Estimation over the collision channel & Observation-driven sensor scheduling

Marcos Vasconcelos

mvasconc@usc.edu

www-bcf.usc.edu/~mvasconc

Nuno Martins (UMD) **Ashutosh Nayyar (USC)** **Urbashi Mitra (USC)**

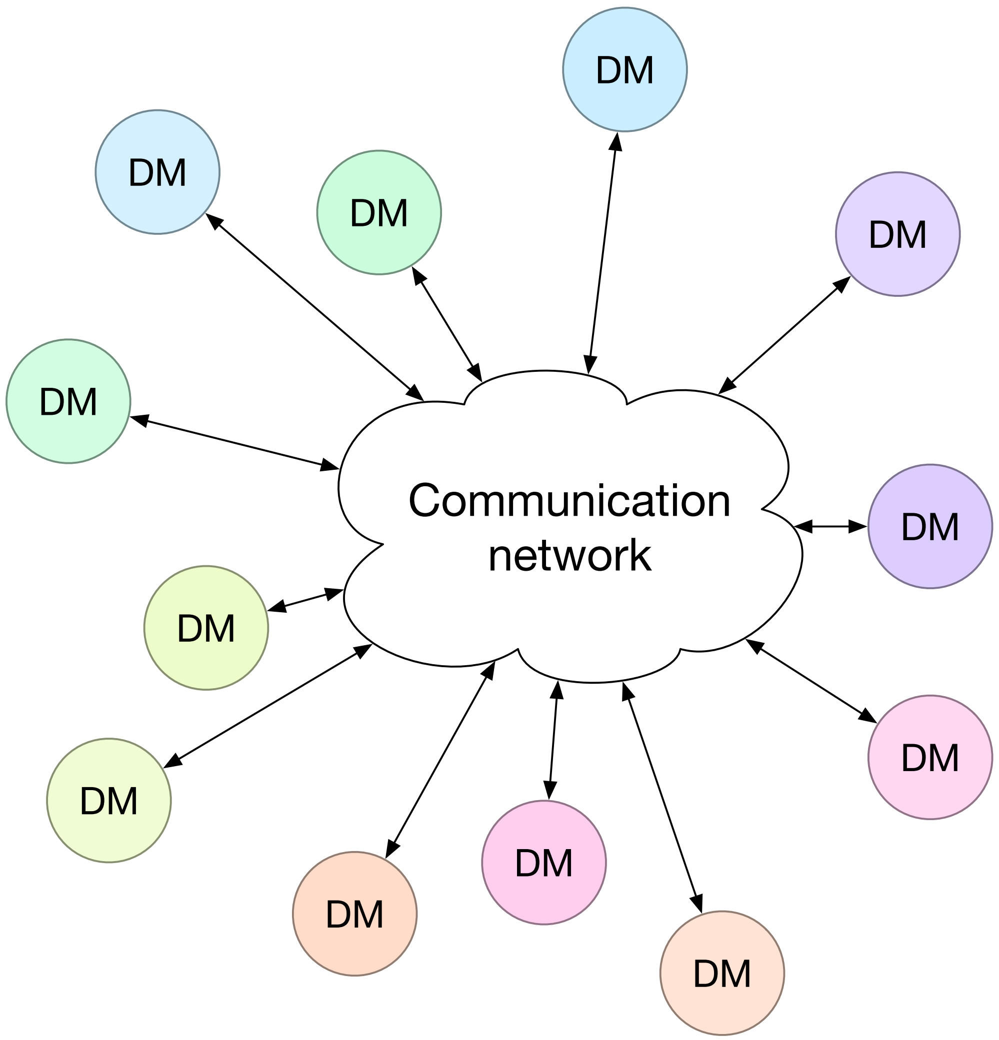
Dept. of Electrical Engineering

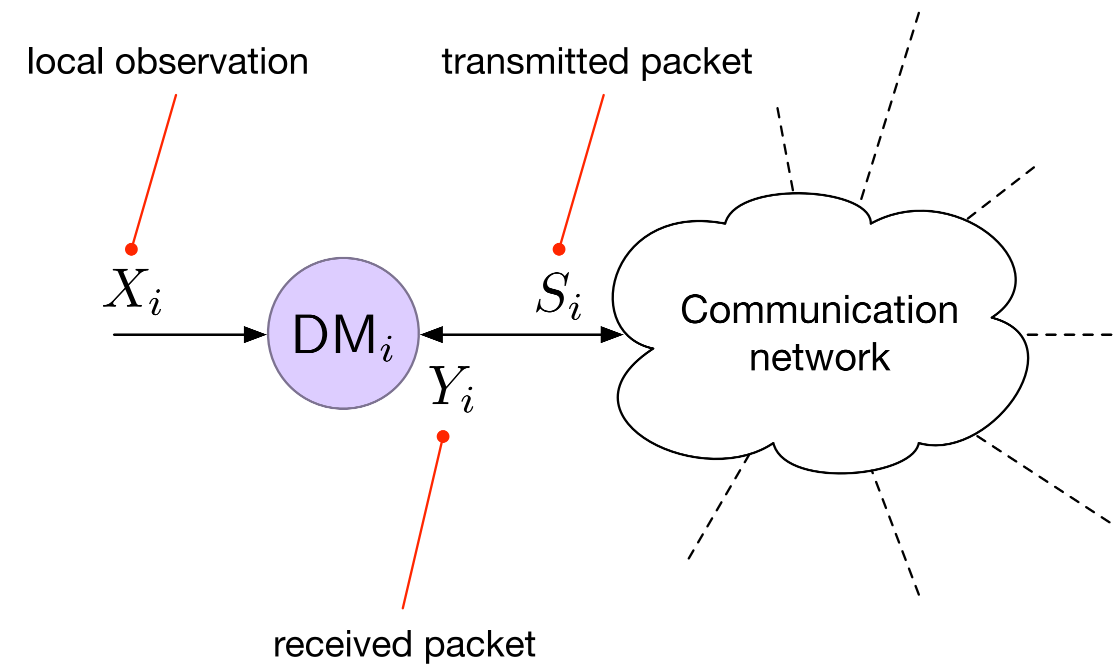
University of Southern California

Carnegie Mellon University

March 20th, 2018 - Pittsburgh, PA

Networked decision systems





Many applications

1. **Networked control**
2. **Remote estimation**
3. **Sensor networks**
4. **Robotic networks**

Many challenges

Communication is imperfect:
Delays, noise, quantization, congestion, packet drops, connectivity and **interference**

Remote sensing

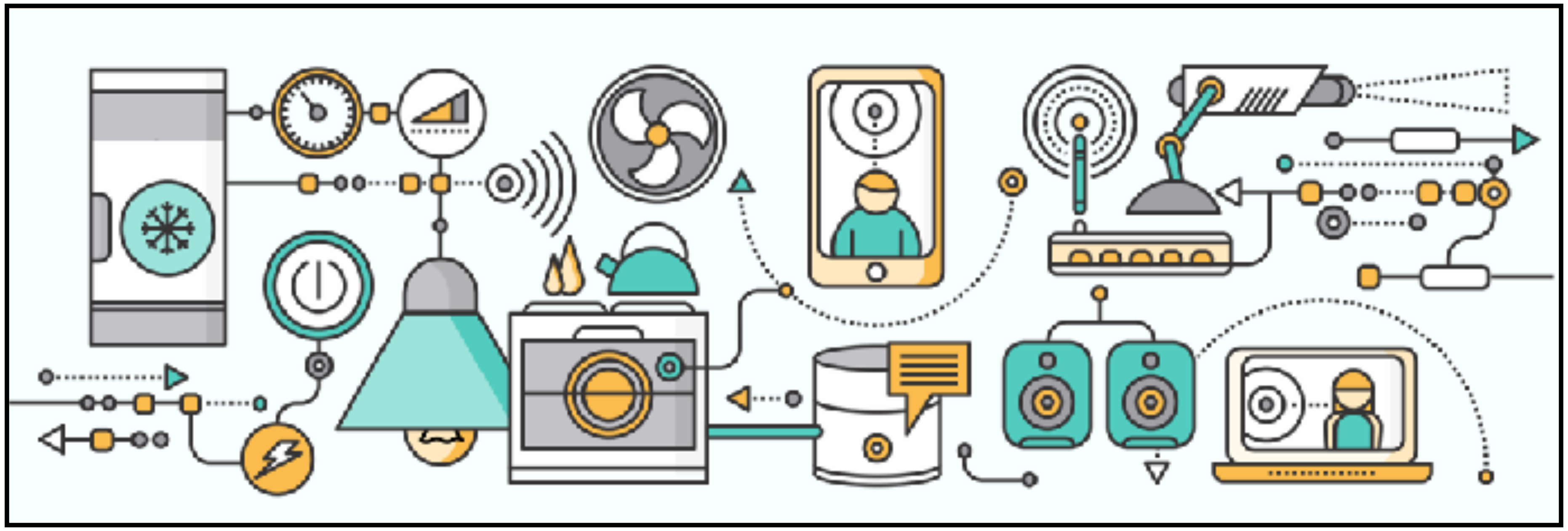
- **Remote monitoring of one-time catastrophic events**

Oil & gas pipelines

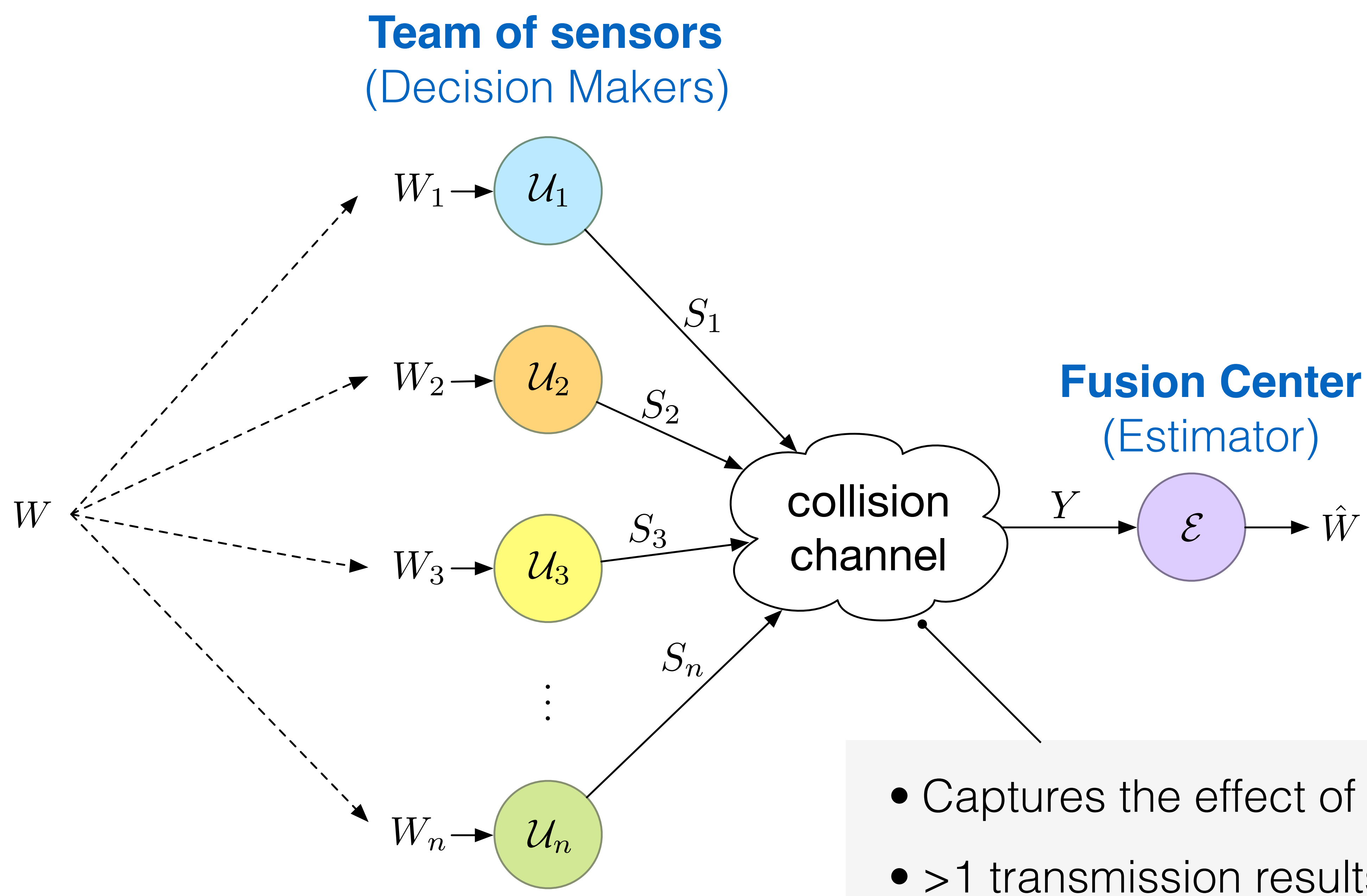
Powerlines

Bridges

- **Real-time wireless networking for the Internet-of-Things**



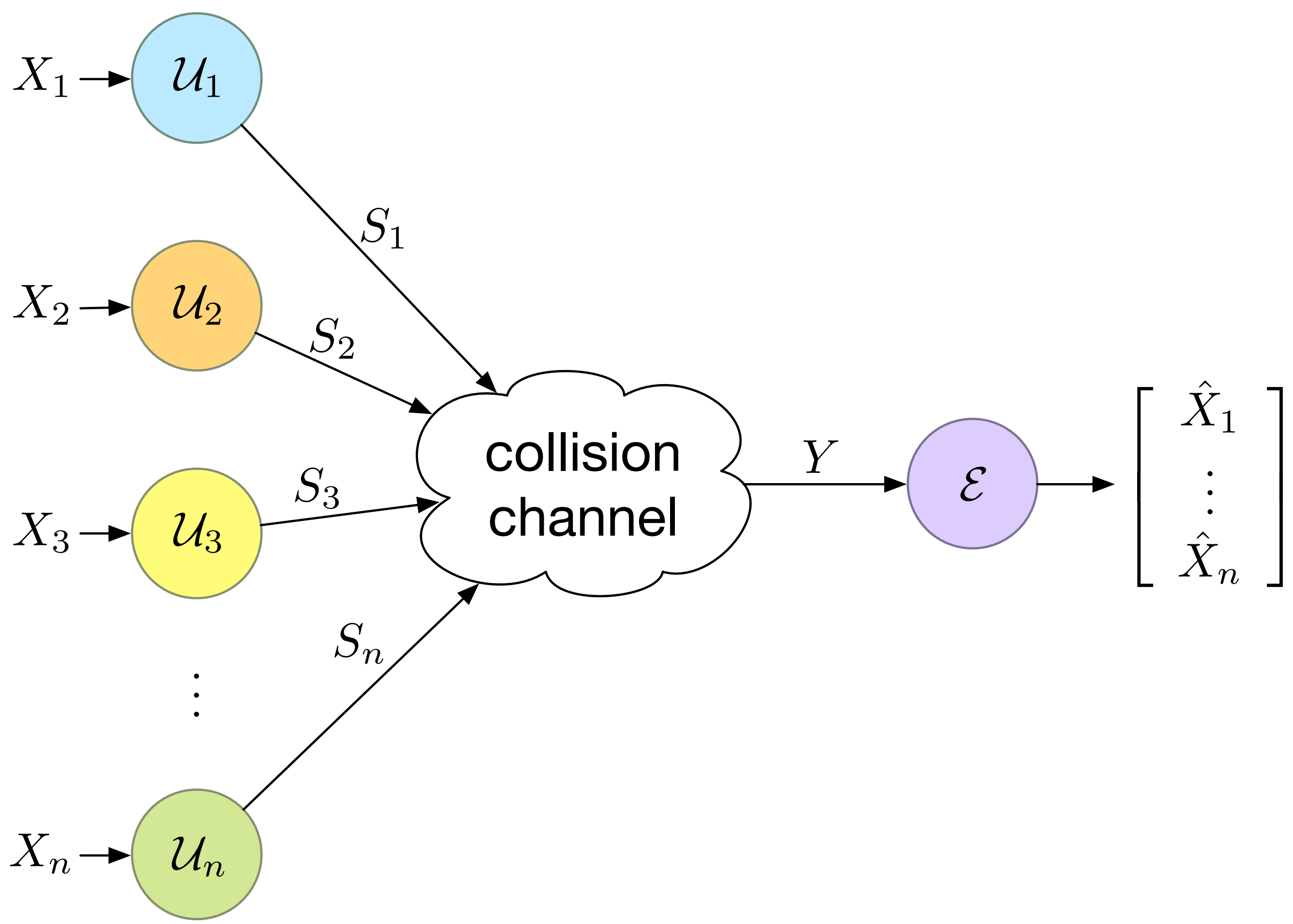
Basic framework



- Captures the effect of **interference**
- >1 transmission results in a **collision**

Design jointly optimal communication and estimation policies

Estimation over the collision channel



Observations

$$X_i \sim f_{X_i}$$

$$X_i \perp\!\!\!\perp X_j$$

$$f_{X_i}(x) > 0, \quad x_i \in \mathbb{R}$$

Decision variables

$$U_i \in \{0, 1\}$$

Stay silent

$$S_i = \varnothing$$

Transmit

$$S_i = (i, X_i)$$

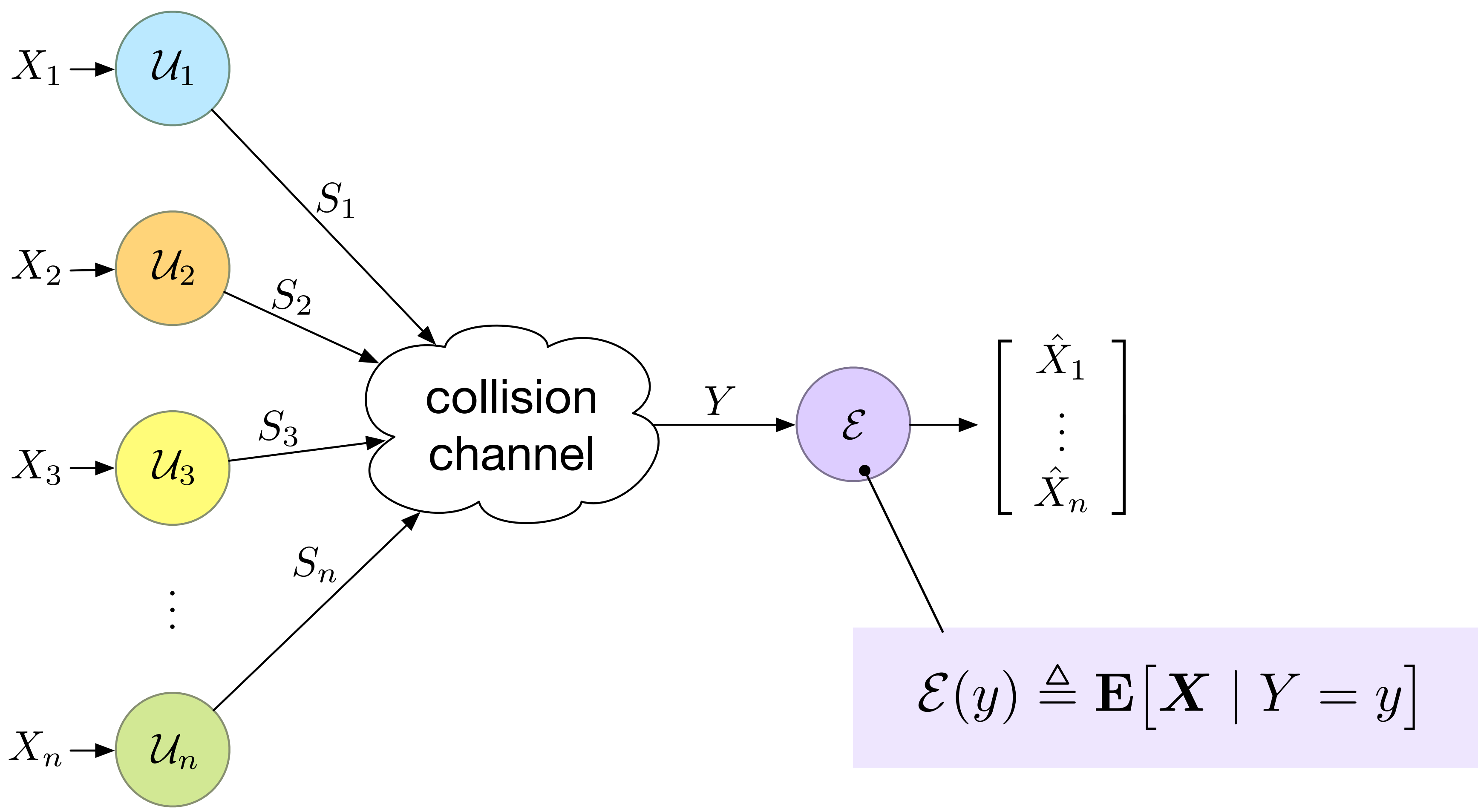
Communication policy

$$\mathbf{P}(U_i = 1 \mid X_i = x_i) = \mathcal{U}_i(x_i)$$

Estimation policy

$$\hat{\boldsymbol{X}} = \mathcal{E}(y)$$

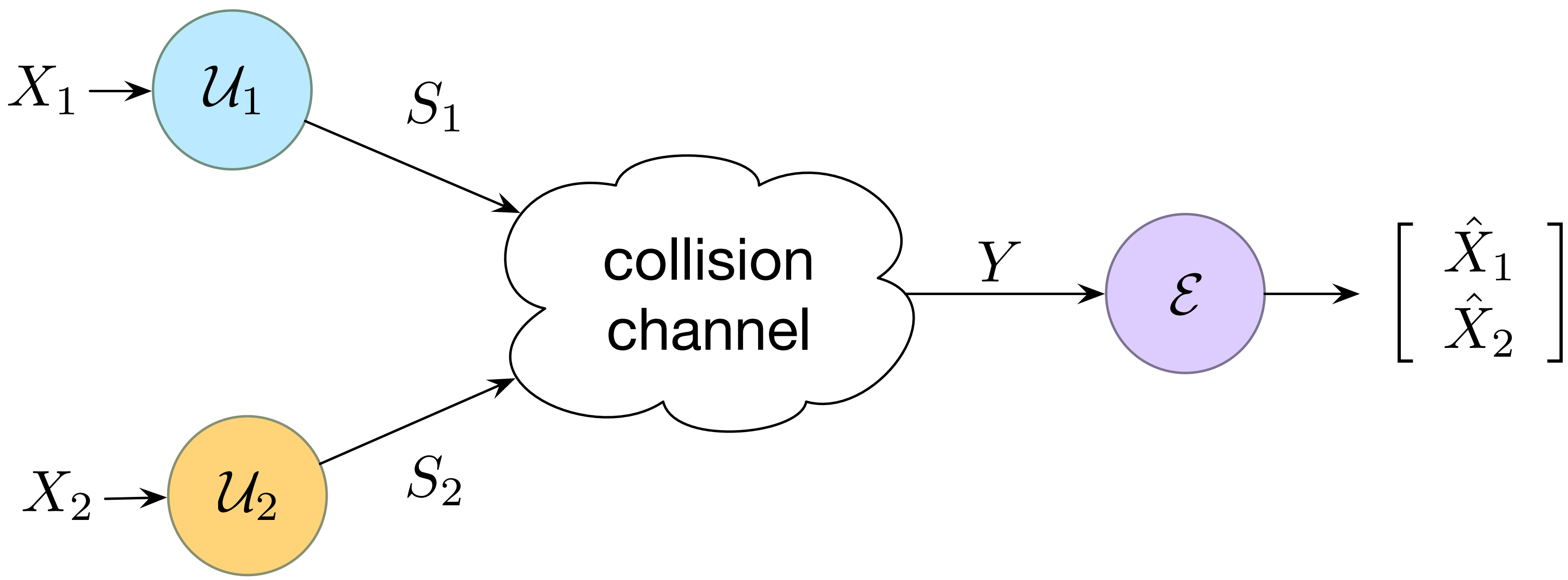
Estimation over the collision channel



Find a strategy $(\mathcal{U}_1^\star, \cdots, \mathcal{U}_n^\star)$ that jointly minimizes the following cost

$$\mathcal{J}(\mathcal{U}_1, \ldots, \mathcal{U}_n) = \mathbf{E}\left[\sum_{i=1}^{n} (X_i - \hat{X}_i)^2\right]$$

Simplest case: two sensors



$$\mathbf{P}(U_i = 1 \mid X_i = x_i) = \mathcal{U}_i(x_i)$$

$$\mathbb{U}_i = \{\mathcal{U} \mid \mathcal{U} : \mathbb{R} \to [0,1]\}, \quad i \in \{1,2\}$$

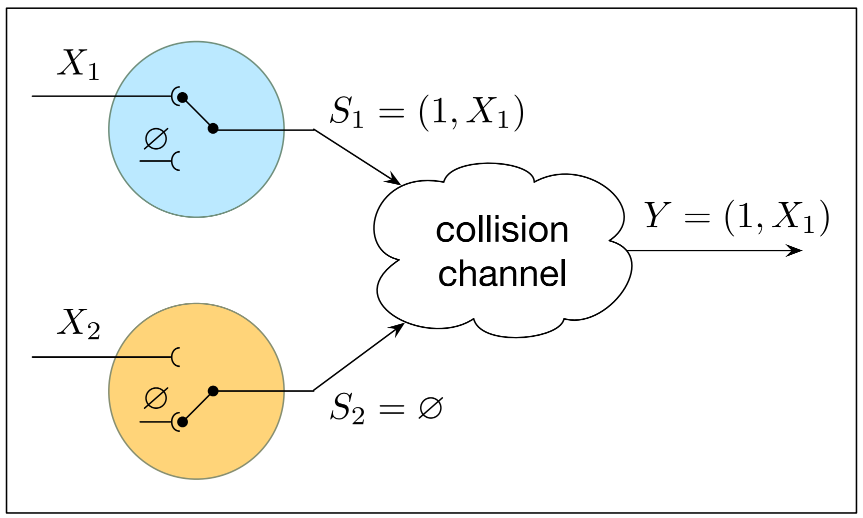
Problem 1

$$\min_{(\mathcal{U}_1,\mathcal{U}_2)\in\mathbb{U}_1\times\mathbb{U}_2} \quad \mathcal{J}(\mathcal{U}_1,\mathcal{U}_2) = \mathbf{E}\left[(X_1 - \hat{X}_1)^2 + (X_2 - \hat{X}_2)^2\right]$$

Collision channel

single transmission

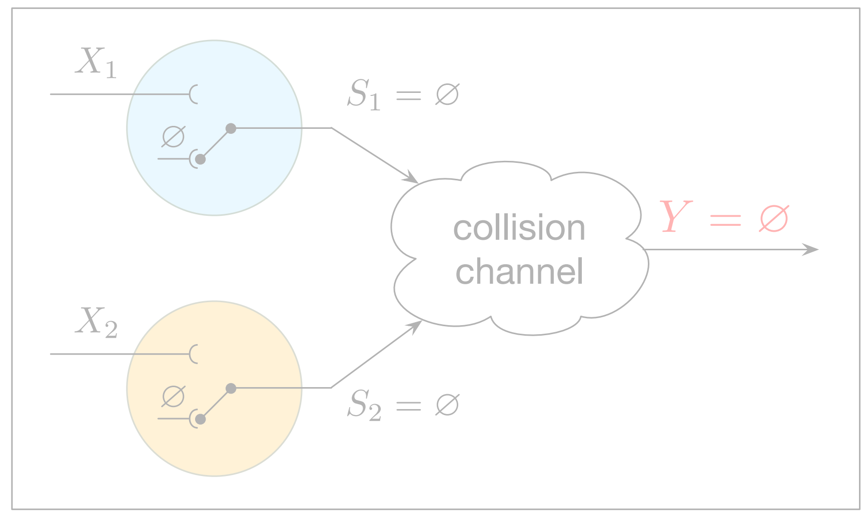
$U_1 = 1, U_2 = 0$



success!

no transmissions

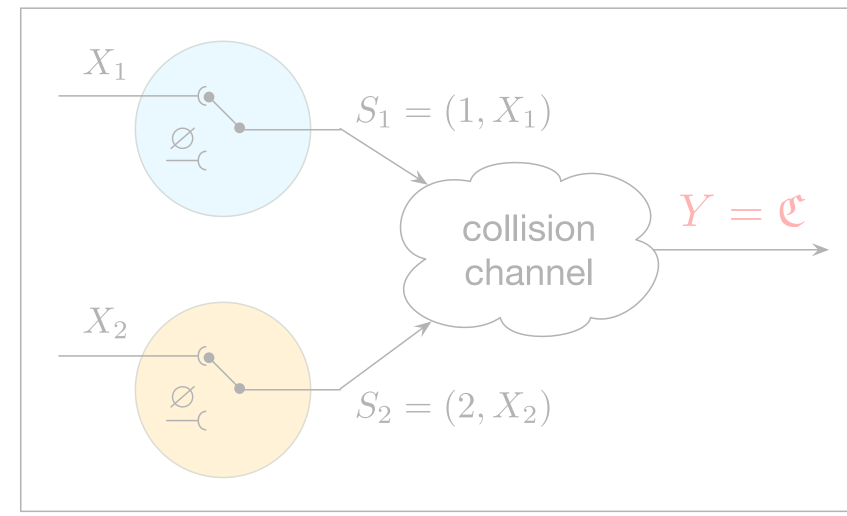
$U_1 = 0, U_2 = 0$



no transmission $\varnothing$

>1 transmissions

$U_1 = 1, U_2 = 1$



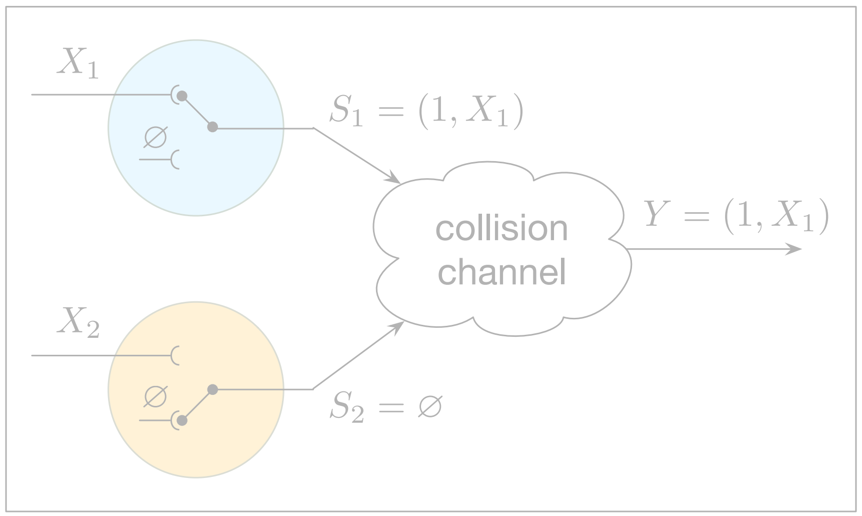
collision $\mathfrak{C}$

From the channel output we can always recover U_1 and U_2

Collision channel

single transmission

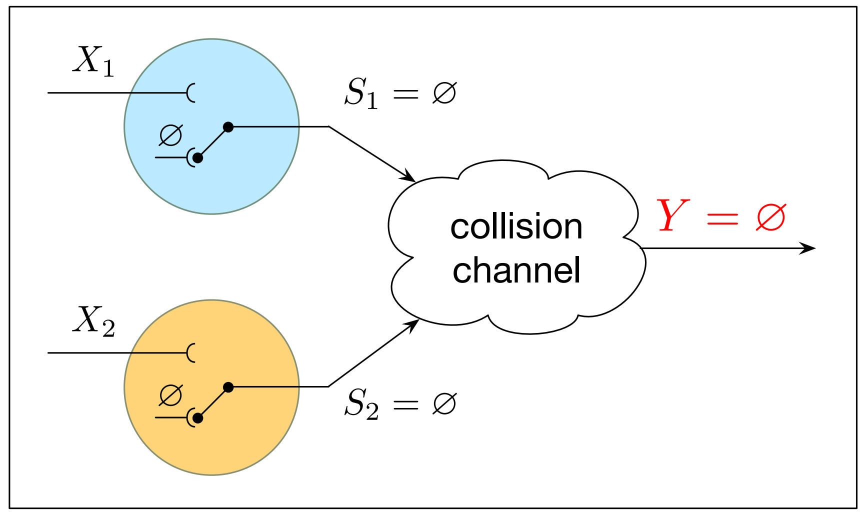
$U_1 = 1, U_2 = 0$



success!

no transmissions

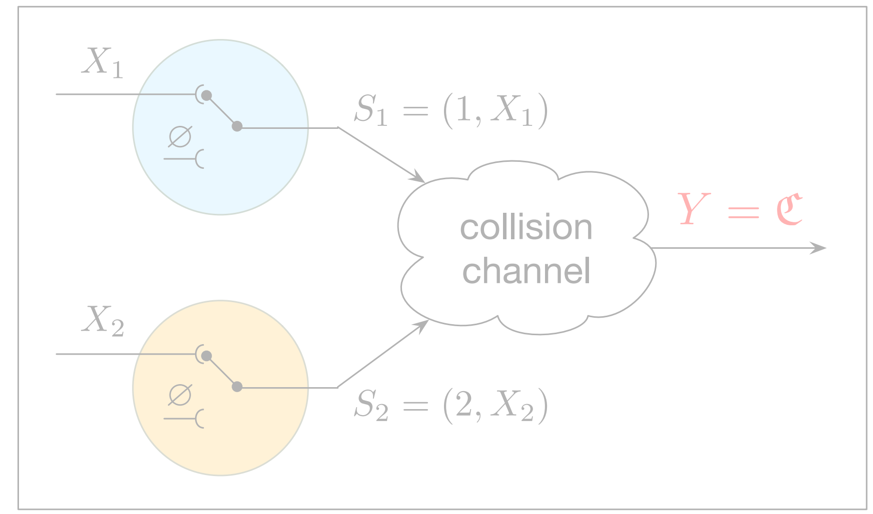
$U_1 = 0, U_2 = 0$



no transmission $\varnothing$

>1 transmissions

$U_1 = 1, U_2 = 1$

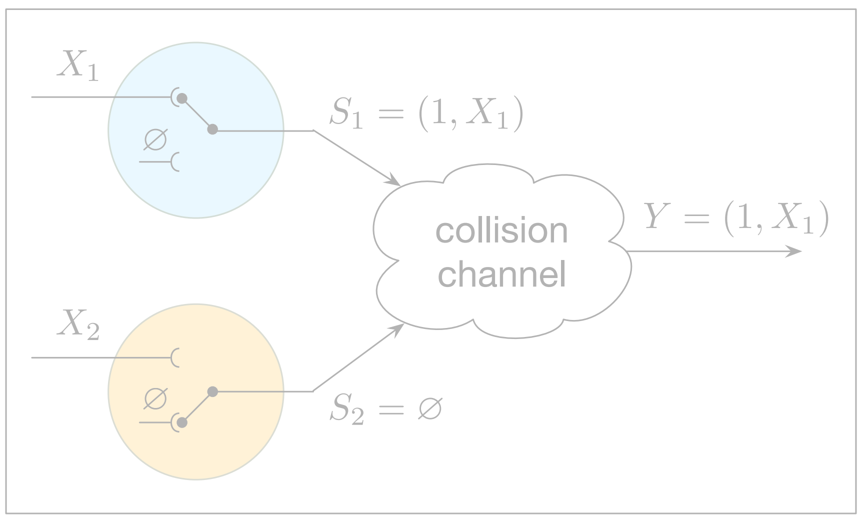


collision $\mathfrak{C}$

From the channel output we can always recover U_1 and U_2

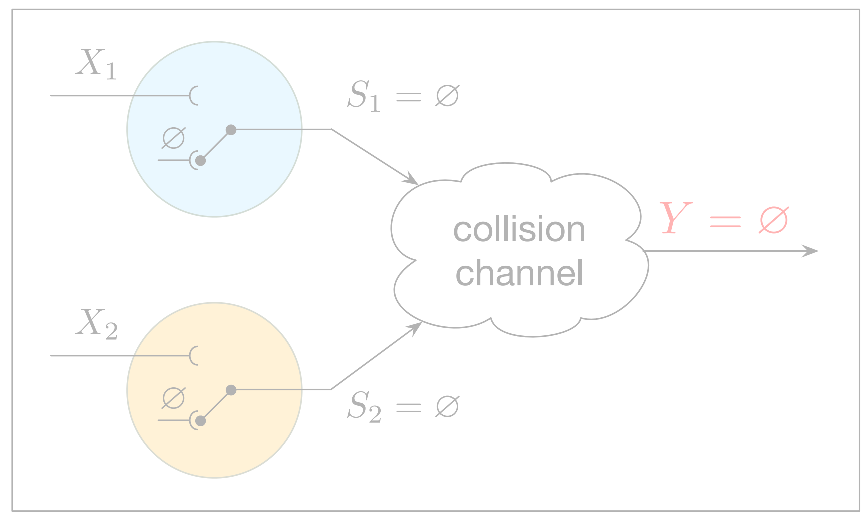
Collision channel

single transmission
$U_1 = 1, U_2 = 0$



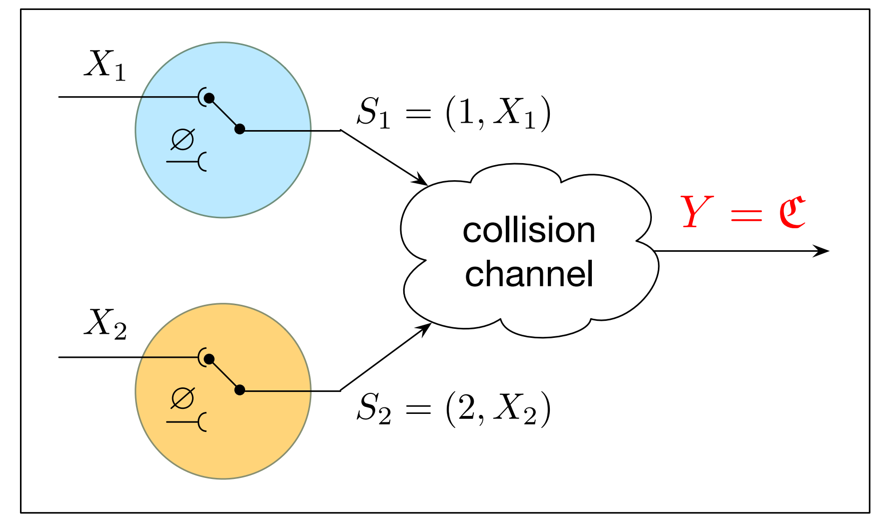
success!

no transmissions
$U_1 = 0, U_2 = 0$



no transmission $\varnothing$

>1 transmissions
$U_1 = 1, U_2 = 1$



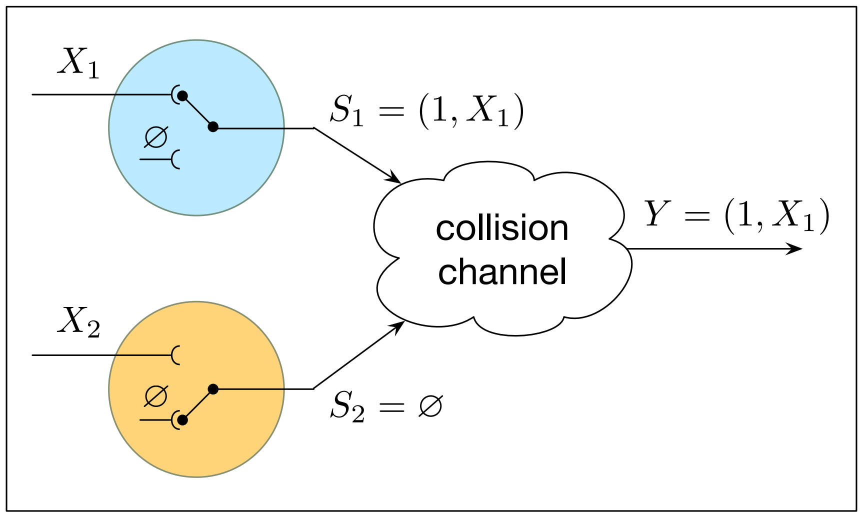
collision $\mathfrak{C}$

From the channel output we can always recover U_1 and U_2

Collision channel

single transmission

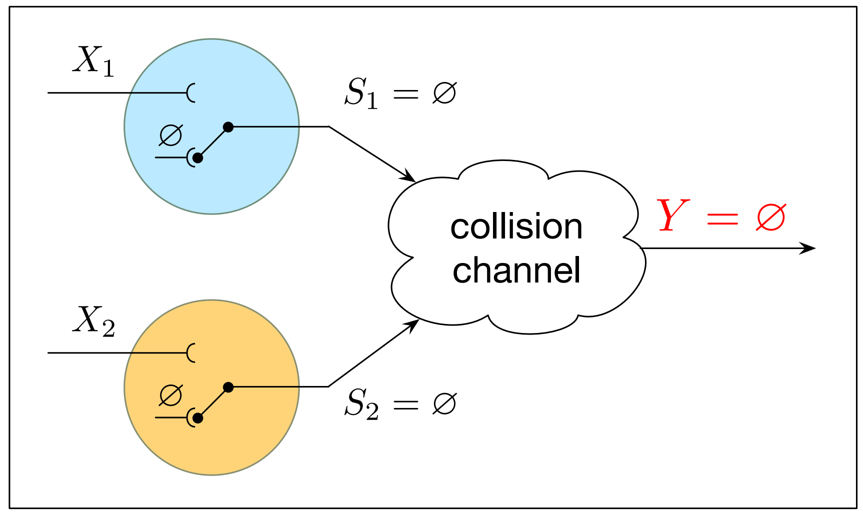
$U_1 = 1, U_2 = 0$



success!

no transmissions

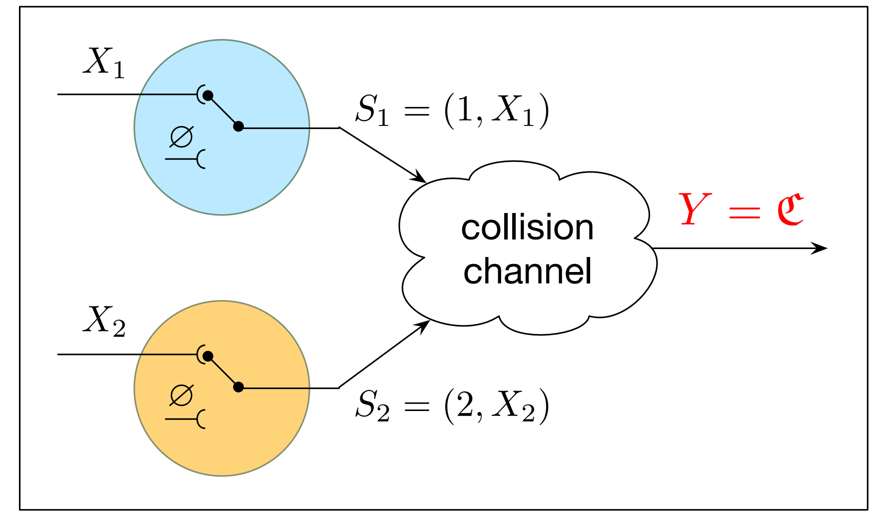
$U_1 = 0, U_2 = 0$



no transmission $\varnothing$

>1 transmissions

$U_1 = 1, U_2 = 1$

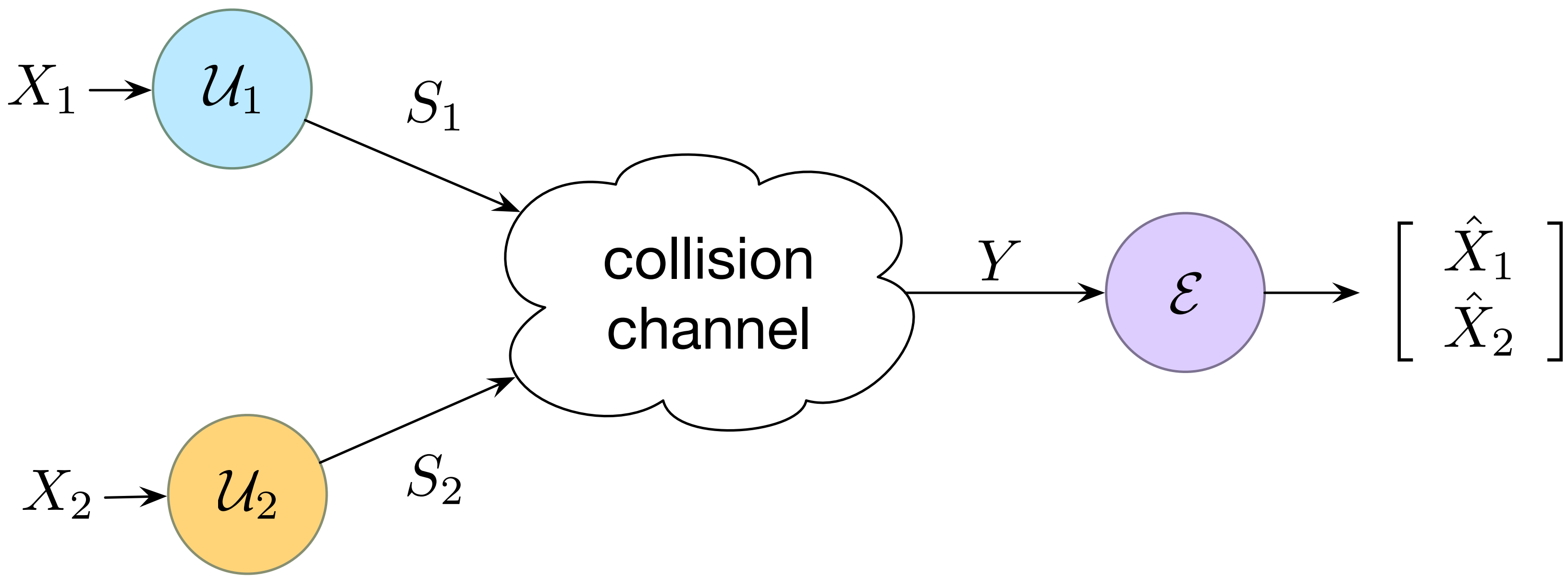


collision $\mathfrak{C}$

The collision channel is fundamentally different from the packet-drop channel[1,2]

1. Sinopoli et al, "Kalman filtering with intermittent observations," *IEEE TAC* 2004
2. Gupta et al, "Optimal LQG control across packet-dropping links," *Systems and Control Letters* 2007

Why is this problem interesting?

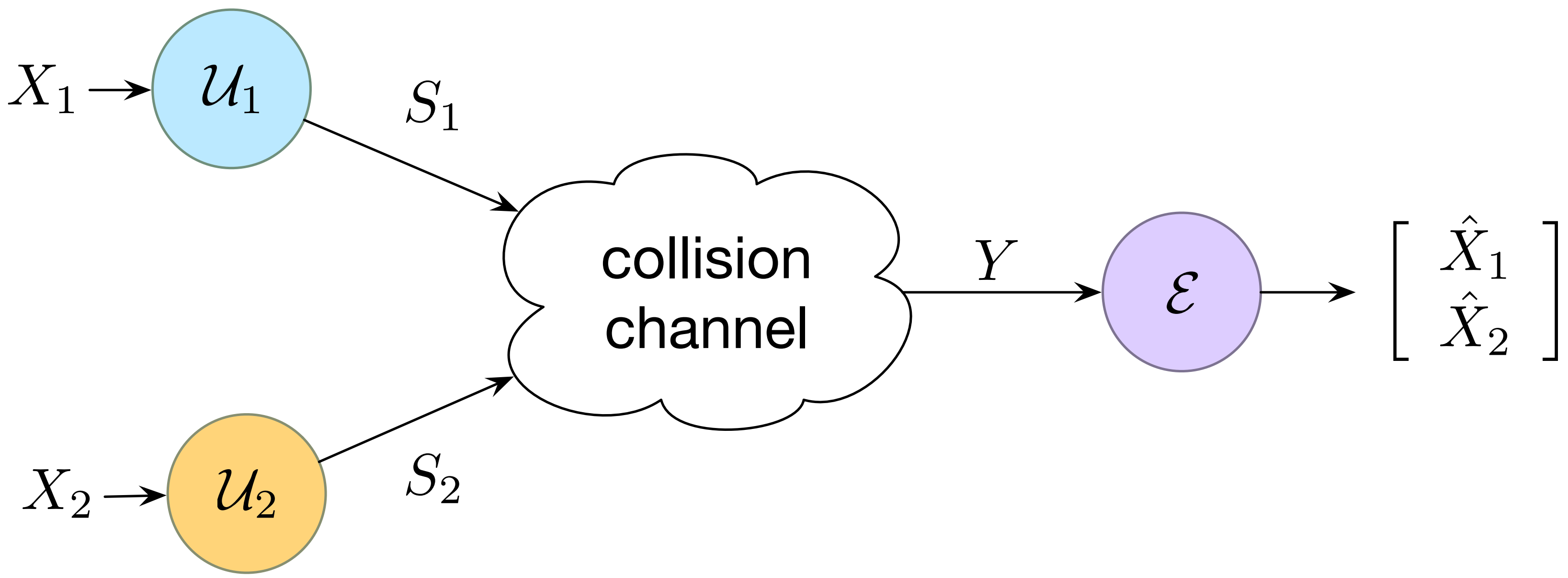


$$\min_{(\mathcal{U}_1,\mathcal{U}_2)\in\mathbb{U}_1\times\mathbb{U}_2} \quad \mathcal{J}(\mathcal{U}_1,\mathcal{U}_2) = \mathbf{E}\left[(X_1-\hat{X}_1)^2 + (X_2-\hat{X}_2)^2\right]$$

Team-decision problem
with **nonclassical information** structure $\Longrightarrow$ **Nonconvex** (in most cases) **intractable**[1,2]

1. Witsenhausen, "A counterexample in optimal stochastic control," *SIAM J. Control* 1968
2. Tsitsiklis & Athans, "On the complexity of decentralized decision making and detection problems," *IEEE TAC* 1985

Why is this problem interesting?

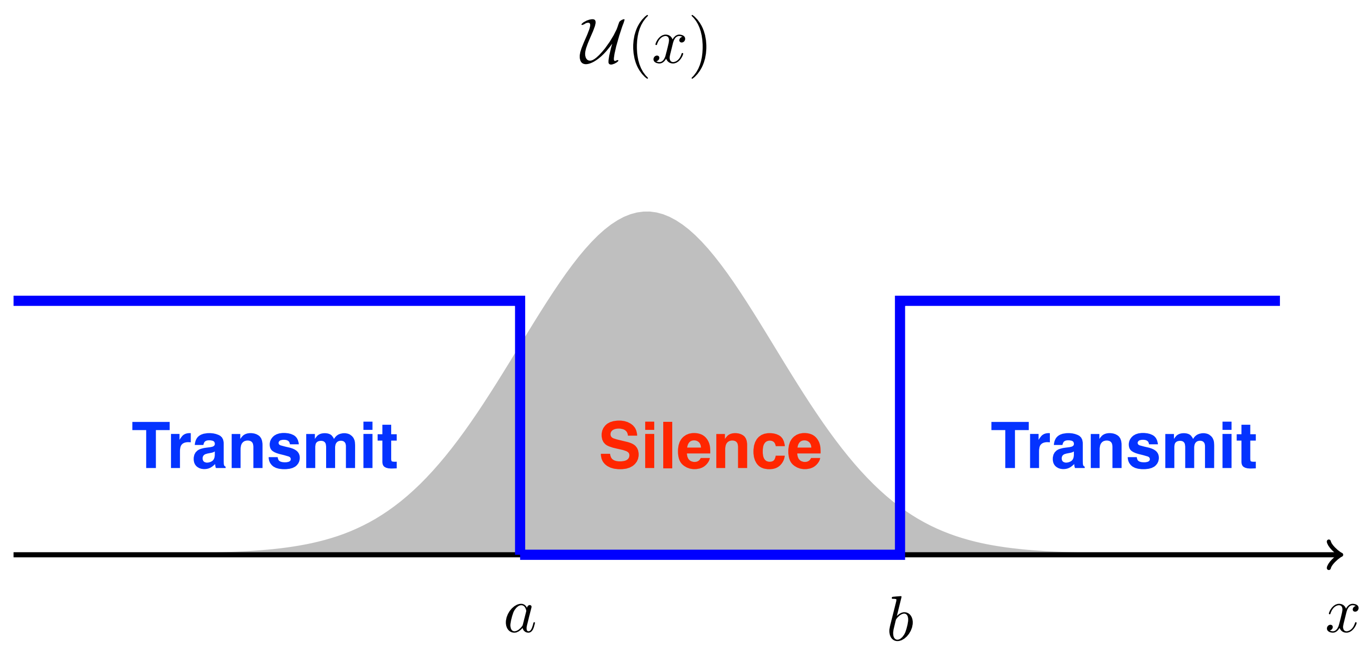


$$\min_{(\mathcal{U}_1,\mathcal{U}_2)\in\mathbb{U}_1\times\mathbb{U}_2} \quad \mathcal{J}(\mathcal{U}_1,\mathcal{U}_2) = \mathbf{E}\left[(X_1-\hat{X}_1)^2 + (X_2-\hat{X}_2)^2\right]$$

Look for a class parametrizable policies that contains an optimal strategy

1. Witsenhausen, "A counterexample in optimal stochastic control," *SIAM J. Control* 1968
2. Tsitsiklis & Athans, "On the complexity of decentralized decision making and detection problems," *IEEE TAC* 1985

Deterministic threshold policies



Threshold policy

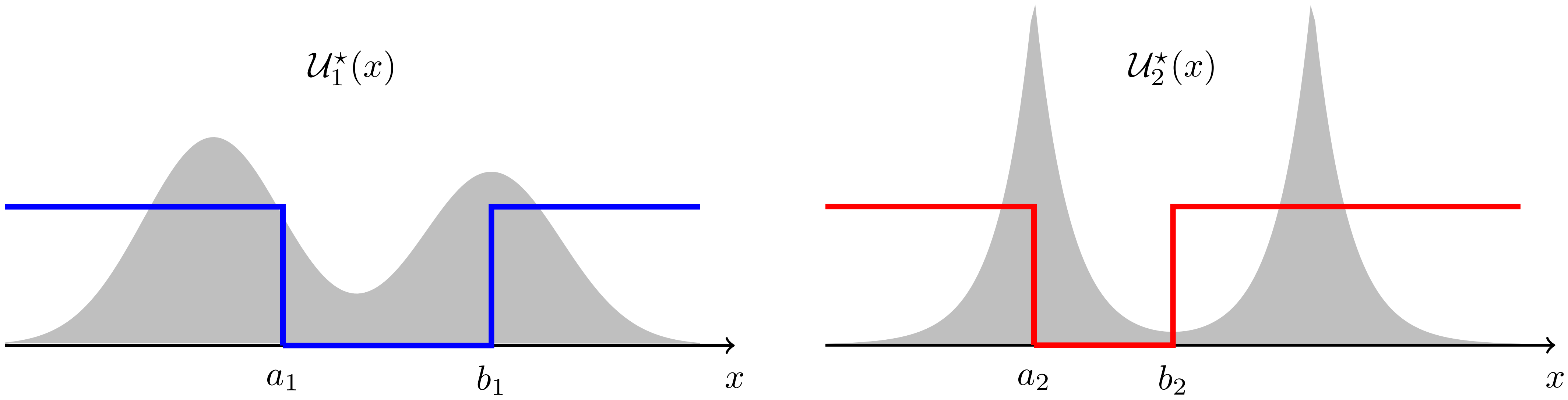
$$\mathcal{U}(x) = \begin{cases} 0 & a \leq x \leq b \\ 1 & \text{otherwise} \end{cases}$$

1. Imer & Basar, "Optimal estimation with limited measurements," *IJSCC* 2010
2. Lipsa & Martins, "Remote state estimation with communication costs for first-order LTI systems," *IEEE TAC* 2011

Characterization of team-optimal policies

Theorem 1

There exists a team optimal pair of **threshold policies** for Problem 1



Sketch of Proof:

- Step 1: Equivalent single DM problem
- Step 2: Lagrange duality for infinite dimensional LPs

Main idea

Team-optimality

$$\mathcal{J}(\mathcal{U}_1^*,\mathcal{U}_2^*) \leq \mathcal{J}(\mathcal{U}_1,\mathcal{U}_2), \quad (\mathcal{U}_1,\mathcal{U}_2) \in \mathbb{U}_1 \times \mathbb{U}_2$$

$$\Longrightarrow$$

$$\not\Longleftarrow$$

Person-by-person optimality

$$\mathcal{J}(\mathcal{U}_1^*,\mathcal{U}_2^*) \leq \mathcal{J}(\mathcal{U}_1,\mathcal{U}_2^*), \quad \mathcal{U}_1 \in \mathbb{U}_1$$

$$\mathcal{J}(\mathcal{U}_1^*,\mathcal{U}_2^*) \leq \mathcal{J}(\mathcal{U}_1^*,\mathcal{U}_2), \quad \mathcal{U}_2 \in \mathbb{U}_2$$

$$(\mathcal{U}_1^*,\mathcal{U}_2^*) \in \mathbb{U}_1 \times \mathbb{U}_2 \quad \longrightarrow \quad (\breve{\mathcal{U}}_1^*,\breve{\mathcal{U}}_2^*) \in \mathbb{U}_1 \times \mathbb{U}_2$$

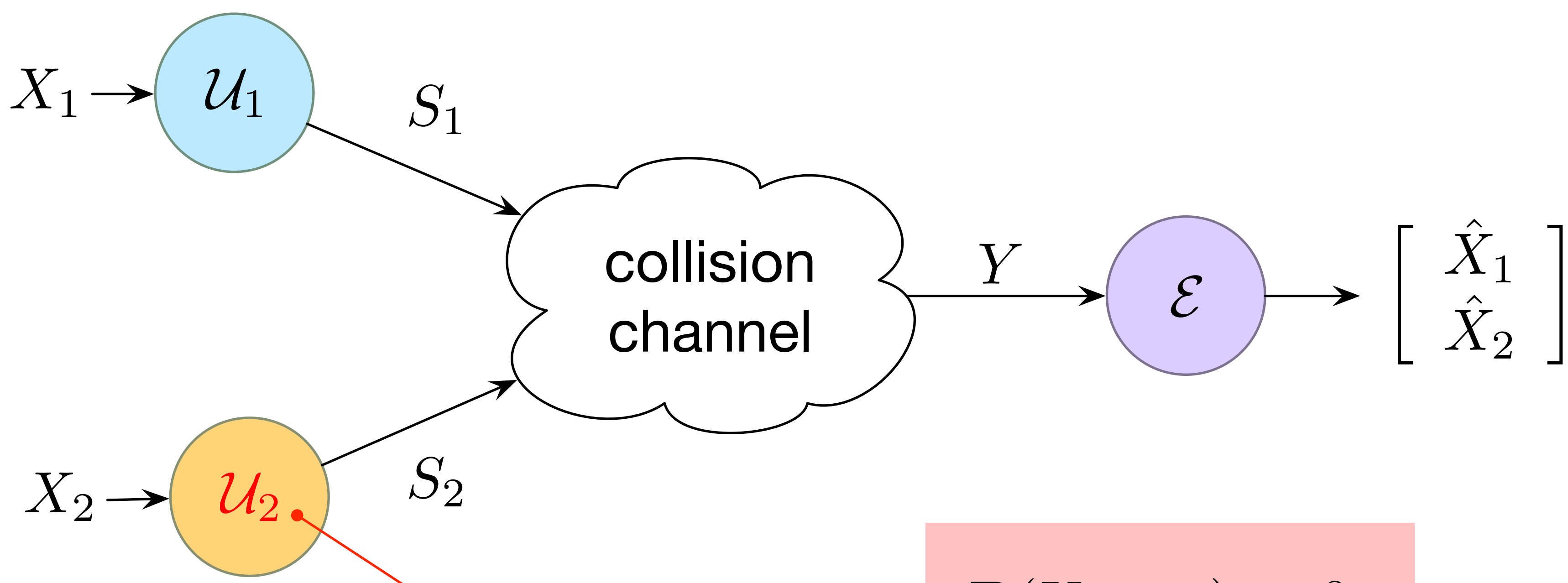
$$\mathcal{J}(\mathcal{U}_1^*,\mathcal{U}_2^*) \geq \mathcal{J}(\breve{\mathcal{U}}_1^*,\breve{\mathcal{U}}_2^*)$$

threshold policies

Given any pair of person-by-person optimal policies
construct a new pair with **equal or better cost**,
where each policy is **threshold**

1. Yuksel & Basar, *Stochastic networked control systems,* Birkhauser 2013
2. Mahajan et al, "Information structures in optimal decentralized control," CDC 2012

Remote estimation with communication costs



$\Rightarrow$ $\mathbf{P}(U_2 = 1) = \beta_2$ **is constant**

Original cost:

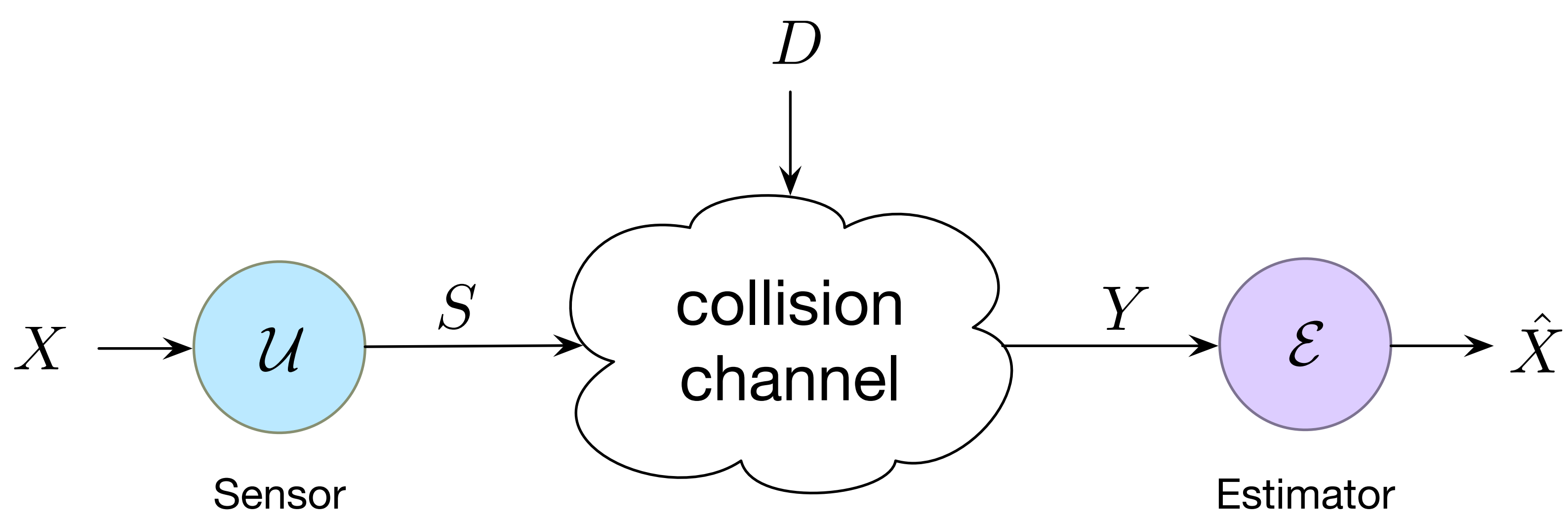
$$\mathcal{J}(\mathcal{U}_1, \mathcal{U}_2) = \mathbf{E}\Big[(X_1 - \hat{X}_1)^2 + (X_2 - \hat{X}_2)^2\Big]$$

Cost from the perspective of DM$_1$:

$$\mathcal{J}_1(\mathcal{U}_1) = \mathbf{E}\Big[(X_1 - \hat{X}_1)^2\Big] + \rho_2 \cdot \mathbf{P}(U_1 = 1) + \theta_2$$

do not depend on $\mathcal{U}_1$

Single DM subproblem



$$D \sim \mathcal{B}(\beta)$$

Determines if the channel is occupied or not

$$X \perp\!\!\!\perp D$$

Problem 2

$$\min_{\mathcal{U} \in \mathbb{U}} \; \mathcal{J}(\mathcal{U}) = \mathbf{E}\big[(X - \hat{X})^2\big] + \rho \cdot \mathbf{P}(U = 1)$$

$$\mathbf{P}(U = 1 \mid X = x) = \mathcal{U}(x) \qquad \mathbb{U} = \{\mathcal{U} \mid \mathcal{U} : \mathbb{R} \to [0,1]\}$$

Lemma

There exists an optimal **threshold policy** for Problem 2

Sketch of Proof

1. Express the cost as

$$\mathcal{J}(\mathcal{U}) = \mathbf{E}\Big[\beta(X - \hat{x}_{\mathfrak{C}})^2 + \rho \mid U = 1\Big] \cdot \mathbf{P}(U = 1) + \mathbf{E}\Big[(X - \hat{x}_{\varnothing})^2 \mid U = 0\Big] \cdot \mathbf{P}(U = 0)$$

$$\hat{x}_{\mathfrak{C}} = \mathbf{E}[X|U = 1] \qquad \hat{x}_{\varnothing} = \mathbf{E}[X|U = 0]$$

2. After **introducing two linear constraints** and a **change of variables**, we have:

$$\mathbf{P}(U = 1) = \alpha$$
$$\mathbf{E}[X|U = 0] = \gamma$$

$$\mathcal{G}(x) = \frac{1 - \mathcal{U}(x)}{1 - \alpha}$$

Sketch of Proof

moment optimization problem
with variable bounds

$$
\begin{aligned}
\underset{\mathcal{G}\in L^2_\mu(\mathbb{R})}{\text{minimize}} \quad & \mathbf{E}\big[X^2\mathcal{G}(X)\big] \\
\text{subject to} \quad & \mathbf{E}\big[X\mathcal{G}(X)\big] = \gamma \\
& \mathbf{E}[\mathcal{G}(X)] = 1 \\
& 0 \leq \mathcal{G}(x) \leq \frac{1}{1-\alpha}
\end{aligned}
$$

convex

1. Akhiezer, *The Classical Moment Problem,* 1965
2. Byrnes & Lindquist, "A convex optimization approach to generalized moment problems," Springer 2003

Sketch of Proof

3. The Lagrange dual function is

$$\mathcal{C}^*(\nu) = -\nu_1 - \nu_0\gamma - \frac{1}{1-\alpha}\mathbf{E}\left[[X^2 + \nu_0 X + \nu_1]^-\right]$$

strong duality holds[1,2]

4. The solution to the primal problem is

$$\mathcal{G}_{\nu^*}(x) = \begin{cases} \frac{1}{1-\alpha} & \text{if } x^2 + \nu_0^* x + \nu_1^* \leq 0 \\ 0 & \text{otherwise} \end{cases}$$

5. In the original optimization variable:

$$\mathcal{U}_{\nu^*}(x) = \begin{cases} 0 & \text{if } x^2 + \nu_0^* x + \nu_1^* \leq 0 \\ 1 & \text{otherwise} \end{cases} \implies \mathcal{U}^*(x) = \begin{cases} 0 & \text{if } a \leq x \leq b \\ 1 & \text{otherwise} \end{cases}$$

■

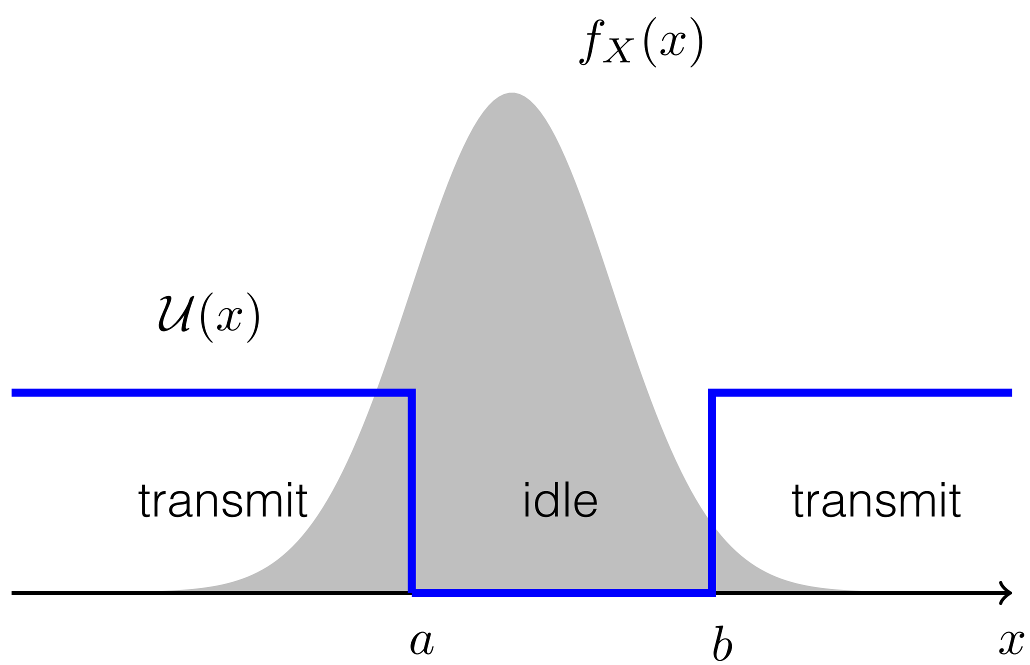
1. Borwein & Lewis, *Math. Prog.* 1992
2. Limber & Goodrich, *JOTA* 1993

Remarks

1. Valid for **any continuous probability distribution**
2. **Vector observations** and **any number of sensors**

Assumption:
Finite 1st and 2nd moments
(req. for strong duality)

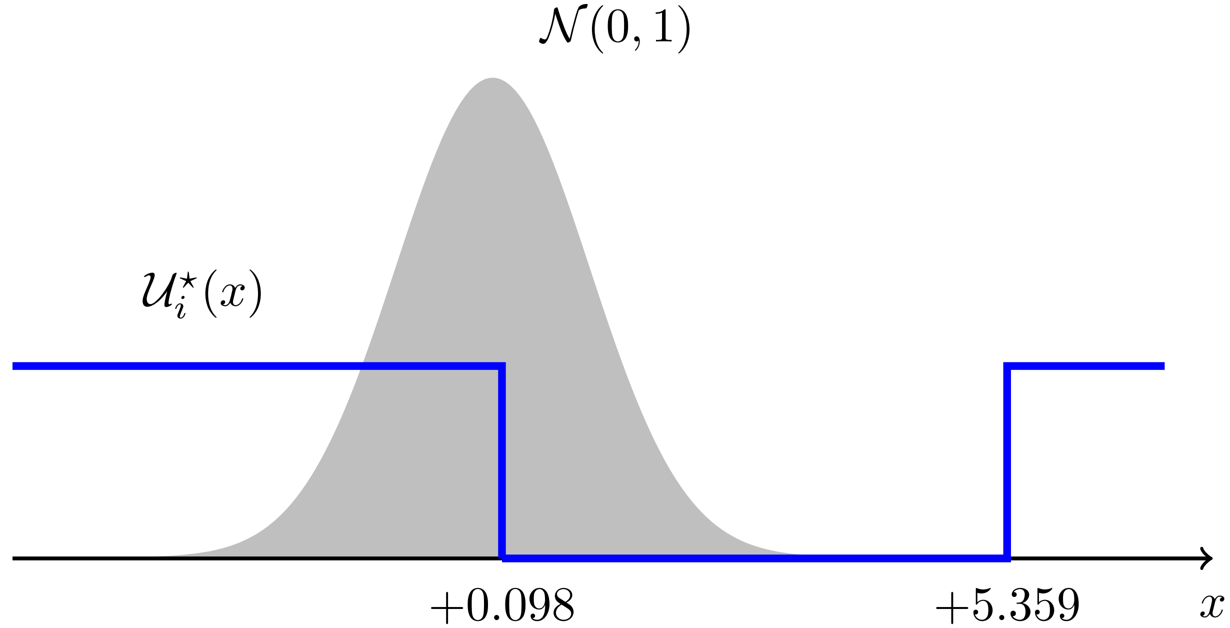
Additional assumption:
The fusion center can decode the indices of all sensors involved in a collision



Vasconcelos & Martins, "Optimal estimation over the collision channel," *IEEE TAC* 2017

Person-by-person optimal threshold policies

$$X_1, X_2 \sim \mathcal{N}(0,1)$$



i.i.d. observations, symmetric pdf

asymmetric thresholds

$$\mathcal{J}(\mathcal{U}_1^\star, \mathcal{U}_2^\star) = 0.54$$

Gain of 46% over
open loop scheduling policies

Vasconcelos & Martins, "Optimal thresholds for remote estimation over the collision channel," *IEEE CDC* 2015

Lipsa & Martins, "Remote state estimation with communication costs for first-order LTI systems". *IEEE TAC* 2011

Drawback

Computing team-optimal thresholds is still a very difficult problem!

We know how to compute person-by-person optimal policies efficiently[1]

Can we provide an optimality guarantee?

1. Vasconcelos & Martins, "Optimal estimation over the collision channel," *IEEE TAC* 2017

Drawback

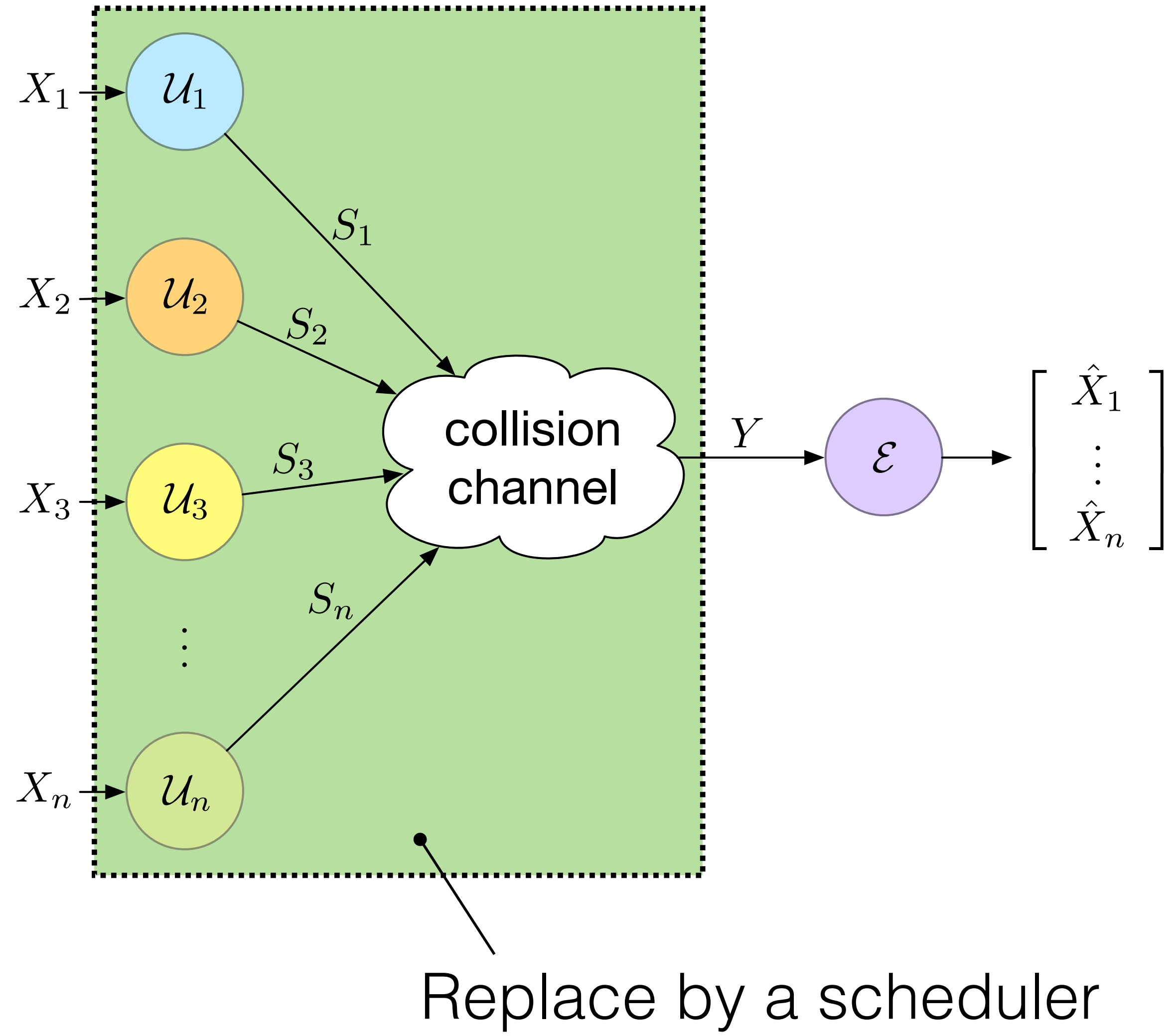
Computing team-optimal thresholds is still a very difficult problem!

We know how to compute person-by-person optimal policies efficiently[1]

Can we find a nontrivial lower bound?

Vasconcelos & Martins, "Optimal estimation over the collision channel," *IEEE TAC* 2017

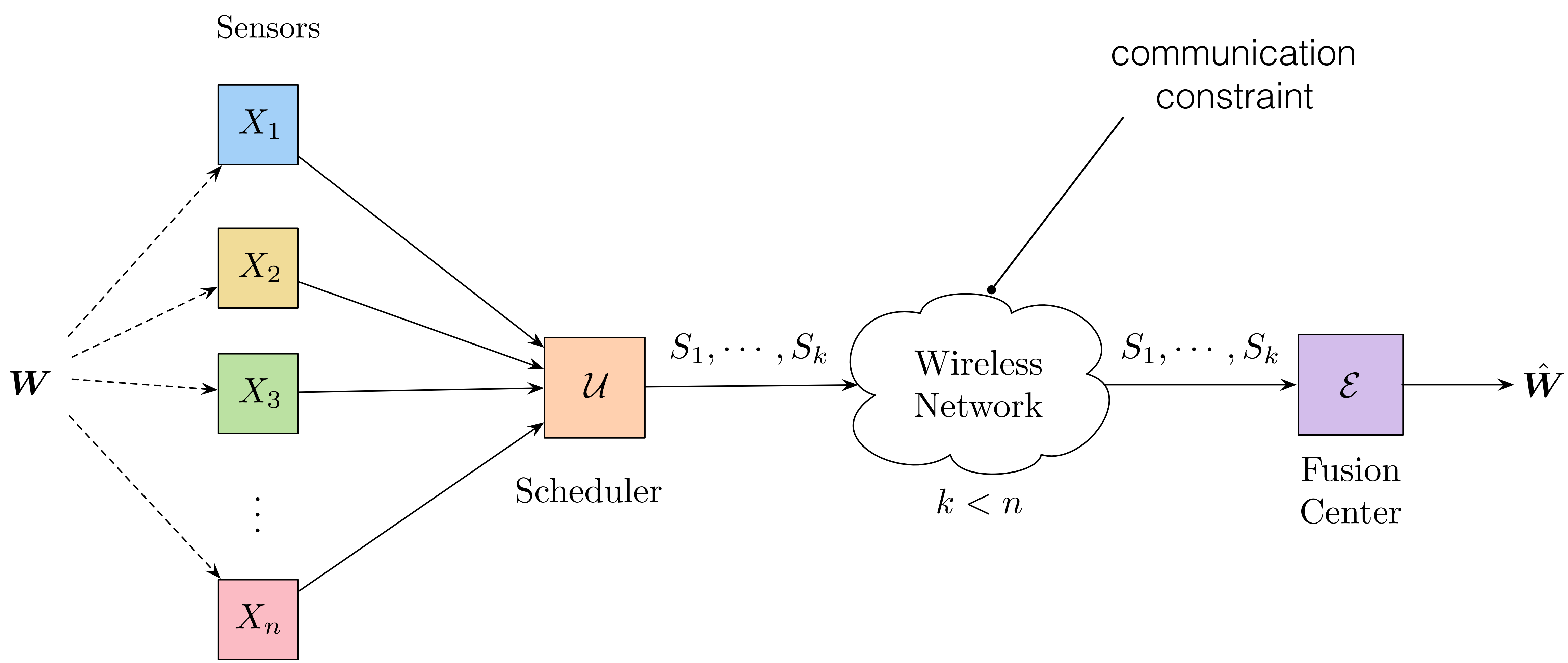
"Centralized" lower bound



The optimal performance of this system is a lower bound to the decentralized problem

Observation-driven sensor scheduling

Basic framework

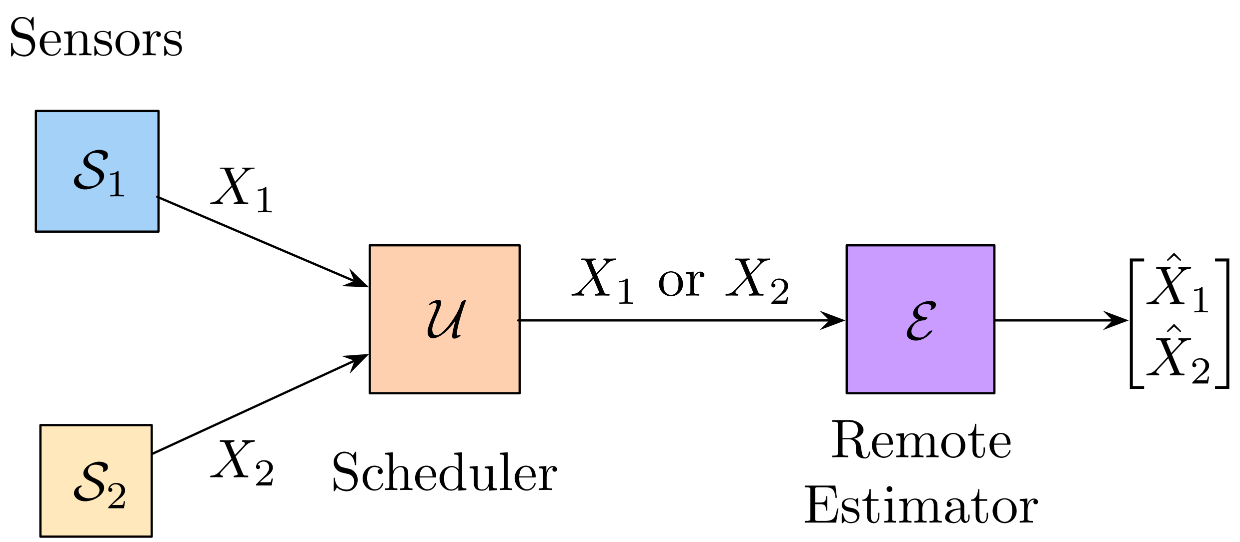


Sensor scheduling problem

Choose k out of n sensors such that
the expected distortion between $\boldsymbol{W}$ and $\hat{\boldsymbol{W}}$ is minimized

1. Athans - *Automatica* 1972
2. Joshi & Boyd - *IEEE TSP* 2009
3. Mo, Ambrosino & Sinopoli - *Automatica* 2011
4. Moon & Basar - *IEEE TSP* 2017

Simplest case: two sensors



Observations

$$X_i \sim \mathcal{N}(0, \sigma_i^2)$$

Decision variable

$$U \in \{1, 2\}$$

Transmit

$$S = (1, X_1)$$

Transmit

$$S = (2, X_2)$$

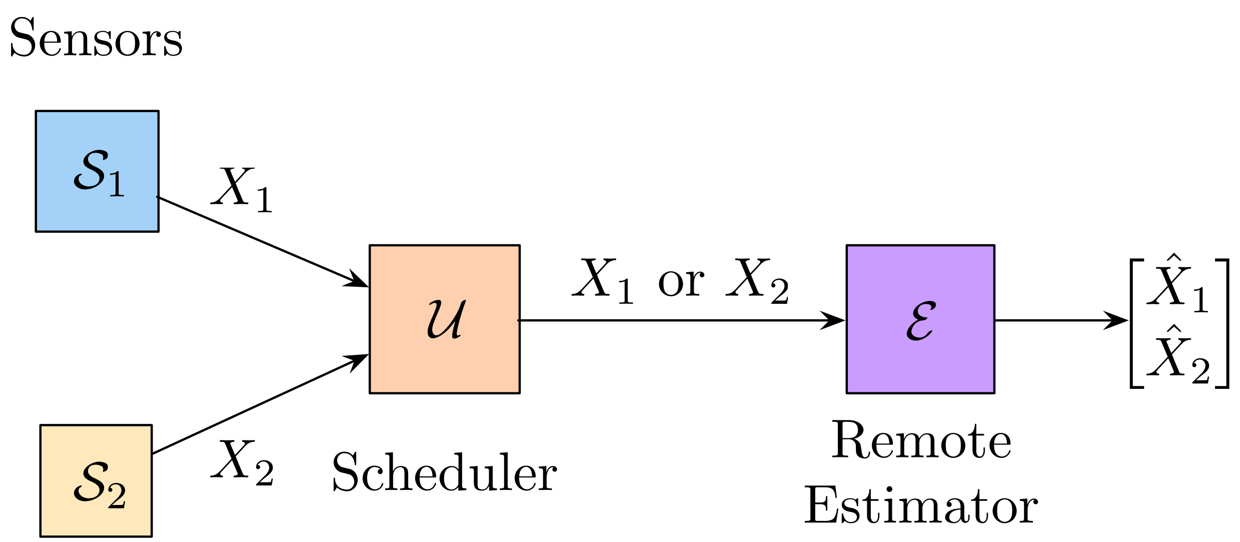
Scheduling policy

$$U = \mathcal{U}(X_1, X_2)$$

Estimation policy

$$\begin{bmatrix}\hat{X}_1\\ \hat{X}_2\end{bmatrix} = \mathcal{E}(Y)$$

Simplest case: two sensors



Observations

$$X_i \sim \mathcal{N}(0, \sigma_i^2)$$

Decision variable

$$U \in \{1, 2\}$$

Scheduling policy

$$U = \mathcal{U}(X_1, X_2)$$

Estimation policy

$$\begin{bmatrix}\hat{X}_1\\ \hat{X}_2\end{bmatrix} = \mathcal{E}(Y)$$

Problem 3

$$\min_{(\mathcal{U},\mathcal{E})\in\mathbb{U}\times\mathbb{E}} \quad \mathcal{J}(\mathcal{U},\mathcal{E}) = \mathbf{E}\left[(X_1 - \hat{X}_1)^2 + (X_2 - \hat{X}_2)^2\right]$$

Notions of optimality

Team-optimality

$$\mathcal{J}(\mathcal{U}^\star, \mathcal{E}^\star) \leq \mathcal{J}(\mathcal{U}, \mathcal{E}), \quad (\mathcal{U}, \mathcal{E}) \in \mathbb{U} \times \mathbb{E}$$

$$\Longrightarrow$$

$$\not\Longleftarrow$$

Person-by-person optimality

$$\mathcal{J}(\mathcal{U}^\star, \mathcal{E}^\star) \leq \mathcal{J}(\mathcal{U}, \mathcal{E}^\star), \quad \mathcal{U} \in \mathbb{U}$$
$$\mathcal{J}(\mathcal{U}^\star, \mathcal{E}^\star) \leq \mathcal{J}(\mathcal{U}^\star, \mathcal{E}), \quad \mathcal{E} \in \mathbb{E}$$

Unfortunately, finding team-optimal optimal solutions is **very difficult**

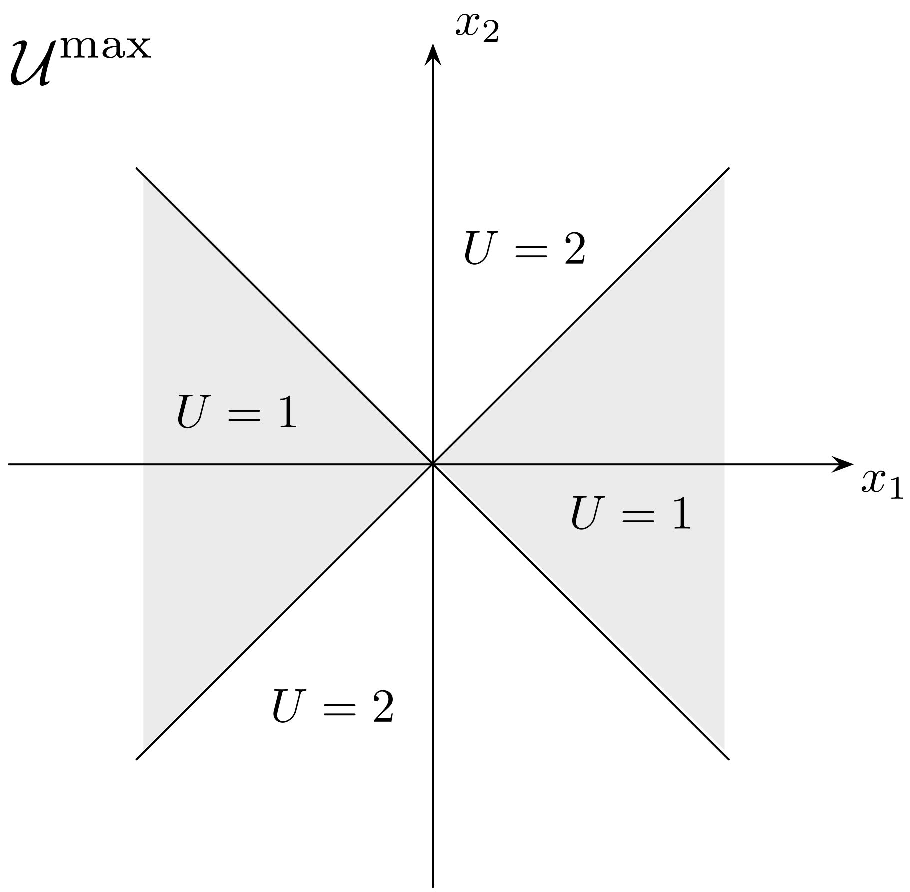
Finding person-by-person optimal solutions is **often much easier***

***depending on the probabilistic model of the source**

Max-scheduling

Max-scheduling policy

$$\mathcal{U}^{\max}(x_1, x_2) = \begin{cases} 1 & \text{if } |x_1| \geq |x_2| \\ 2 & \text{otherwise} \end{cases}$$



Mean-estimation policy

$$\mathcal{E}^{\text{mean}}(1, x_1) = \begin{bmatrix} x_1 \\ 0 \end{bmatrix}$$

$$\mathcal{E}^{\text{mean}}(2, x_2) = \begin{bmatrix} 0 \\ x_2 \end{bmatrix}$$

Independent sources

Theorem 2

$$X_1 \perp\!\!\!\perp X_2 \quad \Longrightarrow \quad (\mathcal{U}^{\max}, \mathcal{E}^{\text{mean}}) \text{ is person-by-person optimal}$$

Open-loop scheduling: let the sensor with the **largest variance** transmit

Observation-driven scheduling[1,2]**:** let the sensor with the **"largest measurement"** transmit

1. Vasconcelos & Mitra - "Observation-driven sensor scheduling" *IEEE ICC* 2017
2. Vasconcelos & Mitra - "Observation-driven scheduling of Gaussian sources" To be submitted to *IEEE TCNS* 2018

Sketch of proof

The MMSE estimator for a given scheduling policy is

$$\mathcal{E}^{\star}_{\mathcal{U}}(1, x_1) = \begin{bmatrix} x_1 \\ \mathbf{E}\big[X_2 \mid U = 1, X_1 = x_1\big] \end{bmatrix}$$

$$\mathcal{E}^{\star}_{\mathcal{U}}(2, x_2) = \begin{bmatrix} \mathbf{E}\big[X_1 \mid U = 2, X_2 = x_2\big] \\ x_2 \end{bmatrix}$$

Suppose that $\mathcal{U} = \mathcal{U}^{\max}$ **then**

$$\mathbf{E}\big[X_2 \mid U = 1, X_1 = x_1\big] = \frac{\int_{\mathbb{R}} x_2 \mathbf{1}(|x_1| \geq |x_2|) f_{X_2|X_1=x_1}(x_2) dx_2}{\int_{\mathbb{R}} \mathbf{1}(|x_1| \geq |x_2|) f_{X_2|X_1=x_1}(x_2) dx_2}$$

Sketch of proof

The MMSE estimator for a given scheduling policy is

$$\mathcal{E}^{\star}_{\mathcal{U}}(1, x_1) = \begin{bmatrix} x_1 \\ \mathbf{E}\big[X_2 \mid U = 1, X_1 = x_1\big] \end{bmatrix}$$

$$\mathcal{E}^{\star}_{\mathcal{U}}(2, x_2) = \begin{bmatrix} \mathbf{E}\big[X_1 \mid U = 2, X_2 = x_2\big] \\ x_2 \end{bmatrix}$$

Suppose that $\mathcal{U} = \mathcal{U}^{\max}$ **then**

Symmetric around zero

$$\mathbf{E}\big[X_2 \mid U = 1, X_1 = x_1\big] = \frac{\int_{-|x_1|}^{|x_1|} x_2 f_{X_2}(x_2) dx_2}{\int_{-|x_1|}^{|x_1|} f_{X_2}(x_2) dx_2} = 0$$

Sketch of proof

Fix an estimation policy of the form:

$$\mathcal{E}(1, x_1) = \begin{bmatrix} x_1 \\ \eta_2(x_1) \end{bmatrix} \qquad \mathcal{E}(2, x_2) = \begin{bmatrix} \eta_1(x_2) \\ x_2 \end{bmatrix}$$

The cost becomes

$$\begin{aligned} \mathcal{J}(\mathcal{U}, \mathcal{E}) = &\int_{\mathbb{R}^2} \big(x_2 - \eta_2(x_1)\big)^2 \mathbf{1}\big(\mathcal{U}(x_1, x_2) = 1\big) f(x_1, x_2) dx_1 dx_2 \\ &+ \int_{\mathbb{R}^2} \big(x_1 - \eta_1(x_2)\big)^2 \mathbf{1}\big(\mathcal{U}(x_1, x_2) = 2\big) f(x_1, x_2) dx_1 dx_2 \end{aligned}$$

$$\mathcal{U}^\star_{\mathcal{E}}(x_1, x_2) = 1 \iff \big(x_1 - \eta_1(x_2)\big)^2 \geq \big(x_2 - \eta_2(x_1)\big)^2$$

Generalized Nearest Neighbor Condition

Sketch of proof

$$\mathcal{U}^\star_{\mathcal{E}}(x_1, x_2) = 1 \iff \big(x_1 - \eta_1(x_2)\big)^2 \geq \big(x_2 - \eta_2(x_1)\big)^2$$

Suppose that $\eta_1(x_2) = \eta_2(x_1) \equiv 0$

then $\mathcal{U}^\star_{\mathcal{E}^{\mathrm{mean}}}(x_1, x_2) = 1 \iff \big(x_1 - 0\big)^2 \geq \big(x_2 - 0\big)^2$

$$|x_1| \geq |x_2|$$

$\blacksquare$

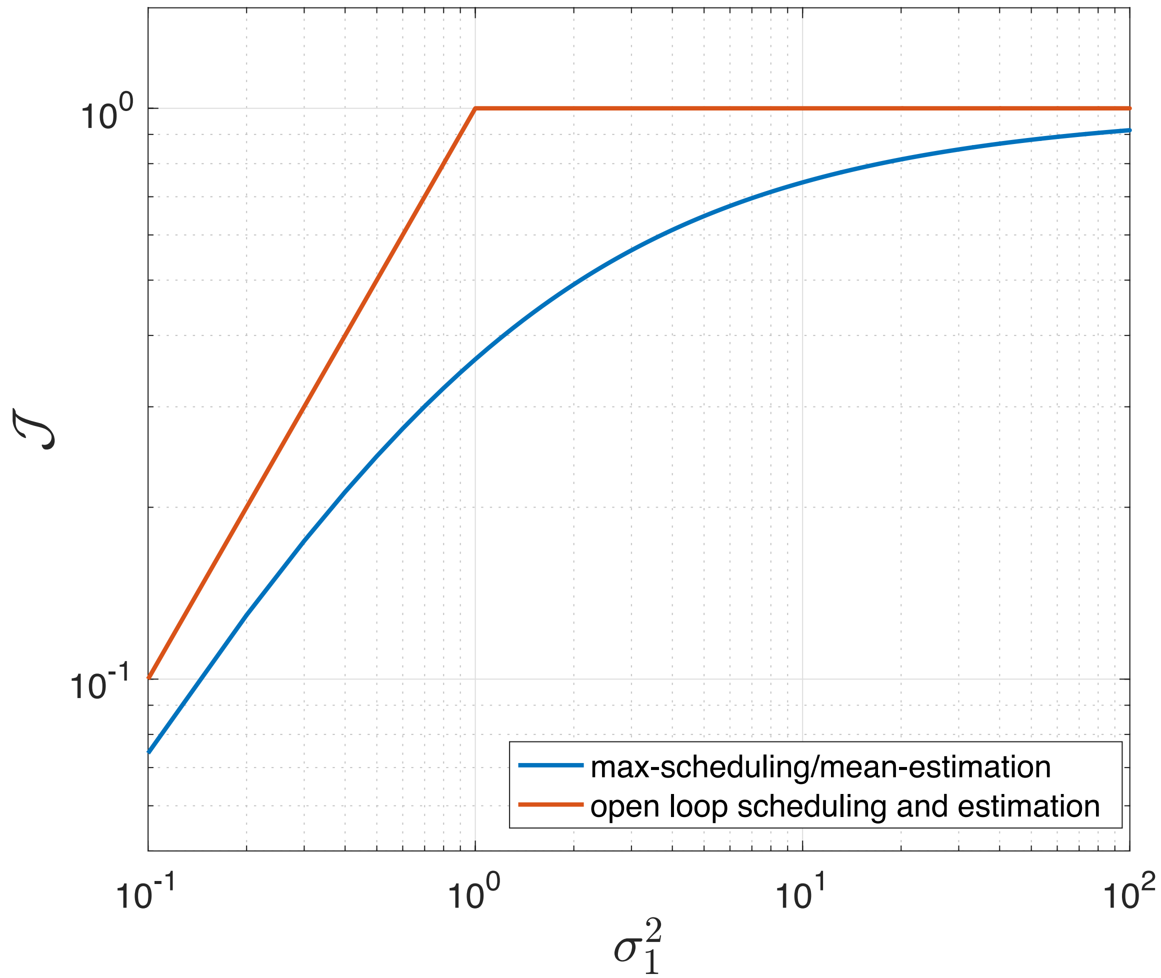
Value of information

$$\mathcal{J}(\mathcal{U}^{\max}, \mathcal{E}^{\text{mean}}) = \mathbf{E}\Big[\min\{X_1^2, X_2^2\}\Big]$$

Observation-driven sensor scheduling

$$\mathcal{J}(\mathcal{U}^{\text{open}}, \mathcal{E}^{\text{mean}}) = \min\{\sigma_1^2, \sigma_2^2\}$$

“Open-loop” sensor scheduling



Remarks

1. Result only depends on the even symmetry of the density
2. Can be extended to **any number of sensors** making **vector observations**[1]

1. Vasconcelos, Nayyar & Mitra - “Optimal scheduling strategies in networked estimation” *IEEE CDC* 2017

Symmetric sources

Theorem 3

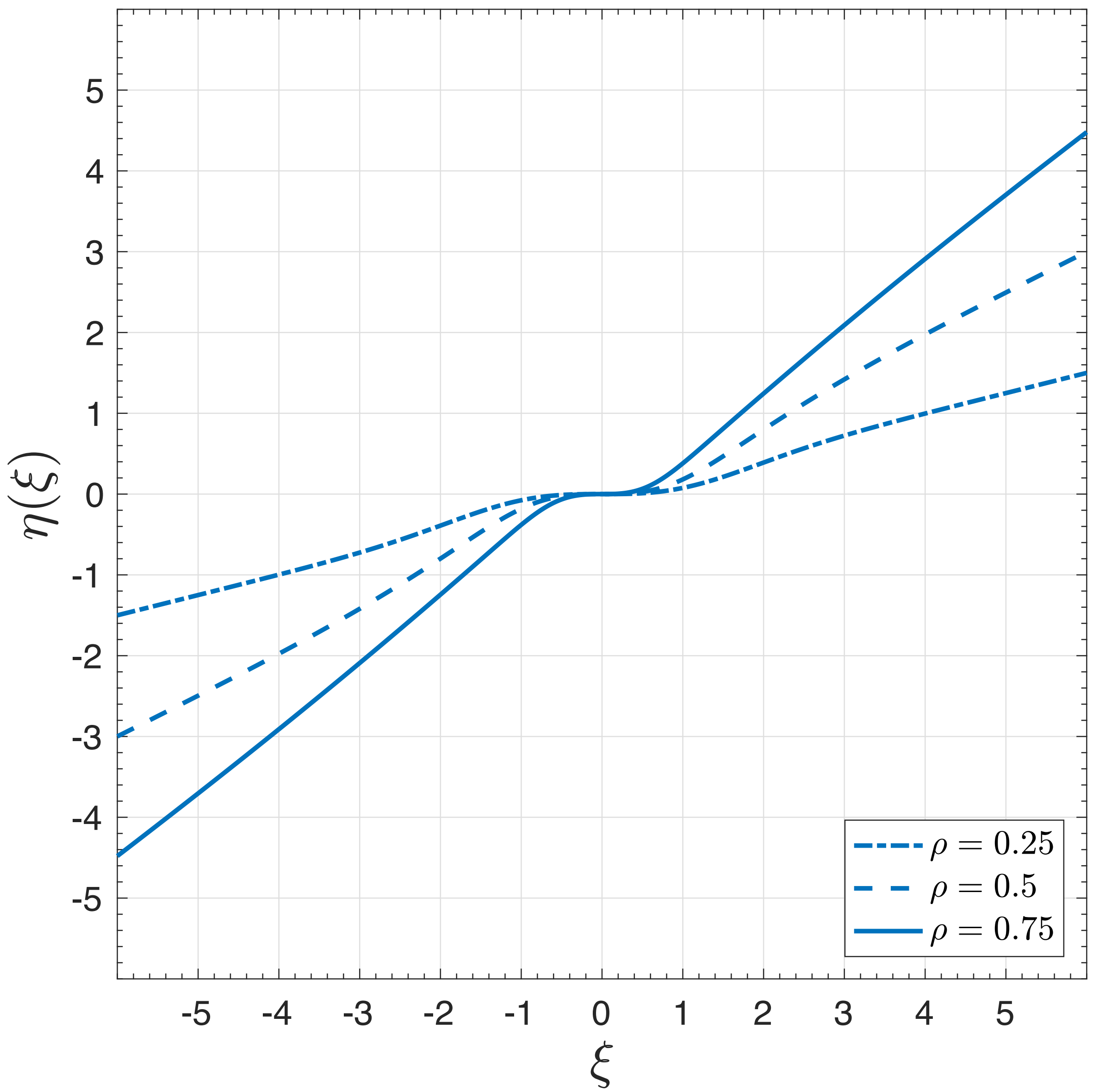
$$\sigma_1^2 = \sigma_2^2 \quad \Longrightarrow \quad (\mathcal{U}^{\max}, \mathcal{E}^{\text{soft}}) \text{ is person-by-person optimal}$$

Soft-threshold estimation policy

$$\mathcal{E}^{\text{soft}}(1, x_1) = \begin{bmatrix} x_1 \\ \eta(x_1) \end{bmatrix}$$

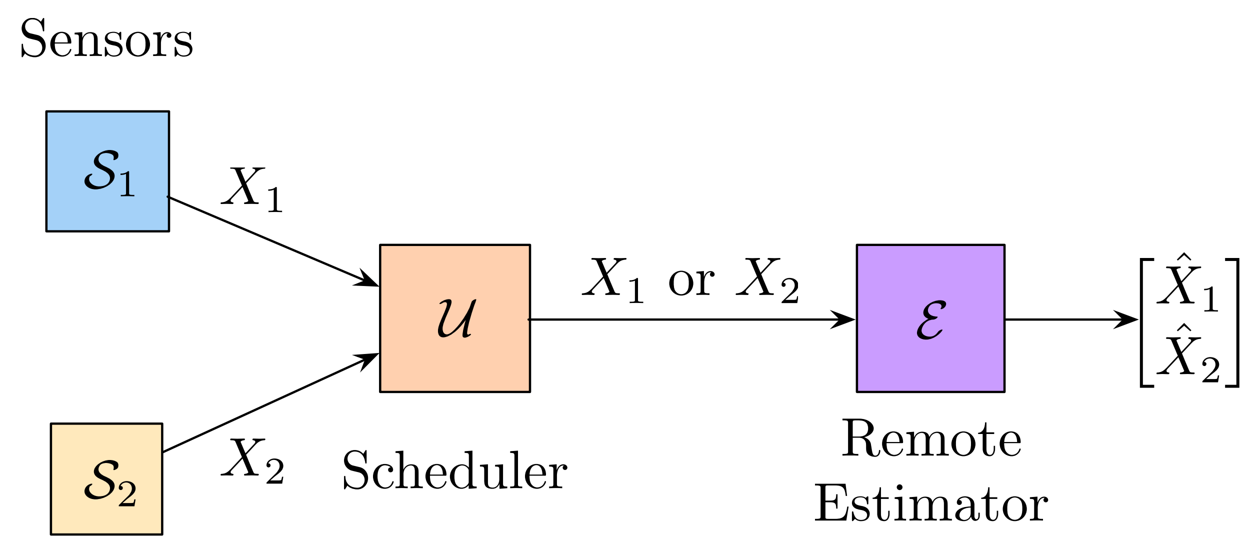
$$\mathcal{E}^{\text{soft}}(2, x_2) = \begin{bmatrix} \eta(x_2) \\ x_2 \end{bmatrix}$$

$$\eta(\xi) = \frac{\int_{-|\xi|}^{|\xi|} \tau \exp\left(-\frac{(\tau-\rho\xi)^2}{2\sigma^2(1-\rho^2)}\right) d\tau}{\int_{-|\xi|}^{|\xi|} \exp\left(-\frac{(\tau-\rho\xi)^2}{2\sigma^2(1-\rho^2)}\right) d\tau}$$



Proof is much more involved...

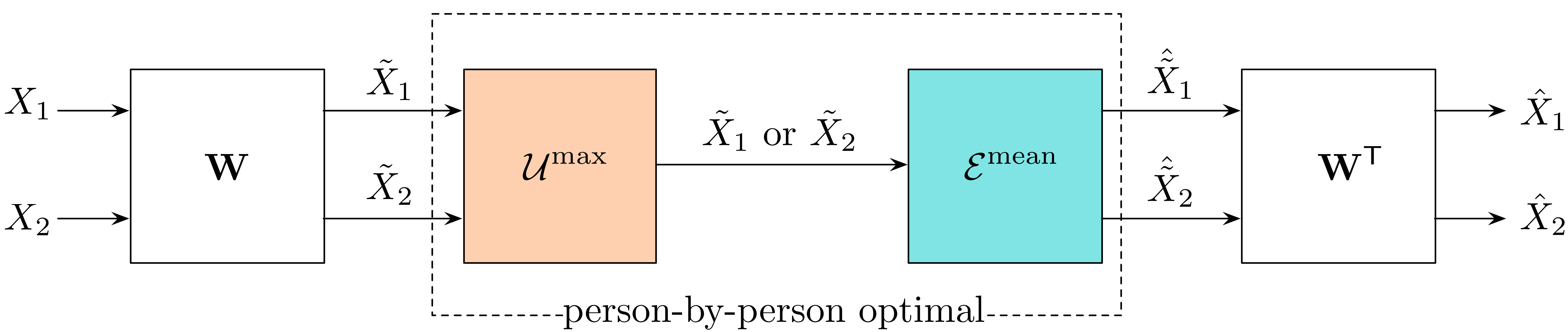
General Gaussian sources



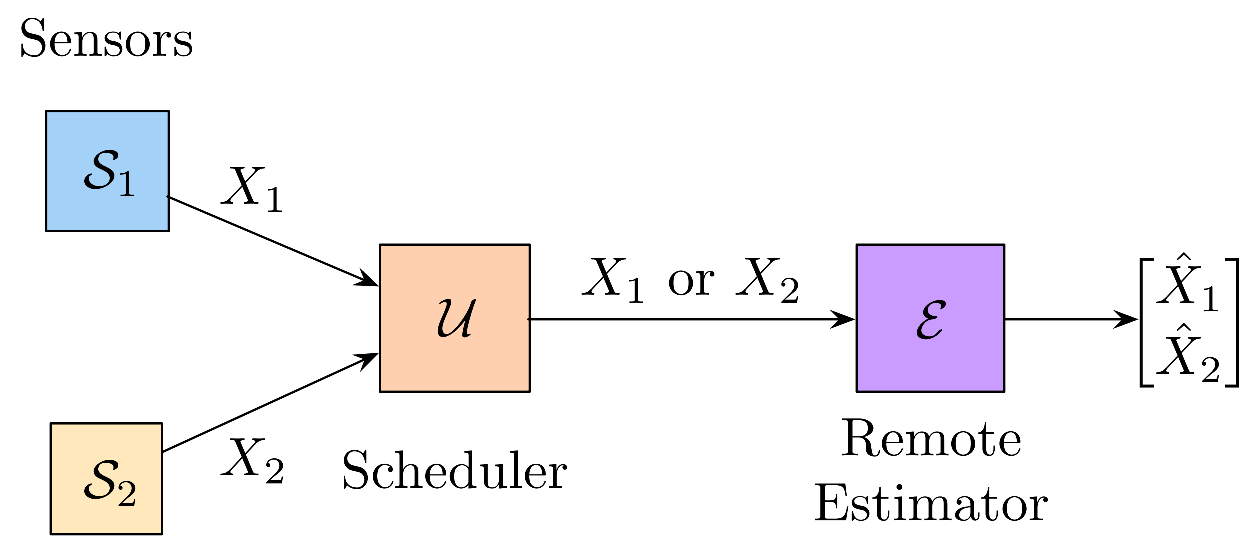
Observations

$$\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma})$$

$$\mathbf{\Sigma} = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix} = \mathbf{W}\mathbf{\Lambda}\mathbf{W}^\mathsf{T}$$



General Gaussian sources

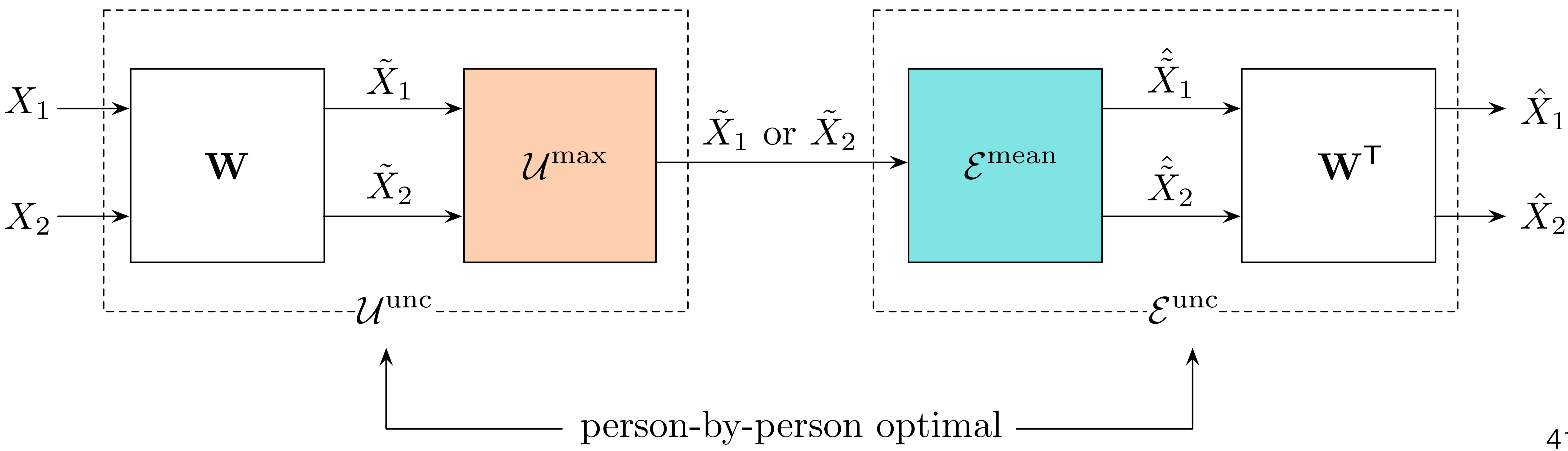


Observations

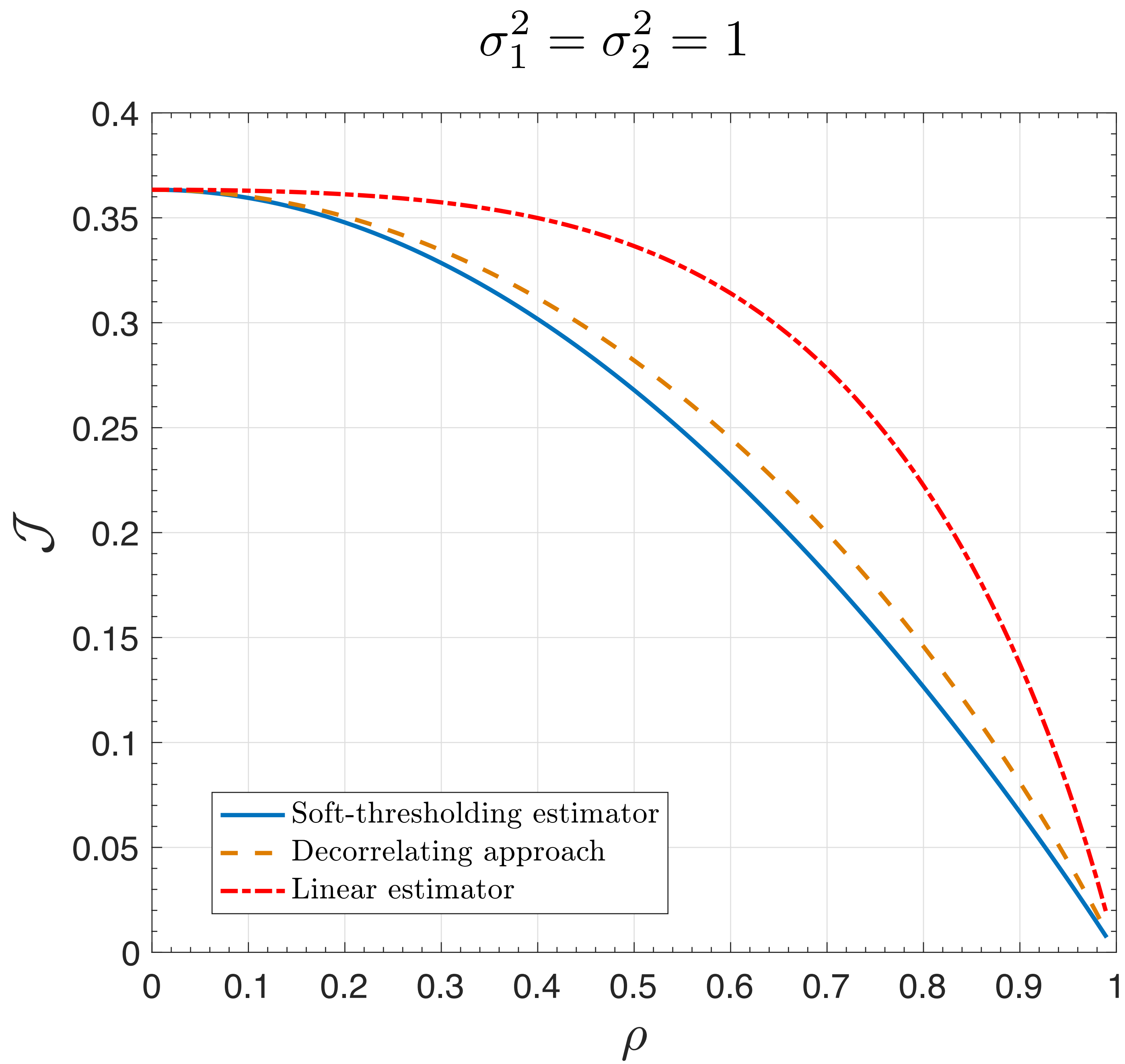
$$\begin{bmatrix}X_1\\ X_2\end{bmatrix} \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma})$$

Theorem 4

$$\mathbf{X} \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}) \implies (\mathcal{U}^{\mathrm{unc}}, \mathcal{E}^{\mathrm{unc}}) \text{ is person-by-person optimal}$$



Performance



$$\eta(\xi) = \frac{\int_{-|\xi|}^{|\xi|} \tau \exp\left(-\frac{(\tau-\rho\xi)^2}{2\sigma^2(1-\rho^2)}\right) d\tau}{\int_{-|\xi|}^{|\xi|} \exp\left(-\frac{(\tau-\rho\xi)^2}{2\sigma^2(1-\rho^2)}\right) d\tau}$$

$$\eta(\xi) = \rho \cdot \xi$$

Scheduling sensors with unknown joint density

Arbitrary joint density

$$(X_1, X_2) \sim f(x_1, x_2)$$

Generalized nearest neighbor condition

$$\mathcal{U}^\star_{\mathcal{E}}(x_1, x_2) = 1 \Longleftrightarrow \big(x_1 - \eta_1(x_2)\big)^2 \geq \big(x_2 - \eta_2(x_1)\big)^2$$

↓

Infinite dimensional optimization

$$\mathcal{J}(\eta_1, \eta_2) = \mathbf{E}\Big[\min\big\{\big(X_1 - \eta_1(X_2)\big)^2, \big(X_1 - \eta_2(X_1)\big)^2\big\}\Big]$$

Arbitrary joint density

$$(X_1, X_2) \sim f(x_1, x_2)$$

Generalized nearest neighbor condition

$$\mathcal{U}^{\star}_{\mathcal{E}}(x_1, x_2) = 1 \iff \big(x_1 - \eta_1(x_2)\big)^2 \geq \big(x_2 - \eta_2(x_1)\big)^2$$

$$\big\downarrow \; \eta_i(x) = a_i x$$

Linear estimators

Finite dimensional optimization

$$\mathcal{J}(\mathbf{a}) = \mathbf{E}\Big[\min\big\{\big(X_1 - a_1 X_2\big)^2, \big(X_1 - a_2 X_1\big)^2\big\}\Big]$$

Nonconvex

Arbitrary joint density

$$(X_1, X_2) \sim f(x_1, x_2)$$

Generalized nearest neighbor condition

$$\mathcal{U}^\star_{\mathcal{E}}(x_1, x_2) = 1 \iff \big(x_1 - \eta_1(x_2)\big)^2 \geq \big(x_2 - \eta_2(x_1)\big)^2$$

$$\Big\downarrow \; \eta_i(x) = a_i x$$

Linear estimators

Finite dimensional optimization

$$\mathcal{J}(\mathbf{a}) = \mathbf{E}\Big[\big(X_1 - a_1 X_2\big)^2 + \big(X_1 - a_2 X_1\big)^2\Big] - \mathbf{E}\Big[\max\big\{\big(X_1 - a_1 X_2\big)^2, \big(X_1 - a_2 X_1\big)^2\big\}\Big]$$

Difference of Convex

Difference of Convex decomposition

$$\mathcal{J}(\mathbf{a}) = \mathcal{F}(\mathbf{a}) - \mathcal{G}(\mathbf{a})$$

$$\mathcal{F}(\mathbf{a}) = \mathbf{E}\Big[\big(X_1 - a_1 X_2\big)^2 + \big(X_1 - a_2 X_1\big)^2\Big]$$

$$\mathcal{G}(\mathbf{a}) = \mathbf{E}\Big[\max\big\{\big(X_1 - a_1 X_2\big)^2, \big(X_1 - a_2 X_1\big)^2\big\}\Big]$$

Convex-concave procedure

Heuristics to find local minimizers[1,2]

$$\mathcal{J}(\mathbf{a}) = \mathcal{F}(\mathbf{a}) - \mathcal{G}(\mathbf{a})$$

$a^{(k)}$

$$\mathcal{G}_{\text{affine}}(\mathbf{a};\mathbf{a}^{(k)}) = \mathcal{G}(\mathbf{a}^{(k)}) + g(\mathbf{a}^{(k)})^{\mathsf{T}}(\mathbf{a} - \mathbf{a}^{(k)})$$

Subgradient of $\mathcal{G}(\mathbf{a})$

$$\mathbf{a}^{(k+1)} = \arg\min_{\mathbf{a}} \left\{ \mathcal{F}(\mathbf{a}) - \mathcal{G}_{\text{affine}}(\mathbf{a};\mathbf{a}^{(k)}) \right\}$$

Converges to a critical point

Convex

[1] Lipp and Boyd - Optim Eng (2016)
[2] Yuille and Rangarajan - Neural Comp (2003)

Unknown density

$$(X_1, X_2) \sim ?$$

Cannot compute expectations

Replace expectations by the **empirical mean**

Data: $\{x_1(k), x_2(k)\}_{k=1}^{K}$

$$\tilde{\mathcal{F}}(\mathbf{a}) = \frac{1}{K} \sum_{k=1}^{K} \left[\big(x_1(k) - a_1 x_2(k)\big)^2 + \big(x_2(k) - a_2 x_1(k)\big)^2 \right]$$

$$\tilde{\mathcal{G}}(\mathbf{a}) = \frac{1}{K} \sum_{k=1}^{K} \left[\max \left\{ \big(x_1(k) - a_1 x_2(k)\big)^2, \big(x_2(k) - a_2 x_1(k)\big)^2 \right\} \right]$$

Approximate convex-concave procedure

$a^{(k)}$

$$\tilde{\mathcal{G}}_{\text{affine}}(\mathbf{a};\mathbf{a}^{(k)}) = \tilde{\mathcal{G}}(\mathbf{a}^{(k)}) + \tilde{g}(\mathbf{a}^{(k)})^{\mathsf{T}}(\mathbf{a}-\mathbf{a}^{(k)})$$

Subgradient of $\tilde{\mathcal{G}}(\mathbf{a})$

$$\mathbf{a}^{(k+1)} = \arg\min_{\mathbf{a}} \left\{ \tilde{\mathcal{F}}(\mathbf{a}) - \tilde{\mathcal{G}}_{\text{affine}}(\mathbf{a};\mathbf{a}^{(k)}) \right\}$$

Convex

Cannot claim convergence

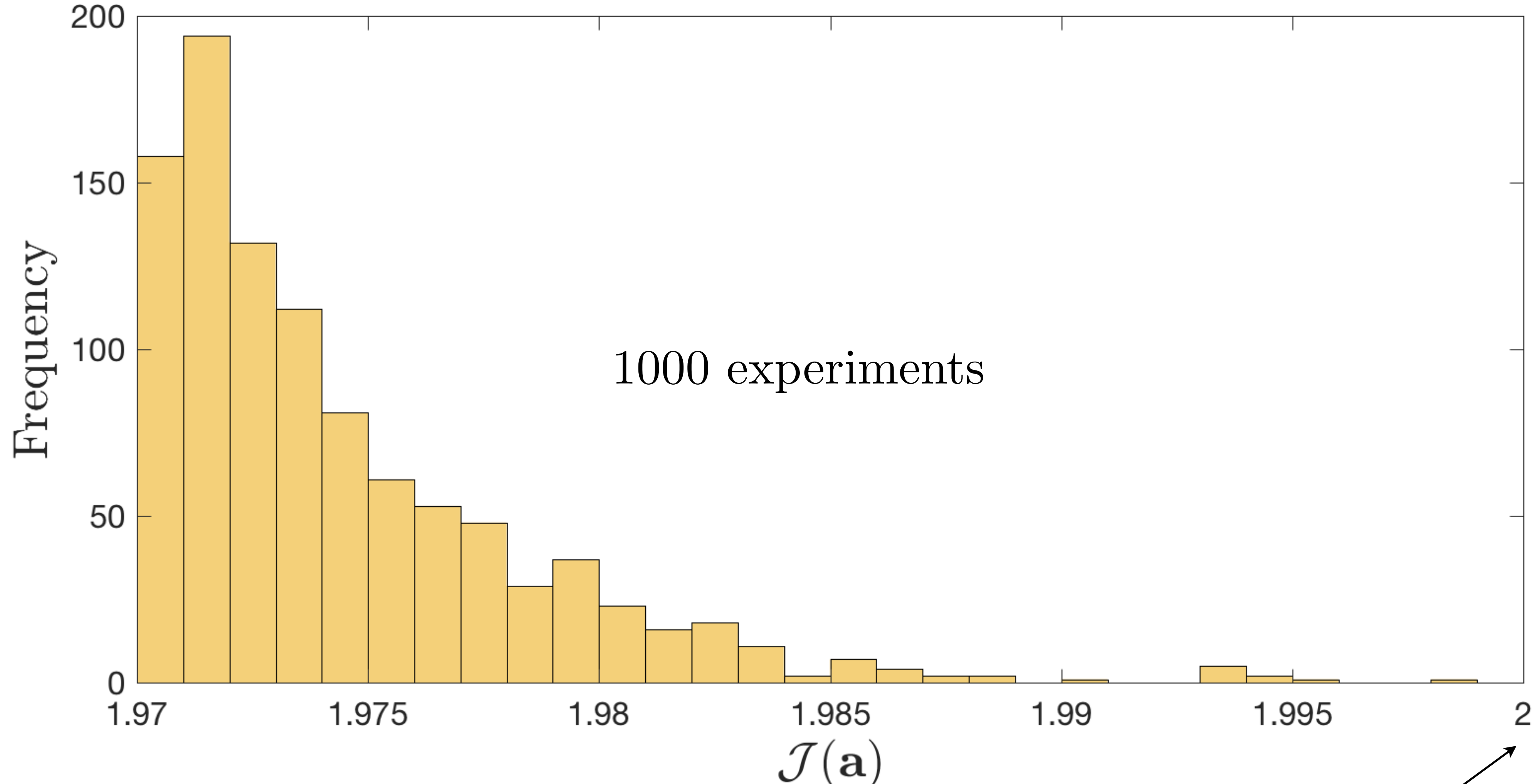
$$\tilde{g}(\mathbf{a}) = \frac{1}{K}\sum_{k=1}^{K}\begin{bmatrix} -2(x_1(k)-a_1x_2(k))x_2(k)\cdot\mathbf{1}\big((x_1(k)-a_1x_2(k))^2 \geq (x_2(k)-a_2x_1(k))^2\big) \\ -2(x_2(k)-a_2x_1(k))x_1(k)\cdot\mathbf{1}\big((x_1(k)-a_1x_2(k))^2 < (x_2(k)-a_2x_1(k))^2\big) \end{bmatrix}$$

Empirical results

$$(X_1, X_2) \sim \mathcal{N}\left(\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}5 & 1.7748\\1.7748 & 7\end{bmatrix}\right)$$

$$K = 1000$$

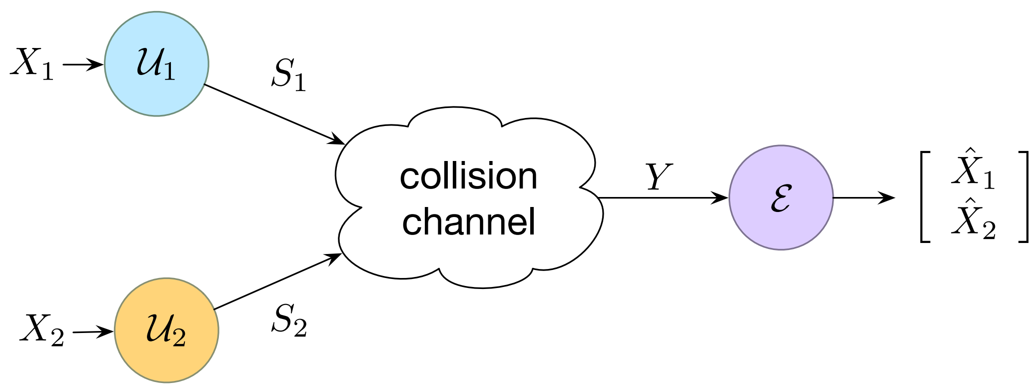
$$L = 100$$

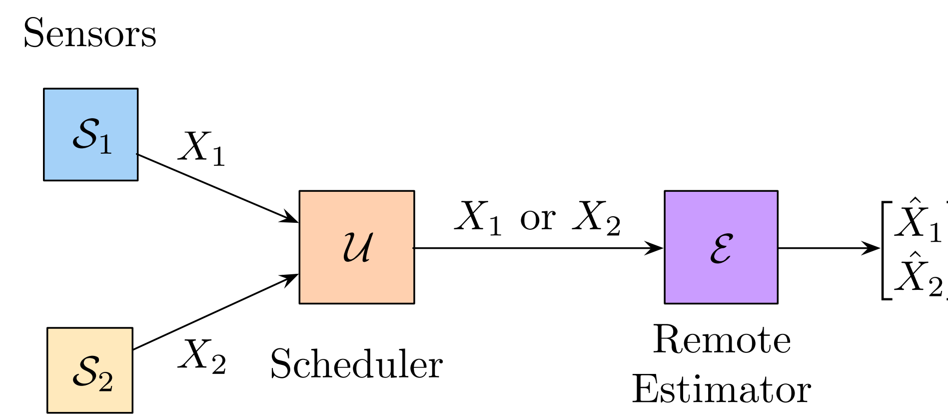


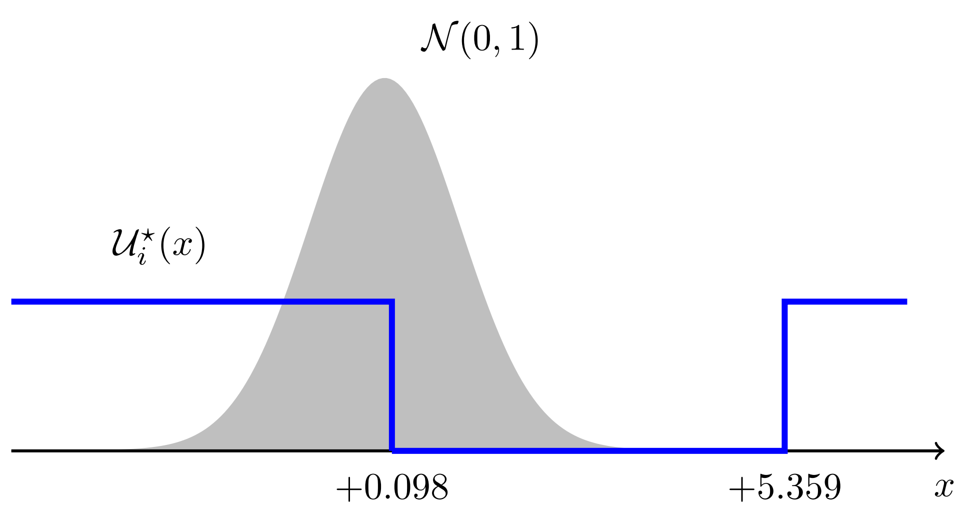
$\mathcal{J}^\star = 1.9704$

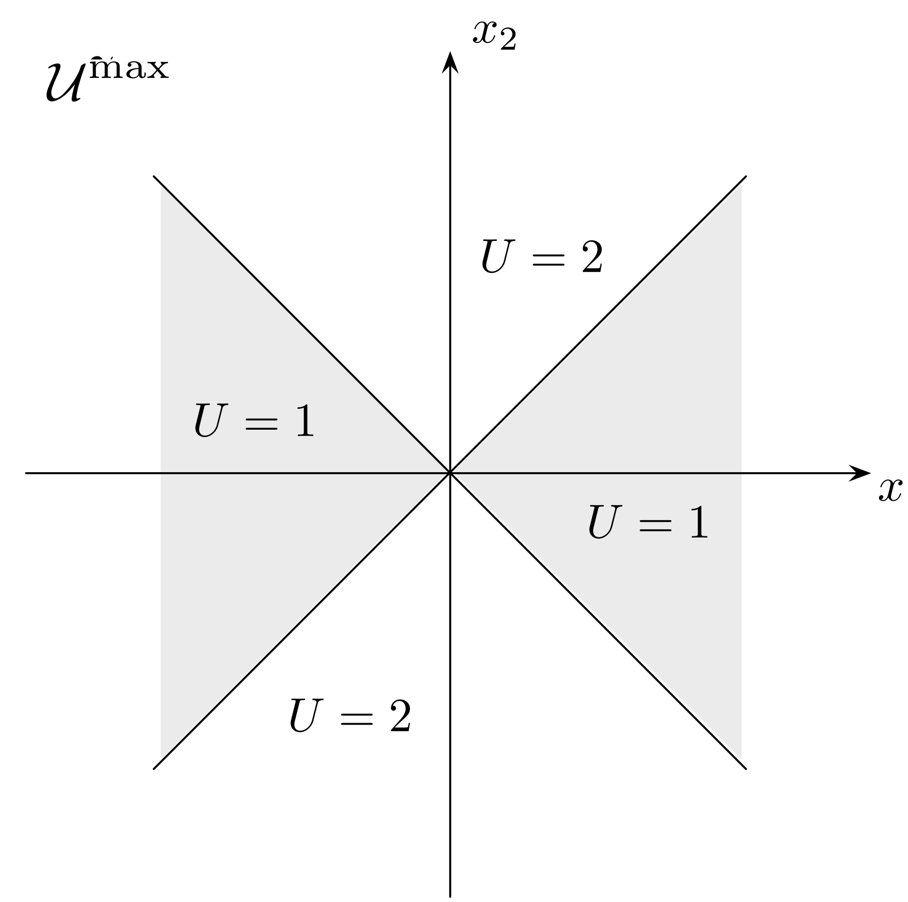
Within 1.5% of the optimal solution

Collision vs. Scheduling









Threshold policies + collision channel = “decentralized max function”

Summary & future work

1. Estimation over the collision channel:

Optimality of threshold policies

Designing globally optimal thresholds is NP-hard

2. Observation-driven scheduling:

Person-by-person optimality results (max-scheduling)

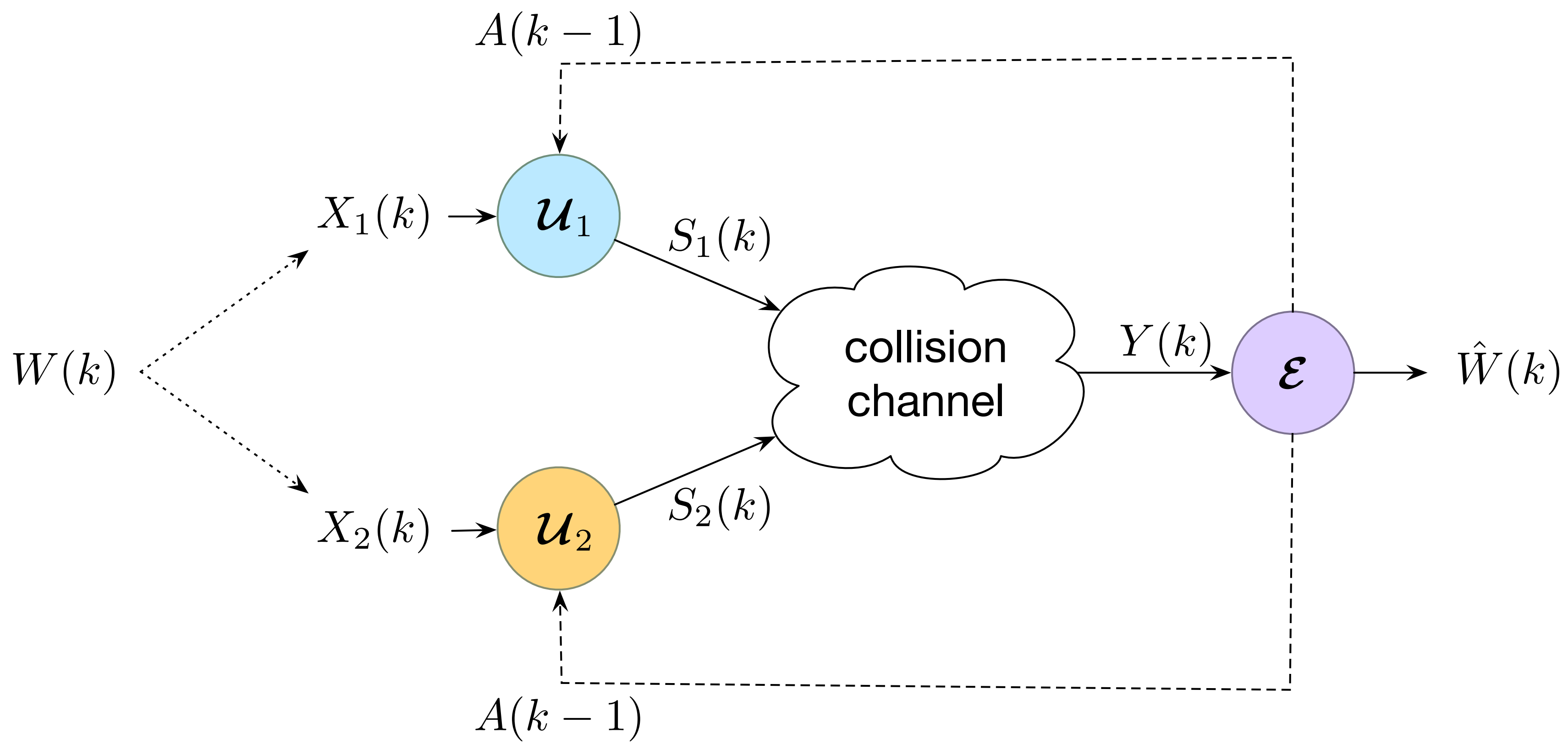
Global optimality results are elusive

Proof of global optimality may come from Information Theory

3. Fundamentals of distributed estimation/scheduling with sensors of unknown (or imprecise) probabilistic models

Future work

The sequential case



$$\mathcal{J}(\boldsymbol{\mathcal{U}}_1, \boldsymbol{\mathcal{U}}_2, \boldsymbol{\mathcal{E}}) = \sum_{k=0}^{T} \mathbf{E}\left[d\big(W(k), \hat{W}(k)\big)\right]$$