



HAL
open science

Information Design for Strategic Coordination of Autonomous Devices with Non-Aligned Utilities

Maël Le Treust, Tristan Tomala

► **To cite this version:**

Maël Le Treust, Tristan Tomala. Information Design for Strategic Coordination of Autonomous Devices with Non-Aligned Utilities. 54th Annual Allerton Conference on Communication, Control, and Computing, Sep 2016, Monticello, Illinois, United States. hal-01383923

HAL Id: hal-01383923

<https://hal.science/hal-01383923>

Submitted on 13 Nov 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Information Design for Strategic Coordination of Autonomous Devices with Non-Aligned Utilities

Maël Le Treust* and Tristan Tomala †

* ETIS, UMR 8051 / ENSEA, Université Cergy-Pontoise, CNRS,
6, avenue du Ponceau, 95014 Cergy-Pontoise CEDEX, FRANCE

Email: mael.le-treust@ensea.fr

† HEC Paris, GREGHEC UMR 2959

1 rue de la Libération, 78351 Jouy-en-Josas CEDEX, FRANCE

Email: tomala@hec.fr

Abstract—In this paper, we investigate the coordination of autonomous devices with non-aligned utility functions. Both encoder and decoder are considered as players, that choose the encoding and the decoding in order to maximize their long-run utility functions. The topology of the point-to-point network under investigation, suggests that the decoder implements a strategy, knowing in advance the strategy of the encoder. We characterize the encoding and decoding functions that form an equilibrium, by using empirical coordination. The equilibrium solution is related to an auxiliary game in which both players choose some conditional distributions in order to maximize their expected utilities. This problem is closely related to the literature on “Information Design” in Game Theory. We also characterize the set of posterior distributions that are compatible with a rate-limited channel between the encoder and the decoder. Finally, we provide an example of non-aligned utility functions corresponding to parallel fading multiple access channels.

I. INTRODUCTION

In this paper, we investigate the coordination of autonomous devices with non-aligned utility functions. We consider a point-to-point network, depicted in Fig. 1, with an i.i.d. information source $\mathcal{P}_u(u)$, an encoder P_1 , a memoryless channel $\mathcal{T}(y|x)$ and a decoder P_2 . The encoder and the decoder are considered as players, endowed with utility functions $\phi_1(u, v) \in \mathbb{R}$ and $\phi_2(u, v) \in \mathbb{R}$. Both utilities depend on the source symbol $u \in \mathcal{U}$ and on the action $v \in \mathcal{V}$ of player P_2 , the decoder. In the n -stage game, the players choose the optimal encoding and the decoding functions. The accumulated utilities are characterized by using the empirical coordination of the random variables (U, V) .

The problem of empirical coordination was investigated in both literatures of Game Theory [1], [2], [3], [4], [5] and Information Theory [6], [7], [8], [9], [10], [11], [12]. The objective is to characterize the set of target empirical distributions that are achievable by using a coding scheme. Optimal solutions have been characterized for lossless decoding [13], for state-dependent source and channel [14], for channel feedback [15], for the two-agent case [16]. Polar coding scheme for empirical coordination has been further investigated in [17], [18], [19], [20]. In [21], the authors measure the secrecy in

communication systems, using a rate-distortion approach that is closely related to empirical coordination. In [22], the authors investigate the connexion between the empirical coordination and the state-leakage induced by a coding scheme. Empirical coordination captures the knowledge of the transmitters, regarding the random variables they don’t observe.

The network topology of Fig. 1 suggests that the decoder P_2 implements a strategy, knowing in advance the strategy of the encoder P_1 . In contrast to the definition of the “Nash Equilibrium” [23], this strategic interaction is not simultaneous and corresponds to the “Stackelberg Equilibrium” [24]. The transmission of strategic information has attracted a lot of attention in the literature of Game Theory [25], [26]. In [27], the authors investigate the problem of “Bayesian Persuasion” in which a sender wants to persuade a receiver to change her action. The state of the nature is a random variable observed by the sender P_1 but not by the receiver P_2 . The sender applies a strategic quantification, designed in order to modify the posterior distributions of the receiver, regarding the state of the nature. The sender chooses an optimal signaling structure, knowing that the receiver implements a best-reply with respect to her posterior belief. This problem is called “Information

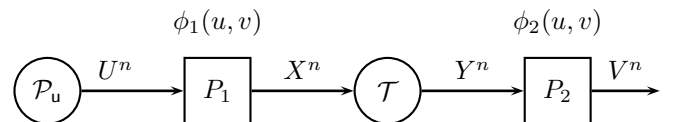


Fig. 1. Strategic Empirical Coordination: The information source is i.i.d. \mathcal{P}_u and the channel \mathcal{T} is memoryless. The encoder P_1 and the decoder P_2 are players, endowed with non-aligned utility functions $\phi_1(u, v) \in \mathbb{R}$ and $\phi_2(u, v) \in \mathbb{R}$, depending on the source U and decoder’s action V .

Design” and relies on the ‘Splitting Lemma’ in the literature on Repeated Games with Incomplete Information [28], [29], [30]. Dynamic versions of this problem were considered in [31], [32], [33], where the informed player announces her strategy in each stage. In [34], the receiver fixes her strategy before the encoder, that react knowing in advance the strategy of the decoder. Strategic communication was considered more recently in the literature of Information Theory, for Gaussian source and channel with quadratic cost functions in [35], [36] and [37].

* Maël Le Treust acknowledges financial support from INS2I CNRS through projects JCJC CoReDe 2015 and PEPS StrategicCoo 2016.

† Tristan Tomala acknowledges financial support from the HEC foundation.

In this paper, we characterize the encoding and decoding functions that form an equilibrium in the long-run game. We introduce an auxiliary game in which both players choose the conditional distributions that maximize their expected utilities. The main result is stated in Sec. II. In Sec. III, we characterize the set of posterior distributions that are compatible with a rate-limited channel. In Sec. IV, we provide an example of non-aligned utility functions corresponding to parallel fading multiple access channels. Conclusion is stated in Sec. V and the proofs of the main result are stated in App. A and B.

II. STRATEGIC EMPIRICAL COORDINATION

A. Problem Statement

We consider the problem of strategic empirical coordination depicted in Fig. 1. Notations U^n, X^n, Y^n, V^n stand for sequences of random variables of information source $u^n = (u_1, \dots, u_n) \in \mathcal{U}^n$, inputs of the channel $x^n \in \mathcal{X}^n$, outputs of the channel $y^n \in \mathcal{Y}^n$ and decoder's output $v^n \in \mathcal{V}^n$, respectively. The sets $\mathcal{U}, \mathcal{X}, \mathcal{Y}, \mathcal{V}$ have finite cardinality. The set of probability distributions over \mathcal{X} is denoted by $\Delta(\mathcal{X})$. The notation $\|\mathcal{Q} - \mathcal{P}\|_1 = \sum_{x \in \mathcal{X}} |\mathcal{Q}(x) - \mathcal{P}(x)|$ stands for the L_1 distance between the probability distributions \mathcal{Q} and \mathcal{P} . With a slight abuse of notation, we denote by $\mathcal{Q}(x) \times \mathcal{Q}(v|x)$, the product of distributions over $\Delta(\mathcal{X} \times \mathcal{V})$. Notation $Y \dashv X \dashv U$ denotes the Markov chain property corresponding to $\mathcal{P}(y|x, u) = \mathcal{P}(y|x)$ for all (u, x, y) . Player P_1 observes a sequence of source symbols $u^n \in \mathcal{U}^n$ and chooses at random a sequence of channel inputs $x^n \in \mathcal{X}^n$. Player P_2 observes a sequence of channel outputs $y^n \in \mathcal{Y}^n$ and chooses at random a sequence of actions $v^n \in \mathcal{V}^n$.

Definition II.1 (Strategies of both players)

• Player P_1 chooses a strategy σ and player P_2 chooses a strategy τ , defined as follows:

$$\sigma : \mathcal{U}^n \longrightarrow \Delta(\mathcal{X}^n), \quad (1)$$

$$\tau : \mathcal{Y}^n \longrightarrow \Delta(\mathcal{V}^n). \quad (2)$$

Both strategies (σ, τ) are stochastic.

• A pair of strategies (σ, τ) induces a joint probability distribution $\mathcal{P}_{\sigma, \tau} \in \Delta(\mathcal{U}^n \times \mathcal{X}^n \times \mathcal{Y}^n \times \mathcal{V}^n)$ over the n -sequences of symbols, defined by:

$$\prod_{i=1}^n \mathcal{P}(U_i) \times \mathcal{P}_\sigma(X^n|U^n) \times \prod_{i=1}^n \mathcal{T}(Y_i|X_i) \times \mathcal{P}_\tau(V^n|Y^n). \quad (3)$$

Definition II.2 (Expected n -stage utilities)

The utilities of the n -stage game Φ_1^n and Φ_2^n are evaluated with respect to the marginal distribution $\mathcal{P}_{\sigma, \tau}$ over the sequences (U^n, V^n) and the utility functions $\phi_1(u, v) \in \mathbb{R}$, $\phi_2(u, v) \in \mathbb{R}$.

$$\begin{aligned} \Phi_1^n(\sigma, \tau) &= \mathbb{E}_{\sigma, \tau} \left[\frac{1}{n} \sum_{i=1}^n \phi_1(U_i, V_i) \right] \\ &= \sum_{u^n, v^n} \mathcal{P}_{\sigma, \tau}(u^n, v^n) \cdot \left[\frac{1}{n} \sum_{i=1}^n \phi_1(u_i, v_i) \right], \end{aligned} \quad (4)$$

$$\Phi_2^n(\sigma, \tau) = \sum_{u^n, v^n} \mathcal{P}_{\sigma, \tau}(u^n, v^n) \cdot \left[\frac{1}{n} \sum_{i=1}^n \phi_2(u_i, v_i) \right]. \quad (5)$$

Definition II.3 (Equilibrium utilities)

We assume that player P_2 knows in advance the strategy σ of player P_1 and chooses the mapping $\sigma \mapsto \tau(\sigma)$. $(\Phi_1^*, \Phi_2^*) \in \mathbb{R}^2$ is a pair of equilibrium utilities if there exists strategies $(\sigma^*, \tau^*(\sigma))$ for both players P_1 and P_2 that satisfy:

- 1) $(\Phi_1^n(\sigma^*, \tau^*(\sigma)), \Phi_2^n(\sigma^*, \tau^*(\sigma)))$ converge to (Φ_1^*, Φ_2^*) , as $n \rightarrow +\infty$,
- 2) for all $\varepsilon > 0$, there exists a $\bar{n} \in \mathbb{N}$ such that for all $n \geq \bar{n}$, the two following equilibrium conditions are satisfied:

$$\forall \sigma, \quad \Phi_2^n(\sigma, \tau^*(\sigma)) \geq \max_{\tilde{\tau}} \Phi_2^n(\sigma, \tilde{\tau}) - \varepsilon, \quad (6)$$

$$\Phi_1^n(\sigma^*, \tau^*(\sigma^*)) \geq \max_{\tilde{\sigma}} \Phi_1^n(\tilde{\sigma}, \tau^*(\tilde{\sigma})) - \varepsilon. \quad (7)$$

Remark II.4 In Definition II.3, player P_2 has access to strategy σ of player P_1 , before choosing her strategy $\sigma \mapsto \tau(\sigma)$. This ‘‘Stackelberg Equilibrium’’ [24] hypothesis comes from the topology of the point-to-point network, Fig. 1.

B. Target Probability Distribution

In this section, we characterize the pair of equilibrium utilities (Φ_1^*, Φ_2^*) , by using a target probability distribution:

$$\mathcal{P}_u(u) \times \mathcal{Q}(v|u) \in \Delta(\mathcal{U} \times \mathcal{V}). \quad (8)$$

Intuitively, the strategies σ and τ of Definition II.1 form a coding scheme. An auxiliary random variable W is used to characterize the message $w^n \in \mathcal{W}^n$ sent by the encoder P_1 and the message $\hat{w}^n \in \mathcal{W}^n$ decoded by the decoder P_2 . The decoding is correct if $w^n = \hat{w}^n$. The objective is to control the empirical distribution $Q^n(u, w, v) \in \Delta(\mathcal{U} \times \mathcal{W} \times \mathcal{V})$ of the sequences of actions and messages (U^n, W^n, V^n) , in order to achieve the following joint distribution:

$$\mathcal{P}_u(u) \times \mathcal{Q}(w|u) \times \mathcal{Q}(v|w) \in \Delta(\mathcal{U} \times \mathcal{W} \times \mathcal{V}), \quad (9)$$

with marginals on (U, V) given by (8). The auxiliary random variable W captures the common information shared by both players P_1 and P_2 . The distribution of (9) satisfies the Markov chain: $U \dashv W \dashv V$. We introduce the sets of target joint probability distributions $\mathcal{P}_u(u) \times \mathcal{Q}(w|u) \times \mathcal{Q}(v|w) \in \Delta(\mathcal{U} \times \mathcal{W} \times \mathcal{V})$ that are achievable \mathbb{Q}_0 and that satisfy the best-reply condition \mathbb{Q}_2 for player P_2 .

Definition II.5 (Achievable target distributions)

We define the set \mathbb{Q}_0 of joint probability distributions $\mathcal{P}_u(u) \times \mathcal{Q}(w|u) \times \mathcal{Q}(v|w)$, that satisfy:

$$\mathbb{Q}_0 = \left\{ \mathcal{P}_u(u) \times \mathcal{Q}(w|u) \times \mathcal{Q}(v|w), \quad \text{s.t.} \right. \\ \left. \max_{\mathcal{P}(x)} I(X; Y) - I(W; U) \geq 0 \right\}. \quad (10)$$

The set \mathbb{Q}_0 is convex since the mutual information $I(W; U)$ is convex in $\mathcal{Q}(w|u)$, for fixed $\mathcal{P}_u(u)$.

The information constraint (10) of \mathbb{Q}_0 does not depend on the conditional distribution $\mathcal{Q}(v|w)$, but only on the product of $\mathcal{P}_u(u) \times \mathcal{Q}(w|u)$ and on the channel capacity $\max_{\mathcal{P}(x)} I(X; Y)$. It ensures that the decoder can correctly recover the sequence $\hat{W}^n = W^n$, with high probability.

Definition II.6 (Strategic compatibility for player P_2)

We define the set \mathbb{Q}_2 of joint probability distributions $\mathcal{P}_u(u) \times \mathcal{Q}(w|u) \times \mathcal{Q}(v|w)$, that are strategically compatible for P_2 .

$$\mathbb{Q}_2 = \left\{ \begin{aligned} &\mathcal{P}_u(u) \times \mathcal{Q}(w|u) \times \mathcal{Q}(v|w) \quad \text{s.t.}, \\ &\mathbb{E}_{\mathcal{Q}} \left[\phi_2(U, V) \right] \geq \mathbb{E}_{\tilde{\mathcal{Q}}} \left[\phi_2(U, V) \right], \\ &\forall \tilde{\mathcal{Q}}(u, w, v) = \mathcal{P}_u(u) \times \mathcal{Q}(w|u) \times \tilde{\mathcal{Q}}(v|w) \end{aligned} \right\} \quad (11)$$

We denote by $\text{BR}_2(\mathcal{Q}(w|u))$, the set of distributions $\mathcal{Q}(v|w)$ that are best-replies of player P_2 , for distribution $\mathcal{Q}(w|u)$. The set \mathbb{Q}_2 is convex since the expectation is linear.

For all joint probability distributions $\mathcal{P}_u(u) \times \mathcal{Q}(w|u)$, the second player P_2 can generate a symbol $v \in \mathcal{V}$, by using another conditional probability distribution $\tilde{\mathcal{Q}}(v|w)$ than the prescribed one $\mathcal{Q}(v|w)$. Definition II.6 ensures that the target distribution is optimal for player P_2 .

Definition II.7 (Set of target distributions)

We define the set \mathbb{Q} of joint distributions $\mathcal{P}_u(u) \times \mathcal{Q}(v|u)$ that satisfy the following conditions:

$$\mathbb{Q} = \left\{ \begin{aligned} &\mathcal{Q}(v|u) \quad \text{s.t.} \quad \exists W \text{ with } U \circlearrowleft W \circlearrowleft V, \\ &\sum_w \mathcal{P}_u(u) \times \mathcal{Q}(w|u) \times \mathcal{Q}(v|w) = \mathcal{P}_u(u) \times \mathcal{Q}(v|u) \\ &\text{and } \mathcal{P}_u(u) \times \mathcal{Q}(w|u) \times \mathcal{Q}(v|w) \in \mathbb{Q}_0 \cap \mathbb{Q}_2 \end{aligned} \right\}. \quad (12)$$

The set \mathbb{Q} is convex since \mathbb{Q}_0 and \mathbb{Q}_2 are convex for any W .

The definition of \mathbb{Q} involves an auxiliary random variable W that satisfies the Markov chain $U \circlearrowleft W \circlearrowleft V$ and the marginal conditions $\mathcal{P}_u(u) \times \mathcal{Q}(v|u)$. The random variable W captures the information of P_2 regarding U and the information of P_1 regarding V . The set \mathbb{Q} characterizes the conditional distributions $\mathcal{Q}(v|u)$ that are achievable and that satisfy a best-reply condition for player P_2 . The conditional distributions $\mathcal{Q}(v|u)$ outside \mathbb{Q} cannot support equilibrium utilities.

Theorem II.8 The equilibrium utility Φ_1^* of player P_1 is:

$$\Phi_1^* = \max_{\mathcal{Q}(v|u) \in \mathbb{Q}} \mathbb{E} \left[\Phi_1(U, V) \right]. \quad (13)$$

The equilibrium utility $\Phi_2^* = \mathbb{E}[\Phi_2(U, V)]$ is given by the expectation with respect to the conditional distribution $\mathcal{Q}^*(v|u) \in \mathbb{Q}$ that achieves the maximum in equation (13).

The proof of Theorem II.8 is stated in App. A and B. The n -stage game of Definition II.3 is reformulated using a one-shot game in which P_1 chooses the optimal achievable distribution $\mathcal{Q}(w|u)$, knowing that P_2 implements a best-reply $\mathcal{Q}(v|w) \in \text{BR}_2(\mathcal{Q}(w|u))$.

III. INFORMATION DESIGN: STRATEGIC COMPRESSION
A. Control of the Posterior Distributions

We consider the binary information source $U \in \{u_1, u_2\}$ with parameter $p \in [0, 1]$, $\mathcal{P}_u(u_1) = p$ and a binary auxiliary random variable $W \in \{w_1, w_2\}$. The set of conditional distributions $\mathcal{Q}(w|u)$ is represented by Fig. 2 where $\mathcal{Q}(w|u)$ involves two parameters $\alpha \in [0, 1]$ and $\beta \in [0, 1]$. The

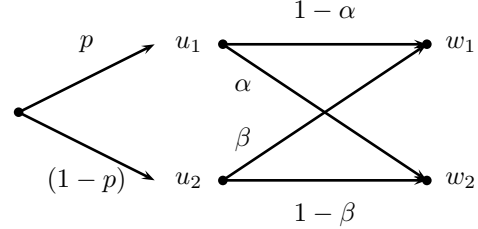


Fig. 2. Signaling $\mathcal{Q}(w|u)$ depending on $\alpha \in [0, 1]$ and $\beta \in [0, 1]$.

posterior distributions of u_1 given w_1 or w_2 are denoted by p_1 and p_2 and write:

$$\mathcal{Q}(u_1|w_1) = \frac{p \cdot (1 - \alpha)}{p \cdot (1 - \alpha) + (1 - p) \cdot \beta} = p_1, \quad (14)$$

$$\mathcal{Q}(u_1|w_2) = \frac{p \cdot \alpha}{p \cdot \alpha + (1 - p) \cdot (1 - \beta)} = p_2. \quad (15)$$

By inverting the system of equations (14) - (15), we obtain the parameters (α, β) corresponding to the pair of target posterior distributions (p_1, p_2) .

$$\alpha = \frac{p_2 \cdot (p_1 - p)}{p \cdot (p_1 - p_2)}, \quad (16)$$

$$\beta = \frac{(1 - p_1) \cdot (p - p_2)}{(1 - p) \cdot (p_1 - p_2)}. \quad (17)$$

Lemma 1 The parameters $(\alpha, \beta, p_1, p_2)$ correspond to Bernoulli distributions if and only if:

- 1) $p \notin \{0, 1\}$ and
- 2) $p_1 < p < p_2$ or $p_2 < p < p_1$.

For each pair of posterior distributions $(p_1, p_2) \in [0, p[\times]p, 1]$, there exists a pair of (α, β) such that the prior distribution p can be splitted according to $p_1 \in [0, p[$ and $p_2 \in]p, 1]$.

B. One-Shot Transmission with a Noisy Channel

We consider a binary symmetric channel $\mathcal{T}(y|x)$ with noise parameter $\varepsilon \in [0, 0.5]$, as represented by Fig. 3. The concatenation of the signaling structure $\mathcal{Q}(w|u)$ with the noisy channel $\mathcal{T}(y|x)$ can be directly expressed with parameters $\alpha \star \varepsilon \in [0, 1]$ and $\beta \star \varepsilon \in [0, 1]$, instead of α, β and ε .

$$\alpha \star \varepsilon = (1 - \alpha) \cdot \varepsilon + \alpha \cdot (1 - \varepsilon). \quad (18)$$

Proposition III.1 A pair of posterior distributions (p_1, p_2) is achievable with the noisy channel $\mathcal{T}(y|x)$ if and only if there exists (α, β) such that:

$$\alpha \star \varepsilon = \frac{p_2 \cdot (p_1 - p)}{p \cdot (p_1 - p_2)}, \quad (19)$$

$$\beta \star \varepsilon = \frac{(1 - p_1) \cdot (p - p_2)}{(1 - p) \cdot (p_1 - p_2)}. \quad (20)$$

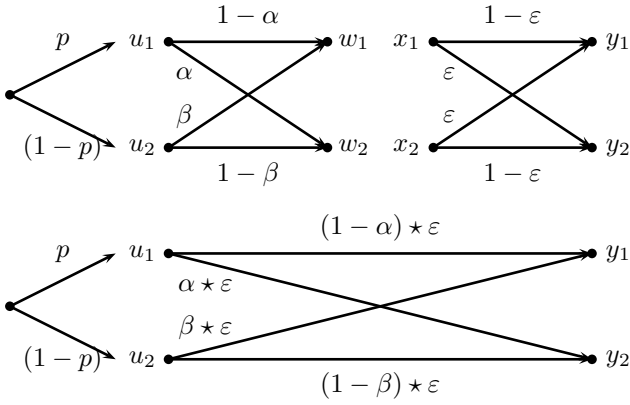


Fig. 3. The concatenation of the conditional distributions $\mathcal{Q}(w|u)$ and $\mathcal{T}(y|x)$ can be expressed using a binary symmetric channel $\mathcal{P}(y|u)$, with noise parameters $\alpha \star \epsilon \in [0, 1]$ and $\beta \star \epsilon \in [0, 1]$.

We can see on Fig. 4 that no α exists such that $\alpha \star \epsilon = \alpha \cdot (1 - 2\epsilon) + \epsilon > 1 - \epsilon$ or $\alpha \star \epsilon < \epsilon$. This imposes a restriction over the set of achievable posteriors (p_1, p_2) , that is represented by the “region of the circle”, in Fig. 5.

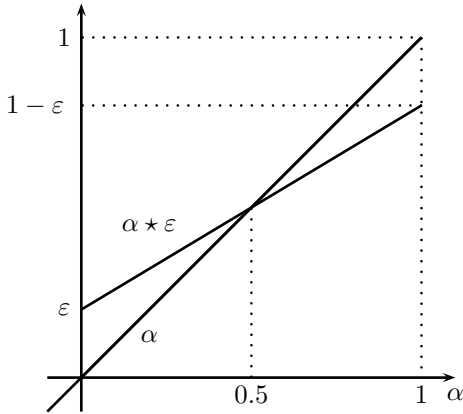


Fig. 4. Function $\alpha \star \epsilon$ depending on $\alpha \in [0, 1]$.

C. Block Transmission with a Noisy Channel

We consider the scenario represented by Fig. 1 where the symbols are encoded by blocks. Theorem II.8 states that the conditional distribution $\mathcal{Q}(w|u)$ defined with (α, β) in Fig. 2, is achievable if and only if:

$$\max_{\mathcal{P}(x)} I(X; Y) - I(U; W) \geq 0, \quad (21)$$

$$\begin{aligned} \iff & 1 - H(\epsilon) - H(\mathcal{P}(W_1)) \\ & + p \cdot H(\alpha) + (1-p) \cdot H(\beta) \geq 0, \quad (22) \\ & \text{with } \mathcal{P}(W_1) = p \cdot (1-\alpha) + (1-p) \cdot \beta. \end{aligned}$$

Fig. 5 represents three regions of posterior distributions (p_1, p_2) . The “region of the square” corresponds to posteriors (p_1, p_2) that satisfy the information constraint of the set \mathbb{Q}_0 . It includes the “region of the circle”, in which the posteriors (p_1, p_2) are achievable in one-shot and satisfy equations (19) - (20) of Proposition III.1. The posteriors (p_1, p_2) that belong to the “region of the cross”, are not achievable.

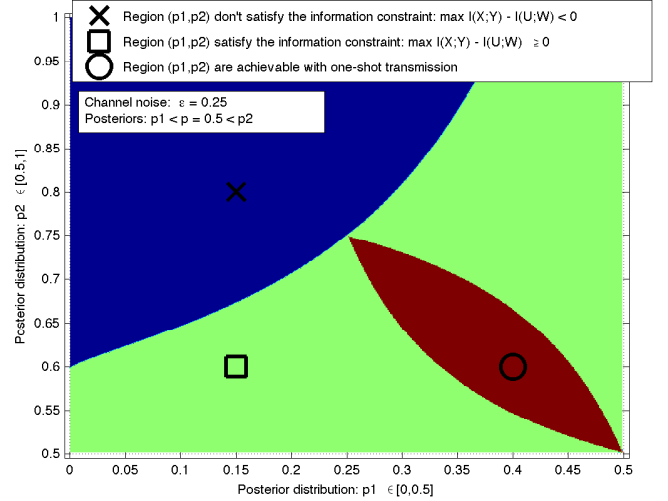


Fig. 5. Regions of achievable posteriors (p_1, p_2) , for channel noise $\epsilon = 0.25$.

IV. NON-ALIGNED UTILITIES: PARALLEL FADING MULTIPLE ACCESS CHANNELS

A. Power Allocation Game

In this section, we consider an example involving two transmitters that communicate with two base stations over parallel fading multiple access channels [38]. The transmissions are simultaneous and cause mutual interferences. Both transmitters have maximal power equal to 1 and the noise variance is $\sigma^2 = 1$. We consider two possible configurations for the random channel gains $G \in \{g_A, g_B\}$, described by the following table and chosen with probability $(p, 1-p)$ for $p = 0.5$. In this section, the channel gains are considered as an information source $U = G$.

	g_A	g_B
g_{11}	1.1878	0.1811
g_{12}	1.1566	1.4475
g_{21}	0.8407	0.0717
g_{22}	0.6293	0.6858

The notation g_{12} corresponds to the channel gain between the first transmitter and the second base station. We assume that the power allocation of the first transmitter is fixed $(a_1, 1 - a_1) = (0.16, 0.84)$. The second transmitter chooses a power allocation $(v, 1 - v)$ from the discrete set $v \in \mathcal{V} = \{0, 0.25, 0.5, 0.75, 1\}$, in order to maximize her expected utility $\mathbb{E}[\phi_2(G, v)]$.

$$\mathbb{E}[\phi_2(G, v)] = p \cdot \phi_2(g_A, v) + (1-p) \cdot \phi_2(g_B, v), \quad (23)$$

$$\begin{aligned} \phi_2(g, v) = & \log_2 \left(1 + \frac{v \cdot g_{21}}{\sigma^2 + a_1 \cdot g_{11}} \right) \\ & + \log_2 \left(1 + \frac{(1-v) \cdot g_{22}}{\sigma^2 + (1-a_1) \cdot g_{12}} \right). \quad (24) \end{aligned}$$

We consider the game between the base station P_1 and the second transmitter P_2 . The base station is informed of the realization of the channel gains $G \in \{g_A, g_B\}$ and wishes to persuade the second transmitter P_2 to choose a favorable

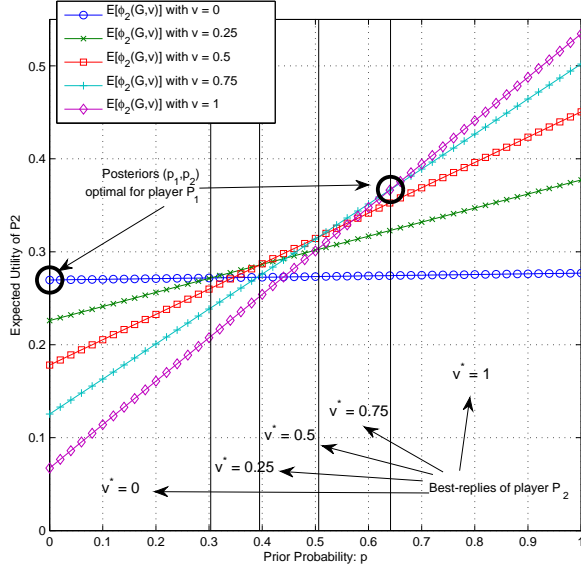


Fig. 6. Expected utility function $\mathbb{E}[\phi_2(G, v)]$ of the second transmitter P_2 , depending on her power allocation $v \in \{0, 0.25, 0.5, 0.75, 1\}$. The best-reply of P_2 is denoted by v^* and depends on her prior probability $p \in [0, 1]$.

power allocation $v \in \mathcal{V}$. The utility of the base station $\phi_1(G, v)$ is equal to the utility of the first transmitter:

$$\begin{aligned} \phi_1(g, v) &= \log_2 \left(1 + \frac{a_1 \cdot g_{11}}{\sigma^2 + v \cdot g_{12}} \right) \\ &+ \log_2 \left(1 + \frac{(1 - a_1) \cdot g_{12}}{\sigma^2 + (1 - v) \cdot g_{22}} \right). \end{aligned} \quad (25)$$

Depending on the realization of the channel gains g_A

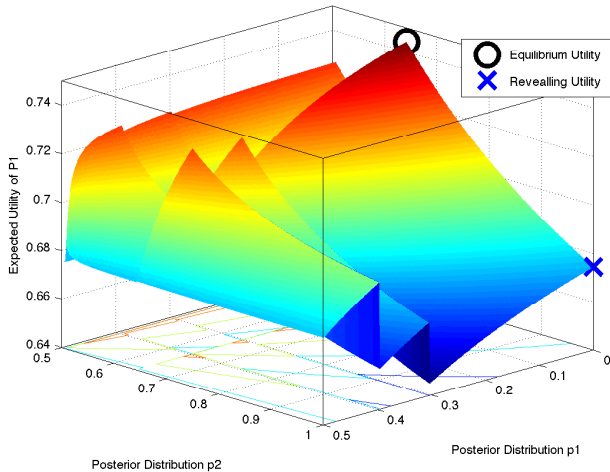


Fig. 7. Expected utility $\mathbb{E}[\phi_1(G, v)]$ depending on the pair of posterior distributions (p_1, p_2) , induced by the signaling $\mathcal{Q}(w|u)$ of Fig. 2. The discontinuities are due to the changes of best-reply v^* of P_2 , see Fig. 6.

or g_B , the base station P_1 sends a strategic signal $W \in \{w_1, w_2\}$ to the second transmitter P_2 using the signaling structure $\mathcal{Q}(w|u)$, depicted in Fig. 2. Given the strategic signal W , player P_2 chooses the power allocation $v \in$

$\{0, 0.25, 0.5, 0.75, 1\}$ that maximizes her own expected utility $\mathbb{E}[\phi_2(G, v)]$.

Fig. 6 represents the expected utility function $\mathbb{E}[\phi_2(G, v)]$ depending on the prior probability $p \in [0, 1]$, for each power allocation $v \in \{0, 0.25, 0.5, 0.75, 1\}$. The best-reply allocation of P_2 is denoted by v^* and depends on the interval to which belongs the prior probability $p \in [0, 1]$. Upon receiving symbol $w \in \mathcal{W}$ and knowing the joint distribution $\mathcal{P}_u(u) \times \mathcal{Q}(w|u)$ of Fig. 2, player P_2 implements a best-reply $\mathcal{Q}(v|w) \in \text{BR}_2(\mathcal{Q}(w|u))$. Theorem II.8 guarantees that the i.i.d. distribution $\mathcal{Q}(v|w)$ induces an ε -best-reply in the long-run game.

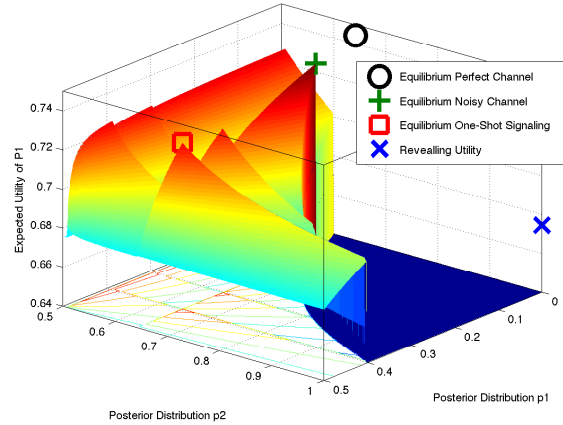


Fig. 8. The “green plus” represents the equilibrium utility $\Phi_1^* = \max_{\mathcal{Q}(v|u) \in \mathcal{Q}} \mathbb{E}[\phi_1(G, v)]$ characterized by Theorem II.8, for channel noise $\varepsilon = 0.25$.

The base station P_1 already knows that P_2 implements a best-reply. It chooses accordingly the signaling structure $\mathcal{Q}(w|u)$ that induces the more favorable response of P_2 . Fig 7 shows the expected utility $\mathbb{E}[\phi_1(G, v)]$ of P_1 , depending on the pair of posterior distributions (p_1, p_2) , induced by the signaling structure $\mathcal{Q}(w|u)$. The discontinuities are due to the changes of best-reply v^* of P_2 , according to the posterior distributions (p_1, p_2) , see Fig. 6. The equilibrium utility $\mathbb{E}[\phi_1] \simeq 0.74$ of P_1 is represented by the “black circle” and corresponds to the posterior distributions $(p_1, p_2) = (0, 0.6415)$, with $(\alpha, \beta) = (1, 0.4424)$. This pair of optimal posteriors for P_1 is also represented by the “black circles”, on Fig. 6. This equilibrium utility provides 9.1% of improvement compared to the revealing strategy (“blue cross” $\mathbb{E}[\phi_1] \simeq 0.67$), *i.e.* when the channel gains are revealed to the second transmitter P_2 , with $\alpha = \beta = 0$. The authors would like to thank Claudio Weidmann for fruitful discussions regarding this section.

B. Rate-Limited Channel between Players P_1 and P_2

We assume that the channel between the base station P_1 and the second transmitter P_2 is rate-limited, *i.e.* there is a noisy channel $\mathcal{T}(y|x)$ between P_1 and P_2 , as depicted in Fig. 1. The rate-limited constraint reduces the set of posterior distributions (p_1, p_2) , see Fig. 5. We determine the equilibrium solutions for a binary symmetric channel with parameter $\varepsilon = 0.25$, for one-shot and block transmission, as depicted in Fig. 9 and

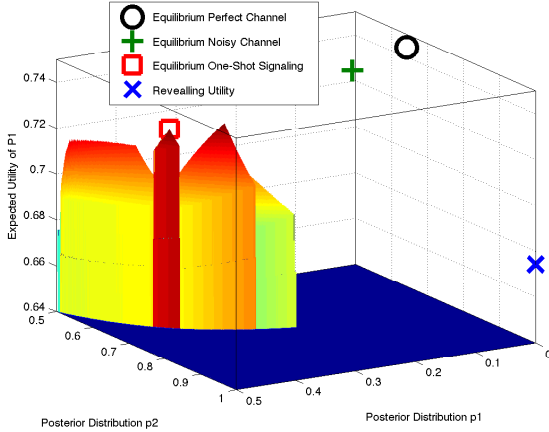


Fig. 9. The “red square” represents the equilibrium utility for one-shot transmission over a noisy channel that satisfies the conditions of Prop. III.1.

Fig. 8. These figures correspond to the utility $\mathbb{E}[\phi_1(G, v)]$ of Fig. 7, restricted to the “region of the circle” and “region of the square”, in Fig. 5. The equilibrium utilities correspond respectively to $\mathbb{E}[\phi_1(G, v)] \simeq 0.72$ and $\mathbb{E}[\phi_1(G, v)] \simeq 0.73$. In both cases, the revealing strategy is not achievable since it is not compatible with the rate-limitation imposed by the noisy channel $\mathcal{T}(y|x)$.

V. CONCLUSION

We investigate the strategic coordination of an encoder and a decoder, endowed with non-aligned utility functions. We characterize the encoding and decoding functions that form an equilibrium, by using empirical coordination. The equilibrium solution is related to an auxiliary game in which both players choose the conditional distributions in order to maximize their expected utilities. We characterize the set of posterior distributions that are compatible with a rate-limited channel between the encoder and the decoder and we provide an example of non-aligned utility functions corresponding to parallel fading multiple access channels.

APPENDIX A

PROOF OF THE UPPER BOUND FOR THEOREM II.8

In this section, we consider that the strategies (σ^*, τ^*) form an equilibrium. $Q^n(u, v)$ denotes the empirical distribution of the sequences (u^n, v^n) .

$$\begin{aligned}
 & \Phi_1^n(\sigma^*, \tau^*) \\
 &= \sum_{\substack{u^n, x^n, \\ y^n, v^n}} \mathcal{P}_{\sigma^* \tau^*}(u^n, x^n, y^n, v^n) \cdot \frac{1}{n} \sum_{i=1}^n \phi_1(u_i, v_i) \\
 &= \sum_{\substack{u^n, x^n, \\ y^n, v^n}} \mathcal{P}_{\sigma^* \tau^*}(u^n, x^n, y^n, v^n) \cdot \sum_{u, v} Q^n(u, v) \cdot \phi_1(u, v) \\
 &= \sum_{u, v} \mathbb{E}_{\sigma^*, \tau^*} \left[Q^n(u, v) \right] \cdot \phi_1(u, v) \quad (26)
 \end{aligned}$$

We denote by $Q^*(u, v)$, the expected empirical distribution $\mathbb{E}_{\sigma^*, \tau^*} [Q^n(u, v)] = Q^*(u, v)$ corresponding to the strategies (σ^*, τ^*) . The marginal distribution over U satisfies

$\sum_{v \in \mathcal{V}} \mathbb{E}_{\sigma^*, \tau^*} [Q^n(u, v)] = \mathcal{P}_u(u)$. We denote by \mathbb{Q}^c the complementary of the set \mathbb{Q} . We show that the conditional distribution $Q^*(v|u) \in \mathbb{Q}$ should belong to the set \mathbb{Q} of Definition II.7. Player P_1 cannot choose a distribution $Q(v|u) \in \mathbb{Q}^c$ that lie outside \mathbb{Q} , in order to maximize her long-run utility:

$$\Phi_1^n(\sigma^*, \tau^*) \leq \max_{Q(v|u) \in \mathbb{Q}} \mathbb{E} \left[\Phi_1(U, V) \right]. \quad (27)$$

A. Information Constraint

We show that there exists an auxiliary random variable W that satisfies the Markov chain $U \dashv W \dashv V$ and the information constraint $\max_{\mathcal{P}(x)} I(X; Y) - I(U; W) \geq 0$ of Definition II.5.

$$0 \leq I(X^n; Y^n) - I(U^n; Y^n) \quad (28)$$

$$\begin{aligned}
 & \leq \sum_{i=1}^n H(Y_i) - \sum_{i=1}^n H(Y_i | X_i) \\
 & \quad - \sum_{i=1}^n H(U_i) + \sum_{i=1}^n H(U_i | Y^n, U^{i-1}) \quad (29)
 \end{aligned}$$

$$= \sum_{i=1}^n I(X_i; Y_i) - \sum_{i=1}^n I(U_i; W_i) \quad (30)$$

$$\leq n \cdot \max_{\mathcal{P}(x)} I(X; Y) - \sum_{i=1}^n I(U_i; W_i) \quad (31)$$

$$= n \cdot \left(\max_{\mathcal{P}(x)} I(X; Y) - I(U; W_T, T) \right) \quad (32)$$

$$= n \cdot \left(\max_{\mathcal{P}(x)} I(X; Y) - I(U; W) \right). \quad (33)$$

Equation (28) comes from the Markov chain $Y^n \dashv X^n \dashv U^n$. Equation (29) comes from the memoryless property of the channel and the i.i.d. property of the source.

Equation (30) comes from the identification of the auxiliary random variable $W_i = (Y^n, U^{i-1})$ that satisfies the Markov chain of the set \mathbb{Q} for all $i \in \{1, \dots, n\}$:

$$U_i \dashv W_i \dashv V_i. \quad (34)$$

Equation (31) comes from taking the maximum over $\mathcal{P}(x)$. Equation (32) comes from the introduction of the uniform random variable T over the indices $\{1, \dots, n\}$, the independence between T and U_T , that implies $I(T; U_T) = 0$ and the i.i.d. property of the source $U_T = U$.

Equation (33) comes from the identification of the auxiliary random variable $W = (W_T, T) = (Y^n, U^{T-1}, T)$ and the Markov chain property:

$$U \dashv (W_T, T) \dashv V_T. \quad (35)$$

Hence, the joint distribution $\mathcal{P}_u(u) \times Q^*(w|u) \times Q^*(v|w)$ induced by the auxiliary random variable $W = (W_T, T) = (Y^n, U^{T-1}, T)$ satisfies:

- 1) the Markov chain condition $U \dashv W \dashv V$,
- 2) the marginal condition $\sum_w \mathcal{P}_u(u) \times Q^*(w|u) \times Q^*(v|w) = \mathcal{P}_u(u) \times Q^*(v|u)$ given by equation (26),
- 3) the positive information constraint (33).

In App. A-B, we prove that the distribution $\mathcal{P}_u(u) \times Q^*(w|u) \times Q^*(v|w)$ belongs to the set \mathbb{Q}_2 , i.e. it satisfies the best-reply condition for P_2 .

B. Profitable Deviation of the Second Player

In this section, we consider a target distribution $\mathcal{P}_u(u) \times \mathcal{Q}(w|u) \times \mathcal{Q}(v|w) \in \mathbb{Q}_0 \cap \mathbb{Q}_2^c$, that lies outside \mathbb{Q}_2 and that is achievable, i.e. $\max_{\mathcal{P}(x)} I(X; Y) - I(W; U) \geq 0$. Hence, there exists a distribution $\tilde{\mathcal{Q}}(v|w) \neq \mathcal{Q}(v|w)$ that increases the utility of player P_2 :

$$\begin{aligned} & \mathbb{E}_{\tilde{\mathcal{Q}}} \left[\phi_2(U, V) \right] > \mathbb{E}_{\mathcal{Q}} \left[\phi_2(U, V) \right], \\ \iff & \sum_{u, w, v} \mathcal{P}_u(u) \times \mathcal{Q}(w|u) \times \tilde{\mathcal{Q}}(v|w) \times \phi_2(u, v) \\ & > \sum_{u, w, v} \mathcal{P}_u(u) \times \mathcal{Q}(w|u) \times \mathcal{Q}(v|w) \times \phi_2(u, v). \end{aligned} \quad (36)$$

Since the target distribution $\mathcal{P}_u(u) \times \mathcal{Q}(w|u) \times \mathcal{Q}(v|w) \in \mathbb{Q}_0$ is achievable, the second player P_2 correctly decodes the sequence W^n , with high probability. There exists a deviating strategy $\tau \neq \tau^*$ based on the i.i.d. distribution $\tilde{\mathcal{Q}}(v|w)$ that is profitable for P_2 . Hence there is a $\delta > 0$ such that the equilibrium condition is not satisfied:

$$\Phi_2^n(\sigma^*, \tau) > \Phi_2^* + \delta. \quad (37)$$

APPENDIX B

PROOF OF THE LOWER BOUND FOR THEOREM II.8

In this section, we provide a coding scheme (σ^*, τ^*) that satisfies both equilibrium conditions of Definition II.3:

$$\Phi_2^n(\sigma^*, \tau^*(\sigma^*)) \geq \max_{\tilde{\tau}} \Phi_2^n(\sigma^*, \tilde{\tau}) - \varepsilon, \quad (38)$$

$$\Phi_1^n(\sigma^*, \tau^*(\sigma^*)) \geq \max_{\tilde{\sigma}} \Phi_1^n(\tilde{\sigma}, \tau^*(\tilde{\sigma})) - \varepsilon. \quad (39)$$

A. Separated Source-Channel Coding

We consider the target joint probability distribution $\mathcal{Q}(u, w, v) = \mathcal{P}_u(u) \times \mathcal{Q}(w|u) \times \mathcal{Q}(v|w) \in \mathbb{Q}_0 \cap \mathbb{Q}_2$ that corresponds to the optimal utility Φ_1^* of equation (13). By definition of \mathbb{Q}_0 , the target distribution $\mathcal{Q}(u, w, v)$ satisfies the information constraint:

$$\max_{\mathcal{P}(x)} I(X; Y) - I(W; U) \geq 0. \quad (40)$$

In this section, we assume that equation (40) is satisfied with strict inequality (41). The case of equality in the information constraint will be treated in App. B-B.

$$\max_{\mathcal{P}(x)} I(X; Y) - I(W; U) > 0. \quad (41)$$

Inequality (41) implies that there exists a small parameter $\delta > 0$ and a rate $R \geq 0$, such that:

$$R \geq I(W; U) + \delta, \quad (42)$$

$$R = \max_{\mathcal{P}(x)} I(X; Y) - \delta. \quad (43)$$

We define a code $c = (f, g) \in \mathcal{C}(n)$ using the following encoding and decoding functions, that involve sequence W^n :

$$f : \mathcal{U}^n \longrightarrow \mathcal{X}^n \times \mathcal{W}^n, \quad (44)$$

$$g : \mathcal{Y}^n \longrightarrow \mathcal{W}^n \times \mathcal{V}^n. \quad (45)$$

We show the existence of an optimal code $c^* \in \mathcal{C}(n)$ such that the empirical distribution $Q^n(u, w, v)$ of symbols

(U^n, W^n, V^n) , is close to the target distribution $\mathcal{Q}(u, w, v) = \mathcal{P}_u(u) \times \mathcal{Q}(w|u) \times \mathcal{Q}(v|w)$, with large probability. More precisely, we prove that for all $\varepsilon > 0$, there exists an $\bar{n} \in \mathbb{N}$, such that for all $n \geq \bar{n}$, there exists a code $c^* \in \mathcal{C}(n)$ that satisfies:

$$\mathcal{P}_{c^*} \left(\left\| Q^n(u, w, v) - \mathcal{Q}(u, w, v) \right\|_1 \geq \varepsilon \right) \leq \varepsilon. \quad (46)$$

The parameter $\varepsilon > 0$ is involved in both the definition of the typical sequences and the upper bound of the error probability.

- *Random codebook.* We generate $|\mathcal{M}| = 2^{nR}$ sequences $W^n(m)$ and $X^n(m)$, drawn from the marginal i.i.d. probability distributions $\mathcal{Q}_w^{\times n}$ and $\mathcal{P}_x^{\times n}$ with index $m \in \mathcal{M}$.
- *Encoding function.* The encoder observes the sequence of symbols of source $U^n \in \mathcal{U}^n$. It finds an index $m \in \mathcal{M}$ such that the sequences $(U^n, W^n(m)) \in A_\varepsilon^{*n}(\mathcal{Q})$ are jointly typical. The encoder sends the sequence $X^n(m)$ corresponding to the index $m \in \mathcal{M}$.
- *Decoding function.* The decoder observes the sequence of channel output $Y^n \in \mathcal{Y}^n$. It returns an index $\hat{m} \in \mathcal{M}$ such that the sequences $(Y^n, X^n(\hat{m})) \in A_\varepsilon^{*n}(\mathcal{Q})$ are jointly typical. It deduces the sequence $W^n(\hat{m})$ and returns V^n , drawn from the conditional probability distribution $\mathcal{Q}_{v|w}^{\times n}$ depending on $W^n(\hat{m})$.
- *Error Event.* An error occurs in the coding process if: 1) the indexes $m \in \mathcal{M}$ and $\hat{m} \in \mathcal{M}$ are not equal, not unique or does not exist, 2) the sequences of symbols $(U^n, W^n, V^n) \notin A_\varepsilon^{*n}(\mathcal{Q})$ are not jointly typical.

Expected error probability. We introduce the parameter $\varepsilon_1 > 0$, in order to provide an upper bound on the expected error probability. For all $\varepsilon_1 > 0$ there exists an $\bar{n} \in \mathbb{N}$ such that for all $n \geq \bar{n}$, the expected probability of the following error events are bounded by ε_1 :

$$\mathbb{E}_c \left[\mathcal{P} \left(\forall m \in \mathcal{M}, (U^n, W^n(m)) \notin A_\varepsilon^{*n}(\mathcal{Q}) \right) \right] \leq \varepsilon_1, \quad (47)$$

$$\mathbb{E}_c \left[\mathcal{P} \left(\exists m' \neq m, (Y^n, X^n(m')) \in A_\varepsilon^{*n}(\mathcal{Q}) \right) \right] \leq \varepsilon_1, \quad (48)$$

$$\mathbb{E}_c \left[\mathcal{P} \left((U^n, W^n(m), V^n) \notin A_\varepsilon^{*n}(\mathcal{Q}) \right) \right] \leq \varepsilon_1. \quad (49)$$

(47) comes from (42) and [39, pp. 208, Covering Lemma].

(48) comes from (43) and [39, pp. 46, Packing Lemma].

(49) comes from the properties of typical sequences, stated in [39, pp. 27], and from (47) and (48).

This proves that for all $\varepsilon_1 > 0$, there exists a $\bar{n} \in \mathbb{N}$ such that for all $n \geq \bar{n}$, there exists a code $c^* = (f^*, g^*) \in \mathcal{C}(n)$ such that:

$$\begin{aligned} & \mathcal{P}_{c^*} \left(\left\| Q^n(u, w, v) - \mathcal{Q}(u, w, v) \right\|_1 \geq \varepsilon \right) \\ & = \mathcal{P}_{c^*} \left((U^n, W^n(m), V^n) \notin A_\varepsilon^{*n}(\mathcal{Q}) \right) \leq \varepsilon_1. \end{aligned} \quad (50)$$

We denote by (σ^*, τ^*) , the strategies of P_1 and P_2 corresponding to the coding scheme $c^* = (f^*, g^*) \in \mathcal{C}(n)$ and we denote

by $\bar{\phi}_1 = \max_{u,v} |\phi_1(u,v)|$, the maximal utility of P_1 .

$$\begin{aligned} & \left| \Phi_1^n(\sigma^*, \tau^*) - \Phi_1^* \right| \\ &= \left| \sum_{u,v} \phi_1(u,v) \times \left(\mathbb{E}[Q^n(u,v)] - Q(u,v) \right) \right| \\ &\leq \bar{\phi}_1 \cdot \sum_{u,v} \left| \mathbb{E}[Q^n(u,v)] - Q(u,v) \right| \leq \bar{\phi}_1 \cdot \varepsilon. \end{aligned}$$

Hence the pair of utilities $(\Phi_1^n(\sigma^*, \tau^*), \Phi_2^n(\sigma^*, \tau^*))$ converges to the utilities (Φ_1^*, Φ_2^*) , as $n \rightarrow +\infty$.

B. Equality in the Information Constraint

We consider a target distribution $\mathcal{P}_u(u) \times \mathcal{Q}(w|u) \times \mathcal{Q}(v|w)$ with equality in the information constraint:

$$\max_{\mathcal{P}(x)} I(X; Y) - I(W; U) = 0. \quad (51)$$

First case: the channel capacity is strictly positive.

$$\max_{\mathcal{P}(x)} I(X; Y) > 0. \quad (52)$$

We consider an auxiliary distribution $\tilde{\mathcal{Q}}(w|u) = \tilde{\mathcal{Q}}(w)$ such that W is independent of U and we denote by $I_{\tilde{\mathcal{Q}}}(W; U) = 0$, the corresponding mutual information. The information constraint for $\mathcal{P}_u(u) \times \tilde{\mathcal{Q}}(w|u)$ is strictly positive:

$$\max_{\mathcal{P}(x)} I(X; Y) - I_{\tilde{\mathcal{Q}}}(W; U) = \max_{\mathcal{P}(x)} I(X; Y) > 0. \quad (53)$$

We construct a sequence $\{\mathcal{Q}^k(u, w, v)\}_{k \in \mathbb{N}^*}$ of convex combination between $\mathcal{Q}(u, w, v)$ and $\mathcal{P}_u(u) \times \tilde{\mathcal{Q}}(w|u) \times \mathcal{Q}(v|w)$:

$$\begin{aligned} \mathcal{Q}^k(u, w, v) &= \frac{1}{k} \left((k-1) \cdot \mathcal{Q}(u, w, v) \right. \\ &\quad \left. + \mathcal{P}_u(u) \times \tilde{\mathcal{Q}}(w) \times \mathcal{Q}(v|w) \right). \quad (54) \end{aligned}$$

The information constraint corresponding to $\mathcal{Q}^k(u, w, v)$ is strictly positive, for all $k \in \mathbb{N}^*$:

$$\begin{aligned} & \max_{\mathcal{P}(x)} I(X; Y) - I_{\mathcal{Q}^k}(W; U) \\ &\geq \frac{1}{n} \cdot \left((n-1) \cdot \left(\max_{\mathcal{P}(x)} I(X; Y) - I_{\mathcal{Q}}(W; U) \right) \right. \\ &\quad \left. + \left(\max_{\mathcal{P}(x)} I(X; Y) - I_{\tilde{\mathcal{Q}}}(W; U) \right) \right) \quad (55) \end{aligned}$$

$$\geq \frac{1}{n} \cdot \left(\max_{\mathcal{P}(x)} I(X; Y) - I_{\tilde{\mathcal{Q}}}(W; U) \right) > 0. \quad (56)$$

Then, for all $k \in \mathbb{N}^*$ the distribution $\mathcal{Q}^k(u, w, v)$ is achievable by using the coding scheme stated in Sec.B-A and converges to the target distribution $\mathcal{Q}(u, w, v)$, as k goes to $+\infty$. This proves that the limit distribution $\mathcal{Q}(u, w, v)$ is achievable.

Second case: the channel capacity is equal to zero. This implies that the random variables U and W of the target distribution $\mathcal{P}_u(u) \times \mathcal{Q}(w|u) \times \mathcal{Q}(v|w)$ are independent, hence the target distribution $\mathcal{Q}(u, w, v)$ decomposes like:

$$\mathcal{P}_u(u) \times \mathcal{Q}(w|u) \times \mathcal{Q}(v|w) = \mathcal{P}_u(u) \times \mathcal{Q}(w) \times \mathcal{Q}(v|w). \quad (57)$$

This target distribution (57) is achievable by the decoder generating (W^n, V^n) with the i.i.d. distribution $\mathcal{Q}(w) \times \mathcal{Q}(v|w)$.

C. Unilateral Deviation of the Second Player

The target joint distribution $\mathcal{Q}(u, w, v) = \mathcal{P}_u(u) \times \mathcal{Q}(w|u) \times \mathcal{Q}(v|w)$ of Definition II.7, satisfies the information constraint of \mathbb{Q}_0 and the strategic compatibility condition of \mathbb{Q}_2 . We consider the coding and decoding functions $c^* = (f^*, g^*)$ presented in App. B-A, for distribution $\mathcal{Q}(u, w, v)$. In this section, we prove that the decoding function g^* combined with a symbol-best-reply $\mathcal{Q}(v|w)$ is ε -optimal for P_2 . We introduce the random event of error $E \in \{0, 1\}$ defined by:

$$E = \begin{cases} 0 & \text{if } (U^n, W^n, V^n) \in A_{\varepsilon}^{*n}(\mathcal{Q}) \text{ and } \hat{M} = M, \\ 1 & \text{if } (U^n, W^n, V^n) \notin A_{\varepsilon}^{*n}(\mathcal{Q}) \text{ or } \hat{M} \neq M. \end{cases}$$

The pair of strategies $c^* = (f^*, g^*) \in \mathcal{C}(n)$, stated in App. B-A, induces a small error probability $\mathcal{P}(E = 1|c^*) \leq \varepsilon$. The expected utility of P_2 is upper bounded by:

$$\begin{aligned} & \sum_{\substack{u^n, x^n, y^n, v^n \\ w^n, E}} \mathcal{P}(u^n, x^n, y^n, w^n, v^n, E|c^*) \cdot \frac{1}{n} \sum_{i=1}^n \phi_2(u_i, v_i) \\ &= \mathcal{P}(E = 0|c^*) \sum_{\substack{u^n, x^n, y^n, \\ v^n, w^n}} \mathcal{P}(u^n, x^n, y^n, w^n, v^n|c^*, E = 0) \\ &\quad \times \frac{1}{n} \sum_{i=1}^n \phi_2(u_i, v_i) \\ &+ \mathcal{P}(E = 1|c^*) \sum_{\substack{u^n, x^n, y^n, \\ v^n, w^n}} \mathcal{P}(u^n, x^n, y^n, w^n, v^n|c^*, E = 1) \\ &\quad \times \frac{1}{n} \sum_{i=1}^n \phi_2(u_i, v_i) \quad (65) \end{aligned}$$

$$\begin{aligned} & \leq \sum_{\substack{u^n, x^n, y^n, \\ v^n, w^n}} \mathcal{P}(u^n, x^n, y^n, w^n, v^n|c^*, E = 0) \\ &\quad \times \frac{1}{n} \sum_{i=1}^n \phi_2(u_i, v_i) + \mathcal{P}(E = 1|c^*) \times \bar{\phi}_2. \quad (66) \end{aligned}$$

We denote by $\bar{\phi}_2 = \max_{u,v} |\phi_2(u,v)|$, the maximal utility of P_2 . In the following, we assume that P_2 chooses the optimal sequence V^n based on her observation Y^n , on the knowledge of the code c^* , on the hypothesis that there is no errors $E = 0$ and on the decoded sequence W^n . We prove that the decoding function g^* presented in App. B-A, is an ε -best-reply for P_2 .

Equation (58) comes from the hypothesis $E = 0$ of correct decoding of the sequence W^n . The decoder maximizes over $\mathcal{P}(v^n|w^n, y^n, c^*, E = 0)$ instead of $\mathcal{P}(v^n|y^n, c^*, E = 0)$.

Equation (59) comes from the Markov chain $U^n \circlearrowleft W^n \circlearrowleft X^n \circlearrowleft Y^n$ of the coding process c^* stated in App. B-A, that induces the following equality $\mathcal{P}(u^n|w^n, x^n, y^n, c^*, E = 0) = \mathcal{P}(u^n|w^n, c^*, E = 0)$.

Equation (60) comes from taking the sum over (x^n, y^n) and removing the sequence y^n from $\mathcal{P}(v^n|w^n, c^*, E = 0)$, since y^n is not involved in the criteria $\sum_{i=1}^n \phi_2(u_i, v_i)$.

Equation (61) is due to Lemmas 2 and 4, stated in App. B-D. Equation (62) comes from Lemma 3, stated in App. B-D.

Equation (63) comes from the hypothesis $E = 0$ that implies the empirical distribution $Q^n(w)$ of sequences w^n is close to the target $\mathcal{Q}(w)$.

Equation (64) comes from the definition of the set

$$\begin{aligned}
& \max_{\mathcal{P}(v^n|y^n, c^*, E=0)} \sum_{u^n, x^n, y^n, v^n, w^n} \mathcal{P}(u^n, x^n, y^n, w^n | c^*, E=0) \times \mathcal{P}(v^n | y^n, c^*, E=0) \cdot \frac{1}{n} \sum_{i=1}^n \phi_2(u_i, v_i) \\
& \leq \max_{\mathcal{P}(v^n|w^n, y^n, c^*, E=0)} \sum_{u^n, x^n, y^n, v^n, w^n} \mathcal{P}(u^n, x^n, y^n, w^n | c^*, E=0) \times \mathcal{P}(v^n | w^n, y^n, c^*, E=0) \cdot \frac{1}{n} \sum_{i=1}^n \phi_2(u_i, v_i) \tag{58}
\end{aligned}$$

$$\begin{aligned}
& = \max_{\mathcal{P}(v^n|w^n, y^n, c^*, E=0)} \sum_{u^n, x^n, y^n, v^n, w^n} \mathcal{P}(x^n, y^n, w^n | c^*, E=0) \times \mathcal{P}(u^n | w^n, c^*, E=0) \times \mathcal{P}(v^n | w^n, y^n, c^*, E=0) \cdot \frac{1}{n} \sum_{i=1}^n \phi_2(u_i, v_i) \tag{59}
\end{aligned}$$

$$\begin{aligned}
& = \sum_{w^n} \mathcal{P}(w^n | c^*, E=0) \times \max_{\mathcal{P}(v^n|w^n, c^*, E=0)} \sum_{v^n} \mathcal{P}(v^n | w^n, c^*, E=0) \times \sum_{u^n} \mathcal{P}(u^n | w^n, c^*, E=0) \cdot \frac{1}{n} \sum_{i=1}^n \phi_2(u_i, v_i) \tag{60}
\end{aligned}$$

$$\begin{aligned}
& \leq \sum_{w^n} \mathcal{P}(w^n | c^*, E=0) \times \max_{\mathcal{P}(v^n|w^n, c^*, E=0)} \sum_{v^n} \mathcal{P}(v^n | w^n, c^*, E=0) \times \sum_{u^n} \left(\prod_{i=1}^n Q(u_i | w_i) \right) \cdot \frac{1}{n} \sum_{i=1}^n \phi_2(u_i, v_i) + 2\sqrt{\ln 2\varepsilon} \cdot \bar{\phi}_2 \tag{61}
\end{aligned}$$

$$\begin{aligned}
& = \sum_{w^n} \mathcal{P}(w^n | c^*, E=0) \sum_w Q^n(w) \max_{\mathcal{P}(v|w)} \sum_v \mathcal{P}(v|w) \sum_{u \in \mathcal{U}} Q(u|w) \cdot \phi_2(u, v) + 2\sqrt{\ln 2\varepsilon} \cdot \bar{\phi}_2 \tag{62}
\end{aligned}$$

$$\begin{aligned}
& \leq \sum_w Q(w) \max_{\mathcal{P}(v|w)} \sum_v \mathcal{P}(v|w) \sum_{u \in \mathcal{U}} Q(u|w) \cdot \phi_2(u, v) + (2\sqrt{\ln 2\varepsilon} + \varepsilon) \cdot \bar{\phi}_2 \tag{63}
\end{aligned}$$

$$\begin{aligned}
& = \sum_{u, v} Q(u, w) \cdot Q(v|w) \cdot \phi_2(u, v) + (2\sqrt{\ln 2\varepsilon} + \varepsilon) \cdot \bar{\phi}_2. \tag{64}
\end{aligned}$$

\mathbb{Q}_2 that requires the distribution $\mathcal{Q}(v|w)$ maximizes $\max_{\mathcal{P}(v|w)} \sum_v \mathcal{P}(v|w) \sum_{u \in \mathcal{U}} Q(u|w) \cdot \phi_2(u, v)$.

We proved that the equilibrium condition is satisfied:

$$\Phi_2^n(\sigma^*, \tau^*(\sigma^*)) \geq \max_{\tilde{\tau}} \Phi_2^n(\sigma^*, \tilde{\tau}) - \varepsilon. \tag{67}$$

D. Lemmas

To simplify the notations of Lemma 2 and 3, we remove the conditioning over the code c^* and the event $E = 0$ in the probabilities.

Lemma 2 *The following expression satisfies:*

$$\begin{aligned}
& \left| \sum_{w^n, v^n} \mathcal{P}(w^n, v^n) \sum_{u^n} \mathcal{P}(u^n | w^n) \cdot \frac{1}{n} \sum_{i=1}^n \phi_2(u_i, v_i) \right. \\
& \left. - \sum_{w^n, v^n} \mathcal{P}(w^n, v^n) \sum_{u^n} \left(\prod_{i=1}^n Q(u_i | w_i) \right) \cdot \frac{1}{n} \sum_{i=1}^n \phi_2(u_i, v_i) \right| \\
& \leq \bar{\phi}_2 \cdot 2\sqrt{\ln 2\varepsilon}. \tag{68}
\end{aligned}$$

Lemma 3 *For all sequence $w^n \in \mathcal{W}^n$, we have this equality:*

$$\begin{aligned}
& \max_{\mathcal{P}(v^n|w^n)} \sum_{v^n} \mathcal{P}(v^n | w^n) \sum_{u^n} \left(\prod_{i=1}^n Q(u_i | w_i) \right) \cdot \frac{1}{n} \sum_{i=1}^n \phi_2(u_i, v_i) \\
& = \sum_w Q^n(w) \max_{\mathcal{P}(v|w)} \sum_v \mathcal{P}(v|w) \sum_{u \in \mathcal{U}} Q(u|w) \cdot \phi_2(u, v).
\end{aligned}$$

The proof of Lemma 2 is based on Lemma 4 and the proof of Lemma 3 comes from the hypothesis $E = 0$, of jointly typical sequences $(u^n, w^n, v^n) \in A_\varepsilon^{*n}(\mathcal{Q})$.

Lemma 4 (Posteriors beliefs) *The coding scheme $c^* = (f^*, g^*) \in \mathcal{C}(n)$ described in App. B-A satisfies:*

$$\begin{aligned}
& \mathbb{E}_{W^n} \left[\frac{1}{n} \cdot \sum_{i=1}^n \left\| \mathcal{P}(U_i | W^n, c^*, E=0) - \mathcal{Q}(U_i | W_i) \right\|_1 \right] \\
& \leq 2\sqrt{\ln 2\varepsilon}. \tag{69}
\end{aligned}$$

Lemma 4 corresponds to the notion of ‘‘Strategic Distance’’ introduced in [5] and in the proof of [3, Lemma 36] that implies the main result of [2] and [4].

Proof. [Lemma 4] We consider the code $c^* = (f^*, g^*) \in \mathcal{C}(n)$, stated in App. B-A and we assume that the sequences $(U^n, W^n, V^n) \in A_\varepsilon^{*n}(\mathcal{Q})$ are jointly typical, i.e. the error event is $E = 0$. We provide an upper bound on the L_1 distance based on Pinsker’s and Jensen’s inequalities. We denote by $D(\mathcal{P}||\mathcal{Q})$ the K-L divergence between distributions \mathcal{P} and \mathcal{Q} .

$$\begin{aligned}
& \mathbb{E}_{W^n} \left[\frac{1}{n} \cdot \sum_{i=1}^n \left\| \mathcal{P}(U_i | W^n, c^*, E=0) - \mathcal{Q}(U_i | W_i) \right\|_1 \right] \\
& = \sum_{w^n} \mathcal{P}(w^n | c^*, E=0) \\
& \times \frac{1}{n} \sum_{i=1}^n \left\| \mathcal{P}(U_i | w^n, c^*, E=0) - \mathcal{Q}(U_i | w_i) \right\|_1 \tag{70}
\end{aligned}$$

$$\begin{aligned}
& \leq \sum_{w^n} \mathcal{P}(w^n | c^*, E=0) \\
& \times \frac{1}{n} \sum_{i=1}^n \sqrt{2 \ln 2 \cdot D\left(\mathcal{P}(U_i | w^n, c^*, E=0) \middle\| \mathcal{Q}(U_i | w_i)\right)} \tag{71}
\end{aligned}$$

$$\begin{aligned}
& \leq \sqrt{2 \ln 2 \sum_{w^n} \mathcal{P}(w^n | c^*, E=0)} \\
& \times \sqrt{\frac{1}{n} \sum_{i=1}^n D\left(\mathcal{P}(U_i | w^n, c^*, E=0) \middle\| \mathcal{Q}(U_i | w_i)\right)} \tag{72}
\end{aligned}$$

$$\begin{aligned}
& \leq \sqrt{2 \ln 2 \cdot \frac{1}{n} \sum_{i=1}^n \mathbb{E}_{W^n} \left[D\left(\mathcal{P}(U_i | W^n, c^*, E=0) \middle\| \mathcal{Q}(U_i | W_i)\right) \right]}. \tag{73}
\end{aligned}$$

In equation (70), the L_1 distance regards U_i . Equation (71) comes from Pinsker’s inequality, [40, pp. 370]. Equation (72) comes from Jensen’s inequality for $x \mapsto \sqrt{x}$.

$$\begin{aligned}
& \frac{1}{n} \sum_{i=1}^n \mathbb{E}_{W^n} \left[D \left(\mathcal{P}(U_i|W^n, c^*, E=0) \parallel \mathcal{Q}(U_i|W_i) \right) \right] \\
&= \frac{1}{n} \sum_{(u^n, w^n) \in A_\varepsilon^{*n}(\mathcal{Q})} \mathcal{P}(u^n, w^n | c^*, E=0) \\
&\times \log_2 \frac{1}{\prod_{i=1}^n \mathcal{Q}(u_i|w_i)} - \frac{1}{n} \sum_{i=1}^n H(U_i|W^n, c^*, E=0) \\
&\leq \frac{1}{n} \sum_{(u^n, w^n) \in A_\varepsilon^{*n}(\mathcal{Q})} \mathcal{P}(u^n, w^n | c^*, E=0) \\
&\times n \cdot \left(H(U|W) + \varepsilon \right) - \frac{1}{n} H(U^n|W^n, c^*, E=0) \quad (75) \\
&= \frac{1}{n} I(U^n; W^n | c^*, E=0) - I(U; W) + \varepsilon \quad (76) \\
&\leq \log |\mathcal{M}| - I(U; W) + \varepsilon \quad (77) \\
&\leq I(U; W) + \varepsilon - I(U; W) + \varepsilon \quad (78) \\
&\leq 2\varepsilon. \quad (79)
\end{aligned}$$

Equation (74) is the definition of the K-L divergence.

Equation (75) comes from the property of typical sequences $(U^n, W^n) \in A_\varepsilon^{*n}(\mathcal{Q})$ in [39, pp. 26] and of the entropy.

Equation (76) comes from the i.i.d. property of the source U .

Equations (77) and (78) come from the cardinality of the codebook $|\mathcal{M}|$, introduced in App. B-A.

Equation (76) involves the information leakage $\frac{1}{n} I(U^n; W^n | c^*, E=0)$ corresponding to the amount of information received by P_2 , regarding the source U^n . The information leakage induced by the coding scheme for empirical coordination, is investigated in [22]. \square

REFERENCES

- [1] O. Gossner, P. Hernandez, and A. Neyman, "Optimal use of communication resources," *Econometrica*, vol. 74, pp. 1603–1636, Nov. 2006.
- [2] O. Gossner and T. Tomala, "Secret correlation in repeated games with imperfect monitoring," *Mathematics of Operation Research*, vol. 32, no. 2, pp. 413–424, 2007.
- [3] O. Gossner and T. Tomala, "Empirical distributions of beliefs under imperfect observation," *Mathematics of Operation Research*, vol. 31, no. 1, pp. 13–30, 2006.
- [4] O. Gossner, R. Laraki, and T. Tomala, "Informationally optimal correlation," *Mathematical Programming*, vol. 116, no. 1-2, pp. 147–172, 2009.
- [5] O. Gossner and N. Vieille, "How to play with a biased coin?," *Games and Economic Behavior*, vol. 41, no. 2, pp. 206–226, 2002.
- [6] G. Kramer and S. Savari, "Communicating probability distributions," *Information Theory, IEEE Transactions on*, vol. 53, no. 2, pp. 518 – 525, 2007.
- [7] P. Cuff, H. Permuter, and T. Cover, "Coordination capacity," *Information Theory, IEEE Transactions on*, vol. 56, no. 9, pp. 4181–4206, 2010.
- [8] P. Cuff and L. Zhao, "Coordination using implicit communication," *Information Theory Workshop (ITW), IEEE*, pp. 467–471, 2011.
- [9] P. Cuff and C. Schieler, "Hybrid codes needed for coordination over the point-to-point channel," in *Communication, Control, and Computing (Allerton), 2011 49th Annual Allerton Conference on*, pp. 235–239, Sept 2011.
- [10] M. Le Treust, A. Zaidi, and S. Lasaulce, "An achievable rate region for the broadcast wiretap channel with asymmetric side information," *IEEE Proc. of the 49th Allerton conference, Monticello, Illinois*, pp. 68 – 75.
- [11] M. Le Treust, "Empirical coordination for the joint source-channel coding problem," *submitted to IEEE Trans. on Information Theory*, <http://arxiv.org/abs/1406.4077>, 2014.
- [12] B. Laroousse, S. Lasaulce, and M. Bloch, "Coordination in distributed networks via coded actions with application to power control," *Submitted to IEEE Transactions on Information Theory*, <http://arxiv.org/abs/1501.03685>, 2014.
- [13] M. Le Treust, "Correlation between channel state and information source with empirical coordination constraint," in *IEEE Information Theory Workshop (ITW)*, pp. 272–276, Nov 2014.
- [14] M. Le Treust, "Empirical coordination with two-sided state information and correlated source and state," in *IEEE International Symposium on Information Theory (ISIT)*, 2015.
- [15] M. Le Treust, "Empirical coordination with channel feedback and strictly causal or causal encoding," in *IEEE International Symposium on Information Theory (ISIT)*, 2015.
- [16] B. Laroousse, S. Lasaulce, and M. Wigger, "Coordinating partially-informed agents over state-dependent networks," *IEEE Information Theory Workshop (ITW)*, 2015.
- [17] R. Blasco-Serrano, R. Thobaben, and M. Skoglund, "Polar codes for coordination in cascade networks," in *Proc. of the International Zurich Seminar on Communication, Zurich, Switzerland*, pp. 55 – 58, March 2012.
- [18] R. Chou, M. Bloch, and J. Kliewer, "Polar coding for empirical and strong coordination via distribution approximation," in *Information Theory Proceedings (ISIT), 2015 IEEE Internat. Symp. on*, June 2015.
- [19] R. Chou, M. Bloch, and J. Kliewer, "Empirical and strong coordination via soft covering with polar codes," *submitted to IEEE Transactions on Information Theory*, <http://arxiv.org/abs/1608.08474>, 2016.
- [20] G. Cervia, L. Luzzi, M. R. Bloch, and M. L. Treust, "Polar coding for empirical coordination of signals and actions over noisy channels," in *in Proc. IEEE Information Theory Workshop (ITW)*, 2016.
- [21] C. Schieler and P. Cuff, "Rate-distortion theory for secrecy systems," *IEEE Trans. on Information Theory*, vol. 60, pp. 7584–7605, Dec 2014.
- [22] M. Le Treust and M. Bloch, "Empirical coordination, state masking and state amplification: Core of the decoder's knowledge," *Proceedings of the IEEE International Symposium on Information Theory (ISIT)*, 2016.
- [23] J. Nash, "Non-cooperative games," *Annals of Mathematics*, vol. 54, pp. 286–295, 1951.
- [24] H. von Stackelberg, *Marktform und Gleichgewicht*. Oxford University Press, 1934.
- [25] V. P. Crawford and J. Sobel, "Strategic Information Transmission," *Econometrica*, vol. 50, no. 6, pp. 1431–1451, 1982.
- [26] F. Forges, "Non-zero-sum repeated games and information transmission," in: *N. Meggido, Essays in Game Theory in Honor of Michael Maschler*, Springer-Verlag, no. 6, pp. 65–95, 1994.
- [27] E. Kamenica and M. Gentzkow, "Bayesian persuasion," *American Economic Review*, vol. 101, pp. 2590 – 2615, 2011.
- [28] R. Aumann and M. Maschler, *Repeated Games with Incomplete Information*. MIT Press, Cambridge, MA, 1995.
- [29] S. Sorin, *A First Course on Zero-Sum Repeated Games*, vol. 37 of *Mathématiques et Applications*. Springer, 2002.
- [30] J. Mertens, S. Sorin, and S. Zamir, *Repeated Games*. Cambridge University Press, 2015.
- [31] R. Laraki, "The splitting game and applications," *International Journal of Game Theory*, vol. 30, pp. 359–376, 2001.
- [32] J. Renault, E. Solan, and N. Vieille, "Optimal dynamic information provision," <http://www.lse.ac.uk/statistics/events/2015-16-Seminar-Series/Optimal-Dynamic-Information-Provision.pdf>, February 2016.
- [33] J. Ely, "Beeps," *Manuscript, Department of Economics, Northwestern University*, 2015.
- [34] M. O. Jackson and H. F. Sonnenschein, "Overcoming incentive constraints by linking decisions," *Econometrica*, vol. 75, pp. 241 – 257, January 2007.
- [35] E. Akyol, C. Langbort, and T. Başar, "Strategic compression and transmission of information," in *Information Theory Workshop - Fall (ITW), 2015 IEEE*, pp. 219–223, Oct 2015.
- [36] E. Akyol, C. Langbort, and T. Başar, "On the role of side information in strategic communication," in *2016 IEEE International Symposium on Information Theory (ISIT)*, pp. 1626–1630, July 2016.
- [37] S. Sarıtaç, S. Yüksel, and S. Gezici, "Dynamic signaling games under nash and stackelberg equilibria," in *2016 IEEE International Symposium on Information Theory (ISIT)*, pp. 1631–1635, July 2016.
- [38] E. V. Belmega, S. Lasaulce, and M. Debbah, "Power allocation games for mimo multiple access channels with coordination," *IEEE Trans. on Wireless Communications*, vol. 8, no. 6, pp. 3182–3192, 2009.
- [39] A. E. Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge University Press, Dec. 2011.
- [40] T. M. Cover and J. A. Thomas, *Elements of information theory*. New York: 2nd. Ed., Wiley-Interscience, 2006.