Optimal remote estimation over the collision channel

Marcos M. Vasconcelos

`marcos@umd.edu`

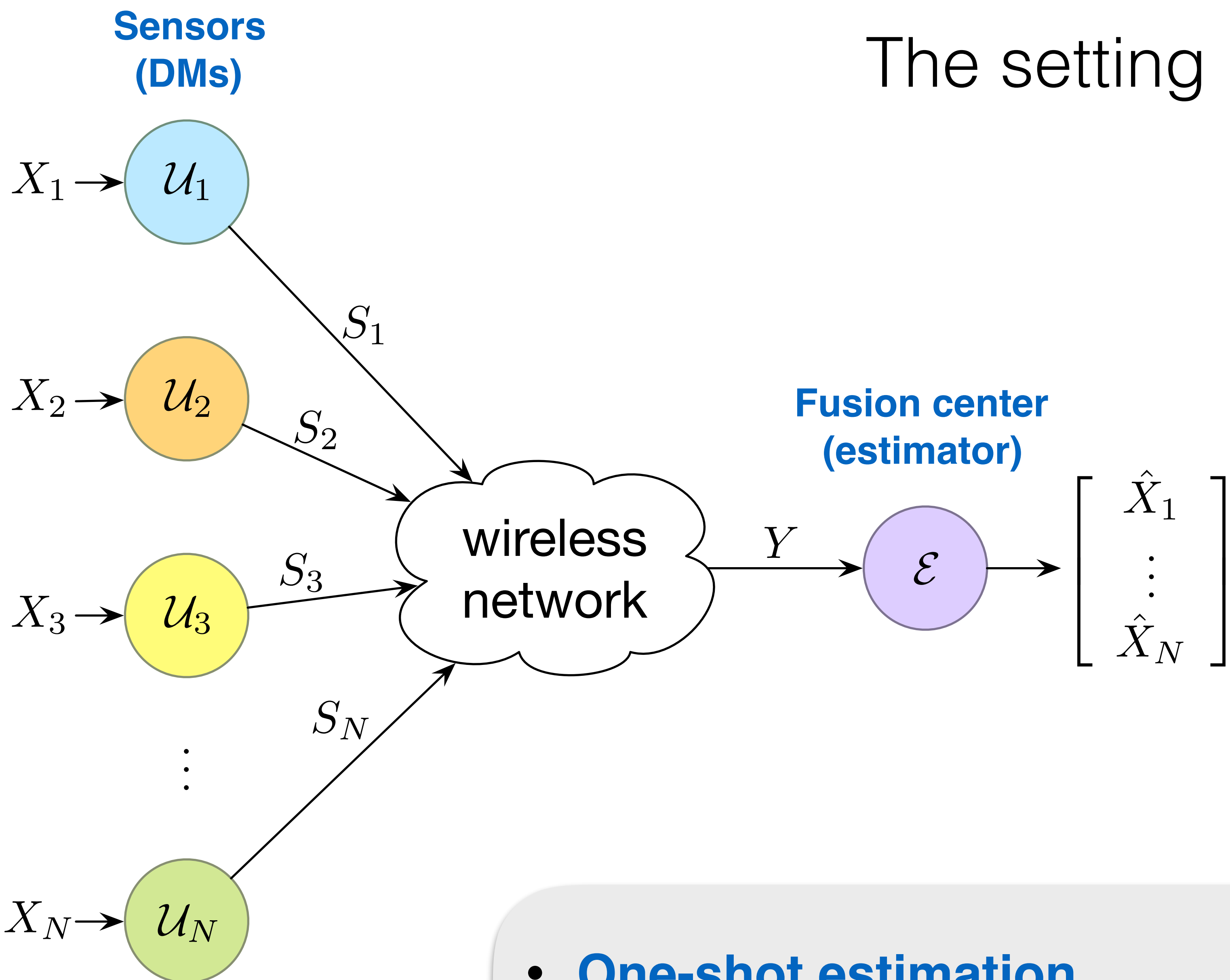
Dept. of Electrical and Computer Engineering

University of Maryland, College Park

Joint work with **Prof. Nuno C. Martins**

PGM seminar, UPenn 4-7-16

The setting



Wireless channel models

- Fading channels
- Rate-limited channels[1]
- Packet drop channels[2]
- **Collision channel**

- **One-shot estimation**
- Sensors and estimator are non-collocated
- Communication over a wireless network
- No coordination between sensors

1. Nair et al, “Feedback control under data rate constraints: An overview,” Proceedings IEEE, 2007.
2. Schenato et al, “Foundations of control and estimation over lossy networks,” Proceedings IEEE, 2007.

Applications

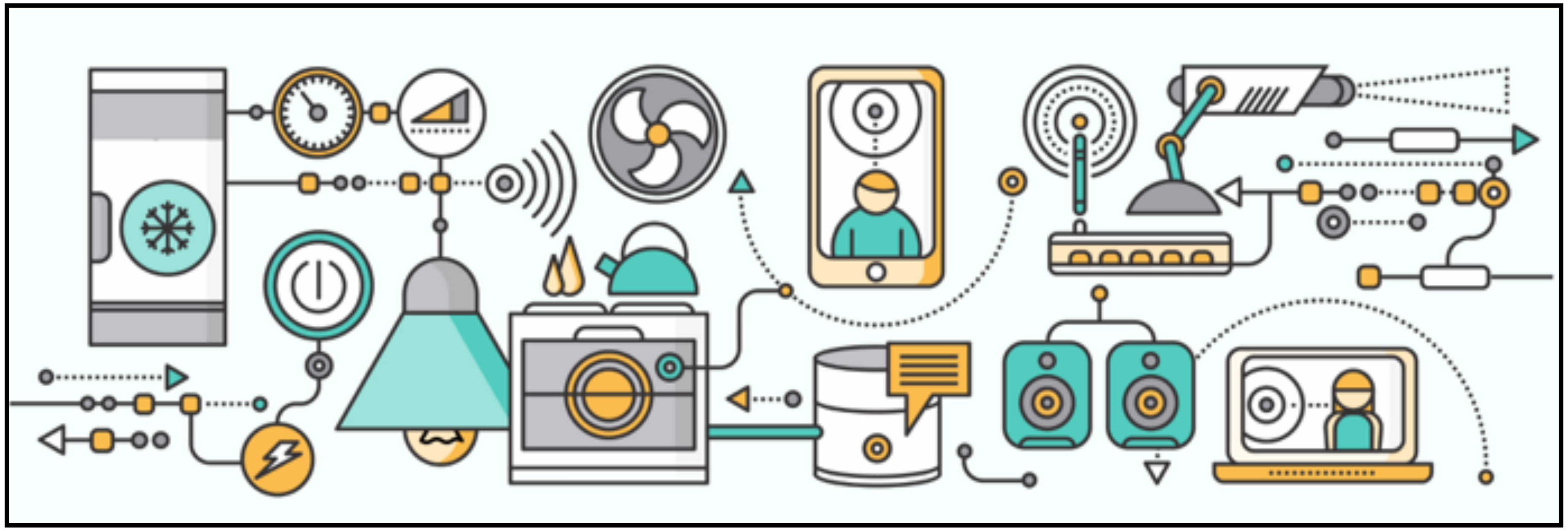
- Cyber-physical systems - **monitoring of large structures**

Oil & gas pipelines

Powergrid

Bridges

- Internet-of-things - **real time wireless networking**



Collision channel

Model of **interference**:

- Sensors decide whether to transmit or not
- Only one transmission can get through
- Widely used in wireless communications[1]

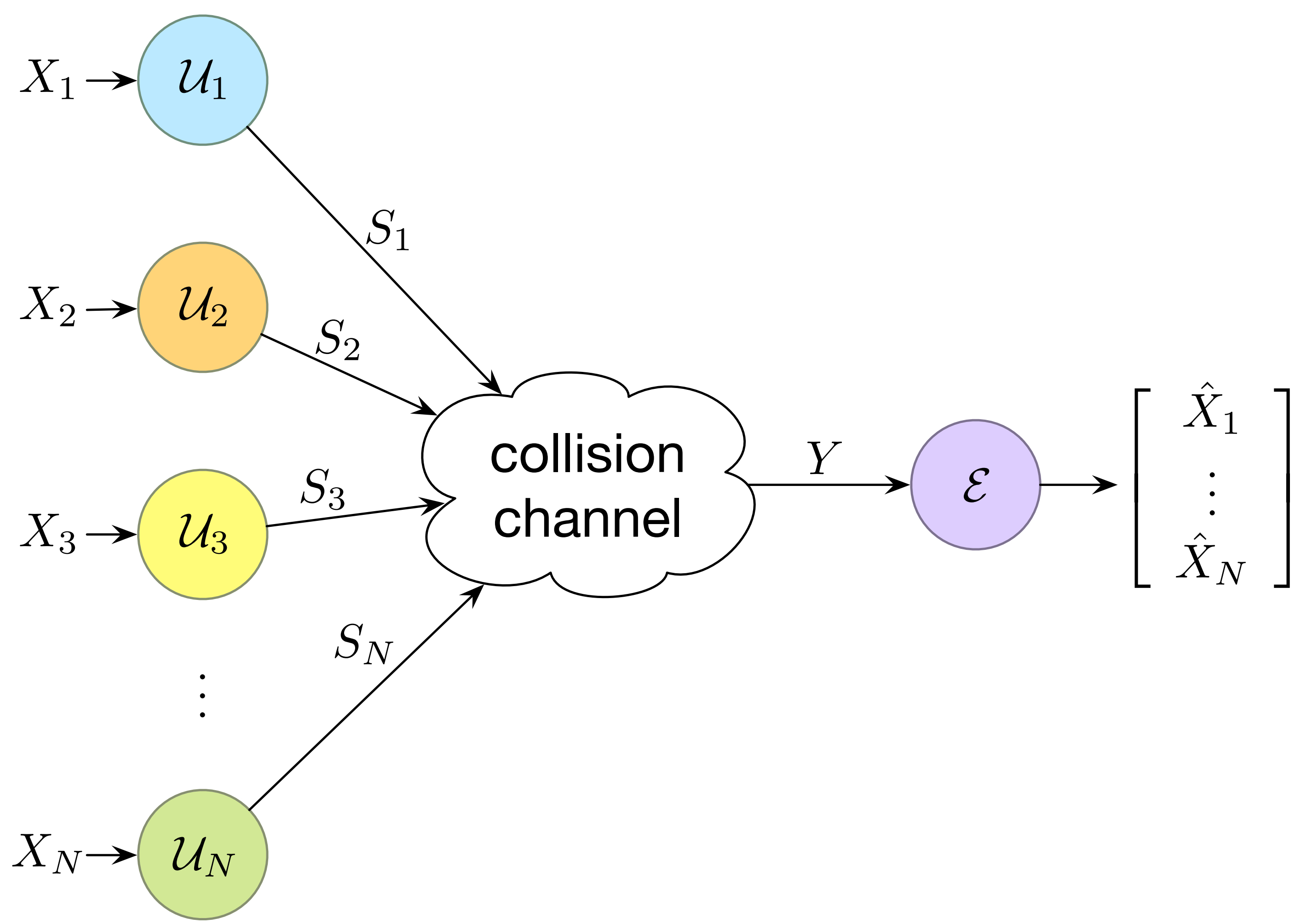
Decision variables: U_i

$$U_i = 1 \quad \Longrightarrow \quad S_i = (i, X_i)$$

(**transmit**)

$$U_i = 0 \quad \Longrightarrow \quad S_i = \varnothing$$

(**stay silent**)



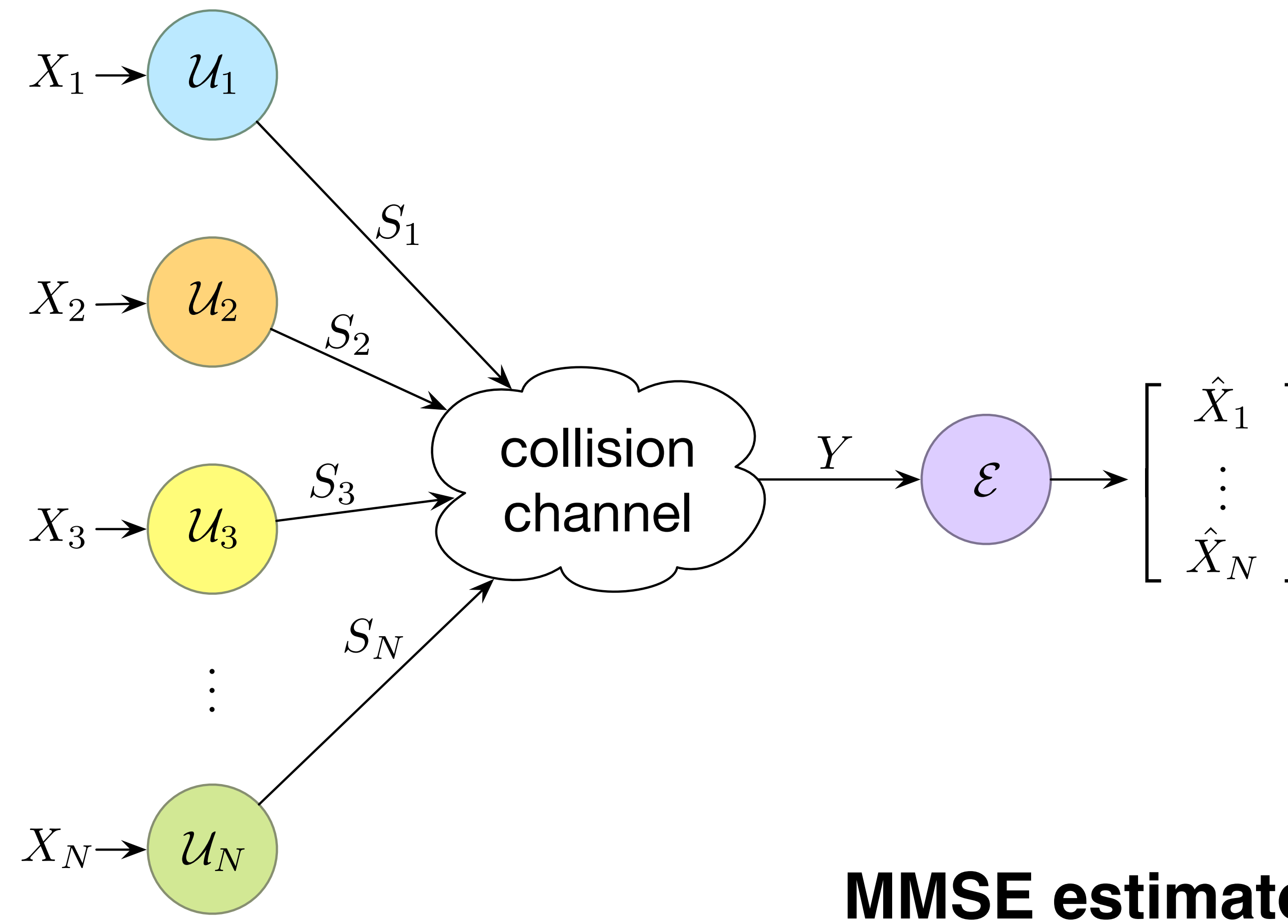
1. Goldsmith, *Wireless Communications*, ch. 14 - Multiuser systems, 2005.

MMSE estimation over the collision channel

$$W = [X_1, X_2, X_3, \cdots, X_N]$$

$$X_i, \quad i = 1, \cdots, N$$

- mutually independent
- continuous rvs
- supported on the real line
- arbitrarily distributed



Stochastic policies

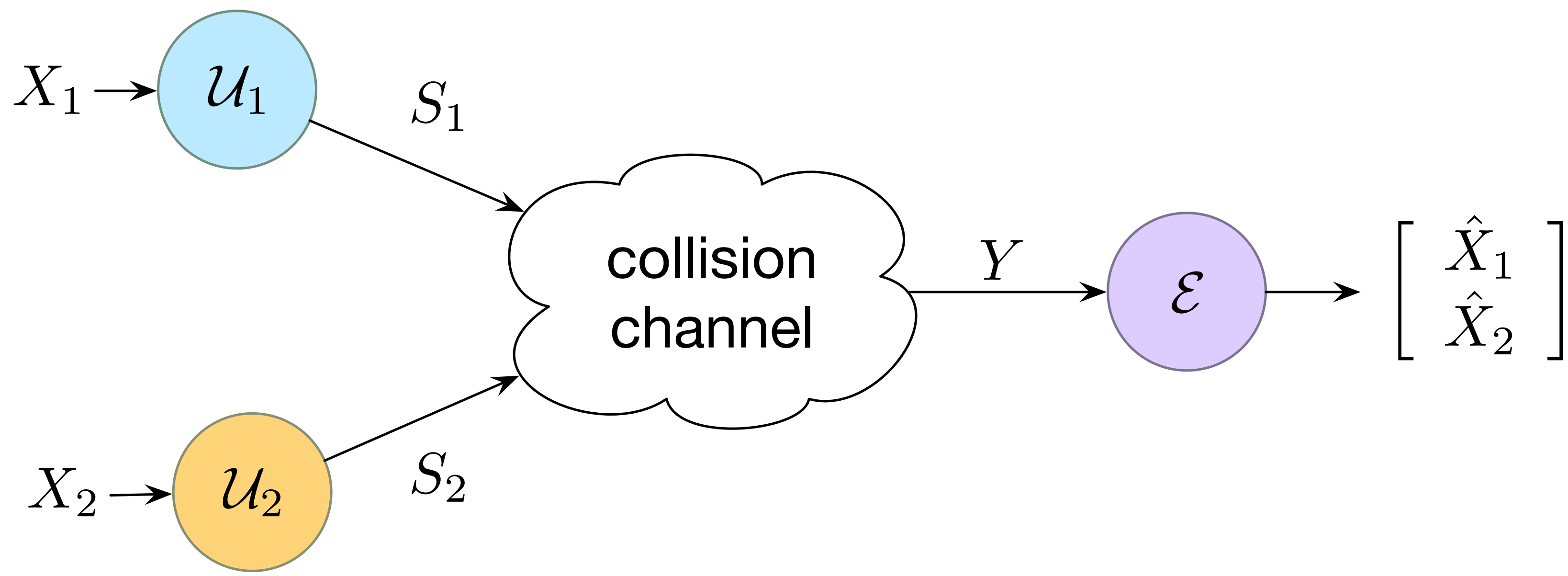
$$\mathbf{prob}(U_i = 1 | X_i = x_i) = \mathcal{U}_i(x_i)$$

MMSE estimator

$$\mathcal{E}(y) = \mathbf{E}\big[W | Y = y\big]$$

$$\begin{aligned} &\text{minimize} && \mathcal{J}(\mathcal{U}_1, \ldots, \mathcal{U}_N) = \mathbf{E}\left[\sum_{i=1}^{N}(X_i - \hat{X}_i)^2\right] \\ &\text{subject to} && \text{the constraint imposed by the collision channel} \end{aligned}$$

Simplest case: two sensors



$$\mathbf{prob}(U_i = 1 | X_i = x_i) = \mathcal{U}_i(x_i)$$

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

Problem 1

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

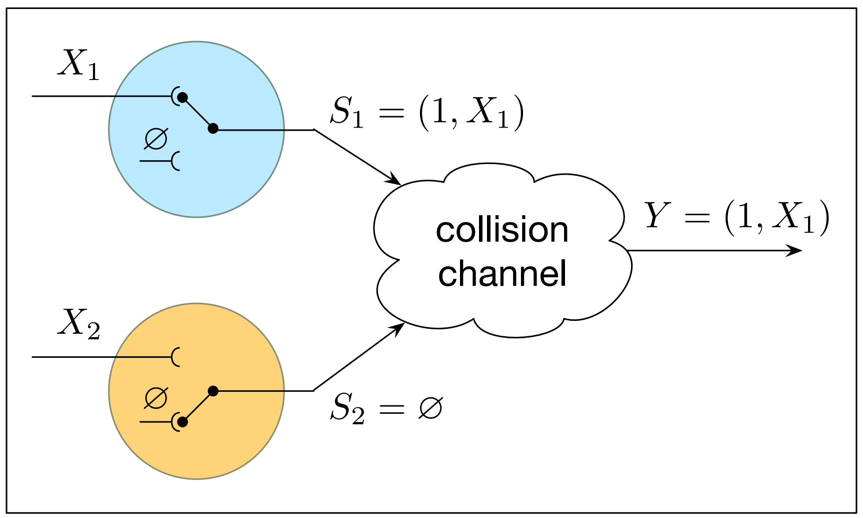
subject to $\quad$ the constraint imposed by the collision channel

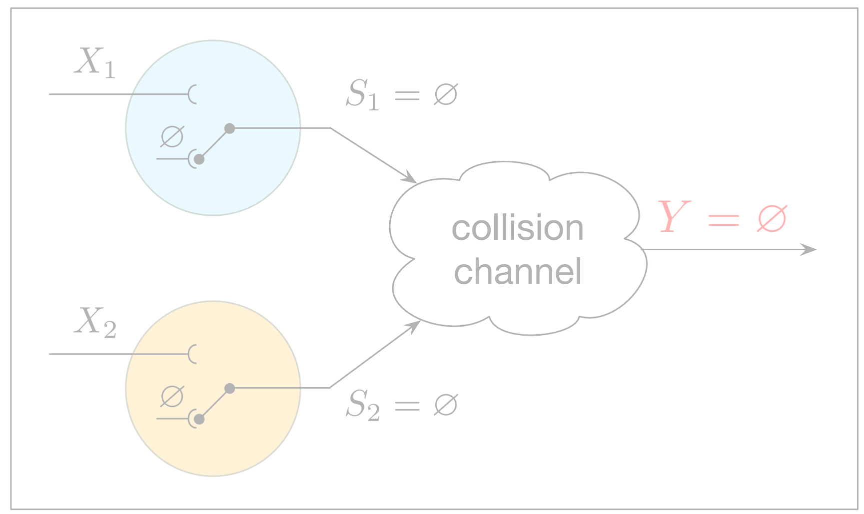
Collision channel

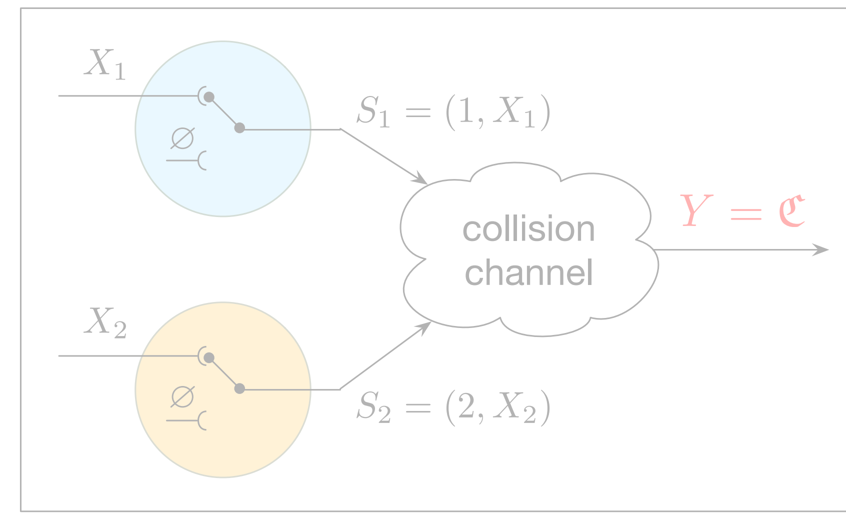
single transmission
$U_1 = 1, U_2 = 0$

no transmissions
$U_1 = 0, U_2 = 0$

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







success!

erasure $\varnothing$

collision $\mathfrak{C}$

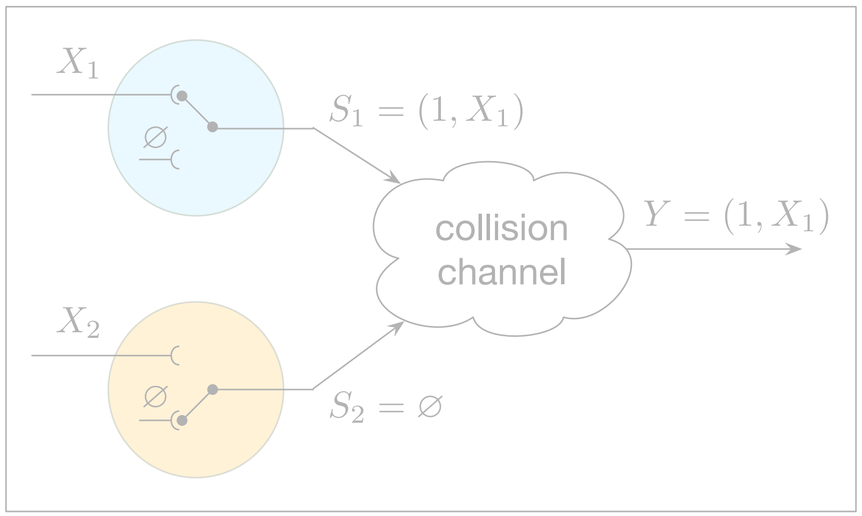
From the channel output we can always recover U_1 and U_2.

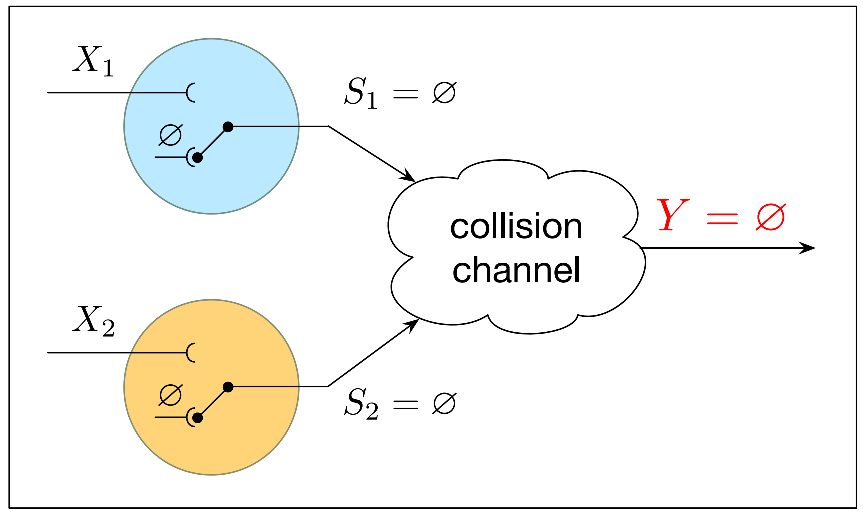
Collision channel

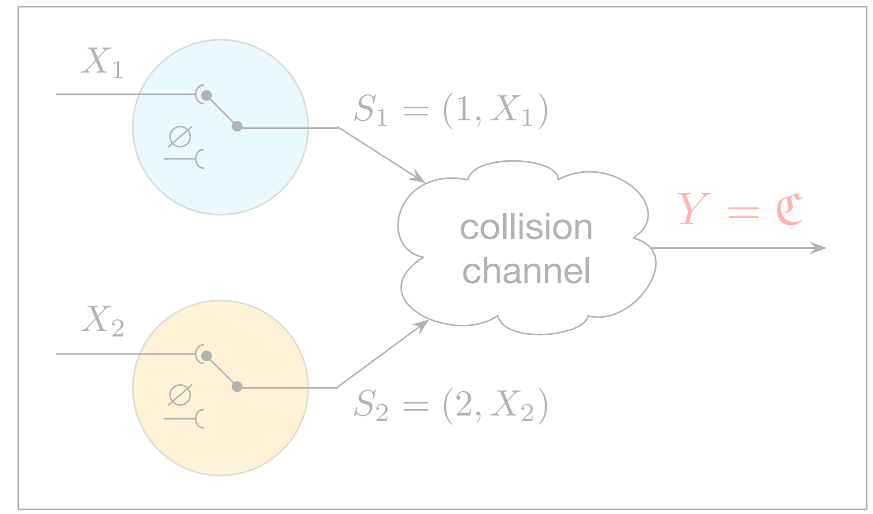
single transmission
$U_1 = 1, U_2 = 0$

no transmissions
$U_1 = 0, U_2 = 0$

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







success!

erasure $\varnothing$

collision $\mathfrak{C}$

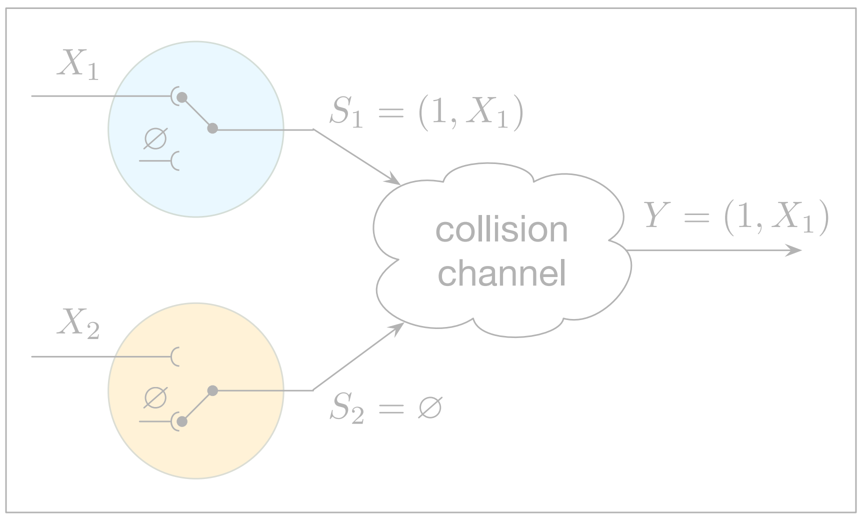
From the channel output we can always recover U_1 and U_2.

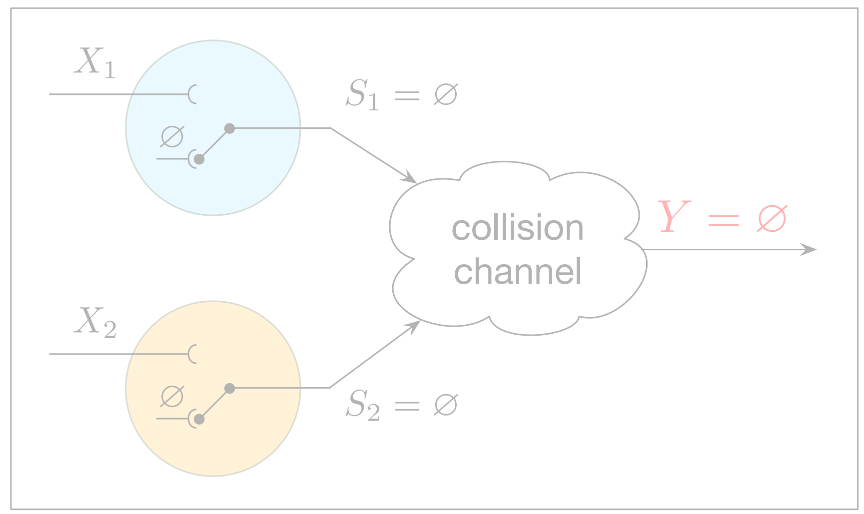
Collision channel

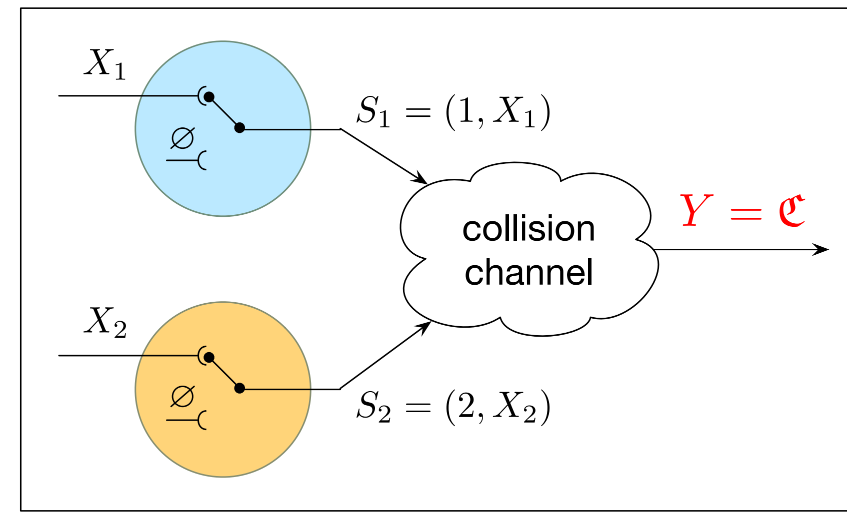
single transmission
$U_1 = 1, U_2 = 0$

no transmissions
$U_1 = 0, U_2 = 0$

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







success!

erasure $\varnothing$

collision $\mathfrak{C}$

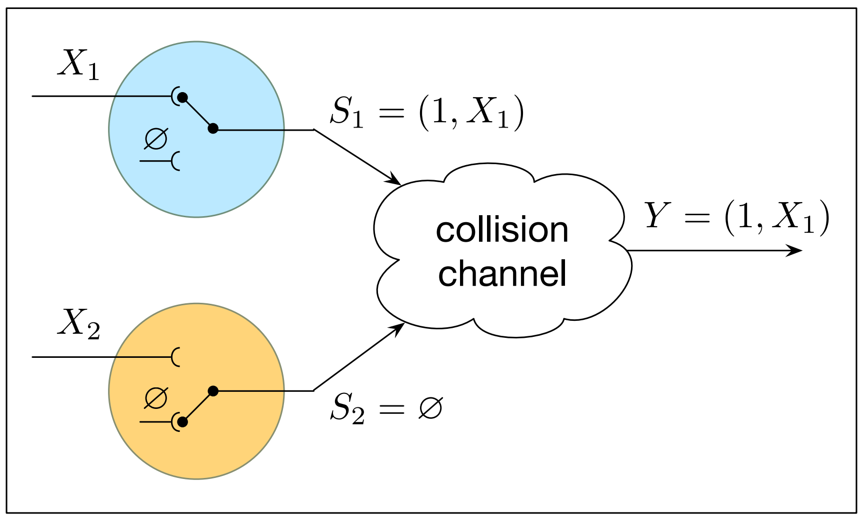
From the channel output we can always recover U_1 and U_2.

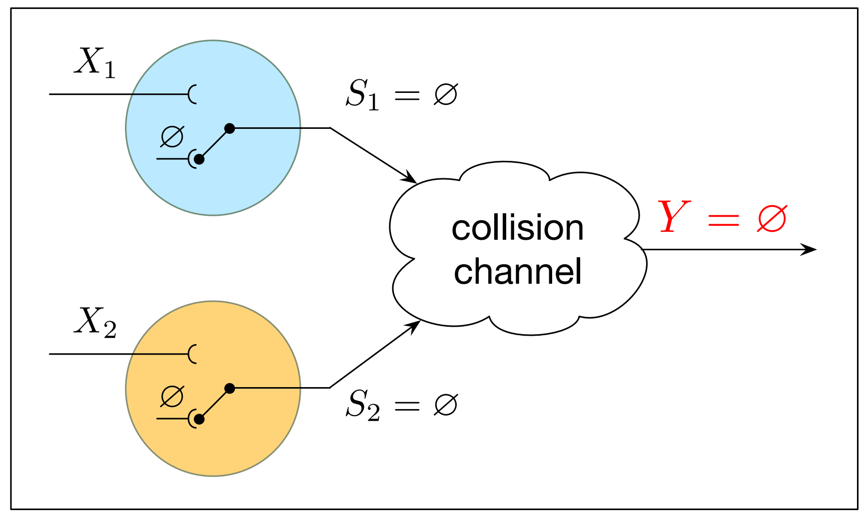
Collision channel

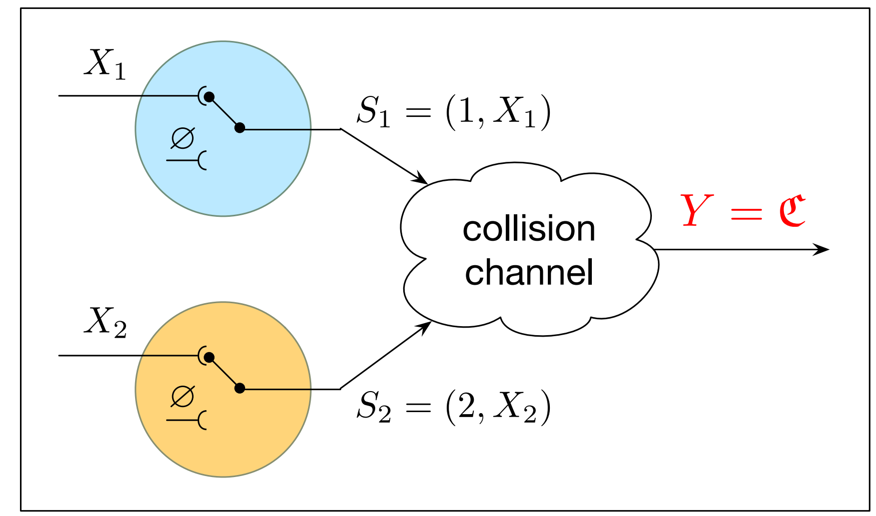
single transmission
$U_1 = 1, U_2 = 0$

no transmissions
$U_1 = 0, U_2 = 0$

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







success!

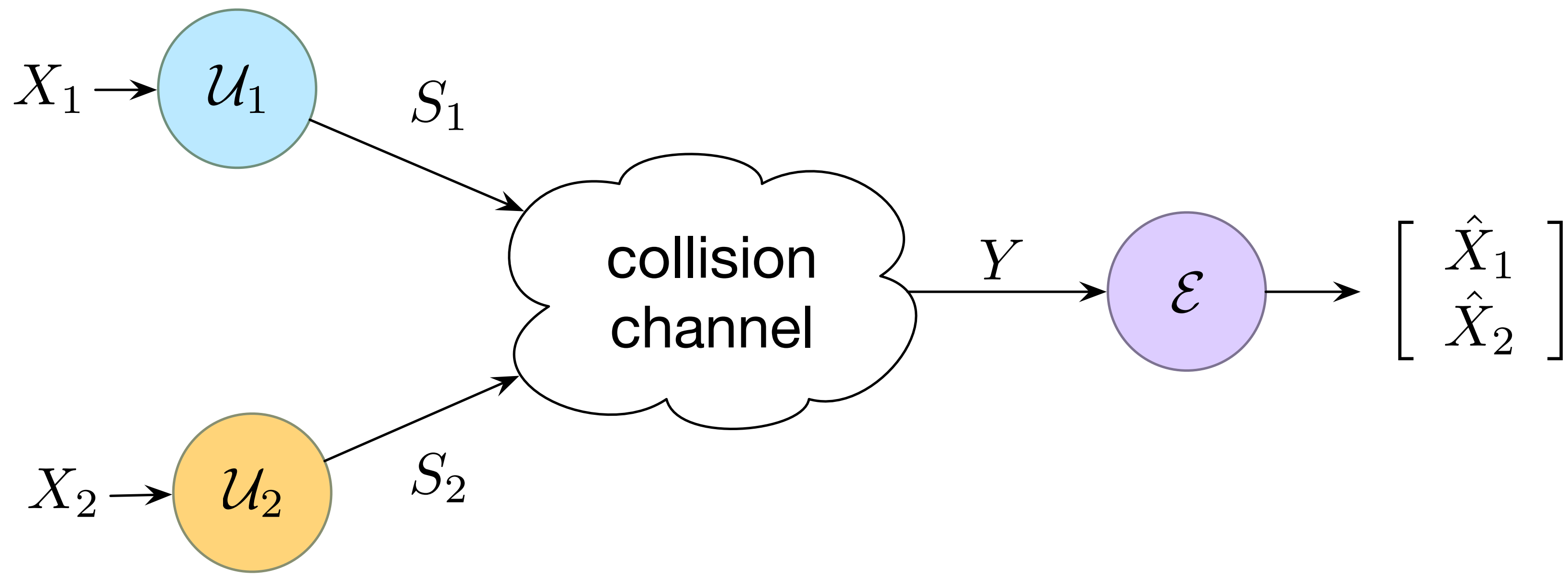
no transmission $\varnothing$

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". *System and Control Letters*, 2007.

Why is this problem interesting?



Problem 1

$$\begin{array}{ll} \text{minimize} & \mathcal{J}(\mathcal{U}_1, \mathcal{U}_2) = \mathbf{E}\left[(X_1 - \hat{X}_1)^2 + (X_2 - \hat{X}_2)^2\right] \\ \text{subject to} & \text{the constraint imposed by the collision channel} \end{array}$$

Static **team-decision problem** with **nonclassical information** structure $\Longrightarrow$ **Non-convex** (in most cases) **intractable**

1. Witsenhausen, "A counterexample in optimal stochastic control." SIAM J. Control '68.
2. Tsitsiklis & Athans, "On the complexity of decentralized decision making and detection problems". IEEE TAC '85.

A one-slide course in team theory

Team-optimality

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

$$\Longrightarrow$$

Person-by-person optimality

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

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

Finding team-optimal solutions is very hard, in general!

Characterization: Find a class of policies with a certain structure which contains the set of team-optimal solutions

Person-by-person approach

Find a structure that **holds for every person-by-person optimal solution**

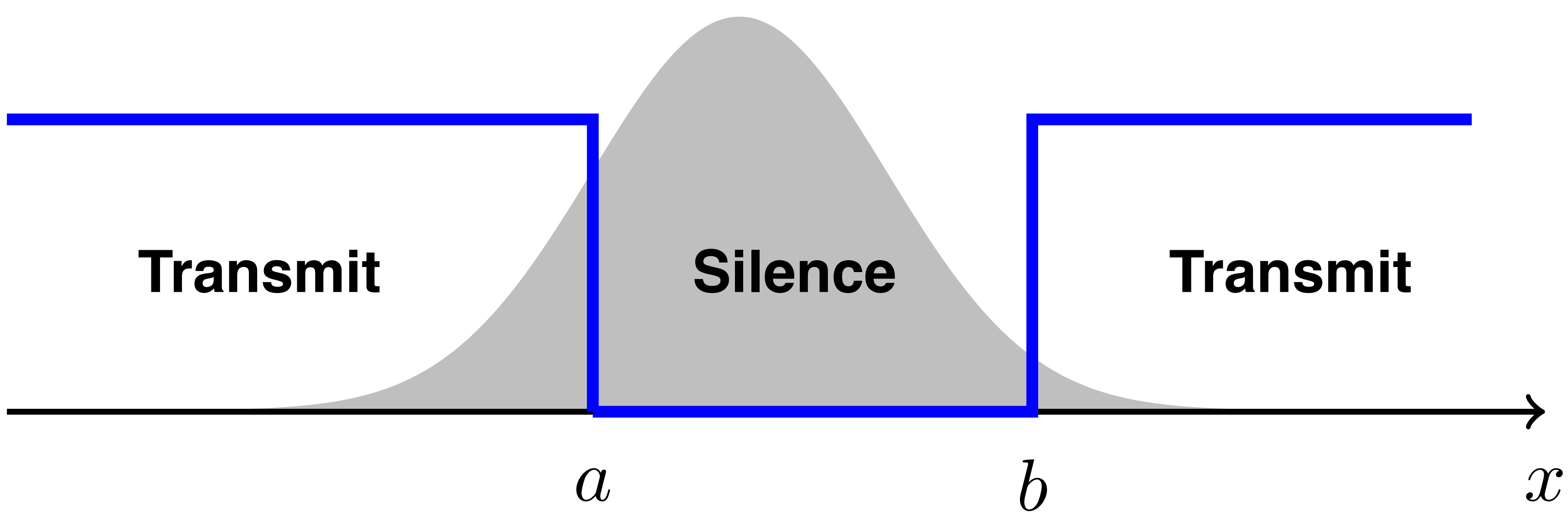
1. Yuksel & Basar, Stochastic networked control systems. Birkhauser 2013.
2. Mahajan et al, "Information structures in optimal decentralized control", CDC 2012.

Deterministic threshold policies

Definition: Threshold policy

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

$$\mathcal{U}(x)$$

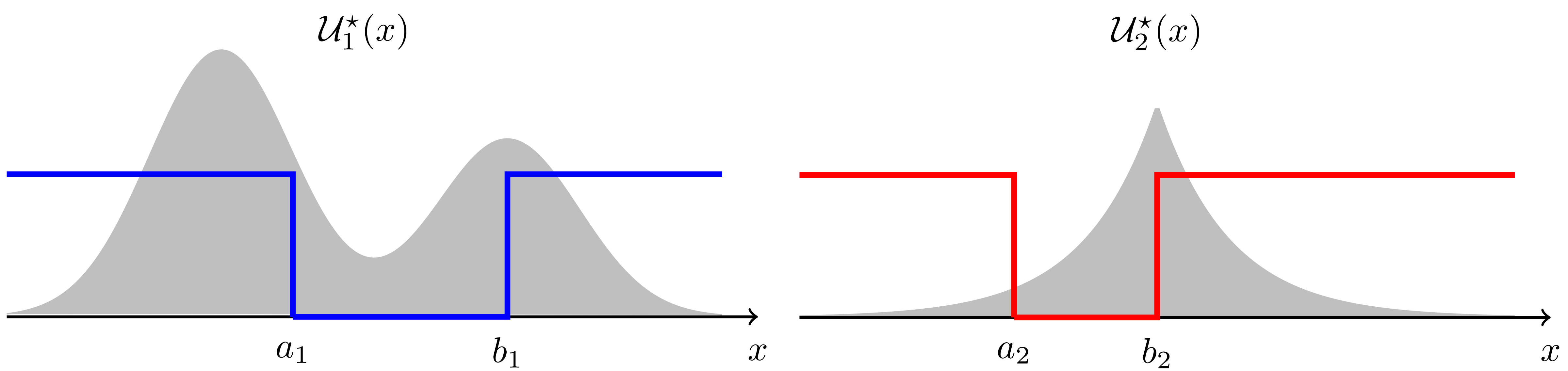


1. Imer & Basar, “Optimal estimation with limited measurements”, *JSCC* 2010
2. Lipsa & Martins, “Remote state estimation with communication costs for first-order LTI systems”. *IEEE TAC* 2011.

Characterization of team-optimal policies

Theorem:

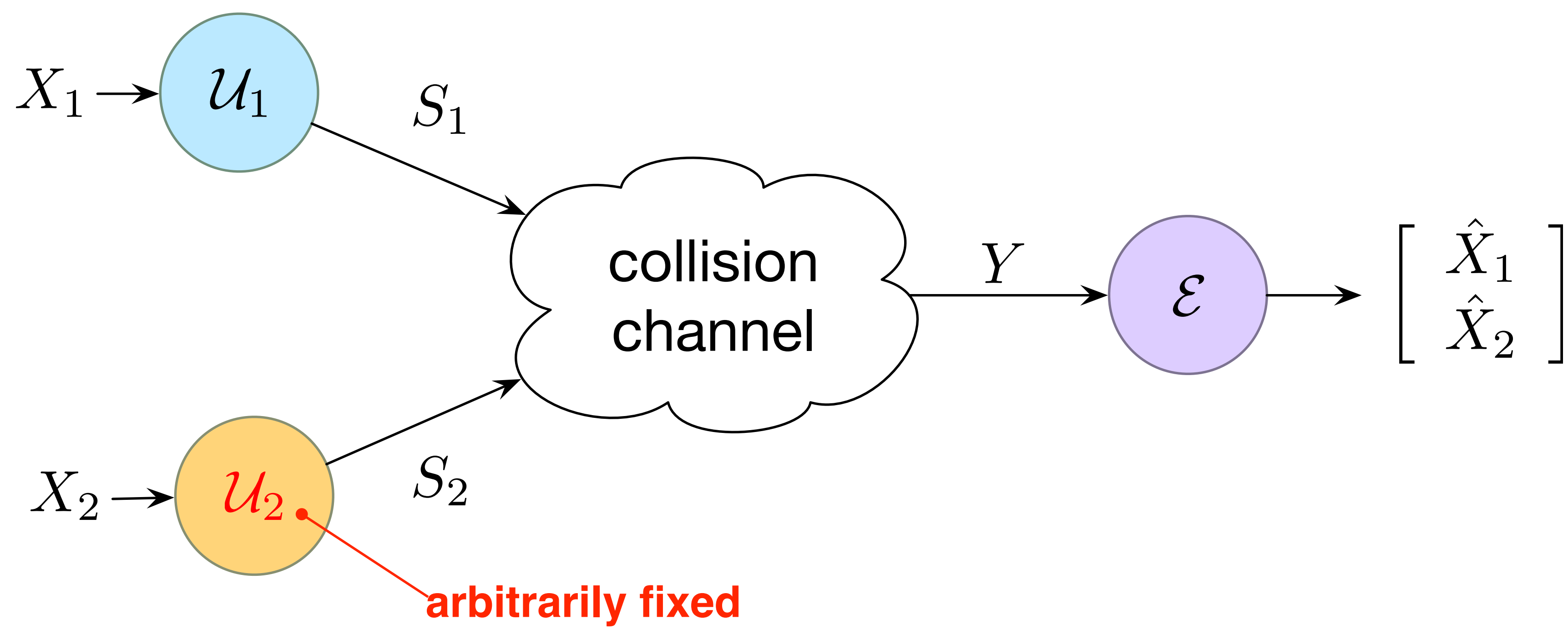
There exists a pair of **threshold policies** that achieves the optimal cost in Problem 1.



Sketch of Proof:

- Step 1: Person-by-person optimality approach
- Step 2: Lagrange duality for infinite dimensional LPs

Remote estimation with communication costs



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$_1$:

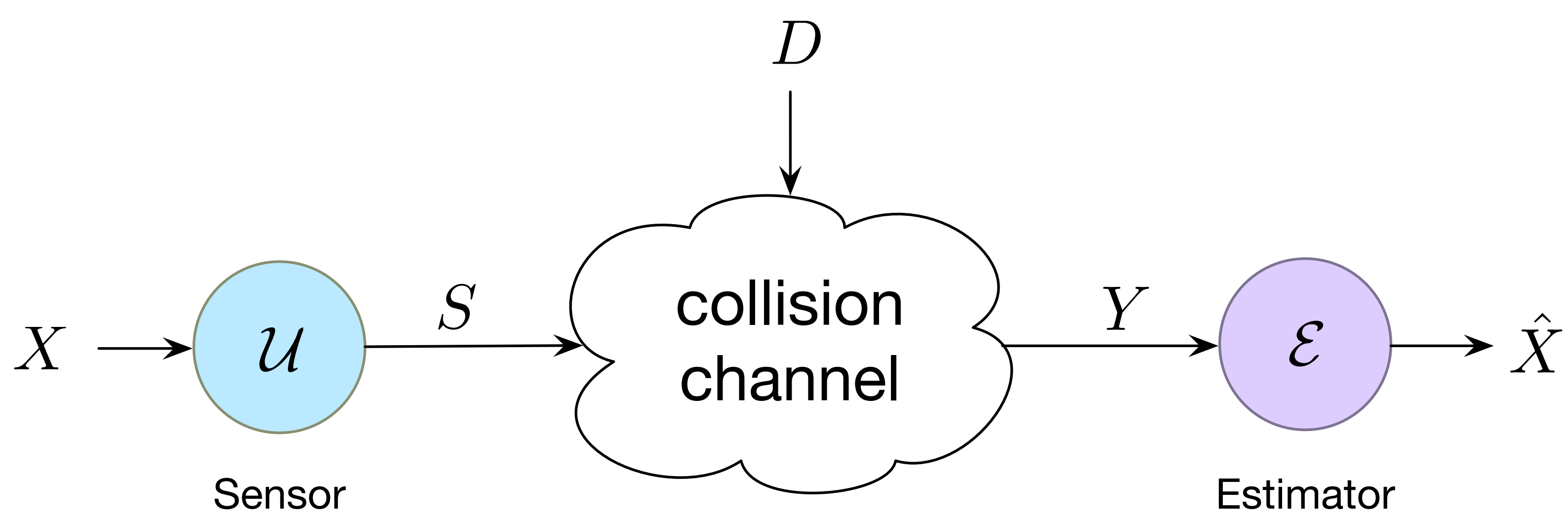
$$\mathcal{J}_1(\mathcal{U}_1) = \mathbf{E}\Big[(X_1 - \hat{X}_1)^2\Big] + \rho_2 \cdot \mathbf{prob}(U_1 = 1) + \theta_2$$

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

Communication cost:

$$\rho_2 = \mathbf{E}\Big[(X_2 - \hat{X}_2)^2 | U_1 = 1\Big] - \mathbf{E}\Big[(X_2 - \hat{X}_2)^2 | U_1 = 0\Big] \geq 0$$

Remote estimation with communication costs



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

Determines if the channel is occupied or not

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

Problem 2

minimize $\quad \mathcal{J}(\mathcal{U}) = \mathbf{E}\Big[(X - \hat{X})^2\Big] + \rho \cdot \mathbf{prob}(U = 1)$

subject to $\quad$ the constraint imposed by the collision channel

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

Lemma:

There exists a **threshold policy** that is optimal for Problem 2.

Sketch of Proof

1. Express the cost as

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

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

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

$$\mathbf{prob}(U=1)=\alpha \qquad \mathbf{E}[X|U=0]=\gamma \qquad \mathcal{G}(x)=\frac{1-\mathcal{U}(x)}{1-\alpha}$$

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

moment optimization problem with variable bounds (convex)

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

Sketch of Proof

4. The Lagrange dual function for this problem is given by

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

strong duality holds

5. Let ν^* be solution to the dual problem, the solution to the primal problem is

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

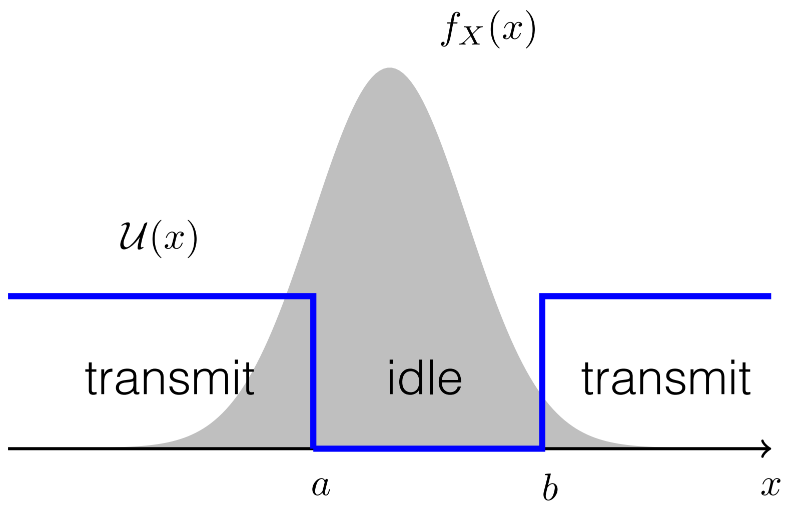
6. In the original optimization variable, we have

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

■

Remarks

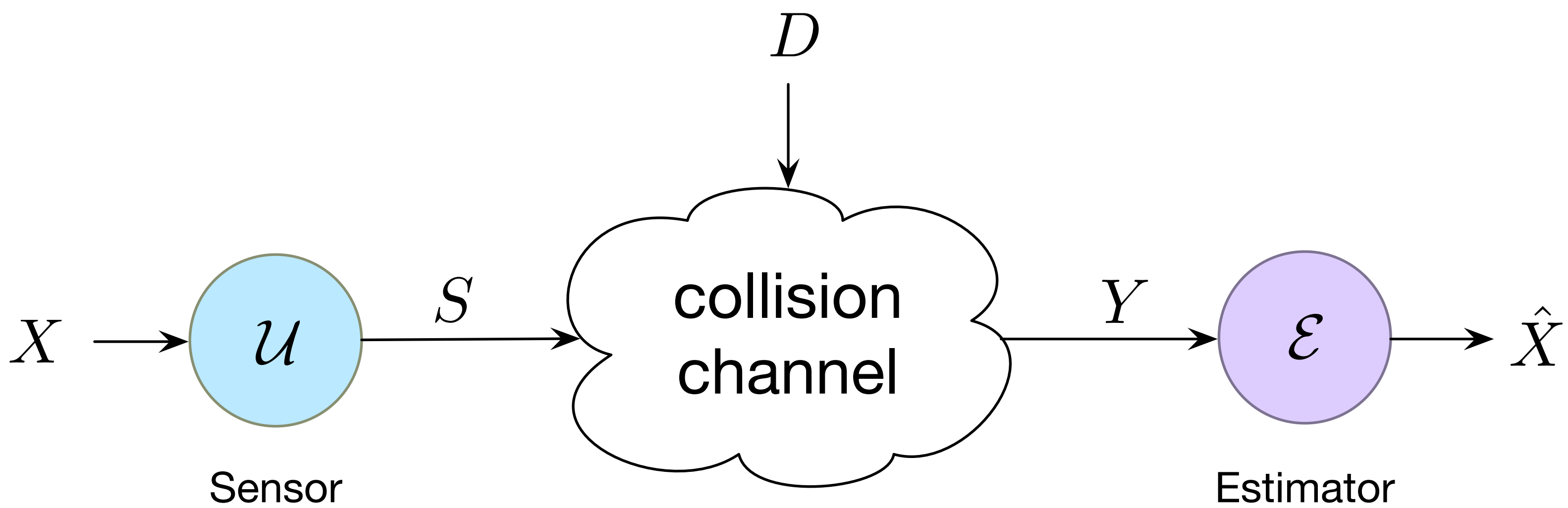
1. The structure is **deterministic**
2. Valid for **any probability distribution**
3. **Vector observations** and **any number of sensors**



How do we compute the optimal thresholds?

1. Vasconcelos & Martins, “Optimal estimation over the collision channel”. IEEE TAC 2017.
2. Vasconcelos & Martins, “Estimation over the collision channel: structural results”. Allerton 2013.

Computing optimal thresholds and representation points



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

$$\mathcal{E}(y) = \begin{cases} x & y = x \\ \hat{x}_{\varnothing} & y = \varnothing \\ \hat{x}_{\mathfrak{C}} & y = \mathfrak{C} \end{cases}$$

$$\mathcal{J}(a, b, \hat{x}_{\varnothing}, \hat{x}_{\mathfrak{C}}) = \int_{[a,b]} (x - \hat{x}_{\varnothing})^2 f_X(x)dx + \int_{\bar{\mathbb{R}}\setminus[a,b]} [\beta(x - \hat{x}_{\mathfrak{C}})^2 + \rho] f_X(x)dx$$

binary quantization with asymmetric distortion

1. Lloyd, "Least squares quantization in PCM", IEEE Trans. Info. Theory, 1982
2. Fleischer, "Sufficient conditions for achieving minimum distortion in quantizer", IEEE Int. Conv, Rec., 1964

Binary quantization with asymmetric distortion

$$\begin{array}{ll} \text{minimize} & \mathcal{J}(a, b, \hat{x}_\varnothing, \hat{x}_\mathfrak{c}) \\ \text{subject to} & a \leq b \end{array}$$

$$x \in [a^*, b^*] \Leftrightarrow (x - \hat{x}_\varnothing)^2 \leq \beta(x - \hat{x}_\mathfrak{c})^2 + \rho$$

necessary optimality condition

Let $\hat{x} = (\hat{x}_\varnothing, \hat{x}_\mathfrak{c})$ denote the pair of estimates in case of erasure and collision

$$a(\hat{x}), b(\hat{x}) = \frac{1}{1-\beta}\left[(\hat{x}_\varnothing - \beta\hat{x}_\mathfrak{c}) \pm \sqrt{\beta(\hat{x}_\varnothing - \hat{x}_\mathfrak{c})^2 + (1-\beta)\rho}\right]$$

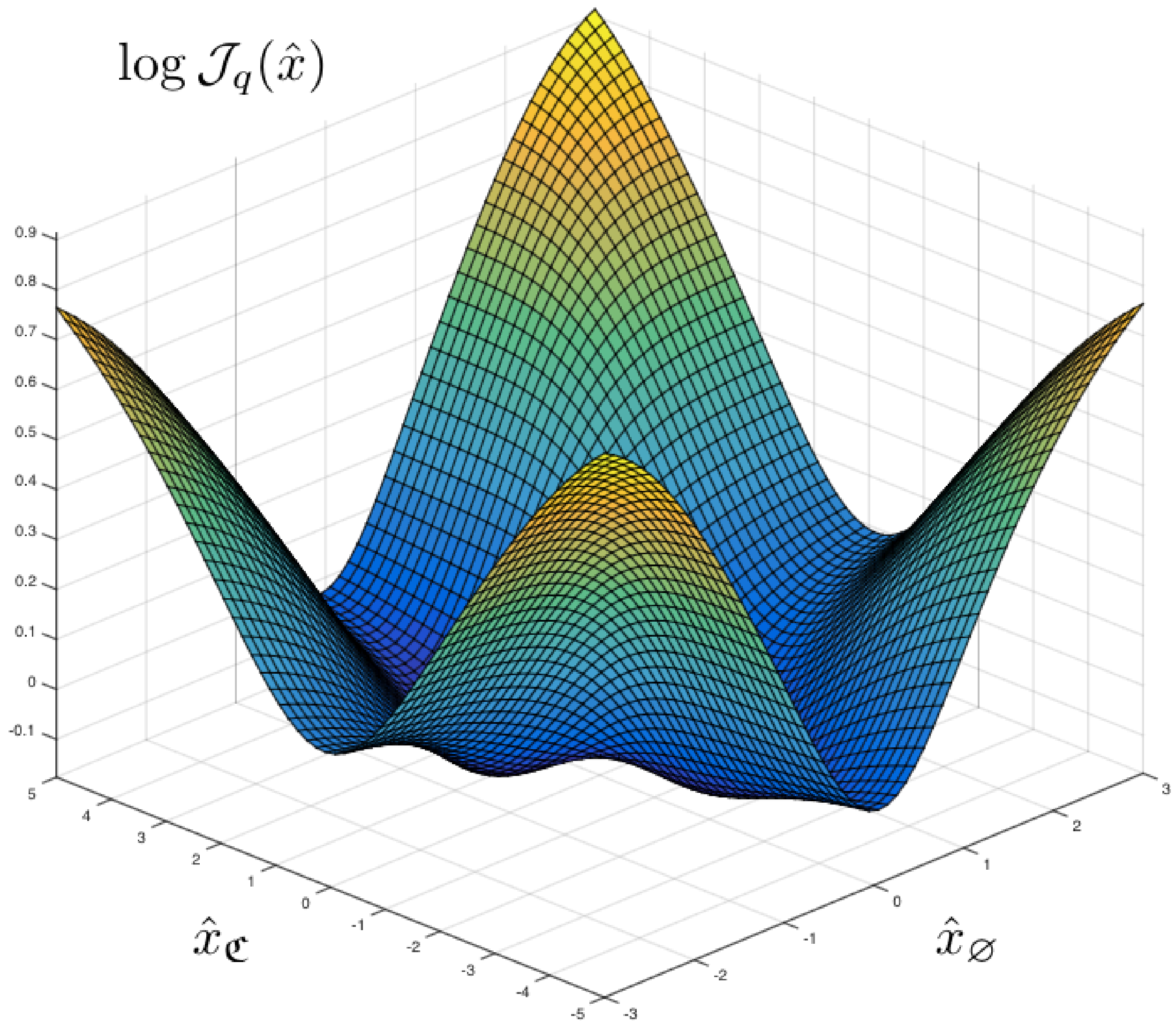
Define a new cost: $\mathcal{J}_q(\hat{x}) = \mathcal{J}(a(\hat{x}), b(\hat{x}), \hat{x}_\varnothing, \hat{x}_\mathfrak{c})$

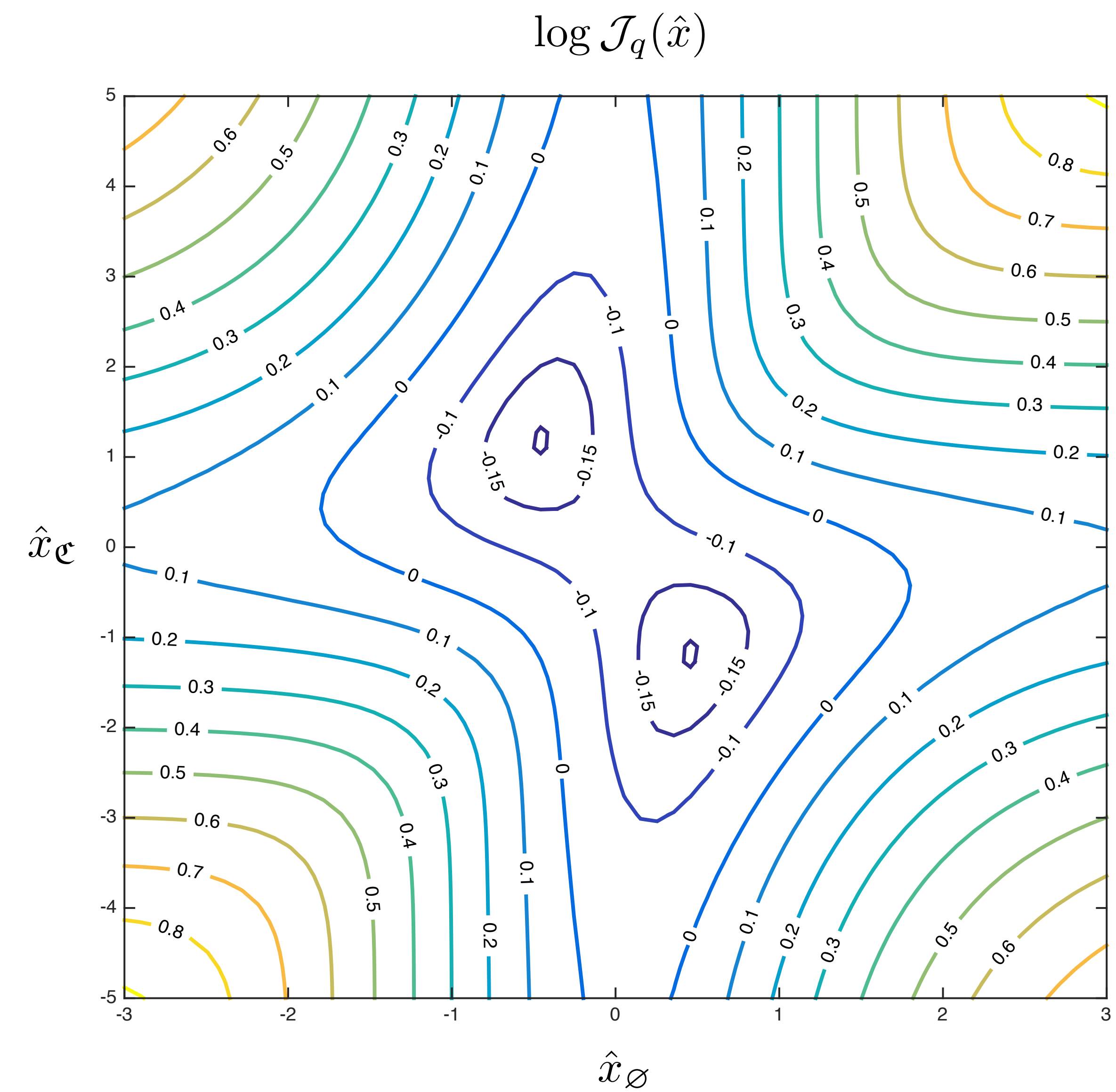
$$\text{minimize} \quad \mathcal{J}_q(\hat{x})$$

unconstrained problem

Quantizer distortion function

$X \sim \mathcal{N}(0,1)$ $\qquad \beta = 0.5$ $\qquad \rho = 1$





Neither convex nor quasi-convex

Continuously differentiable

Modified Lloyd-Max Algorithm

$$\underset{\hat{x}\in\mathbb{R}^2}{\text{minimize}} \quad \mathcal{J}_q(\hat{x})$$

$$\nabla \mathcal{J}_q(\hat{x}) = 0 \qquad \longleftrightarrow \qquad \hat{x} = \mathcal{F}(\hat{x})$$

Lloyd's Map

$$\mathcal{F}(\hat{x}) = \begin{bmatrix} \mathbf{E}\Big[X \,|\, X \in [a(\hat{x}), b(\hat{x})]\Big] \\ \mathbf{E}\Big[X \,|\, X \notin [a(\hat{x}), b(\hat{x})]\Big] \end{bmatrix}$$

Modified Lloyd-Max

$$\hat{x}^{(0)} \neq (0,0)$$

$$\hat{x}^{(k+1)} = \mathcal{F}(\hat{x}^{(k)}), \quad k = 0, 1, \ldots$$

Step 1 From $\hat{x}^{(k)}$ update the thresholds $a(\hat{x}^{(k)})$ and $b(\hat{x}^{(k)})$

Step 2 Compute the centroids of the new quantization regions

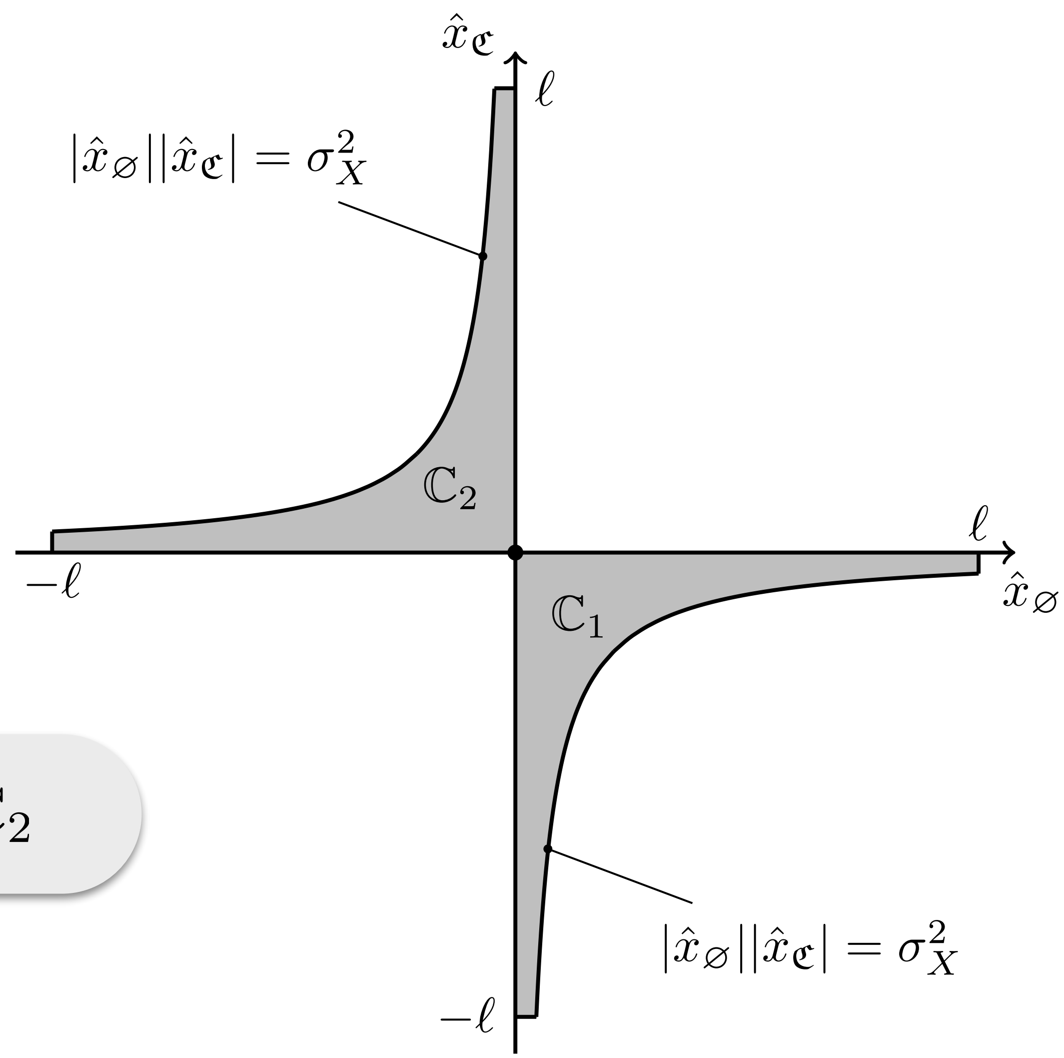
Convergence

Theorem:

For $X \sim \mathcal{N}(0, \sigma_X^2)$ the Modified Lloyd-Max algorithm is **globally convergent** to a critical point of $\mathcal{J}_q(\hat{x})$.

Sketch of Proof:

- Find a compact set $\mathbb{C}$ that contains all the critical points of $\mathcal{J}_q(\hat{x})$
- Show that $\mathcal{F}(\mathbb{C}) \subset \mathbb{C}$

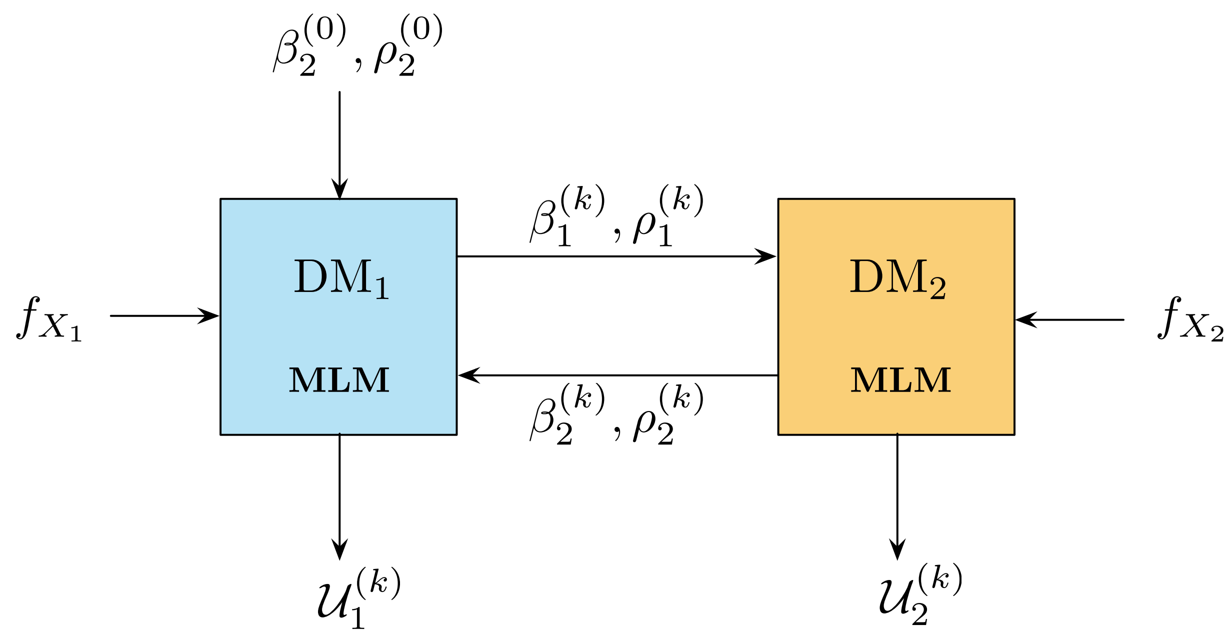


$$\mathbb{C} = \mathbb{C}_1 \cup \mathbb{C}_2$$

■

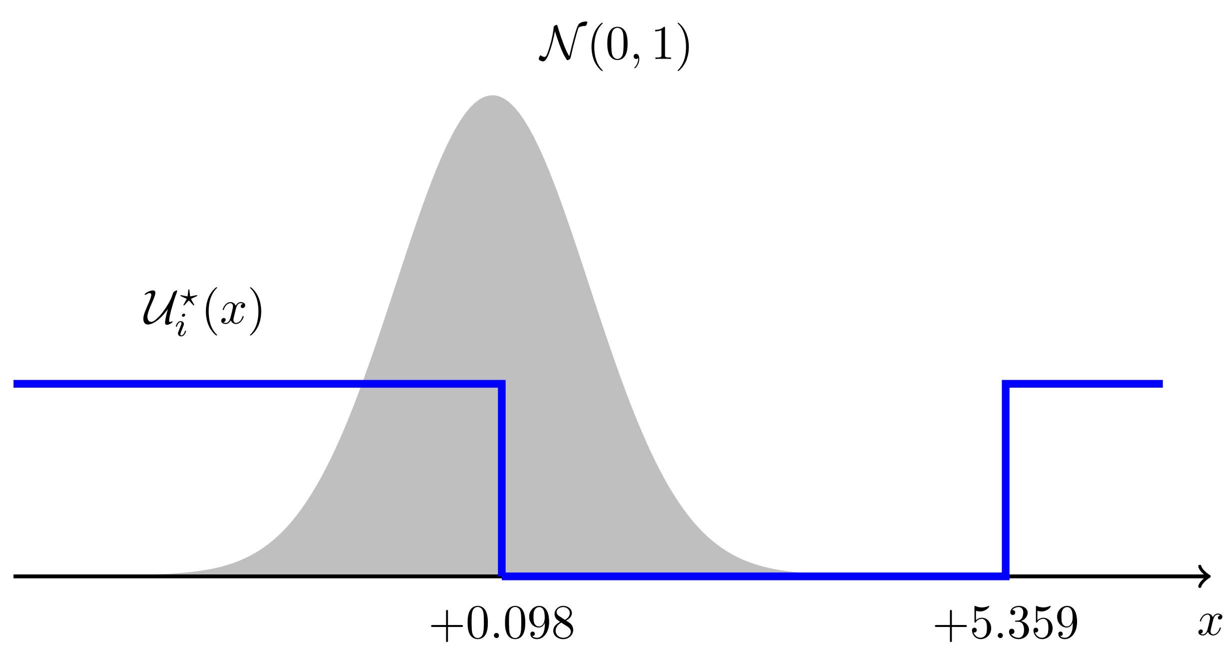
Vasconcelos & Martins, "Optimal thresholds for remote estimation over the collision ch. with comm. costs". CDC 2015.

Numerical procedure



Repeat until the cost cannot be further reduced

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



iid observations, symmetric pdf

asymmetric thresholds

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

Gain of 46% over scheduling policies

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

Collision channel with common and private observations

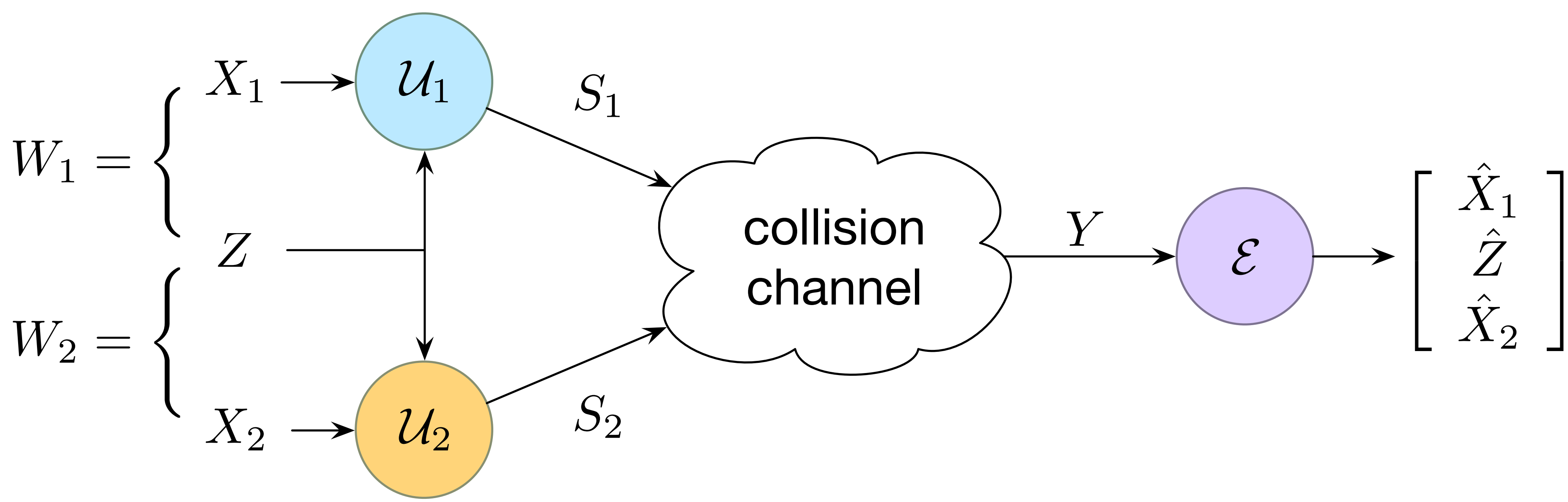
$$W = \begin{bmatrix} X_1 \\ Z \\ X_2 \end{bmatrix} \qquad W_i = \begin{bmatrix} X_i \\ Z \end{bmatrix}$$

private observation

common observation

$$f_W = f_Z \cdot f_{X_1|Z} \cdot f_{X_2|Z}$$

$$X_1 \leftrightarrow Z \leftrightarrow X_2$$



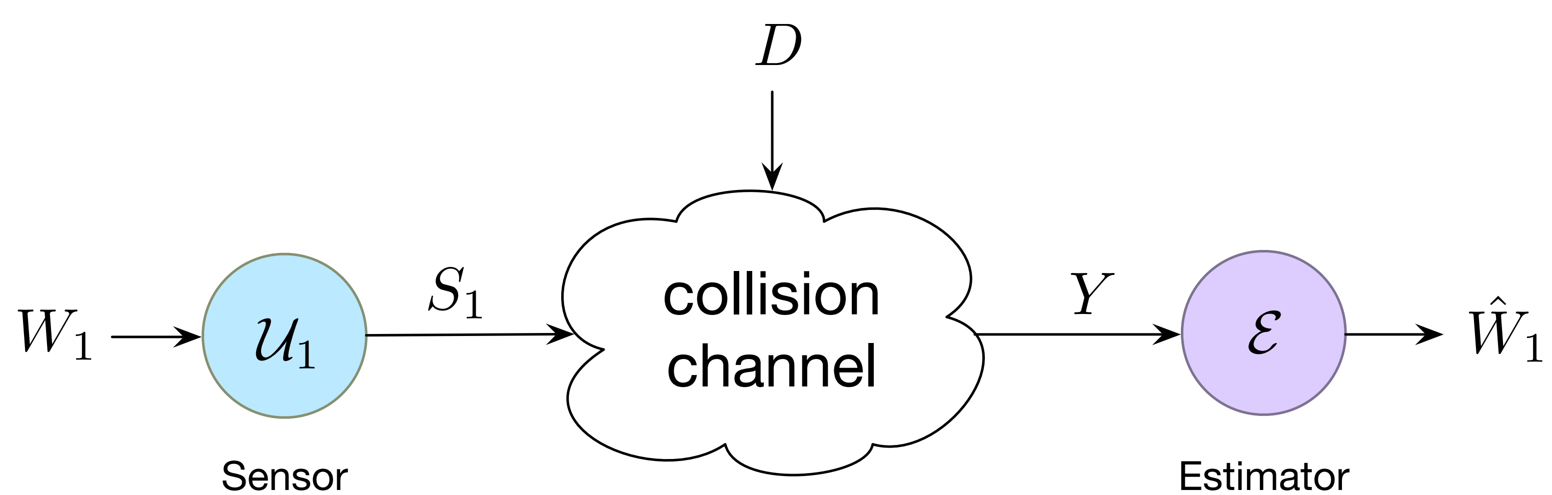
Problem 3

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

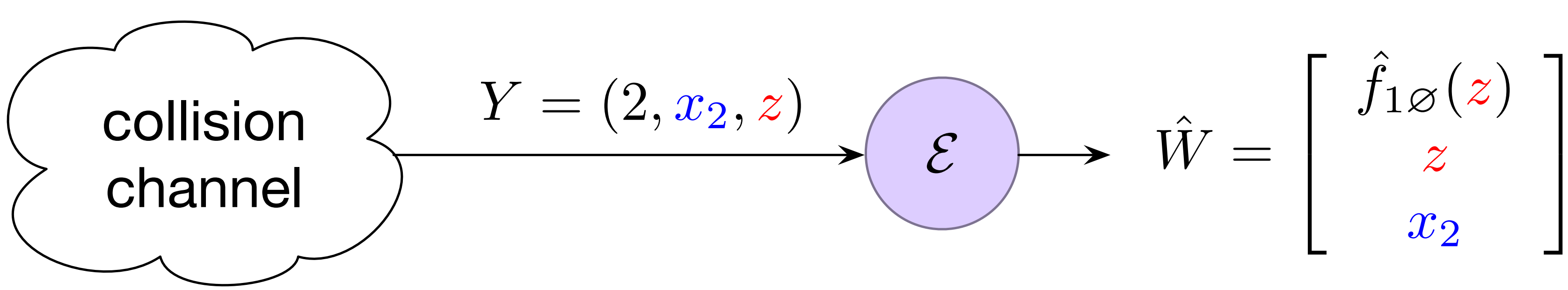
subject to $\quad$ the constraint imposed by the collision channel

Two difficulties

The event that there is a concurring transmission is **not independent** of W_1 .

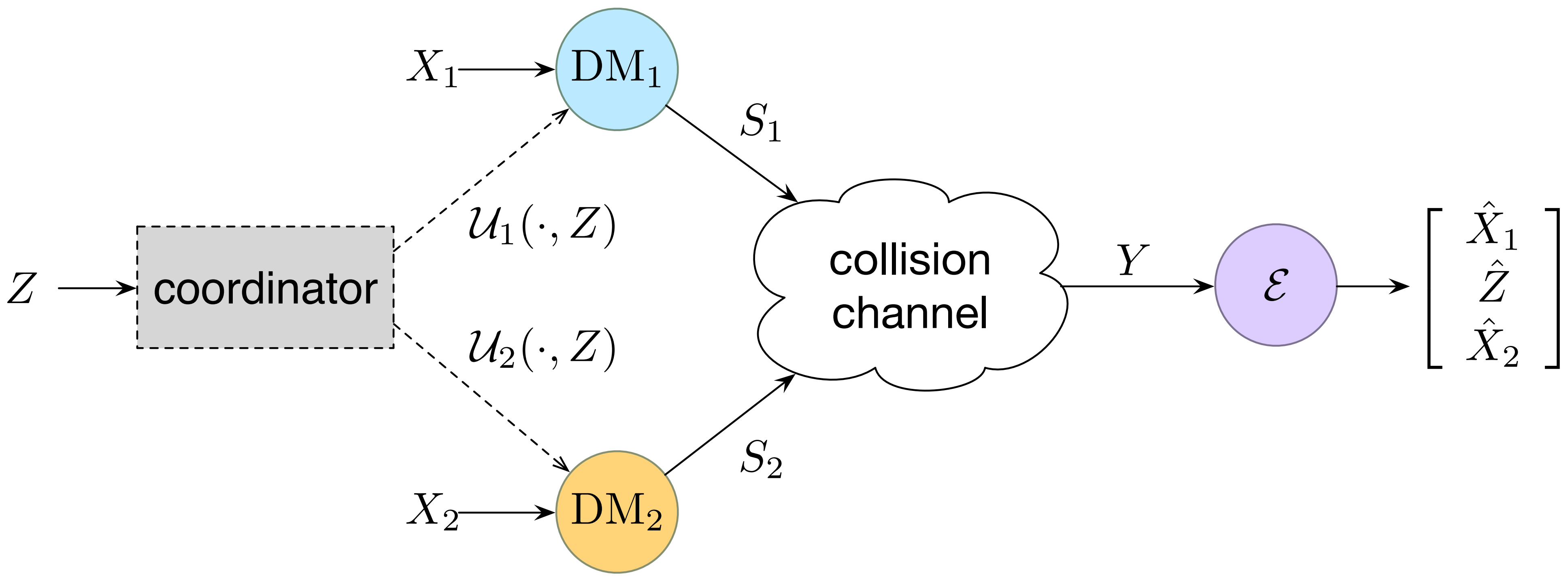


A successful transmission made by sensor 2 **contains side information** for the estimation of X_1.



Common information approach[1]

Common information can be used to **simplify** and **characterize** optimal solutions of team problems.



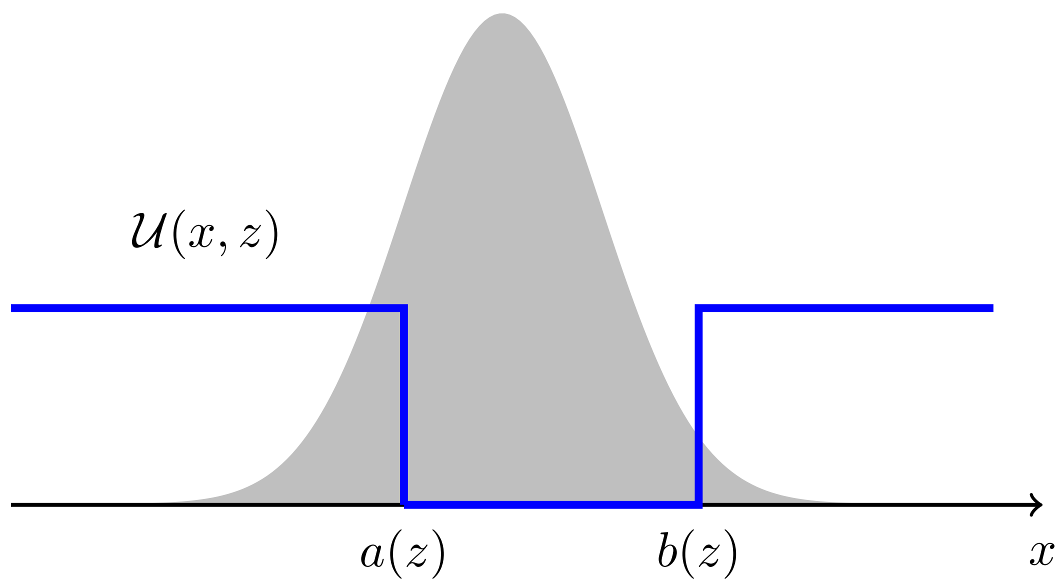
Minimize	$\mathcal{J}^z(\mathcal{U}_1, \mathcal{U}_2) = \mathbf{E}\left[(W - \hat{W})^T (W - \hat{W}) \mid Z = z\right]$
Subject to	the constraint imposed by the collision channel

[1] Nayyar, Mahajan & Teneketzis, "Decentralized Stochastic Control with Partial History Sharing: A Common Information Approach". IEEE TAC 2013.

Threshold structure on private information

Definition: Threshold policy on private information

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

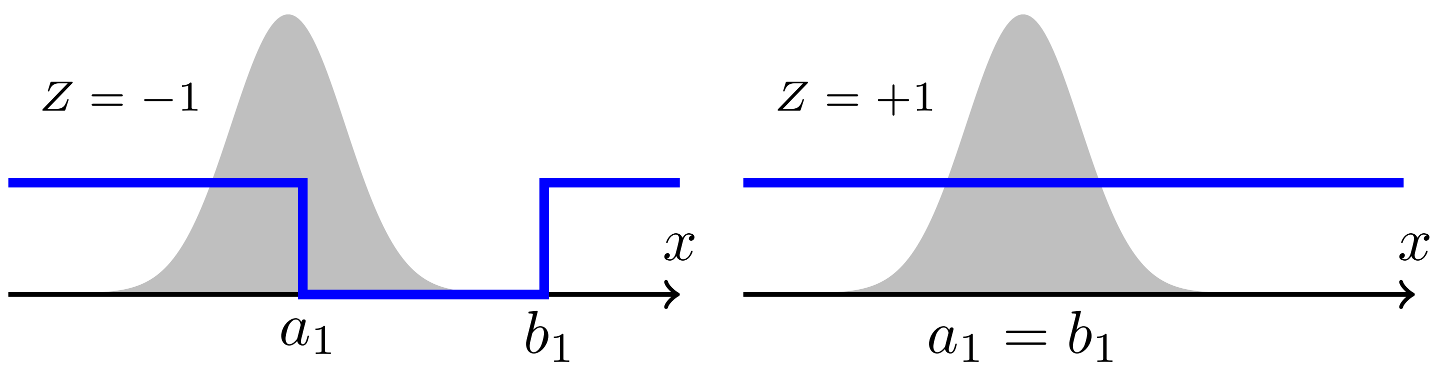


Theorem:

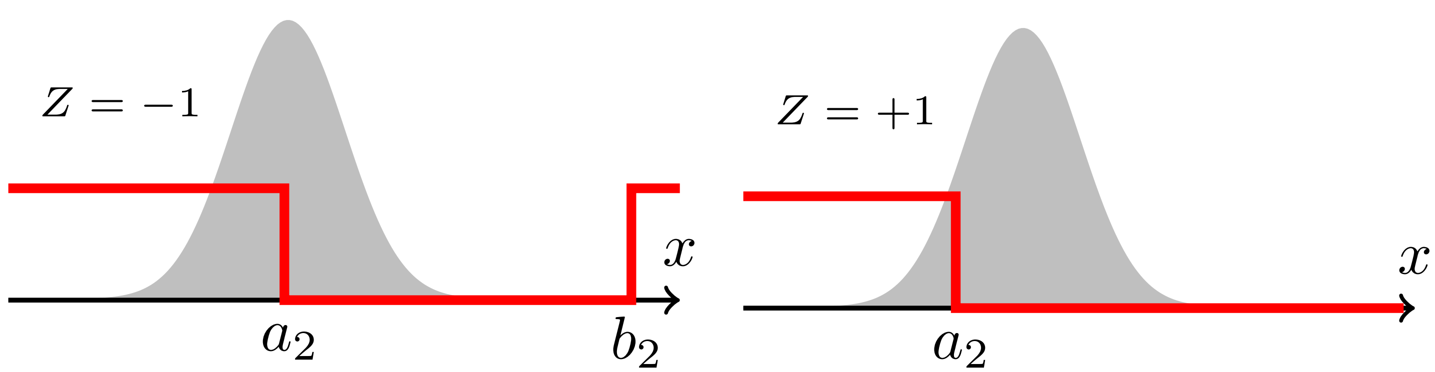
There exists a pair of **threshold policies on private information** that attains the optimal cost in Problem 3.

Scheduling vs. event-based policies

Example $X_1, X_2 \sim \mathcal{N}(0,1)$ $Z = \begin{cases} +1 & \text{w.p. } p \\ -1 & \text{w.p. } 1-p \end{cases}$



(a) Communication policy $\mathcal{U}_1$



(b) Communication policy $\mathcal{U}_2$

p	$\mathcal{J}^\star$
0	0.54
0.1	0.59
0.2	0.63
0.3	0.68
0.4	0.73
0.5	0.78

Combination of **scheduling** and **event-based** policies.

$p = 0.5 \implies \mathcal{J}(\mathcal{U}_1^\star, \mathcal{U}_2^\star) = 0.78$

Gain of 22% over scheduling policies

Vasconcelos & Martins, “The structure of optimal communication policies for remote estimation over the collision channel with private and common observations. Submitted to CDC 2016.

Summary of our talk

1. Structural results of optimal policies for MMSE estimation
2. Numerical method to compute pbp optimal thresholds
3. Extension to dependent observations

Future research directions

1. **Develop a "Kalman" filter** for remote estimation over the collision channel
2. Extend these results for **arbitrarily correlated measurements**
3. **Scalability results**

Recently completed

Vasconcelos & Martins, "Estimation of **discrete random variables** over the collision channel with **minimum probability of error**". *To be submitted to IEEE TAC.*