On Source-Channel Communication in Networks

Michael Gastpar

Department of EECS
University of California, Berkeley
gastpar@eecs.berkeley.edu

DIMACS: March 17, 2003.

Outline

1. Source-Channel Communication seen from the perspective of the separation theorem

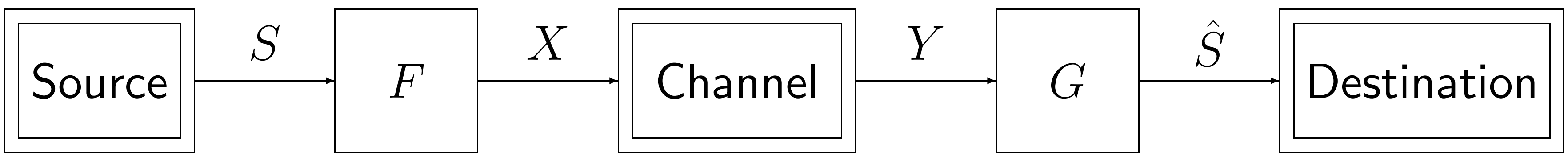
2. Source-Channel Communication seen from the perspective of measure-matching

Acknowledgments

- Gerhard Kramer
- Bixio Rimoldi
- Emre Telatar
- Martin Vetterli

Source-Channel Communication

Consider the transmission of a discrete-time memoryless source across a discrete-time memoryless channel.



$$F(S^n) = X^n \qquad\qquad G(Y^n) = \hat{S}^n$$

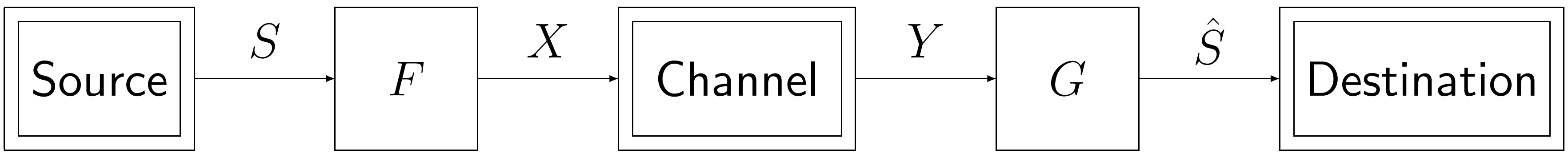
The fundamental trade-off is *cost* versus *distortion*,

$$\begin{aligned} \Delta &= Ed(S^n, \hat{S}^n) \\ \Gamma &= E\rho(X^n) \end{aligned}$$

What is the set of

- *achievable* trade-offs (Γ, Δ)?
- *optimal* trade-offs (Γ, Δ)?

The Separation Theorem



$$F(S^n) = X^n \qquad\qquad G(Y^n) = \hat{S}^n$$

For a fixed source (p_S, d) and a fixed channel $(p_{Y|X}, \rho)$:
A cost-distortion pair (Γ, Δ) is *achievable* if and only if

$$R(\Delta) \ \leq \ C(\Gamma).$$

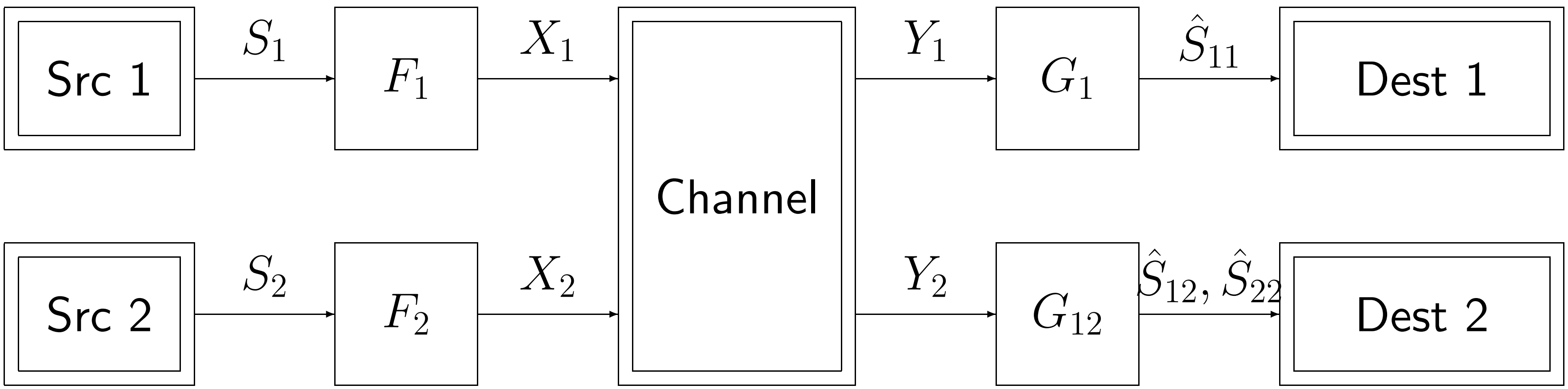
A cost-distortion pair (Γ, Δ) is *optimal* if and only if

$$R(\Delta) \ = \ C(\Gamma),$$

subject to certain technicalities.
Rate-matching: In an optimal communication system, the minimum source rate is *matched* (i.e., equal) to the maximum channel rate.

Source-Channel Communication in Networks

Simple source-channel network:



Trade-off between *cost* (Γ_1, Γ_2) and *distortion* $(\Delta_{11}, \Delta_{12}, \Delta_{22})$.

Achievable cost-distortion tuples? *Optimal* cost-distortion tuples?

For the sketched topology, the (full) answer is unknown.

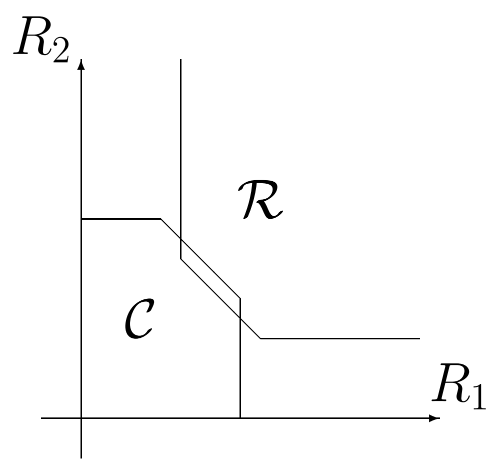
These Trade-offs Are Achievable:

For a fixed network topology and fixed probability distributions and cost/distortion functions:

If a cost-distortion tuple satisfies

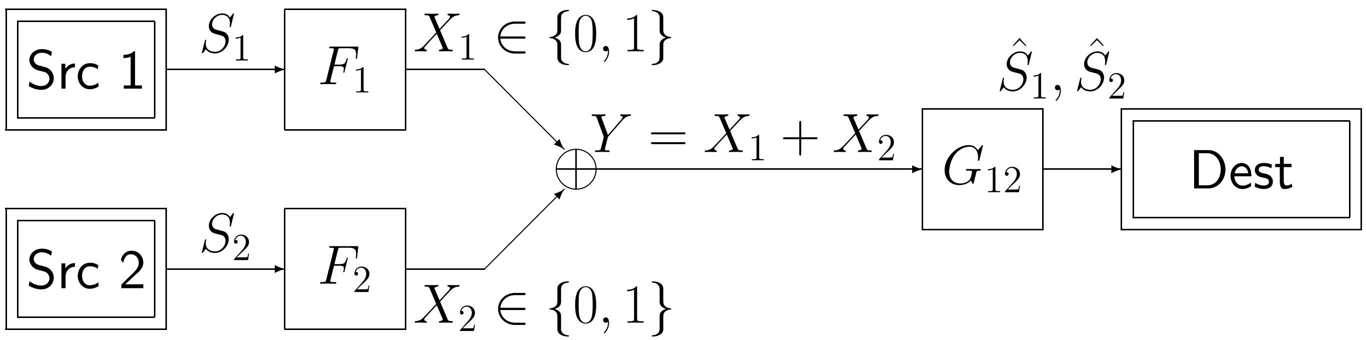
$$\mathcal{R}(\Delta_1, \Delta_2, \ldots) \cap \mathcal{C}(\Gamma_1, \Gamma_2, \ldots) \neq \emptyset,$$

then it is achievable.



When is it optimal?

Example: Multi-access source-channel communication



Capacity region of this channel is contained inside $R_1 + R_2 \leq 1.5$.

Goal: Reconstruct S_1 and S_2 perfectly.

S_1 and S_2 are correlated:

	$S_1 = 0$	$S_1 = 1$
$S_2 = 0$	$1/3$	$1/3$
$S_2 = 1$	0	$1/3$

$R_1 + R_2 \geq \log_2 3 \approx 1.585$.

$\mathcal{R}$ and $\mathcal{C}$ do not intersect.

Yet uncoded transmission works.

This example appears in T. M. Cover, A. El Gamal, M. Salehi, "Multiple access channels with arbitrarily correlated sources." IEEE Trans IT-26, 1980.

So what is capacity?

The capacity region is computed assuming *independent* messages.

In a source-channel context, the underlying sources may be dependent.

MAC example: Allowing arbitrary dependence of the channel inputs, the capacity is $\log_2 3 = 1.585$, "fixing" the example: $\mathcal{R} \cap \mathcal{C} \neq \emptyset$.

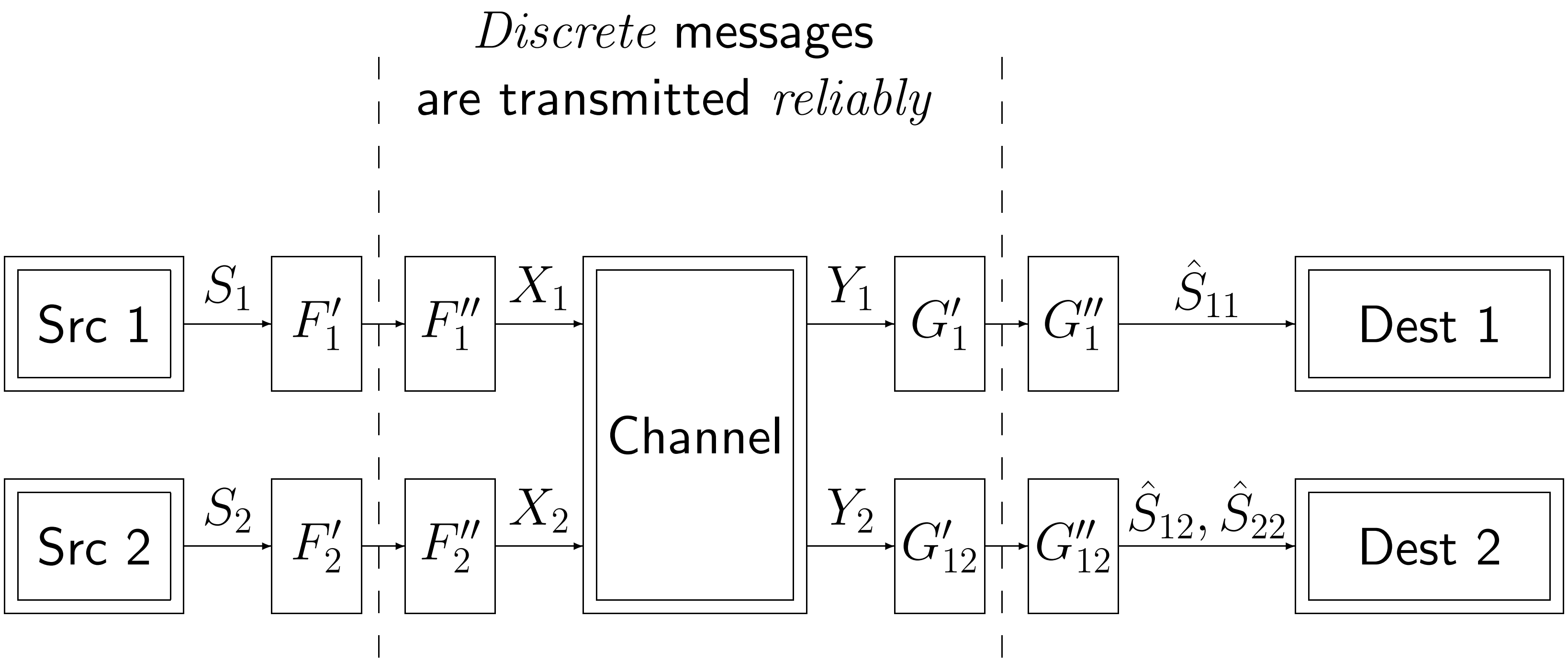
Can we simply redefine capacity appropriately?

Remark: Multi-access with dependent messages is still an open problem.

T. M. Cover, A. El Gamal, M. Salehi, "Multiple access channels with arbitrarily correlated sources." IEEE Trans IT-26, 1980.

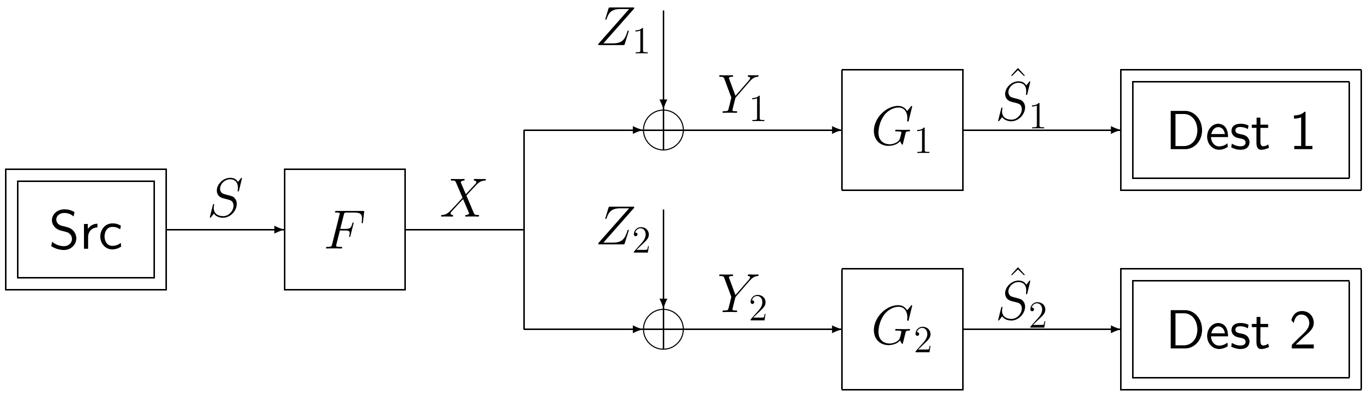
Separation Strategies for Networks

In order to retain a notion of capacity:



What is the best achievable performance for such a system?
— The general answer is unknown.

Example: Broadcast



S, Z_1, Z_2 are i.i.d. Gaussian.

Goal: Minimize the mean-squared errors Δ_1 and Δ_2.

Denote by Δ_1^* and Δ_2^* the single-user minima.

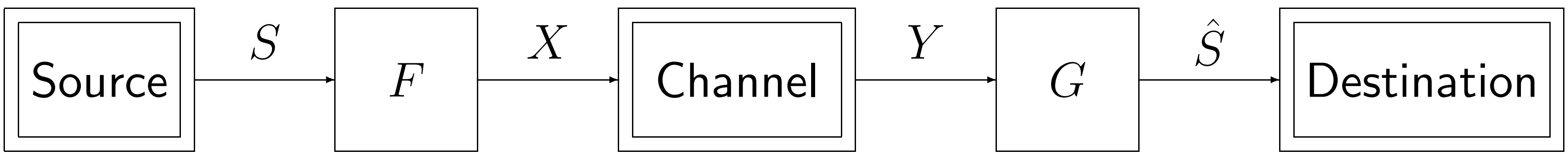
Δ_1^* and Δ_2^* *cannot* be achieved simultaneously by sending messages reliably: The messages disturb one another.

But uncoded transmission achieves Δ_1^* and Δ_2^* simultaneously.

This cannot be fixed by altering the definitions of capacity and/or rate-distortion regions.

Alternative approach

$$E\rho(X^n) \le \Gamma$$



$$F(S^n) = X^n \qquad\qquad G(Y^n) = \hat{S}^n$$

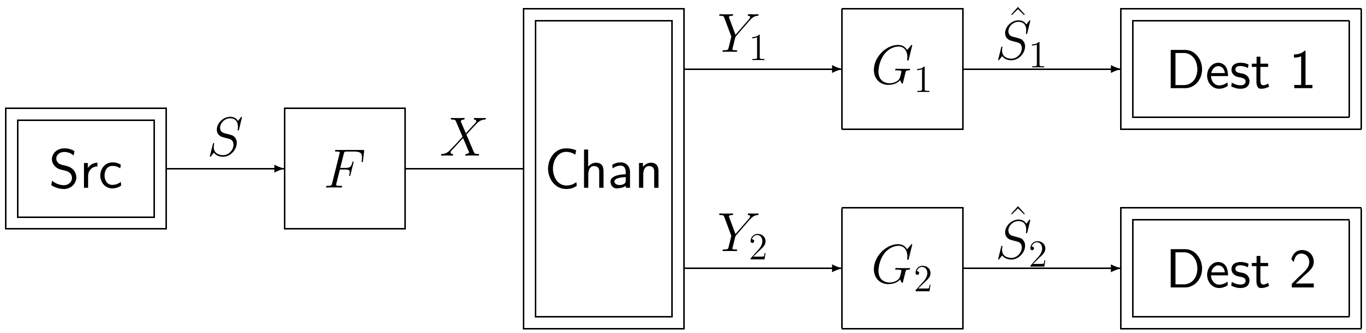
A code (F, G) performs optimally if and only if it satisfies $R(\Delta) = C(\Gamma)$ (subject to certain technical conditions).

Equivalently, a code (F, G) performs optimally if and only if

$$\begin{aligned} \rho(x^n) &= c_1 D(p_{Y^n|x^n} || p_{Y^n}) + \rho_0 \\ d(s^n, \hat{s}^n) &= -c_2 \log_2 p(s^n | \hat{s}^n) + d_0(s) \\ I(S^n; \hat{S}^n) &= I(X^n; Y^n) \end{aligned}$$

We call this the measure-matching conditions.

Single-source Broadcast



Measure-matching conditions for single-source broadcast:

If the single-source broadcast communication system satisfies

$$\rho(x) = c_1^{(1)} D(p_{Y_1|x} || p_{Y_1}) + \rho_0^{(1)} = c_1^{(2)} D(p_{Y_2|x} || p_{Y_2}) + \rho_0^{(2)},$$

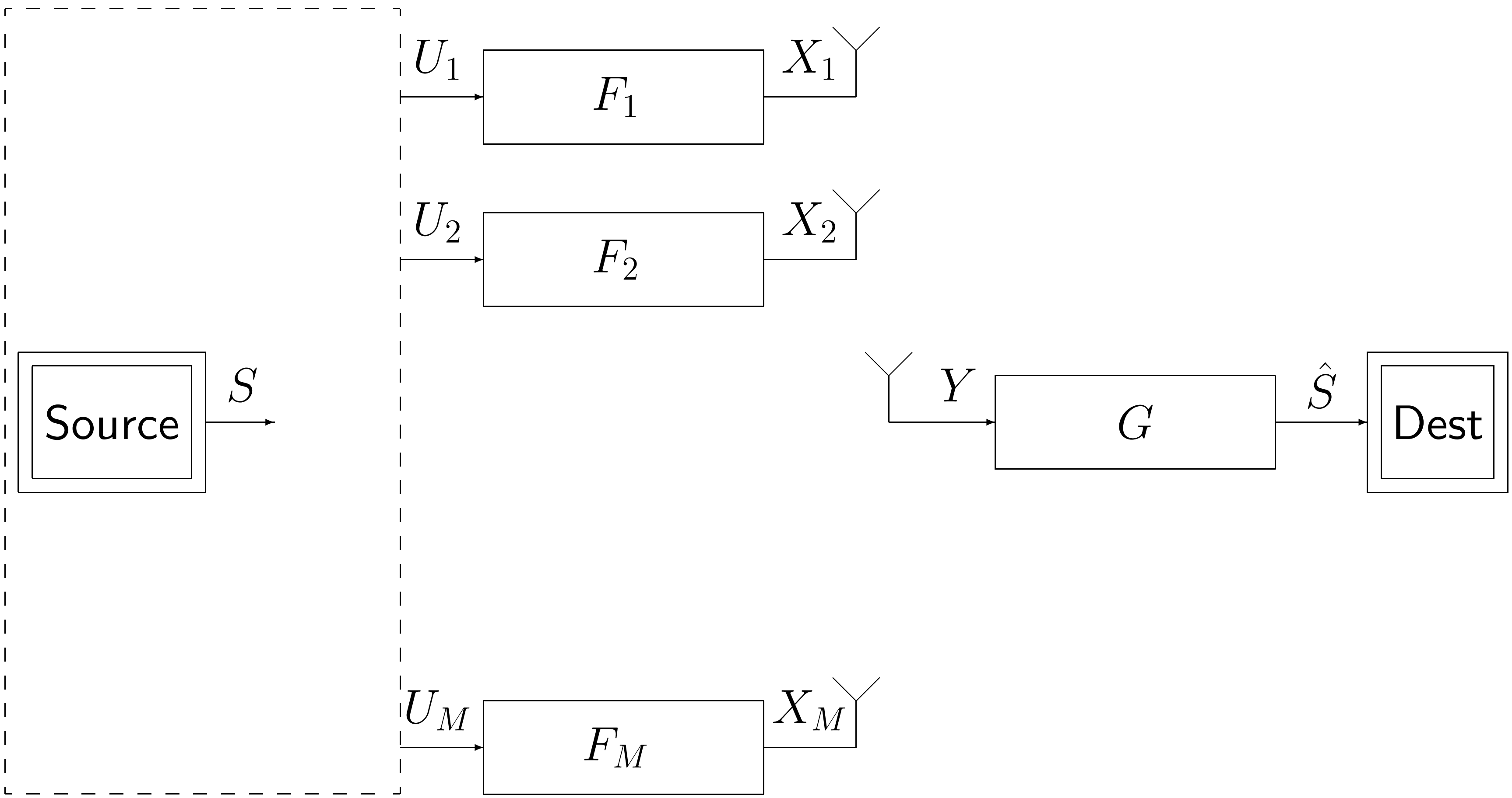
$$\begin{aligned} d_1(s, \hat{s}_1) &= -c_2^{(1)} \log_2 p(s|\hat{s}_1) + d_0^{(1)}(s), \\ d_2(s, \hat{s}_2) &= -c_2^{(2)} \log_2 p(s|\hat{s}_2) + d_0^{(2)}(s), \end{aligned}$$

$$I(X; Y_1) = I(S; \hat{S}_1), \quad \text{and} \quad I(X; Y_2) = I(S; \hat{S}_2),$$

then it performs optimally.

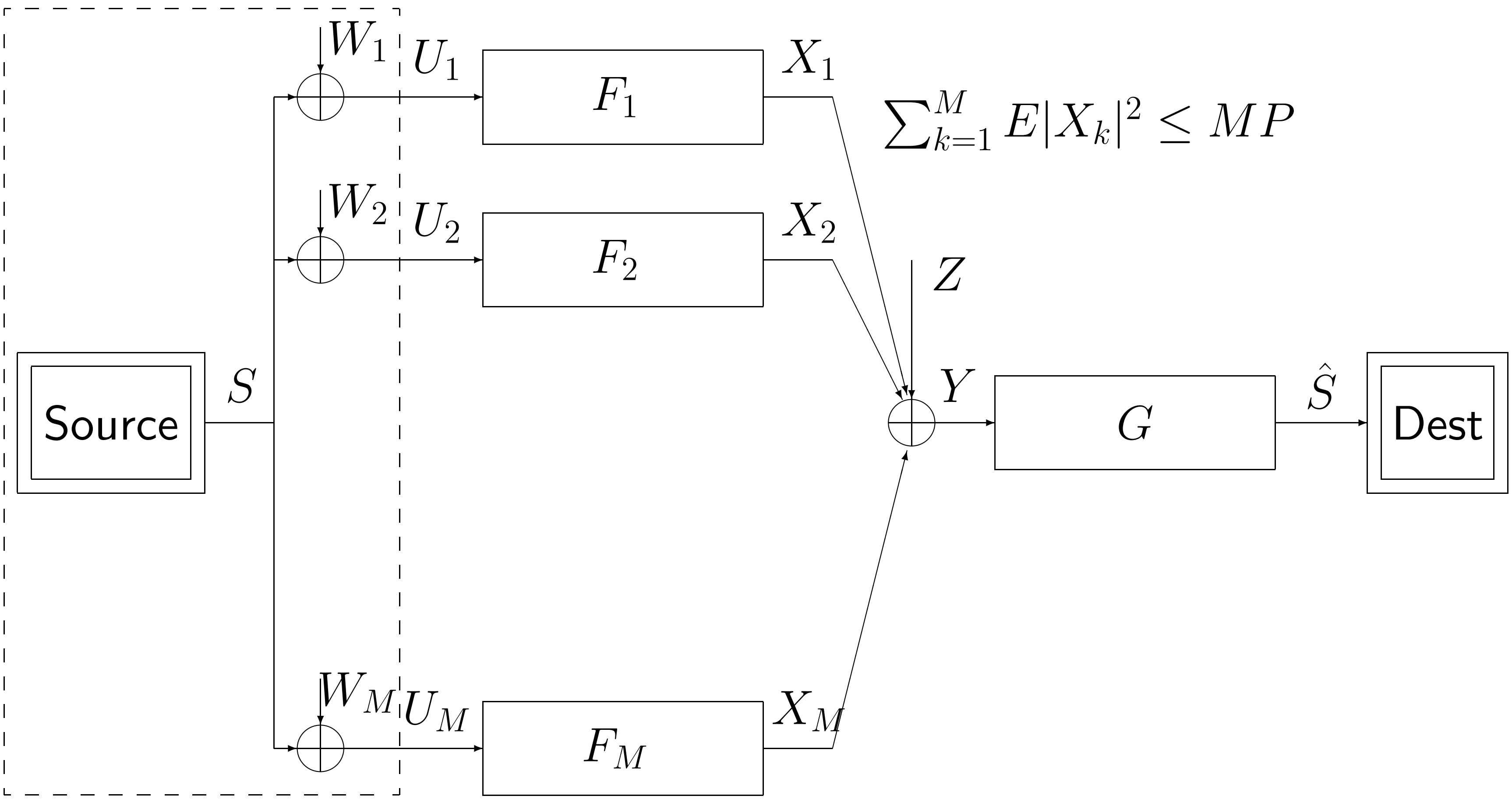
Sensor Network

M wireless sensors measure physical phenomena characterized by S.



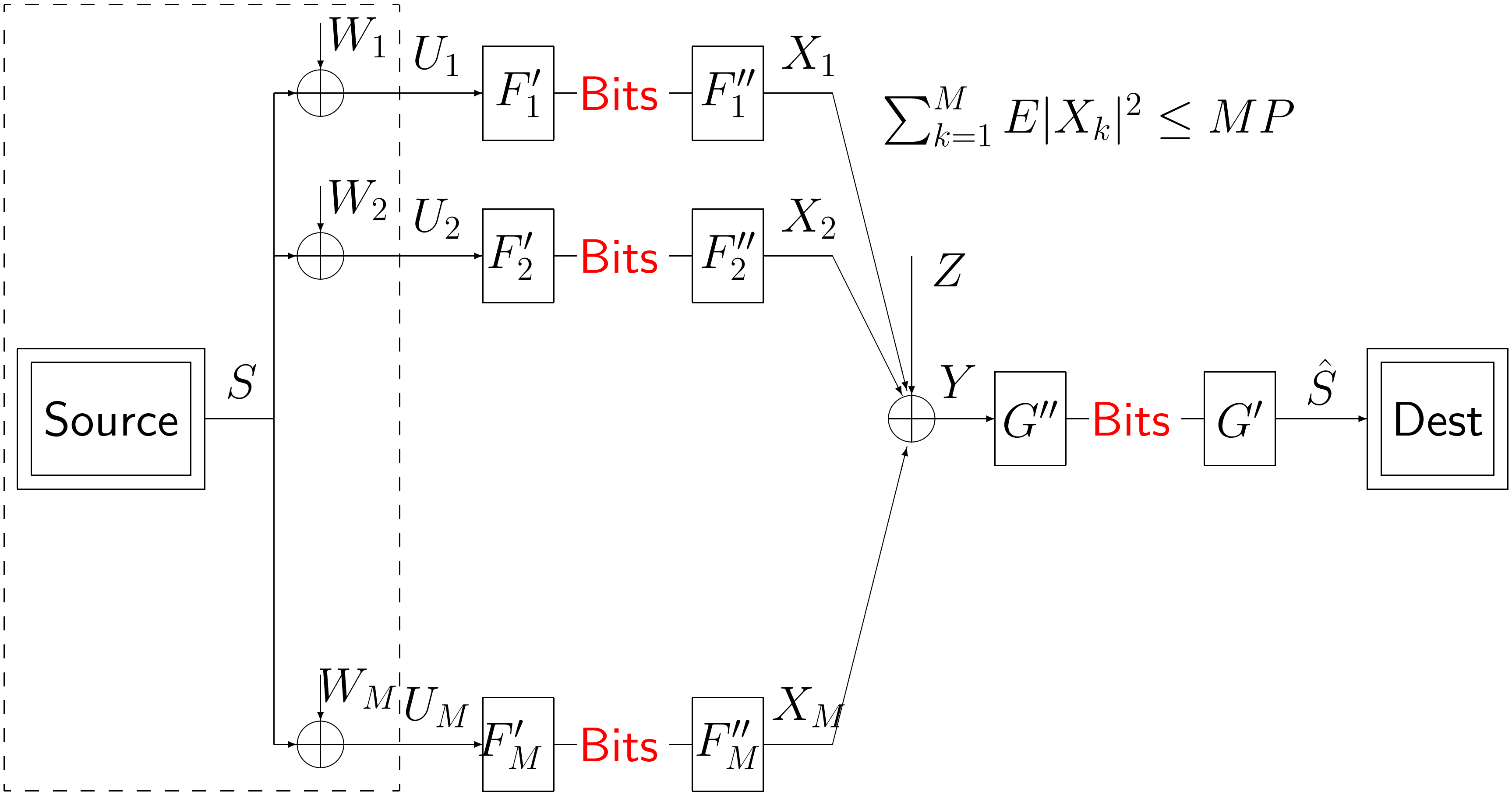
Gaussian Sensor Network

The observations $U_1, U_2, \ldots, U_k$ are noisy versions of S.



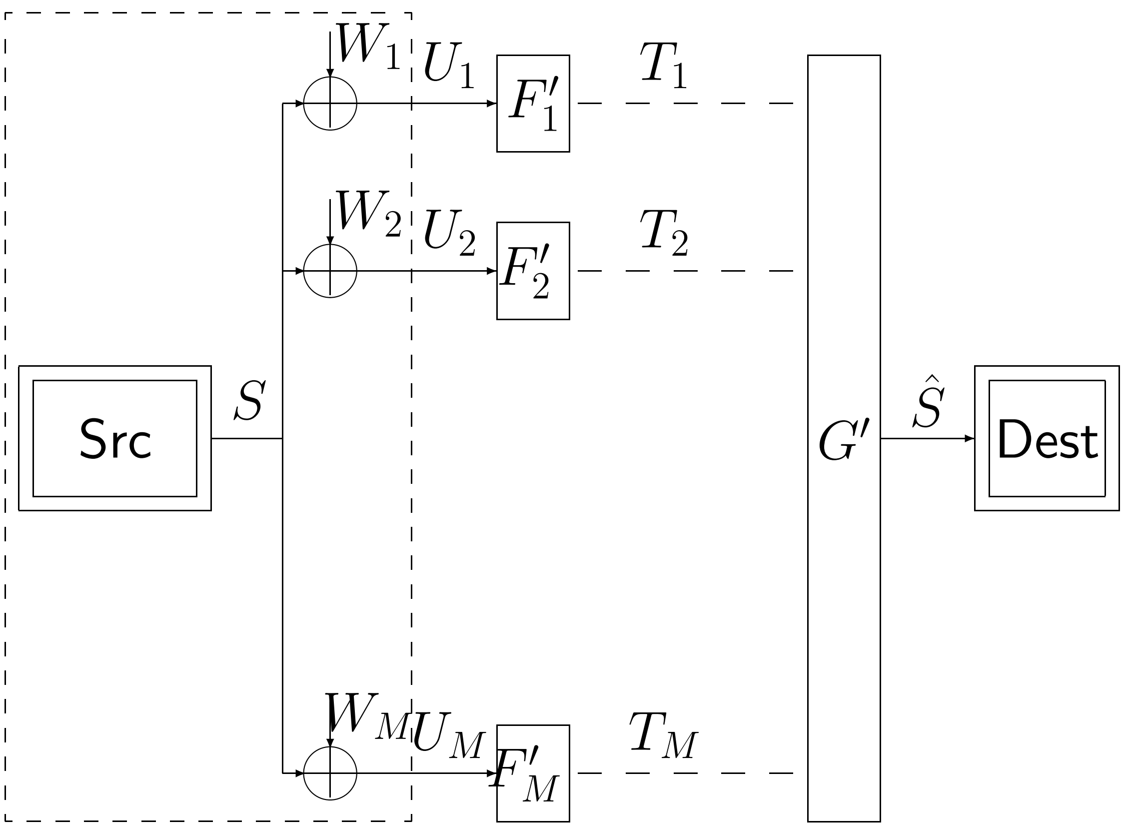
Gaussian Sensor Network: Bits

Consider the following communication strategy:



Gaussian Sensor Network: Bits (1/2)

Source coding part. CEO problem. See Berger, Zhang, Viswanathan (1996); Viswanathan and Berger (1997); Oohama (1998).



$$S \sim \mathcal{N}_c(0, \sigma_S^2)$$

and for $k = 1, \ldots, M$,

$$W_k \sim \mathcal{N}_c(0, \sigma_W^2)$$

For large R_{tot}, the behavior is

$$D_{CEO}(R_{tot}) = \frac{\sigma_W^2}{R_{tot}}.$$

Gaussian Sensor Network: Bits (2/2)

Channel coding part. Additive white Gaussian multi-access channel:

$$R_{sum} \leq \log_2 \left(1 + \frac{MP}{\sigma_Z^2} \right).$$

However, the codewords may be dependent. Therefore, the sum rate may be up to

$$R_{sum} \leq \log_2 \left(1 + \frac{M^2 P}{\sigma_Z^2} \right).$$

$\star \quad \star \quad \star$

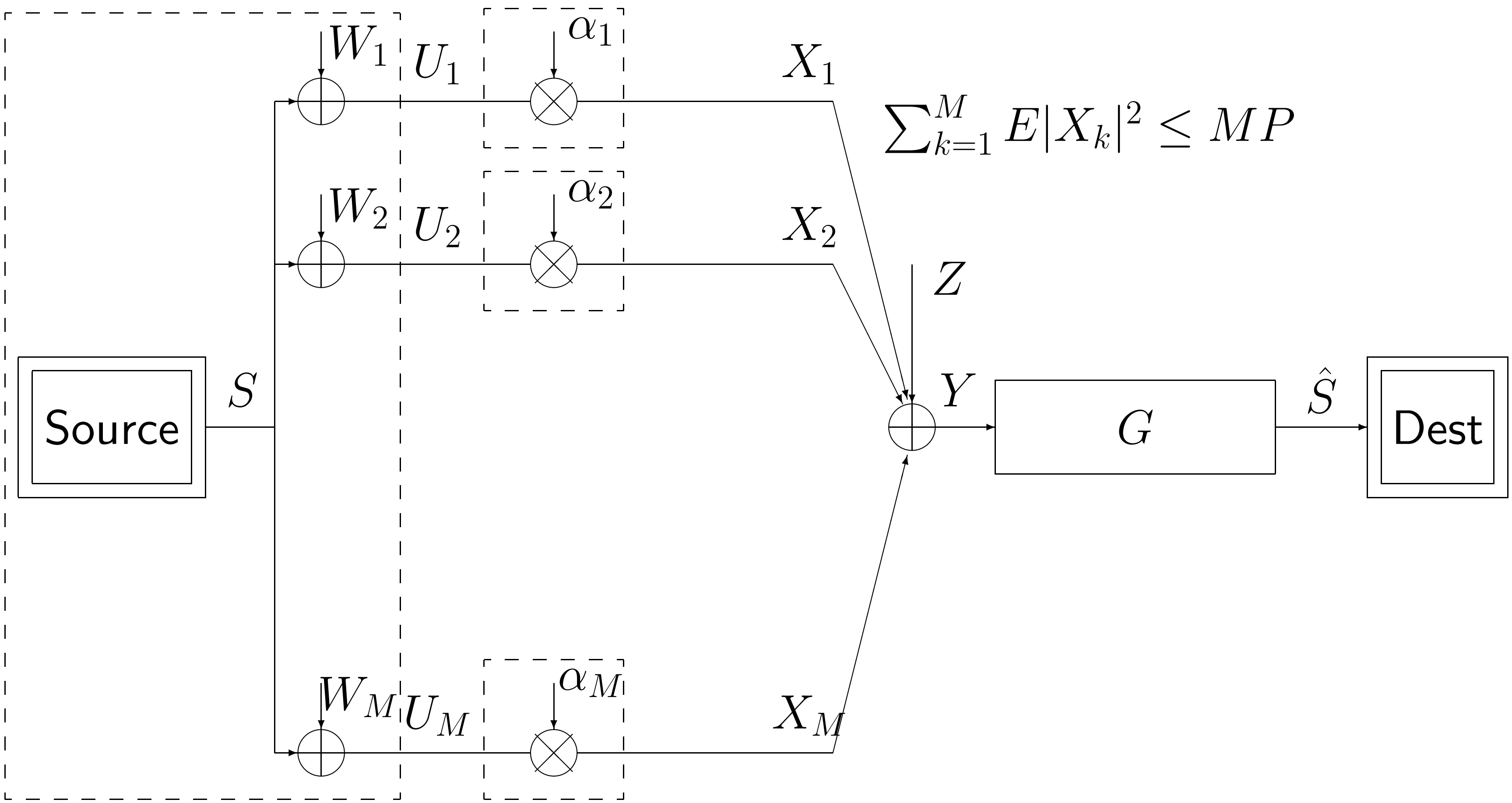
Hence, the distortion for a system that satisfies the rate-matching condition is at least

$$D_{rm}(M) \geq \frac{\sigma_W^2}{\log_2 \left(1 + \frac{M^2 P}{\sigma_Z^2} \right)}$$

Is this optimal?

Gaussian Sensor Network: Uncoded transmission

Consider instead the following "coding" strategy:



Gaussian Sensor Network: Uncoded transmission

Strategy: The sensors transmit whatever they measure, scaled to their power constraint, without any coding at all.

$$Y[n] = \sqrt{\frac{P}{\sigma_S^2 + \sigma_W^2}} \left(MS[n] + \sum_{k=1}^{M} W_k[n] \right) + Z[n].$$

If the "decoder" is the minimum mean-squared error estimate of S based on Y, the following distortion is incurred:

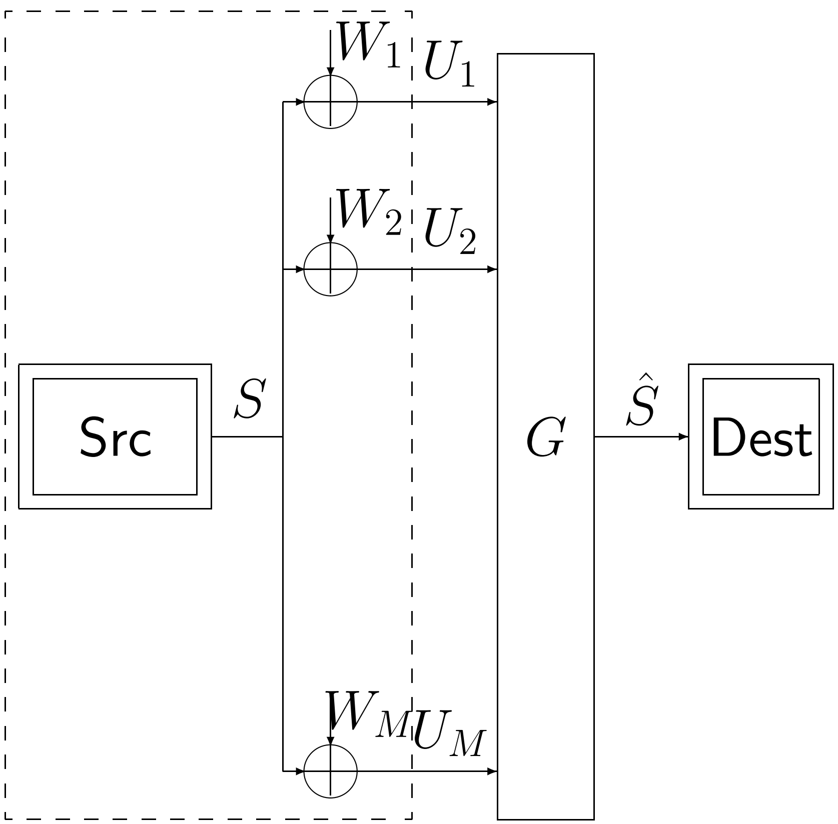
Proposition 1. *Uncoded transmission achieves*

$$D_1(MP) = \frac{\sigma_S^2 \sigma_W^2}{\frac{M^2}{M + (\sigma_Z^2/\sigma_W^2)(\sigma_S^2 + \sigma_W^2)/P} \sigma_S^2 + \sigma_W^2}.$$

This is better than separation ($D_{rm} \propto 1/\log M$). In this sense, uncoded transmission *beats* capacity. Is it optimal?

Gaussian Sensor Network: An outer bound

Suppose the decoder has direct access to $U_1, U_2, \ldots, U_M$.



The smallest distortion for our sensor network *cannot* be smaller than the smallest distortion for the idealization.

$$D_{min,ideal} = \frac{\sigma_S^2 \sigma_W^2}{M\sigma_S^2 + \sigma_W^2}$$

Gaussian Sensor Network: Asymptotic optimum

Rate-matching:

$$D_{rm}(MP) \geq \frac{\sigma_W^2}{\log_2\left(1 + \frac{M^2 P}{\sigma_Z^2}\right)}$$

Uncoded transmission:

$$D_1(MP) = \frac{\sigma_S^2 \sigma_W^2}{\frac{M^2}{M + (\sigma_Z^2/\sigma_W^2)(\sigma_S^2 + \sigma_W^2)/P}\sigma_S^2 + \sigma_W^2}.$$

Proposition 2. *As the number of sensors becomes large, the optimum trade-off is*

$$D(MP) \geq \frac{\sigma_S^2 \sigma_W^2}{M\sigma_S^2 + \sigma_W^2}.$$

Gaussian Sensor Network: Conclusions

Two conclusions from the Gaussian sensor network example:

1. Uncoded transmission is asymptotically optimal.

   - This leads to a general measure-matching condition.

2. Even for finite M, uncoded transmission considerably outperforms the best separation-based coding strategies.

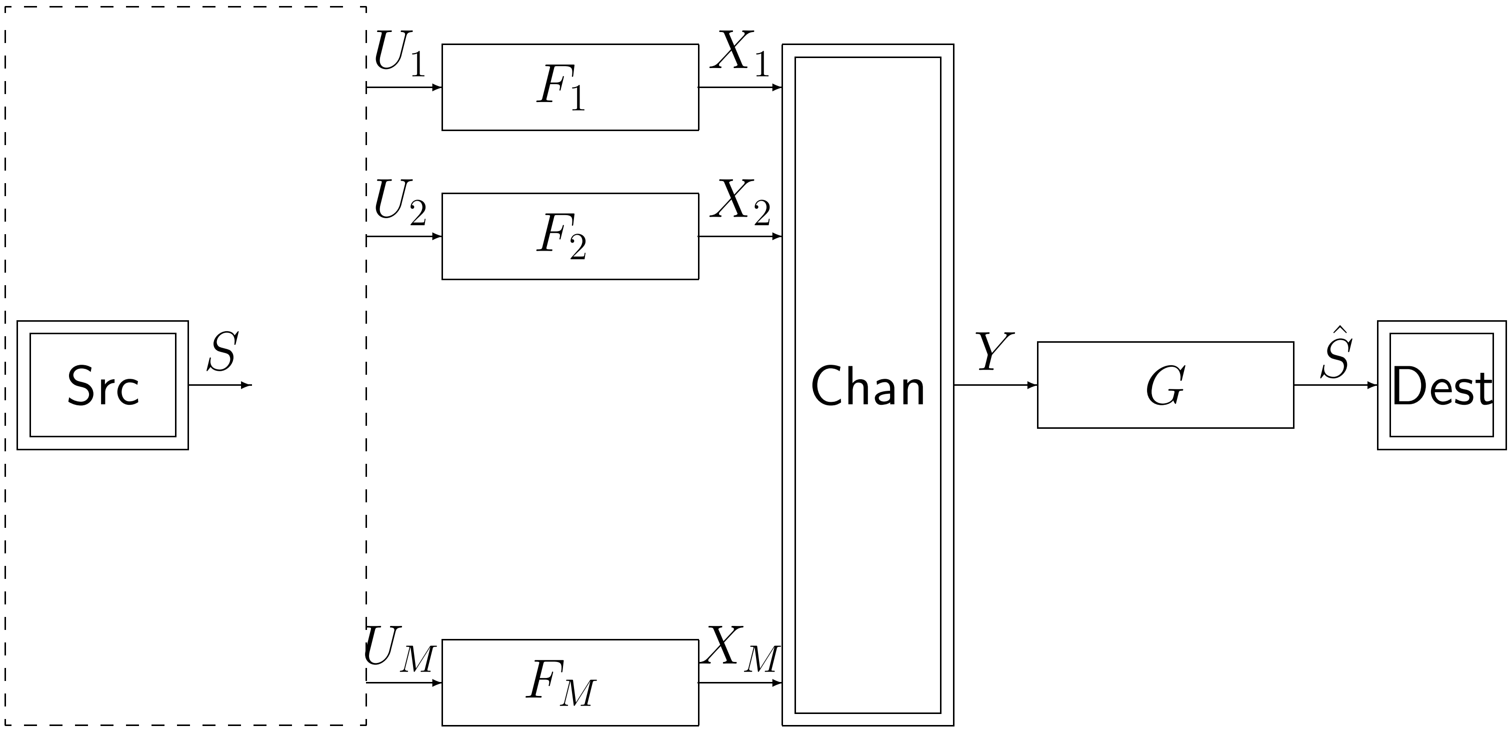
   - This suggests an alternative coding paradigm for source-channel networks.

Sensor Network: Measure-matching

Theorem. *If the coding system* $F_1, F_2, \ldots, F_M, G$ *satisfies the cost constraint* $E\rho(X_1, X_2, \ldots, X_M) \leq \Gamma$, *and*

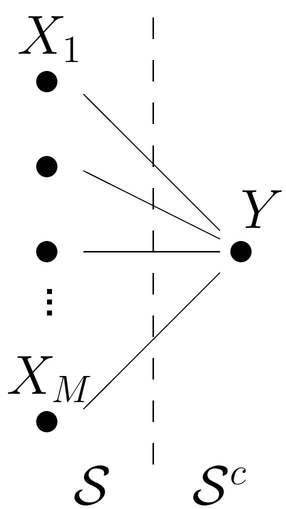
$$d(s, \hat{s}) = -\log_2 p(s|\hat{s})$$
$$I(S; U_1 U_2 \ldots U_M) = I(S; \hat{S}),$$

then it performs optimally.



Proof: **Cut-sets**

Outer bound on the capacity region of a network:



If the rates $(R_1, R_2, \ldots, R_M)$ are achievable, they must satisfy, for *every* cut $\mathcal{S}$:

$$\sum_{\mathcal{S}\to\mathcal{S}^c} R_k \leq \max_{p(x_1,x_2,\ldots,x_M)} I(X_{\mathcal{S}}; Y_{\mathcal{S}^c} | X_{\mathcal{S}^c})$$

Hence, if a scheme satisfies, for some cut $\mathcal{S}$, the above with equality, then it is optimal (with respect to $\mathcal{S}$).

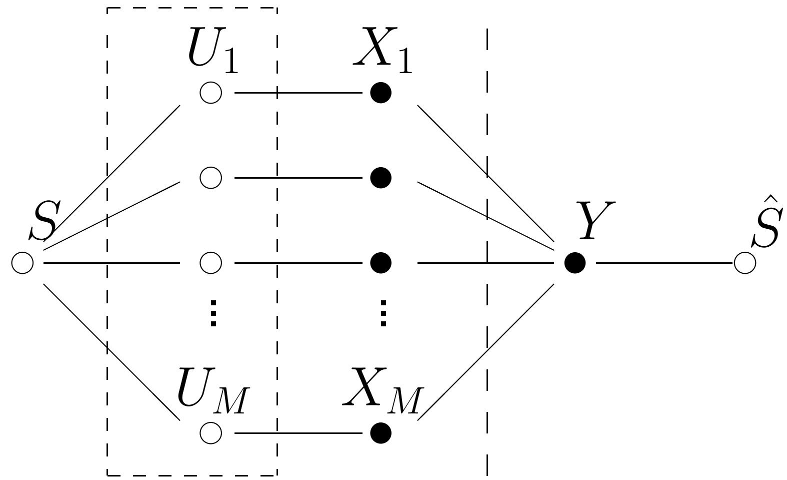
Remark. This can be sharpened.
If the rates $(R_1, R_2, \ldots, R_M)$ are achievable, then there exists some joint probability distribution $p(x_1, x_2, \ldots, x_M)$ such that for *every* cut $\mathcal{S}$:

$$\sum_{\mathcal{S}\to\mathcal{S}^c} R_k \leq I(X_{\mathcal{S}}; Y_{\mathcal{S}^c} | X_{\mathcal{S}^c})$$

Source-channel Cut-sets (1/2)

Fix the coding scheme $(F_1, F_2, \ldots, F_M, G)$. Is it optimal?

Place any "source-channel cut" through the source-channel network.



Sufficient condition for optimality:

$$R_S(\Delta) = C_{(X_1, X_2, \ldots, X_M) \to Y}(\Gamma). \quad \textit{Gaussian:} \quad D \geq \frac{\sigma_S^2 \sigma_Z^2}{M^2 P + \sigma_Z^2}.$$

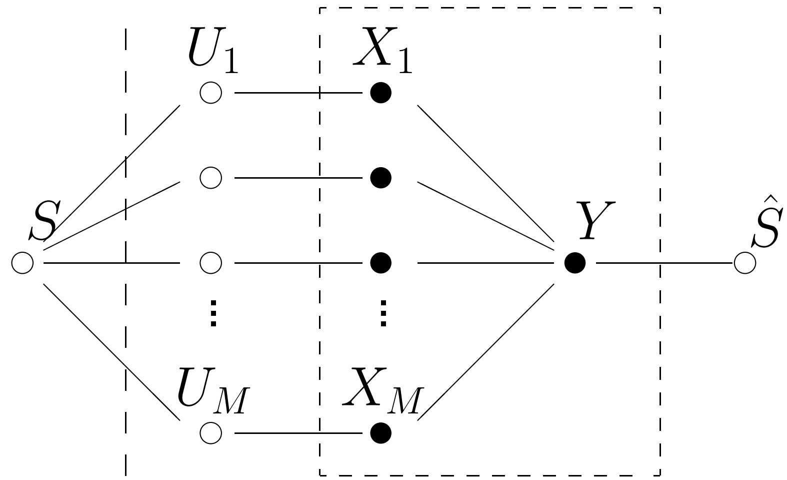
Equivalently, using measure-matching conditions,

$$\begin{aligned} \rho(x_1, x_2, \ldots, x_M) &= D(p_{Y|x_1, x_2 \ldots, x_M} || p_Y) \\ d(s, \hat{s}) &= -\log_2 p(s|\hat{s}) \\ I(S; \hat{S}) &= I(X_1 X_2 \ldots X_M; Y) \end{aligned}$$

Source-channel Cut-sets (2/2)

Fix the coding scheme $(F_1, F_2, \ldots, F_M, G)$. Is it optimal?

Place any "source-channel cut" through the source-channel network.



Sufficient condition for optimality:

$$R_S(\Delta) = C_{S \to (U_1, U_2, \ldots, U_M)}(\Gamma). \quad \textit{Gaussian:} \quad D \geq \frac{\sigma_S^2 \sigma_W^2}{M\sigma_S^2 + \sigma_W^2}.$$

Equivalently, using measure-matching conditions,

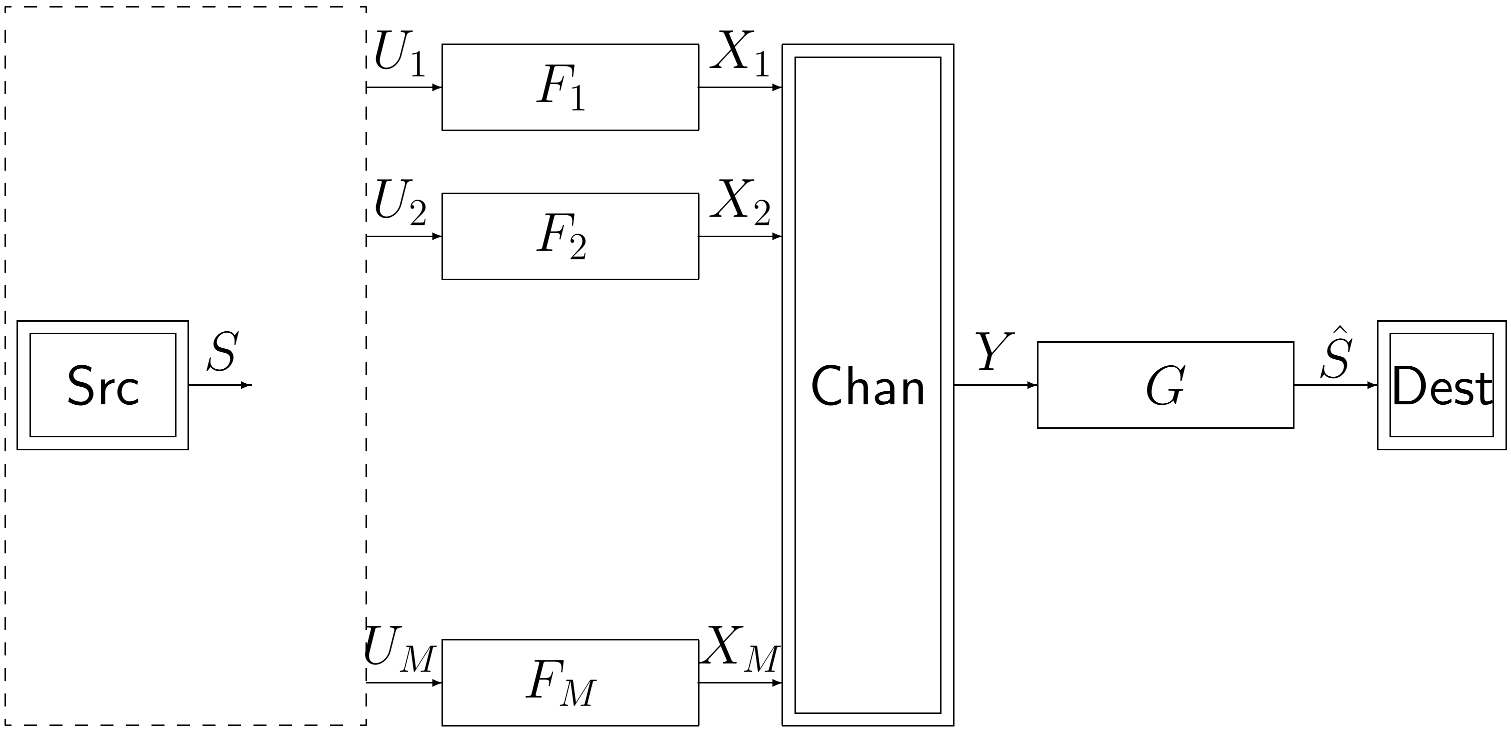
$$\begin{aligned} \rho(s) &= D(p_{U_1, U_2 \ldots, U_M | s} || p_{U_1, U_2 \ldots, U_M}) \\ d(s, \hat{s}) &= -\log_2 p(s|\hat{s}) \\ I(S; \hat{S}) &= I(S; U_1 U_2 \ldots U_M) \end{aligned}$$

Sensor Network: Measure-matching

Theorem. *If the coding system* $F_1, F_2, \ldots, F_M, G$ *satisfies the cost constraint* $E\rho(X_1, X_2, \ldots, X_M) \le \Gamma$, *and*

$$\begin{aligned} d(s, \hat{s}) &= -\log_2 p(s|\hat{s}) \\ I(S; U_1 U_2 \ldots U_M) &= I(S; \hat{S}), \end{aligned}$$

then it performs optimally.



Gaussian Example

1. The uncoded scheme satisfies the condition

$$d(s, \hat{s}) \;=\; -\log_2 p(s|\hat{s})$$

   for any M since $p(s|\hat{s})$ is Gaussian.

   More generally, this is true as soon as the *sum* of the measurement noises W_k, $k = 1, \ldots, M$, is Gaussian.

2. For the mutual information, for large M,

$$I(S; U_1 U_2 \ldots U_M) - I(S; \hat{S}) \;\leq\; c_1 \log_2 \left(1 + \frac{c_2}{M^2}\right)^M,$$

   hence the second measure-matching condition is approached as $M \to \infty$.

Measure-matching as a coding paradigm

Second observation from the Gaussian sensor network example:

2. Even for finite M, uncoded transmission considerably outperforms the best separation-based coding strategies.

Coding Paradigm. *The goal of the coding scheme in the sensor network topology is to approach*

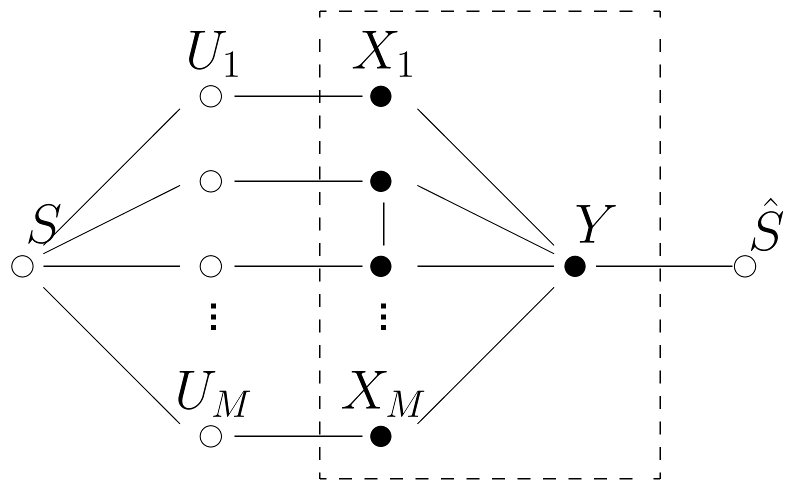
$$\begin{aligned} d(s, \hat{s}) &= -\log_2 p(s|\hat{s}) \\ I(S; U_1 U_2 \ldots U_M) &= I(S; \hat{S}), \end{aligned}$$

as closely as possible.

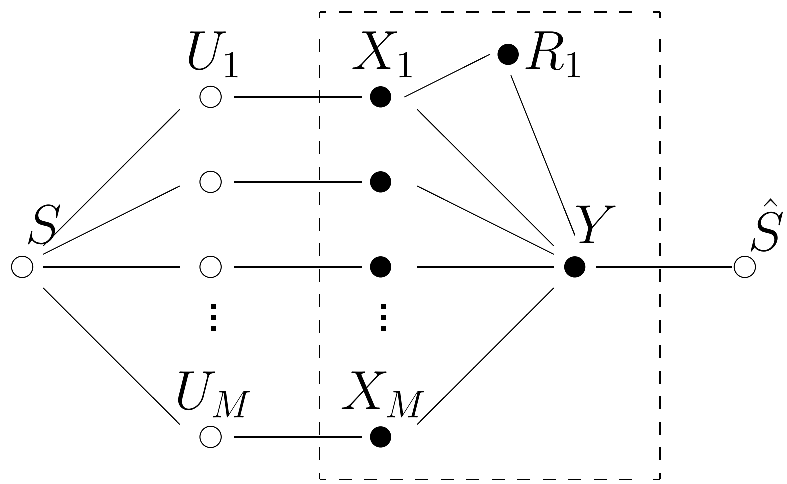
The precise meaning of "as closely as possible" remains to be determined.

Slightly Extended Topologies

- Communication between the sensors



- Sensors assisted by relays



Slightly Extended Topologies

Key insight: The same outer bound applies.

Hence,

- the *same* measure-matching condition applies, and
- in the Gaussian scenario, uncoded transmission, *ignoring*
  - the communication between the sensors, and/or
  - the relay,

  is asymptotically optimal.

But:

- Communication between the sensors simplifies the task of matching the measures.
- Relays simplify the task of matching the measures.

Can this be quantified?

Conclusions

- Rate-matching:

  Yields some *achievable* cost-distortion pairs for arbitrary network topologies.

- Measure-matching:

  Yields some *optimal* cost-distortion pairs for certain network topologies, including

  * single-source broadcast
  * sensor network
  * sensor network with communication between the sensors
  * sensor network with relays

What is Information?

Point-to-point:

"Information = Bits"

Network:

"Information = ???"

References:

1. M. Gastpar and M. Vetterli, "Source-channel communication in sensor networks," IPSN 2003 and *Springer Lecture Notes in Computer Science*, April 2003.

2. M. Gastpar, "Cut-set bounds for source-channel networks," *in preparation.*