Sensor Cooperation via Noiseless and Noisy Feedback

Michael Gastpar
University of California, Berkeley

Joint work with: Gerhard Kramer, Bell Labs.

Boston University: May 25, 2006.

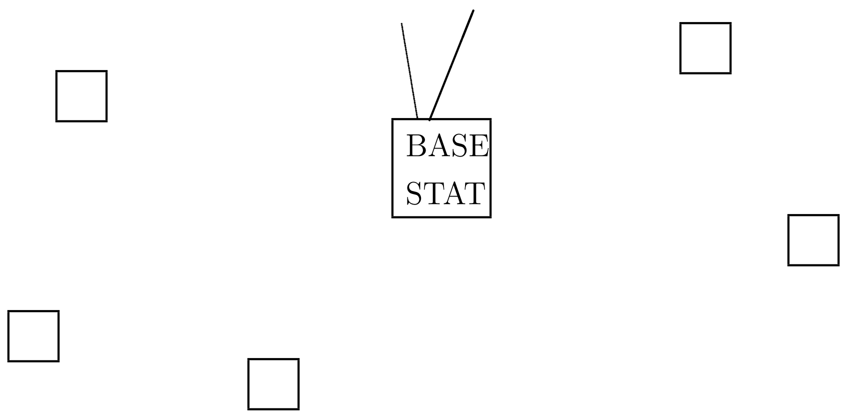
Outline

1. The Use of Noiseless Feedback for Gaussian channels
   (a) Point-to-point
   (b) 2-user Multiple Access

2. Two open problems:
   (a) More than two users
   (b) Noisy feedback

3. Feedback for the Gaussian "Sensor Network 101"

Motivation...

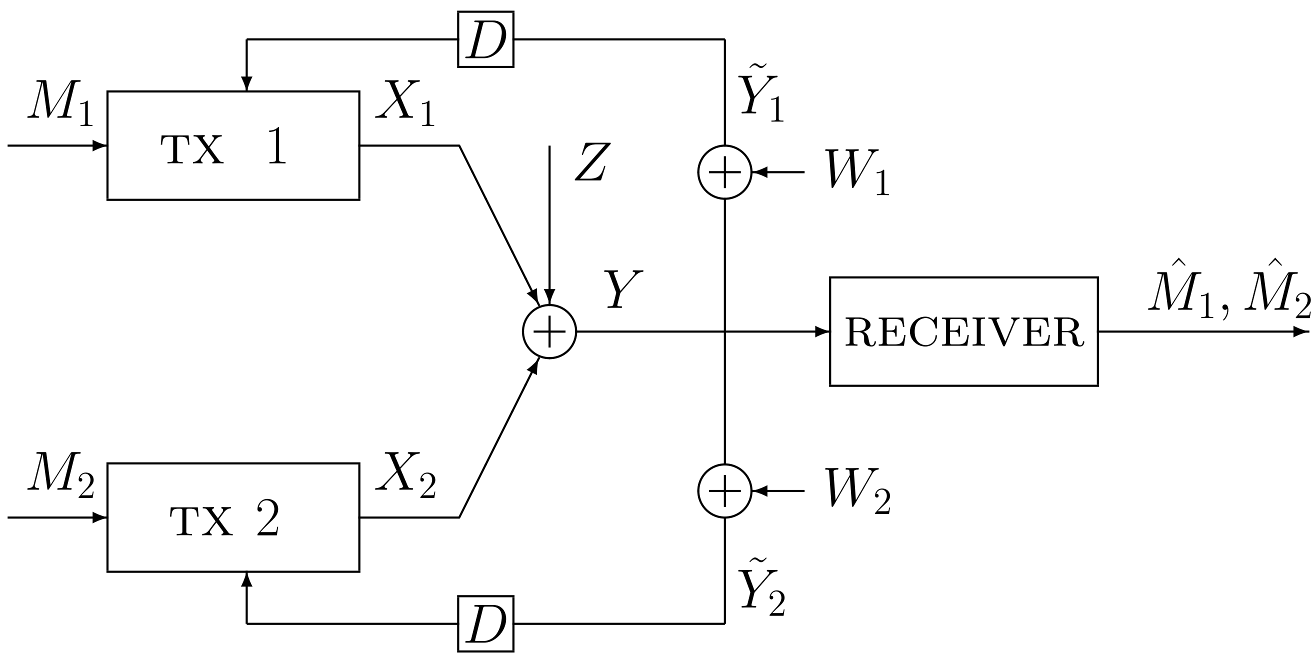


(High-quality) feedback is somewhat realistic in sensor networks.

How much can it buy us?

Motivation...

We will look at (AWGN) Multi-Access with Feedback:



1. Noiseless Feedback and Gaussian channels

Two feedback coding ideas:

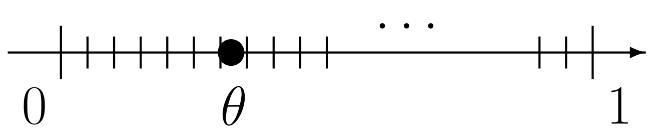
- Transmit only what the decoder does not know yet
  — "Send-innovations-only" approach.

- Predict the noise — or the other users.
  — Hence, make the signals transmitted by the two users *dependent*.

Schalkwijk-Kailath

The transmitter selects one out of 2^{nR} different messages in n channel uses.

Split the unit interval into 2^{nR} equal subintervals.



Identify the midpoint of each subinterval with one of the messages.

Selecting a message is now equivalent to selecting one of the midpoints, call it θ.

Now, consider the strategy: In channel use n, send

$$X_n = \alpha_n \left(\theta - \hat{\theta}_n(Y_1, \ldots, Y_{n-1}) \right),$$

where $\hat{\theta}_n(Y_1, \ldots, Y_{n-1})$ is the destination's estimate of the "message point" θ upon receiving $Y_1, \ldots, Y_{n-1}$.

Multiple Access

Now, we have two (independently selected) message points.

In channel use n, send

$$X_{1,n} = \alpha_{1,n}\left(\theta_1 - \hat{\theta}_{1,n}(Y_1, \ldots, Y_{n-1})\right),$$
$$X_{2,n} = \alpha_{2,n}\left(\theta_2 - \hat{\theta}_{2,n}(Y_1, \ldots, Y_{n-1})\right),$$

where $\hat{\theta}_{1,n}(Y_1, \ldots, Y_{n-1})$ is the destination's estimate of the "message point" θ_1 upon receiving $Y_1, \ldots, Y_{n-1}$.

Multiple Access

Hence, everything is defined — except for the *phase* of $\alpha_{1,n}$ and $\alpha_{2,n}$.

Without loss of generality, $\alpha_{1,n}$ may be selected positive (and real-valued) throughout.

Then, it is easy to show that the best strategy is to *alternate* the sign of $\alpha_{2,n}$.

Can be interpreted like an "orthogonalization" — Fourier basis.

Ozarow, 1984.

Multiple Access: Converse

$$\begin{aligned} R_1 &= H(W_1) \\ &= \frac{1}{n} I(W_1; Y^n | W_2) + \frac{1}{n} H(W_1 | Y^n, W_2) \\ &\leq I(X_1; Y | X_2) + \epsilon_n \end{aligned}$$

$$R_2 \leq I(X_2; Y | X_1) + \epsilon_n$$

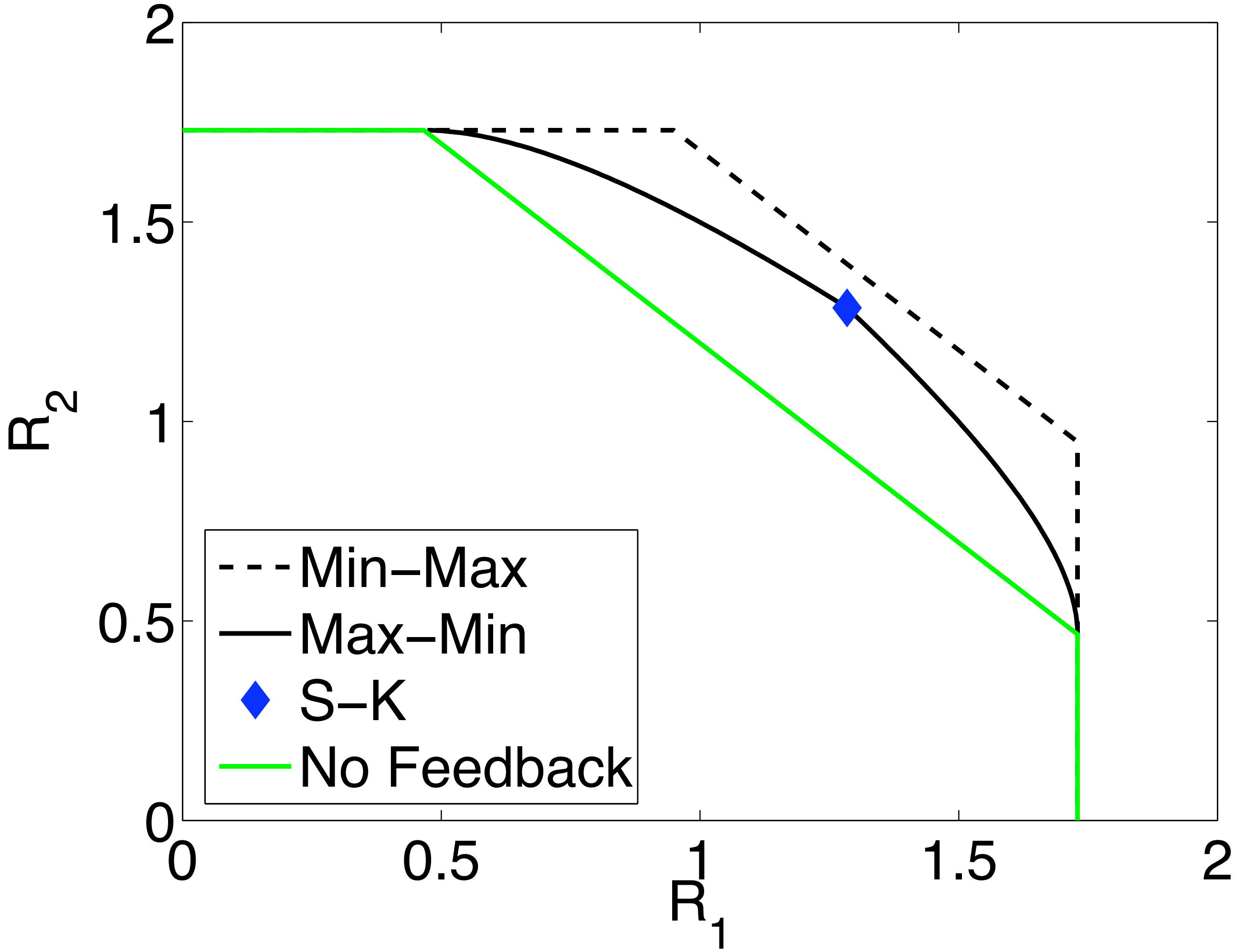
$$\begin{aligned} R_1 + R_2 &= H(W_1, W_2) \\ &\leq I(X_1, X_2; Y) + \epsilon_n \end{aligned}$$

If (R_1, R_2) is achievable, there must exists some joint distribution $p(x_1, x_2)$ such that all three conditions hold.

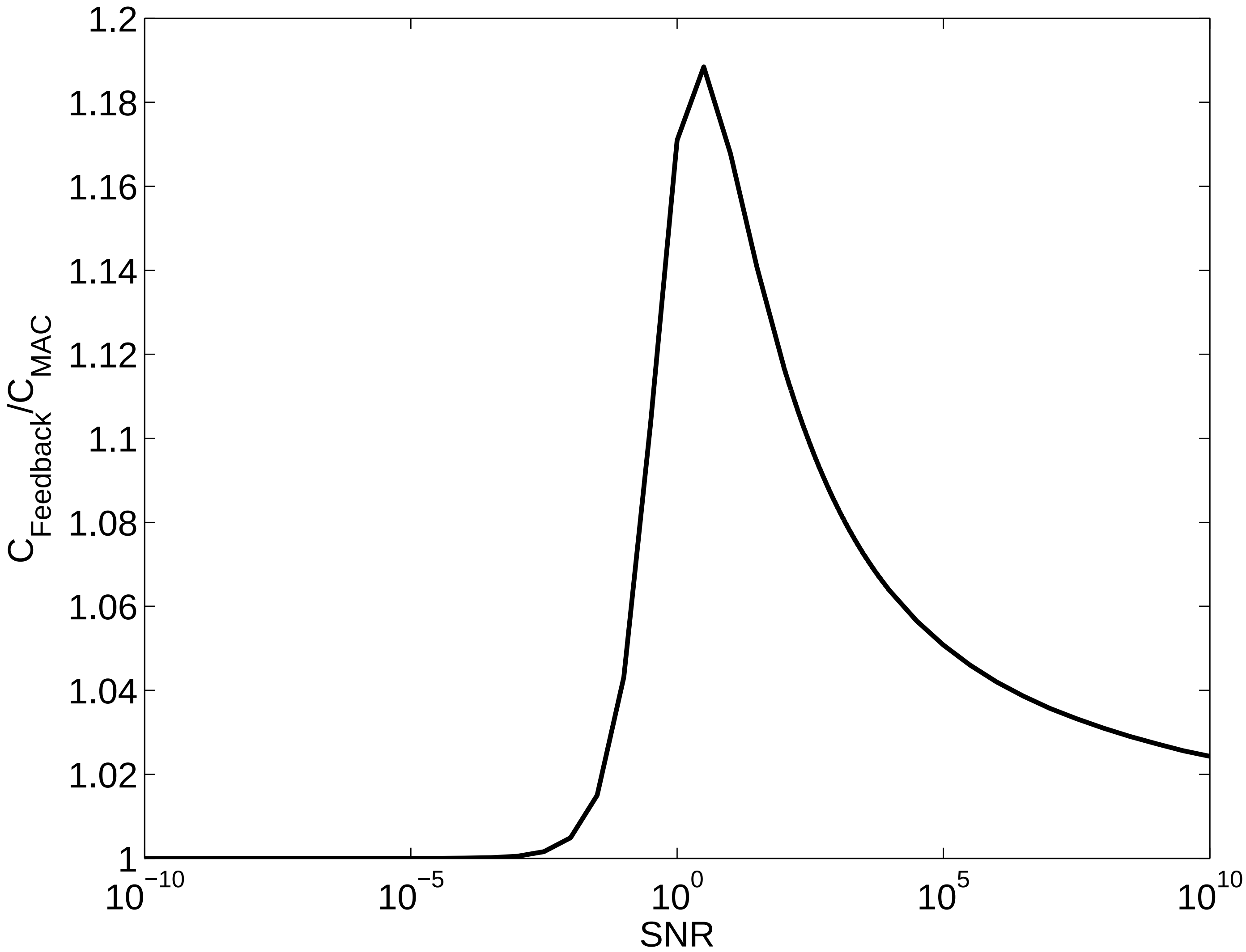
Note: This bound allows for arbitrary (causal) feedback.

Multiple Access Capacity

$P_1 = P_2 = 10:$



Where does feedback really pay off?



2. Open Questions

Our methods permit to shed new light onto two questions:

(a) More than two users

(b) Noisy feedback

More than two users

How to extend the Schalkwijk-Kailath strategy?

In channel use n, send

$$\begin{aligned} X_{1,n} &= \alpha_{1,n}\left(\theta_1 - \hat{\theta}_{1,n}(Y_1,\ldots,Y_{n-1})\right), \\ X_{2,n} &= \alpha_{2,n}\left(\theta_2 - \hat{\theta}_{2,n}(Y_1,\ldots,Y_{n-1})\right), \\ &\vdots \\ X_{M,n} &= \alpha_{M,n}\left(\theta_M - \hat{\theta}_{M,n}(Y_1,\ldots,Y_{n-1})\right) \end{aligned}$$

where $\hat{\theta}_{1,n}(Y_1,\ldots,Y_{n-1})$ is the destination's estimate of the "message point" θ_1 upon receiving $Y_1,\ldots,Y_{n-1}$.

Again, the $\alpha_{m,n}$ are determined up to their phase...

More than two users

Select the phases according to the M-dimensional Fourier basis.

This leads to a sum rate as follows:

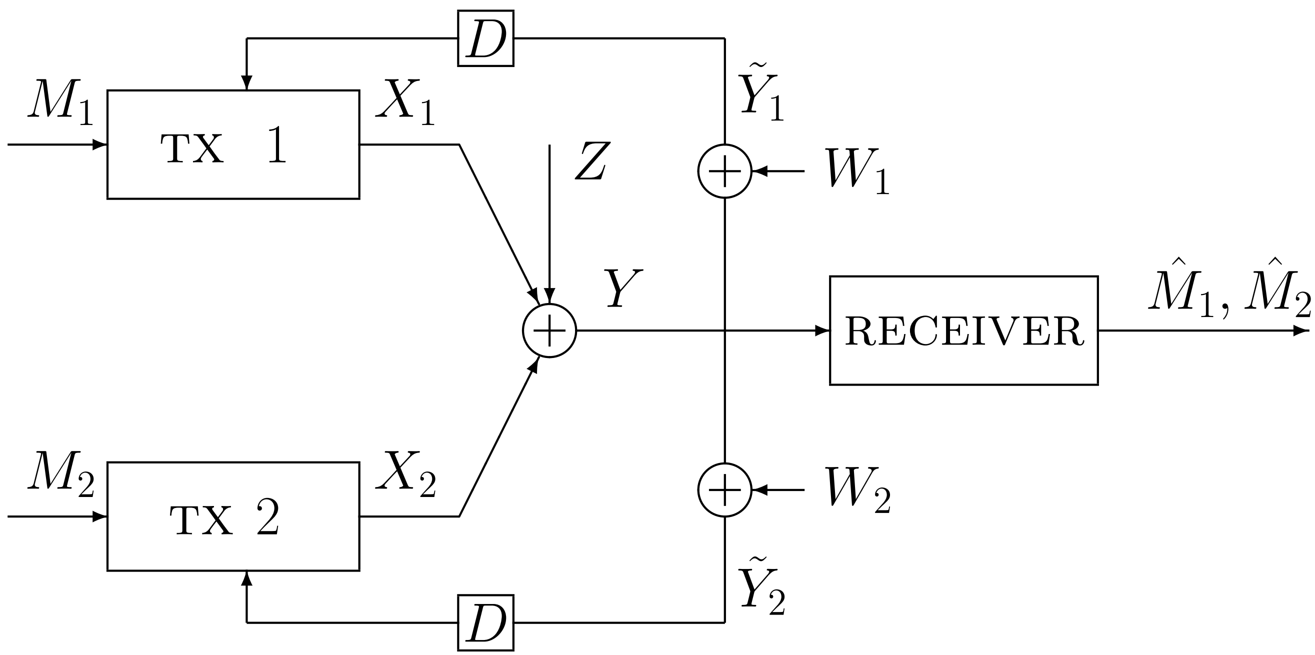
$$\sum_{m=1}^{M} R_m \approx \frac{1}{2} \log M + \frac{1}{2} \log \log M.$$

Kramer, 2002

The known outer bounds only show that the sum rate is upper bounded as

$$\sum_{m=1}^{M} R_m \leq \frac{1}{2} \log \left(1 + M^2 P\right)$$

Noisy Feedback



Noisy Feedback

How to extend the Schalkwijk-Kailath strategy to noisy feedback?

For example, in channel use n, we could send

$$X_{1,n} = \alpha_{1,n}\left(\theta_1 - \hat{\hat{\hat{\theta}}}_{1,n}(Y_1, \ldots, Y_{n-1})\right),$$
$$X_{2,n} = \alpha_{2,n}\left(\theta_2 - \hat{\hat{\hat{\theta}}}_{2,n}(Y_1, \ldots, Y_{n-1})\right).$$

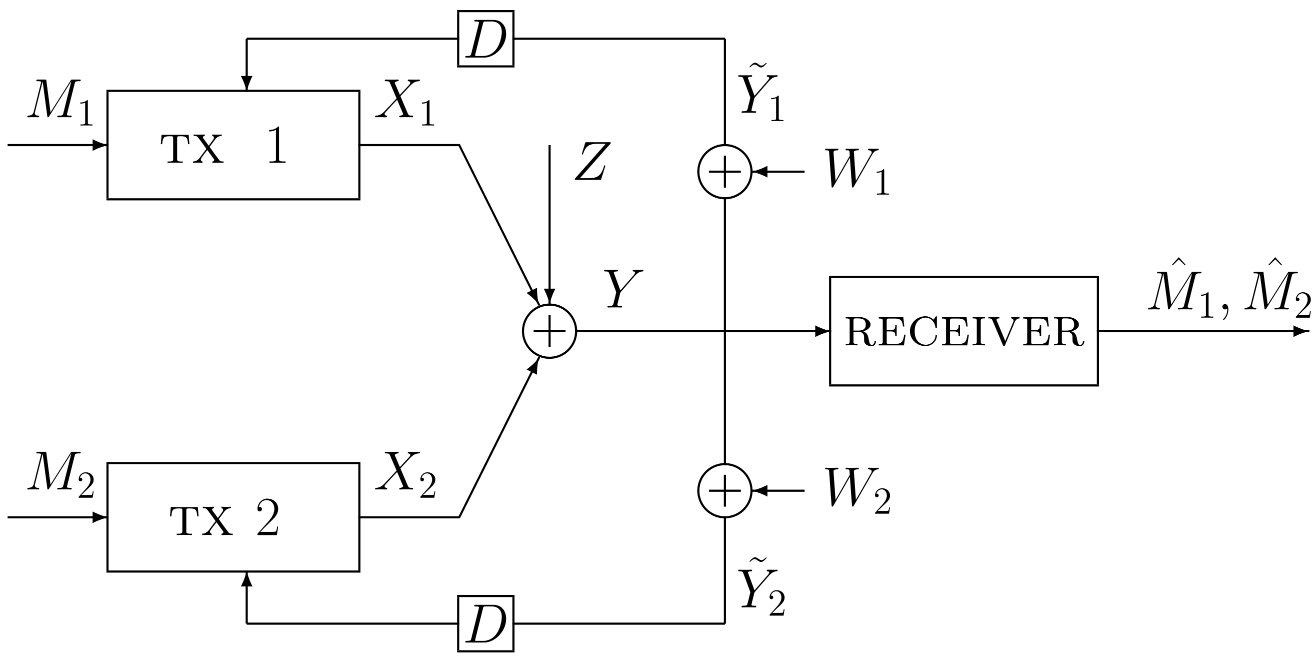
Or, optimize over all linear strategies.

But the main insight is: as n becomes large, this strategy becomes useless.

Instead, use the strategy for n^* channel uses, then start anew.

But there is no reason to believe this should be optimal... we need a non-trivial upper bound!

Dependence-Balance Bound



Cut-set bound revisited:

$$\begin{aligned} R_{\mathcal{S}} &= H(W_{\mathcal{S}}) \\ &= \frac{1}{n} I(W_{\mathcal{S}}; Y^n_{\mathcal{S}^c} | W_{\mathcal{S}^c}) + \frac{1}{n} H(W_{\mathcal{S}} | Y^n_{\mathcal{S}^c}, W_{\mathcal{S}^c}) \\ &\leq \frac{1}{n} \sum_{i=1}^{n} I(X_{\mathcal{S},i}; Y_{\mathcal{S}^c,i} | X_{\mathcal{S}^c,i}, Y^{i-1}) + \epsilon_n \end{aligned}$$

Dependence-Balance Bound

Cut-set bound revisited:

$$
\begin{aligned}
R_{\mathcal{S}} &= H(W_{\mathcal{S}}) \\
&= \frac{1}{n} I(W_{\mathcal{S}}; Y_{\mathcal{S}^c}^n | W_{\mathcal{S}^c}) + \frac{1}{n} H(W_{\mathcal{S}} | Y_{\mathcal{S}^c}^n, W_{\mathcal{S}^c}) \\
&\leq \frac{1}{n} \sum_{i=1}^{n} I(X_{\mathcal{S},i}; Y_{\mathcal{S}^c,i} | X_{\mathcal{S}^c,i}, Y^{i-1}) + \epsilon_n
\end{aligned}
$$

Specifically, for $\mathcal{S} = \{1\}$, we get

$$
R_1 \leq \frac{1}{n} \sum_{i=1}^{n} I(X_{1,i}; Y_i | X_{2,i}, Y^{i-1})
$$

Define S to be a uniform random variable over $\{1, 2, \cdots, n\}$ and $T = (S, Y^{S-1})$. Then,

$$
R_1 \leq I(X_{1,T}; Y_T | X_{2,T}, T),
$$

Dependence-Balance Bound

Hence, if (R_1, R_2) achievable, there must exist $p(t)p(x_1, x_2|t)$ such that

$$\begin{aligned} R_1 &\le I(X_{1,T}; Y_T | X_{2,T}, T) \\ R_2 &\le I(X_{2,T}; Y_T | X_{1,T}, T) \\ R_1 + R_2 &\le I(X_{1,T}, X_{2,T}; Y_T | T) \end{aligned}$$

This region is already convex. Therefore, the T can be dropped, and the region can be expressed as

$$\begin{aligned} R_1 &\le \frac{1}{2}\log_2\left(1 + \frac{P}{\sigma^2}(1-\rho^2)\right) \\ R_2 &\le \frac{1}{2}\log_2\left(1 + \frac{P}{\sigma^2}(1-\rho^2)\right) \\ R_1 + R_2 &\le \frac{1}{2}\log_2\left(1 + \frac{P}{\sigma^2}2(1+\rho)\right) \end{aligned}$$

Dependence-Balance Bound

Instead, let us consider an approach due to Hekstra and Willems ($\textit{IEEE Trans IT}$, submitted 1986, published 1989.)

Trivially, on top of Fano's inequality, any coding scheme must also satisfy

$$0 \leq I(M_1; M_2 | Y^n) - I(M_1; M_2)$$

Along the same expansions, we can express this as

$$0 \leq \sum_{i=1}^{n} \left(\underbrace{I(X_{1,i}; X_{2,i} | Y_i, Y^{i-1})}_{\text{"dependence produced"}} - \underbrace{I(X_{1,i}; X_{2,i} | Y^{i-1})}_{\text{"dependence consumed"}} \right)$$

Dependence-Balance Bound

Along the same expansions, we can express this as

$$0 \le \sum_{i=1}^{n} \left(\underbrace{I(X_{1,i}; X_{2,i}|Y_i, Y^{i-1})}_{\text{“dependence produced”}} - \underbrace{I(X_{1,i}; X_{2,i}|Y^{i-1})}_{\text{“dependence consumed”}} \right)$$

We can rewrite this as the condition

$$I(X_1; X_2|T) \le I(X_1; X_2|Y, T),$$

where $T = (S, Y^{S-1})$ is the same as in the previous outer bound.

Hence, if (R_1, R_2) achievable, there must exist $p(t)p(x_1, x_2|t)$ satisfying the above and

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

Dependence-Balance Bound

To gain understanding, we can also rewrite this bound as

$$I(X_1, X_2; Y|T) \leq I(X_1; Y|X_2, T) + I(X_2; Y|X_1, T),$$

where $T = (S, Y^{S-1})$ is the same as in the previous outer bound.

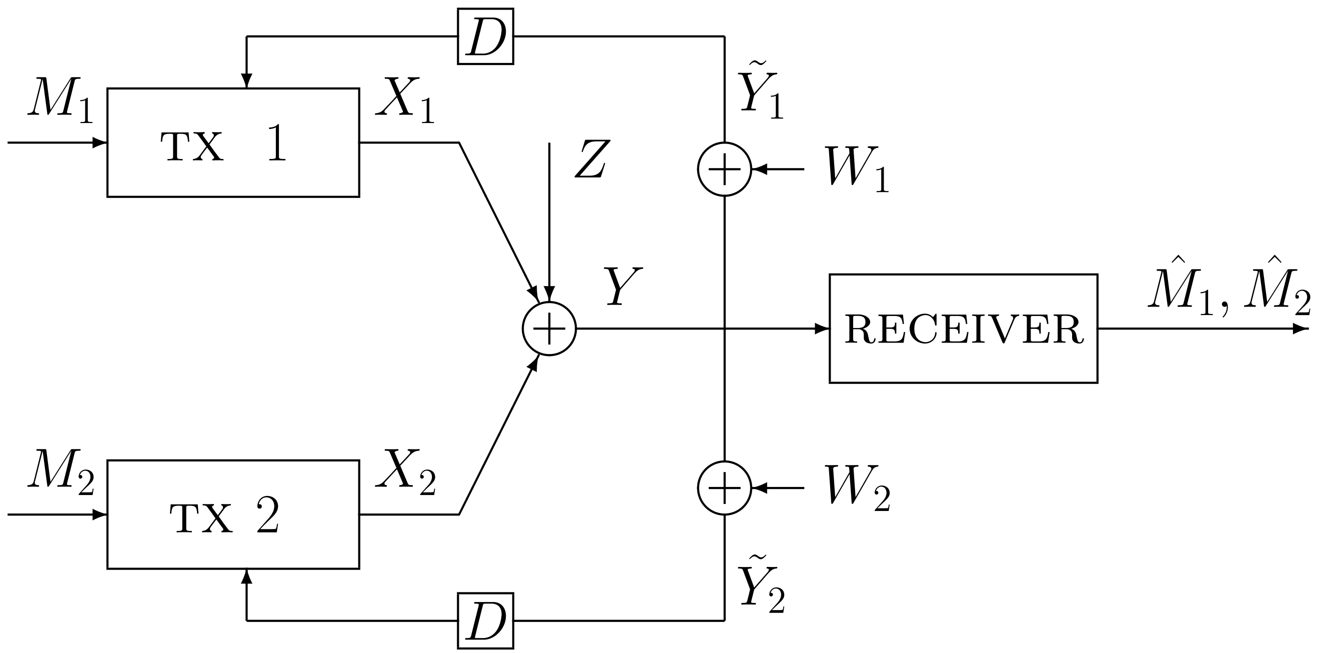
Hence, if (R_1, R_2) achievable, there must exist $p(t)p(x_1, x_2|t)$ satisfying the above and

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

It turns out that the condition is not needed.

However, if we keep the condition, then it becomes tricky to prove that it is sufficient to consider Gaussian distributions.

Dependence-Balance Bound — Noisy Feedback



Dependence-Balance Bound — Noisy Feedback

But here, we can bring in the fact that the FB is noisy!

(Provided it is a degraded version of the channel output Y.)

Namely, by the same token, we can infer that

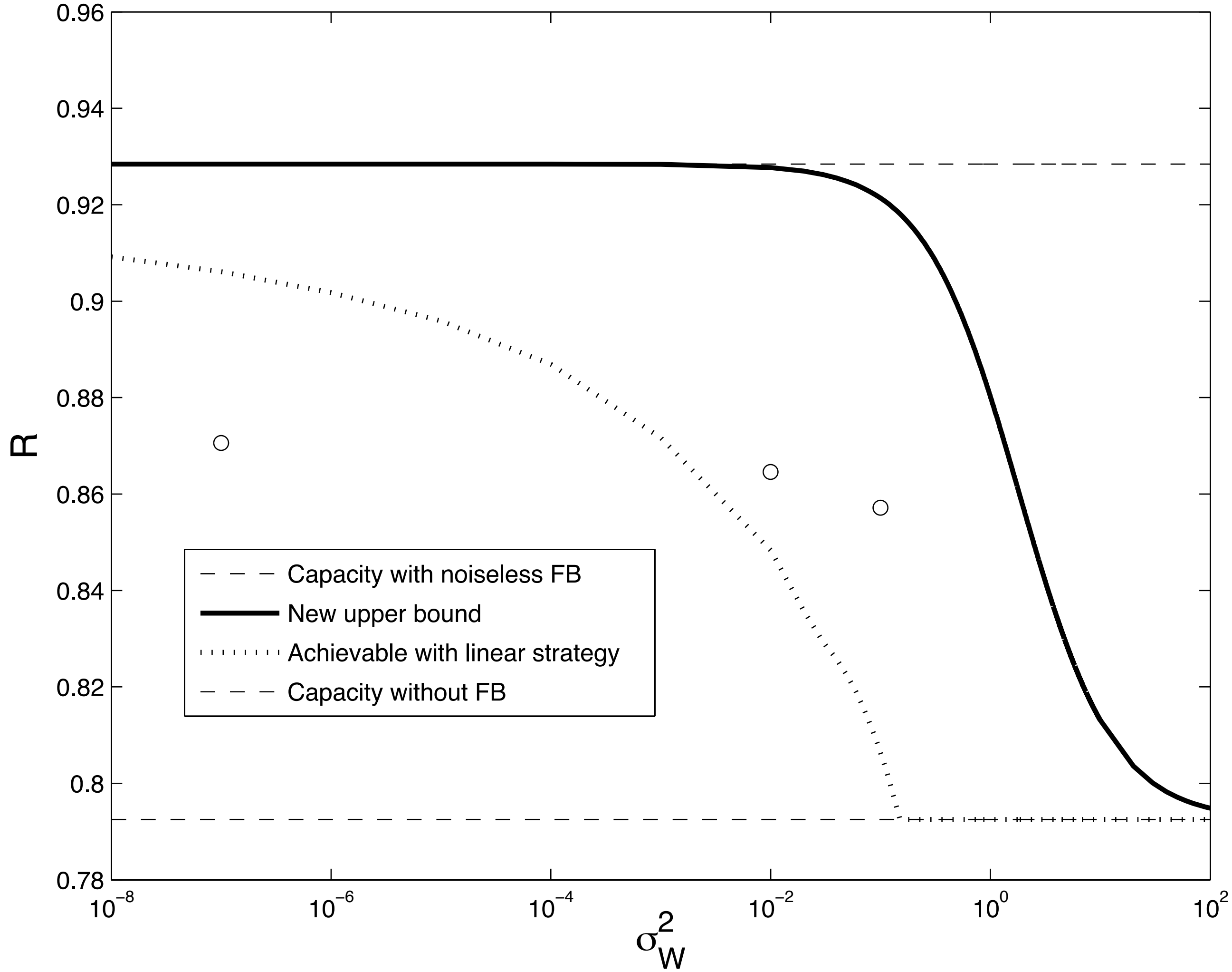
$$0 \leq I(M_1; M_2 | \tilde{Y}_1^n, \tilde{Y}_2^n) - I(M_1; M_2)$$

which leads to

$$I(X_1; X_2 | T) \leq I(X_1; X_2 | \tilde{Y}_1, \tilde{Y}_2, T)$$

The remaining difficulty is to show that it is sufficient to consider only Gaussian distributions.

Dependence-Balance Bound — Noisy Feedback



Gastpar/Kramer, *International Zurich Seminar*, February 2006.

Dependence-Balance Bound — Many Users

For the Gaussian MAC with feedback and M users, the cut-set bound shows that

$$\sum_{m=1}^{M} R_m \le \frac{1}{2}\log\left(1 + M^2 P\right)$$

Ozarow's *send-innovations-only* strategy has been extended to the M-user case by Kramer (*IEEE Trans Info Theory*, 2002). It can be shown to achieve (for large M)

$$\sum_{m=1}^{M} R_m \approx \frac{1}{2}\log M + \frac{1}{2}\log\log M.$$

So, is there a strategy that improves on this?

Dependence-Balance Bound — Many Users

Using Dependence-Balance bound, one can show that the answer is negative — capacity is achieved via Kramer's strategy.

Many different dependence-balance conditions can be written. One of them is

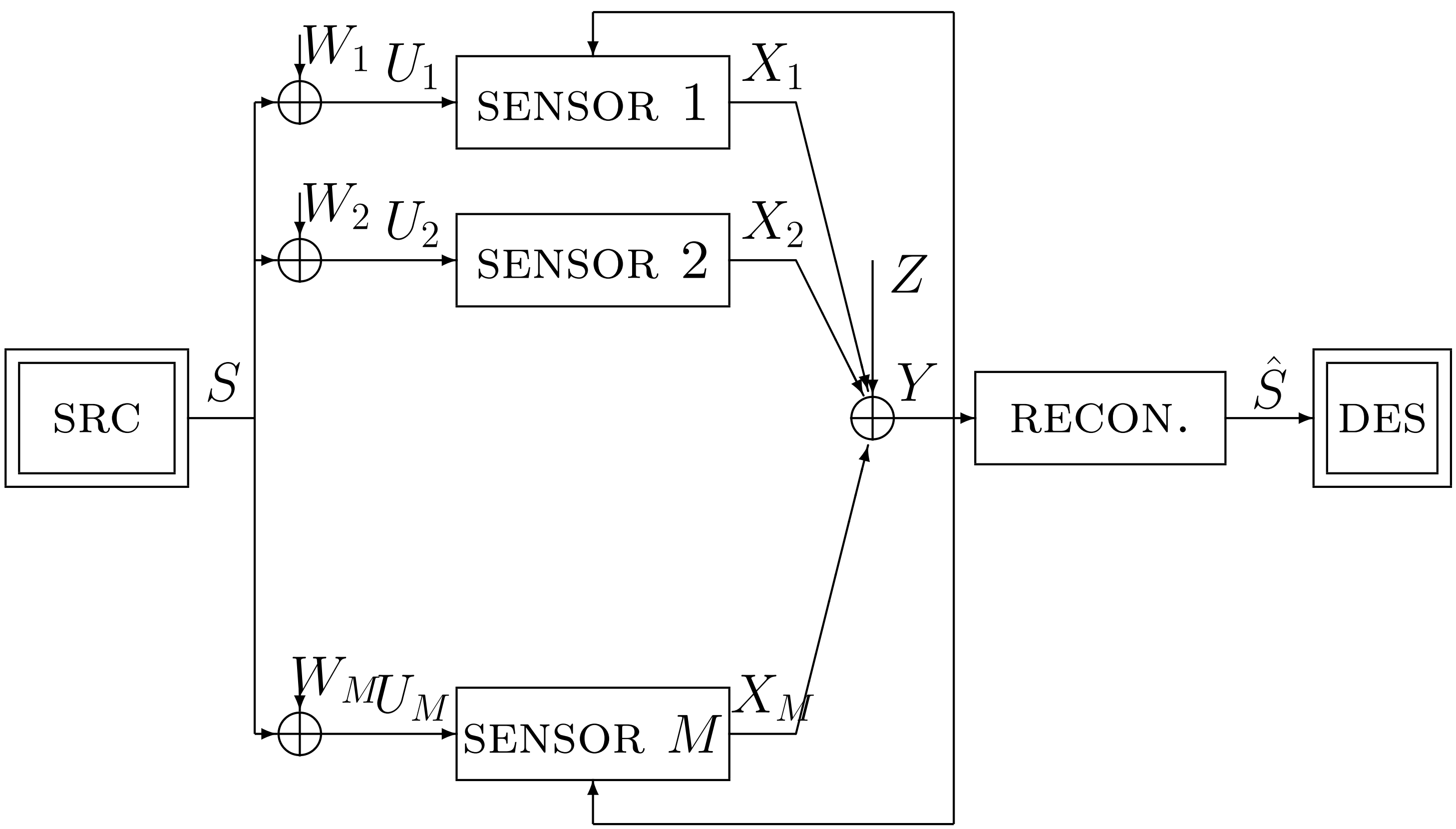
$$I(X_1, X_2, \ldots, X_M; Y|T) \leq \frac{1}{M-1} \sum_{m=1}^{M} I(X_{\setminus m}; Y|X_m, T)$$

Again, the difficulty is to show that Gaussian distributions are sufficient.

Kramer/Gastpar, *IEEE Information Theory Workshop,* March 2006.

3. Sensor Network 101 with Feedback

Instead of independent messages...



Sensor Network 101 with Feedback

- With one channel use per source sample, it is known that:
  - Analog communication gives

$$D = \frac{\sigma_S^2}{1 + \frac{M\sigma_S^2}{\sigma_W^2}} + \frac{\sigma_S^2}{1 + \frac{\sigma_W^2}{M\sigma_S^2}} \cdot \frac{1}{1 + \frac{Q}{\sigma_Z^2}}$$

    and this is optimal. (Gastpar/Vetterli 2002)
  - Feedback is useless.
  - Digital communication (distributed source coding plus channel coding) gives

$$D_{sep} \geq \frac{\sigma_S^2 \sigma_W^2}{\sigma_S^2 \log_2 \left(1 + \frac{Q}{\sigma_Z^2}\right) + \sigma_W^2}.$$

Sensor Network 101 with Feedback

- With *two* channel uses per source sample, we observe:
  - Analog communication with repetition coding gives

$$D = \frac{\sigma_S^2}{1 + \frac{M\sigma_S^2}{\sigma_W^2}} + \frac{\sigma_S^2}{1 + \frac{\sigma_W^2}{M\sigma_S^2}} \cdot \frac{1}{1 + \frac{2Q}{\sigma_Z^2}}$$

    but this **cannot** be shown to be optimal.

  - Digital communication (distributed source coding plus channel coding) gives

$$D_{sep} \geq \frac{\sigma_S^2 \sigma_W^2}{\sigma_S^2 2 \log_2 \left(1 + \frac{Q}{\sigma_Z^2}\right) + \sigma_W^2}.$$

The big mistery is that **without** feedback, we don't know how to exploit the additional channel use...!

Sensor Network 101 with Feedback

- With *two* channel uses per source sample, we observe:
  - Analog communication with repetition coding gives

$$D = \frac{\sigma_S^2}{1+\frac{M\sigma_S^2}{\sigma_W^2}} + \frac{\sigma_S^2}{1+\frac{\sigma_W^2}{M\sigma_S^2}} \cdot \frac{1}{1+\frac{2Q}{\sigma_Z^2}}$$

    but this **cannot** be shown to be optimal.

  - Feedback gets us back to provably optimal performance!

$$D = \frac{\sigma_S^2}{1+\frac{M\sigma_S^2}{\sigma_W^2}} + \frac{\sigma_S^2}{1+\frac{\sigma_W^2}{M\sigma_S^2}} \cdot \left(\frac{1}{1+\frac{Q}{\sigma_Z^2}}\right)^2$$

Conclusions

(Noisy) feedback may be feasible in sensor networks...

We have shown:

- Non-trivial upper bounds on the performance of noisy feedback
- With noiseless feedback, *sending innovations only* is optimal even for M users.