Computing Sequential Equilibria Using Agent Quantal Response Equilibria

Theodore L. Turocy
Department of Economics
Texas A&M University
College Station TX 77843
`turocy@econmail.tamu.edu`

August 26, 2008

Abstract

The limit of any convergent sequence of agent quantal response equilibria is a sequential equilibrium of an extensive game. Using a logarithmic transformation of action probabilities, it is numerically feasible and practical to compute such sequences, and thereby compute good approximations to sequential equilibrium assessments. This paper describes the algorithm to compute the sequences, and outlines the convergence and selection properties of the method.

Keywords: computing Nash equilibrium, quantal response, homotopy methods.

1 Introduction

In Turocy [13] it was shown that the logit quantal response equilibrium (QRE) correspondence, as defined for strategic games by McKelvey and Palfrey [9], has a natural game-theoretic interpretation in terms of the replicator dynamics. The limit, as the precision parameter λ tends to infinity, of any branch of the QRE correspondence is a Nash equilibrium of the strategic game. The

Jacobian of the system of equations defining a branch of the correspondence is bounded, indicating that computing a Nash equilibrium as the limit of a sequence of approximating QREs is numerically feasible. This result is extended here to the task of computing a sequential equilibrium (Kreps and Wilson [5]) of a finite extensive game with perfect recall.

McKelvey and Palfrey [10] define the agent quantal response equilibrium concept for finite games with perfect recall, and show that any limit point of the AQRE correspondence is a sequential equilibrium. The logit specification of AQRE is the most widely-used and computationally convenient. The formulation used in Turocy [13] cannot be applied directly to extensive games, because the elements of the Jacobian matrix need not be bounded for information sets which are reached with small probability. Transforming the system by using logarithms of probabilities results in a bounded Jacobian which can be used in a numerical procedure to trace a branch of AQREs. In addition, the logarithmic transformation results in a straightforward calculation of beliefs at information sets which are not reached in the limiting sequential equilibrium. Therefore, the full sequential equilibrum assessment can be approximately arbitrarily well using this method.

The paper is organized as follows. Section 2 presents the calculations underlying the tracing of a branch of the logit AQRE correspondence. Section 3 illustrates the necessity and practicality of the logarithm transformation of probabilities, and presents some quantitative speed-of-convergence results for selected applications. Section 4 concludes by placing this method within the the literature on equilibrium computation.

2 Tracing the correspondence

2.1 Notation and definition of the path-following problem

The procedure operates on finite extensive games with perfect recall. There are a finite number of players; each player has a finite number of information sets; there are a finite number of actions at each information set. To each terminal node is attached a vector of payoffs, one for each player, and players are expected utility maximizers. Some information sets may belong to the chance player, at which the probabilities of each action are prespecified.

The letters a, b, and c will be used to denote actions. Each action is associated with exactly one information set, denoted $I(a)$. Nodes will typically be denoted by n and m. Information sets are sets of nodes; the notation $n \in h$ means that node n is a member of information set h, and the set of actions available at information set h is $A(h)$. The set of nodes at which action a may be taken as $N(a) \equiv \{n : n \in I(a)\}$. Because the game is perfect recall, any action appears at most once along any path of play. The relation $a \prec n$ indicates that action a precedes node n in the play of the game.

A behavior strategy profile π specifies, for each action a, the probability π_a that the action is played when its information set is reached. Let $u_a(\pi)$ denote the expected payoff to playing action a, conditional on reaching its information set $I(a)$, assuming the behavior strategy profile π is played at all other information sets. Then, a strategy profile π is a logit agent quantal response equilibrium if it satisfies

$$\pi_a = \frac{e^{\lambda u_a(\pi)}}{\sum_{b \in I(a)} e^{\lambda u_b(\pi)}} \tag{1}$$

for all actions a at all information sets for all players. The set of logit AQRE is a correspondence mapping $\lambda \in [0, \infty)$ into the set of (totally mixed) behavior profiles. McKelvey and Palfrey [10] show that the limiting points in this correspondence as $\lambda \to \infty$ form a subset of the set of sequential equilibria. Therefore, tracing a branch of the correspondence amounts to generating (part of) a sequence of totally mixed behavior profiles required for the consistency requirement of sequential equilibrium.

For each information set h, choose some reference action $a \in h$. Then (1) implies that, for all $b \in h$, $b \neq a$,

$$\frac{\pi_a}{\pi_b} = \mathrm{e}^{\lambda[u_a(\pi) - u_b(\pi)]},$$

or equivalently,

$$\ln \pi_a - \ln \pi_b = \lambda[u_a(\pi) - u_b(\pi)].$$

To maintain the normalization of probabilities, each information set s generates a sum-to-one equation

$$\sum_{a \in s} \pi_a = 1.$$

Therefore, the system is defined by (total number of actions - number of information sets) equations of the form

$$H_{ab}(\pi, \lambda) = \ln \pi_a - \ln \pi_b - \lambda \left[u_a(\pi) - u_b(\pi) \right] = 0 \tag{2}$$

plus one equation of the form

$$H_h(\pi, \lambda) = \sum_{a \in h} \pi_a - 1 = 0 \tag{3}$$

for each information set h.

The form of (2) suggests an approach in which profiles are represented according to the logarithm of their probabilities. Introducing the transformation $p_a = \ln \pi_a$ for all actions a, (2) and (3) become

$$H_{ab}(p, \lambda) = p_a - p_b - \lambda \left[u_a(\mathrm{e}^p) - u_b(\mathrm{e}^p) \right] = 0 \tag{4}$$

and

$$H_h(p, \lambda) = \sum_{a \in h} \mathrm{e}^{p_a} \tag{5}$$

respectively, where the notation e^p represents the vector obtained by exponentiating each element of the vector p individually. In the following derivations, π will be written in place of e^p and π_c in place of e^{p_c} for expositional clarity.

Tracing the zeroes of the system of equations defined by (4) and (5) can be done using a predictor-corrector method (e.g., Allgower and Georg [1]). The predictor step uses the Jacobian of the system with respect to (p, λ). For the ratio equations H_{ab}, the Jacobian entries are:

$$\begin{aligned} \frac{\partial H_{ab}}{\partial p_a} &= 1 \\ \frac{\partial H_{ab}}{\partial p_b} &= -1 \\ \forall c \in I(a) - \{a, b\}, \frac{\partial H_{ab}}{\partial p_c} &= 0 \end{aligned}$$

$$\begin{aligned}\forall c \notin I(a), \frac{\partial H_{ab}}{\partial p_c} &= -\lambda \pi_c \left[\frac{\partial u_a(\pi)}{\partial \pi_c} - \frac{\partial u_b(\pi)}{\partial \pi_c} \right] \\ \frac{\partial H_{ab}}{\partial \lambda} &= -\left[u_a(\pi) - u_b(\pi) \right].\end{aligned}$$

The Jacobian entries for the sum-to-one equations are:

$$\begin{aligned}\forall a \in h, \frac{\partial H_h}{\partial p_a} &= \pi_a \\ \forall a \notin h, \frac{\partial H_h}{\partial p_a} &= 0 \\ \frac{\partial H_h}{\partial \lambda} &= 0.\end{aligned}$$

In all cases, the Jacobian is evaluated at a point (p, λ) forming an agent logit quantal response equilibrium.

2.2 Numerical feasibility

In order for tracing the zeroes of the system H to be numerically feasible, the Jacobian of the system must be bounded.

Proposition 1. The Jacobian of H is bounded when the system is expressed in terms of log probabilities.

Proof. The only entries for which boundedness is not straightforward is $\frac{\partial H_{ab}}{\partial p_c}$ for actions $c \notin I(a)$, since these contain terms of the form $\pi_c \frac{\partial u_a(\pi)}{\partial \pi_c}$. It is now shown that this quantity is bounded. Let $P_n(\pi)$ be the probability that a node n is reached when strategy profile π is played,

$$P_n(\pi) = \prod_{a \prec n} \pi_a.$$

The probability information set h is reached can then be written

$$P_h(\pi) = \sum_{n \in h} P_h(\pi).$$

Let $u_{a|n}(\pi)$ be the payoff to playing action a, conditional on node n being reached. By definition,

$$u_a(\pi) = \sum_{n \in N(a)} \frac{P_n(\pi)}{P_h(\pi)} u_{a|n}(\pi).$$

Differentiating with respect to the probability π_c of action $c \notin I(a)$ gives

$$\frac{\partial u_a(\pi)}{\partial \pi_c} = \sum_{n \in N(a)} \frac{\partial}{\partial \pi_c} \left[\frac{P_n(\pi)}{P_h(\pi)} \right] u_{a|n}(\pi) + \frac{P_n(\pi)}{P_h(\pi)} \cdot \frac{\partial u_{a|n}(\pi)}{\partial \pi_c}. \tag{6}$$

In the first term in the sum in (6), the change in beliefs given a change in π_c is

$$\begin{aligned} \frac{\partial}{\partial \pi_c} \left[\frac{P_n(\pi)}{P_h(\pi)} \right] &= \frac{P_h(\pi) \frac{\partial P_n(\pi)}{\partial \pi_c} - P_n(\pi) \frac{\partial P_h(\pi)}{\partial \pi_c}}{P_h(\pi)^2} \\ &= \frac{\frac{\partial P_n(\pi)}{\partial \pi_c}}{P_h(\pi)} - \frac{P_n(\pi)}{P_h(\pi)} \cdot \frac{\frac{\partial P_h(\pi)}{\partial \pi_c}}{P_h(\pi)} \\ &= \frac{\frac{\partial P_n(\pi)}{\partial \pi_c}}{P_h(\pi)} - \frac{P_n(\pi)}{P_h(\pi)} \cdot \sum_{m \in N(a)} \frac{\frac{\partial P_m(\pi)}{\partial \pi_c}}{P_h(\pi)}. \end{aligned} \tag{7}$$

Note that

$$\frac{\partial P_n(\pi)}{\partial \pi_c} = \mathbf{1}_{c \prec n} \frac{P_n(\pi)}{\pi_c} \tag{8}$$

and

$$\frac{\partial P_h(\pi)}{\partial \pi_c} = \sum_{n \in h} \mathbf{1}_{c \prec n} \frac{P_n(\pi)}{\pi_c}, \tag{9}$$

where $\mathbf{1}_{c \prec n}$ is the indicator function taking on the value of 1 when the action c precedes the node n in the game tree, and 0 otherwise. Taking equation (6) and substituting in (7),

$$\begin{aligned} \frac{\partial u_a(\pi)}{\partial \pi_c} &= \sum_{n \in N(a)} \left(\frac{\frac{\partial P_n(\pi)}{\partial \pi_c}}{P_h(\pi)} - \frac{P_n(\pi)}{P_h(\pi)} \cdot \sum_{m \in N(a)} \frac{\frac{\partial P_m(\pi)}{\partial \pi_c}}{P_h(\pi)} \right) u_{a|n}(\pi) + \frac{P_n(\pi)}{P_h(\pi)} \cdot \frac{\partial u_{a|n}(\pi)}{\partial \pi_c} \\ &= \sum_{n \in N(a)} \left[\frac{\frac{\partial P_n(\pi)}{\partial \pi_c}}{P_h(\pi)} u_{a|n}(\pi) - \left(\sum_{m \in N(a)} \frac{\frac{\partial P_m(\pi)}{\partial \pi_c}}{P_h(\pi)} \right) u_{a|n}(\pi) + \frac{P_n(\pi)}{P_h(\pi)} \cdot \frac{\partial u_{a|n}(\pi)}{\partial \pi_c} \right] \\ &= \sum_{n \in N(a)} \frac{\mathbf{1}_{c \prec n} P_n(\pi)}{\pi_c P_h(\pi)} \left[u_{a|n}(\pi) - u_a(\pi) \right] + \sum_{n \in N(a)} \frac{P_n(\pi)}{P_h(\pi)} \cdot \frac{\partial u_{a|n}(\pi)}{\partial \pi_c} \end{aligned}$$

$$\pi_c \frac{\partial u_a(\pi)}{\partial \pi_c} = \sum_{n\in N(a)} \mathbf{1}_{c\prec n} \frac{P_n(\pi)}{P_h(\pi)} \left[u_{a|n}(\pi) - u_a(\pi)\right] + \sum_{n\in N(a)} \pi_c \frac{P_n(\pi)}{P_h(\pi)} \cdot \frac{\partial u_{a|n}(\pi)}{\partial \pi_c}.$$

It is immediate that the right-hand side of this last expression is bounded. **QED**

An implication of the calculation in the proof is that $\frac{\partial u_a(\pi)}{\partial \pi_c}$ need not be bounded. Tracing the AQRE correspondence using untransformed probabilities may be numerically infeasible in some games. The necessity of some transformation is illustration in Section 3.1, which shows that for "Selten's horse" (Selten [12]) the Jacobian of the homotopy system is not bounded using untransformed probabilities. The logarithmic transformation is not the only transformation which would achieve boundedness of the Jacobian of the transformed system. As is shown next, the logarithm is the natural choice given the form belief computations take.

2.3 Computing beliefs

Since the sequential equilibrium refinement is valuable in that it can eliminate certain types of behaviors off the equilibrium path, approximating sequential equilibria by AQRE requires computing beliefs at information sets s for which $P_h(\pi) \to 0$ along the branch of the logit correspondence. For this, the logarithmic implementation of action probabilities is convenient. Define $m \in \arg\max_{n\in h} P_n(\pi)$; this is a node which is reached with maximal probability, among nodes in the information set h. Observe that $\frac{P_m(\pi)}{P_h(\pi)} > |N(s)|^{-1}$. Then the beliefs can be written

$$\begin{aligned}
\frac{P_n(\pi)}{P_h(\pi)} &= \frac{P_n(\pi)}{P_m(\pi)} \cdot \frac{P_m(\pi)}{P_h(\pi)} \\
&= \frac{P_n(\pi)}{P_m(\pi)} \cdot \frac{P_m(\pi)}{\sum_{m'\in I(n)} P_{m'}(\pi)} \\
&= \frac{P_n(\pi)}{P_m(\pi)} \cdot \left(1 + \sum_{m'\in I(n), m'\neq m} \frac{P_{m'}(\pi)}{P_m(\pi)}\right)^{-1}. \qquad (10)
\end{aligned}$$

Recalling that

$$P_n(\pi) = \prod_{a\prec n} \pi_a,$$

it follows that

$$\log P_n(\pi) = \sum_{a \prec n} \log \pi_a$$

or

$$\frac{P_n(\pi)}{P_m(\pi)} = \exp\left[\log P_n(\pi) - \log P_m(\pi)\right] = \exp\left[\sum_{a \prec n} \log \pi_a - \sum_{a \prec m} \log \pi_a\right]. \tag{11}$$

In (10), since m was chosen to be the node reached with maximal probability in the information set, $P_n(\pi) \leq P_m(\pi)$ and $P_{m'}(\pi) \leq P_m(\pi)$. By construction, all ratios in (10) are less than or equal to one, and so (10) can be implemented numerically even when $P_h(\pi)$ is small.

3 Numerical examples

The method described in Section 2 has been implemented in the software package Gambit [7]. This section characterizes the empirical performance of the algorithm. The tracing process requires as an initial condition a point (p, λ) satisfying equations (4) and (5). For all games, the point $\lambda = 0$ with uniform randomization at all information sets satisfies these equations. In this section, results are presented for tracing this "principal branch" of the correspondence.

3.1 Necessity of a transformation: Selten's horse

The first example demonstrates that there are games in which tracing the AQRE correspondence is numerically infeasible without some transformation of the probabilities. Figure 1 shows the extensive form of a game commonly called "Selten's horse," from Selten [12]. Figure 2 plots the probabilities each player plays their strategy L in an AQRE as a function of λ. In the limiting sequential equilibrium, Player I and Player II both choose R. Along the sequence of AQREs, both players choose L with nontrivial probability; in fact, Player II's probability of choosing L is initially increasing in λ. Figure 3 plots Player III's belief μ placed on the left note in his information set as a function of λ. The limiting sequential equilibrium assessment selected by the AQRE correspondence has $\mu = \frac{1}{3}$, and the AQRE values of μ are well-behaved as a function of λ.

In this game, the change in conditional payoffs $\frac{\partial u_a(\pi)}{\partial \pi_c}$ for the two actions belonging to Player III is unbounded along the AQRE correspondence; therefore, without some suitable transformation, tracing the correspondence would fail. As in Figure 1, let ε be the probability Player I chooses L, and δ be the probability Player II chooses L. Since $\varepsilon > 0$ and $\delta > 0$ in an AQRE, Player III's beliefs about at the left node in his information set must be

$$\mu = \frac{\varepsilon}{\varepsilon + (1-\varepsilon)\delta}.$$

Differentiating this with respect to δ gives

$$\frac{\partial \mu}{\partial \delta} = \frac{-\varepsilon(1-\varepsilon)}{[\varepsilon + (1-\varepsilon)\delta)]^2}.$$

Taking the limit as δ gets small,

$$\lim_{\delta \to 0} \frac{\partial \mu}{\partial \delta} = \frac{\varepsilon(1-\varepsilon)}{\varepsilon^2} = \frac{1-\varepsilon}{\varepsilon}.$$

Therefore, when Player I and Player II both rarely play L, then Player III's beliefs μ are arbitrarily sensitive to changes in δ, and therefore Player III's expected payoff to his actions is arbitrarily sensitive to changes in δ. This implies that entries in the Jacobian of the system using untransformed probabilities is unbounded, and, when δ and ε became small enough, entries in the Jacobian would exceed the maximum representable floating-point number. On the other hand, if Player II's choice probability is represented by $\Delta = \log \delta$, the beliefs are

$$\mu = \frac{\varepsilon}{\varepsilon + (1-\varepsilon)\mathrm{e}^{\Delta}}.$$

Differentiation with respect to Δ gives

$$\frac{\partial \mu}{\partial \Delta} = \frac{\varepsilon(1-\varepsilon)e^{\Delta}}{[\varepsilon + (1-\varepsilon)e^{\Delta}]^2} = \frac{\varepsilon(1-\varepsilon)\delta}{[\varepsilon + (1-\varepsilon)\delta]^2},$$

and so $\lim_{\delta \to 0} \frac{\partial \mu}{\partial \Delta} = 0$.

3.2 Usefuless of a transformation: The centipede game

A feature of the logarithmic transformation of probabilities is that it changes the domain on which the path-following problem is expressed from one in which probabilities must lie in $(0, 1)$ to one in which log-probabilities must lie in $(-\infty, 0)$. Removing the lower bound offers a practical benefit when using a predictor-corrector method. The predictor step uses a linear extrapolation based on the Jacobian of the system. However, in the logit specification, the probability a strictly inferior action is chosen decreases exponentially in λ. In the untransformed system, a strictly inferior action being played with small probability at some λ may be extrapolated to be played with a negative probability at the next step λ', which is outside the domain on which the untransformed Jacobian is valid.

As the previous example shows, simply setting small probabilities to zero is not adequate, since it is not the case that small probabilities are negligible in QRE. This is not just true as a limiting phenomenon. Actions played with positive probability in the limiting QRE may be played with arbitrarily small probability at some point along the branch. An application in which this occurs is the centipede game, discussed in McKelvey and Palfrey [10], and illustrated in Figure 4. The centipede game is a two-player, alternating-moves game of perfect information. Initially, the "pot" is set to some amount p. At each move, the player whose turn it is can either take the pot, or pass. If he takes, he earns some fraction $\theta \in (.5, 1]$ of the pot, and the other player receives the rest, $1 - \theta$. If he passes, the pot is multiplied by $m > 1$ and play passes to the other player. The game is finite; after the Nth stage, if the player whose turn it is passes, then at the $N + 1$st stage the other player automatically takes. The top panel in Figure 4 shows the game tree for parameters $p = .5$, $\theta = .8$, $m = 2$, and $N = 6$, which is the subject of experiments reported in McKelvey and Palfrey [8]. In the AQRE correspondence of this game, the probabilities that actions are taken do not converge monotonically to the unique subgame perfect equilibrium in which players take with probability one at each node, a fact used in McKelvey and Palfrey [10] to fit the QRE model to the experimental data.

It is possible to drive take probabilities arbitrarily close to zero in AQRE of games of this class.

Observe that the unique subgame perfect equilibrium remains always to take at every opportunity so long as $m < 4$, holding fixed the other parameters. As $m \to 4$, however, it requires only a small probability that the opponent will err and pass on his next move for passing to become an optimal choice. As a result, passing becomes an attractive strategy until λ becomes large. The bottom panel of Figure 4 shows the centipede game with parameter $m = 3.9$, and Figure 5 plots the AQRE take probabilities for the first player at each of his three turns in this game. For a substantial range of λ, the probability of passing at each of the first two turns is essentially unity. For instance, the probability of taking on the first turn is on the order of 10^{-80} for $\lambda = 33$, and of taking on the second term around 10^{-80} for $\lambda = 2.2$.

3.3 Speed of convergence: Some signaling games

McKelvey and Palfrey also investigate the usefulness of the AQRE concept as an equilibrium selection criterion, using as examples a set of signaling games from Banks, Camerer, and Porter [2] (BCP, games 2, 3, and 4) and Brandts and Holt [3] (BH, games 3 and 4). These games are reproduced in Tables 1 and 2, using the terminology that the informed player is the "sender" and the uninformed player the "receiver." McKelvey and Palfrey [10] graph the AQRE correspondence computed using a grid search method which became infeasible at relatively small values of λ. This limitation is visible as an artifact in their Figures 5 through 7, in which the curves stop short of the limiting pure-strategy equilibria. For the purposes of the maximum likelihood estimation done there, the portion of the correspondence computed was adequate, and it was evident to which equilibrium the branch limits. The path-following method permits precise calculation of the limiting assessment, including beliefs for off-path information sets.

To augment the analysis, the convergence behavior of the algorithm is analyzed for these games. To measure convergence, the Lyapunov function (McKelvey [6]) is used to measure the progress of the tracing towards the limiting equilibrium. Briefly, the Lyapunov function is the sum of the squares of the regrets across all information sets in the game, where the regret is the amount of foregone payoff at each information set relative to the best reply. Figure 6 presents the $\log_{10}$ of the Lyapunov value as a function of the number of steps taken. This figure illustrates three modes of

m_1	a_1	a_2	a_3	m_2	a_1	a_2	a_3	m_3	a_1	a_2	a_3
					BCP2						
t_1	1,2	2,1	0,3	t_1	1,2	1,1	2,1	t_1	3,1	0,0	2,1
t_2	2,2	1,4	3,2	t_2	2,2	0,4	3,1	t_2	2,2	0,0	2,1
					BCP3						
t_1	0,3	2,2	2,1	t_1	1,2	2,1	3,0	t_1	1,6	4,1	2,0
t_2	1,0	3,2	2,1	t_2	0,1	3,1	2,6	t_2	0,0	4,1	0,6
					BCP4						
t_1	4,0	0,3	0,4	t_1	2,0	0,3	3,2	t_1	2,3	1,0	1,2
t_2	3,4	3,3	1,0	t_2	0,3	0,0	2,2	t_2	4,3	0,4	3,0

Table 1: Signaling games from Banks, Camerer, and Porter. t_i are sender types; m_j are messages, and a_k are receiver actions.

$m = I$	C	D	E	$m = S$	C	D	E
				BH3			
A	45,30	15,0	30,15	A	30,90	0,15	45,15
B	30,30	0,45	30,15	B	45,0	15,30	30,15
				BH4			
A	30,30	0,0	50,35	A	45,90	15,15	100,30
B	30,30	30,45	30,0	B	45,0	0,30	0,15

Table 2: Signaling games from Brandts and Holt. A and B are sender types, I and S are messages, and C, D, and E are receiver actions.

convergence behavior, which depend on the characteristics of the limiting equilibrium.

In all five games, both sender types are equally likely, and the equilibria of interest are pooling equilibria. The limiting equilibrium selected by the principal branch in each of games BCP3, BH3, and BH4 is a strict pure-strategy equilibrium. In BCP3, senders send m_1 in the equilibrium selected by AQRE, and receivers place probability one that the sender's type is t_1 if m_2 or m_3 were to be sent. In BH3, senders send $m = I$ in the equilibrium selected by AQRE, and receivers place probability one that the sender's type is B if $m = S$ were to be sent. In BH4, senders send $m = S$ in the equilibrium selected by AQRE, and receivers place probability one that the sender's type is B if $m = I$ were to be sent. Since there is a strict best reply at each information set, the probability with which inferior actions are played decays exponentially in λ. With the logarithmic specification, the log-probabilities decay linearly in λ, so linear extrapolation based on the Jacobian of the transformed system is very accurate, and so the AQRE tracing method converges quickly.

The limiting equilibrium in game BCP4 has senders choosing message m_3. Off the equilibrium path, the receiver randomizes between actions after seeing message m_1, playing C with probability $\frac{1}{4}$ and D with probability $\frac{3}{4}$; randomization is supported by the belief that $P(t_1|m_1) = \frac{1}{4}$. After m_2, the receiver places probability one that the sender was type t_1. Convergence to an equilibrium with randomization is slower, both in terms of the Lyapunov function and in terms of the computed probabilities. The probabilities in the limit at this information set put a weight of $\frac{1}{4}$ on action 1 and $\frac{3}{4}$ on action 2; the AQRE probabilities become correct to 6 digits only after $\lambda \approx 5 \times 10^5$, or about 4 steps before termination at $\lambda = 10^6$. The Lyapunov function decreases more slowly because, along the AQRE component, one of the actions used in the limiting equilibrium is inferior, but is played with a significant probability.

An intermediate case is represented by game BCP2. The equilibrium message is m_3, but the off-path beliefs after both m_1 and m_2 both place equal weight on types t_1 and t_2. With this belief, receivers are indifferent between responses a_2 and a_3 when m_1 is sent. In the selected equilibrium, a_2 is chosen with probability one after both m_1 and m_2. The selected equilibrium is pure but non-strict, and convergence is slower in both the strategy and payoff spaces than when the limiting equilibrium is strict.

This game illustrates a conceptual difference in using logit AQRE as a selection criterion in signaling games versus the traditional literature on refinements. Although there are strict sequential equilibria which have the same behavior strategy profile as the selected equilibrium, logit AQRE selects the assessment with where the receiver does not change his beliefs about the sender's type after observing an off-path signal. Similarly, note that the sequential equilibrium assessment computed for Selten's horse set Player III's beliefs at $\mu = \frac{1}{3}$, the largest value for which Player III's choice of L is a best response.

Evaluating the method in terms of the quality of approximation per number of steps taken is preferred here to timings, because the time per step is closely tied to the quality of the implementation in the computer code. For these games, though, the current Gambit implementation is fast: running time is well under one second on any current (mid-2008) desktop computer.

4 Conclusion

This paper demonstrates the feasibility of computing a sequential Nash equilibrium in an extensive game by constructing a sequence of agent logit quantal response equilibria. Expressing the tracing problem in terms of transformed probabilities is necessary for this procedure to be numerically feasible, and a logarithmic transformation of probabilities is demonstrated to be useful in practice. This is the first implementation of a globally convergent method for computing a sequential equilibrium assessment for any extensive game of perfect recall. The quality of the approximation to the limiting equilibrium measured in the payoff space is good within a small number of steps on a sample of games drawn from McKelvey and Palfrey's original paper introducing AQRE.

This method complements the recent work of MILTERSEN AND SORENSEN [11] in two-player games. They extend the sequence form approach of KOLLER, MEGIDDO, AND VON STENGEL [4] to computing a quasi-perfect equilibrium, which refines the sequential equilibrium concept. In constant-sum games, their method can be expressed in terms of a symbolically perturbed linear program, and can be solved in polynomial time. In general two-player games, the resulting symbolically perturbed linear complementarity program is not polynomial time, but in practice solving such programs is often efficient.

While direct comparisons between the method of this paper and that of Miltersen and Sorensen await mature implementations in code, experience from the related sequence form-based algorithms for suggests that the Miltersen-Sorensen method will in the long run be the recommended choice for computing a sample sequential equilibrium in two-player games. Still, the limiting sequential equilibria of the AQRE correspondence may not be the same as those computed by Miltersen and Sorensen's method in some games, so even when working with two-player games, the two methods will likely be complementary in exploring. questions such as identifying whether a game has multiple sequential equilibria.

Even for applications in which the fact that the limiting equilibrium is sequential is not important, computing directly on the extensive game has a practical advantage. For this algorithm the relevant size of the game is the total number of actions or strategies, since this determines the dimensionality of the system of equations as well as the computational time for expected payoffs. Typically, the reduced strategic form of a game has many more strategies than the extensive game has actions, meaning that the running times for AQRE tracing will be much faster when the game has a nontrivial extensive structure. Since the equilibria selected by QRE on the extensive and strategic representations of the same game may, and often do, differ, having both methods again serves as a numerical approach to establishing multiplicity of Nash equilibria.

References

[1] E. L. Allgower and K. Georg. *Numerical Continuation Methods: An Introduction.* Springer-Verlag, Berlin, 1990.

[2] J. Banks, C. Camerer, and D. Porter. Experimental tests of Nash refinements in signaling games. *Games and Economic Behavior*, 4:1–31, 1992.

[3] J. Brandts and C.A. Holt. Adjustment patterns and equilibrium selection in experimental signaling games. *International Journal of Game Theory*, 22:279–302, 1993.

[4] Daphne Koller, Nimrod Megiddo, and Bernhard von Stengel. Efficient computation of equilibria for extensive two-person games. *Games and Economic Behavior*, 14:247–259, 1996.

[5] David Kreps and Robert Wilson. Sequential equilibrium. *Econometrica*, 50:863–894, 1982.

[6] Richard D. McKelvey. A Liapunov function for Nash equilibria. Caltech Social Science working paper, 1991.

[7] Richard D. McKelvey, Andrew M. McLennan, and Theodore L. Turocy. Gambit: Software Tools for Game Theory. http://gambit.sourceforge.net.

[8] Richard D. McKelvey and Thomas R. Palfrey. An experimental study of the cenitpede game. *Econometrica,* 60: 803-836, 1992.

[9] Richard D. McKelvey and Thomas R. Palfrey. Quantal response equilibria for normal form games. *Games and Economic Behavior*, 10:6–38, 1995.

[10] Richard D. McKelvey and Thomas R. Palfrey. Quantal response equilibria for extensive form games. *Experimental Economics*, 1:9–41, 1998.

[11] Peter Bro Miltersen and Troels Bjerre Sorensen. Computing a quasi-perfect equilibrium of a two-player game. Working paper, 6 October 2006.

[12] Reinhard Selten. Reexamination of the perfectness concept for equilibrium points in extensive games. *International Journal of Game Theory*, 4:25–55, 1975.

[13] Theodore L. Turocy. A dynamic homotopy interpretation of the logistic quantal response equilibrium correspondence. *Games and Economic Behavior*, 51:243–263, 2005.

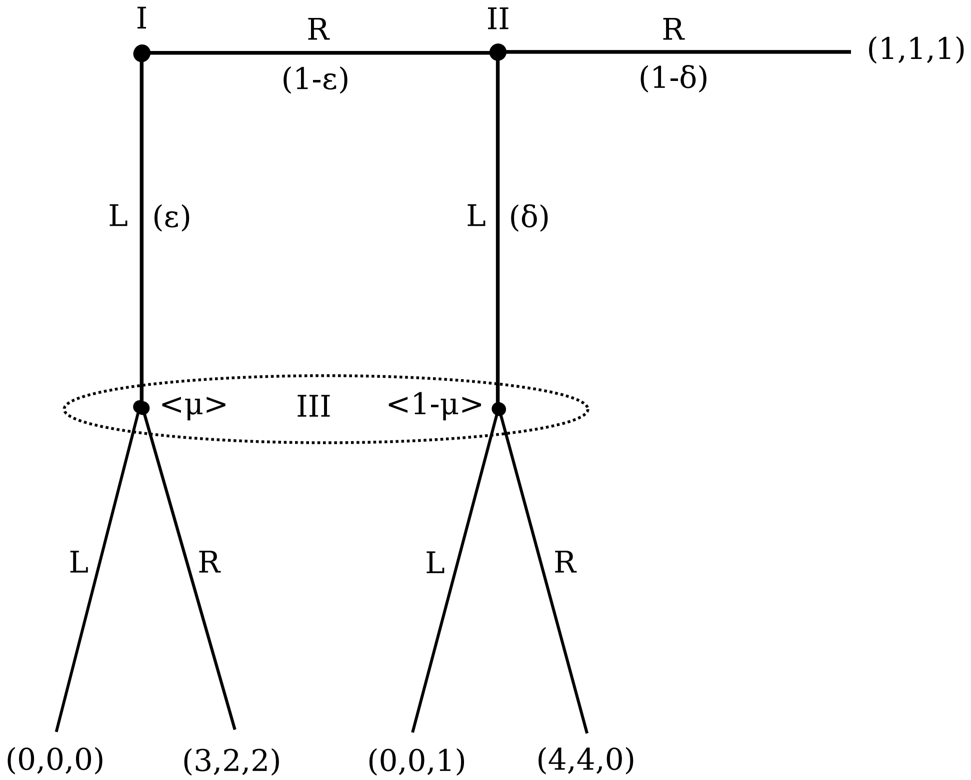

Figure 1: "Selten's horse," an example for which a transformation of probabilities is necessary for tracing logit AQRE to be numerically feasible.

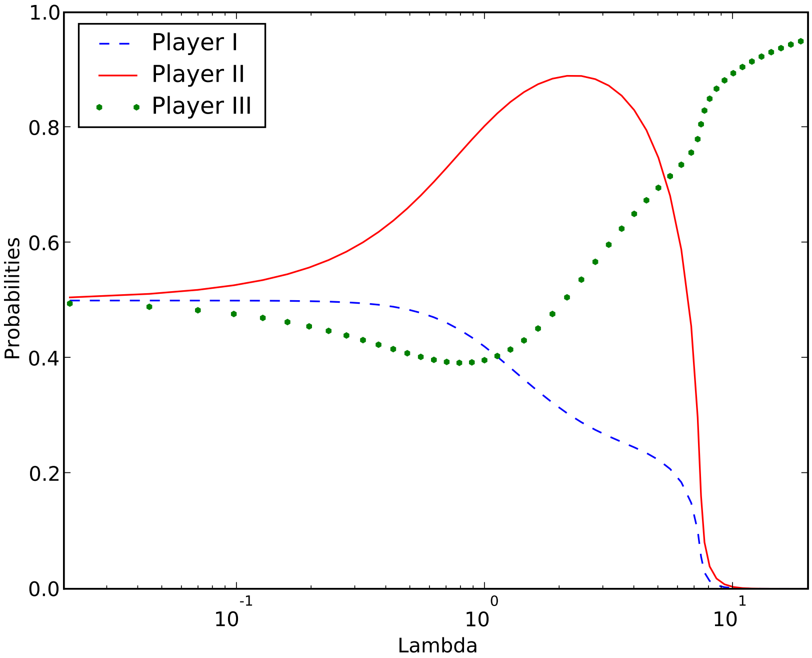

Figure 2: AQRE probabilities of choosing L for each player as a function of λ for Selten's horse.

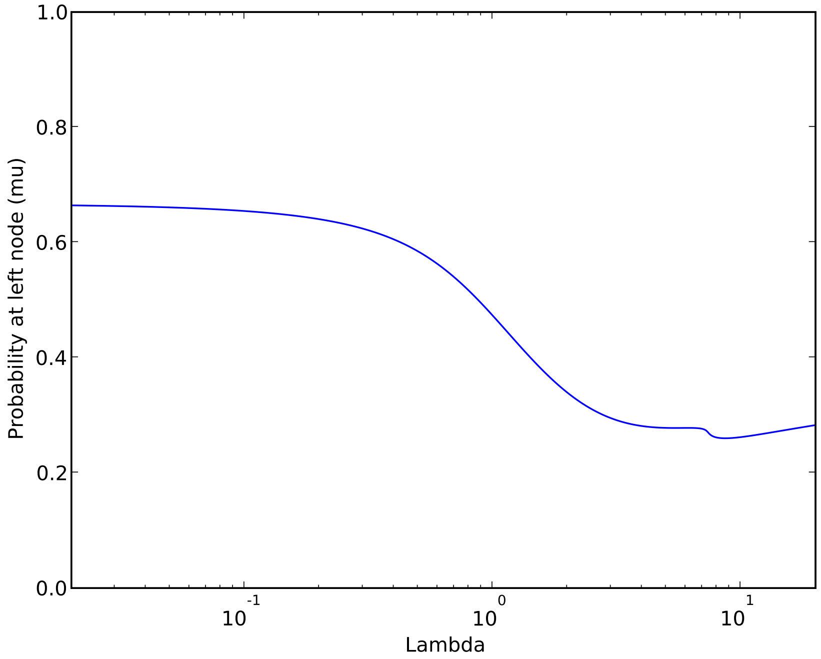

Figure 3: Belief $\mu$of Player III in AQRE as a function of λ for Selten's horse.

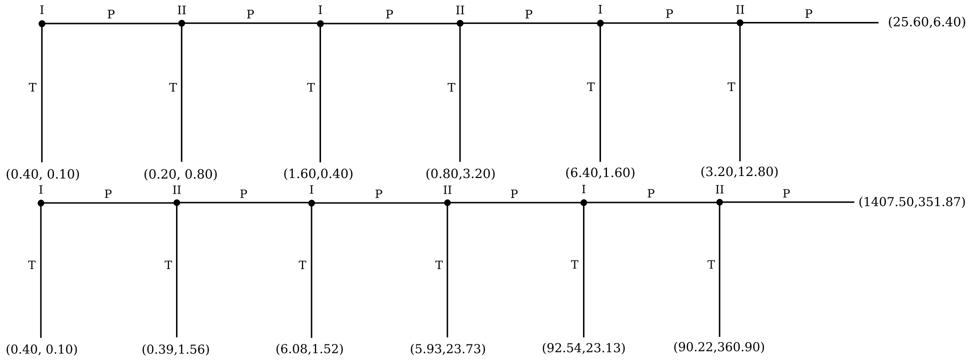

Figure 4: The six-turn centipede game. Top: version studied by McKelvey and Palfrey, with pot multiplier equal to 2. Bottom: version with pot multiplier equal to 3.9.

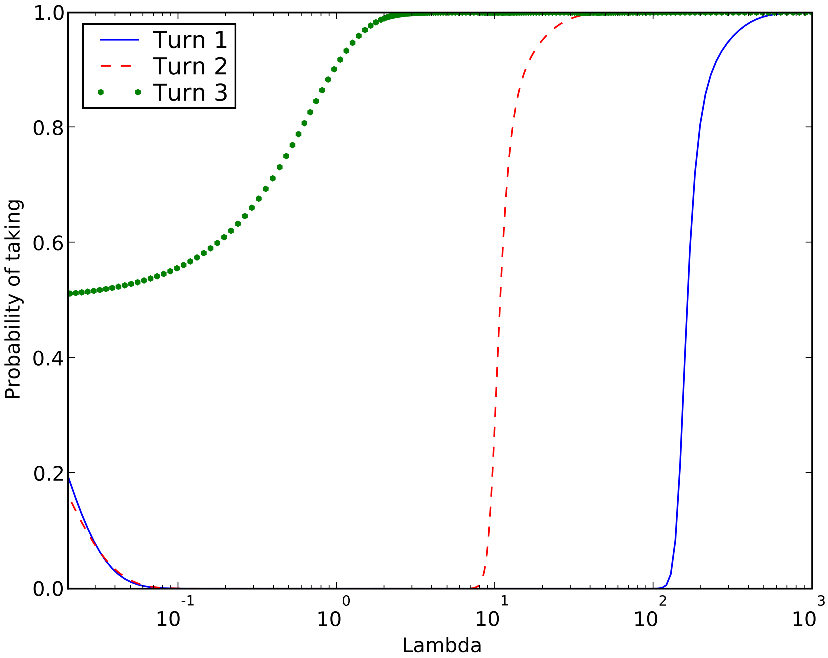

Figure 5: Probability first player takes at each of his turns as a function of λ in AQRE of the centipede game with multiplier $m = 3.9$.

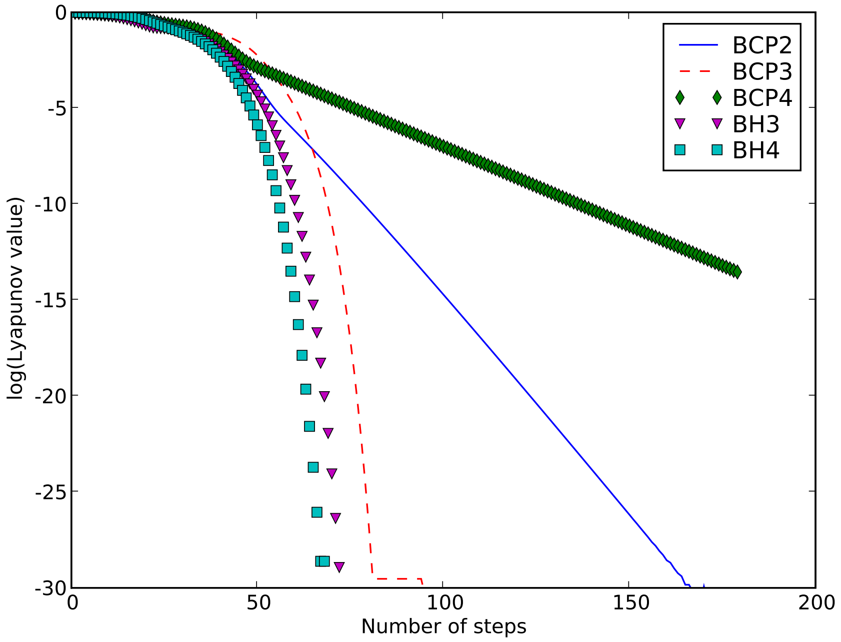

Figure 6: Plot of $\log_{10}$ of the Lyapunov value as a function of the number of steps in tracing AQRE for five signaling games.