Decision making via End-to-End Lossy Distributed Wireless Cooperative Networks

- A Distributed Hypothesis Testing based Formulation -

Full Tutorial: MMMM DD, YYYY @ University of ZZZZ

Tad Matsumoto*, **, ***

IEEE Life Fellow

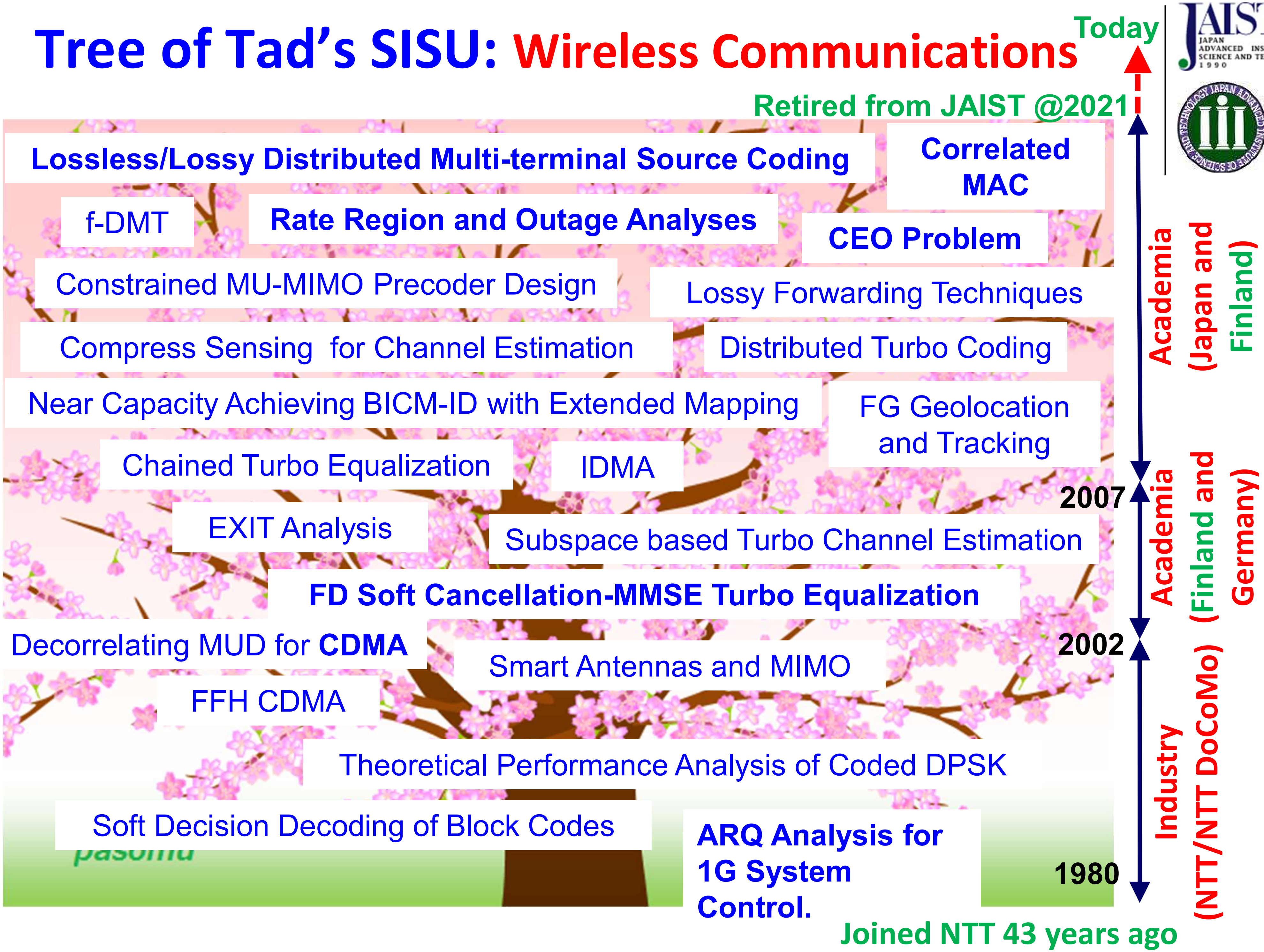
***International Chair, Invited Professor, IMT -Atlantic**
Technopôle Brest-Iroise CS 83818 29238 Brest cedex 03, France
Tadashi.matsumoto@imt-atlantique.fr

** **Professor Emeritus, JAIST**
matumoto@jaist.ac.jp

*** **Professor Emeritus, CWC, University of Oulu**,
tadashi.matsumoto@oulu.fi

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

Tree of Tad's SISU: Wireless Communications



Contents of Full Tutorial:

Copied from March 27 Seminar Top Page Slide

Lossy Multi-terminal Cooperative Networks, Queueing, and Decision Making:

Erlang, Shannon, and Neyman-Pearson Meet in 6G Networks

March 27, 2023 @ Uoulu CWC Research Seminar by Remote

Tad Matsumoto*, **, ***

IEEE Life Fellow

In what fields are they famous as land-make builders?

Home

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

CominLabs

JAIST
JAPAN ADVANCED INSTITUTE OF SCIENCE AND TECHNOLOGY 1990

CWC Oulu

SEVENTH FRAMEWORK PROGRAMME

***International Chair, Invited Professor, IMT -Atlantic**
Technopôle Brest-Iroise CS 83818 29238 Brest cedex 03, France
Tadashi.matsumoto@imt-atlantique.fr

** **Professor Emeritus, JAIST**
matumoto@jaist.ac.jp

*** **Professor Emeritus, CWC, University of Oulu,**
tadashi.matsumoto@oulu.fi

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

CWC
Oulu

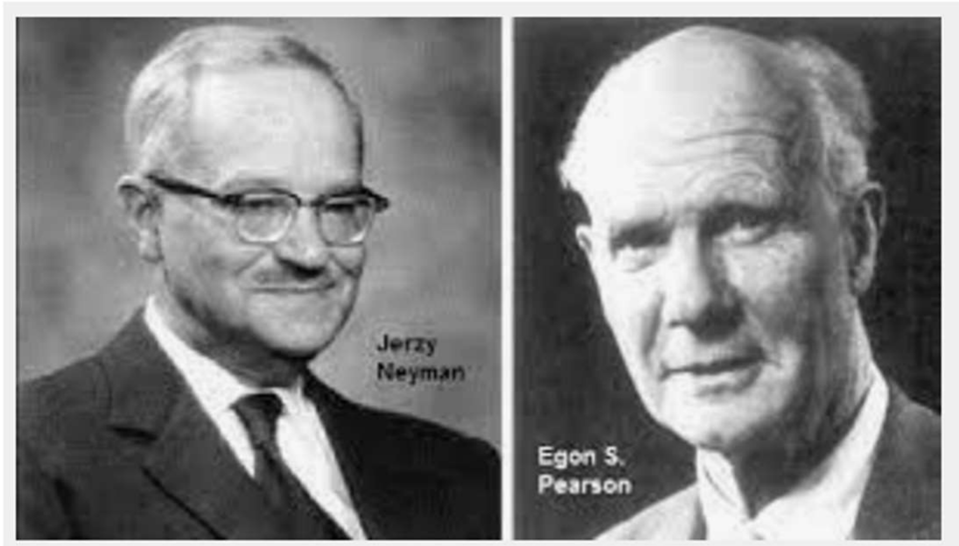
JAIST
JAPAN ADVANCED INSTITUTE OF SCIENCE AND TECHNOLOGY 1990

Erlang:

Queueing Theory

Shannon:

Information Theory



Neyman-Pearson:

Hypothesis Testing

Why “Neyman Pearson” Involved?

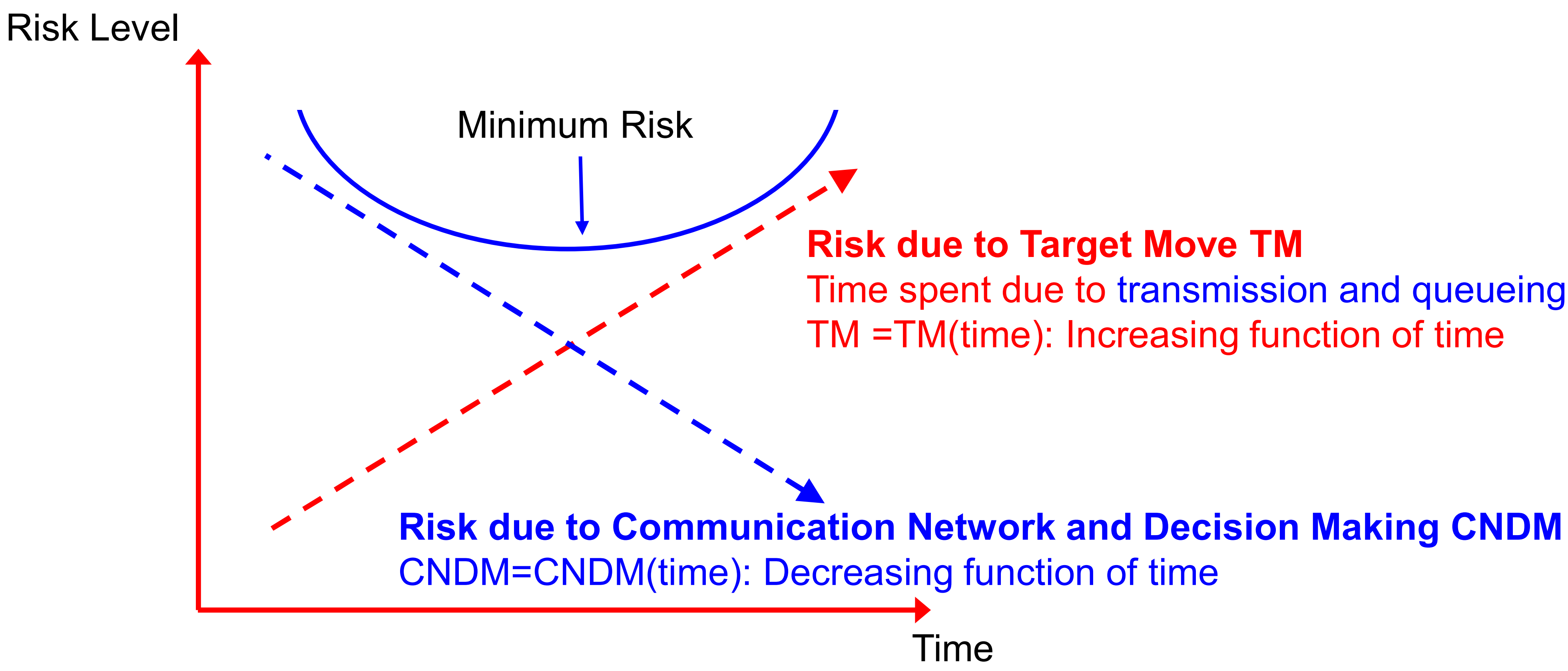
Unclear but still Accident Avoidable
Make left turn or right turn?

Clearer but maybe Too Late to Avoid Accident
Decision: Make left turn!

Making correct decision is most important than lossless reconstruction of the observation for risk avoidance!

Is Lossless reconstruction Needed?

Risk =Risk(TM(time), CNDM(time))



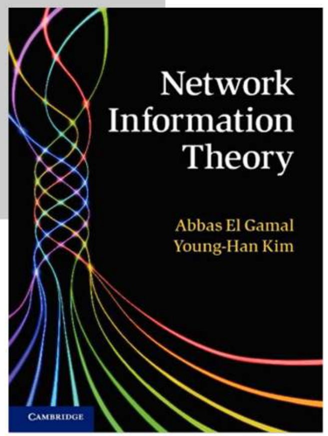
Motivation behind the March 23 Seminar:
Formulate Risk under the Three Landmark Builders' Framework!

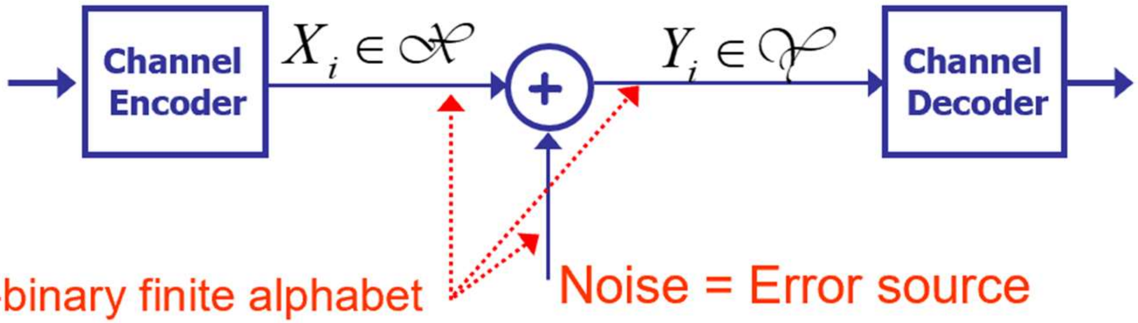
Chapter 0. Revisit of Fundamental Theorems in Network Information Theory

1. Channel Capacity: Point-to-Point Lossless Channel's Maximum Capability

Theorem 3.1 (Channel Coding Theorem). The capacity of the discrete memoryless channel $p(y|x)$ is given by the information capacity formula

$$C = \max_{p(x)} I(X;Y).$$

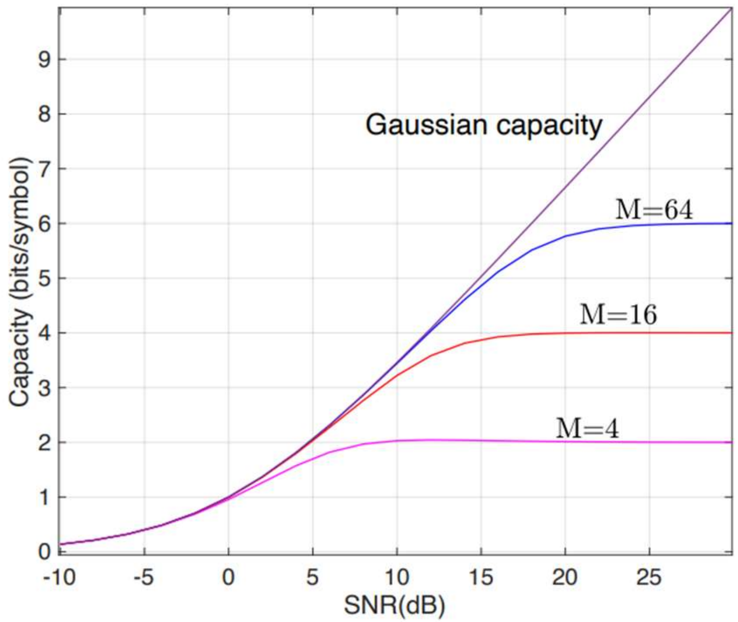


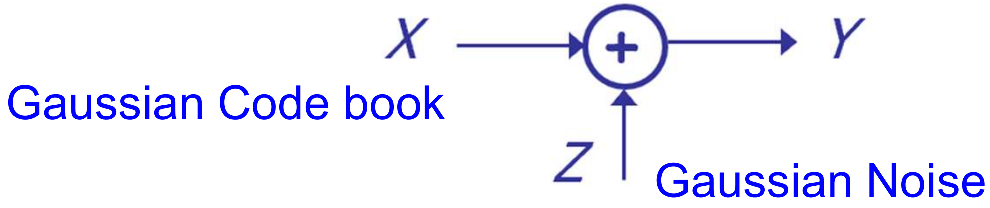


$$C = \max I(X;Y) \geq I(X;Y) \geq 0$$

$C \leq \log|\mathcal{X}|$ because $C = \max I(X;Y) \leq \max H(X) = \log|\mathcal{X}|$

$C \leq \log|\mathcal{Y}|$ because $C = \max I(X;Y) \leq \max H(Y) = \log|\mathcal{Y}|$



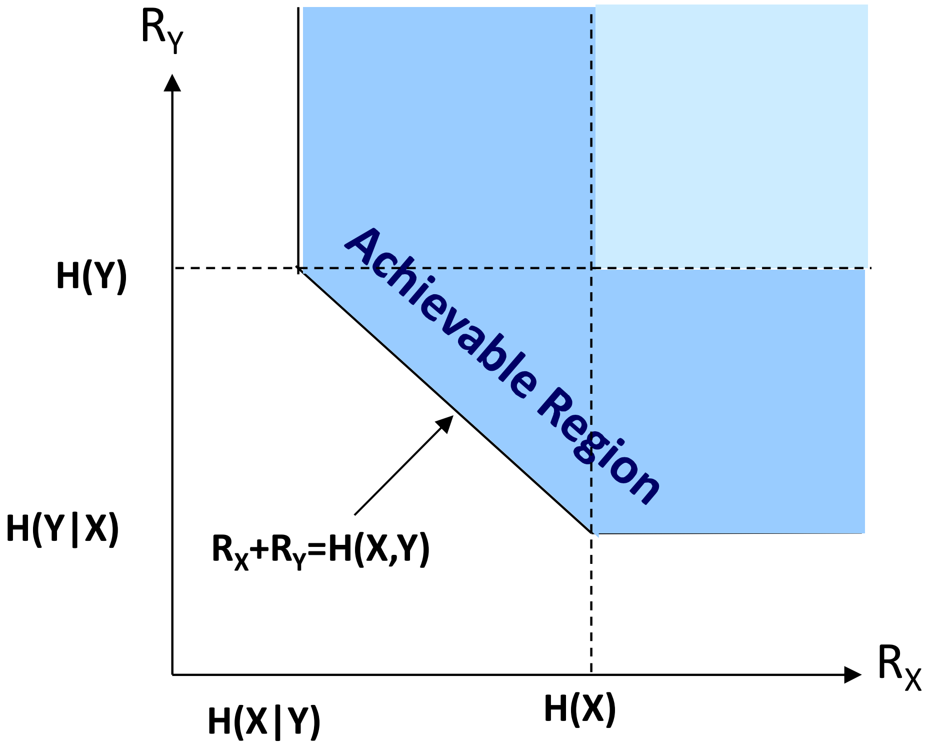


$$C = \max_{p(x):\ EX^2 \le P} I(X;Y) = \log(1 + SNR)$$

2. Rate Region: Distributed Multi-Source Lossless Coding

- Slepian-Wolf Theorem:

$$R_X \ge H(X|Y)$$
$$R_Y \ge H(Y|X)$$
$$R_X + R_Y \ge H(X,Y)$$



IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

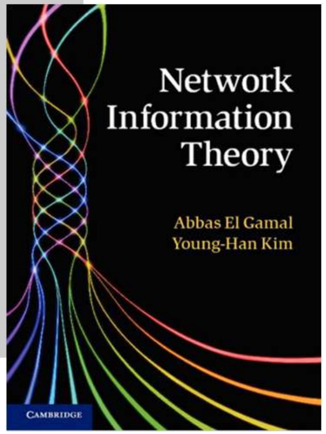
CWC Oulu

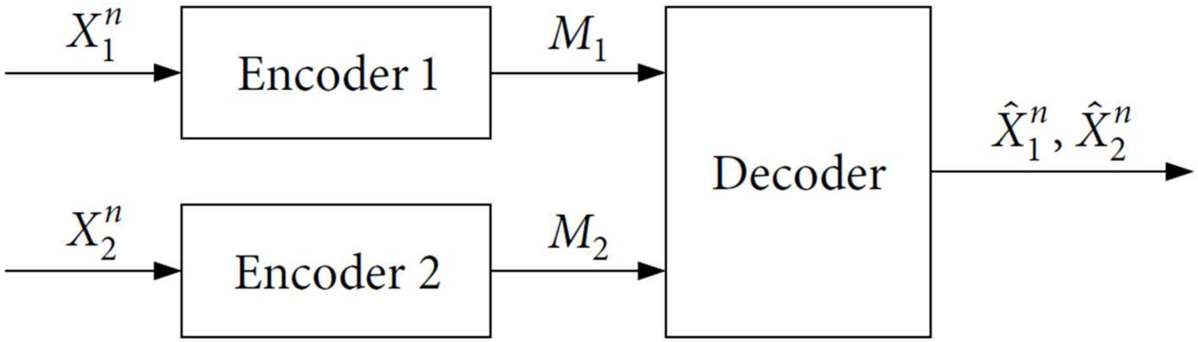
JAIST JAPAN ADVANCED INSTITUTE OF SCIENCE AND TECHNOLOGY 1990

Two Sources:

Theorem 10.1 (Slepian–Wolf Theorem). The optimal rate region $\mathscr{R}^*$ for distributed lossless source coding of a 2-DMS (X_1, X_2) is the set of rate pairs (R_1, R_2) such that

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



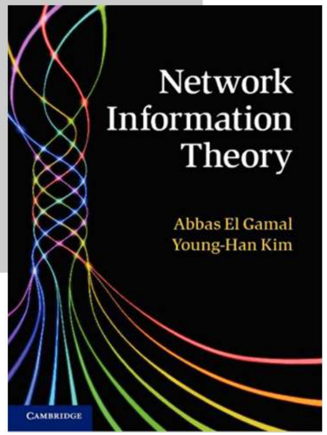


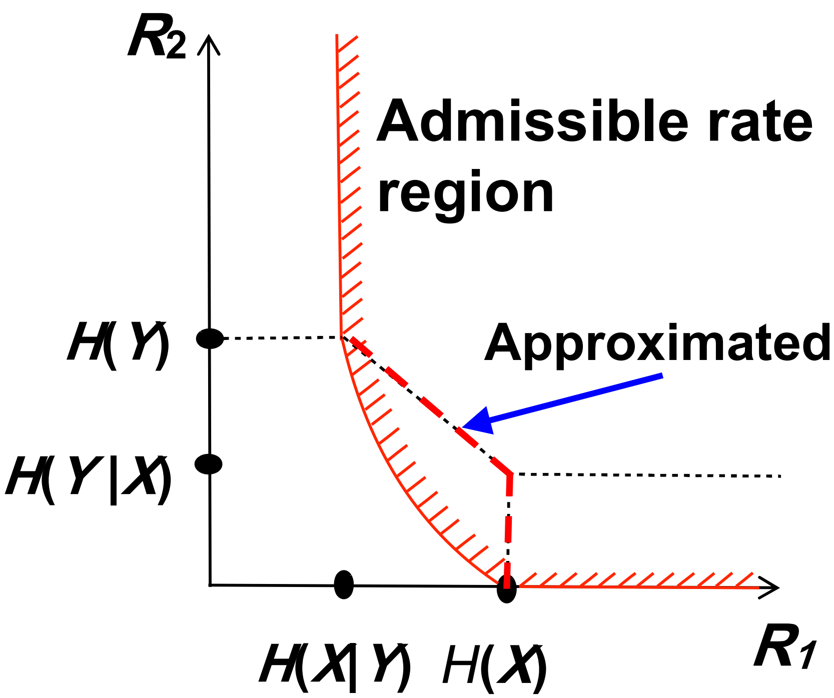
One Source One Helper:

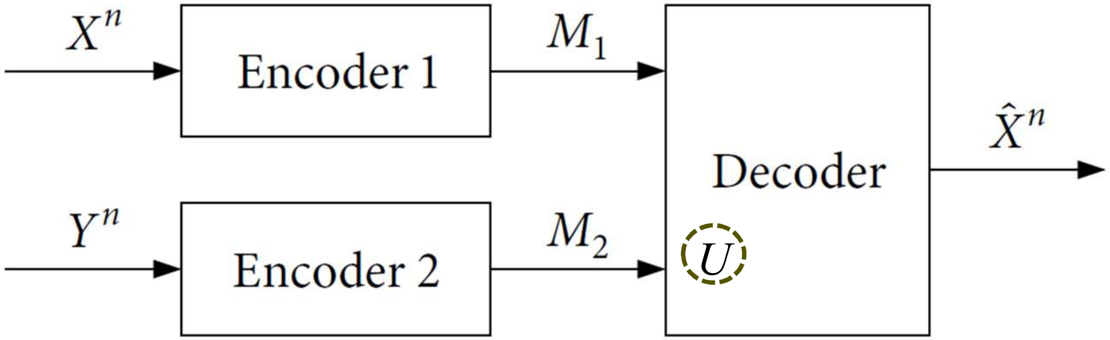
Theorem 10.2. Let (X, Y) be a 2-DMS. The optimal rate region $\mathscr{R}^*$ for lossless source coding of X with a helper observing Y is the set of rate pairs (R_1, R_2) such that

$$R_1 \geq H(X|U),$$
$$R_2 \geq I(Y;U)$$

for some conditional pmf $p(u|y)$, where $|\mathcal{U}| \leq |\mathcal{Y}| + 1$.







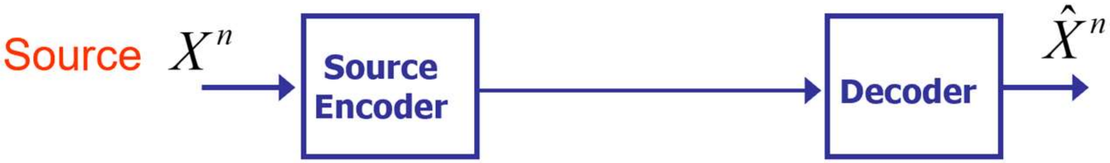
Remember! Coded Side Information=Helper

3. Lossy Source Coding:

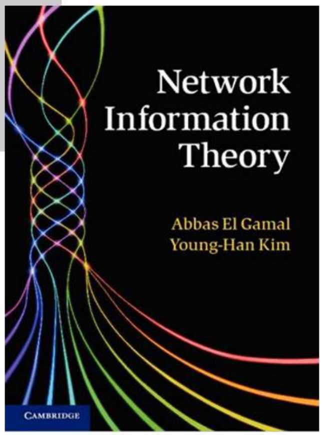
Theorem 3.5 (Lossy Source Coding Theorem). The rate–distortion function for a DMS X and a distortion measure $d(x, \hat{x})$ is

$$R(D) = \min_{p(\hat{x}|x):\mathsf{E}(d(X,\hat{X}))\le D} I(X;\hat{X})$$

for $D \ge D_{\min} = \min_{\hat{x}(x)} \mathsf{E}[d(X, \hat{x}(X))]$.



$$f_n(X^n) \in \{1, \quad 2, \quad \cdots \quad 2^{nR}\} \qquad \hat{X}^n = g_n(f_n(X^n)) \in \{1, \quad 2, \quad \cdots \quad 2^n\}$$



Example 3.4 (Bernoulli source with Hamming distortion). The rate–distortion function for a Bern(p) source X, $p \in [0, 1/2]$, and Hamming distortion measure is

$$R(D) = \begin{cases} H(p) - H(D) & \text{for } 0 \le D < p, \\ 0 & \text{for } D \ge p. \end{cases}$$

$$\begin{aligned} I(X;\hat{X}) &= H(X) - H(X|\hat{X}) \\ &= H(p) - H(X \oplus \hat{X}|\hat{X}) \\ &\ge H(p) - H(X \oplus \hat{X}) \\ &\overset{(a)}{\ge} H(p) - H(D), \end{aligned}$$

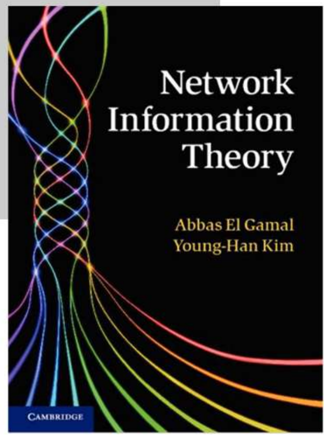
with $\mathsf{P}\{X \ne \hat{X}\} \le D$

4. Source-Channel Separation:

A Connecting point between Source Coding and Channel Coding

Theorem 3.7 (Source–Channel Separation Theorem). Given a DMS U and a distortion measure $d(u, \hat{u})$ with rate–distortion function $R(D)$ and a DMC $p(y|x)$ with capacity C, the following statements hold:

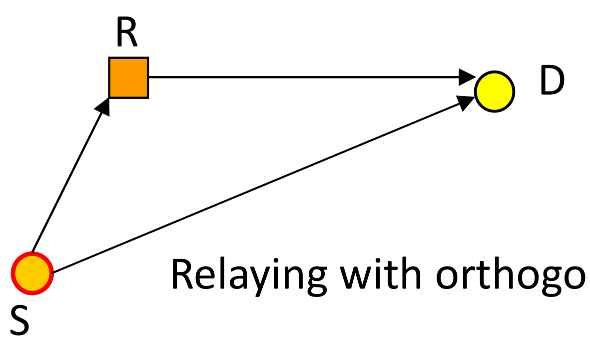
- If $rR(D) < C$, then (r, D) is achievable.
- If (r, D) is achievable, then $rR(D) \leq C$.



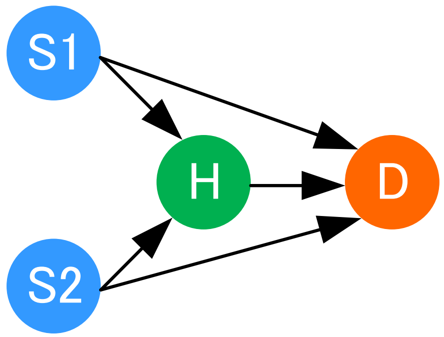
Remark 3.14. As a special case of joint source–channel coding, consider the problem of sending U over a DMC losslessly, i.e., $\lim_{k\to\infty} \mathsf{P}\{\hat{U}^k \neq U^k\} = 0$. The separation theorem holds with the requirement that $rH(U) \leq C$.

Note:
Source-Channel Separation applies to:

- **Orthogonal transmission with multiple Point-to-Point links,**
- **Both Lossless and Lossy, so far as each link is orthogonal,**
- **Helper link,**
- **(Experts say it holds with majority of the cases….)**



Relaying with orthogonal links



Multiple Access Relaying with orthogonal links

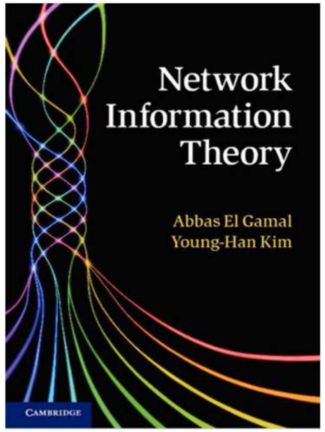
5. Rate Region:
Lossy Source Coding with Side information/a Helper

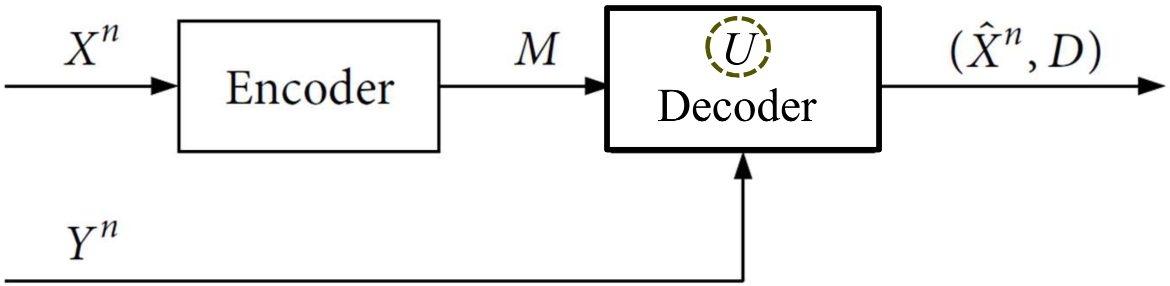
With Side information:

Theorem 11.3 (Wyner–Ziv Theorem). Let (X, Y) be a 2-DMS and $d(x, \hat{x})$ be a distortion measure. The rate–distortion function for X with side information Y available noncausally at the decoder is

$$R_{\text{SI-D}}(D) = \min\big(I(X;U) - I(Y;U)\big) = \min I(X;U|Y) \quad \text{for } D \geq D_{\min},$$

where the minimum is over all conditional pmfs $p(u|x)$ with $|\mathcal{U}| \leq |\mathcal{X}| + 1$ and functions $\hat{x}(u, y)$ such that $\mathsf{E}[d(X, \hat{X})] \leq D$, and $D_{\min} = \min_{\hat{x}(y)} \mathsf{E}[d(X, \hat{x}(Y))]$.





Notice: U→X→Y forms a Markov Chain

$$I(X;U) - I(Y;U) = I(XY;U) - I(Y;U|X) - I(Y;U)$$
$$= I(XY;U) - I(Y;U) = I(X;U|Y).$$

With a Helper:

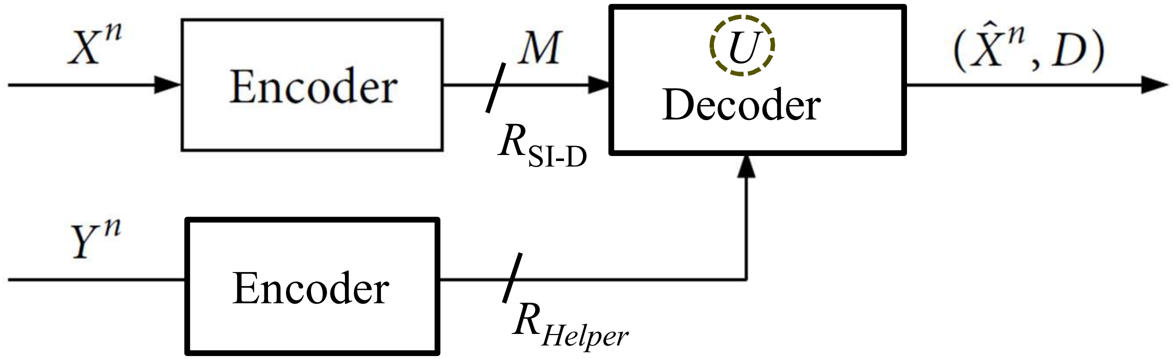
Theorem 11.3 (Wyner–Ziv Theorem). Let (X, Y) be a 2-DMS and $d(x, \hat{x})$ be a distortion measure. The rate–distortion function for X with side information Y available noncausally at the decoder is

$$R_{\text{SI-D}}(D) = \min \big(I(X;U) - I(Y;U) \big) = \min I(X;U|Y) \quad \text{for } D \geq D_{\min},$$

$$R_{Helper} \geq I(Y;U)$$

where the minimum is over all conditional pmfs $p(u|x)$ with $|\mathcal{U}| \leq |\mathcal{X}| + 1$ and functions $\hat{x}(u, y)$ such that $\mathsf{E}[d(X, \hat{X})] \leq D$, and $D_{\min} = \min_{\hat{x}(y)} \mathsf{E}[d(X, \hat{x}(Y))]$.

Tad's Book?



Coded Side Information=Helper

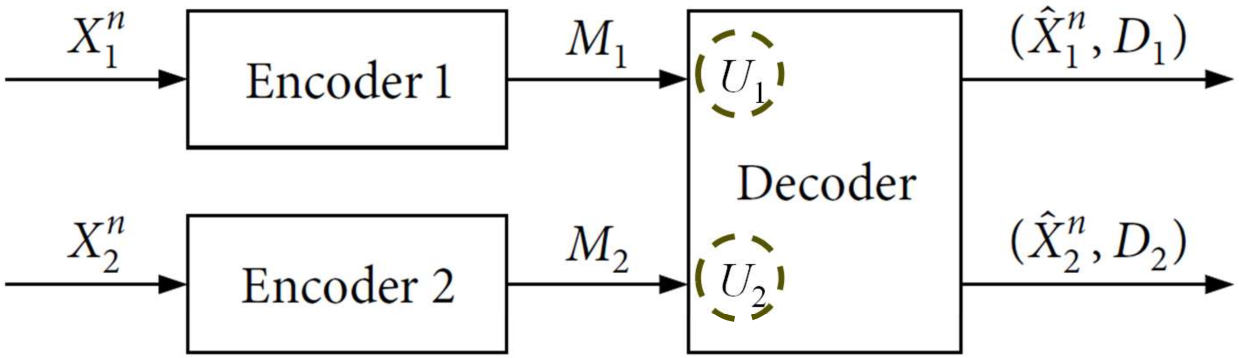
6. Rate Region:
Distributed Multipoint-to-Multipoint Lossy Coding

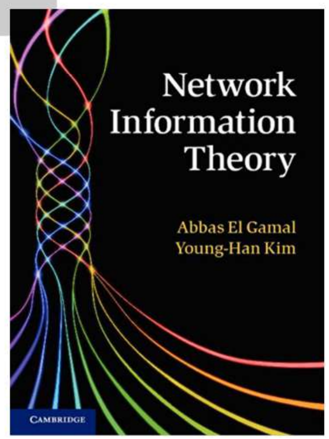
Without Helper:

Theorem 12.1 (Berger–Tung Inner Bound). Let (X_1, X_2) be a 2-DMS and $d_1(x_1, \hat{x}_1)$ and $d_2(x_2, \hat{x}_2)$ be two distortion measures. A rate pair (R_1, R_2) is achievable with distortion pair (D_1, D_2) for distributed lossy source coding if

$$\begin{aligned} R_1 &> I(X_1; U_1 | U_2, Q), \\ R_2 &> I(X_2; U_2 | U_1, Q), \\ R_1 + R_2 &> I(X_1, X_2; U_1, U_2 | Q) \end{aligned}$$

for some conditional pmf $p(q)p(u_1|x_1, q)p(u_2|x_2, q)$ with $|\mathcal{U}_j| \le |\mathcal{X}_j| + 4$, $j = 1, 2$, and functions $\hat{x}_1(u_1, u_2, q)$ and $\hat{x}_2(u_1, u_2, q)$ such that $\mathsf{E}(d_j(X_j, \hat{X}_j)) \le D_j$, $j = 1, 2$.





Notice: $U_1 \to X_1 \to X_2 \to U_2$

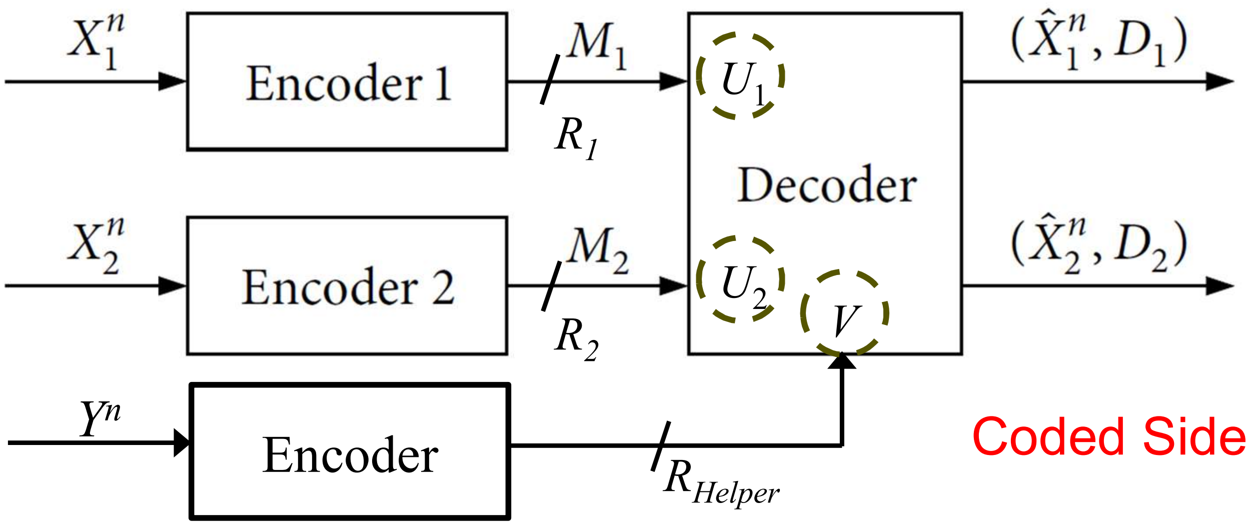
With a Helper:

Theorem 12.1 (Berger–Tung Inner Bound). Let (X_1, X_2) be a 2-DMS and $d_1(x_1, \hat{x}_1)$ and $d_2(x_2, \hat{x}_2)$ be two distortion measures. A rate pair (R_1, R_2) is achievable with distortion pair (D_1, D_2) for distributed lossy source coding if

$$R_1 > I(X_1; U_1 | U_2, V, Q),$$
$$R_2 > I(X_2; U_2 | U_1, V, Q),$$
$$R_1 + R_2 > I(X_1, X_2; U_1, U_2 | V, Q),$$
$$R_{Helper} > I(Y; V)$$

for some conditional pmf $p(q)p(u_1|x_1, q)p(u_2|x_2, q)$ with $|\mathcal{U}_j| \le |\mathcal{X}_j| + 4$, $j = 1, 2$, and functions $\hat{x}_1(u_1, u_2, q)$ and $\hat{x}_2(u_1, u_2, q)$ such that $\mathsf{E}(d_j(X_j, \hat{X}_j)) \le D_j$, $j = 1, 2$.

Tad's Book?



Coded Side Information=Helper

7. Multiple Access Channels (MAC)

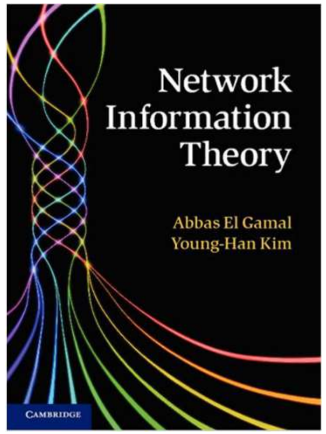
Correlated Sources Transmission over MAC:

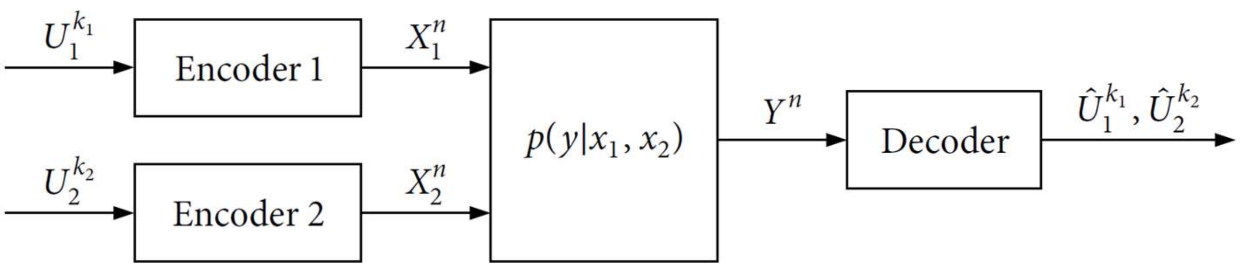
First consider the following sufficient condition for separate source and channel coding. We know that the capacity region $\mathscr{C}$ of the DM-MAC is the set of rate pairs (R_1, R_2) such that

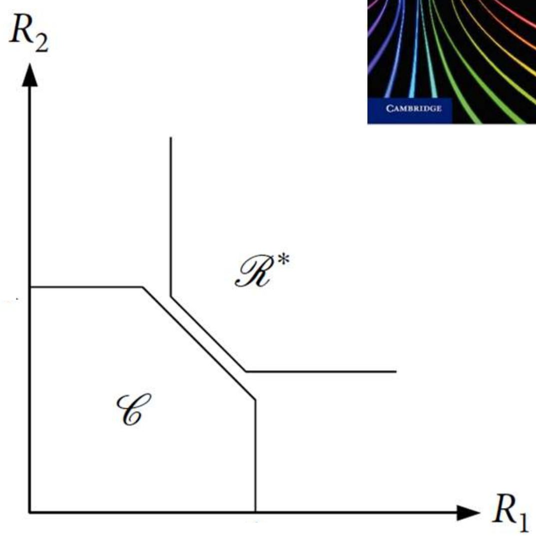
$$\begin{aligned} R_1 &\le I(X_1; Y | X_2, Q), \\ R_2 &\le I(X_2; Y | X_1, Q), \\ R_1 + R_2 &\le I(X_1, X_2; Y | Q) \end{aligned}$$

for some pmf $p(q)p(x_1|q)p(x_2|q)$. We also know from the Slepian–Wolf theorem that the optimal rate region $\mathscr{R}^*$ for distributed lossless source coding is the set of rate pairs (R_1, R_2) such that

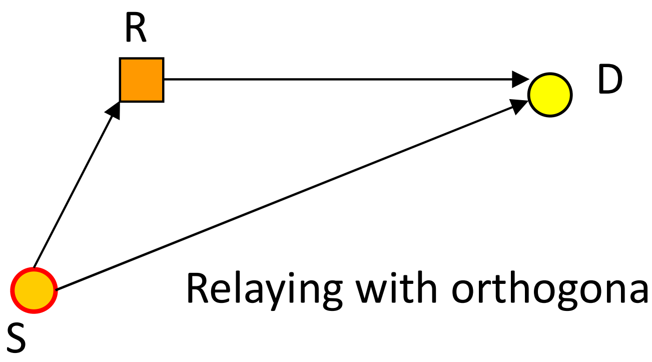
$$\begin{aligned} R_1 &\ge H(U_1 | U_2), \\ R_2 &\ge H(U_2 | U_1), \\ R_1 + R_2 &\ge H(U_1, U_2). \end{aligned}$$

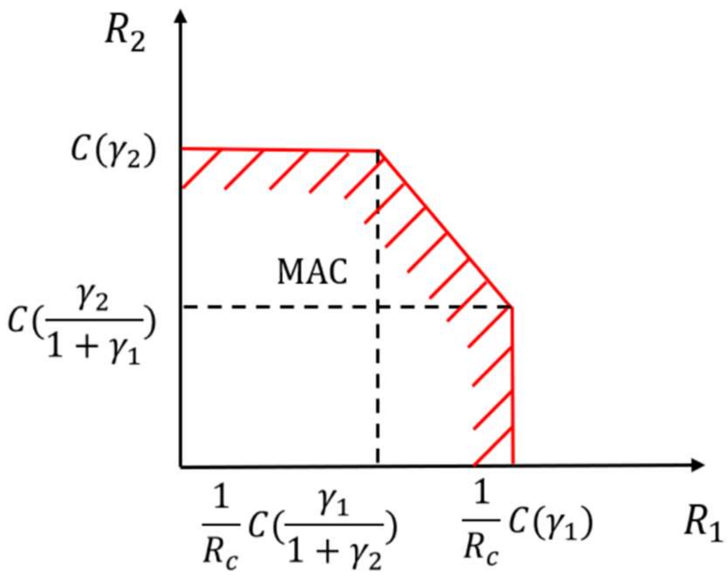


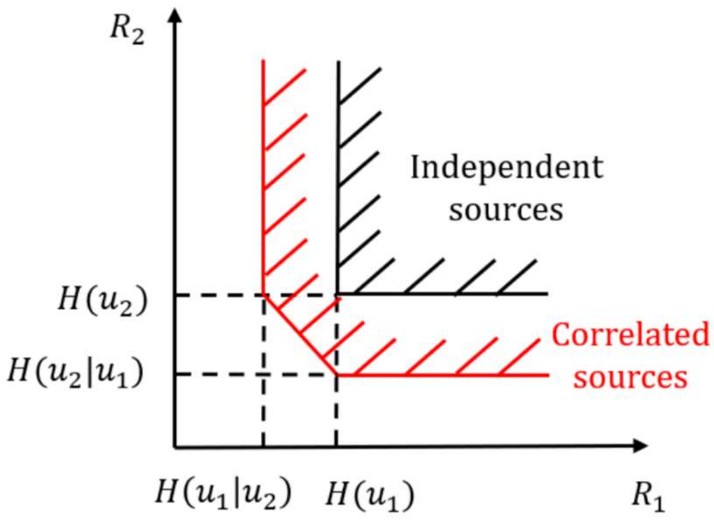




There are two regions in this set up: SW and MAC regions.

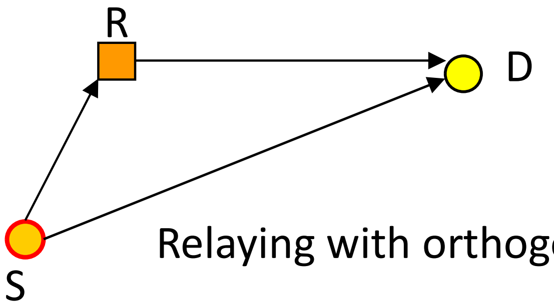




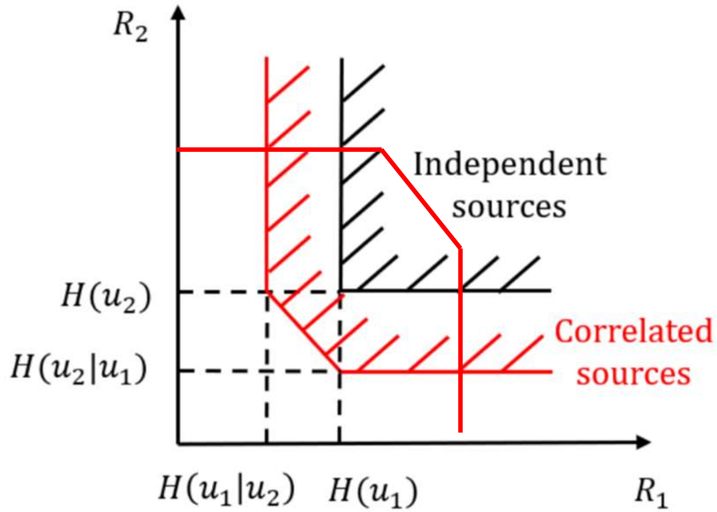


Revisit to Source-Channel Separation:

- Orthogonal transmission with multiple Point-to-Point links,
- Both Lossless and Lossy, so far as orthogonal,
- Helper link,



Region intersection: The rate-pair plot belongs to the both MAC and SW regions.
→ Source-channel separation holds!



Separation holds in:

- MAC transmission when the rates-plot is in intersection (Sufficient condition, NOT optimal. Separation vs. Joint),
- (Experts say it holds in many cases….)

A sufficient condition for Lossless Recovery.

- **Recovery for two sources**

$$H(U_1|U_2) < I(X_1;Y|X_2,Q),$$
$$H(U_2|U_1) < I(X_2;Y|X_1,Q),$$
$$H(U_1,U_2) < I(X_1,X_2;Y|Q) .$$

- **Recovery for one source with one helper**

$$R_1 \le I(X_1;Y|X_2,Q),$$
$$R_2 \le I(X_2;Y|X_1,Q),$$
$$R_1 + R_2 \le I(X_1,X_2;Y|Q)$$

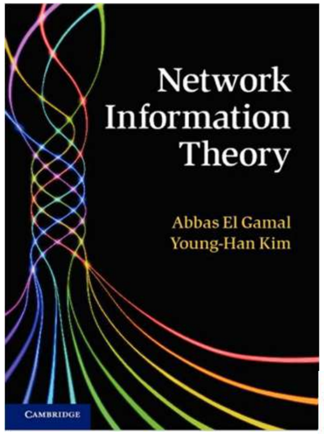
and

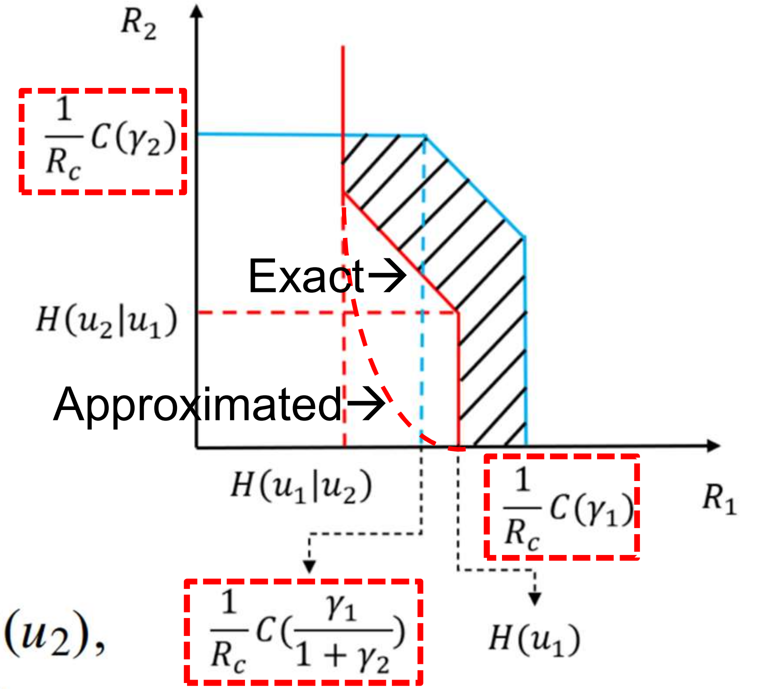
$$R_1 \ge H(X_1|U)$$
$$R_2 \ge I(X_2;U)$$

which can be approximated by

$$R_1 \ge \begin{cases} H(u_1|u_2), & \text{for } R_2 \ge H(u_2), \\ H(u_1,u_2) - R_2, & \text{for } H(u_2|u_1) \le R_2 \le H(u_2), \\ H(u_1), & \text{for } 0 \le R_2 \le H(u_2|u_1), \end{cases}$$

Tad's Book?

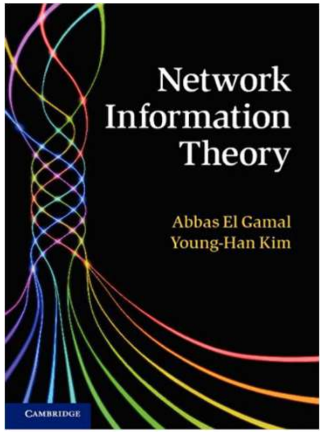




A sufficient condition for <u>Lossy</u> Recovery.

- Recovery for one source with one helper

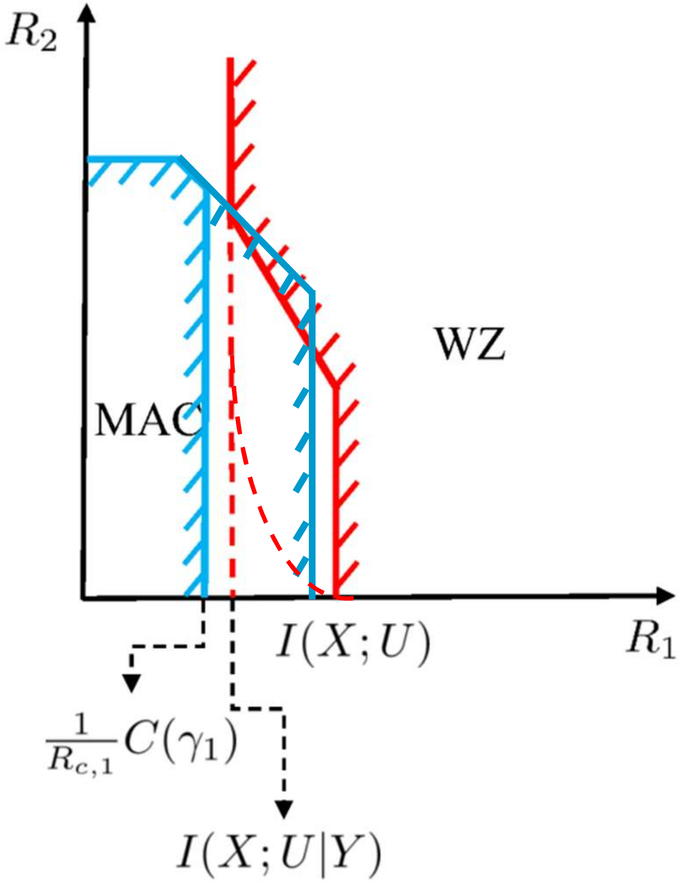
Tad's Book?



$$R_1 \le I(X_1; Y \mid X_2, Q),$$
$$R_2 \le I(X_2; Y \mid X_1, Q),$$
$$R_1 + R_2 \le I(X_1, X_2; Y \mid Q)$$

and

$$R_1 \ge H(X_1; U \mid V)$$
$$R_2 \ge I(X_2; V)$$



Region intersection
→ Source-channel separation holds!

Properties of Binary Convolution:

Binary Convolution

$$x * y = x(1-y) + (1-x)y = x + y - 2xy$$

- For variables $x, t \in [0,0.5]$ and $y \in [0,0.5)$

$$x * y \leq t \;\Rightarrow x \leq \Lambda(t, y) \triangleq \frac{1}{2}(1 - \frac{2t-1}{2y-1})$$

  - $x * y \leq t$ inherently involves $y \leq t$
  - $\Lambda(y, t)$ is a monotonically decreasing function of y, with a maximum

$$\Lambda(y=0, t) = t$$

  - $\Lambda(y, t)$ is a linearly increasing function of t, with a maximum

$$\Lambda(y, t=0.5) = 0.5$$

- **Recursive structure**:

$$y * z \leq s \Rightarrow y \leq \Lambda(s, z)$$

$$x * y * z \leq t \Rightarrow x \leq \Lambda(t, \Lambda(s, z))$$

$$v * w \leq c_1 \Rightarrow v \leq \Lambda(c_1, w)$$

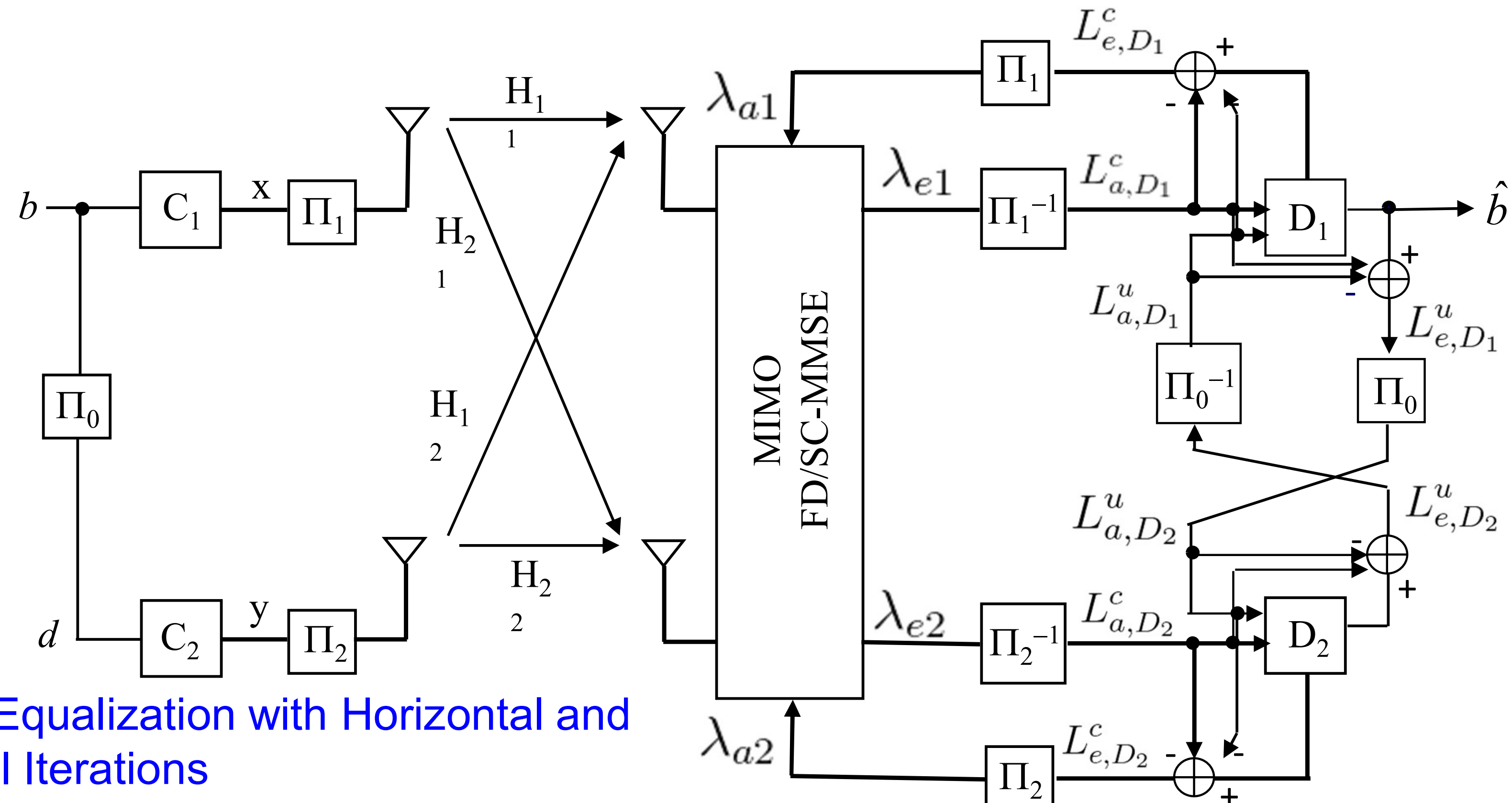
$$\vdots$$

$$x * y * z * \cdots v * w \leq c_n \Rightarrow x \leq \Lambda(c_n, \Lambda(cn_{-1}, \Lambda(\ldots, \Lambda(c_1, w)))$$

Tad's Book?

Chapter 1. End-to-End Lossless Relaying: Slepian Wolf Theorem with Source-Channel Separation

1.1 EXIT Analysis for Source Bit-Flipped MIMO Transmission with Turbo Equalization



Turbo Equalization with Horizontal and Vertical Iterations

- Vertical iteration is expected to improve performance because of space diversity gain from using antenna 1 and 2, and coding gain. Mariella Sarestoniemi, Tad Matsumoto, Kimmo Kansanen, and Jari Iinatti, "Turbo Diversity Based on SC/MMSE Equalization", *IEEE Transactions on Vehicular Technology*, Vol. 54, No. 2, pp. 749-752, March, 2005
- This design is called as **Spatial Turbo Code (STC)** because coded sequences are **multiplexed in the spatial domain**, **not in the time domain** as in the Turbo codes.

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

CominLabs

JAIST
JAPAN ADVANCED INSTITUTE OF SCIENCE AND TECHNOLOGY 1990

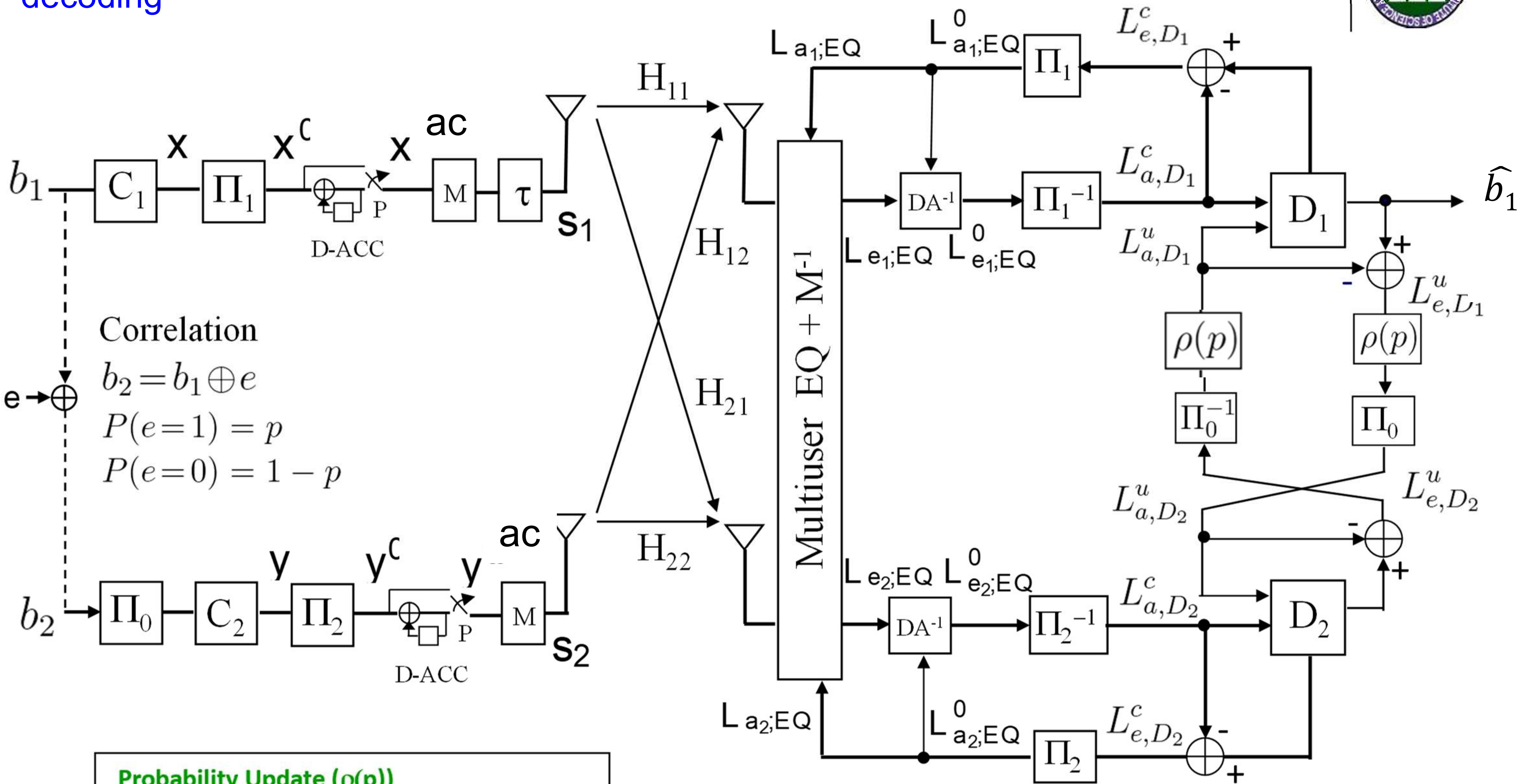
- We* developed Frequency Domain Turbo Equalization Algorithms for single carrier signalling: It requires computational complexity of only ”high school levelmath”!
- Convergence property analysis made significantly easy!

”Do you still spread?” ← Famous words said many times, said by a CWC person.

- Kimmo Kansanen, and Tad Matsumoto, ”An Analytical Method for MMSE MIMO Turbo Equalizer EXIT Chart Computation”, *IEEE Transactions on Wireless Communications*, Vol. 6, No.1, pp.59-63, January, 2007

Source Bit-Flipped MIMO Transmission with Turbo Equalization:

With the bit flipping e between b_1 and b_2, b_1 can be recovered losslssly by joint decoding



Probability Update ($\rho(p)$)

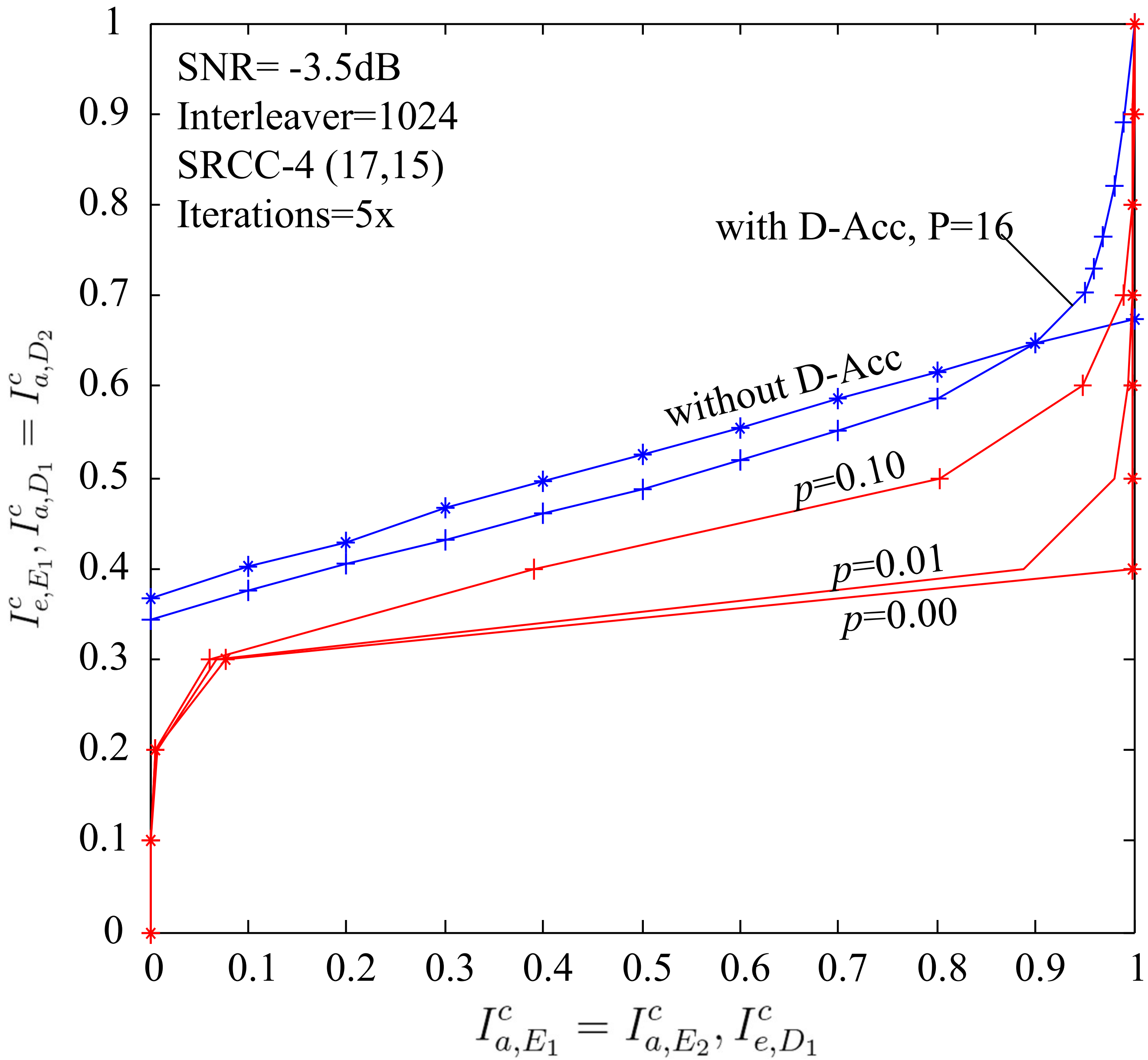
$$P(x=0) = (1-p)P_o(x=0) + pP_o(x=1)$$

$$P(x=1) = (1-p)P_o(x=1) + pP_o(x=0)$$

$$\rho(p) = LLR_{update} = ln\frac{(1-p)e^L + p}{(1-p) + pe^L}$$

K. Anwar and T. Matsumoto, "Spatially Concatenated Codes with Turbo Equalization for Correlated Sources", *IEEE Transaction Signal Processing*, vol. 60, no. 10, Oct. 2012, pp. 5572-5577

EXIT Chart for Source Bit-Flipped MIMO Transmission with Turbo Equalization



Parameters:

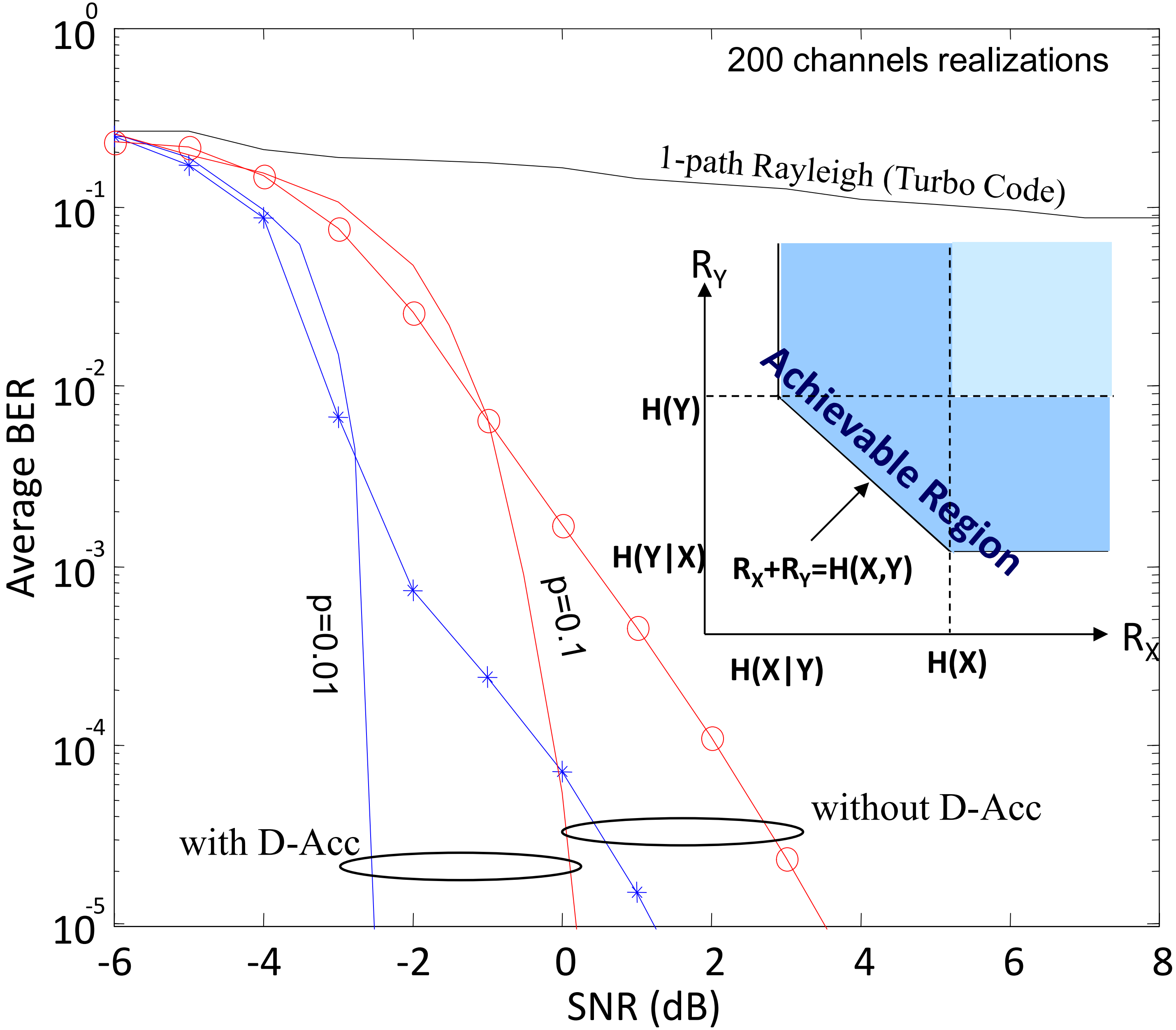
■ **Transmitter:**
Encoder: CC 4(17,15),17
Interleaver=5000 (random)
Correlation Model: Bit-flipping

■ **Channel:**
MIMO 2x2
Equal Power 64-path

■ **Receiver:**
Decoder: BCJR Log-MAP
FFT=512

Average BER in Block Frequency-Selective Block Rayleigh Fading Channels: Source Bit-Flipped MIMO Transmission with Turbo Equalization

- Bit-flipped sequences are correlated sources!



Parameters:

Transmitter:

Encoder: CC 4(17,15),17
Interleaver=1024 (random)
Correlation Model: Bit-flipping

Channel:

MIMO 2x2
Equal Power 64-path

Receiver:

Decoder: BCJR Log-MAP
FFT=512

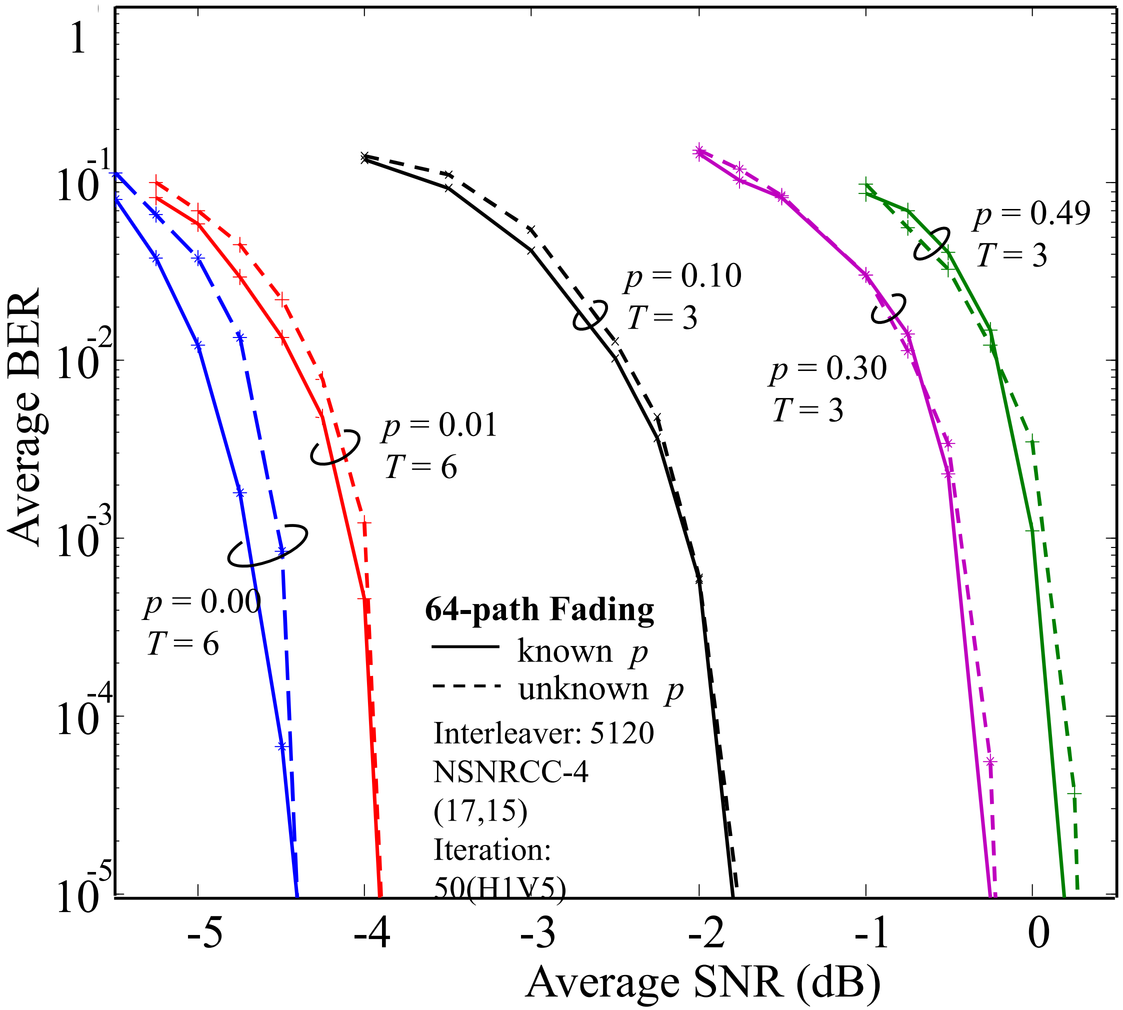
IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

Comparison in SNR to Slepian-Wolf Bound @ BER=10^{-5}

CWC Oulu
JAIST JAPAN ADVANCED INSTITUTE OF SCIENCE AND TECHNOLOGY 1990

p	$\mathcal{H}(b_1, b_2)$	$\mathcal{R}_{SW}$	Bit-Flipped MIMO TEQ		
			SNR_{lim}	$SNR_{BER=10^{-5}}$	Gap
0.00	1.000	0.250	-5.40	-4.38	1.02
0.01	1.081	0.270	-5.05	-3.88	1.17
0.10	1.469	0.367	-3.15	-1.88	1.27
0.30	1.881	0.470	-1.50	-0.25	1.25
0.49	1.999	0.499	-1.15	0.13	1.28

One Source One Helper
Slepian Wolf Theorem



Transmitter:
Encoder: CC 4(17,15)
Interleaver=10000 (random)
Correlation Model: Bit-flipping

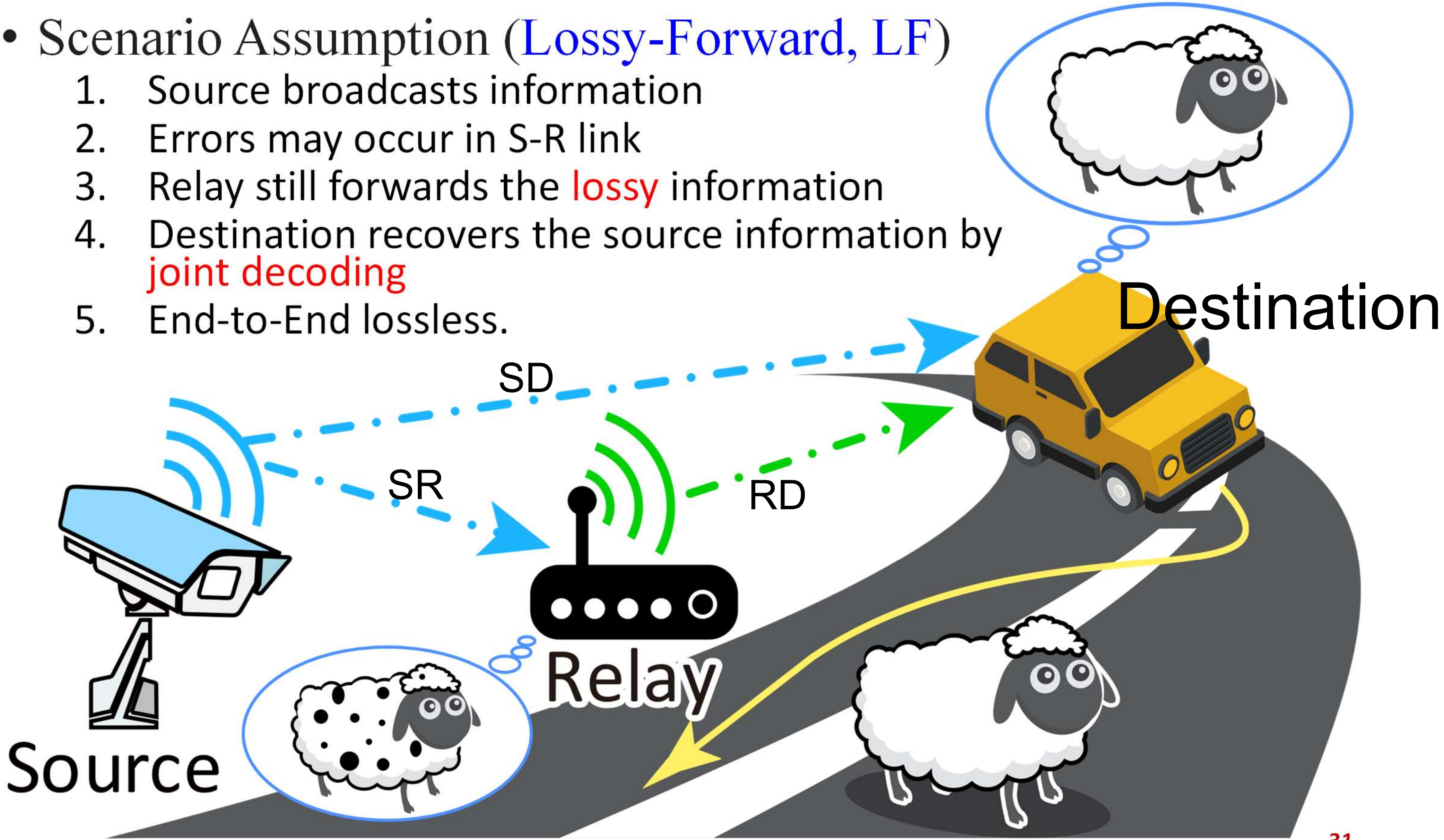
Channel:
MIMO 2x2
Equal Power 64-path

Receiver:
Decoder: BCJR
Log-MAP
FFT=512

1.2 Slepian-Wolf Formulation for Lossless Two-Way Relay Networks

Observation on Bit-flipped MIMO TEQ: BF Model Works as Correlated Sources!

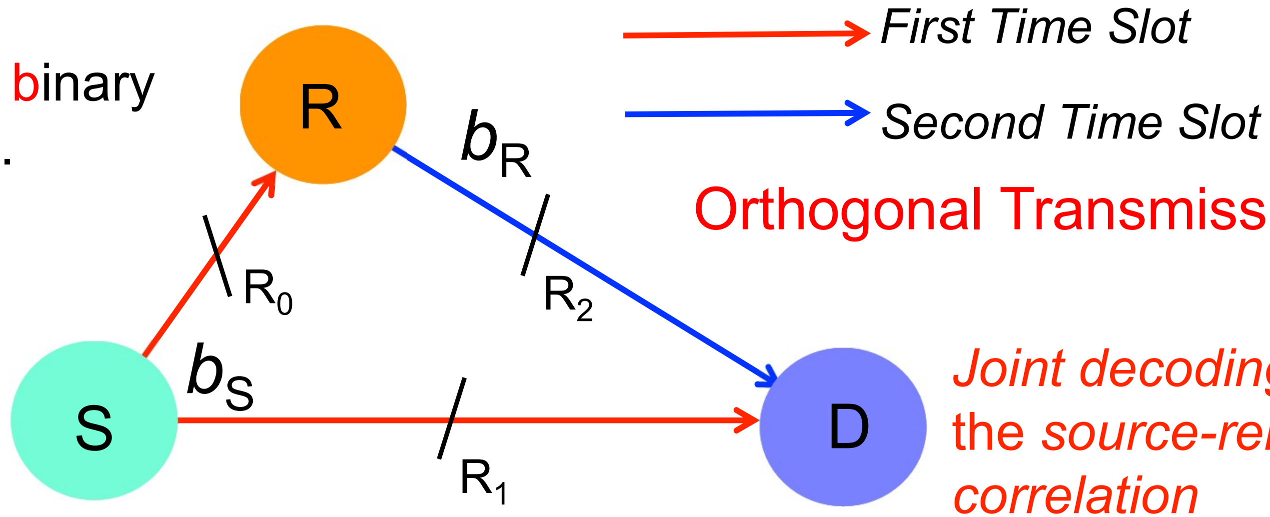
- Scenario Assumption (Lossy-Forward, LF)
  1. Source broadcasts information
  2. Errors may occur in S-R link
  3. Relay still forwards the lossy information
  4. Destination recovers the source information by joint decoding
  5. End-to-End lossless.



Do we need to recover the relay information b_R?

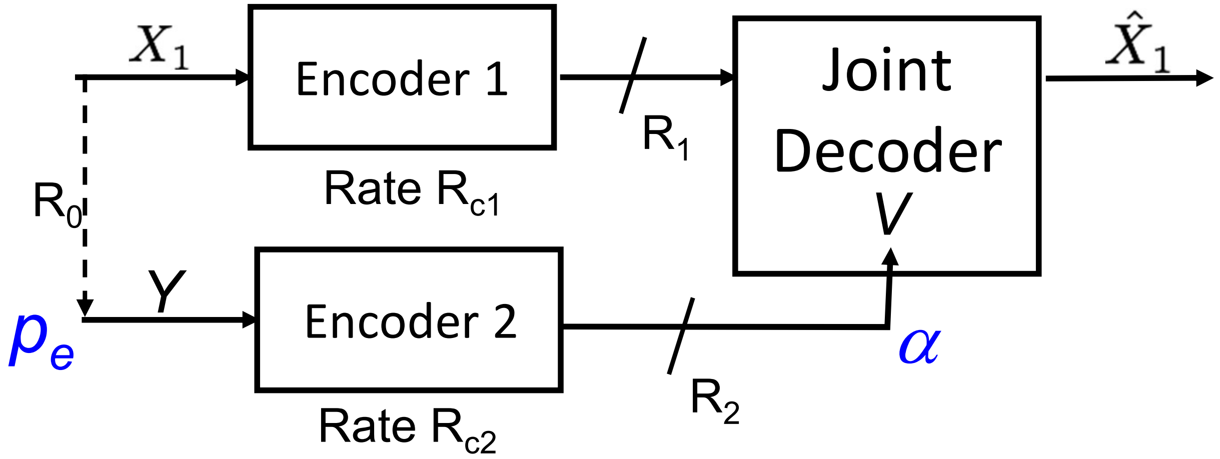
Source-Relay Correlation (bit-flipping model equivalent to BSC, binary symmetric channel.

As in Bit-flipped MIMO TEQ)



Orthogonal Transmission

Joint decoding exploiting the *source-relay correlation*



*We do not care about the decoding result (=V) of b_R, but we can use b_R as a **helper**!*

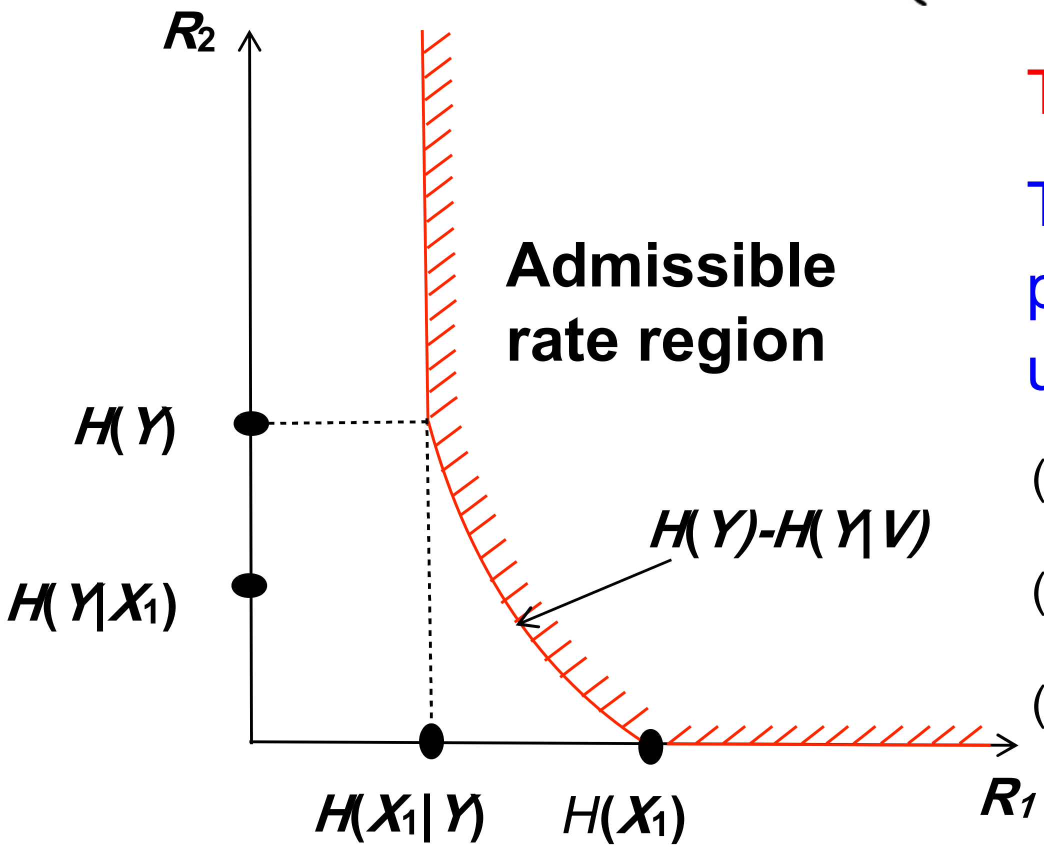
→ One Source One Helper Slepian Wolf Theorem for Lossless Multi-terminal Source Coding

Zhou, Xiaobo, Meng Cheng, Xin He, and Tad Matsumoto, "Exact and approximated outage probability analyses for decode-and-forward relaying system allowing intra-link errors," *IEEE Transactions on Wireless Communications*, vol. 13, no. 12, pp. 7062-7071, Dec. 2014.

LF Rate Region Analysis: *Slepian –Wolf Theorem for Lossless Multi-terminal Source Coding with a helper.*

With LF, the S-R link is lossy, the admissible rate region is given by:

$$\begin{cases} R_1 & \geq H(X_1 | V) \\ R_2 & \geq I(Y; V) = H(Y) - H(Y | V) \end{cases}$$



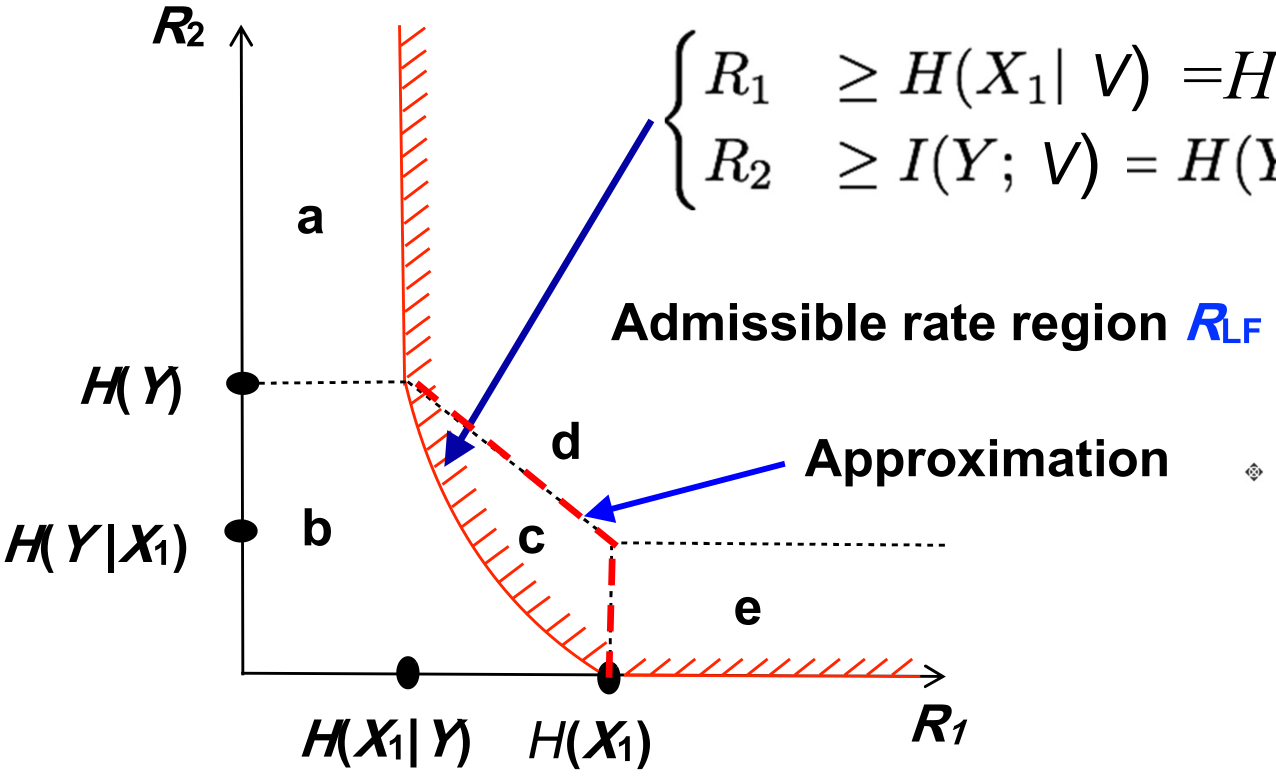
This is a general expression.

To calculate the rate region using parameters related to the links, we use :

(1) Shannon's Source-Channel Separation Theorem,
(2) Test Channel Model of Binary $R(D)$ function to represent each link's threshold, and
(3) Utilization of Markov Chain.
→ Binary Convolution

Rate Region Analysis: we need threefold Integral!

$$\begin{cases} R_1 \geq H(X_1|V) = H_b(\alpha * p_e) \\ R_2 \geq I(Y;V) = H(Y) - H(Y|V) = 1 - H_b(\alpha) \end{cases}$$



- To calculation the outage, we decompose the region into 5 sub-regions.

  - R_a, R_b, R_c, R_d and R_e
  - $R_{LF} = R_c \cup R_d \cup R_e$

- This region is a function of γ_0, γ_1, and γ_2. (SNR of S-R, S-D and R-D links)

  → Threefold integral needed with respect to pdf's of γ_0, γ_1, and γ_2.

LF Rate Region Analysis: ***SR*** *Link*

- How can we combine Shannon's Separation Theorem and the Rate Region?

By using the lossy Separation theorem $R_{c,1} \cdot R(\mathscr{D}) \leq C(\gamma_0)$

and with Inverse $C^{-1}(\gamma_0)$ of the Capacity Function $C(\gamma_0)$, we calculate the binary distortion (D=BER) of the S-R link after decoding, as

$$p_e = \begin{cases} H_b^{-1}[1 - \Phi_1(\gamma_0)], & \text{for } \Phi^{-1}(0) \leq \gamma_0 \leq \Phi_1^{-1}(1) \\ 0, & \text{for } \gamma_0 \geq \Phi_1^{-1}(1), \end{cases}$$

and $\Phi_1(\gamma_0) = \frac{C(\gamma_0)}{R_{c,1}}$ **Separation Theorem**

with $H_b^{-1}(\cdot)$ denotes the inverse function of the binary entropy function $H_b(x) = -x \log_2 x - (1-x)\log_2(1-x)$, and $\Phi_1^{-1}(\cdot)$ is the inverse function of $\Phi_1(\cdot)$.

This means that given the *instantaneous SNR* γ_0 and $R_{c,1}$, we can calculate the binary distortion ($D=BER=p_e$) of S-R link after decoding!

LF Rate Region Analysis: ***Test Channel***

Binary Source

Consider a Binary source $x \in X$, $\quad Prob(x=1)=p \quad , \quad Prob(x=0)=1-p$

Assume that p<1/2. The rate distortion function is given by:

$$R(D)=\begin{cases} H(p)-H(D) \quad , \quad 0 \le D \le \min(p,1-p) \\ 0 \quad , \quad D > \min(p,1-p) \end{cases}$$

where Hamming distortion measure is assumed.

With p=1/2, $R(D)$=1-$H(D)$, hence ***SR*** test channel is BSC with $D=p_e$

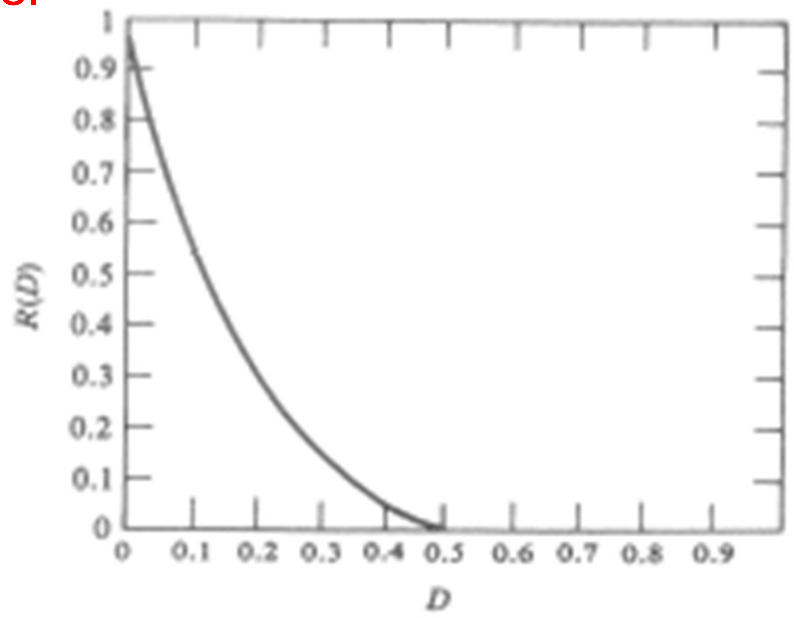
↓

$H_b(p_e)=1-R(p_e)$

↓

$p_e=H_b^{-1}\left(1-\dfrac{C(\gamma_0)}{R_{c,1}}\right)$

Separation Theorem



LF Rate Region Analysis: ***RD Link***

- We also do not need *lossless R-D link*. R-D link's error probability α after decoding can be calculated in the same way, as

$$\alpha = \begin{cases} H_b^{-1}[1 - \Phi_1(\gamma_2)], & \text{for } \Phi_1^{-1}(0) \le \gamma_2 \le \Phi_1^{-1}(1) \\ 0, & \text{for } \gamma_2 \ge \Phi_1^{-1}(1), \end{cases}$$

with $\Phi_1(\gamma_2) = \frac{C(\gamma_2)}{R_{c,1}}$ Separation Theorem

By combining all, we have

$$\begin{cases} R_1 \ge H(X_1 | V) = H_b(\alpha * p_e) \\ R_2 \ge I(Y; V) = H(Y) - H(Y|V) = 1 - H_b(\alpha) \end{cases}$$

, because **V->Y->X_1 forms Markov Chain**.

with $\alpha * p_e = (1-\alpha) p_e + \alpha(1 - p_e)$

We do not know $\gamma_0, \gamma_1, \gamma_2$ but we know their distributions.

To Calculate the Outage, we need threefold Integrals

$$
\begin{aligned}
P_{1,a} &= \Pr\{p=0, R_2 \ge 1, 0 \le R_1 \le H_b(p)\} \\
&= \Pr\{\gamma_0 \ge \Phi_1^{-1}(1), \gamma_2 \ge \Phi_2^{-1}(1), \\
&\qquad \Phi_1^{-1}(0) \le \gamma_1 \le \Phi_1^{-1}(0)\} \\
&= \int_{\Phi_1^{-1}(0)}^{\Phi_1^{-1}(1)} d\gamma_0 \int_{\Phi_2^{-1}(0)}^{\Phi_2^{-1}(1)} d\gamma_2 \\
&\quad \cdot \int_{\Phi_1^{-1}(0)}^{\Phi_1^{-1}(0)} p(\gamma_0)\cdot p(\gamma_1)\cdot p(\gamma_2) d\gamma_1 \\
&= 0,
\end{aligned}
$$

$$
\begin{aligned}
P_{1,b} &= \Pr\{p=0, 0 \le R_2 \le 1, 0 \le R_1 \le H_b(\alpha * p)\} \\
&= \Pr\{\gamma_0 \ge \Phi_1^{-1}(1), \Phi_2^{-1}(0) \le \gamma_2 \le \Phi_2^{-1}(1), \\
&\qquad \Phi_1^{-1}(0) \le \gamma_1 \le \Phi_1^{-1}[1-\Phi_2(\gamma_2)]\} \\
&= \int_{\Phi_1^{-1}(1)}^{\Phi_1^{-1}(\infty)} d\gamma_0 \int_{\Phi_2^{-1}(0)}^{\Phi_2^{-1}(1)} d\gamma_2 \\
&\quad \cdot \int_{\Phi_1^{-1}(0)}^{\Phi_1^{-1}[1-\Phi_2(\gamma_2)]} p(\gamma_0)\cdot p(\gamma_1)\cdot p(\gamma_2) d\gamma_1 \\
&= \frac{1}{\Gamma_2}\exp\left[-\frac{\Phi_1^{-1}(1)}{\Gamma_0}\right] \int_{\Phi_2^{-1}(0)}^{\Phi_2^{-1}(1)} \exp(-\frac{\gamma_2}{\Gamma_2}) \\
&\quad \cdot \left[1-\exp(-\frac{\Phi_1^{-1}[1-\Phi_2(\gamma_2)]}{\Gamma_1})\right] d\gamma_2,
\end{aligned}
$$

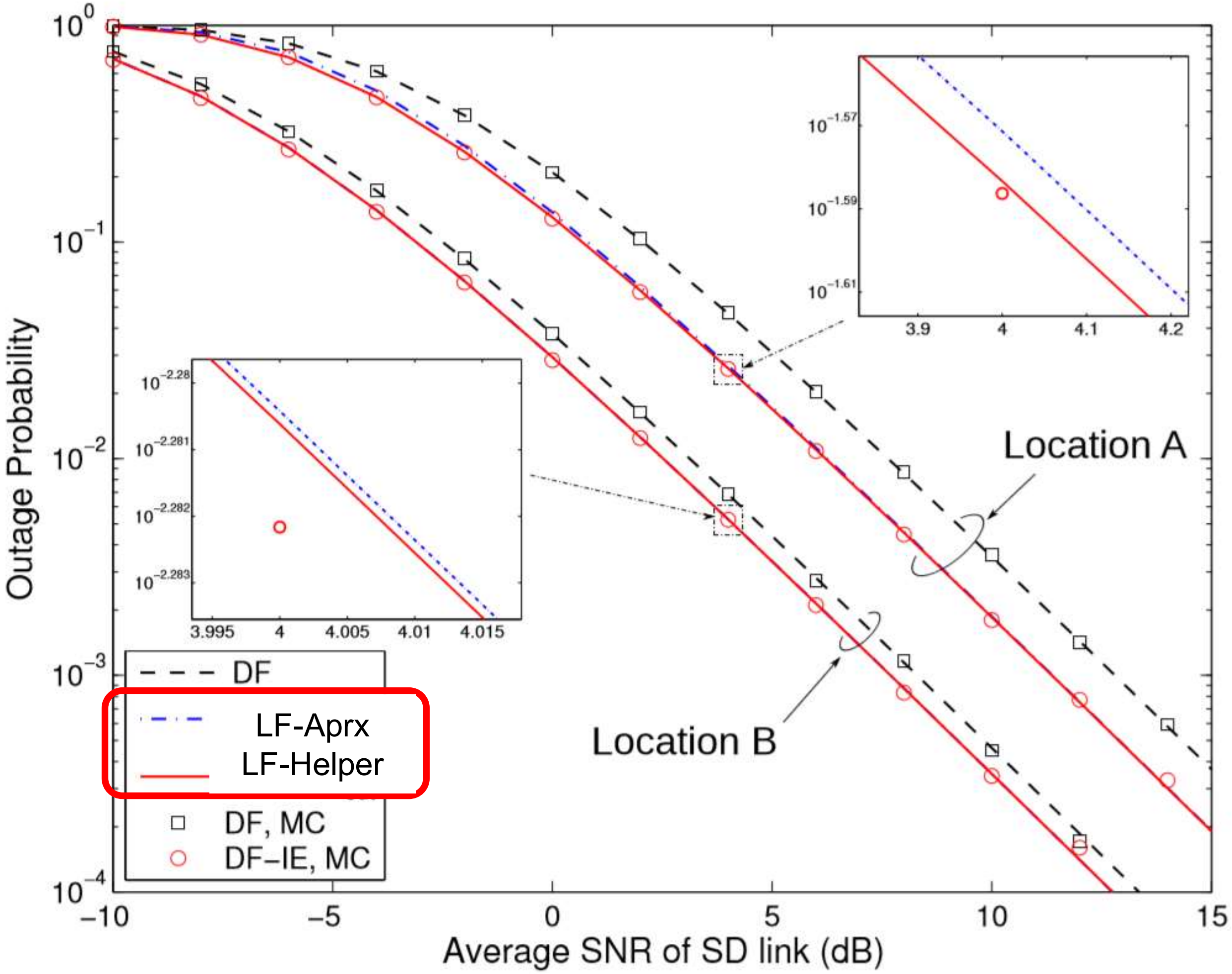
$$
\begin{aligned}
P_{2,a} &= \Pr\{0 < p \le 0.5, R_2 \ge 1, 0 \le R_1 \le H_b(p)\} \\
&= \Pr\{\Phi_1^{-1}(0) \le \gamma_0 \le \Phi_1^{-1}(1), \gamma_2 \ge \Phi_2^{-1}(1), \\
&\qquad \Phi_1^{-1}(0) \le \gamma_1 \le \Phi_1^{-1}[1-\Phi_1(\gamma_0)]\} \\
&= \int_{\Phi_1^{-1}(0)}^{\Phi_1^{-1}(1)} d\gamma_0 \int_{\Phi_2^{-1}(1)}^{\Phi_2^{-1}(\infty)} d\gamma_2 \\
&\quad \cdot \int_{\Phi_1^{-1}(0)}^{\Phi_1^{-1}[1-\Phi_1(\gamma_0)]} p(\gamma_0)\cdot p(\gamma_1)\cdot p(\gamma_2) d\gamma_1 \\
&= \frac{1}{\Gamma_0}\exp\left[-\frac{\Phi_2^{-1}(1)}{\Gamma_2}\right] \int_{\Phi_1^{-1}(0)}^{\Phi_1^{-1}(1)} \exp(-\frac{\gamma_0}{\Gamma_0}) \\
&\quad \cdot \left[1-\exp(-\frac{\Phi_1^{-1}[1-\Phi_1(\gamma_0)]}{\Gamma_1})\right] d\gamma_0,
\end{aligned}
$$

Let's skip the boring threefold integral calculations!

Those who Want to Know the Details of the Calculation, please Come to my Place after the Tutorial (TwWKDCpCmPaT).

$$
\begin{aligned}
P_{2,b} &= \Pr\{0 < p \le 0.5, 0 \le R_2 \le 1, 0 \le R_1 \le H_b(\alpha * p)\} \\
&= \Pr\{\Phi_1^{-1}(0) \le \gamma_0 \le \Phi_1^{-1}(1), \Phi_2^{-1}(0) \le \gamma_2 \le \Phi_2^{-1}(1), \\
&\qquad \Phi_1^{-1}(0) \le \gamma_1 \le \Phi_1^{-1}[\Psi(\gamma_0,\gamma_2)]\} \\
&= \int_{\Phi_1^{-1}(0)}^{\Phi_1^{-1}(1)} d\gamma_0 \int_{\Phi_2^{-1}(0)}^{\Phi_2^{-1}(1)} d\gamma_2 \\
&\quad \cdot \int_{\Phi_1^{-1}(0)}^{\Phi_1^{-1}[\Psi(\gamma_0,\gamma_2)]} p(\gamma_0)\cdot p(\gamma_1)\cdot p(\gamma_2) d\gamma_1 \\
&= \frac{1}{\Gamma_0\Gamma_2} \int_{\Phi_1^{-1}(0)}^{\Phi_1^{-1}(1)} \int_{\Phi_2^{-1}(0)}^{\Phi_2^{-1}(1)} \exp(-\frac{\gamma_0}{\Gamma_0}-\frac{\gamma_2}{\Gamma_2}) \\
&\quad \cdot \left\{1-\exp\left[-\frac{\Phi_1^{-1}[\Psi(\gamma_0,\gamma_2)]}{\Gamma_1}\right]\right\} d\gamma_0 d\gamma_2
\end{aligned}
$$

Comparison of exact and approximated SW region with a helper (Orthogonal Case)



Location A, d0=d1=d2

Location B, d0=(1/4)d1, d2=(3/4)d1

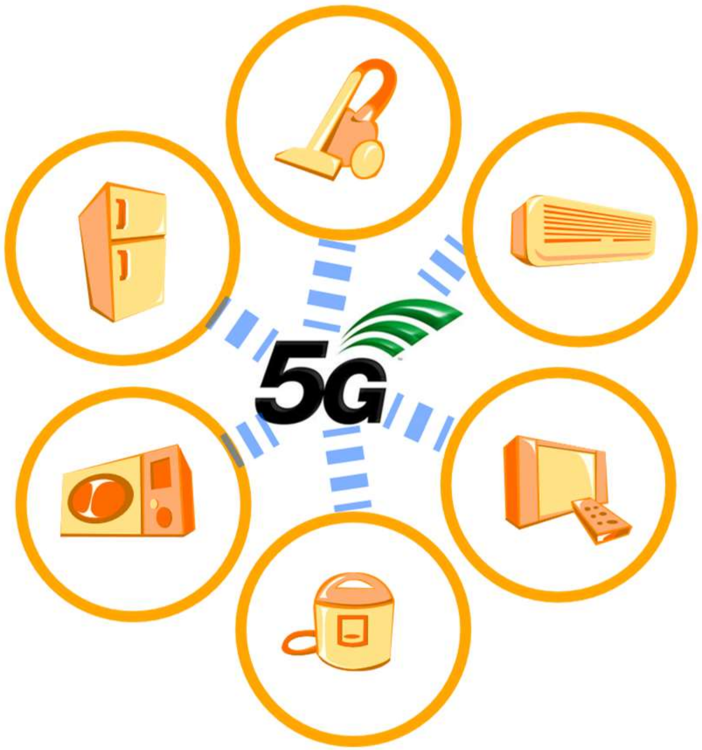
[1] X. Zhou, M. Cheng, X. He and T. Matsumoto, "Exact and Approximated Outage Probability Analyses for Decode-and- Forward Relaying System Allowing Intra-Link Errors," in IEEE Transactions on Wireless Communications, vol. 13, no. 12, pp. 7062-7071, Dec. 2014.

Chapter 2. End-to-End Lossy Distributed Multi-terminal Networks: Rate Distortion Analysis

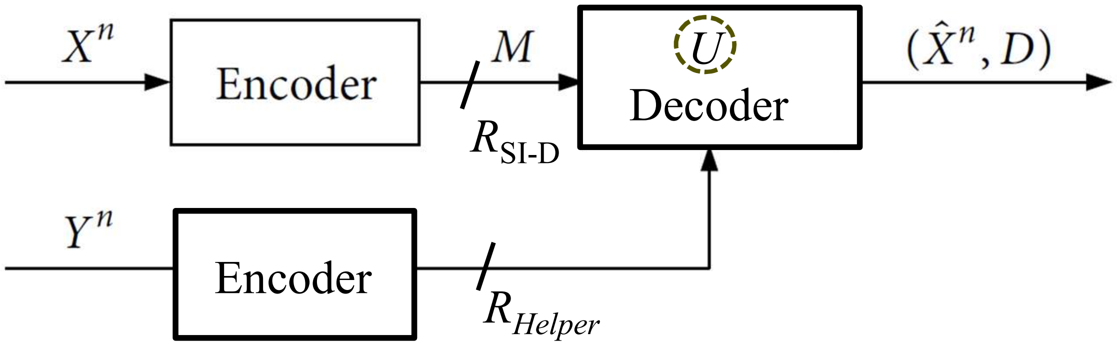
2.1 Wyner-Ziv Formulation for End-to-End Lossy Two-Way Relay Network

- Internet of Things (IoT)

  Connect objects to make Right Decisions→ E2E Lossy

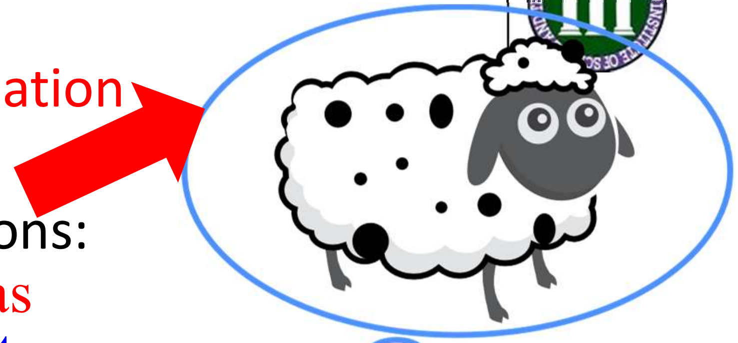


WZ with a Helper:

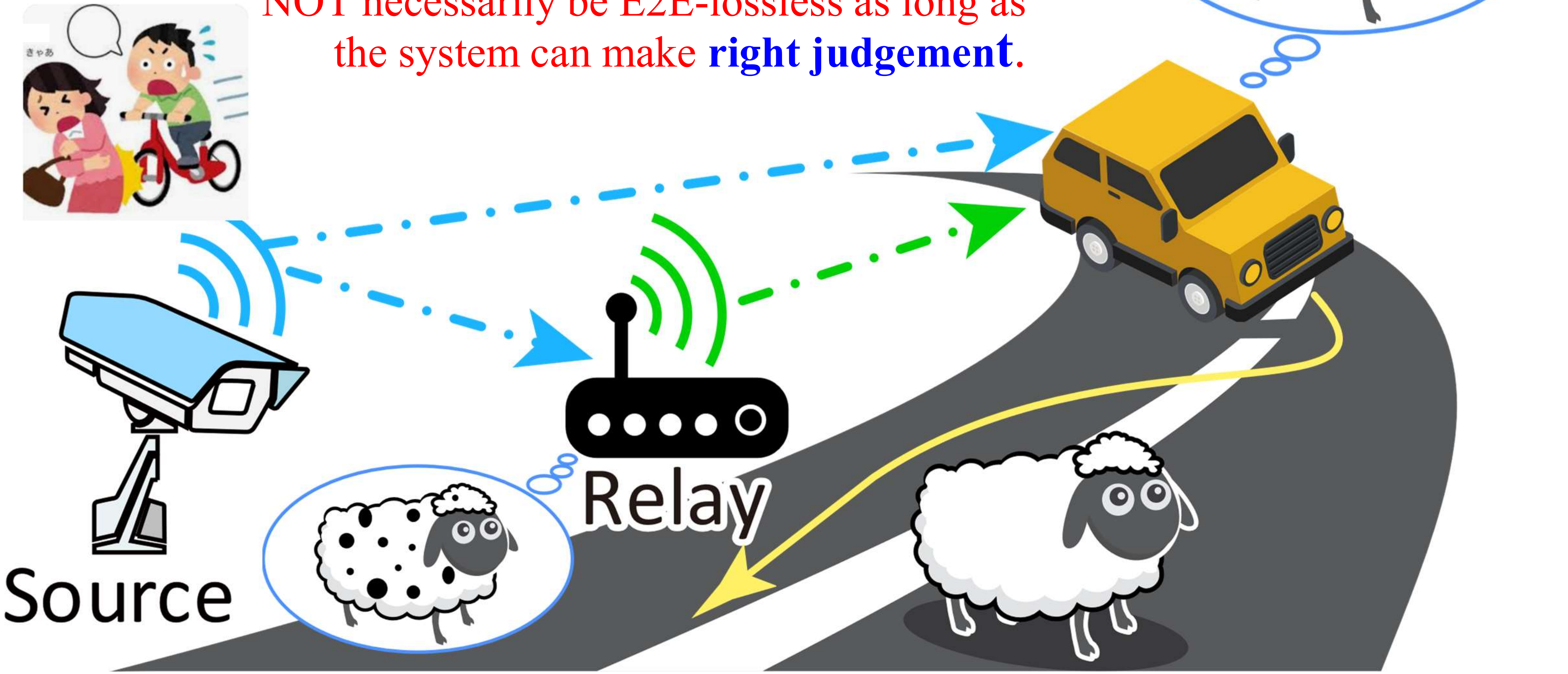


E2E Lossy Communications with LF

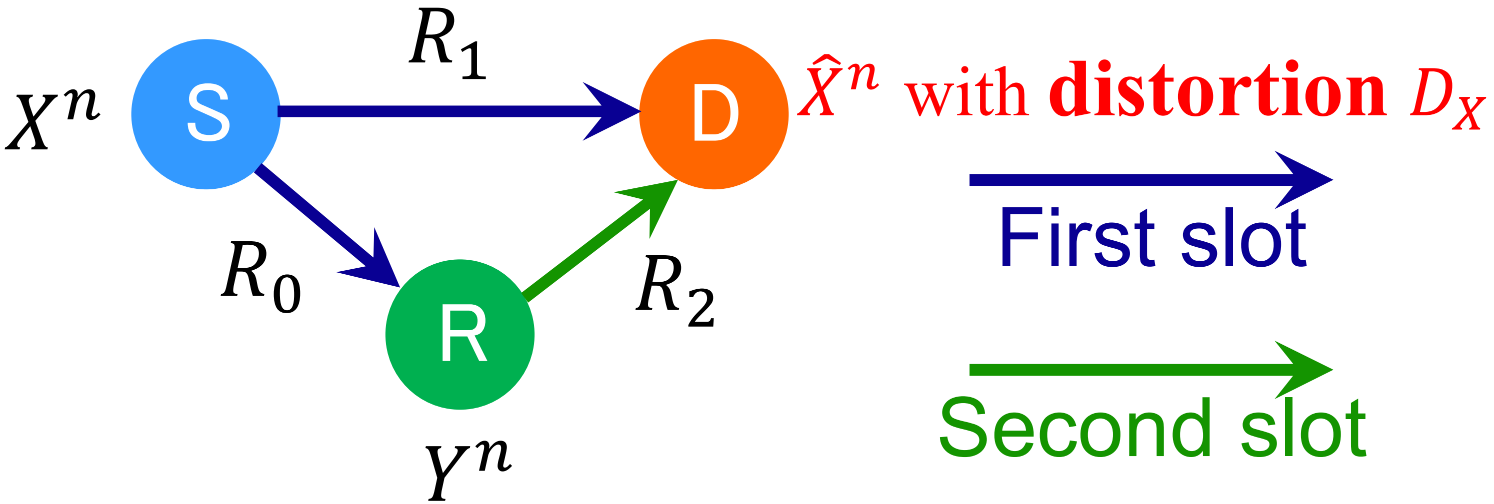
- Objective of IoT:
  - Make a judgement rather than recovering information itself
  - The picture exemplifies LF for lossy communications:



NOT necessarily be E2E-lossless as long as the system can make **right judgement**.

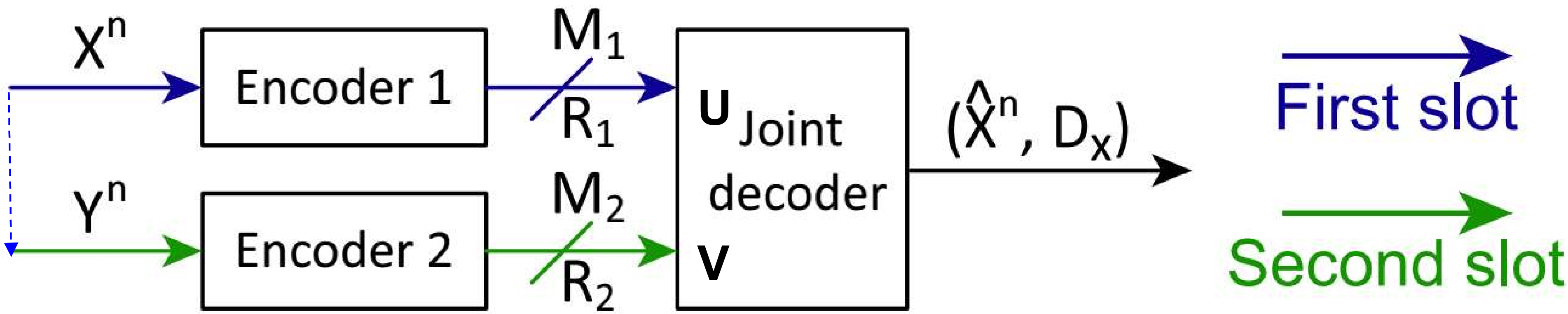


Lossy Distributed Multi-terminal Source Coding



- S-R link: point-to-point communication.
- S-D and R-D links: distributed lossy multi-terminal source coding problem.

→ As a whole, **Wyner-Ziv** Problem



Rate-Distortion Region

- WZ *R*(*D*) function for general sources

$$R_1 > I(X;U|V),$$
$$R_2 > I(Y;V).$$

- For binary sources
  - S-R link

$$R_0 > 1 - H_b(\rho).$$

  - R-D link

$$R_2 > 1 - H_b(\rho').$$

  - S-D link

$$R_1 > H_b(\rho' * \rho * D_X) - H_b(D_X).$$

because V->Y->X->U forms a Markov Chain.

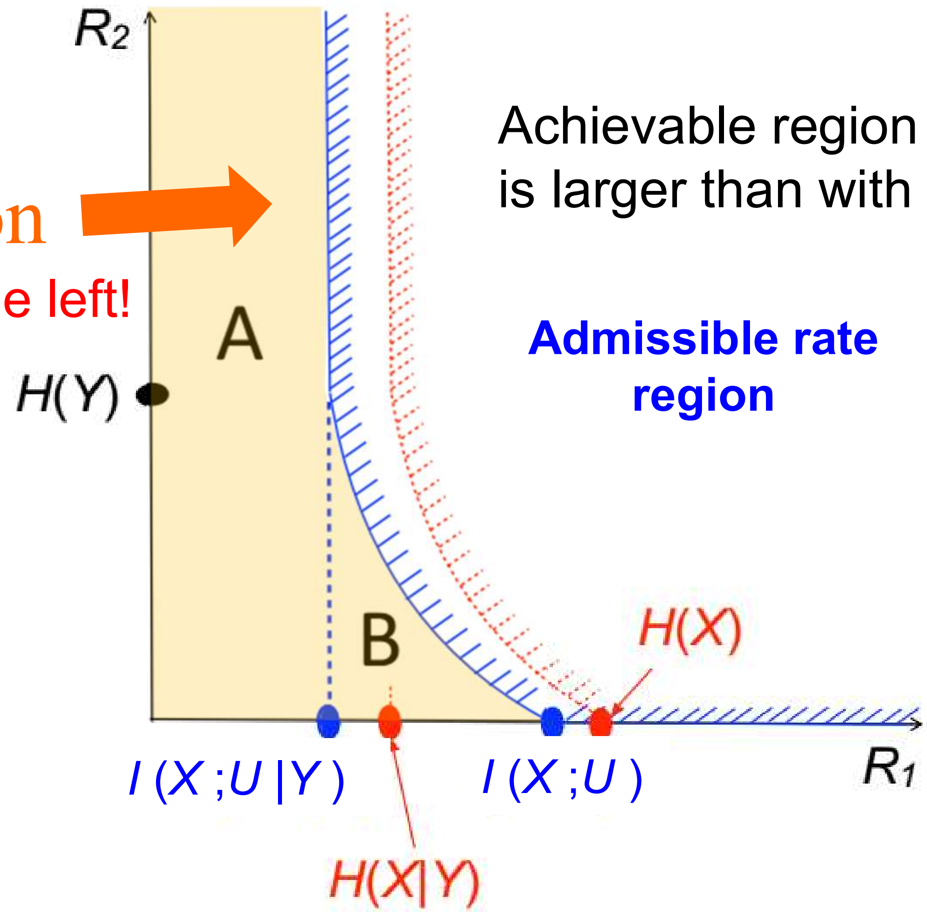
ρ: crossover probability between X and Y

ρ': crossover probability between Y and V

Outage Event

Outage region

The curve is shifted to the left!



Achievable region with **Lossy** WZ
is larger than with **Lossless** helper-SW!

Admissible rate region

- The link rates (R_0, R_1, R_2) supported by channel capacities cannot satisfy the distortion requirement D_X, when they fall outside the achievable rate-distortion region.→ Outage

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

CWC
Oulu

Again, we need multifold Integrals

$$
\begin{aligned}
\mathrm{P_{out}} =& \Pr\{(R_1,R_2)\in\alpha\cup\beta\}\\
=& \Pr\{p=0,(R_1,R_2)\in\alpha\cup\beta\}\\
&+\Pr\{p\in(0,0.5],(R_1,R_2)\in\alpha\cup\beta\}\\
=& \Pr\{p=0,(R_1,R_2)\in\alpha\}\\
&+\Pr\{p=0,(R_1,R_2)\in\beta\}\\
&+\Pr\{p\in(0,0.5],(R_1,R_2)\in\alpha\}\\
&+\Pr\{p\in(0,0.5],(R_1,R_2)\in\beta\}\\
=& \Pr\{0\le R_1\le H_b(p*D_X)-H_b(D_X),0\le R_2,\\
&\quad p=0\}\\
&+\Pr\{H_b(p*D_X)-H_b(D_X)\\
&\quad \le R_1\le H_b(p'*p*D_X)-H_b(D_X),\\
&\quad 0\le R_2\le 1,p=0\}\\
&+\Pr\{0\le R_1\le H_b(p*D_X)-H_b(D_X),0\le R_2,\\
&\quad 0<p\le 0.5\}\\
&+\Pr\{H_b(p*D_X)-H_b(D_X)\\
&\quad \le R_1\le H_b(p'*p*D_X)-H_b(D_X),\\
&\quad 0\le R_2\le 1,0<p\le 0.5\}\\
=& P_{1,\alpha}+P_{1,\beta}+P_{2,\alpha}+P_{2,\beta},
\end{aligned}
$$

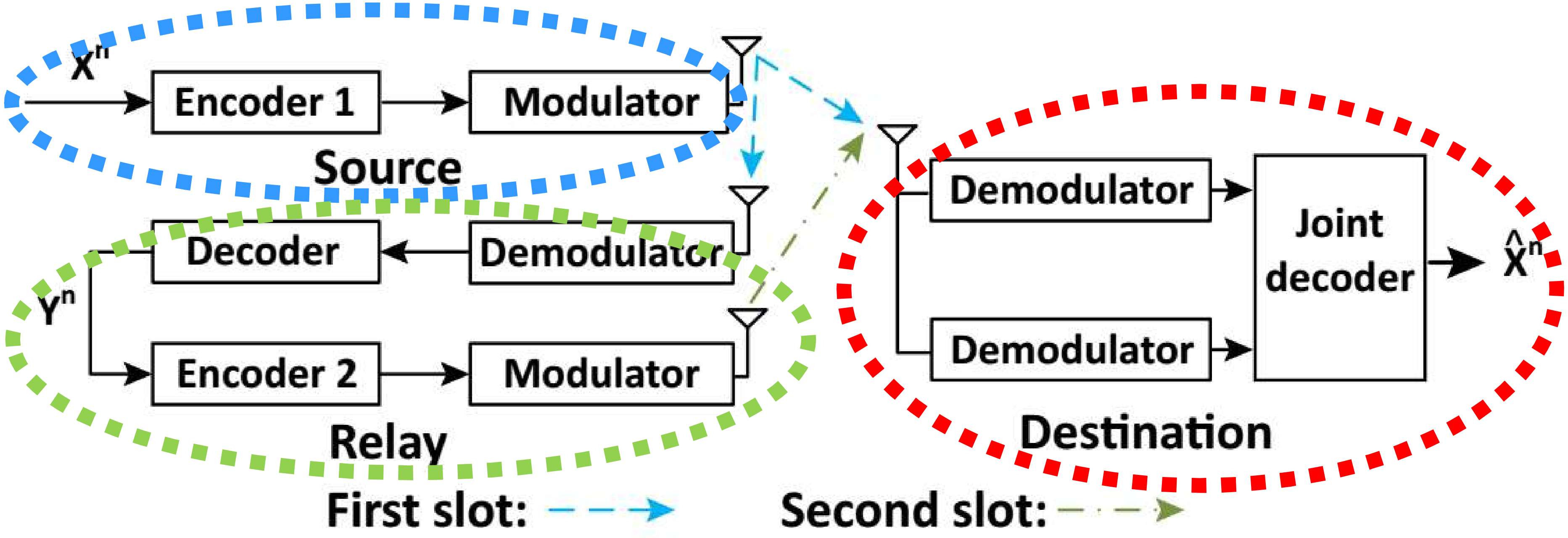
$$
\begin{aligned}
P_{1,\alpha} =& \Pr\{0\le R_1\le H_b(0*D_X)-H_b(D_X),0\le R_2\\
&\quad p=0\}\\
=& \Pr\{0\le R_1\le 0,0\le R_2,p=0\}\\
=& \Pr\{\Theta_1^{-1}(0)\le\gamma_1\le\Theta_1^{-1}(0),\Theta_2^{-1}(0)\le\gamma_2,\\
&\quad \Theta_0^{-1}(1)\le\gamma_0\}\\
=& \int_{\Theta_2^{-1}(0)}^{\infty}d\gamma_2\int_{\Theta_1^{-1}(0)}^{\Theta_1^{-1}(0)}d\gamma_1\\
&\cdot\int_{\Theta_0^{-1}}^{\infty}f(\gamma_0)f(\gamma_1)f(\gamma_2)d\gamma_0,\\
=& 0,
\end{aligned}
$$

$$
\begin{aligned}
P_{1,\beta} =& \Pr\{H_b(0*D_X)-H_b(D_X)\\
&\quad \le R_1\le H_b(p'*0*D_X)-H_b(D_X),\\
&\quad 0\le R_2\le 1,p=0\}\\
=& \Pr\{0\le R_1\le H_b(p'*D_X)-H_b(D_X),\\
&\quad 0\le R_2\le 1,p=0\}\\
=& \Pr\{\Theta_1^{-1}(0)\le\gamma_1\\
&\quad \le\Theta_1^{-1}[H_b(\xi(\gamma_2,D_X))-H_b(D_X)],\\
&\quad \Theta_2^{-1}(0)\le\gamma_2\le\Theta_2^{-1}(1),\Theta_0^{-1}(1)\le\gamma_0\}\\
=& \int_{\Theta_2^{-1}(0)}^{\Theta_2^{-1}(1)}d\gamma_2\int_{\Theta_1^{-1}(0)}^{\Theta_1^{-1}[H_b(\xi(\gamma_2,D_X))-H_b(D_X)]}d\gamma_1\\
&\cdot\int_{\Theta_0^{-1}(1)}^{\infty}f(\gamma_0)f(\gamma_1)f(\gamma_2)d\gamma_0\\
=& \frac{1}{\overline{\gamma}_2}\exp\left(-\frac{\Theta_0^{-1}(1)}{\overline{\gamma}_0}\right)\int_{\Theta_2^{-1}(0)}^{\Theta_2^{-1}(1)}\exp\left(-\frac{\gamma_2}{\overline{\gamma}_2}\right)\cdot\Bigg[1\\
&-\exp\left(-\frac{\Theta_1^{-1}\{H_b[\xi(\gamma_2,D_X)]-H_b(D_X)\}}{\overline{\gamma}_1}\right)\Bigg]d\gamma_2,
\end{aligned}
$$

Again, we need threefold Integrals!

We calculated, but too boring. Let's skip it!

We evaluate the outage probability also by chain simulations using very simple signaling and joint decoding techniques.

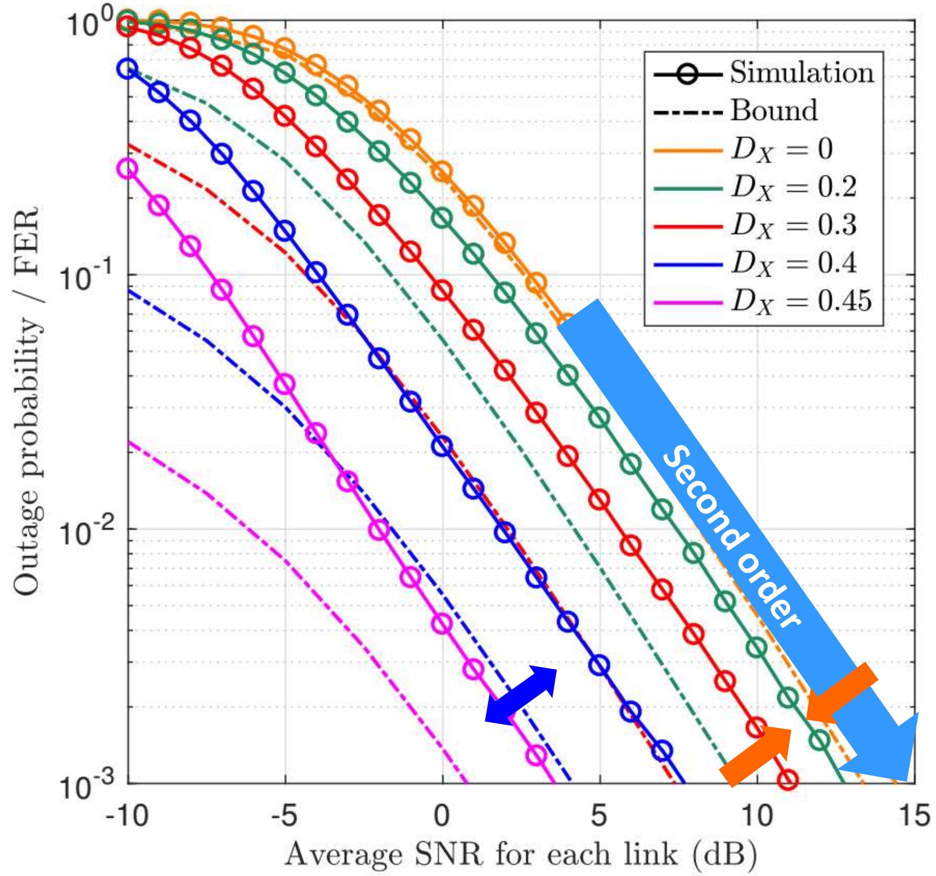


Simulation Results

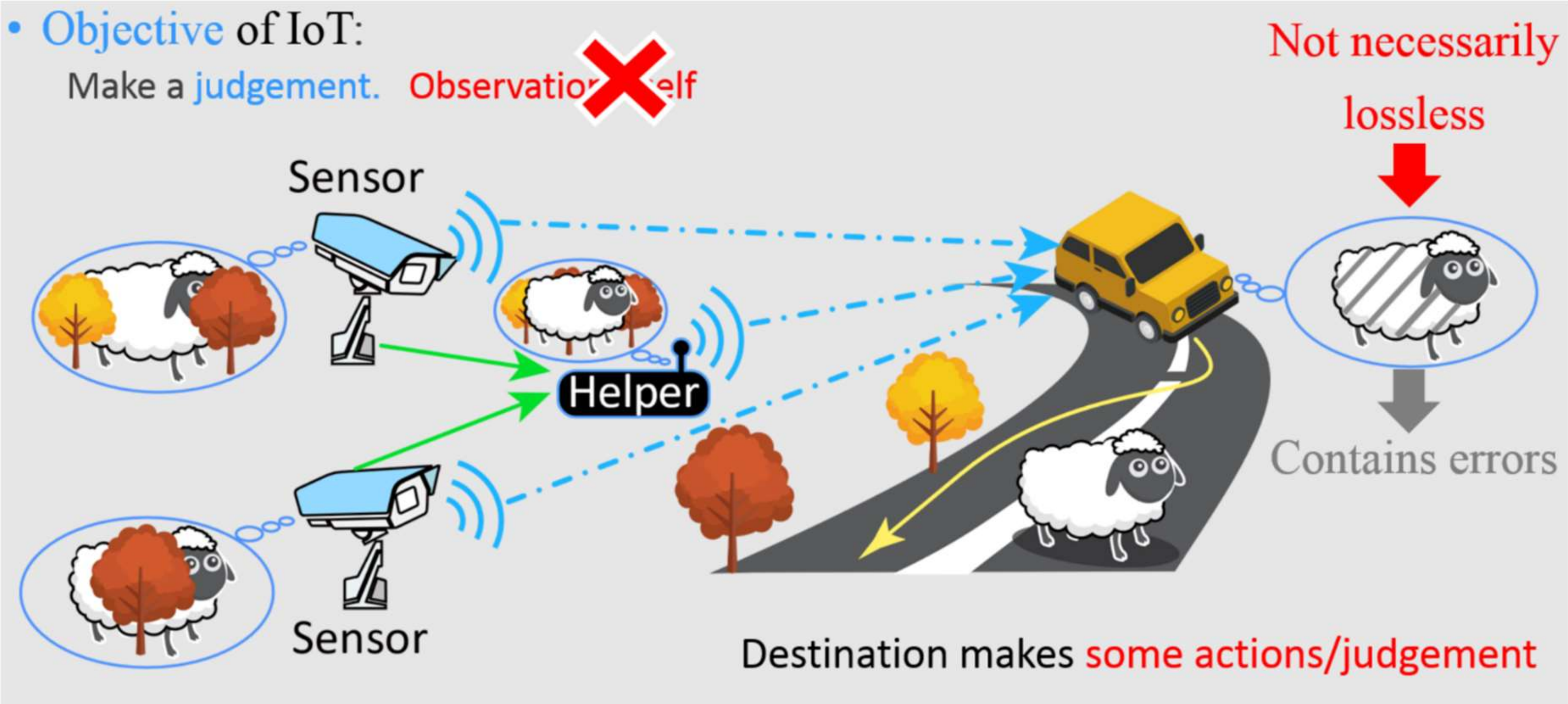
1. Simulation results have the same tendency and the slope decay (=Diversity Order) as the theoretical bound.
2. The gap between the simulation and theoretical results becomes larger as D_X increases. →We need *more efficient rate-distortion code.*

SIMULATION PARAMETERS

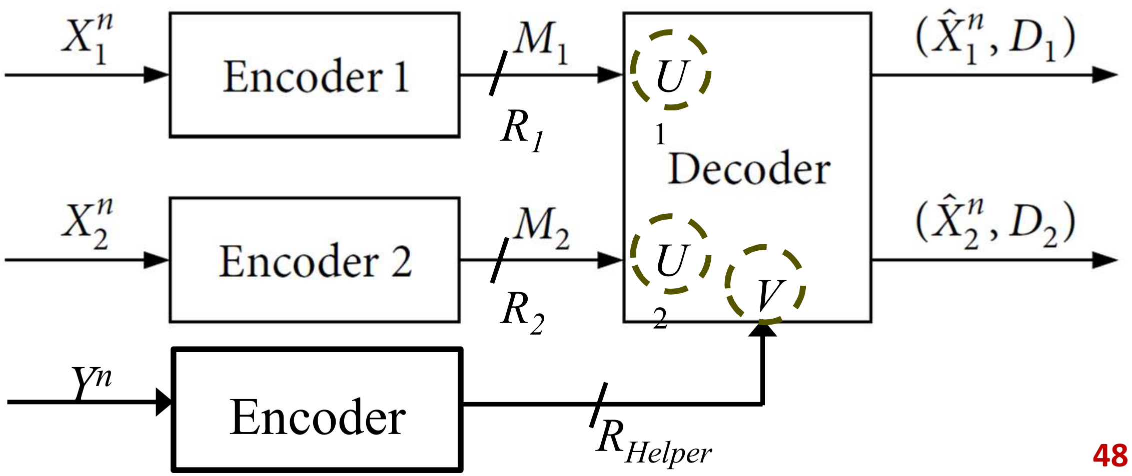
Parameter	Value
Frame length	10^4 bits
Number of frames	5×10^5
Source coding rate	1
Channel coding rate	1/2
Generator polynomial of CC	$G = ([3, 2]3)_8$
Type of interleaver	random interleaver
Modulation method	BPSK
Maximum iteration time	20



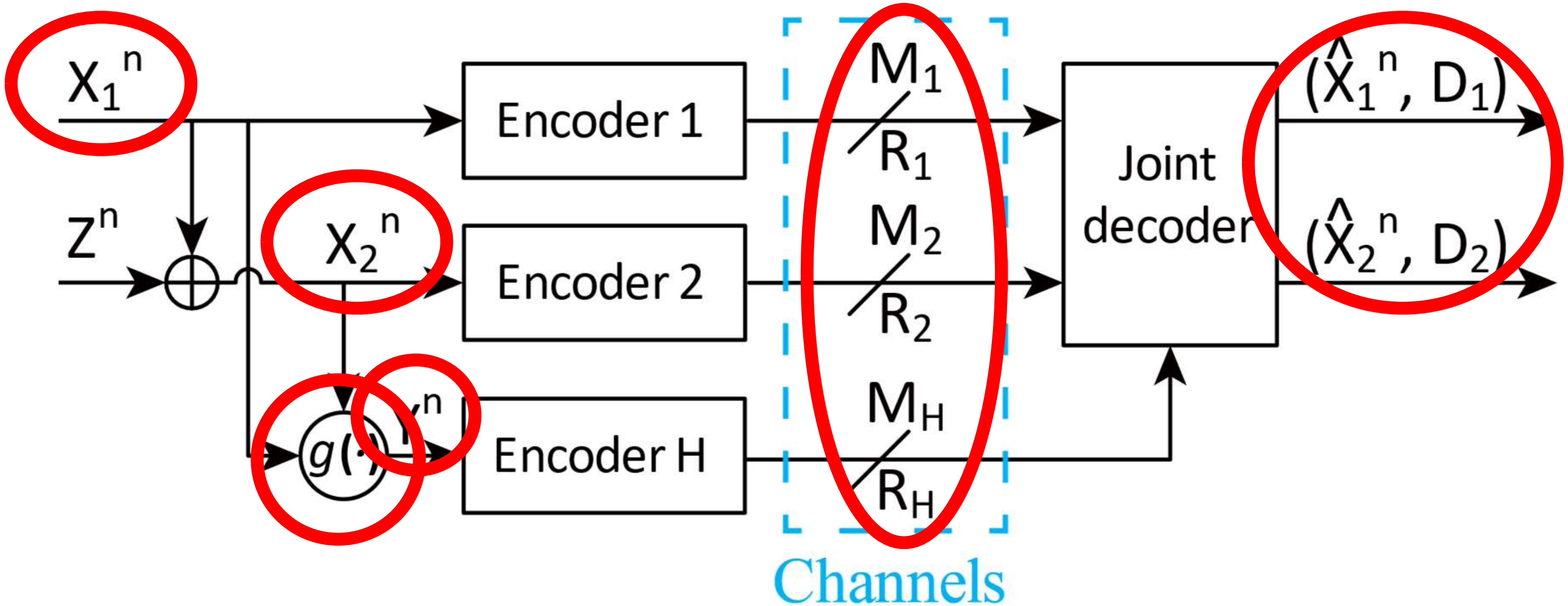
2.2 Berger-Tung (BT) Formulation for Two Source One Helper Network



Multiple Access Relay Channel formulated by BT with a Helper:



System Model



X_1, X_2: two sources M_1, M_2, M_H: codewords $\hat{X}_1^n, \hat{X}_2^n$: recovered sequences

Y : helper information R_1, R_2, R_H: link rates D_1, D_2: distortion requirements

$g(\cdot)$: optimal helper (it utilizes the useful information at the maximum level, since $(U_1, U_2) \to (X_1, X_2) \to Y \to V$ and by data processing theorem, $I(U_1, U_2; V) \le I(Y; V) \le [R_H]^-$)

Lin Wensheng, Qiang Xue, Jiguang He, Markku Juntti, and Tad Matsumoto, "Rate-Distortion and Outage Probability Analyses for Single Helper Assisted Lossy Communications", *IEEE Transactions on Vehicular Technology*, Vol. 68, 2019, No. 11, pp. 10882-10894, **DOI:**10.1109/TVT.2019.2939622

Rate-Distortion Analysis

- What is the achievable rate-distortion region?
  - Condition for reliable communications: link rates can support the transmissions to satisfy the distortion requirements.
- Inner bound on the achievable rate-distortion region, given by Berger Tung Bound

$$
\begin{aligned}
R_1 &> I(X_1; U_1 | U_2, V, Q), \\
R_2 &> I(X_2; U_2 | U_1, V, Q), \\
R_1 + R_2 &> I(X_1, X_2; U_1, U_2 | V, Q), \\
R_{\mathrm{H}} &> I(Y; V),
\end{aligned}
$$

U_i: compressed information of X_i

V: compressed information of Y

Q: an auxiliary variable resulting from time-sharing scheme

Use Inner bound→ Upper bound of the outage probability

- The inner bound for binary sources.
  (a) for some $0 \le \tilde{d} \le D_2$,
  $$\begin{cases} R_1 > H_b(D_1 * \rho * \tilde{d}) - H_b(D_1) - [R_{\mathrm{H}}]^-, \\ R_2 > 1 - H_b(\tilde{d}), \end{cases}$$
  (b) for some $0 \le \tilde{d} \le D_1$,
  $$\begin{cases} R_1 > 1 - H_b(\tilde{d}), \\ R_2 > H_b(\tilde{d} * \rho * D_2) - H_b(D_2) - [R_{\mathrm{H}}]^-, \end{cases}$$
  (c) common case,
  $$R_1 + R_2 > 1 + H_b(D_1 * \rho * D_2) - H_b(D_1) - H_b(D_2) - [R_{\mathrm{H}}]^-,$$

ρ: crossover probability between X_1 and X_2 (representing correlation)

$\tilde{d}$: dummy variable

H_b: binary entropy function

$a * b = a(1-b) + b(1-a)$

$[R_{\mathrm{H}}]^- = \min\{1, R_{\mathrm{H}}\}$

- This is not an inner bound in general. It is only for the case that the following inequality holds with equality.

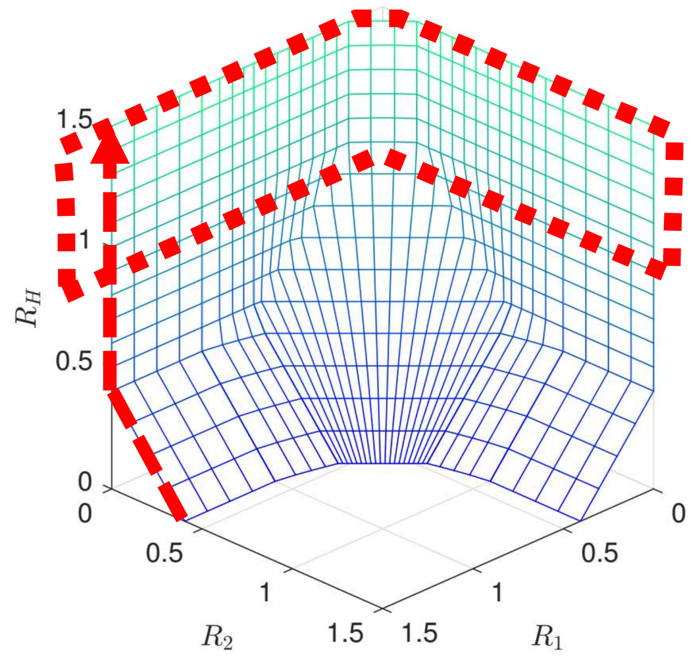
$$I(U_{\boldsymbol{S}}; V | U_{\boldsymbol{S}^c}) \le [R_{\mathrm{H}}]^-,$$

where $\boldsymbol{S} \subseteq \{1,2\}$, and $\boldsymbol{S}^c$ represents the complementary set of $\boldsymbol{S}$.

- The shape of achievable rate-distortion region.

1. The achievable rate-distortion region is expanded as R_{H} increases.

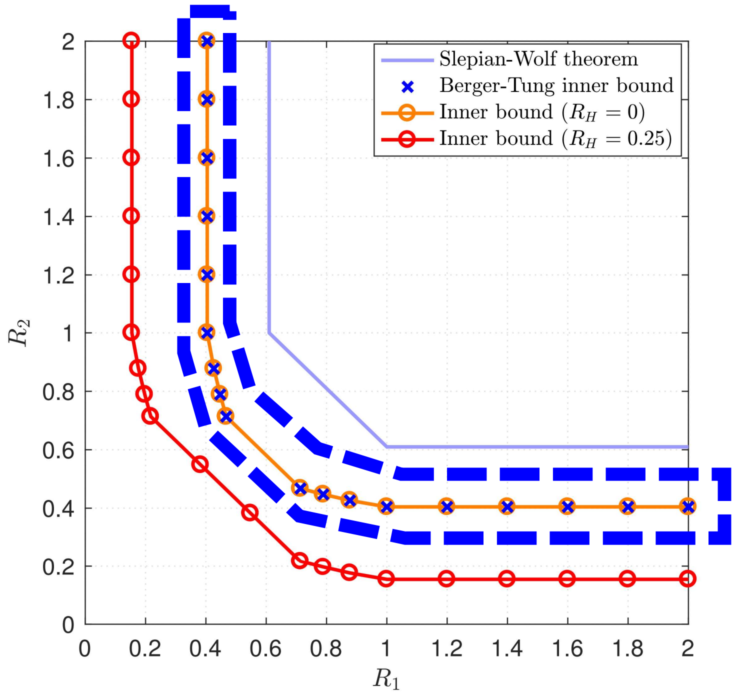
2. However, the above part of the region for $R_{\mathrm{H}} \geq 1$, does not change even if the helper rate continues increasing.



$$\rho = 0.15, D_1 = D_2 = 0.05$$

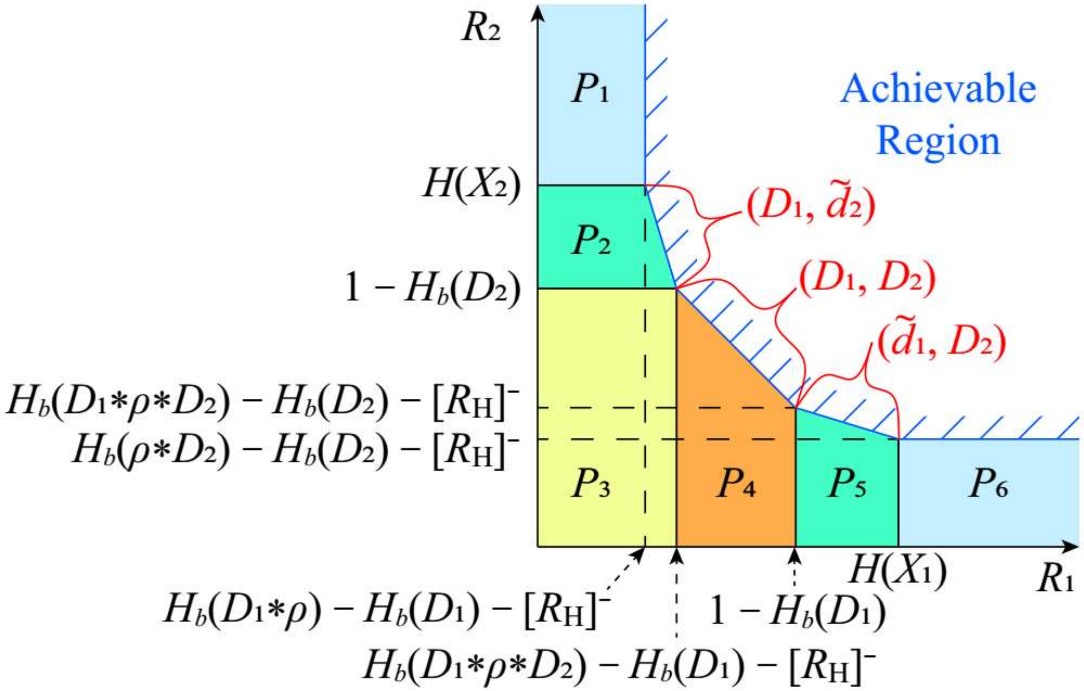
- The achievable rate-distortion region projected on the R_1- R_2 plane by given R_{H}.

The derived inner bound perfectly coincides with the Berger-Tung inner bound when $R_{\mathrm{H}} = 0$ (equivalent to no helper).



$\rho = 0.15$
$D_1 = D_2 = 0.05$

Outage Probability Analysis



Outage event defined as:

- The link rates fall outside the achievable rate-distortion region, i.e., the link rates $(R_1, R_2, R_{\mathrm{H}})$ supported by channel capacities cannot satisfy the distortion requirements (D_1, D_2).

- The outage probability is the multiple integral with respect to $(R_1, R_2, R_{\mathrm{H}})$.

$$P_{\mathrm{out}} = \iiint \cdots dR_1\, dR_2\, dR_{\mathrm{H}}$$

- The instantaneous rates $(R_1, R_2, R_{\mathrm{H}})$ are supported by the instantaneous signal-to-noise ratios (SNRs) $(\gamma_1, \gamma_2, \gamma_{\mathrm{H}})$.

- The outage probability can be calculated by the threefold integral with respect to SNRs $(\gamma_1, \gamma_2, \gamma_{\mathrm{H}})$.

$$P_{\mathrm{out}} = \iiint \cdots d\gamma_1\, d\gamma_2\, d\gamma_{\mathrm{H}}$$

Again, we need multifold Integrals

$$
\begin{aligned}
P_1 &= \Pr\{0 \le R_1 \le H_2(D_1 * \rho) - H_2(D_1) - [R_{\mathrm{H}}]^-, \\
&\qquad H(X_2) \le R_2, 0 \le R_{\mathrm{H}}\} \\
&= \Pr\{0 \le \Theta_1(\gamma_1) \le \lambda_1(0), H(X_2) \le \Theta_2(\gamma_2), \\
&\qquad 0 \le \Theta_{\mathrm{H}}(\gamma_{\mathrm{H}})\} \\
&= \Pr\{\Theta_1^{-1}(0) \le \gamma_1 \le \Theta_1^{-1}[\lambda_1(0)], \\
&\qquad \Theta_2^{-1}[H(X_2)] \le \gamma_2, \Theta_{\mathrm{H}}^{-1}(0) \le \gamma_{\mathrm{H}}\} \\
&= \int_{\Theta_{\mathrm{H}}^{-1}(0)}^{\infty} d\gamma_{\mathrm{H}} \int_{\Theta_1^{-1}(0)}^{\Theta_1^{-1}[\lambda_1(0)]} d\gamma_1 \\
&\quad \cdot \int_{\Theta_2^{-1}[H(X_2)]}^{\infty} p(\gamma_2)p(\gamma_1)p(\gamma_{\mathrm{H}})d\gamma_2 \\
&= \frac{1}{\overline{\gamma}_{\mathrm{H}}} \cdot \exp\left(-\frac{\Theta_2^{-1}[H(X_2)]}{\overline{\gamma}_2}\right) \cdot \int_{\Theta_{\mathrm{H}}^{-1}(0)}^{\Theta_{\mathrm{H}}^{-1}(1)} \exp\left(-\frac{\gamma_{\mathrm{H}}}{\overline{\gamma}_{\mathrm{H}}}\right) \\
&\quad \cdot \left[1 - \exp\left(-\frac{\Theta_1^{-1}[\lambda_1(0)]}{\overline{\gamma}_1}\right)\right] d\gamma_{\mathrm{H}}. \qquad (41)
\end{aligned}
$$

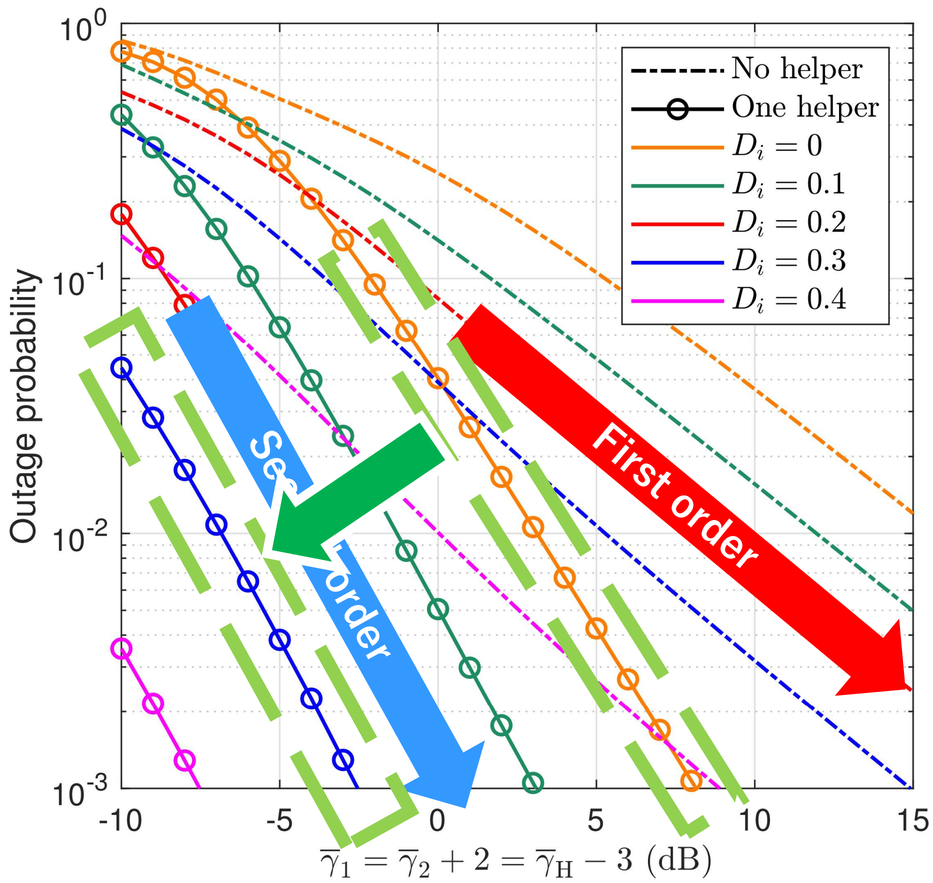
$$
\begin{aligned}
P_2 &= \Pr\{0 \le R_1 \le H_2(D_1 * \rho * \tilde{d}_2) - H_2(D_1) - [R_{\mathrm{H}}]^-, \\
&\qquad 1 - H_2(D_2) \le R_2 < H(X_2), 0 \le R_{\mathrm{H}}\} \\
&= \Pr\{\Theta_1^{-1}(0) \le \gamma_1 \le \Theta_1^{-1}[\lambda_1(\tilde{d}_2)], \\
&\qquad \Theta_2^{-1}[1 - H_2(D_2)] \le \gamma_2 < \Theta_2^{-1}[H(X_2)], \\
&\qquad \Theta_{\mathrm{H}}^{-1}(0) \le \gamma_{\mathrm{H}}\} \\
&= \int_{\Theta_{\mathrm{H}}^{-1}(0)}^{\infty} d\gamma_{\mathrm{H}} \int_{\Theta_2^{-1}[1-H_2(D_2)]}^{\Theta_2^{-1}[H(X_2)]} d\gamma_2 \\
&\quad \cdot \int_{\Theta_1^{-1}(0)}^{\Theta_1^{-1}[\lambda_1(\tilde{d}_2)]} p(\gamma_1)p(\gamma_2)p(\gamma_{\mathrm{H}})d\gamma_1 \\
&= \int_{\Theta_{\mathrm{H}}^{-1}(0)}^{\Theta_{\mathrm{H}}^{-1}(1)} d\gamma_{\mathrm{H}} \int_{\Theta_2^{-1}[1-H_2(D_2)]}^{\Theta_2^{-1}[H(X_2)]} d\gamma_2 \\
&\quad \cdot \int_{\Theta_1^{-1}(0)}^{\Theta_1^{-1}[\lambda_1(\tilde{d}_2)]} p(\gamma_1)p(\gamma_2)p(\gamma_{\mathrm{H}})d\gamma_1 \\
&\quad + \int_{\Theta_{\mathrm{H}}^{-1}(1)}^{\infty} d\gamma_{\mathrm{H}} \int_{\Theta_2^{-1}[1-H_2(D_2)]}^{\Theta_2^{-1}[H(X_2)]} d\gamma_2 \\
&\quad \cdot \int_{\Theta_1^{-1}(0)}^{\Theta_1^{-1}[0]} p(\gamma_1)p(\gamma_2)p(\gamma_{\mathrm{H}})d\gamma_1 \\
&= \int_{\Theta_{\mathrm{H}}^{-1}(0)}^{\Theta_{\mathrm{H}}^{-1}(1)} d\gamma_{\mathrm{H}} \int_{\Theta_2^{-1}[1-H_2(D_2)]}^{\Theta_2^{-1}[H(X_2)]} d\gamma_2 \\
&\quad \cdot \int_{\Theta_1^{-1}(0)}^{\Theta_1^{-1}[\lambda_1(\tilde{d}_2)]} p(\gamma_1)p(\gamma_2)p(\gamma_{\mathrm{H}})d\gamma_1 + 0 \\
&= \frac{1}{\overline{\gamma}_2\overline{\gamma}_{\mathrm{H}}} \cdot \int_{\Theta_{\mathrm{H}}^{-1}(0)}^{\Theta_{\mathrm{H}}^{-1}(1)} d\gamma_{\mathrm{H}} \\
&\quad \cdot \int_{\Theta_2^{-1}[1-H_2(D_2)]}^{\Theta_2^{-1}[H(X_2)]} \exp\left(-\frac{\gamma_2}{\overline{\gamma}_2} - \frac{\gamma_{\mathrm{H}}}{\overline{\gamma}_{\mathrm{H}}}\right)
\end{aligned}
$$

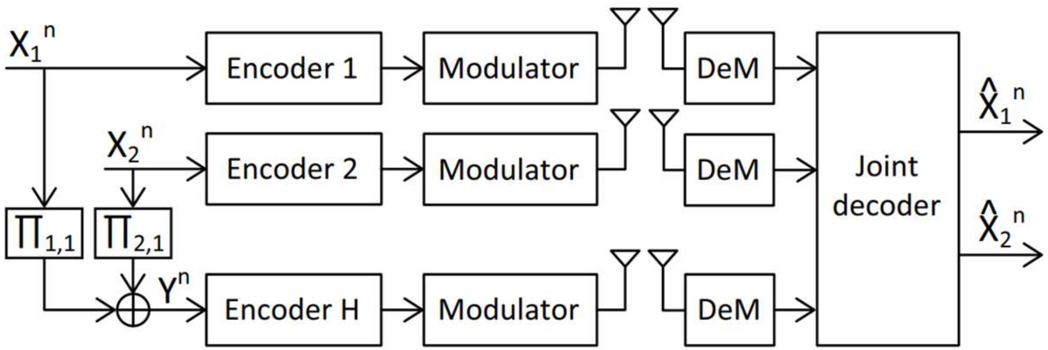
- Numerical Results ($\rho = 0.1$)

1. The larger the acceptable distortion , the smaller the outage probability .
   ($D_j \nearrow \Longrightarrow P_{\text{out}} \searrow$).

2. Without a helper, the curves always exhibit order diversity.

3. With a helper, it can achieve second order diversity (P_{out} decreases faster).





SIMULATION PARAMETERS

Parameter	Value
Frame length	10^4 bits
Number of frames	5×10^5
Source coding rate	1
Channel coding rate	$1/2$
Generator polynomial of CC	$G = ([3, 2]3)_8$
Type of interleaver	random interleaver
Modulation method	BPSK
Maximum iteration time	20

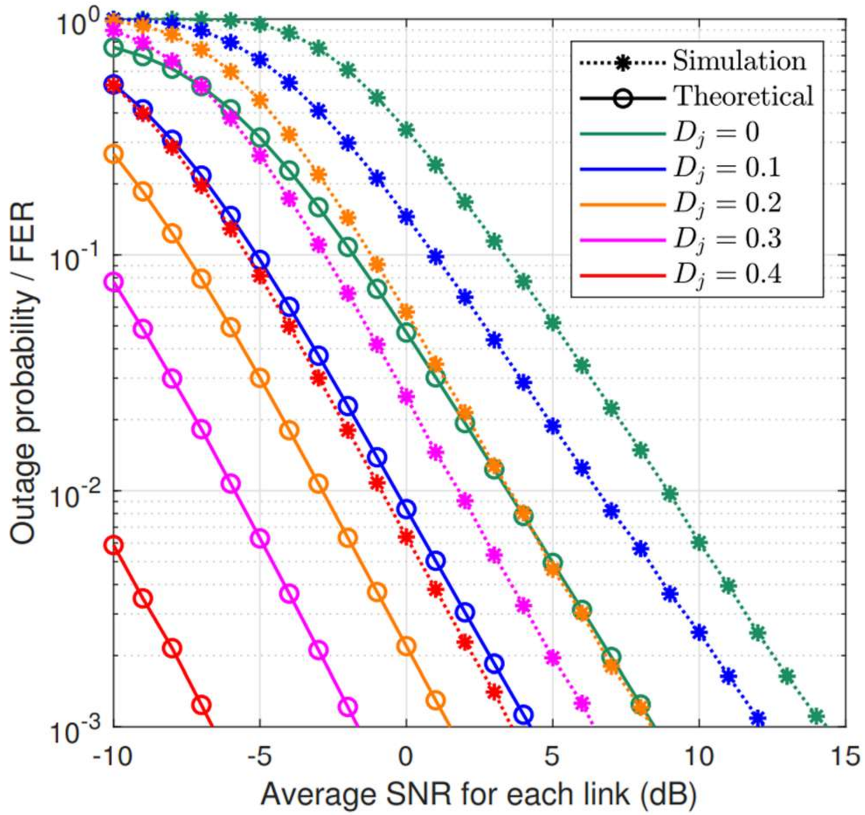


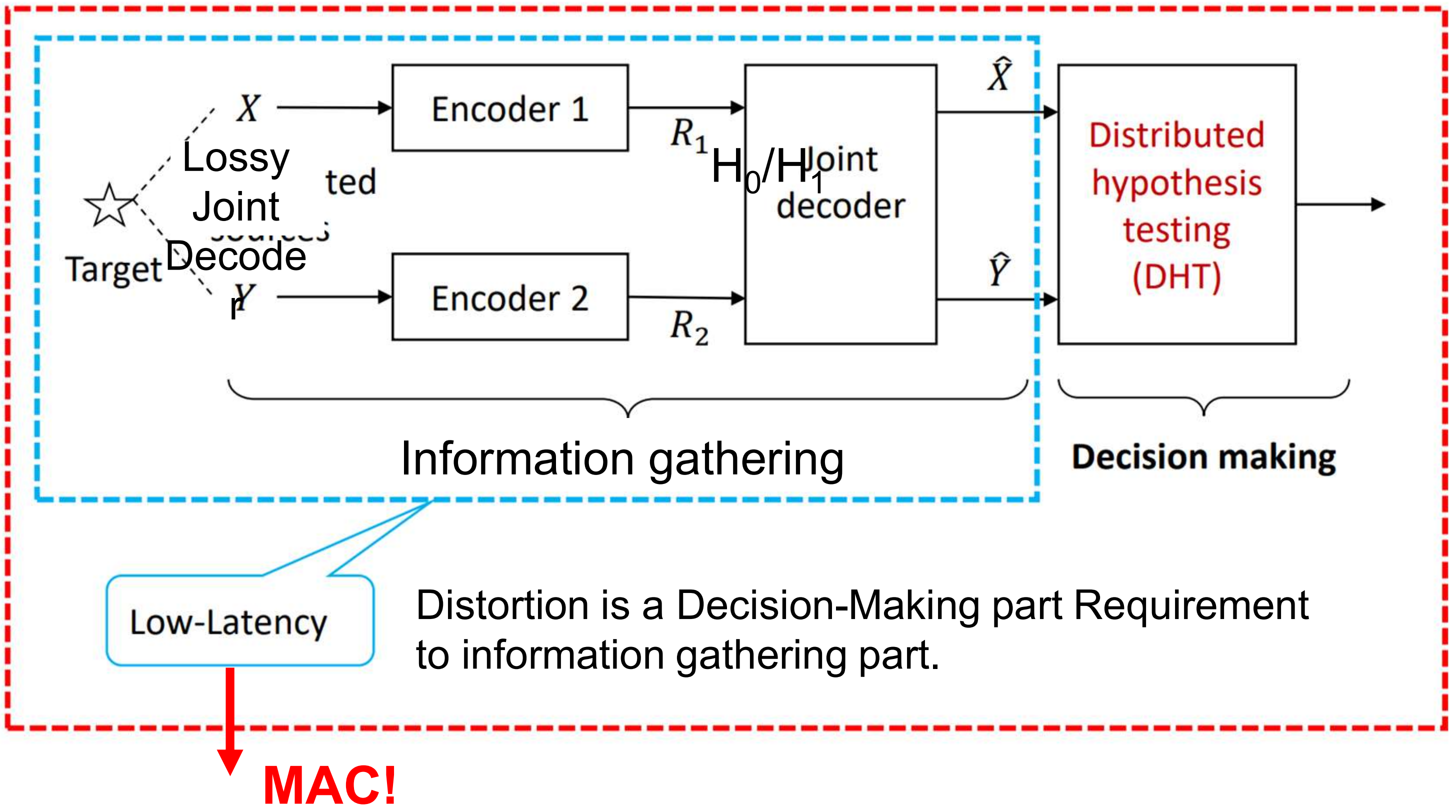
Fig. 13. Simulation results with $\rho = 0.1$.

2.3 End-to-End Lossless and Lossy Multiple Access Channels

An example of wireless sensor network.

59

Internet of things (IoT) system

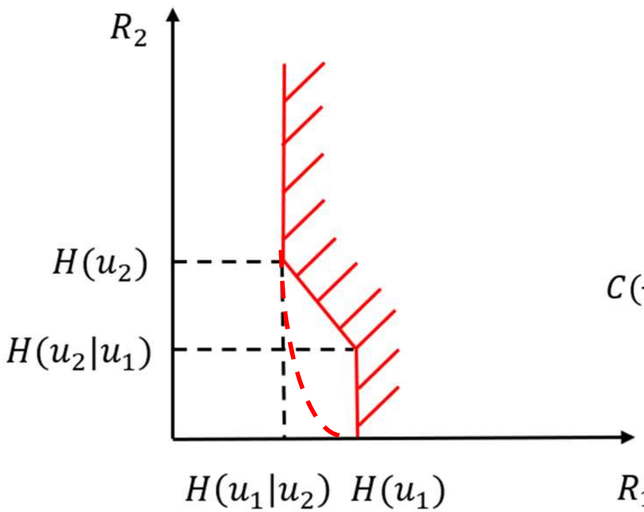


Data gathering should not necessarily be lossless.

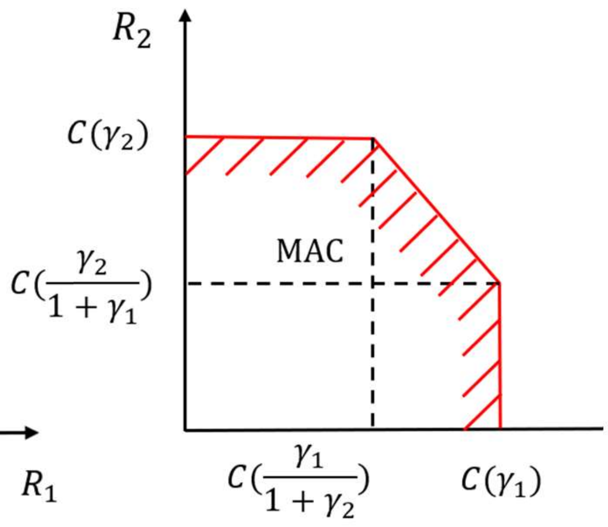
End-to-End Lossless MAC:

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom
CWC
Oulu

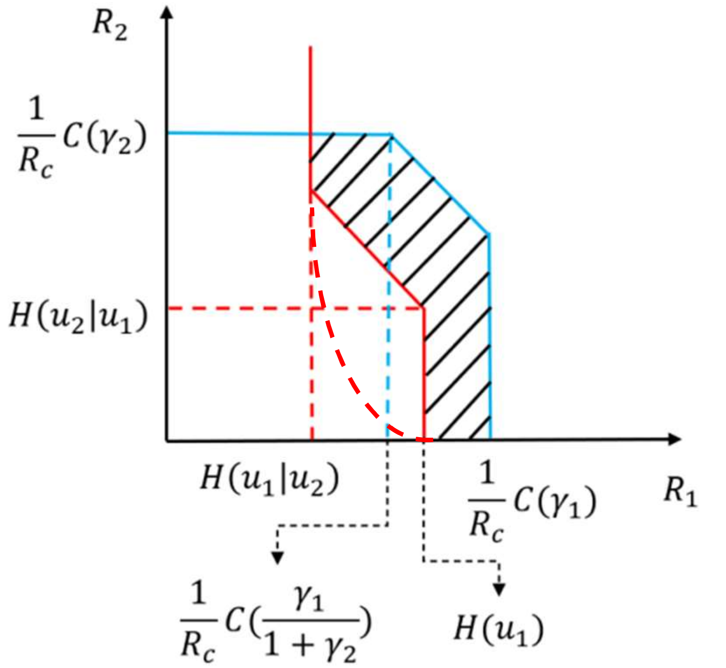
A sufficient condition of successful transmissions is defined as the case Slepian-Wolf region with a helper and MAC rate region intersect, where Source-Channel Separation holds.



Approximated Slepian-Wolf rate region with a helper (h-SW)



MAC rate region



Let the transmission rate of the source and the helper to be R_1 and R_2, respectively.

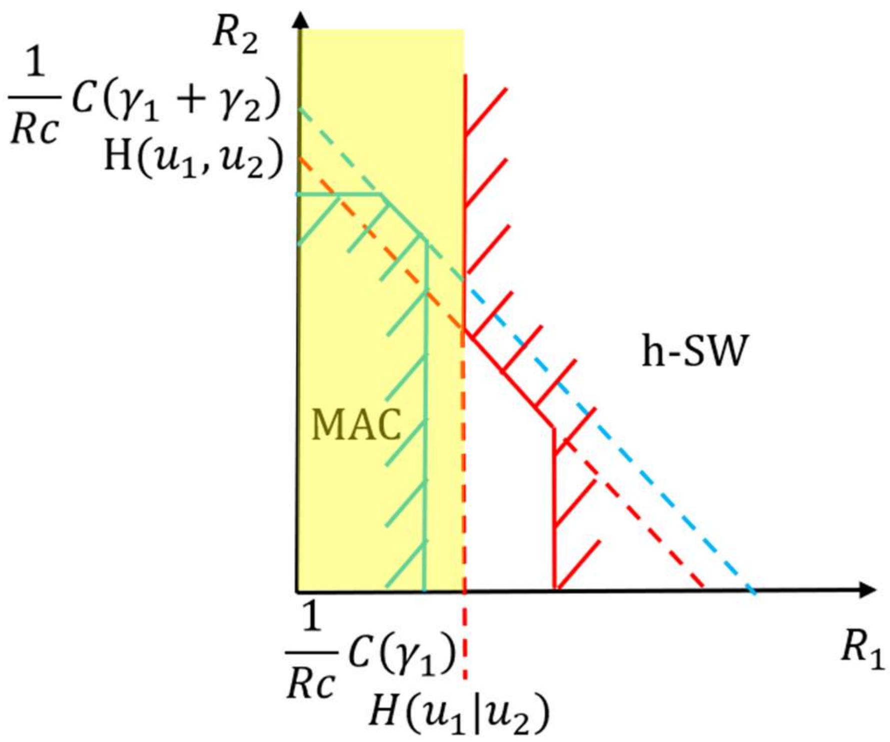
$$R_1 \geq \begin{cases} H(pe), & \text{for} \quad R_2 \geq 1, \\ 1 + H(pe) - R_2, & \text{for} \quad H(pe) \leq R_2 \leq 1, \\ 1, & \text{for} \quad 0 \leq R_2 \leq H(pe). \end{cases}$$

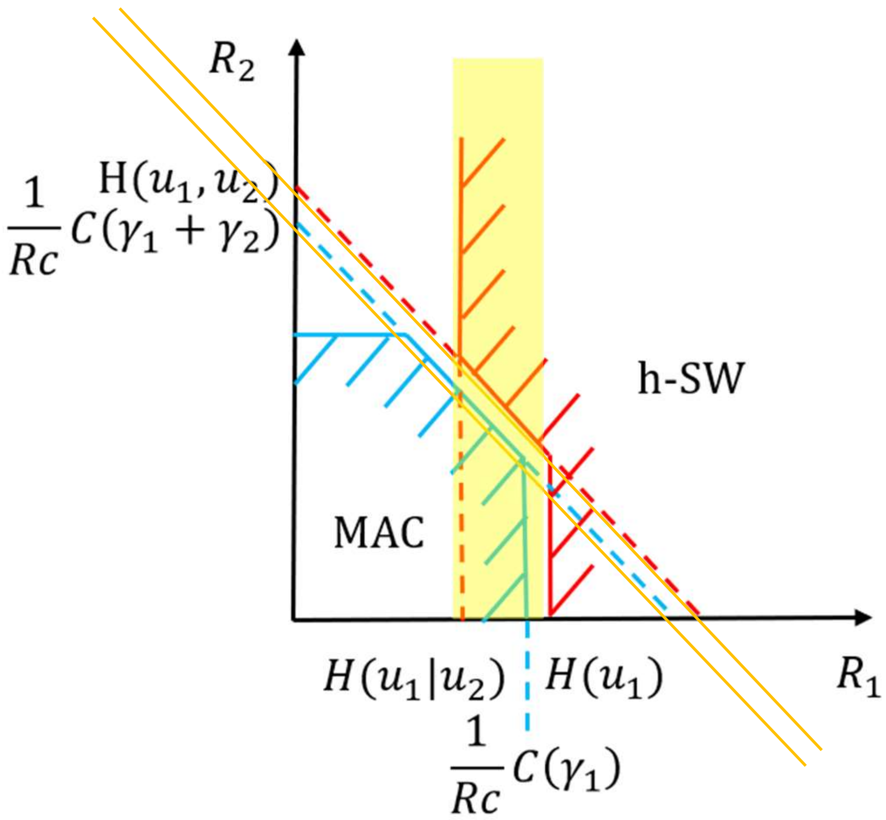
$$\begin{cases} R_1 R_c \leq C(\gamma_1), \\ R_2 R_c \leq C(\gamma_2), \\ R_1 R_c + R_2 R_c \leq C(\gamma_1 + \gamma_2), \end{cases}$$

Source-Channel Separation

Source – Channel separation modeled by Bit–Fipping Model (as in BF MIMO TEQ) with flipping probability p_e

Cases where Outage Happens





$$\begin{cases} 0 < \frac{1}{R_c} C\left(\gamma_1\right) < H\left(u_1 \mid u_2\right) \\ 0 < \frac{1}{R_c} C\left(\gamma_2\right) \end{cases}$$

$$\begin{cases} H\left(u_1 | u_2\right) \leq \frac{1}{Rc} C\left(\gamma_1\right) < H\left(u_1\right) \\ \frac{1}{Rc} C\left(\gamma_1 + \gamma_2\right) < H\left(u_1, u_2\right) \end{cases}$$

Outage Probability Expressions

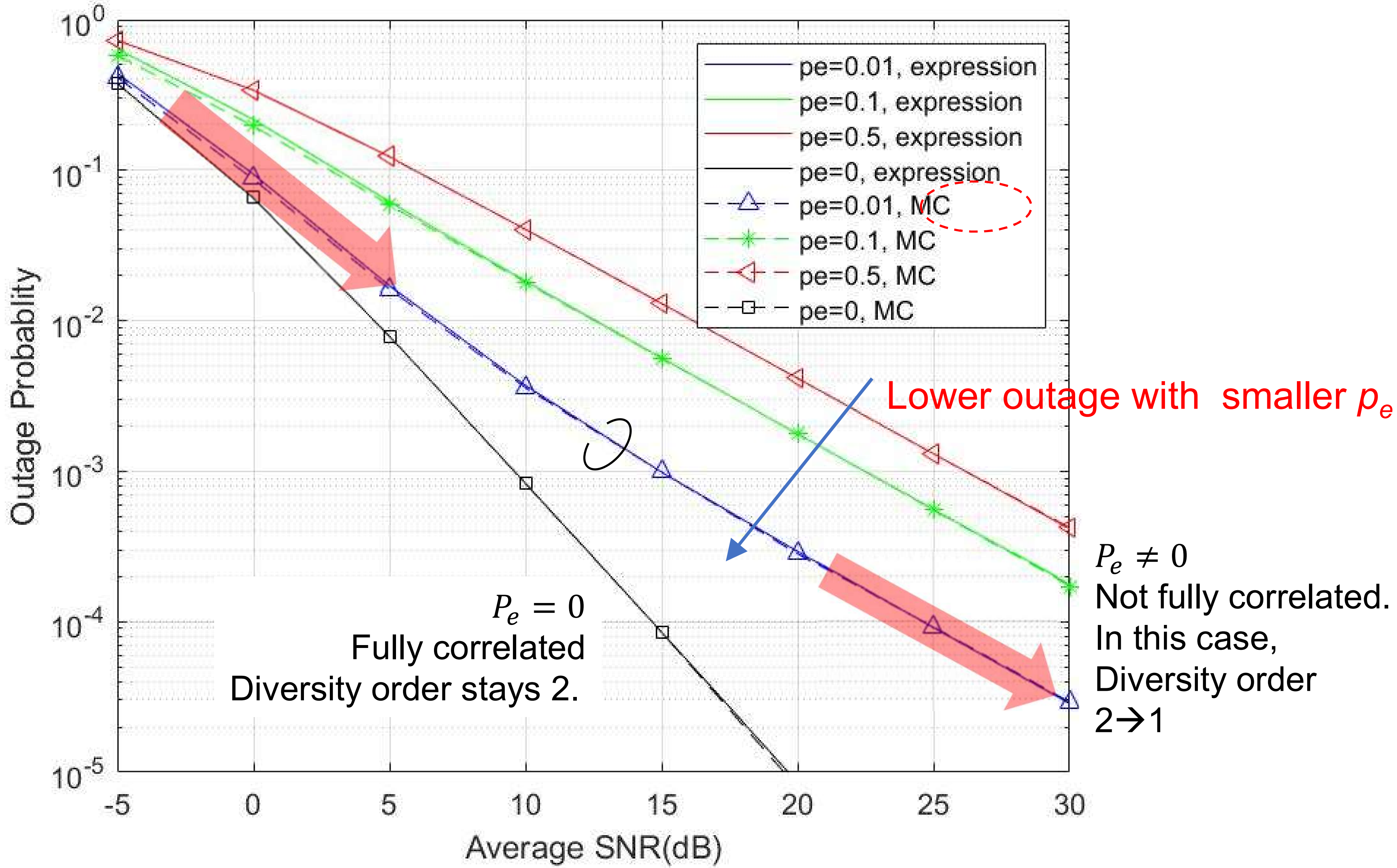
Outage probability can be calculated by multifold integrals with respect to the instantaneous SNR of each link.

$$
\begin{aligned}
P_{out,1} &= \Pr\left\{\frac{1}{R_c}C(\gamma_1) < H(u_1 \mid u_2)\right\} \\
&= \int_{\Phi(0)}^{\Phi[H(u_1|u_2)]}\int_{\Phi(0)}^{\Phi(+\infty)} p(\gamma_1)\,p(\gamma_2)\,d\gamma_2 d\gamma_1 \\
&= \int_{0}^{2^{R_c H(P_e)}-1}\int_{0}^{+\infty} p(\gamma_1)\,p(\gamma_2)\,d\gamma_2 d\gamma_1 \\
&= 1-\exp\left\{-\frac{1}{\Gamma_1}\left[2^{R_c H(P_e)}-1\right]\right\},
\end{aligned}
$$

$$
\begin{aligned}
P_{out,2} &= \Pr\left\{\begin{array}{c} H(u_1|u_2) \le \frac{1}{Rc}C(\gamma_1) < H(u_1); \\ \frac{1}{Rc}C(\gamma_1+\gamma_2) < H(u_1,u_2) \end{array}\right\} \\
&= \int_{\Phi[H(u_1|u_2)]}^{\Phi[H(u_1))]}\int_{\Phi(0)}^{\Phi[H(u_1,u_2)]-\gamma_1} p(\gamma_1)\,p(\gamma_2)\,d\gamma_2 d\gamma_1 \\
&= \int_{2^{R_c H(P_e)}-1}^{2^{R_c}-1}\int_{0}^{2^{R_c[1+H(P_e)]}-1-\gamma_1} p(\gamma_1)\,p(\gamma_2)\,d\gamma_2 d\gamma_1 \\
&= \exp\left\{-\frac{1}{\Gamma_1}\left[2^{R_c H(P_e)}-1\right]\right\} - \exp\left[-\frac{1}{\Gamma_1}\left(2^{R_c}-1\right)\right] \\
&\quad - \frac{\Gamma_2}{\Gamma_1-\Gamma_2}\exp\left[-\frac{1}{\Gamma_1}\left(2^{R_c}-1\right)\right] \\
&\quad \cdot \exp\left\{-\frac{1}{\Gamma_2}\left[2^{R_c H(P_e)+R_c}-2^{R_c}\right]\right\} \\
&\quad + \frac{\Gamma_2}{\Gamma_1-\Gamma_2}\exp\left\{-\frac{1}{\Gamma_1}\left[2^{R_c H(P_e)}-1\right]\right\} \\
&\quad \cdot \exp\left\{-\frac{1}{\Gamma_2}\left[2^{R_c H(P_e)+R_c}-2^{R_c H(P_e)}\right]\right\},
\end{aligned}
$$

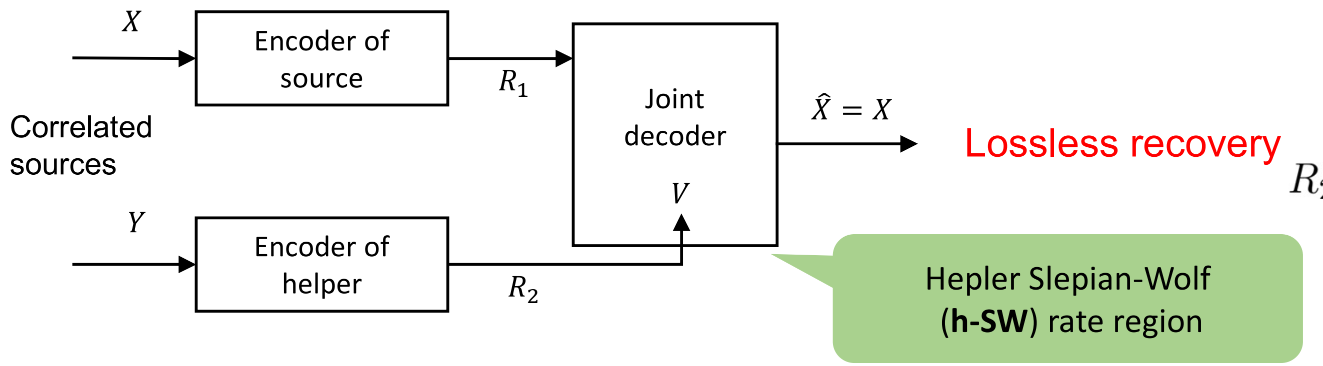
$$
P_{out} = P_{out,1} + P_{out,2}
$$

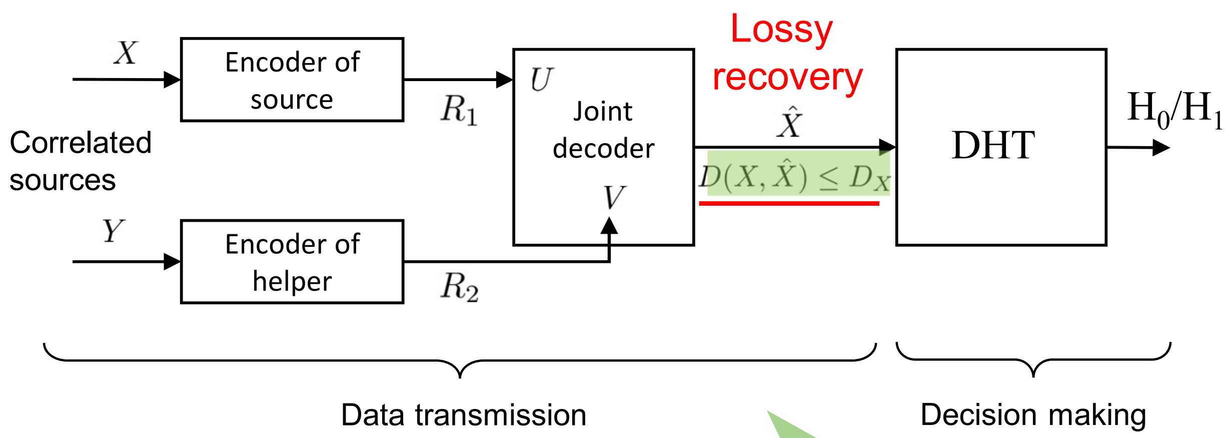
Numerical Result

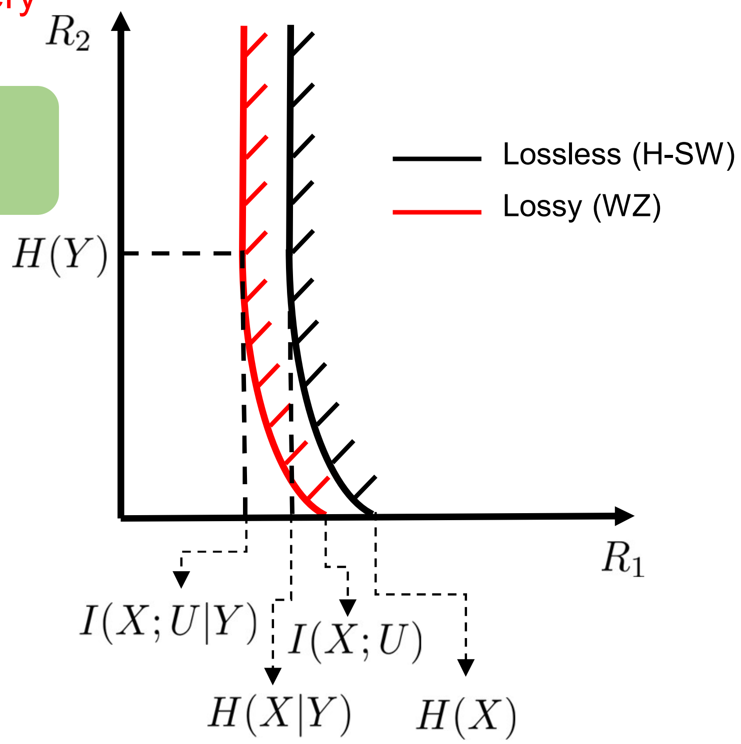


Lossy Case Helper: WZ System

End-to-End Lossy MAC:







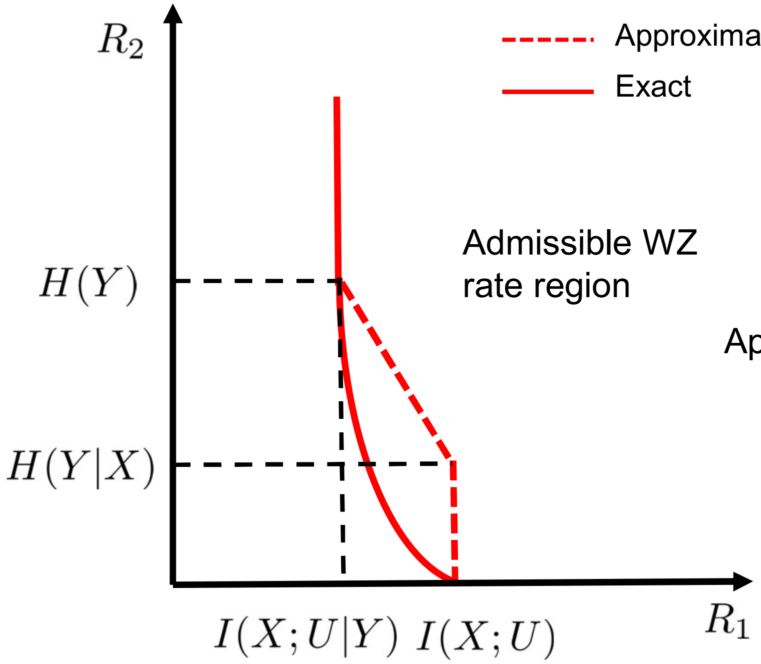
D_X =Maximum Acceptable Distortion

Wyner-Ziv (**WZ**) rate region

In lossy case, admissible rate region is larger.

WZ Region Approximation

Intersection Analysis



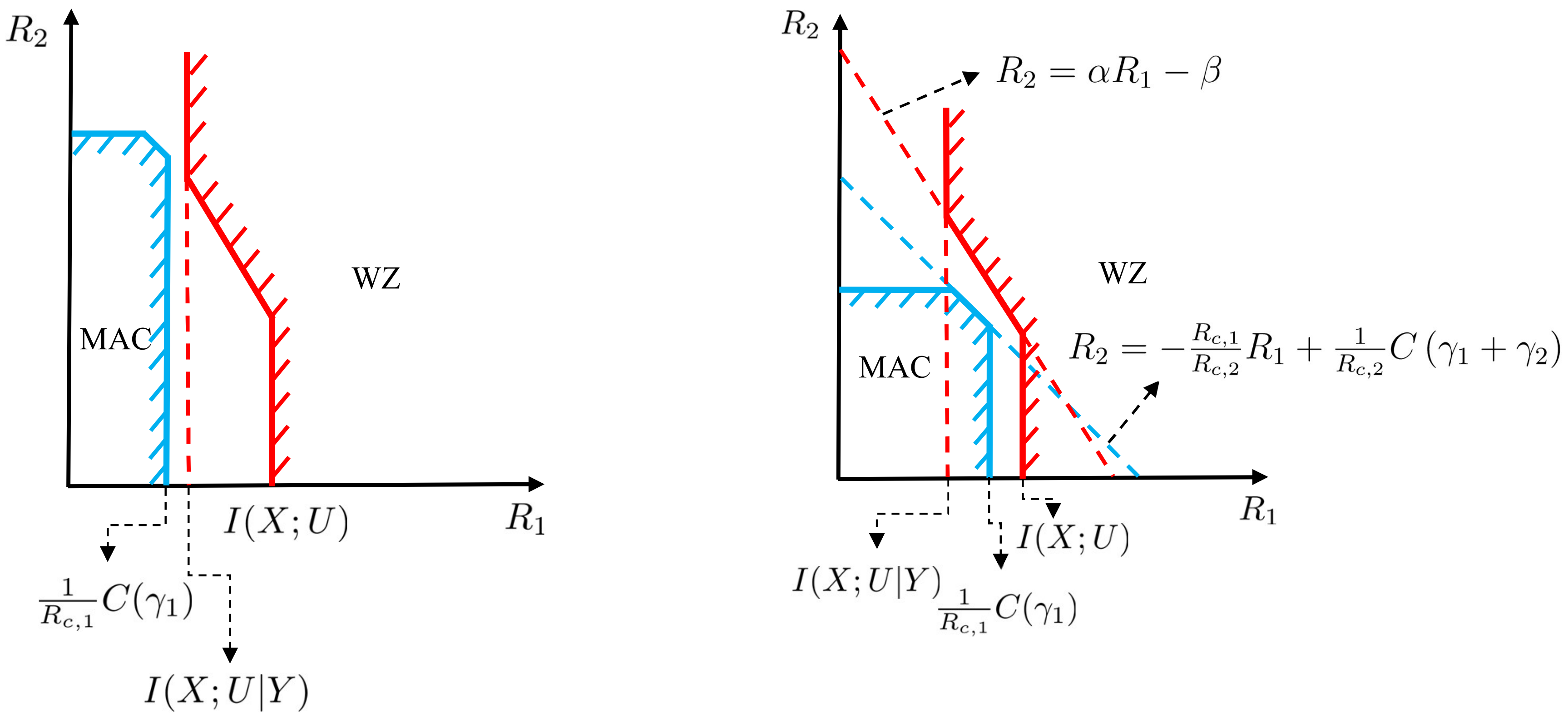
Admissible WZ rate region

Exact: $$\begin{cases} R_1 \geq I(X;U|Y), \\ R_2 \geq I(Y|V), \end{cases}$$

Approximation: $$R_1 \geq \begin{cases} I(X;U|Y), & \text{for} \quad R_2 \geq H(Y), \\ \frac{1}{\alpha}R_2 + \frac{\beta}{\alpha}, & \text{for} \quad H(Y|X) \leq R_2 \leq H(Y), \\ I(X;U), & \text{for} \quad 0 \leq R_2 \leq H(Y|X), \end{cases}$$

with: $$\alpha = \frac{I(X;Y)}{I(X;U|Y) - I(X;U)} \qquad \beta = \frac{I(X;U) - I(X;U|Y) \cdot H(Y|X)}{I(X;U|Y) - I(X;U)}$$

Cases where Outage Happens



Two scenarios that outage happens. $P_{out,total} = P_{out,1} + P_{out,2}$

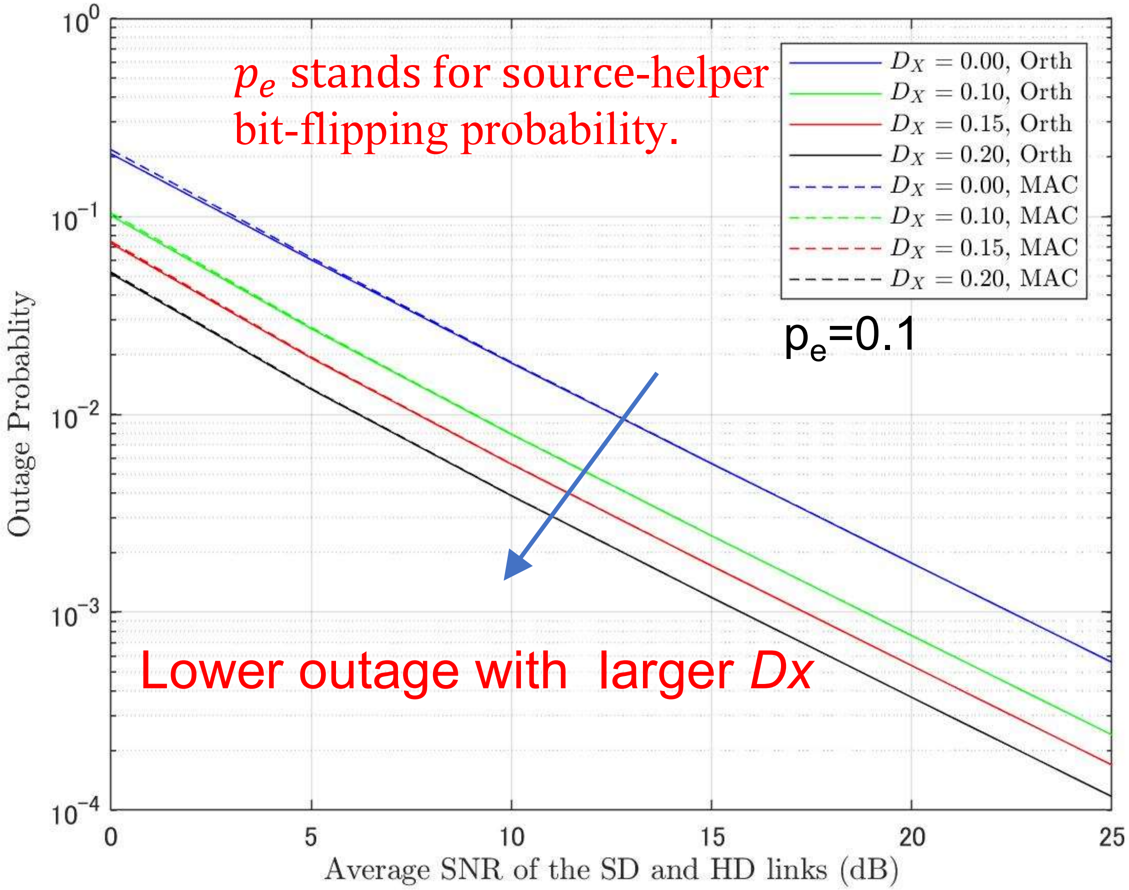
Twofold Integral Similar to Lossless Case

$$
\begin{aligned}
&P_{out,MAC,1}\\
&=\Pr\left\{0 \le \frac{1}{R_{c,1}} C\left(\gamma_1\right) \le I\left(X;U|Y\right);\, \frac{1}{R_{c,2}} C\left(\gamma_2\right) \ge 0\right\}\\
&=\Pr\left\{\Phi\left(0\right) \le \gamma_1 \le \Phi\left[I\left(X;U|Y\right)\right];\, \gamma_2 \ge \Phi\left(0\right)\right\}\\
&=\int_{\Phi(0)}^{\Phi[I(X;U|Y)]} \int_{\Phi(0)}^{+\infty} p\left(\gamma_1\right) p\left(\gamma_2\right) d\gamma_2 d\gamma_1\\
&=\int_{0}^{2^{R_{c,1}[H_b(D_X * P_e) - H_b(D_X)]}-1} \int_{0}^{+\infty} p\left(\gamma_1\right) p\left(\gamma_2\right) d\gamma_2 d\gamma_1
\end{aligned}
$$

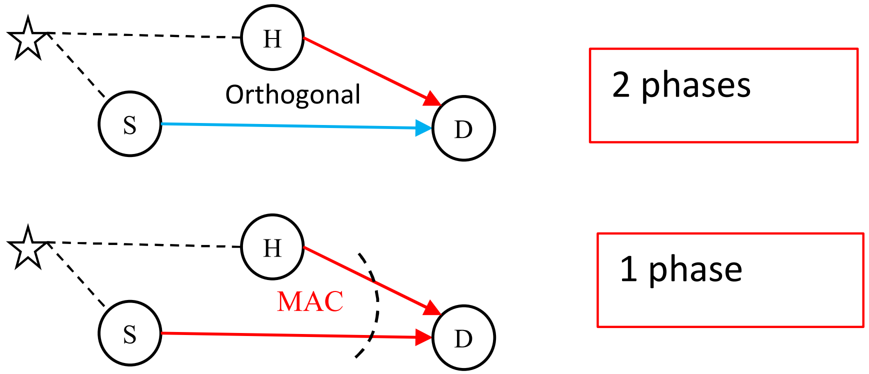
$$
\begin{aligned}
&P_{out,MAC,2}\\
&=\Pr\left\{\begin{array}{c} I\left(X;U|Y\right) \le \frac{1}{R_{c,1}} C\left(\gamma_1\right) \le I\left(X;U\right);\\ 0 \le -\frac{R_{c,1}}{R_{c,2}} R_1 + \frac{1}{R_{c,2}} C\left(\gamma_1 + \gamma_2\right) \le \alpha R_1 - \beta \end{array}\right\}\\
&=\Pr\left\{\begin{array}{c} \Phi\left[I\left(X;U|Y\right)\right] \le \gamma_1 \le \Phi\left[I\left(X;U\right)\right];\\ \Phi\left(0\right) \le \gamma_2 \le M \end{array}\right\}\\
&=\int_{\Phi[I(X;U|Y)]}^{\Phi[I(X;U)]} \int_{\Phi(0)}^{M} p\left(\gamma_1\right) p\left(\gamma_2\right) d\gamma_2 d\gamma_1\\
&=\int_{2^{R_{c,1}[H_b(D_X * P_e) - H_b(D_X)]}-1}^{2^{R_{c,1}[1 - H_b(D_X)]}-1} \int_{0}^{M} p\left(\gamma_1\right) p\left(\gamma_2\right) d\gamma_2 d\gamma_1,
\end{aligned}
$$

with $M = 2^{-\beta R_{c,2}} \left(1+\gamma_1\right)^{\left(\frac{\alpha R_{c,2}}{R_{c,1}}+1\right)} - 1 - \gamma_1.$

WZ MAC Outage Probability



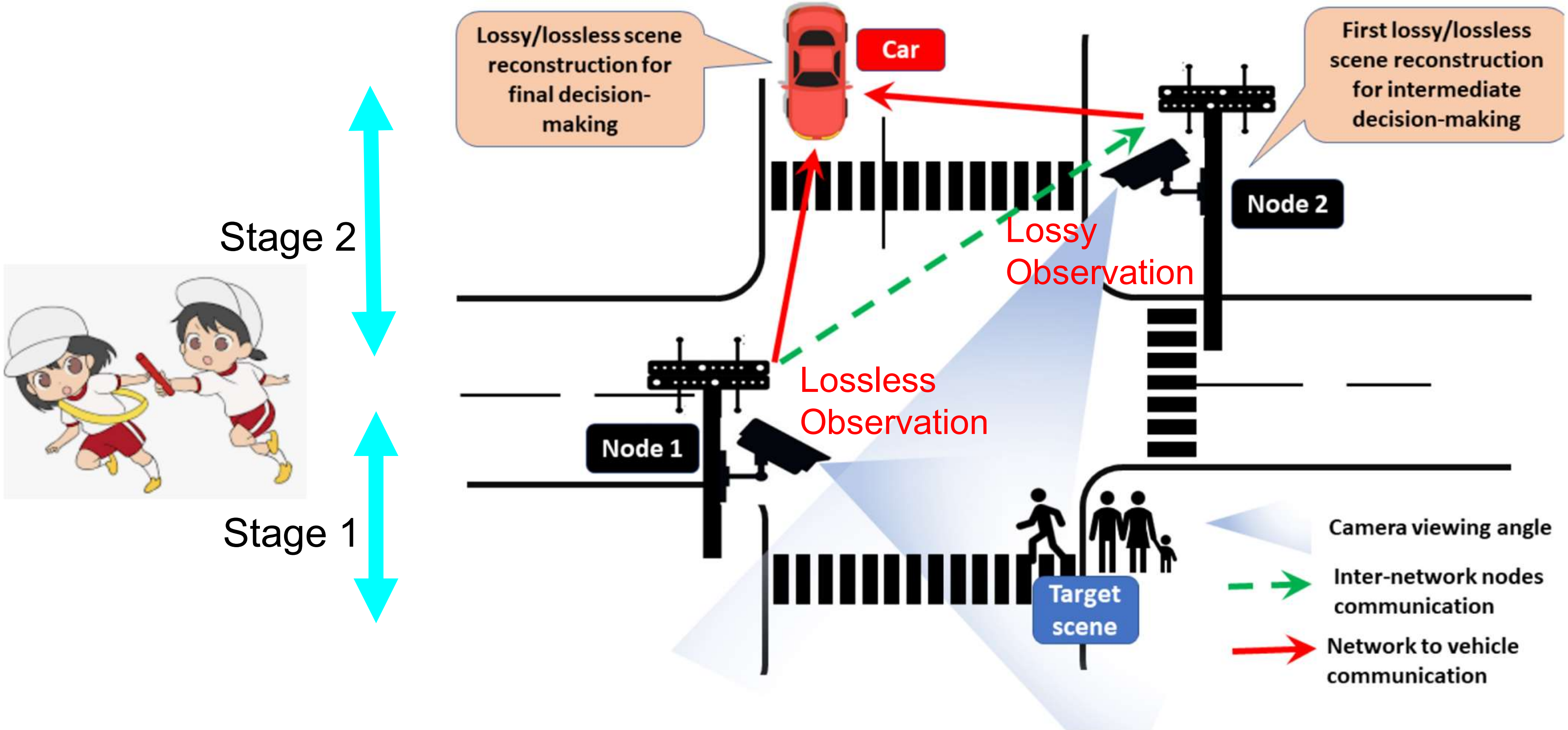
- The results show that the difference between Orthogonal and MAC is negligibly small.
- Transmission efficiency with MAC transmission is twice as large as that in orthogonal transmission.



68

2.4 Two Stage Wyner-Ziv Network: Distortion Transfer Analysis

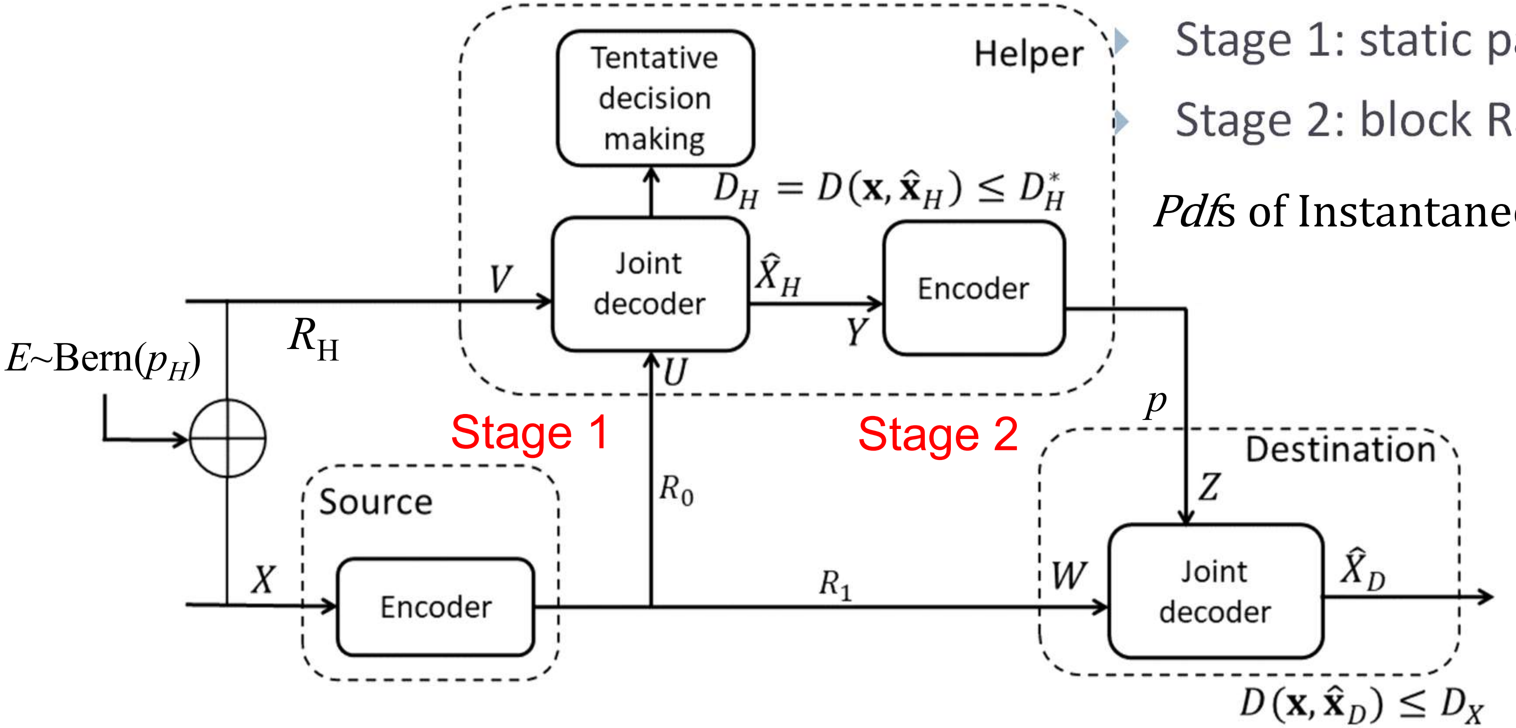
Analyze: How Distortion causing at the previous stage is forwarded to the current stage?
Distortion Transfer Function (DTF)



Block Diagram

- Mathematical assumptions
  - i.i.d Bit-Flipping model for the correlation between the observations
  - Hamming distortion measure
  - Definition of Admissible rate-distortion (RD) region:

$$\mathcal{R}(D_X) = \{(R_0, R_1, R_2): (R_0, R_1, R_2) \text{ is admissible if } \lim_{n \to +\infty} \mathbf{E}(D(\mathbf{x}, \hat{\mathbf{x}}_D) \leq D_X + \epsilon, \forall \epsilon > 0)\}$$



- Stage 1: static parameters
- Stage 2: block Rayleigh fading

Pdfs of Instantaneous SNRs with the links are:

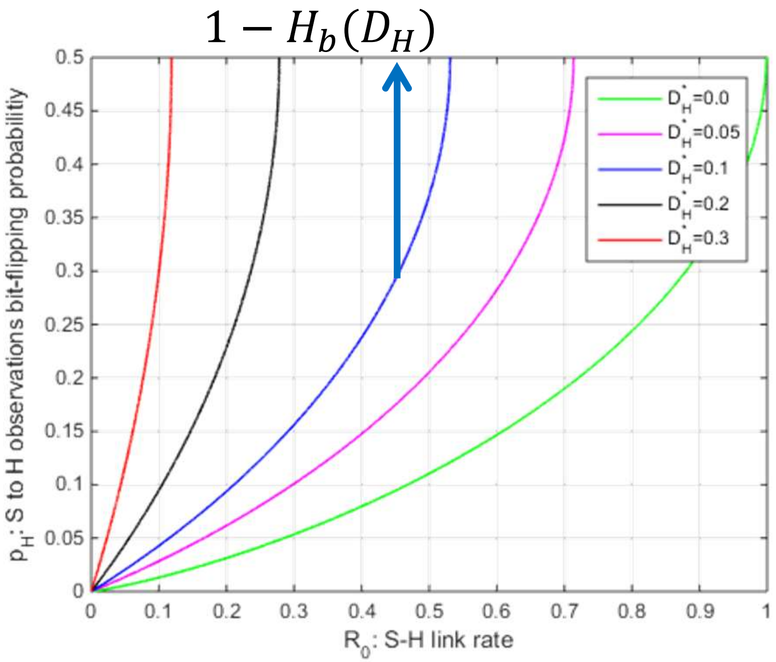
$$p(\gamma_i) = \frac{1}{\Gamma_i} \exp\left(-\frac{\gamma_i}{\Gamma_i}\right), \quad i \in \{SD, HD\}.$$

Stage-Independent RD Analysis

- Stage-by-stage admissible RD region

Stage 1

$$R_0 \geq I(X;U|V) = H_b(D_H * p_H) - H_b(D_H)$$

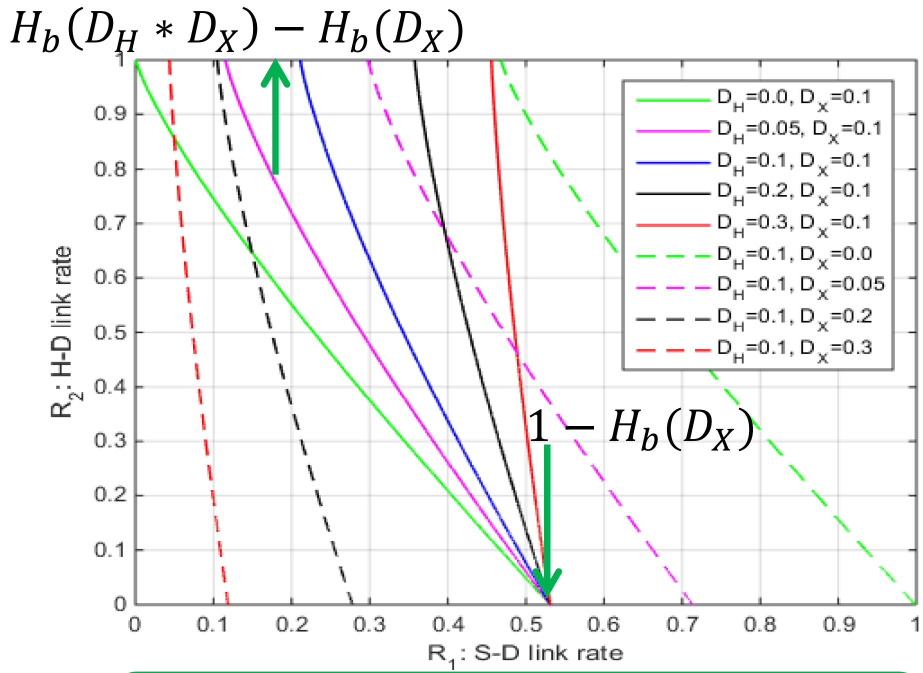


Smaller rate R_0 is required when D_H is larger or p_h is lower

Stage 2

$$R_1 \geq I(X;W|Z) = H_b(p * D_H * D_X) - H_b(D_X)$$

$$R_2 \geq I(Y;Z) = 1 - H_b(p)$$



Smaller rates R_1 and R_2 are required when D_X is larger or D_H is smaller

Stage-Dependent RD Analysis

- The Recursive Structure of the Binary Convolution is Referred to as: Distortion Transfer Function (DTF)

$R(D) = H_b(p) - H_b(D)$

$p = H_b^{-1}(R + H(D))$

DTF Connects the two stages, as

- Stage 1:
  - Assume the required distortion D_H at Helper is given by D_H^*
  - It is found that Bit-Flipping probability between the observations should satisfy

$$p_h \leq \Lambda\left[D_H^*, H_b^{-1}\left(R_0 + H_b(D_H^*)\right)\right]$$

to distortion requirement on Side information

DTF

From distortion requirement at Helper

  - When $R_0 \geq 1 - H_b(D_H^*)$, using $\Lambda(y, t = 0.5) = 0.5$, we have $p_H \leq 0.5$
    - which corresponds to the case no side information is required.
  - When R_0 decreases, $H_b^{-1}(R_0 + H_b(D_H^*))$ also decreases, and hence also p_H decreases,
    - which corresponds to the case higher correlation is needed to satisfy D_H^*

Stage-Dependent RD Analysis: Connecting Stages

- Stage 2
  - Assume the required distortion at Destination is given by D_X
  - It is found that the distortion at Helper should satisfy

$$D_H^* \leq \Lambda\left[p\,,\Lambda\left[D_X, H_b^{-1}\left(R_1 + H_b(D_X)\right)\right]\right]$$

DTF

to distortion requirement at Helper

From distortion requirement at Destination

  - When R_2 is large enough, $p = H_b^{-1}(1 - R_2) = 0$, then using $\Lambda(0, t) = t$:

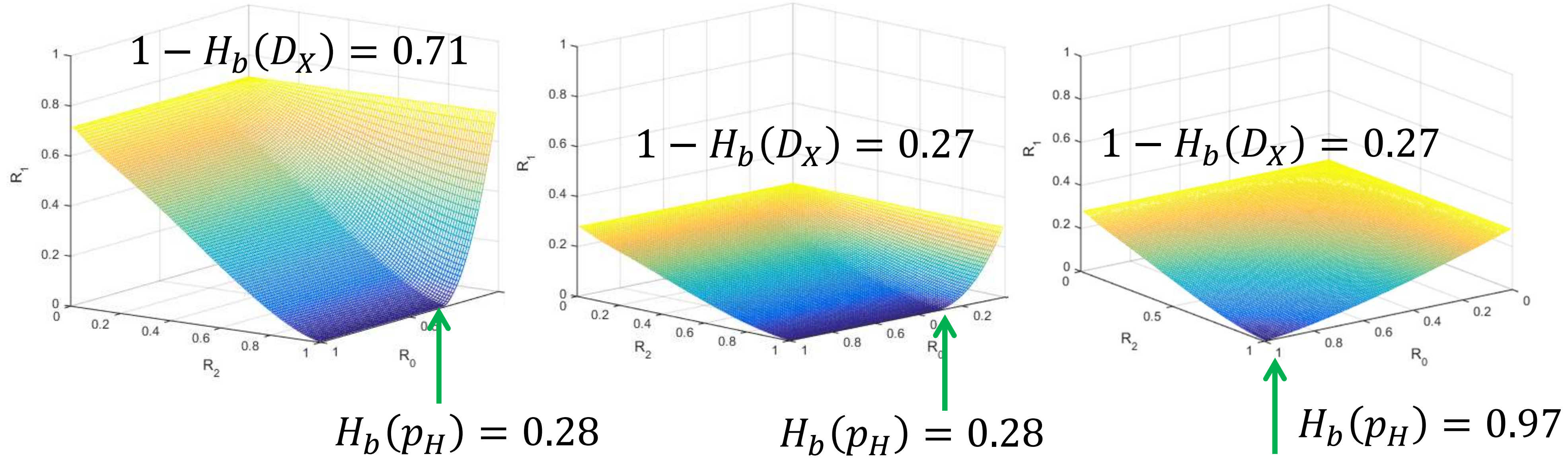
$$D_H^* \leq \Lambda\left[D_X, H_b^{-1}\left(R_1 + H_b(D_X)\right)\right] = D_{SI}$$

  - In this case, Stage 2 is equivalent to Stage 1 with distortion requirement D_{SI} on Side Information.
  - Condition $p \leq D_{SI}$ is required to achieve D_X at Destination

Rate Surface Calculation

- Connecting Stage 1 and Stage 2
  - 3D admissible RD region

$\boldsymbol{D_X = 0.05, p_H = 0.05}$ $\boldsymbol{D_X = 0.2, p_H = 0.05}$ $\boldsymbol{D_X = 0.2, p_H = 0.4}$



Relaxing the requirement with larger D_X expands the admissible RD region

Higher observation correlation reduces distortion requirement at Helper as well as Source-Helper link rate R_0, resulting in expands the admissible RD region

Outage Probability

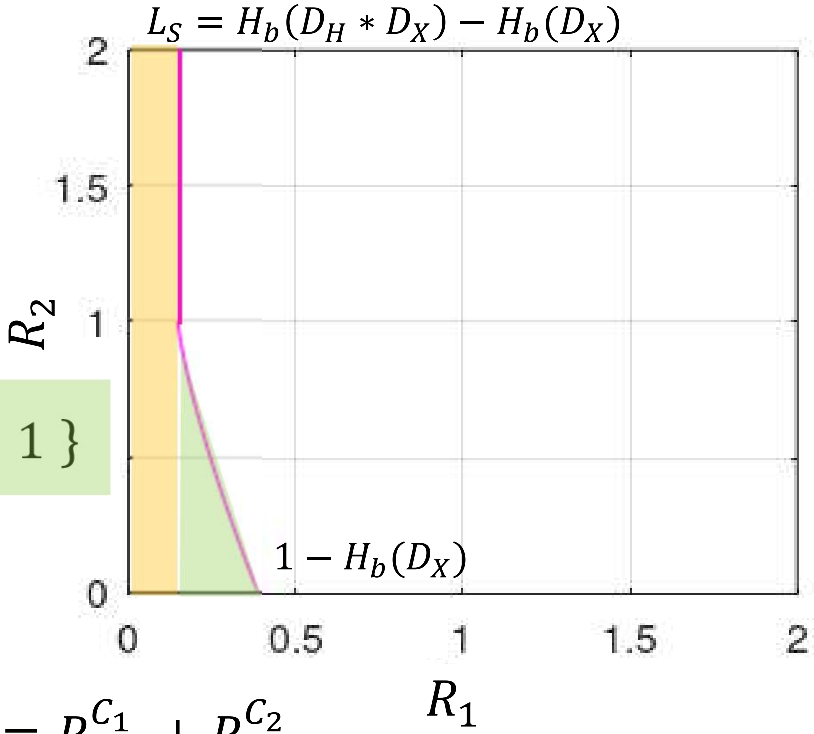
- Assumptions:
  - Static stage 1: D_H is a fixed parameter
  - Destination is moving: fading variation on S-D and H-D links
  - → When rates R_1 and R_2 are in the inadmissible RD region: outage happens

- → Case 1:

$$C_1 \triangleq \{0 \leq R_1 \leq \underbrace{H_b(D_H * D_X) - H_b(D_X)}_{L_S}, R_2 \geq 0\}$$

- → Case 2:

$$C_2 \triangleq \{L_S \leq R_1 \leq \underbrace{H_b(p * D_H * D_X) - H_b(D_X)}_{f(R_2)}, 0 \leq R_2 \leq 1\}$$



$$P_{out} = \Pr\{(R_1, R_2) \in C_1\} + \Pr\{(R_1, R_2) \in C_2\} = P_{out}^{C_1} + P_{out}^{C_2}$$

Twofold Integrals for Outage Probability Calculation

- Utilizing Lossy Source-Channel Separation theorem,

$$R_1 \leq \Phi_S(\gamma_S) \triangleq \frac{C(\gamma_S)}{R_c^{SD}} \qquad R_2 \leq \Phi_S(\gamma_H) \triangleq \frac{C(\gamma_H)}{R_c^{HD}}$$

- $C(\gamma) = \log_2(1+\gamma)$: the channel capacity function with two dimensional signaling

- Case 1 outage probability:

$$\begin{aligned} P_{out}^{C_1} &= \Pr\{0 \leq R_1 \leq L_S, 0 \leq R_2\} \\ &= \Pr\{0 \leq \Phi_S(\gamma_S) \leq L_S, 0 \leq \Phi_H(\gamma_H)\} \\ &= \Pr\{\Phi_S^{-1}(0) \leq \gamma_S \leq \Phi_S^{-1}(L_S), \Phi_H^{-1}(0) \leq \gamma_H\} \\ &= \int_{\Phi_H^{-1}(0)}^{+\infty} \int_{\Phi_S^{-1}(0)}^{\Phi_S^{-1}(L_S)} p(\gamma_S, \gamma_H)\, d\gamma_S\, d\gamma_H \end{aligned}$$

- Case 2 outage probability:

$$\begin{aligned} P_{out}^{C_2} &= \Pr\{L_S \leq R_1 \leq H_S(\gamma_H), 0 \leq R_2 \leq 1\} \\ &= \Pr\{L_S \leq \Phi_S(\gamma_S) \leq H_S(\gamma_H), 0 \leq \Phi_H(\gamma_H) \leq 1\} \\ &= \Pr\{\Phi_S^{-1}(L_S) \leq \gamma_S \leq \Phi_S^{-1}(H_S(\gamma_H)), \Phi_H^{-1}(0) \leq \gamma_H \leq \Phi_H^{-1}(1)\} \\ &= \int_{\Phi_H^{-1}(0)}^{\Phi_H^{-1}(1)} \int_{\Phi_S^{-1}(L_S)}^{\Phi_S^{-1}(H_S(\gamma_H))} p(\gamma_S, \gamma_H)\, d\gamma_S\, d\gamma_H \end{aligned}$$

$$p = H_b^{-1}(1 - \Phi_H(\gamma_H))$$

$$H_S(\gamma_H) = H_b(H_b^{-1}(1 - \Phi_H(\gamma_H)) * D_H * D_X) - H_b(D_X)$$

Twofold Integrals for Outage Probability Calculation

- For **independent fading** on S-D and H-D links

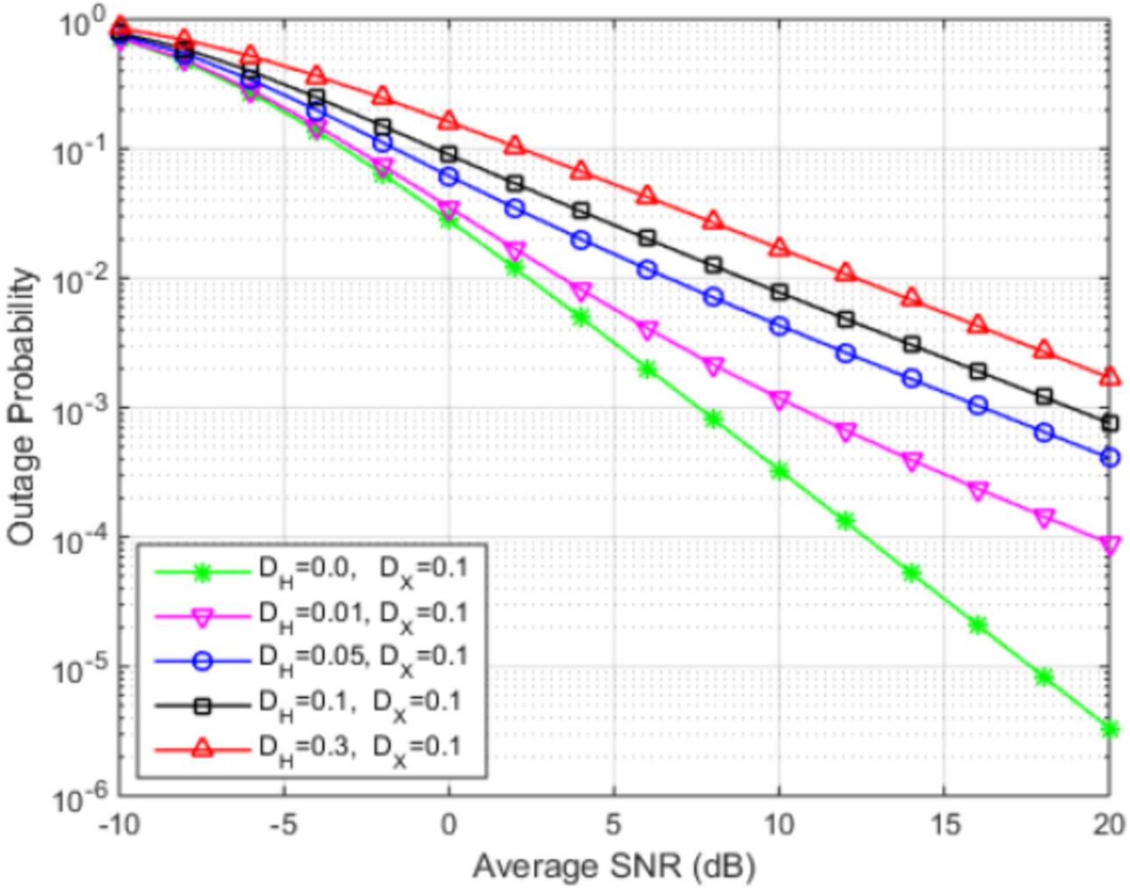
$$
\begin{aligned}
P_{out}^{C_1} &= \frac{1}{\Gamma_S \Gamma_H} \int_{\Phi_H^{-1}(0)}^{+\infty} \int_{\Phi_S^{-1}(0)}^{\Phi_S^{-1}(L_S)} \exp(-\frac{\gamma_S}{\Gamma_S}) \ \exp(-\frac{\gamma_H}{\Gamma_H}) \ d\gamma_S \ d\gamma_H \\
&= 1 - \exp\left(\frac{-\Phi_S^{-1}(L_S)}{\Gamma_S}\right),
\end{aligned}
$$

TwWKDCpCmPaT!

$$
\begin{aligned}
P_{out}^{C_2} &= \frac{1}{\Gamma_S \Gamma_H} \int_{\Phi_H^{-1}(0)}^{\Phi_H^{-1}(1)} \int_{\Phi_S^{-1}(L_S)}^{\Phi_S^{-1}(H_S(\gamma_H))} \exp(-\frac{\gamma_S}{\Gamma_S}) \ \exp(-\frac{\gamma_H}{\Gamma_H}) \ d\gamma_S \ d\gamma_H \\
&= \frac{1}{\Gamma_H} \int_{\Phi_H^{-1}(0)}^{\Phi_H^{-1}(1)} \exp(-\frac{\gamma_H}{\Gamma_H}) \left[\exp\left(-\frac{\Phi_S^{-1}(L_S)}{\Gamma_S}\right) - \exp\left(-\frac{\Phi_S^{-1}(H_S(\gamma_H))}{\Gamma_S}\right)\right] d\gamma_H
\end{aligned}
$$

Outage Probability

- For **independent fading** on the S-D and H-D links



First order diversity when **first stage is lossy** and high D_H at Helper

$$P_{out} \approx P_{out}^{C_1} = 1 - \exp\left(\frac{-\Phi_S^{-1}(L_S)}{\Gamma_S}\right)$$
$$\approx \frac{-\Phi_S^{-1}(L_S)}{\Gamma_S} \propto \frac{1}{\Gamma_S}$$

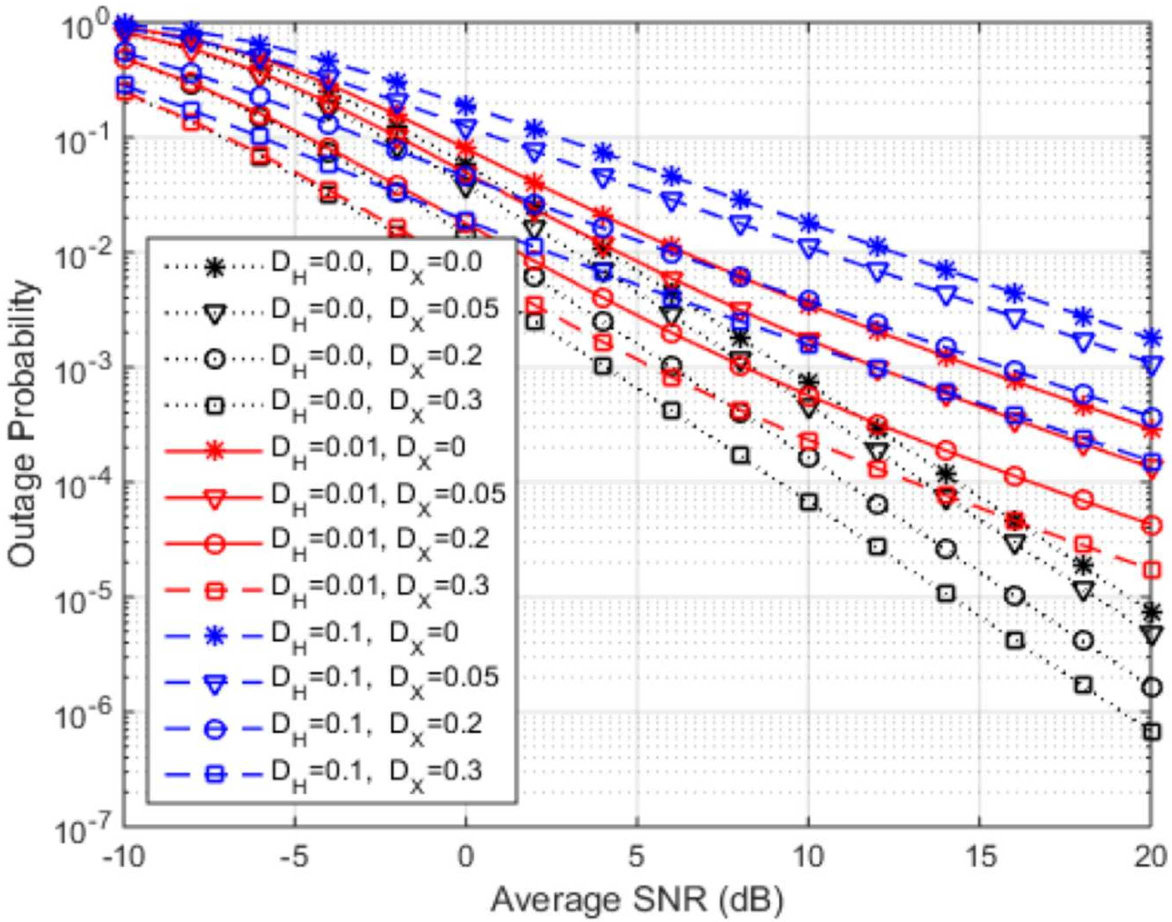
Second-to-first order diversity change when first stage is lossy and D_H is low

Second order diversity achieved when Stage 1 is lossless ($D_H = 0, L_S = 0$)

$$P_{out} = P_{out}^{C_2}$$
$$\approx \frac{1}{\Gamma_H \Gamma_S} \int_{\Phi_H^{-1}(0)}^{\Phi_H^{-1}(1)} \Phi_S^{-1}\left(H_S(\gamma_H)\right) d\gamma_H \propto \frac{1}{\Gamma_H \Gamma_S}$$

Outage Probability

- For **independent fading** on the S-D and H-D links



Increasing the allowed distortion at Destination D_X provides **lower outage probabilities**

However, D_X has no impact on the slope of the outage probability (parallel curves)

Impact of Spatial Correlation on Outage Probability

- For **correlated fading** on the S-D and H-D links
  - $\rho = < h_1, h_2^* >$ the correlation of the complex channel gains h_1 and h_2
  - The joint PDF of the instantaneous SNRs

$$p(\gamma_S, \gamma_H) = \frac{1}{\Gamma_S \Gamma_H (1 - |\rho|^2)} \exp\left(-\frac{1}{1-|\rho|^2} \left(\frac{\gamma_S}{\Gamma_S} + \frac{\gamma_H}{\Gamma_H}\right) \right) \times I_0 \left(\frac{2|\rho|}{1-|\rho|^2} \sqrt{\frac{\gamma_S \gamma_H}{\Gamma_S \Gamma_H}} \right)$$

  - $I_0(x)$ is the zero-order modified Bessel's function of the first kind

$$I_0(x) = \sum_{m=0}^{+\infty} \frac{1}{(m!)^2} \left(\frac{x}{2}\right)^{2m}$$

  - The outage probability of cases 1 and 2 can be written as

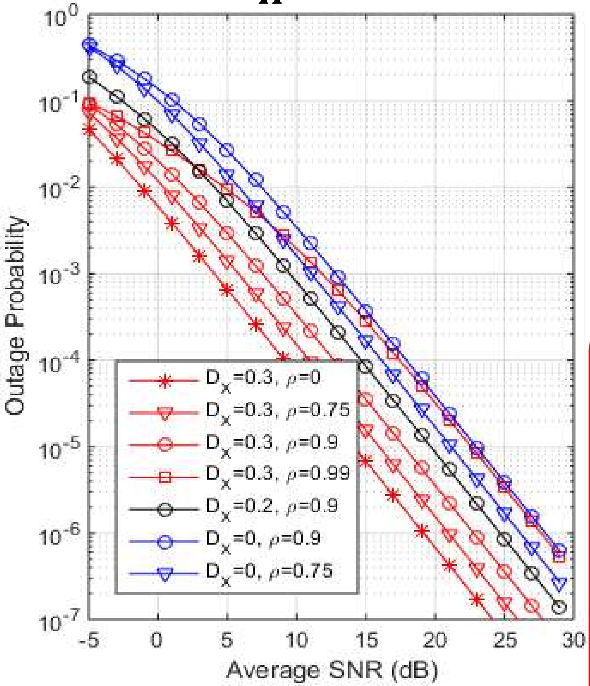
$$P_{out}^{C_1} \approx \sum_{m=0}^{M} A_m \int_{\Phi_H^{-1}(0)}^{+\infty} \int_{\Phi_S^{-1}(0)}^{\Phi_S^{-1}(L_S)} \left(\frac{\gamma_S \gamma_H}{\Gamma_S \Gamma_H} \right)^m \exp\left(-a \frac{\gamma_S}{\Gamma_S}\right) \exp\left(-a \frac{\gamma_H}{\Gamma_H}\right) \, d\gamma_S \, d\gamma_H,$$

$$P_{out}^{C_2} \approx \sum_{m=0}^{M} A_m \int_{\Phi_H^{-1}(0)}^{\Phi_H^{-1}(0)} \int_{\Phi_S^{-1}(L_S)}^{\Phi_S^{-1}(H_S(\gamma_H))} \left(\frac{\gamma_S \gamma_H}{\Gamma_S \Gamma_H} \right)^m \exp\left(-a \frac{\gamma_S}{\Gamma_S}\right) \exp\left(-a \frac{\gamma_H}{\Gamma_H}\right) \, d\gamma_S \, d\gamma_H,$$

Impact of Spatial Correlation on Outage Probability

- For **correlated fading** on the S-D and H-D links

$D_H = 0$



Higher fading **correlation** induces **higher outage** probabilities

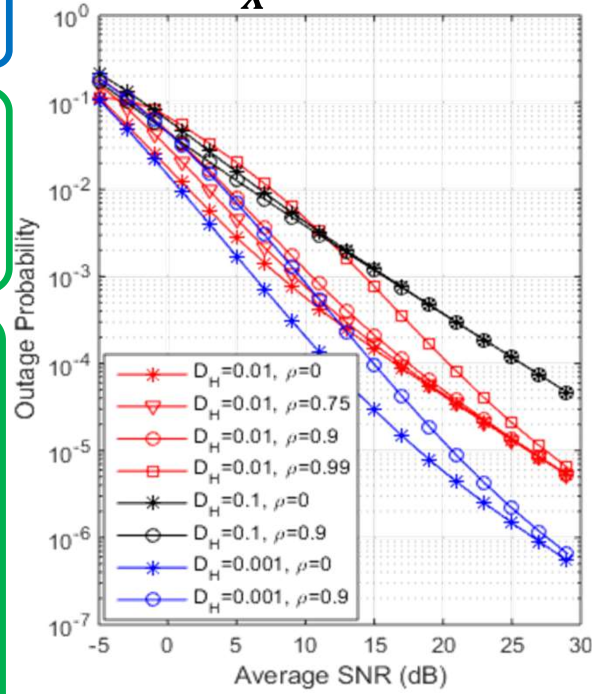
Increasing D_X **reduces** the outage probability

1st order diversity is obtained asymptotically

Second order diversity is achieved with $D_H = 0$ (even when $\rho \simeq 1$), independently of D_X

For small D_H and ρ values, **2nd order diversity** can be achieved at low average SNRs

$D_X = 0.2$



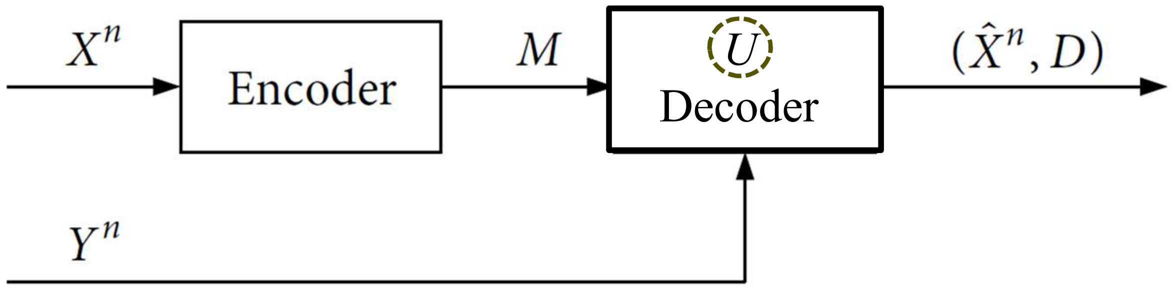
81

Fading correlation has no impact on the asymptotic diversity order

Chapter 3 Wyner-Ziv Formulation for Decision Making Process

3.1 Revisit of Helper-aided Lossy Networks

Wyner Ziv Networks:



Notice: U→X→Y forms a Markov Chain

$$I(X;U) - I(Y;U) = I(XY;U) - I(Y;U|X) - I(Y;U)$$
$$= I(XY;U) - I(Y;U) = I(X;U|Y).$$

We have used U→X→Y in the networks of:

- Outage analysis for wireless End-to-End Lossy Communications networks
- A two-stage wireless communications network based on Distortion Transfer Function
- Extension to two-sources one-helper End-to-End Lossy wireless communications network
- …

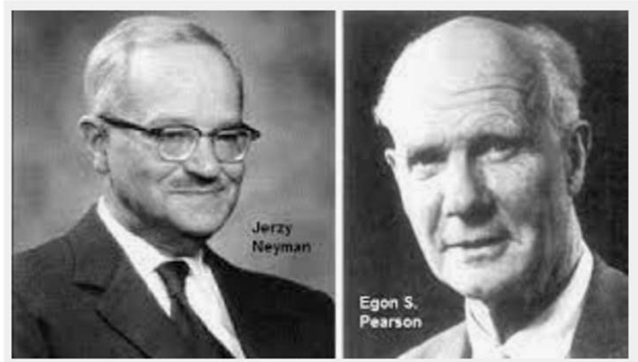
Fact: $I(X;U) - I(Y;U) = I(X;U|Y)$ can be understood as:

- *Y* is training sequence for **Machine Learning**,
- *Y* is training sequence, maybe followed by online observation, used for the knowledge updating of 1st and 2nd order statistics, *pdf* and Markov *dynamics*, in **Semantic Communications**.

Hypothesis Testing HT

3.2 Distributed Hypothesis Testing (DHT)

Landmark Builders for Hypothesis Testing (HT): Neyman-Pearson



Information Theoretic formulation of HT:

Basically, HT is a Code Design problem:

Design $f^{(n)}$ and $g^{(n)}$ such that

Minimize Type II Error Probability β_n subject to

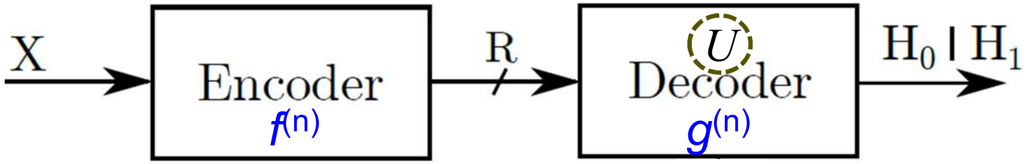
Type I Error Probability $\alpha_n \le \varepsilon$

$$\alpha_n = \mathbb{P}\left[g^{(n)}\left(f^{(n)}\left(\mathbf{X}^n\right)\right) = H_1 \mid H_0 \text{ is true}\right]$$

$$\beta_n = \mathbb{P}\left[g^{(n)}\left(f^{(n)}\left(\mathbf{X}^n\right)\right) = H_0 \mid H_1 \text{ is true}\right]$$

H_0: $X \sim P_{0,X}$

H_1: $X \sim P_{1,X}$



under constraint the rate R being given.

Note: Tradeoff: $\alpha_n \uparrow$, $\beta_n \downarrow$

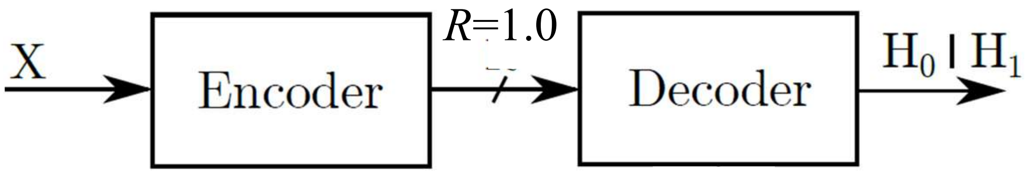
Note further: Decoder does NOT have to really "decode" to obtain *U*".

e.g., by Syndrome check only.

Nyman Pearson Test

With R=1.0, the decision problem boils down to traditional Nyman Pearson test using Likelihood:

$$\frac{P_0(X_1, X_2, \ldots, X_n)}{P_1(X_1, X_2, \ldots, X_n)} > T \,, \quad X_1, X_2, \ldots, X_n \in X, \text{ and} \qquad H_0: X \sim P_{0,X} \qquad H_1: X \sim P_{1,X}$$

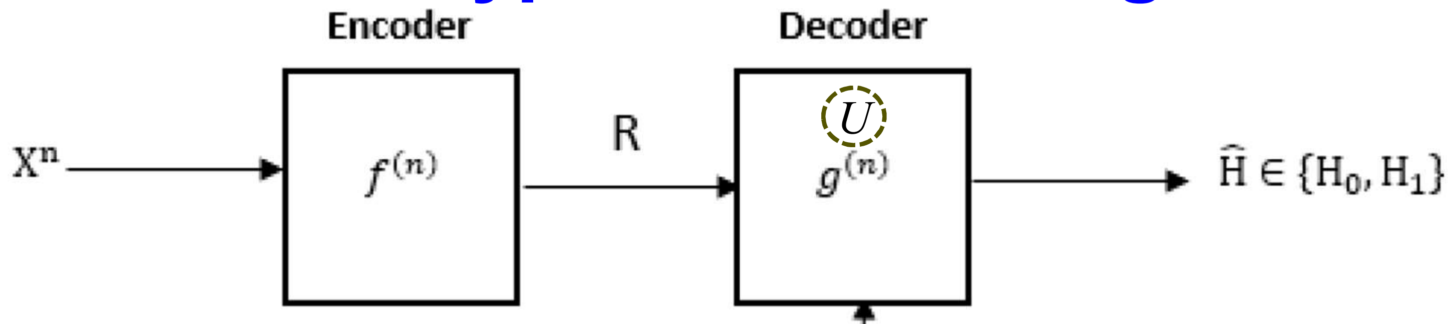


α_n and β_n are given by $\alpha^* = P_1^n(A_n^c(T)), \qquad \beta^* = P_2^n(A_n(T))$,

with $$A_n(T) = \left\{ \frac{P_1(x_1, x_2, \ldots, x_n)}{P_2(x_1, x_2, \ldots, x_n)} > T \right\}.$$

Note: Tradeoff $\alpha_n \uparrow$, $\beta_n \downarrow$ still holds.

Distributed Hypothesis Testing, DHT



$$f^{(n)} : \mathcal{X}^n \longrightarrow \mathcal{M}_n = \{1, 2, \cdots, \mathrm{M}_2\}$$
$$g^{(n)} : \mathcal{M}_n \times \mathcal{Y}^n \longrightarrow \mathcal{H} = \{H_0, H_1\}$$

$$H_0 : (\mathbf{X}^n, \mathbf{Y}^n) \sim P_{0,XY},$$
$$H_1 : (\mathbf{X}^n, \mathbf{Y}^n) \sim P_{1,XY}.$$

Decoder $g^{(n)}$ does NOT have to really "decode" to obtain *U*,
because the objective is to make a decision under the constraint on rate R.

- Objective : find the type-II error exponent θ such that type-I error is imposed and a rate constraint is satisfied

$$\alpha_n = \mathbb{P}\left[g^{(n)}\left(f^{(n)}(\mathbf{X}^n), \mathbf{Y}^n\right) = H_1 \mid H_0 \text{ is true}\right],$$
$$\beta_n = \mathbb{P}\left[g^{(n)}\left(f^{(n)}(\mathbf{X}^n), \mathbf{Y}^n\right) = H_0 \mid H_1 \text{ is true}\right].$$

The objective is that minimize β_n, subject to $\alpha_n \leq \varepsilon$. Tradeoff: $\alpha_n \uparrow$, $\beta_n \downarrow$

Ismaila Salihou Adamou, Elsa Dupraz1, Amin Zribi1, and Tad Matsumoto, "Error-Exponent of Distributed Hypothesis Testing for Gilbert-Elliot Source Models", to be published, Proc of IEEE ISTC 2023, Brest

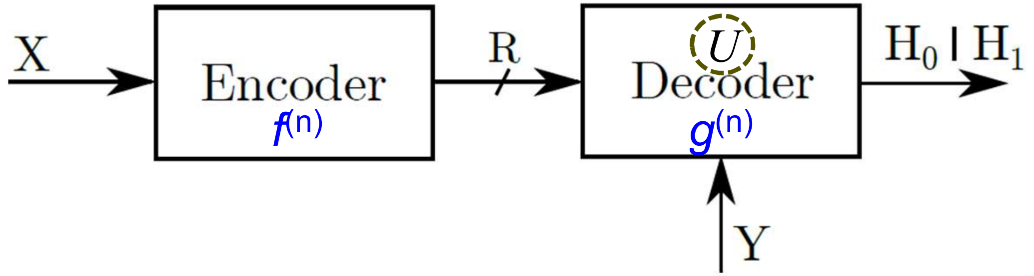
DHT Problems:

(1) In the same way as HT, again, a DHT problem is a Code Design problem:
Design $f^{(n)}$ and $g^{(n)}$ such that

Minimize Type II Error Probability β_n subject to
Type I Error Probability $\alpha_n \leq \varepsilon$ under constraint on the rate R given.

$$\alpha_n = \mathbb{P}\left[g^{(n)}\left(f^{(n)}(\mathbf{X}^n), \mathbf{Y}^n\right) = H_1 \mid H_0 \text{ is true}\right]$$
$$\beta_n = \mathbb{P}\left[g^{(n)}\left(f^{(n)}(\mathbf{X}^n), \mathbf{Y}^n\right) = H_0 \mid H_1 \text{ is true}\right]$$

$$H_0 : (\mathbf{X}^n, \mathbf{Y}^n) \sim P_{0,XY},$$
$$H_1 : (\mathbf{X}^n, \mathbf{Y}^n) \sim P_{1,XY}.$$



(2) Find Type-II Error Exponent θ

$$\limsup_{n\to\infty} \frac{1}{n} \log \frac{1}{\beta_n} \geq \theta$$

Subject to $\alpha_n \leq \epsilon$

and $\limsup_{n\to\infty} \frac{1}{n} \log \mathrm{M}_2 \leq \mathrm{R}$

Home

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

DHT: Type-II Error Exponent θ

$$\limsup_{n\to\infty} \frac{1}{n} \log \frac{1}{\beta_n} \geq \theta$$

It has already been known* that:

$$\theta = \min\{\underbrace{G(P_{0,UXY}, R)}_{\textit{binning-error}}, \underbrace{D(P_{0,UXY} \| P_{1,UXY})}_{\textit{decision-error}}\}$$

With the binning-error-function being:

$$G(P_{0,UXY}, R) = R - [I(X;U) - I(U;Y)]$$

where

$D(P_{0,UXY} \| P_{1,UXY})$ is the KL divergence

We rewite the binning-error-function:

$$G(R, R(D)) = R - R(D)$$

with

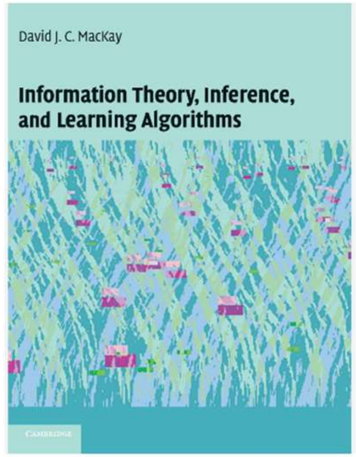
$$R(D) = \min\{I(X;U) - I(U;Y)\}$$

being Wyner Ziv $R(D)$ function! → The DHT Problem (1) boils down to **WZ Coding Problem**! (Left as Open Problem)

English Letters Appearence Probabilities

3.3 Semantic Communications

Copied from:



i	a_i	p_i
1	a	0.0575
2	b	0.0128
3	c	0.0263
4	d	0.0285
5	e	0.0913
6	f	0.0173
7	g	0.0133
8	h	0.0313
9	i	0.0599
10	j	0.0006
11	k	0.0084
12	l	0.0335
13	m	0.0235
14	n	0.0596
15	o	0.0689
16	p	0.0192
17	q	0.0008
18	r	0.0508
19	s	0.0567
20	t	0.0706
21	u	0.0334
22	v	0.0069
23	w	0.0119
24	x	0.0073
25	y	0.0164
26	z	0.0007
27	–	0.1928

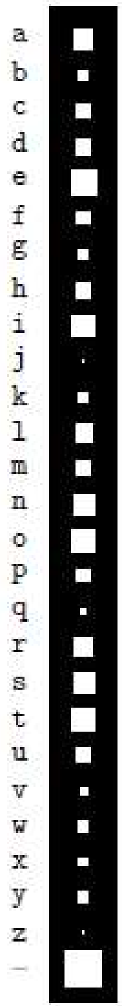
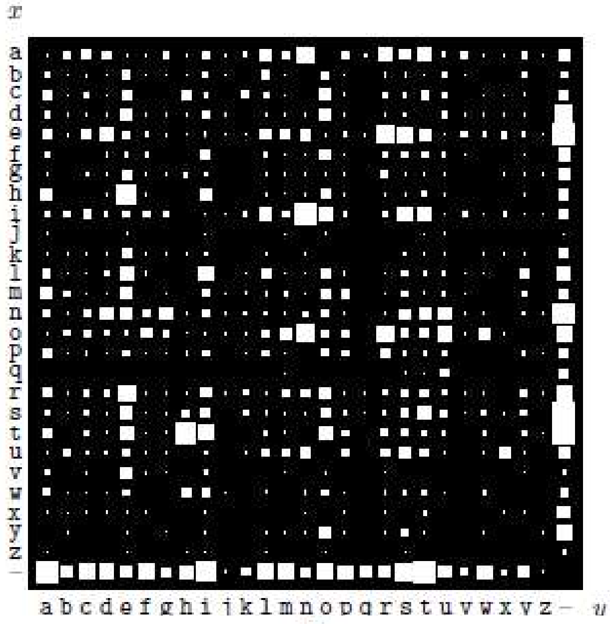
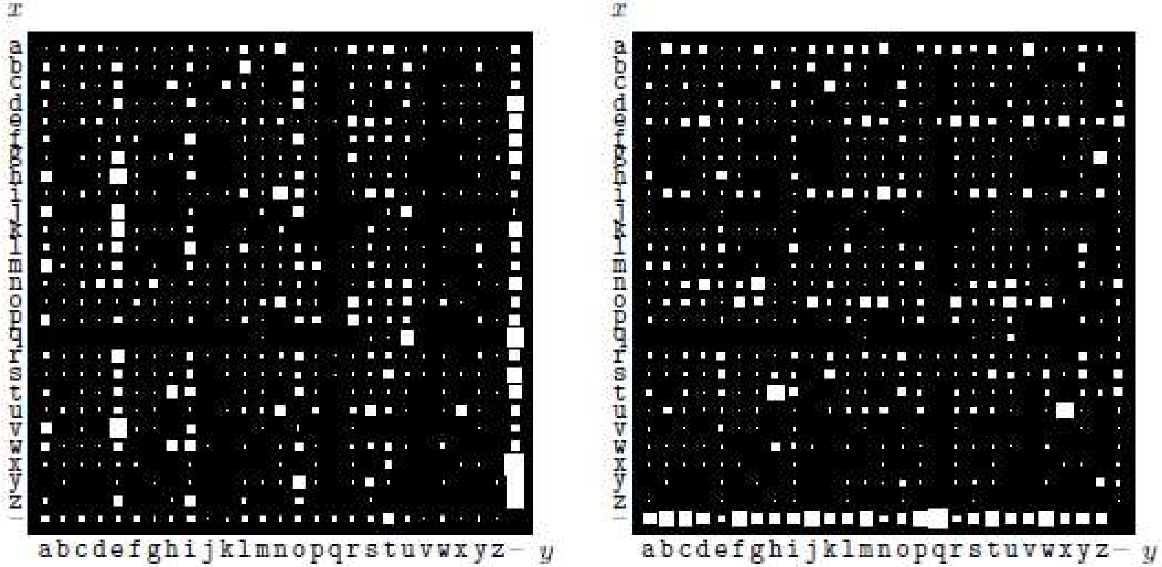


Figure 2.1. Probability distribution over the 27 outcomes for a randomly selected letter in an English language document

$P(x, y)$





(a) $P(y \mid x)$ (b) $P(x \mid y)$

JAIST JAPAN ADVANCED INSTITUTE OF SCIENCE AND TECHNOLOGY 1990

Conditioned English Letter Appearance

If we can evaluate conditional probabilities $p(x_i|x_{i-1})$, $p(x_i|x_{i-1}, x_{i-2})$, …, $p(x_i|x_{i-1}, x_{i-2}, \ldots, x_{i-n})$, empirically or theoretically and create a Markov model of the letter appearances, we can reduce the rate required to encode English.

Shannon's landmark paper presents artificially created English sentences!

Using empirical knowledge $p(x_i)$ →

$p(x_i|x_{i-1})$ →

$p(x_i|x_{i-1}, x_{i-2})$ →

1. *Zero-order approximation.* (The symbols are independent and equiprobable.)

   XFOML RXKHRJFFJUJ ZLPWCFWKCYJ

   FFJEYVKCQSGXYD QPAAMKBZAACIBZLHJQD

2. *First-order approximation.* (The symbols are independent. Frequency of letters matches English text.)

   OCRO HLI RGWR NMIELWIS EU LL NBNESEBYA TH EEI

   ALHENHTTPA OOBTTVA NAH BRL

3. *Second-order approximation.* (The frequency of pairs of letters matches English text.)

   ON IE ANTSOUTINYS ARE T INCTORE ST BE S DEAMY

   ACHIN D ILONASIVE TUCOOWE AT TEASONARE FUSO

   TIZIN ANDY TOBE SEACE CTISBE

4. *Third-order approximation.* (The frequency of triplets of letters matches English text.)

   IN NO IST LAT WHEY CRATICT FROURE BERS GROCID

   PONDENOME OF DEMONSTURES OF THE REPTAGIN IS

   REGOACTIONA OF CRE

Higher-order Conditioning:
Letter and Word Levels

Using empirical knowledge of $p(x_i|x_{i-1}, x_{i-2}, x_{i-3})$ →

5. *Fourth-order approximation.* (The frequency of quadruplets of letters matches English text. Each letter depends on the previous three letters. This sentence is from Lucky's book, *Silicon Dreams* [183].)

THE GENERATED JOB PROVIDUAL BETTER TRAND THE

DISPLAYED CODE, ABOVERY UPONDULTS WELL THE

CODERST IN THESTICAL IT DO HOCK BOTHE MERG.

(INSTATES CONS ERATION. NEVER ANY OF PUBLE AND TO

THEORY. EVENTIAL CALLEGAND TO ELAST BENERATED IN

WITH PIES AS IS WITH THE)

Instead of continuing with the letter models, we jump to word models.

Using empirical knowledge of the word appearance probability $p(w_i)$ →

6. *First-order word model.* (The words are chosen independently but with frequencies as in English.)

REPRESENTING AND SPEEDILY IS AN GOOD APT OR COME

CAN DIFFERENT NATURAL HERE HE THE A IN CAME THE TO

OF TO EXPERT GRAY COME TO FURNISHES THE LINE

MESSAGE HAD BE THESE.

$p(w_i|w_{i-1}, w_{i-2})$ →

7. *Second-order word model.* (The word transition probabilitiee match English text.)

THE HEAD AND IN FRONTAL ATTACK ON AN ENGLISH

WRITER THAT THE CHARACTER OF THIS POINT IS

THEREFORE ANOTHER METHOD FOR THE LETTERS THAT THE

TIME OF WHO EVER TOLD THE PROBLEM FOR AN

UNEXPECTED

With the 4^{th} order model, Shannon showed that 2.8 bits are enough to express one English letter!

Home

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

Adaptive Morse Code

Semantic Communications

Morse Code Alphabet

A	•-	N	-•	0	-----
B	-•••	O	---	1	•----
C	-•-•	P	•--•	2	••---
D	-••	Q	--•-	3	•••--
E	•	R	•-•	4	••••-
F	••-•	S	•••	5	•••••
G	--•	T	-	6	-••••
H	••••	U	••-	7	--•••
I	••	V	•••-	8	---••
J	•---	W	•--	9	----•
K	-•-	X	-••-	.	•-•-•-
L	•-••	Y	-•--	,	--••--
M	--	Z	--••	?	••--••

- The higher the appearance probability, the shorter the code length, following the Huffman coding rule.
- However, the appearance probabilities should change according to the sources, such as Book, Video, File type, …., situation, person, …← **Semantic Dependency**.
- Why not Source and Channel coding ***Jointly* Adaptive**, depending on "**Semantics**"?

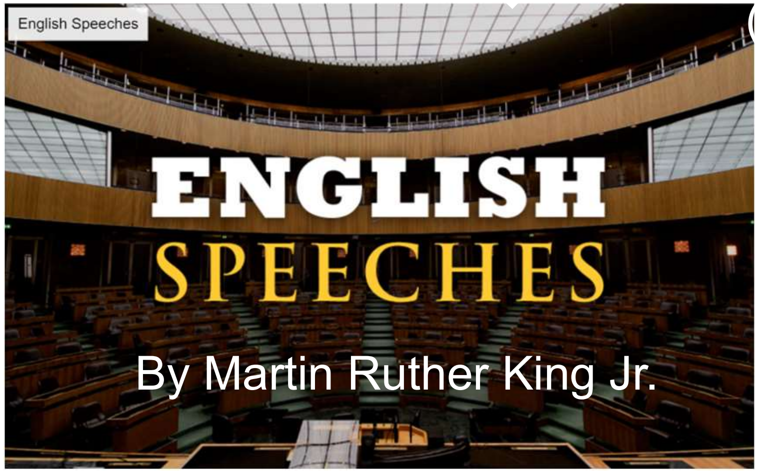
- Joint Coding can achieve "closer to optimal" than Separated Coding?

Knowledge updating for Semantic Communications: a WZ Problem

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

CWC Oulu

JAIST Japan Advanced Institute of Science and Technology 1990

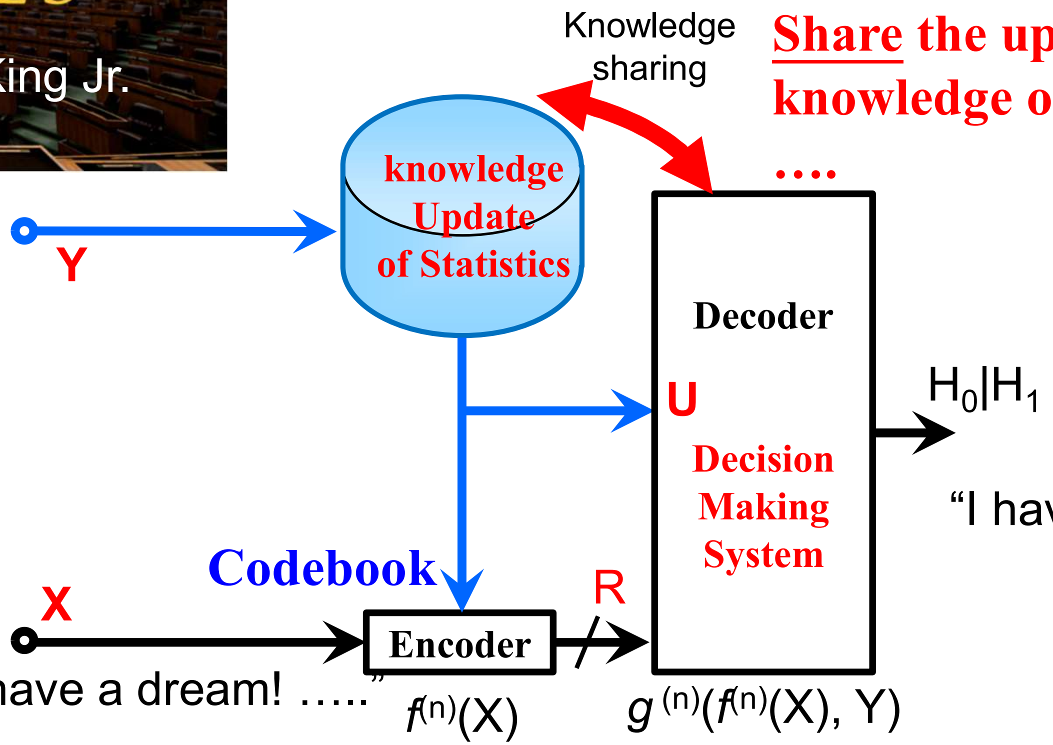


Updating the knowledge to minimize the rate R so that:

β_n, minimized subject to $\alpha_n \leq \varepsilon$

$$Min\ I(X;U) - I(Y;U) = \text{Min}\ I(X;U|Y)$$

Share the updated knowledge of *pdf*, dynamics, ….



“I have a dream! …..”

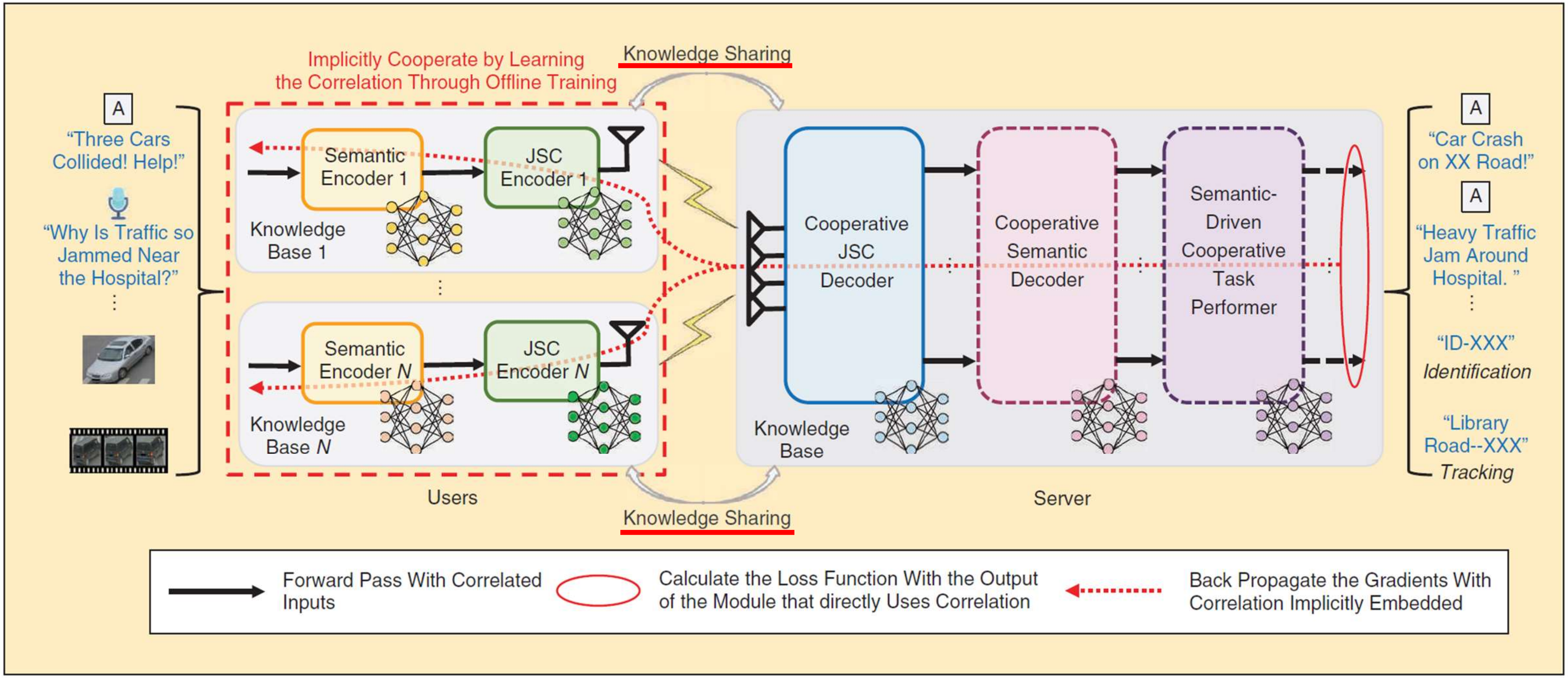
Remembering Martin Luther King Jr. | Tory Daily (toryburch.jp)

SEMANTIC COMMUNICATION FOR THE INTERNET OF VEHICLES

A Multiuser Cooperative Approach

Wenjun Xu, Yimeng Zhang, Fengyu Wang, Zhijin Qin,
Chenyao Liu, and Ping Zhang

IEEE VTS Magazine Volume 18, No. 1, pp. 100-109



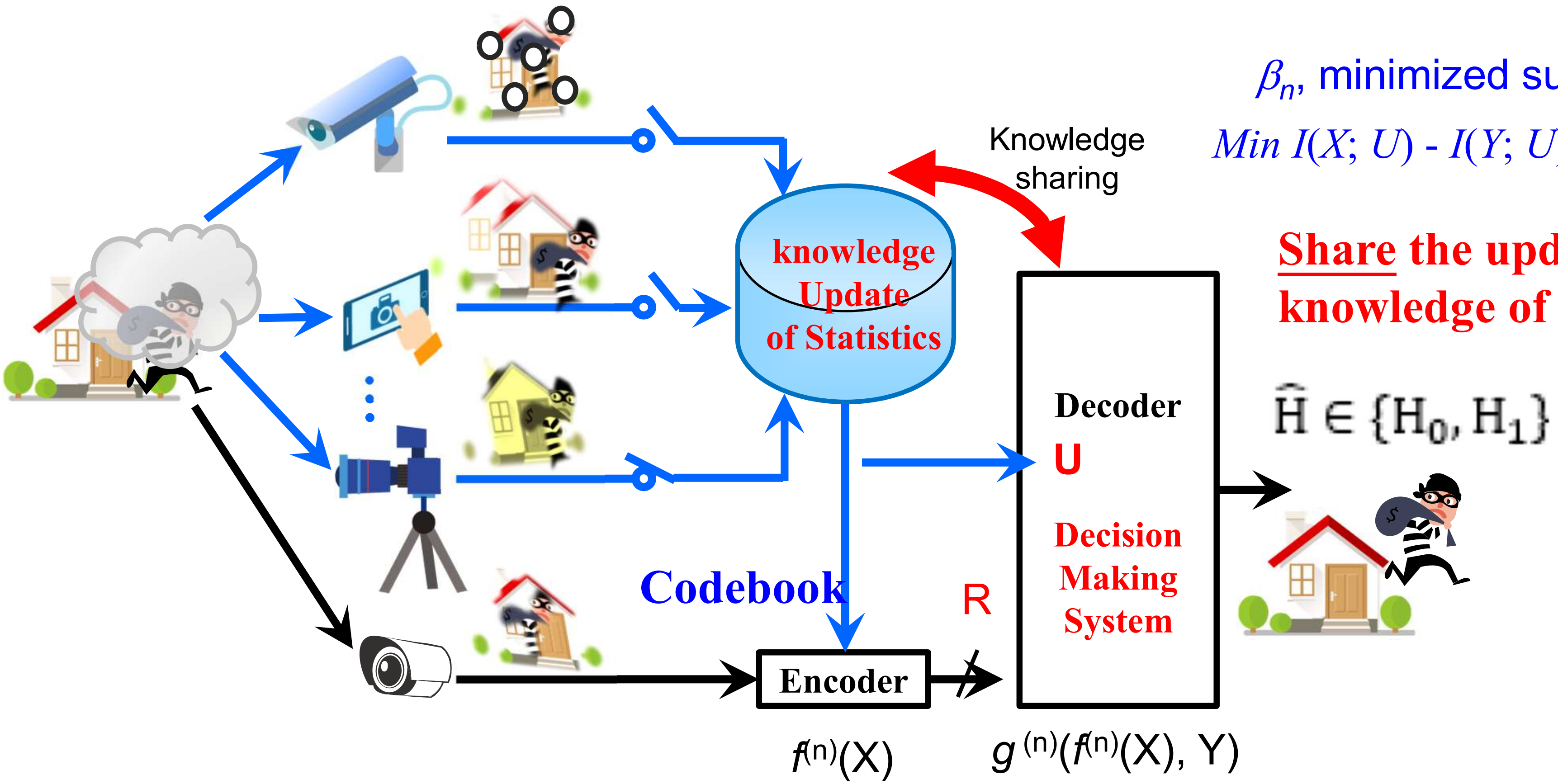
3.4 Training Process in Machine Learning

Updating the knowledge to minimize the rate R so that:

β_n, minimized subject to $\alpha_n \leq \varepsilon$

$$Min\ I(X;U) - I(Y;U) = Min\ I(X;U|Y)$$

Share the updated knowledge of *pdf*, dynamics



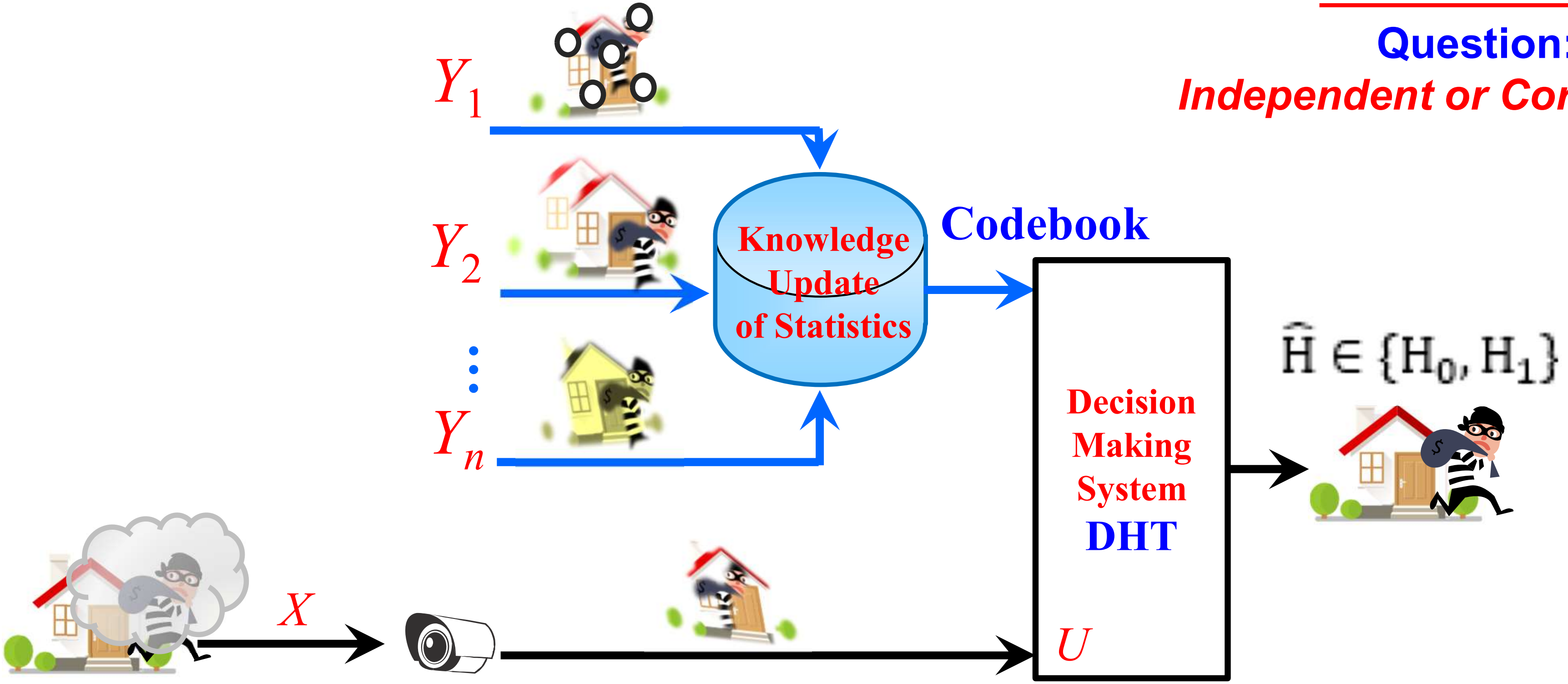
Learning, in the WZ framework

X corresponding to the Current Observation, U to the Lossy Reconstruction, and Y to Data Set for the Learning of Probability Distribution for knowledge updating followed by Codebook generation!

$$Y=(Y_1,Y_2, \ldots , Y_n)$$

$$R(D)=min\{I(X; U) - I(U; Y_1,Y_2, \ldots , Y_n)\}=min\ I(X;U|\ \underline{Y_1,Y_2, \ldots , Y_n})$$

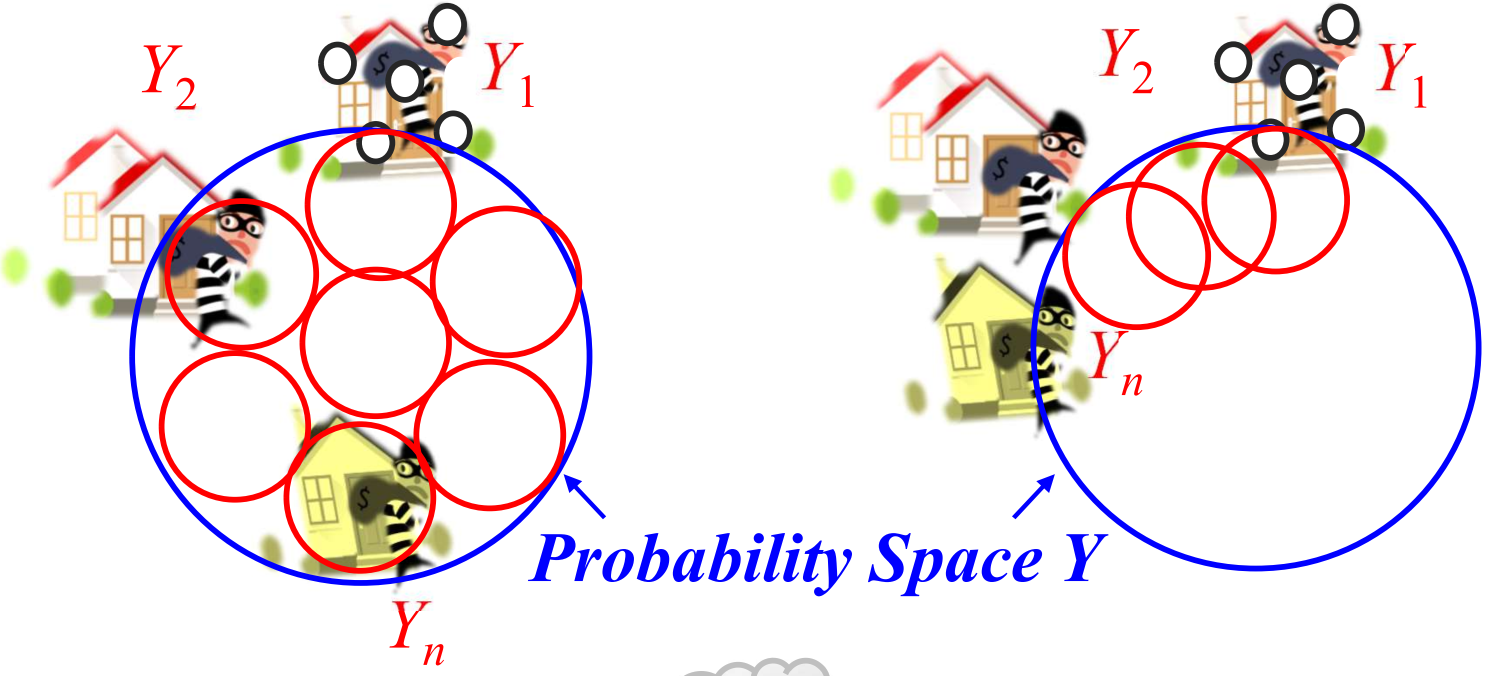
Question:
Independent or Correlated?



Learning in the WZ framework: Open Questions

(1) Y is fully covered by sub-probability space Y_i without overlapping.

(2) Y is NOT fully covered. Y_i are overlapping.



Decision on the observation X=  *being correct or incorrect depends on the generated learning data by* (1) *or* (2).

Ergodic vs. Instantaneous

(1) is suitable when *decision* is *Ergodic* (time average*).*

→Learning requires Large size of training data. *Suitable for pre-training, such as ML.*

(2) is suitable when *decision* is *Instantaneous*. Learning data may require only partial data.

→ *Suitable for online-training by introducing a forgetting factor.*

A similarity to *Information Bottleneck*!

Connection to *Information Bottleneck*

entropy

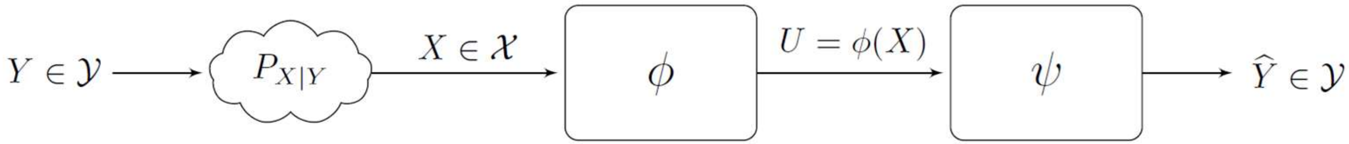
MDPI

Tutorial

On the Information Bottleneck Problems: Models, Connections, Applications and Information Theoretic Views

Abdellatif Zaidi [1,2,*] and Iñaki Estella-Aguerri [2] and Shlomo Shamai (Shitz) [3]

Specifically, IB formulates the problem of extracting the relevant information that some signal $X \in \mathcal{X}$ provides about another one $Y \in \mathcal{Y}$ that is of interest as that of finding a representation U that is maximally informative about Y (i.e., large mutual information $I(U;Y)$) while being minimally informative about X (i.e., small mutual information $I(U;X)$).

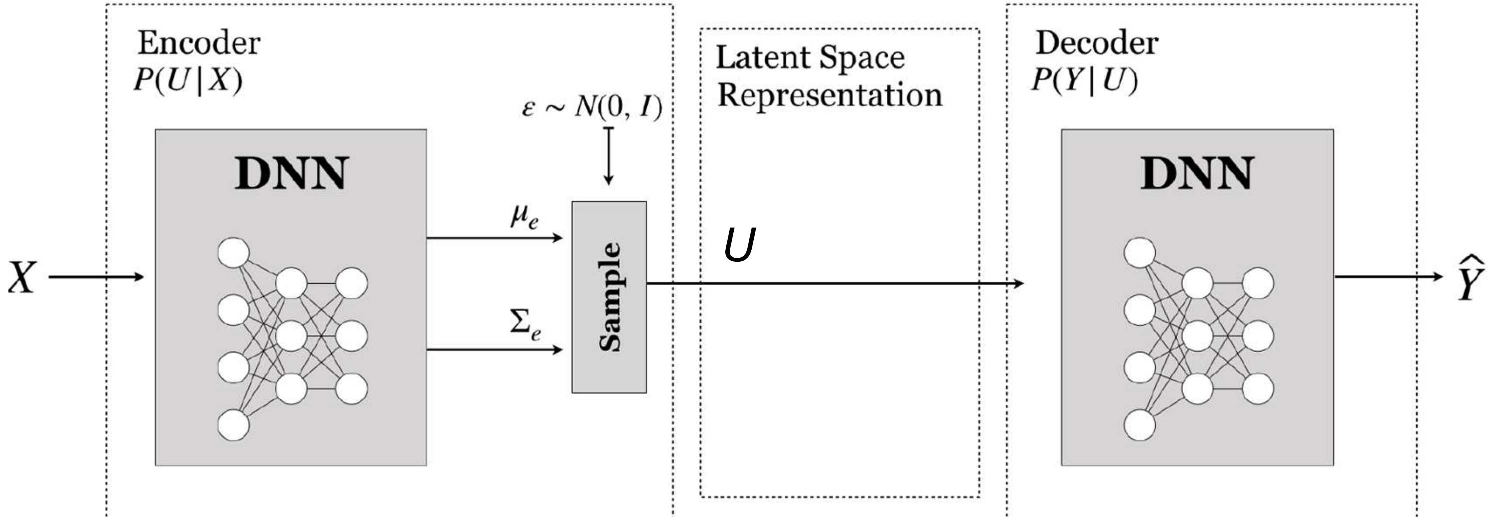


Information Bottleneck: Formulation under WZ Framework!

Accordingly, for a given β and source distribution $P_{X,Y}$, the optimal mapping of the data, denoted by $P^{*,\beta}_{U|X}$, is found by solving the IB problem, defined as

$$\mathcal{L}^{\text{IB}}_{\beta}(P_{U|X}) := I(U;Y) - \beta I(U;X)$$

with $U \rightarrow X \rightarrow Y$ and Lagrange Multiplier β. → We can use some optimization tools.

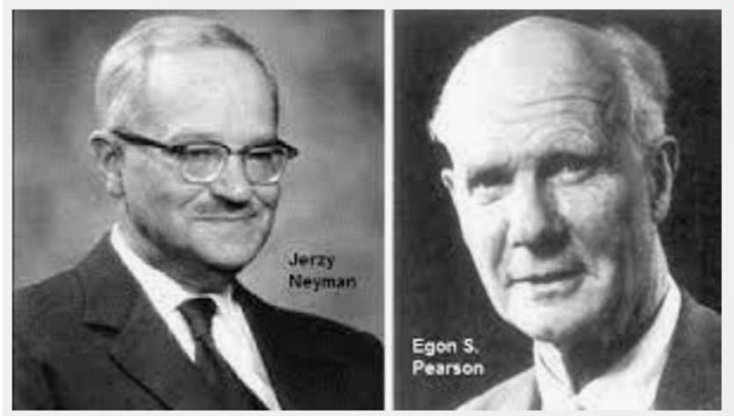


IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

Any Questions?

CWC Oulu

JAIST JAPAN ADVANCED INSTITUTE OF SCIENCE AND TECHNOLOGY 1990



Do they meet in 6G Networks?

My SISU Continues.

Thank you!

See you soon again somewhere in the world!