Sensing Reality and Communicating Bits
A Dangerous Liaison

Michael C. Gastpar

Wireless Foundations Center
Department of Electrical Engineering and Computer Sciences
University of California, Berkeley

January 10, 2007

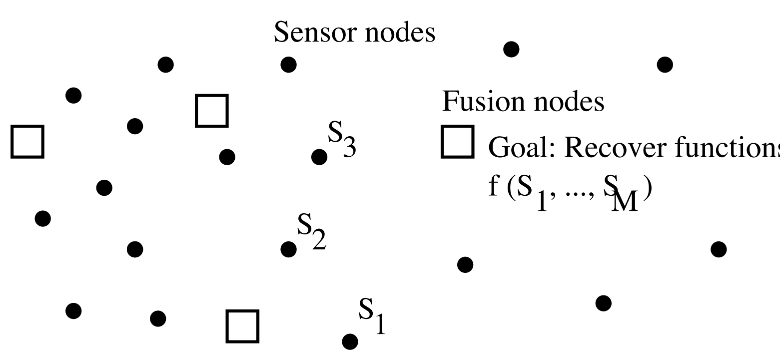
Wireless Foundations

Outline

Acknowledgements

- Martin Vetterli (EPFL)
- Pier Luigi Dragotti (Imperial College)
- Bobak Nazer (Berkeley)
- Krishnan Eswaran (Berkeley)
- Anand Sarwate (Berkeley)

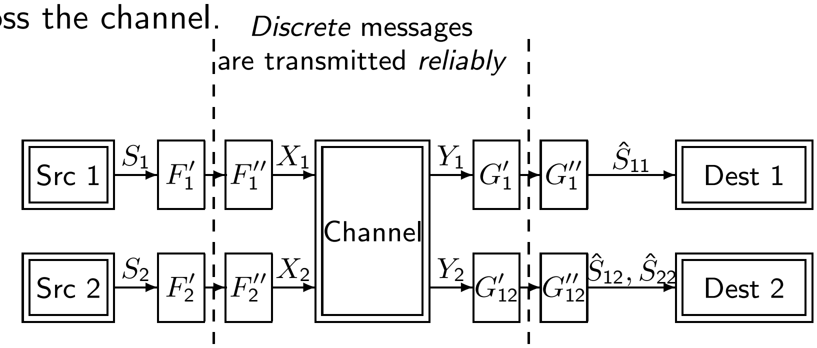
General Problem



- Each destination node wants to get out one or more **functions** of the sensor measurements.
- Goal: Architectural guidelines.

Digital Communication

- One (default?) architectural guideline is digital communication.
- **Definition:** The guideline consists in representing the source information in digital form (bits) and communicating those bits across the channel.



- This leads to a universal interface.

Digital Communication

- Digital communication induces a universal ordering of communication networks and of source representation problems:
  - A communication network is more valuable the **more capacity it has.**
  - A source representation problem is easier the **fewer bits it requires** as a source coding problem.
- Digital communication thus uses bits as a universal currency of information.

Sensor Network Problem Classes

- However, the classification induced by digital communication (bits) is **not** universally correct.
- In other words, a communication network of smaller capacity may be **more valuable** for a certain sensing task.
- A global ordering of sensing problems (think P/NP) is not available at present.
- However, several **partial orderings** can be given. We will briefly discuss two examples.

Partial Orderings

Example Class 1: "Independent Sources."
Consider the sub-class of sensing problems where

1. all sensor observations are independent of each other, and
2. the destination(s) want to recover each source separately (perhaps with respect to a fidelity criterion).

Then, digital communication induces the correct ordering. (Or, as one hears sometimes, "a separation theorem applies.")

This example can be found, e.g., in MG, *To Code or Not To Code,* Ph.D. thesis, 2002.

Partial Orderings

Example Class 2: “Parallel Channels.”
Consider the sub-class of sensor network problems where

1. the connections from each sensor to each destination is a separate, independent point-to-point link (a “wire”).

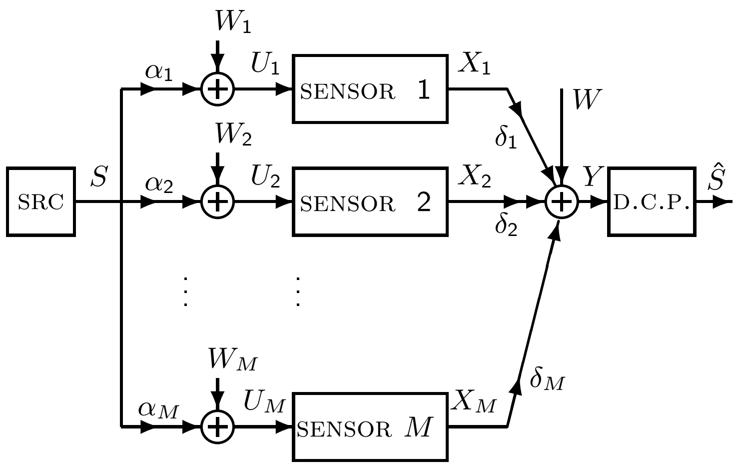
Then, digital communication induces the correct ordering. (Or, as one hears sometimes, “a separation theorem applies.”)

Counterexamples to a universal ordering

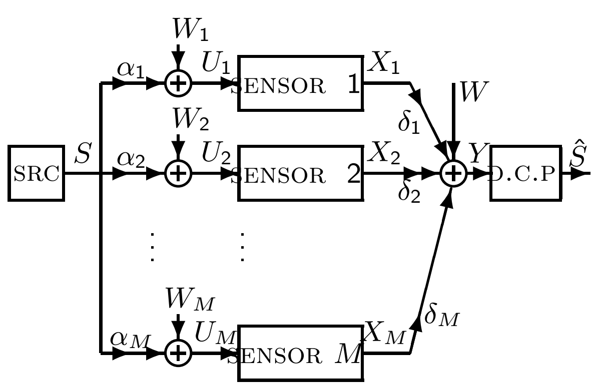
However, it appears that a universal ordering will be hard to come by.
Two case studies will now be discussed:

1. My Sensor Network 101
2. Computation Coding

Case Study 1: "My Sensor Network 101"



Modeling Assumptions



1. *Gaussian Memoryless:* All involved statistics are Gaussian, and the coefficients α and δ are fixed and known.
2. *General Memoryless:* Arbitrary memoryless source, arbitrary noise with fixed second-order statistics.

Achievability

- For the *general memoryless* model, the following distortion can be attained:

$$D = \frac{\sigma_S^2\sigma_W^2}{\sigma_W^2 + \sum_{m=1}^M |\alpha_m|^2\sigma_S^2} \cdot \left(1 + \frac{(\sigma_S^2\sigma_Z^2/\sigma_W^2)\sum_{m=1}^M |\alpha_m|^2}{\sigma_Z^2 + P_{tot}(M)b(M)}\right),$$

where

$$b(M) = \frac{\left(\sigma_W^2 + \sum_{m=1}^M |\alpha_m|^2\sigma_S^2\right)\sum_{m=1}^M |\alpha_m|^2}{\sum_{m=1}^M(|\alpha_m|^2\sigma_S^2 + \sigma_W^2)|\alpha_m|^2/|\delta_m|^2}.$$

- Specifically, for the simplest case ($\alpha_m = \delta_m = 1$), we can simplify as follows:

$$D = \frac{\sigma_S^2\sigma_W^2}{M\sigma_S^2 + \sigma_W^2}\left(1 + \frac{M(\sigma_S^2\sigma_Z^2/\sigma_W^2)}{\frac{M\sigma_S^2+\sigma_W^2}{\sigma_S^2+\sigma_W^2}P_{tot} + \sigma_Z^2}\right) \sim \frac{1}{M}.$$

Achievability via digital communication

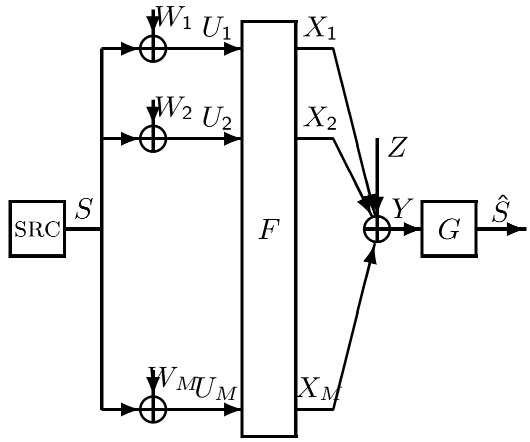
- For the *Gaussian memoryless* model, one can show that digital communication cannot attain a distortion smaller than

$$D \geq \frac{\sigma_S^2\sigma_W^2}{\sigma_S^2 \log_2\left(1 + \frac{MP_{tot}}{\sigma_Z^2}\right) + \sigma_W^2} \sim \frac{1}{\log M}.$$

- Unfortunately, it is hard to extend this to the general case. Gaussian statistics are worst-case, and thus, digital could be much better that the above formula for certain cases.

A simple converse

- For the *Gaussian memoryless* model, we can give the following lower bound:



- We can extend the same bounding technique to the *general memoryless* model.

A simple converse

- For the simplest case ($\alpha_m = \delta_m = 1$), this leads to

$$D \geq \frac{\sigma_S^2 \sigma_W^2}{M\sigma_S^2 + \sigma_W^2} \left(1 + \frac{M(\sigma_S^2 \sigma_Z^2 / \sigma_W^2)}{M P_{tot} + \sigma_Z^2} \right),$$

compared to the achievable distortion of

$$D = \frac{\sigma_S^2 \sigma_W^2}{M\sigma_S^2 + \sigma_W^2} \left(1 + \frac{M(\sigma_S^2 \sigma_Z^2 / \sigma_W^2)}{\frac{M\sigma_S^2 + \sigma_W^2}{\sigma_S^2 + \sigma_W^2} P_{tot} + \sigma_Z^2} \right).$$

- So, could we do (ever so slightly) better? What would the improved code look like? ...

Witsenhausen's argument

- The concept of **maximal correlation:** Consider two sequences $U_1[n]$ and $U_2[n]$ sampled i.i.d. from the joint distribution $p(u_1, u_2)$, and two arbitrary real-valued functions $f_1(\cdot)$ and $g_1(\cdot)$ satisfying

$$E[f_1(U_1)] = E[g_1(U_2)] = 0, \quad \text{and} \quad E[f_1^2(U_1)] = E[g_1^2(U_2)] = 1.$$

The **maximal correlation** is the quantity

$$\rho^* = \sup_{f_1, g_1} E[f_1(U_1)g_1(U_2)].$$

Witsenhausen's argument

Lemma

For any real-valued functions $f_N(\cdot)$ and $g_N(\cdot)$ satisfying

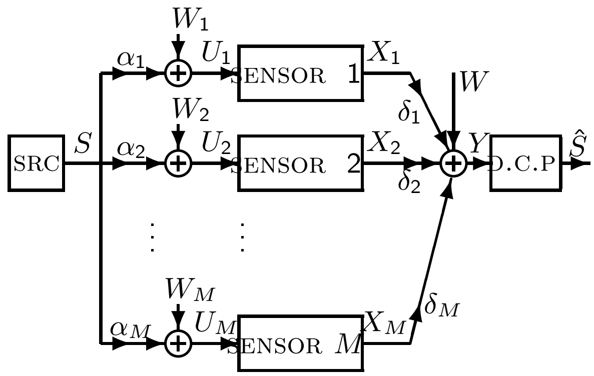
$$E[f_N(U_1^N)] = E[g_N(U_2^N)] = 0,$$
$$E[f_N^2(U_1^N)] = E[g_N^2(U_2^N)] = 1,$$

we have that

$$\sup_{f_N, g_N} E[f_N(U_1^N)g_N(U_2^N)] \leq \rho^*.$$

To prove this...

Witsenhausen's argument



- It remains to determine the *maximal correlation* ρ^* for our case.
- Not unexpectedly, for the Gaussian case, this is merely the *regular correlation* between the sequences.
- To prove this, we can approximate the functions via Hermite-Tchebycheff polynomials. (See e.g. Lancaster, 1957.)

The resulting converse bound...

First, we can upper bound the sum rate

$$I(X_1^N, X_2^N, \ldots, X_M^N; Y^N) \le \sum_{n=1}^{N} \frac{1}{2} \log(1 + \frac{1}{\sigma_Z^2} \delta^T \Sigma_n \delta).$$

But the covariance matrices Σ_n must satisfy Witsenhausen's bound. Therefore, their average satisfies

$$\{\overline{\Sigma}\}_{m,m} \le P_m \quad \text{and} \quad \{\overline{\Sigma}\}_{m,m'} \le \rho^*_{m,m'} \sqrt{P_m P_{m'}}, \text{ for } m \neq m'.$$

The resulting converse bound...

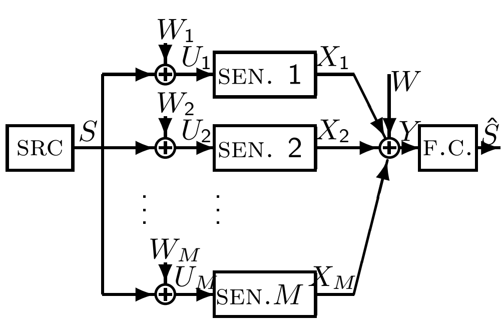
- For our simplest case, we find

$$\rho^*_{m,m'} = \rho^* \stackrel{\text{def}}{=} \frac{\sigma_S^2}{\sigma_S^2 + \sigma_W^2}.$$

- Hence,

$$I(X_1^N, X_2^N, \ldots, X_M^N; Y^N) \leq \frac{1}{2} \log \left(1 + \frac{P_{tot} \frac{M\sigma_S^2 + \sigma_W^2}{\sigma_S^2 + \sigma_W^2}}{\sigma_Z^2} \right).$$

So, for the 101 version, uncoded is exactly optimal...



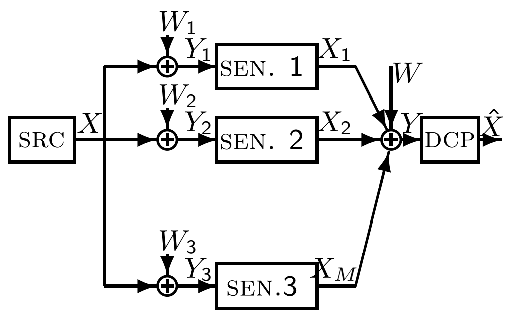
Theorem

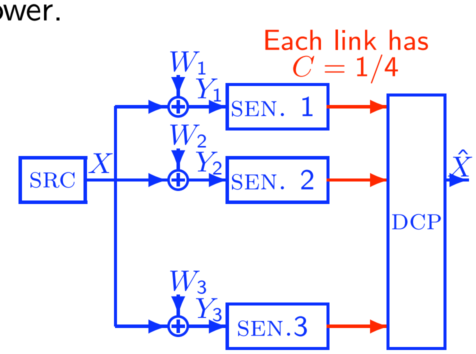
For the Gaussian sensor network 101, uncoded transmission is exactly optimal and attains

$$D = \frac{\sigma_S^2 \sigma_W^2}{M\sigma_S^2 + \sigma_W^2} \left(1 + \frac{M(\sigma_S^2 \sigma_Z^2 / \sigma_W^2)}{\frac{M\sigma_S^2 + \sigma_W^2}{\sigma_S^2 + \sigma_W^2} P_{tot} + \sigma_Z^2} \right).$$

A Riddle...

- Unit variances, unit total transmit power.



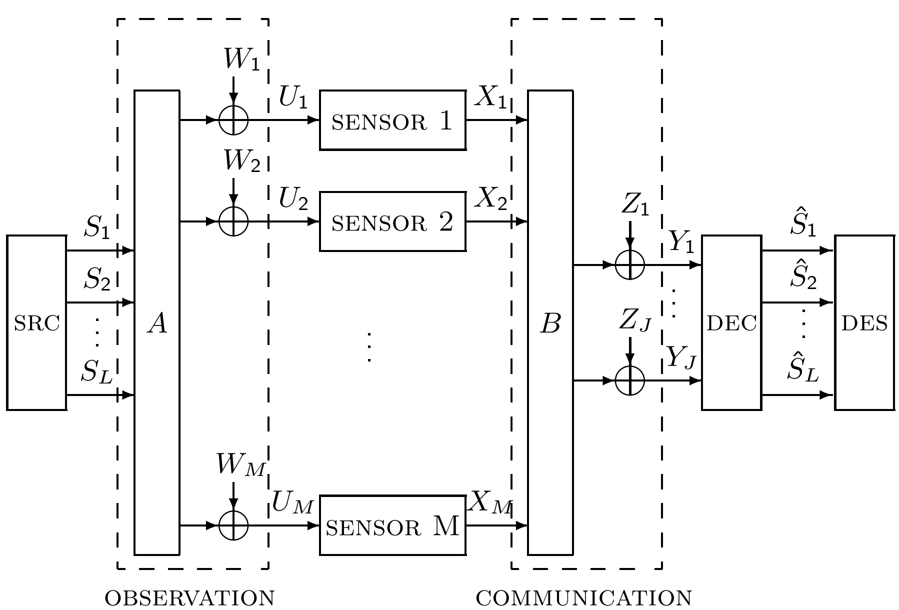


System Capacity: $C = 1/2 \log_2(1+1) = 1/2$. System Capacity: $C = 3/4$.

- However, the attainable overall distortions are:

$$D = \frac{1}{2} \qquad\qquad D \geq \frac{1}{1.75}$$

The Natural Extensions of 101



The Natural Extensions of 101

Several cases have been analyzed:

- *Slowly expanding sensor network,* **with observation noise.**

$$\text{Lower bound: } D \geq \frac{1}{M} \qquad \text{Via Digital: } D \geq \frac{1}{\log M}$$

- *Slowly expanding sensor network,* **no observation noise.**

$$\text{Lower bound: } D \geq \frac{1}{M^2} \qquad \text{Via Digital: } D \geq \frac{1}{M}$$

- *"Linearly" expanding sensor network.*
  Here, the distortion goes to a **constant.**
- Gastpar/Vetterli/Dragotti, *Signal Processing Magazine,* July 2006.

The Natural Extensions of 101

Moreover:

- *Randomly varying observation matrix.*

$$D \geq E_\alpha \left[\frac{\sigma_S^2 \sigma_W^2}{\alpha^2 \sigma_S^2 + \sigma_W^2} \right] + c_1 \left(\frac{1}{c_2 + \frac{P_{tot}}{K \tilde{J} \sigma_Z^2} \mathcal{G}(\beta)} \right)^{K\tilde{J}/L}$$

Gastpar/Vetterli, *JSAC,* 2005.

Case Study 2: Computation Coding

- There was no time to discuss this, but we point to the work of Nazer and Gastpar (see e.g. ISIT 2006).

Some Conclusions

- We illustrated the “danger” of trusting bits as a **universal currency of information** in a sensor network context by the aid of two “paradigmatic” examples, namely
  1. Sensor network 101
  2. Computation coding
     Including several novel code constructions.