Moderate Deviations for Joint Source-Channel Coding of Systems With Markovian Memory

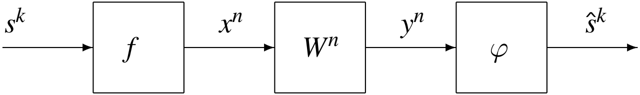
Vincent Y. F. Tan

Department of Electrical and Computer Engineering,
Department of Mathematics,
National University of Singapore

Joint work with Shun Watanabe and Masahito Hayashi

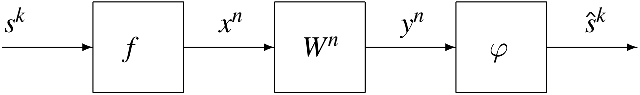
July 2, 2014

Joint Source Channel Coding



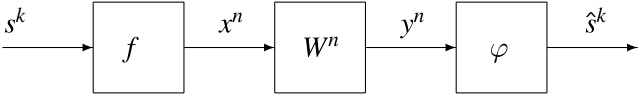
- Consider the fundamental limits of JSCC

Joint Source Channel Coding



- Consider the fundamental limits of JSCC
- Consider Markov sources and additive Markov channels as well discrete memoryless channels

Joint Source Channel Coding



- Consider the fundamental limits of JSCC
- Consider Markov sources and additive Markov channels as well discrete memoryless channels
- Bandwidth expansion ratio r_n satisfies the moderate deviations, i.e.,

$$r_n = \frac{k}{n} = \frac{C(\mathbf{W})}{H(P_{\mathbf{S}})} - \epsilon_n$$

where

$$\epsilon_n \to 0, \quad \text{and} \quad n\epsilon_n^2 \to \infty.$$

Related work in Joint Source Channel Coding

- Error exponents where $\epsilon_n = \epsilon > 0$

  1. Gallager [Problem 5.16 in book] proposed an exponential upper bound on the probability of error
  2. Csiszár (1980) derive bounds on the error exponent which are shown to be tight in some rate regime using the method of types
  3. Zhong-Alajaji-Campbell (2006-2007) related the Csiszár approach to the Gallager approach and extended to stationary, ergodic Markov (SEM) sources and additive SEM channels

Related work in Joint Source Channel Coding

- Error exponents where $\epsilon_n = \epsilon > 0$

  1. Gallager [Problem 5.16 in book] proposed an exponential upper bound on the probability of error
  2. Csiszár (1980) derive bounds on the error exponent which are shown to be tight in some rate regime using the method of types
  3. Zhong-Alajaji-Campbell (2006-2007) related the Csiszár approach to the Gallager approach and extended to stationary, ergodic Markov (SEM) sources and additive SEM channels

- Dispersion where $\epsilon_n = a/\sqrt{n}$ for some $a \in \mathbb{R}$

  1. Wang-Ingber-Kochman (2011) derived the dispersion of JSCC using types
  2. Kostina-Verdú (2012) derived non-asymptotic bounds and dispersion of JSCC for more general settings

Problem Setup

- Finite alphabets $\mathcal{X}, \mathcal{Y}$ and $\mathcal{S}$
- Source $P_{\mathbf{S}} = \{P_{S^k} \in \mathcal{P}(\mathcal{S}^k)\}_{k=1}^{\infty}$
- Channel $\mathbf{W} = \{W^n : \mathcal{X}^n \to \mathcal{Y}^n\}_{n=1}^{\infty}$

Problem Setup

- Finite alphabets $\mathcal{X}, \mathcal{Y}$ and $\mathcal{S}$
- Source $P_{\mathbf{S}} = \{P_{S^k} \in \mathcal{P}(\mathcal{S}^k)\}_{k=1}^{\infty}$
- Channel $\mathbf{W} = \{W^n : \mathcal{X}^n \to \mathcal{Y}^n\}_{n=1}^{\infty}$
- Bandwidth expansion ratio $r_n = k/n$
- Encoder $f_k : \mathcal{S}^k \to \mathcal{X}^n$ and decoder $\varphi_n : \mathcal{Y}^n \to \mathcal{S}^k$
- Error probability

$$e(f_k, \varphi_n) := \sum_{\mathbf{s} \in \mathcal{S}^k} P_{S^k}(\mathbf{s}) W^n(\mathcal{Y}^n \setminus \varphi_n^{-1}(\mathbf{s}) \,|\, f_k(\mathbf{s})).$$

- Minimum error probability

$$e(f_k^*, \varphi_n^*) = \min_{f_k, \varphi_n} e(f_k, \varphi_n).$$

Channel and Source Models

- $\mathbf{W}$ is a DMC with unique capacity-achieving input distribution so

$$W^n(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{n} W(y_i|x_i).$$

Channel and Source Models

- $\mathbf{W}$ is a DMC with unique capacity-achieving input distribution so

$$W^n(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{n} W(y_i|x_i).$$

- $\mathbf{W}$ is a discrete channel with Markov additive noise and alphabet $\mathcal{X} = \mathcal{Y} = \{0, 1, \ldots, b-1\}$ s.t.

$$Y_i = X_i \oplus Z_i \pmod{b}$$

Noise $P_{\mathbf{Z}}$ represents a stationary, ergodic, Markov (SEM) source

$$P_{Z^n}(\mathbf{z}) = P_{Z_1}(z_1) \prod_{i=2}^{n} \Gamma_Z(z_i|z_{i-1})$$

- $P_{\mathbf{S}}$ is a SEM source

Information Quantities for the SEM Source

- Cumulant generating function of any SEM source $P_{\mathbf{S}}$:

$$\phi(\rho; P_{\mathbf{S}}) := \log \sum_{s,s'} \tilde{P}_{S,\rho}(s')\Gamma_S(s|s')^{1-\rho}.$$

where $\tilde{P}_{S,\rho}(s)$ is the normalized eigenvector corresponding to the Perron-Frobenius eigenvalue of the tilted transition matrix $\Gamma_{S,\rho}(s|s')$.

Information Quantities for the SEM Source

- Cumulant generating function of any SEM source $P_{\mathbf{S}}$:

$$\phi(\rho; P_{\mathbf{S}}) := \log \sum_{s,s'} \tilde{P}_{S,\rho}(s') \Gamma_S(s|s')^{1-\rho}.$$

  where $\tilde{P}_{S,\rho}(s)$ is the normalized eigenvector corresponding to the Perron-Frobenius eigenvalue of the tilted transition matrix $\Gamma_{S,\rho}(s|s')$.

- For any SEM source $P_{\mathbf{S}}$, the entropy and varentropy are

$$H(P_{\mathbf{S}}) = \sum_{s'} \tilde{P}(s') \sum_{s} \Gamma_S(s|s') \log \frac{1}{\Gamma_S(s|s')}, \qquad \text{and}$$

$$V(P_{\mathbf{S}}) = \phi''(0; P_{\mathbf{S}}).$$

  where $\tilde{P}$ is the stationary distribution of Γ_S

Information Quantities for the Channel

- For a DMC, the following are the capacity and dispersion

$$C(\mathbf{W}) = \max_P I(P, W), \quad \text{and}$$

$$V(\mathbf{W}) = \sum_x P^*(x) \operatorname{var}\left(\log \frac{W(Y|x)}{P^*W(Y)} \right).$$

Information Quantities for the Channel

- For a DMC, the following are the capacity and dispersion

$$C(\mathbf{W}) = \max_P I(P, W), \quad \text{and}$$
$$V(\mathbf{W}) = \sum_x P^*(x) \operatorname{var}\left(\log \frac{W(Y|x)}{P^*W(Y)}\right).$$

- For the discrete channel with SEM noise,

$$C(\mathbf{W}) = \log b - H(P_{\mathbf{Z}}), \quad \text{and}$$
$$V(\mathbf{W}) = V(P_{\mathbf{Z}}).$$

Main Result

We assume that

$$r_n = \frac{k}{n} = \frac{C(\mathbf{W})}{H(P_{\mathbf{S}})} - \epsilon_n \quad \text{where} \quad \epsilon_n \to 0, \quad \text{and} \quad n\epsilon_n^2 \to \infty.$$

Main Result

We assume that

$$r_n = \frac{k}{n} = \frac{C(\mathbf{W})}{H(P_{\mathbf{S}})} - \epsilon_n \quad \text{where} \quad \epsilon_n \to 0, \quad \text{and} \quad n\epsilon_n^2 \to \infty.$$

Theorem (Moderate Deviations for JSC Coding)

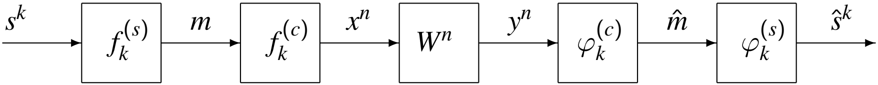
Let $P_{\mathbf{S}}$ be a SEM source and $\mathbf{W}$ be a DMC or additive channel with SEM noise. Then

$$\lim_{n\to\infty} -\frac{1}{n\epsilon_n^2} \log e(f_k^*, \varphi_n^*) = \frac{1}{2V(\mathbf{W}, P_{\mathbf{S}})}$$

where the JSC dispersion is

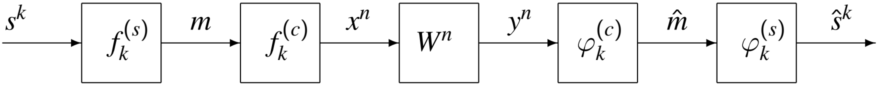
$$V(\mathbf{W}, P_{\mathbf{S}}) := \frac{1}{H(P_{\mathbf{S}})^2} \left[V(\mathbf{W}) + \frac{C(\mathbf{W})}{H(P_{\mathbf{S}})} V(P_{\mathbf{S}}) \right].$$

A Separation Architecture



- Consider MD regime for DMS-DMC when we are forced to design a source and channel code separately

A Separation Architecture

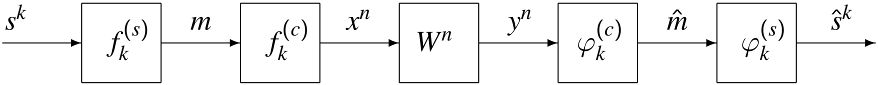


- Consider MD regime for DMS-DMC when we are forced to design a source and channel code separately
- Optimum error probability [Csiszár (1980)] satisfies

$$e\left(f_k^{(s)}, f_k^{(c)}, \varphi_n^{(s)}, \varphi_n^{(c)}\right) \le \exp\left(-n \sup_{R\ge 0} \min\left\{ r_n E_s\left(\frac{R}{r_n}\right), E_c(R)\right\}\right)$$

where $E_s(\cdot)$ and $E_c(\cdot)$ are the source and channel reliability functions resp.

A Separation Architecture



- Consider MD regime for DMS-DMC when we are forced to design a source and channel code separately

- Optimum error probability [Csiszár (1980)] satisfies

$$e\left(f_k^{(s)}, f_k^{(c)}, \varphi_n^{(s)}, \varphi_n^{(c)}\right) \le \exp\left(-n \sup_{R\ge 0} \min\left\{r_n E_s\left(\frac{R}{r_n}\right), E_c(R)\right\}\right)$$

where $E_s(\cdot)$ and $E_c(\cdot)$ are the source and channel reliability functions resp.

- R is the rate of the digital interface m

Loss Due to Separation

- Consider the approximations

$$E_s(H(P_S)+\xi) = \frac{\xi^2}{2V(P_S)} + O(\xi^3), \quad E_c(C(W)-\xi) = \frac{\xi^2}{2V(W)} + O(\xi^3).$$

Loss Due to Separation

- Consider the approximations

$$E_s(H(P_S)+\xi) = \frac{\xi^2}{2V(P_S)} + O(\xi^3), \quad E_c(C(W)-\xi) = \frac{\xi^2}{2V(W)} + O(\xi^3).$$

- Then we can show by optimizing ξ that

$$\liminf_{n\to\infty} -\frac{1}{n\epsilon_n^2} \log e\left(f_k^{(s)}, f_k^{(c)}, \varphi_n^{(s)}, \varphi_n^{(c)}\right) \geq \frac{1}{2V_{\mathrm{sep}}(W,P_S)}$$

where

$$V_{\mathrm{sep}}(W,P_S) = V(W,P_S) + \frac{2\sqrt{\frac{C(W)}{H(P_S)} V(W)V(P_S)}}{H(P_S)^2}$$

Loss Due to Separation

- Consider the approximations

$$E_s(H(P_S)+\xi)=\frac{\xi^2}{2V(P_S)}+O(\xi^3),\quad E_c(C(W)-\xi)=\frac{\xi^2}{2V(W)}+O(\xi^3).$$

- Then we can show by optimizing ξ that

$$\liminf_{n\to\infty}\ -\frac{1}{n\epsilon_n^2}\log e\left(f_k^{(s)},f_k^{(c)},\varphi_n^{(s)},\varphi_n^{(c)}\right)\geq\frac{1}{2V_{\mathrm{sep}}(W,P_S)}$$

where

$$V_{\mathrm{sep}}(W,P_S)=V(W,P_S)+\frac{2\sqrt{\frac{C(W)}{H(P_S)}V(W)V(P_S)}}{H(P_S)^2}$$

- Final term is the cost of separation in the MD regime

Loss Due to Separation

- Consider the approximations

$$E_s(H(P_S)+\xi) = \frac{\xi^2}{2V(P_S)} + O(\xi^3), \quad E_c(C(W)-\xi) = \frac{\xi^2}{2V(W)} + O(\xi^3).$$

- Then we can show by optimizing ξ that

$$\liminf_{n\to\infty} \, -\frac{1}{n\epsilon_n^2} \log e\left(f_k^{(s)}, f_k^{(c)}, \varphi_n^{(s)}, \varphi_n^{(c)}\right) \geq \frac{1}{2V_{\mathrm{sep}}(W,P_S)}$$

where

$$V_{\mathrm{sep}}(W,P_S) = V(W,P_S) + \frac{2\sqrt{\frac{C(W)}{H(P_S)}V(W)V(P_S)}}{H(P_S)^2}$$

- Final term is the cost of separation in the MD regime
- Compare to Wang-Ingber-Kochman (2011) and Kostina-Verdú (2013) for the dispersion analogue of the loss

Achievability Proof for SEM source + DMC

- Now we show how to prove

$$\lim_{n\to\infty} -\frac{1}{n\epsilon_n^2}\log e(f_k^*, \varphi_n^*) = \frac{1}{2V(\mathbf{W}, P_{\mathbf{S}})}$$

where the JSC dispersion is

$$V(\mathbf{W}, P_{\mathbf{S}}) := \frac{1}{H(P_{\mathbf{S}})^2}\left[V(\mathbf{W}) + \frac{C(\mathbf{W})}{H(P_{\mathbf{S}})}V(P_{\mathbf{S}})\right].$$

Achievability Proof for SEM source + DMC

- Now we show how to prove

$$\lim_{n\to\infty} -\frac{1}{n\epsilon_n^2}\log e(f_k^*, \varphi_n^*) = \frac{1}{2V(\mathbf{W}, P_{\mathbf{S}})}$$

  where the JSC dispersion is

$$V(\mathbf{W}, P_{\mathbf{S}}) := \frac{1}{H(P_{\mathbf{S}})^2}\left[V(\mathbf{W}) + \frac{C(\mathbf{W})}{H(P_{\mathbf{S}})}V(P_{\mathbf{S}})\right].$$

- By Problem 5.16 in Gallager's book, exists a (k, n)-JSC code s.t.

$$e(f_k, \varphi_n) \le \left(\sum_{s^k} P_{S^k}(s^k)^{\frac{1}{1+\tau}}\right)^{1+\tau} \sum_{y^n}\left(\sum_{x^n} P^n(x^n)W^n(y^n|x^n)^{\frac{1}{1+\tau}}\right)^{1+\tau}$$

  for any $0 \le \tau \le 1$ and $P \in \mathcal{P}(\mathcal{X})$.

Rényi Entropy and Gallager's function

- Using the reparametrization $\tau = -\theta/(1+\theta)$, we obtain

$$e(f_k, \varphi_n) \le \exp\Big(- \frac{(k-1)H_{1+\theta}(P_{\mathbf{S}}) - \underline{\delta}(\theta)}{1+\theta} \Big) \exp\big(- nE_o(\tau, P, W)\big)$$

where $\underline{\delta}(\theta) = \mathrm{const}$ [bound by Watanabe-Hayashi (2013)]

Rényi Entropy and Gallager's function

- Using the reparametrization $\tau = -\theta/(1+\theta)$, we obtain

$$e(f_k, \varphi_n) \le \exp\Big(- \frac{(k-1) H_{1+\theta}(P_{\mathbf{S}}) - \underline{\delta}(\theta)}{1+\theta}\Big) \exp\big(- n E_o(\tau, P, W)\big)$$

  where $\underline{\delta}(\theta) = \mathrm{const}$ [bound by Watanabe-Hayashi (2013)]

- The Rényi entropy is

$$H_{1+\theta}(P_{\mathbf{S}}) := -\frac{1}{\theta} \sum_{s,s'} \tilde{P}_{S,\theta}(s') \Gamma_{S,\theta}(s|s')$$

Rényi Entropy and Gallager's function

- Using the reparametrization $\tau = -\theta/(1+\theta)$, we obtain

$$e(f_k, \varphi_n) \le \exp\Big(- \frac{(k-1)H_{1+\theta}(P_{\mathbf{S}}) - \underline{\delta}(\theta)}{1+\theta}\Big) \exp\big(- nE_o(\tau, P, W)\big)$$

  where $\underline{\delta}(\theta) = \mathrm{const}$ [bound by Watanabe-Hayashi (2013)]

- The Rényi entropy is

$$H_{1+\theta}(P_{\mathbf{S}}) := -\frac{1}{\theta} \sum_{s,s'} \tilde{P}_{S,\theta}(s') \Gamma_{S,\theta}(s|s')$$

- Gallager's channel function

$$E_o(\tau, P, W) := -\log \sum_x P(x) \Big(\sum_y W(y|x)^{\frac{1}{1+\tau}} \Big)^{1+\tau}$$

Behavior of Rényi entropy and Gallager function

- In the neighborhood of $\theta, \tau = 0$, we have

$$\theta H_{1+\theta}(P_{\mathbf{S}}) = \theta H(P_{\mathbf{S}}) - \frac{\theta^2}{2} V(P_{\mathbf{S}}) + O(\theta^3)$$

$$E_o(\tau, P, W) := \tau I(P, W) - \frac{\tau^2}{2} V(P, W) + O(\tau^3)$$

Behavior of Rényi entropy and Gallager function

- In the neighborhood of $\theta, \tau = 0$, we have

$$\theta H_{1+\theta}(P_{\mathbf{S}}) = \theta H(P_{\mathbf{S}}) - \frac{\theta^2}{2} V(P_{\mathbf{S}}) + O(\theta^3)$$

$$E_o(\tau, P, W) := \tau I(P, W) - \frac{\tau^2}{2} V(P, W) + O(\tau^3)$$

- Plug these into the Gallager bound and optimize τ

$$\tau = \frac{\epsilon_n H(P_{\mathbf{S}})}{V(\mathbf{W}) + \frac{C(\mathbf{W})}{H(P_{\mathbf{S}})} V(P_{\mathbf{S}})}$$

to obtain

$$e(f_k, \varphi_n) \le \exp\left(-n \left[\frac{\epsilon_n^2}{V(\mathbf{W}, P_{\mathbf{S}})} + O(\epsilon_n^3) \right] \right)$$

Converse Proof for SEM source + DMC: Source Part

- $k_{ij}(s^k)$: number of transitions from i to j in s^k with cyclic convention

Converse Proof for SEM source + DMC: Source Part

- $k_{ij}(s^k)$: number of transitions from i to j in s^k with cyclic convention
- $[k_{ij}(s^k)/k]_{i,j}$ is the Markov type [Davisson et al. (1981), Natarajan (1985)] of s^k

Converse Proof for SEM source + DMC: Source Part

- $k_{ij}(s^k)$: number of transitions from i to j in s^k with cyclic convention
- $[k_{ij}(s^k)/k]_{i,j}$ is the Markov type [Davisson et al. (1981), Natarajan (1985)] of s^k
- Type class $\mathcal{T}_P^k$; Set of all Markov types $\mathcal{P}_k^{(2)}(\mathcal{S})$; Restricted Markov type class

$$\mathcal{T}_P^k(i,j) := \{s^k \in \mathcal{T}_P^k : s_1 = i, s_k = j\}$$

$$\mathcal{P}_k^{(2)}(\mathcal{S};R) := \{(i,j,P) \in \mathcal{S}^2 \times \mathcal{P}_k^{(2)}(\mathcal{S}) : |\mathcal{T}_P^k(i,j) \leq 2^{kR}\}$$

Converse Proof for SEM source + DMC: Source Part

- $k_{ij}(s^k)$: number of transitions from i to j in s^k with cyclic convention
- $[k_{ij}(s^k)/k]_{i,j}$ is the Markov type [Davisson et al. (1981), Natarajan (1985)] of s^k
- Type class $\mathcal{T}_P^k$; Set of all Markov types $\mathcal{P}_k^{(2)}(\mathcal{S})$; Restricted Markov type class

$$\mathcal{T}_P^k(i,j) := \{s^k \in \mathcal{T}_P^k : s_1 = i, s_k = j\}$$

$$\mathcal{P}_k^{(2)}(\mathcal{S};R) := \{(i,j,P) \in \mathcal{S}^2 \times \mathcal{P}_k^{(2)}(\mathcal{S}) : |\mathcal{T}_P^k(i,j) \le 2^{kR}\}$$

Lemma (Source coding)

Let $\mathrm{P}_e^*(M, P_S)$ be the smallest error for source P_S with size $\le M$. Then

$$\mathrm{P}_e^*\left(|\mathcal{S}|(k+1)^{|\mathcal{S}|^2} 2^{kR}, P_{S^k}\right) \le \sum_{\substack{(i,j,P)\notin\mathcal{P}_k^{(2)}(\mathcal{S};R) \\ \mathcal{T}_P^k(i,j)\neq\emptyset}} \sum_{s^k\in\mathcal{T}_P(i,j)} P_{S^k}(s^k)$$

Converse Proof: Joint Error + Source Error Approx

Let $\mathrm{P}_e^*(M, W)$ be the smallest error for channel W with M codewords.

Lemma (Source-channel coding)

For any $R \geq 0$, we have

$$e(f_k^*, \varphi_n^*) \geq \mathrm{P}_e^*\left(|\mathcal{S}|(k+1)^{|\mathcal{S}|^2} 2^{kR}, P_{S^k}\right) \cdot \mathrm{P}_e^*(2^{kR}, W^n)$$

Converse Proof: Joint Error + Source Error Approx

Let $\mathrm{P}_e^*(M, W)$ be the smallest error for channel W with M codewords.

Lemma (Source-channel coding)

For any $R \geq 0$, we have

$$e(f_k^*, \varphi_n^*) \geq \mathrm{P}_e^*\big(|\mathcal{S}|(k+1)^{|\mathcal{S}|^2} 2^{kR}, P_{S^k}\big) \cdot \mathrm{P}_e^*(2^{kR}, W^n)$$

First term represents source error; Second term channel error

Converse Proof: Joint Error + Source Error Approx

Let $\mathrm{P}_e^*(M, W)$ be the smallest error for channel W with M codewords.

Lemma (Source-channel coding)

For any $R \geq 0$, we have

$$e(f_k^*, \varphi_n^*) \geq \mathrm{P}_e^*\left(|\mathcal{S}|(k+1)^{|\mathcal{S}|^2} 2^{kR}, P_{S^k}\right) \cdot \mathrm{P}_e^*(2^{kR}, W^n)$$

First term represents source error; Second term represents channel error

Lemma (Approximation of source exponent)

Let $R = H(P_{\mathbf{S}}) + \delta\epsilon_n$. For small $\xi > 0$ and $n \gg 1$,

$$-\log \mathrm{P}_e^*\left(|\mathcal{S}|(k+1)^{|\mathcal{S}|^2} 2^{kR}, P_{S^k}\right) \leq \frac{(\delta+\xi)^2}{2V(P_{\mathbf{S}})} k\epsilon_n^2$$

Converse Proof: Channel Error Approx

By Wolfowitz's strong converse and Haroutunian's sphere-packing bound [cf. Altuğ-Wagner (2014)], we have:

Lemma (Approximation of channel exponent)

For every $\gamma > 0$ *and* $\psi < \infty$*, there exists* $n \gg 1$ *s.t.*

$$-\log \mathrm{P}_e^*(M, W^n) \le \frac{n}{1-\gamma}\left[E\left(\frac{\log M}{n} - \frac{\psi}{\sqrt{n}}\right) + \frac{1}{n}\right]$$

Converse Proof: Channel Error Approx

By Wolfowitz's strong converse and Haroutunian's sphere-packing bound [cf. Altuğ-Wagner (2014)], we have:

Lemma (Approximation of channel exponent)

For every $\gamma > 0$ *and* $\psi < \infty$*, there exists* $n \gg 1$ *s.t.*

$$-\log \mathrm{P}_e^*(M, W^n) \le \frac{n}{1-\gamma}\left[E\left(\frac{\log M}{n} - \frac{\psi}{\sqrt{n}}\right) + \frac{1}{n}\right]$$

- In the rest of the proof, we choose

$$\log M = k\big(H(P_{\mathbf{S}}) + \delta\epsilon_n\big)$$

Converse Proof: Channel Error Approx

By Wolfowitz's strong converse and Haroutunian's sphere-packing bound [cf. Altuğ-Wagner (2014)], we have:

Lemma (Approximation of channel exponent)

For every $\gamma > 0$ and $\psi < \infty$, there exists $n \gg 1$ s.t.

$$-\log \mathrm{P}_e^*(M, W^n) \le \frac{n}{1-\gamma}\left[E\left(\frac{\log M}{n} - \frac{\psi}{\sqrt{n}}\right) + \frac{1}{n}\right]$$

- In the rest of the proof, we choose

$$\log M = k\big(H(P_{\mathbf{S}}) + \delta\epsilon_n\big)$$

- Plug the approximations of the source and channel exponents into the source-channel bound

Converse Proof: Channel Error Approx

By Wolfowitz's strong converse and Haroutunian's sphere-packing bound [cf. Altuğ-Wagner (2014)], we have:

Lemma (Approximation of channel exponent)

For every $\gamma > 0$ and $\psi < \infty$, there exists $n \gg 1$ s.t.

$$-\log \mathrm{P}_e^*(M, W^n) \le \frac{n}{1-\gamma}\left[E\left(\frac{\log M}{n} - \frac{\psi}{\sqrt{n}}\right) + \frac{1}{n}\right]$$

- In the rest of the proof, we choose

$$\log M = k\big(H(P_{\mathbf{S}}) + \delta\epsilon_n\big)$$

- Plug the approximations of the source and channel exponents into the source-channel bound
- Optimize constant δ and take $\xi, \gamma \downarrow 0$ and we are done

Remarks and Conclusion

- Easy to extend analysis to discrete channels with SEM noise using ideas from Watanabe-Hayashi (2013)

Remarks and Conclusion

- Easy to extend analysis to discrete channels with SEM noise using ideas from Watanabe-Hayashi (2013)
- Considered 2 class partition into $\mathcal{P}_k^{(2)}(\mathcal{S};R)$ and its complement; cf. Tauste-Campo et al. (2014), Shkel-Tan-Draper (2014)

Remarks and Conclusion

- Easy to extend analysis to discrete channels with SEM noise using ideas from Watanabe-Hayashi (2013)
- Considered 2 class partition into $\mathcal{P}_k^{(2)}(\mathcal{S}; R)$ and its complement; cf. Tauste-Campo et al. (2014), Shkel-Tan-Draper (2014)
- Zhong-Alajaji-Campbell (2007) considered a finer partition into $(k+1)^{|\mathcal{S}|^2}$ source type classes, that if used here requires

$$\frac{\epsilon_n^2 n}{\log n} \to \infty,$$

a weaker result.

Remarks and Conclusion

- Easy to extend analysis to discrete channels with SEM noise using ideas from Watanabe-Hayashi (2013)
- Considered 2 class partition into $\mathcal{P}_k^{(2)}(\mathcal{S}; R)$ and its complement; cf. Tauste-Campo et al. (2014), Shkel-Tan-Draper (2014)
- Zhong-Alajaji-Campbell (2007) considered a finer partition into $(k+1)^{|\mathcal{S}|^2}$ source type classes, that if used here requires

$$\frac{\epsilon_n^2 n}{\log n} \to \infty,$$

  a weaker result.
- Future work:
  - CLT regime for Markov JSC
  - Lossy JSC