האוניברסיטה העברית בירושלים
THE HEBREW UNIVERSITY OF JERUSALEM

STOCHASTIC GAMES WITH SHORT-STAGE DURATION

By

ABRAHAM NEYMAN

Discussion Paper # 636   April 2013

מרכז לחקר הרציונליות

CENTER FOR THE STUDY OF RATIONALITY

Feldman Building, Givat-Ram, 91904 Jerusalem, Israel
PHONE: [972]-2-6584135   FAX: [972]-2-6513681
E-MAIL: ratio@math.huji.ac.il
URL: http://www.ratio.huji.ac.il/

Stochastic games with short-stage duration

Abraham Neyman*

April 18, 2013

Abstract

We introduce asymptotic analysis of stochastic games with short-stage duration. The play of stage k, $k \geq 0$, of a stochastic game Γ_δ with stage duration δ is interpreted as the play in time $k\delta \leq t < (k+1)\delta$, and therefore the average payoff of the n-stage play per unit of time is the sum of the payoffs in the first n stages divided by $n\delta$, and the λ-discounted present value of a payoff g in stage k is $\lambda^{k\delta} g$. We define convergence, strong convergence, and exact convergence of the data of a family $(\Gamma_\delta)_{\delta>0}$ as the stage duration δ goes to 0, and study the asymptotic behavior of the value, optimal strategies, and equilibrium. The asymptotic analogs of the discounted, limiting-average, and uniform equilibrium payoffs are defined. Convergence implies the existence of an asymptotic discounted equilibrium payoff, strong convergence implies the existence of an asymptotic limiting-average equilibrium payoff, and exact convergence implies the existence of an asymptotic uniform equilibrium payoff.

1 Introduction

Most strategic interactions evolve over time, and are often modeled as a discrete-time multi-stage game. The discrete-time modeling enables us to use the classic theory of extensive form games, which entails no conceptual difficulties. This however comes at implicit costs: players cannot change

*Institute of Mathematics, and Center for the Study of Rationality, The Hebrew University of Jerusalem, Givat Ram, Jerusalem 91904, Israel. *e-mail: aneyman@math.huji.ac.il*
This research was supported in part by Israel Science Foundation grant 1596/10.

their actions within a stage and additional information about others' actions and nature's moves is obtained only at a discrete set of times. An alternative modeling of dynamic interactions is continuous-time games, which avoids the above-mentioned costs, but entails some conceptual difficulties.

The present paper develops a complementary approach that studies the asymptotic behavior of multi-stage games when the stage duration goes to zero. We focus on the theory of stochastic games.

A discrete-time stochastic game, introduced by Shapley (1953), proceeds in stages. The stage payoff is a function $g(z, a)$ of the stage state z and the stage action a, and the transitions to the next state z' are defined by conditional probabilities $P(z' \mid z, a)$ of the next state z' given the present state z and the stage action a. Players' stage-action choices are made simultaneously and are observed by all players following the stage play.

Discrete-time stochastic games are multi-stage game-theoretic models that enable us to account for changes of states between different stages of the interaction, and where the change is impacted by the players' actions. However, no single discrete-time stochastic game can model the case where the probability of a state change in any short time interval can be positive yet arbitrarily small. This feature can be analyzed by studying continuous-time stochastic games, introduced in [14], and studied in, e.g., [14, 2, 3, 4, 8]. An alternative and complementary approach is to study the asymptotic behavior of discrete-time stochastic games, where the individual stage represents short time intervals that converge to zero and the transition probabilities to a new state also converge to zero.

The continuous-time stochastic game model provides us with a tractable analytic model (whose results are neatly stated), but, as mentioned earlier, the model entails some conceptual difficulties. The complementary asymptotic approach builds on the classic discrete-time (well-defined) game model, and therefore avoids the conceptual issues of continuous-time games. The results of the asymptotic approach supplement and cement the conclusions of the analytic continuous-time model.

We consider a family of discrete-time stochastic games Γ_δ, where the positive parameter $\delta > 0$ represents the stage duration. The sets of players N, states S, and actions A are independent of the parameter δ, and the conditional transition probabilities P_δ and the payoff function g_δ depend on the parameter δ. We study the asymptotic behavior of the strategic analysis of Γ_δ as δ goes to zero.

The payoff function g_δ describes the stage payoff in Γ_δ. As the stage duration is δ the stage payoff per unit of time is g_δ/δ. One natural condition, (g.1), on the family of discrete-time stochastic games Γ_δ is that the stage payoff function per unit of time is a function of the current state and action, and independent of δ, i.e., $g_\delta/\delta = g$, where $g : S \times A \to \mathbb{R}^N$. A less restrictive condition, (g.2), is that the stage payoff function per unit of time converges (as δ goes to zero) to a payoff function $g : S \times A \to \mathbb{R}^N$. In the asymptotic results, the distinction between assumptions (g.1) and (g.2) is immaterial.

The *transition rates*, p_δ, are the functions defined on $S \times S \times A$ by $p_\delta(z', z, a) = P_\delta(z' \mid z, a)$ if $z' \neq z$ and $p_\delta(z', z, a) = P_\delta(z' \mid z, a) - 1$ if $z' = z$. The transition rate $p_\delta(z', z, a)$ represents the difference between the probability that the next state will be z' and the probability (0 or 1) that the current state is z' when the current state is z and the current action profile is a. Note that it follows that for every (z, a) the sum of $p_\delta(z', z, a)$ over all states z' is zero and $p_\delta(z', z, a)$ is nonnegative whenever z' and z are two distinct states. It is convenient to express our conditions on the conditional transition probabilities P_δ as conditions on the transition rates p_δ.

There are several natural conditions on the transition rates function p_δ, each reflecting a dependence of p_δ on the stage duration parameter δ. One such condition, (p.1), is that the transition rates per unit of time is constant, i.e., for each $\delta > 0$, $p_\delta/\delta = \mu$, where $\mu : S \times S \times A \to \mathbb{R}$. A weaker asymptotic condition, (p.2), called *convergence*, is that the equality with μ holds in the limit, i.e., for all triples (z', z, a) of states z', z and action profile a, $p_\delta(z', z, a)/\delta$ converges (as δ goes to zero) to a limit $\mu(z', z, a)$. Condition (p.3), called *strong convergence*, requires that condition (p.2) hold and that $p_\delta(z', z, a) > 0$ if and only if $\mu(z', z, a) > 0$. Condition (p.1) implies condition (p.3) and condition (p.3) implies condition (p.2).

An *exact family* of discrete-time stochastic games Γ_δ is one that obeys (g.1) and (p.1). A family of discrete-time stochastic games Γ_δ is said to *converge in data* if it obeys (g.2) and (p.2), and it is said to *converge strongly* if it obeys (g.2) and (p.3).

The above-mentioned convergence conditions on a family $(\Gamma_\delta)_{\delta>0}$ are stated as conditions on the data of the games in the family. The data convergence condition seems natural, and therefore the study of the asymptotic behavior of equilibria of a data-convergent family is of interest. However, one may wonder if the strategic dynamics of some other families of games that do not converge in data have a limit, and therefore such families deserve an asymptotic analysis as well. This leads us to the study of convergence conditions on a family $(\Gamma_\delta)_{\delta>0}$ that depend on the stochastic processes of payoffs and states that are defined by the initial state and a strategy profile σ, in particular, when the strategy profile σ is stationary. This leads to our definition of *stationary convergence*. Roughly speaking, stationary convergence states that for every stationary strategy profile σ and real time t, both the cumulative payoff (in Γ_δ) up to time t and the distribution of the state at time t converge as the stage duration δ goes to zero.

Proposition 1 asserts that stationary convergence is equivalent to data convergence. This result shows that the continuous-time model (see, e.g., [8]) captures all possible limits of "nicely behaved" families of discrete-time stochastic games with short-stage duration.

Data (or its equivalent stationary) convergence is sufficient for our asymptotic results (e.g., Theorem 1 and Theorem 8) on the stationary (as well as the nonstationary) discounted games. In these results we associate with a discount rate ρ and a stage duration δ the discount factor $1-\rho\delta$. These results remain intact if the (δ,ρ)-dependent discount factor $\lambda_{\delta,\rho}$ is such that the limit, as δ goes to zero, of $(1-\lambda_{\delta,\rho})/\delta$ exists and equals ρ. For example, $\lambda_{\delta,\rho}=e^{-\rho\delta}$.

The unnormalized ρ-discounted payoff of a play $(z_0,a_0,z_1,\ldots)$ of the game Γ_δ is $\sum_{m=0}^{\infty}(1-\rho\delta)^m g_\delta(z_m,a_m)$. The corresponding ρ-discounted game is denoted by $\Gamma_{\delta,\rho}$. In the two-person zero-sum case, Section 4.1 shows that, given a converging family $(\Gamma_\delta)_{\delta>0}$ of two-person zero-sum games, 1) the value of $\Gamma_{\delta,\rho}$, denoted by $V_{\delta,\rho}$, converges as δ goes to zero, and 2) there is a stationary strategy σ that is $\varepsilon(\delta)$-optimal in the game $\Gamma_{\delta,\rho}$), where $\varepsilon(\delta$ goes to zero as δ goes to zero.

An *asymptotic ρ-discounted stationary equilibrium strategy* of the family $(\Gamma_\delta)_{\delta>0}$ of non-zero-sum stochastic games is a profile σ of stationary strategies that is an $\varepsilon(\delta)$-equilibrium of Γ_δ, where $\varepsilon(\delta)\to 0$ as δ goes to zero. In the discounted non-zero-sum case, we prove (Theorem 8) that (for every $\rho>0$) a converging family has an asymptotic ρ-discounted stationary equilibrium strategy.

The average (per unit of time) payoff to player i up to time s (in the game Γ_δ) is $g^i_\delta(s):=\frac{1}{s}\sum_{0\le m<s/\delta} g^i_\delta(z_m,a_m)$, where g^i_δ is the i-th coordinate of g_δ. The lim inf, respectively lim sup, game Γ_δ is the game where the payoff to player i is $\underline{g}^i_\delta:=\liminf_{s\to\infty} g^i_\delta(s)$, respectively $\bar{g}^i_\delta:=\limsup_{s\to\infty} g^i_\delta(s)$. The limiting-average value or equilibrium payoff is a payoff v such that for every $\varepsilon>0$, there is a strategy profile such that 1) for every player i, his payoff in the lim inf game is at least $v^i-\varepsilon$, and 2) every unilateral deviation of player i results in a payoff to him in the lim sup game of no more than $v^i+\varepsilon$.

For every $\delta > 0$, $v_{\delta,\rho} := \rho V_{\delta,\rho}$ converges to a limit (denoted by $v_{\delta,0}$) as $\rho \to 0+$ [1]. The limit $v_{\delta,0}$ is the uniform and limiting-average value of Γ_δ [5]. Convergence in data is not sufficient to guarantee the convergence of $v_{\delta,0}$ as δ goes to zero (Remark 10). Strong convergence implies that $v_{\delta,\rho}$ converges as δ goes to zero uniformly in ρ (Theorem 2), and therefore $v_{\delta,0}$ converges as δ goes to zero.

A family $(\Gamma_\delta)_{\delta>0}$ of two-person zero-sum stochastic games has an *asymptotic limiting-average value* v if for every $\varepsilon > 0$ there are strategies σ_δ of player 1 and τ_δ of player 2 and a duration $\delta_0 > 0$, such that for every $0 < \delta < \delta_0$, strategy σ of player 1, and strategy τ of player 2, $\varepsilon + E^z_{\sigma_\delta,\tau}\underline{g}_\delta \geq v(z) \geq -\varepsilon + E^z_{\sigma,\tau_\delta}\bar{g}_\delta$.

A family $(\Gamma_\delta)_{\delta>0}$ of non-zero-sum stochastic games has an *asymptotic limiting-average equilibrium payoff* v if for every $\varepsilon > 0$ there are strategy profiles σ_δ and a duration $\delta_0 > 0$, such that for every $0 < \delta < \delta_0$, player i, and strategy τ^i of player i,

$$\varepsilon + E^z_{\sigma_\delta}\underline{g}^i_\delta \geq v^i(z) \geq -\varepsilon + E^z_{\sigma_\delta^{-i},\tau^i}\bar{g}^i_\delta(s).$$

A family $(\Gamma_\delta)_{\delta>0}$ that converges strongly has an asymptotic limiting-average value in the zero-sum case (Theorem 4), and an asymptotic limiting-average equilibrium payoff in the non-zero-sum case (Theorem 11).

A family $(\Gamma_\delta)_{\delta>0}$ of two-person zero-sum stochastic games has an *asymptotic uniform value* v if for every $\varepsilon > 0$ there are strategies σ_δ of player 1 and τ_δ of player 2, a duration $\delta_0 > 0$, and a time $s_0 > 0$, such that for every $0 < \delta < \delta_0$, $s > s_0$, strategy σ of player 1, and strategy τ of player 2, $\varepsilon + E^z_{\sigma_\delta,\tau} g_\delta(s) \geq v(z) \geq -\varepsilon + E^z_{\sigma,\tau_\delta} g_\delta(s)$.

A family $(\Gamma_\delta)_{\delta>0}$ of non-zero-sum stochastic games has an *asymptotic uniform equilibrium payoff* v if for every $\varepsilon > 0$ there are strategy profiles σ_δ, a duration $\delta_0 > 0$, and a time $s_0 > 0$, such that for every $0 < \delta < \delta_0$, $s > s_0$, player i, and strategy τ^i of player i,

$$\varepsilon + E^z_{\sigma_\delta} g^i_\delta(s) \geq v^i(z) \geq -\varepsilon + E^z_{\sigma_\delta^{-i},\tau^i} g^i_\delta(s).$$

An exact family of games Γ_δ has an asymptotic uniform value in the zero-sum case (Theorem 6), and an asymptotic uniform equilibrium payoff in the non-zero-sum case (Theorem 12).

2 The model and results

Throughout the paper, the set of players N, the set of states S, and the set of actions A, are finite. The set of feasible actions may depend on the state $z \in S$. We denote by $A^i(z)$ the set of actions of player $i \in N$ in state $z \in S$. $A(z)$ is the set of action profiles at state z, $A(z) = \times_{i \in N} A^i(z)$. For notational convenience we set $\mathcal{A} = \{(z,a) : z \in S,\ a \in A(z)\}$.

The data of the stochastic game Γ_δ that depend on the parameter δ are the $\mathbb{R}^N$-valued payoff function g_δ that is defined on $\mathcal{A}$ and the conditional probabilities $P_\delta(z' \mid z,a)$ that are defined for all $z' \in S$ and $(z,a) \in \mathcal{A}$. The payoff function g_δ defines the stage payoff $g_\delta(z,a) \in \mathbb{R}^N$ as a function of the stage state z and the stage action profile a. The i-th coordinate of a vector $g \in \mathbb{R}^N$ is denoted by g^i. The conditional probabilities $P_\delta(z' \mid z,a)$ specify the conditional probability of the next state being z' conditional on playing the action profile a at the current state z.

The conditional probabilities $P_\delta(z' \mid z,a)$ obey $P_\delta(z' \mid z,a) \geq 0$ and $\sum_{z' \in S} P_\delta(z' \mid z,a) = 1$. We describe the conditional probabilities by specifying the function $p_\delta(z',z,a)$ that is defined on $S \times \mathcal{A}$ by $p_\delta(z',z,a) = P_\delta(z' \mid z,a)$ if $z' \neq z$ and $p_\delta(z',z,a) = P_\delta(z' \mid z,a) - 1$ if $z' = z$. Obviously, $p_\delta(z',z,a) \geq 0$ if $z' \neq z$, $p(z,z,a) \geq -1$, and $\sum_{z' \in S} p_\delta(z',z,a) = 0$.

The set H of *plays* of Γ_δ is the set of all sequences $h = (z_0, a_0, \ldots, z_k, a_k, \ldots)$ with $(z_k, a_k) \in \mathcal{A}$. The events are the elements of the minimal σ-algebra $\mathcal{H}$ of subsets of H for which each one of the maps $H \ni h = (z_0, a_0, \ldots) \mapsto (z_k, a_k) \in \mathcal{A}$, $k \geq 0$, is measurable. We denote by $\mathcal{H}_k$ the σ-algebra generated by $(z_0, a_0, \ldots, z_k)$.

The set of strategies in the stochastic game Γ_δ is independent of δ. The transition probabilities, however, do depend on δ. For every strategy profile $\sigma = (\sigma^i)_{i \in N}$ we denote by $P^z_{\delta,\sigma}$ the probability distribution defined by the transition probabilities of the game Γ_δ, the initial state $z_0 = z$, and the strategy profile σ, on the measurable space $(H, \mathcal{H})$ of plays. The expectation with respect to the probability $P^z_{\delta,\sigma}$ is denoted by $E^z_{\delta,\sigma}$. The parameter δ that appears in the probability and expectation above is formally needed as the transition probabilities depend on δ. However, wherever there is an implicit reference to the parameter δ, we suppress (the formally needed) δ. E.g., we write $E^z_{\sigma_\delta}$, for short, instead of the more explicit $E^z_{\delta,\sigma_\delta}$.

2.1 The discounted games

Given a discount factor $0 < \lambda < 1$, the discrete-time stochastic game Γ with a discount factor λ is the game where the (unnormalized) valuation of the stream of payoffs $(g_m = g(z_m, a_m))_{m\geq 0}$ is $\sum_{m=0}^{\infty} \lambda^m g_m$. The normalized valuation is the unnormalized one times $1-\lambda$. The generalization to the case of individual discount factors is straightforward. Given a vector $\overrightarrow{\lambda} = (\lambda_i)_{i\in N}$ of discount factors the game with discount factors $\overrightarrow{\lambda}$ is the game where the unnormalized (respectively, normalized) valuation of player i of the stream of vector payoffs $(g_m)_{m\geq 0}$ is $\sum_{m=0}^{\infty} \lambda_i^m g_m^i$ (respectively, $(1-\lambda_i)\sum_{m=0}^{\infty} \lambda_i^m g_m^i$).

We study the family of discrete-time stochastic games Γ_δ with discount factors λ_δ that depend on the stage duration parameter δ. We require that the limit, as δ goes to zero, of the valuation of a unit payoff per unit of time (i.e., $g_\delta = \delta$ for all $\delta > 0$) with the discount factor λ_δ, exist. This requirement is equivalent to the existence of the limit of $\frac{1-\lambda_\delta}{\delta}$ as δ goes to zero. A family of δ-dependent discount factors λ_δ is called *admissible* if $\lim_{\delta\to 0+} \frac{1-\lambda_\delta}{\delta}$ exists. The limit is called the *asymptotic discount rate* (and is equal to $\lim_{\delta\to 0+} \frac{-\ln \lambda_\delta}{\delta}$). Two examples of admissible δ-dependent discount factors, with asymptotic discount rate $\rho > 0$, are $\lambda_\delta = e^{-\rho\delta}$ and $\lambda_\delta = 1-\rho\delta$.

A family of δ-dependent discount factors, λ_δ, is admissible and has an asymptotic discount rate $\rho > 0$, if and only if for all streams $x_\delta = (g_{\delta,0}, g_{\delta,1}, \ldots)$ of payoffs, with uniformly bounded payoffs per unit of time (i.e., $|g_{\delta,m}| \leq C\delta$), the difference between the valuation of x_δ according to the discount factors λ_δ and its valuation according to the discount factors $e^{-\rho\delta}$ goes to zero as δ goes to zero.

Our asymptotic results on the δ-dependent discounted games depend only on the asymptotic discount rate ρ (and not on the exact choice of the δ-dependent discount factor with asymptotic discount rate ρ). Therefore, it suffices to select, for each $\rho > 0$, an admissible family of δ-dependent discount factors $\lambda_{\delta,\rho}$ with asymptotic discount rate ρ. Our choice of the δ-dependent discount factor with asymptotic discount rate ρ is $\lambda_{\delta,\rho} = 1-\rho\delta$. This simplifies some parts of the presentation.

The ρ-discounted game, denoted by $\Gamma_{\delta,\rho}$, is the game Γ_δ with discount factor $1-\rho\delta$. In the zero-sum case, we say that the family $(\Gamma_\delta)_{\delta>0}$ of two-person zero-sum games[1] has an *asymptotic ρ-discounted value V_ρ* if the values

[1] Henceforth, whenever we discuss a value concept of a family (Γ_δ), we will omit the statement of the implicit condition that it is a family of two-person zero-sum games.

of $\Gamma_{\delta,\rho}$, denoted by $V_{\delta,\rho}$, converge to V_ρ as δ goes to zero. Theorem 1 asserts that a family $(\Gamma_\delta)_{\delta>0}$ that converges in data has an asymptotic ρ-discounted value. In addition, it provides a system of S equations that has a unique solution, which equals V_ρ, and proves the existence of a (δ-independent) stationary strategy that is $\varepsilon(\delta)$-optimal in $\Gamma_{\delta,\rho}$, where $\varepsilon(\delta) \to 0$ as δ goes to zero. In the non-zero-sum case, Theorem 8 asserts that a family $(\Gamma_\delta)_{\delta>0}$ that converges in data has a (δ-independent) stationary strategy that is an $\varepsilon(\delta)$-equilibrium of $\Gamma_{\delta,\rho}$, where $\varepsilon(\delta) \to 0$ as δ goes to zero.

Section 4.1 notes that the map $\rho \mapsto V_\rho$ is semialgebraic and bounded, and therefore $v_\rho := \rho V_\rho = \sum_{k=0}^\infty c_k(z)\rho^{k/M}$ in a right neighborhood of zero. This fact, in conjunction with the covariance properties of v_ρ as a function of (g,μ) (see Section 4.1), is used in the study of the asymptotic uniform value (see Section 4.5). It shows that for an exact family $(\Gamma_\delta)_{\delta>0}$ there is an integrable function $\psi : [0,1] \to \mathbb{R}_+$ and $\delta_0 > 0$ such that $\|\rho V_{\delta,\rho} - \rho' V_{\delta,\rho'}\| \leq \int_\rho^{\rho'} \psi(x)\,dx$ for $0 < \rho < \rho' \leq 1$ and $\delta \leq \delta_0$.

The covariance properties (in conjunction with [10, Theorem 6]) are used in the proof of Theorem 2 that asserts that if Γ_δ converges strongly, then $v_{\delta,\rho}$ ($:= \rho V_{\delta,\rho}$) converges, as δ goes to zero, uniformly on $0 < \rho < 1$.

2.2 The nonstationary discounted games

A time-separable valuation u of streams of payoff is represented by a positive measure w on the nonnegative integers. It is given by the valuation function $u_w(g_0, g_1, \ldots) = \sum_{m=0}^\infty w(m) g_m$. The valuation function u_w is (well) defined over all bounded streams $(g_0, g_1, \ldots)$ of payoffs. The valuation u_w is normalized if the total mass of w equals 1, i.e., $\sum_{m=0}^\infty w(m) = 1$. The generalization to the case of individual time-separable valuations is straightforward. Given a vector $\overrightarrow{w} = (w^i)_{i\in N}$ of positive measures on the nonnegative integers the game with valuation $u_{\overrightarrow{w}}$ is the game where the valuation of player i of the stream of vector payoffs $(g_m)_{m\geq 0}$ is $\sum_{m=0}^\infty w^i(m) g_m^i$. The discrete-time stochastic game Γ with the valuation $u_{\overrightarrow{w}}$ is denoted by $\Gamma_{\overrightarrow{w}}$.

The set of all probability measures on a set $*$ is denoted by $\Delta(*)$. As $A^i(z)$ is finite, the set $X^i(z) := \Delta(A^i(z))$ is a compact subset of a Euclidean space. The set of profiles of Markovian strategies in a discrete-time stochastic game is identified with the cartesian product $\times_{(i,z,n)\in N\times S\times\mathbb{N}} X^i(z)$, which is a compact space in the product topology. Let Γ be a discrete-time stochastic game (with finitely many states and actions). A profile σ of Markovian strategies is an equilibrium of $\Gamma_{\overrightarrow{w}}$ whenever: 1) for every $k \in \mathbb{N}$, $\overrightarrow{w}_k$ is a vector of positive measures on the nonnegative integers, 2) for every $k \in \mathbb{N}$, $\sigma(k)$ is a profile of Markovian strategies that is an equilibrium of $\Gamma_{\overrightarrow{w}_k}$, 3) $\sigma(k) \to_{k\to\infty} \sigma$ (in the product topology), and 4) for every $i \in N$, $\sum_{m=0}^{\infty} |w_k^i(m) - w^i(m)| \to_{k\to\infty} 0$.

By backward induction, if $\overrightarrow{w}$ has finite support, the game $\Gamma_{\overrightarrow{w}}$ has an equilibrium in Markovian strategies. Therefore, the above-mentioned comment implies that a discrete-time stochastic game with individual time-separable evaluations has an equilibrium in Markovian strategies. The discrete-time stochastic game Γ_δ with the individual time-separable valuation $\overrightarrow{w}_\delta$ is denoted by $\Gamma_{\delta,\overrightarrow{w}_\delta}$. In this game, the payoff to player i of a play $(z_0, a_0, \dots)$ is $g_\delta^i(w_\delta^i) := \sum_{m=0}^{\infty} w_\delta^i(m) g_\delta^i(z_m, a_m)$. The discrete-time stochastic game Γ_δ with the common time-separable valuation w_δ, denoted by Γ_{δ,w_δ}, is the game $\Gamma_{\delta,\overrightarrow{w}_\delta}$ with $w_\delta^i = w_\delta$ for every player i.

If $\overrightarrow{w} = (w^i)_{i\in N}$ is a profile of nonnegative measures on $[0,\infty]$, we say that the vector $\overrightarrow{w}_\delta = (w_\delta^i)_{i\in N}$ of N measures on $\mathbb{N}\cup\{\infty\}$ *converges* (as $\delta \to 0+$) to $\overrightarrow{w}$ if 1) $\overrightarrow{w}_\delta(\mathbb{N}\cup\{\infty\})$ converges (as δ goes to 0) to $\overrightarrow{w}([0,\infty])$, and 2) for every $0 \le t < \infty$ there is a family of nonnegative integers m_δ with $\delta m_\delta \to_{\delta\to 0+} t$, and such that $\sum_{m=0}^{m_\delta} \overrightarrow{w}_\delta(m) \to_{\delta\to 0+} \overrightarrow{w}([0,t])$. Note that by identifying the N-vector measure $\overrightarrow{w}_\delta$ with the N-vector measure $\overrightarrow{w}'_\delta$ on $[0,\infty]$ (the one-point compactification of $[0,\infty)$) that is supported on $\{\delta m : m \ge 0\} \cup \{\infty\}$ and satisfies $\overrightarrow{w}'_\delta([\delta m, \delta(m+1))) = \overrightarrow{w}_\delta(m)$ and $\overrightarrow{w}'_\delta(\infty) = \overrightarrow{w}_\delta(\infty)$, our definition of convergence here is equivalent to w^* convergence of measures on compact spaces. Explicitly, $\overrightarrow{w}_\delta$ converges as $\delta \to 0+$ to the N-vector measure $\overrightarrow{w}$ on $[0,\infty]$ if for every continuous function f on $[0,\infty]$, $\int_{[0,\infty]} f(x)\, d\overrightarrow{w}'_\delta(x)$ (which equals $f(\infty)\overrightarrow{w}_\delta(\infty) + \sum_{m=0}^{\infty} f(\delta m)\overrightarrow{w}_\delta(\delta m)$) converges as $\delta \to 0+$ to $\int_{[0,\infty]} f(x)\, d\overrightarrow{w}(x)$.

In this section we focus on the case that $\overrightarrow{w}_\delta$ is supported on $\mathbb{N}$ and $\overrightarrow{w}$ is supported on $[0,\infty)$. The more general convergence definition (above) is used in subsequent parts of the paper.

Of special interest are the nonstationary discounting valuations and their limits. In the discrete-time model, the nonnegative measure w on $\mathbb{N}\cup\{\infty\}$ is called a *nonstationary discounting* valuation (measure) if $w(m) \ge w(m+1)$. The vector measure $\overrightarrow{w}$ is said to be *nonstationary discounting* if each of its components w^i is a nonstationary discounting. A nonnegative measure w on $[0,\infty]$ is said to be *nonstationary discounting* if for every $s > 0$ the function $[0,\infty) \ni t \mapsto w([t, t+s)$ is nonincreasing in t. Note that if the family of nonstationary discounting measures w_δ on $\mathbb{N}$ converges to the nonnegative measure w on $[0,\infty]$, then w is a nonstationary discounting measure.

Let $\overrightarrow{w}$ be a nonstationary discounting N-vector measure on $[0,\infty)$. We say that $v \in \mathbb{R}^{N\times S}$ is an *asymptotic* $\overrightarrow{w}$ *equilibrium payoff* of the family of N-person games $(\Gamma_\delta)_{\delta>0}$, if for every $\varepsilon > 0$ and a family of nonstationary discounting N-vector measures $\overrightarrow{w}_\delta$ on $\mathbb{N}$ that converges to $\overrightarrow{w}$, v is an ε-equilibrium payoff of $\Gamma_{\delta,\overrightarrow{w}_\delta}$ for every $\delta > 0$ sufficiently small.

Let w be a nonstationary discounting measure on $[0,\infty)$. We say that $v \in \mathbb{R}^S$ is an *asymptotic w value* of the family of two-person zero-sum games $(\Gamma_\delta)_{\delta>0}$, if for every $\varepsilon > 0$ and a family of nonstationary discounting measures w_δ on $\mathbb{N}$ that converges to w, the value v_δ of Γ_{δ,w_δ} satisfies $|v_\delta(z)-v(z)| < \varepsilon$ for every $\delta > 0$ sufficiently small and state z. Note that $v \in \mathbb{R}^S$ is an asymptotic w value of the family of two-person zero-sum games $(\Gamma_\delta)_{\delta>0}$ if and only if $(v,-v)$ is an asymptotic (w,w) equilibrium payoff of $(\Gamma_\delta)_{\delta>0}$.

Theorem 9 asserts (in particular) that if $(\Gamma_\delta)_{\delta>0}$ converges in data, then for every nonstationary discounting N-vector measure $\overrightarrow{w}$ on $[0,\infty)$ the family $(\Gamma_\delta)_{\delta>0}$ has an asymptotic $\overrightarrow{w}$ equilibrium payoff. In addition, if the nonstationary discounting N-vector measure $\overrightarrow{w}_\delta$ converges (as δ goes to 0) to the N-vector measure $\overrightarrow{w}$ on $[0,\infty)$, then for every $\varepsilon > 0$ there is $\delta_0 > 0$ and a family of Markovian strategy profiles σ_δ, such that 1) for $0 < \delta < \delta_0$, σ_δ is an ε-equilibrium of $\Gamma_{\delta,\overrightarrow{w}_\delta}$ and its corresponding payoff is within ε of an asymptotic $\overrightarrow{w}$ equilibrium payoff v, and 2) σ_δ converges to a profile of continuous-time Markov strategies[2]. In Section 3.2 we define the convergence of Markovian strategies.

Theorem 9 implies in particular that a finite-horizon continuous-time stochastic game has an ε-equilibrium in Markov strategies. [4] shows that a finite-horizon continuous-time stochastic game need not have an equilibrium in Markov strategies. Therefore, it is impossible to require (in the additional part) that σ_δ be an equilibrium (rather than an ε-equilibrium) of $\Gamma_{\delta,\overrightarrow{w}_\delta}$ and at the same time converge to a profile of continuous-time Markov strategies.

In several dynamic interactions, the game payoff is composed of stage payoffs and a terminal payoff. Such games are also useful in backward induction arguments. For example, in order to find an equilibrium (or an approximate equilibrium) of an extensive form game, a classical procedure is to replace a subgame of the game with a terminal node whose payoff equals an equilibrium (or approximate equilibrium) payoff of the subgame. An equilibrium (or ap-

[2] A continuous-time strategy σ is a mixed-action-valued measurable function defined on $S \times \mathbb{R}$.

proximate equilibrium) of the original game is obtained by patching together an equilibrium (or an approximate equilibrium) of the truncated game with an equilibrium (or approximate equilibrium) of the subgame. This motivates the definition of the following useful family of games.

Let $\overrightarrow{w}_\delta = (w^i_\delta)_{i\in N}$ be a vector of positive measures on $\mathbb{N}$, $m_\delta > 0$, and let $\nu_\delta = (\nu^i_\delta)_{i\in N}$ be a vector of N payoff functions $\nu^i_\delta : \mathcal{A} \to \mathbb{R}$. The game $\Gamma^{m_\delta,\nu_\delta}_{\delta,\overrightarrow{w}_\delta}$ is the game Γ_δ where the valuation of player i of the play $(z_0, a_0, z_1, \ldots)$ is the sum of two terms: $\nu^i_\delta(z_{m_\delta}, a_{m_\delta}) + \sum_{m=0}^{\infty} w^i_\delta(m) g^i_\delta(z_m, a_m)$. The first term accounts for a one-time (e.g., terminal) payoff. This variation enables us to view games like soccer, where the objective is to reach the best score at the end of the game, as stochastic games.

We say that (m_δ, ν_δ) *converges* to (t, ν), where $0 \leq t < \infty$ and $\nu : \mathcal{A} \to \mathbb{R}^N$, if 1) $\nu_\delta(z, a)$ converges to $\nu(z, a)$ for all $(z, a) \in \mathcal{A}$, and 2) δm_δ converges to t as δ goes to zero.

Let $\overrightarrow{w}$ be a nonstationary discounting N-vector measure on $[0, \infty)$, $0 \leq t < \infty$, and $\nu : \mathcal{A} \to \mathbb{R}^N$. The $N \times S$ payoff vector $v \in \mathbb{R}^{N\times S}$ is called an *asymptotic* $(\overrightarrow{w}, t, \nu)$ *equilibrium payoff* of the family $(\Gamma_\delta)_{\delta>0}$, if for every 1) family of nonstationary discounting N-vector measure $\overrightarrow{w}_\delta$ on $\mathbb{N}$ that converges (as δ goes to 0) to $\overrightarrow{w}$, 2) $m_\delta \in \mathbb{N}$ and $\nu_\delta : \mathcal{A} \to \mathbb{R}^N$ such that (m_δ, ν_δ) converges to (t, ν), and 3) $\varepsilon > 0$, there is $\delta_0 > 0$, such that for $0 < \delta < \delta_0$, $\Gamma^{m_\delta,\nu_\delta}_{\delta,\overrightarrow{w}_\delta}$ has an ε-equilibrium payoff within ε of v.

Theorem 9 asserts if 1) $\overrightarrow{w}$ is a nonstationary discounting N-vector measure on $[0, \infty)$, 2) $0 \leq t < \infty$, and 3) $\nu : \mathcal{A} \to \mathbb{R}^N$, then a family $(\Gamma_\delta)_{\delta>0}$ that converges in data has an asymptotic $(\overrightarrow{w}, t, \nu)$ equilibrium payoff. In addition, if 1) $\overrightarrow{w}_\delta$ is a nonstationary discounting N-vector measure on $\mathbb{N}$ that converges (as δ goes to 0) to $\overrightarrow{w}$, and 2) $m_\delta \in \mathbb{N}$ and $\nu_\delta : \mathcal{A} \to \mathbb{R}^N$ are such that (m_δ, ν_δ) converges to (t, ν), then for every $\varepsilon > 0$ there are 1) $\delta_0 > 0$, 2) Markov strategy profiles σ_δ, and 3) a continuous-time Markov strategy profile σ, such that 1) for $0 < \delta < \delta_0$, σ_δ is a ε-equilibrium of $\Gamma^{m_\delta,\nu_\delta}_{\delta,\overrightarrow{w}_\delta}$ with a payoff within ε of an asymptotic $(\overrightarrow{w}, t, \nu)$ equilibrium payoff v, and 2) the Markov strategy profiles σ_δ converge w^* to σ.

2.3 The limiting-average games

The classic limiting-average valuation of a stream $(g_0, g_1, \ldots)$ of payoffs is the limit of the average payoff per stage, $\lim_{n\to\infty} \frac{1}{n} \sum_{0\leq m<n} g_m$, if the limit exists. The interpretation is that the stage duration is one unit of time, and

therefore the average $\frac{1}{n}\sum_{0\le m<n} g_m$ represents the average payoff per unit of time. In studying the limiting-average valuation of streams $(g_{\delta,0}, g_{\delta,1}, \ldots)$ of payoffs in Γ_δ, one has to take into account that the stage duration is δ. Therefore the average payoff per unit of time up to time s is $(g^i_\delta(s))_{i\in N} = g_\delta(s)$ $(= \frac{1}{s}\sum_{m:0\le m\delta<s} g_{\delta,m})$. In the two-person zero-sum case, the set of players is $N = \{1,2\}$ and we write g for g^1 and g_δ for g^1_δ. No confusion should result.

The averages $g^i_\delta(s)$ need not converge as s goes to infinity. Therefore, in defining the limiting-average (value or) equilibrium payoff $v = (v^i)_{i\in N}$, we require that for every $\varepsilon > 0$ the (ε-optimal or) ε-equilibrium strategy result in a distribution on streams of payoffs such that the expectation of $\underline{g}^i_\delta$ $(= \liminf_{\delta\to 0+} g^i_\delta(s))$ is within ε of v, and no unilateral deviation by a player, say player i, can result in a distribution on streams of payoffs with an expectation of $\bar{g}^i_\delta$ $(= \limsup g^i_\delta(s))$ greater than $v^i + \varepsilon$.

Note that if $w_{\delta,s}$ is the probability measure on $\mathbb{N}$ with $w_{\delta,s}(m) = 1/\lceil s/\delta\rceil$ (where $\lceil * \rceil$ denotes the smallest positive integer that is $\ge *$) if $m\delta < s$ and $w_{\delta,s}(m) = 0$ otherwise, then $g^i_\delta(s) = g^i_\delta(w_{\delta,s})$. For each $\delta > 0$, the probability measures $w_{\delta,s}$, $s > 0$, are the extreme points of the convex set $M^1_d(\mathbb{N})$ of nonstationary discounting probability measures w_δ on $\mathbb{N}$. Indeed, $w_\delta = \sum_{m=1}^\infty (w_\delta(m-1) - w_\delta(m))m w_{\delta,m\delta}$ and $\sum_{m=1}^\infty (w_\delta(m-1) - w_\delta(m))m = 1$. As $\sum_{m=1}^k (w_\delta(m-1) - w_\delta(m))m \le w_\delta(0)k^2 \to_{w_\delta(0)\to 0+} 0$, we deduce the following (known) property of the lim inf valuation $\underline{g}^i_\delta$ and the lim sup valuation $\bar{g}^i_\delta$.

$$\begin{aligned} \underline{g}^i_\delta &= \lim_{\eta\to 0+} \inf\{g^i_\delta(w_\delta) : w_\delta \in M^1_d(\mathbb{N}) \text{ with } w_\delta(0) < \eta\}, \text{ and} \\ \bar{g}^i_\delta &= \lim_{\eta\to 0+} \sup\{g^i_\delta(w_\delta) : w_\delta \in M^1_d(\mathbb{N}) \text{ with } w_\delta(0) < \eta\}. \end{aligned}$$

A two-person zero-sum discrete-time stochastic game (with finitely many states and actions) has a limiting-average value [5]. However, this does not imply that a convergent family $(\Gamma_\delta)_{\delta>0}$ has an asymptotic limiting-average value. A non-zero-sum discrete-time stochastic game (with finitely many states and actions) has a limiting-average correlated equilibrium payoff [11], but it is unknown if it has a limiting-average equilibrium payoff.

Recall that $v \in \mathbb{R}^S$ is an asymptotic limiting-average value of the family $(\Gamma_\delta)_{\delta>0}$ if for every $\varepsilon > 0$ there are strategies σ_δ of player 1 and τ_δ of player 2 and a duration $\delta_0 > 0$, such that for every strategy τ of player 2, strategy σ of player 1, and $0 < \delta < \delta_0$, we have

$$\varepsilon + E^z_{\sigma_\delta,\tau}\underline{g}_\delta \ge v(z) \ge -\varepsilon + E^z_{\sigma,\tau_\delta}\bar{g}_\delta.$$

The definition implies that a family $(\Gamma_\delta)_{\delta>0}$ has at most one asymptotic limiting-average value.

Recall that $v \in \mathbb{R}^{N\times S}$ is an asymptotic limiting-average equilibrium payoff of the family $(\Gamma_\delta)_{\delta>0}$ if for every $\varepsilon > 0$ there are strategy profiles σ_δ and a duration $\delta_0 > 0$, such that for every strategy τ^i of player i and every $0 < \delta < \delta_0$, we have

$$\varepsilon + E^z_{\sigma_\delta}\underline{g}^i_\delta \geq v^i(z) \geq -\varepsilon + E^z_{\sigma^{-i}_\delta,\tau^i}\bar{g}^i_\delta.$$

We prove that a family $(\Gamma_\delta)_{\delta>0}$ that converges strongly has an asymptotic limiting-average value in the zero-sum case (Theorem 4), and an asymptotic limiting-average equilibrium payoff in the non-zero-sum case (Theorem 11).

A variation of the limiting-average value, respectively, limiting-average equilibrium payoff, is the *weak limiting-average value*, respectively *weak limiting-average equilibrium payoff*, obtained by exchanging the order of the limiting and the expectation operations. Therefore, we say that $v \in \mathbb{R}^{N\times S}$ is an *asymptotic weak limiting-average equilibrium payoff* of the family $(\Gamma_\delta)_{\delta>0}$ if for every $\varepsilon > 0$ there are strategy profiles σ_δ and a duration $\delta_0 > 0$, such that for every strategy τ^i of player i and every $0 < \delta < \delta_0$, we have

$$\varepsilon + \liminf_{s\to\infty} E^z_{\sigma_\delta} g^i_\delta(s) \geq v^i(z) \geq -\varepsilon + \limsup_{s\to\infty} E^z_{\sigma^{-i}_\delta,\tau^i} g^i_\delta(s).$$

In the general model of repeated games (which includes repeated games with incomplete information), the existence of a limiting-average (value or) equilibrium payoff implies the existence of a weak limiting-average (value or) equilibrium payoff, but not vice versa. In the game models studied in the present paper, all results that we can prove regarding the weak limiting-value hold also for the limiting-average value. Therefore, no special consideration is given to these weaker concepts. It should be noted, however, that in the analogous study of the general model of repeated games, in particular in repeated games with incomplete information, the limiting-average value or equilibrium payoff will typically not exist, while the weak limiting-average value and equilibrium payoff may exist in some of these models.

2.4 The mixed discounting and limiting-average games

The mixed time-separable and the limiting-average (respectively, the weak limiting-average) valuation of payoffs is a positive linear combination of a time-separable valuation u_w and the limiting-average (respectively, the weak limiting-average) valuation. It is represented by a measure w on $\mathbb{N} \cup \{\infty\}$, where $w(\infty)$ represents the weight given to the limiting-average (or weak limiting-average) valuation, and $w(m)$ represents the weight of the payoff at stage $m \in \mathbb{N}$. A normalized mixed time-separable and limiting-average (or weak limiting-average) valuation of payoffs is a convex combination of a normalized time-separable valuation u_w and the limiting-average (or the weak limiting-average) valuation, and is represented by a probability measure on $\mathbb{N} \cup \{\infty\}$.

Let $\overrightarrow{w}_\delta = (w^i)_{i \in N}$ be a vector of positive measures on $\mathbb{N} \cup \{\infty\}$, $m_\delta > 0$, and let $\nu_\delta = (\nu^i_\delta)_{i \in N}$ be a vector of N payoff functions $\nu^i_\delta : \mathcal{A} \to \mathbb{R}$. The game $\Gamma^{m_\delta, \nu_\delta}_{\delta, \overrightarrow{w}_\delta}$ is the game Γ_δ where the valuation of player i of the play $(z_0, a_0, z_1, \ldots)$ is the sum of three terms

$$\nu^i_\delta(z_{m_\delta}, a_{m_\delta}) + w^i_\delta(\infty) \lim_{s \to \infty} g^i_\delta(s) + \sum_{m=0}^{\infty} w^i_\delta(m) g^i_\delta(z_m, a_m),$$

if the limit exists.

The limit of $g^i_\delta(s)$ as $s \to \infty$ need not exist. Therefore, in defining (the value or) an equilibrium payoff v of $\Gamma^{m_\delta, \nu_\delta}_{\delta, \overrightarrow{w}_\delta}$, we require that for every $\varepsilon > 0$ the (ε-optimal or) ε-equilibrium strategy result in a distribution on plays such that the expectation of the $\nu^i_\delta(z_{m_\delta}, a_{m_\delta}) + w^i_\delta(\infty) \underline{g}^i_\delta + \sum_{m=0}^{\infty} w^i_\delta(m) g^i_\delta(z_m, a_m)$ is within ε of v^i, and no unilateral deviation by a player, say player i, can result in a distribution on plays with an expectation of $\nu^i_\delta(z_{m_\delta}, a_{m_\delta}) + w^i_\delta(\infty) \bar{g}^i_\delta + \sum_{m=0}^{\infty} w^i_\delta(m) g^i_\delta(z_m, a_m)$ greater than $v^i + \varepsilon$.

Theorem 13 asserts that if 1) $(\Gamma_\delta)_{\delta > 0}$ is an exact family, 2) the nonstationary discounting N-vector measure $\overrightarrow{w}_\delta$ converges (as δ goes to 0) to the N-vector measure $\overrightarrow{w}$ on $[0, \infty]$, and 3) (m_δ, ν_δ) converges to (t, ν), then for every $\varepsilon > 0$ there are strategy profiles σ_δ, an $N \times S$ vector v, and $\delta_0 > 0$, such that for $0 < \delta < \delta_0$, σ_δ is an ε-equilibrium of $\Gamma^{m_\delta, \nu_\delta}_{\delta, \overrightarrow{w}_\delta}$ with a payoff within ε of v.

2.5 The uniform games

In a uniform (value or) equilibrium payoff v, we require that for every $\varepsilon > 0$ there be a time s_0 and a strategy profile for which for every $s > s_0$ the expectation of $g_\delta(s)$ is within ε of v, and that there be no unilateral deviation by a player, say player i, and a time $s > s_0$ such that the expectation of $g^i_\delta(s)$

is more than $v^i + \varepsilon$. It is known that a uniform value exists in the zero-sum case (with finitely[3] many states and actions) [5]. In the discrete-time non-zero-sum case (with finitely many states and actions), (a uniform correlated equilibrium payoff exists [11], but) it is unknown if a uniform equilibrium payoff exists in this case.

We say that $v \in \mathbb{R}^S$ is an *asymptotic uniform value* of the family $(\Gamma_\delta)_{\delta>0}$ if for every $\varepsilon > 0$ there are 1) a time $s_0 > 0$, 2) a duration $\delta_0 > 0$, and 3) strategies σ_δ of player 1 and τ_δ of player 2, such that for all strategies τ of player 2 and σ of player 1, duration $0 < \delta < \delta_0$, and time $s > s_0$, we have

$$\varepsilon + E^z_{\sigma_\delta,\tau} g_\delta(s) \geq v(z) \geq -\varepsilon + E^z_{\sigma,\tau_\delta} g_\delta(s).$$

The definition implies that a family $(\Gamma_\delta)_{\delta>0}$ has at most one asymptotic uniform value.

Similarly, we say that $v \in \mathbb{R}^{N\times S}$ is an *asymptotic uniform equilibrium payoff* of the family $(\Gamma_\delta)_{\delta>0}$ if for every $\varepsilon > 0$ there are 1) a time $s_0 > 0$, 2) a duration $\delta_0 > 0$, and 3) strategy profiles σ_δ, such that for every player i, strategy τ^i of player i, duration $0 < \delta < \delta_0$, and time $s > s_0$, we have

$$\varepsilon + E^z_{\sigma_\delta} g^i_\delta(s) \geq v^i(z) \geq -\varepsilon + E^z_{\sigma^{-i}_\delta,\tau^i} g^i_\delta(s).$$

An exact family has an asymptotic uniform value in the zero-sum case (Theorem 6), and an asymptotic uniform equilibrium payoff in the non-zero-sum case (Theorem 12).

Remark 1 *The existence of an asymptotic uniform equilibrium payoff has the following corollaries.*

If v is the asymptotic uniform equilibrium payoff of a family $(\Gamma_\delta)_{\delta>0}$ then for every $\varepsilon > 0$ there is $\delta_0 > 0$ such that if $0 < \delta < \delta_0$ and $\overrightarrow{w}_\delta = (w^i)_{i\in N}$ is a profile of nonstationary discounting probability measures on $\mathbb{N}$ with $w^i_\delta(0) < \delta\delta_0$, then the game $\Gamma_{\delta,w}$ has an ε-equilibrium payoff within ε of v.

2.6 The robust nonstationary discounted solutions

Given a nonstationary discounting measure w on $[0,\infty]$, we define $\underline{g}^i_\delta(w)$ by

$$\underline{g}^i_\delta(w) := \liminf_{w_\delta \to w} g^i_\delta(w_\delta) \quad \text{and} \quad \bar{g}^i_\delta(w) := \limsup_{w_\delta \to w} g^i_\delta(w_\delta),$$

[3]Without the assumption of finitely many actions a uniform value need not exist [13]. The assumption of finitely many states is obviously needed.

where the lim inf and lim sup are over all nonstationary discounting measures w_δ on $\mathbb{N}$ that converge to w. If 1_∞ denotes the probability measure on $[0,\infty]$ with $1_\infty(\infty)=1$, then $\underline{g}^i_\delta(1_\infty)=\underline{g}^i_\delta$ and $\bar{g}^i_\delta(1_\infty)=\bar{g}^i_\delta$.

Fix a nonstationary discounting measure w on $[0,\infty]$ and a profile $\overrightarrow{w}=(w^i)_{i\in N}$ of nonstationary discounting measures w^i on $[0,\infty]$.

We say that $v\in\mathbb{R}^S$ is an *asymptotic w-limiting-average value* of the family $(\Gamma_\delta)_{\delta>0}$ if for every $\varepsilon>0$ there are strategies σ_δ of player 1 and τ_δ of player 2, and a duration $\delta_0>0$, such that for every strategy τ of player 2, strategy σ of player 1, and $0<\delta<\delta_0$, we have

$$\varepsilon+E^z_{\sigma_\delta,\tau}\underline{g}_\delta(w)\geq v(z)\geq-\varepsilon+E^z_{\sigma,\tau_\delta}\bar{g}_\delta(w).$$

We say that $v\in\mathbb{R}^S$ is an *asymptotic w-uniform value* of the family $(\Gamma_\delta)_{\delta>0}$ if for every $\varepsilon>0$ there are strategies σ_δ of player 1 and τ_δ of player 2, such that for all strategies τ^*_δ of player 2, strategies σ^*_δ of player 1, and nonstationary discounting measures w_δ on $\mathbb{N}$ that converge (as $\delta\to0+$) to w, we have

$$\varepsilon+\liminf_{\delta\to0+}E^z_{\sigma_\delta,\tau^*_\delta}g_\delta(w_\delta)\geq v(z)\geq-\varepsilon+\limsup_{\delta\to0+}E^z_{\sigma^*_\delta,\tau_\delta}g_\delta(w_\delta).$$

Similarly, we say that $v\in\mathbb{R}^{N\times S}$ is an *asymptotic $\overrightarrow{w}$-limiting-average equilibrium payoff* of the family $(\Gamma_\delta)_{\delta>0}$ if for every $\varepsilon>0$ there are strategy profiles σ_δ ($\delta>0$) and a duration $\delta_0>0$, such that for every player i, strategy τ^i_δ of player i, and $0<\delta<\delta_0$, we have

$$\varepsilon+E^z_{\sigma_\delta}\underline{g}^i_\delta(w^i)\geq v^i(z)\geq-\varepsilon+E^z_{\sigma^{-i}_\delta,\tau^i_\delta}\bar{g}^i_\delta(w^i).$$

We say that $v\in\mathbb{R}^{N\times S}$ is an *asymptotic $\overrightarrow{w}$-uniform equilibrium payoff* of the family $(\Gamma_\delta)_{\delta>0}$ if for every $\varepsilon>0$ there are strategy profiles σ_δ, such that for every player i, all strategies τ^i_δ of player i, and all nonstationary discounting measures w^i_δ on $\mathbb{N}$ that converge (as $\delta\to0+$) to w^i, we have

$$\varepsilon+\liminf_{\delta\to0+}E^z_{\sigma_\delta}g^i_\delta(w^i_\delta)\geq v^i(z)\geq-\varepsilon+\limsup_{\delta\to0+}E^z_{\sigma^{-i}_\delta,\tau^i_\delta}g^i_\delta(w^i_\delta).$$

Note that v is an asymptotic limiting-average, respectively asymptotic uniform, equilibrium payoff of a family $(\Gamma_\delta)_{\delta>0}$ if and only if it is an asymptotic 1_∞-limiting-average, respectively asymptotic 1_∞-uniform, equilibrium payoff of this family. Therefore the results in the paragraph below generalize our results about the existence of an asymptotic limiting-average, respectively asymptotic uniform, equilibrium payoff.

A strongly convergent family $(\Gamma_\delta)_{\delta>0}$ has an asymptotic $\overrightarrow{w}$-limiting-average equilibrium payoff, and an exact family $(\Gamma_\delta)_{\delta>0}$ has an asymptotic $\overrightarrow{w}$-uniform equilibrium payoff.

In what follows we define the asymptotic w-robust value and the asymptotic $\overrightarrow{w}$-robust equilibrium payoff.

We say that $v \in \mathbb{R}^S$ is an *asymptotic w-robust value* of the family $(\Gamma_\delta)_{\delta>0}$ (of two-person zero-sum games) if for every $\varepsilon > 0$ there are strategies σ_δ of player 1 and τ_δ of player 2, such that for all strategies τ^*_δ of player 2, strategies σ^*_δ of player 1, and nonstationary discounting measures w_δ on $\mathbb{N} \cup \{\infty\}$ that converge (as $\delta \to 0+$) to w, we have

$$\varepsilon + \liminf_{\delta\to 0+} E^z_{\sigma^1_\delta,\tau^2_\delta} \underline{g}^i_\delta(w_\delta) \geq v^i(z) \geq -\varepsilon + \limsup_{\delta\to 0+} E^z_{\tau^1_\delta,\sigma^2_\delta,} \bar{g}^i_\delta(w_\delta).$$

We say that $v \in \mathbb{R}^{N\times S}$ is an *asymptotic $\overrightarrow{w}$-robust equilibrium payoff* of the family $(\Gamma_\delta)_{\delta>0}$ if for every $\varepsilon > 0$ there are strategy profiles σ_δ, such that for every player i, all strategies τ^i_δ of player i, and all nonstationary discounting measures w_δ on $\mathbb{N} \cup \{\infty\}$ that converge (as $\delta \to 0+$) to w, we have

$$\varepsilon + \liminf_{\delta\to 0+} E^z_{\sigma_\delta} \underline{g}^i_\delta(w^i_\delta) \geq v^i(z) \geq -\varepsilon + \limsup_{\delta\to 0+} E^z_{\sigma^{-i}_\delta,\tau^i_\delta} \bar{g}^i_\delta(w^i_\delta).$$

An asymptotic $\overrightarrow{w}$-robust equilibrium payoff of a family $(\Gamma_\delta)_{\delta>0}$ is (by definition) an asymptotic w-limiting-average equilibrium payoff and an asymptotic $\overrightarrow{w}$-uniform equilibrium payoff.

Theorem 13 asserts that for every nonstationary discounting N-vector measure $\overrightarrow{w}$ on $[0,\infty]$, an exact family $(\Gamma_\delta)_{\delta>0}$ of N-person games has an asymptotic $\overrightarrow{w}$-robust equilibrium payoff.

2.7 The variable short-stage duration games

The paper states and proves asymptotic results on families $(\Gamma_\delta)_{\delta>0}$ of discrete-time stochastic games. In each game Γ_δ the stage duration is a constant positive number $\delta > 0$. The results remain intact also in the case where the parameter δ is a sequence of stage durations $\delta = (\delta_m)_{m\geq 0}$ with $d_n := \sum_{0\leq m<n} \delta_m \to_{n\to\infty} \infty$, where δ_m is the duration of the m-th stage, the m-th

stage payoff function is $g_{\delta,m}$ (or g_m for short), and the m-th stage transition function is $p_{\delta,m}$ (or p_m for short).[4]

The condition that the constant stage duration is sufficiently small needs to be replaced with the condition that the supremum of the stage durations, $d(\delta) := \sup_{m\geq 0} \delta_m$, is sufficiently small. A family $(\Gamma_\delta)_\delta$ with variable stage duration *converges in data* if $\sup_{m\geq 0} \|g_m/\delta_m - g\|$ and $\sup_{m\geq 0} \|p_m/\delta_m - \mu\|$ converge to zero as $d(\delta)$ goes to zero. It is an exact sequence if $g_m = \delta_m g$ and $p_m = \delta_m \mu$, and it converges strongly if it converges in data and for every δ, $m \geq 0$, $z' \neq z$, and $a \in A(z)$, $p_m(z', z, a) \neq 0$ iff $\mu(z', z, a) \neq 0$.

The ρ-discounted present value of the payoff g_m at stage m is $g_m \prod_{0\leq j<m}(1-\delta_j\rho)$ (where a product over an empty set of indices is zero). Therefore, in the ρ-discounted game Γ_δ, the valuation of a play $(z_0, a_0, \ldots, z_m, a_m \ldots)$ by player i is $\sum_{m=0}^{\infty} g_m(z_m, a_m) \prod_{0\leq j<m}(1 - \delta_j\rho)$.

In the case of a time-separable valuation, w_δ is said to be nonstationary discounting if $\frac{w_\delta(m)}{\delta_m}$ is nonincreasing in m. We assign to the measure w_δ on $\mathbb{N}$ the measure w'_δ on $[0, \infty)$ that is supported on $\{d_n : n \in \mathbb{N}\}$ and $w'_\delta(d_n) = w_\delta(n)$. We say that w_δ converges, as $d(\delta) \to 0+$, to the measure w on $[0, \infty)$ if w'_δ converges w^* to w.

Similarly, in the limiting-average games with variable stage duration δ, we set $g(s) = \frac{1}{s}\sum_{0\leq m:d_m<s} g_m(z_m, a_m)$ and in the definitions of $\underline{g}^i_\delta$ and $\bar{g}^i_\delta$, the condition $w_\delta(m) < \eta$ needs to be replaced with $w_\delta(m) < \eta\delta_m$.

3 Convergence of stochastic games with short-stage duration

We study the "convergence" of the family $(\Gamma_\delta)_{\delta>0}$, and the presentation of the "limit" as a continuous-time stochastic game Γ.

We define various conditions of the dependence of the transition rates p_δ on the stage duration δ. Some of these conditions relate directly to assumptions on the homogeneous Markov chain of states that are defined by an initial state, a stationary strategy, and the stage duration δ. Each one of the conditions can be interpreted as a consistency, or approximate consistency, of the models Γ_δ as δ varies.

Condition (p.0) asserts that the probability of a state change within the

[4]Moreover, the stage-dependent duration δ_m, payoff g_m, and transition function p_m can depend on past history.

first m stages (namely, in a time $t \leq m\delta$) converges to zero as $m\delta$ goes to zero. In particular, the probability of a state change in a single stage converges to zero as δ goes to zero. Condition (p.0) is equivalent to $mp_\delta(z, z, a)$ converging to zero as $m\delta$ goes to zero. Recall that condition (p.2) is $\lim_{\delta \to 0+} p_\delta/\delta = \mu$ where $\mu : S \times \mathcal{A} \to \mathbb{R}$, and note that condition (p.2) implies condition (p.0).

Recall that condition (p.3) requires (p.2) and that $p_\delta(z', z, a) > 0$ if and only if $\mu(z', z, a) > 0$ (where $\mu(z', z, a)$ is the limit, as δ goes to zero, of $p_\delta(z', z, a)/\delta$). Condition (p.3) implies that the ergodic classes of the homogeneous Markov chain that is defined by a stationary strategy and the transition rates p_δ are independent of δ.

Recall that condition (p.1) is $p_\delta = \delta\mu$, condition (p.1) implies condition (p.3), and condition (p.3) implies condition (p.2). Therefore, each asymptotic property that holds in any family $(\Gamma_\delta)_{\delta>0}$ that obeys (g.2) and (p.k) holds also in any family $(\Gamma_\delta)_{\delta>0}$ that obeys (g.1) and (p.k'), where $k' = 3$ if $k = 2$ and $k' = 1$ if $k = 3$.

Recall the following definitions of convergence in data and strong convergence.

Definition 1 (Convergence in data) *We say that* Γ_δ converges in data *(as* $\delta \to 0$*) if the family* $(\Gamma_\delta)_{\delta>0}$ *satisfies conditions (g.2) and (p.2).*

Definition 2 (Strong convergence) *We say that* Γ_δ converges strongly *(as* $\delta \to 0$*) if the family* $(\Gamma_\delta)_{\delta>0}$ *satisfies conditions (g.2) and (p.3).*

Next, we wish to define the "convergence" of the family $(\Gamma_\delta)_{\delta>0}$ as a convergence (as $\delta \to 0+$) of the stochastic process of states and payoffs that is defined by the initial state and a strategy σ. Obviously, in defining the convergence of the stochastic process of states and payoffs one has to take into account the stage duration δ. The state z_n in the play of the discrete-time stochastic game Γ_δ is interpreted as the state at time $n\delta$. Similarly, the sum $\sum_{j=0}^{n-1} g_\delta(z_j, a_j)$ of stage payoffs in stages $0 \leq j < n$ is interpreted as the cumulative payoff in the time interval $[0, n\delta]$.

Definition 3 (Convergence in stationary dynamics) *We say that* Γ_δ converges in stationary dynamics *if for all pure stationary strategies* σ*, states* $z', z \in S$*, times* $t \geq 0$*, and positive integers* n_δ *such that* $n_\delta \delta \underset{\delta \to 0+}{\longrightarrow} t$*, we have*

$$P^z_{\delta,\sigma}(z_{n_\delta} = z') \underset{\delta \to 0+}{\longrightarrow} F^\sigma_{z,z'}(t)$$

and

$$E^z_{\delta,\sigma}\sum_{j=0}^{n_\delta} g_\delta(z_j,a_j) \underset{\delta\to 0+}{\longrightarrow} G_t(z,\sigma),$$

where $(\sigma,z',z,t)\mapsto F^\sigma_{z,z'}(t)\in\mathbb{R}$ *and* $(t,z,\sigma)\mapsto G_t(z,\sigma)\in\mathbb{R}^N$ *are functions that are defined for all pure stationary strategies* σ*, states* $z',z\in S$*, and times* $t\geq 0$*.*

3.1 Stationary convergence

Proposition 1 *The following conditions are equivalent:*
(A) $(\Gamma_\delta)_{\delta>0}$ *converges in stationary dynamics.*
(B) $(\Gamma_\delta)_{\delta>0}$ *converges in data.*

Proof. (A) $\Longrightarrow$ (B). Assume condition (A) holds. Obviously, $\sum_{z'\in S} P^z_{\delta,\sigma}(z_{n_\delta}=z')=1$. Therefore, $\sum_{z'\in S}F^\sigma_{z,z'}(t)=1$. Applying condition (A) to $n_\delta=0$ and $z'=z$, we have $F^\sigma_{z,z}(0)=1$. Applying condition (A) to $t=0$ and all nonnegative integers n_δ with $\delta n_\delta \underset{\delta\to 0+}{\longrightarrow} 0$, we deduce that for every $\varepsilon>0$ there are $t_\varepsilon>0$ and $\delta_\varepsilon>0$, such that for every $0<\delta<\delta_\varepsilon$ and n with $n\delta\leq t_\varepsilon$, we have $P^z_{\delta,\sigma}(z_n=z)>1-\varepsilon$ for all states $z\in S$ and pure stationary strategy profiles σ.

Fix $z\in S$ and $a\in A(z)$, set $K_\delta=K_\delta(z)=\sum_{z'\neq z}p_\delta(z',z,a)$, and let σ be a pure stationary strategy with $\sigma(z)=a$, and $n=n_\delta=[t_{1/3}/\delta]$ (where $[*]$ denotes the largest integer that is less than or equal to $*$). Then, for $\delta<\delta_{1/3}$, $1/3>P^z_{\delta,\sigma}(z_n\neq z)\geq\sum_{m=1}^n P^z_{\delta,\sigma}(\forall j<m\ z_j=z \text{ and } z\neq z_m=z_n)\geq\sum_{m=1}^n(1-K_\delta)^{m-1}K_\delta 2/3=(1-(1-K_\delta)^n)2/3$, which implies the inequality $(1-K_\delta)^n\geq 1/2$. Therefore, $\limsup_{\delta\to 0+}K_\delta/\delta<\infty$. Therefore, there is a positive constant K such that for all $\delta>0$, $z\in S$, and $a\in A(z)$, we have $\sum_{z'\neq z}p_\delta(z',z,a)<K\delta$.

Next, we prove that if for a pair of distinct states $z'\neq z$ and an action profile $a\in A(z)$ we have $\liminf_{\delta\to 0+}p_\delta(z',z,a)/\delta<c$, then, for $t>0$ sufficiently small and a stationary strategy σ with $\sigma(z)=a$, we have $F^\sigma_{z,z'}(t)<ct$. Indeed, the set $\{z_n=z',\ z_0=z\}$ is the union of the disjoint sets $Y_{m,z''}=\{\forall 0\leq j<m,\ \ z_j=z_0, z_m=z'' \text{ and } z_n=z'\}$, where m ranges over the positive integers $1\leq m\leq n$ and z'' ranges over all states $z''\neq z$. Let $\varepsilon>0$ and set $n=n_\delta=[t_\varepsilon/\delta]$. Note that $P^z_{\delta,\sigma}(Y_{m,z''})\leq p_\delta(z',z,a)$ for $z''=z'$ and $\sum_{m=1}^{n-1}\sum_{z\neq z''\neq z'}P^z_{\delta,\sigma}(Y_{m,z''})\leq\varepsilon K\delta n$ for δ sufficiently small. Therefore, if $\delta>0$ is sufficiently small so that, in addition, $p_\delta(z',z,a)/\delta<c$ and for

all $z'' \neq z$ and $a \in A(z)$ we have $p_\delta(z'', z, a) \leq K\delta$, then $P^z_{\delta,\sigma}(z_n = z') \leq \sum_{m=1}^n P^z_{\delta,\sigma}(Y_{m,z'}) + \varepsilon K\delta n \leq (c + K\varepsilon)\delta n$. Therefore for $t > 0$ sufficiently small we have $F^\sigma_{z,z'}(t) < ct$.

Finally, we prove that if for a pair of distinct states $z' \neq z$ and an action profile $a \in A(z)$ we have $\limsup_{\delta \to 0+} p_\delta(z', z, a)/\delta > c$, then, for $t > 0$ sufficiently small and a stationary strategy σ with $\sigma(z) = a$, we have $F^\sigma_{z,z'}(t) > ct$. Indeed, the set $\{z_n = z',\ z_0 = z\}$ contains the disjoint sets $Y_{m,z'} = \{\forall 0 \leq j < m, \quad z_j = z_0, z_m = z' = z_n\}$, where m ranges over the positive integers $1 < m \leq n$. Let $\varepsilon > 0$ and set $n = n_\delta = [t_\varepsilon/\delta]$. Note that $P^z_{\delta,\sigma}(Y_{m,z'}) \geq (1-\varepsilon)^2 p_\delta(z', z, a)$ for δ sufficiently small. Therefore, if $\delta > 0$ is sufficiently small so that, in addition, $p_\delta(z', z, a)/\delta > c$, then $P^z_{\delta,\sigma}(z_n = z') \geq \sum_{m=1}^n P^z_{\delta,\sigma}(Y_{m,z'}) \geq n(1-\varepsilon)^2\delta c$. Therefore for $t > 0$ sufficiently small we have $F^\sigma_{z,z'}(t) > ct$.

We conclude that the $\limsup_{\delta \to 0+} p_\delta(z', z, a)/\delta$ and the $\liminf_{\delta \to 0+} p_\delta(z', z, a)/\delta$ coincide.

We will now prove that the second part of (B) holds. Fix a player $i \in N$ and assume that $\limsup_{\delta \to 0+} \|g^i_\delta\|/\delta < \infty$, where $\|g^i_\delta\| := \max_{z,a} |g^i_\delta(z, a)|$. For $t > 0$ let $\gamma_t(z, \sigma) = \frac{1}{t} G_t(z, \sigma)$. Then, for $\delta > 0$ sufficiently small, $g^i_\delta(z, \sigma(z))/\delta - 2\varepsilon\|g^i_\delta\|/\delta \leq \gamma^i_{t_\varepsilon}(z, \sigma) + \varepsilon$. Therefore

$$\limsup_{\delta \to 0+} g^i_\delta(z, \sigma(z))/\delta \leq \gamma^i_{t_\varepsilon}(z, \sigma) + \varepsilon + 2\varepsilon \limsup_{\delta \to 0+} \|g^i_\delta\|/\delta,$$

and therefore

$$\limsup_{\delta \to 0+} g^i_\delta(z, \sigma(z))/\delta \leq \liminf_{\varepsilon \to 0+} \gamma^i_{t_\varepsilon}(z, \sigma).$$

Similarly, for $\delta > 0$ sufficiently small, $\gamma^i_{t_\varepsilon}(z, \sigma) - \varepsilon \leq g^i_\delta(z, \sigma(z))/\delta + 2\varepsilon\|g^i_\delta\|/\delta$, and therefore $\limsup_{\varepsilon \to 0+} \gamma^z_{t_\varepsilon}(z, \sigma) \leq \liminf_{\delta \to 0+} g^i_\delta(z, \sigma(z))/\delta$. Given $a \in A(z)$ and applying these inequalities to a stationary strategy σ with $\sigma(z) = a$ we conclude that the $\liminf_{\delta \to 0+} g^i_\delta(z, a)/\delta$ and the $\limsup_{\varepsilon \to 0+} g^i_\delta(z, a)/\delta$ coincide.

It remains to prove that condition (A) implies that $\limsup_{\delta \to 0+} \|g^i_\delta\|/\delta < \infty$. For every $1 > \delta > 0$ let $z_\delta \in S$ and $a_\delta \in A(z)$ be such that $|g^i_\delta(z_\delta, a_\delta)| = \|g^i_\delta\|$. Let $\varepsilon > 0$, and let $\sigma = \sigma_\delta$ be a stationary strategy with $\sigma(z_\delta) = a_\delta$. Set $n = n_\delta = [t_\varepsilon/\delta]$ and $z_0 = z_\delta$. If $g^i_\delta(z_\delta, a_\delta) \geq 0$, then, for sufficiently small $\delta > 0$, we have $G^i_{t_\varepsilon}(z_\delta, \sigma) + t_\varepsilon/3 \geq E^{z_\delta}_\sigma \sum_{j=0}^{n-1} g^i_\delta(z_j, a_j) \geq (1 - 2\varepsilon) n g^i_\delta(z_\delta, a_\delta)$. Therefore, if $\varepsilon < 1/3$ we have $g^i_\delta(z_\delta, a_\delta)/\delta \leq 3|\gamma^i_{t_\varepsilon}(z, \sigma)| + 1$ for $\delta > 0$ sufficiently small. If $g^i(z_\delta, a_\delta) < 0$, then, for sufficiently small $\delta > 0$, we have $G^i_{t_\varepsilon}(z_\delta,\sigma) - t_\varepsilon/3 \leq E^{z_\delta}_\sigma \sum_{j=0}^{n-1} g^i_\delta(z_j,a_j) \leq (1-2\varepsilon)ng^i_\delta(z_\delta,a_\delta)$. Therefore, if $\varepsilon < 1/3$ we have $g^i_\delta(z_\delta,a_\delta)/\delta \geq -3|\gamma^i_{t_\varepsilon}(z,\sigma)| - 1$. This proves that $\limsup_{\delta\to 0+} \|g^i_\delta\|/\delta \leq 3|\gamma^i_{t_\varepsilon}(z,\sigma)| + 1 < \infty$.

(B) $\implies$ (A). Let σ be a stationary strategy and let Q be the $S\times S$ matrix whose (z,z')-th entry is $Q_{z,z'} = \mu(z',z,\sigma(z))$. Note that for $\delta > 0$ sufficiently small, $I + \delta Q$ is a transition matrix, where I stands for the identity matrix, and $\|I + \delta Q\| := \max_{z\in S} \sum_{z'\in S} |(I+\delta Q)_{z,z'}| = 1$. In addition, $e^{\delta Q}$ (which equals by definition the convergent sum $\sum_{j=0}^{\infty} \frac{\delta^j Q^j}{j!}$) is an $S \times S$ matrix, and $(e^{\delta Q})^n = e^{n\delta Q}$. Let P_δ be the $S \times S$ transition matrix whose (z,z')-th entry is $(P_\delta)_{z,z'} = I_{z,z'} + p_\delta(z',z,\sigma(z))$. Therefore, if n is a positive integer, then $P^z_{\delta,\sigma}(z_n = z') = (P^n_\delta)_{z,z'}$. By the assumption on p_δ and the definitions of Q and $e^{\delta Q}$, we have $\|e^{\delta Q} - P_\delta\| \leq o(\delta)$ as $\delta \to 0+$.

For any two $S \times S$ matrices (or elements of a norm algebra) A and B we have $A^n - B^n = \sum_{k=1}^{n} A^{n-k}(A-B)B^{k-1}$, implying that $\|A^n - B^n\| \leq \|A - B\| \sum_{j=0}^{n-1} \|A\|^j \|B\|^{n-j}$. Therefore, $\|P^n_\delta - e^{n\delta Q}\| \leq \|P_\delta - e^{\delta Q}\| \sum_{j=0}^{n-1} \|e^{\delta Q}\|^j \leq o(\delta)n$ as $\delta \to 0+$.

Therefore, $\|P^n_\delta - e^{tQ}\| \leq \|P^n_\delta - e^{n\delta Q}\| + \|e^{tQ} - e^{n\delta Q}\| \to 0$ as $\delta \to 0+$ and $n\delta \to t$. We conclude that $P^z_{\delta,\sigma}(z_n = z') \to F^\sigma_{z,z'}(t) = (e^{tQ})_{z,z'} \in \mathbb{R}$ as $\delta \to 0+$.

By assumption (B) we have $g_\delta(z,a) = \delta g(z,a) + o(\delta)$. Therefore, if $\delta \to 0+$ and $n_\delta\delta \to t > 0$, then $|E^z_{\delta,\sigma} \sum_{j=0}^{n_\delta - 1} g^i_\delta(z_j,a_j) - E^z_{\delta,\sigma} \sum_{j=0}^{n_\delta-1} \delta g^i(z_j,a_j)| \to 0$. If $\delta \to 0+$ and $n_\delta\delta \to t > 0$, then, as shown earlier, $P^z_{\delta,\sigma}(z_n = z') \to F^\sigma_{z,z'}(t)$, and, therefore, $E^z_{\delta,\sigma} \sum_{j=0}^{n_\delta-1} \delta g^i(z_j,a_j) \to G_t(z,\sigma) = \int_0^t \sum_{z'\in S} F^\sigma_{z,z'}(s) g(z',\sigma(z'))\,ds$. Therefore, $E^z_{\delta,\sigma} \sum_{j=0}^{n_\delta-1} g^i_\delta(z_j,a_j) \to G_t(z,\sigma)$ as $\delta \to 0+$ and $n_\delta\delta \to t > 0$. $\square$

Remark 2 *The above proof of condition (B) implying condition (A) proves that for every stationary strategy σ, every time $t \geq 0$, all states $z, z' \in S$, and all integers $0 \leq n_\delta$ with $n_\delta\delta \to_{\delta\to 0+} t$, $P^z_\sigma(z_{n_\delta} = z') \to_{\delta\to 0+} F^\sigma_{z,z'}(t) = e^{tQ}_{z,z'}$ where Q is the $S \times S$ matrix whose (z,z')-th entry is $Q_{z,z'} = \mu(z',z,\sigma(z))$.*

Note that every continuous-time stochastic game $\Gamma = \langle N,S,A,\mu,g\rangle$ is a "data limit" of the family of discrete-time stochastic games $\Gamma_\delta = \langle N,S,A,p_\delta,g_\delta\rangle$, where $g_\delta(z,a) = \delta g(z,a)$ and $p_\delta(z',z,a) = \delta\mu(z',z,a)$ for all pairs of distinct states $z' \neq z$ and every action profile $a \in A(z)$.

3.2 Markov convergence

The next proposition gives a sufficient condition for a family of Markov strategies σ_δ in Γ_δ to have a continuous-time limiting dynamics and payoffs as $\delta \to 0+$. In the formulas that follow, we view $\sigma_\delta(z,j)$ ($j \in \mathbb{N}$) as a measure on $A(z)$; i.e., $\sigma_\delta(z,j) \in \Delta(A(z))$, and $\sigma_\delta(j) := (\sigma_\delta(z,j))_{z\in S}$ is an element of $\times_{z\in S}\Delta(A(z))$. Therefore, for any fixed $z \in S$, any linear combination of $\sigma_\delta(z,j)$ is a measure on $A(z)$. Similarly, if $\sigma : S \times \mathbb{R}_+ \to \Delta(A)$ is measurable with $\sigma(z,t) \in \Delta(A(z))$, then, for any function $f \in L_1(\mathbb{R}_+)$, the integral $\int_0^\infty f(t)\sigma(z,t)\,dt$ is well defined.

We say that the Markov strategies σ_δ in Γ_δ *converge* w* if for every continuous function $f : \mathbb{R}_+ \to \mathbb{R}$ with bounded support, the limit of $\sum_{j=0}^\infty f(j\delta)\delta\sigma_\delta(z,j)$ as $\delta \to 0+$ exists. In that case, there is a measurable function $\sigma : S \times \mathbb{R}_+ \to \Delta(A)$ (with $\sigma(z,t) \in \Delta(A(z))$) such that for every $f \in L_1(\mathbb{R}_+)$ the limit of $\int_0^\infty f(t)\sigma_\delta(z,[t/\delta])\,dt$ as $\delta \to 0+$ exists and equals $\int_0^\infty f(t)\sigma(z,t)\,dt$, and we say that the discrete-time Markov strategies σ_δ *converge* w* *to* (the continuous-time Markov correlated strategy) $\sigma : S \times \mathbb{R}_+ \in \Delta(A)$.

Whenever the conditional probability $P^{z_0}_{\delta,\sigma}(E_1 \mid E_2)$ is independent of the initial state z_0, we suppress the superscript of the initial state z_0.

Proposition 2 *If the (correlated) Markov strategies σ_δ in Γ_δ converge w* to $\sigma : S \times \mathbb{R}_+ \to \Delta(A)$ and the family of discrete-time stochastic games $(\Gamma_\delta)_{\delta>0}$ converges in data, then, for every $0 \le s < t$, there are $S \times S$ transition matrices $F^\sigma(s,t)$ such that*

$$P_{\sigma_\delta}(z_n = z' \mid z_k = z) \to F^\sigma_{z,z'}(s,t) \text{ as } \delta \to 0+, k\delta \to s, \text{ and } n\delta \to t,$$

and

$$E^z_{\sigma_\delta} \sum_{0\le m<n} g_\delta(z_m,a_m) \to \int_0^t \sum_{z'\in S} F^\sigma_{z,z'}(0,t) g(z',\sigma(z',t))\,dt \text{ as } \delta \to 0+ \text{ and } n\delta \to t.$$

Proof. As the family of discrete-time stochastic games $(\Gamma_\delta)_{\delta>0}$ converges in data, there is a positive constant $K > 0$ such that for every $(z,a) \in \mathcal{A}$ we have $|p_\delta(z,z,a)| > 1 - K\delta$. Therefore, if $0 \le k < n$, $|P_{\delta,\sigma_\delta}(z_n = z' \mid z_k = z) - I_{z,z'}| < 1 - (1-K\delta)^{n-k} \to 0$ as $n\delta - k\delta \to 0+$. Therefore, it suffices to prove that for every $s < t$ there are sequences $k_\delta < n_\delta$ such that $k_\delta\delta \to s$ and $n_\delta\delta \to t$ such that

$$P_{\delta,\sigma_\delta}(z_{n_\delta} = z' \mid z_{k_\delta} = z) \to F^\sigma_{z,z'}(s,t) \text{ as } \delta \to 0+\,.$$

We will prove it for $n_\delta = [t/\delta]$ and $k_\delta = [s/\delta]$.

Assume that the Markov strategies σ_δ in Γ_δ converge w* to $\sigma : S \times \mathbb{R}_+ \to \Delta(A)$. Let M be the space of all $S \times S$ matrices Q, let M_0 be the subset of all its matrices Q with $\sum_{z' \in S} Q_{z,z'} = 0$ for every $z \in S$ and $Q_{z,z'} \geq 0$ for all $z \neq z'$, and let M_1 be the subset of M of all transition matrices. The space M is a (noncommutative) Banach algebra with the norm $\|Q\| = \max_{z\in S} \sum_{z' \in S} |Q_{z,z'}|$, and M_1 is closed under multiplication. For an ordered list $F_1, \ldots, F_j \in M$ we denote by $\prod_{i=1}^j F_i$ the matrix (ordered) product $F_1 F_2 \ldots F_j$.

Let $Q : [0, \infty) \to M$ be defined by $Q_{z,z'}(u) = \mu(z', z, \sigma(z, u))$, and let $Q^\delta : [0, \infty) \to M$ be defined by $Q^\delta_{z,z'}(u) = p_\delta(z', z, \sigma_\delta(z, [u/\delta]))/\delta$. As $(\Gamma_\delta)_{\delta>0}$ converges in data, $Q^\delta_{z,z'}(u) = \mu(z', z, \sigma_\delta(z, [u/\delta])) + o(1)$ as $\delta \to 0+$. Therefore, $\int_s^t Q^\delta_{z,z'}(u)\, du = \mu(z', z, \int_s^t \sigma_\delta(z, [u/\delta]))\, du + o(1)$ as $\delta \to 0+$, where for a measure α on $A(z)$ we define $\mu(z', z, \alpha) := \sum_{a\in A(z)} \alpha(a)\mu(z', z, a)$. Therefore, as the Markov strategies σ_δ converge w* to σ, for every $s < t$ we have

$$\int_s^t Q^\delta(u)\, du \underset{\delta\to 0+}{\longrightarrow} \int_s^t Q(u)\, du.$$

Let G^δ_j be the transition matrix $(G^\delta_j)_{z,z'} = p_\delta(z', z, \sigma_\delta(z, j)) + I_{z,z'}$, and given $0 \leq s \leq t$ we define $G^\delta(s, t)$ to be the transition matrix $\prod_{j=[s/\delta]}^{[t/\delta]-1} G^\delta_j$, where a product over an empty set of indices is defined as the identity. It suffices to prove that $G^\delta(s, t)$ converges as $\delta \to 0+$.

Let $C = 2 \max_{z,a} |\mu(z, z, a)| < C'$. It follows that for every $t \geq 0$ we have $\|Q(t)\| \leq C$, and for sufficiently small $\delta > 0$ we have $\|Q^\delta(t)\| < C'$. Let $L_\delta(s, t) = [t/\delta] - [s/\delta]$, and note that $\delta L_\delta(s, t) \leq t - s + \delta$.

As M is a Banach algebra, for every finite sequence $Q_1, \ldots, Q_m$ of elements in M, we have

$$\| \prod_{j=1}^m (I + Q_j) - I - \sum_{j=1}^m Q_j \| \leq e^{\sum_{j=1}^m \|Q_j\|} - 1 - \sum_{j=1}^m \|Q_j\|. \qquad (1)$$

Inequality (1) follows from the inequality $e^x \geq 1 + x$, the triangle inequality, and the Banach algebra inequality $\|QQ'\| \leq \|Q\|\|Q'\|$. Indeed, if $\theta_j = \|Q_j\|$, then $\| \prod_{j=1}^m (I+Q_j) - I - \sum_{j=1}^m Q_j \| \leq \prod_{j\in J}(1+\theta_j) - 1 - \sum_{j=1}^m \theta_j \leq e^{\sum_{j=1}^m \theta_j} - 1 - \sum_{j=1}^m \theta_j$.

As $G^\delta_j = I + \int_{j\delta}^{j\delta+\delta} Q^\delta(u)\, du$, $\int_{j\delta}^{j\delta+\delta} \|Q^\delta(u)\|\, du \leq \delta C'$, and $e^x - 1 - x$ is monotonic increasing on $x \geq 0$, for all $0 \leq s < t$, we have

$$\begin{aligned}
\|G^\delta(s,t) - I - \int_s^t Q^\delta(u)\,du\| &\leq \|G^\delta(s,t) - I - \int_{\delta[s/\delta]}^{\delta[t/\delta]} Q^\delta(u)\,du\| + 2\delta C' \\
&\leq e^{C'\delta L_\delta(s,t)} - 1 - L_\delta(s,t)C'\delta + 2\delta C' \\
&\leq e^{(t-s+\delta)C'} - 1 - (t-s+\delta)C' + 2\delta C' \\
&< (t-s)^2 C'^2
\end{aligned}$$

for $(t-s)C' \leq 1$ and $\delta > 0$ sufficiently small.

For every sequence $s = t_0 < t_1 < \ldots < t_k = t$, set $A_j = G^\delta(t_{j-1}, t_j)$, $B_j^\delta = I + \int_{t_{j-1}}^{t_j} Q^\delta(u)\,du$, and $B_j = I + \int_{t_{j-1}}^{t_j} Q(u)\,du$, $j = 1, \ldots, k$. Note that $G^\delta(t_0, t) = \prod_{j=1}^k A_j$ and $\prod_{j=1}^k A_j - \prod_{j=1}^k B_j = \sum_{i=1}^k \prod_{j=1}^{i-1} A_j (A_i - B_i) \prod_{j=i+1}^k B_j$. For $1 \leq j < k$, $\|A_j\| = 1$, and for sufficiently small $\max_{i=1}^k (t_i - t_{i-1})$, $\|B_i\| = 1$ for every $1 \leq i \leq k$. Therefore $\|\prod_{j=1}^k A_j - \prod_{j=1}^k B_j\| \leq \sum_{j=1}^k \|A_j - B_j\| \leq \sum_{j=1}^k \|A_j - B_j^\delta\| + \sum_{j=1}^k \|B_j^\delta - B_j\|$. Therefore, for a sufficiently large k, by setting $t_j = s + j(t-s)/k$ and $F(t_{j-1}, t_j) = I + \int_{t_{j-1}}^{t_j} Q(u)\,du$, there is a (sufficiently small) $\delta(k) > 0$ such that for $0 < \delta < \delta(k)$, we have

$$\|G^\delta(s,t) - \prod_{j=1}^k F(t_{j-1}, t_j)\| \leq 2(t-s)^2 C'^2/k.$$

Therefore, $\sup_{0<\delta,\delta'<\delta(k)} \|G^\delta(s,t) - G^{\delta'}(s,t)\| \leq 4(t-s)^2 C'^2/k$, implying that $\lim_{k\to\infty} \sup_{0<\delta,\delta'<\delta(k)} \|G^\delta(s,t) - G^{\delta'}(s,t)\| = 0$. Therefore, $G^\delta(s,t)$ converges to a limit as $\delta \to 0+$. □

Remark 3 *The result applies in particular to profiles $\sigma_\delta = (\sigma_\delta^i)_{i\in N}$ of (uncorrelated) Markov strategies in Γ_δ that converge w^* to (a continuous-time correlated Markov strategy) $\sigma : S \times \mathbb{R}_+ \to \Delta(A)$. In this case the w^* limit σ need not represent a profile of continuous-time Markov strategies.*

For example, if σ_δ^1 and σ_δ^2 play (T, L) at even stages and (B, R) at odd stages, then the Markov strategy profiles $\sigma_\delta = (\sigma_\delta^1, \sigma_\delta^2)$ converge w^* to (the continuous-time stationary correlated strategy) σ with $\sigma(*)(T, L) = 1/2 = \sigma(*)(B, R)$. Therefore, asymptotic results that involve referral to Markov strategies need special attention. They are not obtained by simply "taking limits." However, if $\sigma : S \times \mathbb{R}_+ \to \Delta(A)$ is a continuous-time correlated Markov strategy, there are profiles of pure (and thus uncorrelated) Markov strategies $\sigma_\delta = (\sigma_\delta^i)_{i\in N}$ such that σ_δ converge w^* to σ.

Remark 4 *Proposition 2 holds also in the model of variable stage duration games. The conditions $\delta \to 0+$, $k\delta \to s$, and $n\delta \to t$, are replaced with $d(\delta) \to 0+$, $d_k \to s$, and $d_n \to t$ respectively, and the term $g_\delta(z_m, a_m)$ is replaced with $g_m(z_m, a_m)$.*

The proof of Remark 4 is obtained by the following (additional) notational modifications in the proof of Proposition 2. The inequality $p_\delta(z, z, a) > 1 - K\delta$ is replaced with $p_m(z, z, a) > 1 - K\delta_m$ for every $m \geq 0$, the term $(1 - K\delta)^{n-k}$ is replaced with $\prod_{k\leq m<n}(1 - K\delta_m)$, and a term of the form $[t/\delta]$ is replaced with the largest integer m such that $d_m \leq t$. The definition (in the proof of Proposition 2) of the $S \times S$ matrix $Q^\delta_{z,z'}(u)$ is modified to $Q^\delta_{z,z'}(u) = p_{[u/\delta]}(z', z, \sigma_\delta(z, [u/\delta]))/\delta_{[u/\delta]}$. The inequality $0 < \delta < \delta(k)$ is interpreted as $0 < d(\delta) < \delta(k)$.

4 Two-person zero-sum stochastic games with short-stage duration

4.1 The discounted case

Fix the sets of player $N = \{1, 2\}$, states S, and actions A, and let $\Gamma_\delta = \langle N, S, A, g_\delta, p_\delta\rangle$, or $\Gamma_\delta = \langle g_\delta, p_\delta\rangle$ for short, be a stochastic game whose stage payoff function g_δ and transitions p_δ depend on the parameter δ that represents the single-stage duration. Recall that $\Gamma_{\delta,\rho}$ denotes the (unnormalized) discounted game Γ_δ with discount factor $1 - \rho\delta$, $V_{\delta,\rho}$ denotes its value, and $V_\rho \in \mathbb{R}^S$ is the asymptotic ρ-discounted value of $(\Gamma_\delta)_{\delta>0}$ if $V_{\delta,\rho} \to_{\delta\to 0+} V_\rho$.

Given a family $(\Gamma_\delta)_{\delta>0}$ that has an asymptotic ρ-discounted value V_ρ, we say that the stationary strategy σ, respectively τ, is *asymptotic ρ-discounted optimal* if for every $\varepsilon > 0$, there is $\delta_0 > 0$, such that for every $0 < \delta < \delta_0$, strategy σ^* of player 1 (in Γ_δ), strategy τ^* of player 2 (in Γ_δ), and state z,

$$\varepsilon + E^z_{\delta,\sigma,\tau^*} \sum_{m=0}^{\infty}(1-\rho\delta)^m g_\delta(z_m, x_m) \geq V_\rho(z) \geq -\varepsilon + E^z_{\delta,\sigma^*,\tau} \sum_{m=0}^{\infty}(1-\rho\delta)^m g_\delta(z_m, x_m).$$

Given a converging family $(\Gamma_\delta)_{\delta>0}$, we denote by g and μ the limits, as $\delta \to 0+$, of g_δ/δ and p_δ/δ respectively.

We denote by $X^i(z)$, respectively $X(z)$, all probability distributions over $A^i(z)$, respectively over $A(z)$ $(= A^1(z) \times A^2(z))$. For $z \in S$ and $x^i \in X^i(z)$

we denote by $x^1 \otimes x^2$ the product distribution $x \in X(z)$ that is given by $x(a) = x^1(a^1)x^2(a^2)$ for $a = (a^1, a^2) \in A^1(z) \times A^2(z)$. For any function $h : a \mapsto h(a)$, that is defined over $A(z)$, e.g., $A(z) \ni a \mapsto g(z, a)$ or $A(z) \ni a \mapsto \mu(z', z, a)$, we denote also by h its linear extension to $X(z)$, i.e., $h(x) = \sum_{a \in A(z)} x(a)h(a)$.

Theorem 1 *Every converging family $(\Gamma_\delta)_{\delta>0}$ has an asymptotic ρ-discounted value, which equals the unique solution $V \in \mathbb{R}^S$ of the system of S equations, $z \in S$,*

$$\rho v(z) = \max_{x^1 \in X^1(z)} \min_{x^2 \in X^2(z)} \left(g(z, x^1 \otimes x^2) + \sum_{z' \in S} \mu(z', z, x^1 \otimes x^2) v(z') \right), \tag{2}$$

and each player has an asymptotic ρ-discounted optimal stationary strategy.

Proof. By the theory of discrete-time stochastic games, $V_{\delta,\rho}$ exists and is the unique solution of the system of equations

$$v(z) = \max_{x^1 \in X^1(z)} \min_{x^2 \in X^2(z)} \left(g_\delta(z, x^1 \otimes x^2) + \sum_{z' \in S} (1 - \rho\delta) P_\delta(z' \mid z, x^1 \otimes x^2) v(z') \right). \tag{3}$$

Since $P_\delta(z' \mid z, a) = p_\delta(z', z, a)$ for $z' \neq z$, and $P_\delta(z' \mid z, a) = 1 + p_\delta(z', z, a)$ for $z' = z$, we can deduce, by subtracting $(1 - \rho\delta)v(z)$ from both sides of the z equation, that $V_{\delta,\rho}$ exists and is the unique solution of the system of equations

$$\rho\delta v(z) = \max_{x^1 \in X^1(z)} \min_{x^2 \in X^2(z)} \left(g_\delta(z, x^1 \otimes x^2) + \sum_{z' \in S} (1 - \rho\delta) p_\delta(z', z, x^1 \otimes x^2) v(z') \right). \tag{4}$$

For $g_\delta = \delta g$ and $p_\delta = \frac{\delta}{1-\rho\delta}\mu$, v solves (4) if and only if it solves (2). For $\delta > 0$ sufficiently small, $p_\delta = \frac{\delta}{1-\rho\delta}\mu$ indeed represents transition probabilities. Therefore the system (2) of equations has a unique solution.

Let V be the unique solution of (2). Let σ be a stationary strategy of player 1 with $\sigma(z)$ maximizing (over all $x^1 \in X^1(z)$)

$$\min_{x^2 \in X^2(z)} g(z, x^1 \otimes x^2) + \sum_{z' \in S} \mu(z', z, x^1 \otimes x^2) V(z'). \tag{5}$$

Therefore, for every $z \in S$ and $x^2 \in X^2(z)$ we have

$$g(z, \sigma(z) \otimes x^2) + \sum_{z' \in S} \mu(z', z, \sigma(z) \otimes x^2) V(z') \geq \rho V(z). \tag{6}$$

Fix $\varepsilon > 0$. We claim that there is $\delta_0 > 0$, such that for every $0 < \delta < \delta_0$, strategy τ of player 2, and state z,

$$E^z_{\delta,\sigma,\tau} \sum_{m=0}^{\infty} (1 - \rho\delta)^m g_\delta(z_m, a_m) \geq V(z) - \varepsilon. \tag{7}$$

Fix an initial history $h_m = (z_0, a_0, \ldots, z_m)$, and let $x^2_m = \tau(h_m)$ and $x_m = \sigma(z_m) \otimes x^2_m$. Let $Y_m := E_{\sigma,\tau} \left(g_\delta(z_m, a_m) + (1 - \rho\delta) V(z_{m+1}) \mid h_m \right)$.

$$\begin{aligned}
Y_m &= g_\delta(z_m, x_m) + (1 - \rho\delta) \sum_{z' \in S} P_\delta(z' \mid z_m, x_m) V(z') \\
&\geq \delta g(z_m, x_m) + \sum_{z' \in S} \delta \mu(z', z_m, x_m) V(z') - \rho\delta V(z_m) + V(z_m) - o(\delta) \\
&\geq V(z_m) - o(\delta).
\end{aligned}$$

Therefore, for every $m \geq 0$, $E^z_{\delta,\sigma,\tau} (1-\rho\delta)^m g_\delta(z_m, a_m) \geq (1-\rho\delta)^m E^z_{\delta,\sigma,\tau} V(z_m) - (1 - \rho\delta)^{m+1} E^z_{\delta,\sigma,\tau} V(z_{m+1}) - o(\delta)(1 - \rho\delta)^m$. Summing over $m = 0, 1, \ldots$, we deduce that

$$E^z_{\delta,\sigma,\tau} \sum_{m=0}^{\infty} (1 - \rho\delta)^m g_\delta(z_m, a_m) \geq V(z) - o(\delta) \sum_{m=0}^{\infty} (1 - \rho\delta)^m \to_{\delta \to 0+} V(z).$$

By duality, if τ is a stationary strategy of player 2 with $\tau(z)$ minimizing (over all $x^2 \in X^2(z)$)

$$\max_{x^1 \in X^1(z)} g(z, x^1 \otimes x^2) + \sum_{z' \in S} \mu(z', z, x^1 \otimes x^2) V(z'), \tag{8}$$

then for every strategy σ of player 1 we have

$$E^z_{\delta,\sigma,\tau} \sum_{m=0}^{\infty} (1 - \rho\delta)^m g_\delta(z_m, a_m) \leq V(z) + o(\delta) \sum_{m=0}^{\infty} (1 - \rho\delta)^m \to_{\delta \to 0+} V(z).$$

□

Denote by $V_\rho(g, \mu)$ the asymptotic ρ-discounted value of the family $(\Gamma_\delta = \langle g_\delta, \mu_\delta \rangle)_{\delta > 0}$ that converges (as δ goes to zero) to $\langle g, \mu \rangle$, and by $V_{\delta,\rho}(g, p)$ the value of the discounted discrete-time stochastic game $\langle g, p \rangle$ with a discount factor $1 - \rho\delta$.

Remark 5 *The above proof of Theorem 1 shows that*

$$V_\rho(g,\mu) = V_{\delta,\rho}(\delta g, \frac{\delta}{1-\rho\delta}\mu) \quad \text{whenever } \delta \leq \frac{1}{\|\mu\|+\rho}, \tag{9}$$

where $\|\mu\| = \max_{z,a} |\mu(z,z,a)|$.

Remark 6 *The proof shows in addition that a stationary strategy σ of player 1, respectively τ of player 2, is asymptotic ρ-discounted optimal if and only if, for every state $z \in S$, $\sigma(z)$ maximizes (5), respectively, $\tau(z)$ minimizes (8).*

Remark 7 *It is worth recalling that a stationary strategy is a (behavioral) strategy whose mixed action at every stage is independent of the stage, past states, and past actions of the players. Therefore, the result holds also in a model where some of the players do not observe past actions, and even in a model where some of the players are unable to recall the current stage and past states.*

Remark 8 *The proof that (2) has a solution was based on the corresponding result from the theory of discounted discrete-time stochastic games. In what follows we prove it directly.*

For a vector $v \in \mathbb{R}^S$ we denote by $\|v\|$ its maximum norm $\|v\| := \max_{z\in S} |v(z)|$.

For every $z \in S$, $a \in A(z)$, $v \in \mathbb{R}^S$, and $x \in X(z)$, $G^z[v](a)$ is defined by

$$G^z[v](a) = \frac{1}{\|\mu\|+\rho}\left(g(z,a) + \sum_{z'\in S}\mu(z',z,a)v(z') + \|\mu\|v(z)\right),$$

and (thus) $G^z[v](x)$ is defined by

$$\begin{aligned} G^z[v](x) &= \sum_{a\in A(z)} x(a)G^z[v](a) \\ &= \frac{1}{\|\mu\|+\rho}\left(g(z,x) + \sum_{z'\in S}\mu(z',z,x)v(z') + \|\mu\|v(z)\right). \end{aligned}$$

Define the operator Q from $\mathbb{R}^S$ to $\mathbb{R}^S$ by

$$Qv(z) = \max_{x\in X^1(z)} \min_{x^2\in X^2(z)} G^z[v](x^1 \otimes x^2).$$

By the minmax theorem we have

$$Qv(z) = \min_{x^2 \in X^2(z)} \max_{x \in X^1(z)} G^z[v](x^1 \otimes x^2)$$

and therefore v is a solution of $Qv = v$ if and only if it is a solution of (2). Therefore, it suffices to prove that Q has a fixed point. Note that $G^z[v + c1_S](x) = G^z[v](x) + \frac{c\,\|\mu\|}{\|\mu\|+\rho}$, and therefore

$$Q(v + c1_S)(z) = Qv + \frac{c\,\|\mu\|}{\|\mu\| + \rho}.$$

In addition, Q is monotonic; i.e., $u \geq v$ implies that $Qu \geq Qv$, and therefore for $v, u \in \mathbb{R}^S$ we have

$$\|Qv - Qu\| \leq \frac{\|\mu\|}{\|\mu\| + \rho}\|v - u\|.$$

Therefore Q is a strict contraction and therefore Q has a unique fixed point. □

Remark 9 *The following (alternative) proof of Theorem 1 is based on results from the theory of continuous-time stochastic games in conjunction with stationary convergence of the family of games Γ_δ.*

We apply notations and inequalities from [8]. First, one recalls that a pair of stationary strategies, σ of player 1 and τ of player 2, where $\sigma(z)$ maximizes (5), and $\tau(z)$ minimizes (8), is a pair of optimal strategies in the continuous-time ρ-discounted game $\Gamma = \langle g, \mu \rangle$, and V is its value. In particular, for every stationary strategy τ^* of player 2 and every stationary strategy σ^* of player 1 we have

$$E^z_{\sigma,\tau^*} \int_0^\infty e^{-\rho t} g(z_t, \sigma(z_t) \otimes \tau^*(z_t))\, dt \geq V(z) \geq E^z_{\sigma^*,\tau} \int_0^\infty e^{-\rho t} g(z_t, \sigma^*(z_t) \otimes \tau(z_t))\, dt.$$

Next, stationary convergence implies that for stationary strategies σ' of player 1 and τ' of player 2 we have

$$E^z_{\delta,\sigma',\tau'} \sum_{m=0}^\infty (1 - \rho\delta)^m g_\delta(z_m, a_m) \to_{\delta \to 0+} E^z_{\sigma',\tau'} \int_0^\infty e^{-\rho t} g(z_t, \sigma'(z_t) \otimes \tau'(z_t))\, dt.$$

Therefore, given $\varepsilon > 0$, for $\delta > 0$ sufficiently small, for every pure stationary strategy τ^* of player 2 and pure stationary strategy σ^* of player 1, we have

$$\varepsilon + E^z_{\delta,\sigma,\tau^*} \sum_{m=0}^{\infty} (1-\rho\delta)^m g_\delta(z_m, a_m) \geq V(z) \geq -\varepsilon + E^z_{\delta,\sigma^*,\tau} \sum_{m=0}^{\infty} (1-\rho\delta)^m g_\delta(z_m, a_m).$$

In a discrete-time discounted game (with finitely many states and actions) there is always a pure stationary strategy that is a best reply to a given stationary strategy. Therefore V is an asymptotic ρ-discounted value and σ and τ are asymptotic ρ-discounted optimal strategies of the converging family $(\Gamma_\delta)_{\delta>0}$. □

The algebraic approach. Fix the finite state space S and the finite action sets $A^i(z)$ ($i = 1, 2$ and $z \in S$), and recall that $\mathcal{A} = \{(z, a) : z \in S,\ a \in A(z)\}$. The set of all $(g, \mu, v, \rho, x^1, x^2)$, where $g \in \mathbb{R}^{\mathcal{A}}$, $\mu \in \mathbb{R}^{S\times\mathcal{A}}$ (with $\mu(z', z, a) \geq 0$ for $S \ni z' \neq z \in S$ and $a \in A(z)$, and $\sum_{z'\in S} \mu(z', z, a) = 0$ for $(z, a) \in \mathcal{A}$), $v \in \mathbb{R}^S$, $0 < \rho < 1$, $x^i \in X^i(z)$, that satisfies the following finite[5] lists of inequalities,

$$\rho v(z) \leq \min_{y^2\in X^2(z)} \left(g(z, x^1 \otimes y^2) + \sum_{z'\in S} \mu(z', z, x^1 \otimes y^2) v(z') \right), \tag{10}$$

$$\rho v(z) \geq \max_{y^1\in X^1(z)} \left(g(z, y^1 \otimes x^2) + \sum_{z'\in S} \mu(z', z, y^1 \otimes x^2) v(z') \right), \tag{11}$$

is semialgebraic. Therefore, for each fixed (g, μ), the graph of the correspondence assigning to each ρ the asymptotic ρ-discounted optimal stationary strategies of each player and the asymptotic ρ-discounted value function V_ρ is semialgebraic. Therefore (see, e.g., [1, 6]), there is a semialgebraic map $\rho \mapsto (V_\rho, \sigma^\rho, \tau^\rho)$, where V_ρ is the ρ-discounted asymptotic value and σ^ρ and τ^ρ are stationary asymptotic ρ-discounted optimal strategies. In particular, the map has a convergent expansion in fractional powers of ρ in a right neighborhood of 0 (and a convergent expansion in fractional powers of ρ in any one-sided neighborhood of a point $0 < \rho_0 < 1$). As V_ρ is the the $\rho\delta$-discounted value of the discrete-time stochastic game with payoff function δg

[5]The finiteness follows from the fact that the minimum and the maximum of a linear function over a simplex is attained in one of the finitely many extreme points of the simplex.

and transitions $p_\delta = \frac{\delta}{1-\rho\delta}\mu$ it is bounded by $\|g\|/\rho$. Therefore $\rho \mapsto v_\rho := \rho V_\rho$ is a bounded semialgebraic function. In particular, there is 1) a positive integer M, 2) real coefficients $c_k(z)$, and 3) a positive discount rate $\bar\rho > 0$, such that for $0 < \rho \leq \bar\rho$ the series $\sum_{k=0}^\infty c_k(z)\rho^{k/M}$ converges and

$$v_\rho(z) = \sum_{k=0}^\infty c_k(z)\rho^{k/M}.$$

If the game is one of perfect information, then each player has for each $1 > \rho > 0$ a pure stationary strategy that is an asymptotic ρ-discounted optimal strategy. Therefore (following the classical argument from discrete-time stochastic games) the value function $\rho \mapsto v_\rho(z)$ is a rational function in ρ in a right neighborhood of 0 (and in any one-sided neighborhood of a point $1 > \rho_0 > 0$). It follows that there are $\bar\rho > 0$ and real coefficients $c_k(z)$, and pure stationary strategies σ^i, $i = 1, 2$, such that for $\rho \leq \bar\rho$ the series $\sum_{k=0}^\infty c_k(z)\rho^k$ converges,

$$v_\rho(z) = \sum_{k=0}^\infty c_k(z)\rho^k,$$

and σ^i is asymptotic ρ-discounted optimal in the family $(\Gamma_\delta)_{\delta>0}$.

Covariance properties. Fix the sets of states S and actions A. Let $V_\rho(g, \mu)$ be the unique solution of the system (2) of S equations. Recall that it equals the asymptotic ρ-discounted value of any family $\langle g_\delta, p_\delta\rangle$ that converges in data to $\langle g, \mu\rangle$. (It is also the value of the continuous-time stochastic game $\langle N, S, A, g, \mu\rangle$, e.g., [8].) Consider the function $V_\rho(g, \mu)$ as a function of ρ, g, and μ. Obviously, the ρ-discounted asymptotic value $V_\rho(g, \mu)$ is monotonic in g and covariant with respect to multiplication of the payoff function g by a positive scalar. Namely, if $g' \geq g$ and α is a nonnegative real number, $V_\rho(g', \mu) \geq V_\rho(g, \mu)$ and $V_\rho(\alpha g, \mu) = \alpha V_\rho(g, \mu)$. For $\alpha > 0$, a vector V satisfies equation (2) if and only if it satisfies the same equation when ρ is replaced by $\alpha\rho$, g is replaced by αg, and μ is replaced by $\alpha\mu$. Therefore, $V_{\alpha\rho}(\alpha g, \alpha\mu) = V_\rho(g, \mu)$. (In the continuous-time game interpretation, this equality is interpreted as, and can be derived by, a simple rescaling of time: $t \mapsto \alpha t$.)

Now we turn to the expression of the ρ-discounted asymptotic value as a value of a discrete-time discounted stochastic game.

If $\|\mu\| \leq 1$, we assign to (the continuous-time game) $\Gamma = \langle N, S, A, g, \mu\rangle$ the discrete-time game $\bar{\Gamma} = \langle N, S, A, g, p = \mu\rangle$. By Remark 5, the value $\bar{V}_\rho(g,\mu)$ of the discrete-time ρ-discounted (with discount factor $1-\rho$) stochastic game $\bar{\Gamma} = \langle \{1,2\}, S, A, g, p = \mu\rangle$ equals $V_\rho(g,(1-\rho)\mu)$ whenever $0 < \rho \leq 1 - \|\mu\|$.

Summarizing,

$$\begin{aligned} V_\rho(g,\mu) &\geq V_\rho(g',\mu) \text{ whenever } g \geq g', && (12)\\ V_\rho(\alpha g, \beta\mu) &= \frac{\alpha}{\beta} V_{\rho/\beta}(g,\mu) \text{ whenever } \alpha \geq 0 \text{ and } \beta > 0, && (13)\\ V_\rho(g,\mu) &= \bar{V}_\rho(g, \frac{1}{1-\rho}\mu) \text{ whenever } 0 < \rho \leq 1 - \|\mu\|; && (14) \end{aligned}$$

equivalently,

$$\bar{V}_\rho(g,\mu) = V_\rho(g,(1-\rho)\mu) \text{ whenever } 0 < \rho < 1 \text{ and } \|\mu\| \leq 1. \quad (15)$$

Note that for a constant payoff function $g = c$, we have $\rho V_\rho(c,\mu) = c$. The normalization $v_\rho := \rho V_\rho$ of the function V_ρ, is a function of (g,μ): $v_\rho(g,\mu) = \rho V_\rho(g,\mu)$. Given two transition rates μ and μ',

$$d(\mu,\mu') := \max\left\{\frac{\mu(z',z,a)}{\mu'(z',z,a)}, \frac{\mu'(z',z,a)}{\mu(z',z,a)} \mid a \in A(z), z, z' \in S\right\} - 1,$$

where by convention $x/0 = \infty$ for $x > 0$, and $0/0 = 1$.

Lemma 1 *For every pair of payoff functions g and g' and every pair of transition rates μ and μ' the following inequality holds:*

$$\|v_\rho(g',\mu') - v_\rho(g,\mu)\|_\infty \leq 4|S| d(\mu,\mu') \min\{\|g\|, \|g'\|\} + \|g - g'\|. \quad (16)$$

Proof. The proof applies [10, Theorem 6] in conjunction with the covariance properties (13) and (14). Fix ρ, g, g', μ, μ'. Let $\beta > 0$, and note that $d(\mu,\mu') = d(\mu/\beta, \mu'/\beta)$. As $\mu = \beta\mu/\beta$, equality (13) implies that $v_\rho(g,\mu) = \frac{\rho}{\beta} V_{\frac{\rho}{\beta}}(g,\mu/\beta) = v_{\frac{\rho}{\beta}}(g,\mu/\beta)$, and similarly, $v_\rho(g',\mu') = v_{\frac{\rho}{\beta}}(g',\mu'/\beta)$. We choose $\beta > 0$ sufficiently large, e.g., $\beta > \rho + \frac{\max\{\|\mu\|,\|\mu'\|\}}{1-\rho}$, so that $\rho/\beta < 1 - \frac{\|\mu\|}{(1-\rho)\beta}$ and $\rho/\beta < 1 - \frac{\|\mu'\|}{(1-\rho)\beta}$. This will enable us to apply equality

(14) in the third equality below. Therefore,

$$\begin{aligned}
\|v_\rho(g',\mu') - v_\rho(g,\mu)\|_\infty &= \|v_{\rho/\beta}(g',\mu'/\beta) - v_{\rho/\beta}(g,\mu/\beta)\|_\infty \\
&= \|\frac{\rho}{\beta}V_{\rho/\beta}(g',\mu'/\beta) - \frac{\rho}{\beta}V_{\rho/\beta}(g,\mu/\beta)\|_\infty \\
&= \|\frac{\rho}{\beta}\bar{V}_{\rho/\beta}(g',\frac{\mu'}{(1-\rho)\beta}) - \frac{\rho}{\beta}\bar{V}_{\rho/\beta}(g,\frac{\mu}{(1-\rho)\beta})\|_\infty \\
&\leq 4|S|d(\mu,\mu')\min\{\|g\|,\|g'\|\} + \|g-g'\|,
\end{aligned}$$

where the first and second equalities follow from (13), the third equality follows from (14), and the last inequality follows from [10, Theorem 6]. □

Recall that the family of discrete-time stochastic games $\Gamma_\delta = \langle N, S, A, g_\delta, p_\delta\rangle$ converges strongly to $\Gamma = \langle N, S, A, g, \mu\rangle$ if for all $(z', z, a) \in S \times \mathcal{A}$, $g_\delta(z,a) = \delta g(z,a) + o(\delta)$ and $p_\delta(z',z,a) = \delta\mu(z',z,a)(1+o(1))$ as $\delta \to 0+$.

Theorem 2 *If* $\Gamma_\delta = \langle g_\delta, p_\delta\rangle$ *converges strongly to* $\Gamma = \langle g, \mu\rangle$ *then* $\rho V_{\delta,\rho} \to_{\delta\to 0+} \rho V_\rho(\mu, g)$ *uniformly in* $0 < \rho < 1$.

Proof. By Remark 5, $V_{\delta,\rho} = V_\rho(g_\delta/\delta, (1-\rho\delta)p_\delta/\delta)$. Therefore, $v_{\delta,\rho} = \rho V_{\delta,\rho} = v_\rho(g'_\delta, \mu_{\delta,\rho}) := \rho V_\rho(g'_\delta, \mu_{\delta,\rho})$, where $g'_\delta = g_\delta/\delta$ and $\mu_{\delta,\rho}(z',z,a) = (1-\rho\delta)p_\delta/\delta$. Therefore, as $\|g' - g\| \to 0$ as $\delta \to 0+$ and $d(\mu, \mu_{\delta,\rho}) \to_{\delta\to 0+} 0$ uniformly in ρ, inequality (16) implies that $\rho V_{\delta,\rho} = v_\rho(g'_\delta, \mu_{\delta,\rho}) \to_{\delta\to 0+} v_\rho(g,\mu)$ uniformly in ρ. □

4.2 The asymptotic nonstationary discounted value

We start with a few simple and useful properties of nonstationary discounting measures. First, if w is a nonstationary discounting measure on $[0,\infty]$ then w has no atoms in $(0,\infty)$, w is absolutely continuous on $(0,\infty)$, and $\frac{dw}{dt}(t)$ is nonincreasing in $0 < t < \infty$. Given a nonstationary discounting measure w on $[0,\infty]$ and a finite sequence $\tilde{t} = (t_0 = 0 < t_1 < \ldots < t_\ell < \infty)$, we define the nonstationary discounting measure $\tilde{w}_{\tilde{t}}$, or $\tilde{w}$ for short, on $[0,\infty]$ by $\tilde{w}([t_j, t_{j+1})) = w([t_j, t_{j+1}))$, $\frac{d\tilde{w}}{dt}(t)$ being a constant (thus, $\frac{d\tilde{w}}{dt}(t) = w([t_j,t_{j+1}))/(t_{j+1}-t_j)$) on each interval $[t_j, t_{j+1})$ $(0 \leq j < \ell)$, and $\tilde{w}$ coincides with w on subsets of $[t_\ell, \infty]$. Set $d(\tilde{t}) := \max_{0\leq j<\ell}(t_{j+1}-t_j)$.

Lemma 2 *Let* w *be a nonstationary measure on* $[0,\infty]$ *and* $\tilde{t} = (t_0 = 0 < t_1 < \ldots < t_\ell < \infty)$ *a finite sequence. Then,*

$$\int_{t_1}^{t_\ell} |\frac{dw}{dt}(t) - \frac{d\tilde{w}}{dt}(t)|\, dt \leq 2\int_{t_1}^{t_1+d(\tilde{t})} \frac{dw}{dt}(t)\, dt, \tag{17}$$

and if the nonstationary discounting measures w_δ on $[0,\infty]$ converge to the measure w on $[0,\infty]$ then

$$\int_{t_1}^{t_\ell} |\frac{dw_\delta}{dt}(t) - \frac{dw}{dt}(t)|\, dt \to_{\delta\to 0+} 0. \tag{18}$$

Proof. As $\frac{dw}{dt}(t)$ is nonincreasing in t, $\int_{t_j}^{t_{j+1}} |\frac{dw}{dt}(t) - \frac{d\tilde{w}}{dt}(t)|\, dt \le 2\int_{t_j}^{t_{j+1}} \frac{dw}{dt}(t) - \frac{dw}{dt}(t+d(\tilde{t}))\, dt$. Therefore $\int_{t_1}^{t_\ell} |\frac{dw}{dt}(t) - \frac{d\tilde{w}}{dt}(t)|\, dt = \sum_{1\le j<\ell} \int_{t_j}^{t_{j+1}} |\frac{dw}{dt}(t) - \frac{d\tilde{w}}{dt}(t)|\, dt \le 2\sum_{1\le j<\ell} \int_{t_j}^{t_{j+1}} \frac{dw}{dt}(t) - \frac{dw}{dt}(t+d(\tilde{t}))\, dt \le 2\int_{t_1}^{t_1+d(\tilde{t})} \frac{dw}{dt}(t)\, dt$, which proves (17).

In order to prove (18), it suffices to prove that for every $\varepsilon > 0$ there is $\delta_0 > 0$ such that for $0 < \delta < \delta_0$, $\int_{t_1}^{t_\ell} |\frac{dw_\delta}{dt}(t) - \frac{dw}{dt}(t)|\, dt < 4\varepsilon$. Fix $\varepsilon > 0$.

For every $d > 0$ and a nonstationary discounting measure ν on $[0,\infty]$, we define the nonstationary discounting measures ν^d on $[0,\infty]$ by $\nu^d([a,b]) = \frac{1}{d}\int_0^d \nu([a+t, b+t])\, dt$. Note that $\frac{dw^d}{dt}(t)$ and $\frac{dw_\delta^d}{dt}(t)$ are continuous at each $t < \infty$ and $\frac{dw_\delta^d}{dt}(t) \to_{\delta\to 0+} \frac{dw^d}{dt}(t)$. Therefore, $\int_{t_1}^{t_\ell} |\frac{dw_\delta^d}{dt}(t) - \frac{dw^d}{dt}(t)|\, dt \to_{\delta\to 0+} 0$. As $\frac{dw}{dt}(t)$ is nonincreasing in t, $\int_{t_1}^{t_\ell} |\frac{dw}{dt}(t) - \frac{dw^d}{dt}(t)|\, dt = \int_{t_1}^{t_\ell} \frac{dw}{dt}(t) - \frac{dw^d}{dt}(t)\, dt \le \int_{t_1}^{t_1+d} \frac{dw}{dt}(t)\, dt - \int_{t_\ell}^{t_\ell+d} \frac{dw}{dt}(t)\, dt \le w([t_1, t_1+d])$. Similarly, $\int_{t_1}^{t_\ell} |\frac{dw_\delta}{dt}(t) - \frac{dw_\delta^d}{dt}(t)|\, dt \le w_\delta([t_1, t_1+d])$. Let $d > 0$ be sufficiently small so that $w([t_1, t_1+d]) < \varepsilon$, and $\delta_0 > 0$ be sufficiently small so that for all $0 < \delta < \delta_0$, $\int_{t_1}^{t_\ell} |\frac{dw_\delta^d}{dt}(t) - \frac{dw^d}{dt}(t)|\, dt < \varepsilon$. Therefore, as $|\frac{dw_\delta}{dt}(t) - \frac{dw}{dt}(t)| \le |\frac{dw_\delta}{dt}(t) - \frac{dw_\delta^d}{dt}(t)| + |\frac{dw_\delta^d}{dt}(t) - \frac{dw^d}{dt}(t)| + |\frac{dw^d}{dt}(t) - \frac{dw}{dt}(t)|$,

$$\int_{t_1}^{t_\ell} |\frac{dw_\delta}{dt}(t) - \frac{dw}{dt}(t)|\, dt < 4\varepsilon.$$

□

Theorem 3 *Let w be a nonstationary discounting measure on $[0,\infty)$, $t \ge 0$, and $\nu : \mathcal{A} \to \mathbb{R}$. Then a family $(\Gamma_\delta)_{\delta>0}$ that converges in data has an asymptotic (w,t,ν) value, and if w_δ, $\delta > 0$, are nonstationary discounting measures on $\mathbb{N}$ that converge to w, and $m_\delta \ge 0$ and $\nu_\delta : \mathcal{A} \to \mathbb{R}$ are such that (m_δ, ν_δ) converges to (t,ν) (as $\delta \to 0+$), then for every $\varepsilon > 0$ there are ε-optimal Markov strategies in $\Gamma_{\delta,w_\delta}^{m_\delta,\nu_\delta}$ that converge to a continuous-time Markov strategy.*

Before turning to the proof of the theorem, we introduce a useful auxiliary lemma.

Fix a payoff function $g : \mathcal{A} \to \mathbb{R}$ and a transition rate function $\mu : S \times \mathcal{A} \to \mathbb{R}$ with $\mu(z', z, a) \geq 0$ if $z' \neq z$ and $\sum_{z' \in S} \mu(z', z, a) = 0$. Let $\|\mu\| := \max_{(z,a) \in \mathcal{A}} |\mu(z, z, a)|$. For every $z \in S$, $\alpha, \beta > 0$, $V \in \mathbb{R}^S$, and $x \in \Delta(A(z))$, $F(z, x, \alpha, \beta, V)$ is defined by

$$F(z, x, \alpha, \beta, V) = \alpha g(z, x) + V(z) + \sum_{z' \in S} \beta \mu(z', z, x) V(z'),$$

and $T(\alpha, \beta, V) \in \mathbb{R}^S$ is defined by

$$T(\alpha, \beta, V)(z) = \max_{x^1 \in X^1(z)} \min_{x^2 \in X^2(z)} F(z, x^1 \otimes x^2, \alpha, \beta, V).$$

Let $V_1 \in \mathbb{R}^S$, $\alpha, \beta > 0$, and define $V_0 \in \mathbb{R}^S$ by $V_0 = T(\alpha, \beta, V_1)$. Given a sequence $\gamma = (0 = \gamma_0 < \ldots < \gamma_m = 1)$, define U_{γ_j}, $0 \leq j \leq m$ (recursively in j) by $U_1 = U_{\gamma_m} = V_1$, and for $0 \leq j < m$ and $z \in S$, $U_{\gamma_j} = T((\gamma_{j+1} - \gamma_j)\alpha, (\gamma_{j+1} - \gamma_j)\beta, U_{\gamma_{j+1}})$. If $d(\gamma) := \max_{0 \leq j < m} \gamma_{j+1} - \gamma_j$ is sufficiently small so that $d(\gamma)\beta\|\mu\| \leq 1$, $\Gamma(\gamma)$ denotes the m-stage game with set of plays $S \times \mathcal{A}^m$, the payoff of a play $z_0, a_0, \ldots, z_m$ is $V_1(z_m) + \sum_{0 \leq j < m}(\gamma_{j+1} - \gamma_j)\alpha g(z_j, a_j)$, and past play is observed by the players, and the "states transitions" are such that the conditional probability of $z_{j+1} = z$, given $z_0, a_0, \ldots, a_j$, is $I_{z_j, z} + (\gamma_{j+1} - \gamma_j)\beta\mu(z, z_j, a_j)$.

Lemma 3 *Assume that $\beta\|\mu\| \leq 1/2$. Then, 1) the game $\Gamma(\gamma)$ is well defined and its value equals U_0, 2) the stationary strategy σ of player 1 (respectively, τ of player 2) that for every state $z \in S$, $\sigma(z)$ maximizes $\min_{x^2 \in X^2(z)} F(z, \sigma(z) \otimes x^2, \alpha, \beta, V_1)$, (respectively, $\tau(z)$ minimizes $\max_{x^1 \in X^1(z)} F(z, x^1 \otimes \tau(z), \alpha, \beta, V_1)$) is $(4\beta\|\mu\|(\alpha\|g\| + 4\beta\|\mu\|\|V_1\|)$-optimal in $\Gamma(\gamma)$, and 3) $\|U_0 - V_0\| \leq 4\beta\|\mu\|(\alpha\|g\| + 4\beta\|\mu\|\|V_1\|)$.*

Proof. For every $(z_j, a_j) \in \mathcal{A}$, the condition $d(\gamma)\beta\|\mu\| \leq 1$ implies that $I_{z_j, z} + (\gamma_{j+1} - \gamma_j)\beta\mu(z, z_j, a_j) \geq 0$, and in addition $\sum_{z \in S}(I_{z_j, z} + (\gamma_{j+1} - \gamma_j)\beta\mu(z, z_j, a_j)) = 1$. Therefore $\Gamma(\gamma)$ is well defined. The recursive formula for the value of the m-stage game $\Gamma(\gamma)$ shows that the value of $\Gamma(\gamma)$ equals U_0.

For every strategy profile σ in $\Gamma(\gamma)$ and state z, $P^z_\sigma(z_0 = z_1 = \ldots = z_m) \geq \prod_{0 \leq j < m}(1 - (\gamma_{j+1} - \gamma_j)\beta\|\mu\|) \geq 1 - \beta\|\mu\|$. Therefore, for every Markov strategy profile σ in $\Gamma(\gamma)$ and state z,

$$E^z_\sigma \sum_{0 \leq j < m} (\gamma_{j+1} - \gamma_j)\alpha g(z_j, a_j) \geq \alpha g(z, \bar{\sigma}(z)) - 2\beta\|\mu\|\alpha\|g\|,$$

where $\bar\sigma(z) = \sum_{0\le j<m}(\gamma_{j+1}-\gamma_j)\sigma(z,j)$.

Let σ^1 be a stationary strategy of player 1 in $\Gamma(\gamma)$ such that for every state $z \in S$, $\sigma^1(z)$ maximizes $\min_{x^2\in X^2(z)} F(z,\sigma(z)\otimes x^2,\alpha,\beta,V_1)$. Then for every Markov strategy σ^2 of player 2 in $\Gamma(\gamma)$, inequality (1) implies that $\sum_{z'\in S}|P^z_\sigma(z_m = z') - I_{z,z'} - \beta\mu(z',z,\bar\sigma(z))| \le e^{2\beta\|\mu\|} - 1 - 2\beta\|\mu\| \le 4\beta^2\|\mu\|^2$, where σ is the strategy profile (σ^1,σ^2) and the last inequality uses the assumption $2\beta\|\mu\| \le 1$. Therefore,

$$E^z_\sigma\left(V_1(z_m) + \sum_{0\le j<m}(\gamma_{j+1}-\gamma_j)\alpha g(z_j,a_j)\right) \ge V_0(z) - 2\beta\|\mu\|(\alpha\|g\| + 4\beta\|\mu\|\|V_1\|).$$

Let τ^2 be a stationary strategy of player 2 in $\Gamma(\gamma)$ such that for every state $z \in S$, $\tau^2(z)$ minimizes $\max_{x^1\in X^1(z)} F(z,\sigma(z)\otimes x^2,\alpha,\beta,V_1)$. Then, by duality, for every Markov strategy τ^1 of player 1 in $\Gamma(\gamma)$,

$$E^z_\tau(V_1(z_m) + \sum_{0\le j<m}(\gamma_{j+1}-\gamma_j)\alpha g(z_j,a_j) \le V_0(z) + 2\beta\|\mu\|(\alpha\|g\| + 4\beta\|\mu\|\|V_1\|),$$

where $\tau = (\tau^1,\tau^2)$.

Therefore, $\|U_0 - V_0\| \le 4\beta\|\mu\|(\alpha\|g\| + 4\beta\|\mu\|\|V_1\|)$ and σ^1 and τ^2 are $(4\beta\|\mu\|(\alpha\|g\| + 4\beta\|\mu\|\|V_1\|))$-optimal. □

Proof of Theorem 3. The first stage of the proof is obtained by associating an extensive form ℓ-stage game $\Gamma(\tilde t)$ with a finite sequence $\tilde t = (t_0 = 0 < t_1 < \ldots < t_k = t < t_{k+1} < \ldots < t_\ell)$ of times (and the triple (w,t,ν)) as follows.

The game $\Gamma(\tilde t)$ is an ℓ-stage "stochastic game" with 1) the same sets of states, actions, and players as in Γ_δ, 2) stage-dependent payoffs (that also incorporate an extra payment in stage k), and 3) stage-dependent transitions. Let $\Delta_j := t_{j+1} - t_j$ and let $\tilde t$ be such that $d(\tilde t)$ is sufficiently small so that $d(\tilde t)\|\mu\| < 1/2$. A play of $\Gamma(\tilde t)$ is a sequence $(\tilde z_0,\tilde a_0,\ldots,\tilde z_\ell)$ with $\tilde a_j \in A(\tilde z_j)$ and the payoff of the play $(\tilde z_0,\tilde a_0,\ldots,\tilde z_\ell)$ is $\nu(\tilde z_k,\tilde a_k) + \sum_{j=0}^{\ell-1} w_j g(\tilde z_j,\tilde a_j)$, where $w_j := w([t_j,t_{j+1}))$.

Past play is observed by the players. Therefore, a strategy of a player chooses his action at stage $j = 0,\ldots,\ell-1$ as a function of $(\tilde z_0,\tilde a_0,\ldots,\tilde z_j)$. The conditional probability, given $\tilde z_0,\tilde a_0,\ldots,\tilde z_j,\tilde a_j$, of $\tilde z_{j+1} = z$ is $\Delta_j\mu(z,\tilde z_j,\tilde a_j) + I_{\tilde z_j,z}$. It is helpful to view the states transitions in $\Gamma(\tilde t)$ as those of an "exact" stochastic game whose j-th stage duration, $0 \le j < \ell$, is Δ_j. The game $\Gamma(\tilde t)$ has a value $\tilde V$ and the players have Markovian optimal strategies.

The value $\tilde{V}$ equals $\tilde{V}_0$, where $\tilde{V}_j \in \mathbb{R}^S$ are defined recursively for $0 \leq j \leq \ell$. For every $z \in S$, $\tilde{V}_\ell(z) = 0$, and for $0 \leq j < \ell$ we define $\tilde{V}_j(z)$ by

$$\tilde{V}_j(z) = \max_{x^1 \in X^1(z)} \min_{x^2 \in X^2(z)} \left(1_{j=k}\nu(z,x) + F(z,x,w_j,\Delta_j,\tilde{V}_{j+1}) \right),$$

where $x = x^1 \otimes x^2$.

Note that for every $j < \ell$, $\|\tilde{V}_j\| \leq 1_{j=k}\|\nu\| + w_j\|g\| + \|\tilde{V}_{j+1}\|$, where $\|\nu\| = \max_{(z,a)\in\mathcal{A}} |\nu(z,a)|$. Therefore, by induction on $0 \leq \ell - j \leq \ell$, $\|\tilde{V}_j\| \leq 1_{j\leq k}\|\nu\| + \sum_{j'\geq j} w_{j'}\|g\| \leq \|\nu\| + w([0,\infty))\|g\|$.

The Markov strategy $\tilde{\sigma}$ of player 1 in $\Gamma(\tilde{t})$ with $\tilde{\sigma}(z,j)$ maximizing (over all $x^1 \in X^1(z)$)

$$\min_{x^2 \in X^2(z)} \left(1_{j=k}\nu(z, x^1 \otimes x^2) + F(z, x^1 \otimes x^2, w_j, \Delta_j, \tilde{V}_{j+1}) \right)$$

is an optimal strategy of player 1 in $\Gamma(\tilde{t})$. Indeed, for every strategy τ of player 2 in $\Gamma(\tilde{t})$ and stage $0 \leq j < \ell$,

$$E^z_{\tilde{\sigma},\tau} \left(1_{j=k}\nu(\tilde{z}_j, \tilde{a}_j) + w_j g(\tilde{z}_j, \tilde{a}_j)\right) \geq E^z_{\tilde{\sigma},\tau} \left(\tilde{V}_j(\tilde{z}_j) - \tilde{V}_{j+1}(\tilde{z}_{j+1}) \right).$$

Therefore, by summing these inequalities over $0 \leq j < \ell$, we have

$$E^z_{\tilde{\sigma},\tau} \left(\nu(\tilde{z}_k, \tilde{a}_k) + \sum_{0\leq j<\ell} w_j g(\tilde{z}_j, \tilde{a}_j) \right) \geq \tilde{V}_0(z).$$

The second stage of the proof is to associate with $\tilde{t}$, $\tilde{\sigma}$, and $\delta > 0$, a sequence $\tilde{m}_\delta = (m_{\delta,0} = 0 < m_{\delta,1} < \ldots < m_{\delta,\ell})$, a Markov strategy σ_δ in Γ_δ, and a nonstationary discounting measure $\tilde{w}_\delta$, as follows.

For $m_{\delta,j} \leq m < m_{\delta,j+1}$, $\sigma_\delta(z,m) = \tilde{\sigma}(z,j)$, for $m \geq m_{\delta,\ell}$, $\sigma_\delta(z,m)$ coincides with an arbitrary stationary strategy, $m_{\delta,k} = m_\delta$, $m_{\delta,j} = [t_j/\delta]$ for $j \neq k$ (thus $\delta m_{\delta,j} \to_{\delta\to 0+} t_j$ for all $0 \leq j < \ell$, $\tilde{w}_\delta(m) = w_\delta(m)$ for $m \geq m_{\delta,\ell}$, and $\tilde{w}_\delta(m) = \frac{1}{m_{\delta,j+1}-m_{\delta,j}} \sum_{m_{\delta,j}\leq m<m_{\delta,j+1}} w_\delta(m)$ for $m_{\delta,j} \leq m < m_{\delta,j+1}$ and $j < \ell$.

Note that $\tilde{w}_\delta$ is a nonstationary discounting measure that converges, as $\delta \to 0+$, to w.

Consider the family of games $\tilde{\Gamma}^{m_\delta,\nu_\delta}_{\delta,\tilde{w}_\delta}$ with $\tilde{g}_\delta = \delta g$ and $\tilde{p}_\delta = \delta\mu$. By Lemma 3, for every $\varepsilon > 0$, there is a sufficiently small $d > 0$ such that if $\tilde{t}$ is such that $d(\tilde{t}) < d$ and $w([t_\ell,\infty)) < d$, then, for sufficiently small $\delta > 0$, the Markov strategy σ_δ guarantees in $\tilde{\Gamma}^{m_\delta,\nu_\delta}_{\delta,\tilde{w}_\delta}$ a payoff that is at least $\tilde{V}-\varepsilon$. Therefore, for sufficiently small $\delta>0$, the Markov strategy σ_δ guarantees in $\Gamma^{m_\delta,\nu_\delta}_{\delta,\tilde{w}_\delta}$ a payoff that is at least $\tilde{V}-2\varepsilon$.

Note that for sufficiently small $\delta>0$, $P^z_\sigma(z_m=z \quad \forall m \leq m_{\delta,1}) \geq 1-d\|\mu\|$ for every strategy profile σ and state z. Therefore, if $d\|\mu\|\|\tilde{V}_1\|<\varepsilon/4$, for sufficiently small $\delta>0$, for every strategy τ of player 2, we have $E^z_{\sigma_\delta,\tau}(\tilde{V}_1(z_{m_{\delta,1}})+\sum_{m<m_{\delta,1}} w_\delta(m)g_\delta(z_m,a_m)) \geq \tilde{V}_0(z)-2d\|\mu\|\|\tilde{V}_1\|-\varepsilon/2 > \tilde{V}_0(z)-\varepsilon$.

By Lemma 2, $\sum_{m\geq m_{\delta,1}}|\tilde{w}_\delta(m)-w_\delta(m)| \to 0$ as $\delta\to 0+$. If $\delta>0$ is sufficiently small so that $\sum_{m\geq m_{\delta,1}}|\tilde{w}_\delta(m)-w_\delta(m)|<\varepsilon$, then σ_δ guarantees in $\Gamma^{m_\delta,\nu_\delta}_{\delta,w_\delta}$ a payoff that is at least $\tilde{V}-3\varepsilon-\varepsilon\|g\|$. By the construction of σ_δ, σ_δ converges to a continuous-time Markov strategy.

Similarly, we associate with the Markov strategy $\tilde{\tau}$ (and $\delta>0$) a Markov strategy τ_δ that for $\delta>0$ sufficiently small guarantees in $\Gamma^{m_\delta,\nu_\delta}_{\delta,w_\delta}$ a payoff that is at most $\tilde{V}+3\varepsilon+\varepsilon\|g\|$ while τ_δ converges to a continuous-time strategy τ. □

4.3 The asymptotic limiting-average value

Recall that the family $(\Gamma_\delta)_{\delta>0}$ has an asymptotic limiting-average value v if for every $\varepsilon>0$ there are $\delta_0>0$ sufficiently small and strategies σ_δ and τ_δ in Γ_δ, such that for every strategy pair (σ^*,τ^*), every initial state z, and every $0<\delta<\delta_0$, we have

$$\varepsilon+E^z_{\sigma_\delta,\tau^*}\underline{g}_\delta \geq v(z) \geq -\varepsilon+E^z_{\sigma^*,\tau_\delta}\bar{g}_\delta. \tag{19}$$

Theorem 4 *A family $(\Gamma_\delta)_{\delta>0}$ that converges strongly has an asymptotic limiting-average value.*

Proof. Let $g=\lim_{\delta\to 0+} g_\delta/\delta$ and $\mu=\lim_{\delta\to 0+} p_\delta/\delta$. As the function $\rho\mapsto v_\rho(g,\mu)$ is semialgebraic and bounded, it converges to a limit v as $\rho\to 0+$. Fix $\varepsilon>0$. As every discrete-time stochastic game with finitely many states and actions has a limiting-average value [5], which is the limit of its ρ-discounted values as ρ goes to 0+, there are strategies σ_δ of player 1 and τ_δ of player 2, such that for every strategy pair (σ^*,τ^*) and every initial state $z\in S$,

$$\varepsilon/2+E^z_{\sigma_\delta,\tau^*}\underline{g}_\delta \geq \lim_{\rho\to 0+} v_{\delta,\rho}(z) \geq -\varepsilon/2+E^z_{\sigma^*,\tau_\delta}\bar{g}_\delta. \tag{20}$$

As $v_{\delta,\rho} \to v_\rho(g,\mu)$ uniformly in ρ, there is $\delta_0 > 0$ such that for every $0 < \delta < \delta_0$ and every state $z \in S$, $|v_{\delta,\rho}(z) - v_\rho(g,\mu)(z)| < \varepsilon/2$. Therefore, for $0 < \delta < \delta_0$, $|\lim_{\rho\to 0+} v_{\delta,\rho}(z) - v(z)| \leq \varepsilon/2$, which together with (20) implies (19). □

Remark 10 *A family $(\Gamma_\delta)_{\delta>0}$ that converges in data need not have an asymptotic limiting-average value.*

For example, consider a game with two states and a single action for each player in each state. The payoff in state one is 1 and in state 2 it is 0. State 2 is absorbing, i.e., $P_\delta(1 \mid 2) = 0$, and the probability of transition from state 1 to state 2, $P_\delta(2 \mid 1)$, equals δ^2 if δ is rational, and it equals 0 if δ is irrational. Then, $v_{\delta,0} = 0$ if δ is rational, and $v_{\delta,0} = 1$ if δ is irrational. Therefore $v_{\delta,0}$ does not converge as δ goes to 0.

4.4 The asymptotic mixed discounting and limiting-average value

For every positive measure w_δ on $\mathbb{N} \cup \{\infty\}$, Γ_{δ,w_δ} is the game Γ_δ where the valuation of a play $(z_0, a_0, z_1, \ldots)$ of Γ_δ is given by $\sum_{m=0}^{\infty} w_\delta(m) g_\delta(z_m, a_m) + w_\delta(\infty) \lim_{s\to\infty} g_\delta(s)$, if the limit exists. Obviously, the limit need not exist.

We say that the two-person zero-sum game Γ_{δ,w_δ} has a value V_{δ,w_δ}, if for every $\varepsilon > 0$ there are strategies σ_δ of player 1 and τ_δ of player 2, such that for every strategy τ of player 2, strategy σ of player 1, and initial state z, we have

$$E^z_{\sigma_\delta,\tau}\left(w_\delta(\infty)\underline{g}_\delta + \sum_{m=0}^{\infty} w_\delta(m) g_\delta(z_m, a_m)\right) \geq V_{\delta,w_\delta}(z) - \varepsilon$$

and

$$E^z_{\sigma,\tau_\delta}\left(w_\delta(\infty)\bar{g}_\delta + \sum_{m=0}^{\infty} w_\delta(m) g_\delta(z_m, a_m)\right) \leq V_{\delta,w_\delta}(z) + \varepsilon.$$

Theorem 5 *If Γ_δ converges strongly and the nonstationary discounting measure w_δ converges to a positive measure w on $[0,\infty]$, and $w_\delta(\infty)$ converges to $w(\infty)$, then V_{δ,w_δ} converges.*

Proof. The proof is obtained by collating the result of Theorem 3 with the result of Theorem 4. Let $0 < \varepsilon < 1$. Let $0 < t < \infty$ be sufficiently large so that $2w([t,\infty))\|g\| < \varepsilon$, and let w_t be the restriction of w to the interval $[0,t)$. Let v be the asymptotic limiting-average value of the family $(\Gamma_\delta)_{\delta>0}$, and define $\nu : \mathcal{A} \to \mathbb{R}$ by $\nu(z,a) = w(\infty)v(z)$. The family $(\Gamma_\delta)_{\delta>0}$ has an asymptotic (w_t, t, ν) value V.

Assume that the nonstationary discounting measure w_δ converges to w and $w_\delta(\infty)$ converges to $w(\infty)$. Let $m_\delta = [t/\delta]$ and let $w_{\delta,t}$ be the restriction of w_δ to $\{0,1,\ldots,m_\delta\}$.

The value $V^{m_\delta,\nu}_{\delta,w_{\delta,t}}$ of the game $\Gamma^{m_\delta,\nu}_{\delta,w_{\delta,t}}$ converges to V. Recall that as Γ_δ converges strongly, the limiting-average value of the game Γ_δ, which is denoted by $v_{\delta,0}$, converges as δ goes to zero to v. Let δ_0 be sufficiently small so that for $0<\delta<\delta_0$, 1) $\|V^{m_\delta,\nu}_{\delta,w_{\delta,t}} - V\| < \varepsilon$, 2) $\|v_{\delta,0} - v\| < \varepsilon$, 3) $\|w_\delta(\infty) - w(\infty)\| < \varepsilon$, and 4) $\|g\| \sum_{m=m_\delta}^\infty w_\delta(m)| < \varepsilon$.

Let σ_δ follow an optimal strategy in $\Gamma^{m_\delta,\nu}_{\delta,w_{\delta,t}}$ up to stage m_δ, and thereafter it "restarts" with an ε-optimal strategy in the limiting-average game Γ_δ. It follows that for every $0<\delta<\delta_0$ and strategy τ of player 2,

$$E^z_{\sigma_\delta,\tau}\left(w_\delta(\infty)\bar{g}_\delta + \sum_{m=0}^{\infty} w_\delta(m) g_\delta(z_m,a_m)\right) \geq V(z) - \varepsilon w(\infty) - 3\varepsilon.$$

Similarly, if τ_δ follows an optimal strategy in $\Gamma^{m_\delta,\nu}_{\delta,w_{\delta,t}}$ up to stage m_δ, and thereafter it "restarts" with an ε-optimal strategy in the limiting-average game Γ_δ, then for every $0<\delta<\delta_0$ and strategy σ of player 1,

$$E^z_{\sigma,\tau_\delta}\left(w_\delta(\infty)\bar{g}_\delta + \sum_{m=0}^{\infty} w_\delta(m) g_\delta(z_m,a_m)\right) \leq V(z) + \varepsilon w(\infty) + 3\varepsilon.$$

□

4.5 The asymptotic uniform and w-robust value

Theorem 6 *An exact family of two-person zero-sum games Γ_δ has an asymptotic uniform value.*

Proof. Let $v = \lim_{\delta\to 0+} v_{\delta,0}$. It is sufficient to prove that for every $\varepsilon > 0$ there are 1) a duration $\delta_0 > 0$, 2) strategies σ_δ of player 1 and τ_δ of player 2, and 3) a positive real number s_ε, such that for every strategy τ of player 2, strategy σ of player 1, $0<\delta<\delta_0$, and $s > s_\varepsilon$ we have

$$E^z_{\sigma_\delta,\tau} g_\delta(s) \geq v(z) - \varepsilon, \tag{21}$$

and

$$E^z_{\sigma,\tau_\delta} g_\delta(s) \leq v(z) + \varepsilon. \tag{22}$$

By duality, it suffices to prove (21).

Let $A = \max\{|g(z,a)| : (z,a) \in \mathcal{A}\}$, and $g_\delta = \delta g$.

The first step is to show that for an exact family Γ_δ the following property holds. There is an integrable function $\psi : [0,1] \to \mathbb{R}_+$ and $\delta_0 > 0$ sufficiently small such that for $0 < \rho < \rho' \leq 1$ and $0 < \delta < \delta_0$, we have

$$\|v_{\delta,\rho} - v_{\delta,\rho'}\| \leq \int_\rho^{\rho'} \psi(x)\,dx. \tag{23}$$

The second step is to show that if the family Γ_δ of two-person zero-sum games satisfies the above-mentioned property, then it has an asymptotic uniform value.

We start with the first step. Fix the payoff function g and the transition rates μ. By the covariance properties, $V_{\delta,\rho} = \bar{V}_{\rho\delta}(\delta g, \delta\mu) = V_{\rho\delta}(\delta g, (1-\rho\delta)\delta\mu) = \frac{\delta}{(1-\rho\delta)\delta} V_{\frac{\delta\rho}{(1-\delta\rho)\delta}}(g,\mu) = \frac{1}{1-\rho\delta} V_{\frac{\rho}{(1-\rho\delta)}}(g,\mu)$. Therefore,

$$v_{\delta,\rho} = v_{\frac{\rho}{(1-\rho\delta)}}(g,\mu).$$

The function $\rho \mapsto v_\rho := v_\rho(g,\mu)$ (is semialgebraic and thus) has a convergent expansion, $v_\rho(z) = \sum_{k=0}^\infty c_k(z)\rho^{k/K}$ (where K is a positive integer), in a right neighborhood of 0. Therefore there is $1/2 > \rho_0 > 0$ such that its derivative, $v'_\rho(z) := \frac{d}{d\rho} v_\rho(z)$, exists in the interval $(0, 2\rho_0]$, and its absolute value is bounded by a positive constant C_1 times $\rho^{1/K-1}$. Therefore, for $\delta < 1/4$, the derivative $\frac{d}{d\rho} v_{\frac{\rho}{(1-\rho\delta)}}$ of the function $(0,\rho_0] \ni \rho \mapsto v_{\frac{\rho}{(1-\rho\delta)}} := v_{\frac{\rho}{(1-\rho\delta)}}(g,\mu)$ equals $\frac{1}{(1-\rho\delta)^2} v'_{\frac{\rho}{(1-\rho\delta)}}$; thus, it is bounded (in the interval $(0,\rho_0]$) by a positive constant C_2 times $\rho^{1/K-1}$. (E.g., $C_2 = 2C_1$.) The function $\rho \mapsto v_{\delta,\rho}$ is $(2A/\rho_0)$-Lipschitz in ρ in the interval $(\rho_0, 1]$ ($\|v_{\delta,\rho} - v_{\delta,\theta}\| \leq 2A|\rho - \theta|/\rho$, e.g., by [5, Lemma 4.2]). The function ψ that is defined by $\psi(x) = 2C_1 x^{1/K-1}$ for $0 < x \leq \rho_0$ and $\psi(x) = 2A/\rho_0$ for $1 \geq x > \rho_0$ is integrable and satisfies (23).

We turn now to the second step. Let Γ_δ be a converging family, $\psi : [0,1] \to \mathbb{R}_+$ be an integrable function, and $\delta_0 > 0$, such that for $0 < \rho < \rho' \leq 1$ and $0 < \delta < \delta_0$, inequality (23) holds.

Fix $\varepsilon > 0$ and w.l.o.g. we assume that $0 < \varepsilon < A$. Fix $\delta_0 > 0$ and $\lambda_0 > 0$ sufficiently small so that for $0 < \delta < \delta_0$ and $0 < \rho < \lambda_0$, $\|v_{\delta,\rho} - v\| < \varepsilon$.

Fix $0 < \delta < \delta_0$. We apply the proof of the existence of a value of the discrete-time stochastic game $\langle \delta g, \delta\mu\rangle$, [5, Section 2]. In what follows we define a strategy σ_δ of player 1 in Γ_δ. We will define a sequence $(\rho_k)_{m=0}^\infty$ so that ρ_k is a function of the past history up to stage $k[1/\delta]$, i.e., measurable with respect to the σ-algebra $\mathcal{F}_k := \mathcal{H}_{k[1/\delta]}$ where $[*]$ stands for the largest integer that is $\leq *$. The $(\rho_k)_{k=0}^\infty$ strategy σ_δ of player 1 is to play a stationary optimal strategy in Γ_{δ,ρ_k} at stage $k[1/\delta] \leq m \leq (k+1)[1/\delta]$. Let

$$\begin{aligned} y_k &= \sum_{k[1/\delta]\leq m<(k+1)[1/\delta]} (1-\delta\rho_k)^{m-k[1/\delta]}\delta g(z_m, a_m) \\ x_k &= \sum_{k[1/\delta]\leq m<(k+1)[1/\delta]} \delta g(z_m, a_m), \text{ and} \\ \bar{z}_k &= z_{k[1/\delta]}. \end{aligned}$$

For every strategy τ of player 2, we have

$$E_{\sigma_\delta,\tau}(\rho_k y_k + (1-\delta\rho_k)^{[1/\delta]} v_{\delta,\rho_k}(\bar{z}_{k+1}) \mid \mathcal{F}_k) \geq v_{\delta,\rho_k}(\bar{z}_k).$$

Note that for every $\varepsilon > 0$ there is $\lambda_0 > 0$ and δ_0 such that for $0 < \rho_k < \lambda_0$ and $0 < \delta < \delta_0$ we have

$$\sum_{k[1/\delta]\leq m<(k+1)[1/\delta]} |(1-\delta\rho_k)^{m-k[1/\delta]} - 1|\delta\rho_k + |(1-\delta\rho_k)^{[1/\delta]} - (1-\rho_k)| \leq \varepsilon\rho_k/A.$$

It follows that for $0 < \delta < \delta_0$ and $0 < \rho_k < \lambda_0$ we have

$$E_{\sigma_\delta,\tau}(v_{\delta,\rho_k}(\bar{z}_{k+1}) - v_{\delta,\rho_k}(\bar{z}_k) + \rho_k(x_k - v_{\delta,\rho_k}(\bar{z}_{k+1})) \mid \mathcal{F}_k) \geq -\varepsilon\rho_k \qquad (24)$$

for every strategy τ of player 2. Now one follows the proof of [5, Section 2] by replacing inequality [5, (2.1)] with inequality (24). The index i in [5, Section 2] is replaced by our stage index k (λ_i by ρ_k, v_λ by $v_{\delta,\rho}$, and z_i by $\bar{z}_k$).

With these substitutions, inequality [5, (2.15)] becomes

$$\sum_{k<n} x_k \geq \sum_{k<n} v_{\delta,\rho_k}(\bar{z}_{k+1}) + s_n - s_0 - 2A\sum_{k<n} I(s_{k+1} = M) - 4n\varepsilon. \qquad (25)$$

Note that the term $-\varepsilon\rho_k$ in inequality (24) does not appear in [5, (2.1)]. It impacts inequality [5, (2.9)] as $-\varepsilon\rho_k$ needs to be added to its right side. Therefore, we have to replace [5, (2.9)] with $E(Y_{k+1} - Y_k \mid \mathcal{F}_k) \geq \varepsilon\rho_k$ (where E stands for $E_{\sigma_\delta,\tau}$), and therefore $E\#\{k : \rho_k \geq \eta\} \leq \frac{A}{\varepsilon\eta}$ (rather than $\leq \frac{2A}{\varepsilon\eta}$

in [5, (2.12)]). Therefore, $E\sum_{k<n} I(s_{k+1}=M) \leq \frac{A}{\varepsilon\lambda(M)}$ and therefore for n sufficiently large $E\sum_{k<n} I(s_{k+1}=M) \leq \varepsilon n/(2A)$.

For δ and ρ sufficiently small, $\|v_{\delta,\rho} - v\| \leq \varepsilon$, where $v = \lim_{\delta\to 0} v_{\delta,0}$. Therefore, inequality (25) implies that

$$E_{\sigma_\delta,\tau} \sum_{k<n} x_k \geq nv(z_0) - 3\varepsilon n - s_0 - \varepsilon n - 4n\varepsilon. \tag{26}$$

□

Remark 11 *Note that the inequality* $E\sum_{k<n} I(s_{k+1}=M) \leq \frac{A}{\varepsilon\lambda(M)}$ *(in the above proof) implies that* $\sum_{k<\infty} I(s_{k+1}=M)$ *is finite a.s. Therefore,*

$$E_{\sigma_\delta,\tau}(\liminf_{n\to\infty} \frac{1}{n} \sum_{m<n} g(z_m,a_m)) = E_{\sigma_\delta,\tau}(\liminf_{n\to\infty} \frac{1}{n} \sum_{k<n} x_k) \geq v(z_0) - 7\varepsilon.$$

This shows that the above-constructed strategy σ_δ *of player 1 is approximate optimal in both the uniform game and the limiting-average game. Therefore an exact family of two-person zero-sum games* Γ_δ *has an asymptotic* 1_∞*-robust value.*

Theorem 7 *For every nonstationary discounting measure* w *on* $[0,\infty]$*, an exact family of two-person zero-sum games* Γ_δ *has an asymptotic* w*-robust value.*

Proof. If $w(\infty)=0$ then an asymptotic w value is a w-robust value. Therefore it suffices to prove the result for w with $w(\infty)>0$. For every $\beta>0$, the family $(\Gamma_\delta)_{\delta>0}$ has an asymptotic w-robust value if and only if it has an asymptotic βw-robust value. Therefore, we may assume that $w(\infty)=1$.

Let ν be the asymptotic 1_∞-robust value of the exact family $(\Gamma_\delta)_{\delta>0}$. Fix $\varepsilon>0$ and let τ_δ be a family of strategy profiles that are ε-optimal in the 1_∞-robust game. Let $t=t_\varepsilon<\infty$ be sufficiently large so that $w([t,\infty) < \varepsilon/\|g\|$. The family $(\Gamma_\delta)_{\delta>0}$ has an asymptotic (w_t,t,ν) value v_ε, where w_t is the restriction of w to the interval $[0,t)$. Let $m_\delta=[t/\delta]$ and let τ_δ be a profile of strategies that is optimal in $\Gamma^{m_\delta,\nu}_{\delta,w_{\delta,t}}$, where $w_{\delta,t}$ is the nonstationary discounting measure that satisfies $w_{\delta,t}(m)=w([m\delta,(m+1)\delta))$ if $m<m_\delta$ and $w_{\delta,t}(m)=0$ otherwise.

The strategy profile σ_δ follows the strategy profile $\tau_{\delta,t}$ in stages $0\leq m<m_\delta$ and in stage m_δ starts following the strategy profile τ_δ (explicitly, $\sigma_\delta(z_0,a_0,\ldots,z_{m_\delta+k}) = \tau_\delta(z_{m_\delta},\ldots,z_{m_\delta+k})$).

Then for every player i, all strategies $\bar{\tau}^i_\delta$ ($\delta > 0$) of player i, and all nonstationary discounting measures w_δ on $\mathbb{N}\cup\{\infty\}$ that converge (as $\delta \to 0+$) to w, we have

$$2\varepsilon + \liminf_{\delta\to 0+} E^z_{\sigma^1_\delta,\bar{\tau}^2_\delta}\underline{g}_\delta(w_\delta) \geq v_\varepsilon(z) \geq -2\varepsilon + \liminf_{\delta\to 0+} E^z_{\bar{\tau}^1_\delta,\sigma^2_\delta,}\bar{g}_\delta(w_\delta).$$

A limit point (as $\varepsilon \to 0+$) of v_ε is an asymptotic w-robust value of the family $(\Gamma_\delta)_{\delta>0}$. □

5 Non-zero-sum stochastic games with short-stage duration: the discounted games

5.1 The asymptotic discounted equilibrium

Fix the sets of players N, states S, and actions A, and let $\Gamma_\delta = \langle N, S, A, g_\delta, p_\delta\rangle$ be a stochastic game whose stage payoff function g_δ and transition function p_δ depend on the parameter $\delta > 0$ that represents the single-stage duration. Let $\Gamma_{\delta,\rho}$ be the (unnormalized) discounted game Γ_δ with discount factor $1 - \rho\delta$. We say that pair (V, σ), where $V \in \mathbb{R}^{N\times S}$ is a payoff vector and σ is a strategy profile, is an *asymptotic ρ-discounted ε-equilibrium* of $(\Gamma_\delta)_{\delta>0}$ if for every $\delta > 0$ sufficiently small, every player $i \in N$, every strategy τ^i of player i in Γ_δ, and every state z,

$$-\varepsilon+E^z_{\delta,\sigma^{-i},\tau^i}\sum_{m=0}^{\infty}(1-\delta\rho)^m g^i_\delta(z_m, a_m) \leq V^i(z) \leq E^z_{\delta,\sigma}\sum_{m=0}^{\infty}(1-\delta\rho)^m g^i_\delta(z_m, a_m)+\varepsilon.$$

The pair (V, σ) is an *asymptotic ρ-discounted equilibrium* if it is an asymptotic ρ-discounted ε-equilibrium for every $\varepsilon > 0$. It is called an *asymptotic ρ-discounted stationary ε-equilibrium*, respectively an *asymptotic ρ-discounted stationary equilibrium*, if, in addition, σ is stationary.

Theorem 8 *Every converging family $(\Gamma_\delta)_{\delta>0}$ has an asymptotic ρ-discounted stationary equilibrium.*

Proof. Let σ be a stationary strategy and $V_\rho \in \mathbb{R}^{N\times S}$ such that for every $z \in S$, $i \in N$, and $a^i \in A^i(z)$, we have

$$\rho V(z) = g(z, \sigma(z)) + \sum_{z'\in S} \mu(z', z, \sigma(z))V(z'),$$

and

$$\rho V^i(z) \geq g^i(z, \sigma(z)^{-i}, a^i) + \sum_{z' \in S} \mu(z', z, \sigma(z)^{-i}, a^i) V(z').$$

The existence of such a pair (V, σ) follows (as in the proof of Theorem 1) from the existence of stationary equilibria in discounted discrete-time stochastic games; alternatively, see, e.g., [8].

Let τ^i be a strategy of player i. Fix an initial history $h_m = (z_0, a_0, \ldots, z_m)$, and let $y_m = \sigma(z_m)$, $x^i_m = \tau^i(h_m)$, and $x_m = \sigma^{-i}(z_m) \otimes x^i_m$. Let

$$\begin{aligned} Y_m &:= E_{\delta,\sigma}\left(g^i_\delta(z_m, a_m) + (1-\rho\delta)V^i_\rho(z_{m+1}) \mid h_m)\right) \\ &= g^i_\delta(z_m, y_m) + (1-\rho\delta)\sum_{z' \in S} P_\delta(z' \mid z_m, y_m) V^i_\rho(z'), \end{aligned}$$

and

$$\begin{aligned} U_m &:= E_{\delta,\sigma^{-i},\tau^i}\left(g^i_\delta(z_m, a_m) + (1-\rho\delta)V^i_\rho(z_{m+1}) \mid h_m)\right) \\ &= g^i_\delta(z_m, x_m) + (1-\rho\delta)\sum_{z' \in S} P_\delta(z' \mid z_m, x_m) V^i_\rho(z'). \end{aligned}$$

It follows that

$$\begin{aligned} Y_m &\geq \delta g^i(z_m, y_m) + \sum_{z' \in S} \delta\mu(z', z, y_m) V^i(z') - \rho\delta V^i(z_m) + V^i(z_m) - o(\delta) \\ &\geq V^i(z_m) - o(\delta). \end{aligned}$$

Therefore,

$$E^z_{\delta,\sigma}\sum_{m=0}^{\infty}(1-\rho\delta)^m g^i_\delta(z_m, a_m) \geq V^i(z_0) - o(\delta)\sum_{m=0}^{\infty}(1-\rho\delta)^m \to_{\delta\to 0+} V^i(z_0).$$

Similarly,

$$\begin{aligned} U_m &\leq \delta g^i(z_m, a_m) + \sum_{z' \in S} \delta\mu(z', z, x_m) V^i(z') - \rho\delta V^i(z_m) + V^i(z_m) + o(\delta) \\ &\leq V^i(z_m) + o(\delta). \end{aligned}$$

Therefore,

$$E^z_{\delta,\sigma^{-i},\tau^i}\sum_{m=0}^{\infty}(1-\rho\delta)^m g^i_\delta(z_m, a_m) \leq V^i(z) + o(\delta)\sum_{m=0}^{\infty}(1-\rho\delta)^m \to_{\delta\to 0+} V^i(z).$$

We conclude that for sufficiently small $\delta > 0$ we have

$$-\varepsilon + E^z_{\delta,\sigma^{-i},\tau^i} \sum_{m=0}^{\infty} (1-\rho\delta)^m g^i_\delta(z_m, a_m) \leq V^i(z) \leq E^z_{\delta,\sigma} \sum_{m=0}^{\infty} (1-\rho\delta)^m g^i_\delta(z_m, a_m) + \varepsilon.$$

□

Remark 12 *The conclusion of Theorem 8 (as well as its proof) holds also for the model with individual discount rates* $\overrightarrow{\rho} = (\rho_i)_{i\in N}$.

Covariance properties. Fix $\alpha, \beta > 0$. A point $(x, V) \in \times_{z\in S, i\in N}(X^i(z) \times [-\|g^i\|/\rho, \|g^i\|/\rho])$ is a stationary equilibrium (strategies and payoffs) of the continuous-time ρ-discounted game $\Gamma = \langle N, S, A, g, \mu\rangle$ if and only if (x, V) is a stationary equilibrium of the continuous-time $\alpha\rho$-discounted game $\Gamma = \langle N, S, A, \alpha g, \alpha\mu\rangle$, and given $0 < \rho < 1$ and $\|\mu\| \leq 1-\rho$, if and only if it is a stationary equilibrium of the discrete-time ρ-discounted game $\bar{\Gamma} = \langle N, S, A, g, \bar{p}\rangle$, where $\bar{p}$ is the transition probability that is given by $\bar{p}(z', z, a) = \frac{1}{1-\rho}\mu(z', z, a)$ for all $z' \neq z$.

5.2 The asymptotic discounted minmax

Fix the sets of players N, states S, and actions A, and let $\Gamma_\delta = \langle N, S, A, g_\delta, p_\delta\rangle$ be a stochastic game whose stage payoff function g_δ and transition function p_δ depend on the parameter $\delta > 0$ that represents the single-stage duration. The (unnormalized) ρ-discounted minmax of the discrete-time game Γ_δ is defined as the (uncorrelated) minmax of the discrete-time stochastic game Γ_δ with discount factor $1-\delta\rho$. It exists and is denoted by $V_{\delta,\rho}$. We say that $V_\rho \in \mathbb{R}^{N\times S}$ is the (unnormalized) *asymptotic ρ-discounted minmax* of the family $(\Gamma_\delta)_{\delta>0}$ if $V_{\delta,\rho} \to V_\rho$ as $\delta \to 0+$.

Using arguments analogous to those used in earlier sections, it follows that 1) $V_{\delta,\rho} = (V^i_{\delta,\rho}(z))_{(i,z)\in N\times S}$ is the unique solution of the following system of $|N \times S|$ equalities,

$$\delta\rho V^i(z) = \min_{x^{-i}\in X^{-i}(z)} \max_{x^i\in X^i(z)} g^i_\delta(z, x^{-i}\otimes x^i) + (1-\delta\rho) \sum_{z'\in S} p_\delta(z', z, x^{-i}\otimes x^i) V^i(z'),$$

where $X^{-i}(z) := \times_{j\neq i} X^i(z)$, 2) a family $(\Gamma_\delta)_{\delta>0}$ ($\Gamma_\delta = \langle g_\delta, p_\delta\rangle$ for short) that converges to $\langle g, \mu\rangle$ has an asymptotic ρ-discounted minmax V_ρ, and 3)

$V_\rho = (V^i_\rho(z))_{(i,z)\in N\times S}$ is the unique solution of the system of $|N \times S|$ equalities,

$$\rho V^i(z) = \min_{x^{-i}\in X^{-i}} \max_{x^i \in X^i(z)} g(z, x^{-i}\otimes x^i) + \sum_{z'\in S} \mu(z', z, x^{-i}\otimes x^i) V^i(z').$$

The normalized ρ-discounted minmax values are $v_\rho = \rho V_\rho$ and $v_{\delta,\rho} = \rho V_{\delta,\rho}$. The semialgebraic and covariance properties of the value of zero-sum games hold for the minmax value of non-zero-sum games as well.

In particular, for fixed g_δ, p_δ, g, and μ, the maps $\rho \mapsto v_\rho$ and $\rho \mapsto v_{\delta,\rho}$ are bounded semialgebraic functions, and thus have a limit as $\rho \to 0+$, the maps $\rho \mapsto V_\rho$ and $\rho \mapsto V_{\delta,\rho}$ are semialgebraic, $v_{\delta,\rho}(g_\delta, p_\delta) = v_\rho(g_\delta/\delta, (1-\rho\delta)p_\delta/\delta)$, inequality (16) holds, and if $\Gamma_\delta = \langle g_\delta, p_\delta\rangle$ converges strongly to $\Gamma = \langle g, \mu\rangle$, then $v_{\delta,\rho}$ converges, as $\delta \to 0+$, uniformly in ρ.

□

5.3 The asymptotic equilibrium of nonstationary discounting games

The following theorem is a generalization of Theorem 3 to the non-zero-sum case. Its proof is analogous to the proof of Theorem 3.

Theorem 9 *If 1) $(\Gamma_\delta)_{\delta>0}$ is a family that converges in data, 2) $\overrightarrow{w}$ is a nonstationary discounting N-vector measure on $[0,\infty)$, 3) $t \in \mathbb{R}$, and 4) $\nu : \mathcal{A} \to \mathbb{R}^N$, then the family $(\Gamma_\delta)_{\delta>0}$ has an asymptotic $(\overrightarrow{w}, t, \nu)$ equilibrium payoff v. If 1) $\overrightarrow{w}_\delta$ is a nonstationary discounting N-vector measure on $\mathbb{N}$ that converges to $\overrightarrow{w}$, and 2) $0 \leq m_\delta \in \mathbb{N}$ and $\nu_\delta : \mathcal{A} \to \mathbb{R}^N$ are such that $(m_\delta, \nu_\delta) \to_{\delta\to 0+} (t,\nu)$, then, for every $\varepsilon > 0$, there are Markov strategy profiles σ_δ and $\delta_0 > 0$ such that 1) for every $0 < \delta < \delta_0$, σ_δ is an ε-equilibrium of $\Gamma^{m_\delta,\nu_\delta}_{\delta,\overrightarrow{w}_\delta}$ with an equilibrium payoff within ε of v, and 2) σ_δ converge w^* to a profile of continuous-time Markov strategies.*

6 Non-zero-sum stochastic games with short-stage duration: the limiting-average and uniform games

Fix the sets of players N, states S, and actions A, and let $\Gamma_\delta = \langle N, S, A, p_\delta, g_\delta\rangle$ be a stochastic game whose stage payoff function g_δ and transition function p_δ depend on the parameter $\delta > 0$ that represents the single-stage duration.

For every strategy profile σ in Γ_δ we set

$$\bar{\gamma}^i_\delta(z,\sigma) = E^z_{\delta,\sigma}\,\bar{g}^i_\delta, \quad \text{and} \quad \underline{\gamma}^i_\delta(z,\sigma) = E^z_{\delta,\sigma}\,\underline{g}^i_\delta.$$

6.1 The asymptotic limiting-average and uniform minmax

We say that the vector $v \in \mathbb{R}^{N\times S}$ is the *asymptotic limiting-average minmax* of the family $(\Gamma_\delta)_{\delta>0}$ if for every $\varepsilon > 0$ there is $\delta_0 > 0$ such that for every player i and $0 < \delta < \delta_0$, 1) there is a strategy profile $\sigma^{-i}_{\delta,\varepsilon}$ of players $N \setminus \{i\}$ such that for every strategy τ^i of player i and every state $z \in S$,

$$\bar{\gamma}^i_\delta(z, \sigma^{-i}_{\delta,\varepsilon}, \tau^i) \leq v^i(z) + \varepsilon,$$

and 2) for every strategy profile σ^{-i}_δ of players $N \setminus \{i\}$ there is a strategy τ^i of player i such that for every state $z \in S$,

$$\underline{\gamma}^i_\delta(z, \sigma^{-i}_\delta, \tau^i) \geq v^i(z) - \varepsilon.$$

We say that the vector $v \in \mathbb{R}^{N\times S}$ is the *asymptotic uniform minmax* of the family $(\Gamma_\delta)_{\delta>0}$ if for every $\varepsilon > 0$ there are $\delta_0 > 0$ and $s_0 > 0$ such that for every player i and $0 < \delta < \delta_0$, 1) there is a strategy profile $\sigma^{-i}_{\delta,\varepsilon}$ of players $N \setminus \{i\}$ such that for every strategy τ^i of player i, state $z \in S$, and duration $s > s_0$,

$$E^z_{\sigma^{-i}_{\delta,\varepsilon},\tau^i} g^i_\delta(s) \leq v^i(z) + \varepsilon,$$

and 2) for every strategy profile σ^{-i}_δ of players $N \setminus \{i\}$ there is a strategy τ^i of player i such that for $s > s_0$,

$$E_{\sigma^{-i}_\delta,\tau^i} g^i_\delta(s) \geq v^i(z) - \varepsilon.$$

We say that the vector $v \in \mathbb{R}^{N\times S}$ is the *asymptotic robust minmax* of the family $(\Gamma_\delta)_{\delta>0}$ if for every $\varepsilon > 0$ there are $\delta_0 > 0$ and $s_0 > 0$ such that for every player i and $0 < \delta < \delta_0$, 1) there is a strategy profile $\sigma^{-i}_{\delta,\varepsilon}$ of players $N \setminus \{i\}$ such that for every strategy τ^i of player i, state $z \in S$, and duration $s > s_0$,

$$E^z_{\sigma^{-i}_{\delta,\varepsilon},\tau^i} g^i_\delta(s) \leq v^i(z) + \varepsilon \quad \text{and} \quad \bar{\gamma}^i_\delta(z, \sigma^{-i}_{\delta,\varepsilon}, \tau^i) \leq v^i(z) + \varepsilon,$$

and 2) for every strategy profile σ_δ^{-i} of players $N \setminus \{i\}$ there is a strategy τ^i of player i, such that for every state $z \in S$ and duration $s > s_0$,

$$E^z_{\sigma_\delta^{-i},\tau^i} g^i_\delta(s) \geq v^i(z) - \varepsilon \quad \text{and} \quad \underline{\gamma}^i_\delta(z, \sigma_\delta^{-i}, \tau^i) \geq v^i(z) - \varepsilon.$$

Theorem 10 *A family $(\Gamma_\delta)_{\delta>0}$ that converges strongly to $\Gamma = \langle \mu, g\rangle$ has a limiting-average minmax $v : S \to \mathbb{R}^N$, which is the limit of ρV_ρ as $\rho \to 0+$, where V_ρ is the unique solution of the following system of equalities:*

$$\rho V^i(z) = \min_{x^{-i}} \max_{y^i} \left(g^i(z, x^{-i}, y^i) + \sum_{z' \in S} \mu(z', z, x^{-i}, y^i) V^i(z) \right), \quad \forall i \in N, z \in S.$$

If the family is exact it has an asymptotic robust minmax (and therefore an asymptotic uniform minmax as well).

Proof. The proof that a family that converges strongly has an asymptotic limiting-average minmax is analogous to the proof of Theorem 4. Let $\tilde{v}_\delta = \lim_{\rho\to 0+} v_{\delta,\rho}$.

As every discrete-time stochastic game with finitely many states and actions has a limiting-average minmax [5, 7], which is the limit of its ρ-discounted minmax as ρ goes to $0+$, it suffices to prove that $\lim_{\delta\to 0+} \tilde{v}_\delta$ exists.

As mentioned in the last section, if $\langle g_\delta, p_\delta\rangle$ converges strongly, then $v_{\delta,\rho}$ converges to v_ρ, as $\delta \to 0$, uniformly in ρ. Therefore, for every $\varepsilon > 0$ there is $\delta_1 > 0$ such that for $0 < \delta, \delta' \leq \delta_1$ we have $\|v_{\delta,\rho} - v_{\delta',\rho}\| < \varepsilon$, and therefore $\|\tilde{v}_\delta - \tilde{v}_{\delta'}\| \leq \varepsilon$.

The proof that an exact family has an asymptotic minmax is analogous to the proof of Theorem 6.

□

6.2 The asymptotic limiting-average equilibrium

We say that $u = (u^i(z))_{i\in N, z\in S} \in \mathbb{R}^{N\times S}$ is an *asymptotic limiting-average ε-equilibrium payoff* of $(\Gamma_\delta)_{\delta>0}$ if for every $\delta > 0$ sufficiently small there is a strategy profile σ_δ, such that for every player $i \in N$, strategy τ^i of player i, and state z,

$$-\varepsilon + \bar{\gamma}^i_\delta(z, \sigma_\delta^{-i}, \tau^i) \leq u^i(z) \leq \underline{\gamma}^i_\delta(z, \sigma_\delta) + \varepsilon.$$

Note that it is an asymptotic limiting-average equilibrium payoff if it is an *asymptotic limiting-average ε-equilibrium payoff* for every $\varepsilon > 0$.

Remark 13 *Note that the existence of a limiting-average equilibrium, respectively ε-equilibrium, payoff in each one of the games Γ_δ does not imply (and is not implied by) the existence of an asymptotic limiting-average equilibrium, respectively ε-equilibrium, payoff of the family $(\Gamma_\delta)_{\delta>0}$.*

Remark 14 *If $u_\varepsilon \in \mathbb{R}^{N\times S}$ is an asymptotic limiting-average ε-equilibrium payoff of the family $(\Gamma_\delta)_{\delta>0}$ and $u \in \mathbb{R}^{N\times S}$, then u is an asymptotic limiting-average ε'-equilibrium payoff of the family $(\Gamma_\delta)_\delta$ whenever $\varepsilon' \geq \varepsilon + \|u - u_\varepsilon\|$. Therefore a limit point, as $\delta \to 0+$, of asymptotic limiting-average ε-equilibrium payoffs is an asymptotic limiting-average ε'-equilibrium payoff whenever $\varepsilon' > \varepsilon$, and a limit point, as $\varepsilon \to 0+$, of asymptotic limiting-average ε-equilibrium payoffs is an asymptotic limiting-average equilibrium payoff.*

Two related equilibrium concepts are the lim sup and the lim inf equilibrium payoffs. We say that $u = (u^i(z))_{i\in N,\, z\in S} \in \mathbb{R}^{N\times S}$ is an *asymptotic* lim sup *ε-equilibrium payoff*, respectively an *asymptotic* lim inf *ε-equilibrium payoff*, of $(\Gamma_\delta)_{\delta>0}$ if for every $\delta > 0$ sufficiently small there is a strategy profile σ_δ, such that for every player $i \in N$, strategy τ^i of player i in Γ_δ, and state z,

$$-\varepsilon + \bar{\gamma}^i_\delta(z, \sigma^{-i}_\delta, \tau^i) \leq u^i(z) \leq \bar{\gamma}^i_\delta(z, \sigma_\delta) + \varepsilon,$$

respectively

$$-\varepsilon + \underline{\gamma}^i_\delta(z, \sigma^{-i}_\delta, \tau^i) \leq u^i(z) \leq \underline{\gamma}^i_\delta(z, \sigma_\delta) + \varepsilon.$$

The corresponding strategies σ_δ are 2ε-equilibrium strategies of Γ_δ with the lim sup, respectively lim inf, payoff function.

We say that $u = (u^i(z))_{i\in N,\, z\in S} \in \mathbb{R}^{N\times S}$ is an *asymptotic* lim sup *equilibrium payoff*, respectively an *asymptotic* lim inf *equilibrium payoff*, if it is an asymptotic lim sup ε-equilibrium payoff, respectively an asymptotic lim inf ε-equilibrium payoff, for every $\varepsilon > 0$.

Remark 15 *Obviously, an asymptotic limiting-average equilibrium payoff is an asymptotic* lim sup *and an asymptotic* lim inf *equilibrium payoff. However, there are stochastic games with countably many states that have both an asymptotic* lim sup *equilibrium payoff and an asymptotic* lim inf *equilibrium payoff, such that, moreover, both payoffs coincide, but have no asymptotic limiting-average equilibrium payoffs.*

Remark 16 *It is unknown whether every stochastic game with finitely many states and actions has a* lim sup*, respectively* lim inf*, equilibrium payoff. In particular, it is unknown whether every stochastic game with finitely many states and actions has a limiting-average equilibrium payoff.*

Theorem 11 *A family $(\Gamma_\delta = \langle g_\delta, p_\delta\rangle)_{\delta>0}$ that converges strongly to $\Gamma = \langle g, \mu\rangle$ has a limiting-average equilibrium payoff.*

Proof. Let $(\Gamma_\delta)_{\delta>0}$ be a family that converges strongly to $\Gamma = \langle \mu, g\rangle$. Then $|g^i_\delta(z,a) - \delta g^i(z,a)| = o(\delta)$, and therefore $|\frac{1}{n\delta}\sum_{0\le m<n} g^i_\delta(z_m,a_m) - \frac{1}{n\delta}\sum_{0\le m<n}\delta g^i(z_m,a_m)| \le \max_{z,a}|g^i_\delta(z,a) - \delta g^i(z,a)|/\delta = o(1)$ as $\delta \to 0+$. Therefore, it suffices to prove the theorem for the special case where $g^i_\delta = \delta g^i$. Note that in this special case

$$\frac{1}{n\delta}\sum_{0\le m<n} g^i_\delta(z_m,a_m) = \frac{1}{n\delta}\sum_{0\le m<n}\delta g^i(z_m,a_m) = \frac{1}{n}\sum_{0\le m<n} g^i(z_m,a_m).$$

Therefore, $\bar{g}^i_\delta$ and $\underline{g}^i_\delta$, as a function of the play $z_0, a_0, \ldots$, are independent of δ. Therefore we write $\bar{g}^i$ and $\underline{g}^i$ for short for $\bar{g}^i_\delta$ and $\underline{g}^i_\delta$. Without loss of generality we may assume that $0 \le g^i \le 1$.

By Remark 14, it suffices to prove that for every $\varepsilon > 0$ there is a vector $u \in \mathbb{R}^{N\times S}$ that is an asymptotic limiting-average ε-equilibrium payoff.

Fix $\varepsilon > 0$ and let u and σ be, respectively, the uniform (and limiting-average) $\varepsilon/8$-equilibrium payoff and the uniform (and limiting-average) $\varepsilon/8$-equilibrium strategy of the continuous-time stochastic game $\Gamma = \langle N, S, A, \mu, g\rangle$ that are constructed in [8]. In particular, for every state $z \in S$, player $i \in N$, and strategy τ^i of player i, we have

$$u^i(z) + \varepsilon/8 \ge E^z_\sigma \bar{g}^i \ge E^z_\sigma \underline{g}^i \ge u^i(z) - \varepsilon/8, \tag{27}$$

where $\bar{g}^i = \limsup_{s\to\infty}\frac{1}{s}\int_0^s g^i(z_t,x_t)\,dt$ and $\underline{g}^i = \liminf_{s\to\infty}\frac{1}{s}\int_0^s g^i(z_t,x_t)\,dt$, and

$$E^z_{\sigma^{-i},\tau^i}\bar{g}^i \le u^i(z) + \varepsilon/8. \tag{28}$$

These inequalities follow from (u,σ) being a limiting-average $\varepsilon/8$-equilibrium payoff and strategy profile. (An additional property that follows from the special construction of σ in [8] is that $\bar{g}^i = \underline{g}^i$ P^z_σ a.e.)

Let $v : S \to \mathbb{R}^N$ be the limit of ρV_ρ as $\rho \to 0+$, where V_ρ is the asymptotic ρ-discounted minmax. Recall that V_ρ is the unique solution of the following system of equalities:

$$\rho V^i(z) = \min_{x^{-i}} \max_{y^i} \left(g^i(z, x^{-i}, y^i) + \sum_{z' \in S} \mu(z', z, x^{-i} \otimes y^i) V^i(z) \right), \quad \forall i \in N, z \in S.$$

As the strategy profile σ (that is constructed in [8]) is a discretimized strategy (namely, there is a strictly increasing sequence of continuous times $t_0 = 0 < t_1 < t_2 < \ldots$, such that $t_\ell \to_{\ell\to\infty} \infty$ and the mixed-action profile selected by σ at time $t_\ell \le t < t_{\ell+1}$ is a function of the play up to time t_ℓ and the state at time t), it follows that for every $\varepsilon' > 0$ and for every player i there is a strategy $\tau^i_{\varepsilon'}$ such that $v^i(z) - \varepsilon' < E^z_{\sigma^{-i},\tau^i_{\varepsilon'}} \underline{g}^i$ ($\le E^z_{\sigma^{-i},\tau^i_{\varepsilon'}} \bar{g}^i$). Therefore, the inequalities $u^i(z) + \varepsilon/8 \ge E^z_{\sigma^{-i},\tau^i_{\varepsilon'}} \bar{g}^i \ge v^i(z) - \varepsilon'$ hold for every $\varepsilon' > 0$. Therefore,

$$u^i(z) \ge v^i(z) - \varepsilon/8. \tag{29}$$

We will prove that u is an asymptotic limiting-average ε-equilibrium payoff of the family $(\Gamma_\delta)_{\delta>0}$. The construction of the corresponding limiting-average ε-equilibrium strategy profile σ_δ is analogous to the construction of σ in [8]. The continuous-time pure-action strategy profiles $\bar{\tau}$ and $\hat{\tau}$, which are used in [8] in the definition of σ, will be adapted to the discrete-time pure strategies $\bar{\tau}_\delta$ and τ_δ respectively.

The continuous-time pure-action strategy profile $\bar{\tau}$ obeys the following property. There is a sequence of continuous times $\mathcal{T} : 0 = t_0 < t_1 < \ldots$ (with $t_k \to_{k\to\infty} \infty$) such that for $t_k \le t < t_{k+1}$, $\bar{\tau}(h,t)$ is a function of t, z_t, and the finite sequence of states $\overrightarrow{z}_k = (z_{t_0}, \ldots, z_{t_k})$. Therefore, for $t_k \le t < t_{k+1}$, we can write $\bar{\tau}(\overrightarrow{z}_k, z_t, t)$ for $\bar{\tau}(h,t)$.

The corresponding discrete-time pure strategy $\bar{\tau}_\delta$ will be such that there is a sequence of stages $\mathcal{T}^\delta : 0 = n_{\delta,0} < \ldots < n_{\delta,k} < \ldots$ such that 1) $\delta n_{\delta,k} \to_{\delta\to0+} t_k$, 2) for $n_{\delta,k} \le m < n_{\delta,k+1}$, $\bar{\tau}_\delta(z_0, a_0, \ldots, z_m)$ is a function of m, z_m, and the finite sequence of states $\overrightarrow{z}^\delta_k = (z_{n_{\delta,0}}, \ldots, z_{n_{\delta,k}})$; thus we can write $\bar{\tau}_\delta(\overrightarrow{z}^\delta_k, z_m, m)$ for $\bar{\tau}(z_0, a_0, \ldots, z_m)$, and 3) for fixed $\overrightarrow{z}_k = \overrightarrow{z}^\delta_k \in S^{k+1}$, the map $[n_{\delta,k}, n_{\delta,k+1}) \ni m \mapsto \bar{\tau}_\delta(\overrightarrow{z}^\delta_k, z_m, m)$, which (given $\overrightarrow{z}^\delta_k$) is a Markov strategy on this interval of stages, converges w^* to the (given $\overrightarrow{z}_k$) Markov strategy $[t_k, t_{k+1}) \ni t \mapsto \bar{\tau}(\overrightarrow{z}_k, z_t, t)$.

This relation between the continuous-time strategy $\bar{\tau}$ and the discrete-time strategy $\bar{\tau}_\delta$ implies, by inductive application of Proposition 2, that for every state z and every positive M,

$$\gamma^i_{\delta,M}(z,\bar\tau_\delta) := E^z_{\bar\tau_\delta}\frac{1}{[M/\delta]}\sum_{0\le m<[M/\delta]} g^i(z_m,a_m) \to_{\delta\to 0+} \gamma^i_M(z,\bar\tau),$$

where

$$\gamma^i_M(z,\bar\tau) := E^z_{\bar\tau}\frac{1}{M}\int_0^M g^i(z_t,x_t)\,dt,$$

and $[*]$ stands for the largest integer that is less than or equal to $*$.

Definition of $\bar\tau_\delta$. We map the discrete-time sequence of states $z_0, z_1, \ldots$ to a continuous-time (step-function) list of states: for any nonnegative real $t \ge 0$ we define $z_{\delta,t} = z_{[t/\delta]}$. Next, we define the profile of strategies $\bar\tau_\delta$ in Γ_δ by $\bar\tau_\delta(z_0, a_m, \ldots, z_m) = \bar\tau((z_{\delta,t})_{t\le m\delta})$.

Properties of $\bar\tau_\delta$. Recall the definition and properties of the positive integer N_0, the sufficiently small $\varepsilon_1 > 0$, the disjoint subset of states, S_1, S_2, and $\bar S$, and the pure-action strategy profile $\bar\tau$ (that were constructed in [8]). One of the properties of $\bar\tau$ is that for every $z \in S_1$ and $z_s \in C_z := \{z' \in S \mid v(z') = v(z)\}$ for all $s \le t$, $\mu(\bar S \cup (S \setminus C_z), z_t, \bar\tau_t) = 0$. Therefore, by the definition of $\bar\tau_\delta$ we have

$$P^z_{\bar\tau_\delta}(z_m \in C_z \setminus \bar S) = 1 \quad \forall z \in S_1,\ m \ge 0. \tag{30}$$

The following inequality[6] is proved in [8]. For $z \in S_1$, for every player i,

$$\gamma^i_{N_0}(z,\bar\tau) \ge v^i(z) - \varepsilon/7,$$

and therefore, for sufficiently small $\delta > 0$,

$$\gamma^i_{\delta,N_0}(z,\bar\tau_\delta) \ge v^i(z) - \varepsilon/6. \tag{31}$$

Definition of τ_δ. We define a stopping time $m_\delta = m_\delta(z_0, a_0, z_1, \ldots)$ as follows. On $z_0 \in S_1$, $m_\delta = [N_0/\delta]$; on $z_0 \in \bar S$, $m_\delta = [1/\delta]$; on $z_0 = z \in S_2$, $m_\delta = \min(\{m : m = [j/\delta], j \in \mathbb{N}, \text{ and } z_m \notin C_z \setminus \bar S\} \cup \{[N_0/\delta]\})$. Define $m_{k,\delta}$, $k \ge 0$ inductively: $m_{0,\delta} = 0$ and $m_{k+1,\delta} = m_{k,\delta} + m_\delta(z_{m_{k,\delta}}, a_{m_{k,\delta}}, z_{m_{k,\delta}+1}, \ldots)$.

The strategy profile τ_δ is defined as follows.

$$\tau_\delta(z_0, a_0, \ldots, z_m) = \bar\tau_\delta(z_{m_{k,\delta}}, z_{m_{k,\delta}}, \ldots, z_m) \text{ if } m_{k,\delta} \le m < m_{k+1,\delta}.$$

[6]The ε in [8] is $\varepsilon/8$ here, and ε_1 there is sufficiently small.

Properties of τ_δ. We define the sequence of states $\bar{z}_k^\delta$, $k \geq 0$, by $\bar{z}_k^\delta = z_{m_{k,\delta}}$. Note that this definition is analogous to that of the sequence of states $\bar{z}_k$, $k \geq 0$, in [8]. Let F^δ, respectively F, be the transition matrix of the homogeneous Markov chain $\bar{z}_0^\delta, \bar{z}_1^\delta, \ldots$ with its $P_{\tau_\delta}^z$ distribution, respectively $\bar{z}_0, \bar{z}_1, \ldots$ with its $P_{\hat{\tau}}^z$ distribution.

By the strong data convergence of $\langle \delta g, p_\delta \rangle$ to $\langle g, \mu \rangle$, $p_\delta(z', z, a) > 0$ if and only if $\mu(z', z, a) > 0$. Therefore (for $\delta > 0$ sufficiently small) $F_{z,z'}^\delta = 0$ if and only if $F_{z,z'} = 0$, and thus the ergodic classes of states of the two homogeneous Markov chains, the one with transition matrix F^δ and the other with transition matrix F, coincide.

Let $\mathcal{E}$ denote the set of ergodic classes of states, and for $E \in \mathcal{E}$ we denote by q_δ^E and q^E the F^δ and F invariant measures that are supported on E, and $q_\delta^z(E)$ (respectively $q^z(E)$) denotes the probability of the F^δ-Markov chain (respectively F-Markov chain) with initial state z entering the ergodic class E. Recall that every ergodic class $E \in \mathcal{E}$ is a subset of S_1, and on $z_0 \in S_1$ we have $m_\delta = [N_0/\delta]$. Therefore,

$$E_{\tau_\delta}^z \underline{g}^i = E_{\tau_\delta}^z \bar{g}^i = \sum_{E \in \mathcal{E}} q_\delta^z(E) \sum_{z \in E} q_\delta^E(z) \gamma_{\delta, N_0}^i(z, \bar{\tau}_\delta).$$

Similarly,

$$E_{\hat{\tau}}^z \underline{g}^i = E_{\hat{\tau}}^z \bar{g}^i = \sum_{E \in \mathcal{E}} q^z(E) \sum_{z \in E} q^E(z) \gamma_{N_0}^i(z, \bar{\tau}).$$

In addition, by Proposition 2 and the w^* convergence of $\bar{\tau}_\delta$ to $\bar{\tau}$, $F^\delta \to F$ as $\delta \to 0+$. Therefore, $q_\delta^z(E) \to_{\delta \to 0+} q^z(E)$ and $q_\delta^E \to_{\delta \to 0+} q^E$. Since for all $z \in S$, $E \in \mathcal{E}$, and $z' \in E$, we have

$$(q_\delta^z(E), q_\delta^E(z'), \gamma_{\delta, N_0}^i(z', \bar{\tau}_\delta)) \to_{\delta \to 0+} (q^z(E), q^E(z'), \gamma_{N_0}^i(z', \bar{\tau})),$$

we deduce that

$$E_{\tau_\delta}^z \underline{g}^i = E_{\tau_\delta}^z \bar{g}^i \to_{\delta \to 0+} E_{\hat{\tau}}^z \bar{g}^i = E_{\hat{\tau}}^z \underline{g}^i.$$

Therefore, for sufficiently small $\delta > 0$ we have

$$u^i(z) - \varepsilon \leq E_{\tau_\delta}^z \underline{g}^i = E_{\tau_\delta}^z \bar{g}^i \leq u^i(z) + \varepsilon/6. \tag{32}$$

Recall the definition of τ, ε_1 and $\tilde{\tau}$ in [8], where it is proved that for every $z \in S$, every player i, and every stopping time T, $E_\tau^z v^i(z_T) \geq v^i(z) - \varepsilon_1/2$.

Therefore, for every $z \in S_2$, $\sum_{z' \in S} F_{z,z'} v^i(z') \geq v^i(z) - \varepsilon_1/2$. For $z \in S_1 \cup \bar{S}$, $\sum_{z' \in S} F_{z,z'} v^i(z') = v^i(z)$. Therefore,

$$\sum_{z' \in S} F_{z,z'} v^i(z') \geq v^i(z) - \frac{\varepsilon_1}{2} 1_{z \in S_2 \cup \bar{S}},$$

where 1_* is the indicator function of $*$. In addition, if we replace the symbols δ and ε in [8] with the symbol η and $\varepsilon/8$, it can be seen that for $\varepsilon_1 < \eta d^2 \frac{\varepsilon}{32}$,

$$E^z_\sigma \sum_{k=0}^{\infty} 1_{\bar{z}_k \notin S_1} \leq \frac{128}{\eta d^2 \varepsilon},$$

and therefore for $\varepsilon < 1$ and $\varepsilon_1 < \eta d^2 \frac{\varepsilon^2}{8} \frac{1}{128}$ $(< \eta d^2 \frac{\varepsilon}{32})$,

$$\varepsilon_1 E^z_\sigma \sum_{k=0}^{\infty} 1_{\bar{z}_k \notin S_1} < \eta d^2 \frac{\varepsilon^2}{8} \frac{1}{128} \frac{128}{\eta d^2 \varepsilon} = \varepsilon/8.$$

Therefore, for sufficiently small δ,

$$\varepsilon_1 E^z_{\tau_\delta} \sum_{k=0}^{\infty} 1_{\bar{z}_k \notin S_1} < \varepsilon/6.$$

Assume that $\varepsilon < 1$ and $\varepsilon_1 < \eta d^2 \frac{\varepsilon^2}{8} \frac{1}{128}$.

Lemma 4 *For sufficiently small $\delta > 0$, for every stopping time T we have*

$$E^z_{\tau_\delta} v^i_T \leq E^z_{\tau_\delta} v^i_\infty + \varepsilon/6, \tag{33}$$

where $v_m = v(z_m)$ and $v^i_\infty = \limsup_{m\to\infty} v^i_m$ (which equals $\lim_{m\to\infty} v^i_m$ $P^z_{\tau_\delta}$ a.e.), and

$$E^z_{\tau_\delta} 1_{T<\infty} v^i_T \leq E^z_{\tau_\delta} 1_{T<\infty} v^i_\infty + \varepsilon/6 \leq E^z_{\tau_\delta} 1_{T<\infty} \bar{g}^i + \varepsilon/3. \tag{34}$$

Proof. The strategy τ defined in [8] obeys $v^i(z_T) \leq E^z_\tau(v^i(z_{T'}) \mid \mathcal{H}_T) + \varepsilon_1/2$ for all finite stopping times $T \leq T'$. Therefore, for all stopping times $T \leq T' \leq N_0$, $E^z_\tau v^i(z_T) \leq E^z_\tau v^i(z_{T'}) + \varepsilon_1/2$, and $E^z_{\bar{\tau}} v^i(z_T) \leq E^z_{\bar{\tau}} v^i(z_{T'}) + 3\varepsilon_1/4$. For $z \in S_1$ we have, $v(z_m) = v(z)$ for all $m \leq m_\delta$, $P^z_{\bar{\tau}_\delta}$ a.e. Therefore, for sufficiently small $\delta > 0$, for every stopping time $T \leq m_\delta$ (in the discrete-time game), we have

$$E^z_{\tau_\delta} v^i_T = E^z_{\bar{\tau}_\delta} v^i_T \leq E^z_{\tau_\delta} v^i(\bar{z}^\delta_1) + \varepsilon_1 1_{z \notin S_1}.$$

Therefore, for $\delta > 0$ sufficiently small, for every stopping time T,

$$E^z_{\tau_\delta} v^i_T \le E^z_{\tau_\delta} v^i_\infty + \varepsilon_1 E^z_{\tau_\delta} \sum_{k=0}^{\infty} 1_{\bar{z}_k \notin S_1} \le E^z_{\tau_\delta} v^i_\infty + \varepsilon/6.$$

This completes the proof of inequality (33).

Since $1_{T=\infty} v_T = 1_{T=\infty} v_\infty$ $P^z_{\tau_\delta}$ a.e., we deduce that

$$E^z_{\tau_\delta} 1_{T<\infty} v_T = E^z_{\tau_\delta} 1_{T<\infty} v_\infty + \varepsilon/6. \tag{35}$$

By inequality (31) we have $v^i_\infty \le \bar{g}^i + \varepsilon/6$, $P^z_{\tau_\delta}$ a.e. Therefore,

$$E^z_{\tau_\delta} 1_{T<\infty} v^i_\infty \le E^z_{\tau_\delta} 1_{T<\infty} \bar{g}^i + \varepsilon/6,$$

which together with (35) implies (34). □

The punishing strategies. Recall that $v : S \to \mathbb{R}^N$ is the asymptotic limiting-average minmax of the family $(\Gamma_\delta)_{\delta>0}$. It follows that for every $\varepsilon > 0$, $z \in S$, $i \in N$, and δ sufficiently small, there is a strategy profile $\sigma^{-i}_{\delta,\varepsilon}$ of players $N \setminus \{i\}$ such that for every strategy τ^i of player i we have

$$\bar{\gamma}^i_\delta(z, \sigma^{-i}_{\delta,\varepsilon}, \tau^i) := E^z_{\delta,\sigma^{-i}_{\delta,\varepsilon},\tau^i} \bar{g}^i \le v^i(z) + \varepsilon/3.$$

The limiting-average ε-equilibrium strategy σ_δ. The strategy profile σ_δ follows the pure strategy profile τ_δ as long as the play coincides with a play that is compatible with the strategy τ_δ, and reverts to punishing (in the lim sup game Γ_δ) a deviating player. A formal definition of σ_δ follows.

Let k_δ be the the first stage m with $a_m \ne \tau_\delta(z_0, a_0, \ldots, z_m)$; $k_\delta = \infty$ if $a_m = \tau_\delta(z_0, a_0, \ldots, z_m)$ for every $m \ge 0$. Fix an order of the player set N, and on $k_\delta < \infty$ let i_δ be the minimal player i with $a^i_{k_\delta} \ne \tau^i_\delta(z_0, a_0, \ldots, z_{k_\delta})$. For every player $i \in N$,

$$\sigma^{-i}_\delta(z_0, a_0, \ldots, z_m) = \begin{cases} \tau^{-i}_\delta(z_0, a_0, \ldots, z_m) & \text{if } k_\delta \ge m \\ \sigma^{-i}_{\delta,\varepsilon}(z_{k_\delta+1}, a_{k_\delta+1}, \ldots, z_m) & \text{if } k_\delta < m \text{ and } i = i_\delta. \end{cases}$$

To complete the definition of the strategy profile σ_δ, there is a need to define $\sigma^i_\delta(z_0, a_0, \ldots, z_m)$ on $k_\delta < m$ and $i = i_\delta$. However, this has no impact on the reasoning that follows. We therefore define it arbitrarily.

Let τ^i be a pure strategy of player i. Note that $(\sigma^{-i}_\delta, \tau^i)$ is a pure strategy profile. Let n_δ be the stopping time of the first stage m such that

$\tau^i(z_0, a_0, \ldots, z_m) \neq \tau^i_\delta(z_0, a_0, \ldots, z_m)$. Note that for every state z, with $P^z_{\tau^{-i}_\delta, \tau^i}$-probability 1, $k_\delta = n_\delta$, and $i_\delta = i$ on $k_\delta < \infty$. Let $\mathcal{H}_{n_\delta}$ be the σ-algebra generated by all $(z_m)_{m \leq n_\delta}$ and $(a_m)_{m<n_\delta}$.

$$\begin{aligned}
\bar\gamma^i(z, \sigma^{-i}_\delta, \tau^i) &= E^z_{\sigma^{-i}_\delta, \tau^i} \bar g^i && (36)\\
&= E^z_{\sigma^{-i}_\delta, \tau^i} E^z_{\sigma^{-i}_\delta, \tau^i}(\bar g^i \mid \mathcal{H}_{n_\delta}) && (37)\\
&= E^z_{\sigma^{-i}_\delta, \tau^i}(1_{n_\delta=\infty} + 1_{n_\delta<\infty}) E^z_{\sigma^{-i}_\delta, \tau^i}(\bar g^i \mid \mathcal{H}_{n_\delta}) && (38)\\
&= E^z_{\sigma_\delta} 1_{n_\delta=\infty}\, \bar g^i + E^z_{\sigma^{-i}_\delta, \tau^i} 1_{n_\delta<\infty}\, E^z_{\sigma^{-i}_\delta, \tau^i}(\bar g^i \mid \mathcal{H}_{n_\delta}) && (39)\\
&\leq E^z_{\sigma_\delta} 1_{n_\delta=\infty}\, \bar g^i + E^z_{\sigma^{-i}_\delta, \tau^i} 1_{n_\delta<\infty}\, v^i(z_{n_\delta+1}) + \varepsilon/3 && (40)\\
&\leq E^z_{\sigma_\delta} 1_{n_\delta=\infty}\, \bar g^i + E^z_{\sigma^{-i}_\delta, \tau^i} 1_{n_\delta<\infty}\, v^i(z_{n_\delta}) + \varepsilon/2 && (41)\\
&\leq E^z_{\sigma_\delta} \bar g^i + 5\varepsilon/6 && (42)\\
&\leq u^i(z) + \varepsilon. && (43)
\end{aligned}$$

Equality (36) follows from the definition of $\bar\gamma^i(z, \sigma^{-i}_\delta, \tau^i)$. Equality (37) follows from one of the basic properties of conditional expectation: that the expectation equals the expectation of the conditional expectation. Equality (38) follows from the rewriting of the constant function 1 as the sum of the two $\{0,1\}$-valued functions $1_{n_\delta=\infty}$ and $1_{n_\delta<\infty}$. Equality (39) follows from the facts that 1) the expectation is additive, 2) $1_{n_\delta=\infty}$ is measurable with respect to σ-algebra $\mathcal{H}_{n_\delta}$ and therefore $E^z_{\sigma^{-i}_\delta, \tau^i} 1_{n_\delta=\infty}\, \bar g^i = E^z_{\sigma^{-i}_\delta, \tau^i} 1_{n_\delta=\infty}\, E^z_{\sigma^{-i}_\delta, \tau^i}(\bar g^i \mid \mathcal{H}_{n_\delta})$, and 3) the $P^z_{\sigma_\delta}$-distribution and the $P^z_{\sigma^{-i}_\delta, \tau^i}$-distribution of $1_{n_\delta=\infty} \bar g^i$ coincide. Inequality (40) follows from the definitions of σ^{-i}_δ and $\sigma^{-i}_{\delta,\varepsilon}$. Inequality (41) follows from the fact that for sufficiently small $\delta > 0$, for every strategy σ and stopping time T, $E^z_\sigma 1_{T<\infty} v^i(z_{T+1}) \leq E^z_\sigma 1_{T<\infty} v^i(z_T) + \varepsilon/6$. Inequality (42) follows from Lemma 4, which asserts that for every stopping time T, $E^z_{\sigma_\delta} 1_{T<\infty}\, v^i(z_T) \leq E^z_{\sigma_\delta} 1_{T<\infty}\, \bar g^i + \varepsilon/3$. Inequality (43) follows from inequality (32) which asserts that $E^z_{\sigma_\delta} \bar g^i \leq u^i(z) + \varepsilon/6$.

By (32), $E^z_{\sigma_\delta} \underline{g}^i \geq u^i(z) - \varepsilon$, which together with the equality $\underline{\gamma}^i(z, \sigma_\delta) = E^z_{\sigma_\delta} \underline{g}^i$ implies that $\underline{\gamma}^i(z, \sigma_\delta) \geq u^i(z) - \varepsilon$. We conclude that (u, σ_δ) is a limiting-average ε-equilibrium payoff and strategy, and therefore u is an asymptotic limiting-average equilibrium payoff. □

Theorem 12 *An exact family $(\Gamma_\delta)_{\delta>0}$ has an asymptotic uniform equilibrium payoff.*

Proof. First, recall that an exact family has an asymptotic uniform minmax. The uniform ε-equilibrium strategy σ_δ follows the pure strategy profile τ_δ (defined in the proof of the previous theorem), and reverts to punishing a deviating player (in the uniform game). □

Theorem 13 *An exact family* $(\Gamma_\delta)_{\delta>0}$ *has an asymptotic* $\overrightarrow{w}$*-robust equilibrium payoff whenever* $\overrightarrow{w} = (w^i)_{i\in N}$ *is a vector of nonstationary discounting measures on* $[0,\infty]$.

Proof. For $\beta = (\beta^i)_{i\in N} \in \mathbb{R}^N_+$ we denote by $\beta * \overrightarrow{w}$ the vector $(\beta^i w^i)_{i\in N}$. Note that if $\beta^i > 0$ for every $i \in N$, then the family $(\Gamma_\delta)_{\delta>0}$ has an asymptotic $\overrightarrow{w}$-robust equilibrium payoff if and only if it has an asymptotic $\beta * \overrightarrow{w}$-robust equilibrium payoff. Therefore we may assume that $w^i(\infty) = 1$.

Fix $\varepsilon > 0$ and an asymptotic 1_∞-robust equilibrium payoff $\nu \in \mathbb{R}^{N\times S}$ of the exact family $(\Gamma_\delta)_{\delta>0}$. Let $0 < t < \infty$ be such that $w^i([t,\infty)) < \varepsilon/\|g\|$ for every $i \in N$, and let $m_\delta = [t/\delta]$ and $\nu_\delta = \nu$. Then, (m_δ, ν_δ) converges to (t,ν). Let $v \in \mathbb{R}^{N\times S}$ be an asymptotic $(\overrightarrow{w}_t, t, \nu)$ equilibrium payoff of the family $(\Gamma_\delta)_{\delta>0}$, where $\overrightarrow{w}_t$ is the restriction of $\overrightarrow{w}$ to the interval $[0,t)$. If $\overrightarrow{w}_\delta$ converges (as δ goes to zero) to $\overrightarrow{w}$, then $\overrightarrow{w}_{t,\delta}$ – the restriction of $\overrightarrow{w}_\delta$ to $\{0,1,2,\ldots,m_\delta\}$ – converges to $\overrightarrow{w}_t$.

If σ_δ is the strategy profile that follows up to stage m_δ an ε-equilibrium strategy profile in $\Gamma^{m_\delta,\nu}_{\delta,\overrightarrow{w}_{t,\delta}}$ with a payoff within ε of v, and thereafter a 1_∞-robust ε-equilibrium with a payoff within ε of ν, then for every player i and all strategies τ^i_δ ($\delta > 0$) of player i in Γ_δ,

$$6\varepsilon + \liminf_{\delta\to 0+} E^z_{\sigma_\delta} g^i_\delta(w^i_\delta) \geq v^i(z) \geq -6\varepsilon + \limsup_{\delta\to 0+} E^z_{\sigma_\delta,\tau^i_\delta} g^i_\delta(w^i_\delta).$$

Therefore, the exact family $(\Gamma_\delta)_{\delta>0}$ has an asymptotic $\overrightarrow{w}$-robust equilibrium payoff. □

References

[1] Bewley, T. and E. Kohlberg (1976), The asymptotic theory of stochastic games, *Mathematics of Operations Research*, 1, 197–208.

[2] Guo, X. and O. Hernandez-Lerma (2005), Nonzero-sum games for continuous-time Markov chains with unbounded discounted payoffs, *Journal of Applied Probability*, 42, 303–320.

[3] Jasso-Fuentes, H. (2005), Noncooperative continuous-time Markov games, *Morfismos*, 9, 39–54.

[4] Levy, Y. (2012), Continuous-time stochastic games of fixed duration, DP 617, Center for the Study of Rationality, Hebrew University of Jerusalem.

[5] Mertens, J.-F. and A. Neyman (1981), Stochastic games, *International Journal of Game Theory*, 10, 53–66.

[6] Neyman, A. (2003), Real algebraic tools in stochastic games, in *Stochastic Games and Applications*, A. Neyman and S. Sorin (eds.), NATO ASI Series, Kluwer Academic Publishers, pp. 57–75.

[7] Neyman, A. (2003), Stochastic games: Existence of the minmax, in *Stochastic Games and Applications*, A. Neyman and S. Sorin (eds.), NATO ASI Series, Kluwer Academic Publishers, pp. 173–193.

[8] Neyman, A. (2012), Continuous-time stochastic games, DP 616, Center for the Study of Rationality, Hebrew University of Jerusalem.

[9] Shapley, L. S. (1953), Stochastic games, *Proceedings of the National Academy of Sciences of the U.S.A.*, 39, 1095–1100.

[10] Solan, E. (2003), Continuity of the value of competitive Markov decision processes, *Journal of Theoretical Probability*, 16, 831–845.

[11] Solan, E. and N. Vieille (2002), Correlated equilibrium in stochastic games, *Games and Economic Behavior*, 38, 362–399.

[12] Solan, E. and R. Vohra (2002), Correlated equilibrium and public signalling in absorbing games, *International Journal of Game Theory*, 31, 91–121.

[13] Vigeral, G. (2012), A zero-sum stochastic game with compact action sets and no asymptotic value, CEREMADE, Universite Paris-Dauphine, preprint.

[14] Zachrisson, L. E. (1964), Markov games, in *Advances in Game Theory*, Princeton: Princeton University Press.