Transforming Monitoring Structures with Resilient Encoders. Application to Repeated Games

Maël Le Treust, Samson Lasaulce

▶ **To cite this version:**

Maël Le Treust, Samson Lasaulce. Transforming Monitoring Structures with Resilient Encoders. Application to Repeated Games. Dynamic Games and Applications, 2013, 3 (1), pp.38-67. 10.1007/s13235-012-0058-3 . hal-00744802v2

HAL Id: hal-00744802
https://hal.science/hal-00744802v2

Submitted on 28 Mar 2018

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.

Noname manuscript No.
(will be inserted by the editor)

Transforming Monitoring Structures with Resilient Encoders. Application to Repeated Games.

Maël Le Treust · Samson Lasaulce

Received: date / Accepted: date

arXiv:1210.6241v1 [cs.IT] 23 Oct 2012

Abstract An important feature of a dynamic game is its monitoring structure namely, what the players effectively see from the played actions. We consider games with arbitrary monitoring structures. One of the purposes of this paper is to know to what extent an encoder, who perfectly observes the played actions and sends a complementary public signal to the players, can establish perfect monitoring for all the players. To reach this goal, the main technical problem to be solved at the encoder is to design a source encoder which compresses the action profile in the most concise manner possible. A special feature of this encoder is that the multi-dimensional signal (namely, the action profiles) to be encoded is assumed to comprise a component whose probability distribution is not known to the encoder and the decoder has a side information (the private signals received by the players when the encoder is off). This new framework appears to be both of game-theoretical and information-theoretical interest. In particular, it is useful for designing certain types of encoders that are resilient to single deviations and provide an equilibrium utility region in the proposed setting; it provides a new type of constraints to compress an information source (i.e., a random variable). Regarding the first aspect, we apply the derived result to the repeated prisoner's dilemma.

Keywords Arbitrarily varying source · Dynamic games · Folk theorem · Games with imperfect monitoring · Information constraint · Observation structure · Source coding

Maël Le Treust
Université Paris-Est Marne-la-Vallée
Institut d'électronique et d'informatique Gaspard-Monge
e-mail : mael.letreust@univ-mlv.fr

Samson Lasaulce
Laboratoire des Signaux et Systèmes
CNRS - Université Paris Sud XI - Supélec
e-mail : samson.lasaulce@lss.supelec.fr

1 Introduction

The set of equilibrium utilities of a non-cooperative dynamic game is strongly related to the observation capabilities of the players. For instance, in a long-run repeated prisoner's dilemma where the two players do not see anything from the played actions (blind players), the only equilibrium point corresponds to the inefficient outcome where both players defect [Aumann(1981a)][Sorin(1992)]. On the other hand, when players perfectly observe all the actions which have been played (perfect monitoring assumption), efficient equilibria can be sustained; in particular, the social optimum is an equilibrium point of the infinitely repeated dilemma or its version with discount factor. This special case illustrates the potential need for being able to transform the monitoring structure of a repeated game into a new one. The relevance of such a transformation may appear in other types of settings such as stochastic games, multi-agent learning, or networked optimization. For example, perfect monitoring (PM) can be targeted to implement the standard fictitious play or best-response algorithms (see e.g., [Peyton(2004)]). The desired final monitoring structure (i.e., after transformation) does not necessarily need to be PM and, for example, ensuring that the players observe (thanks to the transformation) a certain public signal can be sufficient to obtain efficient outcomes for the game. The solution proposed in this paper is to implement this monitoring structure transformation by adding an external agent or encoder (whose role is not strategic but only to encode signals and send them to the players to improve their observation capability) to the initial game. For the sake of clarity and simplicity, the encoder is assumed to perfectly observe the actions played and the desired structure, after transformation, is PM. Note that PM at the encoder is not always necessary to ensure PM for the players (see [LeTreust and Lasaulce(2011a)]). Interestingly, there exist some practical scenarios where assuming PM at a terminal is relevant. In wireless communications, the decentralized multiple access channel case is known to be very important [LeTreust and Lasaulce(2010)]. In this scenario, there are one receiver (e.g., a WiFi access point or a base station) and several transmitters (e.g., mobile terminals) which choose freely their transmission policy (say their power allocation policy) in order to optimize some performance metric such as the individual transmission rate. Considering that the base station has a computational and observation capability much larger than the mobile transmitters is a typical assumption in wireless communications (see e.g., [Kowalewski(2000)][DaSilva et al(2011)DaSilva, Taffin, Lasaulce, and Buljore]). As a consequence, the receiver can, in particular, have the role of an encoder which sends a feedback on the played actions to the transmitters. Another important scenario of practical interest for which the framework proposed in this paper is fully relevant is the case of sensor networks with a fusion center (see e.g., [I. F. Akyildiz and Cayirci(2002)]).

One of the main issues addressed in this paper is the design of an encoder which is capable of transforming a monitoring structure by sending complementary public signals to the players. The problem comes from the fact that

the set of public signals has a fixed cardinality. One of the consequences of this assumption is the existence of an information constraint on the played action profiles and more precisely on their distribution, and therefore on the feasible players' utilities. As explained further, characterizing this information constraint amounts to designing an encoder which represents the information source (namely, the action profile) in a manner as concise as possible. However, to make the source encoder able to operate at equilibrium (and therefore characterize equilibrium utilities), the encoder has to possess a certain property, called the resilience property [LeTreust and Lasaulce(2011b)], which has a cost in terms of compression efficiency. In terms of communication, such a property ensures that, even when one player uses a distribution on his action sequences which is arbitrary and unknown to the encoder, PM remains guaranteed. In strategic terms, if we consider the case of repeated games (which is the case study chosen in this paper), it means that grim-trigger-like plans can be implemented.

The paper is structured as follows. A state of the art on the problem under investigation is done in Sec. 2. Sec. 3 exploits information-theoretic tools to derive one of the central two results of this paper which is the information constraint (9) stated in Theorem 1 and explains how this constraint translates into a set of action profile distributions (and therefore into feasible utilities) that are compatible with the perfect monitoring assumption. Sec. 4 provides the second important result, stated in Theorem 2, which is an achievable equilibrium utility region for encoder-assisted infinitely repeated games with signals. The paper is concluded in Sec. 5.

2 Related works

Before mentioning some relevant works related to the one reported here, it is useful to define at this point a monitoring structure. A monitoring structure is a conditional or transition probability defined by :

$$\urcorner : \mathcal{A} \longrightarrow \Delta(\mathcal{S}) \tag{1}$$

where $\mathcal{A} = \mathcal{A}_1 \times \mathcal{A}_2 \times ... \times \mathcal{A}_K$ is the discrete set of action profiles, K is the number of players, $\mathcal{A}_k$ is the discrete set of actions of player $k \in \mathcal{K} = \{1, 2, ..., K\}$, $\mathcal{S} = \mathcal{S}_1 \times \mathcal{S}_2 \times ... \times \mathcal{S}_K$, $\mathcal{S}_k$ is the discrete set of signals received by player k, and the notation $\Delta(\mathcal{S})$ stands for the set of probability distributions on $\mathcal{S}$ (unit simplex).

The first relevant body of related works comprises papers providing lossless [Shannon(1948)] and zero-errors [Shannon(1956)][Witsenhausen(1976)] source coding theorems. Indeed, the role of the encoder in this paper is to encode a sequence (or block) of action profiles into a sequence of public signals which is observed by the players. As already mentioned, making this in a concise manner is of prime interest to characterize the information constraint. The considered source coding problem has two main features : the decoders (namely, the players) have a side information on the source (the private signal) and we

want the encoder to be resilient to single deviations that is, the past action profiles are decoded reliably even when the probability distribution of the action of a given player varies arbitrarily over time and remains unknown to both the encoder and decoders. Remarkably, the information theory literature provides the right framework to design such encoders. The corresponding framework is the one of arbitrary varying sources (AVS) : the source distribution $\mathcal{P}_v(a) \in \Delta(\mathcal{A})$ can vary from sample to sample, depending on a parameter or state $v \in \mathcal{V}$ which represents, in our setting, the probability distribution of the deviator's action. The most relevant works on AVS is based on graph coloring [Bondy and Murty(1976)] and can be found in [Ahlswede(1979)] and [Ahlswede(1980)]. Indeed, the latter references deal with the scenario of two correlated sources either in the case where the destination is informed with the sequence of states or in the case where it is not known. The work reported in this paper is precisely related to the scenario of two arbitrary varying correlated sources of actions $\boldsymbol{a}$ and private signals $\boldsymbol{s}_k$ with a destination (i.e. player $k \in \mathcal{K}$) uninformed of the state (i.e. strategy of an eventual deviator) ; this scenario is described by Fig. 3.1. One of our contributions, in addition to establishing a link between equilibrium utility regions and the AVS literature, is to show that the entropy positiveness condition (EPC) in [Ahlswede(1979)][Ahlswede(1980)], under which source coding rates (i.e. optimal compression level) can be characterized, can be removed and replaced with another mathematical condition which is of strong game-theoretic interest namely, the resilience property. Additionally, it holds for some useful special cases for which the EPC is not met, the case of deterministic channels in particular.

The second body of works concerns works on folk theorems. The stronger results have been obtained for one of the simplest classes of dynamic games namely, the one of repeated games (see e.g., [Sorin(1992)] for a survey). The standard approach consists in assuming a given monitoring structure (e.g., standard-trivial monitoring [Lehrer(1991)], public monitoring [Fudenberg et al(1994)Fudenberg, Levine, and Maskin], or almost-perfect monitoring [Hörner and Olszewski(2006)]) and, then, deriving a folk theorem. Compared to these works, our approach is different since we do not try to characterize the equilibrium utilities of a repeated game with an arbitrary monitoring structure (which is an open problem [Renault and Tomala(2011)]). Rather, our approach aims at transforming, with an additional encoder, an arbitrary monitoring structure of any dynamic game into a new monitoring structure for which the equilibrium utilities can be fully characterized ; in this paper, PM is the targeted final structure. Even though the final monitoring structure is PM, there are still some differences between a dynamic game with PM (the focus will be on repeated games here) and a dynamic game where players have PM thanks to the encoder :

– there exists an internal information constraint on the action distribution due to the fact that the set of public signals has a fixed cardinality ;

– action profiles are encoded by blocks by the encoder and each player decodes a block of played actions from a whole block of observations. Therefore PM is established with a delay ;
– only i.i.d equilibrium utilities (and convex combinations of them) are studied. This assumption on the action profiles is well motivated in the paper and does not prevent us from deriving useful results which may be extended if needed.

For all of these reasons, we will use the term "virtually perfect monitoring" (VPM) to refer to such a framework.

To conclude on the most relevant references related to the work reported in this paper, we will mention a couple of references at the intersection between game and information theory. For instance, in [Lehrer(1988)], [Bavly and Neyman(2003)], [Peretz(2011)] entropy-based information constraints are used to characterize the individually rational levels of repeated games with bounded recall. In [Gossner and Tomala(2007)], the authors characterize the maximum utility a team can guarantee against another in a class of repeated games with imperfect monitoring by exploiting a constraint on possible correlation schemes expressed in terms of entropy variation. In [Gossner et al(2006)Gossner, Hernandez, and Neyman], the authors are exploiting an information constraint in the sense of the present work that is, the source coding rate has to be less than the channel capacity, although the constraint is not interpreted this way in their work. This leads to a characterization of equilibrium utilities a team of two players can implement when only one player is (noncausally) informed of the i.i.d. sequence of states of the repeated game.

3 Virtual perfect monitoring of an arbitrarily varying information source

3.1 Methodology

The scenario under consideration is as follows (see Figure 3.1). Let us fix a family of probability distributions $\mathcal{P}_k^\star \in \Delta(\mathcal{A}_k)$ with $k \in \mathcal{K}$. When a given action profile $a = (a_1, a_2, ..., a_K) \in \mathcal{A}$ is drawn from the product probability $\mathcal{P}^\star = \mathcal{P}_1^\star \otimes \ldots \otimes \mathcal{P}_K^\star \in \Delta(\mathcal{A})$, player $k \in \mathcal{K}$ receives a symbol $s_k \in \mathcal{S}_k$ with a probability given by the conditional probability

$$\urcorner(s_k|a) = \sum_{s_{-k} \in \mathcal{S}_{-k}} \urcorner(s_k, s_{-k}|a). \tag{2}$$

An encoder $\mathcal{C}$, who perfectly monitors the played actions, encodes the observed action profiles by blocks or sequences of size $n \geq 1$ into a sequence of public signals $s_0 \in \mathcal{S}_0$ which are received by all the players. These public signals form a perfect channel of capacity $\log_2 |\mathcal{S}_0|$, which is orthogonal to the one defined by $\urcorner$ that is, player k actually receives a pair of signals (s_k, s_0) for every played

action profile. Note that player k recall it's own action a_k. The purpose of the encoder is to use the minimal amount of additional information, in order for every player to acquire the information which is missing to have PM. In what follows, we first define a code in our setup. Second, we define the notion of virtually perfect monitoring (VPM) of actions profile $a = (a_1, a_2, ..., a_K) \in \mathcal{A}$ defined as an arbitrarily varying information source (AVS). Third, we prove a theorem which state an information constraint on the action profile distribution which is due to the fact that the communication channel between the encoder and players has a limited capacity. Denote $\mathcal{A}^n$ (resp. $\mathcal{A}^\infty$) the set of sequences $a^n \in \mathcal{A}^n$ of length $n \in \mathbb{N}$ (resp. of sequences $a^\infty \in \mathcal{A}^\infty$ of infinite length).

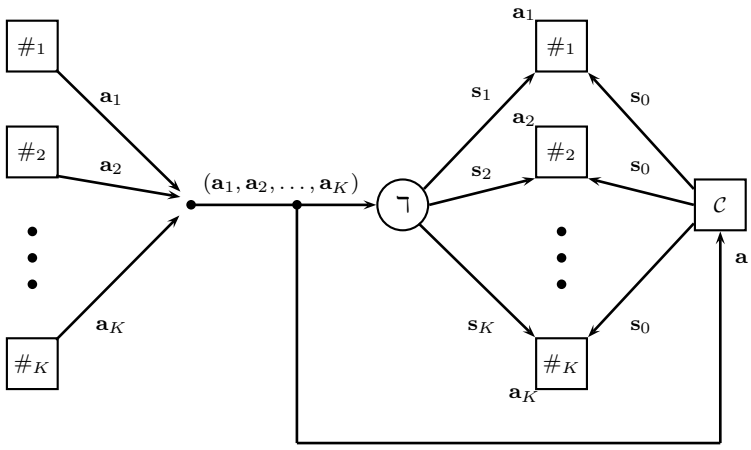

Fig. 1 Each action profile of the game $\mathbf{a} = (\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_K)$ generates a signal profile $(\mathbf{s}_1, \mathbf{s}_2, ...\mathbf{s}_K)$ through a condition probability $\daleth$. Player $\#k$ (represented twice here above) only observes $\mathbf{s}_k$ from this action profile $\mathbf{a}$. The encoder $\mathcal{C}$, who perfectly monitors the played action profiles $\mathbf{a}$, builds a complementary public signal $\mathbf{s}_0$ which is observed by all the players. Each player has to reconstruct virtual perfect monitoring (VPM) from a sequence of pairs of signals $(\mathbf{s}_k, \mathbf{s}_0)$ and the knowledge of the sequence of its individual actions $\mathbf{a}_k$.

3.2 Information constraint for resilient coding with side information at the decoder

Here, we assume that the distribution of the source may vary from stage (or action profile) to stage (or action profile) ; this is the framework of arbitrarily varying source (AVS) coding.

Definition 1 (Arbitrarily Varying Source (AVS)) Let $\mathcal{P}^\star \in \Delta(\mathcal{A})$ a probability distribution (mixed strategy) and $\mathcal{V}$ the set of states of the source:

$$\mathcal{V} = \cup_{k\in\mathcal{K}} \Delta(\mathcal{A}_k^\infty). \tag{3}$$

The arbitrarily varying (AVS) information source $\mathbf{a} \in \mathcal{A}$ is at a certain state $v \in \mathcal{V}$, when one component of the action profile has a distribution which may vary arbitrarily over time and is fully unknown to the coder. For example, when the sequence of the states is $v = \mathcal{Q}_i \in \Delta(\mathcal{A}_i^\infty) \subset \mathcal{V}$, the sequence of actions $\mathbf{a}^n = (\mathbf{a}_1^n, \ldots, \mathbf{a}_K^n)$ is drawn following a probability distribution given by :

$$\mathcal{P}_v(\mathbf{a}_1^n, \ldots, \mathbf{a}_i^n, \ldots, \mathbf{a}_K^n) = \left[\mathcal{P}_1^{\star\otimes n} \otimes \ldots \otimes \mathcal{Q}_i \otimes \ldots \otimes \mathcal{P}_K^{\star\otimes n}\right](\mathbf{a}^n). \tag{4}$$

Now, we formally define the notion of code for the AVS represented by Fig. 3.1.

Definition 2 A code λ of size n for the encoder $\mathcal{C}$ and decoders $\mathcal{K}$ consists of an encoding function f_0 and K decoding functions $(g_k)_{k\in\mathcal{K}}$ defined as :

$$\begin{array}{ll} f_0 : & \mathcal{A}^n \longrightarrow \mathcal{S}_0^n \\ g_k : \mathcal{S}_0^n \times \mathcal{S}_k^n \times \mathcal{A}_k^n \longrightarrow \mathcal{A}^n, & \forall k \in \mathcal{K}. \end{array} \tag{5}$$

Denote by $\Lambda(n)$, the set of codes for which the length $n \in \mathbb{N}$ of the code-words is fixed.

$$\mathcal{P}_e(\lambda) = \sum_{k\in\mathcal{K}} \max_{i\in\mathcal{K}} \max_{v_i\in\Delta(\mathcal{A}_i^\infty)} \mathcal{P}_{v_i}(\mathbf{a}^n \neq g_k(\mathbf{s}_k^n, \mathbf{s}_0^n, \mathbf{a}_k^n)), \tag{6}$$

The error probability $\mathcal{P}_e(\lambda)$ of the code $\lambda \in \Lambda(n)$ is defined by equation (6) and corresponds to the sum of the error probability for each decoder $k \in \mathcal{K}$, considering every possible deviation $v_i \in \Delta(\mathcal{A}_i^\infty)$ of player $i \in \mathcal{K}$ (i.e. any variation of the source).

Definition 3 (Virtually Perfect Monitoring (VPM)) Players $\mathcal{K}$ have a virtually perfect monitoring (VPM) of the information source $\mathbf{a} \in \mathcal{A}$ if for all $\varepsilon > 0$, there exists a parameter $n \in \mathbb{N}$, and a code $\lambda \in \Lambda(n)$ such that:

$$\mathcal{P}_e(\lambda) \leq \varepsilon, \tag{7}$$

The condition (7) means that it is possible to find coding and decoding functions to represent any sequence of n realizations of the $K-$dimensional random variable $\boldsymbol{a}$ with $2^{n\log_2|\mathcal{S}_0|}$ indices or sequences of public signals in such a way that, any decoder k, based on the knowledge of (s_0^n, s_k^n, a_k^n), can find the sequence a^n with an arbitrarily small probability of error. In a game theoretical framework, the players virtually perfect monitor the sequences of past actions played.

At this point, the main issue is to be able to characterize the set of AVS information source that are compatible with the VPM of the players $\mathcal{K}$. The

AVS hypothesis guarantee that the past actions played will be observed by all the players even if one of them deviates, manipulates the coding scheme $\lambda \in \Lambda(n)$ in order to break reliability. Theorem 1 provides an information constraint which guarantee VPM for the AVS information source of player's actions $\mathbf{a}$. To state this theorem, an auxiliary graph needs to be defined first.

Definition 4 (Auxiliary graph) For each player $i \in \mathcal{K}$, an auxiliary graph $\mathcal{G}_i$ is defined as follows $\mathcal{G}_i = (\mathcal{A}_i, \mathcal{E}_i)$. The actions $a_i \in \mathcal{A}_i$ of player $i \in \mathcal{K}$ are the vertices of the graph. There exists an edge $e_i = (a_i, a_i') \in \mathcal{E}_i$ between two actions $a_i \in \mathcal{A}_i$ and $a_i' \in \mathcal{A}_i$ if :

$$\exists\, a_{-i} \in \text{Supp}\, \mathcal{P}^{\star}_{-i},\ \exists k \in \mathcal{K},\ \exists s_k \in \mathcal{S}_k,\ \exists \delta > 0,\ \text{s.t.}$$
$$\min(\urcorner(s_k|a_i, a_{-i}), \urcorner(s_k|a_i', a_{-i})) \geq \delta$$

where Supp $\mathcal{P}^{\star}_{-i}$ is the support of the probability distribution $\mathcal{P}^{\star}_{-i}$ defined by $\mathcal{P}^{\star}_{-i} = \bigotimes_{j\neq i} \mathcal{P}^{\star}_{j} \in \prod_{j\neq i} \Delta(\mathcal{A}_j)$.

Two vertices $a_i \in \mathcal{A}_i$ and $a_i' \in \mathcal{A}_i$ are neighbors in the graph $\mathcal{G}_i$ if the probability that these actions lead, through $\urcorner$, to the same signal $s_k \in \mathcal{S}_k$ for at least one player $k \in \mathcal{K}$ is not zero. Now, to define the chromatic number [Bondy and Murty(1976)] of the graph $\mathcal{G}_i$, we define the notion of coloring in our context.

Definition 5 (Coloring) Let Φ_i a set of colors. A coloring of the graph $\mathcal{G}_i$ is a function $\phi_i : \mathcal{A}_i \longrightarrow \Phi_i$ which satisfies :

$$\forall e_i = (a_i, a_i') \in \mathcal{E}_i,\ \text{we have that } \phi_i(a_i) \neq \phi_i(a_i'). \tag{8}$$

A minimal coloring of the graph $\mathcal{G}_i$ is a coloring ϕ_i for which the cardinality of the set of colors Φ_i is minimal. The chromatic number χ_i of the graph $\mathcal{G}_i$ is the cardinality $|\Phi_i|$ of the set of colors of the minimal coloring of the graph $\mathcal{G}_i$. This is precisely this quantity which is used in the next theorem.

Theorem 1 (Coding result for AVS) *Players $\mathcal{K}$ have a virtually perfect monitoring (VPM) of the arbitrarily varying (AVS) information source $\boldsymbol{a} \in \mathcal{A}$ if the following condition is met :*

$$R^{\star} = \max_{i\in\mathcal{K}} \left[\max_{\substack{k\in\mathcal{K},\\ a_i\in\mathcal{A}_i}} H(\boldsymbol{a}_{-i,k}|\boldsymbol{s}_k(a_i), \boldsymbol{a}_k) + \log_2 \chi_i \right] < \log_2 |\mathcal{S}_0|, \tag{9}$$

where :

- $\boldsymbol{a}_{-i,k}$ *is the action profile without the components i and k. It is distributed as* $\mathcal{P}^{\star}_{-i,k} \in \prod_{\substack{j\neq i,\\ j\neq k}} \Delta(\mathcal{A}_j)$ *;*
- $\boldsymbol{s}_k(a_i)$ *is the signal received by player k when the action a_i is fixed. It is induced by $\boldsymbol{a}_{-i}$ and the transition probability $\urcorner$:*

$$\urcorner_{a_i} : \mathcal{A}_{-i} \longrightarrow \Delta(\mathcal{S}_k) \tag{10}$$
$$a_{-i} \longrightarrow \urcorner_{a_i}(s_k|a_{-i}) = \urcorner(s_k|a_i, a_{-i}) = \sum_{s_{-k}\in\mathcal{S}_{-k}} \urcorner(s_k, s_{-k}|a_i, a_{-i})\ ; \tag{11}$$

- $\log_2 |\mathcal{S}_0|$ *is given by the cardinality of the set of public signals and corresponds to the capacity of the perfect channel between the encoder* $\mathcal{C}$ *and the players* $\mathcal{K}$.

Several comments are in order. First, let us comment on the main assumptions. The i.i.d assumption over time made on the source to be encoded is common in the information theory literature and will only be briefly commented. Solving the i.i.d. case might not only be helpful but even sufficient for solving the case with arbitrary correlation between consecutive source samples. To be more specific, if the source generates B blocks of ℓ correlated symbols for B sufficiently large, $\ell < +\infty$, and i.i.d. blocks, then the information constraint directly follows from the original i.i.d case (concerning i.i.d symbols) by considering vectors of symbols instead of symbols. Beyond this framework, the source coding literature comprises works dealing with refinements such as universal coding [Gallager(1976)] and information-spectrum based coding [Han(2003)]. Now, from a game-theoretic perspective, studying sequences of i.i.d profiles (up to one component) is not only an intermediate case which can be challenging technically (think of repeated games with arbitrary monitoring structures) but also to design implementable equilibrium action plans. As for relaxing the i.i.d assumption over space (over the components), provided the resilience property is relaxed and the joint distribution on the actions is known to the encoder, it only consists in changing scalar quantities into vectors (of size K). When resilience to single deviations is required, the spatial i.i.d assumption is useful to derive information constraint (as advocated by the proof provided in App. A) but studying necessity is a possible extension of this work. At last note that the spatial i.i.d assumption allows one to study mixed strategies which is known to be important.

Now, let us comment on the result i.e., the information constraint defined by (9). The presence of the maximum over i is due to the fact that the location of the component (which corresponds to the deviator in a game), whose distribution is unknown, is itself unknown to $\mathcal{C}$. The second maximum over k and a_i indicates the case where the deviator i chooses the worst action a_i in terms of coding efficiency for to the worst decoder k. The conditioning w.r.t. $(\mathbf{s}_k(a_i), \mathbf{a}_k)$ in the entropy translates the knowledge of the decoder in terms of side information, which therefore reduces the entropy. The isolated term $\log_2 \chi_i$ corresponds to the amount of information needed by $\mathcal{C}$ to encode a component separately ; since the probability distribution of $\mathbf{a}_i$ is unknown, symbol-by-symbol coding is optimal here. Without side information at the decoder i this quantity would be $\log_2 |\mathcal{A}_i|$. At last, the righthandside term $\log_2 |\mathcal{S}_0|$ corresponds to the channel capacity of a broadcast channel with a public message and for which the decoders directly observe the signal sent by the encoder (see e.g., [Cover and Thomas(2006)]).

To conclude this section, let us comment on the proof of this theorem. Although the detailed proof of this theorem is provided in Sec. A, we would like to mention here some technical differences w.r.t the derivation made by Ahlswede in [Ahlswede(1980)]. The imposed condition is totally different. Im-

posing resilience to single deviations to the source encoder requires to transmit without error the sequence of actions of the deviating player. Our proof is based on a sequence of coloring where the vertices are the symbols whereas Ahlswede [Ahlswede(1980)] use a coloring where the vertices are the sequences of symbols. To exploit the law of large numbers for the sequences of symbols, his proof requires an additional condition which is EPC. In our framework, this condition is removed and replaced with a condition over the admissible sequences of states (3) and by the feature that the random signal $\mathbf{s}$ depends on the state v only through the action $\mathbf{a}$. Our result is applicable to the case of deterministic transition probability $\urcorner$ whereas this special type of transition probabilities does not meet EPC.

4 Equilibrium utilities of an encoder-assisted repeated games with signals

The goal of this section is to characterize equilibrium utilities of an infinite repeated game with signals where an additional encoder establishes VPM. To this end, notations, definitions and results of the preceding section are used.

4.1 Game formulation and main result

We consider an encoder-assisted repeated game with signals. The stage or constituent game is given by the triplet $(\mathcal{K}, (\mathcal{A}_k)_{k\in\mathcal{K}}, (u_k)_{k\in\mathcal{K}})$, where $u_k \in \mathbb{R}$ is the utility function of player $k \in \mathcal{K}$. The private monitoring structure is given by the conditional probability $\urcorner(s|a) : \mathcal{A} \longrightarrow \Delta(\mathcal{S})$. The encoder $\mathcal{C}$ is assumed to perfectly monitor the past action profile $a \in \mathcal{A}$ and send a public message $s_0 \in \mathcal{S}_0$ to the players.
A strategy for the encoder (by abuse of language we use the term strategy here even though in this paper the encoder has no utility in the game-theoretic sense) is a sequence of causal functions or mappings $\sigma = (\sigma^t)_{t\geq 1}$ with $\forall t \geq 1, \quad \sigma^t : \mathcal{A}^{t-1} \times \mathcal{S}_0^{t-1} \to \mathcal{S}_0$; t stands for the stage index and a_t is the profile played at stage t ; the set of strategies of the encoder will be denoted by Σ.
A behavior strategy for a player is a sequence of causal functions or mappings $(\tau_k^t)_{t\geq 1}$ with $\forall t \geq 1, \quad \tau_k^t : (\mathcal{A}_k \times \mathcal{S}_k \times \mathcal{S}_0)^{t-1} \to \Delta(\mathcal{A}_k)$; the notation $\tau = (\tau_1, \tau_2, ...\tau_K)$ will stand for a profile of behavior strategies for the repeated game ; the set of behavior strategies will be denoted by $\mathcal{T} = \prod_{k\in\mathcal{K}} \mathcal{T}_k$.
At last, we will denote by $\mathcal{P}_{\sigma,\tau}$ the probability distribution on the infinite sequences of actions, private and public signals $((a_k^\infty)_{k\in\mathcal{K}}, (s_k^\infty)_{k\in\mathcal{K}}, s_0^\infty) \in \mathcal{A}^\infty \times \mathcal{S}^\infty \times \mathcal{S}_0^\infty$ induced by the pair of strategies $(\sigma, \tau) \in \Sigma \times \mathcal{T}$. At this point, one can define a uniform equilibrium of the encoder-assisted repeated game with signal.

Definition 6 (Equilibrium points) A pair of strategies $(\sigma, \tau) \in \Sigma \times \mathcal{T}$ of the encoder $\mathcal{C}$ and the players $\mathcal{K}$ is a uniform equilibrium of the encoder-

assisted repeated game with signals if :

(i) For each player $k \in \mathcal{K}$, the expected utility,

$$\gamma_k^T(\sigma,\tau) = \mathbb{E}_{\sigma,\tau}\left(\frac{1}{T}\sum_{t=1}^{T} u_k(a_t)\right), \tag{12}$$

has a limit when $T \to +\infty$;

(ii) $\forall \varepsilon > 0,\ \exists \bar{T} > 0,\ \forall T \geq \bar{T},\ \forall k \in \mathcal{K},\ \forall \tau_k' \in \mathcal{T}_k$, such that,

$$\gamma_k^T(\sigma,\tau_k',\tau_{-k}) \leq \gamma_k^T(\sigma,\tau_k,\tau_{-k}) + \varepsilon. \tag{13}$$

The point $U^\star = (U_1^\star, U_2^\star, ..., U_K^\star) \in \mathbb{R}^K$ is a vector of equilibrium utilities if there exists a pair of strategies $(\sigma^\star,\tau^\star)$ such that:

$$\forall k \in \mathcal{K},\ \lim_{T\to+\infty} \gamma_k^T(\sigma^\star,\tau^\star) = U_k^\star. \tag{14}$$

The set of the equilibrium points of the encoder-assisted repeated game with signals will be denoted by NE_{enc}^{∞}.

Definition 7 (Individually rational points) The independent min-max level υ_k of player $k \in \mathcal{K}$ is defined by (15) and is also called punishment or defense level. The individually rational IR utilities are defined by (16) and correspond to the utilities that Pareto-dominate the min-max levels defined as

$$\upsilon_k = \min_{\mathcal{P}_{-k} \in \prod_{i\neq j} \Delta(\mathcal{A}_j)} \max_{\mathcal{P}_k \in \Delta(\mathcal{A}_k)} \mathbb{E}_{\mathcal{P}_k,\mathcal{P}_{-k}}\Big[u_k(\mathbf{a}_k,\mathbf{a}_{-k})\Big], \quad k \in \mathcal{K}, \tag{15}$$

$$IR = \left\{(x_k)_{k\in\mathcal{K}} \in \mathbb{R}^K \;\middle|\; x_k \geq \upsilon_k \quad \forall k \in \mathcal{K}\right\}. \tag{16}$$

Definition 8 (Information constraint set) The set $\mathcal{R}$ of mixed actions that satisfy the information constraint (9) is defined by :

$$\mathcal{R} = \left\{\mathcal{P} \in \prod_{k\in\mathcal{K}} \Delta(\mathcal{A}_k) \;\middle|\; \max_{i\in\mathcal{K}}\left[\max_{\substack{k\in\mathcal{K},\\ a_i\in\mathcal{A}_i}} H(\mathbf{a}_{-i,k}|\mathbf{s}_k(a_i),\mathbf{a}_k) + \log_2\chi_i\right] < \log_2|\mathcal{S}_0|\right\}. \tag{17}$$

Theorem 2 (Folk theorem with VPM) *The set of utilities* $\operatorname{conv} u(\mathcal{R}) \cap IR$ *is included in the set of uniform equilibrium utilities for the encoder-assisted repeated game with signals :*

$$\operatorname{conv} u(\mathcal{R}) \cap IR \qquad \subset \qquad NE_{enc}^{\infty}. \tag{18}$$

Moreover, for any utility vector in this set, VPM can be implemented by the encoder $\mathcal{C}$ *and the players* $\mathcal{K}$.

The proof is provided in Sec. B and is based on Theorem 1. The framework of AVS is exploited to characterize the communication possibilities for the encoder. The first feature of the problem is that the coding scheme must be reliable even if one of the players deviates. The second main feature is that the coding scheme must also take into account the private signals received by the players. These two hypotheses allow us to determine the amount of additional information needed from the encoder in order to implement VPM. Interestingly, the proof of Theorem 2 relies on classical grim-trigger strategies but implemented in a blockwise manner and by exploiting VPM and strong typicality [Cover and Thomas(2006)] as a statistical test whose result indicates to every player whether to keep on following the main plan.

4.2 Application to the repeated prisoner's dilemma

We consider a prisoner's dilemma whose matrix form is given by Tab. 1. Let $|\mathcal{A}| = 4$ and $|\mathcal{S}_0| = 3$. Note that the encoder cannot send the action profile profile directly to the players. The goal of this section is to describe the mixed strategies $\mathcal{P}^\star \in \Delta(\mathcal{A})$ that are compatible with the information constraints (9). If this constraint is satisfied, the encoder can compress the sequence of

	L	R
T	(3, 3)	(0, 4)
B	(4, 0)	(1, 1)

Table 1 The prisoner's dilemma in a matrix form.

past actions, encode it into a sequence of public signals and the players can decode the sequence of past action with an error probability that goes to zero when the length of the sequences goes to infinity. Denote $\mathcal{A}_1 = \{T, B\}$ and $\mathcal{A}_2 = \{L, R\}$.

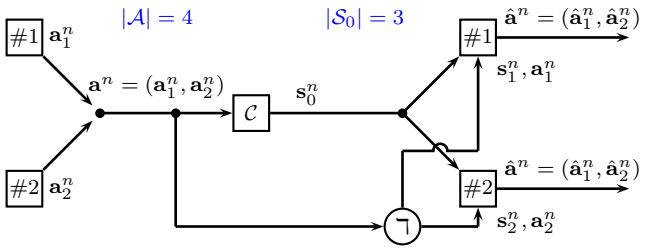

Fig. 2 This figure illustrates the encoder-assisted monitoring structure for the repeated version of the prisoner's dilemma. The Theorem 2 provide a set $\mathcal{R}$ of mixed strategies $\mathcal{P}^\star \in \Delta(\mathcal{A})$ that allow the encoder $\mathcal{C}$ to establish VPM and the players $\mathcal{K}$ to implement an equilibrium strategy.

To have a better understanding on how the results derived in Sec. 3 and 4.1 are exploited here, we consider a particular monitoring structure $\urcorner$ described by Fig. 3 with $\delta \in [0,1]$. This means that if the action $a_{-k} \in \mathcal{A}_{-k}$ was played,

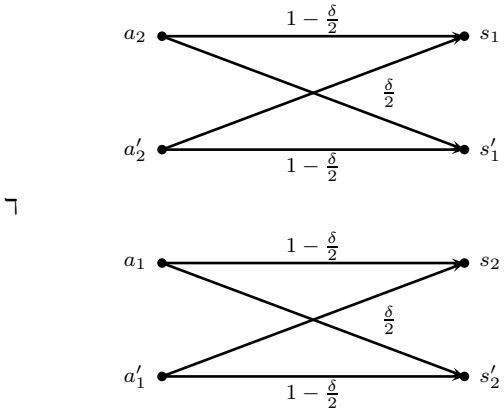

Fig. 3 Private monitoring structure $\urcorner$ that depends on the parameter $\delta \in [0,1]$.

player $k \in \{1,2\}$ observes the right signal $s_k \in \mathcal{S}_k$ with probability $1-\frac{\delta}{2}$ and observes the wrong signal $s_k' \in \mathcal{S}_k$ with probability $\frac{\delta}{2}$. When $\delta = 0$, all the players have perfect monitoring. On the other hand, when $\delta = 1$, they cannot distinguish anything from the signal they observe (trivial monitoring). For this monitoring structure $\urcorner(\mathbf{s}_1, \mathbf{s}_2 | \mathbf{a}_1, \mathbf{a}_2)$ with $\delta \in [0,1]$, we want to determine the set $\text{conv}\, u(\mathcal{R}) \cap IR$ of utility profiles which are compatible with the information constraint (9). For the scenario under investigation, the information constraint (9) for $\delta > 0$ rewrites as :

$$\mathsf{R}^\star = \max_{i\in\mathcal{K}} \left[\max_{\substack{k\in\mathcal{K},\\ a_i\in\mathcal{A}_i}} H(\mathbf{a}_{-i,k}|\mathbf{s}_k(a_i),\mathbf{a}_k) + \log_2 \chi_i \right] \qquad < \log_2 |\mathcal{S}_0|$$

$$\overset{(a)}{\Leftrightarrow} \qquad \max \begin{bmatrix} H(\mathbf{a}_2|\mathbf{s}_1) + \log_2 \chi_1, \\ H(\mathbf{a}_1|\mathbf{s}_2) + \log_2 \chi_2, \end{bmatrix} \qquad < \log_2 |\mathcal{S}_0|$$

$$\overset{(b)}{\Leftrightarrow} \max \begin{bmatrix} \sum_{\substack{a_2\in\mathcal{A}_2,\\ s_1\in\mathcal{S}_1}} \mathcal{P}^\star(a_2)\urcorner(s_1|a_2) \log_2 \frac{\sum_{\tilde{a}_2\in\mathcal{A}_2} \mathcal{P}^\star(\tilde{a}_2)\urcorner(s_1|\tilde{a}_2)}{\mathcal{P}^\star(a_2)\urcorner(s_1|a_2)} + \log_2 \chi_1, \\ \sum_{\substack{a_1\in\mathcal{A}_1,\\ s_2\in\mathcal{S}_2}} \mathcal{P}^\star(a_1)\urcorner(s_2|a_1) \log_2 \frac{\sum_{\tilde{a}_1\in\mathcal{A}_1} \mathcal{P}^\star(\tilde{a}_1)\urcorner(s_2|\tilde{a}_1)}{\mathcal{P}^\star(a_1)\urcorner(s_2|a_1)} + \log_2 \chi_2 \end{bmatrix} \qquad < \log_2 |\mathcal{S}_0|$$

$$\overset{(c)}{\Leftrightarrow} \qquad \max \begin{bmatrix} \mathcal{P}^\star(a_2)(1-\frac{\delta}{2}) \cdot \log_2 \left(\frac{\mathcal{P}^\star(a_2)(1-\frac{\delta}{2}) + \mathcal{P}^\star(a_2')\frac{\delta}{2}}{\mathcal{P}^\star(a_2)(1-\frac{\delta}{2})} \right) \\ + \quad \mathcal{P}^\star(a_2')\frac{\delta}{2} \cdot \log_2 \left(\frac{\mathcal{P}^\star(a_2)(1-\frac{\delta}{2}) + \mathcal{P}^\star(a_2')\frac{\delta}{2}}{\mathcal{P}^\star(a_2')\frac{\delta}{2}} \right) \\ + \quad \mathcal{P}^\star(a_2)\frac{\delta}{2} \cdot \log_2 \left(\frac{\mathcal{P}^\star(a_2)\frac{\delta}{2} + \mathcal{P}^\star(a_2')(1-\frac{\delta}{2})}{\mathcal{P}^\star(a_2)\frac{\delta}{2}} \right) \\ + \mathcal{P}^\star(a_2')(1-\frac{\delta}{2}) \cdot \log_2 \left(\frac{\mathcal{P}^\star(a_2)\frac{\delta}{2} + \mathcal{P}^\star(a_2')(1-\frac{\delta}{2})}{\mathcal{P}^\star(a_2')(1-\frac{\delta}{2})} \right), \\ \mathcal{P}^\star(a_1)(1-\frac{\delta}{2}) \cdot \log_2 \left(\frac{\mathcal{P}^\star(a_1)(1-\frac{\delta}{2}) + \mathcal{P}^\star(a_1')\frac{\delta}{2}}{\mathcal{P}^\star(a_1)(1-\frac{\delta}{2})} \right) \\ + \quad \mathcal{P}^\star(a_1')\frac{\delta}{2} \cdot \log_2 \left(\frac{\mathcal{P}^\star(a_1)(1-\frac{\delta}{2}) + \mathcal{P}^\star(a_1')\frac{\delta}{2}}{\mathcal{P}^\star(a_1')\frac{\delta}{2}} \right) \\ + \quad \mathcal{P}^\star(a_1)\frac{\delta}{2} \cdot \log_2 \left(\frac{\mathcal{P}^\star(a_1)\frac{\delta}{2} + \mathcal{P}^\star(a_1')(1-\frac{\delta}{2})}{\mathcal{P}^\star(a_1)\frac{\delta}{2}} \right) \\ + \mathcal{P}^\star(a_1')(1-\frac{\delta}{2}) \cdot \log_2 \left(\frac{\mathcal{P}^\star(a_1)\frac{\delta}{2} + \mathcal{P}^\star(a_1')(1-\frac{\delta}{2})}{\mathcal{P}^\star(a_1')(1-\frac{\delta}{2})} \right), \end{bmatrix} \qquad < \log_2 |\mathcal{S}_0| - 1$$

where (a) follows from the fact that $\mathbf{a}_2$ and $\mathbf{s}_1$ are independent of $\mathbf{a}_1$ and then the entropy $H(\mathbf{a}_2|\mathbf{s}_1,\mathbf{a}_1)$ reduce to $H(\mathbf{a}_2|\mathbf{s}_1)$. Using the same argument, $H(\mathbf{a}_1|\mathbf{s}_2,\mathbf{a}_2)$ reduce to $H(\mathbf{a}_1|\mathbf{s}_2)$. (b) follow from the definition of the conditional entropy and (c) follow the fact that the chromatic number of the graphs $\mathcal{G}_1$, $\mathcal{G}_2$ of both players are equals to $\chi_1 = \chi_2 = 2$ as soon as $\delta > 0$.

Setting δ to 1, 0.35, 0.31 and 0.2, the above information constraint can be translated into Fig. 4. This figure represents the set of feasible average utility profiles which are both individually rational and compatible with the information constraint (9). Let us interpret these numerical results that depend on the precision parameter $\delta \in [0,1]$ of the private monitoring $\urcorner$.

- *Trivial monitoring:* $\delta = 1$. The players have no information from their private signal, about the actions of their opponent. Theorem 2 show that for some utility vectors represented by the blue hatched region $\mathrm{conv}\, u(\mathcal{R}) \cap IR$, the encoder is able to send to both players, the sequences of past actions (with $|\mathcal{A}| = 4$) using an alphabet of $3 = |\mathcal{S}_0|$ symbols of public signals.
- *Noisy imperfect monitoring:* $0.31 < \delta < 1$. The private signals received by the players reveal a partial information about the past actions of the opponent. Only a portion of the utility region is compatible with the information constraint (9). The virtual perfect monitoring and the equilibrium condition are not always implementable.

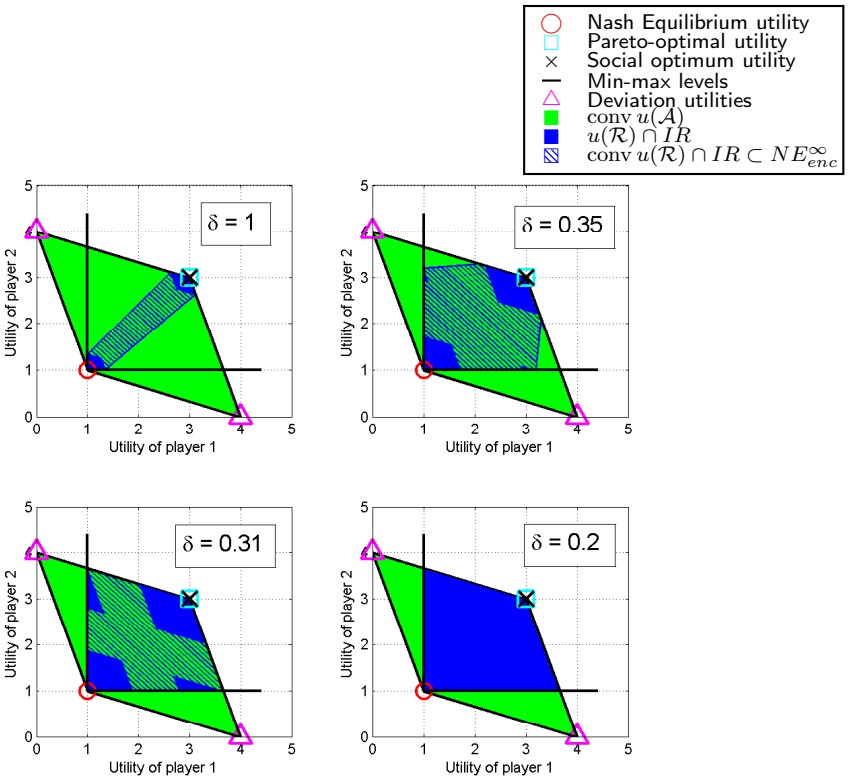

Fig. 4 The repeated version of the prisoner's dilemma is considered. The encoder-assisted monitoring structure of the game is described in Fig. 2 where the private monitoring $\urcorner$ of the players is described by Fig. 3 and depends on a parameter $\delta \in [0,1]$. For $\delta = \{0.2, 0.31, 0.35, 1\}$, the blue region represents the set $u(\mathcal{R}) \cap IR$ of utilities that satisfy the information constraint (9). The hatched blue region represents the convexe hull conv $u(\mathcal{R}) \cap IR$ of the utility that can be supported by a uniform equilibrium strategy (Theorem 2). In that case, the encoder can maintain virtually perfect monitoring even if one of the players deviates. Note that for $\delta \leq 0.31$, the precision of the private monitoring is sufficient to guarantee the same equilibrium utility region as for the Folk theorem with perfect monitoring.

- *Less noisy imperfect monitoring:* $0 < \delta \leq 0.31$. The blue hatched utility region conv $u(\mathcal{R}) \cap IR$ is equal to the utility region of the Folk theorem [Aumann(1981b)] with perfect monitoring conv $u(\mathcal{A}) \cap IR$.
- *Perfect monitoring:* $\delta = 0$. The utility region coincides with the set of feasible and individually rational utilities.

The proposed approach provides an equilibrium strategy $(\sigma^\star, \tau^\star) \in \Sigma \times \mathcal{T}$ that supports any utility profile in the blue hatched region $U^\star = \text{conv}\, u(\mathcal{R}) \cap IR$ of

Fig. 4, while ensuring that the encoder can maintain VPM even in the presence of single deviations.

5 Conclusion

This paper considers games where players have both a private signals they receive through the initial monitoring structure and a public signal which is sent by an encoder. The encoder is assumed to perfectly monitors the played actions and to send a public signal to the players. The purpose of the encoder is to establish virtual perfect monitoring. Technically, the encoder to be designed takes into account the side information at the receiver and possesses the property of resilience to single deviations. It is shown that, the internal information constraint imposes a restriction in terms of feasible utilities in order to establish virtual perfect monitoring and provide an equilibrium utility region (as proved in the case of infinitely repeated games).

The proposed work can be extended in many respects. The targeted monitoring structure can be chosen to be different (e.g., a $2-$connected observation graph or a given public signal). The proposed information constraint might be relaxed by assuming that the encoder sends complementary private signals. An interesting result would be to establish a converse, proving that the information constraint is necessary and sufficient. The i.i.d assumption might be relaxed with the aim to characterize equilibrium utilities which do not assume i.i.d action profiles.

A Proof of Theorem 1

We construct a coding scheme based on graph coloring and statistical tests. Two points have to be considered carefully. First, the side information $s_k^n \in \mathcal{S}_k^n$ may provide some relevant information for player k even if another player $i \in \mathcal{K}$ deviates. Second, the transition probability $\urcorner$ that generates the side information $s_k \in \mathcal{S}_k$ is controlled by the actions $a_k \in \mathcal{A}_k$ of each player $k \in \mathcal{K}$.

Parameter. We choose a parameter $\varepsilon > 0$ such that:

$$\mathsf{R}^\star + 2\varepsilon = \max_{i\in\mathcal{K}} \left[\max_{\substack{k\in\mathcal{K},\\ a_i\in\mathcal{A}_i}} H(\mathbf{a}_{-i,k}|\mathbf{s}_k(a_i),\mathbf{a}_k) + \log_2 \chi_i \right] + 2\varepsilon \leq \log_2 |\mathcal{S}_0|. \tag{19}$$

Encoding function f_0. The encoder proceeds to the statistical test provided by (20) and constructs for a given sequence of actions $a^n = (a_1^n, \ldots, a_K^n) \in \mathcal{A}^n$, the following set :

$$\arg\min_{k\in\mathcal{K}} \sum_{a_{-k}\in\mathcal{A}_{-k}} \left| \frac{N(a_{-k}|a_{-k}^n)}{n} - \mathcal{P}^\star_{-k}(a_{-k}) \right|. \tag{20}$$

It chooses one component $i \in \mathcal{K}$ that minimizes (20). The symbols of the component $i \in \mathcal{K}$ will be encoded using the minimal coloring $\phi_i : \mathcal{A}_i \longrightarrow \Phi_i$ (Def. 5) of the graph $\mathcal{G}_i$ defined by Def. (4). Denote by χ_i the chromatic number of the graph $\mathcal{G}_i$ and

- encode the index of the chosen component $i \in \mathcal{K}$ using $|K|$ sequences $s_0^n \in \mathcal{S}_0^n$ of public signals ;

- encode the sequence of colors $c_i^n \in \Phi_i^n$ that corresponds to the sequence of actions $a_i^n \in \mathcal{A}_i^n$ with at each stage $c_i = \phi_i(a_i)$ using χ_i^n sequences $s_0^n \in \mathcal{S}_0^n$ of public signals.

The other components $a_{-i}^n \in \mathcal{A}_{-i}^n$ will be encoded depending on the transition probability $\top$ and on the sequence $a_i^n \in \mathcal{A}_i^n$. For example, if the symbol $a_i \in \mathcal{A}_i^n$ has been used at a high enough frequency, the sequences of signals $(s_k^n(a_i))_{k\in\mathcal{K}}$, drawn from the transition probability $\top_{a_i} : \mathcal{A}_{-i} \longrightarrow \Delta(\mathcal{S}_k)$, are sufficiently long to use a source coding scheme of the type Slepian and Wolf [Slepian and Wolf(1973)]. Otherwise, the information $a_{-i}^n \in \mathcal{A}_{-i}^n$ should be encoded directly, without any compression. The encoder splits the sequences $s_k^n \in \mathcal{S}_k^n$ into sub-sequences $(s_k^{n_{a_i}})_{a_i\in\mathcal{A}_i}$ indexed by the symbols $a_i \in \mathcal{A}_i$ where $n_{a_i} = N(a_i|a_i^n)$. The sub-sequence $s_k^{n_{a_i}} \in \mathcal{S}_k^{n_{a_i}}$ has length $n_{a_i} \in \mathbb{N}$ and is drawn i.i.d. from the joint probability $\mathcal{P}_{-i}^\star \otimes \top_{a_i} \in \Delta(\mathcal{A}_{-i} \times \mathcal{S})$. The encoder evaluates the partition $(\tilde{\mathcal{A}}_i, \tilde{\mathcal{A}}_i^c)$ of the symbols $a_i \in \mathcal{A}_i$ defined as follows. For each $\varepsilon > 0$, there exists an $\bar{n}_1$ such that the error probability of the Slepian and Wolf [Slepian and Wolf(1973)] coding is upper bounded by $\varepsilon > 0$:

- $a_i \in \tilde{\mathcal{A}}_i$, if $N(a_i|a_i^n) = n_{a_i} \leq \bar{n_1}$ and then the sequence $a_{-i}^{n_{a_i}} \in \mathcal{A}_{-i}^{n_{a_i}}$ is encoded with $|\mathcal{A}_{-i}|^{n_{a_i}}$ sequences $s_0^n \in \mathcal{S}_0^n$ of public signals.
- $a_i \in \tilde{\mathcal{A}}_i^c$ if $N(a_i|a_i^n) = n_{a_i} > \bar{n_1}$ and then the sequence $a_{-i}^{n_{a_i}} \in \mathcal{A}_{-i}^{n_{a_i}}$ is encoded using the "random binning technique" of Slepian and Wolf [Slepian and Wolf(1973)].

The random binning technique [Slepian and Wolf(1973)] consists in randomly assign the $2^{n_{a_i} H(\mathbf{a}_{-i})}$ typical sequences $a_{-i}^{n_{a_i}} \in \mathcal{A}_{-i}^{n_{a_i}}$ to one of the $2^{n_{a_i}(\max_{k\in\mathcal{K}} H(\mathbf{a}_{-i,k}|\mathbf{s}_k(a_i),\mathbf{a}_k)+\varepsilon)}$ bin. Note that $H(\mathbf{a}_{-i,k}|\mathbf{s}_k(a_i),\mathbf{a}_k) = H(\mathbf{a}_{-i}|\mathbf{s}_k(a_i),\mathbf{a}_k)$. Each bin $\mathcal{B}(s_0^n)$ is indexed by a sequence $s_0^n \in \mathcal{S}_0^n$ of public signals and contains $2^{n_{a_i}(\min_{k\in\mathcal{K}} I(\mathbf{a}_{-i};\mathbf{s}_k(a_i),\mathbf{a}_k)-\varepsilon)}$ typical sequences $a_{-i}^{n_{a_i}} \in \mathcal{A}_{-i}^{n_{a_i}}$. The encoder $\mathcal{C}$ observes a sequence of realized actions $a_{-i}^{n_{a_i}} \in \mathcal{A}_{-i}^{n_{a_i}}$. If this sequence is typical, then it send to all the players $\mathcal{K}$, the sequence of public signals $s_0^n \in \mathcal{S}_0^n$ corresponding to the bin containing the sequence $a_{-i}^{n_{a_i}} \in \mathcal{B}(s_0^n)$. If it is not, the encoder $\mathcal{C}$ declares an error.

Decoding function g_k of player $k \in \mathcal{K}$. The decoding player receives the index $i \in \mathcal{K}$ of the player chosen by the statistical test (20). Using the appropriate codebook, it decodes separately the information regarding the component $i \in \mathcal{K}$ and the other components $j \in \mathcal{K}\backslash\{i\}$.

- Knowing component $i \in \mathcal{K}$ chosen by the statistical test, the side information $s_k \in \mathcal{S}_k$ and the color $c_i \in \Phi_i$, the decoding player $k \in \mathcal{K}$ decodes a unique stage symbol $a_i \in \mathcal{A}_i$ for component $i \in \mathcal{K}$.

The decoding player $k \in \mathcal{K}$ knows the entire sequence of actions $a_i^n \in \mathcal{A}_i^n$ and it characterizes the partition $\tilde{\mathcal{A}}_i$ and $\tilde{\mathcal{A}}_i{}^c$ of the set of symbols $\mathcal{A}_i$.

- For the transition $\top_{a_i}$, controlled by the symbol $a_i \in \tilde{\mathcal{A}}_i$, the sequence of symbols $a_{-i}^{n_{a_i}} \in \mathcal{A}_{-i}^{n_{a_i}}$ is directly decoded.
- For the transition $\top_{a_i}$, controlled by the symbol $a_i \in \tilde{\mathcal{A}}_i{}^c$, the sequence of actions $a_{-i}^{n_{a_i}} \in \mathcal{A}_{-i}^{n_{a_i}}$ is decoded using Slepian and Wolf decoding [Slepian and Wolf(1973)]. The decoding player $k \in \mathcal{K}$ find into the bin $\mathcal{B}(s_0^n)$ corresponding to the sequence of public signals $s_0^n \in \mathcal{S}_0^n$, a sequence $a_{-i}^{n_{a_i}} \in \mathcal{A}_{-i}^{n_{a_i}}$ which is jointly typical with the sequence of side information $s_k^n(a_i) \in \mathcal{S}_k^n$ for the probability distribution $\mathcal{P}_{-i} \otimes \top_{a_i} \in \Delta(\mathcal{A}_{-i} \times \mathcal{S}_k)$.

Cardinality of $\mathcal{S}_0^n$. Let $\bar{n_2} > \frac{\log|K| + \bar{n_1}|\mathcal{A}_i|\log|A_{-i}|}{\varepsilon}$. Then for all $n \geq \bar{n_2}$, the cardinality of the set of sequences $|\mathcal{S}_0|^n$ is greater than the number of sequences of the coding scheme

$2^{n(R^\star+3\varepsilon)}$.

$$\frac{\log\left(|\mathcal{K}|\cdot\chi_i^n\cdot|\mathcal{A}_{-i}|^{\sum_{a_i\in\tilde{\mathcal{A}}_i}n_{a_i}}\cdot\prod_{a_i\in\tilde{\mathcal{A}}_i{}^c}2^{n_{a_i}(\max_{k\in\mathcal{K}}H(\mathbf{a}_{-i,k}|\mathbf{s}_k(a_i),\mathbf{a}_k)+\varepsilon)}\right)}{n}$$

$$\leq \frac{\log|K|}{n}+\log\chi_i+\frac{n\bar{1}|\tilde{\mathcal{A}}_i|}{n}\log|\mathcal{A}_{-i}|+\sum_{a_i\in\tilde{\mathcal{A}}_i{}^c}\frac{n_{a_i}}{n}(\max_{k\in\mathcal{K}}H(\mathbf{a}_{-i,k}|\mathbf{s}_k(a_i),\mathbf{a}_k)+\varepsilon)$$

$$\leq \max_{i\in\mathcal{K}}\left[\max_{\substack{k\in\mathcal{K},\\ a_i\in\mathcal{A}_i}}H(\mathbf{a}_{-i,k}|\mathbf{s}_k(a_i),\mathbf{a}_k)+\log\chi_i\right]+\frac{\log|K|+n\bar{1}|\mathcal{A}_i|\log|A_{-i}|}{n}+\varepsilon$$

$$\leq \max_{i\in\mathcal{K}}\left[\max_{\substack{k\in\mathcal{K},\\ a_i\in\mathcal{A}_i}}H(\mathbf{a}_{-i,k}|\mathbf{s}_k(a_i),\mathbf{a}_k)+\log\chi_i\right]+2\varepsilon$$

$$=\mathsf{R}^\star+3\varepsilon \tag{21}$$

$$\leq \log_2|\mathcal{S}_0|. \tag{22}$$

Error probability. Suppose that player $i\in\mathcal{K}$ chooses his sequence of actions $a_i^n\in\mathcal{A}_i^n$ with an arbitrary sequence of distribution. There are two possibilities. First, the statistical test (20) returns the deviating player $i\in\mathcal{K}$. Second the statistical test returns another player $j\neq i$.

- Suppose that the statistical test (20) returns the deviating player $i\in\mathcal{K}$. In that case, the "random binning technique" of Slepian and Wolf [Slepian and Wolf(1973)] guarantees that for all $a_i\in\tilde{\mathcal{A}}_i^c$ the sequence of vectors of actions $a_{-i}^{n_{a_i}}\in\mathcal{A}_{-i}^{n_{a_i}}$ is perfectly reconstructed with large probability. Let us define the following events:

$$E_1=\cup_{\substack{k\in\mathcal{K},\\ a_i\in\mathcal{A}_i}}\left\{(\mathbf{a}_{-i}^{n_{a_i}},\mathbf{s}_k^{n_{a_i}})\notin A_\varepsilon^{\star n}(\mathcal{A}_{-i}\times\mathcal{S}_k)\right\}. \tag{23}$$

  There exists a player $k\in\mathcal{K}$ for which the random sequences of actions and private signals $(\mathbf{a}_{-i}^{n_{a_i}},\mathbf{s}_k^{n_{a_i}})\in\mathcal{A}_{-i}^{n_{a_i}}\times\mathcal{S}_k^{n_{a_i}}$ are not typical.

$$E_2=\cup_{\substack{k\in\mathcal{K},\\ a_i\in\mathcal{A}_i}}\left\{\exists\mathbf{a}_{-i}^{'n_{a_i}}\neq\mathbf{a}_{-i}^{n_{a_i}}\in\mathcal{B}(s_0^{n_{a_i}}),\ (\mathbf{a}_{-i}^{'n_{a_i}},\mathbf{s}_k^{n_{a_i}}(a_i),\mathbf{a}_k^{n_{a_i}})\in A_\varepsilon^{\star n}(\mathcal{A}_{-i}\times\mathcal{S}_k)\right\}. \tag{24}$$

  There exists another sequence $\mathbf{a}_{-i}^{'n_{a_i}}$ in the bin $\mathcal{B}(s_0^{n_{a_i}})$ corresponding to the sequence of public signals $s_0^{n_{a_i}}\in\mathcal{S}_0^{n_{a_i}}$ that is jointly typical with the sequences of private signals $\mathbf{s}_k^{n_{a_i}}(a_i)$ and actions $\mathbf{a}_k^{n_{a_i}}$ of the player $k\in\mathcal{K}$.

  - From Lemma 6 of App. C, the error probability $\mathcal{P}(E_1)$ is lower than $\varepsilon\cdot K\cdot|\mathcal{A}_i|>0$ as soon as n is sufficiently large.
  - From Lemma 7 of App. C, the expected error probability $\mathbb{E}_\lambda[\mathcal{P}(E_2)]$ of the random code $\mu\in\Delta(\Lambda(n))$ is lower than $\varepsilon\cdot K\cdot|\mathcal{A}_i|>0$ as soon as n is sufficiently large and the condition (25) is satisfied.

$$|\mathcal{B}(s_0^{n_{a_i}})|\leq 2^{n(\min_{k\in\mathcal{K}}I(\mathbf{a}_{-i};\mathbf{s}_k(a_i),\mathbf{a}_k)-\varepsilon)} \tag{25}$$

    Lemma 7 applies because the random sequence $\mathbf{a}_{-i}^{'n_{a_i}}$ is generated independently of the random sequences $(\mathbf{s}_k^{n_{a_i}}(a_i),\mathbf{a}_k^{n_{a_i}})$. This ensures the existence of a code $\lambda\in\Lambda(n)$ such that the error probability $\mathcal{P}_\lambda(E_2)\leq 2\varepsilon$ is upper bounded.

- Suppose that the the statistical test returns another player $j \neq i \in \mathcal{K}$. This implies the following inequality:

$$\sum_{a_{-j} \in \mathcal{A}_{-j}} \left| \frac{N(a_{-j}|a^n_{-j})}{n} - \mathcal{P}_{-j}(a_{-j}) \right| \leq \sum_{a_{-i} \in \mathcal{A}_{-i}} \left| \frac{N(a_{-i}|a^n_{-i})}{n} - \mathcal{P}^\star_{-i}(a_{-i}) \right|. \tag{26}$$

  For every player $j \in \mathcal{K}\backslash\{i\}$, the sequence $a^n_j \in \mathcal{A}^n_j$ is drawn i.i.d. from stage to stage with the distribution $\mathcal{P}^\star_j \in \Delta(\mathcal{A}_j)$. From Lemma 6, these action sequences are typical with large probability as n goes to infinity. Then, the sequence of actions $a^n_i \in \mathcal{A}^n_i$ is typical with large probability and then correctly encoded and decoded. There exists n sufficiently large such that the error probability $\mathcal{P}_e(\lambda) \leq \varepsilon$ is upper bounded.

We therefore proved the existence a code $\lambda \in \Lambda(n)$ such that the error probability of the code $\mathcal{P}_e(\lambda) \leq 2\varepsilon \cdot K \cdot |\mathcal{A}_i|$ is upper bounded.

B Proof of Theorem 2

We prove the following inclusion $\text{conv}\, u(\mathcal{R}) \cap IR \subset NE^\infty_{enc}$. First, we consider a utility vector $U \in u(\mathcal{R}) \cap IR$ and provide a pair of strategies for the encoder and the players $(\sigma^\star, \tau^\star) \in \Sigma \times \mathcal{T}$ that forms a uniform equilibrium (see Def. 6. The first condition (i) is satisfied when the asymptotic utility of the strategies $(\sigma^\star, \tau^\star) \in \Sigma \times \mathcal{T}$ converges toward the utility U. The second condition (ii) is satisfied when no unilateral deviation $\tau'_k \in \mathcal{T}_k$ provides to player $k \in \mathcal{K}$ a gain larger than $\epsilon > 0$.

B.1 Construction of strategies $(\sigma^\star, \tau^\star) \in \Sigma \times \mathcal{T}$

B.1.1 Block coding scheme

The $T > 0$ stages of the repeated game are divided into B blocks of stages of length n, represented by Fig. 5. Denote $\mathcal{B}$ the set of blocks, $b \in \mathcal{B}$ the index of one block and $B = |\mathcal{B}| \in \mathbb{N}$ the number of blocks. Denote $s^n_k(b) \in \mathcal{S}^n_k$ the sequence of signals received during the block $b \in \mathcal{B}$. Fix the parameter $\epsilon > 0$ and let us describe the strategies $(\sigma^\star, \tau^\star) \in \Sigma \times \mathcal{T}$ that satisfy both conditions (27) and (28) for all $T \geq \bar{T}$.

$$|\gamma^T_k(\sigma^\star, \tau^\star) - U^\star_k| \leq \epsilon, \qquad \forall k \in \mathcal{K}, \tag{27}$$

$$\gamma^T_k(\sigma^\star, \tau^\star) + \epsilon \geq \gamma^T_k(\sigma^\star, \tau'_k, \tau^\star_{-k}), \qquad \forall k \in \mathcal{K}, \quad \forall \tau'_k \in \mathcal{T}_k. \tag{28}$$

Suppose that the number of blocks $B \in \mathbb{N}$ satisfies condition (29) :

$$B \geq \frac{8 \cdot \max_{a \in \mathcal{A}} |u_k(a)|}{\epsilon}. \tag{29}$$

B.1.2 Strategy of the encoder $\sigma^\star \in \Sigma$

The coding strategy $\sigma^\star \in \Sigma$ consists in sending a sequence of public signals $\mathbf{s}^n_0 \in \mathcal{S}^n_0$ to each player so that they can reconstruct the sequence $\mathbf{a}^n \in \mathcal{A}^n$ of past actions. In order to communicate, the encoder $\mathcal{C}$ and the players $\mathcal{K}$ implement a code $\lambda = (f_0, (g_k)_{k \in \mathcal{K}})$ investigated in Sec. 3 and defined by :

$$\begin{array}{lll} f_0 : & \mathcal{A}^n & \longrightarrow \mathcal{S}^n_0 \\ g_k : \mathcal{S}^n_0 \times \mathcal{S}^n_k \times \mathcal{A}^n_k & \longrightarrow \mathcal{A}^n, & \forall k \in \mathcal{K}. \end{array} \tag{30}$$

Condition $U \in \operatorname{conv} u(\mathcal{R}) \cap IR$ implies that the probability distribution $\mathcal{P}^\star \in \prod_{k\in\mathcal{K}} \Delta(\mathcal{A}_k)$ belong to the set $\mathcal{R}$ described by (17) and satisfies the condition (9) of the Theorem 1. This coding result implies that for all $\varepsilon > 0$, there exists a parameter $n \in \mathbb{N}$ and a code $\lambda \in \Lambda(n)$ with $\lambda = (f_0, (g_k)_{k\in\mathcal{K}})$ such that the error probability of the coding scheme is upper bounded by ε. Denote by $\hat{a}^n(k) \in \mathcal{A}^n$ the sequence of actions obtained as output by the decoder $k \in \mathcal{K}$.

$$\mathcal{P}_e(\lambda) = \sum_{k\in\mathcal{K}} \max_{i\in\mathcal{K}} \max_{v_i \in \Delta(\mathcal{A}_i^\infty)} \mathcal{P}_{v_i}(\mathbf{a}^n \neq g_k(\mathbf{s}_k^n, \mathbf{s}_0^n, \mathbf{a}_k^n)). \tag{31}$$

The strategy of the encoder $\sigma^\star \in \Sigma$ is built as follows. At the beginning of the block $b \in \mathcal{B}$ with $b \geq 2$, the encoder $\mathcal{C}$ observes the sequence of actions $a^n(b-1) \in \mathcal{A}^n$ over the block $b-1 \in \mathcal{B}$ and choose the sequence of public signals $s_0^n(b)$ over block $b \in \mathcal{B}$ using the encoding function f_0 (30) provided by the code λ that satisfies the condition (31).

$$s_0^n(b) = f_0\Big(a^n(b-1)\Big) \in \mathcal{S}_0^n. \tag{32}$$

Over the first block $b_1 \in \mathcal{B}$, the encoder send an arbitrary sequence $s_0(b_1) \in \mathcal{S}_0^n$.

B.1.3 Decoding scheme

At the end of the block $b \in \mathcal{B}$ with $b \geq 3$, player k implements the decoding function g_k (30) provided by the code λ that satisfies the condition (31). The player $k \in \mathcal{K}$ recalls his own actions $a_k^n(b-1) \in \mathcal{A}_k^n$ and observes the sequences of private signals $s_k^n(b-1) \in \mathcal{S}_k^n$ and public signals $s_0^n(b) \in \mathcal{S}_0^n$ sent by the encoder $\mathcal{C}$. The player $k \in \mathcal{K}$ evaluates the sequence $\hat{\mathbf{a}}^n(k, b-1)$ of actions of block $b-1 \in \mathcal{B}$ using the decoding function g_k.

$$\hat{\mathbf{a}}^n(k, b-1) = g_k\Big(s_k^n(b-1), s_0^n(b), a_k^n(b-1)\Big) \in \mathcal{A}^n. \tag{33}$$

Condition (31) guarantees that at the beginning of block $b+1 \in \mathcal{B}$, each player $k \in \mathcal{K}$ observes the sequence of actions $\hat{\mathbf{a}}(b-1) \in \mathcal{A}$ of the other players during the block $b-1 \in \mathcal{B}$ with an error probability arbitrarily low. Over the two first blocks $b_1, b_2 \in \mathcal{B}$, no decoding strategy is implemented.

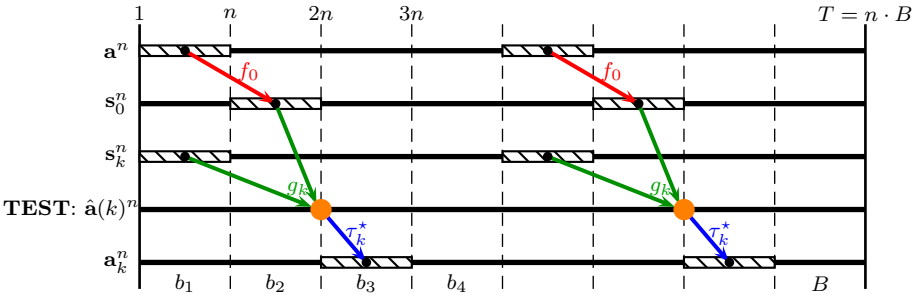

Fig. 5 The strategies of the encoder $\mathcal{C}$ and the players $\mathcal{K}$ $(\sigma^\star, \tau^\star) \in \Sigma \times \mathcal{T}$ are described at sections B.1.2 and B.1.7. The actions $\mathbf{a}^n(b)$ over block $b \in \mathcal{B}$ are encoded over the next block $b+1 \in \mathcal{B}$ into a sequence of public signals $\mathbf{s}_0^n(b+1)$. At the end of block $b+1 \in \mathcal{B}$, player $k \in \mathcal{K}$ decode the sequence of actions $\hat{\mathbf{a}}^n(b)$ over block $b \in \mathcal{B}$ from the sequences of signals $\mathbf{s}_0^n(b+1)$ and $\mathbf{s}_k^n(b)$. Player $k \in \mathcal{K}$ performs a statistical test in order to detect the possible unilateral deviations. The result of this statistical test determines the sequence of actions $\mathbf{a}_k^n(b+2)$ player $k \in \mathcal{K}$ will play during the block $b+2 \in \mathcal{B}$.

B.1.4 Statistical test

Each player $k \in \mathcal{K}$ performs a statistical test at the beginning of each block $b+1 \in \mathcal{B}$. Define the event $\mathbf{E}_k^i(b+1)$ using the set of typical sequences $A_\varepsilon^{\star n}(\mathcal{P}_i^\star)$ stated by the definition 9.

$$\mathbf{E}_k^i(b+1) = \begin{cases} 0 \text{ if } \hat{a}_i^n(b-1) \in A_\varepsilon^{\star n}(\mathcal{P}_i^\star), \\ 1 \text{ if } \hat{a}_i^n(b-1) \notin A_\varepsilon^{\star n}(\mathcal{P}_i^\star). \end{cases} \tag{34}$$

When $\mathbf{E}_k^i(b+1) = 1$, player $k \in \mathcal{K}$ declare player $i \in \mathcal{K}$ deviates from the prescribed strategy $\tau_i^\star \in \mathcal{T}_i$, during block $b-1 \in \mathcal{B}$.

B.1.5 Main plan

The main plan consists in playing the same mixed action $\mathcal{P}^\star$ i.i.d. from stage to stage.

$$\mathcal{P}^\star(a^t) = \mathcal{P}^\star(a_1) \otimes \ldots \otimes \mathcal{P}^\star(a_K) \in \prod_{k \in \mathcal{K}} \Delta(\mathcal{A}_k), \qquad \forall t \geq 1. \tag{35}$$

B.1.6 Punishment plan for player $i \in \mathcal{K}$

The punishment plan $\bar{\mathcal{P}}(i) = (\bar{\mathcal{P}}_k(i))_{k \neq i} \in \prod_{k \neq i} \Delta(\mathcal{A}_k)$ of player $i \in \mathcal{K}$ consists of a vector of mixed actions of other players that minimize the utility of player $i \in \mathcal{K}$.

$$\bar{\mathcal{P}}(i) = \left(\bar{\mathcal{P}}_k(i)\right)_{k \neq i} \in \operatorname{argmin}_{\mathcal{P}_{-i} \in \prod_{k \neq i} \Delta(\mathcal{A}_k)} \left[\max_{\mathcal{P}_i \in \Delta(\mathcal{A}_i)} \mathbb{E}_{\mathcal{P}_i, \mathcal{P}_{-i}} \left[u_k(a_i, a_{-i}) \right] \right], \quad \forall i \in \mathcal{K}. \tag{36}$$

The punishment plan for player $i \in \mathcal{K}$ by player $k \in \mathcal{K}$ is denoted $\bar{\mathcal{P}}_k(i) \in \Delta(\mathcal{A}_k)$ and is given by (36). If all the players $k \neq i$ play the strategy $\bar{\mathcal{P}}(i) = (\bar{\mathcal{P}}_k(i))_{k \neq i}$, the player $i \in \mathcal{K}$ cannot obtain a utility greater than his min-max level $\upsilon_i \in \mathbb{R}$ characterized by (15).

B.1.7 Equilibrium strategy $\tau^\star = (\tau_k^\star)_{k \in \mathcal{K}} \in \mathcal{T}$

At the beginning of each block $b \geq 3 \in \mathcal{B}$, the equilibrium strategy is described as follows:

- Player k implements the decoding scheme (Sec. B.1.3) and reconstructs the actions $\hat{a}_{-k}(b-2) \in \mathcal{A}_{-k}^n$ played by the other players $j \neq k$ during block $b-2 \in \mathcal{B}$.
- Player k implements the statistical test $\mathbf{E}_i^k$ defined section B.1.4, in order to detect possible unilateral deviations.
- If the statistical test is negative, $(\forall b' \leq b,\ \forall i \in \mathcal{K}, \quad \mathbf{E}_k^i(b) = 0)$, then player $k \in \mathcal{K}$ play the main plan $\mathcal{P}_k^\star \in \Delta(\mathcal{A}_k)$ stated section B.1.5 during every stage of block $b \in \mathcal{B}$.
- If the statistical test is positive, $(\exists b' \leq b,\ \exists i \in \mathcal{K}, \quad \mathbf{E}_k^i(b) = 1)$, then player $k \in \mathcal{K}$ play the punishment plan $\bar{\mathcal{P}}_k(i) \in \Delta(\mathcal{A}_k)$ stated section B.1.6 corresponding to the player $i \in \mathcal{K}$ until the end of the last block $B \in \mathcal{B}$. If several deviations are detected simultaneously $\mathbf{E}_k^i(b) = \mathbf{E}_j(b) = 1$, then player $k \in \mathcal{K}$ punishes anyone of those players who is the smaller, according to a total order over $\mathcal{K}$, previously fixed.

Over the first two blocks $b_1, b_2 \in \mathcal{B}$, players $\mathcal{K}$ play the main plan $\mathcal{P}^\star \in \Delta(\mathcal{A})$. The equilibrium strategy $\tau^\star = (\tau_k^\star)_{k \in \mathcal{K}} \in \mathcal{T}$ is defined at each stage $t \geq 1$ as follows:

$$\tau_k^{\star t}(h^t) = \begin{cases} \mathcal{P}_k^\star & \in \Delta(\mathcal{A}_k) \\ \bar{\mathcal{P}}_k(i) & \in \Delta(\mathcal{A}_k) \end{cases} \qquad \begin{array}{ll} \text{while } \mathbf{E}_k^i(b) = 0, & \forall i \neq k,\ \forall b \leq \lfloor \frac{t}{n}, \rfloor \\ \text{otherwise.} & \end{array} \tag{37}$$

B.2 Condition (i) of definition 6: convergence of the utilities

Let us fix a parameter $\epsilon > 0$ and prove that there exists a $\bar{T} > 1$ such that for all $T \geq \bar{T}$, the utilities of the encoder and the players $\mathcal{K}$ $(\sigma^\star, \tau^\star) \in \Sigma \times \mathcal{T}$ defined in Sec. B.1.2 and B.1.7, are ϵ-closed to the utility $U \in \operatorname{conv} u(\mathcal{R}) \cap IR$. Remark that $\epsilon > 0$ and $\varepsilon > 0$ are two distinct parameters. Define the following event:

$$\mathbf{E} = \begin{cases} 1 \text{ if } & \exists b \in \mathcal{B}, \exists i, k \in \mathcal{K} \text{ such that } \mathbf{E}_k^i(b) = 1, \\ 0 \text{ otherwise.} & \end{cases} \tag{38}$$

When $\mathbf{E} = 0$, then no unilateral deviation is detected during the course of the game.

Lemma 1 *Suppose that the encoder $\mathcal{C}$ and the players $\mathcal{K}$ implements the strategies $(\sigma^\star, \tau^\star) \in \Sigma \times \mathcal{T}$. Then for all $\varepsilon > 0$, there exists a block length $n_1 \in \mathbb{N}$, such that for all $n \geq n_1$, the probability of event $\boldsymbol{E} = 1$ is bounded as follows:*

$$\mathcal{P}(\boldsymbol{E} = 1) \leq 2\varepsilon \cdot B \cdot K^2. \tag{39}$$

The result of Lemma 1 is useful for the proof of Lemma 2.

Lemma 2 *Suppose that the encoder $\mathcal{C}$ and the players $\mathcal{K}$ implement the strategies $(\sigma^\star, \tau^\star) \in \Sigma \times \mathcal{T}$. Then for all $\varepsilon > 0$, there exists a block length $n_1 \in \mathbb{N}$, such that for all $n \geq n_1$, the expected utility satisfies the following equation :*

$$\left| \gamma_k^T(\sigma^\star, \tau^\star) - \mathbb{E}_{\mathcal{P}^\star}\Big[u_k(\boldsymbol{a}_k, \boldsymbol{a}_{-k}) \Big] \right| \leq 4\varepsilon \cdot \max_{a \in \mathcal{A}} |u_k(a)| \cdot B \cdot K^2, \qquad \forall k \in \mathcal{K}. \tag{40}$$

For the parameter $\epsilon > 0$ and a fixed number of block $B \in \mathbb{N}$, there exists a parameter $\varepsilon > 0$ and a block length $n_1 \in \mathbb{N}$ such that $4\varepsilon \cdot \max_{a \in \mathcal{A}} |u_k(a)| \cdot B \cdot K^2 \leq \epsilon$. From Lemma 2, the strategy defined over $T = n \cdot B$ stages induce, for each player $k \in \mathcal{K}$, a utility that satisfies:

$$\left| \gamma_k^T(\sigma^\star, \tau^\star) - U_k^\star \right| \leq \epsilon, \qquad \forall k \in \mathcal{K}. \tag{41}$$

By repeating the strategies cyclically, we prove that there exists a $\bar{T} \geq \frac{N \cdot B}{\epsilon}$ such that for all $T' \geq \bar{T}$ and for all players $k \in \mathcal{K}$, the expected T' stage utility $\gamma^{T'}(\sigma^\star, \tau^\star)$ is ϵ-closed of utility $U \in \operatorname{conv} u(\mathcal{R}) \cap IR$. Strategies $(\sigma^\star, \tau^\star) \in \Sigma \times \mathcal{T}$ satisfy the condition (i) of definition 6.

Proof. [Lemma 1] Denote $\hat{\mathbf{a}}_i^k(b)$ the sequence of actions of player $i \in \mathcal{K}$ observed by player $k \in \mathcal{K}$ over block $b \in \mathcal{B}$. For all $\varepsilon > 0$, there exists $n_1 \in \mathbb{N}$ such that for all $n \geq n_1$, we have:

$$\mathcal{P}\left(\mathbf{a}_i^n(b) \notin A_\varepsilon^{\star n}(\mathcal{P}_i^\star) \middle| \cap_{i,k \in \mathcal{K}} \Big\{ \mathbf{E}_k^i(b-1) = 0, \ldots, \mathbf{E}_k^i(b_1) = 0 \Big\} \right) \leq \varepsilon, \ \forall i, k \in \mathcal{K}, \ \forall b \in \mathcal{B}, \tag{42}$$

$$\mathcal{P}\left(\hat{\mathbf{a}}_i^n(k,b) \neq \mathbf{a}_i^n(b) \middle| \cap_{i,k \in \mathcal{K}} \Big\{ \mathbf{E}_k^i(b-1) = 0, \ldots, \mathbf{E}_k^i(b_1) = 0 \Big\} \right) \leq \varepsilon, \ \forall i, k \in \mathcal{K}, \ \forall b \in \mathcal{B}. \tag{43}$$

Equations (42) and (43) come from the definition of strategies $(\sigma^\star, \tau^\star) \in \Sigma \times \mathcal{T}$. When no deviation is detected, the players implement the main plan (Sec. B.1.5) by playing i.i.d. the mixed action $\mathcal{P}^\star \in \Delta(\mathcal{A})$.

Equation (42) is a consequence of Lemma 6 for the typical sequences and (43) is a consequence of the coding result stated by Theorem 1 in Sec. 3.2 for an i.i.d. information source $\mathcal{P}^\star \in \Delta(\mathcal{A})$. More precisely, this inequality is a consequence of (31) that guarantees at the beginning of block $b \in \mathcal{B}$, the players observe the sequence of actions played by the other players over the block $b - 2 \in \mathcal{B}$ with probability $1 - \varepsilon$.

Let us evaluate the probability of event $\mathbf{E} = 1$.

$$\mathcal{P}(\mathbf{E}=1) = \mathcal{P}\Big(\cup_{\substack{b\in\mathcal{B},\\ i,k\in\mathcal{K}}} \Big\{\mathbf{E}_k^i(b)=1\Big\}\Big) \tag{44}$$

$$\leq \sum_{b\in\mathcal{B}} \mathcal{P}\Big(\cup_{i,k\in\mathcal{K}} \mathbf{E}_k^i(b)=1 \Big| \cap_{i,k\in\mathcal{K}} \Big\{\mathbf{E}_k^i(b-1)=0,\dots,\mathbf{E}_k^i(b_1)=0\Big\}\Big) \tag{45}$$

$$\leq \sum_{\substack{b\in\mathcal{B},\\ i,k\in\mathcal{K}}} \mathcal{P}\Big(\Big\{\mathbf{a}_i^n(b)\notin A_\varepsilon^{\star n}(\mathcal{P}_i^\star)\Big\}\cup\Big\{\hat{\mathbf{a}}_i^n(k,b)\neq\mathbf{a}_i^n(b)\Big\}\Big| \cap_{i,k\in\mathcal{K}} \Big\{\mathbf{E}_k^i(b-1)=0,\dots,\mathbf{E}_k^i(b_1)=0\Big\}\Big) \tag{46}$$

$$\leq \sum_{\substack{b\in\mathcal{B},\\ i,k\in\mathcal{K}}} \mathcal{P}\Big(\mathbf{a}_i^n(b)\notin A_\varepsilon^{\star n}(\mathcal{P}_i^\star)\Big| \cap_{i,k\in\mathcal{K}} \Big\{\mathbf{E}_k^i(b-1)=0,\dots,\mathbf{E}_k^i(b_1)=0\Big\}\Big)$$

$$+ \sum_{\substack{b\in\mathcal{B},\\ i,k\in\mathcal{K}}} \mathcal{P}\Big(\hat{\mathbf{a}}_i^n(k,b)\neq\mathbf{a}_i^n(b)\Big| \cap_{i,k\in\mathcal{K}} \Big\{\mathbf{E}_k^i(b-1)=0,\dots,\mathbf{E}_k^i(b_1)=0\Big\}\Big) \tag{47}$$

$$\leq 2\varepsilon\cdot B\cdot K^2. \tag{48}$$

Equality (44) comes from the definition of error event provided by (38).
Inequality (45) comes from the property $\mathcal{P}(A\cup B)=\mathcal{P}(A)+\mathcal{P}(B|A^c)\cdot\mathcal{P}(A^c)$.
Inequalities (46) and (47) comes from the inequality of Boole.
Inequality (48) comes from the inequalities (42) and (43).

As a conclusion, for all $\varepsilon > 0$, there exists a $n_1 \in \mathbb{N}$ such that for all $n \geq n_1$, the condition (39) is satisfied. □

Proof. [Lemma 2] When the event $\mathbf{E} = 0$ occurs, then all the statistical tests of players $\mathcal{K}$ at the beginning of each block $b \in \mathcal{B}$ are negative (i.e. $\mathbf{E}_k^i(b) = 0$, $\forall b \in \mathcal{B}$, $\forall i,k \in \mathcal{K}$). In this case, the strategy $\tau^\star \in \mathcal{T}$ indicate that the sequence of actions are generated with the same mixed strategy $\mathcal{P}^\star \in \prod_{k\in\mathcal{K}} \Delta(\mathcal{A}_k)$ from stage to stage. From the proof of Lemma 1, for all $\varepsilon > 0$, there exists $n_1 \in \mathbb{N}$ such that for all $n \geq n_1$, the sequences of block actions are typical with large probability. More precisely, because $T = n\cdot B \geq n \geq n_1$, the sequences of actions $\mathbf{a}^T$ are typical with large probability.

$$\mathcal{P}\Big(\mathbf{a}^n \in A_\varepsilon^{\star n}(\mathcal{P}^\star)\Big|\mathbf{E}=0\Big) \leq \varepsilon, \tag{49}$$

$$\implies \mathcal{P}\Big(\mathbf{a}^T \in A_\varepsilon^{\star T}(\mathcal{P}^\star)\Big|\mathbf{E}=0\Big) \leq \varepsilon. \tag{50}$$

Suppose that $n \geq n_1$ defined from Lemma 1. Recall the definition of the typical sequences and some implications thereof:

$$a^T \in A_\varepsilon^{\star T}(\mathcal{P}^\star) \tag{51}$$

$$\iff \sum_{a\in\mathcal{A}} \left|\frac{N(a|a^t)}{T} - \mathcal{P}^\star(a)\right| \leq \varepsilon \tag{52}$$

$$\implies \sum_{a\in\mathcal{A}} \left|\frac{N(a|a^t)}{T}u_k(a) - \mathcal{P}^\star(a)u_k(a)\right| \leq \varepsilon\cdot\max_{a\in\mathcal{A}}|u_k(a)| \tag{53}$$

$$\implies \left|\sum_{a\in\mathcal{A}} \frac{N(a|a^t)}{T}u_k(a) - \sum_{a\in\mathcal{A}}\mathcal{P}^\star(a)u_k(a)\right| \leq \varepsilon\cdot\max_{a\in\mathcal{A}}|u_k(a)| \tag{54}$$

$$\implies \left|\frac{1}{T}\sum_{t=1}^{T} u_k(a^t) - \mathbb{E}_{\mathcal{P}^\star}\Big[u_k(\mathbf{a})\Big]\right| \leq \varepsilon\cdot\max_{a\in\mathcal{A}}|u_k(a)|. \tag{55}$$

Inequality (52) comes from the definition 9 of typical sequences.
Inequalities (53) come from the homogeneity property.
Equation (54) comes from the triangle inequality.
Equation (55) is a reformulation of (54) and allows us to obtain the following equations :

$$\left|\gamma_k^T(\sigma^\star,\tau^\star)-\mathbb{E}_{\mathcal{P}^\star}\left[u_k(\mathbf{a})\right]\right| \tag{56}$$

$$=\left|\sum_{a^T\in\mathcal{A}^T}\mathcal{P}_{\sigma^\star,\tau^\star}(a^T)\cdot\frac{1}{T}\sum_{t=1}^{T}u_k(a^t)-\mathbb{E}_{\mathcal{P}^\star}\left[u_k(\mathbf{a})\right]\right| \tag{57}$$

$$=\left|\sum_{a^T\in A_\varepsilon^{\star T}(\mathcal{P}^\star)}\mathcal{P}_{\sigma^\star,\tau^\star}(a^T)\cdot\frac{1}{T}\sum_{t=1}^{T}u_k(a^t)+\sum_{a^T\notin A_\varepsilon^{\star T}(\mathcal{P}^\star)}\mathcal{P}_{\sigma^\star,\tau^\star}(a^T)\cdot\frac{1}{T}\sum_{t=1}^{T}u_k(a^t)-\mathbb{E}_{\mathcal{P}^\star}\left[u_k(\mathbf{a})\right]\right| \tag{58}$$

$$\leq\sum_{a^T\in A_\varepsilon^{\star T}(\mathcal{P}^\star)}\mathcal{P}_{\sigma^\star,\tau^\star}(a^T)\cdot\left|\frac{1}{T}\sum_{t=1}^{T}u_k(a^t)-\mathbb{E}_{\mathcal{P}^\star}\left[u_k(\mathbf{a})\right]\right|+\sum_{a^T\notin A_\varepsilon^{\star T}(\mathcal{P}^\star)}\mathcal{P}_{\sigma^\star,\tau^\star}(a^T)\cdot\frac{1}{T}\sum_{t=1}^{T}|u_k(a^t)| \tag{59}$$

$$\leq\sum_{a^T\in A_\varepsilon^{\star T}(\mathcal{P}^\star)}\mathcal{P}_{\sigma^\star,\tau^\star}(a^T)\cdot\varepsilon\cdot\max_{a\in\mathcal{A}}|u_k(a)|+\sum_{a^T\notin A_\varepsilon^{\star T}(\mathcal{P}^\star)}\mathcal{P}_{\sigma^\star,\tau^\star}(a^T)\cdot\varepsilon\cdot\max_{a\in\mathcal{A}}|u_k(a)| \tag{60}$$

$$\leq\varepsilon\cdot\max_{a\in\mathcal{A}}|u_k(a)|+\mathcal{P}\left(\mathbf{a}^T\notin A_\varepsilon^{\star T}\right)\cdot\varepsilon\cdot\max_{a\in\mathcal{A}}|u_k(a)| \tag{61}$$

$$\leq\max_{a\in\mathcal{A}}|u_k(a)|\cdot\left(\varepsilon+\mathcal{P}(\mathbf{a}^T\notin A_\varepsilon^{\star T}|\mathbf{E}=0)\cdot\mathcal{P}(\mathbf{E}=0)+\mathcal{P}(\mathbf{a}^T\notin A_\varepsilon^{\star T}|\mathbf{E}=1)\cdot\mathcal{P}(\mathbf{E}=1)\right) \tag{62}$$

$$\leq\max_{a\in\mathcal{A}}|u_k(a)|\cdot\left(\varepsilon+\mathcal{P}(\mathbf{a}^T\notin A_\varepsilon^{\star T}|\mathbf{E}=0)+\mathcal{P}(\mathbf{E}=1)\right) \tag{63}$$

$$\leq\max_{a\in\mathcal{A}}|u_k(a)|\cdot\left(2\varepsilon+\mathcal{P}(\mathbf{E}=1)\right) \tag{64}$$

$$\leq\max_{a\in\mathcal{A}}|u_k(a)|\cdot\left(2\varepsilon+2\varepsilon\cdot B\cdot K^2\right) \tag{65}$$

$$\leq 4\varepsilon\cdot\max_{a\in\mathcal{A}}|u_k(a)|\cdot B\cdot K^2. \tag{66}$$

Equalities (57) and (58) come from the definition of the expected T-stages utility, see (12).
Inequality (59) comes from the triangle inequality.
Inequality (60) comes from (55).
Inequalities (61), (62) and (63) are reformulation of (60).
Inequality (64) comes from (50) because by assumption $T\geq n\geq n_1$.
Inequality (65) comes from Lemma 1 because by assumption $n\geq n_1$.
Inequality (66) is a reformulation with $B\cdot K^2\geq 1$ which concludes the proof of Lemma 2.

As a conclusion, for all $\varepsilon>0$, there exists $n_1\in\mathbb{N}$ such that for all $n\geq n_1$, the condition (71) is satisfied. □

B.3 Condition (ii) of definition 6

In order to prove that the strategies $(\sigma^\star,\tau^\star)\in\Sigma\times\mathcal{T}$ support a uniform equilibrium, we suppose that player $k\in\mathcal{K}$ implement a deviating strategy $\tau_k'\neq\tau_k^\star$ and we prove that the deviation gain is less than $\epsilon>0$.

B.3.1 First case: non-typical deviations

Let $b-1 \in \mathcal{B}$ the first action block over which the sequence of actions $a_k(b-1) \notin A_\varepsilon^{\star n}(\mathcal{P}_k^\star)$ of player $k \in \mathcal{K}$ is not typical. Denote $t_1(b)$ and $t_n(b)$ the indexes of the first and the last stage of block $b \in \mathcal{B}$.

Evaluate, for player $k \in \mathcal{K}$, the utilities associated with the strategies $\tau_k^\star$ and τ_k'.

$$\gamma_k^T(\sigma^\star, \tau_k^\star, \tau_{-k}^\star) = \mathbb{E}_{\sigma^\star, \tau_k^\star, \tau_{-k}^\star}\left[\frac{1}{T}\sum_{t=1}^{T} u_k(a^t)\right], \tag{67}$$

$$\gamma_k^T(\sigma^\star, \tau_k', \tau_{-k}^\star) = \frac{1}{T}\mathbb{E}_{\sigma^\star, \tau_k', \tau_{-k}^\star}\left[\sum_{t=1}^{T} u_k(a^t)\right] \tag{68}$$

$$= \frac{1}{T}\mathbb{E}_{\sigma^\star, \tau_k', \tau_{-k}^\star}\left[\sum_{t=1}^{t_n(b-2)} u_k(a^t) + \sum_{t=t_1(b-1)}^{t_n(b)} u_k(a^t) + \sum_{t=t_1(b+1)}^{T} u_k(a^t)\right]. \tag{69}$$

Approximation of the utility associated with the strategies $(\sigma^\star, \tau^\star) \in \Sigma \times \mathcal{T}$ between blocks $b+1 \in \mathcal{B}$ and $B \in \mathcal{B}$.

Lemma 3 *Suppose that the encoder $\mathcal{C}$ and the players $\mathcal{K}$ follow the strategies $(\sigma^\star, \tau^\star) \in \Sigma \times \mathcal{T}$. Then for all $\epsilon > 0$ and for all number of blocks $B \in \mathbb{N}$, there exists a block length n_1 such that for all $n \geq n_1$, the following inequality is satisfied for all $1 \leq b \leq B-1$:*

$$\mathbb{E}_{\sigma^\star, \tau_k^\star, \tau_{-k}^\star}\left[\sum_{t=t_1(b+1)}^{T} u_k(a^t)\right] \geq n(B-b)\cdot(U_k^\star - \frac{\epsilon}{2}). \tag{70}$$

Proof. [Lemma 3] Let us fix the parameter $\epsilon > 0$ and suppose that the encoder $\mathcal{C}$ and the players $\mathcal{K}$ follows the strategies $(\sigma^\star, \tau^\star) \in \Sigma \times \mathcal{T}$. This proof is built on Lemma 2 that prove for all $\varepsilon > 0$, there exists a block length $n_1 \in \mathbb{N}$, such that for all $n \geq n_1$, the expected utility satisfies the following equation :

$$\left|\gamma_k^T(\sigma^\star, \tau^\star) - \mathbb{E}_{\mathcal{P}^\star}\left[u_k(\mathbf{a}_k, \mathbf{a}_{-k})\right]\right| \leq 4\varepsilon \cdot \max_{a\in\mathcal{A}}|u_k(a)| \cdot B \cdot K^2, \qquad \forall k \in \mathcal{K}. \tag{71}$$

For a fixed number of blocks $B \in \mathbb{N}$, we choose $\varepsilon > 0$ such that $\epsilon \geq 4\varepsilon \cdot \max_{a\in\mathcal{A}}|u_k(a)| \cdot B \cdot K^2)$. Using the same reasoning as in Lemma 2, we prove that for all $\varepsilon > 0$, there exists a block length $n_1 \in \mathbb{N}$, such that for all $n \geq n_1$, the expected utility satisfies the following equation for all $k \in \mathcal{K}$:

$$\left|\mathbb{E}_{\sigma^\star, \tau_k^\star, \tau_{-k}^\star}\left[\frac{1}{n(B-b)}\sum_{t=t_1(b+1)}^{T} u_k(a^t)\right] - \mathbb{E}_{\mathcal{P}^\star}\left[u_k(\mathbf{a}_k, \mathbf{a}_{-k})\right]\right| \leq 4\varepsilon \cdot \max_{a\in\mathcal{A}}|u_k(a)| \cdot B \cdot K^2, \tag{72}$$

$$\implies \left|\mathbb{E}_{\sigma^\star, \tau_k^\star, \tau_{-k}^\star}\left[\sum_{t=t_1(b+1)}^{T} u_k(a^t)\right] - n(B-b)\cdot U_k^\star\right| \leq n(B-b)\cdot 4\varepsilon \cdot \max_{a\in\mathcal{A}}|u_k(a)| \cdot B \cdot K^2, \tag{73}$$

$$\implies \mathbb{E}_{\sigma^\star, \tau_k^\star, \tau_{-k}^\star}\left[\sum_{t=t_1(b+1)}^{T} u_k(a^t)\right] \geq n(B-b)\cdot(U_k^\star - 4\varepsilon \cdot \max_{a\in\mathcal{A}}|u_k(a)| \cdot B \cdot K^2). \tag{74}$$

For a fixed block number $B \in \mathbb{N}$, we choose the parameter $\frac{\epsilon}{2} \geq 4\varepsilon \cdot \max_{a\in\mathcal{A}}|u_k(a)| \cdot B \cdot K^2$ and the block length $n_1 \in \mathbb{N}$ that satisfies (42) and (43) of Lemma 1. We obtain the inequality (70) of Lemma 3. □

Upper bound on the deviation gain obtained by player $k \in \mathcal{K}$ using the deviation strategy $\tau_k' \in \mathcal{T}_k$, over blocks $b-1 \in \mathcal{B}$ and $b \in \mathcal{B}$.

Lemma 4 *For every block $b \in \mathcal{B}$, the following inequality is satisfied:*

$$\mathbb{E}_{\sigma^\star,\tau'_k,\tau^\star_{-k}}\left[\sum_{t=t_1(b-1)}^{t_n(b)} u_k(a^t)\right] \leq 2n \cdot \max_{a\in\mathcal{A}} |u_k(a)|, \qquad \forall k \in \mathcal{K}. \tag{75}$$

Proof. [Lemma 4] The proof is direct.

$$\mathbb{E}_{\sigma^\star,\tau'_k,\tau^\star_{-k}}\left[\sum_{t=t_1(b-1)}^{t_n(b)} u_k(a^t)\right] \leq \mathbb{E}_{\sigma^\star,\tau'_k,\tau^\star_{-k}}\left[(t_n(b)-t_1(b-1)+1)\cdot \max_{a\in\mathcal{A}} |u_k(a)|\right] \tag{76}$$

$$\leq 2n \cdot \max_{a\in\mathcal{A}} |u_k(a)|, \qquad \forall k \in \mathcal{K}. \tag{77}$$

The deviation utility satisfies (75). □

Upper bound on the utility of player $k \in \mathcal{K}$ during the punishment phase induced by the prescribed strategies $(\sigma^\star, \tau^\star) \in \Sigma \times \mathcal{T}$.

Lemma 5 *Suppose that $b-1 \in \mathcal{B}$ is the first block on which the sequence of actions $a_k(b-1) \notin A_\varepsilon^{\star n}(\mathcal{P}_k^\star)$ of player $k \in \mathcal{K}$ is not typical. Suppose the block length satisfies $n \geq n_1$, defined by (42) and (43). The prescribed strategies $(\sigma^\star, \tau^\star) \in \Sigma \times \mathcal{T}$, induce the following equation:*

$$\mathbb{E}_{\sigma^\star,\tau'_k,\tau^\star_{-k}}\left[\sum_{t=t_1(b+1)}^{T} u_k(a^t)\right] \leq \varepsilon \cdot \max_{a\in\mathcal{A}} |u_k(a)| + n(B-b)\cdot v_k. \tag{78}$$

Lemma 5 is a consequence of the coding result stated by Theorem 1.
Proof. We suppose that the actions $a_k(b-1) \notin A_\varepsilon^{\star n}(\mathcal{P}_k^\star)$ of player $k \in \mathcal{K}$ over block $b-1 \in \mathcal{B}$ are not typical and the length of the block satisfies $n \geq n_1$, defined by (42) and (43). Define the error event $\mathbf{E}_d$ related with the statistical test $\mathbf{E}_k^i(b+1)$ defined by (34). $\mathbf{E}_d = 0$ means that the statistical test of all the players $j \neq k \in \mathcal{K}$ reveals that player $k \in \mathcal{K}$ deviates during block $b-1 \in \mathcal{B}$.

$$\mathbf{E}_d = \begin{cases} 0 \text{ if } & \forall j \neq k, \quad \mathbf{E}_j^k(b+1) = 1, \\ 1 \text{ otherwise.} \end{cases} \tag{79}$$

The coding result stated by Theorem 1 allow us to bound the probability of event $\mathbf{E}_d = 1$ knowing that $a_k(b-1) \notin A_\varepsilon^{\star n}(\mathcal{P}_k^\star)$. The following inequalities are valid for any deviation strategy $\tau'_k \in \mathcal{T}_k$ of player $k \in \mathcal{K}$.

$$\mathcal{P}_{\sigma^\star,\tau'_k,\tau^\star_{-k}}\left(\mathbf{E}_d = 1 \middle| a_k(b-1) \notin A_\varepsilon^{\star n}(\mathcal{P}_k^\star)\right) \tag{80}$$

$$= \mathcal{P}_{\sigma^\star,\tau'_k,\tau^\star_{-k}}\left(\exists j \neq k \in \mathcal{K}, \quad \mathbf{E}_j^k(b+1) = 0 \middle| a_k(b-1) \notin A_\varepsilon^{\star n}(\mathcal{P}_k^\star)\right) \tag{81}$$

$$\leq \mathcal{P}_{\sigma^\star,\tau'_k,\tau^\star_{-k}}\left(\cup_{j\neq k}\left\{\mathbf{E}_j^k(b+1) = 0\right\} \middle| a_k(b-1) \notin A_\varepsilon^{\star n}(\mathcal{P}_k^\star)\right) \tag{82}$$

$$\leq \mathcal{P}_{\sigma^\star,\tau'_k,\tau^\star_{-k}}\left(\cup_{j\neq k}\left\{\hat{\mathbf{a}}_k^j(b-1) \in A_\varepsilon^{\star n}(\mathcal{P}_k^\star)\right\} \middle| a_k(b-1) \notin A_\varepsilon^{\star n}(\mathcal{P}_k^\star)\right) \tag{83}$$

$$\leq \mathcal{P}_e(\lambda) \tag{84}$$

$$\leq \sum_{k\in\mathcal{K}} \max_{i\in\mathcal{K}} \max_{v_i \in \Delta(\mathcal{A}_i^\infty)} \mathcal{P}(\mathbf{a}^n \neq \hat{\mathbf{a}}^n(k) | v_i) \tag{85}$$

$$\leq \varepsilon, \qquad \forall \tau'_k \in \mathcal{T}_k. \tag{86}$$

Inequalities 81 and 82 come from the definition $\mathbf{E}_d$ and Boole's inequality.
Inequality 83 comes from the definition of the statistical test 34 presented section B.1.4.
Inequality 84 come from the strategy of the encoder $\sigma^\star \in \Sigma$ and the decoding scheme of

player $j \in \mathcal{K}$, described in sections B.1.2 and B.1.3, based on the coding scheme $\lambda \in \Lambda(n)$ with $n \geq n_1$ during block $b-1 \in \mathcal{B}$.
Inequality 85 comes from the definition of the error probability of the code $\lambda \in \Lambda(n)$.
Inequality 86 comes from the coding result given by the Theorem 1 for an arbitrarily varying information source (AVS). When a player deviates, the actions are generated with an incertain probability distribution satisfying the hypothesis (3) and (4) of the definition 1. We suppose the length of a block satisfies $n \geq n_1$ defined by (42) and (43). Therefore, from the Theorem 1, there exists a code $\lambda \in \Lambda(n)$ for which the error probability $\mathcal{P}_e(\lambda)$ is bounded by $\varepsilon > 0$, for every unilateral deviation $\tau'_k \in \mathcal{T}_k$ of player $k \in \mathcal{K}$. This inequality allow us to obtain an upper bound on the utility of player $k \in \mathcal{K}$ during the punishment phase stated by the Lemma 5.

$$\mathbb{E}_{\sigma^\star,\tau'_k,\tau^\star_{-k}}\left[\sum_{t=t_1(b+1)}^{T} u_k(a^t)\right] \tag{87}$$

$$= \sum_{a^T_{t_1(b+1)} \in \mathcal{A}^{t_n(b)}} \cdot \mathcal{P}_{\sigma^\star,\tau'_k,\tau^\star_{-k}}\left(a^T_{t_1(b+1)}, \mathbf{E}_d = 1 \middle| a_k(b-1) \notin A^{\star n}_\varepsilon(\mathcal{P}^\star_k)\right) \cdot \left[\sum_{t=t_1(b+1)}^{T} u_k(a^t)\right]$$

$$+ \sum_{a^T_{t_1(b+1)} \in \mathcal{A}^{t_n(b)}} \cdot \mathcal{P}_{\sigma^\star,\tau'_k,\tau^\star_{-k}}\left(a^T_{t_1(b+1)}, \mathbf{E}_d = 0 \middle| a_k(b-1) \notin A^{\star n}_\varepsilon(\mathcal{P}^\star_k)\right) \cdot \left[\sum_{t=t_1(b+1)}^{T} u_k(a^t)\right] \tag{88}$$

$$\leq \mathcal{P}_{\sigma^\star,\tau'_k,\tau^\star_{-k}}\left(\mathbf{E}_d = 1 \middle| a_k(b-1) \notin A^{\star n}_\varepsilon(\mathcal{P}^\star_k)\right) \cdot \max_{a \in \mathcal{A}} |u_k(a)| + \left[n(B-b) \cdot \upsilon_k\right] \tag{89}$$

$$\leq \varepsilon \cdot \max_{a \in \mathcal{A}} |u_k(a)| + \left[n(B-b) \cdot \upsilon_k\right]. \tag{90}$$

Inequality 88 comes from the definition of the expectation 87 knowing that the sequence of actions $a_k(b-1) \notin A^{\star n}_\varepsilon(\mathcal{P}^\star_k)$ is not typical over the block $b-1 \in \mathcal{B}$.
Inequality 89 comes from the punishment plan stated by the strategy $\tau^\star_{-k}$ when all the players $j \neq k$ detect the deviation of player $k \in \mathcal{K}$ ($\mathbf{E}_d = 1$). For each stage $t_1(b+1) \leq t \leq T$, the utility of player $k \in \mathcal{K}$ is less than the min-max level $u_k(a^t) \leq \upsilon_k$.
Inequality 90 comes from (83). □

Equilibrium condition. Hypothesis (29) over the number of blocks $B \in \mathbb{R}$ and the results of Lemma 3, 4 and 5 allow us to obtain the following inequalities:

$$B \geq \frac{8 \cdot \max_{a\in\mathcal{A}} |u_k(a)|}{\epsilon} \tag{91}$$

$$\implies -\tfrac{B\epsilon}{2} + B\epsilon \geq 4 \cdot \max_{a\in\mathcal{A}} |u_k(a)| \tag{92}$$

$$\implies (B-b)\cdot U_k^\star - \tfrac{B\epsilon}{2} + B\epsilon \geq 4 \cdot \max_{a\in\mathcal{A}} |u_k(a)| + (B-b)\cdot \upsilon_k \tag{93}$$

$$\implies (B-b+2)\cdot U_k^\star - \tfrac{(B-b)\epsilon}{2} + B\epsilon \geq 2 \cdot \max_{a\in\mathcal{A}} |u_k(a)| + \epsilon + \varepsilon \cdot \max_{a\in\mathcal{A}} |u_k(a)| + (B-b)\cdot \upsilon_k \tag{94}$$

$$\implies n(B-b+2)\cdot U_k^\star - n\tfrac{(B-b+2)\epsilon}{2} + nB\epsilon \geq 2n \cdot \max_{a\in\mathcal{A}} |u_k(a)| + n\varepsilon \cdot \max_{a\in\mathcal{A}} |u_k(a)| + n(B-b)\cdot \upsilon_k \tag{95}$$

$$\implies \mathbb{E}_{\sigma^\star,\tau_k^\star,\tau_{-k}^\star}\Big[\textstyle\sum_{t=t_1(b-1)}^{T} u_k(a^t)\Big] + nB\epsilon \geq 2n \cdot \max_{a\in\mathcal{A}} |u_k(a)| + n\varepsilon \cdot \max_{a\in\mathcal{A}} |u_k(a)| + n(B-b)\cdot \upsilon_k \tag{96}$$

$$\implies \mathbb{E}_{\sigma^\star,\tau_k^\star,\tau_{-k}^\star}\Big[\textstyle\sum_{t=t_1(b-1)}^{T} u_k(a^t)\Big] + nB\epsilon \geq \mathbb{E}_{\sigma^\star,\tau_k',\tau_{-k}^\star}\Big[\sum_{t=t_1(b-1)}^{t_n(b)} u_k(a^t)\Big] + n\varepsilon \cdot \max_{a\in\mathcal{A}} |u_k(a)| + n(B-b)\cdot \upsilon_k \tag{97}$$

$$\implies \mathbb{E}_{\sigma^\star,\tau_k^\star,\tau_{-k}^\star}\Big[\textstyle\sum_{t=t_1(b-1)}^{T} u_k(a^t)\Big] + nB\epsilon \geq \mathbb{E}_{\sigma^\star,\tau_k',\tau_{-k}^\star}\Big[\sum_{t=t_1(b-1)}^{t_n(b)} u_k(a^t)\Big] + \mathbb{E}_{\sigma^\star,\tau_k',\tau_{-k}^\star}\Big[\sum_{t=t_1(b+1)}^{T} u_k(a^t)\Big] \tag{98}$$

$$\implies \mathbb{E}_{\sigma^\star,\tau_k^\star,\tau_{-k}^\star}\Big[\textstyle\sum_{t=1}^{T} u_k(a^t)\Big] + T\cdot\varepsilon \geq \mathbb{E}_{\sigma^\star,\tau_k',\tau_{-k}^\star}\Big[\sum_{t=1}^{t_n(b-2)} u_k(a^t) + \sum_{t=t_1(b-1)}^{t_n(b)} u_k(a^t) + \sum_{t=t_1(b+1)}^{T} u_k(a^t)\Big] \tag{99}$$

$$\implies \gamma_k^T(\sigma^\star,\tau_k^\star,\tau_{-k}^\star) + \varepsilon \geq \gamma_k^T(\sigma^\star,\tau_k',\tau_{-k}^\star). \tag{100}$$

Inequality 91 comes from the hypothesis (29) over the number of blocks $B \in \mathbb{R}$.
Inequality 92 comes from the reformulation of inequality (91).
Inequality 93 comes from the hypothesis of individual rationality $U_k^\star \geq \upsilon_k$ stated by the definition 7.
Inequalities 94 and 95 come from the reformulation of inequality (93) with $\epsilon \leq \max_{a\in\mathcal{A}} |u_k(a)|$ and $\varepsilon \leq 1$.
Inequality 96 comes from Lemma 3 which provides an approximation of the utility associated with the strategies $(\sigma^\star,\tau^\star) \in \Sigma\times\mathcal{T}$.
Inequality 97 comes from the Lemma 4 which provides an upper bound on the deviation gain obtained by player $k\in\mathcal{K}$ while playing the strategy $\tau_k' \in \mathcal{T}_k$.
Inequality 98 comes from Lemma 5 which is a consequence of the coding result stated by Theorem 1. This result provides an upper bound over the utility of player $k\in\mathcal{K}$ during the punishment phase.
Inequality 99 comes from the fact that $b-1 \in \mathcal{B}$ is the first block on which the action sequence $a_k(b-1) \notin A_\varepsilon^{\star n}(\mathcal{P}_k^\star)$ of player $k\in\mathcal{K}$ is not typical.
Inequality 100 comes from the definition of the utilities of the T-stages repeated game stated equation (12).
We prove the strategies $(\sigma^\star,\tau^\star) \in \Sigma\times\mathcal{T}$ satisfy the equilibrium condition stated by point (ii) of the definition 6.

B.3.2 Second case: typical deviations

Let us fix $\epsilon > 0$ and suppose player $k\in\mathcal{K}$ uses a deviating strategy $\tau_k' \in \mathcal{T}_k$ such that the action sequence of player $k\in\mathcal{K}$ over each block $b\in\mathcal{B}$ belong to the set of typical sequences $a_k(b) \in A_\varepsilon^{\star n}(\mathcal{P}_k^\star)$. From the proof of Lemma 2, for all $\varepsilon > 0$, there exists $n_1 \in \mathbb{N}$ such that for all $n \geq n_1$ we have the following implication (102). Taking $\varepsilon \cdot \max_{a\in\mathcal{A}} |u_k(a)| \leq \epsilon$, we

obtain the implication (103).

$$a^T \in A_\varepsilon^{\star T}(\mathcal{P}^\star) \tag{101}$$

$$\implies \left| \frac{1}{T} \sum_{t=1}^{T} u_k(a^t) - \mathbb{E}_{\mathcal{P}^\star}\left[u_k(\mathbf{a}) \right] \right| \leq \varepsilon \cdot \max_{a \in \mathcal{A}} |u_k(a)|. \tag{102}$$

$$\implies \gamma_k^T(\sigma^\star, \tau_k', \tau_{-k}^\star) \leq \gamma_k^T(\sigma^\star, \tau_k^\star, \tau_{-k}^\star) + \epsilon. \tag{103}$$

The utility provided by the strategies $(\sigma^\star, \tau^\star) \in \Sigma \times \mathcal{T}$ satisfies the equilibrium condition stated by point (ii) in definition 6.

B.4 Conclusion

We showed that by setting the parameter $\epsilon > 0$, we obtain a condition on the number of blocks $B \in \mathbb{N}$ given by (91), then a condition over the coding parameter $\varepsilon > 0$ given by (71) and then a condition over the block length $n \geq n_1$ given by (42) and (43).
For all $U \in u(\mathcal{R}) \cap IR$, these parameters allow us to construct a pair of strategies $(\sigma^\star, \tau^\star) \in \Sigma \times \mathcal{T}$ over $T = n \cdot B$ stages that satisfies the conditions (27) and (28). By repeating these strategies cyclically, we show that any vector of utility $U \in \operatorname{conv} u(\mathcal{R}) \cap IR$ satisfy both conditions (i) and (ii) (i.e. definition 6) of the uniform equilibrium. The utility $U \in \operatorname{conv} u(\mathcal{R}) \cap IR$ is a uniform equilibrium utility for the infinite repeated game Γ^∞.

C Review of typical sequences

The achievability part of the coding theorems are based on the properties of the typical sequences. This section provides some recall on this notions that can also be found in [Cover and Thomas(2006)] and [Csiszár and Körner(1981)].

Definition 9 (Typical sequences [Csiszár and Körner(1981)]) Let $\mathcal{Q} \in \Delta(\mathcal{X} \times \mathcal{Y})$ a probability distribution over $\mathcal{X} \times \mathcal{Y}$. The typical sequences and the conditional typical sequences are defined as follows:

$$A_\varepsilon^{n\star}(\mathcal{X}) = \left\{ x^n \in \mathcal{X}^n;\ \sum_{x \in \mathcal{X}} \left| \frac{N(x|x^n)}{n} - \mathcal{Q}(x) \right| \leq \varepsilon,\ \forall x \in \mathcal{X},\ \mathcal{Q}(x) = 0 \implies N(x|x^n) = 0 \right\}.$$

$$A_\varepsilon^{n\star}(\mathcal{X}|y^n) = \Bigg\{ x^n \in \mathcal{X}^n;\ \sum_{\substack{x \in \mathcal{X}, \\ y \in \mathcal{Y}}} \left| \frac{N(x,y|x^n,y^n)}{n} - \mathcal{Q}(x,y) \right| \leq \varepsilon,$$

$$\forall (x,y) \in \mathcal{X} \times \mathcal{Y},\ \mathcal{Q}(x,y) = 0 \implies N(x,y|x^n,y^n) = 0 \Bigg\}. \tag{104}$$

Lemma 6 (Properties of the typical sequences [Csiszár and Körner(1981)]) *Let $\mathcal{Q} \in \Delta(\mathcal{X} \times \mathcal{Y})$ a probability distribution, $\mathcal{Q}^{\otimes n}$ a n-product of the probability distribution and $y^n \in A_\varepsilon^{n\star}(\mathcal{Y})$ a typical sequence. For all $\varepsilon > 0$, there exists $n \in \mathbb{N}$ such that:*

$$1 = \mathcal{Q}^{\otimes n}\left(\boldsymbol{x}^n \in A_\varepsilon^{n\star}(\mathcal{X}) \right), \tag{105}$$

$$1 = \mathcal{Q}^{\otimes n}\left(\boldsymbol{x}^n \in A_\varepsilon^{n\star}(\mathcal{X}|y^n) \middle| y^n \right), \tag{106}$$

$$2^{n(H(\boldsymbol{x})-c\varepsilon)} \leq |A_\varepsilon^{n\star}(\mathcal{X})| \leq 2^{n(H(\boldsymbol{x})+c\varepsilon)}, \tag{107}$$

$$2^{n(H(\boldsymbol{x}|\boldsymbol{y})-c\varepsilon)} \leq |A_\varepsilon^{n\star}(\mathcal{X}|y^n)| \leq 2^{n(H(\boldsymbol{x}|\boldsymbol{y})+c\varepsilon)}. \tag{108}$$

where $c = \log\left(\max_{x \in \mathcal{X}} \frac{1}{\mathcal{Q}(x)} \right)$ is constant.

This result states that an i.i.d. sequence of symbols is almost surely typical when n goes to $+\infty$ and it provides an upper and a lower bound on the size of the sets of typical sequences.

Lemma 7 (Packing Lemma [El Gamal and van der Meulen(1981)]) *Let $\mathcal{Q} \in \Delta(\mathcal{U} \times \mathcal{V})$ a correlated probability distribution, $\mathcal{Q}_U$ (resp. $\mathcal{Q}_V$) the marginal induced by $\mathcal{Q}$ over $\mathcal{U}$ (resp. $\mathcal{V}$), $\mathcal{Q}_U^{\otimes n}$ and $\mathcal{Q}_V^{\otimes n}$ the n-product of marginal probability. Let R_I and R_J real numbers,*

- $(u_i^n)_{i\in\{1,\dots,2^{nR_I}\}} \in \mathcal{U}^n$ *a family of sequences drawn with* $\mathcal{Q}_U^{\otimes n}$,
- $(v_j^n)_{j\in\{1,\dots,2^{nR_J}\}} \in \mathcal{V}^n$ *a family of sequences drawn with* $\mathcal{Q}_V^{\otimes n}$,

If the condition (109) is satisfied,

$$R_I + R_J < I_{\mathcal{Q}}(\boldsymbol{u};\boldsymbol{v}), \tag{109}$$

then for all $\varepsilon > 0$, there exists a $\bar{n} \geq 0$ such that for all $n \geq \bar{n}$,

$$\mathcal{P}\left(\cup_{\substack{i\in\mathcal{I},\\ j\in\mathcal{J}}} \left\{ (u_i^n, v_j^n) \in A_\varepsilon^{\star n}(\mathcal{U}\times\mathcal{V}) \right\} \right) \leq \varepsilon. \tag{110}$$

Where $I_{\mathcal{Q}}(\boldsymbol{u};\boldsymbol{v})$ denote the mutual information [Cover and Thomas(2006)] with respect to the probability distribution $\mathcal{Q}$.

References

Ahlswede(1979). Ahlswede R (1979) Coloring hypergraphs: A new approach to multi-user source coding, part 1. Journal of combinatorics, information and system sciences 4(1):76–115

Ahlswede(1980). Ahlswede R (1980) Coloring hypergraphs: A new approach to multi-user source coding, part 2. Journal of combinatorics, information and system sciences 5(3):220–268

Aumann(1981a). Aumann RJ (1981a) Survey of repeated game pp 11–42

Aumann(1981b). Aumann RJ (1981b) Survey of repeated game. In: Essays in Game Theory and Mathematical Economics in Honor of Oskar Morgenstern, Ed. Wissenschaftsverlag, Bibliographisches Institut, Mannheim, Wien, Zurich, pp 11–42

Bavly and Neyman(2003). Bavly G, Neyman A (2003) Online concealed correlation by boundedly rational players. Discussion paper series, The Center for the Study of Rationality, Hebrew University, Jerusalem

Bondy and Murty(1976). Bondy J, Murty U (1976) Graph theory with applications. Elsevier Science Publishing Co.

Cover and Thomas(2006). Cover TM, Thomas JA (2006) Elements of information theory. 2nd. Ed., Wiley-Interscience, New York

Csiszár and Körner(1981). Csiszár I, Körner J (1981) Information Theory: Coding Theorems for Discrete Memoryless Systems

DaSilva et al(2011)DaSilva, Taffin, Lasaulce, and Buljore. DaSilva MJ, Taffin A, Lasaulce S, Buljore S (2011) Closed loop transmit diversity enhancements for UMTS narrowband and wideband TD-CDMA. In: IEEE Proc. of the 53rd Vehicular Technology Conference, vol 3, pp 1963–1967

El Gamal and van der Meulen(1981). El Gamal AA, van der Meulen E (1981) A proof of Marton's coding theorem for the discrete memoryless broadcast channel. IEEE Transactions on Information Theory 27(1):120–122

Fudenberg et al(1994)Fudenberg, Levine, and Maskin. Fudenberg D, Levine D, Maskin E (1994) The folk theorem with imperfect public information. Econometrica 62(5):997–1039

Gallager(1976). Gallager RG (1976) Source coding with side information and universal coding. In: IEEE International Symposium on Information Theory, Renneby, Sweden

Gossner and Tomala(2007). Gossner O, Tomala T (2007) Secret correlation in repeated games with imperfect monitoring. Mathematics of Operation Research 32(2):413–424

Gossner et al(2006)Gossner, Hernandez, and Neyman. Gossner O, Hernandez P, Neyman A (2006) Optimal use of communication resources. Econometrica 74(6):1603–1636

Han(2003). Han TS (2003) Information-spectrum methods in information theory. Springer

Hörner and Olszewski(2006). Hörner J, Olszewski W (2006) The folk theorem for games with private almost-perfect monitoring. Econometrica 74(6):1499–1544

I. F. Akyildiz and Cayirci(2002). I F Akyildiz YS W Su, Cayirci E (2002) Wireless sensor networks : A survey. Computer Networks 38(4):393Ű–422

Kowalewski(2000). Kowalewski F (2000) Joint predistortion and transmit diversity. In: IEEE Proc. of the Global Telecommunications Conference, vol 1, pp 245–249

Lehrer(1988). Lehrer E (1988) Repeated games with stationary bounded recall strategies. Journal of Economic Theory 46:130–144

Lehrer(1991). Lehrer E (1991) Internal correlation in repeated games. International Journal of Game Theory 19(4):431–456

LeTreust and Lasaulce(2010). LeTreust M, Lasaulce S (2010) A repeated game formulation of energy-efficient decentralized power control. IEEE Trans on Wireless Commun 9(9):2860–2869

LeTreust and Lasaulce(2011a). LeTreust M, Lasaulce S (2011a) The price of re-establishing almost perfect monitoring in games with arbitrary monitoring structures. ACM Proc of the 4th International Workshop on Game Theory in Communication Networks (GAMECOMM11), Cachan (Paris), France

LeTreust and Lasaulce(2011b). LeTreust M, Lasaulce S (2011b) Resilient source coding. IEEE Proc of the international conference on network games, control and optimization (NETGCOOP11), Paris, France

Peretz(2011). Peretz R (2011) Correlation through bounded recall strategies. Discussion paper series, The Center for the Study of Rationality, Hebrew University, Jerusalem

Peyton(2004). Peyton Y (2004) Strategic learning and its limits

Renault and Tomala(2011). Renault J, Tomala T (2011) General properties of long-run supergames. Dynamic Games and Applications 1(2):319–350

Shannon(1948). Shannon CE (1948) A mathematical theory of communication. Bell System Technical Journal 27:379–423

Shannon(1956). Shannon CE (1956) The zero error capacity of a noisy channel. IRE Trans Inf Theory 2:8–19

Slepian and Wolf(1973). Slepian D, Wolf JK (1973) Noiseless coding of correlated information sources. IEEE Trans on Information Theory 19:471–480

Sorin(1992). Sorin S (1992) Repeated Games with Complete Information, in Handbook of Game Theory with Economic Applications, vol 1. Elsevier Science Publishers

Witsenhausen(1976). Witsenhausen H (1976) The zero-error side information problem and chromatic numbers (corresp.). Information Theory, IEEE Transactions on 22(5):592 – 593