

Eibelshäuser, Steffen; Poensgen, David

Conference Paper

Markov Quantal Response Equilibrium and a Homotopy Method for Computing and Selecting Markov Perfect Equilibria of Dynamic Stochastic Games

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2019: 30 Jahre Mauerfall - Demokratie und Marktwirtschaft - Session: Economic Theory I, No. F07-V1

Provided in Cooperation with:

Verein für Socialpolitik / German Economic Association

Suggested Citation: Eibelshäuser, Steffen; Poensgen, David (2019) : Markov Quantal Response Equilibrium and a Homotopy Method for Computing and Selecting Markov Perfect Equilibria of Dynamic Stochastic Games, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2019: 30 Jahre Mauerfall - Demokratie und Marktwirtschaft - Session: Economic Theory I, No. F07-V1, ZBW - Leibniz-Informationszentrum Wirtschaft, Kiel, Hamburg

This Version is available at:

<https://hdl.handle.net/10419/203603>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Markov Quantal Response Equilibrium and a Homotopy Method for Computing and Selecting Markov Perfect Equilibria of Dynamic Stochastic Games*

Steffen Eibelshäuser[†] David Poensgen[‡]

January 10, 2019

Abstract

We formally define Markov quantal response equilibrium (QRE) and prove existence for all finite discounted dynamic stochastic games. The special case of logit Markov QRE constitutes a mapping from precision parameter $\lambda \in \mathbb{R}_0^+$ to sets of logit Markov QRE. The limiting points of this correspondence are shown to be Markov perfect equilibria. Furthermore, the logit Markov QRE correspondence can be given a homotopy interpretation. We prove that for all games, this homotopy contains a branch connecting the unique solution at $\lambda = 0$ to a unique limiting Markov perfect equilibrium. This result can be leveraged both for the computation of Markov perfect equilibria, and also as a selection criterion.

Keywords: Homotopy continuation, Stationary equilibrium, Logit choice.

JEL Classification: C63, C73.

*We thank Volker Benndorf, Matthias Blonski, Michael Kosfeld and Eilon Solan as well as seminar participants in Frankfurt and Tel Aviv for helpful comments and discussions.

[†]Goethe University Frankfurt, e-mail: eibelshaeuser@econ.uni-frankfurt.de.

[‡]Goethe University Frankfurt, e-mail: poensgen@econ.uni-frankfurt.de.

Contents

1	Introduction	3
1.1	Related Literature	4
2	Prerequisites	5
2.1	Dynamic Stochastic Games	6
2.2	Markov Perfect Equilibrium	6
2.3	Homotopy Continuation	8
3	Markov Quantal Response Equilibrium	9
3.1	Definition	10
3.2	Logit Markov Quantal Response Equilibrium	11
3.3	Convergence to Markov Perfect Equilibrium	12
4	Homotopy Method	13
4.1	Intuition	13
4.2	Logit Markov QRE Homotopy	14
5	Limiting Equilibrium	15
5.1	Definition	15
5.2	Existence and Uniqueness of Principal Branch	17
5.3	Existence and Uniqueness of Limiting Equilibrium	19
6	Conclusion	20
	References	22
A	Derivation of Homotopy	25

1 Introduction

Economic environments are typically not stable, but highly dynamic. Current choices do not only carry immediate consequences, but also shape the options available in the future. Examples include pricing, the accumulation or depletion of resources, savings or capacities, as well as entering legal obligations through contracts. Intertemporal tradeoffs are clearly reflected in partial and general equilibrium analysis, where dynamic programming is ubiquitous.

However, when one looks at the analysis of strategic interaction, the picture changes. The most prominent models here are either one-shot games, or at best repeated games – incorporating some dynamics between the players, but assuming an essentially state-less world in which the only lasting consequences of actions stem from the reactions of others.

This limitation is not due to a lack of theoretical concepts. Dynamic interaction among forward-looking economic agents can be modeled as a dynamic stochastic game, a broad class of games dating back to [Shapley \(1953\)](#) – generalizing both repeated games (by introducing states) and Markov decision processes (by introducing strategic interaction). The class of games covers essentially any game of complete information. Interestingly, Shapley’s treatment predates Bellman’s famous results on dynamic programming ([Bellman, 1954](#)).

But dynamic stochastic games are generally difficult to solve. Analytical solutions are generally not available. This is true for dynamic programming problems as well – but here, powerful numerical methods are available. Unfortunately, these methods are not readily transferable to dynamic games: They are iterative in nature, and typically do not converge when strategic interaction is present. Developing well-suited numerical methods is thus a crucial step to enable economists to analyze strategic interaction in general dynamic environments.

The most common solution concept is Markov perfect equilibrium ([Maskin and Tirole, 1988a,b](#)), a refinement of subgame perfect Nash equilibrium in which strategies are conditional on the current *state* of the game, independent of past plays.

The most famous algorithms to compute Markov perfect equilibria are due to [Pakes and McGuire \(1994, 2001\)](#). The algorithms are based on value function iteration, but are not guaranteed to converge. In fact, they only allow to find

Markov perfect equilibria in *pure* strategies. If there exists no Markov perfect equilibrium in pure strategies, value function iteration cycles and fails to converge.

Up to now, to the best of our knowledge, there exists only one algorithm able to compute Markov perfect equilibria in potentially *mixed* strategies. It is proposed by [Herings and Peeters \(2004\)](#) and based on homotopy continuation. The method resembles the famous linear tracing procedure ([Harsanyi, 1975](#); [Harsanyi and Selten, 1988](#)). It starts at arbitrary prior beliefs about other players' strategies and gradually transforms beliefs until equilibrium beliefs are obtained. However, convergence is only guaranteed for *generic* games. Convergence fails if, at some intermediate belief along the homotopy path, the set of Nash equilibria is infinite. This may very well happen in applications. Harsanyi and Selten themselves were aware of this problem. In order to ensure uniqueness of best responses, they devise the *logarithmic* tracing procedure which adds a logarithmic penalty term, forcing strategies towards the centroid.

In this paper, we propose a homotopy method to compute Markov perfect equilibria that is guaranteed to converge for *all* finite dynamic stochastic games. As a byproduct, we propose a selection criterion for Markov perfect equilibria that is guaranteed to select a unique equilibrium. The algorithm is based on a logit quantal response framework ([McKelvey and Palfrey, 1995](#)). As a foundation, we provide a formal extension of quantal response equilibrium (QRE) to the domain of dynamic stochastic games. Furthermore, we establish results existence and convergence to Markov perfect equilibria.

1.1 Related Literature

The following paragraphs summarize the most important references for the concepts we draw upon. Stochastic games and equilibria in Markov strategies find extensive formal treatments in [Mertens et al. \(2015\)](#), or in the monographs by [Filar and Vrieze \(1997\)](#) and [Basar and Olsder \(1999\)](#). A comprehensive, general introduction to the homotopy method is given by [Zangwill and Garcia \(1981\)](#), who focus on the mathematical foundation of this technique. [Allgower and Georg \(1990\)](#) complement this with a thorough treatment of its efficient and stable algorithmic implementation. An overview of its applications in computational game theory is due to [Herings and Peeters \(2010\)](#); this includes in particular the tracing homotopy for dynamic stochastic games by [Herings and Peeters \(2004\)](#) themselves.

The homotopy we propose is based on the concept of quantal response equilibrium (QRE), first formulated for normal-form games by (McKelvey and Palfrey, 1995) and subsequently extended to extensive-form games as agent QRE in (McKelvey and Palfrey, 1998). A recent overview over applications and findings is provided by Goeree et al. (2016). While QRE is originally a behavioral solution concept, the classical solution concepts typically arise as limiting cases, which can be utilized for computational purposes. Turocy (2005) is the first to suggest a homotopy based on the QRE correspondence, first for normal-form games. This is extended to extensive-form games in Turocy (2010), suggesting the computation of sequential equilibria using agent QRE as homotopy. While the concept of logit Markov QRE and an according homotopy are first introduced by Breitmoser et al. (2010), an explicit formal treatment is yet to be introduced in the literature. This is done in the present paper. Specifically, we provide a formal definition of Markov QRE and prove its existence for all finite dynamic stochastic games. Furthermore, we show that there exists a unique principal branch connecting the centroid of the strategy simplex to a unique limiting Markov perfect equilibrium.

The remainder of this paper is organized as follows. Section 2 familiarizes the reader with dynamic stochastic games, Markov perfect equilibrium and homotopy continuation. Section 3 introduces the concept of Markov quantal response equilibrium and establishes existence. Particular attention is devoted to the special case of logit Markov QRE. Limiting points of the logit Markov QRE correspondence are shown to be Markov perfect equilibria. Section 4 describes the homotopy method that can be used to trace the logit Markov QRE homotopy. Finally, section 5 states and proves our main result, namely that the homotopy method converges to a *unique* Markov perfect equilibrium for *any* dynamic stochastic game. Section 6 summarizes and concludes the paper.

2 Prerequisites

In this section, we briefly review the fundamentals of *dynamic stochastic games*, *Markov perfect equilibrium* and *homotopy continuation*.

2.1 Dynamic Stochastic Games

A dynamic stochastic game (Shapley, 1953) is a sequence of normal-form games embedded in a dynamic system. Each period, the corresponding game is characterized by a *state* which is determined stochastically. Players obtain payoffs in the current game and exert influence on the likelihood of upcoming states. Finally, all players seek to maximize the expected present values of their streams of payoffs.

Definition 1. *Dynamic stochastic game.*

A dynamic stochastic game \mathcal{G} is a tuple $\mathcal{G} = \langle S, I, \mathbf{A}, \mathbf{U}, \Phi, \Phi_0, \delta \rangle$ with

S : set of states,

I : set of players,

$\mathbf{A} = (\mathbf{A}_s)_{s \in S} = (A_{s,i})_{s \in S, i \in I}$: action sets,

$\mathbf{U} = (u_{s,i}(\mathbf{a}_s))_{\mathbf{a}_s \in \mathbf{A}_s, s \in S, i \in I}$: payoff functions $u_{s,i} : \mathbf{A}_s \rightarrow \mathbb{R}$,

$\Phi = (\phi_{s \rightarrow s'}(\mathbf{a}_s))_{\mathbf{a}_s \in \mathbf{A}_s, s, s' \in S}$: state transition probabilities $\phi_{s \rightarrow s'} : \mathbf{A}_s \rightarrow \Delta(S)$,

$\Phi_0 = (\phi_{s_0})_{s_0 \in S}$: probability distribution $\mathbb{P} \in \Delta(S)$ over initial state s_0 ,

$\delta = (\delta_i)_{i \in I}$: discount factors.

We restrict our attention to *finite* dynamic stochastic games in *discrete time*, time runs in discrete periods $t \in \mathbb{N}_0$ and the sets of states, players and actions are all finite. As usual, payoffs and state transitions extend to mixed strategy profiles. Bold symbols indicate vectors or multidimensional arrays and index $-i$ denotes all players except player i .

2.2 Markov Perfect Equilibrium

The most common solution concept for dynamic stochastic games is Markov perfect equilibrium. Markov perfect equilibrium is a refinement of subgame perfect Nash equilibrium, in which all players are limited to the use of Markov strategies. Markov strategies, in turn, restrict players to condition their responses exclusively on the current state of the game, not on the history of plays.¹

¹Markov perfect equilibrium is formally introduced in Maskin and Tirole (1988a,b, 2001), but the corresponding concept is already present in Shapley (1953) under the name “stationary equilibrium”.

Definition 2. Markov strategy.

A Markov strategy $\sigma_i(s)$ for player $i \in I$ is a function $\sigma_i : S \rightarrow \Delta(A_{s,i})$ on the domain of states, mapping state s to probability distribution \mathbb{P} over state-specific actions $A_{s,i}$ such that $\sigma_i(s, a_{s,i}) = \mathbb{P}(a_{s,i}|s)$.

Definition 3. Markov perfect equilibrium.

A Markov perfect equilibrium $(\sigma_i^*)_{i \in I}$ is a Nash equilibrium in Markov strategies.

The existence of Markov perfect equilibria in dynamic stochastic games has long been established in the literature (Fink, 1964; Takahashi, 1964; Sobel, 1971). Furthermore, by a straightforward application of Bellman's principle of optimality (Bellman, 1954), Markov perfect equilibria admit the following recursive representation.

Theorem 1. Recursive representation of Markov perfect equilibrium.

A Markov strategy profile $(\sigma_i^*)_{i \in I}$ constitutes a Markov perfect equilibrium if and only if

1. for all players $i \in I$, there exist state value functions $V_i^* : S \rightarrow \mathbb{R}$ such that

$$V_i^*(s) = \max_{a_{s,i} \in A_{s,i}} u_{s,i}(a_{s,i}, \sigma_{s,-i}^*) + \delta_i \sum_{s' \in S} \phi_{s \rightarrow s'}(a_{s,i}, \sigma_{s,-i}^*) \cdot V_i^*(s')$$

holds for all states $s \in S$ and

2. for all states $s \in S$, strategy profile $(\sigma_{s,i}^*)_{i \in I}$ constitutes a Nash equilibrium of the normal-form game with action space \mathbf{A}_s and payoffs

$$\tilde{u}_{s,i}(\mathbf{a}_s) = u_{s,i}(\mathbf{a}_s) + \delta_i \sum_{s' \in S} \phi_{s \rightarrow s'}(\mathbf{a}_s) \cdot V_i^*(s')$$

for $\mathbf{a}_s \in \mathbf{A}_s$ and all players $i \in I$.

Proof. See for example (Doraszelski and Escobar, 2010, p. 374). ■

In dynamic stochastic games, decision making is based on

$$\tilde{u}_{s,i}(\sigma_s, \mathbf{V}_i) = u_{s,i}(\sigma_s) + \delta_i \sum_{s' \in S} \phi_{s \rightarrow s'}(\sigma_s) \cdot V_{s',i},$$

the present value of payoffs *including subsequent course of play*. However, the subsequent course of play depends on the strategy profile of all players. Therefore,

strategies and decision-relevant payoffs are interdependent. As a consequence, Markov perfect equilibria can be found by solving for equilibrium strategies σ^* and equilibrium state values V^* *simultaneously*, i.e. by solving the following system of equations for all $s \in S$ and $i \in I$:

$$\begin{aligned}\sigma_{s,i}^* &\in \arg \max_{\sigma_{s,i} \in \Delta(A_{s,i})} \tilde{u}_{s,i}(\sigma_{s,i}, \sigma_{s,-i}^*, V_i^*), \\ V_{s,i}^* &= \max_{\sigma_{s,i} \in \Delta(A_{s,i})} \tilde{u}_{s,i}(\sigma_{s,i}, \sigma_{s,-i}^*, V_i^*).\end{aligned}$$

Due to the maximization operators, the above system is generally very difficult to solve.² We will solve it by first rewriting the equations in terms of quantal response analysis and then applying homotopy continuation.

2.3 Homotopy Continuation

Homotopy continuation methods constitute a numerical solution method suited for high-dimensional, non-linear systems of equations. Compared to most other numerical methods, it has the major advantage of working globally. Iterative Newton-methods for example are only locally convergent, meaning they require a good initial approximation to produce a solution at all. In contrast, homotopy methods arrive at solutions without such a priori knowledge, rendering them an exceptionally powerful tool. In this section, we will briefly sketch the procedure, as a basic understanding is necessary for the following parts of this paper.

The method generally proceeds in two steps: First the formulation of a suitable homotopy function, which implicitly defines a curve from a starting point to the desired solution; and then the numerical traversal along this curve until the solution is obtained. Intuitively, this resembles “bending” the problem until an easy solution is readily available, then reverting it back to the original form, while holding on to the solution.

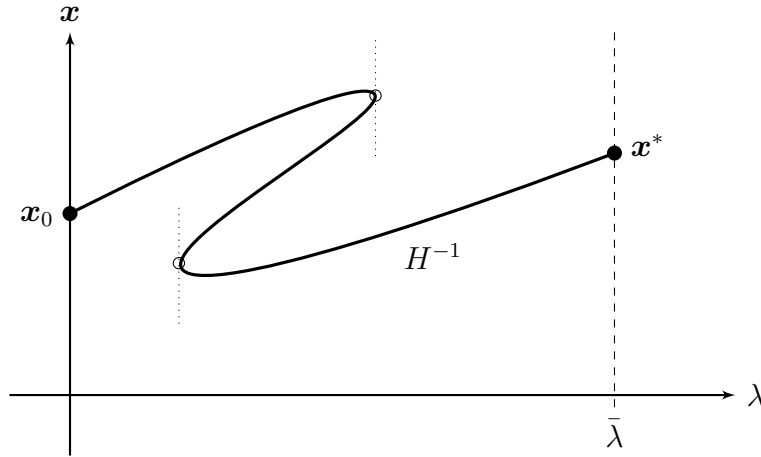
More concretely, suppose one wants to find a solution \mathbf{x}^* to $F(\mathbf{x}) = 0$, where

²Pakes and McGuire (1994, 2001) approach the system by means of value function iteration, i.e. by repeatedly solving for equilibrium strategies and updating the resulting state values. However, the procedure is not guaranteed to converge and, at best, pure-strategy equilibria can be found. Herings and Peeters (2004) solve the system by replacing each optimization with the corresponding Karush-Kuhn-Tucker conditions and performing a transformation of variables to make the resulting equations differentiable. We solve the system by quantal response analysis.

$F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is a high-dimensional, nonlinear mapping. One constructs a function $G : \mathbb{R}^n \rightarrow \mathbb{R}^n$, such that a solution $\mathbf{x}_0 \in G^{-1}(\mathbf{0})$ is known or trivially obtained. Then, a homotopy parameter $\lambda \in [0, \bar{\lambda}]$ with $\bar{\lambda} \in (0, \infty]$ is introduced to construct a homotopy function $H(\mathbf{x}, \lambda) : \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}^n$, with $H(\mathbf{x}, 0) = G(\mathbf{x})$ and $H(\mathbf{x}, \bar{\lambda}) = F(\mathbf{x})$. If H is constructed properly, it thus offers a continuous transformation of the hard problem $F(\mathbf{x}) = 0$ into the trivial one $G(\mathbf{x}) = 0$ and vice versa. The set of solutions $H^{-1} = \{(\mathbf{x}, \lambda) | H(\mathbf{x}, \lambda) = 0\}$ then contains a curve connecting the known solution $(\mathbf{x}_0, 0)$ to the desired solution $(\mathbf{x}^*, \bar{\lambda})$. Tracing this curve to arrive at the latter is the second part of the method. This is done numerically, as described in [Eibelshäuser and Poensgen \(2019\)](#).

The homotopy path might have turning points in the sense that the homotopy parameter λ is not monotonously increasing along the path, as illustrated in [figure 1](#). Therefore, it is generally not possible to follow the path by naively increasing λ . Instead, it is convenient to parameterize the homotopy function in terms of the path length parameter $\tau \in \mathbb{R}_0^+$ such that $H(\mathbf{y}(\tau)) = 0$. For details, see ([Zangwill and Garcia, 1981](#), pp. 25 ff.).

Figure 1: Turning Points of Homotopy Path



3 Markov Quantal Response Equilibrium

In the quantal response framework, players are assumed to perceive payoffs only with some noise. In the resulting *quantal response equilibrium* (QRE) ([McKelvey](#)

and Palfrey, 1995), players' actions appear stochastic and the probability of playing a particular action is related to its true payoff. This idea can be generalized to dynamic games by treating players at different decision nodes as independent agents. The corresponding equilibrium concept is called *agent quantal response equilibrium* (McKelvey and Palfrey, 1998). Finally, in the context of dynamic stochastic games with states as decision nodes, the corresponding equilibrium concept is called *Markov quantal response equilibrium* (Breitmoser et al., 2010).

3.1 Definition

In the context of dynamic stochastic games, players decide on optimal actions based on payoffs *including continuation values*. Specifically, let

$$\tilde{u}_{s,i,a}(\boldsymbol{\sigma}_{s,-i}, \mathbf{V}_i) = u_{s,i}(a, \boldsymbol{\sigma}_{s,-i}) + \delta_i \sum_{s' \in S} \phi_{s \rightarrow s'}(a, \boldsymbol{\sigma}_{s,-i}) \cdot V_{s',i}$$

denote the expected payoff from playing action a for player i in state s , given state values and strategies of the other players. In the quantal response framework, agent $(s, i) \in S \times I$ is assumed to perceive payoffs $\tilde{u}_{s,i,a}(\boldsymbol{\sigma}_{s,-i}, \mathbf{V}_i)$ as

$$\hat{u}_{s,i,a}(\boldsymbol{\sigma}_{s,-i}, \mathbf{V}_i) = \tilde{u}_{s,i,a}(\boldsymbol{\sigma}_{s,-i}, \mathbf{V}_i) + \varepsilon_{s,i,a}$$

with noise $\varepsilon_{s,i,a}$. The error vector $\boldsymbol{\varepsilon}_{s,i} = (\varepsilon_{s,i,a})_{a \in A_{s,i}}$ is assumed to be distributed according to a joint distribution with zero mean and density function $f_{s,i}(\boldsymbol{\varepsilon}_{s,i})$. Let

$$R_{s,i,a} = \left\{ \boldsymbol{\varepsilon}_{s,i} \in \mathbb{R}^{|A_{s,i}|} : \hat{u}_{s,i,a}(\boldsymbol{\sigma}_{s,-i}, \mathbf{V}_i) > \hat{u}_{s,i,a'}(\boldsymbol{\sigma}_{s,-i}, \mathbf{V}_i) \quad \forall a' \in A_{s,i} \right\}$$

denote agent (s, i) 's response set of action a , specifying the realizations of $\boldsymbol{\varepsilon}_{s,i}$ such that agent (s, i) perceives action a as the one with the highest payoff. Then, the probability that player i in state s plays action a is given by the probability mass of the corresponding response set.

Definition 4. Markov quantal response equilibrium.

A Markov quantal response equilibrium is a strategy profile $\boldsymbol{\sigma}^*$ such that

$$\sigma_{s,i,a}^* = \int_{R_{s,i,a}} f_{s,i}(\boldsymbol{\varepsilon}) d\boldsymbol{\varepsilon}.$$

Proving the existence of Markov QRE's in dynamic stochastic games is a straightforward application of Brouwer's fixed-point theorem.

Theorem 2. *Existence of Markov quantal response equilibrium.*

Every dynamic stochastic game \mathcal{G} has a Markov quantal response equilibrium.

Proof. Similar to (McKelvey and Palfrey, 1995, theorem 1), with minor modifications. A Markov quantal response equilibrium σ^* is part of a fixed-point (σ^*, V^*) of the function $g(\sigma, V) = (g^\sigma(\sigma, V), g^V(\sigma, V))$ with

$$\begin{aligned} g_{s,i,a}^\sigma(\sigma, V) &= \int_{R_{s,i,a}} f_{s,i}(\varepsilon) d\varepsilon && \stackrel{!}{=} \sigma_{s,i,a} \\ g_{s,i}^V(\sigma, V) &= \tilde{u}_{s,i}(\sigma_s, V_i) && \stackrel{!}{=} V_{s,i} \end{aligned}$$

for all states s , players i and actions a . Since strategies and state values are bounded, i.e.

$$\begin{aligned} \sigma_{s,i,a} &\in [0, 1] \\ V_{s,i} &\leq \sum_{t=0}^{\infty} \delta_i^t \cdot \max_{s,i,a} \{u_{s,i,a}\} = \frac{1}{1 - \delta_i} \cdot \max_{s,i,a} \{u_{s,i,a}\} < \infty \\ V_{s,i} &\geq \frac{1}{1 - \delta_i} \cdot \min_{s,i,a} \{u_{s,i,a}\} > -\infty \end{aligned}$$

for all s, i and a , the domain of g is compact. Furthermore, since the distribution of noise ε has a density, g is continuous. By Brouwer's fixed-point theorem, g has a fixed-point. ■

For the remainder of this paper, we focus on the logit version of Markov QRE.

3.2 Logit Markov Quantal Response Equilibrium

The most popular special case of quantal response is *logit choice* (Luce, 1959). Logistic rules of choice in the quantal response context arise from noise that is independent and identically distributed according to a Gumbel distribution with parameter $\lambda \in \mathbb{R}_0^+$ (extreme value distribution of type I).³ The corresponding

³The Gumbel distribution has cumulative distribution function $F(\varepsilon) = e^{-e^{-\lambda\varepsilon}}$. The parameter $\lambda \in \mathbb{R}_0^+$ controls the variance of the distribution. For $\lambda = 0$, the variance is infinite, and for $\lambda \rightarrow \infty$, the variance tends to zero.

equilibrium can be expressed in closed form.

Theorem 3. *Logit Markov quantal response equilibrium.*

A logit Markov quantal response equilibrium σ^* is a solution to the following system of equations for all $s \in S$, $i \in I$ and $a \in A_{s,i}$:

$$\begin{aligned}\sigma_{s,i,a}^* &= \frac{e^{\lambda \cdot \tilde{u}_{s,i,a}(\sigma_{s,-i}^*, \mathbf{V}_i^*)}}{\sum_{a' \in A_{s,i}} e^{\lambda \cdot \tilde{u}_{s,i,a'}(\sigma_{s,-i}^*, \mathbf{V}_i^*)}}, \\ V_{s,i}^* &= \sum_{a' \in A_{s,i}} \sigma_{s,i,a'}^* \cdot \tilde{u}_{s,i,a'}(\sigma_{s,-i}^*, \mathbf{V}_i^*).\end{aligned}$$

Proof. The derivation of the logit formula for equilibrium strategies is identical to the famous derivation by [McFadden \(1973\)](#). The formula for equilibrium state values immediately follows from the definitions of $\tilde{u}_{s,i}$ and $\tilde{u}_{s,i,a}$ which can be combined to write

$$\tilde{u}_{s,i}(\sigma_s, \mathbf{V}_i) = \sum_{a' \in A_{s,i}} \sigma_{s,i,a'} \cdot \tilde{u}_{s,i,a'}(\sigma_{s,-i}, \mathbf{V}_i).$$

■

[Theorem 2](#) guarantees the existence of logit Markov QRE for any $\lambda \in \mathbb{R}_0^+$.

3.3 Convergence to Markov Perfect Equilibrium

The noisiness of a logit Markov QRE is parameterized by $\lambda \in \mathbb{R}_0^+$ which can be interpreted as the *precision* of the perception of payoffs. When $\lambda = 0$, the equilibrium is fully noisy and mixing is uniform over actions, i.e. centroid strategies $\sigma_{s,i,a}^* = \frac{1}{|A_{s,i}|}$ for all s, i and a . If a logit Markov QRE converges as $\lambda \rightarrow \infty$, the limit point is a Markov perfect equilibrium.

Theorem 4. *Limiting Markov perfect equilibria.*

Let σ^{**} be the limit point of some logit Markov QRE $\sigma^*(\lambda)$, i.e.

$$\sigma^{**} = \lim_{\lambda \rightarrow \infty} \sigma^*(\lambda).$$

Then σ^{**} is a Markov perfect equilibrium.

Proof. By contradiction, similar to (McKelvey and Palfrey, 1995, theorem 2), with minor modifications. Suppose σ^{**} is not a Markov perfect equilibrium. Then, according to theorem 1, σ^{**} cannot be a Nash equilibrium of the normal-form game with payoffs $\tilde{u}_{s,i,a}(\sigma_{s,-i}^{**}, \mathbf{V}_i^{**})$. Therefore, there exists an agent $(s, i) \in S \times I$ and actions $a, a' \in A_{s,i}$ such that

$$\lim_{\lambda \rightarrow \infty} \tilde{u}_{s,i,a}(\sigma_{s,-i}^*(\lambda), \mathbf{V}_i^*(\lambda)) > \lim_{\lambda \rightarrow \infty} \tilde{u}_{s,i,a'}(\sigma_{s,-i}^*(\lambda), \mathbf{V}_i^*(\lambda))$$

and $\sigma_{s,i,a'}^{**} > 0$. Since $\tilde{u}_{s,i,a}(\sigma_{s,-i}^*(\lambda), \mathbf{V}_i^*(\lambda))$ is continuous in λ , there exist $\varepsilon > 0$ and $\bar{\lambda} \in \mathbb{R}_0^+$ such that for all $\lambda > \bar{\lambda}$ we have

$$\tilde{u}_{s,i,a}(\sigma_{s,-i}^*(\lambda), \mathbf{V}_i^*(\lambda)) > \tilde{u}_{s,i,a'}(\sigma_{s,-i}^*(\lambda), \mathbf{V}_i^*(\lambda)) + \varepsilon.$$

But then we must have $\sigma_{s,i,a'}^*(\lambda) \xrightarrow{\lambda \rightarrow \infty} 0$, which contradicts $\sigma_{s,i,a'}^{**} > 0$. ■

4 Homotopy Method

Theorem 4 suggests that Markov perfect equilibria can be found by starting at any logit Markov QRE and then letting $\lambda \rightarrow \infty$. Provided the limit point exists, it is guaranteed to be a Markov perfect equilibrium. We will prove the existence of the limit point later on in this paper – with arguments based on the homotopy interpretation of the logit Markov QRE correspondence.

The intuition for the homotopy method is as follows. We are interested in the complicated problem of finding Markov perfect equilibria. To do so, we solve the simple problem of finding a logit Markov QRE and then distort the solution into a Markov perfect equilibrium of the complicated problem. Specifically, the homotopy method proposed in this paper takes as starting point the unique logit Markov QRE at $\lambda = 0$ and follows the implied path to the limiting Markov perfect equilibrium.

4.1 Intuition

As noted by Turocy (2005), logit QRE define a homotopy between replicator dynamics at $\lambda = 0$ and best response dynamics at $\lambda \rightarrow \infty$. To see that within

our context of dynamic stochastic games, write a logit Markov QRE σ^* in the following short form

$$\sigma_{s,i,a}^* = \frac{e^{\lambda \cdot \tilde{u}_{s,i,a}}}{\sum_{a' \in A_{s,i}} e^{\lambda \cdot \tilde{u}_{s,i,a'}}}$$

and take the partial derivative with respect to λ to get

$$\frac{\partial \sigma_{s,i,a}^*}{\partial \lambda} = \underbrace{\sigma_{s,i,a}^* \cdot \mathbb{E} \left[\Delta \tilde{u}_{s,i,a} \right]}_{\text{replicator dynamics}} + \underbrace{\lambda \cdot \sigma_{s,i,a}^* \cdot \mathbb{E} \left[\Delta \frac{\partial \tilde{u}_{s,i,a}}{\partial \lambda} \right]}_{\text{best response dynamics}},$$

where

$$\mathbb{E} \left[\Delta \tilde{u}_{s,i,a} \right] := \sum_{a' \in A_{s,i}} \sigma_{s,i,a'}^* \cdot (\tilde{u}_{s,i,a} - \tilde{u}_{s,i,a'})$$

denotes the expected utility difference from playing action a relative to the other actions and analogously for $\mathbb{E} \left[\Delta \frac{\partial \tilde{u}_{s,i,a}}{\partial \lambda} \right]$.

As is standard, the replicator dynamics term updates equilibrium strategies based on payoff differences across actions. On the other hand, the best response dynamics term updates equilibrium strategies based on *anticipated* best response behavior of opponents. As λ grows large, the best response term takes over, leading to a Markov perfect equilibrium.

Our homotopy method is strikingly similar to the famous linear tracing procedure due to [Harsanyi \(1975\)](#); [Harsanyi and Selten \(1988\)](#). This procedure can be interpreted as a homotopy from best responses of prior beliefs to actual best responses. Both the tracing procedure as well as our logit Markov homotopy method start at trivial strategies and follow the equilibrium until best response dynamics are reached. However, the two methods generally differ – even in 2x2 normal-form games ([Zhang and Hofbauer, 2016](#)).

4.2 Logit Markov QRE Homotopy

We call the homotopy between the initial logit Markov QRE and the limiting Markov perfect equilibrium *logit Markov QRE homotopy*. The homotopy path is given by the roots of a suitable homotopy function $H : \mathbb{R}^{n+1} \rightarrow \mathbb{R}^n$, i.e. $H(\mathbf{y}) = \mathbf{0}$,

where $n = \sum_{s \in S} \sum_{i \in I} (|A_{s,i}| + 1)$ denotes the number of equations characterizing logit Markov QRE. The function H can be constructed explicitly based on the system of equations given in [theorem 3](#), i.e.

$$\begin{aligned} H_{s,i,a}^\sigma(\boldsymbol{\sigma}, \mathbf{V}, \lambda) &= -\sigma_{s,i,a} + \frac{e^{\lambda \cdot \tilde{u}_{s,i,a}(\boldsymbol{\sigma}_{s,-i}, \mathbf{V}_i)}}{\sum_{a' \in A_{s,i}} e^{\lambda \cdot \tilde{u}_{s,i,a'}(\boldsymbol{\sigma}_{s,-i}, \mathbf{V}_i)}}, \\ H_{s,i}^V(\boldsymbol{\sigma}, \mathbf{V}, \lambda) &= -V_{s,i} + \sum_{a' \in A_{s,i}} \sigma_{s,i,a'} \cdot \tilde{u}_{s,i,a'}(\boldsymbol{\sigma}_{s,-i}, \mathbf{V}_i) \end{aligned}$$

for all s, i and a .

The Jacobian matrix $J : \mathbb{R}^{n+1} \rightarrow \mathbb{R}^n \times \mathbb{R}^{n+1}$ of the homotopy function is defined by

$$J(\mathbf{y}) = \frac{\partial H(\mathbf{y})}{\partial \mathbf{y}}.$$

An explicit derivation of H and J including a convenient transformation of variables is provided in [appendix A](#).

The functions H and J comprise the main information about the logit Markov homotopy. Based on the explicit construction of the two function, we can make statements about the properties of the homotopy path.

5 Limiting Equilibrium

In this section, we state and prove our main result. First, we define the *limiting logit Markov QRE* in terms of the *principal branch* of the logit Markov QRE homotopy. Secondly, we prove the existence and uniqueness of the principal branch. Finally, we prove that the unique limiting point of the principal branch is indeed a Markov perfect equilibrium for *any* dynamic stochastic game.

5.1 Definition

The limiting logit Markov QRE is defined as the limit point *of the principal branch* of the logit Markov quantal response homotopy. The principal branch \mathcal{H}^0 , in turn, is defined as the part of the homotopy path that passes through the centroid

strategies at $\lambda = 0$ and continues *smoothly* thereafter.

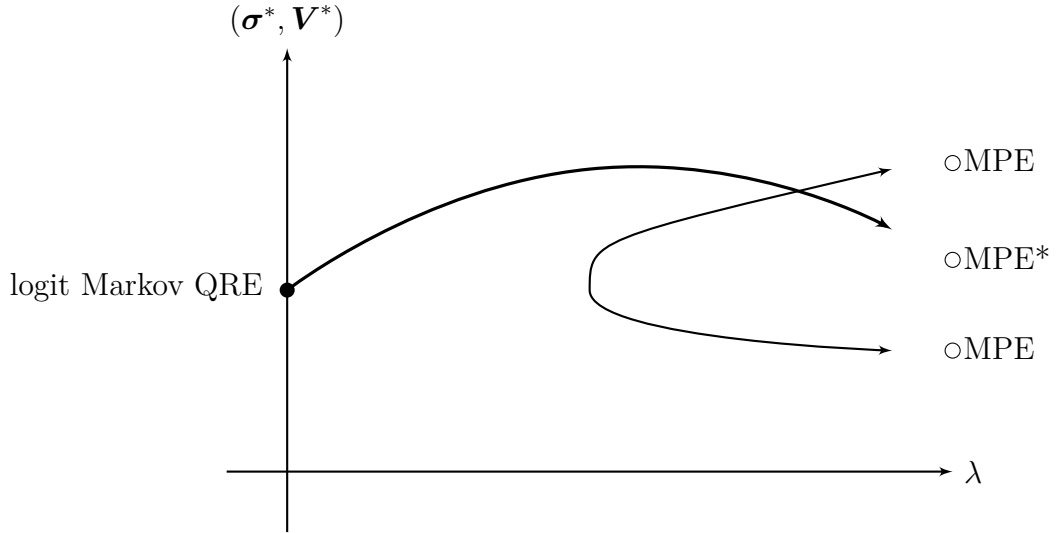
Definition 5. Principal branch.

Let the homotopy function $H(\mathbf{y}(\tau))$ be parameterized by the path length parameter $\tau \in \mathbb{R}_0^+$ with $\lambda(\tau = 0) = 0$ and let $t(\mathbf{y}(\tau))$ denote a tangent at point $\mathbf{y}(\tau)$. Then, the principal branch \mathcal{H}^0 is defined as

$$\mathcal{H}^0 := \left\{ \mathbf{y} \subseteq H^{-1}(\mathbf{0}) \mid \begin{aligned} &\mathbf{y} \text{ smoothly connected to } \mathbf{y}(\tau = 0), \\ &t(\mathbf{y}(\tilde{\tau})) \text{ unique for almost all } \tilde{\tau} \in \mathbb{R}_0^+, \\ &\forall \tilde{\tau} \in \mathbb{R}_0^+ \exists \text{ tangent } t' : \lim_{\tau \rightarrow \tilde{\tau}^+} t'(\mathbf{y}(\tau)) = \pm \lim_{\tau \rightarrow \tilde{\tau}^-} t(\mathbf{y}(\tau)) \end{aligned} \right\}.$$

The definition of the principal branch comprises sufficient properties to guarantee convergence to a unique limiting point. The corresponding principal branch is smooth, as illustrated in [figure 2](#). Provided the principal branch exists and is unique, proving the uniqueness of the limiting equilibrium is fairly straightforward.

Figure 2: Smooth Principal Branch

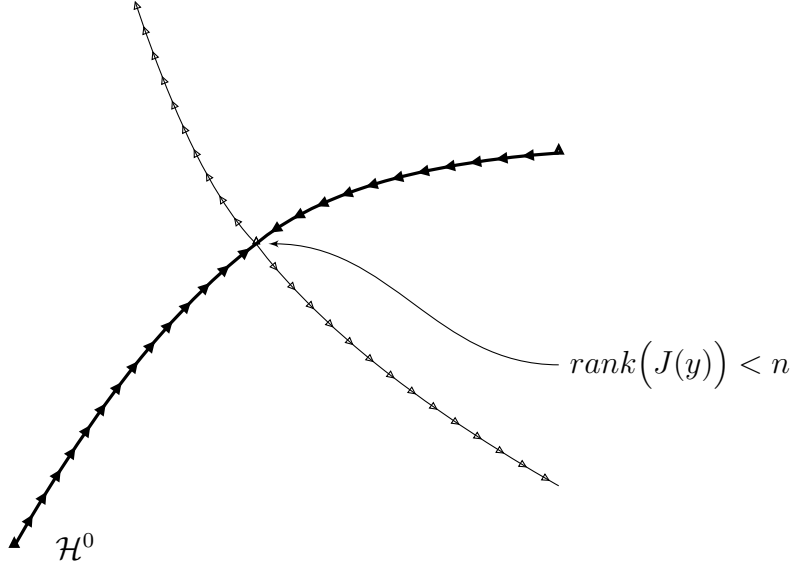


Note that the principal branch is required to have a unique tangent at almost all points. This definition excludes multi-dimensional segments (“areas”) of the path, i.e. the path is required to be one-dimensional. Furthermore, at the isolated points where the tangent is not unique, must exist a smooth continuation with

same or opposite sign. This describes the situation in which the principal branch is crossed by one or more auxiliary paths. At the corresponding crossing point, the tangent of the principal branch is not unique. However, accounting for the potential change in direction, the principal branch continues smoothly after the crossing point, which is illustrated in [figure 3](#).

Being one-dimensional with only isolated crossing points and smooth continuations, the principal branch is guaranteed to lead to a unique limiting point – the limiting Markov perfect equilibrium. Now, we formalize this argumentation.

Figure 3: Simple Bifurcation Point on Principal Branch



Definition 6. *Limiting logit Markov quantal response equilibrium.*

The limiting logit Markov quantal response equilibrium σ^{**} is defined as the limit point of the principal branch, i.e.

$$\sigma^{**} := \lim_{\tau \rightarrow \infty} \sigma^*(\lambda(\tau))$$

for $(\sigma^*(\lambda(\tau)), V^*(\lambda(\tau)), \lambda(\tau)) \in \mathcal{H}^0$.

5.2 Existence and Uniqueness of Principal Branch

Proposition 1. *Existence and uniqueness of principal branch.*

For all finite dynamic stochastic games \mathcal{G} , there exists a unique principal branch \mathcal{H}^0 of the logit Markov QRE homotopy.

Proof. Since H is a continuous function, the homotopy path $H^{-1}(\mathbf{0})$ is upper hemi-continuous and tangents exist. Furthermore, due to the pre-image theorem (Guillemin and Pollack, 2016, p. 21), the homotopy path $H^{-1}(\mathbf{0})$ is a one-dimensional manifold for regular points $y \in \mathbb{R}^{n+1}$ where the Jacobian matrix has full rank $\text{rank}(J(y)) = n$.

The remainder of the proof proceeds in three steps. First, we show that there are only finitely many irregular points $\tilde{y} \in \mathbb{R}^{n+1}$ where the Jacobian matrix has not full rank, i.e. $\text{rank}(J(\tilde{y})) < n$. Secondly, we show that at all irregular points $\tilde{y} \in \mathbb{R}^{n+1}$, exists a unique smooth continuation. Finally, we show that exists a unique starting point of the principal branch at $\lambda = 0$.

Step 1: Irregular points $\tilde{y} \in \mathbb{R}^{n+1}$ with $\text{rank}(J(\tilde{y})) < n$ require that multiple sub-determinants of the Jacobian matrix are zero, i.e. that exist two distinct indices $k \neq l \in \{1, \dots, n\}$ such that $\det(J^{(-k)}(\tilde{y})) = \det(J^{(-l)}(\tilde{y})) = 0$, where $J^{(-k)}$ denotes the Jacobian matrix without its k -th column. However, as is evident from the explicit derivation of the Jacobian in appendix A, all sub-determinants of the Jacobian are finite polynomials in log-strategies β , e^β , \mathbf{V} and λ and, as such, only have a finite number of roots.⁴ Therefore, there can only be finitely many irregular points on the homotopy path $H^{-1}(\mathbf{0})$.

Step 2: Since there are only finitely many irregular points $\tilde{\mathbf{y}} = \mathbf{y}(\tilde{\tau})$ on the homotopy path, all irregular points must be “simple bifurcation points” (Allgower and Georg, 1990, chapter 8) where multiple one-dimensional segments cross. By Lyapunov-Schmidt reduction one can decompose $H^{-1}(\mathbf{0})$ into the different segments and formally show that the tangents on both sides of simple bifurcation points point in same or exactly opposite directions, i.e.

$$\lim_{\tau \rightarrow \tilde{\tau}^+} t(\mathbf{y}(\tau)) = \left| \lim_{\tau \rightarrow \tilde{\tau}^-} t(\mathbf{y}(\tau)) \right|$$

(Allgower and Georg, 1990, theorem 8.1.14). Furthermore, it is not possible that two crossing segments are tangent or that there are multiple smooth continuations.

⁴The fact that e^β appears in the polynomial along with β itself poses no problem because e^β is a strictly increasing function. One can think of the sub-determinants of J as finite polynomials in β , \mathbf{z} , \mathbf{V} and λ . The polynomials clearly have finitely many roots. Now restrict $\mathbf{z} = e^\beta$ and the number of roots cannot increase.

This is because tangents span the kernel of $J(\tilde{\mathbf{y}})$. The dimension $\dim(\ker(J(\tilde{\mathbf{y}})))$ of the kernel corresponds to the number of segments present at point $\tilde{\mathbf{y}}$. Hence tangents must be independent and strictly cross (Keller, 1987, lemma 5.20). Therefore, when accounting for the potential change in direction, the principal branch is well-defined, unique and can be traversed smoothly.

Step 3: Finally, we show that exists a unique starting value $\mathbf{y}(\lambda = 0) = (\boldsymbol{\sigma}^*, \mathbf{V}^*, 0)$. The corresponding equilibrium strategies involve uniform mixing

$$\sigma_{s,i,a}^*(\lambda = 0) = \frac{1}{|A_{s,i}|}.$$

The corresponding state values satisfy

$$V_{s,i}^*(\lambda = 0) = \bar{u}_{s,i} + \delta_i \cdot \bar{\phi}_{s \rightarrow s'} \cdot \sum_{s' \in S} V_{s',i},$$

where

$$\bar{u}_{s,i} = \frac{1}{|\mathbf{A}_s|} \sum_{\mathbf{a}_s \in \mathbf{A}_s} u_{s,i}(\mathbf{a}_s) \quad \text{and} \quad \bar{\phi}_{s \rightarrow s'} = \frac{1}{|\mathbf{A}_s|} \sum_{\mathbf{a}_s \in \mathbf{A}_s} \phi_{s \rightarrow s'}(\mathbf{a}_s)$$

denote average payoff and average transition probabilities, respectively. Accordingly, initial state values can be found by solving a sequence of linear systems of equations, one system for each player, of standard form $\mathbf{M}_i \cdot \mathbf{V}_i^* = \mathbf{b}_i$ with

$$\underbrace{(I_{|S|} - \delta_i \cdot \bar{\boldsymbol{\Phi}})}_{=: \mathbf{M}_i} \cdot \mathbf{V}_i^* = \underbrace{\bar{\mathbf{u}}_i}_{=: \mathbf{b}}.$$

Each system has a unique solution because $\det(\mathbf{M}_i) \neq 0$ for all discount factors $\delta_i \in [0, 1)$ and all transition matrices $\bar{\boldsymbol{\Phi}}$.

This completes the proof. ■

5.3 Existence and Uniqueness of Limiting Equilibrium

The main result of this paper is stated in [theorem 5](#).

Theorem 5. *Uniqueness of limiting logit Markov QRE.*

Every finite dynamic stochastic game \mathcal{G} has a unique limiting logit Markov QRE which constitutes a Markov perfect equilibrium of \mathcal{G} .

Proof. We know from [proposition 1](#) that the logit Markov QRE homotopy of any finite dynamic stochastic game \mathcal{G} has a unique principal branch. The principal branch is one-dimensional and smooth and has a unique starting point for $\lambda = 0$ involving centroid strategies. It remains to show that the limit point of the branch exists and that it is indeed a Markov perfect equilibrium. This is done by proving the existence of a limit point together with $\lambda(\tau) \xrightarrow{\tau \rightarrow \infty} \infty$. Then the result follows from [theorem 4](#).

The proof for existence of a limit point and $\lambda(\tau) \xrightarrow{\tau \rightarrow \infty} \infty$ proceeds in three steps. First, since there is a unique solution $\mathbf{y}(\lambda = 0) = (\boldsymbol{\sigma}^*, \mathbf{V}^*, 0)$, we have

$$\left. \frac{\partial \lambda}{\partial \tau} \right|_{\tau=0} \neq 0.$$

Thus, if the sign of path length parameter τ is chosen appropriately, λ increases initially. Then, we have $\lambda > 0$ on the entire path. Otherwise there would have to be a second solution to $\mathbf{y}(\lambda = 0) = (\boldsymbol{\sigma}^*, \mathbf{V}^*, 0)$. Secondly, the domain of $(\boldsymbol{\sigma}^*, \mathbf{V}^*)$ is bounded (see proof of [theorem 2](#)), so the principal branch may either converge with $\lambda \rightarrow \infty$ or cycle. Thirdly, to exclude cycling, we show that there are finitely many turning points in all variables y_k where $\frac{\partial y_k}{\partial \tau}$ switches sign. Turning points are roots of $\det(J^{(-k)}(y(\tau)))$, but we have already shown in [proposition 1](#) that there are finitely many of them. So the principal branch must converge to a limit point $(\boldsymbol{\sigma}^{**}, \mathbf{V}^{**})$ with $\lambda(\tau) \xrightarrow{\tau \rightarrow \infty} \infty$.

This completes the proof. ■

The main idea of the proof is illustrated in [figure 4](#).

6 Conclusion

In this paper we have given a definition of Markov quantal response equilibria and shown existence for all finite dynamic stochastic games. The according correspondence can be given a homotopy interpretation. As we demonstrate, the solution space includes a uniquely defined principal branch, connecting the unique solution at $\lambda = 0$ to a single limiting Markov perfect equilibrium.

This main result of the current paper opens two avenues. First, the uniqueness of

Figure 4: Limiting Markov perfect equilibrium



the principal branch suggests that it can be used as an equilibrium selection criterion. Second, numerical traversal of the principal branch can be used to efficiently compute at least one Markov perfect equilibrium of any dynamic stochastic game. This traversal is possible since the principle branch, as we have shown, contains as irregular points at most a finite number of simple bifurcations, which pose no obstacle for numerical continuation.

An implementation of the described homotopy method – consisting of the construction of the Markov QRE homotopy function and its subsequent traversal using predictor-corrector steps – is given by the Python program [dsGameSolver](#). Subject to the usual limitations of numerical computation, it can in principle solve any finite game falling under the broad class of dynamic stochastic games; a detailed documentation is given in [Eibelshäuser and Poensgen \(2019\)](#).

References

- ALLGOWER, E. L. AND K. GEORG (1990): *Numerical Continuation Methods: An Introduction*, New York: Springer.
- BASAR, T. AND G. J. OLSDER (1999): *Dynamic Noncooperative Game Theory*, Philadelphia: Society for Industrial and Applied Mathematics.
- BELLMAN, R. (1954): “The Theory of Dynamic Programming,” *Bulletin of the American Mathematical Society*, 60, 503–515.
- BREITMOSER, Y., J. H. TAN, AND D. J. ZIZZO (2010): “Understanding Perpetual R&D Races,” *Economic Theory*, 44, 445–467.
- DORASZELSKI, U. AND J. F. ESCOBAR (2010): “A Theory of Regular Markov Perfect Equilibria in Dynamic Stochastic Games: Genericity, Stability, and Purification,” *Theoretical Economics*, 5, 369–402.
- EIBELSHÄUSER, S. AND D. POENSGEN (2019): “dsGameSolver: A Python Program for Computing Markov Perfect Equilibria of Dynamic Stochastic Games,” *Working paper*.
- FILAR, J. AND K. VRIEZE (1997): *Competitive Markov Decision Processes*, New York: Springer.
- FINK, A. M. (1964): “Equilibrium in a Stochastic n-Person Game,” *Journal of Science of the Hiroshima University, Series A-I (Mathematics)*, 28, 89–93.
- GOEREE, J. K., C. A. HOLT, AND T. R. PALFREY (2016): *Quantal Response Equilibrium: A Stochastic Theory of Games*, Princeton, New Jersey: Princeton University Press.
- GUILLEMIN, V. AND A. POLLACK (2016): *Differential Topology*, Englewood Cliffs, New Jersey: Prentice-Hall.
- HARSANYI, J. C. (1975): “The Tracing Procedure: A Bayesian Approach to Defining a Solution for n-Person Noncooperative Games,” *International Journal of Game Theory*, 4, 61–94.
- HARSANYI, J. C. AND R. SELTEN (1988): *A General Theory of Equilibrium Selection in Games*, Cambridge, Massachusetts: MIT Press.

- HERINGS, P. J.-J. AND R. J. PEETERS (2004): “Stationary Equilibria in Stochastic Games: Structure, Selection, and Computation,” *Journal of Economic Theory*, 118, 32–60.
- (2010): “Homotopy Methods to Compute Equilibria in Game Theory,” *Economic Theory*, 42, 119–156.
- KELLER, H. B. (1987): “Lectures on Numerical Methods in Bifurcation Problems,” *Applied Mathematics*, 217.
- LUCE, R. D. (1959): *Individual Choice Behavior*, New York: Wiley.
- MASKIN, E. AND J. TIROLE (1988a): “A Theory of Dynamic Oligopoly I: Overview and Quantity Competition with Large Fixed Costs,” *Econometrica*, 56, 549–569.
- (1988b): “A Theory of Dynamic Oligopoly II: Price Competition, Kinked Demand Curves, and Edgeworth Cycles,” *Econometrica*, 56, 571–599.
- (2001): “Markov Perfect Equilibrium: I. Observable Actions,” *Journal of Economic Theory*, 100, 191–219.
- McFADDEN, D. L. (1973): “Conditional Logit Analysis of Qualitative Choice Behavior,” in *Frontiers in Econometrics*, ed. by P. Zarembka, New York: Academic Press.
- McKELVEY, R. D. AND T. R. PALFREY (1995): “Quantal Response Equilibria for Normal Form Games,” *Games and Economic Behavior*, 10, 6–38.
- (1998): “Quantal Response Equilibria for Extensive Form Games,” *Experimental Economics*, 1, 9–41.
- MERTENS, J.-F., S. SORIN, AND S. ZAMIR (2015): *Repeated Games*, Cambridge: Cambridge University Press.
- PAKES, A. AND P. MCGUIRE (1994): “Computing Markov Perfect Nash Equilibria: Numerical Implications of a Dynamic Differentiated Product Model,” *RAND Journal of Economics*, 25, 555–589.
- (2001): “Stochastic Algorithms, Symmetric Markov Perfect Equilibrium, and the Curse of Dimensionality,” *Econometrica*, 69, 1261–1281.

- SHAPLEY, L. S. (1953): “Stochastic Games,” *Proceedings of the National Academy of Sciences*, 39, 1095–1100.
- SOBEL, M. J. (1971): “Non-Cooperative Stochastic Games,” *Annals of Mathematical Statistics*, 42, 1930–1935.
- TAKAHASHI, M. (1964): “Equilibrium Points of Stochastic, Noncooperative n-Person Games,” *Journal of Science of the Hiroshima University, Series A-I (Mathematics)*, 28, 95–99.
- TUROC, T. L. (2005): “A Dynamic Homotopy Interpretation of the Logistic Quantal Response Equilibrium,” *Games and Economic Behavior*, 51, 243–263.
- (2010): “Computing Sequential Equilibria Using Agent Quantal Response Equilibria,” *Economic Theory*, 42, 255–269.
- ZANGWILL, W. I. AND C. B. GARCIA (1981): *Pathways to Solutions, Fixed Points, and Equilibria*, Upper Saddle River, New Jersey: Prentice-Hall.
- ZHANG, B. AND J. HOFBAUER (2016): “Quantal Response Methods for Equilibrium Selection in 2x2 Coordination Games,” *Games and Economic Behavior*, 97, 19–31.

A Derivation of Homotopy

The homotopy function H characterizing the logit Markov quantal response homotopy is constructed explicitly based on the system of equations given in [theorem 3](#), i.e.

$$\begin{aligned}\sigma_{s,i,a} &= \frac{e^{\lambda \cdot \tilde{u}_{s,i,a}(\boldsymbol{\sigma}_{s,-i}, \mathbf{V}_i)}}{\sum_{a' \in A_{s,i}} e^{\lambda \cdot \tilde{u}_{s,i,a'}(\boldsymbol{\sigma}_{s,-i}, \mathbf{V}_i)}}, \\ V_{s,i} &= \sum_{a' \in A_{s,i}} \sigma_{s,i,a'} \cdot \tilde{u}_{s,i,a'}(\boldsymbol{\sigma}_{s,-i}, \mathbf{V}_i)\end{aligned}$$

for all s, i and a .

There are two types of equations: strategy equations and state value equations. The state value equations are innocuous. The strategy equations, on the other hand, involve the exponential of the homotopy parameter λ . Since we would like to take λ towards infinity, the corresponding exponentials might become very large. Therefore, we follow [Turocy \(2005, 2010\)](#) and perform two transformations on the strategy equations. First, we introduce a reference action and, secondly, we take the logarithm.

For each state-player tuple (s, i) , we divide the strategy equations by the equation of reference action $a = 0$, leading to

$$\frac{\sigma_{s,i,a>0}}{\sigma_{s,i,0}} = e^{\lambda[\tilde{u}_{s,i,a}(\boldsymbol{\sigma}_{s,-i}, \mathbf{V}_i) - \tilde{u}_{s,i,0}(\boldsymbol{\sigma}_{s,-i}, \mathbf{V}_i)]}.$$

Taking the logarithm yields

$$\log(\sigma_{s,i,a>0}) - \log(\sigma_{s,i,0}) = \lambda[\tilde{u}_{s,i,a}(\boldsymbol{\sigma}_{s,-i}, \mathbf{V}_i) - \tilde{u}_{s,i,0}(\boldsymbol{\sigma}_{s,-i}, \mathbf{V}_i)].$$

Finally, applying the substitution $\beta_{s,i,a} := \log(\sigma_{s,i,a})$, but still writing $\sigma_{s,i,a} = e^{\beta_{s,i,a}}$ for better readability, we define the strategy and state value equations as follows.

$$\begin{aligned}H_{s,i,0}^\sigma(\boldsymbol{\beta}, \mathbf{V}, \lambda) &= 1 - \sum_{a \in A_{s,i}} \sigma_{s,i,a} \\ H_{s,i,a>0}^\sigma(\boldsymbol{\beta}, \mathbf{V}, \lambda) &= -[\beta_{s,i,a} - \beta_{s,i,0}] + \lambda[\tilde{u}_{s,i,a}(\boldsymbol{\sigma}_{s,-i}, \mathbf{V}_i) - \tilde{u}_{s,i,0}(\boldsymbol{\sigma}_{s,-i}, \mathbf{V}_i)] \\ H_{s,i}^V(\boldsymbol{\beta}, \mathbf{V}, \lambda) &= -V_{s,i} + \sum_{a \in A_{s,i}} \sigma_{s,i,a} \cdot \tilde{u}_{s,i,a}(\boldsymbol{\sigma}_{s,-i}, \mathbf{V}_i)\end{aligned}$$

Due to the normalization by reference action $a = 0$, the corresponding strategy equations have been replaced by an equation ensuring that action probabilities sum up to one.

The final homotopy function is obtained by stacking all components, i.e.

$$H(\boldsymbol{\beta}, \mathbf{V}, \lambda) = \begin{bmatrix} H^\sigma(\boldsymbol{\beta}, \mathbf{V}, \lambda) \\ H^V(\boldsymbol{\beta}, \mathbf{V}, \lambda) \end{bmatrix},$$

and given in terms of log-strategies $\boldsymbol{\beta}$.⁵

The components of the Jacobian matrix

$$J(\boldsymbol{\beta}, \mathbf{V}, \lambda) = \begin{bmatrix} \frac{\partial H_{s,i,a}^\sigma(\boldsymbol{\beta}, \mathbf{V}, \lambda)}{\partial \beta_{s',i',a'}}, & \frac{\partial H_{s,i,a}^\sigma(\boldsymbol{\beta}, \mathbf{V}, \lambda)}{\partial V_{s',i'}}, & \frac{\partial H_{s,i,a}^\sigma(\boldsymbol{\beta}, \mathbf{V}, \lambda)}{\partial \lambda} \\ \frac{\partial H_{s,i}^V(\boldsymbol{\beta}, \mathbf{V}, \lambda)}{\partial \beta_{s',i',a'}}, & \frac{\partial H_{s,i}^V(\boldsymbol{\beta}, \mathbf{V}, \lambda)}{\partial V_{s',i'}}, & \frac{\partial H_{s,i}^V(\boldsymbol{\beta}, \mathbf{V}, \lambda)}{\partial \lambda} \end{bmatrix}$$

are given as follows.

Partial derivatives of H^σ :

$$\frac{\partial H_{s,i,0}^\sigma(\boldsymbol{\beta}, \mathbf{V}, \lambda)}{\partial \beta_{s',i',a'}} = \begin{cases} -\sigma_{s,i,a'} & \text{if } s' = s \text{ and } i' = i, \\ 0 & \text{else,} \end{cases}$$

$$\frac{\partial H_{s,i,0}^\sigma(\boldsymbol{\beta}, \mathbf{V}, \lambda)}{\partial V_{s',i'}} = 0,$$

$$\frac{\partial H_{s,i,0}^\sigma(\boldsymbol{\beta}, \mathbf{V}, \lambda)}{\partial \lambda} = 0,$$

$$\frac{\partial H_{s,i,a>0}^\sigma(\boldsymbol{\beta}, \mathbf{V}, \lambda)}{\partial \beta_{s',i',a'}} = \begin{cases} 1 & \text{if } s' = s, i' = i, a' = 0, \\ -1 & \text{if } s' = s, i' = i, a' > 0, \\ \lambda \left[\frac{\partial \tilde{u}_{s,i,a}(\boldsymbol{\sigma}_{s,-i}, \mathbf{V}_i)}{\partial \beta_{s',i',a'}} - \frac{\partial \tilde{u}_{s,i,0}(\boldsymbol{\sigma}_{s,-i}, \mathbf{V}_i)}{\partial \beta_{s',i',a'}} \right] & \text{if } i' \neq i, \\ 0 & \text{else,} \end{cases}$$

⁵For computational purposes, it might be preferable to state the homotopy function in terms of tensor products. The corresponding derivation is done in [Eibelschäuser and Poensgen \(2019\)](#).

$$\frac{\partial H_{s,i,a>0}^\sigma(\beta, \mathbf{V}, \lambda)}{\partial V_{s',i'}} = \lambda \left[\frac{\partial \tilde{u}_{s,i,a}(\sigma_{s,-i}, \mathbf{V}_i)}{\partial V_{s',i'}} - \frac{\partial \tilde{u}_{s,i,0}(\sigma_{s,-i}, \mathbf{V}_i)}{\partial V_{s',i'}} \right],$$

$$\frac{\partial H_{s,i,a>0}^\sigma(\beta, \mathbf{V}, \lambda)}{\partial \lambda} = \tilde{u}_{s,i,a}(\sigma_{s,-i}) - \tilde{u}_{s,i,0}(\sigma_{s,-i}).$$

Partial derivatives of H^V :

$$\frac{\partial H_{s,i}^V(\beta, \mathbf{V}, \lambda)}{\partial \beta_{s',i',a'}} = \begin{cases} \sigma_{s,i,a'} \cdot \tilde{u}_{s,i,a'}(\sigma_{s,-i}, \mathbf{V}_i) & \text{if } s' = s, i' = i, \\ \sum_{a'' \in A_{s,i}} \sigma_{s,i,a''} \cdot \frac{\partial \tilde{u}_{s,i,a''}(\sigma_{s,-i}, \mathbf{V}_i)}{\partial \beta_{s,i',a'}} & \text{if } s' = s, i' \neq i, \\ 0 & \text{else,} \end{cases}$$

$$\frac{\partial H_{s,i}^V(\beta, \mathbf{V}, \lambda)}{\partial V_{s,i}} = \begin{cases} -1 + \sum_{a'' \in A_{s,i}} \sigma_{s,i,a''} \cdot \frac{\partial \tilde{u}_{s,i,a''}(\sigma_{s,-i}, \mathbf{V}_i)}{\partial V_{s,i}} & \text{if } s' = s, i' = i, \\ \sum_{a'' \in A_{s,i}} \sigma_{s,i,a''} \cdot \frac{\partial \tilde{u}_{s,i,a''}(\sigma_{s,-i}, \mathbf{V}_i)}{\partial V_{s',i}} & \text{if } s' \neq s, i' = i, \\ 0 & \text{else,} \end{cases}$$

$$\frac{\partial H_{s,i}^V(\beta, \mathbf{V}, \lambda)}{\partial \lambda} = 0.$$

Underlying partial derivatives of expected payoffs:

$$\frac{\partial \tilde{u}_{s,i,a}(\sigma_{s,-i}, \mathbf{V}_i)}{\partial \beta_{s',i',a'}} = \begin{cases} \sum_{a_{s,-i} \in A_{s,-i}} \prod_{\substack{i'' \in I \\ i'' \neq i}} \sigma_{s,i'',a_{s,i''}} \cdot \tilde{u}_{s,i,a}(\mathbf{a}_{s,-i}, \mathbf{V}_i) & \text{if } s' = s, i' \neq i, \\ 0 & \text{else,} \end{cases}$$

$$\frac{\partial \tilde{u}_{s,i,a}(\sigma_{s,-i}, \mathbf{V}_i)}{\partial V_{s',i'}} = \begin{cases} \sum_{a_{s,-i} \in A_{s,-i}} \prod_{\substack{i'' \in I \\ i'' \neq i}} \sigma_{s,i'',a_{s,i''}} \cdot \delta_i \cdot \phi_{s \rightarrow s'}(a, \mathbf{a}_{s,-i}) & \text{for } i' = i, \\ 0 & \text{else.} \end{cases}$$