HAL
open science

Nash and publicly correlated equilibrium payoffs in non zero sum differential games using mixed strategies

Anne Souquière

To cite this version:

Anne Souquière. Nash and publicly correlated equilibrium payoffs in non zero sum differential games using mixed strategies. 2009. hal-00418629

HAL Id: hal-00418629

https://hal.science/hal-00418629v1

Preprint submitted on 21 Sep 2009

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.

Nash and publicly correlated equilibrium payoffs in non zero sum differential games using mixed strategies

A. SOUQUIERE

1-Institut TELECOM ; TELECOM Bretagne ; UMR CNRS 3192 Lab-STICC,
Technopôle Brest Iroise, CS 83818, 29238 Brest Cedex 3
2- Université de Bretagne Occidentale, Laboratoire de Mathématiques, UMR 6205,
6 Avenue Victor Le Gorgeu, CS 93837, 29238 Brest Cedex 3
3- Université Européenne de Bretagne, France
Anne.Souquiere@telecom-bretagne.eu

September 21st, 2009

Abstract

We consider non zero sum two players differential games. We study Nash equilibrium payoffs and publicly correlated equilibrium payoffs. If players use deterministic strategies, it has been proved that the Nash equilibrium payoffs are precisely the reachable and consistent payoffs. Referring to repeated games, we introduce mixed strategies which are probability distributions on pure strategies. We prove that the set of Nash equilibrium payoffs when using mixed strategies is convex and compact. Unexpectedly, this set is larger than the closed convex hull of the set of Nash equilibrium payoffs using pure strategies. We give a characterization for the Nash equilibrium payoffs using mixed strategies as reachable and consistent, these concepts being adapted to random controls. Finally, still referring to repeated games, we study the set of publicly correlated equilibrium payoffs for differential games and show that it is the same as the set of Nash equilibrium payoffs.

Introduction

We study equilibria for non zero sum differential games. In general, for a given equilibrium concept, existence and characterization of the equilibria highly depend on the strategies used by the players. There are mainly three types of strategies:

- non-anticipative strategies or memory-strategies where the choice of the current control to be played depends on the entire past history of the game (trajectory and controls played so far),
- feed-back strategies where the current control is chosen according only to the actual state of the system,

- open-loop controls where the current control is chosen at the beginning of the game and depends only on time.

In the case of deterministic differential games, there are [13], [14] and [17] existence and characterization results for Nash equilibrium payoffs. Looking for Nash equilibrium payoffs in feedback strategies, one usually computes Nash equilibrium payoffs as a functions of time and space. This leads to a system of non linear partial differential equations for which there is no general result for existence nor uniqueness of a solution. If the system admits regular enough solutions, they allow to compute the optimal feedbacks cf. [11] and [2]. There are few examples for this approach, the results essentially deal with linear quadratic differential games where solutions are sought amongst quadratic functions. For linear quadratic games, there are conditions for existence of Nash equilibria in feedback strategies and for existence and uniqueness of Nash equilibria in open-loops. Some numerical methods can be applied to compute equilibria [10]. The drawback is that feedback equilibria are highly unstable [4], except in some particular cases of one dimensional games [5].
The situation seems somehow better regarding non zero sum stochastic differential games. As for the deterministic case, there is a general result of existence and characterization [6] in case players use non-anticipative strategies. For non degenerate stochastic differential games, there is a general result for existence of a Nash equilibrium in feedback strategies in [3] using existence of smooth enough solutions for the system of partial differential equations defining the equilibrium. Another approach [12] uses BSDEs to check the existence of the solutions, prove the existence of a Nash equilibrium and optimal feedbacks. Note that the equilibria defined through this last approach are in fact equilibria in non-anticipative strategies [16] when they both exist.
Here we deal with general deterministic non zero sum differential games using mixed strategies. In our framework, "mixed strategies" refers to random combination of non anticipative strategies. The disadvantage of non-anticipative strategies is that they lack weak consistency compared to feedback strategies. Their main interest is that they allow to characterize some kind of upper hull of all Nash equilibrium payoffs using reasonable strategies.

We consider a regular non zero sum two players differential game running in finite time and we study usual equilibrium concepts such as Nash equilibrium payoffs and publicly correlated equilibrium payoffs. We are interested in the consequences of using random strategies rather than usual deterministic strategies.
More precisely, we consider a two players non zero sum differential game in $\mathbb{R}^n$ that runs in finite time $t \in [t_0, T]$. For simplicity reasons, we consider only final payoffs, noticing that running payoffs are final payoffs of an extended game.
The dynamics of the game is given by:

$$\begin{cases} \dot{x}(t) = f(x(t), u(t), v(t)) \quad t \in [t_0, T],\ u(t) \in U \text{ and } v(t) \in V \\ x(t_0) = x_0 \end{cases} \tag{1}$$

We first define the open-loop controls: we denote by $\mathcal{U}(t_0)$ (resp. $\mathcal{V}(t_0)$) the set of measurable controls of player I (resp. player II):

$$\mathcal{U}(t_0) := \{u(\cdot) : [t_0, T] \to U,\ u \text{ measurable}\}$$
$$\mathcal{V}(t_0) := \{v(\cdot) : [t_0, T] \to V,\ v \text{ measurable}\}$$

Under suitable regularity assumptions on the dynamics, if controls $u \in \mathcal{U}(t_0)$ and $v \in \mathcal{V}(t_0)$ are played, they define a unique solution of the dynamics (1) denoted by $t \mapsto X_t^{t_0,x_0,u,v}$ defined on $[t_0, T]$.
The payoffs only depend on the terminal position of the system, namely player I's payoff is $g_1(x(T))$ while player II's is $g_2(x(T))$. We assume usual regularity conditions on the payoff functions and assume Isaacs'condition in order to ensure the existence of the value for the zero sum games with payoff function g_1 or g_2. We will denote by V_1 (resp. V_2) the value function of the zero sum game where player I (resp. Player II) aims at maximizing the final payoff $g_1(x(T))$ (resp. $g_2(x(T))$) whereas the opponent aims at minimizing it.
In order to put the game in normal form, we need to define strategies. We first consider deterministic or pure strategies:

Definition 1 (Pure strategy). *A pure strategy for player I at time t_0 is a map $\alpha : \mathcal{V}(t_0) \to \mathcal{U}(t_0)$ which satisfies the following conditions:*

- *α is a measurable map from $\mathcal{V}(t_0)$ to $\mathcal{U}(t_0)$ where $\mathcal{V}(t_0)$ and $\mathcal{U}(t_0)$ are endowed with the Borel σ-field associated with the L^1 distance,*
- *α is non-anticipative with delay, i.e. there exists some delay $\tau > 0$ such that for any $v_1, v_2 \in \mathcal{V}(t_0)$, if $v_1 \equiv v_2$ a.e. on $[t_0, t]$ for some $t \in [t_0, T]$, then $\alpha(v_1) \equiv \alpha(v_2)$ a.e. on $[t_0, (t+\tau) \wedge T]$*

We denote by $\mathcal{A}(t_0)$ (resp. $\mathcal{B}(t_0)$) the set of pure strategies for player I (resp. player II) and by $\tau(\alpha)$ the delay of the strategy $\alpha \in \mathcal{A}(t_0)$.

The point of the paper is to study the impact of introducing mixed strategies on the equilibria. We define mixed strategies as finite probability distributions over pure strategies:

Definition 2 (Mixed strategy). *A mixed strategy for player I at time t_0 is a finite probability space $(\Omega_\alpha, \mathcal{P}(\Omega_\alpha), \mathbf{P}_\alpha)$ associated to a finite collection of pure strategies $(\alpha(\omega_\alpha))_{\omega_\alpha \in \Omega_\alpha}$ such that for all $\omega_\alpha \in \Omega_\alpha$, $\alpha(\omega_\alpha)$ is a pure strategy.*
From now on, mixed strategies $(\alpha, (\Omega_\alpha, \mathcal{P}(\Omega_\alpha), \mathbf{P}_\alpha))$ will be simply denoted α by abuse of notation. We denote by $\mathcal{A}_r(t_0)$ (resp. $\mathcal{B}_r(t_0)$) the set of mixed strategies for player I (resp. player II). Notice that for all mixed strategy α, there exists some delay $\tau > 0$ such that for all $\omega_\alpha \in \Omega_\alpha$, $\alpha(\omega_\alpha)$ is a non-anticipative strategy with delay greater than or equal to τ.

We now are able to recall the two usual equilibrium concepts that we study:

Definition (Nash equilibrium payoff). *The pair $(e_1, e_2) \in \mathbb{R}^2$ is a Nash equilibrium payoff for the initial position (t_0, x_0) if for all $\epsilon > 0$, there exist a strategy of player I denoted by σ_1 and a strategy for player II denoted by σ_2 such that*

- *if the strategies (σ_1, σ_2) are played, then for $i = 1, 2$, the final payoff of player i is ϵ-close to e_i*
- *for $i = 1, 2$, if player i plays a strategy $\sigma \neq \sigma_i$ whereas the opponent sticks to the strategy σ_{3-i}, the final payoff of player i will be less than or equal to the payoff rewarding the pair of strategies (σ_1, σ_2) up to ϵ.*

Definition (Publicly correlated equilibrium payoff). *The payoff* $(e_1, e_2) \in \mathbb{R}^2$ *is a publicly correlated equilibrium payoff if for all* $\epsilon > 0$*, there exist some public filtration* $(\mathcal{F}_t^\epsilon)$ *and some correlated strategies* (σ_1, σ_2) *namely strategies generating* $(\mathcal{F}_t^\epsilon)$*-measurable controls such that:*

- *if the strategies* (σ_1, σ_2) *are played, then for* $i = 1, 2$*, the final payoff of player* i *is* ϵ*-close to* e_i
- *for* $i = 1, 2$*, if player* i *plays a strategy* $\sigma \neq \sigma_i$ *such that* (σ, σ_{3-i}) *generate* $(\mathcal{F}_t^\epsilon)$*-measurable controls whereas the opponent sticks to the strategy* σ_{3-i}*, the final payoff of player* i *will be less than or equal to the payoff rewarding the pair of strategies* (σ_1, σ_2) *up to* ϵ*.*

We denote by $\mathcal{E}_c(t_0, x_0)$ *the set of all equilibrium payoffs for the initial position* (t_0, x_0)*.*

If we study Nash equilibrium payoffs when only pure strategies are allowed and denote by $\mathcal{E}(t_0, x_0)$ the set of all equilibrium payoffs for the initial position (t_0, x_0), then we may recall a well-known theorem. According to [13], [14] and [17], under usual regularity conditions, the Nash equilibrium payoffs in pure strategies are exactly the "reachable and consistent payoffs" $(e_1, e_2) \in \mathbb{R}^2$, namely payoffs satisfying: $\forall \epsilon > 0$, $\exists (u_\epsilon, v_\epsilon) \in \mathcal{U}(t_0) \times \mathcal{V}(t_0)$ such that:

- $\forall i,\ |e_i - g_i(X_T^{t_0,x_0,u_\epsilon,v_\epsilon})| \leq \epsilon$
- $\forall i,\ \forall t \in [t_0, T],\ g_i(X_T^{t_0,x_0,u_\epsilon,v_\epsilon}) \geq V_i(t, X_t^{t_0,x_0,u_\epsilon,v_\epsilon}) - \epsilon$

Furthermore, the set of Nash equilibrium payoffs is non empty: $\mathcal{E}(t_0, x_0) \neq \emptyset$.

We then study Nash equilibrium payoffs when mixed strategies are played. First of all, noticing that any pure strategy can be considered as a mixed strategy whose underlying probability space is trivial, the set of Nash equilibrium payoffs in mixed strategies is a non empty superset of $\mathcal{E}(t_0, x_0)$. We recall that using mixed strategies does not change the value of zero sum differential games. Next, we prove a characterization of Nash equilibrium payoffs in mixed strategies. Our main result (Theorem 2.1 below) states that:
The pair (e_1, e_2) is a Nash equilibrium payoff in mixed strategies iff for all $\epsilon > 0$, there exist random controls (u_ϵ, v_ϵ) on an underlying finite probability space $(\Omega, \mathcal{P}(\Omega), \mathbf{P})$ such that $\forall i = 1,\ 2$:

- $\left|\mathbf{E}[g_i(X_T^{t_0,x_0,u_\epsilon,v_\epsilon})] - e_i\right| \leq \epsilon$
- $\forall t \in [t_0, T]$, denoting by $\mathcal{F}_t = \sigma\left((u_\epsilon, v_\epsilon)(s),\ s \in [t_0, t]\right)$:

$$\mathbf{P}\left\{V_i(t, X_t^{t_0,x_0,u_\epsilon,v_\epsilon}) \leq \mathbf{E}[g_i(X_T^{t_0,x_0,u_\epsilon,v_\epsilon})|\mathcal{F}_t] + \epsilon\right\} \geq 1 - \epsilon$$

It appears that the set of Nash equilibrium payoffs in mixed strategies, denoted by $\mathcal{E}_r(t_0, x_0)$, is in fact compact, convex and generally strictly larger than the closed convex hull of the set $\mathcal{E}(t_0, x_0)$. The proof heavily relies on techniques introduced for repeated games in [1] known as "jointly controlled lotteries" and on the fact that

we work with non-anticipative strategies with delay. Note that the characterization could be given using trajectories following [17] rather than controls, provided the trajectory stems from the dynamics (1).

Finally, studying publicly correlated equilibria, we show that the set of publicly correlated equilibrium payoffs is equal to the set of Nash equilibrium payoffs using mixed strategies. The idea of the proof uses the similarity between correlated equilibrium payoffs and equilibrium payoffs of stochastic non zero sum differential games. Indeed, the characterization of equilibrium payoffs in these games due to [6] is very close to our characterization of Nash equilibrium payoffs.

We complete this introduction by describing the outline of the paper. In section 1, we recall precisely the definitions and assumptions of the differential game we study. In section 2, we give the main properties of the set of Nash equilibrium payoffs in mixed strategies. In section 3, we prove the equivalence between the sets of Nash equilibrium payoffs in mixed strategies and of publicly correlated equilibrium payoffs.

1 Definitions

1.1 Assumptions on the differential game

Throughout the paper, for any $x, y \in \mathbb{R}^n$, we will denote by $x \cdot y$ the scalar product and by $\|x\|$ the euclidian norm. The ball with center x and radius r will be denoted by $B(x, r)$. For any set S, $\mathbf{1}_S$ denotes the indicator function of S: for all $s \in S$, $\mathbf{1}_S(s) = 1$ and for all $s \notin S$, $\mathbf{1}_S(s) = 0$.
We first define more precisely the assumptions on the differential game we are dealing with. The dynamics of the game is given by (1):

$$\begin{cases} \dot{x}(t) = f(x(t), u(t), v(t)) \quad t \in [t_0, T],\ u(t) \in U \text{ and } v(t) \in V \\ x(t_0) = x_0 \end{cases}$$

where

$$\begin{cases} U \text{ and } V \text{ are compact subsets of some finite dimensional spaces} \\ U \text{ and } V \text{ have infinite cardinality,} \\ f : \mathbb{R}^n \times U \times V \to \mathbb{R}^n \text{ is bounded, continuous and uniformly} \\ \qquad \text{Lipschitz continuous with respect to } x \end{cases} \tag{2}$$

These assumptions guarantee existence and uniqueness of the trajectories generated by any pair of controls $(u, v) \in \mathcal{U}(t_0) \times \mathcal{V}(t_0)$. The assumption that U and V have infinite cardinality allows to define correlation procedures relying on as many constant controls as necessary.
We will always assume that players observe the controls played so far.
We will assume that the payoff functions g_1 and g_2 satisfy

$$g_i : \mathbb{R}^n \to \mathbb{R}, i := 1,\ 2 \text{ is Lipschitz continuous and bounded.} \tag{3}$$

In order to guarantee existence of the value functions of the two associated zero sum games, we assume Isaacs'condition: for all $(x, \xi) \in \mathbb{R}^n \times \mathbb{R}^n$

$$H(x, \xi) = \inf_{u \in U} \sup_{v \in V} f(x, u, v) \cdot \xi = \sup_{v \in V} \inf_{u \in U} f(x, u, v) \cdot \xi \tag{4}$$

1.2 Payoffs associated to a pair of strategies

In order to study equilibrium payoffs of this game we have introduced in the previous section the concepts of pure and mixed strategies. The major interest of working with non-anticipative strategies with delay is this following useful result stated in [9]:

Lemma 1.1 (Controls associated to a pair of strategies).

- *For any pair of pure strategies* $(\alpha, \beta) \in \mathcal{A}(t_0) \times \mathcal{B}(t_0)$ *there is a unique pair of controls* $(u_{\alpha\beta}, v_{\alpha\beta}) \in \mathcal{U}(t_0) \times \mathcal{V}(t_0)$ *such that* $\alpha(v_{\alpha\beta}) = u_{\alpha\beta}$ *and* $\beta(u_{\alpha\beta}) = v_{\alpha\beta}$
- *For any pair of mixed strategies* $(\alpha, \beta) \in \mathcal{A}_r(t_0) \times \mathcal{B}_r(t_0)$*, and any* $(\omega_\alpha, \omega_\beta) \in \Omega_\alpha \times \Omega_\beta$*, there is a unique pair of controls* $(u_{\omega_\alpha\omega_\beta}, v_{\omega_\alpha\omega_\beta}) \in \mathcal{U}(t_0) \times \mathcal{V}(t_0)$ *such that* $\alpha(\omega_\alpha)(v_{\omega_\alpha\omega_\beta}) = u_{\omega_\alpha\omega_\beta}$ *and* $\beta(\omega_\beta)(u_{\omega_\alpha\omega_\beta}) = v_{\omega_\alpha\omega_\beta}$*. Furthermore, the map* $(\omega_\alpha, \omega_\beta) \mapsto (u_{\omega_\alpha\omega_\beta}, v_{\omega_\alpha\omega_\beta})$ *is measurable from* $\Omega_\alpha \times \Omega_\beta$ *endowed with* $\mathcal{P}(\Omega_\alpha) \otimes \mathcal{P}(\Omega_\beta)$ *into* $\mathcal{U}(t_0) \times \mathcal{V}(t_0)$ *endowed with the Borel* σ*-field associated with the* L^1 *distance.*

Given any pair of pure strategies $(\alpha, \beta) \in \mathcal{A}(t_0) \times \mathcal{B}(t_0)$, we denote by $(X_t^{t_0,x_0,\alpha,\beta})$ the map $t \mapsto X_t^{t_0,x_0,u_{\alpha\beta},v_{\alpha\beta}}$ defined on $[t_0, T]$ where $X_\cdot^{t_0,x_0,u_{\alpha\beta},v_{\alpha\beta}}$ is the unique solution of dynamics (1).

This allows us to define the payoff associated to any pair of strategies. For $i = 1, 2$, we shall denote by

$$\begin{array}{ll} \mathfrak{J}_i(t, x, \alpha, \beta) := g_i(X_T^{t_0,x_0,\alpha,\beta}) & \text{if } (\alpha, \beta) \in \mathcal{A}(t_0) \times \mathcal{B}(t_0) \\ \mathfrak{J}_i(t, x, \alpha, \beta) := \mathbf{E}_{\alpha\beta}[g_i(X_T^{t_0,x_0,\alpha,\beta})] & \text{if } (\alpha, \beta) \in \mathcal{A}_r(t_0) \times \mathcal{B}_r(t_0) \end{array}$$

with the notation

$$\mathbf{E}_{\alpha\beta}[g_i(X_T^{t_0,x_0,\alpha,\beta})] = \int_{\Omega_\alpha \times \Omega_\beta} g_i(X_T^{t_0,x_0,u_{\omega_\alpha\omega_\beta},v_{\omega_\alpha\omega_\beta}}) \mathrm{d}\mathbf{P}_\alpha \otimes \mathrm{d}\mathbf{P}_\beta(\omega_\alpha, \omega_\beta)$$

Under regularity assumptions (2), (3) and (4), the two-players zero sum game whose payoff function is g_1 (resp. g_2) has a value. We denote by

$$V_1(t, x) := \sup_{\alpha \in \mathcal{A}(t)} \inf_{\beta \in \mathcal{B}(t)} \mathfrak{J}_1(t, x, \alpha, \beta) = \inf_{\beta \in \mathcal{B}(t)} \sup_{\alpha \in \mathcal{A}(t)} \mathfrak{J}_1(t, x, \alpha, \beta)$$

the value of the zero sum game with payoff function g_1 where player I aims at maximizing his payoff and

$$V_2(t, x) := \inf_{\alpha \in \mathcal{A}(t)} \sup_{\beta \in \mathcal{B}(t)} \mathfrak{J}_2(t, x, \alpha, \beta) = \sup_{\beta \in \mathcal{B}(t)} \inf_{\alpha \in \mathcal{A}(t)} \mathfrak{J}_2(t, x, \alpha, \beta)$$

the value of the zero sum game with payoff function g_2 where player II is the maximizer. We recall that these definitions remain unchanged whether $\alpha \in \mathcal{A}(t)$ or $\mathcal{A}_r(t)$ and $\beta \in \mathcal{B}(t)$ or $\mathcal{B}_r(t)$ cf. [8]. The assumptions also guarantee that these value functions are Lipschitz continuous.

It remains to introduce more precisely the concept of publicly correlated strategies. We first have to introduce some publicly observed random signal on the probability space $(\Omega, \mathcal{F}, \mathbf{P})$ leading to some public filtration $(\mathcal{F}_t)$ induced by the evolution of the public signal received until time t. We do not assume that Ω is finite. The following definitions are adapted from [6]. We first introduce admissible controls:

Definition 3 (Admissible control)**.** *An admissible control $\tilde{u}$ for player I is a process taking values in U progressively measurable with respect to $(\mathcal{F}_t)$. The set of admissible controls on $[t_0, T]$ is denoted by $\tilde{\mathcal{U}}(t_0)$ for player I and $\tilde{\mathcal{V}}(t_0)$ for player II.*

We will identify admissible controls and denote it by $\tilde{u}_1 \equiv \tilde{u}_2$ on $[t_0, t]$ as soon as $\mathbf{P}(\tilde{u}_1 = \tilde{u}_2 \text{ a.e. on } [t_0, t]) = 1$.
We define pairs of correlated strategies the following way:

Definition 4 (Publicly correlated strategies)**.** *A pair of correlated strategies is in fact a triplet $((\mathcal{F}_t), \alpha, \beta)$:*

- *The filtration $(\mathcal{F}_t)$ is the filtration generated by the random public signal on the probability space $(\Omega, \mathcal{F}, \mathbf{P})$ and will be referred to as the correlation device. Note that the correlation device is common knowledge for both players.*
- *a correlated strategy for player I is a map $\alpha : \tilde{\mathcal{V}}(t_0) \to \tilde{\mathcal{U}}(t_0)$ which is strongly non-anticipative with delay cf. [6]: there exists $\tau(\alpha) > 0$ such that $\forall (\mathcal{F}_t)$-stopping time S and for all $\tilde{v}_1$, $\tilde{v}_2 \in \tilde{\mathcal{V}}(t_0)$, if $\tilde{v}_1 \equiv \tilde{v}_2$ on $[\![t_0, S]\!]$, then $\alpha(\tilde{v}_1) \equiv \alpha(\tilde{v}_2)$ on $[\![t_0, (S + \tau(\alpha)) \wedge T]\!]$*
- *a correlated strategy for player II is a map $\beta : \tilde{\mathcal{U}}(t_0) \to \tilde{\mathcal{V}}(t_0)$ which is a strongly non-anticipative strategy with delay.*

From now on, we will omit the filtration in the designation of publicly correlated strategies as soon as no confusion is possible. Note that our definition is somehow broader than the usual definition of correlated strategies in repeated games.

Note that Lemma 1.1 still holds for correlated strategies: to any pair of correlated strategies (α, β) one can associate a unique pair of admissible controls $(\tilde{u}, \tilde{v})$ such that: $\alpha(\tilde{v}) = \tilde{u}$ and $\beta(\tilde{u}) = \tilde{v}$. The proof follows the scheme of the proof established for admissible strategies in [6].

Given correlated strategies (α, β), the final payoff of player i is:

$$\mathfrak{J}_i(t, x, \alpha, \beta) := \mathbf{E}[g_i(X_T^{t,x,\alpha,\beta})]$$

The expectation now refers to the probability of the random signals and not to strategies. Notice also that pure strategies are degenerated correlated strategies using some trivial filtration. Finally, note that in a zero sum game, using correlated strategies with a fixed correlation device leads to the same value as using pure strategies. Indeed, fix the filtration $(\mathcal{F}_t)$ and denote by $(\tilde{\alpha}, \tilde{\beta})$ any pair of correlated strategies

using the correlation device $(\mathcal{F}_t)$ and (α, β) any pair of pure strategies. For $i = 1, 2$:

$$\begin{aligned}
\sup_{\tilde{\beta}} \inf_{\tilde{\alpha}} \mathbf{E}[g_i(X_T^{t,x,\tilde{\alpha},\tilde{\beta}})] &\geq \sup_{\beta} \inf_{\tilde{\alpha}} \mathbf{E}[g_i(X_T^{t,x,\tilde{\alpha},\beta})] \\
&= \sup_{\beta} \inf_{\alpha} g_i(X_T^{t,x,\alpha,\beta}) = V_i(t,x) \\
&= \inf_{\alpha} \sup_{\beta} g_i(X_T^{t,x,\alpha,\beta}) \\
&= \inf_{\alpha} \sup_{\tilde{\beta}} \mathbf{E}[g_i(X_T^{t,x,\alpha,\tilde{\beta}})] \\
&\geq \inf_{\tilde{\alpha}} \sup_{\tilde{\beta}} \mathbf{E}[g_i(X_T^{t,x,\tilde{\alpha},\tilde{\beta}})]
\end{aligned}$$

On the other hand we have:

$$\sup_{\tilde{\beta}} \inf_{\tilde{\alpha}} \mathbf{E}[g_i(X_T^{t,x,\tilde{\alpha},\tilde{\beta}})] \leq \inf_{\tilde{\alpha}} \sup_{\tilde{\beta}} \mathbf{E}[g_i(X_T^{t,x,\tilde{\alpha},\tilde{\beta}})]$$

and in the end for any correlated strategies $\tilde{\alpha}, \tilde{\beta}$ using the correlation device $(\mathcal{F}_t)$, for $i = 1, 2$:

$$\sup_{\tilde{\beta}} \inf_{\tilde{\alpha}} \mathbf{E}[g_i(X_T^{t,x,\tilde{\alpha},\tilde{\beta}})] = V_i(t,x) = \inf_{\tilde{\alpha}} \sup_{\tilde{\beta}} \mathbf{E}[g_i(X_T^{t,x,\tilde{\alpha},\tilde{\beta}})]$$

1.3 Equilibrium payoffs

Given the definition of payoffs associated to a pair of strategies, we give precise definitions of equilibrium payoffs:

Definition 5 (Nash equilibrium payoff using pure strategies)**.** *The pair (e_1, e_2) is a Nash equilibrium payoff in pure strategies for the initial position (t_0, x_0) if $\forall \epsilon > 0,\ \exists (\alpha_\epsilon, \beta_\epsilon) \in \mathcal{A}(t_0) \times \mathcal{B}(t_0)$ such that*

$$\begin{cases} \forall i,\ |e_i - \mathfrak{J}_i(t_0, x_0, \alpha_\epsilon, \beta_\epsilon)| \leq \epsilon \\ \begin{cases} \forall \alpha \in \mathcal{A}(t_0),\ \mathfrak{J}_1(t_0, x_0, \alpha_\epsilon, \beta_\epsilon) \geq \mathfrak{J}_1(t_0, x_0, \alpha, \beta_\epsilon) - \epsilon \\ \forall \beta \in \mathcal{B}(t_0),\ \mathfrak{J}_2(t_0, x_0, \alpha_\epsilon, \beta_\epsilon) \geq \mathfrak{J}_2(t_0, x_0, \alpha_\epsilon, \beta) - \epsilon \end{cases} \end{cases}$$

We denote by $\mathcal{E}(t_0, x_0)$ the set of all Nash equilibrium payoffs in pure strategies for the initial position (t_0, x_0).
The characterization of these equilibria when players use pure strategies has already been studied by Kononenko [14]: Nash equilibrium payoffs are equivalent to reachable and consistent payoffs. It is the same characterization as the one known as "folk theorem" for infinitely repeated games where Nash equilibrium payoffs are precisely feasible and individually rational payoffs.

Definition 6 (Nash equilibrium payoff using mixed strategies)**.** *The payoff (e_1, e_2) is a Nash equilibrium payoff in mixed strategies for the initial position (t_0, x_0) if $\forall \epsilon > 0,\ \exists (\alpha_\epsilon, \beta_\epsilon) \in \mathcal{A}_r(t_0) \times \mathcal{B}_r(t_0)$ such that:*

1. $\forall i = 1, 2,\ |e_i - \mathfrak{J}_i(t_0, x_0, \alpha_\epsilon, \beta_\epsilon)| \leq \epsilon$

2.

$$\begin{cases} \forall \alpha \in \mathcal{A}_r(t_0),\ \mathfrak{J}_1(t_0,x_0,\alpha_\epsilon,\beta_\epsilon) \geq \mathfrak{J}_1(t_0,x_0,\alpha,\beta_\epsilon) - \epsilon \\ \forall \beta \in \mathcal{B}_r(t_0),\ \mathfrak{J}_2(t_0,x_0,\alpha_\epsilon,\beta_\epsilon) \geq \mathfrak{J}_2(t_0,x_0,\alpha_\epsilon,\beta) - \epsilon \end{cases}$$

We denote by $\mathcal{E}_r(t_0,x_0)$ the set of all Nash equilibrium payoffs for the initial position (t_0,x_0) when players use mixed strategies. We will call reachable in mixed strategies a payoff which completes only the first part of the above definition. A pair of strategies $(\alpha_\epsilon,\beta_\epsilon)$ satisfying the second point of the definition will be called ϵ-optimal. Note that in the second part of the definition, we just have to check the ϵ-optimality of α_ϵ (resp. β_ϵ) against pure strategies $\beta \in \mathcal{B}(t_0)$ (resp. $\alpha \in \mathcal{A}(t_0)$), for mixed strategies are finite linear combination of pure strategies. Namely, the second part of the definition is equivalent to

$$2. \begin{cases} \forall \alpha \in \mathcal{A}(t_0),\ \mathfrak{J}_1(t_0,x_0,\alpha_\epsilon,\beta_\epsilon) \geq \mathfrak{J}_1(t_0,x_0,\alpha,\beta_\epsilon) - \epsilon \\ \forall \beta \in \mathcal{B}(t_0),\ \mathfrak{J}_2(t_0,x_0,\alpha_\epsilon,\beta_\epsilon) \geq \mathfrak{J}_2(t_0,x_0,\alpha_\epsilon,\beta) - \epsilon \end{cases}$$

Remember we have identified pure strategies with degenerated mixed strategies implying $\mathcal{E}_r(t_0,x_0) \supset \mathcal{E}(t_0,x_0)$.

Definition 7 (Publicly correlated equilibrium payoff)**.** *The payoff $e = (e_1,e_2) \in \mathbb{R}^2$ is a publicly correlated equilibrium payoff if for all $\epsilon > 0$, there exist correlated strategies $(\alpha_\epsilon,\beta_\epsilon)$ using some correlation device $(\mathcal{F}_t^\epsilon)$ such that:*

- *for $i = 1,2$: $|\mathfrak{J}_i(t_0,x_0,\alpha_\epsilon,\beta_\epsilon) - e_i| \leq \epsilon$*
- *for all correlated strategies $((\mathcal{F}_t^\epsilon),\alpha,\beta)$:*

$$\begin{cases} \mathfrak{J}_1(t_0,x_0,\alpha,\beta_\epsilon) \leq \mathfrak{J}_1(t_0,x_0,\alpha_\epsilon,\beta_\epsilon) + \epsilon \\ \mathfrak{J}_2(t_0,x_0,\alpha_\epsilon,\beta) \leq \mathfrak{J}_2(t_0,x_0,\alpha_\epsilon,\beta_\epsilon) + \epsilon \end{cases}$$

The set of all publicly correlated equilibrium payoffs with initial conditions (t_0,x_0) will be denoted by $\mathcal{E}_c(t_0,x_0)$.

2 Nash equilibrium payoffs using mixed strategies

2.1 Characterization

Theorem 2.1 (Characterization of Nash equilibrium payoffs using mixed strategies)**.** *The payoff $e = (e_1,e_2) \in \mathbb{R}^2$ is a Nash equilibrium payoff iff for all $\epsilon > 0$, there exist random controls (u_ϵ,v_ϵ) on an underlying finite probability space $(\Omega,\mathcal{P}(\Omega),\mathbf{P})$ such that $\forall i = 1,\ 2$:*

- *e is ϵ-reachable: $|\mathbf{E}[g_i(X_T^{t_0,x_0,u_\epsilon,v_\epsilon})] - e_i| \leq \epsilon$*
- *(u_ϵ,v_ϵ) are ϵ-consistent: $\forall t \in [t_0,T]$, denoting by $\mathcal{F}_t = \sigma\left((u_\epsilon,v_\epsilon)(s),\ s \in [t_0,t]\right)$:*

$$\mathbf{P}\left\{V_i(t,X_t^{t_0,x_0,u_\epsilon,v_\epsilon}) \leq \mathbf{E}\left[g_i(X_T^{t_0,x_0,u_\epsilon,v_\epsilon})\middle|\mathcal{F}_t\right] + \epsilon\right\} \geq 1 - \epsilon$$

Proof. We start with the proof of the sufficient condition.
Consider a Nash equilibrium payoff $e = (e_1, e_2)$ and a pair of associated $\frac{\epsilon^2}{2}$-optimal mixed strategies $(\alpha_\epsilon, \beta_\epsilon)$. We will consider the random controls defined on $\Omega = \Omega_{\alpha_\epsilon} \times \Omega_{\beta_\epsilon}$ using the probability $\mathbf{P} = \mathbf{P}_{\alpha_\epsilon} \otimes \mathbf{P}_{\beta_\epsilon}$ by $(u_\epsilon, v_\epsilon)(\omega_\alpha, \omega_\beta) = (u_{\omega_\alpha \omega_\beta}, v_{\omega_\alpha \omega_\beta})$. We will denote the associated trajectories by $X^\epsilon_\cdot = X^{t_0, x_0, u_\epsilon, v_\epsilon}_\cdot$.

We have for small ϵ, for all $i = 1, 2$:

$$|\mathbf{E}[g_i(X^\epsilon_T)] - e_i| \leq \frac{\epsilon^2}{2} \leq \epsilon$$

We will prove that these controls are ϵ-consistent. Suppose on the contrary that there exist $\bar{t} \in [t_0, T]$ such that for example:

$$\mathbf{P}\left\{\mathbf{E}\left(g_1(X^\epsilon_T)\middle|\mathcal{F}_{\bar{t}}\right) \geq V_1(\bar{t}, X^\epsilon_{\bar{t}}) - \epsilon\right\} < 1 - \epsilon$$

Denote by

$$\Sigma_\epsilon := \left\{(\omega_\alpha, \omega_\beta)/\ \mathbf{E}\left(g_1(X^\epsilon_T)\middle|\mathcal{F}_{\bar{t}}\right) \geq V_1(\bar{t}, X^\epsilon_{\bar{t}}) - \epsilon\right\}$$

We have to introduce Maximin strategies

Lemma 2.2 (Maximin strategy)**.** *For all $\epsilon > 0$, for all $t \in (t_0, T)$, there exists $\tau > 0$ such that if we denote by $\mathcal{A}_\tau(t) = \{\alpha \in \mathcal{A}(t)/\ \tau(\alpha) \geq \tau\}$ there exists $\alpha^{\epsilon,t}_g$: $B(x_0, (t - t_0)\|f\|_\infty) \to \mathcal{A}_\tau(t)$ such that:*

$$\forall x \in B(x_0, (t - t_0)\|f\|_\infty),\ \inf_{v \in \mathcal{V}(t)} g_1(X^{t,x,\alpha^{\epsilon,t}_g(x)(v),v}_T) \geq V_1(t, x) - \epsilon$$

Proof of Lemma 2.2.
We will build the Maximin strategy $\alpha^{\epsilon,t}_g(\cdot)$ as a collection of finitely many pure strategies with delay. For all $x \in B(x_0, (t-t_0)\|f\|_\infty)$, there exists some pure strategy $\alpha_x \in \mathcal{A}(t)$ such that:

$$\inf_{v \in \mathcal{V}(t)} g_1(X^{t,x,\alpha_x(v),v}_T) \geq V_1(t, x) - \epsilon/2$$

For continuity reasons, there exists a Borelian partition $(O_i)_{i=1,\dots I}$ of the ball $B(x_0, (t-t_0)\|f\|_\infty)$ such that for any i there exists some $x_i \in O_i$ such that

$$\forall z \in O_i,\ \inf_{v \in \mathcal{V}(t)} g_1(X^{t,z,\alpha_{x_i}(v),v}_T) \geq V_1(t, z) - \epsilon$$

and for all $x \in B(x_0, (t - t_0)\|f\|_\infty)$, we define the Maximin strategy $\alpha^{\epsilon,t}_g(x)$ as the strategy that associates to any $v \in \mathcal{V}(t)$ the control:

$$\alpha^{\epsilon,t}_g(x)(v) = \sum_i \alpha_{x_i}(v)\mathbf{1}_{x \in O_i}$$

Note that we have by construction:

$$\forall x \in B(x_0, (t - t_0)\|f\|_\infty),\ \inf_{v \in \mathcal{V}(t)} g_1(X^{t,x,\alpha^{\epsilon,t}_g(x)(v),v}_T) \geq V_1(t, x) - \epsilon$$

As the definition of the Maximin strategy relies on a finite collection of pure strategies with delay, there exists some strictly positive delay τ such that $\forall x \in B(x_0, (t - t_0)\|f\|_\infty)$, $\alpha_g^{\epsilon,t}(x)$ is a pure strategy with delay greater than or equal to τ. $\square$

We now build a mixed strategy α defined on Ω_{α_ϵ} using $\mathbf{P}_{\alpha_\epsilon}$ in the following way: for all $v \in \mathcal{V}(t_0)$

- $\alpha(\omega_{\alpha_\epsilon})(v)(s) \equiv \alpha_\epsilon(\omega_{\alpha_\epsilon})(v)(s)$ for $s \in [t_0, \bar{t})$
- if there exist $\omega \in \Omega$ such that $(u_\epsilon, v_\epsilon)(\omega) \equiv (\alpha_\epsilon(\omega_{\alpha_\epsilon})(v), v)$ on $[t_0, \bar{t})$ and $\omega \in \Sigma_\epsilon$, then go on playing $\alpha(\omega_{\alpha_\epsilon})(v)(s) \equiv \alpha_\epsilon(\omega_{\alpha_\epsilon})(v)(s)$ for $s \in [\bar{t}, T]$
- else, play $\alpha(\omega_{\alpha_\epsilon})(v) = \alpha_g^{\frac{\epsilon}{2},\bar{t}}(X_{\bar{t}}^{t_0,x_0,\alpha(\omega_{\alpha_\epsilon})(v),v})(v|_{[\bar{t},T]})$ for all $t \in [\bar{t}, T]$

Note that $(\alpha(\omega_{\alpha_\epsilon}), \beta_\epsilon(\omega_{\beta_\epsilon}))$ generates the same controls as $(\alpha_\epsilon(\omega_{\alpha_\epsilon}), \beta_\epsilon(\omega_{\beta_\epsilon}))$ for all $(\omega_{\alpha_\epsilon}, \omega_{\beta_\epsilon}) \in \Sigma_\epsilon$ and the same controls as $(\alpha_\epsilon(\omega_{\alpha_\epsilon}), \beta_\epsilon(\omega_{\beta_\epsilon}))$ on $[t_0, \bar{t})$ if $(\omega_{\alpha_\epsilon}, \omega_{\beta_\epsilon}) \notin \Sigma_\epsilon$. Computing the payoff of (α, β_ϵ) and using the fact that Σ_ϵ is $(\mathcal{F}_{\bar{t}})$-measurable:

$$\begin{aligned}
\mathfrak{J}_1(t_0, x_0, \alpha, \beta_\epsilon) &= \mathbf{E}(g_1(X_T^{\bar{t},X_{\bar{t}}^\epsilon,\alpha_g^{\frac{\epsilon}{2},\bar{t}}(X_{\bar{t}}),\beta_\epsilon})\mathbf{1}_{\Sigma_\epsilon^c}) + \mathbf{E}(g_1(X_T^\epsilon)\mathbf{1}_{\Sigma_\epsilon}) \\
&\geq \mathbf{E}(V_1(\bar{t}, X_{\bar{t}}^\epsilon)\mathbf{1}_{\Sigma_\epsilon^c}) - \frac{\epsilon}{2}(1 - \mathbf{P}_\alpha \otimes \mathbf{P}_\beta(\Sigma_\epsilon)) + \mathbf{E}(g_1(X_T^\epsilon)\mathbf{1}_{\Sigma_\epsilon}) \\
&\geq \mathbf{E}(V_1(\bar{t}, X_{\bar{t}}^\epsilon)\mathbf{1}_{\Sigma_\epsilon^c}) - \frac{\epsilon}{2}(1 - \mathbf{P}(\Sigma_\epsilon)) + \mathbf{E}(g_1(X_T^\epsilon)\mathbf{1}_{\Sigma_\epsilon}) \\
&\geq \mathbf{E}(\mathbf{E}(g_1(X_T^\epsilon)|\mathcal{F}_{\bar{t}})\mathbf{1}_{\Sigma_\epsilon^c}) + \mathbf{E}(g_1(X_T^\epsilon)\mathbf{1}_{\Sigma_\epsilon}) + \frac{\epsilon}{2}(1 - \mathbf{P}(\Sigma_\epsilon)) \\
&\geq \mathbf{E}(g_1(X_T^\epsilon)\mathbf{1}_{\Sigma_\epsilon^c}) + \mathbf{E}(g_1(X_T^\epsilon)\mathbf{1}_{\Sigma_\epsilon}) + \frac{\epsilon}{2}(1 - \mathbf{P}(\Sigma_\epsilon)) \\
&> \mathfrak{J}_1(t_0, x_0, \alpha_\epsilon, \beta_\epsilon) + \frac{\epsilon^2}{2}
\end{aligned}$$

This is in contradiction with the $\frac{\epsilon^2}{2}$-optimality of $(\alpha_\epsilon, \beta_\epsilon)$.

We now will prove the necessary condition.

Consider some payoff $e = (e_1, e_2)$ reachable and consistent as in Proposition 2.1. For all $\epsilon > 0$, we will build ϵ-optimal strategies rewarding a payoff ϵ close to e.

Fix $\epsilon > 0$. Set δ small enough such that:

1. $\forall t \in [t_0, T]$, $\forall x \in \mathbb{R}^n$, $\forall y \in B(x, \delta\|f\|_\infty)$, for all $i = 1, 2$:

$$|V_i(t, x) - V_i(t + \delta, y)| \leq \epsilon \tag{5}$$

2. $\forall t \in [t_0, T]$, $\forall x \in \mathbb{R}^n$, $\forall y \in B(x, \delta\|f\|_\infty)$, $\forall (u, v) \in \mathcal{U}(t) \times \mathcal{V}(t)$, for all $i = 1, 2$:

$$|g_i(X_T^{t,x,u,v}) - g_i(X_T^{t,y,u,v})| \leq \epsilon \tag{6}$$

3. $\exists N \in \mathbb{N}^*$ such that $N\delta = T - t_0$

We introduce the time partition $(\theta_0 = t_0, \ldots, \theta_k = t_0 + k\delta, \ldots, \theta_N = T)$.

Set $\eta = \frac{\epsilon}{N}$. Using the assumption, choose random controls (u_η, v_η) on $(\Omega, \mathcal{P}(\Omega), \mathbf{P})$ rewarding a payoff η-close to e and η-consistent, denoting by $(\mathcal{F}_t)$ the filtration $(\mathcal{F}_t) = (\sigma\{(u_\eta, v_\eta)(s),\ s \in [t_0, t]\})$:

$$\mathbf{P}\left\{V_i(t, X_t^{t_0,x_0,u_\eta,v_\eta}) \leq \mathbf{E}[g_i(X_T^{t_0,x_0,u_\eta,v_\eta})|\mathcal{F}_t] + \eta\right\} \geq 1 - \eta \tag{7}$$

We will set $X_\cdot^\eta = X_\cdot^{t_0,x_0,u_\eta,v_\eta}$ and for any $\omega \in \Omega$: $X_\cdot^\eta(\omega) = X_\cdot^{t_0,x_0,(u_\eta,v_\eta)(\omega)}$.
If the random controls are in fact deterministic, we already know a way to build some pure strategies $(\alpha_\epsilon, \beta_\epsilon)$ that are ϵ-optimal and reward a payoff ϵ-close to e (cf. the construction of Proposition 6.1 in [17] for example). If the controls (u_η, v_η) are real random controls, we have to build ϵ-optimal mixed strategies rewarding a payoff ϵ-close to e. The idea of the optimal strategies $(\alpha_\epsilon, \beta_\epsilon)$ is to build "trigger" mixed strategies that are correlated in order to generate controls close to (u_η, v_η). We will use some correlation device depending on a jointly correlated lottery at each "node" of the trajectories generated by (u_η, v_η) and, if the opponent does not play the expected control, the player who detected the deviation swaps to the "punitive strategy". The proof proceeds in several steps. First of all, as we have to build correlation devices, we build random controls on a rational probability space. Then we build correlation devices for each "node". Finally, using these correlations, we build the optimal strategies, and check that they reward a payoff close to e and that they are optimal.

To begin with, we introduce the explosions that are kind of "nodes" in the trajectories generated by (u_η, v_η):

Definition 8 (Explosion)**.** *Consider a pair of random controls (u_ϵ, v_ϵ) on a finite probability space $(\Omega, \mathcal{P}(\Omega), \mathbf{P})$ generating finitely many trajectories. We will call explosion any $\bar{t} \in [t_0, T)$ such that there exists $(\omega_1, \omega_2) \in \Omega^2$ such that $\bar{t} = \sup\{t/\ (u_\epsilon, v_\epsilon)(\omega_1)(s) \equiv (u_\epsilon, v_\epsilon)(\omega_2)(s)$ on $[t_0, t]\}$.*

Assume that (u_η, v_η) generates $\bar{M}$ distinct pairs of deterministic controls with $\bar{M} \geq 2$ and M explosions with $1 \leq M \leq \bar{M}-1$ denoted by $\{\tau_i\}$. We introduce an auxiliary time step τ to be defined later such that $\tau < \min_{j \neq k} |\tau_j - \tau_k|/2$, $\tau < T - \max_j \tau_j$ and $\exists \bar{N} \in \mathbb{N} \backslash \{0, 1\}$ such that $\bar{N}\tau = \delta$. This ensures that there is no explosion on $[T - \tau, T]$. We introduce another time partition $(t_0, \ldots, t_k = t_0 + k\tau, \ldots, t_{N\bar{N}} = T)$.

We will have to build a jointly correlated lottery for each explosion, and we therefore need to approximate the real probability $\mathbf{P}$ with a rational probability $\mathbf{Q}$ for it is easier to build jointly controlled lotteries for rational probabilities:

Rational approximation of the real probability:
Consider some finite probability space $(A, \mathcal{P}(A), \mathbf{P}_A)$. Set $A = \{a_1 \ldots a_M\}$. It is easy to prove that for all $\nu > 0$, there exist M rational numbers $\frac{m_i}{m}$ with $m_i \in \mathbb{N}$ for all i and $m \in \mathbb{N}^*$ such that

- for all $i = 1 \ldots M$, $|\frac{m_i}{m} - \mathbf{P}_A(a_i)| \leq \nu$,
- for all $i = 1 \ldots M$, $m_i \geq 0$
- $\sum_{i=1}^M \frac{m_i}{m} = 1$

Defining $\mathbf{Q}_A$ by $\mathbf{Q}_A(a_i) = \frac{m_i}{m}$ gives a rational approximation of the probability $\mathbf{P}_A$ up to ν on $(A, \mathcal{P}(A))$.

Consider now the probability space $(\Omega, \mathcal{P}(\Omega), \mathbf{P})$ on which the controls (u_η, v_η) are defined. We need a rational approximation $\mathbf{Q}$ of $\mathbf{P}$ generating conditional probabilities $\mathbf{Q}(\cdot|\mathcal{F}_t)$ close to $\mathbf{P}(\cdot|\mathcal{F}_t)$ for all $t \in [t_0, T]$. Consider $\mathcal{F}_T$, the filtration at time T, and assume it is generated by the atoms $\{\Omega_j\}_{j=1\dots\bar{M}}$. Set $p = \min_{j=1\dots\bar{M}} \mathbf{P}(\Omega_j) > 0$. Define $\mathbf{Q}$ on $(\Omega, \mathcal{P}(\Omega), \mathcal{F}_T)$ as the rational approximation of the probability $\mathbf{P}$ up to $\nu\frac{p}{2}$. This means:

$$\text{for all } F \in \mathcal{F}_T,\ |\mathbf{P}(F) - \mathbf{Q}(F)| \leq \frac{\bar{M}\nu p}{2} \leq \bar{M}\nu \tag{8}$$

We will check that this rational probability is close to $\mathbf{P}$ if we consider conditional probabilities: for any $t \in [t_0, T]$, assume that $\mathcal{F}_t = \sigma(\{\Omega_i^t\}_{i\in I})$ where the Ω_i^t are the atoms of the σ-algebra $\mathcal{F}_t$ and take any $F \in \mathcal{F}_T$. For any $i \in I$:

$$\begin{aligned}
|\mathbf{P}(F|\Omega_i^t) - \mathbf{Q}(F|\Omega_i^t)| &= |\frac{\mathbf{P}(F\cap\Omega_i^t)}{\mathbf{P}(\Omega_i^t)} - \frac{\mathbf{Q}(F\cap\Omega_i^t)}{\mathbf{P}(\Omega_i^t)}| + \mathbf{Q}(F\cap\Omega_i^t)|\frac{1}{\mathbf{P}(\Omega_i^t)} - \frac{1}{\mathbf{Q}(\Omega_i^t)}| \\
&\leq \frac{\bar{M}\nu p}{2\mathbf{P}(\Omega_i^t)} + \frac{\bar{M}\nu p}{2}\frac{\mathbf{Q}(F\cap\Omega_i^t)}{\mathbf{Q}(\Omega_i^t)\mathbf{P}(\Omega_i^t)} \\
&\leq \frac{\bar{M}\nu p}{2\mathbf{P}(\Omega_i^t)}(1 + \mathbf{Q}(F|\Omega_i^t)) \\
&\leq \frac{2\bar{M}\nu p}{2p} = \bar{M}\nu
\end{aligned}$$

As the probability space is finite, we have $\mathbf{P}(F|\mathcal{F}_t) = \sum_{i\in I} \mathbf{P}(F|\Omega_i^t)\mathbf{1}_{\Omega_i^t}$ and we immediately get:

$$\text{for all } t \in [t_0, T],\ \text{for all } F \in \mathcal{F}_T:\ |\mathbf{Q}(F|\mathcal{F}_t) - \mathbf{P}(F|\mathcal{F}_t)| \leq \bar{M}\nu \tag{9}$$

Probability change:
From now on, we will consider the same set of random controls (u_η, v_η) but defined on the underlying finite probability space with rational probability $(\Omega, \mathcal{F}_T, \mathbf{Q})$, and we will omit to mention the subscript $\mathbf{Q}$ when writing expectations. We will check that under the probability $\mathbf{Q}$, the controls (u_η, v_η) still reward a payoff close to e and are still consistent:

Lemma 2.3. *For ν small enough,we have for all $i = 1, 2$:*

$$\left|\mathbf{E}_\mathbf{Q}[g_i(X_T^\eta)] - e_i\right| \leq 2\eta$$

and for all $t \in [t_0, T]$:

$$\mathbf{Q}\left\{V_i(t, X_t^\eta) \leq \mathbf{E}_\mathbf{Q}\left(g_i(X_T^\eta)\middle|\mathcal{F}_t\right) + 2\eta\right\} \geq 1 - 2\eta$$

Proof. Let us choose ν such that $\nu\bar{M}^2 \max(\|g_1\|_\infty, \|g_2\|_\infty, 1) \leq \eta$.
For the first inequality, if $\mathcal{F}_T = \sigma(\{\Omega_j\}_{j=1\dots\bar{M}})$ where the Ω_j are the atoms of $\mathcal{F}_T$, we have for all $i = 1, 2$:

$$\left|\mathbf{E}_\mathbf{Q}(g_i(X_T^\eta)) - e_i\right| = \left|\sum_{j=1}^{\bar{M}} g_i(X_T^\eta(\omega_j))\mathbf{Q}(\Omega_j) - e_i\right|$$

where for all j, we choose some $\omega_j \in \Omega_j$. Therefore:

$$\begin{aligned}\left|\mathbf{E_Q}(g_i(X_T^\eta)) - e_i\right| &\leq \left|\sum_{j=1}^{\bar{M}} g_i(X_T^\eta(\omega_j))\mathbf{P}(\Omega_j) - e_i\right| \\ &\quad + \sum_{j=1}^{\bar{M}} \left|g_i(X_T^\eta(\omega_j))(\mathbf{P}(\Omega_j) - \mathbf{Q}(\Omega_j))\right| \\ &\leq \left|\mathbf{E_P}(g_i(X_T^\eta)) - e_i\right| + \sum_{j=1}^{\bar{M}} \|g_i\|_\infty \bar{M}\nu \text{ due to (8)} \\ &\leq \eta + \bar{M}^2\nu\|g_i\|_\infty \leq 2\eta\end{aligned}$$

For the second inequality, we fix $t \in [t_0, T]$ and still assume $\mathcal{F}_T = \sigma(\{\Omega_j\}_{j=1\ldots\bar{M}})$ where the Ω_j are the atoms of $\mathcal{F}_T$. For all $j = 1 \ldots \bar{M}$, we choose some $\omega_j \in \Omega_j$. For $i = 1, 2$, we have thanks to (9):

$$\begin{aligned}&\left|\mathbf{E_Q}\left(g_i(X_T^\eta)\middle|\mathcal{F}_t\right) - \mathbf{E_P}\left(g_i(X_T^\eta)\middle|\mathcal{F}_t\right)\right| \\ &\leq \sum_{j=1}^{\bar{M}} \left|g_i(X_T^\eta(\omega_j))\mathbf{Q}(\Omega_j|\mathcal{F}_t) - g_i(X_T^\eta(\omega_j))\mathbf{P}(\Omega_j|\mathcal{F}_t)\right| \\ &\leq \|g_i\|_\infty \sum_{j=1}^{\bar{M}} \bar{M}\nu \leq \eta\end{aligned}$$

We rewrite:

$$\mathbf{E_P}\left(g_i(X_T^\eta)\middle|\mathcal{F}_t\right) \leq \mathbf{E_Q}\left(g_i(X_T^\eta\middle|\mathcal{F}_t\right) + \eta \tag{10}$$

We will set:

$$\Sigma_t^i(\mathbf{P}) = \left\{V_i(t, X_t^\eta) \leq \mathbf{E_P}\left(g_i(X_T^\eta)\middle|\mathcal{F}_t\right) + \eta\right\}$$

where $\mathbf{P}(\Sigma_t^i(\mathbf{P})) \geq 1 - \eta$ by assumption (7) and

$$\Sigma_t^i(\mathbf{Q}) = \left\{V_i(t, X_t^\eta) \leq \mathbf{E_Q}\left(g_i(X_T^\eta)\middle|\mathcal{F}_t\right) + 2\eta\right\}$$

We have $\Sigma_t^i(\mathbf{P}) \subseteq \Sigma_t^i(\mathbf{Q})$ using (10). Both sets are in $\mathcal{F}_t$ implying due to (8):

$$\mathbf{Q}(\Sigma_t^i(\mathbf{Q})) \geq \mathbf{Q}(\Sigma_t^i(\mathbf{P})) \geq \mathbf{P}(\Sigma_t^i(\mathbf{P})) - \bar{M}\nu \geq 1 - 2\eta$$

□

We now will explain how to correlate the strategies at each explosion using jointly controlled lotteries.

Explosion procedure:

Suppose $\bar{\tau}$ is an explosion with $\bar{\tau} \in [t_k, t_{k+1})$ and $\bar{\tau}$ is associated to $\omega_1, \omega_2 \in \Omega$ as in the definition. We assume that the filtration $\mathcal{F}_{t_k}$ is generated by the atoms $\{\Omega_l\}_{l \in L}$. We have for some l: $\omega_1, \omega_2 \in \Omega_l$. By definition of the delay τ, there is no other explosion on (t_k, t_{k+1}). The definition of an explosion allows us to set $\Omega_l := \bigsqcup_{i=1}^I \Omega_i^l$ with $\Omega_i^l \in \mathcal{F}_{t_{k+1}}$, $2 \leq I \leq \bar{M}$ and for all $\omega_\iota, \omega_j \in \Omega_i^l$, $(u_\eta, v_\eta)(\omega_\iota)(s) \equiv (u_\eta, v_\eta)(\omega_j)(s)$ on $[t_0, t_{k+1})$. We consider the rational conditional probabilities $\mathbf{Q}(\Omega_i^l|\Omega_l) = \frac{q_i(t_k^l)}{q(t_k^l)}$.

We build a jointly controlled lottery as in [1]: consider the auxiliary two players process with outcome matrix G, $q(t_k^l)$ distinct actions $u_a : [t_k, t_{k+1}] \to U$ for player I and $q(t_k^l)$ distinct actions $v_b : [t_k, t_{k+1}] \to V$ for player II. Note that as we assumed that U and V have infinite cardinality, we can define correlation controls (u_a, v_b) as distinct constant controls and use distinct controls for each explosion. The matrix G is build in such a way that the only possible outcomes are $G(a, b) \in \{1 \dots I\}$ and each row and each column of G contains exactly $q_i(t_k^l)$ times the outcome i for all $i \in \{1 \dots I\}$. Note that if player II plays some fixed v_b and player I plays each u_a with equiprobability $\frac{1}{q(t_k^l)}$, then the outcome will be i with probability $\frac{q_i(t_k^l)}{q(t_k^l)} = \mathbf{Q}(\Omega_i^l|\Omega_l)$ and symmetrically, if player I plays some fixed u_a and player II plays each v_b with probability $\frac{1}{q(t_k^l)}$, then the outcome will be i with probability $\mathbf{Q}(\Omega_i^l|\Omega_l)$. Note that this correlation procedure allows the players to correlate their controls on any Ω_i^l with probability $\mathbf{Q}(\Omega_i^l|\Omega_l)$ in such a way that no unilateral cheating in the use of the correlation controls may change the outcome of the correlation matrix G.

We introduce a way to punish the opponent if he is not playing the expected control through punitive strategies:

Lemma 2.4 (Punitive strategy)**.** *For all $\epsilon > 0$, for all $t \in (t_0, T)$, there exists $\tau > 0$ such that if we denote by $\mathcal{A}_\tau(t) = \{\alpha \in \mathcal{A}(t)/\ \tau(\alpha) \geq \tau\}$ there exists $\alpha_p^{\epsilon,t} : B(x_0, (t - t_0)\|f\|_\infty) \to \mathcal{A}_\tau(t)$ such that:*

$$\forall x \in B(x_0, (t - t_0)\|f\|_\infty), \quad \sup_{v \in \mathcal{V}(t)} g_2(X_T^{t,x,\alpha_p^{\epsilon,t}(x)(v),v}) \leq V_2(t, x) + \epsilon$$

Proof of Lemma 2.4.
The proof is similar to the proof of Lemma 2.2. We will build the punitive strategy $\alpha_p^{\epsilon,t}(\cdot)$ as a collection of finitely many pure strategies with delay. For all $x \in B(x_0, (t - t_0)\|f\|_\infty)$, there exists some pure strategy $\alpha_x \in \mathcal{A}(t)$ such that:

$$\sup_{v \in \mathcal{V}(t)} g_2(X_T^{t,x,\alpha_x(v),v}) \leq V_2(t, x) + \epsilon/2$$

For continuity reasons, there exists a Borelian partition $(O_i)_{i=1,\dots I}$ of the ball $B(x_0, (t - t_0)\|f\|_\infty)$ such that for any i there exists some $x_i \in O_i$ such that

$$\forall z \in O_i, \quad \sup_{v \in \mathcal{V}(t)} g_2(X_T^{t,z,\alpha_{x_i}(v),v}) \leq V_2(t, z) + \epsilon$$

and for all $x \in B(x_0, (t - t_0)\|f\|_\infty)$, we define the punitive strategy $\alpha_p^{\epsilon,t}(x)$ as the strategy that associates to any $v \in \mathcal{V}(t)$ the control:

$$\alpha_p^{\epsilon,t}(x)(v) = \sum_i \alpha_{x_i}(v) \mathbf{1}_{x \in O_i}$$

Note that we have by construction:

$$\forall x \in B(x_0, (t - t_0)|f\|_\infty), \quad \sup_{v \in \mathcal{V}(t)} g_2(X_T^{t,x,\alpha_g^{\epsilon,t}(x)(v),v}) \leq V_2(t, x) + \epsilon$$

As for the Maximin strategy, there exists some strictly positive delay τ such that $\forall x \in B(x_0, (t-t_0)\|f\|_\infty)$, $\alpha_p^{\epsilon,t}(x)$ is a pure strategy with delay greater than or equal to τ. $\square$

We now have everything needed to define the ϵ-optimal strategies.
Definition of the strategies $(\alpha_\epsilon, \beta_\epsilon)$:
We recall that the idea of the strategy for Player I is to play the same control as $u_\eta(\omega)$, $\omega \in \Omega$ as long as there is no explosion and as long as player II plays $v_\eta(\omega)$. If an explosion takes place on $[t_k, t_{k+1})$ meaning $\mathcal{F}_{t_{k+1}}$ is generated by the atoms $(\Omega_i)_{i\in I}$, play on this interval some correlation control as defined by the corresponding explosion procedure, then observe at t_{k+1} the correlation control played by the opponent on $[t_k, t_{k+1})$ and deduce from the explosion procedure on which Ω_i the game is now correlated and play $u_\eta(\omega_i)$, $\omega_i \in \Omega_i$ from t_{k+1} on until the next explosion as long as player II plays $v_\eta(\omega_i)$. Player I repeats the same procedure at each explosion. As soon as player I detects that Player II played some unexpected control, he swaps to the punitive strategy.

In order to define the strategy in a more convenient way, we have to introduce some auxiliary random processes depending only on the past, namely $\bar{\Omega}$ keeping the information on which trajectory generated by (u_η, v_η) is currently being followed and $\mathbf{S}$ such that $\mathbf{S} = \emptyset$ if no deviation was observed in the past and $\mathbf{S} = (t_k, x)$ where $t_k \in \{t_0 \dots t_{N\bar{N}}\}$ means that some deviation occurred on $[t_k, t_{k+1})$ and the punitive strategy is to be played from the state (t_{k+2}, x) because there is a delay between the time at which deviation is detected and the time from which punitive strategy is played.

First of all, in order to build the strategy α_ϵ for example, we will define the associated underlying finite probability space. We will define it by induction on the number of explosions. We will always assume that an explosion procedure is defined using constant correlation controls that are not used in any other explosion procedure:

1. If there is no explosion, take any trivial probability space with only one element.

2. Assume that we can associate to any pair of controls generating a number of explosion lower than or equal to n a probability space $(\Omega_n, \mathcal{P}(\Omega_n), \mathbf{P}_n)$.

3. Assume we are facing a couple of random controls generating $n+1$ explosions. Consider the first explosion, that takes place in $[t_k, t_{k+1})$ for some $k \in \{0, \dots, N\bar{N}-1\}$ and such that $\mathcal{F}_{t_{k+1}}$ is generated by the atoms $(\Omega_i)_{i=1\dots I}$. Note that the pairs of random controls defined on each Ω_i generate at most n explosions and can therefore be associated to a probability space $(\Omega_{n_i}, \mathcal{P}(\Omega_{n_i}), \mathbf{P}_{n_i})$. Assume now that the set of correlation controls defined by the first explosion procedure is $\Omega = \{u_a\}_{a=1\dots q(t_k)}$ associated to the probability $\mathbf{P}$ such that for all a, $\mathbf{P}(u_a) = 1/q(t_k)$. We now define the probability space $(\Omega_{n+1}, \mathcal{P}(\Omega_{n+1}), \mathbf{P}_{n+1})$ the following way:

   - $\Omega_{n+1} = \Omega \times \Pi_{i=1}^I \Omega_{n_i}$
   - for all $\omega_{n+1} = (u_a, \omega_{n_1}, \dots \omega_{n_I}) \in \Omega_{n+1}$, $\mathbf{P}_{n+1}(\omega_{n+1}) = \frac{1}{q(t_k)} \times \Pi_{i=1}^I \mathbf{P}_{n_i}(\omega_{n_i})$

It is easy to see that this is a well defined finite probability space.

We now have defined Ω_{α_ϵ}. Note that any $\omega_{\alpha_\epsilon} \in \Omega_{\alpha_\epsilon}$ prescribes one correlation control for any of the possible explosion procedures. Fix any sequence of correlation controls (u_i) possibly leading to the explosion $\bar{\tau} \in [t_k, t_{k+1})$ associated to the atom Ω_l of $\mathcal{F}_{t_k}$. Consider the set of correlation controls $\{u_a\}$ associated to this explosion. Then, the conditional probability of each u_a given (u_i) is by definition $\frac{1}{q(t_k^l)}$:

$$\mathbf{P}_{\alpha_\epsilon}[\omega_{\alpha_\epsilon} \ni u_a | \omega_{\alpha_\epsilon} \ni (u_i)] = \frac{1}{q(t_k^l)} \tag{11}$$

We now define the strategy α_ϵ using auxiliary random processes:

$$\mathbf{S}^{\alpha_\epsilon} : \Omega_{\alpha_\epsilon} \times \mathcal{V}(t_0) \times \{t_k\}_{k=0\ldots N\bar{N}} \to \emptyset \cup ([t_0, T] \times \mathbb{R}^n)$$

and

$$\bar{\Omega}^{\alpha_\epsilon} : \Omega_{\alpha_\epsilon} \times \mathcal{V}(t_0) \times \{t_k\}_{k=0\ldots N\bar{N}} \to \mathcal{F}_T.$$

At time t_0, for any ω_{α_ϵ}, for any control $v \in \mathcal{V}(t_0)$, we set $\mathbf{S}^{\alpha_\epsilon}_{t_0}(\omega_{\alpha_\epsilon}, v) = \emptyset$ and $\bar{\Omega}^{\alpha_\epsilon}_{t_0}(\omega_{\alpha_\epsilon}, v) = \Omega$ and fix $u_0 \in U$. For all $k \in \{0, \ldots N\bar{N} - 1\}$, if $\alpha_\epsilon(\omega_{\alpha_\epsilon})(v)$ is built on $[t_0, t_k)$, we define $\alpha_\epsilon(\omega_{\alpha_\epsilon})(v)$ further by:

1. If $\mathbf{S}^{\alpha_\epsilon}_{t_k}(\omega_{\alpha_\epsilon}, v) \neq \emptyset$, for example $\mathbf{S}^{\alpha_\epsilon}_{t_k}(\omega_{\alpha_\epsilon}, v) = (t_i, x)$, this means that player II did not play the expected control from $t_s \in [t_i, t_{i+1})$ on, then play the punitive strategy $\alpha_\epsilon(v)|_{[t_k, t_{k+1})} = \alpha_p^{\eta, t_{i+2}}(x)(v|_{[t_{i+2}, T]})|_{[t_k, t_{k+1})}$ as defined in Lemma 2.4 and set $\bar{\Omega}^{\alpha_\epsilon}_{t_{k+1}}(\omega_{\alpha_\epsilon}, v) = \emptyset$ and $\mathbf{S}^{\alpha_\epsilon}_{t_{k+1}}(\omega_{\alpha_\epsilon}, v) = \mathbf{S}^{\alpha_\epsilon}_{t_k}(\omega_{\alpha_\epsilon}, v)$.

2. If $\mathbf{S}^{\alpha_\epsilon}_{t_k}(\omega_{\alpha_\epsilon}, v) = \emptyset$, then

   - if there is no explosion on $[t_k, t_{k+1})$ for $(u_\eta, v_\eta)(\omega)$, $\omega \in \bar{\Omega}^{\alpha_\epsilon}_{t_k}(\omega_{\alpha_\epsilon}, v)$, then play $\alpha_\epsilon(\omega_{\alpha_\epsilon})(v)|_{[t_k, t_{k+1})} = u_\eta(\omega)|_{[t_k, t_{k+1})}$ for some $\omega \in \bar{\Omega}^{\alpha_\epsilon}_{t_k}(\omega_{\alpha_\epsilon}, v)$ and set $\bar{\Omega}^{\alpha_\epsilon}_{t_{k+1}}(\omega_{\alpha_\epsilon}, v) = \bar{\Omega}^{\alpha_\epsilon}_{t_k}(\omega_{\alpha_\epsilon}, v)$. If $k \geq 1$ and if $v|_{[t_{k-1}, t_k]} \not\equiv v_\eta(\omega)|_{[t_{k-1}, t_k]}$ for all $\omega \in \bar{\Omega}^{\alpha_\epsilon}_{t_k}(\omega_{\alpha_\epsilon}, v)$ then set $\mathbf{S}^{\alpha_\epsilon}_{t_{k+1}}(\omega_{\alpha_\epsilon}, v) = (t_{k-1}, X^{t_0, x_0, \alpha_\epsilon, v}_{t_{k+1}})$, else set $\mathbf{S}^{\alpha_\epsilon}_{t_{k+1}}(\omega_{\alpha_\epsilon}, v) = \emptyset$
   - if there is an explosion on $[t_k, t_{k+1})$ for $(u_\eta, v_\eta)(\omega)$, $\omega \in \bar{\Omega}^{\alpha_\epsilon}_{t_k}(\omega_{\alpha_\epsilon}, v)$, play on $[t_k, t_{k+1})$ the controls u_a in ω_{α_ϵ} corresponding to the current explosion procedure then consider the control v played by player II on $[t_{k-1} \vee t_0, t_k + \frac{\tau}{2}]$:
     - if $k \geq 1$ and if $v|_{[t_{k-1}, t_k]} \not\equiv v_\eta(\omega)|_{[t_{k-1}, t_k]}$ for all $\omega \in \bar{\Omega}^{\alpha_\epsilon}_{t_k}(\omega_{\alpha_\epsilon}, v)$ then set $\mathbf{S}^{\alpha_\epsilon}_{t_{k+1}}(\omega_{\alpha_\epsilon}, v) = (t_{k-1}, X^{t_0, x_0, \alpha_\epsilon, v}_{t_{k+1}})$ and $\bar{\Omega}^{\alpha_\epsilon}_{t_{k+1}}(\omega_{\alpha_\epsilon}, v) = \emptyset$ and define $\alpha_\epsilon(\omega_{\alpha_\epsilon})$ further using the procedure at step 1, else
     - if $v|_{[t_k, t_k + \frac{\tau}{2}]} \not\equiv v_b$ for any of the v_b prescribed by the explosion procedure, then set $\bar{\Omega}^{\alpha_\epsilon}_{t_{k+1}}(\omega_{\alpha_\epsilon}, v) = \emptyset$ and $\mathbf{S}^{\alpha_\epsilon}_{t_{k+1}}(\omega_{\alpha_\epsilon}, v) = \emptyset$. If $k < N\bar{N} - 1$, play $\alpha_\epsilon(\omega_{\alpha_\epsilon})(v)|_{[t_{k+1}, t_{k+2})} = u_0$ and set $\mathbf{S}^{\alpha_\epsilon}_{t_{k+2}}(\omega_{\alpha_\epsilon}, v) = (t_k, X^{t_0, x_0, \alpha_\epsilon, v}_{t_{k+2}})$ and $\bar{\Omega}^{\alpha_\epsilon}_{t_{k+2}}(\omega_{\alpha_\epsilon}, v) = \emptyset$.
     - else, player II played one of the expected controls for example v_b. Assume $\bar{\Omega}^{\alpha_\epsilon}_{t_k}(\omega_{\alpha_\epsilon}, v) = \Omega_l$. Consider $G(a, b) = \kappa$ and set $\bar{\Omega}^{\alpha_\epsilon}_{t_{k+1}}(\omega_{\alpha_\epsilon}, v) = \Omega^l_\kappa$ and $\mathbf{S}^{\alpha_\epsilon}_{t_{k+1}}(\omega_{\alpha_\epsilon}, v) = \emptyset$. If $k < N\bar{N} - 1$, play $\alpha_\epsilon(\omega_{\alpha_\epsilon})(v)|_{[t_{k+1}, t_{k+2})} = u_\eta(\omega)|_{[t_{k+1}, t_{k+2})}$ for some $\omega \in \bar{\Omega}^{\alpha_\epsilon}_{t_{k+1}}(\omega_{\alpha_\epsilon}, v)$ and set $\bar{\Omega}^{\alpha_\epsilon}_{t_{k+2}}(\omega_{\alpha_\epsilon}, v) = \bar{\Omega}^{\alpha_\epsilon}_{t_{k+1}}(\omega_{\alpha_\epsilon}, v)$ and $\mathbf{S}^{\alpha_\epsilon}_{t_{k+2}}(\omega_{\alpha_\epsilon}, v) = \emptyset$.

Note that this procedure indeed defines a mixed strategy. It also ensures that for all $k = 0, \dots, N\bar{N}$, at time t_k, $\bar{\Omega}^{\alpha_\epsilon}_{t_k}(v)$ is either the empty set or one of the atoms of $\mathcal{F}_{t_k}$ and $\{\mathbf{S}^{\alpha_\epsilon}(v) \in \{t_k\} \times \mathbb{R}^n\} \in \mathcal{F}^{\alpha_\epsilon, v}_{t_k}$ where $\mathcal{F}^{\alpha_\epsilon, v}_{t_k} = \sigma((\alpha_\epsilon(v), v)(s),\ s \in [t_0, t_k])$.

The strategy β_ϵ is built symmetrically using the auxiliary random processes $\bar{\Omega}^{\beta_\epsilon}$ and $\mathbf{S}^{\beta_\epsilon}$.

Payoff of the strategies $(\alpha_\epsilon, \beta_\epsilon)$:
We will first study the controls generated if player I plays α_ϵ and player II plays some pure strategy β with delay $\tau(\beta)$ such that β generates no deviation. We will say that β generates no deviation as soon as for all $k \in \{0, \dots, N\bar{N}\}$, $\mathbf{S}^{\alpha_\epsilon}_{t_k}(\beta) = \emptyset$ (equivalently $\mathbf{S}^{\alpha_\epsilon}_T(\beta) = \emptyset$), even if $\mathbf{S}^{\alpha_\epsilon}_T(\beta) = \emptyset$ does not imply that the control generated by β on $[T - \tau, T]$ is one of the v_η.
We will first consider the values taken by the process $\bar{\Omega}^{\alpha_\epsilon}(\beta)$.

Lemma 2.5. *If the strategies (α_ϵ, β) are played where β is some pure strategy with delay such that for all $k \in \{0, \dots, N\bar{N}\}$, $\mathbf{S}^{\alpha_\epsilon}_{t_k}(\beta) = \emptyset$, then for all $k \in \{0, \dots, N\bar{N}\}$, for all $F \in \mathcal{F}_{t_k}$:*

$$\mathbf{P}_{\alpha_\epsilon}\left[\bar{\Omega}^{\alpha_\epsilon}_{t_k}(\beta) \subset F\right] = \mathbf{Q}(F)$$

Proof. We will prove the Lemma by induction on k for all F such that F is an atom of the filtration $\mathcal{F}_{t_k}$.
For $k = 0$, this is obviously true for the filtration $\mathcal{F}_{t_0}$ is trivial and $\bar{\Omega}^{\alpha_\epsilon}_{t_0}(\beta) = \Omega$.
Assume that the property of the Lemma is true at stage k, $k < N\bar{N} - 1$ and that $\mathcal{F}_{t_k}$ is generated by the atoms $\{\Omega^k_i\}_{i \in I}$. Assume now that $\mathcal{F}_{t_{k+1}} = \sigma\left(\{\Omega^{k+1}_j\}_{j \in J}\right)$ where the Ω^{k+1}_j are the atoms of $\mathcal{F}_{t_{k+1}}$.

- If for some $i \in I$, there exists $j \in J$ such that $\Omega^k_i = \Omega^{k+1}_j$, this means that no explosion takes place on $[t_k, t_{k+1})$ for the controls $(u_\eta, v_\eta)(\omega)$, $\omega \in \Omega^k_i$. Assume that $\bar{\Omega}^{\alpha_\epsilon}_{t_k}(\beta) = \Omega^k_i$. As $\mathbf{S}^{\alpha_\epsilon}_{t_k}(\beta) = \emptyset$, the strategy α_ϵ will generate on $[t_k, t_{k+1})$ the control $u_\eta(\omega)$ for any $\omega \in \Omega^k_i$ and we will get $\bar{\Omega}^{\alpha_\epsilon}_{t_{k+1}}(\beta) = \bar{\Omega}^{\alpha_\epsilon}_{t_k}(\beta) = \Omega^k_i$. This implies $\mathbf{P}_{\alpha_\epsilon}\left[\bar{\Omega}^{\alpha_\epsilon}_{t_{k+1}}(\beta) = \Omega^k_i\right] \geq \mathbf{P}_{\alpha_\epsilon}\left[\bar{\Omega}^{\alpha_\epsilon}_{t_k}(\beta) = \Omega^k_i\right]$. On the other hand, the definition of the process $\bar{\Omega}^{\alpha_\epsilon}(\beta)$ ensures that $\bar{\Omega}^{\alpha_\epsilon}_{t_{k+1}}(\beta) \subseteq \bar{\Omega}^{\alpha_\epsilon}_{t_k}(\beta)$ and $\bar{\Omega}^{\alpha_\epsilon}_{t_k}(\beta) \in \{\Omega^k_i\}_{i \in I} \cup \emptyset$ leading to

$$\mathbf{P}_{\alpha_\epsilon}\left[\bar{\Omega}^{\alpha_\epsilon}_{t_{k+1}}(\beta) = \Omega^k_i\right] = \mathbf{P}_{\alpha_\epsilon}\left[\bar{\Omega}^{\alpha_\epsilon}_{t_k}(\beta) = \Omega^k_i\right] = \mathbf{Q}(\Omega^k_i)$$

- Assume now that for some $i \in I$, $\Omega^k_i \neq \Omega^{k+1}_j$ for all $j \in J$. This means there is an explosion on $[t_k, t_{k+1})$ for the controls $(u_\eta, v_\eta)(\omega)$, $\omega \in \Omega^k_i$ and $\Omega^k_i = \bigsqcup_{j=j_0}^{j_i} \Omega^{k+1}_j$. Assume that we have for some ω_{α_ϵ} $\bar{\Omega}^{\alpha_\epsilon}_{t_k}(\omega_{\alpha_\epsilon}, \beta) = \Omega^k_i$. Recall that $\mathcal{F}^{\alpha_\epsilon, v}_{t_k} = \sigma((\alpha_\epsilon(v), v)(s),\ s \in [t_0, t_k])$. Note that $\mathbf{S}^{\alpha_\epsilon}_{t_k}(\omega_{\alpha_\epsilon}, \beta) = \emptyset$, implying on $[t_k, t_{k+1})$ the strategy α_ϵ will generate one of the correlation control $u_a \in \Omega_{\alpha_\epsilon}$ prescribed by the explosion procedure for Ω^k_i. The conditional probability that the control generated by α_ϵ at time t_k is u_a given all correlation controls played so far is

$$\mathbf{P}_{\alpha_\epsilon}\left[\alpha_\epsilon(\beta)|_{[t_k, t_{k+1})} = u_a \middle| \mathcal{F}^{\alpha_\epsilon, \beta}_{t_k}\right] = \frac{1}{q(t^i_k)} \times \mathbf{P}_{\alpha_\epsilon}\left[\bar{\Omega}^{\alpha_\epsilon}_{t_k}(\beta) = \Omega^k_i \middle| \mathcal{F}^{\alpha_\epsilon, \beta}_{t_k}\right]$$

due to (11) because every correlation control being unique, the only way to play u_a is when $\bar{\Omega}^{\alpha_\epsilon}_{t_k}(\beta) = \Omega^k_i$. Given the controls played on $[t_0, t_k)$ for any trajectory such that $\bar{\Omega}^{\alpha_\epsilon}_{t_k}(\beta) = \Omega^k_i$, at time t_k, the pure strategy β being a strategy with delay will generate on $[t_k, t_k+\tau(\beta)]$ the same control for example v_b whatever the control u_a chosen by player I on $[t_k, t_{k+1})$. Note that we must have $v|_{[t_k, t_k+\tau/2)}$ is equivalent to one of the constant correlation controls, else, player I would detect some deviation at time t_{k+1} and set $\mathbf{S}^{\alpha_\epsilon}_{t_{k+2}}(\beta) \neq \emptyset$. In the end, player II has to play on $[t_k, t_k + \tau/2)$ one of the correlation control v_b, and always plays the same control whatever the control u_a played by player I. Finally, we will get for all $j = j_0 \dots j_i$:

$$\mathbf{P}_{\alpha_\epsilon}\left[\bar{\Omega}^{\alpha_\epsilon}_{t_{k+1}}(\beta) = \Omega^{k+1}_j \middle| \mathcal{F}^{\alpha_\epsilon,\beta}_{t_k}\right] = \mathbf{Q}(\Omega^{k+1}_j|\Omega^k_i) \times \mathbf{P}_{\alpha_\epsilon}\left[\bar{\Omega}^{\alpha_\epsilon}_{t_k}(\beta) = \Omega^k_i \middle| \mathcal{F}^{\alpha_\epsilon,\beta}_{t_k}\right]$$

and finally taking the expectation w.r.t. $\mathbf{P}_{\alpha_\epsilon}$:

$$\begin{aligned}\mathbf{P}_{\alpha_\epsilon}\left[\bar{\Omega}^{\alpha_\epsilon}_{t_{k+1}}(\beta) = \Omega^{k+1}_j\right] &= \mathbf{Q}(\Omega^{k+1}_j|\Omega^k_i)\mathbf{P}_{\alpha_\epsilon}\left[\bar{\Omega}^{\alpha_\epsilon}_{t_k}(\beta) = \Omega^k_i\right] \\ &= \mathbf{Q}(\Omega^{k+1}_j|\Omega^k_i)\mathbf{Q}(\Omega^k_i) = \mathbf{Q}(\Omega^{k+1}_j)\end{aligned}$$

We have proven that for all $k \in \{0, \dots, N\bar{N} - 1\}$, for all atom Ω^k_i of the filtration $\mathcal{F}_{t_k}$, $\mathbf{P}_{\alpha_\epsilon}\left[\bar{\Omega}^{\alpha_\epsilon}_{t_k}(\beta) = \Omega^k_i\right] = \mathbf{Q}(\Omega^k_i)$. Noticing that there is no explosion on $[T - \tau, T]$, we get $\mathcal{F}_T = \mathcal{F}_{t_{N\bar{N}-1}}$ and due to the definition of the strategy and the fact that $\mathbf{S}^{\alpha_\epsilon}_T(\beta) = \emptyset$, we get $\bar{\Omega}^{\alpha_\epsilon}_T(\beta) = \bar{\Omega}^{\alpha_\epsilon}_{t_{N\bar{N}-1}}(\beta)$, hence the result. □

We still assume that player I plays α_ϵ and player II plays some pure strategy β such that β generates no deviation and we will compute the payoff $\mathfrak{J}_i(t_0, x_0, \alpha_\epsilon, \beta)$ for $i = 1, 2$.

Lemma 2.6. *If the strategies (α_ϵ, β) are played where β is some pure strategy with delay such that $\mathbf{S}^{\alpha_\epsilon}_T(\beta) = \emptyset$, then for all $i = 1, 2$:*

$$|\mathfrak{J}_i(t_0, x_0, \alpha_\epsilon, \beta) - e_i| \leq \frac{3\epsilon}{N}$$

Corollary:
The strategies $(\alpha_\epsilon, \beta_\epsilon)$ reward a payoff $\frac{3\epsilon}{N}$ close to e.

Proof of the Corollary. The proof of the Corollary is straightforward. Indeed, as β_ϵ is a mixed strategy, namely a finite probability distribution on finitely many pure strategies $\beta_\epsilon(\omega_{\beta_\epsilon})$ generating $\mathbf{S}^{\alpha_\epsilon}_T(\beta_\epsilon(\omega_{\beta_\epsilon})) = \emptyset$ against α_ϵ, we get for $i = 1, 2$:

$$\begin{aligned}|\mathfrak{J}_i(t_0, x_0, \alpha_\epsilon, \beta_\epsilon) - e_i| &\leq \int_{\Omega_{\beta_\epsilon}} |\mathfrak{J}_i(t_0, x_0, \alpha_\epsilon, \beta_\epsilon(\omega_{\beta_\epsilon})) - e_i| \mathrm{d}\mathbf{P}_{\beta_\epsilon}(\omega_{\beta_\epsilon}) \\ &\leq \int_{\Omega_{\beta_\epsilon}} \mathrm{d}\mathbf{P}_{\beta_\epsilon}(\omega_{\beta_\epsilon}) \frac{3\epsilon}{N} = \frac{3\epsilon}{N} \quad \square\end{aligned}$$

Proof of Lemma 2.6. We recall that the explosions are denoted by τ_i, $i \in I$. For all $i \in I$, there exists $k(\tau_i) \in \{0, \dots, N\bar{N} - 2\}$ such that $\tau_i \in [t_{k(\tau_i)}, t_{k(\tau_i)+1})$. We denote by

$$\Delta := [t_0, T - \tau) \setminus \left(\cup_{i \in I} [t_{k(\tau_i)}, t_{k(\tau_i)+1})\right)$$

Assume that $\mathcal{F}_T = \sigma(\{\Omega_j\}_{j=1\dots\bar{M}})$ where the Ω_j are the atoms of $\mathcal{F}_T$ and players are using (α_ϵ, β) as in the assumptions of the Lemma. Notice that $\forall \omega_{\alpha_\epsilon} \in \{\bar{\Omega}_T^{\alpha_\epsilon}(\beta) = \Omega_j\}$, $\forall \omega_j \in \Omega_j$, the control of Player I generated by $(\alpha_\epsilon(\omega_{\alpha_\epsilon}), \beta)$ satisfies $u_{\alpha_\epsilon(\omega_{\alpha_\epsilon})\beta}(s) \equiv u_\eta(\omega_j)(s)$ $\forall s \in \Delta$. Consequently, $\forall \omega_{\alpha_\epsilon} \in \{\bar{\Omega}_T^{\alpha_\epsilon}(\beta) = \Omega_j\}$, the control of player II generated by $(\alpha_\epsilon(\omega_{\alpha_\epsilon}), \beta)$ satisfies $v_{\alpha_\epsilon(\omega_{\alpha_\epsilon})\beta}(s) \equiv v_\eta(\omega_j)(s)$ $\forall s \in \Delta$, else we would get $\mathbf{S}_T^{\alpha_\epsilon}(\omega_{\alpha_\epsilon}, \beta) \neq \emptyset$. Therefore, for any ω_{α_ϵ} satisfying $\bar{\Omega}_T^{\alpha_\epsilon}(\omega_{\alpha_\epsilon}, \beta) = \Omega_j$ and any $\omega_j \in \Omega_j$, for all $t \in [t_0, T]$:

$$\left\| X_t^{t_0, x_0, (\alpha_\epsilon, \beta)(\omega_{\alpha_\epsilon})} - X_t^\eta(\omega_j) \right\| \leq M\tau(1 + \|f\|_\infty) e^{L_f(T - t_0)}$$

where L_f denotes the Lipschitz constant of f and M the number of explosions. We can choose τ small enough in order that for $i = 1, 2$, for all $t \in [t_0, T]$:

$$\begin{cases} \left| g_i(X_T^{t_0, x_0, (\alpha_\epsilon, \beta)(\omega_{\alpha_\epsilon})}) - g_i\left(X_T^\eta(\omega_j)\right) \right| \leq \eta \\ \left| V_i(t, X_t^{t_0, x_0, (\alpha_\epsilon, \beta)(\omega_{\alpha_\epsilon})}) - V_i(t, X_t^\eta(\omega_j)) \right| \leq \eta \end{cases} \tag{12}$$

leading, for any $j = 1 \dots \bar{M}$ and any $\omega_j \in \Omega_j$, to

$$\begin{aligned} &\left| \mathbf{E}_{\alpha_\epsilon}\left[g_i(X_T^{t_0, x_0, \alpha_\epsilon, \beta}) \mathbf{1}_{\bar{\Omega}_T^{\alpha_\epsilon}(\beta) = \Omega_j} \right] - \mathbf{E}\left[g_i(X_T^\eta) \mathbf{1}_{\Omega_j} \right] \right| \\ &\leq \left| g_i(X_T^\eta(\omega_j)) \mathbf{P}_{\alpha_\epsilon}(\bar{\Omega}_T^{\alpha_\epsilon}(\beta) = \Omega_j) - g_i(X_T^\eta(\omega_j)) \mathbf{Q}(\Omega_j) \right| \\ &\qquad + \eta \mathbf{P}_{\alpha_\epsilon}(\bar{\Omega}_T^{\alpha_\epsilon}(\beta) = \Omega_j) \\ &\leq \eta \mathbf{Q}(\Omega_j) \text{ due to Lemma 2.5} \end{aligned}$$

Finally, for all $i = 1, 2$:

$$\begin{aligned} &|\mathfrak{J}_i(t_0, x_0, \alpha_\epsilon, \beta) - \mathfrak{J}_i(t_0, x_0, u_\eta, v_\eta)| \\ &\quad \leq \sum_{j=1}^{\bar{M}} \left| \mathbf{E}_{\alpha_\epsilon}\left[g_i(X_T^{t_0, x_0, \alpha_\epsilon, \beta}) \mathbf{1}_{\bar{\Omega}_T^{\alpha_\epsilon}(\beta) = \Omega_j} \right] - \mathbf{E}\left[g_i(X_T^\eta) \mathbf{1}_{\Omega_j} \right] \right| \\ &\quad \leq \sum_{j=1}^{\bar{M}} \eta \mathbf{Q}(\Omega_j) = \eta \end{aligned}$$

Using now Lemma 2.3, we have for all $i = 1, 2$:

$$\begin{aligned} |\mathfrak{J}_i(t_0, x_0, \alpha_\epsilon, \beta) - e_i| &\leq |\mathfrak{J}_i(t_0, x_0, \alpha_\epsilon, \beta) - \mathfrak{J}_i(t_0, x_0, u_\eta, v_\eta)| + |\mathfrak{J}_i(t_0, x_0, u_\eta, v_\eta) - e_i| \\ &\leq 3\eta \end{aligned}$$

and the strategies (α_ϵ, β) reward a payoff $\frac{3\epsilon}{N}$ close to e. □

Optimality of the strategies $(\alpha_\epsilon, \beta_\epsilon)$:
It remains to prove that the strategies $(\alpha_\epsilon, \beta_\epsilon)$ are optimal. We will prove it for β_ϵ: there exists some constant C_α satisfying

$$\forall \beta \in \mathcal{B}(t_0),\ \mathfrak{J}_2(t_0, x_0, \alpha_\epsilon, \beta) \leq \mathfrak{J}_2(t_0, x_0, \alpha_\epsilon, \beta_\epsilon) + C_\alpha \epsilon. \tag{13}$$

Consider some pure strategy with delay β. If β generates no deviation ($\mathbf{S}_T^{\alpha_\epsilon}(\beta) = \emptyset$), then we have just proven that:

$$\mathfrak{J}_2(t_0, x_0, \alpha_\epsilon, \beta) \leq e_2 + \frac{3\epsilon}{N} \leq \mathfrak{J}_2(t_0, x_0, \alpha_\epsilon, \beta_\epsilon) + \frac{6\epsilon}{N}. \tag{14}$$

It remains to prove the same kind of result as (13) for any pure strategy β generating some unexpected controls (leading for some ω_{α_ϵ} to $\mathbf{S}_T^{\alpha_\epsilon}(\omega_{\alpha_\epsilon}, \beta) \neq \emptyset$). The idea of the proof is first to build some pure strategy $\tilde{\beta}$ generating the same controls as β against α_ϵ as long as no deviation occurs and generating no deviation against α_ϵ, that is some non-deviating extension of β. We then will compare the payoffs induced by β and $\tilde{\beta}$.

Lemma 2.7 (Non deviating extension $\tilde{\beta}$ of some pure strategy β of player II)**.** *To any pure strategy with delay β, one can associate a pure strategy with delay $\tilde{\beta}$ satisfying:*

- $\mathbf{S}^{\alpha_\epsilon}(\tilde{\beta}) = \emptyset$
- *The pairs of strategies $(\alpha_\epsilon(\omega_{\alpha_\epsilon}), \beta)$ and $(\alpha_\epsilon(\omega_{\alpha_\epsilon}), \tilde{\beta})$ generate the same pairs of controls on $[t_0, T-\tau] \times \{\mathbf{S}_T^{\alpha_\epsilon}(\beta) = \emptyset\} \cup_{k \in \{0,\dots,N\bar{N}\}} [t_0, t_k] \times \{\mathbf{S}_T^{\alpha_\epsilon}(\beta) \in \{t_k\} \times \mathbb{R}^n\}$.*

Proof. In order to build the strategy $\tilde{\beta}$, we will need some auxiliary processes $\mathbf{S}^{\tilde{\beta}} : [t_0, T] \times \mathcal{U}(t_0) \to \emptyset \cup [t_0, T] \cup \{\not\emptyset\}$ and $\bar{\Omega}^{\tilde{\beta}} : [t_0, T] \times \mathcal{U}(t_0) \to \mathcal{F}_T$. We will set $\mathbf{S}^{\tilde{\beta}} = \emptyset$ as long as β generates no deviation against α_ϵ and $\mathbf{S}^{\tilde{\beta}} = t_i$ if β generated some deviation on $[t_i, t_{i+1})$ and $\mathbf{S}^{\tilde{\beta}} = \not\emptyset$ if it is impossible to correlate the trajectories because player I is not playing α_ϵ. The process $\bar{\Omega}^{\tilde{\beta}}$ allows to remind which control $v_\eta(\omega)$ the strategy is currently following. We define the strategy $\tilde{\beta}(u)$ for all $u \in \mathcal{U}(t_0)$ using $\mathbf{S}^{\tilde{\beta}}$ and $\bar{\Omega}^{\tilde{\beta}}$ as follows (omitting to write the dependence of $\mathbf{S}^{\tilde{\beta}}$ and $\bar{\Omega}^{\tilde{\beta}}$ in u):

Set $\mathbf{S}_{t_0}^{\tilde{\beta}} = \emptyset$ and $\bar{\Omega}_{t_0}^{\tilde{\beta}} = \Omega$ and fix $v_0 \in V$ and $u_0 \in U$. For all $k \in \{0 \dots N\bar{N} - 1\}$, if $\tilde{\beta}(u)$ is already built on $[t_0, t_k)$, we define $\tilde{\beta}(u)$ on $[t_k, t_{k+1})$ by:

1. If $\mathbf{S}_{t_k}^{\tilde{\beta}} = \not\emptyset$, play $\tilde{\beta}(u)|_{[t_k, t_{k+1})} = v_0$, set $\mathbf{S}_{t_{k+1}}^{\tilde{\beta}} = \not\emptyset$ and $\bar{\Omega}_{t_{k+1}}^{\tilde{\beta}} = \emptyset$.
2. If $\mathbf{S}_{t_k}^{\tilde{\beta}} \in [t_0, T]$, meaning we are building a non deviating strategy instead of following β then
   - if there is no explosion on $[t_k, t_{k+1})$, play $\tilde{\beta}(u)|_{[t_k, t_{k+1})} = v_\eta(\omega)|_{[t_k, t_{k+1})}$ for some $\omega \in \bar{\Omega}_{t_k}^{\tilde{\beta}}$ and set $\bar{\Omega}_{t_{k+1}}^{\tilde{\beta}} = \bar{\Omega}_{t_k}^{\tilde{\beta}}$ and $\mathbf{S}_{t_{k+1}}^{\tilde{\beta}} = \mathbf{S}_{t_k}^{\tilde{\beta}}$
   - if there is an explosion on $[t_k, t_{k+1})$ for the trajectory generated by $(u_\eta, v_\eta)(\omega)$ for any $\omega \in \bar{\Omega}_{t_k}^{\tilde{\beta}} = \Omega_l$, then play $\tilde{\beta}(u)|_{[t_k, t_{k+1})} = v_1$ where v_1 is the first correlation control of player II expected at time t_k and consider the control $u|_{[t_k, t_k + \frac{\tau}{2}]}$. If it is one of the expected correlation controls for example u_a then consider $G(a, 1) = i$ and set $\bar{\Omega}_{t_{k+1}}^{\tilde{\beta}} = \Omega_i^l$ and $\mathbf{S}_{t_{k+1}}^{\tilde{\beta}} = \mathbf{S}_{t_k}^{\tilde{\beta}}$, else set $\mathbf{S}_{t_{k+1}}^{\tilde{\beta}} = \not\emptyset$ and $\bar{\Omega}_{t_{k+1}}^{\tilde{\beta}} = \emptyset$,
3. If $\mathbf{S}_{t_k}^{\tilde{\beta}} = \emptyset$ then
   - if there is no explosion on $[t_k, t_{k+1})$, play $\tilde{\beta}(u)|_{[t_k, t_{k+1})} = v_\eta(\omega)|_{[t_k, t_{k+1})}$ for some $\omega \in \bar{\Omega}_{t_k}^{\tilde{\beta}}$ and set $\bar{\Omega}_{t_{k+1}}^{\tilde{\beta}} = \bar{\Omega}_{t_k}^{\tilde{\beta}}$. Then, build the control $\tilde{u}$ such that:

$$\begin{cases} \tilde{u}(s) = u(s) & \forall s \in [t_0, t_k) \\ \tilde{u}(s) = u_\eta(\omega)(s) & \forall s \in [t_k, T] \end{cases}$$

and compute the control $\beta(\tilde{u})|_{[t_k,t_{k+1})}$: if $\beta(\tilde{u})|_{[t_k,t_{k+1})} \not\equiv v_\eta(\omega)|_{[t_k,t_{k+1})}$ then set $\mathbf{S}^{\tilde{\beta}}_{t_{k+1}} = t_k$ else set $\mathbf{S}^{\tilde{\beta}}_{t_{k+1}} = \emptyset$

- if there is an explosion on $[t_k, t_{k+1})$ for the trajectory generated by $(u_\eta, v_\eta)(\omega)$ for any $\omega \in \bar{\Omega}^{\tilde{\beta}}_{t_k} = \Omega_l$, then build the control $\tilde{u}$ such that:

$$\begin{cases} \tilde{u}(s) = u(s) & \forall s \in [t_0, t_k) \\ \tilde{u}(s) = u_0 & \forall s \in [t_k, T] \end{cases}$$

  and compute the control $v = \beta(\tilde{u})|_{[t_k,t_k+\tau(\beta))}$:
  - If this control is none of the expected correlation controls v_b, play $\tilde{\beta}(u)|_{[t_k,t_{k+1})} = v_1$ where v_1 is the first expected correlation control and consider the control $u|_{[t_k,t_k+\frac{\tau}{2})}$. If it is one of the expected correlation controls for example u_a then consider $G(a,1) = i$ and set $\bar{\Omega}^{\tilde{\beta}}_{t_{k+1}} = \Omega^l_i$ and $\mathbf{S}^{\tilde{\beta}}_{t_{k+1}} = t_k$, else set $\mathbf{S}^{\tilde{\beta}}_{t_{k+1}} = \not\alpha$ and $\bar{\Omega}^{\tilde{\beta}}_{t_{k+1}} = \emptyset$,
  - If this control v corresponds to one of the correlation controls prescribed by the jointly controlled lottery at t_k for example v_b then play $\tilde{\beta}(u)|_{[t_k,t_k+\frac{\tau}{2}]} = v_b$ and $\tilde{\beta}(u)|_{(t_k+\frac{\tau}{2},t_{k+1})} = \beta(u)|_{(t_k+\frac{\tau}{2},t_{k+1})}$. Consider the control $u|_{[t_k,t_k+\frac{\tau}{2})}$. If it is none of the expected correlation controls, set $\mathbf{S}^{\tilde{\beta}}_{t_{k+1}} = \not\alpha$ and $\bar{\Omega}^{\tilde{\beta}}_{t_{k+1}} = \emptyset$. If it is one of the correlation controls, for example u_a, consider $G(a,b) = i$ and set $\bar{\Omega}^{\tilde{\beta}}_{t_{k+1}} = \Omega^l_i$. Then compute the control $\beta(u)|_{[t_k,t_k+\frac{\tau}{2}]}$. If $\beta(u)|_{[t_k,t_k+\frac{\tau}{2}]} \not\equiv v_b$ then set $\mathbf{S}^{\tilde{\beta}}_{t_{k+1}} = t_k$ else set $\mathbf{S}^{\tilde{\beta}}_{t_{k+1}} = \emptyset$.

It is clear that $\tilde{\beta}$ is a pure strategy with delay. Indeed, $\tilde{\beta}$ is anticipative with respect to β but non anticipative with respect to the control u of the opponent. Furthermore $\tilde{\beta}$ satisfies $\mathbf{S}^{\alpha_\epsilon}_T(\tilde{\beta}) = \emptyset$ and $\bar{\Omega}^{\alpha_\epsilon}_T(\tilde{\beta}) \neq \emptyset$. As long as β generates no deviation, the controls generated by (α_ϵ, β) and $(\alpha_\epsilon, \tilde{\beta})$ are the same. Note that $\{\mathbf{S}^{\tilde{\beta}}_T = t_k\} \in \mathcal{F}^{\alpha_\epsilon\tilde{\beta}}_{t_k}$ where $\mathcal{F}^{\alpha_\epsilon\tilde{\beta}}_{t_k} = \sigma\left((\alpha_\epsilon, \tilde{\beta})(s),\ s \in [t_0, t_k]\right)$. □

We have for any deviating pure strategy β:

$$\begin{aligned} \mathfrak{J}_2(t_0, x_0, \alpha_\epsilon, \beta) = \sum_{i=0}^{N\bar{N}-1} & \mathbf{E}_{\alpha_\epsilon}\left(g_2(X^{t_0,x_0,\alpha_\epsilon,\beta}_T)\mathbf{1}_{\mathbf{S}^{\alpha_\epsilon}_T(\beta)\in\{t_i\}\times\mathbb{R}^n}\right) \\ & + \mathbf{E}_{\alpha_\epsilon}(g_2(X^{t_0,x_0,\alpha_\epsilon,\beta}_T)\mathbf{1}_{\mathbf{S}^{\alpha_\epsilon}_T(\beta)=\emptyset}) \end{aligned} \tag{15}$$

Assume that for example $\mathbf{S}^{\alpha_\epsilon}_T(\beta) = (t_i, x)$. This means that some deviation occurred on $[t_i, t_{i+1})$. There exists $k \in \{1 \dots N\}$ such that $[t_i, t_{i+1}) \subset [\theta_{k-1}, \theta_k)$. Using the

definition of the strategy α_ϵ and introducing the non deviating extension $\tilde{\beta}$ of β:

$$\begin{aligned}
g_2(X_T^{t_0,x_0,\alpha_\epsilon,\beta})\mathbf{1}_{\mathbf{S}_T^{\alpha_\epsilon}(\beta)=(t_i,x)} &= g_2(X_T^{t_{i+2},x,\alpha_p^{\eta,t_{i+2}}(x),\beta})\mathbf{1}_{\mathbf{S}_T^{\alpha_\epsilon}(\beta)=(t_i,x)} \\
&\leq (V_2(t_{i+2},x)+\eta)\mathbf{1}_{\mathbf{S}_T^{\alpha_\epsilon}(\beta)=(t_i,x)} \\
&\leq \left[V_2(t_i, X_{t_i}^{t_0,x_0,\alpha_\epsilon,\beta})+\eta+\epsilon\right]\mathbf{1}_{\mathbf{S}_T^{\alpha_\epsilon}(\beta)=(t_i,x)} \text{ due to (1)} \\
&\leq \left[V_2(t_i, X_{t_i}^{t_0,x_0,\alpha_\epsilon,\tilde{\beta}})+2\epsilon\right]\mathbf{1}_{\mathbf{S}_T^{\alpha_\epsilon}(\beta)=(t_i,x)} \\
&\leq \left[\mathbf{E}_{\alpha_\epsilon}(V_2(\theta_k, X_{\theta_k}^{t_0,x_0,\alpha_\epsilon,\tilde{\beta}})|\mathcal{F}_{t_i}^{\alpha_\epsilon\tilde{\beta}})+3\epsilon\right]\mathbf{1}_{\mathbf{S}_T^{\alpha_\epsilon}(\beta)=(t_i,x)} \\
&\quad \text{due to (1) because } (\theta_k-t_i)\leq\delta
\end{aligned}$$

We introduce this last inequality because our estimate of $V_2(t_i, X_{t_i}^{t_0,x_0,\alpha_\epsilon,\tilde{\beta}})$ induces some error term of length η, therefore we need to sum up at most N such error terms in order to bound the global error to some ϵ.
In the end we have for all $t_i \in [\theta_{k-1},\theta_k)$:

$$g_2(X_T^{t_0,x_0,\alpha_\epsilon,\beta})\mathbf{1}_{\mathbf{S}_T^{\alpha_\epsilon}(\beta)\in\{t_i\}\times\mathbb{R}^n} \leq \left[\mathbf{E}_{\alpha_\epsilon}\left(V_2(\theta_k, X_{\theta_k}^{t_0,x_0,\alpha_\epsilon,\tilde{\beta}})\middle|\mathcal{F}_{t_i}^{\alpha_\epsilon\tilde{\beta}}\right)+3\epsilon\right]\mathbf{1}_{\mathbf{S}_T^{\alpha_\epsilon}(\beta)\in\{t_i\}\times\mathbb{R}^n} \tag{16}$$

The point now is to get an estimate of $V_2(\theta_k, X_{\theta_k}^{t_0,x_0,\alpha_\epsilon,\tilde{\beta}})$. We will prove the following Lemma:

Lemma 2.8. *For all $t \in \{t_k\}_{k=0\ldots N\bar{N}}$, for all pure strategy $\tilde{\beta}$ generating no deviation against α_ϵ, we have:*

$$\mathbf{P}_{\alpha_\epsilon}\left\{V_2(t, X_t^{t_0,x_0,\alpha_\epsilon,\tilde{\beta}}) \leq \mathbf{E}_{\alpha_\epsilon}\left(g_2(X_T^{t_0,x_0,\alpha_\epsilon,\tilde{\beta}})\middle|\mathcal{F}_t^{\alpha_\epsilon\tilde{\beta}}\right)+4\eta\right\} \geq 1-2\eta$$

where $(\mathcal{F}_t^{\alpha_\epsilon\tilde{\beta}}) = \sigma\left((\alpha_\epsilon,\tilde{\beta})(s), s\in[t_0,t]\right)$

Proof. Assume that $t \in \{t_k\}_{k=0\ldots N\bar{N}}$ and $\mathcal{F}_t = \sigma(\{\Omega_i\}_{i\in I})$ where the Ω_i are the atoms of $\mathcal{F}_t$. As the pure strategy $\tilde{\beta}$ generates no deviation, denoting by $\bar{\Omega}^{\alpha_\epsilon}(\tilde{\beta})$ the auxiliary process built in the definition of the strategy α_ϵ as the opponent is playing $\tilde{\beta}$, we have for all $i \in I$, there exist atoms Ω_j^i of the filtration $\mathcal{F}_T$ such that $\Omega_i = \bigsqcup_{j\in J(i)}\Omega_j^i$. Then using (12) for any $\omega_j \in \Omega_j^i$, for any $j \in J(i)$:

$$V_2(t, X_t^{t_0,x_0,\alpha_\epsilon,\tilde{\beta}})\mathbf{1}_{\bar{\Omega}_T^{\alpha_\epsilon}(\tilde{\beta})=\Omega_j^i} \leq \left[V_2(t, X_t^\eta(\omega_j^i))+\eta\right]\mathbf{1}_{\bar{\Omega}_T^{\alpha_\epsilon}(\tilde{\beta})=\Omega_j^i}$$

Using the fact that $\{\bar{\Omega}_t^{\alpha_\epsilon}(\tilde{\beta}) = \Omega_i\} = \cup_{j\in J(i)}\{\bar{\Omega}_T^{\alpha_\epsilon}(\tilde{\beta}) = \Omega_j^i\}$, and that $\forall\omega_j \in \Omega_j^i$, $\forall\omega_i \in \Omega_i$, $V_2(t, X_t^\eta(\omega_j)) = V_2(t, X_t^\eta(\omega_i))$, we get for any $i \in I$ and any $\omega_i \in \Omega_i$

$$V_2(t, X_t^{t_0,x_0,\alpha_\epsilon,\tilde{\beta}})\mathbf{1}_{\bar{\Omega}_T^{\alpha_\epsilon}(\tilde{\beta})=\Omega_i} \leq \left[V_2(t, X_t^\eta(\omega_i))+\eta\right]\mathbf{1}_{\bar{\Omega}_T^{\alpha_\epsilon}(\tilde{\beta})=\Omega_i} \tag{17}$$

Note that a more careful examination of the proof of Lemma 2.5 shows that:

$$\mathbf{P}_{\alpha_\epsilon}\left(\bar{\Omega}_T^{\alpha_\epsilon}(\tilde{\beta}) = \Omega_j^i|\mathcal{F}_t^{\alpha_\epsilon\tilde{\beta}}\right) = \mathbf{Q}(\Omega_j^i|\Omega_i)\mathbf{P}_{\alpha_\epsilon}\left(\bar{\Omega}_t^{\alpha_\epsilon}(\tilde{\beta}) = \Omega_i|\mathcal{F}_t^{\alpha_\epsilon\tilde{\beta}}\right) = \mathbf{Q}(\Omega_j^i|\Omega_i)\mathbf{1}_{\bar{\Omega}_t^{\alpha_\epsilon}(\tilde{\beta})=\Omega_i}$$

Then, using (12):

$$\left|\mathbf{E}_{\alpha_\epsilon}\left(g_2(X_T^{t_0,x_0,\alpha_\epsilon,\tilde{\beta}})\middle|\mathcal{F}_t^{\alpha_\epsilon\tilde{\beta}}\right)\mathbf{1}_{\bar{\Omega}_t^{\alpha_\epsilon}(\tilde{\beta})=\Omega_i} - \mathbf{E}\left(g_2(X_T^\eta)|\Omega_i\right)\mathbf{1}_{\bar{\Omega}_t^{\alpha_\epsilon}(\tilde{\beta})=\Omega_i}\right| \leq \eta \tag{18}$$

We now have to recall that if we denote by

$$\Sigma_t^\eta = \left\{V_2(t, X_t^\eta) \leq \mathbf{E}\left(g_2(X_T^\eta)\big|\mathcal{F}_t\right) + 2\eta\right\}$$

then due to Lemma 2.3:

$$\mathbf{Q}(\Sigma_t^\eta) \geq 1 - 2\eta$$

This implies that for all $i \in I$, for all trajectory satisfying $\bar{\Omega}_t^{\alpha_\epsilon}(\tilde{\beta}) = \Omega_i \subset \Sigma_t^\eta$ and for all $\omega_i \in \Omega_i$:

$$\begin{aligned}
V_2(t, X_t^{t_0,x_0,\alpha_\epsilon,\tilde{\beta}})\mathbf{1}_{\bar{\Omega}_t^{\alpha_\epsilon}(\tilde{\beta})=\Omega_i} &\leq \left[V_2(t, X_t^\eta(\omega_i)) + \eta\right]\mathbf{1}_{\bar{\Omega}_t^{\alpha_\epsilon}(\tilde{\beta})=\Omega_i} \text{ (due to (17))} \\
&\leq \left[\mathbf{E}(g_2(X_T^\eta)|\Omega_i) + 3\eta\right]\mathbf{1}_{\bar{\Omega}_t^{\alpha_\epsilon}(\tilde{\beta})=\Omega_i} \text{ (for } \Omega_i \subset \Sigma_t^\eta) \\
&\leq \left[\mathbf{E}_{\alpha_\epsilon}\left(g_2(X_T^{t_0,x_0,\alpha_\epsilon,\tilde{\beta}})|\mathcal{F}_t^{\alpha_\epsilon\tilde{\beta}}\right) + 4\eta\right]\mathbf{1}_{\bar{\Omega}_t^{\alpha_\epsilon}(\tilde{\beta})=\Omega_i} \text{ (due to (18))}
\end{aligned}$$

We now can check that as announced:

$$\begin{aligned}
&\mathbf{P}_{\alpha_\epsilon}\left\{V_2(t, X_t^{t_0,x_0,\alpha_\epsilon,\tilde{\beta}}) \leq \mathbf{E}_{\alpha_\epsilon}\left(g_2(X_T^{t_0,x_0,\alpha_\epsilon,\tilde{\beta}})\big|\mathcal{F}_t^{\alpha_\epsilon\tilde{\beta}}\right) + 4\eta\right\} \\
&\geq \mathbf{P}_{\alpha_\epsilon}\left\{\bar{\Omega}_t^{\alpha_\epsilon}(\tilde{\beta}) \subset \Sigma_t^\eta\right\} \\
&\geq \mathbf{Q}(\Sigma_t^\eta) \text{ (due to Lemma 2.5)} \\
&\geq 1 - 2\eta
\end{aligned}$$

□

We will denote by:

$$\Sigma_t^{\alpha_\epsilon\tilde{\beta}} = \left\{V_2(t, X_t^{t_0,x_0,\alpha_\epsilon,\tilde{\beta}}) \leq \mathbf{E}_{\alpha_\epsilon}\left(g_2(X_T^{t_0,x_0,\alpha_\epsilon,\tilde{\beta}})\big|\mathcal{F}_t^{\alpha_\epsilon\tilde{\beta}}\right) + 4\eta\right\}$$

We now will compute a more precise estimate of $V_2(\theta_k, X_{\theta_k}^{t_0,x_0,\alpha_\epsilon,\tilde{\beta}})$ denoting by $\|g\|_\infty$ some bound of the payoff functions g_1 and g_2:

$$\begin{aligned}
V_2(\theta_k, X_{\theta_k}^{t_0,x_0,\alpha_\epsilon,\tilde{\beta}}) &\leq \left[\mathbf{E}_{\alpha_\epsilon}\left(g_2(X_T^{t_0,x_0,\alpha_\epsilon,\tilde{\beta}})\big|\mathcal{F}_{\theta_k}^{\alpha_\epsilon\tilde{\beta}}\right) + 4\eta\right]\mathbf{1}_{\Sigma_{\theta_k}^{\alpha_\epsilon\tilde{\beta}}} + \|g\|_\infty\mathbf{1}_{(\Sigma_{\theta_k}^{\alpha_\epsilon\tilde{\beta}})^c} \\
&\leq \mathbf{E}_{\alpha_\epsilon}\left(g_2(X_T^{t_0,x_0,\alpha_\epsilon,\tilde{\beta}})\big|\mathcal{F}_{\theta_k}^{\alpha_\epsilon\tilde{\beta}}\right)\mathbf{1}_{\Sigma_{\theta_k}^{\alpha_\epsilon\tilde{\beta}}} + \|g\|_\infty\mathbf{1}_{(\Sigma_{\theta_k}^{\alpha_\epsilon\tilde{\beta}})^c} + 4\eta \qquad (19) \\
&\leq \mathbf{E}_{\alpha_\epsilon}\left(g_2(X_T^{t_0,x_0,\alpha_\epsilon,\tilde{\beta}})\big|\mathcal{F}_{\theta_k}^{\alpha_\epsilon\tilde{\beta}}\right) + \|g\|_\infty\mathbf{1}_{(\Sigma_{\theta_k}^{\alpha_\epsilon\tilde{\beta}})^c} + 4\eta
\end{aligned}$$

assuming g_2 is non negative, which is possible without lack of generality because this function is bounded.

It remains to introduce this estimate (19) in inequality (16) to get for all $i \in \{0, \ldots, N\bar{N} - 1\}$, if $t_i \in [\theta_{(k-1)}, \theta_k)$:

$$\begin{aligned}
&g_2(X_T^{t_0,x_0,\alpha_\epsilon,\beta})\mathbf{1}_{\mathbf{S}_T^{\alpha_\epsilon}(\beta)\in\{t_i\}\times\mathbb{R}^n} \\
&\leq \left[\mathbf{E}_{\alpha_\epsilon}\left(V_2(\theta_k, X_{\theta_k}^{t_0,x_0,\alpha_\epsilon,\tilde{\beta}})\big|\mathcal{F}_{t_i}^{\alpha_\epsilon\tilde{\beta}}\right) + 3\epsilon\right]\mathbf{1}_{\mathbf{S}_T^{\alpha_\epsilon}(\beta)\in\{t_i\}\times\mathbb{R}^n} \\
&\leq \mathbf{E}_{\alpha_\epsilon}\left(\mathbf{E}_{\alpha_\epsilon}\left(g_2(X_T^{t_0,x_0,\alpha_\epsilon,\tilde{\beta}})\big|\mathcal{F}_{\theta_k}^{\alpha_\epsilon\tilde{\beta}}\right) + \|g\|_\infty\mathbf{1}_{(\Sigma_{\theta_k}^{\alpha_\epsilon\tilde{\beta}})^c} + 4\eta\Big|\mathcal{F}_{t_i}^{\alpha_\epsilon\tilde{\beta}}\right)\mathbf{1}_{\mathbf{S}_T^{\alpha_\epsilon}(\beta)\in\{t_i\}\times\mathbb{R}^n} \\
&\quad + 3\epsilon\mathbf{1}_{\mathbf{S}_T^{\alpha_\epsilon}(\beta)\in\{t_i\}\times\mathbb{R}^n} \text{ thanks to (19)} \\
&\leq \mathbf{E}_{\alpha_\epsilon}\left(g_2(X_T^{t_0,x_0,\alpha_\epsilon,\tilde{\beta}})\big|\mathcal{F}_{t_i}^{\alpha_\epsilon\tilde{\beta}}\right)\mathbf{1}_{\mathbf{S}_T^{\alpha_\epsilon}(\beta)\in\{t_i\}\times\mathbb{R}^n} \\
&\quad + \|g\|_\infty\mathbf{P}_{\alpha_\epsilon}\left((\Sigma_{\theta_k}^{\alpha_\epsilon\tilde{\beta}})^c\big|\mathcal{F}_{t_i}^{\alpha_\epsilon\tilde{\beta}}\right)\mathbf{1}_{\mathbf{S}_T^{\alpha_\epsilon}(\beta)\in\{t_i\}\times\mathbb{R}^n} + 7\epsilon\mathbf{1}_{\mathbf{S}_T^{\alpha_\epsilon}(\beta)\in\{t_i\}\times\mathbb{R}^n}
\end{aligned}$$

Note that at time t_i there is no deviation, implying $\mathcal{F}_{t_i}^{\alpha_\epsilon\tilde{\beta}} = \mathcal{F}_{t_i}^{\alpha_\epsilon\beta}$ and

$$g_2(X_T^{t_0,x_0,\alpha_\epsilon,\beta})\mathbf{1}_{\mathbf{S}_T^{\alpha_\epsilon}(\beta)\in\{t_i\}\times\mathbb{R}^n} \leq \mathbf{E}_{\alpha_\epsilon}\left(g_2(X_T^{t_0,x_0,\alpha_\epsilon,\tilde{\beta}})\big|\mathcal{F}_{t_i}^{\alpha_\epsilon\beta}\right)\mathbf{1}_{\mathbf{S}_T^{\alpha_\epsilon}(\beta)\in\{t_i\}\times\mathbb{R}^n}$$
$$+\|g\|_\infty\mathbf{P}_{\alpha_\epsilon}\left((\Sigma_{\theta_k}^{\alpha_\epsilon\tilde{\beta}})^c\big|\mathcal{F}_{t_i}^{\alpha_\epsilon\beta}\right)\mathbf{1}_{\mathbf{S}_T^{\alpha_\epsilon}(\beta)\in\{t_i\}\times\mathbb{R}^n} + 7\epsilon\mathbf{1}_{\mathbf{S}_T^{\alpha_\epsilon}(\beta)\in\{t_i\}\times\mathbb{R}^n}$$

Using the fact that $\{\mathbf{S}^{\alpha_\epsilon}(\beta)\in\{t_i\}\times\mathbb{R}^n\}$ is $(\mathcal{F}_{t_i}^{\alpha_\epsilon\beta})$-measurable due to the definition of the strategy α_ϵ, we get for $i = 0\ldots N\bar{N}-1$, if $t_i\in[\theta_{k-1},\theta_k)$:

$$g_2(X_T^{t_0,x_0,\alpha_\epsilon,\beta})\mathbf{1}_{\mathbf{S}_T^{\alpha_\epsilon}(\beta)\in\{t_i\}\times\mathbb{R}^n} \leq \mathbf{E}_{\alpha_\epsilon}\left(g_2(X_T^{t_0,x_0,\alpha_\epsilon,\tilde{\beta}})\mathbf{1}_{\mathbf{S}_T^{\alpha_\epsilon}(\beta)\in\{t_i\}\times\mathbb{R}^n}\big|\mathcal{F}_{t_i}^{\alpha_\epsilon\beta}\right)$$
$$+\|g\|_\infty\mathbf{E}_{\alpha_\epsilon}\left(\mathbf{1}_{(\Sigma_{\theta_k}^{\alpha_\epsilon\tilde{\beta}})^c}\mathbf{1}_{\mathbf{S}_T^{\alpha_\epsilon}(\beta)\in\{t_i\}\times\mathbb{R}^n}\big|\mathcal{F}_{t_i}^{\alpha_\epsilon\beta}\right) + 7\epsilon\mathbf{1}_{\mathbf{S}_T^{\alpha_\epsilon}(\beta)\in\{t_i\}\times\mathbb{R}^n} \quad (20)$$

We now use this estimate to compute the expectation of the payoff in case there is some deviation:

$$\begin{aligned}
&\sum_{i=0}^{N\bar{N}-1}\mathbf{E}_{\alpha_\epsilon}\left(g_2(X_T^{t_0,x_0,\alpha_\epsilon,\beta})\mathbf{1}_{\mathbf{S}_T^{\alpha_\epsilon}(\beta)\in\{t_i\}\times\mathbb{R}^n}\right)\\
&\leq \sum_{i=0}^{N\bar{N}-1}\mathbf{E}_{\alpha_\epsilon}\left(\mathbf{E}_{\alpha_\epsilon}\big(g_2(X_T^{t_0,x_0,\alpha_\epsilon,\tilde{\beta}})\mathbf{1}_{\mathbf{S}_T^{\alpha_\epsilon}(\beta)\in\{t_i\}\times\mathbb{R}^n}|\mathcal{F}_{t_i}^{\alpha_\epsilon\beta}\big)\right)\\
&\quad+\sum_{k=1}^{N}\sum_{i=(k-1)\bar{N}}^{k\bar{N}-1}\mathbf{E}_{\alpha_\epsilon}\left(\|g\|_\infty\mathbf{E}_{\alpha_\epsilon}(\mathbf{1}_{(\Sigma_{\theta_k}^{\alpha_\epsilon\tilde{\beta}})^c}\mathbf{1}_{\mathbf{S}_T^{\alpha_\epsilon}(\beta)\in\{t_i\}\times\mathbb{R}^n}|\mathcal{F}_{t_i}^{\alpha_\epsilon\beta})\right) + 7\epsilon \text{ due to (20)}\\
&\leq \mathbf{E}_{\alpha_\epsilon}\left(g_2(X_T^{t_0,x_0,\alpha_\epsilon,\tilde{\beta}})\mathbf{1}_{\mathbf{S}_T^{\alpha_\epsilon}(\beta)\neq\emptyset}\right) + \|g\|_\infty\sum_{k=1}^{N}\mathbf{E}_{\alpha_\epsilon}\left(\mathbf{1}_{(\Sigma_{\theta_k}^{\alpha_\epsilon\tilde{\beta}})^c}\mathbf{1}_{\mathbf{S}_T^{\alpha_\epsilon}(\beta)\in[\theta_{k-1},\theta_k)\times\mathbb{R}^n}\right) + 7\epsilon\\
&\leq \mathbf{E}_{\alpha_\epsilon}\left(g_2(X_T^{t_0,x_0,\alpha_\epsilon,\tilde{\beta}})\mathbf{1}_{\mathbf{S}_T^{\alpha_\epsilon}(\beta)\neq\emptyset}\right) + \|g\|_\infty\sum_{k=1}^{N}\mathbf{P}_{\alpha_\epsilon}\left((\Sigma_{\theta_k}^{\alpha_\epsilon\tilde{\beta}})^c\right) + 7\epsilon\\
&\leq \mathbf{E}_{\alpha_\epsilon}\left(g_2(X_T^{t_0,x_0,\alpha_\epsilon,\tilde{\beta}})\mathbf{1}_{\mathbf{S}_T^{\alpha_\epsilon}(\beta)\neq\emptyset}\right) + \|g\|_\infty\sum_{k=1}^{N}\frac{2\epsilon}{N} + 7\epsilon \text{ thanks to Lemma 2.8}\\
&\leq \mathbf{E}_{\alpha_\epsilon}\left(g_2(X_T^{t_0,x_0,\alpha_\epsilon,\tilde{\beta}})\mathbf{1}_{\mathbf{S}_T^{\alpha_\epsilon}(\beta)\neq\emptyset}\right) + 2\|g\|_\infty\epsilon + 7\epsilon
\end{aligned} \quad (21)$$

Going back to our estimate of $\mathfrak{J}_2(t_0, x_0, \alpha_\epsilon, \beta)$ as in (15) we can write:

$$\begin{aligned}
\mathfrak{J}_2(t_0, x_0, \alpha_\epsilon, \beta) = & \sum_{i=0}^{N\bar{N}-1} \mathbf{E}_{\alpha_\epsilon}\left(g_2(X_T^{t_0,x_0,\alpha_\epsilon,\beta})\mathbf{1}_{\mathbf{S}_T^{\alpha_\epsilon}(\beta)\in\{t_i\}\times\mathbb{R}^n}\right) \\
& + \mathbf{E}_{\alpha_\epsilon}\left(g_2(X_T^{t_0,x_0,\alpha_\epsilon,\tilde{\beta}})\mathbf{1}_{\mathbf{S}_T^{\alpha_\epsilon}(\beta)=\emptyset}\right) \\
\leq & \, \mathbf{E}_{\alpha_\epsilon}\left(g_2(X_T^{t_0,x_0,\alpha_\epsilon,\tilde{\beta}})\mathbf{1}_{\mathbf{S}_T^{\alpha_\epsilon}(\beta)\neq\emptyset}\right) + (2\|g\|_\infty + 7)\epsilon \\
& + \mathbf{E}_{\alpha_\epsilon}\left(g_2(X_T^{t_0,x_0,\alpha_\epsilon,\tilde{\beta}})\mathbf{1}_{\mathbf{S}_T^{\alpha_\epsilon}(\beta)=\emptyset}\right) \text{ due to (21)} \\
\leq & \, \mathbf{E}_{\alpha_\epsilon}(g_2(X_T^{t_0,x_0,\alpha_\epsilon,\tilde{\beta}})) + (2\|g\|_\infty + 7)\epsilon \\
\leq & \, \mathfrak{J}_2(t_0, x_0, \alpha_\epsilon, \beta_\epsilon) + \frac{6\epsilon}{N} + (2\|g\|_\infty + 7)\epsilon \text{ thanks to (14)}
\end{aligned}$$

This proves that β_ϵ is $(13 + 2\|g\|_\infty)\epsilon$ optimal. The proof is symmetric to state that α_ϵ is $(13 + 2\|g\|_\infty)\epsilon$ optimal.

Finally, we have build mixed strategies $(\alpha_\epsilon, \beta_\epsilon)$ rewarding a payoff 3ϵ close to e and $(13 + 2\|g\|_\infty)\epsilon$ optimal. This proves e is a Nash equilibrium payoff. □

2.2 Convexity of the set of Nash equilibrium payoffs using mixed strategies

Proposition 2.9. *The set $\mathcal{E}_r(t_0, x_0)$ of all Nash equilibrium payoffs in mixed strategies is convex and compact in $\mathbb{R}^2$.*

Proof. Let $(e^1, e^2) \in \mathbb{R}^4$ be a pair of Nash equilibrium payoffs in mixed strategies. We will prove that $(\lambda e^1 + (1-\lambda)e^2)$ is a Nash equilibrium payoff in mixed strategies for all $\lambda \in (0,1)$. We will simply build random controls satisfying the characterization property of Theorem 2.1. As for $j = 1, 2$, e^j is a Nash equilibrium payoff, we may choose random controls (u^j, v^j) on an underlying finite probability space $(\Omega^j, \mathcal{P}(\Omega^j), \mathbf{P}^j)$ such that $\forall i, j = 1,\ 2$:

- $|\mathbf{E}^j(g_i(X_T^{t_0,x_0,u^j,v^j})) - e_i^j| \leq \frac{\epsilon}{3}$
- $\forall t \in [t_0, T]$, denoting by $\mathcal{F}_t^j = \sigma\left((u^j, v^j)(s),\ s \in [t_0, t]\right)$:

$$\mathbf{P}^j\left\{V_i(t, X_t^{t_0,x_0,u^j,v^j}) \leq \mathbf{E}^j\left(g_i(X_T^{t_0,x_0,u^j,v^j})\big|\mathcal{F}_t^j\right) + \frac{\epsilon}{3}\right\} \geq 1 - \epsilon$$

We need to build controls close to the initial pairs (u^j, v^j), $j = 1, 2$, but with some tag in order to distinguish them. Set some small delay $\delta > 0$ such that for all $x \in B(x_0, \delta\|f\|_\infty)$, for all $(u, v) \in \mathcal{U}(t_0) \times \mathcal{V}(t_0)$, for all $i = 1, 2$, for all $t \geq t_0 + \delta$:

$$\left\{\begin{array}{l}
\left|V_i(t, X_t^{t_0,x_0,u,v}) - V_i(t-\delta, X_{t-\delta}^{t_0,x,u,v})\right| \leq \frac{\epsilon}{3} \\
\left|g_i(X_T^{t_0,x_0,u,v}) - g_i(X_{T-\delta}^{t_0,x,u,v})\right| \leq \frac{\epsilon}{3}
\end{array}\right. \tag{22}$$

We now choose some $u_1 \neq u_2 \in U$ and $v_1 \neq v_2 \in V$ and set for $j = 1, 2$:

$$\left\{\begin{array}{ll}
\bar{u}^j(s) = u_j & \text{for } s \in [t_0, t_0 + \delta) \\
\bar{u}^j(s) = u^j(s-\delta) & \text{for } s \in [t_0 + \delta, T] \\
\bar{v}^j(s) = v_j & \text{for } s \in [t_0, t_0 + \delta) \\
\bar{v}^j(s) = v^j(s-\delta) & \text{for } s \in [t_0 + \delta, T]
\end{array}\right.$$

We will as usual denote by $\bar{X}^j_\cdot = X^{t_0,x_0,\bar{u}^j,\bar{v}^j}_\cdot$ for $j = 1, 2$. We immediately get thanks to (22) $\forall i, j = 1,\ 2$:

$$|\mathbf{E}^j(g_i(\bar{X}^j_T)) - e^j_i| \leq 2\frac{\epsilon}{3} \leq \epsilon \tag{23}$$

and for all $t \in [t_0, T]$, denoting by $\bar{\mathcal{F}}^j_t = \sigma\left((\bar{u}^j, \bar{v}^j)(s),\ s \in [t_0, t]\right)$:

$$\mathbf{P}^j\left\{V_i(t, \bar{X}^j_t) \leq \mathbf{E}^j\left(g_i(\bar{X}^j_T)\middle|\bar{\mathcal{F}}^j_t\right) + \epsilon\right\} \geq 1 - \epsilon \tag{24}$$

For $i, j = 1, 2$, denote by

$$\Sigma^{ij}_t = \left\{V_i(t, \bar{X}^j_t) \leq \mathbf{E}^j\left(g_i(\bar{X}^j_T)\middle|\bar{\mathcal{F}}^j_t\right) + \epsilon\right\}$$

We now define a new finite random space $\Omega = \{1, 2\} \times \Omega^1 \times \Omega^2$ endowed with the probability $\mathbf{P}$ defined for all $\omega = (j, \omega^1, \omega^2)$ by:

$$\begin{cases} \mathbf{P}(j, \omega^1, \omega^2) = \lambda \mathbf{P}^1(\omega^1)\mathbf{P}^2(\omega^2) & \text{if } j = 1 \\ \mathbf{P}(j, \omega^1, \omega^2) = (1-\lambda)\mathbf{P}^1(\omega^1)\mathbf{P}^2(\omega^2) & \text{if } j = 2 \end{cases}$$

and define on Ω the random controls (u, v) defined by:

$$\begin{cases} (u, v)(j, \omega^1, \omega^2) = (\bar{u}^1, \bar{v}^1)(\omega^1) & \text{if } j = 1 \\ (u, v)(j, \omega^1, \omega^2) = (\bar{u}^2, \bar{v}^2)(\omega^2) & \text{if } j = 2 \end{cases}$$

We will denote by $X_\cdot = X^{t_0,x_0,u,v}_\cdot$. It remains to prove that for $i = 1, 2$:

- $|\mathbf{E}[g_i(X_T)] - \lambda e^1_i - (1-\lambda)e^2_i| \leq \epsilon$
- $\forall t \in [t_0, T]$, denoting by $\mathcal{F}_t = \sigma\left((u, v)(s),\ s \in [t_0, t]\right)$:

$$\mathbf{P}\left\{V_i(t, X_t) \leq \mathbf{E}\left[g_i(X_T)\middle|\mathcal{F}_t\right] + \epsilon\right\} \geq 1 - \epsilon$$

The first relation is easy to prove. For $i = 1, 2$, we have:

$$\begin{aligned} &\left|\mathbf{E}[g_i(X_T)] - \lambda e^1_i - (1-\lambda)e^2_i\right| \\ &\quad = \left|\lambda \mathbf{E}^1[g_i(\bar{X}^1_T)] + (1-\lambda)\mathbf{E}^2[g_i(\bar{X}^2_T)] - \lambda e^1_i - (1-\lambda)e^2_i\right| \\ &\quad \leq \lambda\left|\mathbf{E}^1[g_i(\bar{X}^1_T)] - e^1_i\right| + (1-\lambda)\left|\mathbf{E}^2[g_i(\bar{X}^2_T)] - e^2_i\right| \\ &\quad \leq \epsilon \text{ thanks to (23)} \end{aligned}$$

In order to prove the second inequality, for $i = 1, 2$, we denote by

$$\Sigma^i_t = \left\{V_i(t, X_t) \leq \mathbf{E}\left(g_i(X_T)\middle|\mathcal{F}_t\right) + 3\epsilon\right\}$$

We have for $i = 1, 2$ and $t \in [t_0 + \delta, T]$:

$$\begin{aligned} \mathbf{E}\left(g_i(X_T)\middle|\mathcal{F}_t\right) &= \mathbf{E}\left(g_i(X_T)(\mathbf{1}_{\{1\}\times\Omega^1\times\Omega^2} + \mathbf{1}_{\{2\}\times\Omega^1\times\Omega^2})\middle|\mathcal{F}_t\right) \\ &= \mathbf{E}\left(g_i(\bar{X}^1_T)\middle|\bar{\mathcal{F}}^1_t\right)\mathbf{1}_{\{1\}\times\Omega^1\times\Omega^2} + \mathbf{E}\left(g_i(\bar{X}^2_T)\middle|\bar{\mathcal{F}}^2_t\right)\mathbf{1}_{\{2\}\times\Omega^1\times\Omega^2} \end{aligned}$$

Therefore, assuming as usual that the functions g_i are non negative and using (24):

$$\begin{aligned} \mathbf{E}\left(g_i(X_T)\middle|\mathcal{F}_t\right) &\geq [V_i(t, \bar{X}^1_t) - \epsilon]\mathbf{1}_{\{1\}\times\Sigma^{i1}_t\times\Omega^2} + [V_i(t, \bar{X}^2_t) - \epsilon]\mathbf{1}_{\{2\}\times\Omega^1\times\Sigma^{i2}_t} \\ &\geq V_i(t, X_t)\mathbf{1}_{\{1\}\times\Sigma^{i1}_t\times\Omega^2} + V_i(t, X_t)\mathbf{1}_{\{2\}\times\Omega^1\times\Sigma^{i2}_t} - \epsilon \end{aligned}$$

And finally:

$$\begin{aligned}\mathbf{P}\left(\Sigma_t^i\right) &\geq \mathbf{P}\left[\{1\}\times\Sigma_t^{i1}\times\Omega^2\cup\{2\}\times\Omega^1\times\Sigma_t^{i2}\right]\\ &\geq \lambda(1-\epsilon)+(1-\lambda)(1-\epsilon)\\ &\geq 1-\epsilon\end{aligned}$$

Note that for $t\in[t_0,t_0+\delta]$, the preceding relation is straightforward.

2.3 Comparison between the sets of Nash equilibrium payoffs in pure and mixed strategies

We already know that the set of Nash equilibrium payoffs in mixed strategies contains the convex hull of the set of Nash equilibrium payoffs in pure strategies. We now try to compare these two sets and it appears that in general, they are not equal. This result is not intuitive because the punitive strategies and the guaranteed payoffs of the players are exactly the same whether players use pure or mixed strategies.

Proposition 2.10. *There exist non zero sum differential games such that the set of Nash equilibrium payoffs in mixed strategies is larger than the convex hull of the set of Nash equilibrium payoffs in pure strategies:*

$$\mathcal{E}_r(t_0,x_0)\supsetneq\overline{Conv}\mathcal{E}(t_0,x_0).$$

Proof. We will build a counter-example where a Nash equilibrium payoff in mixed strategies does not belong to the closed convex hull of the Nash equilibrium payoffs in pure strategies.
Consider the simple game in finite time in $\mathbb{R}^2$ with dynamics :

$$\dot{x}=u+v\qquad u,\ v\in[-1/2,1/2]^2$$

starting from the origin $O=(0,0)$ at time $t=0$ and ending at time $t=T=1$ with $x=(x_1,x_2)$.

The payoff functions are the Lipschitz continuous functions defined as follows:

$$g_1:\begin{cases} g_1(x)=1-4|x_2| & \text{for } |x_2|\leq 1/4 \text{ and } |x_2|\geq|x_1|\\ g_1(x)=1-4|x_1| & \text{for } |x_1|\leq 1/4 \text{ and } |x_1|\geq|x_2|\\ g_1(x)=x_2+2|x_1|-1 & \text{for } x_2\geq -2|x_1|+1\\ g_1(x)=0 & \text{elsewhere}\end{cases}$$

In fact g_1 is a non negative function defined on the unit square centered at the origin and its graph looks like a pyramid of height 1 with base the square with length 1/2 centered at the origin, completed with two symmetric triangles joining $(0,1,0)$, $(1,1,2)$ and $(1,-1,0)$ for the first and $(0,1,0)$, $(-1,1,2)$ and $(-1,-1,0)$ for the second.

$$g_2:\begin{cases} g_2(x)=0 & \text{for } x_2\geq 0\\ g_2(x)=-x_2 & \text{for } x_2\leq 0\end{cases}$$

The game clearly fulfill the regularity assumptions listed in the introduction. We will denote by L the greater of the Lipschitz-constants of g_1 and g_2 for the L^1-norm.

The set of all reachable points starting at time 0 from the origin O and stopping at time $T = 1$ is the unit ball in $\mathbb{R}^2$ for the L^1 norm and the set of all reachable payoffs is $[0,2] \times \{0\} \cup \bigcup_{y\in(0,1]}([0,1-y],y)$. It is also clear that

$$\begin{cases} V_1(t,x) = g_1(x) \\ V_2(t,x) = g_2(x) \end{cases}$$

The initial values are $V_1(0,O) = 1$ and $V_2(0,O) = 0$, implying any Nash equilibrium payoff has to reward players I with at least 1 and player II with a non-negative payoff, meaning no trajectory can end up at time T at some x such that $x_2 < 0$ because this would cause player I to earn strictly less than 1. We then have $e_2 = 0$ corresponding to $y \geq 0$ for every Nash equilibrium payoff in pure strategies. Considering Nash equilibrium payoffs in pure strategies, we can easily compute

$$\mathcal{E}(0,O) = [1,2] \times \{0\} = \overline{Conv}\mathcal{E}(0,O).$$

It is the set of all reachable payoffs such that $e_1 \geq 1$.

We now will compute some random controls (u,v) leading to a final payoff of 1 for player I and positive for player II. The controls induce the trajectories joining $(3/4,0)$ at time $t = 3/4$ ($u = v = (1/2,0)$) and then with probability one half:

- from $t = 3/4$ on, join $(1,1/4)$ at $t = 1$ ($u = v = (1/2,1/2)$) and get the payoff $(5/4,0)$
- from $t = 3/4$ on, join $(1,-1/4)$ at $t = 1$ ($u = v = (1/2,-1/2)$) and get the payoff $(3/4,1/4)$

The final payoff will be $(e_1,e_2) = (1,1/8) \notin \overline{Conv}\mathcal{E}(0,O)$. We will as usual denote by $X. = X^{0,O,u,v}$.

It remains to prove that this payoff is a Nash equilibrium payoff in mixed strategies. We use the characterization of the Nash equilibrium payoffs of theorem 2.1 and prove that along the trajectories the condition

$$\mathbf{E}[g_i(X_T)|\mathcal{F}_t] \geq V_i(t,X_t)$$

is satisfied. Indeed, along the trajectories:

$$\begin{array}{llll} \text{for } t \in [0,1/4] & & V_1(t,X_t) = 1-4t & \text{and } \mathbf{E}\left(g_1(X_T)\middle|\mathcal{F}_t\right) = 1 \\ \text{for } t \in [1/4,1/2] & & V_1(t,X_t) = 0 & \text{and } \mathbf{E}\left(g_1(X_T)\middle|\mathcal{F}_t\right) = 1 \\ \text{for } t \in [1/2,3/4] & & V_1(t,X_t) = 2t-1 & \text{and } \mathbf{E}\left(g_1(X_T)\middle|\mathcal{F}_t\right) = 1 \\ \text{for } t \in (3/4,1]: & \text{either} & V_1(t,X_t) = 3t-7/4 & \text{and } \mathbf{E}\left(g_1(X_T)\middle|\mathcal{F}_t\right) = 5/4 \\ & \text{or} & V_1(t,X_t) = t-1/4 & \text{and } \mathbf{E}\left(g_1(X_T)\middle|\mathcal{F}_t\right) = 3/4 \end{array}$$

and

$$\begin{cases} \begin{array}{llll} \text{for } t \in [0,3/4] & & V_2(t,X_t) = 0 & \text{and } \mathbf{E}\left(g_2(X_T)\middle|\mathcal{F}_t\right) = 1/8 \\ \text{for } t \in (3/4,1]: & \text{either} & V_2(t,X_t) = 0 & \text{and } \mathbf{E}\left(g_2(X_T)\middle|\mathcal{F}_t\right) = 0 \\ & \text{or} & V_2(t,X_t) = t-3/4 & \text{and } \mathbf{E}\left(g_2(X_T)\middle|\mathcal{F}_t\right) = 1/4 \end{array} \end{cases}$$

We can see that for $t \leq 3/4$, the condition is satisfied. Now, for each trajectory from $t = 3/4$ on, we have $V_i(t,X_t)$ is either non decreasing or constant so that for $i = 1,2$ $V_i(t,X_t) \leq V_i(T,X_T) = g_i(X_T)$ and the condition is satisfied.
This proves that the final payoff $(e_1,e_2) = (1,1/8)$ is a Nash equilibrium payoff in mixed strategies. □

3 Publicly correlated equilibrium payoffs

We recall that $\mathcal{E}_c(t_0, x_0) \supset \mathcal{E}(t_0, x_0)$. We are going to state some characterization of publicly correlated equilibrium payoffs and compare the set of Nash equilibrium payoffs using mixed strategies and the set of publicly correlated equilibrium payoffs.

Theorem 3.1. *The set of publicly correlated equilibrium payoffs is equal to the set of Nash equilibrium payoffs using mixed strategies.*

Proof. To begin with, we will show that $\mathcal{E}_r(t_0, x_0) \subseteq \mathcal{E}_c(t_0, x_0)$. We will prove that publicly correlated equilibrium payoffs satisfy the same kind of characterization as Nash equilibrium payoffs using mixed strategies:

Proposition 3.2 (Characterization of publicly correlated equilibrium payoffs)**.** *The payoff $e = (e_1, e_2) \in \mathbb{R}^2$ is a publicly correlated equilibrium payoff iff for all $\epsilon > 0$, there exist random controls (u_ϵ, v_ϵ), such that $\forall i = 1,\ 2$:*

- $|\mathbf{E}[g_i(X_T^{t_0,x_0,u_\epsilon,v_\epsilon})] - e_i| \leq \epsilon$
- $\forall t \in [t_0, T]$, *if we denote by* $(\mathcal{F}_t^\epsilon) = \sigma\{(u_\epsilon, v_\epsilon)(s),\ s \in [t_0, t]\}$
  $\mathbf{P}\left\{\mathbf{E}\left[g_i(X_T^{t_0,x_0,u_\epsilon,v_\epsilon})\middle|\mathcal{F}_t^\epsilon\right] \geq V_i(t, X_t^{t_0,x_0,u_\epsilon,v_\epsilon}) - \epsilon\right\} \geq 1 - \epsilon$

Proof of Proposition 3.2.
The proof is a simple adaptation of the proof used in [6]. We rewrite it here for the sake of completeness.

We begin with the necessary condition.

Let e be a publicly correlated equilibrium payoff as in Definition 7. Fix ϵ and choose $\frac{\epsilon^2}{2}$ optimal correlated strategies $((\mathcal{F}_t), \alpha_\epsilon, \beta_\epsilon)$ as in the definition: $(\alpha_\epsilon, \beta_\epsilon)$ generates admissible controls (u_ϵ, v_ϵ) that are $(\mathcal{F}_t)$-adapted. The filtration $(\mathcal{F}_t^\epsilon) = (\sigma\{(u_\epsilon, v_\epsilon)(s),\ s \in [t_0, t]\})$ is a subfiltration of $(\mathcal{F}_t)$. We will denote by $X_\cdot^\epsilon = X_\cdot^{t_0,x_0,u_\epsilon,v_\epsilon}$.

For $\epsilon < 1$, it is clear that $|\mathbf{E}[g_i(X_T^\epsilon)] - e_i| \leq \epsilon^2/2 \leq \epsilon$.

We will prove that (u_ϵ, v_ϵ) satisfy the second part of the Proposition.
Suppose on the contrary that there exist $\tau \in [t_0, T]$ such that:

$$\mathbf{P}\left\{\mathbf{E}\left(g_1(X_T^\epsilon)\middle|\mathcal{F}_\tau^\epsilon\right) \geq V_1(\tau, X_\tau^\epsilon) - \epsilon\right\} < 1 - \epsilon$$

We denote by

$$\Sigma_\epsilon := \left\{\omega/\ V_1(\tau, X_\tau^\epsilon) \leq \mathbf{E}\left(g_1(X_T^\epsilon)\middle|\mathcal{F}_\tau^\epsilon\right) + \epsilon\right\}$$

We have $\mathbf{P}(\Sigma_\epsilon) < 1 - \epsilon$.

We will use the maximin strategy $\alpha_g^{\frac{\epsilon}{2},\tau}(x)$ as in Lemma 2.2. We define the strategy α_ϵ^V: for all $v \in \tilde{\mathcal{V}}(t_0)$

- $\alpha_\epsilon^V(v)|_{[t_0,\tau]} := u_\epsilon|_{[t_0,\tau]}$
- $\alpha_\epsilon^V(v)|_{(\tau,T]} := \alpha_g^{\frac{\epsilon}{2},\tau}(X_\tau^{t_0,x_0,u_\epsilon,v})(v|_{[\tau,T]})$

Consider the pair $(u_\epsilon^V, v_\epsilon^V)$ of admissible controls generated by the $(\mathcal{F}_t)$ publicly correlated strategies $(\alpha_\epsilon^V, \beta_\epsilon)$, and build the admissible control $\tilde{u}_\epsilon$:

- $\tilde{u}_\epsilon = u_\epsilon$ on $([t_0, \tau) \times \Omega_\epsilon) \cup ([\tau, T] \times \Sigma_\epsilon)$
- $\tilde{u}_\epsilon = u_\epsilon^V$ on $[\tau, T] \times \Sigma_\epsilon^c$

Note that $\tilde{u}_\epsilon$ is progressively measurable with respect to $(\mathcal{F}_t)$. We have:

- $\beta_\epsilon(\tilde{u}_\epsilon) \equiv v_\epsilon$ on $[t_0, \tau)$
- $\beta_\epsilon(\tilde{u}_\epsilon) \equiv v_\epsilon$ on $[\tau, T) \times \Sigma_\epsilon$
- $\beta_\epsilon(\tilde{u}_\epsilon) \equiv v_\epsilon^V$ on $[\tau, T) \times \Sigma_\epsilon^c$

and finally:

$$\begin{aligned}
\mathfrak{J}_1(t_0, x_0, \tilde{u}_\epsilon, \beta_\epsilon) &= \mathbf{E}\left(g_1(X_T^\epsilon)\mathbf{1}_{\Sigma_\epsilon}\right) + \mathbf{E}\left(g_1(X_T^{\tau, X_\tau^\epsilon, u_\epsilon^V, v_\epsilon^V})\mathbf{1}_{\Sigma_\epsilon^c}\right) \\
&\geq \mathbf{E}\left(g_1(X_T^\epsilon)\mathbf{1}_{\Sigma_\epsilon}\right) + \mathbf{E}\left(V_1(\tau, X_\tau^\epsilon)\mathbf{1}_{\Sigma_\epsilon^c}\right) - \frac{\epsilon}{2}\mathbf{P}\left(\Sigma_\epsilon^c\right) \\
&\geq \mathbf{E}\left(g_1(X_T^\epsilon)\mathbf{1}_{\Sigma_\epsilon}\right) + \mathbf{E}\left(\mathbf{E}\big(g_1(X_T^\epsilon)\big|\mathcal{F}_\tau^\epsilon\big)\mathbf{1}_{\Sigma_\epsilon^c}\right) + (\epsilon - \frac{\epsilon}{2})\mathbf{P}\left(\Sigma_\epsilon^c\right) \\
&> \mathfrak{J}_1(t_0, x_0, \alpha_\epsilon, \beta_\epsilon) + \frac{\epsilon^2}{2}
\end{aligned}$$

This leads to a contradiction with the $\frac{\epsilon^2}{2}$-optimality of $(\alpha_\epsilon, \beta_\epsilon)$. Therefore, for all $t \in [t_0, T]$, for $i = 1, 2$:

$$\mathbf{P}\left\{\mathbf{E}\left[g_i(X_T^{t_0, x_0, u_\epsilon, v_\epsilon})\big|\mathcal{F}_t^\epsilon\right] \geq V_i(t, X_t^{t_0, x_0, u_\epsilon, v_\epsilon}) - \epsilon\right\} \geq 1 - \epsilon$$

It remains to prove the sufficient condition.

Fix $\epsilon > 0$. Set δ small enough such that:

1. $\forall t \in [t_0, T]$, $\forall x \in \mathbb{R}^n$, $\forall y \in B(x, \delta\|f\|_\infty)$, for all $i = 1, 2$:
$$|V_i(t, x) - V_i(t + \delta, y)| \leq \epsilon$$
2. $\forall t \in [t_0, T]$, $\forall x \in \mathbb{R}^n$, $\forall y \in B(x, \delta\|f\|_\infty)$, $\forall (u, v) \in \mathcal{U}(t) \times \mathcal{V}(t)$, for all $i = 1, 2$:
$$|g_i(X_T^{t,x,u,v}) - g_i(X_T^{t,y,u,v})| \leq \epsilon$$
3. $\exists N \in \mathbb{N}^*$ such that $N\delta = T - t_0$,

in order to build a time partition $t_0, \ldots, t_k = t_0 + k\delta, \ldots, t_N = T$.
Set $\eta = \epsilon/N$. By assumption, we have a filtration $(\mathcal{F}_t^\eta)$ and a pair of correlated controls (u_η, v_η) satisfying the conditions of Proposition 3.2. We will as usual denote by $X_\cdot^\eta = X_\cdot^{t_0, x_0, u_\eta, v_\eta}$. It remains to build a pair of correlated strategies $(\alpha_\epsilon, \beta_\epsilon)$.
We will need "punitive strategies" as in Definition 2.4. The idea of the strategy α_ϵ is a trigger strategy that plays u_η as long as player II plays the control v_η and as soon as a deviation from the expected control of the opponent is detected, player I will use the punitive strategy.

If the control of player II is v, we introduce the stopping times:

$$S(v) := \inf\{t \geq t_0 / \ v|_{[t_0,t]} \not\equiv v_\eta|_{[t_0,t]}\}$$

and

$$\tau(v) = \min\{t_k \geq t_1 / \ t_k > S(v)\}$$

with the convention $\tau(v) = T$ on $\{S(v) = T\}$.
We now define more precisely the strategy α_ϵ:

$$\alpha_\epsilon(v) = \begin{cases} u_\eta & \text{on } [\![t_0, \tau(v)]\!] \\ \alpha_p^{\epsilon,t_k}(X_{t_k}^{t_0,x_0,u_\eta,v})(v|_{(t_k,T]}) & \text{on } (t_k, T] \times \{\tau(v) = t_k\} \end{cases}$$

It is clear that α_ϵ is a correlated strategy for the correlation device $(\mathcal{F}_t^\eta)$. Indeed, we have a strongly non-anticipative strategy by definition, and the delay is the minimum of the delays of the N punitive strategies $\alpha_p^{\epsilon,t_k}(\cdot)$, $k = 1, \ldots, N$. The strategy β_ϵ is defined in a symmetric way.

First of all, we check that the pair of controls associated to the correlated strategies $(\alpha_\epsilon, \beta_\epsilon)$ is (u_η, v_η), and we immediately get for $i = 1, 2$:

$$|\mathfrak{J}_i(t_0, x_0, \alpha_\epsilon, \beta_\epsilon) - e_i| \leq \epsilon/N \leq \epsilon$$

It remains to prove that the correlated strategies we have built are ϵ-optimal.
Consider a strategy β of player II, such that $((\mathcal{F}_t^\eta), \alpha_\epsilon, \beta)$ are correlated strategies. Denote by (u, v) the pair of controls associated to (α_ϵ, β). For any $k = 1, \ldots, N-1$ we have:

$$\begin{aligned} g_2(X_T^{t_0,x_0,u,v})\mathbf{1}_{\{\tau(v)=t_k\}} &= g_2\left(X_T^{t_k, X_{t_k}^{t_0,x_0,u,v}, \alpha_p^{\epsilon,t_k}(X_{t_k}^{t_0,x_0,u_\eta,v}), v}\right)\mathbf{1}_{\{\tau(v)=t_k\}} \\ &\leq V_2(t_k, X_{t_k}^{t_0,x_0,u_\eta,v})\mathbf{1}_{\{\tau(v)=t_k\}} + \epsilon\mathbf{1}_{\{\tau(v)=t_k\}} \\ &\leq V_2\left(t_k, X_{t_k}^{S(v), X_{S(v)}^\eta, u_\eta, v}\right)\mathbf{1}_{\{\tau(v)=t_k\}} + \epsilon\mathbf{1}_{\{\tau(v)=t_k\}} \\ &\leq V_2(t_k, X_{t_k}^\eta)\mathbf{1}_{\{\tau(v)=t_k\}} + 2\epsilon\mathbf{1}_{\{\tau(v)=t_k\}} \text{ due to the choice of } \delta. \end{aligned} \tag{25}$$

Notice also that $\tau(v) = T$ means $\{v(s) \equiv v_\eta(s) \ \forall s \in [t_0, T-\delta]\}$. We now will use these results to compute $\mathfrak{J}_2(t_0, x_0, \alpha_\epsilon, \beta)$:

$$\begin{aligned} \mathfrak{J}_2(t_0, x_0, \alpha_\epsilon, \beta) &= \mathbf{E}\left(g_2(X_T^{t_0,x_0,u_\eta,v})\mathbf{1}_{\tau(v)=T}\right) + \mathbf{E}\left(g_2(X_T^{t_0,x_0,u,v})\mathbf{1}_{\tau(v)<T}\right) \\ &= \mathbf{E}\left(g_2(X_T^{T-\delta, X_{T-\delta}^\eta, u_\eta, v})\mathbf{1}_{\tau(v)=T}\right) + \sum_{k=1}^{N-1}\mathbf{E}\left(g_2(X_T^{t_0,x_0,u,v})\mathbf{1}_{\tau(v)=t_k}\right) \\ &\leq \mathbf{E}\left(g_2(X_T^\eta) + \epsilon)\mathbf{1}_{\tau(v)=T}\right) \\ &\quad + \sum_{k=1}^{N-1}\mathbf{E}\left((V_2(t_k, X_{t_k}^\eta) + 2\epsilon)\mathbf{1}_{\tau(v)=t_k}\right) \text{ due to (25)} \\ &\leq \mathbf{E}\left(g_2(X_T^\eta)\mathbf{1}_{\tau(v)=T}\right) + \sum_{k=1}^{N-1}\mathbf{E}\left(V_2(t_k, X_{t_k}^\eta)\mathbf{1}_{\tau(v)=t_k}\right) + 2\epsilon \end{aligned} \tag{26}$$

At this point we will try to get an estimate of $V_2(t_k, X^\eta_{t_k})$ for all $k = 1, \dots, N-1$. We have to use the assumptions on (u_η, v_η): we denote by

$$\Sigma^k_\eta := \left\{\omega/\ V_2(t_k, X^\eta_{t_k}) \leq \mathbf{E}\left(g_2(X^\eta_T)\middle|\mathcal{F}^\eta_{t_k}\right) + \eta\right\}$$

and we have $\mathbf{P}(\Sigma^k_\eta) \geq 1 - \epsilon/N$. As V_2 is bounded with a constant K we have for all $k = 1, \dots, N-1$:

$$\begin{aligned}
V_2(t_k, X^\eta_{t_k}) &= V_2(t_k, X^\eta_{t_k})\mathbf{1}_{\Sigma^k_\eta} + V_2(t_k, X^\eta_{t_k})\mathbf{1}_{\Sigma^{k\,c}_\eta} \\
&\leq \left(\mathbf{E}\left(g_2(X^\eta_T)\middle|\mathcal{F}^\eta_{t_k}\right) + \eta\right)\mathbf{1}_{\Sigma^k_\eta} + K\mathbf{1}_{\Sigma^{k\,c}_\eta} \\
&\leq \mathbf{E}\left(g_2(X^\eta_T)\middle|\mathcal{F}^\eta_{t_k}\right) + K\mathbf{1}_{\Sigma^{k\,c}_\eta} + \eta
\end{aligned}$$

if we assume as usual that g_2 is non negative. Introducing this inequality in our estimate (26) of $\mathfrak{J}_2(t_0, x_0, \alpha_\epsilon, \beta)$:

$$\begin{aligned}
\mathfrak{J}_2(t_0, x_0, \alpha_\epsilon, \beta) &\leq \mathbf{E}\left(g_2(X^\eta_T)\mathbf{1}_{\tau(v)=T}\right) + \sum_{k=1}^{N-1}\mathbf{E}\left(V_2(t_k, X^\eta_{t_k})\mathbf{1}_{\tau(v)=t_k}\right) + 2\epsilon \\
&\leq \mathbf{E}\left(g_2(X^\eta_T)\mathbf{1}_{\tau(v)=T}\right) + \sum_{k=1}^{N-1}\mathbf{E}\left(\mathbf{E}\left(g_2(X^\eta_T)\middle|\mathcal{F}^\eta_{t_k}\right)\mathbf{1}_{\tau(v)=t_k}\right) \\
&\quad + \sum_{k=1}^{N-1}\mathbf{E}\left(K\mathbf{1}_{\Sigma^{k\,c}_\eta}\mathbf{1}_{\tau(v)=t_k}\right) + 3\epsilon \\
&\leq \mathbf{E}\left(g_2(X^\eta_T)\mathbf{1}_{\tau(v)=T}\right) + \mathbf{E}\left(g_2(X^\eta_T)\mathbf{1}_{\tau(v)<T}\right) \\
&\quad + K\sum_{k=1}^{N-1}\mathbf{P}\left(\Sigma^{k^c}_\eta\right) + 3\epsilon \\
&\leq \mathbf{E}\left(g_2(X^\eta_T)\right) + (3+K)\epsilon \\
&\leq \mathfrak{J}_2(t_0, x_0, \alpha_\epsilon, \beta_\epsilon) + (3+K)\epsilon
\end{aligned}$$

and we have proven that α_ϵ is $(3+K)\epsilon$-optimal. The result is the same for β_ϵ. In the end, we were able to build $(3+K)\epsilon$-optimal correlated strategies that reward a payoff ϵ-close to e, meaning e is a correlated equilibrium payoff. □

End of the proof of Theorem 3.1
This characterization and Theorem 2.1 ensure that any Nash equilibrium payoff using mixed strategies is in fact a publicly correlated equilibrium payoff.

Note that the only difference with the characterization of Nash equilibrium payoffs is that publicly correlated equilibrium payoffs may rely on random controls defined on an infinite underlying probability space, whereas in our definition, mixed strategies are defined only on finite underlying probability spaces.

We now will prove that $\mathcal{E}_r(t_0, x_0) \supseteq \mathcal{E}_c(t_0, x_0)$. We consider some publicly correlated equilibrium payoff satisfying the characterization of proposition 3.2 and we will prove that we are able to build a finite number of random controls satisfying the characterization of theorem 2.1, implying e will be a Nash equilibrium payoff.

Consider some publicly correlated equilibrium payoff e. Fix ϵ and consider the ϵ^2-optimal random controls (u_ϵ, v_ϵ) on the underlying probability space $(\Omega, \mathcal{F}, \mathbf{P})$. Denote as usual by $X^\epsilon_\cdot = X^{t_0,x_0,u_\epsilon,v_\epsilon}_\cdot$ and set for all $\omega \in \Omega$: $X^\epsilon_\cdot(\omega) = X^{t_0,x_0,(u_\epsilon,v_\epsilon)(\omega)}_\cdot$. Note that these controls satisfy:

$$\begin{cases} |\mathbf{E}[g_i(X^\epsilon_T)] - e_i| \leq \epsilon^2 \\ \forall t \in [t_0, T], \text{ if we denote by } \mathcal{F}^\epsilon_t = \sigma\left\{(u_\epsilon, v_\epsilon)(s),\ s \in [t_0, t]\right\} : \\ \quad \mathbf{P}\left\{\mathbf{E}\left[g_i(X^\epsilon_T)\middle|\mathcal{F}^\epsilon_t\right] \geq V_i(t, X^\epsilon_t) - \epsilon^2\right\} \geq 1 - \epsilon^2 \end{cases} \tag{27}$$

If there are finitely many distinct controls (u_ϵ, v_ϵ) there is nothing left to prove. Else, we will build a finite number of random controls rewarding a payoff close to e and consistent.

We set $h > 0$ and $\bar{h} > 0$ to be defined later such that there exist $N_h, N_{\bar{h}} \in \mathbb{N}^*$ such that $T - t_0 = N_h h$ and $(T - t_0)\|f\|_\infty = N_{\bar{h}}\bar{h}$. We build the following time partition $G_h = \{t_k = t_0 + kh\}_{k=0,\dots,N_h}$ and the grid in $\mathbb{R}^n$: $G_{\bar{h}} = \{x_0 + \sum_{i=1}^n k_i \bar{h} e_i\}_{(k_i)\in\{-N_{\bar{h}},\dots,0,\dots,N_{\bar{h}}\}^n}$ where $(e_i)_{i=1\dots n}$ is a basis of $\mathbb{R}^n$. We now introduce a projection on the grid:

$$\Pi : \begin{array}{l} \mathbb{R}^n \to G_{\bar{h}} \\ x \ \mapsto \min\{x_i \in G_{\bar{h}} / \ d_1(x, x_i) = \inf_{x_j \in G_{\bar{h}}} d_1(x, x_j)\} \end{array}$$

where the minimum is taken with respect to the lexicographic order and d_1 is the distance associated to the norm $\|x\|_1 = \max_{i=1\dots n} |x_i|$. From now on we will always use this norm in $\mathbb{R}^n$.

To any $(t_k, x_i, x_j) \in G_h \times G_{\bar{h}} \times G_{\bar{h}}$ we associate, if it exists some $\varphi(t_k, x_i, x_j) = (x, u, v) \in \mathbb{R}^n \times \mathcal{U}(t_k) \times \mathcal{V}(t_k)$ such that $\Pi(x) = x_i$ and $\Pi(X^{t_k,x,u,v}_{t_{k+1}}) = x_j$. We will set $\varphi_x(t_k, x_i, x_j) = x$ and $\varphi_c(t_k, x_i, x_j) = (u, v)$.

We now are able to build a finite number of random controls defined on $(\Omega, \mathcal{F})$. To any $\omega \in \Omega$ we associate $(u_\eta, v_\eta)(\omega)$ in the following way:

- Fix $(u_0, v_0) \in U \times V$
- $(u_\eta, v_\eta)(\omega)|_{[t_0,t_1)} = (u_0, v_0)$
- for all $k = 1 \dots N_h - 1$, for all $s \in [t_k, t_{k+1})$:

$$(u_\eta, v_\eta)(\omega)(s) = \varphi_c\left(t_{k-1}, \Pi(X^\epsilon_{t_{k-1}}(\omega)), \Pi(X^\epsilon_{t_k}(\omega))\right)(s - h)$$

Note that the definition of (u_η, v_η) is non anticipative. From now on, we will denote by $X^\eta_\cdot = X^{t_0,x_0,u_\eta,v_\eta}_\cdot$ and set for all $\omega \in \Omega$: $X^\eta_\cdot(\omega) = X^{t_0,x_0,(u_\eta,v_\eta)(\omega)}_\cdot$.

We now would like to prove that the set of finitely many random controls (u_η, v_η) defined on $(\Omega, \mathcal{F}, \mathbf{P})$ satisfies for $i = 1, 2$, for some constants C_1, C_2, C_3:

- $|\mathbf{E}[g_i(X^\eta_T)] - e_i| \leq C_1\epsilon$
- $\forall t \in [t_0, T]$, if we denote by $\mathcal{F}^\eta_t = \sigma\{(u_\eta, v_\eta)(s),\ s \in [t_0, t]\}$
  $\mathbf{P}\left\{\mathbf{E}\left(g_i(X^\eta_T)\middle|\mathcal{F}^\eta_t\right) \geq V_i(t, X^\eta_t) - C_2\epsilon\right\} \geq 1 - C_3\epsilon$

First of all, we shall prove that the trajectories generated by $(u_\eta, v_\eta)(\omega)$ and $(u_\epsilon, v_\epsilon)(\omega)$ are close for sufficiently small values of h and $\bar{h}$.

For all $k = 0 \dots N_h - 1$, we have

$$\begin{aligned}
\|X^\eta_{t_{k+1}}(\omega) - X^\epsilon_{t_k}(\omega)\| &\leq \left\| X_{t_k}^{t_{k-1}, X^\eta_{t_k}(\omega), \varphi_c(t_{k-1}, \Pi(X^\epsilon_{t_{k-1}}(\omega)), \Pi(X^\epsilon_{t_k}(\omega)))} - X^\epsilon_{t_k}(\omega) \right\| \\
&\leq \left\| X_{t_k}^{t_{k-1}, \varphi(t_{k-1}, \Pi(X^\epsilon_{t_{k-1}}(\omega)), \Pi(X^\epsilon_{t_k}(\omega)))} - X^\epsilon_{t_k}(\omega) \right\| \\
&\quad + \left\| X_{t_k}^{t_{k-1}, X^\eta_{t_k}(\omega), \varphi_c(t_{k-1}, \Pi(X^\epsilon_{t_{k-1}}(\omega)), \Pi(X^\epsilon_{t_k}(\omega)))} \right. \\
&\qquad \left. - X_{t_k}^{t_{k-1}, \varphi(t_{k-1}, \Pi(X^\epsilon_{t_{k-1}}(\omega)), \Pi(X^\epsilon_{t_k}(\omega)))} \right\| \\
&\leq \bar{h} + \left\| \varphi_x(t_{k-1}, \Pi(X^\epsilon_{t_{k-1}}(\omega)), \Pi(X^\epsilon_{t_k})(\omega)) - X^\eta_{t_k}(\omega) \right\| e^{L_f h} \\
&\leq \bar{h} + \left(\left\| X^\eta_{t_k}(\omega) - X^\epsilon_{t_{k-1}}(\omega) \right\| + \bar{h} \right) e^{L_f h}
\end{aligned}$$

because by definition, $\Pi(X^\epsilon_{t_k}(\omega)) = \Pi\left(X_{t_k}^{t_{k-1}, \varphi(t_{k-1}, \Pi(X^\epsilon_{t_{k-1}}(\omega)), \Pi(X^\epsilon_{t_k}(\omega)))} \right)$ and $\Pi(X^\epsilon_{t_{k-1}}(\omega)) = \Pi\left(\varphi_x(t_{k-1}, \Pi(X^\epsilon_{t_{k-1}}(\omega)), \Pi(X^\epsilon_{t_k}(\omega))) \right)$ and points in $B(x_0, (T - t_0)\|f\|_\infty)$ having the same projection on $G_{\bar{h}}$ are at most $\bar{h}$ distant. Using backward induction, and noticing that $\|X^\eta_{t_1}(\omega) - X^\epsilon_{t_0}(\omega)\| \leq \|f\|_\infty h$, we have that for all $k = 0 \dots N_h - 1$:

$$\begin{aligned}
\left\| X^\eta_{t_{k+1}}(\omega) - X^\epsilon_{t_k}(\omega) \right\| &\leq \bar{h}(1 + e^{L_f h}) \sum_{i=0}^{k-1} e^{i L_f h} + h e^{k L_f h} \|f\|_\infty \\
&\leq 2\bar{h} \frac{T - t_0}{h} e^{L_f (T - t_0)} + h e^{L_f (T - t_0)} \|f\|_\infty
\end{aligned}$$

In order to minimize the distance between $X^\epsilon_\cdot(\omega)$ and $X^\eta_\cdot(\omega)$, we set for example $\bar{h} = h^2$ in order to get for all $k = 0 \dots N_h$:

$$\|X^\epsilon_{t_k}(\omega) - X^\eta_{t_k}(\omega)\| \leq h \left[e^{L_f (T - t_0)} (2(T - t_0) + \|f\|_\infty) + \|f\|_\infty \right]$$

and for all $t \in [t_k, t_{k+1})$:

$$\|X^\epsilon_t(\omega) - X^\eta_t(\omega)\| \leq h \left[e^{L_f (T - t_0)} (2(T - t_0) + \|f\|_\infty) + 3\|f\|_\infty \right]$$

Finally choosing h small enough:

$$\|X^\epsilon_\cdot(\omega) - X^\eta_\cdot(\omega)\|_\infty \leq \epsilon \tag{28}$$

Now, it is easy to check that the final payoff using (u_η, v_η) is close to the payoff generated by (u_ϵ, v_ϵ). Indeed for all $i = 1, 2$:

$$\begin{aligned}
|\mathfrak{J}_i(t_0, x_0, u_\epsilon, v_\epsilon) - \mathfrak{J}_i(t_0, x_0, u_\eta, v_\eta)| &\leq \int_\Omega \left| g_i(X^\epsilon_T(\omega)) - g_i(X^\eta_T(\omega)) \right| \mathrm{d}\mathbf{P}(\omega) \\
&\leq \int_\Omega L_g \left\| X^\epsilon_T(\omega) - X^\eta_T(\omega) \right\| \mathrm{d}\mathbf{P}(\omega) \\
&\leq L_g \epsilon
\end{aligned}$$

where L_g is maximum of the Lipschitz constant of the payoff functions g_1 and g_2. Using the assumption (27) on (u_ϵ, v_ϵ), we get for all $i = 1, 2$ and $\epsilon < 1$:

$$\left|\mathbf{E}\left(g_i(X_T^\eta)\right) - e_i\right| \leq L_g\epsilon + \epsilon^2 \leq (L_g + 1)\epsilon$$

It remains to prove that the trajectories generated by (u_η, v_η) are consistent.

For all $t \in [t_0, T]$, for all $i = 1, 2$, using (28) we get:

$$V_i(t, X_t^\eta) \leq \mathbf{E}\left(V_i(t, X_t^\epsilon)\middle|\mathcal{F}_t^\eta\right) + L_V\epsilon \tag{29}$$

where L_V is maximum of the Lipschitz constant of the value functions V_1 and V_2, and

$$\mathbf{E}\left(g_i(X_T^\epsilon)\middle|\mathcal{F}_t^\eta\right) \leq \mathbf{E}\left(g_i(X_T^\eta)\middle|\mathcal{F}_t^\eta\right) + L_g\epsilon \tag{30}$$

We now have to use the assumptions (27) on (u_ϵ, v_ϵ): if we denote by

$$\Sigma_{it}^\epsilon := \left\{\omega/\ V_i(t, X_t^\epsilon) \leq \mathbf{E}\left(g_i(X_T^\epsilon)\middle|\mathcal{F}_t^\epsilon\right) + \epsilon^2\right\}$$

we know that $\mathbf{P}(\Sigma_{it}^\epsilon) \geq 1 - \epsilon^2$. Then, denoting by K an upper bound of the payoff functions, for all $t \in [t_0, T]$, for all $i = 1, 2$, we get:

$$\begin{aligned} V_i(t, X_t^\epsilon) &\leq \mathbf{E}\left(g_i(X_T^\epsilon)\middle|\mathcal{F}_t^\epsilon\right)\mathbf{1}_{\Sigma_{it}^\epsilon} + K\mathbf{1}_{(\Sigma_{it}^\epsilon)^c} + \epsilon^2 \\ &\leq \mathbf{E}\left(g_i(X_T^\epsilon)\middle|\mathcal{F}_t^\epsilon\right) + K\mathbf{1}_{(\Sigma_{it}^\epsilon)^c} + \epsilon \end{aligned}$$

assuming as usual that the functions g_i are non negative.

Going back to our estimate of $V_i(t, X_t^\eta)$ as computed in (29) and noticing that the filtration $(\mathcal{F}_t^\eta)$ is a subfiltration of $(\mathcal{F}_t^\epsilon)$, we can write:

$$\begin{aligned} V_i(t, X_t^\eta) &\leq \mathbf{E}\left[\mathbf{E}\left(g_i(X_T^\epsilon)\middle|\mathcal{F}_t^\epsilon\right)\middle|\mathcal{F}_t^\eta\right] + \mathbf{E}\left[K\mathbf{1}_{(\Sigma_{it}^\epsilon)^c}\middle|\mathcal{F}_t^\eta\right] + \epsilon + L_V\epsilon \\ &\leq \mathbf{E}\left[g_i(X_T^\epsilon)\middle|\mathcal{F}_t^\eta\right] + K\mathbf{P}\left[(\Sigma_{it}^\epsilon)^c\middle|\mathcal{F}_t^\eta\right] + (L_V + 1)\epsilon \\ &\leq \mathbf{E}\left[g_i(X_T^\eta)\middle|\mathcal{F}_t^\eta\right] + K\mathbf{P}\left[(\Sigma_{it}^\epsilon)^c\middle|\mathcal{F}_t^\eta\right] + (L_V + L_g + 1)\epsilon \text{ due to (30)} \end{aligned}$$

We rewrite this last inequality introducing the constant $C^* = \max(L_V, L_g, 1, K)$:

$$V_i(t, X_t^\eta) \leq \mathbf{E}\left[g_i(X_T^\eta)\middle|\mathcal{F}_t^\eta\right] + C^*\mathbf{P}\left[(\Sigma_{it}^\epsilon)^c\middle|\mathcal{F}_t^\eta\right] + 3C^*\epsilon \tag{31}$$

In order to estimate $\mathbf{P}[(\Sigma_{it}^\epsilon)^c|\mathcal{F}_t^\eta]$, we will use the assumption $\mathbf{P}((\Sigma_{it}^\epsilon)^c) \leq \epsilon^2$:

$$\mathbf{E}\left[\mathbf{P}\left((\Sigma_{it}^\epsilon)^c\middle|\mathcal{F}_t^\eta\right)\right] \leq \epsilon^2 \Rightarrow \mathbf{P}\left\{\mathbf{P}\left((\Sigma_{it}^\epsilon)^c\middle|\mathcal{F}_t^\eta\right) \geq \epsilon\right\} \leq \epsilon$$

This implies for all $t \in [t_0, T]$, for all $i = 1, 2$:

$$\begin{aligned} \mathbf{P}\left\{V_i(t, X_t^\eta) \leq \mathbf{E}\left[g_i(X_T^\eta)\middle|\mathcal{F}_t^\eta\right] + 4C^*\epsilon\right\} &\geq \mathbf{P}\left\{\mathbf{P}((\Sigma_{it}^\epsilon)^c|\mathcal{F}_t^\eta) \leq \epsilon\right\} \\ &\geq 1 - \epsilon \end{aligned}$$

Finally, for all $\epsilon > 0$, we have built finitely many random controls (u_η, v_η) defined on an underlying infinite probability space, that are equivalent to random controls defined on an underlying finite probability space satisfying for $\epsilon < 1$ for $i = 1, 2$:

$$\left|\mathbf{E}[g_i(X_T^\eta)] - e_i\right| \leq 2C^*\epsilon$$

and for all $t \in [t_0, T]$ for $i = 1, 2$:

$$\mathbf{P}\left\{V_i(t, X_t^\eta) \leq \mathbf{E}\left[g_i(X_T^{t_0, x_0, u_\eta, v_\eta})\middle|\mathcal{F}_t^\eta\right] + 4C^*\epsilon\right\} \geq 1 - \epsilon$$

This proves that e is a Nash equilibrium payoff in mixed strategies. □

References

[1] Aumann, R. J. & Maschler, M. B. *Repeated games with incomplete information*, MIT Press, Cambrige, MA (1995)

[2] Basar, T. & Olsder, G. J. *Dynamic noncooperative game theory* , 2nd Edition, Academic Press, London (1995)

[3] Bensoussan, A. & Frehse, J. *Stochastic games for N players*, J. Optimisation Theory Appl. 105, no. 3 (2000) pp. 543-565

[4] Bressan, A. & Shen, W. *Semi-cooperative strategies for differential games*, International Journal of Game Theory, 32 (2004) pp. 561-593

[5] Bressan, A. & Shen, W. *Small BV solutions of hyperbolic noncooperative differential games*, SIAM J. Control Optim. 43, no. 1 (2004) pp. 194-215

[6] Buckdahn, R.; Cardaliaguet, P. & Rainer, C. *Nash equilibrium payoffs for nonzero sum stochastic differential games*, SIAM J. Control Optim. 43, no. 2 (2004) pp. 624-642

[7] Cardaliaguet, P. *Differential games with asymmetric information*, SIAM J. Control Optim. 46, no. 3 (2007) pp. 816-838

[8] Cardaliaguet, P. *Representations formulae for some differential games with asymetric information*, J. Optimisation Theory Appl. 138, no. 1 (2008) pp. 1-16

[9] Cardaliaguet, P. & Quincampoix, M. *Deterministic differential games under probability knowledge of initial condition*, International Game Theory Rev., 10, no. 1 (2008) pp. 1-16

[10] Engwerda, J. C. *LQ dynamic optimization and differential games*, Chichester Wiley, New York (2005)

[11] Friedman, A. *Differential games*, Wiley-Interscience, New York (1971)

[12] Hamadène, S.; Lepeltier, J.-P. & Peng, S. *BSDEs with continuous coefficients and stochastic differential games*, in El Karoui et al. eds, Backward stochastic differential equations, Pitman Res. Notes in Math. Series 364 (1997) pp.161-175 Longman, Harlow

[13] Kleimenov, A.F. *Nonantagonist Differential Games*, (in russian) "Nauka" Uralprime skoje Otdelenie Ekaterinburg (1993)

[14] Kononenko, A.F. *Equilibrium positional strategies in non-antagonistic differential games* (Russian) Dokl. Akad. Nauk SSSR 231 (1976), no. 2, pp.285-288. English translation: Soviet Math. Dokl. 17, no. 6 (1976) pp. 1557-1560 (1977).

[15] Krasovskii, N.N. & Subbotin, A.I. *Game-Theoretical Control Problems*, Springer-Verlag (1988)

[16] Rainer, C. *On two different approaches to nonzero sum stochastic differential games*, Appl. Math. Optim. 56 (2007), pp.131-144

[17] Tolwinski, B.; Haurie, A. & Leitmann, G. *Cooperative equilibria in differential games*, Journal of Mathematical Analysis and Applications 119 (1986) pp. 182-202