21st International Symposium on
Mathematical Theory of Networks and Systems
July 7-11, 2014. Groningen, The Netherlands

# Compact zero-sum stochastic games do not have an asymptotic value

Guillaume Vigeral [1]

***Abstract*— We survey two recent results by the author [34], [33] (one in collaboration with Sylvain Sorin) that give examples of zero-sum stochastic games with four states, compact action sets for each player, and continuous payoff and transition functions, such that the discounted value does not converge as the discount factor tends to 0, and the value of the $n-$stage game does not converge as $n$ goes to infinity. In particular we explain how the oscillations of transition functions can lead to oscillations of the values.**

## I. INTRODUCTION

Two person zero-sum stochastic games have been widely studied since Shapley introduced them in [28]. They model interactions repeated in discrete time between two players with opposite interests. The state of nature evolves as a function of the current state and of the actions chosen by each player, and determines which zero-sum game the players are facing at each time period. Hence, the actions of the players have an influence both on the payoff today and on the law of the state of nature tomorrow.

There are several ways of evaluating the payoff of such a stochastic game. For any integer $n \in \mathbf{N}$, one defines the $n-$stage game for which Player 1 (resp. Player 2) maximizes (resp. minimizes) his average gain on the first $n$ stages. For any $\lambda \in ]0,1]$, one defines the $\lambda$-discounted game for which Player 1 (resp. Player 2) maximizes (resp. minimizes) his $\lambda-$discounted payoff. Some of the main questions in the theory of zero-sum stochastic games are related to the asymptotic behavior of the values of these games as players grow more and more patient, either because the finite horizon of the game grows larger and larger or because the discount factor tends to 0. In particular:

- Does the value of the $n-$stage game converge as $n$ tends to infinity ?
- Does the value of the $\lambda-$discounted game converge as $\lambda$ tends to 0 ?
- Are the two limits equal ?

When the answers to these three questions are positive, the game is said to have an asymptotic value. A nice explanation of why the asymptotic value should exist for games regular enough is the following [30]. An $n-$stage game can be seen as a game played in the time interval $[0,1]$, where the payoff is $\int_0^1 g_t$, and in which the players only moves at time $\frac{k}{n}$. Similarly, in a $\lambda$-discounted game, they only play at time $\lambda$, $\lambda+\lambda(1-\lambda)$, and so on. As $n$ goes to infinity and $\lambda$ goes to 0, these games can thus be viewed as some time discretizations of a hypothetic game played in continuous time on $[0,1]$, and thus the values should converge to the value of this "limit game".

Stochastic games were first studied in the case of a finite number of states and when each player has only finitely many actions. Existence and characterization of the values for a fixed $\lambda$ or $n$ is due to Shapley [28] and relies on von Neumann's minmax theorem [19] as well as Banach's fixed point theorem. In this framework, asymptotic value was established first for recursive [12] and absorbing games [15], then in general (see [7], [8] for the original proof using Tarski-Seidenberg's Theorem, or [21] for a recent proof involving linear programming).

Since minmax theorems also hold true for games with compact action sets and continuous payoffs [29], the values exist [16] for fixed $n$ or $\lambda$ for games with finitely many states, compact action sets for each player, and continuous payoff and transitions (called *compacts games* for simplicity in the following). In this framework, asymptotic value was established for recursive [31], [32] and absorbing [26], [32] games (in the latter case there is also a uniform value [18]), and was conjectured to hold true in general [30].

Let us mention that the existence of an asymptotic value was established in the framework of Markov decision processes and dynamic programming [4], [5], [6], [11], [23]; for games with incomplete information [3], [10], [17], [26]; as well as for some stochastic games with incomplete information [22], [24], [25], [27]. Let us also point out the important link discovered recently between zero-sum stochastic games and tropical geometry, see for example [1], [2], [13], [14].

## II. A ZERO-SUM STOCHASTIC GAME WITH COMPACT ACTION SETS AND NO ASYMPTOTIC VALUE

In the paper [34] we answer by the negative to the conjecture in [30] by constructing a game with four states, compact action sets, and continuous payoff and transitions, whose values do not converge as $n$ tends to infinity or $\lambda$ tends to 0. Surprisingly, it is possible to construct a compact game in which Player 1 can guarantee a payoff of 1 in any $10^{2k}-$stage game, while Player 2 can guarantee a payoff of $-1$ in any $10^{2k+1}-$stage game. The idea of the counterexample is to construct transition functions that are continuous but oscillate infinitely often. These oscillations of the transition functions yield oscillations - and thus divergence - of the values.

We first identify a class of finite zero-sum stochastic games, with two absorbing and two non-absorbing states and action-independent payoff, for which no analytical proofs

*This research was supported by grant ANR-10-BLAN 0112 (France)

[1]Guillaume Vigeral is in CEREMADE, University Paris-Dauphine, Place du Marchal de Lattre de Tassigny 75775 Paris Cedex 16, France vigeral@ceremade.dauphine.fr

ISBN: 978-90-367-6321-9

of convergence of the values are known. We then consider a related class of compact zero-sum stochastic games and prove that there exists games in that class with no asymptotic value. To do this, for any functions $s$ and $d$ from $]0,1]$ to $\mathbf{R}$, we construct a zero-sum stochastic game for which the discounted values in the two non-absorbing states $\omega_+$ and $\omega_-$ are exactly $s(\lambda)+d(\lambda)$ and $s(\lambda)-d(\lambda)$, respectively. We then show that the transition functions of this game are continuous as soon as $s$ and $d$ satisfy some properties, that are weak enough not to imply convergence. Precisely, we prove:

*Proposition 2.1:* Let $s$ and $d$ be two continuously differentiable functions from $]0,1]$ to $\mathbf{R}$. Assume that

- $s$ and $\lambda s'(\lambda)$ are bounded.
- $d$ is nonnegative and there exists $\varepsilon > 0$ such that for all $\lambda \in ]0,1]$,
$$\varepsilon \leq \frac{\lambda d'(\lambda)}{d(\lambda)} \leq 1-\varepsilon.$$

Then there exists a zero-sum stochastic game with four states, compact action sets, continuous payoff and transition functions, such that the discounted values in the two non-absorbing states $\omega_+$ and $\omega_-$ are

$$\begin{aligned} v_\lambda(\omega_+) &= s(\lambda)+d(\lambda) \\ v_\lambda(\omega_-) &= s(\lambda)-d(\lambda) \end{aligned}$$

In particular the choice of $s(\lambda)=\sin(\ln\lambda)$ and $d(\lambda)=\sqrt{\lambda}$ yields a compact zero-sum stochastic game in which the discounted values do not converge.

To construct a compact zero-sum stochastic game in which the finite-horison vales $v_n$ do not converge, we next show, using an idea from [20], that under some additional assumptions we can ensure that $v_n$ and $v_\lambda$ share the same asymptotic behavior.

## III. REVERSIBILITY AND OSCILLATIONS IN ZERO-SUM DISCOUNTED STOCHASTIC GAMES

The previous construction does not allow to construct counterexamples when the game is assumed to have some additional structure, for example perfect observation or finiteness of one of the action sets. In the paper [33] (written with S. Sorin), we show that by coupling two well-behaved exit-time problems one can construct two-person zero-sum stochastic games with finite state space having oscillating discounted values. This unifies and generalizes the examples of the previous section, as well as several ones (including one of a compact game with perfect observation and no asymptotic value) due to Ziliotto [35].

To summarize, we construct a family of zero-sum games in discrete time where the purpose is to control the law of a stopping time of exit. For each evaluation of the stream of outcomes (defined by a probability distribution on the positive integers $n = 1, 2, \ldots$), value and optimal strategies are well defined. In particular for a given discount factor $\lambda \in ]0,1]$ optimal stationary strategies define an inertia rate $Q_\lambda$.

When two such configurations (1 and 2) are coupled this induces a stochastic game where the state will move from one to the other in a way depending on the previous rates $Q^i_\lambda, i = 1,2$.
The main observation is that the discounted value is a function of the ratio $\frac{Q^1_\lambda}{Q^2_\lambda}$ that can oscillate as $\lambda$ goes to 0, when both inertia rates converge to 0. This construction reveals a common structure in the counter-examplesconstructed in [34] and in [35], and permits to construct new examples, for example when one of the players has only finitely many actions.

## IV. CONCLUSION

In all the aforementioned examples, the oscillations of the discounted values were due to oscillations of the transition functions. In a way, these examples are a byproduct of the existence of continous functions oscillating infinitely often such as $x\sin(\ln x)$. A way to ensure convergence of the values may be to forbid the oscillations functions to exhibit such behavior, for example working with semialgebraic transition functions, or more generally with transition functions definable in some o-minimal structures. In [9] we indeed prove that if all the parameters of a zero-sum compact stochastic game are definable in some o-minimal structure, the values converge as soon as the game is either with perfect observation or with finitely many actions or one side. Without one of these additional assumptions the problem is still open.

## REFERENCES

[1] X. Allamigeon, P. Benchimol, S. Gaubert, M. Joswig, Combinatorial simplex algorithms can solve mean payoff games. *Preprint arXiv:1309.5925* (2013)
[2] M. Akian, S. Gaubert, and A. Guterman, Tropical polyhedra are equivalent to mean payoff games. *International Journal of Algebra and Computation* **22** (2012)
[3] R. J. Aumann and M. Maschler with the collaboration of R. E. Stearns, *Repeated Games with Incomplete Information.* MIT Press (1995).
[4] J. Bather, Optimal decision procedures for finite Markov chains. Part I: Examples. *Advances in Applied Probability* **5** (1973), 328-339.
[5] J. Bather, Optimal decision procedures for finite Markov chains. Part II: Communicating systems. *Advances in Applied Probability* **5** (1973), 521-540.
[6] J. Bather, Optimal decision procedures for finite Markov chains. Part III: General convex systems. *Advances in Applied Probability* **5** (1973), 541-553.
[7] T. Bewley and E. Kohlberg, The asymptotic theory of stochastic games. *Mathematics of Operations Research* **1** (1976) 197-208.
[8] T. Bewley and E. Kohlberg, The asymptotic solution of a recursion equation occurring in stochastic games. *Mathematics of Operations Research* **1** (1976) 321- 336.
[9] J. Bolte, S. Gaubert and G. Vigeral, Definable zero-sum stochastic games. *to appear in Mathematics of Operation Research* (2014).
[10] P. Cardaliaguet, R. Laraki and S. Sorin, A Continuous Time Approach for the Asymptotic Value in Two-Person Zero-Sum Repeated Games. *SIAM Journal on Control and Optimization* **50** (2012), 15731596.
[11] E. Dynkin and A. Yushkevich, *Controlled Markov Processes.* Controlled Markov Processes, Springer (1979)
[12] H. Everett, Recursive games. In *Contributions to the Theory of Games, III* (H. W. Kuhn and A.W. Tucker, eds.), Annals of Mathematical Studies 39, Princeton University Press (1957) 47-78.
[13] S. Gaubert and J. Gunawardena, The Perron-Frobenius Theorem for Homogeneous, Monotone Functions. In *Transactions of AMS* **356** (2004), 4931-4950.
[14] S. Gaubert and G. Vigeral, A maximin characterization of the escape rate of nonexpansive mappings in metrically convex spaces. In *Mathematical Proceedings of the Cambridge Philosophical Society* **152** (2012), 341-363.

[15] E. Kohlberg, Repeated games with absorbing states. *Annals of Statistics* **2** (1974), 724-738.
[16] A. Maitra and T. Parthasarathy, On stochastic games. *Journal of Optimization Theory and Applications* **5** (1970), 289-300.
[17] J.-F. Mertens and S. Zamir, The value of two player zero sum repeated games with lack of information on both sides. *International Journal of Game Theory* **1** (1971) 39-64.
[18] J.-F. Mertens, A. Neyman and D. Rosenberg, Absorbing games with compact action spaces. *Mathematics of Operation Research* **34** (2009) 257-262.
[19] J. von Neumann, Zur theorie der gesellschaftsspiele. *Mathematische Annalen* **100** (1928) 295-320.
[20] A. Neyman, Stochastic games and nonexpansive maps. Chapter 26 in A. Neyman and S. Sorin (eds), *Stochastic Games and Applications*, Kluwer Academic Publishers (2003).
[21] M. Oliu-Barton, The asymptotic value in finite stochastic games *to appear in Mathematics of Operation Research* (2014).
[22] J. Renault, The value of Markov chain games with lack of information on one side. *Mathematics of Operation Research* **31** (2006) 490-512.
[23] J. Renault, Uniform value in dynamic programming. *Journal of the European Mathematical Society* **13** (2011) 309-330.
[24] J. Renault, The value of Repeated Games with an informed controller. *Mathematics of Operations Research* **37** (2012) 309-330.
[25] D. Rosenberg, Zero-sum absorbing games with incomplete information on one side: Asymptotic analysis. *SIAM Journal of Control and Optimization* **39** (2000) 208-225.
[26] D. Rosenberg and S. Sorin, An operator approach to zero-sum repeated games. *Israel Journal of Mathematics* **121** (2001) 221-246.
[27] D. Rosenberg and N. Vieille, The maxmin of recursive games with lack of information on one side. *Mathematics of Operations Research* **25** (2000) 23-35.
[28] L. S. Shapley, Stochastic games. *Proc. Nat. Acad. Sciences* **39** (1953) 1095-1100.
[29] M. Sion, On general minimax theorems. *Pacific Journal of Mathematics* **8** (1958) 171-176.
[30] S. Sorin, *A First Course on Zero-Sum Repeated Games.* Springer (2002).
[31] S. Sorin, The operator approach to zero-sum stochastic games. Chapter 27 in A. Neyman and S. Sorin (eds), *Stochastic Games and Applications*, Kluwer Academic Publishers (2003).
[32] S. Sorin and G. Vigeral, Existence of the Limit Value of Two Person Zero-Sum Discounted Repeated Games via Comparison Theorems. Journal of Opimization Theory and Applications **157** (2013) 564-576.
[33] S. Sorin and G. Vigeral, Reversibility and oscillations in zero-sum discounted stochastic games. *Preprint, http://hal.archives-ouvertes.fr/hal-00869656* (2013).
[34] G. Vigeral A zero-sum stochastic game with compact action sets and no asymptotic value, *Dynamic Games and Applications*, **3** (2013) 172-186.
[35] B. Ziliotto, Zero-sum repeated games: counterexamples to the existence of the asymptotic value and the conjecture maxmin=$\lim v_n$. *Preprint, http://hal.archives-ouvertes.fr/hal-00824039* (2013).