SCIENCE CHINA
Information Sciences

• RESEARCH PAPER •

Special Topic: Quantum Information

Supplemental material: Towards the ultimate limits of quantum channel discrimination and quantum communication

Kun FANG[1*], Gilad GOUR[2*] & Xin WANG[3*]

[1]School of Data Science, The Chinese University of Hong Kong, Shenzhen, Guangdong, 518172, China
[2]Technion - Israel Institute of Technology, Faculty of Mathematics, Haifa 3200003, Israel
[3]Thrust of Artificial Intelligence, The Hong Kong University of Science and Technology (Guangzhou), Guangdong, China

Appendix A Attempt to solve the strong converse property

An initial attempt to prove the exponentially strong converse property for two general channels, $\mathcal{I}\otimes\mathcal{N}$ and $\mathcal{I}\otimes\mathcal{M}$, was presented in the first arXiv submission (arXiv:2110.14842v1). However, a flaw has been identified in the original proof, rendering it invalid and leaving the problem unresolved. Nevertheless, several preliminary results from our initial analysis remain valid and may hold independent interest. These details are included in this appendix, and we hope they will contribute to resolving the problem in future studies.

The first lemma is an analog of a result of Ogawa and Nagaoka [1] that was originally used to establish the strong converse of quantum Stein's lemma. A similar result was proved by Brandão and Plenio for tensor product states [2]. Here we extend it further to permutation-symmetric states.

Lemma A1. Let $\mu\in\mathbb{R}$ and $\rho_n,\sigma_n\in\mathfrak{D}(A^n)$ be symmetric under permutations of the n subsystems such that $\operatorname{supp}(\rho_n)\subseteq\operatorname{supp}(\sigma_n)$. Then, for any $r\in\mathbb{R}$ and $s\in[0,1]$ the following relation holds

$$\operatorname{Tr}\left(\rho_n-2^{\mu n}\sigma_n\right)_+\leqslant 2^{-nrs+\log\operatorname{Tr}[\rho_n^{1+s}]}+2^{-ns(\mu-r)+s|A|\log(1+n)+\log\operatorname{Tr}[\rho_n\sigma_n^{-s}]}\,. \tag{A1}$$

Proof. Let Π be the projection to the positive part of $\rho_n-2^{\mu n}\sigma_n$ and $\Pi=\sum_{x=1}^{|A|^n}a_x\Pi_x$ be a decomposition of Π into orthogonal rank-one projectors, where $a_x\in\{0,1\}$ and $\sum_x\Pi_x=I_{A^n}$ (i.e. the set $\{\Pi_x\}$ forms a von-Neumann rank-one projective measurement). In general, this decomposition of Π is not unique, and the precise choice of $\{\Pi_x\}$ will be determined later on in the proof. Finally, denote by $p_x:=\operatorname{Tr}[\rho_n\Pi_x]$, $q_x:=\operatorname{Tr}[\sigma_n\Pi_x]$ (note that p_x and q_x depends on n), and let $\mathfrak{I}$ be the set of all x for which $p_x>2^{\mu n}q_x$. Since $(\rho_n-2^{n\mu}\sigma_n)_+=\Pi\left(\rho_n-2^{n\mu}\sigma_n\right)\Pi$, we have

$$\operatorname{Tr}\left(\rho_n-2^{n\mu}\sigma_n\right)_+=\sum_x a_x\left(p_x-2^{\mu n}q_x\right)\leqslant\sum_{x\in\mathfrak{I}}\left(p_x-2^{\mu n}q_x\right)\leqslant\sum_{x\in\mathfrak{I}}p_x=\Pr(\mathfrak{I})\,, \tag{A2}$$

where $\Pr(\mathfrak{I})$ is the probability of the set $\mathfrak{I}$ with respect to the probability distribution $\{p_x\}$. Note that the set $\mathfrak{I}$ can be written as

$$\mathfrak{I}=\left\{x\ :\ \frac{1}{n}\log p_x>\mu+\frac{1}{n}\log q_x\right\}\,. \tag{A3}$$

We would like to replace the set $\mathfrak{I}$ with two sets: one depends solely on p_x, and the other only on q_x. This can be done in the following way. For any $r\in\mathbb{R}$ define the two sets

$$\mathfrak{I}^{(1)}:=\left\{x\ :\ \frac{1}{n}\log p_x\geqslant r\right\}\quad\text{and}\quad\mathfrak{I}^{(2)}:=\left\{x\ :\ \frac{1}{n}\log q_x\leqslant r-\mu\right\}\,. \tag{A4}$$

* Corresponding author (email: kunfang@cuhk.edu.cn, giladgour@technion.ac.il, felixxinwang@hkust-gz.edu.cn)

Note that if $x \in \mathfrak{I}$ then either $x \in \mathfrak{I}^{(1)}$ or $x \in \mathfrak{I}^{(2)}$. We therefore conclude that

$$\operatorname{Tr}\left(\rho_n - 2^{n\mu}\sigma_n\right)_+ \leqslant \Pr\left(\mathfrak{I}^{(1)}\right) + \Pr\left(\mathfrak{I}^{(2)}\right) . \tag{A5}$$

From Cramér's theorem [3] it follows that

$$-\log \Pr\left(\mathfrak{I}^{(1)}\right) \geqslant \sup_{s\in[0,1]} \left\{ nsr - \log \sum_x p_x^{1+s} \right\} \tag{A6}$$

$$-\log \Pr\left(\mathfrak{I}^{(2)}\right) \geqslant \sup_{s\in[0,1]} \left\{ ns(\mu - r) - \log \sum_x p_x q_x^{-s} \right\} . \tag{A7}$$

We first bound (A6) in terms of ρ_n. For this purpose, let $\Delta \in \mathrm{CPTP}(A^n \to A^n)$ be the completely dephasing map (also a pinching map) $\Delta(\omega) \coloneqq \sum_x \Pi_x \omega \Pi_x$ defined on all $\omega \in \mathfrak{D}(A^n)$. Then, the density matrix $\Delta(\rho_n)$ is diagonal (in the basis that the operators $\{\Pi_x\}$ project to) with components $\{p_x\}$ on its diagonal. Hence, denoting by $\pi_{A^n} \coloneqq I_{A^n}/|A|^n$ the completely mixed state in $\mathfrak{D}(A^n)$, and by $\alpha \coloneqq 1+s$, we get by direct calculation that

$$-\log \sum_x p_x^{1+s} = n(\alpha-1)\log|A| - (\alpha-1)\overline{D}_\alpha\left(\Delta(\rho_n)\middle\|\pi_{A^n}\right) . \tag{A8}$$

Since $\overline{D}_\alpha\left(\Delta(\rho_n)\middle\|\pi_{A^n}\right) = \overline{D}_\alpha\left(\Delta(\rho_n)\middle\|\Delta(\pi_{A^n})\right) \leqslant \overline{D}_\alpha\left(\rho_n\middle\|\pi_{A^n}\right)$, we get

$$-\log \sum_x p_x^{1+s} \geqslant n(\alpha-1)\log|A| - (\alpha-1)\overline{D}_\alpha\left(\rho_n\middle\|\pi_{A^n}\right) = -\log \operatorname{Tr}\left[\rho_n^\alpha\right] = -\log \operatorname{Tr}\left[\rho_n^{1+s}\right] . \tag{A9}$$

Together with (A6), this gives the first term on the r.h.s. of (A1).

For the second term, observe that

$$\sum_x p_x q_x^{-s} = \operatorname{Tr}\left[\Delta\left(\rho_n\right)\left(\Delta\left(\sigma_n\right)\right)^{-s}\right] = \operatorname{Tr}\left[\rho_n\left(\Delta\left(\sigma_n\right)\right)^{-s}\right] . \tag{A10}$$

We now estimate this term by utilizing the symmetry of ρ_n and σ_n. Since ρ_n and σ_n are symmetric under permutations of the n subsystems they can be expressed as

$$\rho_n = \bigoplus_{\lambda\in\mathrm{Irr}(\mathcal{S}_n)} I^{B_\lambda} \otimes \rho_\lambda^{C_\lambda} \quad \text{and} \quad \sigma_n = \bigoplus_{\lambda\in\mathrm{Irr}(\mathcal{S}_n)} I^{B_\lambda} \otimes \sigma_\lambda^{C_\lambda} \tag{A11}$$

where λ represents an irrep of the natural representation of the permutation group $\mathcal{S}_n$ on A^n, and $\rho_\lambda, \sigma_\lambda \geqslant 0$. We therefore have

$$\rho_n - 2^{\mu n}\sigma_n = \bigoplus_{\lambda\in\mathrm{Irr}(\mathcal{S}_n)} I^{B_\lambda} \otimes \left(\rho_\lambda^{C_\lambda} - 2^{\mu n}\sigma_\lambda^{C_\lambda}\right) \tag{A12}$$

The condition $\mathrm{supp}(\rho_n) \subseteq \mathrm{supp}(\sigma_n)$ implies that without loss of generality we can assume that $\sigma_n > 0$ (otherwise we can restrict our consideration to the subspace of $\mathrm{supp}(\sigma_n)$ and embed ρ_n in this space). Therefore, under this assumption we have that each $\sigma_\lambda > 0$. Let P_λ be the projector to the support of $(\rho_\lambda - 2^{\mu n}\sigma_\lambda)_+$, and let $P_\lambda \coloneqq \sum_{j=1}^{|C_\lambda|^n} a_{\lambda,j} P_{\lambda,j}$ be a decomposition of P_λ into orthogonal rank-one projectors, where $a_{\lambda,j} \in \{0,1\}$ and $\sum_j P_{\lambda,j} = I^{C_\lambda}$. Moreover, for each $\lambda \in \mathrm{Irr}(\mathcal{S}_n)$ decompose $I^{B_\lambda} \coloneqq \sum_{k=1}^{|B_\lambda|} |\psi_{\lambda,k}\rangle\langle\psi_{\lambda,k}|^{B_\lambda}$, where $\{|\psi_{\lambda,k}\rangle\}_k$ forms an orthonormal basis of B_λ. Finally, we denote by $x \coloneqq \{\lambda, j, k\}$ and take $\Pi_x \coloneqq |\psi_{\lambda,k}\rangle\langle\psi_{\lambda,k}|^{B_\lambda} \otimes P_{\lambda,j}^{C_\lambda}$. With this choice of Π_x we get that

$$\begin{aligned}
\Delta(\sigma_n) &= \bigoplus_{\lambda\in\mathrm{Irr}(\mathcal{S}_n)} \sum_{j,k} |\psi_{\lambda,k}\rangle\langle\psi_{\lambda,k}|^{B_\lambda} \otimes P_{\lambda,j}^{C_\lambda}\sigma_\lambda^{C_\lambda}P_{\lambda,j}^{C_\lambda} \\
&= \bigoplus_{\lambda\in\mathrm{Irr}(\mathcal{S}_n)} I^{B_\lambda} \otimes \sum_j P_{\lambda,j}^{C_\lambda}\sigma_\lambda^{C_\lambda}P_{\lambda,j}^{C_\lambda} \\
&= \bigoplus_{\lambda\in\mathrm{Irr}(\mathcal{S}_n)} I^{B_\lambda} \otimes \Delta_\lambda^{C_\lambda\to C_\lambda}\left(\sigma_\lambda^{C_\lambda}\right)
\end{aligned} \tag{A13}$$

where each $\Delta_\lambda(\cdot) \coloneqq \sum_j P_{\lambda,j}(\cdot)P_{\lambda,j}$ is a completely dephasing map in $\mathrm{CPTP}(C_\lambda \to C_\lambda)$. Therefore,

$$\sum_x p_x q_x^{-s} = \operatorname{Tr}\left[\rho_n\left(\Delta\left(\sigma_n\right)\right)^{-s}\right] = \sum_{\lambda\in\mathrm{Irr}(\mathcal{S}_n)} |B_\lambda| \operatorname{Tr}\left[\rho_\lambda\left(\Delta_\lambda(\sigma_\lambda)\right)^{-s}\right] . \tag{A14}$$

From the pinching inequality, for each $\lambda \in \mathrm{Irr}(\mathcal{S}_n)$ we have $\Delta_\lambda(\sigma_\lambda) \geqslant \frac{1}{|C_\lambda|}\sigma_\lambda$. Moreover, since the function $r \mapsto r^\alpha$ is operator anti-monotone for $\alpha \in [-1,0]$ we get that

$$\left(\Delta_\lambda(\sigma_\lambda)\right)^{-s} \leqslant \left(\frac{1}{|C_\lambda|}\sigma_\lambda\right)^{-s} . \tag{A15}$$

Substituting this into (A14) gives

$$\sum_x p_x q_x^{-s} \leqslant \sum_{\lambda \in \mathrm{Irr}(\mathcal{S}_n)} |C_\lambda|^s |B_\lambda| \operatorname{Tr}\left[\rho_\lambda \sigma_\lambda^{-s}\right] . \tag{A16}$$

Now, since C_λ can be viewed as a subspace of $\mathrm{Sym}^n(A)$, its dimension cannot exceed that of $\mathrm{Sym}^n(A)$ which itself is bounded by $(n+1)^{|A|}$. We therefore conclude that

$$\sum_x p_x q_x^{-s} \leqslant (n+1)^{s|A|} \sum_{\lambda \in \mathrm{Irr}(\mathcal{S}_n)} |B_\lambda| \operatorname{Tr}\left[\rho_\lambda \sigma_\lambda^{-s}\right] = (n+1)^{s|A|} \operatorname{Tr}\left[\rho_n \sigma_n^{-s}\right] . \tag{A17}$$

Together with (A7), this gives the second term on the r.h.s. of (A1).

The next lemma shows that the eigenvalues of the output from n use of a positive definite channel $\mathcal{N} > 0$ (i.e., its Choi matrix is a positive definite operator) are uniformly bounded by an exponential factor.

Lemma A2. Let $\mathcal{N} \in \mathrm{CPTP}(A \to B)$ and $\mathcal{N} > 0$. Then, there exists $b \in (0,1)$ such that for any $n \in \mathbb{N}$

$$\max_{\rho \in \mathfrak{D}(R^n A^n)} \left\|\mathcal{N}^{\otimes n}\left(\rho_{R^n A^n}\right)\right\|_\infty \leqslant b^n . \tag{A18}$$

Proof. Since $\mathcal{N} > 0$ we have its Choi matrix $J_{\mathcal{N}} > 0$. Then there exists $\tau \in \mathfrak{D}(B)$ with $\|\tau\|_\infty < 1$ (e.g. the maximally mixed state) and its associated replacer channel $\mathcal{R}_\tau$ such that $tJ_{\mathcal{N}} > J_{\mathcal{R}_\tau}$ for some $t \in (0,\infty)$. Equivalently, we have $t\mathcal{N} > \mathcal{R}_\tau$. Set $\varepsilon := 1/t$ and then $\mathcal{M} := (\mathcal{N} - \varepsilon\mathcal{R}_\tau)/(1-\varepsilon) > 0$; in particular, $\mathcal{M} \in \mathrm{CPTP}(A \to B)$ and $\mathcal{N} = (1-\varepsilon)\mathcal{M} + \varepsilon\mathcal{R}_\tau$. Observe that

$$\mathcal{N}^{\otimes n} = \sum_{k=0}^n \binom{n}{k} (1-\varepsilon)^k \varepsilon^{n-k} \mathcal{F}_{n,k} \tag{A19}$$

where $\mathcal{F}_{n,k} \in \mathrm{CPTP}(A^n \to B^n)$ is a uniform convex combination of $\binom{n}{k}$ channels all having the form $\mathcal{M}^{\otimes k} \otimes \mathcal{R}_\tau^{\otimes n-k}$ up to permutations of the n channels. Now, observe that

$$\left\|\mathcal{M}^{\otimes k} \otimes \mathcal{R}_\tau^{\otimes n-k}\left(\rho_{R^n A^n}\right)\right\|_\infty = \left\|\mathcal{M}^{\otimes k}\left(\rho_{R^n A^k}\right) \otimes \tau^{\otimes n-k}\right\|_\infty \leqslant \left\|\tau^{\otimes n-k}\right\|_\infty = \|\tau\|_\infty^{n-k} . \tag{A20}$$

Note that the order that $\mathcal{N}$ and $\mathcal{R}_\tau$ appear in the equation above does not effect this upper bound. Therefore, since $\mathcal{F}_{n,k}$ is a convex combination of such channels we conclude that also

$$\left\|\mathcal{F}_{n,k}\left(\rho_{R^n A^n}\right)\right\|_\infty \leqslant \|\tau\|_\infty^{n-k} . \tag{A21}$$

Hence, for any $\rho \in \mathfrak{D}(R^n A^n)$

$$\begin{aligned} \left\|\mathcal{N}^{\otimes n}\left(\rho_{R^n A^n}\right)\right\|_\infty &\leqslant \sum_{k=0}^n \binom{n}{k}(1-\varepsilon)^k \varepsilon^{n-k} \left\|\mathcal{F}_{n,k}\left(\rho^{R^n A^n}\right)\right\|_\infty \\ &\leqslant \sum_{k=0}^n \binom{n}{k}(1-\varepsilon)^k \varepsilon^{n-k} \|\tau\|_\infty^{n-k} \\ &= \left(1 - \varepsilon + \|\tau\|_\infty \varepsilon\right)^n . \end{aligned} \tag{A22}$$

The proof is completed by taking $b := 1 - \varepsilon + \|\tau\|_\infty \varepsilon$ which is clearly in $(0,1)$.

The next lemma shows that by utilizing the permutation symmetry of tensor product channels we can restrict the optimal input states in the discrimination strategies to be symmetric states. This reduces the problem from the most general form to a particular one that can be tackled more easily. A general result is given in [4, Proposition II.4]. Here, we give an alternative proof for the hypothesis testing relative entropy.

Lemma A3. Let $\mathcal{N}, \mathcal{M} \in \mathrm{CPTP}(A \to B)$. For any $n \in \mathbb{N}$ there exists a pure state $|\varphi\rangle \in \mathrm{Sym}^n(\tilde{R}A)$ such that

$$\max_{\psi \in \mathfrak{D}(R^n A^n)} D_H^\varepsilon\left(\mathcal{N}^{\otimes n}(\psi_{R^n A^n}) \middle\| \mathcal{M}^{\otimes n}(\psi_{R^n A^n})\right) = D_H^\varepsilon\left(\mathcal{N}^{\otimes n}(\varphi_{\tilde{R}^n A^n}) \middle\| \mathcal{M}^{\otimes n}(\varphi_{\tilde{R}^n A^n})\right) . \tag{A23}$$

Proof. First recall a variational expression of the hypothesis testing relative entropy [5, Eq.(2)]

$$D_H^\varepsilon(\rho\|\sigma) = -\log \max_{t \geqslant 0} \left\{(1-\varepsilon)t - \operatorname{Tr}(t\rho - \sigma)_+\right\} . \tag{A24}$$

Therefore, we have

$$D_H^\varepsilon\left(\mathcal{N}^{\otimes n} \middle\| \mathcal{M}^{\otimes n}\right) = -\log \max_{t \geqslant 0} \left\{(1-\varepsilon)t - \operatorname{Tr}\left(t\mathcal{N}^{\otimes n}(\psi_{R^n A^n}) - \mathcal{M}^{\otimes n}(\psi_{R^n A^n})\right)_+\right\} , \tag{A25}$$

for some state $\psi_{R^nA^n} \in \mathfrak{D}(R^nA^n)$. Let

$$\omega_{XR^nA^n} = \frac{1}{n!}\sum_{\pi\in\mathcal{S}_n} |\pi\rangle\langle\pi|_X \otimes P_\pi \psi_{R^nA^n} P_\pi^* \tag{A26}$$

where X is a 'flag' system of dimension $|X| = n!$. By construction, the marginal state $\omega_{R^nA^n}$ is symmetric under permutations (i.e. has support on $\mathrm{Sym}^n(RA)$), so there exists a symmetric purification of $\omega_{R^nA^n}$ which we denote by $\varphi_{C^nR^nA^n}$, where $C \cong RA$ [6, Lemma 4.2.2]. Let $\omega_{DXR^nA^n}$ be a purification of $\omega_{XR^nA^n}$ and thus also a purification of $\omega_{R^nA^n}$. Since all purifications of a density matrix are related via isometries, there exists an isometry $V_{C^n\to DX}$ such that

$$\omega_{DXR^nA^n} = (V_{C^n\to DX})\varphi_{C^nR^nA^n}(V_{C^n\to DX})^\dagger. \tag{A27}$$

Taking a partial trace of the system D on both sides gives

$$\omega_{XR^nA^n} = \mathcal{E}_{C^n\to X}\left(\varphi_{C^nR^nA^n}\right) \,, \tag{A28}$$

where $\mathcal{E}(\cdot) = \mathrm{Tr}_D\, V(\cdot)V^\dagger \in \mathrm{CPTP}(C^n \to X)$. Let $\tilde{R} \coloneqq CR$, then $|\varphi_{\tilde{R}^nA^n}\rangle \in \mathrm{Sym}^n(\tilde{R}A)$ and

$$\begin{aligned}
&\mathrm{Tr}\left(t\mathcal{N}^{\otimes n}(\varphi_{\tilde{R}^nA^n}) - \mathcal{M}^{\otimes n}(\varphi_{\tilde{R}^nA^n})\right)_+ \\
&\geqslant \mathrm{Tr}\left(t\mathcal{N}^{\otimes n}(\omega_{XR^nA^n}) - \mathcal{M}^{\otimes n}(\omega_{XR^nA^n})\right)_+ && \text{(A29)}\\
&= \frac{1}{n!}\sum_{\pi\in\mathcal{S}_n} \mathrm{Tr}\left(t\mathcal{N}^{\otimes n}\left(P_\pi\psi_{R^nA^n}P_\pi^*\right) - \mathcal{M}^{\otimes n}(P_\pi\psi_{R^nA^n}P_\pi^*)\right)_+ && \text{(A30)}\\
&= \mathrm{Tr}\left(t\mathcal{N}^{\otimes n}(\psi_{R^nA^n}) - \mathcal{M}^{\otimes n}(\psi_{R^nA^n})\right)_+ && \text{(A31)}
\end{aligned}$$

where the first inequality follows from the data processing inequality of $\mathrm{Tr}(\cdot)_+$ [1)], the first equality follows from the block diagonal structure of $\mathcal{N}^{\otimes n}(\omega_{XR^nA^n}) - t\mathcal{M}^{\otimes n}(\omega_{XR^nA^n})$, the second equality follows because $\mathrm{Tr}(\cdot)_+$ is unitary invariant and $\mathcal{N}^{\otimes n}, \mathcal{M}^{\otimes n}$ commute with permutations. Together with (A25), we can conclude that

$$\begin{aligned}
D_H^\varepsilon\left(\mathcal{N}^{\otimes n}\middle\|\mathcal{M}^{\otimes n}\right) &\leqslant -\log\max_{t\in\mathbb{R}}\left\{(1-\varepsilon)t - \mathrm{Tr}\left(t\mathcal{N}^{\otimes n}(\varphi_{\tilde{R}^nA^n}) - \mathcal{M}^{\otimes n}(\varphi_{\tilde{R}^nA^n})\right)_+\right\} \\
&= D_H^\varepsilon\left(\mathcal{N}^{\otimes n}\left(\varphi_{\tilde{R}^nA^n}\right)\middle\|\mathcal{M}^{\otimes n}\left(\varphi_{\tilde{R}^nA^n}\right)\right).
\end{aligned} \tag{A32}$$

This completes the proof.

Remark 1. We say that a quantum divergence, $\boldsymbol{D}$, satisfies *the direct sum property* if there exists a one-to-one function $f : \mathbb{R}_+ \to \mathbb{R}_+$ such that for any pair of cq-states $\rho, \sigma \in \mathfrak{D}(XA)$ of the form $\rho_{XA} \coloneqq \sum_x p_x|x\rangle\langle x|^X \otimes \rho_x^A$ and $\sigma^{XA} \coloneqq \sum_x p_x|x\rangle\langle x|^X \otimes \sigma_x^A$ where $\{p_x\}$ is a probability distribution and $\rho_x, \sigma_x \in \mathfrak{D}(A)$, we have $f^{-1}\left(\boldsymbol{D}\left(\rho_{XA}\middle\|\sigma_{XA}\right)\right) = \sum_x p_x f^{-1}\left(\boldsymbol{D}\left(\rho_x^A\middle\|\sigma_x^A\right)\right)$. The direct sum property is essentially equivalent to the *general mean property* used by Rényi and Müller-Lennert et al. for its generalization to the quantum case and holds for almost all the quantum divergences studied in the literature. Following a similar proof, we can show that Lemma A3 holds for any quantum divergence with the direct sum property.

References

1 Ogawa T, Nagaoka H. Strong converse and Stein's lemma in quantum hypothesis testing. IEEE Transactions on Information Theory, 2000, 46: 2428–2433

2 Brandao F G, Plenio M B. A generalization of quantum stein's lemma. Communications in Mathematical Physics, 2010, 295: 791–828

3 Dembo A, Zeitouni O. Large deviations techniques and applications. Springer-Verlag, 1998

4 Leditzky F, Kaur E, Datta N, et al. Approaches for approximate additivity of the holevo information of quantum channels. Physical Review A, 2018, 97: 012332

5 Buscemi F, Gour G. Quantum relative lorenz curves. Physical Review A, 2017, 95: 012110

6 Renner R. Security of qkd. Ph.D. thesis, PhD thesis, ETH, 2005, quant-ph/0512258, 2005

[1)] This can be easily seen from the equation $\mathrm{Tr}(X)_+ = (\|X\|_1 + \mathrm{Tr}\,X)/2$ and the data processing inequality of trace norm.