Time-Division Transmission is Optimal for Covert Communication over Some Broadcast Channels

Vincent Y. F. Tan (NUS)

Joint work with Prof. Si-Hyeon Lee (POSTECH)

National University of Singapore (NUS)

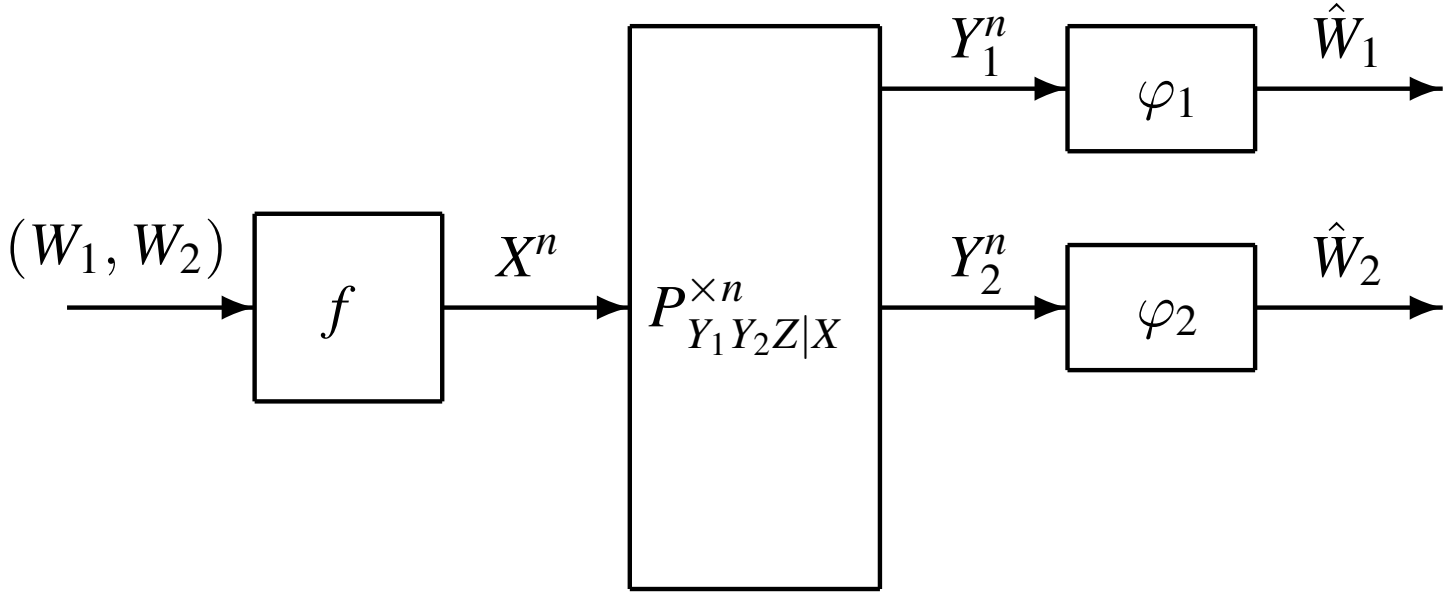
ITW, Guangzhou (Nov 2018)

Problem Setting

- Consider a traditional two-user broadcast channel

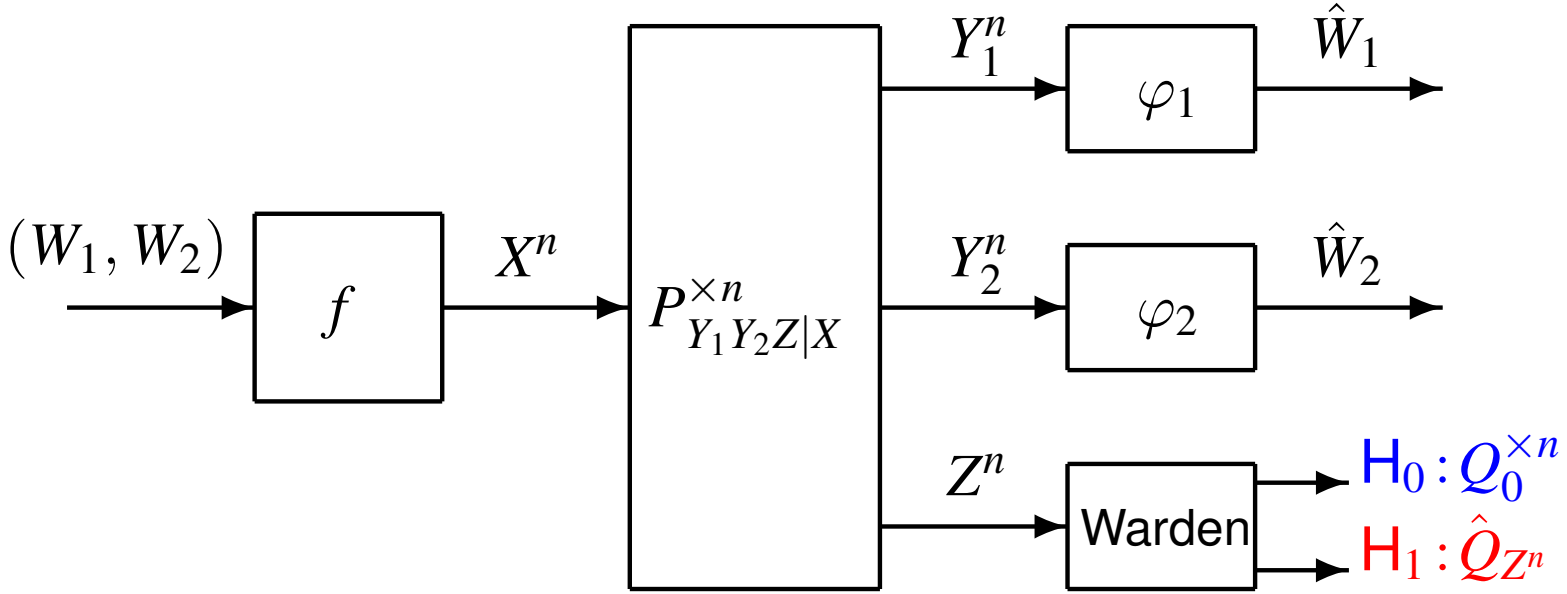
Problem Setting

- Consider a traditional two-user broadcast channel



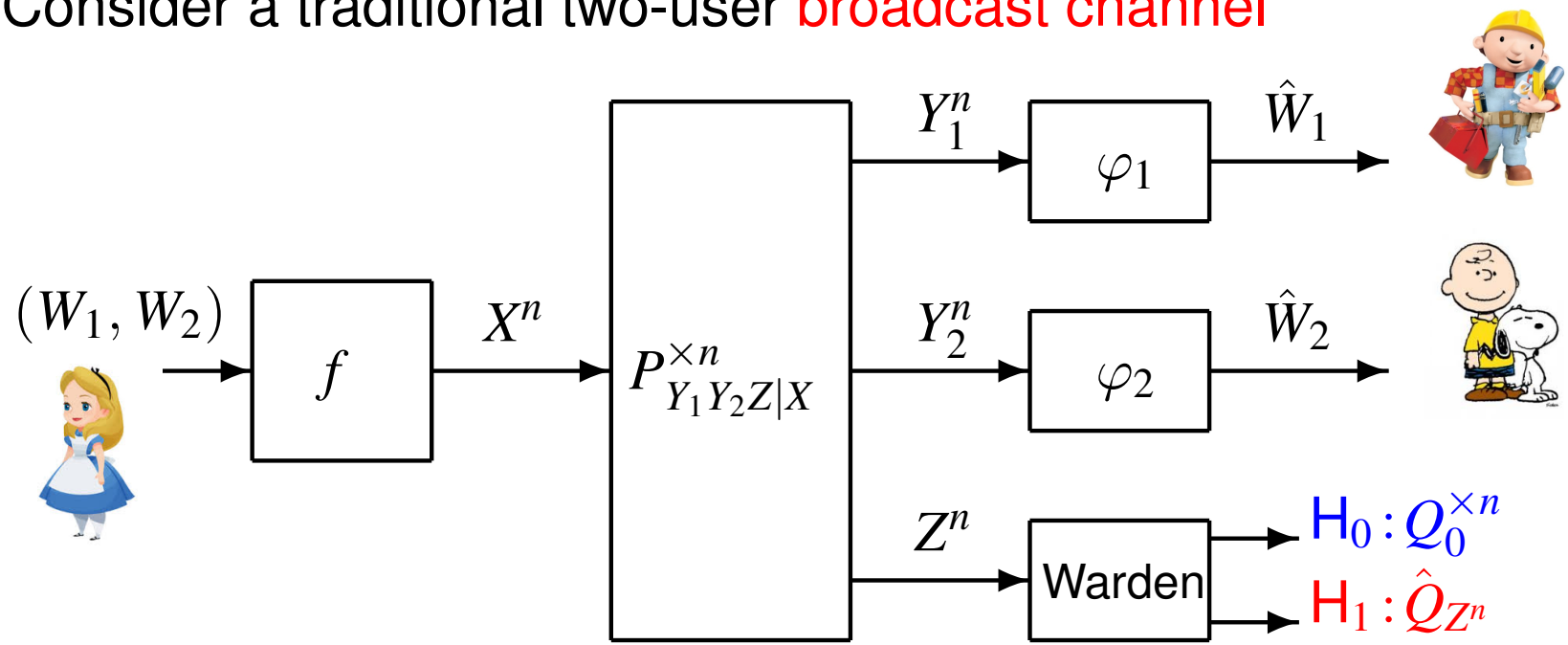
Problem Setting

- Consider a traditional two-user broadcast channel



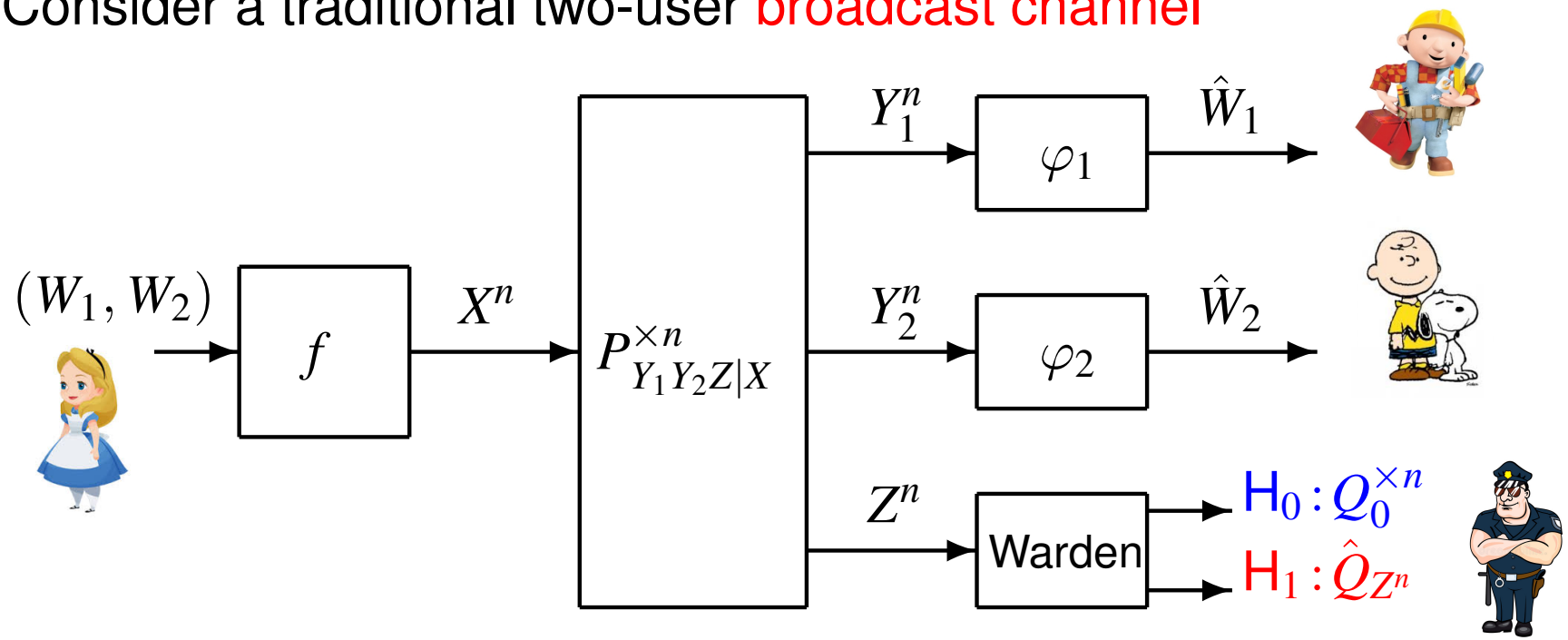
Problem Setting

- Consider a traditional two-user broadcast channel



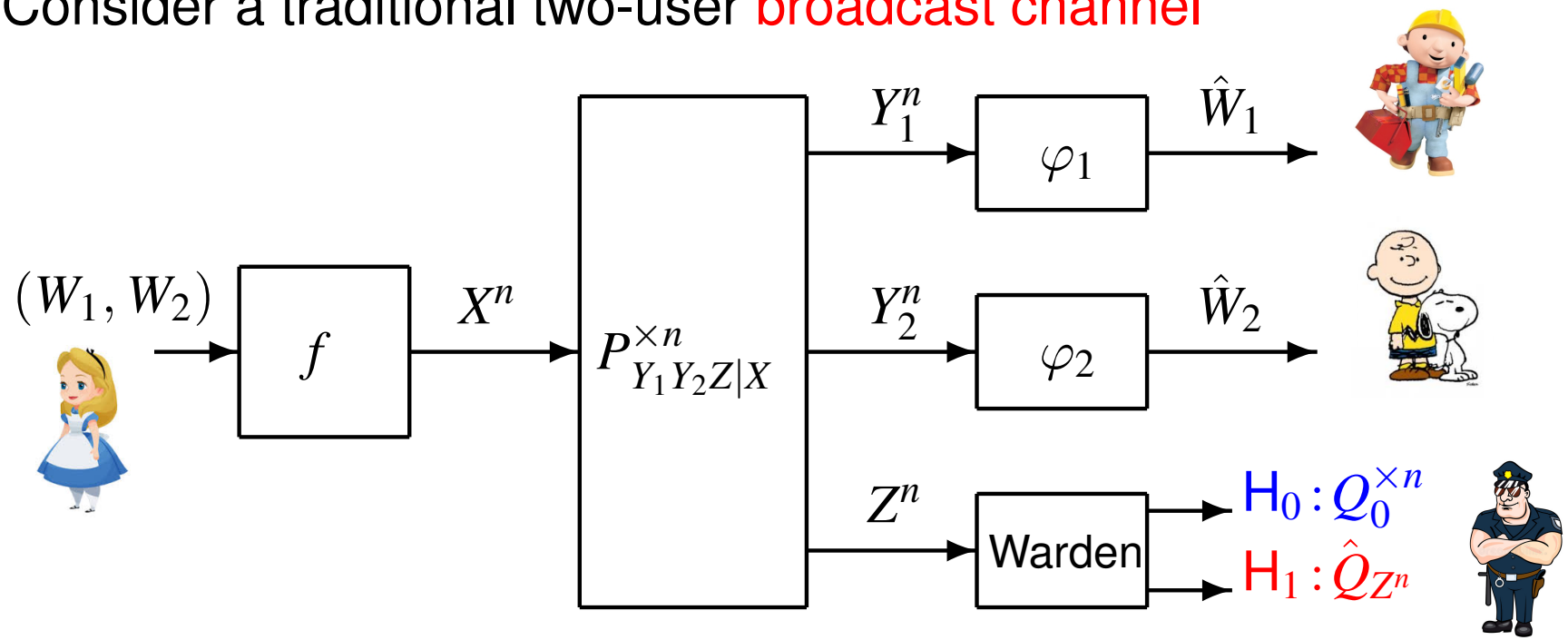
Problem Setting

- Consider a traditional two-user broadcast channel



Problem Setting

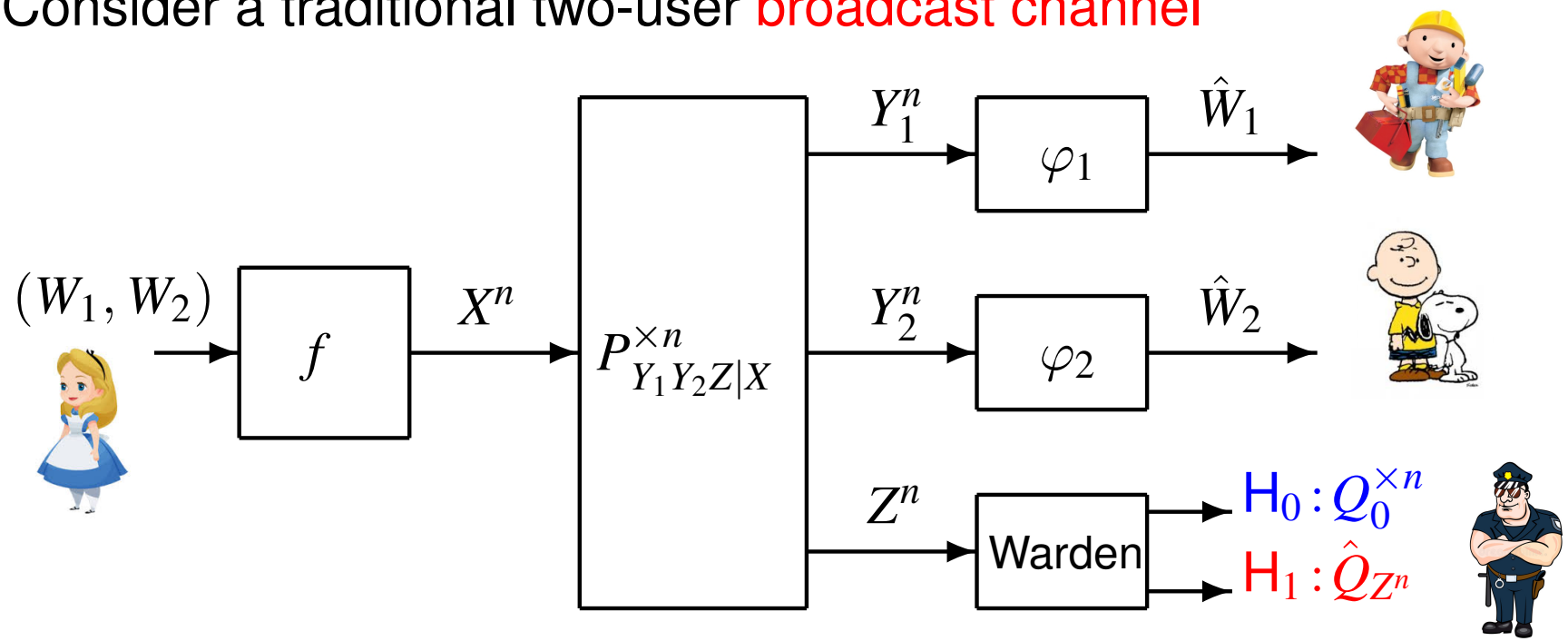
- Consider a traditional two-user broadcast channel



- Covert Communication $\Longleftrightarrow$ Warden should not be able to know whether there is communication

Problem Setting

- Consider a traditional two-user broadcast channel



- Covert Communication $\Longleftrightarrow$ Warden should not be able to know whether there is communication
- Should not be able to distinguish between $\hat{Q}_{Z^n}$ (output dist. induced by a code) and the "no-communication" output dist. $Q_0^{\times n}$.

More on Covert Communications over Noisy Channels

- Innocent symb. $0 \in \mathcal{X}$ inducing warden output dist. $Q_0 = P_{Z|X}(\cdot|0)$

More on Covert Communications over Noisy Channels

- Innocent symb. $0 \in \mathcal{X}$ inducing warden output dist. $Q_0 = P_{Z|X}(\cdot|0)$
- Symb. $1 \in \mathcal{X}$ inducing warden output dist. $Q_1 = P_{Z|X}(\cdot|0)$

More on Covert Communications over Noisy Channels

- Innocent symb. $0 \in \mathcal{X}$ inducing warden output dist. $Q_0 = P_{Z|X}(\cdot|0)$
- Symb. $1 \in \mathcal{X}$ inducing warden output dist. $Q_1 = P_{Z|X}(\cdot|0)$
- Assume $Q_1 \ll Q_0$

More on Covert Communications over Noisy Channels

- Innocent symb. $0 \in \mathcal{X}$ inducing warden output dist. $Q_0 = P_{Z|X}(\cdot|0)$
- Symb. $1 \in \mathcal{X}$ inducing warden output dist. $Q_1 = P_{Z|X}(\cdot|0)$
- Assume $Q_1 \ll Q_0$
- Warden attempts to design optimal detector to distinguish
  - H_0: observed distribution is $Q_0^{\times n}$ (no communication)
  - H_1: observed distribution is $\hat{Q}_{Z^n}$ (communication active)

More on Covert Communications over Noisy Channels

- Innocent symb. $0 \in \mathcal{X}$ inducing warden output dist. $Q_0 = P_{Z|X}(\cdot|0)$
- Symb. $1 \in \mathcal{X}$ inducing warden output dist. $Q_1 = P_{Z|X}(\cdot|0)$
- Assume $Q_1 \ll Q_0$
- Warden attempts to design optimal detector to distinguish
  - H_0: observed distribution is $Q_0^{\times n}$ (no communication)
  - H_1: observed distribution is $\hat{Q}_{Z^n}$ (communication active)
- Optimal performance

$$\pi_{1|0} + \pi_{0|1} = 1 - \frac{1}{2}\left\|Q_0^{\times n} - \hat{Q}_{Z^n}\right\|_1 \geq 1 - \sqrt{D(\hat{Q}_{Z^n}\|Q_0^{\times n})}$$

More on Covert Communications over Noisy Channels

- Innocent symb. $0 \in \mathcal{X}$ inducing warden output dist. $Q_0 = P_{Z|X}(\cdot|0)$
- Symb. $1 \in \mathcal{X}$ inducing warden output dist. $Q_1 = P_{Z|X}(\cdot|0)$
- Assume $Q_1 \ll Q_0$
- Warden attempts to design optimal detector to distinguish
  - H_0: observed distribution is $Q_0^{\times n}$ (no communication)
  - H_1: observed distribution is $\hat{Q}_{Z^n}$ (communication active)
- Optimal performance

$$\pi_{1|0} + \pi_{0|1} = 1 - \frac{1}{2}\left\|Q_0^{\times n} - \hat{Q}_{Z^n}\right\|_1 \geq 1 - \sqrt{D(\hat{Q}_{Z^n}\|Q_0^{\times n})}$$

- For covert communication, we want to make $D(\hat{Q}_{Z^n}\|Q_0^{\times n})$ small.

Context

- Growing concern for privacy and confidentiality

Context

- Growing concern for privacy and confidentiality
- Renewed interest in fundamental limits of covert communications:
  - Secure space-time codes [Hero '03]
  - Secure stegosystems [Korzhik et al. '05]
  - $O(\sqrt{n})$ bits over n ch. uses with $O(\sqrt{n}\log n)$ key [Bash et al. '12]
  - Similar to square-root law in steganography [Cachin '04]

Context

- Growing concern for privacy and confidentiality
- Renewed interest in fundamental limits of covert communications:
  - Secure space-time codes [Hero '03]
  - Secure stegosystems [Korzhik et al. '05]
  - $O(\sqrt{n})$ bits over n ch. uses with $O(\sqrt{n}\log n)$ key [Bash et al. '12]
  - Similar to square-root law in steganography [Cachin '04]
- Several extensions and related results:
  - Constants in $O(\sqrt{n})$ term [Wang, Wornell, Zheng '16 and Bloch '16]
  - Second-order [Tahmasbi-Bloch '16]
  - Error exponents [Tahmasbi-Bloch-Tan '17]
  - Multi-user [Arumugam-Bloch '16, '17]

Definition of a code

An $(n, M_{1n}, M_{2n}, \varepsilon, \delta)$-code for the broadcast channel with a warden $P_{Y_1,Y_2,Z|X}$ consists of

- Two message sets $\mathcal{M}_j := \{1, \ldots, M_{jn}\}$ for $j = 1, 2$;
- Two independent messages uniformly distributed over their respective message sets, i.e., $W_j \sim \mathrm{Unif}(\mathcal{M}_j)$ for $j = 1, 2$;

Definition of a code

An $(n, M_{1n}, M_{2n}, \varepsilon, \delta)$-code for the broadcast channel with a warden $P_{Y_1,Y_2,Z|X}$ consists of

- Two message sets $\mathcal{M}_j := \{1, \ldots, M_{jn}\}$ for $j = 1, 2$;
- Two independent messages uniformly distributed over their respective message sets, i.e., $W_j \sim \mathrm{Unif}(\mathcal{M}_j)$ for $j = 1, 2$;
- One encoder $f : \mathcal{M}_1 \times \mathcal{M}_2 \to \mathcal{X}^n$;
- Two decoders $\varphi_j : \mathcal{Y}_j^n \to \mathcal{M}_j$ for $j = 1, 2$;

Definition of a code

An $(n, M_{1n}, M_{2n}, \varepsilon, \delta)$-code for the broadcast channel with a warden $P_{Y_1,Y_2,Z|X}$ consists of

- Two message sets $\mathcal{M}_j := \{1, \ldots, M_{jn}\}$ for $j = 1, 2$;
- Two independent messages uniformly distributed over their respective message sets, i.e., $W_j \sim \mathrm{Unif}(\mathcal{M}_j)$ for $j = 1, 2$;
- One encoder $f : \mathcal{M}_1 \times \mathcal{M}_2 \to \mathcal{X}^n$;
- Two decoders $\varphi_j : \mathcal{Y}_j^n \to \mathcal{M}_j$ for $j = 1, 2$;

such that the following constraints hold:

$$\text{Reliability:} \qquad \Pr\left(\cup_{j=1}^{2} \{\hat{W}_j \neq W_j\} \right) \leq \varepsilon$$

and

$$\text{Covertness:} \qquad D(\hat{Q}_{Z^n} \| Q_0^{\times n}) \leq \delta.$$

Definition of Covert Capacity Region

- $(L_1, L_2) \in \mathbb{R}_+^2$ is (ε, δ)-achievable if there exists a sequence of $(n, M_{1n}, M_{2n}, \varepsilon_n, \delta)$-codes such that

$$\liminf_{n\to\infty} \frac{1}{\sqrt{n\delta}} \log M_{jn} \geq L_j, \quad j = 1, 2,$$

$$\limsup_{n\to\infty} \varepsilon_n \leq \varepsilon.$$

Definition of Covert Capacity Region

- $(L_1, L_2) \in \mathbb{R}_+^2$ is (ε, δ)-achievable if there exists a sequence of $(n, M_{1n}, M_{2n}, \varepsilon_n, \delta)$-codes such that

$$\liminf_{n\to\infty} \frac{1}{\sqrt{n\delta}} \log M_{jn} \geq L_j, \quad j = 1, 2,$$

$$\limsup_{n\to\infty} \varepsilon_n \leq \varepsilon.$$

- The (ε, δ)-covert capacity region $\mathcal{L}_{\varepsilon,\delta} \subset \mathbb{R}_+^2$ is the closure of all (ε, δ)-achievable pairs of (L_1, L_2).

Definition of Covert Capacity Region

- $(L_1, L_2) \in \mathbb{R}_+^2$ is (ε, δ)-achievable if there exists a sequence of $(n, M_{1n}, M_{2n}, \varepsilon_n, \delta)$-codes such that

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

- The (ε, δ)-covert capacity region $\mathcal{L}_{\varepsilon,\delta} \subset \mathbb{R}_+^2$ is the closure of all (ε, δ)-achievable pairs of (L_1, L_2).

- The δ-covert capacity region

$$\mathcal{L}_\delta := \bigcap_{\varepsilon\in(0,1)} \mathcal{L}_{\varepsilon,\delta} = \lim_{\varepsilon\to 0} \mathcal{L}_{\varepsilon,\delta}.$$

Definition of Covert Capacity

- For simplicity, assume binary-input channels, i.e., $\mathcal{X} = \{0, 1\}$

Definition of Covert Capacity

- For simplicity, assume binary-input channels, i.e., $\mathcal{X} = \{0, 1\}$
- Given a DMC with a warden $P_{Y,Z|X}$, the covert capacity [Wang, Wornell, Zheng '16 and Bloch '16] is

$$L^*(P_{Y,Z|X}) := \sqrt{\frac{2D(W(\cdot|1)\|W(\cdot|0))^2}{\chi_2(Q_1\|Q_0)}}$$

where

$$W(\cdot|x) = P_{Y|X}(\cdot|x) \qquad Q_x := P_{Z|X}(\cdot|x), \quad \forall\, x \in \mathcal{X},$$

and

$$\chi_2(Q_1\|Q_0) := \sum_z \frac{(Q_1(z) - Q_0(z))^2}{Q_0(z)}.$$

Assumption on the BC $P_{Y_1,Y_2,Z|X}$

Assumption

Condition 1: Fix a BC with a warden $P_{Y_1,Y_2,Z|X}$.

Let the covert capacities of $P_{Y_1,Z|X}$ and $P_{Y_2,Z|X}$ be L_1^* and L_2^* respectively.

If $L_1^* \geq L_2^*$ assume that

$$\max_{P_X} \frac{I(X;Y_1)}{I(X;Y_2)} \leq \frac{L_1^*}{L_2^*}.$$

Otherwise, if $L_1^* \leq L_2^*$ assume that

$$\max_{P_X} \frac{I(X;Y_2)}{I(X;Y_1)} \leq \frac{L_2^*}{L_1^*}.$$

Discussion of Condition 1

- Easy to check for binary-input BCs:

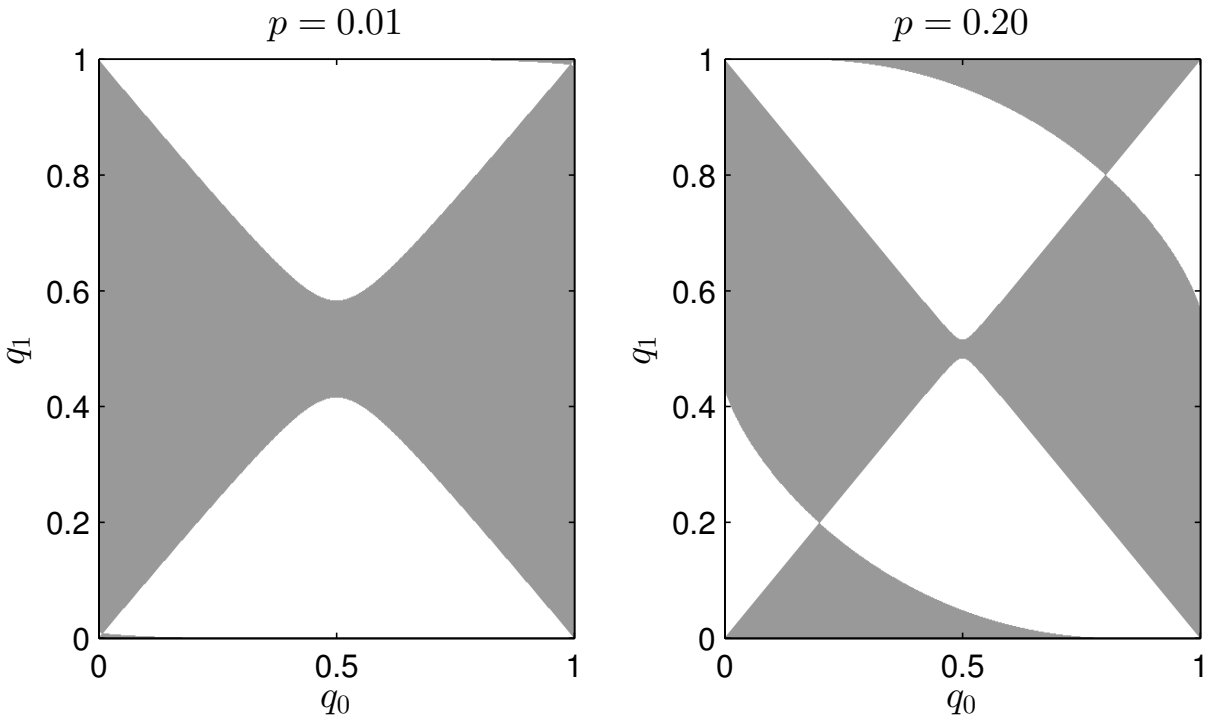
$$\frac{D(W_1\|W_0)}{D(V_1\|V_0)} \leq \min_{\gamma\in[0,1]} \frac{D(W_\gamma\|W_0)}{D(V_\gamma\|V_0)}, \qquad W_\gamma(y) = \sum_{x\in\mathcal{X}} P_\gamma(x)W(y|x).$$

Discussion of Condition 1

- Easy to check for binary-input BCs:

$$\frac{D(W_1\|W_0)}{D(V_1\|V_0)} \leq \min_{\gamma\in[0,1]} \frac{D(W_\gamma\|W_0)}{D(V_\gamma\|V_0)}, \qquad W_\gamma(y) = \sum_{x\in\mathcal{X}} P_\gamma(x)W(y|x).$$

- Let $W = P_{Y_1|X} = \mathrm{BSC}(p)$ and $V = P_{Y_2|X} = \begin{bmatrix} 1-q_0 & q_0 \\ q_1 & 1-q_1 \end{bmatrix}$



Shaded area $\Longrightarrow$ Condition 1 $\surd$

Main Result

Theorem (Tan-Lee (2018))

Assume Condition 1 holds for $P_{Y_1,Y_2,Z|X}$. *For any* $\delta > 0$ *and* $P_{Y_j|X}(\cdot|1) \ll P_{Y_j|X}(\cdot|0)$ *for* $j = 1, 2$,

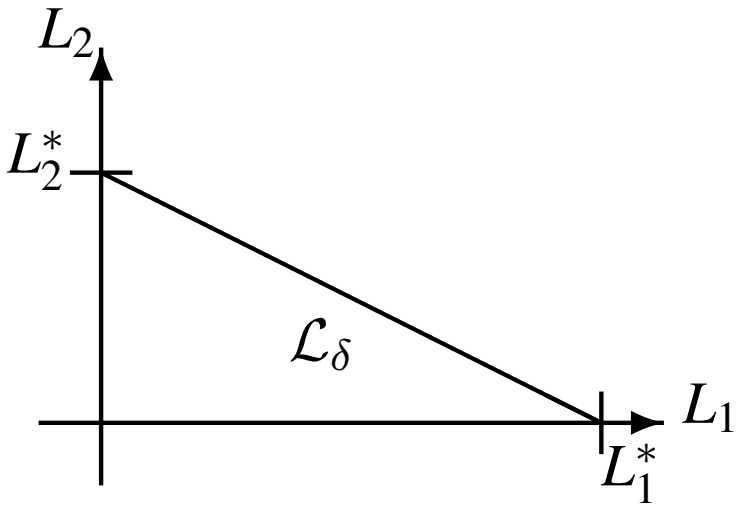
$$\mathcal{L}_\delta = \left\{ (L_1, L_2) \in \mathbb{R}_+^2 : \frac{L_1}{L_1^*} + \frac{L_2}{L_2^*} \leq 1 \right\}.$$

Main Result

Theorem (Tan-Lee (2018))

Assume Condition 1 holds for $P_{Y_1,Y_2,Z|X}$. *For any* $\delta > 0$ *and* $P_{Y_j|X}(\cdot|1) \ll P_{Y_j|X}(\cdot|0)$ *for* $j = 1, 2$,

$$\mathcal{L}_\delta = \left\{ (L_1, L_2) \in \mathbb{R}_+^2 : \frac{L_1}{L_1^*} + \frac{L_2}{L_2^*} \leq 1 \right\}.$$



Effect of Key Size

Theorem (Tan-Lee (2018))

Under Condition 1, the tuple $(L_1, L_2, L_{\text{key}})$ *is achievable if and only if*

$$\frac{L_1}{L_1^*} + \frac{L_2}{L_2^*} \leq 1$$

and

$$L_{\text{key}} \geq \left(\frac{L_1}{L_1^*} + \frac{L_2}{L_2^*} \right) L_Z^* - L_1 - L_2,$$

where

$$L_Z^* = L^*(P_{Z,Z|X})$$

is the self-covert capacity of the channel $X \to Z$.

Effect of Key Size

Theorem (Tan-Lee (2018))

Under Condition 1, the tuple $(L_1, L_2, L_{\text{key}})$ *is achievable if and only if*

$$\frac{L_1}{L_1^*} + \frac{L_2}{L_2^*} \leq 1$$

and

$$L_{\text{key}} \geq \left(\frac{L_1}{L_1^*} + \frac{L_2}{L_2^*} \right) L_Z^* - L_1 - L_2,$$

where

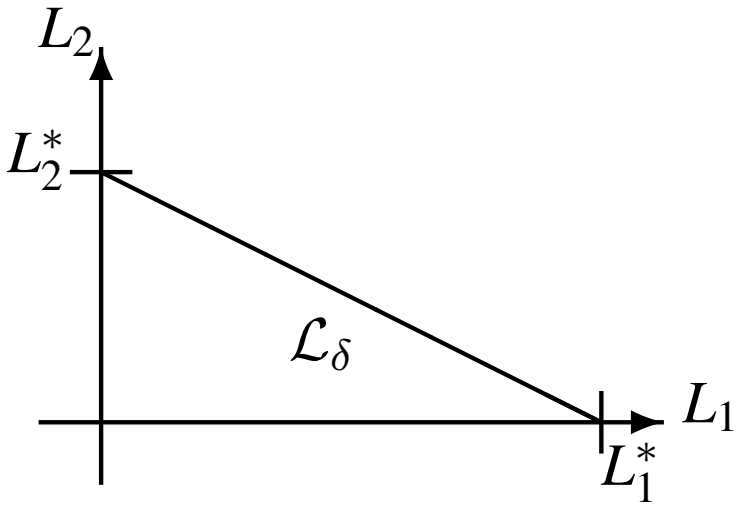
$$L_Z^* = L^*(P_{Z,Z|X})$$

is the self-covert capacity of the channel $X \to Z$.

If we operate on the boundary of the keyless covert capacity region,

$$L_{\text{key}}^* = L_Z^* - L_1 - L_2.$$

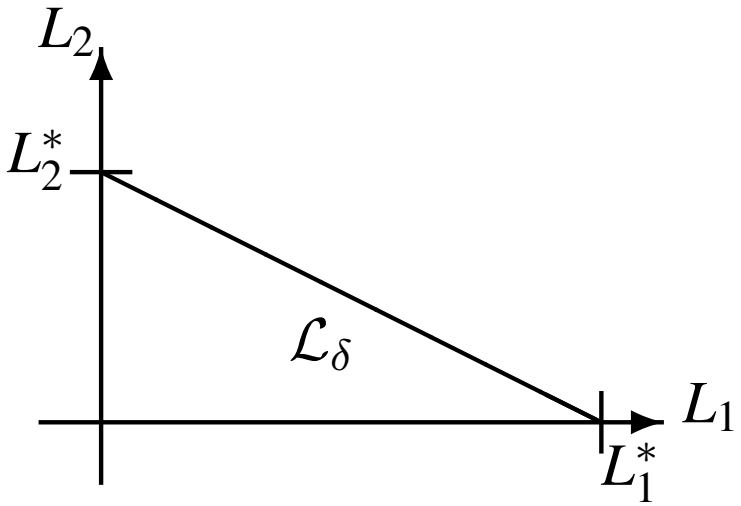
Time-Division is Optimal over Some BCs



- Use an optimal code for $X \to Y_1$ for ρn channel uses. If $\delta' < \delta$,

$$\log M_{1n} \cong \sqrt{\rho n \delta'} L_1^*$$

Time-Division is Optimal over Some BCs



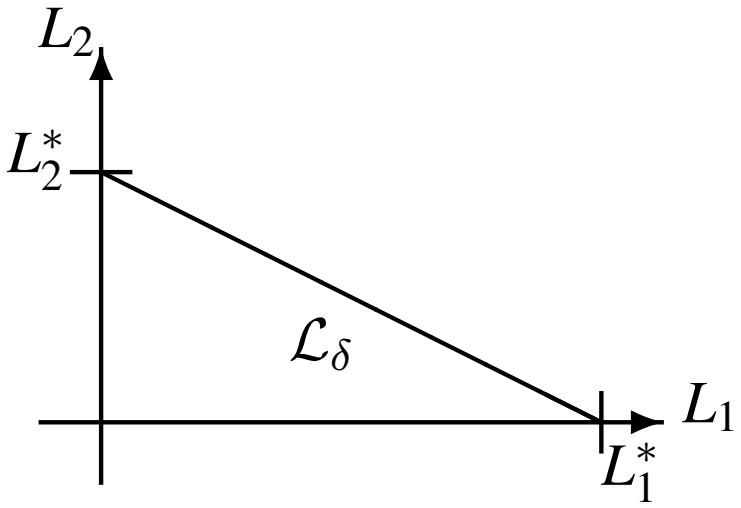
- Use an optimal code for $X \to Y_1$ for ρn channel uses. If $\delta' < \delta$,

$$\log M_{1n} \cong \sqrt{\rho n \delta'} L_1^*$$

- Use another optimal code for $X \to Y_2$ over $(1-\rho)n$ uses. Then,

$$\log M_{2n} \cong \sqrt{(1-\rho) n (\delta - \delta')} L_2^*$$

Time-Division is Optimal over Some BCs



- Use an optimal code for $X \to Y_1$ for ρn channel uses. If $\delta' < \delta$,

$$\log M_{1n} \cong \sqrt{\rho n \delta'} L_1^*$$

- Use another optimal code for $X \to Y_2$ over $(1-\rho)n$ uses. Then,

$$\log M_{2n} \cong \sqrt{(1-\rho)n(\delta-\delta')} L_2^*$$

- Choose $\delta' = (1-\rho)\delta$ achieves the point $(\rho L_1^*, (1-\rho)L_2^*)$.

Time-Division is Optimal for Some BCs: Why?

- Covert communication implies that X^n must have low weight of order $\Theta(\frac{1}{\sqrt{n}})$ [Wang, Wornell, Zheng '16 and Bloch '16], i.e.,

$$|\{i \in [n] : X_i = 1\}| = \Theta(\sqrt{n})$$

Time-Division is Optimal for Some BCs: Why?

- Covert communication implies that X^n must have low weight of order $\Theta(\frac{1}{\sqrt{n}})$ [Wang, Wornell, Zheng '16 and Bloch '16], i.e.,

$$|\{i \in [n] : X_i = 1\}| = \Theta(\sqrt{n})$$

- Hence throughput $\log M_n$ is of the order $\Theta(\sqrt{n})$

Time-Division is Optimal for Some BCs: Why?

- Covert communication implies that X^n must have low weight of order $\Theta(\frac{1}{\sqrt{n}})$ [Wang, Wornell, Zheng '16 and Bloch '16], i.e.,

$$|\{i \in [n] : X_i = 1\}| = \Theta(\sqrt{n})$$

- Hence throughput $\log M_n$ is of the order $\Theta(\sqrt{n})$
- For illustration purposes, consider a BS-BC $P_{Y_1,Y_2|X}$ is such that $P_{Y_j|X}$ for $j = 1, 2$ are BSCs.

Time-Division is Optimal for Some BCs: Why?

- Covert communication implies that X^n must have low weight of order $\Theta(\frac{1}{\sqrt{n}})$ [Wang, Wornell, Zheng '16 and Bloch '16], i.e.,

$$|\{i \in [n] : X_i = 1\}| = \Theta(\sqrt{n})$$

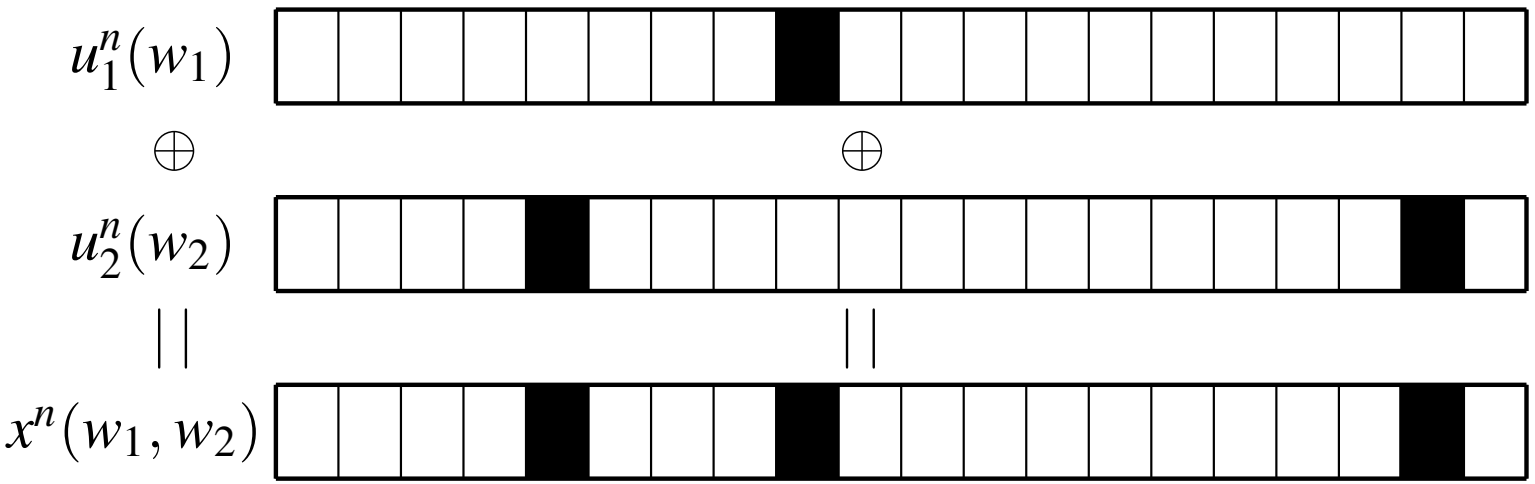
- Hence throughput $\log M_n$ is of the order $\Theta(\sqrt{n})$
- For illustration purposes, consider a BS-BC $P_{Y_1,Y_2|X}$ is such that $P_{Y_j|X}$ for $j = 1, 2$ are BSCs.
- Same intuition for Gaussian broadcast channels
- But use Entropy Power Inequality instead of Mrs. Gerber's Lemma.

Time-Division is Optimal for Some BCs: Why?

- Superposition coding: Cloud center $u_2^n(w_2)$ carries message w_2; Satellite codeword $x^n(w_1, w_2) = u_1^n(w_1) \oplus u_2^n(w_2)$ carries (w_1, w_2)

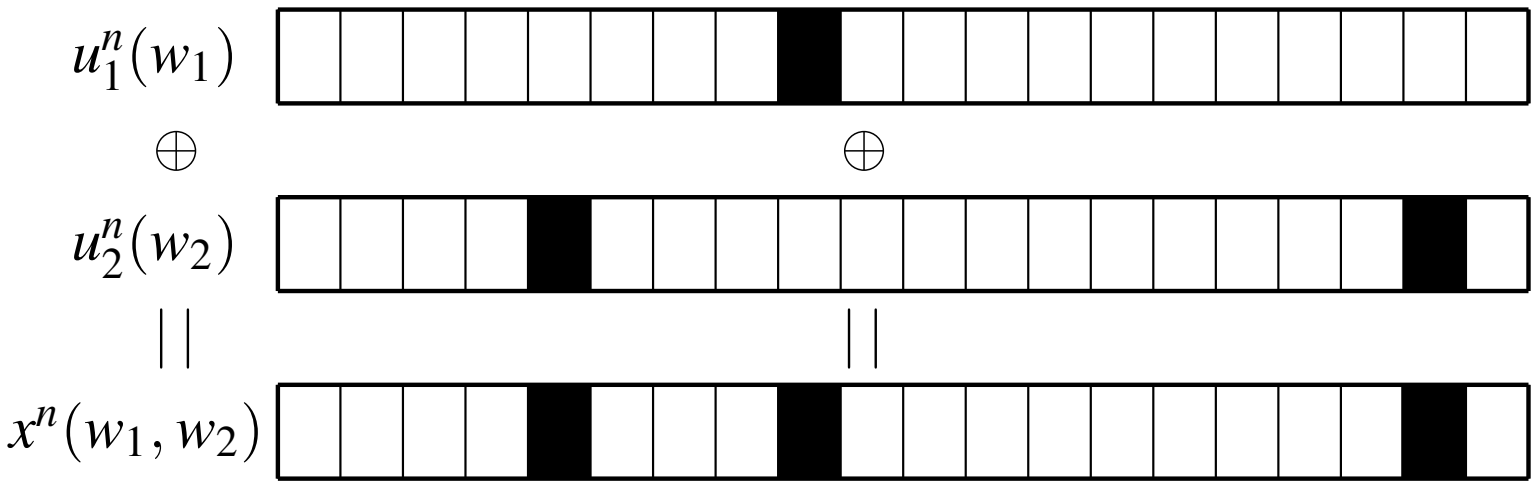
Time-Division is Optimal for Some BCs: Why?

- Superposition coding: Cloud center $u_2^n(w_2)$ carries message w_2; Satellite codeword $x^n(w_1, w_2) = u_1^n(w_1) \oplus u_2^n(w_2)$ carries (w_1, w_2)



Time-Division is Optimal for Some BCs: Why?

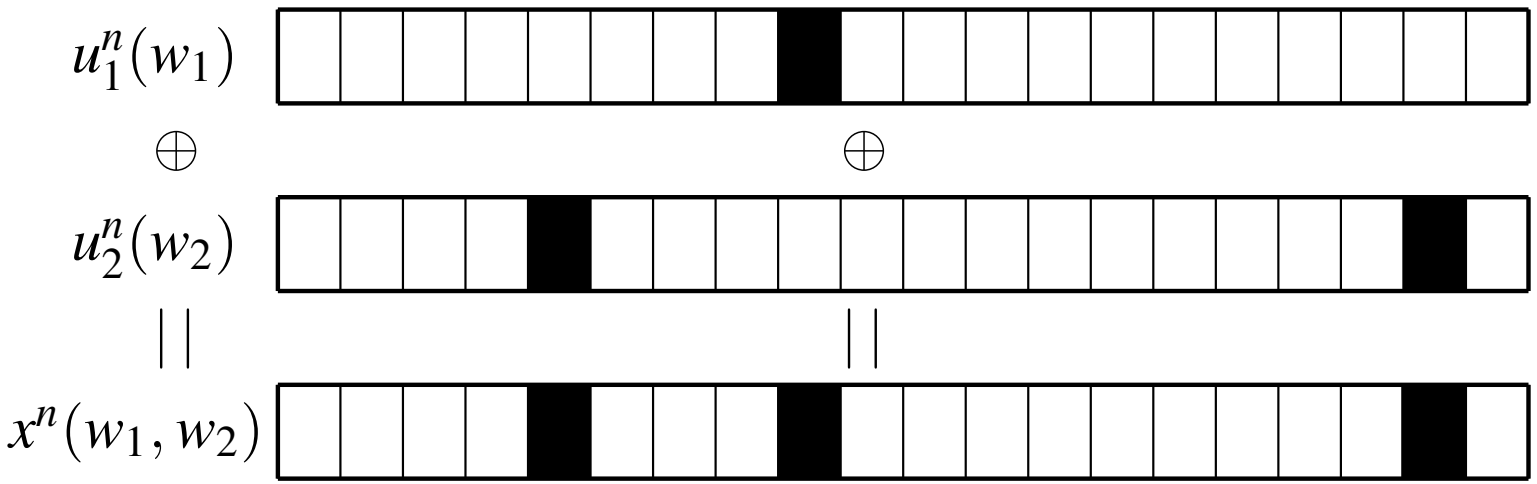
- Superposition coding: Cloud center $u_2^n(w_2)$ carries message w_2; Satellite codeword $x^n(w_1, w_2) = u_1^n(w_1) \oplus u_2^n(w_2)$ carries (w_1, w_2)



- Since $x^n(w_1, w_2)$ has low weight (say α_n) and $u_1^n(w_1)$ and $u_2^n(w_2)$ are randomly chosen, locations of 1's in $u_1^n(w_1)$ and $u_2^n(w_2)$ are not likely to overlap.

Time-Division is Optimal for Some BCs: Why?

- Superposition coding: Cloud center $u_2^n(w_2)$ carries message w_2; Satellite codeword $x^n(w_1, w_2) = u_1^n(w_1) \oplus u_2^n(w_2)$ carries (w_1, w_2)



- Since $x^n(w_1, w_2)$ has low weight (say α_n) and $u_1^n(w_1)$ and $u_2^n(w_2)$ are randomly chosen, locations of 1's in $u_1^n(w_1)$ and $u_2^n(w_2)$ are not likely to overlap.

- Assume weight of $u_1^n(w_1)$ is $\rho\alpha_n$ and that of $u_2^n(w_2)$ is $(1-\rho)\alpha_n$

Time-Division is Optimal for Some BCs: Why?

- Consider BSBCs

$$Y_1 = X \oplus N_1, \qquad Y_2 = X \oplus N_2, \quad N_j \sim \text{Bern}(p_j), \quad p_2 \geq p_1$$

Time-Division is Optimal for Some BCs: Why?

- Consider BSBCs

$$Y_1 = X \oplus N_1, \qquad Y_2 = X \oplus N_2, \quad N_j \sim \text{Bern}(p_j), \quad p_2 \geq p_1$$

- Put

$$\text{wt}(u_1^n(w_1)) = \rho\alpha_n n, \quad \text{and} \quad \text{wt}(u_2^n(w_2)) = (1-\rho)\alpha_n n,$$

the superposition coding inner bound with $X = U_1 \oplus U_2$ reads $(U_2 - X - Y_1 - Y_2)$

$$R_1 \leq I(X; Y_1|U_2) = I(U_1; Y_1 \oplus N_1) \approx \rho\alpha_n L_1^*$$
$$R_2 \leq I(U_2; Y_2) = I(U_2; U_2 \oplus \tilde{N}_2) \approx (1-\rho)\alpha_n L_2^*$$

Time-Division is Optimal for Some BCs: Why?

- Consider BSBCs

$$Y_1 = X \oplus N_1, \qquad Y_2 = X \oplus N_2, \quad N_j \sim \text{Bern}(p_j), \quad p_2 \geq p_1$$

- Put

$$\text{wt}(u_1^n(w_1)) = \rho\alpha_n n, \quad \text{and} \quad \text{wt}(u_2^n(w_2)) = (1-\rho)\alpha_n n,$$

the superposition coding inner bound with $X = U_1 \oplus U_2$ reads $(U_2 - X - Y_1 - Y_2)$

$$R_1 \leq I(X; Y_1|U_2) = I(U_1; Y_1 \oplus N_1) \approx \rho\alpha_n L_1^*$$
$$R_2 \leq I(U_2; Y_2) = I(U_2; U_2 \oplus \tilde{N}_2) \approx (1-\rho)\alpha_n L_2^*$$

- Hence, we can write

$$\frac{R_1}{L_1^*} + \frac{R_2}{L_2^*} \lessapprox \alpha_n$$

El Gamal's Converse for More Capable BCs

El Gamal's Converse for More Capable BCs

Lemma (El Gamal (1979))

Every $(n, M_{1n}, M_{2n}, \varepsilon_n)$-code for any BC satisfies

$$(\log M_{1n})(1-\varepsilon_n) - 1 \leq \sum_{i=1}^{n} I(U_{1i}; Y_{1i})$$

$$(\log M_{2n})(1-\varepsilon_n) - 1 \leq \sum_{i=1}^{n} I(U_{2i}; Y_{2i})$$

El Gamal's Converse for More Capable BCs

Lemma (El Gamal (1979))

Every $(n, M_{1n}, M_{2n}, \varepsilon_n)$-code for any BC satisfies

$$(\log M_{1n})(1-\varepsilon_n) - 1 \leq \sum_{i=1}^{n} I(U_{1i}; Y_{1i})$$

$$(\log M_{2n})(1-\varepsilon_n) - 1 \leq \sum_{i=1}^{n} I(U_{2i}; Y_{2i})$$

$$(\log M_{1n} + \log M_{2n})(1-\varepsilon_n) - 2 \leq \sum_{i=1}^{n} \left[I(X_i; Y_{1i}|U_{2i}) + I(U_{2i}; Y_{2i}) \right]$$

$$(\log M_{1n} + \log M_{2n})(1-\varepsilon_n) - 2 \leq \sum_{i=1}^{n} \left[I(U_{1i}; Y_{1i}) + I(X_i; Y_{2i}|U_{1i}) \right],$$

where U_{1i} and U_{2i} satisfy $(U_{1i}, U_{2i}) - X_i - (Y_{1i}, Y_{2i})$.

Converse Proof: Upper Bound on λ-Sum Rate

- Assume $L_1^* \geq L_2^*$ (wlog) and let

$$\lambda = \frac{L_1^*}{L_2^*} \geq 1.$$

Converse Proof: Upper Bound on λ-Sum Rate

- Assume $L_1^* \geq L_2^*$ (wlog) and let

$$\lambda = \frac{L_1^*}{L_2^*} \geq 1.$$

- Combining previous inequalities and using a standard time-sharing random variable, we obtain

$$\frac{1}{n}\big[\log M_{1n} + \lambda \log M_{2n} - (1+\lambda)\big] \leq \max_{U,X} I(X;Y_1|U) + \lambda I(U;Y_2)$$

Converse Proof: Upper Bound on λ-Sum Rate

- Assume $L_1^* \geq L_2^*$ (wlog) and let

$$\lambda = \frac{L_1^*}{L_2^*} \geq 1.$$

- Combining previous inequalities and using a standard time-sharing random variable, we obtain

$$\frac{1}{n}\big[\log M_{1n} + \lambda \log M_{2n} - (1+\lambda)\big] \leq \max_{U,X} I(X;Y_1|U) + \lambda I(U;Y_2)$$

- Problem: Maximization of $I(X;Y_1|U) + \lambda I(U;Y_2)$ over all (U,X) requires tools specific to the broadcast channel

Converse Proof: Upper Bound on λ-Sum Rate

- Assume $L_1^* \geq L_2^*$ (wlog) and let

$$\lambda = \frac{L_1^*}{L_2^*} \geq 1.$$

- Combining previous inequalities and using a standard time-sharing random variable, we obtain

$$\frac{1}{n}\big[\log M_{1n} + \lambda \log M_{2n} - (1+\lambda)\big] \leq \max_{U,X} I(X;Y_1|U) + \lambda I(U;Y_2)$$

- Problem: Maximization of $I(X;Y_1|U) + \lambda I(U;Y_2)$ over all (U,X) requires tools specific to the broadcast channel
- For the BS-BC, Mrs. Gerber's Lemma [Wyner-Ziv (1973)] helps to simplify

Converse Proof: Concave Envelopes

- Remove U's by exploiting tools from convex analysis

Converse Proof: Concave Envelopes

- Remove U's by exploiting tools from convex analysis
- Note that $U - X - Y_2$ forms a Markov chain so

$$\begin{aligned}
&\max_{P_{U,X}} I(X;Y_1|U) + \lambda I(U;Y_2) \\
&= \max_{P_{U,X}} I(X;Y_1|U) + \lambda[I(X;Y_2) - I(X;Y_2|U)] \\
&= \max_{P_X} \lambda I(X;Y_2) + \max_{P_{U|X}} [I(X;Y_1|U) - \lambda I(X;Y_2|U)]
\end{aligned}$$

Converse Proof: Concave Envelopes

- Remove U's by exploiting tools from convex analysis
- Note that $U - X - Y_2$ forms a Markov chain so

$$\begin{aligned}
&\max_{P_{U,X}} I(X;Y_1|U) + \lambda I(U;Y_2) \\
&= \max_{P_{U,X}} I(X;Y_1|U) + \lambda [I(X;Y_2) - I(X;Y_2|U)] \\
&= \max_{P_X} \lambda I(X;Y_2) + \max_{P_{U|X}} [I(X;Y_1|U) - \lambda I(X;Y_2|U)]
\end{aligned}$$

- Now,

$$\max_{P_{U|X}} [I(X;Y_1|U) - \lambda I(X;Y_2|U)] = \mathbb{C}[I(P_X,W) - \lambda I(P_X,V)]$$

where $W = P_{Y_1|X}$ and $V = P_{Y_2|X}$ and the concave envelope is defined as

$$\mathbb{C}[f](x) := \inf\{g(x) : g \geq f, g \text{ is concave}\}$$

Converse Proof: Concave Envelopes

- The usual superposition coding region is

$$\mathcal{C} = \bigcup_{P_X, P_{U|X}} \left\{ (R_1, R_2) \in \mathbb{R}_+^2 \mid R_1 \leq I(X;Y_1|U), R_2 \leq I(U;Y_2) \right\}$$

- Using the concave envelope representation, we have

$$\mathcal{C} = \bigcap_{\lambda \geq 1} \left\{ (R_1, R_2) \in \mathbb{R}_+^2 \mid R_1 + \lambda R_2 \leq \max_{P_X} \lambda I(X;Y_2) + T_\lambda(X) \right\}$$

where $T_\lambda(X) := \mathbb{C}[I(X;Y_1) - \lambda I(X;Y_2)]$.

Converse Proof: Concave Envelopes

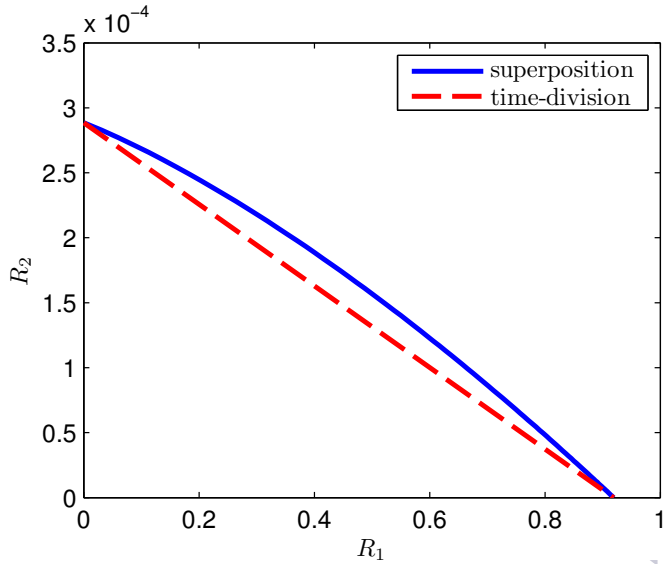
- The usual superposition coding region is

$$\mathcal{C} = \bigcup_{P_X, P_{U|X}} \left\{ (R_1, R_2) \in \mathbb{R}_+^2 \mid R_1 \leq I(X;Y_1|U), R_2 \leq I(U;Y_2) \right\}$$

- Using the concave envelope representation, we have

$$\mathcal{C} = \bigcap_{\lambda \geq 1} \left\{ (R_1, R_2) \in \mathbb{R}_+^2 \mid R_1 + \lambda R_2 \leq \max_{P_X} \lambda I(X;Y_2) + T_\lambda(X) \right\}$$

where $T_\lambda(X) := \mathbb{C}[I(X;Y_1) - \lambda I(X;Y_2)]$.



Converse Proof: Concave Envelopes

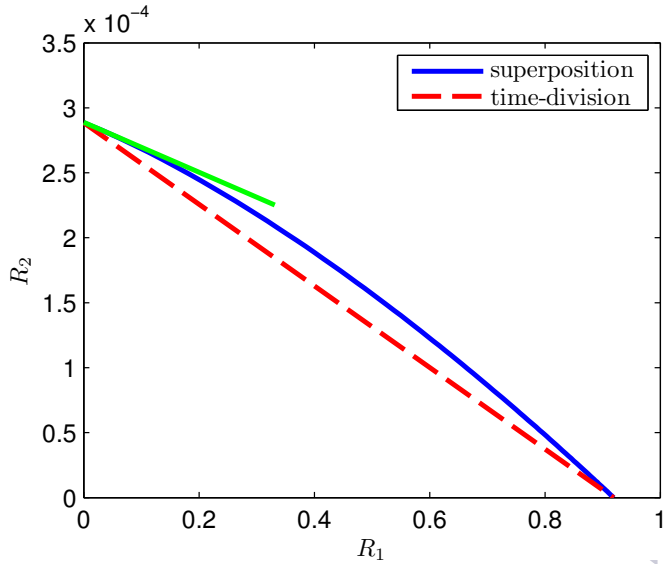
- The usual superposition coding region is

$$\mathcal{C} = \bigcup_{P_X, P_{U|X}} \left\{ (R_1, R_2) \in \mathbb{R}_+^2 \mid R_1 \leq I(X;Y_1|U), R_2 \leq I(U;Y_2) \right\}$$

- Using the concave envelope representation, we have

$$\mathcal{C} = \bigcap_{\lambda \geq 1} \left\{ (R_1, R_2) \in \mathbb{R}_+^2 \mid R_1 + \lambda R_2 \leq \max_{P_X} \lambda I(X;Y_2) + T_\lambda(X) \right\}$$

where $T_\lambda(X) := \mathbb{C}[I(X;Y_1) - \lambda I(X;Y_2)]$.



Converse Proof: Concave Envelopes

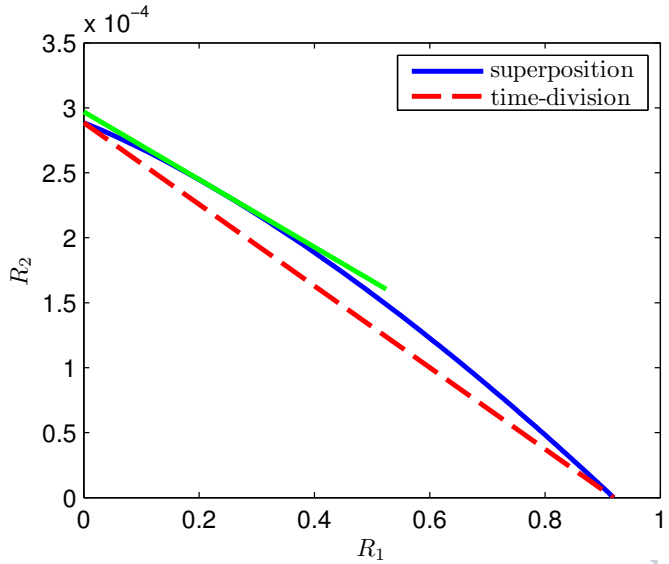
- The usual superposition coding region is

$$\mathcal{C} = \bigcup_{P_X, P_{U|X}} \left\{ (R_1, R_2) \in \mathbb{R}_+^2 \mid R_1 \leq I(X;Y_1|U), R_2 \leq I(U;Y_2) \right\}$$

- Using the concave envelope representation, we have

$$\mathcal{C} = \bigcap_{\lambda \geq 1} \left\{ (R_1, R_2) \in \mathbb{R}_+^2 \mid R_1 + \lambda R_2 \leq \max_{P_X} \lambda I(X;Y_2) + T_\lambda(X) \right\}$$

where $T_\lambda(X) := \mathbb{C}[I(X;Y_1) - \lambda I(X;Y_2)]$.



Converse Proof: Concave Envelopes

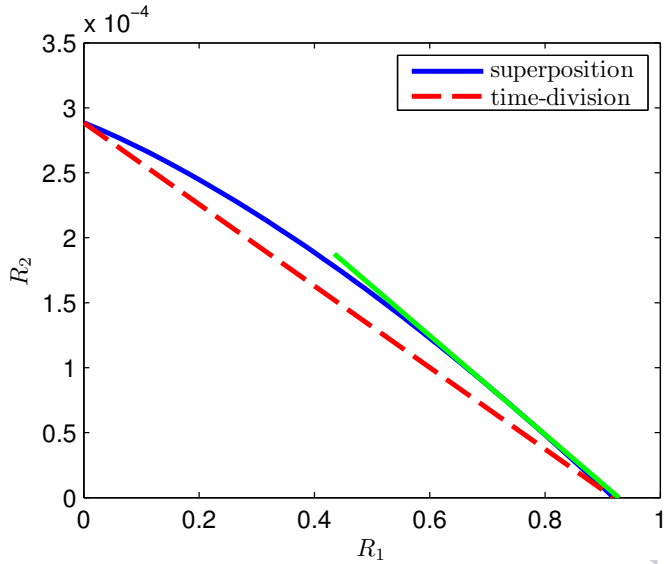
- The usual superposition coding region is

$$\mathcal{C} = \bigcup_{P_X, P_{U|X}} \left\{ (R_1, R_2) \in \mathbb{R}_+^2 \mid R_1 \leq I(X;Y_1|U), R_2 \leq I(U;Y_2) \right\}$$

- Using the concave envelope representation, we have

$$\mathcal{C} = \bigcap_{\lambda \geq 1} \left\{ (R_1, R_2) \in \mathbb{R}_+^2 \mid R_1 + \lambda R_2 \leq \max_{P_X} \lambda I(X;Y_2) + T_\lambda(X) \right\}$$

where $T_\lambda(X) := \mathbb{C}[I(X;Y_1) - \lambda I(X;Y_2)]$.



Converse Proof: Concave Envelopes

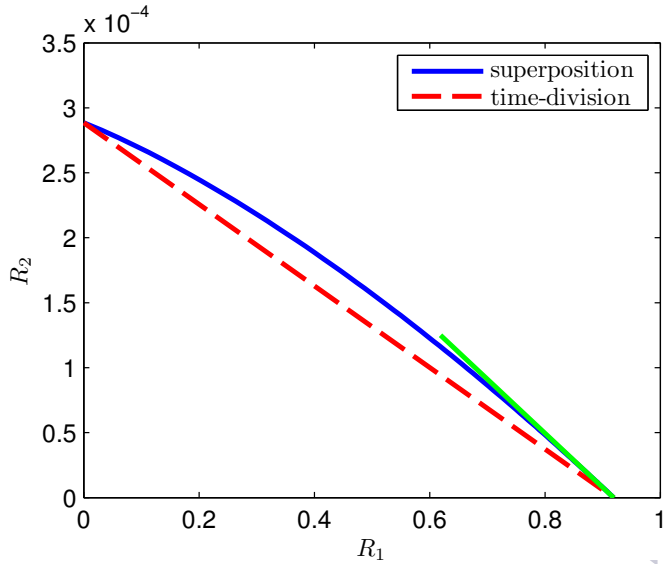
- The usual superposition coding region is

$$\mathcal{C} = \bigcup_{P_X, P_{U|X}} \left\{ (R_1, R_2) \in \mathbb{R}_+^2 \mid R_1 \leq I(X;Y_1|U), R_2 \leq I(U;Y_2) \right\}$$

- Using the concave envelope representation, we have

$$\mathcal{C} = \bigcap_{\lambda \geq 1} \left\{ (R_1, R_2) \in \mathbb{R}_+^2 \mid R_1 + \lambda R_2 \leq \max_{P_X} \lambda I(X;Y_2) + T_\lambda(X) \right\}$$

where $T_\lambda(X) := \mathbb{C}[I(X;Y_1) - \lambda I(X;Y_2)]$.



Converse Proof: Concave Envelopes

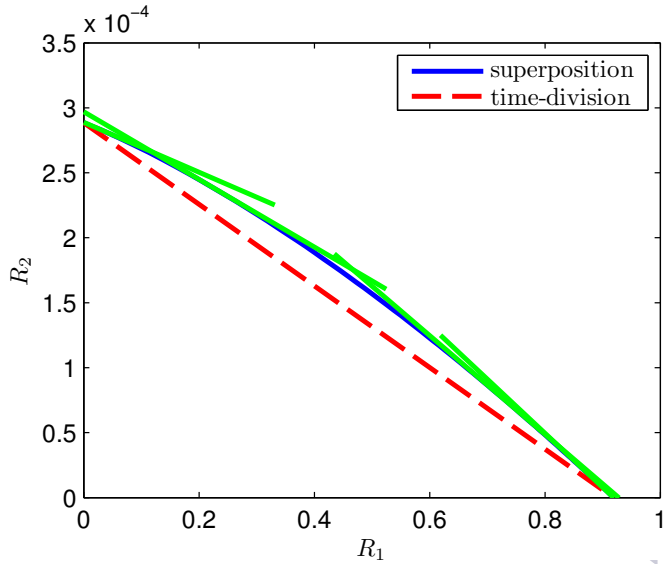
- The usual superposition coding region is

$$\mathcal{C} = \bigcup_{P_X, P_{U|X}} \left\{ (R_1, R_2) \in \mathbb{R}_+^2 \mid R_1 \leq I(X;Y_1|U), R_2 \leq I(U;Y_2) \right\}$$

- Using the concave envelope representation, we have

$$\mathcal{C} = \bigcap_{\lambda \geq 1} \left\{ (R_1, R_2) \in \mathbb{R}_+^2 \mid R_1 + \lambda R_2 \leq \max_{P_X} \lambda I(X;Y_2) + T_\lambda(X) \right\}$$

where $T_\lambda(X) := \mathbb{C}[I(X;Y_1) - \lambda I(X;Y_2)]$.



Converse Proof: A Tiny Bit of Analysis

- Converse bound becomes

$$\frac{\log M_{1n} + \lambda \log M_{2n}}{n} \lessapprox \max_{P_X} \lambda \cdot I(P_X, V) + \mathbb{C}\big[I(P_X, W) - \lambda \cdot I(P_X, V)\big]$$

Converse Proof: A Tiny Bit of Analysis

- Converse bound becomes

$$\frac{\log M_{1n} + \lambda \log M_{2n}}{n} \lessapprox \max_{P_X} \lambda \cdot I(P_X, V) + \mathbb{C}\big[I(P_X, W) - \lambda \cdot I(P_X, V)\big]$$

- Max over $P_X = [1-\alpha_n, \alpha_n]$ is over binary dist. with small mass at 1

Converse Proof: A Tiny Bit of Analysis

- Converse bound becomes

$$\frac{\log M_{1n} + \lambda \log M_{2n}}{n} \lessapprox \max_{P_X} \lambda \cdot I(P_X, V) + \mathbb{C}\big[I(P_X, W) - \lambda \cdot I(P_X, V)\big]$$

- Max over $P_X = [1 - \alpha_n, \alpha_n]$ is over binary dist. with small mass at 1
- Using Condition 1, we have $I(P_X, W) - \lambda \cdot I(P_X, V) \leq 0$ for all P_X

Converse Proof: A Tiny Bit of Analysis

- Converse bound becomes

$$\frac{\log M_{1n} + \lambda \log M_{2n}}{n} \lessapprox \max_{P_X} \lambda \cdot I(P_X, V) + \mathbb{C}\big[I(P_X, W) - \lambda \cdot I(P_X, V)\big]$$

- Max over $P_X = [1 - \alpha_n, \alpha_n]$ is over binary dist. with small mass at 1
- Using Condition 1, we have $I(P_X, W) - \lambda \cdot I(P_X, V) \leq 0$ for all P_X
- $\mathbb{C}\big[I(P_X, W) - \lambda \cdot I(P_X, V)\big] \leq 0$ for all P_X.

Converse Proof: A Tiny Bit of Analysis

- Converse bound becomes

$$\frac{\log M_{1n} + \lambda \log M_{2n}}{n} \lessapprox \max_{P_X} \lambda \cdot I(P_X, V) + \mathbb{C}\big[I(P_X, W) - \lambda \cdot I(P_X, V)\big]$$

- Max over $P_X = [1 - \alpha_n, \alpha_n]$ is over binary dist. with small mass at 1
- Using Condition 1, we have $I(P_X, W) - \lambda \cdot I(P_X, V) \leq 0$ for all P_X
- $\mathbb{C}\big[I(P_X, W) - \lambda \cdot I(P_X, V)\big] \leq 0$ for all P_X.
- Left with $I(P_X, V) \approx \alpha_n D(V(\cdot|1) \| V(\cdot|0))$, which is related to L_2^*.

Converse Proof: A Tiny Bit of Analysis

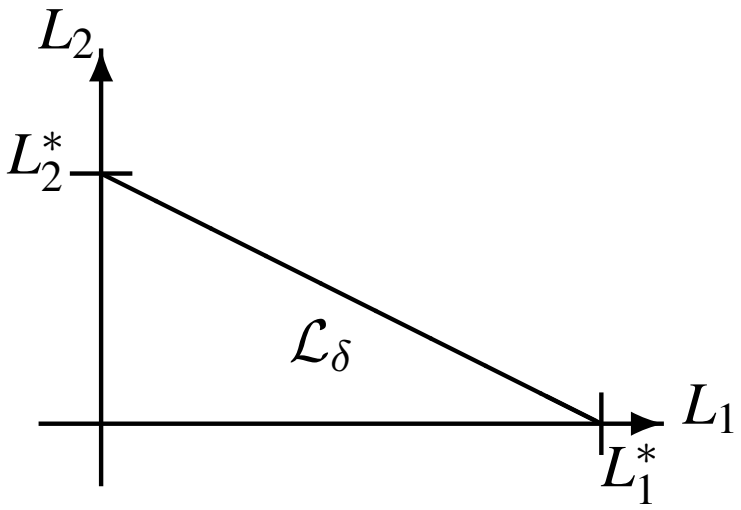
- Converse bound becomes

$$\frac{\log M_{1n} + \lambda \log M_{2n}}{n} \lessapprox \max_{P_X} \lambda \cdot I(P_X, V) + \mathbb{C}\big[I(P_X, W) - \lambda \cdot I(P_X, V)\big]$$

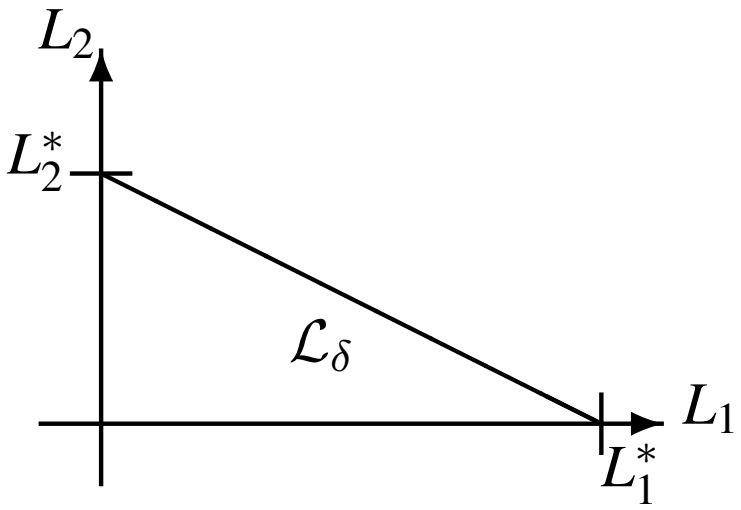
- Max over $P_X = [1-\alpha_n, \alpha_n]$ is over binary dist. with small mass at 1
- Using Condition 1, we have $I(P_X, W) - \lambda \cdot I(P_X, V) \leq 0$ for all P_X
- $\mathbb{C}\big[I(P_X, W) - \lambda \cdot I(P_X, V)\big] \leq 0$ for all P_X.
- Left with $I(P_X, V) \approx \alpha_n D(V(\cdot|1)\|V(\cdot|0))$, which is related to L_2^*.
- Finally, recalling that $\lambda = L_1^*/L_2^*$,

$$L_1 + \frac{L_1^*}{L_2^*} \cdot L_2 \lessapprox \frac{L_1^*}{L_2^*} \cdot L_2^* = L_1^*, \quad \Longrightarrow \quad \frac{L_1}{L_1^*} + \frac{L_2}{L_2^*} \leq 1.$$

Conclusion and Open Problems

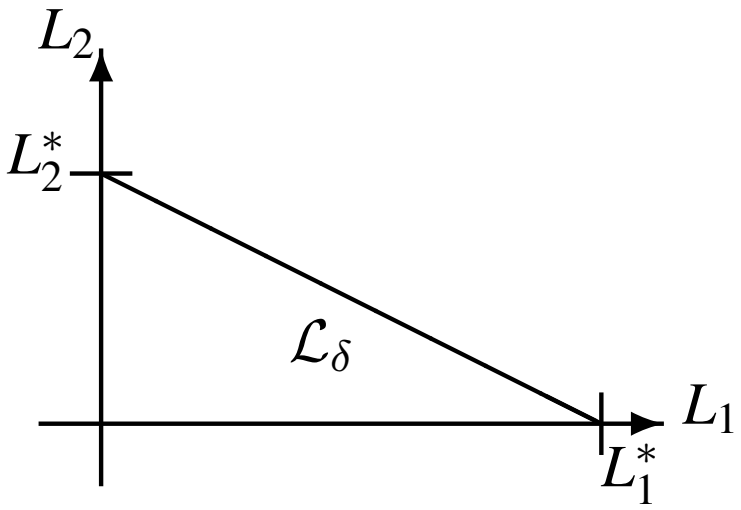


Conclusion and Open Problems



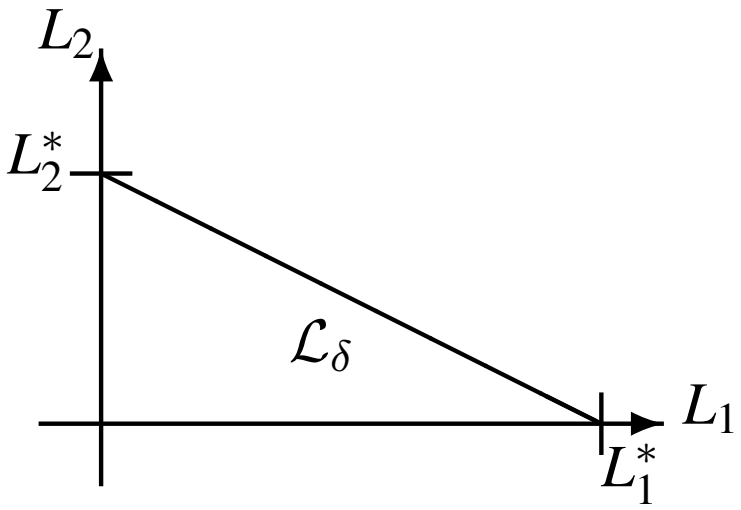
- Concave envelope representation of bounds on capacity region with auxiliary RVs is very useful

Conclusion and Open Problems



- Concave envelope representation of bounds on capacity region with auxiliary RVs is very useful
- What can we say about BCs which don't satisfy Condition 1?

Conclusion and Open Problems



- Concave envelope representation of bounds on capacity region with auxiliary RVs is very useful
- What can we say about BCs which don't satisfy Condition 1?
- More than 2 legitimate receivers?

Multiple Postdoc Positions in IT and ML at NUS

Postdoctoral positions available at the National University of Singapore

JUMP TO OTHER IT SOCIETY WEBSITES: Jump to child site

Recent News

Postdoctoral positions available at the National University of Singapore

Upcoming Events

Subscribe to announcements

Instructions and Guidelines for Submitting Content

Post an announcement

Post an event

The Department of Electrical and Computer Engineering (ECE) at the National University of Singapore is offering positions for postdoctoral fellows who will work in information theory, machine learning and their intersection.

The Department of Electrical and Computer Engineering (ECE) at the National University of Singapore (NUS) is offering positions for postdoctoral fellows who will work closely with Dr. Vincent Tan at the intersection of information theory, statistical signal processing, and machine learning. Some sample topics include:

- Fundamental performance limits (and algorithms) for dictionary learning (e.g., matrix factorization), ranking, and deep learning architectures;
- Learning in the presence of privacy constraints;
- Learning in the large alphabet regime;
- Learning of graphical models and other statistical models.

Working in traditional topics in Shannon's information theory of interest to the PI will also be highly encouraged. Some sample topics include:

- Multi-user information theory;
- Strong converse and second-order asymptotics;
- Error exponent analysis and the method of types;
- Information-theoretic security;