Moderate deviation analysis for c-q channels and quantum hypothesis testing

Joint work with Vincent Y.F. Tan (NUS)
and Marco Tomamichel (USyd/UTS)
arXiv:1701.03114 (to appear in CMP)

Christopher T. Chubb
Centre for Engineered Quantum Systems
University of Sydney

Freie Universität Berlin 2017
Berlin, Germany

physics.usyd.edu.au/~cchubb/

THE UNIVERSITY OF SYDNEY

eQus

Classical communication over a quantum channel

We are going to consider transmitting classical information over a quantum channel.

For channel $\mathcal{W}$, a (n, R, ϵ)-code is an encoder E and decoding POVM $\{D_i\}$ such that

$$\frac{1}{2^{nR}} \sum_{m=1}^{2^{nR}} \mathrm{Tr}\left[\mathcal{W}^{\otimes n}\left(\otimes_{i=1}^{n} E_i(m)\right) D_m\right] \geq 1 - \epsilon$$

We will be concerned with the trade-off between the block-length n, the rate R, and the error probability ϵ. We define the optimal rate/error probability as

$$R^*(\mathcal{W}; n, \epsilon) := \max\left\{R \mid \exists (n, R, \epsilon)\text{-code}\right\},$$
$$\epsilon^*(\mathcal{W}; n, \epsilon) := \min\left\{\epsilon \mid \exists (n, R, \epsilon)\text{-code}\right\}.$$

Asymptotics

For a constant error probability ϵ, the Strong Converse Theorem[1] tells us the rate approaches a constant known as the capacity

$$\lim_{n\to\infty} R^*(\mathcal{W}; n, \epsilon) = C(\mathcal{W}).$$

Equivalently this means that the error probability must to go 0 to 1 either side of the capacity

$$\lim_{n\to\infty} \epsilon^*(\mathcal{W}; n, R) = \begin{cases} 0 & : R < C \\ 1 & : R > C \end{cases}$$

This tells us we can have either $R \to C$ OR $\epsilon \to 0$.

How fast are these convergences? Can we do both?

[1] Winter 1999

Asymptotics

For a constant error probability ϵ, the Strong Converse Theorem[1] tells us the rate approaches a constant known as the capacity

$$\lim_{n\to\infty} R^*(\mathcal{W}; n, \epsilon) = C(\mathcal{W}).$$

Equivalently this means that the error probability must to go 0 to 1 either side of the capacity

$$\lim_{n\to\infty} \epsilon^*(\mathcal{W}; n, R) = \begin{cases} 0 & : R < C \\ 1 & : R > C \end{cases}$$

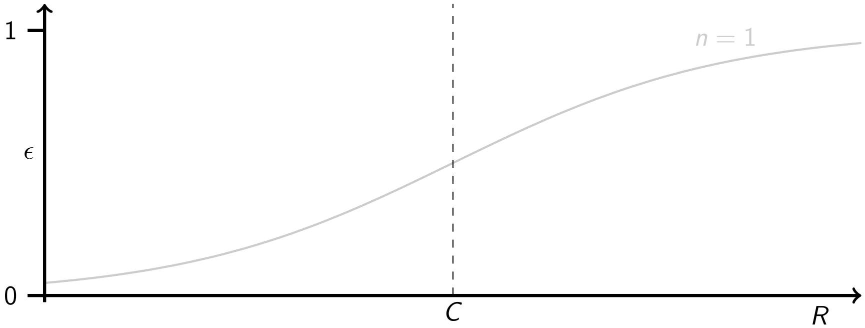
This tells us we can have either $R \to C$ OR $\epsilon \to 0$.

How fast are these convergences? Can we do both?

[1] Winter 1999

Small and large deviations

How fast are the convergences $R \to C$ or $\epsilon \to 0$ as $n \to \infty$?



Small deviation (Tomamichel & Tan 2015)

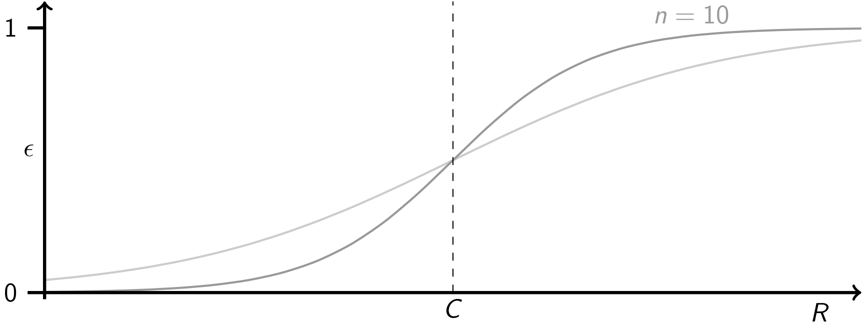
$$R^*(n,\epsilon) = C + \sqrt{\frac{V}{n}}\Phi^{-1}(\epsilon) + o\left(\frac{1}{\sqrt{n}}\right) \quad \epsilon \in (0, \tfrac{1}{2})$$

Large deviation (Partial progress)

$$\ln \epsilon^*(n,R) = -n \cdot E(R) + o(n) \quad R < C$$

Small and large deviations

How fast are the convergences $R \to C$ or $\epsilon \to 0$ as $n \to \infty$?



Small deviation (Tomamichel & Tan 2015)

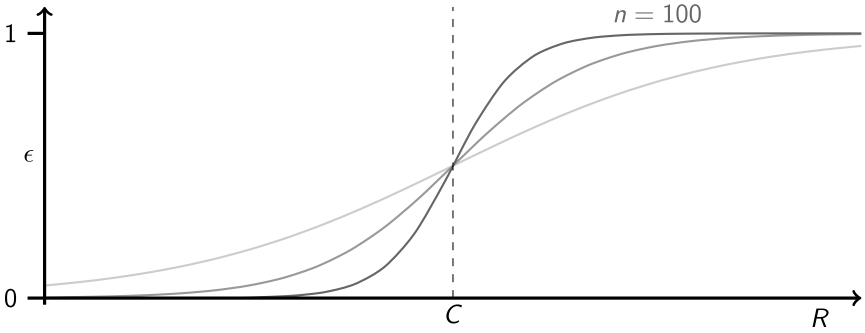
$$R^*(n,\epsilon) = C + \sqrt{\frac{V}{n}}\Phi^{-1}(\epsilon) + o\left(\frac{1}{\sqrt{n}}\right) \quad \epsilon \in (0, \tfrac{1}{2})$$

Large deviation (Partial progress)

$$\ln \epsilon^*(n,R) = -n \cdot E(R) + o(n) \quad R < C$$

Small and large deviations

How fast are the convergences $R \to C$ or $\epsilon \to 0$ as $n \to \infty$?



Small deviation (Tomamichel & Tan 2015)

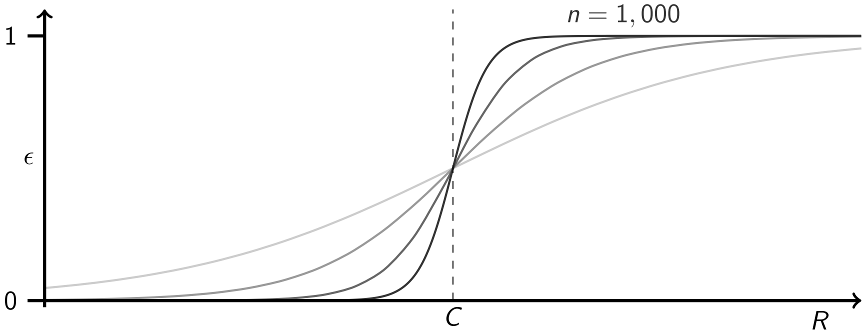
$$R^*(n,\epsilon) = C + \sqrt{\frac{V}{n}}\Phi^{-1}(\epsilon) + o\left(\frac{1}{\sqrt{n}}\right) \quad \epsilon \in (0, \tfrac{1}{2})$$

Large deviation (Partial progress)

$$\ln \epsilon^*(n,R) = -n \cdot E(R) + o(n) \quad R < C$$

Small and large deviations

How fast are the convergences $R \to C$ or $\epsilon \to 0$ as $n \to \infty$?



Small deviation (Tomamichel & Tan 2015)

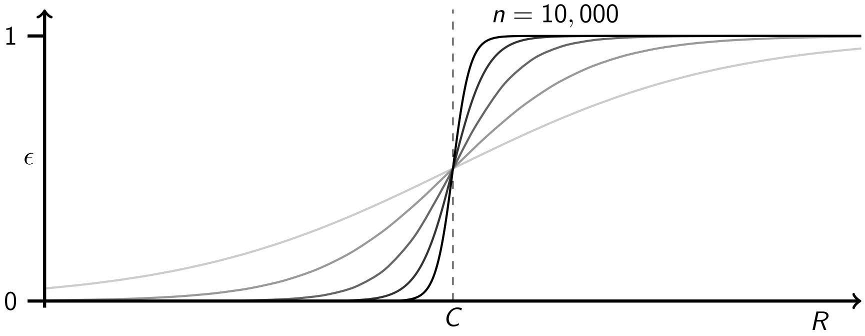
$R^*(n,\epsilon) = C + \sqrt{\frac{V}{n}}\Phi^{-1}(\epsilon) + o\left(\frac{1}{\sqrt{n}}\right) \quad \epsilon \in (0, \tfrac{1}{2})$

Large deviation (Partial progress)

$\ln \epsilon^*(n,R) = -n \cdot E(R) + o(n) \quad R < C$

Small and large deviations

How fast are the convergences $R \to C$ or $\epsilon \to 0$ as $n \to \infty$?



Small deviation (Tomamichel & Tan 2015)

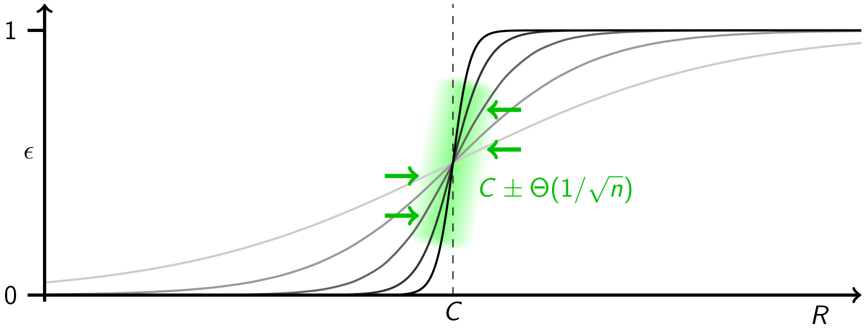
$$R^*(n,\epsilon) = C + \sqrt{\frac{V}{n}}\Phi^{-1}(\epsilon) + o\left(\frac{1}{\sqrt{n}}\right) \quad \epsilon \in (0, \tfrac{1}{2})$$

Large deviation (Partial progress)

$$\ln \epsilon^*(n,R) = -n \cdot E(R) + o(n) \quad R < C$$

Small and large deviations

How fast are the convergences $R \to C$ or $\epsilon \to 0$ as $n \to \infty$?



Small deviation (Tomamichel & Tan 2015)

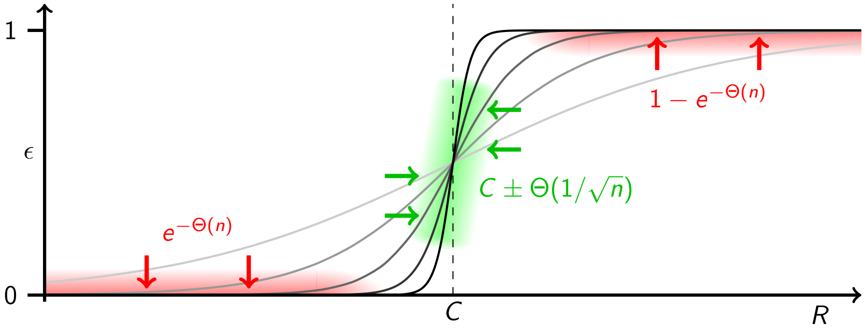
$$R^*(n,\epsilon) = C + \sqrt{\frac{V}{n}}\Phi^{-1}(\epsilon) + o\left(\frac{1}{\sqrt{n}}\right) \quad \epsilon \in (0, \tfrac{1}{2})$$

Large deviation (Partial progress)

$$\ln \epsilon^*(n, R) = -n \cdot E(R) + o(n) \quad R < C$$

Small and large deviations

How fast are the convergences $R \to C$ or $\epsilon \to 0$ as $n \to \infty$?



Small deviation (Tomamichel & Tan 2015)

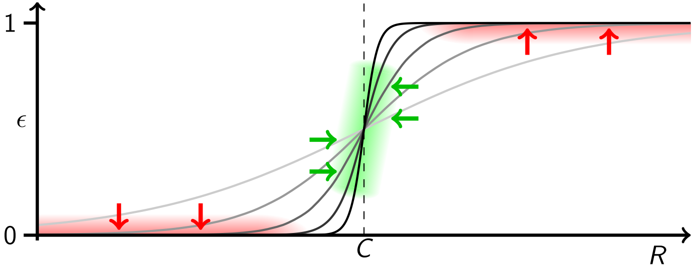
$$R^*(n,\epsilon) = C + \sqrt{\frac{V}{n}}\Phi^{-1}(\epsilon) + o\left(\frac{1}{\sqrt{n}}\right) \quad \epsilon \in (0, \tfrac{1}{2})$$

Large deviation (Partial progress)

$$\ln \epsilon^*(n,R) = -n \cdot E(R) + o(n) \quad R < C$$

Moderate deviations

What if we want $R \to C$ AND $\epsilon \to 0$?



Moderate deviation (This work, Cheng & Hsieh 2017)

For any $\{a_n\}$ such that $a_n \to 0$ and $\sqrt{n}a_n \to \infty$ we have

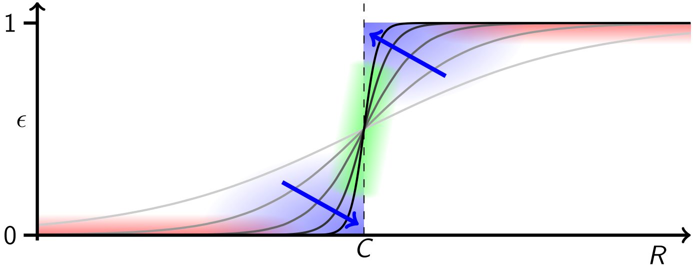
$$R^*(n, \epsilon_n) = C - \sqrt{2V}a_n + o(a_n) \quad \text{for} \quad \epsilon_n = e^{-na_n^2},$$

or equivalently

$$\ln \epsilon^*(n, R_n) = -\frac{na_n^2}{2V} + o(na_n^2) \quad \text{for} \quad R_n = C - a_n.$$

Moderate deviations

What if we want $R \to C$ AND $\epsilon \to 0$?



Moderate deviation (This work, Cheng & Hsieh 2017)

For any $\{a_n\}$ such that $a_n \to 0$ and $\sqrt{n}a_n \to \infty$ we have

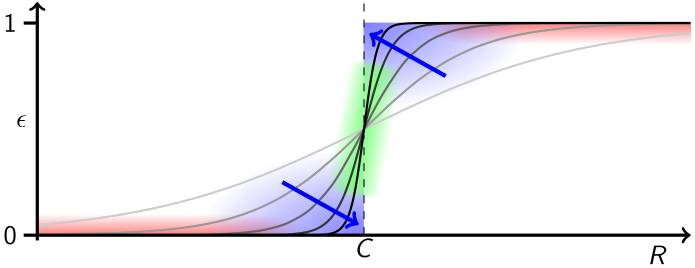
$$R^*(n, \epsilon_n) = C - \sqrt{2V}a_n + o(a_n) \quad \text{for} \quad \epsilon_n = e^{-na_n^2},$$

or equivalently

$$\ln \epsilon^*(n, R_n) = -\frac{na_n^2}{2V} + o(na_n^2) \quad \text{for} \quad R_n = C - a_n.$$

Moderate deviations

What if we want $R \to C$ AND $\epsilon \to 0$?



Moderate deviation (This work, Cheng & Hsieh 2017)

For any $\{a_n\}$ such that $a_n \to 0$ and $\sqrt{n} a_n \to \infty$ we have

$$R^*(n, \epsilon_n) = C - \sqrt{2V} a_n + o(a_n) \quad \text{for} \quad \epsilon_n = e^{-n a_n^2},$$

or equivalently

$$\ln \epsilon^*(n, R_n) = -\frac{n a_n^2}{2V} + o(n a_n^2) \quad \text{for} \quad R_n = C - a_n.$$

Context

	asymmetric binary hypothesis testing	channel coding	quantum hypothesis testing	classical-quantum channel coding
large deviation ($<$)	[16]	[17, 18]	[5, 19]	unknown[3]
moderate deviation ($<$)	[20]	[12, 13]	*this work*	*this work*
small deviation	[21]	[21–23]	[10, 11]	[7]
moderate deviation ($>$)	*this work*	*this work*	*this work*	*this work*
large deviation ($>$)	[24, 25]	[26, 27]	[28, 29]	[30]

Table 1: Exposition of related work on finite resource analysis of hypothesis testing and channel coding problems. The rows correspond to different parameter regimes, labelled by the deviation from the critical rate (i.e., the relative entropy for hypothesis testing and the capacity for channel coding problems).

Concentration inequalities

Take $\{X_i\}$ iid with $\mathbb{E}[X_i] = 0$ and $\mathrm{Var}[X_i] =: V$, and $\bar{X}_n := \frac{1}{n}\sum_{i=1}^n X_i$.

Asymptotic (Law of large numbers)

$$\lim_{n\to\infty} \Pr\left[\bar{X}_n \geq t\right] = \begin{cases} 1 & t < 0, \\ 0 & t > 0. \end{cases}$$

Small deviation (Berry-Esseen)

$$\Pr\left[\bar{X}_n \geq \tfrac{\epsilon}{\sqrt{n}}\right] = Q\left(\tfrac{\epsilon}{\sqrt{V}}\right) + \mathcal{O}\left(\tfrac{1}{\sqrt{n}}\right) \quad \epsilon \in (0,1)$$

Large deviation (Cramér)

$$\ln \Pr\left[\bar{X}_n \geq t\right] = -n \cdot I(t) + o(n) \quad t \geq 0$$

Moderate deviation

For any $\{a_n\}$ such that $a_n \to 0$ and $\sqrt{n}a_n \to \infty$

$$\ln \Pr\left[\bar{X}_n \geq a_n\right] = -\frac{na_n^2}{2V} + o(na_n^2).$$

Concentration inequalities

Take $\{X_i\}$ iid with $\mathbb{E}[X_i] = 0$ and $\mathrm{Var}[X_i] =: V$, and $\bar{X}_n := \frac{1}{n}\sum_{i=1}^n X_i$.

Asymptotic (Law of large numbers)

$$\lim_{n\to\infty} \Pr\left[\bar{X}_n \geq t\right] = \begin{cases} 1 & t < 0, \\ 0 & t > 0. \end{cases}$$

Small deviation (Berry-Esseen)

$$\Pr\left[\bar{X}_n \geq \tfrac{\epsilon}{\sqrt{n}}\right] = Q\left(\tfrac{\epsilon}{\sqrt{V}}\right) + \mathcal{O}\left(\tfrac{1}{\sqrt{n}}\right) \quad \epsilon \in (0,1)$$

Large deviation (Cramér)

$$\ln \Pr\left[\bar{X}_n \geq t\right] = -n \cdot I(t) + o(n) \quad t \geq 0$$

Moderate deviation

For any $\{a_n\}$ such that $a_n \to 0$ and $\sqrt{n}a_n \to \infty$

$$\ln \Pr\left[\bar{X}_n \geq a_n\right] = -\frac{na_n^2}{2V} + o(na_n^2).$$

Concentration inequalities

Take $\{X_i\}$ iid with $\mathbb{E}[X_i] = 0$ and $\mathrm{Var}[X_i] =: V$, and $\bar{X}_n := \frac{1}{n}\sum_{i=1}^n X_i$.

Asymptotic (Law of large numbers)

$$\lim_{n\to\infty} \Pr\left[\bar{X}_n \geq t\right] = \begin{cases} 1 & t < 0, \\ 0 & t > 0. \end{cases}$$

Small deviation (Berry-Esseen)

$$\Pr\left[\bar{X}_n \geq \tfrac{\epsilon}{\sqrt{n}}\right] = Q\left(\tfrac{\epsilon}{\sqrt{V}}\right) + \mathcal{O}\left(\tfrac{1}{\sqrt{n}}\right) \quad \epsilon \in (0,1)$$

Large deviation (Cramér)

$$\ln \Pr\left[\bar{X}_n \geq t\right] = -n \cdot I(t) + o(n) \quad t \geq 0$$

Moderate deviation

For any $\{a_n\}$ such that $a_n \to 0$ and $\sqrt{n}a_n \to \infty$

$$\ln \Pr\left[\bar{X}_n \geq a_n\right] = -\frac{na_n^2}{2V} + o(na_n^2).$$

Hypothesis testing

We want to test between two hypotheses, ρ and σ. For a binary POVM $\{Q, I - Q\}$, we define the type-I and type-II errors as

$$\alpha(Q; \rho, \sigma) := \operatorname{Tr}(I - Q)\rho, \qquad \beta(Q; \rho, \sigma) := \operatorname{Tr} Q\sigma,$$

and the ϵ-hypothesis-testing divergence

$$D_h^\epsilon(\rho\|\sigma) := -\log \min \left\{\beta(Q; \rho, \sigma) \,|\, \alpha(Q; \rho, \sigma) \leq \epsilon\right\}.$$

If we now consider testing between $\rho^{\otimes n}$ and $\sigma^{\otimes n}$, then the asymptotic behaviour is given by Quantum Stein's Lemma.

Asymptotics (Hiai & Petz 1991 / Ogawa & Nagaoka 1999)

$$\lim_{n\to\infty} \frac{1}{n} D_h^\epsilon(\rho^{\otimes n}\|\sigma^{\otimes n}) = D(\rho\|\sigma) \quad \text{where} \quad \epsilon \in (0,1),$$

or equivalently

$$\frac{1}{n} D_h^{\epsilon_n}(\rho^{\otimes n}\|\sigma^{\otimes n}) = R \quad \Longrightarrow \quad \lim_{n\to\infty} \epsilon_n = \begin{cases} 0 & : R < D(\rho\|\sigma), \\ 1 & : R > D(\rho\|\sigma). \end{cases}$$

Hypothesis testing

We want to test between two hypotheses, ρ and σ. For a binary POVM $\{Q, I-Q\}$, we define the type-I and type-II errors as

$$\alpha(Q;\rho,\sigma) := \operatorname{Tr}(I-Q)\rho, \qquad \beta(Q;\rho,\sigma) := \operatorname{Tr} Q\sigma,$$

and the ϵ-hypothesis-testing divergence

$$D_h^\epsilon(\rho\|\sigma) := -\log\min\left\{\beta(Q;\rho,\sigma)\,\middle|\,\alpha(Q;\rho,\sigma)\le\epsilon\right\}.$$

If we now consider testing between $\rho^{\otimes n}$ and $\sigma^{\otimes n}$, then the asymptotic behaviour is given by Quantum Stein's Lemma.

Asymptotics (Hiai & Petz 1991 / Ogawa & Nagaoka 1999)

$$\lim_{n\to\infty}\frac{1}{n}D_h^\epsilon(\rho^{\otimes n}\|\sigma^{\otimes n}) = D(\rho\|\sigma) \quad \text{where} \quad \epsilon\in(0,1),$$

or equivalently

$$\frac{1}{n}D_h^{\epsilon_n}(\rho^{\otimes n}\|\sigma^{\otimes n}) = R \quad\Longrightarrow\quad \lim_{n\to\infty}\epsilon_n = \begin{cases}0 & : R < D(\rho\|\sigma),\\ 1 & : R > D(\rho\|\sigma).\end{cases}$$

Deviation results for hypothesis testing

Small deviation (Tomamichel & Hayashi 2013, Li 2014)

$$\frac{1}{n}D_h^\epsilon(\rho^{\otimes n}\|\sigma^{\otimes n}) = D(\rho\|\sigma) + \sqrt{\frac{V(\rho\|\sigma)}{n}}\Phi^{-1}(\epsilon) + \mathcal{O}\left(\frac{\log n}{n}\right) \quad \text{for} \quad \epsilon \in (0,1).$$

Large deviation (Hayashi 2006, Nagaoka 2006)

$$\ln \epsilon_n = -n \cdot E(R) + o(n) \quad \text{for} \quad \frac{1}{n}D_h^{\epsilon_n}(\rho^{\otimes n}\|\sigma^{\otimes n}) = R < D(\rho\|\sigma).$$

Moderate deviation (This work, Cheng & Hsieh 2017)

For any $\{a_n\}$ such that $a_n \to 0$ and $\sqrt{n}a_n \to \infty$ we have

$$\frac{1}{n}D_h^{\epsilon_n}(\rho^{\otimes n}\|\sigma^{\otimes n}) = D(\rho\|\sigma) - \sqrt{2V(\rho\|\sigma)}a_n + o(a_n) \quad \text{for} \quad \epsilon_n = e^{-na_n^2},$$

or equivalently

$$\ln \epsilon_n = -\frac{na_n^2}{2V(\rho\|\sigma)} + o(na_n^2) \quad \text{for} \quad \frac{1}{n}D_h^{\epsilon_n}(\rho^{\otimes n}\|\sigma^{\otimes n}) = D(\rho\|\sigma) - a_n.$$

Deviation results for hypothesis testing

Small deviation (Tomamichel & Hayashi 2013, Li 2014)

$$\frac{1}{n}D_h^{\epsilon}(\rho^{\otimes n}\|\sigma^{\otimes n}) = D(\rho\|\sigma) + \sqrt{\frac{V(\rho\|\sigma)}{n}}\Phi^{-1}(\epsilon) + \mathcal{O}\left(\frac{\log n}{n}\right) \quad \text{for} \quad \epsilon \in (0,1).$$

Large deviation (Hayashi 2006, Nagaoka 2006)

$$\ln \epsilon_n = -n \cdot E(R) + o(n) \quad \text{for} \quad \frac{1}{n}D_h^{\epsilon_n}(\rho^{\otimes n}\|\sigma^{\otimes n}) = R < D(\rho\|\sigma).$$

Moderate deviation (This work, Cheng & Hsieh 2017)

For any $\{a_n\}$ such that $a_n \to 0$ and $\sqrt{n}a_n \to \infty$ we have

$$\frac{1}{n}D_h^{\epsilon_n}(\rho^{\otimes n}\|\sigma^{\otimes n}) = D(\rho\|\sigma) - \sqrt{2V(\rho\|\sigma)}a_n + o(a_n) \quad \text{for} \quad \epsilon_n = e^{-na_n^2},$$

or equivalently

$$\ln \epsilon_n = -\frac{na_n^2}{2V(\rho\|\sigma)} + o(na_n^2) \quad \text{for} \quad \frac{1}{n}D_h^{\epsilon_n}(\rho^{\otimes n}\|\sigma^{\otimes n}) = D(\rho\|\sigma) - a_n.$$

Reducing hyp. testing to concentration inequalities

To give a moderate deviation analysis of the HTD, we will use concentration bounds. First we see it is related to tail bounds of the Nussbaum-Szkoła distributions[2]

$$P^{\rho,\sigma}(a,b) := r_a |\langle\phi_a|\psi_b\rangle|^2 \quad \text{and} \quad Q^{\rho,\sigma}(a,b) := s_b |\langle\phi_a|\psi_b\rangle|^2,$$

where we have eigendecomposed our states $\rho := \sum_a r_a |\phi_a\rangle\langle\phi_a|$ and $\sigma := \sum_b s_b |\psi_b\rangle\langle\psi_b|$. These reproduce the first two moments of our states

$$D\left(P^{\rho,\sigma} \| Q^{\rho,\sigma}\right) = D(\rho\|\sigma) \qquad \text{and} \qquad V\left(P^{\rho,\sigma} \| Q^{\rho,\sigma}\right) = V(\rho\|\sigma).$$

Specifically for iid $Z_i = \log P^{\rho,\sigma}/Q^{\rho,\sigma}$ and $(a_i, b_i) \sim P^{\rho,\sigma}$, then[3]

$$\frac{1}{n} D_h^{\epsilon_n}\left(\rho^{\otimes n} \middle\| \sigma^{\otimes n}\right) \geq \sup\left\{ R \,\middle|\, \Pr\left[\sum_{i=1}^n Z_i\right] \leq \epsilon_n/2 \right\} - \mathcal{O}(\log 1/\epsilon_n),$$

$$\frac{1}{n} D_h^{\epsilon_n}\left(\rho^{\otimes n} \middle\| \sigma^{\otimes n}\right) \leq \sup\left\{ R \,\middle|\, \Pr\left[\sum_{i=1}^n Z_i\right] \leq 2\epsilon_n \right\} + \mathcal{O}(\log 1/\epsilon_n).$$

[2] Nussbaum & Szkoła 2009.

[3] Tomamichel & Hayashi 2013

Reducing hyp. testing to concentration inequalities

To give a moderate deviation analysis of the HTD, we will use concentration bounds. First we see it is related to tail bounds of the Nussbaum-Szkoła distributions[2]

$$P^{\rho,\sigma}(a,b) := r_a |\langle \phi_a | \psi_b \rangle|^2 \quad \text{and} \quad Q^{\rho,\sigma}(a,b) := s_b |\langle \phi_a | \psi_b \rangle|^2,$$

where we have eigendecomposed our states $\rho := \sum_a r_a |\phi_a\rangle\langle\phi_a|$ and $\sigma := \sum_b s_b |\psi_b\rangle\langle\psi_b|$. These reproduce the first two moments of our states

$$D\left(P^{\rho,\sigma} \middle\| Q^{\rho,\sigma}\right) = D(\rho\|\sigma) \qquad \text{and} \qquad V\left(P^{\rho,\sigma} \middle\| Q^{\rho,\sigma}\right) = V(\rho\|\sigma).$$

Specifically for iid $Z_i = \log P^{\rho,\sigma}/Q^{\rho,\sigma}$ and $(a_i, b_i) \sim P^{\rho,\sigma}$, then[3]

$$\frac{1}{n} D_h^{\epsilon_n}\left(\rho^{\otimes n} \middle\| \sigma^{\otimes n}\right) \geq \sup\left\{ R \,\middle|\, \Pr\left[\sum_{i=1}^n Z_i\right] \leq \epsilon_n/2 \right\} - \mathcal{O}(\log 1/\epsilon_n),$$

$$\frac{1}{n} D_h^{\epsilon_n}\left(\rho^{\otimes n} \middle\| \sigma^{\otimes n}\right) \leq \sup\left\{ R \,\middle|\, \Pr\left[\sum_{i=1}^n Z_i\right] \leq 2\epsilon_n \right\} + \mathcal{O}(\log 1/\epsilon_n).$$

[2] Nussbaum & Szkoła 2009.

[3] Tomamichel & Hayashi 2013

Bounding the rate

For this we can use the one shot bounds

$$R^*(1,\epsilon) \geq \sup_{P_X} D_h^{\epsilon/2}(\pi_{XY} \| \pi_X \otimes \pi_Y) - \mathcal{O}(1), \qquad \text{(Wang \& Renner 2012)}$$

$$R^*(1,\epsilon) \leq \inf_{\sigma} \sup_{\rho \in \mathrm{Im}(\mathcal{W})} D_h^{2\epsilon}(\rho \| \sigma) + \mathcal{O}(1), \qquad \text{(Tomamichel \& Tan 2015)}$$

where $\pi_{XY} = \sum_x P_X(x)\,|x\rangle\langle x|_X \otimes \rho_Y^{(x)}$.

This give n-shot bounds

$$R^*(n,\epsilon_n) \geq \sup_{P_{X^n}} \frac{1}{n} D_h^{\epsilon_n/2}(\pi_{X^nY^n} \| \pi_{X^n} \otimes \pi_{Y^n}) - \mathcal{O}(1/n),$$

$$R^*(n,\epsilon_n) \leq \inf_{\sigma^n} \sup_{\rho^n \in \mathrm{Im}(\mathcal{W}^{\otimes n})} \frac{1}{n} D_h^{2\epsilon_n}(\rho^n \| \sigma^n) + \mathcal{O}(1/n).$$

We show that a moderate deviation analysis of the rate follows from that of the hypothesis testing divergence.

Bounding the rate

For this we can use the one shot bounds

$$R^*(1,\epsilon) \geq \sup_{P_X} D_h^{\epsilon/2}(\pi_{XY} \| \pi_X \otimes \pi_Y) - \mathcal{O}(1), \qquad \text{(Wang \& Renner 2012)}$$

$$R^*(1,\epsilon) \leq \inf_{\sigma} \sup_{\rho \in \mathrm{Im}(\mathcal{W})} D_h^{2\epsilon}(\rho \| \sigma) + \mathcal{O}(1), \qquad \text{(Tomamichel \& Tan 2015)}$$

where $\pi_{XY} = \sum_x P_X(x) \, |x\rangle\langle x|_X \otimes \rho_Y^{(x)}$.

This give n-shot bounds

$$R^*(n,\epsilon_n) \geq \sup_{P_{X^n}} \frac{1}{n} D_h^{\epsilon_n/2}(\pi_{X^n Y^n} \| \pi_{X^n} \otimes \pi_{Y^n}) - \mathcal{O}(1/n),$$

$$R^*(n,\epsilon_n) \leq \inf_{\sigma^n} \sup_{\rho^n \in \mathrm{Im}(\mathcal{W}^{\otimes n})} \frac{1}{n} D_h^{2\epsilon_n}(\rho^n \| \sigma^n) + \mathcal{O}(1/n).$$

We show that a moderate deviation analysis of the rate follows from that of the hypothesis testing divergence.

Bounding the rate

For this we can use the one shot bounds

$$R^*(1,\epsilon) \geq \sup_{P_X} D_h^{\epsilon/2}(\pi_{XY}\|\pi_X \otimes \pi_Y) - \mathcal{O}(1), \qquad \text{(Wang \& Renner 2012)}$$

$$R^*(1,\epsilon) \leq \inf_{\sigma} \sup_{\rho\in\mathrm{Im}(\mathcal{W})} D_h^{2\epsilon}(\rho\|\sigma) + \mathcal{O}(1), \qquad \text{(Tomamichel \& Tan 2015)}$$

where $\pi_{XY} = \sum_x P_X(x)\,|x\rangle\langle x|_X \otimes \rho_Y^{(x)}$.

This give n-shot bounds

$$R^*(n,\epsilon_n) \geq \sup_{P_{X^n}} \frac{1}{n} D_h^{\epsilon_n/2}(\pi_{X^nY^n}\|\pi_{X^n} \otimes \pi_{Y^n}) - \mathcal{O}(1/n),$$

$$R^*(n,\epsilon_n) \leq \inf_{\sigma^n} \sup_{\rho^n\in\mathrm{Im}(\mathcal{W}^{\otimes n})} \frac{1}{n} D_h^{2\epsilon_n}(\rho^n\|\sigma^n) + \mathcal{O}(1/n).$$

We show that a moderate deviation analysis of the rate follows from that of the hypothesis testing divergence.

Achieveability

In general we have

$$R^*(n,\epsilon_n) \gtrsim \sup_{P_{X^n}} \frac{1}{n} D_h^{\epsilon_n/2}(\pi_{X^nY^n} \| \pi_{X^n} \otimes \pi_{Y^n})$$

where $\pi_{X^nY^n} = \sum_{\vec{x}} P_{X^n}(\vec{x}) \, |\vec{x}\rangle\langle\vec{x}|_{X^n} \otimes \rho_{Y^n}^{(\vec{x})}$.

If P_{X^n} is iid:

$$R^*(n,\epsilon_n) \gtrsim \sup_{P_X} \frac{1}{n} D_h^{\epsilon_n/2}\left(\pi_{XY}^{\otimes n} \middle\| (\pi_X \otimes \pi_Y)^{\otimes n}\right)$$
$$\gtrsim D(\pi_{XY} \| \pi_X \otimes \pi_Y) - \sqrt{2V(\pi_{XY} \| \pi_X \otimes \pi_Y)}\, a_n.$$

Tomamichel & Tan 2015 showed there exists a distribution P_X such that

$$D(\pi_{XY} \| \pi_X \otimes \pi_Y) = C \qquad \text{and} \qquad V(\pi_{XY} \| \pi_X \otimes \pi_Y) = V,$$

and so for $\epsilon_n := e^{-na_n^2}$ this gives

$$R^*(n,\epsilon_n) \gtrsim C - \sqrt{2V}\, a_n.$$

Achieveability

In general we have

$$R^*(n,\epsilon_n) \gtrsim \sup_{P_{X^n}} \frac{1}{n} D_h^{\epsilon_n/2}(\pi_{X^nY^n} \| \pi_{X^n} \otimes \pi_{Y^n})$$

where $\pi_{X^nY^n} = \sum_{\vec{x}} P_{X^n}(\vec{x}) \, |\vec{x}\rangle\langle\vec{x}|_{X^n} \otimes \rho_{Y^n}^{(\vec{x})}$.

If P_{X^n} is iid:

$$\begin{aligned} R^*(n,\epsilon_n) &\gtrsim \sup_{P_X} \frac{1}{n} D_h^{\epsilon_n/2}\left(\pi_{XY}^{\otimes n} \middle\| (\pi_X \otimes \pi_Y)^{\otimes n}\right) \\ &\gtrsim D(\pi_{XY} \| \pi_X \otimes \pi_Y) - \sqrt{2V(\pi_{XY} \| \pi_X \otimes \pi_Y)} a_n. \end{aligned}$$

Tomamichel & Tan 2015 showed there exists a distribution P_X such that

$$D(\pi_{XY} \| \pi_X \otimes \pi_Y) = C \qquad \text{and} \qquad V(\pi_{XY} \| \pi_X \otimes \pi_Y) = V,$$

and so for $\epsilon_n := e^{-na_n^2}$ this gives

$$R^*(n,\epsilon_n) \gtrsim C - \sqrt{2V} a_n.$$

Achieveability

In general we have

$$R^*(n,\epsilon_n) \gtrsim \sup_{P_{X^n}} \frac{1}{n} D_h^{\epsilon_n/2}(\pi_{X^nY^n} \| \pi_{X^n} \otimes \pi_{Y^n})$$

where $\pi_{X^nY^n} = \sum_{\vec{x}} P_{X^n}(\vec{x}) \, |\vec{x}\rangle\langle\vec{x}|_{X^n} \otimes \rho_{Y^n}^{(\vec{x})}$.

If P_{X^n} is iid:

$$\begin{aligned} R^*(n,\epsilon_n) &\gtrsim \sup_{P_X} \frac{1}{n} D_h^{\epsilon_n/2}\left(\pi_{XY}^{\otimes n} \middle\| (\pi_X \otimes \pi_Y)^{\otimes n}\right) \\ &\gtrsim D(\pi_{XY} \| \pi_X \otimes \pi_Y) - \sqrt{2V(\pi_{XY} \| \pi_X \otimes \pi_Y)} a_n. \end{aligned}$$

Tomamichel & Tan 2015 showed there exists a distribution P_X such that

$$D(\pi_{XY} \| \pi_X \otimes \pi_Y) = C \qquad \text{and} \qquad V(\pi_{XY} \| \pi_X \otimes \pi_Y) = V,$$

and so for $\epsilon_n := e^{-na_n^2}$ this gives

$$R^*(n,\epsilon_n) \gtrsim C - \sqrt{2V} a_n.$$

Optimality

We start with

$$R^*(n, \epsilon_n) \lesssim \inf_{\sigma^n} \sup_{\rho^n \in \mathrm{Im}(\mathcal{W}^{\otimes n})} \frac{1}{n} D_h^{2\epsilon_n}(\rho^n \| \sigma^n).$$

As $\mathcal{W}$ is c-q we have that $\rho^n := \otimes_{i=1}^n \rho_i$, so

$$R^*(n, \epsilon_n) \lesssim \inf_{\sigma^n} \sup_{\{\rho_i\} \subset \mathrm{Im}(\mathcal{W})} \frac{1}{n} D_h^{2\epsilon_n} \left(\bigotimes_{i=1}^n \rho_i \middle\| \sigma^n \right).$$

We then separate into 'good' and 'bad' sequences.

$$\text{Good}: \ \frac{1}{n} \sum_{i=1}^n D(\rho_i \| \bar{\rho}_n) > C - \eta \qquad \text{Bad}: \ \frac{1}{n} \sum_{i=1}^n D(\rho_i \| \bar{\rho}_n) \leq C - \eta$$

where $\bar{\rho}_n := \frac{1}{n} \sum_{j=1}^n \rho_j$.
We find a choice of σ^n that gives the desired bound for 'good' and 'bad', then take a mixture.

Optimality

We start with

$$R^*(n,\epsilon_n) \lesssim \inf_{\sigma^n} \sup_{\rho^n \in \mathrm{Im}(\mathcal{W}^{\otimes n})} \frac{1}{n} D_h^{2\epsilon_n}(\rho^n \| \sigma^n).$$

As $\mathcal{W}$ is c-q we have that $\rho^n := \otimes_{i=1}^n \rho_i$, so

$$R^*(n,\epsilon_n) \lesssim \inf_{\sigma^n} \sup_{\{\rho_i\} \subset \mathrm{Im}(\mathcal{W})} \frac{1}{n} D_h^{2\epsilon_n}\left(\bigotimes_{i=1}^n \rho_i \middle\| \sigma^n \right).$$

We then separate into 'good' and 'bad' sequences.

$$\text{Good}: \ \frac{1}{n}\sum_{i=1}^n D\left(\rho_i \| \bar{\rho}_n\right) > C - \eta \qquad \text{Bad}: \ \frac{1}{n}\sum_{i=1}^n D\left(\rho_i \| \bar{\rho}_n\right) \leq C - \eta$$

where $\bar{\rho}_n := \frac{1}{n}\sum_{j=1}^n \rho_j$.
We find a choice of σ^n that gives the desired bound for 'good' and 'bad', then take a mixture.

Optimality

We start with

$$R^*(n, \epsilon_n) \lesssim \inf_{\sigma^n} \sup_{\rho^n \in \mathrm{Im}(\mathcal{W}^{\otimes n})} \frac{1}{n} D_h^{2\epsilon_n}(\rho^n \| \sigma^n).$$

As $\mathcal{W}$ is c-q we have that $\rho^n := \otimes_{i=1}^n \rho_i$, so

$$R^*(n, \epsilon_n) \lesssim \inf_{\sigma^n} \sup_{\{\rho_i\} \subset \mathrm{Im}(\mathcal{W})} \frac{1}{n} D_h^{2\epsilon_n} \left(\bigotimes_{i=1}^n \rho_i \middle\| \sigma^n \right).$$

We then separate into 'good' and 'bad' sequences.

$$\text{Good} : \; \frac{1}{n} \sum_{i=1}^n D\left(\rho_i \| \bar{\rho}_n\right) > C - \eta \qquad \text{Bad} : \; \frac{1}{n} \sum_{i=1}^n D\left(\rho_i \| \bar{\rho}_n\right) \leq C - \eta$$

where $\bar{\rho}_n := \frac{1}{n} \sum_{j=1}^n \rho_j$.
We find a choice of σ^n that gives the desired bound for 'good' and 'bad', then take a mixture.

Conclusion and further work

- We have give a moderate deviation analysis of the capacity of c-q channels, and asymmetric binary hypothesis testing.
- Our proof covers the strong converse and $V = 0$ cases which had not been considered in the classical literature.
- This proof nicely extends to image-additive channels (separable encodings) and infinite input alphabets.

- Can we improve the $o(a_n)$ error terms? It seems they might actually be $\mathcal{O}(a_n^2 + \log n)$.
- Can we extend this to a moderate deviation analysis of other capacities such as the quantum or entanglement-assisted?
- What other channels such as entanglement-breaking?

ArXiv: 1701.03114 (to appear in CMP)