Decentralized Processing: An Information Theoretic Perspective

Shlomo Shamai, EE. Dept. Technion, Israel

sshlomo@ee.technion.ac.il

Communications Sciences Institute

Department of Electrical Engineering

USC Viterbi School of Engineering

February 2, 2007

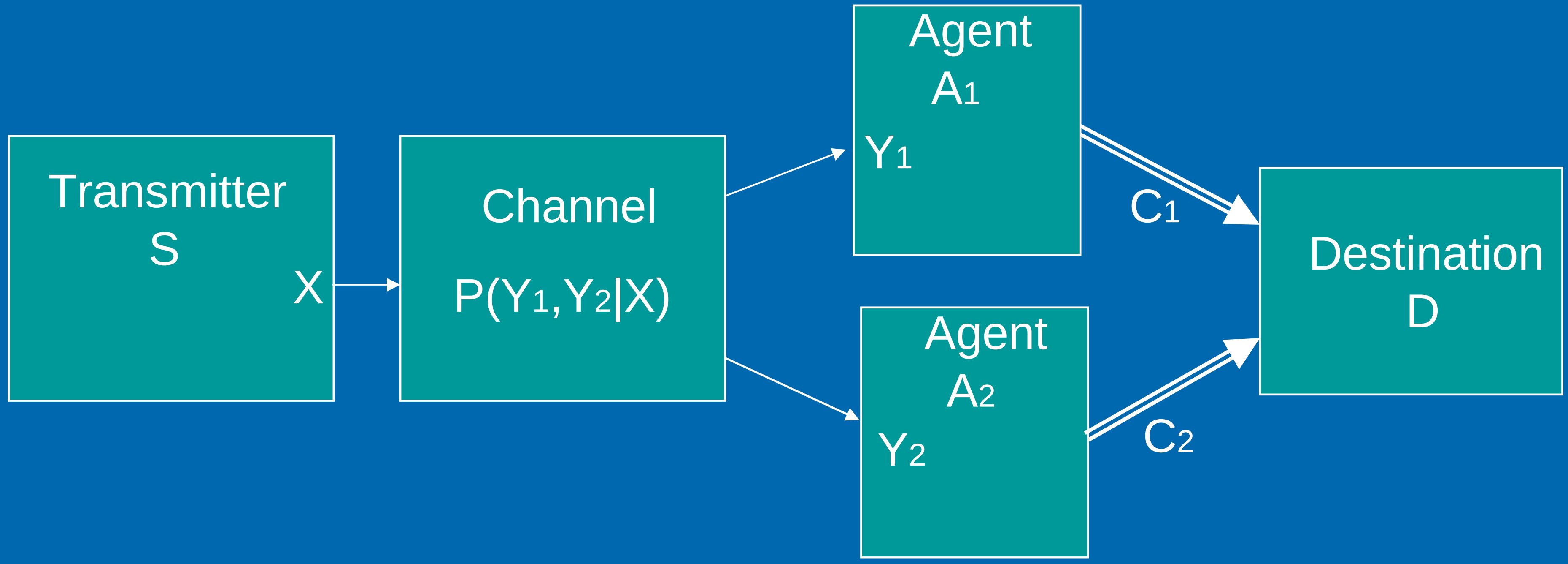
Joint work with Amichai Sanderovich, Yossef Steinberg and Michael Peleg, EE, Technion.

Overview of presentation

- Decentralized Processing: Overview
  - Nomadic transmitter and remote destination.
  - Simple access points (agents) that encode but do not decode.
- Information theoretic upper and lower bounds.
- The Gaussian scalar channel – capacity through EPI.
- Decentralized MIMO: Multiple antenna Tx:
  - Two schemes: Simple and Wyner-Ziv compression.
  - Multiplexing gain of the schemes.
  - Upper bounds for block fading.
- Impact of common feedback – cooperation mode.
  - Two nomadic schemes.
  - Limited benefits of the feedback – due to the nomadic setting.

Decentralized Processing: Overview

- Nomadic transmitter



Scheme description

- Memoryless broadcast channel without feedback: $P(Y_1^n,\ldots,Y_T^n \mid X^n) = \prod_{i=1}^{n} P(Y_{1,i},\ldots,Y_{T,i} \mid X_i)$

  Connected via T (here T=2) non-interfering and reliable links with bandwidths $C_1,\ldots,C_T$ bits/channel use.
- One shot transmission (one block), but extends to steady-state operation.
- Nomadic transmitter can not adapt to the different access points (agents), where each may use different ways of forwarding.
- Nomadic transmitter → the agents lack the codebook knowledge.

Scheme description

Agents:

- Agents observe a memoryless noisy source → similar to the CEO problem.
- Each agent encodes $n \leq k$ channel outputs into an index $V_t \in \{1, \ldots, 2^{nC_t}\}$, $V_t = \phi_t(Y_t^n)$.
- Each agent sends the k/n indices to the destination.

Final destination:

- Decodes the transmitted message from the received indices $\{V_{1,\ldots,T}^{k/n}\}$.

Problem statement

- Nomadic transmitters:
  - Agents are ignorant of code-book structure.

- Objective: maximize the total achievable rate of the scheme over
  - Agents encoding functions.
  - Single letter distribution used for codebook selection.
  - Decoding function.

Some Relevant Literature

- Distributed source coding are directly applicable:
  - The CEO problem [Berger, Zhang and Viswanathan 1996].
  - Upper bound on the sum rate distortion function for the CEO problem [Chen, Zhang, Berger and Wicker 2004].
  - Distributed Wyner-Ziv source coding over tree structure [Draper and Wornell 2004].
  - Rate Region of Quadratic Gaussian CEO [Oohama 2005].
  - Network Vector Quantization [Fleming, Zhao and Effros 2004].
  - Rate Region of the Quadratic Gaussian two terminals source coding [Wagner, Tavildar and Viswanath2006].
  - Backward channels, Distributed WZ [Servetto 2006].

Relevant literature

- Relevant channel coding problems:
  - Parallel relay scheme [Schein/Gallager 2001].
  - Relay channel [Cover and El-Gamal 1979].
  - Multihopping for relay networks [Kramer, Gastpar and Gupta 2004].
  - User cooperation [Sendonaris, Erkip and Aazhang 2003].
  - MIMO relay problems [Wang, Zhang,Høst-Madsen 2005].

Achievable rate

- Agents compress their received signal:

$$\left(U_t, Y_t\right) \in \mathrm{T}_\varepsilon$$

where U_t depends only on Y_t, i.e. the following are Markov chains:

$$U_t - Y_t - \left(X, Y_{1,\ldots,t-1,t+1,\ldots,T}, U_{1,\ldots,t-1,t+1,\ldots,T}\right)$$

- Transmit $\left\{U_1, \ldots, U_T\right\}$ to the destination, save rate by using the correlations between them (Wyner-Ziv coding).

Achievable rate

- Achievable rate

$$R < \max_{P(X)P(Y_1,\ldots,Y_T|X)\prod_{t=1}^{T} P(U_t|Y_t)} I(X;U_1,\ldots,U_T)$$

- Where $\sum_{t\in S} C_t > I(U_S;Y_S|U_{S^C})$

for all $S \subseteq \{1,\ldots,T\}$

The Gaussian channel

- For a Gaussian channel with:

$$Y_t = X + N_t, \quad E|X|^2 = P_X, \; E|N_t|^2 = P_{N_t}$$

- The capacity is

$$R \le \max_{0 \le r_t \le C_t} \min_{S \subseteq \{1,\ldots,T\}} \sum_{t \in S} [C_t - r_t] + \frac{1}{2}\log_2\left(1 + P_X \sum_{t \in S^C} \frac{1 - 2^{-2r_t}}{P_{N_t}}\right)$$

- Proof: through entropy power inequality and the contra-polymatroid region for the achievable rates [Tse et. al. 2004, Oohama 2005].
- r_t can be interpreted as the bandwidth used for noise quantization.

The Gaussian channel

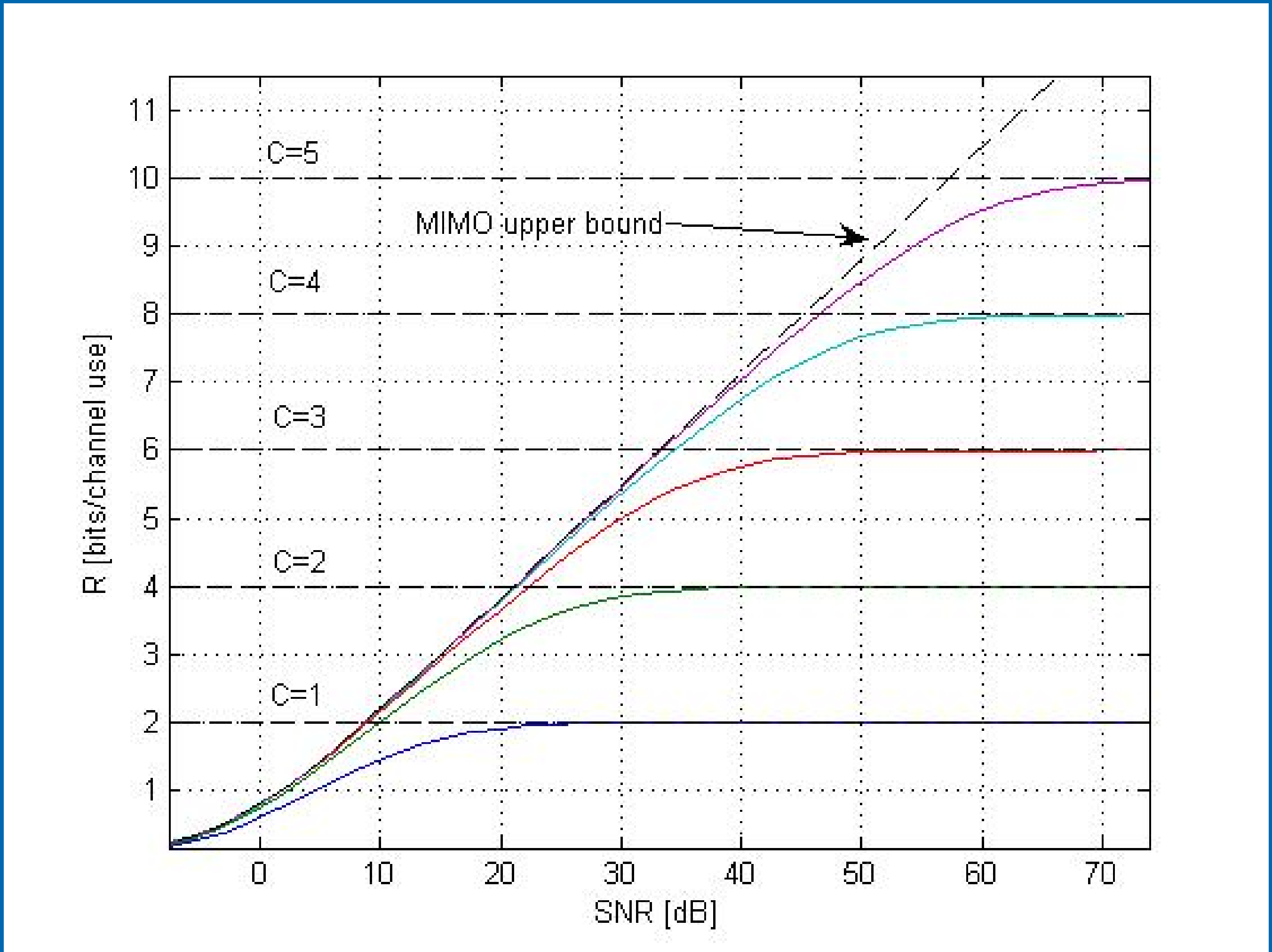
- For symmetric agents, analytic solution gives:

$$R = \frac{1}{2}\log_2\left(1+2SNR\left(1-\frac{\sqrt{SNR^2+2^{4C}(1+2SNR)}-SNR}{2^{4C}}\right)\right)$$

- Asymptotics:

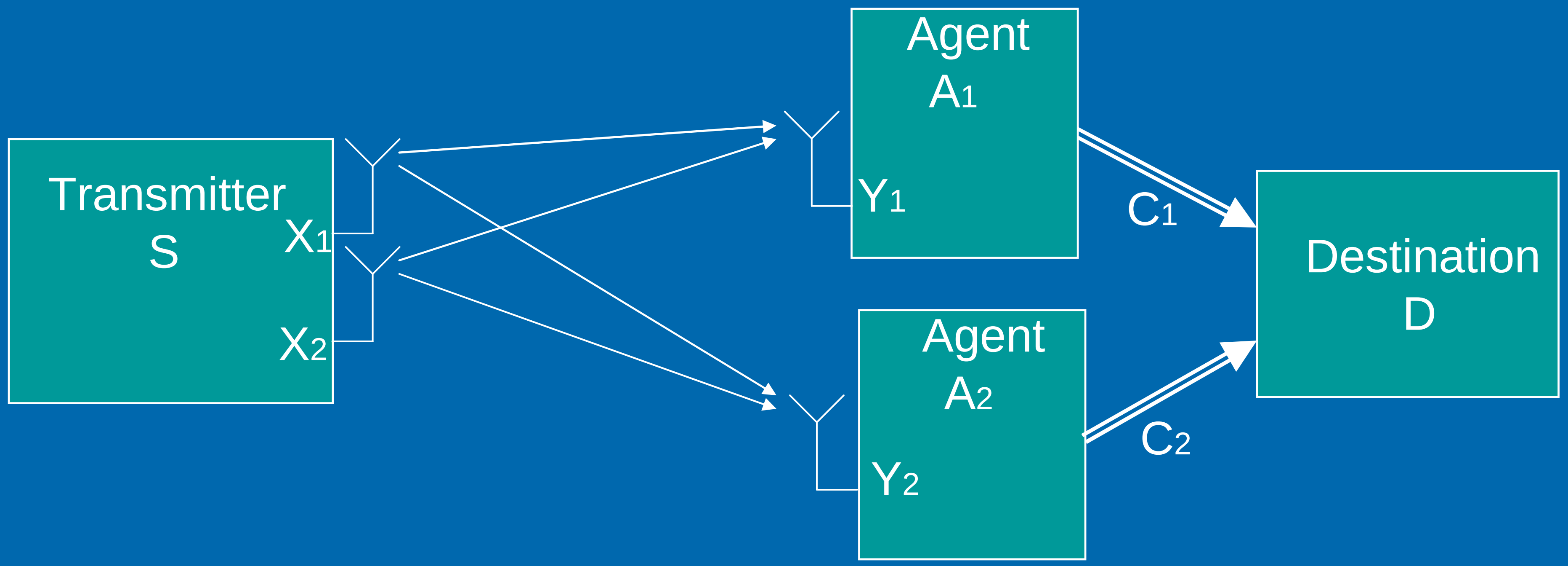
$$\begin{cases} SNR \to \infty : \quad R = 2C \\ C \to \infty : \quad R = \frac{1}{2}\log_2\left(1+2SNR\right) \end{cases}$$

The Gaussian channel



Decentralized MIMO: Scheme description

- Nomadic multi antenna transmitter



Decentralized MIMO: Scheme description

- MIMO channel:
  t Tx antennas and r agents with single antenna.
- Rayleigh fading channel:
  where $Y = HX + N$

$$Y \in \mathbb{C}^{[r,1]},\ X \in \mathbb{C}^{[t,1]},\ \ E\left[X^* X\right] \le P,\ N \sim \mathbb{C}N\left(0, I_r\right)$$

$$h_i \sim \mathbb{C}N\left(0, I_t\right),\ \ i = 1,\ldots,r, \qquad H^T = \left[h_1, h_2, \ldots, h_r\right]$$

- H fully known to both agents and remote destination, can be fast fading (Shannon capacity),
  and block fading channel (average rate).
- Connected to the remote destination via orthogonal reliable links with bandwidths $C_1,\ldots,C_r$ bits/channel use.

Relevant literature – Distributed MIMO

- Relevant channel coding problems:
  - Gaussian MIMO channels [Telatar 1999].
  - MIMO broadcast [H. Weingarten, Y. Steinberg, and S. Shamai 2004].
  - Capacity scaling laws in MIMO relay networks [Bolcskei, Nabar, Oyman and Paulraj 2006].
  - Capacity-achieving input covariance for single-user MIMO [Tulino, Lozano and Verdu 2006].
  - Quadratic gaussian two-terminal source-coding [Wagner, Tavildar and Viswanath 2005].
  - Gaussian Many-Help-One Problem [Tavildar, Viswanath and Wagner 2007].

Achievable rate – simple compression

- Each agent needs to know only its own fading h_i. Destination knows H.
- Agents compress their received signal (i is the agent index): Y_i^n to the codewords U_i^n,

Codebook size is 2^{nC_i}.

U_i is defined by: $U_i = Y_i + d_i, \quad d_i \sim \mathrm{N}\left(0, P_{D_i}\right)$

$$P_{D_i} = \frac{|h_i|^2 P/t + 1}{2^C - 1}$$

Achievable rate – simple compression

- Transmit $\{U_1,\ldots,U_r\}$ to the destination.
- Simple compression: no use of dependencies of $\{U_1,\ldots,U_r\}$.
- Then uses $\{U_1,\ldots,U_r\}$ to decode X.
- The achievable rate is:

$$R_{SC} = E_H \log_2 \det\left(I_r + \frac{P}{t}\begin{pmatrix} \frac{1}{P_{D_1}+1} & & \\ & \ddots & \\ & & \frac{1}{P_{D_r}+1} \end{pmatrix} HH^* \right)$$

Achievable rate WZ compression

- Use the dependencies between receptions to increase compression efficiency.
- Requires the knowledge of H in all agents.
- Define $U_i = Y_i + d_i$,

  Only now: $d_i \sim \mathrm{N}\left(0, P_{D_i}\right)$, $\frac{1}{P_{D_i}+1} = 1 - 2^{-r_i(H)}$

- $r_i\left(H\right)$ is bandwidth wasted on noise compression (this is due to nomadic assumption: no decoding, so only source coding possible [Oohama2005]).

Achievable rate WZ compression – cont.

- The achievable rate:

$$R_{WZ} = E_H\left[\max_{0\le r_i(H)\le C}\ \min_{S\subseteq\{1,\ldots,r\}} R_S\left(\left\{r_i(H)\right\}_{i=1}^{r}, H_S, C\right)\right]$$

$$R_S\left(\left\{r_i(H)\right\}_{i=1}^{r}, H_S, C\right) \triangleq \sum_{i\in S^C}\left[C - r_i(H)\right]$$

$$+\log_2\det\left(I_{|S|} + \frac{P}{t}\,\mathrm{diag}\left(\left\{\frac{1}{P_{D_i}+1}\right\}_{i\in S}\right)H_S H_S^*\right)$$

- Notice that $C - r_i(H)$ is the bandwidth left for signal transmission.

$$H_S \triangleq H(i,j), \qquad i\in S,\ j=1,\ldots,t$$

Achievable rate WZ compression – cont.

- The optimization over $\left\{r_i\left(H\right)\right\}_{i=1}^{r}$ in WZ compression needs to be performed for any H
- The optimization is concave, and thus can be performed efficiently.
- As $t \to \infty$, the channel matrix hardens, and no need for H in the agents, since it is required only for the WZ binning resolution.

Multiplexing Gain

- Multiplexing gain = $\lim_{P\to\infty} \frac{R(P)}{\log_2(P)}$

- For simple compression, when $t \geq r$ and $C = \log_2(P)$ the compression noise d_i in the definition of U_i is with power

$$P_{D_i} = \frac{|h_i|^2 P/t + 1}{2^C - 1} \underset{C=\log(P)}{=} \frac{|h_i|^2 P/t + 1}{P - 1}$$

Upper bounded by:

$$P_{D^*} = \max_i \frac{|h_i|^2 P/t + 1}{2^C - 1} \underset{C=\log(P)}{=} \max_i \frac{|h_i|^2 P/t + 1}{P - 1}$$

Multiplexing Gain

- Thus full multiplexing Gain is attained (inequality since we use the largest compression noise, there are others smaller)

$$R_{SC} \geq E_H \log_2 \det\left(I_r + \frac{P}{t}\frac{1}{1+P_{D^*}} HH^* \right) = \sum_{i=1}^{r} E_{\lambda_i} \log_2\left(1 + \frac{P\lambda_i / t}{1+P_{D^*}} \right)$$

$$un-ordered \ \ \text{eignvalues}\left(HH^* \right) = \left\{ \lambda_i \right\}$$

$$\lim_{P\to\infty} \frac{R_{SC}\left(P\right)}{\log_2\left(P\right)} \geq r$$

- When $t < r$, only subset of t agents can be used.

Multiplexing Gain

- Wyner-Ziv compression is always better than the simple compression, and thus, trivially:

$$\lim_{P\to\infty} \frac{R_{WZ}\left(P\right)}{\log_2\left(P\right)} \geq m, \qquad m \triangleq \min\left\{r, t\right\}$$

Nomadic joint UB

- For upper bounds on average rate with block fading (not valid for fast fading):
- Assume that the agents can fully cooperate.
- Equivalent to m parallel channels, with signal to noise ratios: $\{\lambda_i P/t\}$
  results in almost simple MIMO.
- In [SSSK2005] there was no H. We copy that result for a fixed H:

$$R(H) \le \max_{\substack{\sum B_i \le C_{total} \\ B_i \ge 0}} \sum \log_2 \left(1 + P\lambda_i / t \frac{2^{B_i} - 1}{2^{B_i} + P\lambda_i / t} \right)$$

Nomadic separated UB

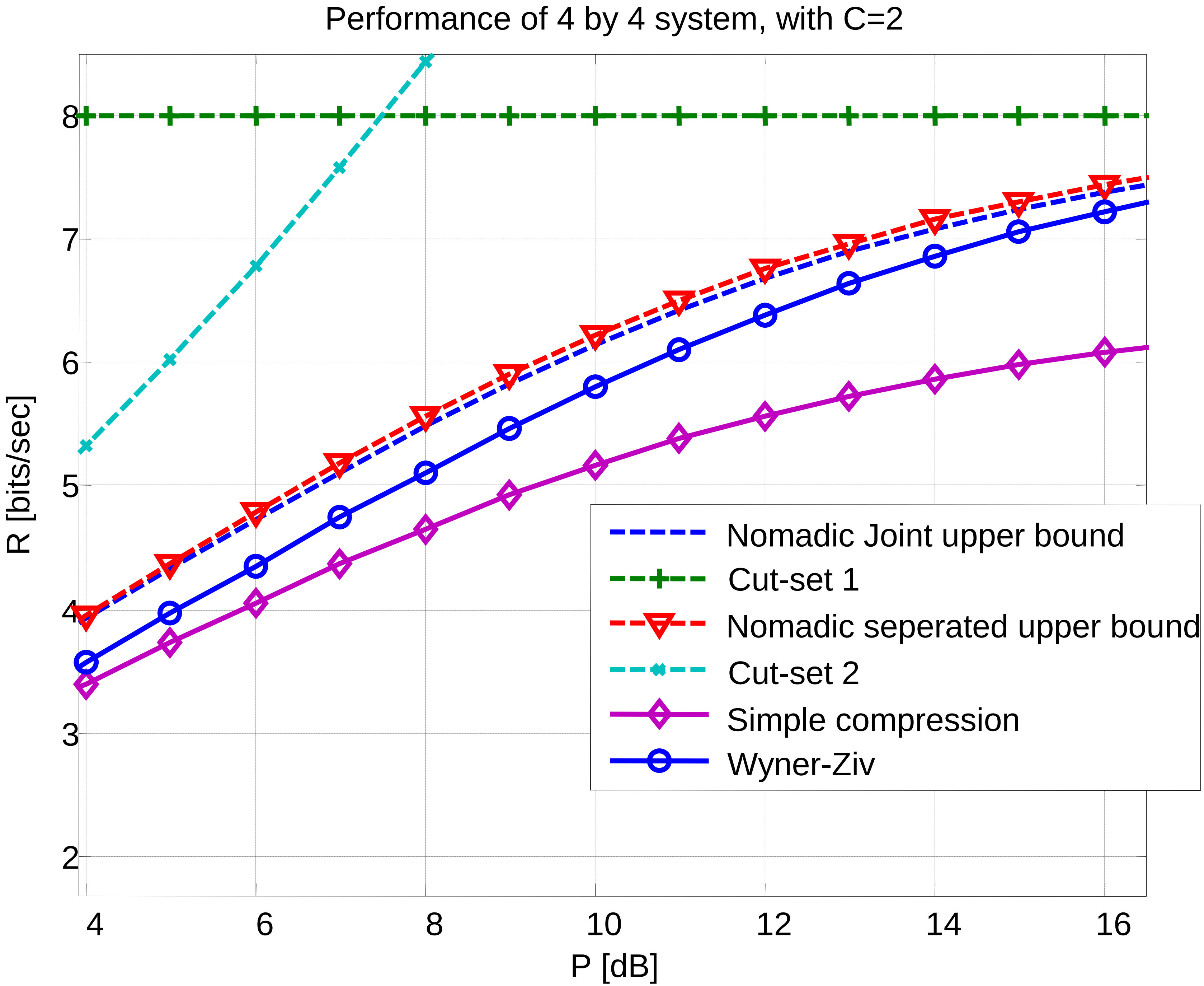
- Another upper bound is based on:

$$R \le \frac{1}{n} I\left(X^n ; V_{\{1,\ldots,r\}} \mid H \right) \le \frac{1}{n} \sum_{i=1}^{r} I\left(X^n ; V_i \mid H \right)$$

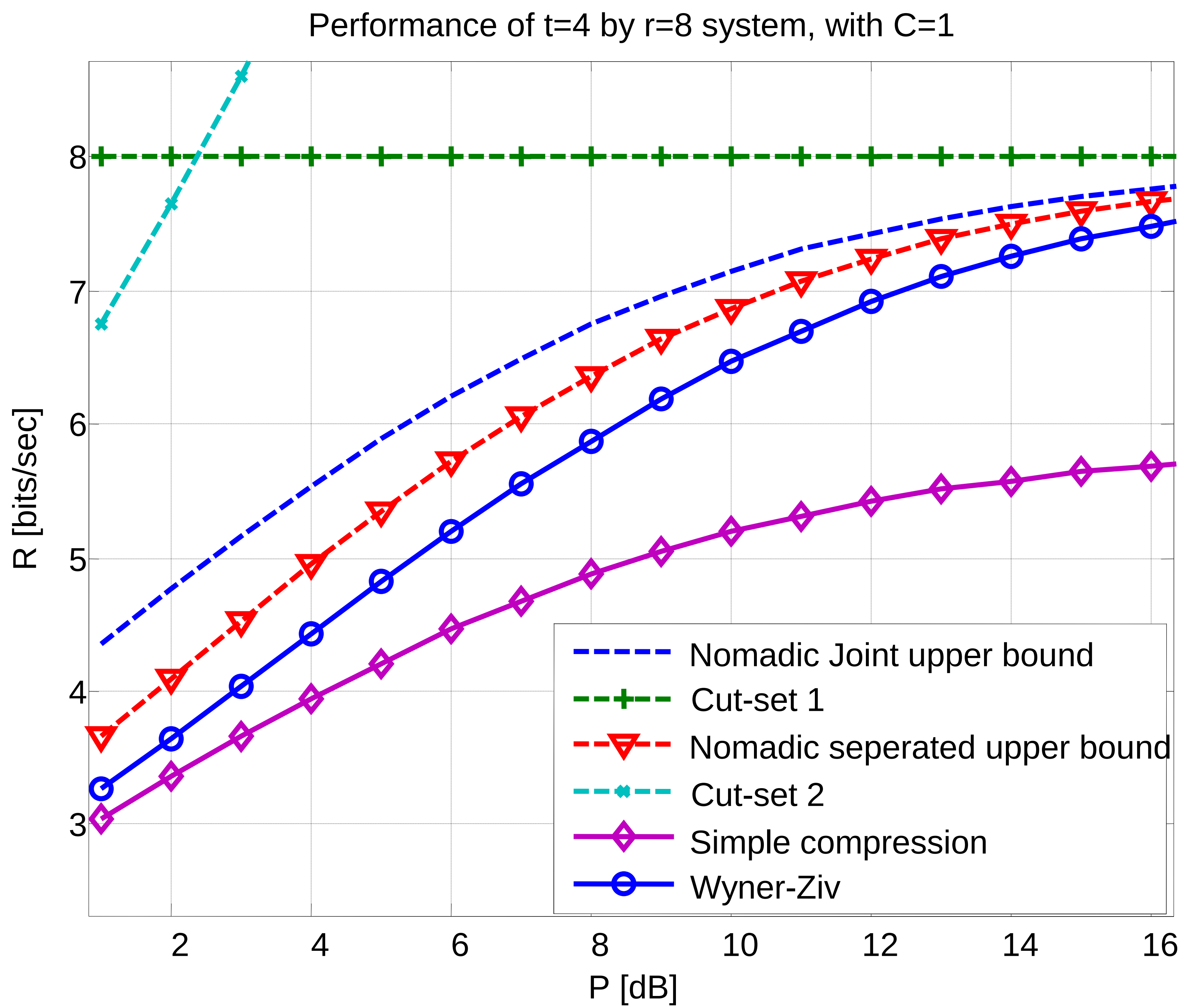
V_i is the message sent form the i^{th} agent.

- Every term in the sum can be treated as a single-agent model, number of antennas at the source is not important. Invoking results of [SSSK2005] for single antenna, for every term of the above sum (loose: not accounting for the single source setting):

$$R(H) \le \sum_{i=1}^{r} \log_2 \left(1 + P/t \left| h_i \right|^2 \frac{2^{C_i} - 1}{2^{C_i} + P/t \left| h_i \right|^2} \right)$$



Performance of 4 by 4 system, with C=2
R [bits/sec]
P [dB]
2
3
4
5
6
7
8
4
6
8
10
12
14
16
Nomadic Joint upper bound
Cut-set 1
Nomadic seperated upper bound
Cut-set 2
Simple compression
Wyner-Ziv



Performance of t=4 by r=8 system, with C=1
R [bits/sec]
P [dB]
3
4
5
6
7
8
2
4
6
8
10
12
14
16
Nomadic Joint upper bound
Cut-set 1
Nomadic seperated upper bound
Cut-set 2
Simple compression
Wyner-Ziv

Asymptotics – simple compression

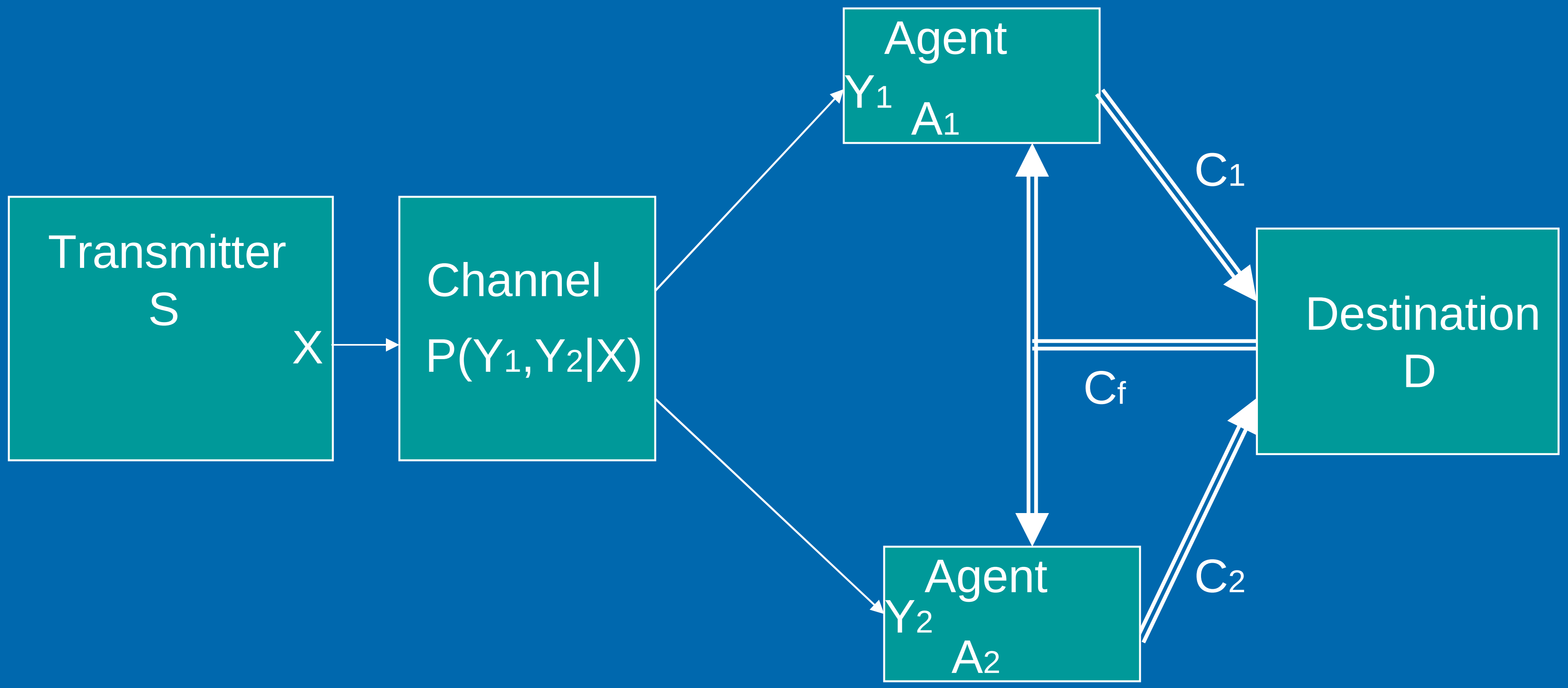
- Consider the case of: $r, t \to \infty$
  while $\tau = \frac{r}{t}$ and $C = \frac{C_{total}}{r}$, C_{total} fixed, ind. of r.
- Here we take min/max on P_{D_i} to get upper/lower bounds:
  → identical asymptotically: $\forall i : \frac{|h_i|^2}{t} \xrightarrow{a.s.} 1$
- simple compression:

$$E_H \log_2 \det\left(I_r + \frac{P}{t} \frac{1}{1+P_{D_*}} HH^* \right) \to mE_H \log_2\left(1 + \frac{P\lambda / t}{1 + (P+1) \big/ \left(2^{C_{total}/r} - 1 \right)} \right)$$

- So $\lim_{r\to\infty} R_{SC} = \frac{C_{total} P}{1+P}$ (no dependence on $\tau = \frac{r}{t}$).

- Notice that $E_H \log_2 \det\left(I_r + \frac{P}{t} HH^* \right) \xrightarrow[P \ll 1]{} rP$

Feedback link: $t=1,\ r=2.$



Relevant literature – Feedback - Cooperation

- Feedback to one agent, in asymmetric case [Schein-Gallager 2001].
- Coding for interactive communication [Schulman 1996].
- Cooperative relay broadcast channels [Liang-Veeravalli 2005].
- Cooperative receivers [Dabora-Servetto 2005\6].
- Network-Coding in Interference Networks [Smith-Vishwanath 2006].

Feedback link - setting

- The final destination has finite capacity C_f feedback – common to the two agents.
- Cooperation is limited to two phases.
- The final destination does not reveal codebook to agents.
- Feedback link is used to improve compression quality.
- Network coding can be used.

Feedback link - setting

- First phase: the agents (1,2) transmit M_1 and M_2, respectively.
- The final destination receives these messages and forwards M_f through the feedback link, involving network coding.
- Second phase: the first agent sends M'_1 and the second M'_2, through the loss less links.

Feedback link - setting

- We have that:

$$\frac{1}{n}\log_2\left(\mathrm{M'_t M_t}\right) \le C_\mathrm{t}, \qquad \mathrm{t}=1,2$$

$$\frac{1}{n}\log_2\left(\mathrm{M_f}\right) \le C_f$$

- where

$$\mathrm{M_t} = \phi_\mathrm{t}\left(Y_\mathrm{t}^n\right), \qquad \mathrm{t}=1,2$$

$$\mathrm{M'_t} = \phi'_\mathrm{t}\left(Y_\mathrm{t}^n, \mathrm{M_f}\right)$$

$$\mathrm{M_f} = \phi\left(\mathrm{M_1}, \mathrm{M_2}\right)$$

Feedback link – achievable rate #1

- First phase: each agent uses WZ to compresses Y_t^n into U_t^n.
- Then sends the corresponding bin by M_t.
- The destination uses M_1, M_2 to decode $\left(U_1^n, U_2^n\right)$ and then sends $M_f = \phi\left(M_1, M_2\right)$.
- Each agent then decodes the compressed vector U_{3-t}^n, from M_f with the SI Y_t^n.
- Network coding effect example:

$$M_f = \phi\left(M_1\right) \oplus \phi\left(M_2\right)$$

Feedback link – achievable rate #1

- Second phase: each agent uses WZ (with SI at agents & destination) to compresses Y_t^n into Z_t^n, given U_{3-t}^n.
- Then sends the corresponding bin: M'_t.
- The destination finally decodes $\left(Z_1^n, Z_2^n\right)$ from the received M'_1, M'_2.
- Since agents have better SI than D, they can decode U_{3-t}^n.

Feedback link – achievable rate #1

- Achievable rate:

$$R < \max I(X;U_1,U_2,Z_1,Z_2)$$

With the constraints:

$$\left.\begin{aligned} &I(U_1;Y_1|U_2) \le C_f \\ &I(U_2;Y_2|U_1) \le C_f \\ &I(Y_1,Y_2;U_1,U_2) \le 2C_f \end{aligned}\right\} \text{Standard WZ}$$

$$I(Z_1;Y_1|U_1,U_2,Z_2) \le C - I(U_1;Y_1|U_2)$$

$$I(Z_2;Y_2|U_1,U_2,Z_1) \le C - I(U_2;Y_2|U_1)$$

$$I(Z_1,Z_2;Y_1,Y_2|U_1,U_2) \le 2C - I(Y_1,Y_2;U_1,U_2)$$

Conditional-WZ on remaining BW

Feedback link – achievable rate #1

- For Gaussian setting, if side information $\left(U_1^n, U_2^n\right)$ is known at decoder (destination): no rate gain by sending it to the encoder (agents).

  WZ = Conditional Rate-Distortion.
- This mean that for Gaussian setting, the scheme can do the same as without feedback.

- For other setting (such as binary), the scheme can use the feedback for rate improvement.

Feedback link – achievable rate #2

- The destination does not decode $\left(U_1^n, U_2^n\right)$ at the first phase.
- For decoding at the agents, first phase requires:
  $I\left(U_1, Y_1 | Y_2\right) \le I\left(U_1, Y_1 | U_2\right)$ (since $U_1 - Y_1 - Y_2, U_2$).
  → A1 and A2 have better SI than D.
- Second phase: each agent uses WZ to compress $\left(Y_t^n, U_{3-t}^n\right)$ into Z_t^n.
- Then sends the corresponding bin: M'_t.
- The destination finally jointly decodes $\left(U_1^n, Z_1^n, U_2^n, Z_2^n\right)$ from the received M_1, M'_1, M_2, M'_2.
- Better than decoding $\left(U_1^n, U_2^n\right)$ and $\left(Z_1^n, Z_2^n\right)$ separately.

Feedback link – achievable rate #2

- Achievable rate is:
  $R < \max I(X;U_1,U_2,Z_1,Z_2)$

  But, with the constraints:

$$\left.\begin{aligned} I(U_1;Y_1|Y_2) \le C_f \\ I(U_2;Y_2|Y_1) \le C_f \end{aligned}\right\} \text{Standard WZ}$$

$$I(Z_1;Y_1|U_1,U_2,Z_2) + \max\{I(U_1;Y_1|Y_2), I(U_2;Y_2|Z_2,U_1)\} \le C$$

$$I(Z_2;Y_2|U_1,U_2,Z_1) + \max\{I(U_2;Y_2|Y_1), I(U_1;Y_1|Z_1,U_2)\} \le C$$

$$I(Z_1,Z_2;Y_1,Y_2|U_1,U_2) + I(Y_1,Y_2;U_1,U_2) \le 2C$$

Conditional-WZ

Remaining BW under decoding at both agents and D

Feedback link – achievable rate #2

- For Gaussian symmetric channel, the achievable rate with no feedback is with equality in the non-diagonal constraint:

$$I\left(U_1,U_2;Y_1,Y_2\right)\le 2C$$

- The achievable rate in the second scheme, is limited by the same constraint:

$$I\left(Z_1,Z_2;Y_1,Y_2\,|\,U_1,U_2\right)+I\left(Y_1,Y_2;U_1,U_2\right)\le 2C$$

- Thus, for the Gaussian symmetric setting, no benefit from the feedback.

Feedback link – achievable rate #3

- An increase in the achievable rate, for the Gaussian setting: by not requiring the destination to decode $\left(U_1^n, U_2^n\right)$, at all.
- This means an improved compression, on the expense of wasted bandwidth.
- Nomadic setting is not improved even by considering Layering (since, again Gaussian WZ equals conditional MI).

Feedback link – achievable rate #3

- Achievable rate:

$$R < \max I(X;Z_1,Z_2)$$

With the constraints:

$$\left.\begin{aligned} I\left(U_1;Y_1 \mid Y_2\right) &\le C_f \\ I\left(U_2;Y_2 \mid Y_1\right) &\le C_f \end{aligned}\right\} \text{ Standard WZ}$$

$$I\left(Z_1;Y_1,U_2 \mid Z_2\right) \le C - I\left(U_1;Y_1 \mid Y_2\right)$$

$$I\left(Z_2;Y_2,U_1 \mid Z_1\right) \le C - I\left(U_2;Y_2 \mid Y_1\right)$$

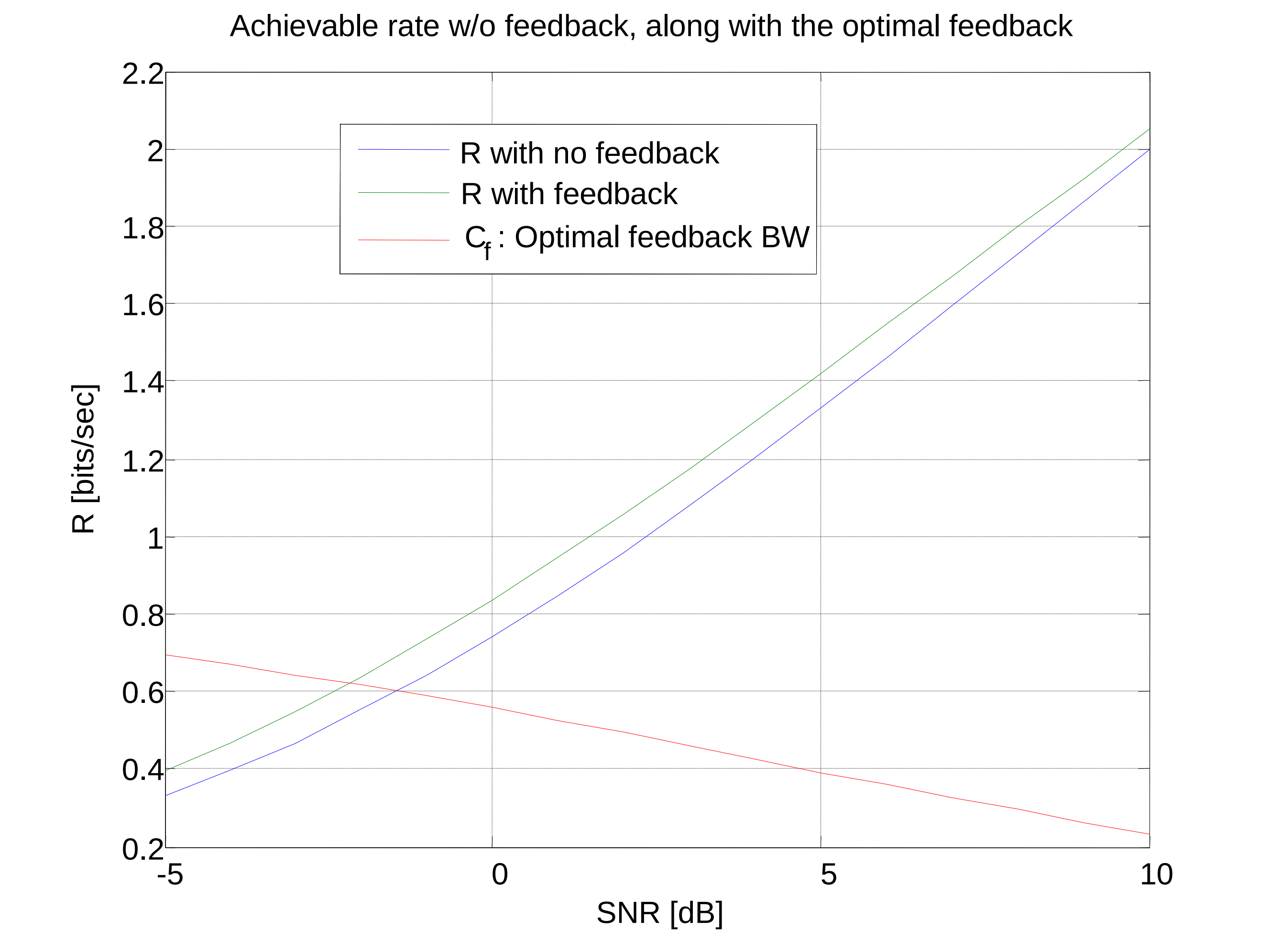
$$I\left(Z_1,Z_2;Y_1,Y_2,U_1,U_2\right) \le 2C - \sum_{i=1,2} I\left(U_i;Y_i \mid Y_{3-i}\right)$$

WZ on remaining BW

Achievable with feedback – numerical example

- U's not decoded by Destination.

- C_f required for the best rate enhancement.



Conclusions

- A framework for communication schemes which includes a nomadic transmitter communicating through agents.
- The Gaussian case is solved provided the agents do not know/use the code book of the nomadic transmitter.
- Explicit solution is given for two equivalent agents.
- Simple and Wyner-Ziv compression schemes for nomadic multi antenna transmitter communicating through agents.
- The full multiplexing gain is demonstrated for both compression schemes, no CSI at transmitter!

Conclusions – Cont.

- Upper bounds for average rate (block fading).
- Numerical results demonstrate the tightness of the bounds and the effectiveness of the WZ approach.
- For the single transmit antenna: the impact of a feedback link is investigated, in three variations.
- A nomadic setting substantially limits the gain from the feedback link.
- For efficient use of the feedback, the agents need decoding ability.

Outlook

- Distributed MIMO: Improved upper bounds, based on EPI vector versions.
- Distributed MIMO: Multiple antennas agents.
- Distributed MIMO: Including decoding agents (broadcast MIMO channel).
- Distributed MIMO: Multi cell sites joint processing.
- The impact of feedback with decoding agents.

Thank you!