frac{1}{4}$ gives a desired bound.