MATHEMATICS OF OPERATIONS RESEARCH
Articles in Advance, pp. 1–24
ISSN 0364-765X (print), ISSN 1526-5471 (online)
https://pubsonline.informs.org/journal/moor

Marginal Values of a Stochastic Game

Luc Attia,[a,*] Miquel Oliu-Barton,[a] Raimundo Saona[b]

[a] Université Paris-Dauphine, Paris Sciences and Lettres Research University, Centre National de la Recherche Scientifique, Centre de Recherche en Mathématiques de la Décision, 75775 Paris, France; [b] Institute of Science and Technology Austria, 3400 Klosterneuburg, Austria
*Corresponding author

Contact: luc.attia@dauphine.eu, https://orcid.org/0000-0001-5518-9881 (LA); oliubarton@ceremade.dauphine.fr, https://orcid.org/0000-0003-1793-6350 (MO-B); raimundo.saona@gmail.com, https://orcid.org/0000-0001-5103-038X (RS)

Received: September 25, 2023
Revised: December 28, 2023
Accepted: February 15, 2024
Published Online in Articles in Advance: March 12, 2024

MSC2020 Subject Classifications: Primary: 91A15

https://doi.org/10.1287/moor.2023.0297

Abstract. Zero-sum stochastic games are parameterized by payoffs, transitions, and possibly a discount rate. In this article, we study how the main solution concepts, the discounted and undiscounted values, vary when these parameters are perturbed. We focus on the marginal values, introduced by Mills in 1956 in the context of matrix games—that is, the directional derivatives of the value along any fixed perturbation. We provide a formula for the marginal values of a discounted stochastic game. Further, under mild assumptions on the perturbation, we provide a formula for their limit as the discount rate vanishes and for the marginal values of an undiscounted stochastic game. We also show, via an example, that the two latter differ in general.

Funding: This work was supported by Fondation CFM pour la Recherche; the European Research Council [Grant ERC-CoG-863818 (ForM-SMArt)]; and Agence Nationale de la Recherche [Grant ANR-21-CE40-0020].

Keywords: stochastic games • marginal value • asymptotic value • uniform value

1. Introduction

Introduced by Shapley [26], finite zero-sum stochastic games are a central model in dynamic games and model situations where the environment evolves in response to players' actions in a stationary manner. They are parameterized by a payoff function, a transition function, and possibly a discount rate, which describe, respectively, the current payoff and transition probabilities for each state and each action profile and the probability that the game stops after each stage. The number of states and possible actions at each state are assumed to be finite.

The literature on stochastic games is abundant, both in theory and applications. For this reason, we focus on the finite zero-sum case and present only the results that are most directly connected to our findings. We refer the reader to Solan and Vieille [30] for a summary of the historical context and the impact of Shapley's seminal paper and to Amir [1] for a review of applications, which include resource economics, industrial organization, market games, and empirical economics, among others.

The main solution concept of finite zero-sum stochastic games—henceforth, stochastic games—is the discounted value. Its existence and characterization go back to Shapley [26]. The convergence of the discounted values as the discount rate vanishes was established by Bewley and Kohlberg [6] using the theory of semialgebraic sets. An alternative, probabilistic proof was obtained by Oliu-Barton [18]. The existence of the undiscounted value was established by Mertens and Neyman [15]. The perturbation analysis goes back to Filar and Vrieze [9], who provided a modulus of continuity for the discounted value function in terms of the parameters of the game (i.e., the payoffs, the transition probabilities, and the discount rate). Solan [29] obtained an analogous result for the undiscounted value function under certain conditions on the perturbation of the transition probabilities. A formula for the discounted and undiscounted values was recently obtained by Attia and Oliu-Barton [2], together with algorithms to compute them Oliu-Barton [20] which are polynomial in the number of pure stationary strategies. Robustness results for the undiscounted value function were provided, among others, by Neyman and Sorin [17], Ziliotto [32], Catoni et al. [8], Oliu-Barton [19], and Oliu-Barton [21].

Of particular interest are the directional derivatives of the discounted and undiscounted value functions with respect to a given perturbation, referred to as the *marginal values* of the game. This concept was introduced by Mills [16] in the context of matrix games and linear programming, together with an elegant characterization: for all two matrices A and B of the same size, the directional derivative of the value of A in the direction B—that is, $\lim_{\varepsilon\to 0}\frac{1}{\varepsilon}(\mathrm{val}(A+\varepsilon B)-\mathrm{val}(B))$—is equal to the value of the matrix game B, where players are constrained to play optimal strategies of A. An extension of this result to compact-continuous games—that is, games where the action sets are arbitrary compact metric sets and the payoff and transition functions are continuous—was proved by Rosenberg and Sorin [25]. Notably, this result provides a formula for the marginal value of a stochastic game in the discounted case when only the payoffs are perturbed.

In this article, we investigate the marginal values of stochastic games. Our contribution is three-fold. First, we obtain a formula for the marginal values of a discounted stochastic game, when whichever combination of its defining parameters are perturbed. Further, we show that this formula also holds in the compact-continuous framework—that is, when action sets are compact, and the payoff and transition functions are continuous. Second, we obtain a formula for the limit of the marginal value of a discounted stochastic game as the discount rate vanishes. Further, we show, by example, that the marginal value of an undiscounted stochastic game can exist and differ from the limit of the marginal value of the discounted stochastic game as the discount rate vanishes. Third, we provide a formula for the marginal value of an undiscounted stochastic game, under mild regularity assumptions.

These characterizations provide additional quantitative and qualitative insights into the model of stochastic games. They describe the dependence of the value function on the various parameters of the game, requiring no specific functional form for the payoff or the transition functions. Moreover, the formula for the marginal discounted value implies tractable algorithms for its approximation. Indeed, we show that the complexity of computing an approximation of the marginal value is at most polynomial-time in the number of pure stationary strategies of the game. This complexity coincides with the algorithms described in Oliu-Barton [20] for computing an approximation of the discounted value.

The paper is organized as follows. Section 2 introduces stochastic games and perturbed stochastic games, in both discounted and undiscounted cases. Some useful notation is gathered at the end of this section. Section 3 is devoted to the statements of our main results. Section 4 illustrates our results through two examples of perturbed stochastic games. Section 5 recalls some useful results from the literature, which will be used to prove our results. Finally, Section 6 is devoted to the proofs of our main contributions. Some open questions are provided in Section 7, and some additional material is given in the appendix.

2. Stochastic Games

We start by presenting the standard model of finite two-player, zero-sum stochastic games—henceforth, stochastic games—as introduced by Shapley [26]. Then, we recall the definition and some properties of the auxiliary matrices introduced by Attia and Oliu-Barton [2]. Lastly, we introduce perturbed stochastic games.

2.1. Classical Framework

A stochastic game is described by a tuple $\Gamma = (K, k, I, J; g, q, \lambda)$, where K is a finite set of states; $k \in K$ is the initial state; I and J are the finite action sets, respectively, of players 1 and 2; $g : K \times I \times J \to \mathbb{R}$ is the payoff function; $q : K \times I \times J \to \Delta(K)$ is the transition function; and $\lambda \in [0, 1]$ is the discount rate. We refer to the case $\lambda \in (0, 1]$ as the *discounted case* and $\lambda = 0$ as the *undiscounted case*. Both cases are described below. In the sequel, K, k, I, and J will be fixed, while g, q, λ will be parameters. When useful, we will highlight the parameters of Γ, especially the discount rate λ, by using the notation Γ_λ.

2.1.1. Outline of the Game.

The game proceeds in stages as follows. At each stage $m \geq 1$, both players are informed of the current state $k_m \in K$. Then, independently and simultaneously, player 1 chooses an action $i_m \in I$, and player 2 chooses an action $j_m \in J$. Both players may choose their actions using randomization. The pair (i_m, j_m) is then observed by both players, from which they can infer the stage payoff $g_m = g(k_m, i_m, j_m)$. A new state k_{m+1} is then chosen according to the probability distribution $q(\cdot | k_m, i_m, j_m)$, and the game proceeds to stage $m+1$. A play thus produces a sequence of payoffs $(g_m)_{m \geq 1}$, and the aim of player 1 is to maximize, in expectation,

- $\sum_{m \geq 1} \lambda (1-\lambda)^{m-1} g_m$ in the λ-*discounted case*;
- $\liminf_{L \to \infty} \frac{1}{L} \sum_{m=1}^{L} g_m$ in the *undiscounted case*.

Note that, by the Tauberian theorem for real sequences (Hardy and Littlewood [12]), the undiscounted objective is equal to $\liminf_{\lambda \to 0} \sum_{m \geq 1} \lambda (1-\lambda)^{m-1} g_m$. The game is zero-sum, so player 2 minimizes the same amount.

2.1.2. Strategies.

A *strategy* is a decision rule from the set of possible observations of a player to the set of probabilities over their action set. A strategy for player 1 is thus a sequence of mappings $\sigma = (\sigma_m)_{m \geq 1}$, where $\sigma_m : (K \times I \times J)^{m-1} \times K \to \Delta(I)$. Similarly, a strategy for player 2 is a sequence of mappings $\tau = (\tau_m)_{m \geq 1}$, where $\tau_m : (K \times I \times J)^{m-1} \times K \to \Delta(J)$. The sets of strategies are denoted, respectively, by Σ and $\mathcal{T}$. A *stationary strategy* depends only on the current state, so $x : K \to \Delta(I)$ is a stationary strategy for player 1, and $y : K \to \Delta(J)$ is a stationary strategy for player 2. The sets of stationary strategies are $\Delta(I)^K$ and $\Delta(J)^K$, respectively. A *pure stationary strategy* is deterministic, so the sets of pure stationary strategies are I^K and J^K, respectively.

2.1.3. The Value. For all pairs of strategies $(\sigma,\tau)\in\Sigma\times\mathcal{T}$, the unique probability distribution over the sets of plays $(K\times I\times J)^{\mathbb{N}}$ induced by (σ,τ), the initial state k, and the transition function q is denoted by $\mathbb{P}_{\sigma,\tau}$. The existence and uniqueness of this probability follow from Kolmogorov's extension theorem on the sigma-algebra generated by cylinders. The expectation with respect to this probability is denoted by $\mathbb{E}_{\sigma,\tau}$. Recall that we distinguish discounted and undiscounted stochastic games, depending on whether $\lambda>0$ or $\lambda=0$.

- *Discounted stochastic games.* Let $\Gamma_\lambda=(K,k,I,J;g,q,\lambda)$ be a stochastic game with $\lambda>0$. For all pairs of strategies (σ,τ), set

$$\gamma_\lambda(\sigma,\tau):=\mathbb{E}_{\sigma,\tau}\left[\sum_{m\geq 1}\lambda(1-\lambda)^{m-1}g(k_m,i_m,j_m)\right].$$

By Shapley [26], the discounted stochastic game admits a value, $v(\Gamma_\lambda)$—that is:

$$v(\Gamma_\lambda):=\sup_{\sigma\in\Sigma}\ \inf_{\tau\in\mathcal{T}}\ \gamma_\lambda(\sigma,\tau)=\inf_{\tau\in\mathcal{T}}\ \sup_{\sigma\in\Sigma}\ \gamma_\lambda(\sigma,\tau).$$

Furthermore, the vector of discounted values, as the initial state ranges over the set K, was shown to be the unique fixed point of a contracting map of $\mathbb{R}^K$. This characterization implies, in particular, that both players have optimal stationary strategies, denoted by $O_1(\Gamma)$ and $O_2(\Gamma)$.

- *Undiscounted stochastic games.* Let $\Gamma_0=(K,k,I,J;g,q,\lambda=0)$ be an *undiscounted* stochastic game. For all pairs of strategies (σ,τ), set

$$\gamma_0(\sigma,\tau):=\liminf_{\lambda\to 0}\ \gamma_\lambda(\sigma,\tau).$$

By Mertens and Neyman [15] the undiscounted stochastic games has a value, $v(\Gamma_0)$, that is:

$$v(\Gamma_0):=\sup_{\sigma\in\Sigma}\ \inf_{\tau\in\mathcal{T}}\ \gamma_0(\sigma,\tau)=\inf_{\tau\in\mathcal{T}}\ \sup_{\sigma\in\Sigma}\ \gamma_0(\sigma,\tau).$$

Moreover, the equality $\lim_{\lambda\to 0} v(\Gamma_\lambda)=v(\Gamma_0)$ holds.

2.2. Auxiliary Matrices

We now introduce two auxiliary matrices and a parameterized matrix game that can be associated to the discounted stochastic game $\Gamma_\lambda=(K,k,I,J;g,q,\lambda)$ with $\lambda>0$. Their definition goes as follows.

If the players play a fixed pair of pure stationary strategies $(\mathbf{i},\mathbf{j})\in I^K\times J^K$, then the state follows a Markov chain. Its transition matrix is denoted by $Q(\mathbf{i},\mathbf{j})\in\mathbb{R}^{K\times K}$. Furthermore, the stage payoffs depend only on the current state, so let $g(\mathbf{i},\mathbf{j})\in\mathbb{R}^K$ denote this fixed payoff vector. Let $\rho_\lambda(\mathbf{i},\mathbf{j})\in\mathbb{R}^K$ be the vector of expected λ-discounted payoffs as the initial state ranges over the set of states. In particular, $\rho_\lambda^k(\mathbf{i},\mathbf{j})=\gamma_\lambda(\mathbf{i},\mathbf{j})$ because k is the initial state. By stationarity, $Q(\mathbf{i},\mathbf{j})$, $g(\mathbf{i},\mathbf{j})$ and $r_\lambda(\mathbf{i},\mathbf{j})$ satisfy the following recursive relation:

$$\rho_\lambda(\mathbf{i},\mathbf{j})=\gamma_\lambda(\mathbf{i},\mathbf{j})=\lambda g(\mathbf{i},\mathbf{j})+(1-\lambda)Q(\mathbf{i},\mathbf{j})\rho_\lambda(\mathbf{i},\mathbf{j}).$$

The matrix $\mathrm{Id}-(1-\lambda)Q(\mathbf{i},\mathbf{j})$ is invertible because $Q(\mathbf{i},\mathbf{j})$ is a stochastic matrix and $\lambda\in(0,1]$. Then, by Cramer's rule, for the initial state k, one has:

$$\gamma_\lambda(\mathbf{i},\mathbf{j})=\rho_\lambda^k(\mathbf{i},\mathbf{j})=\frac{\Delta^k(\mathbf{i},\mathbf{j})}{\Delta^0(\mathbf{i},\mathbf{j})},$$

where $\Delta^0(\mathbf{i},\mathbf{j})=\det(\mathrm{Id}-(1-\lambda)Q(\mathbf{i},\mathbf{j}))$ and $\Delta^k(\mathbf{i},\mathbf{j})$ is the determinant of the matrix obtained by replacing the k-th column of $\mathrm{Id}-(1-\lambda)Q(\mathbf{i},\mathbf{j})$ with the vector $\lambda g(\mathbf{i},\mathbf{j})$. Ranging over $(\mathbf{i},\mathbf{j})$, one thus defines the two auxiliary matrices Δ^0 and Δ^k of size $|I^K|\times|J^K|$, whose entries are polynomials in (λ,q) and (λ,g,q), respectively. In addition, every entry of Δ^0 is larger than or equal to $\lambda^{|K|}$. This is the case as for any stochastic matrix $M\in\mathbb{R}^{d\times d}$ and any $r\in(0,1]$, the matrix $S:=\mathrm{Id}-(1-r)M$ satisfies $S_{\ell,\ell}\geq\sum_{\ell'\neq\ell}|S_{\ell,\ell'}|+r$ for every $1\leq\ell\leq d$, which, by Ostrowski [23], implies

$\det(S) \geq r^d$. Finally, for every $z \in \mathbb{R}$, we define the following matrix game:

$$W(z) := \Delta^k - z\,\Delta^0.$$

This parameterized matrix game characterizes the value according to Attia and Oliu-Barton [2, theorem 1]—that is, $v(\Gamma_\lambda)$ is the unique $z \in \mathbb{R}$ satisfying $\mathrm{val}(W(z)) = 0$. It is worth noting that this characterization holds for a fixed initial state and involves matrices of size $|I|^K \times |J|^K$. In contrast, Shapley's [26] approach characterizes the vector of values (that is, for all possible initial positions) via a $|K|$-dimensional system of equations involving matrices of size $I \times J$. Lastly, it is worth noting that the use of these auxiliary matrices was refined in Attia and Oliu-Barton [3] and generalized to the nonzero sum case in Attia and Oliu-Barton [4].

2.3. Perturbed Games and Marginal Value

Consider a stochastic game $\Gamma = (K, k, I, J; g, q, \lambda)$ with $\lambda > 0$ (discounted) or $\lambda = 0$ (undiscounted). An *admissible perturbation* is a triplet $H = (\tilde{g}, \tilde{q}, \tilde{\lambda})$, where

- $\tilde{g} : K \times I \times J \to \mathbb{R}$;
- $\tilde{q} : K \times I \times J \to \mathbb{R}$ and, for all $\varepsilon \geq 0$ sufficiently small, $q + \varepsilon\tilde{q}$ is a transition function;
- For all $\varepsilon \geq 0$ sufficiently small, $\lambda + \varepsilon\tilde{\lambda} \geq 0$.

The *perturbed stochastic game* Γ *in the direction* H, denoted by $\Gamma + \varepsilon H$, is defined as

$$\Gamma + \varepsilon H := (K, k, I, J; g + \varepsilon\tilde{g}, q + \varepsilon\tilde{q}, \lambda + \varepsilon\tilde{\lambda}).$$

Its value is denoted by $v(\Gamma + \varepsilon H)$. Note that the perturbed game $\Gamma + \varepsilon H$ is a discounted stochastic game if $\lambda + \tilde{\lambda} > 0$. In this case, the auxiliary matrices can be defined as in Section 2.2 and are denoted by Δ^0_ε, Δ^k_ε, and $W_\varepsilon(z)$, respectively.

The *marginal value of the stochastic game* Γ *in the perturbation direction* H is then defined as follows, provided that the limit exists:

$$\partial_H v(\Gamma) := \lim_{\varepsilon \to 0^+} \frac{v(\Gamma + \varepsilon H) - v(\Gamma)}{\varepsilon}.$$

This notion was introduced by Mills [16] in the context of matrix games and extended by Rosenberg and Sorin [25] to more general two-player games—that is, with arbitrary compact action sets and a continuous payoff function.

The marginal value of stochastic games has not been characterized yet. However, their existence can be deduced by using the theory of semialgebraic sets, and bounds can be derived from Filar and Vrieze [9, chapter 4] for the discounted case and from Solan [29, theorem 6] for the undiscounted case under mild assumptions. The mild assumptions in Solan [29, theorem 6] consist of $\tilde{\lambda} = 0$ and that $\tilde{q}$ *does not introduce new transitions*—that is, there is no $(\ell, \ell', i, j) \in K^2 \times I \times J$ such that $\ell \neq \ell'$, $q(\ell' \mid \ell, i, j) = 0$, and $\tilde{q}(\ell' \mid \ell, i, j) > 0$. The regularity of the perturbed value function, as well as the existence of the marginal values under mild conditions, is recalled in Section 5.3 for completeness.

2.4. Notation

The following notation is used in the sequel.

- *Simplices.* For any finite set E, the set of probabilities over E is denoted by $\Delta(E)$. Formally,

$$\Delta(E) := \left\{ x : E \to [0,1],\ \sum_{e \in E} x(e) = 1 \right\}.$$

- *Matrix games.* For any matrix $M \in \mathbb{R}^{I \times J}$ and any pair of mixed strategies $(x, y) \in \Delta(I) \times \Delta(J)$, the expected payoff in the matrix game M when players play (x, y) is given by $x^\top M y$. The value of M exists by the minmax theorem and is denoted by

$$\mathrm{val}(M) := \max_{x \in \Delta(I)} \min_{y \in \Delta(J)} x^\top M y.$$

The sets of optimal strategies are denoted by $O_1(M) \subset \Delta(I)$ and $O_2(M) \subset \Delta(J)$, respectively, which are then nonempty, convex, compact sets. Finally, we set $O^*(M) := O_1(M) \times O_2(M)$.

- *Marginal value of a matrix game.* For any pair of matrices $M, N \in \mathbb{R}^{I \times J}$, we consider the matrix game N, in which the players are restricted to play an optimal strategy of M. Its value exists by Sion [28] and is denoted by:

$$\mathrm{val}_{O^*(M)}N = \max_{x\in O_1(M)} \min_{y\in O_2(M)} x^\top N y.$$

- *A canonical set inclusion.* For any two finite sets E and F, the product set $\Delta(E)^F$ can be injected into the simplex $\Delta(E^F)$ via the product measure map as follows:

$$w = (w^1,\dots,w^{|F|}) \in \Delta(E)^F \longmapsto \hat{w} = w^1 \otimes \cdots \otimes w^{|F|} \in \Delta(E^F),$$

where $\otimes$ denotes the direct product. That is, $\hat{w}(e_1,\dots,e_{|F|}) = \prod_{\ell=1}^{|F|} w^\ell(e_\ell)$, for all $(e_1,\dots,e_{|F|}) \in E^F$. The canonical inclusion is strict as soon as $|E| \geq 2$ and $|F| \geq 2$.

- *Strategies in the stochastic game.* For any stochastic game $\Gamma = (K,k,I,J,g,q,\lambda)$ and any stationary strategy $x = (x^\ell)_{\ell\in K} \in \Delta(I)^K$, the canonical inclusion gives $\hat{x} := \otimes_{\ell\in K} x^\ell \in \Delta(I^K)$, which is a probability distribution over their set of pure stationary strategies. The same is true for player 2.
- *An important set of strategies.* Consider a discounted stochastic game Γ. Its auxiliary matrix game is of size $|I^K| \times |J^K|$ and plays an important role in the sequel. We set:

$$\begin{cases} O_1^*(\Gamma) := O_1(W(v(\Gamma))), \\ O_2^*(\Gamma) := O_2(W(v(\Gamma))), \\ O^*(\Gamma) := O_1^*(\Gamma) \times O_2^*(\Gamma). \end{cases}$$

These sets are nonempty, convex, and compact. Moreover, $O_1^*(\Gamma) \subset \Delta(I^K)$ and $O_2^*(\Gamma) \subset \Delta(J^K)$. Hence, the elements of $O^*(\Gamma)$ are probabilities over the sets of pure stationary strategies, rather than (mixed) stationary strategies, of the game Γ. However, by Attia and Oliu-Barton [2], $(\hat{x},\hat{y}) \in O^*(\Gamma)$ for all pairs of optimal stationary strategies $(x,y) \in \Delta(I)^K \times \Delta(J)^K$. In this sense, $O(\Gamma) \subset O^*(\Gamma)$.

- *Perturbed value function.* In the sequel, when a stochastic game Γ and an admissible perturbation H are clear from the context, we use the notation $v_\varepsilon := v(\Gamma + \varepsilon H)$ in general and $v_{\lambda,\varepsilon}$ when the discount rate λ varies.

3. Main Contributions

In the sequel, we consider a stochastic game $\Gamma = (K,k,I,J;g,q,\lambda)$ for some fixed $\lambda \in [0,1]$ and an admissible perturbation $H = (\tilde{g},\tilde{q},\tilde{\lambda})$. Our main contributions are the characterizations of the marginal discounted values, of their limits as the discount rate vanishes, and of the undiscounted marginal values.

3.1. A Formula for the Discounted Marginal Value

Our first result is a characterization of the marginal values of a discounted stochastic game (when $\lambda > 0$). By Section 2.2, to every discounted stochastic game, one can associate auxiliary matrices Δ^0 and Δ^k, so, in particular, this is the case for the perturbed game $\Gamma + \varepsilon H$. We set

$$\begin{cases} \partial_H \Delta^0 := \lim\limits_{\varepsilon\to 0^+} \dfrac{1}{\varepsilon}(\Delta^0_\varepsilon - \Delta^0), \\ \partial_H \Delta^k := \lim\limits_{\varepsilon\to 0^+} \dfrac{1}{\varepsilon}(\Delta^k_\varepsilon - \Delta^k), \end{cases}$$

where the limits exist because, by construction, all entries of Δ^k_ε and Δ^0_ε are polynomial in ε. We can now state our first result.

Theorem 1. *Consider a stochastic game $\Gamma = (K,k,I,J;g,q,\lambda)$ with $\lambda > 0$ and an admissible perturbation $H = (\tilde{g},\tilde{q},\tilde{\lambda})$. Then, $\partial_H v(\Gamma)$ is the unique $z \in \mathbb{R}$, satisfying*

$$D(z) := \mathrm{val}_{O^*(\Gamma)}(\partial_H \Delta^k - v(\Gamma)\,\partial_H \Delta^0 - z\,\Delta^0) = 0.$$

We now discuss the intuition behind this result, its implications regarding the computation complexity of the discounted marginal value, the particular case where only the payoffs are perturbed, and its extension to infinite action sets.

1. **Heuristics.** The formula of Theorem 1 can be interpreted as a derivation under the value operator of the formula $\mathrm{val}(\Delta^k - v(\Gamma)\Delta^0) = 0$ obtained in Attia and Oliu-Barton [2]. Indeed, it follows that $\mathrm{val}(\Delta^k_\varepsilon - v_\varepsilon \Delta^0_\varepsilon) = 0$ for all ε

sufficiently small. If allowed, entry-wise differentiation inside the value operator would yield

$$0 = \frac{\partial}{\partial \varepsilon}(\mathrm{val}(\Delta^k_\varepsilon - v_\varepsilon \Delta^0_\varepsilon))_{|\varepsilon=0} = \mathrm{val}\left(\frac{\partial}{\partial \varepsilon}(\Delta^k_\varepsilon - v_\varepsilon \Delta^0)_{|\varepsilon=0}\right),$$
$$= \mathrm{val}(\partial_H \Delta^k - v(\Gamma)\,\partial_H \Delta^0 - \partial_H v(\Gamma)\,\Delta^0).$$

It thus makes sense to have $z = \partial_H v(\Gamma)$ as the unique solution of $\mathrm{val}(\partial_H \Delta^k - v(\Gamma)\partial_H \Delta^0 - z\Delta^0) = 0$. This heuristic can be turned into a formal statement by restricting the strategy domain in the right-hand-side expressions to $O^*(\Gamma)$.

2. **Computational complexity.** From Theorem 1, one can derive an algorithm to approximate the marginal discounted values, whose computational complexity is polynomial $|I^K|$ and $|J^K|$ for rational data. To do so, we proceed in two steps. First, compute $v(\Gamma)$ using the algorithm from Oliu-Barton [20]. Second, use the map $z \longmapsto D(z)$ to do a dichotomic search and find an approximation of the marginal value. At every step, $D(z)$ is the value of a linear program proposed by Mills [16] and can be computed efficiently by Beling [5]. See Appendix A for more details.

3. **Perturbation of the payoff only.** When only the payoffs are perturbed—that is, for $H = (\tilde{g}, 0, 0)$—Theorem 1 boils down to the following alternative formula, already obtained in Rosenberg and Sorin [25]:

$$\partial_H v(\Gamma) = \mathrm{val}_{O(\Gamma)} \tilde{\gamma}(x, y),$$

where $O(\Gamma) := O_1(\Gamma) \times O_2(\Gamma)$ is the set of optimal stationary strategies of Γ and $\tilde{\gamma}_\lambda$ is the discounted payoff function of the game $\tilde{\Gamma} = (K, k, I, J; \tilde{g}, q, \lambda)$. Indeed, the condition for $\partial_H v(\Gamma)$ simplifies to $\mathrm{val}_{O^*(\Gamma)}(\tilde{\Delta}^k - \partial_H v(\Gamma)\tilde{\Delta}^0) = 0$. Therefore, considering optimal strategies for each player separately, one can reverse the linearization used in constructing the auxiliary matrices and, thus, recover $\tilde{\gamma}(x, y)$.

4. **Extension to the compact-continuous case.** Theorem 1 can be extended to the compact-continuous case—that is, when action sets are compact and the payoff and transition functions are continuous. This extension is proved in Appendix C.

3.2. A Formula for the Limit of the Discounted Marginal Values

Let $\Gamma_\lambda = (K, k, I, J; g, q, \lambda)$ be a discounted stochastic game, where the subindex $\lambda > 0$ is used to better track the variations in λ, while the rest of the parameters remain fixed. Similarly, we denote the auxiliary matrix games by Δ^0_λ and Δ^k_λ.

Theorem 2. *Let Γ_λ be a λ-discounted stochastic game and $H = (\tilde{g}, \tilde{q}, \tilde{\lambda})$ an admissible perturbation. Assume that $(\partial_H v(\Gamma_\lambda))_\lambda$ remains bounded. Then, $\lim_{\lambda\to 0} \partial_H v(\Gamma_\lambda)$ exists and is the unique $w \in \mathbb{R}$, where the following map (over the extended reals) changes sign*:

$$z \in \mathbb{R} \longmapsto F(z) := \lim_{\lambda \to 0} \lambda^{-|K|} \mathrm{val}_{O^*(\Gamma_\lambda)}(\partial_H \Delta^k_\lambda - v_\lambda\, \partial_H \Delta^0_\lambda - z\,\Delta^0_\lambda) \in [-\infty, +\infty].$$

This result calls for several observations, too.

1. **The boundedness assumption.** Sufficient conditions ensuring that $(\partial_H v(\Gamma_\lambda))_\lambda$ remains bounded follow from Solan [29, theorem 6]. For example, it is enough that $\tilde{q}$ *does not introduce new transitions* in the following sense: for all $(i, j) \in I \times J$ and $\ell \neq \ell'$, if $q(\ell' \mid \ell, i, j) = 0$, then $\tilde{q}(\ell' \mid \ell, i, j) = 0$. See Section 5.3 for details.

2. **Applicability.** Similarly to Theorem 1, the present formula suggests a dichotomic algorithm to compute an approximation of the limit of the discounted marginal values, based on the successive computation of the sign of $F(z)$. To compute the latter, we would need to determine an explicit $\lambda_0 > 0$ such that the sign of $F(z)$ is given by the sign of $\mathrm{val}_{O^*(\Gamma_\lambda)}(\partial_H \Delta^k_\lambda - v_\lambda\, \partial_H \Delta^0_\lambda - z\,\Delta^0_\lambda)$ at $\lambda = \lambda_0$, for all the considered z. This approach was followed in Oliu-Barton [20], where the similar map $z \longmapsto \lim_{\lambda\to 0} \lambda^{|K|} \mathrm{val}(\Delta^k_\lambda - z\Delta^0_\lambda)$ was considered to compute the undiscounted value of a stochastic game and where an explicit lower bound for λ_0 was obtained whose size is polynomial in $|I^K|$ and $|J^K|$.

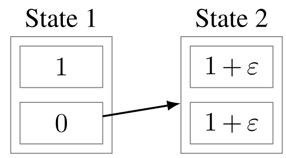
Note, however, that additional difficulties may arise in determining an explicit lower bound for λ_0 in the present case, notably because of the presence of the term v_λ, which is a Puiseux series near zero unlike all other terms, which are polynomials in λ.

3. **Extension.** This result cannot be extended to the compact-continuous case, as explained in Appendix C.

At this stage, a natural question is whether the operators $\lim_{\lambda\to 0}$ and ∂_H commute, so that the limit of the marginal discounted values is equal to the marginal undiscounted value. The answer is no, except in very particular cases. More precisely, the following assertions hold (see Section 6.4 for more details).

i. The operators ∂_H and $\lim_{\lambda\to 0}$ commute when either $|K| = 1$, or $|I| = |J| = 1$ and $H = (\tilde{g}, 0, 0)$.

Figure 1. A perturbed stochastic game with $\lim_{\lambda\to 0} \partial_H v(\Gamma_\lambda) \neq \partial_H v(\Gamma_0)$.



ii. There exists a (minimal) example with $|K| = |I| = 2$, $|J| = 1$ and $H = (\tilde{g}, 0, 0)$, where

$$\lim_{\lambda\to 0} \partial_H v(\Gamma_\lambda) \neq \partial_H \lim_{\lambda\to 0} v(\Gamma_\lambda) = \partial_H v(\Gamma_0).$$

For example, see the perturbed stochastic game in Figure 1, where the arrow indicates a deterministic transition from state 1 to state 2 when the bottom action is played and where there are no other positive transition probabilities. Indeed, for $\lambda \in [0,1]$ and $\varepsilon \geq 0$, a direct computation yields:

$$v_{\lambda,\varepsilon} = \begin{cases} (1-\lambda)(1+\varepsilon) & \text{if } \varepsilon \geq \dfrac{\lambda}{1-\lambda}, \\ 1 & \text{otherwise.} \end{cases}$$

Hence, for all $\lambda > 0$, there exists $\varepsilon_0 > 0$ small enough such that $v_{\lambda,\varepsilon} = 1$ for all $\varepsilon \in [0, \varepsilon_0]$. Consequently,

$$\lim_{\lambda\to 0} \partial_H v(\Gamma_\lambda) = \lim_{\varepsilon\to 0} \frac{v_{\lambda,\varepsilon} - v_{\lambda,0}}{\varepsilon} = 0.$$

On the other hand, for all $\varepsilon \geq 0$, $v_{0,\varepsilon} = \lim_{\lambda\to 0} v_{\lambda,\varepsilon} = 1 + \varepsilon$, so

$$\partial_H v(\Gamma_0) = \lim_{\varepsilon\to 0} \frac{v_{0,\varepsilon} - v_{0,0}}{\varepsilon} = 1.$$

Consequently, both $\lim_{\lambda\to 0} \partial_H v(\Gamma_\lambda)$ and $\partial_H v(\Gamma_0)$ exist, but differ.

In view of the (minimal) example exhibited in (ii), it is worth noting that Theorem 2 characterizes the limit of the marginal discounted values, but not the marginal undiscounted values.

3.3. A Formula for the Marginal Undiscounted Value

We explore the undiscounted case, $\lambda = 0$. We assume that $H = (\tilde{g}, \tilde{q}, 0))$ because for $\tilde{\lambda} > 0$, the marginal undiscounted value may fail to exist. Indeed, Kohlberg [13, p. 120] exhibited an example of a discounted stochastic game where $v(\Gamma_\lambda) = \frac{1-\sqrt{\lambda}}{1-\lambda} = 1 - \sqrt{\lambda} + o(\sqrt{\lambda})$, so that for $H = (0,0,1)$, the marginal undiscounted value $\partial_H v(\Gamma_0) = \lim_{\lambda\to 0} \frac{1}{\lambda}(v_\lambda - v_0)$ exists, but is unbounded.

We introduce a family of perturbed discounted stochastic games and define a finite set of bivariate polynomials denoted $\mathcal{P}_{(\Gamma,H)}$. For every $\beta > 0$ and $\varepsilon \geq 0$ sufficiently small, let $\Gamma_{\beta,\varepsilon} := (K, k, I, J; g + \varepsilon\tilde{g}, q + \varepsilon\tilde{q}, \beta)$ and let $W_{\beta,\varepsilon}(z)$ denote its corresponding parameterized matrix game. For every square submatrix of $W_{\beta,\varepsilon}(z)$, consider the polynomial $P(\beta, \varepsilon, z)$ obtained by taking the determinant of this submatrix. There exist $(R_m)_m \in \mathbb{R}[\varepsilon, z]$ such that $P(\beta,\varepsilon,z) = \sum_{m\geq 0} R_m(\varepsilon,z)\beta^m$, and let $s \geq 0$ be the smallest integer such that $R_s \neq 0$. Define, then, the projection $\Phi(P) := R_s \in \mathbb{R}[\varepsilon, z]$. Running over all possible square submatrices of $W_{\beta,\varepsilon}(z)$, one thus defines the desired finite set of polynomials:

$$\mathcal{P}_{(\Gamma,H)} := \{\Phi(P) : P(\beta,\varepsilon,z) = \det(\overline{W}_{\beta,\varepsilon}(z)),\ \overline{W}_{\beta,\varepsilon}(z) \text{ square sub-matrix of } W_{\beta,\varepsilon}(z)\}.$$

Our next results will justify the introduction of this set.

Proposition 1. *Let $\Gamma = (K, k, I, J; g, q, 0)$ be an undiscounted stochastic game, and let $H = (\tilde{g}, \tilde{q}, 0)$ be an admissible perturbation (not perturbing the discount rate). Then, there exists $\varepsilon_0 > 0$ and $P \in \mathcal{P}_{(\Gamma,H)}$ such that*

$$P \not\equiv 0, \quad \text{and} \quad P(\varepsilon, v(\Gamma + \varepsilon H)) = 0, \quad \forall \varepsilon \in (0, \varepsilon_0].$$

This result may call for some clarifications, notably as the existence of *some* polynomial P satisfying $P(\varepsilon, v_\varepsilon) = 0$ for all ε near zero follows from the theory of semialgebraic sets (see Section 5.2 for more details).

1. **Novelty of the result.** The novelty of Proposition 1 is the identification of a finite, computable set of candidates for the polynomial P. This step is crucial to obtain an explicit formula for the undiscounted marginal values in Theorem 3.

2. **Computability of P.** Although, in general, a suitable polynomial P may be hard to find, information about (the support of) optimal stationary strategies in the auxiliary discounted stochastic games $\Gamma_{\beta,\varepsilon}$ greatly simplifies the search. For example, if both players have a unique optimal stationary strategy whose support is constant for all (ε, λ) sufficiently small, and of equal size, then it is enough to define the submatrix $\overline{W}_{\beta,\varepsilon}(z)$ over these supports and set $P(\varepsilon, z) = \Phi(\det(\overline{W}_{\beta,\varepsilon}(z))$. This result, proved in Oliu-Barton and Vigeral [22, lemma 6], gives a practical way to determine P in (generic) applications.

Our last result is an explicit formula for the undiscounted marginal values, relying on the set of polynomials identified in Proposition 1.

Theorem 3. *Let $\Gamma = (K, k, I, J; g, q, 0)$ be an undiscounted stochastic game and $H = (\tilde{g}, \tilde{q}, 0)$. Let $P \in \mathbb{R}[\varepsilon, z]$ be a polynomial satisfying $P(\varepsilon, v_\varepsilon) = 0$ for all $\varepsilon > 0$ sufficiently small (which can be determined using Proposition 1). Suppose that $\frac{\partial P}{\partial z}(0, v(\Gamma)) \neq 0$ and that $\varepsilon \longmapsto v(\Gamma + \varepsilon H)$ is continuous at $\varepsilon = 0$. Then, the marginal undiscounted value exists and is given by*

$$\partial_H v(\Gamma) = -\frac{\frac{\partial P}{\partial \varepsilon}(0, v(\Gamma))}{\frac{\partial P}{\partial z}(0, v(\Gamma))}.$$

Note that Theorem 3 requires the continuity of $\varepsilon \longmapsto v(\Gamma + \varepsilon H)$ at zero. Like for Theorem 2, a simple sufficient condition ensuring this property is that $\tilde{q}$ does not introduce new transitions.

4. Examples

We now illustrate our contributions via two known examples. The first one is a perturbed Big Match; the second is a perturbed version of a game introduced by Bewley and Kohlberg [7]. For both examples, we will compute the discounted marginal values, their limit as the discount rate goes to zero, and the undiscounted marginal values, using Theorems 1–3.

4.1. The Big Match

We start with the classical Big Match introduced by Gillette [11]. This is a three-state stochastic game, where two states are absorbing with payoffs of zero and one, respectively. The Big Match is thus an absorbing game. It can be represented by Figure 2, where a^* indicates a stage payoff of a followed by a deterministic transition to an absorbing state with payoff a.

For all $\lambda \in (0, 1]$, we denote by Γ_λ the λ-discounted Big Match with initial state $k = 1$. Its value satisfies $v_\lambda = 1/2$ for all λ.

4.1.1. The Auxiliary Matrices. Because the game is absorbing, the set of pure stationary strategies can be identified with the set $I \times J$. Consequently, up to removing redundant rows and columns, the auxiliary matrices are of size $I \times J$. Let states 2 and 3 be, respectively, absorbing states with payoff one and zero. We follow Section 2.2 to compute the auxiliary matrices, denoted by Δ_λ^0, Δ_λ^k, and the parameterized matrix game $W_\lambda(z) := \Delta_\lambda^k - z\Delta_\lambda^0$. Set $I = \{T, B\}$ and $J = \{L, R\}$, where T, B, L, and R refer to "top," "bottom," "left," and "right," respectively. When the pure stationary strategy (T, L) is played, every time states 1, 2, and 3 are visited, the stage payoffs are $g(1, T, L) = 1$, $g(2, T, L) = 1$, and $g(3, T, L) = 0$, respectively. On the other hand, the transitions are δ_2 in states 1 and 2 and δ_3 in state 3. Hence, in matrix form, one has:

$$g(T, L) = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \quad \text{and} \quad Q(T, L) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Figure 2. The Big Match.

State 1

1^*	0^*
0	1

Recall that $\Delta_\lambda^0(T,L) = \det(\mathrm{Id} - (1-\lambda)Q(T,L))$ and $\Delta_\lambda^k(I,J)$ is the determinant of the matrix obtained by replacing the k-th column of $\mathrm{Id} - (1-\lambda)Q(T,L)$ with $\lambda g(T,L)$. Hence, for all $\lambda \in (0,1]$,

$$\begin{cases} \Delta_\lambda^0(T,L) & = \det\begin{pmatrix} 1 & -(1-\lambda) & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix} = \lambda^2, \\ \Delta_\lambda^k(T,L) & = \det\begin{pmatrix} \lambda & -(1-\lambda) & 0 \\ \lambda & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix} = \lambda^2. \end{cases}$$

Hence, $W_\lambda(z)(T,L) = \lambda^2(1-z)$ for all $z \in \mathbb{R}$. Computations for the other action profiles are similar. Overall, one obtains:

$$\Delta_\lambda^0 = \lambda^2\begin{pmatrix} 1 & 1 \\ \lambda & \lambda \end{pmatrix}, \ \Delta_\lambda^k = \lambda^2\begin{pmatrix} 1 & 0 \\ 0 & \lambda \end{pmatrix}, \quad \text{and} \quad W_\lambda(z) = \lambda^2\begin{pmatrix} 1-z & -z \\ -\lambda z & \lambda(1-z) \end{pmatrix}.$$

In particular, for $z = v_\lambda = 1/2$,

$$\mathrm{val}(W_\lambda(1/2)) = \lambda^2 \ \mathrm{val}\begin{pmatrix} 1/2 & -1/2 \\ -\lambda/2 & \lambda/2 \end{pmatrix} = 0.$$

Like for the Big Match itself, the auxiliary game $W_\lambda(1/2)$ has a unique pair of optimal strategies, $x_\lambda = \left(\frac{\lambda}{1+\lambda}, \frac{1}{1+\lambda}\right)$ and $y_\lambda = \left(\frac{1}{2}, \frac{1}{2}\right)$. This is not a coincidence: for any discounted absorbing game Γ, the set of optimal stationary strategies of the game coincides with $O^*(\Gamma)$.

4.1.2. The Marginal Discounted Values and Their Limit. As already argued, $O^*(\Gamma_\lambda) = O(W_\lambda(v_\lambda)) = \{(x_\lambda, y_\lambda)\}$ is a singleton, so the value operator $\mathrm{val}_{O^*(\Gamma_\lambda)}$ is trivialized. By Theorem 1, for any perturbation $H = (\tilde{g}, \tilde{q}, \tilde{\lambda})$, the marginal value is thus the unique $z \in \mathbb{R}$ satisfying

$$D_\lambda(z) := x_\lambda^\top(\partial_H \Delta_\lambda^k - v_\lambda \,\partial_H \Delta_\lambda^0 - z\Delta_\lambda^0)y_\lambda = 0.$$

Or, equivalently, the marginal discounted values satisfy the following explicit formula:

$$\partial_H v_\lambda = \frac{x_\lambda^\top(\partial_H \Delta_\lambda^k - v_\lambda \partial_H \Delta_\lambda^0)y_\lambda}{x_\lambda^\top \Delta_\lambda^0 y_\lambda}. \tag{1}$$

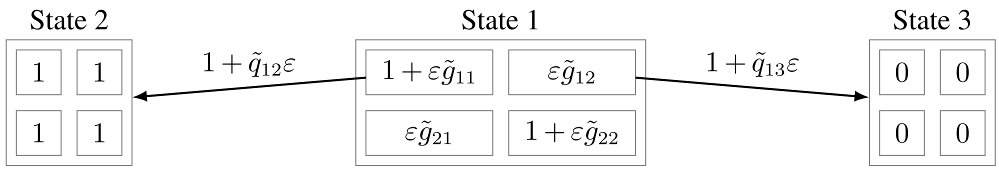
The computation of the marginal discounted values is thus straightforward for any perturbation H, as it boils down to computing the matrices $\partial_H \Delta_\lambda^k$ and $\partial_H \Delta_0$. Similarly, their limit can be obtained directly from Theorem 2, where, again, the value operator is trivialized.

Remark 1. The previous explicit Equation (1) holds for every discounted absorbing game where both players have a unique pair of optimal stationary strategies, denoted by (x_λ, y_λ).

4.1.3. The Undiscounted Marginal Values. Consider now a perturbed version of the Big Match, for an admissible perturbation $H = (\tilde{g}, \tilde{q}, 0)$. The assumption $\tilde{\lambda} = 0$ is not needed here to ensure the existence of the marginal values, but is kept for simplicity and to apply Theorem 3. The perturbed game is illustrated in Figure 3, where the arrows indicate transitions from state 1 to states 2 and 3, respectively, with the indicated probabilities.

More precisely, we consider an arbitrary perturbation of the stage payoffs at state 1 and an arbitrary perturbation of the two positive transition probabilities of the game. To be more precise of the latter, we consider $\tilde{q} \equiv 0$, except

Figure 3. The perturbed Big Match.



for $\tilde{q}_{12} := \tilde{q}(2|1,T,L) \leq 0$ and $\tilde{q}_{12} := \tilde{q}(2|1,T,L) \leq 0$, and also $\tilde{q}(1|1,T,L) = -\tilde{q}_{12}$ and $\tilde{q}(3|1,T,L) = -\tilde{q}_{13}$, so that $q + \varepsilon\tilde{q}$ is a transition probability for all ε sufficiently small.

Let v_0 and v_ε denote the value of the undiscounted Big Match and its perturbed version, respectively. To apply Theorem 3, we proceed as follows. First, we check the continuity of $\varepsilon \longmapsto v_\varepsilon$ at zero. Second, we find a polynomial P such that $P(\varepsilon, v_\varepsilon) = 0$ for all ε sufficiently small. Lastly, we check $\frac{\partial P}{\partial z}(0, v_0) \neq 0$.

Step 1. *Regularity of the perturbed value functions.* Note that the perturbation $\tilde{q}$ does not introduce new transitions. The continuity of $\varepsilon \longmapsto v_\varepsilon$ at $\varepsilon = 0$ follows then directly from Solan [29, theorem 6]. See also Section 5.3 for more details.

Step 2. *Determine the desired polynomial.* By Proposition 1, we have a finite set of candidates, $\mathcal{P}_{(\Gamma,H)}$, obtained by applying the operator $\Phi : \mathbb{R}[\beta, \varepsilon, z] \to \mathbb{R}[\varepsilon, z]$ to the determinants of all possible square submatrices of $W_{\beta,\varepsilon}(z)$, where $\beta > 0$ is an auxiliary discount rate. Because this matrix is of size 2×2, the set $\mathcal{P}_{(\Gamma,H)}$ contains five polynomials, the four corresponding to the entries of the matrix, denoted by $P_{T,L}, P_{T,R}, P_{B,L}$ and $P_{B,R}$, and the one corresponding to the entire matrix, denoted by P. To compute them explicitly, one needs to find $W_{\beta,\varepsilon}(z)$. We proceed as for the classical Big Match, one strategy profile at a time, to obtain the auxiliary matrices for all $\beta > 0$ and $\varepsilon \geq 0$:

$$\begin{cases} \Delta^0_{\beta,\varepsilon} = \beta^2 \begin{pmatrix} 1 + (1-\beta)\tilde{q}_{12}\varepsilon & 1 + (1-\beta)\tilde{q}_{13}\varepsilon \\ \beta & \beta \end{pmatrix}, \\ \Delta^k_{\beta,\varepsilon} = \beta^2 \begin{pmatrix} 1 + \beta\varepsilon(\tilde{g}_{11} - \tilde{q}_{12}) + \varepsilon\tilde{q}_{12} & \beta\varepsilon\tilde{g}_{12} \\ \beta\varepsilon\tilde{g}_{21} & \beta(1 + \varepsilon\tilde{g}_{22}) \end{pmatrix}. \end{cases}$$

Set then $W_{\beta,\varepsilon}(z) := \Delta^k_{\beta,\varepsilon} - z\Delta^0_{\beta,\varepsilon}$ for all $z \in \mathbb{R}$. The set $\mathcal{P}_{(\Gamma,H)}$ is then composed of the following five polynomials:

$$\begin{cases} P_{T,L}(\varepsilon, z) = \Phi(W_{\beta,\varepsilon}(z)(T,L)) = 1 - z + \varepsilon\tilde{q}_{12}(1-z), \\ P_{T,R}(\varepsilon, z) = \Phi(W_{\beta,\varepsilon}(z)(T,R)) = -z + \varepsilon\tilde{q}_{13}z, \\ P_{B,L}(\varepsilon, z) = \Phi(W_{\beta,\varepsilon}(z)(B,L)) = -z + \varepsilon\tilde{g}_{12}, \\ P_{B,R}(\varepsilon, z) = \Phi(W_{\beta,\varepsilon}(z)(B,R)) = 1 - z + \varepsilon\tilde{g}_{22}, \\ P(\varepsilon, z) = \Phi(\det(W_{\beta,\varepsilon}(z)) = 1 - 2z + \varepsilon(\tilde{g}_{22}(1-z) + \tilde{g}_{21}z + \tilde{q}_{12}(1-z)^2 - \tilde{q}_{13}z^2) + o(\varepsilon). \end{cases}$$

Note that we have abbreviated P by grouping the multiple of ε^2, as these terms do not matter for the sequel. To determine the appropriate polynomial, we use the practical remark after Theorem 3. For every $\beta > 0$ and $\varepsilon \geq 0$, the optimal strategy of player 2 in the perturbed discounted Big Match, $\Gamma_{\beta,\varepsilon}$, has full support, by continuity. Hence, as already noted, this property implies that P is the desired polynomial by Oliu-Barton and Vigeral [22, lemma 6].

Step 3. *A formula for the undiscounted marginal values.* One easily checks that $\frac{\partial P}{\partial z}(0, v_0) = -2 \neq 0$ for all the considered perturbations $(\tilde{g}, \tilde{q})$. Theorem 3 can thus be applied to obtain the desired formula:

$$\partial_H v_0 = -\frac{\frac{\partial P}{\partial \varepsilon}(0, v_0)}{\frac{\partial P}{\partial z}(0, v_0)} = \frac{\tilde{q}_{12} - \tilde{q}_{13}}{8} + \frac{\tilde{g}_{21} + \tilde{g}_{22}}{4}.$$

4.1.4. A Final Remark. Consider the perturbation $H = (\tilde{g}, \tilde{q}, 0)$ described in Figure 3: do the undiscounted marginal values and the limit discounted marginal values coincide in this example? From the expressions of $\Delta^0_{\beta,\varepsilon}$ and $\Delta^k_{\beta,\varepsilon}$, one easily determines $\partial_H \Delta^0_\lambda$ and $\partial_H \Delta^0_\lambda$ by setting $\beta := \lambda$ and then taking the entry-wise derivatives in ε at zero. Consequently,

$$\partial_H \Delta^0_\lambda = \lambda^2 \begin{pmatrix} (1-\lambda)\tilde{q}_{12} & (1-\lambda)\tilde{q}_{13} \\ 0 & 0 \end{pmatrix},$$

$$\partial_H \Delta^k_\lambda = \lambda^2 \begin{pmatrix} \lambda(\tilde{g}_{11} - \tilde{q}_{12}) + \tilde{q}_{12} & \lambda\tilde{g}_{12} \\ \lambda\tilde{g}_{21} & \lambda\tilde{g}_{22} \end{pmatrix}.$$

Replacing these matrices in (1), together with (x_λ, y_λ) and v_λ, one thus obtains

$$\partial_H v_\lambda = \frac{(1-\lambda)(\tilde{q}_{12} - \tilde{q}_{13})}{8} + \frac{\lambda \tilde{g}_{11} + \lambda \tilde{g}_{12} + \tilde{g}_{21} + \tilde{g}_{22}}{4}.$$

Therefore, the equality $\partial_H v_0 = \lim_{\lambda \to 0} \partial_H v_\lambda$ holds for this example.

4.2. The Bewley-Kohlberg Stochastic Game

Our next example is a (nonabsorbing) stochastic game introduced by (Bewley and Kohlberg [7, p. 120], which was also considered in Vigeral [31]. A representation of this game is given in Figure 4, where $a, b \geq 0$ are fixed parameters:

The Bewley-Kohlberg stochastic game has four states, two of which are absorbing with payoffs 1 and -1. We label these states as states 3 and 4, respectively. In Figure 4, the arrows indicate a deterministic transition from state 1 to state 2, and vice versa, and a stage payoff of zero. The initial state is fixed to state $k = 1$. Let v_λ denote the value of the λ-discounted game, and let $v_0 := \lim_{\lambda \to 0} v_\lambda$ be the undiscounted value. From the literature, we know that, for each $\lambda \in (0, 1]$, both players have a unique optimal stationary strategy in the λ-discounted game, which is of full support. We also know that $v_0 = \frac{a-b}{a+b+2}$.

4.2.1. The Marginal Discounted Values. For every admissible perturbation $H = (\tilde{g}, \tilde{q}, \tilde{\lambda})$, the marginal discounted values can be obtained from Theorem 1—for example, using the algorithm we described after this statement to obtain an approximation of $\partial_H v$ to a desired level of accuracy. Similarly, their limit as λ goes to zero can be derived from Theorem 2.

On the other hand, it is worth noting that an explicit expression like (1) cannot be obtained here, even when both players have a unique optimal stationary strategy for all $\lambda > 0$. This is because the game is not absorbing, and, thus, $O^*(\Gamma_\lambda)$ contains, but may not be equal to, the set of optimal stationary strategies of Γ_λ.

4.2.2. The Marginal Undiscounted Values. We now focus on the undiscounted marginal values for a given admissible perturbation $H = (\tilde{g}, \tilde{q}, 0)$. For the sake of simplicity, we assume that some of the entries of $\tilde{g}$ and $\tilde{q}$ are zero; namely, we consider four perturbation parameters, denoted by $(\tilde{g}_1, \tilde{g}_2) \in \mathbb{R}^2$ and $(\tilde{q}_1, \tilde{q}_2) \in \mathbb{R}^2_+$ and set:

- $\tilde{g}(1, T, L) := \varepsilon \tilde{g}_1$ and $\tilde{g}(2, T, L) := \varepsilon \tilde{g}_2$;
- $\tilde{q}(1 | 1, B, R) := \varepsilon \tilde{q}_1$ and $\tilde{q}(3 | 1, B, R) := -\varepsilon \tilde{q}_1$;
- $\tilde{q}(2 | 2, B, R) := \varepsilon \tilde{q}_2$ and $\tilde{q}(4 | 2, B, R) := -\varepsilon \tilde{q}_2$;
- All other entries of $\tilde{g}$ and $\tilde{q}$ are equal to zero.

In other words, the perturbations consist in replacing the stage payoffs a and $-b$ with $a + \varepsilon \tilde{g}_1$ and $-b + \varepsilon \tilde{g}_2$, respectively, and the deterministic absorption probabilities with $1 - \varepsilon \tilde{q}_1$ and $1 - \varepsilon \tilde{q}_2$. This perturbed Bewley-Kolhberg stochastic game is illustrated in Figure 5.

The value of this perturbed undiscounted game is denoted by $v_{0,\varepsilon}$. Note that $\tilde{q}$ does not introduce new (nonloop) transitions, so that the map $\varepsilon \longmapsto v_{0,\varepsilon}$ is continuous at $\varepsilon = 0$, and the marginal undiscounted values $\partial_H v_0$ exist. We determine $\partial_H v_0$ in three steps.

Step 1. *The auxiliary matrices.* We consider the auxiliary matrices of a β-discounted version of the game, denoted by $\Delta^0_{\beta,\varepsilon}$ and $\Delta^k_{\beta,\varepsilon}$. To compute them, we follow the construction described in Section 2.2. A direct calculation gives:

$$\Delta^0_{\beta,\varepsilon}(z) = \beta^2 \begin{pmatrix} \beta^2 & \beta & \beta & \beta(2-\beta) \\ \beta & \beta(1 - \varepsilon \tilde{q}_2(1-\beta)) & \beta(2-\beta) & 1 - \varepsilon \tilde{q}_2(1-\beta) \\ \beta & \beta(2-\beta) & \beta(1 - \varepsilon \tilde{q}_1(1-\beta)) & 1 - \varepsilon \tilde{q}_1(1-\beta) \\ \beta(2-\beta) & 1 - \varepsilon \tilde{q}_2(1-\beta) & 1 - \varepsilon \tilde{q}_1(1-\beta) & (1 - \varepsilon \tilde{q}_1(1-\beta))(1 - \varepsilon \tilde{q}_2(1-\beta)) \end{pmatrix},$$

Figure 4. The Bewley-Kohlberg stochastic game with parameters $a, b \geq 0$. (The original game corresponds to the case $a = b = 1$.)

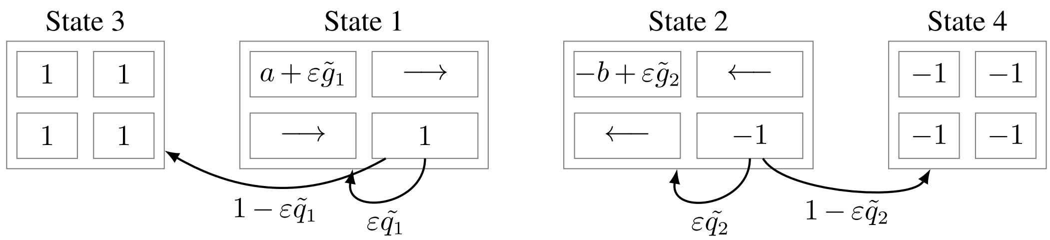
State 1

a	$\longrightarrow$
$\longrightarrow$	1^*

State 2

$-b$	$\longleftarrow$
$\longleftarrow$	-1^*

Figure 5. The Bewley-Kohlberg stochastic game with parameters $a, b \geq 0$.



and

$$\Delta^1_{\beta,\varepsilon}(z) = \beta^2 \begin{pmatrix} \beta^2(a+\varepsilon\tilde{g}_1) & \beta(a+\varepsilon\tilde{g}_1) & \beta(1-\beta)(\varepsilon\tilde{g}_2 - b) & 0 \\ \beta(a+\varepsilon\tilde{g}_1) & \beta(1-\varepsilon\tilde{q}_2(1-\beta))(a-\varepsilon\tilde{g}_1) & 0 & (1-\beta)(1-\varepsilon\tilde{q}_2(1-\beta)) \\ \beta(1-\beta)(\varepsilon\tilde{g}_2 - b) & 0 & \beta(1-\varepsilon\tilde{q}_1(1-\beta)) & 1-\varepsilon\tilde{q}_1(1-\beta) \\ 0 & -(1-\beta)(1-\varepsilon\tilde{q}_2(1-\beta)) & 1-\varepsilon\tilde{q}_1(1-\beta) & (1-\varepsilon\tilde{q}_1(1-\beta))(1-\varepsilon\tilde{q}_2(1-\beta)) \end{pmatrix}.$$

One then sets $W_{\beta,\varepsilon}(z) := \Delta^k_{\beta,\varepsilon} - z\Delta^0_{\beta,\varepsilon}$, for all $z \in \mathbb{R}$.

Step 2. *The desired polynomial.* We now look for a polynomial P satisfying $P(\varepsilon, v_{0,\varepsilon}) = 0$ for all $\varepsilon \geq 0$ small enough. To do so, we consider the possible candidates, which are in the set $\mathcal{P}_{(\Gamma,H)}$ by Proposition 1. To determine the suitable one, we first note that for every $\beta \in (0,1]$, the unperturbed β-discounted stochastic game admits a unique pair of optimal stationary strategies, which is of full support. This property was already noted by Bewley and Kohlberg [7]. By continuity, this property still holds in the perturbed auxiliary discounted game $\Gamma_{\beta,\varepsilon}$ for $\varepsilon \geq 0$ sufficiently small. As already noted (see comment (2) after Proposition 1), this property implies that the desired polynomial is $P(\varepsilon, z) := \Phi(\det(W_{\beta,\varepsilon}(z)))$. A direct computation using Matlab gives then $P(\varepsilon, z) = p(\varepsilon, z)^2$, where

$$p(\varepsilon, z) = a - b - z(a+b+2) + \varepsilon((1-z)(\tilde{g}_1 - (z-a)\tilde{q}_1) + (1+z)(\tilde{g}_2 + (z+b)\tilde{q}_2) + o(\varepsilon).$$

Here, again, we have abbreviated the polynomial, as the terms in ε^2 play no role in the sequel.

Step 3. *A formula for the undiscounted marginal values.* Lastly, we use Theorem 3 to determine the undiscounted marginal values. By construction, p satisfies $p(\varepsilon, v_{0,\varepsilon}) = 0$ for all ε sufficiently small. We then check that $\frac{\partial p}{\partial z}(0, v_0) = -(a+b+2) \neq 0$. As the map $\varepsilon \longmapsto v_\varepsilon$ is continuous at $\varepsilon = 0$ by the choice of $\tilde{q}$, we can apply Theorem 3 to obtain the following explicit expression (recall that $v_0 = \frac{a-b}{a+b+2}$):

$$\partial_H v_0 = -\frac{\frac{\partial p}{\partial \varepsilon}(0, v_0)}{\frac{\partial p}{\partial z}(0, v_0)} = \frac{(1-v_0)\tilde{g}_1 + (1+v_0)\tilde{g}_2 + (v_0 - a)(1-v_0)\tilde{q}_1 + (v_0 + b)(1+v_0)\tilde{q}_2}{a+b+2}.$$

In particular, for $a = b = 1$ (and, thus, $v_0 = 0$) the following simpler formula holds:

$$\partial_H v_0 = \frac{\tilde{g}_1 + \tilde{g}_2 - \tilde{q}_1 + \tilde{q}_2}{4}.$$

Remark 2. In the Bewley-Kohlberg game, the polynomial $P(\varepsilon, z) := \Phi(\det(W_{\beta,\varepsilon}(z)))$ satisfies $\frac{\partial P}{\partial z}(0, v_0) = 0$. However, we could replace P by its divider, p, to obtain the desired formula from Theorem 3.

Remark 3. The two considered examples share a common regularity property: there exists a polynomial $P(\varepsilon, z)$ such that $P(\varepsilon, v_\varepsilon) = 0$ for all $\varepsilon \geq 0$ and $\frac{\partial P}{\partial z}(0, v_0)$ is independent from the perturbation (though it could be equal to zero). This property holds as long as there exists a pair of optimal stationary strategies in the perturbed discounted stochastic game $\Gamma_{\beta,\varepsilon}$ with a fixed and equal support for all $\beta > 0$ and $\varepsilon \geq 0$ sufficiently small. The reason for this stability is as follows: the fixed support implies there exists a constant submatrix that defines the optimal strategies, and this submatrix can then be chosen to define P, independently of the perturbation.

5. Preliminary Tools

Before going into the proofs of our main contributions, we provide some useful preliminary results. We start with the theory of matrix games developed by Shapley and Snow [27] to then briefly present the theory of semialgebraic sets. Lastly, we present some known results regarding the regularity of the perturbed value function of a stochastic game following Filar and Vrieze [9, chapter 4] for the discounted case and Solan [29] for the undiscounted case.

5.1. Shapley-Snow Theory

We follow the notation used in Shapley and Snow [27]. For any square matrix M, we denote $S(M)$, and $\mathrm{co}(M)$ the sum of its entries and its cofactor matrix, respectively.

Definition 1. For any matrix M, a *Shapley-Snow kernel* is a square submatrix $\overline{M}$ satisfying:

- $S(\mathrm{co}(\overline{M})) \neq 0$.
- The strategies $\overline{x} = \frac{\mathrm{co}(\overline{M})}{S(\mathrm{co}(\overline{M}))}\overline{1}$ and $\overline{y} = \frac{\mathrm{co}(\overline{M})^\top}{S(\mathrm{co}(\overline{M}))}\overline{1}$, when completed by zeros, are optimal strategies of M, respectively

for Player 1 and Player 2.

Shapley-Snow kernels characterize the extreme points of the set of optimal strategies, denoted $O^*(M)$. In particular, every matrix admits at least one and, at most, finitely many of these kernels. Moreover, for any Shapley-Snow kernel $\overline{M}$ of M the following properties hold:

i. $\mathrm{val}(M) = \frac{\det \overline{M}}{S(\mathrm{co}(\overline{M}))}$, so, in particular, if $\mathrm{val}(M) = 0$, then $\det(\overline{M}) = 0$.
ii. All entries of $\mathrm{co}(\overline{M})$ are of same sign, and not all zero.

5.2. Semialgebraic Sets

A set $A \subset \mathbb{R}^d$ is basic semialgebraic if it is defined by finitely many polynomial equalities or strict inequalities—that is, there exists $L \in \mathbb{N}$ and polynomials $p_0, p_1, \ldots, p_L$ in $\mathbb{R}[x_1, \ldots, x_d]$ such that

$$A = \{(x_1, \ldots, x_d) \in \mathbb{R}^d : p_0(x_1, \ldots, x_d) = 0,\, p_1(x_1, \ldots, x_d) > 0, \ldots, p_L(x_1, \ldots, x_d) > 0\}.$$

A semialgebraic set is the finite union or intersection of basic semialgebraic sets. A semialgebraic function is a function whose graph is a semialgebraic set.

The following result, referred to as the Tarski-Seidenberg elimination theorem, establishes the stability of semialgebraic by projection to lower dimensions.

Theorem 4 (Tarski-Seidenberg Theorem). *Consider $d \in \mathbb{N}$ and $A \subset \mathbb{R}^{d+\ell}$ a semialgebraic set. Then, $\{x \in \mathbb{R}^d : \exists w \in \mathbb{R}^\ell\ (x, w) \in A\}$ is a semialgebraic set of $\mathbb{R}^d$.*

As a consequence, any set that can be expressed using first-order formulas over the real field is semialgebraic.

Local expansions of semialgebraic functions were investigated by Puiseux [24]. A Puiseux series is a function of the form

$$f(\varepsilon) = \sum_{m \geq m_0} c_m \varepsilon^{m/M},$$

where $M \in \mathbb{N}$, $m_0 \in \mathbb{Z}$ and coefficients $(c_m)_{m \geq m_0} \subset \mathbb{R}$. Then, we have the following result.

Theorem 5 (Puiseux [24]). *Let $P \in \mathbb{R}[x, y]$ be a nonzero polynomial and $a > 0$. Suppose that $f : (0, a) \to \mathbb{R}$ is a continuous function satisfying $P(\varepsilon, f(\varepsilon)) = 0$, for all $\varepsilon \in (0, a)$. Then, $\varepsilon \longmapsto f(\varepsilon)$ admits a Puiseux expansion near zero.*

5.3. Regularity of the Perturbed Value Function

Consider a discounted stochastic game Γ with $\lambda > 0$ and an admissible perturbation $H = (\tilde{g}, \tilde{q}, \tilde{\lambda})$. Then, $v_\varepsilon := \mathrm{val}(\Gamma + \varepsilon H)$ is the $(\lambda + \varepsilon\tilde{\lambda})$-discounted value of some stochastic game. For any map $\phi : K \times I \times J \to \mathbb{R}^K$, let

$$\|\phi\|_1 := \max_{\ell \in K, (i,j) \in I \times J} \sum_{\ell' \in K} |\phi(\ell' \mid \ell, i, j)|.$$

Recall that $\tilde{q} : K \times I \times J \to \mathbb{R}^K$, but it is not a transition function (rather, $q + \varepsilon\tilde{q}$ is a transition function for all ε sufficiently small), so $\|\tilde{q}\|_1$ is not necessarily equal to one. Let $\|\cdot\|_\infty$ denote the standard L^∞-norm of $\mathbb{R}^d$, and let $z_+ := \max(0, z)$ for all $z \in \mathbb{R}$. The following property is a rephrasing of Filar and Vrieze [9, chapter 4, equation (4.19)] with the notation of the present paper.

Lemma 1. *The following inequality holds for all $\varepsilon, \varepsilon' \geq 0$ sufficiently small*:

$$|v_\varepsilon - v_{\varepsilon'}| \leq |\varepsilon - \varepsilon'| \left(\|\tilde{g}\|_\infty + \frac{1-(\lambda+\varepsilon\tilde{\lambda})}{\lambda+\varepsilon\tilde{\lambda}} \|\tilde{q}\|_1 \|g+\varepsilon'\tilde{g}\|_\infty + 2\frac{|\tilde{\lambda}|}{\lambda+\varepsilon\tilde{\lambda}} \|g+\varepsilon'\tilde{g}\|_\infty \right).$$

Next, consider an undiscounted stochastic game Γ and an admissible perturbation $H = (\tilde{g}, \tilde{q}, 0)$. By definition, $v_\varepsilon := \mathrm{val}(\Gamma + \varepsilon H)$ is then the undiscounted value of some stochastic game. For all pairs of transition functions q_1 and q_2 with the same domain—that is, $q_1, q_2 : K \times I \times J \to \Delta(K)$—let

$$r(q_1, q_2) := \max_{\substack{(i,j)\in I\times J \\ (\ell,\ell')\in K^2, \ell\neq\ell'}} \left\{ \max\left\{ \frac{q_1(\ell'|\ell,i,j)}{q_2(\ell'|\ell,i,j)}, \frac{q_2(\ell'|\ell,i,j)}{q_1(\ell'|\ell,i,j)} \right\} - 1 \right\},$$

with the convention $z/0 = +\infty$ for all $z > 0$, and $0/0 = 1$. The next result is a rephrasing of Solan [29, theorem 6] with the notation of the present paper.

Lemma 2. *Let Γ be a discounted or undiscounted stochastic game, and let $H = (\tilde{g}, \tilde{q}, 0)$ be an admissible perturbation. Then, for all $\varepsilon, \varepsilon' > 0$ sufficiently small,*

$$|v_{\varepsilon'} - v_\varepsilon| \leq \frac{4|K| r(q+\varepsilon\tilde{q}, q+\varepsilon'\tilde{q})}{(1-2|K| r(q+\varepsilon\tilde{q}, q+\varepsilon'\tilde{q}))_+} \|g+\varepsilon\tilde{g}\|_\infty + |\varepsilon-\varepsilon'| \|\tilde{g}\|_\infty.$$

Note that the perturbed value function is Lipschitz continuous as soon as $r(q, q+\varepsilon\tilde{q}) = O(\varepsilon)$. By Solan [29], this condition holds if $\tilde{q}$ does not introduce new transitions—that is, if there is no $(i,j) \in I \times J$ and $(\ell,\ell') \in K^2$ with $\ell \neq \ell'$ so that $q(\ell'|\ell,i,j) = 0$ and $(q+\varepsilon\tilde{q})(\ell'|\ell,i,j) > 0$—or the converse, for all ε sufficiently small. Before we continue, a technical observation on the definition of r is needed.

Remark 4. In the original statement of Solan [29], the definition of r was slightly different: there was no condition $\ell \neq \ell'$ in the maximum. However, this additional condition comes for free, as the bound in Solan [29, theorem 6] relies on the limit discounted occupation times for a Markov chain, which by Freidlin and Wentzell [10] depend only on the transitions between pairs of different states. Therefore, we have added the condition $\ell \neq \ell'$ in the definition of r.

The following three results are direct consequences of Lemma 1, Lemma 2, and the theory of semialgebraic sets (see Section 5.2). Their proof is provided in the appendix for completeness.

Proposition 2. *Consider a discounted or undiscounted stochastic game Γ and an admissible perturbation $H = (\tilde{g}, \tilde{q}, \tilde{\lambda})$. Then, there exists $\varepsilon_0 > 0$ so that $\varepsilon \longmapsto v(\Gamma + \varepsilon H)$ is continuous on $(0, \varepsilon_0]$. Moreover, continuity at $\varepsilon = 0$ holds if either*: (i) $\lambda > 0$; *or* (ii) $\lambda = 0$ *and* $r(q, q+\varepsilon\tilde{q}) = O(\varepsilon)$.

Proposition 3. *Consider a discounted or undiscounted stochastic game Γ and an admissible perturbation $H = (\tilde{g}, \tilde{q}, \tilde{\lambda})$. The marginal values exist if either* (i) $\lambda > 0$; *or* (ii) $\lambda = 0$ *and* $r(q, q+\varepsilon\tilde{q}) = O(\varepsilon)$.

Proposition 4. *Let Γ_λ be a discounted stochastic game, and $H = (\tilde{g}, \tilde{q}, \tilde{\lambda})$ an admissible perturbation. Then, $(\partial_H v(\Gamma_\lambda))_\lambda$ is uniformly bounded if $r(q, q+\varepsilon\tilde{q}) = O(\varepsilon)$.*

6. Proofs of Main Results

We are now ready to prove our main results. In the sequel, consider a fixed stochastic game $\Gamma = (K, k, I, J; g, q, \lambda)$ and an admissible perturbation $H = (\tilde{g}, \tilde{q}, \tilde{\lambda})$. To alleviate the notation, we set $v := v(\Gamma)$ and $\partial_H v := \partial_H v(\Gamma)$ throughout this section. Also, note that we omitted the subindex λ in Theorem 1, as the discount rate is fixed. We follow the same convention in its proof.

6.1. Proof of Theorem 1

The proof is divided into three steps. First, note that the map $z \longmapsto D(z)$ is well-defined. This is the case as the set of admissible strategies $O^*(\Gamma)$ is compact and convex (by the minmax theorem), and the payoff function is bilinear. Hence, $D(z)$ is well-defined by Sion [28, theorem 3.4]. Second, we argue the existence of at most one $z \in \mathbb{R}$ such that $D(z) = 0$. On the one hand, all the entries of Δ^0 are positive, as noted in Section 2.2. On the other hand, the monotonicity and continuity of the value operator imply that $z \longmapsto D(z)$ is a strictly decreasing, continuous, and bijective map from $\mathbb{R}$ to $\mathbb{R}$. Our second step follows. Lastly, we claim that $D(\partial_H v) = 0$. We present two alternative proofs.

6.1.1. First Approach. Let $(\varepsilon_m, x_m, y_m) \in (0,1) \times \Delta(I)^K \times \Delta(J)^K$ be a sequence converging to $(0, x_0, y_0)$ such that, for every m, $\varepsilon_m > 0$ and (x_m, y_m) is a pair of optimal stationary strategies of the perturbed game $\Gamma + \varepsilon_m H$, whose value is denoted by v_{ε_m}. To establish the desired result—that is, $D(\partial_H v) = 0$—it is enough to prove the following relations, where, as before, $\widehat{x_0} = \otimes_{\ell \in K} x_0^\ell \in \Delta(I^K)$ and $\widehat{y_0} = \otimes_{\ell \in K} y_0^\ell \in \Delta(J^K)$:

$$\begin{cases} \widehat{x_0}^\top (\partial_H \Delta^k - v\,\partial_H \Delta^0 - (\partial_H v)\,\Delta^0)\,y \geq 0, & \forall y \in O_2^*(\Gamma), \\ x^\top (\partial_H \Delta^k - v\,\partial_H \Delta^0 - (\partial_H v)\,\Delta^0)\,\widehat{y_0} \leq 0, & \forall x \in O_1^*(\Gamma). \end{cases}$$

For all $m \in \mathbb{N}$ and $y \in O_2^*(\Gamma)$, the following holds:

$$\begin{aligned} 0 &\leq \hat{x}_m^\top (\Delta_{\varepsilon_m}^k - v_{\varepsilon_m} \Delta_{\varepsilon_m}^0) y, \\ &= \hat{x}_m^\top (\Delta^k - v\Delta^0 + (\partial_H \Delta^k - v\,\partial_H \Delta^0 - (\partial_H v)\Delta^0)\varepsilon_m + o(\varepsilon_m)) y, \\ &= \hat{x}_m^\top (\Delta^k - v\Delta^0) y + \hat{x}_m^\top (\partial_H \Delta^k - v\,\partial_H \Delta^0 - (\partial_H v)\Delta^0) y\, \varepsilon_m + o(\varepsilon_m), \\ &\leq \varepsilon_m\, \hat{x}_m^\top (\partial_H \Delta^k - v\,\partial_H \Delta^0 - (\partial_H v)\Delta^0) y + o(\varepsilon_m). \end{aligned}$$

Indeed, the first inequality follows from Attia and Oliu-Barton [2, theorem 1]—that is, for any discounted stochastic game with value v, one has $\operatorname{val}(\Delta^k - v\Delta^0) = 0$—and from the fact that the optimality of x_m in the stochastic game $\Gamma + \varepsilon_m H$ implies the optimality of $\hat{x}_m$ in the auxiliary matrix game $\Delta_{\varepsilon_m}^k - v_{\varepsilon_m} \Delta_{\varepsilon_m}^0$. The second line is a standard asymptotic development for $\Delta_{\varepsilon_m}^k$, v_{ε_m} and $\Delta_{\varepsilon_m}^0$ at zero. The third line is a consequence of linearity, as we are dealing with mixed strategies and matrix games. The fourth line follows from the choice of y as an optimal strategy of $\Delta^k - v\Delta^0$ and from the fact that the latter has value zero by Attia and Oliu-Barton [2, theorem 1]. Dividing the last inequality by ε_m and then taking $m \to \infty$, one thus obtains

$$\widehat{x_0}^\top (\partial_H \Delta^k - v\,\partial_H \Delta^0 - (\partial_H v)\Delta^0) y \geq 0, \quad \forall y \in O_2^*(\Gamma).$$

By reversing the roles of the players, one similarly obtains the desired relation for $\hat{y}_0$, which proves the claim. □

6.1.2. Second Approach. Consider a matrix game $M(\varepsilon)$ whose entries depend on ε and are differentiable in a neighborhood of zero. Let $\frac{\partial}{\partial \varepsilon} M(\varepsilon)$ be its entry-wise derivative, and let $O^*(M)$ be the set of optimal strategies of $M(0)$. By Mills [16], if there exist two matrices of the same size M and N such that $M(\varepsilon) = M + \varepsilon N$, then the value operator commutes with differentiation up to a restriction of the strategy domain, namely:

$$\frac{\partial}{\partial \varepsilon} (\operatorname{val} M(\varepsilon))_{|\varepsilon=0} = \operatorname{val}_{O^*(M)} \left(\frac{\partial}{\partial \varepsilon} M(\varepsilon)_{|\varepsilon=0} \right).$$

The extension of this result to a nonlinear differentiable dependency on ε is straightforward. Indeed, in this case, $N := \frac{\partial}{\partial \varepsilon} M(\varepsilon)_{|\varepsilon=0}$ satisfies $M(\varepsilon) = M(0) + \varepsilon N + E(\varepsilon)$, where $E(\varepsilon)$ is a matrix of same size as $M(\varepsilon)$ such that all its entries are $o(\varepsilon)$ as ε goes to zero. The extension of Mills' result follows from the monotonicity and continuity of the value operator. Next, set $M(\varepsilon) := W^k(v_\varepsilon) = \Delta_\varepsilon^k - v_\varepsilon \Delta_\varepsilon^0$ for all ε. Then, by definition, $O^*(M) = O(W(v)) = O^*(\Gamma)$. On the other hand, $\operatorname{val}(M(\varepsilon)) = 0$ for all ε sufficiently small by Attia and Oliu-Barton [2, theorem 1]. The desired result follows then from the differentiable extension of Mills' result, applied to $M(\varepsilon)$. Indeed,

$$\begin{aligned} 0 = \frac{\partial}{\partial \varepsilon} (\operatorname{val}(\Delta_\varepsilon^k - v_\varepsilon \Delta_\varepsilon^0))_{|\varepsilon=0} &= \operatorname{val}_{O^*(\Gamma)} \left(\frac{\partial}{\partial \varepsilon} (\Delta_\varepsilon^k - v_\varepsilon \Delta_\varepsilon^0)_{|\varepsilon=0} \right) \\ &= \operatorname{val}_{O^*(\Gamma)} (\partial_H \Delta^k - v\,\partial_H \Delta^0 - (\partial_H v)\Delta^0), \\ &= D(\partial_H v). \end{aligned}$$

These three steps give the desired result. □

6.2. Simplified Formula for Perturbed Payoffs

We consider Theorem 1 for the particular case of a perturbation of the payoffs only—that is, $H = (\tilde{g}, 0, 0)$. We prove that, in this case, our formula is equivalent to the following simple expression:

$$\partial_H v(\Gamma) = \operatorname{val}_{O(\Gamma)} \tilde{\gamma}(x, y),$$

where $\tilde{\gamma}$ is the expected payoff function of $\tilde{\Gamma} = (K, k, I, J; \tilde{g}, q, \lambda)$, and $O(\Gamma)$ is the set of optimal stationary strategies of Γ (not to be confused with $O^*(\Gamma)$, which is the set of optimal strategies of the matrix game $W(v)$). We proceed in three steps. First, let $\tilde{\Delta}^0$ and $\tilde{\Delta}^k$ denote the auxiliary matrices of $\tilde{\Gamma}$. By construction, $\Delta^0 = \tilde{\Delta}^0$ because the 0-auxiliary matrix depends on (q, λ), but not on the payoffs, as noted in Section 2.2. For the same reason $\partial_H \Delta^0 = 0$. Second, by the linearity of the determinant, the definition of the k-auxiliary matrix implies $\partial_H \Delta^k = \tilde{\Delta}^k$. Theorem 1 thus boils down to $\partial_H v(\Gamma)$ being the unique solution to

$$z \in \mathbb{R}, \quad \mathrm{val}_{O^*(\Gamma)}(\tilde{\Delta}^k - z\tilde{\Delta}^0) = 0.$$

Third, we prove that $O^*(\Gamma)$ can be replaced with $O(\Gamma)$ in the previous equation. In the proof of Theorem 1 (first approach, replacing $\partial_H \Delta^k - v\partial_H \Delta^0 - (\partial_H v)\Delta^0$ with $\tilde{\Delta}^k - (\partial_H v)\tilde{\Delta}^0$), we proved the existence of $x_0 \in \Delta(I)^K$ such that

$$\hat{x}_0^\top (\tilde{\Delta}^k - \partial_H v\, \tilde{\Delta}^0) y \geq 0, \quad \forall y \in O_2^*(\Gamma).$$

Using the positivity of $\tilde{\Delta}^0$ and taking the minimum over y, one thus obtains:

$$\min_{y \in O_2^*(\Gamma)} \frac{\hat{x}_0^\top \tilde{\Delta}^k y}{\hat{x}_0^\top \tilde{\Delta}^k y} \geq \partial_H v.$$

By Attia and Oliu-Barton [2], $O_2(\Gamma) \subset O_2^*(\Gamma)$ holds via the canonical inclusion $y \longmapsto \hat{y}$. Hence,

$$\min_{y \in O_2(\Gamma)} \frac{\hat{x}_0^\top \tilde{\Delta}^k \hat{y}}{\hat{x}_0^\top \tilde{\Delta}^k \hat{y}} \geq \partial_H v.$$

On the other hand, the definition of the auxiliary matrices (see Section 2.2) and the multilinearity of the following map:

$$\begin{aligned} &\Delta(I)^K \times \Delta(J)^K \times \mathbb{R} \to \mathbb{R} \\ &\qquad (x, \quad y, \quad z) \longmapsto \hat{x}^\top \tilde{\Delta}^k \hat{y} - z\,\hat{x}^\top \tilde{\Delta}^0 \hat{y}, \end{aligned}$$

imply that, for any pair of stationary strategies (x, y),

$$\tilde{\gamma}(x, y) = \frac{\hat{x}^\top \tilde{\Delta}^k \hat{y}}{\hat{x}^\top \tilde{\Delta}^0 \hat{y}}.$$

Putting the last two equations together, one thus obtains

$$\min_{y \in O_2(\Gamma)} \tilde{\gamma}(x_0, y) \geq \partial_H v.$$

Therefore, $\max_{x \in O_1(\Gamma)} \min_{y \in O_2(\Gamma)} \tilde{\gamma}(x, y) \geq \partial_H v$. Finally, we reverse the roles of the players to obtain the reverse inequality and conclude because the max min is always smaller or equal to the min max—that is,

$$\partial_H v \geq \min_{y \in O_2(\Gamma)} \max_{x \in O_1(\Gamma)} \tilde{\gamma}(x, y) \geq \max_{x \in O_1(\Gamma)} \min_{y \in O_2(\Gamma)} \tilde{\gamma}(x_0, y) \geq \partial_H v.$$

This concludes the proof. □

6.3. Proof of Theorem 2

For all $\lambda \in (0, 1]$, recall that $\Gamma_\lambda := (K, k, I, J; g, q, \lambda)$, and let $v_\lambda \in \mathbb{R}$ denote its value. Let Δ_λ^0, Δ_λ^k be the corresponding auxiliary matrices, and $W_\lambda(z) := \Delta_\lambda^k - z\Delta_\lambda^0$ for all $z \in \mathbb{R}$. Lastly, for all $z \in \mathbb{R}$, we set

$$\begin{cases} D_\lambda(z) := \mathrm{val}_{O^*(\Gamma_\lambda)}(\partial_H \Delta_\lambda^k - v_\lambda\, \partial_H \Delta_\lambda^0 - z\,\Delta_\lambda^0), \\ F(z) := \lim\limits_{\lambda \to 0} \dfrac{1}{\lambda^{|K|}} D_\lambda(z) \in [-\infty, +\infty]. \end{cases}$$

We start by assuming the following claims: (1) F is well-defined; (2) F is strictly decreasing, provided it is not constant; and (3) F is not constant. Together, these statements prove that there is a unique point where F changes sign. Let w be the unique point where $F(z) > 0$ for all $z < w$ and $F(z) < 0$ for all $z > w$. Then, for all $\delta > 0$, we have that

$F(w-\delta) > 0$. Therefore, by the definition of F, there exists $\lambda_0 > 0$ so that

$$D_\lambda(w-\delta) > 0, \quad \forall \lambda \in (0, \lambda_0).$$

By Theorem 1, this implies $\partial_H v_\lambda > w - \delta$ for all $\lambda \in (0, \lambda_0)$. Because δ is arbitrary, we deduce that $\liminf_{\lambda\to 0} \partial_H v_\lambda \geq w$. Similarly, we deduce that $\limsup_{\lambda\to 0} \partial_H v_\lambda \leq w$. Therefore, $\lim_{\lambda\to 0} \partial_H v_\lambda$ exists and satisfies the desired property. The rest of the proof proves the three above-mentioned claims.

Claim 1. Recall that, by Puiseux [24], v_λ is a Puiseux series near zero. Also, the entries of $\partial_H \Delta_\lambda^k$, $\partial_H \Delta_\lambda^0$, and Δ_λ^0 are polynomials in λ. Fix $z \in \mathbb{R}$. By Mills [16], $D_\lambda(z)$ is the value of a linear program (see Appendix A for more details). Therefore, as λ goes to zero, all the rest of parameters being fixed, $D_\lambda(z)$ is a piece-wise rational fraction in λ and v_λ. Hence, $\lambda \longmapsto D_\lambda(z)$ is a Puiseux series near zero, which gives the desired result.

Claim 2. Consider $(z_1, z_2) \in \mathbb{R}^2$ such that $z_2 > z_1$. Then, for all $\lambda \in (0,1]$ and $(x,y) \in O^*(\Gamma_\lambda)$,

$$x^\top(\partial_H \Delta_\lambda^k - v_\lambda \partial_H \Delta_\lambda^0 - z_1 \Delta_\lambda^0) y - x^\top(\partial_H \Delta_\lambda^k - v_\lambda \partial_H \Delta_\lambda^0 - z_2 \Delta_\lambda^0) y = (z_2 - z_1) x^\top \Delta_\lambda^0 y.$$

Moreover, as all the entries of Δ_λ^0 are larger than or equal to $\lambda^{|K|}$, as noted in Section 2.2, and because x and y are probabilities, one has

$$x^\top \Delta_\lambda^0 y \geq \lambda^{|K|}.$$

Maximization over $x \in O^*(\Gamma)$ and minimization over $y \in O_2^*(\Gamma)$ give then

$$D_\lambda(z_1) \geq D_\lambda(z_2) + (z_2 - z_1)\lambda^{|K|}.$$

Division by $\lambda^{|K|}$ and then taking λ to zero yields then:

$$F(z_1) \geq F(z_2) + z_2 - z_1.$$

The function F is thus strictly decreasing, provided that it is not constant and equal to $+\infty$ or $-\infty$.

Claim 3. By assumption, $\partial_H v_\lambda$ is uniformly bounded, so there exists $C > 0$ such that $-C \leq \partial_H v_\lambda \leq C$ for all $\lambda \in (0,1]$. By Theorem 1, one then has $D_\lambda(C) \leq 0 \leq D_\lambda(-C)$ for all λ. Division by $\lambda^{|K|}$ and then taking λ to zero yields $F(C) \leq 0 \leq F(-C)$, so, in particular, F is not constant. □

6.4. Differentiation and Limit Operators Do Not Commute

When both the limit of the marginal discounted values and the marginal undiscounted values exist, it is natural to ask whether they are equal. In this section, first, we show that this is the case in two very particular cases: (a) when $|K| = 1$; and (b) when $|I| = |J| = 1$ and $H = (\tilde{g}, 0, 0)$. Then, we show by example that these two notions differ in general.

In the sequel, consider a discounted stochastic game $\Gamma_\lambda = (K, k, I, J; g, q, \lambda)$ with $\lambda > 0$ and an admissible perturbation $H = (\tilde{g}, \tilde{q}, \tilde{\lambda})$ and set $v_{\lambda,\varepsilon} := \operatorname{val}(\Gamma_{\lambda,\varepsilon})$ for all (λ, ε) sufficiently small.

a. Assume $|K| = 1$. The transition function is trivial, and g and $\tilde{g}$ are matrix games of size $I \times J$. Moreover, $v_{\lambda,\varepsilon} = \operatorname{val}(g + \varepsilon\tilde{g})$ for all ε, so, in particular, it is independent from λ, and thus equal to $\lim_{\lambda\to 0} v_{\lambda,\varepsilon}$, too. By Mills [16], the marginal value is equal to $\partial_H v(\Gamma_\lambda) = \operatorname{val}_{O^*(g)} \tilde{g}$, and this expression is both equal to $\partial_H \lim_{\lambda\to 0} v(\Gamma_\lambda)$ and to $\lim_{\lambda\to 0} \partial_H v(\Gamma_\lambda)$.

b. Assume $|I| = |J| = 1$ and $H = (\tilde{g}, 0, 0)$. The game $\Gamma_{\lambda,\varepsilon}$ is then a discounted Markov chain with perturbed payoffs. Indeed, the transition function can be identified with a stochastic matrix $Q \in \mathbb{R}^{K\times K}$, and the payoffs are state-dependent and given by $g_\varepsilon := g + \varepsilon \tilde{g} \in \mathbb{R}^K$. Let Π_λ be the k-th column of $\sum_{m\geq 1} \lambda(1-\lambda)^{m-1} Q^{m-1}$, and $\Pi_0 := \lim_{\lambda\to 0} \Pi_\lambda$ where the limit exists by semialgebraicity. Then, $v_{\lambda,\varepsilon} = \Pi_\lambda g_\varepsilon$ for all $\lambda \in (0,1]$ and $\varepsilon \geq 0$ sufficiently small, and $v_{0,\varepsilon} := \lim_{\lambda\to 0} v_{\lambda,\varepsilon} = \Pi_0 g_\varepsilon$. Hence,

$$\partial_H \lim_{\lambda\to 0} v(\Gamma_\lambda) = \lim_{\varepsilon\to 0} \frac{\Pi_0(g + \varepsilon\tilde{g}) - \Pi_0 g}{\varepsilon} = \Pi_0 \tilde{g}.$$

On the other hand, for all fixed λ,

$$\partial_H v(\Gamma_\lambda) = \lim_{\varepsilon\to 0} \frac{v_{\lambda,\varepsilon} - v_{\lambda,0}}{\varepsilon} = \lim_{\varepsilon\to 0} \frac{\Pi_\lambda(g + \varepsilon\tilde{g}) - \Pi_\lambda g}{\varepsilon} = \Pi_\lambda \tilde{g}.$$

Taking λ to zero gives the desired result—that is, $\lim_{\lambda\to 0} \partial_H v(\Gamma_\lambda) = \partial_H \lim_{\lambda\to 0} v(\Gamma_\lambda)$.

6.4.1. A Minimal Counterexample. Consider now the example of Section 3.2. For $\lambda \in [0,1]$ and $\varepsilon \geq 0$, a direct computation yields:

$$v_{\lambda,\varepsilon} = \begin{cases} (1-\lambda)(1+\varepsilon) & \text{if } \varepsilon \geq \dfrac{\lambda}{1-\lambda}, \\ 1 & \text{otherwise.} \end{cases}$$

Hence, for all $\lambda > 0$, there exists $\varepsilon_0 > 0$ small enough such that $v_{\lambda,\varepsilon} = 1$ for all $\varepsilon \in [0, \varepsilon_0]$. Consequently,

$$\lim_{\lambda \to 0} \partial_H v(\Gamma_\lambda) = \lim_{\varepsilon \to 0} \frac{v_{\lambda,\varepsilon} - v_{\lambda,0}}{\varepsilon} = 0.$$

On the other hand, for all $\varepsilon \geq 0$, $v_{0,\varepsilon} = \lim_{\lambda \to 0} v_{\lambda,\varepsilon} = 1 + \varepsilon$, so

$$\partial_H v(\Gamma_0) = \lim_{\varepsilon \to 0} \frac{v_{0,\varepsilon} - v_{0,0}}{\varepsilon} = 1.$$

In other words, $\lim_{\lambda \to 0} \partial_H v(\Gamma_\lambda)$ and $\partial_H v(\Gamma_0)$ both exist and differ.

6.5. Proof of Proposition 1

We start by proving an analog version of Proposition 1 for the case $\lambda + \tilde{\lambda} > 0$—that is, that there exists a square submatrix of $W_\varepsilon(z)$, denoted by $\overline{W}_\varepsilon(z)$, so that $P(\varepsilon, z) := \det(\overline{W}_\varepsilon(z))$ is nonzero and $P(\varepsilon, v_\varepsilon) = 0$ for all $\varepsilon \geq 0$ sufficiently small. As in Proposition 1, the novelty of this remark is not the existence of such a polynomial, which follows from the theory of semialgebraic sets (see Section 5.2), but the identification of a tractable set of possible candidates.

For every $\varepsilon > 0$ sufficiently small, the assumption $\lambda + \tilde{\lambda} > 0$ implies that the perturbed game $\Gamma + \varepsilon H$ is a $(\lambda + \varepsilon\tilde{\lambda})$-discounted stochastic game. Consider a Shapley-Snow kernel $\overline{W}_\varepsilon(v_\varepsilon)$ of $W_\varepsilon(v_\varepsilon)$, and let $\overline{\Delta}_\varepsilon^k$ and $\overline{\Delta}_\varepsilon^0$ be the corresponding auxiliary matrices so that $\overline{W}_\varepsilon(z) = \overline{\Delta}_\varepsilon^k - z\overline{\Delta}_\varepsilon^0$ for all z. Let us show that $P(\varepsilon, z) := \det(\overline{W}_\varepsilon(z))$ is nonzero and satisfies $P(\varepsilon, v_\varepsilon) = 0$. For the former, we use Jacobi's formula (i.e., for all two square matrices M and N of same size, the derivative of $z \longmapsto \det(M - zN)$ is $-\mathrm{tr}(\mathrm{co}(M)^\top N)$) and the definition of $W_\varepsilon(z) = \Delta_\varepsilon^k - z\Delta_\varepsilon^0$, to get:

$$\frac{\partial P}{\partial z}(\varepsilon, v_\varepsilon) = -\mathrm{tr}(\mathrm{co}(\overline{W}_\varepsilon(v_\varepsilon))^\top \overline{\Delta}_\varepsilon^0).$$

On the other hand, by Shapley and Snow [27], for any Shapley-Snow kernel $\overline{M}$ of M, all the entries of $\mathrm{co}(\overline{M})$ are of the same sign, and not all zero (see Section 5.1). Because all entries of $\overline{\Delta}_\varepsilon^0$ are strictly positive, as noted in Section 2.2, it follows that $\frac{\partial P}{\partial z}(\varepsilon, v_\varepsilon) \neq 0$, so, in particular, $P \not\equiv 0$. The relation $P(\varepsilon, v_\varepsilon) = 0$ follows from Attia and Oliu-Barton [2, theorem 1], which implies $\mathrm{val}(W(v_\varepsilon)) = 0$, and from the theory of Shapley and Snow (see Section 5.1), which implies that for any Shapley-Snow kernel $\overline{W}_\varepsilon(v_\varepsilon)$ of $W(v_\varepsilon)$, one has $\det(\overline{W}_\varepsilon(v_\varepsilon) = 0$. As ε goes to zero, the Shapley-Snow kernels may vary, as well as the finitely many polynomials defined as their determinants. Because they are finitely many, there exists $\delta > 0$ so that for all two polynomials in this set, either they are equal or they never cross in the interval $(0, \delta]$. If $\overline{W}_\delta(v_\delta)$ is a Shapley-Snow kernel of $W_\delta(v_\delta)$, then $P(\varepsilon, z) := \det(\overline{W}_\delta(z))$ is the desired polynomial satisfying $P(\varepsilon, v_\varepsilon) = 0$ for all $\varepsilon \in (0, \delta]$. Moreover, the equality at $\varepsilon = 0$ follows from continuity.

We now prove Proposition 1. In this case, $\lambda = 0$, and also $\tilde{\lambda} = 0$ by the choice of $H = (\tilde{g}, \tilde{q}, 0)$. The perturbed game $\Gamma + \varepsilon H$ is thus an undiscounted stochastic game for all ε sufficiently small. Let $\varepsilon > 0$ be fixed. For all $\beta \in (0,1]$, consider the auxiliary β-discounted stochastic game $\Gamma_{\beta,\varepsilon} := (K, k, I, J; g + \varepsilon\tilde{g}, q + \varepsilon\tilde{q}, \beta)$, whose value and auxiliary parameterized game are denoted, respectively, by $v_{\beta,\varepsilon}$ and $W_{\beta,\varepsilon}(z)$. To this game, we can apply the property that we just proved—that is, statement (i) at the end of Section 3.2. Hence, there exists $\delta > 0$ and a fixed square submatrix $\overline{W}_{\beta,\varepsilon}(v_{\beta,\varepsilon})$ of $W_{\beta,\varepsilon}(v_{\beta,\varepsilon})$ so that $\det(\overline{W}_{\beta,\varepsilon}(z))$ is a nonzero polynomial in the variables (β, z) and $\det(\overline{W}_{\beta,\varepsilon}(v_{\beta,\varepsilon})) = 0$ for all $\beta \in [0, \delta]$. For this fixed submatrix, as ε varies, $P(\varepsilon, \beta, z) := \det(\overline{W}_{\beta,\varepsilon}(z))$ defines a polynomial in (β, ε, z). Its projection $\Phi(P) \in \mathbb{R}[\varepsilon, z]$ belongs to the set $\mathcal{P}_{(\Gamma,H)}$ by the definition of this set. By definition of $\Phi(P)$, there exist $s \geq 0$ and polynomials $(R_m)_{m>s}$ in $\mathbb{R}[\varepsilon, z]$ such that $P(\varepsilon, \beta, z) = \Phi(P)(\varepsilon, z)\beta^s + \sum_{m>s} R_m(\varepsilon, z)\beta^m$. This equality holds, as well as $P(\varepsilon, \beta, v_{\beta,\varepsilon}) = 0$, for all $\beta \in [0, \delta]$. Hence, dividing by β^s and letting β go to zero, we have that $\Phi(P)(\varepsilon, v_{0,\varepsilon}) = 0$, where $v_{0,\varepsilon} := \lim_{\beta \to 0} v_{\beta,\varepsilon}$ exists as the limit value of a discounted stochastic game by Bewley and Kohlberg [6]. Let $\varepsilon > 0$ go to zero. Although the choice of the polynomial P, and thus $\Phi(P)$, varies with ε, because there are only finitely many candidates for this polynomial, there exist $\varepsilon_0 > 0$ and $\Phi(P) \in \mathcal{P}_{(\Gamma,H)}$ such that $\Phi(P)(\varepsilon, v_{0,\varepsilon}) = 0$ for all $\varepsilon \in (0, \varepsilon_0]$. Note that this relation can be extended to $\varepsilon = 0$ as soon as the map $\varepsilon \longmapsto v_{0,\varepsilon}$ is continuous at zero, which gives the desired result.

6.6. Proof of Theorem 3

By Proposition 1, there exists a polynomial P such that, for $\varepsilon > 0$ sufficiently small, $P(\varepsilon, v_\varepsilon) = 0$. Moreover, the map $\varepsilon \longmapsto v_\varepsilon$ is continuous at $\varepsilon = 0$ by assumption, so $P(0, v_0) = 0$ holds, too. The desired result follows then from the implicit function theorem (see Krantz and Parks [14]) applied to $P \in \mathbb{R}[\varepsilon, z]$ at $(0, v_0)$, which holds as P is continuous and differentiable, and satisfies $P(0, v_0) = 0$ and $\frac{\partial}{\partial z} P(0, z)\big|_{z=v_0} \neq 0$.

7. Open Questions

Our article contributes to the theory of stochastic games by providing explicit formulas for the marginal values, in the discounted and the undiscounted case. On the way, however, some questions were raised, but only partially solved. The following points, for example, deserve further research.

- **Complexity of the marginal values.** In the present paper, we provided an algorithm to compute approximations of the marginal discounted values, which is polynomial in $|I^K|$ and $|J^K|$ (and the size of the input, assuming it is rational). To address their exact computation, however, the following questions arise. Are the marginal discounted values algebraic? And, if so, what is a tight bound for their degree and for the size of coefficients of their minimal polynomial?
- **Computation of the limit discounted marginal values.** As already noted, Theorem 2 suggests a dichotomic search algorithm for the limit marginal values based on the computation of (the sign of) $F(z) = \lim_{\lambda \to 0} \lambda^{|K|} D_\lambda(z)$, where $D_\lambda(z) = \mathrm{val}_{O^*(\Gamma_\lambda)}(\partial_H \Delta_\lambda^k - v_\lambda\, \partial_H \Delta_\lambda^0 - z\, \Delta_\lambda^0)$. The map $\lambda \longmapsto D_\lambda(z)$ is a Puiseux series near zero (see Claim 1 of Section 6.3), so there exists $\lambda_0 > 0$ such that its sign is constant in $(0, \lambda_0)$. Hence, determining an explicit lower bound for λ_0 is required to derive a bound on the computational complexity of the above-mentioned algorithm. A similar problem was obtained, and solved, in Oliu-Barton [20], but there, the map $\lambda \longmapsto D_\lambda(z)$ was a rational fraction near zero. Additional research is thus needed to extend this result to the present case.
- **Computability of *P* in Proposition 1.** As noted in the comments after Proposition 1, determining a nonzero polynomial $P \in \mathcal{P}_{(\Gamma, H)}$ such that $P(\varepsilon, v_\varepsilon) = 0$ for all ε sufficiently small may be hard in general. A practical condition was given—namely, the existence of a pair of optimal stationary strategies in the auxiliary discounted game $\Gamma_{\beta,\varepsilon}$, whose support is fixed for all $\beta > 0$ and $\varepsilon \geq 0$ sufficiently small. Determining the desired polynomial efficiently in the absence of this simplifying property requires further research. Further, suppose that the desired polynomial satisfies $\frac{\partial}{\partial z} P(0, z)\big|_{z=v_0} \neq 0$. Is it always the case that a divider p of P exists satisfying (1) $p(\varepsilon, v_\varepsilon) = 0$ for all ε sufficiently small, and (2) $\frac{\partial}{\partial z} p(0, z)\big|_{z=v_0} \neq 0$, as was the case in the Bewley-Kohlberg example? This property would have two important implications: first, that the formula of Theorem 3 can be applied to every stochastic game; second, that the undiscounted marginal values are algebraic of the same degree as the undiscounted values.
- **Derivation and limit operators.** As noted via a minimal example, the derivation and limit operators do not commute, in general. However, they do commute in certain cases, beyond the ones identified. For example, they commute in the Big Match. It is thus natural to look for necessary and sufficient conditions for the two operators to commute.

Acknowledgments

The authors are particularly grateful to Eilon Solan, Krishnendu Chatterjee, and Sylvain Sorin for their insightful remarks on an earlier version of this draft; and to the editors and reviewers for their important suggestions.

Appendix A. Computational Complexity of Theorem 1

Recall that Theorem 1 gives a characterization of the marginal values of a discounted stochastic game as the unique $z \in \mathbb{R}$ such that $D(z) = 0$, where $D(z)$ is a constrained value problem. We now explain how to derive an algorithm from this characterization to compute an approximation of the marginal values to a desired level of accuracy. The algorithm is a dichotomic search on $D(z)$, which is possible because $z \longmapsto D(z)$ is strictly monotonous. Consequently, its computational complexity is reduced to (1) bounds on the marginal values, and (2) the complexity of the query $D(z)$.

In the sequel, we assume that the input data are all rational numbers—that is, all entries of $(g, q, \tilde{g}, \tilde{q})$ as well as λ and $\tilde{\lambda}$ are rational numbers—and their complexity refers to the size of their binary representation.

Step 1: *A bound for the marginal values.* By Lemma 1,

$$|\partial_H v(\Gamma)| \leq \|\tilde{g}\|_\infty + \frac{1 - \lambda}{\lambda} \|\tilde{q}\|_1 \, \|g\|_\infty + 2 \frac{|\tilde{\lambda}|}{\lambda} \|g\|_\infty.$$

We can therefore start the dichotomic search in the interval $[-C, C]$, where C is an integer upper bound of the previous expression. This choice ensures that z is a rational number during the dichotomic search, whose size is polynomial in the input.

Step 2: *Computational complexity of D(z).* Consider the monotonic operator

$$z \longmapsto D(z) = \mathrm{val}_{O^*(\Gamma)}(\partial_H \Delta^k - v(\Gamma)\, \partial_H \Delta^0 - z\, \Delta^0).$$

The computation of $D(z)$ calls for two key observations: first, $D(z)$ is the value of a matrix game, where players are restricted to playing optimal strategies of another matrix game; second, the input of $D(z)$ involves rational numbers and the algebraic number $v(\Gamma)$. To address them, we use, respectively, Mills [16], who proved that a constrained value problem can be written as a linear program, and Beling [5], who provided a polynomial-time algorithm to solve linear programs involving algebraic numbers.

Step 2(a): *Writing D(z) as a linear program.* Let us start by considering a standard linear program—that is, maximize $X^\top c$ subject to $b + AX \geq 0$ and $X \geq 0$, where $X, c \in \mathbb{R}^n$, $b \in \mathbb{R}^m$ and $A \in \mathbb{R}^{m \times n}$—and its canonical representation:

$$\left[\begin{array}{c|c} 0 & c \\ \hline b & A \end{array}\right].$$

The value of a matrix game $M \in \mathbb{R}^{I \times J}$ is the solution of the following linear program, where $\mathbf{1}_J$ and $\mathbf{1}_J$ denote column vectors of ones in $\mathbb{R}^I$ and $\mathbb{R}^J$, respectively:

$$\begin{aligned} \max_{x,z} \quad & z \\ \text{s.t.} \quad & x^\top M \geq z \mathbf{1}_J, \\ & x^\top \mathbf{1}_I = 1, \\ & (z, x) \in \mathbb{R} \times \mathbb{R}^I_+. \end{aligned}$$

Its canonical representation is then as follows:

$$\tilde{M} := \left[\begin{array}{c|cccccc} 0 & 1 & -1 & 0 & \dots & 0 \\ \hline 1 & 0 & 0 & -1 & \dots & -1 \\ -1 & 0 & 0 & 1 & \dots & 1 \\ 0 & -1 & 1 & & & \\ \vdots & \vdots & \vdots & & M^\top & \\ 0 & -1 & 1 & & & \end{array}\right].$$

By Mills [16], for all matrices $M, N \in \mathbb{R}^{I \times J}$ with $\mathrm{val}(M) = 0$, the constrained value problem $\mathrm{val}_{O(M)} N$ is a linear program whose canonical representation can be easily derived from the following block matrix:

$$\left[\begin{array}{c|c} \tilde{N} & \tilde{M} \\ \hline \tilde{M} & 0 \end{array}\right].$$

The first line and column of this block matrix correspond, respectively, to the objective vector $c \in \mathbb{R}^{2(|I|+3)}$ and to the constraint vector $b \in \mathbb{R}^{2(|J|+3)}$. Explicitly, the constrained value problem $\mathrm{val}_{O(M)} N$ is the solution of the following linear program:

$$\begin{aligned} \max_{z, x, u_0, u, \sigma} \quad & z + u \\ \text{s.t.} \quad & x^\top N + \sigma^\top M \geq (z + u) \mathbf{1}_J, \\ & x^\top M \geq z \mathbf{1}_J, \\ & x^\top \mathbf{1}_I = 1, \\ & \sigma^\top \mathbf{1}_I = u_0, \\ & (z, x, u_0, u, \sigma) \in \mathbb{R} \times \mathbb{R}^I_+ \times \mathbb{R}_+ \times \mathbb{R} \times \mathbb{R}^I_+. \end{aligned}$$

To compute $D(z)$ is thus sufficient to set $M := \Delta^k - v(\Gamma)\Delta^0$ and $N := \partial_H \Delta^k - v(\Gamma)\, \partial_H \Delta^0 - z\, \Delta^0$, so that $D(z) = \mathrm{val}_{O(M)} N$. The matrices M and N are of equal size, and $\mathrm{val}\, M = 0$. Consequently, we can compute $D(z)$ by solving the previously described linear program.

Step 2(b): *Computational complexity of D(z).* We claim that the computational complexity of $D(z)$ is polynomial in the input size, notably on the size of λ, and on $|I^K|$ and $|J^K|$.

Consider an approximation of the marginal $\partial_H v(\Gamma)$ up to some additive $\delta > 0$. A dichotomic search requires computing $D(z)$ on values of z that are rational numbers, whose size is polynomial in the input description and $1/\delta$. On the other hand, $D(z)$ is a linear program involving the value $v(\Gamma)$. By Oliu-Barton [20, proposition 1], this is an algebraic number of degree at most $d := \min(|I^K|, |J^K|)$ and whose defining polynomial has coefficients whose size is polynomial in the input size and d. By Beling [5, theorem 21], because all the numbers involved in the program belong to the extension of the rational numbers by the algebraic number $v(\Gamma)$, this linear program can be solved in polynomial time with respect to the input size, $1/\delta$, and its size $2(|I^K|+3)$ and $2(|J^K|+3)$. We thus conclude that a δ-approximation of $\partial_H v(\Gamma)$ can be obtained in polynomial time with respect to the input size, $1/\delta$, $|I^K|$ and $|J^K|$. This proves the desired result.

Appendix B. Regularity of the Perturbed Value Function

For completeness, we provide proofs for Propositions 2 and 3. As already mentioned, these two results follow from Solan [29] (i.e., Lemma 1), Filar and Vrieze [9] (i.e., Lemma 2), and the theory of semialgebraic sets.

B.1. Proof of Proposition 2

Recall that this statement is about the continuity of the perturbed value function. Consider $\varepsilon_0 > 0$ such that the perturbation is well-defined on $[0, \varepsilon_0]$. We distinguish three cases, $\lambda > 0$, $\lambda = \tilde{\lambda} = 0$, and $\tilde{\lambda} > \lambda = 0$.

i. Assume $\lambda > 0$. The Lipschitz continuity of $\varepsilon \longmapsto v_\varepsilon$ in the interval $[0, \varepsilon_0]$ follows directly from Lemma 1.

ii. Assume $\lambda = \tilde{\lambda} = 0$ and $\lim_{\varepsilon \to 0} r(q, q + \varepsilon\tilde{q}) = 0$. By Lemma 2, for all $(\varepsilon, \varepsilon') \in [0, \varepsilon_0]^2$,

$$|v_{\varepsilon'} - v_\varepsilon| \le \frac{4|K| r(q + \varepsilon\tilde{q}, q + \varepsilon'\tilde{q})}{(1 - 2|K| r(q + \varepsilon\tilde{q}, q + \varepsilon'\tilde{q}))_+} \|g + \varepsilon\tilde{g}\|_\infty + |\varepsilon - \varepsilon'| \|\tilde{g}\|_\infty.$$

For all fixed $\varepsilon > 0$, the map $\varepsilon' \longmapsto r(q + \varepsilon\tilde{q}, q + \varepsilon'\tilde{q})$ is continuous and equal to zero at $\varepsilon' = \varepsilon$. Therefore, $\varepsilon \longmapsto v_\varepsilon$ is continuous for all $\varepsilon > 0$ sufficiently small. The continuity at $\varepsilon = 0$ follows from $\lim_{\varepsilon \to 0} r(q, q + \varepsilon\tilde{q}) = 0$.

iii. Assume $\lambda = 0 < \tilde{\lambda}$ and $\lim_{\varepsilon \to 0} r(q, q + \varepsilon\tilde{q}) = 0$. Fix $\eta \in [0, \varepsilon_O] > 0$. For all $\varepsilon \in [\eta, \varepsilon_0]$, one can write $\Gamma + \varepsilon H = \Gamma + \eta H + (\varepsilon - \eta)H$, where the game $\Gamma + \eta H$ is a discounted stochastic game with discount rate $\varepsilon\tilde{\lambda}$. Therefore, the case $\lambda > 0$ can be applied to the game $\Gamma + \eta H$, which is perturbed by $(\varepsilon - \eta)H$. We thus deduce that $\varepsilon \longmapsto v_\varepsilon$ is right-continuous at η. The left-continuity can be obtained similarly by writing $\Gamma + \varepsilon H = \Gamma + \eta H + (\varepsilon_1 - \varepsilon)(-H)$ for all $\varepsilon \in (0, \eta]$. Hence, $\varepsilon \to v_\varepsilon$ is continuous for all $\varepsilon \in (0, \varepsilon_0]$. To establish the continuity at zero, consider, for all $\varepsilon \in [0, \varepsilon_0]$,

$$v_\varepsilon - v_0 = (v_\varepsilon - w_\varepsilon) + (w_\varepsilon - v_0),$$

where w_ε is the value of the game $(K, k, I, J; g, q, \varepsilon\tilde{\lambda})$. The first term, $v_\varepsilon - w_\varepsilon$, converges to zero when ε goes to zero by Solan [29, theorem 6]). The second term, $w_\varepsilon - v_0$, converges to zero when ε goes to zero because $w_\varepsilon \to v_0$ by the theory of semialgebraic sets, as shown by Bewley and Kohlberg [6]. Consequently, $\lim_{\varepsilon \to 0+} v_\varepsilon = v_0$, which proves the desired continuity at $\varepsilon = 0$. □

B.2. Proof of Proposition 3

Recall that this statement is about the existence of the marginal values. Distinguish the cases $\lambda + \tilde{\lambda} > 0$ and $\lambda + \tilde{\lambda} = 0$, although the proof relies on both cases in the theory of semi-algebraic sets—namely, in the fact that $\varepsilon \longmapsto v_\varepsilon$ admits a Puiseux series expansion near zero. Note that the marginal value $\partial_H v(\Gamma)$ is the right-derivative of this map.

i. Assume $\lambda + \tilde{\lambda} > 0$. Then, $\Gamma + \varepsilon H$ is a discounted stochastic game for all $\varepsilon \ge 0$. Let $(\varepsilon, x, y, v) \in \mathbb{R}^d$ be such that $\varepsilon \in [0, \varepsilon_0]$ is the perturbation parameter, (x, y) is a pair of optimal stationary strategies in the discounted game $\Gamma + \varepsilon H$, and v is its value. These conditions define finitely many polynomial equalities and inequalities whose coefficients depend on the entries of $(g, \tilde{g}, q, \tilde{q}, \lambda, \tilde{\lambda})$. Hence, the set of such tuples is semialgebraic. By Tarksi-Seidenberg (see Section 5.2), the projection onto the set (ε, v) is semialgebraic, so the map $\varepsilon \longmapsto v_\varepsilon$ is a semialgebraic function. By Proposition 2, it is also continuous, so it admits a Puiseux expansion near zero. In particular, $\varepsilon \longmapsto v_\varepsilon$ is right-differentiable at zero, and, thus, the marginal value exists (possibly being unbounded). A bound on the marginal value follows from Filar and Vrieze [9, equation (4.19)]—that is, Lemma 1—as this result implies that, for all $\varepsilon \in (0, \varepsilon_0]$,

$$\frac{|v_\varepsilon - v_0|}{\varepsilon} \le \|\tilde{g}\|_\infty + \frac{1 - \lambda}{\lambda} \|\tilde{q}\|_1 \|g + \varepsilon\tilde{g}\|_\infty + 2\frac{|\tilde{\lambda}|}{\lambda} \|g + \varepsilon\tilde{g}\|_\infty.$$

ii. Assume $\lambda = \tilde{\lambda} = 0$ and $\lim_{\varepsilon \to 0} r(q, q + \varepsilon\tilde{q}) = 0$. Consider the pairs $(\varepsilon, w) \in \mathbb{R}^2$ so that w is the value of $\Gamma + \varepsilon H$. These pairs can be expressed using first-order formulas over the real field, as follows: $\forall \delta > 0$, $\exists \lambda_0 > 0$, $\exists x$, $\exists v$ so that (x, y) and v are, respectively, a pair of optimal stationary strategies and the value of the λ-discounted version of $\Gamma + \varepsilon H$, and $|v - w| \le \delta$ for all $\lambda \in (0, \lambda_0)$. These conditions require only finitely many polynomial equations. By the theory of semialgebraic sets, the set containing such pairs is semialgebraic, so, by Tarski-Seidenberg, the graph of the map $\varepsilon \longmapsto v_\varepsilon$ is semialgebraic. By Proposition 2, this map is also continuous, so it admits a Puiseux expansion near zero. In particular, $\varepsilon \longmapsto v_\varepsilon$ is right-differentiable at zero, so the marginal value exists. Set $r_\varepsilon := r(q, q + \varepsilon\tilde{q})$. Then, by Solan [29, theorem 6]—that is, Lemma 2 setting $\varepsilon' = 0$—for all $\varepsilon \in [0, \varepsilon_0]$, one has

$$|v_\varepsilon - v_0| \le \frac{4|K| r_\varepsilon}{(1 - 2|K| r_\varepsilon)_+} \|g\|_\infty + \varepsilon \|\tilde{g}\|_\infty.$$

A bound for $\varepsilon \longmapsto \frac{v_\varepsilon - v_0}{\varepsilon}$ near zero follows then from the assumption $r_\varepsilon = O(\varepsilon)$. □

B.3. Proof of Proposition 4

Recall that this result is about a uniform bound for the discounted marginal values. Let us show that this follows directly from Solan [29, theorem 6]. By assumption, one has $r_\varepsilon = O(\varepsilon)$. Therefore, there exists $L > 0$ such that, for all $\lambda \in (0,1]$ and all ε sufficiently small,

$$\frac{|v_{\lambda,\varepsilon} - v_{\lambda,0}|}{\varepsilon} \le \frac{4|K|L}{(1-2|K|L\varepsilon)_+}\|g\|_\infty + \|\tilde{g}\|_\infty.$$

Taking ε to zero gives the desired result—that is,

$$|\partial_H v_\lambda| \le 4|K|L + \|\tilde{g}\|_\infty, \quad \forall \lambda \in (0,1]. \quad \square$$

Appendix C. Compact-Continuous Stochastic Games

In this section, we consider the extension of Theorem 1 to the compact-continuous case—that is, to discounted stochastic games over a finite state space, but where the sets of actions are compact metric sets, and the payoff and the transition functions are continuous. The existence of a value for compact-continuous discounted stochastic games follows from Shapley [26] and Sion [28]. The extension of Theorem 2 to the compact-continuous case fails. This is the case because the discounted values may fail to have a limit, as shown by Vigeral [31].

C.1. Preliminaries

In the sequel, let $\Gamma = (K, k, I, J; g, q, \lambda)$ be a discounted compact-continuous stochastic game with value v. Because action sets I and J are compact metric spaces and $\Delta(I)$ and $\Delta(J)$ denote, respectively, the sets of probability distributions over I and J, these sets are compact and convex when endowed with the weak* topology. For any $\varepsilon \ge 0$ sufficiently small, and any perturbation $H = (\tilde{g}, \tilde{q}, \tilde{\lambda})$, let $\Gamma_\varepsilon := \Gamma + \varepsilon H$ denote the perturbed game, and let v_ε denote its value.

The auxiliary games Δ_ε^0 and Δ_ε^k and $W_\varepsilon(z)$ are defined as in the finite case (see Section 2.2). In other words, recall that each pair of pure stationary strategies $(\mathbf{i}, \mathbf{j}) \in I^K \times J^K$ induces a Markov chain $Q_\varepsilon(\mathbf{i}, \mathbf{j}) \in \mathbb{R}^{K \times K}$ with payoffs $g_\varepsilon(\mathbf{i}, \mathbf{j}) \in \mathbb{R}^K$. Then, $\Delta_\varepsilon^0(\mathbf{i}, \mathbf{j}) := \det(\mathrm{Id} - (1-\lambda)Q_\varepsilon(\mathbf{i}, \mathbf{j}))$ and $\Delta_\varepsilon^k(\mathbf{i}, \mathbf{j})$ is the determinant of the matrix obtained by replacing the k-th column of $\mathrm{Id} - (1-\lambda)Q_\varepsilon(\mathbf{i}, \mathbf{j})$ with the vector $\lambda g_\varepsilon(\mathbf{i}, \mathbf{j})$. Finally, for each $z \in \mathbb{R}$, one sets:

$$W_\varepsilon[z](\mathbf{i}, \mathbf{j}) := \Delta_\varepsilon^k(\mathbf{i}, \mathbf{j}) - z\Delta_\varepsilon^0(\mathbf{i}, \mathbf{j}).$$

For $\varepsilon = 0$, we use the notation Δ^0, Δ^k, and $W(z)$.

Contrary to the finite case, the auxiliary games are no longer matrix games, but compact-continuous zero-sum games. By Sion [28, theorem 3.4], they have a value (in mixed strategies), and the sets of optimal strategies are compact and convex. In particular, this is the case for $O_1(\Gamma^*)$ and $O_2(\Gamma^*)$, the set of optimal strategies of $W(v)$.

Note that the games $\partial_H \Delta^0$ and $\partial_H \Delta^k$ are well-defined, as for each $(\mathbf{i}, \mathbf{j})$, the maps $\varepsilon \longmapsto \Delta_\varepsilon^0(\mathbf{i}, \mathbf{j})$ and $\varepsilon \to \Delta_\varepsilon^k(\mathbf{i}, \mathbf{j})$ are polynomials in ε. Hence, the following Taylor expansions hold:

$$\Delta_\varepsilon^0 = \Delta^0 + \varepsilon\,\partial_H \Delta^0 + o(\varepsilon) \quad \text{and} \quad \Delta_\varepsilon^k = \Delta^k + \varepsilon\,\partial_H \Delta^k + o(\varepsilon),$$

where, by a slight abuse in the notation, $o(\varepsilon)$ denotes a matrix of size $I^K \times J^K$ that depends on ε, satisfying $\lim_{\varepsilon \to 0} \|o(\varepsilon)\|_\infty / \varepsilon = 0$. The two appearances of $o(\varepsilon)$ do not denote the same matrix, but possibly two matrices with the same asymptotic property.

C.2. Existence of the Marginal Values

In the finite case, the existence proof of the marginal values relied on the theory of semi-algebraic sets—namely, on the fact that the map $\varepsilon \longmapsto v(\Gamma + \varepsilon H)$ is a Puiseux series near zero. However, this theory does not apply to the compact-continuous framework. Therefore, we describe an alternative approach not relying on the existence of the marginal value and extend the proof of Theorem 1 to the compact-continuous case.

C.3. Extension of Theorem 1

Let $\Gamma = (K, k, I, J; g, q, \lambda)$ be a compact-continuous discounted stochastic game with value $v := v(\Gamma)$. For each $\varepsilon > 0$, set $t_\varepsilon := \frac{1}{\varepsilon}(v_\varepsilon - v)$. We claim that $\partial_H v := \lim_{\varepsilon \to 0} t_\varepsilon$ exists and is equal to the unique $z \in \mathbb{R}$ so that

$$D_\lambda(z) := \mathrm{val}_{O^*(\Gamma)}(\partial_H \Delta^k - v\,\partial_H \Delta^0 - z\Delta^0) = 0.$$

Proof. First, we claim that the existence of at most one $z \in \mathbb{R}$ such that $D_\lambda(z) = 0$. This is the case as, for all two games M and N over the same action sets, the value function (unconstrained or constrained over some convex compact subset of strategies) satisfies:

- *Monotonicity*. $|\mathrm{val}(M) - \mathrm{val}(N)\| \le \|M - N\|_\infty$;
- *Continuity*. The map $z \longmapsto \mathrm{val}(M + zN)$ is continuous.

Further, $\|\Delta^0\|_\infty \ge \lambda^{|K|} > 0$, like in the finite case described in Section 2.2. The extension to the compact case holds because, for every stationary action profile $(\mathbf{i}, \mathbf{j}) \in I^K \times J^K$, the matrix $Q(\mathbf{i}, \mathbf{j})$ is still a Markov chain over a finite state space, and

$\Delta^0(\mathbf{i},\mathbf{j}) = \det(\mathrm{Id} - (1-\lambda)Q(\mathbf{i},\mathbf{j})) \geq \lambda^{|K|}$ like in the finite case. Altogether, these properties imply that $z \longmapsto D_\lambda(z)$ is strictly decreasing and bijective from $\mathbb{R}$ to $\mathbb{R}$, which proves the claim.

Now, let $L \in \mathbb{R}$ be the unique solution of $D_\lambda(z) = 0$. To prove that $\lim_{\varepsilon\to 0} t_\varepsilon$ exists and is equal to L, we proceed in five steps.

Step 1. A simple algebraic rearrangement yields $v_\varepsilon \Delta_\varepsilon^0 = v\Delta^0 + \varepsilon\, t_\varepsilon \Delta_\varepsilon^0 + v(\Delta_\varepsilon^0 - \Delta^0)$. Together with the Taylor expansions of Δ_ε^0 and Δ_ε^k, one obtains:

$$\begin{aligned}\Delta_\varepsilon^k - v_\varepsilon \Delta_\varepsilon^0 &= \Delta^k + \varepsilon\,\partial_H \Delta^k + o(\varepsilon) - (v\,\Delta^0 + \varepsilon\, t_\varepsilon\, \Delta_\varepsilon^0 + v\,(\varepsilon\,\partial_H \Delta^0 + o(\varepsilon)),\\ &= \Delta^k - v\Delta^0 + \varepsilon(\partial_H\Delta^k - v\,\partial_H\Delta^0 - t_\varepsilon\Delta_\varepsilon^0) + o(\varepsilon).\end{aligned}$$

Rearranging terms, and adding the term $-\varepsilon L\Delta^0$ to both sides, one thus obtains:

$$\Delta_\varepsilon^k - v_\varepsilon\Delta_\varepsilon^0 + \varepsilon(\,t_\varepsilon\Delta_\varepsilon^0 - L\Delta^0) + o(\varepsilon) = \Delta^k - v\Delta^0 + \varepsilon(\partial_H\Delta^k - v\,\partial_H\Delta^0 - L\Delta^0).$$

Step 2. By the extension of Attia and Oliu-Barton [2, theorem 1] to the compact-continuous case, $\mathrm{val}(\Delta_\varepsilon^k - v_\varepsilon\ \Delta_\varepsilon^0) = 0$ for all $\varepsilon \geq 0$. Applying Rosenberg and Sorin [25, proposition 4] to the compact-continuous game $\Delta^k - v\Delta^0$ perturbed in the direction $\partial_H\Delta^k - v\partial_H\Delta^0 - L\Delta^0$ yields

$$\lim_{\varepsilon\to 0}\frac{\mathrm{val}(\Delta^k - v\Delta^0 + \varepsilon(\partial_H\Delta^k - v\partial_H\Delta^0 - L\Delta^0))}{\varepsilon} = D_\lambda(L).$$

Step 3. By the definition of L, the right-hand side of the previous equation equals zero. Therefore,

$$\mathrm{val}(\Delta^k - v\Delta^0 + \varepsilon(\partial_H\Delta^k - v\partial_H\Delta^0 - L\Delta^0)) = o(\varepsilon).$$

Together with the last equation of Step 1, and the monotonicity of the value function, this implies:

$$\mathrm{val}(\Delta_\varepsilon^k - v_\varepsilon\Delta_\varepsilon^0 + \varepsilon(t_\varepsilon\Delta_\varepsilon^0 - L\Delta^0)) = o(\varepsilon).$$

Step 4. Next, the equality $\Delta^0 = \Delta_\varepsilon^0 + o(1)$ implies:

$$\begin{aligned}\varepsilon(t_\varepsilon\Delta_\varepsilon^0 - L\Delta^0) &= \varepsilon(t_\varepsilon\Delta_\varepsilon^0 - L(\Delta_\varepsilon^0 + o(1)))\\ &= \varepsilon\,(t_\varepsilon - L)\Delta_\varepsilon^0 + o(\varepsilon).\end{aligned}$$

Replacing in the last equation of Step 3 one thus gets:

$$\mathrm{val}(\Delta_\varepsilon^k - v_\varepsilon\Delta_\varepsilon^0 + \varepsilon(t_\varepsilon - L)\Delta_\varepsilon^0) = o(\varepsilon).$$

Step 5. Lastly, $\|\Delta_\varepsilon^0\|_\infty \geq (\lambda + \varepsilon\tilde{\lambda})^{|K|}$ for any $\varepsilon \geq 0$, as long as $q + \varepsilon\tilde{q}$ is a transition function, as noted earlier in the proof (i.e., the construction of Section 2.2 applies to the function $(\mathbf{i},\mathbf{j}) \in I^K \times J^K \longmapsto \Delta_\varepsilon^0(\mathbf{i},\mathbf{j}) \in \mathbb{R}$). Also, recall that $\mathrm{val}(\Delta_\varepsilon^k - v_\varepsilon\Delta_\varepsilon^0) = 0$ for all $\varepsilon \geq 0$ by Step 2. Hence,

$$|\mathrm{val}(\Delta_\varepsilon^k - v_\varepsilon\Delta_\varepsilon^0 + \varepsilon(t_\varepsilon - L)\Delta_\varepsilon^0)| \geq \varepsilon\,|t_\varepsilon - L|(\lambda + \varepsilon\tilde{\lambda})^{|K|}.$$

Combining with the last equation of Step 4, one obtains $|t_\varepsilon - L|(\lambda + \varepsilon\tilde{\lambda})^{|K|} = o(1)$. Hence, in particular, $\lim_{\varepsilon\to 0} t_\varepsilon = L$, which is the desired result. □

References

[1] Amir R (2003) Stochastic games in economics and related fields: An overview. Neyman A, Sorin S, eds. *Stochastic Games and Applications* (Springer, Dordrecht, Netherlands), 455–470.

[2] Attia L, Oliu-Barton M (2019) A formula for the value of a stochastic game. *Proc. Natl. Acad. Sci. USA* 116(52):26435–26443.

[3] Attia L, Oliu-Barton M (2021) Shapley-Snow kernels, multiparameter eigenvalue problems and stochastic games. *Math. Oper. Res.* 46(3):1181–1202.

[4] Attia L, Oliu-Barton M (2023) Stationary equilibria in discounted stochastic games. *Dynam. Games Appl.*, ePub ahead of print March 5, https://doi.org/10.1007/s13235-023-00495-x.

[5] Beling PA (2001) Exact algorithms for linear programming over algebraic extensions. *Algorithmica* 31(4):459–478.

[6] Bewley T, Kohlberg E (1976) The asymptotic theory of stochastic games. *Math. Oper. Res.* 1:197–208.

[7] Bewley T, Kohlberg E (1978) On stochastic games with stationary optimal strategies. *Math. Oper. Res.* 3:104–125.

[8] Catoni O, Oliu-Barton M, Ziliotto B (2021) Constant payoff in zero-sum stochastic games. *Ann. Inst. Henri Poincare Probab. Statist* 57(4):1888–1900.

[9] Filar J, Vrieze K (1997) *Competitive Markov Decision Processes* (Springer New York, New York).

[10] Freidlin MI, Wentzell AD (1984) *Random Perturbations of Dynamical Systems* (Springer Berlin, Heidelberg, Germany).

[11] Gillette D (1957) Stochastic games with zero stop probabilities. Dresher M, Tucker AW, Wolfe P, eds. *Contributions to the Theory of Games*, vol. 3 (Princeton University Press, Princeton, NJ), 179–187.

[12] Hardy GH, Littlewood JE (1914) Tauberian theorems concerning power series and Dirichlet's series whose coefficients are positive. *Proc. London Math. Soc.* 13:174–191.
[13] Kohlberg E (1974) Repeated games with absorbing states. *Ann. Statist.* 2:724–738.
[14] Krantz SG, Parks HR (2013) *The Implicit Function Theorem: History, Theory, and Applications* (Birkhäuser, New York).
[15] Mertens J-F, Neyman A (1981) Stochastic games. *Internat. J. Game Theory* 10:53–66.
[16] Mills HD (1956) Marginal values of matrix games and linear programs. Kuhn HW, Tucker AW, eds. *Linear Inequalities and Related Systems*, *Annals of Mathematics Studies*, vol. 38 (Princeton University Press, Princeton, NJ), 183–193.
[17] Neyman A, Sorin S (2010) Repeated games with public uncertain duration process. *Internat. J. Game Theory* 39:29–52.
[18] Oliu-Barton M (2014) The asymptotic value in stochastic games. *Math. Oper. Res.* 39:712–721.
[19] Oliu-Barton M (2018) The splitting game: Value and optimal strategies. *Dynam. Games Appl.* 8:157–179.
[20] Oliu-Barton M (2021) New algorithms for solving zero-sum stochastic games. *Math. Oper. Res.* 46(1):255–267.
[21] Oliu-Barton M (2022) Weighted-average stochastic games with constant payoff. *Oper. Res.* 22(3):1675–1696.
[22] Oliu-Barton M, Vigeral G (2023) Absorbing games with irrational values. *Oper. Res. Lett.* 51:555–559.
[23] Ostrowski A (1937) Sur la détermination des bornes inférieures pour une classe des determinants. *Bull. Sci. Math.* 61:19–32.
[24] Puiseux V (1850) Recherches sur les fonctions algébriques. *J. Math. Pures Appl.* (15):365–480.
[25] Rosenberg D, Sorin S (2001) An operator approach to zero-sum repeated games. *Israel J. Math.* 121(1):221–246.
[26] Shapley LS (1953) Stochastic games. *Proc. Natl. Acad. Sci. USA* 39(10):1095–1100.
[27] Shapley LS, Snow RN (1950) Basic solutions of discrete games. Kuhn HW, Tucker AW, eds. *Contributions to the Theory of Games*, vol. I, *Annals of Mathematics Studies*, vol. 24 (Princeton University Press, Princeton, NJ), 27–35.
[28] Sion M (1958) On general minimax theorems. *Pacific J. Math.* 8:171–176.
[29] Solan E (2003) Continuity of the value of competitive Markov decision processes. *J. Theoret. Probab.* 16(4):831–845.
[30] Solan E, Vieille N (2015) Stochastic games. *Proc. Natl. Acad. Sci. USA* 112(45):13743–13746.
[31] Vigeral G (2013) A zero-sum stochastic game with compact action sets and no asymptotic value. *Dynam. Games Appl.* 3:172–186.
[32] Ziliotto B (2016) A Tauberian theorem for nonexpansive operators and applications to zero-sum stochastic games. *Math. Oper. Res.* 41:1522–1534.