The Dispersion of Slepian-Wolf Coding

Vincent Y. F. Tan[*], Oliver Kosut[†]

[*] Institute for Infocomm Research (I2R) and National University of Singapore (NUS)
[†] LIDS, Massachusetts Institute of Technology

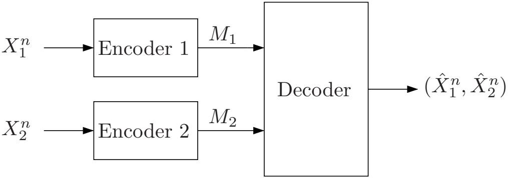
July 3, 2012

Joint work with

Oliver Kosut
Massachusetts Institute of Technology

The Slepian-Wolf Problem

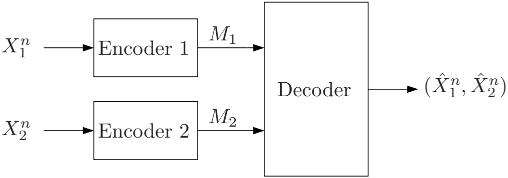
The optimal rate region for the Slepian-Wolf problem $\mathcal{R}^*_{\mathrm{SW}}$ is



$$R_1 \geq H(X_1|X_2)$$
$$R_2 \geq H(X_2|X_1)$$
$$R_1 + R_2 \geq H(X_1, X_2)$$

The Slepian-Wolf Problem

The optimal rate region for the Slepian-Wolf problem $\mathcal{R}^*_{\mathrm{SW}}$ is



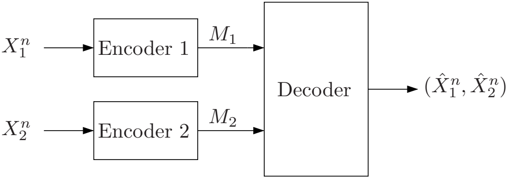
$$R_1 \geq H(X_1|X_2)$$
$$R_2 \geq H(X_2|X_1)$$
$$R_1 + R_2 \geq H(X_1, X_2)$$

How does the region change if we impose that the blocklength n but allow the error probability to be no larger than $\epsilon > 0$?

The error probability is defined as

$$\mathbb{P}((\hat{X}_1^n, \hat{X}_2^n) \neq (X_1^n, X_2^n))$$

The optimal rate region for the Slepian-Wolf problem $\mathcal{R}^*_{\mathrm{SW}}$ is



$$R_1 \geq H(X_1|X_2)$$
$$R_2 \geq H(X_2|X_1)$$
$$R_1 + R_2 \geq H(X_1, X_2)$$

How does the region change if we impose that the blocklength n but allow the error probability to be no larger than $\epsilon > 0$?

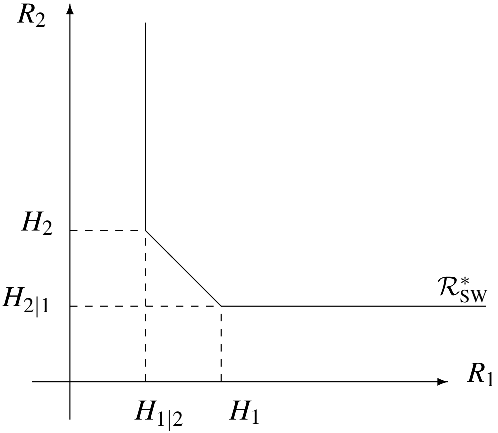
The error probability is defined as

$$\mathbb{P}((\hat{X}_1^n, \hat{X}_2^n) \neq (X_1^n, X_2^n))$$

The (n, ϵ)-region $\mathcal{R}^*_{\mathrm{SW}}(n, \epsilon)$ is the set of all (n, ϵ)-achievable rates.

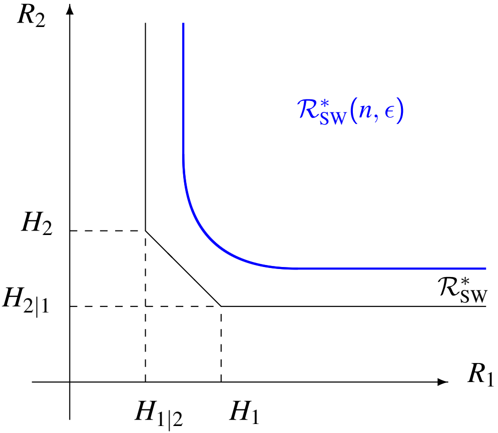
Main Contribution

$$R_1 \geq H(X_1|X_2)$$
$$R_2 \geq H(X_2|X_1)$$
$$R_1 + R_2 \geq H(X_1, X_2)$$



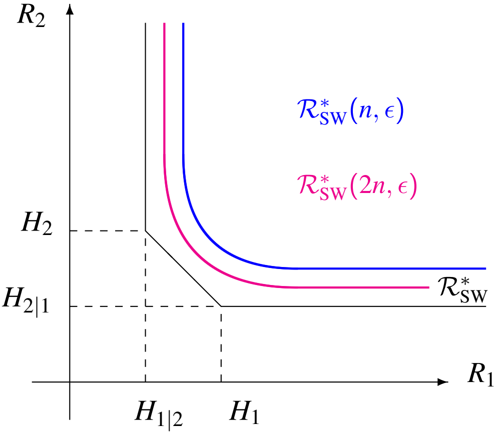
Main Contribution

$$R_1 \geq H(X_1|X_2)$$
$$R_2 \geq H(X_2|X_1)$$
$$R_1 + R_2 \geq H(X_1, X_2)$$



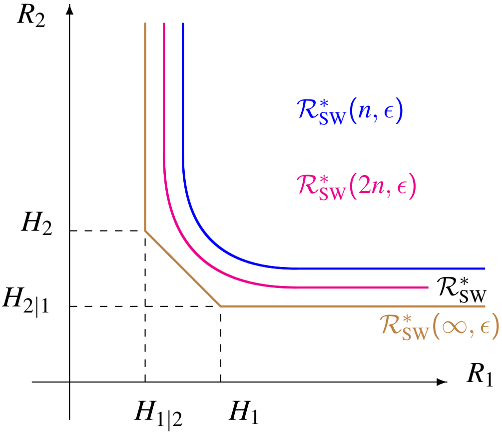
Main Contribution

$$R_1 \geq H(X_1|X_2)$$
$$R_2 \geq H(X_2|X_1)$$
$$R_1 + R_2 \geq H(X_1, X_2)$$



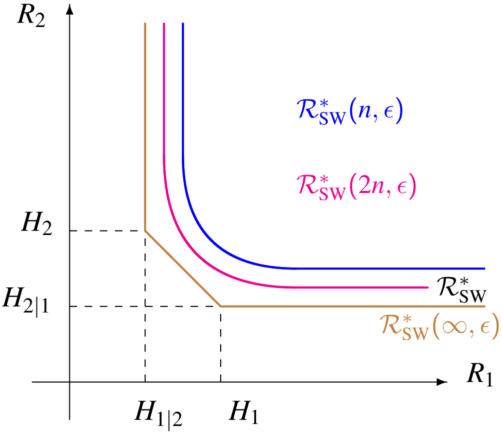
Main Contribution

$$R_1 \geq H(X_1|X_2)$$
$$R_2 \geq H(X_2|X_1)$$
$$R_1 + R_2 \geq H(X_1, X_2)$$



Main Contribution

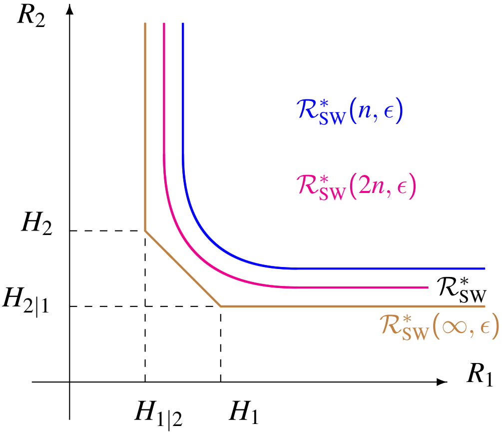
$$R_1 \geq H(X_1|X_2)$$
$$R_2 \geq H(X_2|X_1)$$
$$R_1 + R_2 \geq H(X_1, X_2)$$



For $n < \infty$, corners are curved!

Main Contribution

$$R_1 \geq H(X_1|X_2)$$
$$R_2 \geq H(X_2|X_1)$$
$$R_1 + R_2 \geq H(X_1, X_2)$$



For $n < \infty$, corners are curved!

Figure "correct" up to $O(\frac{\log n}{n})$ terms.

Prior Work on (n, ϵ)-regions

- For fixed-length lossless source coding [Strassen (1964), Kontoyiannis (1997), Hayashi (2008)],

$$R^*(n,\epsilon) = H(X) + \sqrt{\frac{V_1}{n}} Q^{-1}(\epsilon) + O\left(\frac{\log n}{n}\right)$$

where the source coding dispersion $V_1 := \mathrm{Var}(-\log p(X))$.

Prior Work on (n, ϵ)-regions

- For fixed-length lossless source coding [Strassen (1964), Kontoyiannis (1997), Hayashi (2008)],

$$R^*(n, \epsilon) = H(X) + \sqrt{\frac{V_1}{n}} Q^{-1}(\epsilon) + O\left(\frac{\log n}{n}\right)$$

where the source coding dispersion $V_1 := \mathrm{Var}(-\log p(X))$.

- Recent second-order coding rate results for
  1. Intrinsic randomness and resolvability (Hayashi 2008)
  2. Channel coding (Polyanskiy et al. 2010 and Hayashi 2009)
  3. Rate distortion (Ingber-Kochman 2011, Kostina-Verdú 2011)
  4. Joint source-channel coding (Wang-Ingber-Kochman 2011, Kostina-Verdú 2012)

Prior Work on (n, ϵ)-regions

- For fixed-length lossless source coding [Strassen (1964), Kontoyiannis (1997), Hayashi (2008)],

$$R^*(n,\epsilon) = H(X) + \sqrt{\frac{V_1}{n}} Q^{-1}(\epsilon) + O\left(\frac{\log n}{n}\right)$$

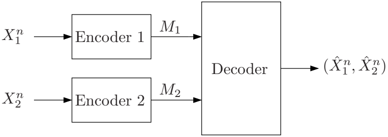
where the source coding dispersion $V_1 := \mathrm{Var}(-\log p(X))$.

- Recent second-order coding rate results for

1. Intrinsic randomness and resolvability (Hayashi 2008)
2. Channel coding (Polyanskiy et al. 2010 and Hayashi 2009)
3. Rate distortion (Ingber-Kochman 2011, Kostina-Verdú 2011)
4. Joint source-channel coding (Wang-Ingber-Kochman 2011, Kostina-Verdú 2012)

- First non-trivial network information theory problem (where there is more than one rate) is the Slepian-Wolf problem

(n, ϵ)-region for the Slepian-Wolf Problem



An $(n, 2^{nR_1}, 2^{nR_2}, \epsilon)$ code for the source (X_1, X_2) is characterized by

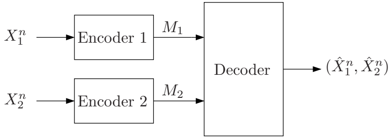
1. Two Encoders: $f_{j,n} : \mathcal{X}_j^n \to [1 : 2^{nR_j}]$ for $j = 1, 2$
2. Decoder: $\varphi_n : [1 : 2^{nR_1}] \times [1 : 2^{nR_2}] \to \mathcal{X}_1^n \times \mathcal{X}_2^n$

such that $\mathbb{P}((\hat{X}_1^n, \hat{X}_2^n) \neq (X_1^n, X_2^n)) \leq \epsilon$.

Definition

Rate pair (R_1, R_2) is (n, ϵ)-achievable if there exists an $(n, 2^{nR_1}, 2^{nR_2}, \epsilon)$ code for the source (X_1, X_2).

(n, ϵ)-region for the Slepian-Wolf Problem



An $(n, 2^{nR_1}, 2^{nR_2}, \epsilon)$ code for the source (X_1, X_2) is characterized by

1. Two Encoders: $f_{j,n} : \mathcal{X}_j^n \to [1 : 2^{nR_j}]$ for $j = 1, 2$
2. Decoder: $\varphi_n : [1 : 2^{nR_1}] \times [1 : 2^{nR_2}] \to \mathcal{X}_1^n \times \mathcal{X}_2^n$

such that $\mathbb{P}((\hat{X}_1^n, \hat{X}_2^n) \neq (X_1^n, X_2^n)) \leq \epsilon$.

Definition

Rate pair (R_1, R_2) is (n, ϵ)-achievable if there exists an $(n, 2^{nR_1}, 2^{nR_2}, \epsilon)$ code for the source (X_1, X_2).
The (n, ϵ)-optimal rate region $\mathcal{R}^*_{\mathrm{SW}}(n, \epsilon)$ is the set of all (n, ϵ)-achievable rate pairs.

Related Work

- Baron et al. (2004): Slepian-Wolf with perfect side-information

$$R_1 \approx H(X_1|X_2) + \sqrt{\frac{V}{n}} Q^{-1}(\epsilon)$$

where $V = \mathrm{Var}(-\log p_{X_1|X_2}(X_1|X_2))$

Related Work

- Baron et al. (2004): Slepian-Wolf with perfect side-information

$$R_1 \approx H(X_1|X_2) + \sqrt{\frac{V}{n}} Q^{-1}(\epsilon)$$

where $V = \mathrm{Var}(-\log p_{X_1|X_2}(X_1|X_2))$

- Sarvotham-Baron-Baraniuk (2005): Slepian-Wolf for correlated Bernoulli(1/2) sources

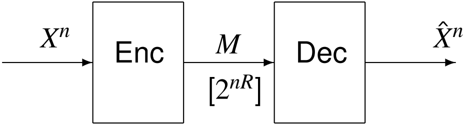
$$R_1 \geq H(X_1|X_2) + \sqrt{\frac{V_1}{n}} Q^{-1}(\epsilon)$$

$$R_2 \geq H(X_2|X_1) + \sqrt{\frac{V_2}{n}} Q^{-1}(\epsilon)$$

$$R_1 + R_2 \geq H(X_1, X_2) + \sqrt{\frac{V_3}{n}} Q^{-1}(\epsilon)$$

where $V_1 = \mathrm{Var}(-\log p_{X_1|X_2}(X_1|X_2))$, $V_2 = \mathrm{Var}(-\log p_{X_2|X_1}(X_2|X_1))$ and $V_3 = \mathrm{Var}(-\log p_{X_1,X_2}(X_1, X_2))$.

Point-To-Point Source Coding Revisited



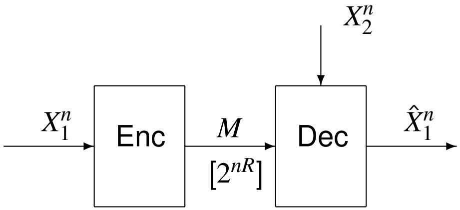
Rate R is (n, ϵ)-achievable if [Strassen (1964) and Kontoyiannis (1997)]

$$R \geq H(X) + \sqrt{\frac{V}{n}} Q^{-1}(\epsilon)$$

Define entropy density $h(x) = -\log p_X(x)$, then

- $H(X) = \mathbb{E}[h(X)]$
- $V = \mathrm{Var}[h(X)]$

Slepian-Wolf with Perfect Side Information



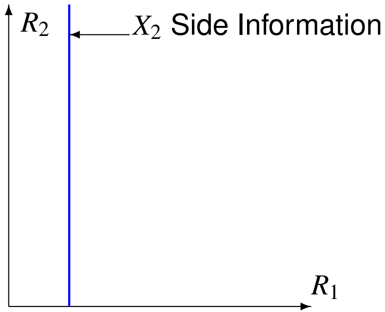
Baron et al. (2004): Rate R is (n, ϵ)-achievable if

$$R \geq H(X_1|X_2) + \sqrt{\frac{V}{n}} Q^{-1}(\epsilon)$$

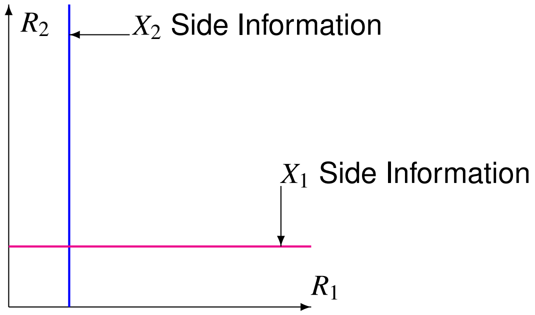
Define conditional entropy density $h(x_1|x_2) = -\log p_{X_1|X_2}(x_1|x_2)$, then

- $H(X_1|X_2) = \mathbb{E}[h(X_1|X_2)]$
- $V = \mathrm{Var}[h(X_1|X_2)]$

Simple Slepian-Wolf Outer Bound

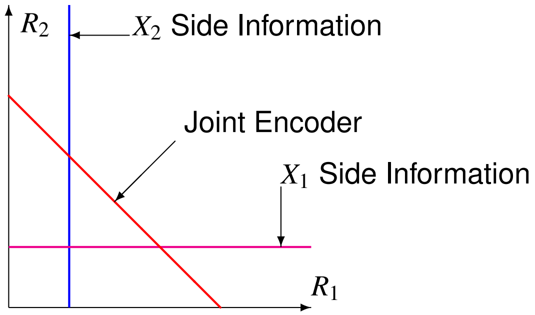


$$R_1 \geq H(X_1|X_2) + \sqrt{V_1/n}\, Q^{-1}(\epsilon), \qquad V_1 := \mathrm{Var}[-\log p(X_1|X_2)]$$



$$R_1 \geq H(X_1|X_2) + \sqrt{V_1/n}\, Q^{-1}(\epsilon), \qquad V_1 := \mathrm{Var}[-\log p(X_1|X_2)]$$
$$R_2 \geq H(X_2|X_1) + \sqrt{V_2/n}\, Q^{-1}(\epsilon), \qquad V_2 := \mathrm{Var}[-\log p(X_2|X_1)]$$

Simple Slepian-Wolf Outer Bound

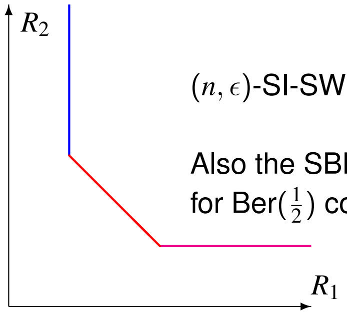


$$R_1 \geq H(X_1|X_2) + \sqrt{V_1/n}\, Q^{-1}(\epsilon), \qquad V_1 := \mathrm{Var}[-\log p(X_1|X_2)]$$

$$R_2 \geq H(X_2|X_1) + \sqrt{V_2/n}\, Q^{-1}(\epsilon), \qquad V_2 := \mathrm{Var}[-\log p(X_2|X_1)]$$

$$R_1 + R_2 \geq H(X_1, X_2) + \sqrt{V_3/n}\, Q^{-1}(\epsilon), \qquad V_3 := \mathrm{Var}[-\log p(X_1, X_2)]$$

Simple Slepian-Wolf Outer Bound



$$R_1 \geq H(X_1|X_2) + \sqrt{V_1/n}\, Q^{-1}(\epsilon), \qquad V_1 := \mathrm{Var}[-\log p(X_1|X_2)]$$

$$R_2 \geq H(X_2|X_1) + \sqrt{V_2/n}\, Q^{-1}(\epsilon), \qquad V_2 := \mathrm{Var}[-\log p(X_2|X_1)]$$

$$R_1 + R_2 \geq H(X_1, X_2) + \sqrt{V_3/n}\, Q^{-1}(\epsilon), \qquad V_3 := \mathrm{Var}[-\log p(X_1, X_2)]$$

Transforming the point-to-point result

$$R \geq H(X) + \sqrt{\frac{V}{n}} Q^{-1}(\epsilon)$$

Transforming the point-to-point result

$$R \geq H(X) + \sqrt{\frac{V}{n}} Q^{-1}(\epsilon) \quad \Leftrightarrow \quad R \in H(X) + \frac{1}{\sqrt{n}} \mathcal{S}_V(\epsilon)$$

where the set

$$\mathcal{S}_V(\epsilon) = \{z \in \mathbb{R} : \mathbb{P}(Z \leq z) \geq 1 - \epsilon\}$$

and the random variable $Z \sim \mathcal{N}(0, V)$.

Transforming the point-to-point result

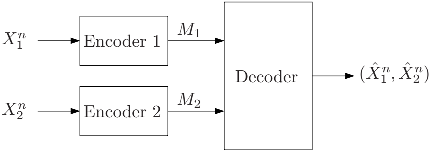
$$R \geq H(X) + \sqrt{\frac{V}{n}} Q^{-1}(\epsilon) \quad \Leftrightarrow \quad R \in H(X) + \frac{1}{\sqrt{n}} \mathcal{S}_V(\epsilon)$$

where the set

$$\mathcal{S}_V(\epsilon) = \{z \in \mathbb{R} : \mathbb{P}(Z \leq z) \geq 1 - \epsilon\}$$

and the random variable $Z \sim \mathcal{N}(0, V)$.

What's the analogue of this result for the two-encoder Slepian-Wolf problem?



Slepian-Wolf Dispersion Result

- Define entropy density vector

$$\mathbf{h}(x_1, x_2) := \begin{bmatrix} -\log p_{X_1|X_2}(x_1|x_2) \\ -\log p_{X_2|X_1}(x_2|x_1) \\ -\log p_{X_1,X_2}(x_1, x_2) \end{bmatrix}$$

- Note that $\mathbb{E}[\mathbf{h}(X_1, X_2)] = [H(X_1|X_2), H(X_2|X_2), H(X_1, X_2)]^T$

Slepian-Wolf Dispersion Result

- Define entropy density vector

$$\mathbf{h}(x_1, x_2) := \begin{bmatrix} -\log p_{X_1|X_2}(x_1|x_2) \\ -\log p_{X_2|X_1}(x_2|x_1) \\ -\log p_{X_1,X_2}(x_1, x_2) \end{bmatrix}$$

- Note that $\mathbb{E}[\mathbf{h}(X_1, X_2)] = [H(X_1|X_2), H(X_2|X_2), H(X_1, X_2)]^T$
- Define the dispersion matrix $\mathbf{V} := \mathrm{Cov}_{X_1,X_2}(\mathbf{h}(X_1, X_2))$

Slepian-Wolf Dispersion Result

- Define entropy density vector

$$\mathbf{h}(x_1,x_2) := \begin{bmatrix} -\log p_{X_1|X_2}(x_1|x_2) \\ -\log p_{X_2|X_1}(x_2|x_1) \\ -\log p_{X_1,X_2}(x_1,x_2) \end{bmatrix}$$

- Note that $\mathbb{E}[\mathbf{h}(X_1,X_2)] = [H(X_1|X_2), H(X_2|X_2), H(X_1,X_2)]^T$
- Define the dispersion matrix $\mathbf{V} := \mathrm{Cov}_{X_1,X_2}(\mathbf{h}(X_1,X_2))$

Theorem (T.-Kosut (2012))

The (n,ϵ)-optimal rate region for the Slepian-Wolf problem is

$$\begin{bmatrix} R_1 \\ R_2 \\ R_1+R_2 \end{bmatrix} \in \begin{bmatrix} H(X_1|X_2) \\ H(X_2|X_1) \\ H(X_1,X_2) \end{bmatrix} + \frac{1}{\sqrt{n}} \mathcal{S}_{\mathbf{V}}(\epsilon) \pm O\left(\frac{\log n}{n}\right)\mathbf{1}$$

and $\mathcal{S}_{\mathbf{V}}(\epsilon) := \{\mathbf{z} \in \mathbb{R}^3 : \mathbb{P}(\mathbf{Z} \le \mathbf{z}) \ge 1-\epsilon\}$ *with* $\mathbf{Z} \sim \mathcal{N}(\mathbf{0},\mathbf{V})$.

Slepian-Wolf Dispersion: Proof of Direct Part

Direct part:

- Random binning + Joint minimum empirical entropy decoding similar to error exponent analysis in Csiszár and Körner
- Multidimensional Berry-Essèen Theorem

Slepian-Wolf Dispersion: Proof of Direct Part

Direct part:

- Random binning + Joint minimum empirical entropy decoding similar to error exponent analysis in Csiszár and Körner
- Multidimensional Berry-Essèen Theorem

Theorem (Bentkus 2003)

Let $\mathbf{X}_1, \ldots, \mathbf{X}_n$ be independent and identically distributed random vectors in $\mathbb{R}^d$ with

$$\mathbb{E}(\mathbf{X}_1) = \mathbf{0}_d, \qquad \mathrm{Cov}(\mathbf{X}_1) = \mathbf{I}_d, \qquad \mathbb{E}\|\mathbf{X}_1\|_2^3 = \xi$$

Slepian-Wolf Dispersion: Proof of Direct Part

Direct part:

- Random binning + Joint minimum empirical entropy decoding similar to error exponent analysis in Csiszár and Körner
- Multidimensional Berry-Essèen Theorem

Theorem (Bentkus 2003)

Let $\mathbf{X}_1, \ldots, \mathbf{X}_n$ *be independent and identically distributed random vectors in* $\mathbb{R}^d$ *with*

$$\mathbb{E}(\mathbf{X}_1) = \mathbf{0}_d, \qquad \mathrm{Cov}(\mathbf{X}_1) = \mathbf{I}_d, \qquad \mathbb{E}\|\mathbf{X}_1\|_2^3 = \xi$$

Then, for every $n \in \mathbb{N}$ *with* $\mathbf{Z} \sim \mathcal{N}(\mathbf{0}_d, \mathbf{I}_d)$,

$$\sup_{\substack{\mathcal{C} \subset \mathbb{R}^d: \\ \mathcal{C} \text{ measurable,convex}}} \left| \mathbb{P}\left(\frac{1}{\sqrt{n}} \sum_{i=1}^n \mathbf{X}_i \in \mathcal{C} \right) - \mathbb{P}(\mathbf{Z} \in \mathcal{C}) \right| \le \frac{400 d^{1/4} \xi}{\sqrt{n}}$$

Slepian-Wolf Dispersion: Proof of Converse Part

Converse part:

- Strong converse in Han's information spectrum book – A result by Miyake and Kanaya (1995)

Slepian-Wolf Dispersion: Proof of Converse Part

Converse part:

- Strong converse in Han's information spectrum book – A result by Miyake and Kanaya (1995)
- Every (n, M_1, M_2, ϵ) SW code has to satisfy

$$\epsilon \geq \mathbb{P}\Bigg[-\frac{1}{n}\log p_{X_1^n|X_2^n}(X_1^n|X_2^n) \geq \frac{1}{n}\log M_1 + \gamma$$

$$\text{or } -\frac{1}{n}\log p_{X_2^n|X_1^n}(X_2^n|X_1^n) \geq \frac{1}{n}\log M_2 + \gamma$$

$$\text{or } -\frac{1}{n}\log p_{X_1^n,X_2^n}(X_1^n,X_2^n) \geq \frac{1}{n}\log(M_1M_2) + \gamma \Bigg] - 3(2^{-n\gamma})$$

$$= 1 - \mathbb{P}\left(\mathbf{h}(X_1^n, X_2^n) \leq \mathbf{R} + \gamma\mathbf{1}\right) - 3(2^{-n\gamma}), \qquad \mathbf{R} = [R_1, R_2, R_1 + R_2]^T$$

Slepian-Wolf Dispersion: Proof of Converse Part

Converse part:

- Strong converse in Han's information spectrum book – A result by Miyake and Kanaya (1995)
- Every (n, M_1, M_2, ϵ) SW code has to satisfy

$$\epsilon \geq \mathbb{P}\bigg[-\frac{1}{n}\log p_{X_1^n|X_2^n}(X_1^n|X_2^n) \geq \frac{1}{n}\log M_1 + \gamma$$
$$\text{or } -\frac{1}{n}\log p_{X_2^n|X_1^n}(X_2^n|X_1^n) \geq \frac{1}{n}\log M_2 + \gamma$$
$$\text{or } -\frac{1}{n}\log p_{X_1^n,X_2^n}(X_1^n,X_2^n) \geq \frac{1}{n}\log (M_1 M_2) + \gamma \bigg] - 3(2^{-n\gamma})$$
$$= 1 - \mathbb{P}\left(\mathbf{h}(X_1^n, X_2^n) \leq \mathbf{R} + \gamma\mathbf{1}\right) - 3(2^{-n\gamma}), \qquad \mathbf{R} = [R_1, R_2, R_1 + R_2]^T$$

- Set $\gamma = O(\frac{\log n}{n})$ and apply the Multidim. Berry-Essèen Theorem

Our Main Result

The (n,ϵ)-optimal rate region for the Slepian-Wolf problem is

$$\begin{bmatrix} R_1 \\ R_2 \\ R_1 + R_2 \end{bmatrix} \in \begin{bmatrix} H(X_1|X_2) \\ H(X_2|X_1) \\ H(X_1,X_2) \end{bmatrix} + \frac{1}{\sqrt{n}} \mathcal{S}_{\mathbf{V}}(\epsilon) \pm O\left(\frac{\log n}{n}\right) \mathbf{1}$$

and $\mathcal{S}_{\mathbf{V}}(\epsilon) := \{\mathbf{z} \in \mathbb{R}^3 : \mathbb{P}(\mathbf{Z} \le \mathbf{z}) \ge 1 - \epsilon\}$ with $\mathbf{Z} \sim \mathcal{N}(\mathbf{0}, \mathbf{V})$.

Our Main Result

The (n, ϵ)-optimal rate region for the Slepian-Wolf problem is

$$\begin{bmatrix} R_1 \\ R_2 \\ R_1 + R_2 \end{bmatrix} \in \begin{bmatrix} H(X_1|X_2) \\ H(X_2|X_1) \\ H(X_1, X_2) \end{bmatrix} + \frac{1}{\sqrt{n}} \mathcal{S}_{\mathbf{V}}(\epsilon) \pm O\left(\frac{\log n}{n}\right) \mathbf{1}$$

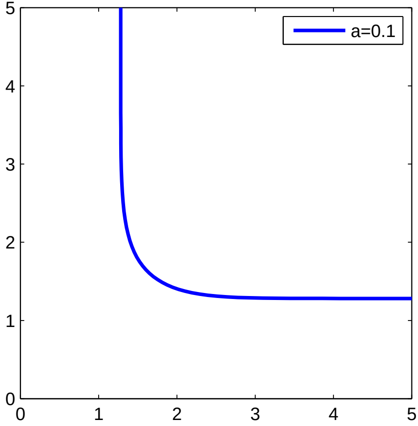
and $\mathcal{S}_{\mathbf{V}}(\epsilon) := \{\mathbf{z} \in \mathbb{R}^3 : \mathbb{P}(\mathbf{Z} \le \mathbf{z}) \ge 1 - \epsilon\}$ with $\mathbf{Z} \sim \mathcal{N}(\mathbf{0}, \mathbf{V})$.

What's this special set $\mathcal{S}_{\mathbf{V}}(\epsilon)$?

What's this special set $\mathcal{S}_{\mathbf{V}}(\epsilon)$?

$\mathcal{S}_{\mathbf{V}}(\epsilon) := \{\mathbf{z} \in \mathbb{R}^3 : \mathbb{P}(\mathbf{Z} \le \mathbf{z}) \ge 1 - \epsilon\}$ where $\mathbf{Z} \sim \mathcal{N}(\mathbf{0}, \mathbf{V})$

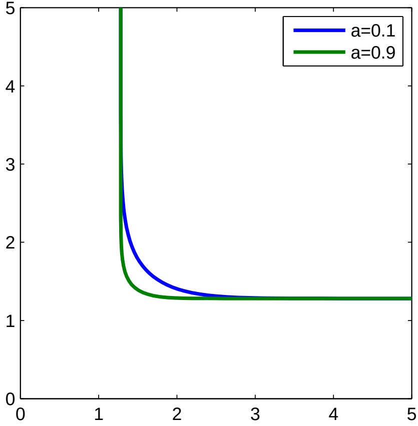
$$\mathbf{V} = \begin{bmatrix} 1 & a \\ a & 1 \end{bmatrix}, \quad a = 0.1$$



What's this special set $\mathcal{S}_{\mathbf{V}}(\epsilon)$?

$\mathcal{S}_{\mathbf{V}}(\epsilon) := \{\mathbf{z} \in \mathbb{R}^3 : \mathbb{P}(\mathbf{Z} \le \mathbf{z}) \ge 1 - \epsilon\}$ where $\mathbf{Z} \sim \mathcal{N}(\mathbf{0}, \mathbf{V})$

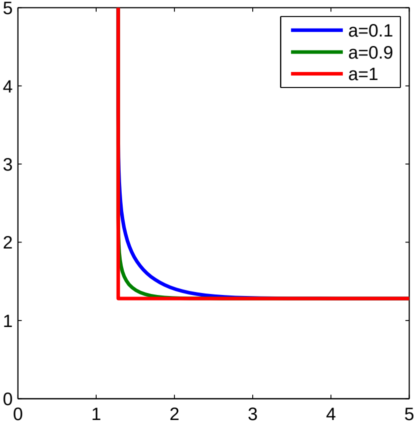
$$\mathbf{V} = \begin{bmatrix} 1 & a \\ a & 1 \end{bmatrix}, \quad a = 0.9$$



What's this special set $\mathcal{S}_{\mathbf{V}}(\epsilon)$?

$\mathcal{S}_{\mathbf{V}}(\epsilon) := \{\mathbf{z} \in \mathbb{R}^3 : \mathbb{P}(\mathbf{Z} \le \mathbf{z}) \ge 1 - \epsilon\}$ where $\mathbf{Z} \sim \mathcal{N}(\mathbf{0}, \mathbf{V})$

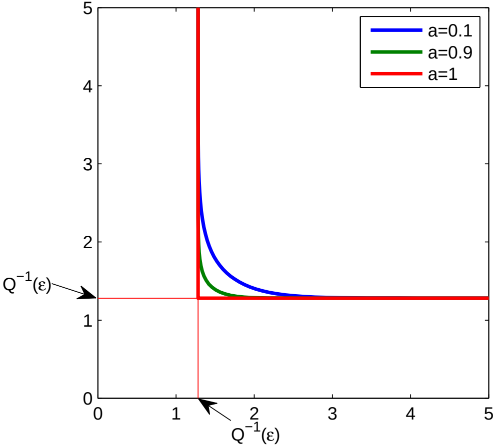
$$\mathbf{V} = \begin{bmatrix} 1 & a \\ a & 1 \end{bmatrix}, \quad a = 1.0$$



What's this special set $\mathcal{S}_{\mathbf{V}}(\epsilon)$?

$\mathcal{S}_{\mathbf{V}}(\epsilon) := \{\mathbf{z} \in \mathbb{R}^3 : \mathbb{P}(\mathbf{Z} \le \mathbf{z}) \ge 1 - \epsilon\}$ where $\mathbf{Z} \sim \mathcal{N}(\mathbf{0}, \mathbf{V})$

$$\mathbf{V} = \begin{bmatrix} 1 & a \\ a & 1 \end{bmatrix}$$

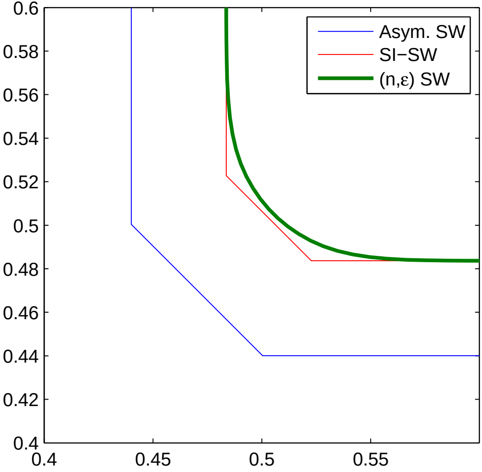


(n, ϵ)-Slepian-Wolf Region

$$p(x_1, x_2) = \begin{bmatrix} 0.7 & 0.1 \\ 0.1 & 0.1 \end{bmatrix}$$

$$\epsilon = 0.1$$

$$n = 300$$

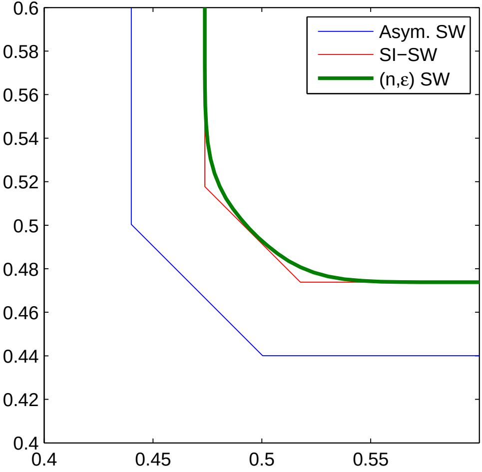


(n, ϵ)-Slepian-Wolf Region

$$p(x_1, x_2) = \begin{bmatrix} 0.7 & 0.1 \\ 0.1 & 0.1 \end{bmatrix}$$

$$\epsilon = 0.1$$

$$n = 500$$

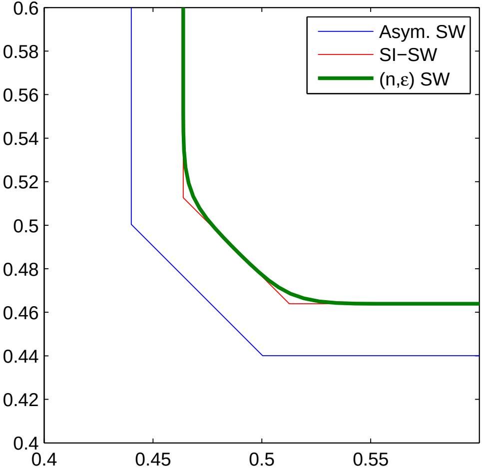


(n, ϵ)-Slepian-Wolf Region

$$p(x_1, x_2) = \begin{bmatrix} 0.7 & 0.1 \\ 0.1 & 0.1 \end{bmatrix}$$

$$\epsilon = 0.1$$

$$n = 1000$$

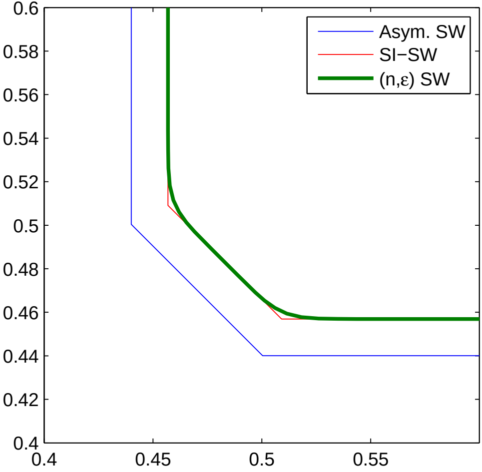


(n, ϵ)-Slepian-Wolf Region

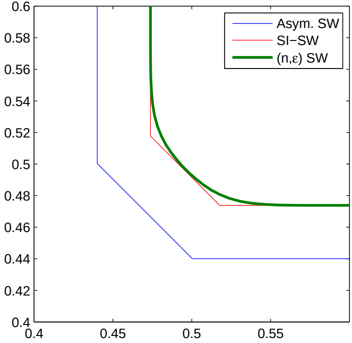
$$p(x_1, x_2) = \begin{bmatrix} 0.7 & 0.1 \\ 0.1 & 0.1 \end{bmatrix}$$

$$\epsilon = 0.1$$

$$n = 2000$$

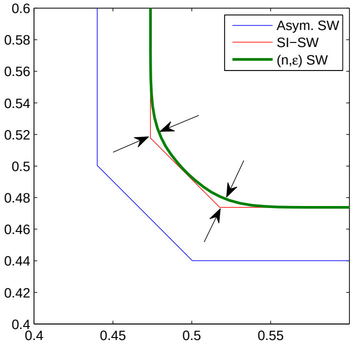


(n, ϵ)-Slepian-Wolf vs (n, ϵ)-side-information SW



Proposition (T.-Kosut (2012))

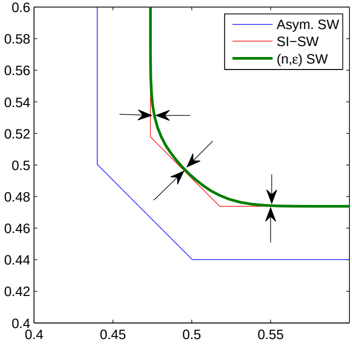
(n, ϵ)-Slepian-Wolf vs (n, ϵ)-side-information SW



Proposition (T.-Kosut (2012))

- *At corner points, distance is* $\frac{c}{\sqrt{n}}$ *for some* $0 < c < \infty$.

(n,ϵ)-Slepian-Wolf vs (n,ϵ)-side-information SW



Proposition (T.-Kosut (2012))

- *At corner points, distance is* $\frac{c}{\sqrt{n}}$ *for some* $0 < c < \infty$*.*
- *Away from corner points, distance is* $\exp(-\Theta(n))$

Comparison to Error Exponent Analysis

- Given a pair of rates in the interior of the SW region, we can ask what blocklength is needed to achieve a tolerable error probability

Comparison to Error Exponent Analysis

- Given a pair of rates in the interior of the SW region, we can ask what blocklength is needed to achieve a tolerable error probability
- Gallager (1976) and Kosholev (1977) proved that

$$P_{\mathrm{e}}^{(n)} \le 3 \exp\{-nE(R_1, R_2)\}$$

where $E(R_1, R_2) > 0$ iff (R_1, R_2) is in the interior of the SW region

Comparison to Error Exponent Analysis

- Given a pair of rates in the interior of the SW region, we can ask what blocklength is needed to achieve a tolerable error probability
- Gallager (1976) and Kosholev (1977) proved that

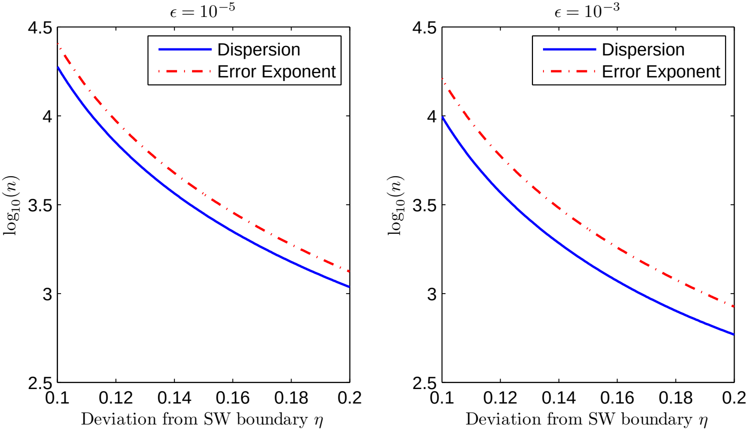
$$P_{\mathrm{e}}^{(n)} \le 3 \exp\{-nE(R_1, R_2)\}$$

  where $E(R_1, R_2) > 0$ iff (R_1, R_2) is in the interior of the SW region
- Compare the two methods

Dispersion: $$n_{\mathrm{D}} := \min\left\{n : \mathbb{P}(\mathbf{Z} \le \sqrt{n}(\mathbf{R} - \mathbf{H})) \ge 1 - \epsilon\right\}$$

Error Exponent: $$n_{\mathrm{E}} := \left\lceil \frac{1}{E(R_1, R_2)} \log\left(\frac{3}{\epsilon}\right) \right\rceil$$

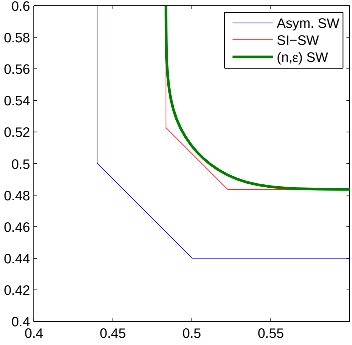
Comparison to Error Exponent Analysis



Prediction of required blocklength n by dispersion analysis (the Gaussian approximation) is much smaller than error exponent analysis!

Summary

- Characterization of (n, ϵ)-Slepian-Wolf rate region up to $O(\frac{\log n}{n})$
- All existing Slepian-Wolf results are corollaries of our theorem
- Required use of dispersion matrix
- Region differs substantively from the SI-SW only at corner points



Extensions

- Achievability technique is exceedingly general
- Can be applied to get inner bounds to the discrete memoryless
  1. Multiple-access channel
  2. Asymmetric broadcast channel (with degraded message sets)
  3. Interference channel
  4. Transmitting correlated sources over a MAC (Cover-El Gamal-Salehi, 1980)
- See preprint on arXiv.

Extensions

- Achievability technique is exceedingly general
- Can be applied to get inner bounds to the discrete memoryless
  1. Multiple-access channel
  2. Asymmetric broadcast channel (with degraded message sets)
  3. Interference channel
  4. Transmitting correlated sources over a MAC (Cover-El Gamal-Salehi, 1980)
- See preprint on arXiv.
- There are other inner bounds for the (n, ϵ)-region for the MAC:

Extensions

- Achievability technique is exceedingly general
- Can be applied to get inner bounds to the discrete memoryless
  1. Multiple-access channel
  2. Asymmetric broadcast channel (with degraded message sets)
  3. Interference channel
  4. Transmitting correlated sources over a MAC (Cover-El Gamal-Salehi, 1980)
- See preprint on arXiv.
- There are other inner bounds for the (n, ϵ)-region for the MAC:
  - Huang and Moulin (ISIT 2012)
  - Molavianjazi and Laneman (ISIT 2012)

Extensions

- Achievability technique is exceedingly general
- Can be applied to get inner bounds to the discrete memoryless
  1. Multiple-access channel
  2. Asymmetric broadcast channel (with degraded message sets)
  3. Interference channel
  4. Transmitting correlated sources over a MAC (Cover-El Gamal-Salehi, 1980)
- See preprint on arXiv.
- There are other inner bounds for the (n, ϵ)-region for the MAC:
  - Huang and Moulin (ISIT 2012)
  - Molavianjazi and Laneman (ISIT 2012)
- However, other than the Slepian-Wolf problem, it is difficult to use information spectrum methods to get converses for other (n, ϵ)-regions in network information theory.