Correlation and Cooperation in Wireless Communications

Şennur Ulukuş

Department of ECE

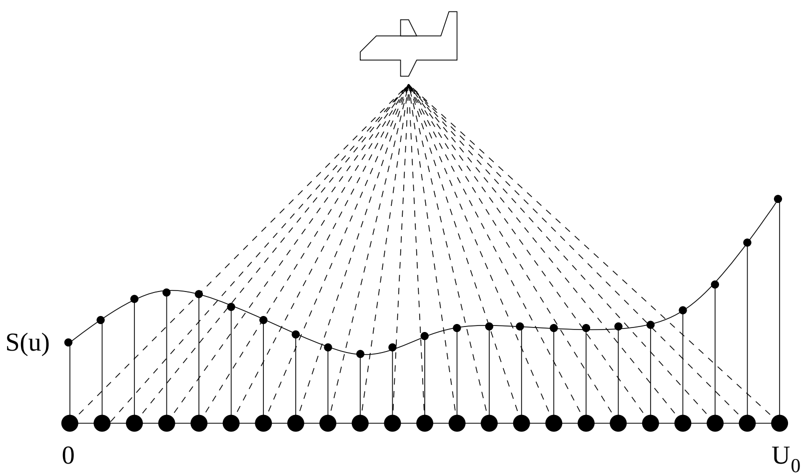
University of Maryland

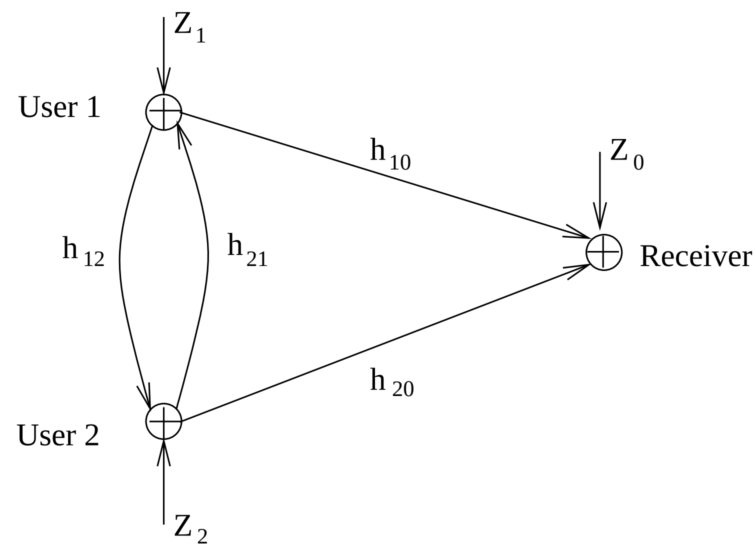
ulukus@umd.edu

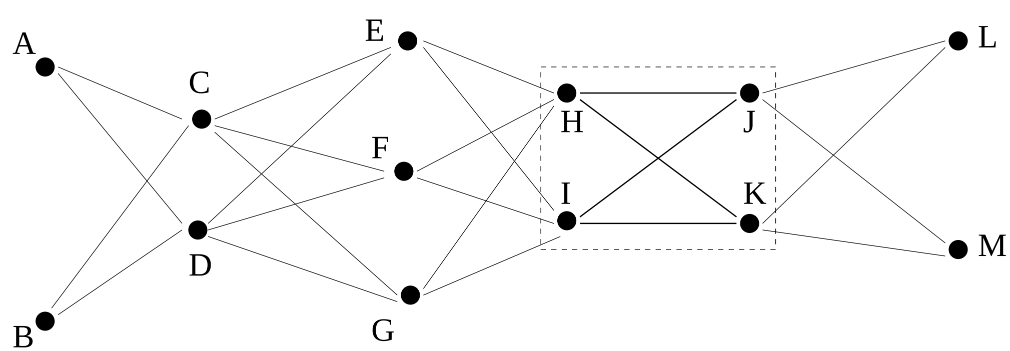
Joint work with Wei Kang and Nan Liu.

Correlated Data

- Correlated data arises in wireless communications for many reasons.
  - Measured data may be correlated as in sensor networks,
  - Correlated data may be created by communication between users as in user cooperation,
  - Correlated data may result from relaying and multi-hopping in large networks.





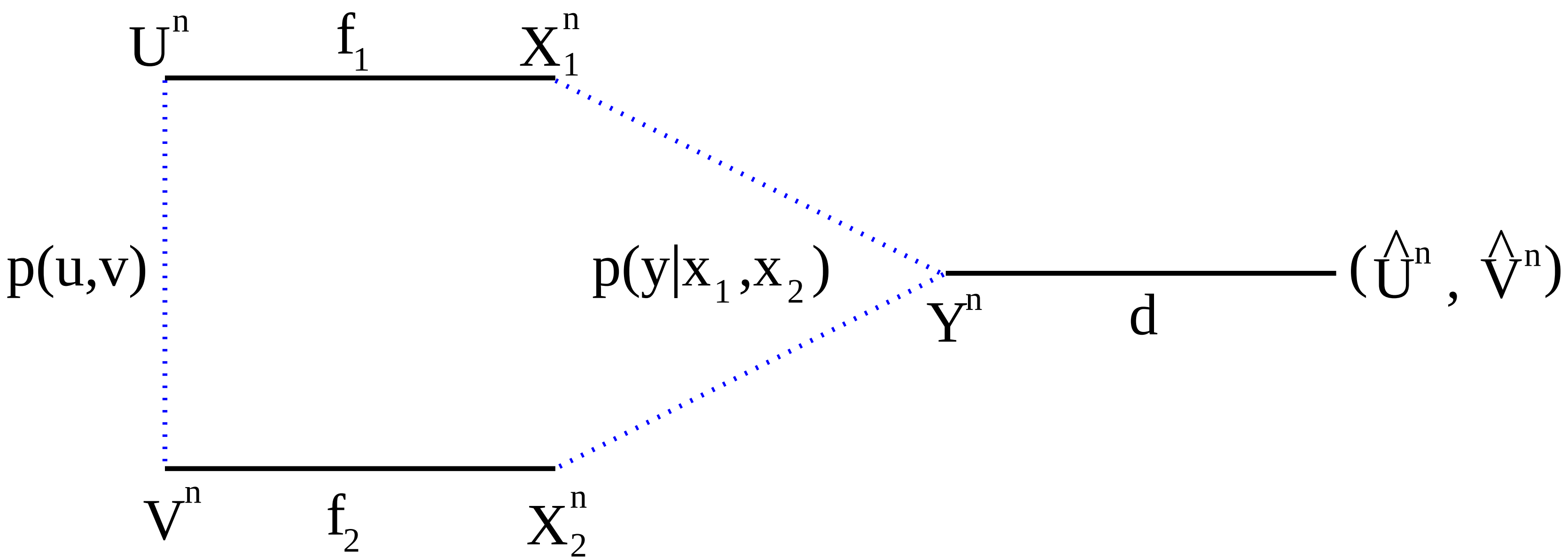


Outline of the Talk

- Two-user multiple access channel with arbitrarily correlated data
  - The capacity region is still an open problem
  - Currently there are
    - single-letter sub-optimal achievable schemes [Cover, El Gamal, Salehi], [Dueck]
    - n-letter incomputable tight outer bounds [Cover, El Gamal, Salehi]
  - We develop a computable single-letter outer bound using our new data processing inequality
- N-user multiple access channel with correlated data (a.k.a. dense sensor network)
  - Correlated Gaussian random process, Gaussian channel with cooperation, $N \to \infty$
  - The general problem is still open
  - We develop lower and upper bounds for a certain class of Gaussian random processes
  - We identify cases where our lower and upper bounds meet and yield **order-optimal** schemes

Two-user Multiple Access Channel

- A joint source-channel coding problem
  - Sources: (U^n, V^n), i.i.d., $p(u,v)$.
  - Channel: discrete memoryless, $p(y|x_1,x_2)$.
  - Encoding/decoding: $X_1^n = f_1(U^n)$, $X_2^n = f_2(V^n)$, $(\hat{U}^n, \hat{V}^n) = d(Y^n)$.
  - Necessary and sufficient conditions for reliable transmission?



Existing Single-letter Results

- Independent sources: $p(u,v) = p(u)p(v)$

$$
\begin{aligned}
H(U) &< I(X_1;Y|X_2) \\
H(V) &< I(X_2;Y|X_1) \\
H(U)+H(V) &< I(X_1,X_2;Y)
\end{aligned}
$$

  for some $p(x_1,x_2) = p(x_1)p(x_2)$.

- Correlated sources: arbitrary $p(u,v)$
  - Cover, El Gamal and Salehi, 1980: A single-letter achievability result,

$$
\begin{aligned}
H(U|V) &< I(X_1;Y|X_2,V) \\
H(V|U) &< I(X_2;Y|X_1,U) \\
H(U,V) &< I(X_1,X_2;Y)
\end{aligned}
$$

    for some $X_1 \rightarrow U \rightarrow V \rightarrow X_2$.
  - Not optimal, counter example by Dueck, 1981.

Arbitrarily Correlated Sources: Capacity Result in n-letters

- Correlated sources: arbitrary $p(u,v)$
  - Cover, El Gamal and Salehi, 1980:

$$H(U|V) < \frac{1}{n} I(X_1^n; Y^n | X_2^n, V^n)$$
$$H(V|U) < \frac{1}{n} I(X_2^n; Y^n | X_1^n, U^n)$$
$$H(U,V) < \frac{1}{n} I(X_1^n, X_2^n; Y^n)$$

    where $X_1^n \rightarrow U^n \rightarrow V^n \rightarrow X_2^n$.
  - necessary and sufficient condition;
  - n-letter form, incomputable for sufficiently large n;
  - case-by-case.

Single-letter Upper Bound

- Our contribution: a computable single-letter outer bound
- In this talk, we will focus on the sum-rate point
- Steps to our upper bound:
  - n-letter converse for sum rate: incomputable

$$H(U,V) \le \frac{1}{n} I(X_1^n, X_2^n; Y^n) \quad \text{where } X_1^n \to U^n \to V^n \to X_2^n$$

  - A usual way to upper bound the above n-letter mutual information

$$\frac{1}{n} I(X_1^n, X_2^n; Y^n) \le \frac{1}{n} \sum_{i=1}^{n} I(X_{1i}, X_{2i}; Y_i) \le \max_{\substack{p(x_1,x_2) \\ X_1 \to U^n \to V^n \to X_2}} I(X_1, X_2; Y)$$

  - $B \triangleq \{p(x_1,x_2) : X_1 \to U^n \to V^n \to X_2\}$. If $B' \supseteq B$, then

$$\max_{B} \; I(X_1, X_2; Y) \le \max_{B'} \; I(X_1, X_2; Y)$$

  - Find a larger set B', i.e., find a necessary condition for $X_1 \to U^n \to V^n \to X_2$.

Necessary Condition for the Markov Chain

- Need a single-letter necessary condition on $p(x_1,x_2)$ for $X_1 \to U^n \to V^n \to X_2$.
- Intuitively, using the data processing inequality: if $X_1 \to U^n \to V^n \to X_2$

$$I(X_1;X_2) \leq I(U^n;V^n) = nI(U;V)$$

  and therefore

$$B' = \{p(x_1,x_2) : I(X_1;X_2) \leq nI(U;V)\}$$

  is a larger set.
- This becomes trivial when n is large, as all $p(x_1,x_2) \in B'$.
- Data processing inequality limits correlation, and limits feasible set of probability distributions.
- Usual data processing inequality seems useless when $n \to \infty$.
- We need a data processing inequality that will be useful when $n \to \infty$.
- We need a new data processing inequality on new measures of correlation.

New Measures of Correlation

- For any joint distribution matrix P_{XY} with diagonal marginal distributions P_X and P_Y, there is a one-to-one transformation
$$\tilde{P}_{XY} = P_X^{-\frac{1}{2}} P_{XY} P_Y^{-\frac{1}{2}}$$
- Spectral properties of $\tilde{P}_{XY}$
$$\tilde{P}_{XY} = \mathbf{W}\Lambda\mathbf{Z}^T = p_X^{\frac{1}{2}}(p_Y^{\frac{1}{2}})^T + \sum_{i=2}^{l} \lambda_i \mathbf{w}_i \mathbf{z}_i^T$$
where $\mathbf{w}_1 = p_X^{\frac{1}{2}} = P_X^{\frac{1}{2}}\mathbf{e}$, $\mathbf{z}_1 = p_Y^{\frac{1}{2}} = P_X^{\frac{1}{2}}\mathbf{e}$, and $\lambda_1 = 1 \geq \lambda_2 \geq \cdots \geq \lambda_l \geq 0$.
- $\lambda_2, \ldots, \lambda_l$ are measures of correlation.
  - $(\lambda_2, \ldots, \lambda_l) = \begin{cases} (1, \ldots, 1) & X = Y \\ \ldots & \text{correlated} \\ (0, \ldots, 0) & X \perp Y \end{cases}$
  - Witsenhausen showed that common data $\Leftrightarrow \lambda_2 = 1$.

A New Data Processing Inequality

- If $X \to Y \to Z$

$$\tilde{P}_{XZ} = \tilde{P}_{XY}\tilde{P}_{YZ}$$

  and

$$\tilde{P}_{XZ} - p_X^{\frac{1}{2}}(p_Z^{\frac{1}{2}})^T = \left(\tilde{P}_{XY} - p_X^{\frac{1}{2}}(p_Y^{\frac{1}{2}})^T\right)\left(\tilde{P}_{YZ} - p_Y^{\frac{1}{2}}(p_Z^{\frac{1}{2}})^T\right)$$

- For matrices A and B, $\lambda_i(AB) \leq \lambda_i(A)\lambda_1(B)$.

- New Data Processing Inequality: If $X \to Y \to Z$, then

$$\lambda_i(\tilde{P}_{XZ}) \leq \lambda_i(\tilde{P}_{XY})\lambda_2(\tilde{P}_{YZ}) \leq \lambda_i(\tilde{P}_{XY}), \quad i = 2, \ldots, \mathrm{rank}(\tilde{P}_{XZ})$$

- Processing from Y to Z scales correlation measures $\lambda_i(\tilde{P}_{XY})$ down by a factor less than $\lambda_2(\tilde{P}_{YZ})$.

- Our data processing inequality versus usual data processing inequality: if $X \to Y \to Z$

$$\lambda_i(\tilde{P}_{XZ}) \leq \lambda_i(\tilde{P}_{XY}) \qquad \text{versus} \qquad I(X;Z) \leq I(X;Y)$$

I.i.d. Sequences

- If (U^n, V^n), i.i.d., then

$$P_{U^nV^n} = \underbrace{P_{UV} \otimes P_{UV} \otimes \cdots \otimes P_{UV}}_{n} = P_{UV}^{\otimes n}$$

  and correspondingly,

$$\tilde{P}_{U^nV^n} = \tilde{P}_{UV}^{\otimes n}$$

- Applying SVD to $\tilde{P}_{U^nV^n}$,

$$\tilde{P}_{U^nV^n} = \mathbf{W}_n \Lambda_n \mathbf{Z}_n^T$$

  Then, $\mathbf{W}_n = \mathbf{W}^{\otimes n}$, $\Lambda_n = \Lambda^{\otimes n}$ and $\mathbf{Z}_n = \mathbf{Z}^{\otimes n}$.

- The ordered singular values of $\tilde{P}_{U^nV^n}$ are

$$1 \geq \underbrace{\lambda_2(\tilde{P}_{U^nV^n}) = \cdots = \lambda_{n+1}(\tilde{P}_{U^nV^n})}_{\substack{\shortparallel \\ \lambda_2(\tilde{P}_{UV})}} \geq \lambda_{n+2}(\tilde{P}_{U^nV^n}) \geq \cdots$$

A Necessary Condition

- If $X_1 \to U^n \to V^n \to X_2$, then (our new data processing inequality)

$$\lambda_i(\tilde{P}_{X_1X_2}) \le \lambda_2(\tilde{P}_{X_1U^n})\lambda_i(\tilde{P}_{U^nV^n})\lambda_2(\tilde{P}_{V^nX_2}), \quad i = 2, \ldots, \min(|X_1|, |X_2|)$$

- From our previous discussion, we have
  - $\lambda_i(\tilde{P}_{U^nV^n}) \le \lambda_2(\tilde{P}_{UV})$, for any $i \ge 2$. (i.i.d. sequence)
  - $\lambda_2(\tilde{P}_{X_1U^n}) \le 1$. (property of $\tilde{P}$)

- Therefore, if $X_1 \to U^n \to V^n \to X_2$, then

$$\lambda_i(\tilde{P}_{X_1X_2}) \le \lambda_2(\tilde{P}_{UV}), \quad i = 2, \ldots, \min(|X_1|, |X_2|)$$

- Compare this with the usual data processing inequality: $I(X_1;X_2) \le nI(U;V)$.

- Now we have, $B = \{p(x_1,x_2) : X_1 \to U^n \to V^n \to X_2\}$

$$B' = \{p(x_1,x_2) : \lambda_i(\tilde{P}_{X_1X_2}) \le \lambda_2(\tilde{P}_{UV}), \quad i = 2, \ldots, \min(|X_1|, |X_2|)$$

  and

$$\max_{B} \quad I(X_1,X_2;Y) \le \max_{B'} \quad I(X_1,X_2;Y)$$

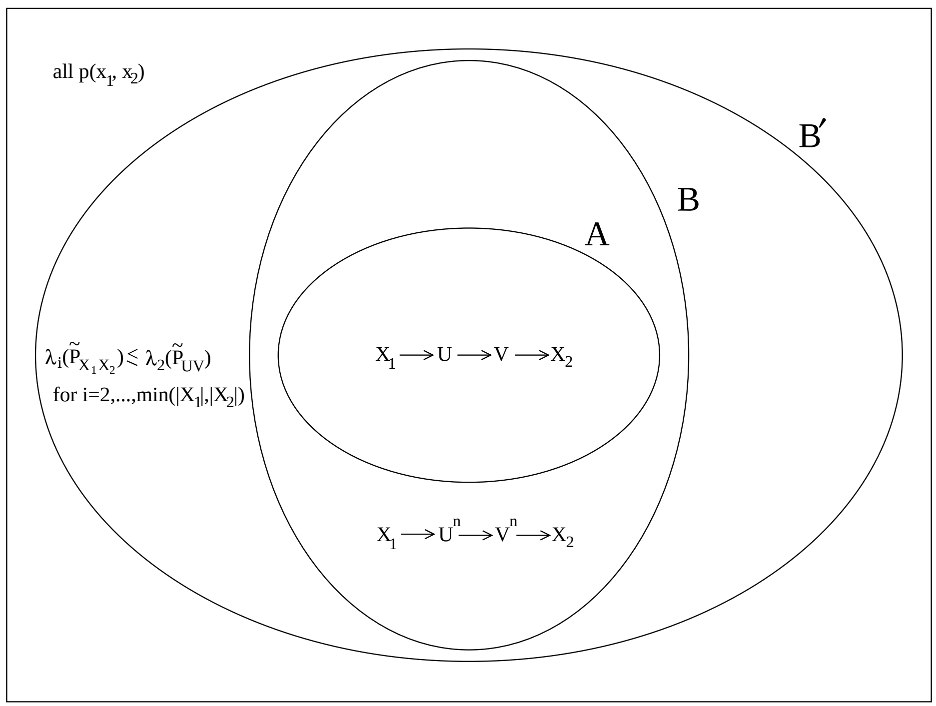
Our Main Result

- If a pair of i.i.d. sources (U,V) with joint distribution $p(u,v)$ can be transmitted reliably through a discrete, memoryless, multiple access channel characterized by $p(y|x_1,x_2)$, then

$$H(U,V) < I(X_1,X_2;Y)$$

for some (X_1,X_2) satisfying

$$\lambda_i(\tilde{P}_{X_1X_2}) \leq \lambda_2(\tilde{P}_{UV}), \quad i = 2,\ldots,\min(|X_1|,|X_2|).$$



Some Examples: The Channel

- Channel inputs X_1 and X_2 and output Y are all binary, and $p(y|x_1,x_2)$ is

$Y\backslash X_1X_2$	11	10	01	00
1	1	1/2	1/2	0
0	0	1/2	1/2	1

- Trivial upper bound

$$C_0 \triangleq \max_{p(x_1,x_2)} I(X_1,X_2;Y) = 1$$

- Our upper bound

$$C_1 \triangleq \max_{\substack{p(x_1,x_2) \\ \lambda_2(\tilde{P}_{X_1X_2}) \le \lambda_2(\tilde{P}_{UV})}} I(X_1,X_2;Y)$$

- The sub-optimal single-letter achievability result by Cover, El Gamal and Salehi

$$C_2 \triangleq \max_{\substack{p(x_1,x_2) \\ X_1 \to U \to V \to X_2}} I(X_1,X_2;Y)$$

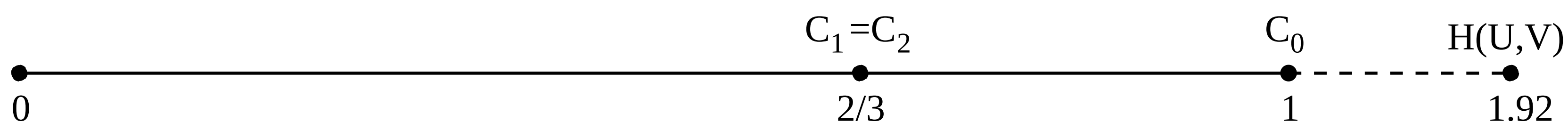
- We have: $C_0 \ge C_1 \ge C_2$

Example 1

- Joint source distribution $p(u,v)$ is

$U \backslash V$	1	0
1	$1/3$	$1/6$
0	$1/6$	$1/3$

- Trivial upper bound shows that reliable transmission is impossible.

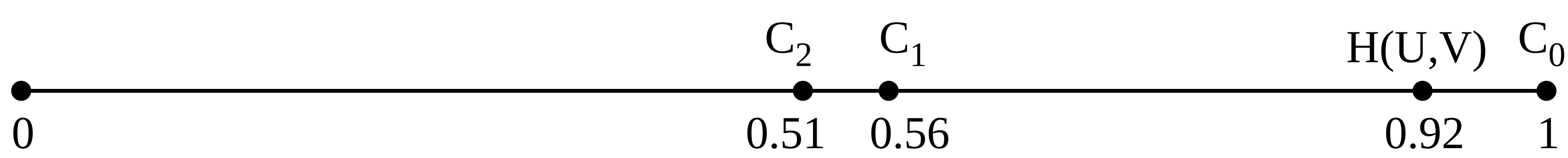


Example 2

- Joint source distribution $p(u,v)$ is

$U \backslash V$	1	0
1	0	0.1
0	0.1	0.8

- Our bound shows that reliable transmission is impossible.
- The trivial upper bound fails to show whether reliable transmission is possible or not.

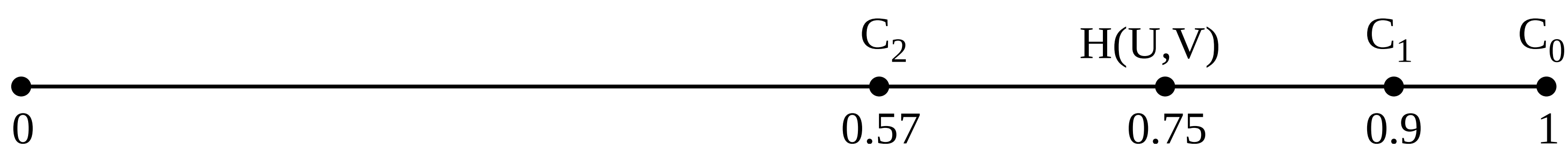


Example 3

- Joint source distribution $p(u,v)$ is

$U \backslash V$	1	0
1	0	0.85
0	0.1	0.05

- Both our upper bound and the achievability expression of Cover, El Gamal and Salehi fail to show whether reliable transmission is possible or not.



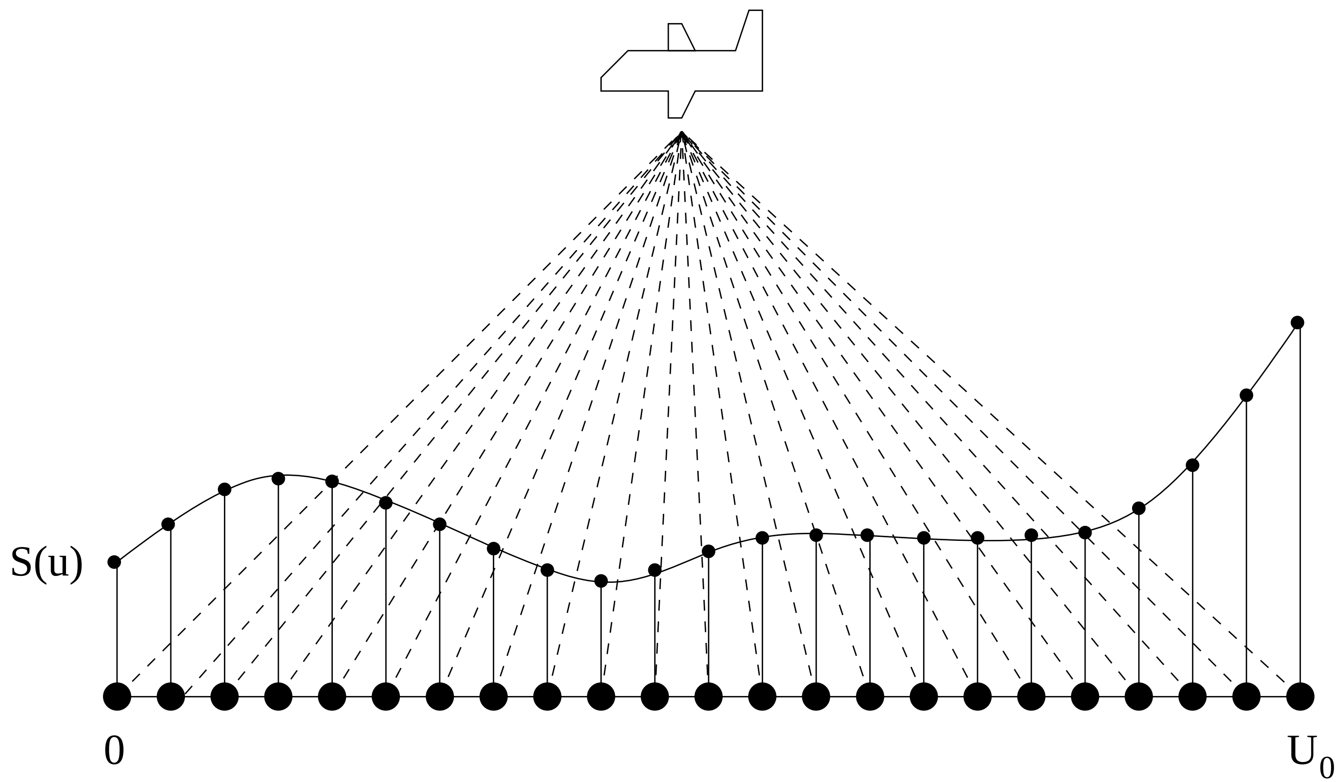
Conclusions for the First Part

- The capacity region of the MAC with correlated sources remains an open problem.
- The current converse is in n-letter form; and is incomputable.
- We proposed a new data processing inequality.
- By using this inequality, we obtained a new outer bound for the capacity region.
- Our outer bound is loose but is in single-letter and computable.

N-user Multiple Access Channel with Correlated Data

- Underlying one-dimensional spatial Gaussian random process $S(u), 0 \leq u \leq U_0$
- N sensors equally placed on $[0, U_0]$, sampling $S(u)$ without noise
- Collector node wishes to reconstruct the entire random process with minimum MSE

$$d(s(u), \hat{s}(u)) = \frac{1}{U_0} \int_0^{U_0} (s(u) - \hat{s}(u))^2 du$$



Channel Model

- All nodes are equipped with one transmit and one receive antenna
- All nodes hear a linear combination of the signals transmitted by all other nodes

$$Y_i = \sum_{j=0, j \neq i}^{N} h_{ji} X_j + Z_i, \qquad i = 0, 1, 2, \cdots, N$$

  where

  - $\{Z_i\}_{i=0}^{N}$ is a vector of $N+1$ random variables, i.i.d. $\mathcal{N}(0,1)$
  - Channel gains h_{ij} are bounded, i.e.

$$\overline{h_l} \leq h_{ij} \leq \overline{h_u}, \qquad i, j = 0, 1, \cdots, N$$

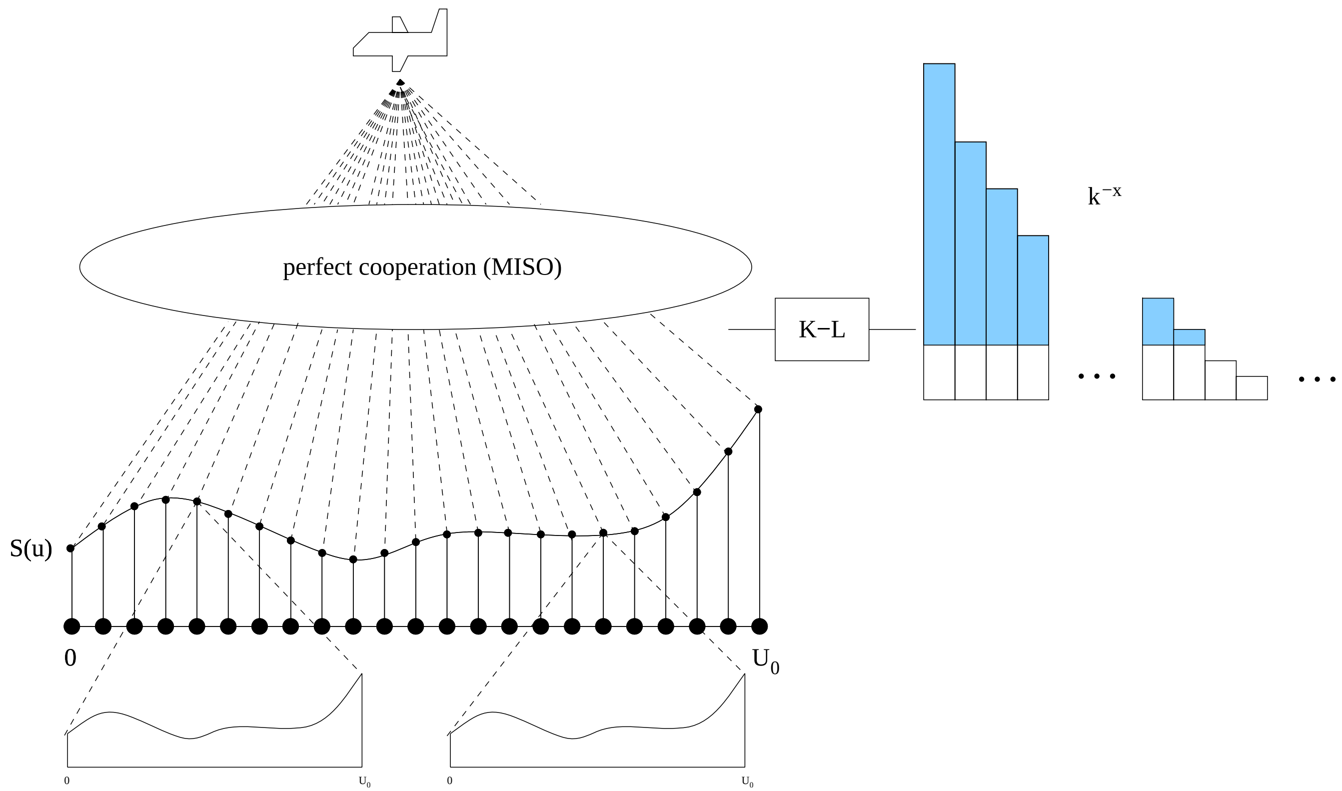
- All sensors share a sum power constraint $P(N)$, e.g., $P(N) = NP_{\text{ind}}$ and $P(N) = P_{\text{tot}}$
- Multiple access channel with potential cooperation and (imperfect) feedback

The Nature of the Problem

- Finding minimum distortion is difficult: joint source-channel coding problem with distortion
  - Source coding: indirect observation, distributed correlated data
  - Channel coding: MAC with potential cooperation and (imperfect) feedback
- **Goal:** Understanding the order optimal performance of such network
  - Minimum achievable expected distortion as a function of $P(N)$?
  - Rate at which distortion goes to zero with N?
  - Order-optimal achievability scheme?
- Approach:
  - Lower bound on distortion
    - Perfect cooperation: sensors know $S(u)$; point-to-point joint source-channel coding
  - Upper bound on distortion
    - Separation-based scheme:
      - Achievable rate of cooperative MAC [Gastpar & Vetterli, 05]
      - Achievable distributed rate-distortion point [Flynn & Gray, 87]

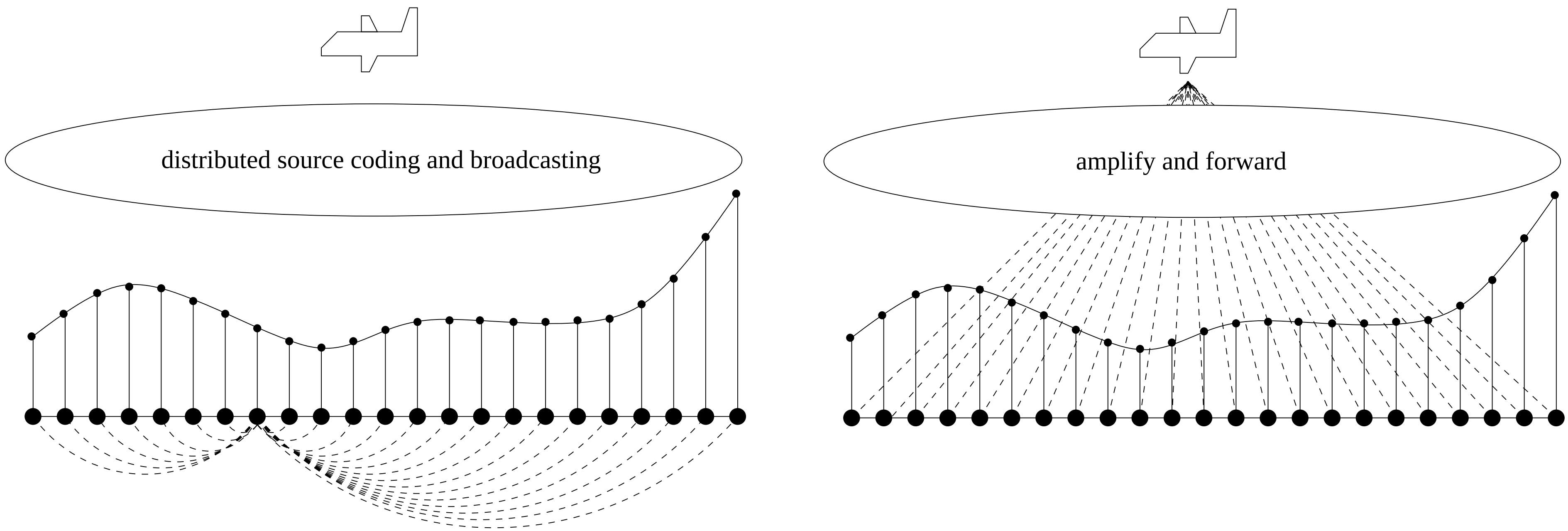
Lower Bound: Perfect Cooperation

- All sensors know the entire random process perfectly
- Rate-distortion for a Gaussian random process (inverse waterfilling over the Karhunen-Loeve eigenvalues of the random process) & capacity of $N \times 1$ point-to-point MISO channel



Upper Bound based on an Achievable Scheme

- Separation-based achievability scheme
- Source coding [Flynn & Gray, 87]:
  - Each sensor performs single-user rate-distortion coding
  - Slepian-Wolf distributed lossless encoding on the index of the rate-distortion code
- Channel coding—amplify and forward [Gastpar & Vetterli, 05]:
  - All sensors take turns broadcasting their own data, using single-user channel coding
  - All sensors amplify and forward to the collector node



The Class A of Gaussian Random Processes

- Gaussian random process with autocorrelation $K(u,s)$
  - Mercer's theorem:

$$K(u,s) = \sum_{k=0}^{\infty} \lambda_k \phi_k(u) \phi_k(s)$$

  - λ_k are the eigenvalues, $\phi_k(t)$ are the orthonormal eigenfunctions
- Eigenvalues:

$$\lambda_k \sim \frac{1}{k^x}, \quad x > 1$$

- Autocorrelation function and the eigenfunctions satisfy some Lipschitz-like conditions

$$|K(u_1,s_1) - K(u_2,s_2)| \le B\left(\sqrt{(u_1-u_2)^2 + (s_1-s_2)^2}\right)^{\alpha} \quad \forall u_1, s_1, u_2, s_2 \in [0, U_0]$$

$$|\phi_k(s_1) - \phi_k(s_2)| \le B_3(k+B_4)^{\tau}|s_1 - s_2|^{\gamma} \quad \forall u \in [0, U_0]$$

$$|K(u,s_1)\phi_k(s_1) - K(u,s_2)\phi_k(s_2)| \le B_2(k+B_1)^{\tau}|s_1 - s_2|^{\beta} \quad \forall s_1, s_2 \in [0, U_0]$$

- Parameters $x > 1$, $1/2 < \alpha \le 1$, $0 \le \beta, \gamma \le 1$, $\tau \ge 0$
- E.g., Gauss-Markov process: $K(u,s) = \frac{\sigma^2}{2\eta} e^{\eta|u-s|}$ satisfies above with $x = 2, \alpha = \beta = \tau = \gamma = 1$.

Results for Gaussian Processes in A

- Lower bound:
  - Rate-distortion for a Gaussian random process with $\lambda_k \sim k^{-x}$

$$D_l(R) = \Theta\left(\frac{1}{R^{x-1}}\right), \quad R > 0$$

  - Capacity of $N \times 1$ point-to-point MISO channel [Telatar, 99]

$$C_l(P(N)) = \Theta\left(\log(NP(N))\right), \quad P(N) > N^{-1}$$

  - Lower bound on minimum achievable distortion: $D_l(C_l)$
- Upper bound:
  - Distributed achievable rate-distortion

$$D_u(R) = \Theta\left(\frac{1}{R^{x-1}}\right), \quad 0 < R < N^{\min\left(\frac{\gamma}{2\tau}, \frac{2\alpha-1}{2(x-1)}, \frac{\beta}{x+\tau+1}\right)}$$

  - Achievable sum rate using amplify and forward

$$C_u(P(N)) = \Theta\left(\log(NP(N))\right), \quad P(N) > N^{-\frac{1}{2}}$$

  - Upper bound on minimum achievable distortion: $D_u(C_u)$

Comparison of the Lower and Upper Bounds for Gaussian Processes in A

- Large sum power
  - $P(N)$ is larger than $\frac{e^{N^{\min\left(\frac{\gamma}{2\tau}, \frac{2\alpha-1}{2(x-1)}, \frac{\beta}{x+\tau+1}\right)}}}{N}$
  - Lower bound: $\Theta\left(\frac{1}{(\log(NP(N)))^{x-1}}\right)$
  - Upper bound: we do not have an explicit upper bound
  - Whether we can develop an upper bound that meets the lower bound remains open
  - However, this region of sum power constraints is not of practical interest

Comparison of the Lower and Upper Bounds for Gaussian Processes in A

- Medium sum power
  - $P(N)$ is in the wide range of $N^{-\frac{1}{2}}$ to $\frac{e^{N^{\min\left(\frac{\gamma}{2\tau}, \frac{2\alpha-1}{2(x-1)}, \frac{\beta}{x+\tau+1}\right)}}}{N}$
  - Our lower and upper bounds meet
  - Minimum achievable expected distortion is

$$\Theta\left(\frac{1}{(\log(NP(N)))^{x-1}}\right)$$

  - Order-optimal achievability scheme: separation-based scheme
    * Distributed rate-distortion code [Flynn & Gray, 87]
    * Single-user channel code with amplify and forward cooperation [Gastpar & Vetterli, 05]
  - The practically interesting cases: $P(N) = P_{\text{tot}}$ and $P(N) = NP_{\text{ind}}$
    * For both cases, the distortion decreases to zero at the rate of

$$\frac{1}{(\log N)^{x-1}}$$

    * Therefore, we would prefer $P(N) = P_{\text{tot}}$ over $P(N) = NP_{\text{ind}}$
    * In fact, we would prefer $P(N) = N^{-1/3}$!

Comparison of the Lower and Upper Bounds for Gaussian Processes in A

- Small sum power
  - $P(N)$ ranges from N^{-1} to $N^{-\frac{1}{2}}$
  - Our lower and upper bounds do not meet
    * Lower bound decreases to zero as $\Theta\left(\frac{1}{(\log N)^{x-1}}\right)$
    * Upper bound is a non-zero constant
  - Main discrepancy: sum rate of a cooperative MAC with feedback

$$C_l(P(N)) = \Theta\left(\log(NP(N))\right), \quad P(N) > N^{-1}$$

$$C_u(P(N)) = \Theta\left(\log(NP(N))\right), \quad P(N) > N^{-\frac{1}{2}}$$

  - This region is of practical interest
    * Sum power constraint is low
    * Potentially good performance based on lower bound on the distortion
  - More effort is required to find an order-optimal achievability scheme

Comparison of the Lower and Upper Bounds for Gaussian Processes in A

- Very small sum power
  - $P(N)$ is less than N^{-1}
  - Our lower and upper bounds meet
  - Minimum achievable expected distortion: $\Theta(1)$
  - Power consumption is good, but not of practical interest, unacceptable distortion

Conclusions for the Second Part

- We investigated the achievable distortion in dense Gaussian sensor networks.
- We provided lower and upper bounds on the minimum achievable expected distortion.
- For a class of random process, we showed that the lower and upper bounds meet **order-wise**.
- For these Gaussian random process, the order-optimal scheme is separation-based:
  - Distributed rate-distortion coding [Flynn & Gray, 87]
  - Amplify and forward [Gastpar & Vetterli, 05] method for cooperative multiple access
- Interplay between:
  - **external** correlation in the observed data, and
  - **internal** correlation created through cooperation
- Internal correlation created through cooperation converts the MAC into a MISO channel.