Two Applications of the Gaussian Poincaré Inequality in the Shannon Theory

Vincent Y. F. Tan (Joint work with Silas L. Fong)

National University of Singapore (NUS)

2016 International Zurich Seminar on Communications

Gaussian Poincaré Inequality

Theorem

For $Z^n \stackrel{iid}{\sim} \mathcal{N}(0,1)$ *and any differentiable mapping* f *such that*

$$\mathbb{E}[(f(Z^n))^2] < \infty, \quad \textit{and} \quad \mathbb{E}[\|\nabla f(Z^n)\|^2] < \infty$$

we have

$$\mathrm{var}[f(Z^n)] \le \mathbb{E}[\|\nabla f(Z^n)\|^2].$$

Gaussian Poincaré Inequality

Theorem

For $Z^n \overset{iid}{\sim} \mathcal{N}(0,1)$ *and any differentiable mapping* f *such that*

$$\mathbb{E}[(f(Z^n))^2] < \infty, \quad \textit{and} \quad \mathbb{E}[\|\nabla f(Z^n)\|^2] < \infty$$

we have

$$\mathrm{var}[f(Z^n)] \le \mathbb{E}[\|\nabla f(Z^n)\|^2].$$

- Controlling the variance of $f(Z^n)$ that is a function of i.i.d. random variables in terms of the gradient of $f(Z^n)$

Gaussian Poincaré Inequality

Theorem

For $Z^n \overset{iid}{\sim} \mathcal{N}(0,1)$ and any differentiable mapping f such that

$$\mathbb{E}[(f(Z^n))^2] < \infty, \quad \textit{and} \quad \mathbb{E}[\|\nabla f(Z^n)\|^2] < \infty$$

we have

$$\mathrm{var}[f(Z^n)] \le \mathbb{E}[\|\nabla f(Z^n)\|^2].$$

- Controlling the variance of $f(Z^n)$ that is a function of i.i.d. random variables in terms of the gradient of $f(Z^n)$
- Using the Gaussian Poincaré inequality for appropriate f,

$$\mathrm{var}[f(Z^n)] = O(n).$$

Gaussian Poincaré Inequality in Shannon Theory

- Polyanskiy and Verdú (2014) bounded the KL divergence between the empirical output distribution of AWGN channel codes P_{Y^n} and the n-fold product of the CAOD P_Y^*, i.e.,

$$D(P_{Y^n} \| (P_Y^*)^n)$$

Gaussian Poincaré Inequality in Shannon Theory

- Polyanskiy and Verdú (2014) bounded the KL divergence between the empirical output distribution of AWGN channel codes P_{Y^n} and the n-fold product of the CAOD P_Y^*, i.e.,

$$D(P_{Y^n} \| (P_Y^*)^n)$$

- Often we need to bound the variance of certain log-likelihood ratios (dispersion)

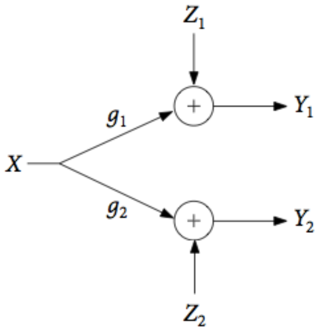
Gaussian Poincaré Inequality in Shannon Theory

- Polyanskiy and Verdú (2014) bounded the KL divergence between the empirical output distribution of AWGN channel codes P_{Y^n} and the n-fold product of the CAOD P_Y^*, i.e.,

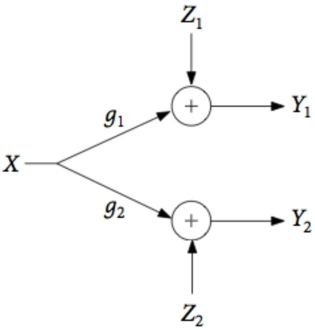
$$D(P_{Y^n} \| (P_Y^*)^n)$$

- Often we need to bound the variance of certain log-likelihood ratios (dispersion)
- Demonstrate its utility by establishing
  - Strong converse for the Gaussian broadcast channels
  - Properties of the empirical output distribution of delay-limited codes for quasi-static fading channels

Gaussian Broadcast Channel

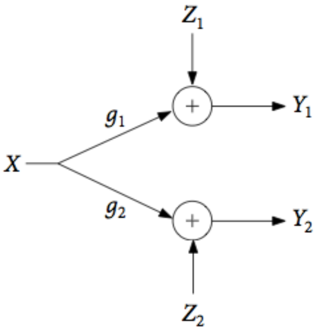


Gaussian Broadcast Channel



- Assume $g_1 = g_2 = 1$ and $\sigma_2^2 > \sigma_1^2$

Gaussian Broadcast Channel



- Assume $g_1 = g_2 = 1$ and $\sigma_2^2 > \sigma_1^2$
- Input X^n must satisfy

$$\|X^n\|_2^2 = \sum_{i=1}^{n} X_i^2 \leq nP$$

Capacity Region

- An $(n, M_{1n}, M_{2n}, \varepsilon_n)$-code consists of
  - an encoder $f : \{1, \ldots, M_{1n}\} \times \{1, \ldots, M_{2n}\} \to \mathbb{R}^n$ such that the power constraint is satisfied;
  - two decoders $\varphi_j : \mathbb{R}^n \to \{1, \ldots, M_{jn}\}$ for $j = 1, 2$;

  such that the average error probability

$$P_e^{(n)} := \Pr(\hat{W}_1 \neq W_1 \text{ or } \hat{W}_2 \neq W_2) \leq \varepsilon_n.$$

Capacity Region

- An $(n, M_{1n}, M_{2n}, \varepsilon_n)$-code consists of
  - an encoder $f : \{1, \ldots, M_{1n}\} \times \{1, \ldots, M_{2n}\} \to \mathbb{R}^n$ such that the power constraint is satisfied;
  - two decoders $\varphi_j : \mathbb{R}^n \to \{1, \ldots, M_{jn}\}$ for $j = 1, 2$;

  such that the average error probability

$$P_e^{(n)} := \Pr(\hat{W}_1 \neq W_1 \text{ or } \hat{W}_2 \neq W_2) \leq \varepsilon_n.$$

- (R_1, R_2) is achievable $\Leftrightarrow \exists$ a sequence of $(n, M_{1n}, M_{2n}, \varepsilon_n)$-codes s.t.

$$\liminf_{n\to\infty} \frac{1}{n} \log M_{jn} \geq R_j, \quad j = 1, 2, \quad \text{and}$$

$$\lim_{n\to\infty} \varepsilon_n = 0.$$

Capacity Region

- An $(n, M_{1n}, M_{2n}, \varepsilon_n)$-code consists of
  - an encoder $f : \{1, \dots, M_{1n}\} \times \{1, \dots, M_{2n}\} \to \mathbb{R}^n$ such that the power constraint is satisfied;
  - two decoders $\varphi_j : \mathbb{R}^n \to \{1, \dots, M_{jn}\}$ for $j = 1, 2$;

  such that the average error probability

$$P_e^{(n)} := \Pr(\hat{W}_1 \neq W_1 \text{ or } \hat{W}_2 \neq W_2) \leq \varepsilon_n.$$

- (R_1, R_2) is achievable $\Leftrightarrow \exists$ a sequence of $(n, M_{1n}, M_{2n}, \varepsilon_n)$-codes s.t.

$$\liminf_{n\to\infty} \frac{1}{n} \log M_{jn} \geq R_j, \quad j = 1, 2, \quad \text{and}$$
$$\lim_{n\to\infty} \varepsilon_n = 0.$$

- Capacity region $\mathcal{C}$ is the set of all achievable rate pairs

Capacity Region

- Cover (1972) and Bergmans (1974) showed that

$$\mathcal{C} = \mathcal{R}_{\mathrm{BC}} = \bigcup_{\alpha \in [0,1]} \mathcal{R}(\alpha)$$

where

$$\mathcal{R}(\alpha) = \left\{ (R_1, R_2) : R_1 \le \mathrm{C}\left(\frac{\alpha P}{\sigma_1^2}\right), R_2 \le \mathrm{C}\left(\frac{(1-\alpha)P}{\alpha P + \sigma_2^2}\right) \right\}$$

and

$$\mathrm{C}(x) = \frac{1}{2} \log(1 + x).$$

Capacity Region

- Cover (1972) and Bergmans (1974) showed that

$$\mathcal{C} = \mathcal{R}_{\mathrm{BC}} = \bigcup_{\alpha \in [0,1]} \mathcal{R}(\alpha)$$

where

$$\mathcal{R}(\alpha) = \left\{ (R_1, R_2) : R_1 \le \mathrm{C}\left(\frac{\alpha P}{\sigma_1^2} \right), R_2 \le \mathrm{C}\left(\frac{(1-\alpha)P}{\alpha P + \sigma_2^2} \right) \right\}$$

and

$$\mathrm{C}(x) = \frac{1}{2} \log(1 + x).$$

- Direct part: Random coding + Superposition coding

Capacity Region

- Cover (1972) and Bergmans (1974) showed that

$$\mathcal{C} = \mathcal{R}_{\mathrm{BC}} = \bigcup_{\alpha\in[0,1]} \mathcal{R}(\alpha)$$

where

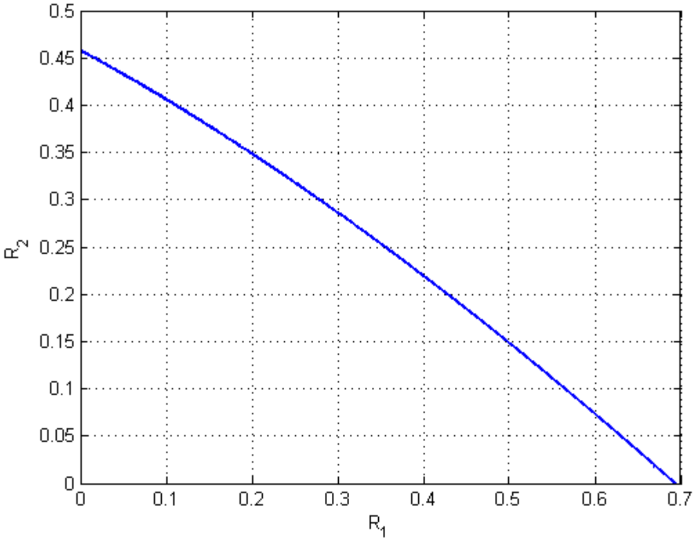
$$\mathcal{R}(\alpha) = \left\{ (R_1, R_2) : R_1 \le \mathrm{C}\left(\frac{\alpha P}{\sigma_1^2}\right), R_2 \le \mathrm{C}\left(\frac{(1-\alpha)P}{\alpha P + \sigma_2^2}\right) \right\}$$

and

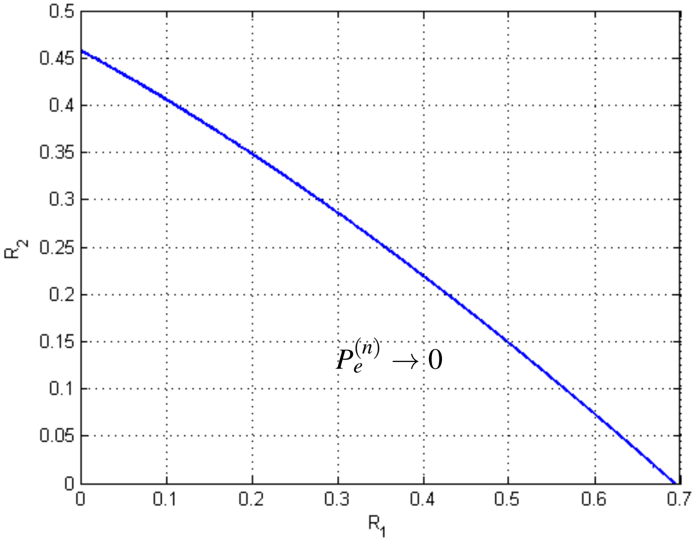
$$\mathrm{C}(x) = \frac{1}{2}\log(1+x).$$

- Direct part: Random coding + Superposition coding
- Converse part: Fano's inequality + Entropy power inequality

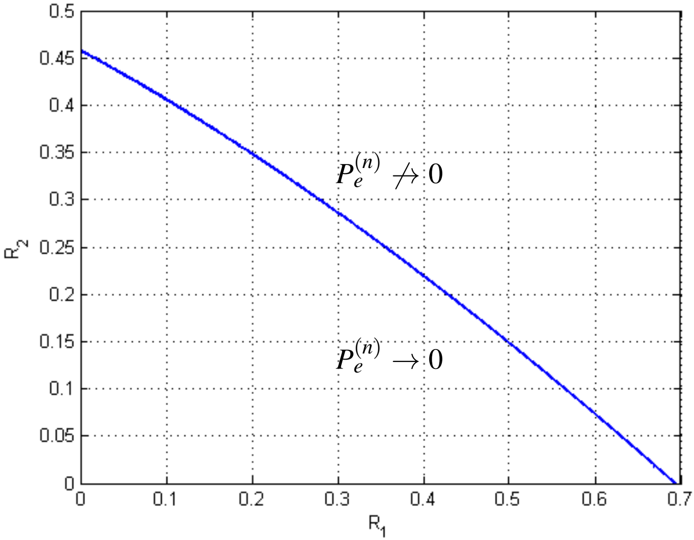
Capacity Region



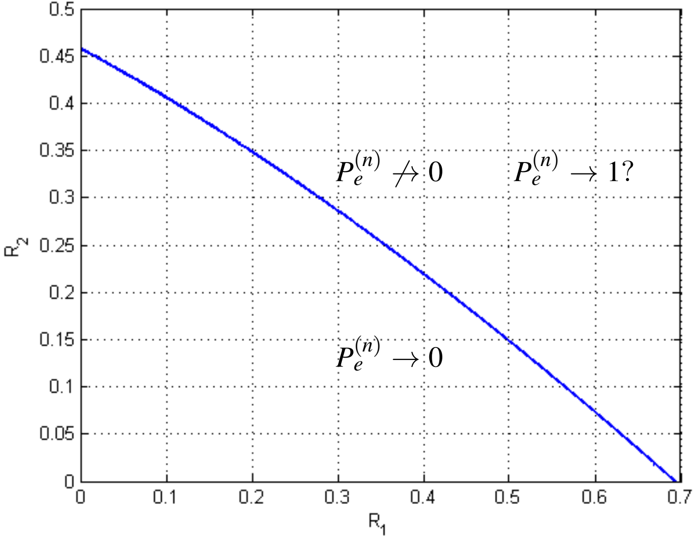
Capacity Region



Capacity Region



Capacity Region



Strong converse vs weak converse

- Can we claim that if $(R_1, R_2) \notin \mathcal{C}$, then

$$P_e^{(n)} \to 1?$$

Strong converse vs weak converse

- Can we claim that if $(R_1, R_2) \notin \mathcal{C}$, then

$$P_e^{(n)} \to 1? \qquad \text{Indeed!}$$

Strong converse vs weak converse

- Can we claim that if $(R_1, R_2) \notin \mathcal{C}$, then

$$P_e^{(n)} \to 1? \qquad \text{Indeed!}$$

- Sharp phase transition between what's possible and what's not

Strong converse vs weak converse

- Can we claim that if $(R_1, R_2) \notin \mathcal{C}$, then

$$P_e^{(n)} \to 1? \qquad \text{Indeed!}$$

- Sharp phase transition between what's possible and what's not
- The strong converse has been established for only degraded discrete memoryless BC

Strong converse vs weak converse

- Can we claim that if $(R_1, R_2) \notin \mathcal{C}$, then

$$P_e^{(n)} \to 1? \qquad \text{Indeed!}$$

- Sharp phase transition between what's possible and what's not
- The strong converse has been established for only degraded discrete memoryless BC
- Ahlswede, Gács and Körner (1976) used the blowing-up lemma
  - BUL doesn't work for continuous alphabets [but see Wu and Özgür (2015)]

Strong converse vs weak converse

- Can we claim that if $(R_1, R_2) \notin \mathcal{C}$, then

$$P_e^{(n)} \to 1? \qquad \text{Indeed!}$$

- Sharp phase transition between what's possible and what's not
- The strong converse has been established for only degraded discrete memoryless BC
- Ahlswede, Gács and Körner (1976) used the blowing-up lemma
  - BUL doesn't work for continuous alphabets [but see Wu and Özgür (2015)]
- Oohama (2015) uses properties of the Rényi divergence
  - Good bounds between the Rényi divergence $D_\alpha(P\|Q)$ and the relative entropy $D(P\|Q)$ exist for finite alphabets

ε-Capacity Region

- (R_1, R_2) is ε-achievable $\Leftrightarrow \exists$ a sequence of $(n, M_{1n}, M_{2n}, \varepsilon_n)$-codes s.t.

$$\liminf_{n\to\infty} \frac{1}{n} \log M_{jn} \geq R_j, \quad j = 1, 2, \quad \text{and}$$
$$\limsup_{n\to\infty} \varepsilon_n \leq \varepsilon.$$

- Capacity region $\mathcal{C}_\varepsilon$ is the set of all achievable rate pairs
- Strong converse holds iff $\mathcal{C}_\varepsilon$ does not depend on ε.
- We already know that

$$\mathcal{R}_{\text{BC}} = \mathcal{C}_0 \subset \mathcal{C}_\varepsilon$$

Strong converse

Theorem

The Gaussian BC satisfies the strong converse property:

$$\mathcal{C}_\varepsilon = \mathcal{R}_{\mathrm{BC}}, \qquad \forall\, \varepsilon \in [0, 1)$$

Key ideas in proof:

- Derive an appropriate information spectrum converse bound
- Use the Gaussian Poincaré inequality to bound relevant variances

Weak Converse for GBC [Bergmans (1974)]

- Step 1: Invoke Fano's inequality to assert that for any sequence of codes with vanishing error probability $\varepsilon_n \to 0$,

$$R_j \le \frac{1}{n} I(W_j; Y_j^n) + o(1), \qquad \forall j \in \{1, 2\}.$$

Weak Converse for GBC [Bergmans (1974)]

- Step 1: Invoke Fano's inequality to assert that for any sequence of codes with vanishing error probability $\varepsilon_n \to 0$,

$$R_j \leq \frac{1}{n} I(W_j; Y_j^n) + o(1), \qquad \forall j \in \{1, 2\}.$$

- Step 2: Single-letterize and entropy power inequality

$$I(W_1; Y_1^n) \leq nI(X; Y_1|U) \overset{EPI}{\leq} n\mathrm{C}\left(\frac{\alpha P}{\sigma_1^2}\right)$$

$$I(W_1; Y_1^n) + I(W_2; Y_2^n) \leq nI(U; Y_2) \overset{EPI}{\leq} n\mathrm{C}\left(\frac{(1-\alpha)P}{\alpha P + \sigma_2^2}\right)$$

Strong Converse for DM-BC [Ahlswede et al. (1976)]

- Step 1: Invoke the blowing-up lemma to assert that for any sequence of codes with non-vanishing error probability $\varepsilon \in [0,1)$,

$$R_j \le \frac{1}{n} I(W_j; Y_j^n) + o(1), \qquad \forall j \in \{1,2\}.$$

Strong Converse for DM-BC [Ahlswede et al. (1976)]

- Step 1: Invoke the blowing-up lemma to assert that for any sequence of codes with non-vanishing error probability $\varepsilon \in [0,1)$,

$$R_j \le \frac{1}{n} I(W_j; Y_j^n) + o(1), \qquad \forall j \in \{1,2\}.$$

- Step 2: Single-letterize

$$I(W_1; Y_1^n) \le nI(X; Y_1 | U),$$
$$I(W_1; Y_1^n) + I(W_2; Y_2^n) \le nI(U; Y_2)$$

where $U_i := (W_2, Y_1^{i-1})$. One also uses the degradedness condition here:

$$I(W_2, Y_2^{i-1}, Y_1^{i-1}; Y_{2i}) = I(U_i; Y_{2i}).$$

Our Strong Converse Proof for Gaussian BC

- Convert code defined based on avg error prob $\leq \varepsilon$ to one based on max error prob $\leq \sqrt{\varepsilon} =: \varepsilon'$ w/o loss in rate [Telatar]

Our Strong Converse Proof for Gaussian BC

- Convert code defined based on avg error prob $\le \varepsilon$ to one based on max error prob $\le \sqrt{\varepsilon} =: \varepsilon'$ w/o loss in rate [Telatar]
- Establish information spectrum bound. For every (w_1, w_2), every code with max error prob $\le \varepsilon'$ satisfies

$$\varepsilon' \ge \Pr\left(\log \frac{P(Y_1^n | w_1)}{P(Y_1^n)} \le nR_1 - \gamma_1(w_1, w_2) \right) - n^2 e^{-\gamma_1(w_1, w_2)} - \mathbf{1}\left\{ 2^{n(R_1+R_2)} \int_{\mathcal{D}_1(w_1)} P(y_1^n) P(w_2 | w_1, y_1^n) \, dy_1^n > n^2 \right\}$$

Our Strong Converse Proof for Gaussian BC

- Convert code defined based on avg error prob $\le \varepsilon$ to one based on max error prob $\le \sqrt{\varepsilon} =: \varepsilon'$ w/o loss in rate [Telatar]
- Establish information spectrum bound. For every (w_1, w_2), every code with max error prob $\le \varepsilon'$ satisfies

$$\varepsilon' \ge \Pr\left(\log \frac{P(Y_1^n | w_1)}{P(Y_1^n)} \le nR_1 - \gamma_1(w_1, w_2) \right) - n^2 e^{-\gamma_1(w_1, w_2)}$$

$$- \mathbf{1}\left\{ 2^{n(R_1+R_2)} \int_{\mathcal{D}_1(w_1)} P(y_1^n) P(w_2 | w_1, y_1^n) \, dy_1^n > n^2 \right\}$$

- Indicator term is often negligible (by Markov's inequality)

Our Strong Converse Proof for Gaussian BC

- Convert code defined based on avg error prob $\le \varepsilon$ to one based on max error prob $\le \sqrt{\varepsilon} =: \varepsilon'$ w/o loss in rate [Telatar]

- Establish information spectrum bound. For every (w_1, w_2), every code with max error prob $\le \varepsilon'$ satisfies

$$\varepsilon' \ge \Pr\left(\log \frac{P(Y_1^n|w_1)}{P(Y_1^n)} \le nR_1 - \gamma_1(w_1,w_2)\right) - n^2 e^{-\gamma_1(w_1,w_2)}$$
$$- \mathbf{1}\left\{2^{n(R_1+R_2)} \int_{\mathcal{D}_1(w_1)} P(y_1^n) P(w_2|w_1,y_1^n)\, dy_1^n > n^2\right\}$$

- Indicator term is often negligible (by Markov's inequality)

- Establish a bound on the coding rate

$$nR_1 \le \mathbb{E}\left[\log \frac{P(Y_1^n|w_1)}{P(Y_1^n)}\right] + \sqrt{\frac{2}{1-\varepsilon'} \operatorname{var}\left[\log \frac{P(Y_1^n|w_1)}{P(Y_1^n)}\right]} + 3\log n$$

Our Strong Converse Proof for Gaussian BC

- Use the Gaussian Poincaré inequality with careful identification of f and peak power constraint $\|X^n\|^2 \le nP$ to assert that

$$\mathrm{var}\left[\log \frac{P(Y_1^n|w_1)}{P(Y_1^n)}\right] = O(n).$$

Our Strong Converse Proof for Gaussian BC

- Use the Gaussian Poincaré inequality with careful identification of f and peak power constraint $\|X^n\|^2 \le nP$ to assert that

$$\mathrm{var}\left[\log \frac{P(Y_1^n|w_1)}{P(Y_1^n)}\right] = O(n).$$

- Thus we conclude that

$$nR_1 \le I(W_1;Y_1^n) + O(\sqrt{n}), \qquad \forall\, \varepsilon \in [0,1).$$

Our Strong Converse Proof for Gaussian BC

- Use the Gaussian Poincaré inequality with careful identification of f and peak power constraint $\|X^n\|^2 \le nP$ to assert that

$$\mathrm{var}\left[\log \frac{P(Y_1^n|w_1)}{P(Y_1^n)}\right] = O(n).$$

- Thus we conclude that

$$nR_1 \le I(W_1; Y_1^n) + O(\sqrt{n}), \qquad \forall\, \varepsilon \in [0,1).$$

- Backoff term is of the order $O(1/\sqrt{n})$, i.e.,

$$\lambda \log M_{1n}^* + (1-\lambda)\log M_{2n}^* = n\mathrm{C}_\lambda + O(\sqrt{n})$$

where

$$\mathrm{C}_\lambda := \max_{\alpha\in[0,1]} \left\{ \lambda \mathrm{C}\Big(\frac{\alpha P}{\sigma_1^2}\Big) + (1-\lambda)\mathrm{C}\Big(\frac{(1-\alpha)P}{\alpha P + \sigma_2^2}\Big) \right\}$$

but nailing down the constant (dispersion) seems challenging.

Another Application: Quasi-Static Fading Channels

- Consider the channel model

$$Y_i = \sqrt{H} X_i + Z_i, \qquad i = 1, \ldots, n$$

where Z_i are independent standard normal random variables and

$$\mathbb{E}[1/H] \in (0, \infty)$$

Another Application: Quasi-Static Fading Channels

- Consider the channel model

$$Y_i = \sqrt{H} X_i + Z_i, \qquad i = 1, \ldots, n$$

where Z_i are independent standard normal random variables and

$$\mathbb{E}[1/H] \in (0, \infty)$$

- If fading state is h and message is w, codeword is $f_h(w) \in \mathbb{R}^n$.

Another Application: Quasi-Static Fading Channels

- Consider the channel model

$$Y_i = \sqrt{H} X_i + Z_i, \qquad i = 1, \dots, n$$

where Z_i are independent standard normal random variables and

$$\mathbb{E}[1/H] \in (0, \infty)$$

- If fading state is h and message is w, codeword is $f_h(w) \in \mathbb{R}^n$.
- Long-term power constraint

$$\frac{1}{M_n} \sum_{w=1}^{M_n} \int_{\mathbb{R}_+} P_H(h) \|f_h(w)\|^2 \, \mathrm{d}h \le nP$$

Another Application: Quasi-Static Fading Channels

- Consider the channel model

$$Y_i = \sqrt{H} X_i + Z_i, \qquad i = 1, \ldots, n$$

where Z_i are independent standard normal random variables and

$$\mathbb{E}[1/H] \in (0, \infty)$$

- If fading state is h and message is w, codeword is $f_h(w) \in \mathbb{R}^n$.

- Long-term power constraint

$$\frac{1}{M_n} \sum_{w=1}^{M_n} \int_{\mathbb{R}_+} P_H(h) \|f_h(w)\|^2 \, \mathrm{d}h \le nP$$

- Delay-limited capacity [Hanly and Tse (1998)], i.e., the maximum transmission rate under the constraint that the maximal error probability over all $H > 0$ vanishes

Vanishing Normalized Relative Entropy

- The delay-limited capacity [Hanly and Tse (1998)] is

$$\mathrm{C}(P_{\mathrm{DL}}), \quad \text{where} \quad P_{\mathrm{DL}} := \frac{P}{\mathbb{E}[1/H]}$$

- For any sequence of capacity-achieving codes with vanishing maximum error probability

$$\lim_{n\to\infty} \frac{1}{n} D\big(P_{Y^n} \| (P_Y^*)^n\big) \to 0 \quad \text{where} \quad P_Y^* = \mathcal{N}(0, P_{\mathrm{DL}}).$$

Vanishing Normalized Relative Entropy

- The delay-limited capacity [Hanly and Tse (1998)] is

$$\mathrm{C}(P_{\mathrm{DL}}), \quad \text{where} \quad P_{\mathrm{DL}} := \frac{P}{\mathbb{E}[1/H]}$$

- For any sequence of capacity-achieving codes with vanishing maximum error probability

$$\lim_{n\to\infty} \frac{1}{n} D\big(P_{Y^n} \| (P_Y^*)^n\big) \to 0 \quad \text{where} \quad P_Y^* = \mathcal{N}(0, P_{\mathrm{DL}}).$$

- Every good code is s.t. the induced output distribution “looks like” the n-fold CAOD.

Vanishing Normalized Relative Entropy

- The delay-limited capacity [Hanly and Tse (1998)] is

$$\mathrm{C}(P_{\mathrm{DL}}), \quad \text{where} \quad P_{\mathrm{DL}} := \frac{P}{\mathbb{E}[1/H]}$$

- For any sequence of capacity-achieving codes with vanishing maximum error probability

$$\lim_{n\to\infty} \frac{1}{n} D\big(P_{Y^n} \| (P_Y^*)^n\big) \to 0 \quad \text{where} \quad P_Y^* = \mathcal{N}(0, P_{\mathrm{DL}}).$$

- Every good code is s.t. the induced output distribution “looks like” the n-fold CAOD.
- Extend to the case where the error probability is non-vanishing

Vanishing Normalized Relative Entropy

- The delay-limited capacity [Hanly and Tse (1998)] is

$$\mathrm{C}(P_{\mathrm{DL}}), \quad \text{where} \quad P_{\mathrm{DL}} := \frac{P}{\mathbb{E}[1/H]}$$

- For any sequence of capacity-achieving codes with vanishing maximum error probability

$$\lim_{n\to\infty} \frac{1}{n} D\big(P_{Y^n} \| (P_Y^*)^n\big) \to 0 \quad \text{where} \quad P_Y^* = \mathcal{N}(0, P_{\mathrm{DL}}).$$

- Every good code is s.t. the induced output distribution “looks like” the n-fold CAOD.
- Extend to the case where the error probability is non-vanishing
- Control a variance term

$$\mathrm{var}\left[\log \frac{P_{Y^n|X^n,H}(Y^n|X^n,h)}{P_{Y^n|H}(Y^n|h)}\right] = O(n).$$

Concluding Remarks

- Gaussian Poincaré inequality is useful for Shannon-theoretic problems with uncountable alphabets

$$\mathrm{var}[f(Z^n)] \le \mathbb{E}[\|\nabla f(Z^n)\|^2].$$

Concluding Remarks

- Gaussian Poincaré inequality is useful for Shannon-theoretic problems with uncountable alphabets

$$\mathrm{var}[f(Z^n)] \le \mathbb{E}[\|\nabla f(Z^n)\|^2].$$

- Allows us to establish strong converses and properties of good codes by controlling variance of log-likelihood ratios

Concluding Remarks

- Gaussian Poincaré inequality is useful for Shannon-theoretic problems with uncountable alphabets

$$\mathrm{var}[f(Z^n)] \le \mathbb{E}[\|\nabla f(Z^n)\|^2].$$

- Allows us to establish strong converses and properties of good codes by controlling variance of log-likelihood ratios
- For more details, please see
  - Arxiv: 1509.01380 (Strong converse for Gaussian broadcast)
  - Arxiv: 1510.08544 (Empirical output distribution of good codes for fading channels)