Optimal estimation of discrete random variables over the collision channel

Marcos Vasconcelos

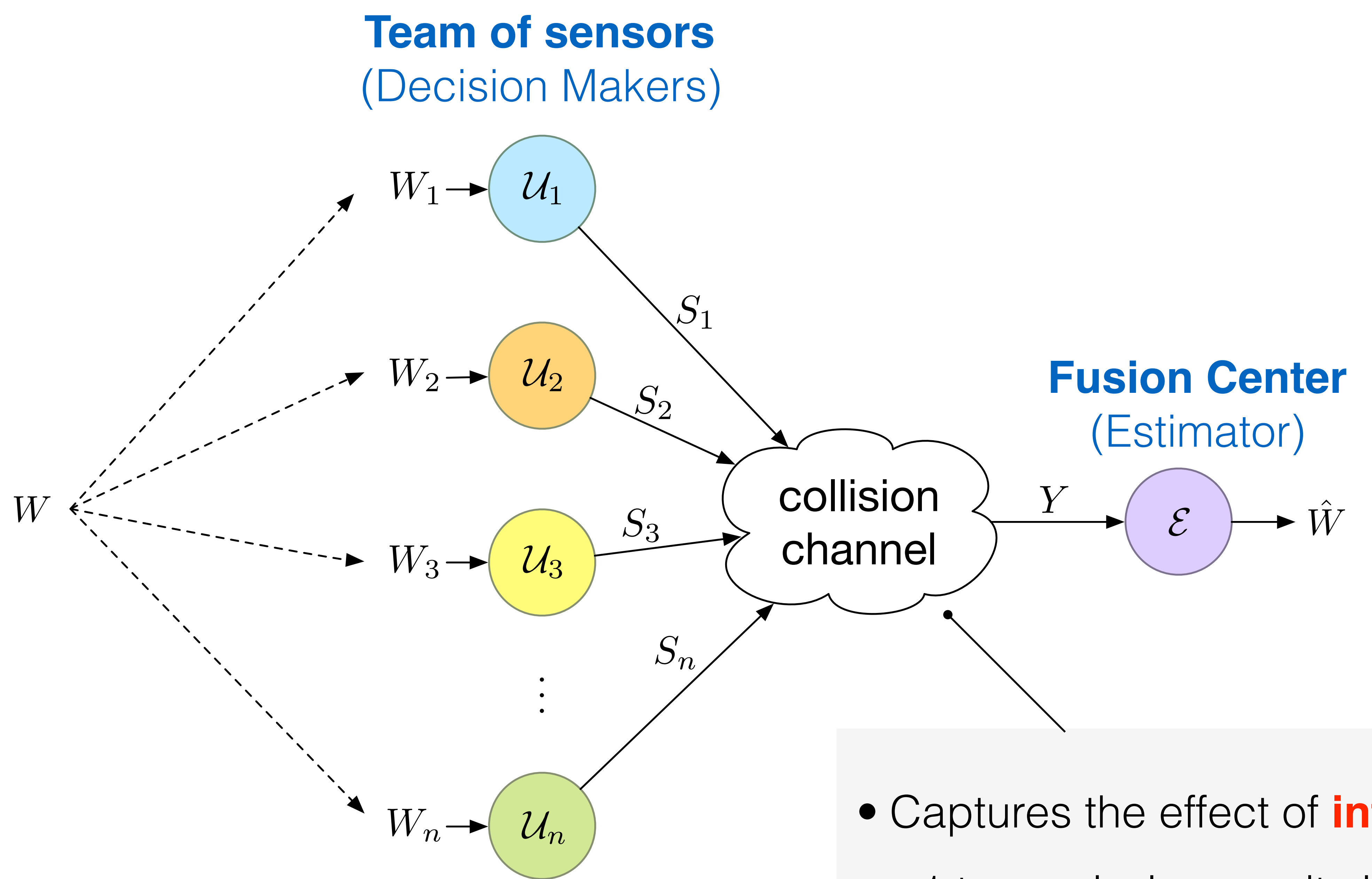
marcos@terpconnect.umd.edu

Nuno C. Martins

Dept. of Electrical and Computer Engineering
University of Maryland, College Park

52th Annual Conference on Information Sciences and Systems
21-23 March 2018 - Princeton, NJ

Basic framework

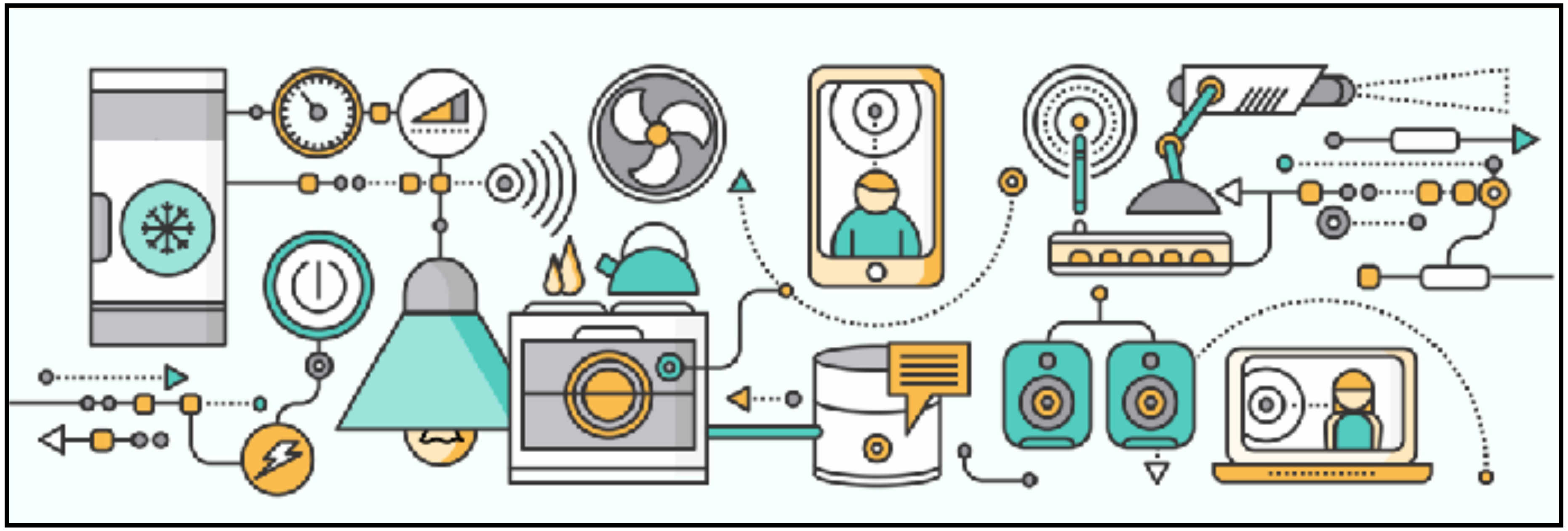


- Captures the effect of **interference**
- >1 transmission results in a **collision**

Design jointly optimal communication and estimation policies

Potential application

Internet-of-things



Real-time wireless networking

ARQ schemes require feedback & introduce undesirable delays

Explicitly deal with collision events

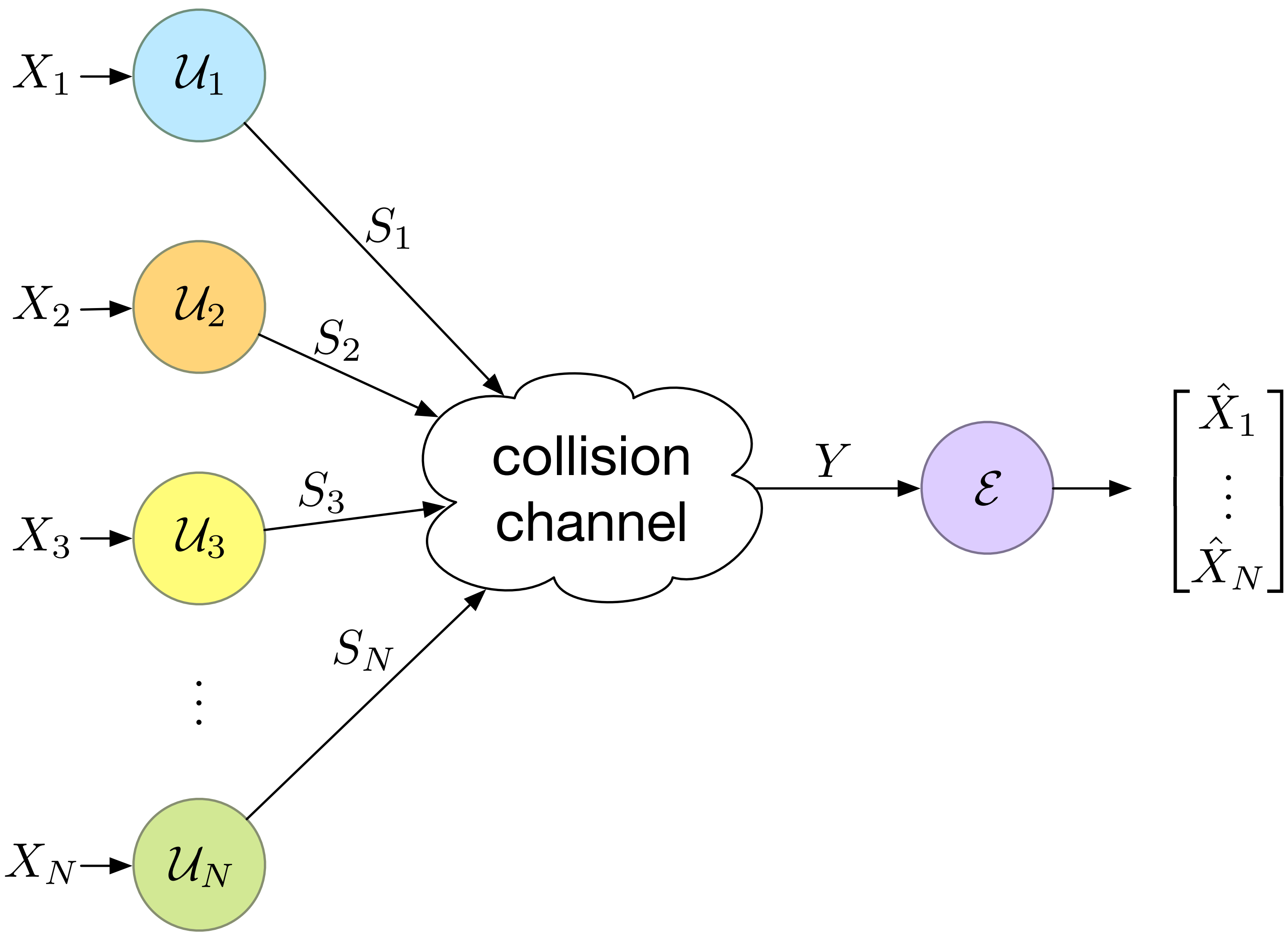
Estimation over the collision channel

Observations

$$X_i \sim p_i$$

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

$$|\mathbb{X}_i| < \infty$$



Decision variables

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

Stay silent

$$S_i = \varnothing$$

Transmit

$$S_i = (i, X_i)$$

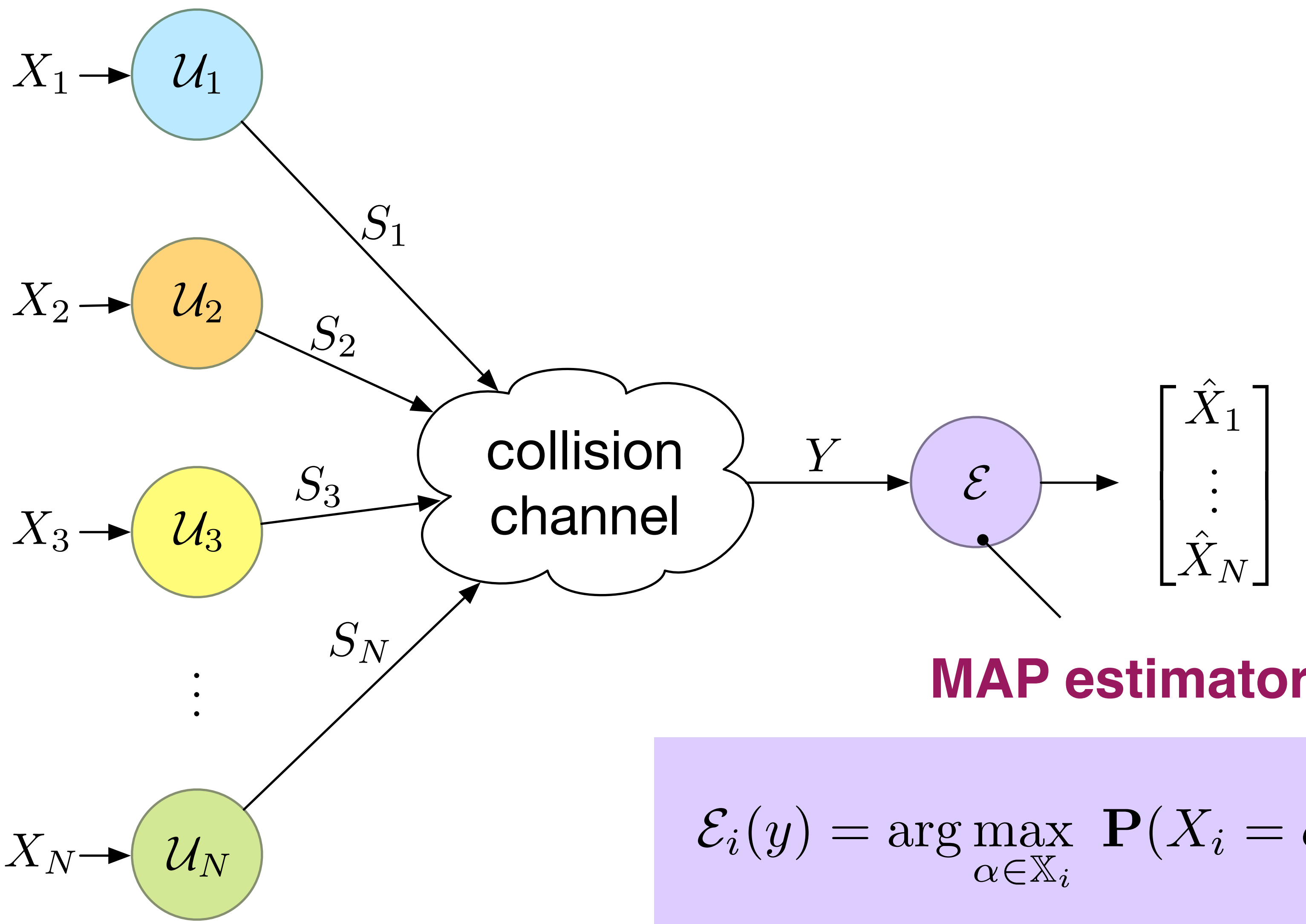
Communication policies

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

Estimation policies

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

MAP estimation over the collision channel



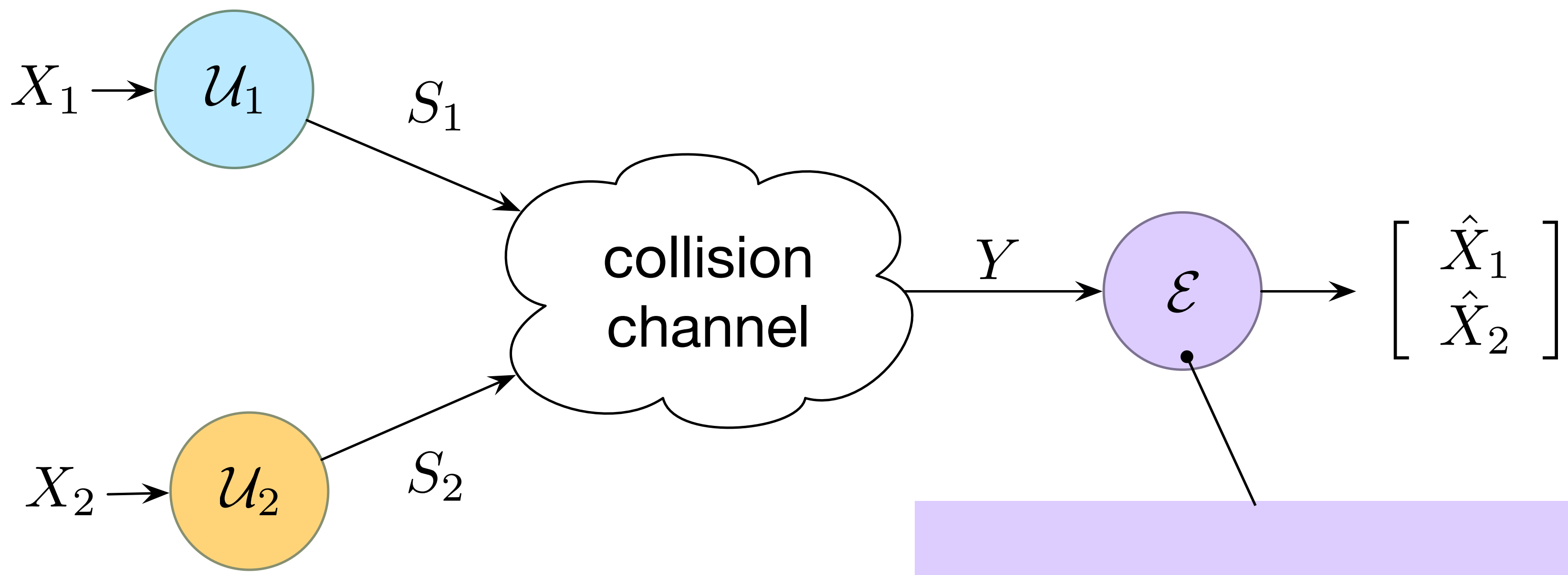
$$\mathcal{E}_i(y) = \arg\max_{\alpha \in \mathbb{X}_i} \; \mathbf{P}(X_i = \alpha \mid Y = y)$$

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

Aggregate prob. of error

$$\mathcal{J}(\mathcal{U}_1, \cdots, \mathcal{U}_N) = \sum_{k=1}^{N} \eta_k \mathbf{P}(X_k \neq \hat{X}_k)$$

Simplest problem: two sensors



$$\mathcal{E}_i(y) = \arg\max_{\alpha \in \mathbb{X}_i} \mathbf{P}(X_i = \alpha \mid Y = y)$$

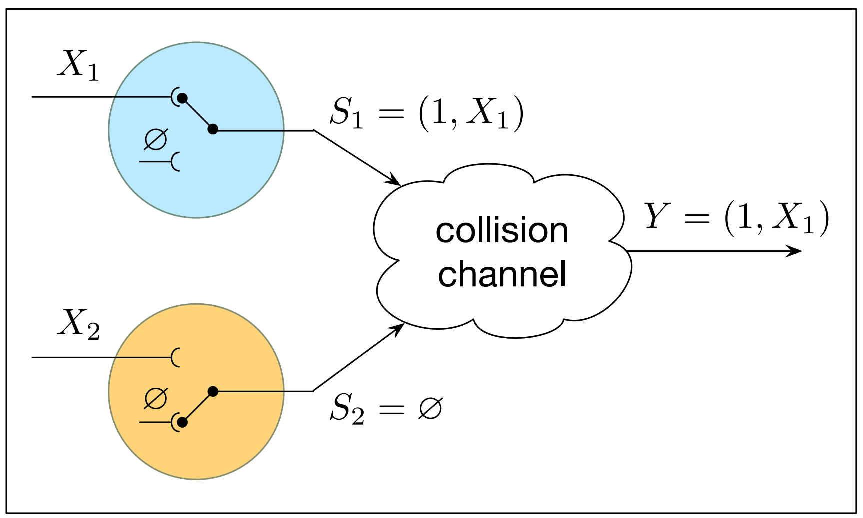
$$\mathbf{P}(U_i = 1 \mid X_i = x_i) = \mathcal{U}_i(x_i)$$

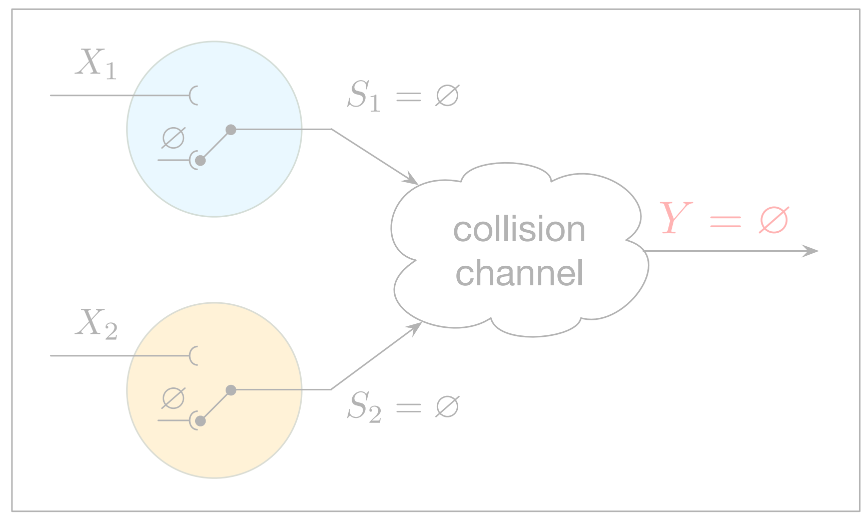
$$\mathbb{U}_i = \{\mathcal{U}_i \mid \mathbb{X}_i \to [0,1]\}$$

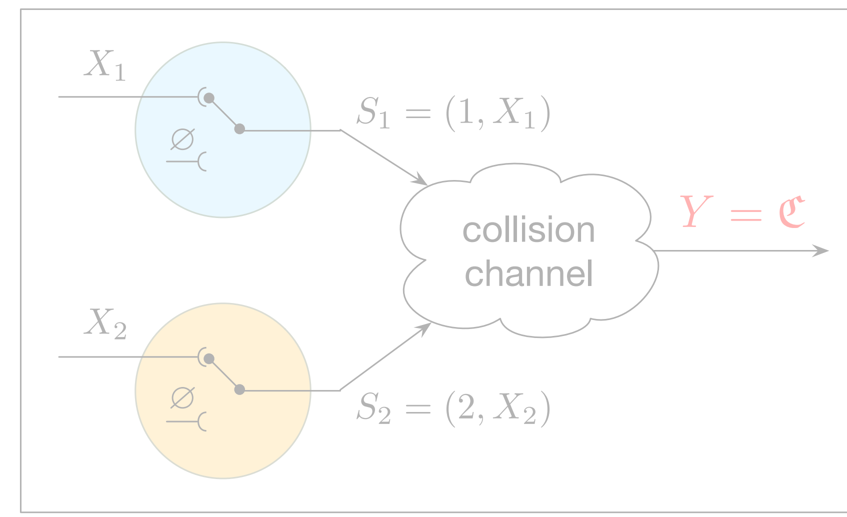
$$\min_{(\mathcal{U}_1,\mathcal{U}_2)\in\mathbb{U}_1\times\mathbb{U}_2} \quad \mathcal{J}(\mathcal{U}_1,\mathcal{U}_2) = \eta_1\mathbf{P}(X_1 \neq \hat{X}_1) + \eta_2\mathbf{P}(X_2 \neq \hat{X}_2)$$

Collision channel

single transmission	no transmissions	>1 transmissions
$U_1 = 1, U_2 = 0$	$U_1 = 0, U_2 = 0$	$U_1 = 1, U_2 = 1$







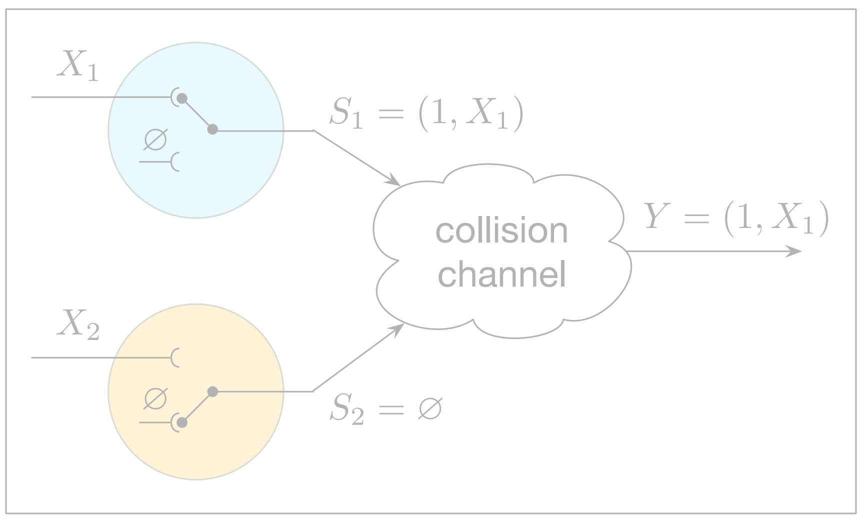
success! | no transmission $\varnothing$ | collision $\mathfrak{C}$

From the channel output we can always recover U_1 and U_2

Collision channel

single transmission

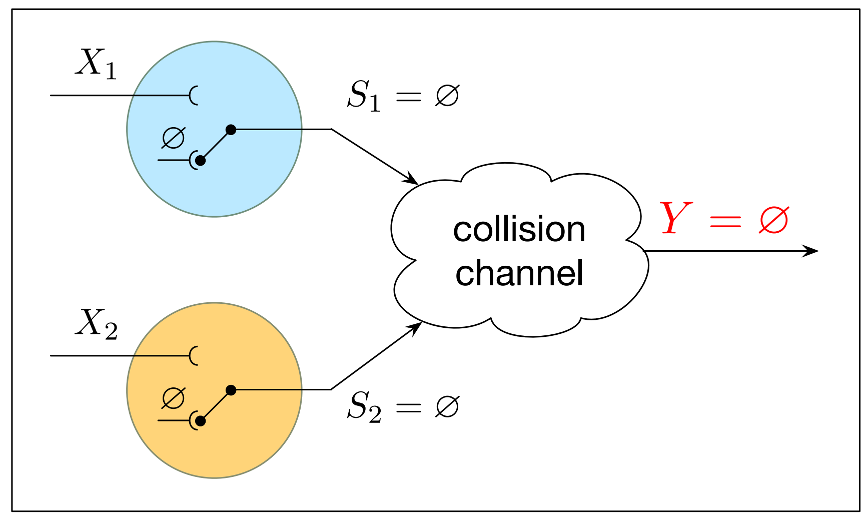
$U_1 = 1, U_2 = 0$



success!

no transmissions

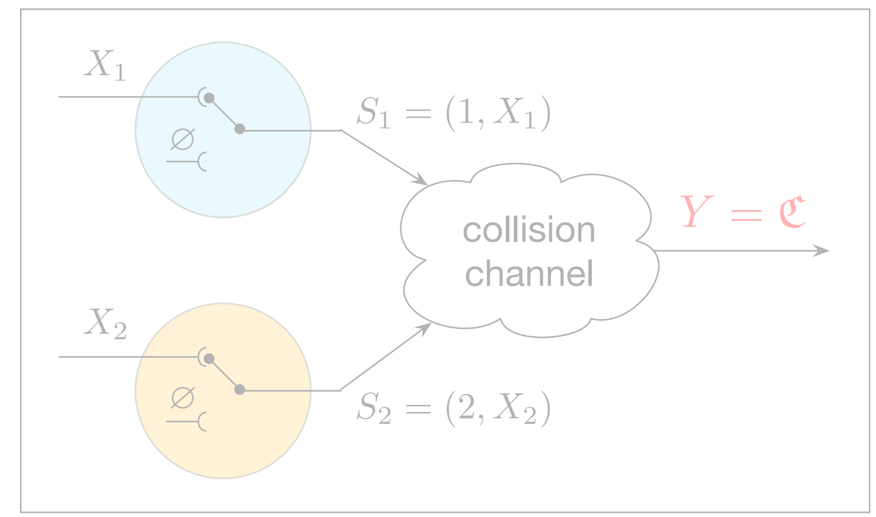
$U_1 = 0, U_2 = 0$



no transmission $\varnothing$

>1 transmissions

$U_1 = 1, U_2 = 1$

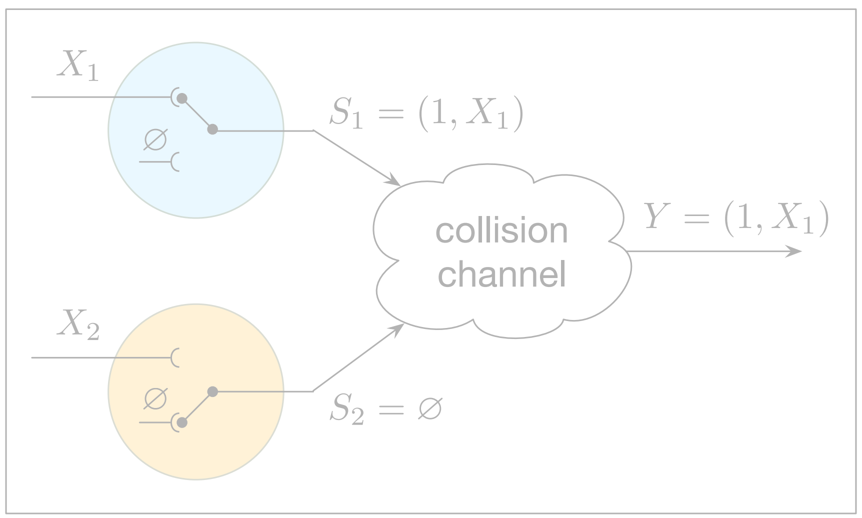


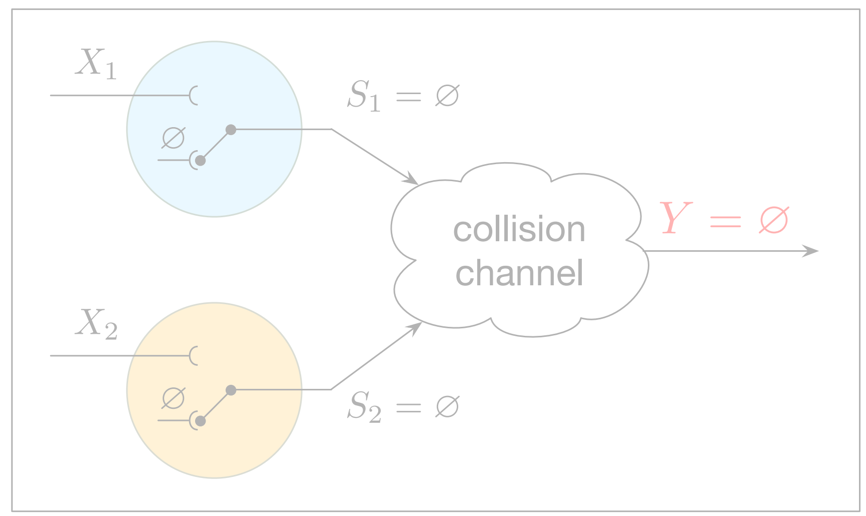
collision $\mathfrak{C}$

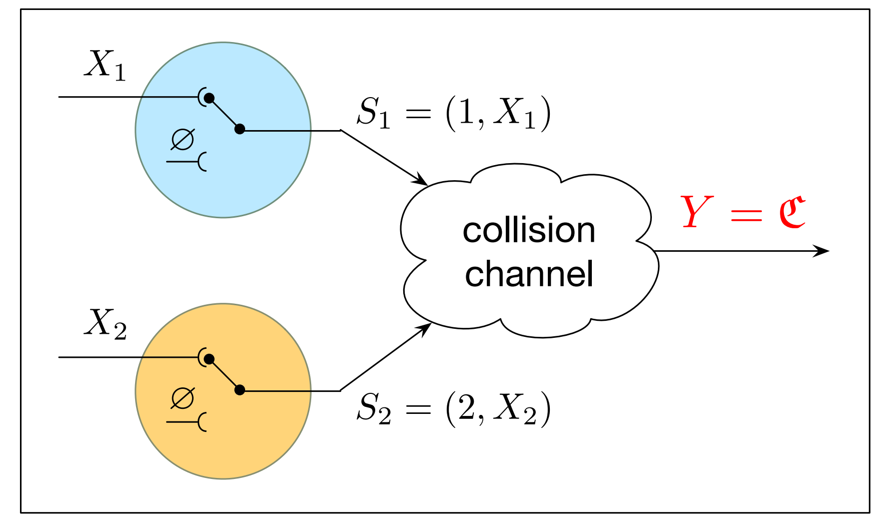
From the channel output we can always recover U_1 and U_2

Collision channel

single transmission	no transmissions	>1 transmissions
$U_1 = 1, U_2 = 0$	$U_1 = 0, U_2 = 0$	$U_1 = 1, U_2 = 1$







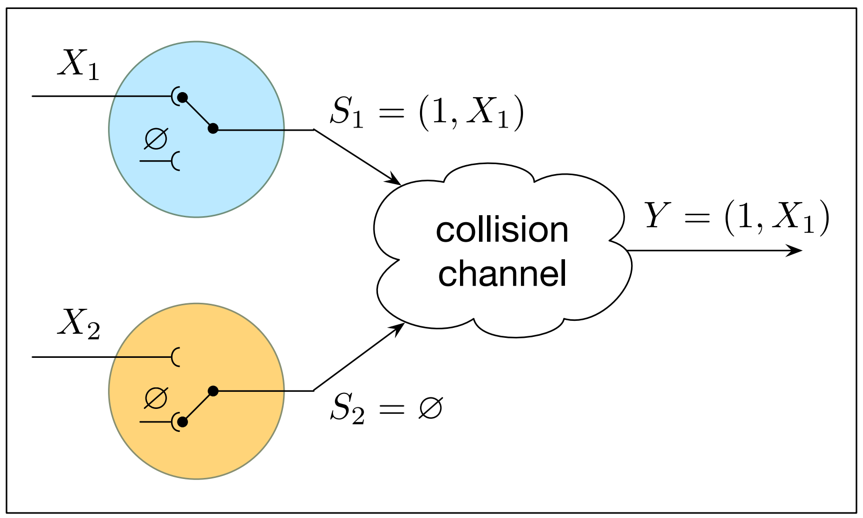
success!	no transmission $\varnothing$	collision $\mathfrak{C}$

From the channel output we can always recover U_1 and U_2

Collision channel

single transmission

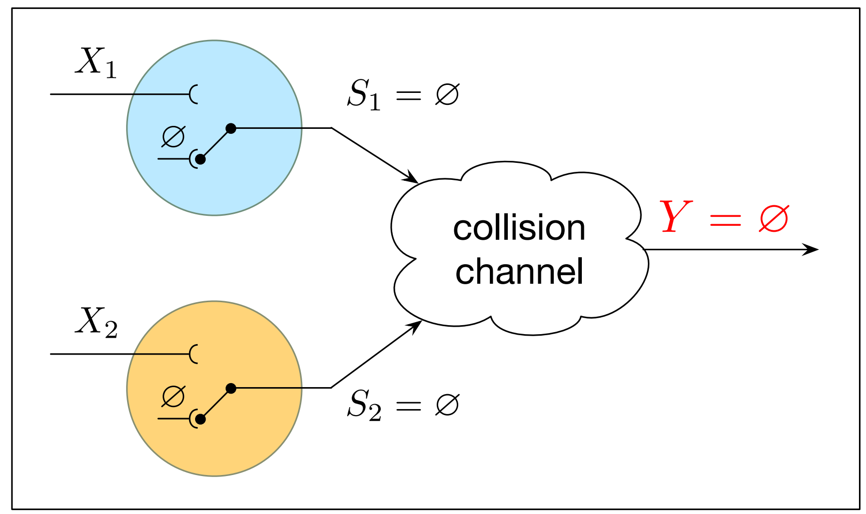
$U_1 = 1, U_2 = 0$



success!

no transmissions

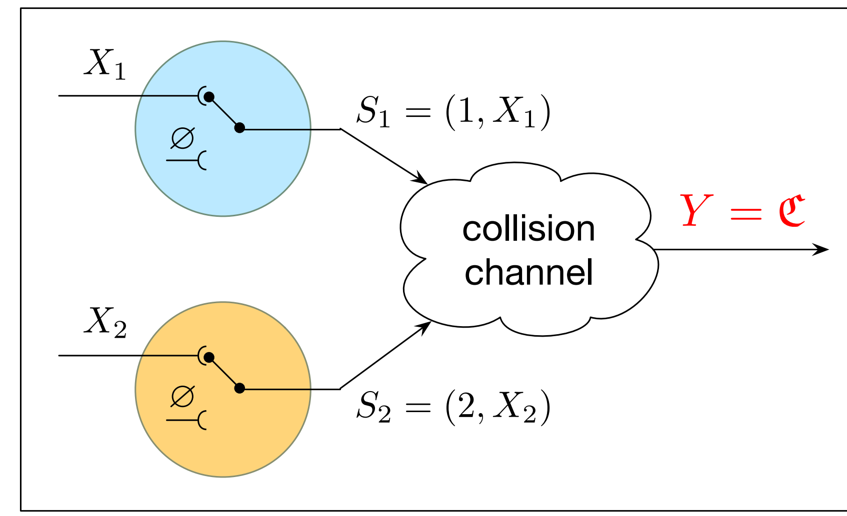
$U_1 = 0, U_2 = 0$



no transmission $\varnothing$

>1 transmissions

$U_1 = 1, U_2 = 1$

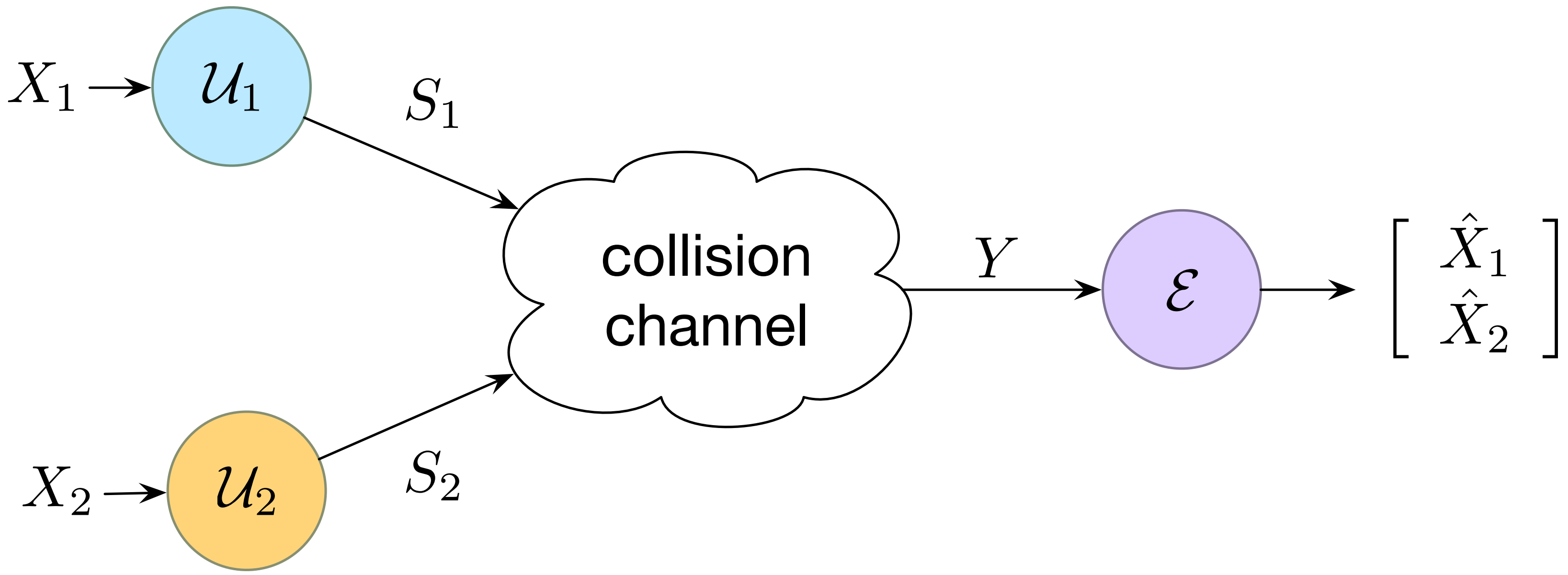


collision $\mathfrak{C}$

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?

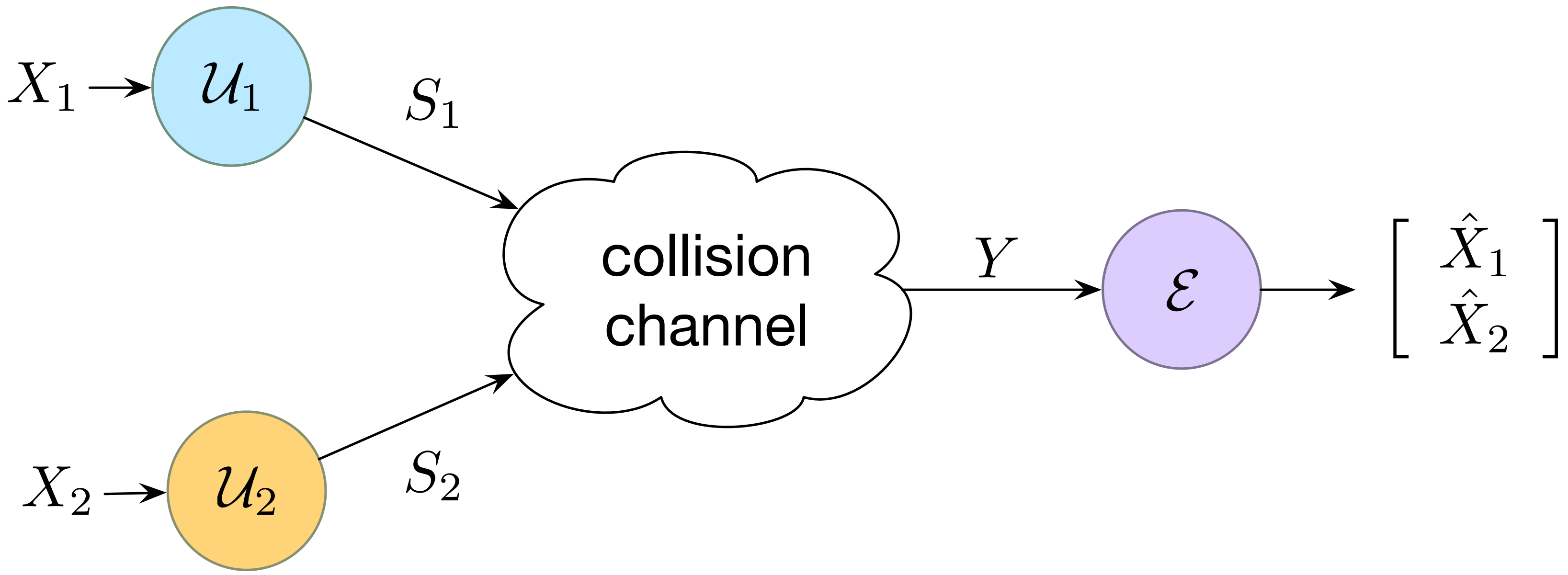


$$\min_{(\mathcal{U}_1,\mathcal{U}_2)\in\mathbb{U}_1\times\mathbb{U}_2} \quad \mathcal{J}(\mathcal{U}_1,\mathcal{U}_2) = \eta_1\mathbf{P}(X_1 \neq \hat{X}_1) + \eta_2\mathbf{P}(X_2 \neq \hat{X}_2)$$

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?



$$\min_{(\mathcal{U}_1,\mathcal{U}_2)\in\mathbb{U}_1\times\mathbb{U}_2} \quad \mathcal{J}(\mathcal{U}_1,\mathcal{U}_2) = \eta_1 \mathbf{P}(X_1 \neq \hat{X}_1) + \eta_2 \mathbf{P}(X_2 \neq \hat{X}_2)$$

NP-hard!

Find a structured subset of policies that contains an optimal solution

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

Main result

Theorem 1

There exists a **globally optimal** solution $(\mathcal{U}_1^\star, \mathcal{U}_2^\star)$ where:

$$\mathcal{U}_i^\star \in \left\{ \mathcal{V}_i^0, \mathcal{V}_i^1, \mathcal{V}_i^2 \right\}$$

Main result

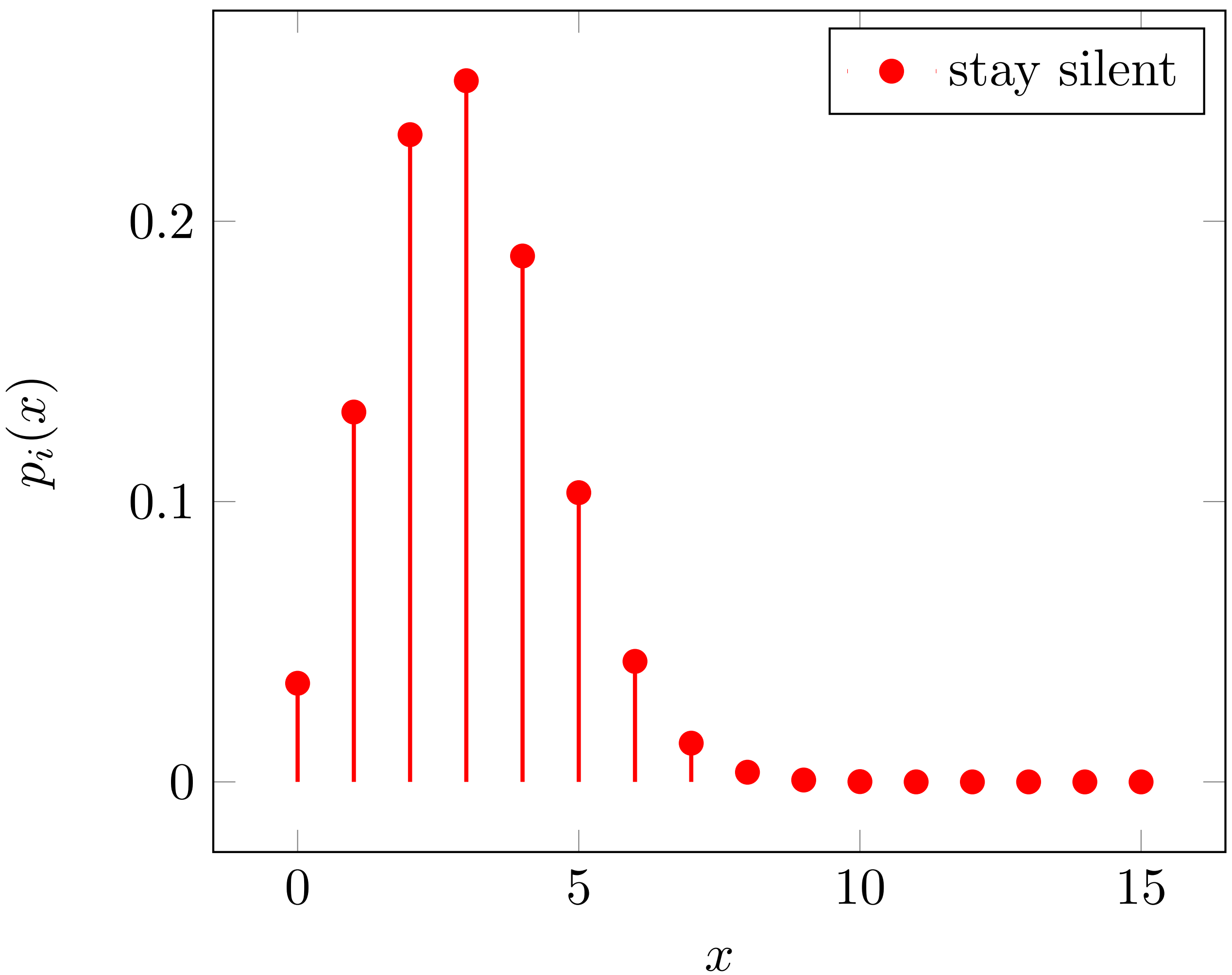
Theorem 1

There exists a **globally optimal** solution $(\mathcal{U}_1^\star, \mathcal{U}_2^\star)$ where:

$$\mathcal{U}_i^\star \in \left\{ \mathcal{V}_i^0, \mathcal{V}_i^1, \mathcal{V}_i^2 \right\}$$

Never transmit

Policy $\mathcal{V}_i^0(x)$



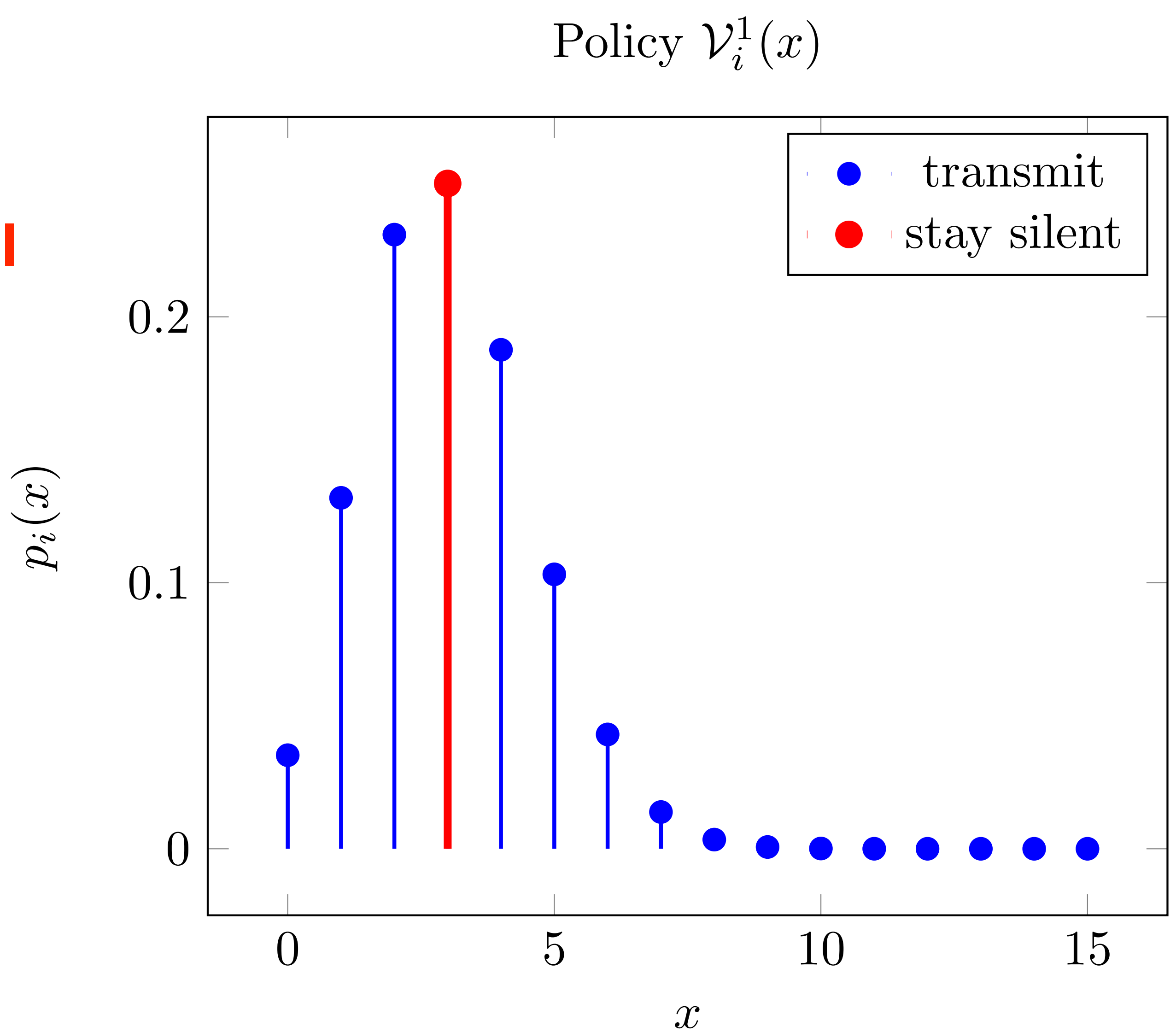
Main result

Theorem 1

There exists a **globally optimal** solution $(\mathcal{U}_1^\star, \mathcal{U}_2^\star)$ where:

$$\mathcal{U}_i^\star \in \left\{\mathcal{V}_i^0, \mathcal{V}_i^1, \mathcal{V}_i^2\right\}$$

Transmit all, except the most likely symbol



Main result

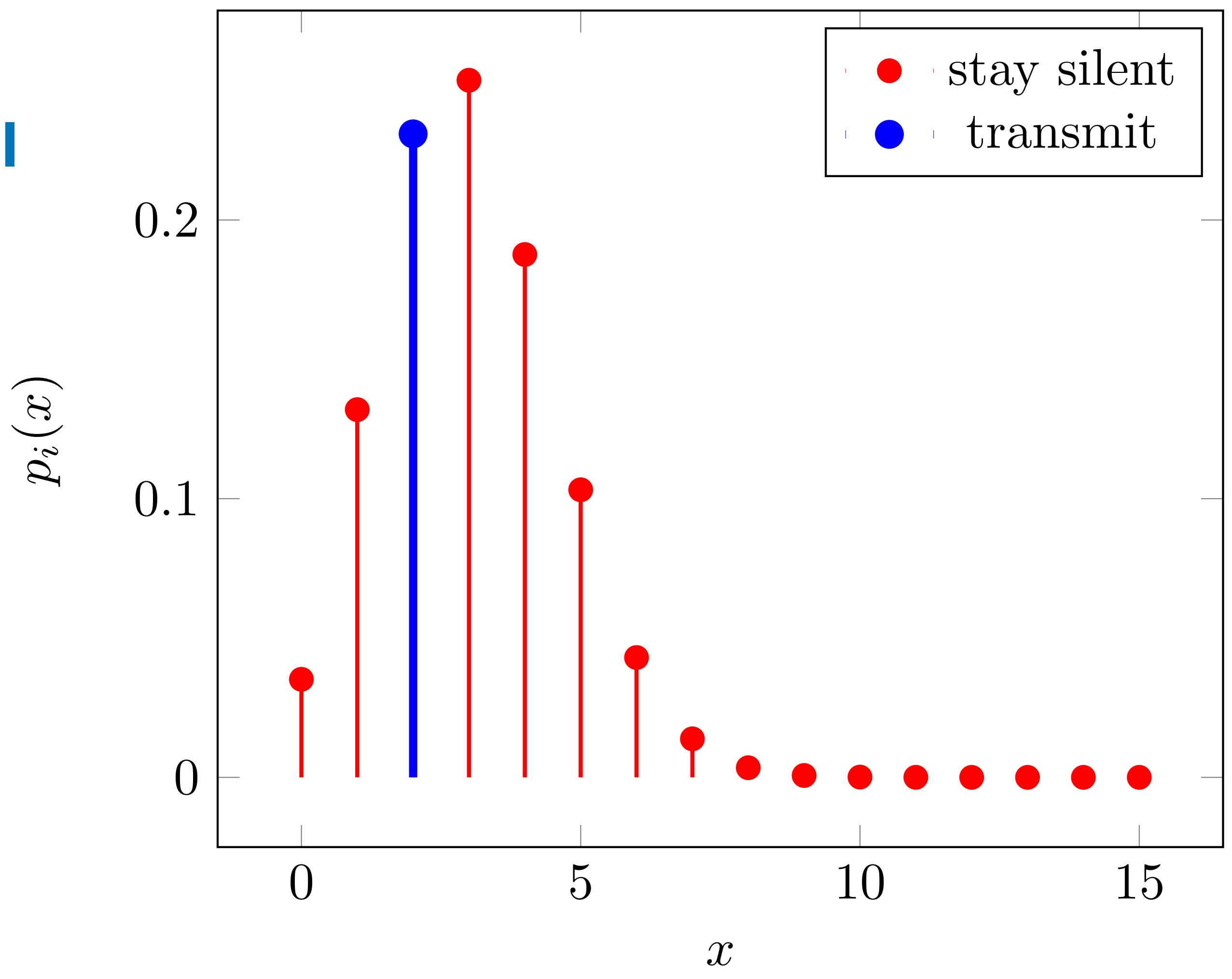
Theorem 1

There exists a **globally optimal** solution $(\mathcal{U}_1^\star, \mathcal{U}_2^\star)$ where:

$$\mathcal{U}_i^\star \in \left\{ \mathcal{V}_i^0, \mathcal{V}_i^1, \mathcal{V}_i^2 \right\}$$

Transmit only the second most likely symbol

Policy $\mathcal{V}_i^2(x)$



Main result

Theorem 1

There exists a **globally optimal** solution $(\mathcal{U}_1^\star, \mathcal{U}_2^\star)$ where:

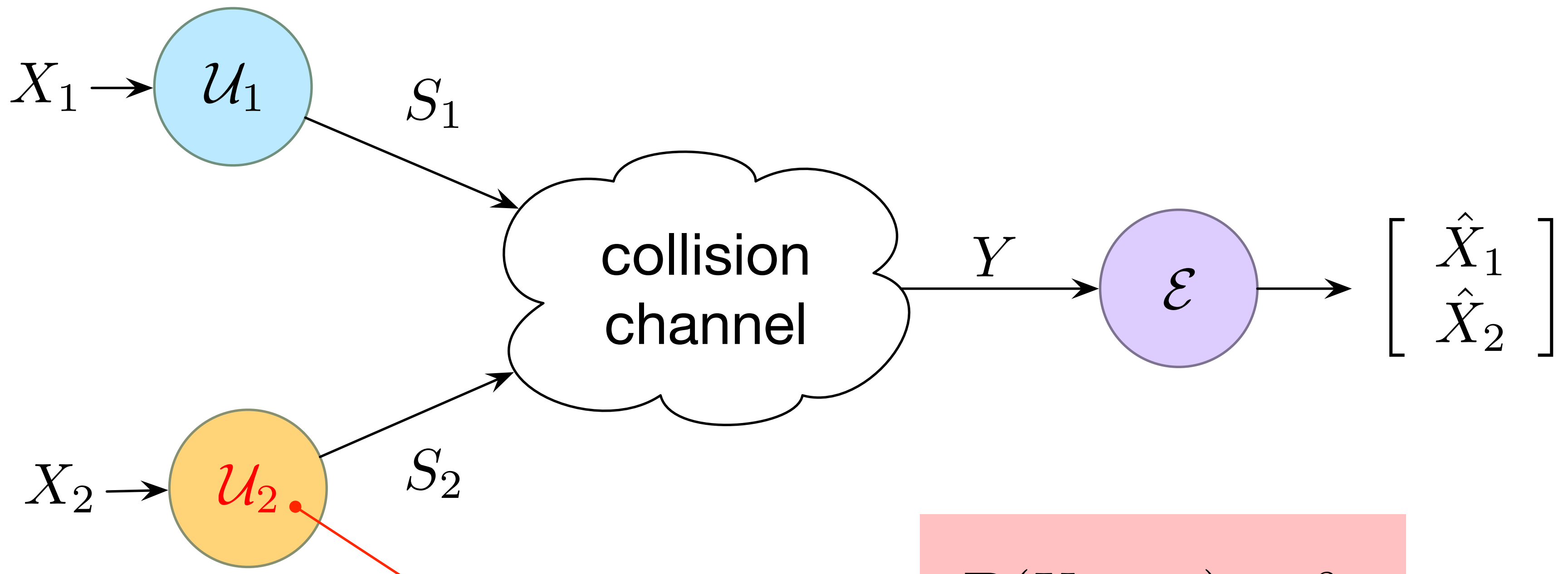
$$\mathcal{U}_i^\star \in \left\{\mathcal{V}_i^0, \mathcal{V}_i^1, \mathcal{V}_i^2\right\}$$

This structural result is independent of the pmfs!

complexity reduction

$$2^{|\mathbb{X}_1|+|\mathbb{X}_2|} \longrightarrow 3^2$$

Proof sketch



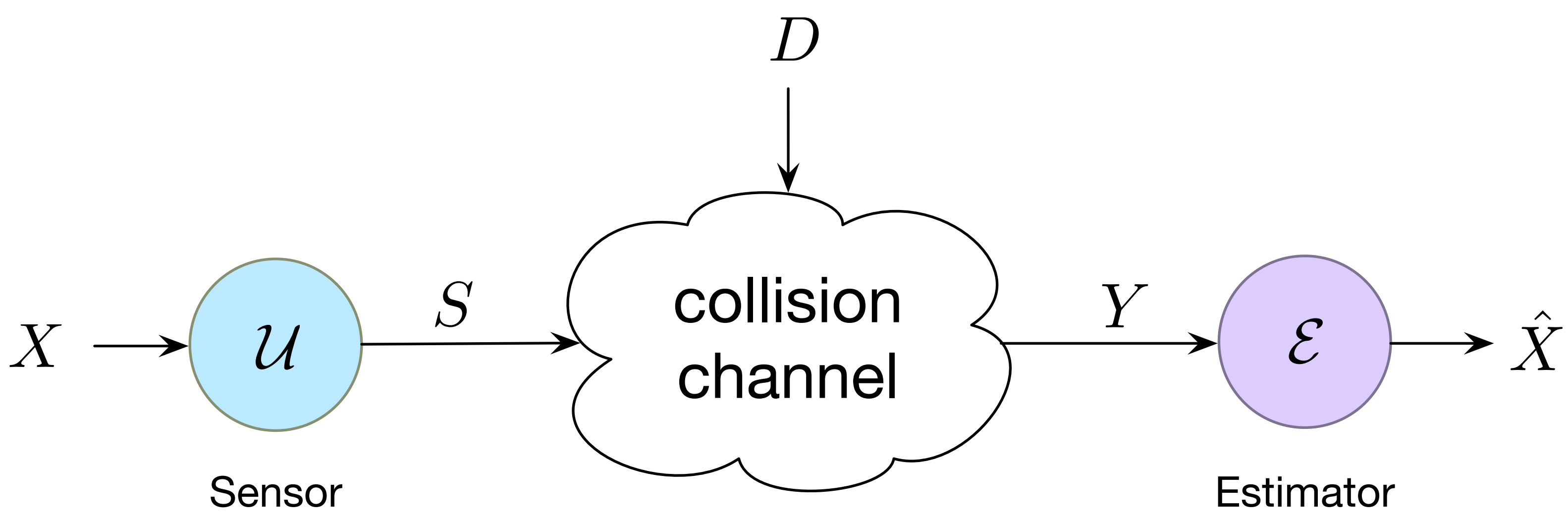
$\Rightarrow$ $\mathbf{P}(U_2 = 1) = \beta_2$ **is constant**

$$\mathcal{J}(\mathcal{U}_1, \mathcal{U}_2) = \eta_1 \mathbf{P}(X_1 \neq \hat{X}_1) + \eta_2(\rho_{\mathcal{U}_2}\mathbf{P}(U_1 = 1) + \theta_{\tilde{\mathcal{U}}_2})$$

Cost from the perspective of DM_1:

$$\mathcal{J}(\mathcal{U}_1, \mathcal{U}_2) \propto \mathbf{P}(X_1 \neq \hat{X}_1) + \frac{\eta_2}{\eta_1}\rho_{\mathcal{U}_2}\mathbf{P}(U_1 = 1)$$

Proof sketch



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

Determines if the channel is occupied or not

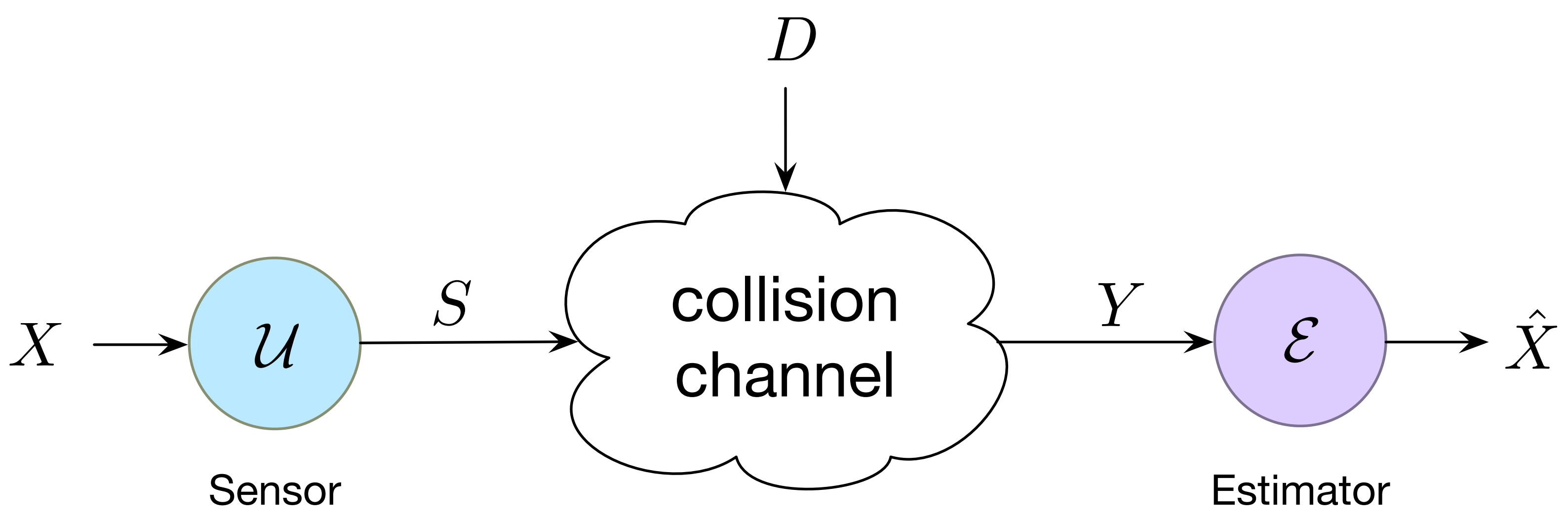
$X \perp\!\!\!\perp D$

$$\mathcal{J}(\mathcal{U}) = \mathbf{P}(X \neq \hat{X}) + \varrho \mathbf{P}(U = 1)$$

$$\mathcal{E}(y) = \arg \max_{\alpha \in \mathbb{X}} \mathbf{P}(X = \alpha \mid Y = y)$$

1. Imer & Basar, "Optimal estimation with limited measurements," *IJSCC* 2010

Proof sketch



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

Determines if the channel is occupied or not

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

$$\mathcal{J}(\mathcal{U}) = \mathbf{P}(X \neq \hat{X}) + \varrho \mathbf{P}(U = 1)$$

$$\mathcal{J}(\mathcal{U}) = \ 1 + (\varrho + \beta - 1) \sum_{\alpha \in \mathbb{X}} \mathcal{U}(\alpha) p(\alpha) - \max_{\alpha \in \mathbb{X}} \Big\{ \big(1 - \mathcal{U}(\alpha)\big) p(\alpha) \Big\} - \beta \max_{\alpha \in \mathbb{X}} \Big\{ \mathcal{U}(\alpha) p(\alpha) \Big\}$$

Proof sketch

For $\beta \in [0,1]$ and $\varrho \geq 0$, solve:

$$\begin{array}{ll} \underset{\mathcal{U}}{\text{minimize}} & \mathcal{J}(\mathcal{U}) \\ \text{subject to} & 0 \leq \mathcal{U}(x) \leq 1, \quad x \in \mathbb{X} \end{array}$$

$$\mathcal{J}(\mathcal{U}) = 1 + (\varrho + \beta - 1) \sum_{\alpha \in \mathbb{X}} \mathcal{U}(\alpha) p(\alpha) - \max_{\alpha \in \mathbb{X}} \Big\{ \big(1 - \mathcal{U}(\alpha)\big) p(\alpha) \Big\} - \beta \max_{\alpha \in \mathbb{X}} \Big\{ \mathcal{U}(\alpha) p(\alpha) \Big\}$$

concave function

1. Constrain to **deterministic policies**

$\mathbb{X} \to \mathbb{X}_0^{\mathcal{U}}$ — **no transmission set**

$\mathbb{X} \to \mathbb{X}_1^{\mathcal{U}}$ — **transmission set**

Sketch of Proof

The cost becomes:

$$\mathcal{J}(\mathcal{U}) = 1 + (\varrho + \beta - 1) \sum_{\alpha \in \mathbb{X}_1^{\mathcal{U}}} p(\alpha) - \max_{\alpha \in \mathbb{X}_0^{\mathcal{U}}} p(\alpha) - \beta \max_{\alpha \in \mathbb{X}_1^{\mathcal{U}}} p(\alpha)$$

Lemma

$$\mathcal{U}^{\star}_{\beta,\varrho} = \begin{cases} \mathcal{V}^1 & \text{if } \; 0 \leq \varrho \leq 1 - \beta \\ \mathcal{V}^2 & \text{if } \; 1 - \beta < \varrho \leq 1 \\ \mathcal{V}^0 & \text{otherwise} \end{cases}$$

Step 1: "Converse result" - Lower bound satisfied by every policy

Step 2: "Achievability" - The policy above meets the lower bound with equality

1. Vasconcelos and Martins, "Estimation over the collision channel: the discrete case," Submitted to *IEEE TAC*

$0 \leq \varrho \leq 1 - \beta$

$$\mathcal{J}(\mathcal{U}) = \; 1 + (\varrho + \beta - 1) \sum_{\alpha \in \mathbb{X}_1^{\mathcal{U}}} p(\alpha) - \max_{\alpha \in \mathbb{X}_0^{\mathcal{U}}} p(\alpha) - \beta \max_{\alpha \in \mathbb{X}_1^{\mathcal{U}}} p(\alpha)$$

$$\sum_{\alpha \in \mathbb{X}_1^{\mathcal{U}}} p(\alpha) \leq 1 - \max_{\alpha \in \mathbb{X}_0^{\mathcal{U}}} p(\alpha)$$

$$\mathcal{J}(\mathcal{U}) \geq (\varrho + \beta) - (\varrho + \beta) \max_{\alpha \in \mathbb{X}_0^{\mathcal{U}}} p(\alpha) - \beta \max_{\alpha \in \mathbb{X}_1^{\mathcal{U}}} p(\alpha)$$

$$\mathcal{J}(\mathcal{U}) \geq (\varrho + \beta) - (\varrho + \beta) \max_{\alpha \in \mathbb{X}_0^{\mathcal{U}}} p(\alpha) - \beta \max_{\alpha \in \mathbb{X}_1^{\mathcal{U}}} p(\alpha)$$

$$x_{[1]} \longrightarrow \mathbb{X}_0^{\mathcal{U}}$$

$$x_{[2]} \longrightarrow \mathbb{X}_1^{\mathcal{U}}$$

“Converse”

$$\mathcal{J}(\mathcal{U}) \geq (\varrho + \beta) - (\varrho + \beta) p_{[1]} - \beta p_{[2]}$$

“Achievability”

$$\mathcal{J}(\mathcal{V}^1) = (\varrho + \beta) - (\varrho + \beta) p_{[1]} - \beta p_{[2]}$$

$\blacksquare$

Extension

Theorem 2

There exists a globally optimal solution $(\mathcal{U}_1^\star, \cdots, \mathcal{U}_N^\star)$ where:

$$\mathcal{U}_i^\star \in \left\{ \mathcal{V}_i^0, \mathcal{V}_i^1, \mathcal{V}_i^2 \right\}$$

Additional assumption[1]

When a collision occurs,
the receiver decodes the indices of the transmitting nodes

$Y = \mathfrak{C}$ $\longrightarrow$ $U = (U_1, \cdots, U_N)$ is available to the estimator

1. Whitehouse et al. (2005), "Exploiting the capture effect for collision detection and recovery"

Problem

Theorem 2

There exists a globally optimal solution $(\mathcal{U}_1^\star, \cdots, \mathcal{U}_N^\star)$ where:

$$\mathcal{U}_i^\star \in \left\{\mathcal{V}_i^0, \mathcal{V}_i^1, \mathcal{V}_i^2\right\}$$

independent of the support of the pmfs

complexity reduction

$$2^{|\mathbb{X}_1|+\cdots+|\mathbb{X}_N|} \longrightarrow 3^N$$

exponential in *N*

Identically distributed case

$$X_i \sim p(x)$$

$$\eta_1 \geq \eta_2 \geq \cdots \geq \eta_N$$

Theorem 3

There exists a globally optimal solution $\mathcal{U}^\star$ where:

$$\mathcal{U}^\star = (\underbrace{\mathcal{V}^1, \cdots, \mathcal{V}^1}_{n_1^\star \text{ tuple}}, \underbrace{\mathcal{V}^2, \cdots, \mathcal{V}^2}_{n_2^\star \text{ tuple}}, \underbrace{\mathcal{V}^0, \cdots, \mathcal{V}^0}_{(N - n_1^\star - n_2^\star) \text{ tuple}})$$

$$n_1^\star, n_2^\star \geq 0$$

$$n_1^\star + n_2^\star \leq N$$

$$\frac{1}{2}(N+1)(N+2)$$

complexity is quadratic!

Identical weights

$$X_i \sim p(x)$$

$$\eta_1 = \eta_2 = \cdots = \eta_N$$

Theorem 4

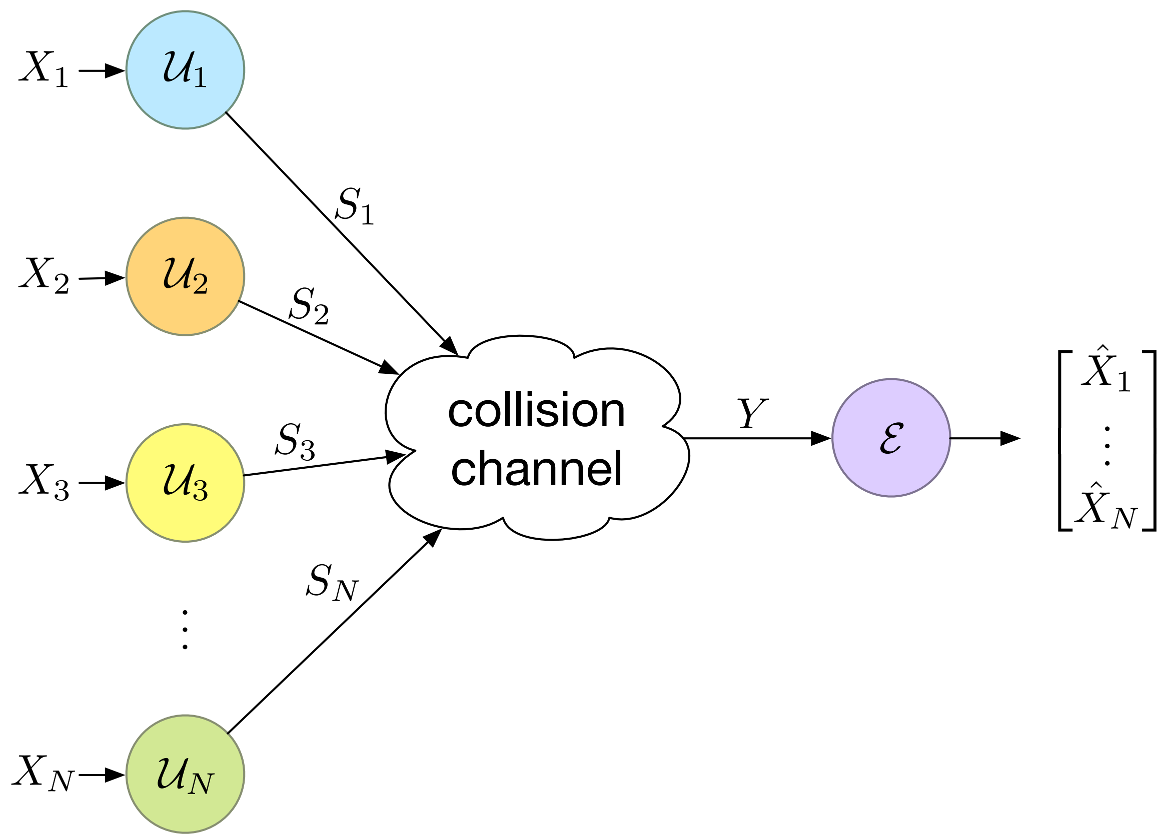
There exists a globally optimal solution $\mathcal{U}^\star$ where:

$$\mathcal{U}^\star = (\underbrace{\mathcal{V}^1, \cdots, \mathcal{V}^1}_{n^\star \text{ tuple}}, \underbrace{\mathcal{V}^2, \cdots, \mathcal{V}^2}_{(N-n^\star) \text{ tuple}})$$

$$n^\star = \min\left\{ \left\lfloor \frac{p_{[1]}}{1 - p_{[1]} - p_{[2]}} \right\rfloor + 1, N \right\}$$

exact solution!

Final remarks



Structural results

1. Two sensor case
2. N sensor case (**3^N**)
3. Identically distributed (**N^2**)
4. Identically dist. with identical weights (**closed form solution!**)

Vasconcelos and Martins, "Estimation over the collision channel: the discrete case," Submitted to *IEEE TAC*

Future work

1. Correlated observations

2. Lift the "id. of colliding nodes" assumption

$$Y = \mathfrak{C} \not\Leftrightarrow (U_1, \cdots, U_N) \text{ available at the estimator}$$

3. The sequential case