Markov Perfect Equilibria in Stochastic Revision Games

Stefano Lovo and Tristan Tomala*

August 20, 2015

Abstract

We introduce the model of Stochastic Revision Games where a finite set of players control a state variable and receive payoffs as a function of the state at a terminal deadline. There is a Poisson clock which dictates when players are called to choose to revise their actions. This paper studies the existence of Markov perfect equilibria in those games. We give an existence proof assuming some form of correlation. We deduce the existence of subgame perfect equilibria (without correlation).

Key words: Stochastic games, Poisson clock, Markov perfect equilibrium.

MSC 2010 Subject Classification: Primary 91A15. Secondary 91A25.

OR/MS subject classification: Games/group decisions, Stochastic.

*HEC Paris and GREGHEC, 78351 Jouy-en-Josas, France. `lovo@hec.fr, tomala@hec.fr`. The authors thank Attila Ambrus, Fabien Gensbittel, Yuichiro Kamada, Yehuda "John" Levy, Sofia Moroni, Jérôme Renault, Thomas Rivera, Sylvain Sorin for useful discussions and comments. The support of the HEC Foundation and of the CNRS (projet PEPS) are gratefully acknowledged.

1 Introduction

We define stochastic revision games to model situations where players' payoffs depend on the value of a state variable at a future deadline. Before the deadline, players prepare actions in order to affect the state. Precisely, players enter the game at an initial time and in a given state. They choose actions and then wait for these actions to affect the state and for payoffs to be distributed. Between the initial time and the deadline, players observe the evolution of the state and may have the opportunity to reconsider and revise their choices. As the deadline gets closer and closer, the probability to get another revision opportunity vanishes and the last choices become more and more binding and relevant for the final state.

Stochastic revision games can be used to model a number of actual situations. An example is the pre-opening phase in stock markets, such as Nasdaq or Euronext for example. Half an hour before the opening of the market, traders can submit orders which can be withdrawn and changed until the opening time. Only the orders recorded in the system at the opening time are binding and payoff relevant. Another example are armies deploying their forces on a battleground. The redeployment time is uncertain and the outcome of the battle depends on the location of each party's forces at the moment of impact. Sports also provide many applications for stochastic revision games. The composition of teams in team games such as football or rugby needs to be made before the match, the choice of the tires and car setting before the race starts. Another sport example is the count-down preceding the start of a sailing race. In this phase each boat maneuvers as to be in the best position and at the highest speed on the starting line when the starting gun is fired. Available trajectories depend on the current direction of wind at the boats' positions.

Stochastic revision games generalize the revision games that have been introduced by Kamada and Kandori, 2013 and further studied in Calcagno et al., 2014. In these papers players receive revision opportunities at the beats of a Poisson clock. At a revision opportunity a player selects an action within a fixed set of actions. Player payoffs depends on the action profile last selected before the deadline. Like in revision games, in stochastic revision games players are allowed to choose actions only at the tick of a random (Poisson) clock, and they only care about the terminal payoff. Like in a stochastic game, there is a state space and the evolution of the state is governed by the actions of the players. A state may include the last prepared action but also some exogenous random factor and may affect the set of players' available actions. The action prepared at revisions affect the evolution of the state. In the sailing count-down example a state is the positioning of each boat and the direction of wind. A boat currently in location A and wishing to reach location B must choose one of many trajectories. The set of available trajectories depends on the current locations of the other boats and the current direction of wind. The choice of a trajectory by a boat might limit the ability of a competitor to maneuver.

The objective of the current paper is to study the existence of Markov perfect equilibria in such games. We assume complete information and full monitoring: states, actions and revision times are publicly observed, payoff functions are fully known. Under these assumptions, we prove the existence of Markov perfect *correlated* equilibria. Namely, there is a correlated equilibrium in strategies that depend only on the current state and the time remaining before the deadline. For this, we assume that players have access to a correlation device that delivers messages at each revision time. We consider both private correlation *à la* Aumann where players receive private messages, and public correlation where players receive messages from a public randomization device.

There is a number of papers which use the revision game framework to study economic models such as bargaining (Ambrus and Lu, 2015) or auctions (Ambrus, Burns and Ishii, 2014; Moroni, 2014). In those papers, as in Kamada and Kandori (2013) and Calcagno et al. (2014), the existence of equilibrium is derived from the specific structures of the models and equilibria are actually characterized. However, those works do not prove an existence result for revision games.

Recently, Moroni (2015) has proved the existence of trembling hand perfect equilibria (and thus of Nash equilibria also) for revision games with imperfect information. There are two important differences with the current paper, one is about the class of games and the second is about the solution concept. About the model, Moroni considers imperfect information: players do not necessarily observe their opponent's moves or revision opportunities. They might also be uncertain about the revision rates of the opponents. By contrast, we assume complete information and perfect observation of states, actions and revision times, but extend the model to stochastic games. We also allow revision rates to depend on the state so that players can influence the speed of arrivals. About the solution concept, we study Markov perfect equilibria, that is we impose strategies to depend only on the current state and on the remaining time before the deadline, whereas Moroni considers strategies that depend on the full history. While both papers imply the existence of Nash equilibria for the original revision games of Kamada and Kandori (2013) and Calcagno et al. (2014), they provide different additional findings.

Our proof strategy is inspired by the literature on existence of stationary equilibria for discounted stochastic games. We take the same view point as Fink (1964) and Takahashi (1964) and consider the bi-variate correspondence from strategies and continuations payoffs which associates: to a strategy profile its continuations payoffs, and to continuation payoffs, the set of strategy profiles that are Nash equilibria for those payoffs. A stationary equilibrium is a fixed point of this correspondence, so the exercise is to check that the Kakutani-Fan-Glicksberg Theorem applies. A key property that we show is that the transition probabilities and payoffs are governed by differential equations. This implies that the set of continuation payoff functions is compact for the strong topology, which is crucial in our proof.

Our results are closely related to the paper of Levy (2013) who considers stochastic games in continuous time with flow and terminal payoffs. Levy also derives differential equations satisfied by transitions and payoffs. His existence proof is based on differential inclusions rather than on fixed points. A key difference with our work is that in Levy's game (following Neyman, 2012), players change actions continuously, only the states jump in discrete time. Also, in our model, transition between states depend on the action played at the revision time, whereas in Neyman and Levy, transition rates depend on the average action (viewed as a mixed action) played over a time interval, so a sudden change of action does not have high impact on the transitions.

The results on stationary equilibria for infinite discounted stochastic games (Mertens and Parthasarathy 1987, 1991, Nowak and Raghavan 1992, Duffie et al. 1994) are also closely related. In particular (as in Levy's paper), it is a common feature of those works to prove existence of some form of correlated equilibria. This is because strategy sets are compact for the weak-$*$ topology and the weak-$*$ limit of a sequence functions with values in a set S, ranges in the convex hull of S. Thus, some form of convexity of the equilibrium concept is to be expected. Note that a revision game can be interpreted as a stochastic game on a continuous state space, interpreting the revision times as part of the state. In all those papers, the transition probability is assumed to be norm-continuous with respect to the state and actions. This assumption is not satisfied for revision games. The distribution at time t of the next revision time is the convex combination of a Dirac (with some probability there is no further revision) and of an absolutely continuous measure on the remaining time interval. For $t \neq t'$, the associated Dirac measures are far away in the norm on probabilities. Yet, they are close-by in the weak-$*$ sense.

Stochastic revision games are related to asynchronous repeated games as in Lagunoff and Matsui (1997). The variant that is closest to our model is when players revision dates follow a

Poisson process. These models differ from ours in two important perspectives. First, the infinite horizon implies that the distribution of the next revision opportunity does not depend on calendar time whereas in stochastic revision games the probability of having another revision opportunity shrinks as the deadline approaches. Second, players receive continuous flow payoffs rather than terminal payoffs. Existence of Markov equilibria in these repeated games is not an issue (see for instance Haller and Laugunoff, 2000). For these games Markov equilibria strategies only depend on the current state and translate into a stationary Markov transition matrix across state (see for instance Ambrus and Ishii (2015) for a study of Markov equilibria in coordination repeated games when players' moves follow asynchronous Poissons processes).

The paper is organized as follows. The model is introduced in Section 2. Section 3 provides a detailed example. The main result and the related mathematical tools for studying Markov perfect equilibria are presented in section 4. In section 5, we show how to construct a subgame perfect equilibrium from a public correlated Markov perfect equilibrium. Some proofs are relegated to the appendix.

2 Model

We consider a stochastic game with finite set of states and actions, played in continuous time over a compact time interval where players maximize their terminal payoff. Actions choices and jumps between states occur only at the beat of a random Poisson clock. We first define the game formally.

2.1 The game

The data of such a game are as follows. To avoid trivialities, all sets considered below are assumed to be non-empty.

- There is a finite set of players N and a finite set of states K.
- For each state k and each player i, there is a set of actions $A_i(k)$ available to player i in state k.
- There is a reward function $r : K \to \mathbb{R}^N$, r_i denoting the reward of player i.
- There is a transition probability $q : K \times A \to \Delta(K)$.
- For each state k, there is a positive number λ_k.
- The time interval is $[0, T]$.

The game is played at the beats of a clock over the interval $[0, T]$. The game $\Gamma(0, k_0)$ starts at time 0 in an initial state k_0, and if at time $t < T$ and the state is k, the game $\Gamma(t, k)$ is played.

For each $t \in [0, T)$, the game $\Gamma(t, k)$ unfolds as follows. First, a random time τ is drawn from the exponential distribution $\mathcal{E}(\lambda_k)$. Let $t' := t + \tau$. If $t' > T$, the game is over, the final state is $k_T = k$ and the payoff is $r(k_T)$. If $t' < T$, then t' is said to be a *revision time* and each player i selects an available action in $A_i(k)$, these choices being simultaneous. Let a be the resulting action profile. Second, a new state k' is selected with probability $q(k'|k, a)$. Then the game $\Gamma(t', k')$ is played. All players observe the current and past *revision times*, the actions selected and the current state. The dynamics are described more formally in Section 2.2.

This model is a generalization of the revision games as defined in Kandori and Kamada (2013) and in Calcagno et al. (2014). In both of these papers, a normal form game is considered. At the beat of the clock, players have opportunities to revise their actions. At the deadline, actions freeze and payoffs are distributed. Kandori and Kamada (2013) consider synchronous games where all players revise their actions at each revision time. Calcagno et al. (2014) consider asynchronous games where each player is equipped with a Poisson clock independent of the clocks of other players, so that almost surely, only one player revises her action at a time.

Let's represent a revision game as a special stochastic revision game. Take finite sets of actions A_i and a payoff function $u : A \to \mathbb{R}^N$. Let $\mathcal{N}$ be the set of non empty subsets of N and let denote with η an element of $\mathcal{N}$. Define $K = A \times \mathcal{N}$, that is, let a state $k = (a, \eta)$ specify a current action profile $a \in A$ for all players and a set of players $\eta \subseteq N$ who are given an opportunity to revise their actions. If at the tick of the clock the current state is $k = (a, \eta)$, then all players in η can choose an action, that is $A_i(k) = A_i$, for $i \in \eta$, and is the singleton a_i for $i \notin \eta$. If player $i \in \eta$ chooses b_i, then the next state is $k' = ((b_i)_{i\in\eta}, (a_i)_{i\notin\eta}, \eta'))$ with probability $\mu(\eta')$. Revision opportunities follow a poisson clock with intensity λ. The reward function is $r(a, \eta) = u(a)$, that is, players actual payoff only depends on the last action profile prepared before time T.

If we let $\mu(N) = 1$, we get a synchronous revision game. If we assume $\mu(\eta) > 0$ if and only if η is a singleton, we get an asynchronous revision game where the revision intensity of player i is $\lambda\mu(\{i\})$.

Note that we make the seemingly restrictive assumption that rewards depend only on states and not directly on actions. This is innocuous since it is possible to augment the state space to $K' = K \times A$ in order to encompass models where payoffs depend directly on actions.

2.2 Strategies and equilibria

Although the game is in continuous time, it admits a natural tree structure. A history is a list of past revision times, states and action profiles ending with the current time. This writes,

$$h = (k_0, t_1, k_1, a_1, \ldots, t_n, k_n, a_n, t).$$

If $t > T$, the history is terminal, the final state k_T is equal to k_n and the payoff is $r(k_T)$. Otherwise, players play the subgame $\Gamma(t, k_n)$.

A (behavioral) strategy for player i is then a mapping from non-terminal histories to mixed actions. It is assumed to be measurable with respect to the natural σ-algebra over histories. Namely, we endow $[0, T]$ with the Borel sets and the set of histories with the product σ-algebra. A profile σ of strategies induces a probability distribution $\mathbb{P}_\sigma$ over histories in the usual way. A Nash equilibrium is a strategy profile such that for each player i and each strategy σ'_i,

$$\mathbb{E}_\sigma r(k_T) \geq \mathbb{E}_{\sigma'_i, \sigma_{-i}} r(k_T).$$

A subgame perfect equilibrium is a strategy profile which is a Nash equilibrium in every subgame $\Gamma(t, k)$.

We are interested in the existence of a Markov Perfect equilibrium, namely a subgame perfect equilibrium that depends only on the current "state" of the system and the remaining time. At a revision time t, the minimal amount of information that players need for playing this game is the current state k and the amount of time $T - t$ remaining before the deadline. So we define a Markov strategy for player i as a measurable mapping $\sigma_i : [0, T] \times K \to \Delta(A_i(k))$. This prescribes a distribution over actions, for each possible current revision time $t < T$ and current state k. A Markov Perfect equilibrium (MPE) is then a profile of Markov strategies which is a subgame perfect equilibrium. Let us provide an alternative definition of MPE.

Let σ be a Markov strategy profile. Consider the expected terminal payoff given that the state is k at time t (denote $\mathbb{E}_\sigma^{k,t}$ the conditional expectation):

$$U_i^\sigma(t,k) = \mathbb{E}_\sigma^{k,t} r_i(k_T).$$

This is the continuation payoff at time t in state k, i.e., in subgame $\Gamma(k,t)$. The mapping $U^\sigma : t \mapsto U^\sigma(t) = (U_i^\sigma(t,k))_{i\in,N,k\in K}$ is *the continuation payoff mapping* associated to σ.

Now, consider a revision time t and let k be the current state. Players have to simultaneously select their actions in $A(k)$. Then the players play the one-shot continuation game with payoff function,

$$u_i^\sigma(t,k,a) := \sum\nolimits_{l\in K} q(l|k,a) U_i^\sigma(t,l).$$

Lemma 2.1 *A profile of Markov strategies σ is a MPE if and only if for every state k and almost every $t \in [0,T]$, $\sigma(t,k)$ is a Nash equilibrium of $u^\sigma(t,k,\cdot)$.*

This property is similar to the one-shot deviation principle valid for subgame perfect equilibria of discounted dynamic games.

Proof. If σ is a MPE, then for each t and k, no player i should have an incentive to deviate, given that no other player is going to deviate in the future. That is, $\sigma(t,k)$ must be a Nash equilibrium of $u^\sigma(t,k,\cdot)$.

Conversely, take a profile of Markov strategies σ which is not subgame perfect. We will show that there exists a set of times with positive Lebesgue measure at which there is a state and a player who has a profitable one-shot deviation.

There exists $\varepsilon > 0$, a time $s < T$, a state k, a player i and a strategy τ_i such that,

$$\mathbb{E}^{k,s}_{\tau_i,\sigma_{-i}} r_i(k_T) > \mathbb{E}^{k,s}_{\sigma} r_i(k_T) + \varepsilon.$$

For each integer n, let τ_i^n be the strategy that coincides with τ_i at the first n revision times and with σ_i at subsequent revisions. The probability of having more than n revisions is at most $\mathbb{P}(X > n)$ where X is a Poisson random variable with intensity $\lambda^* = T \max_k \lambda_k$. Thus, this probability vanishes as n goes to infinity. It follows that for n large enough,

$$|\mathbb{E}^{k,s}_{\tau_i,\sigma_{-i}} r_i(k_T) - \mathbb{E}^{k,s}_{\tau_i^n,\sigma_{-i}} r_i(k_T)| \leq \varepsilon/2,$$

and therefore,

$$\mathbb{E}^{k,s}_{\tau_i^n,\sigma_{-i}} r_i(k_T) > \mathbb{E}^{k,s}_{\sigma} r_i(k_T) + \varepsilon/2.$$

So we deduce the existence of a profitable deviation (by $\varepsilon/2$) which deviates from σ only at the first n revisions. Now assume that,

$$\mathbb{E}^{k,s}_{\tau_i^{n-1},\sigma_{-i}} r_i(k_T) < \mathbb{E}^{k,s}_{\tau_i^n,\sigma_{-i}} r_i(k_T),$$

so that deviating at the first n revisions improves upon deviating only at the first $n-1$ revisions. In particular, this implies that the probability of getting at least n revisions is positive. Consider the set of $(l,t) \in K \times [s,T]$ such that t is the n-th revision time in $\Gamma(s,k)$ and $k_t = l$. From the previous inequality, the probability of this set is positive under $(\tau_i^{n-1}, \sigma_{-i})$. This implies that there exists a set of times t with positive Lebesgue measure, such that on this set there is a state l and a player i who has a profitable one-shot deviation.

If the converse inequality holds,

$$\mathbb{E}^{k,s}_{\tau_i^{n-1},\sigma_{-i}} r_i(k_T) \geq \mathbb{E}^{k,s}_{\tau_i^{n},\sigma_{-i}} r_i(k_T) > \mathbb{E}^{k,s}_{\sigma} r_i(k_T) + \varepsilon/2,$$

we deduce the existence of a profitable deviation which deviates from σ only at the first $n-1$ revisions. We conclude by induction. $\square$

We now define notions of Markov perfect correlated equilibria. A Markov correlated strategy is a measurable mapping $\sigma : [0,T] \times K \to \Delta(\prod_i A_i(k))$.

Definition 2.2 *A Markov correlated strategy σ is a* Markov perfect correlated equilibrium *(MPCE) if for every state k and almost every $t \in [0,T]$, $\sigma(t,k)$ is a correlated equilibrium of $u^\sigma(t,k,\cdot)$.*

A Markov correlated strategy σ is a Markov perfect public correlated equilibrium *(MPPCE) if for every state k and almost every $t \in [0,T]$, $\sigma(t,k)$ is a convex combination of Nash equilibria of $u^\sigma(t,k,\cdot)$.*

Based on Lemma 2.1, we define a MPE as a strategy profile such that the current mixed actions is an equilibrium for the continuation payoffs, at almost every time. We extend this principle to correlated equilibria: at almost every time, the correlated distributions of actions is a correlated equilibrium for the continuation payoffs. This ensures that, at almost all revision times, all players have incentives to play the actions recommended by the correlation device. This is a natural adaptation of the *extensive form correlated equilibrium* of Forges (1986).

Now, a public correlated equilibrium is such that all players receive a public signal before choosing actions. For one-shot games, this boils down to taking convex combinations of Nash equilibria. Alternatively, this can be represented by a public randomisation device which announces draws from the uniform distribution over $[0,1]$ at each beat of the clock.

3 Example: sailing race countdown

To illustrate our model, we consider the following zero-sum stochastic game. Two sailing boats, 1 and 2 are approaching the race start line during the start countdown. The starting line is upwind and at the beginning of the countdown boat 1 is approaching from the right side and boat 2 from the left side. Each boat gains by being the fastest on the start line. During this phase wind can change direction and favour either side. Players have to choose the side of the start line, left or right, and might have opportunities to change side during the countdown. Changing side require the boat to "tack" and this temporarily slows down the boat. Also, a boat ability to change course to the other side depends on the current side of the opponent as illustrated in Figure 1. More formally:

Available actions: There are two independent Poisson clocks with intensity λ_1 and λ_2 for players 1 and 2 respectively. Each boat can tack only at the tick of its own clock: moves are asynchronous.[1] As long as they are on different sides, each player can choose to keep the current side or to tack to the other side. However, when both boats are on the left side, only boat 1 can tack and switch to the right side whereas boat 2 cannot. Symmetrically, as long as both boats are on the right side, then it is player 2 who can switch side whereas player 1 cannot.

At the tick of each of the Poisson clocks and just before the player makes a decision, the wind may change randomly. We assume that the wind favours the right side of the line with probability q. Let $w_t \in \{l, r\}$ denote the side favoured by the wind at time t.

Payoffs: Payoffs depend on the boat positions and on the windy side at the end of the countdown. Overall there are four states $K = \{LL, RL_l, RL_r, RR\}$. If both boats are on the left side

[1] Note that this is equivalent to having a single Poisson clock and one of the dimension of the state specifying the identity of the player who can move in the current state. The two independent poisson clock can then be replicated by assuming that the transition probability over this dimension of the state does not depend on the current state and on players' action.

($k = LL$), then it must be that player 1 has tacked at least one more time than player 2 and so he is behind and have a payoffs is -1. Symmetrically, if both boats are on the right side ($k = RR$), player 1 is ahead and his payoff is 1. If player 1 is on the right side and player 2 on the left side, then the fastest boat is the one positioned on the side favoured by wind and the speed difference will be larger. Namely, if there is more wind on the left side ($k = RL_l$), player 1 payoff's is -2, and it is 2 if the most windy side is the right one ($k = RL_r$).

When matching player 2 on the left side, player 1 is slower than player 2 but he keeps the possibility to tack to the right side. When matched by player 2 on the right side, player 1 is the fastest but he cannot tack to the left side until player 2 has done so. Player 2 faces the symmetric trade-off between speed and ability to choose the windy side. The following figure represents player 1's terminal payoffs and players' available actions.

	L	R
L	-1	
R	-2	1

$w_t = l$

	L	R
L	-1	
R	2	1

$w_t = r$

The following lemma shows that the unique equilibrium of this stochastic revision game consists of two phases. Toward the very end of the countdown, each boat will seek the most windy side to maximize speed. However, to be able to do so, they need to remain manoeuvrable, implying that at the beginning of the countdown players will seek room to manoeuver at the expenses of speed and regardless of the wind conditions.

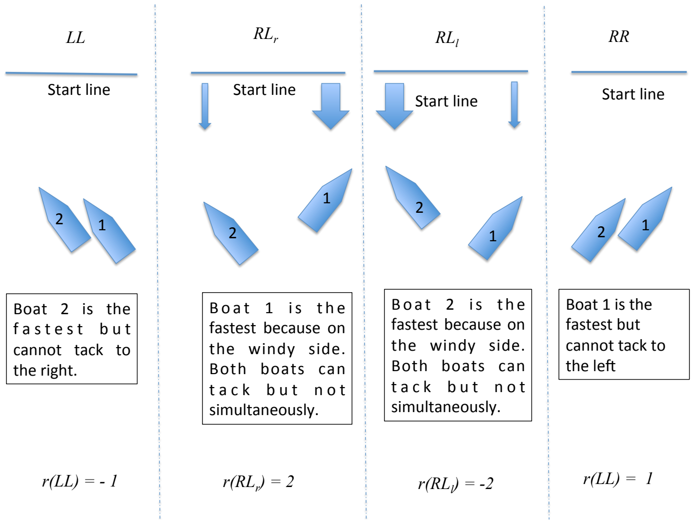

Figure 1: Sailing race

Lemma 3.1 *Assume that* $\lambda_1 = \lambda_2 = 1$ *and* $q = 1/2$*, there is a unique SPE equilibrium. The equilibrium is Markovian, in pure strategies and can be described as follows.*

1. If the first revision opportunity arrives when the remaining time is more than $t^* := 2\ln(4)$*, then the player who can tack will do it irrespectively of* w_t *and lock the other player until* t^**.*

2. If a player has a revision opportunity when the remaining time is less than t^**, then if the player can tack will choose the side that is currently favoured by the wind.*

The formal proof is in the Appendix. The idea is to write a backward differential equation satisfied by the continuation payoffs at equilibrium and to solve it. As is seen in the next section, the differential equation induced by a strategy profile is an important tool.

4 Existence of MPE

We are ready to state our main existence result.

Theorem 4.1 *The game* $\Gamma(0, k_0)$ *admits a Markov perfect (public) correlated equilibrium.*

An interesting class are the games with *perfect information* where in each state k, there is a single active player i such that $A_i(k)$ is not a singleton. That is, in a game with perfect information, only one player at a time can choose an action. Asynchronous revision games have this property. In this case, the continuation game $u^\sigma(t, k, \cdot)$ has a single active player and therefore, the (public) correlated equilibria coincide with the Nash equilibria in mixed strategies. Thus, this theorem implies that stochastic revision games with perfect information admit MPE in mixed strategies.

The proof strategy is simple. On one hand, we map a strategy profile σ to the continuation payoff mapping U^σ. On the other hand, to each continuation payoff mapping V, we associate the set of strategy profiles that play a public correlated equilibrium for the continuation payoff given by V at almost every time. The purpose is to show that this correspondence admits a fixed point, using the Fan-Glicksberg theorem. So we simply need to check that the assumptions of this theorem are satisfied. This is the objective of the remainder of this section.

4.1 Transitions and payoffs

We fix a Markov correlated strategy σ which is a measurable mapping $\sigma : [0, 1] \times K \to \Delta(\prod_i A_i(k))$. We can view it as a measurable function $\sigma : [0, 1] \to \Delta(K \times A)^K$ by setting

$$\sigma(t)[k', a|k] = \sigma(t)[a|k]q(k'|k, a).$$

This is the probability that if at time t the state is k and the players' strategy is σ, then the played action is a and the next state is k'. Let then $\sigma(t)[k'|k] = \sum_a \sigma(t)[k', a|k]$ be the probability of moving from k to k'.

We denote $P^\sigma_{kj}(s, t) = \mathbb{P}_\sigma(k_t = j|k_s = k)$, so that $P^\sigma(s, t) = (P^\sigma_{kj}(s, t))_{kj}$ is the transition matrix from time s to time t. Denote $M^{\sigma}_{kj}(t) = \lambda_k \sigma(t)[j|k] - \lambda_k 1_{\{k=j\}}$.

Lemma 4.2 *The function* $(s,t) \mapsto P^{\sigma}(s,t)$ *satisfies the following differential equation for every* $t \in (0,T]$ *and for almost every* $s \in [0,t)$*:*

$$\forall t \in (0,T],\ \frac{d}{ds}P(s,t) = -M^{\sigma}(s)P(s,t),\ s - a.e.,\ P(t,t) = I$$

where I *denotes the identity matrix.*

Note that this is a *backward* equation. A similar *forward* equation can be obtained.

Proof. The proof method can be seen in Levy (2013), following Miller (1967). Fix $s < t$ and $h > 0$ and write,

$$\mathbb{P}_{\sigma}(k_t = j|k_{s-h} = k) = \sum\nolimits_{l} \mathbb{P}_{\sigma}(k_t = j|k_s = l)\mathbb{P}_{\sigma}(k_s = l|k_{s-h} = k)$$

For $l \neq k$,

$$\mathbb{P}_{\sigma}(k_s = l|k_{s-h} = k) = \int_0^h \lambda_k \exp(-\lambda_k z)\sigma(s-h+z)[l|k]dz$$

For $l = k$,

$$\mathbb{P}_{\sigma}(k_s = k|k_{s-h} = k) = \int_0^h \lambda_k \exp(-\lambda_k z)\sigma(s-h+z)[k|k]dz + \exp(-\lambda_k h)$$

So,

$$\frac{P^{\sigma}_{kj}(s-h,t) - P^{\sigma}_{kj}(s,t)}{h} =$$

$$\sum\nolimits_{l} P^{\sigma}_{lj}(s,t)\frac{1}{h}\int_0^h \lambda_k \exp(-\lambda_k z)\sigma(s-h+z)[l|k]dz - P^{\sigma}_{kj}(s,t)\frac{1-\exp(-\lambda_k h)}{h}$$

From Lebesgue's differentiation theorem, for almost every s,

$$-\frac{d}{ds}P^{\sigma}_{kj}(s,t) = \lim_{h\to 0}\frac{P^{\sigma}_{kj}(s-h,t)-P^{\sigma}_{kj}(s,t)}{h} = \sum\nolimits_l P^{\sigma}_{lj}(s,t)\lambda_k\sigma(s)[l|k] - P^{\sigma}_{kj}(s,t)\lambda_k$$

as desired. $\square$

We deduce from Lemma 4.2 that the mapping $\sigma \mapsto P^{\sigma}$ is continuous for the weak-$*$ topology on the set of strategies. Recall that a sequence of strategies σ_n weakly converges to σ if for any $(a,k) \in A \times K$ and any Lebesgue integrable function g, $\int_0^T \sigma_n(t)[a|k]g(t)dt$ converges to $\int_0^T \sigma(t)[a|k]g(t)dt$.

Lemma 4.3 *If σ_n weakly converges to σ, then for each $0 \leq s < t \leq 1$, $P^{\sigma_n}(s,t)$ converges to $P^{\sigma}(s,t)$.*

Proof. To simplify notations, denote $P^n = P^{\sigma_n}$ and $M^n = M^{\sigma_n}$. Clearly, M^n weakly converges to $M := M^{\sigma}$. From Lemma 4.2, for each $s < t$,

$$P_n(s,t) = I + \int_s^t M_n(z)P_n(z,t)dz.$$

Fix t. The sequence of functions $g_n(z) = M_n(z)P_n(z,t)$ is uniformly bounded and thus admits a weakly converging subsequence and a weak limit g. Along this subsequence, $P_n(s,t) = I + \int_s^t g_n(z)dz$ converges to $I + \int_s^t g(z)dz := Q(s)$ for each s. We want to obtain $Q(s) = P^{\sigma}(s,t)$. To this end, we argue that $g(z) = M(z)Q(z)$ for almost all $z \in [s,t]$, which holds if and only if

$$\int_s^t \phi(z)g(z)dz = \int_s^t \phi(z)M(z)Q(z)dz$$

for all integrable $\phi(\cdot)$. For any such ϕ, $\int_s^t \phi(z)g_n(z)dz$ converges to $\int_s^t \phi(z)g(z)dz$. Now,

$$\int_s^t \phi(z)M_n(z)P_n(z,t)dz = \int_s^t \phi(z)M_n(z)Q(z)dz + \int_s^t \phi(z)M_n(z)(P_n(z,t) - P^\sigma(z,t))dz$$

By weak convergence of M_n, the first term converges to $\int_s^t \phi(z)M(z)Q(z)dz$. For the second term,

$$\left|\int_s^t \phi(z)M_n(z)(P_n(z,t) - Q(z))dz\right| \le C\int_s^t |\phi(z)|.|P_n(z,t) - Q(z)|dz$$

where we have used that $||M_n||$ is bounded by $C = \max_k \lambda_k$. Since $P_n(z,t)$ converges to $Q(z)$ for each z, the RHS converges to 0 from the dominated convergence theorem.

We conclude that $g(z) = M(z)Q(z)$ for almost all $z \in [s,t]$, and thus $I + \int_s^t M(z)Q(z)dz = Q(s)$ for each s. By uniqueness of the solution of the ODE, $Q(s) = P^\sigma(s,t)$ for each s.

So far, we have shown that if σ_n weakly converges to σ, then P^{σ_n} converges pointwise to P^σ *along a subsequence*. If P^{σ_n} does not converges pointwise to P^σ, then we can find $s < t$ and a subsequence $\psi(n)$ such that $P^{\psi(n)}(s,t)$ converges to $R(s,t) \neq P^\sigma(s,t)$. Since $\sigma_{\psi(n)}$ weakly converges to σ, from the previous reasoning, we get that $P^{\psi(n)}$ converges pointwise to P^σ along a sub-subsequence, a contradiction. $\square$

A first consequence we derive is that continuation payoffs are continuous for the weak-$*$ topology.

Corollary 4.4 *If σ_n weakly converges to σ, then for each player i, each state k, and each time t in $[0,T)$, the continuation payoff $U_i^{\sigma_n}(t,k)$ converges to $U_i^\sigma(t,k)$.*

Proof. The expected terminal payoff can be written as,

$$U_i^{\sigma_n}(t,k) = \mathbb{E}_{\sigma_n}^{k,t} r_i(k_T) = \sum\nolimits_l r_i(l) P_{kl}^{\sigma_n}(t,T)$$

thus, convergence follows directly from Lemma 4.3 □

Another important consequence is that the continuation payoff $U_i^\sigma(t,k)$ is a Lipschitz continuous function of time. Precisely, since $P^\sigma(t,T) = I + \int_t^T M^\sigma(z)P^\sigma(z,T)dz$, we can write,

$$U_i^\sigma(t,k) = \sum\nolimits_l P_{kl}^\sigma(t,T)r(l) = r(k) + \int_t^T \langle M^\sigma(z)P^\sigma(z,T), r\rangle dz$$

where $\langle M^\sigma(z)P^\sigma(s,z), r\rangle$ is a shorthand for $\sum_l \sum_m M_{km}^\sigma(z)P_{ml}^\sigma(s,z)r(l)$. Denote $||r|| = \max_{i,k}|r_i(k)|$ the greatest payoff in the game and let $C = \max_k \lambda_k \times ||r||$. We have,

$$\forall (s,t),\ \max_{i,k} |U_i^\sigma(t,k) - U_i^\sigma(s,k)| \leq C|t-s|. \tag{1}$$

We conclude this section by making precise the topologies we consider for strategies and continuation payoffs.

Proposition 4.5 *Let Σ be the set of correlated Markov strategies and let $\mathcal{V}$ be the set of functions $v : [0,T] \to [-||r||, ||r||]^{N\times K}$ such that $\forall (s,t),\ \max_{i,k}|v_i(t,k) - v_i(s,k)| \leq C|t-s|$.*

Then, Σ is convex and compact for the weak-$*$ topology, $\mathcal{V}$ is convex and compact both for the weak-$*$ topology and for the strong topology of uniform convergence. The continuation payoff mapping $\sigma \mapsto U^\sigma$ maps Σ into $\mathcal{V}$.

Proof. Σ and $\mathcal{V}$ are convex sets of uniformly bounded measurable functions defined on $[0,T]$, thus from Alaoglu's Theorem, they are compact for the weak-$*$ topology. Further, the functions in $\mathcal{V}$ are all Lipschitz continuous with the same constant. This set is thus equicontinuous and from Ascoli's Theorem it is compact for the topology of uniform convergence. From Equation 1, $U^\sigma \in \mathcal{V}$. $\square$

4.2 The fixed point

We first prove the existence of a MPCE without the public qualification. Recall that Σ is the set of Markov correlated strategies and $\mathcal{V}$ is the set of continuation payoff functions. We define a correspondence $\Phi : \Sigma \times \mathcal{V} \to \Sigma \times \mathcal{V}$ such that a fixed point is an equilibrium and its associated continuation payoff. The first component is the mapping $\sigma \mapsto U^\sigma$ associating a strategy with its continuation payoff function. The second component is a correspondence $v \mapsto B(v)$ mapping a continuation payoff function to the strategies that are correlated equilibria for the one-shot game induced by v at almost every time.

Precisely, a strategy σ belong to $B(v)$ if for each state k, each player i, each pair of actions a_i, b_i and almost every t,

$$\sum\nolimits_{a_{-i},l} \sigma(t)[a_i, a_{-i}|k]q(l|a_i, a_{-i}, k)v_{i,l}(t) \geq \sum\nolimits_{a_{-i},l} \sigma(t)[a_i, a_{-i}|k]q(l|b_i, a_{-i}, k)v_{i,l}(t). \tag{2}$$

The LHS is the expected payoff of player i when he should play a_i and he does so, the RHS is the expected payoff of player i when he should play a_i and plays b_i instead. In other words, $\sigma(t)[\cdot|k]$ is a correlated equilibrium of the game played at the revision time t in state k, and payoffs are determined by the transition and the function v.

Proposition 4.6 *The correspondence $\Phi : \Sigma \times \mathcal{V} \to \Sigma \times \mathcal{V}$ defined by $\Phi(\sigma, v) = B(v) \times \{U(\sigma)\}$ admits a fixed point.*

Proof. We apply the Fan-Glicksberg fixed point theorem. $\Sigma \times \mathcal{V}$ is convex and compact for the weak-$*$ topology. For each t, the set of $\sigma(t)$ that satisfy Equation (2) for all k, i, a_i, b_i is non-empty and convex, being the set of correlated equilibria of a finite game. The set $B(v)$ is thus convex, and Φ is convex valued.

Claim 4.7 *$\Phi(v)$ is weak-$*$ compact.*

Proof. $\Phi(v)$ is clearly closed for the strong topology of uniform convergence. Since it is convex, it is also closed for the weak convergence. $\square$

Φ is thus convex and compact valued. Now, we prove that Φ has a weakly closed graph. Since $\sigma \mapsto U^\sigma$ is continuous, it is enough to show the following:

Lemma 4.8 *B has a closed graph.*

Proof. Take a sequence (σ^n, v^n) such that for each n, $\sigma^n \in B(v^n)$ assume that (σ^n, v^n) weakly converges to (σ, v).

Fix k, i, a_i, b_i. For each n, we have

$$\sum\nolimits_{a_{-i},l} \sigma^n(t)[a_i, a_{-i}|k]q(l|a_i, a_{-i}, k)v^n_{i,l}(t) \geq \sum\nolimits_{a_{-i},l} \sigma^n(t)[a_i, a_{-i}|k]q(l|b_i, a_{-i}, k)v^n_{i,l}(t),\ t - a.e.$$

This is equivalent to:

$$\int \sum\nolimits_{a_{-i},l} \sigma^n(t)[a_i, a_{-i}|k]q(l|a_i, a_{-i}, k)v^n_{i,l}(t)f(t)dt \geq$$

$$\int \sum\nolimits_{a_{-i},l} \sigma^n(t)[a_i, a_{-i}|k]q(l|b_i, a_{-i}, k)v^n_{i,l}(t)f(t)dt$$

for all bounded measurable $f \geq 0$. To get the desired conclusion, we would like to take limits on both sides. To justify that it is legitimate, we have the following claim.

Claim 4.9 *Let f_n, g_n be two sequences of uniformly bounded functions defined on $[0,T]$. Assume that f_n weakly converges to f and g_n weakly converges to g. Assume also that there exists $C > 0$ such that g_n is C-Lipschitz for each n. Then the product $f_n g_n$ weakly converges to fg along a subsequence.*

Proof. From Ascoli's Theorem, g_n uniformly converges to g along a subsequence. Take ϕ Lebesgue integrable and write,

$$\int \phi(t) f_n(t) g_n(t) dt = \int \phi(t) f_n(t) g(t) dt + \int \phi(t) f_n(t)(g_n(t) - g(t)) dt$$

By weak convergence of f_n, the first term of the RHS converges to $\int \phi(t) f(t) g(t) dt$. The absolute value of the second term is bounded from above by a constant times $\int |\phi(t)|.|g_n(t) - g(t)| dt$ which converges to 0 from the dominated convergence theorem. $\square$

Thanks to this claim, we can take a subsequence and obtain at the limit,

$$\int \sum\nolimits_{a_{-i},l} \sigma(t)[a_i, a_{-i}|k] q(l|a_i, a_{-i}, k) v_{i,l}(t) f(t) dt \geq \int \sum\nolimits_{a_{-i},l} \sigma(t)[a_i, a_{-i}|k] q(l|b_i, a_{-i}, k) v_{i,l}(t) f(t) dt$$

This holds for all k, i, a_i, b_i, so $\sigma \in B(v)$ as desired.

Finally, the correspondence Φ has a fixed point which is a MPCE. $\square$

4.3 Public correlation

In this section, we explain how to adapt the previous proof to public correlation. This follows the argument of Nowak and Raghavan (1992). Implicitly, in the definition of MPCE, at each tick of the clock, there is a correlation mechanism that draws correlatively a recommended action profile and informs each player i of the action he should play. With public correlation, at each tick of the clock, a random number (say uniform in $\Omega = [0,1]$) is drawn and publicly announced, and players choose actions conditional on this number. A Markov public correlated strategy is then a measurable mapping σ from time and Ω into mixed actions. Letting,

$$\sigma(t)[a|k] = \int_0^1 \prod_i \sigma_i(t,\omega)[a|k]d\omega,$$

we see that this induces a particular Markov correlated strategy. The evolution of transition probabilities and payoffs is the same as before.

The only adjustment needed from the previous proof is the definition of the best-reply correspondence $B(v)$. Given a continuation payoff function v, let $B^p(v)$ be the set of Markov correlated strategies such that for almost every t and every state k, $\sigma(t)[\cdot|k]$ is a convex combination of Nash equilibria of the one-shot game with payoff function $u_i(t,k,a) = \sum_l q(l|a,k)v_{i,l}(t)$. Similarly as for $B(v)$, $B^p(v)$ is convex and compact (for the weak-$*$ topology). The only point to check is that its graph is closed. Consider thus a sequence $\sigma^n \in B^p(v^n)$ with $\sigma^n \to \sigma$ and $v^n \to v$. As before, up to extracting a sub-sequence, we can assume that v^n converges to v uniformly. Then, there is a sub-sequence of convex combinations of the σ_n's which converges to σ almost everywhere. Thus, for every t in a set of full measure, we have a $\hat{\sigma}^n(t)$ converging to $\sigma(t)$, which is in the convex hull of the Nash equilibria of the game $u^n(t,k,a)$. We get the desired conclusion by sending n to infinity, by upper-semi-continuity of Nash equilibria.

5 Existence of SPE

We can deduce from Theorem 4.1 the existence of (uncorrelated) subgame perfect equilibria.

Theorem 5.1 *A stochastic revision game $\Gamma(0, k_0)$ admits a subgame perfect equilibrium. In the case of perfect information, there exists a SPE in pure strategies.*

The main argument is as follows. From Theorem 4.1, there exists a Markov perfect public correlated equilibrium. This entails an external public randomization device which sends a public signal at each tick of the clock. The idea is that the players can use the exact value of the next revision time to replicate the public correlation device. The formal proof uses the same arguments as Mertens and Parthasarathy (1991).

Proof. Fix σ a Markov perfect public correlated equilibrium. Suppose that there is a tick of the clock at some time t, that players choose actions according to σ and that the selected state is k. The remaining time interval is $[t, T]$. Let $\tilde{T} = t + \tau$ be the next revision time with τ exponentially distributed with parameter λ_k. Let us write the probability distribution of the action profile at the next revision time, computed at time t in state k, conditional on the event $\{\tilde{T} < T\}$. This is,

$$\mathbb{P}_\sigma(a_{\tilde{T}} = a|t, k) = \int_0^1 \int_t^T \prod_i \sigma_i(z, \omega)[a_i|k] \frac{\lambda_k e^{-\lambda_k z}}{1 - e^{-\lambda_k (T-t)}} dz d\omega$$

where ω is the draw of the public randomization device and $\frac{\lambda_k e^{-\lambda_k z}}{1-e^{-\lambda_k (T-t)}} dz$ is the probability of having a revision between z and $z + dz$, conditional on $\{\tilde{T} < T\}$. We have also for each ω,

$$\mathbb{P}_\sigma(a_{\tilde{T}} = a|t, k, \omega) = \int_t^T \prod_i \sigma_i(z, \omega)[a_i|k] \frac{\lambda_k e^{-\lambda_k z}}{1 - e^{-\lambda_k (T-t)}} dz$$

If we let $D_\sigma(a_{\tilde{T}}|t, k) = (\mathbb{P}_\sigma(a_{\tilde{T}} = a|t, k))_{a \in A(k)}$ be the distribution of the action profile at the next revision time and $D_\sigma(a_{\tilde{T}}|t,k,\omega) = (\mathbb{P}_\sigma(a_{\tilde{T}} = a|t,k,\omega))_{a\in A(k)}$, we have,

$$D_\sigma(a_{\tilde{T}}|t,k) = \int_0^1 D_\sigma(a_{\tilde{T}}|t,k,\omega)d\omega. \tag{3}$$

Remark that for each ω, $\sigma(\cdot,\omega)$ is an uncorrelated strategy profile, so from Equation 3, the distribution $D_\sigma(a_{\tilde{T}}|t,k)$ induced by a public correlated strategy profile is a convex combination of the distributions induced by uncorrelated strategy profiles. Now, the distribution $\frac{\lambda_k e^{-\lambda_k z}}{1-e^{-\lambda_k(T-t)}}dz$ on $[t,T]$ is purely non-atomic, so from Lyapunov's convexity theorem, the set of distributions $D_s(a_{\tilde{T}}|t,k)$ generated by all uncorrelated profiles s is convex. Thus, there exists an uncorrelated strategy profile s such that $D_s(a_{\tilde{T}}|t,k) = D_\sigma(a_{\tilde{T}}|t,k)$. Since the distributions of the next actions are the same under σ and s, the best-replies at time t are also the same, thus changing σ into s remains an equilibrium.

Observe that s is not Markov, as the strategy at time $z \in [t,T]$ depends not only on z but also on t, through the distribution of $\tilde{T}$. So we obtain a SPE but not necessarily a MPE.

Finally, in the case of perfect information, the same argument allows to construct a pure SPE from a mixed MPE. □

6 Appendix: Proof of Lemma 3.1

Proof. There are four payoff relevant states, i.e., $K = \{RL_l, LL, RR, RL_r\}$, the two states where the two boats are on the same side and only one can tack, i.e., $\{LL, RR\}$ and the two states where boat 1 is on the right and boat 2 is on the left, both can tack, and the windy side is either the left or right side, i.e., $\{RL_l, RL_r\}$. The payoff function for player 1 is $\{r(RL_l), r(LL), r(RR), r(RL_r)\} = \{-2,-1,1,2\}$. Because the game is zero-sum and symmetric it is sufficient to focus on the equilibrium strategies of player 1. Moreover because player 1 cannot move in states RR it is sufficient to consider his strategy in the other states.

Take any SPE equilibrium and let σ be the equilibrium strategy, not necessarily Markov. Let $U^{\sigma}(k, h_t)$ denote player 1's expected equilibrium continuation payoff, where t is the remaining time, h_t is the past history and k is the current state. In any equilibrium, if player 1 has a revision opportunity at time t and can tack (i.e. player 2 is on the left side), then his strategy must satisfy:

$$\sigma_1(h_t) \in \arg \max_{x \in \{LL, RL_{w_t}\}} U^{\sigma}(h_t, x) \tag{4}$$

We first characterize the unique equilibrium strategy when the remaining time is small enough. Note that because the state does not change with probability $e^{-2\lambda t}$, we have that

$$r(k)e^{-2\lambda t} - 2(1 - e^{-2\lambda t}) \leq U^{\sigma}(h_t, k) \leq r(k)e^{-2\lambda t} + 2(1 - e^{-2\lambda t})$$

Because $r(RL_l) < r(LL) < r(RR) < r(RL_r)$ it follows from the above inequality that for $t > 0$ close enough to 0, for any strategy σ and history h_t,

$$U^{\sigma}(h_t, RL_l) < U^{\sigma}(h_t, LL) < U^{\sigma}(h_t, RR) < U^{\sigma}(h_t, RL_r).$$

This implies that close enough to the end, in any equilibrium, a player who can tack will choose the side that is currently the most windy. This shows that towards the end, the equilibrium is unique, Markov, pure and as described in point 2. of the Lemma. Given this Markov strategy, the continuation payoffs evolve according to the following backward ODE:

$$
\begin{cases}
-\frac{\partial U^\sigma(t,RL_l)}{\partial t} = \lambda_1(1-q)\left(U^\sigma(t,LL) - U^\sigma(t,RL_l)\right) + \lambda_2 q\left(U^\sigma(t,RR) - U^\sigma(t,RL_l)\right) \\
\\
-\frac{\partial U^\sigma(t,LL)}{\partial t} = \lambda_1(1-q)\left(U^\sigma(t,RL_r) - U^\sigma(t,LL)\right) \\
\\
-\frac{\partial U^\sigma(t,RR)}{\partial t} = \lambda_2 q\left(U^\sigma(t,RL_l) - U^\sigma(t,RR)\right) \\
\\
-\frac{\partial U^\sigma(t,RL_r)}{\partial t} = \lambda_1(1-q)\left(U^\sigma(t,LL) - U^\sigma(t,RL_r)\right) + \lambda_2 q\left(U^\sigma(t,RR) - U^\sigma(t,RL_r)\right)
\end{cases}
\tag{5}
$$

These expression can be interpreted as follows. The instantaneous variation of $U^\sigma(k,t)$ can be written as the instantaneous probability of a change of state, multiplied by the increment in payoffs due to the change of state. Take for example the first differential equation which provides the evolution of player 1's equilibrium payoff in state RL_l as the deadline increases. Suppose that the current state is RL_l and that player 1 has a revision opportunity, an event that arrives with intensity λ_1. If the windy side remains the left side, which occurs with probability $1-q$, then player 1 will tack and the new state will be LL. With intensity λ_2, it is player 2 who has a revision opportunity, and he will tack only if the windy side swings to the right, which occurs with probability q. In this case the new state is RR.

For $\lambda_1 = \lambda_2 = 1$ and $q = 1/2$, the solution of this differential equation together with the terminal conditions $U^\sigma(T,k) = r(k)$ is,

$$\begin{aligned} U^\sigma(RL_l,t) &= -2e^{t-T} \\ U^\sigma(LL,t) &= e^{(t-T)/2}(1-2e^{(t-T)/2}) \\ U^\sigma(RR,t) &= -e^{(t-T)/2}(1-2e^{(t-T)/2}) \\ U^\sigma(RL_r,t) &= 2e^{t-T}. \end{aligned}$$

Note that as long as $T-t \le t^*$ one has $U^\sigma(t,RL_l) < U^\sigma(t,LL) \le U^\sigma(t,RL_r)$ and $U^\sigma(t,RL_l) \le U^\sigma(t,RR) < U^\sigma(t,RL_r)$, with equality for $T-t = t^*$. This implies that when the remaining time is strictly less than t^*, the unique equilibrium of the stochastic game is the one described in point 2.

Now observe that the same strategy cannot be an equilibrium when the remaining time is slightly larger than t^*. If it was, then for $T-t > t^*$ then applying (5), one would have $U^\sigma(t,RL_l) > U^\sigma(t,RR)$ and $U^\sigma(t,RL_r) < U^\sigma(t,LL)$. This contradicts the fact that players prefer to tack to the windy side whenever they can tack. Let consider now the strategy described in point 1. This is clearly optimal at $t = T-t^*$ because

$$U^\sigma(T-t^*,RR) = U^\sigma(T-t^*,RL_l) < U^\sigma(T-t^*,LL) = U^\sigma(T-t^*,RL_r) = -U^\sigma(T-t^*,RR) = -\frac{1}{8}. \tag{6}$$

If the same strategy is used when the remaining time is greater than t^*, then the backward evolution of the continuation payoffs is described by the following ODE system:

$$
\begin{cases}
-\frac{\partial U^\sigma(t,RL_l)}{\partial t} = \lambda_1 \left(U^\sigma(t,LL) - U^\sigma(t,RL_l)\right) + \lambda_2 \left(U^\sigma(t,RR) - U^\sigma(t,RL_l)\right) \\
\\
-\frac{\partial U^\sigma(t,LL)}{\partial t} = 0 \\
\\
-\frac{\partial U^\sigma(t,RR)}{\partial t} = 0 \\
\\
-\frac{\partial U^\sigma(t,RL_r)}{\partial t} = \lambda_1 \left(U^\sigma(t,LL) - U^\sigma(t,RL_r)\right) + \lambda_2 \left(U^\sigma(t,RR) - U^\sigma(t,RL_r)\right)
\end{cases}
\tag{7}
$$

For $\lambda_1 = \lambda_2 = 1$ and $q = 1/2$, using the condition at t^* given by expression (6), one has that for $t < T - t^*$:

$$
\begin{aligned}
U^\sigma(t,RL_l) &= -\frac{1}{8}e^{-2(T-t^*-t)} \\
U^\sigma(t,LL) &= \frac{1}{8} \\
U^\sigma(t,RR) &= -\frac{1}{8} \\
U^\sigma(t,RL_r) &= \frac{1}{8}e^{-2(T-t^*-t)}.
\end{aligned}
$$

For any $t < T - t^*$ one has $U^\sigma(t,RR) < U^\sigma(t,RL_l) < U^\sigma(t,RL_r) < U^\sigma(t,LL)$. This implies that when the remaining time is more than t^*, player 1 (resp. player 2) strictly prefers tacking to the left (resp. to the right) no matter what is the windy side. To see that this is the unique equilibrium, it is sufficient to argue that starting from the payoff at t^*, any strategy which satisfies the equilibrium condition (4) has also to satisfy the ODE, whose solution is unique. $\square$

References

[1] Ambrus, A., J. Burns, and Y. Ishii, "Gradual bidding in ebay-like auctions". Working Paper, pages 1-50, 2014.

[2] Ambrus, A. and S. Lu, "A continuous model of multilateral bargaining with random arrival times". American Economics Journal: Microeconomics, 7, 208-249, 2015.

[3] Ambrus, A. and Y. Ishii, "On Asynchronicity of Moves and Coordination", working paper, 2015.

[4] Calcagno, R., S. Lovo, Y. Kamada and T. Sugaya, "Asynchronicity and Coordination in Common and Opposing Interest Games", Theoretical Economics, 2014, 9 (2) 409–434.

[5] Duffie, D., J. Geanakoplos, A. McLennan and A. Mas-Colell, "Stationary Markov Equilibria", 1994, Econometrica, 62, 745–781.

[6] Fink A.M., "Equilibrium in a Stochastic n-Person Game", 1964, J. Sci. Hiroshima Univ., 28, 89–93.

[7] Forges, F. "An approach to communication equilibria", 1986, Econometrica, **54**, 1375–1385.

[8] Haller, H., R. Lagunoff, 'Genericity and Markovian Behavior in Stochastic Games", 2000, Econometrica, 68, 1231–1248.

[9] Kamada, Y. and M. Kandori, "Revision Games", working paper, 2013.

[10] Lagunoff, R., A. Matsui, "Asynchronous Choice in Repeated Coordiation Games", 1997, Econometrica, 65, 1467–1477.

[11] Levy, J., "Continuous-Time Stochastic Games of Fixed Duration", Dynamic Games and Applications, 2013, 3, 279–312.

[12] Mertens J.F. and T. Parthasarathy, "Equilibria for Discounted Stochastic Games", 1987, CORE Discussion Paper No. 8750.

[13] Mertens J.F. and T. Parthasarathy, "Non Zero Sum Stochastic Games", 1991, In Stochastic Games and Related Topics, Raghavan T.E.S. et al. (eds.), Kluwer Academic Publishers, 145–148.

[14] Miller, B., "Finite State Continuous-Time Markov Decision Pro cesses with Applications to a Class of Optimization Problems in Queueing Theory", 1967, Technical Report 15, Stanford University.

[15] Moroni, S., "Proxy auctions with a deadline". Working Paper, 2014.

[16] Moroni, S., "Existence of trembling hand equilibrium in revision games with imperfect information". Working Paper, 2014.

[17] Neyman, A., "Continuous Time Stochastic Games", 2012, DP 616, Hebrew University, Jerusalem.

[18] Nowak, A.S. and Raghavan, T.E.S., "Existence of Stationary Correlated Equilibria with Symmetric Information for Discounted Stochastic Games", 1992, Mathematics of Operations Research, 17, 519-526.

[19] Takahashi, M., "Equilibrium points of stochastic non-cooperative n-person game", 1964, J. Sci. Hiroshima Univ., 28, 95–99.