Using Quantal Response to Compute Nash and Sequential Equilibria

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

13 November 2006

Abstract

The limit of any convergent sequence of logit quantal response equilbria is a Nash equilibrium in a strategic game, and 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 practical to compute such sequences, and thereby compute good approximations to Nash and sequential equilibria. 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 [19] it was shown that the logit quantal response equilibrium (QRE) correspondence, as defined by McKelvey and Palfrey [13] for strategic games, 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 representation of the Jacobian of the system of equations defining each branch suggests that computing a Nash equilibrium as the limit of a sequence of approximating QREs would be numerically feasible. This paper shows that, by using a logarithmic transformation of the probabilities in the strategy profile, it is possible to trace the correspondence accurately and efficiently for arbitrarily large values of the precision parameter λ.

The logarithmic representation is essential for the practical implementation for two reasons. First, it is possible to construct games for which a strategy played with positive probability in the limiting equilibrium of a branch may be played with arbitrarily small probability at some point along the branch. One such class of games is the strategic representations of the centipede game illustrated in Section 3.2. The probability of the strategy played with probability one by the first mover in any equilibrium becomes small along the correspondence, and can attain levels on the order of 10^{-300} for even modest lengths of the game.

The ability to accurately represent small probabilities permits application of the same method to computing a sequential equilibrium in extensive games using the agent logit quantal response equilibrium (AQRE) defined in MCKELVEY AND PALFREY [14]. Any limit point of the AQRE correspondence is a sequential equilibrium (KREPS AND WILSON [8]). The sequential equilibrium refinement specifies a notion of "reasonable" beliefs for information sets that are not reached in equilibrium; therefore, the methods must compute beliefs at information sets which are not reached in the limiting equilibrium. The specification in TUROCY [19] cannot be directly used for extensive games, because the elements of the Jacobian matrix need not be bounded for information sets which are reached with small probability. The logarithmic transformation results in a bounded Jacobian which can be used in a numerical procedure to trace a branch of AQREs.

This development complements the recent work of MILTERSEN AND SORENSEN [15] in two-player games. They extend the sequence form approach of KOLLER, MEGIDDO, AND VON STENGEL [7] 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 be superior for computing a sequential equilibrium in two-player games. In addition to applying to games with any number of players, the method of this paper offers the advantage of computing successively better approximations to the limiting equilibrium, which a pivoting approach does not. Also, the limiting sequential equilibria of the AQRE correspondence may not be the same as those computed by Miltersen and Sorensen's method, so even in two-player games, the two methods may be complementary in answering whether a game has a unique sequential equilibrium.

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. 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 games may, and often do, differ, having both methods also serves as a numerical approach to establishing non-uniqueness.

The paper is organized as follows. Section 2 presents the calculations underlying the tracing of a branch of the logit AQRE correspondence, which can be specialized to the case of strategic games by treating them as extensive games with simultaneous moves. Section 3 discusses performance of the algorithm in terms of speed of convergence, gives some examples of games where following the logit QRE correspondence is challenging, and discusses which equilibria the logit QRE correspondence selects out of positive-dimensional equilibrium components. Section 4 concludes with a brief discussion.

2 Tracing the correspondence

2.1 Notation and definition of the path-following problem

For the purposes of exposition, notation will be developed for extensive games, with the observation that strategic games can be viewed as a special case. The procedure operates on standard, 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 particular 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 $A(h)$. Define the set of nodes at which action a may be taken as $N(a) \equiv \{n: n \in I(a)\}$.

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. For each action a, $u_a(\pi)$ denotes 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 agent quantal response equilibria is a correspondence mapping $\lambda \in [0, \infty)$ into the set of (totally mixed) behavior profiles. McKelvey and Palfrey [14] 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) the 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

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

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

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

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:

$$\begin{aligned}
\frac{\partial H_{ab}}{\partial p_a} &= 1\\
\frac{\partial H_{ab}}{\partial p_b} &= -1\\
\text{for all other } c \in I(a), \frac{\partial H_{ab}}{\partial p_c} &= 0\\
\text{for all } c \notin I(a), \frac{\partial H_{ab}}{\partial p_c} &= -\lambda\pi_c\left[\frac{\partial u_a(\pi)}{\partial \pi_c} - \frac{\partial u_b(\pi)}{\partial \pi_c}\right]\\
\frac{\partial H_{ab}}{\partial \lambda} &= -[u_a(\pi) - u_b(\pi)].
\end{aligned}$$

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

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

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

2.2 Computation of the Jacobian entries

The only quantity in these equations that is not straightforward to compute is $\pi_c \frac{\partial u_a(\pi)}{\partial \pi_c}$ for actions a and c at different information sets. Let $P_n(\pi)$ be the probability that a node n is reached, if strategy profile π is being played, and let $u_{a|n}(\pi)$ be the payoff to playing action a, conditional on node n being reached. Then, 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 gives

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

Implementing equation (6) for extensive games requires some consideration. For a simultaneous-move game, the quantity $\frac{\partial u_a}{\partial \pi_c}(\pi)$, when the actions a and c belong to different players $P(a)$ and $P(c)$, respectively, is simply the payoff to player $P(a)$ of playing strategy a when player $P(c)$ plays strategy c, and all other players play according to π. The same observation can be used to compute the conditional action value $\frac{\partial u_{a|n}(\pi)}{\partial \pi_c}$: compute the expected payoff to player $P(a)$ in the subtree rooted at n, assuming that play is according to π, except player $P(c)$ plays action c at information set $I(c)$. If no nodes in the subtree rooted at n are in information set $I(c)$, then $\frac{\partial u_{a|n}(\pi)}{\partial \pi_c} = 0$.

The quantity $\frac{\partial}{\partial \pi_c}\left[\frac{P_n(\pi)}{P_h(\pi)}\right]$ is the change in the beliefs at node n as the probability of playing action c changes. Game-theoretically, this concept is not well-defined; in order to increase (decrease) the probability of action c being played, the probability of some other action or actions at c's information set must be decreased (increased). Consider the two game trees in Figure 1. The first game is a simultaneous-move game represented in extensive form; the second game is identical, except Player I has an outside option X which ends the game without Player II playing. In the first game, changing the probability of actions U or D for Player I do not change the realization probability of Player II's information set. In the second game, how should one define the change in the realization probability of Player II's information set as the probability of one of Player I's actions, for example π_U, is changed?

In the calculation that follows,

$$\frac{\partial P_n(\pi)}{\partial \pi_c} = \mathbf{1}_{c \prec n} \frac{P_n(\pi)}{\pi_c} \tag{7}$$

and

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

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. Equation (7) is based on the observation that

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

and, therefore, the change in the realization probability of node n with a change in π_c is the same as realization probability of node n assuming c is played with probability one, and actions at all other information sets are played with probabilities as specified in π. Equation (8) then implies that, for the simultaneous-move game in Figure 1, the change in the probability that Player II's information set is reached is defined to be positive, even though, game-theoretically, this cannot be.

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

$$\begin{aligned}
\frac{\partial}{\partial \pi_c}\left[\frac{P_n(\pi)}{P_h(\pi)}\right] &= \frac{P_h(\pi)\frac{\partial P_n(\pi)}{\partial \pi_c} - P_n(\pi)\frac{\partial P_h(\pi)}{\partial \pi_c}}{P_h(\pi)^2} \\
&= \frac{\frac{\partial P_n(\pi)}{\partial \pi_c}}{P_h(\pi)} - \frac{P_n(\pi)}{P_h(\pi)} \cdot \frac{\frac{\partial P_h(\pi)}{\partial \pi_c}}{P_h(\pi)} \\
&= \frac{\frac{\partial P_n(\pi)}{\partial \pi_c}}{P_h(\pi)} - \frac{P_n(\pi)}{P_h(\pi)} \cdot \sum_{m \in N(a)} \frac{\frac{\partial P_m(\pi)}{\partial \pi_c}}{P_h(\pi)}. \qquad (9)
\end{aligned}$$

Taking equation (6) and substituting in (9),

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

Returning to the definition in (7), it is now seen that the choice of a logarithmic representation for tracing the agent QRE correspondence is essential. The presence of the π_c factor in $\pi_c \frac{\partial u_a(\pi)}{\partial \pi_c}$ comes from the application of the chain rule, due to the representation of probabilities by their logarithms. Were one to compute $\frac{\partial u_a(\pi)}{\partial \pi_c}$ alone, there would be terms of the form $\frac{P_n(\pi)}{\pi_c P_h(\pi)}$ present in the entries in the Jacobian arising from $\frac{\partial H_{ab}}{\partial p_c}$. When both π_c and $P_h(\pi)$ are small, this ratio may be unbounded.

To illustrate this possibility and to understand why the logarithmic transformation neutralizes the problem effectively, consider the game tree in Figure 2, from SELTEN [18]. The labels in parentheses on the branches indicate action probabilities. The probability that Player III's information set is reached is $(1-\varepsilon)\delta+\varepsilon$, and Player III's beliefs about at the left node in his information set should be

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

Differentiating this with respect to δ gives

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

Suppose that both ε and δ are positive and small, and that δ goes to 0 much faster than ε. Then this quantity becomes approximately

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

which is unbounded as ε becomes small.

However, if δ is instead represented by its logarithm, $\Delta = \ln \delta$, then:

$$\begin{aligned} \mu &= \frac{(1-\varepsilon)\mathrm{e}^{\Delta}}{(1-\varepsilon)\mathrm{e}^{\Delta}+\varepsilon} \\ \frac{\partial \mu}{\partial \Delta} &= \frac{\varepsilon(1-\varepsilon)\mathrm{e}^{\Delta}}{[(1-\varepsilon)\mathrm{e}^{\Delta}+\varepsilon)]^2}, \\ \lim_{\delta\to 0} \frac{\partial \mu}{\partial \Delta} &= 0. \end{aligned}$$

In TUROCY [19] it was shown that the Jacobian could be represented in a way in which no terms diverged even as $\lambda \to \infty$; this assertion assumed that the quantity $\frac{\partial u_a(\pi)}{\partial \pi_c}$ was bounded. For strategic games, the focus of that paper, this assumption is always true; for extensive games, this may not hold due to unboundedness of the change in beliefs. However, with the logarithmic transformation, the well-behavedness of the Jacobian is recovered.

2.3 Computing beliefs

Finally, the essence of computing a sequential equilibrium is to compute accurately the beliefs for 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 ideal. Define $m \in \text{argmax}_{n \in h} P_n(\pi)$; 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

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

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

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

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 constriction, 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 experiments

The method described in Section 2 has been implemented in the software package Gambit [12]. 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 Speed of convergence

The performance of a preliminary implementation of a predictor-corrector method for tracing logit QREs was analyzed in TUROCY [19]. Those results, for strategic games only, did not use the logarithmic transformation described here, and used a simplification in which small probabilities (less than about 10^{-10}) were set to zero. Therefore, new experiments are in order to verify that the steplength acceleration properties asserted previously still hold in practice for this new representation.

The process proposed here computes a sequence of strategy profiles which converge to an equilibrium in the limit. Thus, this method belongs to the class of methods which can be thought of as "approximation" methods, which include simplicial subdivision (VAN DE LAAN ET AL [20]), Lyapunov function minimization (MCKELVEY [11]), and the method of YAMAMOTO [21] for computing a proper equilibrium. These can be contrasted with enumeration methods, including MANGASARIAN [10], LEMKE AND HOWSON [9], and support enumeration methods such as PORTER ET AL [17] and HERINGS AND PEETERS [6]. Since an approximation method typically never produces an exact equililbrium,[1] evaluation should be based on the relationship between the number of steps required and the quality of the approximation. The quality of the equilibrium approximations are evaluated here using the Lyapunov function of MCKELVEY [11]. In brief, the Lyapunov function is computed as the sum of the squares of the regrets over all information sets. Thus, it is a non-negative quantity which is zero exactly at an equilibrium.

The first results are for the signaling games considered in MCKELVEY AND PALFREY [14]. In that paper, the graphs shows were computed using a grid search method which became infeasible at relatively small values of λ; this can be seen 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.

The games used are signaling games from BANKS, CAMERER, AND PORTER [2] (BCP, games 2, 3, and 4) and BRANDTS AND HOLT [3] (BH, games 3 and 4). For each game, the principal branch of the correspondence is traced to $\lambda = 10^6$. Figure 3 presents the $\log_{10}$ of the Lyapunov value as a function of the number of steps taken.[2] This figure illustrates the three modes of convergence behavior, which depend on the characteristics of the limiting equilibrium. A first observation is that an approximation of good quality is computed within about 50 steps.

1. The method of tracing QREs only produces an exact equilibrium in the special case where uniform randomization at all information sets forms an equilibrium, in which case uniform randomization is a QRE for all values of λ.

2. Since the magnitude of the Lyapunov value depends on the size of the payoffs and the total number of actions in the game, the reported Lyapunov values are normalized such that the Lyapunov value of uniform randomization at all information sets is 1.

The limiting equilibrium on the principal branch in games BCP3, BHG3, and BHG4 is a strict pure-strategy equilibrium. 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. This means that the predictor portion of the predictor-corrector method is very accurate, since it extrapolates linearly based on the Jacobian of the system. Therefore, the AQRE tracing method converges quickly to such an equilibrium. For a strategic game, the fact that the method converges quickly would not be particularly interesting, since simple enumeration of contingencies is a much faster method in practice for finding pure-strategy equilibria. In extensive games, pure-strategy equilbria generally result in unreached information sets; the AQRE generates a sequential equilibrium assessment including beliefs at such information sets.

The game BCP4 has at the limit of the principal branch an equilibrium in which there is randomization at one information set, which follows "Message 1" in the game. In general, 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.

Finally, BCP2 represents the intermediate case, in which the equilibrium is in pure strategies, but the equilibrium is not strict at all information sets. Thus, the probability of the action which is not inferior but is not played in equilibrium does not decay exponentially, so convergence as measured in the action space is slow. The Lyapunov function does decrease faster, though, because the weight being put on the unplayed strategy does not generate significant regret.

The relatively slow convergence in mixed strategies suggests a hybrid approach. Once one is clearly on the asymptotic portion of the branch of QREs, one can polish a QRE strategy profile into a Nash profile using the Lyapunov function method. The Lyapunov function method, in practice, is very non-globally convergent from a random starting point. However, close to an equilibrium, gradient descent should be very effective in computing mixed strategy probabilities to greater accuracy.

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 about 0.3s for the BCP games and 0.2s for the BH games (which are smaller) on the author's workstation (with a 2.4GHz Xeon processor). The relevant measure of the size of the game for this algorithm is the total number of actions (or strategies). Running times remain feasible for at least some games of nontrivial size; the principal branch for the average cost pricing game in CHEN ET AL [5], which is a four-player game in which each player has 21 strategies, is traced in under 7 minutes on the same workstation.

3.2 Challenging cases

While the logit QRE correspondence is well-behaved, in the sense that branches are differentiable except at bifurcation points, branches of the correspondence may exhibit significant nonmonotonicity. In the context of the application of QRE to explaining behavior in laboratory experiments, these nonmonotonicities are a positive feature. For computational purposes, they may present difficulties.

As already discussed, one problem is that, when an information set is reached with small probability, beliefs may depend sensitively on probabilities of actions played on the path to members of the information set. In the game in Figure 2, this does occur on the principal branch of the AQRE correspondence. In the sequential equilibrium to which the principal branch limits, both Players I and II play R with probability one, and Player III's information set is not realized. The assessment justified by a sequence of AQREs puts probability $\mu = \frac{2}{3}$ on the left node in that information set, and Player III chooses action L with probability one.[3] Along this branch, the quantity $\frac{\partial\mu}{\partial\delta}$ increases without bound. The left panel of Figure 4 plots $\ln \frac{\partial\mu}{\partial\delta}$ as a function of λ on this branch. Without the logarithmic translation of probabilities, the Jacobian of the system would be unbounded. Using the logarithmic representation, the relevant quantity is $\delta\frac{\partial\mu}{\partial\delta}$, which is plotted in the right panel of Figure 4; this quantity is well-behaved.

A second feature of QRE correspondences is that a probability which is used with positive probability – or even probability one – in the limiting equilibrium may be used with arbitrarily small probability along the same branch. Examples of this can be constructed for any desired probability using the reduced normal form of centipede games of increasing length. Following the centipede game examples in McKelvey and Palfrey [14], consider a family of centipede games parameterized by the number of innings M; an inning consists of one turn for each player. The pot starts at 0.50. Players alternate turns. If a player chooses to take the pot, he receives 80% of the value of the pot, and the other player receives 20%. If he passes, the pot is doubled, and it becomes the other players turn.

In the reduced normal form, each player has $M + 1$ pure strategies. M of these strategies are of the form "take in inning d," and the last strategy is "always pass." For example, the reduced normal form for the two-inning is shown in Table 1.

	Take in 1	Take in 2	Always pass
Take in 1	0.40, 0.10	0.40, 0.10	0.40, 0.10
Take in 2	0.10, 0.80	1.60, 0.40	1.60, 0.40
Always pass	0.10, 0.80	0.80, 3.20	6.40, 1.60

Table 1. Reduced normal form of the two-inning centipede game.

There is a unique subgame perfect, and therefore sequential, equilibrium in the extensive form. In the reduced normal form, there is a convex component of Nash equilibria in

3. Note that at $\mu = \frac{2}{3}$, Player III is indifferent between his actions.

which the first player takes in inning 1 with probability 1. Since this makes the second player indifferent among his strategies, he may randomize over them, subject to the constraint that "Take in 1" remains a best-reply for the first player.

In these games, along the principal branch of the QRE correspondence in the normal form, the probability with which the first player plays "Take in 1" initially decreases, and only increases towards the equilibrium value of 1 for large λ. Intuitively, this happens because, when player 2 is playing randomly further down the tree, it is attractive to pass in the first inning and allow the pot to grow. Because the pot doubles with each pass, the pot grows quite large as the number of innings increases. Table 2 presents the smallest probability with which "Take in 1" is played along the principal branch, and approximately the value of λ at which the minimum is attained. The minimum probability decreases rapidly in the number of innings. One can make this minimum as small as desired by adding innings.

Innings	Minimum probability	occurs at λ
2	0.00153	7.7
3	3.6×10^{-23}	45.3
4	1.6×10^{-88}	170.7
5	1.9×10^{-292}	562.5

Table 2. Minimum of the probability the first player plays "Take in 1" in a QRE of the reduced normal form of the centipede game, as a function of the number of innings.

3.3 Selection from a positive-dimensional component of equilibria

As noted, the game in Figure 1 has a convex set of equilibria. The equilibria form a polytope with four vertices. All equilibria have the first player playing "Take in 1" with probability one. The four vertices have the second player playing his three strategies with probabilities $(1, 0, 0)$, $\left(\frac{6}{7}, \frac{1}{7}, 0\right)$, $\left(\frac{30}{31}, 0, \frac{1}{31}\right)$, and $\left(\frac{6}{7}, \frac{6}{49}, \frac{1}{49}\right)$. The limit of the principal branch of the logit QRE correspondence is the last vertex listed, which is also the one which maximizes the entropy of the mixed strategy distribution.

This observation is perhaps not surprising, since entropy is intimately connected with logit QRE. Suppose, in a logit QRE, a player obtains an expected payoff of $\bar{u}$. Then, in the logit QRE, the player is playing the mixed strategy that maximizes entropy among all mixed strategies that obtain an expected payoff of $\bar{u}$, holding fixed the strategies of the other players. The parameter λ is the Lagrange multiplier on the expected payoff constraint.

Convexity of the equilibrium component is not necessary for this property to obtain. Table 3 presents a three-player game, with each player having two strategies, from NAU ET AL [16]. As part of its set of Nash equilibria, there is a one-dimensional curve of totally mixed equilibria. From this curve, logit QRE selects the equilibrium in which the row and column players choose A with probability of about 0.55051, and the table player chooses A with probability 0.25. This is the entropy-maximizing equilibrium from this set.

	A		B	
	A	B	A	B
A	0,0,2	0,3,0	1,1,0	0,0,0
B	3,0,0	0,0,0	0,0,0	0,0,3

Table 3. A game where the limit of logit QREs maximizes entropy along a nonconvex equilibrium component. Payoff vectors are (row player, column player, table player).

However, the conjecture that limit points of the QRE correspondence are always entropy-maximizing within a component is false. Consider the nongeneric normal form game in Table 4. This game has an isolated pure strategy Nash equilibrium (s_2, t_1), and a continuum of equilibria where the row player plays s_1 with probability 1, and the column player plays strategy t_1 with probability no more than 5/6.

In this game, the principal branch selects the equilibrium $(s_1, \frac{1}{2}[t_1] + \frac{1}{2}[t_2])$. Meanwhile, the pure strategy equilibrium $(s_2,\ t_1)$ is connected in another branch to the equilibrium $(s_1,\ \frac{5}{6}[t_1] + \frac{1}{6}[t_2])$, where the latter is the unstable branch under perturbed best-reply dynamics. This equilibrium is (locally) entropy-*minimizing* within the set of Nash equilibria.

	t_1	t_2
s_1	$6,7$	$6,7$
s_2	$7,6$	$1,4$

Table 4. A game in which the limit of logit QREs is locally entropy-minimizing within the set of Nash equilibria.

3.4 A note on genericity

McKelvey and Palfrey show that, for both the strategic and agent logit QRE specifications, for generic games there is a unique, one-dimensional branch of the correspondence which connects the centroid at $\lambda = 0$ with a Nash (sequential, for the agent version) equilibrium as $\lambda \to \infty$. The limiting equilibrium is called the "logit solution" of the game.

This genericity condition has bite in some common classes of games, which are generic in the sense of the properties of their set of Nash equilibria, but not in terms of the uniqueness of the logit solution. The most common example is that of symmetric coordination games. For example, consider the battle of the sexes game from CHEN, FRIEDMAN, AND THISSE [4], presented in Table 5. Chen et al assert that this game has a unique logit QRE which limits to the mixed-strategy equilibrium. In fact, this game is not generic, in that the logit solution is not well-defined. The branch of the correspondence emanating from the centroid at $\lambda = 0$ encounters a pitchfork bifurcation at $\lambda \approx 1.372$, at which each player chooses his or her preferred activity with probability about .5841.

	Ballet	Tennis
Ballet	3,2	1,1
Tennis	1,1	2,3

Table 5. A battle of the sexes game.

4 Conclusion

This paper demonstrates the feasibility of computing a Nash equilibrium in a strategic game or a sequential Nash equilibrium in an extensive game by constructing a sequence of logit quantal response equilibria. 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.

Bibliography

[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] Hsiao-Chi Chen, James W. Friedman, and Jacques-Francois Thisse. Boundedly rational Nash equilibrium: A probabilistic choice approach. *Games and Economic Behavior*, 18:32–54, 1997.

[5] Yan Chen, Laura Razzolini, and Theodore L. Turocy. Congestion Allocation for Distributed Networks: An Experimental Study. Working paper, 1 September 2005.

[6] P.J.J. Herings and R.J.A.P. Peeters. A globally convergent algorithm to compute all Nash equilibria for n-Person games. *Annals of Operations Research*, 137:349–368, 2005.

[7] 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.

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

[9] C. E. Lemke and Jr. Howson, J. T. Equilibrium points of bimatrix games. *Journal of the Society of Industrial and Applied Mathematics*, 12:413–423, 1964.

[10] Oscar Mangasarian. Equilibrium points of bimatrix games. *Journal of the Society for Industrial and Applied Mathematics*, 12:778–780, 1964.

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

[12] Richard D. McKelvey, Andrew M. McLennan, and Theodore L. Turocy. Gambit: Software Tools for Game Theory. Version 0.2006.10.31.

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

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

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

[16] Robert Nau, Sabrina Gomez Canovas, and Pierre Hansen. On the Geometry of Nash Equilibria and Correlated Equilibria. *International Journal of Game Theory*, 2005. Forthcoming.

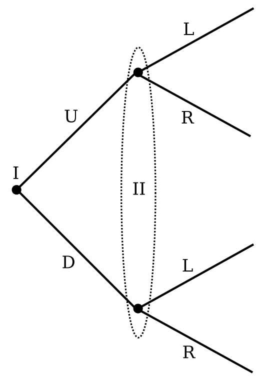
[17] Ryan W. Porter, Eugene Nudelman, and Yoav Shoham. Simple search methods for finding a Nash equilibrium. *Games and Economic Behavior*, 2006.

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

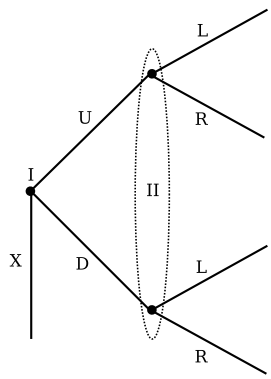
[19] Theodore L. Turocy. A dynamic homotopy interpretation of the logistic quantal response equilibrium correspondence. *Games and Economic Behavior*, 51:243–263, 2005.

[20] G. van de Laan, A. J. J. Talman, and L. van Der Heyden. Simplicial variable dimension algorithms for solving the nonlinear complementarity problem on a product of unit simplices using a general labelling. *Mathematics of Operations Research*, 12:377–397, 1987.

[21] Yoshitsugo Yamamoto. A path-following procedure to find a proper equilibrium of finite games. *International Journal of Game Theory*, 22:249–259, 1993.



(a) A simultaneous-move game



(b) A game where Player I may opt out

Figure 1. Two game trees which illustrate the considerations in computing the derivative of beliefs with respect to the probability with which an action is played.

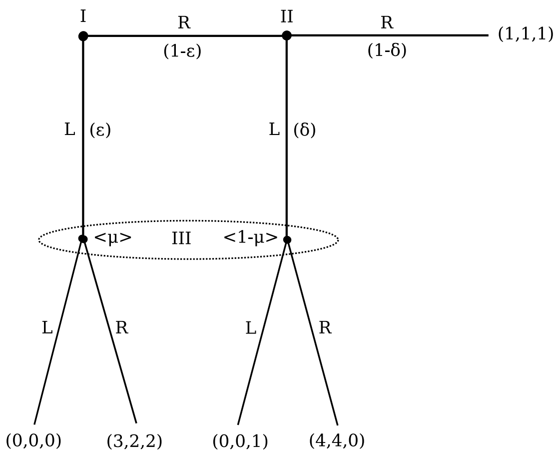


Figure 2. A game in which the change in beliefs as a function of the change in probabilities may be unbounded.

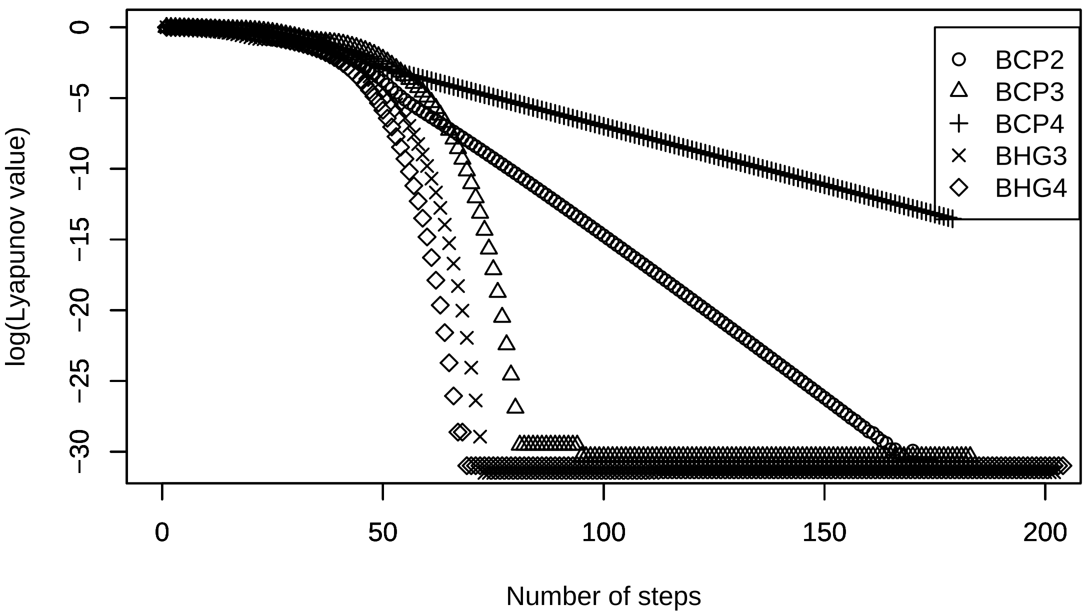


Figure 3. Plot of $\log_{10}$(Lyapunov value) as a function of the number of steps for five signaling games.

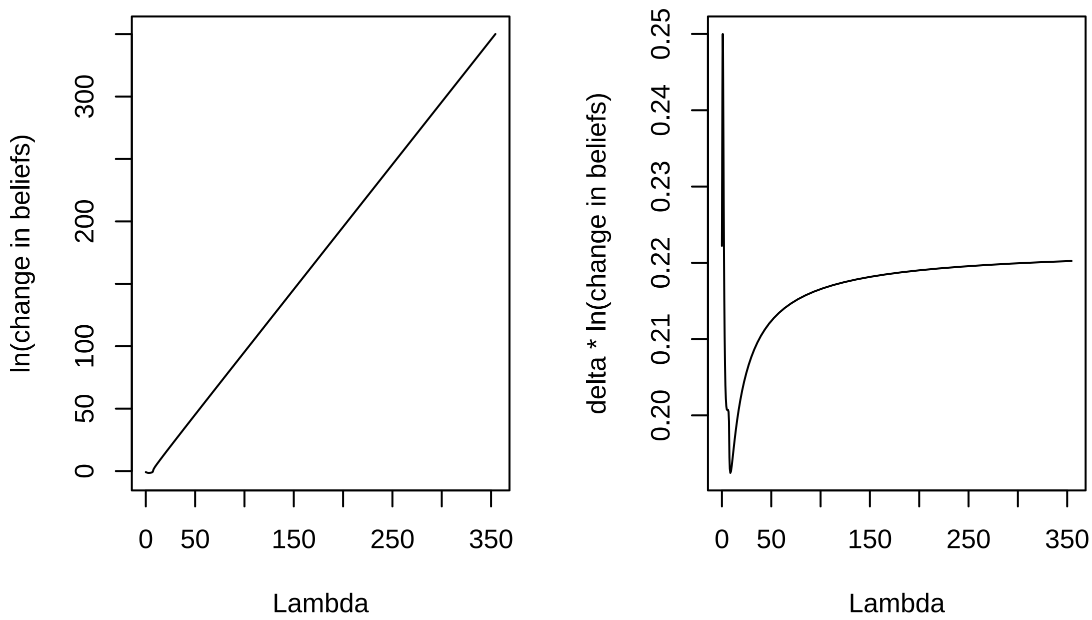


Figure 4. Plots of $\frac{\partial\mu}{\partial\delta}$ and $\delta\frac{\partial\mu}{\partial\delta}$ along the principal branch of the AQRE correspondence of the game in Figure 2.