Persuasion with limited communication capacity

Maël Le Treust

ETIS UMR 8051, CY Cergy Paris Université, ENSEA, CNRS

joint work with **Tristan Tomala** (HEC Paris, GREGHEC UMR 2959)

”ReLaX” Workshop on Games - Chennai Mathematical Institute
3rd February 2021

Overview

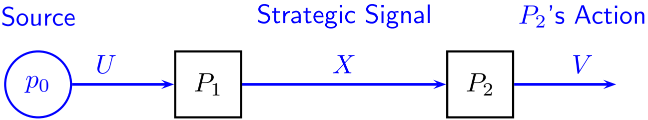
- Autonomous Devices with **Non-Aligned** Objectives/Utilities

- Transmission of Strategic Information in **Economics** Literature
  → **Bayesian Persuasion** : [Kamenica Gentzkow (AER) 2011]

- We characterize the solution when the **communication channel is noisy**
  → [Le Treust and Tomala, Journal of Economic Theory, Vol. 184, Nov. 2019]
  https ://arxiv.org/abs/1711.04474

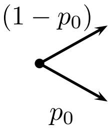
- We extend the solution by considering **decoder side information**
  → [Le Treust and Tomala, draft 2018]
  https ://arxiv.org/abs/1807.05147

Bayesian Persuasion [Kamenica Gentzkow (AER) 2011]



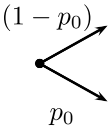
Distinct Utility Functions

$\phi_1(u,v)$



$$\begin{array}{c|c|c|}
 & v_0 & v_1 \\ \hline
u_0 & 0 & 1 \\ \hline
u_1 & 0 & 1 \\ \hline
\end{array}$$

$\phi_2(u,v)$



$$\begin{array}{c|c|c|}
 & v_0 & v_1 \\ \hline
u_0 & 1 & 0 \\ \hline
u_1 & 0 & 1 \\ \hline
\end{array}$$

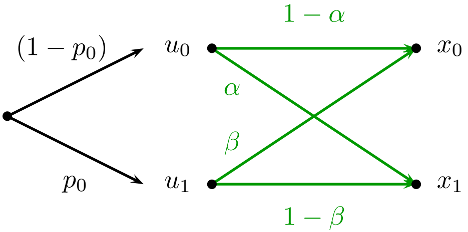
P_2 : **Decision Maker/Decoder**

P_1 : **Persuader/Encoder**

1. P_1 chooses/announces $\mathbb{P}(x|u)$
2. (U,X) is drawn with $\mathbb{P}(u) \times \mathbb{P}(x|u)$
3. P_2 observes X and chooses V
4. Player k's payoff is $\phi_k(u,v)$

Control of the posterior beliefs

P_1's signaling strategy : $\mathbb{P}(x|u)$ with $\alpha \in [0,1]$, $\beta \in [0,1]$



Posterior Distributions (p_1, p_2) :

$$\mathbb{P}(u_1|x_0) = \frac{p_0 \cdot \beta}{p_0 \cdot \beta + (1-p_0) \cdot (1-\alpha)} = p_1$$

$$\mathbb{P}(u_1|x_1) = \frac{p_0 \cdot (1-\beta)}{p_0 \cdot (1-\beta) + (1-p_0) \cdot \alpha} = p_2$$

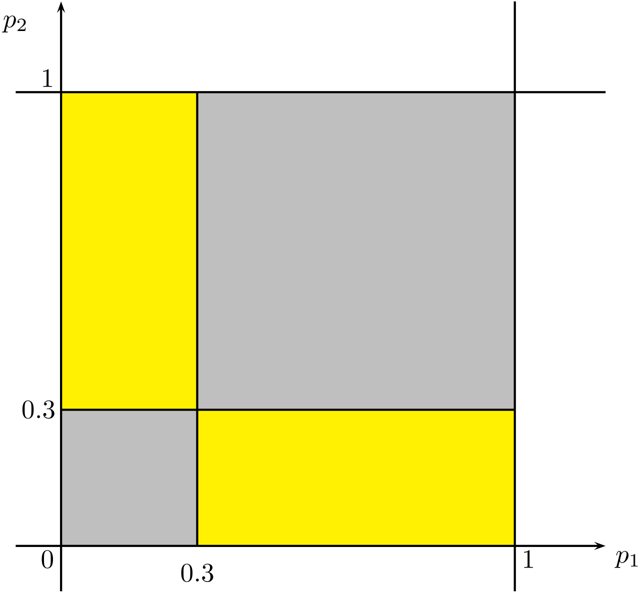
$p_0 = \mathbb{P}(u_1)$: Prior Distribution

Given any prior p_0 and posteriors (p_1, p_2), the signaling parameters (α, β) are :

$$\begin{aligned} \alpha &= \frac{(1-p_2) \cdot (p_0 - p_1)}{(1-p_0) \cdot (p_2 - p_1)} \\ \beta &= \frac{p_1 \cdot (p_2 - p_0)}{p_0 \cdot (p_2 - p_1)} \end{aligned}$$

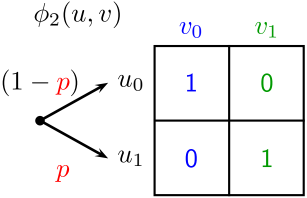
Under condition : $0 \leq p_1 \leq p_0 \leq p_2 \leq 1$ or $0 \leq p_2 \leq p_0 \leq p_1 \leq 1$

Region of Posterior Beliefs (p_1, p_2) for prior $p_0 = 0.3$



Best-Response $BR_2(p)$ of P_2 given the belief p

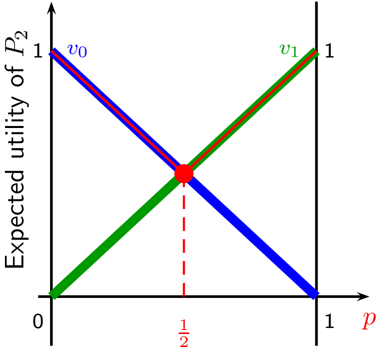
- Denote by $p = \mathbb{P}(u_1)$ the **belief** of P_2



$(1-p)$

p

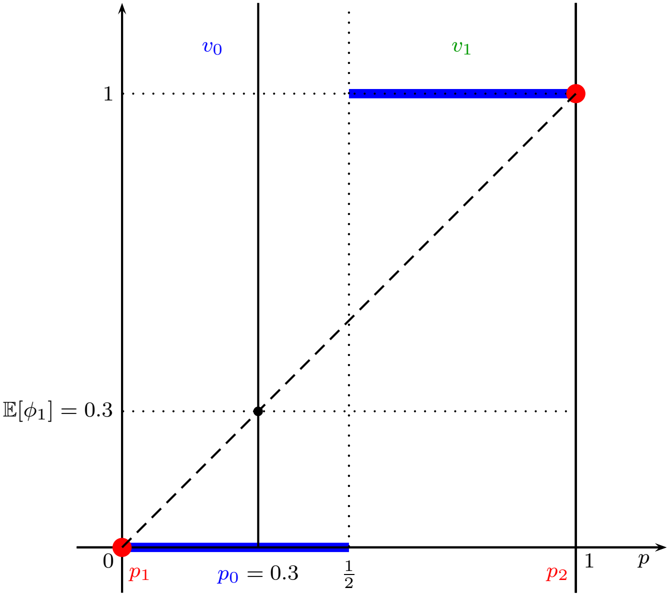
$\phi_2(u,v)$	v_0	v_1
u_0	1	0
u_1	0	1



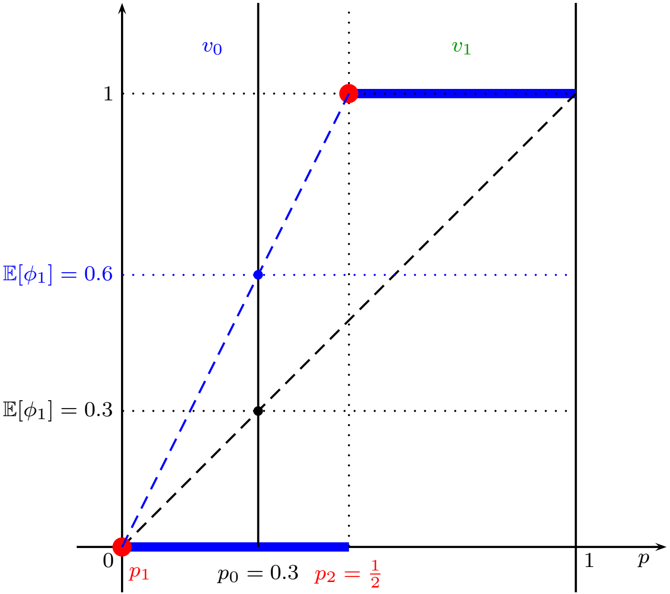
$BR_2(p)$ Best-Response of P_2 depending on his belief p

- Player P_2 plays v_1 iff his **belief** $p \geq \frac{1}{2}$
- Player P_1 wants player P_2 to play v_1
- Player P_1 choose the signalling $\mathbb{P}(x|u)$

Utility of P_1 and full revelation



Full revelation is not optimal for P_1



Characterization

Given decoder's belief $\mathbb{P}(u)$, the Best-Reply $v^\star$ is :

$$v^\star \in \operatorname{argmax} \sum_u \mathbb{P}(u)\phi_2(u,v) = BR_2(\mathbb{P}(u))$$

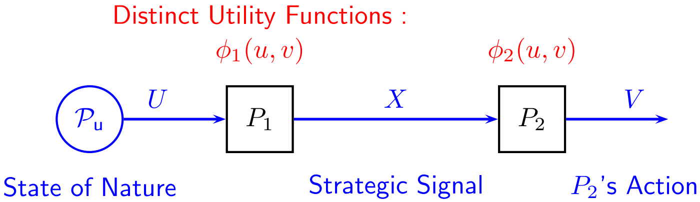
and the sender's payoff is : $\Psi\big(\mathbb{P}(u)\big) = \sum_u \mathbb{P}(u)\phi_1(u, v^\star)$.

Concavification [Kamenica Gentzkow 2011]

$$\begin{aligned} \Phi_1^\star \quad &= \quad \sup \sum_w \mathbb{P}(w)\Psi\big(\mathbb{P}(U|w)\big) \\ \text{s.t.} \quad & \qquad \sum_w \mathbb{P}(w)\mathbb{P}(u|w) = \mathcal{P}(u) \end{aligned}$$

This solution relies on Splitting Lemma [Aumann Maschler 1995]

Persuasion with noisy communication ?



Distinct Utility Functions :

$\phi_1(u,v)$ $\phi_2(u,v)$

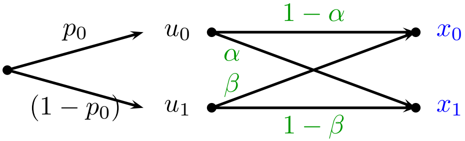
$\mathcal{P}_{\mathsf{u}}$ $\xrightarrow{U}$ P_1 $\xrightarrow{X}$ P_2 $\xrightarrow{V}$

State of Nature

Strategic Signal

P_2's Action

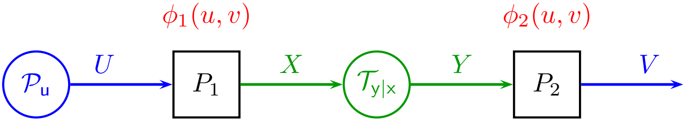
- Player P_1 sends a strategic signal X in order to control posteriors (p_1, p_2),
- Player P_2 implements a best-reply to her belief $p \in [0,1]$



$$\mathcal{P}(u_0|x_0) = \frac{p_0 \cdot (1-\alpha)}{p_0 \cdot (1-\alpha) + (1-p_0) \cdot \beta} = p_1$$

$$\mathcal{P}(u_0|x_1) = \frac{p_0 \cdot \alpha}{p_0 \cdot \alpha + (1-p_0) \cdot (1-\beta)} = p_2$$

Persuasion with noisy communication ?



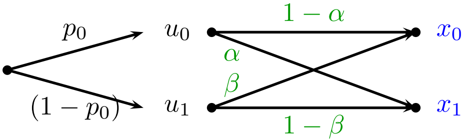
Distinct Utility Functions :

State of Nature

Noisy Channel $\mathcal{T}(y|x)$

P_2's Action

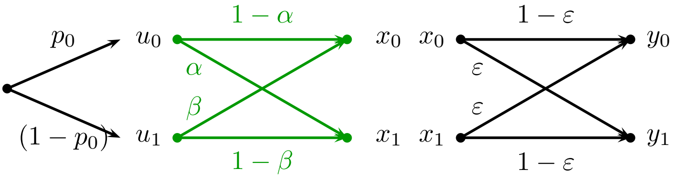
- Player P_1 sends a strategic signal X in order to control posteriors (p_1, p_2),
- Player P_2 implements a best-reply to her belief $p \in [0,1]$

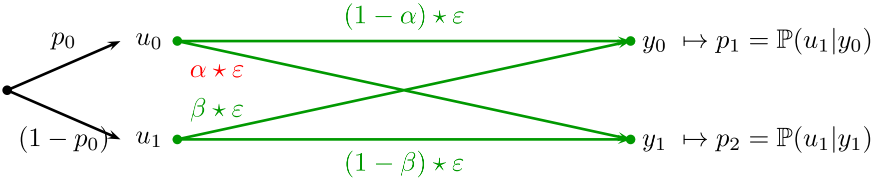


$$\mathcal{P}(u_0|x_0) = \frac{p_0 \cdot (1-\alpha)}{p_0 \cdot (1-\alpha) + (1-p_0) \cdot \beta} = p_1$$

$$\mathcal{P}(u_0|x_1) = \frac{p_0 \cdot \alpha}{p_0 \cdot \alpha + (1-p_0) \cdot (1-\beta)} = p_2$$

One-shot Persuasion with Noisy Channel $\varepsilon \in [0, 0.5]$





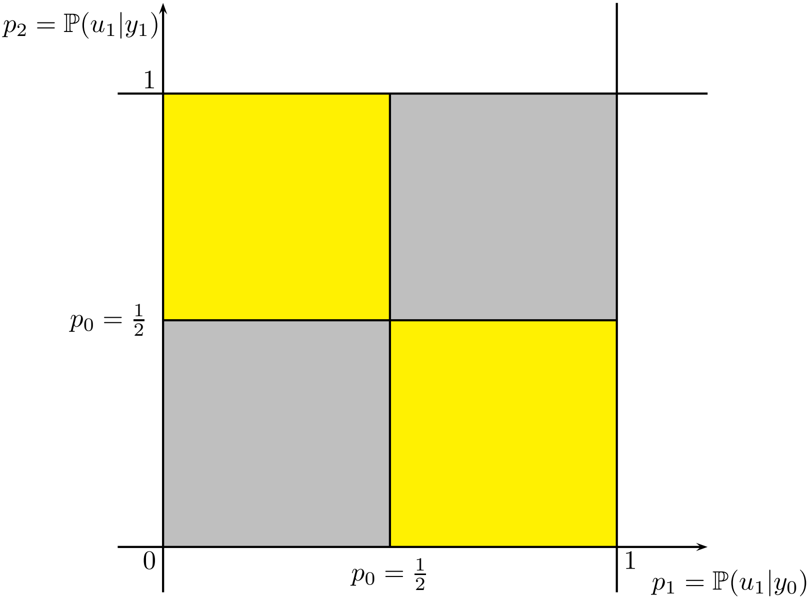
$$\alpha\star\varepsilon = \alpha\cdot(1-\varepsilon)+(1-\alpha)\cdot\varepsilon \quad \in \quad [\varepsilon, 1-\varepsilon]$$

Lemma MLT and Tomala - IEEE Allerton Conference 2016

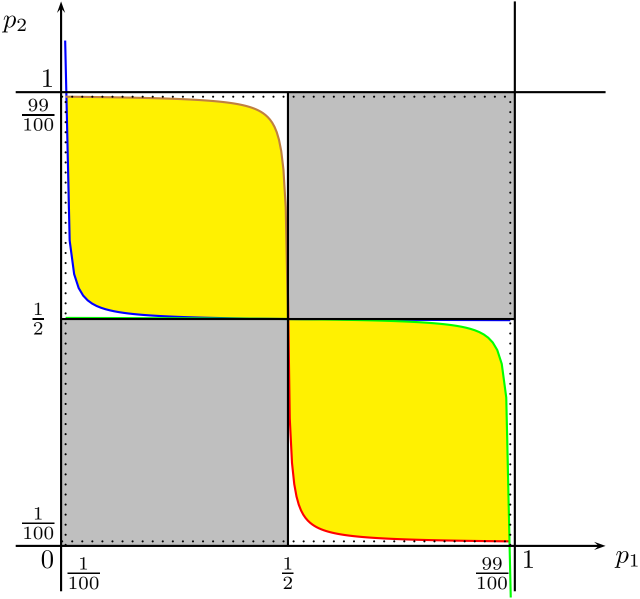
Posterior distributions (p_1, p_2) *are achievable if and only if* $\exists\,(\alpha,\beta)$ *s.t. :*

$$\alpha\star\varepsilon = \frac{(1-p_2)\cdot(p_0-p_1)}{(1-p_0)\cdot(p_2-p_1)}, \qquad \beta\star\varepsilon = \frac{p_1\cdot(p_2-p_0)}{p_0\cdot(p_2-p_1)}.$$

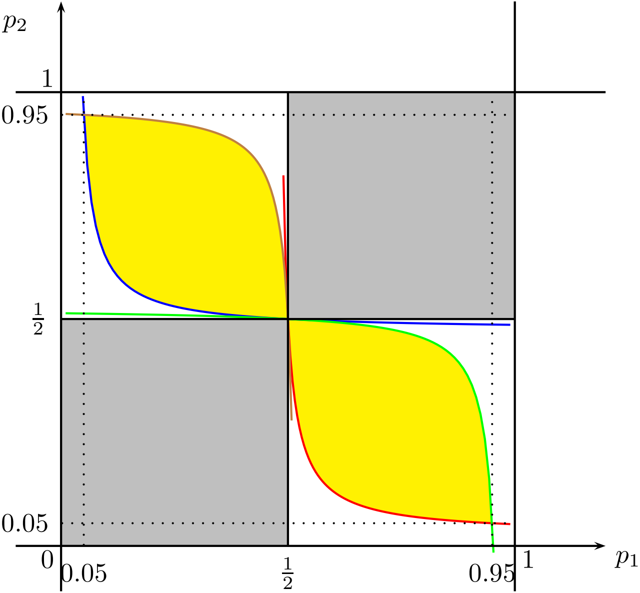
Region of posteriors (p_1, p_2) without channel noise $\varepsilon = 0$



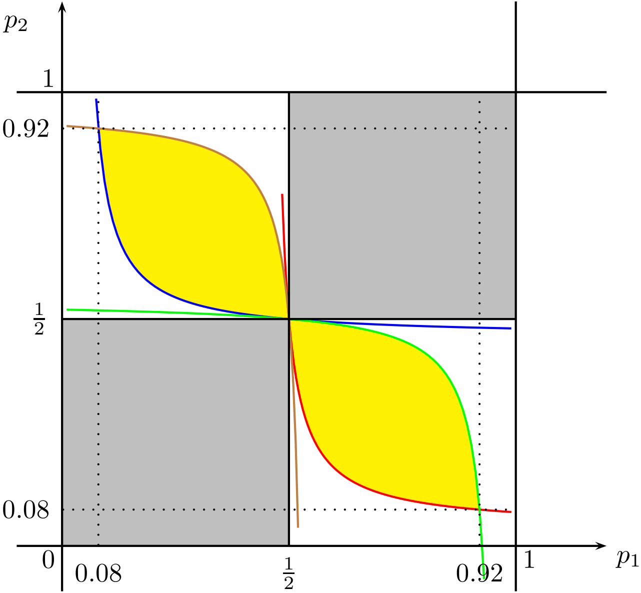
Region of posteriors (p_1, p_2) for channel noise $\varepsilon = 0.01$



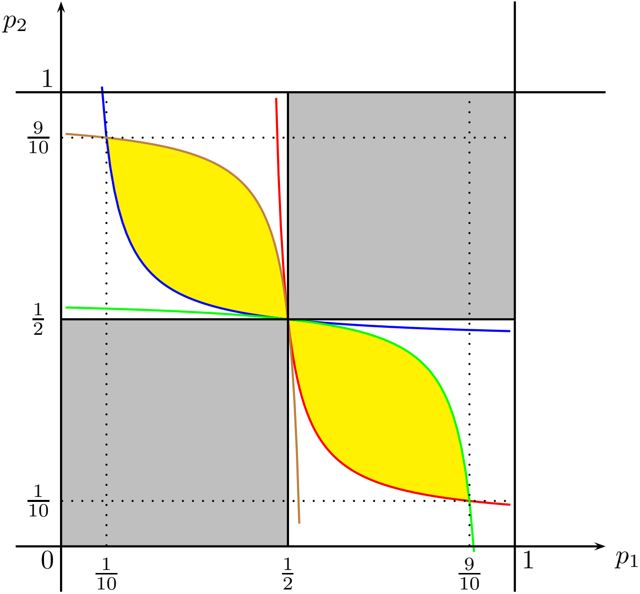
Region of posteriors (p_1, p_2) for channel noise $\varepsilon = 0.05$



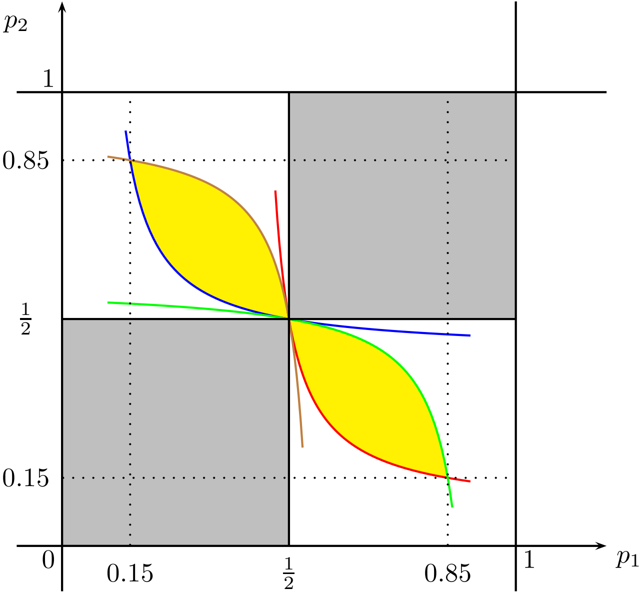
Region of posteriors (p_1, p_2) for channel noise $\varepsilon = 0.08$



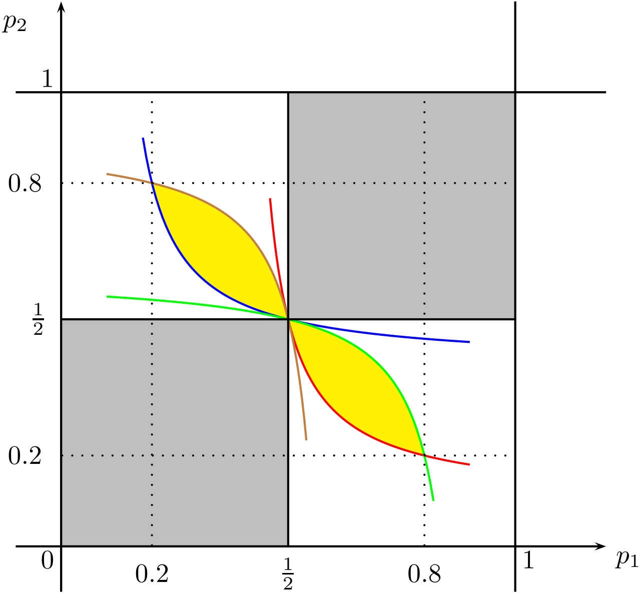
Region of posteriors (p_1, p_2) for channel noise $\varepsilon = 0.1$



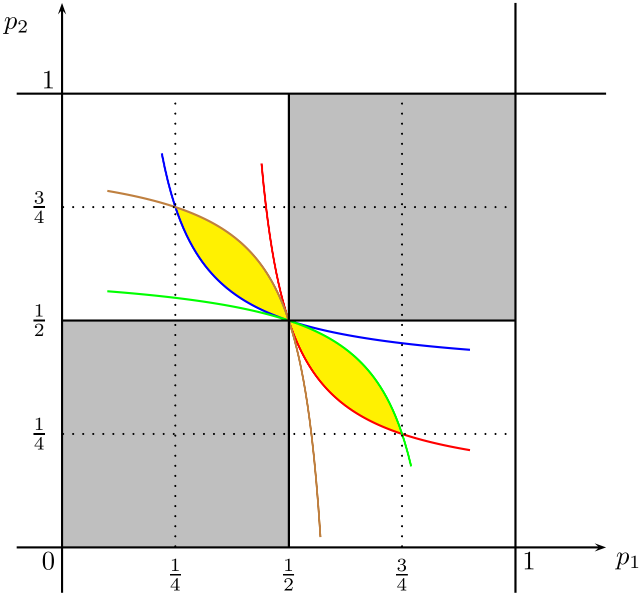
Region of posteriors (p_1, p_2) for channel noise $\varepsilon = 0.15$



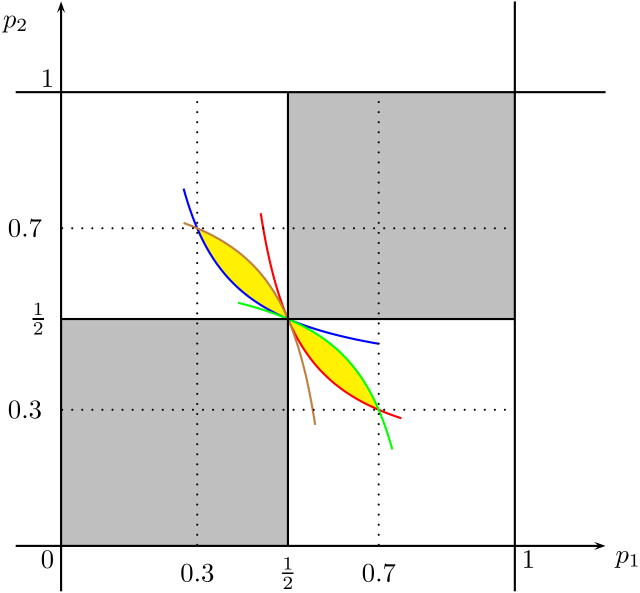
Region of posteriors (p_1, p_2) for channel noise $\varepsilon = 0.20$



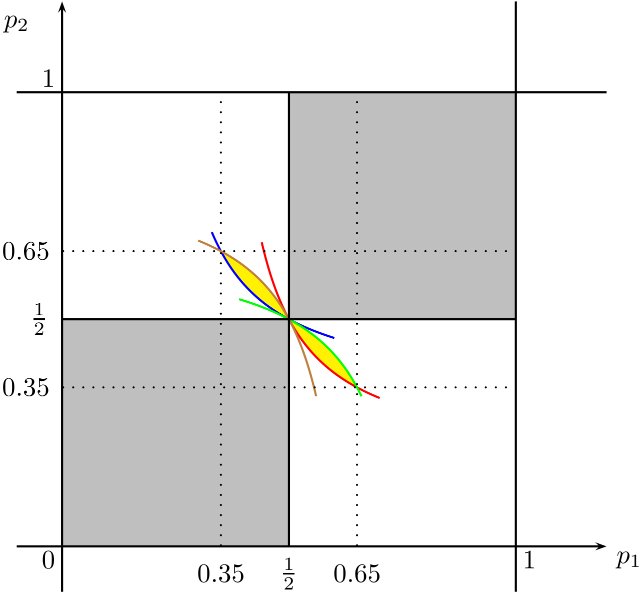
Region of posteriors (p_1, p_2) for channel noise $\varepsilon = 0.25$



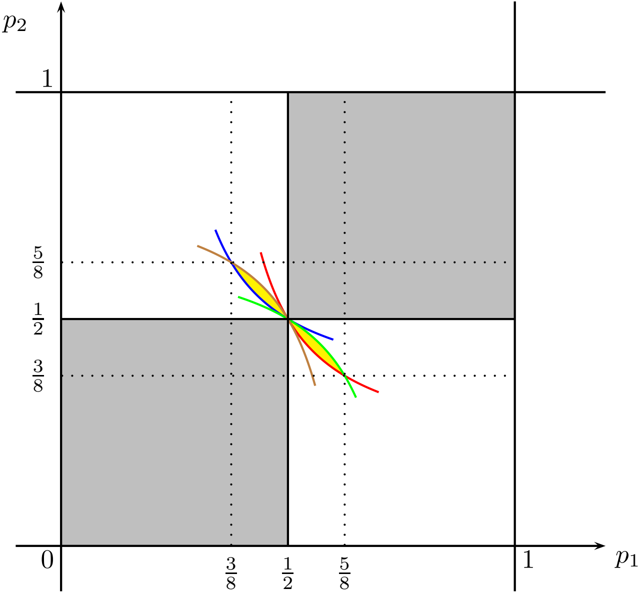
Region of posteriors (p_1, p_2) for channel noise $\varepsilon = 0.30$



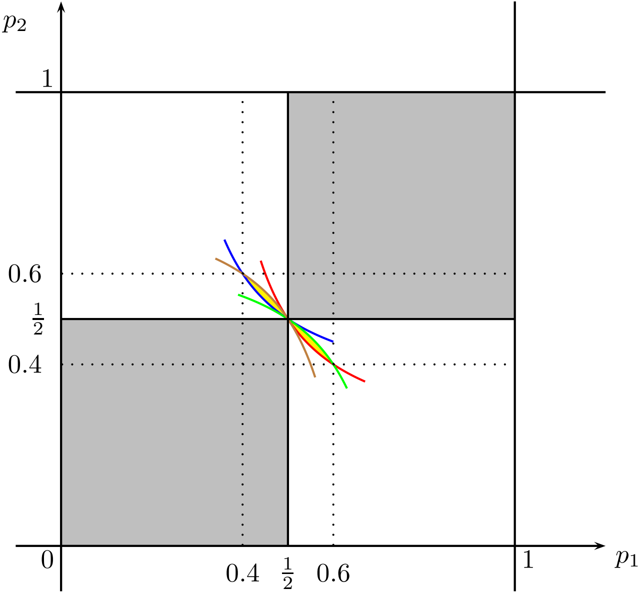
Region of posteriors (p_1, p_2) for channel noise $\varepsilon = 0.35$



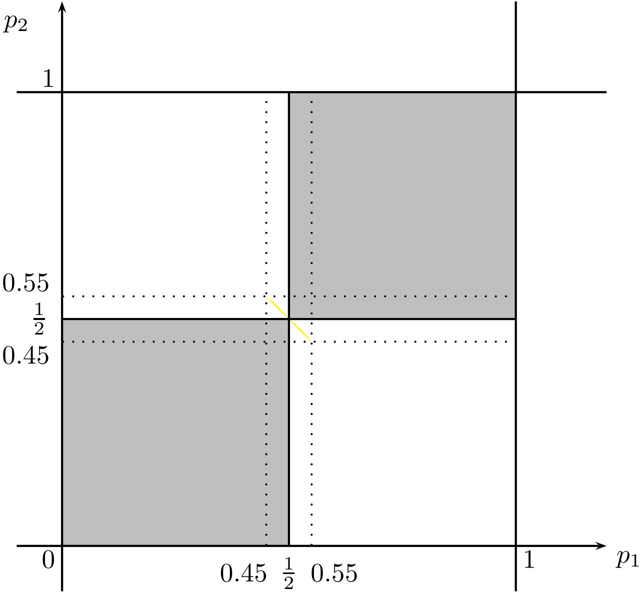
Region of posteriors (p_1, p_2) for channel noise $\varepsilon = 0.375$



Region of posteriors (p_1, p_2) for channel noise $\varepsilon = 0.40$



Region of posteriors (p_1, p_2) for channel noise $\varepsilon = 0.45$



[Le Treust and Tomala, 2019]

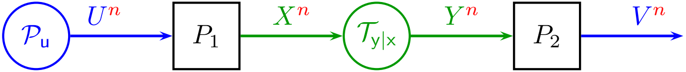
"Persuasion with limited communication capacity"

Journal of Economic Theory, Vol. 184, pp. 104940, Nov. 2019
https ://arxiv.org/abs/1711.04474

Joint source-channel coding

Long-run

Utility Functions : $\Phi_k(\sigma,\tau) = \mathbb{E}_{\sigma,\tau}\left[\frac{1}{n}\sum_{i=1}^{n}\phi_k(u_i,v_i)\right]$ for player $k \in \{1,2\}$



Source i.i.d. Noisy Channel i.i.d. P_2's Action

Player P_1's Strategy : $\sigma : \mathcal{U}^n \to \Delta(\mathcal{X}^n)$

Player P_2's Strategy : $\tau : \mathcal{Y}^n \to \mathcal{V}^n$

- Player P_1 chooses and announces strategy $\sigma(x^n|u^n)$ (commitment power).
- (U^n, X^n, Y^n) are drawn with $\prod_{i=1}^{n}\mathcal{P}_{\mathsf{u}}(u_i) \times \sigma(x^n|u^n) \times \prod_{i=1}^{n}\mathcal{T}(y_i|x_i) = \mathbb{P}_\sigma$
- Sequence Y^n is announced to P_2.
- Player P_2 chooses a sequence of actions with $v^n = \tau(y^n)$.

Information Theory Tools

The entropy $H(U)$ and the mutual information $I(U;W)$ are defined by :

$$H(U) = \mathbb{E}_u\Big[\log_2 \frac{1}{p(u)}\Big] = \sum_u p(u)\log_2 \frac{1}{p(u)}$$

$$I(U;W) = \mathbb{E}_{uw}\Big[\log_2 \frac{p(u,w)}{p(u)p(w)}\Big] = \sum_{uw} p(u,w)\log_2 \frac{p(u,w)}{p(u)p(w)}$$

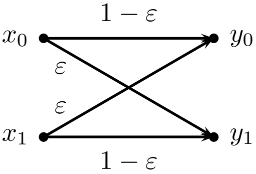
The capacity C of a noisy channel $\mathcal{T}(y|x)$ is defined by :

$$C = \max_{p(x)\in\Delta(\mathcal{X})} I(X;Y) = \max_{p(x)\in\Delta(\mathcal{X})} \mathbb{E}\Big[\log_2 \frac{\mathcal{T}(y|x)}{\sum_x p(x)\cdot\mathcal{T}(y|x)}\Big]$$

Ex : $C = 1 - \Big(\varepsilon\cdot\log_2\frac{1}{\varepsilon} + (1-\varepsilon)\cdot\log_2\frac{1}{1-\varepsilon}\Big)$,

for $\varepsilon = 0.25$, $C \simeq 0.19$ bits,

$|M| = 2^C \simeq 1.14$: average number of messages correctly transmitted per channel use



Characterization

We define $\Psi\big(\mathbb{P}(u)\big)$ the robust payoff at belief $\mathbb{P}(u)$:

$$\Psi\big(\mathbb{P}(u)\big) = \min_{v \in \operatorname{argmax} \sum_u \mathbb{P}(u)\phi_2(u,v)} \sum_u \mathbb{P}(u)\phi_1(u,v)$$

Splitting with information constraint - reformulation of Allerton 2016

$$\begin{aligned} \Phi_1^\star \quad &= \quad \sup \sum_w \mathbb{P}(w)\Psi\big(\mathbb{P}(U|w)\big) \\ \text{s.t.} \quad & \quad \sum_w \mathbb{P}(w)\mathbb{P}(u|w) = \mathcal{P}(u), \\ \text{and} \quad & \quad \sum_w \mathbb{P}(w)H\Big(\mathbb{P}(U|w)\Big) \geq H(U) - C \end{aligned}$$

Information constraint $I(U;W) \leq C$ with auxiliary random variable W.

Main Result

- Player 2's Best Replies $BR_2(\sigma) = \arg\max_\tau \Phi_2(\sigma, \tau)$:

$$BR_2(\sigma) = \arg \max_{v^n = \tau(y^n)} \mathbb{E}_{\sigma,\tau}\left[\frac{1}{n}\sum_i^n \phi_2(u_i, v_i)\right]$$

Theorem MLT and Tomala, JET 2019

We characterize the best payoff player P_1 can secure :

$$1) \qquad \forall n \in \mathbb{N},\ \forall \sigma, \qquad \min_{\tau \in BR_2(\sigma)} \Phi_1(\sigma, \tau) \leq \Phi_1^\star,$$

$$2) \qquad \forall \varepsilon > 0,\ \exists \bar{n},\ \forall n \geq \bar{n},\ \exists \sigma, \qquad \min_{\tau \in BR_2(\sigma)} \Phi_1(\sigma, \tau) \geq \Phi_1^\star - \varepsilon.$$

Comments

Lemma MLT and Tomala, JET 2019

The optimal splitting for $\Phi_1^\star$ has a number of posteriors $|\mathcal{W}|$ restricted to :

$$|\mathcal{W}| = \min\Big(|\mathcal{V}|, |\mathcal{U}| + 1\Big).$$

1) Extend the mapping $\Psi\big(\mathbb{P}(u)\big)$ on the domain :

$$\mathcal{D} = \{(p,h) \in \Delta(\mathcal{U}) \times \mathbb{R} : 0 \leq h \leq H(U)\}$$

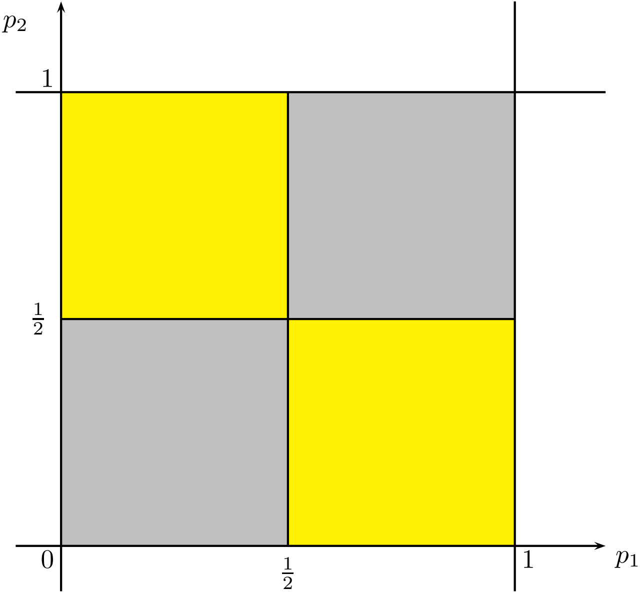
by $\Psi\big(\mathbb{P}(u), h\big) = \Psi\big(\mathbb{P}(u)\big)$. Then, $\Phi_1^\star = \mathsf{Cav}_{\,p,h}\,\Psi\Big(\mathcal{P}(u), H(U) - C\Big)$

2) Lagrangian of the concavification :

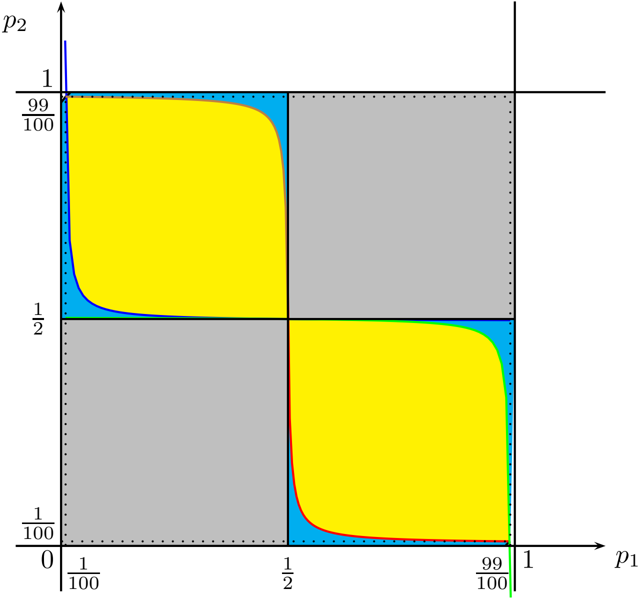
$$\Phi_1^\star = \inf_{t\geq 0}\Big\{\mathsf{Cav}\,(\Psi + tH)(p_0) - t(H(U) - C)\Big\}$$

related with the **Cost of Information** in [Kamenica Gentzkow (AER) 2014]

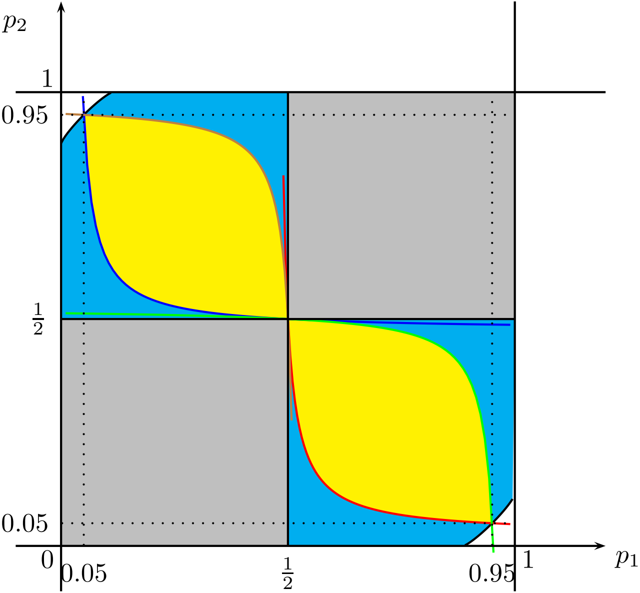
Region of Posteriors (p_1, p_2) without Noise $\varepsilon = 0$



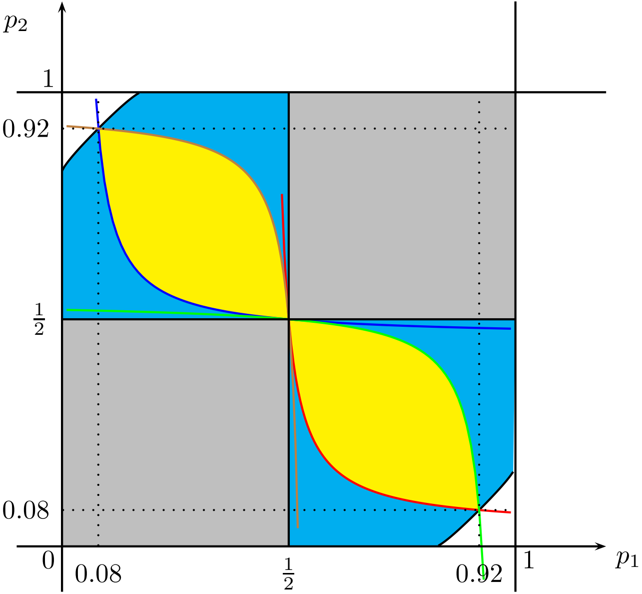
Region of Posteriors (p_1, p_2) for Channel Noise $\varepsilon = 0.01$



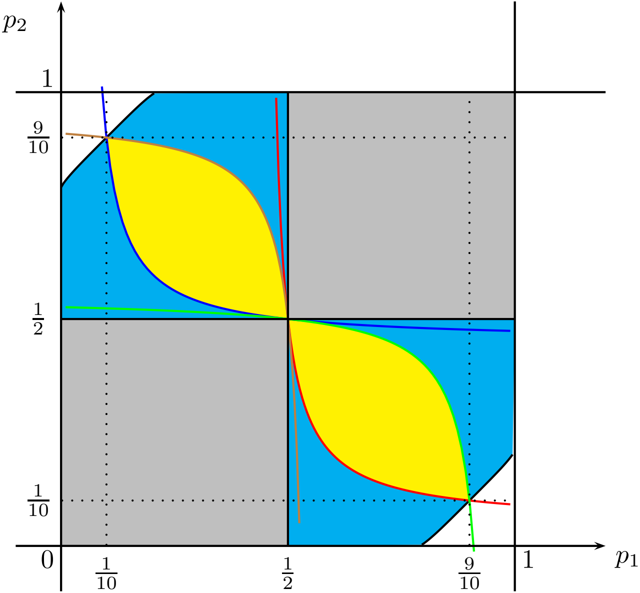
Region of Posteriors (p_1, p_2) for Channel Noise $\varepsilon = 0.05$



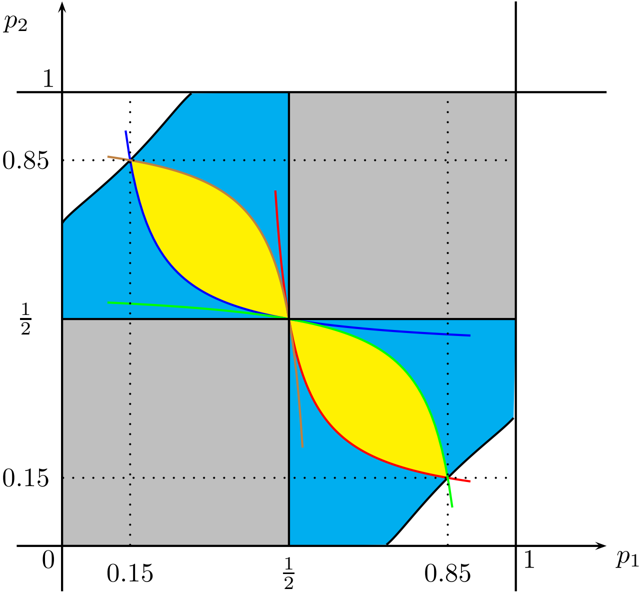
Region of Posteriors (p_1, p_2) for Channel Noise $\varepsilon = 0.08$



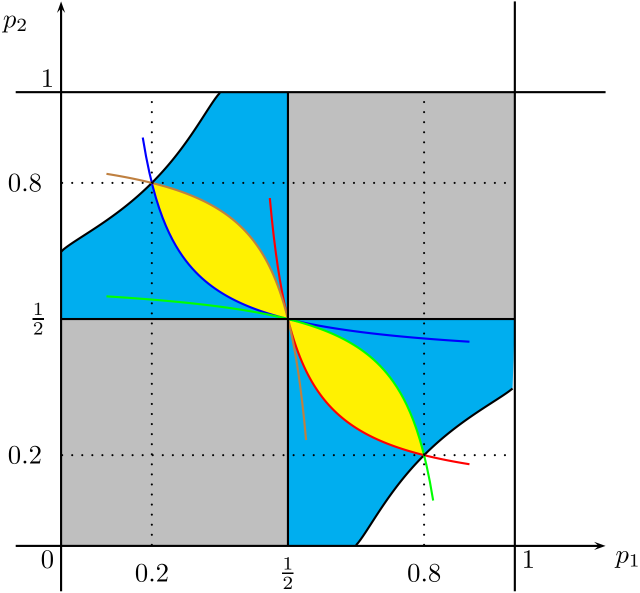
Region of Posteriors (p_1, p_2) for Channel Noise $\varepsilon = 0.1$



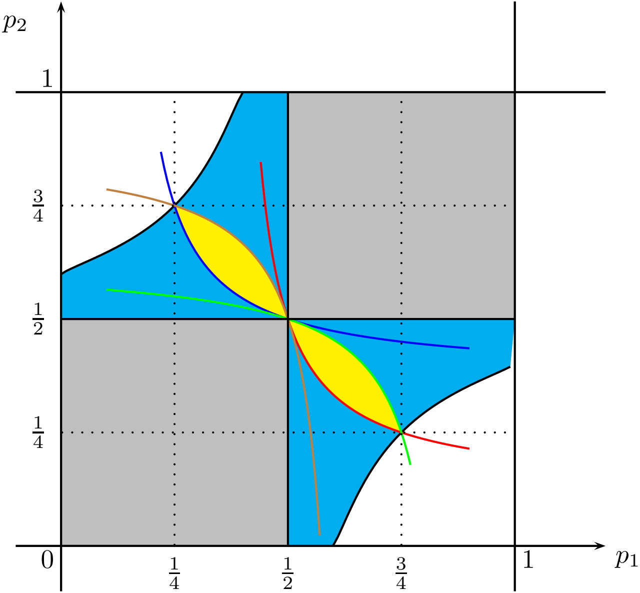
Region of Posteriors (p_1, p_2) for Channel Noise $\varepsilon = 0.15$



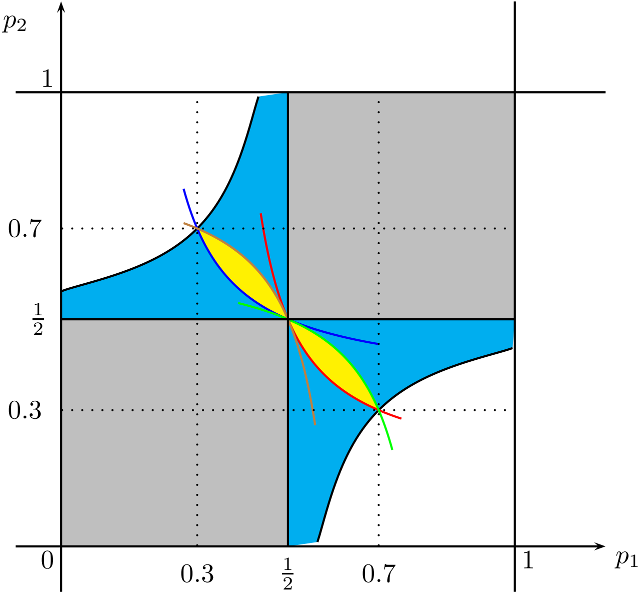
Region of Posteriors (p_1, p_2) for Channel Noise $\varepsilon = 0.20$



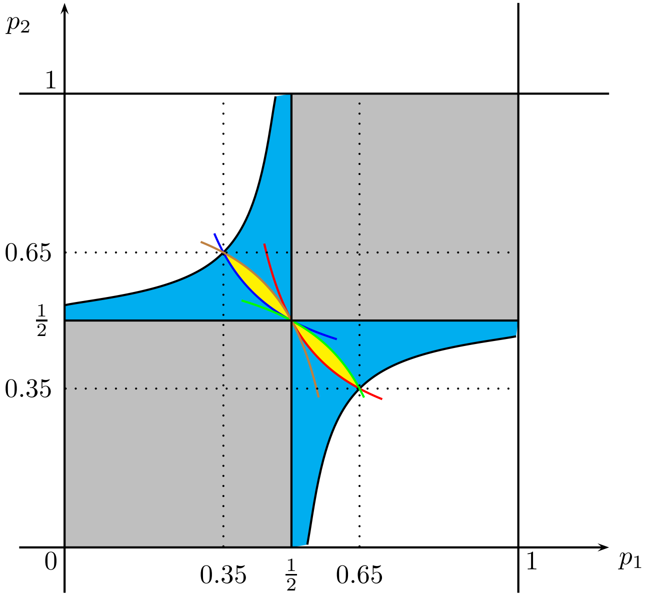
Region of Posteriors (p_1, p_2) for Channel Noise $\varepsilon = 0.25$



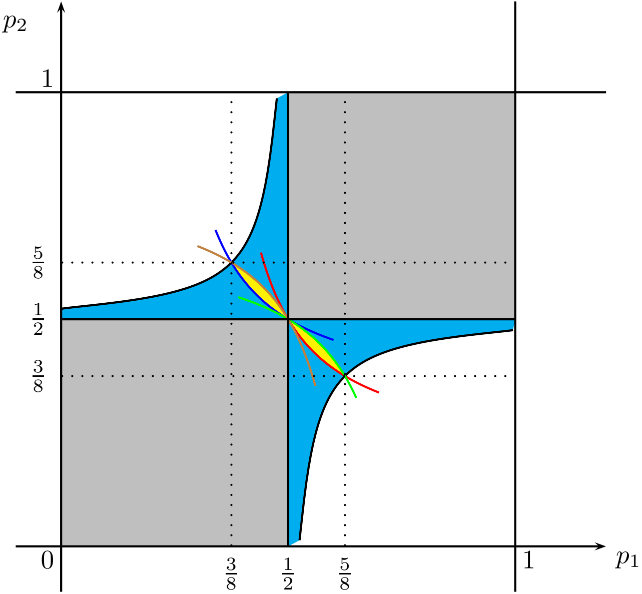
Region of Posteriors (p_1, p_2) for Channel Noise $\varepsilon = 0.30$



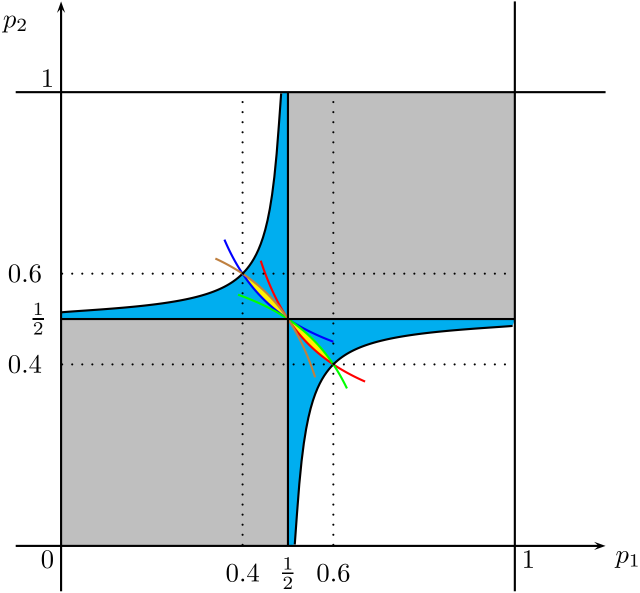
Region of Posteriors (p_1, p_2) for Channel Noise $\varepsilon = 0.35$



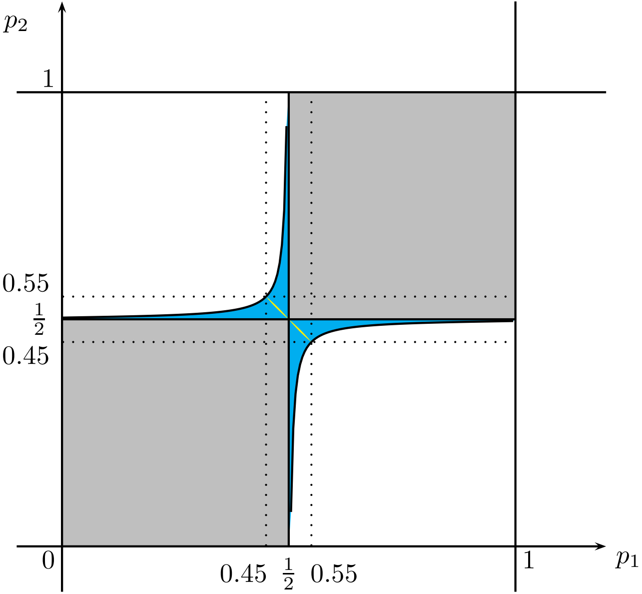
Region of Posteriors (p_1, p_2) for Channel Noise $\varepsilon = 0.375$



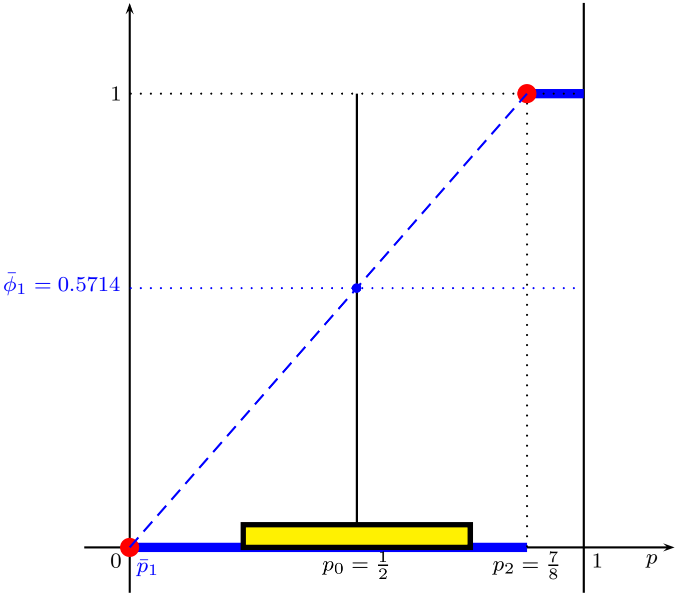
Region of Posteriors (p_1, p_2) for Channel Noise $\varepsilon = 0.40$



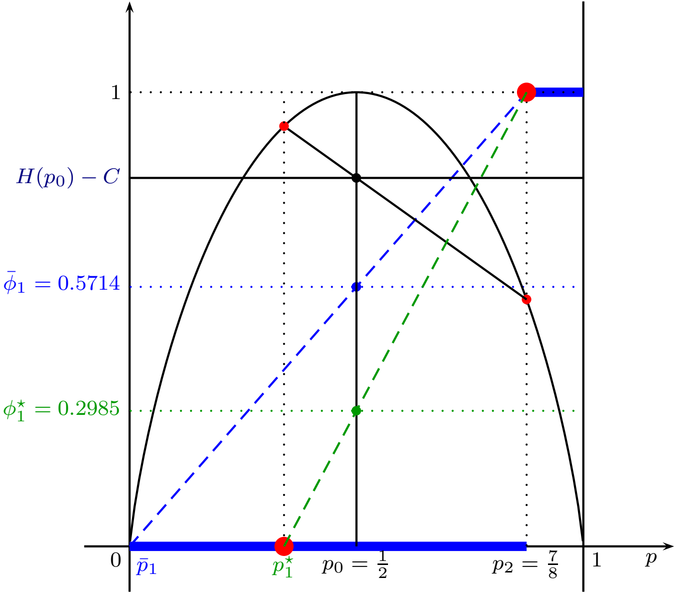
Region of Posteriors (p_1, p_2) for Channel Noise $\varepsilon = 0.45$



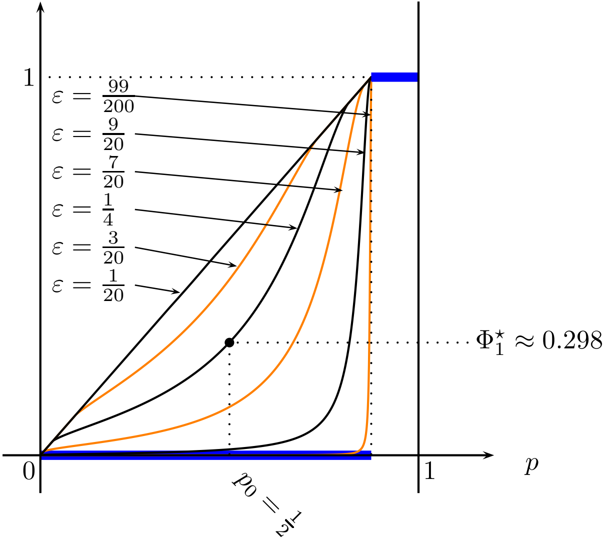
Example : One-sided investment for channel noise $\varepsilon = 0.25$



Example : One-sided investment for channel noise $\varepsilon = 0.25$



Optimal value for noise parameters ε



Sketch of proof

- Converse proof follows from identification of the auxiliary R.V. $W_T = (Y^n, T)$.

- Achievability proof

1) Sequences are jointly typical : Shannon 1948 "random coding scheme"

$(U^n, W^n) \in A(\mathcal{Q})$, for the target probability distribution $\mathbb{P}(u) \times \mathcal{Q}(w|u)$,

2) Control of the Posterior Beliefs induced by the coding process :

$$\mathbb{P}_\sigma(U_i|Y^n) \sim \mathcal{Q}(U_i|W_i)$$

$$\mathbb{E}_\sigma\left[\frac{1}{n}\sum_{i=1}^{n} D\Big(\mathbb{P}_\sigma(U_i|Y^n, E_\delta^0)\Big|\Big|\mathcal{Q}(U_i|W_i)\Big)\right] \leq \varepsilon$$

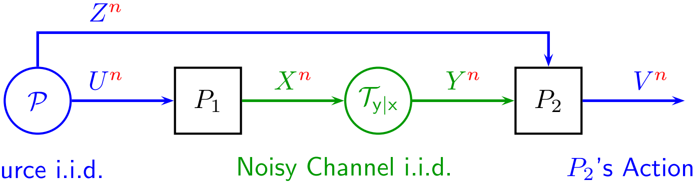
[Le Treust and Tomala, 2018]

"Information-Theoretic Limits of Strategic Communication"

https ://arxiv.org/abs/1807.05147

Joint Wyner-Ziv and channel codings

Utility Functions : $\Phi_k(\sigma,\tau) = \mathbb{E}_{\sigma,\tau}\left[\frac{1}{n}\sum_{i=1}^{n}\phi_k(u_i,v_i)\right]$ for player $k \in \{1,2\}$



Player P_1's Strategy :

$\sigma : \mathcal{U}^n \to \Delta(\mathcal{X}^n)$

Player P_2's Strategy :

$\tau : \mathcal{Y}^n \times \mathcal{Z}^n \to \mathcal{V}^n$

- Player P_1 chooses and announces strategy $\sigma(x^n|u^n)$ (commitment power).
- $(U^n, Z^n, X^n, Y^n) \sim \prod_{i=1}^{n}\mathcal{P}(u_i,z_i) \times \sigma(x^n|u^n) \times \prod_{i=1}^{n}\mathcal{T}(y_i|x_i) = \mathbb{P}_\sigma$
- Sequences (Y^n, Z^n) are observed by Player P_2.
- Player P_2 chooses a sequence of actions with $v^n = \tau(y^n, z^n)$.

Solution

Auxiliary random variable W with $|\mathcal{W}| = \min\left(|\mathcal{U}|+1, |\mathcal{V}|^{|\mathcal{Z}|}\right)$

$$\mathbb{Q}_0 = \left\{ \mathcal{P}_{\mathsf{uz}}(u,z) \times \mathcal{Q}(w|u), \quad \text{s.t.,} \quad \max_{\mathcal{P}(x)} I(X;Y) - I(U;W|Z) \geq 0 \right\}$$

$$\mathbb{Q}_2\big(\mathcal{Q}(u,z,w)\big) = \text{argmax}_{\mathcal{Q}(v|z,w)} \mathbb{E}_{\substack{\mathcal{Q}(u,z,w) \\ \times \mathcal{Q}(v|z,w)}} \Big[\phi_2(U,Z,V)\Big]$$

Define the optimal utility $\Phi_1^\star$ for Player P_1 :

$$\Phi_1^\star = \sup_{\mathcal{Q}(u,z,w) \in \mathbb{Q}_0} \min_{\substack{\mathcal{Q}(v|z,w) \in \\ \mathbb{Q}_2\left(\mathcal{Q}(u,z,w)\right)}} \mathbb{E}_{\substack{\mathcal{Q}(u,z,w) \\ \times \mathcal{Q}(v|z,w)}} \Big[\phi_1(U,Z,V)\Big]$$

Reformulation as an optimal splitting problem

Markov chain $W -\!\!\ominus\!\!- U -\!\!\ominus\!\!- Z \quad \Longleftrightarrow \quad \mathbb{P}(u,z,w) = \mathcal{P}(u,z)\mathbb{P}(w|u), \ \forall (u,z,w)$

$$\Longrightarrow \quad \mathbb{P}(u|z,w) = \frac{\mathbb{P}(u|w)\mathcal{P}(z|u)}{\sum_{u'} \mathbb{P}(u'|w)\mathcal{P}(z|u')}, \ \forall (u,z,w)$$

Encoder's utility reformulates as a function $\Psi_1(p)$ of decoder's belief $p(u)$:

$$\Psi_1(p) = \sum_{u,z} p(u) \cdot \mathcal{P}(z|u) \cdot \psi_1\left(\frac{p(u) \cdot \mathcal{P}(z|u)}{\sum_{u'} p(u') \cdot \mathcal{P}(z|u')}\right)$$

Conditional entropy $H(U|Z)$ reformulates as a function $h(p)$ of belief $p(u)$:

$$h(p) = \sum_{u,z} p(u) \cdot \mathcal{P}(z|u) \cdot \log_2 \frac{\sum_{u'} p(u') \cdot \mathcal{P}(z|u')}{p(u) \cdot \mathcal{P}(z|u)}$$

Reformulation as an Optimal Splitting

Concavification with information constraint

$$\begin{aligned} \Phi_1^\star \quad &= \quad \sup \sum_w \lambda_w \cdot \Psi_1(p_w) \\ \text{s.t.} \quad & \quad \sum_w \lambda_w \cdot p_w(u) = \mathcal{P}(u) \\ \text{and} \quad & \quad \sum_w \lambda_w \cdot h(p_w) \geq H(U|Z) - C \end{aligned}$$

- Information constraint : $H(U|W,Z) \geq H(U|Z) - C \Longleftrightarrow I(U;W|Z) \leq C$
- Auxiliary RV W is the index of the posterior beliefs
- Problem's dimension is $|\mathcal{U}|$, Caratheodory implies : $|\mathcal{W}| = |\mathcal{U}| + 1$.

Main Result

- Player 2's Best Replies $BR_2(\sigma) = \arg\max_{\tau} \Phi_2(\sigma, \tau)$:

$$BR_2(\sigma) = \arg \max_{v^n = \tau(y^n, z^n)} \mathbb{E}_{\sigma,\tau}\left[\frac{1}{n}\sum_{i}^{n} \phi_2(u_i, v_i)\right]$$

Theorem MLT and Tomala 2018

We characterize the best payoff player P_1 can secure :

$$1) \qquad \forall n \in \mathbb{N},\ \forall \sigma, \qquad \min_{\tau \in BR_2(\sigma)} \Phi_1(\sigma, \tau) \leq \Phi_1^{\star},$$

$$2) \qquad \forall \varepsilon > 0,\ \exists \bar{n},\ \forall n \geq \bar{n},\ \exists \sigma, \qquad \min_{\tau \in BR_2(\sigma)} \Phi_1(\sigma, \tau) \geq \Phi_1^{\star} - \varepsilon.$$

Comments

Lemma MLT and Tomala 2018

The optimal splitting for $\Phi_1^\star$ *has a number of posteriors* $|\mathcal{W}|$ *restricted to :*

$$|\mathcal{W}| = \min\Big(|\mathcal{U}| + 1, |\mathcal{V}|^{|\mathcal{Z}|}\Big).$$

1) Extend the mapping $\Psi_1(p)$ on the domain :

$$\mathcal{D} = \{(p,h) \in \Delta(\mathcal{U}) \times \mathbb{R} : 0 \leq h \leq H(U|Z)\}$$

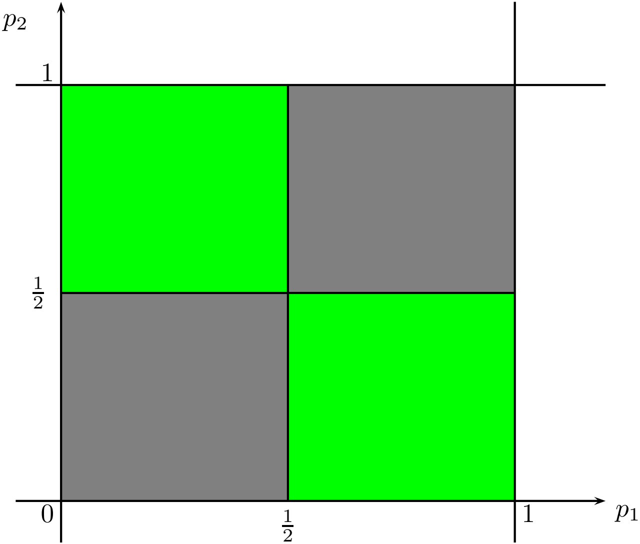
by $\Psi\big(p,h\big) = \Psi(p)$. Then, $\Phi_1^\star = \text{Cav}_{\,p,h}\,\Psi\Big(\mathcal{P}(u), H(U|Z) - C\Big)$

2) Lagrangian of the concavification :

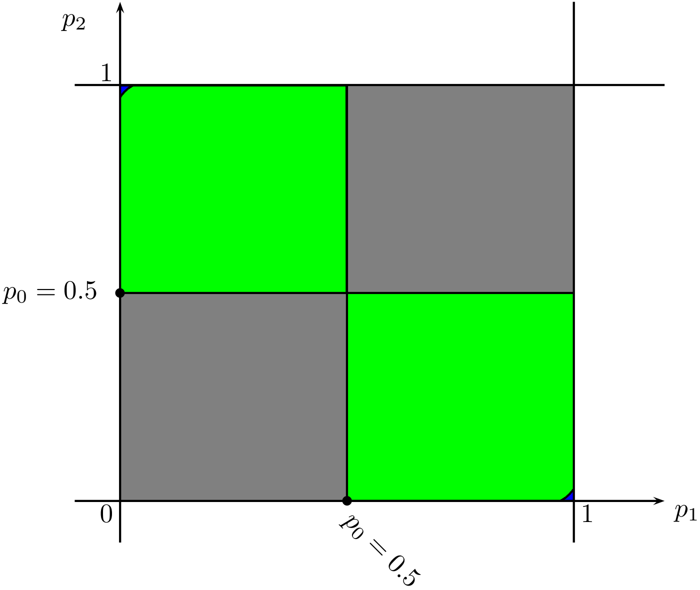
$$\Phi_1^\star = \inf_{t \geq 0}\Big\{\text{Cav}\,\big(\Psi + th\big)\big(\mathcal{P}(u)\big) - t\big(H(U|Z) - C\big)\Big\}$$

related with the **Cost of Information** in [Kamenica Gentzkow (AER) 2014]

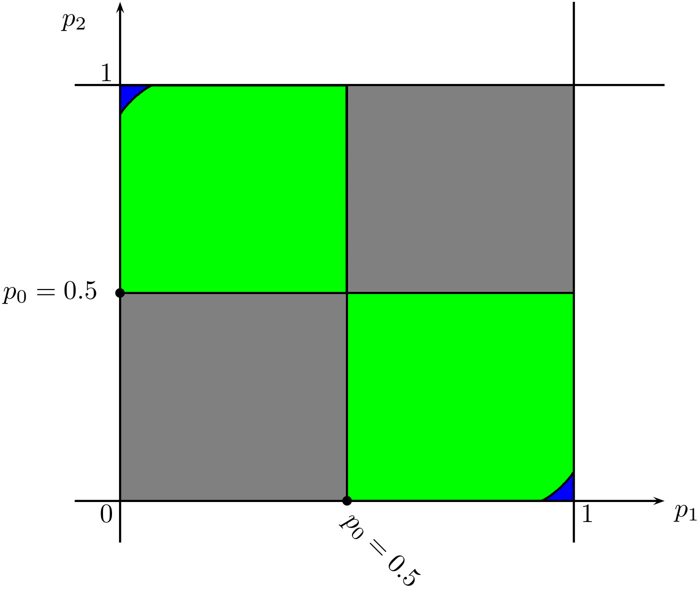
Region of Posteriors (p_1, p_2) for Capacity $C = 1$



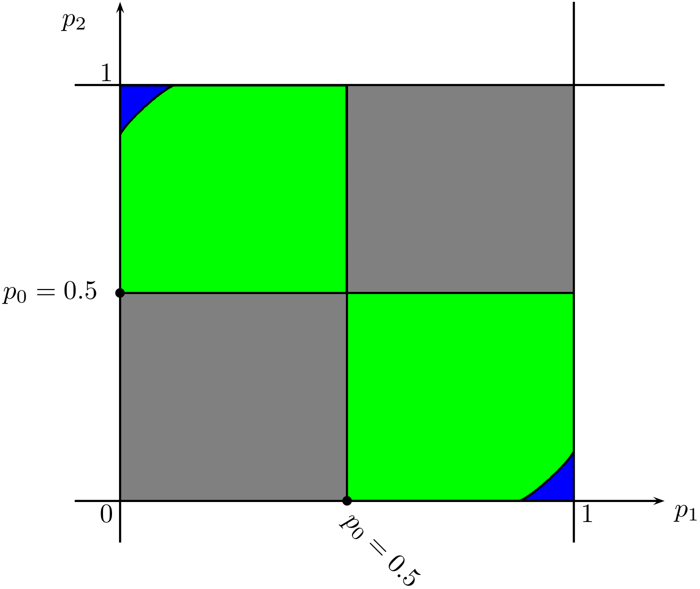
Region of Posteriors (p_1, p_2) for Capacity $C = 0.9$



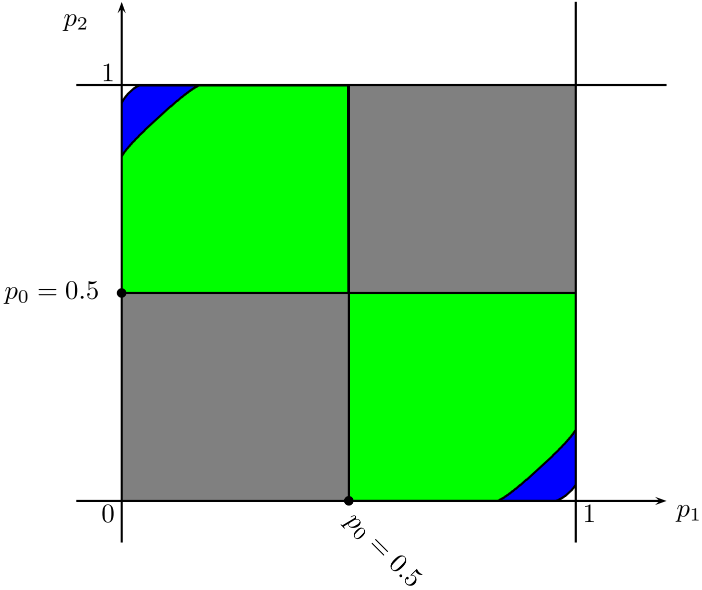
Region of Posteriors (p_1, p_2) for Capacity $C = 0.8$



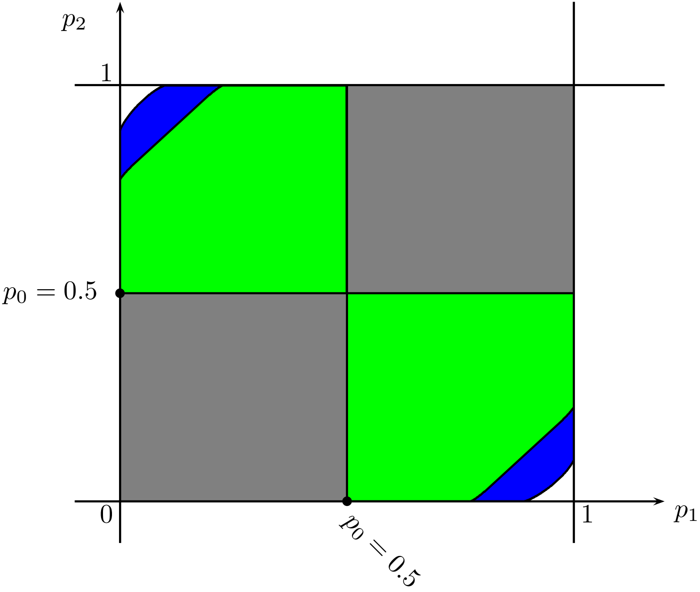
Region of Posteriors (p_1, p_2) for Capacity $C = 0.7$



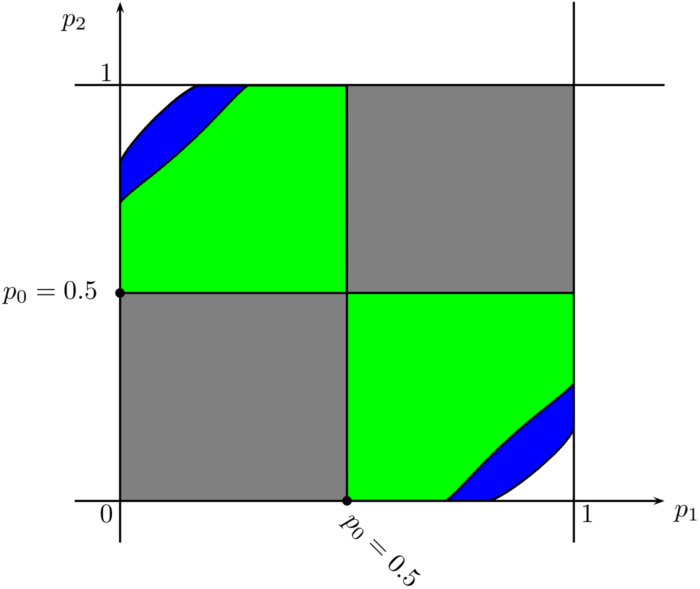
Region of Posteriors (p_1, p_2) for Capacity $C = 0.6$



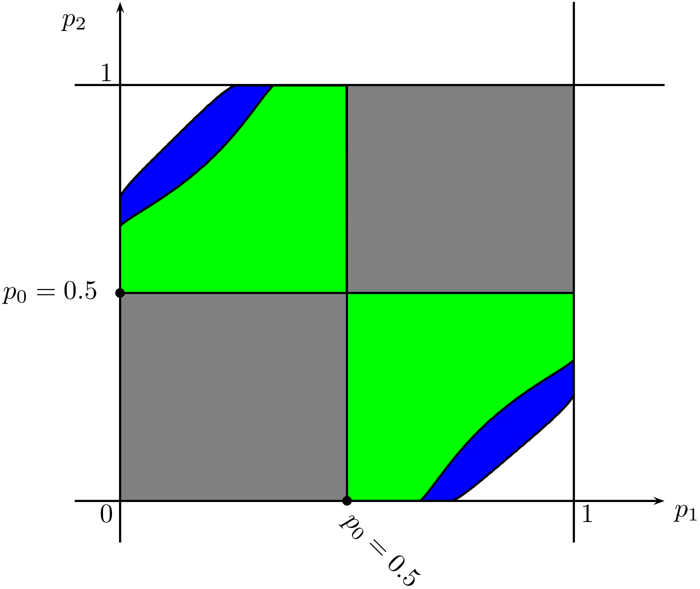
Region of Posteriors (p_1, p_2) for Capacity $C = 0.5$



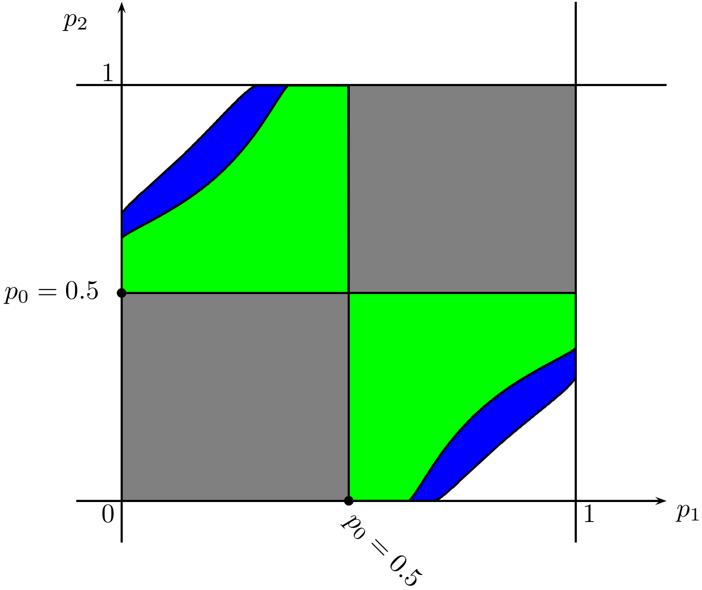
Region of Posteriors (p_1, p_2) for Capacity $C = 0.4$



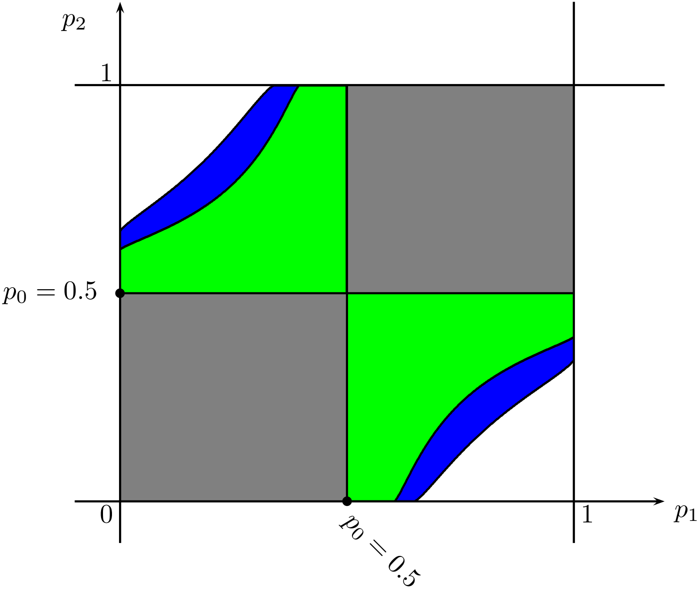
Region of Posteriors (p_1, p_2) for Capacity $C = 0.3$



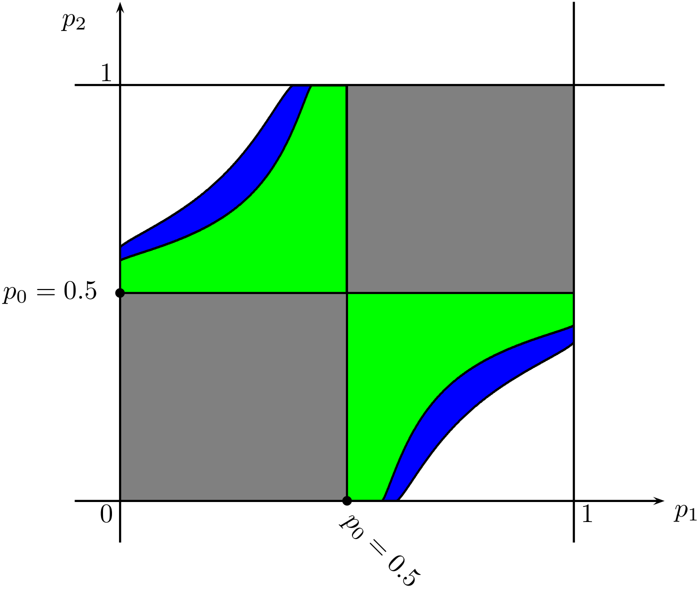
Region of Posteriors (p_1, p_2) for Capacity $C = 0.25$



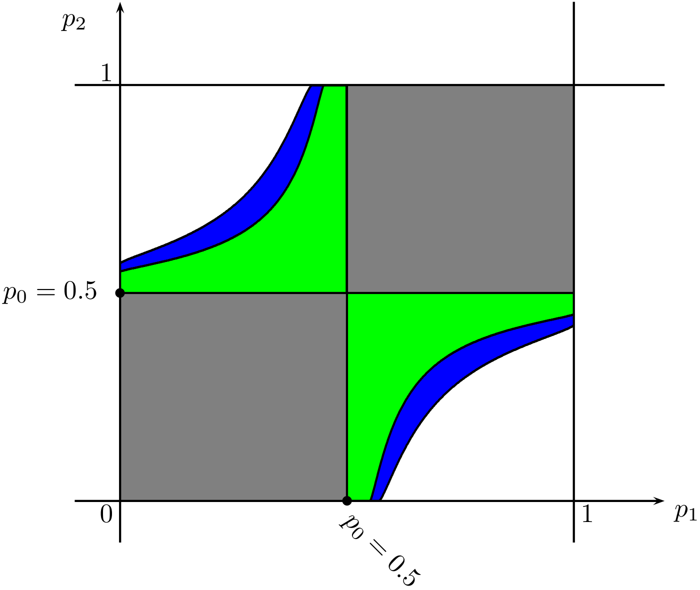
Region of Posteriors (p_1, p_2) for Capacity $C = 0.2$



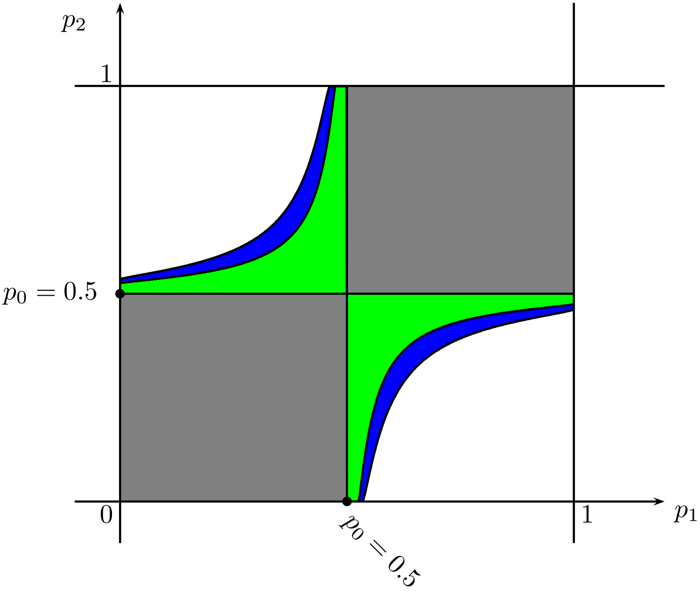
Region of Posteriors (p_1, p_2) for Capacity $C = 0.15$



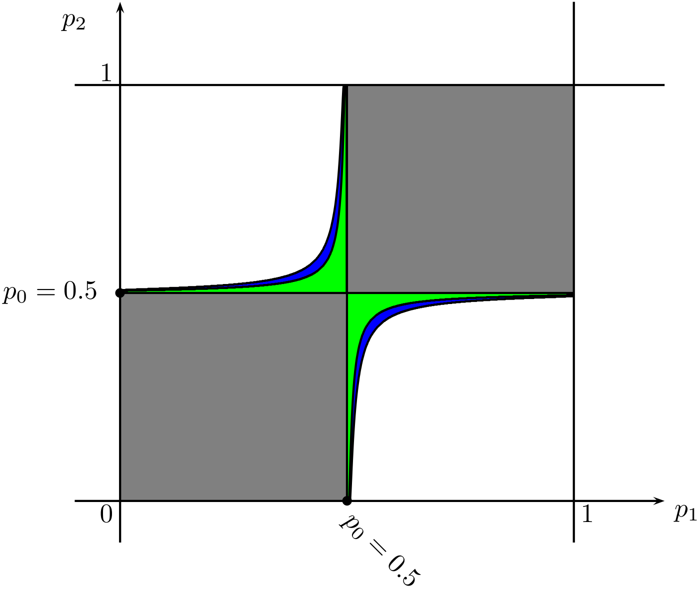
Region of Posteriors (p_1, p_2) for Capacity $C = 0.1$



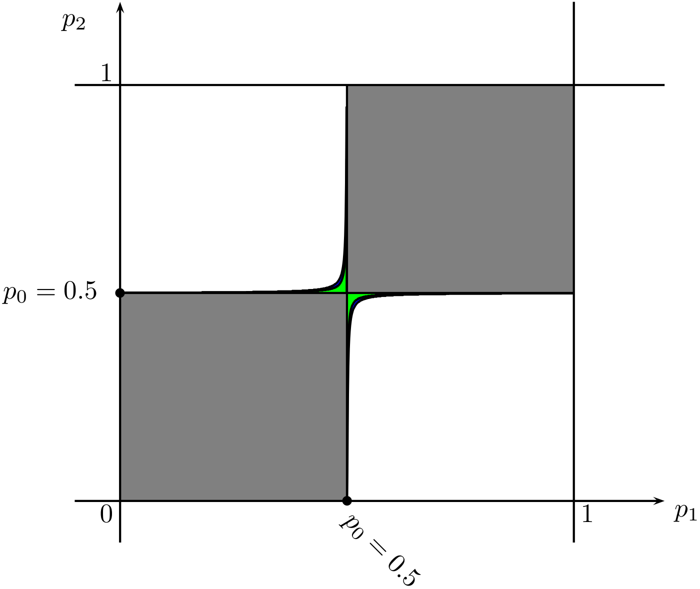
Region of Posteriors (p_1, p_2) for Capacity $C = 0.05$



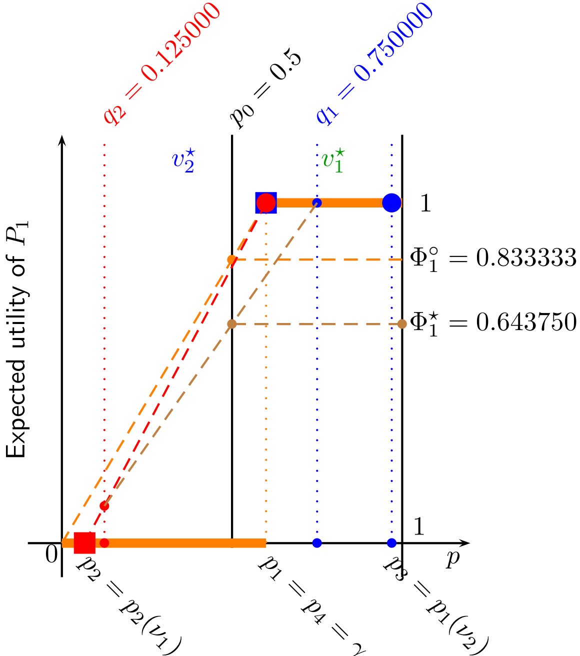
Region of Posteriors (p_1, p_2) for Capacity $C = 0.01$



Region of Posteriors (p_1, p_2) for Capacity $C = 0.001$



Example


$q_2 = 0.125000$
$p_0 = 0.5$
$q_1 = 0.750000$
$v_2^\star$
$v_1^\star$
1
$\Phi_1^\circ = 0.833333$
$\Phi_1^\star = 0.643750$
Expected utility of P_1
1
0
p
$p_2 = p_2(\nu_1)$
$p_1 = p_4 = \gamma$
$p_3 = p_1(\nu_2)$

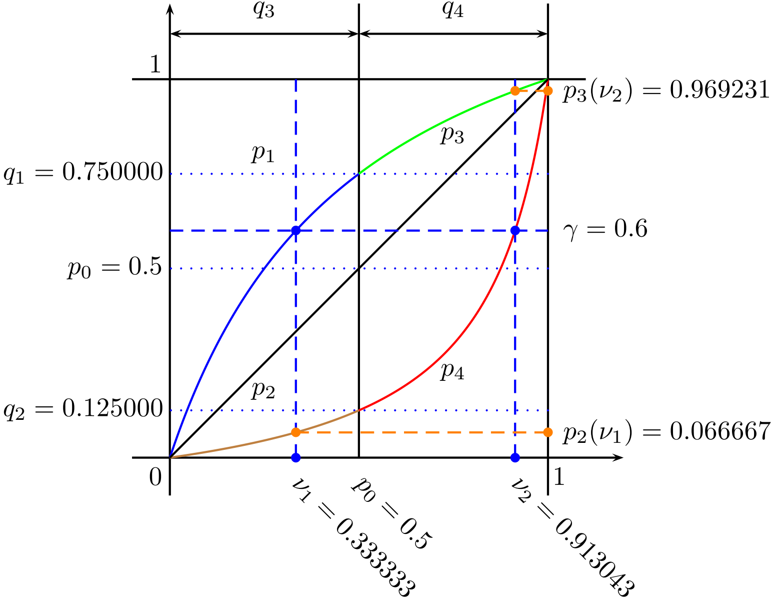
Example

We let $q_1 = \mathbb{P}(u_2|w_1)$, $q_2 = \mathbb{P}(u_2|w_2)$, $\delta_1 = \mathbb{P}(z_2|u_1)$, $\delta_2 = \mathbb{P}(z_1|u_2)$,

$$\begin{aligned}
p_1 = \mathcal{Q}(u_2|w_1, z_1) &= \frac{q_1 \cdot \delta_2}{(1-q_1)\cdot(1-\delta_1) + q_1 \cdot \delta_2}, \\
p_2 = \mathcal{Q}(u_2|w_1, z_2) &= \frac{q_1 \cdot (1-\delta_2)}{(1-q_1)\cdot\delta_1 + q_1 \cdot (1-\delta_2)}, \\
p_3 = \mathcal{Q}(u_2|w_2, z_1) &= \frac{q_2 \cdot \delta_2}{(1-q_2)\cdot(1-\delta_1) + q_2 \cdot \delta_2}, \\
p_4 = \mathcal{Q}(u_2|w_2, z_2) &= \frac{q_2 \cdot (1-\delta_2)}{(1-q_2)\cdot\delta_1 + q_2 \cdot (1-\delta_2)}.
\end{aligned}$$

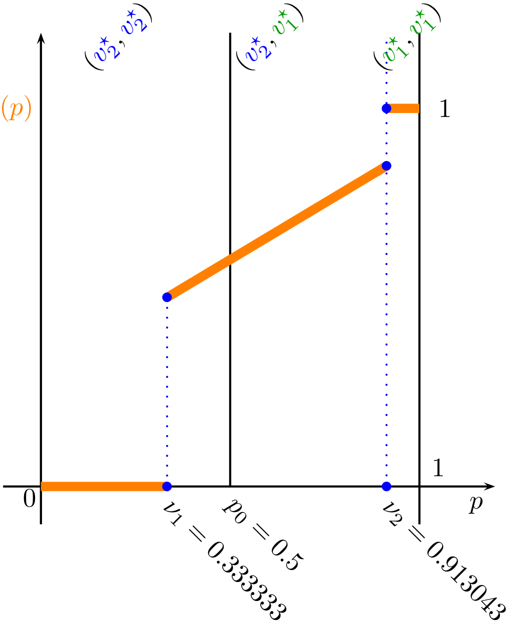
$$\begin{aligned}
p_1(q) &= \frac{q \cdot \delta_2}{(1-q)\cdot(1-\delta_1) + q \cdot \delta_2}, \\
p_2(q) &= \frac{q \cdot (1-\delta_2)}{(1-q)\cdot\delta_1 + q \cdot (1-\delta_2)}.
\end{aligned}$$

Example


q_3
q_4
1
$p_3(\nu_2) = 0.969231$
p_3
p_1
$q_1 = 0.750000$
$\gamma = 0.6$
$p_0 = 0.5$
p_4
p_2
$q_2 = 0.125000$
$p_2(\nu_1) = 0.066667$
0
1
$\nu_1 = 0.333333$
$p_0 = 0.5$
$\nu_2 = 0.913043$

Example

Encoder's Utility: $\Psi_1(p)$

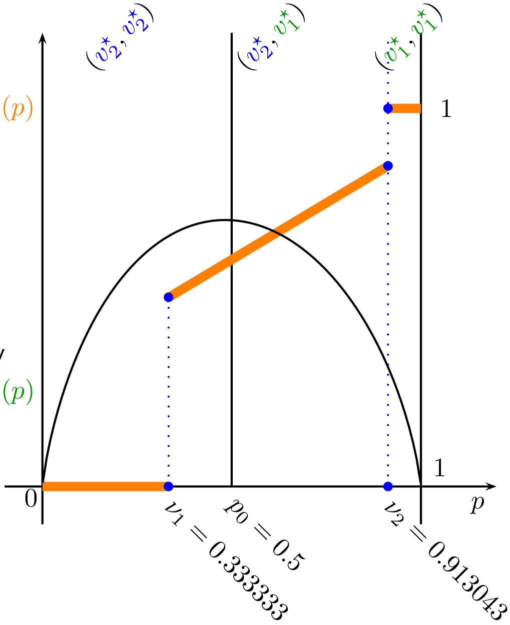


Example

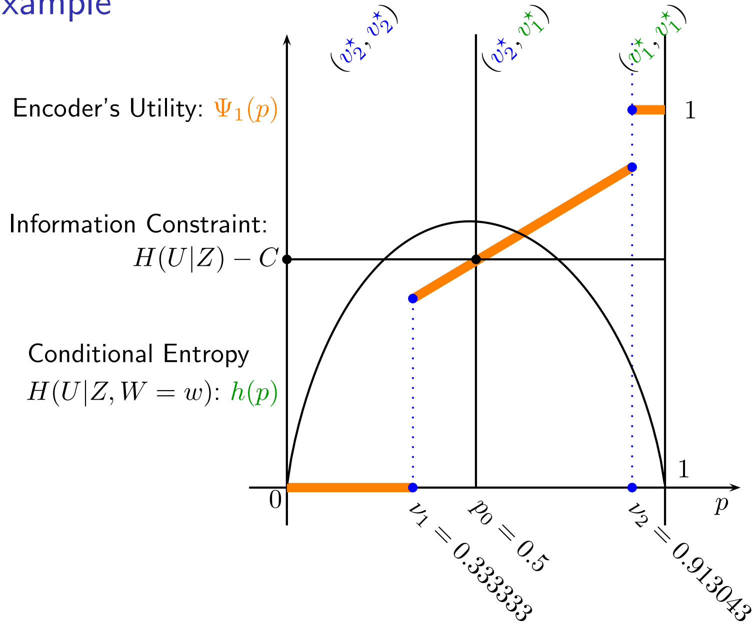
Encoder's Utility: $\Psi_1(p)$

Conditional Entropy

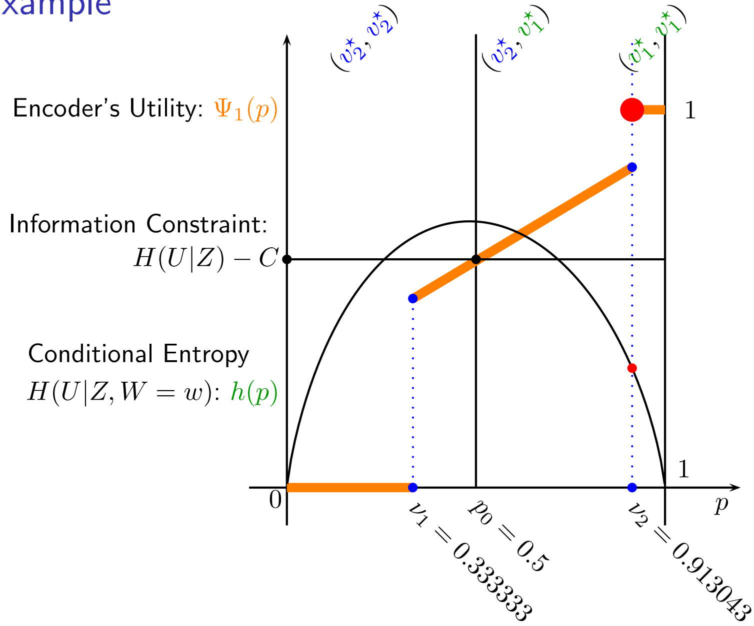
$H(U|Z, W = w)$: $h(p)$



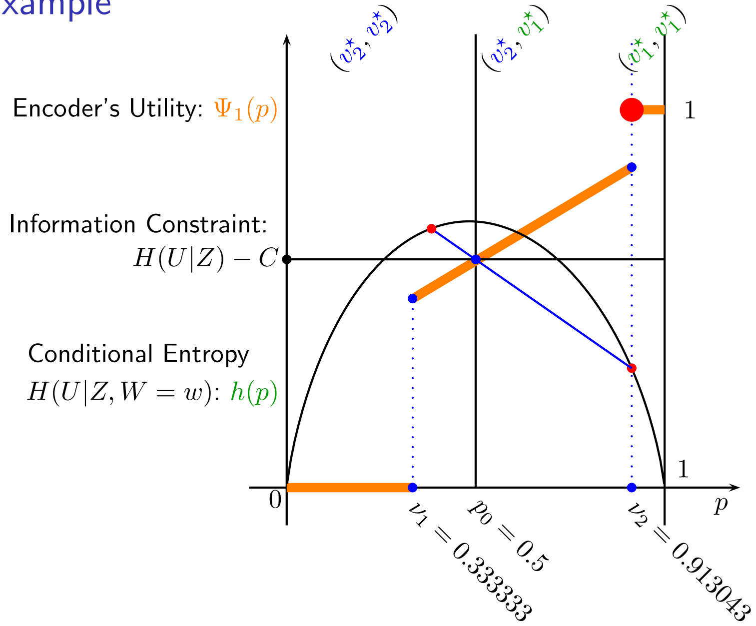
Example


$(v_2^\star, v_2^\star)$
$(v_2^\star, v_1^\star)$
$(v_1^\star, v_1^\star)$
Encoder's Utility: $\Psi_1(p)$
1
Information Constraint:
$H(U|Z) - C$
Conditional Entropy
$H(U|Z, W = w)$: $h(p)$
1
0
p
$\nu_1 = 0.333333$
$p_0 = 0.5$
$\nu_2 = 0.913043$

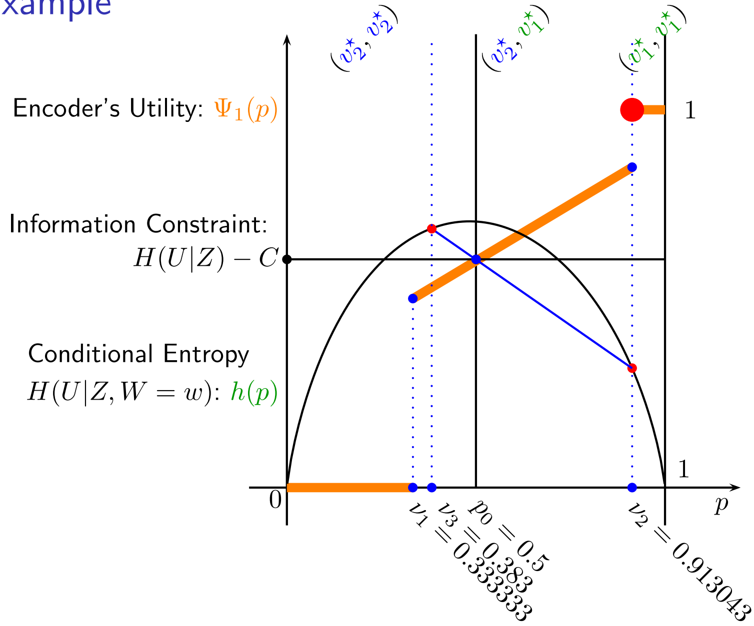
Example


$(\upsilon_2^\star, \upsilon_2^\star)$
$(\upsilon_2^\star, \upsilon_1^\star)$
$(\upsilon_1^\star, \upsilon_1^\star)$
Encoder's Utility: $\Psi_1(p)$
1
Information Constraint:
$H(U|Z) - C$
Conditional Entropy
$H(U|Z, W = w)$: $h(p)$
1
0
p
$\nu_1 = 0.333333$
$p_0 = 0.5$
$\nu_2 = 0.913043$

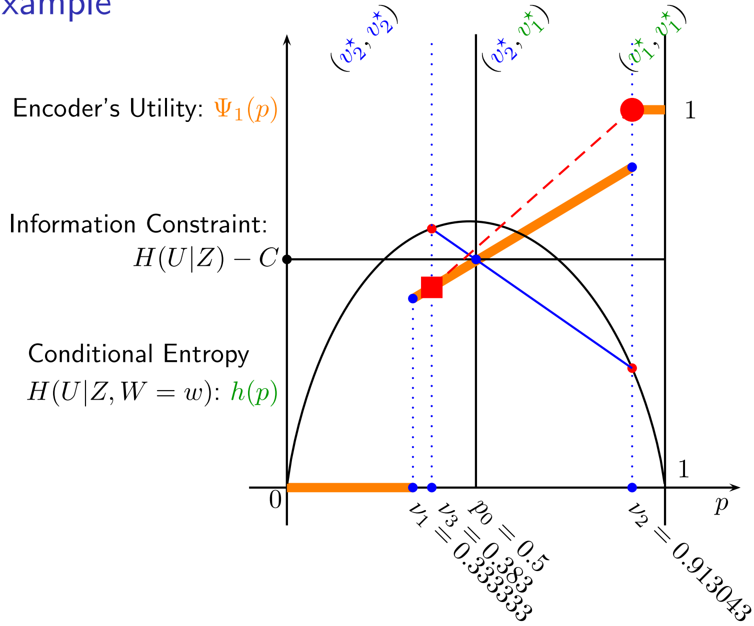
Example


$(v_2^\star, v_2^\star)$
$(v_2^\star, v_1^\star)$
$(v_1^\star, v_1^\star)$
Encoder's Utility: $\Psi_1(p)$
1
Information Constraint:
$H(U|Z) - C$
Conditional Entropy
$H(U|Z, W = w)$: $h(p)$
1
0
p
$\nu_1 = 0.333333$
$p_0 = 0.5$
$\nu_2 = 0.913043$

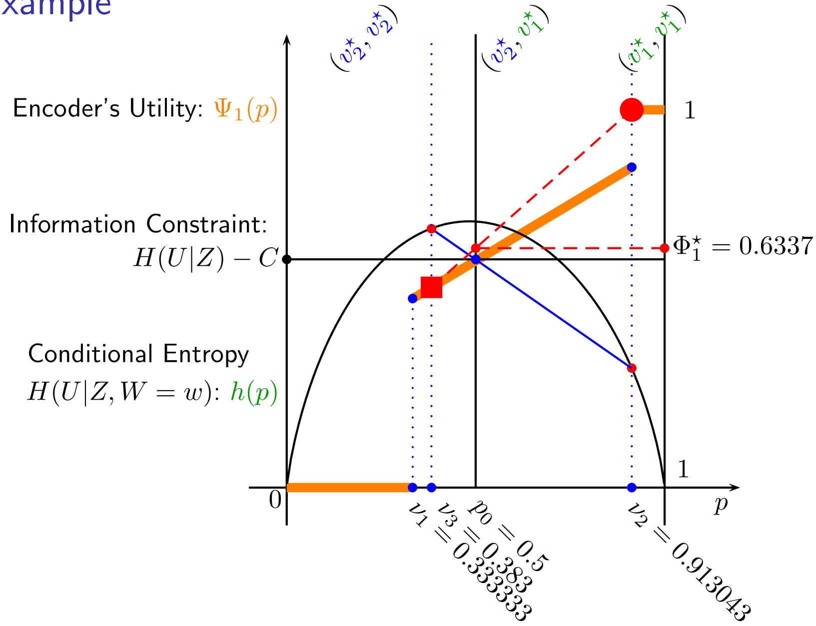
Example


Encoder's Utility: $\Psi_1(p)$
Information Constraint:
$H(U|Z) - C$
Conditional Entropy
$H(U|Z, W = w)$: $h(p)$
$(\nu_2^\star, \nu_2^\star)$
$(\nu_2^\star, \nu_1^\star)$
$(\nu_1^\star, \nu_1^\star)$
1
1
0
p
$\nu_1 = 0.333333$
$\nu_3 = 0.383$
$p_0 = 0.5$
$\nu_2 = 0.913043$

Example

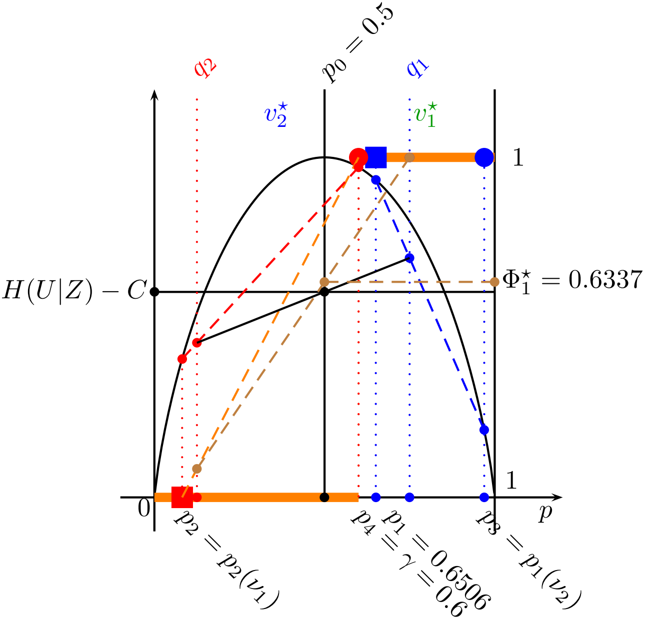

Encoder's Utility: $\Psi_1(p)$
Information Constraint:
$H(U|Z) - C$
Conditional Entropy
$H(U|Z, W = w)$: $h(p)$
$(v_2^\star, v_2^\star)$
$(v_2^\star, v_1^\star)$
$(v_1^\star, v_1^\star)$
1
1
0
p
$\nu_1 = 0.333333$
$\nu_3 = 0.383$
$p_0 = 0.5$
$\nu_2 = 0.913043$

Example



$(\nu_2^\star, \nu_2^\star)$
$(\nu_2^\star, \nu_1^\star)$
$(\nu_1^\star, \nu_1^\star)$
Encoder's Utility: $\Psi_1(p)$
1
Information Constraint:
$H(U|Z) - C$
$\Phi_1^\star = 0.6337$
Conditional Entropy
$H(U|Z, W = w)$: $h(p)$
1
0
p
$\nu_1 = 0.333333$
$\nu_3 = 0.383$
$p_0 = 0.5$
$\nu_2 = 0.913043$

Example


$p_0 = 0.5$
q_2
q_1
$v_2^\star$
$v_1^\star$
1
$H(U|Z) - C$
$\Phi_1^\star = 0.6337$
1
0
p
$p_2 = p_2(\nu_1)$
$p_4 = \gamma = 0.6$
$p_1 = 0.6506$
$p_3 = p_1(\nu_2)$

Achievability : Wyner-Ziv and Control of Posterior Beliefs

$$\mathbb{E}_\sigma\left[\frac{1}{n}\sum_{i=1}^{n} D\left(\mathcal{P}_\sigma(U_i|Y^n,Z^n,E_\delta^0)\middle\|\mathcal{Q}(U_i|W_i,Z_i)\right)\right]$$

$$= \frac{1}{n}\sum_{\substack{(u^n,z^n,\\ w^n,y^n)\in A_\delta}} \mathcal{P}_\sigma(u^n,z^n,w^n,y^n|E_\delta^0)\log_2\frac{1}{\prod_{i=1}^n \mathcal{Q}(u_i|w_i,z_i)} - \frac{1}{n}\sum_{i=1}^{n} H(U_i|Y^n,Z^n,E_\delta^0)$$

$$\leq H(U|W,Z)+\delta-\frac{1}{n}H(U^n|W^n,Y^n,Z^n,E_\delta^0)$$

$$\leq H(U|W,Z)-\frac{1}{n}H(U^n|W^n,Z^n,E_\delta^0)+\delta$$

$$\left[Z^n -\!\!\ominus\!\!- U^n -\!\!\ominus\!\!- W^n -\!\!\ominus\!\!- Y^n \Longrightarrow H(U^n|W^n,Z^n)=H(U^n|W^n,Y^n,Z^n)\right]$$

$$= H(U|W,Z)-\frac{1}{n}H(U^n|E_\delta^0)+\frac{1}{n}I(U^n;W^n|E_\delta^0)+\frac{1}{n}H(Z^n|W^nE_\delta^0)-\frac{1}{n}H(Z^n|U^nW^nE_\delta^0$$

$$= H(U|W,Z)-H(U)+I(U;W)+H(Z|W)-H(Z|U,W)+5\delta$$

$$= -I(U;W,Z)+I(U;W)+I(U;Z|W)+5\delta=5\delta$$

i.i.d. source, codebook size, typical sequences, i.i.d. source + $Z^n -\!\!\ominus\!\!- U^n -\!\!\ominus\!\!- W^n$

Sketch of Converse Proof :

Markov chain $Y^n -\!\!\!\!\circ\!\!\!\!- X^n -\!\!\!\!\circ\!\!\!\!- (U^n, Z^n)$ implies :

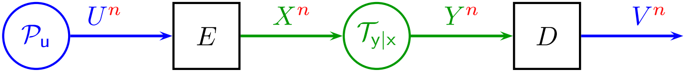
$$\begin{aligned}
0 &\leq I(X^n;Y^n) - I(U^n,Z^n;Y^n) \leq I(X^n;Y^n) - I(U^n;Y^n|Z^n) \\
&= H(Y^n) - H(Y^n|X^n) - H(U^n|Z^n) + H(U^n|Y^n,Z^n) \\
&\leq \sum_{i=1}^{n} H(Y_i) - \sum_{i=1}^{n} H(Y_i|X_i) - \sum_{i=1}^{n} H(U_i|Z_i) + \sum_{i=1}^{n} H(U_i|Y^n, Z^{-i}, Z_i) \\
&= \sum_{i=1}^{n} I(X_i;Y_i) - \sum_{i=1}^{n} I(U_i;W_i|Z_i) \\
&\leq n \cdot \max_{\mathcal{P}(x)} I(X;Y) - \sum_{i=1}^{n} I(U_i;W_i|Z_i) \\
&= n \cdot \left(\max_{\mathcal{P}(x)} I(X;Y) - I(U;W_T,T|Z) \right) \\
&= n \cdot \left(\max_{\mathcal{P}(x)} I(X;Y) - I(U;W|Z) \right).
\end{aligned}$$

Identification $W = (Y^n, Z^{-T}, T)$ satisfies both Markov chains :

$$Z -\!\!\!\!\circ\!\!\!\!- U -\!\!\!\!\circ\!\!\!\!- W \text{ and } V -\!\!\!\!\circ\!\!\!\!- (Z,W) -\!\!\!\!\circ\!\!\!\!- U.$$

Mismatched distortion functions

Joint source-channel coding scheme

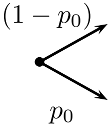


Source i.i.d.

Noisy Channel i.i.d.

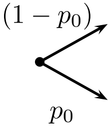
Mismatched distortion functions

$d_{\mathsf{e}}(u, v)$



	v_0	v_1
u_0	0	1
u_1	1	0

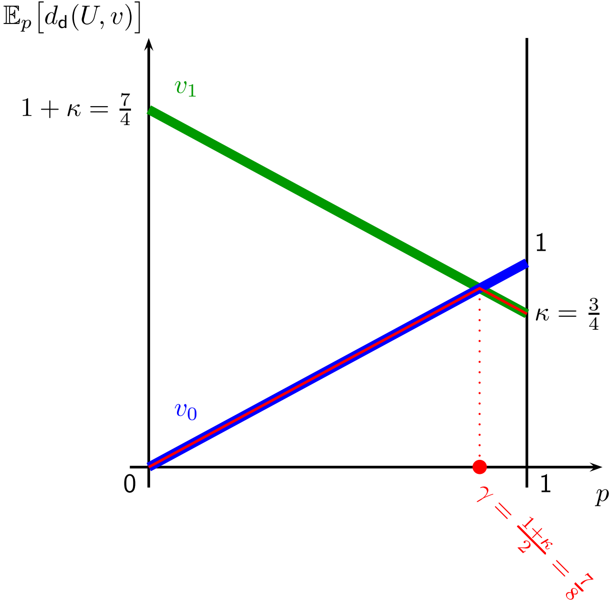
$d_{\mathsf{d}}(u, v)$



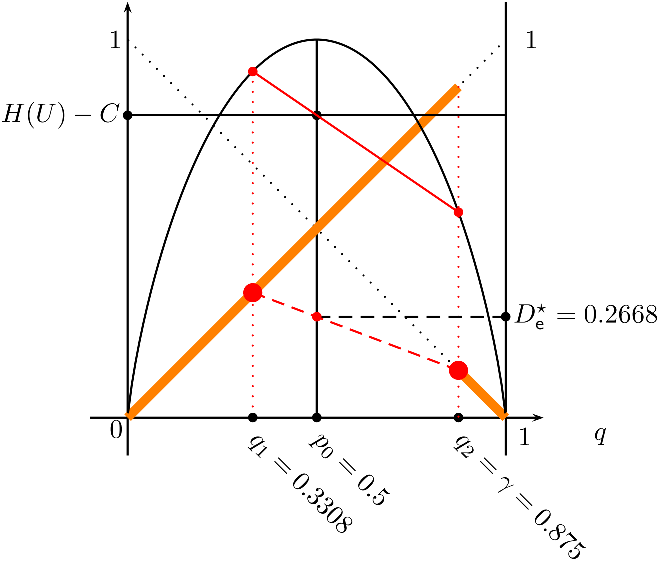
	v_0	v_1
u_0	0	$1 + \kappa$
u_1	1	κ

$\kappa \geq 0$ is an **extra cost**, e.g. energy, computing, symbol preference

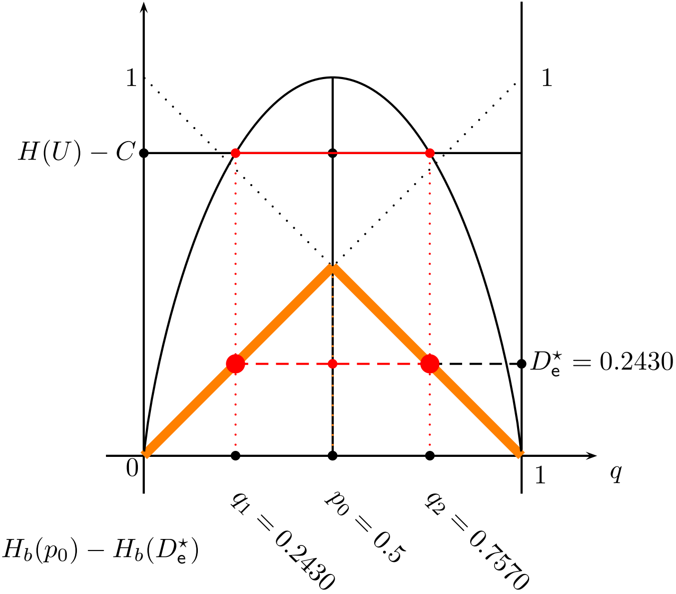
Decoder's best-reply symbol



Encoder's optimal distortion for $C = 0.2$ and $\kappa = \frac{3}{4}$



Shannon's rate-distortion function, $C = 0.2$ and $\kappa = 0$

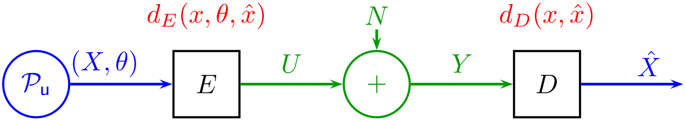


$C(D_{\mathsf{e}}^{\star}) = H_b(p_0) - H_b(D_{\mathsf{e}}^{\star})$

[Akyol Langbort Başar in Proc. IEEE 2017]

→ Information Theoretical view of Persuasion - Decentralized Stoch. Control

Distortions measures :



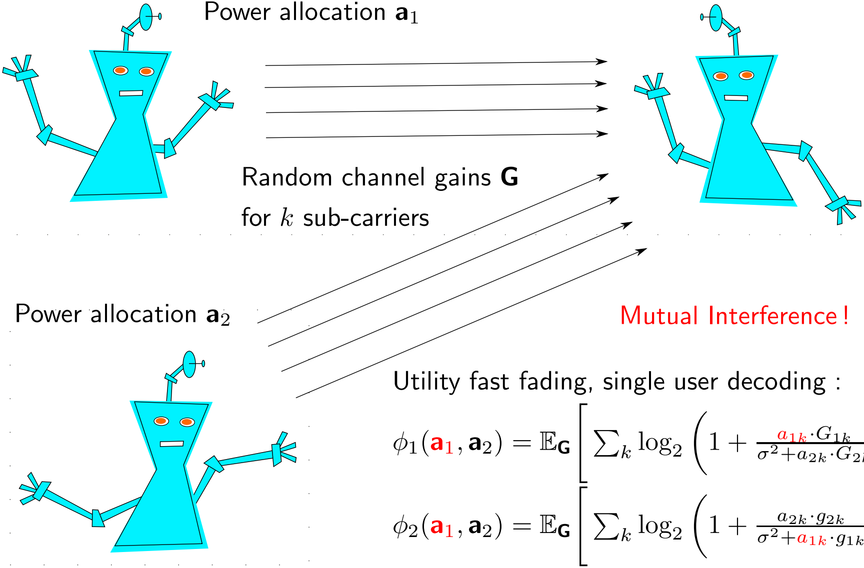
- **Gaussian** Source $(X,\theta) \sim \mathcal{N}(0, R_{X\theta})$ with $R_{X\theta} = \sigma_X^2 \begin{bmatrix} 1 & \rho \\ \rho & r \end{bmatrix}$, $\mathbb{E}[X^2] \le P_T$,
- **Quadratic** Distortions functions for encoder d_E and decoder d_D :

$$d_E(x,\theta,\hat{x}) = (x+\theta-\hat{x})^2, \qquad d_D(x,\hat{x}) = (x-\hat{x})^2$$

Theorem 7 [Akyol Langbort Başar in Proc. IEEE 2017]

$$\begin{aligned} U^\star &= \sqrt{\frac{P_T}{\sigma_X^2(1+2\alpha\rho+\alpha^2 r)}}(X+\alpha\theta), \\ \hat{X}^\star &= \mathbb{E}[X|Y], \qquad \text{with } \alpha = (-1+\sqrt{1+4(r+\rho)})/2(r+\rho) \end{aligned}$$

Example : Power Allocation Game for Parallel MAC



Power allocation $\mathbf{a}_1$

Random channel gains $\mathbf{G}$

for k sub-carriers

Power allocation $\mathbf{a}_2$

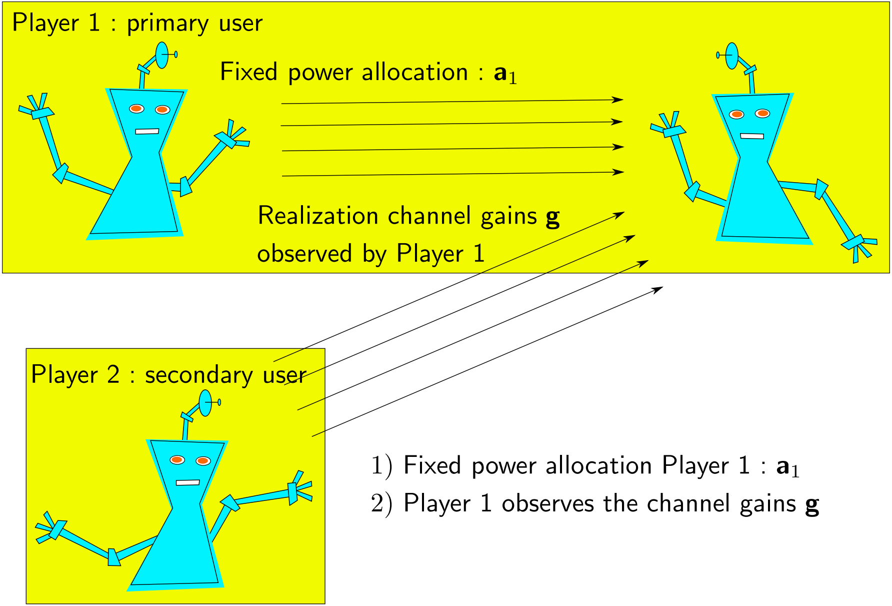
Mutual Interference !

Utility fast fading, single user decoding :

$$\phi_1(\mathbf{a}_1, \mathbf{a}_2) = \mathbb{E}_{\mathbf{G}}\left[\sum_k \log_2\left(1 + \frac{a_{1k} \cdot G_{1k}}{\sigma^2 + a_{2k} \cdot G_{2k}}\right)\right]$$

$$\phi_2(\mathbf{a}_1, \mathbf{a}_2) = \mathbb{E}_{\mathbf{G}}\left[\sum_k \log_2\left(1 + \frac{a_{2k} \cdot g_{2k}}{\sigma^2 + a_{1k} \cdot g_{1k}}\right)\right]$$

Strategic Transmission of Information



Player 1 : primary user

Fixed power allocation : $\mathbf{a}_1$

Realization channel gains $\mathbf{g}$

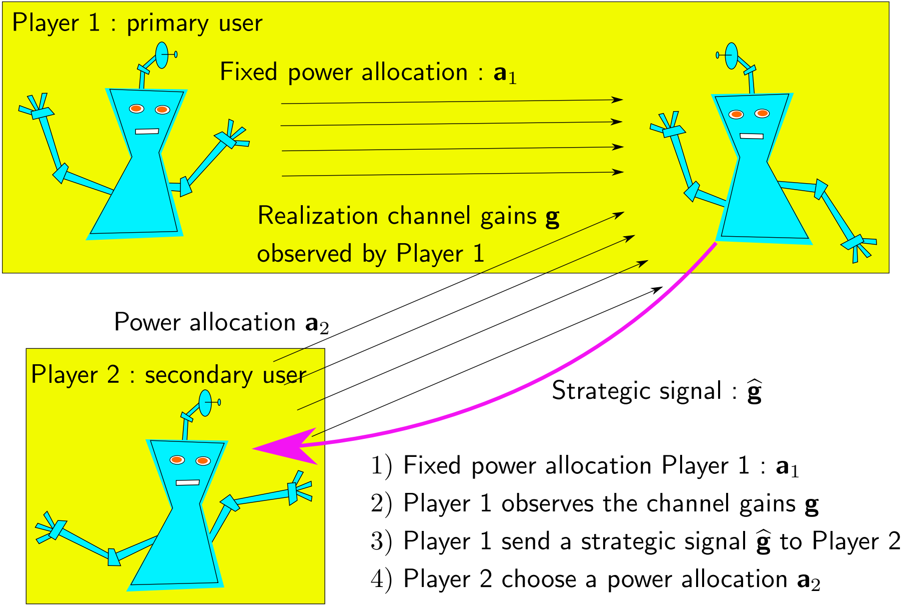
observed by Player 1

Player 2 : secondary user

$1)$ Fixed power allocation Player 1 : $\mathbf{a}_1$

$2)$ Player 1 observes the channel gains $\mathbf{g}$

Strategic Signaling of Channel Gains



Player 1 : primary user

Fixed power allocation : $\mathbf{a}_1$

Realization channel gains $\mathbf{g}$

observed by Player 1

Power allocation $\mathbf{a}_2$

Player 2 : secondary user

Strategic signal : $\widehat{\mathbf{g}}$

$1)$ Fixed power allocation Player 1 : $\mathbf{a}_1$

$2)$ Player 1 observes the channel gains $\mathbf{g}$

$3)$ Player 1 send a strategic signal $\widehat{\mathbf{g}}$ to Player 2

$4)$ Player 2 choose a power allocation $\mathbf{a}_2$

Example with Two Configurations

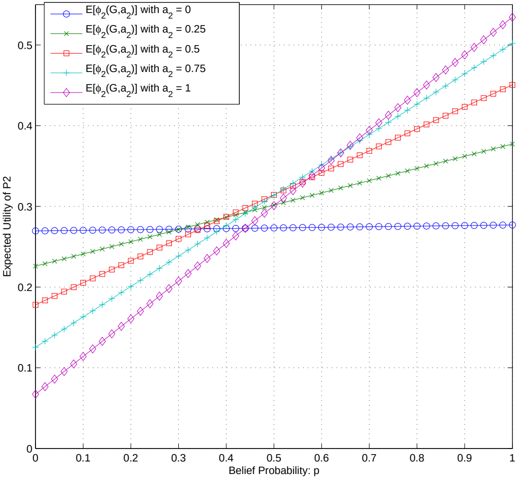
Power allocation game
with **two** parallel MACs :

- Belief proba. : $p \in [0, 1]$
- Channel gains :

	u_0	u_1
g_{11}	1.1878	0.1811
g_{12}	1.1566	1.4475
g_{21}	0.8407	0.0717
g_{22}	0.6293	0.6858

- Allocations : $a_1 = 0.16$ fixed
- $v \in \{0, 0.25, 0.5, 0.75, 1\}$

$$\phi_2(u, v) = \log_2\left(1 + \frac{v \cdot g_{21}}{\sigma^2 + a_1 \cdot g_{11}}\right) + \log_2\left(1 + \frac{(1 - v) \cdot g_{22}}{\sigma^2 + (1 - a_1) \cdot g_{12}}\right).$$



Player P_2's Best-Reply to her Belief

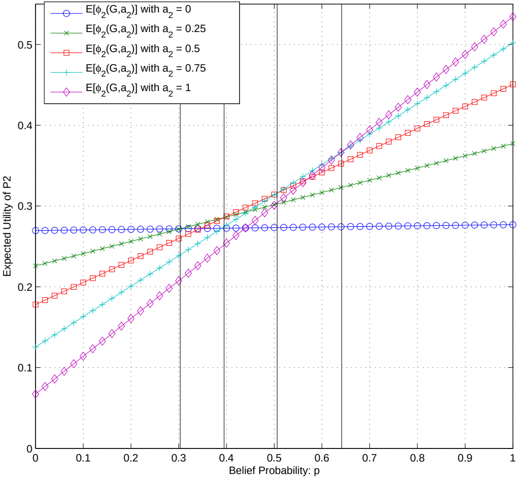
Power allocation game
with **two** parallel MACs :

- Belief proba. : $p \in [0, 1]$
- Channel gains :

	u_0	u_1
g_{11}	1.1878	0.1811
g_{12}	1.1566	1.4475
g_{21}	0.8407	0.0717
g_{22}	0.6293	0.6858

- Allocations : $a_1 = 0.16$ fixed
- $v \in \{0, 0.25, 0.5, 0.75, 1\}$

$$\phi_2(u, v) = \log_2 \left(1 + \frac{v \cdot g_{21}}{\sigma^2 + a_1 \cdot g_{11}}\right)$$
$$+ \log_2 \left(1 + \frac{(1 - v) \cdot g_{22}}{\sigma^2 + (1 - a_1) \cdot g_{12}}\right).$$



Expected Utility : $\mathbb{E}\left[\phi_2(U, v)\right] \;=\; p \cdot \phi_2(u_0, v) + (1 - p) \cdot \phi_2(u_1, v2)$

Player P_2's Best-Reply to her Belief

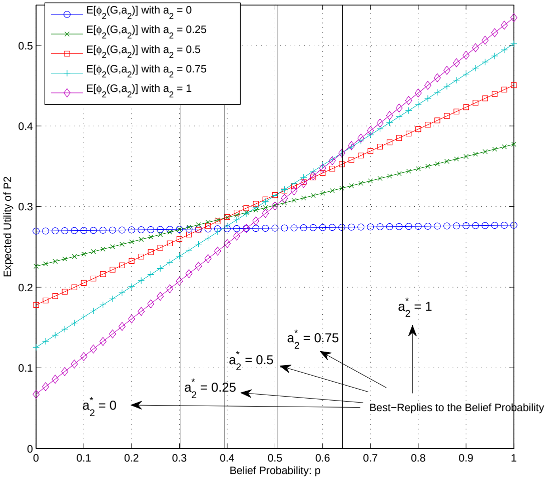
Power allocation game
with **two** parallel MACs :

- Belief proba. : $p \in [0, 1]$
- Channel gains :

	u_0	u_1
g_{11}	1.1878	0.1811
g_{12}	1.1566	1.4475
g_{21}	0.8407	0.0717
g_{22}	0.6293	0.6858

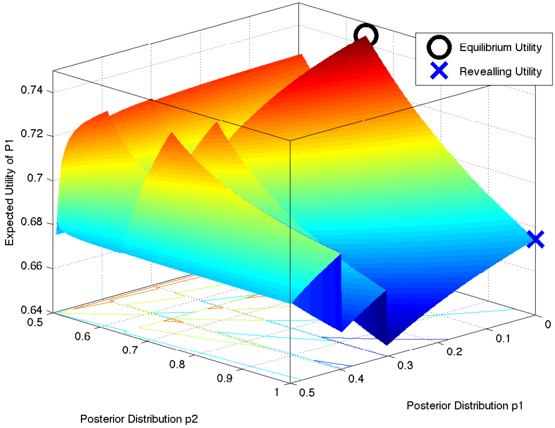
- Allocations : $a_1 = 0.16$ fixed
- $v \in \{0, 0.25, 0.5, 0.75, 1\}$

$$\phi_2(u, v) = \log_2\left(1 + \frac{v \cdot g_{21}}{\sigma^2 + a_1 \cdot g_{11}}\right) + \log_2\left(1 + \frac{(1 - v) \cdot g_{22}}{\sigma^2 + (1 - a_1) \cdot g_{12}}\right).$$



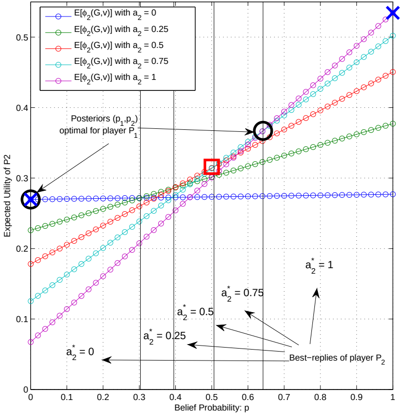
Expected Utility : $\mathbb{E}\left[\phi_2(U, v)\right] = p \cdot \phi_2(u_0, v) + (1 - p) \cdot \phi_2(u_1, v2)$

Expected Utility of Player P_1 for Posteriors (p_1, p_2)

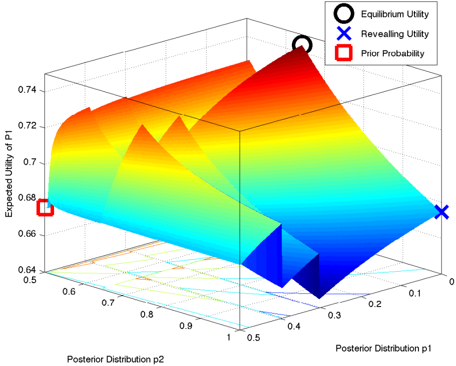


$$\mathbb{E}\left[\phi_1(U,v)\right] = \lambda\cdot\Big(p_1\cdot\phi_1(g_A,v^\star)+(1-p_1)\cdot\phi_1(g_B,v^\star)\Big)$$
$$+\quad (1-\lambda)\cdot\Big(p_2\cdot\phi_1(g_A,v^\star)+(1-p_2)\cdot\phi_1(g_B,v^\star)\Big)$$

Equilibrium Solution without Channel Noise



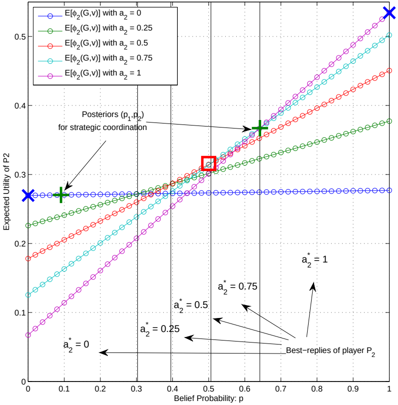
Utility of P_2

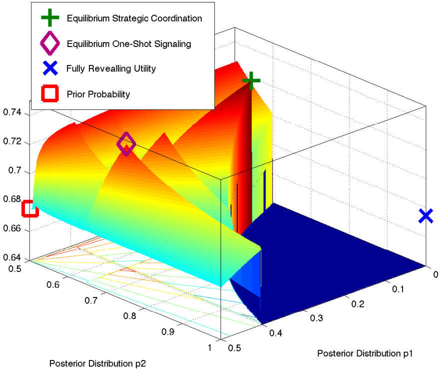


Utility of P_1

Equilibrium parameters : $(p_1, p_2) = (0, 0.6415), \quad (\alpha, \beta) = (1, 0.4424)$

Equilibrium Solution for Noisy Channel





Utility of P_2 Utility of P_1

$$\varepsilon = 0.25, \qquad (p_1, p_2) = (0.0910, 0.6420), \qquad (\alpha, \beta) = (0.9531, 0.4685)$$

Thank you !

https ://sites.google.com/site/maelletreust/

Bibliography

Emir Kamenica and Matthew Gentzkow.
Bayesian persuasion.
American Economic Review, 101 :2590 – 2615, 2011.

E. Akyol, C. Langbort, and T. Başar.
Information-theoretic approach to strategic communication as a hierarchical game.
Proceedings of the IEEE, 105(2) :205–218, 2017.

Mael Le Treust and Tristan Tomala.
Information design for strategic coordination of autonomous devices with non-aligned utilities.
IEEE Proc. of the 54th Allerton conference, Monticello, Illinois, pages 233–242, 2016.

Mael Le Treust and Tristan Tomala.
Persuasion with limited communication capacity.
draft available on ArXiv, Dec. 2017.

Mael Le Treust and Tristan Tomala.
Information-Theoretic Limits of Strategic Communication .
draft available on ArXiv, July 2018.

P.W. Cuff, H.H. Permuter, and T.M. Cover.
Coordination capacity.
Information Theory, IEEE Transactions on, 56(9) :4181–4206, 2010.

Mael Le Treust.
Joint empirical coordination of source and channel.
IEEE Transactions on Information Theory, 63(8) :5087–5114, Aug 2017.

Mael Le Treust.
Empirical coordination with two-sided state information and correlated source and state.
In *IEEE International Symposium on Information Theory (ISIT)*, 2015.

S. Sarıtaş, S. Yüksel, and S. Gezici.
Dynamic signaling games under nash and stackelberg equilibria.
In *2016 IEEE International Symposium on Information Theory (ISIT)*, pages 1631–1635, July 2016.